The Role of Killer Immunoglobulin-like Receptors (KIR) and Specific Human Leukocyte Antigen (HLA) Class I Molecules in Control of HIV-1 Infection

Nikki Lynne Gentle

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

Johannesburg 2015

Declaration

I, Nikki Lynne Gentle, hereby declare that:

a) The research reported in this thesis, except where otherwise indicated is my original work.
b) This thesis has not been submitted for any degree or examination to any other University
c) This thesis does not contain other people's data, pictures, graphs or information, unless specifically stated as being of sourced from other people.
d) This thesis does not contain another person's writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
    I. Their words have been re-written but the general information attributed to them has been referenced;
    II. Where their exact words have been used, their writing has been placed inside quotation marks and referenced.
e) Where I have reproduced a publication of which I am an author, co-author or editor, I have indicated in detail which part of the publication was written by myself alone and have fully referenced such publications.
f) This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being listed in the thesis and in the References section.

Signed_____ Date_____

Acknowledgements

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

My supervisor, Prof. Caroline Tiemessen, and co-supervisor, Dr Maria Paximadis, without whose guidance, patience and encouragement this would not have been possible.

The National Research Foundation and University of the Witwatersrand for funding this project.

Shayne Loubser and Heather Hong for their assistance with the *KIR* genotyping assays.

Prof. Adrian Puren and Dr Neil Martinson for allowing us access to their cohorts.

The doctors, nurses and, most importantly, participants in this study, whose contributions made this work possible.

My family, for their patience and unwavering support.

Presentations and Publications

Presentations:

1. *8th Conference of the Federation of African Immunological Societies (FAIS) Conference* (2012) Durban, South Africa. Genetic variability in markers of HLA-C expression in two diverse South African populations. Gentle, N.L., Paximadis, M., Puren, A. and Tiemessen, C.T.

2. *Joint International Conference of the African and Southern African Societies of Human Genetics* (2011) Cape Town, South Africa. KIR AA1 homozygotes and HIV-specific NK cell responses. Tiemessen CT, Gentle N, Paximadis M, Shalekoff S, Schramm DB, Minevich G, Winchester R, Gray GE, Martinson N, Kuhn L.

3. *4th MASIR (Measuring of Antigen-Specific Immune Responses) Conference* (2010) Mykonos, Greece. KIR AA1 Homozygotes and Understanding the Development of HIV-specific NK Cell Responses. Tiemessen C.T., Gentle N., Paximadis M., Shalekoff S., Schramm D.B., Minevich G., Winchester R., Gray G.E., Martinson N., Kuhn L.

Publications:

1. Loubser, S., Paximadis, M., Gentle, N.L., Puren, A., Gray, C.M. and Tiemessen, C.T. (2014) Frequencies of immune hypersensitivity reaction-associated HLA class I alleles in healthy South African Indian and mixed ancestry populations determined by a novel real-time PCR assay. *Tissue Antigens* **84**: 389 - 397

2. Gentle, N.L., Paximadis, M., Puren, A. and Tiemessen, C.T. Genetic variability in markers of HLA-C expression in two diverse South African populations. *PLoS One* **8**: e67780

3. Paximadis, M., Mathebula, T.Y., Vardas, E., Colvin, M., Gentle, N.L., Gray, C., Tiemessen, C.T. and Puren, A. (2012) HLA Class I and II (DRB1*) diversity in the Black and Caucasian South African population. *Human Immunology* **73**: 80-92

Abstract

The classical human leukocyte antigen (HLA) class I molecules are important regulators of both the adaptive and innate immune responses to viral infection. Genetic variability within these loci determines the nature of the interaction between both the T cell receptor (TCR) on $CD8^+$ T cells and specific killer immunoglobulin-like receptors (KIR) on the surface of natural killer (NK cell). Both of these interactions have previously been demonstrated to be important in determining the course of HIV-1 disease outcome. We therefore examined patterns of genetic variability within both of these complex gene families in individuals from the Black South African population group, contrasting them with genetic variability observed within the corresponding loci in Caucasian South Africans and demonstrated associations between specific genetic variants within the *HLA* and *KIR* gene complexes and HIV-1 control in the Black South African population.

Examination of genetic diversity within the *KIR* gene complex in the Black and Caucasian South African population groups revealed these two population groups differed significantly with respect to their *KIR2DS1* and *KIR3DS1* gene frequencies, as well as with respect to the full-length (*KIR2DS4f*) and truncated (*KIR2DS4v*) forms of *KIR2DS4*. Like *KIR2DS1* and *KIR3DS1*, *KIR2DS4v* was most frequently observed in the Caucasian population group, while *KIR2DS4f* was more frequently observed within the Black population group. These differences could be attributed to the different frequency distributions of specific telomeric KIR haplotype motifs within these two population groups. These findings are of particular importance in the South African context, given the associations of *KIR2DS4* and *KIR3DS1* with both HIV-1 transmission and disease progression.

An insertion-deletion (indel) polymorphism within the 3' untranslated region (UTR) of *HLA-C* has also been shown to be involved in the regulation of HLA-C expression. Individuals who carry a deletion at this position exhibit increased HLA-C expression, which associates with lower viral set point in HIV-1 infected individuals. This *263* indel (rs67384697) is reported to be in strong linkage disequilibrium (LD) with a single nucleotide polymorphism (SNP) 35 kilobases upstream of *HLA-C* (*-35T/C*; rs9264942) in Caucasian individuals, making this SNP a potential marker for both HLA-C expression and HIV-1 disease progression. We therefore examined genetic variation within the UTRs of the *HLA-C* alleles present in Black and Caucasian South Africans and identified two overlapping haplotypes encompassing the *263* indel and another indel at position 230 in both populations, which we propose may act in concert to regulate levels of HLA-C expression. Concomitant evaluation of variability at the *-35* SNP revealed this polymorphism to be an inappropriate marker for either indel in these populations.

Recently, individual polymorphic amino acids within the classical *HLA* class I loci, located predominantly within the peptide binding groove, have been shown to be strongly associated with HIV-1 control. We, therefore, examined patterns of genetic variability within and across the *HLA* class I loci in Black South African HIV-1 progressors and –controllers. Our findings confirmed those from other populations, demonstrating the importance of HLA-B residues 67, 70, 97 and 116 in determining disease outcome, while also identifying additional residues in HLA-A and -B that may potentially contribute to determining differential disease outcome in this population. Variability at these residues likely impacts the specificity of the peptide bound by the HLA molecule, resulting in differential regulation of both cytotoxic T lymphocyte (CTL) and natural killer (NK) cell responses. No significant associations were observed between HIV-1 control and variability within either the HLA-C peptide binding groove or the 3’ UTR.

Finally, we examined the role of genetic variability within the *KIR* gene complex in regulating HIV-1 control by examining patterns of genetic variability within this locus in Black South African HIV-1 progressors and –controllers. We found loss of control to be significantly associated with specific *KIR* haplotype motifs lacking *KIR2DS4* and *KIR3DL1*, while maintenance of viral control was found to be associated with possession of *KIR* haplotypes containing the centromeric cB01 motif. Furthermore, elite controllers were more frequently found to be in possession of cB01 motifs containing *KIR2DS5*, rather than *KIR2DS3*. In light of the strong linkage disequilibrium observed across this region, *KIR2DS3* and *KIR2DS5* are thought act as markers for specific allelic variants of the inhibitory receptors KIR2DL1 and KIR2DL2, which are known to mediate differential inhibition of NK cell function.

Collectively, these data represent the first comprehensive description of genetic variability within the *KIR* gene complex in Black South Africans and provide the valuable insights into the role of these receptors in mediating control of HIV-1 infection through interaction with their *HLA* class I -encoded ligands.

List of Abbreviations

ADCC	Antibody-dependent cellular cytotoxicity
AIDS	Acquired immunodeficiency syndrome
ART	Antiretroviral treatment
BGB	Beta-globin
CI	Confidence interval
CTL	Cytotoxic T lymphocyte
DNA	Deoxyribonucleic acid
GWAS	Genome-wide association study
ELISA	Enzyme-linked immunosorbent assay
HWE	Hardy-Weinberg equilibrium
HLA	Human leukocyte antigen
HIV-1	Human immunodeficiency virus type 1
Indel	Insertion-deletion
kb	kilobase
KIR	Killer cell immunoglobulin-like receptors
LD	Linkage disequilibrium
LILRs	Leukocyte immunoglobulin-like receptors
LNA	Lock nucleic acid
LRC	Leukocyte receptor complex
MHC	Major histocompatibility complex
µl	Microlitre
µM	Micromolar
mRNA	Messenger RNA
miRNA	microRNA
ml	Millilitre
mm^3	Millimetre cubed
ng	Nanogram
NGS	Next-generation sequencing
NK	Natural killer
OR	Odds ratio
PBMC	Peripheral blood mononuclear cell
PCR	Polymerase chain reaction
pmol	Picomole
qPCR	Quantitative Polymerase Chain Reaction
RNA	Ribonucleic acid
SBT	Sequence based typing
SNP	Single nucleotide polymorphism
SSP-PCR	Site-specific primer-polymerase chain reaction
TCR	T cell receptor
UTR	Untranslated region

Table of Contents

Declaration ii
Acknowledgements iii
Presentations and Publications iv
Abstract v
List of Abbreviations viii
List of Tables xii
List of Figures xiv
Chapter 1 1
Introduction 1
1.1 Genetic Variation in Africans 2
1.2 The Human Leukocyte Antigen (HLA) Complex 3
1.2.1 HLA Class I Polymorphism 3
1.2.2 HLA Class I Function 4
1.2.3 The Dominant Influence of HLA-B on HIV-1 Control 5
1.3 The Killer Immunoglobulin-like Receptor (KIR) Family 7
1.3.1 The KIR Gene Complex 7
1.3.2 KIR Structure 7
1.3.3 KIR Signalling 8
1.3.4 KIR Diversity 8
1.3.5 KIR Ligands 10
1.3.6 The Role of KIR-HLA Ligand Interactions in HIV-1 Control 12
1.4 Study Objectives 13
Chapter 2 15
The Distribution of Killer Cell Immunoglobulin-like Receptor (KIR) Genes and Their HLA Class I Ligands in Two Ethnically Distinct South African Populations 15
2.1 Introduction 15
2.2 Materials and Methods 17
2.2.1 Study Populations 17
2.2.2 KIR Genotyping 17
2.2.3 KIR Genotype Assignment 18
2.2.4 KIR Haplotype Estimation 18
2.2.5 HLA Class I Ligand Determination 19
2.2.6 Statistical Analyses 19
2.3 Results 20
2.3.1 KIR Gene Diversity 20

2.3.2 KIR Genotypes Distributions 23
2.3.3 KIR Haplotype Estimation 23
2.3.4 Estimation of Linkage Disequilibrium 27
2.3.5 Association between Specific KIR and Their HLA Class I Ligands 30
2.3.5.1 HLA-B and KIR3DL1/DS1 30
2.3.6 HLA-C and its Associated KIR Receptors 33
2.4 Discussion 33
Chapter 3 39
Genetic Variability in Markers of HLA-C Expression Show Similarity between Two Diverse South African Populations 39
3.1 Introduction 39
3.2 Materials and Methods 40
3.2.1 Study Populations 40
3.2.2 HLA-C Genotyping 41
3.2.3 HLA-C 3' UTR DNA Sequencing 41
3.2.4 -35 SNP Genotyping 42
3.2.5 Statistical Analyses 43
3.3.1 Genetic Variation within the HLA-C 3' UTR 43
3.3.2 Estimation of Haplotype Structure within the HLA-C 3' UTR 47
3.3.3 Linkage Disequilibrium between Specific HLA-C alleles and the -35 SNP 49
3.3.4 Linkage Disequilibrium between the -35 SNP and the 263 Indel 51
3.4 Discussion 51
Chapter 4 57
Genetic Variability at Residues within HLA-A and HLA-B that Determine Peptide Specificity are Associated with HIV-1 Control in Black South Africans 57
4.1 Introduction 57
4.2 Materials and Methods 58
4.2.1 Study Populations 58
4.2.2 DNA Extraction 59
4.2.3 HLA Class I Genotyping 60
4.2.4 Resolution of HLA Ambiguity 60
4.2.5 Evaluation of the Polymorphic Positions within HLA Class I Loci 61
4.2.6 HLA-C 3' UTR DNA Sequencing 61
4.2.7 Statistical Analyses 62
4.3 Results 62
4.3.1 Association of specific HLA Class I alleles with HIV-1 Control 62
4.3.2 Association of Polymorphic Positions within the HLA Class I Peptide-binding Groove with HIV-1 Control 63

4.3.3 Association of Polymorphic Positions within the HLA-C 3' UTR with HIV-1 Control 68
4.4 Discussion 70
Chapter 5 75
Genetic Variability within the KIR Gene Complex is Associated with HIV-1 Control and Distinguishes Elite Controllers from Viraemic Controllers 75
5.1 Introduction 75
5.2 Materials and Methods 76
5.2.1 Study Populations 76
5.2.2 DNA Extraction 77
5.2.3 KIR Genotyping 78
5.2.4 KIR Copy Number Determination 78
5.2.5 HLA Class I Ligand Determination 79
5.2.6 Statistical Analyses 79
5.3 Results 80
5.3.1 KIR Gene Diversity 80
5.3.2 KIR Genotype Distributions 82
5.3.3 Estimation of KIR Haplotypes 82
5.3.4 Examination of KIR Haplotype Motifs 85
5.3.5 Association between Specific KIR and their HLA Class I Ligands 88
5.3.5.1 HLA-B and KIR3DL1 88
5.3.5.2 HLA-C and its Associated KIR Receptors 89
5.4 Discussion 90
Chapter 6 95
Concluding Remarks 95
References 105
Appendix A 123
Appendix B 124

List of Tables

Table 2.1: The gene carrier frequencies of the *KIR* genes present in the Black and Caucasian South African population groups……..…21

Table 2.2: The *KIR* genotype profiles identified within the South African population and their frequencies in the Black and Caucasian population groups…….....24

Table 2.3: The frequencies of the *KIR* haplotypes occurring in the Black and Caucasian South African population groups…….....26

Table 2.4: The telomeric *KIR* haplotype motifs present in the South African population and their frequencies within the Black and Caucasian population groups……..27

Table 2.5: The estimated linkage disequilibrium across the *KIR* gene complex in the Black South African population…….....28

Table 2.6: The estimated linkage disequilibrium across the *KIR* gene complex in the Caucasian South African population…….....29

Table 2.7: The frequency with which *KIR3DL1* occurs in combination with the HLA Bw4 and Bw6 allotypes in the Black and Caucasian South African population groups…….31

Table 3.1: The polymorphic positions present within the *HLA-C* 3'UTR and their minor allele frequencies within the Black and Caucasian population groups…….….45

Table 3.2: LD between the -35 SNP (rs9624942) and the *HLA-C* alleles present in the Black and Caucasian South African population groups…….50

Table 4.1: A summary of the clinical characteristics used to classify individuals as either HIV-1 progressors or –controllers……..……..59

Table 4.2: *HLA* class I alleles associated with differential HIV-1 control…….…….63

Table 4.3: The polymorphic positions within the *HLA-A* locus associated with differential HIV-1 control……64

Table 4.4: The polymorphic positions within the *HLA-B* locus associated with differential HIV-1 control……...…….66

Table 4.5: The proportion of HIV-1 progressors and –controllers in possession of specific *230* (rs35877659) and *263* (rs67384697) indel alleles and genotypes……69

Table 5.1: The frequencies of the *KIR* genotype profiles present within HIV-1 progressors and controllers……...……..83

Table 5.2: The frequencies of the *KIR* haplotypes occurring within HIV-1 progressors, viraemic controllers and elite controllers……85

Table 5.3: The frequency with which *KIR3DL1* occurs in combination with *HLA-A* and *HLA-B* alleles encoding its Bw4 ligand in HIV-1 progressors, viraemic controllers and elite controllers……...……..…89

Table 5.4: The frequencies with which *KIR2DL1*, *KIR2DL2*, *KIR2DL3* and *KIR2DS1* occur in combination with HLA-C alleles encoding the C1 and C2 epitopes in HIV-1 progressors, viraemic controllers and elite controllers……...……..…90

List of Figures

Figure 1.1: The TCR and KIR binding sites occur within overlapping regions of the HLA class I α domains……....4

Figure 1.2: The domain structure of KIR receptors……....9

Figure 1.3: A schematic representation of the most commonly observed *KIR* haplotype motifs……10

Figure 2.1: A two-dimensional heatmap based on hierarchical cluster of the carrier frequencies of the non-framework *KIR* genes in African and Caucasian populations.……22

Figure 2.2: The frequency with which HLA-Bw4 and -Bw6 allotypes occur in individuals from the Black and Caucasian South African population groups in possession of *KIR3DS1*……..31

Figure 2.3a: The frequency with which *KIR2DL1*, *KIR2DL2*, *KIR2DL3* and *KIR2DS1* occur in combination with the HLA-C1 in individuals from the Black and Caucasian South African population groups…….32

Figure 2.3b: The frequency with which *KIR2DL1*, *KIR2DL2*, *KIR2DL3* and *KIR2DS1* occur in combination with the HLA-C2 in individuals from the Black and Caucasian South African population groups…….32

Figure 3.1: The association between the polymorphic variants observed within the *HLA-C* 3' UTR and specific *HLA-C* alleles……..…….....…...46

Figure 3.2a: The haplotypes occurring within the 3' UTRs of the *HLA-C* alleles present in the Black and Caucasian South African population groups……...48

Figure 3.2b: The combinations of the 230H and 263H haplotypes observed within the Black and Caucasian South African population groups…….……48

Figure 4.1: The relationship between individual polymorphisms within *HLA-A* associated with differential HIV-1 control and specific *HLA-A* alleles…….65

Figure 4.2: The relationship between individual polymorphisms within *HLA-B* associated with differential HIV-1 control and specific *HLA-B* alleles…….67

Figure 5.1: The gene carrier frequencies of the *KIR* genes present within HIV-1 progressors, viraemic controllers and elite controllers……..……...81

Figure 5.2: The *KIR* genotypes present in HIV-1 viraemic- and elite controllers and their frequencies……..……..84

Figure 5.3: The frequencies of the centromeric *KIR* haplotype motifs observed within HIV-1 progressors, viraemic controllers and elite controllers…….……86

Figure 5.4: The frequencies of the centromeric *KIR* haplotype motif combinations observed in HIV-1 progressors, viraemic controllers and elite controllers……..……....87

Chapter 1

Introduction

To date it is estimated that more than 60 million people worldwide have been infected with HIV-1, with 35 million people currently thought to be living with the virus (UNAIDS, 2015). Most individuals are thought to exhibit a similar disease course, characterized by an asymptomatic period of eight to ten years following the initial phase of acute infection, during which ongoing viral replication induces immune dysregulation and eventually death (Deeks and Walker, 2007). However, subsets of individuals who exhibit extreme rates of disease progression have been identified.

One such group, termed HIV-1 controllers, includes individuals able to spontaneously suppress viral replication and maintain very low levels of viraemia (< 2000 RNA copies/ml) in the absence of highly active antiretroviral therapy (Walker and Yu, 2013). Within this group, a still smaller percentage of individuals, known as elite controllers, can be further distinguished from other HIV-1 controllers by their ability to suppress viral replication to virtually undetectable levels (Deeks and Walker, 2007). While the ability of these individuals to suppress viral replication has been shown to be associated with a number of both viral and host genetic characteristics - suggesting that control of HIV-1 viral replication is mediated by multiple immunological mechanisms - it is hoped that identification and characterization of immune signatures that distinguish controllers from other HIV-1 infected individuals will aid in the identification of future therapeutic targets (Deeks and Walker, 2007; Walker and Yu, 2013).

1.1 Genetic Variation in Africans

The importance of differences in genetic variation within and between individuals in determining disease outcome is becoming increasingly apparent (Hirschhorn *et al*., 2002). What is also becoming clear, is that geographically distinct populations with differing demographic histories often display marked differences in patterns of genetic variability (Rosenberg *et al*., 2002; Conrad *et al.*, 2006). These differences are often most pronounced when comparisons are drawn between African and non-African populations, as African populations often exhibit increased levels of genetic variability relative to their non-African counterparts (Tishkoff and Williams, 2002). Africans also typically display lower levels of linkage disequilibrium (LD) than non-African populations (Reich *et al*., 2001), with the result that African populations exhibit greater haplotypic diversity (Stephens *et al*., 2001; Conrad *et al*., 2006). This has been attributed to the fact that African populations have consistently maintained a larger effective population size than non-African populations, which have been subject to extensive genetic drift as a consequence of having undergone bottleneck and founder effects following the migration out of Africa (Quintana-Murci *et al*., 1999).

It is now well established that genetic variation plays an important role in determining an individual's susceptibility to disease (Tishkoff and Verrelli, 2003). Differential disease susceptibilities and rates of disease progression between African and non-African populations, stemming from differences in patterns of genetic variability between these population groups, have been described for a number of complex diseases (Campbell and Tishkoff, 2008). Differences in patterns of genetic variability between African and non-African populations are also becoming increasingly apparent within genes influencing HIV-1 transmission and disease outcome (Gonzalez *et al*., 2001; Han *et al*., 2008; Pelak *et al*., 2010). Many of these genes encode products that are involved in viral entry into the cell, immune recognition and antigen

presentation, and as a result, polymorphic variations in these genes can have profound effects on host-pathogen interactions (Winkler *et al.*, 2004; Kaslow *et al.*, 2005; Brass *et al.*, 2008).

1.2 The Human Leukocyte Antigen (HLA) Complex

Genetic variability within the *human leukocyte antigen* (*HLA*) complex has been identified as one of the major determinants of HIV-1 control (Fellay *et al*., 2007; The International HIV Controller Study, 2010). The *HLA* class I loci, *HLA-A*, *HLA-B* and *HLA-C*, are located within the *major histocompatibility complex* (*MHC*), situated on the short arm of chromosome 6 (The MHC sequencing consortium, 1999) and are the most polymorphic genes in the human genome (Robinson *et al*., 2014). These genes encode membrane-bound glycoproteins, found on the surface of all nucleated cells and platelets (Fleming *et al*., 1981), which mediate immune responses through their interactions with multiple cell types (Salter *et al*., 1990; Colonna and Samaridis, 1995).

1.2.1 HLA Class I Polymorphism

HLA molecules expressed at the cell surface are typically non-covalently linked heterodimers, consisting of a highly polymorphic heavy (H) chain in complex with an invariant β_2-microglobulin light chain and a short (generally 8 – 10 amino acids in length) bound peptide (Bjorkman *et al*., 1987a). The extracellular domain of the H chain is encoded within exons 2, 3 and 4 of each gene, and is comprised of three α domains that collectively constitute the peptide-binding groove (Bjorkman *et al*., 1987b). Genetic variability within these α domains determines both the specificity of the peptide bound by the HLA molecule (Bjorkman *et al*., 1987b), as well as the nature of the interaction between the peptide-bound HLA molecule and its cognate receptors (Salter *et al*., 1990; Fadda *et al*., 2011).

1.2.2 HLA Class I Function

HLA class I molecules bind both the αβ T cell receptor (TCR) found on the surface of $CD8^+$ T cells (Salter *et al*., 1990), as well as specific killer immunoglobulin-like receptors (KIR) expressed (predominantly) on the surface of natural killer (NK) cells (Lanier, 1998). The binding sites for the TCR and inhibitory KIR receptors within HLA class I molecules are encoded within overlapping regions of the highly polymorphic α1 domain (Figure 1.1; Boyington and Sun, 2002). This overlap is anticipated to result in competing selective pressures being exerted on the *HLA* class I loci by the adaptive and innate immune responses (Older Aguilar *et al*., 2010). As a counter-measure to this, human-specific evolution of the *KIR* loci encoding inhibitory receptors has resulted in reduced affinity for HLA-A and –B, while diversifying the interaction between these receptors and HLA-C (Older Aguilar *et al*., 2010; Parham, 2012). Consequently, while HLA-B has emerged as the dominant regulator of cytotoxic T lymphocyte (CTL) responses (Bihl *et al*., 2006), binding of HLA-C to inhibitory KIR is thought to be the interaction primarily responsible for regulating NK cell function (Moesta and Parham, 2012).

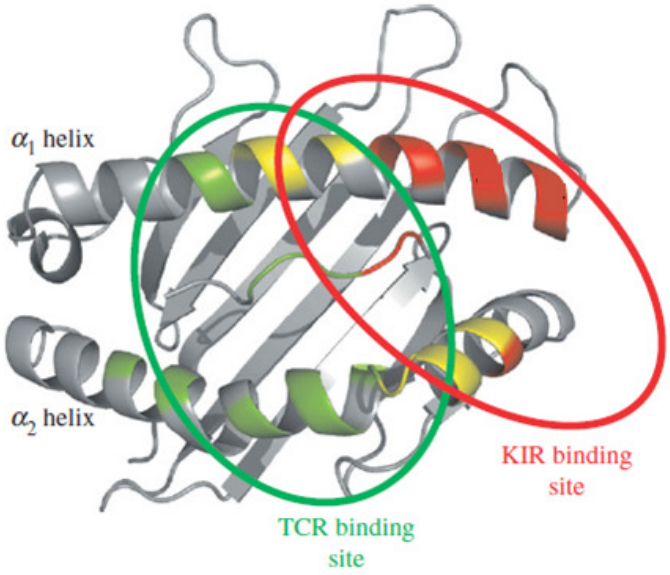


Figure 1.1: The TCR and KIR binding sites occur within overlapping regions of the HLA class 1 α domains. The KIR and TCR binding sites are indicated in red and green, respectively. Image modified from Parham (2012).

1.2.3 The Dominant Influence of HLA-B on HIV-1 Control

Allelic variability within *HLA-B* has repeatedly been shown to be strongly associated with the rate of HIV-1 disease progression. While multiple alleles have been shown to be associated with differential rates of HIV-1 disease progression (Costello *et al*., 1999; Kiepiela *et al*., 2004; Frahm *et al*., 2005; Lazaryan *et al*., 2006; Honeybourne *et al*., 2007; Lazaryan *et al*., 2010; Leslie *et al*., 2010; Tang *et al*., 2011), the *HLA-B*57* alleles (*B*57:01:01*, *B*57:02:01* and *B*57:03:01*) in particular have consistently been shown to be strongly associated with HIV-1 control (Kiepiela *et al*., 2004; Costello *et al*., 2008; Leslie *et al*., 2010; The International HIV Controller Study, 2010; McLaren *et al.*, 2012). These associations are thought to be predominantly as a result of targeting of specific immunodominant epitopes in Gag early in infection (Zuñiga *et al*., 2006).

Consistent with the dominant influence of HLA-B in the regulation of CTL responses is the recent identification of specific residues encoded within the *HLA-B* α domains that have been shown to strongly associate with HIV-1 control (The International HIV Controller Study, 2010; McLaren *et al*., 2012). These residues are involved in determining the structural integrity of the peptide-binding groove, as well as mediating the interaction with the bound peptide (Bjorkman *et al*., 1987b). Specificity of the peptide bound by HLA class I molecules not only determines the nature of their interactions with their cognate receptors (Fadda *et al*., 2011), but also limits their capacity to elicit CTL responses due to reductions in their surface expression (Rizvi *et al*., 2014).

The roles of HLA-A and HLA-C in mediating control of HIV-1 infection can also not be discounted. As a result of the presence of strong LD across the *HLA* gene complex, *HLA-B* alleles associated with differential HIV-1 disease progression are commonly found to occur in combination with specific *HLA-A* and *HLA-C* alleles (Kiepiela *et al*., 2004; Leslie *et al*., 2010;

Tang *et al*., 2011) and while the effects of these haplotypes are often attributed solely to functional interactions mediated through HLA-B, there is evidence to support the idea that these allelic combinations may act in concert and contribute additively to differentially regulate HIV-1 disease progression (Leslie *et al*., 2010).

1.2.4 The Influence of HLA-C Expression on HIV-1 Control

Genome-wide association studies (GWAS) have shown a single nucleotide polymorphism (SNP) 35 kilobases (kb) upstream of *HLA-C* to be strongly associated with HIV-1 control (Fellay *et al*., 2007; Fellay *et al*., 2009. This SNP (*-35T/C*; rs9264942) also strongly associates with differences in HLA-C mRNA (Stranger *et al*., 2005; Fellay *et al*., 2007) and cell surface expression levels (Thomas *et al*., 2009). However, while these associations have consistently been shown in Caucasian cohorts (Fellay *et al*., 2007; Dalmasso *et al*., 2008; Fellay *et al*., 2009; Thomas *et al*., 2009; van Manen *et al*., 2009; The International HIV Controller Study, 2010; Ballana *et al*., 2012), the association between the *-35 SNP* and HIV-1 viral set point has not been shown to be significant in African American cohorts - despite the presence of this SNP in this population (Han *et al*., 2008; Shrestha *et al*., 2009; The International HIV Controller Study, 2010; Apps *et al*., 2013). This has lead to the suggestion that the *-35 SNP* is not the causative variant responsible for the alteration in HLA-C expression and viral set point, but rather acts as a marker, in Caucasian populations, for another functionally relevant polymorphism.

A single base pair insertion-deletion (indel) polymorphism at position 263 of the *HLA-C* 3' untranslated region (UTR) has since been identified that has been shown to be in strong LD with the *-35 SNP* in Caucasian individuals (Kulkarni *et al*., 2011). Genetic variability at this position (*263* indel; rs67384697) affects the binding of a regulatory microRNA (miRNA148a) to the *HLA-C* 3' UTR, with a deletion at this position (*263del*) abolishing miRNA148a binding and leading to increased HLA-C expression (Kulkarni *et al*., 2011; Kulkarni *et al*., 2013). The *263del* allele at this position has also been shown to be over-represented in HIV-1 controllers of

European descent (Kulkarni *et al*., 2011). Given that increased HLA-C expression is associated with HIV-1 viral control (Apps *et al*., 2013), it is thought that the changes in HLA-C expression associated with this allele could help to provide long-term protection against HIV-1 disease progression (Kulkarni *et al*., 2011).

1.3 The Killer Immunoglobulin-like Receptor (KIR) Family

1.3.1 The KIR Gene Complex

KIR are type I integral membrane glycoproteins that are usually expressed as monomers on the surface of NK cells and a subset of T cells (Lanier, 1998). They are encoded by a family of rapidly evolving genes found on chromosome 19q13.4 (Wilson *et al*., 2000; Martin *et al*., 2000). To date, 14 *KIR* genes (*KIR2DL1, KIR2DL2, KIR2DL3, KIR2DL4, KIR2DL5, KIR2DS1, KIR2DS2, KIR2DS3, KIR2DS4, KIR2DS5, KIR3DL1, KIR3DS1, KIR3DL2,* and *KIR3DL3*) and two pseudogenes (*KIR2DP1* and *KIR3DP1*) have been identified in humans. They are arranged in a head-to-tail manner, approximately 2 kb apart, within a 150 kb region called the *leukocyte receptor complex* (*LRC*) (Wilson *et al*., 2000).

1.3.2 KIR Structure

KIR genes typically consist of nine exons, which encode a leader sequence (exons 1 and 2), two or three extracellular immunoglobulin (Ig)-like domains called D0, D1 and D2 (exons 3-5), a stem (exon 6), transmembrane region (exon 7) and cytoplasmic tail (exons 8 and 9) (Vilches *et al*., 2000). The receptors are classified as either two-domain (KIR2D) or three-domain receptors (KIR3D), based on how many extracellular domains are present (Figure 1.2). They are further classified as either short-tailed (S) or long-tailed (L) based on the length of their cytoplasmic tails. Based on the presence or absence of these domain structures, human KIR can be classified into four lineages (I, II, III and V; Sambrook *et al*., 2005). Lineage I KIR, namely KIR2DL4 and KIR2DL5, contain the D0 and D2 extracellular domains; while KIR3DL1 and KIR3DL2

are classified as lineage II KIR. The remaining two- and three-domain receptors are classified as lineage III KIR, with KIR3DL3 being the only lineage V receptor.

1.3.3 KIR Signalling

KIR regulate NK cell function by modulating a delicate balance between activating and inhibitory signals (Lanier, 2004). Inhibitory KIR receptors have two immunoreceptor tyrosine-based inhibition motifs (ITIMs) in their cytoplasmic tail that become phosphorylated at specific tyrosine residues upon ligand binding (Fry *et al*., 1996). These then recruit the phosphatases SHP-1 and SHP-2, which dephosphorylate several tyrosine kinases and inhibit NK cell-mediated killing and the secretion of cytokines (Yusa and Campbell, 2003). Activating KIR receptors, on the other hand, associate with the adapter protein DAP-12 via a charged residue within their transmembrane region, which is subsequently phosphorylated at tyrosine residues within its immunoreceptor tyrosine-based activation motif (ITAM) domain (Campbell *et al*., 1998). DAP-12 then recruits several tyrosine kinases that activate NK cell-mediated killing and signal the cell to secrete cytokines (Carr *et al*., 2007; Mulrooney *et al*., 2013). This generalization appears to hold true for all KIR receptors, except KIR2DL4, which has a long cytoplasmic tail containing an ITIM domain, but also has a charged residue within its transmembrane region that allows the receptor to non-covalently associate with the γ chain of the FcεRI when bound to HLA-G (Rajagopalan *et al*., 2001; Faure and Long, 2002; Kikuchi-Maki *et al*., 2003).

1.3.4 KIR Diversity

KIR exhibit extensive genetic diversity with respect to gene content, copy number and allelic representation within individual genes (Uhrberg *et al*., 1997; Schilling *et al*., 2002, Hou *et al*., 2012; Jiang *et al*., 2012). Variability in terms of *KIR* gene content is determined by the presence of two main *KIR* haplotypes (Uhrberg *et al*., 2002, Martin *et al*., 2004), termed A and B, which are themselves the product of varying combinations of a limited number of centromeric and

telomeric gene content motifs (Figure 1.3; Pyo *et al*., 2010, Jiang *et al*., 2012; Vierra-Green *et al*., 2012, Pyo *et al*., 2013).

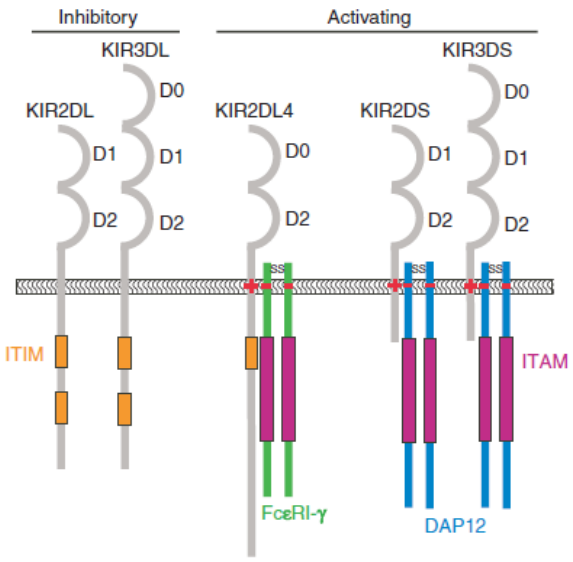


Figure 1.2: The domain structure of KIR receptors. ITIM sequences are indicated in yellow, while ITAM sequences are shown in purple. Taken from Campbell and Purdy (2011).

The A haplotype is characterized by the presence of the genes encoding the inhibitory receptors KIR2DL1 and KIR3DL1, the activating receptor KIR2DS4 and the pseudogene, *KIR2DP1*, while the B haplotype is characterized by the presence of *KIR2DL5* and the genes encoding the activating receptors KIR2DS1, KIR2DS2, KIR2DS3, KIR2DS5 and KIR3DS1 (Uhrberg *et al*., 2002; Martin *et al*., 2004). Four genes, namely *KIR2DL4*, *KIR3DL2*, *KIR3DL3* and *KIR3DP1*, are common to both haplotypes and are referred to as framework genes, as they demarcate the boundaries of the centromeric and telomeric haplotype motifs (Wilson *et al*., 2000; Pyo *et al*., 2010). *KIR* genotypes resulting as a consequence of combinations of A haplotypes are designated as AA genotypes, while those composed of combinations of both A and B haplotypes are collectively designated as Bx genotypes. The A and B haplotypes (and by extension the AA and Bx genotypes) are present at different frequencies in different human populations (Gonzalez-Galarza *et al*., 2011).

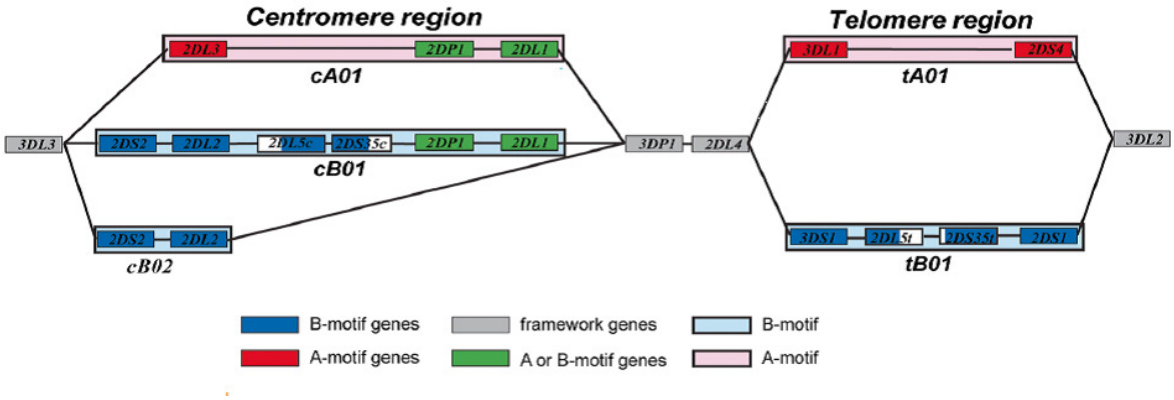


Figure 1.3: A schematic representation of the most commonly observed *KIR* haplotypic motifs. Image taken from Pyo *et al*., (2013).

KIR diversity is also regulated at the transcriptional level. Promoter allelic polymorphism (Li *et al*., 2008) and alternative methylation patterns within promoters (Chan *et al*., 2003) synergize to mediate allele-specific KIR expression, while the presence of multiple (Trompeter *et al*., 2005), bi-directional promoters (Davies *et al*., 2007) allow for alternative modes of transcription (Stulberg *et al*., 2006). Collectively, these mechanisms aid in inducing clonal and stochastic expression of KIR on NK cells (Valiante *et al*., 1997), with the result that possession of a particular gene does not necessarily translate into expression of the receptor.

1.3.5 KIR Ligands

Multiple KIR receptors have been shown to bind HLA class I, with allelic polymorphism in both the receptor and the ligand determining the nature of the interaction. Lineage II KIR receptors bind HLA-A (Döhring *et al.*, 1996; Hansasuta *et al*., 2004; Thananchai *et al.*, 2007, Stern *et al*., 2008) and –B (Cella *et al*., 1994; Gumperz *et al*., 1995), while lineage III KIR bind predominantly HLA-C (Colonna *et al*., 1993; Winter and Long, 1997; Winter *et al*., 1998; Graef *et al*., 2009; Moesta *et al*., 2010) and to a lesser extent, HLA-A (Katz *et al*., 2001; Graef *et al*., 2009; Liu *et al*., 2014). With respect to the two lineage I KIR, KIR2DL4 binds HLA-G (Rajagopalan and Long, 1999), while the ligand for KIR2DL5 has yet to be described. Variants

of this receptor have, however, been shown to be expressed at the cell surface (Estefanía *et al*., 2007) and are capable of inhibiting NK cell activity (Yusa *et al*., 2004).

HLA-A and -B molecules containing the Bw4 epitope serve as ligands for KIR3DL1 (Cella *et al*., 1994; Gumperz *et al*., 1995, Thananchai *et al.*, 2007, Stern *et al*., 2008). This epitope is defined by the amino acids at positions 77-83 of the HLA α1 domain. The affinity with which KIR3DL1 binds HLA-Bw4 allotypes depends on whether they have a threonine (HLA-Bw4-80T) or isoleucine (HLA-Bw4-80I) at position 80 of their amino acid sequence (Cella *et al*., 1994). HLA-Bw4-80I allotypes bind KIR3DL1 with higher affinity than HLA-Bw4-80T allotypes (Martin *et al*., 2002a). Binding affinity is also determined by residues within KIR3DL1 (Carr *et al*., 2005). Despite a high degree of sequence similarity with KIR3DL1 and evidence of the ability to trigger NK cell activity (Carr *et al*., 2007), KIR3DS1 has not been shown to bind HLA-Bw4 (Gillespie *et al.*, 2007). KIR3DL2 has also been shown to bind particular HLA-A antigens in a peptide-specific fashion (Döhring *et al.*, 1996; Hansasuta *et al*., 2004).

The lineage III inhibitory receptors, KIR2DL1, KIR2DL2 and KIR2DL3 (Colonna *et al*., 1993; Winter and Long, 1997, Winter *et al*., 1998), and the activating receptor KIR2DS1 (Stewart *et al*., 2005; Moesta *et al*., 2010) are known to bind to HLA-C allotypes. The specificities of these interactions are determined at positions 77 and 80 in HLA-C (Biassoni *et al.*, 1995; Winter *et al*., 1998) and position 44 within the KIR receptors (Winter and Long, 1997). HLA-C molecules with an asparagine at position 80 are designated as HLA-C1; whereas those with a lysine at position 80 are referred to as HLA-C2 (Biassoni *et al*., 1995; Winter *et al*., 1998). KIR with a methionine residue at position 44 (KIR2DL1) will preferentially bind HLA-C2, while KIR with lysine at this position (KIR2DL2 or KIR2DL3) will preferentially bind HLA-C1. However, while binding of KIR2DL1 to HLA-C is limited to HLA-C2, KIR2DL2 and KIR2DL3 can bind both HLA-C1 and –C2 epitopes. However, the binding of HLA-C1 to these receptors occurs

with reduced avidity and induces less effective inhibition of NK cell-mediated cytotoxicity than that of HLA-C2 binding (Moesta *et al*., 2008; Hilton *et al*., 2012). While KIR2DS2 is not known to bind HLA-C (Moesta and Parham, 2012), is has recently been shown to bind HLA-A2 (Liu *et al*., 2014).

1.3.6 The Role of KIR-HLA Ligand Interactions in HIV-1 Control

A role for KIR-HLA interactions in determining the course of HIV-1 disease outcome was first described following the observation that the co-occurrence of KIR3DS1 and HLA-Bw4-80I is associated with delayed HIV-1 disease progression (Martin *et al*., 2002a). Despite there being no evidence to support binding of HLA-Bw4-80I to KIR3DS1 (Carr *et al*., 2007; Gillespie *et al*., 2008), there is an expansion of KIR3DS1$^+$ cells during acute HIV-1 infection (Alter *et al*., 2009) and KIR3DS1$^+$ cells have been shown to induce lysis of HIV-1 infected cells (Alter *et al.*, 2007).

KIR3DL1 was similarly shown to be associated with control of HIV-1 infection in the presence of HLA-Bw4-80I (Martin *et al*., 2007). Based on their surface expression, *KIR3DL1* alleles can be classified as either high expressing (*KIR3DL1*h*), low expressing (*KIR3DL1*l*) or non-expressed (*KIR3DL1*004*) (Gardiner *et al*., 2001; Pando *et al.*, 2003; Thomas *et al*., 2008). *KIR3DL*h/KIR3DL*h* and *KIR3DL1*h/KIR3DL1*004* genotypes (collectively referred to as *KIR3DL1*h/*y*) have been shown to be associated with protection against HIV-1 disease progression in the presence of HLA-Bw4-80I, particularly when the HLA-Bw4 molecule in question is HLA-B*57 (Martin *et al*., 2007).

NK cells from HIV-1 slow-progressors with KIR3DL1 in combination with HLA-Bw4 show greater functionality than those from individuals without Bw4 (Kamya *et al*., 2011). This is largely thought to be due to Bw4 licensing of KIR3DL1$^+$ NK cell licensing, with the result that individuals with HLA-Bw4 who express KIR3DL1 on the surface of their NK cells exhibit an

increased capacity for antibody-dependent cellular cytotoxicity (ADCC) relative to individuals without HLA-Bw4 (Parsons *et al*., 2010; Parsons *et al*., 2012). This increase in functional potential is highest in individuals with the *KIR3DL1*h/*y/B*57* genotype (Boulet *et al*., 2010; Kamya *et al*., 2011) and NK cells from these individuals are better able to inhibit HIV-1 replication in autologous $CD4^+$ T cells than those from individuals homozygous for Bw6 or in possession of *KIR3DL1*l* alleles (Song *et al.*, 2014). The *KIR3DL1*h/*y* genotype in combination with HLA-Bw4-80I has also be shown to be associated with strong NK responses in elite controllers with reduced HIV-1 Gag-specific $CD8^+$ T cell responses (Tomescu *et al*., 2012).

Additional associations between KIR receptors recognizing HLA-C and HIV-1 disease course have since been noted (Gaudieri *et al*., 2005; Tiemessen *et al*., 2010; Tiemessen *et al*., 2011; Jennes *et al*., 2011; Merino *et al*., 2014). *KIR* B haplotypes containing *KIR2DL2* have been found to be associated with lower $CD4^+$ T cell counts in both Caucasian Australians (Gaudieri *et al*., 2005) and West Africans (Jennes *et al*., 2011), while expression of functional KIR2DS4 has been seen to promote HIV-1 pathogenesis during chronic infection (Merino *et al*., 2014). In addition, individuals who exhibit NK cell responses to HIV-1 peptides are more likely to possess KIR2DL3 in combination with HLA-C1 (Tiemessen *et al*., 2011) and these NK responses have been shown to be associated with lower viral loads in Black South African women (Tiemessen *et al*., 2010).

1.4 Study Objectives

While patterns of genetic variability within *HLA* class I loci in the Black and Caucasian South African population groups had previously been reported prior to the commencement of this study (Paximadis *et al*., 2012), no corresponding data regarding variability within the genes encoding their cognate KIR receptors were available. Furthermore, while the extent of genetic diversity of *HLA-C* alleles present in these population groups had been established, their

relationship to markers of HLA-C expression had not. Given the demonstrated roles for these molecules in controlling HIV-1 infection in other populations (Martin *et al*., 2007; Kulkarni *et al*., 2011), it was further necessary to assess the association between the observed patterns of genetic variability within these loci and HIV-1 control in the Black South African population – a population group heavily affected by the HIV-1 pandemic. The objectives of this study were therefore:

1. To characterize patterns of genetic variability within the *KIR* gene complex in the Black and Caucasian South African populations
2. To describe patterns of genetic variability at polymorphic positions associated with differential HLA-C expression in the Black and Caucasian South African populations
3. To identify genetic variants within *HLA* class I loci associated with HIV-1 control in the Black South African population
4. To identify genetic variability within the *KIR* gene complex associated with HIV-1 control in the Black South African population

Chapter 2

The Distribution of Killer Cell Immunoglobulin-like Receptor (KIR) Genes and Their HLA Class I Ligands in Two Ethnically Distinct South African Populations

2.1 Introduction

The KIR receptors are a family of glycoprotein receptors involved in the regulation of NK cell function. These receptors are encoded by 14 genes (*KIR2DL1*, *KIR2DL2*, *KIR2DL3*, *KIR2DL4*, *KIR2DL5*, *KIR2DS1*, *KIR2DS2*, *KIR2DS3*, *KIR2DS4*, *KIR2DS5*, *KIR3DL1*, *KIR3DL2*, *KIR3DL3* and *KIR3DS1*) and two pseudogenes (*KIR2DP1* and *KIR3DP1*) located on chromosome 19 (Wilson *et al*., 2000; Martin *et al*., 2000), which exhibit extensive genetic diversity with respect to gene content, copy number and allelic representation within individuals genes (Uhrberg *et al*., 1997; Hou *et al*., 2012; Jiang *et al*., 2012).

Variability in terms of *KIR* gene content is determined by the presence of two main *KIR* haplotypes (Uhrberg *et al*., 2002; Martin *et al*., 2004), which are themselves the product of varying combinations of a limited number of centromeric and telomeric gene content motifs (Jiang *et al*., 2012; Vierra-Green *et al*., 2012; Pyo *et al*., 2010; Pyo *et al*., 2013). The A haplotype is defined by the presence of *KIR2DL1*, *KIR2DL3*, *KIR2DS4*, *KIR3DL1* and *KIR2DP1*, while the more variable B haplotype includes *KIR2DL2*, *KIR2DL5*, *KIR2DS1*, *KIR2DS2*, *KIR2DS3*, *KIR2DS5* and *KIR3DS1* in varying combinations. These two human-specific haplotypes are subject to balancing selection (Parham *et al*., 2012) and are present in all human populations (Yawata *et al*., 2002; Gendzekhadze *et al*., 2006), thereby maintaining the equilibrium between immune defense and reproductive success (Parham *et al*., 2012).

KIR regulate NK cell activity through interaction with their specific HLA class I ligands - the predominant interactions involving the binding of the HLA-C1 epitope to KIR2DL2 and KIR2DL3 and the binding of HLA-C2 to KIR2DL1 (Colonna *et al*., 1993; Winter and Long, 1997; Winter *et al*., 1998) and KIR2DS1 (Moesta *et al*., 2010). In addition, KIR3DL1 is known to bind HLA-Bw4 epitopes (Cella *et al*., 1994; Gumperz *et al*., 1995; Thananchai *et al*., 2007; Stern *et al*., 2008), while HLA-A*03 and A*11 have been shown to bind KIR3DL2 (Döhring *et al*., 1996; Hansasuta *et al*., 2004). HLA-A*11 is also known to interact with both KIR2DS2 (Liu *et al*., 2014) and KIR2DS4 (Graef *et al*., 2009) – which also recognizes a subset of HLA-C alleles (Katz *et al*., 2001; Graef *et al*., 2009).

Genetic diversity at the *KIR* loci has been shown to be associated with disease susceptibility and outcome (Khakoo *et al*., 2004; Dring *et al*., 2011; Hirayasu *et al*., 2012), autoimmunity (Martin *et al*., 2002a; Hou *et al*., 2009), complications during pregnancy (Hiby *et al*., 2004; Hiby *et al*., 2008) and transplantation success (Giebel *et al*., 2003; Hsu *et al*., 2005). Of particular interest in a South African context, is the association seen between KIR-HLA receptor-ligand interactions and HIV-1 transmission (Paximadis *et al*., 2011; Hong *et al*., 2013) and disease outcome (Martin *et al*., 2002a; Martin *et al*., 2007).

However, the distributions of both these gene families are known to vary, often dramatically, between geographically and ethnically distinct populations (Yawata *et al*., 2002; Gendzekhadze *et al*., 2006; Norman *et al*., 2013; Guinan *et al*., 2010), a factor, which if unrecognized, could confound later association studies. Therefore, in order to better understand patterns of genetic diversity within the *KIR* gene complex in the South African population, we examined genetic variability within these genes and the genes encoding their associated *HLA* class I ligands in 167 Black and 97 Caucasian South Africans.

2.2 Materials and Methods

2.2.1 Study Populations

A total of 264 unrelated individuals were selected in order to describe patterns of genetic variability within the *KIR* gene cluster in the South African population. These 167 Black and 97 Caucasian South Africans were selected from a larger previously described cohort (Paximadis *et al*., 2012) on the basis of a non-reactive HIV enzyme-linked immunosorbent assay (ELISA) test (Genscreen HIV1/2 version 2; Bio-Rad, Marnes-La-Coquette, France). The DNA used for genotyping these individuals was extracted from buffy coat samples using the PEL-FREEZ DNA Isolation Kit (DYNAL Invitrogen Corporation, Carlsbad, California, USA). Informed consent was obtained from all study participants and the study was approved by the University of the Witwatersrand Committee for Research on Human Subjects (Appendix B).

2.2.2 KIR Genotyping

Individuals were genotyped for the presence or absence of the 14 *KIR* genes (*KIR2DL1*, *KIR2DL2*, *KIR2DL3*, *KIR2DL4*, *KIR2DL5*, *KIR2DS1*, *KIR2DS2*, *KIR2DS3*, *KIR2DS4*, *KIR2DS5*, *KIR3DL1*, *KIR3DL2*, *KIR3DL3* and *KIR3DS1*) and the 2 pseudogenes (*KIR2DP1* and *KIR3DP1*) using a previously described real-time PCR assay (Hong *et al*., 2011). Additional primers (Jiang *et al*., 2012) were included to distinguish between *KIR2DS4* alleles encoding the full length form of the receptor (*KIR2DS4*0010101-00103* – designated as *KIR2DS4f*) and alleles encoding a deleted form of the gene (*KIR2DS4*003*, **004*, **006*, **007* and **009* – designated as *KIR2DS4v*) - which differ on the basis of the presence or absence of a 22 bp deletion in exon 5.

Reactions were performed in a 5 µl volume, containing 2x Maxima SYBR Green/ROX qPCR Master Mix (Fermentas, Burlington, Canada), 0.2 µM of *KIR*-specific primers, 0.2 µM of

galactosylceramidase-specific primers and 5 ng of DNA. Thermocycling was performed using the Applied Biosystems 7500 Real-Time PCR system (Applied Biosystems, Foster City, California, USA), under the following conditions: 95°C for 10 minutes, followed by 30 cycles of 95°C for 15 seconds and 60°C for 1 minute, with subsequent melt curve analysis.

2.2.3 KIR Genotype Assignment

Individuals were assigned either AA or Bx *KIR* genotypes, according to the guidelines and nomenclature stipulated by the Allele Frequency Net Database (Gonzalez-Galarza *et al*., 2011). AA *KIR* genotypes were defined as consisting of the four framework genes (*KIR2DL4*, *KIR3DL2*, *KIR3DL3* and *KIR3DP1*), in combination with *KIR2DL1*, *KIR2DL3*, *KIR2DS4*, *KIR2DP1* and *KIR3DL1*; while Bx genotypes were defined by the presence of any one or more of the following genes: *KIR2DL5*, *KIR2DS1*, *KIR2DS2*, *KIR2DS3*, *KIR2DS5* and *KIR3DS1*.

2.2.4 KIR Haplotype Estimation

In order to further elucidate the haplotypic composition of the *KIR* genotypes identified, individual *KIR* haplotypes were inferred using HAPLO-IHP (Yoo *et al*., 2007). This program was developed to allow for the inference of haplotypic structure from presence-absence genotype data, specifically in instances where information regarding gene copy number may be limited or entirely unavailable. This software also allows for the optional incorporation of information regarding previously identified haplotypes and haplotype patterns. As all known *KIR* haplotypes are derived from varying combinations of a limited number of centromeric and telomeric haplotype motifs (Pyo *et al*., 2010; Jiang *et al*., 2012; Vierra-Green *et al*., 2012; Nakimuli *et al*., 2013; Norman *et al*., 2013; Pyo *et al*., 2013), an *a priori* list of haplotypes consisting of all possible combinations of previously identified *KIR* haplotype motifs were included in these analyses.

The availability of gene copy number information for *KIR2DS4*, *KIR3DL1* and *KIR3DS1* allowed haplotypes containing the tA01 telomeric haplotype motif to be resolved with a fair degree of certainty. However, because *KIR2DL5*, *KIR2DS3* and *KIR2DS5* can occur on both centromeric and telomeric *KIR* haplotype motifs, in the absence of copy number information for these genes it was impossible to distinguish between the cA01|tB01 and cB03|tB01 haplotypes using our haplotype estimation method. These two haplotypes are identical in gene content and differ only with respect to the location of the three aforementioned genes, with *KIR2DL5*, *KIR2DS3* and *KIR2DS5* located within the telomeric region of the cA01|tB01 haplotype and found within the centromeric portion of cB03|tB01 (Pyo *et al*., 2010; Pyo *et al*., 2013). The frequencies of these haplotypes are thus reported as cA01/cB03|tB01.

2.2.5 HLA Class I Ligand Determination

HLA-A, *-B* and *-C* genotype data were available for all 264 individuals (Paximadis *et al*., 2012; Chapter 3). HLA-A and -B Bw4 isotopes within alleles were identified based on the amino acid residues present at positions 77, 80, 81, 82 and 83 of the protein sequence (Martin *et al*., 2002a). A further distinction was made between *HLA-B* alleles with isoleucine at amino acid position 80 (*Bw4-80I*) and those with threonine at this position (*Bw4-80T*), as these epitopes differ in the avidity of their interaction with KIR3DL1 (Cella *et al*., 1994; Martin *et al*., 2002a). *HLA-C* alleles were similarly classified as having either HLA-C1 or –C2 epitopes on the basis of the presence of asparagine or lysine at position 80, respectively (Colonna *et al*., 1993; Winter and Long, 1997).

2.2.6 Statistical Analyses

KIR gene carrier- and *HLA* class I ligand frequencies were determined by direct counting. *KIR* gene carrier frequencies observed within the two South African population groups were

compared to those previously reported for other African and Caucasian populations (Gonzalez-Galarza *et al*., 2011) by complete linkage hierarchical clustering, as implemented in the hclust R package (www.r-project.org). The significance of differences in the frequencies observed between the Black and Caucasian population groups was assessed using a two-sided Fisher's exact test, as implemented in R v3.1.1 (www.r-project.org). Pairwise LD between individual *KIR* loci was quantified in both population groups using the measures D' (Lewontin, 1964) and r^2 (Hill and Robertson, 1968) and the significance of LD between loci was assessed by an exact test for LD (Slatkin, 1994). LD analyses were performed using Arlequin v3.5.1.2 (Excoffier and Lisher, 2010). All measures were considered significant at $p < 0.05$.

2.3 Results

2.3.1 KIR Gene Diversity

We examined *KIR* gene diversity in 167 Black and 97 Caucasian South African individuals and found the four framework *KIR* genes to be present in all 264 individuals genotyped. *KIR2DL1*, *KIR2DS4*, *KIR3DL1* and the pseudogene, *KIR2DP1*, were also present in more than 98 percent of individuals in both population groups (Table 2.1). However, when a distinction was made between alleles encoding the full-length (*KIR2DS4f*) and truncated (*KIR2DS4v*) forms of *KIR2DS4*, the two population groups exhibited significantly different distributions of these two isoforms; with *KIR2DS4f* being the more frequent isoform present in the Black population group ($p = 1.23 \times 10^{-13}$) and *KIR2DS4v* being the most frequent in the Caucasian population group ($p = 8.79 \times 10^{-9}$). The two groups also displayed significantly different distributions of *KIR2DS1* ($p = 1.20 \times 10^{-6}$) and *KIR3DS1* ($p = 5.43 \times 10^{-10}$). While there was a tendency for *KIR2DL2* to be more commonly observed in the Black population group, and for *KIR2DL3* to occur with higher frequency in the Caucasian population group, these differences did not reach statistical significance (Table 2.1).

Comparison of the *KIR* gene carrier frequencies observed in the two South African population groups with those from other African and Caucasian populations (Figure 2.1) revealed the South African Caucasian population clustered most closely with other Caucasian populations, most notably those from England and the USA. Similarly, the Black South African population group clustered with other African populations from Uganda, Kenya and Gabon. Differences in gene frequency between African and Caucasian population groups were again most clearly apparent for *KIR2DS1* and *KIR3DS1*, as already seen in the South African population comparison (Table 2.1) and consistent with previous reports (Norman *et al*., 2013; Single *et al*., 2007; Nakimuli *et al*., 2013).

Table 2.1: The gene carrier frequencies of the *KIR* genes present in the Black and Caucasian South African population groups

***KIR* Gene**[1]	**Black Individuals (n = 167)**		**Caucasian Individuals (n = 97)**		**p-value**[2]
	n	**Frequency (%)**	**n**	**Frequency (%)**	
KIR2DL1	166	99	96	99	-
KIR2DL2	114	68	56	58	-
KIR2DL3	137	82	86	89	-
KIR2DL5	102	61	49	51	-
KIR2DS1	21	13	38	39	<0.001
KIR2DS2	100	60	55	57	-
KIR2DS3	44	26	23	24	-
KIR2DS4f	139	84	37	38	<0.001
KIR2DS4v	83	50	82	85	<0.001
KIR2DS5	72	43	38	39	-
KIR2DP1	166	99	96	99	-
KIR3DL1	166	99	95	98	-
KIR3DS1	12	7	38	39	<0.001

[1] The four framework genes (*KIR2DL4*, *KIR3DL2*, *KIR3DL3* and *KIR3DP1*) were present in all individuals in both population groups and their frequencies are thus not reported here.

[2] p-value is for a two-sided Fisher's exact test. Only significant values ($p < 0.05$) are indicated.

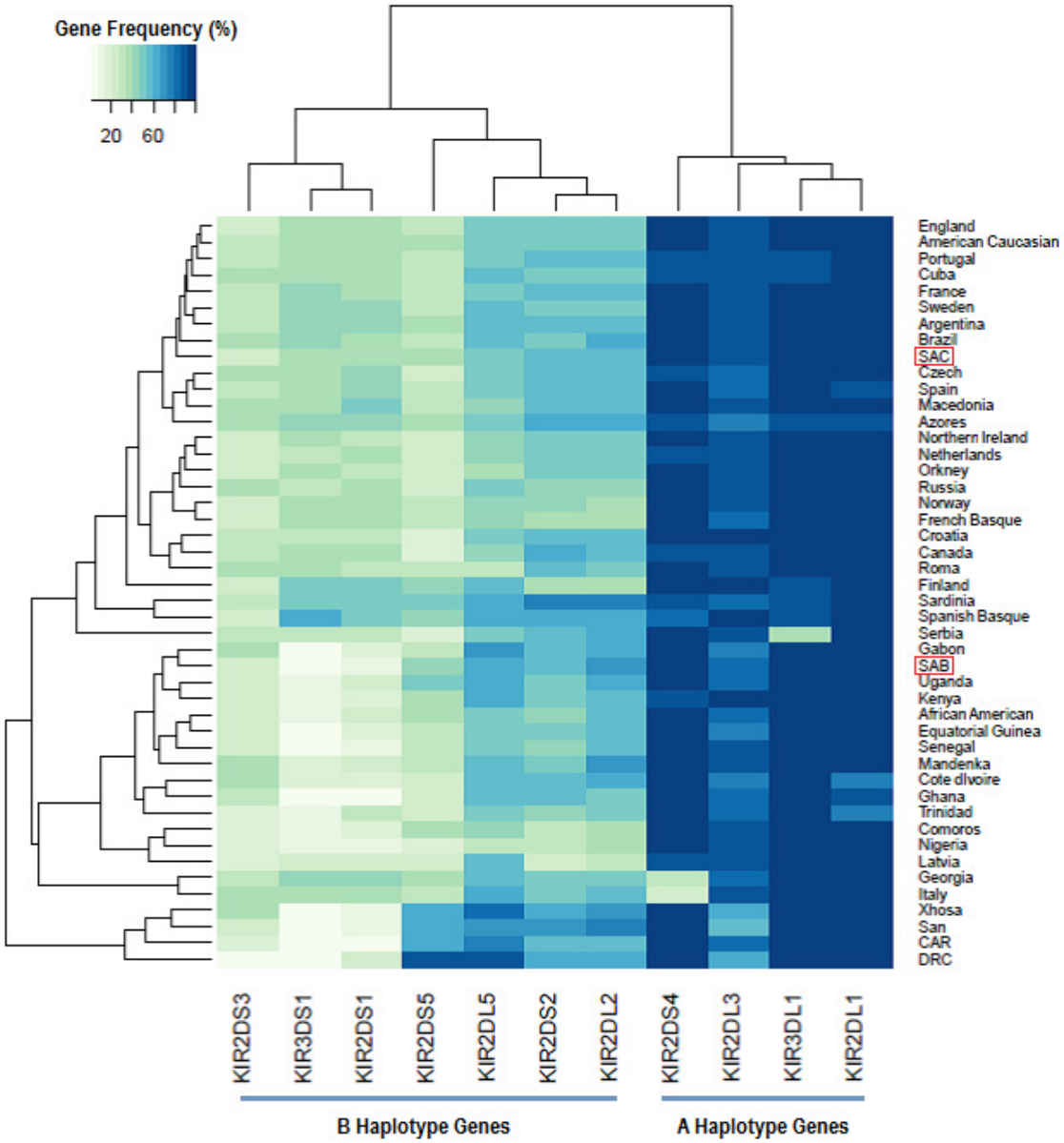


Figure 2.1: A two-dimensional heatmap based on hierarchical cluster of the carrier frequencies of the non-framework *KIR* genes in 17 African and 29 Caucasian populations. Population groups are clustered along the y-axis, with the South African Black (SAB) and Caucasian (SAC) groups indicated in red. Frequencies are indicated by the color scale, with higher frequencies indicated by increased color intensity. *KIR* loci are clustered along the x-axis, where A and B haplotype genes are indicated. Abbreviations: CAR = Central African Republic, DRC = Democratic Republic of the Congo.

2.3.2 KIR Genotypes Distributions

Analysis of the *KIR* genotypes present in the South African population revealed the presence of 33 *KIR* genotypes in the Black population group (Table 2.2). Thirteen of these were found to be present in more than one individual and collectively, these accounted for more than 88% of the genotypes observed in this population group. A total of 24 genotypes were observed in the Caucasian population group (Table 2.2). One of these genotypes could not be assigned an AA or Bx identifier according to the current nomenclature, as it had not previously been reported in other populations (Gonzalez-Galarza *et al*., 2011). Consequently, it was referred to as unknown (Table 2.2). However, based on its gene content, it resembled the Bx genotypes, particularly Bx 58. The AA 1 genotype was the most prevalent in both population groups and was also present at similar frequencies in both groups. The Bx 2 ($p = 9.33 \times 10^{-5}$) and Bx 3 ($p = 2.50 \times 10^{-4}$) genotypes were significantly more frequent in the Caucasian population group, while Bx 5 ($p = 0.035$), Bx 20 ($p = 0.029$), Bx 21 ($p = 0.001$) and Bx 228 ($p = 0.035$) were observed at significantly higher frequencies in the Black population group.

2.3.3 KIR Haplotype Estimation

Estimation of the *KIR* gene content haplotypes present in the South African population identified a total of 24 haplotypes across both population groups (Table 2.3). Of these, five were unique to the Black population group, while three were only observed in Caucasian individuals. The cA01|tA01 haplotype was the most common haplotype observed in both population groups and was present at similar frequencies in both groups. The 49% frequency seen in the South African Caucasian population was slightly lower than that reported for other populations of European descent (Pyo *et al*., 2010; Jiang *et al*., 2012; Vierra-Green *et al*., 2012), while the frequency in the Black population group was also lower than that reported in Ghanaians (48% versus 53%; Norman *et al*., 2013).

Table 2.2 The *KIR* genotype profiles identified within the South African population and their frequencies in the Black and Caucasian population groups

KIR Genotype[1]	KIR2DL4	KIR3DL2	KIR3DL3	KIR3DP1	KIR2DP1	KIR2DL1	KIR2DL3	KIR2DS4	KIR3DL1	KIR2DL2	KIR2DL5	KIR2DS1	KIR2DS2	KIR2DS3	KIR2DS5	KIR3DS1	Frequency Black Individuals (n = 167)	Frequency Caucasian Individuals (n = 97)
AA 1	■	■	■	■	■	■	■	■	■								0.269 (45)[2]	0.278 (27)
Bx 21	■	■	■	■	■	■	■	■	■	■	■		■		■		0.114 (19)	0.010 (1)
Bx 4	■	■	■	■	■	■	■	■	■	■			■				0.090 (15)	0.155 (15)
Bx 5	■	■	■	■	■	■	■	■	■	■	■		■	■			0.090 (15)	0.021 (2)
Bx 228	■	■	■	■	■	■		■	■	■	■		■		■		0.072 (12)	0.010 (1)
Bx 20	■	■	■	■	■	■	■	■	■	■	■				■		0.054 (9)	-
Bx 71	■	■	■	■	■	■		■	■	■	■		■	■			0.042 (7)	0.041 (4)
Bx 9	■	■	■	■	■	■	■	■	■	■	■	■	■		■		0.036 (6)	-
Bx 6	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	0.030 (5)	0.062 (6)
Bx 112	■	■	■	■	■	■		■	■	■	■		■	■	■		0.030 (5)	-
Bx 32	■	■	■	■	■	■	■	■	■		■				■		0.024 (4)	-
Bx 51	■	■	■	■	■	■	■	■	■	■	■			■			0.018 (3)	-
Bx 31	■	■	■	■	■	■	■	■	■	■	■		■				0.012 (2)	-
Bx 2	■	■	■	■	■	■	■	■	■		■	■			■	■	0.006 (1)	0.113 (11)
Bx 3	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	0.006 (1)	0.103 (10)
Bx 7	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	0.006 (1)	0.021 (2)
Bx 11	■	■	■	■	■	■	■	■	■	■	■	■	■	■			0.006 (1)	-
Bx 19	■	■	■	■	■	■	■	■	■	■							0.006 (1)	-
Bx 25	■	■	■	■	■	■	■	■	■	■	■		■	■	■		0.006 (1)	-
Bx 30	■	■	■	■	■	■	■	■	■		■			■			0.006 (1)	-
Bx 44	■	■	■	■	■	■	■	■	■	■			■		■		0.006 (1)	-
Bx 48	■	■	■	■	■	■	■	■	■		■		■		■		0.006 (1)	0.010 (1)
Bx 62	■	■	■	■	■	■	■	■	■	■			■	■			0.006 (1)	-
Bx 64	■	■	■	■	■	■	■	■	■	■	■	■		■		■	0.006 (1)	0.010 (1)
Bx 73	■	■	■	■	■	■		■	■	■	■	■	■	■	■	■	0.006 (1)	0.010 (1)
Bx 81	■	■	■	■	■	■				■	■	■	■	■	■	■	0.006 (1)	-
Bx 89	■	■	■	■	■	■		■	■	■			■				0.006 (1)	0.031 (3)
Bx 92	■	■	■	■	■	■		■	■	■	■	■	■		■		0.006 (1)	0.010 (1)
Bx 169	■	■	■	■	■		■		■	■	■	■	■		■	■	0.006 (1)	-
Bx 172	■	■	■	■		■		■	■	■			■				0.006 (1)	0.010 (1)
Bx 175	■	■	■	■	■	■		■	■	■	■				■		0.006 (1)	-
Bx 270	■	■	■	■	■	■	■	■	■	■	■			■	■		0.006 (1)	-
Bx 393	■	■	■	■	■	■	■	■	■		■	■	■		■		0.006 (1)	-
Bx 13	■	■	■	■	■	■	■	■	■	■	■		■	■		■	-	0.021 (2)
Bx 70	■	■	■	■	■	■	■			■	■	■	■	■	■	■	-	0.021 (2)
Bx 10	■	■	■	■	■	■	■	■	■				■				-	0.010 (1)
Bx 15	■	■	■	■	■	■	■	■	■			■					-	0.010 (1)
Bx 24	■	■	■	■	■	■	■	■	■	■	■		■		■	■	-	0.010 (1)
Bx 58	■	■	■	■	■	■	■	■	■	■	■	■		■	■	■	-	0.021 (1)
Bx 382	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		-	0.010 (1)
Unknown	■	■	■	■	■		■	■	■	■	■	■		■	■	■	-	0.010 (1)

[1] *KIR* genotypes are ordered based on their frequency within the Black South African population group.

[2] The number of individuals in possession of each genotype, in each population group, is given in brackets.

The two population groups displayed significant differences in frequency for six other haplotypes (Table 2.3), including cB01|tA01 and four other haplotypes derived from it, following gene deletions or insertions (Pyo *et al*., 2013). While it was not possible to distinguish between the cA01|tB01 and cB03|tB01 haplotypes using our haplotype estimation method, based on the frequencies of these haplotype motifs observed in other populations (Pyo *et al*., 2010; Jiang *et al*., 2012; Vierra-Green *et al*., 2012; Norman *et al*., 2013; Pyo *et al*., 2013), we would anticipate the majority of haplotypes with these gene content configurations to be cA01|tB01. Nonetheless, frequencies for these haplotypes were reported as cA01/cB03|tB01 (Table 2.3). These haplotypes were found at significantly greater frequency in the Caucasian population group.

While the centromeric cA01 motif was found to occur at similar frequencies in the Black (53%) and Caucasian (51%) population groups, at frequencies comparable with the centromeric B motif (cB01, cB02 and cB03), the telomeric motifs tA01 and tB01 displayed significantly different frequencies within the two groups (Table 2.4). While tA01 was clearly the dominant telomeric motif present in both groups, an increased frequency of tB01 in the Caucasian population distinguished the two groups. In addition, significantly more Caucasian individuals were found to possess tA01 motifs bearing non-functional *KIR2DS4* alleles (tA01v; Table 2.4). tA01 motifs lacking any form of KIR2DS4 were more prevalent in the Black population group (tA01-hybd1; Table 2.4).

Table 2.3 The frequencies of the *KIR* haplotypes occurring in the Black and Caucasian South African population groups

***KIR* Haplotype**	**Black Individuals (n = 167)**		**Caucasian Individuals (n = 97)**		**p-value[1]**
	n	**Frequency**	**n**	**Frequency**	
cA01|tA01[2]	159	0.476	95	0.490	-
cA01|tA01-hybd1	9	0.027	1	0.005	-
cA01|tA01-del5	8	0.024	2	0.010	-
cA01/cB03|tB01[3]	2	0.006	25	0.129	< 0.001
cB01|tA01	57	0.171	10	0.052	< 0.001
cB01|tA01-hybd1	14	0.042	0	0.000	0.003
cB01|tA01-ins4	0	0.000	5	0.026	0.006
cB01|tA01-del3	14	0.042	0	0.000	0.003
cB01|tA01-del3-ins4	0	0.000	1	0.005	-
cB01|tA01-del3-del5	1	0.003	0	0.000	-
cB01|tA01-del3-del5-del9	1	0.003	0	0.000	-
cB01|tA01-del9	32	0.096	33	0.170	0.012
cB01|tA01-del4-del9	2	0.006	5	0.026	-
cB01|tA01-del10	2	0.006	0	0.000	-
cB01|tB01	9	0.027	11	0.057	-
cB01|tB01-del5	1	0.003	0	0.000	-
cB01|tB01-del7	8	0.024	1	0.005	-
cB02|tA01-ins5	1	0.003	1	0.005	-
cB02|tA01-hybd1	3	0.009	1	0.005	-
cB03|tA01	5	0.015	1	0.005	-
cB03|tA01-hybd1	3	0.009	0	0.000	-
H1[4]	2	0.006	0	0.000	-
H2	1	0.003	1	0.005	-
H3	0	0.000	1	0.005	-

[1] p-value is for a two-sided Fisher's exact test. Only significant values ($p < 0.05$) are indicated.

[2] Major haplotypes comprised of the classic centromeric and telomeric *KIR* motifs are indicated in grey. Minor haplotypes derived from these following gene insertions, deletions and duplications are named in accordance with the nomenclature of Pyo *et al*. (2013).

[3] cA01\|tB01 and cB03\|tB01 could not be distinguished from each other and are reported as cA01/cB03\|tB01.

[4] Haplotypes that could not be named according to the aforementioned nomenclature were assigned arbitrary designations based on their frequencies within the Black population group.

Table 2.4 The telomeric *KIR* haplotype motifs present in the South African population and their frequencies within the Black and Caucasian population groups

Telomeric Motif[1]	Black Individuals (n = 167)		Caucasian Individuals (n = 97)		p-value[2]
	n	Frequency	n	Frequency	
tA01	314	0.940	156	0.813	< 0.001
tA01f	193	0.619	45	0.288	< 0.001
tA01v	89	0.285	108	0.692	< 0.001
tA01-hybd1	30	0.096	3	0.019	0.002
tB01	20	0.060	38	0.198	< 0.001

[1] tA01f, tA01v and tA01-hybrd1 refer to tA01 motifs carrying *KIR2DS4f* alleles, *KIR2DS4v* alleles and no *KIR2DS4*, respectively.

[2] p-value is for a two-sided Fisher's exact test. Only significant values ($p < 0.05$) are indicated.

2.3.4 Estimation of Linkage Disequilibrium

In order to better understand the organization of the *KIR* haplotypes present in the South African population, pairwise LD between individual *KIR* loci was quantified in both population groups using the measures D' (Lewontin, 1964) and r^2 (Hill and Robertson, 1968). As was expected, strong positive associations were detected between genes known to be present on the same haplotype motifs (Tables 2.5 and 2.6). In the Black population group, strong positive associations were observed between *KIR2DL1*, *KIR2DL3*, *KIR2DS4* and *KIR3DL1*, as well as between *KIR2DL2*, *KIR2DS2*, *KIR2DS1* and *KIR3DS1* (Table 2.5). In the Caucasian population group a clear positive association was seen between *KIR2DL2*, *KIR2DS2* and *KIR2DS3*, as well as between *KIR2DS1*, *KIR2DS5* and *KIR3DS1* (Table 2.6). These observations were consistent with previous reports in other Black and Caucasian population groups (Single *et al*., 2007; Hollenbach *et al*., 2012; Nakimuli *et al*., 2013).

Table 2.5 The estimated linkage disequilibrium across the *KIR* gene complex in the Black South African population

		KIR2DL1	*KIR2DL2*	*KIR2DL3*	*KIR2DL5*	*KIR2DS1*	*KIR2DS2*	*KIR2DS3*	*KIR2DS4*	*KIR2DS5*	*KIR2DP1*	*KIR3DL1*
	D'	-0.650										
KIR2DL2	**r^2**	0.027										
	p-value[1]	0.006										
	D'	0.647	-1.000									
KIR2DL3	**r^2**	0.026	0.976									
	p-value	0.004	< 0.001									
	D'	-0.691	0.163	-0.805								
KIR2DL5	**r^2**	0.041	0.807	0.462								
	p-value	< 0.001	< 0.001	< 0.001								
	D'	-0.787	0.750	-0.748	1.000							
KIR2DS1	**r^2**	0.434	0.050	0.049	0.123							
	p-value	< 0.001	< 0.001	< 0.001	< 0.001							
	D'	-0.674	0.959	-0.986	0.613	0.767						
KIR2DS2	**r^2**	0.034	0.774	0.798	0.327	0.063						
	p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001						
	D'	0.482	0.919	-0.918	0.964	0.344	0.811					
KIR2DS3	**r^2**	0.002	0.167	0.162	0.251	0.054	0.154					
	p-value	NS	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001					
	D'	1.000	-0.475	0.470	-0.629	-1.000	-0.511	-0.070				
KIR2DS4	**r^2**	0.267	0.053	0.051	0.128	0.381	0.073	0.004				
	p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	NS				
	D'	-0.647	0.808	-0.806	0.981	0.874	0.583	-0.337	-0.629			
KIR2DS5	**r^2**	0.060	0.284	0.276	0.573	0.157	0.176	0.006	0.214			
	p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	NS	< 0.001			
	D'	0.693	-1.000	1.000	0.485	-0.287	-0.963	1.000	0.118	0.259		
KIR2DP1	**r^2**	0.149	0.203	0.198	0.020	0.036	0.224	0.022	0.012	0.003		
	p-value	< 0.001	< 0.001	< 0.001	0.003	< 0.001	< 0.001	< 0.001	0.040	NS		
	D'	0.709	-0.376	0.369	-0.558	-0.896	-0.302	-0.180	0.622	-0.574	0.177	
KIR3DL1	**r^2**	0.258	0.017	0.016	0.052	0.589	0.013	0.020	0.201	0.093	0.019	
	p-value	< 0.001	0.023	0.023	< 0.001	< 0.001	0.017	0.003	< 0.001	< 0.001	0.019	
	D'	-0.239	0.682	-0.679	1.000	1.000	0.556	0.687	-1.000	0.880	1.000	-1.000
KIR3DS1	**r^2**	0.041	0.021	0.020	0.062	0.508	0.017	0.109	0.193	0.081	0.005	0.372
	p-value	< 0.001	0.003	0.004	< 0.001	< 0.001	0.006	< 0.001	< 0.001	< 0.001	NS	< 0.001

[1] p-value is for an exact test of linkage disequilibrium (Slatkin, 1994).

Table 2.6 The estimated linkage disequilibrium across the *KIR* gene complex in the Caucasian South African population

		KIR2DL1	*KIR2DL2*	*KIR2DL3*	*KIR2DL5*	*KIR2DS1*	*KIR2DS2*	*KIR2DS3*	*KIR2DS4*	*KIR2DS5*	*KIR2DP1*	*KIR3DL1*
	D'	0.596										
KIR2DL2	**r^2**	0.032										
	p-value[1]	0.011										
	D'	-0.625	-1.000									
KIR2DL3	**r^2**	0.040	0.892									
	p-value	0.001	< 0.001									
	D'	-0.953	0.293	-0.265								
KIR2DL5	**r^2**	0.399	0.073	0.053								
	p-value	< 0.001	< 0.001	0.001								
	D'	-0.916	-0.015	0.086	0.965							
KIR2DS1	**r^2**	0.591	0.000	0.001	0.581							
	p-value	< 0.001	NS	NS	< 0.001							
	D'	0.710	0.928	-0.977	0.223	-0.059						
KIR2DS2	**r^2**	0.049	0.804	0.912	0.039	0.001						
	p-value	< 0.001	< 0.001	< 0.001	0.007	NS						
	D'	0.438	1.000	-1.000	1.000	0.398	0.801					
KIR2DS3	**r^2**	0.005	0.273	0.244	0.323	0.082	0.164					
	p-value	NS	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001					
	D'	0.958	-0.005	-0.063	-0.899	-1.000	0.035	-0.390				
KIR2DS4	**r^2**	0.607	0.000	0.001	0.537	0.940	0.000	0.074				
	p-value	< 0.001	NS	NS	< 0.001	< 0.001	NS	< 0.001				
	D'	-0.914	0.084	-0.048	1.000	0.871	0.063	0.269	-0.815			
KIR2DS5	**r^2**	0.521	0.004	0.001	0.705	0.671	0.002	0.033	0.626			
	p-value	< 0.001	NS	NS	< 0.001	< 0.001	NS	0.002	< 0.001			
	D'	-0.631	-1.000	1.000	0.903	0.861	-1.000	1.000	-0.736	0.874		
KIR2DP1	**r^2**	0.016	0.447	0.399	0.075	0.042	0.417	0.030	0.033	0.049		
	p-value	NS	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.010	0.006	< 0.001		
	D'	0.872	0.062	0.002	-0.865	-0.936	0.035	-0.390	0.878	-0.784	-0.868	
KIR3DL1	**r^2**	0.504	0.001	0.000	0.497	0.824	0.000	0.074	0.772	0.580	0.046	
	p-value	< 0.001	NS	NS	< 0.001	< 0.001	NS	< 0.001	< 0.001	< 0.001	< 0.001	
	D'	-0.871	0.152	-0.118	1.000	0.903	0.028	0.606	-0.846	0.824	1.000	-0.846
KIR3DS1	**r^2**	0.473	0.014	0.007	0.705	0.722	0.000	0.169	0.674	0.678	0.065	0.674
	p-value	< 0.001	NS	NS	< 0.001	< 0.001	NS	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

[1] p-value is for an exact test of linkage disequilibrium (Slatkin, 1994).

2.3.5 Association between Specific KIR and Their HLA Class I Ligands

2.3.5.1 HLA-B and KIR3DL1/DS1

The HLA-Bw4 and Bw6 epitopes are known to occur at comparable frequencies within the Black and Caucasian South African population groups (Paximadis *et al*., 2012). However, the frequencies with which they appear in combination with KIR3DL1, the NK cell receptor for HLA-Bw4, have yet to be reported for these groups. We therefore examined the proportion of individuals in the possession of KIR3DL1 in combination with both HLA-Bw4 and Bw6 in each of these population groups and found them to be similar (Table 2.7). However, Black individuals with alleles for both *KIR3DL1* and *HLA-Bw4* were more likely to carry *HLA-Bw4* alleles with isoleucine at position 80 (*Bw4-80I*), while Caucasian individuals more frequently carried alleles for *HLA-Bw4* epitopes with threonine at position 80 (*Bw4-80T*; Table 2.7).

Although KIR3DS1 - with the exception of the receptor encoded by one allele (O'Connor *et al*., 2011) - has not been shown to bind HLA-Bw4 (Carr *et al*., 2007; Gillespie *et al*., 2007; O'Connor *et al*., 2007), co-occurrence of the alleles encoding these molecules has previously been associated with differential disease prognosis (Martin *et al*., 2002b; López-Vázquez *et al*., 2005). We therefore also investigated the frequency with which individuals in the South African population possessed alleles encoding *KIR3DS1* in combination with either the HLA-Bw4 or Bw6 epitopes (Figure 2.2). The small number of Black individuals in possession of *KIR3DS1* was most often found in possession of *HLA-Bw4* alleles, while the corresponding Caucasian individuals were more likely to possess *HLA-Bw6* alleles – although this observation did not reach statistical significance. In the Black population groups, as seen with *KIR3DL1*, *KIR3DS1* was more likely to be found in combination with alleles with HLA-Bw4-80I motifs.

Table 2.7 The frequency with which *KIR3DL1* occurs in combination with the HLA Bw4 and Bw6 allotypes in the Black and Caucasian South African population groups

	Black Individuals (n = 166)		Caucasian Individuals (n = 95)		p-value[1]
	n	Frequency[2]	n	Frequency	
***KIR3DL1* + Bw4**	118	0.711	67	0.705	-
***KIR3DL1* + Bw4-80I**	85	0.512	29	0.305	0.001
***KIR3DL1* + Bw4-80T**	22	0.133	35	0.368	< 0.001
***KIR3DL1* + Bw6**	141	0.849	78	0.821	-

[1] p-value is for a two-sided Fisher's exact test. Only significant values ($p < 0.05$) are indicated.

[2] Frequencies are given relative to the number of individuals in each population group in possession of *KIR3DL1*.

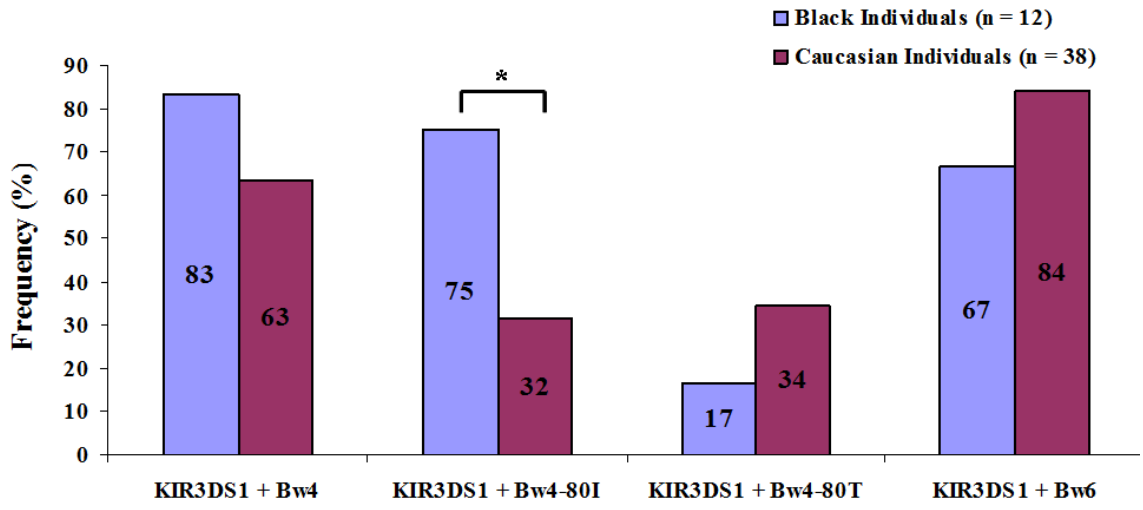
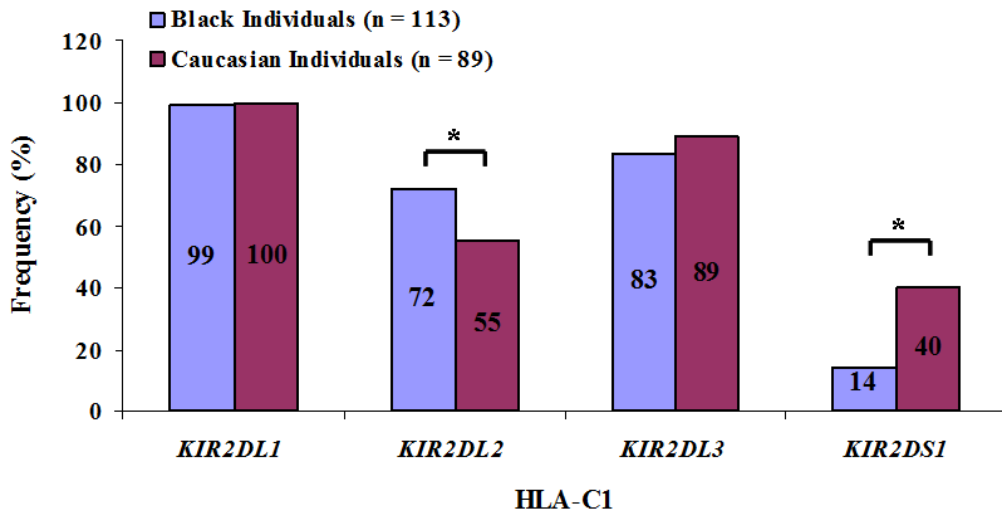


Figure 2.2: The frequency with which HLA-Bw4 and -Bw6 allotypes occur in individuals from the Black and Caucasian South African population groups in possession of *KIR3DS1*. Frequencies are given relative to the number of individuals in each population group in possession of at least one copy of *KIR3DS1*. Significantly differences in frequency ($p < 0.05$) between the two groups, as measured by a two-sided Fisher's exact test, are indicated with an asterisk.

(a)



(b)

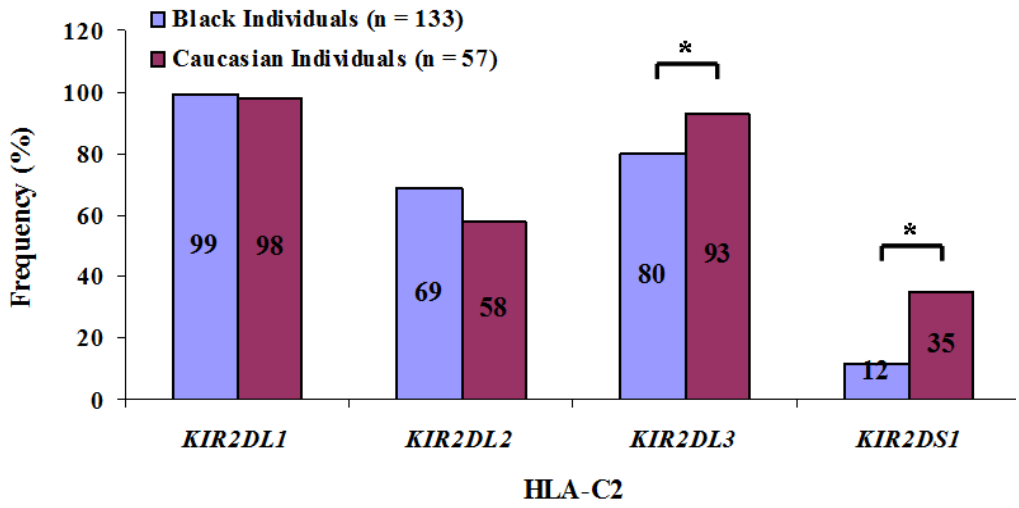


Figure 2.3: **(a)** The frequency with which *KIR2DL1*, *KIR2DL2*, *KIR2DL3* and *KIR2DS1* occur in combination with the HLA-C1 in individuals from the Black and Caucasian South African population groups. **(b)** The frequency with which *KIR2DL1*, *KIR2DL2*, *KIR2DL3* and *KIR2DS1* occur in combination with the HLA-C2 in individuals from the Black and Caucasian South African population groups. In both (a) and (b), frequencies are given relative to the number of individuals in each population group in possession of the specific HLA-C epitope of interest. Significant differences in frequency ($p < 0.05$) between the two groups, as measured by a two-sided Fisher's exact test, are indicated with an asterisk.

2.3.6 HLA-C and its Associated KIR Receptors

The HLA-C1 and -C2 epitopes present in virtually all *HLA-C* alleles serve as ligands for KIR2DL2, KIR2DL3, KIR2DL1 and KIR2DS1 (Colonna *et al*., 1993; Winter and Long, 1997; Winter *et al*., 1998; Moesta *et al*., 2010), and occur at significantly different frequencies in the Black and Caucasian population groups (Paximadis *et al*., 2012). We therefore also examined the proportion of individuals in each population group who possessed the genes encoding these KIR receptors in the presence or absence of their respective HLA-C ligands (Figure 2.3). In terms of individuals found to be in possession of *HLA-C1* alleles, a significantly higher percentage of individuals in the Black population group were also seen to possess *KIR2DL2* (Figure 2.3a). Conversely, Caucasian individuals in possession of *HLA-C2* alleles were significantly more likely to also be in possession of *KIR2DL3* (Figure 2.3b). *KIR2DS1* was also seen to be significantly more common in Caucasian individuals with both *HLA-C1* (Figure 2.3a) and *–C2* (Figure 2.3b) than in their counterpart in the Black population group.

2.4 Discussion

We examined genetic diversity within the *KIR* gene complex in 167 Black and 97 Caucasian South African individuals and found these two population groups differed significantly with respect to their *KIR2DS1* and *KIR3DS1* gene frequencies in a manner consistent with previous reports describing *KIR* gene diversity in other African and Caucasian populations (Norman *et al*., 2007; Single *et al*., 2007; Nakimuli *et al*., 2013; Petitdemange *et al*., 2014). Hierarchical clustering analysis of gene frequencies revealed the patterns of *KIR* gene diversity observed in the Caucasian South African population closely resembled that of other Caucasian populations; most notably those from England and the USA. The Black South African population group, on the other hand, was found to most closely resemble the populations from Gabon and Uganda, as well as the population of predominantly Bantu-speakers from Kenya. The genetic similarity

observed between these populations is consistent with the proposed migration of West African pastoralists into southern Africa, through East Africa (Tishkoff and Williams, 2002; Henn *et al*., 2008; Tishkoff *et al*., 2009).

The differences in gene frequency observed between the Black and Caucasian population groups were reflected by the differences in *KIR* genotype frequencies observed between the two population groups, where genotypes including *KIR2DS1* and *KIR3DS1* were more commonly observed in the Caucasian population group and genotypes lacking these genes were most frequently observed in Black individuals. The commonalities in frequency distributions seen between these two *KIR* genes could be attributed to the strong pairwise LD present between these loci in both population groups, consistent with their co-localization on the telomeric tB01 *KIR* haplotype motif (Pyo *et al*., 2010; Jiang *et al*., 2012; Vierra-Green *et al*., 2012; Pyo *et al*., 2013). This motif has been found to be present at substantially lower frequencies than the tA01 motif in all populations examined to date, however is reported to be present at higher frequencies in Caucasian populations than in African population groups (Pyo *et al*., 2010; Jiang *et al*., 2012; Vierra-Green *et al*., 2012; Norman *et al*., 2013; Pyo *et al*., 2013) - an observation that is consistent with our findings in the South African population.

KIR2DS1 and *KIR3DS1* were also seen to be in strong LD with *KIR2DS5* in both population groups. This gene, along with *KIR2DL5* and *KIR2DS3*, may be present on either (or both) the centromeric and telomeric regions of *KIR* B haplotypes (Pyo *et al*., 2010; Jiang *et al*., 2012; Vierra-Green *et al*., 2012; Norman *et al*., 2013; Pyo *et al*., 2013). Given the strong association seen between *KIR2DS5*, *KIR2DS1* and *KIR3DS1* in the Caucasian population group, coupled with correspondingly low levels of observed LD between *KIR2DS5* and the centromeric genes, *KIR2DL2*, *KIR2DL3* and *KIR2DS2*, it seems likely *KIR2DS5* occurs predominantly on the telomeric region of B haplotypes, as part of the tB01 motif, in this population group. In contrast, strong LD between *KIR2DS3* and *KIR2DL2* and *KIR2DS2* in this population group would

suggest this gene occurs predominantly within centromeric motifs. This is consistent with previous findings in other Caucasian population groups (Single *et al*., 2008; Jiang *et al*., 2012).

High levels of LD between *KIR2DS3* and *KIR2DL2*, *KIR2DL3* and *KIR2DS2* in the Black population group, in conjunction with the observation of reduced levels of LD with *KIR2DS1*, suggest this gene also likely occurs predominantly within centromeric haplotype motifs in this population. Conversely, *KIR2DS5* exhibited comparable level of LD with both centromeric (*KIR2DL2* and *KIR2DL3*) and telomeric (*KIR2DS1* and *KIR3DS1*) genes in the Black population group. This could suggest that this gene occurs within both the centromeric and telomeric motifs present in the Black population group at comparable frequencies, but this would be inconsistent with previous observations in other African populations, where *KIR2DS5* was rarely seen to occur within the telomeric motif (Nakimuli *et al*., 2013; Norman *et al*., 2013). Rather, this observation most likely reflects the strong LD observed between the centromeric and telomeric B motifs in this population group. This positional discrepancy could potentially be resolved with additional information regarding *KIR2DS5* copy number.

The Black and Caucasian South African population groups also differed with respect to variability in their tA01 telomeric haplotype motifs. While the majority of Black individuals possessed tA01 motifs containing functional, full-length *KIR2DS4* alleles, most individuals in the Caucasian population group carried tA01 motifs with *KIR2DS4v* alleles. *KIR2DS4v* alleles encode a truncated receptor lacking a transmembrane domain that is not expressed at the cell surface (Maxwell *et al*., 2002). An additional 2% of Caucasians and nearly 10% of Black individuals were also found to possess tA01 motifs from which *KIR2DS4* had been entirely deleted, with the result that nearly 40% of Black individuals and more than 70% of Caucasian individuals were in possession of tA01 motifs lacking the ability to encode a functional KIR2DS4 receptor and thus devoid of an activating *KIR* receptor.

While the functional significance of KIR2DS4v has yet to be elucidated, this isoform has been maintained at variable frequencies in populations across the world (Middleton *et al*., 2007) and represents the dominant form of the receptor present in not only the South African Caucasian, but also most other Caucasian populations examined to date (van der Slik *et al*., 2003; Denis *et al*., 2005; Du *et al*., 2007; Flores *et al*., 2007; Karlsen *et al*., 2007). Binding between KIR2DS4v and known KIR2DS4 HLA ligands has not been demonstrated (Graef *et al*., 2009). However, the observation that low levels of transcription of this variant do occur in NK cells (Middleton *et al*., 2007), as well as that possession of this isoform has been associated with disease susceptibility (Taniguchi and Kawabata, 2009; Zhuang *et al*., 2012), would suggest that its role in NK cell function warrants further investigation. Elucidation of KIR2DS4v function is of particular importance in a South Africa context, as possession of the alleles encoding these receptors has been associated with transmission of HIV-1 (Hong *et al*., 2013).

The roles of other KIR receptors and their associated HLA ligands in the regulation of NK cell function, however, are more clear. The co-occurrence of KIR3DL1 and HLA-Bw4 epitopes have been shown to be associated with increased NK cell functional potential (Boulet *et al*., 2010), while binding of HLA-C to KIR is thought to represent the major receptor-mediated mechanism responsible for the regulation of NK cell function (Moesta and Parham, 2012). We therefore examined the co-occurrence of genetic variability in the genes encoding these KIR receptors and their associated HLA ligands in the Black and Caucasian South African population groups. We observed that while the Black and Caucasian South African population groups do not differ with respect to co-occurrence of *HLA-Bw4* alleles with either *KIR3DL1* or *KIR3DS1*, Black individuals were significantly more likely to possess alleles containing HLA-Bw4-80I epitopes when in possession of either *KIR* gene than Caucasian individuals. The interaction between HLA-Bw4-80I and KIR3DL1 is stronger than that between the receptor and HLA-Bw4-80T epitopes (Cella *et al*., 1994), suggesting a preference for a more inhibitory interaction between the KIR receptor and its HLA ligand in the Black South African population

group. These findings are consistent with those of a previous study examining co-evolution of KIR and HLA, which noted a general positive correlation between the allele frequencies of inhibitory *KIR* and their *HLA* ligands across 30 populations world-wide (Single *et al*., 2007).

The correlations observed by Single *et al*., (2007) were also present between *HLA-C* alleles and alleles encoding their associated KIR receptors. The Black and Caucasian South African population groups are known to differ significantly with respect to their *HLA-C1* and *–C2* allele frequency distributions, with *HLA-C1* being more prevalent in the Caucasian population and *HLA-C2* being more frequently observed in Black individuals (Paximadis *et al*., 2012). We further noted a trend towards an increased frequency of *KIR2DL2* in the Black population group and *KIR2DL3* in the Caucasian group. It was, thus, unsurprising to find increased co-occurrence of *KIR2DL2* and *HLA-C2* in Black individuals and *KIR2DL3* and *HLA-C1* in Caucasian individuals (data not shown). Therefore, in an effort to control for potentially confounding associations arising from these differences in the *HLA-C* allele frequency distributions of these two population groups, KIR and HLA-C ligand interactions were evaluated separately for individuals in possession of *HLA-C1* and those in possession of *HLA-C2*. This allowed us to observe that Black individuals in possession of *HLA-C1* were significantly more likely to also be in possession of *KIR2DL2* than their Caucasian counterparts, while Caucasian individuals with *HLA-C2* were more likely to carry *KIR2DL3* than their Black counterparts.

While binding of KIR2DL1 to HLA-C is limited to HLA-C2, KIR2DL2 and KIR2DL3 can bind both HLA-C1 and –C2 epitopes. However, the binding of HLA-C1 to these receptors occurs with reduced avidity and induces less effective inhibition of NK cell-mediated cytotoxicity than that of HLA-C2 binding (Moesta *et al*., 2008; Hilton *et al*., 2012). This, coupled with the observation that KIR2DL2 binds both HLA-C1 and –C2 with greater avidity than KIR2DL3 (Moesta *et al*., 2008; Hilton *et al*., 2012; Frazier *et al*., 2013), would suggest a tendency in both population groups towards reduced avidity between HLA-C and its KIR receptors. *KIR2DS1*

was also seen to be significantly more common in Caucasian individuals with both *HLA-C1* and *–C2* than in their counterparts from the Black population group. However it was not possible to establish whether this observation was representative of a true association between these genes, stemming from their functional interaction, or simply a consequence of the differences in *KIR2DS1* gene frequency observed between the two population groups

Collectively, these data describe patterns of genetic diversity within the *KIR* gene complex in the Black and Caucasian South African population groups, providing the first description of *KIR* haplotype content in these populations. The haplotypic differences observed between these two groups are of particular importance because they affect the frequency distributions of the disease-associated genes, *KIR2DS1*, *KIR2DS4* and *KIR3DS1* within these populations. If unaccounted for, these differences could potentially confound future disease association studies conducted within these groups. The findings regarding *KIR2DS4* and *KIR3DS1* are especially important in the South African context, given the association between these genes and HIV-1 transmission (Merino *et al*., 2011; Hong *et al*., 2013) and disease progression (Martin *et al*., 2002a; Martin *et al*., 2007).

Chapter 3

Genetic Variability in Markers of HLA-C Expression Show Similarity between Two Diverse South African Populations[1]

[1] ***Published***: Gentle *et al*., (2012) *PLoS One* **8**: e67780

3.1 Introduction

A SNP 35 kb upstream of *HLA-C* has been shown to be associated with differences in HIV-1 viral set point (Fellay *et al*., 2007) - a key early determinant of the rate of HIV-1 disease progression. This SNP (*-35T/C*; rs9264942) also strongly associates with differences in HLA-C mRNA (Stranger *et al*., 2005; Fellay *et al*., 2007) and cell surface expression levels (Thomas *et al*., 2009). However, while these associations have consistently been shown in Caucasian cohorts (Fellay *et al*., 2007; Dalmasso *et al*., 2008; Fellay *et al*., 2009; Thomas *et al*., 2009; van Manen *et al*., 2009; Ballana *et al*., 2012), the association between the *-35 SNP* (rs9264942) and HIV-1 viral set point has not been shown to be significant in African American cohorts - despite the presence of this SNP in this population (Han *et al*., 2008; Shrestha *et al*., 2009; Apps *et al*., 2013). This has lead to the suggestion that the *-35 SNP* is not the causative variant responsible for the alteration in HLA-C expression and viral set point, but rather acts as a marker in Caucasian populations for another functionally relevant polymorphism.

This view is supported by the findings of Kulkarni *et al*. (2011), who identified a single base pair indel polymorphism at position 263 of the *HLA-C* 3' UTR that has been shown to be in LD with the *-35 SNP* in Caucasian individuals. They found that this variant (*263* indel; rs67384697) affects the binding of a regulatory microRNA (miRNA148a) to the 3' UTR, with a deletion at this position (*263del*) abolishing miRNA148a binding and leading to increased HLA-C expression (Kulkarni *et al*., 2011; Kulkarni *et al*., 2013). They also found an over-representation of the *263del* allele in HIV-1 controllers relative to non-controllers in a cohort of HIV-infected

individuals of European ancestry, suggesting that the change in HLA-C expression associated with this allele could help to provide long-term protection against HIV-1 disease progression (Kulkarni *et al*., 2011). However, these data are again only representative of individuals of Caucasian ancestry and have yet to be confirmed in other populations.

That differences exist in patterns of genetic variation, and specifically in patterns of LD, between Black and Caucasian populations has become increasingly apparent. A recent study described *HLA* class I diversity in both Black and Caucasian South African populations (Paximadis *et al*., 2012) and outlined the patterns of LD that characterise this region, highlighting the key differences in *HLA-C* allelic representation between these two population groups. However, no data are yet available regarding the *-35* SNP or describing variation in the *HLA-C* 3' UTR in these populations. Here we report the first description of these data in 264 unrelated South Africans from both the Black and Caucasian population groups.

3.2 Materials and Methods

3.2.1 Study Populations

A total of 264 HIV-1 negative, unrelated South African individuals were used to describe genetic variation and patterns of LD within and between the coding and regulatory regions of the *HLA-C* locus. These 167 Black and 97 Caucasian South Africans were selected from a larger previously described cohort (Paximadis *et al*., 2012) on the basis of a non-reactive HIV ELISA test (Genscreen HIV1/2 version 2; Bio-Rad, Marnes-La-Coquette, France). Informed consent was obtained from all study participants and the study was approved by the University of the Witwatersrand Committee for Research on Human Subjects (Appendix B).

The genomic DNA used for genotyping was extracted from buffy coat samples using the PEL-FREEZ DNA Isolation Kit (DYNAL Invitrogen Corporation, Carlsbad, California, USA).

3.2.2 HLA-C Genotyping

HLA-C genotyping was performed at both low and high resolution, using a combination of single specific primer-polymerase chain reaction (SSP-PCR) and sequence based typing (SBT) methods, as previously described (Paximadis *et al*., 2012). However, because the genotyping was performed prior to 2005, it was possible that alleles that had not yet been identified at the time of genotyping may have been present within the sample population. Misclassification of these alleles could potentially confound LD analyses. *HLA-C*02:10* has previously been found at relatively high frequencies in other Black populations (Gonzalez-Galarza *et al*., 2011), however, it was not observed in the Black South African population during the prior genotyping (Paximadis *et al*., 2012). *HLA-C*02:02* and *HLA-C*02:10* differ by only a single amino acid (T211C) in exon 4 (Robinson *et al*., 2014), a difference that would not be detected by the SSP-PCR genotyping kit previously employed to genotype these samples.

Therefore, all Black individuals who had originally been typed as having *HLA-C*02:02* alleles; as well as all individuals who were initially typed as homozygous at the *HLA-C* locus, were re-genotyped using the AlleleSEQR HLA-C PLUS Sequence-Based Typing (SBT) Kit (Abbot Molecular, Des Plaines, Illinois, USA) according to the manufacturer's instructions. Sequencing analysis and allele assignment were performed using Assign™ SBT v3.5.1 software (Conexio Genomics, Fremantle, Western Australia, Australia), with the IMGT/HLA July 2011 (v3.6.0) references. Because at least one of the alleles in any genotype combination was already known (based on the previously available genotyping data), none of the retyped samples were regarded as ambiguous.

3.2.3 HLA-C 3' UTR DNA Sequencing

The *HLA-C* 3' UTR was amplified from genomic DNA using the PCR primers described by Kulkarni *et al*. (2011) and the following thermocycling conditions: 94°C for 2 minutes, followed by 30 cycles of 94°C for 15 seconds, 65°C for 30 seconds, 72°C for 90 seconds and

72°C for 7 minutes. The amplicons were sequenced in both directions by capillary electrophoresis using an ABI 3100 DNA analyzer (Applied Biosystems, Foster City, California, USA) and the sequencing primers: $^{5'}$-GTGAGATTCTGGGGAGCTGA-$^{3'}$ and $^{5'}$-TCTGGAAGGTTCTCAGGTC-$^{3'}$. The chromatograms obtained were analysed using Sequencher v4.2 (Genes Codes Corporation, Ann Arbor, Michigan, USA) and sequences were aligned with an available reference sequence (GenBank Accession Number NG_029422) to identify polymorphic positions.

3.2.4 -35 SNP Genotyping

A real-time PCR assay was designed to genotype the *-35* SNP (rs9264942). PCR amplicons were amplified from genomic DNA using a common forward primer ($^{5'}$-GCCCATACCTGTTTATACATCCA-$^{3'}$) and one of two allele-specific reverse primers ($^{5'}$-CAGAAAGTCCCACAGTGCCTG-$^{3'}$ and $^{5'}$-CAGAAAGTCCCACAGTGCCTA-$^{3'}$). Both allele-specific primers were designed with a lock nucleic acid (LNA) modified 3'-end base. The assay was performed using the Applied Biosystems 7500 Real-Time PCR system (Applied Biosystems, Foster City, California, USA) and the following thermocycling conditions: 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds, 60°C for 15 seconds and 70°C for 1 minute. Reactions were performed in a 10 µl volume, containing 1x Maxima SYBR Green/ROX qPCR Master Mix (Fermentas, Burlington, Canada), 10 pmol of each primer and 20-60 ng of DNA. PCR analysis was based on calculation of ΔC_T (difference in cycle threshold; the difference between the C_T of the *-35T* and *-35C* reactions). Heterozygous individuals had ΔC_T values of between 0 and 0.3, while homozygous *TT* and *CC* individuals had ΔC_T values of ≥ 4 or ≤ -4, respectively.

3.2.5 Statistical Analyses

Allele frequencies at all polymorphic positions were determined by direct counting. The significance of differences in allele frequency observed between the Black and Caucasian population groups were performed using a two-sided Fisher exact test (SISA: Simple Interactive Statistical Analysis). Deviations from Hardy-Weinberg equilibrium (HWE) were assessed using the Markov chain exact test for HWE (Guo and Thompson, 1992). The Excoffier-Laval-Balding (ELB) algorithm (Excoffier *et al*., 2003) was used to estimate the gametic phase of the genotypic data generated at all polymorphic positions, LD to be quantified as the LD coefficients D' (Lewontin, 1964) and r^2 (Hill and Robertson, 1968). The significance of pairwise LD between all polymorphic positions was estimated using an exact test for LD (Slatkin, 1994). All HWE, LD and haplotypic analyses were implemented through Arlequin v3.5.1.2 (Excoffier and Lisher, 2010). All statistical measures were considered significant at $p<0.05$.

3.3 Results

3.3.1 Genetic Variation within the HLA-C 3' UTR

Direct sequencing of the 3'UTRs of the *HLA-C* alleles present in the Black and Caucasian South African population revealed this region contained 33 polymorphisms within the approximately 400 bp of sequence analysed, including the indel at position 263 (Table 3.1). Thirty-one positions (including the *263* indel) were found to be polymorphic in both the Black and the Caucasian population groups, while positions 84 (*84G/A*; rs139211788) and 285 (*285ACTT/-*; rs60637457) were only polymorphic within the Black population. Minor allele frequencies differed significantly between the two population groups at 10 of these positions; however, these did not include the *263* indel (Table 3.1). Seven other SNPs previously reported in this region were not observed in either population. The Black population group deviated significantly from HWE at positions 46 ($p = 0.026$), 110 ($p < 0.001$), 138 ($p = 0.001$), 267 ($p =$

0.023) and 303 ($p < 0.001$), while positions 263 ($p = 0.033$), 266 ($p = 0.032$), 294 ($p = 0.032$), 299 ($p = 0.031$), 300 ($p = 0.032$), 307 ($p = 0.024$), 345 ($p = 0.032$) and 346 ($p = 0.032$) deviated significantly from HWE in the Caucasian population group.

A number of the polymorphisms present in the 3' UTR were found to only be associated with specific *HLA-C* alleles (Figure 3.1). For the majority of alleles, these allele-specific polymorphisms were associated with the same *HLA-C* allele (or alleles) in both population groups (when the given *HLA-C* allele was present in both population groups). However, this was not the case for *C*02:02* and *C*02:05*, which were found to be associated with different 3' UTR polymorphisms, depending on the population group under consideration. *C*02:10* was also unique in that, unlike the other *HLA-C* alleles that were uniquely associated with either one *263* indel variant or the other, this allele was found to associated with both the insertion and the deletion variants. This was also true for the indel at position 230 (Figure 3.1).

Table 3.1 The polymorphic positions present within the *HLA-C* 3'UTR and their minor allele frequencies within the Black and Caucasian population groups

Polymorphic Position[1]	NCBI dbSNP ID	Alleles	Minor Allele Frequency		p-value[2]
			Black Individuals (n = 167)	Caucasian Individuals (n = 96)	
46	rs1049853	C/T	0.067	0.089	-
84	rs1049724	A/G	0.003	0.000	-
92	rs1049709	A/G	0.052	0.177	<0.001
93	rs1065711	C/T	0.061	0.151	0.001
101	rs3176007	C/T	0.046	0.083	-
110	rs41289069	C/T	0.135	0.047	0.001
125	rs1049668	A/G	0.117	0.005	<0.001
133	rs1049663	G/T	0.117	0.005	<0.001
138	rs1049650	C/G	0.065	0.094	-
146	rs116229144	C/T	0.117	0.005	<0.001
179	rs1049579	C/T	0.117	0.005	<0.001
224	rs1094	A/G	0.379	0.401	-
230	rs35877659	G/-	0.379	0.401	-
256	rs1130592	A/C	0.256	0.373	-
259	rs3207555	C/T	0.307	0.260	-
261	rs3207561	C/T	0.307	0.260	-
263	rs67384697	G/-	0.354	0.313	-
266	rs1130586	C/T	0.351	0.313	-
267	rs1130580	A/G	0.253	0.286	-
278	rs1130576	A/G	0.122	0.005	<0.001
285	rs60637457	ACTT/-	0.003	0.000	-
294	rs1130559	A/C	0.351	0.313	-
299	rs1130558	A/G	0.351	0.313	-
300	rs1130554	A/T	0.351	0.313	-
303	rs1130552	A/G	0.065	0.135	0.006
307	rs1071643	C/G/T	0.003	0.010	-
324	rs1130538	G/T	0.393	0.401	-
345	rs116302614	A/G	0.351	0.313	-
346	rs115906458	A/G	0.351	0.313	-
347	rs3189472	C/G	0.393	0.401	-
356	rs115510686	A/G	0.122	0.005	<0.001
375	rs114027487	C/T	0.074	0.078	-
379	rs1049281	A/G	0.393	0.401	-

[1] Positions are given relative to the start of the *HLA-C* 3' UTR, as seen in Kulkarni *et al.* (2011).

[2] p-values are for a two-sided Fisher's exact test, only significant values ($p < 0.05$) are shown.

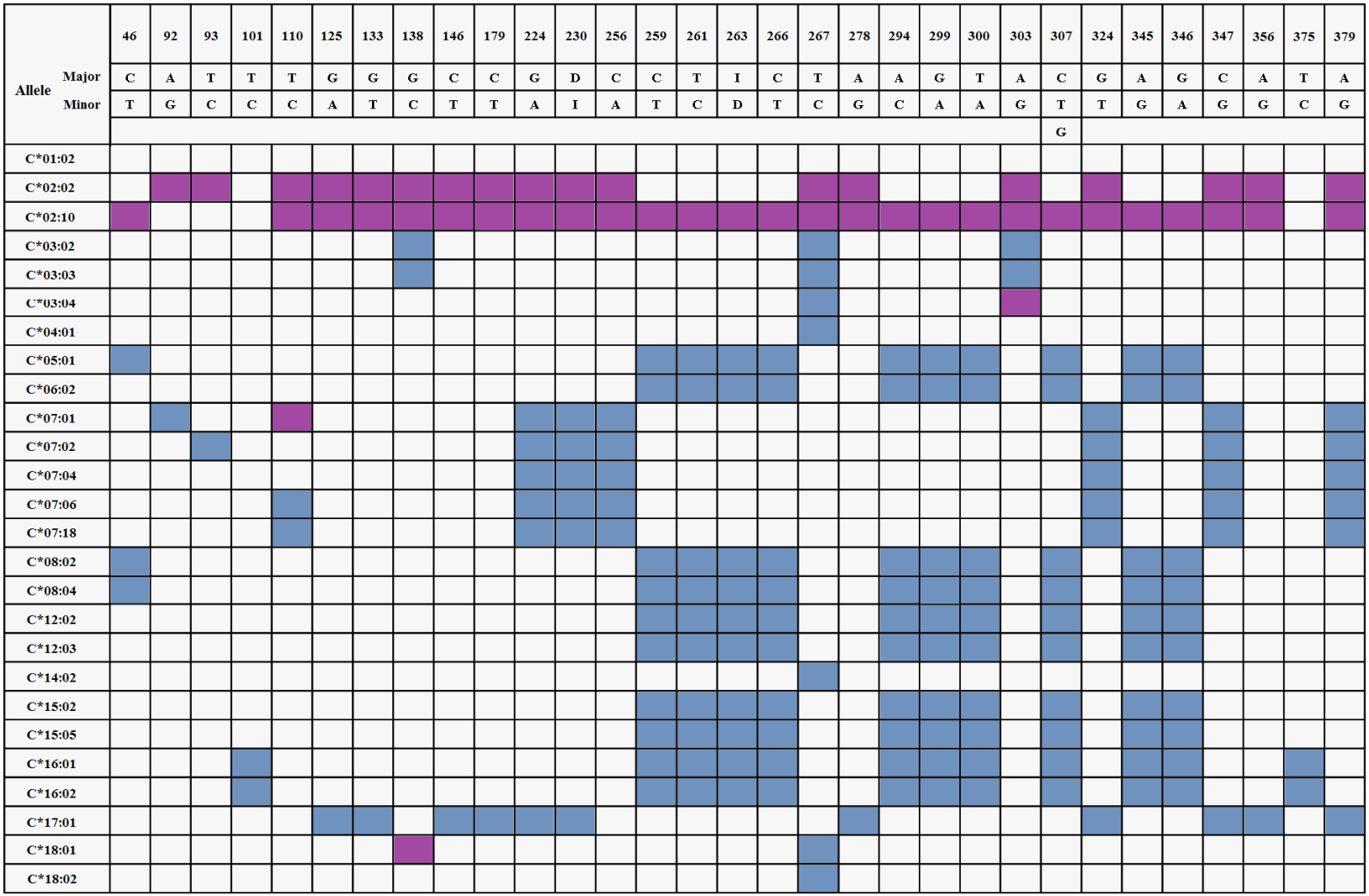


Figure 3.1: The association between the polymorphic variants observed within the *HLA-C* 3' UTR and specific *HLA-C* alleles. Only alleles observed more than once in either population group have been included. The major and minor alleles at each polymorphic position are given, while the presence of the minor allele within a given *HLA-C* allele is indicated in blue. When both allelic variants have been observed to be associated with a particular HLA-C allele, it is indicated in purple. At position 307, which is tri-allelic, the presence of the T allele has been indicated. The G allele at this position is uniquely associated with *HLA-C*14:02*.

3.3.2 Estimation of Haplotype Structure within the HLA-C 3' UTR

In order to further investigate the relationships between the genetic variants identified within the *HLA-C* 3' UTR, the ELB algorithm (Excoffier *et al*., 2003) was used to estimate the gametic phase of the genotyping data generated at all the polymorphic positions within this region, allowing LD between these positions to be quantified. These LD analyses allowed for the identification of two overlapping haplotypes, involving multiple positions across the *HLA-C* 3' UTR. The larger of the haplotypes spanned across ten positions (263H - Figure 3.2a) and included the indel at position 263. The other extended across five positions (230H - Figure 3.2a) and encompassed the indel at position 230. Only positions with both D' and r^2 measures of pairwise LD equal to 1 were considered as these haplotypes. Both haplotypes were present at similar frequencies in both population groups.

Because of strong linkage disequilibrium between the alleles at positions 230 and 263 (D' = 1 and $r^2 = 1$, $p < 0.001$), the two haplotypes were only found to occur in three possible combinations (Figure 3.2b). The most commonly observed haplotypic combination in both groups was that of the 230H minor allele haplotype with the 263H major allele haplotype (*230I/263I*). Slightly less frequent in both population groups was the combination of the 230H major allele haplotype with the 263H minor allele haplotype (*230D/263D*). The combination of both major allele haplotypes (*230D/263I*) was the least common in both groups, while the combination of both minor allele haplotypes was never observed. The *230D/263D* haplotype was found to occur at a slightly higher frequency in the Black population group, but this comparison did not reach statistical significance.

a.

	Polymorphic Position																															
	46	84	92	93	101	110	125	133	138	146	179	224	230	256	259	261	263	266	267	278	285	294	299	300	303	307	324	345	346	347	356	375	379
Major Allele	C	G	A	T	T	T	G	G	G	C	C	G	I	C	C	T	I	C	A	A	I	A	G	T	A	C	G	A	G	C	A	T	A
Minor Allele	T	A	G	C	C	C	A	T	C	T	T	A	D	A	T	C	D	T	G	G	D	C	A	A	G	T	T	G	A	G	G	C	G
																										G							
230H																																	
263H																																	

b.

	Polymorphic Position																																F_B	F_C
	46	84	92	93	101	110	125	133	138	146	179	224	230	256	259	261	263	266	267	278	285	294	299	300	303	307	324	345	346	347	356	375	379	
230I/263I																																		0.383	0.401
230D/263I																																		0.267	0.286
230D/263D																																		0.350	0.313

Figure 3.2: **(a)** The haplotypes occurring within the 3' UTRs of the *HLA-C* alleles present in the Black and Caucasian South African population groups. The haplotypes were named 230H and 263H, for their association with the two indels present in the region. The major and minor alleles at each polymorphic position are indicated. **(b).** The combinations of the 230H and 263H haplotypes observed within the Black and Caucasian South African population groups. F_B refers to the frequency of the given combination in the Black population group, while F_C refers to the frequency within the Caucasian population group. The presence of the major allele at each polymorphic position is indicated in blue, while the presence of the minor allele is indicated in purple.

3.3.3 Linkage Disequilibrium between Specific HLA-C alleles and the -35 SNP

A SNP 35 kb upstream of *HLA-C* has previously been shown to be associated with both HLA-C expression and HIV-1 viral set point (Stranger *et al*., 2005; Fellay *et al*., 2007; Thomas *et al*., 2009). However, while repeatedly observed within Caucasian populations (Fellay *et al*., 2007; Dalmasso *et al*., 2008; Thomas *et al*., 2009; van Manen *et al*., 2009), this association has not been observed in populations of African descent (Han *et al*., 2008; Shrestha *et al*., 2009; Apps *et al*., 2013). We therefore examined genetic variability at the -35 SNP (rs9264942) and evaluated the extent of LD between this polymorphism and the *HLA-C* alleles present in the Black and Caucasian population groups (Table 3.2).

The frequencies of the *-35C* allele in the Black (0.348) and Caucasian (0.284) population groups were not found to be significantly different ($p = 0.13$), and neither population group deviated significantly from HWE at this position. In the Black population group, the *HLA-C* alleles *C*02:02*, *C*03:02*, *C*06:02*, *C*07:06*, *C*07:18*, *C*12:03* and *C*14:02* were found to be in complete linkage with this allele, while *C*03:04*, *C*04:01*, *C*17:01* and *C*18:02* were in complete LD with the *-35T* allele. While not complete, *C*02:10* ($D' = 0.59$; $p < 0.05$), *C*07:02* ($D' = 0.86$; $p < 0.01$) and *C*16:01* ($D' = 0.76$; $p < 0.01$) were also found to be in significant LD with the *-35T* allele. Of the 26 *HLA-C* alleles present in the Caucasian population, 15 were found to be in significant LD with the *-35* SNP (Table 3.2). *C*02:02*, *C*06:02*, *C*07:06*, *C*07:18*, *C*08:04*, *C*12:02*, *C*12:03* and *C*14:02* were in complete linkage with the *-35C* allele, while *C*03:03*, *C*03:04*, *C*04:01*, *C*07:01*, *C*07:02* and *C*16:01* were in complete LD with the *-35T* allele. Although not complete, *C*01:02* was also found to be in significant LD with the *-35C* allele ($D' = 0.65$, $p < 0.05$).

Table 3.2 Linkage disequilibrium between the *-35* SNP (rs9624942) and the *HLA-C* alleles present in the Black and Caucasian South African population groups

HLA-C Allele	Black Individuals (n = 168)				Caucasian Individuals (n = 97)			
	Allele	**Frequency**	**D'**	**p-value[1]**	**Allele**	**Frequency**	**D'**	**p-value[1]**
01:02	-	-	-	-	C	0.0155 (3)	0.65	<0.05
02:02	C	0.0149 (5^5)	1.00	<0.05	C	0.0723 (14)	1.00	<0.01
02:05	T	0.0030 (1)	1.00	NS	C	0.0052 (1)	1.00	NS
02:10	T	0.0655 (22)	0.59	<0.05	-	-	-	-
03:02	C	0.0149 (5)	1.00	<0.01	-	-	-	-
03:03	C	0.0030 (1)	1.00	NS	T	0.0619 (12)	1.00	<0.05
03:04	T	0.0446 (15)	1.00	<0.01	T	0.0567 (11)	1.00	<0.05
03:16	-	-	-	-	T	0.0052 (1)	1.00	NS
04:01	T	0.1250 (42)	1.00	<0.01	T	0.0876 (17)	1.00	<0.01
04:04	T	0.0030 (1)	1.00	NS	-	-	-	-
04:08	-	-	-	-	T	0.0052 (1)	1.00	NS
05:01	C	0.0060 (2)	0.49	NS	T	0.0412 (8)	0.28	NS
06:02	C	0.1488 (50)	1.00	<0.01	C	0.0773 (15)	1.00	<0.01
06:06	C	0.0030 (1)	1.00	NS	-	-	-	-
06:11	C	0.0030 (1)	1.00	NS	C	0.0052 (1)	1.00	NS
07:01	T	0.0506 (17)	0.16	NS	T	0.1753 (34)	1.00	<0.01
07:02	T	0.0595 (20)	0.86	<0.01	T	0.1443 (28)	1.00	<0.01
07:04	T	0.0149 (5)	1.00	NS	T	0.0103 (2)	1.00	NS
07:06	C	0.0387 (13)	1.00	<0.01	C	0.0206 (4)	1.00	<0.01
07:11	T	0.0030 (1)	1.00	NS	-	-	-	-
07:18	C	0.0417 (14)	0.89	<0.01	C	0.0155 (3)	1.00	<0.01
08:01	T	0.0030 (1)	1.00	NS	T	0.0052 (1)	1.00	NS
08:02	T	0.0119 (4)	0.42	NS	C	0.0103 (2)	0.31	NS
08:04	C	0.0149 (5)	0.24	NS	C	0.0103 (2)	1.00	<0.05
12:02	-	-	-	-	C	0.0103 (2)	1.00	<0.05
12:03	C	0.0149 (5)	1.00	<0.01	C	0.0206 (4)	1.00	<0.01
14:02	C	0.0060 (2)	1.00	<0.05	C	0.0103 (2)	1.00	<0.05
15:02	T	0.0060 (2)	1.00	NS	T	0.0258 (5)	1.00	NS
15:05	T	0.0089 (3)	1.00	NS	-	-	-	-
16:01	T	0.0655 (22)	0.76	<0.01	T	0.0670 (13)	1.00	<0.05
16:02	-	-	-	-	T	0.0103 (2)	1.00	NS
17:01	T	0.1131 (38)	1.00	<0.01	T	0.0052 (1)	1.00	NS
18:01	T	0.0149 (5)	1.00	NS	-	-	-	-
18:02	T	0.0298 (10)	1.00	<0.05	-	-	-	-

[1] P-values are for an exact test for linkage disequilibrium (Slatkin, 1994) and are significant at $p < 0.05$.

3.3.4 Linkage Disequilibrium between the -35 SNP and the 263 Indel

It has been suggested that the association observed between the *-35* SNP and HLA-C expression observed in populations of European descent is due strong LD between this SNP and the indel at position 263 of the HLA-C 3' UTR (Kulkarni *et al*., 2011). We therefore evaluated the extent of LD between these two polymorphisms in the Black and Caucasian South African population groups and found that in contrast to previous observations in other Caucasian populations (D' = 0.75; Kulkarni *et al*., 2011), the *-35* SNP was found to be only weakly associated with the *263* indel in both the Black (D' = 0.46; $p < 0.001$) and the Caucasian (D' = 0.34; $p < 0.001$) population groups. In both groups, the *-35C* allele was associated with *263del* and the *-35T* allele with *263ins*.

3.4 Discussion

We examined genetic variation within the 3' UTRs of the *HLA-C* alleles present in the Black and Caucasian South African population groups and investigated the relationship between these variants and a SNP 35 kb upstream of the *HLA-C* locus. The data confirmed the presence of the *263* indel (rs67384697) in both population groups, as well as that of other polymorphisms previously identified within this region in other populations (Xu *et al*., 2011; Kulkarni *et al*., 2011; O'hUigin *et al*., 2011). The persistence of strong LD was observed between particular 3' UTR polymorphisms in both population groups and the underlying haplotypic structure of the region was described. Finally, these data demonstrated that the *-35* SNP (rs9624942) is not in strong LD with the *263* indel in either the Black or the Caucasian South African population; and as such, is not an appropriate marker for this indel in either of these groups.

Previous investigations of genetic variability within the *HLA-C* 3' UTR in other populations have identified numerous polymorphisms other than the *263* indel in this region (Xu *et al*.,

2011; Kulkarni *et al*., 2011; O'hUigin *et al*., 2011). Consistent with their findings, our analysis of the same region identified 33 of these polymorphisms in the Black South African population. The variation identified within the Caucasian South African population represented only a subset of that seen in the Black population group. However, the additional polymorphisms observed in the Black population group were not unique to this population, having been previously described in other studies (Xu *et al*., 2011; Kulkarni *et al*., 2011; O'hUigin *et al*., 2011). A number of these polymorphisms were found to deviate from HWE in both population groups - although different polymorphisms were seen to deviate from HWE between the Black and Caucasian population groups. While deviation from HWE can often be attributed to genotyping errors, this is unlikely to be the case in this instance, as all genotyping was performed by direct sequencing. In light of the finding that a subset of these polymorphisms are unique to specific *HLA-C* alleles and given the differences observed in the *HLA-C* allele distributions between the two population groups, these deviations from HWE most likely reflect the influence of random genetic drift.

As previously observed (Xu *et al*., 2011; Kulkarni *et al*., 2011; O'hUigin *et al*., 2011), the majority *HLA-C* alleles observed within the Caucasian population group were represented by a single, distinctive 3' UTR sequence. However, contrary to prior observations, several different sequences were observed for the *C*02:02* allele in this population group – none of which corresponded to the sequence previously reported for this allele. Nonetheless, despite this variability, none of the 3' UTR sequences observed for *C*02:02* differed with respect to their *263* indel allele. A similar pattern was observed for the 3' UTR sequences of the *HLA-C* alleles present in the Black population group – with *C*02:02* again showing a high degree of variability within this population group, but without differing with respect to the *263* indel allele observed. However, one allele present only in the Black population group (*HLA C*02:10*), was found to be associated with both the insertion and the deletion at position 263. Coupled with

the finding that the 3' UTR sequence observed for *C*02:05* differed between the two population groups, these data may suggest that a diversity of genetic variability exists within the HLA Cw 02 alleles that cannot be adequately described using the current commonly used *HLA-C* SBT genotyping methods.

LD analysis between all the polymorphisms present in the *HLA-C* 3' UTR revealed the *263* indel to be at the centre of a large haplotype involving the same ten polymorphic positions in both the Black and Caucasian populations. This indel is thought to regulate differential *HLA-C* expression by disrupting a putative miRNA binding site (Kulkarni *et al*., 2011; Kulkarni *et al*., 2013). It has previously been shown to be in complete linkage with the SNPs at positions 256, 261 and 266; and as none of these SNPs produced any significant change in luciferase activity when analysed independently, the indel is thought to be responsible for any differences seen in HLA-C expression (Kulkarni *et al*., 2011). However, all analyses of alterations in luciferase activity involving the indel included concomitant modifications at all four positions, and all four polymorphisms occur within the putative miR-148a/miR-148b binding site (Kulkarni *et al*., 2011), allowing for the possibility that concomitant changes at these positions may act synergistically to disrupt miRNA binding.

Similarly, the SNP at position 307 was also previously shown to be potentially disruptive of a putative miRNA binding site, but when analysed independently did not produce any significant changes in luciferase activity (Kulkarni *et al*., 2011). However, if analysed in haplotypic combination with the SNPs at positions 299, 300 and 303, which occur within the same putative mi-657 binding site (Kulkarni *et al*., 2011); a significant difference in expression may potentially be observed, as the presence of multiple mismatched bases within the same binding site would most likely increase the probability of disruption of miRNA binding relative to modification at a single position. Additionally, the observation that these variants always occur

in combination with each other, and could thus collectively potentially prevent binding of more than one miRNA, would further suggest that the presence of the haplotype (rather than any one single polymorphism) is responsible for the differential expression of *HLA-C* alleles observed. However, further functional studies would be required to test this hypothesis.

A second large haplotype encompassing the indel at position 230 and four other SNPs was also observed in both populations. The presence of an insertion at position 230, coupled with a guanine at 224, introduces at additional putative miR-181a binding site into the *HLA-C* 3' UTR (Kulkarni *et al*., 2011). Because of strong LD between the two indels, these haplotypes occur in three overlapping configurations within all the *HLA-C* alleles observed within the Black and Caucasian South African population groups, each of which is potentially subject to differential degrees of miRNA-mediated gene regulation. *HLA-C* 3' UTR sequences with deletions at both positions 230 and 263 would therefore potentially disrupt the binding of four miRNAs (miR-181a, miR-148a/miR-148b, miR-657 and miR-181a*; Kulkarni *et al*., 2011), while those sequences with insertions would be inhibited at the same positions. Sequences with a deletion at position 230 and insertion at 263 would disrupt two putative miRNA binding sites (miR-148a/miR-148b and miR-657; Kulkarni *et al*., 2011), and thus potentially exhibit a phenotype of intermediate expression. The presence of these alternative haplotype combinations could thus account for the variances previously observed in HLA-C expression (Thomas *et al*., 2009; Corrah *et al*., 2011; Kulkarni *et al*., 2011). Again, however, further functional studies are necessary to test this hypothesis.

The suggestion that the overlap of the two haplotypes described could potentially account for observed variances in HLA-C expression is consistent with the previous findings of Corrah *et al*. (2011), who prior to the identification of the *263* indel, used monoclonal antibodies to examine HLA-C expression in relation to *-35* SNP genotype (Corrah *et al*., 2011). They attributed the association between the *-35T* allele and higher HIV-1 viral set point observed in

Caucasian individuals (Fellay *et al*., 2007; Dalmasso *et al*., 2008; Fellay *et al*., 2009; Thomas *et al*., 2009; van Manen *et al*., 2009) to the especially low expression of HLA-Cw*07 alleles, which are particularly common in individuals of European descent (Gonzalez-Galarza *et al*., 2011). This association is not seen in African-American populations (Han *et al*., 2008; Shrestha *et al*., 2009), where HLA-Cw*07 is less prevalent. All the HLA-Cw*07 sequences analysed during the course of our investigation had insertions at both indel positions, and would thus (under the aforementioned model) be subject to especially strong miRNA-mediated inhibition. This could also partially account for variances in the distribution of HLA-C expression levels observed in other functional studies (Kulkarni *et al*., 2011; Apps *et al*., 2013).

This association between the *-35* SNP and HLA-C expression (Thomas *et al*., 2009), was previously thought to be responsible for the differences in HIV-1 viral set point observed in Caucasian individuals. However, the *263* indel is now considered to be the actual variant responsible for the differential regulation of HLA-C expression, which, in turn, is thought to be responsible for the variation seen in HIV-1 viral set point (Kulkarni *et al*., 2011). The association observed between the *-35* SNP and in HIV-1 viral set point has since been attributed to strong LD between the SNP and indel (Kulkarni *et al*., 2011).

Examination of the extent of LD between the *-35* SNP and *263* indel in both the Black and Caucasian populations revealed LD between them (and all the other SNPs in the haplotype) to be weak in both groups. However when linkage between the *-35* SNP and specific *HLA-C* alleles was examined, LD was found to be complete and significant for fourteen alleles in the Caucasian population. This was consistent with previous reports in other Caucasian populations (Thomas *et al*., 2009; Kulkarni *et al*., 2011). LD was not found to be significant for more than 50% of the *HLA-C* alleles in the Black population group. This is consistent with observations in African-American populations, where common *HLA-C* alleles show no significant LD with the

-35 SNP and for alleles that do show significant LD, linkage is not complete (The International HIV Controllers Study *et al*., 2010).

Thus while the *-35* SNP may be indicative of specific *HLA-C* alleles in the Caucasian population, it cannot be regarded as an effective marker for the *263* indel in either the Black or Caucasian South African population groups. The *-35* SNP was also not in strong LD with the *230* indel in either population group (data not shown). Thus, as has been shown in the African-American population (The International HIV Controllers Study *et al*., 2010), this SNP is unlikely to associate with either HLA-C expression or HIV-1 viral set point in these groups. However, this is yet to be confirmed. Whether the *-35* SNP associates with any other as yet unidentified polymorphisms that influence HIV-1 disease progression by alternative mechanisms may warrant further investigation.

In conclusion, these data provide the first description of variation in the regulatory regions of *HLA-C* for the Black and Caucasian South African populations. Furthermore, the data from the Black population are the first for a sub-Saharan African population. These data allow for a description of the haplotypic patterns within the *HLA-C* 3' UTR and identify two overlapping haplotypes within this region - which we hypothesize may act independently and synergistically to influence miRNA regulation of HLA-C expression. Concomitantly, we demonstrate that the *-35* SNP is not in strong LD with either haplotype (in either population) and as such is unlikely to be an appropriate marker for HLA-C expression in either of these populations. Even in the absence of supporting expression data, these findings provide important insights into genetic variability within the regulatory regions of *HLA-C*, that have potential implications for our understanding of the regulation of HLA-C expression and its impact on HIV-1 disease progression in the populations occupying the region worst affected by this epidemic.

Chapter 4

Genetic Variability at Residues within HLA-A and HLA-B that Determine Peptide Specificity are Associated with HIV-1 Control in Black South Africans

4.1 Introduction

The rate of HIV-1 disease progression is known to be highly variable between individuals. A small subset of HIV-1 infected individuals, termed HIV-1 controllers, are able to spontaneously suppress viral replication and maintain very low levels of viraemia (< 2000 RNA copies/ml) in the absence of highly active antiretroviral therapy. Genetic variability within the *HLA* gene complex is well established as one of the major determinants of HIV-1 control (Fellay *et al*., 2007; The International HIV Controllers Study *et al*., 2010), with associations between specific *HLA* class I alleles and haplotypes (Goulder and Walker, 2012), as well as variation within non-coding regulatory regions (Fellay *et al*., 2007; Dalmasso *et al*., 2008; Kulkarni *et al*., 2011; Ballana *et al*., 2012) and HIV-1 control having been well documented.

Recently, association testing of individual polymorphic amino acids within the classical *HLA* class I loci has also identified key residues, located predominantly within the peptide binding groove, that are strongly associated with HIV-1 control (The International HIV Controllers Study *et al*., 2010; McLaren *et al*., 2012). However, because ethnically and geographically distinct populations often display marked differences in their *HLA* allele frequency spectrums and the patterns of LD they exhibit across the *HLA* gene complex, associations observed at these loci may not always be consistently replicable in all population groups (Fellay *et al*., 2007; Shrestha *et al*., 2009; van Manen *et al*., 2009; The International HIV Controllers Study *et al*., 2010; McLaren *et al*., 2012).

Associations between specific *HLA* class I alleles and HIV-1 control have been well documented within the Black South African population (Kiepiela *et al*., 2004; Honeyborne *et al*., 2007; Leslie *et al*., 2010). However, the impact of genetic variability at individual amino acids within *HLA* class I loci, as well as at positions outside of the protein coding region, on HIV-1 control have yet to be assessed in this population. We therefore examined patterns of genetic variability within and across the *HLA-A*, *-B* and *-C* loci in 111 treatment-naïve HIV-1 infected Black South Africans exhibiting differing degrees of viral control.

Our results demonstrate that genetic variability at specific residues within the *HLA-A* and *–B* peptide binding grooves is strongly associated with HIV-1 control in this population group. In accordance with previous reports (The International HIV Controllers Study *et al*., 2010; McLaren *et al*., 2012), this association is not seen for the corresponding residues within *HLA-C*. In addition, we further show that, as seen in African-Americans (McLaren *et al*., 2012), genetic variability within the *HLA-C* 3'UTR, thought to influence HLA-C expression (Kulkarni *et al*., 2011; Kulkarni *et al*., 2013) is not significantly associated with HIV-1 control in the Black South African population.

4.2 Materials and Methods

4.2.1 Study Populations

A total of 111 treatment-naïve HIV-1 infected Black South Africans were included in this study in order to investigate the impact of genetic variability within the *HLA* gene complex on HIV-1 control. Sixty-nine of these individuals were classified as HIV-1 progressors based on a $CD4^+$ T cell count that declined to below 300 cells/mm^3 prior to the administration of antiretroviral (ART) therapy. As decisions to commence ART are based solely on $CD4^+$ T cell count, which may then include HIV-1 controllers with declining counts, we included the criterion of high

viraemia (i.e. viral loads greater than 10 000 RNA copies/ml) to identify progressors (Table 4.1). The remaining 42 HIV-1 positive individuals were classified as HIV-1 controllers on the basis of their having $CD4^+$ T counts greater than 300 cells/mm^3, while also suppressing viral replication to levels below 2000 RNA copies/ml.

Table 4.1 A summary of the clinical characteristics used to classify individuals as either HIV-1 progressors or -controllers

	$CD4^+$ T Cell Count (cells/mm^3)		**Log Viral Load (RNA copies/ml)**	
	Median	**Range**	**Median**	**Range**
HIV-1 Progressors (n = 69)	178	3 - 296	4.57	4.01 - 5.70
HIV-1 Controllers (n = 42)	699	306 - 1380	2.54	1.7 - 3.24

HIV-1 RNA levels were quantitated using either the Roche Amplicor RNA Monitor assay (lower detection limit of 400 HIV-1 RNA copies/ml) or, where necessary, the COBAS AmpliPrep/COBAS Taqman HIV-1 Test, v2.0 (Roche Molecular Diagnostics, Pleasanton, California, USA), which has a lower detection limit of 50 RNA copies/ml. $CD4^+$ T cell counts were determined using the FACSCount System (BD Biosciences, San Jose, California, USA). Informed consent was obtained from all study participants and the study was approved by the University of the Witwatersrand Committee for Research on Human Subjects (Appendix B).

4.2.2 DNA Extraction

The DNA obtained from HIV-1 controllers was extracted from whole blood samples, while that obtained from HIV-1 progressors was derived from frozen buffy coat samples. All DNA extractions were performed by silica spin-column purification using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany).

4.2.3 HLA Class I Genotyping

HLA class I genotyping was performed at high resolution for the HIV-1 controllers included in this study using the AlleleSEQR *HLA-A*, *HLA-B* and *HLA-C* PLUS SBT kits (Abbot Molecular, Des Plaines, Illinois, USA). This SBT method involved initial amplification of a 2 kb fragment of the gene, followed by bidirectional direct sequencing of exons 2, 3 and 4 by capillary electrophoresis using the ABI 3100 DNA analyzer (Applied Biosystems, Foster City, California, USA). Allele assignment was performed using Assign™ SBT v3.5.1 (Conexio Genomics, Fremantle, Western Australia, Australia), based on the IMGT/HLA July 2014 (v3.17.0) references.

The *HLA class I* genotyping for the 69 HIV-1 progressors was performed by next-generation sequencing (NGS) using the GS FLX System (454 Life Sciences, Branford, Connecticut, USA). This system involves multiplexed, clonal, amplicon-based pyrosequencing of exons 2 and 3 of all three *HLA class I* loci, using a combination of multiple locus-specific primers, each with unique multiplex identifiers (MIDs). Allele assignment was carried out using IIID Allele Caller, in-house software designed and used by the Institute for Immunology and Infectious Diseases at Murdoch University, Murdoch, Western Australia, Australia.

4.2.4 Resolution of HLA Ambiguity

The allele assignment algorithms employed during *HLA* class I genotyping assign individual genotypes based on the nucleotide sequences observed at specific key positions within exons 2, 3 and 4 of each gene. However, the presence of specific heterozygous combinations at these positions can often lead to ambiguity in the genotype assignment. This is particularly problematic in (the not uncommon) instances where two or more *HLA* alleles have identical sequences at these positions (Robinson *et al*., 2014). While these ambiguities can be resolved experimentally (Paximadis *et al*., 2012), the high costs and time consuming nature of *HLA* class

I genotyping make it more prudent to find alternative algorithmic solutions to this problem (Nunes *et al*., 2014).

Therefore, in the case of the 111 treatment-naïve HIV-1 infected Black South Africans included in this study, genotypic ambiguities were not resolved experimentally. Rather - given the high levels of LD observed across the *MHC* region and increased likelihood of observing common alleles at these loci - where ambiguities were observed, individuals were assigned the genotype regarded as being the most likely given the population-specific allele frequencies and patterns of LD previously observed across these loci by Paximadis *et al.* (2012) and others (Gonzalez-Galarza *et al*., 2011). In the case of the *HLA-C* locus, where necessary, additional data regarding allele-specific polymorphic variants within the *HLA-C* 3' UTR were included in allele assignment.

4.2.5 Evaluation of the Polymorphic Positions within HLA Class I Loci

SNPs within exons 2, 3 and 4 of each the three *HLA* class I loci were identified following multiple sequence alignment of the sequence data obtained from *HLA* allele typing with the appropriate available reference sequences (GenBank Accession Numbers NM_001242758, NM_005514 and NM_002117). The amino acid residues corresponding to all non-synonymous positions were determined based on the sequence information available within the IMGT/HLA database (Robinson *et al*., 2014).

4.2.6 HLA-C 3' UTR DNA Sequencing

The *HLA-C* 3' UTR was amplified from genomic DNA using the PCR primers described by Kulkarni *et al*. (2011), as previously described (Chapter 3; Section 3.2.3). Bi-directional sequencing of the resulting amplicons was performed by capillary electrophoresis using an ABI

3100 DNA analyzer (Applied Biosystems, Foster City, California, USA). All sequence analysis was performed using Sequencher v4.2 (Genes Codes Corporation, Ann Arbor, Michigan, USA) and sequences were aligned with an available reference sequence (GenBank Accession Number NG_029422) to identify polymorphic positions.

4.2.7 Statistical Analyses

HLA class I allele and SNP frequencies were determined by direct counting. The association between specific *HLA* class I alleles and HIV-1 control was evaluated using multivariate logistic regression, as implemented in R v3.1.1 (www.r-project.org). The significance of differences in the allele frequencies at individual polymorphic positions within *HLA* loci between HIV-1 progressors and -controllers was assessed using the Fisher's exact test implemented in PLINK v1.07 (Purcell *et al*., 2007). *HLA* class I allelic haplotypes were estimated using Arlequin v3.5.1.2 (Excoffier and Lisher, 2010), while individual SNP haplotypes were identified using PLINK v1.07 (Purcell *et al*., 2007). The significance of allele and genotype frequency differences observed between HIV-1 progressors and –controllers at polymorphic positions within the *HLA-C* 3' UTR was also assessed using a two-sided Fisher's exact test, as implemented in R v3.1.1 (www.r-project.org). All statistical measures were considered significant at $p < 0.05$.

4.3 Results

4.3.1 Association of specific HLA Class I alleles with HIV-1 Control

We evaluated the association of *HLA* class I alleles with HIV-1 control in Black South African HIV-1 progressors and controllers and found four alleles to be significantly associated with differential HIV-1 control (Table 4.2). Consistent with previous studies conducted within the Black South African population (Kiepiela *et al*., 2004; Leslie *et al*., 2010; Tang *et al*., 2011) and

others (Lazaryan *et al*., 2006; Lazaryan *et al*., 2010; McLaren *et al*., 2012), *B*57:03:01*, *B*58:01:01* and *B*39:10:01* were identified as the alleles most strongly associated with HIV-1 control. *A*03:01:01* has also previously been shown to be associated with reduced absolute $CD4^+$ T cell count in Black South Africans (Leslie *et al*., 2010). An association was not seen for any individual *HLA-C* alleles, however the association observed for *B*39:10:01* was only seen when found in haplotypic combination with *C*12:03:01*; while the *B*81:01:01|C*04:01:01* haplotype was found to be associated with HIV-1 control ($p = 0.020$, $OR = \infty$), despite no association being seen for either individual allele. These results were again consistent with previous observations within the Black South African population (Leslie *et al*., 2010).

Table 4.2 *HLA* class I alleles associated with differential HIV-1 control

***HLA* Allele**	**HIV-1 Progressors (n = 69)**		**HIV-1 Controllers (n = 42)[1]**		**OR (95% Confidence Interval)[2]**	**p-value[2]**
	n	**Frequency**	**n**	**Frequency**		
B*57:03:01	3	0.022	5	0.060	13.7 (1.5 - 159.3)	0.024
B*58:01:01	4	0.029	6	0.071	9.8 (1.3 - 99.5)	0.037
B*14:02:01	2	0.014	3	0.036	9.0 (0.8 - 13.3)	0.080
B*39:10:01	4	0.029	5	0.060	8.5 (1.2 - 70.3)	0.037
B*14:01:01	2	0.014	3	0.036	7.0 (0.7 - 8.5)	0.098
B*42:01:01	9	0.065	9	0.107	4.6 (0.9 - 28.4)	0.080
A*23:01:01	14	0.101	4	0.059	0.1 (0.0 - 1.0)	0.066
A*03:01:01	11	0.080	1	0.015	0.0 (0.0 - 0.3)	0.012

[1] Only 34 HIV-1 controllers were genotyped at the HLA-A locus.

[2] Odds ratios and p-values based on multivariate logistic regression, as implemented in R v 3.1.1 (www.r-project.org).

4.3.2 Association of Polymorphic Positions within the HLA Class I Peptide-binding Groove with HIV-1 Control

The allele frequencies of the individual polymorphic positions within the *HLA* class I loci were compared between HIV-1 progressors and –controllers, in order to assess their potential

association with HIV-1 control. Only variants within exons 2, 3 and 4 were considered, as these were the exons for which sequencing data were available. In addition, as the functional implications of genetic variability within these exons were anticipated to arise as a consequence of alterations in the resulting protein structure, only non-synonymous polymorphisms were considered – reducing the number of SNPs under consideration to 58 in *HLA-A*, 47 in *HLA-B* and 43 in *HLA-C*.

Table 4.3 The polymorphic positions within the *HLA-A* locus associated with differential HIV-1 control

Position	Amino Acid	HIV-1 Progressors (n = 69)	HIV-1 Controllers (n = 34)	OR (95% Confidence Interval)[1]	p-value[2]
17	S	0.167	0.324	2.4 (1.2 - 4.7)	0.013
56	R	0.167	0.324	2.4 (1.2 - 4.7)	0.013
97	M	0.355	0.191	0.4 (0.2 - 0.9)	0.016
114	E	0.167	0.324	2.4 (1.2 - 4.7)	0.013
116	H	0.167	0.324	2.4 (1.2 - 4.7)	0.013
144	K	0.486	0.324	0.5 (0.3 - 0.9)	0.036
149	T	0.065	0.162	2.8 (1.1 - 7.0)	0.042
152	V	0.688	0.515	0.4 (0.3 - 0.9)	0.021

[1] Odds ratio

[2] p-values are for a two-sided Fisher's exact test, as implemented in R v3.1.1 (www.r-project.org).

Eight SNPs within *HLA-A* were found to be present at significantly different frequencies in HIV-1 progressors and –controllers (Table 4.3), four of which (positions 17, 56, 114 and 116) were found to be in complete LD ($r^2 = 1$; $p < 0.001$). The Ser^{17}-Arg^{56}-Glu^{114}-His^{116} haplotype was significantly associated with HIV-1 control ($p = 0.013$; OR = 2.4) and was found to be unique to *HLA-A*30* alleles (*A*30:01:01*, *A*30:02:01*, *A*30:04:01* and *A*30:09*; Figure 4.1). An independent association with HIV-1 control was observed at position 149, while variability at positions 97, 144 and 152 was associated with loss of control (Table 4.3).

	17	56	97	114	116	144	149	152
A*01:01:01	R	G	I	R	D	K	A	A
A*01:03	R	G	M	R	D	K	A	A
A*02:01:01	R	G	R	H	Y	K	A	V
A*02:02	R	G	R	H	Y	K	A	V
A*02:05:01	R	G	R	H	Y	K	A	V
A*03:01:01	R	G	I	R	D	K	A	E
A*23:01:01	R	G	M	H	Y	Q	A	V
A*24:02:01	R	G	M	H	Y	K	A	V
A*26:01:01	R	G	R	Q	D	Q	T	E
A*29:01:01	R	G	M	R	D	Q	A	V
A*29:02:01	R	G	M	R	D	Q	A	V
A*29:11	R	G	M	R	D	Q	A	V
A*30:01:01	S	R	I	E	H	Q	A	W
A*30:02:01	S	R	I	E	H	Q	A	R
A*30:04:01	S	R	I	E	H	Q	A	V
A*30:09	S	R	I	E	H	Q	A	V
A*32:01:01	R	G	M	Q	D	Q	A	V
A*33:01:01	R	G	M	Q	D	Q	A	V
A*33:03:01	R	G	M	Q	D	Q	A	V
A*34:02:01	R	G	I	R	D	Q	T	E
A*36:01	R	G	I	R	D	K	T	A
A*43:01	R	G	R	Q	D	Q	T	E
A*66:01:01	R	G	R	Q	D	Q	A	E
A*68:01:01	R	G	M	R	D	K	A	V
A*68:02:01	R	G	R	H	D	K	A	V
A*74:01	R	G	M	Q	D	Q	A	V
A*80:01:01	R	E	I	R	D	K	A	R

Figure 4.1: The relationship between individual polymorphisms within *HLA-A* associated with differential HIV-1 control and specific *HLA-A* alleles. Positions favourably associated with HIV-1 control are shaded in green, while those associated with loss of control are shaded in red. Only the *HLA-A* alleles present in HIV-1 progressors and –controllers are represented.

Twelve SNPs within *HLA-B* were found to be present at significantly different frequencies in HIV-1 progressors and –controllers (Table 4.4). However, due to strong LD between these variants only a limited number of independent associations with differential HIV-1 control were observed. Alleles at position 70 were found to be in strong LD with those at positions 69 and 71, positions 24 and 41 were seen to be in strong LD with position 67, and alleles at position 32

were strongly associated with those at position 45. Consistent with previous reports (The International HIV Controllers Study *et al*., 2010; McLaren *et al*., 2012), Val^{97} - unique to *HLA-B*57* alleles (Figure 4.2) – was most strongly associated with HIV-1 control (p = 0.026; OR = 4.7). The Ala^{69}-Gln^{70}-Ala^{71} and Ala^{69}-Ser^{70}-Ala^{71} haplotypes were also seen to be associated with HIV-1 control, while variants at positions 9, 24, 32, 41, 45, 67, 116 and 167 were associated with loss of control. Despite strong LD between the variants identified within both *HLA-A* and *–B*, significant LD was not observed between these variants across loci (data not shown).

Table 4.4 The polymorphic positions within the *HLA-B* locus associated with differential HIV-1 control

Position	Amino Acid	HIV-1 Progressors (n = 69)	HIV-1 Controllers (n = 42)	OR (95% Confidence Interval)[1]	p-value[2]
9	Y	0.790	0.558	2.9 (1.2 - 8.3)	0.014
24	T	0.246	0.058	0.3 (0.1 - 0.7)	0.005
32	L	0.210	0.095	0.4 (0.2 - 0.9)	0.027
41	T	0.246	0.058	0.3 (0.1 - 0.7)	0.005
45	K	0.210	0.095	0.4 (0.2 - 0.9)	0.027
67	S	0.384	0.138	0.5 (0.2 - 0.9)	0.018
69	A	0.333	0.488	1.9 (1.1 - 3.3)	0.024
70	N	0.667	0.312	0.5 (0.3 - 0.9)	0.024
71	A	0.333	0.488	1.9 (1.1 - 3.3)	0.024
97	V	0.022	0.058	4.7 (1.1 - 28.3)	0.026
116	L	0.145	0.022	0.2 (0.4 - 0.8)	0.011
167	S	0.210	0.095	0.4 (0.2 - 0.9)	0.027

[1] Odds ratio

[2] p-values are for a two-sided Fisher's exact test, as implemented in R v3.1.1 (www.r-project.org).

No associations were observed at polymorphic positions within *HLA-C*, again consistent with previous studies that found genetic variation outside the peptide-binding domain of this locus to be associated with HIV-1 control (The International HIV Controllers Study *et al*., 2010; McLaren *et al*., 2012).

	9	24	32	41	45	67	69	70	71	97	116	167
B*07:02:01	Y	S	Q	A	E	Y	A	Q	A	S	Y	W
B*07:06	Y	S	Q	A	E	Y	A	Q	A	S	Y	W
B*08:01:01	D	S	Q	A	E	F	T	N	T	S	Y	W
B*13:02:01	Y	T	Q	T	M	S	T	N	T	T	L	W
B*13:03	Y	T	Q	T	M	S	T	N	T	T	L	W
B*14:01:01	Y	S	Q	A	E	C	T	N	T	W	F	W
B*14:02:01	Y	S	Q	A	E	C	T	N	T	W	F	W
B*15:01:01	Y	A	Q	A	M	S	T	N	T	R	S	W
B*15:03:01	Y	S	Q	A	E	S	T	N	T	R	S	W
B*15:10:01	Y	S	Q	A	E	C	T	N	T	R	Y	W
B*15:16:01	Y	A	Q	A	M	M	A	S	A	R	S	W
B*18:01:01	H	S	Q	A	T	S	T	N	T	R	S	W
B*35:01:01	Y	A	Q	A	T	F	T	N	T	R	S	W
B*39:10:01	Y	S	Q	A	E	Y	T	N	T	R	F	W
B*42:01:01	Y	S	Q	A	E	Y	A	Q	A	S	Y	W
B*42:02	H	S	Q	A	E	Y	A	Q	A	S	Y	W
B*44:03:01	Y	T	L	T	K	S	T	N	T	R	D	S
B*44:03:02	Y	T	L	T	K	S	T	N	T	R	D	S
B*45:01:01	H	T	L	T	K	S	T	N	T	R	L	S
B*49:01:01	H	T	L	T	K	S	T	N	T	R	L	W
B*51:01:01	Y	A	Q	A	T	F	T	N	T	T	Y	W
B*53:01:01	Y	A	Q	A	T	F	T	N	T	R	S	W
B*57:02:01	Y	A	Q	A	M	M	A	S	A	V	Y	W
B*57:03:01	Y	A	Q	A	M	M	A	S	A	V	Y	W
B*58:01:01	Y	A	Q	A	T	M	A	S	A	R	S	W
B*58:02	Y	A	Q	A	T	M	A	S	A	W	S	W
B*81:01	Y	S	Q	A	E	Y	A	Q	A	S	Y	W
B*82:02:01	Y	S	Q	A	E	Y	A	Q	A	R	L	S

Figure 4.2: The relationship between individual polymorphisms within *HLA-B* associated with differential HIV-1 control and specific *HLA-B* alleles. Positions favourably associated with HIV-1 control are shaded in green, while those associated with loss of control are shaded in red. Only the *HLA-B* alleles present in HIV-1 progressors and –controllers are represented.

4.3.3 Association of Polymorphic Positions within the HLA-C 3' UTR with HIV-1 Control

Allelic variability at an indel polymorphism at position 263 (rs67384697) within the *HLA-C* 3' UTR has previously been shown to be associated with HIV-1 control (Kulkarni *et al*., 2011). The presence of a deletion at this position is associated with increased HLA-C expression, due to disruption of an miRNA binding site (Kulkarni *et al*., 2011; Kulkarni *et al*., 2013), and increased HLA-C expression has been shown to be associated with HIV-1 viral control (Apps *et al*., 2013). We therefore, examined the association between genetic variability at this position (rs67384697) and HIV-1 control in the Black South African population, but found no difference in the frequency distribution of the protective *263D* allele between HIV-1 progressors and –controllers (Table 4.5).

We have previously reported that the *263* indel is in strong LD with another indel at position 230 of the *HLA-C* 3' UTR (rs35877659) in the Black South African population (Chapter 3). Due to the haplotypic structure present within the *HLA-C* 3' UTR, we hypothesized that genetic variability at this position may have the potential to impact HLA-C expression (and thus HIV-1 control) in a manner similar to that seen at position 263 (Chapter 3). Examination of genetic variability at this position within HIV-1 progressors and –controllers revealed that while possession of a deletion at position 230 was not found to be significantly associated with HIV-1 control in the Black South Africa, HIV-1 controllers were considerably less likely than progressors to possess two copies of the potentially deleterious *230I* allele at this position (Table 4.5).

Table 4.5 The proportion of HIV-1 progressors and –controllers in possession of specific *230* (rs35877659) and *263* (rs67384697) indel alleles and genotypes

	230 Indel					**263 Indel**				
	I	**D**	**II**	**ID**	**DD**	**I**	**D**	**II**	**ID**	**DD**
HIV-1 Progressors (n = 62)	0.468	0.806	0.210	0.258	0.548	0.758	0.565	0.435	0.323	0.242
HIV-1 Controllers (n = 37)	0.405	0.919	0.081	0.324	0.595	0.730	0.595	0.405	0.324	0.270
OR (95% CI)[1]	0.8 (0.3 - 1.9)	2.7 (0.7 - 16.0)	0.3 (0.1 - 1.4)	1.4 (0.5 - 3.7)	1.2 (0.5 - 3.0)	0.9 (0.3 - 2.5)	1.1 (0.5 - 2.8)	0.9 (0.4 - 2.2)	-	1.2 (0.4 - 3.2)
p-value[2]	0.68	0.16	0.16	0.50	0.68	0.81	0.84	0.84	1.00	0.81

[1] Odds ratio

[2] p-value is for a two-sided Fisher's exact test, as implemented in R v3.1.1 (www.r-project.org). Comaprisons are considered significant at $P < 0.05$.

4.4 Discussion

We examined genetic variability within and across the *HLA* class I loci in 111 treatment-naïve HIV-1 infected Black South Africans exhibiting differing degrees of viral control and in this way identified the residues within the *HLA-A* and *HLA-B* peptide binding grooves that strongly associate with HIV-1 control in this population group. Our findings confirm those from other populations, demonstrating the importance of HLA-B residues 67, 70, 97 and 116 in determining disease outcome (The International HIV Controllers Study *et al*., 2010; McLaren *et al*., 2012), while also identifying additional residues in HLA-A and -B that may potentially contribute to determining differential disease outcome in this population. We further demonstrate a lack of association between an indel polymorphism within the *HLA-C* 3' UTR (rs67384697) and HIV-1 control in Black South Africans, a finding consistent with previous observations in African-Americans (McLaren *et al*., 2012).

Recent evidence suggests that the effects of the most strongly protective *HLA* class I alleles may be lost in sub-Saharan African populations with high HIV-1 seroprevalence, as a consequence of viral adaptation to these alleles at the population level (Payne *et al*., 2014). Therefore, before analyzing the effects of individual SNPs within *HLA* class I loci on HIV-1 control, it was necessary to first establish whether the effects of these protective alleles were still of relevance in our cohort of HIV-1 progressors and controllers. Multivariate logistic regression analysis revealed *B*57:03:01* and *B*58:01:01* to be the alleles most strongly associated with HIV-1 control in our cohort, consistent with previous reports in this population (Kiepiela *et al*., 2004; Leslie *et al*., 2010; Tang *et al*., 2011). Both these alleles, and particularly *B*57:03:01*, have been shown to limit viral replication through induction of strong CTL responses (Kiepiela *et al*., 2004; Miura *et al*., 2009), as well as through their interaction with KIR3DL1 receptors NK cells (Martin *et al*., 2007; Kamya *et al*., 2011).

Positive associations with HIV-1 control were also observed for the *B*39:10:01|C*12:03:01* and *B*81:01:01|C*04:01:01* haplotypes, both of which have previously been shown to be protective in the Black South African population (Kiepiela *et al*., 2004; Leslie *et al*., 2010). However, the mechanisms mediating the additive protective effects associated with these haplotypes have yet to be established. *HLA-A*03:01:01*, a known ligand for KIR3DL2 (Döhring *et al*., 1996; Hansasuta *et al*., 2004), was the only allele seen to be significantly associated with a loss of viral control. Given that the interaction between A*03:01:01 and KIR3DL2 is peptide-specific (Hansasuta *et al*., 2004), presentation of viral peptides by A*03:01:01 may result in alterations in NK cell activity that cause a loss of viral control. This allele has previously been shown to be associated with reduced absolute $CD4^+$ T cell count in Black South Africans (Leslie *et al*., 2010).

Examination of individual SNPs within the *HLA* class I alleles observed within our cohort of HIV-1 progressors and controllers revealed 20 positions to be associated with differential HIV-1 control. However, strong LD between SNPs within HLA loci resulted in only a limited number of independent associations. Residues in strong LD were also generally found to occur within the same structural features, making it difficult to establish whether the associations observed were attributable to a single residue or were, in fact, the result of the additive effects of all the residues comprising the haplotypes observed.

Of the twelve positions identified within *HLA-B*, four (positions 67, 70, 97 and 116) had previously been observed to be associated with HIV-1 control in African-American (McLaren *et al*., 2012) and European-American (The International HIV Controllers Study *et al*., 2010) HIV-1 progressors and -controllers. Of the eight additional residues observed to be associated with differential HIV-1 control in our cohort, alleles at position 70 were found to be in strong LD with those at positions 69 and 71, positions 24 and 41 were seen to be in strong LD with position 67, and alleles at position 32 were strongly associated with those at position 45.

Positions 24, 45, 67 and 70 all occur within the HLA-B α1 helix, the region of the molecule most critical to determining peptide specificity (Barber *et al*., 1997). Residue 67, in particular, is known to play an important role in determining the conformation of the B-pocket. The presence of bulky, hydrophobic, aromatic amino acids at this position limits the size of the B-pocket, sterically precluding binding of peptides containing amino acids with large side chains at position P2 (Smith *et al*., 1996). A similar effect has been observed at position 45 (Badrinath *et al*., 2012). Positions 97 and 116 within the β-strand of the α2 domain of HLA-B (Barber *et al*., 1997; Sanjanwala *et al*., 2008; The International HIV Controllers Study *et al*., 2010; McLaren *et al*., 2012) and positions 149 and 152 within HLA-A α2 domain (Cowan *et al*., 1987; Hogan *et al*., 1988a; Hogan *et al*., 1988b) are also involved in modulating peptide specificity. These differences in peptide specificity are anticipated to result in differential regulation of CTL responses (Barber *et al*., 1997).

While variability at position 116 has been shown to be important for regulating interactions with the C-terminal of the bound peptide (Barber *et al*., 1997; Macdonald *et al*., 2003; Zernich *et al*., 2004), this residue has also been demonstrated to be vitally important in determining the nature of the interaction between HLA and tapasin (Rizvi *et al*., 2014). This interaction determines the dynamics of both HLA protein folding and peptide loading (Thammavongsa *et al*., 2006; Chen and Bouvier, 2007), such that more structurally stable tapasin-independent HLA allotypes may be loaded with low-affinity peptides and thus exhibit suboptimal antigen presentation to $CD8^+$ T cells (Rizvi *et al*., 2014). Position 114 has also been implicated in determining tapasin dependence (Park *et al*., 2003). Both residues 114 and 116 in HLA-A were found to be associated with HIV-1 control. Tapasin-independent allotypes have previously been linked to rapid disease progression following HIV-1 infection (Rizvi *et al*., 2014). These positions were found to be in strong LD with each other, as well as positions 17 and 56. The resulting protective Ser^{17}-Arg^{56}-Glu^{114}-His^{116} haplotype was found be unique to HLA-A30.

The differences in peptide specificity observed between *HLA* class I alleles are also known to differentially regulate the interactions of peptide-bound HLA to their cognate receptors on cell types other $CD8^+$ T cells (Hansasuta *et al*., 2004; Sanjanwala *et al*., 2008; Jones *et al*., 2011). Variability at HLA-B residues 67, 97 and 116 has been shown to modulate the binding of HLA-Bw4 molecules to KIR3DL1 (Sanjanwala *et al*., 2008). The interaction between KIR3DL1 and HLA-Bw4 is associated with increased NK responsiveness (Boulet *et al*., 2010; Kamya *et al*., 2011; Parsons *et al*., 2012) and has been shown to be strongly associated with HIV-1 control (López-Vázquez *et al*., 2005; Martin *et al*., 2007; Kamya *et al*., 2011).

Variability within the classical HLA class I molecules has also been found to differentially regulate their interaction with leukocyte immunoglobulin-like receptors (LILRs) (Jones *et al*., 2011). Position 9 within all three classical HLA class I molecules has been shown to modulate their binding to LILRB2, in the absence of β_2-microglobulin and peptide, as has position 144 within HLA-A (Jones *et al*., 2011). High avidity binding of LILRB2 to HLA-B*35-Px molecules is thought to impede dendritic cell function and has been associated with more rapid HIV-1 disease progression (Huang *et al*., 2009). Conversely, *HLA-B* alleles with reduced LILRB2 binding affinities are associated with HIV-1 control (Bashirova *et al*., 2014).

Although KIR (Moesta and Parham, 2012) and LILR (Jones *et al*., 2011) are both known to bind HLA-C in a peptide-specific manner, an association was not observed between HIV-1 control and any of the residues examined within this molecule. This was consistent with results observed in African-Americans, where only variation occurring within the HLA-C α3 domain was shown to be associated with viral control (McLaren *et al*., 2012). The influence of genetic variability within *HLA-C* on HIV-1 control has also been attributed to an indel polymorphism at position 263 of the HLA 3' UTR (rs67384697). Variability at this position has been shown to impact HLA-C expression by differentially regulating binding of specific miRNA to the *HLA-C* 3' UTR (Kulkarni *et al*., 2011; Kulkarni *et al*., 2013). Increased HLA-C expression has been

seen to be associated with HIV-1 control (Apps *et al*., 2013), and the deletion allele at this indel has been reported to be over-represented in HIV-1 controllers of European descent (Kulkarni *et al*., 2011).

We, however, did not observe an association between possession of the *263D* allele and HIV-1 control, consistent with observations in African-American HIV-1 controllers (McLaren *et al*., 2012). We did however note an increased frequency of the deletion allele at position 230 (rs35877659) within the *HLA-C* 3' UTR in HIV-1 controllers relative to progressors. We have previously hypothesized that due to strong LD between the two indels, variability at this position may also impact HLA-C expression through regulation of miRNA-binding, most dramatically when found in combination with the corresponding deletion allele at position 263 (Chapter 3). Although this comparison did not reach statistical significance, we cannot exclude the possibility that variability at this position may influence HLA-C expression and thus HIV-1 control. Our study may simply lack sufficient power to statistically validate this association, which may influence HIV-1 control in a manner more subtle than the associations observed in *HLA-A* and *HLA-B*.

Collectively these data provide the first description of the impact of genetic variability at individual polymorphisms with the classical *HLA* class I loci on HIV-1 control in the Black South African population. These findings highlight the importance of HLA peptide specificity in determining immune responses to HIV-1 infection and suggest variability at key residues within the HLA molecule determine both the sequence of the peptide bound, as well as the avidity of the peptide-HLA interaction. The nature of these peptide-HLA interactions are likely to differentially impact the function of multiple immune cell types and thus have implications for future vaccine studies. Finally, we demonstrate that while the impact of HLA-C on HIV-1 control is less subtle than that of HLA-A and HLA-B, the importance of this molecule in understanding antiviral immune responses should not be overlooked.

Chapter 5

Genetic Variability within the KIR Gene Complex is Associated with HIV-1 Control and Distinguishes Elite Controllers from Viraemic Controllers

5.1 Introduction

The rate of HIV-1 disease progression is known to be highly variable between individuals. HIV-1 controllers are a small subset of infected individuals capable of spontaneously suppressing viral replication and maintaining very low levels of viraemia (< 2000 RNA copies/ml) in the absence of highly active antiretroviral therapy. This ability to suppress viral replication has been shown to be associated with a number of both viral and host genetic characteristics, suggesting control of HIV-1 viral replication can be achieved through multiple immunological mechanisms (Deeks and Walker, 2007; Walker and Yu, 2013).

The argument for the existence of multiple mechanisms of HIV-1 control is further supported by the observation that a small percentage of HIV-1 controllers, known as elite controllers, can be further distinguished from other controllers by their ability to suppress viral replication to levels undetectable by currently available commercial assays (<50 RNA copies/ml; Deeks and Walker, 2007). Whether elite controllers employ entirely different immune mechanisms of viral control, or simply more effectively utilize the same mechanisms as other HIV-1 controllers, remains unclear. Therefore, identification and characterization of the immune signatures that distinguish elite controllers from other HIV-1 controllers is necessary.

Genetic variability within the *HLA* gene complex has been identified as one of the major determinants of HIV-1 control (Fellay *et al*., 2007; The International HIV Controllers Study, 2010). *HLA* class I alleles function as ligands for both the TCR found on the surface of $CD8^+$ T cells, and specific KIR expressed predominantly on the surface of NK cells. While the impact of

HLA-restricted CTL responses on HIV-1 control has been well established (Migueles *et al*., 2003; Kiepiela *et al*., 2007; Sáez-Cirión *et al*., 2007), the role of NK cell-mediated immunity in modulating viral control during chronic HIV-1 infection is only just beginning to be understood (Alter *et al*., 2011).

KIR receptors are encoded by a family of highly polymorphic genes located on chromosome 19 (Wilson *et al*., 2000; Martin *et al*., 2000). Diversity in terms of *KIR* gene content is determined by two *KIR* haplotypes (Uhrberg *et al*., 2002; Martin *et al*., 2004), which are themselves the product of varying combinations of centromeric and telomeric gene content motifs (Pyo *et al*., 2010; Jiang *et al*., 2012; Vierra-Green *et al*., 2012; Pyo *et al*., 2013). Genetic variability within *KIR* loci has previously been shown to be associated with both HIV-1 transmission (Jennes *et al*., 2006; Merino *et al*., 2011; Paximadis *et al*., 2011; Hong *et al*., 2013) and disease progression (Martin *et al*., 2002a; Martin *et al*., 2007; Jennes *et al*., 2011). We therefore examined patterns of genetic variability within the *KIR* and *HLA* gene complexes in 111 treatment-naïve HIV-1 infected Black South Africans exhibiting differing degrees of viral control. These data reveal an association between common *KIR* haplotypes and HIV-1 control in a population heavily affected by the HIV/AIDS pandemic.

5.2 Materials and Methods

5.2.1 Study Populations

A total of 111 treatment-naïve HIV-1 infected Black South Africans were included in this study in order to investigate the impact of genetic variability within the *KIR* gene complex on control of HIV-1 infection. Sixty-nine of these individuals were classified as HIV-1 progressors based on a $CD4^+$ T cell count that declined to below 300 cells/mm^3 prior to the administration of ART. As decisions to commence ART are based solely on $CD4^+$ T cell count, which may then include HIV-1 controllers (viraemic or elite) with declining counts, we included the criterion of high

viraemia (i.e. viral loads greater than 10 000 RNA copies/ml) to identify progressors (Table 4.1).

The remaining 42 HIV-1 positive individuals were classified as HIV-1 controllers on the basis of their having $CD4^+$ T counts greater than 300 cells/mm^3, while also suppressing viral replication to levels below 2000 RNA copies/ml. Eleven of these HIV-1 controllers could be categorized as elite controllers as a consequence of their ability to maintain viral loads below 50 RNA copies/ml - based on counts taken on two or more occasions, recorded at least 12 months apart. The remaining 31 HIV-1 controllers exhibited viral loads between 50 and 2000 RNA copies/ml and were, for the purposes of this study, regarded as viraemic controllers.

HIV-1 RNA levels were quantitated using either the Roche Amplicor RNA Monitor assay (lower detection limit of 400 HIV-1 RNA copies/ml) or, where necessary, the COBAS AmpliPrep/COBAS Taqman HIV-1 Test, v2.0 (Roche Molecular Diagnostics, Pleasanton, California, USA), which has a lower detection limit of 50 RNA copies/ml. $CD4^+$ T cell counts were determined using the FACSCount System (BD Biosciences, San Jose, California, USA). Informed consent was obtained from all study participants and the study was approved by the University of the Witwatersrand Committee for Research on Human Subjects (Appendix B).

5.2.2 DNA Extraction

The DNA obtained from HIV-1 controllers was extracted from whole blood samples, while that obtained from HIV-1 progressors was derived from frozen buffy coat samples. DNA extractions from all HIV-1 infected samples were performed by silica spin-column purification using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany).

5.2.3 KIR Genotyping

Individuals were genotyped for the presence or absence of 16 *KIR* genes using a real-time PCR assay (Hong *et al*., 2011), as previously described (Chapter 2, Section 2.2.2). Additional primers (Jiang *et al*., 2012) were included to distinguish between *KIR2DS4* alleles encoding the full length form of the receptor (*KIR2DS4f*) and alleles encoding a deleted form of the gene (*KIR2DS4v*). All individuals were assigned either AA or Bx *KIR* genotypes, according to the guidelines and nomenclature stipulated by the Allele Frequency Net Database (Gonzalez-Galarza *et al*., 2011); and individual *KIR* haplotypes were inferred using HAPLO-IHP (Yoo *et al*., 2007), as previously described (Chapter 2, Section 2.2.4).

5.2.4 KIR Copy Number Determination

KIR3DL1 and *KIR3DS1* copy number was determined using a quantitative real-time PCR assay, as described by Pelak *et al*. (2011). Standard curves were derived from serial dilution of genomic DNA obtained from the CEPH cell lines GM11840 and GM12752 (which are known to possess single copies of *KIR3DL1* and *KIR3DS1*), with concentrations ranging between 20 and 0.625 ng/µl. Primers specific for the dual copy housekeeping gene, *beta-globin* (*BGB*; Shostakovich-Koretskaya *et al*., 2009), were included as an internal control. *KIR2DS4f* and *KIR2DS4v* copy number was similarly determined using the primers and probes of Jiang *et al*. (2012).

For both assays, reactions were performed in a 5 µl volume, containing 2x Lightcycler 480 Probes Master (Roche, Mannheim, Germany), 0.5 µM of each primer, 0.1µM of VIC®-labelled gene-specific probe, 0.1µM of 6-carboxyfluorescein (FAM)-labelled BGB-specific probe (Life Technologies, Carlsbad, California, USA) and 5 ng of DNA. Both assays was performed using the Applied Biosystems 7500 Real-Time PCR system (Applied Biosystems, Foster City, California, USA), under the following thermocycling conditions: 95°C for 10 minutes, followed

by 40 cycles of 95°C for 15 seconds and 60°C for 60 seconds. Reactions were performed in duplicate for all samples, or triplicate, if the initial results were discordant.

5.2.5 HLA Class I Ligand Determination

HLA-A, *-B* and *-C* genotype data were available for all 111 individuals included in this study (Paximadis *et al*., 2012; Chapter 3, Section 3.2.2; Chapter 4, Section 4.2.3). HLA-A and -B Bw4 isotopes within alleles were identified based on the amino acid residues present at positions 77, 80, 81, 82 and 83 of the protein sequence (Martin *et al*., 2002a). A further distinction was made between *HLA-B* alleles with isoleucine at amino acid position 80 (*Bw4-80I*) and those with threonine at this position (*Bw4-80T*), as these epitopes differ in the avidity of their interaction with KIR3DL1 (Cella *et al*., 1994; Martin *et al*., 2002a). *HLA-C* alleles were similarly classified as having either HLA-C1 or –C2 epitopes on the basis of the presence of asparagine or lysine at position 80, respectively (Colonna *et al*., 1993; Winter and Long, 1997).

5.2.6 Statistical Analyses

KIR gene carrier- and *HLA* class I ligand frequencies were determined by direct counting. The significance of differences in the frequencies observed between groups was assessed using a two-sided Fisher's exact test and the associated odds ratios were estimated. Odds ratios greater than 1 indicate a positive association with HIV-1 control. Differences in gene copy number between groups were assessed using a non-parametric Mann-Whitney U-test. All statistical analyses were conducted using R v3.1.1 (www.r-project.org) and all measures were considered significant at $p < 0.05$.

5.3 Results

5.3.1 KIR Gene Diversity

In order to investigate the association between genetic variability within the *KIR* gene complex and HIV-1 control, we examined gene diversity at this locus in HIV-1 progressors, viraemic controllers and elite controllers (Figure 5.1). HIV-1 progressors were found to differ from viraemic controllers in that individuals within this group were less likely to be in possession of *KIR2DL5* ($p = 0.036$; OR = 3.3; CI = 1.1 – 9.8) and *KIR2DS5* ($p = 0.031$; OR = 2.7; CI = 1.1 – 6.7), but displayed *KIR2DL3* frequencies greater than that seen in the viraemic controller group ($p = 0.046$; OR = 0.4; CI = 0.1 – 0.9). Viraemic controllers, in turn, were distinguished from elite controllers, by displaying a greater frequency of *KIR2DS5* ($p = 0.033$; OR = 0.2; CI = 0.0 – 0.8) and a decreased likelihood of being in possession of *KIR2DL3* ($p = 0.018$; OR = ∞). HIV-1 progressors were also found to possess significantly fewer copies of both *KIR2DS4* ($p = 0.035$) and *KIR3DL1* ($p = 0.029$) per individual than HIV-1 controllers (i.e. individuals with viral loads below 2000 RNA copies/ml).

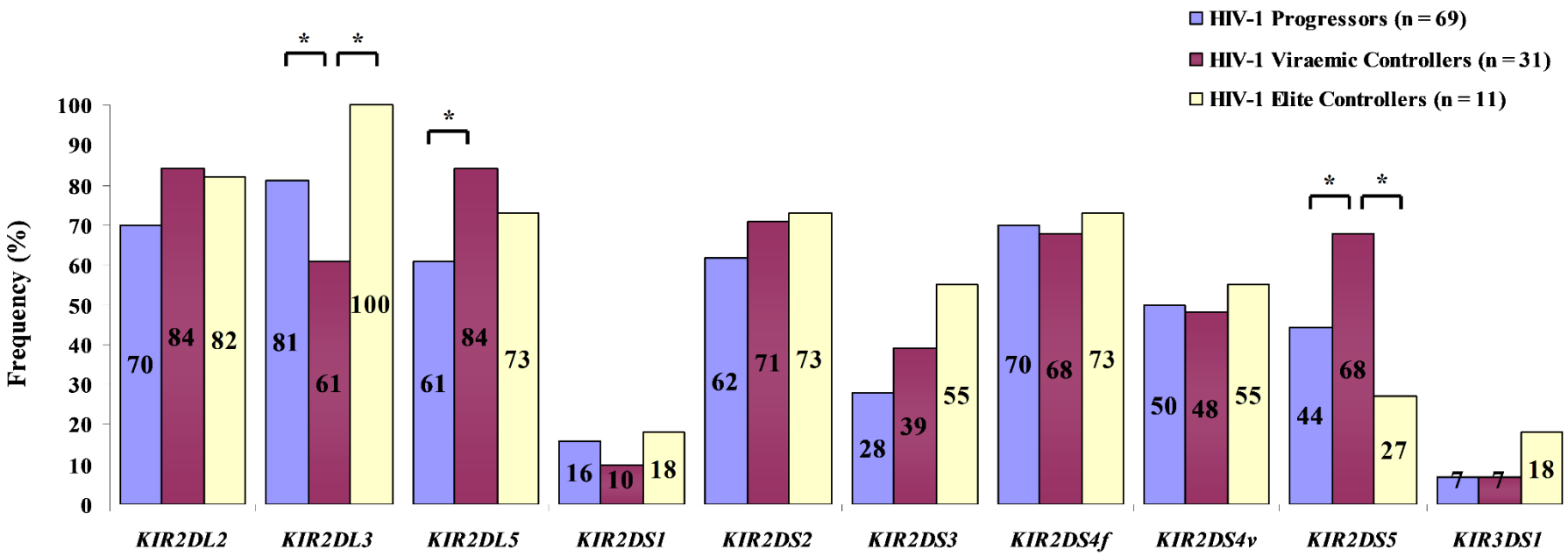


Figure 5.1: The gene carrier frequencies of the *KIR* genes present within HIV-1 progressors, viraemic controllers and elite controllers. The four framework KIR genes (*KIR2DL4*, *KIR3DL2*, *KIR3DL3* and *KIR3DP1*) were present in all individuals and are thus not shown. *KIR2DL1*, *KIR2DP1* and *KIR3DL1* were also present in more than 98% of individuals in all groups considered and their frequencies are thus also not shown. The frequencies of *KIR2DS4* are reported for its two isoforms, the cell surface receptor, KIR2DS4f and the secreted receptor, KIR2DS4v. Significant comparisons ($p < 0.05$), as measured by a two-sided Fisher's exact test, are indicted with an asterisk.

5.3.2 KIR Genotype Distributions

The differences in gene frequencies seen between HIV-1 progressors and -controllers (i.e. all individuals with viral loads below 2000 RNA copies/ml) translated into noticeable differences in their *KIR* genotype profiles (Table 5.2), with the Bx 5 and Bx 112 genotypes being observed more frequently in HIV-1 controllers than progressors. While these comparisons did not reach statistical significance, it is worth noting that the frequencies of the Bx 5 ($p = 0.032$) and Bx 112 ($p = 0.010$) genotypes observed in HIV-1 controllers were significantly higher than those generally observed in the Black South African population (Chapter 2, Table 2.2). Furthermore, the Bx 5 genotype was found to be especially over-represented within elite controllers, while the Bx 112 was found to be absent from this group all together (Figure 5.2). These two observations were indicative of a pattern that emerged between the two groups, whereby genotypes containing *KIR2DS5* were more commonly observed in HIV-1 viraemic controllers than in elite controllers, who most often possessed *KIR* genotypes containing *KIR2DS3*.

5.3.3 Estimation of KIR Haplotypes

Further analysis of the *KIR* genotypes present in Black South African HIV-1 progressors and controllers revealed they were comprised of a total of 23 *KIR* haplotypes (Table 5.3), which were themselves the result of varying combinations of three centromeric and two telomeric haplotype motifs (Pyo *et al*., 2010; Pyo *et al*., 2013). The canonical cA01|tA01 haplotype was found to be the most frequently observed haplotype within both HIV-1 progressors and elite controllers, while viraemic controllers were distinguished from progressors ($p = 3.46 \times 10^{-4}$; OR = 3.7; CI = 1.9 – 7.5) and elite controllers ($p = 0.035$; OR = 0.3; CI = 0.1 – 0.9) by an unusually high frequency of the archetypal cB01|tA01 haplotype (Table 5.3).

Table 5.1 The frequencies of the *KIR* genotype profiles present within HIV-1 progressors and -controllers

KIR Genotype	KIR2DL4	KIR3DL2	KIR3DL3	KIR3DP1	KIR2DP1	KIR2DL1	KIR2DL3	KIR2DS4	KIR3DL1	KIR2DL2	KIR2DL5	KIR2DS1	KIR2DS2	KIR2DS3	KIR2DS5	KIR3DS1	HIV-1 Progressors (n = 69)	HIV-1 Viremic Controllers (n = 31)	HIV-1 Elite Controllers (n = 11)
AA 1	■	■	■	■	■	■	■	■	■								0.290 (20)	0.161 (5)	0.182 (2)
Bx 21	■	■	■	■	■	■	■	■	■	■	■		■		■		0.130 (9)	0.194 (6)	-
Bx 5	■	■	■	■	■	■	■	■	■	■	■		■	■			0.116 (8)	0.129 (4)	0.455 (5)
Bx 4	■	■	■	■	■	■	■	■	■	■			■				0.087 (6)	-	0.091 (1)
Bx 20	■	■	■	■	■	■	■	■	■	■	■				■		0.072 (5)	0.097 (3)	0.091 (1)
Bx 9	■	■	■	■	■	■	■	■	■	■	■	■	■		■		0.058 (4)	-	-
Bx 71	■	■	■	■	■	■		■	■	■	■		■	■			0.058 (4)	-	-
Bx 112	■	■	■	■	■	■		■	■	■	■		■	■	■		0.058 (4)	0.194 (6)	-
Bx 6	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	0.029 (2)	-	0.091 (1)
Bx 92	■	■	■	■	■	■		■	■	■	■	■	■		■		0.029 (2)	-	-
Bx 2	■	■	■	■	■	■	■	■	■		■	■			■	■	0.014 (1)	-	-
Bx 3	■	■	■	■	■	■	■	■	■	■	■	■	■		■	■	0.014 (1)		0.091 (1)
Bx 73	■	■	■	■	■	■		■	■	■	■	■	■	■	■	■	0.014 (1)	0.065 (2)	-
Bx 172	■	■	■	■		■		■	■	■			■				0.014 (1)	-	-
Bx 228	■	■	■	■	■	■		■	■	■	■		■		■		0.014 (1)	0.097 (3)	-
Bx 38	■	■	■	■	■	■	■	■	■	■	■						-	0.032 (1)	-
Unknown	■	■	■	■					■	■	■	■	■		■		-	0.032 (1)	-

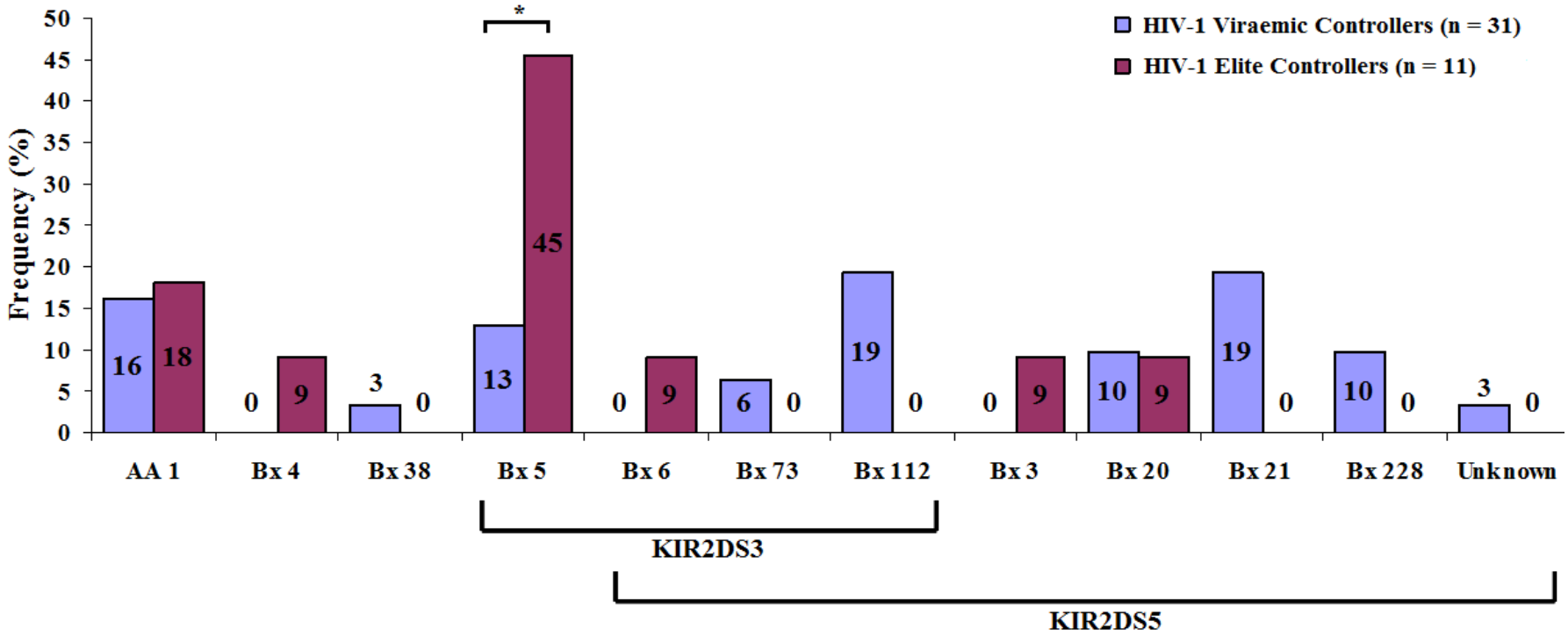


Figure 5.2: The *KIR* genotypes present in HIV-1 viraemic- and elite controllers and their frequencies. The presence of *KIR2DS3* and *KIR2DS5* within genotypes is highlighted. Significantly differences in frequency ($p < 0.05$) between the two groups, as measured by a two-sided Fisher's exact test, are indicated with an asterisk.

Table 5.2 The frequencies of the *KIR* haplotypes occurring within HIV-1 progressors, -viraemic controllers and -elite controllers

	HIV-1 Progressors (n = 69)		HIV-1 Viraemic Controllers (n = 31)		HIV-1 Elite Controllers (n = 11)	
	n	**Frequency**	**n**	**Frequency**	**n**	**Frequency**
cA01|tA01[1]	52	0.377	20	0.323	10	0.455
cA01|tA01-hybd1[2]	13	0.094	2	0.032	2	0.091
cA01|tA01-del5	11	0.080	1	0.016	1	0.045
cA01/cB03|tB01[3]	1	0.007	0	0.000	0	0.000
cB01|tA01	20	0.145	24	0.387	3	0.136
cB01|tA01-hybd1	2	0.014	2	0.032	1	0.045
cB01|tA01-del3	4	0.029	5	0.081	1	0.045
cB01|tA01-del5	11	0.080	2	0.032	1	0.045
cB01|tA01-del9	8	0.058	0	0.000	1	0.045
cB01|tA01-del3-del5	1	0.007	0	0.000	0	0.000
cB01|tA01-del3-del5-del9	0	0.000	1	0.016	0	0.000
cB01|tB01	4	0.029	2	0.032	1	0.045
cB01|tB01-del7	6	0.043	1	0.016	0	0.000
cB02|tA01-hybd1	5	0.036	1	0.016	0	0.000
cB02|tB01	0	0.000	0	0.000	1	0.045
H2[4]	0	0.000	1	0.016	0	0.000

[1] Major haplotypes comprised of the classic centromeric and telomeric *KIR* motifs are indicated in grey.

[2] Minor haplotypes including gene insertions, deletions, duplications and hybrid genes are named in accordance with the nomenclature of Pyo *et al*. (2013).

[3] cA01|tB01 and cB03|tB01 could not be distinguished from each other and are reported as cA01/cB03|tB01.

[4] Haplotypes that could not be named according to the aforementioned nomenclature were assigned arbitrary designations based on their frequencies within the Black South African population (Table 2.2).

5.3.4 Examination of KIR Haplotype Motifs

The high frequencies of the cA01|tA01 and cB01|tA01 haplotypes observed in all three groups were found to be a reflection of the high frequencies of the cA01 and cB01 centromeric haplotype motifs (Pyo *et al*., 2010; Pyo *et al*., 2013) within these groups. Collectively, these two motifs accounted for the centromeric component of more than 95% of the haplotypes occurring in all three groups. However, while the cA01 motif was the centromeric motif most frequently observed in HIV-1 progressors and -elite controllers, the cB01 motif was the centromeric motif seen at the highest frequency in viraemic controllers (Figure 5.3).

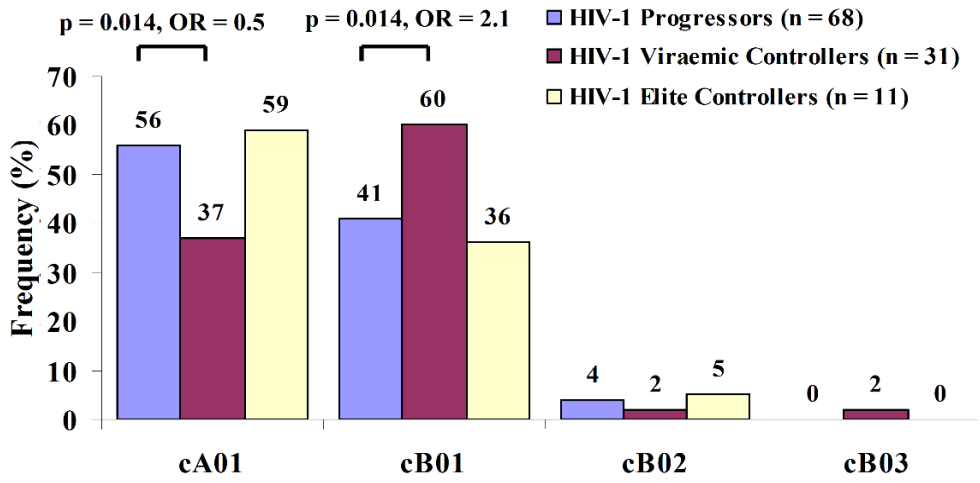


Figure 5.3: The frequencies of the centromeric *KIR* haplotype motifs observed within HIV-1 progressors, viraemic controllers and elite controllers. The cA01/cB01ltB01 haplotype from one HIV-1 progressor has been omitted from the analyses. The odds ratios and p-values associated with significant differences in frequency between groups ($p < 0.05$), as measured by a two-sided Fisher's exact test, are indicated.

The canonical cB01 centromeric *KIR* haplotype motif may contain either *KIR2DS3* or *KIR2DS5* (Pyo *et al.*, 2010). A truncated motif lacking either of these genes has also been identified (cB01-del9; Pyo *et al.*, 2013). Given the variability in *KIR2DS3* and *KIR2DS5* frequencies displayed across the groups included in this study, we examined the distribution of cB01 gene content variants within the three groups. Elite controllers were distinguished from both progressors ($p = 0.048$; OR = 5.8; CI = 1.1 – 31.8) and viraemic controllers ($p = 0.055$; OR = 5.5; CI = 1.0 – 31.5) by their higher frequency of cB01 motifs containing *KIR2DS3* rather than *KIR2DS5*. Conversely, cB01 motifs containing *KIR2DS5* were more commonly observed within viraemic controllers than elite controllers ($p = 0.007$; OR = 0.1; CI = 0.0 – 0.7).

Elite controllers could also be distinguished from the other two groups by their high degree of heterozygosity with respect to centromeric *KIR* haplotype motifs, as more than 80% of elite

controllers were found to possess both cA and cB haplotype motifs (Figure 5.4). Viraemic controllers, on the other hand, were most likely to possess two copies of a cB haplotype motif, a characteristic not seen in any of the elite controllers.

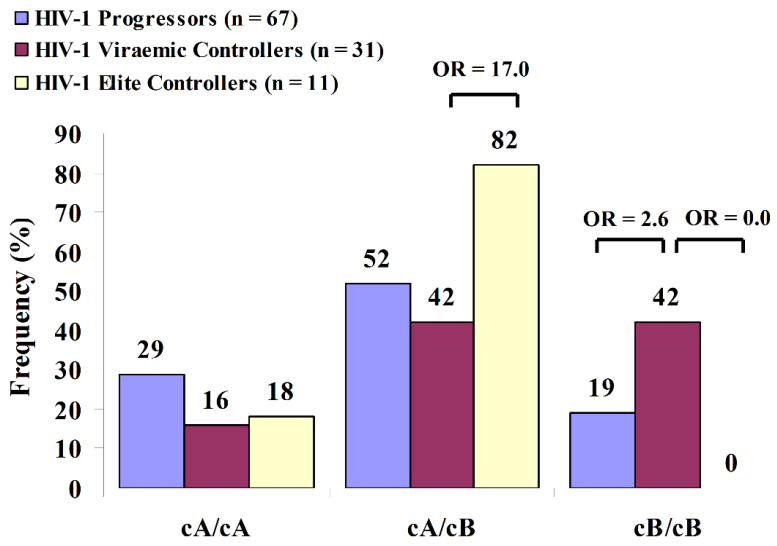


Figure 5.4: The frequencies of the centromeric KIR haplotype motif combinations observed in HIV-1 progressors, viraemic controllers and elite controllers. The odds ratios associated with significant differences in frequency between groups ($p < 0.05$), as measured by a two-sided Fisher's exact test, are indicated.

In contrast with centromeric B motifs, the telomeric tB01 motif was present at similarly low frequencies in HIV-1 progressors and -controllers (data not shown); with the result that the tA01 motif was present in more than 90% of haplotypes in all three groups. Telomeric tA01 motifs lacking *KIR3DL1* (tA01-del5) were found to occur at a significantly higher frequency in progressors than HIV-1 controllers ($p = 0.046$; OR = 0.4; CI = 0.1 – 1.0). tA01 motifs lacking any form of *KIR2DS4* (tA01-hybd1) were also found to occur at a greater frequency in progressors than controllers, but this comparison did not reach statistical significance.

5.3.5 Association between Specific KIR and their HLA Class I Ligands

5.3.5.1 HLA-B and KIR3DL1

KIR receptors are involved in the regulation of NK cell function, predominantly through their interaction with HLA class I molecules (Lanier, 1998). One of these receptors is KIR3DL1, which is known to bind the HLA-Bw4 epitopes present within a subset of both HLA-A (Thananchai *et al.*, 2007, Stern *et al*., 2008) and HLA-B (Cella *et al*., 1994; Gumperz *et al*., 1995, Vivian *et al*., 2011) molecules. Interactions between KIR3DL1 (Martin *et al*., 2007) - and the activating form of this receptor, KIR3DS1 (Martin *et al*., 2002a) - and Bw4 have previously been shown to be associated with control of HIV-1 infection. However, *KIR3DS1* was found to occur at very low frequencies in all the groups included in the current study (Figure 5.1) - an observation consistent with previous studies conducted in other African populations (Norman *et al*., 2007; Single *et al*., 2007). Given our sample size, this made it difficult to realistically assess the impact of *KIR3DS1* co-occurrence with Bw4 on HIV-1 control in the Black South African population.

We therefore turned our attention to the co-occurrence of *KIR3DL1* and the *HLA-A* and *–B* alleles encoding its Bw4 ligand in HIV-1 progressors and controllers. *KIR3DL1* was generally found to occur in combination with *Bw4* at lower frequencies in HIV-1 controllers than in progressors. However, when a distinction was made between alleles encoding Bw4 epitopes with isoleucine at position 80 (*Bw4-80I*) and those with threonine at position 80 (*Bw4-80T*), it became clear that *Bw4-80T* occurs in combination with *KIR3DL1* at higher frequencies in progressors and viraemic controllers than in elite controllers (Table 5.3). *Bw4-80I*, on the other hand was found to occur in combination with *KIR3DL1* at higher frequencies in progressors than in HIV-1 controllers. This was particularly true of the Bw4-80I epitopes present in *HLA-A* alleles, which were not found to occur in elite controllers (Table 5.3). However, despite these

seemingly clear distinctions, none of these comparisons were found to reach statistical significance.

Table 5.3 The frequency with which *KIR3DL1* occurs in combination with *HLA-A* and *HLA-B* alleles encoding its Bw4 ligand in HIV-1 progressors, viraemic controllers and elite controllers

	HIV-1 Progressors (n = 69)		HIV-1 Viraemic Controllers (n = 31)[1]		HIV-1 Elite Controllers (n = 11)	
	n	**Frequency**	**n**	**Frequency**	**n**	**Frequency**
KIR3DL1 + Bw4	50	0.725	19	0.679	7	0.636
KIR3DL1 + HLA-A Bw4	20	0.290	4	0.174	0	0.000
KIR3DL1 + HLA-B Bw4	41	0.594	17	0.548	7	0.636
KIR3DL1 + Bw4-80I	43	0.623	15	0.536	5	0.455
KIR3DL1 + HLA-A Bw4-80I	20	0.290	4	0.174	0	0.000
KIR3DL1 + HLA-B Bw4-80I	27	0.391	15	0.536	5	0.455
KIR3DL1 + Bw4-80T	19	0.275	7	0.226	2	0.182

[1] Only 23 viraemic controllers were genotyped at the *HLA-A* locus.

5.3.5.2 HLA-C and its Associated KIR Receptors

The HLA-C1 and -C2 epitopes present in virtually all *HLA-C* alleles also serve as ligands for the KIR receptors, KIR2DL2, KIR2DL3, KIR2DL1 and KIR2DS1 (Colonna *et al*., 1993; Winter and Long, 1997; Winter *et al*., 1998, Moesta *et al*., 2010). We therefore examined the co-occurrence of these *KIR* genes and the *HLA-C* alleles encoding the C1 and C2 epitopes in HIV-1 progressors and controllers (Table 5.4). The progressor group was found to differ from viraemic controllers in having a higher frequency of *KIR2DL3* in combination with *HLA-C1*, while viraemic controllers in turn, differed from elite controllers by having a lower frequency of *KIR2DL3* in combination with *HLA-C2*; although neither of these comparisons were found to reach statistical significance.

Table 5.4 The frequencies with which *KIR2DL1*, *KIR2DL2*, *KIR2DL3* and *KIR2DS1* occur in combination with *HLA-C* alleles encoding the C1 and C2 epitopes in HIV-1 progressors, -viraemic controllers and -elite controllers

	HIV-1 Progressors (n = 69)		HIV-1 Viraemic Controllers (n = 31)		HIV-1 Elite Controllers (n = 11)	
	n	**Frequency**	**n**	**Frequency**	**n**	**Frequency**
HLA-C1						
KIR2DL1 + C1	49	0.710	21	0.677	6	0.545
KIR2DL2 + C1	33	0.478	17	0.548	4	0.364
KIR2DL3 + C1	40	0.580	12	0.387	6	0.545
KIR2DS1 + C1	8	0.116	2	0.065	1	0.091
HLA-C2						
KIR2DL1 + C2	56	0.812	23	0.742	9	0.818
KIR2DL2 + C2	38	0.551	20	0.645	9	0.818
KIR2DL3 + C2	45	0.652	15	0.484	9	0.818
KIR2DS1 + C2	10	0.145	3	0.097	2	0.182

5.4 Discussion

Epidemiological studies have previously shown the presence of specific KIR receptors to be associated with HIV-1 control; both in conjunction with and independently of their HLA encoded ligands (Martin *et al*., 2002a; Martin *et al*., 2007; Jennes *et al*., 2011; Pelak *et al*., 2011). Further functional studies (Alter *et al*., 2007; Merino *et al*., 2014) have since supported these findings by demonstrating the impact of these receptors on NK cell function, in the context of the immune dysregulation brought about as a consequence of HIV-1 infection (Wong *et al*., 2010). In light of these observations, we sought to examine the association between the presence of specific KIR receptors, and their HLA class I ligands, and control of HIV-1 infection in the Black South African population - one of the populations most heavily impacted by the HIV/AIDS pandemic.

In order to do this, we examined genetic variability within the *KIR* and *HLA* gene complexes in 111 HIV-1 infected, treatment-naïve, Black South Africans with differing levels of viral control.

This cohort was comprised of 69 HIV-1 progressors and 42 HIV-1 controllers, 11 of whom were elite controllers, able to suppress viral replication to levels undetectable by commercially available assays. Our analysis revealed these groups could be clearly distinguished from each other based on a number of differences in *KIR* gene content and copy number. HIV-1 progressors were most notably distinguished from HIV-1 controllers by having significantly fewer copies of *KIR2DS4* and *KIR3DL1*. These differences could be attributed to the increased frequency of truncated telomeric tA01 *KIR* haplotype motifs lacking either *KIR2DS4* (tA01-hybd1) or *KIR3DL1* (tA01-del5) in this group.

While an association between *KIR2DS4* copy number and HIV-1 control has not previously been reported, this gene has been shown to be associated with mother-to-child transmission of HIV-1 in the Black South African population (Hong *et al.*, 2013). *KIR2DS4*, and in particular allelic variants encoding receptors capable of cell surface expression, have also been shown to be associated with increased viral loads and likelihood of HIV-1 transmission in heterosexual Zambian couples (Merino *et al.*, 2011). While the precise mechanism responsible for this observed association has yet to be ascertained, it has been suggested that this receptor may promote HIV-1 pathogenesis by contributing to the maintenance of a pro-inflammatory state during chronic infection (Merino *et al.*, 2014). Our data, however, conflict with these previous observations in that we find a significant underrepresentation of *KIR2DS4* in HIV-1 progressors that suggests the lack of *KIR2DS4* to be associated with unfavorable HIV-1 disease outcomes. These conflicting findings would suggest that further investigation of the role of KIR2DS4 in both HIV-1 transmission and disease progression is warranted.

In contrast with *KIR2DS4*, the association between copy number variation across the *KIR3DL1* locus and HIV-1 control has previously been examined in individuals of European descent - where an inverse correlation between HIV-1 set point and increased *KIR3DL1* copy number was observed in individuals who were also in possession of *KIR3DS1* and a corresponding HLA-

Bw4 ligand (Pelak *et al*., 2011). We similarly report a positive correlation between lower *KIR3DL1* copy number and an inability to suppress HIV-1 viral replication. However, we found this association to be independent of the presence of either *KIR3DS1* or the HLA Bw4 epitope.

Unlike previous reports (Martin *et al*., 2002a; Long *et al*., 2008; Jiang *et al*., 2013), we did not observe an association between the presence of *KIR3DS1* and HIV-1 control; and due to the low frequency of this gene in the Black South African population, our study lacked the power to adequately assess the impact of *KIR3DS1* co-occurrence with Bw4 on HIV-1 control. While we cannot rule out a potential role for KIR3DS1 in determining HIV-1 disease outcome, particularly in light of the findings of studies conducted in other populations (Martin *et al*., 2002a; Long *et al*., 2008; Jiang *et al*., 2013), it could be suggested that, at least at the greater population level, the impact of this receptor on HIV-1 control in the Black South African population may not be as profound as that seen in other non-African populations.

HIV-1 controllers could also be distinguished on the basis of their *KIR* gene content, as they were found to possess *KIR2DL2*, *KIR2DL5* and *KIR2DS3* more frequently than HIV-1 progressors. Viraemic- and elite controllers could be further distinguished by their differing frequencies of *KIR2DL3* and *KIR2DS5*. These differences in gene content could be attributed to the increased frequency of the cB01|tA01 *KIR* haplotype seen in HIV-1 controllers. But while, viraemic controllers in possession of these haplotypes most frequently carried haplotypes with cB01 motifs containing KIR2DS5, elite controllers were most often found to carry cB01 motifs containing KIR2DS3.

The differing gene content of the cB01 haplotype motifs observed in viraemic- and elite controllers could also account for the differences in *KIR* genotype frequencies observed between these two groups; where viraemic controllers displayed a high frequency of Bx 112, elite controllers predominantly possessed the Bx 5 genotype. The presence/absence *KIR* gene profile

associated wit the Bx 5 genotype can most easily be achieved as a result of having a *KIR2DS3*-containing cB01|tA01 haplotype in combination with cA01|tA01 haplotype, while the Bx 112 genotype is easily derived from the combination of *KIR2DS3*- and *KIR2DS5*-containing cB01|tA01 haplotypes.

While these observations point to a clear association between possession of cB01 (particularly *KIR2DS3*-containing) haplotype motifs and HIV-1 control, the functional rationale for this correlation is less apparent. Despite KIR2DS3 and KIR2DS5 both being lineage III KIR, and thus structurally most closely resembling those KIR that recognize HLA-C (Moesta and Parham, 2012), as seen with KIR2DS2, there is no functional evidence to support recognition of HLA-C by either of these receptors (Chiesa *et al*., 2008). Both receptors are also known to display markedly reduced levels of surface expression relative to the HLA-C-recognizing receptor, KIR2DS1 (Steiner *et al*., 2008; VandenBussche *et al*., 2008; Steiner *et al*., 2014). However, more highly expressed KIR2DS5 variants have been shown to be capable of triggering both NK cytotoxicity and cytokine production (Chiesa *et al*., 2008; Steiner *et al*., 2014), which could indicate that like KIR2DS4 (Katz *et al*., 2004), KIR2DS5 may engage with as yet unidentified non-HLA ligands in order to activate NK function.

Alternatively, the presence of *KIR2DS3* or *KIR2DS5* may act as markers for specific allelic variants of the inhibitory receptors KIR2DL1 and/or KIR2DL2, both of which are known to regulate NK cell function through interaction with HLA-C (Colonna *et al*., 1993; Winter and Long, 1997). *KIR2DS3* and *KIR2DS5*, like *KIR2DL2* and *KIR2DL3*, effectively represent alleles of the same locus and strong linkage disequilibrium between *KIR2DL5*, *KIR2DS3/2DS5* and *KIR2DL2/2DL3* has previously been reported in this (Chapter 2, Section 2.3.4) and other African populations (Hou *et al*., 2010; Nakimuli *et al*., 2013; Norman *et al*., 2013). Supporting this explanation are the observations that a *KIR2DS3*-containing haplotype has previously been associated with increased risk of chronic HCV infection (Dring *et al*., 2011), while *KIR2DS5*-

containing haplotypes are associated with a reduced risk of developing pre-eclampsia in African women (Nakimuli *et al*., 2015).

Allelic variability is known to modify the structure of both KIR2DL1 and KIR2DL2 in a manner that could potentially impact on their function (VandenBussche *et al*., 2006; Bari *et al*., 2009; Hilton *et al*., 2012; Frazier *et al*., 2013). These variations may affect the levels of receptor expression at the cell surface (VandenBussche *et al*., 2006) or modify the avidity of receptor-ligand interactions (Hilton *et al*., 2012; Frazier *et al*., 2013). The same is true of KIR3DL1 (Pando *et al*., 2003; Carr *et al*., 2005), for which allelic variability has already been shown to be associated with HIV-1 control in individuals in possession of a corresponding HLA-Bw4 ligand (Lopez Vasquez *et al*., 2005; Martin *et al*., 2007). The involvement of *KIR* allelic polymorphism in HIV-1 control could thus account for why we did not see a significant difference in the frequencies of KIR-HLA ligand pairs between HIV-1 controllers and progressors. If differences in the immune mechanisms employed by individuals in either group are indeed mediated by allelic micro-alterations in KIR receptor structure, it would be beyond the power of the present study to detect these associations.

Regardless, our data provide the first description of genetic variability within the *KIR* gene complex in Black South African HIV-1 controllers and progressors. We demonstrate a clear association between telomeric *KIR* tA01 haplotype motifs lacking either *KIR2DS4* and *KIR3DL1* and the loss of control of HIV-1 replication, while concurrently demonstrating an association between the presence of specific cB01 centromeric haplotype motifs and the ability to spontaneously suppress viral replication. We further observed the presence of *KIR2DS3* (rather than *KIR2DS5*) within these cB01 motifs and concomitant heterozygosity with cA01 motifs to be significantly associated with HIV-1 elite controller status. Collectively, these data point to a potentially important role for NK cell-mediated immunity in HIV-1 control in the Black South African population.

Chapter 6

Concluding Remarks

The rate of HIV-1 disease progression is known to vary, often dramatically, between individuals. A small subset of infected individuals, termed HIV-1 controllers, are able to maintain very low levels of viraemia (< 2000 RNA copies/ml) in the absence of highly active antiretroviral therapy. A still smaller percentage of individuals, known as elite controllers, can be further distinguished from other controllers by their ability to suppress viral replication to levels undetectable by currently available commercial assays (Deeks and Walker, 2007). It is thought that an understanding of the genetic and immunological characteristics that define these individuals may provide insights pertinent to the development of an HIV-1 vaccine and/or other potential therapeutic interventions (Walker and Yu, 2013).

However, attempts to characterize the host genetic, viral and immunological signatures associated with natural control of HIV infection have revealed HIV-1 controllers to be a heterogeneous group and have suggested that multiple immune mechanisms may be employed in order to control HIV-1 replication (Walker and Yu, 2013). Genetic variability within the *HLA* class I loci is widely recognized as one of the major determinants of HIV-1 control (Goulder and Walker, 2012). GWAS studies have revealed polymorphic variants within the *HLA* gene complex to be the genetic factors most strongly associated with viral control (Fellay *et al*., 2007; The International HIV Controller Study, 2010) and a number of *HLA* alleles have been associated with differential HIV-1 disease outcome (Kiepiela *et al*., 2004; Lazaryan *et al*., 2006; Leslie *et al*., 2010; Tang *et al*., 2011). More recently, individual polymorphisms within both the coding (The International HIV Controller Study, 2010; McLaren *et al*., 2012) and regulatory regions surrounding these genes (Kulkarni *et al*., 2011; Ballana *et al*., 2012) have also been shown to be associated with viral control.

HLA class I molecules are crucial regulators of the host response to viral infection, mediating both adaptive and innate immune responses through their interactions with both the TCR and KIR on the surface of NK cells. As a result, the loci encoding these receptors have evolved to be the most polymorphic in the human genome (Robinson *et al*., 2014). This high level of variability often lead to differences in the allele distributions and patterns of linkage disequilibrium across these loci being observed between populations with distinct demographic histories (Cao *et al*., 2001; Cao *et al*., 2004; Paximadis *et al*., 2012). These differences, if unaccounted for, have the potential to confound disease association studies. Therefore, a thorough prior understanding of the patterns of genetic variability present within ones population of interest is necessary in order to fully understand the nature of potential correlations observed during disease association studies.

While allelic patterns of genetic variability across the *HLA* class I loci within the Black and Caucasian South African population groups had previously been reported (Paximadis *et al*., 2012) prior to the commencement of this study, a description of the patterns of variation within the non-coding regulatory regions upstream (*-35* SNP; rs9264942) and downstream (*263* indel; rs67384697) of *HLA-C* in the South African population had not yet been reported. The *-35* SNP was initially identified as one of the genetic variants most strongly associated with lower HIV-1 viral set point (Fellay *et al*., 2007). The *-35C* allele at this position had also previously been shown to be strongly associated with increased HLA-C mRNA expression in Epstein-Barr virus-transformed B-cell lines (Stranger *et al*., 2005). Therefore, the association observed between the *-35C* allele and HIV-1 control was thought to stem from differential regulation of HLA-C expression at this position (Fellay *et al*., 2007).

However, the association seen between this SNP and HIV-1 viral set point in individuals of European descent (Fellay *et al*., 2007), was not observed in African-Americans (Shrestha *et al*., 2009); suggesting that variability at this SNP was not responsible for the observed alterations in

HLA-C expression, but rather acted as marker for the true causal variant (Corrah *et al*., 2009). The *263* indel was subsequently identified as the variant responsible for mediating differential HLA-C expression (Kulkarni *et al*., 2011; Kulkarni *et al*., 2013) and was shown to be in strong LD with the *-35* SNP in European-Americans (Kulkarni *et al*., 2011). We therefore examined and characterized variability at these two positions in individuals from the Black and Caucasian South African population groups, in order to both confirm their presence in the South African population, determine their allelic representation if present and to evaluate LD between them in both population groups (Chapter 3).

While *-35* SNP and *263* indel were both found to occur in the Black and Caucasian population groups, we demonstrated that these polymorphisms were not in strong LD in either population; indicating that the *-35* SNP was not an appropriate marker for the *263* indel in either group. Our findings provided the first description of variability at these polymorphic positions in a sub-Saharan African population; but more importantly, highlight the necessity of exercising caution when applying findings observed in populations of European descent to individuals from African populations. African populations are known to exhibit higher levels of genetic variability and lower levels of LD than their non-African counterparts (Tishkoff and Williams, 2002). These genetic characteristics have important implications for medical association studies conducted in these populations (Gomez *et al*., 2014), particularly GWAS studies, which are heavily reliant on LD. Our observation that, unlike in Caucasian-Americans (Fellay *et al*., 2007), the *-35* SNP and *263* indel were not found to be in strong LD in the Caucasian South African population, also demonstrates the need to consider the role of demographic history in shaping patterns of genetic variability and LD (Rosenberg *et al*., 2002) when designing association studies.

The *-35* SNP (Stranger *et al*., 2005) and *263* indel (Kulkarni *et al*., 2011) were first identified as a result of their association with HLA-C expression in individuals of European descent. We

further identified an additional indel at position 230 within the *HLA-C* 3' UTR (rs35877659), which we hypothesized could also impact HLA-C expression through modulation of miRNA-binding (Chapter 3). Given that HIV-1 Nef selectively downregulates HLA-A and –B to avoid lysis of primarily infected cells by cytotoxic T lymphocytes (Collins *et al*., 1998), but does not significantly alter HLA-C expression in order to avoid NK cell-mediated lysis (Cohen *et al*., 1999), and that an inverse correlation has been observed between HLA-C expression and HIV-1 viral load (Apps *et al*., 2013), we evaluated the impact of genetic variability at the two indel positions on HIV-1 control in the Black South African population (Chapter 5).

We found HIV-1 progressors and –controllers to exhibit very similar patterns of genetic variability at the *263* indel; suggesting that either this variant does not significantly alter HLA-C expression in the Black South African population, or that alterations in HLA-C expression are not associated with HIV-1 control in this population. Our observation that progressors and controllers exhibit (albeit not significant) differences in allele frequencies at the *230* indel would suggest that the former explanation seems more likely. We had previously proposed a model in which variability at both the *230* and *263* indels acted in concert to differentially regulate HLA-C expression (Chapter 3). While we did not functionally verify the validity of this model, the observation by others (Corrah *et al*., 2009; Kulkarni *et al*., 2011; Apps *et al*., 2013) that the *HLA-C* alleles with the lowest levels of expression correspond to those in our population with insertion alleles at both indel positions, would support our findings.

Although we did not observe a significant association between markers of increased HLA-C expression and HIV-1 control (Chapter 4), we can not exclude the possibility that due to limitations in our sample size, our study may lack sufficient power to detect associations with a small effect size. In this way, the more subtle effects of genetic variability within *HLA-C* on HIV-1 control may be masked by the dominance of those associations detected within the other *HLA* class I loci. The dominant influence of HLA-B in mediating HIV-1 control, particularly in

the Black South African population, is well established (Kiepiela *et al*., 2004; Leslie *et al*., 2010). Consistent with this, we have noted strong associations between HIV-1 control and genetic variability at the *HLA-B* locus, involving both specific *HLA-B* alleles and individual polymorphic positions within the locus (Chapter 5).

Our observation that *HLA-B*57:03:01* and *B*58:01:01* are the alleles most strongly associated with HIV-1 control is consistent with multiple reports in the Black South African (Kiepiela *et al*., 2004; Leslie *et al*., 2010; Tang *et al*., 2011) and other African (Lazaryan *et al*., 2006; Costello *et al*., 2008; Lazaryan *et al*., 2010) populations. While our findings identifying the HLA-B residues associated with HIV-1 control in the Black South African population also confirm those observed in both European- (The International HIV Controller Study, 2010) and African-American (McLaren *et al.*, 2012) populations and support the idea that peptide selection is an important factor in determining the host response to HIV-1 infection (The International HIV Controller Study, 2010). Collectively these data (Chapter 4) support the well established assertion that a large degree of the genetic variability within *HLA-B* impacts on HIV-1 control as a consequence of its influence on HIV-specific CTL responses (Kiepiela *et al*., 2007; Miura *et al*., 2009; Berger *et al*., 2011).

However, our identification of additional residues within both HLA-A and HLA-B that have previously been shown to modulate their interaction with receptors on both NK- (Sanjanwala *et al*., 2008) and dendritic cells (Jones *et al*., 2011), suggests that HLA class I-mediated regulation of innate immune function is also involved in HIV-1 control in the Black South African population. As innate immune effector cells, NK cells are well established as being critical to the early immune response to viral infection (Caligiuri, 2008). However recent findings supporting the existence of immunological memory in NK cells (Cooper *et al*., 2009; Sun *et al*., 2009), as well as the identification of KIR-associated amino-acid polymorphisms in the HIV-1

sequences of chronically infected individuals (Alter *et al*., 2011), have led to a re-evaluation of the role of these cells in chronic HIV-1 infection.

Accordingly, an association with HIV-1 control has been observed between the NK cell receptor, KIR3DL1, and its HLA-Bw4 ligand (Martin *et al*., 2007). Based on their surface expression, *KIR3DL1* alleles can be classified as either high expressing (*KIR3DL1*h*), low expressing (*KIR3DL1*l*) or non-expressed (*KIR3DL1*004*) (Gardiner *et al*., 2001: Pando *et al*., 2003; Thomas *et al*., 2008). *KIR3DL*h/KIR3DL*h* and *KIR3DL1*h/KIR3DL1*004* genotypes (collectively referred to as *KIR3DL1*h/*y*) have been shown to be associated with protection against HIV-1 disease progression in the presence of HLA-Bw4 molecules with isoleucine at position 80 (Bw4-80I) and this protective effect is most pronounced when the HLA-Bw4 molecule in question is HLA-B57 (Martin *et al*., 2007).

NK cells from HIV-1 slow-progressors with KIR3DL1 in combination with HLA-Bw4 show greater functionality than those from individuals without Bw4 (Kamya *et al*., 2011). This can largely been attributed to the fact that, due to NK cell licensing, individuals with HLA-Bw4 who express KIR3DL1 on the surface of their NK cells exhibit an increased capacity for ADCC relative to individuals without HLA-Bw4 (Parsons *et al*., 2010; Parsons *et al*., 2012). Again, the highest functional potential is demonstrated by individuals with the *KIR3DL1*h/*y/B*57* genotype (Boulet *et al*., 2010; Kamya *et al*., 2011) and NK cells from these individuals are better able to inhibit HIV-1 replication in autologous $CD4^+$ T cells than those from individuals homozygous for Bw6 or in possession of *KIR3DL1*l* alleles (Song *et al*., 2014).

The *KIR3DL1*h/*y* genotype in combination with HLA-Bw4-80I has also be shown to be associated with strong NK responses in elite controllers with reduced HIV-1 Gag-specific $CD8^+$ T cell responses (Tomescu *et al*., 2012). Given that elite controllers often elicit strong HLA-B57-restricted CTL responses to Gag early in infection (Miura *et al*., 2009) and that CTL escape

mutations within an immunodominant HLA-B57-restricted Gag epitope are known to abrogate KIR3DL1 binding to HLA-B57 (Fadda *et al.*, 2011), it is not unreasonable to expect that CTL and NK responses in elite controllers with both *KIR3DL1* and *HLA-B*57* may act synergistically, possibly at different points during infection, to limit viral replication (Fadda *et al*., 2011).

While our study could not directly demonstrate this, as we did not examine allelic variability within *KIR3DL1* and our cohort did not include sufficient individuals in possession of *HLA-B*57* alleles to adequately assess characteristics uniquely associated with these individuals, we observed that *B*57:03:01* is the allele most strongly associated with HIV-1 control and identified of an association between HIV-1 control and amino acids within HLA-B known to affect the interaction between KIR3DL1 and its HLA-Bw4 ligands (Chapter 4) - likely through differential peptide binding (Sanjanwala *et al*., 2008). This could suggest that immune mechanisms mediated by the *KIR3DL1*h/*y/B*57* genotype may influence HIV-1 control in the Black South African population and therefore warrant further functional study.

A role for KIR3DL1 in controlling HIV-1 infection in the Black South African population is further demonstrated by the observation that HIV-1 progressors possess significantly fewer copies of *KIR3DL1* per individual than either HIV-1 controllers (Chapter 5) or HIV-1 uninfected individuals within the Black South African population (Chapter 2). Information regarding the patterns of genetic variability within the *KIR* gene complex in either the Black or Caucasian South African populations was not available prior to the commencement of this study. Therefore, before we could adequately assess the impact of genetic variability at this locus, it was necessary to first describe the *KIR* gene and haplotype content within the general South African population (Chapter 2).

As anticipated, based on similar studies conducted in other African populations (Norman *et al*., 2007; Single *et al*., 2007; Nakimuli *et al*., 2013), we found *KIR3DS1* and *KIR2DS1* to be severely under-represented in the Black South African population group relative to Caucasians. This observation impacted on our ability to assess the role of variability within the *KIR* gene complex on HIV-1 control, as the co-occurrence of *KIR3DS1* and HLA-Bw4 has been reported to be favorably associated with HIV-1 disease progression in individuals of European descent (Martin *et al*., 2002a). However, despite this limitation, we were still able to identify patterns of *KIR* genetic variability associated with HIV-1 control (Chapter 5).

Variability in terms of *KIR* gene content is determined by the presence of two main *KIR* haplotypes (Uhrberg *et al*., 2002, Martin *et al*., 2004), which are themselves the product of varying combinations of a limited number of centromeric and telomeric gene content motifs (Pyo *et al*., 2010, Jiang *et al*., 2012; Vierra-Green *et al*., 2012, Pyo *et al*., 2013). Our examination of the distribution of these motifs and their associated haplotypes within HIV-1 progressors and –controllers revealed that while HIV-1 progression was associated with telomeric motifs lacking *KIR3DL1* (and to a lesser extent *KIR2DS4*), HIV-1 control was associated with the heterozygous combination of the centromeric cA01 and cB01 haplotype motifs (Chapter 5).

These findings are consistent with previous reports demonstrating an over-representation on the Bx 21 *KIR* genotype and increased frequency of *KIR2DS2* and *KIR2DS5* in individuals who exhibit NK responses to HIV-1 peptides (Tiemessen *et al*., 2011), as the Bx 21 genotype is representative of the heterozygous combination of the cA01|tA01 haplotype and a cB01|tA01 haplotype containing *KIR2DS5*. These HIV-specific NK responses have previously been shown to be associated with HIV-1 control in Black South African women (Tiemessen *et al*., 2010).

Furthermore, we noted that elite controllers could be distinguished from other HIV-1 controllers based on an increased likelihood of being in possession of cB01 haplotype motifs containing *KIR2DS3* rather than *KIR2DS5*. As both KIR2DS3 and KIR2DS5 have been demonstrated to show markedly reduced levels of surface expression relative to other KIR receptors (Steiner *et al*., 2008; VandenBussche *et al*., 2009; Steiner *et al*., 2014) and are not known to recognize HLA class I (Chiesa *et al*., 2008), we hypothesized that the association observed between these loci and HIV-1 control was not a consequence of any functional interaction on their part, but rather as result of their acting as markers for specific allelic variants of the inhibitory receptors KIR2DL1 and/or KIR2DL2 - both of which are known to regulate NK cell function through interaction with HLA-C (Colonna *et al*., 1993; Winter and Long, 1997).

We could not, unfortunately, verify this association experimentally, as we did not examine allelic variability with the loci encoding inhibitory KIR receptors. However, it is known that KIR2DL1 allelic variants exhibit extensive heterogeneity with respect to their ability to inhibit NK cell-mediated cytotoxicity and cytokine secretion (Bari *et al*., 2009). KIR2DL1 expression is also known to be reduced in chronically infected African individuals (Eller *et al*., 2011) and has been shown to inversely correlate with increased ADCC responses in elite controllers (Jia *et al*., 2013). We therefore propose, in light of its potential role in HIV-1 control, that further examination of the *KIR2DL1* allelic variability present within the Black South African population is warranted.

Collectively the data presented within this study provide new insights into the impact of genetic variability with the *HLA* class I loci in the Black South African population - a group severely affected by the HIV-1 pandemic - and highlight its influence on HIV-1 control in this population. This work also represents the first comprehensive description of genetic variability within the *KIR* gene complex in the Black South African population and identifies the genetic variants within this locus associated with HIV-1 control in this population. Even in the absence

of functional verification of key findings, these data strongly suggest an important role for these molecules in mediating the immune response to HIV-1 and demonstrate the need to better understand the role of NK cell-mediated immunity in the control of HIV-1 infection in Black South Africans.

References

Alter, G., Martin, M.P., Teigen, N., Carr, W.H., Suscovich, T.J., Schneidewind, A., Streeck, H., Waring, M., Meier, A., Brander, C., et al. (2007). Differential natural killer cell–mediated inhibition of HIV-1 replication based on distinct KIR/HLA subtypes. J. Exp. Med. *204*, 3027–3036.

Alter, G., Rihn, S., Walter, K., Nolting, A., Martin, M., Rosenberg, E.S., Miller, J.S., Carrington, M., and Altfeld, M. (2009). HLA Class I Subtype-Dependent Expansion of KIR3DS1+ and KIR3DL1+ NK Cells during Acute Human Immunodeficiency Virus Type 1 Infection. J. Virol. *83*, 6798–6805.

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

Apps, R., Qi, Y., Carlson, J.M., Chen, H., Gao, X., Thomas, R., Yuki, Y., Del Prete, G.Q., Goulder, P., Brumme, Z.L., et al. (2013). Influence of HLA-C Expression Level on HIV Control. Science *340*, 87–91.

Badrinath, S., Huyton, T., Schumacher, H., Blasczyk, R., and Bade-Doeding, C. (2012). Position 45 influences the peptide binding motif of HLA-B*44:08. Immunogenetics *64*, 245–249.

Ballana, E., Ruiz-de Andres, A., Mothe, B., Ramirez de Arellano, E., Aguilar, F., Badia, R., Grau, E., Clotet, B., del Val, M., Brander, C., et al. (2012). Differential prevalence of the HLA-C −35 CC genotype among viremic long term non-progressor and elite controller HIV+ individuals. Immunobiology *217*, 889–894.

Barber, L.D., Percival, L., Arnett, K.L., Gumperz, J.E., Chen, L., and Parham, P. (1997). Polymorphism in the alpha 1 helix of the HLA-B heavy chain can have an overriding influence on peptide-binding specificity. J. Immunol. *158*, 1660–1669.

Bari, R., Bell, T., Leung, W.-H., Vong, Q.P., Chan, W.K., Das Gupta, N., Holladay, M., Rooney, B., and Leung, W. (2009). Significant functional heterogeneity among KIR2DL1 alleles and a pivotal role of arginine245. Blood *114*, 5182–5190.

Bashirova, A.A., Martin-Gayo, E., Jones, D.C., Qi, Y., Apps, R., Gao, X., Burke, P.S., Taylor, C.J., Rogich, J., Wolinsky, S., et al. (2014). LILRB2 Interaction with HLA Class I Correlates with Control of HIV-1 Infection. PLoS Genet. *10*, e1004196.

Berger, C.T., Frahm, N., Price, D.A., Mothe, B., Ghebremichael, M., Hartman, K.L., Henry, L.M., Brenchley, J.M., Ruff, L.E., Venturi, V., et al. (2011). High-Functional-Avidity Cytotoxic T Lymphocyte Responses to HLA-B-Restricted Gag-Derived Epitopes Associated with Relative HIV Control. J. Virol. *85*, 9334–9345.

Biassoni, R., Falco, M., Cambiaggi, A., Costa, P., Verdiani, S., Pende, D., Conte, R., Di Donato, C., Parham, P., and Moretta, L. (1995). Amino acid substitutions can influence the natural killer (NK)-mediated recognition of HLA-C molecules. Role of serine-77 and lysine-80 in the target cell protection from lysis mediated by "group 2" or "group 1" NK clones. J. Exp. Med. *182*, 605–609.

Bihl, F., Frahm, N., Di Giammarino, L., Sidney, J., John, M., Yusim, K., Woodberry, T., Sango, K., Hewitt, H.S., Henry, L., et al. (2006). Impact of HLA-B Alleles, Epitope Binding Affinity, Functional Avidity, and Viral Coinfection on the Immunodominance of Virus-Specific CTL Responses. J. Immunol. *176*, 4094–4101.

Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. (1987a). Structure of the human class I histocompatibility antigen, HLA-A2. Nature *329*, 506–512.

Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. (1987b). The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. Nature *329*, 512–518.

Boulet, S., Song, R., Kamya, P., Bruneau, J., Shoukry, N.H., Tsoukas, C.M., and Bernard, N.F. (2010). HIV Protective KIR3DL1 and HLA-B Genotypes Influence NK Cell Function Following Stimulation with HLA-Devoid Cells. J. Immunol. *184*, 2057–2064.

Boyington, J.C., and Sun, P.D. (2002). A structural perspective on MHC class I recognition by killer cell immunoglobulin-like receptors. Mol. Recognit. Immuno-Recept. Struct. Overv. *38*, 1007–1021.

Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Xavier, R.J., Lieberman, J. and Elledge, S.J. (2008). Identification of Host Proteins Required for HIV Infection Through a Functional Genomic Screen. Science *319*, 921-926

Caligiuri, M.A. (2008). Human natural killer cells. Blood *112*, 461–469.

Campbell, M.C., and Tishkoff, S.A. (2008). African Genetic Diversity: Implications for Human Demographic History, Modern Human Origins, and Complex Disease Mapping. Annu. Rev. Genomics Hum. Genet. *9*, 403–433.

Campbell, K.S., Cella, M., Carretero, M., López-Botet, M., and Colonna, M. (1998). Signaling through human killer cell activating receptors triggers tyrosine phosphorylation of an associated protein complex. Eur. J. Immunol. *28*, 599–609.

Cao, K., Hollenbach, J., Shi, X., Shi, W., Chopek, M., and Fernández-Viña, M.A. (2001). Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. Hum. Immunol. *62*, 1009–1030.

Cao, K., Moormann, A.M., Lyke, K.E., Masaberg, C., Sumba, O.P., Doumbo, O.K., Koech, D., Lancaster, A., Nelson, M., Meyer, D., et al. (2004). Differentiation between African populations is evidenced by the diversity of alleles and haplotypes of HLA class I loci. Tissue Antigens *63*, 293–325.

Carr, W.H., Pando, M.J., and Parham, P. (2005). KIR3DL1 Polymorphisms That Affect NK Cell Inhibition by HLA-Bw4 Ligand. J. Immunol. *175*, 5222–5229.

Carr, W.H., Rosen, D.B., Arase, H., Nixon, D.F., Michaelsson, J., and Lanier, L.L. (2007). Cutting Edge: KIR3DS1, a Gene Implicated in Resistance to Progression to AIDS, Encodes a DAP12-Associated Receptor Expressed on NK Cells That Triggers NK Cell Activation. J. Immunol. *178*, 647–651.

Cella, M., Longo, A., Ferrara, G.B., Strominger, J.L., and Colonna, M. (1994). NK3-specific natural killer cells are selectively inhibited by Bw4-positive HLA alleles with isoleucine 80. J. Exp. Med. *180*, 1235–1242.

Chan, H.-W., Kurago, Z.B., Stewart, C.A., Wilson, M.J., Martin, M.P., Mace, B.E., Carrington, M., Trowsdale, J., and Lutz, C.T. (2003). DNA Methylation Maintains Allele-specific KIR Gene Expression in Human Natural Killer Cells. J. Exp. Med. *197*, 245–255.

Chen, M., and Bouvier, M. (2007). Analysis of interactions in a tapasin/class I complex provides a mechanism for peptide selection. EMBO J. *26*, 1681–1690.

Chiesa, M.D., Romeo, E., Falco, M., Balsamo, M., Augugliaro, R., Moretta, L., Bottino, C., Moretta, A., and Vitale, M. (2008). Evidence that the KIR2DS5 gene codes for a surface receptor triggering natural killer cell function. Eur. J. Immunol. *38*, 2284–2289.

Cohen, G.B., Gandhi, R.T., Davis, D.M., Mandelboim, O., Chen, B.K., Strominger, J.L., and Baltimore, D. (1999). The Selective Downregulation of Class I Major Histocompatibility Complex Proteins by HIV-1 Protects HIV-Infected Cells from NK Cells. Immunity *10*, 661–671.

Collins, K.L., Chen, B.K., Kalams, S.A., Walker, B.D., and Baltimore, D. (1998). HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. Nature *391*, 397–401.

Colonna, M., and Samaridis, J. (1995). Cloning of immunoglobulin-superfamily members associated with HLA-C and HLA-B recognition by human natural killer cells. Science *268*, 405–408.

Colonna, M., Borsellino, G., Falco, M., Ferrara, G.B., and Strominger, J.L. (1993). HLA-C is the inhibitory ligand that determines dominant resistance to lysis by NK1- and NK2-specific natural killer cells. Proc. Natl. Acad. Sci. *90*, 12000–12004.

Conrad, D.F., Jakobsson, M., Coop, G., Wen, X., Wall, J.D., Rosenberg, N.A., and Pritchard, J.K. (2006). A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. Nat Genet *38*, 1251–1260.

Cooper, M.A., Elliott, J.M., Keyel, P.A., Yang, L., Carrero, J.A., and Yokoyama, W.M. (2009). Cytokine-induced memory-like natural killer cells. Proc. Natl. Acad. Sci. *106*, 1915–1919.

Corrah, T.W., Goonetilleke, N., Kopycinski, J., Deeks, S.G., Cohen, M.S., Borrow, P., McMichael, A., and Brackenridge, S. (2011). Reappraisal of the Relationship between the HIV-1-Protective Single-Nucleotide Polymorphism 35 Kilobases Upstream of the HLA-C Gene and Surface HLA-C Expression. J. Virol. *85*, 3367–3374.

Costello, C., Tang, J., Rivers, C., Karita, E., Meizen-Derr, J., Allen, S., and Kaslow, R.A. (1999). HLA-B*5703 independently associated with slower HIV-1 disease progression in Rwandan women. AIDS *13*, 1990.

Cowan, E.P., Jelachich, M.L., Coligan, J.E., and Biddison, W.E. (1987). Site-directed mutagenesis of an HLA-A3 gene identifies amino acid 152 as crucial for major-histocompatibility-complex-restricted and alloreactive cytotoxic-T-lymphocyte recognition. Proc. Natl. Acad. Sci. *84*, 5014–5018.

Dalmasso, C., Carpentier, W., Meyer, L., Rouzioux, C., Goujard, C., Chaix, M.-L., Lambotte, O., Avettand-Fenoel, V., Le Clerc, S., de Senneville, L.D., et al. (2008). Distinct Genetic Loci Control Plasma HIV-RNA and Cellular HIV-DNA Levels in HIV-1 Infection: The ANRS Genome Wide Association 01 Study. PLoS ONE *3*, e3907.

Davies, G.E., Locke, S.M., Wright, P.W., Li, H., Hanson, R.J., Miller, J.S., and Anderson, S.K. (2007). Identification of bidirectional promoters in the human KIR genes. Genes Immun *8*, 245–253.

Deeks, S.G., and Walker, B.D. (2007). Human Immunodeficiency Virus Controllers: Mechanisms of Durable Virus Control in the Absence of Antiretroviral Therapy. Immunity *27*, 406–416.

Denis, L., Sivula, J., Gourraud, P.-A., Kerdudou, N., Chout, R., Ricard, C., Moisan, J.-P., Gagne, K., Partanen, J., and Bignon, J.-D. (2005). Genetic diversity of KIR natural killer cell markers in populations from France, Guadeloupe, Finland, Senegal and Réunion. Tissue Antigens *66*, 267–276.

Döhring, C., Scheidegger, D., Samaridis, J., Cella, M., and Colonna, M. (1996). A human killer inhibitory receptor specific for HLA-A1,2. J. Immunol. *156*, 3098–3101.

Dring, M.M., Morrison, M.H., McSharry, B.P., Guinan, K.J., Hagan, R., Irish HCV Research Consortium, O'Farrelly, C., and Gardiner, C.M. (2011). Innate immune genes synergize to predict increased risk of chronic disease in hepatitis C virus infection. Proc. Natl. Acad. Sci. *108*, 5736–5741.

Du, Z., Gjertson, D., Reed, E., and Rajalingam, R. (2007). Receptor-ligand analyses define minimal killer cell Ig-like receptor (KIR) in humans. Immunogenetics *59*, 1–15.

Eller, M.A., Koehler, R.N., Kijak, G.H., Eller, L.A., Guwatudde, D., Marovich, M.A., Michael, N.L., de Souza, M.S., Wabwire-Mangen, F., Robb, M.L., et al. (2011). Human Immunodeficiency Virus Type 1 Infection Is Associated with Increased NK Cell Polyfunctionality and Higher Levels of KIR3DL1+ NK Cells in Ugandans Carrying the HLA-B Bw4 Motif. J. Virol. *85*, 4802–4811.

Estefanía, E., Flores, R., Gómez-Lozano, N., Aguilar, H., López-Botet, M., and Vilches, C. (2007). Human KIR2DL5 Is an Inhibitory Receptor Expressed on the Surface of NK and T Lymphocyte Subsets. J. Immunol. *178*, 4402–4410.

Excoffier, L., and Lisher, H. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. *10*, 564–567.

Excoffier, L., Laval, G., and Balding, D. (2003). Gametic phase estimation over large genomic regions using an adaptive window approach. Hum. Genomics *1*, 7.

Fadda, L., O'Connor, G.M., Kumar, S., Piechocka-Trocha, A., Gardiner, C.M., Carrington, M., McVicar, D.W., and Altfeld, M. (2011). Common HIV-1 Peptide Variants Mediate Differential Binding of KIR3DL1 to HLA-Bw4 Molecules. J. Virol. *85*, 5970–5974.

Faure, M., and Long, E.O. (2002). KIR2DL4 (CD158d), an NK Cell-Activating Receptor with Inhibitory Potential. J. Immunol. *168*, 6208–6214.

Fellay, J., Shianna, K.V., Ge, D., Colombo, S., Ledergerber, B., Weale, M., Zhang, K., Gumbs, C., Castagna, A., Cossarizza, A., et al. (2007). A Whole-Genome Association Study of Major Determinants for Host Control of HIV-1. Science *317*, 944–947.

Fellay, J., Ge, D., Shianna, K.V., Colombo, S., Ledergerber, B., Cirulli, E.T., Urban, T.J., Zhang, K., Gumbs, C.E., Smith, J.P., et al. (2009). Common Genetic Variation and the Control of HIV-1 in Humans. PLoS Genet. *5*, e1000791.

Fleming, K.A., McMichael, A., Morton, J.A., Woods, J., and McGee, J.O. (1981). Distribution of HLA class 1 antigens in normal human tissue and in mammary cancer. J. Clin. Pathol. *34*, 779–784.

Flores, A.C., Marcos, C.Y., Paladino, N., Capucchio, M., Theiler, G., Arruvito, L., Pardo, R., Habegger, A., Williams, F., Middleton, D., et al. (2007). KIR genes polymorphism in Argentinean Caucasoid and Amerindian populations. Tissue Antigens *69*, 568–576.

Frazier, W.R., Steiner, N., Hou, L., Dakshanamurthy, S., and Hurley, C.K. (2013). Allelic Variation in KIR2DL3 Generates a KIR2DL2-like Receptor with Increased Binding to its HLA-C Ligand. J. Immunol. *190*, 6198–6208.

Fry, A.M., Lanier, L.L., and Weiss, A. (1996). Phosphotyrosines in the killer cell inhibitory receptor motif of NKB1 are required for negative signaling and for association with protein tyrosine phosphatase 1C. J. Exp. Med. *184*, 295–300.

Gardiner, C.M., Guethlein, L.A., Shilling, H.G., Pando, M., Carr, W.H., Rajalingam, R., Vilches, C., and Parham, P. (2001). Different NK Cell Surface Phenotypes Defined by the DX9 Antibody Are Due to KIR3DL1 Gene Polymorphism. J. Immunol. *166*, 2992–3001.

Gaudieri, S., DeSantis, D., McKinnon, E., Moore, C., Nolan, D., Witt, C.S., Mallal, S.A., and Christiansen, F.T. (2005). Killer immunoglobulin-like receptors and HLA act both independently and synergistically to modify HIV disease progression. Genes Immun *6*, 683–690.

Gendzekhadze, K., Norman, P., Abi-Rached, L., Layrisse, Z., and Parham, P. (2006). High KIR diversity in Amerindians is maintained using few gene-content haplotypes. Immunogenetics *58*, 474–480.

Giebel, S., Locatelli, F., Lamparelli, T., Velardi, A., Davies, S., Frumento, G., Maccario, R., Bonetti, F., Wojnar, J., Martinetti, M., et al. (2003). Survival advantage with KIR ligand incompatibility in hematopoietic stem cell transplantation from unrelated donors. Blood *102*, 814–819.

Gillespie, G.M.A., Bashirova, A., Dong, T., McVicar, D.W., Rowland-Jones, S.L., and Carrington, M. (2007). Lack of KIR3DS1 Binding to MHC Class I Bw4 Tetramers in Complex with CD8+ T Cell Epitopes. AIDS Res. Hum. Retroviruses *23*, 451–455.

Gomez, F., Hirbo, J., and Tishkoff, S.A. (2014). Genetic Variation and Adaptation in Africa: Implications for Human Evolution and Disease. Cold Spring Harb. Perspect. Biol. *6*.

Gonzalez, E., Dhanda, R., Bamshad, M., Mummidi, S., Geevarghese, R., Catano, G., Anderson, S.A., Walter, E.A., Stephan, K.T., Hammer, M.F., et al. (2001). Global survey of genetic variation in CCR5, RANTES, and MIP-1α: Impact on the epidemiology of the HIV-1 pandemic. Proc. Natl. Acad. Sci. *98*, 5199–5204.

Gonzalez-Galarza, F.F., Christmas, S., Middleton, D., and Jones, A.R. (2011). Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. Nucleic Acids Res. *39*, D913–D919.

Goulder, P.J.R., and Walker, B.D. (2012). HIV and HLA Class I: An Evolving Relationship. Immunity *37*, 426–440.

Graef, T., Moesta, A.K., Norman, P.J., Abi-Rached, L., Vago, L., Older Aguilar, A.M., Gleimer, M., Hammond, J.A., Guethlein, L.A., Bushnell, D.A., et al. (2009). KIR2DS4 is a product of gene conversion with KIR3DL2 that introduced specificity for HLA-A*11 while diminishing avidity for HLA-C. J. Exp. Med. *206*, 2557–2572.

Guethlein, L.A., Flodin, L.R., Adams, E.J., and Parham, P. (2002). NK Cell Receptors of the Orangutan (Pongo pygmaeus): A Pivotal Species for Tracking the Coevolution of Killer Cell Ig-Like Receptors with MHC-C. J. Immunol. *169*, 220–229.

Guinan, K.J., Cunningham, R.T., Meenagh, A., Gonzalez, A., Dring, M.M., McGuinness, B.W., Middleton, D., and Gardiner, C.M. (2010). Signatures of natural selection and coevolution between killer cell immunoglobulin-like receptors (KIR) and HLA class I genes. Genes Immun *11*, 467–478.

Gumperz, J.E., Litwin, V., Phillips, J.H., Lanier, L.L., and Parham, P. (1995). The Bw4 public epitope of HLA-B molecules confers reactivity with natural killer cell clones that express NKB1, a putative HLA receptor. J. Exp. Med. *181*, 1133–1144.

Guo, S.W., and Thompson, E.A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 361–372.

Han, Y., Lai, J., Barditch-Crovo, P., Gallant, J.E., Williams, T.M., Siliciano, R.F., and Blankson, J.N. (2008). The role of protective HCP5 and HLA-C associated polymorphisms in the control of HIV-1 replication in a subset of elite suppressors. AIDS *22*, 541–544.

Hansasuta, P., Dong, T., Thananchai, H., Weekes, M., Willberg, C., Aldemir, H., Rowland-Jones, S., and Braud, V.M. (2004). Recognition of HLA-A3 and HLA-A11 by KIR3DL2 is peptide-specific. Eur. J. Immunol. *34*, 1673–1679.

Henn, B.M., Gignoux, C., Lin, A.A., Oefner, P.J., Shen, P., Scozzari, R., Cruciani, F., Tishkoff, S.A., Mountain, J.L., and Underhill, P.A. (2008). Y-chromosomal evidence of a pastoralist migration through Tanzania to southern Africa. Proc. Natl. Acad. Sci. *105*, 10693–10698.

Hiby, S.E., Walker, J.J., O'Shaughnessy, K.M., Redman, C.W.G., Carrington, M., Trowsdale, J., and Moffett, A. (2004). Combinations of Maternal KIR and Fetal HLA-C Genes Influence the Risk of Preeclampsia and Reproductive Success. J. Exp. Med. *200*, 957–965.

Hiby, S.E., Regan, L., Lo, W., Farrell, L., Carrington, M., and Moffett, A. (2008). Association of maternal killer-cell immunoglobulin-like receptors and parental HLA-C genotypes with recurrent miscarriage. Hum. Reprod. *23*, 972–976.

Hill, W.G., and Robertson, A. (1968). Linkage disequilibrium in finite populations. Theor. Appl. Genet. *38*, 226–231.

Hilton, H.G., Vago, L., Older Aguilar, A.M., Moesta, A.K., Graef, T., Abi-Rached, L., Norman, P.J., Guethlein, L.A., Fleischhauer, K., and Parham, P. (2012). Mutation at Positively Selected Positions in the Binding Site for HLA-C Shows That KIR2DL1 Is a More Refined but Less Adaptable NK Cell Receptor Than KIR2DL3. J. Immunol. *189*, 1418–1430.

Hirayasu, K., Ohashi, J., Kashiwase, K., Hananantachai, H., Naka, I., Ogawa, A., Takanashi, M., Satake, M., Nakajima, K., Parham, P., et al. (2012). Significant Association of KIR2DL3-HLA-C1 Combination with Cerebral Malaria and Implications for Co-evolution of KIR and HLA. PLoS Pathog. *8*, e1002565.

Hirschhorn, J.N., Lohmueller, K., Byrne, E., and Hirschhorn, K. (2002). A comprehensive review of genetic association studies. Genet Med *4*, 45–61.

Hogan, K.T., Shimojo, N., Walk, S.F., Engelhard, V.H., Maloy, W.L., Coligan, J.E., and Biddison, W.E. (1988a). Mutations in the alpha 2 helix of HLA-A2 affect presentation but do not inhibit binding of influenza virus matrix peptide. J. Exp. Med. *168*, 725–736.

Hogan, K.T., Clayberger, C., Bernhard, E.J., Walk, S.F., Ridge, J.P., Parham, P., Krensky, A.M., and Engelhard, V.H. (1988b). Identification by site-directed mutagenesis of amino acid residues contributing to serologic and CTL-defined epitope differences between HLA-A2.1 and HLA-A2.3. J. Immunol. *141*, 2519–2525.

Hollenbach, J., Nocedal, I., Ladner, M., Single, R., and Trachtenberg, E. (2012). Killer cell immunoglobulin-like receptor (KIR) gene content variation in the HGDP-CEPH populations. Immunogenetics *64*, 719–737.

Honeyborne, I., Prendergast, A., Pereyra, F., Leslie, A., Crawford, H., Payne, R., Reddy, S., Bishop, K., Moodley, E., Nair, K., et al. (2007). Control of Human Immunodeficiency Virus Type 1 Is Associated with HLA-B*13 and Targeting of Multiple Gag-Specific CD8+ T-Cell Epitopes. J. Virol. *81*, 3667–3672.

Hong, H.A., Loubser, A.S., de Assis Rosa, D., Naranbhai, V., Carr, W., Paximadis, M., Lewis, D.A., Tiemessen, C.T., and Gray, C.M. (2011). Killer-cell immunoglobulin-like receptor genotyping and HLA killer-cell immunoglobulin-like receptor-ligand identification by real-time polymerase chain reaction. Tissue Antigens *78*, 185–194.

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

Hou, L., Chen, M., Jiang, B., Wu, D., Ng, J., and Hurley, C. (2010). Thirty allele-level haplotypes centered around KIR2DL5 define the diversity in an African American population. Immunogenetics *62*, 491–498.

Hou, L., Chen, M., Ng, J., and Hurley, C.K. (2012). Conserved KIR allele-level haplotypes are altered by microvariation in individuals with European ancestry. Genes Immun *13*, 47–58.

Hou, Y.-F., Zhang, Y.-C., Jiao, Y.-L., Wang, L.-C., Li, J.-F., Pan, Z.-L., Yang, Q.-R., Sun, H.-S., and Zhao, Y.-R. (2009). Disparate distribution of activating and inhibitory killer cell immunoglobulin-like receptor genes in patients with systemic lupus erythematosus. Lupus.

Hsu, K.C., Keever-Taylor, C.A., Wilton, A., Pinto, C., Heller, G., Arkun, K., O'Reilly, R.J., Horowitz, M.M., and Dupont, B. (2005). Improved outcome in HLA-identical sibling hematopoietic stem-cell transplantation for acute myelogenous leukemia predicted by KIR and HLA genotypes. Blood *105*, 4878–4884.

Huang, J., Goedert, J.J., Sundberg, E.J., Cung, T.D.H., Burke, P.S., Martin, M.P., Preiss, L., Lifson, J., Lichterfeld, M., Carrington, M., et al. (2009). HLA-B*35-Px–mediated acceleration of HIV-1 infection by increased inhibitory immunoregulatory impulses. J. Exp. Med. *206*, 2959–2966.

Jennes, W., Verheyden, S., Demanet, C., Menten, J., Vuylsteke, B., Nkengasong, J.N., and Kestens, L. (2011). Low CD4+ T Cell Counts among African HIV-1 Infected Subjects with Group B KIR Haplotypes in the Absence of Specific Inhibitory KIR Ligands. PLoS ONE *6*, e17043.

Jia, M., Li, D., He, X., Zhao, Y., Peng, H., Ma, P., Hong, K., Liang, H., and Shao, Y. (2013). Impaired natural killer cell-induced antibody-dependent cell-mediated cytotoxicity is associated with human immunodeficiency virus-1 disease progression. Clin. Exp. Immunol. *171*, 107–116.

Jiang, W., Johnson, C., Jayaraman, J., Simecek, N., Noble, J., Moffatt, M.F., Cookson, W.O., Trowsdale, J., and Traherne, J.A. (2012). Copy number variation leads to considerable diversity for B but not A haplotypes of the human KIR genes encoding NK cell receptors. Genome Res. *22*, 1845–1854.

Jiang, Y., Chen, O., Cui, C., Zhao, B., Han, X., Zhang, Z., Liu, J., Xu, J., Hu, Q., Liao, C., et al. (2013). KIR3DS1/L1 and HLA-Bw4-80I are associated with HIV disease progression among HIV typical progressors and long-term nonprogressors. BMC Infect. Dis. *13*, 405–405.

Jones, D.C., Kosmoliaptsis, V., Apps, R., Lapaque, N., Smith, I., Kono, A., Chang, C., Boyle, L.H., Taylor, C.J., Trowsdale, J., et al. (2011). HLA Class I Allelic Sequence and Conformation Regulate Leukocyte Ig-Like Receptor Binding. J. Immunol. *186*, 2990–2997.

Kamya, P., Boulet, S., Tsoukas, C.M., Routy, J.-P., Thomas, R., Côté, P., Boulassel, M.-R., Baril, J.-G., Kovacs, C., Migueles, S.A., et al. (2011). Receptor-Ligand Requirements for Increased NK Cell Polyfunctional Potential in Slow Progressors Infected with HIV-1 Coexpressing KIR3DL1*h/*y and HLA-B*57. J. Virol. *85*, 5949–5960.

Karlsen, T.H., Boberg, K.M., Olsson, M., Sun, J.-Y., Senitzer, D., Bergquist, A., Schrumpf, E., Thorsby, E., and Lie, B.A. (2007). Particular genetic variants of ligands for natural killer cell receptors may contribute to the HLA associated risk of primary sclerosing cholangitis. J. Hepatol. *46*, 899–906.

Kaslow, R.A., Dorak, T. and Tang, J. (2005). Influence of Host Genetic Variation on Susceptibility to HIV Type 1 Infection. J. Infect. Dis. *191*, S68-S77

Katz, G., Markel, G., Mizrahi, S., Arnon, T.I., and Mandelboim, O. (2001). Recognition of HLA-Cw4 but Not HLA-Cw6 by the NK Cell Receptor Killer Cell Ig-Like Receptor Two-Domain Short Tail Number 4. J. Immunol. *166*, 7260–7267.

Katz, G., Gazit, R., Arnon, T.I., Gonen-Gross, T., Tarcic, G., Markel, G., Gruda, R., Achdout, H., Drize, O., Merims, S., et al. (2004). MHC Class I-Independent Recognition of NK-Activating Receptor KIR2DS4. J. Immunol. *173*, 1819–1825.

Khakoo, S.I., Thio, C.L., Martin, M.P., Brooks, C.R., Gao, X., Astemborski, J., Cheng, J., Goedert, J.J., Vlahov, D., Hilgartner, M., et al. (2004). HLA and NK Cell Inhibitory Receptor Genes in Resolving Hepatitis C Virus Infection. Science *305*, 872–874.

Kiepiela, P., Leslie, A.J., Honeyborne, I., Ramduth, D., Thobakgale, C., Chetty, S., Rathnavalu, P., Moore, C., Pfafferott, K.J., Hilton, L., et al. (2004). Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. Nature *432*, 769–775.

Kiepiela, P., Ngumbela, K., Thobakgale, C., Ramduth, D., Honeyborne, I., Moodley, E., Reddy, S., de Pierres, C., Mncube, Z., Mkhwanazi, N., et al. (2007). CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. Nat Med *13*, 46–53.

Kikuchi-Maki, A., Yusa, S., Catina, T.L., and Campbell, K.S. (2003). KIR2DL4 Is an IL-2-Regulated NK Cell Receptor That Exhibits Limited Expression in Humans but Triggers Strong IFN-γ Production. J. Immunol. *171*, 3415–3425.

Kulkarni, S., Savan, R., Qi, Y., Gao, X., Yuki, Y., Bass, S.E., Martin, M.P., Hunt, P., Deeks, S.G., Telenti, A., et al. (2011). Differential microRNA regulation of HLA-C expression and its association with HIV control. Nature *472*, 495–498.

Kulkarni, S., Qi, Y., O'hUigin, C., Pereyra, F., Ramsuran, V., McLaren, P., Fellay, J., Nelson, G., Chen, H., Liao, W., et al. (2013). Genetic interplay between HLA-C and MIR148A in HIV control and Crohn disease. Proc. Natl. Acad. Sci. *110*, 20705–20710.

Lanier, L.L. (2004). NK Cell Recognition. Annu. Rev. Immunol. *23*, 225–274.

Lazaryan, A., Lobashevsky, E., Mulenga, J., Karita, E., Allen, S., Tang, J., and Kaslow, R.A. (2006). Human Leukocyte Antigen B58 Supertype and Human Immunodeficiency Virus Type 1 Infection in Native Africans. J. Virol. *80*, 6056–6060.

Lazaryan, A., Song, W., Lobashevsky, E., Tang, J., Shrestha, S., Zhang, K., Gardner, L.I., McNicholl, J.M., Wilson, C.M., Klein, R.S., et al. (2010). Human Leukocyte Antigen Class I Supertypes and HIV-1 Control in African Americans. J. Virol. *84*, 2610–2617.

Leslie, A., Matthews, P.C., Listgarten, J., Carlson, J.M., Kadie, C., Ndung'u, T., Brander, C., Coovadia, H., Walker, B.D., Heckerman, D., et al. (2010). Additive Contribution of HLA Class I Alleles in the Immune Control of HIV-1 Infection. J. Virol. *84*, 9879–9888.

Lewontin, R.C. (1964). The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. Genetics *49*, 49–67.

Li, H., Pascal, V., Martin, M.P., Carrington, M., and Anderson, S.K. (2008). Genetic Control of Variegated KIR Gene Expression: Polymorphisms of the Bi-Directional KIR3DL1 Promoter Are Associated with Distinct Frequencies of Gene Expression. PLoS Genet. *4*, e1000254.

Liu, J., Xiao, Z., Ko, H.L., Shen, M., and Ren, E.C. (2014). Activating killer cell immunoglobulin-like receptor 2DS2 binds to HLA-A*11. Proc. Natl. Acad. Sci. *111*, 2662–2667.

Long, B.R., Ndhlovu, L.C., Oksenberg, J.R., Lanier, L.L., Hecht, F.M., Nixon, D.F., and Barbour, J.D. (2008). Conferral of Enhanced Natural Killer Cell Function by KIR3DS1 in Early Human Immunodeficiency Virus Type 1 Infection. J. Virol. *82*, 4785–4792.

López-Vázquez, A., Rodrigo, L., Martínez-Borra, J., Pérez, R., Rodríguez, M., Fdez-Morera, J.L., Fuentes, D., Rodríguez-Rodero, S., González, S., and López-Larrea, C. (2005). Protective Effect of the HLA-Bw4I80 Epitope and the Killer Cell Immunoglobulin-Like Receptor 3DS1 Gene against the Development of Hepatocellular Carcinoma in Patients with Hepatitis C Virus Infection. J. Infect. Dis. *192*, 162–165.

Macdonald, W.A., Purcell, A.W., Mifsud, N.A., Ely, L.K., Williams, D.S., Chang, L., Gorman, J.J., Clements, C.S., Kjer-Nielsen, L., Koelle, D.M., et al. (2003). A Naturally Selected Dimorphism within the HLA-B44 Supertype Alters Class I Structure, Peptide Repertoire, and T Cell Recognition. J. Exp. Med. *198*, 679–691.

Van Manen, D., Kootstra, N.A., Boeser-Nunnink, B., Handulle, M.A., van't Wout, A.B., and Schuitemaker, H. (2009). Association of HLA-C and HCP5 gene regions with the clinical course of HIV-1 infection. AIDS *23*, 19–28.

Martin, A.M., Freitas, E.M., Witt, C.S., and Christiansen, F.T. (2000). The genomic organization and evolution of the natural killer immunoglobulin-like receptor (KIR) gene cluster. Immunogenetics *51*, 268–280.

Martin, A.M., Kulski, J.K., Gaudieri, S., Witt, C.S., Freitas, E.M., Trowsdale, J., and Christiansen, F.T. (2004). Comparative genomic analysis, diversity and evolution of two KIR haplotypes A and B. Gene *335*, 121–131.

Martin, M.P., Nelson, G., Lee, J.-H., Pellett, F., Gao, X., Wade, J., Wilson, M.J., Trowsdale, J., Gladman, D., and Carrington, M. (2002a). Cutting Edge: Susceptibility to Psoriatic Arthritis: Influence of Activating Killer Ig-Like Receptor Genes in the Absence of Specific HLA-C Alleles. J. Immunol. *169*, 2818–2822.

Martin, M.P., Gao, X., Lee, J.-H., Nelson, G.W., Detels, R., Goedert, J.J., Buchbinder, S., Hoots, K., Vlahov, D., Trowsdale, J., et al. (2002b). Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. Nat Genet *31*, 429–434.

Martin, M.P., Qi, Y., Gao, X., Yamada, E., Martin, J.N., Pereyra, F., Colombo, S., Brown, E.E., Shupert, W.L., Phair, J., et al. (2007). Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. Nat Genet *39*, 733–740.

Maxwell, L.D., Wallace, A., Middleton, D., and Curran, M.D. (2002). A common KIR2DS4 deletion variant in the human that predicts a soluble KIR molecule analogous to the KIR1D molecule observed in the rhesus monkey. Tissue Antigens *60*, 254–258.

McLaren, P.J., Ripke, S., Pelak, K., Weintrob, A.C., Patsopoulos, N.A., Jia, X., Erlich, R.L., Lennon, N.J., Kadie, C.M., Heckerman, D., et al. (2012). Fine-mapping classical HLA variation associated with durable host control of HIV-1 infection in African Americans. Hum. Mol. Genet. *21*, 4334–4347.

Merino, A., Malhotra, R., Morton, M., Mulenga, J., Allen, S., Hunter, E., Tang, J., and Kaslow, R.A. (2011). Impact of a Functional KIR2DS4 Allele on Heterosexual HIV-1 Transmission among Discordant Zambian Couples. J. Infect. Dis. *203*, 487–495.

Merino, A.M., Dugast, A.-S., Wilson, C.M., Goepfert, P.A., Alter, G., Kaslow, R.A., and Tang, J. (2014). KIR2DS4 Promotes HIV-1 Pathogenesis: New Evidence from Analyses of Immunogenetic Data and Natural Killer Cell Function. PLoS ONE *9*, e99353.

Middleton, D., Gonzalez, A., and Gilmore, P.M. (2007). Studies on the Expression of the Deleted KIR2DS4*003 Gene Product and Distribution of KIR2DS4 Deleted and Nondeleted Versions in Different Populations. Hum. Immunol. *68*, 128–134.

Migueles, S.A., Laborico, A.C., Imamichi, H., Shupert, W.L., Royce, C., McLaughlin, M., Ehler, L., Metcalf, J., Liu, S., Hallahan, C.W., et al. (2003). The Differential Ability of HLA B*5701+ Long-Term Nonprogressors and Progressors To Restrict Human Immunodeficiency Virus Replication Is Not Caused by Loss of Recognition of Autologous Viral gag Sequences. J. Virol. *77*, 6889–6898.

Miura, T., Brockman, M.A., Schneidewind, A., Lobritz, M., Pereyra, F., Rathod, A., Block, B.L., Brumme, Z.L., Brumme, C.J., Baker, B., et al. (2009). HLA-B57/B*5801 Human Immunodeficiency Virus Type 1 Elite Controllers Select for Rare Gag Variants Associated with Reduced Viral Replication Capacity and Strong Cytotoxic T-Lymphotye Recognition. J. Virol. *83*, 2743–2755.

Moesta, A.K., and Parham, P. (2012). Diverse functionality among human NK cell receptors for the C1 epitope of HLA-C: KIR2DS2, KIR2DL2, and KIR2DL3. Front. Immunol. *3*, 336.

Moesta, A.K., Norman, P.J., Yawata, M., Yawata, N., Gleimer, M., and Parham, P. (2008). Synergistic Polymorphism at Two Positions Distal to the Ligand-Binding Site Makes KIR2DL2 a Stronger Receptor for HLA-C Than KIR2DL3. J. Immunol. *180*, 3969–3979.

Moesta, A.K., Graef, T., Abi-Rached, L., Older Aguilar, A.M., Guethlein, L.A., and Parham, P. (2010). Humans Differ from Other Hominids in Lacking an Activating NK Cell Receptor That Recognizes the C1 Epitope of MHC Class I. J. Immunol. *185*, 4233–4237.

Mulrooney, T.J., Posch, P.E., and Hurley, C.K. (2013). DAP12 impacts trafficking and surface stability of killer immunoglobulin-like receptors on natural killer cells. J. Leukoc. Biol. *94*, 301–313.

Nakimuli, A., Chazara, O., Farrell, L., Hiby, S., Tukwasibwe, S., Knee, O., Jayaraman, J., Traherne, J., Elliott, A., Kaleebu, P., et al. (2013). Killer cell immunoglobulin-like receptor (KIR) genes and their HLA-C ligands in a Ugandan population. Immunogenetics *65*, 765–775.

Nakimuli, A., Chazara, O., Hiby, S.E., Farrell, L., Tukwasibwe, S., Jayaraman, J., Traherne, J.A., Trowsdale, J., Colucci, F., Lougee, E., et al. (2015). A KIR B centromeric region present in Africans but not Europeans protects pregnant women from pre-eclampsia. Proc. Natl. Acad. Sci. *112*, 845–850.

Norman, P.J., Abi-Rached, L., Gendzekhadze, K., Korbel, D., Gleimer, M., Rowley, D., Bruno, D., Carrington, C.V.F., Chandanayingyong, D., Chang, Y.-H., et al. (2007). Unusual selection on the KIR3DL1/S1 natural killer cell receptor in Africans. Nat Genet *39*, 1092–1099.

Norman, P.J., Hollenbach, J.A., Nemat-Gorgani, N., Guethlein, L.A., Hilton, H.G., Pando, M.J., Koram, K.A., Riley, E.M., Abi-Rached, L., and Parham, P. (2013). Co-evolution of Human Leukocyte Antigen (HLA) Class I Ligands with Killer-Cell Immunoglobulin-Like Receptors (KIR) in a Genetically Diverse Population of Sub-Saharan Africans. PLoS Genet. *9*, e1003938.

Nunes, J.M., Buhler, S., Roessli, D., Sanchez-Mazas, A., and the HLA-net 2013 collaboration (2014). The HLA-net GENE[RATE] pipeline for effective HLA data analysis and its application to 145 population samples from Europe and neighbouring areas. Tissue Antigens *83*, 307–323.

O'Connor, G.M., Guinan, K.J., Cunningham, R.T., Middleton, D., Parham, P., and Gardiner, C.M. (2007). Functional Polymorphism of the KIR3DL1/S1 Receptor on Human NK Cells. J. Immunol. *178*, 235–241.

O'Connor, G.M., Yamada, E., Rampersaud, A., Thomas, R., Carrington, M., and McVicar, D.W. (2011). Analysis of Binding of KIR3DS1*014 to HLA Suggests Distinct Evolutionary History of KIR3DS1. J. Immunol. *187*, 2162–2171.

O'hUigin, C., Kulkarni, S., Xu, Y., Deng, Z., Kidd, J., Kidd, K., Gao, X., and Carrington, M. (2011). The Molecular Origin and Consequences of Escape from miRNA Regulation by HLA-C Alleles. Am. J. Hum. Genet. *89*, 424–431.

Older Aguilar, A.M., Guethlein, L.A., Adams, E.J., Abi-Rached, L., Moesta, A.K., and Parham, P. (2010). Coevolution of Killer Cell Ig-Like Receptors with HLA-C To Become the Major Variable Regulators of Human NK Cells. J. Immunol. *185*, 4238–4251.

Pando, M.J., Gardiner, C.M., Gleimer, M., McQueen, K.L., and Parham, P. (2003). The Protein Made from a Common Allele of KIR3DL1 (3DL1*004) Is Poorly Expressed at Cell Surfaces due to Substitution at Positions 86 in Ig Domain 0 and 182 in Ig Domain 1. J. Immunol. *171*, 6640–6649.

Parham, P., Norman, P.J., Abi-Rached, L., and Guethlein, L.A. (2012). Human-specific evolution of killer cell immunoglobulin-like receptor recognition of major histocompatibility complex class I molecules. Philos. Trans. R. Soc. B Biol. Sci. *367*, 800–811.

Park, B., Lee, S., Kim, E., and Ahn, K. (2003). A Single Polymorphic Residue Within the Peptide-Binding Cleft of MHC Class I Molecules Determines Spectrum of Tapasin Dependence. J. Immunol. *170*, 961–968.

Parsons, M.S., Zipperlen, K., Gallant, M., and Grant, M. (2010). Killer cell immunoglobulin-like receptor 3DL1 licenses CD16-mediated effector functions of natural killer cells. J. Leukoc. Biol. *88*, 905–912.

Parsons, M.S., Wren, L., Isitman, G., Navis, M., Stratov, I., Bernard, N.F., and Kent, S.J. (2012). HIV Infection Abrogates the Functional Advantage of Natural Killer Cells Educated through KIR3DL1/HLA-Bw4 Interactions To Mediate Anti-HIV Antibody-Dependent Cellular Cytotoxicity. J. Virol. *86*, 4488–4495.

Paximadis, M., Minevich, G., Winchester, R., Schramm, D.B., Gray, G.E., Sherman, G.G., Coovadia, A.H., Kuhn, L., and Tiemessen, C.T. (2011). KIR-HLA and Maternal-Infant HIV-1 Transmission in Sub-Saharan Africa. PLoS ONE *6*, e16541.

Paximadis, M., Mathebula, T.Y., Gentle, N.L., Vardas, E., Colvin, M., Gray, C.M., Tiemessen, C.T., and Puren, A. (2012). Human leukocyte antigen class I (A, B, C) and II (DRB1) diversity in the black and Caucasian South African population. Hum. Immunol. *73*, 80–92.

Payne, R., Muenchhoff, M., Mann, J., Roberts, H.E., Matthews, P., Adland, E., Hempenstall, A., Huang, K.-H., Brockman, M., Brumme, Z., et al. (2014). Impact of HLA-driven HIV adaptation on virulence in populations of high HIV seroprevalence. Proc. Natl. Acad. Sci. *111*, E5393–E5400.

Pelak, K., Goldstein, D.B., Walley, N.M., Fellay, J., Ge, D., Shianna, K.V., Gumbs, C., Gao, X., Maia, J.M., Cronin, K.D., et al. (2010). Host Determinants of HIV-1 Control in African Americans. J. Infect. Dis. *201*, 1141–1149.

Pelak, K., Need, A.C., Fellay, J., Shianna, K.V., Feng, S., Urban, T.J., Ge, D., De Luca, A., Martinez-Picado, J., Wolinsky, S.M., et al. (2011). Copy Number Variation of KIR Genes Influences HIV-1 Control. PLoS Biol. *9*, e1001208.

Petitdemange, C., Wauquier, N., Jacquet, J.-M., Theodorou, I., Leroy, E., and Vieillard, V. (2014). Association of HLA Class-I and Inhibitory KIR Genotypes in Gabonese Patients Infected by Chikungunya or Dengue Type-2 Viruses. PLoS ONE *9*, e108798.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. Am. J. Hum. Genet. *81*, 559–575.

Pyo, C.-W., Guethlein, L.A., Vu, Q., Wang, R., Abi-Rached, L., Norman, P.J., Marsh, S.G.E., Miller, J.S., Parham, P., and Geraghty, D.E. (2010). Different Patterns of Evolution in the Centromeric and Telomeric Regions of Group A and B Haplotypes of the Human Killer Cell Ig-Like Receptor Locus. PLoS ONE *5*, e15115.

Pyo, C.-W., Wang, R., Vu, Q., Cereb, N., Yang, S.Y., Duh, F.-M., Wolinsky, S., Martin, M.P., Carrington, M., and Geraghty, D.E. (2013). Recombinant structures expand and contract inter and intragenic diversification at the KIR locus. BMC Genomics *14*, 89–89.

Quintana-Murci, L., Semino, O., Bandelt, H.-J., Passarino, G., McElreavey, K., and Santachiara-Benerecetti, A.S. (1999). Genetic evidence of an early exit of Homo sapiens sapiens from Africa through eastern Africa. Nat Genet *23*, 437–441.

Rajagopalan, S., and Long, E.O. (1999). A Human Histocompatibility Leukocyte Antigen (HLA)-G–specific Receptor Expressed on All Natural Killer Cells. J. Exp. Med. *189*, 1093–1100.

Rajagopalan, S., Fu, J., and Long, E.O. (2001). Cutting Edge: Induction of IFN-γ Production but Not Cytotoxicity by the Killer Cell Ig-Like Receptor KIR2DL4 (CD158d) in Resting NK Cells. J. Immunol. *167*, 1877–1881.

Reich, D.E., Cargill, M., Bolk, S., Ireland, J., Sabeti, P.C., Richter, D.J., Lavery, T., Kouyoumjian, R., Farhadian, S.F., Ward, R., et al. (2001). Linkage disequilibrium in the human genome. Nature *411*, 199–204.

Rizvi, S.M., Salam, N., Geng, J., Qi, Y., Bream, J.H., Duggal, P., Hussain, S.K., Martinson, J., Wolinsky, S.M., Carrington, M., et al. (2014). Distinct Assembly Profiles of HLA-B Molecules. J. Immunol. *192*, 4967–4976.

Robinson, J., Halliwell, J.A., Hayhurst, J.D., Flicek, P., Parham, P., and Marsh, S.G.E. (2014). The IPD and IMGT/HLA database: allele variant databases. Nucleic Acids Res.

Rosenberg, N.A., Pritchard, J.K., Weber, J.L., Cann, H.M., Kidd, K.K., Zhivotovsky, L.A., and Feldman, M.W. (2002). Genetic Structure of Human Populations. Science *298*, 2381–2385.

Sambrook, J.G., Bashirova, A., Palmer, S., Sims, S., Trowdale, J., Abi-Rached, L., Parham, P., Carrington, M. and Beck, S. (2005). Single haplotype analysis demonstrates rapid evolution of the killer immunoglobulin-like receptor (*KIR*) loci in primates. Genome Res. *15*, 25-35

Sáez-Cirión, A., Lacabaratz, C., Lambotte, O., Versmisse, P., Urrutia, A., Boufassa, F., Barré-Sinoussi, F., Delfraissy, J.-F., Sinet, M., Pancino, G., et al. (2007). HIV controllers exhibit potent CD8 T cell capacity to suppress HIV infection ex vivo and peculiar cytotoxic T lymphocyte activation phenotype. Proc. Natl. Acad. Sci. *104*, 6776–6781.

Salter, R.D., Benjamin, R.J., Wesley, P.K., Buxton, S.E., Garrett, T.P.J., Clayberger, C., Krensky, A.M., Norment, A.M., Littman, D.R., and Parham, P. (1990). A binding site for the T-cell co-receptor CD8 on the [alpha]3 domain of HLA-A2. Nature *345*, 41–46.

Sanjanwala, B., Draghi, M., Norman, P.J., Guethlein, L.A., and Parham, P. (2008). Polymorphic Sites Away from the Bw4 Epitope That Affect Interaction of Bw4+ HLA-B with KIR3DL1. J. Immunol. *181*, 6293–6300.

Shostakovich-Koretskaya, L., Catano, G., Chykarenko, Z.A., He, W., Gornalusse, G., Mummidi, S., Sanchez, R., Dolan, M.J., Ahuja, S.S., Clark, R.A., et al. (2009). Combinatorial content of CCL3L and CCL4L gene copy numbers influence HIV-AIDS susceptibility in Ukrainian children. AIDS Lond. Engl. *23*, 679–688.

Shrestha, S., Aissani, B., Song, W., Wilson, C.M., Kaslow, R.A., and Tang, J. (2009). Host Genetics and HIV-1 Viral Load Set-point in African-Americans. AIDS Lond. Engl. *23*, 673–677.

Single, R., Martin, M., Meyer, D., Gao, X., and Carrington, M. (2008). Methods for assessing gene content diversity of KIR with examples from a global set of populations. Immunogenetics *60*, 711–725.

Single, R.M., Martin, M.P., Gao, X., Meyer, D., Yeager, M., Kidd, J.R., Kidd, K.K., and Carrington, M. (2007). Global diversity and evidence for coevolution of KIR and HLA. Nat Genet *39*, 1114–1119.

Slatkin, M. (1994). Linkage disequilibrium in growing and stable populations. Genetics *137*, 331–336.

Van der Slik, A.R., Koeleman, B.P.C., Verduijn, W., Bruining, G.J., Roep, B.O., and Giphart, M.J. (2003). KIR in Type 1 Diabetes: Disparate Distribution of Activating and Inhibitory Natural Killer Cell Receptors in Patients Versus HLA-Matched Control Subjects. Diabetes *52*, 2639–2642.

Smith, K.J., Reid, S.W., Harlos, K., McMichael, A.J., Stuart, D.I., Bell, J.I., and Jones, E.Y. (1996). Bound Water Structure and Polymorphic Amino Acids Act Together to Allow the Binding of Different Peptides to MHC Class I HLA-B53. Immunity *4*, 215–228.

Song, R., Lisovsky, I., Lebouché, B., Routy, J.-P., Bruneau, J., and Bernard, N.F. (2014). HIV Protective KIR3DL1/S1-HLA-B Genotypes Influence NK Cell-Mediated Inhibition of HIV Replication in Autologous CD4 Targets. PLoS Pathog. *10*, e1003867.

Steiner, N., Dakshanamurthy, S., VandenBussche, C., and Hurley, C. (2008). Extracellular domain alterations impact surface expression of stimulatory natural killer cell receptor KIR2DS5. Immunogenetics *60*, 655–667.

Steiner, N.K., Dakshanamurthy, S., Nguyen, N., and Hurley, C.K. (2014). Allelic variation of killer cell immunoglobulin-like receptor 2DS5 impacts glycosylation altering cell surface expression levels. Hum. Immunol. *75*, 124–128.

Stephens, J.C., Schneider, J.A., Tanguay, D.A., Choi, J., Acharya, T., Stanley, S.E., Jiang, R., Messer, C.J., Chew, A., Han, J.-H., et al. (2001). Haplotype Variation and Linkage Disequilibrium in 313 Human Genes. Science *293*, 489–493.

Stern, M., Ruggeri, L., Capanni, M., Mancusi, A., and Velardi, A. (2008). Human leukocyte antigens A23, A24, and A32 but not A25 are ligands for KIR3DL1. Blood *112*, 708–710.

Stewart, C.A., Laugier-Anfossi, F., Vély, F., Saulquin, X., Riedmuller, J., Tisserant, A., Gauthier, L., Romagné, F., Ferracci, G., Arosa, F.A., et al. (2005). Recognition of peptide–MHC class I complexes by activating killer immunoglobulin-like receptors. Proc. Natl. Acad. Sci. U. S. A. *102*, 13224–13229.

Stranger, B.E., Forrest, M.S., Clark, A.G., Minichiello, M.J., Deutsch, S., Lyle, R., Hunt, S., Kahl, B., Antonarakis, S.E., Tavaré, S., et al. (2005). Genome-Wide Associations of Gene Expression Variation in Humans. PLoS Genet. *1*, e78.

Stulberg, M.J., Wright, P.W., Dang, H., Hanson, R.J., Miller, J.S., and Anderson, S.K. (2006). Identification of distal KIR promoters and transcripts. Genes Immun *8*, 124–130.

Sun, J.C., Beilke, J.N., and Lanier, L.L. (2009). Adaptive immune features of natural killer cells. Nature *457*, 557–561.

Tang, J., Cormier, E., Gilmour, J., Price, M.A., Prentice, H.A., Song, W., Kamali, A., Karita, E., Lakhi, S., Sanders, E.J., et al. (2011). Human Leukocyte Antigen Variants B*44 and B*57 Are Consistently Favorable during Two Distinct Phases of Primary HIV-1 Infection in Sub-Saharan Africans with Several Viral Subtypes. J. Virol.

Taniguchi, M., and Kawabata, M. (2009). KIR3DL1/S1 genotypes and KIR2DS4 allelic variants in the AB KIR genotypes are associated with Plasmodium-positive individuals in malaria infection. Immunogenetics *61*, 717–730.

Thammavongsa, V., Raghuraman, G., Filzen, T.M., Collins, K.L., and Raghavan, M. (2006). HLA-B44 Polymorphisms at Position 116 of the Heavy Chain Influence TAP Complex Binding via an Effect on Peptide Occupancy. J. Immunol. *177*, 3150–3161.

Thananchai, H., Gillespie, G., Martin, M.P., Bashirova, A., Yawata, N., Yawata, M., Easterbrook, P., McVicar, D.W., Maenaka, K., Parham, P., et al. (2007). Cutting Edge: Allele-Specific and Peptide-Dependent Interactions between KIR3DL1 and HLA-A and HLA-B. J. Immunol. *178*, 33–37.

The International HIV Controllers Study, Writing team, Pereyra, F., Jia, X., McLaren, P.J., Telenti, A., de Bakker, P.I., Walker, B.D., Analysis team, Jia, X., et al. (2010). The Major Genetic Determinants of HIV-1 Control Affect HLA Class I Peptide Presentation. Science *330*, 1551–1557.

The MHC sequencing consortium (1999). Complete sequence and gene map of a human major histocompatibility complex. Nature *401*, 921–923.

Thomas, R., Yamada, E., Alter, G., Martin, M.P., Bashirova, A.A., Norman, P.J., Altfeld, M., Parham, P., Anderson, S.K., McVicar, D.W., et al. (2008). Novel KIR3DL1 Alleles and Their Expression Levels on NK Cells: Convergent Evolution of KIR3DL1 Phenotype Variation? J. Immunol. *180*, 6743–6750.

Thomas, R., Apps, R., Qi, Y., Gao, X., Male, V., O'hUigin, C., O'Connor, G., Ge, D., Fellay, J., Martin, J.N., et al. (2009). HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. Nat Genet *41*, 1290–1294.

Tiemessen, C.T., Shalekoff, S., Meddows-Taylor, S., Schramm, D.B., Papathanasopoulos, M.A., Gray, G.E., Sherman, G.G., Coovadia, A.H., and Kuhn, L. (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.

Tiemessen, C.T., Paximadis, M., Minevich, G., Winchester, R., Shalekoff, S., Gray, G.E., Sherman, G.G., Coovadia, A.H., and Kuhn, L. (2011). Natural Killer Cell Responses to HIV-1 Peptides are Associated With More Activating KIR Genes and HLA-C Genes of the C1 Allotype. J. Acquir. Immune Defic. Syndr. 1999 *57*, 181–189.

Tishkoff, S.A., and Williams, S.M. (2002). Genetic analysis of African populations: human evolution and complex disease. Nat Rev Genet *3*, 611–621.

Tishkoff, S.A. and Verrelli, B.C. (2003) Patterns of human genetic diversity: Implications for human evolutionary history and disease. Ann. Rev. Genom. Hum. Genet. *4*, 293-340

Tishkoff, S.A., Reed, F.A., Friedlaender, F.R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J.B., Awomoyi, A.A., Bodo, J.-M., Doumbo, O., et al. (2009). The Genetic Structure and History of Africans and African Americans. Science *324*, 1035–1044.

Tomescu, C., Duh, F.-M., Hoh, R., Viviani, A., Harvill, K., Martin, M.P., Carrington, M., Deeks, S.G., and Montaner, L.J. (2012). Impact of protective killer inhibitory receptor/human leukocyte antigen genotypes on natural killer cell and T-cell function in HIV-1-infected controllers. AIDS *26*, 1869–1878.

Trompeter, H.-I., Gómez-Lozano, N., Santourlidis, S., Eisermann, B., Wernet, P., Vilches, C., and Uhrberg, M. (2005). Three Structurally and Functionally Divergent Kinds of Promoters Regulate Expression of Clonally Distributed Killer Cell Ig-Like Receptors (KIR), of KIR2DL4, and of KIR3DL3. J. Immunol. *174*, 4135–4143.

Uhrberg, M., Valiante, N.M., Shum, B.P., Shilling, H.G., Lienert-Weidenbach, K., Corliss, B., Tyan, D., Lanier, L.L., and Parham, P. (1997). Human Diversity in Killer Cell Inhibitory Receptor Genes. Immunity *7*, 753–763.

Uhrberg, M., Parham, P., and Wernet, P. (2002). Definition of gene content for nine common group B haplotypes of the Caucasoid population: KIR haplotypes contain between seven and eleven KIR genes. Immunogenetics *54*, 221–229.

Valiante, N.M., Uhrberg, M., Shilling, H.G., Lienert-Weidenbach, K., Arnett, K.L., D'Andrea, A., Phillips, J.H., Lanier, L.L., and Parham, P. (1997). Functionally and Structurally Distinct NK Cell Receptor Repertoires in the Peripheral Blood of Two Human Donors. Immunity *7*, 739–751.

VandenBussche, C.J., Dakshanamurthy, S., Posch, P.E., and Hurley, C.K. (2006). A Single Polymorphism Disrupts the Killer Ig-Like Receptor 2DL2/2DL3 D1 Domain. J. Immunol. *177*, 5347–5357.

VandenBussche, C.J., Mulrooney, T.J., Frazier, W.R., Dakshanamurthy, S., and Hurley, C.K. (2008). Dramatically reduced surface expression of NK cell receptor KIR2DS3 is attributed to multiple residues throughout the molecule. Genes Immun *10*, 162–173.

Vierra-Green, C., Roe, D., Hou, L., Hurley, C.K., Rajalingam, R., Reed, E., Lebedeva, T., Yu, N., Stewart, M., Noreen, H., et al. (2012). Allele-Level Haplotype Frequencies and Pairwise Linkage Disequilibrium for 14 KIR Loci in 506 European-American Individuals. PLoS ONE *7*, e47491.

Vilches, C., Rajalingam, R., Uhrberg, M., Gardiner, C.M., Young, N.T., and Parham, P. (2000). KIR2DL5, a Novel Killer-Cell Receptor with a D0-D2 Configuration of Ig-Like Domains. J. Immunol. *164*, 5797–5804.

Walker, B.D., and Yu, X.G. (2013). Unravelling the mechanisms of durable control of HIV-1. Nat Rev Immunol *13*, 487–498.

Wilson, M.J., Torkar, M., Haude, A., Milne, S., Jones, T., Sheer, D., Beck, S., and Trowsdale, J. (2000). Plasticity in the organization and sequences of human KIR/ILT gene families. Proc. Natl. Acad. Sci. *97*, 4778–4783.

Winkler, C., An, P. and O'Brien, S.J. (2004). Patterns of ethnic diversity among the genes that influence AIDS. Hum. Mol. Genet. *13*, R9-R19

Winter, C.C., and Long, E.O. (1997). A single amino acid in the p58 killer cell inhibitory receptor controls the ability of natural killer cells to discriminate between the two groups of HLA-C allotypes. J. Immunol. *158*, 4026–4028.

Winter, C.C., Gumperz, J.E., Parham, P., Long, E.O., and Wagtmann, N. (1998). Direct Binding and Functional Transfer of NK Cell Inhibitory Receptors Reveal Novel Patterns of HLA-C Allotype Recognition. J. Immunol. *161*, 571–577.

Xu, Y., Deng, Z., O'hUigin, C., Wang, D., Gao, S., Zeng, J., Yang, B., Jin, S., and Zou, H. (2011). Characterization and polymorphic analysis of 4.5 kb genomic full-length HLA-C in the Chinese Han population. Tissue Antigens *78*, 102–114.

Yawata, M., Yawata, N., McQueen, K., Cheng, N., Guethlein, L., Rajalingam, R., Shilling, H., and Parham, P. (2002). Predominance of group A KIR haplotypes in Japanese associated with diverse NK cell repertoires of KIR expression. Immunogenetics *54*, 543–550.

Yoo, Y.J., Tang, J., Kaslow, R.A., and Zhang, K. (2007). Haplotype inference for present–absent genotype data using previously identified haplotypes and haplotype patterns. Bioinformatics *23*, 2399–2406.

Yusa, S., and Campbell, K.S. (2003). Src Homology Region 2-Containing Protein Tyrosine Phosphatase-2 (SHP-2) Can Play a Direct Role in the Inhibitory Function of Killer Cell Ig-Like Receptors in Human NK Cells. J. Immunol. *170*, 4539–4547.

Yusa, S., Catina, T.L., and Campbell, K.S. (2004). KIR2DL5 Can Inhibit Human NK Cell Activation Via Recruitment of Src Homology Region 2-Containing Protein Tyrosine Phosphatase-2 (SHP-2). J. Immunol. *172*, 7385–7392.

Zernich, D., Purcell, A.W., Macdonald, W.A., Kjer-Nielsen, L., Ely, L.K., Laham, N., Crockford, T., Mifsud, N.A., Bharadwaj, M., Chang, L., et al. (2004). Natural HLA Class I Polymorphism Controls the Pathway of Antigen Presentation and Susceptibility to Viral Evasion. J. Exp. Med. *200*, 13–24.

Zhuang, Y.-L., Zhu, C.-F., Zhang, Y., Song, Y.-H., Wang, D.-J., Nie, X.-M., Liu, Y., and Ren, G.-J. (2012). Association of KIR2DS4 and its variant KIR1D with syphilis in a Chinese Han population. Int. J. Immunogenet. *39*, 114–118.

Zuñiga, R., Lucchetti, A., Galvan, P., Sanchez, S., Sanchez, C., Hernandez, A., Sanchez, H., Frahm, N., Linde, C.H., Hewitt, H.S., et al. (2006). Relative Dominance of Gag p24-Specific Cytotoxic T Lymphocytes Is Associated with Human Immunodeficiency Virus Control. J. Virol. *80*, 3122–3125.

Appendix A

Turnitin Report

Turnitin Originality Report
turnitin.docx by Nikki Gentle
From PhD_Research
(9T5NEoFKOF6336g28aNnCeg30dOsUdRa91SaP18F5PMsvNimlyM44l5a4kyMuzKwR1M3lV3d8wUdr5SYgJ6y7Kv6021K6br6KOM)

- Processed on 27-Feb-2015 4:56 AM SAST
- ID: 510231497
- Word Count: 29926

Similarity Index 26%
Similarity by Source
Internet Sources: 6%
Publications: 24%
Student Papers: 4%

sources:

1

15% match (publications)
Gentle, Nikki L., Maria Paximadis, Adrian Puren, and Caroline T. Tiemessen. "Genetic Variability in Markers of HLA-C Expression in Two Diverse South African Populations", PLoS ONE, 2013.

2

1% match (publications)
Stephens, Henry A.F.. "Immunogenetic surveillance of HIV/AIDS", Infection Genetics and Evolution, 2012.

3

1% match (publications)
Chen, Haoyan, Genki Hayashi, Olivia Y. Lai, Alexander Dilthey, Peter J. Kuebler, Tami V. Wong, Maureen P. Martin, Marcelo A. Fernandez Vina, Gil McVean, Matthias Wabl, Kieron S. Leslie, Toby Maurer, Jeffrey N. Martin, Steven G. Deeks, Mary Carrington, Anne. "Psoriasis Patients Are Enriched for Genetic Variants That Protect against HIV-1 Disease", PLoS Genetics, 2012.

4

1% match (publications)
Paximadis, M.. "Human leukocyte antigen class I (A, B, C) and II (DRB1) diversity in the black and Caucasian South African population", Human Immunology, 201201

5

1% match (student papers from 01-Feb-2011)
Submitted to University of KwaZulu-Natal on 2011-02-01

Appendix B

Ethics Clearance

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M130872

NAME: (Principal Investigator) Ms Nikkie Gentle

DEPARTMENT: Centre for HIV and STIs/Virology
National Institue for Communicable Diseases

PROJECT TITLE: The Role of Killer-Immunoglobulin-like Receptors and Specific Human Leukocyte Antigen clss 1 Molecules in Control of Human Immunodeficiency Virus-1 Infection

DATE CONSIDERED: Ad hoc

DECISION: Approved unconditionally

CONDITIONS:

SUPERVISOR: Prof C Tiemessen

APPROVED BY:

Professor PE Cleaton-Jones, Chairperson, HREC (Medical)

DATE OF APPROVAL: 04/09/2013

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

DECLARATION OF INVESTIGATORS

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

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

Principal Investigator Signature M130872Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES