

---

# HIV-1 CORECEPTOR CCR5: GENE CHARACTERIZATION AND EXPRESSION

---

Anabela Correia Pereira Picton

A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand,  
Johannesburg, in fulfillment of the requirements for the degree of Doctor of Philosophy

Johannesburg 2013

---

---

## DECLARATION

---

I declare that this thesis is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other university.

---

(Anabela Correia Pereira Picton)

\_\_\_\_\_ day of \_\_\_\_\_ 2013

---

## ABSTRACT

---

Genetic variability within both the HIV-1 coreceptor, *CCR5*, and its ligand, *CCL3L*, has been shown to contribute towards differences between individuals in their susceptibility to HIV-1 infection and rate of disease progression. In this study we investigated the extent of genetic variation within the *CCR5* gene as well as *CCL3L*, *CCL3La* and *CCL3Lb* gene copy number distribution in two healthy HIV-1 uninfected South African populations, South African Africans (SAA) and South African Caucasians (SAC). The impact of variations within these genes on the expression of CCR5 and CCL3 was subsequently assessed. Furthermore, *CCR5* genetic variability, *CCL3L* gene copy number distribution and the expression of CCR5 and CCL3, was assessed in a similar way in a small cohort of HIV-1 infected long term nonprogressors (LTNPs).

Genotyping of the *CCR5* gene in SAA (n=41) and SAC (n=46) HIV-1 uninfected individuals revealed a high degree of genetic variation between the two population groups, both in terms of single nucleotide polymorphism (SNP) profiles and *CCR5* haplotype distribution. Seven complex putative haplotypes spanning the length of the sequenced region were identified with only one of the identified haplotypes, SAA/C-HHC, common to both study populations. The effect of genetic variability on promoter activity of four different *CCR5* promoter regions for three *CCR5* haplotypes, SAA-HHA, SAA/C-HHC and SAC-HHE, were evaluated. Results showed variability in (i) promoter activity between different promoter regions tested, (ii) results obtained with different cells used for analysis, and (iii) the haplotype being analysed, thereby highlighting that both the cellular environment as well as genetic variability within the promoter region, have the capacity to influence the efficiency of a promoter and consequently CCR5 expression levels. Haplotype-specific promoter analysis demonstrated the SAA-HHA haplotype to have the strongest promoter activity in THP-1 and K562 cells for both P1A (downstream) and P2 (upstream) promoter regions, while in the other cell lines tested (Jurkat and U937), HHA demonstrated intermediate promoter strength. Differences seen between the haplotypes tested in this study and other published studies may be attributable to additional SNPs being tested in the promoter constructs used in this study.

The two population groups differed significantly with regards to cell activation levels, as measured by HLA-DR expression, in CD4<sup>+</sup> T cell ( $P=0.002$ ) and CD56<sup>+</sup> NK cell subsets ( $P<0.001$ ). CCR5 expression, determined both as the number of CCR5 molecules per cell (density) and the percentage of CCR5-expressing cells, was found to differ between SAA and SAC individuals across all peripheral blood cell types. SAA individuals had larger proportions of CCR5-expressing natural killer (NK) cell subsets ( $P<0.01$ ) but lower CCR5 molecules per cell density on CCR5<sup>+</sup>CD8<sup>+</sup> T cell and CCR5<sup>+</sup> NK cell subsets (CD56<sup>+</sup>, CD16<sup>+</sup>CD56<sup>+</sup> and CD56<sup>dim</sup>) (all  $P<0.05$ ) compared to SAC individuals. These differences were maintained even after *CCR3Δ32*

heterozygous SAC individuals were included in the analyses. Furthermore, the previously described haplotypes, HHA and HHC, associated with differences in CCR5 expression on different cell subsets between individuals within the same population group. SAA individuals with the HHA haplotype had significantly lower percentages of CCR5-expressing CD8+ T cells compared to SAA individuals that lacked the haplotype ( $P=0.001$ ). SAC individuals with the HHC haplotype had significantly higher density on NK (CD56+) and CD16+CD56+ NK cell subsets ( $P=0.030$  and  $P=0.024$ , respectively) compared to SAC individuals without this haplotype. The latter observation suggests that the protective effect of the HHC haplotype in Caucasians might be explained by higher density of CCR5 expression on NK cells that is not evident in HHC+ SAA individuals, thus highlighting the potential role of CCR5-expressing cells other than CD4+ T cells in protection from HIV-1 acquisition and disease progression.

Despite significant differences in *CCL3La* (*CCL3L* chemokine coding) and *CCL3Lb* (non-chemokine coding) copy number between SAA and SAC populations, no difference in CCL3 production by peripheral blood mononuclear cells (PBMCs) was noted between the two study populations. Assuming equal contribution of *CCL3* and each copy of *CCL3La* to CCL3 production, we found that SAC individuals produced higher levels of CCL3 per functional copy of *CCL3La* compared to SAA individuals ( $P<0.001$ ). Although, when SAA and SAC individuals with comparable *CCL3La* and *CCL3Lb* gene copy numbers were compared, there was no difference in production per functional copy between the two groups ( $P=0.974$ ). We also determined *CCL3La* and *CCL3Lb* gene copy number for a previously established cohort of HIV-1 intrapartum-infected (IP) and exposed uninfected (EU) infants and found that differences previously seen in cord blood mononuclear cell (CBMC) CCL3 production between IP and EU infants with comparable *CCL3L* copy numbers could not be attributed to differences in *CCL3Lb* copy number.

The potential role of differences in *CCR5* genotype, CCR5 expression, *CCL3* genotypes and CCL3 production levels in the control of HIV-1 infection was then examined by comparing a small group (10 SAA and 4 SAC) of LTNPs to the respective background population. No polymorphisms in the *CCR5* open reading frame were detected in these LTNP individuals. However, the HHA haplotype frequency was significantly higher in SAC LTNP individuals compared to SAC control individuals ( $P=0.010$ ). Interestingly, CCR5 density on CD4+ T cells and monocytes was significantly lower in SAA LTNP individuals ( $P=0.025$  and  $P=0.022$ , respectively) with a trend towards a similar relationship in CD8+ T cells ( $P=0.058$ ), while the proportions of CCR5-expressing CD8+ T cells were elevated compared to SAA controls ( $P=0.043$ ). This latter finding reflects the increased immune activation in these individuals compared to uninfected individuals, as evidenced by increased proportions of HLA-DR-expressing T cells (CD8+ and CD4+,  $P<0.0001$ ). In addition, PHA-induced CCL3 production by PBMCs was significantly lower in LTNP (SAA and SAC

combined) compared to control individuals ( $P=0.004$ ). SAA LTNP individuals had higher proportions of CD8+ T cells ( $P<0.0001$ ) and lower proportions of natural killer cells (CD56+,  $P=0.002$ ) compared to control SAA individuals. Thus, CCL3 production differences may be partially explained by differences in the distribution of immune cell subsets between the two study groups. Furthermore, PBMCs of LTNP individuals with low viral loads (<400 copies/ml) produced CCL3 at lower levels than those from individuals with higher viral loads, irrespective of whether or not the cells were stimulated ( $P=0.005$  and  $P=0.035$ , respectively).

In summary, this study demonstrates that: (i) two ethnically divergent populations show marked differences in *CCR5* genetic variability, cell activation and CCR5 expression which are likely to impact on both susceptibility to HIV-1 infection and the rate of HIV-1 disease progression, (ii) both *CCR5* genotypic differences and differences in baseline cellular activation levels appear to be contributing towards the observed differences in CCR5 protein expression, and (iii) the two study populations do not differ with respect to CCL3 production by PBMC cultures which suggests that either the two copy per diploid genome gene, *CCL3*, may play a significant role in CCL3 production and/or that as yet undefined mechanisms regulate production of CCL3 from variable *CCL3L* copy number. In addition, a pilot study conducted in a small group of LTNP individuals demonstrates that two major determinants of HIV-1 disease progression, CCR5 and CCL3, are both expressed at lower levels in LTNPs individuals compared to healthy uninfected controls and has identified *CCR5* haplotypes which are potentially associated with disease progression.

---

## ACKNOWLEDGEMENTS

---

I would like to express my sincere thanks to the following people and institutions for their contribution to the work presented in this thesis:

The South African AIDS Vaccine Initiative (SAAVI), Poliomyelitis Research Foundation (PRF), Wellcome Trust (Caroline T. Tiemessen is the recipient of a Wellcome Trust International Senior Research Fellowship (076352/Z/05/Z)), Technology Innovation Agency [South African HIV/AIDS Research (and Innovation) Platform (SHARP)] and University of the Witwatersrand (Faculty Research Committee) for funding this research.

My supervisor, Prof. Caroline Tiemessen, for her mentorship, guidance and positive attitude, no matter what. My co-supervisor, Dr Maria Paximadis, whose advice, help, encouragement, and above all friendship, throughout my thesis has been invaluable. Dr Sharon Shalekoff for training in flow cytometry data acquisition and analysis. Ms. Avani Bharuthram for conducting the PBMC isolations and production assays. My colleagues in the Cell Biology division of the Centre for HIV and STIs, for providing a wonderful environment in which to work.

Prof. Sunil Ahuja, whose enthusiasm for research is highly contagious, for hosting me in his laboratory for the period of my CU-South African Fogarty AITRP fellowship. Although the work conducted in his laboratory does not form part of this thesis, the experience and knowledge I gained while there was invaluable. The members of Prof. Ahuja's laboratory for their help during my fellowship. In particular to Dr Srinivas Mummidi for his advice on promoter expression analysis, Dr Weijing He, Dr John Castiblanco-Quinche and Ms. Komathy Jayasekar.

The Columbian University Mailman School of Public Health for funding my traineeship fellowship in Prof. Sunil Ahuja's laboratory.

Dr Mike Greyling for assistance with Modified one-way ANOVA statistical analysis and for statistical advice.

My family for their love and encouragement throughout the years, particularly my husband, Chris, for his unwavering support in so many ways, and my daughter, Daniela, for her patience.

Last, but not least, I'd like to thank the volunteer participants in this study.

---

# TABLE OF CONTENTS

---

<b>Declaration</b> .....	<b>ii</b>
<b>Abstract</b> .....	<b>iii</b>
<b>Acknowledgements</b> .....	<b>vi</b>
<b>Table of Contents</b> .....	<b>vii</b>
<b>List of Tables</b> .....	<b>xiii</b>
<b>List of Figures</b> .....	<b>xv</b>
<b>Abbreviations</b> .....	<b>xix</b>
<b>Preface</b> .....	<b>xxi</b>
<b>CHAPTER 1 Introduction</b> .....	<b>1</b>
1.1. Mechanism of cellular entry of HIV .....	3
1.2. Chemokine receptors .....	4
1.2.1. CCR5 .....	5
1.2.1.1. CCR5 antagonists for HIV therapy.....	5
1.2.1.2. The CCR5 protein .....	6
1.2.1.3. The role of CCR5 in the immune system .....	8
1.2.1.4. <i>CCR5</i> gene .....	9
1.2.1.5. Genetic variation within the <i>CCR5</i> gene .....	10
1.2.1.5.1. ORF polymorphisms .....	10
1.2.1.5.2. Promoter polymorphisms .....	16
1.2.1.5.3. <i>CCR5</i> Haplotypes .....	17
1.2.1.6. <i>CCR2V64I</i> mutation .....	18
1.2.1.7. CCR5 expression .....	19
1.2.1.7.1. CCR5 expression and HIV-1 infection .....	20
1.3. Chemokines.....	21
1.3.1. CCL3 (MIP-1 $\alpha$ ) .....	22
1.3.1.1. <i>CCL3</i> genes .....	23
1.3.1.2. <i>CCL3</i> protein isoforms .....	24
1.3.1.3. CCL3 and HIV-1 .....	25
1.4. The <i>CCR5-CCL3L</i> coreceptor-ligand axis and HIV-1 .....	26
1.5. Rationale for the current study .....	27
1.6. Study objectives.....	29

<b>CHAPTER 2 Genetic variation within the gene encoding the HIV-1 CCR5 coreceptor in two South African populations.....</b>	<b>30</b>
Abstract .....	31
2.1. Introduction .....	31
2.2. Materials and methods.....	32
2.2.1. Study population.....	32
2.2.2. PCR and sequencing of <i>CCR5</i> .....	32
2.2.3. Sequence analysis .....	33
2.2.4. <i>CCR2-V64I</i> .....	33
2.2.5. Determination of wild type (WT) reference sequence .....	34
2.2.6. Inference of putative haplotypes .....	34
2.2.7. Characterization of indels .....	34
2.2.8. Hardy-Weinberg equilibrium .....	34
2.2.9. Linkage disequilibrium between haplotype SNPs .....	35
2.2.10. Fisher exact test .....	35
2.3. Results.....	35
2.3.1. Single nucleotide polymorphisms .....	35
2.3.2. Indels.....	40
2.3.3. Haplotypes .....	41
2.3.4. Hardy-Weinberg equilibrium .....	44
2.4. Discussion .....	44
<b>CHAPTER 3 <i>In vitro</i> functional analysis of different <i>CCR5</i> promoter regions and the impact of genetic variability present within the SAA/C-HHC, SAA-HHA and SAC-HHE haplotypes, on promoter function.....</b>	<b>49</b>
Abstract .....	50
3.1. Introduction.....	50
3.2. Materials and methods.....	53
3.2.1. Construction of <i>CCR5</i> Promoter Constructs .....	53
3.2.1.1. P1A and P2.....	54
3.2.1.2. P1B and P1+P2 .....	54
3.2.2. Cell lines.....	55
3.2.3. Transfection of cell lines and Luciferase assays.....	55
3.2.4. Statistical analysis .....	56
3.2.5. Bioinformatics analysis .....	56
3.3. Results.....	56
3.3.1. Efficiency of the different promoter constructs .....	56

3.3.2. Comparison of haplotype-specific promoter efficiencies .....	59
3.3.2.1. P1A: <i>CCR5</i> downstream promoter region .....	59
3.3.2.2. P1B: <i>CCR5</i> downstream promoter region + <i>CCR5</i> intron 2 .....	60
3.3.2.3. P2: <i>CCR5</i> upstream promoter .....	61
3.3.2.4. P1+P2 .....	62
3.3.3. The newly identified SNP, -2150A/G, negatively influences promoter activity .....	63
3.3.4. Evaluation of -2150A/G SNP on putative transcription factor binding.....	64
3.4. Discussion .....	65
<b>CHAPTER 4 Marked differences in <i>CCR5</i> expression and activation levels in two South African populations.....</b>	<b>70</b>
Abstract .....	71
4.1. Introduction .....	71
4.2. Materials and methods.....	73
4.2.1. Study participants .....	73
4.2.2. Whole blood surface staining and flow cytometry .....	73
4.2.3. Statistical analysis .....	74
4.3. Results.....	74
4.3.1. <i>CCR5</i> Expression on Lymphocyte populations.....	74
4.3.2. <i>CCR5</i> expression differs between SAA and SAC individuals within all cell subsets.....	75
4.3.3. Relationship between <i>CCR5</i> expression and age of individuals.....	77
4.3.4. SAA and SAC individuals differ in cell activation levels.....	78
4.3.5. Influence of <i>CCR5</i> $\Delta$ 32 allele on <i>CCR5</i> expression within SAC population.....	79
4.4. Discussion .....	81
<b>CHAPTER 5 <i>CCR5</i> promoter haplotypes differentially influence <i>CCR5</i> expression on Natural Killer and T cell subsets in ethnically divergent HIV-1 uninfected South African populations.....</b>	<b>86</b>
Abstract .....	87
5.1. Introduction .....	87
5.2. Materials and methods.....	88
5.2.1. Cohort .....	88
5.2.2. Whole blood surface staining and flow cytometry .....	89
5.2.3. <i>CCR5</i> genotyping .....	89
5.2.3.1. <i>CCR5</i> gene .....	89
5.2.3.2. <i>CCR2-V64I</i> .....	89
5.2.4. Statistical analysis .....	89

5.3. Results.....	90
5.3.1. <i>CCR5</i> genotype distribution within the two study populations .....	91
5.3.2. Influence of haplotypes on <i>CCR5</i> expression .....	92
5.3.3. HHA and T cells in the SAA population.....	93
5.3.4. HHC and T cell subsets.....	94
5.3.5. HHC and NK cell subsets .....	96
5.3.6. HHD and HHE haplotypes .....	98
5.4. Discussion .....	98
<b>CHAPTER 6 Contribution of variable <i>CCL3L</i> copy number to <i>CCL3</i> protein production in two ethnically divergent South African populations .....</b>	<b>103</b>
Abstract .....	104
6.1. Introduction .....	104
6.2. Materials and Methods.....	106
6.2.1. Study populations.....	106
6.2.1.1. Healthy participants .....	106
6.2.1.2. HIV-1 exposed-uninfected and intrapartum-infected infants .....	106
6.2.2. Nomenclature.....	106
6.2.3. Isolation of Peripheral Blood Mononuclear Cells.....	107
6.2.4. <i>CCL3</i> production assays.....	107
6.2.5. DNA isolation.....	107
6.2.6. Evaluation of gene copy numbers by real-time quantitative PCR assay.....	107
6.2.7. Copy number assay validation.....	108
6.2.7.1. <i>CCL3Lb</i> PCR .....	108
6.2.8. <i>CCL3</i> quantification .....	108
6.2.9. Statistical analysis .....	109
6.3. Results.....	109
6.3.1. Assay validation .....	109
6.3.1.1. Sum of gene copy numbers.....	109
6.3.1.2. <i>CCL3Lb</i> PCR .....	109
6.3.2. Copy number variation in the two population groups .....	110
6.3.3. <i>CCL3</i> production does not correlate with <i>CCL3La</i> copy number.....	111
6.3.4. <i>CCL3</i> production per functional <i>CC3La</i> copy is higher in SAC individuals .....	112
6.3.5. <i>CCL3La</i> copies do not differ functionally between SAA and SAC individuals.....	114
6.3.6. <i>CCL3La</i> , <i>CCL3Lb</i> copy number variation and <i>CCL3</i> production in the context of mother to child transmission.....	115
6.4. Discussion .....	117

<b>CHAPTER 7 CCR5 and CCL3L genotypic and phenotypic differences between HIV-1 infected long term nonprogressors and healthy uninfected individuals .....</b>	<b>122</b>
Abstract .....	123
7.1. Introduction .....	123
7.2. Materials and Methods.....	125
7.2.1. Cohort description .....	125
7.2.1.1. Long term nonprogressors (LTNP) .....	125
7.2.1.2. HIV-1 infected progressors .....	126
7.2.2. Plasma viraemia quantification and CD4 T cell determination .....	126
7.2.3. Genotyping.....	127
7.2.4. CCR5 quantification.....	127
7.2.5. Absolute counts of cells in blood samples .....	127
7.2.6. CCL3 quantification .....	128
7.3. Results.....	128
7.3.1. <i>CCR5</i> gene polymorphisms.....	128
7.3.2. <i>CCR5</i> haplotypes .....	130
7.3.2.1. The HHA haplotype may be overrepresented in SAC LTNP individuals .....	130
7.3.2.2. Preliminary indications that HHC and HHE haplotypes may be associated with disease progression .....	130
7.3.3. <i>CCR5</i> expression on lymphocyte populations of LTNP individuals .....	132
7.3.4. <i>CCR5</i> expression relative to healthy HIV-1 uninfected controls .....	133
7.3.4.1. LTNP individuals have lower <i>CCR5</i> density on CD4+ T cells and monocytes compared to healthy HIV-1 uninfected individuals.....	133
7.3.4.2. LTNPs have higher proportions of <i>CCR5</i> -expressing CD8+ T cells relative to control individuals .....	135
7.3.4.3. Lower <i>CCR5</i> density on monocytes of HIV-1-infected LTNP individuals, in comparison to controls, maintained when individuals were matched for <i>CCR5</i> genotype.....	136
7.3.5. Influence of haplotypes on <i>CCR5</i> expression .....	138
7.3.5.1. HHA.....	138
7.3.5.2. HHC .....	138
7.3.6. Cell activation was higher in LTNP individuals, compared to control individuals, in T cell but not NK cell subsets .....	139
7.3.7. <i>CCL3L</i> , <i>CCL3La</i> and <i>CCL3Lb</i> copy number distribution in LTNP individuals.....	140
7.3.8. PHA-stimulated <i>CCL3</i> production in LTNP individuals compared to control individuals .....	141
7.3.9. Absolute counts of monocytes and NK cell subsets in whole blood differ between SAA and SAC individuals .....	143

7.3.10. SAA LTNP PBMCs had higher proportions of T cells and lower proportions of NK cells compared to SAA control PBMCs.....	145
7.4. Discussion .....	147
<b>CHAPTER 8 Concluding Remarks.....</b>	<b>155</b>
<b>REFERENCES.....</b>	<b>167</b>
<b>APPENDIX A Supplementary Data for Chapter 2 .....</b>	<b>190</b>
<b>APPENDIX B Supplementary Data for Chapter 3 .....</b>	<b>193</b>
<b>APPENDIX C Supplementary Data for Chapter 5 .....</b>	<b>194</b>
<b>APPENDIX D Supplementary Data for Chapter 6 .....</b>	<b>195</b>
<b>APPENDIX E Supplementary Data for Chapter 7 .....</b>	<b>196</b>
<b>APPENDIX F Amino acid codes .....</b>	<b>199</b>
<b>APPENDIX G Ethical clearance .....</b>	<b>200</b>

---

## LIST OF TABLES

---

<b>Table 1.1.</b> Expression of human CCR5 on various cell types (LEDERMAN <i>et al.</i> 2006; MENTEN <i>et al.</i> 2002).....	7
<b>Table 1.2.</b> Distribution of <i>CCR5</i> mutations resulting in amino acid substitutions in different populations.....	14
<b>Table 2.1.</b> Frequencies of identified polymorphisms within the South African African (SAA) and Caucasian (SAC) study populations .....	37
<b>Table 3.1.</b> Mean relative luciferase activity of the various <i>CCR5</i> promoter constructs transfected into K562, U937, THP-1 and Jurkat cell lines .....	58
<b>Table 4.1.</b> Differences in CCR5 expression, both as percentage of CCR5-expressing cells and CCR5 density, between two South African population groups, South African African (SAA) (n=22) and South African Caucasian (SAC) (n=31) across all peripheral blood cell populations* .....	76
<b>Table 4.2.</b> Correlation between cell activation, as measured by HLA-DR percentage expression, and CCR5 expression (percentage CCR5-expressing cells and density) on T cell subsets* .....	79
<b>Table 7.1.</b> Characteristics of HIV-1-infected long term nonprogressing and HIV-1-infected progressing cohorts*.....	126
<b>Table 7.2.</b> Frequencies of identified polymorphisms within the South African African (SAA) and Caucasian (SAC) study populations .....	129
<b>Table 7.3.</b> CCR5 expression, both as percentage of CCR5-expressing cells and CCR5 density, in HIV-1 infected long term nonprogressors (LTNP) from two South African population groups, South African African (SAA) (n=10) and South African Caucasian (SAC) (n=4) across all peripheral blood cell populations.....	133
<b>Table 7.4.</b> Genotype matched comparison of CCR5 expression, both as percentage of CCR5-expressing cells and CCR5 density, between HIV-1 infected long term nonprogressors (LTNP) and healthy HIV-1 uninfected South African African (SAA) individuals across all peripheral blood cell populations .....	137

<b>Table A.1.</b> <i>CCR5</i> primer sets and experimental conditions for polymerase chain reaction amplification of overlapping regions of gene.....	190
<b>Table A.2.</b> Flanking sequences of all SNPs detected in this study .....	191
<b>Table B.1.</b> Primer sets and experimental conditions for polymerase chain reaction (PCR) amplification of <i>CCR5</i> promoter construct regions used to assess promoter efficiency.....	193
<b>Table B.2.</b> Primer pairs used to screen clones for directional cloning of promoter fragments.....	193
<b>Table D.1.</b> Comparison of CCL3 production in South African African (SAA) and Caucasian (SAC) individuals as measured in CCL3 production assays .....	195
<b>Table D.2.</b> Comparison of CCL3 production of phytohaemagglutinin stimulated cord blood mononuclear cells (CBMCs) isolated from HIV-1 intrapartum infected (IP) and exposed uninfected (EU) infants.....	195
<b>Table E.1.</b> Characteristics of individual HIV-1-infected long term nonprogressors and HIV-1-infected progressing study participants.....	196
<b>Table E.2.</b> Comparison of CD4 counts determined by FACSCount with those determined with FACSCalibur .....	197
<b>Table E.3.</b> Comparison of CD4 and CD8 counts and CD4:CD8 ratio between control uninfected and HIV-1-infected long term nonprogressing (LTNP) individuals in the South African African (SAA) and South African Caucasian (SAC) populations .....	197

---

## LIST OF FIGURES

---

<b>Figure 1.1.</b> The process by which HIV-1 gains entry into a host target cell. ....	4
<b>Figure 1.2.</b> The two dimensional topology of the human CCR5 illustrating the amino acid sequence and seven transmembrane structure. ....	7
<b>Figure 1.3.</b> Structure of the <i>CCR5</i> gene and its messenger RNA transcripts. ....	10
<b>Figure 1.4.</b> Chemokines are classified into four different classes based on the spacing of their first two cysteine residues. ....	22
<b>Figure 1.5.</b> Genomic organisation and mRNA products of <i>CCL3</i> and <i>CCL3L</i> genes. ....	24
<b>Figure 2.1.</b> Schematic representation of <i>CCR5</i> gene region sequences indicating polymorphism positions and nucleotide base changes at these points. ....	39
<b>Figure 2.2.</b> (A) Schematic representation of haplotypes previously defined by Gonzalez <i>et al.</i> (1999) and the frequency at which they were detected within our two study populations. (B) Haplotypes identified within the <i>CCR5</i> gene in the South African African (SAA) and South African Caucasian (SAC) study populations. ....	42
<b>Figure 3.1.</b> Schematic representation of <i>CCR5</i> gene regions used for promoter expression analysis. ....	53
<b>Figure 3.2.</b> Promoter expression strength of the different constructs, P1A, P1B, P2 and P1+P2, relative to each other. ....	59
<b>Figure 3.3.</b> Haplotype differences in promoter efficiency between P1A promoter constructs. ....	60
<b>Figure 3.4.</b> Haplotype differences in promoter efficiency between P1B, i.e., <i>CCR5</i> downstream promoter and intron2, promoter constructs. ....	61
<b>Figure 3.5.</b> Haplotype differences in promoter efficiency between P2, i.e., <i>CCR5</i> upstream promoter, promoter constructs. ....	62
<b>Figure 3.6.</b> Haplotype differences in promoter efficiency between promoter constructs spanning the entire <i>CCR5</i> promoter region, i.e., P1+P2. ....	63
<b>Figure 3.7.</b> Effect of the -2150A/G polymorphism on promoter activity of SAA-HHA promoter constructs in four different cell lines tested. ....	64

<b>Figure 3.8.</b> Putative MOK-2 transcription factor binding site located within the <i>CCR5</i> exon 3 region of the SAA-HHA allele containing the -2150G mutation. ....	65
<b>Figure 4.1.</b> <i>CCR5</i> expression in South African African (SAA) (n=22) and South African Caucasian (SAC) (n=24) individuals within natural killer (NK) cell subsets. ....	76
<b>Figure 4.2.</b> Correlation between <i>CCR5</i> density (mean number of molecules per cell) on monocyte cells and age of SAA study participants (n=22). ....	77
<b>Figure 4.3.</b> Activation levels, as demonstrated by percentage HLA-DR expression in the two study population groups, South African African (SAA) (n=16) and South African Caucasian (SAC) (n=23) within natural killer (NK) and T cell subsets. ....	78
<b>Figure 4.4.</b> Influence of <i>CCR5</i> $\Delta$ 32 bearing alleles on <i>CCR5</i> expression in South African Caucasian (SAC) population within all cell populations examined. ....	80
<b>Figure 5.1.</b> Distribution of <i>CCR5</i> haplotypes and genotypes within our two study populations, South African African (SAA) and South African Caucasian (SAC). ....	90
<b>Figure 5.2.</b> Influence of the <i>CCR5</i> HHA haplotype on <i>CCR5</i> expression in SAA individuals within the CD8+ (A) and CD4+ (B) T cell subsets. ....	93
<b>Figure 5.3.</b> Influence of the HHC haplotype on the percentage of <i>CCR5</i> -expressing cells in T cell subsets, CD8+ (i) and CD4+ (ii), in South African Caucasian (SAC) individuals (A), SAC population with the exclusion of individuals with the <i>CCR5</i> $\Delta$ 32 allele (B), and in South African African (SAA) individuals (C). ....	95
<b>Figure 5.4.</b> Influence of the HHC haplotype on <i>CCR5</i> density on natural killer cell subsets, CD56+ NK (i) and CD16+CD56+ NK (ii), within the South African Caucasian (SAC) population (A), the SAC population with the exclusion of individuals with the <i>CCR5</i> $\Delta$ 32 allele (B) and the SAA population (C). ....	97
<b>Figure 6.1.</b> Concordance between different real time assays to assess gene copy number, <i>CCL3L</i> and <i>CCL3La</i> + <i>CCL3Lb</i> illustrated using both rounded (A) and unrounded values (B). ....	110
<b>Figure 6.2.</b> Differences in <i>CCL3L</i> , <i>CCL3La</i> and <i>CCL3Lb</i> gene copy number distribution between healthy HIV-1 uninfected individuals from the two study populations, South African African (SAA) (n=89) and South African Caucasian (SAC) (n=54). ....	111

<b>Figure 6.3.</b> Phytohaemagglutinin-induced CCL3 production of South African African (SAA) (A) and South African Caucasian (SAC) (B) individuals stratified around the population specific <i>CCL3La</i> copy number median (four and two, respectively).....	112
<b>Figure 6.4.</b> CCL3 production in South African African (SAA) and South African Caucasian (SAC) individuals as measured in CCL3 production assays. ....	113
<b>Figure 6.5.</b> Correlation between calculated CCL3 production per <i>CCL3La</i> copy and <i>CCL3La</i> copy number for the two populations, South African African (SAA) and South African Caucasian (SAC), examined independently (A and B) or combined (C).....	114
<b>Figure 6.6.</b> Comparison of CCL3 production per <i>CCL3La</i> copy number between South African African (SAA) and South African Caucasian (SAC) individuals with comparable <i>CCL3La</i> and <i>CCL3Lb</i> copy numbers.....	115
<b>Figure 6.7.</b> CCL3 production of phytohaemagglutinin stimulated cord blood mononuclear cells (CBMCs) isolated from HIV-1 intrapartum infected (IP) and exposed uninfected (EU) infants. ....	116
<b>Figure 7.1.</b> Haplotype and genotype distribution in HIV-1 infected long term non-progressing (A) and HIV-1 infected progressing individuals (B).....	131
<b>Figure 7.2.</b> CCR5 density on CD4+ (i) and CD8+ (ii) T cells in healthy HIV-1 uninfected and HIV-1 infected long term nonprogressing (LTNP) South African African (SAA) (A) and South African Caucasian (SAC) (B) individuals. ....	134
<b>Figure 7.3.</b> CCR5 density on monocytes in healthy HIV-1 uninfected and HIV-1 infected long term nonprogressing (LTNP) South African African (SAA) (A) and South African Caucasian (SAC) (B) individuals. ....	135
<b>Figure 7.4.</b> Percentage of CCR5-expressing cells within T cell subsets, CD4+ (i) and CD8+ (ii), in healthy HIV-1 uninfected and HIV-1 infected long term nonprogressing (LTNP) South African African (SAA) (A) and South African Caucasian (SAC) (B) individuals. ....	136
<b>Figure 7.5.</b> CCR5 density on monocytes in healthy HIV-1 uninfected and HIV-1 infected long term nonprogressing (LTNP) South African African (SAA) individuals matched for <i>CCR5</i> genotype.....	138

<b>Figure 7.6.</b> Cell activation levels, as measured by the percentage of HLA-DR-expressing cells, in CD4+ (i) and CD8+ (ii) T cells, in healthy HIV-1 uninfected and HIV-1 infected long term nonprogressing (LTNP) South African African (SAA) (A) and South African Caucasian (SAC) (B) individuals. ....	140
<b>Figure 7.7.</b> CCL3 production by peripheral blood mononuclear cells (PBMCs) isolated from HIV-1-infected long term nonprogressors (LTNPs) and healthy HIV-1 uninfected individuals as measured in CCL3 production assays. ....	142
<b>Figure 7.8.</b> CCL3 production by peripheral blood mononuclear cells (PBMCs) isolated from HIV-1 infected long term nonprogressors (LTNPs) stratified according to low (<400 copies/ml) and higher (>400 copies/ml) viral loads. ....	143
<b>Figure 7.9.</b> Comparison of absolute counts of monocytes and lymphocyte cell subsets from South African African (SAA) (n=22) and South African Caucasian (SAC) (n=31) healthy HIV-1 uninfected individuals. ....	144
<b>Figure 7.10</b> Distribution of different cell types within peripheral blood mononuclear cells (PBMCs) isolated from South African African (SAA) HIV-1 uninfected control (n=22) and SAA long term nonprogressing (LTNP) (n=10) individuals. ....	146
<b>Figure C.1.</b> Influence of the HHC haplotype on the percentage of CCR5-expressing cells on natural killer cell subsets, NK (i) and CD16+CD56+ NK (ii), within the SAC population (A), the SAC population with the exclusion of individuals with the <i>CCR5Δ32</i> allele (B) and the SAA population (C). ....	194
<b>Figure E.1.</b> Comparison of extrapolated absolute cell counts between control and HIV-1 infected long term nonprogressing (LTNP) individuals from two populations, SAA and SAC. ....	198

---

## ABBREVIATIONS

---

A	Adenine
Ab	Antibody
ABC	Antibodies bound per cell
AIDS	Acquired Immunodeficiency Syndrome
AP-1	Activator protein 1
APC	Allophycocyanin
APOBEC3G	Apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like-3G
ART	Antiretroviral therapy
bp	base pair
°C	Degrees Celcius
C	Cytosine
CBMC	Cord blood mononuclear cell
CCL3	Chemokine (C-C motif) ligand 3 or MIP-1 $\alpha$
CCL3L	C-C motif chemokine ligand 3-like
CCL4	Chemokine (C-C motif) ligand 4 or MIP-1 $\beta$
CCL4L	C-C motif chemokine ligand 4-like
CCL5	Chemokine (C-C motif) ligand 5 or RANTES
CCL8	Chemokine (C-C motif) ligand 8 or MCP-2
CCR	C-C chemokine receptor
CCR5	C-C chemokine receptor 5
CCR5 $\Delta$ 32	CCR5 delta 32 (32 bp deletion)
CD	Cluster of differentiation
CT	Cycle threshold
CXCR4	CXC-chemokine receptor 4
DC-SIGN	Dendritic cell-specific intercellular adhesion molecular 3-grabbing non-integrin
DNA	Deoxyribonucleic acid
DPBS	Dulbecco's Phosphate Buffered Saline
EC	Elite controller
EDTA	Ethylenediaminetetraacetic acid
eg.	Example
Egr-1	Early growth response protein 1
Env	HIV viral envelope glycoprotein
et al.	And others
EU	Exposed uninfected
FACS $\textcircled{R}$	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FITC	Fluorescein isothiocyanate
FL2	Fluorescence 2
FOXC1	Forkhead box C1
FSC	Forward Scatter
G	Guanine
g	g-force
GATA-1	Erythroid transcription factor or GATA-binding factor 1
GBV	GB type C virus, hepatitis G virus
h	Hours
HAART	Highly active antiretroviral therapy
HBB	Human beta-globin
HESN	Highly exposed seronegative
HHA-G*2	Human haplotype A to G*2 on CCR5 gene
HIV-1	Human Immunodeficiency Virus type 1
HIV-2	Human Immunodeficiency Virus type 2
HLA	Human Leukocyte Antigens
Hoxb-9	Homeobox B9
HSP70	70 kDa Heat shock protein
HSV-2	Herpes simplex virus 2
HTLV-2	Human T-cell leukemia virus 2
i.e.	That is

IMDM	Iscove's Modified Dulbecco's Medium
IP	Intrapartum infected
kb	Kilo bases
kDa	Kilo Daltons
KLF2	Krueppel-like factor 2
LNA	Lock nucleic acid
LPS	Lipopolysaccharide
LTNP	Long term nonprogressor
M	Molar
mAB	Monoclonal antibody
MCP-2	Monocytic chemotactic protein-2 or CCL8
MFI	Mean/median fluorescence intensity
min	Minutes
MHC	Major histocompatibility complex
MIP-1 $\alpha$	Macrophage inflammatory protein-1 $\alpha$ or CCL3
MIP-1 $\beta$	Macrophage inflammatory protein-1 $\beta$ or CCL4
MOX-2	Mesenchyme homeobox 2 or Homeobox protein MOX-2
mRNA	Messenger RNA
MUT	Mutant
NF-AT	Nuclear factor of activated T cells
NI	Newly Identified
NF $\kappa$ B	Nuclear factor of kappa-light-chain-enhancer of activated B cells
NK	Natural killer
Oct-1	Octamer transcription factor 1 or POU2F1
Oct-1	Octamer transcription factor 2 or POU2F2
ORF	Open Reading Frame
p65(RelA)	Transcription factor p65 or nuclear factor NF $\kappa$ B p65 subunit
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PD or P1	CCR5 downstream promoter
pdg	Per diploid genome
PE	Phosphatidylethanolamine
PerCP	Peridinin chlorophyll protein
PHA	Phytohaemagglutinin
PRT	Paralogue ratio test
PU or P2	CCR5 upstream promoter
qPCR	Quantitative real-time polymerase chain reaction
RANTES	Regulated upon Activation, Normal T-cell Expressed, and Secreted or CCL5
RNA	Ribonucleic acid
RPMI	Roswell Park Memorial Institute medium
SAA	South African Africans
SAC	South African Caucasians
SIV	Simian Immunodeficiency Virus
SNP	Single nucleotide polymorphism
SSC	Side Scatter
STAT	Signal-transducer and activator of transcription
SV40	Simian Virus 40
T	Thymine
T	Temperature
TK	Thymidine kinase
TRIM5a	Tripartite motif-containing protein 5 $\alpha$
UTR	Untranslated region
vs	Versus
WT	Wild type

---

## PREFACE

---

Some aspects of the work conducted for this thesis have been published or presented in the form of papers and poster presentations:

### **Publications:**

#### *Chapter 2:*

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

#### *Chapter 4:*

Picton, A. C. P., Shalekoff, S., Paximadis, M. and Tiemessen, C. T. (2012). Marked differences in CCR5 expression and activation levels in two South African populations. *Immunology* **136**: 397-407.

#### *Chapter 5:*

Picton, A. C. P., Paximadis, M. and Tiemessen, C. T. (2012). CCR5 promoter haplotypes differentially influence CCR5 expression on natural killer and T cell subsets in ethnically divergent HIV-1 uninfected South African populations. *Immunogenetics* **64**: 795-806.

#### *Chapter 6:*

Picton, A. C. P., Paximadis, M. and Tiemessen, C. T. (2013). Contribution of variable CCL3L copy number to CCL3 protein production in two ethnically divergent South African populations. *Infection, Genetics and Evolution* **14**: 347-356.

### **Presentations:**

#### *Chapter 2:*

Picton, A. C. P., Paximadis, M. and Tiemessen, C. T. Characterization of *CCR5* genes within South African individuals. NICD Academic day, 11 November 2008, James Gear Auditorium, NICD, Johannesburg. (Poster presentation)

Picton, A. C. P, Paximadis, M. and Tiemessen, C. T. Genetic variation within the gene encoding the HIV-1 CCR5 coreceptor in two South African population groups. AIDS Vaccine 2009. 19-22 October 2009. Marriott Rive Gauche Hotel and Conference Center, Paris, France. (Poster presentation)

*Chapter 4:*

Picton, A. C. P, Shalekoff, S., Paximadis, M. and Tiemessen, C. T. CCR5 Expression on Peripheral Blood Cells Differs between South African Caucasian and African Populations. AIDS Vaccine 2011. 12-15 September 2011. Patumwan, Bangkok, Thailand. (Poster presentation)

Picton, A. C. P, Shalekoff, S., Paximadis, M. and Tiemessen, C. T. CCR5 Expression on Peripheral Blood Cells Differs between South African Caucasian and African Populations. Virology Africa 2011. 29 November-2 December 2011. UCT Graduate School of Business, V&A Waterfront, Cape Town, South Africa. (Poster presentation)

---

# CHAPTER 1

## Introduction

---

In 1981, the first reports of an immune disorder characterized by a decline of immune function and of T cells, most notably CD4+ T cells, were made (GOTTLIEB *et al.* 1981; SIEGAL *et al.* 1981). Shortly afterwards (1983-1984) the lentivirus, human immunodeficiency virus (HIV), was identified in patients with acquired immunodeficiency syndrome (AIDS) (BARRE-SINOUSI *et al.* 1983; GALLO *et al.* 1984; POPOVIC *et al.* 1984). Two types of HIV have been described, HIV-1 and HIV-2, both of which are thought to have originated in Africa from non-human primates. HIV-1 infection is more prevalent and also more pathogenic than HIV-2. HIV-1 is closely related to the simian immunodeficiency virus (SIV) which naturally infects chimpanzees (SIV<sub>CPZ</sub>), while HIV-2 is more closely related to the SIV infecting sooty mangabees (SIV<sub>SM</sub>) (GAO *et al.* 1999; HIRSCH *et al.* 1989).

It is estimated that since the beginning of the AIDS epidemic, more than 60 million individuals have been infected with HIV and nearly 30 million have died of HIV-related causes ([www.unaids.org](http://www.unaids.org)). Although the survival of patients infected with HIV has considerably improved since the introduction of highly active antiretroviral therapy (HAART) (MICHAELS *et al.* 1998; PALELLA *et al.* 1998), and progress has been made in numerous preventive approaches including behavioral interventions, male circumcision, pre- and post- exposure prophylaxis, vaccines, and microbicides, HIV infection remains a major global health problem, with an estimated 34 million people currently infected worldwide (UNAIDS 2012a).

Extensive variation exists between individuals, in their susceptibility, or vulnerability, to HIV-1 infection and the rate of disease progression to AIDS. Although rare, there are a number of individuals who, despite repeated exposure to HIV-1, remain uninfected. These individuals are referred to as HIV-exposed seronegative (HESN) or exposed uninfected (EU) individuals and include individuals exposed to HIV through a number of routes, including sexual intercourse (discordant couples and high risk sex practitioners) (BERETTA *et al.* 1996; CLERICI *et al.* 1992; FOWKE *et al.* 1996), direct blood exposure (intravenous drug users and health care workers (PINTO *et al.* 1995; SCOTT-ALGARA *et al.* 2003), and infants born to infected mothers (TIEMESSEN *et al.* 2009).

After HIV-1 infection, individuals remain asymptomatic for an average of 8-10 years. However, the period of asymptomatic disease varies greatly between infected individuals and has led to the use of various definitions to describe patient subsets, ranging from rapid progressors to elite controllers. These differences between individuals can be partly attributed to genetic variation at loci associated with either cellular entry of HIV-1, the expression of ligands competing for the same receptors, immune recognition or antigen presentation (Reviewed in ARENZANA-SEISDEDOS and PARMENTIER 2006). However, no single gene or polymorphism is likely to be responsible for these differences. A number of genes and/or gene mutations which have either a positive or negative effect on infection and disease progression have been identified. These include genes coding for

cytokines, chemokines, chemokine receptors, human leukocyte antigens (HLAs) and intracellular “restriction factors” such as Tripartite motif-containing protein 5 $\alpha$  (TRIM5 $\alpha$ ) and Apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like-3G (APOBEC3G) (Reviewed in HEENEY *et al.* 2006; NOLAN *et al.* 2004; O'BRIEN and NELSON 2004). Thus, variation in the genes involved in both the adaptive and innate arms of the immune system seem to be contributing to differential HIV-1 susceptibility and disease progression.

Until 1996, the mechanism through which HIV-1 gained entry into target cells remained unclear. Cocchi *et al.* (1995) demonstrated that the  $\beta$ -chemokines, macrophage inflammatory protein-1 $\alpha$  (MIP-1 $\alpha$ , or chemokine ligand 3 (CCL3)), MIP-1 $\beta$  (CCL4) and RANTES (Regulated upon Activation Normal T-cell Expressed and Secreted, or CCL5) can mediate antiviral effects by blocking HIV-1 infection of CD4+ lymphocytes. This led to the discovery that HIV-1 must interact with CD4 (primary receptor) and a second chemokine receptor, either CCR5 or CXCR4, in order for HIV-1 to gain entry into cells (DENG *et al.* 1996; DRAGIC *et al.* 1996; FENG *et al.* 1996).

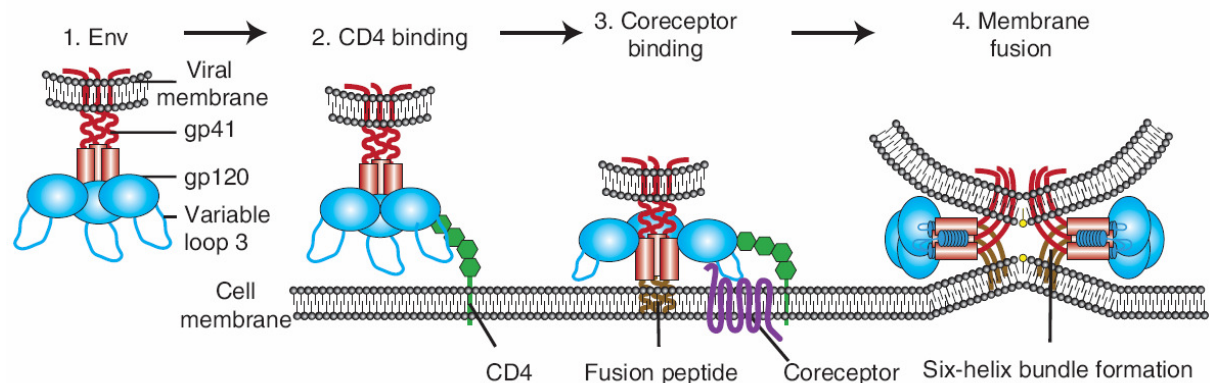
### **1.1. MECHANISM OF CELLULAR ENTRY OF HIV**

HIV-1 has been identified to have a wide cellular host range, infecting several cell types of the brain, bowel, heart, kidney, liver, testes, prostate and other organs in addition to haematopoietic cells (Reviewed in LEVY 2006).

The first step of the HIV-1 replication cycle involves binding to a host cell and subsequent entry into the target cell (Figure 1.1). This is initiated with the adhesion of virus to the host cell. Initial adhesion to a target cell can be mediated through the interaction of the viral envelope glycoproteins (Env) with a number of different molecules such as negatively charged cell-surface heparin sulfate proteoglycans,  $\alpha$ 4 $\beta$ 7 integrin, or pattern recognition receptors such as dendritic cell-specific intercellular adhesion molecular 3-grabbing non-integrin (DC-SIGN) (Reviewed by WILEN *et al.* 2012). HIV-1 virions enter cells via interaction of Env with the primary receptor, CD4, and a coreceptor, CCR5 or CXCR4, so as to infect cells. Env is highly glycosylated and is organized into trimers on virions consisting of the gp120 surface and gp41 transmembrane subunits. HIV-1 entry into target cells is initiated by a high affinity interaction between the gp120 subunit and CD4 (MADDON *et al.* 1986; MCDUGAL *et al.* 1986). This induces a conformational change in gp120 which exposes the chemokine receptor binding site of gp120, thus allowing for coreceptor engagement (KWONG *et al.* 1998). Attached virions may then move along the cell surface to specific destinations where membrane fusion can occur (Reviewed in WILEN *et al.* 2012). Binding to the coreceptor is followed by membrane fusion mediated by Env. The HIV-1 virion is brought into close contact with the target cell and the hydrophobic gp41 fusion peptide is inserted into the

host cell membrane, followed by the formation of a six-helix bundle, complete membrane fusion and subsequent entry of the HIV-1 viral capsid into the host cell (CHAN and KIM 1998).

CCR5 coreceptor binding HIV-1 viruses are termed R5 viruses, those which use the CXCR4 coreceptor are termed X4 viruses, and viruses which use both coreceptors are called R5X4 viruses (BERGER *et al.* 1998). Genetic sequences within the HIV gp120 subunits, particularly within the highly variable and structurally flexible V3 region involved in coreceptor binding, determine the coreceptor selectivity of HIV (BRIGGS *et al.* 2000; FOUCHIER *et al.* 1992). HIV-1 infection is predominantly established by R5 strains.



**Figure 1.1.** The process by which HIV-1 gains entry into a host target cell. The HIV viral envelope glycoprotein (Env), comprised of the gp120 and gp41 subunits (1), binds the primary host receptor, CD4 (2). This induces a conformational change in Env which allows the virus to bind to a coreceptor, CCR5 or CXCR4 (3) which, in turn, initiates the membrane fusion process whereby the fusion peptide of gp41 inserts into the host cell membrane, followed by the formation of a six-helix bundle (formed by each of the three gp41 molecules folding upon itself) and complete membrane fusion (4). Image source: (WILEN *et al.* 2012).

## 1.2. CHEMOKINE RECEPTORS

Chemokine receptors are a group of transmembrane proteins that belong to the superfamily of G-protein coupled receptors and have attracted interest as they play a role in the cellular entry of the human immunodeficiency viruses (HIV-1 and HIV-2) and other related simian or feline retroviruses. In addition to the CD4 receptor, HIV-1 R5 and X4 strains require that the chemokine receptors, namely CCR5 and CXCR4 respectively, be expressed on the surface of cells in order to gain entry into the target cell (DENG *et al.* 1996; DRAGIC *et al.* 1996). CCR5 and CXCR4 are the main HIV-1 coreceptors, however, numerous other alternate coreceptors are able to support HIV-1 infection and replication. These include, among others, CCR1, CCR2b, CCR3, CCR4, CCR6, CCR8, CXCR6, CXCR3, CXCR7, GPR1 and GPR15 (Reviewed in POLLAKIS and PAXTON 2012). The

ability of modified ligands to bind to chemokine receptors promoting the clearance of HIV binding sites from the cell surface has opened a new field of vaccination strategies (LEHNER 2002).

### 1.2.1. CCR5

Binding of chemokines to the CCR5 receptor (also known as CD195) induces migration of the receptor-bearing cells along a gradient of increasing chemokine concentration and results in the recruitment of leukocytes to the site of inflammation. CCR5 has multiple ligands. Of these, CCL3, CCL4, CCL5 and CCL8 (or monocyte chemotactic protein-2, MCP-2), demonstrate the most suppressive activities in HIV-1 infection assays (BLANPAIN *et al.* 1999c). The ligands, CCL3 and CCL5, bind to other chemokine receptors, while CCL4 is known to be the most specific chemokine for CCR5.

Shortly after the identification of CCR5 as an HIV-1 coreceptor, an allele consisting of a 32 base pair (bp) deletion within the *CCR5* open reading frame (*CCR5 $\Delta$ 32*) was identified in a small number of individuals that had remained seronegative despite being at high risk for acquiring infection (DEAN *et al.* 1996; LIU *et al.* 1996; SAMSON *et al.* 1996b). HIV-1 infected individuals heterozygous for this mutation were also demonstrated to progress slower to AIDS and have significantly higher representation in cohorts of HIV-1 infected long term nonprogressors compared to normal progressors (EUGEN-OLSEN *et al.* 1997). The central role of CCR5 in HIV-1 infection has been further highlighted in a report of an HIV-1 infected man (known as “The Berlin patient”) who, after receiving a transplant of haematopoietic stem cells from a donor homozygous for the *CCR5 $\Delta$ 32* mutation and ceasing HIV therapy, remained free of detectable HIV for now more than five years (DEEKS *et al.* 2012; HUTTER *et al.* 2009).

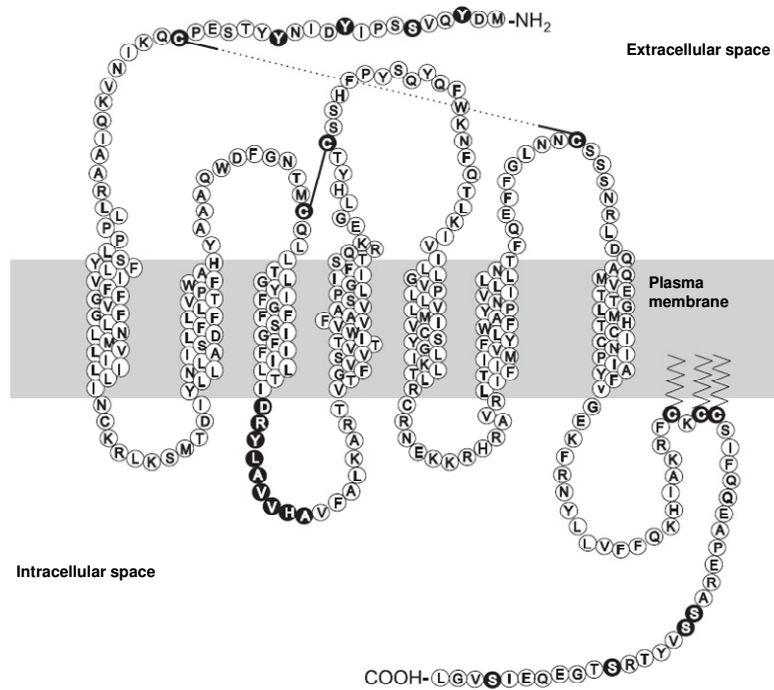
#### 1.2.1.1. CCR5 antagonists for HIV therapy

Although antiretroviral therapies have proven effective at reducing HIV-1 RNA to undetectable levels, HIV-1 infection persists due to viral integration into the host genome and the establishment of viral reservoirs. Antiretroviral therapy (ART) can be costly, requires ongoing medical care, can have toxic side effects and may also result in the emergence of resistant viral strains. Thus, a number of drugs targeting the CCR5 coreceptor, CCR5 antagonists, have been developed. To date, although several antagonists have been developed, maraviroc (Celsentri<sup>®</sup>, Pfizer) is the only CCR5 blocker which has been approved for HIV-1 treatment in combination with ART, the treatment of patients experiencing virological failure due to resistance to other classes of antiretroviral drugs and for the treatment of ART naïve patients (Reviewed in BOESECKE and PETT 2012; CHOI and AN 2011).

### 1.2.1.2. The CCR5 protein

The chemokine receptor, CCR5, belongs to the G-protein-coupled receptor superfamily and is composed of 7 hydrophobic trans-membrane domains, an extracellular N-terminus, an intracellular cytoplasmic C-terminal and alternating extracellular and intracellular loops (SAMSON *et al.* 1996c) (Figure 1.2). The protein consists of 352 amino acids with a calculated molecular mass of 40.6 kDa (SAMSON *et al.* 1996c) and is expressed on a number of cells, including immune effector cells and antigen presenting cells, as listed in Table 1.1. The C-terminus contains structural motifs which are critical for ligand-dependent signaling, desensitization and receptor trafficking (Reviewed in OPPERMANN 2004). The amino-terminal domain of CCR5 is essential for chemokine binding and also plays a role in its ability to support HIV-1 entry (BLANPAIN *et al.* 1999a). Two disulphide bonds, between cysteine residues at amino acid positions 20 and 269 and positions 178 and 101, which are essential for maintaining the structural integrity of the protein as well as for ligand binding, have been identified (BLANPAIN *et al.* 1999b) (Figure 1.2).

Although the structure of CCR5 has to date not been resolved, it is thought to resemble that of rhodopsin (PALCZEWSKI *et al.* 2000). Two distinct CCR5 protein forms with differing molecular mass, 62 and 42 kDa, have been found to exist in both human lymphocytic cell lines and monkey peripheral mononuclear cells (SUZUKI *et al.* 2002). These two protein forms also differ in their cellular localization, the 62 kDa CCR5 resides mainly on the cell membrane while the 42 kDa CCR5 is exclusively found in cellular cytoplasm (SUZUKI *et al.* 2002). Through study of the binding affinity of anti-CCR5 antibodies which recognize different CCR5 epitopes, it has been demonstrated that CCR5 exists in multiple conformational states on the surface of cells (BERRO *et al.* 2011; LEE *et al.* 1999a).



**Figure 1.2.** The two dimensional topology of the human CCR5 illustrating the amino acid sequence and seven transmembrane structure. The grey box indicates the approximate position of the membrane bilayer. Amino acids critical for CCR5 function are highlighted in black. The position of the two disulphide bonds are indicated with solid and dotted lines (Image adapted from OPPERMANN 2004).

**Table 1.1.** Expression of human CCR5 on various cell types (LEDERMAN *et al.* 2006; MENTEN *et al.* 2002)

Cell Type	CCR5	Cell Type	CCR5
Monocytes/macrophages	+	CD34+ bone marrow cells	+
Lymphocytes		Neutrophils	-
CD3+	+	Eosinophils	-
CD4+	+	Basophils	+
CD8+	+	Dendritic cells	+
Memory	+	Langerhans cells	+
Naïve	± <sup>a</sup>	Platelets	-
NK cells	+	Neurons	+
NK T cells	+	Astrocytes	+
B cells	±	Fibroblasts	+
Th1	+	Smooth muscle cells	+
Th2	-	Capillary endothelial cells	+
Thymocytes	+		

<sup>a</sup>(±) Low expression

### 1.2.1.3. The role of CCR5 in the immune system

CCR5 plays an important role in linking the innate and the adaptive immune system. At the site of microbial invasion, binding of certain microbial products (such as endotoxin, lipopolysaccharide, flagellin, peptidoglycan, RNA and DNA) to toll-like receptors on macrophages, immature dendritic cells, and other cells, induces the activation of these cells and subsequent release of cytokines, which include chemokine ligands for CCR5 (LEDERMAN *et al.* 2006; LUSTER 2002). Activation of CCR5-expressing immature dendritic cells, leads to the maturation of the dendritic cells accompanied by downmodulation of CCR5 expression and simultaneous upregulation of CCR7 expression (SOZZANI *et al.* 1998). The CCR7 receptor promotes homing of these cells to lymphoid tissue where mature dendritic cells are able to present foreign antigens to T and B cells, thus initiating specific immune responses.

In adaptive immune responses, the release of CCR5 ligands induces migration of CCR5-expressing cells towards the site of microbial infection where they may contribute towards the control and/or elimination of the foreign microbe. Inflammation results in the upregulated expression of CCR5 molecules by antigen-naïve CD8<sup>+</sup> T cells in lymph nodes (CASTELLINO *et al.* 2006). This allows for these cells to be attracted actively to sites of productive CD4<sup>+</sup> T cell-dendritic cell interaction where the chemokines, CCL3 and CCL4, are produced and thus promoting the generation of optimally activated antigen-specific CD8<sup>+</sup> T cells, i.e., development of long-term CD8<sup>+</sup> T cell memory (CASTELLINO *et al.* 2006).

During T cell stimulation, the chemokine receptors, CCR5 and CXCR4, are recruited to and accumulate at the immunological synapse (MOLON *et al.* 2005). Furthermore, simultaneous expression and cooperation between the two coreceptors is required for chemokine-induced T cell costimulation at the immunological synapse and it has been suggested that these two molecules hetero-oligomerize to allow signal versatility in T cells (CONTENTO *et al.* 2008). Thus, CCR5 appears to be important for both CD8<sup>+</sup> T cell recruitment and for costimulatory signals required for efficient T cell priming.

Interestingly, CCR5 can act as a receptor for microbial 70 kDa heat shock protein (HSP70) (WHITTALL *et al.* 2006). Binding of either HSP70 or CCL3 to CCR5 has been demonstrated to result in the *in vitro* upregulation of the APOBEC3G enzyme in CD4<sup>+</sup> T and dendritic cells (PIDO-LOPEZ *et al.* 2007). APOBEC3G inhibits viral replication, particularly against retroviruses, thus making it an important component of intracellular innate immunity. This demonstrates another important role for CCR5 in the innate immune system. However, HIV-1 has evolved the virion infectivity factor (Vif) protein to counteract the effect of APOBEC3G and thus APOBEC3G is only effective at restricting HIV-1 replication in the absence of Vif (DONAHUE *et al.* 2008).

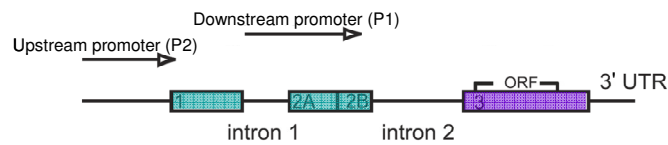
#### 1.2.1.4. CCR5 gene

The *CCR5* gene is located in the short arm of the third chromosome (3p21.3 region) of the human genome in close proximity to other chemokine receptor genes, *CCR1*, *CCR2*, *CCR3* and *CCR4* (LIU *et al.* 1998; SAMSON *et al.* 1996c). It is composed of four exons and two introns, where exons 2A and 2B are not interrupted by an intron (MUMMIDI *et al.* 1997) (Figure 1.3A). Exons 1 and 2 contain most of the 5' untranslated regions. Exon 3 contains 11 base pairs of the 5'-untranslated region (UTR) and the intronless open reading frame (ORF) (MORIUCHI *et al.* 1997; SAMSON *et al.* 1996a).

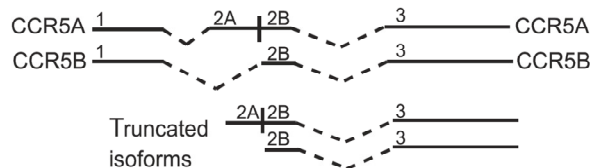
Two *CCR5* promoters have been described for the *CCR5* gene, an upstream promoter ( $P_U$ ) which lies upstream of exon 1 and a downstream promoter ( $P_D$ ) which lies in the region between intron 1 and intron 2 (GUIGNARD *et al.* 1998; LIU *et al.* 1998; MORIUCHI *et al.* 1997; MUMMIDI *et al.* 1997) (Figure 1.3A). Alternative splicing, as well as multiple transcription start sites, results in the production of several *CCR5* mRNA transcripts differing in their 5'-untranslated regions (UTR) (Figure 1.3B) (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 P2 promoter 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 1.3B) (MUMMIDI *et al.* 1997). Transcripts generated from the downstream promoter are referred to as “truncated” isoforms (Figure 1.3B) (MUMMIDI *et al.* 1997).

Promoter expression studies conducted in transformed cell lines have demonstrated the downstream promoter ( $P_1$ ) to be the stronger of the two promoters (LIU *et al.* 1998; MUMMIDI *et al.* 1997). If unstimulated, both isolated peripheral blood mononuclear cells (PBMCs) and purified CD4<sup>+</sup> T cells contain a greater proportion of  $P_1$  initiated *CCR5* mRNA transcripts than  $P_2$  initiated *CCR5* mRNA transcripts (MUMMIDI *et al.* 2007). However, the  $P_2$ -driven transcripts, *CCR5A* and *CCR5B*, have been demonstrated as important determinants of *CCR5* expression on activated primary T cells, i.e., stimulation of T cells results in the induction of the  $P_2$  promoter and consequently the expression of exon 1-containing mRNA transcripts (MUMMIDI *et al.* 2007). Importantly, increased expression of *CCR5A* and *CCR5B* transcripts in memory T cells was found only in those cells that also expressed *CCR5* on the surface (MUMMIDI *et al.* 2007).

### A. *CCR5* gene:



### B. *CCR5* mRNA:



**Figure 1.3.** Structure of the *CCR5* gene and its messenger RNA transcripts. **(A)** Diagrammatic representation of the *CCR5* gene organization. *CCR5* is comprised of three exons represented by coloured boxes. Exons 1, 2A, 2B, and 11 bp of exon 3 encode the 5' untranslated region (UTR), while exon 3 encodes the open reading frame (ORF) and the 3'UTR. Approximate positions of the two promoter regions, P1 and P2, are indicated. **(B)** Genomic organization of *CCR5* mRNA. *CCR5* mRNA transcripts generated from the upstream promoter, P2, undergo alternative splicing, giving rise to two transcripts, CCR5A and CCR5B which differ by the presence or absence of the untranslated exon 2A, respectively (MUMMIDI *et al.* 1997). P1 drives the expression of the transcripts that originate in exons 2A or 2B which results in the formation of several “truncated” transcripts. Image adapted from MUMMIDI *et al.* 2007.

#### 1.2.1.5. Genetic variation within the *CCR5* gene

To date, numerous single nucleotide polymorphisms (SNPs) have been identified in both the coding and noncoding regions of the *CCR5* gene. In addition to the well-studied *CCR5* $\Delta$ 32 mutation, several studies have assessed the potential influence of individual mutations in *CCR5* on HIV-1 susceptibility and/or disease progression with a number of these having been identified as important genetic factors capable of influencing susceptibility to HIV-1 infection or to affect the rate of disease progression. However, more recent studies have focused on the effect of *CCR5* SNPs inherited as a block in the form of haplotypes.

##### 1.2.1.5.1. ORF polymorphisms

The functional consequences of several naturally occurring ORF mutations of the *CCR5* gene have been studied. In addition to deletion mutations and those which generate premature stop codons, mutations which result in amino acid alterations in regions of the gene encoding transmembrane, intracellular, as well as extracellular domains of the CCR5 molecule, have been identified. However, it is important to note that not all mutations which affect receptor function deleteriously

influence HIV-1 binding (BLANPAIN *et al.* 2000; DONG *et al.* 2005; HOWARD *et al.* 1999). A few of the ORF polymorphisms studied are described below.

#### **a. CCR5 $\Delta$ 32**

*CCR5 $\Delta$ 32* is characterized by a 32-base pair deletion in the region encoding the second extracellular loop of CCR5 (LIU *et al.* 1996). This deletion results in truncation of the expressed protein (215 amino acid protein vs 352 amino acid residues in wild type) and prevents the expression of CCR5 on the cell surface (LIU *et al.* 1996). Individuals homozygous for this mutation are resistant to R5 HIV infection. Heterozygous individuals exhibit decreased HIV susceptibility and delayed disease progression most likely due to decreased CCR5 expression. Homozygosity for this allele does not, however, provide complete resistance to infection. Shortly after the discovery of *CCR5 $\Delta$ 32*, separate reports documenting evidence for HIV-1 infection in *CCR5 $\Delta$ 32* homozygous individuals emerged (BITI *et al.* 1997; O'BRIEN *et al.* 1997; THEODOROU *et al.* 1997), likely due to infection by viruses which use an alternate coreceptor for entry.

The geographical distribution of this allele has been studied in detail. The *CCR5 $\Delta$ 32* allele is restricted mainly to individuals of European, North African and West Asian descent and is virtually absent in American Indian, East Asian other African populations (NOVEMBRE *et al.* 2005; STEPHENS *et al.* 1998). There exists a north-to-south geographic cline in allele frequency, with the highest population frequency occurring in northern European populations, particularly those in the Baltic region (16%), lower frequencies in Italian and Greek populations (6% and 4%, respectively), and a 0% allele frequency among Saudi and East Asian populations (NOVEMBRE *et al.* 2005; STEPHENS *et al.* 1998). Both smallpox and the bubonic plague have been suggested as possible selective agents driving the observed allele frequency distribution (GALVANI and SLATKIN 2003; STEPHENS *et al.* 1998). The detection of this allele in skeletons dating back to the bronze age suggests that this allele is at least 2900 years old (HUMMEL *et al.* 2005).

In addition to the association with HIV-1 infection, the *CCR5 $\Delta$ 32* allele is associated with various chronic inflammatory diseases, such as delayed onset and lower risk of recurrent clinical disease activity of multiple sclerosis (BARCELLOS *et al.* 2000; SELLEBJERG *et al.* 2000), protection against advanced atherosclerosis (ZERNECKE *et al.* 2008), decreased risk of type 1 diabetes (SMYTH *et al.* 2008) and decreased severity of rheumatoid arthritis (PRAHALAD 2006; ZAPICO *et al.* 2000) compared to individuals lacking this allele. Individuals who are homozygous *CCR5 $\Delta$ 32* and who receive renal transplants demonstrate longer transplant survival times than individuals who are *CCR5* wild type (FISCHEREDER *et al.* 2001). However, a higher frequency of this allele has been observed in patients with more severe pulmonary sarcoidosis (PETREK *et al.* 2000) and primary sclerosing cholangitis, a chronic liver disease (ERI *et al.* 2004).

The general good health of individuals homozygous for the *CCR5* $\Delta$ 32 mutation, which renders the CCR5 protein dysfunctional, is thought to be attributed to the redundancy of chemokine receptors and their ligands and has presented CCR5 as a target for therapeutic treatment of HIV-1 infection. However, although CCR5 might be a logical target for new drug development in HIV/AIDS, the benefits of blocking CCR5 could carry the cost of an increased risk of other diseases, such as infection with flaviviruses, specifically, tickborne encephalitis virus and the re-emerging pathogen West Nile Virus (KLEIN 2008).

#### **b. Other ORF SNPs and deletion mutations resulting in premature stop codons**

In addition to the well-studied *CCR5* $\Delta$ 32 deletion, several other deletion mutations have been detected in various populations. A 24 base pair deletion, *hCCR5* $\Delta$ 24, was detected in Rwanda (MASQUELIER *et al.* 2007). The effect of this deletion on CCR5 expression has not yet been determined.

The CCR5-893(-) mutation, exclusively observed in Asian populations, is characterized by a single-nucleotide deletion in codon 299 which results in a frameshift and hence a truncated protein (308 amino acids) lacking the C-terminal cytoplasmic tail (ANSARI-LARI *et al.* 1997). Shioda *et al.* (2001) demonstrated that this mutation results in reduced coreceptor activity and suggest that, like *CCR5* $\Delta$ 32, the protein products of CCR5-893(-) are retained in the endoplasmic reticulum and as a result this mutation affects HIV-1 transmission and disease progression (SHIODA *et al.* 2001).

A mutation (C101X, T303A or m303) creating a stop codon, and hence a truncated protein, before the third transmembrane domain results in an inability of the CCR5 coreceptor to mediate the entry of HIV-1 (BLANPAIN *et al.* 2000). This mutation is reported to, along with another variant (A29S), to be the most common CCR5 ORF mutation in Central African populations. When found in addition to the *CCR5* $\Delta$ 32 allele (i.e., heterozygous), C101X provides complete protection from infection with R5 isolates (CARRINGTON *et al.* 1997; QUILLEN *et al.* 1998).

FS299, found in Asian populations (3.5% frequency), is the result of a single base pair deletion which causes a frame shift and hence the premature termination of the CCR5 protein within the seventh transmembrane region and the absence of an intracellular C-terminal tail (BLANPAIN *et al.* 2000). Although this mutant lacks the ability to bind or respond to chemokines, it can still bind HIV-1 (BLANPAIN *et al.* 2000).

CCR5 deletions have also been found in non-human primates such as red-capped and sooty mangabees (CHEN *et al.* 1998; PALACIOS *et al.* 1998).

### **c. Amino acid substitutions within the ORF**

A number of nucleotide sequence variants resulting in altered CCR5 amino acid sequences have been identified and studied by different groups (ANSARI-LARI *et al.* 1997; BLANPAIN *et al.* 2000; CAPOULADE-METAY *et al.* 2004; CARRINGTON *et al.* 1997; HOWARD *et al.* 1999; MAGIEROWSKA *et al.* 1999a; PETERSEN *et al.* 2001; ZHAO *et al.* 2005). These variants may result in conformational changes in the protein, alter the binding affinity of CCR5 to some, or all of its ligands, alter the binding affinity to HIV-1 and/or affect the signaling ability of CCR5. However, some variants may have little or no effect on the protein function. The frequency of these polymorphisms differs according to the ethnicity of a population (Table 1.2). With the exception of L55Q (4.1% frequency in Caucasians (CARRINGTON *et al.* 1997)), R223Q (4.6% in the Chinese population (ZHAO *et al.* 2005)) and Y339F (2.6% in African Americans (CARRINGTON *et al.* 1997)), all the identified ORF non-synonymous polymorphisms are reported to occur at relatively low frequencies ( $\leq 1.6\%$ ) within different population groups (Table 1.2).

A report by Blanpain *et al.* (2000), in which the functional responses of several such mutants to their chemokine ligands and to HIV-1 binding was assessed, highlights how the effects of CCR5 ORF mutations on these factors can vary considerably. The C20S amino acid substitution results in a disruption of the disulphide bond between the N-terminal region and the third extracellular loop of the protein (C269) shown in Figure 1.2. This disulphide bond, along with a second disulphide bond between the cysteine residues at position 178 and 101, has been demonstrated to be important in maintaining the structural integrity of the CCR5 protein in addition to being essential for ligand binding (BLANPAIN *et al.* 1999b). Although this disruption causes markedly reduced CCR5 surface expression and an absence of ligand binding to the receptor, the C20S CCR5 mutant still functions as an HIV-1 coreceptor (BLANPAIN *et al.* 2000). Similarly, another mutation at a cysteine position involved in linking the first and second extracellular loops, C178R, results in reduced surface expression and an inability to respond to chemokines but retains its coreceptor function (BLANPAIN *et al.* 2000). Contrastingly, cell surface expression of the I12L mutant is similar to that of 'wild type' CCR5 and there appears to be no functional impairment (BLANPAIN *et al.* 2000). The A29S mutation does not affect CCR5 cell surface expression, however binding affinity for two of the CCR5 ligands, CCL3 and CCL4, is lost while that for a third ligand, CCL8, is retained (BLANPAIN *et al.* 2000).

**Table 1.2.** Distribution of *CCR5* mutations resulting in amino acid substitutions in different populations

SNP*	Amino acid substitution*	Allele frequency in different populations, where detected <sup>†</sup>	References
+5A/T	D2V	0.7% South African Coloureds	(PETERSEN <i>et al.</i> 2001)
+25A/C	I12L	0.3% Caucasians	(CARRINGTON <i>et al.</i> 1997)
+58T/A	C20S	0.3% Caucasians	(CARRINGTON <i>et al.</i> 1997)
		-	(BLANPAIN <i>et al.</i> 2000)
		-	(HOWARD <i>et al.</i> 1999)
+85G/T	A29S	1.5% African Americans	(CARRINGTON <i>et al.</i> 1997)
		-	(BLANPAIN <i>et al.</i> 2000)
		-	(HOWARD <i>et al.</i> 1999)
+124A/T	I42F	0.1% Caucasians	(CARRINGTON <i>et al.</i> 1997)
+164T/A	L55Q	0.4% Hispanics	(ANSARI-LARI <i>et al.</i> 1997)
		4.1% Caucasians, 0.7% African Americans	(CARRINGTON <i>et al.</i> 1997)
		-	(BLANPAIN <i>et al.</i> 2000)
		-	(PETERSEN <i>et al.</i> 2001)
+180G/T	R60S	1.3% African Americans	(CARRINGTON <i>et al.</i> 1997)
		-	(DONG <i>et al.</i> 2005)
+218C/T	A73V	0.2% Caucasians	(CARRINGTON <i>et al.</i> 1997)
		-	(HOWARD <i>et al.</i> 1999)
+316G/A	G106R	1.35% Vietnamese	(CAPOULADE-METAY <i>et al.</i> 2004)
		0.25% Chinese	(ZHAO <i>et al.</i> 2005)
+319C/T	L107F	1.6% South African Africans	(PETERSEN <i>et al.</i> 2001)
+513A/G	W513C	0.06% Chinese	(ZHAO <i>et al.</i> 2005)
+532T/C	C178R	1.09% Vietnamese	(MAGIEROWSKA <i>et al.</i> 1999a)
+553A/C	S185R	1.35% Vietnamese	(CAPOULADE-METAY <i>et al.</i> 2004)
+582G/T	Q194H	0.16% Chinese (HIV-1 infected cohort)	(ZHAO <i>et al.</i> 2005)
+664C/T	S215L	0.2% African Americans	(ANSARI-LARI <i>et al.</i> 1997)
+668G/A	R223Q	0.4% Hispanics, 1.0 % Chinese, 0.8% Japanese	(ANSARI-LARI <i>et al.</i> 1997)
		1.6% Caucasians	(CARRINGTON <i>et al.</i> 1997)
		1.35% Vietnamese	(CAPOULADE-METAY <i>et al.</i> 2004)
		4.6% Chinese	(ZHAO <i>et al.</i> 2005)
		0.7% South African Coloureds	(PETERSEN <i>et al.</i> 2001)
		-	(BLANPAIN <i>et al.</i> 2000)
		-	(DONG <i>et al.</i> 2005)
		-	(DONG <i>et al.</i> 2005)
+674G/A	R225Q	0.7% South African Coloureds	(PETERSEN <i>et al.</i> 2001)
+758T/C	I254T	1.06% Vietnamese (HIV-1 infected cohort)	(CAPOULADE-METAY <i>et al.</i> 2004)
+806G/T	C269F	1.35% Vietnamese	(CAPOULADE-METAY <i>et al.</i> 2004)
+902G/T	G301V	1.1% Caucasians	(CARRINGTON <i>et al.</i> 1997)
		-	(DONG <i>et al.</i> 2005)

+1004C/T	A335V	0.6% African Americans, 0.2% Hispanic 0.6% Caucasians 1.6% South African Africans, 1.4% South African Coloureds -	(ANSARI-LARI <i>et al.</i> 1997) (CARRINGTON <i>et al.</i> 1997) (PETERSEN <i>et al.</i> 2001) (DONG <i>et al.</i> 2005)
+1007G/T	S336I	0.32% Chinese	(ZHAO <i>et al.</i> 2005)
+1016A/T	Y339F	0.2% African Americans 2.6% African Americans 0.8% South African Africans -	(ANSARI-LARI <i>et al.</i> 1997) (CARRINGTON <i>et al.</i> 1997) (PETERSEN <i>et al.</i> 2001) (DONG <i>et al.</i> 2005)

\*The first nucleotide of the translational start site is designated as +1.

♦See Appendix F for amino acid codes.

†The study conducted by Carrington *et al.* (1997) was restricted to HIV-1 infected cohorts and thus frequencies may be slightly biased. Where an allele frequency of a particular SNP is not shown (“-“), it is an indication that although detected in the referenced study, the frequency at which it occurs was not determined.

#### d. Open reading frame mutations in South African populations

Although the *CCR5* $\Delta$ 32 allele is virtually absent in African populations in South Africa (0.1%) (WILLIAMSON *et al.* 2000), several other ORF mutations have been reported. Four mutations, D2V, L107F, R225Q and R225X, predicted to result in non-synonymous amino acid changes within the CCR5 protein, were identified as unique to South African Coloured and African populations, occurring at low frequencies (Table 1.2) (PETERSEN *et al.* 2001). One of these mutations introduced a premature stop codon at amino acid position 225 (R225X). *In vitro* functional analysis demonstrated lack of CCR5 expression in cells transfected with the R225X CCR5 mutant construct (FOLEFOC *et al.* 2010). The other three *CCR5* mutants were demonstrated to have varying effects on CCR5 function: L107F was similar to CCR5 “wild type”, R225Q was partially constitutively active and D2V demonstrated reduced chemokine binding affinity, lower cell surface expression in addition to decreased Env-dependent cell fusion (FOLEFOC *et al.* 2010).

In addition to these newly identified mutations, a further four ‘silent’ mutations (three of which were newly identified) and four non-synonymous mutations (L55Q, R223Q, A335V and Y339F), which had been previously identified in other populations, were detected in South African populations (African and Coloured) (PETERSEN *et al.* 2001). The A335V mutation was demonstrated to associate significantly with delayed disease progression, i.e., long term nonprogression, in a study of HIV-1 infected individuals of African ethnicity in the Western Cape region of South Africa (HAYES *et al.* 2002).

In a further South African study conducted by Williamson *et al.* (2000), the C101X mutation was reported to be absent amongst Africans and present at a frequency of 0.9 % in the South African Caucasian population (WILLIAMSON *et al.* 2000).

#### 1.2.1.5.2. Promoter polymorphisms

Expression levels of proteins are likely to be affected by mutations within the regulatory region of the corresponding gene and polymorphisms within the *CCR5* regulatory region have been found to affect susceptibility to HIV-1 (BAMSHAD *et al.* 2002). Furthermore, a mutation (V64I) within *CCR2* ORF has been associated with delayed disease progression in African Americans (MUMMIDI *et al.* 1998). This mutation is however in complete linkage disequilibrium with a SNP, in the *CCR5* regulatory region (-1835T) (MUMMIDI *et al.* 1998) making it unclear which particular SNP or SNP interaction is conferring protection. The *CCR2* gene is found in close proximity (17.5 kb upstream) to *CCR5* and the nucleic acid and amino acid sequences of the two receptors share 75% identity (SAMSON *et al.* 1996a).

#### **a. -2459A/G [also known as: 303 (MARTIN *et al.* 1998), 59029 A/G (McDERMOTT *et al.* 1998)]**

McDermott *et al.* (1998) identified an A/G polymorphism within the downstream promoter associated with the rate of progression to AIDS. Both alleles (-2459A and -2459G) are present in high frequencies in all racial groups (McDERMOTT *et al.* 1998). This SNP is linked to *CCR5* expression, where individuals homozygous or heterozygous for the -2459G allele have fewer *CCR5* receptors present on the surface of CD14<sup>+</sup> monocytes (SALKOWITZ *et al.* 2003). Thus, individuals homozygous for -2459A progress to AIDS and death more rapidly when compared to those homozygous for -2459G (KNUDSEN *et al.* 2001; McDERMOTT *et al.* 1998). Also, this polymorphism predicts the magnitude of HIV-1 propagation *in vitro* with low, medium and high viral propagation in association with G/G, A/G and A/A genotypes, respectively (SALKOWITZ *et al.* 2003).

#### **b. -2135C/T [59353T (McDERMOTT *et al.* 1998) or 627C/T (MARTIN *et al.* 1998)]**

The -2135C allele is in very strong linkage disequilibrium with the -2459A allele and is thus also associated with an accelerated rate of HIV-1 disease progression (CLEGG *et al.* 2000; OMETTO *et al.* 2001). Neither -2135C, nor -2459A have been shown to have an influence on the risk of acquiring HIV (CLEGG *et al.* 2000). However, homozygosity for the C allele at this position has been associated with a trend towards slower rate of decline in CD4<sup>+</sup> T cell numbers (EASTERBROOK *et al.* 1999).

#### **c. -1835C/T [59653-C/T (McDERMOTT *et al.* 1998) or 927C/T (MARTIN *et al.* 1998)]**

The point mutation (-1835T) within intron 2 of *CCR5* with the single amino acid substitution is in complete linkage disequilibrium with the single amino acid substitution, CCR2V64I (KOSTRIKIS *et*

*al.* 1998; MUMMIDI *et al.* 1998). This combination is protective against disease progression but not on disease transmission (KOSTRIKIS *et al.* 1998).

**d. -1951G/A [59537-G/A (McDERMOTT *et al.* 1998)]**

Although not expressed on the cell surface, CCR5 $\Delta$ 32 protein does get translated. In a study investigating the lack of CCR5 $\Delta$ 32 protein expression in HIV-1 infected individuals homozygous for the CCR5 $\Delta$ 32, it was demonstrated that the -1951A mutation was associated with low translational efficiency of CCR5 $\Delta$ 32 in comparison to the -1951G allele (JIN *et al.* 2008).

1.2.1.5.3. CCR5 Haplotypes

Association studies assessing the effects of a particular SNP on a phenotype of the gene, such as disease outcome, are often complicated as it is not always clear whether that SNP is the true causal variant. Generally, SNPs do not arise independently, as there exists an intrinsic dependency of SNPs with one another due to the history of their entry into a population, and thus genetic variation within a population is structured into haplotypes which are likely to be transmitted as a unit (CLARK 2004). Haplotypes are defined as a combination of allelic variants or SNPs that are statistically associated on the same chromosome and which are inherited as a block.

Two main haplotype grouping systems, based on SNPs within the CCR5 regulatory region, have been described by different groups and have been shown to influence HIV-1 disease progression (GONZALEZ *et al.* 1999; MARTIN *et al.* 1998). Although there is some overlap between the SNPs within haplotypes studied by the different groups, i.e., -2554G/T, -2135T/C, -2132C/T and -2086A/G polymorphisms, consensus or clarity is lacking, which illustrates the complexity involved in studying these relationships. The HHA through HHG\*2 CCR5 haplotype-defining system which includes the CCR5 $\Delta$ 32 and CCR2V64I polymorphisms (GONZALEZ *et al.* 1999; MUMMIDI *et al.* 1998) has been favoured in subsequent studies.

Martin *et al.* (1998) identified ten CCR5 haplotypes (CCR5P1 through CCR5P10) based on observed linkage disequilibrium patterns of ten CCR5 SNP positions spanning exons 1, 2A and 2B and intron 1 of the gene. Four of these haplotypes (CCR5P1 through CCR5P4) were reported as common, while the remaining six haplotypes were found to be rare, in cohorts of HIV-1 infected individuals (MARTIN *et al.* 1998). Individuals homozygous for the CCR5P1 allele were reported to progress to AIDS more rapidly than those with other CCR5 promoter haplotypes (MARTIN *et al.* 1998).

A second group of haplotypes consisting of nine evolutionarily distinct groups (HHA through HHG\*2) has been defined based on seven SNPs at positions -2733, -2554, -2459, -2135, -2132, -2086 and -1835 within the CCR5 5' regulatory region (GONZALEZ *et al.* 1999; MUMMIDI *et al.* 1998).

The *CCR2V64I* and *CCR5Δ32* mutations were demonstrated to be in linkage disequilibrium with the HHF and HHG haplotypes, respectively, forming haplotypes HHF\*2 and HHG\*2 (GONZALEZ *et al.* 1999; MUMMIDI *et al.* 2000). Through comparison of human *CCR5* haplotypes to those of non-human primates, including apes, New World Monkeys and Old World Monkeys, the HHA haplotype was identified as ancestral to all other haplotypes (MUMMIDI *et al.* 2000). These haplotypes have striking population-specific distributions and, furthermore, appear to have differing disease modifying effects in different population groups (i.e., Caucasian vs African Americans) (GONZALEZ *et al.* 1999; GONZALEZ *et al.* 2001). For example, in a study conducted by Gonzalez *et al.* (1999), the HHC haplotype was shown to be associated with disease retardation in Caucasian and Hispanic populations, whereas it is associated with disease acceleration in African Americans. In addition to the influence that *CCR5* haplotypes have on HIV-1 pathogenesis through modification of *CCR5* coreceptor activity, it has been demonstrated that *CCR5* haplotypes may also influence immune mechanisms independent of its function as a coreceptor, i.e., T cell mediated immunity (CATANO *et al.* 2011).

Importantly, although the presence or absence of a single haplotype, for example HHG\*2 (*CCR5Δ32*-containing allele), can have a dramatic influence on an individual's susceptibility to HIV-1 infection and rate of disease progression to AIDS, it is the genotype of the individual which is a strong determining factor. The disease-retarding effects of heterozygosity for *CCR5Δ32* are mainly attributable to the HHA/HHG\*2, HHC/HHG\*2 and HHF\*2/HHG\*2 genotypes (CATANO *et al.* 2011). In contrast the HHE/HHG\*2 genotype is associated with detrimental effects (CATANO *et al.* 2011).

The HIV-1 resistance afforded by *CCR5* haplotypes within the promoter region are thought to be the result of adaptive changes to pathogens older than HIV (BAMSHAD *et al.* 2002).

#### 1.2.1.6. *CCR2V64I* mutation

A non-synonymous G/A polymorphism at nucleotide position 190 within the coding region of the CC chemokine receptor 2 (*CCR2*), a minor coreceptor, substitutes the "wild type" valine amino acid residue with an isoleucine amino acid at position 64 (SMITH *et al.* 1997). The *CCR2* 64I mutation does not affect acquisition of HIV-1 infection (MICHAEL *et al.* 1997; SMITH *et al.* 1997). Controversial data exists on the influence of the *CCR2* 64I mutation on HIV-1 disease progression. Although the presence of this mutation has been shown to result in a delay in progression to AIDS (EASTERBROOK *et al.* 1999; RIZZARDI *et al.* 1998; SMITH *et al.* 1997), others have failed to show the same (EUGEN-OLSEN *et al.* 1998; HLADIK *et al.* 2005; MAGIEROWSKA *et al.* 1999b; MICHAEL *et al.* 1997; SCHINKEL *et al.* 1999). Furthermore, the protective effects of this 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 an association with the *CCR2* 64I mutation and

delayed disease progression to AIDS in Caucasian, but not African HIV-1 infected individuals, while Mummidi *et al.* (1998) reported the opposite. A meta-analysis of all studies available in 2001 concluded that this allele (or haplotype) provides a significant protective effect on disease progression (IOANNIDIS *et al.* 2001). Furthermore, a subsequent meta-analysis concluded that this same mutation is associated with a decreased risk of death in perinatally infected infants (IOANNIDIS *et al.* 2003). As mentioned above, due to complete linkage disequilibrium with the -1835T SNP within the *CCR5* 5' flanking region, it is not clear whether the association between *CCR2V64I* and HIV-1 disease progression is due to the polymorphism in the *CCR2* or in the *CCR5* gene.

#### 1.2.1.7. *CCR5* expression

*CCR5* is expressed by a spectrum of cell types of both human and macaque origin (ROTTMAN *et al.* 1997). These include subpopulations of lymphocytes and monocytes/macrophages in blood, primary and secondary lymphoid organs and noninflamed tissues (ROTTMAN *et al.* 1997) (Table 1.1). Other *CCR5* expressing cell types, found in a variety of tissues, include neurons, astrocytes, microglia, epithelium, endothelium, vascular smooth muscle and fibroblasts (ROTTMAN *et al.* 1997). The ability of a leukocyte to respond to a given chemokine is dependent on surface expression of the appropriate receptor. Thus, *CCR5* density is a determinant of the efficiency of *CCR5* in chemotactic response to its ligands which implies that this is an important determinant of T-cell migration in many biological situations where chemokines that bind *CCR5* play a role (DESMETZ *et al.* 2006). *CCR5* receptor expression has been shown to be stable within individuals over multiple time points despite the wide range of variability that exists between individuals (OSTROWSKI *et al.* 1998; REYNES *et al.* 2000).

*CCR5* expression is increased where there is chronic inflammation (ROTTMAN *et al.* 1997). For most donors, the most distinctive expression of *CCR5* has been reported as occurring on CD3+ T cells (10-20%) (WU *et al.* 1997). In T lymphocytes, expression of *CCR5* is largely restricted to subsets corresponding to activated or memory subsets (CD3+ CD26+ CD45RA<sup>low</sup> CD45RO+ (BLEUL *et al.* 1997; WU *et al.* 1997); HLA-DR+ T cells (OSTROWSKI *et al.* 1998)).

Differences in *CCR5* expression have been noted between different population groups as well as between individuals of the same population exposed to different environmental factors (CLERICI *et al.* 2000; KALINKOVICH *et al.* 2001; KALINKOVICH *et al.* 1999). Both monocytes and lymphocytes obtained from Ethiopian individuals have demonstrated increased expression of *CCR5*, manifested by significantly higher proportions of *CCR5*-expressing CD4+ cells and higher numbers of *CCR5* molecules per cells, with decreased  $\beta$ -chemokine secretion compared to non-Ethiopian Caucasian Israeli individuals (KALINKOVICH *et al.* 2001; KALINKOVICH *et al.* 1999). It was initially suggested that

this difference in expression was mainly a result of environmental factors resulting in chronic immune activation in Ethiopian individuals since these individuals appeared to display higher HLA-DR frequencies which are accompanied by higher CCR5 expression. However, increased CCR5 expression was also found in the nonactivated compartment of CD4<sup>+</sup> cells, suggesting that factors other than immune activation are also contributing to the differences observed between the two population groups (KALINKOVICH *et al.* 2001). Therefore both genetic factors and environmental stimuli appear to influence CCR5 expression.

#### 1.2.1.7.1. CCR5 expression and HIV-1 infection

Previous work has indicated that the amount of CCR5 expressed on the cell surface can directly influence an individual's susceptibility to HIV-1 with the discovery of the *CCR5Δ32* mutant found in high risk individuals resisting HIV infection (DEAN *et al.* 1996; LIU *et al.* 1998; SAMSON *et al.* 1996b). Individuals homozygous for this mutant fail to express a detectable CCR5 receptor on cell surfaces (LIU *et al.* 1996). This discovery prompted further studies of the gene and how its naturally occurring mutations may influence the outcome of HIV-1 exposure and infection.

Interestingly, in addition to the observed reduced cell surface expression of CCR5 due to truncation of the protein in individuals with the *CCR5Δ32* mutation, it has been suggested that *CCR5Δ32* protein may form heterodimers with both wild type CCR5 and CXCR4 protein, which as a result are retained within the endoplasmic reticulum and thus contribute to reduced cell surface expression of the HIV-1 wild type coreceptors (AGRAWAL *et al.* 2004; BENKIRANE *et al.* 1997).

Variation in CCR5 expression levels between individuals can have an impact on several different aspects of HIV-1 pathogenesis, such as HIV-1 entry into a cell (PAXTON *et al.* 1998; PLATT *et al.* 1998; WU *et al.* 1997), susceptibility to HIV-1 infection (KASLOW *et al.* 2005; PAXTON *et al.* 1998; REYNES *et al.* 2003; THOMAS *et al.* 2006), disease progression (GERVAIX *et al.* 2002), viral load (DE RODA HUSMAN *et al.* 1999; GERVAIX *et al.* 2002), immune reconstitution during antiretroviral therapy (GERVAIX *et al.* 2002; VINCENT *et al.* 2006), the efficacy of CCR5 antagonists and entry inhibitors (HEREDIA *et al.* 2007), as well as the neutralizing capacity of HIV-1-specific antibodies (CHOUDHRY *et al.* 2006). CCR5 coreceptors are coupled to two proteins, Gai and Gαq. The signaling through the Gai protein once R5 virions have bound CCR5 is responsible for the difference in HIV-1 infectability between high or low CCR5 expressing CD4<sup>+</sup> T cells (LIN *et al.* 2006). In addition to the role of CCR5 expression in HIV-1 infection, it has also been shown to influence other diseases. For example, *CCR5* deficient mice demonstrate increased severity of brain injury compared to wild type mice (SORCE *et al.* 2010). Furthermore, as discussed previously, *CCR5Δ32* heterozygosity, which influences CCR5 expression levels, is associated either positively or negatively with various diseases (Section 1.2.1.5.1a). Cell surface expression of CCR5 is highly variable, even in

individuals homozygous for the wild type ORF region. This may be explained by differences in the promoter region of *CCR5* which result in differing *CCR5* expression levels (MENTEN *et al.* 2002).

HIV-1 infection has been shown to alter the expression of *CCR5*: HIV-1 infected individuals have significantly greater percentages of *CCR5*-expressing CD4<sup>+</sup> T cells when compared to healthy controls (NICHOLSON *et al.* 2001; OSTROWSKI *et al.* 1998). Patients not infected with HIV-1 but who have other forms of chronic immune activation, such as asymptomatic chronic hepatitis C virus infection, have also demonstrated upregulation of *CCR5* expression (OSTROWSKI *et al.* 1998). Furthermore, primary T cells productively infected *in vitro* by HIV-1 have increased *CCR5* expression (MARODON *et al.* 1999). Interestingly, HIV-2 infection results in downregulation of *CCR5* expression on CD4<sup>+</sup> T cells to levels comparable to individuals harbouring the *CCR5*Δ32 deletion (SHEA *et al.* 2004).

However, given the important role played by *CCR5* expression in HIV-1 infection, it is important to note that cells which are putatively susceptible to infection with macrophage-tropic strains of HIV-1 (CD4<sup>+</sup>; *CCR5*<sup>+</sup> phenotype) constitute a relatively small percentage of the total cellular population within lymph node, blood and other lymphoid tissues (ROTTMAN *et al.* 1997).

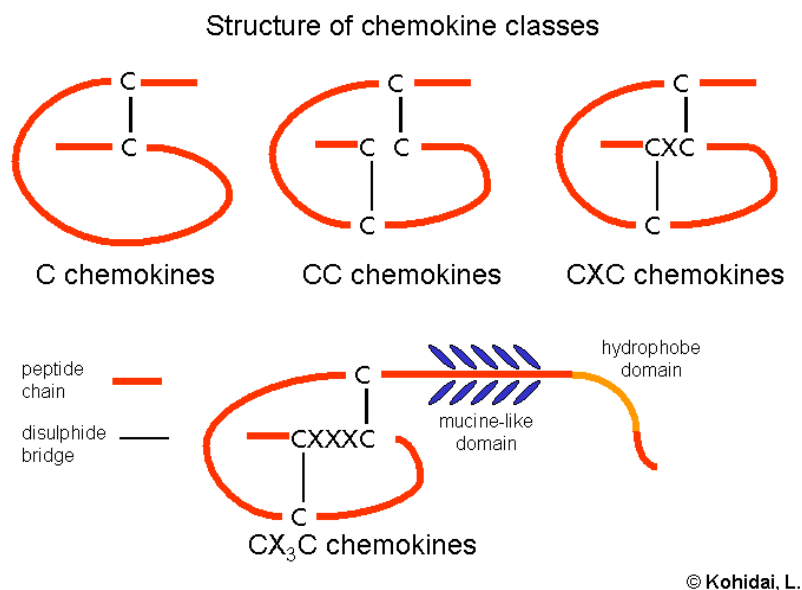
### 1.3. CHEMOKINES

Chemokines are small structurally related cytokines which play a major role in the chemotaxis of various leukocyte cell types to sites of injury or infection or to secondary lymphoid organs (Reviewed in FURIE and RANDOLPH 1995). These molecules are also important in other processes including inflammatory response, angiogenesis, haematopoiesis, foetal development, and in pathological conditions such as HIV-1 infection, autoimmune diseases and tumor growth (BALKWILL 2004; MENTEN *et al.* 2002; ROLLINS 2006).

Any single leukocyte cell type has receptors for a number of different chemokines and will respond to a number of different molecules. Also, most chemokines interact with more than one chemokine receptor and most receptors bind more than one chemokine. Thus, when one chemokine or chemokine receptor displays a function altering mutation or is absent, there are usually alternate molecules which will perform the same biological function (Reviewed in COLOBRAN *et al.* 2007).

Based on the positions of two conserved cysteine residues in their amino (N)-termini, chemokines can be classed into four subfamilies, namely: C chemokine family, CC chemokine family, CXC chemokine family and CX<sub>3</sub>C chemokine family (ZLOTNIK and YOSHIE 2000) (Figure 1.4). Human genes coding for members of the 4 subfamilies tend to be clustered together on the same chromosome. Two main clusters have been identified, the CXC cluster on the long arm of chromosome 4 (4q12-21) and the CC cluster on the long arm of chromosome 17 (17q11.2)

(COLOBRAN *et al.* 2007). CXC chemokines are primarily involved in the recruitment of neutrophils, whereas CC chemokines recruit other leukocytes such as monocytes, lymphocytes and basophils. The chemokines, CCL3, CCL4 and CCL5 also known as MIP-1 $\alpha$ , MIP-1 $\beta$  and RANTES, respectively, are members of the CC-subfamily of chemokines.



**Figure 1.4.** Chemokines are classified into four different classes based on the spacing of their first two cysteine residues. (Image source: <http://en.wikipedia.org/wiki/Chemokines>).

CCL3, CCL4 and CCL5, are the natural ligands for the CCR5 receptor and are known to inhibit replication of CCR5-restricted (i.e., R5) HIV-1 variants (COCCHI *et al.* 1995; TRKOLA *et al.* 1998). This is achieved mainly by preventing virus entry. Three mechanisms for this have been postulated, by directly blocking the HIV envelope binding site (i.e., steric hindrance), downmodulation of CCR5 after binding thereby reducing the number of available entry points, and CCR5 dimerization which would result in a conformational change of the receptor (Reviewed in MENTEN *et al.* 2002). CCR5 internalized as a result of ligand binding has been demonstrated to recycle to the plasma membrane after the removal of the ligand (SIGNORET *et al.* 2000).

### 1.3.1. CCL3 (MIP-1 $\alpha$ )

Although *CCL3* deletion in mice has no discernible effect on development, it has profound effects on immunological responses in models of viral, bacterial as well as yeast infection (COOK *et al.* 1995; OLSZEWSKI *et al.* 2001), graft-versus-host disease (SERODY *et al.* 2000), central nervous system demyelination (MCMAHON *et al.* 2001) and type 1 diabetes (CAMERON *et al.* 2000). CCL3 acts as a pro-inflammatory cytokine, inducible in a number of haematopoietic cells, particularly in

those involved in the adaptive immune responses (macrophages, dendritic cells, B and T lymphocytes). CCL3 is capable of activating a wide variety of immune cells, particularly CD8+ T and immature dendritic cells, and is inhibited by IL-4, IL-10 and IL-13 (BERKMAN *et al.* 1995; OHTA *et al.* 1998; SCHALL *et al.* 1993; STANDIFORD *et al.* 1993). It is secreted in response to stimulation and functions by attracting lymphocytes and macrophages to the sites of infection and inflammation. Besides binding to CCR5, CCL3 also binds to the receptors CCR1 and CCR3.

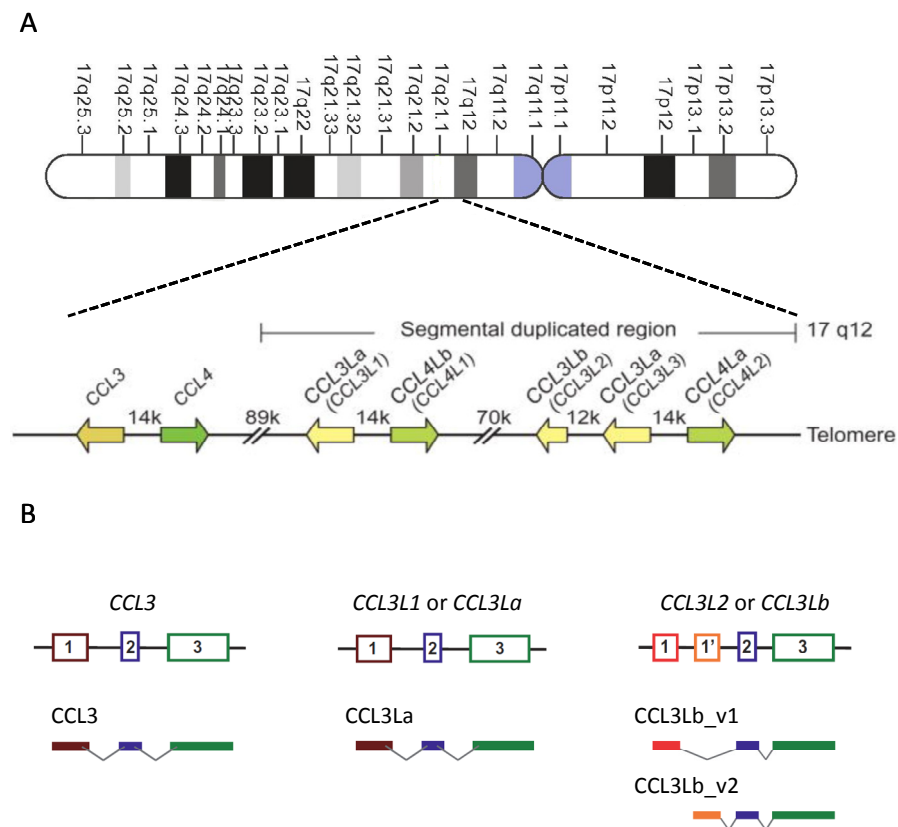
#### 1.3.1.1. CCL3 genes

The *CCL3* gene family is localized on chromosome 17 in a hotspot for segmental duplications, 17q12 (HIRASHIMA *et al.* 1992) (Figure 1.5A). The *CCL3* gene is found in close proximity to *CCL4* and it is thought that the two chemokines were formed by the duplication of a common ancestral gene (MENTEN *et al.* 2002). Through sequencing of a BAC clone located in this region, two complete nonallelic copies (*CCL3L1* and *CCL3L3*) and a truncated copy (*CCL3L2*) were identified (MODI 2004). Both the allelic (*CCL3*) and non-allelic (*CCL3L1* and *CCL3L3*) genes are composed of three exons. Similarly to *CCL3*, the *CCL4* gene has a second non-allelic copy, *CCL4L*. Both *CCL3L* and *CCL4L* are often preserved in a single unit of segmental duplication (Figure 1.5A), however, there exists only a modest correlation between *CCL3L* and *CCL4L* gene copy number, i.e., the two non-allelic genes are not always found at a 1:1 ratio within individuals, which suggests that these two genes are not always duplicated simultaneously (SHAO *et al.* 2007).

Another related gene, *CCL3Lb* (*CCL3L2* or *LD78y*), was initially identified as a 5'-truncated pseudogene, as it contains only 2 exons homologous to exons 2 and 3 of the *CCL3L1* and *CCL3L3* genes (HIRASHIMA *et al.* 1992). However, novel 5' exons have been recently identified within *CCL3Lb* which give rise to two alternatively spliced transcripts (SHOSTAKOVICH-KORETSKAYA *et al.* 2009) (Figure 1.5B). These mRNA transcripts contain chemokine-like domains but are not predicted to code for classical chemokines (SHOSTAKOVICH-KORETSKAYA *et al.* 2009).

*CCL3* is present at two copies per diploid genome (pdg), whereas *CCL3La* and *CCL3Lb* are present at variable copy number pdg. Copy number polymorphism is present between populations as well as individuals within a population. *CCL3L* (*CCL3La* + *CCL3Lb*) gene copy number varies from 0 to approximately 14 in humans (GONZALEZ *et al.* 2005). Sub-Saharan populations display the highest number of *CCL3L* gene copy numbers while European populations have the lowest *CCL3L* gene copy numbers (median of 6 and 2, respectively) (GONZALEZ *et al.* 2005). Few individuals (<5%) have no *CCL3L* genes (BUGEJA *et al.* 2004; GONZALEZ *et al.* 2005). Non-human primates, chimpanzees (*Pan troglodytes*), African green monkeys (*Chlorocebus sabaeus*) and rhesus macaques of both Indian and Chinese origin (*Mucaca mulatta*), have higher *CCL3L* gene

copy numbers compared to humans (DEGENHARDT *et al.* 2009; GONZALEZ *et al.* 2005; GORNALUSSE *et al.* 2009; LIM *et al.* 2010).



**Figure 1.5.** Genomic organisation and mRNA products of *CCL3* and *CCL3L* genes. **(A)** Schematic representation of the allelic, *CCL3* and *CCL4*, and non-allelic, *CCL3L* and *CCL4L*, genes located in close proximity to each other on chromosome 17. The arrows indicate the orientation of the genes and the distances between genes are indicated. **(B)** Genomic organization, showing introns and exons (boxes), and transcription pattern of the *CCL3*, *CCL3L* (*CCL3La* and *CCL3Lb*) mRNA. The *CCL3Lb* gene, previously thought to be a pseudogene, gives rise to two alternatively spliced transcripts. (Figure modified from figures in COLOBRAN *et al.* 2010; GRUNHAGE *et al.* 2010; SHOSTAKOVICH-KORETSKAYA *et al.* 2009).

### 1.3.1.2. *CCL3* protein isoforms

The highly related functional genes (94% nucleic acid homology) (NAKAO *et al.* 1990), *CCL3* (*LD78 $\alpha$*  or *MIP-1 $\alpha$* ) and *CCL3La* (*LD78 $\beta$*  or *MIP-1 $\alpha$ P*, *CCL3L1* and *CCL3L3*), are known to code for two *CCL3* isoforms. Although these two isoforms of the *CCL3* chemokine, *CCL3* (*MIP-1 $\alpha$* ) and *CCL3L1* (*MIP-1 $\alpha$ P*), differ only in three amino acids at positions 3, 39 and 47 (MENTEN *et al.* 1999; MODI *et al.* 2001), they appear to differ in their biological function. *CCL3* binds the chemokine receptors CCR1 and CCR5 while *CCL3L1* also binds CCR3 efficiently (STRUYF *et al.* 2001).

Functionally, CCL3L1, encoded by the *CCL3* non-allelic copies, is more effective at downregulating CCR5 expression and has greater antiviral activity in monocytes/macrophages (AQUARO *et al.* 2001).

Stimulation experiments have shown a positive relationship between *CCL3L* copy number and CCL3 production, i.e., low *CCL3L* copy numbers show association with reduced CCL3 levels (GONZALEZ *et al.* 2005; MEDDOWS-TAYLOR *et al.* 2006; TOWNSON *et al.* 2002). It was recently reported that there is no association with *CCL3L* copy number and the peripheral circulating levels of CCL3 in both HIV-1 infected and uninfected individuals (LARSEN *et al.* 2012).

The CCL3L1 chemokine appears to be the most potent agonist of CCR5 and inhibitor of R5 HIV-1 strains in comparison to the allelic CCL3 chemokine (MENTEN *et al.* 1999; NIBBS *et al.* 1999).

### 1.3.1.3. *CCL3* and HIV-1

Given that CCL3 is a ligand for the HIV-1 coreceptor, CCR5, numerous studies have investigated the importance of CCL3 in the context of HIV-1 infection. Production of both CCL3 and CCL4 by CD8+ T cells has been demonstrated to be significantly higher in asymptomatic HIV+ individuals compared to both HIV+ individuals who have progressed to AIDS and uninfected individuals suggesting that these two chemokines play an important role of the outcome of HIV-1 infection (COCCHI *et al.* 2000). Several reports have demonstrated associations between *CCL3L* (and *CCL4L*) copy number variation on HIV-1 disease progression to AIDS, however this has been disputed.

The first report demonstrating an association between *CCL3L* gene copy number and HIV-1 susceptibility and disease progression to AIDS emerged in 2005 (GONZALEZ *et al.* 2005). Individuals with *CCL3L* gene copy numbers above the population-specific median were demonstrated to be less susceptible to acquiring HIV-1 infection (both children and adults) and progress to AIDS at a slower rate compared to individuals with *CCL3L* gene copy numbers below the population-specific median. Variation in *CCL3L* copy number has been associated with a number of HIV-related outcomes, including vertical and horizontal transmission of HIV-1 (GONZALEZ *et al.* 2005; KUHN *et al.* 2007; MEDDOWS-TAYLOR *et al.* 2006), plasma viral load (DOLAN *et al.* 2007; SHALEKOFF *et al.* 2008), progression to AIDS (DOLAN *et al.* 2007), HIV-1 specific gag responses (SHALEKOFF *et al.* 2008), and response to treatment of HIV-1/AIDS (AHUJA *et al.* 2008). However, there are reports disputing these findings by demonstrating a lack of association between *CCL3L* copy number variation and any HIV-1 disease phenotypes (BHATTACHARYA *et al.* 2009; SHAO *et al.* 2007; URBAN *et al.* 2009). This controversy may be partly attributed to differences in the methods used to quantify gene copy number (Reviewed in SHRESTHA *et al.* 2009). In a study assessing *CCL3L* gene copy number in more than 12 000 British individuals,

differences were noted in *CCL3L* copy number values obtained through two different assays, quantitative real-time polymerase chain reaction (qPCR) and the paralogue ratio test (PRT) (FIELD *et al.* 2009). A second report demonstrated variation in gene copy number distribution in two populations (European Americans and African Americans) when copy numbers were determined by means of qPCR using four different primer and probe sets, and suggested that some of the inconsistencies in association studies could be attributable to assays which provide heterogeneous results (SHRESTHA *et al.* 2010). Subsequently, a meta-analysis of nine published studies indicated that *CCL3L* copy number variability, relative to the population-specific median, is associated with susceptibility to HIV-1 infection, with higher copy number being associated with lower risk of HIV-1 acquisition, and vice versa (LIU *et al.* 2010).

In rhesus macaques, the most widely used non-human primate model of HIV/AIDS, SIV-infected monkeys with lower *CCL3L* copy numbers have been similarly shown to progress to simian-AIDS than those with higher copy numbers (DEGENHARDT *et al.* 2009). However, a later study demonstrated that rhesus macaques positive for *Mamu-A\*01*, an MHC class I allele associated with efficient control of SIV replication, possess higher *CCL3L* copy numbers than monkeys lacking this allele, and suggested that *CCL3L* copy number serves as a surrogate for the *Mamu-A\*01* allele and the association of simian-AIDS pathogenesis might be a consequence of expression of *Mamu-A\*01* and not *CCL3L* (LIM *et al.* 2010).

#### 1.4. THE *CCR5-CCL3L* CORECEPTOR-LIGAND AXIS AND HIV-1

Since both *CCR5* expression and *CCL3L* gene copy number variation are associated with outcomes of HIV-1 infection and that *CCR5* surface expression can also be influenced by gene copy number of its ligand, *CCL3L*, it follows that the effects of both factors on HIV-1 disease phenotypes may be more apparent when assessed in combination (GONZALEZ *et al.* 2005). When individuals were stratified into genetic risk groups according to whether they had high or low *CCL3L* copy number and detrimental or non-detrimental *CCR5* genotypes, relative to population-specific observations, HIV-1-infected individuals with high *CCL3L* gene copy number and non-detrimental *CCR5* genotypes (low genetic risk group) were shown to progress to AIDS-defining illnesses at much slower rates than other genetic risk groups (GONZALEZ *et al.* 2005). *CCL3L-CCR5* genetic risk groups can also influence viral burden, as seen in individuals in high genetic risk groups (*CCL3L<sup>low</sup>CCR5<sup>det</sup>*) who achieve steady-state viraemia post seroconversion as much as a year later than low genetic risk groups and also demonstrate higher viral burdens during the course of HIV-1 infection (DOLAN *et al.* 2007). Furthermore, HIV-1-infected individuals demonstrate differences in the extent of immune reconstitution (CD4+ T cell counts) after initiating highly active antiretroviral therapy (HAART), according to their *CCL3L-CCR5* genetic risk group, i.e., the low

genetic risk group ( $CCL3L^{high} CCR5^{non-det}$ ) had higher CD4+ T cells gains in comparison to the other genetic risk groups (AHUJA *et al.* 2008).

## 1.5. RATIONALE FOR THE CURRENT STUDY

Epidemiological data from 2011 estimated that 69% (23.5 million) of all HIV-infected individuals reside in sub-Saharan Africa (UNAIDS 2012b). With approximately 5.6 million South Africans living with HIV (prevalence of 17.3%), South Africa continues to have the largest HIV epidemic in the world (UNAIDS 2012b). Thus, given the central role of CCR5 and its ligand, *CCL3L*, in HIV-1 infection and that the effect of genetic polymorphisms and gene copy number variation, respectively, of these two molecules on susceptibility to HIV-1 and rate at which individuals progress to AIDS may differ according to the population groups, it is important to assess the extent of variability within South African populations. However, there is little information available with respect to *CCR5* and *CCL3L* variation within South African populations.

Many of the identified *CCR5* genetic variants are population specific (ANSARI-LARI *et al.* 1997; LIU *et al.* 2012; MAGIEROWSKA *et al.* 1999a; MASQUELIER *et al.* 2007; SHIODA *et al.* 2001). Population studies of *CCR5Δ32* show that it is present at an average allele frequency of 10% in Europe and yet it is very rare or absent in Africans (GALVANI and NOVEMBRE 2005; PETERSEN *et al.* 2001). Similarly, the prevalence of numerous other *CCR5* polymorphisms, with a potential impact on CCR5 expression and hence HIV-1 infection, differs according to the study population. Few studies have investigated the prevalence of *CCR5* gene polymorphisms in South African populations and these studies have been mainly restricted to polymorphisms within the ORF region (PETERSEN *et al.* 2001; WILLIAMSON *et al.* 2000). No studies have been conducted assessing the variability of the *CCR5* 5'- and 3'-UTR regions in South African populations.

Recent studies have highlighted the importance of the combined effect of receptor (*CCR5*) and ligand (*CCL3L* copy number) on HIV-1 disease pathogenesis (DOLAN *et al.* 2007; GONZALEZ *et al.* 2005). In addition, we and others have demonstrated that, not only do these molecules play a role in restricting viral entry and replication, but that they are also major determinants in cell mediated immunity (CATANO *et al.* 2011; DOLAN *et al.* 2007; SHALEKOFF *et al.* 2008). Individuals with more copies of *CCL3L* than their population median have been found to be less susceptible to HIV-1 infection (GONZALEZ *et al.* 2005). We have demonstrated that deficient CCL3 production is associated with intrapartum transmission, and that *CCL3L* copy number is independently associated with maternal-infant transmission (MEDDOWS-TAYLOR *et al.* 2006). Higher numbers of infant, but not maternal, *CCL3L* gene copies were associated with reduced HIV-1 transmission overall but the association was attenuated if mothers took single-dose nevirapine or if maternal viral load was low (KUHN *et al.* 2007).

In a study conducted on 71 HIV-1 chronically infected women, an influence of host *CCL3L* gene copy number (used as a surrogate for CCL3 production capacity) on disease progression and integrity of HIV-1 specific T cell responses was demonstrated (SHALEKOFF *et al.* 2008). *CCL3L* gene copy number correlated negatively with viral load, and *CCL3L* copy number greater than or equal to the population median of 5 was significantly associated with increased magnitude of CD4<sup>+</sup> Gag responses. Interestingly, women who had both CD4<sup>+</sup> and CD8<sup>+</sup> Gag-specific responses had significantly lower viral loads and higher *CCL3L1* copy number than those women with only CD8<sup>+</sup> Gag-specific responses. These data suggest that the integrity of HIV-specific CD4<sup>+</sup> T cell responses that are associated with control of infection are influenced by host gene duplications of *CCL3L*.

The insights gained over many years using maternal-infant HIV-1 transmission for studying the role of CC chemokines in protective immunity to HIV-1 have recently been published (TIEMESSEN and KUHN 2007). This commentary, in addition to outlining the complexities in studying the gene-protein relationships of these molecules, highlights the importance of ensuring that innate immunity becomes a more integral part of studies of HIV-1 vaccines, as understanding the interaction between innate and adaptive immunity may hold the key to understanding what constitutes protective immunity to HIV-1. In order to start to address how variation of *CCL3* and *CCL3L* genes from different hosts may impact on CCL3 protein production/infection outcome, the two functional genes *CCL3* and *CCL3L1* genes from 86 black and from 29 Caucasian South African individuals have been characterized by complete gene sequencing (PAXIMADIS *et al.* 2009).

It is within this context that this study was undertaken. The overall aim of this study was to determine the extent of *CCR5* gene variability within two ethnically divergent South African populations and to determine how host gene variation of *CCR5* translates to different expression levels of cell surface *CCR5* on various immune cell types. Additionally, the influence of differences in *CCL3L* gene copy number between populations on CCL3 chemokine production was investigated using the same cohort. Since HIV-1 infection can influence both *CCR5* and *CCL3* expression, these factors were assessed in healthy HIV-1 uninfected individuals.

The study of HIV-1 infected long term nonprogressors (LTNPs) has led to insights into various genetic determinants of HIV-1/AIDS susceptibility. Having had access to a small group of South African LTNPs cohort, a study of the potential role of differences in *CCR5* genotype, *CCR5* expression, *CCL3* genotypes and *CCL3* production levels in the control of HIV-1 infection was undertaken.

## 1.6. STUDY OBJECTIVES

The specific objectives of this study were as follows:

- i. To characterize *CCR5* genes within Caucasian and African HIV-1 uninfected individuals in South Africa.
- ii. To quantitate CCR5 receptor expression on different mononuclear cell types of the same individuals
- iii. To identify genetic determinants which translate to differential CCR5 expression by comparing data from i and ii above.
- iv. To use an *in vitro* promoter function assay to determine the effect of promoter polymorphisms on promoter function and, indirectly, expression of CCR5.
- v. To determine *CCL3L* gene copy numbers for the same individuals and further investigate the influence of *CCL3L* copy number on the production of CCL3, a ligand for CCR5.
- vi. To describe genotypic variants of *CCR5* and *CCL3L* in the context of HIV-1 infection in a cohort of long term nonprogressors.

---

## CHAPTER 2

### Genetic variation within the gene encoding the HIV-1 CCR5 coreceptor in two South African populations

---

---

*Published: Infection, Genetics and Evolution 10 (2010): 487-494.*<sup>1</sup>

---

<sup>1</sup> Data published was based on the genotyping of 35 South African African (SAA) and 35 South African Caucasian (SAC) individuals. These numbers were expanded to 41 SAA and 47 SAC individuals and values have been adjusted accordingly within this chapter. Furthermore, *CCR2V64I* genotyping was not included in this publication. The methodology for *CCR2V64I* genotyping was published as part of a later publication (Immunology 64 (2012): 795-806).

## ABSTRACT

Polymorphisms within the open reading frame as well as the promoter and regulatory regions can influence the amount of CCR5 expressed on the cell surface and hence an individual's susceptibility to HIV-1. In this study we characterize *CCR5* genes within the South African African (SAA) and Caucasian (SAC) populations by sequencing a 9.2kb continuous region encompassing the *CCR5* open reading frame, its two promoters and the 3' untranslated region. Full length *CCR5* sequences were obtained for 88 individuals (41 SAA and 46 SAC) and sequences were analyzed for the presence of single nucleotide polymorphisms (SNPs), indels and intragenic haplotypes. A novel SNP (+258G/C) within the ORF leading to a non-synonymous amino acid (Trp → Cys) change was detected in one Caucasian individual. Results demonstrate a high degree of genetic variation: 68 SNP positions, 4 indels, as well as the  $\Delta 32$  deletion mutant, were detected. Seven complex putative haplotypes spanning the length of the sequenced region have been identified. These haplotypes appear to be extensions of haplotypes previously described within *CCR5*. Two haplotypes, SAA-HHE and SAC-HHE were found in high frequency in the SAA and SAC population groups studied (18.29% and 20.21%, respectively) and share 4 SNP positions suggesting an evolutionary link between the two haplotypes. Only one of the identified haplotypes, SAA/C-HHC, is common to both study populations but the haplotype frequency differs markedly between the two groups (10.98% in SAA and 45.74% in SAC). The two population groups show differences in both haplotype arrangement as well as SNP profile.

## 2.1. INTRODUCTION

The human coreceptor, CCR5, acts as the principal coreceptor required for macrophage-tropic (R5) human immunodeficiency virus type 1 (HIV-1) virions to gain entry to a cell (DENG *et al.* 1996; DRAGIC *et al.* 1996). Shortly after the role of CCR5 was discovered, the *CCR5* $\Delta 32$  mutant and its association with protection to HIV-1 infection in individuals homozygous for this allele, was found (SAMSON *et al.* 1996b). This discovery provided the first genetic evidence of protection to HIV-1 infection and prompted further studies of the gene and how its naturally occurring mutations may influence the outcome of HIV-1 exposure and infection.

The *CCR5* gene is composed of four exons and two introns, where exons 2A and 2B are not interrupted by an intron (MUMMIDI *et al.* 1997). Exon 3 contains an intronless open reading frame (ORF). Two *CCR5* promoters have been described, a weak upstream promoter ( $P_U$  or P2) and a stronger downstream promoter ( $P_D$  or P1) (MUMMIDI *et al.* 1997). Cell surface expression of CCR5 is highly variable, even in individuals homozygous for the wild type ORF region. This may be explained by differences in the promoter region of *CCR5* which result in differing CCR5 expression levels (MENTEN *et al.* 2002). To date, several mutations and single nucleotide polymorphisms

(SNPs) in the HIV-1 coreceptor gene, *CCR5*, have been found to be important genetic factors capable of influencing susceptibility to HIV-1 infection or affecting the rate of disease progression.

Striking ethnic or population differences in SNPs frequencies of *CCR5* exist. The most studied polymorphism exhibiting this is the *CCR5* $\Delta$ 32 mutant. The *CCR5* $\Delta$ 32 allele occurs at a variable frequency of 4-15% in Caucasian populations, with an average of 10% in Europe (Reviewed in GALVANI and NOVEMBRE 2005) and yet is rarely found in Asian or African populations. In a South African context, Petersen et al. (2001) identified seven novel mutations within the *CCR5* ORF in African and Coloured populations, however polymorphisms within the promoter region have not been studied. Thus, a detailed descriptive study of SNPs and/or haplotypes within the *CCR5* receptor gene was carried out in two South African populations, South African Africans (SAA) and South African Caucasians (SAC), with the aim of providing a baseline study of the prevalence of polymorphisms that exist in *CCR5* within these populations and to determine whether all previously defined haplotypes are represented within these populations.

## 2.2. MATERIALS AND METHODS

### 2.2.1. Study population

Characterization of the *CCR5* gene was carried out on 88 healthy, HIV-1 uninfected adult volunteers, 41 were South African Africans (SAA) and 47 were South African Caucasians (SAC). This study was approved by the University of the Witwatersrand Committee for Research on Human Subjects, and informed written consent was obtained from all participants.

### 2.2.2. PCR and sequencing of *CCR5*

Genomic DNA was extracted from blood samples anticoagulated with ethylenediaminetetraacetic acid (EDTA) using the QIAamp DNA Mini Kit (QIAGEN, Dusseldorf, Germany). A ~9.2kb continuous region encompassing the *CCR5* open reading frame (ORF), its two promoters and the 3' untranslated region (UTR) was polymerase chain reaction (PCR) amplified in 5 overlapping sections using the Expand High Fidelity PCR System (Roche, Mannheim, Germany). PCR and sequencing primers were designed using PRIMER DESIGNER for Windows (v. 2.0) (Table A.1 in Appendix) using the published sequences for *CCR5* (GenBank accession: U95626, AF017632 (MORIUCHI *et al.* 1997), AF031236 & AF031237 (MUMMIDI *et al.* 1997)) as reference sequences.

All sequencing reactions were carried out using BigDye Terminator version 3.1 chemistry (Applied Biosystems, Foster City, CA, USA). Amplified fragments were sequenced using the automated 3100 Genetic Analyzer (Applied Biosystems).

### 2.2.3. Sequence analysis

Sequence data was assembled and analyzed for the presence of SNPs and indels using SEQUENCHER software version 4.5 (Gene Codes Corporation, Ann Arbor, MI, USA). Assembled sequences were aligned with each other and the published GenBank sequence, U95626, using SEQUENCHER, to identify polymorphisms. The GenBank NCBI SNP database (dbSNP) was searched for all reported SNPs in the *CCR5* gene to determine whether polymorphisms detected in this study had been previously reported.

The *CCR5* numbering system used in this study is as described by Mummidi et al. (2000) where the first nucleotide of the translational start site is designated as +1 and the nucleotide immediately upstream from that is -1. A composite of the reference sequences AF031236 & AF031237 (MUMMIDI *et al.* 1997) was used as a basis for determining SNP positions as these sequences appeared to be closer to the wild type (WT) or more 'ancestral' gene. It must be noted however that when all *Homo sapiens* reference sequences used in this study were aligned, a number of differences between them were noted, including base insertions or deletions (indels), which would affect the SNP position values. Also, AF031236 & AF031237 do not encompass the entire region sequenced in our study. Thus, using the sequences flanking the various SNPs may be a more reliable means of identifying SNP positions (Table A.2 in Appendix).

### 2.2.4. *CCR2-V64I*

We developed a real-time PCR assay to detect a single nucleotide polymorphism (SNP) in the *CCR2* coding region which forms part of the described HHF\*2 haplotype (G190A; *CCR2-V64I*) (GONZALEZ *et al.* 1999). The assay consists of allele-specific PCR with two allele-specific reverse primers (designed with their 3'-end bases complementary to one of the two SNP variants present), 5'-TTTGCAGTTTATTAAGATGAGGAY-3', and a common forward primer, 5'-AACGAGAGCGGTGAAGAAG-3'. The reverse primers each had a lock nucleic acid (LNA) modified 3'-end base. Two PCR reactions were thus conducted for each sample, each with the common forward primer and one of the reverse primers, wild type or mutant. Each 10 µl reaction conducted contained 1×Maxima SYBR Green/ROX qPCR master mix (Fermentas, Ontario, Canada), 0.3 µM solution of each forward and reverse primer, and ~30-80 ng DNA template. The PCR reactions were run in an Applied Biosystems 7500 Real-Time PCR system. An initial 10 min incubation at 95 °C was followed by 40 cycles: 15 s denaturation at 95 °C, annealing at 60 °C for 40 s followed by extension at 72 °C for 1 minute. To analyze the PCR data, the threshold, ( $C_T$ , the minimal fluorescence above which a sample is determined positive) was determined (using the default threshold setting) and used to calculate  $\Delta C_T$  values by subtracting the  $C_T$  of the 'wild type reaction' from the 'mutant reaction'. Heterozygous (G/A) individuals had  $\Delta C_T$  between 0 and 0.5,

whereas homozygous wild type (G/G) and mutant (A/A) individuals had  $\Delta C_T$  values of 7 to 10 and -7 to -8, respectively. Melt curve analysis was conducted to ensure that PCR amplified a single product, i.e., that there was no non-specific binding, and that there were no primer dimers formed.

#### *2.2.5. Determination of wild type (WT) reference sequence*

Once polymorphisms within *CCR5* were identified, it was necessary to determine which nucleotide to deem as the WT nucleotide. Generally, the most prevalent nucleotide in our combined populations was considered to be the WT or ancestral nucleotide/allele. In addition, to identify the WT nucleotide where it was not apparent which nucleotide/allele was most prevalent, the human *CCR5* sequences were aligned with those of the chimpanzee found on the sequence available for *Pan troglodytes* chromosome 3 (GenBank accession number: NW\_001232822.1) and the *Pan troglodytes CCR5* sequence (GenBank accession numbers: NM\_001009046 & AF005663).

#### *2.2.6. Inference of putative haplotypes*

Analysis of the sequence data generated for *CCR5* revealed certain obvious patterns wherein the presence of a polymorphism at one position was consistently associated with polymorphisms at one or more other positions. These associations were identified as putative intragene haplotypes. The HAPLOTYPER software which uses a Bayesian algorithm for haplotypes inference (NIU *et al.* 2002) was also used to infer haplotypes for *CCR5*.

The frequencies of putative intragenic haplotypes were calculated by counting the number of alleles harbouring the haplotypes and dividing by the total number of alleles. Counting of the haplotypes was irrespective of the presence of additional SNPs not forming part of the haplotypes in question.

#### *2.2.7. Characterization of indels*

Seven SAA and one SAC individual appeared to have a previously unidentified indel downstream from the open reading frame (+2772). Characterization of the putative indel as well as verification of Intron 2 indels was carried out by TA cloning of PCR amplicons into the pCR®4-TOPO® cloning vector using the TOPO-TA Cloning Kit (Invitrogen, Carlsbad, CA, USA). Recombinant plasmids were screened for the allele with the putative indel by sequencing. Sequences were aligned with reference sequences, using SEQUENCHER, in order to characterize the indel.

#### *2.2.8. Hardy-Weinberg equilibrium*

All polymorphic loci detected within the characterized *CCR5* gene region were tested for deviation from Hardy-Weinberg equilibrium using the conventional Monte Carlo exact test of Guo and

Thompson (1992) implemented through the computer program TFGA (Tools for Population Genetic Analyses version 1.3) (GUO and THOMPSON 1992; MILLER 1997). The two population groups were tested independently.

#### 2.2.9. Linkage disequilibrium between haplotype SNPs

To test whether the SNPs forming part of the putative intragenic haplotypes were in complete or strong linkage, linkage disequilibrium between every two SNP combination in each haplotype was estimated using the method described by Lewontin (1964) where the linkage disequilibrium coefficient  $D$  was calculated ( $D_{ij} = HF_{ij} - p_i p_j$ ).  $D$  was subsequently normalized ( $D'$ ) or standardized by the maximum value it can take ( $D_{max}$ ) using the formula  $D'_{ij} = D_{ij} / D_{max}$  where  $HF_{ij}$  is the frequency of the haplotypes carrying SNPs  $i$  and  $j$ ,  $p_i$  and  $p_j$  are the frequencies of SNPs  $i$  and  $j$ , respectively and  $D_{max}$  is either  $\min [p_i p_j, (1 - p_i)(1 - p_j)]$  if  $D_{ij} < 0$ , or  $\min [(1 - p_i) p_j, p_i(1 - p_j)]$  if  $D_{ij} > 0$  (LEWONTIN 1964).  $D'$  values are defined in the range  $[-1, 1]$  with a value of '1' representing perfect disequilibrium. The statistical significance of the linkage disequilibrium between each of the SNP pairs was evaluated by the approximate chi-square (LIAU *et al.* 1984).

#### 2.2.10. Fisher exact test

Fisher exact tests were performed using the Simple Interactive Statistical Analysis software (UITENBROOK 1997) to test whether there was any significant difference in SNP frequencies between this and other studies.

## 2.3. RESULTS

#### 2.3.1. Single nucleotide polymorphisms

Assembled sequences of the *CCR5* gene including promoter, coding and 3' UTR regions, from 88 HIV-1 uninfected individuals were analyzed for DNA polymorphisms, SNPs and indels. Across the entire 9.2kb region sequenced, 69 SNPs were identified. The positions and nucleotide (nt) changes are indicated in Figure 2.1. The identified polymorphisms were found across the entire sequenced region with the exclusion of exon 2B, a small region spanning 54 nucleotides. The majority of the polymorphisms were located in the intron and untranslated regions (UTR) of the gene and only 6 were located in the ORF (Figure 2.1). With regards to the two study populations, 60 and 38 polymorphisms were found in the SAA and SAC populations, respectively. Of the 69 identified SNPs, 46 have been previously described in the GenBank dbSNP database and by Petersen *et al.* (2001). Their corresponding accession numbers, where available, are shown in Table 2.1. To the best of our knowledge, with comparison to the GenBank dbSNP database and literature reports, 24 polymorphisms are newly identified and have been designated as newly identified (NI) in Table 2.1.

These NI polymorphisms were found in both population groups. Newly identified polymorphisms are also distributed across the entire gene although the majority are located in the 5' and 3' UTRs. Most NI SNPs (12/24, 50%) were found to be rare polymorphisms present in only one individual. The NI polymorphisms, -4808G/A, -4223C/T, -3886C/T, -2454G/A, -451C/T and +2066G/A, were detected in higher numbers (3 or more individuals each, all of which were heterozygous for the polymorphisms) in either/both populations.

The alignment of reference sequences (GenBank accession numbers: U95626, AF031236, AF031237, AF017632 and NT\_022517.17) did not demonstrate 100% homology at many of the SNP positions detected, as well as at other potential polymorphic positions not detected in this study. Also, it was not always apparent which the most predominant base at certain positions was. For instance, with the -2554G/T SNP, 100 (56.8%) and 61 (43.2%) alleles in this study contained a G and T nucleotide, respectively. Although the G allele was more frequent overall, in the SAC population neither nucleotide was predominant (51.1% G and 48.9% T) and in SAA individuals the major allele was a G (63.4%). Caution was necessary in the selection of a WT nucleotide as the population sizes used in this study were of a size where bias could be introduced and the apparent WT allele (most frequent) may not correspond to the ancestral allele. Thus, reference *Homo sapiens* sequences were aligned with *Pan troglodytes* sequences. Due to the low mutation rate since the human-chimpanzee divergence, the human allele almost always corresponds to the allele present in chimpanzees (CARGILL *et al.* 1999). Where an allele which was obviously the major allele in *Homo sapiens* (this study and sequences used as references) was not in agreement with that of the *Pan troglodytes* sequences at the same position, the former was selected as WT. This was the case for the indels, CTAT/-, AG/- and ACAA/G where published sequences for *Pan troglodytes*, some of the *Homo sapiens* reference sequences, as well as data for SAC would indicate the minor allele contained nucleotide insertions at these indel positions, yet overall the most frequent alleles detected in this study indicated that the minor allele for all 3 indel positions was the allele containing a nucleotide deletion at those positions. The same was observed for SNPs -2852A/G and -113G/T, where the *Pan troglodytes* base at that position was the equivalent of the minor allele/nucleotide in humans. The minor allele frequencies of the *CCR5* SNPs and indels in both the study populations are shown in Table 2.1.

Of the polymorphisms detected within the *CCR5* ORF, one has not been reported previously (Table 2.1). This novel SNP (+258G/C) within the ORF leads to a non-synonymous amino acid change (Trp → Cys) at codon 86 and was detected in one SAC individual heterozygous for that mutation. Four mutations in the open reading frame were detected in the SAA population: +225T/C; +319C/T; +673C/T and +1004C/T (S75S; L107F; R225X and A335V, respectively). One individual was found to be heterozygous for previously described mutations at both codon 107 and

225 (PETERSEN *et al.* 2001). In a study conducted by Petersen *et al.* (2001), these two mutations were reported as occurring simultaneously in an individual and at low frequencies in SAA and South African Coloureds. The codon 335 amino acid substitution mutation was detected within the SAA population at a frequency of 0.073 (n=6) and only one individual in the SAC population was found to harbour this mutation (Table 2.1). Previous reports looking at African American (ANSARI-LARI *et al.* 1997; CARRINGTON *et al.* 1999; CARRINGTON *et al.* 1997) and SAA (PETERSEN *et al.* 2001) populations found the mutation present at a frequency of approximately 3% and 2%, respectively. Although representation of this SNP appears higher (7.3%) in our study, this did not differ statistically from frequencies reported in these studies (P>0.05).

**Table 2.1.** Frequencies of identified polymorphisms within the South African African (SAA) and Caucasian (SAC) study populations

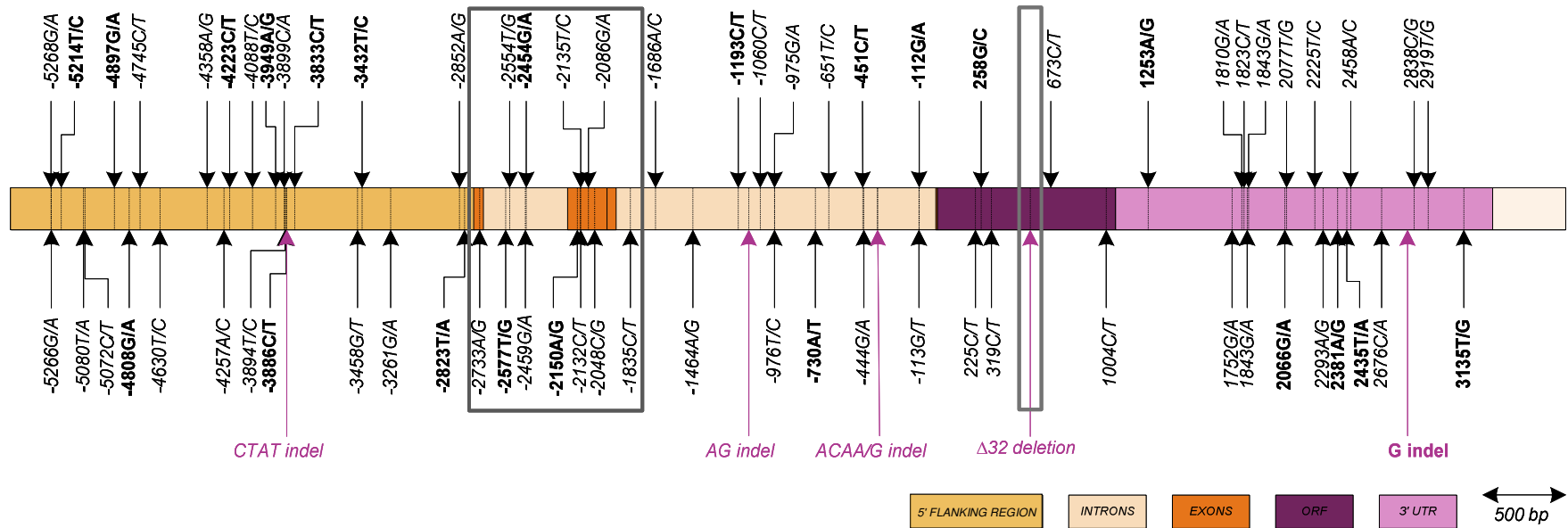
Location on gene	SNP Position	Base change (wt/mut)	Accession number <sup>a</sup>	n (population frequency) <sup>b</sup>	
				SAA <sup>c</sup>	SAC <sup>c</sup>
5' Flanking region (2762 bp)	-5268	G/A	rs3136535	3 (0.037)	12 (0.128)
	-5266	G/A	rs6776227	3 (0.037)	0
	-5214	T/C	NI	2 (0.024)	0
	-5080	T/A	rs41429449	4 (0.049)	1 (0.011)
	-5072	C/T	rs35078594	4 (0.049)	0
	-4897	G/A	NI	0	1 (0.011)
	-4808	G/A	NI	2 (0.024)	6 (0.064)
	-4745	C/T	rs3136536	10 (0.122)	0
	-4630	T/C	NI	0	1 (0.011)
	-4358	A/G	rs7637813	2 (0.024)	23 (0.245)
	-4257	A/C	rs41490645	0	12 (0.128)
	-4223	C/T	NI	4 (0.049)	0
	-4088	T/C	rs41499550	18 (0.220)	0
	-3949	A/G	NI	0	1 (0.011)
	-3899	A/C	rs72622924	14 (0.171)	49 (0.521)
	-3894	T/C	rs41395049	18 (0.220)	0
	-3886	C/T	NI	0	3 (0.032)
	-3868	CTAT/-	rs10577983	14 (0.171)	47 (0.500)
	-3833	C/T	NI	1 (0.012)	0
	-3458	G/T	rs2734225	13 (0.159)	46 (0.489)
-3432	T/C	NI	2 (0.024)	0	
-3261	G/A	rs41475349	18 (0.220)	0	
-2852	A/G	rs2227010	20 (0.244)	38 (0.298)	
-2823	T/A	NI	1 (0.012)	0	
Exon 1 (57 bp)	-2733	A/G	rs2856758	3 (0.043)	13 (0.138)
Intron 1 (501bp)	-2577	T/G	NI	1 (0.012)	0
	-2554	G/T	rs2734648	30 (0.366)	46 (0.489)
	-2459	G/A	rs1799987	33 (0.402)	44 (0.468)
	-2454	G/A	NI	3 (0.037)	1 (0.011)

Exon 2A (235bp)	-2150	A/G	NI	1 (0.012)	0
	-2135	T/C	rs1799988	33 (0.402)	44 (0.468)
	-2132	C/T	rs41469351	17 (0.207)	0
	-2086	A/G	rs1800023	13 (0.159)	46 (0.489)
	-2048	C/G	rs41355345	0	1 (0.011)
Intron 2 (1903bp)	-1835	C/T	rs1800024	13 (0.159)	6 (0.064)
	-1686	A/C	rs9282632	18 (0.220)	0
	-1464	A/G	rs3181037	18 (0.220)	0
	-1193	C/T	NI	2 (0.024)	0
	-1130	AG/-	rs3054375	13 (0.159)	46 (0.489)
	-1060	C/T	rs2856762	0	4 (0.043)
	-976	C/T	rs2254089	13 (0.159)	46 (0.489)
	-975	G/A	rs41395249	6 (0.073)	0
	-730	A/T	NI	2 (0.024)	0
	-651	C/T	rs2856764	13 (0.159)	46 (0.489)
	-451	C/T	NI	3 (0.037)	1 (0.011)
	-444	G/A	rs2856765/ rs35046662	13 (0.159)	46 (0.489)
	-362	ACAA/G	rs71619644	13 (0.159)	46 (0.489)
	-113	G/T	rs3176763	18 (0.220)	0
-112	G/A	rs41352147	1 (0.012)	0	
Exon 3/ORF (1059bp)	+225	T/C	rs1800941	1 (0.012)	0
	+258	G/C	NI	0	1 (0.011)
	+319	C/T	Petersen <i>et al.</i> (2001)	1 (0.012)	0
	+554	Δ32	rs333	0	10 (0.106)
	+673	C/T	Petersen <i>et al.</i> (2001)	1 (0.012)	0
	+1004	C/T	rs1800944	6 (0.073)	1 (0.011)
3' UTR (2651bp)	+1253	A/G	NI	1 (0.012)	0
	+1752	G/A	rs41495153	20 (0.244)	0
	+1810	G/A	NI	1 (0.012)	0
	+1823	C/T	rs17765882	0	4 (0.043)
	+1843	G/A	rs41418945	9 (0.110)	0
	+1846	G/A	rs41466044	9 (0.110)	0
	+2066	G/A	NI	4 (0.049)	0
	+2077	G/T	rs1800874	12 (0.146)	44 (0.468)
	+2225	T/C	rs41535253	4 (0.049)	0
	+2293	A/G	rs41526948	0	2 (0.021)
	+2381	A/G	NI	2 (0.024)	0
	+2435	T/A	NI	1 (0.012)	0
	+2458	A/C	rs3188094	7 (0.085)	0
	+2676	C/A	rs41442546	0	7 (0.074)
	+2772	G insertion	NI	7 (0.085)	1 (0.011)
	+2838	C/G	rs41512547	3 (0.037)	0
	+2919	T/G	rs746492	31 (0.378)	27 (0.287)
	+3132	T/G	NI	0	1 (0.011)

<sup>a</sup>Accession numbers of SNPs detected in this study which have been previously reported in the SNP database (dbSNP) or reference to report not in database are listed here; NI indicates newly identified polymorphisms not found in dbSNP.

<sup>b</sup>Frequency was calculated for both populations using total number of alleles, i.e., n=82 for SAA and n=94 for SAC.

<sup>c</sup>Grey shading highlights polymorphisms which were found to be restricted to either the SAA or the SAC population group.



**Figure 2.1.** Schematic representation of *CCR5* gene region sequences indicating polymorphism positions and nucleotide base changes at these points. The structure of the gene (MUMMIDI *et al.* 1997) is indicated by colour coded boxes. Nucleotide base changes at SNP positions are described by stating the WT base first followed by the base found on the minor allele. Grey boxes delineate regions of the *CCR5* gene which have been previously used to describe *CCR5* haplotypes. Newly identified (NI) polymorphisms and indels are indicated in bold font. UTR: untranslated region, ORF: open reading frame.

Several polymorphisms were found to be restricted to either the SAA or the SAC population group (frequencies highlighted in grey in Table 2.1). The SNPs, -3894T/C, -3261G/A, -2132C/T, -1686A/C, -1464A/G, -113G/T and +1752G/A, are all restricted to the SAA population at a frequency of >20% and all form part of the putative haplotypes, SAA-HHA and SAA-HHD, identified in this study. In addition, the SNPs, -4745C/T, +1843G/A and +1846G/A, were detected at a reasonably high frequency ( $\geq 11.0\%$  in SAA individuals but not in SAC individuals). Only one SNP, -4257A/C, found exclusively in the SAC population was also present at relatively high frequencies (i.e., at a frequency >10%) in that population. Where the SNP frequency is low in one population, absence in the other population cannot be used to state that that particular SNP is only prevalent in one population due to the sample size used in this study.

### 2.3.2. Indels

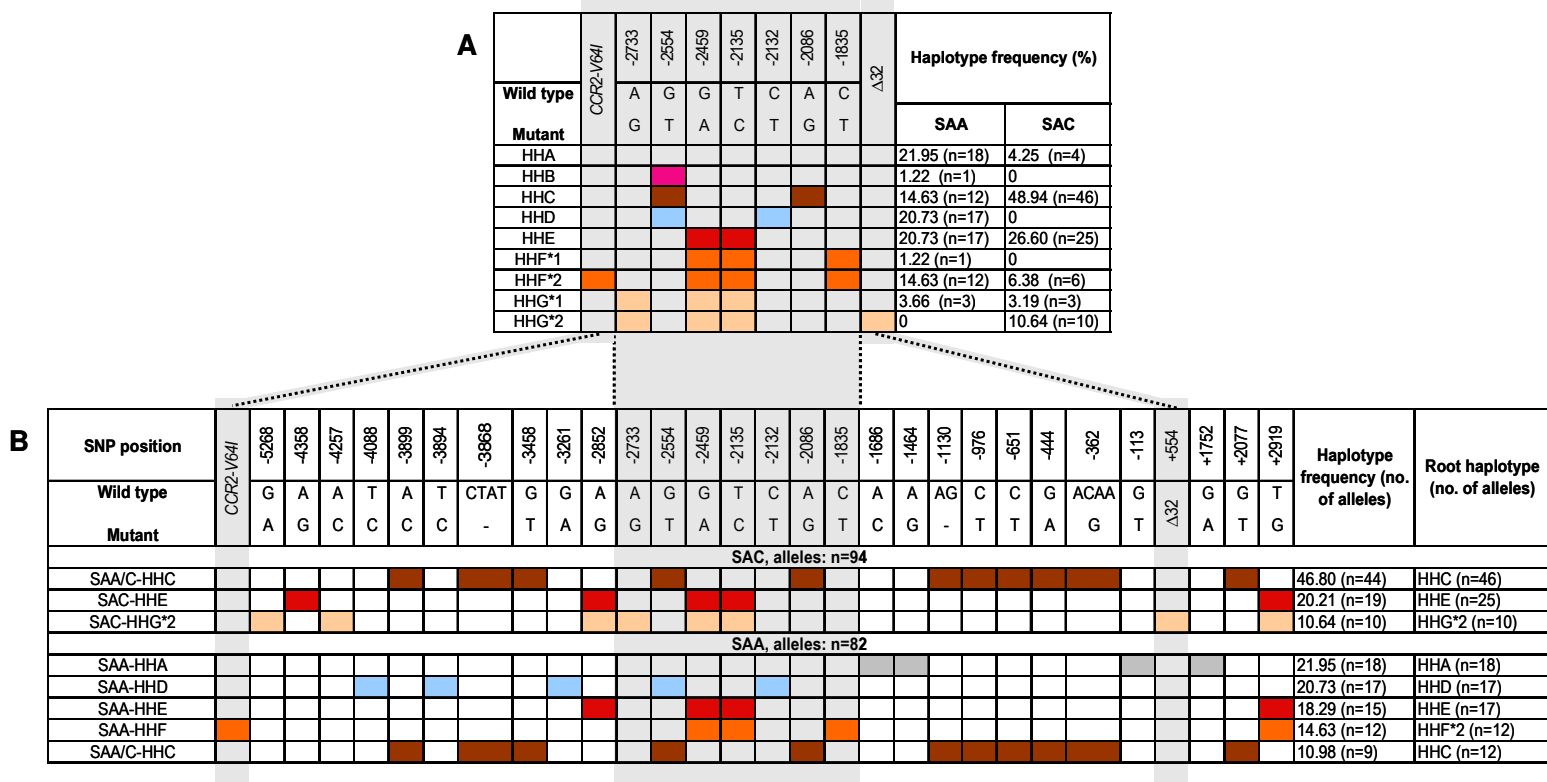
Five indels were detected across the entire *CCR5* gene (Figure 2.1). Four of the five detected indels have been previously described (DEAN *et al.* 1996; MUMMIDI *et al.* 1997; SAMSON *et al.* 1996b). The  $\Delta 32$  deletion indel was found in 9/47 (allelic frequency of 10.6%) SAC individuals, with one individual homozygous for this mutation, but not in SAA individuals. In a previous South African study, the *CCR5* $\Delta 32$  allele was detected at a frequency of 9.4% and 0.1% in SAC and SAA individuals, respectively (WILLIAMSON *et al.* 2000). Comparison of *CCR5* $\Delta 32$  allele frequencies observed in SAC populations in the two studies showed no significant difference between them ( $P=0.69$ ). The other three previously reported deletion indels (CTAT/-, AG/- and ACAA/G) appear to be in very strong linkage disequilibrium ( $D' = 1.0$ ,  $P < 0.0005$  for all 3 indel associations in both population groups).

Two indels are located within Intron 2 of the *CCR5* gene. The indel located at position -362 has been reported differently. In a report characterizing the *CCR5* gene this is shown to be a CAA indel (MUMMIDI *et al.* 1997). Within the dbSNP database there are two polymorphism reports for that location: accession number rs41515644 reports an A/G SNP at that position, whereas accession number rs71615644 reports that the four nucleotide sequence, ACAA, is substituted with a single guanine nucleotide. Within our study group, only the latter polymorphism was observed. The Intron 2 PCR amplicons from 2 individuals heterozygous for the ACAA indel were cloned and sequenced. Sequencing demonstrated that on the alleles where there was an ACAA deletion, there would be a G substitution at that point. Thus, in our report, the deletion and base substitution have been treated as a single polymorphism. It is possible though that the ACAA/G indel may have arisen as two separate events which became evolutionarily linked, i.e., an A to G substitution and a CAA deletion immediately downstream from the substitution.

The indel downstream from the open reading frame (Figure 2.1) consists of a single guanine insertion. Alleles containing the indel have a string of nine guanine bases in that region, whereas the WT alleles have eight guanine nucleotides. The exact position of the single base insertion within the eight consecutive guanine bases of the WT sequence cannot be precisely determined. Thus, the position of +2772, at the end of the 8 guanine bases has been selected.

### 2.3.3. Haplotypes

Individuals within the SAA and SAC populations were assigned to previously described haplogroups (GONZALEZ *et al.* 1999) based on SNPs at positions -2733, -2554, -2459, -2135, -2132, -2086 and -1835 as well as the presence of *CCR5* $\Delta$ 32 and *CCR2*V64I (Figure 2.2A). One SAA individual was found to be heterozygous for a haplotype allele which could not be classed into any of the haplotypes defined by Martin *et al.* (1998) or Gonzalez *et al.* (1999). Similar trends in haplotype frequency to that reported in a larger study conducted by Gonzalez *et al.* (1999 and 2001) were observed. In SAC, HHA appears to be underrepresented (4.3% vs. 10% reported in Caucasians (GONZALEZ *et al.* 2001)) and HHC as overrepresented (48.9% vs. 35% in Caucasians (GONZALEZ *et al.* 2001)) and in SAA, HHF appears underrepresented (15.9% vs. 24% reported in African non-pygmyes (GONZALEZ *et al.* 2001)) and HHG\*1 overrepresented at 3.7% (2% reported in African non-pygmyes (GONZALEZ *et al.* 2001)). Fisher exact test shows significant difference in HHC frequencies ( $P=0.006$ ) but no significant difference in the HHF, HHG\*1 and HHA frequencies between the two studies ( $P>0.05$ ). The overrepresentation of HHC haplotype frequency in the SAC population in this study could potentially be attributed to differences in Caucasian population ancestry in the two studies. In the larger Gonzalez *et al.* (2001) study, their Caucasian study group (n=959) is comprised of HIV-1 uninfected individuals from Finland, France and Poland and European American individuals of mixed infection status (i.e., both HIV-positive and HIV-negative individuals) (GONZALEZ *et al.* 2001). Another possible explanation would be the presence of a greater amount of admixture within the Caucasian study group in the Gonzalez *et al.* (2001) study. These previously defined haplotypes however are located in a relatively small region of the *CCR5* gene (898 bp of the regulatory region of *CCR5*, in addition to presence/absence of *CCR5* $\Delta$ 32 in ORF and *CCR2* V64I upstream on the same chromosome). This study has identified putative haplotypes which extend over the entire gene in both directions.



**Figure 2.2.** (A) Schematic representation of haplotypes previously defined by Gonzalez *et al.* (1999) and the frequency at which they were detected within our two study populations. (B) Haplotypes identified within the *CCR5* gene in the South African African (SAA) and Caucasian (SAC) study populations. Haplotypes were named by prefixing the root haplotype with the population within which it was found. Polymorphic positions as well as the wild type (WT) to mutant change are shown. Coloured boxes indicate SNPs or indels which form part of the haplotype and the frequency of occurrence in that particular population group is indicated. The root haplotype (GONZALEZ *et al.* 1999) is also indicated.

Seven complex putative haplotypes spanning the length of the sequenced region have been identified (Figure 2.2B). These haplotypes appear to be extensions of haplotypes previously described within *CCR5* (GONZALEZ *et al.* 1999) (HHA, HHC, HHD, HHE and HHG\*2). Haplotypes were named by prefixing the root haplotype name with the population within which it was found. Thus, a distinction can be made when haplotypes with the same root differ in SNP composition between study population groups (eg. SAA-HHE and SAC-HHE which are both rooted on the HHE haplotype but differ between SAA and SAC individuals bearing the HHE haplotype). Where haplotypes were found to be identical in both study populations, the prefix SAA/C- was used. All the haplotypes described in this study occurred at a frequency greater than 5% and one was found at a frequency of 45.74% (SAA/C-HHC in SAC individuals). Five predominant putative haplotypes were identified in the SAA population whereas only three were identified in the SAC population (Figure 2.2B). Only one haplotype appears to be shared by both study populations. This haplotype, SAA/C-HHC, is comprised of 3 indels and 8 SNPs and is the most frequent haplotype in SAC individuals (45.74%) and the least frequent in SAA individuals (10.98%) (Figure 2.2B).

Linkage disequilibrium analysis between every two SNP combination in haplotypes identified in this study demonstrated strong linkage disequilibrium between SNPs with a statistical significance greater than 95%.

All HHA, HHF, HHG\*2 and HHD haplotypes were found to be associated with further SNPs, forming haplotypes SAA-HAA, SAA-HHF, SAC-HHG\*2 and SAA-HHD, respectively (Figure 2.2B). The majority of HHC and HHE alleles are associated with further SNPs forming SAA/C-HHC, SAC-HHE and SAA-HHE. SAC-HHE and SAA-HHE are two putative haplotypes rooted on the HHE haplotype but differ at the inclusion of an additional SNP (-4358A/G) in SAC-HHE. The corresponding haplotypes occur at similar frequencies in the two populations (SAC-HHE: 20.21% in SAC and SAA-HHE: 18.29% in SAA).

All 17 alleles (SAA) classed as HHD can also be classed as SAA-HHD. Hence, a further 3 polymorphisms (-4088T/C; -3894T/C and -3261G/A) can be said to be associated with the HHD haplotype ( $D' = 1.0$ ,  $P < 0.0005$ , for all SNP associations within the haplotype). No HHD haplotypes were found in the SAC population. The HHF\*2 haplotype appears to be linked to the +2919T/G SNP forming SAA-HHF.

In the SAC population 44/46 alleles classed as haplotype HHC, could also be classed as the HHC extended haplotype, SAA/C-HHC, by far the most predominant haplotype in that population (Figure 2.2B). The remaining 2 HHC-bearing alleles occurred in two individuals homozygous for ten of the eleven polymorphism sites comprising SAA/C-HHC and were heterozygous for the +2077T/G SNP. Nine out of twelve HHC alleles in the SAA population exhibited the SAA/C-HHC

polymorphism pattern, two alleles lack the SNP at position +2077 and the third lacked the SNP at -3458.

The HHA haplotype, which comprises WT bases at all SNPs positions used in the Gonzalez *et al.* (1999) classification system, is present at a frequency of 21.95% and 4.25% in the SAA and SAC populations, respectively (Figure 2.2A). Although this may appear to imply that individuals harbouring the HHA haplotype are WT across the entire *CCR5* gene, this is not the case as the HHA haplotype appears to be associated with different SNPs in the extended haplotypes in the different populations (Figure 2.2B). In SAA individuals, HHA is associated with SNPs: -1686A/C, -1464A/G; -113G/T and +1752G/A forming SAA-HHA, whereas in SAC, HHA alleles demonstrate no association with those SNPs but are instead linked in a haplotype to C/T SNPs at positions -3886; -1060 and +1823, all of which are polymorphisms not detected within the SAA study group. It must be noted, however, that in the SAC population this is a rare haplotype (4.25%) and so has not been shown as one of the predominant haplotypes in Figure 2.2B and caution must be taken in assuming this is a true association.

The haplotype HHG can be subdivided into HHG\*1 (alleles not containing  $\Delta 32$  deletion in ORF) and HHG\*2 (alleles containing  $\Delta 32$  deletion in ORF). All HHG\*2 haplotypes detected in this study (n=10) were found to contain an additional 4 SNPs (-5268G/A, -4257A/C, -2852A/G and +2919T/G), forming SAA-HHG\*2 (Figure 2.2B). The one HHG\*2 homozygous individuals was also homozygous for the additional SNPs. Two of the three SAC HHG\*1 alleles were identical to SAA-HHG\*2 with one individual lacking both the -5268G/A and the -4257A/C polymorphisms. In contrast, the three SAA HHG\*1 alleles only had the additional polymorphisms, -5268G/A and -2852A/G, in common with the SAA-HHG\*2 haplotype.

#### 2.3.4. Hardy-Weinberg equilibrium

No significant deviations from Hardy-Weinberg equilibrium were noted for any of the indels or SNP loci detected in this study in both the SAA and SAC population groups.

## 2.4. DISCUSSION

In this study we have characterized polymorphisms (SNPs and indels) and intragenic haplotypes found within *CCR5* for two South African populations, SAA and SAC. This provides a baseline study for the *CCR5* polymorphism and haplotype profiles within these two populations. Previously unreported polymorphisms have been identified and previously defined haplotypes within the *CCR5* gene have been expanded upon.

There exists greater genetic diversity and low levels of linkage disequilibrium within African populations in comparison to European-originating populations (TISHKOFF *et al.* 2009; TISHKOFF and VERRELLI 2003). Hence, it is not unexpected to have found a greater number of polymorphisms in SAA in comparison to SAC individuals, as also observed in a recent study reported by Paximadis *et al.* (2009). Full length sequencing of the *CCR5* gene allows for identification of SNPs which would normally not be detected. Although a number of NI SNPs were identified in our study population, these may have been missed in previous studies which look at a smaller portion of the gene or which use other means of identifying specific polymorphisms. Also, most of the NI polymorphisms identified in this study were detected exclusively in the SAA population. Owing to the sample sizes in this study, detection of a SNP in only one of the two study population groups cannot be used to conclude that that SNP is absent in the other population, but it can be used as an indication of overall prevalence and diversity.

At SNP positions where the major allele (WT) in our study and that of other human reference sequences differed from that of the chimpanzee sequence, the chimpanzee sequence was found to correspond to the minor allele in humans (CTAT/-, AG/- & ACAA/G indels as well as -2852A/G & -113G/T SNPs). In a study characterizing the SNPs in 106 human genes, Cargill *et al.* (1999) noted that in a significant fraction of cases, the minor chimpanzee allele had become the major human allele and hence the minor human allele was in fact the older allele. This has also been observed in a study reporting variants in the *CCL3* and *CCL3L* genes which code for CCR5 ligands (PAXIMADIS *et al.* 2009).

The codon 335 mutation resulting in an alanine to valine (A335V) substitution was previously reported by Ansari-Lari *et al.* (1997) and has been found to be present at a higher frequency in African American populations in comparison to Caucasians (ANSARI-LARI *et al.* 1997; CARRINGTON *et al.* 1999; CARRINGTON *et al.* 1997). Although this mutation occurs in the ORF and could be thought to potentially affect protein structure and/or function, in a disease association study, this mutation has been found to have no effect on the rate of progression to AIDS (CARRINGTON *et al.* 1997). In a study conducted in South African populations, the A335V mutation was detected in African and Coloured populations but not in Caucasians (PETERSEN *et al.* 2001). Our study indicates that this mutation is in fact present in the SAC population but as a very rare polymorphism (only 1/47 individuals harboured this allele). In SAA individuals, this mutation was detected at a much higher frequency than that reported elsewhere (ANSARI-LARI *et al.* 1997; CARRINGTON *et al.* 1999; CARRINGTON *et al.* 1997; PETERSEN *et al.* 2001). However, comparison of A335V mutation frequencies in SAA individuals from this study to that of healthy SAA individuals in another South African study (PETERSEN *et al.* 2001) indicated no significant difference, although a trend was noted, between observed frequencies (7.3% vs. 1.6%;  $P=0.058$ ). This and the Petersen

*et al.* (2001) study were conducted in two widely separated geographical regions within South Africa, the Gauteng and Western Cape provinces, respectively.

A novel non-synonymous mutation has been detected within the ORF at codon 86. It is unclear whether this tryptophan to cysteine amino acid substitution will have an impact on chemokine receptor function. Amino acid alignment of chemokine receptors, CCR5, CCR2B, CCR1, CCR3, CCR4 and CXCR4 (CARRINGTON *et al.* 1997) shows that this mutation occurs within a highly conserved region (second transmembrane region) between the receptors at a point where all aligned proteins contain a tryptophan residue. This high level of conservation implies that this region is important to the structure or function of the protein and hence indicates that the significance of this novel mutation warrants further study. Both tryptophan and cysteine residues are hydrophobic molecules but cysteine is considerably smaller than tryptophan. Also, tryptophan residues positioned near lipid bilayers, as with residue 86, tend to form hydrogen bonds with the lipid head groups (SCHIFFER *et al.* 1992) whereas cysteine residues are likely to form disulphide bonds. Thus, it is possible that this amino acid change may have an impact on the folding of the peptide chain and hence its function as a receptor.

Several SNPs located in the *CCR5* promoter have been previously reported to affect the expression of CCR5. One such polymorphism is the -2459G/A polymorphism located within the downstream promoter (P1). This polymorphism has been linked to differences in *CCR5* expression levels on CD14<sup>+</sup> monocytes (SALKOWITZ *et al.* 2003) and has known association with the rate of progression to AIDS (MCDERMOTT *et al.* 1998). Individuals homozygous for the -2459G allele exhibit lower *CCR5* receptor density in CD14<sup>+</sup> monocytes (SALKOWITZ *et al.* 2003) and have been linked to slower disease progression (CLEGG *et al.* 2000; KNUDSEN *et al.* 2001; MCDERMOTT *et al.* 1998). Both WT and mutant alleles have been found to be present at high frequencies in all racial groups, with reported frequencies of 43% and 57% in African and Caucasian populations, respectively (MCDERMOTT *et al.* 1998). In this study, the -2459A allele was detected at a frequency of 40.2% and 46.8% in SAA and SAC population groups, respectively. Fisher exact analysis of the frequencies observed in the two studies has shown no statistical difference between the African population ( $P=0.75$ ) but there was a trend towards statistical difference between the Caucasian populations ( $P=0.063$ ). This is not unexpected as the WT alleles, -2459G and -2135T, which are in very strong linkage disequilibrium with each other (CLEGG *et al.* 2000; GONZALEZ *et al.* 1999), form part of the SAA/C-HHC haplotype which was present at a higher than expected frequency in the SAC population. Thus, it follows that the minor/mutant alleles at those positions will also be underrepresented. These two SNPs form part of the HHE, HHF and HHG haplotypes defined by Gonzalez *et al.* (1999) and the predominant putative haplotypes, SAC-HHE and SAC-HHG\*2 described in the SAC population.

Although studies looking at individual polymorphisms on susceptibility to HIV-1 infection and the rate at which individuals progress to AIDS, do provide useful information, it is not always possible to pinpoint the cause of the observed effect of a particular polymorphism when studying the regulatory region of the gene as different combinations of polymorphisms may be in linkage disequilibrium forming haplotypes. Thus, it is important to look at haplotypes and their prevalence across the breadth of a gene.

When examining the effects of *CCR5* haplotypes on HIV-1 disease in different population groups, Gonzalez *et al.* (1999) observed that haplotype diversity is greatest in African populations. This was also reflected in our study where five major (frequency >5%) haplotypes were observed in SAA individuals, whereas only three were found in SAC individuals.

In this report there are a number of polymorphisms where the minor allele in the SAA population has been shown to be the predominant or 'major' allele in the SAC population. This is most evident with the SNPs comprising the SAA/C-HHC haplotype. While SAA/C-HHC, and hence its associated polymorphisms, is by far the most prevalent haplotype detected in the SAC population (45.74%), in the SAA population this haplotype is much less prevalent (10.98%). The most prevalent haplotype in the SAA population is SAA-HHA which is an extension of the HHA haplotype reported to be the ancestral *CCR5* haplotype (MUMMIDI *et al.* 2000). This is likely to be due to different evolutionary pressures being exerted on the two populations or a genetic bottleneck where a significant number of members of one of the populations was unable to reproduce.

Previous reports have defined haplotypes within the *CCR5* gene (GONZALEZ *et al.* 1999; MARTIN *et al.* 1998). Martin *et al.* (1998) described 10 haplotypes, CCR5P1 – P10, comprising 10 SNP positions within the region starting at Exon 1 and ending in Exon 2B of the gene. The nine haplotypes described by Gonzalez *et al.* (1999) comprise 7 SNP positions within a similar region but extending slightly into Intron 2, the presence/absence of *CCR5* $\Delta$ 32 in addition to the *CCR2-64I* mutation. In this study we have expanded upon this and have linked previously defined haplotypes to SNPs both upstream and downstream of these regions forming haplotypes which extend over a larger region of the gene with SNPs linked in a haplotype being as much as ~8.1 kb apart (SNP positions -5268 and +2919 in the haplotype SAC-HHG\*2).

A better understanding of the role played by host genes in response to human immunodeficiency virus (HIV-1) exposure will contribute towards a better understanding of the protective immunity to HIV-1 and of the disease process in HIV-1 infected individuals. *CCR5* is increasingly being shown to play a critical and central role in HIV-1 infection and to date a number of genetic mutations within the gene have been found to positively or negatively influence an individual's susceptibility and rate

of disease progression. Thus, studies such as these which provide valuable new information regarding the genetic diversity within this gene, are important to the further understanding of the impact of CCR5 expression on host susceptibility to HIV-1.

---

## CHAPTER 3

***In vitro* functional analysis of different *CCR5* promoter regions and the impact of genetic variability present within the SAA/C-HHC, SAA-HHA and SAC-HHE haplotypes, on promoter function**

---

## ABSTRACT

Polymorphisms in the *CCR5* noncoding regions are among the factors known to influence *CCR5* expression. We selected three *CCR5* haplotypes, SAA-HHA, SAA/C-HHC and SAC-HHE, found at high frequencies in South African African (SAA) and South African Caucasian (SAC) populations, so as to evaluate the effect of single nucleotide polymorphisms (SNPs) on promoter activity, and hence, potential differences in *CCR5* expression. The promoter activity of four different *CCR5* promoter regions (P1A, P1B, P2 and P1+P2), for each haplotype being tested was evaluated. Promoter activity of the upstream promoter region (P2) was higher than that of the downstream promoter (P1A). When intron 2 was included with the downstream promoter (P1B), promoter efficiency decreased relative to that observed for the downstream promoter (P1A) on its own, indicating a possible negative regulatory role of intron 2. Haplotype-specific promoter analysis demonstrated the SAA-HHA haplotype to have the strongest promoter activity in THP-1 and K562 cells for both P1A and P2 promoter regions. However, results were variable for other cell lines tested (Jurkat and U937). A newly identified SNP, -2150A/G, was shown to downregulate promoter activity in the P1A, P1B and P1+P2 promoter regions in three of the four cell lines tested. This could potentially be attributed to binding of a MOK-2 transcription factor, as determined through a predictive bioinformatics tool. Contrastingly, in THP-1 cells, the -2150A/G SNP resulted in increased promoter activity. The observed variability in results between the cell lines tested highlights the importance of the cellular environment used for promoter expression analysis studies. The findings in this study demonstrate that polymorphisms across the entire *CCR5* promoter region, both in the downstream and the less well studied upstream promoter region, have the capacity to influence the efficiency of a promoter and consequently *CCR5* expression levels.

## 3.1. INTRODUCTION

*CCR5* gene expression can be regulated at several different levels, i.e., through transcriptional, translational, post-translational or through epigenetic mechanisms. A recent report has demonstrated that epigenetic modifications of the *CCR5* P1 promoter region contribute towards the control of *CCR5* expression in immune effector cells (WIERDA *et al.* 2012). At the protein level, *CCR5* expression is thought to be regulated, through interaction with its  $\beta$ -chemokine ligands, CCL3, CCL4 and CCL5, which results in subsequent internalization of the receptor (MACK *et al.* 1998; RILEY *et al.* 1997). The *CCR5* $\Delta$ 32 polymorphism, resulting in the truncation of the expressed protein, is the most well-known *CCR5* polymorphism shown to have an impact on cell surface expression of the *CCR5* coreceptor. Individuals bearing this mutation have been demonstrated to have reduced or no cell *CCR5* surface expression if they are heterozygous or homozygous for this mutation, respectively (DE RODA HUSMAN *et al.* 1999; LIU *et al.* 1996; PICTON *et al.* 2012b; SHALEKOFF and TIEMESSEN 2003; WU *et al.* 1997). In addition to polymorphisms within the coding

region, several *CCR5* polymorphisms in the noncoding region of the gene have been shown to have an impact on *CCR5* expression. Polymorphisms in the *CCR5* promoter region can influence *CCR5* promoter activity at the transcriptional level and thus affect *CCR5* expression. Differences in *CCR5* surface expression could explain the reported associations between various promoter polymorphisms and/or haplotypes and differences in HIV-1 susceptibility and disease progression (CLEGG *et al.* 2000; GONZALEZ *et al.* 1999; KASLOW *et al.* 2005; KNUDSEN *et al.* 2001; MCDERMOTT *et al.* 1998), virus replication (KAWAMURA *et al.* 2003; NGUYEN *et al.* 2004; SALKOWITZ *et al.* 2003; TANG *et al.* 2002), virus transmission (KOSTRIKIS *et al.* 1999; MALHOTRA *et al.* 2011; MANGANO *et al.* 2001; SINGH *et al.* 2008) and the effect of antiviral therapy on immune reconstitution (AHUJA *et al.* 2008).

*CCR5* transcription is initiated at multiple sites via two promoters, an upstream promoter ( $P_U$  or  $P_2$ ) and a downstream promoter ( $P_D$  or  $P_1$ ) (MUMMIDI *et al.* 1997). The use of alternate promoters provides important mechanisms by which either tissue- or cell-type specific expression, expression levels and translational efficiencies of resultant mRNAs can be regulated (Reviewed in AYOUBI and VAN DE VEN 1996). 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  $P_2$  promoter undergo alternative splicing, giving rise to two transcripts, *CCR5A* and *CCR5B*, which differ by the presence or absence of the untranslated exon 2A, respectively (MUMMIDI *et al.* 1997). Interestingly, both *CCR5A* and *CCR5B* transcripts contain AUG start codons upstream of the start codon for the *CCR5* open reading frame (ORF). It has been suggested that these elements may result in the generation of short upstream ORFs, thus leading to a negative effect on the efficiency of translation (MUMMIDI *et al.* 1997). This inhibitory effect may potentially be counteracted by the  $P_1$  promoter, since AUG triplets are absent in mRNA transcripts initiated at the  $P_1$  promoter (MUMMIDI *et al.* 1997). Furthermore, it has been suggested that differences in the secondary structures of the 5' untranslated region of the distinct *CCR5* mRNA transcripts may influence translation efficiency (MUMMIDI *et al.* 1997).

Promoter expression studies conducted in transformed cell lines, THP-1 and Jurkat, demonstrated the downstream promoter,  $P_1$ , to be the strongest promoter of the two alternate *CCR5* promoters (LIU *et al.* 1998; MUMMIDI *et al.* 1997). Furthermore, in an unstimulated state, *CCR5* mRNA transcripts isolated from both isolated peripheral blood mononuclear cells (PBMCs) and purified CD4+ T cells, are comprised mainly of  $P_1$  initiated transcripts (MUMMIDI *et al.* 2007). However, the  $P_2$ -driven transcripts, *CCR5A* and *CCR5B*, have been demonstrated as important determinants of *CCR5* expression on activated primary T cells, i.e., stimulation of T cells results in the induction of the  $P_2$  promoter and consequently the expression of exon 1-containing mRNA transcripts (MUMMIDI *et al.* 2007). Importantly, the gain in expression of *CCR5A* and *CCR5B* transcripts in

memory T cells was found only in those cells that also expressed CCR5 on the surface (MUMMIDI *et al.* 2007).

A number of studies have examined the role of various transcription factors in *CCR5* expression and several transcription factors binding to either the P1 or the P2 promoters have been identified. The KLF2 transcription factor, which has binding sites within both P1 and P2 *CCR5* promoter regions, modulates CCR5 expression in CD4+ T cells and has been shown to influence susceptibility of these cells to infection by HIV-1 R5 strains *in vitro* (RICHARDSON *et al.* 2012). The P1 promoter contains potential binding sites for several transcription factors including STAT, NF $\kappa$ B, AP-1, NF-AT and CD28RE (LIU *et al.* 1998). Mutation within the binding sites for these transcription factors negatively affects the expression of CCR5 (LIU *et al.* 1998). Transcription factors thought to be upregulated upon T-cell activation, NF $\kappa$ B, AP-1, Oct-1 and Oct-2, have been shown to bind to the *CCR5* P2 region (MUMMIDI *et al.* 2007). The Oct-1 and Oct-2 transcription factors bind to identical *cis* motifs, however they have opposing effects on CCR5 surface expression: Oct-2 results in upregulation, while Oct-1 transcription results in downregulation of CCR5 expression (MORIUCHI and MORIUCHI 2001; MUMMIDI *et al.* 2007). The presence of an Egr-1 binding site in P2, which may mediate CCR5 upregulation in THP-1 cells upon treatment with human amyloid peptides, has also been identified (GIRI *et al.* 2004). Two transcription factors, GATA-1 and p65(RelA), have been shown to stimulate CCR5 expression in transformed T lymphoid cell lines (LIU *et al.* 1998; MORIUCHI *et al.* 1999). In primary human T and dendritic cells, however GATA-1 represses CCR5 expression (SUNDRUD *et al.* 2005). Furthermore, a recent report has identified differences in the transcription factors which bind to the *CCR5* HHA and HHE haplotypes, thus providing evidence that CCR5 variation in expression can be regulated at the transcriptional level by the binding of specific transcription factors to SNPs within a portion of the *CCR5* promoter region (JIANG *et al.* 2011).

Previously, we identified three haplotypes, SAA-HHA, SAA/C-HHC and SAC-HHE, as collectively the most prevalent haplotypes in two populations, South African Africans (SAA) and South African Caucasians (SAC) (PICTON *et al.* 2010, Chapter 2). These haplotypes span the entire *CCR5* gene and are extensions of previously defined haplotypes restricted to the P1 region (GONZALEZ *et al.* 1999; PICTON *et al.* 2010). Thus, we undertook the current study in order to evaluate differences in promoter activity between these three haplotypes in our two study populations. Previous studies have assessed promoter expression differences between *CCR5* haplotypes described by Gonzalez *et al.* (1999), by assessing the promoter activity of the P1 region. We sought to extend this by investigating the potential influence of other SNPs, either upstream or downstream of the P1 region, found to be in complete linkage disequilibrium with the SNPs which define the *CCR5* haplotypes, HHA, HHC and HHE. Furthermore, we conducted these experiments in several

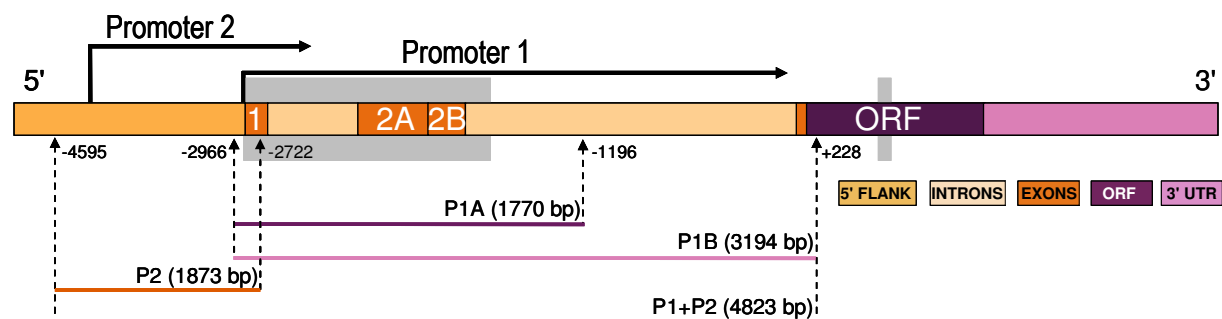
different cell lines which potentially differ in their transcription factor composition, a factor which is likely to determine the efficiency of a particular promoter.

## 3.2. MATERIALS AND METHODS

### 3.2.1. Construction of CCR5 Promoter Constructs

Four different promoter regions were used for generating promoter constructs of three of the most prevalent haplotypes found in the South African African and South African Caucasian populations, namely, SAA/C-HHC, SAA-HHA and SAC-HHE, with the aim of examining the effects of polymorphisms in these regions on promoter activity (Figure 3.1). Prepared promoter fragments of interest were cloned into the promoterless pGL4.10 vector (Promega, Fitchburg, Wisconsin, USA), upstream of a firefly luciferase gene, *luc2* (*Photinus pyralis*). The induction of the luciferase gene serves as an indicator of promoter function. Two strategies, depending on the promoter regions, were employed for generating promoter constructs into the promoterless pGL4.10 vector (Promega) upstream of the *luc2* gene.

Primers were designed using PRIMER DESIGNER for Windows (v. 2.0) using the published sequences for *CCR5* (GenBank accession: U95626, AF017632 (MORIUCHI *et al.* 1997), AF031236 and AF031237 (MUMMIDI *et al.* 1997)) as reference sequences (Table B.1 in Appendix). The *CCR5* numbering system used in this study is as described previously (MUMMIDI *et al.* 2000). The first nucleotide of the translational start site is numbered +1 and the nucleotide immediately upstream from that is -1.



**Figure 3.1.** Schematic representation of *CCR5* gene regions used for promoter expression analysis. Black arrows above the gene indicate positions of the two identified *CCR5* promoters, Promoter 1 (or downstream promoter, Pd) and Promoter 2 (or upstream promoter). Grey boxes delineate regions (-2733 to -1835 and *CCR5* $\Delta$ 32) which have been previously used to describe *CCR5* haplotypes (GONZALEZ *et al.* 1999). Coloured bars under the gene indicate promoter regions used to construct *CCR5* promoter constructs P1A, P1B, P2 and P1+P2. Promoter construct sizes (base pair, bp) and gene structure are shown. Figure is not to scale.

### 3.2.1.1. P1A and P2

The smaller promoter regions, P1A and P2, (Figure 3.1), were PCR amplified (Expand High Fidelity PCR System, Roche, Basel, Switzerland) using the primers described in Supplementary Table B.1, and amplicons were gel purified (QIAquick gel extraction kit, Qiagen, Hilden, Germany) from low melting temperature agarose gels. Purified amplicons were first TA cloned into the pcDNA3.1/V5 vector (Invitrogen, Carlsbad, California, USA) and transformed into XL1-Blue Supercompetent cells (Stratagene, La Jolla, California, USA) according to manufacturer's instructions. To select for clones containing the desired insert in the correct orientation, clones were initially screened by PCR using a plasmid-specific primer (BGH reverse) and an insert/promoter-specific primer (Supplementary Table B.2). Positive clones were randomly selected, cultured and plasmids extracted from these were sequenced to confirm nucleotide fidelity. The promoter regions were reisolated by sequential digestion with *KpnI* and *XhoI* enzymes (Promega) and cloned into the pGL4.10 vector (Invitrogen) prepared by digestion with the same enzymes.

### 3.2.1.2. P1B and P1+P2

The larger promoter inserts, P1B and P1+P2, were amplified using PrimeSTAR™ HS DNA polymerase (Takara, Otsu, Shiga, Japan) using primers designed with built-in restriction sites at their 5' ends (Table B.1 in Appendix). The forward primer included a *KpnI* restriction site whereas the reverse primer included a *Sall* restriction site. Although the recognition sequences for the *XhoI* and *Sall* enzymes differ, the nucleotide overhangs generated upon digestion are complementary allowing for ligation of DNA fragments digested with the two different enzymes. *Sall* was used instead of *XhoI* in the reverse primer as these constructs contained internal *XhoI* restriction sites. Amplified fragments were sequentially digested with *KpnI* and *Sall* enzymes (Promega), gel purified from low melting temperature agarose gels, ligated into pGL4.10 (Promega) as for the smaller fragments and transformed into XL1-Blue Supercompetent cells (Stratagene).

Clones were screened for the presence of the desired plasmid insert using two primers binding within the cloned promoter regions (Table B.2). Randomly selected clones were then cultured and plasmids extracted from these were sequenced to confirm nucleotide fidelity. All sequencing reactions were carried out using BigDye Terminator version 3.1 chemistry (Applied Biosystems, Foster City, CA, USA). Amplified fragments were sequenced using the automated 3100 Genetic Analyzer (Applied Biosystems).

Plasmids harbouring all promoter constructs (P1A, P1B, P2 and P1+P2), verified to be error-free, were then extracted using EndoFree Plasmid Maxi Kits (Qiagen) to avoid any possible introduction of bacterial endotoxins into the cell lines that plasmids were subsequently transfected into.

### 3.2.2. Cell lines

Four human cell lines, K562, U937, Jurkat E6-1 and THP-1 (ATCC, Manassas, VA, USA), were used to test the promoter efficiency of the different constructs. The K562 cell line was grown in IMDM supplemented with 10% fetal bovine serum (FBS), U937 and Jurkat cells were maintained in RPMI-1640 supplemented with 10% FBS and THP-1 cells were maintained in RPMI-1640 supplemented with 10% FBS and 0.05mM 2-mercaptoethanol.

### 3.2.3. Transfection of cell lines and Luciferase assays

We employed the dual Firefly and Renilla Luciferase Reporter Gene Assay to evaluate the activity of the different *CCR5* promoter constructs. Transfections were performed using Nucleofector™ Solutions V and C (Lonza, Basel, Switzerland) and the Nucleofector™ 2b device (Lonza), according to the manufacturer's instructions. An internal control reporter vector, pGL4.73 (Promega), containing a different luciferase gene *Renilla hRluc* (*Renilla reniformis*) downstream of a Simian Virus 40 (SV40) promoter was cotransfected with the *CCR5* promoter vectors. The SV40 promoter allows for strong constitutive production of the *Renilla* luciferase gene, thus providing an internal control value to which the experimental firefly luciferase gene may be normalised. Use of an internal control reporter controls for variability in the amount of DNA transfected into cells. A ratio of 1:1 test vector:control vector was used and the total amount of DNA used in transfections was kept constant (500 ng for THP-1 cells, 2.5 µg for other cell lines).

After 24h incubation at 37° C and 5% CO<sub>2</sub>, transfected cells were pelleted (200 x g, 10 min, room T), washed in phosphate buffered saline (PBS) and lysed in 1 x passive lysis buffer (Promega). Twenty microliters from each lysate was used to measure Firefly and Renilla luciferase activities. The *luc2* luminescence was measured with the luciferase assay reagent II (Promega). Quenching of the Firefly luciferase and concomitant activation of Renilla luciferase was achieved by adding Stop & Glo® reagent (Promega). Luminescence was measured in a luminometer (Victor X2, PerkinElmer) according to manufacturer's instructions (Dual Luciferase Reporter Assay System, Promega).

In each replicate, a control transfection of a promoterless vector (pGL4.10) cotransfected with the pGL4.73 internal control plasmid was included. This was included as a control for background promoter activity of the plasmid into which the test promoter constructs were cloned.

Transcriptional activity was expressed as relative luciferase units. The relative promoter/luciferase activity was determined as follows:

$$\frac{\text{CCR5 promoter construct firefly luciferase activity}}{\text{cotransfected pGL4.73 renilla luciferase activity}} \div \frac{\text{firefly luciferase of promoterless vector pGL4.10}}{\text{cotransfected pGL4.73 renilla luciferase activity}}$$

Results presented are based on five separate replicates for the K562 and THP-1 cell lines and six replicates for the THP-1 and Jurkat cell lines.

#### 3.2.4. Statistical analysis

Comparisons between the relative luciferase activity of each of the test promoter constructs were analysed using Mixed Linear Model for a randomised block design, with the blocking variable (repetition) treated as a random effect (Modified one-way ANOVA).

Bonferroni correction was used to control for multiple comparisons of the different constructs. Thus, relationships for the P1A, P1B and P1+P2 constructs were considered significant if  $P < 0.008$ , based on multiple comparisons between the SAA-HHA “wild type”, SAA-HHA “mutant”, SAA/C-HHC and SAC-HHE constructs. Comparisons within the P2 region were considered significant if  $P < 0.017$ , based on multiple comparisons between the SAA-HHA “wild type”, SAA/C-HHC and SAC-HHE constructs.

#### 3.2.5. Bioinformatics analysis

Identification of potential transcription factors and their binding sites within the *CCR5* promoter regions used to generate promoter constructs was performed using the application MATINSPECTOR (Genomatix, Munich, Germany) (CARTHARIUS *et al.* 2005).

### 3.3. RESULTS

#### 3.3.1. Efficiency of the different promoter constructs

CCR5 haplotypes are associated with altered CCR5 expression. Both *in vitro* and *in vivo* experiments have shown this to be cell-type specific (MOORE *et al.* 1997; MUMMIDI *et al.* 2007; PICTON *et al.* 2012a). Thus, we selected four different cell lines, K562, THP-1, U937 and Jurkat E6-1, to examine the effect of haplotype-associated SNPs on CCR5 expression. Both THP-1 and U937 cells are monocytic cell lines. However, THP-1 and U937 cells represent different stages of monocytic development, monocytes and promonocytes, respectively (TSUCHIYA *et al.* 1980). K562 is an erythroleukemia cell line and Jurkat is a T lymphocyte cell line. The cell lines chosen

represent cell types known to express CCR5 and are thus likely to have the requisite transcription factors.

Fragments of interest, P1A (1770 bp), P1B (3194 bp), P2 (1873 bp) and P1+P2 (4823 bp) (Figure 3.1), were amplified from genomic DNA extracted from individuals homozygous for three haplotypes, SAA/C-HHC, SAA-HHA and SAC-HHE, which we previously described (PICKTON *et al.* 2010, Chapter 2). These haplotypes are extensions of previously defined haplotypes (GONZALEZ *et al.* 1999). The SAA/C-HHC haplotype is an extension of the HHC haplotype found to be associated with the same SNPs in both SAA and SAC individuals. SAC-HHE was found in SAC individuals and differs from SAA-HHE in that it lacks the -4358A/G SNP. The third haplotype selected, SAA-HHA, was found in SAA individuals. Since the HHA haplotype occurs at low frequencies in the SAC population, we were unable to verify whether the extended haplotype, SAA-HHA, is also found in SAC individuals. Where present in SAC individuals, the HHA haplotype appears to be in linkage disequilibrium with different SNPs from those found in the extended SAA-HHA haplotype (PICKTON *et al.* 2010).

The P1A region encompasses the *CCR5* downstream promoter (P1) and includes the haplotype region as defined by Gonzalez *et al.* (1999). The P1B promoter region includes P1A and *CCR5* intron 2 which is thought to have a negative regulatory effect on transcription (LIU *et al.* 1998). Fragment P2 encompasses the upstream promoter. Lastly, the entire promoter region was also assessed in the form of fragment P1+P2. PCR amplified promoter fragments were cloned upstream of a *luc2* (firefly luciferase) gene. Thus, the luminescence from the firefly luciferase reaction serves as an indication of the promoter strength of the test promoter.

To compare promoter strengths of the different promoter regions, we selected the SAA-HHA haplotype, which is an extension of the HHA haplotype, and compared the promoter strengths of different promoter regions derived from this haplotype. The human HHA haplotype has been defined as ancestral to all other *CCR5* haplotypes through comparison to *CCR5* haplotypes of non-human primates (MUMMIDI *et al.* 2000). The promoter expression levels of the different constructs relative to each other are shown in Figure 3.2. Results were consistent in the K562, U937, THP-1 and Jurkat cells, with P2 fragments having the strongest promoter activity and P1+P2 fragments exhibiting the lowest promoter activity (Figure 3.2). Among the cell lines tested, promoter activity was strongest in THP-1 cells, followed by U937 cells (Figure 3.2). The weakest promoter activity was observed in Jurkat cells (Figure 3.2).

The percentage relative standard deviation, calculated as the percentage standard deviation relative to the mean, was used as a measure of the variability, or precision, between replicates. Inter-experimental variability, as measured by percentage relative standard deviation, was highest

in the U937 and Jurkat cell transfections and lowest in THP-1 cells (Table 3.1). It is important to note that where the mean relative luciferase activity of a particular assay is low, the relative standard deviation is more likely to be high.

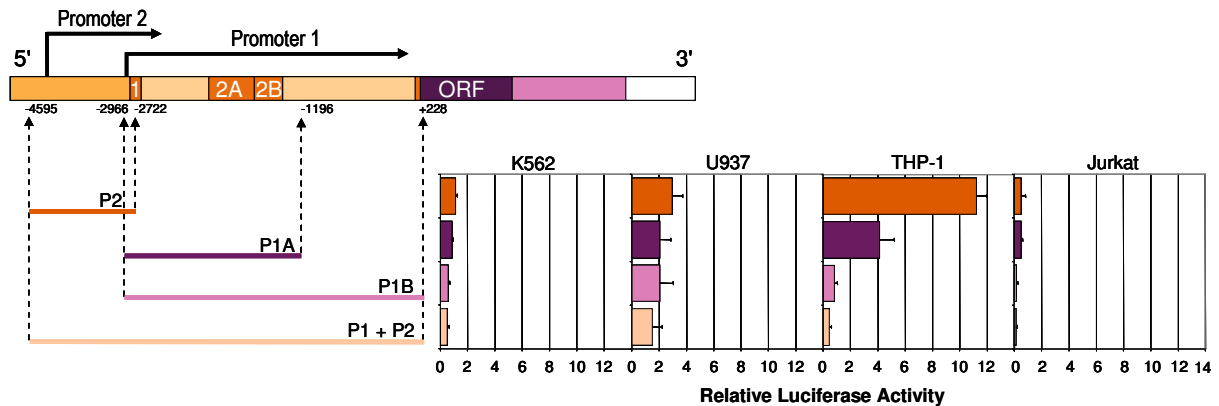
The upstream promoter, P2, demonstrated stronger promoter activity compared to that of the downstream promoter (Figure 3.2). The inclusion of *CCR5* intron 2 in downstream promoter constructs resulted in reduced promoter activity (Figure 3.2).

**Table 3.1.** Mean relative luciferase activity of the various *CCR5* promoter constructs transfected into K562, U937, THP-1 and Jurkat cell lines

Cell line	Promoter region	Haplotype					
		(Mean relative luciferase activity $\pm$ SD, % <i>Relative standard deviation</i> )*					
		SAA-HHA		SAA/C-HHC		SAC-HHE	
K562	P1A	0.85 $\pm$ 0.08	9.19	0.79 $\pm$ 0.19	24.50	0.64 $\pm$ 0.04	5.79
	P1B	0.56 $\pm$ 0.13	22.97	0.62 $\pm$ 0.08	12.13	0.64 $\pm$ 0.16	25.25
	P2	1.11 $\pm$ 0.14	12.45	0.85 $\pm$ 0.06	7.44	1.00 $\pm$ 0.17	17.44
	P1+P2	0.47 $\pm$ 0.18	38.29	0.31 $\pm$ 0.15	48.35	0.62 $\pm$ 0.29	46.86
U937	P1A	2.03 $\pm$ 0.82	40.50	1.71 $\pm$ 0.39	22.50	2.45 $\pm$ 1.30	53.08
	P1B	2.06 $\pm$ 0.97	47.17	1.58 $\pm$ 0.41	26.11	2.06 $\pm$ 0.82	39.60
	P2	2.96 $\pm$ 0.75	25.47	3.34 $\pm$ 1.77	52.89	3.41 $\pm$ 1.97	57.68
	P1+P2	1.49 $\pm$ 0.72	48.46	1.88 $\pm$ 0.67	35.77	2.10 $\pm$ 1.02	48.30
THP-1	P1A	4.13 $\pm$ 1.09	26.35	3.18 $\pm$ 0.63	19.93	4.00 $\pm$ 0.71	17.86
	P1B	0.81 $\pm$ 0.22	27.29	0.87 $\pm$ 0.13	14.95	0.77 $\pm$ 0.23	29.38
	P2	11.21 $\pm$ 0.75	6.72	7.98 $\pm$ 0.49	6.10	9.03 $\pm$ 1.37	15.16
	P1+P2	0.44 $\pm$ 0.16	35.54	0.40 $\pm$ 0.12	28.60	0.47 $\pm$ 0.15	31.74
Jurkat	P1A	0.49 $\pm$ 0.12	25.31	0.50 $\pm$ 0.15	30.62	0.38 $\pm$ 0.10	25.40
	P1B	0.16 $\pm$ 0.10	67.36	0.13 $\pm$ 0.06	45.29	0.13 $\pm$ 0.04	32.66
	P2	0.48 $\pm$ 0.34	70.62	0.44 $\pm$ 0.38	86.48	0.51 $\pm$ 0.28	54.53
	P1+P2	0.12 $\pm$ 0.09	77.43	0.08 $\pm$ 0.05	63.47	0.13 $\pm$ 0.04	33.88

\* Relative standard deviation was determined as follows:

$$100 \times [\text{standard deviation (SD)} \div \text{mean relative luciferase activity}]$$



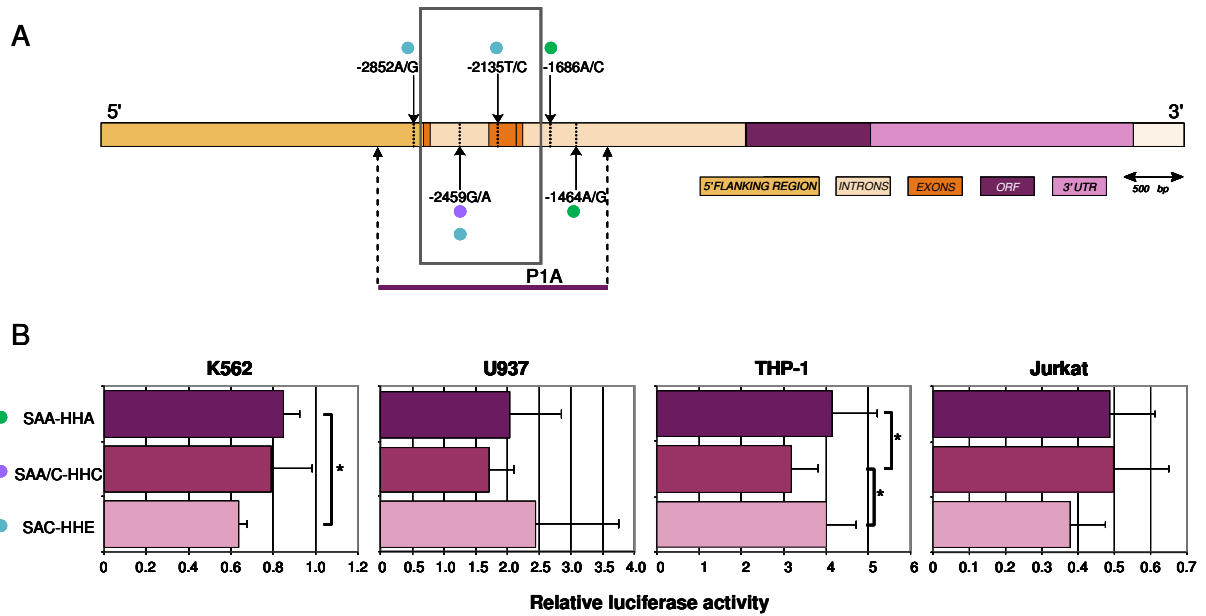
**Figure 3.2.** Promoter expression strength of the different constructs, P1A, P1B, P2 and P1+P2, relative to each other. Constructs were generated from the ancestral haplotype, SAA-HHA, and were tested in four different cell lines, K562, U937, THP-1 and Jurkat. The *CCR5* gene structure and the extent of the promoter constructs are shown. The *CCR5* exons 1, 2A and 2B and open reading frame (ORF) are also illustrated. Results presented are means representative of five separate replicates conducted in K562 and THP-1 cells and six replicates in U937 and Jurkat cells. Data represent the mean and standard deviation for 5 or 6 replicates.

### 3.3.2. Comparison of haplotype-specific promoter efficiencies

We next assessed the effect of different haplotypes on promoter function of the four different promoter regions. We compared the promoter efficiency of constructs generated from three different haplotypes, SAA/C-HHC, SAA-HHA and SAC-HHE (Table 3.1).

#### 3.3.2.1. P1A: *CCR5* downstream promoter region

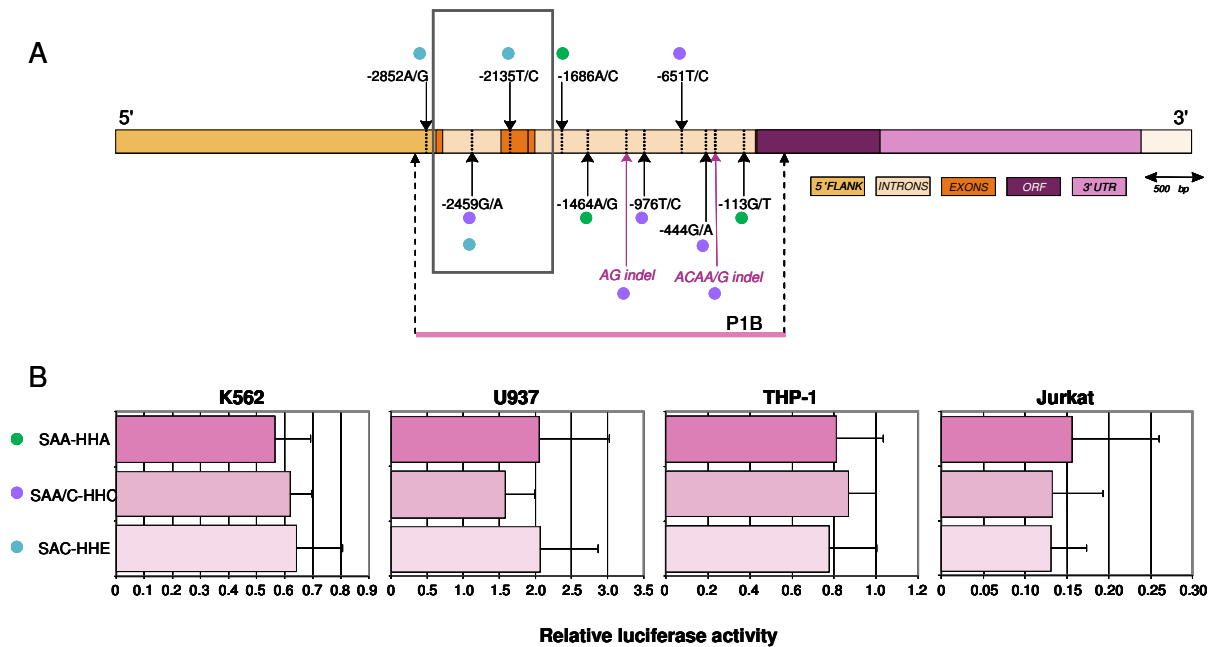
The promoter activity, or efficiencies, of different haplotype constructs relative to one another differed, depending on which cell line the promoter constructs were transfected into. Promoter efficiency for the P1A region is shown in Figure 3.3. In both the U937 and THP-1 cell lines, SAA/C-HHC promoter activity was the weakest of the three haplotypes tested. Promoter efficiency of the SAC-HHE haplotype was significantly higher than that of SAA/C-HHC in THP-1 cell transfections (Figure 3.3). Similarly, the SAA/C-HHC haplotype exhibited significantly higher promoter activity compared to the SAC-HHE haplotype (Figure 3.3). Thus, in THP-1 cells, P1A promoter strength could be ranked as: SAA-HHA>SAC-HHE>SAA/C-HHC. However, in the K562 and Jurkat cell lines, the SAC-HHE haplotype had the weakest promoter activity. This relationship was significant in K562 cells but not in Jurkat cells (Figure 3.3).



**Figure 3.3.** Haplotype differences in promoter efficiency between P1A promoter constructs. **(A)** Schematic representation of the *CCR5* gene and the extent of the P1A region. SNPs forming part of the three haplotypes tested are indicated with colour coded dots. The region bounded by a grey box delineates the region of the *CCR5* gene which, in addition to *CCR5* $\Delta$ 32, is used to describe *CCR5* haplotypes defined by Gonzalez *et al.* (1999). **(B)** Relative luciferase activity of P1A promoter construct derived from SAA-HHA, SAA/C-HHC and SAC-HHE haplotypes, transfected into four different cell lines. Data represent the mean and standard deviation for 5 or 6 replicates. Where significant, differences in promoter efficiency between constructs are indicated with (\*). Note that different scales were used for the different cell lines.

### 3.3.2.2. P1B: *CCR5* downstream promoter region + *CCR5* intron 2

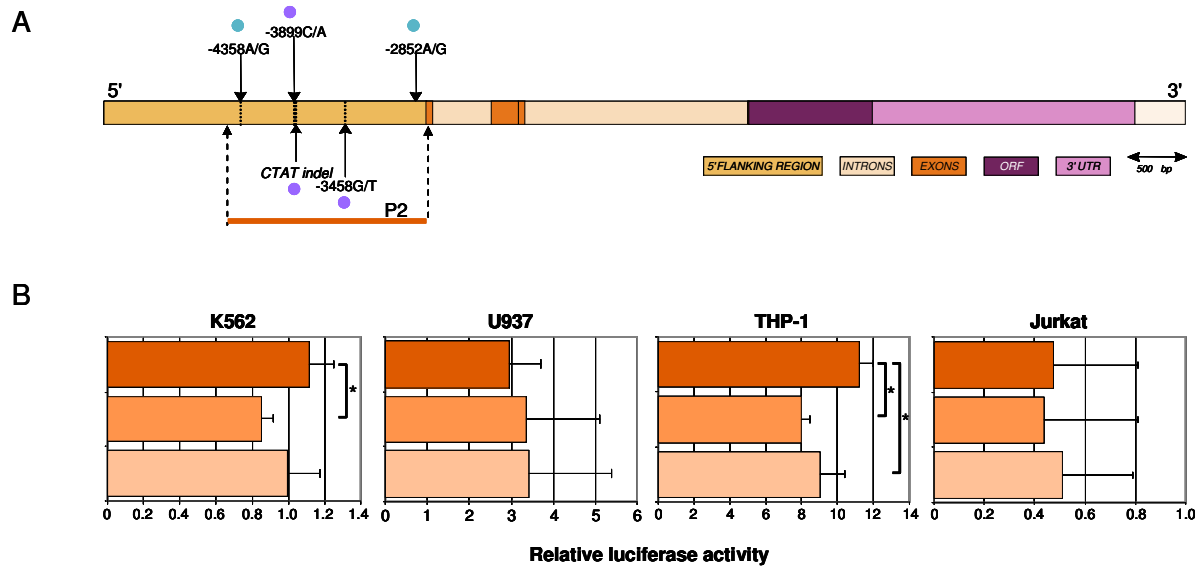
When Intron 2 was included in the promoter construct, i.e., P1B which incorporates both the downstream promoter (P1A) and *CCR5* intron 2, promoter efficiency decreased relative to that observed in the P1A constructs (Figure 3.2). No significant differences between P1B constructs of the different haplotypes were noted (Figure 3.4). Variation in luciferase activity for these constructs was high, with standard deviation from 12.13 to 67.36 % from the mean (Table 3.1).



**Figure 3.4.** Haplotype differences in promoter efficiency between P1B, i.e., *CCR5* downstream promoter and intron2, promoter constructs. **(A)** Schematic representation of the *CCR5* gene and the extent of the P1B region. SNPs forming part of the three haplotypes tested are indicated with colour coded dots. The region bounded by a grey box delineates the region of the *CCR5* gene which, in addition to *CCR5* $\Delta$ 32, is used to describe *CCR5* haplotypes defined by Gonzalez *et al.* (1999). **(B)** Relative luciferase activity of P1B promoter construct derived from SAA-HHA, SAA/C-HHC and SAC-HHE haplotypes, transfected into four different cell lines. Data represent the mean and standard deviation for 5 or 6 replicates. Where significant, differences in promoter efficiency between constructs are indicated with (\*). Note that different scales were used for the different cell lines.

### 3.3.2.3. P2: *CCR5* upstream promoter

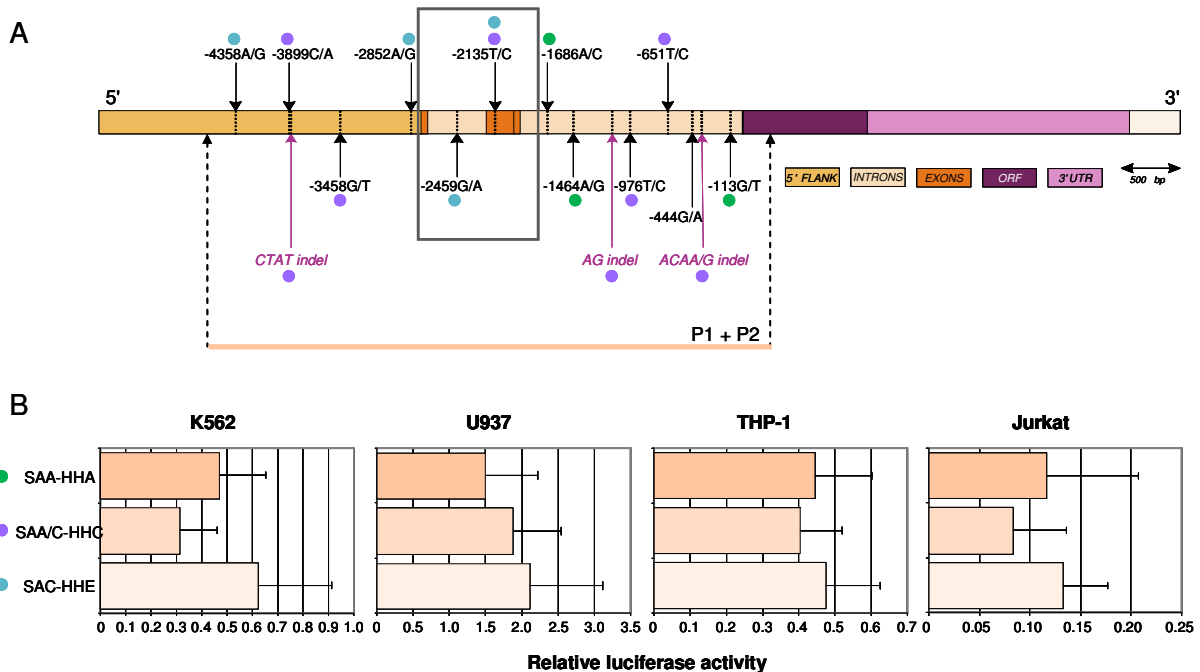
The relative luciferase activity of the different P2 constructs is shown in Figure 3.5. In THP-1 cells, the SAA-HHA P2 promoter efficiency was significantly higher than that of both the SAC-HHE and SAA/C-HHC P2 promoters (Figure 3.5). Thus, the strongest P2 promoter activity was observed for the SAA-HHA haplotype, SAC-HHE had intermediate promoter strength and the weakest promoter activity was observed for the SAA/C-HHC haplotype (Figure 3.5). This was consistent with the relationships observed for P1A constructs in the same cell line (Figure 3.3). The same relationship was observed for P2 constructs in K562 cells, although significant difference was only noted between the SAA-HHA and SAA/C-HHC haplotype within this cell line (Figure 3.5). Thus, P2 promoter strength in THP-1 and K562 cells could be ranked as follows: SAA-HHA>SAC-HHE>SAA/C-HHC. Interestingly this ranking differs slightly than that observed for the P1A construct in K562 cells although SAA-HHA remains the strongest promoter. No significant relationships were observed in either U937 or Jurkat cell line promoter activity.



**Figure 3.5.** Haplotype differences in promoter efficiency between P2, i.e., *CCR5* upstream promoter, promoter constructs. **(A)** Schematic representation of the *CCR5* gene and the extent of the P2 region. SNPs forming part of the three haplotypes tested are indicated with colour coded dots. The region bounded by a grey box delineates the region of the *CCR5* gene which, in addition to *CCR5Δ32*, is used to describe *CCR5* haplotypes defined by Gonzalez *et al.* (1999). **(B)** Relative luciferase activity of P2 promoter constructs derived from SAA-HHA, SAA/C-HHC and SAC-HHE haplotypes, transfected into four different cell lines. Data represent the mean and standard deviation for 5 or 6 replicates. Where significant, differences in promoter efficiency between constructs are indicated with (\*). Note that different scales were used for the different cell lines.

#### 3.3.2.4. P1+P2

Given that polymorphisms within both the upstream and downstream promoter regions appear to have an impact on promoter activities, assessing the synergy between the two promoters becomes important. Thus, the promoter activity of the entire *CCR5* promoter region (downstream and upstream promoter) was also assessed (P1+P2) (Figure 3.6). To our knowledge, only one other study has assessed promoter activity of the combined promoter regions (LIU *et al.* 1998). Addressing the impact that variability across the entire promoter region may have on promoter function was, however, not investigated by Liu *et al.* (1998). We noted no significant differences between constructs of the different haplotypes. Promoter activity for P1+P2 promoters was considerably lower than that of the other promoter regions thus indicating that the effects seen in P1A and P2 separately may be diluted when the *CCR5* promoter region is considered as a whole (P1+P2).

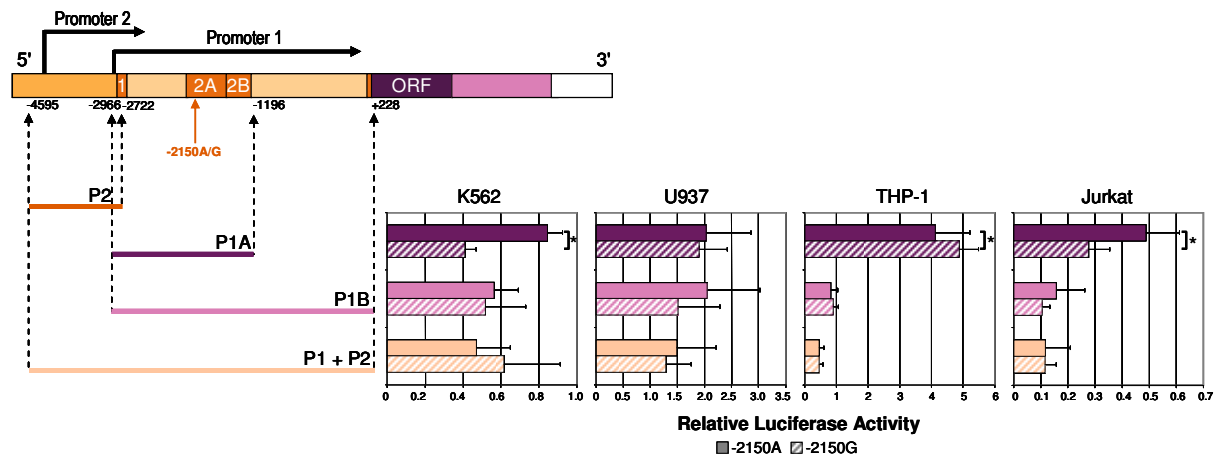


**Figure 3.6.** Haplotype differences in promoter efficiency between promoter constructs spanning the entire *CCR5* promoter region, i.e., P1+P2. **(A)** Schematic representation of the *CCR5* gene and the extent of the P1A region. SNPs forming part of the three haplotypes tested are indicated with colour coded dots. The region bounded by a grey box delineates the region of the *CCR5* gene which, in addition to *CCR5*<sub>Δ32</sub>, is used to describe *CCR5* haplotypes. **(B)** Relative luciferase activity of P1+P2 promoter constructs derived from SAA-HHA, SAA/C-HHC and SAC-HHE haplotypes, transfected into four different cell lines. Data represent the mean and standard deviation for 5 or 6 replicates. Where significant, differences in promoter efficiency between constructs are indicated with (\*). Note that different scales were used for the different cell lines.

### 3.3.3. The newly identified SNP, -2150A/G, negatively influences promoter activity

We previously identified a new *CCR5* SNP, -2150A/G, within *CCR5* exon 2A which forms part of the downstream promoter (P1) (PICKTON *et al.* 2010). This was found to be a rare SNP in SAA individuals (detected as a heterozygous SNP in only one of the 41 SAA individuals genotyped) and absent in SAC individuals (PICKTON *et al.* 2010). The individual from whom the SAA-HHA promoter regions were amplified harboured was homozygous “wild type” across the entire *CCR5* gene sequenced (~9.2 kb) with the exception of this single SNP position. Given its position within the haplotype-defining region as described by Gonzalez *et al.* (1999), we investigated the effect of this polymorphism on promoter activity. Only promoter constructs, P1A, P1B and P1+P2, were tested as the SNP was located within the downstream promoter region.

Promoter activity of the -2150G bearing constructs was remarkably reduced compared to the “wild type” -2150A constructs when transfected into K562, U937 and Jurkat cells (Figure 3.7). This relationship was significant with the P1A constructs transfected into both the K562 and Jurkat cell lines (Figure 3.7). However, this relationship was reversed in THP-1 cells. P1A constructs with the -2150G mutation exhibited significantly higher promoter efficiency compared to the “wild type” (Figure 3.7) thus highlighting the important role of the choice of cells used to assess promoter strengths.



**Figure 3.7.** Effect of the -2150A/G polymorphism on promoter activity of SAA-HHA promoter constructs in four different cell lines tested. The *CCR5* gene structure and the extent of the promoter constructs are indicated. Relative luciferase activity of “wild type” promoter constructs for the different promoter regions is represented by solid bars while that of constructs bearing a G nucleotide at position -2150 is represented by bars with pattern shading. Data represent the mean and standard deviation for 5 or 6 replicates. Significant difference between “wild type” and “mutant” constructs is indicated with (\*). Note the different scales used for the different cell lines.

### 3.3.4. Evaluation of -2150A/G SNP on putative transcription factor binding

To determine whether any of the observed promoter efficiency differences between SAA-HHA promoter transcripts, with or without the -2150A/G mutation, could be attributed to putative transcription binding site differences, we used predictive tools to identify transcription factors binding at that site.

Two transcription factors, FOXC1 and Hoxb-9, were identified as potentially binding to *CCR5* regions including the -2150 site for both “wild type” and “mutant” alleles. However, the SNP position was not located within the core sequence of the transcription factors’ binding sites. A third transcription factor, human MOK-2, with a binding site spanning the -2150 SNP was identified as

putatively binding to the DNA minus strand (Figure 3.8). MOK-2 is expected to only bind the “mutant” allele, with the -2150G nucleotide forming part of the core binding sequence.

```

-2200    CTGTGTAGTG GGATGAGCAG AGAACAAAAA CAAAATAATC
-2160    CAGTGAGAAa ggCCCGTAAA TAAACTTTCA GACCAGAGAT
          -2150                    -2135 -2132
-2120    CTATTCTCTA GCTTATTTTA AGCTCAACTT AAAAGGAAGA
                                         -2086

```

**Figure 3.8.** Putative MOK-2 transcription factor binding site located within the *CCR5* exon 3 region of the SAA-HHA allele containing the -2150G mutation. The transcription factor binding site is illustrated with the coloured box. The -2150G mutation is shown in pink font and the core sequence is in lower case. The nearby SNPs, which among others, are used to define previously described *CCR5* haplotypes (GONZALEZ *et al.* 1999) are indicated in bold font and their positions, relative to the translational start site, are also shown.

The zinc-finger Krüppel-like MOK-2 proteins bind both DNA and RNA through their zinc-finger motifs (ARRANZ *et al.* 1996; ERNOULT-LANGE *et al.* 1995). Furthermore, human MOK-2 acts as a transcriptional repressor (DREUILLET *et al.* 2002). Overlap between the MOK-2 binding site (-2156 to -2136) with those of the other two putative transcription factors, FOXC1 and Hoxb-9, could potentially result in repression of transcription through competitive binding. However, it is important to note that the two latter transcription factors are predicted to bind to the positive DNA strand while MOK-2 is predicted to bind to the negative strand. Whether MOK-2 is a true *CCR5* transcription factor remains to be tested.

### 3.4. DISCUSSION

The *CCR5* haplotypes investigated in this study, SAA-HHA, SAA/C-HHC and SAC-HHE, are extensions of the previously defined haplotypes (PICTON *et al.* 2010, Chapter 2). These haplotypes span the *CCR5* promoter region and thus are likely to influence *CCR5* expression and, as a result, an individuals’ susceptibility to HIV-1 infection and rate of disease progression to AIDS. The root haplotypes, HHA, HHC and HHE, have been associated with, sometimes contrasting, disease modifying effects in ethnically divergent populations. HHC is associated with disease retardation, particularly a delayed progression to death, in HIV-1 infected Caucasians but is associated with disease acceleration in African Americans (GONZALEZ *et al.* 1999). The HHA haplotype is associated with delayed disease progression in African Americans and the HHE haplotype is associated with an increased risk of acquiring HIV-1 in addition to faster disease progression in both Caucasian and African American individuals (MALHOTRA *et al.* 2011; MANGANO *et al.* 2001;

OMETTO *et al.* 2001; TANG *et al.* 2002). This study explored the promoter strength of various *CCR5* promoter regions as well as the influence that SNPs found within three haplotypes, occurring at high frequencies in SAA and SAC populations, have on promoter function. The promoter strengths of the four different promoter regions tested varied considerably between the cell lines tested.

Although *CCR5* is primarily expressed on monocytes/macrophage and T cells, most transformed T or monocytoid cells lines do not express detectable levels of *CCR5* mRNA or protein (LIU *et al.* 1998). Furthermore, cell line subclones may also differ in their ability to express *CCR5*. For example, Liu *et al.* (1998) state that although untransfected Jurkat subclone, Jurkat-LH, expresses detectable *CCR5* mRNA, two other Jurkat derivatives contained no detectable *CCR5* mRNA, however which subclones these are was not specified. Therefore, we tested promoter activity in a range of different cell lines. It is possible that cell lines in which promoter activity was the weakest, i.e., Jurkat E6-1 and K562, may lack the requisite transcription factors for *CCR5* expression. Assessment of endogenous *CCR5* gene expression in different cell lines prior to selecting cells for transfections may prove useful in selecting the optimal representative cell line to test the strengths of gene promoters.

The inclusion of intron 2 in promoter activity reporter plasmids has been shown to reduce promoter activity by as much as 80-90% in Jurkat-LH cells (LIU *et al.* 1998). In agreement with this, we found that constructs with the downstream promoter and intron 2 region (P1B) were weaker than those spanning the downstream promoter region only. This serves to validate the previous study suggesting that *CCR5* intron 2 has a negative regulatory effect on *CCR5* expression (LIU *et al.* 1998).

However, our results differed from published reports both with regards to the relative strengths of the P1 and P2 promoters and the strength of P1 promoter haplotypes in relation to each other (LIU *et al.* 1998; MUMMIDI *et al.* 1997; MUMMIDI *et al.* 2000). Mummidi *et al.* (1997) have demonstrated P1 as the stronger of the two promoters in promoter expression studies conducted in THP-1, Jurkat and K562 cells, whereas our results indicate P2 to be the stronger promoter in the same cell lines in addition to U937. A previous study has demonstrated HHA to be the weakest *CCR5* promoter, through the analysis of downstream promoter (P1) constructs transfected into K562 and Jurkat cells and promoter activity determined in a manner similar to that used in this study (MUMMIDI *et al.* 2000). Of the three constructs tested in our study, HHE was demonstrated as having the strongest promoter activity by Mummidi *et al.* (2000). Our data suggests the opposite. Although relationships were not significant, in the same cells, Jurkat and K562, the ranking of P1A promoter strengths were: SAA-HHA>SAA/C-HHC>SAC-HHE. Although the methodology used in the two studies was similar, they differed in the plasmid vectors selected. Furthermore, the downstream promoter constructs differ in size between the two studies: constructs analysed in the

report by Mummidi *et al.* (2000) were made from the region spanning -2761 to -1814 (927 bp), while we assessed a larger region (-2966 to -1196, 1770 bp). Although the HHA haplotype is “wild type” at all positions in the promoter region assessed by Mummidi *et al.* (2000), the larger region investigated in our study includes two SNPs, -1686A/C and -1464A/G, found in complete linkage disequilibrium with the HHA haplotype (PICTON *et al.* 2010). Similarly, the -2852A/G SNP associated with SAC-HHE was included in our SAC-HHE P1A construct but was not part of the HHE promoter construct investigated by Mummidi *et al.* (2000). It is therefore likely that these additional polymorphisms contribute towards the differing results between the two studies and to test this hypothesis would require the generation and testing of additional constructs of varying length and composition.

An earlier study assessing the influence of two SNPs, -2459G/A and -2135T/C, on promoter function in Jurkat cells, demonstrated that the presence of the -2459G and -2135T (“wild type”) nucleotides resulted in 45% lower promoter activity compared to constructs with the -2459A and -2135C mutations (McDERMOTT *et al.* 1998). These two SNP positions form part of the HHE, HHF and HHG haplotypes, indicating that for the promoter region analysed (-2591 to -1809, 782 bp), these haplotypes are stronger than those which are “wild type” at these positions, i.e., HHA, HHB, HHC and HHD. However, in the HHF haplotypes, these same SNPs are in complete linkage disequilibrium with a further SNP, -1835C/T, in the region analysed. This implies that for the region analysed by McDermott *et al.* (1998), the promoter region of haplotypes HHE and HHG are stronger than the HHA, HHB, HHC and HHD haplotypes. These results are contradictory to our analysis of the downstream promoter region and the two studies differ in the methodology employed to analyse promoter strength in addition to the size of the promoter region analysed, making comparison between the studies difficult. Taken together, the reports by Mummidi *et al.* (2000), McDermott *et al.* (1998) and our results, illustrate that SNPs associated with *CCR5* haplotypes within the downstream promoter region are able to affect *CCR5* expression and hence, probably HIV-1 pathogenesis. However, results appear to differ according to the experimental conditions as well as regions selected to test these promoters. What the true *in vivo* scenario is requires considerably more investigation.

Our findings suggest that the newly identified SNP, -2150A/G, results in reduced promoter activity. This is a rare SNP which was detected in SAA individuals at a haplotype frequency of 1.2% and was found to be absent in SAC individuals (PICTON *et al.* 2010). Although this is not likely to be a major determinant of *CCR5* expression given its low prevalence, the results highlight the importance of this region in transcriptional regulation of *CCR5*. Interestingly, promoter activity of constructs bearing the -2150G nucleotide, i.e., “mutant” alleles, were lower than that of the “wild type” constructs in all cell lines tested with the exception of THP-1, where the opposite effect was

observed. This may be an indication of differences in the transcription factors available in THP-1 and those in the other cells tested, once again highlighting the importance of the cells selected for analysis of promoter function and that CCR5 expression and transcriptional control thereof is likely to differ according to cell type *in vivo*. Predictive tools identified a potential binding site for a transcription factor, MOK-2, spanning the -2150G region. However, this result would have to be validated by means of functional assays.

With the exception of transfections into U937 cells and P1A and P2 transfections into THP-1 cells, the luciferase signal ratio of the control promoterless vector to cotransfected pGL4.73 internal control vector was greater than that of the test promoter to cotransfected pGL4.73 internal control vector (data not shown). This would suggest two things: first, the *CCR5* P1A, P1B, P2 and P1+P2 promoter regions are relatively weak promoters and, secondly, the plasmid into which the promoter constructs were cloned, pGL4.10, is competing with the test promoter regions for transcription factors and thus contributing to a large proportion of background noise in the luciferase assay which would account for the high variability observed between replicates. This could be an indication that another plasmid containing fewer consensus regulatory sequences and hence reduced background and a decreased risk of anomalous transcription (eg. pGL4.16), may be more suitable to studying promoter activity of *CCR5* promoter regions.

Another potential problem with the assay could be the strength of the internal control vector. We used a SV40 promoter containing *Renilla hRluc* as our internal control. The *hRluc*-induced luminescence was substantially higher than that of *luc2* test promoter-induced luminescence. Although this would imply that future assays using this same plasmid as an internal control vector should be assessed using smaller amounts of control plasmid, i.e., higher vector:control vector ratio, in a trial experiment where we assessed the effect of different vector:control vector ratios (10:1, 20:1 and 100:1) on relative luciferase activity values, the ranking of P1A promoter strengths were the same as that obtained for 1:1 ratios, i.e., SAA-HHA>SAA/C-HHC>SAC-HHE (data not shown). The SV40 promoter provides high-level transcription. Thus, in future studies assessing promoter strengths of the weak *CCR5* promoters, an internal control vector which contains a relatively weak promoter, such as the TK (Thymidine kinase) promoter, may prove to provide less competition and hence provide more accurate results, with less variability.

With the advent of new technologies, transfections into PBMCs and isolated cell populations have become easier, which may provide a more accurate reflection of the true influence that SNPs associated with various haplotypes have on gene expression. However, the use of primary cells may also add further confusion due to potential differences between individual donors. Furthermore, promoter expression assays using a reporter gene system, such as *luc2*, provide

very variable results. Thus, in future studies the use of other methods to assess promoter function may provide important validation of results.

Although promoter expression studies conducted in transformed cell lines have shown P1 to be the stronger of the two *CCR5* promoter regions (LIU *et al.* 1998; MUMMIDI *et al.* 1997), the importance of the P2 promoter has been highlighted when isolated PBMCs were transfected with *CCR5* promoter constructs (MUMMIDI *et al.* 2007). T cell activation results in the enrichment of P2-induced *CCR5* mRNA transcripts which have been demonstrated as important determinants of *CCR5* expression on activated primary T cells (MUMMIDI *et al.* 2007). Therefore, SNPs within the *CCR5* P2 region are likely to play an important role determining *CCR5* expression *in vivo* which is in turn an important determinant of HIV-1 pathogenesis. To date, to our knowledge, no other studies have assessed the impact of SNPs within this promoter region on *CCR5* expression. Here, we provide preliminary evidence indicating that differences in the three study haplotypes do influence P2 promoter efficiency.

In conclusion, this study shows that variability within both the upstream and downstream *CCR5* promoter regions influences promoter strength. The inherent limitations of the assay employed indicate that the results of this study should be verified in isolated PBMCs, or even purified cell subsets. This may provide a more accurate representation of the *in vivo* influence of genetic variability of the *CCR5* promoter on promoter strength. Nonetheless, this study has provided new insight into what is known about the *CCR5* promoters on which future studies can be built.

---

## **CHAPTER 4**

### **Marked differences in CCR5 expression and activation levels in two South African populations**

---

## ABSTRACT

The chemokine receptor, CCR5, is pivotal in determining an individual's susceptibility to HIV-1 infection and rate of disease progression. To establish whether population-based differences exist in cell surface expression of CCR5 we evaluated the extent of CCR5 expression across all peripheral blood cell types in individuals from two populations, South African Africans (SAA) and South African Caucasians (SAC). Significant differences in CCR5 expression, both in number of CCR5 molecules per cell (density) and the percentage of CCR5-expressing cells, were observed between the two study groups, within all cell subsets. Most notably, the percentage of all CCR5+ cell subsets were significantly lower in SAC compared to SAA individuals ( $P < 0.01$ ) among natural killer (NK) cell subsets (CD56+, CD16+CD56+ and CD56<sup>dim</sup>) whereas CCR5 density was significantly higher in SAC compared to SAA individuals in CCR5+CD8+ T cell and CCR5+ NK cell subsets (CD56+, CD16+CD56+ and CD56<sup>dim</sup>) (all  $P < 0.05$ ). These relationships were maintained post-exclusion of *CCR5* $\Delta$ 32 heterozygous individuals ( $n=7$ ) from the SAC dataset. SAA individuals exhibited significantly higher cell activation levels, as measured by HLA-DR expression, compared to SAC individuals in CD4+ T cell ( $P=0.002$ ) and CD56+ NK cell subsets ( $P < 0.001$ ). Thus, this study serves to demonstrate that ethnically divergent populations show marked differences in both cell activation and CCR5 expression which are likely to impact on both susceptibility to HIV-1 infection and the rate of HIV-1 disease progression.

## 4.1. INTRODUCTION

Shortly after the discovery of CCR5 as a HIV-1 coreceptor, it was identified as one of the host cell proteins which play an important role in the transmission and pathogenesis of HIV infection. Numerous studies have demonstrated the importance of CCR5 receptor density in the context of HIV-1 infection in that the amount of CCR5 expressed on the cell surface can directly influence an individual's susceptibility to HIV-1 (LEE *et al.* 1999; LIU *et al.* 1996; PLATT *et al.* 1998). *In vitro* studies by Platt *et al.* have demonstrated that a coreceptor density threshold of between  $7 \times 10^2$  and  $2 \times 10^3$  CCR5 molecules/cell is required for efficient replication of R5 HIV-1 in cell lines expressing CD4 levels similar to those on primary T cells (PLATT *et al.* 1998). Furthermore, the density of CCR5 molecules on CD4 T cells correlates positively with the replication of R5 HIV-1 (HEREDIA *et al.* 2007; LIN *et al.* 2002). Increased CCR5 density, determined as the mean number of molecules per cell, in HIV-1 infected individuals correlates with high viral loads (REYNES *et al.* 2000), faster disease progression (GERVAIX *et al.* 2002; REYNES *et al.* 2001), as well as poorer response to antiretroviral treatment (GERVAIX *et al.* 2002; HEREDIA *et al.* 2008; KETAS *et al.* 2007). In addition, CCR5 density is also a determinant of the efficiency of CCR5 in chemotactic response to its ligands (DESMETZ *et al.* 2006).

The best possible example of the importance of CCR5 in HIV-1 infection has been demonstrated by a 32-base pair deletion in the *CCR5* open reading frame (ORF), *CCR5Δ32*, which was first discovered in high risk individuals resisting HIV infection (DEAN *et al.* 1996; LIU *et al.* 1996; SAMSON *et al.* 1996b). Individuals heterozygous for the *CCR5Δ32* allele have a marked reduction in CCR5 surface expression in comparison to individuals lacking this allele (WU *et al.* 1997) and individuals homozygous for this mutant fail to express detectable CCR5 protein on cell surfaces (LIU *et al.* 1996). Population studies of *CCR5Δ32* show that it is present at an average allele frequency of 10% in Europe, however it is very rare or absent in Africans (GALVANI and NOVEMBRE 2005; PETERSEN *et al.* 2001) suggesting that this allele is fairly recent in terms of human evolution (GALVANI and NOVEMBRE 2005). There is also considerable individual to individual variability in surface expression on blood lymphocytes in CCR5 'wild type' individuals, i.e., individuals lacking polymorphisms in the *CCR5* ORF (LEE *et al.* 1999b; MOORE *et al.* 1997; WU *et al.* 1997).

In addition to *CCR5* genetic polymorphisms, CCR5 surface expression can also be influenced by its chemokine ligands. For example, inverse associations between gene copy number of the CCR5 ligand, CCL3L, and CCR5 expression levels have been reported (GONZALEZ *et al.* 2005; KETAS *et al.* 2007). Similarly, CCL5, the CCR5 ligand most abundant in human plasma, regulates CCR5 density by inducing internalization of the receptor (LIN *et al.* 2008). Other molecules, such as interleukin-2, interleukin-12 and interferon- $\alpha$  have also been shown to upregulate CCR5 expression (WEISSMAN *et al.* 2000; YANG *et al.* 2001a; YANG *et al.* 2001b; ZOU *et al.* 1999). An increase in cell activation levels has also been associated with increased CCR5 expression (MEDITZ *et al.* 2011; MORIUCHI *et al.* 1997; OSTROWSKI *et al.* 1998; WU *et al.* 1997). Furthermore, HIV-1 infected individuals have significantly greater percentages of CCR5-expressing CD4+ T cells when compared to healthy controls (OSTROWSKI *et al.* 1998; SHALEKOFF *et al.* 2001). No study to date has considered CCR5 expression across different peripheral blood immune cell subsets between ethnically divergent populations, and earlier studies have largely focused on CD4+ T cells and have assessed CCR5 expression mainly in the context of HIV-1 infection. To gain further insight into the many roles that can be attributed to CCR5 in the immune response and its role as an HIV co-receptor, requires an in-depth look at how this molecule is distributed across immune cell types in the absence of any chronic infections or immune disorders, and how this might vary between individuals and ethnic groups. Given that CCR5 expression plays an important role in HIV-1 infection and rate of disease progression, the overall CCR5 expression profile may predispose to these infection/disease outcomes. We therefore performed a cross sectional study in which we evaluated the expression of the receptor CCR5, as both percentage of CCR5-expressing cells and CCR5 density, on various cell types in whole blood samples taken from healthy, HIV-1 uninfected individuals to evaluate baseline expression in two South African population groups, South African Africans (SAA) and South African Caucasians (SAC).

## 4.2. MATERIALS AND METHODS

### 4.2.1. Study participants

This study cohort comprised of 22 SAA and 31 SAC healthy, HIV-1 uninfected individuals (a subset of the cohort described previously, Chapter 2). The SAA cohort had a median age of 33.5 years (range: 23-62 years) and comprised 14 (63.6%) females and 8 (36.4%) males. The SAC cohort had a median age of 40.0 years (range: 25-67 years) and comprised 20 (64.5%) females and 11 (35.5%) males. There were no statistical differences among the median ages or the male:female ratio of the two groups (Mann-Whitney  $P=0.129$  and Fisher's exact  $P=1$ , respectively). Since the *CCR5Δ32* allele has been shown to impact upon CCR5 expression, participants in this study were genotyped as described previously (PICTON *et al.* 2010). Seven SAC (22.6%) individuals were heterozygous for the *CCR5Δ32* allele. This allele was absent in the SAA population. If *CCR5Δ32* allele-bearing individuals were removed from the analysis, the population group remained age ( $P=0.209$ : SAA median 33.5 years; SAC median 39.5 years) and gender matched ( $P=1$ ; SAA: 14 females and 8 males; SAC: 16 females and 8 males). This study was approved by the University of the Witwatersrand Committee for Research on Human Subjects, and informed written consent was obtained from all participants.

### 4.2.2. Whole blood surface staining and flow cytometry

Ethylenediaminetetraacetic acid (EDTA)-anticoagulated whole blood obtained from each of the study participants was stained within one hour of blood collection. The following antibody (Ab) panels were used for each donor: (i) T cells: CD3-APC, CD4-FITC, CD8-PerCP, CCR5-PE; (ii) B cells: CD19-APC, CCR5-PE; (iii) Natural Killer (NK) cells: CD3-PerCP, CD16-FITC, CD56-APC, CCR5-PE; and (iv) granulocytes and monocytes: CD45-APC, CD14-FITC, CCR5-PE. A HLA-DR marker to assess the extent of cell activation (i.e., HLA-DR) was also included in a subset of the cohort in the following panel: (v) CD3-PerCP, CD8-FITC, CD56-APC, HLA-DR-PE. All antibodies were obtained from BD Pharmigen (BD Biosciences, San Jose, CA). The CCR5 antibody used was conjugated to PE at a ratio of 1:1, thereby allowing for CCR5 quantification, as the mean number of CCR5 molecules per cell (CCR5 density), in addition to the percentage of CCR5-expressing cells within a cell subset. Quantification was carried out using the QuantiBRITE system (BD BioSciences) which is a set of four precalibrated beads to calibrate the fluorescence 2 (FL2) axis in terms of PE molecules.

Fifty microliters of whole blood was used for each antibody panel described above. After incubating stained samples for 15 minutes, the red blood cells were lysed with 2 ml FACS@ lysing solution (BD Biosciences) for 7 minutes. The cells were then pelleted by centrifugation at 100 g for 5 minutes at room temperature and washed with FACS flow (BD Biosciences). The stained cells

were suspended in 150 µl paraformaldehyde (Electron Microscopy Services, Pretoria, South Africa) and stored at 4°C until flow cytometric acquisition within 6 hours of sample collection. All incubations were performed at room temperature in the dark.

Flow cytometric acquisition and analysis was performed on the FACSCalibur (BD Biosciences). The flow cytometer was set up by running FACSComp 5.2 (four-colour lyse-wash) software (BD Biosciences) with CaliBRITE beads (BD Biosciences). Daily compensation using whole blood stained with a single antibody was conducted to optimize instrument settings for the assay. The lymphocyte population was identified on the basis of forward (FSC) and side (SSC) scatter parameters. T cells were defined as CD3 expressing lymphocytes and were further classified as CD4+ or CD8+ T cells. Lymphocytes expressing CD19 were defined as B cells. Natural killer (NK) cells were defined as lymphocytes negative for CD3 and positive for CD56 expression. NK cell subsets analysed were, CD56+, i.e., total CD56-expressing NK cells, CD56<sup>dim</sup>, CD56<sup>bright</sup> and CD16+CD56+. Monocyte and granulocyte populations were identified on the basis of SSC and CD14 parameters. Monocytes were identified by the presence of CD14 markers, whereas granulocytes are negative for CD14. The mean number of CCR5 receptors per cell was determined for CCR5+ cell subsets. Data was analysed using FloJo 7.6.1 (Tree Star, San Carlos, CA).

#### *4.2.3. Statistical analysis*

Mann-Whitney U tests were conducted to compare CCR5 density between individuals, grouped by population or by the presence or absence of the *CCR5Δ32* allele. Correlations between CCR5 density and age of individuals were calculated by bivariate Spearman's rank coefficients. All statistical analyses were performed using PASW Statistics 18 software (SPSS Inc., Chicago, Illinois, USA).

Fisher exact tests were performed using the Simple Interactive Statistical Analysis software (UITENBROOK 1997) to test whether there was any significant difference in composition of population groups.

### **4.3. RESULTS**

#### *4.3.1. CCR5 Expression on Lymphocyte populations*

Isolation of peripheral blood mononuclear cells (PBMCs) by Ficoll purification and a delay in whole blood sample processing is documented to result in acute downregulation of CCR5 expression on various cell types (BERHANU *et al.* 2003; NARANBHAI *et al.* 2011; SHALEKOFF and TIEMESSEN 2001). Therefore, flow cytometry was performed on whole blood samples. Data presented as

mean/medium fluorescence intensity (MFI) or as the proportion of positive cells are always relative to controls that are specific for any given experiment. Thus, we used a method for enumerating CCR5 molecules which allowed better comparison between individuals using a CCR5 antibody conjugated to PE at a ratio of 1:1 in combination with antibodies that define different cell types. This allowed for CCR5 quantification as the mean number of CCR5 molecules per cell, i.e., CCR5 density, in addition to the percentage of CCR5-expressing cells within a cell subset.

In agreement with previous studies, a large interindividual variability on CCR5 expression was observed (LEE *et al.* 1999b; REYNES *et al.* 2000). The mean number of CCR5 molecules per cell differed by as much as 6-fold between individuals, most notably in the CD8<sup>+</sup> T cell and CD56<sup>bright</sup> cell subsets.

The potential influence of gender upon expression was examined. Female participants demonstrated significantly higher percentage CCR5-expressing cells in the CD56<sup>bright</sup> subset than males ( $P=0.042$ ) when the whole cohort was examined. However, this significance was lost when the population groups were examined separately. No differences in expression between male and female participants were observed in all other cell subsets (data not shown).

It is interesting to note that a high percentage of CCR5-expressing cells within a subset does not necessarily correlate with high CCR5 density, i.e., some individuals have a small percentage of CCR5 expressing cells within a cell subpopulation but express CCR5 at high density on this small proportion of cells and vice versa. This is in agreement with reports by others (REYNES *et al.* 2000).

The cell subsets with the highest CCR5 density were T cells and monocytes: T cells (ranging from 1369 to 4820 molecules/cell); CD4<sup>+</sup> T cells (ranging from 945 to 3678 molecules/cell); CD8<sup>+</sup> T cells (ranging from 1055 to 5953 molecules/cell) and CCR5<sup>+</sup> monocytes (ranging from 1628 to 8773 molecules/cell). The subsets with the greatest percentage of cells expressing CCR5 were CD8<sup>+</sup> T cells (CD3<sup>+</sup>CD8<sup>+</sup>, 57.8% median) and CD56<sup>bright</sup> NK cells (64.8% median). Although the CD56<sup>bright</sup> NK cell and B cell populations were observed as high CCR5 expressing subsets, these cell types were present in low numbers which may affect the accuracy of the CCR5 density values.

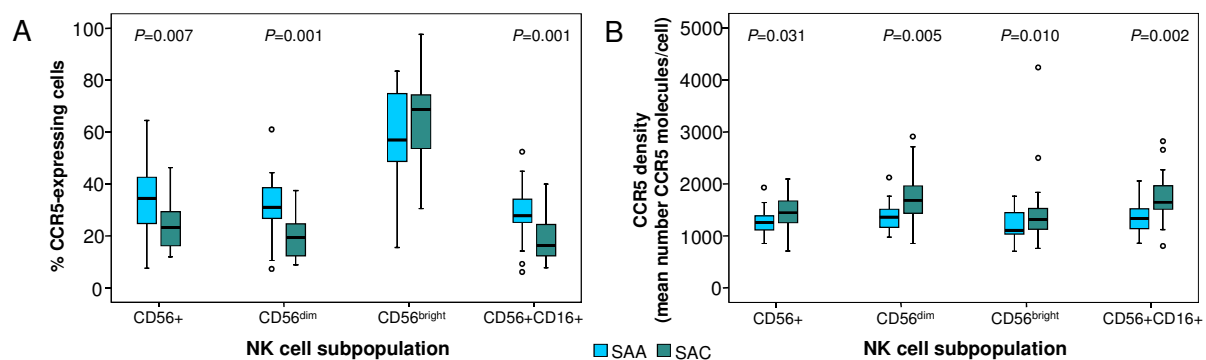
#### *4.3.2. CCR5 expression differs between SAA and SAC individuals within all cell subsets*

Significant differences in CCR5 expression were noted between SAA and SAC individuals, both in terms of CCR5 density as well as the percentage of individual cell types that express CCR5 in the different lymphocyte subpopulations (Table 4.1). When *CCR5* $\Delta$ 32 bearing SAC individuals (n=7) were removed from the analysis, these differences were maintained or even strengthened (Table 4.1).

**Table 4.1.** Differences in CCR5 expression, both as percentage of CCR5-expressing cells and CCR5 density, between two South African population groups, South African African (SAA) (n=22) and South African Caucasian (SAC) (n=31) across all peripheral blood cell populations\*

Cell subset	Median percentage CCR5-expressing cells (%)					Median number of CCR5 molecules per cell						
	Population		P	Population (excluding <i>CCR5Δ32</i> )		P (excluding <i>CCR5Δ32</i> )	Population		P	Population (excluding <i>CCR5Δ32</i> )		P (excluding <i>CCR5Δ32</i> )
	SAA	SAC		SAA	SAC		SAA	SAC		SAA	SAC	
<b>B cells</b>												
CD19+CCR5+	26.0	19.7	<0.001	26.0	20.1	0.001	1761	1492	0.028	1761	1474	0.012
<b>T cells</b>												
CD3+CCR5+	40.8	38.4	0.610	40.8	42.8	0.775	2435	2901	0.066	2435	3045	0.004
CD4+CCR5+	29.0	26.5	0.698	29.0	31.4	0.652	2150	2252	0.626	2150	2357	0.113
CD8+CCR5+	55.8	59.1	0.773	55.8	60.5	0.524	2150	3217	0.008	2150	3502	0.001
<b>Monocytes</b>												
CD14+CCR5+	10.1	12.4	0.093	10.1	12.5	0.050	4802	4318	0.220	4802	4646	0.403
<b>Natural killer cells</b>												
CD56+CCR5+	34.6	22.8	0.001	34.6	24.6	0.007	1265	1450	0.017	1265	1440	0.031

\*Values are indicated for analysis including and excluding individuals heterozygous for the *CCR5Δ32* allele present in the SAC population. P values are indicated. Where significant differences are observed ( $P < 0.05$ ), the highest value in the comparison is highlighted.



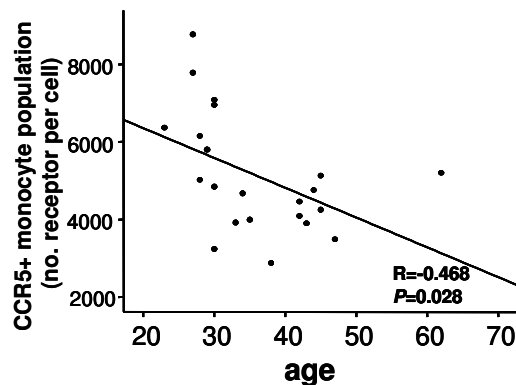
**Figure 4.1.** CCR5 expression in South African African (SAA) (n=22) and South African Caucasian (SAC) (n=24) individuals within natural killer (NK) cell subsets. **(A).** Percentage of CCR5-expressing cells within NK cell subsets. **(B).** CCR5 density on CCR5-expressing NK cell subsets. Heterozygous *CCR5Δ32* individuals have been excluded. Box-whisker plots depicting the median (horizontal black line), 25<sup>th</sup> and 75<sup>th</sup> percentiles (margins of the box) and the 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers). Outliers are indicated with (○). Where significant, P values have been indicated.

Since NK cells showed significant differences in CCR5 expression, we examined expression on various NK cell subsets in detail (Figure 4.1). Overall, SAC individuals tend to express a greater number of CCR5 receptor molecules per cell, compared to SAA individuals, but on a significantly smaller proportion of CCR5 positive cells.

#### 4.3.3. Relationship between CCR5 expression and age of individuals

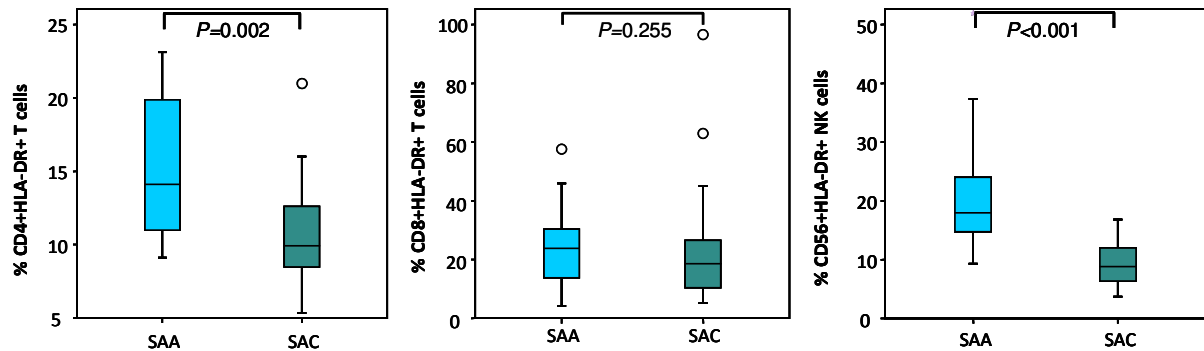
The age range of participants in our study was broad, ranging between 23 and 67 years of age. Thus, we investigated the relationship between CCR5 expression and age on all studied PBMC subsets. A significant negative correlation with age and the percentage of CCR5-expressing cells was seen in the B cell subset in SAA individuals ( $R=-0.478$ ,  $P=0.024$ ) but not in CCR5 'wild type' SAC individuals ( $R=-0.352$ ,  $P=0.092$ ). When grouped together (SAA+SAC), a significant positive correlation was noted on CD8+ T cell subsets ( $R=0.320$ ,  $P=0.020$ ). This relationship was lost when the populations were analysed separately, although a trend was still maintained in the SAC population excluding *CCR5* $\Delta$ 32 allele-bearing individuals ( $R=0.374$ ,  $P=0.072$ ). Significant negative correlations were observed in the SAC population in CD56+ NK cell subsets whether the population was examined as a whole or with the exclusion of *CCR5* $\Delta$ 32 heterozygous individuals ( $R=-0.461$ ,  $P=0.009$  and  $R=-0.448$ ,  $P=0.028$ , respectively). The relationship appears to be due to the CD16+CD56+ NK cell subset ( $R=-0.417$ ,  $P=0.043$  in SAC CCR5 'wild type' individuals), since no significant correlations were observed in the other NK cell subsets. No similar relationships were observed in SAA individuals. Thus, the proportion of CCR5-expressing NK cells appear to decrease with age in SAC individuals.

Although correlations between the proportion of CCR5-expressing cell subsets and age were observed in B, T and NK cell subsets, these relationships were not observed when correlations with CCR5 density and age were analysed in the same cell subsets (data not shown). In monocyte cells, however, a negative correlation between CCR5 density and age only in SAA individuals was observed ( $R=-0.468$ ,  $P=0.028$ ) (Figure 4.2).



**Figure 4.2.** Correlation between CCR5 density (mean number of molecules per cell) on monocyte cells and age of South African African (SAA) study participants (n=22). *P* values are indicated.

#### 4.3.4. SAA and SAC individuals differ in cell activation levels



**Figure 4.3.** Activation levels, as demonstrated by percentage HLA-DR expression in the two study population groups, South African African (SAA) (n=16) and South African Caucasian (SAC) (n=23) within natural killer (NK) and T cell subsets. Box-whisker plots depicting the median (horizontal black line), 25<sup>th</sup> and 75<sup>th</sup> percentiles (margins of the box) and the 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers). Outliers are indicated with (○). *P* values have been indicated.

The levels of activation for a subset of individuals from both population groups, SAA (n=16) and SAC (n=24), was measured for T cells and NK cells using HLA-DR as a marker of activation. SAA individuals exhibited significantly higher expression of HLA-DR, measured as the percentage of HLA-DR-expressing cells, in CD4<sup>+</sup> T cell ( $P=0.002$ ) and NK cell (CD56<sup>+</sup>,  $P<0.001$ ) subsets (Figure 4.3). When we investigated the correlations between HLA-DR and CCR5 expression levels, a positive correlation was observed in SAA individuals, between HLA-DR expression (%) and the percentage of CCR5-expressing cells on T cell (CD3<sup>+</sup>) and CD4<sup>+</sup> T cell subsets ( $P=0.021$  and  $0.037$ , respectively, Table 4.2). However, the proportion of HLA-DR-expressing cells did not correlate with CCR5 density on these same subsets (Table 4.2). SAC individuals, on the other hand, demonstrated a positive relationship between proportions of HLA-DR-expressing cells and CCR5 density, i.e., individuals with higher proportions of T cell subsets expressing HLA-DR, had a higher number of mean CCR5 molecules per cell on CD3<sup>+</sup>, CD4<sup>+</sup> and CD8<sup>+</sup> T cell subsets (Table 4.2) but no significant correlation was observed when analyzing the relationship between the percentage of CCR5-expressing and of HLA-DR-expressing subsets. No significant correlations were observed in any NK cell subsets (data not shown).

To investigate possible relationships between CCR5 density and HLA-DR density on the same cell subsets, we used the geometric mean of HLA-DR<sup>+</sup> subsets as a substitute measure for the mean number of molecules per cell as we did not use an HLA-DR antibody conjugated to PE at a 1:1 ratio. No significant differences in HLA-DR mean fluorescence intensity (MFI) between the two populations were observed (data not shown). Furthermore, no correlations between HLA-DR MFI and CCR5 density were observed.

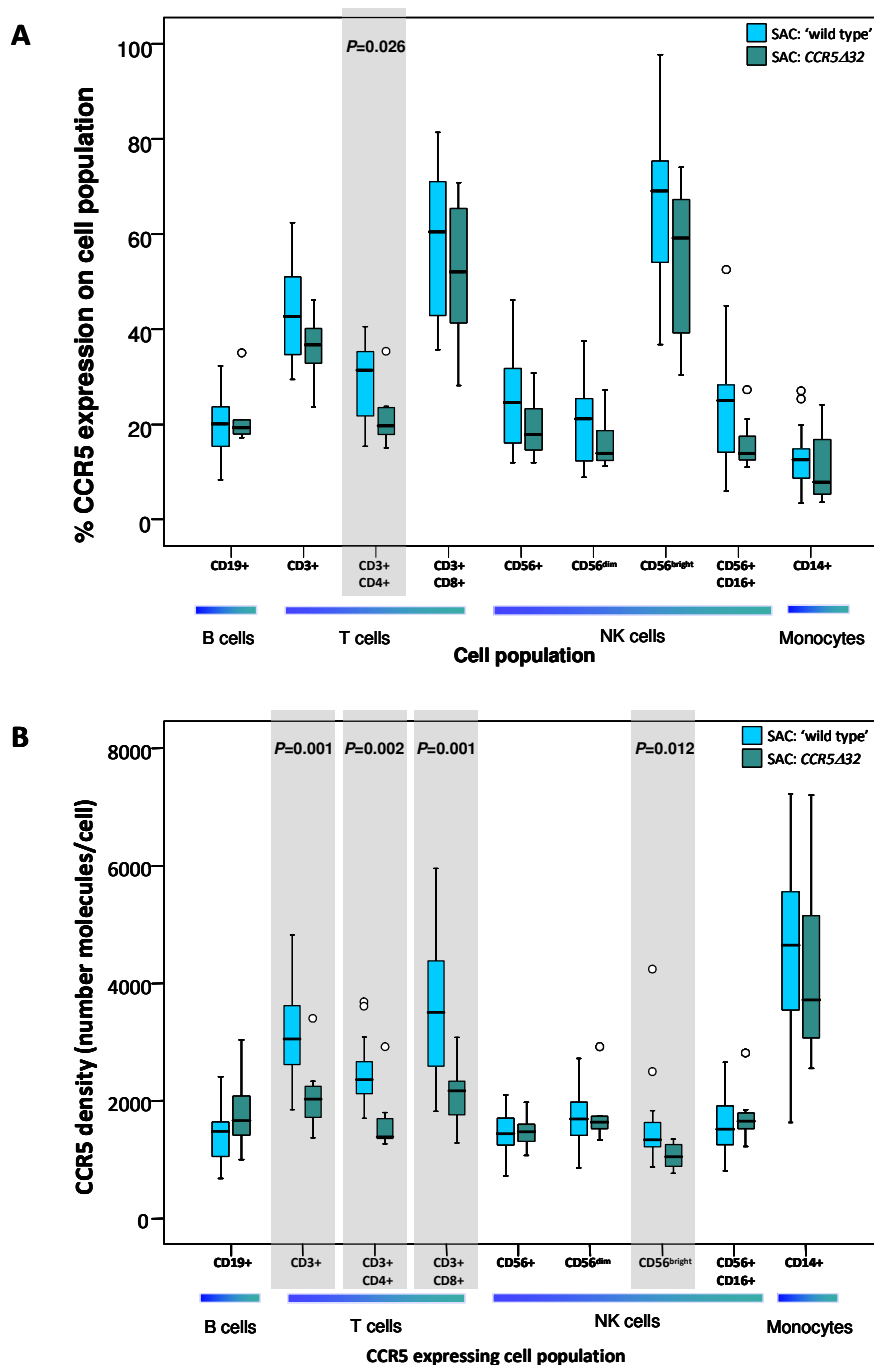
**Table 4.2.** Correlation between cell activation, as measured by HLA-DR percentage expression, and CCR5 expression (percentage CCR5-expressing cells and density) on T cell subsets\*

Cell subset	% HLA-DR+CD4+ T cells*			% HLA-DR+ CD8+ T cells		
	SAA + SAC (n=39)	SAA (n=16)	SAC (n=23)	SAA + SAC (n=39)	SAA (n=16)	SAC (n=23)
% CCR5-expressing cells						
CD3+	R=0.367 <i>P</i> =0.022	R=0.569 <i>P</i> =0.021	R=0.254 <i>P</i> =0.242	R=0.159 <i>P</i> =0.334	R=0.141 <i>P</i> =0.602	R=0.172 <i>P</i> =0.434
CD3+CD4+	R=0.308 <i>P</i> =0.056	R=0.525 <i>P</i> =0.037	R=0.071 <i>P</i> =0.746	R=0.273 <i>P</i> =0.093	R=0.200 <i>P</i> =0.457	R=0.315 <i>P</i> =0.144
CD3+CD8+	R=0.407 <i>P</i> =0.010	R=0.366 <i>P</i> =0.164	R=0.569 <i>P</i> =0.005	R=0.172 <i>P</i> =0.294	R=0.349 <i>P</i> =0.185	R=0.106 <i>P</i> =0.630
Number of CCR5 molecules/cell						
CD3+	R=-0.058 <i>P</i> =0.725	R=0.242 <i>P</i> =0.367	R=-0.020 <i>P</i> =0.929	R=0.334 <i>P</i> =0.038	R=-0.058 <i>P</i> =0.832	R=0.526 <i>P</i> =0.010
CD3+CD4+	R=-0.033 <i>P</i> =0.840	R=0.164 <i>P</i> =0.544	R=-0.051 <i>P</i> =0.817	R=0.286 <i>P</i> =0.077	R=-0.077 <i>P</i> =0.777	R=0.444 <i>P</i> =0.034
CD3+CD8+	R=-0.145 <i>P</i> =0.377	R=0.320 <i>P</i> =0.227	R=-0.044 <i>P</i> =0.843	R=0.317 <i>P</i> =0.050	R=-0.005 <i>P</i> =0.984	R=0.484 <i>P</i> =0.019

\* Where significant, *P* values have been highlighted.

#### 4.3.5. Influence of *CCR5*Δ32 allele on *CCR5* expression within SAC population

Individuals bearing the *CCR5*Δ32 allele tended to have a lower percentage of CCR5-expressing cell subsets in all cell populations examined, yet statistical significance was only observed in CD4 T cells (*P*=0.026) (Figure 4.4A). Individuals heterozygous for the *CCR5*Δ32 allele demonstrated lower CCR5 density on T cell populations relative to individuals that lack the allele (*P*<0.005) (Figure 4.4B). This is consistent with other reports (Wu *et al.* 1997). However, it is interesting to note that differences in expression in other cell populations are not significant with the exception of CD56<sup>bright</sup> NK cells (Figure 4.4B). In addition, some SAC individuals heterozygous for the *CCR5*Δ32 allele had comparable CCR5 expression, both in the percentage of CCR5-expressing cells and CCR5 density, compared to homozygous 'wild type' individuals (Figure 4.4A and B).



**Figure 4.4.** Influence of *CCR5Δ32* bearing alleles on CCR5 expression in the South African Caucasian (SAC) population within all cell populations examined. **(A)** Percentage of CCR5-expressing cells in the different cell subsets examined in SAC individuals heterozygous for the *CCR5Δ32* allele ( $n=7$ ) and CCR5 'wild type' SAC individuals ( $n=24$ ). **(B)** CCR5 density (mean number of molecules/cell) on the different cell subsets examined in SAC individuals heterozygous for the *CCR5Δ32* allele ( $n=7$ ) and CCR5 'wild type' SAC individuals ( $n=24$ ). Where significant,  $P$  values have been indicated. Box-whisker plots depicting the median (horizontal black line), 25<sup>th</sup> and 75<sup>th</sup> percentiles (margins of the box) and the 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers). Outliers are indicated with (○).

#### 4.4. DISCUSSION

Extensive variation exists between individuals, in their susceptibility to HIV-1 and the rate of disease progression to AIDS. Cell surface expression of the HIV-1 coreceptor, CCR5, is highly variable, even in individuals homozygous for the wild type ORF region (LEE *et al.* 1999b; MOORE *et al.* 1997; WU *et al.* 1997). Although the influence of the *CCR5* $\Delta$ 32 mutation on CCR5 expression on T cells has been studied extensively in the context of HIV-1 infection, we lack information on how healthy HIV-1 uninfected individuals may differ in terms of individual CCR5 expression levels. Given that the extent of expression of CCR5 plays a pivotal role in determining HIV-1 susceptibility and that striking ethnic or population differences in SNP frequencies of CCR5 exist, it is surprising that very few studies have questioned how populations may differ in terms of their CCR5 expression levels.

Transsexual men receiving oestrogens and antiandrogens have been reported to have increased CCR5 expression on T cells (GILTAY *et al.* 2000) and female mice have higher CCR5 expression on CD4<sup>+</sup> T cells compared to their male counterparts (MO *et al.* 2005). Since oestrogen thus appears to influence CCR5 expression, we looked at potential differences in expression between genders in the current study. On the whole, no differences in expression were observed between male and female study participants, whether examined collectively or separately in the two population groups.

Previous studies have demonstrated CCR5 receptor expression to be stable within individuals over multiple time points despite the wide range of variability that exists between individuals (KIVISAKK *et al.* 2003; OSTROWSKI *et al.* 1998; REYNES *et al.* 2000). However, aging has been associated with increased CCR5 expression on T cells both in mice through gene expression studies (MO *et al.* 2003) and in humans using ribonuclease protection assays and Western blots (YUNG and MO 2003). Thus, we investigated the relationship between CCR5 expression and age on all studied cell subsets in the two population groups (age and sex matched). The proportion of CCR5-expressing CD8<sup>+</sup> T cells was observed to significantly increase with age in the pooled cohort but no correlations were observed in T cells when the two populations were analyzed separately. In a human study in which CCR5 expression on CD4 T cells was shown to be higher in older individuals, CCR5 expression was compared between two groups, young adults (aged 18 to 40 years) and older subjects ( $\geq$ 60 years, mean age 73 years) (YUNG and MO 2003). Only three of our participants were aged  $\geq$ 60 years (1 SAA individual, aged 62, and 2 SAC individuals, aged 64 and 67), thus it is likely that our study did not have sufficient 'older' individuals to show this relationship. Also, levels of expression were determined differently, where we have looked at cell surface expression by means of flow cytometry, Yung and Mo (2003) used ribonuclease protection assays and Western blots. However, interestingly, in SAA individuals, but not SAC individuals, we

observed a decrease in CCR5 density on CCR5+ monocytes in older individuals. Further studies with a larger cohort would be necessary to unequivocally establish the effects of age on CCR5 expression in healthy adults.

Although only a few studies have compared CCR5 expression in different population groups, differences have been noted. For example, Ethiopians living in Israel have been shown to express CCR5 at higher levels, both in terms of proportions of CCR5-expressing CD4+ cells and the number of molecules per CD4+ cell, than Israeli Caucasians, despite comparable activation levels as measured by HLA-DR expression (KALINKOVICH *et al.* 2001). However, the prevalence, and hence potential influence upon expression, of the *CCR5Δ32* allele in this study was not indicated. In the present study, we show significant expression differences, both in terms of percentage of CCR5-expressing cells and the CCR5 density, between our two study populations, in all cell subsets. SAA individuals tended to express fewer CCR5 molecules per cell but on a larger proportion of cells, compared to SAC individuals. The most significant differences in CCR5 expression between the two study groups were seen in the NK cell subsets. The role of NK cells in HIV-1 infection is complex and is not fully understood. While NK cells are able to directly mediate the killing of HIV-1 infected cells, they are also thought to be involved in the killing of uninfected CD4+ T cells and thus contribute towards CD4+ T cell decline (Reviewed by FUNKE *et al.* 2011). Cytotoxic NK cell subsets, expressing CD4, CCR5 and/or CXCR4 have been identified (BERNSTEIN *et al.* 2009; HARADA *et al.* 2007; VALENTIN *et al.* 2002). Furthermore, *in vitro* studies have shown HIV-1 to infect stimulated CD4+ NK cells expressing HIV-1 coreceptors (BERNSTEIN *et al.* 2009; HARADA *et al.* 2007; VALENTIN *et al.* 2002), an indication that these cells could serve as HIV-1 reservoirs. Thus, it would follow that, similar to T cell subsets, both the proportion of CCR5-expressing and CCR5 density of NK cells could impact on the progression of HIV-1 infection.

Immune activation levels have previously been reported as environmentally driven (CLERICI *et al.* 2000). Individuals living in Africa, of Italian (Caucasian) and Ugandan origin, were demonstrated to have higher activation levels when compared to individuals from the same population groups living in Italy (CLERICI *et al.* 2000). Interestingly, comparable activation levels, as measured by HLA-DR expression, were seen in both the Italian and Ugandan group living in Uganda (CLERICI *et al.* 2000), contrasting the results in our study where SAA and SAC individuals, largely controlled for environment, differ significantly in HLA-DR expression, i.e., activation levels, in CD4+ T cell and NK cell subsets. These differences could be indicative of either the two South African population groups being exposed to differing environmental/activation triggers or that corresponding geographically distinct population groups (SAA versus Ugandans and SAC versus Italians) differ with respect to their genetic background. Similarly, in a study whose participants were of Ethiopian and Israeli origin, Ethiopian individuals that had recently moved to Israel had higher CD4+HLA-

DR+ and CD8+CD38+ levels compared to Ethiopian individuals who had been living in Israel for >7 years and Caucasian Israelis (KALINKOVICH *et al.* 2001). Ethiopians who had been living in Israel for >7 years still had higher activation levels than their Caucasian Israeli counterparts, implying that there must be factors other than environmental, i.e., genetic, to which the higher activation levels can be attributed (KALINKOVICH *et al.* 2001). Furthermore, Kenyan women negative for HIV-1 and other sexually transmitted infections, have reportedly increased activated mucosal T cells (CD4+CD69+ T, CD4+CD69+CCR5+ T and CD8+CD69+ T cell subsets) compared to women of mixed ethnicity in the United States (COHEN *et al.* 2010). It was postulated that these differences could be attributed to differences in host genetics or from immune activation caused by chronic systemic infections which may be more prevalent in Africa (COHEN *et al.* 2010). Given that the participants in our study were all living in the same environment and had no overt infections or known immune conditions, we postulate that differences in activation levels are more likely to be driven by genetic factors. These studies collectively highlight the complexity of understanding what influences the level of immune activation measurable in a healthy individual and point toward an interplay between the genetics and environment of the individual.

Cell activation results in an increase in CCR5 expression (MEDITZ *et al.* 2011; MORIUCHI *et al.* 1997; OSTROWSKI *et al.* 1998; WU *et al.* 1997). Activated (HLA-DR+) CD4+ cells have been shown to express CCR5 at higher levels compared to nonactivated (HLA-DR-) CD4+ cells (BEGAUD *et al.* 2006; KALINKOVICH *et al.* 2001; REYNES *et al.* 2000; WU *et al.* 1997). Our results in healthy SAC individuals are in keeping with this; higher proportions of HLA-DR-expressing CD4+ and CD8+ T cells correlated with correspondingly higher density of CCR5 expression. Furthermore, the low expression of activation markers HLA-DR and CD38 on CD4+ T cells is linked to low HIV-1 susceptibility in HIV-1 exposed uninfected individuals (BEGAUD *et al.* 2006; KONING *et al.* 2005). It is interesting to note that in a report by Reynes *et al.* (2000), the observation is made that although activation, measured by HLA-DR expression on CD4+ T cells, leads to a significant increase in both the number of CCR5 molecules and the proportion of CCR5-expressing cells, the difference in percentage CCR5-expressing cells between HLA-DR- and HLA-DR+ cell subsets is much greater than differences seen in CCR5 density (REYNES *et al.* 2000). A recent report has demonstrated *in vivo* that activated (HLA-DR+CD38+) T cells in lymphoid tissues are highly susceptible to infection by R5 tropic HIV-1 viruses and postulated that this is likely due to the observed higher expression of CCR5 on these cells in comparison to other CD4+ T cell subsets (MEDITZ *et al.* 2011).

Consistent with the significantly higher activation levels observed in T and NK cells, SAA individuals had greater proportions of CCR5-expressing subsets compared to SAC individuals. However, an inverse relationship was observed in terms of CCR5 density. Several studies have shown that CCR5 density is capable of influencing HIV-1 susceptibility, with higher density

associated with increased susceptibility to infection (LEE *et al.* 1999b; LIU *et al.* 1996; PLATT *et al.* 1998). Furthermore, it has been reported that CCR5 density on CD4+ T cells of slow progressors versus normal progressors, but not the percentage of CCR5-expressing cells, correlate positively with viral loads (REYNES *et al.* 2001). Interestingly, CCR5 density was also lower in the slow progressors compared to the uninfected controls (REYNES *et al.* 2001). However, Meditz *et al.* (2011) have recently demonstrated that both CCR5 density and the percentage of CCR5-expressing cells are important determinants of HIV-1 susceptibility *in vivo*. Thus, both density and the percentage of CCR5-expressing cells appear to be important determinants, coupled with the extent of immune activation, and it is the combinatory effects of these that may predispose to different infection and disease outcome phenotypes. It can be envisaged that larger proportions of CCR5-expressing cells in the periphery would result in larger numbers of cells entering an immune response at the site of infectious encounter; if these are CD4+ cells then more target cells would be available for HIV-1 infection, if among these there are CCR5-expressing NK cells, for example, then those co-expressing CD4 may be vulnerable to infection but others may be more effective killer cells – providing a counter response to control the extent of infection. The overall outcome would be dependent on many factors, including also the concentration of the natural chemokines that bind CCR5, which aside from their role as chemoattractants would also prevent HIV-1 entry by blocking or down-regulating CCR5 expression.

Although increased cellular activation in CD4+ T cells subsets is associated with increased susceptibility to HIV-1 infection, it is not clear what the functional significance of HLA-DR expression on NK cells may be. The expression of class II major histocompatibility complex (MHC) proteins such as HLA-DR, is usually considered to be restricted to professional antigen presenting cells (Reviewed by CHAPLIN 2010). However, HLA-DR expression on both NK and T cell subsets has been documented and in both these subsets, HLA-DR expression is often used as a biomarker of activation (EVANS *et al.* 2011; EVANS *et al.* 1978; KO *et al.* 1979). HLA-expressing NK cells have been shown to present antigen to trigger T cell proliferation and IL-2 production (HANNA *et al.* 2004; RONCAROLO *et al.* 1991). In addition, HLA-DR-expressing NK cells have significantly stronger cytolytic activity in comparison to other NK cells (EVANS *et al.* 2011). Evans *et al.* (2011) demonstrate that the magnitude of an individual's NK cell IFN- $\gamma$  response triggered by the model pathogen, *Mycobacterium bovis* BCG, was associated with the initial proportion of HLA-DR-expressing NK cells in PBMCs (EVANS *et al.* 2011). Thus, HLA-DR-expressing NK cells appear to play an important role in differences in individual responses to BCG and potentially other pathogens. Although SAA individuals appear to have a larger proportion of activated CD4+ T cells in comparison to SAC individuals, and therefore potentially increased risk of acquiring HIV-1 if exposed, we hypothesise that the concomitant higher NK cell activation may counteract this

potential risk either entirely or to some extent through enhanced killing of activated/infected CD4+ T cells.

Across all subsets examined, the influence of *CCR5Δ32* on CCR5 expression was most notable in the T cell subsets. We observed individuals heterozygous for the *CCR5Δ32* allele to have significantly lower percentages of CCR5-expressing CD4+ T cells in comparison to CCR5 'wild type' individuals, an observation which is in agreement with previous studies (DE RODA HUSMAN *et al.* 1999; SHALEKOFF and TIEMESSEN 2003; WU *et al.* 1997). SAC individuals bearing this allele had significantly reduced receptor density on T cell (CD4+ and CD8+ subsets) and CD56bright NK subsets. Interestingly, the statistical relationships between CCR5 expression of *CCR5Δ32* and 'wild type' individuals were stronger when analyzing CCR5 density than when analyzing the percentage of CCR5-expressing cells. Similar results have been reported by Shalekoff and Tiemessen (SHALEKOFF and TIEMESSEN 2003). Although *CCR5Δ32* heterozygosity does have an impact on CCR5 expression, this was not observed across all subsets and furthermore, in agreement with previous studies, some 'wild type' individuals lacking *CCR5Δ32* had CCR5 densities comparable to individuals bearing this allele (DE RODA HUSMAN *et al.* 1999; SHALEKOFF and TIEMESSEN 2003; WU *et al.* 1997). Thus, although a significant factor, heterozygosity for the *CCR5Δ32* allele is not the only determinant of CCR5 cell surface density.

Although several studies have described differences in CCR5 expression levels between individuals, these have been mainly restricted to a single lymphocyte population (eg. CD4+ T cells), and have mainly addressed only the role of CCR5 as co-receptor for entry of HIV-1 into CD4 T cells. Importantly, we should not overlook that CCR5 also plays a role in T-cell immunity (CAMARGO *et al.* 2009; DOLAN *et al.* 2007). In this study we analyzed CCR5 expression on T cells, NK cells, B cells and monocytes. Although HIV-1 mainly infects CD4+ T lymphocytes, it is known to have a broad cellular host range infecting a number of different cell populations of the brain, bowel, heart, kidney, liver in addition to haematopoietic cells (LEVY 2006). Aside from the ability of HIV-1 to infect cell types other than CD4+ T cells, the role of differential CCR5 expression on other cell types and their role in immune response on initial encounter with HIV-1, or in existing chronic infection, need to be better understood to establish the overall impact of differing degrees of CCR5 expression *in vivo*.

We found CCR5 expression to differ substantially between SAA and SAC individuals and that these differences can be attributed to factors other than the *CCR5Δ32* allele. In addition, these two groups differ substantially in their activation levels as measured by HLA-DR. Further understanding of what contributes towards the large interindividual and interpopulation variability in CCR5 expression, and understanding the implications thereof, will assist in the design of prophylactic and therapeutic strategies that work in conjunction with the host's immune response to HIV-1.

---

## **CHAPTER 5**

***CCR5* promoter haplotypes differentially influence *CCR5* expression on Natural Killer and T cell subsets in ethnically divergent HIV-1 uninfected South African populations**

---

## ABSTRACT

CCR5 plays a critical and central role in HIV-1 infection and, to date, a number of genetic mutations and haplotypes within the gene have been found to positively or negatively influence an individual's susceptibility and rate of disease progression. In this study, we have evaluated the influence of *CCR5* haplotypes, HHA, HHC, HHD, and HHE, on CCR5 expression in healthy HIV-1 uninfected individuals from two populations, South African Africans (SAA, n=22) and South African Caucasians (SAC, n=31). *CCR5* haplotypes were determined through sequencing and real time polymerase chain reaction. Flow cytometry was used to quantitate CCR5 surface expression, as both CCR5 density and percentage of CCR5-expressing cells, on B, T, natural killer (NK) cells and monocytes. SAA individuals positive for the HHA haplotype had significantly lower percentages of CCR5-expressing CD8+ T cells in comparison to individuals without HHA ( $P=0.001$ ). HHC+ SAC individuals had significantly higher CCR5 molecules per cell (density) on NK (CD56+) and CD16+CD56+ NK cell subsets ( $P=0.030$  and  $P=0.024$ , respectively) compared to HHC- SAC individuals. Haplotypes HHD and HHE had no impact on CCR5 expression. Overall, our data highlight that the protective effect of the HHC haplotype in Caucasians might be explained by higher density of CCR5 expression on NK cells that is not evident in HHC+ SAA individuals. Findings raise the question as to the role of CCR5-expressing cells other than CD4+ T cells in protection from HIV-1 acquisition and disease progression.

## 5.1. INTRODUCTION

CCR5 is the major coreceptor for infection used by macrophage-tropic (M-tropic) strains of HIV-1 (DENG *et al.* 1996; DRAGIC *et al.* 1996) and thus plays a crucial role in the infection and pathogenesis of these viruses. The discovery that individuals homozygous for the *CCR5* $\Delta$ 32 allele, characterized by a 32 base pair deletion in the *CCR5* open reading frame (ORF), are highly resistant to HIV-1 infection (DEAN *et al.* 1996; LIU *et al.* 1996; SAMSON *et al.* 1996b) was a landmark finding demonstrating how host genetic factors can influence both susceptibility to HIV-1 infection and the rate of disease progression to AIDS. The importance of CCR5 receptor density in the context of HIV-1 infection has been highlighted in numerous studies which have shown that the amount of CCR5 expressed on the cell surface can directly influence an individual's susceptibility to HIV-1 (LEE *et al.* 1999; LIU *et al.* 1996; PLATT *et al.* 1998). CCR5 density is also a determinant of the efficiency of CCR5 in response to its chemokine ligands and therefore is an important determinant of capacity for cell migration in response to inflammatory signals (DESMETZ *et al.* 2006).

To date, numerous other *CCR5* ORF mutations, as well as single nucleotide polymorphisms (SNPs) and/or haplotypes in the *CCR5* regulatory or promoter region, capable of influencing HIV-1

susceptibility and the rate of disease progression have been reported by different groups (GONZALEZ *et al.* 1999; HOWARD *et al.* 1999; KOSTRIKIS *et al.* 1998; MANGANO *et al.* 2001; MARTIN *et al.* 1998; MCDERMOTT *et al.* 1998; QUILLEN *et al.* 1998). Striking ethnic or population differences in SNP frequencies of *CCR5* exist (ANSARI-LARI *et al.* 1997; CARRINGTON *et al.* 1997; GALVANI and NOVEMBRE 2005; LIU *et al.* 2012; MAGIEROWSKA *et al.* 1999a; MASQUELIER *et al.* 2007; PETERSEN *et al.* 2001; PICTON *et al.* 2010; SHIODA *et al.* 2001). Nine evolutionarily distinct *CCR5* haplotypes, comprising seven SNP positions in the *CCR5* regulatory region, the *CCR5* $\Delta$ 32 ORF mutation, in addition to a *CCR2* SNP, *CCR2V64I*, have been described (GONZALEZ *et al.* 1999; MUMMIDI *et al.* 2000). These haplotypes however appear to have differing disease-modifying effects in the different population groups studied, for example, in a study conducted by Gonzalez *et al.* (1999), the HHC haplotype was shown to be associated with disease retardation in the Caucasian and Hispanic populations, whereas it was associated with disease acceleration in African Americans (GONZALEZ *et al.* 1999).

The presence and severity of HIV-1 infection influences the extent of expression of *CCR5* (OSTROWSKI *et al.* 1998; SHALEKOFF *et al.* 2001), and is therefore a confounder in establishing the influence of host genes directly on *CCR5* surface expression. To our knowledge, the influence of *CCR5* haplotypes other than *CCR5* $\Delta$ 32 on *CCR5* expression in healthy HIV-1 uninfected individuals has not been investigated. Thus, we performed a cross sectional study to evaluate the influence of *CCR5* genotypes prevalent in two South African population groups, South African Africans (SAA) and South African Caucasians (SAC), on *CCR5* expression. Contrary to expectations that particular *CCR5* haplotypes (other than *CCR5* $\Delta$ 32) that reportedly influence protection/susceptibility might affect proportions of *CCR5*-expressing CD4<sup>+</sup> T cells, or densities of *CCR5* on CD4<sup>+</sup> T cells, altered *CCR5* expression on NK cells in particular was found to be a contender.

## 5.2. MATERIALS AND METHODS

### 5.2.1. Cohort

This study cohort comprised of 22 SAA and 31 SAC healthy, HIV-1 uninfected individuals recruited from within our institute as described previously (PICTON *et al.* 2012b, Chapter 4, Section 4.2.1). There were no statistical differences among the median ages or the male:female ratio of the two groups (Mann-Whitney  $P=0.129$  and Fisher's exact  $P=1$ , respectively). Sixteen SAA and 20 SAC individuals used in this study had been previously genotyped (PICTON *et al.* 2010, Chapter 2), a further six SAA and 11 SAC individuals were newly recruited. This study was approved by the University of the Witwatersrand Committee for Research on Human Subjects, and informed written consent was obtained from all participants.

### 5.2.2. Whole blood surface staining and flow cytometry

Ethylenediaminetetraacetic acid-anticoagulated whole blood obtained from each study participant was stained within one hour of blood collection. Four antibody panels were used for each donor to quantify CCR5 expression on different peripheral blood cells: (i) T cells: CD3-APC, CD4-FITC, CD8-PerCP, CCR5-PE; (ii) B cells: CD19-APC, CCR5-PE; (iii) Natural Killer (NK) cells: CD3-PerCP, CD16-FITC, CD56-APC, CCR5-PE; and (iv) Granulocytes and monocytes: CD45-APC, CD14-FITC, CCR5-PE. All antibodies were obtained from BD Pharmigen (BD Biosciences, Franklin Lakes, NJ, USA). Flow cytometric acquisition, using the FACSCalibur (BD Biosciences, San Jose, CA, USA), quantification and analysis was performed as described previously (PICTON *et al.* 2012b, Chapter 4).

### 5.2.3. CCR5 genotyping

#### 5.2.3.1. CCR5 gene

Participants in this study were genotyped as described previously (PICTON *et al.* 2010). Briefly, genomic DNA was extracted from blood samples; a continuous region encompassing the CCR5 open reading frame (ORF) and the promoter 1 region was polymerase chain reaction (PCR) amplified in overlapping sections using Expand High Fidelity PCR System (Roche, Mannheim, Germany) and sequenced using BigDye Terminator version 3.1 chemistry (Applied Biosystems, Foster City, CA, USA). Sequenced fragments were electrophoresed using the automated 3100 Genetic Analyzer (Applied Biosystems) and HAPLOTYPYER software was used to infer haplotypes (NIU *et al.* 2002).

#### 5.2.3.2. CCR2-V64I

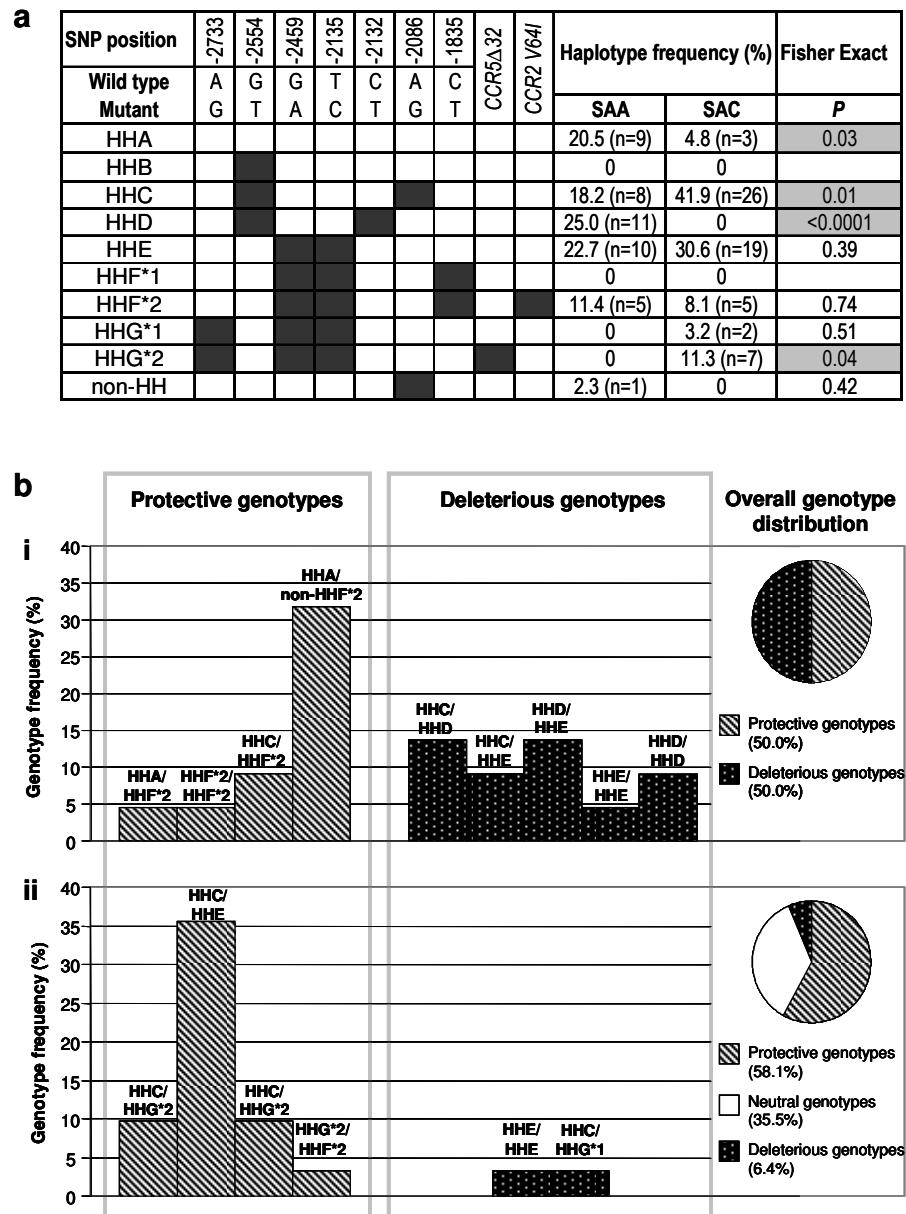
We developed a real-time PCR assay to detect a SNP in the *CCR2* coding region which forms part of the described HHF\*2 haplotype (G190A; *CCR2-V64I*) (GONZALEZ *et al.* 1999). The assay was performed as described previously (Section 2.2.4).

### 5.2.4. Statistical analysis

Mann-Whitney U tests were conducted to compare CCR5 density between individuals, grouped by the presence or absence of certain *CCR5* haplotypes. All statistical analyses were performed using PASW Statistics 18 software (SPSS Inc., Chicago, Illinois, USA).

Fisher exact tests were performed using the Simple Interactive Statistical Analysis software (UITENBROOK 1997) to test whether there was any significant difference in composition of population groups.

### 5.3. RESULTS



**Figure 5.1.** Distribution of *CCR5* haplotypes and genotypes within our two study populations, South African African (SAA) and Caucasian (SAC). **(A)** Schematic representation of frequency at which haplotypes previously defined (GONZALEZ *et al.* 1999) were detected within SAA and SAC individuals. Polymorphic positions (SNP) as well as the wild type mutant change are shown. Shaded boxes indicate SNPs or indels (*CCR5* $\Delta$ 32) which form part of the haplotype and the frequencies of occurrence in the two population groups are indicated. Statistical differences between the two population groups in terms of the representation of haplotypes are indicated, where significant, i.e.,  $P < 0.05$ , these values are highlighted. **(B)** Distribution of genotypes, categorized according to phenotypic effects in the context of HIV-1 infection, within SAA (i) and SAC (ii) populations (GONZALEZ *et al.* 1999; GONZALEZ *et al.* 2005; MALHOTRA *et al.* 2011; MANGANO *et al.* 2001; TANG *et al.* 2002). Genotypes with no such described phenotypes, i.e., neutral genotypes, detected in the SAC population were: HHA/HHE, HHE/HHF\*2, HHF\*2/HHF\*2, HHE/HHG\*1, and HHC paired with HHA, HHC, and HHF\*2.

### 5.3.1. *CCR5* genotype distribution within the two study populations

Full *CCR5* gene sequences were used to determine the *CCR5* genotype of individuals (Figure 5.1A). Individuals within the SAA and SAC populations were assigned to previously described haplogroups (GONZALEZ *et al.* 1999) based on *CCR5* gene SNPs at positions -2733, -2554, -2459, -2135, -2132, -2086 and -1835 as well as the presence of *CCR5* $\Delta$ 32 and the *CCR2* V64I SNP. No HHB and HHF\*1 haplotypes were detected in either population. One individual harboured an allele with a SNP pattern inconsistent with any of the defined haplotypes for *CCR5* (labeled non-HH in Figure 5.1A). Among the other detected haplotypes, there were significant differences ( $P < 0.05$ ) in haplotype frequency between SAA and SAC individuals, in HHA, HHC, HHD, and HHG\*2 haplotypes (Figure 5.1A). The distribution of haplotypes within the two study populations did not differ from what we reported previously (PICTON *et al.* 2010).

In the SAA population, the most prevalent genotypes were HHA/HHE (13.6%), HHC/HHD (13.6%), HHD/HHE (13.6%), HHC/HHF\*2 (9.1%), HHD/HHD (9.1%), and HHC/HHE (9.1%). In the SAC population, the most prevalent genotypes were HHC/HHE (35.5%), HHC/HHC (12.9%), HHC/HHG\*2 (9.7%), and HHE/HHG\*2 (9.7%). Since a number of *CCR5* genotypes are associated either positively or negatively with susceptibility to HIV-1 infection and/or the rate of disease progression to AIDS, individuals within each population were categorized into one of the following categories: (i) those with protective genotypes, (ii) those with deleterious genotypes, and (iii) those with genotypes for which there is no described phenotype in the context of HIV-1 infection (Figure 5.1B).

Protective genotypes were defined as haplotype pairs which have been associated with decreased HIV-1 susceptibility and/or slower disease progression in relation to specific population groups. The following genotypes were designated as protective in the SAA population according to disease modifying associations observed in African Americans: HHA/HHF\*2, non-HHA/HHF\*2 (i.e., HHF\*2/HHF\*2 and HHC/HHF\*2 in our population) and HHA/non-HHF\*2 (non-HHF\*2 haplotypes paired with HHA included HHA, HHC, HHD, HHE, and non-HH in our population) (GONZALEZ *et al.* 1999) (Figure 5.1Bi). Haplotype pairs considered as protective in SAC individuals were: HHC/HHG\*2, HHC/HHE and non-HHC/HHG\*2 (i.e., HHC/HHG\*2 and HHF\*2/HHG\*2 detected in SAC) (GONZALEZ *et al.* 1999) (Figure 5.1Bii). Deleterious genotypes were defined as haplotype pairs which have been associated with faster disease progression to AIDS and death and/or increased susceptibility to HIV-1 infection in populations of African and Caucasian ancestry. These include HHE/HHE and HHC/HHG\*1 detected in the SAC population (GONZALEZ *et al.* 1999; GONZALEZ *et al.* 2005) (Figure 5.1Bii) and HHC/HHD, HHC/HHE, HHD/HHE and HHD/HHD in the SAA population (GONZALEZ *et al.* 1999; JOHN *et al.* 2001; KOSTRIKIS *et al.* 1999; MALHOTRA *et al.* 2011) (Figure 5.1Bi). Although, in populations of African ancestry, the HHE/HHE genotype has not

been specifically shown to be deleterious, i.e., associate with increased susceptibility to HIV-1 or accelerated disease progression, a trend towards this has been reported, thus this genotype was designated as deleterious in our study (TANG *et al.* 2002). Haplotype pairs for which there is no reported disease modifying effects were designated as neutral.

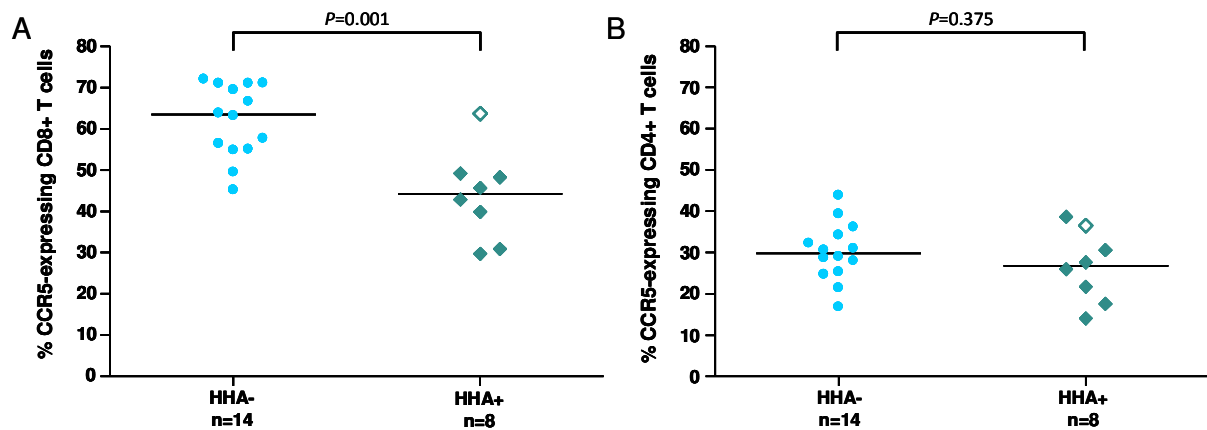
SAA individuals were equally divided between those with protective and those with deleterious genotypes (50.0% each, Figure 5.1Bi), whereas SAC individuals had a greater proportion of individuals with protective genotypes compared to those with deleterious genotypes (58.1 and 3.2%, respectively, Figure 5.1Bii) and a large proportion of genotypes designated as neutral (35.5%).

Although it is an individuals' genotype, i.e., the combination of the *CCR5* haplotypes, that is likely to determine the overall *CCR5* expression, given the number of possible *CCR5* haplotype combinations, it was not possible to assess the effect of individual genotypes on expression. Instead, we investigated what the effect the presence or absence of particular haplotypes had on *CCR5* expression.

### 5.3.2. Influence of haplotypes on *CCR5* expression

*CCR5* haplotypes were chosen for analysis of their impact on *CCR5* expression based on the following criteria: (i) haplotypes present at high frequencies within one or both of the study populations and (ii) haplotypes with known positive or negative association to HIV-1 susceptibility and/or disease progression. The most common haplotypes are HHA, HHC, HHD, and HHE in SAA, and HHC, HHE, and HHG\*2 in SAC individuals (PICTON *et al.* 2010). The HHA and HHC haplotypes are associated with slower disease progression in HIV-1 infected African American and Caucasian individuals, respectively (GONZALEZ *et al.* 1999). The HHE haplotype is associated with an increased risk of acquiring HIV-1 infection as well as faster disease progression in ethnically divergent populations (GONZALEZ *et al.* 2005; MALHOTRA *et al.* 2011; MANGANO *et al.* 2001; OMETTO *et al.* 2001; TANG *et al.* 2002). However, this association appears to be dependent on an individual's *CCR5* genotype and not solely on the presence or absence of this haplotype, i.e., the combinatory effect of HHE and its allelic pair (GONZALEZ *et al.* 1999; MALHOTRA *et al.* 2011). The HHG\*2 haplotype, or *CCR5*<sub>Δ32</sub> allele, present at a haplotype frequency of 11.3% in the SAC population (Figure 5.1) has been studied extensively and shown to reduce *CCR5* expression. We recently demonstrated HHG\*2 to impact upon *CCR5* expression within the same cohort (PICTON *et al.* 2012b), with individuals with HHG\*2 having both reduced percentage of *CCR5*-expressing cells and *CCR5* density, a result consistent with previous studies (DE RODA HUSMAN *et al.* 1999; SHALEKOFF and TIEMESSEN 2003; WU *et al.* 1997). Lastly, the presence of a SNP unique to the HHD haplotype has been linked with increased perinatal HIV-1 acquisition in infants, and

accelerated maternal HIV-1 disease progression (JOHN *et al.* 2001; KOSTRIKIS *et al.* 1999). However, a more recent report has shown sub-Saharan infants with this same SNP to have a lower rate of HIV-1 acquisition (SINGH *et al.* 2008). Thus, the potential impact of the presence or absence of the HHA, HHC, HHD, and HHE haplotypes found in the two study populations, upon CCR5 expression was investigated. Individuals were stratified according to the presence of a particular haplotype (heterozygous or homozygous) or absence of that same haplotype. Since it has been shown that the disease modifying effect of a particular haplotype may differ according to the population being studied and that populations differ substantially in the prevalence of CCR5 haplotypes (GONZALEZ *et al.* 1999), we analyzed the two study populations independently. CCR5 expression, both as percentage of CCR5-expressing cells and mean CCR5 density on CCR5-expressing cells, was determined across B, T, NK, and monocyte cell subsets. The haplotypes investigated were found to have no impact upon CCR5 expression on B cells or monocytes (data not shown); however, differences in CCR5 expression were noted within T and NK cell subsets.



**Figure 5.2.** Influence of the *CCR5* HHA haplotype on CCR5 expression in South African African (SAA) individuals within the CD8+ (A) and CD4+ (B) T cell subsets. Individuals with or without an HHA allele are designated as HHA+ (♦) and HHA- (●), respectively. The HHA homozygous individual is indicated with (◊). P values and the number of individuals in each group are indicated.

### 5.3.3. HHA and T cells in the SAA population

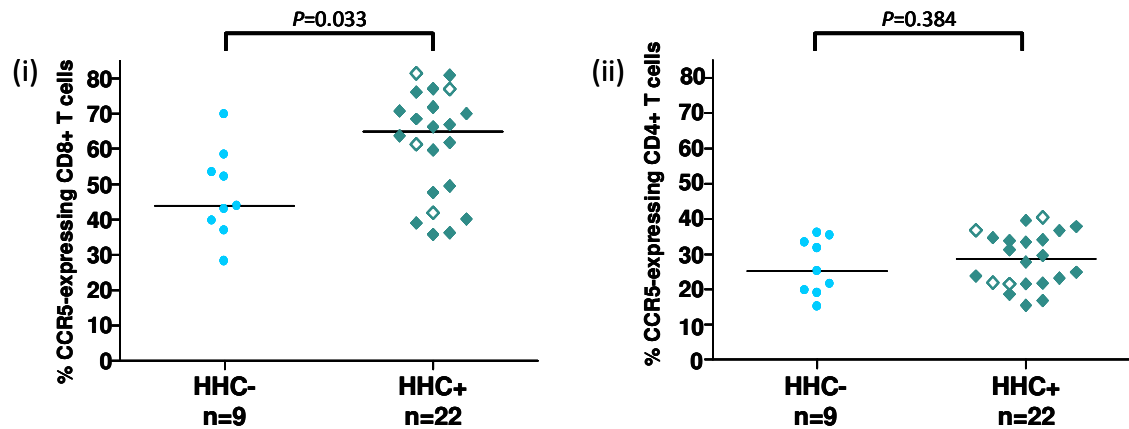
The HHA haplotype, present in 36.4% (8/22) of SAA individuals at a haplotype frequency of 20.5% (n=9 alleles, one homozygous individual) (Figure 5.1) was found to influence CCR5 expression (Figure 5.2). SAA individuals whose *CCR5* genotype contained the HHA haplotype had significantly lower percentages of CCR5-expressing CD8+ T cells, in comparison to SAA individuals without the HHA haplotype ( $P=0.001$ , Figure 5.2A). However, no significant relationship was observed for CD4+ T cells ( $P=0.375$ , Figure 5.2B). Furthermore, no difference in CCR5

density was noted between HHA+ and HHA- individuals in either CD4+ or CD8+ T cell subsets ( $P=0.815$  and  $P=1.0$ , respectively). SAC individuals have a low HHA haplotype representation (4.8%, Figure 5.1) and thus the impact of HHA on CCR5 expression in this population could not be determined.

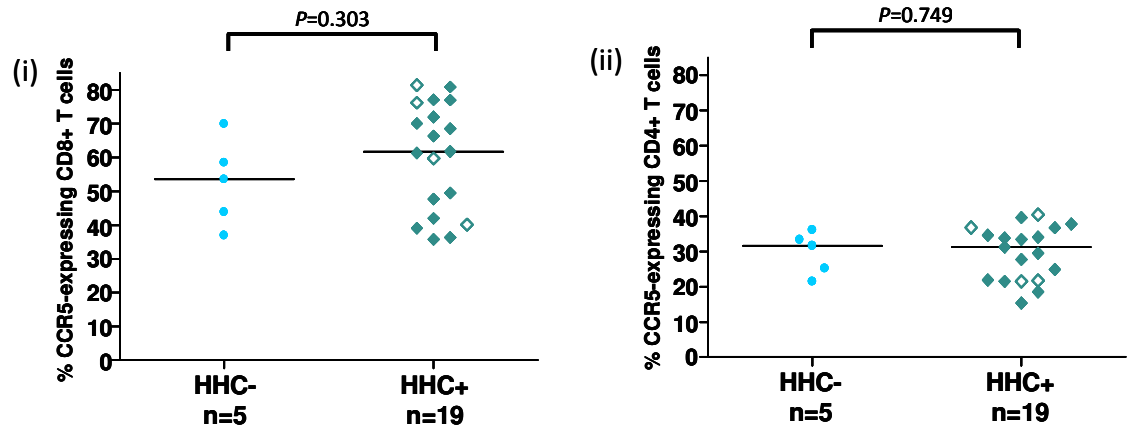
#### 5.3.4. HHC and T cell subsets

The HHC haplotype, present in 71.0% of the SAC population (22/31) at a haplotype frequency of 41.9% ( $n=26$  alleles, four homozygous individuals) was also found to have an impact on CCR5 expression in this population group (Figure 5.3 and Figure 5.4). SAC individuals whose *CCR5* genotype contained the HHC haplotype had significantly higher percentages of CCR5-expressing CD8+ T cell subsets in comparison to SAC individuals lacking this haplotype ( $P=0.033$ , Figure 5.3Ai), however this was not observed for CD4+ T cells ( $P=0.384$ , Figure 5.3Aii). Among the seven HHG\*2 (*CCR5* $\Delta$ 32) haplotype bearing individuals, the genotype distribution was as follows: HHC/HHG\*2 ( $n=3$ ), HHE/HHG\*2 ( $n=3$ ) and HHF\*2/HHG\*2 ( $n=1$ ), i.e., three were HHC positive and four were HHC negative. When *CCR5* $\Delta$ 32 heterozygous individuals were excluded from the analysis, the relationship between the HHC haplotype and the percentage of CCR5-expressing CD8+ T cells was lost ( $P=0.303$ , Figure 5.3Bi), indicating that the differences observed between HHC+ and HHC- SAC individuals could be attributed to the influence of the *CCR* $\Delta$ 32 mutation upon expression. Percentages of CCR5-expressing CD8+ T cells did not differ between HHC+ and HHC- SAA individuals ( $P=0.759$ , Figure 5.3Ci). No difference in the percentage of CCR5-expressing CD4+ T cells was noted between individuals with or without the HHC haplotype in either population ( $P=0.749$ , Figure 5.3Bii and  $P=0.495$ , Figure 5.3Cii).

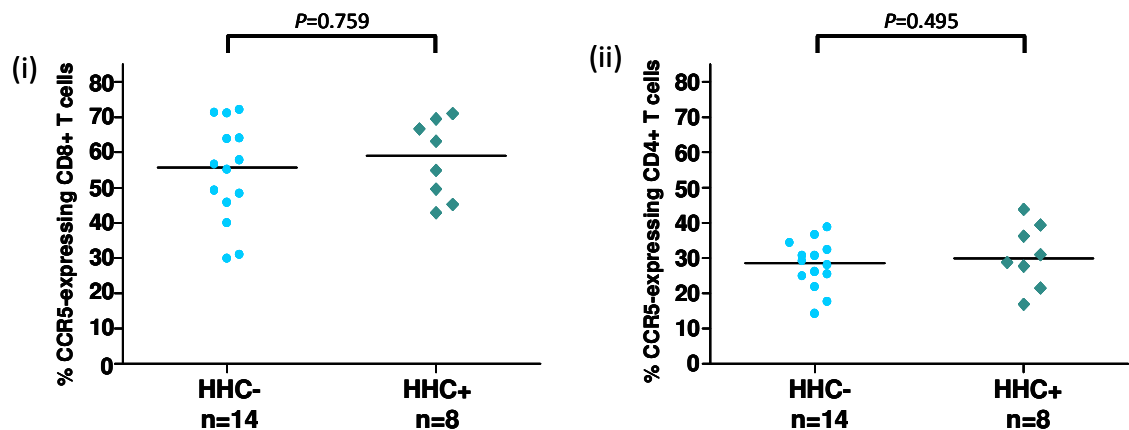
**A. SAC:**



**B. SAC (without *CCR5Δ32*):**



**C. SAA:**



**Figure 5.3.** Influence of the HHC haplotype on the percentage of CCR5-expressing cells in T cell subsets, CD8+ (i) and CD4+ (ii), in South African Caucasian (SAC) individuals (**A**), SAC population with the exclusion of individuals with the *CCR5Δ32* allele (**B**), and in South African African (SAA) individuals (**C**). Individuals with or without an HHC allele are indicated as HHC+ and HHC-, respectively. The HHC homozygous individuals are indicated with (◊). P values and the number of individuals in each group are indicated.

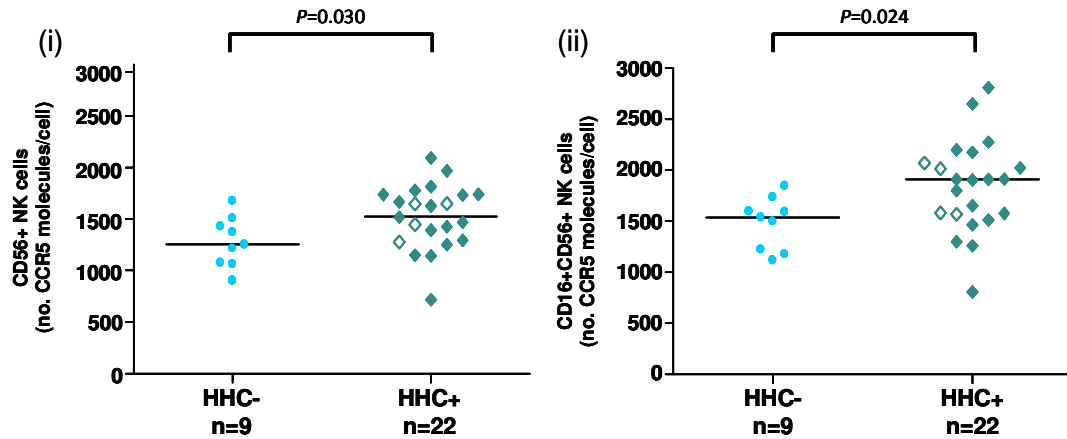
### 5.3.5. HHC and NK cell subsets

The two groups within the SAC population, i.e., HHC positive and HHC negative, differed in terms of CCR5 density on NK (CD3-CD56+) and CD16+CD56+ NK cell subsets (Figure 5.4) but not in terms of percentage of CCR5-expressing cells within the same subsets (Figure C.1). No significant differences were noted within the CD56<sup>bright</sup> and CD56<sup>dim</sup> NK cell subsets (data not shown). HHC+ SAC individuals had significantly higher CCR5 density on NK (CD56+) and CD16+CD56+ NK cell subsets ( $P=0.030$  and  $P=0.024$ , Figure 5.4Ai and ii, respectively) but not in the percentage of the respective cell subsets that express CCR5 ( $P=0.587$  and  $P=0.601$ , respectively, Figure C.1). Importantly, the relationship between HHC haplotype and CCR5 density was maintained in NK cells when *CCR5Δ32* heterozygotes were excluded from the analysis ( $P=0.043$  and  $P=0.025$  on NK cell and CD16+CD56+ NK cell subsets, respectively, Figure 5.4Bi and ii). No relationship between CCR5 density and the presence of the HHC haplotype was seen in SAA individuals for the NK cell and CD16+CD56+ NK cell subsets ( $P=0.682$  and  $P=0.539$ , Figure 5.4Ci and ii, respectively). It is interesting to note that the outlier within the SAA HHC- group, with the highest CCR5 density on CD56+ and CD16+CD56+ NK cell subsets within that population (1930 and 2063 average CCR5 molecules/cell, respectively), was the only SAA individual in the study group with the HHE/HHE genotype. If this individual was removed from analysis within the SAA population, CCR5 density remained comparable between the HHC- and HHC+ group (data not shown).

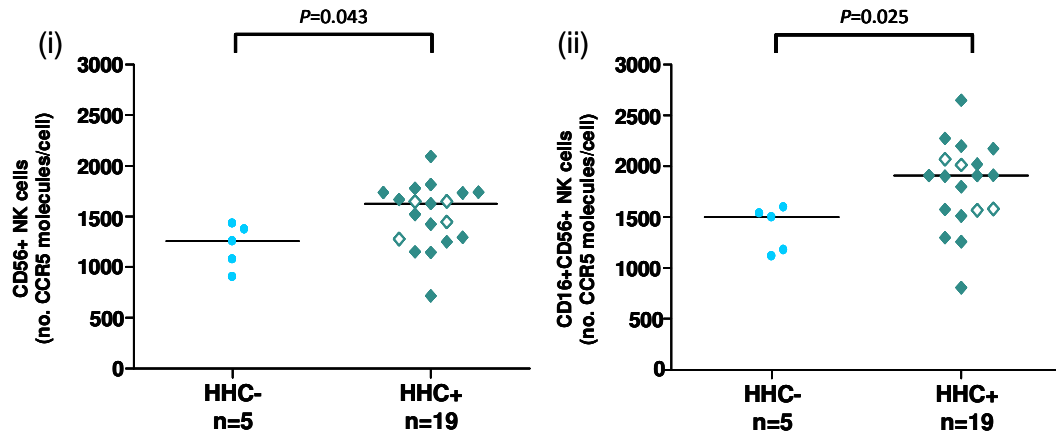
HHC- and HHC+ groups were compared across the two study populations, i.e., SAA HHC- vs SAC HHC- and SAA HHC+ vs SAC HHC+. SAA and SAC (without *CCR5Δ32* or “wild type CCR5 ORF”) individuals demonstrated similar densities between the HHC- groups ( $P=0.781$  and  $P=0.308$  on CD56+ and CD16+CD56+ NK cell subsets, respectively). However, the HHC+ groups differed between the two populations. HHC+ SAC “wild type CCR5 ORF” individuals ( $n=19$ ) had significantly higher CCR5 density on CD16+CD56+ NK cells ( $P=0.019$ ) and a trend towards higher density on CD56+ NK cells ( $P=0.063$ ) than HHC+ SAA individuals.

The haplotype pairs for the eight HHC+ SAA individuals were as follows: one HHA, three HHD, two HHE and two HHF\*2. Thus, the majority of HHC+ SAA individuals (5/8, 62.5%) were paired with HHD and HHE, genotypes which have been associated with accelerated disease progression in HIV-1 infection in African Americans (GONZALEZ *et al.* 1999). In contrast, the HHC/HHE genotype has been shown to be “protective” in Caucasian populations and has a demonstrated association with a delay in time to death (GONZALEZ *et al.* 1999). Within the SAC population, 11/19 (57.9%) HHC+ CCR5 ORF wild type individuals had the genotype HHC/HHE (Figure 5.1B).

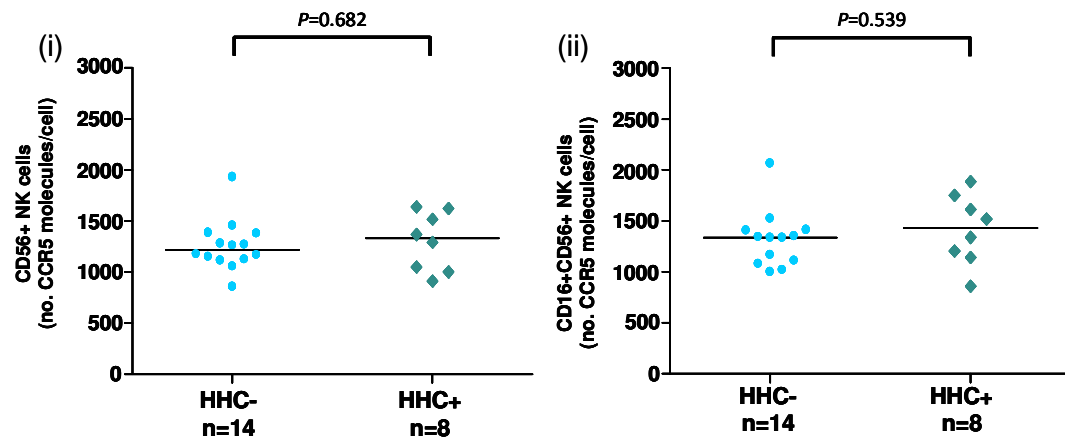
**A. SAC:**



**B. SAC (without *CCR5Δ32*):**



**C. SAA:**



**Figure 5.4.** Influence of the HHC haplotype on CCR5 density on natural killer cell subsets, CD56+ NK (i) and CD16+CD56+ NK (ii), within the South African Caucasian (SAC) population (A), the SAC population with the exclusion of individuals with the *CCR5Δ32* allele (B) and the SAA population (C). Individuals with or without an HHC allele are indicated as HHC+ and HHC-, respectively. The HHC homozygous individuals are indicated with (◇). *P* values and the number of individuals in each group are indicated.

### 5.3.6. HHD and HHE haplotypes

The HHE haplotype, found at comparable frequencies in the two populations ( $P=0.39$ , Figure 5.1), was found to have no significant impact upon CCR5 expression levels in either SAA or SAC individuals (data not shown). The HHD haplotype, found in SAA but not in SAC individuals (PICTON *et al.* 2010), had no impact on CCR5 expression levels on any of the cell populations studied, both in terms of the number of molecules expressed on cells and the proportion of CCR5 expressing cells (data not shown). Furthermore, if individuals with the HHD/HHE genotype ( $n=3$ ), a genotype associated with increased susceptibility to HIV-1 infection and higher viral loads (MALHOTRA *et al.* 2011), were excluded from the analysis, no association with presence or absence of the HHD haplotype was found (data not shown).

## 5.4. DISCUSSION

Although a significant factor, heterozygosity for the *CCR5* $\Delta$ 32 allele is not the only determinant of variability in CCR5 cell surface expression, both in terms of CCR5 density and the percentage of CCR5-expressing cells. CCR5 expression can differ by as much as 20-fold in individuals that lack this mutation (CCR5 ORF wild type) (LEE *et al.* 1999; MOORE *et al.* 1997; REYNES *et al.* 2000; WU *et al.* 1997). Furthermore, CCR5 ORF wild type individuals can have CCR5 expression comparable to that of *CCR5* $\Delta$ 32 heterozygotes (WU *et al.* 1997). Thus, it follows that other genetic variants within the regulatory region of *CCR5* may influence CCR5 expression. *CCR5* haplotypes have been demonstrated to associate with haplotype-specific differences in transcriptional activity within K562 and Jurkat cell lines (MUMMIDI *et al.* 2000). We have previously demonstrated two South African populations to have marked differences in CCR5 haplotype prevalence as well as CCR5 expression (PICTON *et al.* 2010; PICTON *et al.* 2012b, Chapter 2 and 4). Therefore, in this study, we investigated whether specific haplotypes found at high frequencies in these two populations, contribute towards differences in CCR5 expression in healthy HIV-1 uninfected individuals.

In adults, the *CCR5* HHE haplotype has been associated with acceleration to AIDS in different populations including Caucasian, African American, Asian, and African populations (AN *et al.* 2000; GONZALEZ *et al.* 1999; LI *et al.* 2005; MARTIN *et al.* 1998; NGUYEN *et al.* 2004; TANG *et al.* 2002). HIV-1 infected infants with this haplotype progress to AIDS faster than infants without the HHE haplotype (MANGANO *et al.* 2001; OMETTO *et al.* 2001). Furthermore, this haplotype has also been associated with increased risk of acquisition of HIV-1 infection in both adults and HIV-1 exposed infants (MALHOTRA *et al.* 2011; MANGANO *et al.* 2001; TANG *et al.* 2002) and with higher viral loads in infected adults (NGUYEN *et al.* 2004; TANG *et al.* 2002). Thus, we sought potential associations with CCR5 expression in individuals positive for this haplotype in the two study populations. The HHE haplotype did not appear to influence CCR5 expression in either study population. However,

it has been reported that HHE homozygosity, but not heterozygosity in Caucasians, and pairing of HHE with HHC in African Americans is associated with disease acceleration, indicating that it is the haplotype pairing (genotype) that is important here (GONZALEZ *et al.* 1999). Thus, it is to be expected that the presence of the HHE haplotype alone will not affect CCR5 expression. With only one HHE homozygous individual in each of the study populations and two HHC/HHE SAA individuals, it was not possible to independently assess the effect of these haplotype combinations on CCR5 expression.

In a cohort of 90 HIV-negative SAA female sex workers, those possessing the HHD allele were shown to have higher CCR5 density on CD4+ HLA-DR+ T cells compared to those that lack this allele (MAMTANI *et al.* 2011). We found that the presence of the HHD haplotype alone was not sufficient to alter CCR5 expression levels in SAA individuals, in any of the subsets examined. However, our analysis was conducted on CD4+ and CD8+ T cell subsets as a whole and did not distinguish between activated and non-activated subsets. Furthermore, the two cohorts differ: our study group is comprised of laboratory workers who one would expect to have lower environmental exposure to potential pathogens, and hence lower cellular activation, compared to the cohort of female sex workers. Since activated CD4+ T cells express CCR5 at higher levels compared to nonactivated CD4+ T cells (BEGAUD *et al.* 2006; KALINKOVICH *et al.* 2001; REYNES *et al.* 2000; WU *et al.* 1997), it is possible that HHD haplotype induced expression differences may become more apparent with higher activation states. A recent study showed that haplotype HHD, largely restricted to African populations (GONZALEZ *et al.* 2001), associates with increased susceptibility to HIV-1 infection and higher viral loads when present in combination with HHE, i.e., HHD/HHE genotype (MALHOTRA *et al.* 2011). Only three of our HHD+ SAA individuals had the HHD/HHE genotype, thus it was not possible to investigate whether this association may be related to increased CCR5 expression.

Two haplotypes, HHA and HHC, both associated with delayed disease progression in HIV-1 infected African American and Caucasian individuals, respectively (GONZALEZ *et al.* 1999), were demonstrated to have an impact upon CCR5 expression in the corresponding populations. Promoter expression studies have shown the ancestral HHA haplotype to have the lowest transcriptional activity *in vitro* (MUMMIDI *et al.* 2000). This and the reported protective effects of the HHA allele in individuals with African ancestry where possession of a single allele confers disease retardation (GONZALEZ *et al.* 1999), led us to expect lower CCR5 expression on CD4+ T cells in HHA positive individuals. However, HHA positive SAA individuals had lower percentages of CCR5-expressing CD8+ T cells compared to individuals lacking this allele. Lower percentages of CCR5-expressing CD8+ T cells could mean that fewer cells are able to be recruited from the periphery

into an immune response, and if so this would suggest that a small deficit in these cells is not disadvantageous.

Overall, the HHC haplotype was not found to influence expression on T cell subsets in either study population. However, in the SAC population, individuals with the HHC haplotype demonstrated higher CCR5 density in NK cell subsets (CD56+ and CD16+CD56+) compared to individuals without this haplotype. This relationship was not influenced by the presence of the *CCR5* $\Delta$ 32 mutation (HHG\*2) and was absent in SAA individuals. HHC is associated with disease retardation, particularly a delayed progression to death in Caucasians but is associated with disease acceleration in African Americans (GONZALEZ *et al.* 1999). Given the reported contrasting influence of the HHC haplotype in the context of HIV-1 infection, it is likely that the observed increased density on NK cell subsets is a factor contributing to the delay in disease progression afforded by the HHC haplotype in Caucasian populations, possibly by enhancing their migration and their killer capability in response to CCR5 ligands released through immune activation in chronic HIV-1 infection, or at the site of a primary infectious insult.

In murine studies, CCR5-expressing NK cells have been demonstrated to play critical roles in other diseases. For example, CCR5 “null” mice, infected with *Toxoplasma gondii*, exhibit a decreased inflammatory response, resulting in higher parasite burden and ultimate mortality as a result of decreased migration of NK cells into infected tissues (KHAN *et al.* 2006). Similarly, in the context of infection with an unrelated pathogen, CCR5 ‘null’ mice exhibit increased sensitivity to infection and a diminished capacity to control herpes simplex virus 2 (HSV-2) infection as a result of reduced NK cell recruitment (THAPA *et al.* 2007). It is interesting to note that HSV-2 is the most common co-pathogen that enhances the susceptibility of individuals to other sexually transmitted diseases, including HIV-1 (FREEMAN *et al.* 2006). In another mouse study, CCR5 “null” mice were shown to have reduced numbers of NK cells in the bone marrow, spleen, blood, and liver compared to wild type mice in the absence of infection (WEISS *et al.* 2011). Interestingly, this was not observed in mice lacking CCR1, a chemokine receptor which binds two of the three CCR5 chemokine ligands, CCL3, and RANTES (WEISS *et al.* 2011). NK cells in the bone marrow of CCR5 “null” mice were further shown to proliferate to a lesser extent compared to wild type mice and it was suggested that this may be due to improper localization of these cells to their specific niche or the deficiency in co-stimulatory signals provided by CCR5 (WEISS *et al.* 2011). Therefore, it follows that differences in the levels of NK cell CCR5 expression, may have an impact on NK cell trafficking in similar infections in humans.

However, despite the observed associations between *CCR5* haplotypes and CCR5 expression, either in terms of percentage of CCR5-expressing cells or CCR5 density, it is important to bear in mind that the phenotypic effect of an individual’s *CCR5* genotype is not only determined by the

presence or absence of a single HIV-disease modifying allele but by the combination of the individual's two alleles (GONZALEZ *et al.* 1999).

We recently demonstrated CCR5 expression to differ substantially between SAA and SAC individuals and that these differences can be attributed to factors other than contributions made by *CCR5Δ32* (haplotype HHG\*2), found in SAC but not SAA individuals (PICTON *et al.* 2012b, Chapter 4). Differences between the two populations in CCR5 expression were seen across all cell subsets analyzed, i.e., T cells, NK cells, B cells, and monocytes. The significant difference in haplotype distribution and frequency between SAA and SAC individuals may explain some of the observed differences. For example, the HHC haplotype found at significantly higher haplotype frequencies in the SAC population, or more specifically the predominant genotype HHC/HHE associated with delayed time to death in HIV-1 infected Caucasians (GONZALEZ *et al.* 1999), would explain the higher CCR5 densities seen on NK cells in SAC individuals compared to SAA individuals (PICTON *et al.* 2012b). However, although the HHA haplotype, present in SAA individuals at significantly higher frequencies compared to SAC individuals, was shown to associate with a lower percentage of CCR5-expressing CD8+ T cells, SAA individuals did not have lower percentages of CCR5-expressing CD8+ T cells compared to SAC individuals (PICTON *et al.* 2012b).

We observed the genotype distribution, in terms of representation of protective versus deleterious genotypes, according to specific populations, to vary between SAA and SAC individuals. The SAA population had a greater proportion of individuals with deleterious genotypes. Although this would suggest that the two populations may differ from each other with respect to susceptibility to infection and/or rate of disease progression, we would caution against making such a conclusion. With the exception of the HHG\*2 allele (*CCR5Δ32*), found predominantly in Caucasian populations, there are no data comparing the combinatory effects of protective and deleterious genotypes between population groups infected with HIV-1. In addition, to our knowledge, there are no data relating *CCR5* genotypes to HIV-1 risk of infection or progression to AIDS in South African populations, thus it is unknown whether genotypes classed as either “protective” or “deleterious” in other geographically distinct populations can be inferred to have the same phenotypic effect in the context of HIV-1 infection in South African individuals. Furthermore, it should be kept in mind that many other host genetic factors aside from *CCR5* influence HIV-1 infection.

CCR5 surface expression can be influenced by a number of factors, including levels of its chemokine ligands (GONZALEZ *et al.* 2005; KETAS *et al.* 2007; LIN *et al.* 2008), cell activation levels (MEDITZ *et al.* 2011; MORIUCHI *et al.* 1997; OSTROWSKI *et al.* 1998; WU *et al.* 1997) as well as other molecules, such as interleukin-2, interleukin-12, and interferon- $\alpha$  (WEISSMAN *et al.* 2000; YANG *et al.* 2001a; YANG *et al.* 2001b; ZOU *et al.* 1999). In this study, we demonstrate how CCR5 expression can also be influenced differently, i.e., in terms of percentage of CCR5-expressing cell

subsets or the density of CCR5, by some of the more prevalent haplotypes of the *CCR5* regulatory region. We have previously discussed the implications of measures of CCR5 expression, i.e., percentage of CCR5-expressing cells and CCR5 density, in the context of HIV-1 infection (PICTON *et al.* 2012b). It should be kept in mind that the complement and distribution of CCR5-expressing immune cells at mucosal surfaces is quite different from that in peripheral blood, and studies directly assessing how CCR5 haplotypes influence these parameters will shed further light on host susceptibility/resistance to HIV-1.

The most striking observation was the higher CCR5 density on NK cell subsets of HHC positive SAC individuals, providing an explanation for the differences in expression observed between SAA and SAC individuals (PICTON *et al.* 2012b), and the reported divergent outcomes of disease progression based on ethnicity. Based on our data, we propose that NK cells with higher CCR5 density would traffic to the mucosa upon antigenic challenge to a greater extent than NK cells with lower CCR5 density. This would potentially offset the number of CCR5+ CD4+ T cells vulnerable to infection. Since NK cells are increasingly being found to be important players with respect to HIV-1 susceptibility and disease progression (ALTER and ALTFELD 2009; SCOTT-ALGARA *et al.* 2003; TIEMESSEN *et al.* 2009; TIEMESSEN *et al.* 2010), investigation into the possible role of differential CCR5 expression on these cells that is afforded by different *CCR5* haplotypes should form the basis of future studies aimed at understanding HIV-1 protective immunity.

---

## CHAPTER 6

Contribution of variable *CCL3L* copy number to CCL3 protein production in two ethnically divergent South African populations

---

## ABSTRACT

When accounting for the specific population, *CCL3L* copy number, a measure of the sum of chemokine- and non-chemokine-producing genes (*CCL3La* and *CCL3Lb*, respectively), has been reported to associate with risk of HIV-1 infection. In this study, we have described the distribution of *CCL3La* and *CCL3Lb* copy number variation in two populations, South African Africans (SAA) and South African Caucasians (SAC), and investigated the impact of these variations upon CCL3 protein production. Despite significant differences in *CCL3La* and *CCL3Lb* copy number, no differences in CCL3 production were noted between the two populations. Assuming equal contribution of CCL3 and each copy of *CCL3La* to CCL3 production, we found that SAC individuals produced higher levels of CCL3 per functional copy of *CCL3La* compared to SAA individuals ( $P < 0.001$ ). However, when individuals with comparable *CCL3La* and *CCL3Lb* gene copy numbers were compared, no difference in production per functional copy between SAA and SAC individuals was noted. Furthermore, we demonstrate that differences noted in cord blood mononuclear cell CCL3 production between HIV-1 intrapartum-infected (IP) and exposed uninfected (EU) infants with comparable *CCL3L* copy numbers could not be attributed to differences in *CCL3Lb* copy number. Collectively, our findings suggest that either the *CCL3* gene may play a significant role in CCL3 production and/or that as yet undefined mechanisms regulate production of CCL3 from variable *CCL3L* copy number.

## 6.1. INTRODUCTION

Chemokines are small structurally related cytokines which play a major role in the chemotaxis of various leukocyte cell types to sites of injury or infection (Reviewed in FURIE and RANDOLPH 1995). These molecules are also important in other processes including inflammatory response, angiogenesis and haematopoiesis as well as in pathological conditions such as HIV-1 infection, autoimmune diseases and carcinogenesis (BALKWILL 2004; MENTEN *et al.* 2002; ROLLINS 2006). Chemokines are classified into four subfamilies based upon the spacing of the two N-terminal cysteine residues, with C-chemokines having only one N-terminal cysteine residue (ZLOTNIK and YOSHIE 2000). The chemokine, CCL3, also known as MIP-1 $\alpha$  (macrophage inflammatory protein-1 $\alpha$ ), is a member of the CC-subfamily of chemokines.

Genes coding for many chemokines are located within defined chromosomal clusters. The *CCL3* gene family is localized on chromosome 17q21 (HIRASHIMA *et al.* 1992). Through sequencing of a BAC clone located in this region, two complete nonallelic copies (*CCL3L1* and *CCL3L3*) and a truncated copy (*CCL3L2*) were identified (MODI 2004).

Highly related functional genes (94% nucleic acid homology) (NAKAO *et al.* 1990), *CCL3* (*LD78α* or *MIP-1α*) and *CCL3La* (*LD78β* or *MIP-1αP*, *CCL3L1* and *CCL3L3*), are known to code for two CCL3 isoforms. These two isoforms of the CCL3 chemokine, CCL3 (MIP-1α) and CCL3L1 (MIP-1αP), differ only in three amino acids (MENTEN *et al.* 1999; MODI *et al.* 2001). Functionally, the CCL3L chemokine, encoded by the *CCL3L* non-allelic copies, appears to be the most potent agonist of CCR5 and inhibitor of R5 HIV-1 strains in comparison to the allelic CCL3 chemokine (AQUARO *et al.* 2001; MENTEN *et al.* 1999; NIBBS *et al.* 1999). Another related gene, *CCL3Lb* (*CCL3L2* or *LD78γ*), was initially identified as a 5'-truncated pseudogene, as it contains only 2 exons homologous to exons 2 and 3 of the *CCL3L1* and *CCL3L3* genes (HIRASHIMA *et al.* 1992). However, novel 5' exons have been recently identified within *CCL3Lb* which give rise to two alternatively spliced transcripts (SHOSTAKOVICH-KORETSKAYA *et al.* 2009). These mRNA transcripts contain chemokine-like domains but are not predicted to code for classical chemokines (SHOSTAKOVICH-KORETSKAYA *et al.* 2009). To our knowledge, a functional role for these *CCL3Lb* mRNA transcripts has not been reported.

*CCL3* is present at two copies per diploid genome (pdg), whereas *CCL3La* and *CCL3Lb* are present at variable copy number pdg. Copy number variation is found between populations as well as individuals within a population. Stimulation experiments have shown a positive relationship between *CCL3L* copy number and CCL3 production, i.e., higher *CCL3L* copy numbers associate with increased mRNA levels (CARPENTER *et al.* 2012; TOWNSON *et al.* 2002; URBAN *et al.* 2009), as well as CCL3/CCL3L1 protein levels (GONZALEZ *et al.* 2005; MEDDOWS-TAYLOR *et al.* 2006; TOWNSON *et al.* 2002). A recent study failed to demonstrate a significant correlation between mean CCL3 plasma concentration and *CCL3L* copy number in a cohort of HIV-1 infected individuals (LARSEN *et al.* 2012). Previous studies investigating the association between *CCL3L* copy number and protein production have measured CCL3 production post stimulation of isolated cells, while the latter study assessed circulating CCL3 levels in plasma. However, the assays used did not distinguish between *CCL3La* and *CCL3Lb*. Therefore, we investigated the effects of *CCL3La* gene copy number on CCL3 protein levels independently of *CCL3Lb*. Furthermore, we sought to determine whether South African African (SAA) and South African Caucasian (SAC) populations differ in their CCL3 production given that these two populations differ in their *CCL3L* copy number distribution. In addition, we genotyped infants from a cohort used to study mother to child transmission to test our hypothesis that a higher representation of the 'non-functional' *CCL3Lb* gene might account for the reduced CCL3 production in HIV-1 intrapartum-infected (IP) compared to exposed-uninfected (EU) infants with higher *CCL3L* (*CCL3La*+*CCL3Lb*) copy number.

## 6.2. MATERIALS AND METHODS

### 6.2.1. Study populations

#### 6.2.1.1. Healthy participants

One hundred and forty three healthy HIV-1 uninfected individuals, 89 SAA and 54 SAC, were genotyped for *CCL3L*, *CCL3La* and *CCL3Lb* copy number in order to determine the copy number distribution within the two populations. From this cohort, a subset of 22 SAA and 31 SAC individuals, age and sex matched as described previously (PICTON *et al.* 2012b, Chapter 4, Section 4.2.1.), were selected to assess the ability of their peripheral blood mononuclear cells (PBMCs) to produce CCL3 upon stimulation.

#### 6.2.1.2. HIV-1 exposed-uninfected and intrapartum-infected infants

Fifty four infants born to HIV-1 infected SAA women were genotyped for *CCL3L*, *CCL3La* and *CCL3Lb* copy number to determine whether the *CCL3Lb* gene was contributing towards previously detected differences in CCL3 production between exposed-uninfected (EU) (n=41) and intrapartum-infected (IP) (n=13) infants (MEDDOWS-TAYLOR *et al.* 2006). This cohort has been described by Meddows-Taylor *et al.* (2006).

This study was approved by the University of the Witwatersrand Committee for Research on Human Subjects, and informed written consent was obtained from all participants.

### 6.2.2. Nomenclature

Different nomenclature has been used in various studies in reference to both the genes coding for CCL3 as well as the chemokine itself. In this study, CCL3 refers to both isoforms of the chemokine protein, i.e., CCL3 + CCL3L, as the two, almost identical, isoforms cannot be distinguished from each other by means of enzyme-linked immunosorbent assays (ELISA). With regards to the coding genes, *CCL3* refers to the allelic gene present as 2 copies pdg, whereas *CCL3L* refers to the sum of the copy number variable genes, *CCL3La* and *CCL3Lb*, definitions described by Shostokovich-Koretskaya *et al.* (2009). The human genome reference assembly contains two full syntenic copies of the non-allelic *CCL3L* gene which have been designated as *CCL3L1* and *CCL3L3*. Thus, *CCL3La* is used to denote the cumulative copies of *CCL3L1* and *CCL3L3*. Although *CCL3L1* and *CCL3L3* are separate genes, both genes have three identical exons and encode identical proteins (MENTEN *et al.* 2002; MODI 2004).

### 6.2.3. Isolation of Peripheral Blood Mononuclear Cells

PBMCs were isolated from blood samples anticoagulated with ethylenediaminetetraacetic acid (EDTA) by centrifugation on Ficoll-Paque™ Plus (GE Healthcare, Sweden) using standard methods. PBMCs were resuspended in RPMI medium (Gibco) containing 10% fetal bovine serum (FBS, Gibco) at a concentration of  $2 \times 10^6$  cells/ml.

### 6.2.4. CCL3 production assays

Equal numbers of isolated PBMCs ( $1.4 \times 10^6$  cells) were aliquoted per well in 24 well plates (Corning, NY, USA) with or without the addition of phytohaemagglutinin (PHA) at a final concentration of 12.5 µg/ml (stimulated and unstimulated, respectively), and incubated at 37°C, 5 % CO<sub>2</sub>. Following 20 hours of incubation, culture supernatants were harvested and stored at -70°C until analyses were performed.

### 6.2.5. DNA isolation

Genomic DNA was extracted from whole blood using QIAamp DNA Mini Kit (QIAGEN, Dusseldorf, Germany). Extracted DNA was quantified using a NanoDrop 2000C (Thermo Scientific, MA, USA) and all samples were diluted to a concentration of 5 ng/µl for copy number determination.

### 6.2.6. Evaluation of gene copy numbers by real-time quantitative PCR assay

Genotyping of *CCL3L*, *CCL3La* and *CCL3Lb* genes was performed by means of real-time TaqMan assays as described by Shostavich-Koretskaya *et al.* (2009). Primers and probes used were as described by Townson *et al.* (2002) and Shostavich-Koretskaya *et al.* (2009) and have been designed to specifically detect and distinguish between the versions of *CCL3L*. Standard curves of Ct (threshold cycle) value against log[DNA, ng] were generated for each probe from duplicate reactions of six doubling dilutions of genomic DNA (25ng to 0.78ng) extracted from A431 and K562 cell lines (ATCC, Manassas, VA, USA), on each 96-well reaction plate (Applied Biosystems, Foster City, CA, USA). A431 cells are known to have exactly two copies of the *CCL3L* and *CCL3La* genes pdg (GONZALEZ *et al.* 2005; TOWNSON *et al.* 2002) and K562 cells have one *CCL3Lb* gene pdg (SHOSTAKOVICH-KORETSKAYA *et al.* 2009).

For each test sample, duplicate wells were set up on a plate for every probe being tested. The Ct value was determined and converted into template quantity using the standard curves specific to that plate. Copy number is the ratio of the template quantity for the gene of interest to the template quantity for the housekeeping gene (human *beta-globin* gene, *HBB*), multiplied by two (with the exception of *CCL3Lb*, where K562 DNA used as a standard contains only one copy pdg). To our knowledge, all humans carry two copies of the *HBB* gene. Copy number values were determined

for each test replicate and the mean of these values determined. If the standard deviation between replicates was >10% relative to the mean, copy number quantitation was repeated. The mean copy number value was rounded to the nearest integer and used to estimate copy number.

PCR was conducted in an Applied Biosystems 7500 Real-Time PCR system detecting emitted fluorescence as FAM (6-carboxyfluorescein, 6-FAM) or VIC from the probes during amplification. The TaqMan Universal Master Mix (Applied Biosystems, Foster City, CA, USA) was used for all reactions. The amount of input DNA per test PCR was 5ng.

#### *6.2.7. Copy number assay validation*

The importance of accurately assessing copy number has been highlighted when addressing the influence of *CCL3L1* copy number on HIV-1 susceptibility (BHATTACHARYA *et al.* 2009; FIELD *et al.* 2009; HE *et al.* 2009; URBAN *et al.* 2009). Determining the sum of gene copy numbers for *CCL3La* and *CCL3Lb* serves as an internal validation of the *CCL3L* assay (HE *et al.* 2009).

##### *6.2.7.1. CCL3Lb PCR*

As a further quality control to verify the real-time results, we designed primers to amplify the *CCL3Lb* gene to confirm the presence or absence of the *CCL3Lb* gene by means of conventional PCR (Expand High Fidelity PCR System, Roche, Mannheim, Germany). The *CCL3Lb* gene contains 2 exons homologous to exons 2 and 3 of the *CCL3L1* and *CCL3L3* genes (HIRASHIMA *et al.* 1992). The primers used to screen for *CCL3Lb* were designed such that the forward primer (5' CACACAAACATATGCACACATACA 3') binds just upstream to this region of homology and the reverse primer (5' CTCGGTCGTCACCAGACACAC 3') binds within exon 3, at a point where there are three GC mismatches between *CCL3L1* and *CCL3Lb* at its 3' end, resulting in a 937 bp amplicon.

##### *6.2.8. CCL3 quantification*

CCL3 was quantified in unstimulated and PHA-stimulated PBMCs by means of ELISA (DuoSet ELISA Development Systems: R&D Systems, Minneapolis, Minnesota, USA). Assay conditions were as recommended by the manufacturer. The minimal detectable levels were <10 pg/ml. A duplicate doubling serial dilution (500 to 7.8 pg/ml) of a CCL3 control of known concentration was used to generate a standard curve for each plate which was used for the calculation of chemokine concentration. Measurements for each sample were conducted in triplicate and the mean optical density (OD) was used to determine CCL3 concentrations from the standard curve. In instances where the OD was greater than that obtained for the 500 pg/ml standard (n=3), ELISAs were repeated with diluted sample. CCL3 production values were calculated as follows: CCL3

production of PHA stimulated PBMCs less CCL3 production of unstimulated PBMCs from the same individual.

Chemokine production assays and CCL3 quantification had previously been conducted on cord-blood mononuclear cells (CBMCs) isolated from infants born to HIV-1 positive women, similar to the method described above (MEDDOWS-TAYLOR *et al.* 2006).

#### 6.2.9. Statistical analysis

Differences in peripheral CCL3 chemokine levels and gene copy number were assessed using the nonparametric Mann Whitney U test. Correlations between chemokine levels were calculated by bivariate Spearman's rank coefficients. All statistical analyses were performed using PASW Statistics 18 software (SPSS Inc., Chicago, Illinois, USA).

Fisher exact tests were performed using the Simple Interactive Statistical Analysis software (UITENBROOK 1997) to test for significant difference between populations used in this study.

### 6.3. RESULTS

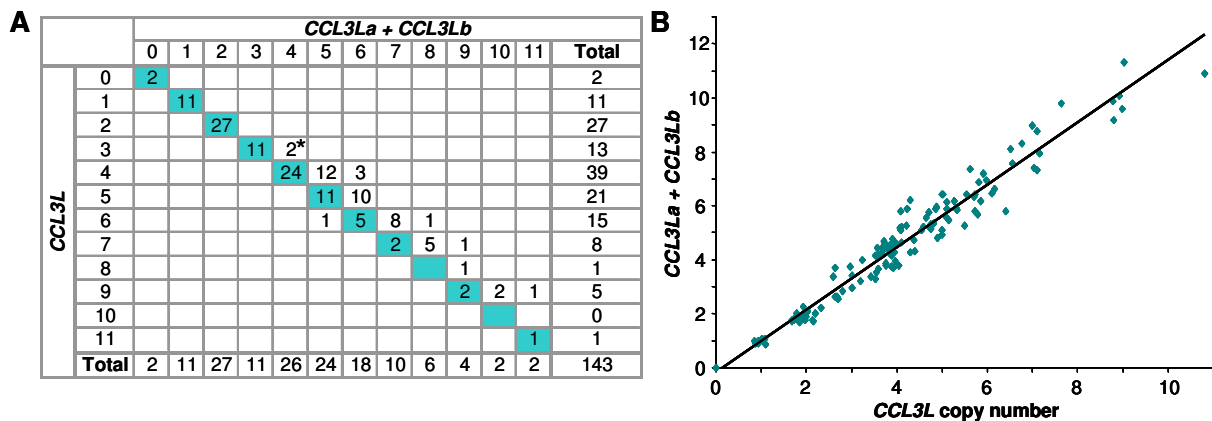
#### 6.3.1. Assay validation

##### 6.3.1.1. Sum of gene copy numbers

*CCL3L*, *CCL3La* and *CCL3Lb* gene copy number was determined for 143 healthy HIV-1 uninfected individuals, 89 SAA and 54 SAC individuals. The real time assay for *CCL3L* quantifies both *CCL3La* and *CCL3Lb*. Although *CCL3L* copy number is meant to be a measure of the sum of *CCL3La* and *CCL3Lb* copy numbers, in some individuals these values were not in agreement (Figure 6.1A). This occurred with higher gene copy numbers, i.e., *CCL3L* gene copy numbers  $\geq 3$ . In the individuals where there was discordance between the values, the '*CCL3La* + *CCL3Lb*' value tended to be larger than that of *CCL3L* (Figure 6.1A). There was at most a difference of two between values. An internal validation of the assay, using unrounded copy numbers, demonstrated satisfactory concordance between the assays (Figure 6.1B;  $y = 1.154x - 0.159$ ;  $R^2 = 0.958$ ;  $P < 0.0001$ ).

##### 6.3.1.2. CCL3Lb PCR

Sixty four of the individuals in the entire healthy HIV-1 uninfected cohort were shown to have a copy number of zero for the *CCL3Lb* gene by real time PCR. Although the conventional PCR cannot verify copy number values, there was 100% agreement between the PCR and the *CCL3Lb* real time assay in terms of presence or absence of the gene (results not shown).



**Figure 6.1.** Concordance between different real time assays to assess gene copy number, *CCL3L* and *CCL3La + CCL3Lb* illustrated using both rounded (**A**) and unrounded values (**B**). Numbers within shaded blocks in (**A**) indicate the number of individuals where  $CCL3L = CCL3La + CCL3Lb$ , while numbers above or below the shaded boxes represent the instances where these values were discordant (eg. two individuals had a copy number of '3' for *CCL3L* and '4' for the sum of *CCL3La + CCL3Lb* (\*)). The regression line for the scatter plot relating unrounded average *CCL3L* values to the unrounded sum of *CCL3La* and *CCL3Lb* is indicated (B,  $y = 1.154x - 0.159$ ;  $R^2 = 0.958$ ;  $P < 0.0001$ ). N=143.

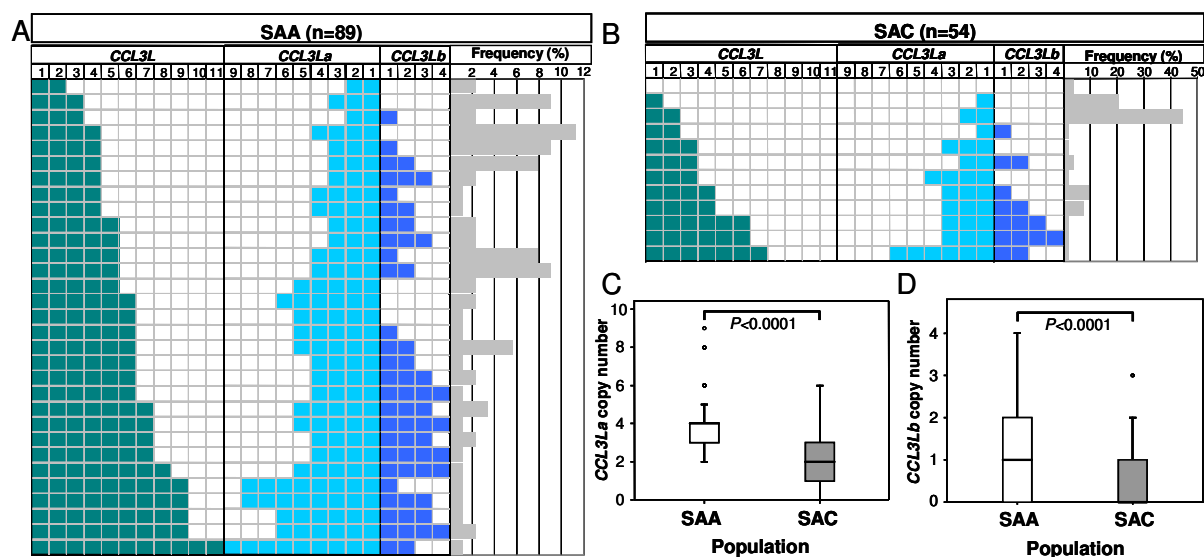
### 6.3.2. Copy number variation in the two population groups

There exists significant inter-individual variation in gene copy numbers. The number of total *CCL3L* copies ranged from two to eleven in SAA, with a median of five, and ranged from zero to seven in SAC, with a median of two. These values are comparable to those previously observed in these two populations (GONZALEZ *et al.* 2005; SHALEKOFF *et al.* 2008). The median copy number for *CCL3La* and *CCL3Lb* genes were calculated to be four and one in SAA, and two and zero in SAC, respectively.

The total *CCL3L* copy number distribution within the two test populations as well as the frequency at which different patterns of *CCL3La* and *CCL3Lb* occur in the two populations are shown in Figure 6.2A and B. In the SAA population, 46.1% (41/89) of individuals had gene copy numbers below the *CCL3L* population median and 37.1% (33/89) had copy numbers below the *CCL3La* population median, whereas in SAC, only 24.1% (13/54) and 25.9% (14/54) were below the population median for *CCL3L* and *CCL3La*, respectively. The distribution of *CCL3L*, *CCL3La* and *CCL3Lb* gene copy number differed significantly ( $P < 0.0001$ ) between the SAA and SAC populations (Figure 6.2C and D).

It is interesting to note that, although *CCL3Lb* is present in the SAC population (15/54; 27.8%), the median copy number for *CCL3Lb* is zero. The *CCL3Lb* gene tended to be present in individuals

with a *CCL3La* gene copy number of at least 2, however, one SAC individual with a *CCL3La* copy of 1 was found to have the *CCL3Lb* gene. *CCL3Lb* copy number correlated positively with *CCL3La* copy number ( $P < 0.0001$ ,  $R = 0.497$ , SAA and SAC combined).



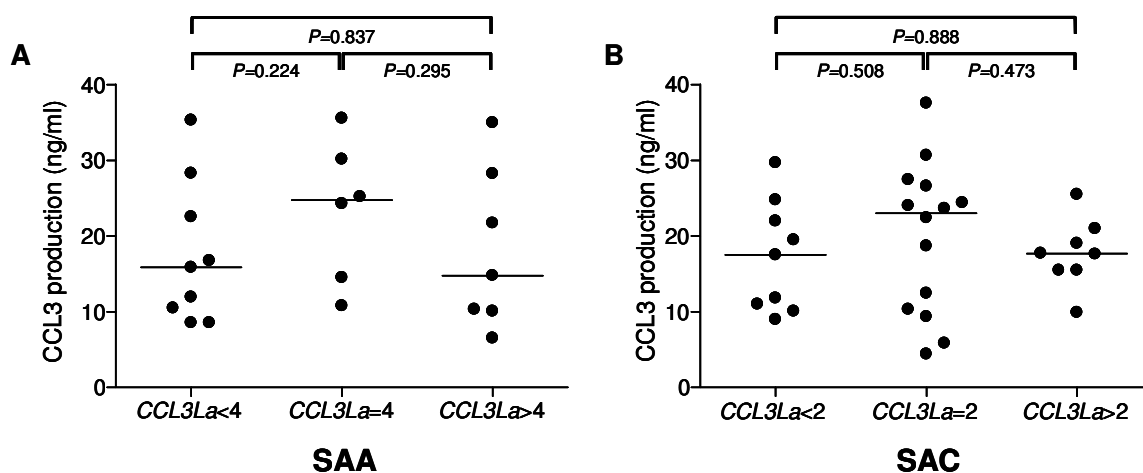
**Figure 6.2.** Differences in *CCL3L*, *CCL3La* and *CCL3Lb* gene copy number distribution between healthy HIV-1 uninfected individuals from the two study populations, South African African (SAA) ( $n=89$ ) and South African Caucasian (SAC) ( $n=54$ ). **(A)** and **(B)**: Each shaded block represents a single copy of a gene. Frequencies at which individual copy number patterns occur within a population are indicated. Box-whisker plots depicting the median (horizontal black line), 25<sup>th</sup> and 75<sup>th</sup> percentiles (margins of the box) and the 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) illustrate *CCL3La* **(C)** and *CCL3Lb* **(D)** gene copy number pdg differences between SAA and SAC individuals. Outliers are indicated with ( $\circ$ ). Significant differences between groups are indicated.

### 6.3.3. *CCL3* production does not correlate with *CCL3La* copy number

We measured the production of CCL3 in the supernatants of unstimulated and PHA-stimulated PBMCs obtained from 22 SAA and 31 SAC individuals by ELISA. The antibody used in the ELISA recognizes both CCL3 and CCL3L proteins. To our knowledge, these two proteins cannot be immunologically distinguished by ELISA. Therefore, comparisons are based on total CCL3 production.

We investigated whether individuals with higher *CCL3L* and *CCL3La* copy number produce CCL3 at higher levels compared to individuals with lower copy number within the same population. No correlation between both *CCL3L* and *CCL3La* copy number and CCL3 production was observed in either population, or when stratified according to less than, equal to, or greater than, the population

median (Figure 6.3). Furthermore, the median production of corresponding groups stratified around the *CCL3La* population specific copy number mean did not differ between the two population groups, i.e., SAA individuals with *CCL3La*<4 vs SAC individuals with *CCL3La*<2; SAA individuals with *CCL3La*=4 vs SAC individuals with *CCL3La*=2 and SAA individuals with *CCL3La*>4 vs SAC individuals with *CCL3La*>2 demonstrated similar CCL3 production ( $P=0.863$ ,  $P=0.387$  and  $P=0.867$ , respectively). Interestingly, although not statistically significant, both populations revealed a pattern of higher production at the population specific median copy number (i.e., *CCL3La*=4 for SAA and *CCL3La*=2 for SAC, Figure 6.3).



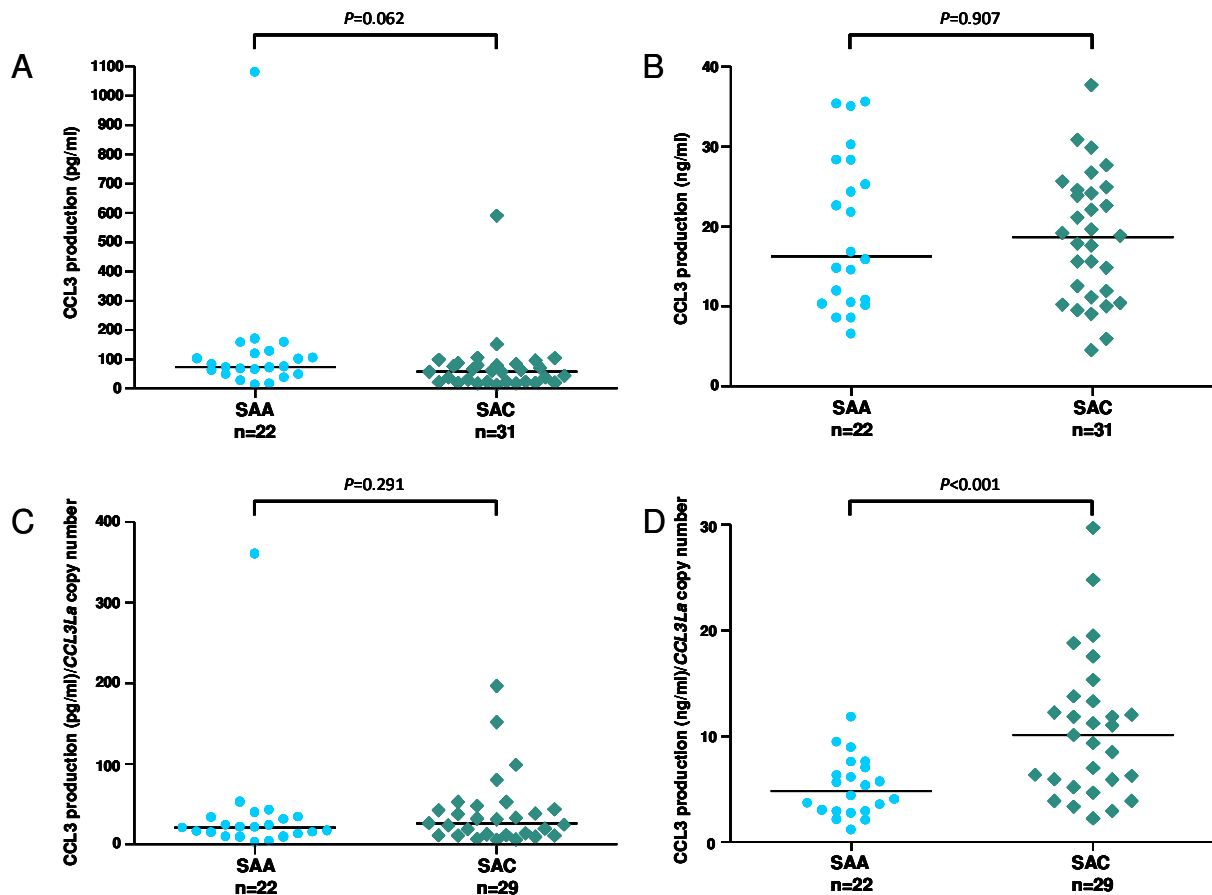
**Figure 6.3.** Phytohaemagglutinin-induced CCL3 production of South African African (SAA) (A) and South African Caucasian (SAC) (B) individuals stratified around the population specific *CCL3La* copy number median (four and two, respectively).  $P$  values are indicated.

#### 6.3.4. CCL3 production per functional *CCL3La* copy is higher in SAC individuals

We next compared CCL3 production between the two study populations (who have significantly different *CCL3L*, *CCL3La* and *CCL3Lb* copy numbers). In the absence of PHA stimulation, SAA individuals tended to have marginally higher CCL3 production compared to SAC individuals with a median production level of 72.5 and 57.5 pg/ml, respectively ( $P=0.062$ , Figure 6.4A). If the two outliers were removed from the analysis, this trend was maintained ( $P=0.060$ ). Upon PHA stimulation, production of CCL3 did not differ between the two population groups ( $P=0.907$ , Figure 6.4B).

To control for differences in functional gene copy number (*CCL3La*) between individuals, we next examined CCL3 production per gene copy. Given that *CCL3Lb* does not produce CCL3, *CCL3La* copy number was used as the divisor in determining CCL3 production per gene copy. Unstimulated

PBMC production per copy number did not differ between the study population groups ( $P=0.291$ , Figure 6.4C). Not unexpectedly, PHA-stimulated PBMCs from SAC individuals, who did not differ in stimulated CCL3 levels compared to SAA individuals (Figure 6.4B), produced significantly higher levels of CCL3 per *CCL3La* copy ( $P=0.0007$ , Figure 6.4D).



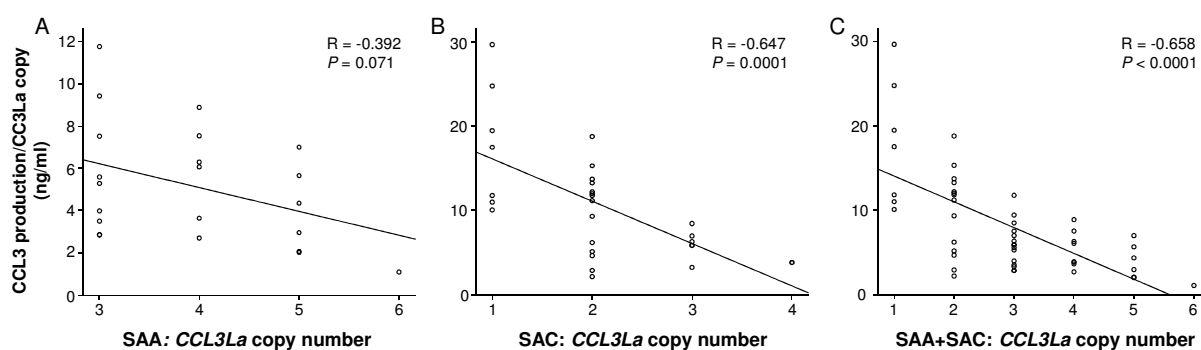
**Figure 6.4.** CCL3 production in South African African (SAA) and Caucasian (SAC) individuals as measured in CCL3 production assays. Overall CCL3 production is shown for unstimulated (A) and PHA stimulated (B) isolated PBMCs. Production per *CCL3La* copy number for unstimulated and PHA stimulated isolated PBMCs are shown in (C) and (D), respectively.  $P$  values and numbers (n) in each group are indicated.

Similar to above, when  $\log_{10}$ -transformed unstimulated, PHA-stimulated and unstimulated CCL3 production per *CCL3La* copy from the two populations were compared using t-tests, there was no significant difference between the two populations (Table D.1). However,  $\log_{10}$ -transformed PHA-induced production per *CCL3La* copy was significantly higher in SAC compared to SAA individuals (Table D.1,  $P=0.0002$ ).

Two SAC individuals had a *CCL3La* copy number of zero and thus production per copy number for these individuals could not be determined. Despite the absence of *CCL3La* genes, production was

markedly increased in these individuals: 57.53 pg/ml (unstimulated) to 22.03 ng/ml (PHA-stimulated) in one individual; and 31.09 pg/ml (unstimulated) to 9.02 ng/ml (PHA-stimulated) in the second individual. The SAC population unstimulated and PHA-stimulated CCL3 production medians were 57.53 pg/ml and 18.77 ng/ml, respectively. This highlights the importance of the unaccounted for contribution of *CCL3* to overall CCL3 production levels, a factor which may mask production levels based on copy number of *CCL3La*.

If the contribution to production by each copy of a gene were similar, one would expect the CCL3 production per *CCL3La* gene copy to yield a constant value irrespective of variations in *CCL3La* copy numbers. However, a significant negative correlation was observed between CCL3 production per *CCL3La* copy and *CCL3La* copy number, i.e., individuals with higher *CCL3La* copy number demonstrated lower production per copy (Figure 6.5).

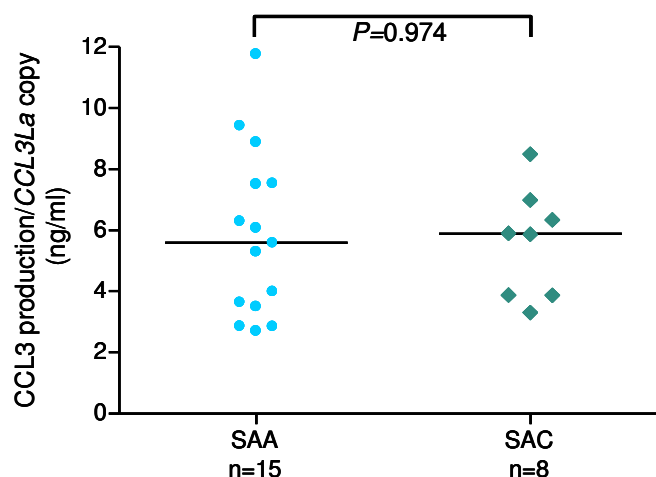


**Figure 6.5.** Correlation between calculated CCL3 production per *CCL3La* copy and *CCL3La* copy number for the two populations, South African African (SAA) and Caucasian (SAC), examined independently (**A** and **B**) or combined (**C**). *P* and *R* values are indicated

### 6.3.5. *CCL3La* copies do not differ functionally between SAA and SAC individuals

As reported above, SAA and SAC individuals differ in the number of *CCL3b* genes, with a median of two and zero *CCL3b* genes, respectively. A subset of individuals from each population group with comparable *CCL3L* genotypes was selected and their CCL3 production per *CCL3La* copy was compared. The subset consisted of individuals with a *CCL3La* copy number of three or four, matched for both the number of individuals with the *CCL3b* gene (Fisher's exact  $P=0.466$ ) and median *CCL3b* copy number (Mann-Whitney  $P=0.454$ ). No CCL3 production difference was observed between the two study groups when individuals matched for *CCL3La* copy number, *CCL3b* presence and *CCL3b* dose ( $P=0.974$ ) were compared (Figure 6.6). This suggests that the contribution to production by each functional *CCL3La* gene does not differ between SAA and SAC individuals when one compares individuals with similar *CCL3L* genotypes, thus correcting for the unknown influence of *CCL3b*.

Since we found *CCL3La* copy number to correlate positively with *CCL3Lb* copy number, it is possible that the *CCL3Lb* gene may function to downregulate CCL3L protein production which will reflect as a decrease on CCL3 production per functional *CCL3La* copy number. Therefore, to control for the possible contribution of *CCL3Lb* we compared CCL3 production per *CCL3La* copy number within a subset of SAA individuals who did not possess *CCL3Lb*. Despite the absence of *CCL3Lb*, individuals with *CCL3La* copy number of one (n=6) had significantly higher CCL3 production per *CCL3La* copy compared to those with a *CCL3La* copy number of two (n=13) ( $P=0.039$ ). This would indicate that *CCL3Lb* is not responsible for the observed negative correlation between CCL3 production per functional copy and *CCL3La* copy number (Figure 6.5).



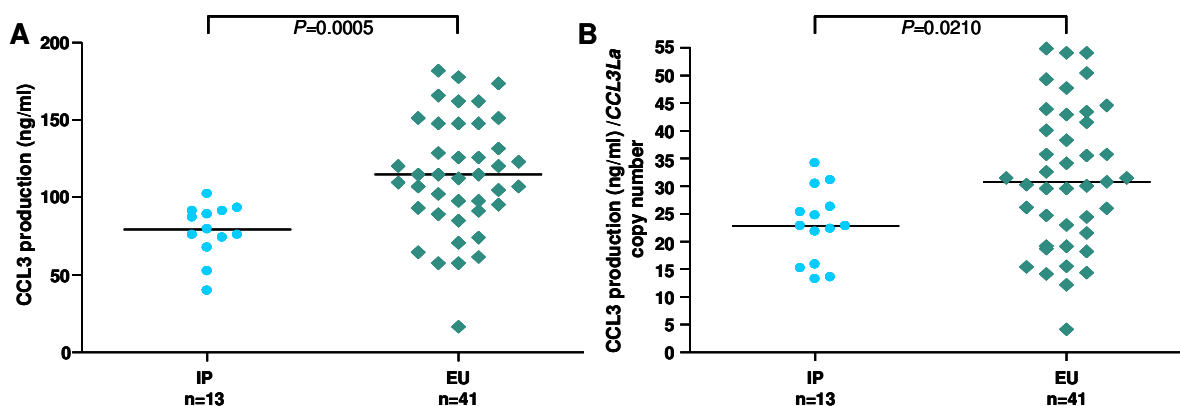
**Figure 6.6.** Comparison of CCL3 production per *CCL3La* copy number between South African African (SAA) and Caucasian (SAC) individuals with comparable *CCL3La* and *CCL3Lb* copy numbers. A subset of individuals matched for *CCL3La* copy number, *CCL3Lb* presence and dose, was selected from both population groups. The  $P$  value and number of individuals in each group are indicated.

### 6.3.6. *CCL3La*, *CCL3Lb* copy number variation and CCL3 production in the context of mother to child transmission

We previously demonstrated that HIV-1 exposed uninfected (EU) infants had significantly higher production compared to infants who acquired infection through the intrapartum (IP) route, which was attributed to differences in *CCL3L* copy number (MEDDOWS-TAYLOR *et al.* 2006). However, at equivalent *CCL3L* copy numbers, EU infants produced CCL3 at higher levels than IP infants which led to the hypothesis that the latter group of infants may harbour more non-functional copies of this gene. We therefore asked whether the lower CCL3 production in IP infants (MEDDOWS-TAYLOR *et al.* 2006) was due to a higher prevalence of *CCL3Lb*. Thirteen IP and 41 EU SAA infants from the same cohort (MEDDOWS-TAYLOR *et al.* 2006) were genotyped for *CCL3L*, *CCL3La* and *CCL3Lb*

gene copy numbers. These infants had previously only been genotyped for *CCL3L* since the genotyping assay used at that time did not distinguish between *CCL3La* and *CCL3Lb*.

Both groups had a median copy number of 4 and 2 for *CCL3L* and *CCL3Lb* genes, respectively. Although the median *CCL3La* copy number differed between the two groups, EU infants had a median of 4 while IP infants had a median of 3, this was not statistically significant ( $P=0.473$ ). The number of infants positive for *CCL3Lb* gene/s did not differ between EU and IP groups ( $P=0.750$ ). The copy number range differed between the two groups (EU: 0-7 *CCL3Lb* copies; IP: 0-3 *CCL3Lb* copies), however, there was no statistical difference in *CCL3Lb* copy number between EU and IP infants ( $P=0.646$ ), and in fact, as expected, the EU infants had a broader range of *CCL3Lb* copy number compared to the IP infants consistent with the higher *CCL3La* copy number within this group.



**Figure 6.7.** CCL3 production of phytohaemagglutinin stimulated cord blood mononuclear cells (CBMCs) isolated from HIV-1 intrapartum infected (IP) and exposed uninfected (EU) infants. Overall production is shown in (A) and production normalized for differences in *CCL3La* copy number is shown in (B).  $P$  values and number of infants in each group are indicated.

As previously seen, CCL3 levels differed significantly between EU and IP infants ( $P<0.001$ , Figure 6.7A). Since the *CCL3La* median copy numbers were different for the two groups, with EUs having a higher median than IPs, we compared CCL3 production per *CCL3La* copy number between the two groups (Figure 6.7B) to assess whether this was potentially influencing the observed CCL3 production difference between the two groups. CCL3 production remained significantly different between EU and IP upon correction for copy number ( $P=0.021$ , Figure 6.7B), indicating that the difference in production between the two groups cannot be attributed to differences in dose of functional gene copy number. Similarly,  $\log_{10}$ -transformed CCL3 production data was significantly different between the two groups (Table D.2).

## 6.4. DISCUSSION

Copy number variation, the deletion and duplication of DNA strands ranging from a few hundred to several million base pairs in length, is increasingly being found to play a role in numerous diseases (Reviewed in WAIN *et al.* 2009). Variation in *CCL3L* copy number has been associated with a number of HIV-related outcomes, including vertical and horizontal transmission of HIV-1 (GONZALEZ *et al.* 2005; KUHN *et al.* 2007), plasma viral load (DOLAN *et al.* 2007; SHALEKOFF *et al.* 2008), progression to AIDS (DOLAN *et al.* 2007) and response to treatment of HIV-1/AIDS (AHUJA *et al.* 2008). These findings have however been disputed (BHATTACHARYA *et al.* 2009; URBAN *et al.* 2009). This controversy may be partly due to differences in *CCL3L* gene quantifying methods and differentiating these genes from *CCL3* and from *CCL3Lb* (Reviewed in SHRESTHA *et al.* 2009). Subsequently, a meta-analysis of nine published studies indicated that *CCL3L* copy number variability, relative to the population-specific median, is associated with susceptibility to HIV-1 infection, with higher copy number being associated with lower risk of HIV-1 acquisition, and vice versa (LIU *et al.* 2010). Given the support for an important role in several, although not all, studies that variation in *CCL3L* copy number plays an important role in HIV-1 infection and that the protective role of *CCL3L* copy number variation is relative to the population specific copy number median, we investigated whether populations known to have significantly different *CCL3L* copy numbers differ in their production of CCL3 protein.

Individuals infected with HIV-1 have been shown to have increased levels of CCL3 (ULLUM *et al.* 1998). Similarly, other infectious (eg. *Mycobacterium leprae* infections (leprosy) and influenza A infection), and non-infectious diseases (eg. Gaucher disease) are associated with altered CCL3 levels in comparison to control, healthy individuals (MENDONCA *et al.* 2007; VAN BREEMEN *et al.* 2007). In addition, the maturation state of the immune system can also have an impact on levels of CCL3 with adults producing lower levels of CCL3 than infants (TIEMESSEN and KUHN 2007). This highlights the complexity of the regulation of chemokine production. Thus, understanding the relationship between *CCL3La* gene copy number and levels of CCL3 would be confounded by existing HIV-1, as well as other infections. For this reason, in assessing whether CCL3 production differs between the study populations with significantly different *CCL3L* copy number, we studied this relationship in healthy HIV-1 uninfected individuals.

Although protein expression upon lipopolysaccharide (LPS) stimulation of isolated monocytes has been shown to correlate with *CCL3L* copy number (TOWNSON *et al.* 2002), a recent report failed to demonstrate the same relationship (CARPENTER *et al.* 2012). In this study, we too failed to show an association between CCL3 production upon stimulation and *CCL3L* copy number. All three studies differ experimentally making it difficult to directly compare results between studies; stimulation times are different between all three studies and, furthermore, we stimulated isolated PBMCs (not

monocytes) with PHA, whereas the other two studies used LPS. We have previously demonstrated that PHA stimulation of CBMCs isolated from EU infants with higher *CCL3L* copy numbers ( $\geq 4$ ) produce CCL3 at higher levels than CBMCs from EU infants with lower *CCL3L* copy number ( $< 4$ ) (MEDDOWS-TAYLOR *et al.* 2006). This relationship, however, was absent in IP infants whose CBMCs were stimulated in the same way (MEDDOWS-TAYLOR *et al.* 2006). Thus, this highlights that in trying to ascertain the true relationship between variable *CCL3L* copy number and CCL3 production, results can be influenced both by cohort selection and experimental parameters.

Given the differences in *CCL3L* copy number in the two study populations, and that the associations with *CCL3L* copy number and HIV-associated outcomes have been observed in relation to the population specific median, we investigated the influence of copy number on CCL3 production in two ethnically divergent South African populations. Furthermore, in assessing *CCL3L* copy number, we sought to distinguish between the functional *CCL3La* genes coding for the CCL3L chemokine and the *CCL3Lb* genes which do not code for a chemokine but which are not always excluded in assays used for quantifying *CCL3L* gene copy number. The median *CCL3La* copy number is four and two for SAA and SAC individuals, respectively. Thus, one might expect overall higher CCL3 production in SAA individuals compared to SAC individuals. However, no significant differences in CCL3 production were observed between the two groups. When CCL3 production was stratified around the population median, no difference in production was noted between subgroups within each population (i.e., lower than median, median and higher than median) and more importantly no difference was observed between SAA and SAC individuals in corresponding groups, with CCL3 production in both populations being highest around the copy number median (although this lacked statistical significance). This result suggests that there may be an optimal functional amount of CCL3L, centered around the population specific median, that is the same for the two population groups, irrespective of whether the *CCL3La* median is four or two.

We did, however, observe significantly higher CCL3 production per functional *CCL3La* copy in SAC compared to SAA individuals. When individuals with equivalent *CCL3La* and *CCL3Lb* copy numbers were compared, no differences in CCL3 production per *CCL3La* copy were noted, indicating the differences noted between SAA and SAC individuals are likely due to the differences in gene copy numbers between the two populations (i.e., a greater copy number (denominator) will result in lower CCL3 production if looking at production per copy). This result suggests that populations with higher *CCL3La* copy numbers may have developed mechanisms to regulate CCL3 production, although this begs the question as to why evolution would maintain more copies if functionally these do not result in the production of more protein. It could also point to the concept of 'not all copies are created equal' and possible differences in promoter activity between copies (PAXIMADIS *et al.* 2009; TIEMESSEN and KUHN 2007).

Another possible mechanism of CCL3 production regulation could be the *CCL3Lb* gene. *CCL3Lb*, originally thought to be a 5'-truncated pseudogene, has been demonstrated to generate two alternatively spliced mRNA transcripts containing chemokine-like domains but which are not predicted to code for classical chemokines (HIRASHIMA *et al.* 1992; SHOSTAKOVICH-KORETSKAYA *et al.* 2009). The function of these transcripts is as yet unknown but it has been suggested that 'this gene may influence HIV-AIDS pathogenesis by as yet unidentified means' (SHOSTAKOVICH-KORETSKAYA *et al.* 2009). *CCL3Lb* copy number was found to correlate positively with *CCL3La* copy number. Thus it is conceivable that if *CCL3Lb* had a regulatory function upon CCL3 production, this would be more apparent in individuals with higher *CCL3La* copy numbers. However, in the absence of any *CCL3Lb* genes, SAC individuals with a *CCL3La* copy number of 2 still demonstrated a significantly lower production per *CCL3La* copy compared to SAC individuals with a *CCL3La* copy number of 1.

We have previously demonstrated that at equivalent *CCL3L* copy numbers, infants who acquired HIV-1 infection still produced less CCL3 relative to their exposed uninfected counterparts (MEDDOWS-TAYLOR *et al.* 2006). This led to the hypothesis that at equivalent *CCL3L* copy numbers, infants with the lower CCL3 production had greater numbers of non-functional genes, i.e., *CCL3Lb*, or that this CCL3 production difference could be attributed to differences in *CCL3*, the other gene that contributes to CCL3 production (MEDDOWS-TAYLOR *et al.* 2006; PAXIMADIS *et al.* 2009). Despite the vast differences in CCL3 production between IP and EU infants, these two groups did not differ in the presence or dose of the *CCL3Lb* gene. This suggests that there may be further genetic differences, within either the allelic *CCL3* or any of the non-allelic *CCL3La* genes, contributing towards CCL3 production differences.

Although a lack of association between *CCL3L* copy number and CCL3 protein expression upon monocyte stimulation was recently reported, the *CCL3L* mRNA as well as the ratio of *CCL3L:CCL3* mRNA was shown to correlate significantly with copy number in the same study (CARPENTER *et al.* 2012). Since it was not possible to distinguish between CCL3 and CCL3L protein, for this study it was necessary to assume that these two genes would contribute equally to CCL3 production. In two separate studies, both monocytes and PBMCs, derived from control seronegative individuals, stimulated with LPS and PHA, respectively, have been shown to produce *CCL3* mRNA transcripts at much higher levels than *CCL3L* transcripts, indicating that the same may be reflected in the respective protein, i.e., more CCL3 protein produced compared to CCL3L (CARPENTER *et al.* 2012; PILOTTI *et al.* 2007). Carpenter *et al.* (2012) hypothesized that the variation observed with CCL3 is thus potentially more biologically consequential than *CCL3L* copy number variation (CARPENTER *et al.* 2012). However, given that *in vitro*, the CCL3L chemokine appears to be the most potent agonist of CCR5 and inhibitor of R5 HIV-1 strains in comparison to the allelic CCL3 chemokine and

that a much lower concentration of CCL3L, compared to CCL3, is needed to illicit antiviral activity (AQUARO *et al.* 2001; MENTEN *et al.* 1999; NIBBS *et al.* 1999), it is likely that small changes in *CCL3L* mRNA may in fact have a large functional consequence despite the predominance of *CCL3* mRNA, i.e., “a little may go a long way”. Furthermore, Pilotti *et al.* (2007) made the observation that in an unstimulated state, PBMCs isolated from HIV-1 infected long term nonprogressing individuals produced *CCL3L1* mRNA at higher levels than that of *CCL3* (PILOTTI *et al.* 2007), in contrast to the study by Carpenter *et al.* (2012), thus suggesting that HIV-1 infected individuals may produce CCL3L protein at higher levels than CCL3. Expression of CCL3L mRNA may also be regulated or influenced by other factors, such as infection with human T-cell leukemia virus 2 (HTLV-2) (PILOTTI *et al.* 2007).

If *CCL3* is the major contributor to total CCL3 production, then the effect of differential *CCL3La* copy number on CCL3 production would be very difficult to measure. Without a means of distinguishing between CCL3 and CCL3L, it is not possible to elucidate the true effect that higher or lower *CCL3La* copy numbers will have on CCL3L production. The difference in CCL3 production in the two SAC individuals that lack *CCL3L* genes highlights the large interindividual variation in CCL3 production and thus the complexity in determining what influences production. It is important to keep in mind that genetic variations, such as promoter polymorphisms and/or haplotypes in both *CCL3* and *CCL3L* genes, may also play a role in determining differences in CCL3 protein production. A number of single nucleotide polymorphisms within *CCL3* have been reported to associate with HIV-1 infection outcomes (GONZALEZ *et al.* 2001; HU *et al.* 2012; MODI *et al.* 2006; SHRESTHA *et al.* 2006). The two population groups investigated in this study have also been shown to differ with respect to polymorphisms, intragenic haplotypes and indels found in the *CCL3* and *CCL3L* genes (PAXIMADIS *et al.* 2009). Furthermore, in the context of mother to child transmission, it has recently been shown that *CCL3* haplotypes play a role in infant susceptibility to HIV-1 infection (PAXIMADIS *et al.* 2012). Therefore, in addition to *CCL3L* copy number variation, the contribution made by *CCL3* in the context of HIV-1 infection cannot be ignored. The ability to distinguish between these two proteins would greatly aid in determining whether any genetic variations apart from copy number variation are influencing CCL3 production.

A limitation of our study is the relatively small samples sizes, particularly in the SAC population. Given the broad range of *CCL3L* copy number distribution, it is possible that some of the less frequent *CCL3L*, *CCL3La* and *CCL3Lb* copy number combinations were not represented within our populations. Sample size could also provide an explanation for the observed lack of association between *CCL3L/CCL3La* gene copy number and the amount of CCL3 produced upon stimulation. Furthermore, there is no correction for population ancestry. The SAA population is comprised of Bantu speaking individuals, however, the distribution of the different language/tribal groups found

in South Africa within our study group is unknown to us. Thus, it is possible that these groups are not proportionally represented. We found median *CCL3L* gene copy number to not differ significantly between individuals stratified according to the different language/tribal groups found in South Africa, i.e., Ndebele, Pedi, South Sotho, Swazi, Tsonga, Tswana, Venda, Xhosa and Zulu (S. Donninger, unpublished data), suggesting that our results are representative of diversity within South African populations but not necessarily of other African populations. Validations from independent populations are warranted to solidify the current findings.

In this study we have provided the first report of the *CCL3La* and *CCL3Lb* copy number distribution in South African populations. We have further demonstrated that despite the differences between populations with regards to the *CCL3La* copy number mean, there is no difference in CCL3 production between the two study populations and that in fact CCL3 production per functional copy number decreases with an increase in *CCL3La* copy number. This reflects the observed absence of correlation between functional gene copy number and CCL3 production. The contribution of *CCL3La* to protein production may be masked by that of *CCL3*. Furthermore, polymorphisms within *CCL3* itself, or other as yet undefined mechanisms, may be impacting on overall CCL3 production as would be suggested by the IP and EU infants having differential CCL3 protein production despite similar *CCL3La* and *CCL3Lb* copy number. In addition to recent literature, this study further supports or illustrates the complexity surrounding this gene and its protein expression.

---

## **CHAPTER 7**

**CCR5 and CCL3L genotypic and phenotypic differences between HIV-1 infected long term nonprogressors and healthy uninfected individuals**

---

## ABSTRACT

The outcome of HIV-1 infection and the course to AIDS is highly variable, with some individuals progressing to disease more rapidly than others. Both CCR5, the major HIV-1 coreceptor, and copy number variation of its ligand CCL3L, are major determinants in HIV-1 disease progression. Having previously established the extent of genotypic and phenotypic variation of these molecules among South African African (SAA) and South African Caucasian (SAC) individuals, we compared these population-specific findings to a small cohort of HIV-1 infected long term nonprogressing (LTNP) individuals (n=10 SAA and n=4 SAC). No new polymorphisms were detected in the *CCR5* gene. Furthermore, polymorphisms within the *CCR5* open reading frame, including *CCR5Δ32*, were absent in LTNP individuals. The HHA haplotype frequency was significantly higher in SAC LTNP individuals compared to SAC control individuals ( $P=0.010$ ). CCR5 density on CD4+ T cells and monocytes was significantly lower in SAA LTNP individuals compared to SAA controls ( $P=0.025$  and  $P=0.022$ , respectively) with a trend towards a similar relationship in CD8+ T cells ( $P=0.058$ ). However, the proportions of CCR5-expressing CD8+ T cells were elevated in SAA LTNP individuals compared to SAA controls ( $P=0.043$ ). Although the copy number distribution of *CCL3L*, *CCL3La* and *CCL3Lb* genes did not differ between control and LTNP individuals, PHA-induced CCL3 production was significantly lower in LTNP compared to control individuals ( $P=0.004$ ). Peripheral blood mononuclear cells from LTNP individuals with low viral loads (<400 copies/ml) produced CCL3 at lower levels than those from individuals with higher viral loads, irrespective of whether or not the cells were stimulated ( $P=0.005$  and  $P=0.035$ , respectively). It is possible that observed differences in CCL3 production may partially be explained by differences in cell subset distribution between study groups: SAA LTNP individuals had higher proportions of CD8+ T cells ( $P<0.0001$ ) and lower proportions of natural killer cells (CD56+,  $P=0.002$ ) compared to control SAA individuals. This pilot study demonstrates that two major determinants of HIV-1 disease progression, CCR5 and CCL3, are both expressed at lower levels in LTNPs individuals compared to healthy uninfected controls and has identified *CCR5* haplotypes which are potentially associated with disease progression. These associations will be assessed in future studies with larger numbers of HIV-1 infected LTNP and progressing individuals.

## 7.1. INTRODUCTION

HIV-1 infection is characterized by a broad spectrum of disease progression rates, an observation which has led to various definitions of clinical progression ranging from rapid progressors to elite controllers. Following HIV-1 infection, individuals whose CD4+ T cell counts rapidly decline and progress to AIDS within a period of three to five years are termed rapid progressors (Reviewed in MIKHAIL *et al.* 2003). Furthermore, this group of individuals maintain high plasma viral loads which

do not decline after primary infection (MIKHAIL *et al.* 2003). Slow progressors (SP) or long-term survivors are defined as having lower CD4+ T cell counts and higher viral loads than HIV-1 long term nonprogressors (LTNPs) (POROPATICH and SULLIVAN 2011). Individuals referred to as LTNPs are a group of HIV-1 seropositive individuals who have been infected with HIV-1 for several years and in the absence of antiretroviral therapy have maintained normal CD4+ T cell counts (mostly >500 cells/ml), control viral replication and maintain healthy clinical parameters, in some cases for more than two decades (MIKHAIL *et al.* 2003). A small subset of LTNPs, termed Elite Controllers (ECs), maintain their plasma HIV RNA levels below 50 copies/ml for an unspecified period of time (PEREYRA *et al.* 2008; POROPATICH and SULLIVAN 2011).

It is important to note that although the majority of EC and LTNP individuals maintain their CD4+ T cell counts within normal levels for many years, with time several will experience immunodeficiency and opportunistic infections (HUNT 2009). Furthermore, there is no standardized definition of the various clinical progression profiles and different studies vary in their definition of these profiles. For example, some research groups have defined clinical progression, i.e., chronic progressors, controllers, and elite controllers, solely on viral loads whereas others have largely based clinical progression on the length of survival without AIDS in the absence of antiretroviral medication. This, at times, makes the interpretation of data and comparisons between studies difficult. Nonetheless, the study of individuals who are able to control their HIV-1 infection has yielded important information regarding correlates of protection from disease progression and in addition, may provide important information for the development of novel vaccine therapies and/or immunological treatment strategies.

HIV-1-infected progressors and nonprogressors have been described to differ in their host gene complement, viral strains as well as their immunological responses (POROPATICH and SULLIVAN 2011). Where host genetics has been experimentally associated with the control of HIV-1 progression, this has mainly implicated a role for allelic variants of human leukocyte antigens (HLA) and genetic polymorphisms that interfere with the binding of virus to target cells through impaired cell surface receptor expression or through an increase in the production of competitive inhibitors. The role of CCR5 coreceptor density in an individual's susceptibility to HIV-1 has been well established. The *CCR5Δ32* allele has been reported as overrepresented within HIV-1 infected groups of patients progressing to disease at slower than normal rates (COHEN *et al.* 1998; SMITH *et al.* 1997; STEWART *et al.* 1997). The observation that high CCR5 expression levels on CD4+ T cells are associated with high viral loads and accelerated disease progression further highlights the importance of CCR5 in determining the rate of disease progression in an HIV-1 infected individual (REYNES *et al.* 2001; REYNES *et al.* 2000). However, the lack of association with the *CCR5Δ32* allele with disease progression in other study groups (SCHINKEL *et al.* 1999; WILKINSON *et al.* 1998)

demonstrates that variation in CCR5 alone is not the only determining factor. The  $\beta$ -chemokines, CCL3, CCL4 and CCL5 (RANTES), have also been identified as HIV-suppressive factors (COCCHI *et al.* 1995). When bound to CCR5, the  $\beta$ -chemokines induce internalization of the receptor thus making it unavailable for HIV-1 infection. Therefore both CCR5 and CCL3 affect viral entry and replication. Thus, we undertook a pilot study to examine the potential role of differences in *CCR5* genotype, CCR5 expression, *CCL3* genotypes and CCL3 production levels on HIV-1 disease progression by examining these factors in a small group of HIV-1-infected long term nonprogressing individuals and comparing these to data obtained for healthy HIV-1 uninfected individuals.

## 7.2. MATERIALS AND METHODS

### 7.2.1. Cohort description

#### 7.2.1.1. Long term nonprogressors (LTNP)

The LTNP individuals used in this study were from a cohort of HIV-1-infected individuals with long term follow up that have been prospectively recruited. Criteria for LTNP selection were: absence of antiretroviral treatment since diagnosis and consistently high CD4+ T cell counts and/or low viral loads. Although some of the individuals selected as part of our LTNP cohort have been infected with HIV-1 for less than 10 years, these individuals are being monitored for disease progression.

This study cohort comprised of ten South African African (SAA) and four South African Caucasian (SAC) HIV-1 infected individuals. At the time at which the experiments were conducted, the characteristics of the cohort were as described in Table 7.1. The SAA LTNP cohort had a median age of 40.5 years (range: 33-54 years) and comprised four (40%) females and six (60%) males. The SAC LTNP cohort had a median age of 47 years (range: 33-67 years) and comprised three (75%) females and one (25%) male. The SAA LTNP individuals differed significantly in age from SAA healthy individuals (median age of 33.5 years) ( $P=0.023$ ), however, there was no difference in the male:female ratio between the two groups ( $P=0.267$ ).

Among the group of LTNPs, two SAA individuals (LTNP6 and LTNP7) and one SAC individual (LTNP13) met the criteria of elite controllers (i.e., patients with plasma HIV RNA levels of <50 copies/ml (PEREYRA *et al.* 2008)). It is important to note that ECs do differ from LTNPs both in CD4+ T cell counts (often higher in LTNPs) and viral loads (lower in ECs), thus grouping EC with LTNPs may sometimes mask certain mechanisms of control (POROPATICH and SULLIVAN 2011). However, the favourable clinical outcome of the two groups indicates that they share some mechanisms of control.

**Table 7.1.** Characteristics of HIV-1-infected long term nonprogressing and HIV-1-infected progressing cohorts\*

	LTNP		HIV-1 infected progressors
	SAA (n=10)	SAC (n=4)	(n=10)
	Median (range)	Median (range)	Median (range)
Age (years)	40.5 (33 - 54)	47(33 - 67)	28 (19 - 36)
Male:Female	4:6	3:1	0:10
Viral load (RNA copies/ml)	777 (<40 - 23 300)	399 (<40 - 5 420)	128 500 (1 590 - 491 000)
Log(viral load)	2.83 (1.59 - 4.37)	2.6 (1.59 - 3.73)	5.1 (3.20 - 5.69)
CD4 count (cells/ $\mu$ l)	665 (327 - > 2 000)	603.5 (327 - 877)	195 (40 - 295)
Time since diagnosis (years)	9 (4 – 13)	5.5 (5 – 18)	unknown

\* The characteristics of individual participants are shown in Table E.1 and Table E.2 within the Appendices

Subsequent to performing the experiments detailed below, two individuals forming part of our LTNP cohort, one SAA (LTNP2) and one SAC (LTNP12) individual, were started on antiretroviral treatment. These two individuals had been infected for seven and six years, respectively, prior to starting treatment. The current median number of years the remaining individuals have been infected with HIV-1 without the need for antiretroviral treatment is 11 years (range: 6-15) for SAA and 8 years (range: 7-20) for SAC individuals.

#### 7.2.1.2. HIV-1 infected progressors

For a preliminary investigation into whether there are differences in the frequencies of CCR5 haplotype or genotype between HIV-1 infected individuals who progress to AIDS and LTNPs, ten SAA individuals with high viral loads and low CD4+ T cells counts were selected from an existing separate cohort, to compare with the group of SAA LTNP individuals described above. The characteristics of the cohort are outlined in Table 7.1. It is important to note that what we have included as a “progressor” group is simply a selection of individuals from a cross-sectional chronic infection cohort with poor clinical features. They were included for preliminary comparison purposes only. However, in future studies as the LTNP cohort is expanded, rapid progressing individuals with known time of infection will also be recruited for comparison.

#### 7.2.2. Plasma viraemia quantification and CD4 T cell determination

HIV-1 RNA levels were quantified using one of two methods: (i) the COBAS® AmpliPrep/COBAS® TaqMan® HIV-1 test v2.0 (Roche Diagnostic Systems, Indianapolis, IN, USA) with a lower detection limit of 20 HIV-1 RNA copies/ml or (ii) the Roche Amplicor RNA Monitor Assay (Roche) with a lower detection limit of 400 HIV-1 RNA copies/ml. CD4 T cell counts were determined using the commercially available FACSCount System (Becton Dickinson, Franklin Lakes, NJ, USA).

### 7.2.3. Genotyping

The full length *CCR5* gene sequence (~9.2 kb) was determined for the 14 LTNP individuals as described previously (PICTON *et al.* 2010, Chapter 2). To genotype HIV-1 infected progressors, a smaller portion of the *CCR5* gene, i.e., the *CCR5* Promoter 1 region (nucleotide position -3384 to -1196), was sequenced as described previously (PICTON *et al.* 2010) to allow for genotyping according to the haplotypes described by Gonzalez *et al.* (1999).

### 7.2.4. *CCR5* quantification

EDTA-anticoagulated whole blood obtained from each of the study participants was stained within one hour of blood collection as described previously (PICTON *et al.* 2012b, Chapter 4, Section 4.2.2). Four antibody panels were used for each donor to assess *CCR5* expression on T cells, B cells, NK cells and granulocytes and monocytes. Furthermore, a HLA-DR marker was included in a fifth panel to assess the extent of cell activation (i.e., percentage of HLA-DR-expressing cells).

### 7.2.5. Absolute counts of cells in blood samples

TruCOUNT™ Tubes (BD Biosciences, San Jose, CA) were used for determining absolute counts of lymphocytes and monocytes in blood. Assay conditions were as recommended by the manufacturer. Briefly, monoclonal antibodies were added with 50 µl whole blood to the lyophilized pellet containing a known number of fluorescent beads and samples were prepared using a lyse/no wash procedure. Flow cytometric acquisition was performed on the FACSCalibur (BD Biosciences). Data was analysed using FloJo 7.6.1 (Tree Star, San Carlos, CA). Lymphocytes and monocytes were gated on a FL3(CD45) versus SSC dot plot, while beads were gated on a FL1 versus FL2 dot plot. Absolute counts were determined as follows:

$$\text{absolute cell count (cells/}\mu\text{l)} = \frac{\text{no. of events in cell-containing region}}{\text{no. of events in bead-containing region}} \times \frac{\text{no. beads per test (lot specific)}}{\text{test volume (50 }\mu\text{l)}}$$

T, B and NK cell counts were determined based on the proportions of total lymphocytes the respective cell subsets comprised, as determined from appropriate antibody panels used to quantitate *CCR5* expression.

Absolute counts of monocytes and lymphocyte subpopulations within PBMC cultures ( $1.4 \times 10^6$  cells total) were calculated based on the monocyte:lymphocyte ratio determined from the absolute counts.

### 7.2.6. *CCL3* quantification

*CCL3* was quantified in unstimulated and PHA-stimulated PBMCs by means of ELISA (DuoSet ELISA Development Systems; R&D Systems, Minneapolis, Minnesota, USA) as described previously (PICTON *et al.* 2013, Chapter 6, Section 6.2.8).

## 7.3. RESULTS

### 7.3.1. *CCR5* gene polymorphisms

Assembled sequences of the *CCR5* gene including promoter, coding and 3' UTR regions, from 14 HIV-1 infected LTNP individuals were analyzed for DNA polymorphisms, SNPs and indels. Across the entire 9.2 kb region sequenced, 39 SNPs and 4 indels were identified. No polymorphisms were found within the ORF. All polymorphisms identified had been previously reported, four of which we previously reported as new (-4223C/T, -3886C/T, +2066G/A and +2772 G insertion) in SAA and SAC individuals (PICTON *et al.* 2010, Chapter 2). Frequencies at which polymorphisms were found in SAA and SAC LTNP individuals, along with the background population frequencies are indicated in Table 7.2. We had previously reported several polymorphisms (n=48) that appeared to be restricted to either the uninfected SAA (n=41) or the SAC (n=47) population (PICTON *et al.* 2010). However, in the LTNP group, 18 of these were not restricted to one or other population as seen in the uninfected controls (Table 7.2).

The allelic frequencies of five of the SNPs detected in the SAC LTNP individuals, -3894T/C, -1686A/C, -1464A/G, -113G/T and +1752G/A, differed significantly from the healthy SAC cohort (Fishers  $P=0.0054$  for all SNPs, Table 7.2). With the exception of the -3894T/C SNP, these SNPs are in linkage disequilibrium and form part of the SAA-HHA haplotype (PICTON *et al.* 2010).

**Table 7.2.** Frequencies of identified polymorphisms within the South African African (SAA) and Caucasian (SAC) study populations

Location on gene	SNP Position	Base change (wt/mut)	Accession number <sup>a</sup>	n (LTNP frequency), population frequency <sup>b</sup>			
				SAA <sup>c</sup>		SAC <sup>c</sup>	
5' Flanking region (2762 bp)	-5268	G/A	rs3136535	1 (0.050)	<i>0.037</i>	0 (0)	<i>0.128</i>
	-5080	T/A	rs41429449	2 (0.100)	<i>0.049</i>	0 (0)	<i>0.011</i>
	-5072	C/T	rs35078594	3 (0.150)	<i>0.049</i>	0 (0)	<i>0</i>
	-4745	C/T	rs3136536	2 (0.100)	<i>0.122</i>	1 (0.125)	<i>0</i>
	-4257	A/C	rs41490645	1 (0.050)	<i>0</i>	0 (0)	<i>0.128</i>
	-4223	C/T	NI	3 (0.150)	<i>0.049</i>	0 (0)	<i>0</i>
	-4088	T/C	rs41499550	3 (0.150)	<i>0.220</i>	1 (0.125)	<i>0</i>
	-3899	A/C	rs72622924	5 (0.250)	<i>0.171</i>	3 (0.375)	<i>0.521</i>
	-3894	T/C	rs41395049	2 (0.100)	<i>0.220</i>	2 (0.250)	<i>0</i>
	-3886	C/T	NI	1 (0.050)	<i>0</i>	1 (0.125)	<i>0.032</i>
	-3868	CTAT/-	rs10577983	5 (0.250)	<i>0.171</i>	3 (0.375)	<i>0.500</i>
	-3458	G/T	rs2734225	5 (0.250)	<i>0.159</i>	3 (0.375)	<i>0.489</i>
	-3261	G/A	rs41475349	3 (0.150)	<i>0.220</i>	1 (0.125)	<i>0</i>
	-2852	A/G	rs2227010	2 (0.100)	<i>0.244</i>	1 (0.125)	<i>0.298</i>
Exon 1 (57 bp)	-2733	A/G	rs2856758	1 (0.050)	<i>0.043</i>	0 (0)	<i>0.138</i>
Intron 1 (501bp)	-2554	G/T	rs2734648	8 (0.400)	<i>0.366</i>	4 (0.500)	<i>0.489</i>
	-2459	G/A	rs1799987	6 (0.300)	<i>0.402</i>	1 (0.125)	<i>0.468</i>
Exon 2A (235bp)	-2135	T/C	rs1799988	6 (0.300)	<i>0.402</i>	1 (0.125)	<i>0.468</i>
	-2132	C/T	rs41469351	3 (0.150)	<i>0.207</i>	1 (0.125)	<i>0</i>
	-2086	A/G	rs1800023	5 (0.250)	<i>0.159</i>	3 (0.375)	<i>0.489</i>
Intron 2 (1903bp)	-1835	C/T	rs1800024	4 (0.200)	<i>0.159</i>	0 (0)	<i>0.064</i>
	-1686	A/C	rs9282632	5 (0.250)	<i>0.220</i>	2 (0.250)	<i>0</i>
	-1464	A/G	rs3181037	5 (0.250)	<i>0.220</i>	2 (0.250)	<i>0</i>
	-1130	AG/-	rs3054375	5 (0.250)	<i>0.159</i>	3 (0.375)	<i>0.489</i>
	-1060	C/T	rs2856762	1 (0.050)	<i>0</i>	1 (0.125)	<i>0.043</i>
	-976	C/T	rs2254089	5 (0.250)	<i>0.159</i>	3 (0.375)	<i>0.489</i>
	-975	G/A	rs41395249	2 (0.100)	<i>0.073</i>	1 (0.125)	<i>0</i>
	-651	C/T	rs2856764	5 (0.250)	<i>0.159</i>	3 (0.375)	<i>0.489</i>
	-444	G/A	rs2856765/ rs35046662	5 (0.250)	<i>0.159</i>	3 (0.375)	<i>0.489</i>
	-362	ACAA/G	rs71619644	5 (0.250)	<i>0.159</i>	3 (0.375)	<i>0.489</i>
	-113	G/T	rs3176763	5 (0.250)	<i>0.220</i>	2 (0.250)	<i>0</i>
-112	G/A	rs41352147	1 (0.050)	<i>0.012</i>	0 (0)	<i>0</i>	
3' UTR (2651bp)	+1752	G/A	rs41495153	4 (0.200)	<i>0.244</i>	2 (0.250)	<i>0</i>
	+1823	C/T	rs17765882	1 (0.050)	<i>0</i>	1 (0.125)	<i>0.043</i>
	+1843	G/A	rs41418945	2 (0.100)	<i>0.110</i>	1 (0.125)	<i>0</i>
	+1846	G/A	rs41466044	2 (0.100)	<i>0.110</i>	1 (0.125)	<i>0</i>
	+2066	G/A	NI	1 (0.050)	<i>0.049</i>	0 (0)	<i>0</i>
	+2077	G/T	rs1800874	5 (0.250)	<i>0.146</i>	3 (0.375)	<i>0.468</i>
	+2225	T/C	rs41535253	0 (0)	<i>0.049</i>	1 (0.125)	<i>0</i>
	+2458	A/C	rs3188094	<b>5 (0.250)<sup>d</sup></b>	<i>0.085</i>	0 (0)	<i>0</i>
	+2772	G insertion	NI	1 (0.050)	<i>0.085</i>	0 (0)	<i>0.011</i>
	+2838	C/G	rs41512547	0 (0)	<i>0.037</i>	1 (0.125)	<i>0</i>
	+2919	T/G	rs746492	7 (0.350)	<i>0.378</i>	1 (0.125)	<i>0.287</i>

<sup>a</sup>Accession numbers of SNPs detected in this study which have been previously reported in the SNP database (dbSNP).

NI indicates newly identified polymorphisms (Picton *et al.* 2010).

<sup>b</sup>Frequency was calculated for both LTNP populations using total number of alleles, i.e., n=20 for SAA and n=8 for SAC.

Background population frequencies are indicated in italicized font.

<sup>c</sup>Grey shading highlights polymorphisms where significant differences in SNP distribution between LTNP and respective healthy population exists ( $P=0.0054$  for all), bold font indicates a trend towards significance ( $P=0.055$ ).

### 7.3.2. CCR5 haplotypes

#### 7.3.2.1. The HHA haplotype may be overrepresented in SAC LTNP individuals

Individuals within the LTNP cohort were assigned to previously described haplogroups (GONZALEZ *et al.* 1999). Figure 7.1 illustrates the genotype distribution within the LTNP cohort. No LTNP individuals were found to possess the *CCR5* $\Delta$ 32 (HHG\*2) allele. All haplotypes, with the exception of HHB, found in our cohort of control SAA individuals were also found in the group of ten SAA LTNP individuals. Haplotypes HHA, HHC, HHD, HHE, HHF\*1, HHF\*2 and HHG\*1 were present in LTNP SAA individuals at frequencies of 30% (6/20), 25% (5/20), 15% (3/20), 5% (1/20), 5% (1/20), 15% (3/20) and 5% (1/20), respectively. Haplotype frequencies did not differ between control and HIV-1 infected LTNP SAA individuals. Haplotypes HHA, HHC, HHD and HHE were found in LTNP SAC individuals at a haplotype frequency of 37.5% (3/8), 37.5% (3/8), 12.5% (1/8), and 12.5% (1/8), respectively. Three of the four (75.0%) SAC LTNP individuals were heterozygous for the HHA allele. Fisher exact analysis indicates that the frequency of the HHA haplotype in the SAC group of LTNP individuals is significantly higher than that of SAC controls ( $P=0.010$ ). This allele is relatively rare (4.29%) in the HIV-1 uninfected SAC population (PICTON *et al.* 2010).

#### 7.3.2.2. Preliminary indications that HHC and HHE haplotypes may be associated with disease progression

To gauge whether there may be any differences in haplotype frequencies between HIV-1 infected individuals that progress to AIDS and those which remain asymptomatic for long periods (LTNP), a group of ten progressing SAA women, part of a mother-to-child-transmission cohort (KUHN *et al.* 2007), were also *CCR5* genotyped according to haplotypes described by Gonzalez *et al.* (1999) (Figure 7.1B).

As we have previously discussed, haplotypes HHA and HHC were designated as 'protective' in SAA and SAC individuals, respectively, HHC was considered 'deleterious' in the SAA population and HHE was designated 'deleterious' in both SAA and SAC individuals (PICTON *et al.* 2012a, Chapter 5). The HHF\*2 haplotype was designated as 'protective' in the SAA population where evidence of association with disease retardation appears to be strongest (MUMMIDI *et al.* 1998). This haplotype is distinguished from the HHF\*1 haplotype by a non-synonymous G/A SNP within the *CCR2* gene which results in a valine-to-isoleucine amino acid substitution, CCR2V64I (SMITH *et al.* 1997). The CCR2 64I mutation does not affect acquisition of HIV-1 infection (MICHAEL *et al.* 1997; SMITH *et al.* 1997) and controversial data exists on the influence of the CCR2 64I mutation on HIV-1 disease progression. However, a meta-analysis collecting all studies available in 2001 concluded that this allele (or haplotype) provides a significant protective effect on disease progression (IOANNIDIS *et al.* 2001). Subsequently, a second meta-analysis concluded that this

same mutation is associated with a decreased risk of death in perinatally infected infants (IOANNIDIS *et al.* 2003). The protective effects of this mutation have been shown to be population specific, although here too results seem to be contradictory (MUMMIDI *et al.* 1998; SMITH *et al.* 1997).

**A**

	CCR2-V64I	-2733	-2554	-2459	-2135	-2132	-2086	-1835	CCR5Δ32	Haplotypes		Genotype
Wild type	G	A	G	G	T	C	A	C				
Mutant	A	G	T	A	C	T	G	T		1	2	
<b>HIV-1 infected long term non-progressors</b>												
Patient ID	SAA (n=10)											
LTNP1	G/A	A	G	G/A	T/C	C	A	C/T		HHA	HHF*2	HHA/HHF*2
LTNP2	G/A	A	G	G/A	T/C	C	A	C/T		HHA	HHF*2	HHA/HHF*2
LTNP3	G	A	G	G	T	C	A	C		HHA	HHA	HHA/HHA
LTNP4	G	A	T	G	T	T	A	C		HHD	HHD	HHD/HHD
LTNP5	G/A	A	G	A	C	C	A	C/T		HHE	HHF*2	HHE/HHF*2
*LTNP6	G	A	G/T	G	T	C	A/G	C		HHA	HHC	HHA/HHC
*LTNP7	G	A	T	G	T	C	G	C		HHC	HHC	HHC/HHC
LTNP8	G	A/G	G	A	C	C	A	C/T		HHF*1	HHG*1	HHF*1/HHG*1
LTNP9	G	A	G/T	G	T	C	A/G	C		HHA	HHC	HHA/HHC
LTNP10	G	A	T	G	T	C/T	A/G	C		HHC	HHD	HHC/HHD
SAC (n=4)												
LTNP11	G	A	G	G/A	T/C	C	A	C		HHA	HHE	HHA/HHE
LTNP12	G	A	T	G	T	C	G	C		HHC	HHC	HHC/HHC
*LTNP13	G	A	G/T	G	T	C	A/G	C		HHA	HHC	HHA/HHC
LTNP14	G	A	G/T	G	T	C/T	A	C		HHA	HHD	HHA/HHD

**B**

HIV-1 infected progressors												
Patient ID	SAA (n=10)											
P1	G	A	G/T	G	T	C	A/G	C		HHA	HHC	HHA/HHC
P2	G	A	G/T	G	T	C/T	A	C		HHA	HHD	HHA/HHD
P3	G/A	A	G	G/A	T/C	C	A	C/T		HHA	HHF*2	HHA/HHF*2
P4	G	A	T	G	T	C/T	A/G	C		HHC	HHD	HHC/HHD
P5	G/A	A	G	A	C	C	A	C/T		HHE	HHF*2	HHE/HHF*2
P6	G	A	G/T	G/A	T/C	C/T	A	C		HHD	HHE	HHD/HHE
P7	G	A	G	A	C	C	A	C		HHE	HHE	HHE/HHE
P8		A	G	A	C	C	A	C/T		HHE	HHF*2	HHE/HHF*2
P9	G	A	G	G	T	C	A	C		HHA	HHA	HHA/HHA
P10	G/A	A	G	A	C	C	A	C/T		HHE	HHF*2	HHE/HHF*2

**Figure 7.1.** Haplotype and genotype distribution in HIV-1 infected long term non-progressing (A) and HIV-1 infected progressing individuals (B). Nucleotides found at haplotype defining positions are indicated (GONZALEZ *et al.* 1999). Wild type to mutant changes at these positions are also shown. Shaded boxes indicate where a heterozygous (dark grey) or homozygous (light grey) SNP was detected in a particular individual. Haplotypes and/or genotypes 'protective' and 'deleterious', from published data, are highlighted in blue and pink, respectively. Individuals with patient IDs: LTNP6, LTNP7 and LTNP13 met the criteria of elite controllers (i.e., patients with plasma HIV RNA levels of <50 copies/ml (patient ID indicated with \*)) (PEREYRA *et al.* 2008)).

No statistical differences were found in haplotype frequencies between the group of progressors and LTNPs. Furthermore, there were no differences in the number of 'protective' and 'deleterious' haplotypes between the two groups (Figure 7.1). However, it is interesting to note that the HHE haplotype appears to be overrepresented in the HIV-1 infected progressor group (6/20 haplotypes, 30.0%) in comparison to the LTNP group (2/28 haplotypes, 7.1%, 1 SAA and 1 SAC). The HHE haplotype is associated with an increased risk of acquiring HIV-1 infection and faster disease progression in ethnically divergent populations (GONZALEZ *et al.* 2005; MALHOTRA *et al.* 2011; MANGANO *et al.* 2001; OMETTO *et al.* 2001). Furthermore, the HHC haplotype, considered as potentially deleterious in SAA individuals was at a haplotype frequency of 25% (5/20) in SAA LTNP individuals but was only present at a frequency of 10% (2/20) in SAA progressors (Figure 7.1).

Since it is an individuals' genotype that is likely to influence the effect of CCR5 expression on disease progression, we have also indicated which individuals within the LTNP and progressor groups are considered to be protective and deleterious in this context (Figure 7.1). Similar to haplotype distribution, there was no difference in the number of 'protective' and 'deleterious' genotypes in the two groups.

### *7.3.3. CCR5 expression on lymphocyte populations of LTNP individuals*

CCR5 expression, measured both as density and percentage of CCR5-expressing cells within a cell subset was determined for granulocytes, monocytes and B, T, and NK cells (Table 7.3). We have previously shown CCR5 expression levels to differ considerably between SAA and SAC individuals (PICTON *et al.* 2012b, Chapter 4), however, given the small number in our SAC LTNP group, we did not compare expression between SAA LTNP and SAC LTNP individuals. Future studies in which our cohort of LTNP individuals will be expanded, will allow such comparisons to be made. Although data is presented for the NK cell subset CD56<sup>bright</sup>, it is important to note that this data was obtained from a very small subset for all LTNP individuals (<100 cells) and thus is likely to be an inaccurate representation.

**Table 7.3.** CCR5 expression, both as percentage of CCR5-expressing cells and CCR5 density, in HIV-1 infected long term nonprogressors (LTNP) from two South African population groups, South African African (SAA) (n=10) and South African Caucasian (SAC) (n=4) across all peripheral blood cell populations

Cell subset	Percentage of CCR5-expressing cells (%)				Mean number of CCR5 molecules per cell			
	SAA		SAC		SAA		SAC	
	median	range	median	range	median	range	median	range
<b>B cells</b>								
CD19+CCR5+	26.0	7.85-90.6	35.7	14.3-78.7	3424	1379-6886	1596	829-11112
<b>T cells</b>								
CD4+CCR5+	26.8	19.9-70.4	51.1	27.1-93.4	1718	1223-2503	2569	1409-4344
CD8+CCR5+	65.6	42.3-84.3	84.8	77.6-88.3	1822	1422-2035	2704	1803-3112
<b>NK cells</b>								
CD56+CCR5+	34.3	13.9-49.0	30.2	26.7-43.1	1244	1017-3447	1716	1266-3566
CD56 <sup>dim</sup> CCR5+	32.3	18.1-61.2	21.1	12.9-24.4	1114	701-2591	1819	1355-6271
CD56 <sup>bright</sup> CCR5+	60.5	10.0-75.0	43.8	29.9-71.5	1254	503-1810	1465	1016-3890
CD56+CD16+	25.3	13.3-52.2	19.5	14.6-40.8	1152	839-2200	1849	1297-2102
<b>Monocytes</b>								
CD14+CCR5+	10.9	2.13-12.9	10.5	5.7-24.5	3769	2792-5146	5170	3032-5695

#### 7.3.4. CCR5 expression relative to healthy HIV-1 uninfected controls

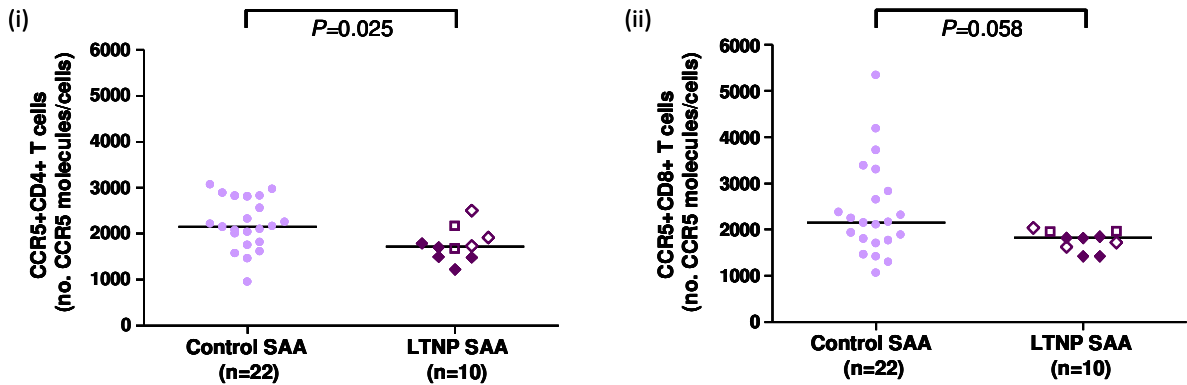
We next compared CCR5 expression of LTNP individuals to control healthy HIV-1 uninfected individuals. Since CCR5 expression differs between the two study populations (PICKTON *et al.* 2012b), comparisons between LTNP and control individuals were made across the same ethnic groups. Given the small number of individuals in the SAC LTNP cohort (n=4), we acknowledge the need for caution in making conclusions from these comparisons. However, we have included data and figures comparing expression levels between control SAC and LTNP SAC individuals to provide a preliminary indication of potential differences between control and LTNP individuals that can be investigated in larger cohort numbers. Control individuals heterozygous for the *CCR5Δ32* allele were excluded from comparisons because of the distinctive associated reduction in CCR5 expression and the absence of this genotype in the LTNPs (DE RODA HUSMAN *et al.* 1999; PICKTON *et al.* 2012b; SHALEKOFF and TIEMESSEN 2003; WU *et al.* 1997).

##### 7.3.4.1. LTNP individuals have lower CCR5 density on CD4+ T cells and monocytes compared to healthy HIV-1 uninfected individuals

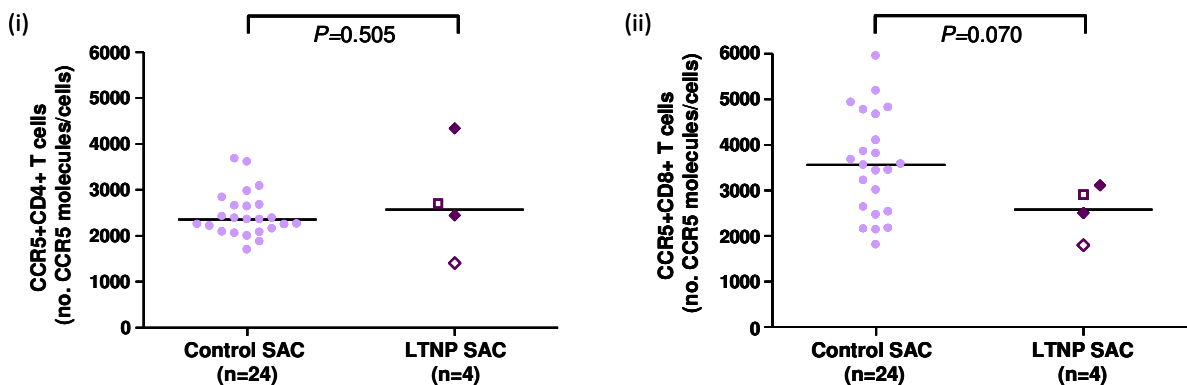
CCR5 density was significantly lower on CD4+ T cells and on monocytes of SAA LTNP individuals compared to SAA control individuals ( $P=0.025$  and  $P=0.022$ , respectively, Figure 7.2A(i) and Figure 7.3A). Furthermore, there was a trend towards SAA LTNP individuals expressing CCR5 at lower densities than SAA control individuals on CD8+ T cells ( $P=0.058$ , Figure 7.2A(ii)). It is

interesting to note that with the SAA LTNP group, the range of CCR5 density on CD8+ T cells, in comparison to control individuals, was very narrow (1422-2035 molecules/cell in LTNP, Table 7.3, vs 1055-5339 CCR5 molecules/cell in control SAA individuals). In the SAC group, there was no difference in CCR5 density on CD4+ or CD8+ T cells between healthy and LTNP individuals ( $P=0.505$ , Figure 7.2B(i) and ( $P=0.070$ , Figure 7.2B(ii), respectively). No differences in CCR5 density were noted in NK cell subsets between SAA LTNP and SAA control individuals.

A. SAA:

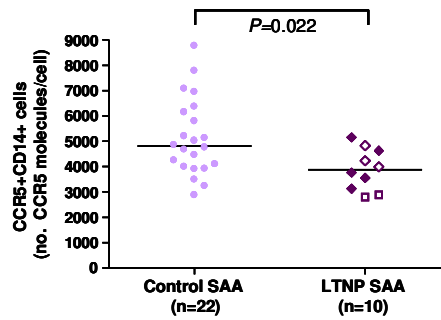


B. SAC (without *CCR5Δ132*):

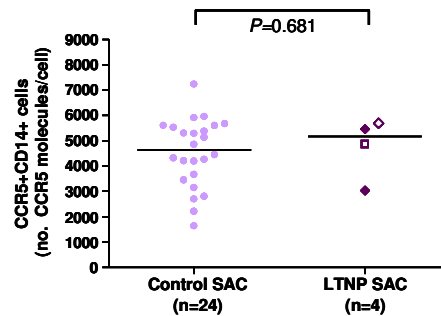


**Figure 7.2.** CCR5 density on CD4+ (i) and CD8+ (ii) T cells in healthy HIV-1 uninfected and HIV-1 infected long term nonprogressing (LTNP) South African African (SAA) (A) and South African Caucasian (SAC) (B) individuals.  $P$  values and the number of individuals in each group are indicated. HIV-1 infected LTNP individuals with viral loads <400 copies/ml are indicated with ( $\diamond$ ), elite controllers (viral load <50 copies/ml) are indicated with ( $\square$ ).

A. SAA:



B. SAC (without *CCR5Δ32*):

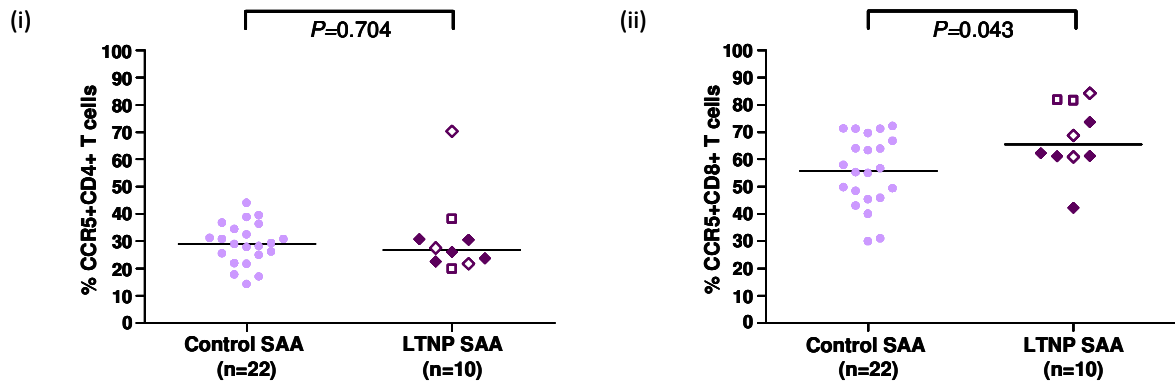


**Figure 7.3.** CCR5 density on monocytes in healthy HIV-1 uninfected and HIV-1 infected long term nonprogressing (LTNP) South African African (SAA) (A) and South African Caucasian (SAC) (B) individuals. *P* values and the number of individuals in each group are indicated. HIV-1 infected LTNP individuals with viral loads <400 copies/ml are indicated with (◇), elite controllers (viral load <50 copies/ml) are indicated with (□).

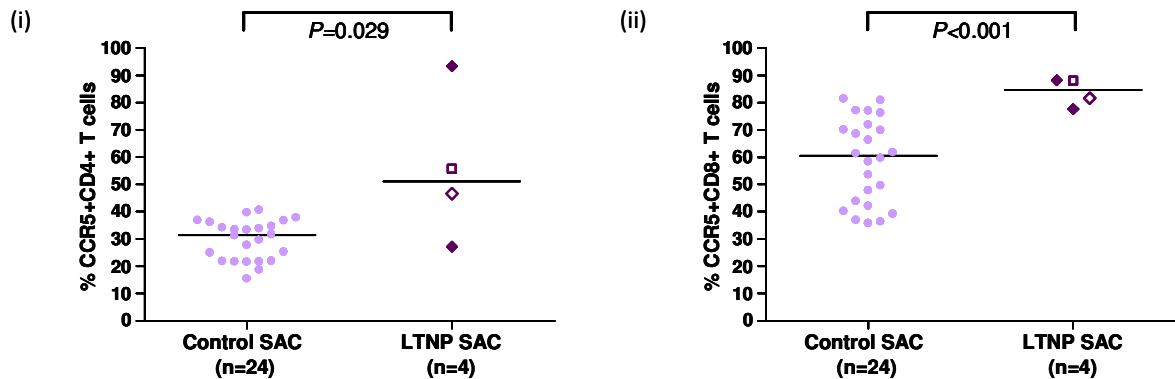
#### 7.3.4.2. LTNPs have higher proportions of CCR5-expressing CD8+ T cells relative to control individuals

Differences in the percentage of CCR5-expressing cells were also observed between LTNP and control uninfected individuals (Figure 7.4). The proportion of CCR5-expressing CD4+ T cells did not differ between control and LTNP SAA individuals ( $P=0.704$ , Figure 7.4A(i)). However, LTNP individuals (both SAA and SAC) had significantly higher proportions of CCR5-expressing CD8+ T cells in comparison to control uninfected individuals (Figure 7.4A(ii) and B(ii)). Although it appears that this may also be the case for CD4+ T cells of SAC individuals ( $P=0.029$ , Figure 7.4B(i)), the range of expression is very broad (27.1 – 93.4% CCR5-expressing CD4+ T cells) and the outlier is likely to be influencing statistics ( $P=0.114$ , however this analysis only includes 3 individuals in the LTNP SAC group).

A. SAA:



B. SAC (without *CCR5Δ32*):



**Figure 7.4.** Percentage of CCR5-expressing cells within T cell subsets, CD4+ (i) and CD8+ (ii), in healthy HIV-1 uninfected and HIV-1 infected long term nonprogressing (LTNP) South African African (SAA) (A) and South African Caucasian (SAC) (B) individuals. *P* values and the number of individuals in each group are indicated. HIV-1 infected LTNP individuals with viral loads <400 copies/ml are indicated with (◇), elite controllers (viral load <50 copies/ml) are indicated with (□).

7.3.4.3. Lower *CCR5* density on monocytes of HIV-1-infected LTNP individuals, in comparison to controls, maintained when individuals were matched for *CCR5* genotype

To control for the possible influence of individual haplotypes and/or genotypes, we next selected a subgroup of individuals from the SAA control and SAA LTNP cohorts that were matched for *CCR5* genotype and investigated whether there were any *CCR5* expression differences between the two groups. These subgroups consisted of the following genotypes (control:LTNP): HHA/HHA (1:1); HHA/HHC (2:1); HHA/HHF\*2 (2:1); HHC/HHD (1:2) and HHD/HHD (1:2). The subgroups of control and LTNP individuals did not differ with respect to age (median age of 41 and 38 years, respectively, *P*=0.805).

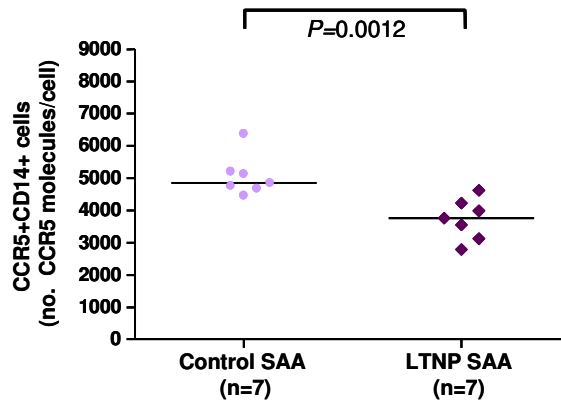
**Table 7.4.** Genotype matched comparison of CCR5 expression, both as percentage of CCR5-expressing cells and CCR5 density, between HIV-1 infected long term nonprogressors (LTNP) and healthy HIV-1 uninfected South African African (SAA) individuals across all peripheral blood cell populations

Cell subset	Percentage of CCR5-expressing cells (%)					Mean number of CCR5 molecules per cell				
	Control SAA (n=7)		LTNP SAA (n=7)		P	Control SAA (n=7)		LTNP SAA (n=7)		P
	median	range	median	range		median	range	median	range	
<b>B cells</b>										
CD19+CCR5+	24.2	21.0-40.7	20.6	7.85-90.6	0.530	1653	1326-5091	4260	1379-6886	0.202
<b>T cells</b>										
CD4+CCR5+	27.7	21.5-36.6	26.1	19.9-70.4	0.902	2154	1815-2804	1787	1223-2503	0.073*
CD8+CCR5+	57.8	42.9-69.5	62.3	42.3-84.3	0.318	2314	1413-4183	1824	1422-2035	0.209
<b>NK cells</b>										
CD56+CCR5+	28.0	7.52-48.5	34.7	13.9-49.0	0.620	1177	912-1294	1306	1118-1810	0.097
CD56 <sup>dim</sup> CCR5+	29.5	7.26-43.9	37.9	18.1-61.2	0.456	1296	976-1422	1068	701-2058	0.710
CD56 <sup>bright</sup> CCR5+	65.0	15.5-78.1	60.9	10.0-75.0	0.456	1038	701-1199	1317	1042-1810	0.017
CD56+CD16+	26.6	6.02-44.5	29.6	13.9-99.2	0.456	1203	860-1614	1036	839-1483	0.456
<b>Monocytes</b>										
CD14+CCR5+	7.03	3.48-11.3	10.7	8.67-12.9	0.073*	4847	4458-6371	3762	2792-4625	0.001

\* Indicates a trend towards significance.

Table 7.4 shows comparisons between these two groups. Although it was noted that CCR5 density on CD56<sup>bright</sup> NK cells was significantly higher in LTNP SAA individuals than in control SAA individuals ( $P=0.017$ , Table 7.4), this result was based on a small number of cells within this subset and is likely to be inaccurate. The differences noted in CCR5 expression on T cell subsets between HIV-1 uninfected SAA control and LTNP individuals (Figure 7.2) lost significance when individuals matched for *CCR5* genotype were compared (Table 7.4). This could potentially be attributed to the small number of individuals in each group ( $n=7$  in each). However, a trend towards LTNP individuals expressing CCR5 at lower densities on CD4+ T cells compared to control individuals was maintained ( $P=0.073$ , Table 7.4).

It is interesting to note that the difference in CCR5 density on monocyte cells between control and LTNP SAA individuals became more significant in the subgroup of genotype matched individuals compared to that seen when SAA individuals were not genotype-matched ( $P=0.0012$ , Figure 7.5 and  $P=0.022$ , Figure 7.3A, respectively). A trend towards lower CCR5 density on CD4+ T cells on LTNP individuals compared to controls was noted ( $P=0.073$ , Table 7.4).



**Figure 7.5.** CCR5 density on monocytes in healthy HIV-1 uninfected and HIV-1 infected long term nonprogressing (LTNP) South African African (SAA) individuals matched for *CCR5* genotype. The *P* value and the number of individuals in each group are indicated.

### 7.3.5. Influence of haplotypes on *CCR5* expression

#### 7.3.5.1. HHA

We have previously demonstrated that the healthy HIV-1 uninfected individuals, positive for the HHA haplotype, have significantly lower percentages of CCR5-expressing CD8<sup>+</sup> T cells compared to individuals that lack this haplotype (PICTON *et al.* 2012a). Given that within our SAA LTNP cohort, 5/10 individuals were HHA<sup>+</sup> (with one homozygous individual, Figure 7.1), we investigated whether the presence of this haplotype had an impact on CCR5 expression within this group. The HHA haplotype had no impact on CCR5 expression levels on any of the cell populations studied, both in terms of CCR5 density and the percentage of CCR5-expressing cells within a subset.

#### 7.3.5.2. HHC

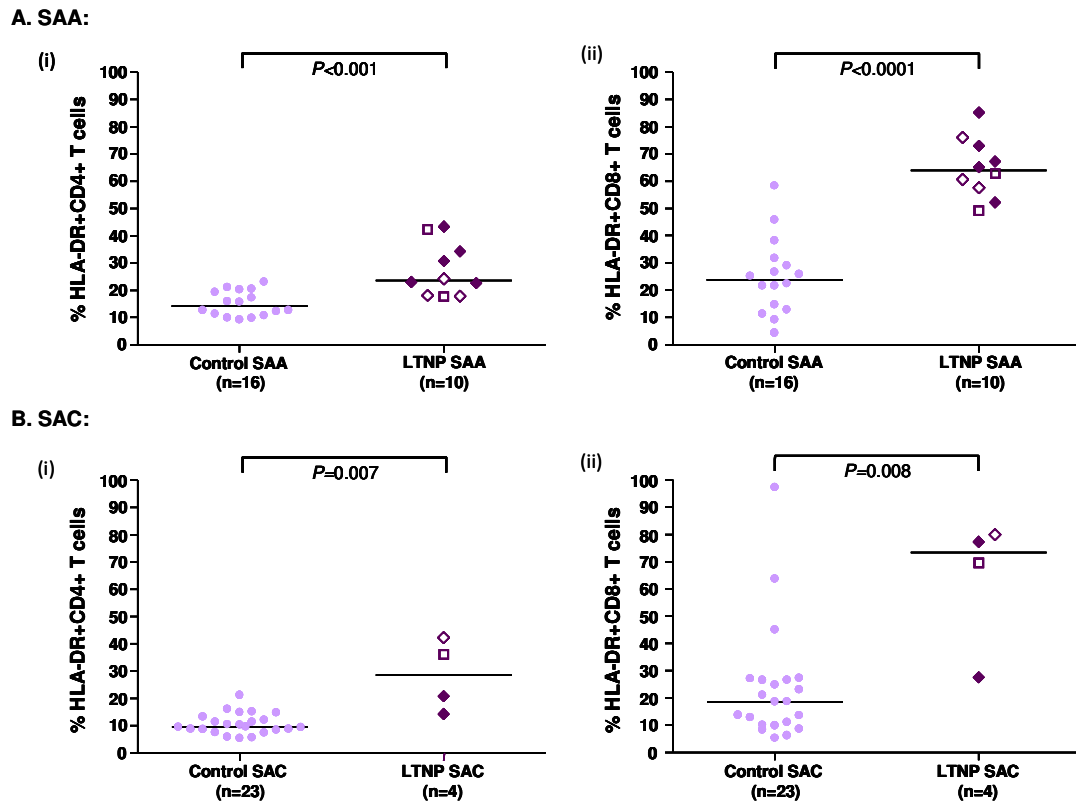
The second most prevalent haplotype within the SAA LTNP group was HHC. We investigated whether there were any CCR5 expression differences between individuals with (HHC<sup>+</sup> SAA LTNP, n=4) and those lacking the haplotype (HHC<sup>-</sup> SAA LTNP individuals, n=6). HHC<sup>+</sup> individuals had a greater percentage of CCR5-expressing CD56<sup>dim</sup> NK cells compared to HHC<sup>-</sup> individuals (39.7% and 20.7% median, respectively, *P*=0.019). However, HHC<sup>+</sup> individuals demonstrated lower activation levels (median of 14.7% HLA-DR<sup>+</sup>CD56<sup>+</sup>CD3<sup>-</sup>, range: 3.23-16.7%) compared to HHC<sup>-</sup> individuals (median of 37.5% HLA-DR<sup>+</sup>CD56<sup>+</sup>CD3<sup>-</sup>, range: 24.0-42.6%) on NK cells (*P*=0.010). A similar relationship was observed in CD8<sup>+</sup> T cells where HHC<sup>+</sup> individuals demonstrated lower activation levels (median of 55.0% HLA-DR<sup>+</sup>CD8<sup>+</sup>CD3<sup>+</sup>, range: 49.3-62.8%) compared to HHC<sup>-</sup> individuals (median of 70.2% HLA-DR<sup>+</sup>CD8<sup>+</sup>CD3<sup>+</sup>, range: 60.7-85.3%) (*P*=0.019). Given the small number of HHC<sup>+</sup> and HHC<sup>-</sup> SAA LTNP individuals in this analysis, these relationships will have to be verified on a larger cohort of LTNP individuals. However, it is interesting to note that in healthy

HIV-1 uninfected individuals we have shown that HHC+ SAC individuals have significantly higher CCR5 density on NK (CD56+) cells and CD56+CD16+ NK cell subsets compared to SAC individuals that lack this haplotype (PICTON *et al.* 2012a). The HHC haplotype did not have an impact on the percentage of CCR5-expressing cells in SAA individuals (PICTON *et al.* 2012a).

*7.3.6. Cell activation was higher in LTNP individuals, compared to control individuals, in T cell but not NK cell subsets*

The levels of activation were also measured for T and NK cells. Cell activation levels, as measured by the percentage of cells expressing HLA-DR, were significantly higher on CD4+ and CD8+ T cells in SAA LTNP individuals compared to control individuals from the same population ( $P < 0.001$  for both cell subsets, Figure 7.6A). No differences in cell activation were observed in NK cells ( $P = 0.363$ ). Similarly, cell activation levels in both CD4+ and CD8+ T cells were higher in SAC LTNP individuals compared to SAC control individuals (Figure 7.6B).

In the LTNP individuals (SAA and SAC), cell activation levels were higher in CD8+ than in CD4+ T cells. The median percentage of HLA-DR-expressing CD4+ T cells in SAA LTNP individuals was 23.7%, whereas this value was considerably higher in CD8+ T cells, i.e., 64.0% ( $P < 0.0001$ ). Four of the five LTNP SAA individuals with viral loads  $< 400$  copies/ml had HLA-DR expression levels at or below that of the group median for both CD4+ and CD8+ T cell subsets, indicating an association between viral load and activation levels in the T cell subsets. The median percentage of HLA-DR-expressing CD4+ T cells and CD8+ T cells in SAC LTNP individuals was 28.5% and 73.4%, respectively. LTNP individuals from both population groups tend to express HLA-DR at higher levels (% HLA-DR-expressing cells) compared to control individuals. This is expected as a result of HIV-1 infection.



**Figure 7.6.** Cell activation levels, as measured by the percentage of HLA-DR-expressing cells, in CD4+ (i) and CD8+ (ii) T cells, in healthy HIV-1 uninfected and HIV-1 infected long term nonprogressing (LTNP) South African African (SAA) (A) and South African Caucasian (SAC) (B) individuals.  $P$  values and the number of individuals in each group are indicated. HIV-1 infected LTNP individuals with viral loads  $<400$  copies/ml are indicated with ( $\diamond$ ), elite controllers (viral load  $<50$  copies/ml) are indicated with ( $\square$ ).

### 7.3.7. *CCL3L*, *CCL3La* and *CCL3Lb* copy number distribution in LTNP individuals

Variation in *CCL3L* copy number has been associated with HIV-1 disease progression (DOLAN *et al.* 2007). Thus, *CCL3L*, *CCL3La* and *CCL3Lb* gene copy number was determined for all LTNP individuals. The number of total *CCL3L* copies ranged from 2 to 8 in SAA LTNPs, with a median of 5, and ranged from 0 to 6 in SAC LTNPs, with a median of 1.5. The median copy number for *CCL3La* and *CCL3Lb* genes were calculated to be 4.5 and 2 in SAA LTNPs, and 1.5 and 0 in SAC LTNPs, respectively. The copy number distribution of *CCL3L*, *CCL3La* and *CCL3Lb* did not differ between LTNP individuals and healthy control individuals from the same population group. Although it appears that the proportion of SAA LTNP individuals that have copy numbers above the population-specific median for the *CCL3L*, *CCL3La* and *CCL3Lb* genes are greater than that observed in the background control population (70.0% vs 53.9%, 80.0% vs 62.9% and 60.0% vs 49.4%, respectively), these relationships lack statistical significance ( $P=0.505$ ,  $P=0.487$  and  $P=0.741$ , respectively). Given that the LTNP study participant numbers are small and there exists a

broad variation in range of gene copy numbers, it will be necessary to assess copy number distribution in a larger cohort of LTNP individuals in order to determine whether there is a bias towards individuals having copy numbers above the population-specific median in this group.

### 7.3.8. PHA-stimulated CCL3 production in LTNP individuals compared to control individuals

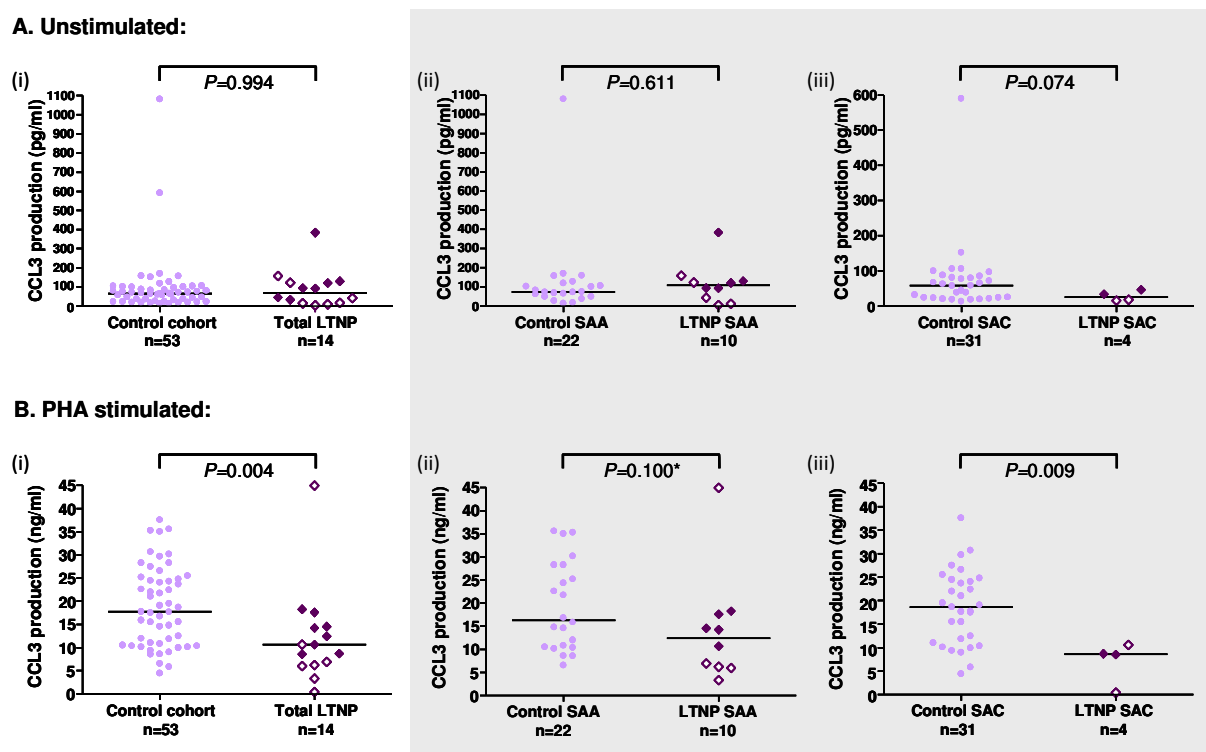
We measured the production of CCL3 in the supernatants of both unstimulated and PHA-stimulated PBMCs obtained from 10 SAA and 4 SAC HIV-1-infected LTNP individuals by ELISA. The median CCL3 production upon PHA stimulation, was determined to be 12.42 ng/ml in SAA LTNP and 8.64 ng/ml in SAC LTNP individuals. CCL3 production did not correlate with either HIV-1 viral load or CD4+ T cell counts ( $P=0.994$  and  $P=0.938$ , respectively).

We next compared the CCL3 production of isolated PBMCs from control individuals and LTNPs. We have shown that SAA and SAC individuals do not produce CCL3 at different levels despite the differences in *CCL3L* copy number (PICTON *et al.* 2013, Chapter 6) and have thus also included analyses using SAA and SAC individuals as a combined group (Figure 7.7). In the absence of PHA stimulation, no differences were noted between control and LTNP individuals (Figure 7.7A), although CCL3 production was marginally lower in LTNP SAC individuals compared to control SAC individuals ( $P=0.074$ , Figure 7.7A(iii)), but not in the SAA population ( $P=0.611$ , Figure 7.7A(ii)).

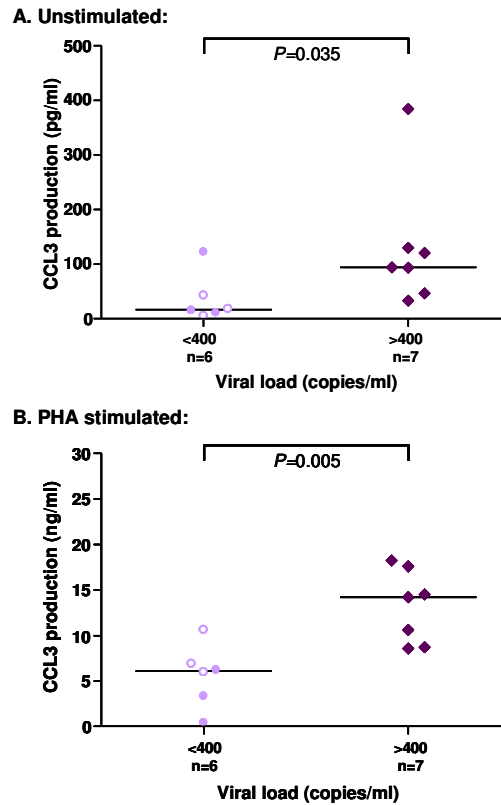
Within the LTNP cohort, there was one SAA individual (LTNP3), considered as an outlier, whose CCL3 production was well above that of the control and LTNP cohort combined (44.95 ng/ml compared to range of 4.45 – 37.58 ng/ml in control individuals and 0.41 - 18.24 ng/ml in remaining LTNP individuals). However, the *CCL3L* copy number of this individual is equal to the SAA population median, i.e., five. This individual had low viral load (183 copies/ml), a CD4+ T cell count of 968 cells/ml and, at the time at which this experiment was conducted, had been HIV-1 positive for nine years. It is interesting to note that in a similar experiment conducted to determine CCL4 levels within the same cohorts, this same individual had very high CCL4 production in comparison to all other individuals tested (A. Bharuthrum, unpublished data).

Upon PHA stimulation, CCL3 production differed significantly between control and LTNP individuals (Figure 7.7B). This was true whether or not the individuals were stratified according to ethnicity ( $P=0.004$  (Figure 7.7B(i)) for the combined group,  $P=0.100$  (Figure 7.7B(ii)) for SAA individuals, and  $P=0.009$  (Figure 7.7B(iii)) for SAC individuals). Although the PHA-induced CCL3 production difference between control SAA and LTNP SAA individuals does not appear to be significant ( $P=0.100$ , Figure 7.7B(ii)), if the LTNP outlier is removed from the analysis, this does become significant ( $P=0.025$ ). Thus, LTNP individuals produced significantly lower CCL3 upon stimulation in comparison to control uninfected individuals. Furthermore, upon exclusion of the LTNP outlier from analysis, LTNP (SAA and SAC) individuals with viral loads <400 copies/ml ( $n=6$ )

had significantly lower PHA-induced CCL3 production compared to those with viral loads >400 copies/ml (n=7), with median CCL3 production of 6.10 and 14.22 ng/ml, respectively ( $P=0.005$ , Figure 7.8B). In addition, unstimulated production of CCL3 was significantly lower in LTNP individuals with viral loads <400 copies/ml compared to those with viral loads >400 copies/ml (n=7) (medians of 16.4 vs 94.4 pg/ml,  $P=0.035$ , Figure 7.8A).



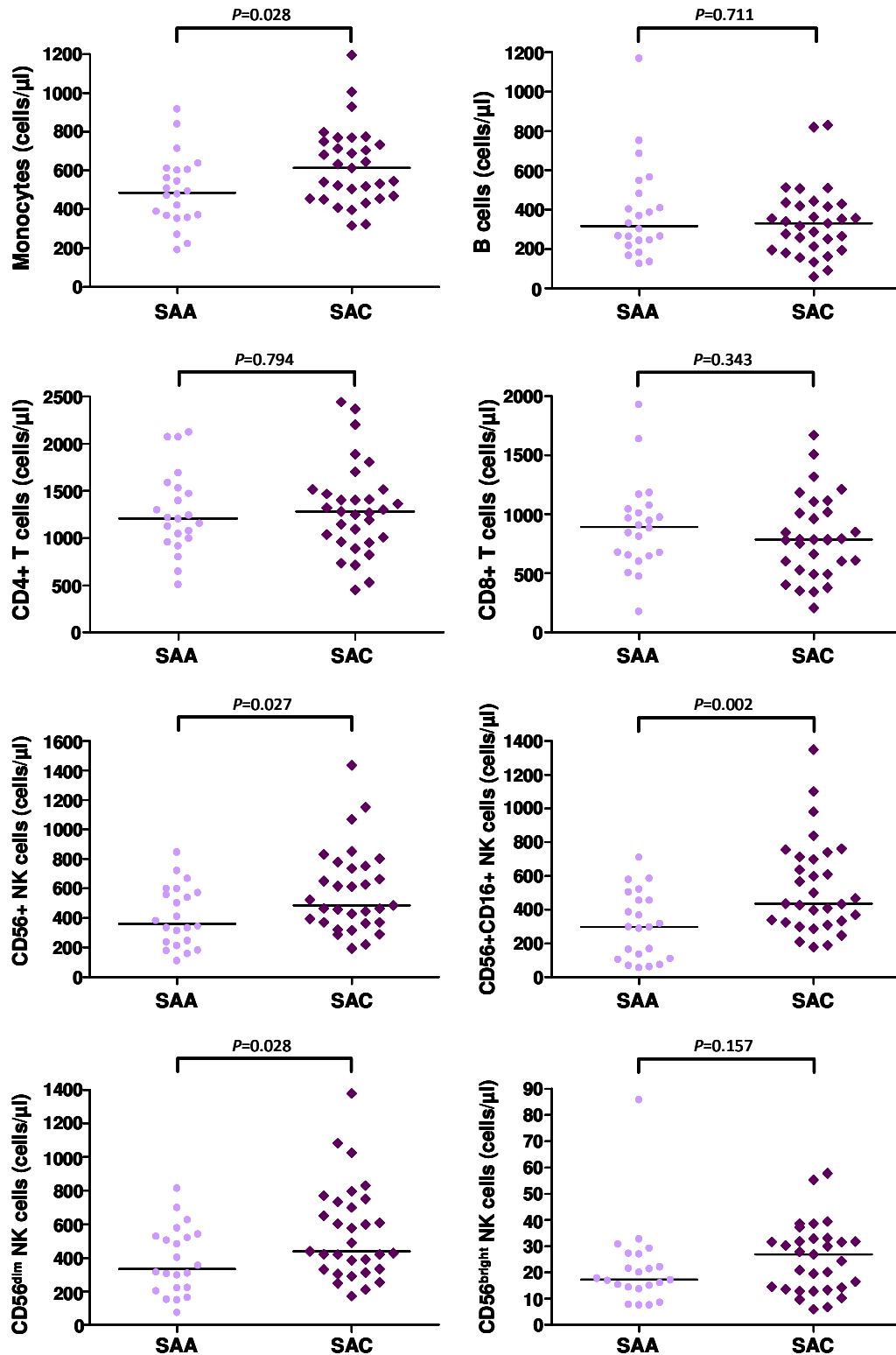
**Figure 7.7.** CCL3 production by peripheral blood mononuclear cells (PBMCs) isolated from HIV-1-infected long term nonprogressors (LTNPs) and healthy HIV-1 uninfected individuals as measured in CCL3 production assays. Overall CCL3 production is shown for unstimulated (**A**) and PHA stimulated (**B**) isolated PBMCs for the entire cohort (i), South African African (SAA) (ii) and South African Caucasian (SAC) (iii).  $P$  values and numbers (n) in each group are indicated. HIV-1 infected LTNP individuals with viral loads <400 copies/ml are indicated with ( $\diamond$ ). \*  $P=0.025$  upon exclusion of LTNP outlier.



**Figure 7.8.** CCL3 production by peripheral blood mononuclear cells (PBMCs) isolated from HIV-1 infected long term nonprogressors (LTNPs) stratified according to low (<400 copies/ml) and higher (>400 copies/ml) viral loads. CCL3 production is shown for unstimulated (**A**) and PHA stimulated (**B**) isolated PBMCs. P values and numbers (n) in each group are indicated. One LTNP individual, considered to be an outlier was excluded from the analysis. HIV-1 infected elite controllers (viral load <50 copies/ml) are indicated with (O).

### 7.3.9. Absolute counts of monocytes and NK cell subsets in whole blood differ between SAA and SAC individuals

It is unclear to what extent the contribution to CCL3 production by various cell subsets varies. Differences in CCL3 production could potentially be attributed to variation in the distribution of CCL3-producing cells between control and HIV-1-infected LTNP individuals. Thus, we determined absolute counts of various cell types within whole blood samples of the study participants. Absolute counts of lymphocytes and monocytes were determined by means of flow cytometry with the inclusion of fluorescent beads. Counts of T, B and NK cell subsets were then determined based on the proportions of total lymphocytes these cell subsets comprised.



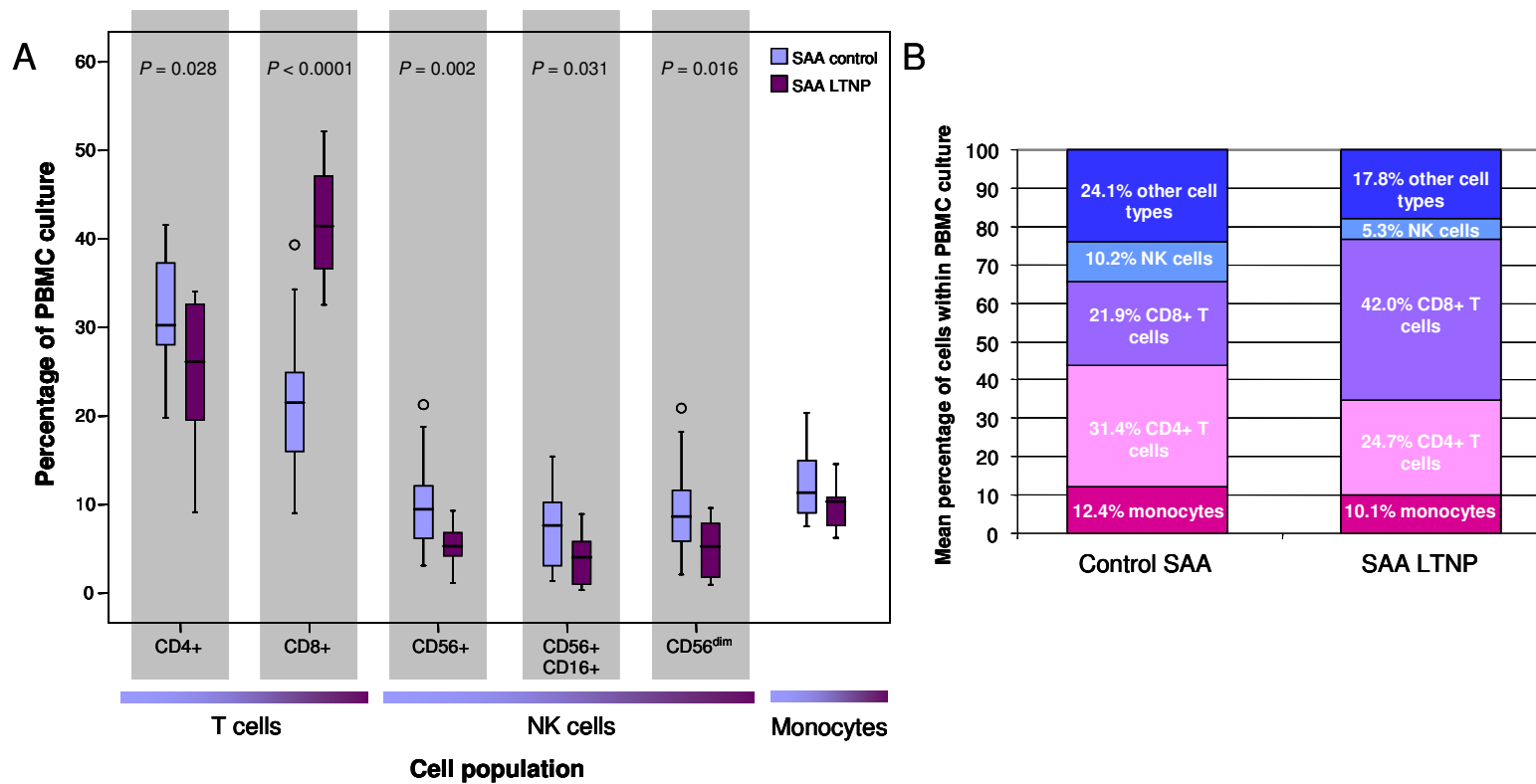
**Figure 7.9.** Comparison of absolute counts of monocytes and lymphocyte cell subsets from South African African (SAA) (n=22) and South African Caucasian (SAC) (n=31) healthy HIV-1 uninfected individuals. *P* values are indicated.

We first determined whether control SAA and SAC individuals differ in absolute cell counts of different cell types found in whole blood. Healthy control SAA individuals had lower counts of both monocytes ( $P=0.028$ ) and NK cells (CD56+, CD56+CD16+ and CD56<sup>dim</sup> subsets) ( $P=0.027$ ,  $P=0.002$  and  $P=0.028$ , respectively, Figure 7.9) compared to healthy control SAC individuals. However, it is important to note that despite differences in absolute counts of various cell subsets, SAA and SAC individuals were shown to not differ in CCL3 production in PBMC cultures (PICTON *et al.* 2013). Figure E.1 within the Appendices illustrates absolute cell count differences between control and LTNP individuals for both population groups.

#### *7.3.10. SAA LTNP PBMCs had higher proportions of T cells and lower proportions of NK cells compared to SAA control PBMCs*

Isolated PBMCs are comprised of monocytes and lymphocytes, including T (CD4+ and CD8+), B and NK cells. We calculated the expected number of each of these cell types within the isolated PBMCs used in CCL3 production assays by extrapolating from absolute cell counts. We then compared percentage composition of PBMCs between SAA control and LTNP individuals. Analysis was restricted to SAA individuals since absolute counts in whole blood were found to differ between SAA and SAC healthy uninfected individuals (Figure 7.9).

SAA LTNPs had significantly higher proportions of T cells within isolated PBMCs compared to SAA control individuals ( $P=0.0002$ ). Not unexpectedly, LTNP individuals had significantly fewer CD4+ T cells compared to control individuals ( $P=0.028$ , Figure 7.10A). This CD4+ T cell deficit was offset by the significantly higher number of CD8+ T cells LTNP individuals had in culture compared to control individuals ( $P<0.0001$ , Figure 7.10A). CD8+ T cell expansion is expected in HIV-1-infected individuals. LTNP individuals were also found to have lower numbers of NK cells (CD56+, CD16+CD56+ and CD56<sup>dim</sup>) than control individuals ( $P=0.002$ ,  $P=0.031$ ,  $P=0.016$ , respectively, Figure 7.10A). No differences in monocyte cell numbers were noted between the two groups ( $P=0.163$ ). Figure 7.10B indicates the expected proportional representation of different cell types within PBMCs isolated from uninfected and LTNP SAA individuals.



**Figure 7.10.** Distribution of different cell types within peripheral blood mononuclear cells (PBMCs) isolated from South African African (SAA) HIV-1 uninfected control (n=22) and SAA long term nonprogressing (LTNP) (n=10) individuals. **(A)** Percentage of PBMCs comprised of monocytes and T and NK cell subsets in the two study groups. Where significant, *P* values have been indicated. Box-whisker plots depicting the median (horizontal black line), 25<sup>th</sup> and 75<sup>th</sup> percentiles (margins of the box) and the 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers). Outliers are indicated with (○). **(B)** Proportional representation of different cell types within PBMCs isolated from control (n=22) and SAA LTNP (n=10) individuals determined by using mean values from the respective groups. Cells referred to as “other cell types” include B cells, double negative T cells, basophils and dendritic cells.

## 7.4. DISCUSSION

The vast variation in the rates of HIV-1 disease progression among individuals can be attributed to viral, genetic and immunological factors. A number of host genetic factors associated with delayed disease progression have been identified (Reviewed in LAMA and PLANELLES 2007). In this study we have assessed the possible contribution to delayed disease progression by two of these factors, the CCR5 receptor and its chemokine ligand, CCL3. We have characterized *CCR5* gene polymorphisms (both SNPs and indels) and intragenic haplotypes, determined CCR5 expression on different blood cell types, determined *CCL3L*, *CCL3La* and *CCL3Lb* gene copy numbers as well as CCL3 production levels for 14 HIV-1 infected long term nonprogressing (ten SAA and four SAC) individuals. This was conducted as a pilot study to assess to what extent LTNP individuals differ from control uninfected individuals both genotypically and phenotypically, when considering the HIV-1 coreceptor-ligand (CCR5-CCL3) axis. In addition, we included a small group of HIV-1 infected progressors for genetic comparison. Although we were restricted to low sample numbers, we were able to identify some interesting relationships which can be investigated further upon future expansion of cohorts.

Although 39 SNPs and four indels were identified within the LTNP cohort, these were restricted to the noncoding regions of the *CCR5* gene and all had previously been identified. Polymorphisms within the *CCR5* ORF, such as the *CCR5Δ32* and m303 (303A/T) mutations, introduce premature stop codons and hence prevent expression of a functional protein (LIU *et al.* 1996; QUILLEN *et al.* 1998). Further genetic variants within the *CCR5* coding region result in non-synonymous protein mutations which may alter the conformation of the CCR5 receptor and hence its function (CARRINGTON *et al.* 1999; FOLEFOC *et al.* 2010). Nonetheless, other than *CCR5Δ32*, polymorphisms in the *CCR5* ORF have not been reported to play a dominant role in nonprogression among HIV-1 infected individuals (COHEN *et al.* 1998), which may be due to these variants occurring at relatively low frequencies. Since neither the *CCR5Δ32* allele, nor any other polymorphism within the *CCR5* ORF, were detected in our LTNP cohort, nonprogression in these individuals cannot be attributed to possible structural changes within the CCR5 protein. However, haplotypes within the *CCR5* promoter region have been associated with slower disease progression (GONZALEZ *et al.* 1999).

The HHA haplotype was previously only detected at a frequency of 4.25% in healthy SAC individuals and thus is relatively rare in that population (PICTON *et al.* 2010, Chapter 2). However, 3/4 LTNP SAC individuals were heterozygous for this allele indicating that the HHA haplotype might be playing an important role in Caucasians in delaying disease progression. In African American individuals, the HHA haplotype is associated with HIV-1 disease retardation, where even

the presence of a single allele is sufficient to confer disease retardation (GONZALEZ *et al.* 1999). However, in Caucasians this allele has been shown to not have a dominant effect, in fact, the HHA/HHA genotype is associated with disease acceleration (GONZALEZ *et al.* 1999). It is not clear, however, whether HHA, when present as a single allele in Caucasians will be beneficial or detrimental in the context of HIV-1 infection. It is likely that, as with other haplotypes such as HHE, it may be the haplotype pairing that determines its effect. As additional SAC individuals are recruited into our cohort of LTNP individuals, it will be interesting to determine whether this allele remains overrepresented in SAC LTNPs in comparison to SAC control individuals and whether particular haplotype pairings with HHA are more common.

According to population studies including African American individuals, the HHC haplotype is associated with accelerated disease progression in this population and thus we have considered it as deleterious in SAA individuals (GONZALEZ *et al.* 1999; PICTON *et al.* 2012a). However, in the current study, although not statistically significant, the HHC haplotype appears to be overrepresented in the SAA LTNP group compared to the SAA progressor group. Interestingly, preliminary data indicates that LTNP individuals with the HHC haplotype appear to have lower proportions of CCR5-expressing NK cells (CD56<sup>dim</sup>) compared to those that lack this allele. Although, when assessing the influence of the HHC haplotype on CCR5 expression we noted higher coreceptor density on NK cells of SAC individuals with the HHC haplotype than those that lacked the allele, we failed to see this relationship in SAA individuals (PICTON *et al.* 2012a, Chapter 5). Furthermore, no differences in the percentage of CCR5-expressing NK cells were noted between control groups with or without HHC (PICTON *et al.* 2012a). As with the HHA allele, the relative frequency of the HHC allele in SAA LTNP, rapid progressing and uninfected individuals will need to be further assessed as the cohort is expanded.

In a study of HIV-1 infected Caucasian individuals, the frequency of the HHE haplotype was found to be significantly higher in the group of rapid progressors compared to slow progressors (LI *et al.* 2005). The HHE haplotype is associated with disease acceleration in ethnically divergent populations including Caucasian, African American, Asian and African populations (AN *et al.* 2000; GONZALEZ *et al.* 1999; LI *et al.* 2005; MARTIN *et al.* 1998; NGUYEN *et al.* 2004; OMETTO *et al.* 2001). Although there was no statistical significance, the HHE haplotype was more prevalent in our group of SAA progressors than in the SAA LTNPs (30.0% vs 5.0%,  $P=0.091$ ).

It has previously been reported that active HIV-1 replication, as manifested by detectable plasma HIV-1 viraemia, up-regulates CCR5 expression on NK cells (CD3-CD56+) with higher percentages of CCR5-expressing NK cells seen in viraemic individuals compared to HIV-1 uninfected individuals (JIANG *et al.* 2008; KOTTILIL *et al.* 2004). In the current study, no differences in CCR5 expression on NK cells were noted between control and LTNP individuals. However, given that the

percentage of CCR5-expressing NK cells between individuals without viraemia (<50 copies/ml) and HIV-1 uninfected individuals do not differ (KOTTILIL *et al.* 2004), the absence of any observed significant differences in CCR5 expression on NK cells between LTNP and uninfected control individuals might be explained by the inclusion of three EC (two SAA and one SAC) with viral loads of <50 copies/ml and a further four individuals with viral loads of <400 copies/ml (three SAA and one SAC EC) in our LTNP cohort. The percentage of CCR5-expressing NK cells has also been shown to positively correlate with HIV-1 viral loads (JIANG *et al.* 2008; KOTTILIL *et al.* 2004), a relationship which was not observed in our study. Both studies used cohorts with a much broader range and higher viral loads compared to the patients in our study. It is likely that with smaller differences in viral loads, changes in CCR5 expression will be less evident. Furthermore, since HIV-1 uninfected individuals from different population groups differ in their levels of CCR5 expression on NK cells (PICTON *et al.* 2012b), an observation that might be explained by differences in CCR5 haplotype distribution in different population groups (PICTON *et al.* 2012a), it is possible that the ethnicity of individuals used in the other studies may influence results.

Although HIV-1 infection is known to upregulate CCR5 density on T cells, particularly CD4+ T cells (NICHOLSON *et al.* 2001; OSTROWSKI *et al.* 1998), we observed the opposite in that HIV-1 infected LTNP individuals expressed CCR5 at lower densities compared to healthy individuals within these same cell subsets. Our observation is in agreement with a previous report showing lower CCR5 density on CD4+ T cells of slow progressors compared to uninfected controls (REYNES *et al.* 2001). The beneficial effects of reduced CCR5 expression on CD4+ T cells has been highlighted in Simian Immunodeficiency Virus (SIV) infection models where it has been demonstrated that natural SIV hosts, that do not develop AIDS-like symptoms, express CCR5 at remarkably lower proportions in CD4+ T cells isolated from blood, lymph nodes, and mucosal tissues compared to humans and non-natural SIV hosts such as rhesus macaques (PAIARDINI *et al.* 2009).

In addition to *CCR5* genetic variants, other mechanisms have been identified which influence CCR5 function. Although the reason for reduced CCR5 density on T cells in LTNP individuals is unknown, a possible cause could be the presence of anti-CCR5 antibodies. Antibodies to CCR5 which reduce HIV-1 infectivity of cells through CCR5 downregulation, have been found in both HIV-exposed seronegative individuals and HIV-1 infected LTNP individuals leading to the hypothesis that anti-CCR5 antibodies protect against HIV-1 infection (LOPALCO *et al.* 2000; PASTORI *et al.* 2006). Anti-CCR5 antibodies, which were shown to protect against infection with HIV-1 R5 strains *in vitro*, have also been detected in breast milk of both HIV-seronegative (66%) and seropositive (83%) women (BOUHLAL *et al.* 2005). Similarly, natural anti-CCR5 antibodies, capable of inhibiting the infection of macrophages and dendritic cells by HIV-1, have also been detected in the cervicovaginal secretions of both HIV-infected and HIV-uninfected women (ESLAHPAZIR *et al.*

2008). Collectively, these findings suggest that some degree of protection against vertical transmission of HIV-1 may be mediated by the presence of anti-CCR5 antibodies.

Among the factors associated with HIV-1 disease progression, coinfection with other viruses has been associated with delayed disease progression. Coinfection with either the GB type C (GBV-C, also known as hepatitis G virus) and Human T-cell leukemia virus 2 (HTLV-2) viruses has been associated with delayed HIV-1 disease progression. *In vitro* studies have shown that both viruses upregulate CCL3 production with a concomitant reduction in CCR5 expression (PILOTTI *et al.* 2007; XIANG *et al.* 2004). Thus, the observed reduction in CCR5 expression might also be explained by the presence of coinfections within these individuals. In future studies, we will also be assessing the possible presence of both antibodies to CCR5 and the presence of coinfecting viruses, and comparing the prevalence of these between uninfected control and HIV-1 infected LTNP individuals.

CCR5 expression correlates directly with the differentiation of monocytes to macrophages (NAIF *et al.* 1998; TUTTLE *et al.* 1998). Tissue macrophages are derived exclusively from blood monocytes, which as monocyte-derived macrophages support HIV-1 replication. Thus, it is interesting that when SAA control and LTNP individuals were matched for *CCR5* genotypes, the lower CCR5 density on monocytes became more significant compared to when the groups were analysed as a whole, thus indicating that the reduction in CCR5 density was not a result of the effect of haplotype/genotype induced expression differences but rather as a result of another mechanism. We have previously reported a negative correlation between age and CCR5 density on monocytes in SAA but not SAC individuals (PICTON *et al.* 2012b). Given that the SAA LTNP cohort comprises a group of individuals who are older than the SAA control group (median age difference of 7 years), it could be argued that age may be a contributing factor to the observed monocyte CCR5 expression differences between the two groups. However, when genotype matched, SAA control and LTNP individuals did not differ in age, therefore ruling out this possibility. As these CCR5 expression assays are conducted in whole blood, anti-CCR5 antibodies, chemokines or other factors mentioned earlier (coinfection with GBV-C, HTLV-2) may play a role. These factors will all in future be tested.

Macrophages play a central role in HIV-1 pathogenesis. These are long-lived cells which are relatively resistant to the cytopathic effects of HIV-1, making them ideal HIV-1 reservoirs, and are involved in trafficking HIV-1 from the periphery to the brain (Reviewed in GRAS and KAUL 2010; IGLESIAS-USSEL and ROMERIO 2011). Macrophages are also widely distributed throughout the body and are able to spread HIV-1 to CD4+ T cells (GROOT *et al.* 2008). Lower CCR5 expression on macrophages of LTNP individuals may make them less permissive to HIV-1 infection. We have previously demonstrated the proportion of CCR5-expressing monocytes to be significantly higher in

HIV-1-infected SAA individuals (viral load:  $50418 \pm 13336$  copies/ml, CD4 counts:  $355 \pm 62$  cells/ $\mu$ l) compared to uninfected healthy control individuals (SHALEKOFF *et al.* 2001). However in the current study, the percentage of CCR5-expressing monocytes did not differ between SAA control and SAA LTNP individuals ( $P=0.325$ ) which may be due to the LTNP status of the study participants.

When adjusting for the influence of *CCR5* genotypes on CCR5 expression, the assumption is made that this is the only factor contributing towards CCR5 expression differences. However, in addition to the role played by anti-CCR5 antibodies and coinfection with other viruses, CCR5 expression can also be influenced by a number of other factors, including the levels of its chemokine ligands, cell activation levels and other cytokines as discussed previously (PICTON *et al.* 2012a). In addition, the effects of particular haplotypes and/or genotypes on CCR5 expression may also be contradictory. Thus, to accurately assess differences in CCR5 expression on different cell types in the context of nonprogression and progression of disease, a study would ideally be conducted where genotypes were matched in much higher numbers.

As expected, HIV-1 infected individuals demonstrated higher T cell activation levels, as measured by the percentage of HLA-DR-expressing cells, compared to control HIV-1 uninfected individuals. Given that increased cell activation is known to result in an increase in CCR5 expression (MEDITZ *et al.* 2011; MORIUCHI *et al.* 1997; OSTROWSKI *et al.* 1998; WU *et al.* 1997), it is intriguing that these same individuals expressed CCR5 at densities lower than that of control individuals on the same cells. This suggests one of two possible hypotheses: (i) upon infection, despite the high levels of activation, this group of individuals is able to downregulate CCR5 expression through unknown mechanisms or (ii) LTNP individuals are comprised of individuals who, prior to infection with HIV-1, would be stratified with the lower CCR5-expressing individuals and their immune cells are not as “activatable” as other individuals who progress more rapidly. CD38 is widely used as a marker of T and B cell activation. In a cross-sectional study of normal progressing HIV-1 infected individuals we previously demonstrated significantly higher percentage of CD38+CCR5+ lymphocytes compared to normal donors (SHALEKOFF *et al.* 2001). Upon reexamination of this same data, the percentage of CCR5-expressing CD8+ T cells from HIV-1-infected individuals was found to correlate positively with the percentage of CD38-expressing CD8+ T cells ( $P=0.005$ ,  $r=0.54$ ; S. Shalekoff, unpublished data). This serves to indicate that the percentage of CCR5 expression could be used as a surrogate for a cell activation measure. The percentage of CCR5-expressing CD8+ T cells, but not CD4+ T cells, from LTNP individuals was higher compared to control individuals which would reflect the higher cellular activation in LTNP individuals.

Although several studies have investigated the influence of HIV-1 infection and disease progression on CCL3 chemokine levels, results have been contradictory. There is a lack of

consensus on whether HIV-1 infected individuals produce CCL3 at higher, lower or equivalent levels compared to HIV-1 uninfected individuals in plasma or serum samples (KAKKANAIH *et al.* 1998; KROWKA *et al.* 1997; YE *et al.* 2004), with the most recent report demonstrating that CCL3 plasma levels in healthy HIV-1 uninfected individuals are higher than that of HIV-1 controllers (viraemic and elite controllers combined), which in turn produce CCL3 at higher levels compared to HIV-1 infected noncontrollers (CARD *et al.* 2012). Furthermore, there is a lack of consensus on whether or not cellular CCL3 production and/or circulating CCL3 levels correlate with disease progression (CARD *et al.* 2012; GARZINO-DEMO *et al.* 1999; KROWKA *et al.* 1997; MCKENZIE *et al.* 1996; POLO *et al.* 1999; ULLUM *et al.* 1998; YE *et al.* 2004). These results are difficult to interpret due to a lack of homogeneity in patient selection as well as differences in study design.

Stimulation with PHA resulted in lower CCL3 production by LTNP individuals compared to controls. This is in contrast to two other studies investigating CCL3 production upon PHA stimulation. In the first study, PHA stimulation of whole blood demonstrated that HIV-1 infected individuals produce CCL3 at higher levels than both HIV-1 uninfected individuals and HIV-1 infected individuals who have progressed to AIDS (ULLUM *et al.* 1998). The second study showed that CD8+ T cells isolated from asymptomatic HIV-1 infected individuals produced CCL3 at higher levels than control individuals, but that individuals who have progressed to AIDS produce CCL3 at levels comparable to uninfected persons (COCCHI *et al.* 2000). However, the two previous reports as well as the current study differ experimentally both in the cells selected to analyse and in the selection criteria of study individuals. We included individuals known to control infection for a number of years, whereas the selection criteria for both other studies was based on whether or not HIV-1 infected individuals had progressed to AIDS, without the knowledge of the duration of symptom-free infection (COCCHI *et al.* 2000; ULLUM *et al.* 1998). It is likely that CCL3 production may differ considerably between LTNP individuals and asymptomatic HIV-1 chronically infected patients who lack sufficient control of their infection and will progress to AIDS more rapidly.

Interestingly, PBMCs from LTNP individuals with low viral loads produced CCL3 at lower levels than those from individuals with higher viral loads, irrespective of whether or not the cells were stimulated. The former also demonstrated relatively lower activation levels, as measured by the percentage of cells expressing HLA-DR, on CD4+ and CD8+ T cells. If activation levels correlate with CCL3 production (MENTEN *et al.* 2002), one would expect CCL3 production of unstimulated PBMCs isolated from LTNPs to be higher than that of healthy control individuals, however, CCL3 production was comparable between the two groups. Stimulation with PHA resulted in lower production by LTNP individuals compared to controls. In a study assessing CCL4 production upon stimulation within the same study groups, CCL4 production was also found to be significantly lower in LTNP individuals compared to control individuals (A. Bharuthrum, unpublished data). This raises

the question as to whether there is generalized lower production of chemokines and/or cytokines in LTNP individuals compared to HIV-1 infected rapid progressors. This also begs the question as to whether HIV-1-infected individuals attain LTNP status due to an ability to downmodulate immune cell activation.

Differences in PBMC CCL3 production between groups could potentially be attributed to differences in the distribution of CCL3 producing cells. Although CCL3 production is inducible in most haematopoietic cells (Reviewed in MENTEN *et al.* 2002), it is not clear whether different cell types within a PBMC culture will produce CCL3 at different levels, relative to each other, upon PHA stimulation. In addition, HIV-1 infection may differentially affect CCL3 production of different cell types. For example, it has been shown that the production of CCL3 by NK cells isolated from HIV-1 infected individuals correlates negatively with viral loads, i.e., HIV viraemia impairs the ability of NK cells to secrete CC chemokines (KOTTILIL *et al.* 2003), whereas higher production of CCL3 by CD8+ T cells is associated with nonprogression (COCCHI *et al.* 2000; KOTTILIL *et al.* 2003). Furthermore, CCL3 production by CD8+ T cells is higher in HIV-1 infected individuals, without AIDS, compared to healthy uninfected individuals (COCCHI *et al.* 2000). Both T cell (CD4+ and CD8+) and NK cell subsets, which together constitute the majority of cells found in PBMCs, differed substantially between control and LTNP individuals with respect to the proportions of cell types. Thus, we would recommend assessing CCL3 production from single cell types in future studies.

In addition to the effect of gene polymorphisms, CCR5 expression can also be influenced by *CCL3L* copy numbers (GONZALEZ *et al.* 2005). HIV-1 plasma viral loads have been demonstrated to correlate negatively with viral load (DOLAN *et al.* 2007; SHALEKOFF *et al.* 2008). Furthermore, SAA HIV-1-infected women with *CCL3L* copy numbers above the population-specific median of 5 were demonstrated to have significantly higher magnitude of Gag-specific CD4+ T cell responses than those with lower copy numbers (SHALEKOFF *et al.* 2008). Although lacking in significance, the proportion of SAA LTNP individuals with *CCL3L*, *CCL3La* and *CCL3Lb* copy numbers above the population-specific median was generally higher than that of the background control population. These differences are likely to become significant in a larger cohort of LTNPs.

When assessing the influence of genetic risk groups determined according to whether individuals have detrimental or non-detrimental *CCR5* genotypes and whether they had high or low *CCL3L* copy numbers, it has been shown that the group of individuals with high *CCL3L* copy numbers relative to their population median and who also possessed non-detrimental *CCR5* genotypes were the slowest to progress to AIDS (GONZALEZ *et al.* 2005). The opposite was observed in individuals with a combination of low *CCL3L* copy numbers and detrimental *CCR5* genotypes (GONZALEZ *et al.*

2005). Thus the combinatorial effect of CCL3 and CCR5 may be more important in the context of HIV-1 infection than that of the contribution made by the individual molecules.

In conclusion, we found both CCR5 expression and CCL3 production was lower in LTNP individuals compared to healthy controls even though no *CCR5* coding region polymorphisms were found in any of our LTNP participants and *CCL3L*, *CCL3La* and *CCL3Lb* copy numbers did not differ between LTNP and control individuals. However, haplotypes HHA, HHC and HHE were identified as potentially associated with disease progression. The results of this preliminary study will be tested in future studies with larger numbers of LTNPs and rapid progressors.

---

**CHAPTER 8**  
**Concluding Remarks**

---

There is increasing evidence demonstrating that genetic differences between ethnically divergent populations has an impact on the susceptibility of individuals to both infectious (communicable) and non-communicable diseases. In addition, population-based host genetic polymorphisms have been shown to influence immune responses to vaccines. This provides an important rationale for determining genetic variation of disease-associated genes within different populations.

In order to gain entry into a target cell, HIV-1 must bind to CD4 and a coreceptor, CCR5 or CXCR4 (DENG *et al.* 1996; DRAGIC *et al.* 1996; FENG *et al.* 1996). The vast majority of new HIV-1 infections are initiated by CCR5-utilising virions (R5 HIV-1). Thus, CCR5 plays an important role in both the immune system, by mediating chemotaxis of CCR5-expressing cells in response to its ligands, and in mediating HIV-1 infection. The functional efficiency of CCR5 is in part determined by the level of CCR5 expression on the surface of cells which can be influenced by several factors including genetic polymorphisms in both the coding and noncoding regions of the genes and interaction with its  $\beta$ -chemokine ligands, CCL3, CCL4 or CCL5. Since the binding of the  $\beta$ -chemokines to CCR5 induces internalization, CCR5 expression levels may also be influenced by differences in the circulating levels of its ligands. *CCL3* and *CCL3L* code for two isoforms of a CCR5 ligand, i.e., CCL3 and CCL3L. CCL3L is the most potent known agonist of CCR5 and inhibitor of R5 HIV-1 strains. Individuals who have *CCL3L* gene copy numbers above their population-specific median are reported to be less susceptible to HIV-1 infection and to exhibit slower progression compared to individuals within the same population who have *CCL3L* gene copy numbers below the population-specific median (DOLAN *et al.* 2007; GONZALEZ *et al.* 2005; KUHN *et al.* 2007; MEDDOWS-TAYLOR *et al.* 2006).

Although two previous studies have assessed the distribution of *CCR5* polymorphisms in South African populations (PETERSEN *et al.* 2001; WILLIAMSON *et al.* 2000), these studies were restricted to the *CCR5* ORF and did not assess the extent of variability within the noncoding regions of the gene. Genetic variability within the *CCR5* promoter region can significantly impact CCR5 expression and thus determine both an individual's susceptibility to HIV-1 infection and rate of disease progression to AIDS. The overall aim of this study was thus to determine the extent of *CCR5* gene variability within two ethnically divergent South African populations and to determine how host gene variation of CCR5 translates to different expression levels of cell surface CCR5 on various immune cell types. We also sought to determine whether significant differences in median *CCL3L* gene copy numbers, as seen in the same two populations, is associated with differences in protein production. In the last component of this study, genotypic variants of *CCR5* and *CCL3L* in the context of HIV-1 infection were investigated in a small cohort of long term nonprogressors (LTNPs).

In order to assess factors that affect expression levels of CCR5 it was first necessary to determine genetic variation within the target populations. We believe that one of the strengths of this work is that genetic variation was not restricted to a small region of the gene but was determined for the entire length of the gene. We demonstrated substantial variation with regards to *CCR5* polymorphisms (Chapter 2) as well as CCR5 cell surface expression (Chapter 4) between SAA and SAC individuals. The selection of the cohort (i.e., HIV-1 uninfected individuals) ensured that there would be no skewing of the genetic variability data due to the presence of HIV-1 infected individuals. In order to evaluate the impact of genetic variability on CCR5 expression it is imperative to do so in the absence of HIV-1 infection since infection in its own right is known to alter expression levels of a range of proteins and receptors, including CCR5 and CCL3 (NICHOLSON *et al.* 2001; OSTROWSKI *et al.* 1998; ULLUM *et al.* 1998).

An expected, greater genetic diversity and a low level of linkage disequilibrium within African populations in comparison to populations of European descent (TISHKOFF *et al.* 2009; TISHKOFF and VERRELLI 2003) was reflected in our study (Chapter 2). Of the *CCR5* polymorphisms identified, 60 were found in SAA individuals, while only 38 were found in SAC individuals. Twenty three SNPs and one indel were newly identified in this study. The HHC haplotype frequency was found to be significantly higher in SAC individuals compared to that observed in other studies of Caucasian populations. This, in addition to evidence of a north-to-south geographic cline in *CCR5* $\Delta$ 32 (HHG\*2) allele frequency in Caucasian European populations (NOVEMBRE *et al.* 2005; STEPHENS *et al.* 1998), highlights the importance of assessing genetic variability within populations which are geographically separated even though they are of the same ethnicity.

Although associations between particular SNPs and the differential outcome of disease provide valuable information, it is not always possible to pinpoint the SNP responsible for the observed effect as SNPs and/or indels are often in linkage disequilibrium with each other resulting in haplotypes. In this study, we identified haplotypes that were extensions of previously defined haplotypes restricted to a small region of the gene, i.e., approximately 800 bp in addition to the *CCR5* $\Delta$ 32 and *CCR2V64I* mutations (GONZALEZ *et al.* 1999). The haplotypes defined in this study thus span across the entire *CCR5* gene (Chapter 2). Interestingly, only one of these haplotypes, SAA/C-HHC, was common to the two populations. The HHE haplotype provides an example where the SNPs making up HHE are in linkage disequilibrium with different SNPs in the SAA and SAC populations thus forming the extended haplotypes, SAA-HHE and SAC-HHE, respectively. This is also likely to be the case for HHA, although the number of SAC individuals positive for this haplotype was low and consequently we were unable to confirm this. If extended haplotypes differ between populations, this could, in part explain some of the differing HIV-1 disease associated

phenotypic effects observed when the smaller root haplotypes were investigated in different populations (GONZALEZ *et al.* 1999).

Having identified the three haplotypes, SAA-HHA, SAA/C-HHC and SAC-HHE, as collectively the most prevalent haplotypes in the two study populations, we next evaluated the effect of polymorphisms within various *CCR5* promoter regions of these haplotypes on promoter efficiency or strength (Chapter 3). Furthermore, we also evaluated differences in promoter activity between four regions of the promoter. Although the impact of variability (in the form of different *CCR5* haplotypes) within the P1 (downstream) promoter region on promoter activity has previously been assessed (MUMMIDI *et al.* 2000), having identified additional SNPs to be in complete linkage disequilibrium with these haplotypes necessitated the investigation of the role of these additional SNPs on promoter activity.

In a complex gene such as *CCR5*, which has two alternate promoters, assessing the impact of genetic variability within promoter regions on protein expression is complicated because it is unknown to what extent the two promoter regions interact with, and influence each other. For example, polymorphisms in one of the promoter regions which have either a positive or negative impact on protein expression may be counteracted by polymorphisms in the second promoter. To date, analysis of the influence of SNPs on promoter activity has been restricted to the *CCR5* downstream promoter (referred to as P1A in our study). Furthermore, although the importance of the upstream (P2) promoter region in determining *CCR5* expression on activated primary T cells has been demonstrated (MUMMIDI *et al.* 2007), the influence of *CCR5* SNPs within this region on promoter activity has, to our knowledge, not been studied. Thus, in addition to independent assessment of the downstream and upstream promoters (P1A and P2, respectively), we assessed the effect of polymorphisms of the entire *CCR5* promoter region (P1+P2) on the promoter activity (Chapter 3). Interestingly, although lacking statistical significance, the SAC-HHE P1+P2 promoter proved to be the strongest of the three haplotypes evaluated in all the cell lines tested. However, we found results of the SAC-HHE P1A promoter strength relative to the other haplotypes to vary according to the cell line used. HHE, the root haplotype of SAC-HHE, is associated with both an increased risk of acquisition of HIV-1 infection in adults and exposed infants, and acceleration to AIDS in ethnically divergent populations (discussed in Chapter 5). One would thus expect *CCR5* expression to be higher in HHE-haplotype bearing individuals compared to individuals with other *CCR5* haplotypes.

Promoter expression of the various constructs tested was highest in the monocytic cell lines, THP-1 and U937, which corresponds with the observed highest *CCR5* density on monocytes in comparison to all other immune cells studied (Chapter 4). SAC-HHE P1A promoter activity was ranked highest (although lacking in significance) in U937 cells and second highest (by a small

margin) in THP-1 cells, while in the T lymphocyte cell line, Jurkat, SAC-HHE promoter activity ranked lowest of the haplotypes tested. Suspension cells, such as those used in this study, can be difficult to transfect when traditional lipofection is used. The technology used in this study (Nucleofector<sup>®</sup> Technology, Lonza) has made huge strides in increasing transfection efficiency, particularly in cells that are difficult to transfect. According to the technical specifications of the Amaxa<sup>®</sup> Cell Line Nucleofector<sup>®</sup> kits used in this study, transfection efficiency for K562, THP-1 and U937 can be expected to be very similar ( $\approx 70\%$ , 24 h post Nucleofection<sup>®</sup>) and that of Jurkat E6-1 cells is much higher ( $\approx 90\%$ ) (<http://www.lonza.com>). Assuming that the transfection efficiencies that we attained in our experiments were in accordance with the reported Amaxa<sup>®</sup> efficiency, the results obtained suggest evidence of CCR5 expression being differentially regulated in different cell types, probably as a result of differences in transcription factor repertoires between cell types. This was highlighted when CCR5 cell surface expression on different immune cells in whole blood samples was analysed relative to the presence or absence of particular CCR5 haplotypes (Chapter 5).

The assay we used to evaluate promoter strengths does have some inherent problems which makes it difficult to compare our results to those obtained by other studies. Variation between studies is likely to be attributable to differences in experimental conditions and the regions selected to test promoters, as well as the cell lines used. Thus, although our data does indicate that SNPs across the entire promoter region are influencing *CCR5* promoter strength, further investigation is necessary in order to determine the true effect of this variability on promoter strength. Repeating these assays in isolated PBMCs, or even purified cell subsets may provide a more accurate representation of promoter strength differences between *CCR5* haplotypes. However, this may at the same time introduce further assay variability as a result of donor-to-donor variation.

An interesting phenomenon is that some haplotypes appear to have differing disease-modifying effects which are dependent on the population being studied. For example, the HHC haplotype has been shown to be associated with disease retardation in Caucasian and Hispanic populations, whereas it was associated with disease acceleration in African Americans (GONZALEZ *et al.* 1999). Unlike other haplotypes (eg. HHE), we found SNPs comprising the HHC haplotype to be in complete linkage equilibrium with the same SNPs, upstream and downstream from the HHC haplotype defining region, in both study populations (thus termed SAA/C-HHC, Chapter 2). This implies that if, as in African American and Caucasian American populations, the HHC haplotype is associated with different HIV-1 disease phenotypes in SAA and SAC populations, these differences cannot be attributed to genetic polymorphisms within the *CCR5* gene. Assuming that these phenotypic differences are due to variations in CCR5 cell surface expression, differences in

other *CCR5* gene expression regulating factors may be contributing towards differential *CCR5* expression.

The high degree of genetic variability between the two study populations served to indicate that this may be accompanied by differences in *CCR5* surface expression between the two populations. Previous studies assessing *CCR5* expression in the context of HIV-1 infection have focused on the role of *CCR5* as a coreceptor and hence have only looked at select cell types, in particular CD4+ T cells. There has never been any consideration of the potential role of other *CCR5*-expressing cells and their possible role in protection from infection acquisition or control of disease progression. We assessed *CCR5* expression, measured as both the proportions of cells within a particular cell subset which express *CCR5* and the density of *CCR5*, on T cells, NK cells, B cells and monocytes of a subset of individuals which had been *CCR5* genotyped (Chapter 4). Generally, *CCR5* density was higher on cells of SAC individuals, while SAA individuals had higher proportions of *CCR5*-expressing cells. CD4+ T cells, the main HIV-1 permissive cell type, did not differ between SAC and SAA individuals – either in *CCR5* proportions or density. Interestingly, *CCR5* density on NK cells was found to differ between SAA and SAC individuals. We postulate that increased *CCR5* expression on NK cells may have an impact on NK cell trafficking and, although this may make them more permissive to HIV-1 infection and consequently may serve as HIV-1 reservoirs, enhanced migration may serve to counterbalance this by improving their killer capacity in response to *CCR5* ligands released through immune activation in chronic HIV-1 infection. Given that NK cells are increasingly being found to play an important role in HIV-1 infection control, we believe this provides a different perspective on how HIV-1 might be prevented or controlled through altered *CCR5* expression that is not necessarily explained by CD4+ T cell infection alone.

Since *CCR5* expression is known to increase with cell activation, we investigated whether there were any differences in cell activation levels, measured as the percentage of HLA-DR-expressing cells, between the two study populations (Chapter 4). Few studies have assessed cell activation level differences between healthy HIV-1 uninfected individuals from different ethnicities or population groups living in the same environment. We observed higher cell activation levels in SAA individuals compared to SAC individuals. Immune activation levels have previously been reported to be environmentally driven (CLERICI *et al.* 2000). However, in this study, the participants were largely controlled for environment, since the cohort comprised laboratory workers who were all living in the same geographical region, which led us to postulate that differences in immune activation levels between the two populations is more likely to be a result of genetic differences. A limitation of our study, however, was that the technique used to determine *CCR5* expression levels (four colour flow cytometry) did not allow us to include sufficient antibodies in the panels to differentiate between cell subsets, to detect *CCR5* and also to determine activation levels on these

same cell sets. Thus, we were unable to determine the extent of CCR5 expression on activated cell subsets and compare CCR5 expression on these to cell subsets which lacked expression of cellular activation markers. This could provide an indication of the extent to which CCR5 expression differences seen between the population groups are a result of higher activation levels. However, given that SAA individuals demonstrated higher proportions of CCR5-expressing cells, but not higher CCR5 density, compared to SAC individuals, in whom activation levels were lower, differences in CCR5 expression between the two populations cannot be attributed to activation alone.

Having genotyped as well as determined CCR5 expression of select individuals, we investigated the effect of the presence of particular *CCR5* haplotypes on CCR5 expression (Chapter 4 and 5). As with all genetic based assays, large cohorts are required to unequivocally establish the influence of genetic differences on particular biological phenotypes. Thus, our relatively small sample numbers is a limitation of our study. However, our findings, if validated in a bigger population or in other studies, could further our understanding of CCR5 expression particularly on cells other than T cells.

There has been a long-standing interest in studying CCR5 expression differences between individuals and we have provided further evidence that the extent of CCR5 expression on immune cells can vary according to individuals' *CCR5* haplotypes. Importantly, we demonstrated this in cell types other than CD4+ T cells (Chapter 5). Interestingly, the presence of the *CCR5* $\Delta$ 32 allele (HHG\*2 haplotypes) was found to have a greater impact on CCR5 density than on the percentage of CCR5-expressing cells (Chapter 4). The two main observations, with regards to the influence of *CCR5* haplotypes on expression, were that the presence of HHA results in lower percentages of CCR5-expressing CD8+ T cells in SAA individuals and that HHC increases CCR5 density in NK cells of SAC individuals (Chapter 5). The latter observation is likely to explain some of the CCR5 density differences noted in Chapter 4, given the high prevalence of the HHC haplotype in the SAC population. In Caucasian populations, the HHC haplotype is associated with delayed disease progression. Thus, one would expect the presence of the HHC haplotype to result in reduced CCR5 expression on CD4+ T cells. Instead SAC individuals with the HHC haplotype had higher CCR5 density on NK cell subsets compared to those that lacked the allele. This result suggests that CCR5-expressing NK cells are likely to be playing a beneficial role in HIV-1 infection control.

Comparison of CCR5 expression between individuals with and without the HHE haplotype did not reveal any significant differences between the two groups (Chapter 5) which was surprising, given its reported association with deleterious HIV-1 disease associated outcomes (AN *et al.* 2000; GONZALEZ *et al.* 1999; LI *et al.* 2005; MARTIN *et al.* 1998; OMETTO *et al.* 2001), as well as the promoter strength analysis results (Chapter 3). Trying to assess the impact of a particular

haplotype on CCR5 expression can be complicated since the influence of the partner allele/haplotype is difficult to exclude. Thus, this lack of association between the HHE haplotype and CCR5 expression could be ascribed to either insufficient sample numbers and/or the influence of the partner allele as most (with the exception of one) individuals were heterozygous for HHE. Ideally, in a study designed to assess the impact of genetic variability on CCR5 expression one would like to analyse and compare individuals homozygous for *CCR5* haplotypes, however the high number of possible *CCR5* haplotype combinations would require a very large study population in order to achieve this.

It has been demonstrated that the effect of CCR5 on HIV-1 disease pathogenesis may be more apparent when assessed in combination with that of its ligand CCL3L. Thus, we assessed *CCL3L* gene copy distribution and CCL3 (CCL3 and CCL3L combined) protein levels in the two study populations (Chapter 6) and investigated three different aspects considering *CCL3L*: (i) the differences in CCL3 production between SAA and SAC individuals given that they differ in *CCL3L* gene copy numbers (along with that of its components *CCL3La* and *CCL3Lb*), (ii) the effects of *CCL3La* gene copy number on CCL3 chemokine production upon PHA stimulation, and (iii) *CCL3La* and *CCL3Lb* gene copy numbers in exposed uninfected (EU) and intrapartum (IP) infected infants born to HIV-1 infected mothers.

Previous reports have determined the median *CCL3L* gene copy number for various populations and populations of Caucasian and African descent are known to differ in gene copy number. In addition to *CCL3L*, we determined the median *CCL3La* and *CCL3Lb* copy numbers for the SAA and SAC populations. Although *CCL3L* copy number provides a good surrogate measure for that of the functional, or chemokine-coding, *CCL3La* gene, knowledge of the *CCL3La* gene copy number is a more accurate measure when trying to relate protein production levels to functional copy number. Our study provides the first report assessing CCL3 production relative to *CCL3La* gene copy number. No correlation between CCL3 production and *CCL3La* copy number was observed. However, given the number of study participants, we cannot conclusively state that there is a lack of association between the two measures. This association needs to be tested with much larger sample numbers and ideally in purified cell subsets to allow a better understanding of cell-specific CCL3 production in relation to these genetic variants.

Against expectations, we observed no difference in CCL3 production between SAA and SAC individuals, despite the significant difference in *CCL3La* gene copy number between the two populations (median copy numbers of 4 and 2, respectively). Furthermore, when CCL3 production was stratified around the population-specific median, no differences were observed between subgroups within each population. Equivalent *CCL3L* gene copy numbers do not associate with equivalent HIV-1 disease-associated phenotypic effects when comparing individuals from different

populations with different population-specific medians (GONZALEZ *et al.* 2005). For example, HIV-1 infected Caucasian and African American individuals with two *CCL3L* copies (the population median in Caucasians but half-median in African Americans), have been shown to differ with regards to rates of disease progression, whereas when African Americans with *CCL3L* copy number of four were compared to Caucasians with *CCL3L* copy number of two, comparable rate of disease progression were observed between the two groups (GONZALEZ *et al.* 2005). Thus, our study offers a plausible explanation for this observation, if the phenotypic effects associated with differences in *CCL3L* gene copy number are in fact attributable to the amount of protein produced. This also raises the question as to how and why CCL3 production is regulated regardless of differences in population *CCL3L* copy number median.

Although *CCL3Lb* has been demonstrated to code for two alternately spliced mRNA transcripts, its function *in vivo* has not yet been determined. It has been suggested that these mRNA transcripts may have a regulatory function (SHOSTAKOVICH-KORETSKAYA *et al.* 2009). We therefore investigated whether differences in *CCL3Lb* gene copy numbers could explain why we had previously observed that at equivalent *CCL3L* copy numbers, IP infants produced less CCL3 relative to EU infants. However, no differences in *CCL3Lb* dose were noted in the two groups which implies that there may be other genetic differences between the two infant groups which are contributing towards the CCL3 production difference. Further studies need to be conducted to assess whether *CCL3Lb* is in fact involved in regulating CCL3 protein production and, if so, by which mechanism.

Due to the high amino acid sequence homology between the CCL3 and CCL3L chemokines, it is not possible to distinguish between the two proteins in ELISA assays. Thus, the necessary assumption that the 2 copy pdg *CCL3* gene results in the production of an equivalent and baseline amount of protein between individuals and that any variation in CCL3 production between individuals is a result of variable *CCL3La* gene copy number, needs to be tested. Given that our two study populations have been shown to differ with respect to polymorphisms across the *CCL3* (and *CCL3L*) genes, it is important to note that this assumption is very likely incorrect (PAXIMADIS *et al.* 2009). Furthermore, some genetic variants in the *CCL3* gene have also been associated with both HIV-1 susceptibility as well as disease control (GONZALEZ *et al.* 2001; HU *et al.* 2012; MODI *et al.* 2006; PAXIMADIS *et al.* 2012; SHRESTHA *et al.* 2006), which adds further evidence for variation in CCL3 protein between individuals. An assay which is able to distinguish between the two chemokines, such as mass spectrometry, may help to unravel the complexity around CCL3/CCL3L production and the regulation thereof.

Since no studies have been conducted in South African populations to investigate the role of *CCR5* haplotypes/genotypes in HIV/AIDS pathogenesis in South Africa, we had to extrapolate from

reports based on populations that are geographically separated but of similar ethnicity, eg. African Americans. However, it is important to note that it is unknown whether genotypes classed as either 'protective' or 'deleterious' in other geographically distinct populations can be inferred to have the same phenotypic effect in the context of HIV-1 infection in South African individuals. Thus, having determined the frequency of *CCR5* haplotypes in healthy, HIV-1 uninfected South African populations, the prevalence of these haplotypes in cohorts of both rapid progressing and LTNP HIV-1 infected individuals can be assessed and compared in future studies to those found in the background population.

Some of the insights gained regarding the genetic determinants of HIV-1/AIDS susceptibility, including *CCR5*, have come from studies of LTNPs. Furthermore, the study of extreme phenotypes of HIV-1 infection, i.e., rapid progressors and LTNPs, allows for the use of smaller numbers in studies aimed at evaluating whether particular immune factors are involved in disease progression. In a pilot study in which we assessed *CCR5* and *CCL3L* genotypic and phenotypic variation in a small group of LTNPs and compared this to that found in our healthy control individuals, we identified some interesting relationships which may serve as markers of disease progression and can be tested in future studies with larger cohorts (Chapter 7). For example, the HHA haplotype, found to be a 'beneficial' haplotype in African Americans in the context of HIV-1 infection was detected in 3/4 SAC LTNP individuals. Since this is a rare allele in SAC individuals (4.25%), it is interesting that it was present at such a high frequency in the SAC LTNP group, thereby suggesting a possible role for this haplotype in delaying disease progression in SAC individuals. Both *CCR5* and *CCL3* expression was found to be lower in LTNP individuals compared to control individuals. This observation suggests that LTNP individuals may, by an unknown mechanism, be able to downregulate expression of these two molecules or be less "activatable" than individuals who progress rapidly. However, comparison to rapid progressing individuals is necessary. Interestingly, PBMCs from LTNP individuals with lower HIV-1 viral loads produced *CCL3* at lower levels compared to individuals with higher viral loads, upon stimulation.

The individuals in the LTNP cohort used in this study are part of a larger cohort of LTNPs and ECs which are currently being recruited by our laboratory. Although all LTNP individuals have had long term follow-up and have been showing markers of favourable disease progression, i.e., consistently high CD4+ T cell counts and/or low viral loads in the absence of antiretroviral treatment, some have been HIV-1 infected for a period of less than 10 years, a time period usually considered necessary for classification as LTNPs (however some studies have used a cut-off of 7 years). In future, as our cohort of South African LTNP individuals is expanded, we will be able to validate our results with larger numbers and with more clearly defined clinical characteristics. The small number of LTNP participants in our study, as well as the absence of rapid progressors to

compare phenotypic measures to, did not allow the assessment of the combined influence of *CCR5* genotypes and *CCL3L* gene copy number on disease progression. Nonetheless, this small study on LTNPs and *CCR5* and *CCL3L*, has validated the importance of these two molecules in HIV-1 disease.

In summary, we have demonstrated that: (i) SAA and SAC individuals differ with respect to both *CCR5* genetic variability and *CCR5* haplotype structure; (ii) genetic variability within both the P1 and P2 promoter regions can affect *CCR5* promoter strength; (iii) *CCR5* expression differs between SAA and SAC individuals in different cell subsets; (iv) *CCR5* haplotypes can differentially affect *CCR5* expression on a range of immune cell types and these differences may be evident as either differences in *CCR5* density or percentage of *CCR5*-expressing cells within a cell subset; (v) SAA and SAC individuals differ with regards to cell activation levels; (vi) despite differences in *CCL3L* gene copy numbers, total CCL3 production by PBMCs does not differ between SAA and SAC individuals; (vii) CCL3 production differences between IP and EU infants cannot be attributed to *CCL3Lb* gene copy number differences; and (viii) LTNP individuals express *CCR5* and CCL3 at significantly lower levels compared to healthy uninfected individuals. Furthermore, in our study of LTNP individuals, we have identified several *CCR5* haplotypes as potential prognostic markers of HIV-1 disease progression. The combinatorial influence of the *CCR5-CCL3L* chemokine-ligand axis will be assessed in future studies using larger sample numbers.

In conclusion, work from this study has contributed to the study of *CCR5* genetic variability and its impact on *CCR5* surface expression, particularly in the context of whether or not a particular haplotype is present in individuals within the same population group. In addition to the impact that polymorphisms within the *CCR5* gene may have on *CCR5* expression levels and thus the number of coreceptors available on the cell surface for virus binding, *CCR5* polymorphisms may also indirectly influence HIV-1 susceptibility and pathogenesis. *CCR5* haplotypes (in combination with *CCL3L* gene copy number) can influence immune reconstitution once HAART has been initiated (AHUJA *et al.* 2008). Furthermore, SNPs within the *CCR5* ORF that result in conformational changes could also potentially influence the binding ability of possible HIV-1 therapeutic agents targeting *CCR5*, such as modified *CCR5* ligands, anti-*CCR5* antibodies, small interfering RNA (siRNA) and zinc finger nucleases, which have been investigated by various groups (Reviewed by SYMONDS *et al.* 2010). Thus, given that many *CCR5* genetic polymorphisms are population specific, studies that assess this genetic variability provide important and useful information. Furthermore, since both *CCR5* density and the percentage of *CCR5*-expressing cells have been shown to be important determinants of *in vivo* HIV-1 susceptibility (MEDITZ *et al.* 2011) and that *CCR5* haplotypes affect HIV-1 disease pathogenesis (GONZALEZ *et al.* 1999), our findings demonstrating that *CCR5* haplotypes may influence both *CCR5* density and the percentage of

CCR5-expressing cells add important information to this pool of knowledge. Importantly, before extrapolating CCR5 expression observed in whole blood studies into risk factors for acquiring HIV-1 infection, it should be kept in mind that the complement and distribution of CCR5-expressing immune cells at mucosal surfaces is quite different from that in peripheral blood, and studies directly assessing how CCR5 haplotypes influence these parameters will shed further light on host susceptibility/resistance to HIV-1.

---

## REFERENCES

---

- AGRAWAL, L., X. LU, J. QINGWEN, Z. VANHORN-ALI, I. V. NICOLESCU *et al.*, 2004 Role for CCR5Delta32 protein in resistance to R5, R5X4, and X4 human immunodeficiency virus type 1 in primary CD4+ cells. *J Virol* **78**: 2277-2287.
- AHUJA, S. K., H. KULKARNI, G. CATANO, B. K. AGAN, J. F. CAMARGO *et al.*, 2008 CCL3L1-CCR5 genotype influences durability of immune recovery during antiretroviral therapy of HIV-1-infected individuals. *Nat Med* **14**: 413-420.
- ALTER, G., and M. ALTFELD, 2009 NK cells in HIV-1 infection: evidence for their role in the control of HIV-1 infection. *J Intern Med* **265**: 29-42.
- AN, P., M. P. MARTIN, G. W. NELSON, M. CARRINGTON, M. W. SMITH *et al.*, 2000 Influence of CCR5 promoter haplotypes on AIDS progression in African-Americans. *Aids* **14**: 2117-2122.
- ANSARI-LARI, M. A., X. M. LIU, M. L. METZKER, A. R. RUT and R. A. GIBBS, 1997 The extent of genetic variation in the CCR5 gene. *Nat Genet* **16**: 221-222.
- AQUARO, S., P. MENTEN, S. STRUYF, P. PROOST, J. VAN DAMME *et al.*, 2001 The LD78beta isoform of MIP-1alpha is the most potent CC-chemokine in inhibiting CCR5-dependent human immunodeficiency virus type 1 replication in human macrophages. *J Virol* **75**: 4402-4406.
- ARENZANA-SEISDEDOS, F., and M. PARMENTIER, 2006 Genetics of resistance to HIV infection: Role of co-receptors and co-receptor ligands. *Semin Immunol* **18**: 387-403.
- ARRANZ, V., M. KRESS and M. ERNOULT-LANGE, 1996 Localization of zinc finger Mok2 gene to mouse chromosome 6, a new region of homology with human chromosome 19. *Mamm Genome* **7**: 77-78.
- AYOUBI, T. A., and W. J. VAN DE VEN, 1996 Regulation of gene expression by alternative promoters. *Faseb J* **10**: 453-460.
- BALKWILL, F., 2004 Cancer and the chemokine network. *Nat Rev Cancer* **4**: 540-550.
- BAMSHAD, M. J., S. MUMMIDI, E. GONZALEZ, S. S. AHUJA, D. M. DUNN *et al.*, 2002 A strong signature of balancing selection in the 5' cis-regulatory region of CCR5. *Proc Natl Acad Sci U S A* **99**: 10539-10544.
- BARCELLOS, L. F., A. M. SCHITO, J. B. RIMMLER, E. VITTINGHOFF, A. SHIH *et al.*, 2000 CC-chemokine receptor 5 polymorphism and age of onset in familial multiple sclerosis. Multiple Sclerosis Genetics Group. *Immunogenetics* **51**: 281-288.
- BARRE-SINOSSI, F., J. C. CHERMANN, F. REY, M. T. NUGEYRE, S. CHAMARET *et al.*, 1983 Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**: 868-871.
- BEGAUD, E., L. CHARTIER, V. MARECHAL, J. IPERO, J. LEAL *et al.*, 2006 Reduced CD4 T cell activation and in vitro susceptibility to HIV-1 infection in exposed uninfected Central Africans. *Retrovirology* **3**: 35.

- BENKIRANE, M., D. Y. JIN, R. F. CHUN, R. A. KOUP and K. T. JEANG, 1997 Mechanism of transdominant inhibition of CCR5-mediated HIV-1 infection by ccr5delta32. *J Biol Chem* **272**: 30603-30606.
- BERETTA, A., L. FURCI, S. BURASTERO, A. COSMA, M. E. DINELLI *et al.*, 1996 HIV-1-specific immunity in persistently seronegative individuals at high risk for HIV infection. *Immunol Lett* **51**: 39-43.
- BERGER, E. A., R. W. DOMS, E. M. FENYO, B. T. KORBER, D. R. LITTMAN *et al.*, 1998 A new classification for HIV-1. *Nature* **391**: 240.
- BERHANU, D., F. MORTARI, S. C. DE ROSA and M. ROEDERER, 2003 Optimized lymphocyte isolation methods for analysis of chemokine receptor expression. *J Immunol Methods* **279**: 199-207.
- BERKMAN, N., M. JOHN, G. ROESEMS, P. J. JOSE, P. J. BARNES *et al.*, 1995 Inhibition of macrophage inflammatory protein-1 alpha expression by IL-10. Differential sensitivities in human blood monocytes and alveolar macrophages. *J Immunol* **155**: 4412-4418.
- BERNSTEIN, H. B., G. WANG, M. C. PLASTERER, J. A. ZACK, P. RAMASASTRY *et al.*, 2009 CD4+ NK cells can be productively infected with HIV, leading to downregulation of CD4 expression and changes in function. *Virology* **387**: 59-66.
- BERRO, R., P. J. KLASSE, D. LASCANO, A. FLEGLER, K. A. NAGASHIMA *et al.*, 2011 Multiple CCR5 conformations on the cell surface are used differentially by human immunodeficiency viruses resistant or sensitive to CCR5 inhibitors. *J Virol* **85**: 8227-8240.
- BHATTACHARYA, T., J. STANTON, E. Y. KIM, K. J. KUNSTMAN, J. P. PHAIR *et al.*, 2009 CCL3L1 and HIV/AIDS susceptibility. *Nat Med* **15**: 1112-1115.
- BITI, R., R. FFRENCH, J. YOUNG, B. BENNETTS, G. STEWART *et al.*, 1997 HIV-1 infection in an individual homozygous for the CCR5 deletion allele. *Nat Med* **3**: 252-253.
- BLANPAIN, C., B. J. DORANZ, J. VAKILI, J. RUCKER, C. GOVAERTS *et al.*, 1999a Multiple charged and aromatic residues in CCR5 amino-terminal domain are involved in high affinity binding of both chemokines and HIV-1 Env protein. *J Biol Chem* **274**: 34719-34727.
- BLANPAIN, C., B. LEE, M. TACKOEN, B. PUFFER, A. BOOM *et al.*, 2000 Multiple nonfunctional alleles of CCR5 are frequent in various human populations. *Blood* **96**: 1638-1645.
- BLANPAIN, C., B. LEE, J. VAKILI, B. J. DORANZ, C. GOVAERTS *et al.*, 1999b Extracellular cysteines of CCR5 are required for chemokine binding, but dispensable for HIV-1 coreceptor activity. *J Biol Chem* **274**: 18902-18908.
- BLANPAIN, C., I. MIGEOTTE, B. LEE, J. VAKILI, B. J. DORANZ *et al.*, 1999c CCR5 binds multiple CC-chemokines: MCP-3 acts as a natural antagonist. *Blood* **94**: 1899-1905.
- BLEUL, C. C., L. WU, J. A. HOXIE, T. A. SPRINGER and C. R. MACKAY, 1997 The HIV coreceptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes. *Proc Natl Acad Sci U S A* **94**: 1925-1930.
- BOESECKE, C., and S. L. PETT, 2012 Clinical studies with chemokine receptor-5 (CCR5)-inhibitors. *Curr Opin HIV AIDS* **7**: 456-462.
- BOUHLAL, H., V. LATRY, M. REQUENA, S. AUBRY, S. V. KAVERI *et al.*, 2005 Natural antibodies to CCR5 from breast milk block infection of macrophages and dendritic cells with primary R5-tropic HIV-1. *J Immunol* **174**: 7202-7209.

- BRIGGS, D. R., D. L. TUTTLE, J. W. SLEASMAN and M. M. GOODENOW, 2000 Envelope V3 amino acid sequence predicts HIV-1 phenotype (co-receptor usage and tropism for macrophages). *Aids* **14**: 2937-2939.
- BUGEJA, M. J., D. R. BOOTH, B. H. BENNETTS, J. GUERIN, J. M. KALDOR *et al.*, 2004 Analysis of the CCL3-L1 gene for association with HIV-1 susceptibility and disease progression. *Aids* **18**: 1069-1071.
- CAMARGO, J. F., M. P. QUINONES, S. MUMMIDI, S. SRINIVAS, A. A. GAITAN *et al.*, 2009 CCR5 expression levels influence NFAT translocation, IL-2 production, and subsequent signaling events during T lymphocyte activation. *J Immunol* **182**: 171-182.
- CAMERON, M. J., G. A. ARREAZA, M. GRATTAN, C. MEAGHER, S. SHARIF *et al.*, 2000 Differential expression of CC chemokines and the CCR5 receptor in the pancreas is associated with progression to type I diabetes. *J Immunol* **165**: 1102-1110.
- CAPOULADE-METAY, C., L. MA, L. X. TRUONG, Y. DUDOIT, P. VERSMISSE *et al.*, 2004 New CCR5 variants associated with reduced HIV coreceptor function in southeast Asia. *Aids* **18**: 2243-2252.
- CARD, C. M., Y. KEYNAN, J. LAJOIE, C. P. BELL, M. DAWOOD *et al.*, 2012 HIV controllers are distinguished by chemokine expression profile and HIV-specific T-cell proliferative potential. *J Acquir Immune Defic Syndr* **59**: 427-437.
- CARGILL, M., D. ALTSHULER, J. IRELAND, P. SKLAR, K. ARDLIE *et al.*, 1999 Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Genet* **22**: 231-238.
- CARPENTER, D., R. S. MCINTOSH, R. J. PLEASS and J. A. ARMOUR, 2012 Functional effects of CCL3L1 copy number. *Genes Immun* **13**: 374-379.
- CARRINGTON, M., M. DEAN, M. P. MARTIN and S. J. O'BRIEN, 1999 Genetics of HIV-1 infection: chemokine receptor CCR5 polymorphism and its consequences. *Hum Mol Genet* **8**: 1939-1945.
- CARRINGTON, M., T. KISSNER, B. GERRARD, S. IVANOV, S. J. O'BRIEN *et al.*, 1997 Novel alleles of the chemokine-receptor gene CCR5. *Am J Hum Genet* **61**: 1261-1267.
- CARTHARIUS, K., K. FRECH, K. GROTE, B. KLOCKE, M. HALTMEIER *et al.*, 2005 MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* **21**: 2933-2942.
- CASTELLINO, F., A. Y. HUANG, G. ALTAN-BONNET, S. STOLL, C. SCHEINECKER *et al.*, 2006 Chemokines enhance immunity by guiding naive CD8<sup>+</sup> T cells to sites of CD4<sup>+</sup> T cell-dendritic cell interaction. *Nature* **440**: 890-895.
- CATANO, G., Z. A. CHYKARENKO, A. MANGANO, J. M. ANAYA, W. HE *et al.*, 2011 Concordance of CCR5 genotypes that influence cell-mediated immunity and HIV-1 disease progression rates. *J Infect Dis* **203**: 263-272.
- CHAN, D. C., and P. S. KIM, 1998 HIV entry and its inhibition. *Cell* **93**: 681-684.
- CHAPLIN, D. D., 2010 Overview of the immune response. *J Allergy Clin Immunol* **125**: S3-23.
- CHEN, Z., D. KWON, Z. JIN, S. MONARD, P. TELFER *et al.*, 1998 Natural infection of a homozygous delta24 CCR5 red-capped mangabey with an R2b-tropic simian immunodeficiency virus. *J Exp Med* **188**: 2057-2065.

- CHOI, W. T., and J. AN, 2011 Biology and clinical relevance of chemokines and chemokine receptors CXCR4 and CCR5 in human diseases. *Exp Biol Med* (Maywood) **236**: 637-647.
- CHOUDHRY, V., M. Y. ZHANG, I. HARRIS, I. A. SIDOROV, B. VU *et al.*, 2006 Increased efficacy of HIV-1 neutralization by antibodies at low CCR5 surface concentration. *Biochem Biophys Res Commun* **348**: 1107-1115.
- CLARK, A. G., 2004 The role of haplotypes in candidate gene studies. *Genet Epidemiol* **27**: 321-333.
- CLEGG, A. O., L. J. ASHTON, R. A. BITI, P. BADHWAR, P. WILLIAMSON *et al.*, 2000 CCR5 promoter polymorphisms, CCR5 59029A and CCR5 59353C, are under represented in HIV-1-infected long-term non-progressors. The Australian Long-Term Non-Progressor Study Group. *Aids* **14**: 103-108.
- CLERICI, M., S. BUTTO, M. LUKWIYA, M. SARESELLA, S. DECLICH *et al.*, 2000 Immune activation in africa is environmentally-driven and is associated with upregulation of CCR5. Italian-Ugandan AIDS Project. *Aids* **14**: 2083-2092.
- CLERICI, M., J. V. GIORGI, C. C. CHOU, V. K. GUDEMAN, J. A. ZACK *et al.*, 1992 Cell-mediated immune response to human immunodeficiency virus (HIV) type 1 in seronegative homosexual men with recent sexual exposure to HIV-1. *J Infect Dis* **165**: 1012-1019.
- COCCHI, F., A. L. DEVICO, A. GARZINO-DEMO, S. K. ARYA, R. C. GALLO *et al.*, 1995 Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells. *Science* **270**: 1811-1815.
- COCCHI, F., A. L. DEVICO, R. YARCHOAN, R. REDFIELD, F. CLEGHORN *et al.*, 2000 Higher macrophage inflammatory protein (MIP)-1alpha and MIP-1beta levels from CD8+ T cells are associated with asymptomatic HIV-1 infection. *Proc Natl Acad Sci U S A* **97**: 13812-13817.
- COHEN, C. R., A. B. MOSCICKI, M. E. SCOTT, Y. MA, S. SHIBOSKI *et al.*, 2010 Increased levels of immune activation in the genital tract of healthy young women from sub-Saharan Africa. *Aids* **24**: 2069-2074.
- COHEN, O. J., S. PAOLUCCI, S. M. BENDE, M. DAUCHER, H. MORIUCHI *et al.*, 1998 CXCR4 and CCR5 genetic polymorphisms in long-term nonprogressive human immunodeficiency virus infection: lack of association with mutations other than CCR5-Delta32. *J Virol* **72**: 6215-6217.
- COLOBRAN, R., E. PEDROSA, L. CARRETERO-IGLESIA and M. JUAN, 2010 Copy number variation in chemokine superfamily: the complex scene of CCL3L-CCL4L genes in health and disease. *Clin Exp Immunol* **162**: 41-52.
- COLOBRAN, R., R. PUJOL-BORRELL, M. P. ARMENGOL and M. JUAN, 2007 The chemokine network. I. How the genomic organization of chemokines contains clues for deciphering their functional complexity. *Clin Exp Immunol* **148**: 208-217.
- CONTENTO, R. L., B. MOLON, C. BOULARAN, T. POZZAN, S. MANES *et al.*, 2008 CXCR4-CCR5: a couple modulating T cell functions. *Proc Natl Acad Sci U S A* **105**: 10101-10106.
- COOK, D. N., M. A. BECK, T. M. COFFMAN, S. L. KIRBY, J. F. SHERIDAN *et al.*, 1995 Requirement of MIP-1 alpha for an inflammatory response to viral infection. *Science* **269**: 1583-1585.

- DE RODA HUSMAN, A. M., H. BLAAK, M. BROUWER and H. SCHUITEMAKER, 1999 CC chemokine receptor 5 cell-surface expression in relation to CC chemokine receptor 5 genotype and the clinical course of HIV-1 infection. *J Immunol* **163**: 4597-4603.
- DEAN, M., M. CARRINGTON, C. WINKLER, G. A. HUTTLEY, M. W. SMITH *et al.*, 1996 Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CCR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science* **273**: 1856-1862.
- DEEKS, S. G., B. AUTRAN, B. BERKHOUT, M. BENKIRANE, S. CAIRNS *et al.*, 2012 Towards an HIV cure: a global scientific strategy. *Nat Rev Immunol* **12**: 607-614.
- DEGENHARDT, J. D., P. DE CANDIA, A. CHABOT, S. SCHWARTZ, L. HENDERSON *et al.*, 2009 Copy number variation of CCL3-like genes affects rate of progression to simian-AIDS in Rhesus Macaques (*Macaca mulatta*). *PLoS Genet* **5**: e1000346.
- DENG, H., R. LIU, W. ELLMEIER, S. CHOE, D. UNUTMAZ *et al.*, 1996 Identification of a major co-receptor for primary isolates of HIV-1. *Nature* **381**: 661-666.
- DESMETZ, C., Y. L. LIN, C. METTLING, P. PORTALES, H. RABESANDRATANA *et al.*, 2006 The strength of the chemotactic response to a CCR5 binding chemokine is determined by the level of cell surface CCR5 density. *Immunology* **119**: 551-561.
- DOLAN, M. J., H. KULKARNI, J. F. CAMARGO, W. HE, A. SMITH *et al.*, 2007 CCL3L1 and CCR5 influence cell-mediated immunity and affect HIV-AIDS pathogenesis via viral entry-independent mechanisms. *Nat Immunol* **8**: 1324-1336.
- DONAHUE, J. P., M. L. VETTER, N. A. MUKHTAR and R. T. D'AQUILA, 2008 The HIV-1 Vif PPLP motif is necessary for human APOBEC3G binding and degradation. *Virology* **377**: 49-53.
- DONG, H. F., K. WIGMORE, M. N. CARRINGTON, M. DEAN, J. A. TURPIN *et al.*, 2005 Variants of CCR5, which are permissive for HIV-1 infection, show distinct functional responses to CCL3, CCL4 and CCL5. *Genes Immun* **6**: 609-619.
- DRAGIC, T., V. LITWIN, G. P. ALLAWAY, S. R. MARTIN, Y. HUANG *et al.*, 1996 HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* **381**: 667-673.
- DREUILLET, C., J. TILLIT, M. KRESS and M. ERNOULT-LANGE, 2002 In vivo and in vitro interaction between human transcription factor MOK2 and nuclear lamin A/C. *Nucleic Acids Res* **30**: 4634-4642.
- EASTERBROOK, P. J., T. ROSTRON, N. IVES, M. TROOP, B. G. GAZZARD *et al.*, 1999 Chemokine receptor polymorphisms and human immunodeficiency virus disease progression. *J Infect Dis* **180**: 1096-1105.
- ERI, R., J. R. JONSSON, N. PANDEYA, D. M. PURDIE, A. D. CLOUSTON *et al.*, 2004 CCR5-Delta32 mutation is strongly associated with primary sclerosing cholangitis. *Genes Immun* **5**: 444-450.
- ERNOULT-LANGE, M., V. ARRANZ, M. LE CONIAT, R. BERGER and M. KRESS, 1995 Human and mouse Kruppel-like (MOK2) orthologue genes encode two different zinc finger proteins. *J Mol Evol* **41**: 784-794.
- ESLAHPAZIR, J., M. A. JENABIAN, H. BOUHLAL, H. HOCINI, C. CARBONNEIL *et al.*, 2008 Infection of macrophages and dendritic cells with primary R5-tropic human immunodeficiency virus type

- 1 inhibited by natural polyreactive anti-CCR5 antibodies purified from cervicovaginal secretions. *Clin Vaccine Immunol* **15**: 872-884.
- EUGEN-OLSEN, J., A. K. IVERSEN, T. L. BENFIELD, U. KOPPELHUS and P. GARRED, 1998 Chemokine receptor CCR2b 64I polymorphism and its relation to CD4 T-cell counts and disease progression in a Danish cohort of HIV-infected individuals. Copenhagen AIDS cohort. *J Acquir Immune Defic Syndr Hum Retrovirol* **18**: 110-116.
- EUGEN-OLSEN, J., A. K. IVERSEN, P. GARRED, U. KOPPELHUS, C. PEDERSEN *et al.*, 1997 Heterozygosity for a deletion in the CCR5 gene leads to prolonged AIDS-free survival and slower CD4 T-cell decline in a cohort of HIV-seropositive individuals. *Aids* **11**: 305-310.
- EVANS, J. H., A. HOROWITZ, M. MEHRABI, E. L. WISE, J. E. PEASE *et al.*, 2011 A distinct subset of human NK cells expressing HLA-DR expand in response to IL-2 and can aid immune responses to BCG. *Eur J Immunol* **41**: 1924-1933.
- EVANS, R. L., T. J. FALDETTA, R. E. HUMPHREYS, D. M. PRATT, E. J. YUNIS *et al.*, 1978 Peripheral human T cells sensitized in mixed leukocyte culture synthesize and express Ia-like antigens. *J Exp Med* **148**: 1440-1445.
- FENG, Y., C. C. BRODER, P. E. KENNEDY and E. A. BERGER, 1996 HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* **272**: 872-877.
- FIELD, S. F., J. M. HOWSON, L. M. MAIER, S. WALKER, N. M. WALKER *et al.*, 2009 Experimental aspects of copy number variant assays at CCL3L1. *Nat Med* **15**: 1115-1117.
- FISCHEREDER, M., B. LUCKOW, B. HOCHER, R. P. WUTHRICH, U. ROTHENPIELER *et al.*, 2001 CCR5 chemokine receptor 5 and renal-transplant survival. *Lancet* **357**: 1758-1761.
- FOLEFOC, A. T., B. J. FROMME, A. A. KATZ and C. A. FLANAGAN, 2010 South African mutations of the CCR5 coreceptor for HIV modify interaction with chemokines and HIV Envelope protein. *J Acquir Immune Defic Syndr* **54**: 352-359.
- FOUCHIER, R. A., M. GROENINK, N. A. KOOTSTRA, M. TERSMETTE, H. G. HUISMAN *et al.*, 1992 Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule. *J Virol* **66**: 3183-3187.
- FOWKE, K. R., N. J. NAGELKERKE, J. KIMANI, J. N. SIMONSEN, A. O. ANZALA *et al.*, 1996 Resistance to HIV-1 infection among persistently seronegative prostitutes in Nairobi, Kenya. *Lancet* **348**: 1347-1351.
- FREEMAN, E. E., H. A. WEISS, J. R. GLYNN, P. L. CROSS, J. A. WHITWORTH *et al.*, 2006 Herpes simplex virus 2 infection increases HIV acquisition in men and women: systematic review and meta-analysis of longitudinal studies. *Aids* **20**: 73-83.
- FUNKE, J., R. DURR, U. DIETRICH and J. KOCH, 2011 Natural killer cells in HIV-1 infection: a double-edged sword. *AIDS Rev* **13**: 67-76.
- FURIE, M. B., and G. J. RANDOLPH, 1995 Chemokines and tissue injury. *Am J Pathol* **146**: 1287-1301.
- GALLO, R. C., S. Z. SALAHUDDIN, M. POPOVIC, G. M. SHEARER, M. KAPLAN *et al.*, 1984 Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224**: 500-503.

- GALVANI, A. P., and J. NOVEMBRE, 2005 The evolutionary history of the CCR5-Delta32 HIV-resistance mutation. *Microbes Infect* **7**: 302-309.
- GALVANI, A. P., and M. SLATKIN, 2003 Evaluating plague and smallpox as historical selective pressures for the CCR5-Delta 32 HIV-resistance allele. *Proc Natl Acad Sci U S A* **100**: 15276-15279.
- GAO, F., E. BAILES, D. L. ROBERTSON, Y. CHEN, C. M. RODENBURG *et al.*, 1999 Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* **397**: 436-441.
- GARZINO-DEMO, A., R. B. MOSS, J. B. MARGOLICK, F. CLEGHORN, A. SILL *et al.*, 1999 Spontaneous and antigen-induced production of HIV-inhibitory beta-chemokines are associated with AIDS-free status. *Proc Natl Acad Sci U S A* **96**: 11986-11991.
- GERVAIX, A., J. NICOLAS, P. PORTALES, K. POSFAY-BARBE, C. A. WYLER *et al.*, 2002 Response to treatment and disease progression linked to CD4+ T cell surface CC chemokine receptor 5 density in human immunodeficiency virus type 1 vertical infection. *J Infect Dis* **185**: 1055-1061.
- GILTAY, E. J., J. C. FONK, B. M. VON BLOMBERG, H. A. DREXHAGE, C. SCHALKWIJK *et al.*, 2000 In vivo effects of sex steroids on lymphocyte responsiveness and immunoglobulin levels in humans. *J Clin Endocrinol Metab* **85**: 1648-1657.
- GIRI, R. K., V. RAJAGOPAL and V. K. KALRA, 2004 Curcumin, the active constituent of turmeric, inhibits amyloid peptide-induced cytochemokine gene expression and CCR5-mediated chemotaxis of THP-1 monocytes by modulating early growth response-1 transcription factor. *J Neurochem* **91**: 1199-1210.
- GONZALEZ, E., M. BAMSHAD, N. SATO, S. MUMMIDI, R. DHANDA *et al.*, 1999 Race-specific HIV-1 disease-modifying effects associated with CCR5 haplotypes. *Proc Natl Acad Sci U S A* **96**: 12004-12009.
- GONZALEZ, E., R. DHANDA, M. BAMSHAD, S. MUMMIDI, R. GEEVARGHESE *et al.*, 2001 Global survey of genetic variation in CCR5, RANTES, and MIP-1alpha: impact on the epidemiology of the HIV-1 pandemic. *Proc Natl Acad Sci U S A* **98**: 5199-5204.
- GONZALEZ, E., H. KULKARNI, H. BOLIVAR, A. MANGANO, R. SANCHEZ *et al.*, 2005 The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* **307**: 1434-1440.
- GORNALUSSE, G., S. MUMMIDI, W. HE, G. SILVESTRI, M. BAMSHAD *et al.*, 2009 CCL3L Copy number variation and the co-evolution of primate and viral genomes. *PLoS Genet* **5**: e1000359.
- GOTTLIEB, M. S., R. SCHROFF, H. M. SCHANKER, J. D. WEISMAN, P. T. FAN *et al.*, 1981 *Pneumocystis carinii* pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. *N Engl J Med* **305**: 1425-1431.
- GRAS, G., and M. KAUL, 2010 Molecular mechanisms of neuroinvasion by monocytes-macrophages in HIV-1 infection. *Retrovirology* **7**: 30.
- GROOT, F., S. WELSCH and Q. J. SATTENTAU, 2008 Efficient HIV-1 transmission from macrophages to T cells across transient virological synapses. *Blood* **111**: 4660-4663.

- GRUNHAGE, F., J. NATTERMANN, O. A. GRESSNER, H. E. WASMUTH, C. HELLERBRAND *et al.*, 2010 Lower copy numbers of the chemokine CCL3L1 gene in patients with chronic hepatitis C. *J Hepatol* **52**: 153-159.
- GUIGNARD, F., C. COMBADIÈRE, H. L. TIFFANY and P. M. MURPHY, 1998 Gene organization and promoter function for CC chemokine receptor 5 (CCR5). *J Immunol* **160**: 985-992.
- GUO, S. W., and E. A. THOMPSON, 1992 Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**: 361-372.
- HARADA, H., Y. GOTO, T. OHNO, S. SUZU and S. OKADA, 2007 Proliferative activation up-regulates expression of CD4 and HIV-1 co-receptors on NK cells and induces their infection with HIV-1. *Eur J Immunol* **37**: 2148-2155.
- HAYES, V. M., D. C. PETERSEN, T. J. SCRIBA, M. ZEIER, A. GRIMWOOD *et al.*, 2002 African-based CCR5 single-nucleotide polymorphism associated with HIV-1 disease progression. *Aids* **16**: 2229-2231.
- HE, W., H. KULKARNI, J. CASTIBLANCO, C. SHIMIZU, U. ALUYEN *et al.*, 2009 Reply to: "Experimental aspects of copy number variant assays at CCL3L1". *Nat Med* **15**: 1117-1120.
- HEENEY, J. L., A. G. DALGLEISH and R. A. WEISS, 2006 Origins of HIV and the evolution of resistance to AIDS. *Science* **313**: 462-466.
- HEREDIA, A., B. GILLIAM, A. DEVICO, N. LE, D. BAMBA *et al.*, 2007 CCR5 density levels on primary CD4 T cells impact the replication and Enfuvirtide susceptibility of R5 HIV-1. *Aids* **21**: 1317-1322.
- HEREDIA, A., O. LATINOVIC, R. C. GALLO, G. MELIKYAN, M. REITZ *et al.*, 2008 Reduction of CCR5 with low-dose rapamycin enhances the antiviral activity of vicriviroc against both sensitive and drug-resistant HIV-1. *Proc Natl Acad Sci U S A* **105**: 20476-20481.
- HIRASHIMA, M., T. ONO, M. NAKAO, H. NISHI, A. KIMURA *et al.*, 1992 Nucleotide sequence of the third cytokine LD78 gene and mapping of all three LD78 gene loci to human chromosome 17. *DNA Seq* **3**: 203-212.
- HIRSCH, V. M., R. A. OLMSTED, M. MURPHEY-CORB, R. H. PURCELL and P. R. JOHNSON, 1989 An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature* **339**: 389-392.
- HLADIK, F., H. LIU, E. SPEELMON, D. LIVINGSTON-ROSANOFF, S. WILSON *et al.*, 2005 Combined effect of CCR5-Delta32 heterozygosity and the CCR5 promoter polymorphism -2459 A/G on CCR5 expression and resistance to human immunodeficiency virus type 1 transmission. *J Virol* **79**: 11677-11684.
- HOWARD, O. M., A. K. SHIRAKAWA, J. A. TURPIN, A. MAYNARD, G. J. TOBIN *et al.*, 1999 Naturally occurring CCR5 extracellular and transmembrane domain variants affect HIV-1 Co-receptor and ligand binding function. *J Biol Chem* **274**: 16228-16234.
- HU, L., W. SONG, I. BRILL, J. MULENGA, S. ALLEN *et al.*, 2012 Genetic variations and heterosexual HIV-1 infection: analysis of clustered genes encoding CC-motif chemokine ligands. *Genes Immun* **13**: 202-205.
- HUMMEL, S., D. SCHMIDT, B. KREMEYER, B. HERRMANN and M. OPPERMANN, 2005 Detection of the CCR5-Delta32 HIV resistance gene in Bronze Age skeletons. *Genes Immun* **6**: 371-374.

- HUNT, P. W., 2009 Natural control of HIV-1 replication and long-term nonprogression: overlapping but distinct phenotypes. *J Infect Dis* **200**: 1636-1638.
- HUTTER, G., D. NOWAK, M. MOSSNER, S. GANEPOLA, A. MUSSIG *et al.*, 2009 Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med* **360**: 692-698.
- IGLESIAS-USSEL, M. D., and F. ROMERIO, 2011 HIV reservoirs: the new frontier. *AIDS Rev* **13**: 13-29.
- IOANNIDIS, J. P., D. G. CONTOPOULOS-IOANNIDIS, P. S. ROSENBERG, J. J. GOEDERT, A. DE ROSSI *et al.*, 2003 Effects of CCR5-delta32 and CCR2-64I alleles on disease progression of perinatally HIV-1-infected children: an international meta-analysis. *Aids* **17**: 1631-1638.
- IOANNIDIS, J. P., P. S. ROSENBERG, J. J. GOEDERT, L. J. ASHTON, T. L. BENFIELD *et al.*, 2001 Effects of CCR5-Delta32, CCR2-64I, and SDF-1 3'A alleles on HIV-1 disease progression: An international meta-analysis of individual-patient data. *Ann Intern Med* **135**: 782-795.
- JIANG, D., S. MUMMIDI, S. K. AHUJA and H. W. JARRETT, 2011 CCR5 promoter haplotype transcription complex characterization. *J Health Care Poor Underserved* **22**: 73-90.
- JIANG, Y., Z. ZHANG, Y. DIAO, X. JIN, W. SHI *et al.*, 2008 Expression of chemokine receptors on natural killer cells in HIV-infected individuals. *Cell Immunol* **251**: 19-24.
- JIN, Q., L. AGRAWAL, L. MEYER, R. TUBIANA, I. THEODOROU *et al.*, 2008 CCR5Delta32 59537-G/A promoter polymorphism is associated with low translational efficiency and the loss of CCR5Delta32 protective effects. *J Virol* **82**: 2418-2426.
- JOHN, G. C., T. BIRD, J. OVERBAUGH, R. NDUATI, D. MBORI-NGACHA *et al.*, 2001 CCR5 promoter polymorphisms in a Kenyan perinatal human immunodeficiency virus type 1 cohort: association with increased 2-year maternal mortality. *J Infect Dis* **184**: 89-92.
- KAKKANAIH, V. N., E. A. OJO-AMAIZE and J. B. PETER, 1998 Concentrations of circulating beta-chemokines do not correlate with viral load in human immunodeficiency virus-infected individuals. *Clin Diagn Lab Immunol* **5**: 499-502.
- KALINKOVICH, A., G. BORKOW, Z. WEISMAN, A. TSIMANIS, M. STEIN *et al.*, 2001 Increased CCR5 and CXCR4 expression in Ethiopians living in Israel: environmental and constitutive factors. *Clin Immunol* **100**: 107-117.
- KALINKOVICH, A., Z. WEISMAN, Q. LENG, G. BORKOW, M. STEIN *et al.*, 1999 Increased CCR5 expression with decreased beta chemokine secretion in Ethiopians: relevance to AIDS in Africa. *J Hum Virol* **2**: 283-289.
- KASLOW, R. A., T. DORAK and J. J. TANG, 2005 Influence of host genetic variation on susceptibility to HIV type 1 infection. *J Infect Dis* **191 Suppl 1**: S68-77.
- KAWAMURA, T., F. O. GULDEN, M. SUGAYA, D. T. MCNAMARA, D. L. BORRIS *et al.*, 2003 R5 HIV productively infects Langerhans cells, and infection levels are regulated by compound CCR5 polymorphisms. *Proc Natl Acad Sci U S A* **100**: 8401-8406.
- KETAS, T. J., S. E. KUHMANN, A. PALMER, J. ZURITA, W. HE *et al.*, 2007 Cell surface expression of CCR5 and other host factors influence the inhibition of HIV-1 infection of human lymphocytes by CCR5 ligands. *Virology* **364**: 281-290.

- KHAN, I. A., S. Y. THOMAS, M. M. MORETTO, F. S. LEE, S. A. ISLAM *et al.*, 2006 CCR5 is essential for NK cell trafficking and host survival following *Toxoplasma gondii* infection. *PLoS Pathog* **2**: e49.
- KIVISAKK, P., C. TREBST, J. C. LEE, B. H. TUCKY, R. A. RUDICK *et al.*, 2003 Expression of CCR2, CCR5, and CXCR3 by CD4+ T cells is stable during a 2-year longitudinal study but varies widely between individuals. *J Neurovirol* **9**: 291-299.
- KLEIN, R. S., 2008 A moving target: the multiple roles of CCR5 in infectious diseases. *J Infect Dis* **197**: 183-186.
- KNUDSEN, T. B., T. B. KRISTIANSEN, T. L. KATZENSTEIN and J. EUGEN-OLSEN, 2001 Adverse effect of the CCR5 promoter -2459A allele on HIV-1 disease progression. *J Med Virol* **65**: 441-444.
- KO, H. S., S. M. FU, R. J. WINCHESTER, D. T. YU and H. G. KUNKEL, 1979 Ia determinants on stimulated human T lymphocytes. Occurrence on mitogen- and antigen-activated T cells. *J Exp Med* **150**: 246-255.
- KONING, F. A., S. A. OTTO, M. D. HAZENBERG, L. DEKKER, M. PRINS *et al.*, 2005 Low-level CD4+ T cell activation is associated with low susceptibility to HIV-1 infection. *J Immunol* **175**: 6117-6122.
- KOSTRIKIS, L. G., Y. HUANG, J. P. MOORE, S. M. WOLINSKY, L. ZHANG *et al.*, 1998 A chemokine receptor CCR2 allele delays HIV-1 disease progression and is associated with a CCR5 promoter mutation. *Nat Med* **4**: 350-353.
- KOSTRIKIS, L. G., A. U. NEUMANN, B. THOMSON, B. T. KORBER, P. MCHARDY *et al.*, 1999 A polymorphism in the regulatory region of the CC-chemokine receptor 5 gene influences perinatal transmission of human immunodeficiency virus type 1 to African-American infants. *J Virol* **73**: 10264-10271.
- KOTTILIL, S., T. W. CHUN, S. MOIR, S. LIU, M. MCLAUGHLIN *et al.*, 2003 Innate immunity in human immunodeficiency virus infection: effect of viremia on natural killer cell function. *J Infect Dis* **187**: 1038-1045.
- KOTTILIL, S., K. SHIN, M. PLANTA, M. MCLAUGHLIN, C. W. HALLAHAN *et al.*, 2004 Expression of chemokine and inhibitory receptors on natural killer cells: effect of immune activation and HIV viremia. *J Infect Dis* **189**: 1193-1198.
- KROWKA, J. F., M. L. GESNER, M. S. ASCHER and H. W. SHEPPARD, 1997 Lack of associations of chemotactic cytokines with viral burden, disease progression, or lymphocyte subsets in HIV-infected individuals. *Clin Immunol Immunopathol* **85**: 21-27.
- KUHN, L., D. B. SCHRAMM, S. DONNINGER, S. MEDDOWS-TAYLOR, A. H. COOVADIA *et al.*, 2007 African infants' CCL3 gene copies influence perinatal HIV transmission in the absence of maternal nevirapine. *Aids* **21**: 1753-1761.
- KWONG, P. D., R. WYATT, J. ROBINSON, R. W. SWEET, J. SODROSKI *et al.*, 1998 Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**: 648-659.
- LAMA, J., and V. PLANELLES, 2007 Host factors influencing susceptibility to HIV infection and AIDS progression. *Retrovirology* **4**: 52.

- LARSEN, M. H., L. W. THORNER, R. ZINYAMA, J. AMSTRUP, P. KALLESTRUP *et al.*, 2012 CCL3L gene copy number and survival in an HIV-1 infected Zimbabwean population. *Infect Genet Evol* **12**: 1087-1093.
- LEDERMAN, M. M., A. PENN-NICHOLSON, M. CHO and D. MOSIER, 2006 Biology of CCR5 and its role in HIV infection and treatment. *JAMA* **296**: 815-826.
- LEE, B., M. SHARRON, C. BLANPAIN, B. J. DORANZ, J. VAKILI *et al.*, 1999a Epitope mapping of CCR5 reveals multiple conformational states and distinct but overlapping structures involved in chemokine and coreceptor function. *J Biol Chem* **274**: 9617-9626.
- LEE, B., M. SHARRON, L. J. MONTANER, D. WEISSMAN and R. W. DOMS, 1999b Quantification of CD4, CCR5, and CXCR4 levels on lymphocyte subsets, dendritic cells, and differentially conditioned monocyte-derived macrophages. *Proc Natl Acad Sci U S A* **96**: 5215-5220.
- LEHNER, T., 2002 The role of CCR5 chemokine ligands and antibodies to CCR5 coreceptors in preventing HIV infection. *Trends Immunol* **23**: 347-351.
- LEVY, J. A., 2006 HIV pathogenesis: knowledge gained after two decades of research. *Adv Dent Res* **19**: 10-16.
- LEWONTIN, R. C., 1964 The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. *Genetics* **49**: 49-67.
- LI, M., R. SONG, S. MASCIOTRA, V. SORIANO, T. J. SPIRA *et al.*, 2005 Association of CCR5 human haplogroup E with rapid HIV type 1 disease progression. *AIDS Res Hum Retroviruses* **21**: 111-115.
- LIAU, S. W., R. MICKEY, P. ROMANO and T. D. LEE, 1984 Study of the HLA system in the Haitian population. *Tissue Antigens* **23**: 308-313.
- LIM, S. Y., T. CHAN, R. S. GELMAN, J. B. WHITNEY, K. L. O'BRIEN *et al.*, 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 Genet* **6**: e1000997.
- LIN, Y. L., C. METTLING, P. PORTALES, B. REANT, V. ROBERT-HEBMANN *et al.*, 2006 The efficiency of R5 HIV-1 infection is determined by CD4 T-cell surface CCR5 density through G alpha i-protein signalling. *Aids* **20**: 1369-1377.
- LIN, Y. L., C. METTLING, P. PORTALES, J. REYNES, J. CLOT *et al.*, 2002 Cell surface CCR5 density determines the postentry efficiency of R5 HIV-1 infection. *Proc Natl Acad Sci U S A* **99**: 15590-15595.
- LIN, Y. L., C. METTLING, P. PORTALES, R. ROUZIER, J. CLOT *et al.*, 2008 The chemokine CCL5 regulates the in vivo cell surface expression of its receptor, CCR5. *Aids* **22**: 430-432.
- LIU, R., W. A. PAXTON, S. CHOE, D. CERADINI, S. R. MARTIN *et al.*, 1996 Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* **86**: 367-377.
- LIU, R., X. ZHAO, T. A. GURNEY and N. R. LANDAU, 1998 Functional analysis of the proximal CCR5 promoter. *AIDS Res Hum Retroviruses* **14**: 1509-1519.
- LIU, S., L. YAO, D. DING and H. ZHU, 2010 CCL3L1 copy number variation and susceptibility to HIV-1 infection: a meta-analysis. *PLoS ONE* **5**: e15778.

- LIU, Y., F. L. LIU, Y. HE, L. LI, S. LI *et al.*, 2012 The genetic variation of CCR5, CXCR4 and SDF-1 in three Chinese ethnic populations. *Infect Genet Evol* **12**: 1072-1078.
- LOPALCO, L., C. BARASSI, C. PASTORI, R. LONGHI, S. E. BURASTERO *et al.*, 2000 CCR5-reactive antibodies in seronegative partners of HIV-seropositive individuals down-modulate surface CCR5 in vivo and neutralize the infectivity of R5 strains of HIV-1 In vitro. *J Immunol* **164**: 3426-3433.
- LUSTER, A. D., 2002 The role of chemokines in linking innate and adaptive immunity. *Curr Opin Immunol* **14**: 129-135.
- MACK, M., B. LUCKOW, P. J. NELSON, J. CIHAK, G. SIMMONS *et al.*, 1998 Aminooxypentane-RANTES induces CCR5 internalization but inhibits recycling: a novel inhibitory mechanism of HIV infectivity. *J Exp Med* **187**: 1215-1224.
- MADDON, P. J., A. G. DALGLEISH, J. S. MCDOUGAL, P. R. CLAPHAM, R. A. WEISS *et al.*, 1986 The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* **47**: 333-348.
- MAGIEROWSKA, M., V. LEPAGE, T. X. LIEN, N. T. LAN, M. GUILLOTTEL *et al.*, 1999a Novel variant of the CCR5 gene in a Vietnamese population. *Microbes Infect* **1**: 123-124.
- MAGIEROWSKA, M., I. THEODOROU, P. DEBRE, F. SANSON, B. AUTRAN *et al.*, 1999b Combined genotypes of CCR5, CCR2, SDF1, and HLA genes can predict the long-term nonprogressor status in human immunodeficiency virus-1-infected individuals. *Blood* **93**: 936-941.
- MALHOTRA, R., L. HU, W. SONG, I. BRILL, J. MULENGA *et al.*, 2011 Association of chemokine receptor gene (CCR2-CCR5) haplotypes with acquisition and control of HIV-1 infection in Zambians. *Retrovirology* **8**: 22.
- MAMTANI, M., S. MUMMIDI, V. RAMSURAN, M. H. PHAM, R. MALDONADO *et al.*, 2011 Influence of variations in CCL3L1 and CCR5 on tuberculosis in a northwestern Colombian population. *J Infect Dis* **203**: 1590-1594.
- MANGANO, A., E. GONZALEZ, R. DHANDA, G. CATANO, M. BAMSHAD *et al.*, 2001 Concordance between the CC chemokine receptor 5 genetic determinants that alter risks of transmission and disease progression in children exposed perinatally to human immunodeficiency virus. *J Infect Dis* **183**: 1574-1585.
- MARODON, G., N. R. LANDAU and D. N. POSNETT, 1999 Altered expression of CD4, CD54, CD62L, and CCR5 in primary lymphocytes productively infected with the human immunodeficiency virus. *AIDS Res Hum Retroviruses* **15**: 161-171.
- MARTIN, M. P., M. DEAN, M. W. SMITH, C. WINKLER, B. GERRARD *et al.*, 1998 Genetic acceleration of AIDS progression by a promoter variant of CCR5. *Science* **282**: 1907-1911.
- MASQUELIER, C., J. Y. SERVAIS, E. RUSANGANWA, F. ROMAN, E. HAVUGA *et al.*, 2007 A novel 24-base pair deletion in the coding region of CCR5 in an African population. *Aids* **21**: 111-113.
- MCDERMOTT, D. H., P. A. ZIMMERMAN, F. GUIGNARD, C. A. KLEEGERGER, S. F. LEITMAN *et al.*, 1998 CCR5 promoter polymorphism and HIV-1 disease progression. Multicenter AIDS Cohort Study (MACS). *Lancet* **352**: 866-870.

- MCDUGAL, J. S., M. S. KENNEDY, J. M. SLIGH, S. P. CORT, A. MAWLE *et al.*, 1986 Binding of HTLV-III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule. *Science* **231**: 382-385.
- MCKENZIE, S. W., G. DALLALIO, M. NORTH, P. FRAME and R. T. MEANS, JR., 1996 Serum chemokine levels in patients with non-progressing HIV infection. *Aids* **10**: F29-33.
- MCMAHON, E. J., D. N. COOK, K. SUZUKI and G. K. MATSUSHIMA, 2001 Absence of macrophage-inflammatory protein-1alpha delays central nervous system demyelination in the presence of an intact blood-brain barrier. *J Immunol* **167**: 2964-2971.
- MEDDOWS-TAYLOR, S., S. L. DONNINGER, M. PAXIMADIS, D. B. SCHRAMM, F. S. ANTHONY *et al.*, 2006 Reduced ability of newborns to produce CCL3 is associated with increased susceptibility to perinatal human immunodeficiency virus 1 transmission. *J Gen Virol* **87**: 2055-2065.
- MEDITZ, A. L., M. K. HAAS, J. M. FOLKVORD, K. MELANDER, R. YOUNG *et al.*, 2011 HLA-DR+ CD38+ CD4+ T lymphocytes have elevated CCR5 expression and produce the majority of R5-tropic HIV-1 RNA in vivo. *J Virol* **85**: 10189-10200.
- MENDONCA, V. A., L. C. MALAQUIAS, G. E. BRITO-MELO, A. CASTELO-BRANCO, C. M. ANTUNES *et al.*, 2007 Differentiation of patients with leprosy from non-infected individuals by the chemokine eotaxin/CCL11. *Am J Trop Med Hyg* **77**: 547-550.
- MENTEN, P., S. STRUYF, E. SCHUTYSER, A. WUYTS, E. DE CLERCQ *et al.*, 1999 The LD78beta isoform of MIP-1alpha is the most potent CCR5 agonist and HIV-1-inhibiting chemokine. *J Clin Invest* **104**: R1-5.
- MENTEN, P., A. WUYTS and J. VAN DAMME, 2002 Macrophage inflammatory protein-1. *Cytokine Growth Factor Rev* **13**: 455-481.
- MICHAEL, N. L., L. G. LOUIE, A. L. ROHRBAUGH, K. A. SCHULTZ, D. E. DAYHOFF *et al.*, 1997 The role of CCR5 and CCR2 polymorphisms in HIV-1 transmission and disease progression. *Nat Med* **3**: 1160-1162.
- MICHAELS, S. H., R. CLARK and P. KISSINGER, 1998 Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *N Engl J Med* **339**: 405-406.
- MIKHAIL, M., B. WANG and N. K. SAKSENA, 2003 Mechanisms involved in non-progressive HIV disease. *AIDS Rev* **5**: 230-244.
- MILLER, M. P., 1997 Tools for population genetic analyses version 1.3.
- MO, R., J. CHEN, A. GROLLEAU-JULIUS, H. S. MURPHY, B. C. RICHARDSON *et al.*, 2005 Estrogen regulates CCR gene expression and function in T lymphocytes. *J Immunol* **174**: 6023-6029.
- MODI, W. S., 2004 CCL3L1 and CCL4L1 chemokine genes are located in a segmental duplication at chromosome 17q12. *Genomics* **83**: 735-738.
- MODI, W. S., J. BERGERON and M. SANFORD, 2001 The human MIP-1beta chemokine is encoded by two paralogous genes, ACT-2 and LAG-1. *Immunogenetics* **53**: 543-549.
- MODI, W. S., J. LAUTENBERGER, P. AN, K. SCOTT, J. J. GOEDERT *et al.*, 2006 Genetic variation in the CCL18-CCL3-CCL4 chemokine gene cluster influences HIV Type 1 transmission and AIDS disease progression. *Am J Hum Genet* **79**: 120-128.

- MOLON, B., G. GRI, M. BETTELLA, C. GOMEZ-MOUTON, A. LANZAVECCHIA *et al.*, 2005 T cell costimulation by chemokine receptors. *Nat Immunol* **6**: 465-471.
- MOORE, J. P., A. TRKOLA and T. DRAGIC, 1997 Co-receptors for HIV-1 entry. *Curr Opin Immunol* **9**: 551-562.
- MORIUCHI, H., M. MORIUCHI and A. S. FAUCI, 1997 Cloning and analysis of the promoter region of CCR5, a coreceptor for HIV-1 entry. *J Immunol* **159**: 5441-5449.
- MORIUCHI, M., and H. MORIUCHI, 2001 Octamer transcription factors up-regulate the expression of CCR5, a coreceptor for HIV-1 entry. *J Biol Chem* **276**: 8639-8642.
- MORIUCHI, M., H. MORIUCHI and A. S. FAUCI, 1999 GATA-1 transcription factor transactivates the promoter for CCR5, a coreceptor for human immunodeficiency virus type 1 entry. *Blood* **93**: 1433-1435.
- MUMMIDI, S., L. M. ADAMS, S. E. VANCOMPENOLLE, M. KALKONDE, J. F. CAMARGO *et al.*, 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. *J Immunol* **178**: 5668-5681.
- MUMMIDI, S., S. S. AHUJA, E. GONZALEZ, S. A. ANDERSON, E. N. SANTIAGO *et al.*, 1998 Genealogy of the CCR5 locus and chemokine system gene variants associated with altered rates of HIV-1 disease progression. *Nat Med* **4**: 786-793.
- MUMMIDI, S., S. S. AHUJA, B. L. MCDANIEL and S. K. AHUJA, 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. *J Biol Chem* **272**: 30662-30671.
- MUMMIDI, S., M. BAMSHAD, S. S. AHUJA, E. GONZALEZ, P. M. FEUILLET *et al.*, 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. *J Biol Chem* **275**: 18946-18961.
- NAIF, H. M., S. LI, M. ALALI, A. SLOANE, L. WU *et al.*, 1998 CCR5 expression correlates with susceptibility of maturing monocytes to human immunodeficiency virus type 1 infection. *J Virol* **72**: 830-836.
- NAKAO, M., H. NOMIYAMA and K. SHIMADA, 1990 Structures of human genes coding for cytokine LD78 and their expression. *Mol Cell Biol* **10**: 3646-3658.
- NARANBHAI, V., P. BARTMAN, D. NDLOVU, P. RAMKALAWON, T. NDUNG'U *et al.*, 2011 Impact of blood processing variations on natural killer cell frequency, activation, chemokine receptor expression and function. *J Immunol Methods* **366**: 28-35.
- NGUYEN, L., M. LI, T. CHAOWANACHAN, D. J. HU, S. VANICHSENI *et al.*, 2004 CCR5 promoter human haplogroups associated with HIV-1 disease progression in Thai injection drug users. *Aids* **18**: 1327-1333.
- NIBBS, R. J., J. YANG, N. R. LANDAU, J. H. MAO and G. J. GRAHAM, 1999 LD78beta, a non-allelic variant of human MIP-1alpha (LD78alpha), has enhanced receptor interactions and potent HIV suppressive activity. *J Biol Chem* **274**: 17478-17483.

- NICHOLSON, J. K., S. W. BROWNING, R. L. HENGEL, E. LEW, L. E. GALLAGHER *et al.*, 2001 CCR5 and CXCR4 expression on memory and naive T cells in HIV-1 infection and response to highly active antiretroviral therapy. *J Acquir Immune Defic Syndr* **27**: 105-115.
- NIU, T., Z. S. QIN, X. XU and J. S. LIU, 2002 Bayesian haplotype inference for multiple linked single-nucleotide polymorphisms. *Am J Hum Genet* **70**: 157-169.
- NOLAN, D., S. GAUDIERI, M. JOHN and S. MALLAL, 2004 Impact of host genetics on HIV disease progression and treatment: new conflicts on an ancient battleground. *Aids* **18**: 1231-1240.
- NOVEMBRE, J., A. P. GALVANI and M. SLATKIN, 2005 The geographic spread of the CCR5 Delta32 HIV-resistance allele. *PLoS Biol* **3**: e339.
- O'BRIEN, S. J., and G. W. NELSON, 2004 Human genes that limit AIDS. *Nat Genet* **36**: 565-574.
- O'BRIEN, T. R., C. WINKLER, M. DEAN, J. A. NELSON, M. CARRINGTON *et al.*, 1997 HIV-1 infection in a man homozygous for CCR5 delta 32. *Lancet* **349**: 1219.
- OHTA, T. M., T. KASAMA, M. HANYUUDA, Y. HATANO, K. KOBAYASHI *et al.*, 1998 Interleukin-13 down-regulates the expression of neutrophil-derived macrophage inflammatory protein-1 alpha. *Inflamm Res* **47**: 361-368.
- OLSZEWSKI, M. A., G. B. HUFFNAGLE, T. R. TRAYNOR, R. A. McDONALD, D. N. COOK *et al.*, 2001 Regulatory effects of macrophage inflammatory protein 1alpha/CCL3 on the development of immunity to *Cryptococcus neoformans* depend on expression of early inflammatory cytokines. *Infect Immun* **69**: 6256-6263.
- OMETTO, L., R. BERTORELLE, M. MAINARDI, M. ZANCHETTA, S. TOGNAZZO *et al.*, 2001 Polymorphisms in the CCR5 promoter region influence disease progression in perinatally human immunodeficiency virus type 1-infected children. *J Infect Dis* **183**: 814-818.
- OPPERMANN, M., 2004 Chemokine receptor CCR5: insights into structure, function, and regulation. *Cell Signal* **16**: 1201-1210.
- OSTROWSKI, M. A., S. J. JUSTEMENT, A. CATANZARO, C. A. HALLAHAN, L. A. EHLER *et al.*, 1998 Expression of chemokine receptors CXCR4 and CCR5 in HIV-1-infected and uninfected individuals. *J Immunol* **161**: 3195-3201.
- PAIARDINI, M., I. PANDREA, C. APETREI and G. SILVESTRI, 2009 Lessons learned from the natural hosts of HIV-related viruses. *Annu Rev Med* **60**: 485-495.
- PALACIOS, E., L. DIGILIO, H. M. McCLURE, Z. CHEN, P. A. MARX *et al.*, 1998 Parallel evolution of CCR5-null phenotypes in humans and in a natural host of simian immunodeficiency virus. *Curr Biol* **8**: 943-946.
- PALCZEWSKI, K., T. KUMASAKA, T. HORI, C. A. BEHNKE, H. MOTOSHIMA *et al.*, 2000 Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**: 739-745.
- PALELLA, F. J., JR., K. M. DELANEY, A. C. MOORMAN, M. O. LOVELESS, J. FUHRER *et al.*, 1998 Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *N Engl J Med* **338**: 853-860.
- PASTORI, C., B. WEISER, C. BARASSI, C. UBERTI-FOPPA, S. GHEZZI *et al.*, 2006 Long-lasting CCR5 internalization by antibodies in a subset of long-term nonprogressors: a possible protective effect against disease progression. *Blood* **107**: 4825-4833.

- PAXIMADIS, M., N. MOHANLAL, G. E. GRAY, L. KUHN and C. T. TIEMESSEN, 2009 Identification of new variants within the two functional genes CCL3 and CCL3L encoding the CCL3 (MIP-1alpha) chemokine: implications for HIV-1 infection. *Int J Immunogenet* **36**: 21-32.
- PAXIMADIS, M., D. B. SCHRAMM, G. E. GRAY, G. SHERMAN, A. COOVADIA *et al.*, 2012 Influence of intragenic CCL3 haplotypes and CCL3L copy number in HIV-1 infection in a sub-Saharan African population. *Genes Immun*.
- PAXTON, W. A., R. LIU, S. KANG, L. WU, T. R. GINGERAS *et al.*, 1998 Reduced HIV-1 infectability of CD4+ lymphocytes from exposed-uninfected individuals: association with low expression of CCR5 and high production of beta-chemokines. *Virology* **244**: 66-73.
- PEREYRA, F., M. M. ADDO, D. E. KAUFMANN, Y. LIU, T. MIURA *et al.*, 2008 Genetic and immunologic heterogeneity among persons who control HIV infection in the absence of therapy. *J Infect Dis* **197**: 563-571.
- PETERSEN, D. C., M. J. KOTZE, M. D. ZEIER, A. GRIMWOOD, D. PRETORIUS *et al.*, 2001 Novel mutations identified using a comprehensive CCR5-denaturing gradient gel electrophoresis assay. *Aids* **15**: 171-177.
- PETREK, M., J. DRABEK, V. KOLEK, J. ZLAMAL, K. I. WELSH *et al.*, 2000 CC chemokine receptor gene polymorphisms in Czech patients with pulmonary sarcoidosis. *Am J Respir Crit Care Med* **162**: 1000-1003.
- PICTON, A. C., M. PAXIMADIS and C. T. TIEMESSEN, 2010 Genetic variation within the gene encoding the HIV-1 CCR5 coreceptor in two South African populations. *Infect Genet Evol* **10**: 487-494.
- PICTON, A. C., M. PAXIMADIS and C. T. TIEMESSEN, 2012a CCR5 promoter haplotypes differentially influence CCR5 expression on natural killer and T cell subsets in ethnically divergent HIV-1 uninfected South African populations. *Immunogenetics* **64**: 795-806.
- PICTON, A. C., M. PAXIMADIS and C. T. TIEMESSEN, 2013 Contribution of variable CCL3L copy number to CCL3 protein production in two ethnically divergent South African populations. *Infect Genet Evol* **14**: 347-356.
- PICTON, A. C., S. SHALEKOFF, M. PAXIMADIS and C. T. TIEMESSEN, 2012b Marked differences in CCR5 expression and activation levels in two South African populations. *Immunology* **136**: 397-407.
- PIDO-LOPEZ, J., T. WHITTALL, Y. WANG, L. A. BERGMEIER, K. BABAAHMADY *et al.*, 2007 Stimulation of cell surface CCR5 and CD40 molecules by their ligands or by HSP70 up-regulates APOBEC3G expression in CD4(+) T cells and dendritic cells. *J Immunol* **178**: 1671-1679.
- PILOTTI, E., L. ELVIRI, E. VICENZI, U. BERTAZZONI, M. C. RE *et al.*, 2007 Postgenomic up-regulation of CCL3L1 expression in HTLV-2-infected persons curtails HIV-1 replication. *Blood* **109**: 1850-1856.
- PINTO, L. A., J. SULLIVAN, J. A. BERZOFSKY, M. CLERICI, H. A. KESSLER *et al.*, 1995 ENV-specific cytotoxic T lymphocyte responses in HIV seronegative health care workers occupationally exposed to HIV-contaminated body fluids. *J Clin Invest* **96**: 867-876.
- PLATT, E. J., K. WEHRLY, S. E. KUHMANN, B. CHESEBRO and D. KABAT, 1998 Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1. *J Virol* **72**: 2855-2864.

- POLLAKIS, G., and W. A. PAXTON, 2012 Use of (alternative) coreceptors for HIV entry. *Curr Opin HIV AIDS* **7**: 440-449.
- POLO, S., F. VEGLIA, M. S. MALNATI, C. GOBBI, P. FARCI *et al.*, 1999 Longitudinal analysis of serum chemokine levels in the course of HIV-1 infection. *Aids* **13**: 447-454.
- POPOVIC, M., M. G. SARNGADHARAN, E. READ and R. C. GALLO, 1984 Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science* **224**: 497-500.
- POROPATICH, K., and D. J. SULLIVAN, JR., 2011 Human immunodeficiency virus type 1 long-term non-progressors: the viral, genetic and immunological basis for disease non-progression. *J Gen Virol* **92**: 247-268.
- PRAHALAD, S., 2006 Negative association between the chemokine receptor CCR5-Delta32 polymorphism and rheumatoid arthritis: a meta-analysis. *Genes Immun* **7**: 264-268.
- QUILLET, C., E. OBERLIN, J. BRAUN, D. ROUSSET, G. GONZALEZ-CANALI *et al.*, 1998 HIV-1-resistance phenotype conferred by combination of two separate inherited mutations of CCR5 gene. *Lancet* **351**: 14-18.
- REYNES, J., V. BAILLAT, P. PORTALES, J. CLOT and P. CORBEAU, 2003 Low CD4+ T-cell surface CCR5 density as a cause of resistance to in vivo HIV-1 infection. *J Acquir Immune Defic Syndr* **34**: 114-116.
- REYNES, J., P. PORTALES, M. SEGONDY, V. BAILLAT, P. ANDRE *et al.*, 2001 CD4 T cell surface CCR5 density as a host factor in HIV-1 disease progression. *Aids* **15**: 1627-1634.
- REYNES, J., P. PORTALES, M. SEGONDY, V. BAILLAT, P. ANDRE *et al.*, 2000 CD4+ T cell surface CCR5 density as a determining factor of virus load in persons infected with human immunodeficiency virus type 1. *J Infect Dis* **181**: 927-932.
- RICHARDSON, M. W., J. JADLOWSKY, C. A. DIDIGU, R. W. DOMS and J. L. RILEY, 2012 Kruppel-like Factor 2 Modulates CCR5 Expression and Susceptibility to HIV-1 Infection. *J Immunol* **189**: 3815-3821.
- RILEY, J. L., R. G. CARROLL, B. L. LEVINE, W. BERNSTEIN, D. C. ST LOUIS *et al.*, 1997 Intrinsic resistance to T cell infection with HIV type 1 induced by CD28 costimulation. *J Immunol* **158**: 5545-5553.
- RIZZARDI, G. P., R. A. MORAWETZ, E. VICENZI, S. GHEZZI, G. POLI *et al.*, 1998 CCR2 polymorphism and HIV disease. *Swiss HIV Cohort. Nat Med* **4**: 252-253.
- ROLLINS, B. J., 2006 Inflammatory chemokines in cancer growth and progression. *Eur J Cancer* **42**: 760-767.
- ROTTMAN, J. B., K. P. GANLEY, K. WILLIAMS, L. WU, C. R. MACKAY *et al.*, 1997 Cellular localization of the chemokine receptor CCR5. Correlation to cellular targets of HIV-1 infection. *Am J Pathol* **151**: 1341-1351.
- SALKOWITZ, J. R., S. E. BRUSE, H. MEYERSON, H. VALDEZ, D. E. MOSIER *et al.*, 2003 CCR5 promoter polymorphism determines macrophage CCR5 density and magnitude of HIV-1 propagation in vitro. *Clin Immunol* **108**: 234-240.

- SAMSON, M., O. LABBE, C. MOLLEREAU, G. VASSART and M. PARMENTIER, 1996a Molecular cloning and functional expression of a new human CC-chemokine receptor gene. *Biochemistry* **35**: 3362-3367.
- SAMSON, M., F. LIBERT, B. J. DORANZ, J. RUCKER, C. LIESNARD *et al.*, 1996b Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* **382**: 722-725.
- SAMSON, M., P. SOULARUE, G. VASSART and M. PARMENTIER, 1996c The genes encoding the human CC-chemokine receptors CC-CKR1 to CC-CKR5 (CMKBR1-CMKBR5) are clustered in the p21.3-p24 region of chromosome 3. *Genomics* **36**: 522-526.
- SCHALL, T. J., K. BACON, R. D. CAMP, J. W. KASPARI and D. V. GOEDEL, 1993 Human macrophage inflammatory protein alpha (MIP-1 alpha) and MIP-1 beta chemokines attract distinct populations of lymphocytes. *J Exp Med* **177**: 1821-1826.
- SCHIFFER, M., C. H. CHANG and F. J. STEVENS, 1992 The functions of tryptophan residues in membrane proteins. *Protein Eng* **5**: 213-214.
- 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 64l/+ mutations on human immunodeficiency virus (HIV)-1 disease progression among HIV-1-infected injecting drug users. *J Infect Dis* **179**: 825-831.
- SCOTT-ALGARA, D., L. X. TRUONG, P. VERSMISSE, A. DAVID, T. T. LUONG *et al.*, 2003 Cutting edge: increased NK cell activity in HIV-1-exposed but uninfected Vietnamese intravascular drug users. *J Immunol* **171**: 5663-5667.
- SELLEBJERG, F., H. O. MADSEN, C. V. JENSEN, J. JENSEN and P. GARRED, 2000 CCR5 delta32, matrix metalloproteinase-9 and disease activity in multiple sclerosis. *J Neuroimmunol* **102**: 98-106.
- SERODY, J. S., S. E. BURKETT, A. PANOSKALTSIS-MORTARI, J. NG-CASHIN, E. MCMAHON *et al.*, 2000 T-lymphocyte production of macrophage inflammatory protein-1alpha is critical to the recruitment of CD8(+) T cells to the liver, lung, and spleen during graft-versus-host disease. *Blood* **96**: 2973-2980.
- SHALEKOFF, S., S. MEDDOWS-TAYLOR, D. B. SCHRAMM, S. L. DONNINGER, G. E. GRAY *et al.*, 2008 Host CCL3L1 gene copy number in relation to HIV-1-specific CD4+ and CD8+ T-cell responses and viral load in South African women. *J Acquir Immune Defic Syndr* **48**: 245-254.
- SHALEKOFF, S., S. PENDLE, D. JOHNSON, D. J. MARTIN and C. T. TIEMESSEN, 2001 Distribution of the human immunodeficiency virus coreceptors CXCR4 and CCR5 on leukocytes of persons with human immunodeficiency virus type 1 infection and pulmonary tuberculosis: implications for pathogenesis. *J Clin Immunol* **21**: 390-401.
- SHALEKOFF, S., and C. T. TIEMESSEN, 2001 Duration of sample storage dramatically alters expression of the human immunodeficiency virus coreceptors CXCR4 and CCR5. *Clin Diagn Lab Immunol* **8**: 432-436.
- SHALEKOFF, S., and C. T. TIEMESSEN, 2003 CCR5 delta32 heterozygosity is associated with an increase in CXCR4 cell surface expression. *AIDS Res Hum Retroviruses* **19**: 531-533.

- SHAO, W., J. TANG, W. SONG, C. WANG, Y. LI *et al.*, 2007 CCL3L1 and CCL4L1: variable gene copy number in adolescents with and without human immunodeficiency virus type 1 (HIV-1) infection. *Genes Immun* **8**: 224-231.
- SHEA, A., D. A. SARR, N. JONES, L. PENNING, G. EISEN *et al.*, 2004 CCR5 receptor expression is down-regulated in HIV type 2 infection: implication for viral control and protection. *AIDS Res Hum Retroviruses* **20**: 630-635.
- SHIODA, T., E. E. NAKAYAMA, Y. TANAKA, X. XIN, H. LIU *et al.*, 2001 Naturally occurring deletional mutation in the C-terminal cytoplasmic tail of CCR5 affects surface trafficking of CCR5. *J Virol* **75**: 3462-3468.
- SHOSTAKOVICH-KORETSKAYA, L., G. CATANO, Z. A. CHYKARENKO, W. HE, G. GORNALUSSE *et al.*, 2009 Combinatorial content of CCL3L and CCL4L gene copy numbers influence HIV-AIDS susceptibility in Ukrainian children. *Aids* **23**: 679-688.
- SHRESTHA, S., M. NYAKU and J. C. EDBERG, 2010 Variations in CCL3L gene cluster sequence and non-specific gene copy numbers. *BMC Res Notes* **3**: 74.
- SHRESTHA, S., S. A. STRATHDEE, N. GALAI, T. OLEKSYK, M. D. FALLIN *et al.*, 2006 Behavioral risk exposure and host genetics of susceptibility to HIV-1 infection. *J Infect Dis* **193**: 16-26.
- SHRESTHA, S., J. TANG and R. A. KASLOW, 2009 Gene copy number: learning to count past two. *Nat Med* **15**: 1127-1129.
- SIEGAL, F. P., C. LOPEZ, G. S. HAMMER, A. E. BROWN, S. J. KORNFELD *et al.*, 1981 Severe acquired immunodeficiency in male homosexuals, manifested by chronic perianal ulcerative herpes simplex lesions. *N Engl J Med* **305**: 1439-1444.
- SIGNORET, N., A. PELCHEN-MATTHEWS, M. MACK, A. E. PROUDFOOT and M. MARSH, 2000 Endocytosis and recycling of the HIV coreceptor CCR5. *J Cell Biol* **151**: 1281-1294.
- SINGH, K. K., M. D. HUGHES, J. CHEN, K. PHIRI, C. ROUSSEAU *et al.*, 2008 Associations of chemokine receptor polymorphisms With HIV-1 mother-to-child transmission in sub-Saharan Africa: possible modulation of genetic effects by antiretrovirals. *J Acquir Immune Defic Syndr* **49**: 259-265.
- SMITH, M. W., M. DEAN, M. CARRINGTON, C. WINKLER, G. A. HUTTLEY *et al.*, 1997 Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC), ALIVE Study. *Science* **277**: 959-965.
- SMYTH, D. J., V. PLAGNOL, N. M. WALKER, J. D. COOPER, K. DOWNES *et al.*, 2008 Shared and distinct genetic variants in type 1 diabetes and celiac disease. *N Engl J Med* **359**: 2767-2777.
- SORCE, S., J. BONNEFONT, S. JULIEN, N. MARQ-LIN, I. RODRIGUEZ *et al.*, 2010 Increased brain damage after ischaemic stroke in mice lacking the chemokine receptor CCR5. *Br J Pharmacol* **160**: 311-321.
- SOZZANI, S., P. ALLAVENA, G. D'AMICO, W. LUINI, G. BIANCHI *et al.*, 1998 Differential regulation of chemokine receptors during dendritic cell maturation: a model for their trafficking properties. *J Immunol* **161**: 1083-1086.

- STANDIFORD, T. J., S. L. KUNKEL, J. M. LIEBLER, M. D. BURDICK, A. R. GILBERT *et al.*, 1993 Gene expression of macrophage inflammatory protein-1 alpha from human blood monocytes and alveolar macrophages is inhibited by interleukin-4. *Am J Respir Cell Mol Biol* **9**: 192-198.
- STEPHENS, J. C., D. E. REICH, D. B. GOLDSTEIN, H. D. SHIN, M. W. SMITH *et al.*, 1998 Dating the origin of the CCR5-Delta32 AIDS-resistance allele by the coalescence of haplotypes. *Am J Hum Genet* **62**: 1507-1515.
- STEWART, G. J., L. J. ASHTON, R. A. BITI, R. A. FFRENCH, B. H. BENNETTS *et al.*, 1997 Increased frequency of CCR-5 delta 32 heterozygotes among long-term non-progressors with HIV-1 infection. The Australian Long-Term Non-Progressor Study Group. *Aids* **11**: 1833-1838.
- STRUYF, S., P. MENTEN, J. P. LENAERTS, W. PUT, A. D'HAESE *et al.*, 2001 Diverging binding capacities of natural LD78beta isoforms of macrophage inflammatory protein-1alpha to the CC chemokine receptors 1, 3 and 5 affect their anti-HIV-1 activity and chemotactic potencies for neutrophils and eosinophils. *Eur J Immunol* **31**: 2170-2178.
- SUNDRUD, M. S., S. E. VANCOMPERNOLLE, K. A. EGER, T. C. BRUNO, A. SUBRAMANIAM *et al.*, 2005 Transcription factor GATA-1 potently represses the expression of the HIV-1 coreceptor CCR5 in human T cells and dendritic cells. *Blood* **106**: 3440-3448.
- SUZUKI, S., T. MIYAGI, L. F. CHUANG, P. M. YAU, R. H. DOI *et al.*, 2002 Chemokine receptor CCR5: polymorphism at protein level. *Biochem Biophys Res Commun* **296**: 477-483.
- SYMONDS, G. P., H. A. JOHNSTONE, M. L. MILLINGTON, M. P. BOYD, B. P. BURKE *et al.*, 2010 The use of cell-delivered gene therapy for the treatment of HIV/AIDS. *Immunol Res* **48**: 84-98.
- TANG, J., B. SHELTON, N. J. MAKHATADZE, Y. ZHANG, M. SCHAEN *et al.*, 2002 Distribution of chemokine receptor CCR2 and CCR5 genotypes and their relative contribution to human immunodeficiency virus type 1 (HIV-1) seroconversion, early HIV-1 RNA concentration in plasma, and later disease progression. *J Virol* **76**: 662-672.
- THAPA, M., W. A. KUZIEL and D. J. CARR, 2007 Susceptibility of CCR5-deficient mice to genital herpes simplex virus type 2 is linked to NK cell mobilization. *J Virol* **81**: 3704-3713.
- THEODOROU, I., L. MEYER, M. MAGIEROWSKA, C. KATLAMA and C. ROUZIOUX, 1997 HIV-1 infection in an individual homozygous for CCR5 delta 32. Seroco Study Group. *Lancet* **349**: 1219-1220.
- THOMAS, S. M., D. B. TSE, D. S. KETNER, G. ROCHFORD, D. A. MEYER *et al.*, 2006 CCR5 expression and duration of high risk sexual activity among HIV-seronegative men who have sex with men. *Aids* **20**: 1879-1883.
- TIEMESSEN, C. T., and L. KUHN, 2007 CC chemokines and protective immunity: insights gained from mother-to-child transmission of HIV. *Nat Immunol* **8**: 219-222.
- TIEMESSEN, C. T., S. SHALEKOFF, S. MEDDOWS-TAYLOR, D. B. SCHRAMM, M. A. PAPATHANASOPOULOS *et al.*, 2009 Cutting Edge: Unusual NK cell responses to HIV-1 peptides are associated with protection against maternal-infant transmission of HIV-1. *J Immunol* **182**: 5914-5918.
- TIEMESSEN, C. T., S. SHALEKOFF, S. MEDDOWS-TAYLOR, D. B. SCHRAMM, M. A. PAPATHANASOPOULOS *et al.*, 2010 Natural killer cells that respond to human immunodeficiency virus type 1 (HIV-1) peptides are associated with control of HIV-1 infection. *J Infect Dis* **202**: 1444-1453.

- TISHKOFF, S. A., F. A. REED, F. R. FRIEDLAENDER, C. EHRET, A. RANCIARO *et al.*, 2009 The genetic structure and history of Africans and African Americans. *Science* **324**: 1035-1044.
- TISHKOFF, S. A., and B. C. VERRELLI, 2003 Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annu Rev Genomics Hum Genet* **4**: 293-340.
- TOWNSON, J. R., L. F. BARCELLOS and R. J. NIBBS, 2002 Gene copy number regulates the production of the human chemokine CCL3-L1. *Eur J Immunol* **32**: 3016-3026.
- TRKOLA, A., W. A. PAXTON, S. P. MONARD, J. A. HOXIE, M. A. SIANI *et al.*, 1998 Genetic subtype-independent inhibition of human immunodeficiency virus type 1 replication by CC and CXC chemokines. *J Virol* **72**: 396-404.
- TSUCHIYA, S., M. YAMABE, Y. YAMAGUCHI, Y. KOBAYASHI, T. KONNO *et al.*, 1980 Establishment and characterization of a human acute monocytic leukemia cell line (THP-1). *Int J Cancer* **26**: 171-176.
- TUTTLE, D. L., J. K. HARRISON, C. ANDERS, J. W. SLEASMAN and M. M. GOODENOW, 1998 Expression of CCR5 increases during monocyte differentiation and directly mediates macrophage susceptibility to infection by human immunodeficiency virus type 1. *J Virol* **72**: 4962-4969.
- UITENBROOK, D. G., 1997 SISA-Binomial. <http://quantitativeskills.com/sisa/distributions/binomial>.
- ULLUM, H., A. COZZI LEPRI, J. VICTOR, H. ALADDIN, A. N. PHILLIPS *et al.*, 1998 Production of beta-chemokines in human immunodeficiency virus (HIV) infection: evidence that high levels of macrophage inflammatory protein-1beta are associated with a decreased risk of HIV disease progression. *J Infect Dis* **177**: 331-336.
- UNAIDS, 2012a 2012 UNAIDS Report on the Global AIDS Epidemic. Joint United Nations Programme on HIV/AIDS (UNAIDS).
- UNAIDS, 2012b Regional fact sheet: sub-Saharan Africa. Geneva: Joint United Nations Program on HIV/AIDS.
- URBAN, T. J., A. C. WEINTROB, J. FELLAY, S. COLOMBO, K. V. SHIANNA *et al.*, 2009 CCL3L1 and HIV/AIDS susceptibility. *Nat Med* **15**: 1110-1112.
- VALENTIN, A., M. ROSATI, D. J. PATENAUE, A. HATZAKIS, L. G. KOSTRIKIS *et al.*, 2002 Persistent HIV-1 infection of natural killer cells in patients receiving highly active antiretroviral therapy. *Proc Natl Acad Sci U S A* **99**: 7015-7020.
- VAN BREEMEN, M. J., M. DE FOST, J. S. VOERMAN, J. D. LAMAN, R. G. BOOT *et al.*, 2007 Increased plasma macrophage inflammatory protein (MIP)-1alpha and MIP-1beta levels in type 1 Gaucher disease. *Biochim Biophys Acta* **1772**: 788-796.
- VINCENT, T., P. PORTALES, V. BAILLAT, A. EDEN, J. CLOT *et al.*, 2006 The immunological response to highly active antiretroviral therapy is linked to CD4+ T-cell surface CCR5 density. *J Acquir Immune Defic Syndr* **43**: 377-378.
- WAIN, L. V., J. A. ARMOUR and M. D. TOBIN, 2009 Genomic copy number variation, human health, and disease. *Lancet* **374**: 340-350.
- WEISS, I. D., H. SHOHAM, O. WALD, H. WALD, K. BEIDER *et al.*, 2011 Ccr5 deficiency regulates the proliferation and trafficking of natural killer cells under physiological conditions. *Cytokine* **54**: 249-257.

- WEISSMAN, D., M. DYBUL, M. B. DAUCHER, R. T. DAVEY, JR., R. E. WALKER *et al.*, 2000 Interleukin-2 up-regulates expression of the human immunodeficiency virus fusion coreceptor CCR5 by CD4+ lymphocytes in vivo. *J Infect Dis* **181**: 933-938.
- WHITTALL, T., Y. WANG, J. YOUNSON, C. KELLY, L. BERGMEIER *et al.*, 2006 Interaction between the CCR5 chemokine receptors and microbial HSP70. *Eur J Immunol* **36**: 2304-2314.
- WIERDA, R. J., H. F. KUIPERS, M. C. VAN EGGERMOND, A. BENARD, J. C. VAN LEEUWEN *et al.*, 2012 Epigenetic control of CCR5 transcript levels in immune cells and modulation by small molecules inhibitors. *J Cell Mol Med* **16**: 1866-1877.
- WILEN, C. B., J. C. TILTON and R. W. DOMS, 2012 HIV: cell binding and entry. *Cold Spring Harb Perspect Med* **2**.
- WILKINSON, D. A., E. A. OPERSKALSKI, M. P. BUSCH, J. W. MOSLEY and R. A. KOUP, 1998 A 32-bp deletion within the CCR5 locus protects against transmission of parenterally acquired human immunodeficiency virus but does not affect progression to AIDS-defining illness. *J Infect Dis* **178**: 1163-1166.
- WILLIAMSON, C., S. A. LOUBSER, B. BRICE, G. JOUBERT, T. SMIT *et al.*, 2000 Allelic frequencies of host genetic variants influencing susceptibility to HIV-1 infection and disease in South African populations. *Aids* **14**: 449-451.
- WU, L., W. A. PAXTON, N. KASSAM, N. RUFFING, J. B. ROTTMAN *et al.*, 1997 CCR5 levels and expression pattern correlate with infectability by macrophage-tropic HIV-1, in vitro. *J Exp Med* **185**: 1681-1691.
- XIANG, J., S. L. GEORGE, S. WUNSCHMANN, Q. CHANG, D. KLINZMAN *et al.*, 2004 Inhibition of HIV-1 replication by GB virus C infection through increases in RANTES, MIP-1alpha, MIP-1beta, and SDF-1. *Lancet* **363**: 2040-2046.
- YANG, Y. F., M. TOMURA, M. IWASAKI, T. MUKAI, P. GAO *et al.*, 2001a IL-12 as well as IL-2 upregulates CCR5 expression on T cell receptor-triggered human CD4+ and CD8+ T cells. *J Clin Immunol* **21**: 116-125.
- YANG, Y. F., M. TOMURA, M. IWASAKI, S. ONO, J. P. ZOU *et al.*, 2001b IFN-alpha acts on T-cell receptor-triggered human peripheral leukocytes to up-regulate CCR5 expression on CD4+ and CD8+ T cells. *J Clin Immunol* **21**: 402-409.
- YE, P., P. KAZANJIAN, S. L. KUNKEL and D. E. KIRSCHNER, 2004 Lack of good correlation of serum CC-chemokine levels with human immunodeficiency virus-1 disease stage and response to treatment. *J Lab Clin Med* **143**: 310-319.
- YUNG, R. L., and R. MO, 2003 Aging is associated with increased human T cell CC chemokine receptor gene expression. *J Interferon Cytokine Res* **23**: 575-582.
- ZAPICO, I., E. COTO, A. RODRIGUEZ, C. ALVAREZ, J. C. TORRE *et al.*, 2000 CCR5 (chemokine receptor-5) DNA-polymorphism influences the severity of rheumatoid arthritis. *Genes Immun* **1**: 288-289.
- ZERNECKE, A., E. SHAGDARSUREN and C. WEBER, 2008 Chemokines in atherosclerosis: an update. *Arterioscler Thromb Vasc Biol* **28**: 1897-1908.
- ZHAO, X. Y., S. S. LEE, K. H. WONG, K. C. CHAN, F. NG *et al.*, 2005 Functional analysis of naturally occurring mutations in the open reading frame of CCR5 in HIV-infected Chinese patients and healthy controls. *J Acquir Immune Defic Syndr* **38**: 509-517.

- ZLOTNIK, A., and O. YOSHIE, 2000 Chemokines: a new classification system and their role in immunity. *Immunity* **12**: 121-127.
- ZOU, W., A. FOUSSAT, S. HOUBOU, I. DURAND-GASSELIN, A. DULIOUST *et al.*, 1999 Acute upregulation of CCR-5 expression by CD4+ T lymphocytes in HIV-infected patients treated with interleukin-2. ANRS 048 IL-2 Study Group. *Aids* **13**: 455-463.

## APPENDIX A

### Supplementary Data for Chapter 2

**Table A.1.** *CCR5* primer sets and experimental conditions for polymerase chain reaction amplification of overlapping regions of gene

Fragment	Primer sequence	Binding position	Amplicon size (bp)	Annealing temperature
Promoter 2	f: 5' AAG AAG CCG CCT ACA GAA TGG 3'	-5523 to -5503	2412	65°C (first 10 cycles) 60°C (last 20 cycles)
	r: 5' CCC TCA GCT TTC TCG TCT GG 3'	-3131 to -3111		
Promoter 1	f: 5' CCA AGC ACC AGC AAT TAG C 3'	-3384 to -3366	2189	60°C
	r: 5' TGC CAC CAC AGA TGA ATG TC 3'	-1215 to -1196		
Intron 2	f: 5' GTA ACC TCT CAG CTG CTT G 3'	-1910 to -1892	2036	55°C
	r: 5' GAT GAA CAC CAG TGA GTA GAG C 3'	+105 to +126		
ORF	f: 5' GCA CCA TGC TTG ACC CAG TT 3'	-340 to -321	2381	60°C
	r: 5' CAT ATG CTG CAC GAA TAC CTC 3'	+2020 to +2041		
3' UTR	f: 5' GAG ATC CTG GTT GGT GTT GC 3'	+1844 to +1863	1866	60°C
	r: 5' GTA AGT GAC CAG GCC ATG AC 3'	+3689 to +3709		

<sup>a</sup>bp, base pair

**Table A.2.** Flanking sequences of all SNPs detected in this study

	SNP Position	Base change (wt/mut)	5' flanking sequence	3' flanking sequence
5' Flanking region (2762 bp)	-5268	G/A	GTC ACC CAG GTT GGA	TGC AAT GGC CCA ATC
	-5266	G/A	CAC CCA GGT TGG ART	CAA TGG CCC AAT CTT
	-5214	T/C	CCT GGG TTC AAG AGA	TCT CCC ACT TCA GCC
	-5080	T/A	CTC CTG ACC TCA AGT	ATC CAC TCG CCT TGG
	-5072	C/T	CCT CAA GTA TCC ACT	GCC TTG GCT TCC CAA
	-4897	G/A	AAC AGC TGC TGT ATG	CAG GGT TTC TGC TCA
	-4808	G/A	TGT TGA CGA TGC TCT	AAA ATA TGG TCC AGA
	-4745	C/T	CTT GGA ACA TAG GTG	AGT GAC TAG ACA TGG
	-4630	T/C	TGA TCA AAA GTT CAT	TCC TAT GGG GTG TCC
	-4358	A/G	CAA AAG TGG AGT AAC	CAC ACT GCA AAG CTG
	-4257	A/C	GCC ATA GAA TCA TGT	GTA TTT AGG GTG GAA
	-4223	C/T	GCC CCA GGT CTA GCA	GTC ATT TAA CAG ATG
	-4088	T/C	CTC TGG GGG TGA GTA	GTC TTC ACA TCC TAA
	-3949	A/G	TCCGGGGTCTGC	CAAGTGGATTACCAG
	-3899	A/C	CTT TCG AAA AAC CAA	GTT GYA TTT ATG CTA
	-3894	T/C	CGA AAA ACC AAM TTG	ATT TAT GCT ATC TAT
	-3886	C/T	AAM GTT GYA TTT ATG	TAT (CTAT) TTT CTA TA
	-3868	CTAT/-	TTG YAT TTA TGC TAT	TTT CTA TAA AAT TTT
	-3833	C/T	ACC TAT TTT TGA ACT	TTT CAA AAG CAC ACT
	-3458	G/T	CTA TAT GGG GCG GGG	TGG GGG TGT CTT GAT
-3432	T/C	TTG ATC GCT GGG CTA	TTC TAT ACT GTT CTG	
-3261	G/A	ACT CTC TCT GAC AAA	GAC TGC TCA AAG AGT	
-2852	A/G	CAG ATG TCA CCA ACC	CCA AGA GAG CTT GAT	
-2823	T/A	TAT GAC TGT ATA TAG	ATA GTC ATA AAG AAC	
Exon 1 (57 bp)	-2733	A/G	ATA TCT GGA GTG AAG	ATC CTG CCA CCT ATG
Intron 1 (501bp)	-2577	T/G	CCG TGA GCC CAT AGT	AAA ACT CTT TAG ACA
	-2554	G/T	TTT AGA CAA CAG GTT	TTT CCG TTT ACA GAG
	-2459	G/A	GTG GAG AAA AAG GGG	CAC ARG GTT AAT GTG
	-2454	G/A	GAA AAA GGG GRC ACA	GGT TAA TGT GAA GTC
Exon 2A (235bp)	-2150	A/G	TAA TCC AGT GAG AAA	GCC CGT AAA TAA ACT
	-2135	T/C	AGC CCG TAA ATA AAC	TTY AGA CCA GAG ATC
	-2132	C/T	CCG TAA ATA AAC YTT	AGA CCA GAG ATC TAT
	-2086	A/G	AAG CTC AAC TTA AAA	GAA GAA CTG TTC TCT
	-2048	C/G	TTC GCC TTC AAT ACA	TTA ATG ATT TAA CTC

Intron 2 (1903bp)	-1835	C/T	GAA ACC CAT AGA AGA	ATT TGG CAA ACA CCA
	-1686	A/C	TTT AAA GGG AGC AAT	GTA TTT TAA TAA CTA
	-1464	A/G	ATG GTC AAA ATT AAT	TTA AAT TAC AAA CGC
	-1193	C/T	ATC TGT GGT GGC AGA	GAA ACA TTT TTT ATT
	-1130	AG/-	TAA TTG TGG CAA CTC	AAA CTA CAA ACA CAA
	-1060	C/T	TCA TCT ATG ACC TTC	CTG GGA CTT GGG CAC
	-976	C/T	TTT TAA TTC TCT TTT	RAG GAC TGA GAG GGA
	-975	G/A	TTT AAT TCT CTT TTY	AGG ACT GAG AGG GAG
	-730	A/T	CCC TCG AGG CCT CTT	ATT ATT ACT GGC TTG
	-651	C/T	AGA CTG AGT TGC AGC	GGG CAT GGT GGC TCA
	-451	C/T	AGG ATT GCT TGA GCC	GGG ATG RTC CAG GCT
	-444	G/A	CTT GAG CCY GGG ATG	TCC AGG CTG CAG TGA
	-362	ACAA/G	CTC ACA ACA ACA ACA	CAA CAA AAA GGC TGA
	-113	G/T	GTA GAC ATC TAT GTA	RCA ATT AAA AAC CTA
	-112	G/A	TAG ACA TCT ATG TAK	CAA TTA AAA ACC TAT
	Exon 3/ORF (1059bp)	+225	T/C	CAA CCT GGC CAT CTC
+258		G/C	TAC TGT CCC CTT CTG	GCT CAC TAT GCT GCC
+319		C/T	CAA CTC TTG ACA GGG	TCT ATT TTA TAG GCT
+554		Δ32	CTC ATT TTC CAT ACA	TTA AAG ATA GTC ATC
+673		C/T	ACT CTG CTT CGG TGT	GAA ATG AGA AGA AGA
+1004		C/T	AGG CTC CCG AGC GAG	AAG CTC AGT TTA CAC
3' UTR (2651bp)	+1253	A/G	CC CAT CAA TTA TAG	AAG CCA AAT CAA AAT
	+1752	G/A	CTG TAG AAG GAG ACA	AGC TGG TTG GGA AGA
	+1810	G/A	ATG AAG AAC CTT GAC	GCA TTG CTC CGT CTA
	+1823	C/T	ACR GCA TTG CTC CGT	TAA GTC ATG AGC TGA
	+1843	G/A	TCA TGA GCT GAG CAG	GAR ATC CTG GTT GGT
	+1846	G/A	TGA GCT GAG CAG RGA	ATC CTG GTT GGT GTT
	+2066	G/A	AGT CAG CAG AAC TGG	GTG GAT TTG GKT TGG
	+2077	G/T	CTG GRG TGG ATT TGG	TTG GAA GTG AGG GTC
	+2225	T/C	AAA GGA TGG GTC TGG	TTG CAG AGC TTG AAC
	+2293	A/G	TGC TTC TGA CTT CAT	GAT TTC CTT CCC ATC
	+2381	A/G	ATG AGG TCT AGG AAC	TAC TTC AGC TCA CAC
	+2435	T/A	TAC CTA GTA GTC ATT	CAT GGG TTG TTG GGA
	+2458	A/C	TGT TGG GAG GAT TCT	TGA GGC AAC CAC AGG
	+2676	C/A	AAA GGG GGG AAG GGA	ATA TTC ATT TGG AAA
	+2772	G insertion	TGG GGG TGG GGG GGG	CGC CTT AGG TAC TTA
	+2838	C/G	GAA AAA ATC GTC TCT	CCT CCC TTT GAA ATG
	+2919	T/G	GTT TTT TTC TGT TCT	TCT CAT ATG ATT GTG
+3132	T/G	CAA CGA AGG GAA ATG	CTT TCC TTT TGC TCT	

## APPENDIX B

### Supplementary Data for Chapter 3

**Table B.1.** Primer sets and experimental conditions for polymerase chain reaction (PCR) amplification of *CCR5* promoter construct regions used to assess promoter efficiency

Fragment	Primer sequence <sup>†</sup>	Binding position	Amplicon size (bp)	PCR conditions
P1A	f: 5' ATT CTA GAG CCA AGG TCA CG 3'	-2966 to -2947	1770	Expand High Fidelity PCR System (Roche): 60°C annealing temperature.
	r: 5' TGC CAC CAC AGA TGA ATG TC 3'	-1215 to -1196		
P2	f: 5' GAG CCA TGC AAA CTG AGA G 3'	-4595 to -4577	1873	
	r: 5' GGT GGC AGG ATT CTT CAC T 3'	-2740 to -2722		
P1B	f: 5' AAG <b>GTA CCA</b> TTC TAG AGC CAA GGT CAC G 3'	-2966 to -2947	3194	PrimeSTAR™ HS DNA polymerase (Takara). Cycling conditions: 94°C: 1 min Followed by 30 cycles: 98°C: 10 s, 68°C: 5 min
	r: 5' TAA <b>GTC GAC</b> GTC AGA GAT GGC CAG GTT G 3'	+210 to +228		
P1+P2	f: 5' AAG <b>GTA CCG</b> AGC CAT GCA AAC TGA GAG 3'	-4595 to -4577	4823	
	r: 5' TAA <b>GTC GAC</b> GTC AGA GAT GGC CAG GTT G 3'	+210 to +228		

bp, base pair

<sup>†</sup>Built-in restriction sites are indicated in coloured font, red for *Sall* and blue for *KpnI*.

**Table B.2.** Primer pairs used to screen clones for directional cloning of promoter fragments

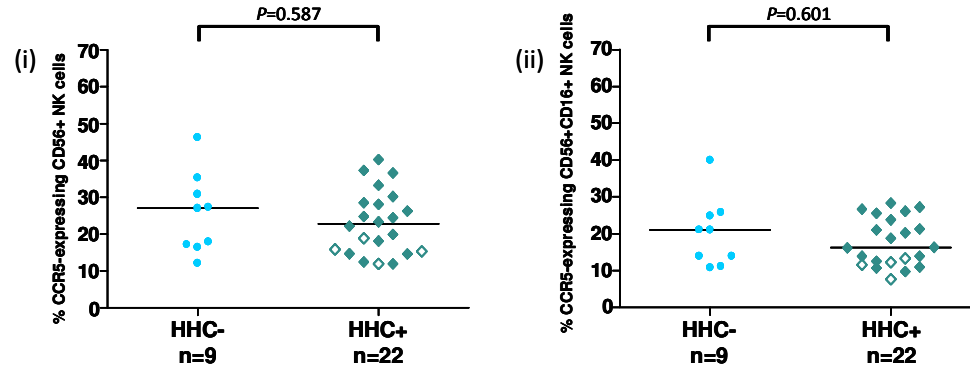
Fragment	Primer sequence	Screening method
P1A	f: 5' CTA ACA GAT TCT GTG TAG TGG 3'	Directional
	r: 5' TAG AAG GCA CAG TCG AGG 3' (BGH reverse)	
P2	f: 5' CCA AGC ACC AGC AAT TAG C 3'	Directional
	r: 5' TAG AAG GCA CAG TCG AGG 3' (BGH reverse)	
P1B*	f: 5' GGA TGA TCC AGG CTG CAG TG 3'	Non-directional
	r: 5' GTC AGA GAT GGC CAG GTT GAG 3'	
P1+P2*	f: 5' GGA TGA TCC AGG CTG CAG TG 3'	Non-directional
	r: 5' GTC AGA GAT GGC CAG GTT GAG 3'	

\*P1B and P1+P2 constructs were prepared with sticky ends prior to cloning, therefore directional insert screening was not necessary

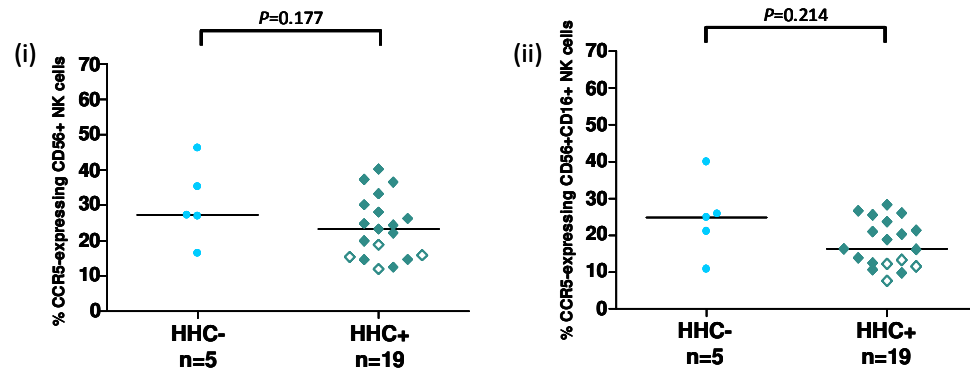
# APPENDIX C

## Supplementary Data for Chapter 5

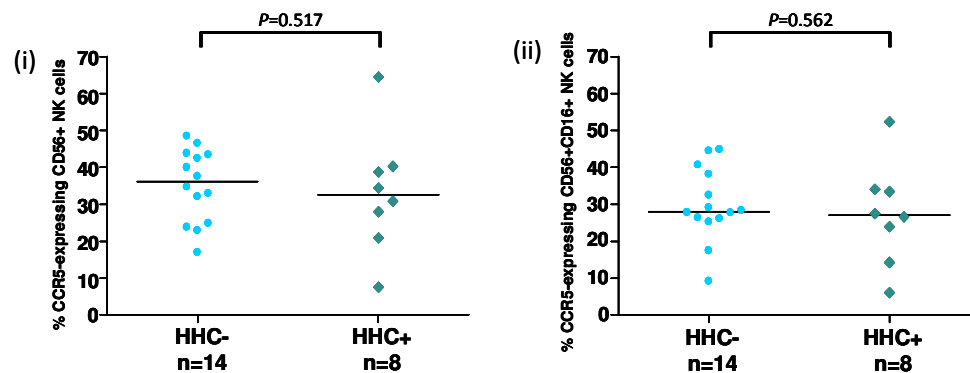
### a. SAC:



### b. SAC (without *CCR5Δ32*):



### c. SAA:



**Figure C.1.** Influence of the HHC haplotype on the percentage of CCR5-expressing cells on natural killer cell subsets, NK (i) and CD16+CD56+ NK (ii), within the South African Caucasian (SAC) population (A), the SAC population with the exclusion of individuals with the *CCR5Δ32* allele (B) and the SAA population (C). Individuals with or without an HHC allele are indicated as HHC+ and HHC-, respectively. The HHC homozygous individuals are indicated with (◊). P values and the number of individuals in each group are indicated.

## APPENDIX D

### Supplementary Data for Chapter 6

**Table D.1.** Comparison of CCL3 production in South African African (SAA) and Caucasian (SAC) individuals as measured in CCL3 production assays

<i>Comparison</i>	<i>Population*</i>	<i>log<sub>10</sub>(CCL3 production)</i>			<i>p<sup>†</sup></i> <i>(t-test)</i>
		<b>Mean</b>	<b>Range</b>	<b>Standard deviation</b>	
Unstimulated production of PBMCs	SAA (n=22)	1.882	1.100 - 3.034	0.395	0.0886
	SAC (n=31)	1.698	1.106 – 2.772	0.355	
Production of PHA-stimulated PBMCs	SAA (n=22)	4.231	3.814 – 4.551	0.229	0.8802
	SAC (n=31)	4.222	3.644 – 4.575	0.218	
Unstimulated production per <i>CCL3La</i> copy	SAA (n=22)	1.297	0.4001 – 2.557	0.427	0.3080
	SAC (n=29)	1.417	0.770 – 2.294	0.397	
PHA-stimulated production per <i>CCL3La</i> copy	SAA (n=22)	3.646	3.036 – 4.070	0.255	0.0002
	SAC (n=29)	3.941	3.343 – 4.472	0.286	

\* When calculating CCL3 production per copy number, two SAC individuals were excluded as these had a *CCL3La* copy number of zero.

† T-tests (UITENBROOK 1997) used to compare means yielded similar results to those of untransformed data analysed using nonparametric Mann-Whitney U tests.

**Table D.2.** Comparison of CCL3 production of phytohaemagglutinin stimulated cord blood mononuclear cells (CBMCs) isolated from HIV-1 intrapartum infected (IP) and exposed uninfected (EU) infants

<i>Comparison</i>	<i>Population</i>	<i>log<sub>10</sub>(CCL3 production)</i>			<i>p<sup>†</sup></i> <i>(t-test)</i>
		<b>Mean</b>	<b>Range</b>	<b>Standard deviation</b>	
Production of PHA-stimulated CBMCs	IP (n=13)	4.882	4.599 – 5.010	0.113	0.0018
	EU (n=41)	5.028	4.219 – 5.260	0.186	
PHA-stimulated production per <i>CCL3La</i> copy	IP (n=13)	4.339	4.122 – 4.533	0.139	0.0358
	EU (n=41)	4.452	3.617 – 4.739	0.222	

† T-tests (UITENBROOK 1997) used to compare means yielded similar results to those of untransformed data analysed using nonparametric Mann-Whitney U tests.

## APPENDIX E

### Supplementary Data for Chapter 7

**Table E.1.** Characteristics of individual HIV-1-infected long term nonprogressors and HIV-1-infected progressing study participants

Patient ID <sup>†</sup>	Lab ID	Age (years)	Sex	Viral load (copies/ml)	Log(viral load) <sup>*</sup>	CD4 count (cells/ $\mu$ l)	Time since diagnosis (years) <sup>*</sup>
<b>HIV-1 infected long term nonprogressors</b>							
<b>SAA (n=10)</b>							
LTNP1	TG1	38	♂	6 070	3.78	334	9
<b>LTNP2</b>	TG2	47	♀	5 780	3.76	400	6
LTNP3	TG4	35	♂	183	2.26	910	9
LTNP4	TG7	54	♂	23 300	4.37	968	4
LTNP5	TG9	46	♀	<400	2.60	327	9
LTNP6	TG11	32	♀	<40	1.59	693	7
LTNP7	Pru1	54	♀	<40	1.59	>2000	14
LTNP8	Pru2	43	♂	1 155	3.06	637	14
LTNP9	Pru3	36	♀	1 410	3.15	775	11
LTNP10	Pru4	38	♀	124	2.09	379	13
<b>SAC (n=4)</b>							
LTNP11	TG3	53	♂	<400	2.60	877	5
<b>LTNP12</b>	TG5	33	♀	5 420	3.73	396	5
LTNP13	TG8	67	♂	<40	1.59	811	6
LTNP14	TG10	40	♂	<400	2.60	327	18
<b>HIV-1 infected progressors</b>							
<b>SAA (n=10)</b>							
P1	M006	34	♀	491 000	5.69	195	-
P2	M019	31	♀	63 800	4.80	253	-
P3	M027	27	♀	56 900	4.76	40	-
P4	M032	22	♀	154 000	5.19	?	-
P5	M043	20	♀	460 000	5.66	81	-
P6	M124	36	♀	103 000	5.01	253	-
P7	M279	20	♀	55 600	4.75	219	-
P8	M282	29	♀	466 000	5.67	170	-
P9	M310	30	♀	1 590	3.20	86	-
P10	M415	19	♀	258 000	5.41	295	-

<sup>†</sup>LTNP individuals in bold font were started on antiretroviral treatment approximately one year after experiments were performed.

<sup>\*</sup>Where viral load was determined to be <40 or <400 copies/ml, log(viral load) was determined for values of 39 and 399, respectively.

<sup>\*</sup>Subsequent to performing these experiments, HIV-1-infected LTNPs have remained treatment naïve for a further two years. Time since diagnosis is unknown for the progressor group.

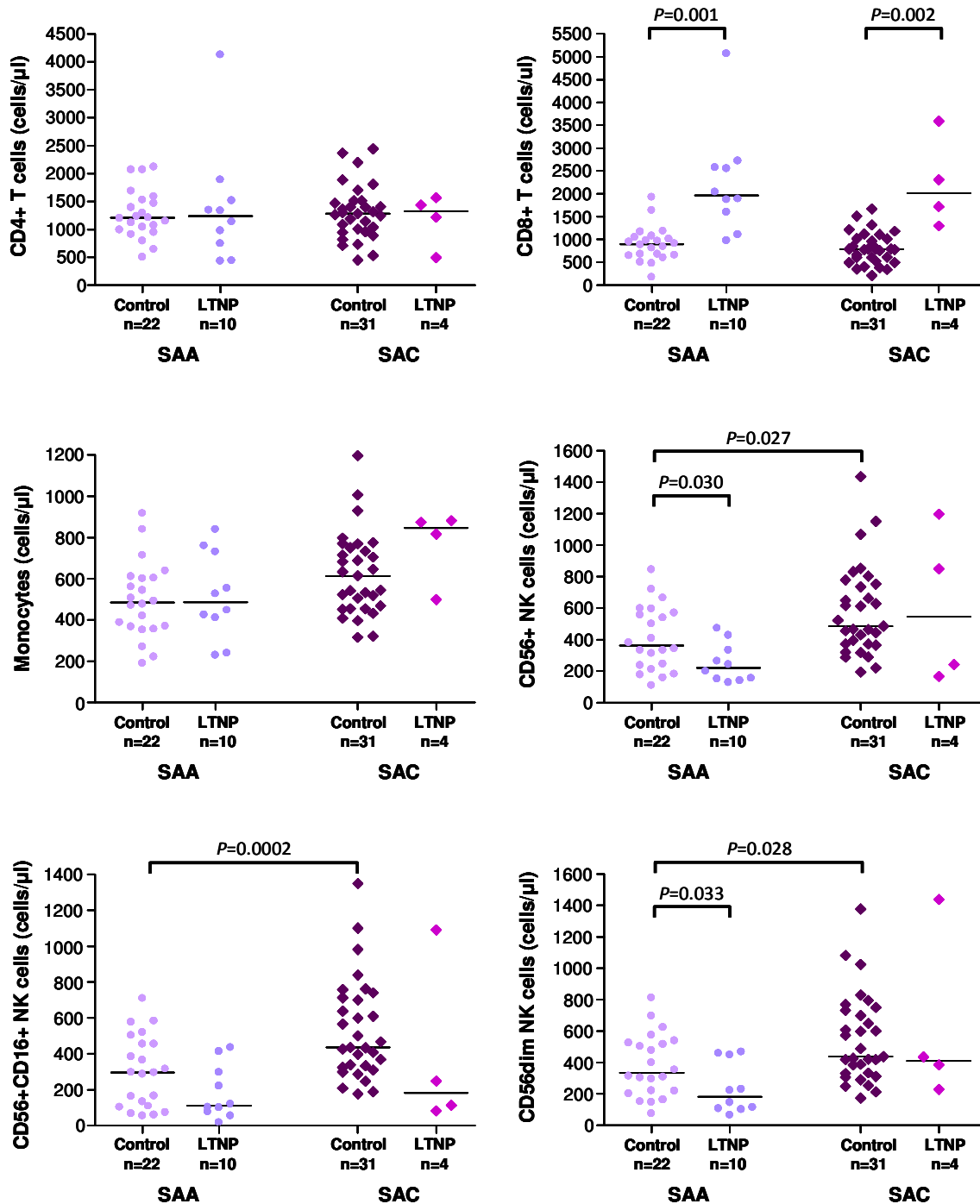
**Table E.2.** Comparison of CD4 counts determined by FACSCCount with those determined with FACSCalibur

Patient ID <sup>†</sup>	Lab ID	CD4 count (cells/μl)	Flow CD4 count (cells/ μl)	Flow CD8 count (cells/ μl)	CD4:CD8
<b>HIV-1 infected long term nonprogressors</b>					
<b>SAA (n=10)</b>					
LTNP1	TG1	334	433	1100	0.393
<b>LTNP2</b>	TG2	400	747	1597	0.468
LTNP3	TG4	910	1338	2036	0.657
LTNP4	TG7	968	445	2548	0.175
LTNP5	TG9	327	1517	2572	0.590
LTNP6	TG11	693	1138	2715	0.419
LTNP7	Pru1	>2000	4126	5068	0.814
LTNP8	Pru2	637	1348	1876	0.719
LTNP9	Pru3	775	1891	1891	1.000
LTNP10	Pru4	379	976	974	1.003
<b>SAC (n=4)</b>					
LTNP11	TG3	877	1439	3585	0.401
<b>LTNP12</b>	TG5	396	496	1721	0.288
LTNP13	TG8	811	1220	2307	0.529
LTNP14	TG10	327	1568	1293	1.212

**Table E.3.** Comparison of CD4 and CD8 counts\* and CD4:CD8 ratio between control uninfected and HIV-1-infected long term nonprogressing (LTNP) individuals in the South African African (SAA) and South African Caucasian (SAC) populations

Study group	CD4 count (cells/μl)	CD8 count (cells/μl)	CD4:CD8	<i>P</i> (CD4:CD8 control <i>cf.</i> LTNP)
<b>SAA</b>				
Control (n=22)	1208 (504 - 2120)	892 (174 - 1928)	1.47 (0.62 - 3.21)	<0.0001
LTNP (n=10)	1238 (433 - 4126)	1963 (974 - 5038)	0.62 (0.17 - 1.00)	
<b>SAC</b>				
Control (n=31)	1283 (453 - 2443)	786 (208 - 1672)	1.80 (0.63 - 4.00)	0.0047
LTNP (n=4)	1330 (496 - 1568)	2014 (1293 - 3585)	0.47 (0.29 - 1.21)	

\* Counts are based on extrapolations from flow cytometric bead analysis as outlined in materials and methods section



**Figure E.1.** Comparison of extrapolated absolute cell counts between control and HIV-1 infected long term nonprogressing (LTNP) individuals from two populations, South African African (SAA) and Caucasian (SAC). Where significant, *P* values are indicated.

---

## APPENDIX F

### Amino acid codes

---

Code	Amino acid
A	alanine
C	cysteine
D	aspartate
E	glutamate
F	phenylalanine
G	glycine
H	histidine
I	isoleucine
K	lysine
L	leucine
M	methionine
N	asparagine
P	praline
Q	glutamine
R	arginine
S	serine
T	threonine
V	valine
W	tryptophan
Y	tyrosine

---

---

**APPENDIX G**  
Ethical clearance

---