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 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:


Publications:


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 \textit{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 \textit{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 \textit{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 \textit{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 \textit{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
mm³    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
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 (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 β2-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 \(\alpha_1\) 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 \(\alpha\) 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).
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), it 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). KIR3DL1*h/KIR3DL1*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 been 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 \textit{KIR2DS4}, \textit{KIR3DL1} and \textit{KIR3DS1} allowed haplotypes containing the \textit{tA01} telomeric haplotype motif to be resolved with a fair degree of certainty. However, because \textit{KIR2DL5}, \textit{KIR2DS3} and \textit{KIR2DS5} can occur on both centromeric and telomeric \textit{KIR} haplotype motifs, in the absence of copy number information for these genes it was impossible to distinguish between the \textit{cA01ltB01} and \textit{cB03ltB01} 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 \textit{KIR2DL5}, \textit{KIR2DS3} and \textit{KIR2DS5} located within the telomeric region of the \textit{cA01ltB01} haplotype and found within the centromeric portion of \textit{cB03ltB01} (Pyo \textit{et al.}, 2010; Pyo \textit{et al.}, 2013). The frequencies of these haplotypes are thus reported as \textit{cA01/cB03ltB01}.

\subsection*{2.2.5 HLA Class I Ligand Determination}

\textit{HLA-A}, \textit{-B} and \textit{-C} genotype data were available for all 264 individuals (Paximadis \textit{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 \textit{et al.}, 2002a). A further distinction was made between HLA-B alleles with isoleucine at amino acid position 80 (\textit{Bw4-80I}) and those with threonine at this position (\textit{Bw4-80T}), as these epitopes differ in the avidity of their interaction with \textit{KIR3DL1} (Cella \textit{et al.}, 1994; Martin \textit{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 \textit{et al.}, 1993; Winter and Long, 1997).

\subsection*{2.2.6 Statistical Analyses}

\textit{KIR} gene carrier- and \textit{HLA} class I ligand frequencies were determined by direct counting. \textit{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 r² (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 x 10⁻¹³) and KIR2DS4v being the most frequent in the Caucasian population group (p = 8.79 x 10⁻⁹). The two groups also displayed significantly different distributions of KIR2DS1 (p = 1.20 x 10⁻⁶) and KIR3DS1 (p = 5.43 x 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).

20
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

<table>
<thead>
<tr>
<th>KIR Gene</th>
<th>Black Individuals (n = 167)</th>
<th>Caucasian Individuals (n = 97)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Frequency (%)</td>
<td>n</td>
</tr>
<tr>
<td>KIR2DL1</td>
<td>166</td>
<td>99</td>
<td>96</td>
</tr>
<tr>
<td>KIR2DL2</td>
<td>114</td>
<td>68</td>
<td>56</td>
</tr>
<tr>
<td>KIR2DL3</td>
<td>137</td>
<td>82</td>
<td>86</td>
</tr>
<tr>
<td>KIR2DL5</td>
<td>102</td>
<td>61</td>
<td>49</td>
</tr>
<tr>
<td>KIR2DS1</td>
<td>21</td>
<td>13</td>
<td>38</td>
</tr>
<tr>
<td>KIR2DS2</td>
<td>100</td>
<td>60</td>
<td>55</td>
</tr>
<tr>
<td>KIR2DS3</td>
<td>44</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>KIR2DS4f</td>
<td>139</td>
<td>84</td>
<td>37</td>
</tr>
<tr>
<td>KIR2DS4v</td>
<td>83</td>
<td>50</td>
<td>82</td>
</tr>
<tr>
<td>KIR2DS5</td>
<td>72</td>
<td>43</td>
<td>38</td>
</tr>
<tr>
<td>KIR2DP1</td>
<td>166</td>
<td>99</td>
<td>96</td>
</tr>
<tr>
<td>KIR3DL1</td>
<td>166</td>
<td>99</td>
<td>95</td>
</tr>
<tr>
<td>KIR3DS1</td>
<td>12</td>
<td>7</td>
<td>38</td>
</tr>
</tbody>
</table>

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 cA01tA01 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

<table>
<thead>
<tr>
<th>KIR Genotype1</th>
<th>KIR2DL1</th>
<th>KIR2DL2</th>
<th>KIR2DL3</th>
<th>KIR2DL5</th>
<th>KIR2DS1</th>
<th>KIR2DS2</th>
<th>KIR2DS3</th>
<th>KIR2DS4</th>
<th>KIR2DS5</th>
<th>KIR2DL4</th>
<th>KIR2DL6</th>
<th>KIR2DL7</th>
<th>KIR3DL1</th>
<th>KIR3DL2</th>
<th>KIR3DL3</th>
<th>KIR3DL4</th>
<th>KIR3DL5</th>
<th>KIR3DL6</th>
<th>Frequency Black Individuals (n = 167)</th>
<th>Frequency Caucasian Individuals (n = 97)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.269 (45)2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.278 (27)</td>
</tr>
<tr>
<td>Bx 21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.114 (19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010 (1)</td>
</tr>
<tr>
<td>Bx 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.090 (15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.155 (15)</td>
</tr>
<tr>
<td>Bx 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.090 (15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.021 (2)</td>
</tr>
<tr>
<td>Bx 228</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.072 (12)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010 (1)</td>
</tr>
<tr>
<td>Bx 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.054 (9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.042 (7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.041 (4)</td>
</tr>
<tr>
<td>Bx 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.036 (6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.030 (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.062 (6)</td>
</tr>
<tr>
<td>Bx 112</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.030 (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.024 (4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.018 (3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.012 (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.113 (11)</td>
</tr>
<tr>
<td>Bx 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.103 (10)</td>
</tr>
<tr>
<td>Bx 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010 (1)</td>
</tr>
<tr>
<td>Bx 62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010 (1)</td>
</tr>
<tr>
<td>Bx 73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010 (1)</td>
</tr>
<tr>
<td>Bx 81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 89</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.031 (3)</td>
</tr>
<tr>
<td>Bx 92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010 (1)</td>
</tr>
<tr>
<td>Bx 169</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 172</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010 (1)</td>
</tr>
<tr>
<td>Bx 175</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 393</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.006 (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bx 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.021 (2)</td>
</tr>
<tr>
<td>Bx 70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.021 (2)</td>
</tr>
<tr>
<td>Bx 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010 (1)</td>
</tr>
<tr>
<td>Bx 15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010 (1)</td>
</tr>
<tr>
<td>Bx 24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010 (1)</td>
</tr>
<tr>
<td>Bx 58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.021 (1)</td>
</tr>
<tr>
<td>Bx 382</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010 (1)</td>
</tr>
<tr>
<td>Unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.010 (1)</td>
</tr>
</tbody>
</table>

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 \textit{KIR} haplotypes occurring in the Black and Caucasian South African population groups

<table>
<thead>
<tr>
<th>KIR Haplotype</th>
<th>Black Individuals (n = 167)</th>
<th>Caucasian Individuals (n = 97)</th>
<th>p-value$^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Frequency</td>
<td>n</td>
</tr>
<tr>
<td>cA01</td>
<td>tA01$^2$</td>
<td>159</td>
<td>0.476</td>
</tr>
<tr>
<td>cA01</td>
<td>tA01-hybd1</td>
<td>9</td>
<td>0.027</td>
</tr>
<tr>
<td>cA01</td>
<td>tA01-del5</td>
<td>8</td>
<td>0.024</td>
</tr>
<tr>
<td>cA01/cB03</td>
<td>tB01$^3$</td>
<td>2</td>
<td>0.006</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01</td>
<td>57</td>
<td>0.171</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-hybd1</td>
<td>14</td>
<td>0.042</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-ins4</td>
<td>0</td>
<td>0.000</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-del3</td>
<td>14</td>
<td>0.042</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-del3-ins4</td>
<td>0</td>
<td>0.000</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-del3-del5</td>
<td>1</td>
<td>0.003</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-del3-del5-del9</td>
<td>1</td>
<td>0.003</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-del9</td>
<td>32</td>
<td>0.096</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-del4-del9</td>
<td>2</td>
<td>0.006</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-del10</td>
<td>2</td>
<td>0.006</td>
</tr>
<tr>
<td>cB01</td>
<td>tB01</td>
<td>9</td>
<td>0.027</td>
</tr>
<tr>
<td>cB01</td>
<td>tB01-del5</td>
<td>1</td>
<td>0.003</td>
</tr>
<tr>
<td>cB01</td>
<td>tB01-del7</td>
<td>8</td>
<td>0.024</td>
</tr>
<tr>
<td>cB02</td>
<td>tA01-ins5</td>
<td>1</td>
<td>0.003</td>
</tr>
<tr>
<td>cB02</td>
<td>tA01-hybd1</td>
<td>3</td>
<td>0.009</td>
</tr>
<tr>
<td>cB03</td>
<td>tA01</td>
<td>5</td>
<td>0.015</td>
</tr>
<tr>
<td>cB03</td>
<td>tA01-hybd1</td>
<td>3</td>
<td>0.009</td>
</tr>
<tr>
<td>H1$^4$</td>
<td>2</td>
<td>0.006</td>
<td>0</td>
</tr>
<tr>
<td>H2</td>
<td>1</td>
<td>0.003</td>
<td>1</td>
</tr>
<tr>
<td>H3</td>
<td>0</td>
<td>0.000</td>
<td>1</td>
</tr>
</tbody>
</table>

$^1$ 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 \textit{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 \textit{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

<table>
<thead>
<tr>
<th>Telomeric Motif</th>
<th>Black Individuals (n = 167)</th>
<th>Caucasian Individuals (n = 97)</th>
<th>p-value$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Frequency</td>
<td>n</td>
</tr>
<tr>
<td>tA01</td>
<td>314</td>
<td>0.940</td>
<td>156</td>
</tr>
<tr>
<td>tA01f</td>
<td>193</td>
<td>0.619</td>
<td>45</td>
</tr>
<tr>
<td>tA01v</td>
<td>89</td>
<td>0.285</td>
<td>108</td>
</tr>
<tr>
<td>tA01-hybrd1</td>
<td>30</td>
<td>0.096</td>
<td>3</td>
</tr>
<tr>
<td>tB01</td>
<td>20</td>
<td>0.060</td>
<td>38</td>
</tr>
</tbody>
</table>

$^1$ 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

<table>
<thead>
<tr>
<th></th>
<th>KIR2DL1</th>
<th>KIR2DL2</th>
<th>KIR2DL3</th>
<th>KIR2DL5</th>
<th>KIR2DS1</th>
<th>KIR2DS2</th>
<th>KIR2DS3</th>
<th>KIR2DS4</th>
<th>KIR2DS5</th>
<th>KIR2DP1</th>
<th>KIR3DL1</th>
</tr>
</thead>
<tbody>
<tr>
<td>KIR2DL2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D' )</td>
<td>-0.650</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r^2 )</td>
<td>0.027</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p\text{-value})</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIR2DL3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D' )</td>
<td>0.647</td>
<td>-1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r^2 )</td>
<td>0.026</td>
<td>0.976</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p\text{-value})</td>
<td>0.004</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIR2DL5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D' )</td>
<td>-0.691</td>
<td>0.163</td>
<td>-0.805</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r^2 )</td>
<td>0.041</td>
<td>0.807</td>
<td>0.462</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p\text{-value})</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIR2DS1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D' )</td>
<td>-0.787</td>
<td>0.750</td>
<td>-0.748</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r^2 )</td>
<td>0.434</td>
<td>0.059</td>
<td>0.049</td>
<td>0.123</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p\text{-value})</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIR2DS2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D' )</td>
<td>-0.674</td>
<td>0.959</td>
<td>-0.986</td>
<td>0.613</td>
<td>0.767</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r^2 )</td>
<td>0.034</td>
<td>0.774</td>
<td>0.798</td>
<td>0.327</td>
<td>0.063</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p\text{-value})</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIR2DS3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D' )</td>
<td>0.482</td>
<td>0.919</td>
<td>-0.918</td>
<td>0.964</td>
<td>0.344</td>
<td>0.811</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r^2 )</td>
<td>0.002</td>
<td>0.167</td>
<td>0.162</td>
<td>0.251</td>
<td>0.054</td>
<td>0.154</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p\text{-value})</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>KIR2DS4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D' )</td>
<td>1.000</td>
<td>-0.475</td>
<td>0.470</td>
<td>-0.629</td>
<td>-1.000</td>
<td>-0.511</td>
<td>-0.070</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r^2 )</td>
<td>0.267</td>
<td>0.053</td>
<td>0.051</td>
<td>0.128</td>
<td>0.381</td>
<td>0.073</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p\text{-value})</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>KIR2DS5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D' )</td>
<td>-0.647</td>
<td>0.808</td>
<td>-0.806</td>
<td>0.981</td>
<td>0.874</td>
<td>0.583</td>
<td>-0.337</td>
<td>-0.629</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r^2 )</td>
<td>0.060</td>
<td>0.284</td>
<td>0.276</td>
<td>0.573</td>
<td>0.157</td>
<td>0.176</td>
<td>0.006</td>
<td>0.214</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p\text{-value})</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>NS</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>KIR2DP1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D' )</td>
<td>0.693</td>
<td>-1.000</td>
<td>1.000</td>
<td>0.485</td>
<td>-0.287</td>
<td>-0.963</td>
<td>1.000</td>
<td>0.118</td>
<td>0.259</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( r^2 )</td>
<td>0.149</td>
<td>0.203</td>
<td>0.198</td>
<td>0.020</td>
<td>0.036</td>
<td>0.224</td>
<td>0.022</td>
<td>0.012</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p\text{-value})</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.040</td>
<td>NS</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>KIR3DL1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D' )</td>
<td>0.709</td>
<td>-0.376</td>
<td>0.369</td>
<td>-0.558</td>
<td>-0.896</td>
<td>-0.302</td>
<td>-0.180</td>
<td>0.622</td>
<td>-0.574</td>
<td>0.177</td>
<td></td>
</tr>
<tr>
<td>( r^2 )</td>
<td>0.258</td>
<td>0.017</td>
<td>0.016</td>
<td>0.052</td>
<td>0.589</td>
<td>0.013</td>
<td>0.020</td>
<td>0.201</td>
<td>0.093</td>
<td>0.019</td>
<td></td>
</tr>
<tr>
<td>( p\text{-value})</td>
<td>&lt; 0.001</td>
<td>0.023</td>
<td>0.023</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.017</td>
<td>0.003</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.019</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>KIR3DS1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D' )</td>
<td>-0.239</td>
<td>0.682</td>
<td>-0.679</td>
<td>1.000</td>
<td>0.556</td>
<td>0.687</td>
<td>-1.000</td>
<td>0.880</td>
<td>1.000</td>
<td>-1.000</td>
<td></td>
</tr>
<tr>
<td>( r^2 )</td>
<td>0.041</td>
<td>0.021</td>
<td>0.020</td>
<td>0.062</td>
<td>0.508</td>
<td>0.017</td>
<td>0.109</td>
<td>0.193</td>
<td>0.081</td>
<td>0.005</td>
<td>0.372</td>
</tr>
<tr>
<td>( p\text{-value})</td>
<td>&lt; 0.001</td>
<td>0.003</td>
<td>0.004</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>NS</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

\(^1\) \text{p-value is for an exact test of linkage disequilibrium (Slatkin, 1994).}
<table>
<thead>
<tr>
<th></th>
<th>KIR2DL1</th>
<th>KIR2DL2</th>
<th>KIR2DL3</th>
<th>KIR2DL5</th>
<th>KIR2DS1</th>
<th>KIR2DS2</th>
<th>KIR2DS3</th>
<th>KIR2DS4</th>
<th>KIR2DS5</th>
<th>KIR2DP1</th>
<th>KIR3DL1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>KIR2DL2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D'$</td>
<td>0.596</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>0.032</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>0.011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KIR2DL3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D'$</td>
<td></td>
<td>-0.625</td>
<td>-1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td></td>
<td>0.040</td>
<td>0.892</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td></td>
<td>0.001</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KIR2DL5</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D'$</td>
<td></td>
<td>-0.953</td>
<td>0.293</td>
<td>-0.265</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td></td>
<td>0.399</td>
<td>0.073</td>
<td>0.053</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td></td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KIR2DS1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D'$</td>
<td></td>
<td>-0.916</td>
<td>-0.015</td>
<td>0.086</td>
<td>0.965</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td></td>
<td>0.591</td>
<td>0.000</td>
<td>0.001</td>
<td>0.581</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td></td>
<td>&lt; 0.001</td>
<td>NS</td>
<td>NS</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KIR2DS2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D'$</td>
<td></td>
<td>0.710</td>
<td>0.928</td>
<td>-0.977</td>
<td>0.223</td>
<td>-0.059</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td></td>
<td>0.049</td>
<td>0.804</td>
<td>0.912</td>
<td>0.039</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td></td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>NS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KIR2DS3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D'$</td>
<td></td>
<td>0.438</td>
<td>1.000</td>
<td>-1.000</td>
<td>1.000</td>
<td>0.398</td>
<td>0.801</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td></td>
<td>0.005</td>
<td>0.273</td>
<td>0.244</td>
<td>0.323</td>
<td>0.082</td>
<td>0.164</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td></td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KIR2DS4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D'$</td>
<td></td>
<td>0.958</td>
<td>-0.005</td>
<td>-0.063</td>
<td>-0.899</td>
<td>-1.000</td>
<td>0.035</td>
<td>-0.390</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td></td>
<td>0.607</td>
<td>0.000</td>
<td>0.001</td>
<td>0.537</td>
<td>0.940</td>
<td>0.000</td>
<td>0.074</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td></td>
<td>&lt; 0.001</td>
<td>NS</td>
<td>NS</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>NS</td>
<td>&lt; 0.001</td>
<td>NS</td>
<td>NS</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td><strong>KIR2DS5</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D'$</td>
<td></td>
<td>-0.914</td>
<td>0.084</td>
<td>-0.048</td>
<td>1.000</td>
<td>0.871</td>
<td>0.063</td>
<td>0.269</td>
<td>-0.815</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td></td>
<td>0.521</td>
<td>0.004</td>
<td>0.001</td>
<td>0.705</td>
<td>0.671</td>
<td>0.002</td>
<td>0.033</td>
<td>0.626</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td></td>
<td>&lt; 0.001</td>
<td>NS</td>
<td>NS</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>NS</td>
<td>0.002</td>
<td>&lt; 0.001</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td><strong>KIR2DP1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D'$</td>
<td></td>
<td>-0.631</td>
<td>-1.000</td>
<td>1.000</td>
<td>0.903</td>
<td>0.861</td>
<td>1.000</td>
<td>0.736</td>
<td>0.874</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td></td>
<td>0.016</td>
<td>0.447</td>
<td>0.399</td>
<td>0.075</td>
<td>0.042</td>
<td>0.417</td>
<td>0.030</td>
<td>0.033</td>
<td>0.049</td>
<td></td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td></td>
<td>NS</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>0.010</td>
<td>0.006</td>
<td>&lt; 0.001</td>
<td>NS</td>
</tr>
<tr>
<td><strong>KIR3DL1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D'$</td>
<td></td>
<td>0.872</td>
<td>0.062</td>
<td>0.002</td>
<td>-0.865</td>
<td>-0.936</td>
<td>0.035</td>
<td>-0.390</td>
<td>0.878</td>
<td>-0.784</td>
<td>-0.868</td>
</tr>
<tr>
<td>$r^2$</td>
<td></td>
<td>0.504</td>
<td>0.001</td>
<td>0.000</td>
<td>0.497</td>
<td>0.824</td>
<td>0.000</td>
<td>0.074</td>
<td>0.772</td>
<td>0.580</td>
<td>0.046</td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td></td>
<td>&lt; 0.001</td>
<td>NS</td>
<td>NS</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>NS</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td><strong>KIR3DS1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D'$</td>
<td></td>
<td>-0.871</td>
<td>0.152</td>
<td>-0.118</td>
<td>1.000</td>
<td>0.903</td>
<td>0.028</td>
<td>0.606</td>
<td>-0.846</td>
<td>0.824</td>
<td>1.000</td>
</tr>
<tr>
<td>$r^2$</td>
<td></td>
<td>0.473</td>
<td>0.014</td>
<td>0.007</td>
<td>0.705</td>
<td>0.722</td>
<td>0.000</td>
<td>0.169</td>
<td>0.674</td>
<td>0.678</td>
<td>0.065</td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td></td>
<td>&lt; 0.001</td>
<td>NS</td>
<td>NS</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>Black Individuals (n = 166)</th>
<th>Caucasian Individuals (n = 95)</th>
<th>p-value(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Frequency(^2)</td>
<td>n</td>
</tr>
<tr>
<td>KIR3DL1 + Bw4</td>
<td>118</td>
<td>0.711</td>
<td>67</td>
</tr>
<tr>
<td>KIR3DL1 + Bw4-80I</td>
<td>85</td>
<td>0.512</td>
<td>29</td>
</tr>
<tr>
<td>KIR3DL1 + Bw4-80T</td>
<td>22</td>
<td>0.133</td>
<td>35</td>
</tr>
<tr>
<td>KIR3DL1 + Bw6</td>
<td>141</td>
<td>0.849</td>
<td>78</td>
</tr>
</tbody>
</table>

\(^1\) 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.
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 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'-TCTGGGAAGTCTCAGGTC-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 \( \Delta C_T \) (difference in cycle threshold; the difference between the \( C_T \) of the -35T and -35C reactions). Heterozygous individuals had \( \Delta C_T \) values of between 0 and 0.3, while homozygous TT and CC individuals had \( \Delta C_T \) values of \( \geq 4 \) or \( \leq -4 \), 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

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

<sup>1</sup>Positions are given relative to the start of the HLA-C 3’ UTR, as seen in Kulkarni et al. (2011).

<sup>2</sup>p-values are for a two-sided Fisher's exact test, only significant values (p < 0.05) are shown.
Table: 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.

![Table and diagram]
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 (230H/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.
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

<table>
<thead>
<tr>
<th>HLA-C Allele</th>
<th>Black Individuals (n = 168)</th>
<th>Caucasian Individuals (n = 97)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Allele</td>
<td>Frequency</td>
</tr>
<tr>
<td>01:02</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>02:02</td>
<td>C</td>
<td>0.0149 (5(^5))</td>
</tr>
<tr>
<td>02:05</td>
<td>T</td>
<td>0.0030 (1)</td>
</tr>
<tr>
<td>02:10</td>
<td>T</td>
<td>0.0655 (22)</td>
</tr>
<tr>
<td>03:02</td>
<td>C</td>
<td>0.0149 (5)</td>
</tr>
<tr>
<td>03:03</td>
<td>C</td>
<td>0.0030 (1)</td>
</tr>
<tr>
<td>03:04</td>
<td>T</td>
<td>0.0446 (15)</td>
</tr>
<tr>
<td>03:16</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>04:01</td>
<td>T</td>
<td>0.1250 (42)</td>
</tr>
<tr>
<td>04:04</td>
<td>T</td>
<td>0.0030 (1)</td>
</tr>
<tr>
<td>04:08</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>05:01</td>
<td>C</td>
<td>0.0060 (2)</td>
</tr>
<tr>
<td>06:02</td>
<td>C</td>
<td>0.1488 (50)</td>
</tr>
<tr>
<td>06:06</td>
<td>C</td>
<td>0.0030 (1)</td>
</tr>
<tr>
<td>06:11</td>
<td>C</td>
<td>0.0030 (1)</td>
</tr>
<tr>
<td>07:01</td>
<td>T</td>
<td>0.0506 (17)</td>
</tr>
<tr>
<td>07:02</td>
<td>T</td>
<td>0.0595 (20)</td>
</tr>
<tr>
<td>07:04</td>
<td>T</td>
<td>0.0149 (5)</td>
</tr>
<tr>
<td>07:06</td>
<td>C</td>
<td>0.0387 (13)</td>
</tr>
<tr>
<td>07:11</td>
<td>T</td>
<td>0.0030 (1)</td>
</tr>
<tr>
<td>07:18</td>
<td>C</td>
<td>0.0417 (14)</td>
</tr>
<tr>
<td>08:01</td>
<td>T</td>
<td>0.0030 (1)</td>
</tr>
<tr>
<td>08:02</td>
<td>T</td>
<td>0.0119 (4)</td>
</tr>
<tr>
<td>08:04</td>
<td>C</td>
<td>0.0149 (5)</td>
</tr>
<tr>
<td>12:02</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>12:03</td>
<td>C</td>
<td>0.0149 (5)</td>
</tr>
<tr>
<td>14:02</td>
<td>C</td>
<td>0.0060 (2)</td>
</tr>
<tr>
<td>15:02</td>
<td>T</td>
<td>0.0060 (2)</td>
</tr>
<tr>
<td>15:05</td>
<td>T</td>
<td>0.0089 (3)</td>
</tr>
<tr>
<td>16:01</td>
<td>T</td>
<td>0.0655 (22)</td>
</tr>
<tr>
<td>16:02</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>17:01</td>
<td>T</td>
<td>0.1131 (38)</td>
</tr>
<tr>
<td>18:01</td>
<td>T</td>
<td>0.0149 (5)</td>
</tr>
<tr>
<td>18:02</td>
<td>T</td>
<td>0.0298 (10)</td>
</tr>
</tbody>
</table>

\(^1\)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.,
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 an 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/mm³ 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/mm³, 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

<table>
<thead>
<tr>
<th></th>
<th>CD4⁺ T Cell Count (cells/mm³)</th>
<th>Log Viral Load (RNA copies/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HIV-1 Progressors</strong></td>
<td>Median</td>
<td>Range</td>
</tr>
<tr>
<td>(n = 69)</td>
<td>178</td>
<td>3 - 296</td>
</tr>
<tr>
<td><strong>HIV-1 Controllers</strong></td>
<td>Median</td>
<td>Range</td>
</tr>
<tr>
<td>(n = 42)</td>
<td>699</td>
<td>306 - 1380</td>
</tr>
</tbody>
</table>

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 = ∞), 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

<table>
<thead>
<tr>
<th>HLA Allele</th>
<th>HIV-1 Progressors (n = 69)</th>
<th>HIV-1 Controllers (n = 42)¹</th>
<th>OR (95% Confidence Interval)²</th>
<th>p-value²</th>
</tr>
</thead>
<tbody>
<tr>
<td>B*57:03:01</td>
<td>3 0.022</td>
<td>5 0.060</td>
<td>13.7 (1.5 - 159.3)</td>
<td>0.024</td>
</tr>
<tr>
<td>B*58:01:01</td>
<td>4 0.029</td>
<td>6 0.071</td>
<td>9.8 (1.3 - 99.5)</td>
<td>0.037</td>
</tr>
<tr>
<td>B*14:02:01</td>
<td>2 0.014</td>
<td>3 0.036</td>
<td>9.0 (0.8 - 13.3)</td>
<td>0.080</td>
</tr>
<tr>
<td>B*39:10:01</td>
<td>4 0.029</td>
<td>5 0.060</td>
<td>8.5 (1.2 - 70.3)</td>
<td>0.037</td>
</tr>
<tr>
<td>B*14:01:01</td>
<td>2 0.014</td>
<td>3 0.036</td>
<td>7.0 (0.7 - 8.5)</td>
<td>0.098</td>
</tr>
<tr>
<td>B*42:01:01</td>
<td>9 0.065</td>
<td>9 0.107</td>
<td>4.6 (0.9 - 28.4)</td>
<td>0.080</td>
</tr>
<tr>
<td>A*23:01:01</td>
<td>14 0.101</td>
<td>4 0.059</td>
<td>0.1 (0.0 - 1.0)</td>
<td>0.066</td>
</tr>
<tr>
<td>A*03:01:01</td>
<td>11 0.080</td>
<td>1 0.015</td>
<td>0.0 (0.0 - 0.3)</td>
<td>0.012</td>
</tr>
</tbody>
</table>

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

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

\(^1\) 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).
**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 \(\neg B\), 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 |  
|---|---|---|---|---|---|---|
| 9 | Y | HIV-1 Progressors (\(n = 69\)) | 0.790 | HIV-1 Controllers (\(n = 42\)) | 0.558 | OR (95% Confidence Interval) \(^{1}\) | 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).
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

<table>
<thead>
<tr>
<th></th>
<th>230 Indel</th>
<th>263 Indel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>D</td>
</tr>
<tr>
<td>HIV-1 Progressors (n = 62)</td>
<td>0.468</td>
<td>0.806</td>
</tr>
<tr>
<td>HIV-1 Controllers (n = 37)</td>
<td>0.405</td>
<td>0.919</td>
</tr>
<tr>
<td>OR (95% CI)</td>
<td>0.8</td>
<td>2.7</td>
</tr>
<tr>
<td>p-value</td>
<td>0.68</td>
<td>0.16</td>
</tr>
</tbody>
</table>

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). Comparisons 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 \textit{B*39:10:01|C*12:03:01} and \textit{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. \textit{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 \textit{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 \textit{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 Ser17-Arg56-Glu114-His116 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 β2-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 \textit{HLA} gene complex has been identified as one of the major determinants of HIV-1 control (Fellay \textit{et al.}, 2007; The International HIV Controllers Study, 2010). \textit{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/mm³ 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 \((p = 0.036; \text{ OR } = 3.3; \text{ CI } = 1.1 – 9.8)\) and KIR2DS5 \((p = 0.031; \text{ OR } = 2.7; \text{ CI } = 1.1 – 6.7)\), but displayed KIR2DL3 frequencies greater than that seen in the viraemic controller group \((p = 0.046; \text{ OR } = 0.4; \text{ CI } = 0.1 – 0.9)\). Viraemic controllers, in turn, were distinguished from elite controllers, by displaying a greater frequency of KIR2DS5 \((p = 0.033; \text{ OR } = 0.2; \text{ CI } = 0.0 – 0.8)\) and a decreased likelihood of being in possession of KIR2DL3 \((p = 0.018; \text{ OR } = \infty)\). 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 cA01tA01 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 x 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 cB01tA01 haplotype (Table 5.3).
<table>
<thead>
<tr>
<th>KIR Genotype</th>
<th>HIV-1 Progressors (n = 69)</th>
<th>HIV-1 Viremic Controllers (n = 31)</th>
<th>HIV-1 Elite Controllers (n = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA 1</td>
<td>0.290 (20)</td>
<td>0.161 (5)</td>
<td>0.182 (2)</td>
</tr>
<tr>
<td>Bx 21</td>
<td>0.130 (9)</td>
<td>0.194 (6)</td>
<td>-</td>
</tr>
<tr>
<td>Bx 5</td>
<td>0.116 (8)</td>
<td>0.129 (4)</td>
<td>0.455 (5)</td>
</tr>
<tr>
<td>Bx 4</td>
<td>0.087 (6)</td>
<td>-</td>
<td>0.091 (1)</td>
</tr>
<tr>
<td>Bx 20</td>
<td>0.072 (5)</td>
<td>0.097 (3)</td>
<td>0.091 (1)</td>
</tr>
<tr>
<td>Bx 9</td>
<td>0.058 (4)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bx 71</td>
<td>0.058 (4)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bx 112</td>
<td>0.058 (4)</td>
<td>0.194 (6)</td>
<td>-</td>
</tr>
<tr>
<td>Bx 6</td>
<td>0.029 (2)</td>
<td>-</td>
<td>0.091 (1)</td>
</tr>
<tr>
<td>Bx 92</td>
<td>0.029 (2)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bx 2</td>
<td>0.014 (1)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bx 3</td>
<td>0.014 (1)</td>
<td>0.091 (1)</td>
<td></td>
</tr>
<tr>
<td>Bx 73</td>
<td>0.014 (1)</td>
<td>0.065 (2)</td>
<td>-</td>
</tr>
<tr>
<td>Bx 172</td>
<td>0.014 (1)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bx 228</td>
<td>0.014 (1)</td>
<td>0.097 (3)</td>
<td>-</td>
</tr>
<tr>
<td>Bx 38</td>
<td>-</td>
<td>0.032 (1)</td>
<td>-</td>
</tr>
<tr>
<td>Unknown</td>
<td>-</td>
<td>-</td>
<td>0.032 (1)</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th></th>
<th>HIV-1 Progressors (n = 69)</th>
<th>HIV-1 Viraemic Controllers (n = 31)</th>
<th>HIV-1 Elite Controllers (n = 11)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n  Frequency</td>
<td>n  Frequency</td>
<td>n  Frequency</td>
</tr>
<tr>
<td>cA01</td>
<td>tA01(^1)</td>
<td>52 0.377</td>
<td>20 0.323</td>
</tr>
<tr>
<td>cA01</td>
<td>tA01-hybd1(^2)</td>
<td>13 0.094</td>
<td>2 0.032</td>
</tr>
<tr>
<td>cA01</td>
<td>tA01-del5(^3)</td>
<td>11 0.080</td>
<td>1 0.016</td>
</tr>
<tr>
<td>cA01/cB03</td>
<td>tB01(^3)</td>
<td>1 0.007</td>
<td>0 0.000</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01</td>
<td>20 0.145</td>
<td>24 0.387</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-hybd1</td>
<td>2 0.014</td>
<td>2 0.032</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-del3</td>
<td>4 0.029</td>
<td>5 0.081</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-del5</td>
<td>11 0.080</td>
<td>2 0.032</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-del9</td>
<td>8 0.058</td>
<td>0 0.000</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-del3-del5</td>
<td>1 0.007</td>
<td>0 0.000</td>
</tr>
<tr>
<td>cB01</td>
<td>tA01-del3-del5-del9</td>
<td>0 0.000</td>
<td>1 0.016</td>
</tr>
<tr>
<td>cB01</td>
<td>tB01</td>
<td>4 0.029</td>
<td>2 0.032</td>
</tr>
<tr>
<td>cB01</td>
<td>tB01-del7</td>
<td>6 0.043</td>
<td>1 0.016</td>
</tr>
<tr>
<td>cB02</td>
<td>tA01-hybd1</td>
<td>5 0.036</td>
<td>1 0.016</td>
</tr>
<tr>
<td>cB02</td>
<td>tB01</td>
<td>0 0.000</td>
<td>0 0.000</td>
</tr>
<tr>
<td>H2(^4)</td>
<td>0 0.000</td>
<td>1 0.016</td>
<td>0 0.000</td>
</tr>
</tbody>
</table>

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/cB01tB01 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

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

\(^1\) 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 \textit{KIR2DL1}, \textit{KIR2DL2}, \textit{KIR2DL3} and \textit{KIR2DS1} occur in combination with \textit{HLA-C} alleles encoding the C1 and C2 epitopes in HIV-1 progressors, -viraemic controllers and -elite controllers.

\begin{tabular}{|c|c|c|c|c|}
\hline
 & HIV-1 Progressors \hspace{1cm} (n = 69) & HIV-1 Viraemic Controllers \hspace{1cm} (n = 31) & HIV-1 Elite Controllers \hspace{1cm} (n = 11) \\
\hline
\textbf{HLA-C1} & & & \\
KIR2DL1 + C1 & 49 & 0.710 & 21 & 0.677 \\
KIR2DL2 + C1 & 33 & 0.478 & 17 & 0.548 \\
KIR2DL3 + C1 & 40 & 0.580 & 12 & 0.387 \\
KIR2DS1 + C1 & 8 & 0.116 & 2 & 0.065 \\
\hline
\textbf{HLA-C2} & & & \\
KIR2DL1 + C2 & 56 & 0.812 & 23 & 0.742 \\
KIR2DL2 + C2 & 38 & 0.551 & 20 & 0.645 \\
KIR2DL3 + C2 & 45 & 0.652 & 15 & 0.484 \\
KIR2DS1 + C2 & 10 & 0.145 & 3 & 0.097 \\
\hline
\end{tabular}

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 \textit{KIR} and \textit{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 with the Bx 5 genotype can most easily be achieved as a result of having a KIR2DS3-containing cB01tA01 haplotype in combination with cA01tA01 haplotype, while the Bx 112 genotype is easily derived from the combination of KIR2DS3- and KIR2DS5-containing cB01tA01 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). KIR3DL1*h/KIR3DL1*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 cA01ltA01 haplotype and a cB01ltA01 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


Appendix A

Turnitin Report

**Turnitin Originality Report**
turnitin.docx by Nikki Gentle
From PhD Research
(9T5NEoFKOF6336g28aNnCeg30dOsUdRa91SaP18F5PMsvNimlyM44I5a4kyMuzKwR1M3lV3d8wUdr5SYgJ6y7Kv6021K6br6KOM)

- 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)

2
1% match (publications)

3
1% match (publications)

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: Ms Nikkie Gentle
(Principal Investigator)

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

PROJECT TITLE: The Role of Killer-Immunoglobulin-like
Receptors and Specific Human Leukocyte
Antigen cls 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

M130872 Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES