Genetic aspects of HIV-1 risk in

an African setting

Desiree C. Petersen

Dissertation presented in fulfillment of the requirements for the degree of

Doctor of Philosophy in Health Sciences (Medical Virology) at the

University of Stellenbosch

Co-promoter and study supervisor: Dr Vanessa M. Hayes

Co-promoter: Dr Michael Dean

Co-promoter: Prof. Estrelita Janse van Rensburg

Promoter: Dr Richard H. Glashoff

April 2006

Declaration

I, the undersigned, hereby declare that the work contained in this dissertation is my own original work and that I have not previously in its entirety or in part submitted it at any University for a degree.

Signature ……………………   **Date** ……………………

Summary

Host susceptibility to human immunodeficiency virus-1 (HIV-1) infection and disease progression to acquired immunodeficiency syndrome (AIDS) varies widely amongst individuals. This observation led to the identification of host genetic factors playing a vital role in HIV-1 pathogenesis. Previous studies mainly focusing on Caucasian-based populations have indicated possible associations between genetic variants and host susceptibility to HIV-1/AIDS. The limited studies performed on African-based populations have emphasised the need for extensive investigation of both previously reported and particularly novel genetic variants within the older and genetically diverse Sub-Saharan African populations.

In this study, the case-control samples were represented by African individuals of Xhosa descent, all residing in the Western Cape Province of South Africa. This included 257 HIV-1 seropositive patients and 110 population-matched HIV-1 seronegative controls. Mutational screening was performed in a subset of individuals for the entire coding regions of the CC chemokine receptor 5 (*CCR5)* and CC chemokine receptor 2 (*CCR2*) genes, and the 3' untranslated region of the CXC chemokine ligand (*CXCL12)* gene, as previously reported (Petersen, 2002). Further analysis of these genes in a larger study sample involved the genotyping of previously identified mutations and single nucleotide polymorphisms (SNPs), which forms part of the present study. In addition, mutational screening was performed for the entire coding region of the CXC chemokine receptor 4 (*CXCR4*) gene, partial coding region of the

mannose binding lectin (*MBL*) gene, and the promoter regions of interleukin 4 (*IL4*), interleukin 10 (*IL10*) and the solute carrier 11A1 (*SLC11A1)* genes. This was followed by genotyping of SNPs occurring in *CCR5*, *CCR2*, *CXCL12, MBL, IL4, IL10,* CX3C chemokine receptor 1 (*CX3CR1*), CC chemokine ligand 5 (*CCL5*) and tumour necrosis factor alpha (*TNFα*) genes. Significant associations were observed with HIV-1 susceptibility in the Xhosa population of South Africa. These included the *CCR5*-2733A>G*,* *CX3CR1*V249I, *IL10*-819C>T and *IL10*-592C>A SNPs being associated with a reduced risk for HIV-1 infection, while the *CCR5*-2135C>T and *SDF1*-3'G>A (*CXCL12*-3'G>A) SNPs were associated with increased susceptibility to HIV-1 infection. Furthermore, certain haplotypes for *IL4* and *IL10* showed association with reduced risk for HIV-1 infection. This included the identification of a novel *IL4* haplotype restricted to the HIV-1 seronegative control group.

This study emphasises the importance of considering genetic diversity across all populations, as certain HIV-1/AIDS associations appear to be restricted to specific ethnic groups. These findings have also provided an understanding for further elucidating the functional roles of genetic variants in determining HIV-1/AIDS susceptibility. Ultimately, such genetic association studies will contribute to establishing HIV-1/AIDS risk profiles for African-based populations from pandemic-stricken Sub-Saharan Africa.

Opsomming

Die vatbaarheid van gashere vir menslike immuniteitsgebrek virus (MIV-1) infeksie en siekteprogressie na verworwe immuniteitsgebreksindroom (VIGS) varieer baie in individue. Hierdie waarneming het gelei tot die identifikasie van genetiese faktore in gashere wat 'n belangrike rol speel in die patogenese van MIV-1. Vorige studies, wat meestal gefokus het op Kaukasier-gebaseerde populasies, het moontlike assosiasies getoon tussen genetiese faktore en gasheervatbaarheid vir MIV-1/VIGS. Die beperkte studies wat gedoen is op Afrikaan-gebaseerde populasies het die behoefte beklemtoon vir omvattende navorsing van reeds geidentifiseerde en veral nuwe genetiese variante wat in die ouer en geneties diverse populasies van Sub-Sahara Afrika voorkom.

In hierdie studie is alle monsters afkomstig van Afrikane van Xhosa afkoms wat almal in die Wes-Kaap Provinsie van Suid-Afrika woon. Dit sluit 257 MIV-1 seropositiewe pasiënte en 110 MIV-1 seronegatiewe kontroles van dieselfde populasie in. In 'n vorige studie is mutasie sifting gedoen in 'n groep individue vir die volledige koderende areas van die CC chemokien reseptor 5 (*CCR5*) en CC chemokien reseptor 2 (*CCR2*) gene, en die 3' streek van die CXC chemokien ligand 12 (*CXCL12*) geen waar translasie nie plaasgevind nie (Petersen, 2002). Verdere analise van hierdie gene in 'n groter studiegroep vorm deel van die huidige studie en het die genotipering van reeds geidentifiseerde mutasies en enkel nukleotied polimorfisme (SNP) setels ingesluit. Addisioneel is mutasie sifting gedoen vir die volledige koderende area van die CXC chemokien reseptor 4 (*CXCR4*) geen, die gedeeltelike koderende

area van die mannose bindingslektien (*MBL*) geen, en die promotor areas van die interleukien 4 (*IL4*), interleukien 10 (*IL10*) en oplosbare draer 11A1 (*SLC11A1)* gene. Dit is gevolg deur genotipering van SNPs wat in die *CCR5, CCR2, CX3C* chemokien reseptor 1 (*CX3CR1*), *CXCL12,* CC chemokien ligand 5 (*CCL5)* en tumor nekrose faktor alfa (*TNFα*) gene voorkom. Betekenisvolle assosiasies met MIV-1 vatbaarheid het in die Xhosa populasie van Suid-Afrika voorgekom.   Dit sluit in *CCR5*-2733A>G*, CX3CR1*V249I, *IL10*-819C>T en *IL10*-592C>A SNPs wat geassosieer word met 'n verlaagde risiko vir MIV-1 infeksie, terwyl die *CCR5*-2135C>T en *SDF1*-3'G>A (*CXCL12*-3'G>A) SNPs geassosieer word met 'n verhoogde vatbaarheid vir MIV-1 infeksie. Verder kon sekere haplotipes van *IL4* en *IL10* geassosieer word met 'n verlaagde risiko vir MIV-1 infeksie. Dit sluit in die identifikasie van 'n nuwe *IL4* haplotipe wat uitsluitlik by MIV-1 seronegatiewe kontroles voorgekom het.

Hierdie studie beklemtoon die belangrikheid om genetiese diversiteit in alle populasies in aanmerking te neem omdat dit blyk dat sekere MIV-1/VIGS assosiasies slegs in spesifieke etniese groepe voorkom**.** Die bevindings het ook die weg gebaan vir die verdere ondersoek van die funksionele rolle van genetiese variante in die bepaling van vatbaarheid vir MIV-1/VIGS. Sulke studies sal uiteindelik bydra tot die daarstelling van risikoprofiele vir Afrikaan-gebaseerde populasies van Sub-Sahara Afrika waar die pandemie heers.

Acknowledgements

I wish to extend my sincere thanks to the following people and institutions:

For my Mother and Father

List of abbreviations

AIDS	acquired immunodeficiency syndrome
AIM	ancestral informative marker
Ala (A)	alanine
Arg (R)	arginine
Asn (N)	asparagine
Asp (D)	aspartic acid
bp	base pair
CCL	CC chemokine ligand
CCR	CC chemokine receptor
CXCL	CXC chemokine ligand
CXCR	CXC chemokine receptor
CX3CL	CX3C chemokine ligand
CX3CR	CX3C chemokine receptor
Cys (C)	cysteine
DGGE	denaturing gradient gel electrophoresis
DNA	deoxyribonucleic acid
env	envelope
GC-clamp	guanine and cytosine clamp
Gln (Q)	glutamine

Glu (E)	glutamic acid
Gly (G)	glycine
gp	glycoprotein
His (H)	histidine
HIV	human immunodeficiency virus
HIV –	HIV seronegative
HIV +	HIV seropositive
HLA	human leukocyte antigen
HWE	Hardy-Weinberg equilibrium
IL	interleukin
Ile (I)	isoleucine
kb	kilobase
Leu (L)	leucine
LD	linkage disequilibrium
LTNP	long-term non-progressor
Lys (K)	lysine
MBL	mannose binding lectin
MCP	monocyte chemotactic protein
Met (M)	methionine
MIM#	OMIM database reference
MIP	macrophage inflammatory protein

M-tropic	macrophage tropic
NRAMP1	natural resistance-associated macrophage protein 1
NSI	non-syncytium inducing
ORF	open reading frame
PCR	polymerase chain reaction
Phe (F)	phenylalanine
Pro (P)	proline
RANTES	regulated on activation normal T cell expressed and secreted
RNA	ribonucleic acid
SDF1	stromal derived factor 1
Ser (S)	serine
SI	syncytium inducing
SLC11A1	solute carrier 11A1
SNP	single nucleotide polymorphism
TDT	transmission disequilibrium test
Th	T helper
Thr (T)	threonine
T lymphocytes	thymus-derived lymphocytes
Trp (W)	tryptophan

T-tropic   T cell line tropic

Tyr (Y)   tyrosine

UF   urea/formamide

UTR   untranslated region

Valine (V)   valine

Contents

Chapter 2: Chemokine and chemokine receptors

Chapter 3: Th1 and Th2 cytokines

Chapter 4: Immunoregulatory proteins

Chapter 1

Introduction

1.1. Host genetic susceptibility to HIV-1/AIDS

Individual susceptibility to HIV-1 infection and disease progression to AIDS varies extensively. Most people are susceptible to HIV-1 infection, although there are uninfected groups who have experienced high-risk or repeated exposure. Presently, these exposed uninfected groups are mainly defined by unprotected sexual encounters (commercial sex workers and discordant couples), intravenous drug usage, contact with contaminated blood or blood products, and the absence of perinatal transmission. Furthermore, those individuals who do become infected display diverse clinical outcomes and have different rates of disease progression. The median interval from the time of HIV-1 seroconversion to the development of AIDS is approximately eight to ten years, but long-term non-progressors (remain healthy for periods longer than 10 years) and rapid progressors (develop AIDS within five years) have been observed. It is well-established that host susceptibility to HIV-1 infection and the disease course to AIDS is determined by the complex interaction of certain parameters, including viral characteristics, socio-economic/environmental aspects, host immunological and host genetic factors. Many studies have focused on the role of host genetic factors as gene variants, including SNPs (occur at allele frequencies greater than 0.01 or 1%), influencing HIV-1/AIDS susceptibility have been identified. HIV-1/AIDS risk profiles for individuals from various populations are therefore being established. The candidate genes are mainly selected based on the known or hypothesised function of their gene product (protein) in the presence of HIV-1. This chapter subsection will provide a comprehensive understanding of host genetic factors that contribute to elucidating the complexity of HIV-1/AIDS pathogenesis.

1.1.1. Chemokine and chemokine receptors

Chemokines are chemoattractant cytokines that act via G-protein-coupled chemokine receptors. The direct ligand-receptor interaction between chemokine and chemokine receptors is vital for the regulation and maintenance of a functional host immune system [Reviewed in Dong et al., 2003]. The discovery of chemokines playing a role in HIV-1 infection originated from a study showing that CC chemokines, secreted by CD8[+] cells (cytotoxic T-lymphocytes), act as potent inhibitors preventing HIV-1 infection by macrophage-tropic (M-tropic) or non-syncytium inducing (NSI) viruses (mainly infect macrophages, monocytes and T-lymphocytes) [Cocchi et al., 1995]. These chemokines included CCL5 (RANTES), CCL3 (MIP-1α) and CCL4 (MIP-1β), which are the natural ligands for the CC chemokine receptor, CCR5 [Samson et al., 1996a]. Further studies indicated that a CXC chemokine, CXCL12 (SDF1), acts as an inhibitor of T-cell tropic (T-tropic) or syncytium inducing (SI) viruses (mainly infect T-lymphocytes) and also potentially influences HIV-1 replication. CXCL12 serves as the only known natural ligand for the CXC chemokine receptor, CXCR4, previously referred to as 'fusin' [Bleul et al., 1996a; Oberlin et al., 1996].

The CD4 molecule on the host cell surface was initially identified as the primary co-receptor for HIV-1 [Dalgeish et al., 1984; Klatzman et al., 1984], but subsequent studies found specific chemokine receptors to serve as additional cellular host co-receptors for virus entry [Alkhatib et al., 1996; Choe et al., 1996; Deng et al., 1996; Doranz et al., 1996; Dragic et al., 1996; Feng et al., 1996; Rucker et al., 1997; Combadiere et al., 1998]. HIV-1 infection is therefore facilitated by the binding of viral env glycoprotein (gp) 120 to the CD4 molecule,

which results in the formation of a CD4-gp120 complex [Dalgeish et al., 1984; Klatzman et al., 1984; Maddon et al., 1986; Lasky et al., 1987]. This is followed by a conformational change within the viral envelope that enables the gp120 to bind to the chemokine receptor [Trkola et al., 1996; Wu et al., 1996; Speck et al., 1997; Kwong et al., 1998; Rizzuto et al., 1998], resulting in the exposure of the viral env gp41 peptide for ultimate virus-host cell fusion [Moore et al., 1993; Sattentau et al., 1995; Lapham et al., 1996] (see Figure 1).
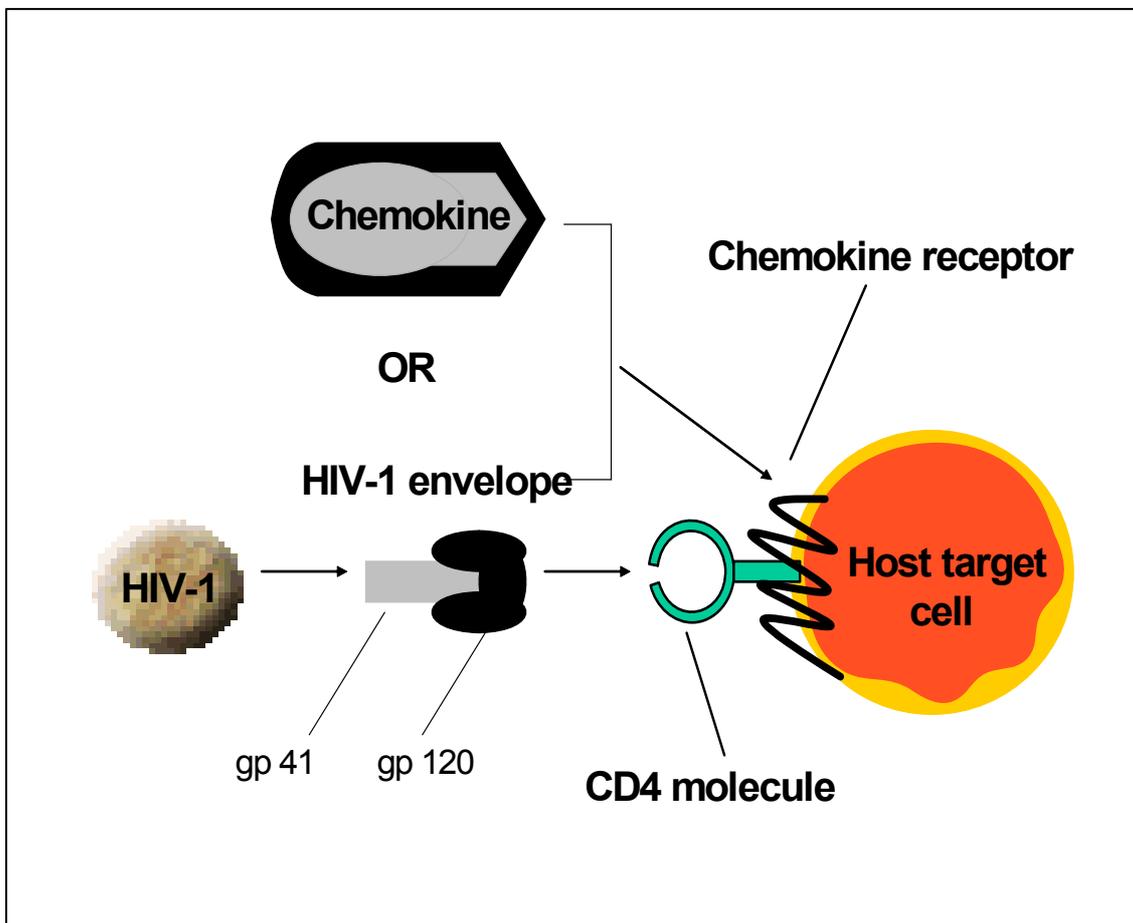
Figure 1. HIV-1 entry into the host target cell. HIV-1 infection is facilitated by the interaction of the virion envelope glycoproteins, gp120 and gp41, with two cellular host receptors, of which one is a CD4 molecule and the other, a chemokine receptor, whose natural ligands are specific chemokines.

The unravelling of the HIV-1 entry process accentuated the importance of investigating the exact role of chemokines in the presence of HIV. Chemokines were found to suppress and/or prevent HIV-1 infection by directly competing with the virus for binding to chemokine receptors or by down-modulation of chemokine receptor expression [Bleul et al., 1996a; Combadiere et al., 1996; Oberlin et al., 1996; Samson et al., 1996a; Raport et al., 1996; Amara et al., 1997, Combadiere et al., 1998] (see Figure 1). This was followed by elevated levels of chemokines being observed in 'exposed yet uninfected' individuals and associated with delayed progression to AIDS [Paxton et al., 1996; Paxton et al., 1998; Ullum et al., 1998, Paxton et al., 1999; Paxton et al., 2001].

A model for specific co-receptor usage by different HIV-1 strains does exist. M-tropic/NSI viruses preferentially utilise CCR5 or "less efficient" co-receptors such as CCR2, CX3CR1 and are therefore termed R5 strains [Alkhatib et al., 1996, Choe et al., 1996, Deng et al., 1996; Dragic et al., 1996; Combadiere et al., 1998]. T-tropic/SI viruses utilise primarily CXCR4 as co-receptors for entry and are therefore termed X4 strains [Feng et al., 1996]. There are however dual tropic viruses utilising both CXCR4 and CCR5, but also additional co-receptors such as CCR2 and CX3CR1 for entry, and are therefore termed R5X4 strains. [Choe et al., 1996; Doranz et al., 1996; Simmons et al., 1996; Berger et al., 1998; Combadiere et al, 1998]. R5 viruses are normally present during transmission and the early asymptomatic stages, while the more cytopathic X4 viruses are generally present during the later symptomatic stages. Most HIV-1 infected individuals with the onset of rapid disease progression and approximately 50% of all individuals progressing to AIDS experience a shift in

viral tropism that results in the conversion of NSI to SI phenotype [Tersmette et al., 1988; Tersmette et al., 1989; Roos et al., 1992; Schuitemaker et al., 1992; Connor et al., 1993, Zhu et al., 1993; Jansson et al., 1999] (see Figure 2).
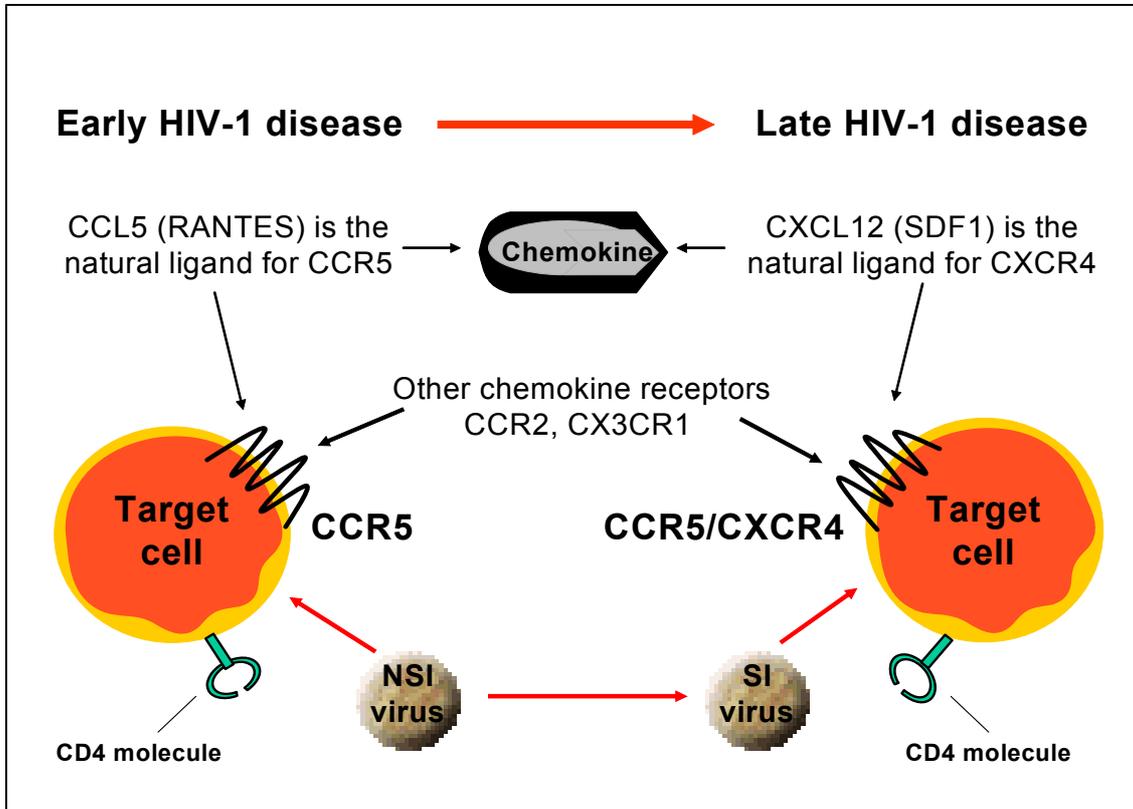
Figure 2. Schematic illustration of the model for co-receptor usage by different viruses in the presence of chemokine ligands.

The utilisation of specific chemokine receptors by different virus strains at various stages of HIV-1 infection, together with the inhibitory roles of their chemokine ligands, could therefore be affected by the presence of variants in the encoding genes (see Figure 2). These genetic variants could alter protein cell surface expression or its biological function, further influencing susceptibility to HIV-1/AIDS. Numerous studies have identified and investigated the role of chemokine and chemokine receptor gene variants of several candidates including, *CCR5, CCR2, CXCR4, CX3CR1, CCL5 and CXCL12*, in HIV-1 susceptibility and/or disease progression to AIDS.

1.1.1.1. CC chemokine receptor 5 (CCR5)

CCR5 (MIM# 601373) serves as the principle co-receptor for NSI/R5 strains of HIV-1 [Alkhatib et al., 1996; Choe et al., 1996; Deng et al., 1996; Dragic et al., 1996]. Several studies have shown that multiple domains of CCR5 and its amino terminus play a vital role in mediating co-receptor activity [Atchison et al., 1996; Rucker et al., 1996; Alkhatib et al., 1997; Bieniasz et al., 1997; Doranz et al., 1997; Picard et al., 1997; Wang et al., 1999]. The natural ligands for CCR5 are CCL5 (RANTES), CCL3 (MIP-1α) and CCL4 (MIP-1β) [Samson et al., 1996a], which were found to block cell fusion mediated by the virion envelope glycoprotein and thereby inhibit HIV-1 infection [Deng et al., 1996].

The *CCR5* gene, located at chromosomal position 3p21 [Liu et al., 1996], comprises four exons and two introns, with exon 4 containing the entire coding region. The CCR5 protein consists of 352 amino acids [Mummidi et al., 1997]. *CCR5* has dual promoter usage with the presence of a weak promoter upstream of exon 1 and a strong downstream promoter, which includes the intronic region between exon 1 and 3 [Mummidi et al., 1997].

The most commonly studied *CCR5* variant is a well-documented 32bp deletion mutation (*CCR5Δ32*) [Dean et al., 1996; Liu et al., 1996; Samson et al., 1996b]. This mutation results in the formation of a truncated protein that is not expressed at the cell surface. The HIV-1 virus is therefore unable to bind and infect host target cells. The *CCR5Δ32* comprises nucleotides 794 to 825 of the coding region, resulting in a frameshift after amino acid 174 and a premature stop codon at 182 [Liu et al., 1996]. Individuals homozygous for *CCR5Δ32* are

highly protective against HIV-1 infection [Dean et al., 1996; Huang et al., 1996; Liu et al., 1996; Samson et al., 1996b; Zimmerman et al., 1997], although a few exceptions to the rule have been observed. These include individuals who have been infected with virus strains that utilise additional or other co-receptors [Balotta et al., 1997; Biti et al., 1997; O'Brien et al., 1997; Theodorou et al., 1997]. Heterozygosisty for *CCR5*Δ32 has not been markedly associated with protection against HIV-1 infection, but does offer delayed progression to AIDS by two to four years [Dean et al., 1996; Huang et al., 1996; Liu et al., 1996; Samson et al., 1996b; Zimmerman et al., 1997]. *CCR5*Δ32 is however largely confined to the Caucasian population (allele frequencies are 12% -14% in Northern Europeans and 4%-6% in Southern Europeans) and rarely observed or completely absent in Africans [Martinson et al., 1997; Libert et al., 1998; Stephens et al., 1998; Petersen et al., 2001; Dean et al., 2002] (see Chapter 2.1 and Appendix A).

Additional *CCR5* variants occurring in both the coding [Dean et al., 1996; Ansari-Lari et al. 1997; Carrington et al., 1997; Quillent et al., 1998; Carrington et al., 1999; Petersen et al., 2001; Hayes et al., 2002] and strong downstream promoter regions [Mummidi S et al., 1997; Kostrikis et al., 1998; Martin M et al., 1998a; McDermottt DH et al., 1998; Mummidi et al., 1998] have been identified in various population groups. A few of these variants occur as polymorphisms or as part of extended haplotypes (P1 to P10 or HHA to HHG) that are associated with influencing susceptibility to HIV-1/AIDS [Martin M et al., 1998a; Carrington et al., 1999; Gonzalez et al., 1999; An et al., 2000; Ramaley et al., 2002] (see Chapter 2.1, Appendix A and B). Functional studies conducted have

shown that a few of the coding variants are associated with disrupted co-receptor activity and altered ligand binding affinities or expression [Howard et al., 1999; Blanpain et al., 2000], while some of the promoter variants influenced promoter activity and differential nuclear factor binding [Mummidi et al., 1997; Bream et al., 1999]. Due to inconsistent functional findings for *CCR5* mutations, ongoing research is required for confirming the underlying effects to provide a clear understanding of observed HIV-1/AIDS associations.

1.1.1.2. CC chemokine receptor 2 (CCR2)

CCR2 (MIM# 601267) is recognised as an additional co-receptor for a minority of dual-tropic/R5X4 HIV-1 strains, but has a lower efficiency compared to CCR5 and CXCR4 [Doranz et al., 1996]. The amino terminus of CCR2 is considered essential for co-receptor function [Rucker et al., 1996; Frade et al., 1997]. Natural ligands for CCR2 include CCL2 (MCP-1), CCL8 (MCP-2), CCL7 (MCP-3), CCL13 (MCP-4), and CCL12 (MCP-5). These chemokines have been found to inhibit HIV-1 replication of both NSI/R5 and SI/X4 viruses [Kalinkovich et al., 1999; Lee and Montaner, 1999].

Two isoforms exist for CCR2 (CCR2A, CCR2B) [Charo et al., 1994], which is the result of alternative splicing a single gene [Wong et al., 1997] localised to chromosome 3p21 [Daugherty and Springer, 1997]. The gene consists of three exons and two introns with the coding region for *CCR2A* found in exon 2 and part of exon 3, while exon 2 contains the coding region for *CCR2B*. CCR2A, consisting of 374 amino acids, is mainly found in the cytoplasm due to retention

signals, while the predominant CCR2B, consisting of 360 amino acids, is expressed in the cytoplasm and at the cell surface [Wong et al., 1997].

A polymorphism identified in the coding region of *CCR2* is characterised by a conservative amino acid change from valine to isoleucine at codon 64 (*CCR2*V64I G>A) [Smith et al., 1997a]. The *CCR2*V64I SNP (rs1799864) in both the homozygous and heterozygous state delayed the onset of AIDS by two to four years, although no decreased susceptibility risk for HIV-1 infection was conferred [Smith et al., 1997a; Smith et al., 1997b; Anzala et al., 1998; Kostrikis et al., 1998; Rizzardi et al., 1998]. The influence of *CCR2*V64I on disease progression seemed more apparent in Africans compared to Caucasians [Anzala et al., 1998; Mummidi et al., 1998, O'Brien and Moore, 2000]. However, the reported effect has not been confirmed in all studies [Michael et al., 1997; Eugen-Olsen et al., 1998; Ioannidis et al., 1998; Petersen et al., 2002; Ramaley et al., 2002]. Allele frequencies for *CCR2*V64I range from 10 to 25% in different population groups, including African, Caucasian, Asian and admixed ethnic groups [Michael et al., 1997; Smith et al., 1997a; Mummidi et al., 1998; Kostrikis et al., 1998; Williamson et al., 2000; Petersen et al., 2002; Ramaley et al., 2002] (see Chapter 2.1 and Appendix C). A significant functional effect for *CCR2*V64I was recently reported and involves a change in CCR2A isoform stability, which results in increased down-modulation of CCR5, a principle HIV-1 co-receptor [Nakayama et al., 2004].

The *CCR5* and *CCR2* genes, which display high sequence homology and are in close proximity (approximately 10kb apart), show a high degree of linkage

disequilibrium (LD) [Smith et al., 1997b]. It was initially suggested that *CCR2*V64I is in LD with *CCR5*Δ32, but it was shown that the *CCR2*V64I SNP occurs invariably with the wildtype *CCR5* allele [Smith et al., 1997b; Kostrikis et al., 1998]. The *CCR2*V64I SNP is however in strong LD with a *CCR5* promoter variant, *CCR5*-1835C>T (rs1800024) [Kostrikis et al., 1998; Mummidi et al., 1998]. Previously described haplotypes comprising the *CCR2*V64I and *CCR5* polymorphisms are associated with either various risks for HIV-1/AIDS or having no influence within specific population groups (see Chapter 2.1). Other *CCR2* gene variants have also been reported, although to date none of these have indicated significant associations with susceptibility for HIV-1/AIDS [Petersen et al., 2001, Petersen et al., 2002] (see Appendix A and C).

1.1.1.3. CX3C chemokine receptor 1 (CX3CR1)

CX3CR1 (MIM# 601470) is considered a minor HIV-1 co-receptor for a limited number of dual-tropic/R5X4 viruses, having a lower fusion activity compared to CCR5 and CXCR4 [Rucker et al., 1997; Combadiere et al., 1998]. The amino-terminal domain of CX3CR1 was found to play a crucial role in determining co-receptor activity [Garin et al., 2003]. CX3CL1 (fractalkine) is the natural ligand for CX3CR1 and can effectively block its ability to serve as a HIV-1 co-receptor [Combadiere et al., 1998].

The *CX3CR1* gene, localised to chromosome 3p21.3 [Maho et al., 1999], consists of four exons and three introns with its coding region contained within exon 4 and encoding a protein of 355 amino acids [Raport et al., 1995; Garin et al., 2002]. Three gene transcripts are however produced by the splicing of

three untranslated exons (1 – 3) with exon 4. The predominant gene transcript corresponds to the splicing of exon 2 with exon 4. Three different functional promoter regions therefore control the expression of the individual gene transcripts, which only differ by their untranslated regions [Garin et al., 2002].

Previous studies investigating the possibility of *CX3CR1* variants influencing susceptibility to HIV-1/AIDS have identified two coding region SNPs, *CX3CR1*V249I (rs3732379) and *CX3CR1*T280M (rs3732378). The first SNP occurring at codon 249 (G>A) is characterised by a conservative amino acid change from valine to isoleucine, while the second SNP at codon 280 (C>T) involves a non-conservative amino acid change from threonine to methionine. *CX3CR1*T280M has been associated with inconsistent effects on disease progression [Faure et al., 2000; McDermott et al., 2000a; Hendel et al., 2001, Faure et al., 2003; Kwa et al., 2003a] (see Chapter 2.1), while *CX3CR1*V249I showed an association with reduced susceptibility for HIV-1 infection (see Chapter 2.1). The allele frequencies reported indicate a higher occurrence of *CX3CR1*V249I compared to *CX3CR1*T280M in various populations [Faure et al., 2000; Faure et al., 2003; Kwa et al., 2003a; Singh et al., 2005] (see Chapter 2.1). Functional studies have shown reduced ligand binding affinity and impaired HIV-1 co-receptor activity when considering *CX3CR1*V249I and *CX3CR1*T280M occurring together [Faure et al., 2000; McDermott et al., 2000a].

The *CX3CR1* gene is located in close proximity to the *CCR5* and *CCR2* genes [Maho et al., 1999]. No linkage disequilibrium was however observed when analysing the distribution of *CX3CR1*V249I and *CX3CR1*T280 in the presence

of *CCR2*V64I, *CCR5*Δ32 and the *CCR5* promoter variants. This finding further ensured that the functional effects being observed were not attributed to any of the known neighbouring gene variants, but rather the *CX3CR1* SNPs themselves [Faure et al., 2000], which requires ongoing investigation.

1.1.1.4. CXC chemokine receptor 4 (CXCR4)

CXCR4, (MIM# 162643) is utilised as a HIV-1 co-receptor by SI/X4 strains to facilitate virus entry [Bleul et al., 1996a, Oberlin et al., 1996]. The CXCR4 amino-terminal domain together with its extracellular loop 2 was found to be important for co-receptor activity [Brelot et al., 2000]. CXCL12 (SDF1) is the natural ligand for this co-receptor and suppresses HIV-1 infection by down-regulation of CXCR4 surface expression [Amara et al., 1997; Signoret et al., 1997].

The *CXCR4* gene is located at position 2q21 [Federsppiel et al., 1993; Herzog et al., 1993] and consists of a single intron and two exons. Both Exon 1 and 2 contain parts of the coding region that encodes for a protein of 352 amino acids [Feng et al., 1996; Caruz et al., 1998; Wegner et al., 1998].

Investigations aimed at determining the role of *CXCR4* gene variants in susceptibility to HIV-1/AIDS in different population groups have resulted in no significant findings [Cohen et al., 1998; Alvarez et al., 1998; Martin et al., 1998b]. This is mainly due to the fact that *CXCR4* is a highly conserved gene [Moriuchi et al., 1997] and therefore the presence and effect of a few rare mutations remains questionable. Previously reported *CXCR4* variants occurring

at low allele frequencies include two silent mutations, *CXCR4*-I261I (C>T) [Martin et al., 1998b] and *CXCR4*-K68K (A>G), and a non-conservative mutation, *CXCR4*-F93S (T>C), involving an amino acid change from phenylalanine to serine [Cohen et al., 1998]. All of these mutations have displayed insignificant functional effects [Cohen et al., 1998]. A more recent study confirmed that *CXCR4* is conserved in a genetically older African ethnic group and furthermore the absence of HIV-1/AIDS associations (see Chapter 2.2).

1.1.1.5. CC chemokine ligand 5 (CCL5)

CCL5 (MIM# 187011), more commonly known as RANTES, is a natural ligand for the principle HIV-1 co-receptor, CCR5 [Samson et al., 1996a]. CCL5 has been found to suppress HIV-1 infection of NSI/R5 strains by directly competing with the virion envelope gp120 for binding to CCR5 or by down-regulation of CCR5, which limits its cell surface expression [Cocchi et al., 1995; Arenzana-Seisdedos et al., 1996; Deng et al., 1996; Mack et al., 1998; Abdelwahab et al., 2003; Pastore et al., 2003]. Elevated levels of CCL5 have been observed in 'exposed yet uninfected' individuals [Paxton et al., 1996; Paxton et al., 1998; Zagury et al., 1998; Garzino-Demo et al., 1999; Paxton et al., 1999]. CCL5 production and circulating levels have also been inversely correlated with rates of disease progression to AIDS [Aukrust et al., 1998; Paxton et al., 2001]. Some studies have found no significant change in CCL5 production in the presence of HIV-1 [McKenzie et al., 1996; Moriuchi et al., 1996; Mazzoli et al., 1997]. Other studies have suggested that CCL5 may actually up-regulate virus replication [Schmidtmayerova et al., 1996a; Schmidtmayerova et al., 1996b;

Gordon et al., 1999]. These inconsistent results could be attributed to a more recent finding where the antiviral activity of CCL5 at initial infection was the same in macrophages and lymphocytes, but it appeared that these cells differentially modulated the inhibitory ability of CCL5 during virus replication [Gross et al., 2003]

The *CCL5* gene found at chromosome position 17q11.2-q12 [Donlon et al., 1990] comprises three exons and two introns. All three exons contain partial coding regions that encode for 91 amino acids, including a signal peptide of 23 amino acids (Exon 1) and a mature protein of 68 amino acids (Exon1, 2, and 3) [Nelson et al., 1993; Nomiyama et al., 1999].

There are four previously identified *CCL5* SNPs that have been associated with influencing susceptibility to HIV-1/AIDS. These include two promoter variants at positions -403G>A (rs2107538) and -28C>G (rs2280788) (relative to the transcription start site) [Liu et al., 1999a] and a variant in both the first intron, designated In1.1T>C (rs2280789) and 3' untranslated region, designated 3'222T>C [An et al., 2002]. Initial studies showed the *CCL5*-28C>G SNP to be associated with delayed progression to AIDS in Japanese [Liu et al., 1999a], while the *CCL5*-403G>A SNP was found to offer an increased risk for HIV-1 infection, but also faster progression to AIDS in Caucasians [McDermott et al., 2000b]. Using reporter assays, both promoter *CCL5* variants have been found to upregulate gene transcription [Liu et al. 1999a; Nickel et al., 2000]. The *CCL5* SNPs, -403G>A (5'UTR-403G>A), In1.1T>C (IVS1+307T>C) and 3'222T>C (3'UTR+222T>C, relative to stop codon) in various populations are

associated with increased risk for HIV-1 infection, while the *CCL5* In1.1T>C also accounts for a more rapid disease progression to AIDS [An et al., 2002]. The *CCL5* In1.1T>C SNP lies within an enhancing regulatory element and was found to down-regulate gene transcription [An et al., 2002].

The four *CCL5* SNPs are in strong linkage disequilibrium and form four common haplotypes, R1 to R4 [An et al., 2002]. Therefore these SNPs individually or as part of derived haplotypes have been analysed in different populations for confirming or establishing associations with HIV-1/AIDS susceptibility [Liu et al., 1999b; Gonzalez et al., 2001; An et al., 2002] (see Chapter 2.1). More recently, it was suggested that *CCL5* variants may down-regulate gene expression and thereby increase initial HIV-1 plasma levels [Duggal et al., 2005].

1.1.1.6. CXC chemokine ligand 12 (CXCL12)

CXCL12 (MIM# 600835), previously called SDF1, is an extremely efficacious chemokine and the only known natural ligand for CXCR4, a major HIV-1 co-receptor [Bleul et al., 1996a; Bleul et al., 1996b; Oberlin et al., 1996]. CXCL12 was found to suppress HIV-1 infection of SI/X4 strains by down-regulation of CXCR4 surface expression and thereby interfering with virus fusion and entry [Bleul et al., 1996a; Oberlin et al., 1996; Amara et al., 1997; Signoret et al., 1997]. Reduced levels of CXCL12 were observed in persons infected with SI/X4 viruses compared to those infected with NSI/R5 viruses. Increased CXCL12 expression could therefore explain why the more cytopathic SI/X4 viruses do not appear in certain individuals [Llano et al., 2001]. A direct correlation between CXCL12 level and CD4$^+$ cell count has also been reported

[Derdeyn et al., 1999], which suggested an association between lower CXCL12 levels and progression to AIDS as found in another study [Soriano et al., 2002]. This finding was however in contrast to other studies where higher CXCL12 levels were found in HIV-1 infected individuals compared to their uninfected counterparts and an inverse correlation was observed with the CD4$^+$ cell count [Ikegawa et al., 2001; Shalekoff and Tiemessen, 2003].

The *CXCL12* gene, located at position 10q11.1, consists of four exons and three introns [Shirozu et al., 1995]. *CXCL12* encodes 2 isoforms, CXCL12α (89 amino acids) and CXCL12β (93 amino acids) [Tashiro et al., 1993; Nagasawa et al., 1994; Shirozu et al., 1995] due to alternative splicing of a single gene. The first 21 amino acids of both CXCL12α and CXCL12β form a signal peptide [Bleul et al., 1996b]. The coding regions for *CXCL12α* and *CXCL12β* are found within exons 1 to 3 and exons 1 to 4, respectively [Shirozu et al., 1995].

A *CXCL12β* polymorphism, designated *SDF1*-3'A (rs1801157), has been identified in the 3' untranslated region at position +801 (relative to the start codon) and involves a G to A transition [Winkler et al., 1998]. Initially it was found that the recessive state of *SDF1*-3'A is associated with slower disease progression to AIDS [Hendel et al., 1998; Martin et al., 1998a; Winkler et al., 1998] and hypothesised that the SNP up-regulates CXCL12 biosynthesis, which blocks infection of T-tropic/SI viruses that utilise CXCR4 as a HIV-1 co-receptor [Winkler et al., 1998]. Other studies have however shown inconsistent findings including, association with faster progression to death [Mummidi et al., 1998; van Rij et al., 1998]; prolonged [van Rij et al., 1998] or decreased [Brambilla et

al., 2000] survival after AIDS is diagnosed; low CD4$^+$ cell counts [Balotta et al., 2000]; or no effect on disease progression [Meyer et al., 1999; Mangano et al., 2000; Ioannidis et al., 2001a].   Dominant effects of the *SDF1*-3'A SNP have also been reported where association with increased vertical transmission from mother to child in an African study [John et al., 2000]; rapid disease progression in HIV-1 infected children born to seropositive mothers [Tresoldi et al., 2002]; and resistance to HIV-1 infection in seronegative high-risk individuals [Tiensiwakul, 2004] or the absence thereof [Liu et al., 2004].   The *SDF1*-3'A SNP has been identified in various population groups, but occurs more commonly in Caucasians compared to Africans.   Based on the findings to date, the HIV-1/AIDS associations observed with *SDF1*-3'A appear to be population specific (see Chapter 2.2).

Furthermore, plasma levels of CXCL12 in relation to *SDF1*-3'A genotypes have been considered in HIV-1 seropositive patients, exposed high-risk HIV-1 seronegative individuals and healthy HIV-1 seronegative controls [Llano et al., 2001; Soriano et al., 2002; Tiensiwakul, 2004] (see Chapter 2.2).   However, inconsistent results were found including the *SDF1*-3'A homozygous genotype being associated with higher [Tiensiwakul, 2004] and lower [Soriano et al., 2002] CXCL12 levels in exposed uninfected individuals.   In a recent study, it was shown that other polymorphisms in linkage disequilibrium with the *SDF1*-3'A SNP are responsible for altered gene transcription, rather than *SDF1*-3'A itself [Kimura et al., 2005].   This emphasises the need for investigating derived *CXCL12* haplotypes in various population groups for determining associations with susceptibility to HIV-1/AIDS.

1.1.2. Th1 and Th2 cytokines

Cytokines are a large group of small proteins that mediate immune responses by forming a signaling network between host cells. CD4$^+$ T lymphocytes can be classified into T-helper (Th) cell subsets (Th1 and Th2) based on the cytokines they secrete. Th1 cells produce cytokines such as INFγ, IL2 and IL12 that are important for driving cell-mediated immunity by stimulating cytotoxic T cell development. Th2 cells produce cytokines such as IL4, IL5 and IL13 that activate the humoral immune response by promoting antibody production. Additionally there are cytokines such as TNFα and IL10 that are secreted by T cells, but more predominantly by other cell types (e.g. macrophages). Many researchers have however combined inflammatory cytokines such as TNFα with the characteristic Th1 cytokines, while IL10 is considered to be a Th2 cytokine [Reviewed in Kidd et al., 2003].

A hypothesis by Clerici and Shearer in 1993 suggested that an imbalance in Th1-type and Th2-type responses occurs during HIV-1 infection resulting in immune dysregulation. It was further proposed that resistance to HIV-1 infection and disease progression to AIDS is largely dependant on a Th1>Th2 dominance (see Figure 3). This 'Th1 to Th2 switch' model was based on findings that HIV-1 exposed uninfected individuals generate strong Th1-type responses to HIV-1 antigens and that those individuals who progressed to AIDS showed reduced IL2 and INFγ production together with an increase in IL4 and IL10 levels [Clerici et al., 1993; Clerici and Shearer, 1994; Clerici et al., 1997]. Additional studies have shown a significant decrease in Th1 cytokines IL12 and INFγ and a significant increase in Th2 cytokines IL4, IL5 and IL10 during HIV-1

infection [Maggi et al., 1994; Autran et al., 1995; Barker et al., 1995; Meroni et al., 1996; Klein et al., 1997; Wasik et al., 1997].
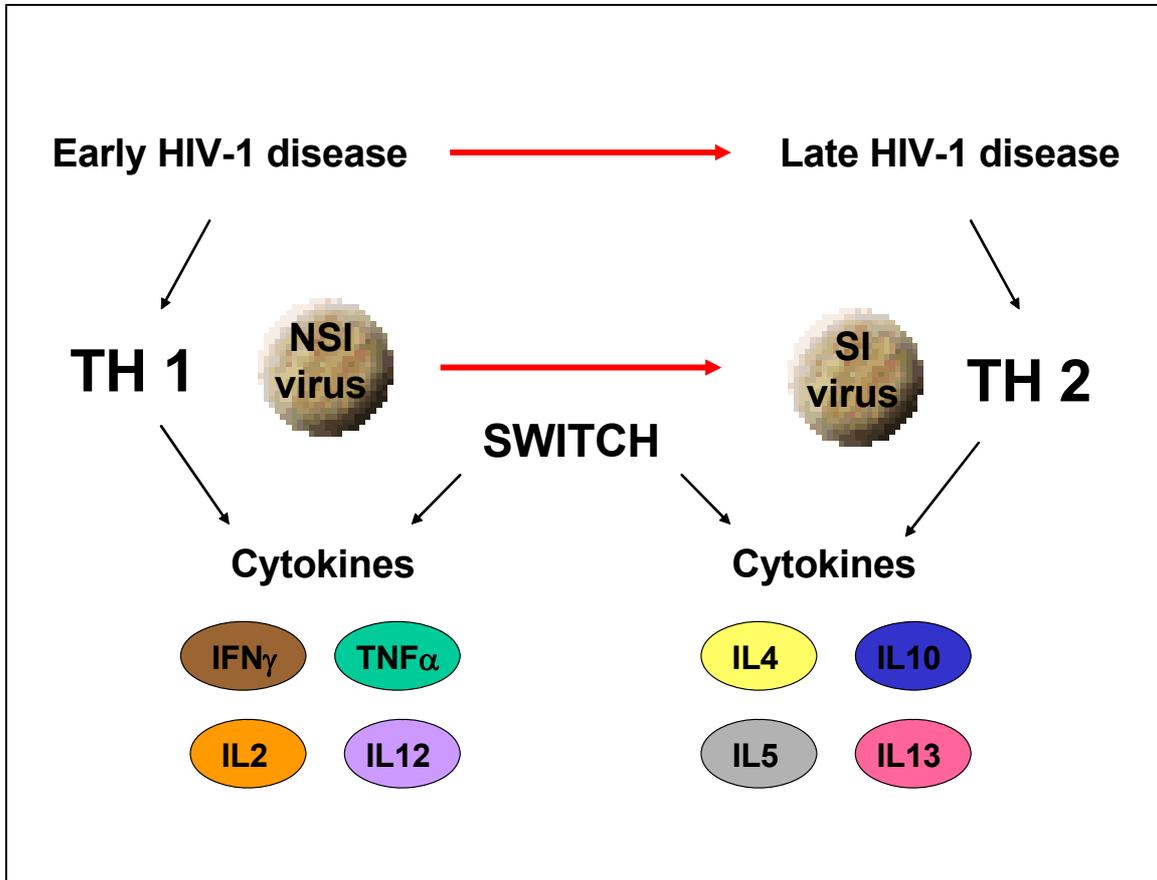
Figure 3. Schematic illustration of the 'Th1 to Th2 switch' model and phenotypic conversion from NSI to SI in the presence of specific cytokines.

There are however studies in disagreement with the 'Th1 to Th2 switch' model. These include the suggestion that HIV-1 replication preferentially occurs in Th2-type cells rather than a switch from a Th1 to Th2 cytokine profile [Maggi et al., 1994; Romagnani et al., 1994]. Other findings were the nearly undetectable expression levels of IL2 and IL4 irrespective of disease stage; CD8[+] cells expressing large and stable levels of IFNγ and IL10 [Graziosi et al., 1994]; and

a direct correlation between decreasing IL2 producing cells and reduced CD4[+] counts [Tanaka et al., 1999].

It has been proposed that the phenotypic conversion from NSI to SI is also associated with a switch in Th1 to Th2 cytokine profile (see Figure 3). A study of HIV-1 infected individuals who converted to the SI virus phenotype showed baseline significant lower levels of IL2 and higher levels of IL4 when compared to those infected persons who did not acquire SI variants of HIV-1. Shortly after SI-conversion, the HIV-1 infected individuals were characterised by significantly high levels of IL4 and low levels of IFNγ [Torres et al., 1998]. Additional studies have further investigated the presence of specific cytokines being linked to the emergence of HIV-1 virus strains with distinct tropisms [Suzuki et al., 1999; Galli et al., 2001].

It is evident that cytokine production during HIV-1 exposure serves as a mediator of virus-host interactions and influences the rates of disease progression to AIDS in those individuals who do become infected. The effects of cytokines on HIV-1 infection in cells of the macrophage lineage have also now been classified as being suppressive (e.g. IL10), stimulatory (e.g. TNFα) or bifunctional, i.e. both suppressive and stimulatory (e.g. IL4) [Reviewed in Kedzierska et al., 2003]. Therefore based on previous findings it is apparent why the genes encoding specific cytokines, such as *TNFα*, *IL10* and *IL4*, have been selected as candidates for determining genetic variants that may influence HIV-1/AIDS susceptibility.

1.1.2.1. Tumour necrosis factor alpha (TNFα)

TNFα (MIM# 191160) is a multifunctional pro-inflammatory cytokine also serving as a potent inducer of HIV-1 replication. It activates a cellular transcription factor, NF-κB, that enhances virus expression by binding to the HIV-1 long terminal repeat within the viral promoter [Nabel and Baltimore, 1987; Duh et al., 1989; Folks et al., 1989; Osborn et al., 1989; Matsuyma et al., 1991; Mellors et al., 1991]. Recently, the signal transduction pathway for the TNFα and NF-κB interacting protein components was mapped [Bouwmeester et al., 2004]. Increased levels of TNFα have been reported in persons who had progressed to AIDS [Brinkman et al., 1997]. It has been suggested that differences among HIV-1 strains in their ability to activate secretion of TNFα could be related to different rates of disease progression [Khanna et al., 2000]. Elevated TNFα activation by HIV-1 subtype C found in Southern Africa was associated with the presence of at least three NF-κB sites. This number of NF-κB sites is more than for other subtypes having only one or two. The significance of this finding on HIV-1 pathogenesis requires further study [Montano et al., 2000].

The *TNFα* gene lies within the highly polymorphic major histocompatibility complex (MHC) region at chromosomal position 6p21.3 [Nedwin et al., 1985; Spies et al., 1986]. It comprises four exons and three introns with a coding region encoding 233 amino acids. This includes a putative signal peptide of 76 amino acids (within Exon 1 and 2) and a mature protein of 157 amino acids (within Exon 2, 3 and 4) [Nedwin et al., 1985].

Many polymorphisms have been identified in the promoter region of the *TNFα* gene. These include two G-A transitions at positions –308 (rs1800629) and –238 (rs361525) relative to the transcription start site [Wilson et al., 1992; D'Alfonso et al., 1994; Hamann et al., 1995]. Both these SNPs have been analysed in HIV-1/AIDS association studies [Brinkman et al., 1997; Knuchel et al., 1998; Smolnikova and Konenkov, 2002] (see Chapter 3.1). Previous significant findings in Caucasian-based studies include a weak recessive effect of the *TNFα*-308G>A SNP associated with long-term non-progression [Knuchel et al., 1998] and the *TNFα*-308G/A heterozygous genotype associated with rapid progression to AIDS [Smolnikova and Konenkov, 2002]. The *TNFα*-308G>A SNP is more commonly observed than the *TNFα*-238G>A in different populations [McGuire et al., 1994; Conway et al., 1997; Baena et al., 2002] (see Chapter 3.1). The functional significance of both *TNFα*-308G>A and *TNFα*-238G>A is not clearly understood due to inconsistent findings regarding its influence on gene transcription and protein production [Reviewed in Reynard et al., 2000 and Hajeer et al., 2001].

The close proximity of the *TNFα* gene to the human leukocyte antigen (*HLA*) genes has resulted in the analysis of derived *TNFα–HLA* haplotypes for association with HIV-1/AIDS susceptibility [Wilson et al., 1993]. *HLA* alleles have been previously found to influence risk for HIV-1 infection and disease progression to AIDS [Reviewed in Carrington and O'Brien, 2003]. Furthermore, a few *HLA* haplotypes containing specific alleles have been associated with varying levels of TNFα production [Bendtzen et al., 1988; Jacob et al., 1990; Abraham et al., 1993].

1.1.2.2. Interleukin 4 (IL4)

IL4 (MIM# 147780) is a pleiotropic cytokine responsible for differentially regulating CC chemokine receptor 5 (CCR5) and CXC chemokine receptor 4 (CXCR4), two major HIV-1 co-receptors. This involves down-regulation of CCR5 with inhibition of early stage NSI/R5 virus replication in macrophages and T lymphocytes [Valentin et al., 1998; Wang et al., 1998a, Jinquan et al., 2000] and up-regulation of the CXCR4 with enhanced replication of the later emerging SI/X4 viruses in T lymphocytes [Valentin et al., 1998; Wang et al., 1998b]. It has been suggested that resistance to HIV-1 infection among African commercial sex workers is associated with reduced IL4 HIV-1 specific responses, independent of changes in other Th2 cytokines [Trivedi et al., 2001]

The *IL4* gene, located at position 5q31.1 [Sutherland et al., 1988; Le Beau et al., 1993], consists of four exons and three introns [Arai et al., 1989]. The coding region encodes for 153 amino acids, including a putative signal peptide of 24 amino acids (Exon 1) and a mature protein of 129 amino acids (Exon 1,2,3 and 4) [Yokota et al., 1986].

Two *IL4* promoter SNPs, *IL4*-589C>T (rs2243250) [Rosenwasser et al., 1995] and *IL4*-33C>T (rs2070874) [Takabayashi et al., 1999] (positions relative to the translation start site), and their derived haplotypes have been associated with influencing susceptibility to HIV-1/AIDS [Nakayama et al., 2000; Vasilescu et al., 2003; Wang et al., 2004] (see Chapter 3.2). The *IL4*-589C>T SNP, which is in complete linkage disequilibrium with the *IL4*-33C>T SNP in Japanese, was associated with decreased susceptibility to HIV-1 infection. However, the

IL4-589C>T SNP in its homozygous state was correlated with a more rapid emergence of SI/X4 virus strains and possibly faster disease progression to AIDS [Nakayama et al., 2000]. In contrast, the *IL4*-589C>T SNP was associated with delayed acquisition of SI/X4 virus strains in Caucasians, but no overall effect on disease progression was observed [Kwa et al., 2003b]. Another Caucasian-based study found the *IL4*-589C>T SNP offers slower progression to AIDS and death [Nakayama et al., 2002]. This finding was confirmed in an additional study where a specific haplotype associated with delayed disease progression carries the *IL4*-589 T allele [Vasilescu et al., 2003]. In an African-based population, homozygosity for the *IL4*-589C>T and *IL4*-33C>T SNPs has been associated with slower disease progression and increased risk for HIV-1 infection, respectively [Wang et al., 2004].

Functional studies have shown significant findings for *IL4*-589C>T, including increased promoter activity and transcription, suggesting that increased IL4 levels are expressed in the presence of the SNP [Rosenwaser et al., 1995; Song et al., 1996]. Previous HIV-1/AIDS associations observed with the *IL4* promoter variants and derived haplotypes do however appear to differ between populations (see Chapter 3.2).

1.1.2.3.   Interleukin 10 (IL10)

IL10 (MIM# 124092) is a vital regulatory cytokine that inhibits HIV-1 replication in macrophages. [Kollmann et al., 1996; Schols and De Clercq, 1996] This control of virus proliferation is presumably due to restricting the amount of macrophages available for HIV-1 replication [Edelman et al., 1996; Pataraca et

al., 1996; Than et al., 1997; Muller et al., 1998]. IL10 was previously found to block secretion of pro-inflammatory cytokines such as TNFα and IL6, thus further inhibiting HIV-1 replication [Weissman et al., 1994]. IL10 has also been reported to differentially regulate CCR5 and CXCR4 expression in various cell types, which is another indication of its important role in susceptibility to HIV-1 infection [Houle et al., 1999; Patterson et al., 1999; Jinquan et al., 2000, Torres et al., 2001, Wang et al., 2002]

The *IL10* gene was localised to the chromosomal position 1q31-q32 [Kim et al., 1992; Eksdale et al., 1997]. The gene is comprised of four exons and three introns with a coding region encoding 178 amino acids. This includes a putative signal peptide of 18 amino acids (Exon 1) and a mature protein of 160 amino acids (Exon 1, 2, 3, and 4) [Vieira et al., 1991].

Polymorphisms in the *IL10* promoter region that have been associated with influencing susceptibility to HIV-1/AIDS individually or as part of extended haplotypes are located at positions -1082A>G (rs1800896), -819C>T (rs1800871) and -592C>A (rs1800872) relative to transcription start site [Turner et al., 1997; Shin et al., 2000; Vasilescu et al., 2003; Wang et al., 2004] (see Chapter 3.2). In a Caucasian-based study it was found that *IL10*-592C>A, in complete linkage disequilibrium with *IL10*-819C>T and in strong linkage disequilibrium with *IL10*-1082A>G, is associated with accelerated disease progression to AIDS, particularly during the late disease stage [Shin et al., 2000]. Another study involving Caucasians however showed a haplotype that contains the *IL10*-592C allele to be associated with faster disease progression

[Vasilescu et al., 2003]. The homozygous *IL10*-1082A/A genotype in Hispanics and an *IL10* haplotype comprised of 5 alleles, including -592C, -819C and -1082G, in an African-based population was also associated with a higher risk for HIV-1 infection [Wang et al., 2004].

The *IL10*-592C>A SNP is functionally significant by reducing gene transcription and decreasing IL10 production [Rosenwasser et al., 1997; Crawley et al., 1999; Shin et al., 2000], while the functional effect of *IL10*-1082A>G remains debatable [Hoffman et al., 2001; Rees et al., 2002]. The associations between *IL10* promoter SNPs and haplotypes with HIV-1/AIDS susceptibility have been observed in distinct population groups and further research is required for confirmation of previous findings (see Chapter 3.2).

1.1.3. Immunoregulatory proteins

Additional immunoregulatory proteins that play a functional role in providing effective immune responses have been discovered. These include MBL and SLC11A1 (NRAMP1), which are both essential proteins acting during host-pathogen interactions. MBL is vital in immune defence, particularly during the stage of primary contact with microorganisms [Reviewed in Turner, 2003; Klein, 2005]. SLC11A1 is responsible for the transport of iron into bacterium-containing phagosomes and thereby regulates intracellular pathogen proliferation and macrophage inflammatory responses [Reviewed in Forbes and Gros, 2001; Blackwell et al., 2003]. Identifying genetic variants in both the *MBL* and *SLC11A1* genes for possible associations further advanced the study of host genetic factors influencing HIV-1/AIDS pathogenesis.

1.1.3.1. Mannose binding lectin (MBL)

MBL (MIM# 154545) is a calcium dependent serum protein produced by the liver [Kawasaki et al., 1983] as an acute phase response [Thiel et al., 1992] and binds to pathogens, including HIV-1 [Ezekowitz et al., 1989; Haurum et al., 1993; Saifuddin et al., 2000]. The importance of MBL in innate immunity therefore involves binding to the carbohydrate-rich domains on pathogens for destruction by either opsonophagocytosis [Kuhlman et al., 1989] or activation of the lectin complement pathway [Matsushita and Fujita, 2001]. It remains debatable as to whether MBL-binding results in virus neutralisation or enhances infection by providing another mode for virus entry [Sölder et al., 1989; Holmskov et al., 1994; Thielens et al., 2002]. Inconsistent associations also exist between MBL levels and susceptibility to HIV-1/AIDS [Senaldi et al., 1995; Prohászka et al., 1997] (see Chapter 4).

The *MBL* gene located at 10q11.2-21 [Sastry et al., 1989; Schuffenecker et al., 1991] consists of four exons and three introns [Taylor et al., 1989]. The coding region encodes 248 amino acids, including a putative signal peptide of 20 amino acids (Exon 1) and mature protein of 228 amino acids (Exon 1, 2, 3, 4) [Ezekowitz et al., 1988; Taylor et al., 1989]

Three *MBL* SNPs occurring in the coding region at codons C52R (C>T) (rs5030737) [Madsen et al., 1994], D54G (G>A) (rs1800450) [Sumiya et al., 1991] and E57G (G>A) (rs1800451) [Lipscombe et al., 1992] have been previously implicated in HIV-1/AIDS pathogenesis [Garred et al., 1997; Maas et al., 1998; Pastinen et al., 1998; Mombo et al., 2003] (see Chapter 4). These

SNPs are also known as the D (*MBL*C52R), B (*MBL*D54G) and C (*MBL*E57G) alleles with A being the wild-type allele. The SNPs represent non-conservative amino acid changes that disrupt oligomerisation and result in impaired protein function [Sumiya et al., 1991; Wallis and Cheng, 1999] and have also been linked to reduced MBL levels [Garred et al., 1992a; Garred et al., 1992b; Lipscombe et al., 1992; Madsen et al., 1994; Turner, 1996]. Previous Caucasian-based findings include homozygosity for any combination of *MBL* SNPs associated with increased susceptibility to HIV-1 infection [Garred et al., 1997; Pastinen et al., 1998] and *MBL* variants resulting in slower disease progression [Maas et al., 1998] and shorter survival after AIDS diagnosis [Garred et al., 1997]. *MBL*E57G in both the homozygous and compound heterozygous state was associated with higher risk for HIV-1 infection in an African population, while individuals heterozygous for the SNP were less susceptible than those homozygous for the wild-type allele [Mombo et al., 2003].

The *MBL* SNPs are associated with reduced serum MBL levels [Garred et al., 1992a; Garred et al., 1992b; Lipscombe et al., 1992; Madsen et al., 1994; Turner, 1996] that could result in opsonisation impairment [Super et al., 1989] and failure to defend against HIV-1 infection. The effect of functional *MBL* SNPs and their derived genotypes on susceptibility to HIV-1/AIDS therefore requires further investigation to fully elucidate the population-based associations that have been previously observed (see Chapter 4).

1.1.3.2. Solute Carrier 11A1 (SLC11A1)

SLC11A1 (MIM# 600266), more commonly known as NRAMP1, is a divalent cation transporter that plays an important role in iron metabolism and recycling, thereby regulating susceptibility to infectious and autoimmune disease [Reviewed in Blackwell et al., 2003]. The monocyte/macrophage cell lineage has a key function during HIV-1 infection and the macrophage-expressed SLC11A1 protein was therefore considered for possibly modulating individual risk for HIV-1/AIDS [Marquet et al., 1999].

The *SLC11A1* gene is localised to chromosome position 2q35 [Blackwell et al., 1995; Liu et al., 1995; Marquet et al., 2000] and comprises 15 exons separated by 14 introns. The coding region is contained within Exons 1 to 15 and encodes for 550 amino acids [Cellier et al., 1994; Blackwell et al., 1995].

Two previous studies have analysed the influence of *SLC11A1* variants on susceptibility to HIV-1/AIDS [Marquet et al., 1999; Donninger et al., 2004] (see Appendix D). The genotypes of four markers, including a GT repeat sequence in the *SLC11A1* promoter region [Liu et al., 1995; Blackwell et al., 1995] has been associated with altered risk of HIV-1 infection in a Caucasian-based population [Marquet et al., 1999]. The length of the GT repeat was shown to have a functional effect on *SLC11A1* promoter activity and expression levels [Searle and Blackwell, 1999; Blackwell et al., 2003] (see Appendix D). A recent African-based study further investigated the presence of *SLC11A1* promoter variants and although no association was observed for previously reported markers or three novel mutations and susceptibility to HIV-1 infection, gene

expression studies showed enhanced promoter activity for both a previously reported SNP and two novel mutations [Donninger et al., 2004] (see Appendix D).

Limited studies have focused on the role of *SLC11A1* in determining risks for HIV-1/AIDS.   Establishing the functional significance of known *SLC11A1* variants has provided an explanation for how *SLC11A1* variants could possibly influence susceptibility to HIV-1/AIDS.   The functional findings will also contribute to further identifying and confirming associations between *SLC11A1* and HIV-1/AIDS in distinct population groups.

1.2. Genetic association studies

Complex traits such as host genetic susceptibility to HIV-1/AIDS are likely to display genetic heterogeneity (different mutations in numerous genes resulting in the same effect) or have a polygenic nature (combination of mutations in multiple genes acting collectively) when determining infection/disease risk profiles. A number of associations implicating various candidate genes in HIV-1/AIDS have therefore been observed. Genetic association studies can be performed on families or the general population depending on the specific outcome being analysed. Family-based studies can involve either a multigenerational pedigree for locating candidate genes using linkage analysis or the case-control design where relatives of the cases are used as the controls. It is relatively difficult to collect large informative families and this is particularly true for determining HIV-1/AIDS susceptibility where the number of infected cases within a family is either low or generally unknown due to the stigma and discrimination that may result from the disclosure of HIV-1 status amongst relatives. The population-based analysis therefore serves as a more feasible approach for HIV-1/AIDS association studies. It includes the testing of a genetic variant for 1) an increased occurrence in either the cases or their unrelated population-matched controls (risk for HIV-1 infection) and 2) correlation with a defined phenotype within a study cohort (rates of disease progression to AIDS). The candidate gene approach, SNPs, family-based versus population-based association studies, criteria for ensuring successful population-based association studies, together with the sample group presented in this dissertation (see Chapters 2 to 4), are discussed in this chapter subsection.

1.2.1. Candidate gene approach

It is evident that a genetic component exists for many complex traits, but often the underlying mechanisms of these genetic factors remain unknown. Various strategies for identifying genes with a distinct contribution to a complex trait have shown only a certain degree of success. This is mainly due to the fact that several genes each with relatively weak effects and strongly interacting with both other genes and the environment are often involved in determining a specific trait outcome. A candidate gene is defined by evidence of its possible role in the trait that is being investigated. The candidate gene approach is therefore either 1) based on the location of the gene within a previously determined region of linkage or 2) focuses on genes selected for their protein product having a plausible function in a biological pathway or in an interaction appropriate for the trait of interest [Lander and Schork, 1994; Taylor et al., 2001; Hirschhorn and Daly, 2005; Suh and Vijg, 2005]. Although the latter does rely on limited existing knowledge of candidate genes with hypothesised functional variants, it has formed the basis of many successful HIV-1/AIDS association studies [Reviewed in Carrington et al., 2001; Hogan and Hammer, 2001; Dean et al., 2002; Anastassopoulou and Kostrikis, 2003; O'Brien and Nelson, 2004; Winkler et al., 2004; Kaslow et al., 2005] (see Chapters 2, 3 and 4).

Although the most comprehensive analysis of a candidate gene is obtained by resequencing of the entire gene in cases and controls and searching for genetic variants with heterogeneity between the two groups, this process is laborious and costly. Association studies focusing on commonly occurring genetic variants therefore offer a simpler and more cost-effective strategy to elucidate

complex traits [Hirschhorn and Daly, 2005]. The functional genetic variants selected for genotyping could occur in exons, introns, untranslated regions and promoter regions. They can therefore be classified as being either coding variants resulting in amino acid changes (non-synonymous) and thereby altering the structure and reducing optimal functioning of the resulting protein product or as regulatory variants in non-coding regions, that modulate gene expression and influence RNA stability and splicing by altering regulatory elements [Rebbeck et al., 2004; Hirschhorn and Daly, 2005; Knight 2005; Newton-Cheh and Hirschhorn, 2005].

Single nucleotide non-synonymous missense genetic variants in the coding region are either conservative (replacement of an amino acid by another with similar chemical properties) or non-conservative (replacement of an amino acid by another with different chemical properties) with the non-conservative changes usually having a more potent effect on altering the protein product. There are also other non-synonymous genetic variants (insertions, deletions and nonsense mutations involving premature stop codons) and dynamic/unstable mutations (repeat nucleotides) that occur in the coding region and that could possibly result in nucleotide reading frameshifts [Mueller and Young, 1998; Hirschhorn and Daly, 2005]. Regulatory genetic variants (single nucleotide, insertions, deletions and repeat nucleotides) can be classified as being either cis-acting (variant located either in or close to the gene that it affects) or trans-acting (variant in one gene affecting another gene) with presently only potential cis-acting variation being considered for complex trait studies [Mueller and Young, 1998; Knight, 2005].

There are synonymous mutations occurring in the coding regions that do not result in amino acid changes, but were also shown to be regulatory variants [D'Souza et al., 1999; Lorson et al., 1999]. This further emphasises the importance of additional functional studies for determining the role of all significant trait-associated genetic variants before assuming that it may only act as a marker for another known or yet unidentified functional genetic variant in the same or a closely lying gene.

1.2.2. Single nucleotide polymorphisms (SNPs)

The completion of the human genome project [Lander et al., 2001; Venter et al., 2001], the distribution of SNPs throughout the human genome deposited into public databases (e.g. dbSNP, SNP Consortium, SNPper) [Sachidanandam et al., 2001], and the more recent initiation of the HapMap Project [Gibbs et al., 2003] has largely contributed to the designing of genetic association studies. The SNP has been identified as the most common genetic variant and together with other sequence variation (e.g. insertions, deletions and repeat nucleotide polymorphisms) offers the possibility of identifying the genes that directly influence complex trait outcomes [Gray et al., 2000; Sachidanandam et al., 2001; Taylor et al., 2001; Venter et al., 2001].

A SNP is a substitution of a single nucleotide base that occurs at a frequency of more than 0.01 or 1%. The two alleles for each individual SNP are designated as "major" and "minor" based on their observed frequency in the general population. These diallelic SNPs have a low rate of recurrent mutation and are therefore stable compared to other more mutable multiallelic variants such as

short tandem di, tri or tetranucleotide repeats (microsatellites). On average SNPs occur every 1000 – 2000 bases of genomic sequence and are present in high density for applying extensive genetic measures (linkage disequilibrium and haplotype analysis). The SNP densities do however differ among specific chromosomal regions and within genes for various population groups. In comparison to SNPs, other types of genetic variants are relatively sparse, making SNPs the ideal genetic markers appropriate for large-scale association studies of complex traits using high-throughput SNP genotyping methods [Gray et al., 2000; Kwok, 2001; Sachidanandam et al., 2001; Taylor et al., 2001; Crawford and Nickerson, 2005; Hirschhorn and Daly, 2005; Suh and Vijg, 2005].

1.2.3. Family-based versus population-based association studies

Family-based linkage studies involve multigenerational pedigrees (diagram showing ancestral relationships and transmission of genetic traits in a family) as models to define the physical link between genetic variants and thereby identify the location of a trait-causing gene. These studies are particularly useful for simple Mendelian recessive or dominant traits involving rare highly penetrant alleles (all those with the predisposing allele will manifest the trait), but are confounded by incomplete penetrance (those with the predisposing allele may not manifest the trait) and by phenocopy (those without the predisposing allele still manifest the trait due to environmental factors). It is therefore difficult to identify large families to efficiently apply linkage analysis to complex traits where common low penetrant alleles with moderate effects and possibly interacting with the environment will not provide a suitable model explaining a definite pattern of inheritance [Lander and Schork, 1994; Taylor et al., 2001].

Association studies of complex traits using the family-based case-control design have been attempted. This involves the use of parents, siblings and cousins of the cases as controls to avoid the effect of ethnic confounding due to population stratification. Family members with common environments are also particularly useful when considering gene-environment interactions. The design is however restricted by the availability of parents and siblings and often the difficulty in recruiting extended family members, who must belong to the same ethnic group. Statistical power is also reduced as the cases and their relatives are more likely to have the same genotype compared to the cases and the unrelated controls used in population-based association studies [Schaid and Rowland, 1998; Gauderman et al., 1999]

The family-based case-control design using pseudosibling controls is another approach where no actual controls are selected but rather genotypes of the parents are considered when comparing the genotype of the case to the three genotypes (pseudosiblings) that were not transmitted to the case. The transmission disequilibrium test (TDT) can be used to determine association with a specific allele or genotype occurring more commonly in cases than in their pseudosiblings by testing many cases and their parents. Although this design also protects against population stratification, the genotype data needs to be available for both parents to prevent bias in the TDT and a sufficient number of families are needed for extensive analysis of complex traits [Lander and Schork, 1994; Gauderman et al., 1999; Witte et al., 1999]

Population-based case-control and cohort study designs are widely used and due to the size limitations of family-based study designs, they have been the preferred option for investigating associations between variants in candidate genes and complex traits. The case-control approach involves affected cases directly compared to their population-matched unaffected controls, while the study cohort approach involves defined groups of affected cases followed over time for comparison of outcomes [Lander and Schork, 1994; Taylor et al., 2001; Hirschhorn and Daly, 2005; Wang et al., 2005].

Therefore in HIV-1/AIDS population-based association studies, it is essential that all individuals (cases and controls) are drawn from the same well-defined population to avoid spurious findings arising from population stratification. A successful HIV-1/AIDS association study investigating allele and genotype distributions in a specific population is also dependant on, various other confounding factors; extended measures (linkage disequilibrium and haplotypes); computational software for statistical analysis and reproducibility of findings in matching ethnic groups [O'Brien et al. 2000; Huber et al., 2003; Clark and Dean, 2004], which are further discussed in 1.2.3.1. to 1.2.3.5.

1.2.3.1. Confounding factors of population-based association studies

A well-designed population-based study is characterised by its ability to overcome all confounding factors for the correct interpretation of results and valid association findings. These factors include size of study sample, ensuring no selection bias by obtaining study samples in a random and blinded manner, choice of genetic variants, genotyping methods used for screening and the

statistical tests selected for data analysis [Ioannidis et al., 2001b; Corfield and Brink, 2002; Thomas and Witte, 2002; Ioannidis et al., 2003a; Ioannidis 2003b; Rebbeck et al., 2004; Newton-Cheh and Hirschhorn, 2005; Suh and Vijg, 2005]. Population stratification is an additional factor that is of widespread concern for causing spurious associations in case-control studies of admixed populations. Population stratification refers to heterogeneity of allele frequencies observed between cases and controls that are from a population containing a number of genetically diverse subpopulations where the prevalence of a specific trait may differ [Lander and Schork, 1994; Thomas and Witte, 2002; Cardon and Palmer, 2003]. The likelihood for population stratification increases with larger sample sizes, but presently there are various methods to detect and adjust for this confounding factor [Pritchard and Rosenberg, 1999; Pritchard et al., 2000; Pritchard and Donnelly, 2001; Falush et al., 2003; Hoggart et al., 2003; Freedman et al., 2004; Hinds et al., 2004; Marchini et al., 2004].

Association studies have however been performed using admixture mapping where genetic markers informative for ancestry form the basis of this approach. It involves the use of samples from recently admixed populations (e.g. African Americans) to identify susceptibility gene regions as some trait-influencing alleles have different frequencies amongst ancestral populations that have contributed to the admixture. The admixed populations should have an increased probability of inheriting the alleles from the ethnic group with the increased number of trait-influencing alleles. This approach is therefore particularly suitable for traits that largely differ in incidence between the ethnic groups contributing to the admixture [Patterson et al., 2004; Smith et al., 2004;

Montana and Pritchard, 2004; McKeigue 2005]. The Cape Coloured population from South Africa is of recent admixed descent, including San, Khoi, African Negro, Madagascan, Javanese and Caucasian [Hayes, 2003] and could also therefore be considered for identifying ancestral informative markers (AIMs).

The comparison of allele and genotype distributions in cases versus controls during association studies is important when considering consistency with Hardy-Weinberg equilibrium (HWE) expectations. Deviations from the HWE in unaffected controls from the general population could be indicative of non-random mating, unbalanced mutation rates, selection bias, small population sizes, gene flow, population stratification or even problematic genotyping methods [Schaid and Jacobson, 1999; Salanti et al., 2005; Wigginton et al., 2005]. Differences between the observed and HWE expected genotype frequencies could also be used to identify alleles that are more likely to be functionally associated with influencing a specific trait outcome. An example is deviation from the HWE in cases where the alleles and genotypes that are trait-causative would be over-represented when compared to the controls [Schaid and Jacobson, 1999; Wigginton et al., 2005]. This was observed in an African population from South Africa for a SNP occurring in the *IL4* promoter region, *IL4*-589C>T. The heterozygous genotype (CT) was over-represented in the HIV-1 infected cases and therefore showed deviation from the HWE ($P = 0.0305$) (see Chapter 3.2). It is recommended that HWE should only be tested in controls. An approach that corrects for HWE deviations in the general population and thereby decreases the chance of false-positive association has been reported [Schaid and Jacobson, 1999].

1.2.3.2. Linkage disequilibrium (LD)

LD is the tendency of alleles at linked loci to occur together on the same chromosome more often than expected by chance and is the result of factors that disturb HWE. The extent of LD between susceptibility and marker alleles disperses as future generations of the population originate and recombination occurs. The absence of population association with alleles at two or more loci is therefore referred to as linkage equilibrium and is dependant on new generations since the trait allele arose and the genetic distance between the trait and marker alleles [Huttley et al., 1999; Olson et al., 1999]. The extent of LD varies across populations and ethnic groups (see Figure 4).

Several LD measures have been developed, including pairwise LD analysis based on Lewontins D' [Lewontin, 1964] and the r^2 coefficient [Hill and Robertson, 1968] where a value of 1 for both measures indicates complete LD (see Figure 4). D' equals 1 in the presence of two or three out of the four possible allele combinations and will therefore be less than 1 if four allele combinations are present, while r^2 equals 1 in the presence of two out of four allele combinations and will therefore be less than 1 if three or four allele combinations are present. D' is directly related to the frequency of recombination (exchange of alleles between parent chromosomes during meiosis) indicating the physical extent of LD over time, while r^2 is more informative in association studies as it is inversely related to the sample size required for determining significance. Although both D' and r^2 are subjected to biases in small sample sizes and with rare genetic variants, r^2 has more reliable properties in such instances [Reviewed in Jorde 2000; Wang et al., 2005].

CCR2-CCR5 region (chromosome 3p21)
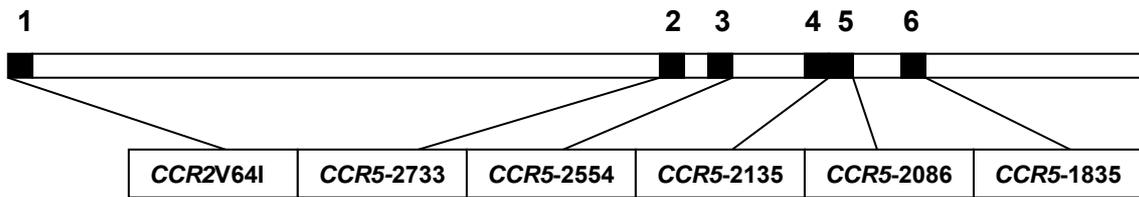
African

	1	2	3	4	5	6	
1		0.08	0.29	0.46	0.17	**0.91**	r^2
2	0.66		0.17	0.26	0.10	0.05	
3	1.00	1.00		0.65	0.60	0.33	
4	1.00	1.00	1.00		0.39	0.51	
5	1.00	1.00	1.00	1.00		0.19	
6	1.00	0.36	1.00	1.00	1.00		

D'

Cape Coloured

	1	2	3	4	5	6	
1		0.06	0.24	0.37	0.21	**0.94**	r^2
2	0.62		0.19	0.28	0.17	0.01	
3	0.87	1.00		0.70	**0.89**	0.28	
4	0.93	1.00	1.00		0.62	0.40	
5	0.87	1.00	1.00	1.00		0.24	
6	0.94	0.15	1.00	1.00	1.00		

D'

Caucasian

	1	2	3	4	5	6	
1		0.01	0.23	0.13	0.22	**0.94**	r^2
2	0.18		0.19	0.22	0.18	0.01	
3	1.00	1.00		**0.85**	**0.96**	0.25	
4	0.47	1.00	1.00		**0.80**	0.17	
5	1.00	1.00	1.00	1.00		0.24	
6	1.00	0.15	1.00	0.56	1.00		

D'

Figure 4. The extent of pairwise linkage disequilibrium for the *CCR2-CCR5* region shown for control samples representative of various South African populations that were studied by our research group. The values in bold indicate strong linkage disequilibrium.

LD extends over shorter distances in older populations and is more apparent in younger isolated populations [Lander and Schork, 1994]. The older African populations therefore have higher levels of genetic diversity and less LD compared to other younger non-African populations [Tishkoff and Williams, 2002; Kittles and Weiss, 2003; Tishkoff and Verrelli, 2003; Bamshad et al., 2004; Sawyer et al., 2005]. This was evident when measuring LD in various South African population groups where the African population displayed less LD than the Caucasian population for SNPs occurring within the *CCR2-CCR5* region. Although D' appears to show stronger LD in the Africans compared to the Caucasians, r^2 is considered more reliable when taking into account the sample sizes. The Cape Coloured population of admixed descent, including African and Caucasian, displayed varying LD values (see Figure 4).

Population-based association studies involving genes occurring in regions of high LD are cost effective as it is unnecessary to genotype SNPs in strong LD with others, but rather analyse subsets of tag SNPs reflecting most of the allelic variation [Hirschhorn and Daly, 2005; Wang et al., 2005]. The identification of significant associations would however rely on identifying the underlying LD patterns in the specific population being studied [Tishkoff and Williams, 2002].

1.2.3.3. Haplotype analysis

The significance of LD in a population is shown using chromosomal haplotype analysis. A haplotype is therefore a combination of alleles at different loci along the same chromosome that are inherited as a unit. Haplotypes provide information regarding recombination and are important for locating trait-causing

alleles [Clark et al., 2004; Crawford and Nickerson, 2005]. Both molecular and automated methods are available for constructing haplotypes in population-based studies. The molecular methods (allele-specific polymerase chain reaction and somatic cell hybrids) are however costly and time-consuming and therefore the automated method using statistical inference software programs are generally used. These programs are classified into groups depending on the use of specific algorithms (Clark algorithm, expectation-maximization algorithms and Bayesian algorithms) for determining haplotypes [Reviewed in Crawford and Nickerson, 2005].

In an African population from South Africa, the haplotype analysis of two SNPs located in the promoter region of the *IL4* gene (*IL4*-589C>T and *IL4*-33C>T) showed a different number of allele combinations in HIV-1 infected cases when compared to their population-matched uninfected controls. Using the expectation-maximization algorithm, which allows for assignment of all alleles to haplotypes with high probability, it was estimated that the cases had three haplotypes (CC, TC, and TT) versus the four haplotypes (CC, TC, TT and CT) of the controls. The fourth haplotype (CT) occurred exclusively in the controls and indicated statistical significance for an association with a decreased risk for HIV-1 infection (see Figure 5). To date, this fourth haplotype occurring in an older African population group has not been previously reported for any other populations (see Chapter 3.2).

It has been shown that the human haplotype structure for various populations can be defined as haplotype blocks. These are large regions with little evidence

for recombination and therefore only a few common haplotypes are observed. The advantage of haplotype blocks is the genotyping of a small number of tagSNPs to represent haplotypes for population-based association studies [Gabriel et al., 2002]. These findings resulted in the initiation of the International HapMap project aimed at describing the common patterns of human variation in different populations by identifying chromosomal regions with haplotypes and the tagSNPs that represent them for future genome-wide association studies [Gibbs et al., 2003; Wang et al., 2005; Hirschhorn and Daly, 2005].

Locus 1   **Locus 2**

IL4-589C>T   *IL4*-33C>T

— C/T ———————————————— C/T —

Allele combinations

— C ———————————————— C —
— T ———————————————— C —
— C ———————————————— T —
— T ———————————————— T —

| | Frequency (%) | | |
Haplotype	Cases	Controls	*P* value
C – C	26.1	24.3	0.53
T – C	19.5	21.4	0.57
C – T	0	3.4	**0.0013***
T – T	54.3	50.9	0.35
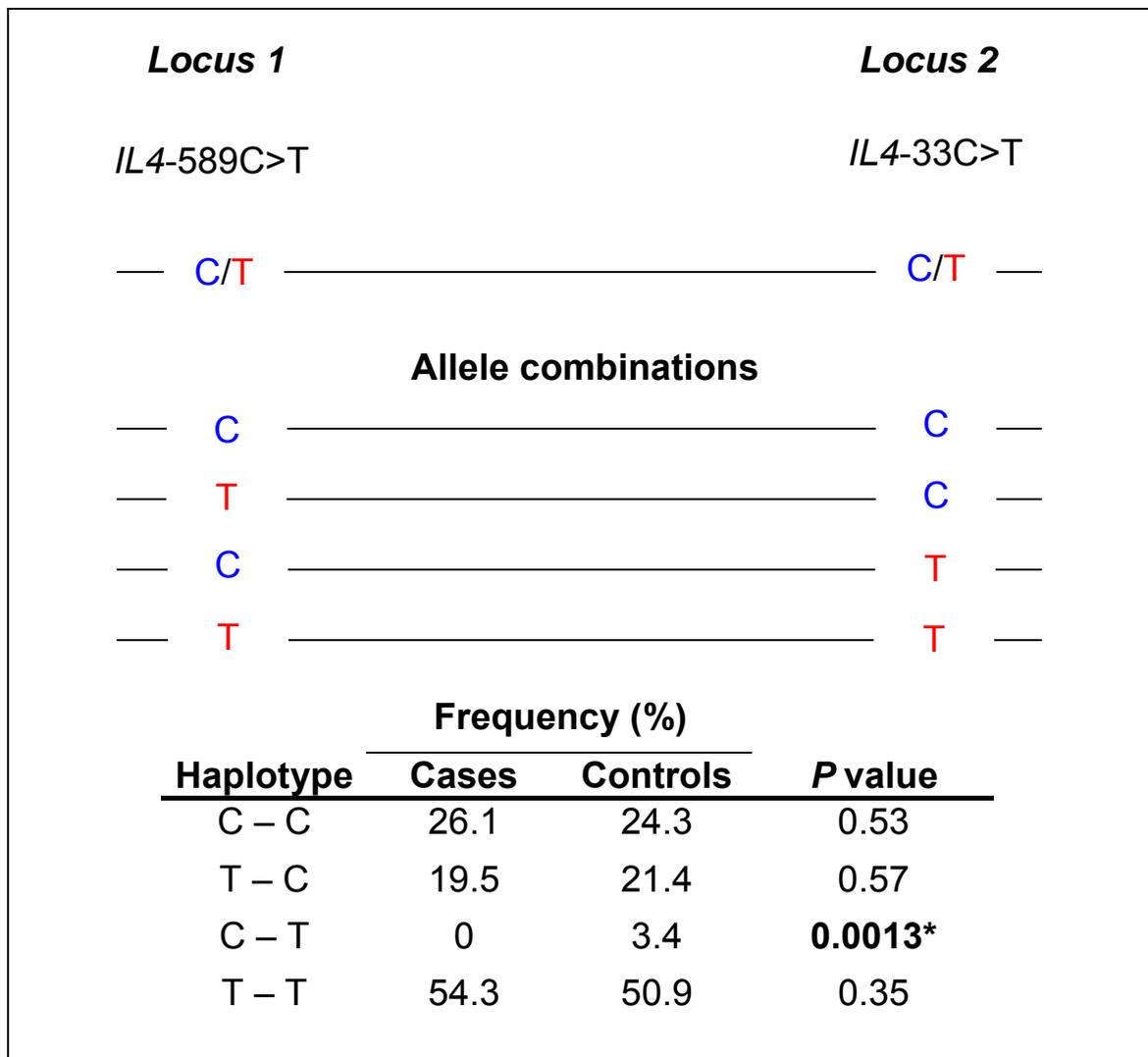
Figure 5. The various allele combinations for two *IL-4* promoter SNPs are shown for an African population from South Africa. One haplotype (CT) occurs exclusively in HIV-1 uninfected controls and indicates an association with decreased HIV-1 risk having a significant *P* value of 0.0013 (see Chapter 3.2).

1.2.3.4. Computational programs for statistical analysis

The validity of genetic association studies is largely dependant on statistical analysis, which involves the assembling, presentation and interpretation of the sample data. The increasing availability of information on genetic variation being used for extensive association studies has prompted the development of reliable computational programs (e.g. SAS/Genetics, Stata, GraphPad InStat, PS and R). These programs include the application of various statistical tests for comprehensive analysis of population-based case-control data [Fallin et al., 2001; Tsai et al., 2003]. The statistical analysis performed for the studies presented in this dissertation included the calculation of allele and genotype distributions and estimations of linkage disequilibrium (LD) and haplotype measures for genetic variants in HIV-1 seropositive cases versus HIV-1 seronegative controls. These analyses were mainly achieved using the SAS/Genetics (SAS Institute Inc. software, Cary, North Carolina) and GraphPad InStat (GraphPad Software Inc, San Diego, California) computational programs and the threshold set for statistically significant P values was less than 0.05 (see Chapters 2 to 4).

The SAS/Genetics program provides optimal advanced statistical testing for the data obtained from large sample numbers. It creates data sets and controls analysis with certain procedures. The ALLELE procedure is used for preliminary analysis of the study sample and genetic variants. It includes the calculation of allele and genotype frequencies by determining the proportion of samples carrying a particular allele or genotype when considering the allele and genotype distributions for the entire sample group. Genotype distributions are

also tested for consistency with Hardy Weinberg equilibrium (HWE) using the chi-square goodness-of-fit test. The extent of LD among alleles at paired loci is estimated using Lewontins D' and the r^2 coefficient measures. The CASECONTROL procedure is designed to compare the allele and genotype frequencies between case and control samples in populations of unrelated individuals. The allele frequencies are compared using the chi-square test for 2x2 contingency tables, while genotype comparison is based on the chi-square test using 2x3 contingency tables. A significant difference in heterogeneity for allele and/or genotype distribution indicates association with the presence or absence of a trait. The HAPLOTYPE procedure estimates the multilocus haplotype frequencies with the highest probability based on the analysis of observed sample genotypes under the assumption of HWE using the expectation-maximization algorithm. It also allows for association testing with a large number of haplotypes for trait outcomes by using data available for case-control studies. The P value is calculated by permutation testing which empirically assesses the probability of significance arising by chance (www.sas.com)

The GraphPad InStat program is designed to perform standard statistical analysis. This includes the Fisher exact test for 2x2 contingency tables that is used for determining significant heterogeneity in allele frequencies between cases and controls. The P value calculated for the Fisher exact test is more accurate for the study sample sizes presented in this dissertation (Chapters 2 to 4) and was therefore chosen above the chi-square test, which is more suitable for larger study sizes with thousands of samples (www.graphpad.com).

The PS computational program for power and sample size calculations can determine the level of significance detected with a given power and specified sample size [Dupont and Plummer, 1990]. In a case-control study consisting of 257 HIV-1 seropositive patients and 113 HIV-1 seronegative controls, as presented in this dissertation, there is sufficient power (80% or higher at 0.05 significance level) to detect a protective association of Odds Ratio (OR) = 0.5 or a susceptible association of OR = 1.9 with common variants (assuming an allele frequency of 45%), and a protective association of OR = 0.2 or susceptible association of OR = 2.6 with less common variants (assuming an allele frequency of 10%).

Adjustments of statistical significance when performing multiple testing (e.g., the Bonferroni adjustment) in exploratory studies are considered unnecessary as data are collected with an objective rather than a predefined hypothesis. Multiple testing is however essential in confirmatory studies where the aim is to prove a predefined hypothesis and the results from the multiple statistical tests are combined for determining the final outcome [Perneger et al., 1998; Bender and Lange, 2001]. The description of the analysis performed and the reasons given for selecting specific statistical tests are therefore sufficient for the exploratory studies presented in this dissertation (Chapters 2 to 4).

1.2.3.5. Reproducibility of population-based association studies

The inconsistencies in results generated between different studies investigating the same trait are not always indicative of biases as genetic associations may show various levels of significance in different populations [Ioannidis et al.,

2003a]. Although the frequencies of the genetic variants associated with a trait could vary between populations, the effect on the trait outcome is suggested to be generally consistent amongst different population groups [Ioannidis et al., 2004]. Significant associations could arise from: 1) functional effect of the variant on gene expression; 2) confounding factors; and 3) the variant being in LD with the causative allele located nearby [Cardon and Palmer, 2003].

Testing for reproducibility has been achieved using a meta-analysis approach [Ioannidis et al., 2001b; Lohmueller et al., 2003]. One of these meta-analyses involved a total of 370 studies investigating 36 various genetic associations for different disease outcomes and showed that significant differences are frequently observed between studies for the same trait. It was found that the findings for the first study performed correlated only moderately with further research of the same association. The first study also often suggested a stronger genetic effect. It was suggested that this could be due to the fact that the earlier studies may have included more bias and subpopulation diversity that resulted in an overestimation of association with a genetic variant [Ioannidis et al., 2001b]. A second meta-analysis of 301 studies investigating 25 different reported associations showed the reproducibility of the first report to be statistically significant for less than half of these associations. It was suggested that false negative and underpowered studies contribute to inconsistent association findings. They also concluded that definite effects of common variants would be confirmed in studies using large study samples [Lohmueller et al., 2003].

Another review of associations between genetic variants and complex genetic traits indicated that most previously reported associations are not robust. It was found that for 166 associations studied three or more times, only 6 have been consistently reproducible (>75% of the studies were positive). This further suggested that caution should be taken when considering single previous reports of an association between a genetic variant and trait outcome [Hirschhorn et al., 2002].

There has been major concern and debate over the failure to reproduce previously observed associations [Thomas and Clayton, 2004], particularly in population-based studies for complex traits [Witte et al., 1999; Ioannidis et al., 2001; Thomas and Witte, 2002]. The possible reasons for lack of reproducibility has been reported extensively and are mainly due to: 1) a false-positive association being correctly not reproducible; 2) valid associations not reproduced in follow-up studies of inadequate statistical power or when the initial report of association is due to non-stringent statistical analysis used; and 3) heterogeneity caused by genetic or environmental factors that results in the presence of an association in one population, but lack of reproducibility in another population [Cardon and Palmer, 2003; Newton-Cheh and Hirschhorn, 2005]. The more rapid publication of statistically significant findings compared to non-significant findings could also provide a misrepresentation for reproducibility of studies [Ioannidis et al., 2001b].

Future population-based association studies should therefore be designed using, relatively large sample sizes collected randomly with no prior knowledge

of genetic variation for both cases and controls; genetic variants with preferably known functional consequences; reliable screening methods to ensure high-sensitivity mutation detection; appropriate analysis of genetic marker measures for determining significance; and relatively homogenous populations or adjustment for population stratification in admixed ethnic groups.

1.2.4. HIV-1 infection and AIDS in South Africa

The estimated number of people globally living with HIV/AIDS at the end of 2004 was 39.4 million. This includes 25.4 million individuals from regions of Sub-Saharan Africa, which accounts for more than 60% of all infections [UNAIDS/WHO, December 2004 (www.unaids.org)]. The HIV/AIDS epidemic in South Africa continues to grow with an estimated 6.5 million infected people in 2004, the highest for any one country in the world, and a sharp increase from the previous 2003 estimate of 5.6 million. The latest figure reflects that approximately 13.8% of a total 47 million South Africans have HIV/AIDS [Department of Health, South Africa, 2005].

The Department of Health survey for women attending antenatal clinics across South Africa form the main basis for both regional and national HIV/AIDS estimates. The overall HIV/AIDS prevalence among pregnant women was 27.9% in 2003 and increased to 29.5% in 2004. The 2004 survey also indicated that different geographical regions (provinces) of South Africa show significant variation in HIV/AIDS prevalence for pregnant women. The highest prevalence of 40.7% is for the Kwazulu Natal Province, while the Western Cape Province

has the lowest prevalence of 15.4% (see Figure 6) [Department of Health, South Africa, 2005].
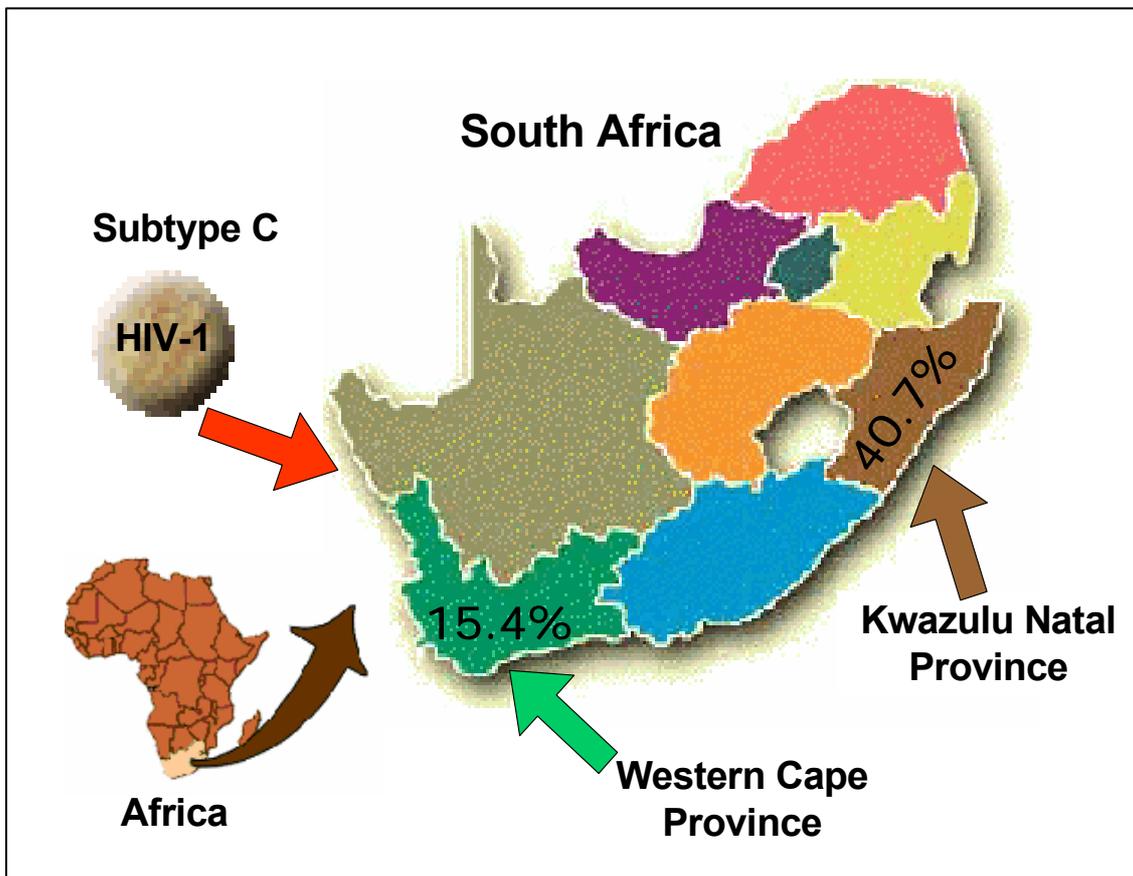
Figure 6. HIV/AIDS subtype C prevalence in South Africa based on the survey of women attending antenatal clinics. The HIV/AIDS prevalence rate varies for specific geographical regions of South Africa, with the highest in the Kwazulu Natal Province and the lowest in the Western Cape Province [Department of Health, South Africa, 2005]. The HIV-1 Subtype C is the most commonly found subtype in all the nine provinces (represented by different colours) of South Africa.

Two independent and distinct HIV-1 epidemics exist for South Africa. The first originated in the early 1980's and has remained mainly restricted to homosexual males (HIV-1 subtypes B and D). The more predominant epidemic started in the late 1980's and spread rapidly among heterosexuals with still no signs of a

decline (HIV-1 subtype C) [Engelbrecht et al., 1995; Williamson et al., 1995; van Harmelen et al., 1997; Moodley et al., 1998]. The expansion of HIV-1 subtypes and their distribution for Africa and the world indicates extensive genetic diversity [Papathanasopoulos et al., 2003; Wainberg, 2004]. Although there is evidence for diversity of HIV-1 subtypes in South Africa, subtype C is most commonly found in all the different provinces (see Figure 6) [Engelbrecht et al., 1995; Engelbrecht et al., 1998; Moodley et al., 1998; Engelbrecht et al., 1999; van Harmelen et al., 1999; Bredell et al., 2000; Gordon et al., 2003; van Harmelen et al., 2003; Bessong et al., 2005; Loxton et al., 2005].

In South Africa, the main mode of HIV-1 transmission is via heterosexual intercourse [van Harmelen et al., 1997]. There are however distinct factors that contribute to the rampant epidemic, including the large migrant labor workforce, high incidence of sexually transmitted diseases, thriving commercial sex worker industry and poverty in particularly the rural regions [Bredell et al., 1998; Moodley et al., 1998; van Harmelen et al., 1999; Williams et al., 2003; Dunkle et al., 2004; Bessong et al., 2005, Ramjee et al., 2005; Zuma et al., 2005].

1.2.4.1. Study sample

South Africa represents a complex and rich diversity of ethnic backgrounds with 11 official languages and additional regional dialects. According to the most recent census in 2001, the South African population is classified as being 79% African, 9.6% Caucasian, 8.9% Cape Coloured and 2.5% Indian/Asian. Although the major part of the population is African, this does not refer to a culturally or linguistically homogenous group. The African population is divided

into four main ethnic groups, namely Nguni, Sotho, Shangaan-Tsonga and Venda. There are numerous subgroups of which the Xhosa and Zulu, both subgroups of the Nguni, are the largest. The Xhosa ethnic group forms approximately 22.3% of the total African population. African individuals form 26.7% of the population in the Western Cape Province with at least 23.7% belonging to the Xhosa ethnic group [Statistics South Africa (www.statssa.gov.za)].

The population-based case-control study sample presented in Chapters 2 to 4 of this dissertation are represented by Africans residing in the Western Cape Province of South Africa. These individuals are predominantly of Xhosa ethnic descent and therefore represent a relatively homogenous group. The HIV-1 seropositive individuals are patients of Tygerberg Hospital, Woodstock Chapel Street Community Health Clinic or the Langa Clinic, all in the Western Cape Province. The population-matched HIV-1 seronegative control individuals are healthy blood donors from the Western Province Blood Transfusion Service. The study sample also consisted of an additional group of South African Caucasians. These individuals mainly of German, Dutch, French or British descent were included in the study for the confirmation of African-based genetic variants and HIV-1/AIDS associations occurring exclusively in the Xhosa population (see Chapters 2 to 4).

1.3. Methodologies

Mutation detection forms a major part of all genetic association studies and the selection of appropriate genotyping methodologies can largely contribute to the success of a study. In addition to direct automated sequencing, a variety of methodologies do exist, including sequence non-specific or indirect pre-screening methods (e.g. cleavage, electrophoretic mobility shift and liquid chromatography assays) and sequence specific or direct screening methods (e.g. allelic discrimination assays involving hybridization, nucleotide incorporation, oligonucleotide ligation and invasive cleavage). These mutation detection methods differ in many ways with regards to their simplicity and reliability to detect genetic variants. An ideal mutation detection method should address the following aspects: 1) assays that are easily designed; 2) reagents and equipment that are cost-effective; 3) not require time-consuming or intensive manual labour; 4) have 100% sensitivity for detection of genetic variants and have high specificity for eliminating false positives or negatives; 5) yield reproducible results to ensure reliability; and 6) allow for easy visualisation for interpretation of results and accurate data analysis. It is however difficult to identify a single mutation detection method as being ideal and the final choice made is therefore largely based on the approach taken to ensure that all the necessary requirements of a particular study aim are met. The reasons for selecting certain methodologies for the detection of previously reported or novel mutations (denaturing gradient gel electrophoresis) and high-throughput SNP genotyping (TaqMan allelic discrimination method) in the studies presented in Chapters 2, 3 and 4 of this dissertation are discussed in this chapter subsection.

1.3.1. Denaturing gradient gel electrophoresis (DGGE)

DGGE, developed by Fischer and Lerman in 1983, is a pre-screening mutation detection method with a theoretical framework explaining the principles of the technique (see Figure 7). It involves the differential melting behaviour of double-stranded (ds) DNA molecules in an increasing concentration gradient of denaturants (urea and formamide, UF) at a fixed elevated temperature. The melting behaviour is highly sequence dependant and determined by the composition and order of nucleotide base pairs within a DNA fragment. As the dsDNA fragment passes through the denaturing gel it melts and undergoes a conformational change resulting in reduced electrophoretic mobility. This is shown in the schematic representation of a DGGE time-travel gel, which is loaded with the same amplified product at hourly intervals (see Figure 7B). The addition of a guanine and cytosine (GC)-rich fragment, known as a GC-clamp, to the 5' end of either the forward or reverse primers prevents complete strand dissociation during fragment amplification [Myers et al., 1985a]. The GC-clamp increases the mutation detection sensitivity to theoretically 100%, including the detection of single nucleotide bases [Sheffield et al., 1989; Abrams et al., 1990]. An additional heteroduplexing step, involving denaturation and renaturation of the wild-type and mutant DNA and the formation of mismatched heteroduplexes, further contributes to the high mutation detection sensitivity [Myers et al., 1985b] (see Figure 7A). The lower stability of the mismatched heteroduplexes causes it to melt earlier than the corresponding homoduplexes within the denaturing gel. Heterozygous mutations are therefore visualized as four bands (two heteroduplexes and two homoduplexes), while homozygous mutations are represented by a single "shifted" DGGE band (see Figure 7C).
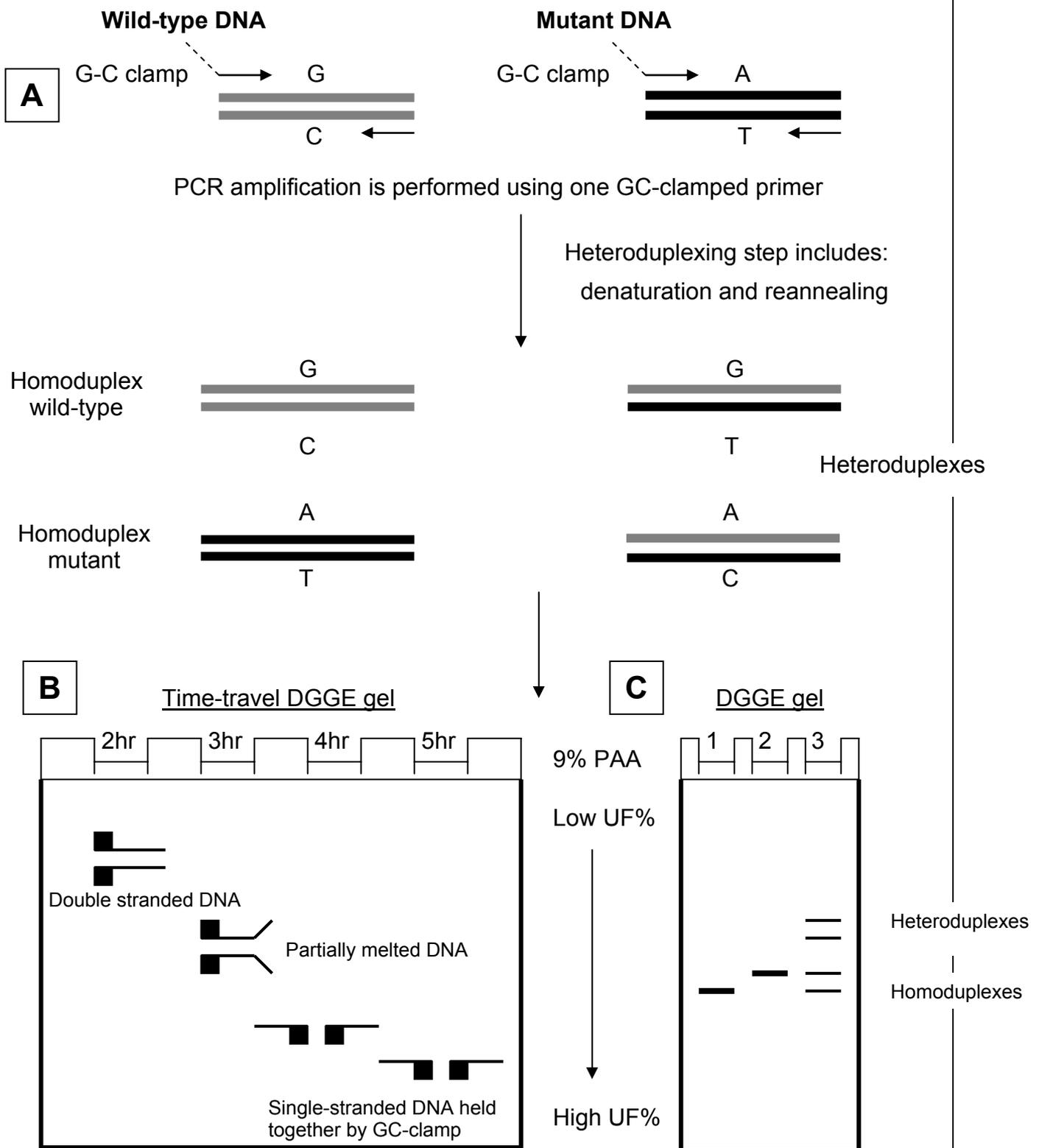
Figure 7. **A.** PCR amplification of DNA fragments for DGGE using a GC-clamped primer, followed by a heteroduplexing step, which involves denaturation and reannealing to form a wild-type homoduplex, a mutant homoduplex and heteroduplexes. **B.** The principles of DGGE are depicted as a time-travel DGGE-gel (Myers et al., 1987). **C.** The detection of mutations by electrophoresis through an increasing denaturing gradient (urea and formamide, UF), where the wild-type and mutant DNA have different melting profiles: 1) wild-type, 2) homozygous mutant, 3) heterozygous mutant (Adapted from Hayes, 1999).

DGGE is a PCR-based method allowing for optimal mutation detection of fragments up to 500bp in length [Hayes, 1999]. A successful DGGE-assay is dependant on the melting profile of the DNA and the choice of appropriate primers (see Figure 8). The melt 87 computer program [Lerman and Silverstein, 1987] is used to design DGGE primers and identify a single melting domain for the DNA fragment. There are however instances where the single melting domain can only be achieved by considering alternative fragment selection, changing the position of the GC-clamp or certain primer modifications such as the addition of T/A or G/C tails [Wu et al., 1998].
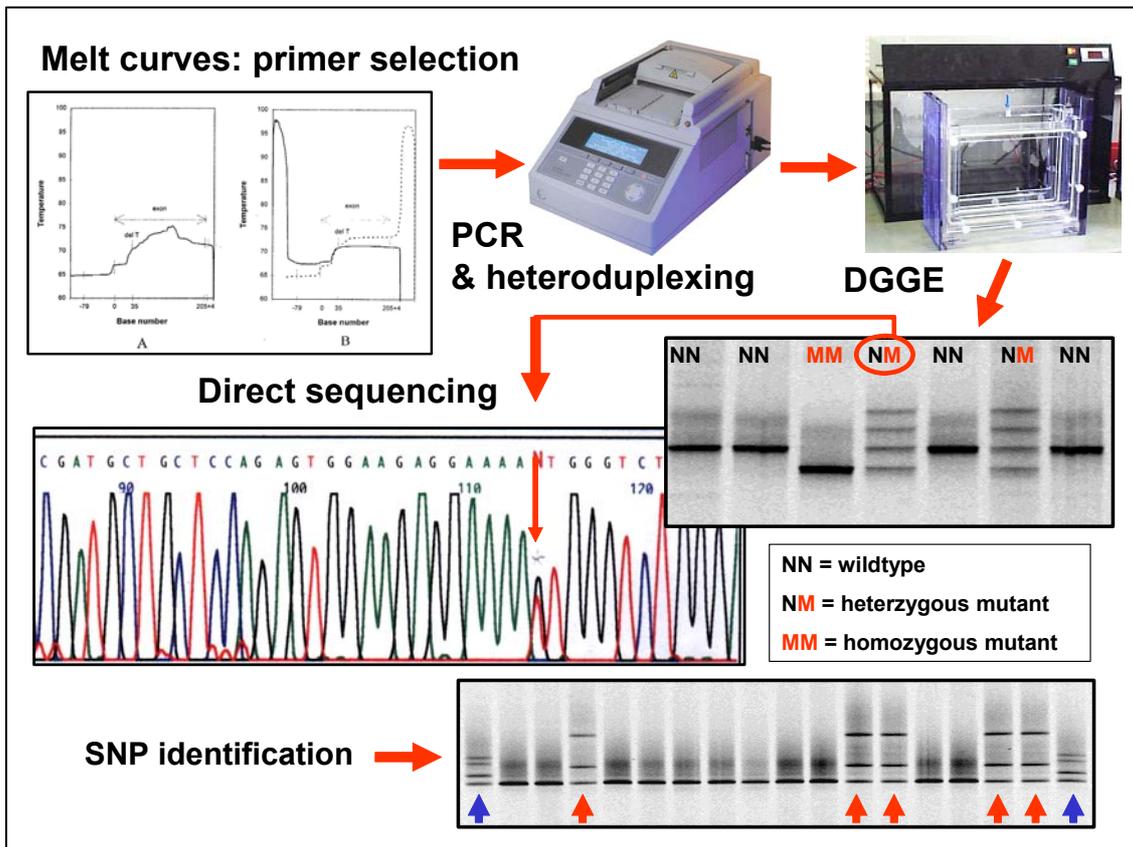
Figure 8. The DGGE methodology is depicted. An optimal DGGE assay is based on the melting profile of the DNA, the GC-clamped primer pair used for amplification and an additional heteroduplexing step. These all contribute to the detection of mutations identified by their specific banding pattern on a denaturing gel. Heterozygous mutations are visualised as four bands while homozygous mutations are visualised as a single shifted band compared to the wildtype. Samples showing aberrant banding patterns are subjected to direct sequencing and verification of commonly occurring mutations are achieved by the identification of the same banding pattern on a denaturing gel. (Partially adapted from Wu et al., 1998; Hayes et al., 2003; Hayes and Gardiner-Garden 2003).

A formula: %UF = [(melting temperature – buffer temperature) x 100 / 32] is used to calculate the amount of denaturant required for optimal melting of the DNA fragment. Although factors including the buffer composition, buffer concentration and electrophoretic voltage are not considered when applying the formula, they may additionally influence the melting behaviour of the DNA fragment. Furthermore, the use of a specific gel system (see Figure 8) and a single set of experimental conditions (gel composition, temperature and voltage), based on conditions previously described for improving broad-range mutation detection analysis, contributes to the reliability of the DGGE assay. These conditions include the use of 9% polyacrylamide with a denaturing gradient of 30 to 50% between the lowest and highest concentration of UF denaturant, and 0.5 X TAE buffer for electrophoresis at a voltage between 100 and 200 volts [Hayes, 1999; Hayes et al., 1999]. Adaptations of the DGGE method do exist, including constant denaturing gel electrophoresis (CDGE) [Børresen et al., 1991; Smith-Sørensen et al., 1993], temperature gradient gel electrophoresis (TGGE) [Sliutz et al., 1997] and two-dimensional (2-D) DNA electrophoresis [Rines et al., 1998].

The heteroduplexing step is vital for detecting single base deletions or insertions and C/G to G/C transversions as the homoduplexes may have a similar melting behaviour resulting in the same banding pattern being visualised, which means that mutations are undetected [Myers et al., 1985b]. The various mutations are therefore identified by a specific banding pattern (see Figure 8). Verification of commonly occurring mutations is achieved by the mixture of samples with similar banding patterns followed by heteroduplexing

and electrophoresis on a denaturing gel. Only samples showing additional heteroduplex bands are subjected to sequencing [Guldberg and Guttler, 1993]. Repeated automated direct sequencing of the same mutation is therefore eliminated, which reduces cost and is a huge advantage for SNP analysis within a large study sample (see Figure 8).

The designing of DGGE primers may be time-consuming and more costly with the GC-clamps attached to the primers. There are also cleavage-based pre-screening methods allowing for the detection of mutations in DNA fragments more than 1kb. The use of specific gel systems together with expertise for optimal standardisation of DGGE assays is an additional requirement for successfully mastering the technique. These disadvantages are however outweighed by the theoretical concept of high mutation sensitivity achieved using an optimised experimental condition, which allows for rapid mutation detection and easy visualisation of both novel and commonly occurring mutations. It has been recently reported that certain aspects of DGGE to some extent are regarded as being more sensitive for detecting mutations than electrophoresis based sequencing, which shows difficulty detecting heterozygotes unambiguously and is not 100% accurate for a specific base due to compression in GC rich regions [Edwards et al., 2005]. Considering the advantages above other pre-screening and direct gel-based mutation detection methods, DGGE was an appropriate method chosen for wide-range genetic variant analysis in the studies presented in Chapters 2, 3 and 4 of this dissertation (see Figure 9).
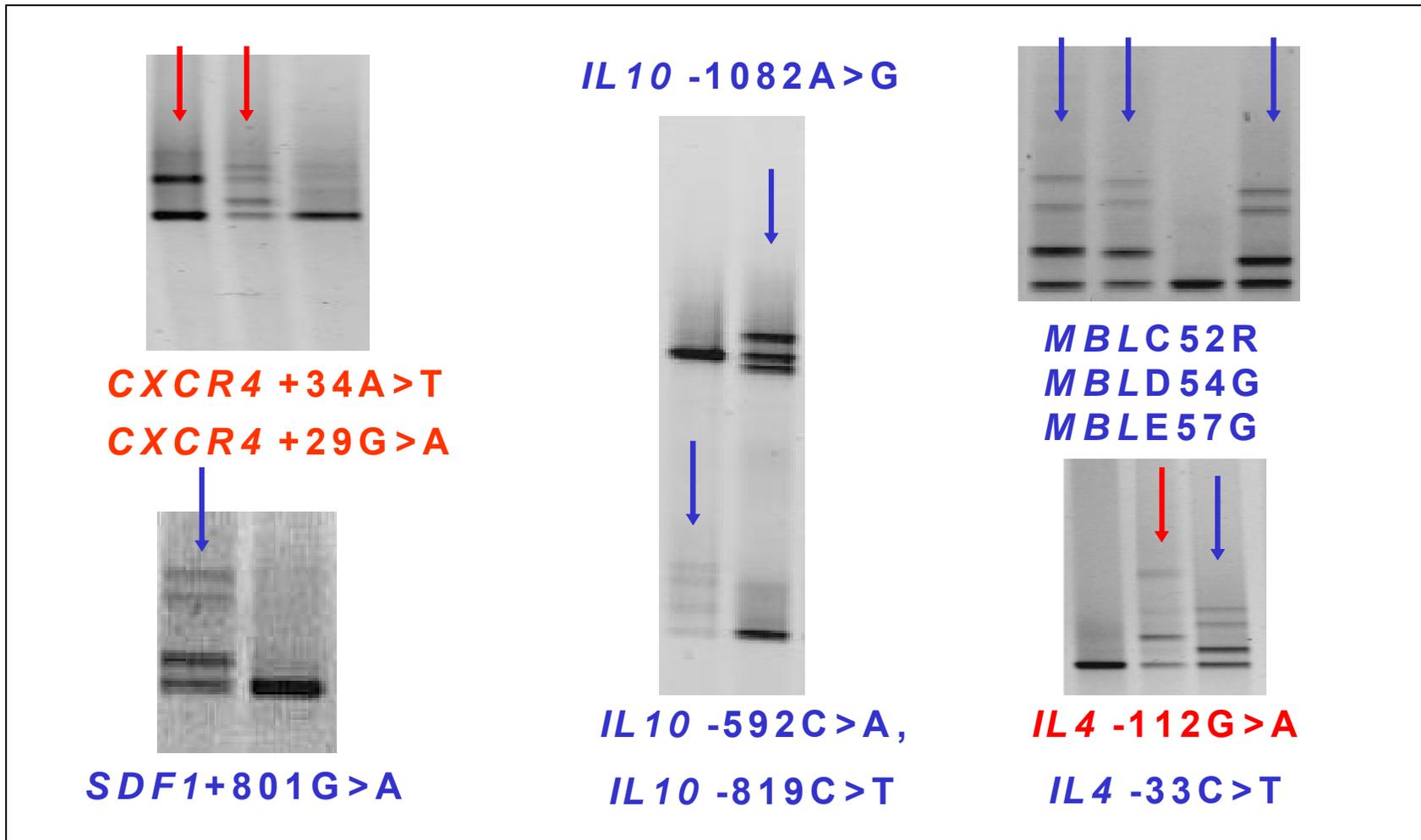
Figure 9. DGGE banding patterns are shown for a few mutations identified in the *CXCR4, CXCL12 (SDF1), IL10, IL4* and *MBL* genes (see Chapters 2 to 4). Previously reported mutations are indicated in blue and the novel mutations are indicated in red, according to the order in which they occur on the gel from left to right. The single bands represent wild-type control samples.

1.3.2. TaqMan allelic discrimination method

TaqMan allelic discrimination is a high-throughput genotyping method for previously identified SNPs and combines PCR amplification and mutation detection into a single step by using fluorogenic oligonucleotide probes in the 5' nuclease assay [Livak et al., 1995a; Livak et al., 1999] (see Figure 10). The 5' nuclease assay is based on the exonuclease activity of the *Taq* DNA polymerase and the use of a probe in the PCR reaction together with forward and reverse primers. If the target sequence of the probe is present during amplification, the probe will hybridise to the target during the annealing/extension PCR cycling step. The *Taq* DNA polymerase acts upon the template surface by removing obstacles downstream of the growing amplified product. Therefore when the *Taq* DNA polymerase encounters the hybrised probe, the probe is displaced and cleaved. Cleavage is however dependant on the hybridisation of the probe to its specific target sequence [Holland et al., 1991] (see Figure 10). The use of fluorogenic probes allows for cleavage to be detected without post-PCR processing and the need for electrophoresis. The fluorescent probe consists of a high-energy reporter dye at the 5' end and a low-energy quencher dye at the 3' end. When the probe is intact, the proximity of the quencher to the reporter reduces the fluorescent signal observed from the reporter dye. Cleavage by the *Taq* DNA polymerase results in increased fluorescence of the reporter dye as it is separated from the quencher dye. An increase in the reporter dye's characteristic fluorescence is therefore indicative of amplification for the probe-specific target sequence and nonspecific amplification is undetected [Lee et al., 1993; Livak et al., 1995b] (see Figure 10).
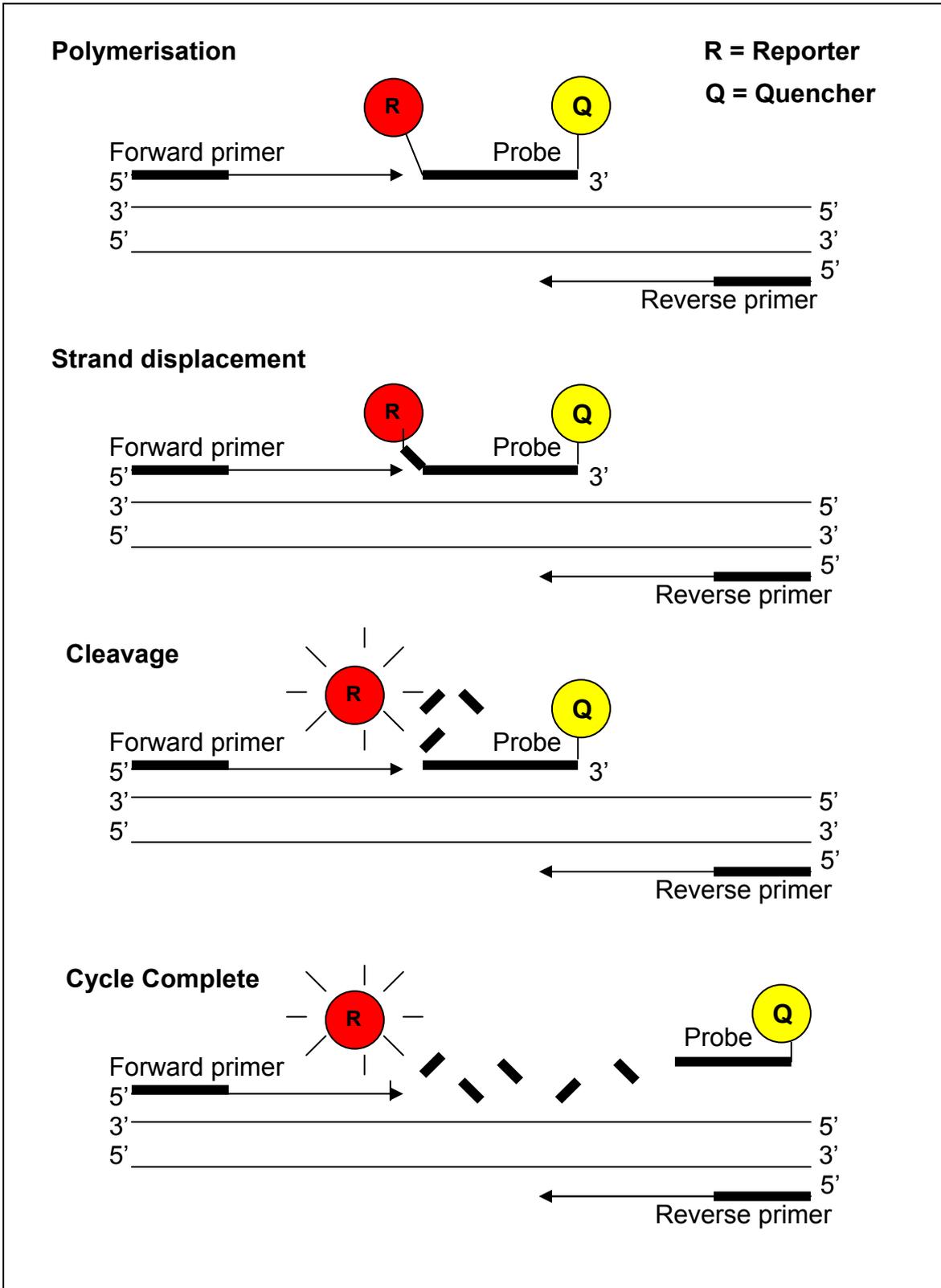
Figure 10. Schematic representation of the polymerisation-associated 5' nuclease activity of *Taq* DNA polymerase acting on a fluorogenic probe during amplification. A fluorescent reporter dye and a quencher dye are attached to probe at the 5' and 3', ends, respectively. When the probe is intact, the reporter's emission is quenched. During each PCR extension cycle, the *Taq* DNA polymerase cleaves the reporter from the probe. Once separated from the quencher, the reporter emits its characteristic fluorescence (Adapted from Livak et al., 1999).

The allelic discrimination method involves the designing of two fluorogenic probes, for the 5' nuclease assay. These probes are allele-specific with one matching the wildtype sequence and the other the mutant sequence. Each probe is distinguished by being labeled with different fluorescent reporter dyes (usually FAM and VIC) at the 5' end and a quencher dye at the 3' end (TAMRA). A mismatch between the probe and target sequence largely reduces the efficiency of probe hybridisation and cleavage. The probes therefore only emit fluorescence in the presence of their respective complementary target sequence. An increase in fluorescence from only one probe indicates homozygosity for either the wildtype or mutant allele, while an increase in fluorescence from both probes indicates heterozygosity [Livak et al., 1995a, Livak et al., 1999] (see Figure 11).

Primers and probes are designed using the Primer Express program (Applied Biosystems, Foster City, California). The guidelines for assay design for both primers and probes include a GC content ranging between 30 to 80% and no runs of more than three consecutive G's. The melting temperature (Tm) for the primers should be 58 to 60^{0}C and the Tm of the probes should be 7^{0}C higher than that of the primer. The probe lengths should be adjusted to have the same Tm. The five nucleotides at the 3' end of the primer should have only one to two G's and C's. The strand that gives the probe more C's than G's should be selected and there must not be a G on the 5' end of the probe. The primers should be placed as close as possible to the probes, but not overlapping. Amplicons ranging between 75 and 150bp provide consistent results as larger amplicons might require extensive optimisation. The limitations set for amplicon

design allows for all reactions to use a single buffer referred to as the TaqMan Universal PCR Master Mix. This PCR master mix contains all the necessary reaction components, except the primers and probes, to perform the fluorogenic 5' nuclease assay. The AmpliTaq Gold enzyme is a thermal stable DNA polymerase and also forms part of the PCR master mix. The enzyme is active only after incubation at elevated temperatures and the use of this enzyme introduces an invisible Hot Start to any amplification reaction, which reduces primer dimer formation and further improves amplification of specific target sequences [Livak et al., 1995b; Livak et al., 1999].
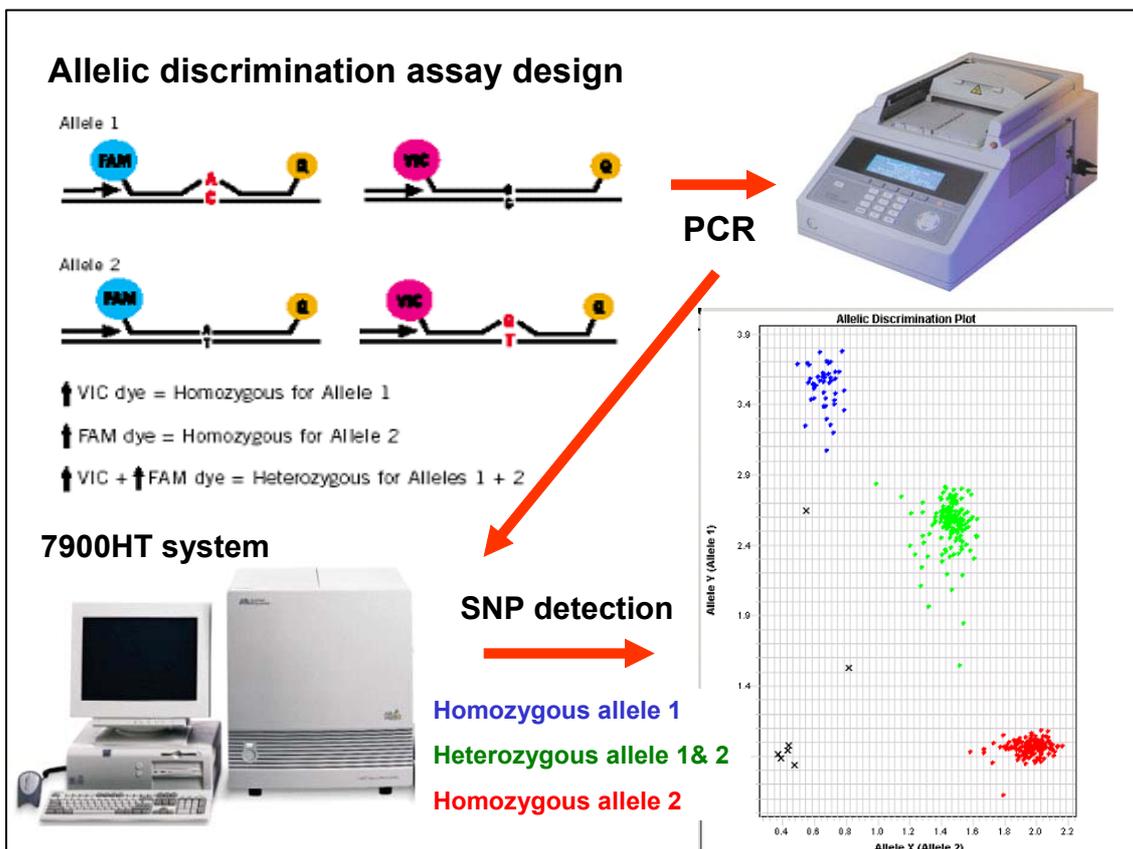
Figure 11. The allelic discrimination methodology is depicted. The method requires two allele-specific fluorogenic probes in a 5' nuclease assay. Genotyping involves a PCR step followed by end-point fluorescence measured on a sequence detection system. The results are obtained using the data analysis software. The allelic discrimination viewer shows three different clusters representing the three different genotypes that can be observed, while the crosses indicate either blank controls or samples, which did not amplify. (Partially adapted from Applied Biosystems: www.appliedbiosystems.com)

Once the allelic discrimination assay is designed, genotyping is performed using a PCR step followed by end-point fluorescence being measured on a sequence detection system such as the 7900 high-throughput system (Applied Biosystems, Foster City, California) offering 384 well plate compatibility and robotic loading. The sequence detection software automatically processes fluorescence data and provides the genotype calls in an allelic discrimination viewer. Each sample is represented by a single point and the various cluster colours indicate the different genotypes. The presence of crosses indicates either blanc template controls or samples that were not amplified [Livak et al., 1999; Ranade et al., 2001; Sevall et al., 2001] (see Figure 11).

Although the fluorogenic probes are costly, the cost per sample is low when large sample numbers are genotyped. The allelic discrimination method therefore allows for higher throughout compared to the pre-screening mutation detection methods for the same time period and is still more cost-effective than automated direct sequencing. The presence of closely lying SNPs may have an influence on the genotyping of a specific SNP. This is clearly visible with the distortion of the three clusters representing the different genotypes or the presence of additional sub-clusters where one cluster appears to be split in two. This finding is seldom observed with the accuracy of this genotyping method generally being estimated at an error rate of less than 1 in 2000 genotypes [Ranade et al., 2001]. The allelic discrimination method therefore provides reliable and reproducible results that are rapidly generated and was the chosen method for large-scale SNP genotyping in the studies presented in Chapters 2 and 3 of this dissertation.

1.4. References

- Abdelwahab SF, Cocchi F, Bagley KC, et al. HIV-1 supressive factors are secreted by CD4$^+$ T cells during primary immune responses. *Proc Natl Acad Sci USA* 2003; 100:15006-15010.

- Abraham LJ, French MAH and Dawkins RL. Polymorphic MHC ancestral haplotypes affect the activity of tumour necrosis factor-alpha. *Clin Exp Immunol* 1993; 92:14-18.

- Abrams ES, Murdaugh SE and Lerman LS. Comprehensive detection of single base changes in human genomic DNA using denaturing gradient gel electrophoresis and a GC-clamp. *Genomics* 1990; 7: 463-475.

- Alkhatib G, Combadiere C, Broder CC, et al. CC CKR5: a RANTES, MIP-1α, MIP-1β receptor as a fusin cofactor for the macrophage-tropic HIV-1. *Science* 1996; 272:1955-1958.

- Alkhatib G, Ahuja SS, Light D, Mummidi S, Berger EA and Ahuja SK. CC chemokine receptor 5-mediated signalling and HIV-1 co-receptor activity share common structural determinants. Critical residues in the third extracelllular loop support HIV-1 fusion. *J Biol Chem* 1997; 271:19771-19776.

- Alvarez V, Lopez-Larrea C and Coto E. Mutational analysis of the CCR5 and CXCR4 genes (HIV-1 co-receptors) in resistance to HIV-1 infection and AIDS development among intravenous drug users. *Hum Genet* 1998; 102:483-486.

- Amara A, Gall SL, Schwartz O, et al. HIV coreceptor downregulation as antiviral principle: SDF-1 alpha-dependant internalisation of the chemokine receptor CXCR4 contributes to inhibition of HIV replication. *J Exp Med* 1997; 186:139-146.

- An P, Martin MP, Nelson GW, et al. Influence of CCR5 promoter haplotypes on AIDS progression in African-Americans. *AIDS* 2000; 14:2117-2122.

- An P, Nelson GW, Wang L, et al. Modulating influence on HIV/AIDS by interacting *RANTES* gene variants. *Proc Natl Acad Sci USA* 2002; 99:10002-10007.

- Anastassopoulou CG and Kostrikis LG. The impact of human allelic variation on HIV-1 disease. *Curr HIV Res* 2003; 1:185-203.

- Ansari-Lari MA, Liu X, Metzker ML, Rut AR and Gibbs RA. The extent of genetic variation in the *CCR5* gene. *Nat Genet* 1997; 16:221-222.

- Anzala AO, Ball TB, Rostron T, O'Brien SJ, Plummer FA and Rowland-Jones SL. CCR2-64I allele and genotype association with delayed AIDS progression on African women. *Lancet* 1998; 351:1632-1633.

- Applied Biosystems, Foster City, California, (www.appliedbiosystems.com)

- Arai N, Nomura D, Villaret D, et al. Complete nucleotide sequence of chromosomal gene for human IL-4 and its expression. *J Immun* 1989; 142:274-282.

- Arenzana-Seisdedos F, Virelizier JL, Rousset D, et al. HIV blocked by chemokine antagonist. *Nature* 1996; 383:400.

- Atchison RE, RE, Gosling J, Monteclaro FS et al. Multiple extracellular elements of CCR5 and HIV-1 entry: dissociation from response to chemokines. *Science* 1996; 274:1924-1926.

- Aukrust P, Muller F and Froland SS. Circulating levels of RANTES in human immunodeficiency virus type 1 infection: effect of potent antiretroviral therapy. *J Infect Dis* 1998; 177:1091-1096.

- Autran B, Legac E, Blanc C and Debré P. A Th0/Th2-like function of CD4+CD7- T helper cells from normal donors and HIV-infected patients. *J Immunol* 1995; 154:1408-17.

- Baena A, Leung JY, Sullivan AD, et al. TNF-alpha promoter single nucleotide polymorphisms are markers of human ancestry. *Genes Immun* 2002; 3:482-487.

- Balotta C, Bagnarelli P, Violin M, et al. Homozygous Δ32 deletion of the *CCR-5* chemokine receptor gene an HIV-1 infected patient. *AIDS* 1997; 11:F67-F71.

- Balotta C, Bagnarelli P, Corvasce S et al. Identification of two subsets of long-term nonprogressors with divergent viral activity by stromal-derived factor 1 chemokine gene polymorphism analysis. *J Infect Dis* 1999; 180:285-289.

- Bamshad M, Wooding S, Salisbury BA and Stephens JC. Deconstructing the relationship between genetics and race. *Nat Rev Genetics* 2004; 5:598-609.

- Barker E, Mackewicz CE and Levy JA. Effects of T_H1 and T_H2 cytokines on $CD8^+$ cell response against human immunodeficiency virus: Implications for long-term survival. *Proc Natl Acad Sci USA* 1995; 92:11135-11139.

- Bender R and Lange S. Adjusting for multiple testing – when and how? *J Clin Epidemiol* 2001; 54:343-349.

- Bendtzen K, Morling N, Fomsgaard A, et al. Association between HLA-DR2 and production of tumour necrosis factor alpha and interleukin 1 by mononuclear cells activated by lipopolysaccharide. *Scand J Immunol* 1998; 28:599-606.

- Berger EA, Doms RW, Fenyo EM, et al. A new classification for HIV-1. *Nature* 1998; 391:240.

- Bessong PO, Obi CL, Cilliers T et al. Characterization of human immunodeficiency virus type 1 from a previously unexplored region of South Africa with a high HIV prevalence. *AIDS Res Hum Retroviruses* 2005; 21:103-109.

- Bieniasz PD, Fridell RA, Aramori L, Ferguson SS, Caron M and Cullen BR. HIV-1-induced cell fusion is mediated by multiple regions within both the viral envelope and the CCR-5 co-receptor. *EMBO Journal* 1997; 16:2599-2609.

- Biti R, French R, Young J, Bennets B, Stewart G and Liang T. HIV-1 infection in an individual homozygous for the CCR5 deletion allele. *Nat Med* 1997; 3:252-253.

- Blackwell JM, Barton CH, White JK, et al. Genomic organization and sequence of the human NRAMP gene: identification and mapping of a promoter region polymorphism. *Molec Med* 1995; 1:194-205.

- Blackwell JM, Searle S, Mohamed H and White JK. Divalent cation transport and susceptibility to infectious and autoimmune disease: continuation of the Ity/Lsh/Bcg/Nramp1/Slc11a1 gene story. *Immunol Lett* 2003; 85:197-203.

- Blanpain C, Lee B, Tackoen M, et al. Multiple non-functional alleles of CCR5 are frequent in various human populations. *Blood* 2000, 96:1638-1645.

- Bleul CC, Farzan M, Choe H, et al. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusion and blocks HIV-1 entry. *Nature* 1996a; 382:829-833.

- Bleul CC, Fuhlbrigge RC, Casasnovas JM, Aiuti A and Springer TA. A highly efficacious lymphocyte chemoattractant, stromal cell-derived factor 1 (SDF-1) *J Exp Med* 1996b; 184:799-802.

- Børresen A-L, Hovig E, Smith-Sørensen B, et al. Constant denaturing gradient gel electrophoresis as a rapid screening technique for p53 mutations. *Proc Natl Acad Sci USA* 1991; 88:8405-8409.

- Bouwmeester T, Bauch A, Ruffner H, et al. A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat Cell Biol* 2004; 6:97-105.

- Brambilla A, Villa C, Rizzardi G et al. Shorter survival of *SDF1-3' A/3 'A* homozygotes linked to CD4+ T cell decrease in advanced human immunodeficiency virus type 1 infection. *J Infect Dis* 2000; 182: 311-315.

- Bream JH, Young HA, Rice N, et al. *CCR5* promoter alleles and specific DNA binding factors. *Science* 1999; 284:223a.

- Bredell H, Williamson C, Sonnenberg P, Martin DJ and Morris L. Genetic characterisation of HIV type 1 from migrant workers in three South African gold mines. *AIDS Res Hum Retroviruses* 1998; 14:677-684.

- Bredell H, Hunt G, Morgan B, Tiemessen CT, Martin DJ and Morris L. Identification of HIV type 1 intersubtype recombinants in South Africa using *env* and *gag* heteroduplex mobility assays. *AIDS Res Hum Retroviruses* 2000; 16:493-497.

- Brelot A, Heveker N, Montes M and Alizon M. Identification of residues of CXCR4 critical for human immunodeficiency virus co-receptor and chemokine receptor activities. *J Biol Chem* 2000; 275:23736-23744.

- Brinkman BMN, Keet IPM, Miedema F, Verweij CL and Klein MR. Polymorphisms within the human tumor necrosis factor-α promoter region in human immunodeficiency virus type 1-seropositive persons. *J Infect Dis* 1997; 175:188-190.

- Cardon LR and Palmer LJ. Population stratification and spurious allele association. *Lancet* 2003; 361:598-604.

- Carrington M, Kissner T, Gerrard B, Ivanov S, O'Brien SJ and Dean M. Novel alleles of the chemokine-receptor gene *CCR5*. *Am J Hum Genet* 1997; 61:1261-1267.

- Carrington M, Dean M, Martin MP and O'Brien SJ. Genetics of HIV-1 infection: chemokine receptor CCR5 polymorphism and it's consequences. *Hum Molec Genet* 1999; 8:1939-1945.

- Carrington M, Nelson G and O'Brien SJ. Considering genetic profiles in functional studies of immune responsiveness to HIV-1. *Immunol Lett* 2001; 79:131-40.

- Carrington M and O'Brien S. The influence of *HLA* genotype on AIDS. *Annu Rev Med* 2003; 54:535-551.

- Caruz A, Samson M, Alonso JM, et al. Genomic organization and promoter characterization of human CXCR4 gene. *FEBS Lett* 1998; 426: 271-278.

- Cellier M, Govoni G, Vidal S, et al. Human natural resistance-associated macrophage protein: cDNA cloning, chromosomal mapping, genomic organization, and tissue-specific expression. *J Exp Med* 1994; 180:1741-1752.

- Charo IF, Myers SJ, Herman A, Franci C, Connolly AJ and Coughlin SR. Molecular cloning and functional expression of two monocyte chemoattractant protein 1 receptors reveals alternative spilcing of the carboxyl-terminal tails. *Proc Natl Acad Sci USA* 1994; 91:2752-2756.

- Choe H, Farzan M, Sun Y, et al. The β-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell* 1996; 85:1135-1148.

- Clark AG. The role of haplotypes in candidate gene studies. *Genet Epidemiol* 2004; 27:321-333.

- Clark VJ and Dean M. Haplotype structure and linkage disequilibrium in chemokine and chemokine receptor genes. *Hum Genomics* 2004; 1:255-273.

- Clerici M and Shearer GM. A TH1→TH2 switch is a critical step in the etiology of HIV infection. *Immunol Today* 1993; 14:107-111.

- Clerici M and Shearer GM. The Th1-Th2 hypothesis of HIV infection: new insights. *Immunol Today* 1994; 15:575-581.

- Clerici M, Fusi ML, Ruzzante S, et al. Type 1 and type 2 cytokines in HIV infection – a possible role in apoptosis and disease progression. *Ann Med* 1997; 29:185-188.

- Cocchi F, De Vico A, Garzino-Demo A, Arya S, Gallo R and Lusso P. Identification of RANTES, MIP-1α, and MIP-β as the major HIV suppressive factors produced by CD8+ T cells. *Science* 1995; 270:1811-1815.

- Cohen OJ, Paolucci S, Bende SM et al. CXCR4 and CCR5 genetic polymorphsms in long-term nonprogressive human immunodeficiency virus infection: Lack of association with mutations other than CCR5-Δ32. *J Virol* 1998; 72:6215-6217.

- Combadiere C, Ahuja S, Tiffany H and Murphy P. Cloning and functional expression of CC CKR5, a human monocyte CC chemokine receptor selective for MIP-1 (alpha), MIP-1 (beta), and RANTES. *J Leukoc Biol* 1996; 60:147-152.

- Combadiere C, Salzwedel K, Smith ED, Tiffany HL, Berger EA and Murphy PM. Identification of CX3CR1. A chemotactic receptor for the human CX3C chemokine fractalkine and a fusion coreceptor for HIV-1. *J Biol Chem* 1998; 273:23799-23804.

- Connor RI, Mohri H, Cao Y and Ho DD. Increased viral burden and cytopathicity correlate temporally with CD4+ T-lymphocyte decline and clinical progression in human immunodeficiency virus type 1-infected individuals. *J Virol* 1993; 67:1772-1777.

- Conway DJ, Holland MJ, and Bailey RL. Scarring trachoma is associated with polymorphism in the tumor necrosis factor alpha (TNF-α) gene promoter and with elevated TNF-α levels in tear fluid. *Infec Immun* 1997, 65, 1003.

- Corfield VA and Brink PA. Case:control association studies of complex disease: is the search for genetic determinants a fishing expedition? *Cardiovasc J S Afr* 2002; 13:93-95.

- Crawford DC and Nickerson DA.   Definition and clinical importance of haplotypes. *Ann Rev Med* 2005; 56:303-320.

- Crawley E, Kay R, Sillibourne J, Patel P, Hutchinson I and Woo P. Polymorphic haplotypes of the interleukin-10 5' flanking region determine variable interleukin-10 transcription and are associated with particular phenotypes of juvenile rheumatoid arthritis. *Arthritis Rheum* 1999; 42:1101-1108.

- D'Alfonso SD, and Richiardi PM.   A polymorphic variation in a putative regulation box of the TNFα promoter region. *Immunogenetics* 1994; 39:150-154.

- Dalgeish AG, Beverley PCL, Clapham PR, Crawford DH, Greaves MF and Weiss RA. The CD4 (T4) antigen is an essential of the receptor of the AIDS retrovirus. *Nature* 1984; 312:763-767.

- Daugherty BL and Springer MS. The β-chemokine receptor genes CCR1 (CMKBR1), CCR2 (CMKBR2), and CCR3 (CMBRK3) cluster within 285kb on human chromosome 3p21. *Genomics* 1997; 41: 294-297.

- Dean M, Carrington M, Winkler C, et al.   Genetic restriction of HIV-1 infection by a deletion allele of the *CKR5* structural gene. *Science* 1996; 273:1856-1862.

- Dean M, Carrington M and O'Brien SJ. Balanced polymorphism selected by genetic versus infectious human disease. *Annu Rev Genomics Hum Genet* 2002; 3:263-292.

- Deng H, Liu R, Ellmeier W, et al. Identification of a major co-receptor for primary isolates of HIV-1. *Nature* 1996; 381:661-666.

- Department of Health. National HIV and Syphilis antenatal sero-prevalence survey in South Africa 2004. Pretoria, South Africa, 2005.

- Derdeyn CA, Costello C, Kilby JM et al. Correlation between circulating stromal cell-derived factor 1 levels and CD4+ cell count in human immunodeficiency virus type 1-infected individuals. *AIDS Res Hum Retroviruses* 1999; 15:1063-1071.

- Dong VM, McDermott DH and Abdi R. Chemokines and disease. *Eur J Dermatol* 2003; 13:224-230.

- Donlon TA, Krensky AM, Wallace MR, Collins FS, Lovett M and Clayberger C. Localization of a human T-cell-specific gene, RANTES (D17S136E), to chromosome 17q11.2-q12. *Genomics* 1990; 5:548-553.

- Donninger H, Cashmore TJ, Scriba T, Petersen DC, Janse van Rensburg E and Hayes VM. Functional analysis of novel *SLC11A1 (NRAMP1)* promoter variants in susceptibility to HIV-1. *J Med Genet* 2004; 41:e49.

- Doranz BJ, Rucker J, Yi Y, et al. A dual-tropic primary HIV-1 isolate that uses fusin and the β-chemokine receptors CKR-5, CKR-3, and CKR-2 as fusin cofactors. *Cell* 1996; 85:1149-1158.

- Doranz BJ, Lu ZH, Rucker J, et al. Two distinct CCR5 domains can mediate coreceptor usage by human immunodeficiency virus type 1. *J Virol* 1997; 71:6305-6314.

- Dragic T, Litwin V, Allaway GP, et al. HIV-1 entry into CD4$^+$ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* 1996; 381:667-673.

- D'Souza I, Poorkaj P, Hong M, et al. Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements. *Proc Natl Acad Sci USA* 1999; 96:5598-5603.

- Duggal P, Winkler CA, An P, et al. The effect of RANTES chemokine genetic variants on early HIV-1 plasma RNA among African American injection drug users *J Acquir Immune Defic Syndr* 2005; 38:584-589.

- Duh EJ, Maury WJ, Folks TM, Fauci AS, and Rabson AB. Tumor necrosis factor alpha activates human immunodeficiency virus type 1 through induction of nuclear factor binding to the NF-kappa B sites in the long terminal repeat. *Proc Natl Acad Sci USA* 1989; 86:5974-5978.

- Dunkle KL, Jewkes RK, Brown HC, Gray GE, McIntryre JA and Harlow SD. Transactional sex among women in Soweto, South Africa: prevalence, risk factors and association with HIV infection. *Soc Sci Med* 2004; 59:1581-1592.

- Dupont WD and Plummer WD Jr. Power and sample size calculations. A review and computer program. *Control Clin Trials* 1990; 11:116-128.

- Edelman L, Deveau C, Raphael M, et al. Serum interleukin-10 in acquired immunodeficiency syndrome lymphoma patients. Seroco-Hemoco Study Group. *Eur Cytokine Netw* 1996; 7:785-91.

- Edwards J, Ruparel H and Ju J. Mass-spectrometry DNA sequencing. *Mutat Res* 2005; 573:3-12.

- Engelbrecht S, Laten JD, Smith TL and van Rensburg EJ. Identification of env subtypes in fourteen HIV type 1 isolates from South Africa. AIDS *Res Hum Retroviruses* 1995; 11:1296-1271.

- Engelbrecht S, Koulinska I, Smith TL, Bareto J and van Rensburg EJ. Variation in the HIV type V3 region env sequences from Mozambique. *AIDS Res Hum Retroviruses* 1998; 14:803-805.

- Engelbrecht S, Smith TL, Kasper P, et al. HIV type 1 V3 domain serotyping and genotyping in Gauteng, Mpumalanga, KwaZulu-Natal, and Western Cape provinces of South Africa. *AIDS Res Hum Retroviruses* 1999; 15:325-328.

- Eskdale J, Kube D, Tesch H and Gallagher G. Mapping of the human IL10 gene and further characterisation of the 5' flanking sequence. *Immunogenetics* 1997; 46:120-128.

- Eugen-Olsen J, Iversen AK, Benfield TL, Koppelhus U and Garred P. Chemokine receptor CCR2b 64I polymorphism and it's relation to CD4 cell counts and disease progression in a Danish cohort of HIV-infected individuals. Copenhagen AIDS cohort. *J Acquir Immun Defic Syndr Hum Retrovirol* 1998; 18:110-116.

- Ezekowitz RA, Day LE and Herman GA. A human mannose-binding protein is an acute phase reactant that shares sequence homology with other vertebrate lectins. *J Exp Med* 1988; 167:1034-1046.

- Ezekowitz RAB, Kuhlman M, Groopman JE and Bryn RA. A human serum mannose-binding protein inhibits in vitro infection by human immunodeficiency virus. *J Exp Med* 1989; 169:185-196.

- Fallin D, Cohen A, Essioux L, et al. Genetic analysis of case/control data using estimated haplotype frequencies: Application to APOE locus variation and Alzheimer's disease. *Genome Res* 2001; 11:143-151.

- Falush D, Stephens M and Pritchard JK. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 2003; 164:1567-1587.

- Faure S, Meyer L, Costagliola D, et al. Rapid progression to AIDS in HIV+ individuals with a structural variant of the chemokine receptor CX_3CR1. *Science* 2000; 287:2274-2277.

- Faure S, Meyer L, Genin E, et al. Deleterious genetic influence of CX3CR1 genotypes on HIV-1 disease progression. *J Acquir Immune Defic Syndr* 2003; 32:335-337.

- Federsppiel B, Melhado IG, Duncan AMV, et al. Molecular cloning of the cDNA and chromosomal localization of the gene for a putative seven-transmembrane segment (7-TMS) receptor isolated from human spleen. *Genomics* 1993; 16:707-712.

- Feng Y, Broder CC, Kennedy PE and Berger E. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 1996; 272:827-877.

- Fischer SG and Lerman LS. DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: correspondence with melting theory. *Proc Natl Acad Sci USA* 1983; 80:1579-1583.

- Folks TM, Clouse KA, Justement J, et al. Tumor necrosis factor alpha induces expression of human immunodeficiency virus in chronically infected T-cell clone. *Proc Natl Acad Sci USA* 1989; 86:2365-2368.

- Forbes JR and Gros P. Divalent-metal transport by NRAMP proteins at the interface of host-pathogen interactions. *Trends Microbiol* 2001; 9:397-403.

- Frade JMR, Llorente M, Mellado M, et al. The amino-terminal domain of the CCR2 chemokine receptor acts as a coreceptor for HIV-1 infection. *J Clin Invest* 1997; 100:497-502.

- Freedman ML, Reich D, Penney KL, et al. Assessing the impact of population stratification on genetic association studies. *Nat Genet* 2004; 36:388-393.

- Gabriel SB, Schaffner SF, Nguyen H et al. The structure of haplotype blocks in the human genome. *Science* 2002; 296:2225-2229.

- Galli G, Annunziato F, Cosmi L, Manetti R, Maggi E and Romagnani S. Th1 and th2 responses, HIV-1 coreceptors, and HIV-1 infection. *J Biol Regul Homeost Agents* 2001; 15:308-313.

- Garin A, Pellet P, Deterre P, Debré P and Combadiere C. Cloning and functional characterization of human fractalkine receptor promoter regions. *Biochem J* 2002; 368:753-760.

- Garin A, Tarantino N, Faure S, et al. Two novel fully functional isoforms of CX3CR1 are potent HIV coreceptors. *J Immunol* 2003; 171:5305-5312.

- Garred P, Thiel S, Madsen HO, Ryder LP, Jensenius JC and Svegaard A. Gene frequency and partial protein characterization of an allelic variant of mannan binding protein associated with low serum concentrations. *Clin Exp Immunol* 1992a; 90:517-521.

- Garred P, Madsen HO, Kurtzhals JAL, et al. Diallelic polymorphism may explain variations of blood concentration of mannan binding protein in Eskimos, but not in Black Africans. *Eur Journal Immunogenet* 1992b; 19:403-412.

- Garred P, Madsen HO, Balslev U, et al. Susceptibility to HIV infection and progression of AIDS in relation to variant alleles of mannose binding lectin. *Lancet* 1997; 349:236-240.

- Garzino-Demo A, Moss RB, Margolick JB, et al. Spontaneous and antigen-induced production of HIV-inhibitory beta-chemokines are associated with AIDS-free status. *Proc Natl Acad Sci USA* 1999; 96:11986-11991.

- Gauderman WJ, Witte JS and Thomas DC. Family-based association studies. *Monogr Natl Cancer Inst* 1999; 26:31-37.

- Gibbs RA, Belmont JW, Hardenbol P, et al. The International HapMap project. *Nature* 2003; 426:789-796.

- Gonzalez E, Bamshad M, Sato N, et al. Race-specific HIV-1 disease modifying effects associated with CCR5 haplotypes. *Proc Natl Acad Sci USA* 1999; 96:12004-12009.

- Gonzalez E, Dhanda R, Bamshad M, et al. Global survey of genetic variation in CCR5, RANTES, and MIP-1α: Impact on the epidemiology of the HIV-1 pandemic. *Proc Natl Acad Sci USA* 2001; 98:5199-5204.

- Gordon CJ, Muesing MA, Proudfoot AE, Power CA, Moore JP and Trkola A. Enhancement of human immunodeficiency virus type 1 infection by the CC-chemokine Rantes is independent of the mechanism of virus-cell fusion. *J Virol* 1999; 73:684-694.

- Gordon M, De Oliveira T, Bishop K et al. Molecular characteristics of human immunodeficiency immunodeficiency virus type 1 subtype C viruses from KwaZulu-Natal, South Africa: Implications for vaccine and antiretroviral control strategies. *J Virol* 2003; 77:2587-2599.

- GraphPad Software Inc, San Diego, California, (www.graphpad.com).

- Gray IC, Campbell DA and Spurr NK. Single nucleotide polymorphisms as tools in human genetics. *Hum Molec Genet* 2000; 9:2403-2408.

- Graziosi C, Pantaleo G, Gantt KR, et al. Lack of evidence for the dichotomy of TH1 and TH2 predominance in HIV-infected individuals. *Science* 1994; 265:248-252.

- Gross E, Amelia CA, Pompucci L, Franchin G, Sherry B and Schmidtmayerova H. Macrophages and lymphocytes differentially modulate the ability of RANTES to inhibit HIV-1 infection. *J Leukoc Biol* 2003; 74:781-790.

- Guldberg P and Guttler F. A simple method for identification of point mutations using denaturing gradient gel electrophoresis. *Nucleic Acids Res* 1993; 21:2261-2262.

- Hajeer AH, and Hutchinson IV. Influence of TNFα gene polymorphisms on TNFα production and disease. *Hum Immunol* 2001; 62:1191-1199.

- Hamann A, Mantzoros C, Vidal-Puig A, and Flier JS. Genetic variability in the TNF-α promoter is not associated with type II diabetes mellitus (NIDDM). *Biochem Biophys Res Commun* 1995; 211:833-839.

- Haurum JS, Thiel S, Jones IM, Fischer PB, Laursen SB and Jensenius JC. Complement activation upon binding of mannan-binding protein to HIV envelope glycoproteins. *AIDS* 1993; 7:1307-1313.

- Hayes VM. Mutation screening of the TP53 and KRAS genes in tumors of the colon, lung, brain and cervix. *Thesis* 1999.

- Hayes VM, Wu Y, Osinga J, et al. Improvements in gel composition and electrophoretic conditions for broad-range mutation analysis by denaturing gradient gel electrophoresis. *Nucleic Acids Res* 1999; 27:e29.

- Hayes VM, Petersen DC, Scriba TJ, Zeier M, Grimwood A and Janse van Rensburg E. Analysis of African-based *CCR5*-single nucleotide polymorphisms and association with HIV-1 disease progression. *AIDS* 2002, 16:2229 -2238.

- Hayes VM. Genetic diversity of the alpha-1-antitrypsin gene in Africans identified using a novel genotyping assay. *Hum Mut* 2003; 22:59-66.

- Hayes VM and Gardiner-Garden M. Are polymorphic markers in the α-1-antitrypsin gene associated with risk of human immunodefficiency virus disease? *J Infect Dis* 2003; 188:1205-1208.

- Hendel H, Henon N, Lebuanec H, et al. Distinctive effects of *CCR5, CCR2*, and *SDF1* genetic polymorphisms in AIDS progression. *J Acquir Immune Defic Syndr Hum Retrovirol* 1998; 19:381-386.

- Hendel H, Winkler C, An P et al. Validation of genetic case-control studies in AIDS and application to the CX3CR1 polymorphism. *J Acquir Immune Defic Syndr* 2001; 26:507-11.

- Herzog H, Hort YJ, Shine J, et al. Molecular cloning, characterization, and localization of the human homolog to the reported bovine NPY Y3 receptor: lack of NPY binding and activation. *DNA Cell Biol* 1993; 12: 465-471.

- Hill WG and Robertson A. The effects of inbreeding at loci with heterozygote advantage. *Genetics* 1968; 60:615-628.

- Hinds DA, Stokowski RP, Patil N, et al. Matching strategies for genetic association studies in structured populations. *Am J Hum Genet* 2004; 74:317-325.

- Hirschhorn JN, Lohmueller K, Byrne E and Hirschhorn K. A comprehensive review of genetic association studies. *Genet Med* 2002; 4:45-61.

- Hirschhorn JN and Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005; 6:95-108.

- Hoffmann S, Stanley EM, Cox ED, et al. Association of cytokine polymorphic inheritance and in vitro cytokine production in anti-CD3/CD28-stimulated peripheral blood lymphocytes. *Transplantation* 2001; 72:1444-1450.

- Hogan CM and Hammer SM. Host determinants in HIV infection and disease, Part 2: Genetic factors and implications for antiretroviral therapeutics. *Ann Intern Med* 2001; 134:978-996.

- Hoggart CJ, Parra EJ, Shriver MD, et al. Control of confounding of genetic associations in stratified populations. *Am J Hum Genet* 2003; 72:1492-1504.

- Holland PM, Abramson RD, Watson R and Gelfand DH. Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proc Natl Acad Sci USA* 1991; 88:7276-7280.

- Holmskov U, Malhotra R, Sim RB and Jensenius JC. Collectins: collagenous C-type lectins of the innate immune defense system. *Immunol Today* 1994; 15:67-73.

- Houle M, Thivierge M, Le Gouill C, Stanková J and Rola-Pleszczynski M. IL-10 up-regulates CCR5 gene expression in human monocytes. *Inflammation* 1999; 23:241-251.

- Howard OMZ, Shirakawa A-K, Turpin JA, et al. Naturally occurring CCR5 extracellular and transmembrane domain variants affect HIV-1 co-receptor and ligand binding function. *J Biol Chem* 1999; 274:16228-16234.

- Huang Y, Paxton WA, Wolinsky SM, et al. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. *Nat Med* 1996; 2:1240-1243.

- Huber C, Pons O, Hendel H, et al. Genomic studies in AIDS: problems and answers. Development of a statistical model integrating both longitudinal cohort studies and transversal observations of extreme cases. *Biomed Pharmacother* 2003; 57:25-33.

- Huttley G, Smith MW, Carrington M and O'Brien. A scan for linkage disequilibrium across the human genome. *Genetics* 1999; 152:1711-1722.

- Ikegawa M, Yuan J, Matsumoto K, et al. Elevated plasma stromal cell-derived factor 1 protein level in the progression of HIV type 1 infection/AIDS. *AIDS Res Hum Retroviruses* 2001; 17:587-595.

- Ioannidis JP, O'Brien TR, Rosenberg PS, Contopoulos-Ioannidis DG and Goedert JJ. Genetic effects on HIV disease progression. *Nat Med* 1998; 4:536.

- Ioannidis JP, Rosenberg PS, Goedert JJ, et al. Effects of CCR5-Delta32, CCR2-64I and the SDF-1 3'A alleles on HIV-1 disease progression: An international meta-analysis of individual-patient data. *Ann Intern Med* 2001a; 135: 782-795.

- Ioannidis JPA, Ntzani EE, Trikalinos TA and Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet* 2001b; 29:306-309.

- Ioannidis JPA. Genetic associations: false or true? *Trends Molec Med* 2003a; 9:135-138.

- Ioannidis JPA, Trikalinos TA, Ntzani EE and Contopoulos-Ioannidis DG. Genetic associations in large versus small studies: an empirical assessment. *Lancet* 2003b; 361:567-571.

- Ioannidis JP, Ntzani EE and Trikalinos TA. 'Racial' differences in genetic effects for complex diseases. *Nat Genet* 2004; 36:1312-1318.

- Jacob CO, Fronek Z, Lewis GD, Koo M, Hansen JA, and McDevitt HO. Heritable major histocompatibility complex class II-associated differences in production of tumor necrosis factor alpha: relevance to genetic predisposition to systemic lupus erythematosus. *Proc Natl Acad Sci USA* 1990; 87:1233-1237.

- Jansson M, Backström E, Björedal A, et al. Coreceptor usage and RANTES sensitivity of non-syncytium-inducing HIV-1 isolates obtained from patients with AIDS. *J Hum Virol* 1999; 2:325-338.

- Jinquan T, Quan S, Jacobi HH, et al. CXC chemokine receptor 4 expression and stromal cell-derived factor 1α-induced chemotaxis in CD4$^+$ T lymphocytes are regulated by interleukin-4 and interleukin-10. *Immunology* 2000; 99:402-410.

- John GC, Rousseau C, Dong T, et al. Maternal SDF1 3' A polymorphism is associated with increased perinatal human immunodeficiency virus type 1 transmission. *J Virol* 2000; 74:5736-5739.

- Jorde LB. Linkage disequilibrium and the search for complex disease genes. *Genome Res* 2000; 10:1435-1444.

- Kalinkovich A, Weisman Z and Bentwich Z. Chemokine and chemokine receptors: role in HIV infection. *Immunol Lett* 1999; 68:281-287.

- Kaslow RA, Dorak T and Tang J. Influence of host genetic variation on susceptibility to HIV type 1 infection. *J Infect Dis* 2005; 191: S68-S77.

- Kawasaki N, Kawasaki T and Yamashina I. Isolation and characterization of a mannan-binding protein from human serum. *J Biochem* 1983; 94:937-945.

- Kedzierska K, Crowe SM, Turville S and Cunningham AL. The influence of cytokines, chemokines and their receptors on HIV-1 replication in monocytes and macrophages. *Rev Med Virol* 2003; 13:39-56.

- Khanna KV, Yu XF, Ford DH, Ratner L, Hildreth JK and Markham RB. Differences among HIV-1 variants in their ability to elicit secretion of TNF-alpha. *J Immunol* 2000; 164:1408-1415.

- Kidd P. Th1/Th2 balance: The hypothesis, it's limitations, and implications for health and disease. *Altern Med Rev* 2003; 8:223-246.

- Kittles RA and Weiss KM. Race, ancestry, and genes: Implications for defining disease risk. *Ann Rev Genomics Hum Genet* 2003; 4:33-67.

- Kim JM, Brannan CI, Copeland NG, Jenkins NA, Khan TA and Moore KW. Structure of the mouse IL10-gene and chromosomal localization of the mouse and human genes. *J Immunol* 1992; 148:3618-3623.

- Kimura R, Nishioka T, Soemantri A and Ishida T. Allele-specific transcript quantification detects haplotypic variation in the levels of the SDF-1 transcripts. *Hum Molec Genet* 2005; 14:1579-1585.

- Klatzman D, Champagne E, Chamaret S, et al. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* 1984; 312:767-768.

- Klein SA, Dobmeyer JM, Dobmeyer TS, et al. Demonstration of the Th1 to Th2 cytokine shift during the course of HIV-1 infection using cytoplasmic

cytokine detection on single cell level by flow cytometry. *AIDS* 1997; 11:1111-1118.

- Klein NJ. Mannose-binding lectin: do we need it? *Mol Immunol* 2005; 42:919-924.

- Knight JC. Regulatory polymorphisms underlying complex disease traits. *J Mol Med* 2003; 83:97-109.

- Knuchel MC, Spira TJ, Neumann AU, et al. Analysis of a biallelic polymorphism in the tumor necrosis factor α promoter and HIV type 1 disease progression. *AIDS Res Hum Retroviruses* 1998; 14:305-309.

- Kollmann TR, Pettoello-Mantovani M, Katopodis NF, et al. Inhibition of acute in vivo human immunodeficiency virus infection by human interleukin 10 treatment of SCID mice implanted with human fetal thymus and liver. *Proc Natl Acad USA Sci* 1996; 93:3126-3131.

- Kostrikis LG, Huang Y, Moore JP, et al. A chemokine receptor *CCR2* allele delays HIV-1 disease progression and is associated with a *CCR5* promoter mutation. *Nat Med* 1998; 4: 350-355.

- Kuhlman M, Joiner K and Ezekowitz RAB. The human mannose-binding protein functions as an opsonin. *J Exp Med* 1989; 169:1733-1745.

- Kwa D, Boeser-Nunnink B and Schuitemaker H. Lack of evidence for an association between a polymorphism in CX3CR1 and the clinical course of HIV infection or virus phenotype evolution. *AIDS* 2003a; 17:759-761.

- Kwa D, van Rij RP, Boeser-Nunnink B, Vingerhoed J and Schuitemaker H. Association between an interleukin-4 promoter polymorphism and acquisition of CXCR4 using HIV-1 variants. *AIDS* 2003b; 17:981-985.

- Kwok P-Y. Methods for genotyping single nucleotide polymorphisms. *Ann Rev Genomics Hum Genet* 2001; 2:235-258.

- Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J and Hendrickson WA. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and neutralizing human antibody. *Nature* 1998; 393:648-659.

- Lander ES and Schork NJ. Genetic dissection of complex traits. *Science* 1994; 265:2037-2048.

- Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature* 2001; 409: 860-921.

- Lapham CK, Ouyang J, Chandrasekhar B, Nguyen NY, Dimitrov DS and Golding H. Evidence for cell-surface association between fusin and the CD4-gp12 complex in human cell lines. *Nature* 1996; 274:602-605.

- Lasky LA, Nakamura G, Smith DH, et al. Delineation of a region of the human immunodeficiency virus type 1 gp120 glycoprotein critical for interaction with the CD4 receptor. *Cell* 1987; 50:975-985.

- Le Beau MM, Espinosa R, Neuman WL, et al. Cytogenetic and molecular delineation of the smallest commonly deleted region of chromosome 5 in malignant myeloid diseases. *Proc Natl Acad Sci USA* 1993; 90:5484-5488.

- Lee LG, Connell CR and Bloch W. Allelic discrimination by nick-translation PCR with fluorogenic probes. *Nucleic Acids Res* 1993; 21:3761-3766.

- Lee B and Montaner LJ. Chemokine immunobiology in HIV-1 pathogenesis. *J Leukoc Biol* 1999; 65:552-565.

- Lerman LS and Silverstein K. Computational simutation of the DNA melting and its application to denaturing gel electrophoresis. *Meth Enzymol* 1987; 155:482-501.

- Lewontin RC. The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* 1964; 49:49-67.

- Libert F, Cochaux P, Beckman G, et al. The Delta ccr5 mutation conferring protection against HIV-1 in Caucasian populations has a single and recent origin in Northeastern Europe. *Hum Mol Genet* 1998; 7:399-406.

- Lipscombe RJ, Sumiya M, Hill AVS, et al. High frequencies in African and non-African populations of independent mutations in the mannose binding protein gene. *Hum Mol Genet* 1992; 1:709-715.

- Liu J, Fujiwara TM, Buu NT, et al. Identification of polymorphisms and sequence variants in the human homologue of the mouse natural resistance-associated macrophage protein gene. *Am J Hum Genet* 1995; 56:845-853.

- Liu R, Paxton WA, Choe S et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply exposed individuals to HIV-1 infection. *Cell* 1996; 86:367-377.

- Liu H, Chao D, Nakayama EE, et al. Polymorphism in RANTES chemokine promoter affects HIV-1 disease progression. *Proc Natl Acad Sci USA* 1999a; 96:4581-4585.

- Liu H, Shioda T, Nagai Y, et al. Distribution of HIV-1 disease modifying regulated on activation normal T cell expressed and secreted haplotypes in Asian, African and Caucasian individuals. *AIDS* 1999b; 13:2602-2603.

- Livak KJ, Marmaro J and Todd JA. Towards fully automated genomewide polymorphism screening. *Nat Genet* 1995a; 9:341-342.

- Livak KJ, Flood SJ, Marmaro J, Giusti W and Deetz K. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Appl* 1995b; 4:357-362.

- Livak KJ. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genet Anal* 1999; 14:143-149.

- Llano A, Barretina J, Blanco J, et al. Stromal cell-derived factor 1 prevents the emergence of syncytium-inducing phenotype of HIV-1 in vivo. *AIDS* 2001; 15:1890-1892.

- Lohmueller KE, Pearce CL, Pike M, Lander ES and Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 2003; 33:177-82.

- Lorson CL, Hahnen E, Androphy EJ and Wirth B. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc Natl Acad Sci USA* 1999; 96:6307-6311.

- Loxton AG, Treurnicht F, Laten A, van Rensburg EJ and Engelbrecht S. Sequence analysis of near full-length HIV type 1 subtype D primary strains isolated in Cape Town, South Africa, from 1984 to 1986. *AIDS Res Hum Retroviruses* 2005; 21:410-413.

- Maas J, de Roda Husman AM, Brouwer M, et al. Presence of the variant mannose-binding lectin alleles associated with slower progression to AIDS. *AIDS* 1998; 12:2275-2280.

- Mack M, Luckow B, Nelson PJ, et al. Aminooxypentane-RANTES induces CCR5 internalization but inhibits recycling: A novel inhibitory mechanism of HIV infectivity. *J Exp Med* 1998; 187:1215-1224.

- Maddon PJ, Dalgleish AG, McDougal JS, Clapham PR, Weiss RA and Axel R. The T4 encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* 1986; 47:333-348.

- Madsen HO, Garred P, Kurtzhals JAL, et al. A new frequent allele is the missing link in the structural polymorphism of the human mannan-binding protein. *Immunogenetics* 1994; 40:37-44.

- Maggi EM, Giudizi G, Biagiotti F, et al. Th2-like CD8$^+$ T cells showing B cell helper function and reduced cytolytic activity in human immunodeficiency virus type 1 infection. *J Exp Med* 1994; 180:489-495.

- Maho A, Bensimon A, Vassart G and Parmentier M. Mapping of the CCXCR1, CX3CR1, CCBP2 and CCR9 genes to the CCR cluster within the 3p21.3 region of the human genome. *Cytogenet Cell Genet* 1999; 87:265-268.

- Mangano A, Kopka J, Batalla M, et al. Protective effect of CCR2-64I and not CCR5-delta32 and SDF1-3'A in pediatric HIV-1 infection. *J Acquir Immune Defic Syndr* 2000; 23: 52-57.

- Marchini J, Cardon LR, Phillips MS and Donnelly P. The effects of human population structure on large genetic association studies. *Nat Genet* 2004; 36:512-517.

- Martin MP, Dean M, Smith MW, et al. Genetic acceleration of AIDS progression by a promoter variant of *CCR5*. *Science* 1998a; 282:1907-1911.

- Martin MP, Carrington M, Dean M, et al. *CXCR4* polymorphisms and HIV-1 pathogenesis. *J Acquir Immune Defic Syndr* 1998b; 19:430.

- Martinson JJ, Chapman NH, Rees DC, Liu YT and Clegg JB. Global distribution of the *CCR5* gene 32-basepair deletion. *Nat Genet* 1997; 16:100-103.

- Marquet S, Sánchez FO, Arias M, et al. Variants of the Human *NRAMP1* gene and altered human immunodeficiency virus infection susceptibility. *J Infect Dis* 1999; 180:1521-1525.

- Marquet S, Lepage P, Hudson TJ, Musser JM and Schurr E. Complete nucleotide sequence and genomic structure of the human NRAMP1 gene region on chromosome region 2q35. *Mammalian Genome* 2000; 11:755-762.

- Matsushita M and Fujita T. Ficolins and the lectin complement pathway. *Immunol Rev* 2001; 180:78-85.

- Matsuyma T, Kobayashi N, and Yamamoto N. Cytokines and HIV infection: is AIDS a tumor necrosis factor disease? *AIDS* 1991; 5:1405-1417.

- Mazzoli S, Trabattoni D, Lo Caputo S, et al. HIV-specific mucosal and cellular immunity in HIV-seronegative partners of HIV-seropositive individuals. *Nat Med* 1997; 3:1250-1257.

- McDermott DH, Zimmerman PA, Guignard F, Kleeberger CA, Leitman SF and Murphy PM. *CCR5* promoter polymorphism and HIV-1 disease progression. *Lancet* 1998; 352:866-870.

- McDermott DH, Colla JS, Kleeberger CA, et al. Genetic polymorphism in CX_3CR1 and risk of HIV disease. *Science* 2000a; 290:2031a.

- McDermott DH, Beecroft MJ, Kleeberger CA, et al. Chemokine RANTES promoter polymorphism affects risk of both HIV infection and disease progression in the Multicenter AIDS Cohort Study. *AIDS* 2000b; 14:2671-2678.

- McKeigue PM. Prospects for admixture mapping of complex traits. *Am J Hum Genet* 2005; 76:1-7.

- McKenzie SW, Dallalio G, North M, Frame P and Means RT. Serum chemokine levels in patients with non-progressing HIV infection. *AIDS* 1996; 10:F29-33.

- McQuire W, Hill AV, Allsopp CE, Greenwood BM, and Kwiatkowski D. Variation in the TNF-alpha promoter region associated with susceptibility to celebral malaria. *Nature* 1994; 371:508-510.

- Mellors JW, Griffith BP, Oritz MA, Landry ML and Ryan JL. Tumor necrosis factor α/cachectin enhances human immunodeficiency virus type 1 replicationin primary macrophages. *J Infect Dis* 163:78-82.

- Meroni L, Trabattoni D, Balotta C, et al. Evidence for type 2 cytokine production and lymphocyte activation in the early phases of HIV-1 infection. *AIDS* 1996; 10:23-30.

- Meyer L, Magierowska M, Hubert JB, et al. CC-chemokine receptor variants, SDF-1 polymorphism, and disease progression in 720 HIV-infected patients. SEROCO cohort. *AIDS* 1999; 13:624-626.

- Michael NL, Louie LG, Rohrbaugh AL, et al. The role of *CCR5* and *CCR2* polymorphisms in HIV-1 transmission and disease progression. *Nat Med* 1997; 3:1160-1162.

- Mombo LE, Lu CY, Ossari S, et al. Mannose-binding lectin alleles in sub-Saharan Affricans and relation with susceptibility to infections. *Genes Immun* 2003; 4:362-367.

- Montana G and Pritchard JK. Statistical tests for admixture mapping with case-control and cases-only data. *Am J Hum Genet* 2004; 75:771-789.

- Moriuchi H, Moriuchi M, Combadiere C, Murphy PM and Fauci AS. CD8[+] T-cell-derived soluble factor(s), but not β-chemokines RANTES, MIP-1α, and MIP-1β, suppress HIV-1 replication in monocyte/macrophages. *Proc Natl Acad Sci USA* 1996; 93:15341-15345.

- Moriuchi M, Moriuchi H, Turner W and Fauci AS. Cloning and analysis of the promoter region of CXCR4, a co-receptor for HIV-1 entry. *J Immunol* 1997; 159:4322-4329.

- Moodley D, Smith TL, Rensburg EJ, Moodley J and Engelbrecht S. HIV type 1 V3 region subtyping in Kwazulu Natal, a high-seroprevalence South African region. *AIDS Res Hum Retroviruses* 1998; 14:1015-1018.

- Moore JP, Jameson BA, Weiss RA and Sattentau QJ. The HIV-cell fusion reaction. In Viral Fusion Mechanisms (J. Bentz, ed.), Boca Raton, Fla: CRC Press 1993; 233-289.

- Montano MA, Nixon CP, Ndung'u T, et al. Elevated tumor necrosis factor-alpha activation of human immunodeficiency virus type 1 subtype C in Southern Africa is associated with NF-kappaB enhancer gain-of-function. *J Infect Dis* 2000; 181:76-81.

- Mueller RF and Young ID. Emery's elements of medical genetics. Tenth Edition. Churchill Livingstone, Edinburgh, UK, 1998.

- Muller F, Aukrust P, Lien E, Huag CJ, Espevik T and Froland SS. Enhanced interleukin-10 production in response to *Mycobacterium avium* products in mononuclear cells from patients with human immunodeficiency virus infection. *J Infect Dis* 1998; 177:586-594.

- Mummidi S, Ahuja SS, McDaniel BL and Ahuja SK. The human CC chemokine receptor 5 (CCR5) gene. *J Biol Chem* 1997; 272:30662-30671.

- Mummidi S, Ahuja SS, Gonzalez SA, et al. Genealogy of the *CCR5* locus and chemokine system gene variants associated with altered rates of HIV-1 disease progression. *Nat Med* 1998; 4:786-793.

- Myers RM, Fischer SG, Lerman LS and Maniatis T. Nearly all single base substitutions in DNA fragments joined to a GC-clamp can be detected by denaturing gradient gel electrophoresis. *Nucleic Acids Res* 1985a; 13:3131-3145.

- Myers RM, Lumelsky N, Lerman LS and Maniatis T. Detection of single base substitutions in total genomic DNA. *Nature* 1985b; 313:495-497.

- Myers RM, Maniatis T and Lerman LS. Detection and localization of single base changes by denaturing gradient gel electrophoresis. *Methods Enzymol* 1987; 155:501-527.

- Nabel G and Baltimore D. An inducible transcription factor activates expression of human immunodeficiency virus in T-cells. *Nature* 1987; 326:711-713.

- Nagasawa T, Kikutani H and Kishimoto T. Molecular cloning and structure of a pre-B-cell growth-stimulating factor. *Proc Natl Acad Sci USA* 1994; 91:2305-2309.

- Nakayama EE, Hoshino Y, Xin X, et al. Polymorphism in the interleukin-4 promoter affects acquisition of human immunodeficiency virus type 1 syncytium-inducing phenotype. *J Virol* 2000; 74:5452-5459.

- Nakayama EE, Meyer L, Iwamoto A, et al. Protective effect of   Interleukin-4-589T polymorphism on human immunodeficiency virus type 1 disease progression: relationship with virus load. *J Infect Dis* 2002; 185:1183-1186.

- Nakayama EE, Tanaka Y, Nagai Y, Iwamoto A and Shioda T. A *CCR2*-V64I polymorphism affects stability of CCR2A isoform. *AIDS* 2004; 18:729-738.

- Nedwin GE, Naylor SL, Sakaguchi AY, et al. Human lymphotoxin and tumor necrosis factor genes: structure, homology and chromosomal localization. *Nucleic Acids Res* 1985; 13:6361-6373.

- Nelson PJ, Kim HT, Manning WC, Goralski TJ and Krensky AM. Genomic organization and transcriptional regulation of the RANTES chemokine gene. *J Immunol* 1993; 151:2601-2612.

- Newton-Cheh C and Hirschhorn JN. Genetic association studies of complex traits: design and analysis issues. *Mut Res* 2005; 573:54-69.

- Nickel RG, Casolaro V, Wahn U, et al. Atopic dermatitis is associated with a functional mutation in the promoter of the C-C chemokine RANTES. *J Immunol* 2000; 164:1612-1616.

- Nomiyama H, Fukuda S, Ho M, Tanase S, Miura R and Yoshie O. Organization of the chemokine gene cluster on human chromosome 17q11.2 containing the genes for CC chemokine MPIF-1, HCC-2, HCC-1, LEC, and RANTES. *J Interferon and Cytokine Res* 1999; 19:227-234.

- Oberlin E, Amara A, Bachelerie F, et al. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature* 1996; 382:833-835.

- O'Brien TR, Winkler C, Dean M, et al. HIV-infection in a man homozygous for *CCR5-∆32*. *Lancet* 1997; 349:1219.

- O'Brien SJ and Moore JP. The effect of genetic variation in chemokines and their receptors on HIV transmission and progression to AIDS. *Immunol Rev* 2000; 177: 99-111.

- O'Brien SJ, Nelson GW, Winkler CA and Smith MW. Polygenic and multifactorial disease gene association in man: Lessons from AIDS. *Annu Rev Genet* 2000; 34:563-591.

- O'Brien SJ and Nelson GW. Human genes that limit AIDS. *Nature Genet* 2004; 36: 565-574.

- Olson JM, Witte JS and Elston RC. Tutorial in biostatistics. Genetic mapping of complex traits. *Stat Med* 1999; 18:2961-2981.

- Osborn L, Kunkel S, and Nabel GJ. Tumor necrosis factor alpha and interleukin-1 stimulate the human immunodeficiency virus enhancer by activation of the nuclear factor kappa B. *Proc Natl Acad Sci USA* 1989; 86:2336-2340.

- Pastinen T, Liitsola K, Niini P, Salminen M and Syvänen AC. Contribution of the CCR5 and MBL genes to susceptibility to HIV type 1 infection in the Finnish population. *AIDS Res Hum Retroviruses* 1998; 14:695-698.

- Pastore C, Picchio GR, Galimi E, et al. Two mechanisms for human immunodeficiency virus type 1 inhibition by N-terminal modifications of RANTES. *Antimicrob Agents Chemother* 2003; 47:509-517.

- Patarca R, Sandler D, Maher K, et al. Immunological correlates of disease severity in pediatric slow progressors with human immunodeficiency virus type 1 infection. *AIDS Res Hum Retroviruses* 1996; 12:1063-1068.

- Patterson BK, Czerniewski M, Andersson J, et al. Regulation of CCR5 and CXCR4 expression by type 1 and type 2 cytokines: CCR5 expression is downregulated by IL-10 in CD4-positive lymphocytes. *Clin Immunol* 1999; 91:254-262.

- Patterson N, Hattangadi N, Lane B, et al. Methods for High-Density Admixture Mapping of Disease Genes. *Am J Hum Genet* 2004; 74:979-1000.

- Papathanasopoulos MA, Hunt GM and Tiemessen CT. Evolution and diversity of HIV-1 in Africa – a review. *Virus Genes* 2003; 26:151-163.

- Paxton WA, Martin SR, Tse D, et al. Relative resistance to HIV-1 infection of CD4 lymphocytes from persons who remain uninfected despite multiple high-risk sexual exposure. *Nat Med* 1996; 2:412-417.

- Paxton WA, Liu R, Kang S, et al. Reduced HIV-1 infectibility of CD4+ lymphocytes from exposed-uninfected individuals: association with low expression of CCR5 and high production of β-chemokines. *Virology* 1998; 244:66-73.

- Paxton WA, Kang S, Liu R, et al. HIV-1 infectability of CD4+ lymphocytes with relation to beta-chemokines and the CCR5 coreceptor. *Immunol Lett* 1999; 66:71-5.

- Paxton WA, Neumann AU, Kang S, et al. RANTES production from CD4+ lymphocytes correlates with host genotype and rates of human immunodeficiency virus type 1 disease progression. *J Infect Dis* 2001; 183:1678-1681.

- Perneger TV. What's wrong with Bonferroni adjustments. *BMJ* 1998; 316:1236-1238.

- Petersen DC, Kotze MJ, Zeier MD, et al. Novel mutations identified using a comprehensive *CCR5*-denaturing gradient gel electrophoresis assay. *AIDS* 2001; 15:171-177.

- Petersen DC. The role of chemokine and chemokine receptor genes in genetic susceptibility to HIV infection in South Africa. *Thesis* 2002.

- Petersen DC, Laten A, Zeier MD, Grimwood A, Janse van Rensburg E and Hayes VM. Novel Mutations and SNPs identified in *CCR2* using a new comprehensive denaturing gradient gel electrophoresis assay. *Hum Mut* 2002; 20:253-259.

- Picard L, Simmons G, Power CA, Meyer A, Weiss RA and Clapham PR. Multiple extracellular domains of CCR-5 contribute to human immunodeficiency type 1 entry and fusion. *J Virol* 1997; 71:5003-5011.

- Pritchard JK and Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* 1999; 65:220-228.

- Pritchard JK, Stephens M and Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000; 155:945-959.

- Pritchard JK and Donnelly P. Case-control studies of association in structured or admixed populations. *Theor Popul Biol* 2001; 60: 227-237.

- Prohászka Z, Thiel S, Ujhelyi E, Szlávik J, Bánhegyi D and Füst G. Mannan- binding lectin serum concentrations in HIV-infected patients are influenced by the stage of disease. *Immunol Lett* 1997; 58:171-175.

- Quillent C, Oberlin E, Braun J, et al. HIV-1-resistance phenotype conferred by combination of two separate inherited mutations of the CCR5 gene. *Lancet* 1998, 351:14-18.

- Ramaley PA, French N, Kaleebu P, et al. Chemokine-receptor genes and AIDS risk. *Nature* 2002; 417:140.

- Ramjee G, Williams B, Gouws E, van Dyck E, de Deken B and Karim SA. The impact of incident and prevalent herpes simplex virus-2 infection on the incidence of HIV-1 infection among commercial sex workers in South Africa. *J Acquir Immune Defic Syndr* 2005; 39:333-339.

- Ranade K, Chang M-S, Ting C-T et al. High-throughput genotyping with single nucleotide polymorphisms. *Genome Res* 2001; 11:1262-1268.

- Raport CJ, Schweickart VL, Eddy RL, Shows TB and Gray PW. The orphan G-protein-coupled receptor-encoding gene V28 is closely related to genes for chemokine receptors and is expressed in lymphoid and neural tissues. *Gene* 1995; 163:295-299.

- Raport C, Gosling J, Schweickart V, Gray P and Charo I. Molecular cloning and functional characterization of a novel human CC chemokine receptor (CCR5) for RANTES, MIP-1beta and, and MIP-1 alpha. *J Biol Chem* 1996; 271:17161-17166.

- Rebbeck TR, Spitz M and Wu X. Assessing the function of genetic variants in candidate gene association studies. *Nat Rev Genet* 2004; 5:589-597.

- Rees LEN, Wood NAP, Gillespie KM, Lai KN, Gaston K and Mathieson PW. The interleukin-10 -1082G/A polymorphism: allele frequency in different populations and functional significance. *Cell Mol Life* Sci 2002; 59:560-569.

- Reynard MP, Turner D, and Navarrete CV. Allele frequencies of polymorphisms of tumour necrosis factor-α, interleukin-10, interferon-γ and interleukin-2 genes in a North European Caucasoid group from the UK. *Eur J of Immunogenetics* 2000; 27:241-249.

- Rines RD, van Orsouw NJ, Sigalas I, Li FP, Eng C and Vijg J. Comprehensive mutational scanning of the p53 coding region by two-dimensional gene scanning. *Carcinogenesis* 1998; 19:979-984.

- Rizzardi GP, Morawetz RA, Vicenzi E, et al. CCR2 polymorphism and HIV disease. Swiss HIV cohort. *Nat Med* 1998; 4:252-253.

- Rizzuto CD, Wyatt R, Hernandez-Ramos N, et al. A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. *Science* 1998; 280:1949-1953.

- Romagnani S, Del Prete G, Manetti R, et al. Role of TH1/TH2 cytokines in HIV infection. *Immunol Rev* 1994; 140:73-92.

- Roos MT, Lange JM, de Goede RE, et al. Viral phenotype and immune response in primary human immunodeficiency virus type 1 infection. *J Infect Dis* 1992; 165:427-432.

- Rosenwasser LJ, Klemm DJ, Dresback JK, et al. Promoter polymorphisms in the chromosome 5 gene cluster in asthma and atopy. *Clin Exp Allergy* 1995; 25 (Suppl 2): 74-78.

- Rosenwasser LJ and Borish L. Genetics of atopy and asthma: the rationale behind promoter-based candidate gene studies (IL-4 and IL-10). *Am J Resp Crit Care Med* 1997; 156:S152-S155.

- Rucker J, Samson M, Doranz BJ, et al.   Regions in beta-chemokine receptors CCR5 and CCR2b that determine HIV-1 cofactor specificity. *Cell* 1996; 87:437-446.

- Rucker J, Edinger AL, Sharron M, et al. Utilization of chemokine receptors, orphan receptors, and herpesvirus-encoded receptors by diverse human and simian immunodeficiency viruses. *J Virol* 1997; 71:8999-9007.

- Sachidanandam R, Weissman D, Schmidt SC, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 2001; 409:928-933.

- Saifuddin M, Hart ML, Gewurz H, Zhang Y and Spear GT. Interaction of mannose-binding lectin with primary isolates of human immunodeficiency virus type 1. *J Gen Virol* 2000; 81:949-955.

- Salanti G, Amountza G, Ntzani EE and Ioannidis JP. Hardy-Weinberg equilibrium in genetic association studies: an empirical evaluation of reporting, deviations, and power. *Eur J Hum Genet* 2005; 13:840-848.

- Samson M, Labbe O, Mollereau C, Vassart G and Parmentier M. Molecular cloning and functional expression of a new human CC-chemokine receptor gene. *Biochemistry* 1996a; 35:3362-3367.

- Samson M, Libert F, Doranz BJ, et al. Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 1996b; 382:722-725.

- SAS Institute Inc. software, Cary, North Carolina, (www.sas.com).

- Sastry K, Herman GA, Day L, et al. The human mannose- binding protein gene: exon structure reveals its evolutionary relationship to a human pulmonary surfactant gene and localization to chromosome 10. *J Exp Med* 1989; 170:1175-1189.

- Sattentau OJ, Zolla-Pazner S and Poignard P. Epitope exposure on functional, oligomeric HIV-1 gp41 molecules. *Virology* 1995; 206:713-717.

- Sawyer SL, Mukherjee N, Pakstis AJ, et al. Linkage disequilibrium patterns vary substantially among populations. *Eur J Hum Genet* 2005; 13:677-686.

- Schaid DJ and Rowland C. Use of parents, sibs and unrelated controls for detection of associations between genetic markers and disease. *Am J Hum Genet* 1998; 63:1492-1506.

- Schaid DJ and Jacobsen SJ. Biased tests of association: Comparisons of allele frequencies when departing from Hardy-Weinberg proportions. *Am J Epidemiol* 1999; 149:706-711.

- Schmidtmayerova H, Sherry B and Bukrinsky M. Chemokines and HIV replication. *Nature* 1996a; 382:767.

- Schmidtmayerova H, Nottet HS, Nuovo G, et al. Human immunodeficiency virus type 1 infection alters chemokine beta peptide expression in human monocytes: implications for recruitment of leukocytes into brain and lymph nodes. *Proc Natl Acad Sci USA* 1996b; 93:700-704.

- Schols D and De Clercq E. Human immunodeficiency virus type 1 gp120 induces anergy in human peripheral blood lymphocytes by inducing interleukin-10 production. *J Virol* 1996; 70:4953-4960.

- Schuffenecker I, Narod SA, Ezekowitz RAB, Sobol H, Feunteun J and Lenoir GM. The gene for mannose-binding protein maps to chromosome 10 and is a marker for multiple endocrine neoplasia type 2. *Cytogenet Cell Genet* 1991; 56:99-102.

- Schuitemaker H, Koot M, Koostra NA, et al. Biological phenotype of human immunodeficiency virus type 1 clones at different stages of infection: Progression of disease is associated with shift from monocytotropic to T-cell tropic populations. *J Virol* 1992; 66:1354-1360.

- Searle S and Blackwell JM. Evidence for a functional repeat polymorphism in the promoter of the human NRAMP1 gene that correlates with autoimmune versus infectious diseases. *J Med Genet* 1999; 36:295-299.

- Senaldi G, Davies ET, Mahalingam M, et al. Circulating levels of mannose binding protein in human immunodeficiency virus infection. *J Infect* 1995; 31:145-148.

- Sevall JS. Rapid allelic discrimination from real-time DNA amplification. *Methods* 2001; 25:452-455.

- Shalekoff S and Tiemessen CT. Circulating levels of stromal cell-derived factor 1α and Interleukin 7 in HIV type 1 infection in pulmonary tuberculosis are reciprocally related to CXCR4 expression on peripheral blood leukocytes. *AIDS Res Hum Retroviruses* 2003; 19:461-468.

- Sheffield VC, Cox DR, Lerman LS and Myers RM. Attachment of a 40 base-pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single base changes. *Proc Natl Acad Sci USA* 1989; 86:232-236.

- Shin HO, Winkler C, Stephens JC, et al. Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of IL10. *Proc Natl Acad Sci USA* 2000; 97:14467-72.

- Shirozu M, Nakano T, Inazawa J, et al. Structure and chromosomal localization of the human stromal cell-derived factor (SDF1) gene. *Genomics* 1995; 28:495-500.

- Signoret N, Oldridge J, Pelchen-Matthews A, et al. Phorbol esters and SDF-1 induce rapid endocytosis and down modulation of the chemokine receptor CXCR4. *J Cell Biol* 1997; 139:651-664.

- Simmons G, Wilkinson D, Reeves JD, et al. Primary, syncytium-inducing human immunodeficiency virus type 1 isolates are dual-tropic and most can use either Lestr or CCR5 as coreceptors for virus entry. *J Virol* 1996; 70:8355-8360.

- Singh KK, Hughes MD, Chen J and Spector SA. Genetic polymorphisms in *CX3CR1* predict HIV-1 disease progression in children independently of CD4+ lymphocyte count and HIV-1 RNA load. *J Infect Dis* 2005; 191:1971-1980.

- Sliutz G, Schmidt W, Tempfer C, et al. Detection of p53 point mutations in primary human vulvar cancer by PCR and temperature gradient gel electrophoresis. *Gynecol Oncol* 1997; 64:93-98.

- Smith MW, Dean M, Carrington M, et al. Contrasting genetic influence of *CCR2* and *CCR5* variants on HIV-1 infection and disease progression. *Science* 1997a; 277: 959-965.

- Smith MW, Carrington M, Winkler C, et al. CCR2 chemokine receptor and AIDS progression. *Nat Med* 1997b; 3:1052-1053.

- Smith MW, Patterson N, Lautenberger JA, et al. A high-density admixture map for disease gene discovery in African Americans. *Am J Hum Genet* 2004; 74:1001-1013.

- Smith-Sørensen B, Gebhardt MC, Kloen P, et al. Screening for TP53 mutations in osteosarcomas using constant denaturant gel electrophoresis (CDGE). *Hum Mutat* 1993; 2:274-285.

- Smolnikova MV and Konenkov VI. Association of IL2, TNFA, IL4 and IL10 promoter gene polymorphisms with the rate of progression of HIV infection. *Russ J Immunol* 2002; 7:349-356.

- Sölder BM, Schultz TF, Hengster P, et al. HIV and HIV-infected cells differentially activate the human complement system independent of antibody. *Immunol Lett* 1989; 22:135-146.

- Song Z, Casolaro V, Chen R, Georas SN, Monos D and Ono SJ. Polymorphic nucleotides within the human IL-4 promoter that mediate overexpression of the gene. *J Immunol* 1996; 156:424-429.

- Soriano A, Martinez C Garcia F, et al. Plasma stromal cell-derived factor (SDF)-1 levels, SDF1-3' A genotype, and expression of CXCR4 on T lymphocytes: Their impact on resistance to human immunodeficiency virus type 1 infection and its progression. *J Infect Dis* 2002; 186:922-931.

- Speck RF, Wehrly K, Platt EJ, et al. Selective employment of chemokine receptors as human immunodeficiency virus type 1 coreceptors determined by individual amino acids within the envelope V3 loop. *J Virol* 1997; 71:7136-7139.

- Spies T, Morton CC, Nedospasov SA, Fiers W, Pious D, and Strominger JL. Genes for the tumor necrosis factors alpha and beta are linked to the human major histocompatibility complex. *Proc Natl Acad Sci USA* 1986; 83:8699-8702.

- Statistics South Africa (www.statssa.gov.za). Census 2001.

- Stephens JC, Reich DE, Goldstein DB, et al. Dating the origin of the CCR5-Delta 32 AIDS-resistance allele by the coalescence of haplotypes. *Am J Hum Genet* 1998; 62:1507-1515.

- Suh Y and Vijg J. SNP discovery in associating genetic variation with human disease phenotypes. *Mut Res* 2005; 573:41-53.

- Sumiya M, Super M, Tabona P, et al. Molecular basis of opsonic defect in immunodeficient children. *Lancet* 1991; 337:1569-1570.

- Super M, Thiel S, Lu J, Levinsky RJ and Turner MW. Association of low levels of mannan-binding protein with a common defect of opsonisation. *Lancet* 1989; 2:1236-1239.

- Sutherland GR, Baker E, Callen DF, et al. Interleukin 4 is at 5q31 and interleukin 6 is at 7p15. *Hum Genet* 1988; 79:335-337.

- Suzuki Y, Koyanagi Y, Tanaka Y, et al. Determinant in human immunodeficiency virus type 1 for efficient replication under cytokine-induced CD4$^+$ T-helper 1 (Th1) - and Th2-type conditions. *J Virol* 1999; 73:316-324.

- Takabayashi A, Ihara K, Sasaki Y, Kusuhara K, Nishima S and Hara T. Novel polymorphism in the 5'untranslated region of the interleukin-4 gene. *J Hum Genet* 1999; 44:352-53.

- Tanaka M, Hirabayashi Y, Gatanaga H, et al. Reduction in interleukin-2-producing cells but not Th1 to Th2 shift in moderate and advanced stages of human immunodeficiency virus type-1-infection: direct analysis of intracellular cytokine concentrations in CD4+CD8-T cells. *Scand J Immunol* 1999; 50:550-554.

- Tashiro K, Tada H, Heilker R, Shirozu M, Nakano T and Honjo T. Signal sequence trap: A cloning strategy for secreted protein and type I membrane proteins. *Science* 1993; 261:600-603.

- Taylor ME, Brickell PM, Craig RK and Summerfield JA. Structure and evolutionary origin of the gene encoding a human serum mannose-binding protein. *Biochem J* 1989; 262:763-771.

- Taylor JG, Choi EH, Foster CB and Chanock SJ. Using genetic variation to study human disease. *Trends Molec Med* 2001; 7:507-512.

- Tersmette M, deGoede REY, Ai BJM, et al. Differential syncytium-inducing capacity of human immunodeficiency virus isolates: Frequent detection of syncytium-inducing isolates in patients with acquired immunodeficiency virus syndrome (AIDS) and AIDS-related complex. *J Virol* 1988; 62:2026-2032.

- Tersmette M, Gruters R, Wolf F, et al. Evidence for a role of virulent human immunodeficiency virus (HIV) variants in the pathogenesis of acquired immunodeficiency virus syndrome: Studies on sequential isolates. *J Virol* 1989; 63:2118-2125.

- Than S, Hu R, Oyaizu N, et al. Cytokine pattern in relation to disease progression in human immunodeficiency virus-infected children. *J Infect Dis* 1997; 175:47-56.

- Theodorou I, Meyer L, Magierowska M, Katlama C and Rouzioux C. HIV-1 infection in an individual homozygous for *CCR5*-Δ32. *Lancet* 1997; 349:1219-1220.

- Thiel S, Holmskov U, Hviid L, Laursen SB and Jensenius JC. The concentration of the C-type lectin, mannan-binding protein, in human plasma increases during acute phase response. *Clin Exp Immunol* 1992; 90:31-35.

- Thielens NM, Tacnet-Delorme P and Arlaud GJ. Interaction of C1q and mannan-binding lectin with viruses. *Immunobiology* 2002; 205:563-574.

- Thomas DC and Witte JS. Point: Population stratification: A problem for case-control studies of candidate-gene associations. *Cancer Epidemiol Biomark and Prevent* 2002; 11:505-512.

- Thomas DC and Clayton DG. Betting odds and genetic associations. *J Natl Cancer Inst* 2004; 96:421-423.

- Tishkoff SA and Williams SM. Genetic analysis of African populations: Human Evolution and complex disease. *Nat Rev Genet* 2002; 3:611-621.

- Tishkoff SA and Verrelli BC. Patterns of human genetic diversity: Implications for human evolutionary history and disease. *Annu Rev Genomics Hum Genet* 2003; 4:293-340.

- Tiensiwakul P. Stromal cell-derived factor (SDF) 1-3'A polymorphism may play a role in resistance to HIV-1 infection in seronegative high-risk Thais. *Intervirology* 2004; 47:87-92.

- Torres Y, Medrano FJ, Rey C, Calderon EJ, et al. Evidence for a role of T-helper type 2 cytokines in the acquisition of human immunodeficiency virus syncytium-inducing phenotype. *Eur J Clin Invest* 1998; 28:930-936.

- Torres G, Garcia V, Sanchez E, Segarra A, Patterson BK and Melendez-Guerrero LM. Expression of the HIV-1 co-receptors CCR5 and CXCR4 on placental macrophages and the effect of IL-10 on their expression. *Placenta* 2001; 22 (Suppl A):S29-S33.

- Tresoldi E, Romiti ML, Boniotto M, et al. Prognostic value of the stromal cell-derived factor 1 3'A mutation in pediatric human immunodeficiency virus type 1 infection. *J Infect Dis* 2002; 185:696-700.

- Trivedi HN, Plummer FA, Anzala AO, et al. Resistance to HIV-1 infection among African sex workers is associated with global hyporesponsiveness in interleukin 4 production. *FASEB J* 2001; 15:1795-1797.

- Trkola A, Dragic T, Arthos J, et al. CD4-dependant antibody-sensitive interactions between HIV-1 and its co-receptor CCR5. *Nature* 1996; 384:184-187.

- Tsai C-T, Fallin D, Chiang F-T, et al. Angiotensinogen gene haplotype and hypertension. Interaction with ACE gene I allele. *Hypertension* 2003; 41:9-15.

- Turner MW. Mannose binding lectin: the pluripotent molecule of the innate immune system. *Immunol Today* 1996; 17:532-540.

- Turner DM, Williams DM, Sankaran D, Lazarus M, Sinnott PJ and Hutchinson IV. An investigation of polymorphism in the interleukin 10 gene promoter. *Eur J Immunogenet* 1997; 24:1-8.

- Turner MW. The role of mannose-binding lectin in health and disease. *Mol Immunol* 2003; 40:423-429.

- Ullum H, Cozzi Lepri A, Victor J, et al. Production of beta-chemokines in human immunodeficiency (HIV) infection: evidence that high levels of macrophage inflammatory protein-1beta are associated with a decreased risk of HIV disease progression. *J Infect Dis* 1998; 177:331-336.

- UNAIDS (www.unaids.org). AIDS epidemic update, December 2004.

- Valentin A, Lu W, Rosati M, et al. Dual effect of interleukin 4 on HIV-1 expression: Implications for viral phenotypic switch and disease progression. *Proc Natl Acad Sci USA* 1998; 95:8886-8891.

- van Harmelen JH, Wood R, Lambrick M, Rybicki EP, Williamson AL and Williamson C. An association between HIV-1 subtypes and mode of transmission in Cape Town, South Africa. *AIDS* 1997; 11:81-87.

- van Harmelen JH, van der Ryst E, Loubser AS et al., A predominantly HIV type 1 C-restricted epidemic in South African urban populations. *AIDS Res Hum Retroviruses* 1999; 15:395-398.

- van Harmelen JH, Shephard E, Thomas R, Hanke T, Williamson AL and Williamson C. Construction and characterisation of a candidate HIV-1 subtype C DNA vaccine for South Africa. *Vaccine* 2003; 21:4380-4389.

- van Rij RP, Broersen S, Goudsmit J, et al. The role of a stromal cell-derived factor-1 chemokine gene variant in the clinical course of HIV-1 infection. *AIDS* 1998; 12:F85-F90.

- Vasilescu A, heath SC, Ivanova R, et al. Genomic analysis of Th1-Th2 cytokine genes in an AIDS cohort: identification of IL4 and IL10 haplotypes associated with the disease progression. *Gene Immun* 2003; 4:441-449.

- Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science* 2001; 291:1304-1351.

- Vieira P, de Waal-Malefyt R, Dang MN, et al. Isolation and expression of human cytokine synthesis inhibitory factor cDNA clones: homology to Epstein-Barr virus open reading frame BCRF1. *Proc Natl Acad Sci USA* 1991; 88:1172-1176.

- Wainberg MA. HIV-1 subtype distribution and problem of drug resistance. *AIDS* 2004; 18:S63-S68.

- Wallis R and Cheng JYT. Molecular defects in variant forms of mannose-binding protein associated with immunodeficiency. *J Immunol* 1999; 163:4953-4959.

- Wang J, Roderiquez G, Oravecz T and Norcross MA. Cytokine regulation of human immunodeficiency virus type 1 entry and replication in human monocytes/macrophages through modulation of CCR5 expression. *J Virol* 1998a; 72:7642-7647.

- Wang J, Harada A, Matsushita S, et al. IL-4 and a glucocorticoid up-regulate CXCR4 expression on human CD4$^+$ T lymphocytes and enhance HIV-1 replication. *J Leukoc Biol* 1998b; 64:642-649.

- Wang Z, Lee B, Murray JL, et al. CCR5 HIV-1 coreceptor activity. *J Biol Chem* 1999; 274:28413-28419.

- Wang J, Crawford K, Yuan M, Wang H, Gorry PR and Gabuzda D. Regulation of CC chemokine receptor 5 and CD4 expression and human immunodeficiency virus type 1 replication in human macrophages and microglia by T helper type 2 cytokines. *J Infect Dis* 2002; 185:885-897.

- Wang C, Song W, Lobashevsky E, et al. Cytokine and chemokine gene polymorphisms among ethnically diverse North Americans with HIV-1 infection. *J Acquir Immune Defic Syndr* 2004; 35:446-54.

- Wang WYS, Barratt BJ, Clayton DG and Todd JA. Genome-wide association studies: Theoretical and practical concerns. *Nat Rev Genet* 2005; 6:109-118.

- Wasik TJ, Jagodzinski PP, Hyjek EM, et al. Diminished HIV-specific CTL activity is associated with lower type 1 and enhanced type 2 responses to HIV-specific peptides during perinatal HIV infection. *J Immunol* 1997; 158:6029-6036.

- Wegner SA, Ehrenberg PK, Chang G, et al. Genomic organization and functional characterization of the chemokine receptor CXCR4, a major entry co-receptor for human immunodeficiency virus type 1. *J Biol Chem* 1998; 273: 4754-4760.

- Weissman D, Poli G and Fauci AS. Interleukin 10 blocks HIV replication in macrophages by inhibiting the autocrine loop of tumor necrosis factor alpha and interleukin 6 induction of virus. *AIDS Res Hum Retroviruses* 1994; 10:1199-1206.

- Wigginton JE, Cutler DJ and Abecasis GR. A note on exact tests of Hardy Weinberg equilibrium. *Am J Hum Genet* 2005; 76:887-893.

- Williams BG, Taljaard D, Campbell CM, et al. Changing patterns of knowledge, reported behaviour and sexually transmitted infections in South African gold mining community. *AIDS* 2003; 17:2099-2107.

- Williamson C, Engelbrecht S, Lambrick M, et al. HIV-1 subtypes in different risk groups in South Africa. *Lancet* 1995; 346:782.

- Williamson C, Loubser SA, Brice B, et al. Allelic frequencies of host genetic variants influencing susceptibility to HIV-1 infection and disease in South African populations. *AIDS* 2000; 14: 449-451.

- Wilson AG, di Giovine FS, Blakemore ALF, and Duff GW. Single base polymorphism in the human tumour necrosis factor alpha (TNFα) gene detectable by *NcoI* restriction of PCR product. *Hum Mol Genet* 1992; 1:353.

- Winkler C, Modi W, Smith MW, et al. Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. *Science* 1998; 279:389-393.

- Winkler C, An P and O'Brien SJ. Patterns of ethnic diversity among the genes that influence AIDS. *Hum Mol Genet* 2004; 1:R9-19.

- Witte JS, Gauderman WJ and Thomas DC. Asymptomatic bias and efficiency in case-control studies of candidate genes and gene-environment interactions: basic family designs. *Am J Epidemiol* 1999; 149:693-705.

- Wong L-M, Myers SJ, Tsou C-L, Gosling J, Arai H and Charo IF. Organization and differential expression of the human monocyte chemoattractant protein 1 receptor gene. *J Biol Chem* 1997; 272:1038-1045.

- Wu L, Gerard NP, Wyatt R, et al. CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature* 1996, 384:179-183.

- Wu Y, Hayes VM, Osinga J, et al. Improvement of fragment and primer selection for mutation detection by denaturing gradient gel electrophoresis. *Nucleic Acids Res* 1998; 26:5432-5440.

- Yokota T, Otsuka T, Mosmann T, et al. Isolation and characterization of a human interleukin cDNA clone, homologous to mouse B-cell stimulatory factor 1, that expresses B-cell and T-cell-stimulating activities. *Proc Natl Acad Sci USA* 1986; 83:5894-5898.

- Zagury D, Lachgar A, Chams V, et al. C-C chemokines, pivotal in protection against HIV type 1 infection. *Proc Natl Acad Sci USA* 1998; 95:3857-3861.

- Zhu T, Mo H, Wang N, et al. Genotypic and phenotypic characterization of HIV-1 in patients with primary infection. *Science* 1993, 261:1179-1181.

- Zimmerman PA, Buckler-White A, Alkhatib G, et al. Inherited resistance to HIV-1 conferred by an inactivating mutation in the CC chemokine receptor 5: studies in populations with contrasting clinical phenotypes, defined racial background and quantified risk. *Molec Med* 1997, 3:23-26.

- Zuma K, Lurie MN, Williams BG, Mkaya-Mwamburi D, Garnett GP and Sturm AW. Risk factors of sexually transmitted infections among migrant and non-migrant sexual partnerships from rural South Africa. *Epidemiol Infect* 2005; 133:421-428.

Chapter 2.1.

The effect of *CCR5*, *CCR2*, *CX3CR1* and *CCL5 (RANTES)* SNPs on susceptibility to HIV-1 infection in an African population

Desiree C. Petersen[1], Sadeep Shrestha[2,3], Julie Bergeron[3], Bert Gold[2], Vanessa M. Hayes[4,5] and Michael Dean[2]

Department of Medical Virology[1], University of Stellenbosch, Tygerberg Medical School, South Africa; Laboratory of Genomic Diversity[2], National Cancer Institute, Frederick, Maryland, USA; SAIC[3], Frederick, Maryland, USA; Cancer Research Program[4], Garvan Institute of Medical Research, Darlinghurst, Sydney, NSW, Australia; University of New South Wales [5], Sydney, Australia.

(Full article)

ABSTRACT

Mutations in genes that encode chemokines and chemokine receptors have been implicated in host susceptibility to human immunodeficiency virus-1 (HIV-1) infection, and subsequent progression to acquired immunodeficiency syndrome (AIDS). More recently, the associations between these genetic variants and HIV-1/AIDS pathogenesis have emerged as being population-specific. Limited research exists on the effects of previously reported HIV-1/AIDS-associated polymorphisms in African-based ethnic groups, particularly those residing in the pandemic-stricken regions of the world. This prompted our case-control study of Sub-Saharan Africans of Xhosa descent, consisting of 215 HIV-1 seropositive and 113 HIV-1 seronegative individuals. We screened *CCR5*, *CCR2*, *CX3CR1* and *CCL5* (*RANTES*), with significance being observed for three single nucleotide polymorphisms (SNPs), including *CCR5*-2733A>G, *CCR5*-2135C>T and *CX3CR1*V249I (G>A). The *CCR5*-2733 G allele (P = 0.0409) together with the heterozygous genotype, *CCR5*-2733 AG (P = 0.0270), was associated with a reduced risk for HIV-1 infection. The *CX3CR1*249I allele (P = 0.0479) was also associated with reduced risk to HIV-1 infection. The opposite was observed for the *CCR5*-2135 T allele (P = 0.0024) and homozygosity for *CCR5*-2135 TT (P = 0.0096), which were associated with increased susceptibility to HIV-1 infection. All these associations with susceptibility to HIV-1 infection appear to be unique to the Xhosa ethnic group. This result emphasises the importance of elucidating the specific effect of known HIV-1/AIDS candidate gene variants in understudied populations.

INTRODUCTION

Host susceptibility for human immunodeficiency virus-1 (HIV-1) infection and rates of disease progression to acquired immunodeficiency syndrome (AIDS) varies widely amongst individuals. This observation was advanced by the discovery of HIV suppressive chemokines. These chemokines modulate the efficiency of HIV-1 infections by serving as the natural ligands for specific chemokine receptors, which together with the CD4$^+$ molecule act as co-receptors for HIV-1 cell entry. Elevated levels of chemokines and low expression of chemokine receptors have been associated with relative resistance to HIV-1 infection in exposed high-risk individuals, and delayed disease progression to AIDS [Reviewed in Hogan and Hammer, 2001; Anastassopoulou and Kostrikis, 2003; O'Brien and Nelson, 2004; Winkler et al., 2004; Kaslow et al., 2005].

The study of host genetic factors interacting with other parameters (viral, environmental/socio-economic and host immunological factors) in determining susceptibility to HIV-1/AIDS pathogenesis was therefore accentuated by the identification of mutations in genes encoding specific chemokines and chemokine receptors. CC chemokine receptor 5 (CCR5) (MIM# 601373), CC chemokine receptor 2 (CCR2) (MIM# 601267) and CX3C chemokine receptor 1 (CX3CR1) (MIM# 601470) have all been identified as co-receptors for HIV-1 infection. CC chemokine ligand 5 (CCL5)/RANTES (MIM# 187011), a natural ligand for CCR5, can directly compete with the virus for receptor binding. A large number of polymorphisms and/or haplotypes influencing host susceptibility to HIV-1 infection and/or disease progression to AIDS have been

identified in the *CCR5, CCR2, CX3CR1* and *CCL5* genes [Reviewed in Carrington et al., 2001; Hogan and Hammer, 2001; Dean et al., 2002; Anastassopoulou and Kostrikis, 2003; O'Brien and Nelson, 2004; Winkler et al., 2004; Kaslow et al., 2005]. These previously reported studies however provide evidence that certain HIV-1/AIDS associations with specific genetic polymorphisms are population-based and thus restricted to certain ethnic groups. This emphasises the need for further investigation, in particularly of the understudied African populations.

Our study focused on the screening and analysis of genetic variants in a Xhosa ethnic group from Sub-Saharan Africa where the HIV-1/AIDS pandemic is most pronounced. We performed genotyping for single nucleotide polymorphisms (SNPs) in the promoter region of *CCR5* (positions -2733A>G, -2554G>T, -2135C>T, -2086A>G, and -1835C>T, relative to the translation start site) [Mummidi et al., 1997; Martin et al., 1998; McDermottt et al., 1998; Mummidi et al., 1998], the coding regions of *CCR2* (*CCR2*V64I G>A) [Smith et al., 1997] and *CX3CR1* (*CX3CR1*V249I G>A and *CX3CR1*T280M C>T) [Faure et al., 2000], and the promoter and intronic region of *CCL5* (positions -403G>A, relative to the transcription start site and IVS1+307T>C, previously designated In1.1T>C) [Liu et al., 1999a; An et al., 2002]. Furthermore, we assessed allele, genotype and haplotype distributions within our case-control sample group for associations with HIV-1 susceptibility.

MATERIALS AND METHODS

Sample group

Our Sub-Saharan African sample group was represented by 328 individuals of Xhosa descent who are all residents of the same geographical region within South Africa. The 215 HIV-1 seropositive individuals were patients at Tygerberg Hospital, Woodstock Chapel Street Community Health Clinic or the Langa Clinic, all in the Western Cape Province of South Africa. The 113 population-matched HIV-1 seronegative controls were healthy blood donors from the Western Province Blood Transfusion Service of South Africa. Clinical information regarding disease staging of HIV-1 seropositive patients is limited and thus rates of progression to AIDS remain largely unknown. Informed consent was obtained from all the study participants and the Ethics Review Committee of the University of Stellenbosch approved the study protocol (#98/158).

Genotyping

The 5' nuclease or TaqMan allelic discrimination method [Livak et al., 1999] allowed for the complete screening of SNPs using the probes and primers as listed in Table 1. The TaqMan Universal PCR Master Mix (Applied Biosystems, Foster City, California) was used for amplification and each PCR reaction contained 5 ng of DNA (detailed PCR reaction mix protocol available on request). Cycling conditions included an initial denaturation of $95^{\circ}C$ for 10 minutes, followed by 50 cycles of denaturation at $95^{\circ}C$ for 15 seconds and annealing for 1 minute (annealing temperatures shown in Table 1). The ABI7900 high throughput sequence detection system (Applied Biosystems,

Foster City, California) was used to measure end-point fluorescence and generate data for allelic discrimination and genotype determination.

Statistical analysis

Allele and genotype distributions were calculated and tested for consistency with Hardy Weinberg equilibrium expectations. The Fisher's exact or chi-square (χ^2) tests for 2x2 contingency tables were used for determining significant heterogeneity in allele frequencies between cases and controls. The significance of genotype frequencies was also considered using the χ^2 test for independence. Pairwise linkage disequilibrium (LD) analysis was based on the Lewontins D' (D' = 1 in the presence of two or three haplotypes) and/or r^2 coefficient (r^2 = 1 in the presence of two haplotypes) measures. Haplotypes were estimated using the expectation-maximization (E-M) algorithm (GraphPad Software Inc, San Diego, California and SAS Institute Inc. software, Cary, North Carolina).

RESULTS

CCR5 and _CCR2_ SNPs

Allele and genotype frequencies for the _CCR5_ promoter variants (-2733A>G, -2554G>T, -2135C>T, -2086A>G and -1835C>T) and _CCR2_V64I (G>A) polymorphism in the cases versus controls are shown in Table 2. Significant heterogeneity was present for the allele and genotype frequencies of two _CCR5_ promoter SNPs, _CCR5_-2733A>G and _CCR5_-2135C>T. The _CCR5_-2733 G allele occurred at a higher frequency in the HIV-1 seronegative controls compared to the HIV-1 seropositive patients (OR = 2.269, 95% CI = 1.060 -

4.858, *P* = 0.0409).   The genotype distribution was also significant, with the *CCR5*-2733 AG heterozygous genotype being more prevalent in the HIV-1 uninfected group (*P* = 0.0270).   Analysis of the *CCR5*-2135C>T SNP showed that the *CCR5*-2135 T allele occurs at a higher frequency in the HIV-1 seropositive patients when compared to the HIV-1 seronegative controls (OR = 0.5985, 95% CI = 0.4303 – 0.8324, *P* = 0.0024). This significance in allele frequency distribution is driven by the increase of the *CCR5*-2135 TT homozygous genotype in the HIV-1 infected group (*P* = 0.0096).   No significance was observed for the *CCR2*V64I SNP when comparing both the allele and genotype frequencies in the cases versus the controls.   Pairwise D' measures indicated strong LD, D' = 1, for the majority of the closely lying *CCR2* and *CCR5* SNPs.   The exceptions include *CCR2*V64I with *CCR5*-2733A>G (D' = 0.66) and *CCR5*-2733A>G with *CCR5*-1835C>T (D' = 0.36). The more stringent r^2 coefficient measures, where complete pairwise LD is indicated by r^2 = 1, suggested strong LD for *CCR2*V64I and *CCR5*-1835C>T exclusively (r^2 = 0.91). Pairwise LD based on D' measures are presented in Figure 1, as determined by the Haploview software [Barrett et al., 2005].   Estimated haplotype analysis (data not shown) for specific *CCR2 and CCR5* allele combinations and susceptibility to HIV-1 infection were not significant.

CX3CR1 SNPs

The frequencies of allele and genotype distributions for the *CX3CR1*V249I (G>A) and *CX3CR1*T280M (C>T) SNPs are shown in Table 2. The increased occurrence of the *CX3CR1*249I allele in the HIV-1 seronegative controls when compared to the HIV-1 seropositive patients was significant (OR = 1.630, 95%

CI = 1.023 – 2.595, *P* = 0.0479). Overall, the *CX3CR1*V249I SNP (0.138) occurred more commonly than the *CX3CR1*T280M SNP (0.024). The r^2 coefficient measure provided no support for strong pairwise LD between *CX3CR1* alleles (r^2 = 0.37). No further associations with HIV-1 susceptibility were observed, as no significance was present for genotype frequencies and haplotype estimations (data not shown) of the *CX3CR1*V249I and *CX3CR1*T280M SNPs.

CCL5 SNPs

The *CCL5* SNPs (-403G>A and *IVS1+307T>C*) together with their observed allele and genotype frequencies are shown in Table 2. No significant heterogeneity was identified between the allele distribution of the *CCL5*-403G>A and *CCL5 IVS1+307T>C* SNPs in the HIV-1 seropositive patients versus the HIV-1 seronegative controls. *CCL5*-403G>A (0.409) did, however, generally occur at a higher allele frequency compared to *CCL5 IVS1+307T>C* (0.222). Pairwise LD analysis based on the r^2 coefficient measure suggested no LD between the *CCL5* alleles (r^2 = 0.43). Independent genotype analysis and estimated haplotypes (data not shown) also showed no significant association with HIV-1 susceptibility.

DISCUSSION

Various studies have indicated that the influence of many gene variants on host susceptibility to HIV-1 infection and disease progression to AIDS may be population-specific [Reviewed in O'Brien and Nelson, 2004; Winkler et al., 2004; Kaslow et al., 2005]. This encouraged the validation of known HIV-1/AIDS

candidate gene polymorphisms in a Sub-Saharan African ethnic group consisting of individuals of Xhosa descent. We have previously established that the commonly studied Caucasian-based HIV-1/AIDS resistance-associated variant, *CCR5*Δ32, is not present in this Xhosa sample group [Petersen et al., 2001]. In this study, we therefore continued with the screening of previously reported SNPs occurring in the promoter or coding regions of the *CCR5*, *CCR2*, *CX3CR1* and *CCL5* genes using the TaqMan allele discrimination method. No significant deviations from the Hardy-Weinberg equilibrium estimations were observed for both the cases and controls. Since the majority of HIV-1 seropositive individuals have no clinical staging (3 slow, 34 normal, 11 fast and 167 unknown progressors), associations with disease progression to AIDS were not considered.

Allele frequencies reported here and previously observed for *CCR5* promoter variants in different African-based populations [John et al., 2001; Dean et al., 2002; Ramaley et al., 2002] have not been directly comparable for all SNPs. This further suggests that the influence of host genetic diversity on susceptibility to HIV-1/AIDS is distinct for specific African subpopulations. There have been studies suggesting a significant role for *CCR5* SNPs, individually and as part of haplotypes, in determining HIV-1/AIDS risks for various population groups. These included the *CCR5*-2459A>G SNP associated with influencing disease progression to AIDS in Caucasians [McDermott et al., 1998, Clegg et al., 2000, Knudsen et al., 2001]. In addition a haplotype referred to as *CCR5*P1, which is in complete LD with the *CCR5*-2459 A allele, is associated with rapid disease progression, although the effect is recessive in Caucasians and dominant in

African Americans [Martin et al., 1998; Carrington et al., 1999; An et al., 2000]. Other *CCR5* haplotype associations with progression to AIDS have also been reported for African Americans [Gonzalez et al., 1999]. A case-control study of Africans from Uganda has shown that *CCR5* promoter SNPs and derived haplotypes are not associated with HIV-1/AIDS pathogenesis [Ramaley et al., 2002]. In the Xhosa ethnic group, we found the *CCR5*-2733 G allele (P = 0.0409) and the *CCR5*-2733 AG heterozygous genotype (P = 0.0270) to be associated with decreased risk for HIV-1 infection. In contrast, the *CCR5*-2135 T allele (P = 0.0024) and *CCR5*-2135 TT homozygous genotype (P = 0.0096) were associated with an increased risk for HIV-1 infection. It has been reported that *CCR5*-2135C>T is in strong LD with the *CCR5*-2459A>G influential AIDS SNP [Gonzalez et al., 1999], however no previously identified disease associations could be confirmed in our study sample. The *CCR2*V64I SNP frequency (0.156) in our study was consistent with previously reported African-based studies [Smith et al., 1997; Kostrikis et al., 1998; Mummidi et al., 1998; Ramaley et al 2002]. We found the *CCR2*V64I SNP to be in strong LD with *CCR5*-1835C>T, as observed previously [Kostrikis et al., 1998; Mummidi et al., 1998]. There are studies indicating significance for the *CCR2V*64I and *CCR5*-1835 T alleles with slower disease progression in African Americans [Mummidi et al., 1998, Gonzalez et al., 1999] and Kenyan Africans [Anzala et al., 1998]. This was not confirmed in other African ethnic groups [Petersen et al., 2002; Ramaley et al., 2002]. Our findings for the *CCR2-CCR5* derived haplotypes and association with HIV-1 infection risk lack significance, but are in agreement with the Ugandan study [Ramaley et al., 2002]. It has been suggested that the effects of *CCR5* and *CCR2* genetic variants are most

pronounced during early HIV-1 pathogenesis [Mulherin et al., 2003; Winkler et al., 2004]. Functional analysis has found a lower promoter activity for the *CCR5*-2459 G allele compared to the *CCR5*-2459 A allele [Mummidi et al., 1997] and differences in nuclear factor binding for the two alleles representing *CCR5*-2554G>T [Bream et al., 1999]. The functional effect reported for *CCR2*V64I involves an influence on CCR2A isoform stability and the resulting increased down-modulation of CCR5 by CCR2A [Nakayama et al., 2004].

The allele frequencies for the *CX3CR1*V249I (0.138) and *CX3CR1*T280M (0.024) SNPs are relatively low and similar to values reported for African Americans [Singh et al., 2005]. Caucasian-based studies have indicated that homozygosity for *CX3CR1*280M is associated with more rapid disease progression to AIDS [Faure et al., 2000; Faure et al., 2003]. Another study involving Caucasians however, showed conflicting findings as the *CX3CR1*280TM heterozygous genotype was associated with a delay in the rate of progression to AIDS, while the *CX3CR1*280M homozygous genotype had no effect [McDermott et al., 2000a]. Two additional studies did not confirm a significant association with the *CX3CR1*T280M SNP and rate of disease progression [Hendel et al., 2001; Kwa et al., 2003]. We identified an association with the *CX3CR1*249I allele and decreased risk for HIV-1 infection in the Xhosa (P = 0.0479). No other associations were present for both the genotype and haplotype analysis. Functional analysis of these SNPs showed a combined association with reduced ligand binding affinity and impaired HIV-1 co-receptor activity [Faure et al., 2000; McDermott et al., 2000a].

The *CCL5*-403G>A (0.409) and *CCL5 IVS1+307T>C* (0.222) SNP frequencies were comparable with previous reports for African Americans [An et al., 2002; Fernandez et al., 2003]. A known association with the *CCL5*-403 A allele includes an increased risk for HIV-1 infection [McDermott et al., 2000b; An et al., 2002], while another study showed resistance to AIDS in various populations [McDermott et al., 2000b]. The *CCL5 IVS1+307* C allele has been associated with adversely influencing HIV-1 susceptibility and also disease progression to AIDS, particularly in African-Americans [An et al. 2002, Wang et al., 2004]. Several studies in different population groups have also investigated and identified HIV-1/AIDS associations with the presence of specific *CCL5* SNP genotypes or haplotypes [Liu et al., 1999a, Liu et al., 1999b; McDermott et al., 2000b; Gonzalez et al., 2001; An et al., 2002; Duggal et al., 2005]. Functional effects reported for *CCL5* variants include the up and down-regulation of transcription by the *CCL5*-403 A allele and *CCL5 IVS1+307* C allele, respectively [Liu et al., 1999a; Nickel et al., 2000; An et al., 2002]. No significant associations with both *CCL5* SNPs and HIV-1 susceptibility were observed when considering the allele, genotype and haplotype distributions in the Xhosa ethnic group.

Based on previous HIV-1/AIDS association studies and our findings for *CCR5*, *CCR2*, *CX3CR1*, *CCL5* SNPs it has become evident that the distribution and effects of these genetic variants are specific for well-defined populations. The ongoing investigation of large case-control studies for many diverse ethnic groups is aimed at establishing HIV-1/AIDS risk profiles for individuals, particularly those living in regions of the world where the pandemic continues to

grow. Our study findings for the Xhosa from pandemic-stricken Sub-Saharan Africa therefore contribute to the quest for clarifying the effects of all known HIV-1/AIDS candidate gene variants in understudied African populations.

ACKNOWLEDGEMENTS

REFERENCES

- An P, Martin MP, Nelson GW, et al. Influence of CCR5 promoter haplotypes on AIDS progression in African-Americans. *AIDS* 2000; 14:2117-2122.

- An P, Nelson GW, Wang L, et al. Modulating influence on HIV/AIDS by interacting *RANTES* gene variants. *Proc Natl Acad Sci USA* 2002; 99:10002-10007.

- Anastassopoulou CG and Kostrikis LG. The impact of human allelic variation on HIV-1 disease. *Curr HIV Res* 2003; 1:185-203.

- Anzala AO, Ball TB, Rostron T, O'Brien SJ, Plummer FA, Rowland-Jones SL. CCR2-64I allele and genotype association with delayed AIDS progression on African women. *Lancet* 1998, 351:1632-1633.

- Barret JC, Fry B, Maller J and Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; 21:263-265.

- Bream JH, Young HA, Rice N, et al. *CCR5* promoter alleles and specific DNA binding factors. *Science* 1999; 284:223a.

- Carrington M, Dean M, Martin MP and O'Brien SJ. Genetics of HIV-1 infection: chemokine receptor CCR5 polymorphism and its consequences. *Hum Mol Genet* 1999; 8:1939-1945.

- Carrington M, Nelson G and O'Brien SJ. Considering genetic profiles in functional studies of immune responsiveness to HIV-1. *Immunol Lett* 2001; 79:131-40.

- Clegg AO, Ashton LJ, Biti RA, et al. *CCR5* promoter polymorphisms, *CCR5* 59029A and *CCR5* 59353C are under represented in HIV-1 infected long-term non-progressors. *AIDS* 2000; 14:103-108.

- Dean M, Carrington M and O'Brien SJ. Balanced polymorphism selected by genetic versus infectious human disease*. Annu Rev Genomics Hum Genet* 2002; 3:263-292.

- Duggal P, Winkler CA, An P, et al. The effect of RANTES chemokine genetic variants on early HIV-1 plasma RNA among African American injection drug users *J Acquir Immune Defic Syndr* 2005; 38:584-589.

- Faure S, Meyer L, Costagliola D, et al. Rapid progression to AIDS in HIV+ individuals with a structural variant of the chemokine receptor CX$_3$CR1. *Science* 2000; 287:2274-2277.

- Faure S, Meyer L, Genin E, et al. Deleterious genetic influence of CX3CR1 genotypes on HIV-1 disease progression. *J Acquir Immune Defic Syndr* 2003; 32:335-337.

- Fernandez RM, Borrego S, Marcos I, Rubio A, Lissen E and Antinolo G. Fluorescence resonance energy transfer analysis of the RANTES polymorphisms -403G --> A and -28G --> C: evaluation of both variants as susceptibility factors to HIV type 1 infection in the Spanish population. *AIDS Res Hum Retroviruses* 2003; 19:349-352.

- Gonzalez E, Bamshad M, Sato N, et al. Race-specific HIV-1 disease modifying effects associated with CCR5 haplotypes. *Proc Natl Acad Sci USA* 1999; 96:12004-12009.

- Gonzalez E, Dhanda R, Bamshad M, et al. Global survey of genetic variation in CCR5, RANTES, and MIP-1α: Impact on the epidemiology of the HIV-1 pandemic. *Proc Natl Acad Sci USA* 2001; 98:5199-5204.

- Hendel H, Winkler C, An P et al. Validation of genetic case-control studies in AIDS and application to the CX3CR1 polymorphism. *J Acquir Immune Defic Syndr* 2001; 26:507-11.

- Hogan CM and Hammer SM. Host determinants in HIV infection and disease, Part 2: Genetic factors and implications for antiretroviral therapeutics. *Ann Intern Med* 2001; 134:978-996.

- John GC, Bird T, Overbaugh J, et al. CCR5 promoter polymorphisms in a Kenyan perinatal human immunodeficiency virus type 1 cohort: Association with increased 2-year maternal mortality. *J Infect Dis* 2001; 184:89-92.

- Kaslow RA, Dorak T and Tang J. Influence of host genetic variation on susceptibility to HIV type 1 infection. *J Infect Dis* 2005; 191(Suppl 1):S68-77.

- Knudsen TB, Kristiansen TB, Katzenstein TL and Eugen-Olsen J. Adverse effect of the CCR5 promoter –2459A allele on HIV-1 disease progression. *J Med Virol* 2001; 65:441-444.

- Kostrikis LG, Huang Y, Moore JP, et al. A chemokine receptor *CCR2* allele delays HIV-1 disease progression and is associated with a *CCR5* promoter mutation. *Nature Med* 1998; 4:350-353.

- Kwa D, Boeser-Nunnink B, Schuitemaker H. Lack of evidence for an association between a polymorphism in CX3CR1 and the clinical course of HIV infection or virus phenotype evolution. *AIDS* 2003; 17:759-761.

- Liu H, Chao D, Nakayama EE, et al. Polymorphism in RANTES chemokine promoter affects HIV-1 disease progression. *Proc Natl Acad Sci USA* 1999a; 96:4581-4585.

- Liu H, Shioda T, Nagai Y, et al. Distribution of HIV-1 disease modifying regulated on activation normal T cell expressed and secreted haplotypes in Asian, African and Caucasian individuals. *AIDS* 1999b; 13:2602-2603.

- Livak KJ. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genet Anal* 1999; 14:143-149.

- Martin M, Dean M, Smith MW, et al. Genetic acceleration of AIDS progression by a promoter variant of *CCR5*. *Science* 1998; 282: 1907-1911.

- McDermottt DH, Zimmerman PA, Guignard F, Kleeberger CA, Leitman SF and Murphy PM. CCR5 promoter polymorphism and HIV-1 disease progression. Multicenter AIDS Cohort Study (MACS). *Lancet* 1998; 352:866-870.

- McDermott DH, Colla JS, Kleeberger CA et al. Genetic polymorphism in CX$_3$CR1 and risk of HIV disease. *Science* 2000a; 290:2031a.

- McDermott DH, Beecroft MJ, Kleeberger CA, et al. Chemokine RANTES promoter polymorphism affects risk of both HIV infection and disease progression in the Multicenter AIDS Cohort Study. *AIDS* 2000b; 14:2671-2678.

- Mulherin SA, O'Brien TR, Ioannidis JPA, et al. Effects of *CCR5-Δ32* and CCR264I alleles on HIV-1 disease progression: the protection varies with duration of infection. *AIDS* 2003, 17:377-387.

- Mummidi S, Ahuja SS, McDaniel BL and Ahuja SK. The human CC chemokine receptor 5 (CCR5) gene. Multiple transcripts with 5'-end heterogeneity, dual promoter usage, and evidence for polymorphisms within the regulatory regions and noncoding exons. *J Biol Chem* 1997; 272:30662-30671.

- Mummidi S, Ahuja SS, Gonzalez E, et al. Genealogy of the *CCR5* locus and chemokine system gene variants associated with altered rates of HIV-1 disease progression. *Nature Med* 1998; 4:786-793.

- Nakayama EE, Tanaka Y, Nagai Y, Iwamoto A and Shioda T. A *CCR2*-V64I polymorphism affects stability of CCR2A isoform. *AIDS* 2004; 18:729-738.

- Nickel RG, Casolaro V, Wahn U, et al. Atopic dermatitis is associated with a functional mutation in the promoter of the C-C chemokine RANTES. *J Immunol* 2000; 164:1612-1616.

- O'Brien SJ and Nelson GW. Human genes that limit AIDS. *Nature Genet* 2004; 36: 565-574.

- Petersen DC, Kotze MJ, Zeier MD, et al. Novel mutations identified using a comprehensive CCR5-denaturing gradient gel electrophoresis assay. *AIDS* 2001; 15:171-177.

- Petersen DC, Laten A, Zeier MD, Grimwood A, Janse van Rensburg E and Hayes VM. Novel Mutations and SNPs identified in *CCR2* using a new comprehensive denaturing gradient gel electrophoresis assay. *Hum Mut* 2002; 20:253-259.

- Ramaley PA, French N, Kaleebu P, et al. Chemokine-receptor genes and AIDS risk. *Nature* 2002; 417:140.

- Singh KK, Hughes MD, Chen J, Spector SA. Genetic polymorphisms in CX_3CR1 predict HIV-1 disease progression in children independently of CD4+ lymphocyte count and HIV-1 RNA load. *J Infect Dis* 2005; 191:1971-1980.

- Smith M, Dean M, Carrington M, et al. Contrasting genetic influence of *CCR2* and *CCR5* variants on HIV-1 infection and disease progression. *Science* 1997; 277:959-965.

- Wang C, Song W, Lobashevsky E, et al. Cytokine and chemokine gene polymorphisms among ethnically diverse North Americans with HIV-1 infection. *J Acquir Immune Defic Syndr* 2004; 35:446-54.

- Winkler C, An P and O'Brien SJ. Patterns of ethnic diversity among the genes that influence AIDS. *Hum Mol Genet* 2004; 1:R9-19.

- Winkler CA, Hendel H, Carrington M, et al. Dominant effects of *CCR2-CCR5* haplotypes in HIV-1 disease progression. *J Acquir Immune Defic Syndr* 2004; 37:1534-1538.

Table 1. Primers and probes for the TaqMan allelic discrimination assays

Gene – SNP	NCBI ID	Primers (5'- 3')	Probes (5'- 3')	T_M (^{0}C)
CCR5-2733 (A/G)	rs2856758	TCATGTGGAAAATTTCTCATAGCTTCAGA TCACACTATGCCAGATACGTAGGT	AGTGAAGAATCCTGCC AGTGAAGGATCCTGCC	59
CCR5-2554 (G/T)	rs2734648	CCGTGAGCCCATAGTTAAAACTCTT CACAGATGCTCACCACCCAATATTA	ACAACAGGTTTTTTCCGT CAACAGGTTGTTTCCGT	59
CCR5-2135 (T/C)	rs1799988	GGGATGAGCAGAGAACAAAAACAAA TGTATTGAAGGCGAAAAGAATCAGAGA	CCCGTAAATAAACCTT CCCGTAAATAAACTTT	59
CCR5-2086 (A/G)	rs1800023	GGGATGAGCAGAGAACAAAAACAAA TGTATTGAAGGCGAAAAGAATCAGAGA	CAACTTAAAAGGAAGAAC CTCAACTTAAAAAGAAGAAC	59
CCR5-1835 (C/T)	rs1800024	CCTGTTAGTTAGCTTCTGAGATGAGTAAA CCAAACTGTGACCCTTTCCTTATCT	TTTGCCAAATATCTTCT TTTGCCAAATGTCTTCT	59
*CCR2*V64I (G/A)	rs1799864	CCGCTCTACTCGCTGGTGTT AAATGTCAGTCAAGCACTTCAGCT	CAACATGCTGGTCGTCCTCATCTTAATAA CAACATGCTGGTCATCCTCATCTTAATAAACT	59
*CX3CR1*V249I(G/A)	rs3732379	TGGTCATCGTGTTTTTCCTCTT GGGAAAGAAGTCATAGAGCTTAAGC	ACACCCTACAACGTTATGATTTTCC ACACCCTACAACATTATGATTTTCCT	60
*CX3CR1*T280M(C/T)	rs3732378	CCCAGCAAATGCATAGATGA TCCCAGTTGTGACATGAGGA	TAAATGCAACCGTCTCAGTCACACT TAAATGCAACCATCTCAGTCACACTG	60
CCL5-403(A/G)	rs2107538	TCCAGAGGACCCTCCTCAATAA CTGAGTCACTGAGTCTTCAAAGTTCC	AAAGGAGGTAAGATCTGTAAT AAAGGAGATAAGATCTGTAATG	60
CCL5 IVS1+307(T/C)	rs2280789	TGCTTCATGGCAGGGATCTC GTGAACACCTGTAGGCCTTGAG	CTGTCTTCAAGGTCTAC TTTTTCTGTCTTTAAGGTCTAC	59

Table 2. Allelic and genotypic distribution of *CCR5*, *CCR2*, *CX3CR1* and *CCL5* SNPs in cases versus controls

Gene[a] SNP	NCBI ID	Minor Allele Frequency			Genotype Frequency [b]						
		HIV+	HIV-	P	-/-		+/-		+/+		
					HIV+	HIV-	HIV+	HIV-	HIV+	HIV-	P
CCR5											
-2733 (A/G)	rs2856758	0.031	0.067	**0.0409***	0	0	0.061	0.134	0.939	0.866	**0.0270***
-2554 (G/T)	rs2734648	0.329	0.292	0.3743	0.120	0.115	0.418	0.354	0.462	0.531	0.4706
-2135 (T/C)	rs1799988	0.374	0.500	**0.0024***	0.150	0.243	0.449	0.514	0.401	0.243	**0.0096***
-2086 (A/G)	rs1800023	0.174	0.131	0.1747	0.028	0.027	0.291	0.207	0.681	0.766	0.2588
-1835 (C/T)	rs1800024	0.162	0.203	0.2284	0.039	0.036	0.245	0.333	0.716	0.631	0.2461
CCR2											
V64I (G/A)	rs1799864	0.140	0.173	0.2963	0.024	0.036	0.232	0.273	0.744	0.691	0.5551
CX3CR1											
V249I(G/A)	rs3732379	0.109	0.167	**0.0479***	0.009	0.018	0.200	0.297	0.791	0.685	0.1045
T280M(C/T)	rs3732378	0.021	0.027	0.5930	0	0	0.0421	0.055	0.958	0.946	0.6124
CCL5											
-403(A/G)	rs2107538	0.412	0.405	0.8641	0.190	0.133	0.446	0.543	0.365	0.324	0.2203
IVS1+307(T/C)	rs2280789	0.221	0.223	1	0.056	0.055	0.329	0.336	0.615	0.609	0.9894

HIV+, seropositive; HIV -, seronegative
[a]The second base pair indicates the minor allele
[b](-/-), (+/-) and (+/+) represents the minor allele homozygotes, heterozygotes and major allele homozygotes, respectively.
* Indicates significant *P* values
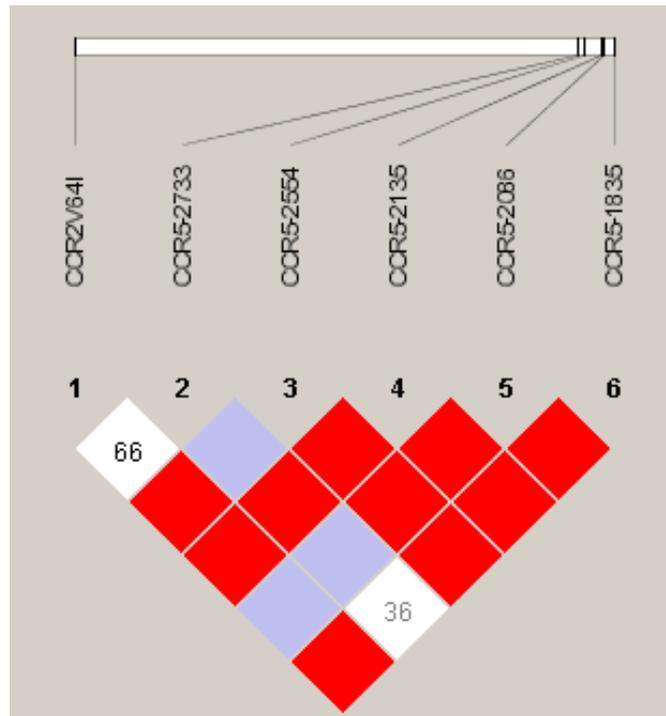
Figure 1. Pairwise linkage disequilibrium based on D' measures for the SNPs in the *CCR2-CCR5* region created using Haploview. Red: D' = 1 (LOD ≥ 2); Blue: D' = 1 (LOD < 2); White: D' < 1 (LOD < 2).

Chapter 2.2.

Risk for HIV-1 infection associated with a common *CXCL12* (*SDF1*) polymorphism and *CXCR4* variation in an African population

Desiree C. Petersen[1], Richard H. Glashoff[1], Sadeep Shrestha[3,4],

Julie Bergeron[4], Annette Laten[1], Bert Gold[3],

Estrelita Janse van Rensburg[1], Michael Dean[3]

and Vanessa M. Hayes [2,5]

Departments of Medical Virology[1] and Urology[2], University of Stellenbosch, Tygerberg Medical School, South Africa; Laboratory of Genomic Diversity[3], National Cancer Institute, Frederick, Maryland, USA; Basic Research Program[4], SAIC-Frederick, Maryland USA; Cancer Research Program[5] Garvan Institute of Medical Research, Darlinghurst, Sydney, NSW, Australia.

(Full article)

(References cited in the format of the journal where this manuscript is accepted for publication)

ABSTRACT

CXC chemokine ligand 12 (CXCL12), or stromal cell-derived factor 1 (SDF1), is the only known natural ligand for the human immunodeficiency virus-1 (HIV-1) co-receptor, CXC chemokine receptor 4 (CXCR4). A single nucleotide polymorphism (SNP) in the *CXCL12* gene (*SDF1*-3'A) has been associated with disease progression to acquired immune deficiency syndrome (AIDS) in some studies, but not others. Mutations in the *CXCR4* gene are generally rare and have not been implicated in HIV-1/AIDS pathogenesis. In this study, we analysed the *SDF1*-3'A SNP and performed mutation screening for polymorphic markers in the *CXCR4* gene to determine the presence/absence of significant associations with susceptibility to HIV-1 infection. Our study consisted of 257 HIV-1 seropositive patients and 113 HIV-1 seronegative controls representing a Sub-Saharan African population belonging to the Xhosa ethnic group of South Africa. The *SDF1*-3'A SNP was associated with an increased risk for HIV-1 infection ($P = 0.0319$) while no significant association between the occurrence of the *SDF1*-3'A SNP and increased/decreased plasma levels of CXCL12 was observed. Comprehensive mutation analysis of the *CXCR4* gene confirmed a high degree of genetic conservation within the coding region of this ancient population.

Keywords: CXC chemokine ligand 12 (CXCL12); CXC chemokine receptor 4 (CXCR4); *SDF1*-3'A SNP; HIV-1 infection risk; African population.

INTRODUCTION

The role of specific chemokines acting as inhibitors of human immunodeficiency virus-1 (HIV-1) infection and also possibly influencing viral replication[1] was highlighted by the discovery that chemokines are natural ligands for chemokine receptors. These chemokine receptors, together with the CD4 molecule, serve as necessary co-factors for HIV-1 entry.[2-5] The CXC chemokine ligand 12 (CXCL12), known more commonly as stromal cell-derived factor 1 (SDF1), inhibits infection of T cell line tropic (T-tropic) or syncytium-inducing (SI) viruses normally found during late-stage HIV disease[6,7] by down-regulating the surface expression of the HIV-1 co-receptor, CXC chemokine receptor 4 (CXCR4).[8-9] The demonstration that mice, deficient for either CXCL12 or CXCR4, die perinatally [10] further promoted the understanding of this ligand-receptor interaction, which appears to be vital in physiological processes.

It has been found that the *CXCL12* gene, previously known as *SDF1* or *PBSF* (MIM# 600835), is located at band q11 on chromosome 10 and encodes for two isoforms, CXCL12α and CXCL12β, which are the result of alternative splicing of a single gene.[11-13] The coding regions for *CXCL12α* and *CXCL12β* are composed of three and four exons, respectively. The *CXCL12β* gene transcript has an extra exon, which encodes for four additional amino acids.[13] The *CXCR4* gene, also known as *NPY3R*, *FUSIN* and *LESTR* (MIM# 162643)*,* is located at band q21 on chromosome 2[14,15] and consists of an intron separating 2 exons in which lies the open reading frame (ORF).[16-17]

A single nucleotide polymorphism (SNP), designated *SDF1*-3'UTR-801G>A and abbreviated *SDF1*-3'A (rs1801157), was identified in the 3' untranslated region (3'UTR) of the *CXCL12β* gene transcript and involves a G to A transition at nucleotide position +801 relative to the start codon.[18] Although the SNP in the recessive state was initially associated with delayed onset of AIDS,[18,19] other studies suggested an association with accelerated progression to death;[20,21] prolonged[21] or decreased[22] survival after AIDS diagnosis; or no effect on disease progression.[23-25] Another study showed an association between the SNP in the heterozygous state and increased vertical transmission from mother to child in an African study,[26] while an association with rapid disease progression and the SNP occurring heterozygously was observed in HIV-1 infected children born to seropositive mothers.[27] More recently, the SNP has been found to play a role in resistance to HIV-1 infection in seronegative high-risk individuals,[28] although this association was absent in a study involving repeatedly exposed HIV-1 seronegatives.[29] Studies investigating plasma CXCL12 protein levels in HIV-1 seropositive patients; exposed high-risk HIV-1 seronegative individuals and healthy HIV-1 seronegative controls[28,30-34] with consideration of *SDF1*-3'A genotypes[32,33] have also reported inconsistent associations.

Previous studies investigating the role of *CXCR4* in host susceptibility to HIV-1/AIDS in Caucasian-based and African American populations have shown a relatively low occurrence of *CXCR4* mutations and therefore their significance is unclear.[35-37] The *CXCR4* genetic variants reported include silent mutations *CXCR4*-I261I[35] and *CXCR4*-K68K; and a non-conservative mutation, *CXCR4*-

F93S.[36] Both the *CXCR4*-K68K and *CXCR4*-F93S mutations were further considered for their possible influence on HIV-1 entry, with the results being comparable to what was found for wildtype CXCR4.[36] More recently, mutations in the cytoplasmic tail domain of *CXCR4* were identified as being causative for WHIM syndrome, an immunodeficiency disorder characterised by warts, hypogammaglobulinemia, infections and myelokathexis.[38] These mutations are however familial and rare.

Controversy with regards to the role of *CXCL12 (SDF1*-3'A SNP) and *CXCR4* mutations in HIV-1/AIDS pathogenesis has accentuated the need for additional studies within ethnically distinct populations. In this study we genotyped the *SDF1*-3'A SNP and performed comprehensive mutational analysis of the *CXCR4* coding region. Plasma CXCL12 levels were measured to assess possible functional correlation between the *SDF1*-3'A SNP and protein levels. Our results indicate the importance of investigating the genetic basis for HIV-1/AIDS within specific ethnic groups, particularly populations from understudied pandemic-stricken Sub-Saharan Africa.

MATERIALS AND METHODS

Study population

The population group represented in this study are Sub-Saharan Africans defined as individuals of Xhosa descent all residing in the Western Cape Province of South Africa. According to recent consensus, the South African population is 79% African, with the Xhosa ethnic group forming approximately

22% of the total African population and 90% of the African population residing in the Western Cape (Statistics South Africa, 2001; www.statssa.gov.za). The Xhosa are from the early clan of the Nguni, the most southern group of Bantu migrants from central Africa. The HIV-1 seropositive (HIV+) individuals were patients of Tygerberg Hospital, Woodstock Chapel Street Community Health Clinic or the Langa Clinic, which at the time of DNA extraction totalled 1035 ethnically diverse patients, as previously described.[39] Individuals were included in this study if they were from Xhosa descent and blood was available for DNA extraction. The HIV-1 seronegative (HIV-) controls were population-matched blood donors from the Western Province Blood Transfusion Service. The only criteria for exclusion, was a positive HIV-1 status. Disease progression for most of the HIV+ patients remains unknown and was therefore not assessed in this study. Informed consent was obtained from all the study participants and the Ethics Review Committee of the University of Stellenbosch approved the study protocol (#98/158). The sample size for genotyping the *SDF1*-3'A SNP consisted of 257 HIV+ (66% female, 34% male) and 113 HIV- (62% female, 38% male) individuals. CXCL12 protein levels were determined for samples where plasma was available (131 HIV+, 63 HIV-). Comprehensive mutational analysis of the entire coding region of *CXCR4* gene was performed on 57 HIV+ and 39 HIV- Xhosas. An additional 30 HIV+ and 22 HIV- samples were further screened to determine allele frequencies of identified mutations.

Genetic analysis

Genotyping: Genomic DNA was extracted from whole blood and genotyped in a blinded manner. Two methods were used to genotype the *SDF1-3'A* SNP

due to availability of technologies at the time. To confirm genotyping specificity of the two methods, 155 samples were genotyped using both methods. The first utilised denaturing gradient gel electrophoresis (DGGE).[40] A single DGGE primer set, including a GC-clamp (GC-rich fragment) on the 5'end of the reverse primer, was designed for partial analysis of the 3'UTR of the *CXCL12β* gene transcript: 5'-GTGAAGGCTTCTCTCTGTGG-3' and 5'-[40GC]GTGGACACACA TGATGATGG-3'. Amplification and heteroduplexing was performed as described below (mutation detection) using an annealing temperature of 56^0C. Amplified PCR products were electrophoresed in a 9% polyacrylamide gel with a denaturing gradient of 45% to 85% urea and formamide (100% UF = 7mol/L urea per 40% deionised formamide), at 60^oC for 110 volts overnight, using the Ingeny phorU-2 system (INGENY, Goes, The Netherlands, www.ingeny.com). The 5' nuclease or TaqMan allelic discrimination method[41] was also used for genotyping the *SDF1*-3'A SNP using the following primers: 5'-CAAAGCCTAGTGAAGGCTTCTCTC-3' and 5'-TCAGGGTAGCCCTGCTGC-3'; and probes: 5'-FAM-TGGGAGCC**G**GGTCTGC CTCT-TAMRA-3' and 5'-VIC-ACATGGGAGCC**A**GGTCTGCCTCTT-TAMRA-3'. PCR reactions each containing 5 ng of DNA and TaqMan Universal PCR Master Mix (Applied Biosysytems) were used for amplification (detailed PCR reaction mix protocol available on request) with the cycling conditions including a initial denaturation of 95^oC for 10 minutes, followed by 50 cycles of denaturation at 95^oC for 15 seconds and annealing at 58^oC for 1 minute. Allele discrimination and genotype determination was based on the end-point fluorescence measured by the ABI7900 high throughput sequence detection system (Applied

Biosysytems). A total of 199 and 215 samples were genotyped using DGGE and TaqMan allelic discrimination, respectively.

Mutation Detection: DGGE primers were designed for the entire coding region, including the intron/exon boundaries of the *CXCR4* gene. The *CXCR4* coding region within exon 1 (codons 1 to 5) and exon 2 (codons 6 to 352) was divided into seven overlapping amplicons (A – G) (Table 1). Each PCR reaction contained 50 ng of DNA (detailed PCR reaction mixture protocol available on request) and amplification was performed using a 9600 thermocycler (Applied Biosysytems). PCR cycling conditions included: an initial denaturation at 96°C for 3 minutes, followed by 32 cycles of denaturation at 96°C for 45 seconds, annealing for 1 minute (annealing temperatures are shown in Table 1) and elongation at 72°C for 1 minute 20 seconds. Following the last cycle was an additional extension step of 72°C for 7 minutes. For heteroduplex formation, the PCR products were subjected to denaturation at 96°C for 10 minutes, followed by renaturation for 45 minutes at 56°C. Electrophoresis was used to check the amplified products, where 10% of each sample was resolved on 2% agarose gel. Optimal DGGE analysis was achieved using previously described conditions for broad-range mutation detection by DGGE.[42] Amplified PCR products were electrophoresed in a 9% polyacrylamide gel with a denaturing gradient of 30% to 75% UF, at 60°C for 110 volts overnight. The seven amplicons for *CXCR4* were electrophoresed in five lanes (Fragments B and E; Fragments C and G were pooled) and allowed for the complete analysis of 6 patients per denaturing gel. The gels were stained with ethidium bromide and photographed under an UV transilluminator. Samples showing aberrant DGGE

banding patterns were purified using the high pure PCR product purification kit of Roche (Roche Diagnostics, Mannheim, Germany) and subjected to automated sequencing using the non-GC-clamped primer and the dye terminator sequencing kit (Applied Biosystems, Foster City, U.S.A., www.appliedbiosystems.com).

Determining plasma CXCL12 protein levels by ELISA

Plasma was isolated by centrifugation of EDTA-anticoagulated blood samples at 2000 rpm for 10 minutes. Plasma samples were stored at -80°C before being thawed for analysis. A CXCL12 enzyme-linked immunosorbent assay (ELISA) was developed using commercial monoclonal antibodies (R&D systems, Minneapolis, Minnesota, USA) according to manufacturer's recommendations. Flat-bottom 96-well microtiter plates with high-binding capacity (Nunc Maxisorp, Nunc, Denmark) were coated with capture antibody (mouse anti-human CXCL12) prior to the addition of plasma samples. Biotinylated mouse anti-human CXCL12β was used as detection antibody. Recombinant human CXCL12β was included as a standard. Each sample was run in duplicate and the mean concentration (pg/ml) of plasma CXCL12 protein was determined from the standard curve using ELISA software (Bio-Tek KC4, Bio-Tek Instruments, Winooski, Vermont, USA).

Statistical analysis

The allele and genotype distributions, including Hardy Weinberg equilibrium estimations were calculated. Testing of HIV-1 seropositives versus HIV-1 seronegatives for significance of heterogeneity in allele and genotype

frequencies was based on the two-sided Fisher's exact test for 2x2 contingency tables and the chi-squared (x^2) test for independence, respectively. The Mann Whitney U test was used to compare mean plasma CXCL12 protein levels between the case and control groups (GraphPad Software Inc, San Diego, and SAS Institute Inc., Cary, North Carolina, U.S.A.).

RESULTS

Analysis of the *SDF1*-3'A SNP

The commonly reported *SDF1*-3'A SNP was detected using gel-based DGGE and a TaqMan allelic discrimination assay (Table 2). The 100% concordance observed for 155 samples screened with both methods reflects the reliability of the two *SDF1*-3'A SNP assays for generating valid and reproducible results. The presence of the SNP in the Xhosa population was observed at a significantly higher allelic frequency in the HIV-1 seropositive patients (0.037; 19/514) compared to their uninfected counterparts (0.009; 2/226) with a *P* value of 0.0319. No significance was found for the independent genotype analysis (*P* = 0.1191), however, a significant association between the presence of the A allele (AA and GA) and HIV-1 infection was observed (*P* = 0.0454). There was no significant deviation from the expected Hardy-Weinberg equilibrium in either the cases or controls. There was no significant associations between specific CXCL12 plasma levels between *SDF1*-3'A genotype among cases and controls.

Analysis of *CXCR4* and identification of mutations

We identified one previously known and three novel mutations in *CXCR4* gene. The previously found silent mutation occurs in the coding region at codon 138 (rs2228014). The *CXCR4*-I138I mutation was detected using the amplicon C primers set and was observed in one HIV-1 seropositive patient and one HIV-1 seronegative control. The three novel mutations found in the 3'UTR at nucleotide positions +29 (G-A), +34 (A-T) and +46 (deletion T), relative to the stop codon, were all identified using the primer set for amplicon G. Further screening in additional samples resulted in the novel mutations occurring at allele frequencies ranging from 0.008 to 0.011 in the Xhosa population. No associations with susceptibility to HIV-1 infection or disease progression to AIDS were found.

DISCUSSION

Our study focused on the analysis of the *SDF1*-3'A SNP and the *CXCR4* gene within the Xhosa ethnic group from South Africa. The importance of determining population-specific genetic variants influencing HIV-1 susceptibility in the understudied African populations is evident.[43] A previous study by Ramaley *et al.*, 2003 suggested that caution should be taken when considering an association observed in one population to be present in another and showed that the *CCR5* alleles previously identified and significantly associated with influencing susceptibility to HIV-1/AIDS in Caucasian populations did not have the same effect in Africans.[44] We have previously screened for the well-documented *CCR5*-Δ32 HIV-1/AIDS resistance-associated mutation and found it to be completely absent in the Xhosa population.[45] Although genetic markers

may vary in frequency across populations, a recent study has suggested that their biological impact on the risk for the disease may usually be consistent across traditional "racial" boundaries.[46] In the ancient Xhosa population, these functional genetic markers may as yet, not have been identified in the majority of studies that have focused on younger populations.

Although the allele frequency observed for the *SDF1*-3'A SNP in our Xhosa group (0.028) is within the range previously reported for African populations, [18,20,47,48] it is slightly higher than that reported for another South African study (0.010).[47] The study participants reported in Williamson *et al.*, resided in the Free State Province of South Africa and are predominantly of Sotho ethnic descent, which could explain this bias in allele frequencies.

The *SDF1*-3'A SNP analysis resulted in an association being observed between the presence of the A allele and an increased rate of infection both at the level of allele frequency ($P = 0.0319$) and A allele carriage frequency ($P = 0.0454$), although significance was not found for independent genotype analysis. Previous studies, focusing on predominantly Caucasian-based populations have not reported similar findings, but rather found associations with disease progression to AIDS. An association with the *SDF1*-3'A SNP occurring heterozygously and increased vertical transmission from mother to child was however previously reported in Africans.[26] Controversies in genetic association studies have been significantly addressed in the literature, with the most compelling short-coming of this study being the relatively small study numbers

which could result in possible bias. Therefore, these findings require independent elucidation.

Plasma CXCL12 protein levels were not significantly different between the wild-type and the *SDF1*-3'A SNP within cases and controls. Due to the relatively small sample number (as well as lack of homozygous individuals), in the study population, we determined the CXCL12 protein levels in an additional 48 HIV seronegative (27 wild-type, 15 heterozygous and 6 AA homozygous) Caucasian samples (data not shown), but this also resulted in no significant association. Our findings require further investigation in larger sample numbers and replication in other cohorts supporting the effect of the SNP in ligand expression. This may however be limited by the low occurrence of homozygotes in the Xhosa population. An advantage of our study was the utilisation of the recombinant human CXCL12β for the ELISA assay as opposed to recombinant human CXCL12α used in other studies where associations were reported. The latter studies included low plasma CXCL12 levels found in uninfected persons homozygous for the *SDF1*-3'A SNP[33] and a significant increase in CXCL12 levels being observed in the HIV-1 seropositive individuals when compared to the HIV-1 seronegative control group.[34]

The identification of three novel genetic variants within the 3' UTR of *CXCR4*, a potential regulatory region, could have an effect on the expression or functioning of the protein. CXCR4 serves as the co-receptor for T-tropic or SI viruses normally emerging during late-stage HIV disease[5] and thus a possible effect of mutations occurring in *CXCR4* is more likely to be seen if they

influence disease progression rather than susceptibility to HIV-1 infection. We screened an additional 51 Caucasian controls (data not shown) which excluded the presence of the three novel *CXCR4* mutations. Therefore, although no associations could be made, these novel variants being undetected in the Caucasian population suggests that these mutations may be African-based.

Controversy exists in the literature between the association of the *SDF1*-3'A SNP and HIV-1 infection and/or disease progression to AIDS. It was recently reported that other polymorphisms in linkage disequilibrium with the *SDF1*-3'A SNP, rather than the SNP itself, are responsible for altered levels of *SDF1* transcripts. The inconsistent findings for the *SDF1*-3'A SNP amongst various populations may therefore be attributed to different haplotype structures, including or excluding functional variants, for specific ethnic groups.[49] In this study, we found an association between the presence of the *SDF1*-3'A SNP and risk to HIV-1 infection in a Sub-Saharan African population. Our results emphasise the need for investigating HIV-1/AIDS candidate genes in many diverse ethnic groups, and particularly in the populations most affected by the HIV-1/AIDS pandemic. Although this study only focused on a relatively small number of individuals, its findings contribute to the growing evidence that the presence and effects of genetic variants in the understudied African populations are important when predicting host susceptibility to HIV-1/AIDS within Sub-Saharan Africa.

ACKNOWLEDGEMENTS

REFERENCES

1. Cocchi F, De Vico A, Garzino-Demo A, et al. Identification of RANTES, MIP-1α, and MIP-1β as the major HIV suppressive factors produced by CD8+ T cells. *Science* 1995;270:1811-5.

2. Deng H, Liu R, Ellmeier W, et al. Identification of a major co-receptor for primary isolates of HIV-1. *Nature* 1996;381:661-6.

3. Dragic T, Litwin V, Allaway GP, et al. HIV-1 entry into CD4$^+$ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* 1996;381:667-73.

4. Doranz BJ, Rucker J, Yi Y, et al. A dual-tropic primary HIV-1 isolate that uses fusin and the β-chemokine receptors CKR-5, CKR-3, and CKR-2 as fusin cofactors. *Cell* 1996;85:1149-58.

5. Feng Y, Broder CC, Kennedy PE, et al. HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 1996;272:827-7.

6. Bleul CC, Farzan M, Choe H, et al. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusion and blocks HIV-1 entry. *Nature* 1996;382:829-33.

7. Oberlin E, Amara A, Bachelerie F, et al. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature* 1996;382:833-5.

8. Amara A, Gall SL, Schwartz O, et al. HIV coreceptor downregulation as antiviral principle: SDF-1 alpha-dependant internalisation of the chemokine receptor CXCR4 contributes to inhibition of HIV replication. *J Exp Med* 1997;186:139-46.

9. Signoret N, Oldridge J, Pelchen-Matthews A, et al. Phorbol esters and SDF-1 induce rapid endocytosis and down modulation of the chemokine receptor CXCR4. *J Cell Biol* 1997;139:651-64.

10. Ma Q, Jones D, Borghesani PR, et al. Impaired B-lymphopoiesis, myelopoiesis, and derailed cerebellar neuron migration in CXCR4 and SDF-1-deficient mice. *Proc Natl Acad Sci USA* 1998;95:9448-53.

11. Tashiro K, Tada H, Heilker R, et al. Signal sequence trap: A cloning strategy for secreted protein and type I membrane proteins. *Science* 1993;261:600-3.

12. Nagasawa T, Kikutani H, Kishimoto T. Molecular cloning and structure of a pre-B-cell growth-stimulating factor. *Proc Natl Acad Sci USA* 1994;91:2305-9.

13. Shirozu M, Nakano T, Inazawa J, et al. Structure and chromosomal localization of the human stromal cell-derived factor (SDF1) gene. *Genomics* 1995;28:495-500.

14. Herzog H, Hort YJ, Shine J, et al. Molecular cloning, characterization, and localization of the human homolog to the reported bovine NPY Y3 receptor: lack of NPY binding and activation. *DNA Cell Biol* 1993;12:465-71.

15. Federsppiel B, Melhado IG, Duncan AMV, et al. Molecular cloning of the cDNA and chromosomal localization of the gene for a putative seven-transmembrane segment (7-TMS) receptor isolated from human spleen. *Genomics* 1993;16:707-12.

16. Wegner SA, Ehrenberg PK, Chang G, et al. Genomic organization and functional characterization of the chemokine receptor CXCR4, a major entry co-receptor for human immunodeficiency virus type 1. *J Biol Chem* 1998;273:4754-60.

17. Caruz A, Samson M, Alonso JM, et al. Genomic organization and promoter characterization of human CXCR4 gene. *FEBS Lett* 1998;426:271-8.

18. Winkler C, Modi W, Smith MW, et al. Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. *Science* 1998;279:389-93.

19. Hendel H, Henon N, Lebuanec H, et al. Distinctive effects of *CCR5, CCR2,* and *SDF1* genetic polymorphisms in AIDS progression. *J Acquir Immune Defic Syndr Hum Retrovirol* 1998;19:381-6.

20. Mummidi S, Ahuja SS, Gonzalez SA, et al. Genealogy of the *CCR5* locus and chemokine system gene variants associated with altered rates of HIV-1 disease progression. *Nat Med* 1998;4:786-93.

21. van Rij RP, Broersen S, Goudsmit J et al. The role of a stromal cell-derived factor-1 chemokine gene variant in the clinical course of HIV-1 infection. *AIDS* 1998;12:F85-F90.

22. Brambilla A, Villa C, Rizzardi G et al. Shorter survival of *SDF1-3' A/3 'A* homozygotes linked to CD4+ T cell decrease in advanced human immunodeficiency virus type 1 infection. *J Infect Dis* 2000;182:311-5.

23. Meyer L, Magierowska M, Hubert JB, et al. CC-chemokine receptor variants, SDF-1 polymorphism, and disease progression in 720 HIV-infected patients. SEROCO cohort. AIDS 1999;13:624-6.

24. Mangano A, Kopka J, Batalla M et al. Protective effect of CCR2-64I and not CCR5-delta32 and SDF1-3'A in pediatric HIV-1 infection. *J Acquir Immune Defic Syndr* 2000;23:52-7.

25. Ioannidis JP, Rosenberg PS, Goedert JJ, et al. Effects of CCR5-Delta32, CCR2-64I and the SDF-1 3'A alleles on HIV-1 disease progression: An international meta-analysis of individual-patient data. *Ann Intern Med* 2001;135:782-95.

26. John GC, Rousseau C, Dong T et al. Maternal SDF1 3' A polymorphism is associated with increased perinatal human immunodeficiency virus type 1 transmission. *J Virol* 2000;74:5736-9.

27. Tresoldi E, Romiti ML, Boniotto M et al. Prognostic value of the stromal cell-derived factor 1 3'A mutation in pediatric human immunodeficiency virus type 1 infection. *J Infect Dis* 2002;185:696-700.

28. Tiensiwakul P. Stromal cell-derived factor (SDF) 1-3'A polymorphism may play a role in resistance to HIV-1 infection in seronegative high-risk Thais. *Intervirology* 2004;47:87-92.

29. Liu H, Hwangbo Y, Holte S et al. Analysis of genetic polymorphisms in CCR5, CCR2, stromal cell-derived factor-1, RANTES, and dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin in seronegative individuals repeatedly exposed to HIV-1. *J Infect Dis* 2004;190:1055-8.

30. Derdeyn CA, Costello C, Kilby JM et al. Correlation between circulating stromal cell-derived factor 1 levels and CD4+ cell count in human immunodeficiency virus type 1-infected individuals. *AIDS Res Hum Retroviruses* 1999;15:1063-71.

31. Ikegawa M, Yuan J, Matsumoto K et al. Elevated plasma stromal cell-derived factor 1 protein level in the progression of HIV type 1 infection/AIDS. *AIDS Res Hum Retroviruses* 2001;17:587-95.

32. Llano A, Barretina J, Blanco J et al. Stromal cell-derived factor 1 prevents the emergence of syncytium-inducing phenotype of HIV-1 in vivo. *AIDS* 2001;15:1890-2.

33. Soriano A, Martinez C Garcia F, et al. Plasma stromal cell-derived factor (SDF)-1 levels, SDF1-3' A genotype, and expression of CXCR4 on T lymphocytes: Their impact on resistance to human immunodeficiency virus type 1 infection and its progression. *J Infect Dis* 2002;186:922-31.

34. Shalekoff S, Tiemessen CT. Ciculating levels of stromal cell-derived factor 1α and Interleukin 7 in HIV type 1 infection in pulmonary tuberculosis are

reciprocally related to CXCR4 expression on peripheral blood leukocytes. *AIDS Res Hum Retroviruses* 2003;19:461-8.

35. Martin MP, Carrington M, Dean M et al. *CXCR4* polymorphisms and HIV-1 pathogenesis. *J Acquir Immune Defic Syndr* 1998;19:430.

36. Cohen OJ, Paolucci S, Bende SM et al. CXCR4 and CCR5 genetic polymorphsms in long-term nonprogressive human immunodeficiency virus infection: Lack of association with mutations other than CCR5-Δ32. *J Virol* 1998;72:6215-7.

37. Alvarez V, Lopez-Larrea C, Coto E. Mutational analysis of the CCR5 and CXCR4 genes (HIV-1 co-receptors) in resistance to HIV-1 infection and AIDS development among intravenous drug users. *Hum Genet* 1998;102:483-6.

38. Hernandez PA, Gorlin RJ, Lukens JN. Mutations in the chemokine receptor gene *CXCR4* are associated with WHIM syndrome, a combined immunodeficiency disease. *Nat Genet* 2003;34:70-4.

39. Hayes VM, Petersen DC, Scriba TJ,et al. African-based CCR5 single-nucleotide polymorphism associated with HIV-1 disease progression. *AIDS* 2003;16:2229-2231.

40. Wu Y, Hayes VM, Osinga J, et al. Improvement of fragment and primer selection for mutation detection by denaturing gradient gel electrophoresis. *Nucleic Acids Res* 1998;26:5432-40.

41. Livak KJ. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genet Anal* 1999;14:143-9.

42. Hayes VM, Wu Y, Osinga J et al. Improvements in gel composition and electrophoretic conditions for broad-range mutation analysis by denaturing gradient gel electrophoresis. *Nucleic Acids Res* 1999;27:e29.

43. Tishkoff SA, Williams SM. Genetic analysis of African populations: Human evolution and complex disease. *Nat Rev Genet* 2002;3:611-21.

44. Ramaley PA, French N, Kaleebu P et al. Chemokine-receptor genes and AIDS risk. *Nature* 2002;417:140.

45. Petersen DC, Kotze MJ, Zeier MD et al. Novel mutations identified using a comprehensive CCR5-denaturing gradient gel electrophoresis assay. *AIDS* 2001;15:171-7.

46. Ioannidis JP, Ntzani EE, Trikalinos TA. 'Racial' differences in genetic effects for complex diseases. *Nat Genet* 2004;36:1312-8.

47. Williamson C, Loubser SA, Brice B et al. Allelic frequencies of host genetic variants influencing susceptibility to HIV-1 infection and disease in South African populations. *AIDS* 2000;14:449-51.

48. Su B, Sun G, Lu D et al. Distribution of three HIV-1 resistance-conferring polymorphisms (SDF1-3'A, CCR2-64I, and CCR5-delta32) in global populations. *Eur J Hum Genet* 2000;8:975-9.

49. Kimura R, Nishioka T, Soemantri A, Ishida T. Allele-specific transcript quantification detects haplotypic variation in levels of the SDF-1 transcripts. *Hum Mol Genet* 2005;14:1579-85.

Table 1: *CXCR4* primers sets and experimental conditions for PCR amplification and DGGE.

Fragment	Amplimers, 5'-3'	Size (bp)	Temperature (°C) Melting	Temperature (°C) Annealing
A	CTCCAGTAGCCACCGCATCT	154	73	62
	[40GC]GCTGCGCTCTAAGTTCAAACG			
B	[40GC]GAATGTCCATTCCTTTGCCTCT	286	73	60
	GCCTGTACTTGTCCGTCATGC			
C	[40GC]CCACCATCTACTCCATCATC	397	66	55
	AGACGCCAACATAGACCAC			
D	CACGCCACCAACAGTCAGA	278	71	60
	[40GC]AGCAGGACAGGATGACAATACC			
E	[40GC]CAGTTTCAGCACATCATGGT	180	66	55
	AGGATGAGGATGACTGTGGT			
F	CATCTCCAAGCTGTCACACT	445	66	54
	[40GC]TTACATCTGTGTTAGCTGGAGT			
G	TCCACTGAGTCTGAGTCTTCAA	282	67	54
	[40GC]TCCTGCCTAGACACACATCA			

bp, base pair
GC-clamp used was as follows: [40GC] CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCCG

Table 2: Allele and genotype distribution of the *SDF1*-3'G>A SNP in 257 HIV-1 infected cases versus 113 HIV-1 seronegative controls from the Xhosa population.

SNP	HIV+ (%) (n = 257)[*]	HIV- (%) (n = 113)[*]	*P* value[#]
***SDF1*-3'G>A**			
G	495 (96.3)	224 (99.1)	**0.0319**
A	19 (3.7)	2 (0.9)	
GG	239 (93)	111 (98.2)	0.1191
GA	17 (6.6)	2 (1.8)	
AA	1 (0.4)	0	
	$P_{HWE} = 0.26^{\delta}$	$P_{HWE} = 0.92^{\delta}$	

HIV+, HIV-1 seropositive; HIV-, HIV-1 seronegative
[*] n, number of individuals. Number of alleles, HIV+ (2n = 514) and HIV- (2n = 226)
[#] Two-sided Fisher's exact or chi-squared (x^2). P < 0.05 was required for statistical significance and is presented in bold.
[δ] Test for Hardy-Weinberg equilibrium.

Chapter 3.1.

Lack of association with *TNFα* promoter SNPs and susceptibility to HIV-1 infection in an African population

Desiree C. Petersen[1], Sadeep Shrestha[2,3], Julie Bergeron[3], Bert Gold[2], Vanessa M. Hayes[4,5] and Michael Dean[2]

Department of Medical Virology[1], University of Stellenbosch, Tygerberg Medical School, South Africa; Laboratory of Genomic Diversity[2], National Cancer Institute, Frederick, Maryland, USA; SAIC[3], Frederick, Maryland, USA; Cancer Research Program[4], Garvan Institute of Medical Research, Darlinghurst, Sydney, NSW, Australia; University of New South Wales [5], Sydney, Australia.

(Submitted)

(Short Communication)

(References cited in the format of the journal to which this manuscript is submitted)

ABSTRACT

Tumor necrosis factor α (TNFα) is a pro-inflammatory cytokine functioning as a mediator of host response to pathogens and is considered to be the most potent of the human immunodeficiency virus-1 (HIV-1) inducing cytokines. Previous studies investigating the effect of *TNFα* promoter single nucleotide polymorphisms (SNPs) on disease progression to acquired immunodeficiency syndrome (AIDS) have resulted in inconsistent findings. Our study included the genotyping of two *TNFα* promoter SNPs, both involving guanine (G) to adenosine (A) transitions, at positions –308 and –238, to determine association with HIV-1 susceptibility in an African population. The study cohort, originating from South Africa, consisted of 215 HIV-1 seropositive individuals and 113 HIV-1 seronegative controls representing Africans belonging to the Xhosa ethnic group. Statistical testing, including allele, genotype and haplotype analysis, of the *TNFα* promoter SNPs in cases versus controls revealed no significant associations with risk for HIV-1 infection (all *P* values > 0.05). This study resulted in a lack of association with *TNFα* promoter SNPs directly influencing host susceptibility to HIV-1 infection in the Xhosa. However, the possibility that these SNPs in combination with other closely linked alleles could influence HIV-1/AIDS pathogenesis in African populations still exists.

INTRODUCTION

Cytokines are part of a complex network of proteins that mediate inflammatory and immune responses. Their stimulatory and inhibitory effects were found to influence human immunodeficiency virus-1 (HIV-1) replication.[1] Tumor necrosis

163

factor α (TNFα), a pro-inflammatory cytokine, is an important and potent inducer of HIV-1 replication. It activates a cellular transcription factor, NF-κB, which enhances virus expression.[2-4] The *TNFα* gene (MIM# 191160) is located within the highly polymorphic major histocompatibility complex (MHC) region on chromosome 6p21.3.[5,6] Genetic variants occurring within the promoter region of *TNFα* have been suggested to alter the production and regulation of this cytokine. A number of *TNFα* promoter single nucleotide polymorphisms (SNPs), including two G-A transitions at positions −308 (rs1800629) and −238 (rs361525) relative to the transcription start site, have been identified.[7-9] Allele frequencies previously reported in African populations for the *TNFα*-308G>A and *TNFα*-238G>A and SNPs range from 0.08 − 0.16 and 0 − 0.11, respectively.[10-12] *TNFα*-308G>A is the most commonly studied SNP in the *TNFα* promoter, although its functional significance remains questionable. Its influence on gene transcription and/or TNFα production differs between studies. Conflicting findings have also been reported for the effect of *TNFα*-238G>A on TNFα production.[13, 14] *TNFα* is a well-documented candidate gene for many diseases, with inconsistent findings between the presence of *TNFα*-308G>A and/or *TNFα*-238G>A and HIV-1 disease susceptibility in Caucasians.[15-17] These include a weak association with homozygosity for *TNFα*-308A and long-term non-progression[16]; and an association with the *TNFα*-308G/A heterozygous genotype and faster disease progression to acquired immunodeficiency syndrome (AIDS).[17] A third study indicated no significant associations with any of the *TNFα* promoter SNPs and rate of disease progression, however elevated levels of TNFα were observed in HIV-1 seropositive individuals who had developed AIDS.[15] The *TNFα*-308G>A and

TNFα-238G>A SNPs have also been associated with influencing the development of HIV-related dementia[18] and lipodystrophy[19;20], respectively.

MATERIALS AND METHODS

It is important to evaluate HIV-1/AIDS candidate gene SNPs with possible functional consequences in diverse ethnic groups to assess disease associations in specific populations. Our study aimed at determining the role of *TNFα*-308G>A and *TNFα*-238G>A SNPs in host susceptibility to HIV-1 infection within the understudied Africans who belong to the Xhosa ethnic group. The 215 HIV-1 seropositive patients were from three locations, Tygerberg Hospital; Woodstock Chapel Street Community Health Clinic; and the Langa Clinic, all in the Western Cape Province of South Africa. Disease progression for the many of these individuals is unknown as insufficient clinical information is available due to infrequent follow-up. The control group consisted of 113 HIV-1 seronegative population-matched samples from the Western Cape Province Blood Transfusion Service of South Africa. Informed consent was obtained from all the study participants and the Ethics Review Committee of the University of Stellenbosch approved the study protocol (98/158). Genotyping of both *TNFα* SNPs was performed using the 5' nuclease or TaqMan allelic discrimination method.[21] The assay primers and probes for *TNFα*-308G-A were: 5'-CCTGCATCCTGTCTGGAAGTTAGAAG-3 and 5'-TGGGCCACTGACTGATT TGTGTGT-3', 5'-FAM-AACCCCGTCC**T**CATGCCCCTCAA-TAMRA-3' and 5'-VIC-AACCCCGTCC**C**CATGCCCCTC-TAMRA-3' and for *TNFα*-238G-A were: 5'-CAGTGGCCCAGAAGACCC-3' and 5'-AGCATCAAGGATACCCCTC AC-3', 5'-FAM-AATC**A**GAGCAGGGAGGATGGGGA-TAMRA-3' and 5'-VIC-AA

TC**G**GAGCAGGGAGGATGGG-TAMRA-3'. The TaqMan Universal PCR Master Mix (Applied Biosysytems, Foster City, California) was used for amplification (detailed PCR reaction mix protocol available on request) with the cycling conditions including a initial denaturation of $95^{o}C$ for 10 minutes, followed by 50 cycles of denaturation at $95^{o}C$ for 15 seconds and annealing at $58^{o}C$ for 1 minute. Measurement of end-point fluorescence for allelic discrimination and genotype determination was performed on the 7900 high throughput sequence detection system (Applied Biosysytems, Foster City, California). The comparison of allele frequencies in HIV-1 seropositives versus HIV-1 seronegatives and testing for significance of heterogeneity was achieved using the two-sided Fisher's exact test for 2x2 contingency tables (GraphPad Software Inc, San Diego, California). Allele and genotype distributions, including consistency of the genotypes observed with Hardy Weinberg equilibrium, were determined for the different population groups. Genotype frequencies between cases and controls were also compared for significance using the chi-squared (x^2) test for independence. Haplotypes were estimated using the expectation-maximization (E-M) algorithm and the frequencies between cases and controls were assessed. (SAS Institute Inc. software, Cary, North Carolina).

RESULTS AND DISCUSSION

Using the *TNFα*-308G>A and *TNFα*-238G>A SNP allelic discrimination assays we determined the allele, genotype and haplotype frequencies. We further tested for significant associations between the presence of the two polymorphic

markers (individually and combined), and susceptibility to HIV-1 infection in our Xhosa sample group (Tables 1 and 2).

Genotyping was successful for 319 (207 HIV+, 112 HIV-) and 325 (213 HIV+, 112 HIV-) samples for the *TNFα*-308G>A and *TNFα*-238G>A SNPs, respectively. The *TNFα* promoter SNPs in the case and control groups displayed expected genotype distributions for the Hardy Weinberg equilibrium (all *P* values > 0.25). No significance was observed between allele frequencies in the HIV-1 seropositives versus the HIV-1 seronegatives for both the *TNFα*-308G>A and *TNFα*-238G>A SNPs (Table 1). The *TNFα*-308G>A SNP (14.26%) did however occur at an overall higher allele frequency compared to the *TNFα*-238G>A SNP (5.54%), which is in agreement with findings previously reported for African populations.[10-12] Pairwise linkage disequilibrium (LD) analysis based on the r^2 coefficient measure (data not shown) provided no strong LD, which is similar to previous reports.[9] Furthermore, the independent genotype (Table 1) and estimated haplotype (Table 2) analysis showed no significant association with susceptibility to HIV-1 infection. Since the majority of HIV seropositive individuals have no clinical staging (3 slow, 34 normal, 11 fast and 167 unknown progressors), associations with disease progression to AIDS were not considered.

Previous findings for *TNFα* promoter SNPs and disease progression to AIDS emerge as being generally inconsistent[15-17], but the role of other genetic variants at neighbouring loci has also been investigated. This includes a lymphotoxin alpha (*LTα*, previously called *TNF-beta)* microsatellite

polymorphism designated *TNF*c2 allele, which has been associated with slower disease progression to AIDS among Caucasians[22], but when occurring homozygously in Africans, was found to offer a decreased risk for HIV-1 infection.[23] It has also been shown that the *TNFα*-308A allele is part of an extended Caucasian human leukocyte antigen (*HLA*) haplotype, *HLA*-A1-B8-DR3[24], which has been associated with increased TNFα production[25,26] and faster disease progression to AIDS.[27]   All these findings thus suggest ethnic-based differences in association to HIV-1 disease.

This study indicates that the commonly studied *TNFα*-308G>A and *TNFα*-238G>A SNPs do not directly influence host susceptibility to HIV-1 infection in our Xhosa ethnic group. The crucial promoter region for *TNFα* transcriptional control has been characterised and it was found that both *TNFα*-308G>A and *TNFα*-238G>A are not within an important regulatory region.[28-30] These *TNFα* promoter SNPs may therefore together with other *TNFα* variants or alleles of the neighbouring *LTα* gene have an effect on HIV/AIDS outcomes in Africans. Furthermore these *TNFα* promoter SNPs in combination with the closely located *HLA* alleles, which have been previously implicated in African-based HIV/AIDS pathogenesis[31], could form specific haplotypes that determine risks for HIV-1 infection in the understudied African populations. Our unique study of a Xhosa ethnic group from pandemic stricken Sub-Sahara accentuates the importance for ongoing population-based HIV/AIDS association studies in large well-defined sample groups.

ACKNOWLEDGEMENTS

REFERENCES

1. Vincenzi E, Biswas P, Mengozzi, and Poli G: Role of pro-inflammatory cytokines and beta-chemokines in controlling HIV replication. J Leukoc Biol 1997; 62:34-40.

2. Duh EJ, Maury WJ, Folks TM, Fauci AS, and Rabson AB: Tumor necrosis factor alpha activates human immunodeficiency virus type 1 through induction of nuclear factor binding to the NF-kappa B sites in the long terminal repeat. Proc Natl Acad Sci USA 1989; 86:5974-5978.

3. Osborn L, Kunkel S, and Nabel GJ: Tumor necrosis factor alpha and interleukin-1 stimulate the human immunodeficiency virus enhancer by activation of the nuclear factor kappa B. Proc Natl Acad Sci USA 1989; 86:2336-2340.

4. Matsuyma T, Kobayashi N, and Yamamoto N: Cytokines and HIV infection: is AIDS a tumor necrosis factor disease? AIDS 1991; 5:1405-1417.

5. Nedwin GE, Naylor SL, Sakaguchi AY, *et al*.: Human lymphotoxin and tumor necrosis factor genes: structure, homology and chromosomal localization. Nucleic Acids Res 1985; 13:6361-6373.

6. Spies T, Morton CC, Nedospasov SA, Fiers W, Pious D, and Strominger JL: Genes for the tumor necrosis factors alpha and beta are linked to the human major histocompatibility complex. Proc Natl Acad Sci USA 1986; 83:8699-8702.

7. Wilson AG, di Giovine FS, Blakemore ALF, and Duff GW: Single base polymorphism in the human tumor necrosis factor alpha (TNFα) gene detectable by *Nco*I restriction of PCR product. Hum Mol Genet 1992; 1:353.

8. D'Alfonso SD, and Richiardi PM: A polymorphic variation in a putative regulation box of the TNFα promoter region. Immunogenetics 1994; 39:150-154.

9. Hamann A, Mantzoros C, Vidal-Puig A, and Flier JS: Genetic variability in the TNF-α promoter is not associated with type II diabetes mellitus (NIDDM). Biochem Biophys Res Commun 1995; 211:833-839.

10. McQuire W, Hill AV, Allsopp CE, Greenwood BM, and Kwiatkowski D: Variation in the TNF-alpha promoter region associated with susceptibility to celebral malaria. Nature 1994; 371:508-510.

11. Conway DJ, Holland MJ, and Bailey RL: Scarring trachoma is associated with polymorphism in the tumor necrosis factor alpha (TNF-α) gene promoter and with elevated TNF-α levels in tear fluid. Infec Immun 1997, 65, 1003.

12. Baena A, Leung JY, Sullivan AD, *et al*.: TNF-alpha promoter single nucleotide polymorphisms are markers of human ancestry. Genes Immun 2002; 3:482-487.

13. Reynard MP, Turner D, and Navarrete CV: Allele frequencies of polymorphisms of tumour necrosis factor-α, interleukin-10, interferon-γ and interleukin-2 genes in a North European Caucasoid group from the UK. Eur J of Immunogenetics 2000; 27:241-249.

14. Hajeer AH, and Hutchinson IV: Influence of TNFα gene polymorphisms on TNFα production and disease. Hum Immunol 2001; 62:1191-1199.

15. Brinkman BMN, Keet IPM, Miedema F, Verweij CL, and Klein MR: Polymorphisms within the human tumor necrosis factor-α promoter region in human immunodeficiency virus type 1-seropositive persons. J Infect Dis 1997; 175:188-190.

16. Knuchel MC, Spira TJ, Neumann AU, *et al*.: Analysis of a biallelic polymorphism in the tumor necrosis factor α promoter and HIV type 1 disease progression.

17. Smolnikova MV and Konenkov VI: Association of IL2, TNFA, IL4 and IL10 promoter gene polymorphisms with the rate of progression of HIV infection. Russ J Immunol 2002; 7:349-356.

18. Quasney MW, Zhang Q, Sargent S, Mynatt M, Glass J, and McArthur J: Increased frequency of the tumor necrosis factor-alpha-308 A allele in adults with human immunodeficiency virus dementia. Ann Neurol 2001; 50: 157-162.

19. Maher B, Alfirevic A, Vilar FJ, Wilkins EG, Park BK, and Pirmohamed M: TNF-alpha promoter region gene polymorphisms in HIV-positive patients with lipodystrophy. AIDS 2002; 16: 2013-2018.

20. Nolan D, Moore C, Castley A, *et al*.: Tumour necrosis factor-alpha gene   -238G/A promoter polymorphism associated with a more rapid onset of lipodystrophy. AIDS 2003; 17:121-123.

21. Livak KJ: Allelic discrimination using fluorogenic probes and the 5' nuclease assay. Genet Anal 1999; 14:143-149.

22. Khoo SH, Pepper L, Snowden N, *et al*.:   Tumour necrosis factor c2 microsatellite allele is associated with the rate of HIV disease progression. AIDS 1997; 11:423-428.

23. Tang J, Myracle AD, Allen S, *et al*.: Novel alleles at the Lymphotoxin Alpha (LTα) locus mark extended HLA haplotypes in Native Africans.   Hum Immunol; 62:269-278.

24. Wilson AG, de Vries N, Pociot F, di Giovine FS, van der Putte LBA, and Duff GW: An allelic polymorphism within the tumor necrosis factor α promoter region is strongly associated with HLA A1, B8 and DR3 alleles. J of Exp Med 1993; 177:557-560.

25. Jacob CO, Fronek Z, Lewis GD, Koo M, Hansen JA, and McDevitt HO: Heritable major histocompatibility complex class II-associated differences in production of tumor necrosis factor alpha: relevance to genetic predisposition   to systemic lupus erythematosus. Proc Natl Acad Sci USA 1990; 87:1233-1237.

26. Abraham LJ, French MAH, and Dawkins RL: Polymorphic MHC ancestral haplotypes affect the activity of tumour necrosis factor-alpha. Clin Exp Immunol 1993; 92:14-18.

27. Kaslow RA, Duquesnoy R, Van Raden M, *et al*: A1, Cw7, B8, DR3 HLA antigen combination associated with rapid decline of T-helper lymphocytes in HIV-1 infection. A report from the Multicenter AIDS Cohort Study. Lancet 1990; 335: 927-930.

28. Falvo JV, Uglialoro AM, Brinkman BM, *et al*.: Stimulus-specific assembly of enhancer complexes on the tumor necrosis factor alpha gene promoter. Mol Cell Biol 2000; 20: 2239-2247.

29. Tsai EY, Falvo JV, Tsytsykova AV, *et a*l: A lipopolysaccharide-specific enhancer complex involving ets, elk-1, sp1, and CREB binding protein and p300 is recruited to the tumor necrosis factor alpha promoter *in vivo*. Mol Cell Biol 2000; 20:6084-6094.

30. Tsytsykova AV and Goldfeld AE: Inducer-specific enhanceosome formation controls TNF-alpha gene expression in T lymphocytes. Mol Cell Biol 2002; 22:2620-2631.

31. Carrington M and O'Brien S: The influence of *HLA* genotype on AIDS. Annu Rev Med 2003; 54:535-551.

Table 1. Allele and genotype distributions of the *TNFα* promoter SNPs in cases versus controls

Gene[a] SNP	NCBI ID	Minor Allele Frequency (%)			Genotype Frequency (%) [b]						
					-/-		+/-		+/+		
		HIV+	HIV-	P	HIV+	HIV-	HIV+	HIV-	HIV+	HIV-	P
TNFα											
-308 (G/A)	rs 1800629	13.77	15.18	0.637	2.42	3.57	22.71	23.21	74.88	73.21	0.827
-238 (G/A)	rs 361525	4.93	6.70	0.370	0	0.89	9.86	11.61	90.14	87.50	0.337

HIV+, seropositive; HIV -, seronegative
[a]The second base pair indicates the minor allele
[b](-/-), (+/-) and (+/+) represents the minor allele homozygotes, heterozygotes and major allele homozygotes, respectively.

Table 2. Haplotype analysis of the *TNFα* promoter SNPs in cases versus controls

Gene	Haplotype -308G>A and -238G>A	Frequency (%)		P
		HIV+	HIV-	
TNFα	A – A	0.01	0.19	0.646
	G – A	4.93	6.49	0.438
	A – G	13.78	14.99	0.709
	G – G	81.28	78.33	0.409

HIV+, seropositive; HIV -, seronegative

Chapter 3.2.

The influence of *IL4* and *IL10* promoter SNPs and haplotypes on HIV-1 infection risk in Sub-Saharan Africans

Desiree C. Petersen[1], Elizabeth A. Tindall[2], Sadeep Shrestha[3,4], Annette Laten[1], Estrelita Janse van Rensburg[1] and Vanessa M. Hayes[2,5]

Departments of Medical Virology[1], University of Stellenbosch, Tygerberg Medical School, South Africa; Cancer Research Program[2], Garvan Institute of Medical Research, St Vincent's Hospital, Darlinghurst, Sydney, NSW, Australia; Laboratory of Genomic Diversity[3], National Cancer Institute, Frederick, Maryland, USA; SAIC[4], Frederick, Maryland USA; University of NSW[5], Sydney, NSW, Australia.

(Brief report)

ABSTRACT

The cytokines Interleukin-4 (IL4) and Interleukin-10 (IL10) both influence HIV-1 replication. Previous studies determining the role of *IL4* and *IL10* promoter variants in HIV-1 pathogenesis have resulted in an array of associations for diverse populations. We analysed the influence of *IL4* and *IL10* promoter SNPs and extended haplotypes on HIV-1 infection risk within Sub-Saharan Africans of Xhosa descent. Significance was observed for two linked *IL10* SNPs (-819C>T and -592C>A) and specific haplotypes for *IL4* and *IL10* (all *P* < 0.02). This includes a novel *IL4* haplotype being associated with a decreased risk for HIV-1 infection.

INTRODUCTION

Cytokines such as Interleukin-4 (IL4) and Interleukin-10 (IL10) play a vital role in modulating host immune responses with both having been found to influence human immunodeficiency virus-1 (HIV-1) replication. IL4 is responsible for down-regulation of CC chemokine receptor 5 (CCR5), with inhibition of early stage R5 virus replication, and up-regulation of the CXC chemokine receptor 4 (CXCR4), with enhanced replication of the later emerging X4 viruses [Nakayama et al., 2000]. IL10 has been found to inhibit HIV-1 replication and it is believed that this control of virus proliferation is due to restriction of the amount of macrophages available for replication [Reviewed in Carrington et al., 2001].

Previously identified single nucleotide polymorphisms (SNPs) are well documented for the promoter regions of *IL4* (MIM# 147780) and *IL10* (MIM# 147780) and have been associated with functional consequences that influence

177

gene expression and protein production [Reviewed in Anastassopoulou and Kostrikis, 2003]. Commonly occurring promoter variants include two *IL4* SNPs, -589C>T (rs2243250) [Rosenwasser et al., 1995] and -33C>T (rs2070874) [Takabayashi et al., 1999] (relative to the translation start site), and three *IL10* SNPs, -1082A>G (rs1800896), -819C>T (rs1800871) and -592A>C (rs1800872) (relative to transcription start site) [Turner et al., 1997]. These SNPs form three major haplotypes for both *IL4* (CC, TC, TT) and *IL10* (ATA, ACC, GCC) in most populations, including African-based ethnic groups [Shin et al., 2000; Meenagh et al., 2002; Vasilescu et al., 2003; Basehore et al., 2004; Wang et al., 2004].

IL4 and *IL10* promoter SNPs and extended haplotypes have previously been implicated in HIV-1/AIDS pathogenesis. In Japanese individuals, *IL4*-589C>T is in complete linkage disequlibrium with *IL4*-33C>T and has been associated with decreased risk for HIV-1 infection. However, the *IL4*-589 TT homozygous genotype was correlated with possibly faster progression to AIDS due to the more rapid emergence of X4 variants during late disease [Nakayama et al., 2000]. Different findings were reported for Dutch Caucasians where the effect of the *IL4*-589C>T SNP involved delayed acquisition of X4 variants and no overall influence on disease progression [Kwa et al., 2003]. A French Caucasian-based study, where *IL4*-589 TT was rare, indicated that the *IL4*-589C>T SNP is protective against disease progression to AIDS and death by reducing viral load [Nakayama et al., 2002]. This protective effect in French Caucasians was confirmed in another study where a specific haplotype that carries the *IL4*-589 T allele was associated with slower disease progression [Vasilescu et al., 2003]. More recently, a study of African American individuals

showed that *IL4*-589 TT is associated with delaying disease progression to AIDS, while *IL4*-33 CC is associated in increased risk for HIV-1 infection [Wang et al., 2004]. In North American Caucasians, *IL10*-592C>A, which is in complete linkage disequilibrium with *IL10*-819C>T and strong linkage disequilibrium with *IL10*-1082A>G, has been associated with dominantly increasing risk for HIV-1 infection and faster disease progression to AIDS, particularly evident during late stages of infection [Shin et al., 2000]. Another French Caucasian-based study did not yield similar findings, but was indicative of a haplotype, including the *IL10*-592 C allele, being associated with rapid disease progression [Vasilescu et al., 2003]. A recent study showed association between the *IL10*-1082 AA genotype and increased susceptibility to HIV-1 infection in Hispanics. An *IL10* haplotype comprised of 5 alleles, including -1082G, -819C, -592C, and was also associated with increased risk for HIV-1 infection in African Americans [Wang et al., 2004].

These various HIV-1/AIDS associations observed within specific ethnic groups emphasise the importance of determining the potential role of *IL4* and *IL10* promoter SNPs and haplotypes within well-defined understudied African populations. We investigated the promoter regions of *IL4* and *IL10* for variants and determined the role of *IL4*-589C>T, *IL4*-33C>T, *IL10*-1082A>G, *IL10*-819C>T and *IL10*-592C>A in susceptibility to HIV-1 infection within Africans from pandemic stricken Sub-Saharan Africa.

SUBJECTS AND METHODS

Our study sample consisted of 310 African individuals of Xhosa descent who all reside in the Western Cape Province of South Africa. The Xhosa are from the early clan of the Nguni, the most southern group of Bantu migrants from central Africa. In the Western Cape the Xhosa form 90% of the African population (Statistics South Africa, 2001; www.statssa.gov.za). The 197 HIV-1 seropositives (69% females, 31% males) are all patients of Tygerberg Hospital, Woodstock Chapel Street Community Health Clinic or the Langa Clinic. Due to the lack of updated clinical information, the disease progression for most of these patients remains unknown. Blood donors for the Western Province Blood Transfusion Service of South Africa represented our 113 HIV-1 seronegative (62% females, 38% males) population-matched controls Informed consent was obtained from all the study participants and the Ethics Review Committee of the University of Stellenbosch approved the study protocol (#98/158).

Assays based on DGGE were utilised for the mutation screening of *IL4* and *IL10* promoter SNPs in a blinded manner. Two DGGE PCR primer sets, including a GC-clamp (GC-rich fragment) on the 5'end of either the forward or reverse primer, were designed for each of the partial promoter regions of *IL4* (nucleotide (nt) -549 to -640 and codon 8 to nt -169) and *IL10* (nt -878 to -1113 and nt -569 to -860) (Table 1). Detailed PCR reaction mix protocols and amplification conditions are available on request. The two amplicons for both *IL4* (pooled P1 and P2) and *IL10* (pooled P1 and P2) were electrophoresed in a single lane of a 9% polyacrylamide gel containing a 30% to 70% urea and formamide (UF) denaturing gradient (100% UF = 7mol/L urea per 40%

deionised formamide), at 60°C for 110 volts overnight. Gels were stained with ethidium bromide before being photographed for mutation analysis. This was followed by sequencing of samples that showed aberrant banding patterns.

Testing for significance of heterogeneity in cases versus controls was achieved using the Fischer's exact or chi-square (x^2) test for 2x2 contingency tables. The Hardy-Weinberg equilibrium (HWE) principle was applied to measure the distribution of allele and genotype frequencies within the sample group. Using the x^2 test for independence, significance of genotype frequencies between cases and controls were determined. Haplotype analysis was performed using the expectation-maximization (E-M) algorithm. (GraphPad Software Inc, San Diego, California and SAS Institute Inc. software, Cary, North Carolina).

RESULTS

The *IL4* and *IL10* DGGE assays were specifically designed for the partial analysis of the promoter regions of these genes.   This resulted in the identification three novel variants within the *IL4* promoter region at nucleotide positions -594A>C, -142A>G and -112G>A (relative to the translation start site), as well as the identification of the five previously reported *IL4* and *IL10* promoter SNPs.   Allele frequencies of the novel variants (0.005 to 0.021) did not warrant further analysis in this study. Genotyping the known promoter variants showed distributions in accordance to the HWE principle for all SNPs, except *IL4*-589C>T in the HIV-1 seropositives (*P* = 0.03).   The *IL10*-819C>T and *IL10*-592C>A SNPs were in complete linkage disequlibrium as reflected by the allele and genotype frequencies shown in Table 1. No significance was found for allele frequency comparison between HIV-1 seropositives and HIV-1

seronegatives for both the *IL4* and *IL10* SNPs (Table 1). The allele carriage frequency analysis did however indicate significance for the *IL10*-819 T allele (CT and TT) and *IL10*-592 A allele (CA and AA) occurring more commonly in the uninfected controls when compared to the HIV-1 infected patients (*P* = 0.02). Furthermore, the independent genotype analysis resulted in a significant association being observed for *IL10*-819C>T and *IL10*-592C>A in the cases versus the controls (*P* = 0.02) (Table 1). The haplotype analysis for the *IL4*-589T>C and *IL4*-33T>C SNPs resulted in an uncommon allele combination, namely CT, being present exclusively in the uninfected controls (*P* = 0.0013) (Table 2). Haplotype analysis for the *IL10*-1082A>G, *IL10*-819C>T and *IL10*-592C>A SNPs also showed marginal significance for the ACC allele combination occurring more commonly in the HIV-1 seronegatives (*P* = 0.04) (Table 2). No associations were considered regarding progression to AIDS due to the large number of HIV-1 infected patients with unknown disease status.

DISCUSSION

The aim of our study was to determine the significance of *IL4* and *IL10* promoter variants in HIV-1 susceptibility for a well-defined Sub-Saharan African population, represented by Xhosa individuals. Previous inconsistent HIV-1/AIDS associations with *IL4* and *IL10* SNPs and haplotypes across various ethnic groups [Nakayama et al., 2000; Nakayama et al., 2002; Shin et al., 2000; Kwa et al., 2003; Vasilescu et al., 2003; Wang et al., 2004] have emphasised the importance of further investigating the role of these genes, particularly in a population severely infected by this devastating pandemic. It has been shown that not all candidate gene variants and haplotypes

previously implicated in HIV-1/AIDS are either present or have the same effect in the understudied African-based populations [Petersen et al., 2001; Ramaley et al., 2002]. Allele frequencies for the known *IL4* and *IL10* promoter SNPs in our Xhosa sample group were similar to those previously reported for African-based populations [Shin et al., 2000; Meenagh et al., 2002; Basehore et al., 2004; Wang et al., 2004].

The HIV-1 infected patients showed deviation from the expected HWE genotype distributions for the *IL4*-589C>T SNP due to over representation of the heterozygous genotype (*IL4*-589 CT), which could be indicative of an association with increased susceptibility to HIV-1 infection ($P = 0.03$) (data not shown). Further analysis showed that an uncommon haplotype consisting of the *IL4* -589 C and *IL4* -33 T alleles (CT) is associated with decreased risk for HIV-1 infection as it was observed exclusively in 3.4% of the HIV-1 seronegatives ($P = 0.0013$) (Table 2). The CT haplotype has not been previously reported in disease association studies, including those for African-based populations where the *IL4*-589C>T and *IL4*-33C>T promoter SNPs form only three haplotypes (CC, TC, TT) [Basehore et al., 2004; Wang et al., 2004], as observed in the HIV-1 seropositives.

A significant association was found with the presence of the T (CT and TT) and A (CA and AA) alleles of the *IL10*-819C>T and *IL10*-592C>A linked SNPs, respectively, and decreased risk for HIV-1 infection ($P = 0.02$) (data not shown). The *IL10*-819 CT and *IL10*-592 CA heterozygous genotypes were found at an increased frequency in the HIV-1 seronegatives (56%) versus the HIV-1

seropositives (39%) (P = 0.02) (Table 1). This further supports an association between the presence of *IL10*-819 T and *IL10*-592 A alleles and reduced susceptibility to HIV-1 infection. A marginal significance was observed for a haplotype consisting of the *IL10*-1082 A*, IL10*-819 C and *IL10*-592 C alleles (ACC) occurring at a higher frequency in the uninfected controls compared to the HIV-1 infected patients (P = 0.04) (Table 2). It is therefore only suggestive of a possible association with decreased risk for HIV-1 infection The presence of only three allele combinations (ATA, ACC, GCC) in both the cases and controls indicates strong linkage disequilibrium, which is in agreement with previous African-based studies [Shin et al., 2000; Meenagh et al., 2002].

The three novel mutations identified in the *IL4* promoter region, *IL4*-594A>C, -*IL4*142A>G and *IL4*-112G>A could possibly have a functional significance by influencing regulation of gene transcription and/or protein expression. Due to the rare occurrence of these genetic variants and lack of clinical information for HIV-1 seropositive patients, no HIV-1/AIDS associations could be made. Further screening of 43 Caucasian control individuals did however provide confirmation of these mutations being exclusive to the Xhosa ethnic group.

Previous studies of the *IL4* and *IL10* promoter variants and HIV-1/AIDS have resulted in an array of inconsistent findings. A major strength of our study is the assessment of these markers in a population most affected by HIV-1/AIDS. Our study indicates associations with HIV-1 susceptibility that are specific for the Xhosa ethnic group, in particular a novel *IL4* promoter haplotype being significantly associated with resistance to HIV-1 infection. Although the

IL4-589C>T SNP has been associated with increased gene transcription [Rosenwasser et al., 1995], further investigation is needed for determining the functional relevance of the *IL4*-33C>T SNP and more specifically the newly identified African-based *IL4* haplotype. Future studies with larger sample numbers from understudied Sub-Saharan African populations are also required.

ACKNOWLEDGMENTS

REFERENCES

- Anastassopoulou CG and Kostrikis LG. The impact of human allelic variation on HIV-1 disease. *Curr HIV Res* 2003; 1:185-203.

- Basehore MJ, Howard TD, Lange LA, et al. A comprehensive evaluation of IL4 variants in ethnically diverse populations: Association of total serum IgE levels and asthma in white subjects. *J Allergy Clin Immunol* 2004; 114:80-87.

- Carrington M, Nelson G and O'Brien S. Considering genetic profiles in functional studies of immune responsiveness to HIV-1. *Immunol Lett* 2001: 131-140.

- Kwa D, van Rij RP, Boeser-Nunnink B, Vingerhoed J and Schuitemaker H. Association between an interleukin-4 promoter polymorphism and acquisition of CXCR4 using HIV-1 variants. *AIDS* 2003; 17:981-985.

- Meenagh A, Williams F, Ross OA, et al. Frequency of cytokine polymorphisms in populations from Western Europe, Africa, Asia, the Middle East and South America. *Human Immunol* 2002; 63: 1055-1061.

- Nakayama EE, Hoshino Y, Xin X, et al. Polymorphism in the interleukin-4 promoter affects acquisition of human immunodeficiency virus type 1 syncytium-inducing phenotype. *J Virol* 2000; 74:5452-5459.

- Nakayama EE, Meyer L, Iwamoto A, et al. Protective effect of Interleukin-4-589T polymorphism on human immunodeficiency virus type 1 disease progression: relationship with virus load. *J Infect Dis* 2002; 185:1183-1186.

- Petersen DC, Kotze MJ, Zeier MD et al. Novel mutations identified using a comprehensive CCR5-denaturing gradient gel electrophoresis assay. *AIDS* 2001, 15:171-177.

- Ramaley PA, French N, Kaleebu P et al. Chemokine-receptor genes and AIDS risk. *Nature* 2002; 417:140.

- Rosenwasser LJ, Klemm DJ, Dresback JK, et al. Promoter polymorphisms in the chromosome 5 gene cluster in asthma and atopy. *Clin Exp Allergy* 1995; 25 (Suppl 2): 74-78.

- Shin HO, Winkler C, Stephens JC, et al. Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of IL10. *Proc Natl Acad Sci USA* 2000; 97:14467-14472.

- Takabayashi A, Ihara K, Sasaki Y, Kusuhara K, Nishima S and Hara T. Novel polymorphism in the 5'untranslated region of the interleukin-4 gene. *J Hum Genet* 1999; 44:352-353.

- Turner DM, Williams DM, Sankaran D, Lazarus M, Sinnott PJ and Hutchinson IV 1997; An investigation of polymorphism in the interleukin 10 gene promoter. *Eur J Immunogenet* 1997; 24:1-8.

- Vasilescu A, Heath SC, Ivanova R, et al. Genomic analysis of Th1-Th2 cytokine genes in an AIDS cohort: identification of IL4 and IL10 haplotypes associated with the disease progression. *Gene Immun* 2003; 4:441-449.

- Wang C, Song W, Lobashevsky E, et al. Cytokine and chemokine gene polymorphisms among ethnically diverse North Americans with HIV-1 infection. *J Acquir Immune Defic Syndr* 2004; 35:446-454.

Table 1. Allelic and genotypic distribution of *IL4 and IL10* promoter SNPs in cases versus controls

Gene[a] SNP	NCBI ID	DGGE AMPLICON	Minor Allele Frequency (%)			Genotype Frequency (%) [b]						
			HIV+	HIV-	P	-/-		+/-		+/ +		
						HIV+	HIV-	HIV+	HIV-	HIV+	HIV-	P
IL4												
-589 (T/C)	rs2243250	IL4P1	28	30	0.52	5	11	46	39	49	50	0.09
-33 (T/C)	rs2070874	IL4P2	46	43	0.65	21	20	50	47	29	33	0.83
IL10												
-1082(A/G)	rs1800896	IL10P1	30	31	0.61	9	8	42	47	49	45	0.67
-819(C/T)[#]	rs1800871	IL10P2	32	37	0.14	12	10	39	56	49	34	**0.02***
-592(C/A)[#]	rs1800872	IL10P2	32	37	0.14	12	10	39	56	49	34	**0.02***

HIV+, seropositive; HIV -, seronegative
[a]The second base pair indicates the minor allele
[b](-/-), (+/-) and (+/+) represents the minor allele homozygotes, heterozygotes and major allele homozygotes, respectively.
[#]*IL10*-819C>T and *IL10*-592C>A SNPs are in complete linkage disequilibrium
* Indicates significant *P* value less than 0.05.
IL4P1: 5'-[40GC] [10AT] ACCTGATACGACCTGTCCTT-3' and 5'-GGCAGAATAACAGGCAGACT-3'
IL4P2: 5'-[40GC] [10AT] CCAAGTGACTGACAATCTGGT-3' and 5'-AGCAGTTGGGAGGTGAGAC-3'
IL10P1: 5'-CCAAGACAACACTACTAAGGCT-3' and 5'-[40GC] ACTGTACACCATCTCCAGCA-3'
IL10P2: 5'-[40GC] TTCTCAGTTGGCACTGGTGT-3' and 5'-TTCCAGAGACTGGCTTCCTA-3'
GC-clamp used was as follows: [40GC] CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCGCCCG
AT-stretch used was as follows [10AT] TATAATATTA

Table 2. Haplotype analysis of *IL4* and *IL10* promoter SNPs in cases versus controls

| Gene | Haplotype | Frequency (%) | | P |
		HIV+	HIV-	
IL4	C – C	26.1	24.3	0.53
	T – C	19.5	21.4	0.57
	C – T	0	3.4	**0.0013***
	T – T	54.3	50.9	0.35
IL10	A – T – A	37.6	31.4	0.12
	A – C – C	31.0	39.2	**0.04***
	G – C – C	31.4	29.4	0.58

HIV+, seropositive; HIV -, seronegative
IL4 haplotypes are for the *IL4*-589C>T, *IL4*-33C>T SNPs
IL10 haplotypes are for the *IL10*-1082A>G, *IL10*-819C>T and *IL10*-592C>A SNPs
*Indicates significant *P* value

Chapter 4

Common *MBL* dimorphic markers associated with population-based HIV-1 susceptibility

Desiree C. Petersen[1], Mariska Botha[1], Gianluca Severi[2], Annette Laten[1], Estrelita Janse van Rensburg[1] and Vanessa M. Hayes[3,4]

Department of Medical Virology[1], University of Stellenbosch, Tygerberg Medical School, South Africa; Cancer Epidemiology Centre[2], The Cancer Council Victoria, Carlton, Melbourne, VIC, Australia; Cancer Research Program[3], Garvan Institute of Medical Research, Darlinghurst, Sydney, NSW, Australia; Department of Medicine[4], St Vincent's Hospital Clinical School, University of New South Wales, Sydney, Australia.

(Full article)

ABSTRACT

Mannose-binding lectin (MBL), a calcium dependant serum protein synthesised by the liver, plays an important role in innate immunity and binds to the gp120 envelope protein of HIV-1. Three single nucleotide polymorphisms (SNPs) identified in the coding region of the *MBL* gene (C52R, D54G, E57G) have been associated with lower serum MBL levels, as well as influencing susceptibility to HIV-1 infection, disease progression to AIDS and AIDS free survival. In this study, we genotyped these functional SNPs within two diverse populations from South Africa, including Africans (Xhosa ethnic group) and Caucasians, to assess for association with HIV-1 susceptibility in 135 HIV-1 seropositive cases (114 Africans; 21 Caucasians) versus 109 HIV-1 seronegative controls (60 Africans; 49 Caucasians). In our study, no significant association between the presence of one or more of the functional *MBL* SNPs and HIV-1 susceptibility was observed in Africans or Caucasians. A meta-analysis showed a marginally significant association between HIV-1 seropositivity and the presence of the *MBL*D54G mutant allele (pooled OR = 1.4, 95% CI = 1.0 − 2.0) in Caucasians, while this significance increased when considering homozygote/compound heterozygote association (pooled OR = 6.5, 95% CI = 2.1 − 20.1). In conclusion, the role of the common functional *MBL* SNPs in conferring risk to HIV-1 infection may only be true for Caucasian-based populations.

Keywords: Mannose binding lectin (MBL); HIV-1 susceptibility; meta-analysis; polymorphisms; diverse populations.

INTRODUCTION

The study of host genetics and its role in human immunodeficiency virus-1 (HIV-1) infection and progression to acquired immunodeficiency syndrome (AIDS) has been advanced by the identification of genetic variants in candidate genes encoding for proteins that have a functional role in the maintenance of an effective immune response. Mannose-binding lectin (MBL), also known as mannose-binding or mannan-binding protein, is a calcium dependant serum lectin synthesised by the liver [Kawasaki et al., 1983] as part of the acute phase response following the primary invasion of various microorganisms [Thiel et al., 1992]. The specific role of MBL in the innate immune response involves binding to the carbohydrate-rich domains on pathogens for destruction by either opsonisation (recognition by phagocytic cells) [Kuhlman et al., 1989] or activation of the lectin complement pathway (evolutionary precursor to the classical pathway) [Matsushita and Fujita, 2001].

The human *MBL* gene, also known as *MBL2, MBP1* and *COLEC1* (MIM# 154545) is located on the long arm of chromosome 10 at band position q11.2-21 and is represented by four exons and three introns. The exons code for distinct protein domains, with exon 1 coding for the signal peptide and the NH_2 terminal cysteine-rich, collagen-like domain [Sastry et al., 1989; Taylor et al., 1989]. Exon 1 *MBL* single nucleotide polymorphisms (SNPs), namely *MBL*C52R (rs5030737) [Madsen et al., 1994], *MBL*D54G (rs1800450) [Sumiya et al., 1991] and *MBL*E57G (rs1800451) [Lipscombe et al., 1992], also referred to as the D, B and C alleles, respectively, with A representing the wild-type allele, result in non-conservative amino acid changes disrupting oligomerisation

and resulting in impaired protein function [Sumiya et al., 1991; Wallis and Cheng, 1999]. They are believed to contribute to MBL deficiency in various populations [Madsen et al., 1994; Turner, 1996]. The population-specific allele frequencies for the three functional *MBL* SNPs are summarised in Table 1.

Associations have been found with low serum MBL levels and increased susceptibility to a number of infectious diseases, including HIV-1/AIDS [Turner, 2003; Eisen and Minchinton, 2003]. MBL has been found to bind the gp120 envelope protein of HIV-1 and inhibit entry of the virus into cells *in vitro* [Ezekowitz et al., 1989]. Controversy does however exist as to whether MBL-binding leads to viral lysis or promotes infection by supplying the virus with an additional mode of entry [Sölder et al., 1989; Holmskov et al., 1994]. A range of findings has also been reported for association between MBL and susceptibility to HIV-1/AIDS. One study showed the prevalence of MBL deficiency in HIV-1 seropositives to be similar to that of the normal population and the serum MBL levels in HIV-1 infected individuals were elevated at all stages of disease [Senaldi et al., 1995]. In contrast, a more recent study found that HIV-1 seropositives had significantly lower serum MBL levels when compared to HIV-1 seronegative controls, with MBL levels being linked to disease progression [Prohászka et al., 1997]. Individuals homozygous for the *MBL* SNPs were at increased risk for HIV-1 infection [Garred et al., 1997; Pastinen et al., 1998] and shorter AIDS-free survival [Garred et al., 1997]. Furthermore the variant *MBL* alleles were also associated with slower disease progression to AIDS [Maas et al., 1998]. A recent study in a Gabonese population found homozygosity for *MBL*E57G, as well as compound

heterozygosity for *MBL*E57G (*MBL*E57G was found together with either *MBL*C52R or *MBL*D54G) to be associated with increased susceptibility to HIV-1 infection. The observation that persons heterozygous for this variant appeared to be less susceptible when compared to wildtype and mutant homozygous individuals could be suggestive of heterozygous advantage [Mombo et al., 2003].

Current investigations are therefore not consistent with regards to the contribution of functional *MBL* SNPs influencing HIV-1/AIDS. In this study the *MBL* SNPs were genotyped in diverse populations from South Africa and assessed for ethnic-specific association with HIV-1 susceptibility. Meta-analysis was performed to increase the power of this study. Therefore we aimed to quantify population-based differences in the role of *MBL* functional SNPs in HIV-1 susceptibility.

MATERIALS AND METHODS

Study population

Bloods were obtained from 135 HIV-1 seropositive patients who attended Tygerberg Hospital, Woodstock Chapel Street Community Health Clinic or the Langa Clinic in the Western Cape Province of South Africa. The clinical information for these individuals is limited and therefore disease progression is mostly unknown. The Western Province Blood Transfusion Service of South Africa provided blood samples of 109 HIV-1 seronegative controls that are donors. The two diverse population groups in this study are represented by Africans, predominantly of Xhosa descent, and Caucasians, defined as being of

Dutch, French, German or British origin. The various South African populations have been previously described [Hayes, 2003].

Genotyping

A single denaturing gradient gel electrophoresis (DGGE) amplicon for the detection of the *MBL* SNPs was designed for partial analysis of exon 1 of the *MBL* gene. The primers used were 5'-GTGATTGCCTGTAGCTCTC-3' and 5'-GACATCAGTCTCCTCATATCC-3' with a 40bp GC-rich-fragment (GC-clamp, (5'-CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCCG-3')added to the 5' end of the reverse primer, thus preventing complete strand dissociation. Genomic DNA was extracted using the QiAmp extraction kit (Qiagen). Each PCR reaction had a final volume of 50μl and contained 50ng of genomic DNA, 0.1mM of each deoxyribonucleoside triphosphate (dNTP), 40pmol of each primer, 2.5mM of a 10x Mg^{2+} reaction buffer and 1 unit of DNA *Taq* polymerase (Boehringer Mannheim). Amplification was performed using a 9600 thermocycyler (Applied Biosysytems) and the following PCR cycling conditions; an initial denaturation at 96^0C for 3 minutes, followed by 32 cycles of denaturation at 96^0C for 45 seconds, annealing at 50^0C for 1 minute and elongation at 72^0C for 1 minute 20 seconds. The final cycle was followed by an additional extension step of 72^0C for 7 minutes. Prior to electrophoresis, heteroduplexing was performed via denaturation at 96^0C for 10 minutes, followed by renaturation for 45 minutes at 50^0C. The 204 base pair fragment was electrophoresed in a 9% polyacrylamide gel containing a 30% to 70% urea and formamide (UF) denaturing gradient (100% UF = 7mol/L urea per 40% deionised formamide), at 60^oC for 110 volts overnight, using the Ingeny phorU-2

system (www.ingeny.com). DGGE gels were stained with ethidium bromide and visualised using a UV transilluminator. The *MBL* SNPs, including a previously reported *MBL*N44 SNP [Gabolde et al., 1999], were visualised as differing banding patterns. All SNPs were confirmed via automated sequencing using the dye terminator sequencing kit (Applied Biosystems, www.appliedbiosystems.com).

Studies for Meta-analysis

All studies that examined the association of the *MBL*C52R, *MBL*D54G and *MBL*E57G SNPs with HIV-1 susceptibility were included in the meta-analysis. The search strategy was based on combinations of the following key words, "MBL", "HIV susceptibility", "allele frequency" and "polymorphism" using MEDLINE and EMBASE. The studies were eligible if the aim was to identify HIV-1/AIDS disease risk associations within a specific population group. Studies were characterised by ethnic origin into African or Caucasian.

Statistical Analysis

Estimates and comparisons of allele frequencies were carried out using standard procedures based on the two-sided Fisher's exact test. Comparisons of genotype and allelic distribution were performed using unconditional logistic regression, including population groups and HIV-1 status and their interaction as covariates. For the meta-analysis, pooled odds ratios (OR) and heterogeneity in the associations across the different studies were calculated using fixed effect models. Heterogeneity in the estimates of allelic frequency was tested using logistic regression models with allele as outcome and study as

covariate. The statistical analysis was performed using the statistical program "R" (www.insighful.com) including its package "rmeta" for the meta-analysis.

RESULTS

A single step DGGE assay was designed for the identification of all three functional *MBL* SNPs (C52R, D54G and E57G) and a previously reported silent polymorphism, *MBL*N44 [Gabolde et al., 1999]. Using this assay we determined the allele frequencies within two diverse populations from South Africa to determine the role of these markers in susceptibility to HIV-1. Allele and genotype frequencies are presented in Table 2 and 3, respectively.

South African Population

The distribution of *MBL* variant alleles in both the HIV-1 infected and control group were in Hardy-Weinberg equilibrium across population groups (all P > 0.05). The silent *MBL*N44 SNP was rare in both populations with allele frequencies below 2%. *MBL*C52R was observed in the Caucasians at a frequency of 5%, but was absent in the Africans. The distribution of these two rare SNPs did not differ by HIV-1 status (all P > 0.4). The allelic frequency of *MBL*D54G varied across the two population groups (P < 0.0001), being higher amongst the Caucasians (15%) compared to the Africans (<2%). The allelic frequency did not differ by HIV-1 status in any of the population groups (all P > 0.1). *MBL*E57G was rare in the Caucasian population while the allelic frequency was 18% in the Africans. No significant difference in allelic frequency by HIV-1 status was observed in Caucasians and Africans (both P > 0.7).

Homozygosity was observed for *MBL*D54G in one HIV-1 seronegative Caucasian and for *MBL*E57G in four HIV-1 seropositive and two HIV-1 seronegative Africans. No individual homozygous for *MBL*C52R was detected. Compound heterozygotes included only *MBL*D54G/*MBL*E57G observed in two HIV-1 seropositive Africans. Analysis for the combined functional *MBL* variants and the presence of homozygous/compound heterozygous alleles revealed no significant associations in both population groups studied (all P values> 0.05). Linkage disequilibrium and haplotype measures were not calculated for the three SNPs as the occurrence of compound heterozygotes was generally rare.

Meta-analysis

A meta-analysis of studies (including our own) reporting the three functional *MBL* markers in relation to HIV-1 susceptibility was performed. These studies are described in Table 4 and have been grouped according to ethnic background into Caucasian (three studies with a total of 228 HIV+ and 366 HIV) or African (two studies with a total of 182 HIV+ and 180 HIV-). In these studies, HIV-1 seropositivity was identified by CD4+ T cell counts or western blotting, while all HIV-1 seronegative controls were blood donors tested for their HIV-1 status. As the study of Maas *et al*., 1998 included only HIV-1 seropositive individuals to determine association with progression to AIDS, it was not included in the meta-analysis. Genotyping was performed using restriction fragment length polymorphism analysis, allele-specific oligonucleotide hybridisation, site-directed mutagenesis, DGGE and automated sequencing.

No evidence for heterogeneity in the distribution of the alleles across the populations grouped was found (all P > 0.2). In the African populations (our study combined with the Gabonese study) no significant associations were found for either *MBL*D54G (pooled OR = 2.3, 95% CI = 0.6 - 9.1) or *MBL*E57G (pooled OR = 1.3, 95% CI = 0.9 -1.9). A marginally increased frequency of *MBL*D54G was found in HIV-1 seropositives in the pooled Caucasian population (pooled OR = 1.4, 95% CI = 1.0 - 2.0, *P* = 0.05). This pattern was even more evident when combining the *MBL* variants and considering homozygotes and compound heterozygotes (pooled OR = 6.5, 95% CI = 2.1-20.1).

DISCUSSION

In this study, we examined three well-characterised functional polymorphisms of the *MBL* gene and their relationship to HIV-1 susceptibility in two diverse populations from South Africa. A meta-analysis was performed to determine population-based associations.

Allele frequencies for all *MBL* dimorphic markers within the South African Caucasian population were highly comparable with previous reports. Although the allele frequencies of these markers were comparable to most of the African studies reported, allele frequencies within the African (Xhosa) population from South Africa most closely resembled those reported in the Gabonese study determining association with HIV-1 susceptibility [Mombo et al., 2003]. A South to North/North-West trend of increased allele frequency for the *MBL*E57G SNP within Africa was observed.

No significant associations with disease progression were observed for *MBL*C52R, *MBL*D54G and *MBL*E57G in the two populations represented in our South African study. We did however, find four out of five Caucasian HIV-1 seropositives who are heterozygous for *MBL*D54G to display slow progression, which should be taken into account considering the majority of our HIV-1 infected individuals have unknown disease status. This would support the findings by Maas *et al.* in 1998, where variant *MBL* alleles have been associated with slower progression to AIDS. A meta-analysis of the Caucasian populations however revealed a trend towards association with the *MBL*D54G variant and increased susceptibility to HIV-1 infection. *MBL*D54G is the most common of the three functional variants occurring in Caucasians and the lack of previous associations specifically with this SNP may be due to low sample numbers used for individual studies.

No significant association was observed with the *MBL*E57G SNP and risk for HIV-1 infection in the African and extended African population meta-analysis, although the latter only consisted of two studies. The previous Gabonese study showed that homozygosity or compound heterozygosity for *MBL*E57G is associated with increased risk for HIV-1 infection, while heterozygosity is linked to protection [Mombo et al., 2003]. When we assessed for *MBL*E57G homozygotes and heterozygotes, no significant associations with either increased or decreased susceptibility for HIV-1 infection or disease progression were observed. It has been shown that *MBL*E57G results in lower serum MBL concentrations that could lead to impairment of opsonisation [Super et al., 1989] and therefore an inability to destroy pathogens, including HIV-1. Our findings

thus suggest a hypothesis for selection against the *MBL*E57G variant, as lower allele frequencies are being observed for African populations of Sub-Saharan Africa where the HIV-1/AIDS pandemic is most severe.

All previously reported studies investigating the role of variant *MBL* alleles and HIV-1/AIDS susceptibility involved the combined analysis of the three functional variants to determine association. When considering the combined effect of *MBL* SNPs within our two diverse populations, no significant associations were observed. The meta-analysis of the combined dimorphic markers showed no effect in the pooled African population, but a significant association was observed in the collective Caucasian population for an increased risk for HIV-1 infection, in agreement with two previous studies [Garred et al., 1997; Pastinen et al., 1998].

The controversy that exists in determining the exact role of these functional *MBL* SNPs in HIV-1 susceptibility may be explained by the small study numbers, which may result in spurious findings [Ioannidis et al., 2003]. Therefore the multigenetic nature of the association requires large study numbers to clarify relatively small ORs. Obtaining such well-defined case-control studies in Africa has proved difficult to date [Hayes et al., 2002] and therefore additional studies are required to combine data from populations with the same or similar ethnic heritage. In doing this, one cannot exclude the population divergence that occurs even within closely related populations, as seen throughout Africa. It is also important to note that other polymorphisms within the *MBL* gene, for example the previously reported promoter variants at

positions –550 (H/L variant) and –221 (X/Y variant) [Madsen et al., 1995], could together with the coding functional variants, collectively determine an individual's susceptibility for HIV-1/AIDS. The exact interaction between MBL and HIV-1 is still largely unknown and thus further studies are required to verify the associations observed across well-defined populations.

In this study we have implicated a strong population influence with regards to the role of the functional *MBL* dimorphic markers in HIV-1 susceptibility. We demonstrated a trend towards increasing HIV-1 susceptibility in Caucasian population groups, while no association with susceptibility was observed in Africans even from the same geographical location. These findings emphasise the need for large-scale studies across diverse ethnic groups to determine population-specific association with HIV-1 susceptibility.

REFERENCES

- Dahl M, Tybjaerg-Hansen A, Schnohr P and Nordestgaard BG. A population-based study of morbidity and mortality in mannose-binding lectin deficiency. *J Exp Med* 2004; 199:1391-1399.

- Eisen DP and Minchinton RM. Impact of mannose-binding lectin on susceptibility to infectious diseases. *Clin Infect Dis* 2003; 37:1496-1505.

- Ezekowitz RAB, Kuhlman M, Groopman JE and Bryn RA. A human serum mannose-binding protein inhibits in vitro infection by human immunodeficiency virus. *J Exp Med* 1989; 169:185-196.

- Gabolde M, Muralitharan S and Besmond C. Genotyping of the three major allelic variants of the human mannose-binding lectin gene by denaturing gradient gel electrophoresis. *Hum Mut* 1999; 14:80-83.

- Garred P, Thiel S, Madsen HO, Ryder LP, Jensenius JC and Svegaard A. Gene frequency and partial protein characterization of an allelic variant of mannan binding protein associated with low serum concentrations. *Clin Exp Immunol* 1992a; 90:517-521.

- Garred P, Madsen HO, Kurtzhals JAL, et al. Diallelic polymorphism may explain variations of blood concentration of mannan binding protein in Eskimos, but not in Black Africans. *Eur Journal Immunogenet* 1992b; 19:403-412.

- Garred P, Madsen HO, Balslev U, et al. Susceptibility to HIV infection and progression of AIDS in relation to variant alleles of mannose binding lectin. *Lancet* 1997; 349:236-240.

- Hayes VM, Petersen DC, Scriba TJ, Zeier M, Grimwood A and Janse van Rensburg E. African-based CCR5 single nucleotide polymorphism associated with HIV-1 disease progression. *AIDS* 2002; 16:2229-2230.

- Hayes VM. Genetic diversity of the alpha-1-antitrypsin gene in Africans identified using a novel genotyping assay. *Hum Mut* 2003; 22:59-66.

- Hoal-Van Helden EG, Epstein J, Victor TC, et al. Mannose-binding protein B allele confers protection against tuberculous meningitis. *Pediatr Res* 1999; 45:459-464.

- Holmskov U, Malhotra R, Sim RB and Jensenius JC. Collectins: collagenous C-type lectins of the innate immune defense system. *Immunol Today* 1994; 15:67-73.

- Ioannidis JP, Trikalinos TA, Ntzani EE and Contopoulos-Ioannidis DG. Genetic associations in large versus small studies: an empirical assessment. *Lancet* 2003; 361:567-571.

- Kawasaki N, Kawasaki T and Yamashina I. Isolation and characterization of a mannan-binding protein from human serum. *J Biochem* 1983; 94:937-945.

- Kuhlman M, Joiner K and Ezekowitz RAB. The human mannose-binding protein functions as an opsonin. *J Exp Med* 1989; 169:1733-1745.

- Lipscombe RJ, Sumiya M, Hill AVS, et al. High frequencies in African and non-African populations of independent mutations in the mannose binding protein gene. *Hum Mol Genet* 1992; 1:709-715.

- Lipscombe RJ, Beatty DW, Ganczakowski M, et al. Mutations in the Human Mannose-binding gene: frequencies in several populations. *Eur J Hum Genet* 1996; 4:13-19.

- Maas J, de Roda Husman AM, Brouwer M, et al. Presence of the variant mannose-binding lectin alleles associated with slower progression to AIDS. *AIDS* 1998 12:2275-2280.

- Madsen HO, Garred P, Kurtzhals JAL, et al. A new frequent allele is the missing link in the structural polymorphism of the human mannan-binding protein. *Immunogenetics* 1994; 40:37-44.

- Madsen HO, Garred P, Thiel S, et al. Interplay between promoter and structural gene variants control basal serum level of mannan-binding protein. *J Immunol* 1995; 155:3013-3020.

- Madsen HO, Satz ML, Hogh B, Svejgaard A and Garred P. Different molecular events result in low protein levels of mannan-binding lectin in populations from Southeast Africa and South America. *J Immunol* 1998; 161:3169-3175.

- Malik S, Arias M, Di Flumeri C, Garcia LF and Schurr E. Absence of association between mannose-binding lectin gene polymorphisms and HIV-1 infection in a Columbian population. *Immunogenetics* 2003; 55:49-52.

- Matsushita M and Fujita T. Ficolins and the lectin complement pathway. *Immunol Rev* 2001; 180:78-85.

- Mead R, Jack D, Pembrey M, Tyfield L and Turner M. Mannose-binding lectin alleles in a prospectively recruited UK population. The ALSPAC study team. Avon longitudinal study of pregnancy and childhood. *Lancet* 1997; 349:1669-1670.

- Minchinton RM, Dean MM, Clark TR, Heatley S and Mullighan CG. Analysis pf the relationship between mannose-binding lectin (MBL) gentotype, MBL

levels and function in an Australian blood donor population. *Scand J Immunol* 2002; 56:630-641.

- Mombo LE, Lu CY, Ossari S, et al. Mannose-binding lectin alleles in sub-Saharan Africcans and relation with susceptibility to infections. *Genes Immun* 2003; 4:362-367.

- Pastinen T, Liitsola K, Niini P, Salminen M and Syvänen AC. Contribution pf the CCR5 and MBL genes to susceptibility to HIV type 1 infection in the Finnish population. *AIDS Res Hum Retroviruses* 1998; 14:695-698.

- Prohászka Z, Thiel S, Ujhelyi E, Szlávik J, Bánhegyi D and Füst G. Mannan-binding lectin serum concentrations in HIV-infected patient s are influenced by the stage of disease. *Immunol Lett* 1997; 58:171-175.

- Sastry K, Herman GA, Day L, et al. The human mannose- binding protein gene: exon structure reveals its evolutionary relationship to a human pulmonary surfactant gene and localization to chromosome 10. *J Exp Med* 1989; 170:1175-1189.

- Senaldi G, Davies ET, Mahalingam M, et al. Circulating levels of mannose binding protein in human immunodeficiency virus infection. *J Infect* 1995; 31:145-148.

- Sölder BM, Schultz TF, Hengster P, et al. HIV and HIV-infected cells differentially activate the human complement system independent of antibody. *Immunol Lett* 1989; 22:135-146.

- Sumiya M, Super M, Tabona P, et al. Molecular basis of opsonic defect in immunodeficient children. *Lancet* 1991; 337:1569-1570.

- Super M, Thiel S, Lu J, Levinsky RJ, and Turner MW. Association of low levels of mannan-binding protein with a common defect of opsonisation. *Lancet* 1989; 2:1236-1239.

- Taylor ME, Brickell PM, Craig RK and Summerfield JA. Structure and evolutionary origin of the gene encoding a human serum mannose-binding protein. *Biochem J* 1989; 262:763-771.

- Thiel S, Holmskov U, Hviid L, Laursen SB and Jensenius JC. The concentration of the C-type lectin, mannan-binding protein, in human plasma increases during acute phase response. *Clin Exp Immunol* 1992; 90:31-35.

- Turner MW. Mannose binding lectin: the pluripotent molecule of the innate immune system. *Immunol Today* 1996; 17:532-540.

- Turner MW, Dinan L, Heatley S, et al. Restricted polymorphism of the mannose-binding lectin gene of indigenous Australians. *Hum Molec Genet* 2000; 9:1481-1486.

- Turner MW. The role of mannose-binding lectin in health and disease. *Mol Immunol* 2003; 40:423-429

- Wallis R and Cheng JYT. Molecular defects in variant forms of mannose-binding protein associated with immunodeficiency. *J Immunol* 1999; 163:4953-4959.

Table 1. Minor allele frequencies of *MBL* SNPs within various populations.

Population	Sample size	C52R	D54G	E57G	Reference
Caucasian					
Danish*	123	0.049	0.130	0.020	Garred et al., 1992a, Madsen et al., 1994
Danish [a] *	255	0.045	0.143	0.022	Garred et al., 1997
Danish	9245	0.076	0.145	0.017	Dahl et al., 2004
British	98	0.020	0.168	0	Lipscombe et al., 1992; 1996
British	302	0.066	0.144	0.015	Mead et al., 1997
Finnish [a]	305	0.026	0.121	0.015	Pastinen et al., 1998
Dutch [a]	131	0.069	0.145	0.008	Maas et al., 1998
Australians	236	0.076	0.144	0.030	Minchinton et al., 2002
Caucasian [a] (South Africa)	70	0.050	0.150	0.014	This study
African					
Kenyans**	56	-	0.009	-	Garred et al., 1992b
Kenyans**	66	0.045	0.030	0.235	Madsen et al., 1994
Gambians	99 Babies 100 Adults	- 0.010	0.005 0	0.232 0.288	Lipscombe et al., 1992; 1996
San (Namibia)	58	0	0.034	0.069	Lipscombe et al., 1996
Gabonese [a] ***	188	0	0.016	0.157	Mombo et al., 2003
Gabonese***	214	0	0.026	0.187	Mombo et al., 2003
Central Africans***	266	0	0.007	0.202	Mombo et al., 2003
West Africans	311	0	0.024	0.307	Mombo et al., 2003
Southeast Africans (Mozambique)	154	0	0	0.240	Madsen et al., 1998
Xhosa (South Africa)	45	0	0	0.267	Lipscombe et al., 1996
Xhosa [a] (South Africa)	174	0	0.017	0.175	This study
Other					
Eskimos	73	0	0.130	0	Garred et al., 1992b, Madsen et al., 1994
Chinese	123	0.010	0.114	0	Lipscombe et al., 1996
Papua New Guinea (SW Pacific)	49	0	0.071	0	Lipscombe et al., 1996
Vanuatu (SW Pacific)	112	0	0.013	0	Lipscombe et al., 1996
Chiniguanos Indians (South America)	43	0	0.420	0.010	Madsen et al., 1998
Mapuche Indians (South America)	25	0	0.460	0.010	Madsen et al., 1998
Cape Coloured (South Africa)	101	0.020	0.110	0.060	Hoal-Van Helden et al., 1999
Walpiri Australians	190	0.003	0	0	Turner et al., 2000
Central Desert Australians	103	0.005	0	0	Turner et al., 2000
Columbian	278	0.036	0.120	0.036	Malik et al., 2003

[a] Indicates studies for the screening of *MBL* SNPs to determine association with HIV-1/AIDS.

*/**/*** Indicates that the same individuals may have been included for both studies.

Table 2. Allele frequencies of *MBL* variants among 135 HIV-1 seropositive and 109 HIV-1 seronegative individuals from two South African population groups ordered according to the codon at which they occurred.

MBL SNP	Africans		Caucasians	
	HIV + (n = 228)	HIV - (n = 120)	HIV + (n = 42)	HIV - (n = 98)
N44	2 (0.009)	2 (0.017)	0	1 (0.010)
C52R	0	0	3 (0.071)	4 (0.041)
D54G	6 (0.026)	0	7(0.167)	14 (0.143)
E57G	42 (0.184)	19 (0.158)	0	2 (0.020)

n, number of alleles; HIV+, HIV-1 seropositive; HIV-, HIV-1 seronegative

Table 3. Distribution of *MBL* genotypes among 135 HIV-1 seropositve and 109 HIV-1 seronegative individuals from two South African population groups.

Genotype No. of individuals	HIV-1 seropositive		HIV-1 seronegative	
	Africans 114	Caucasians 21	Africans 60	Caucasians 49
Homozygous normal (52C/C, 54G/G, 57G/G)	72	11	43	30
Heterozygous mutant	36	10	15	18
52C/T	0	3	0	4
54G/A	4	7	0	12
57G/A	32	0	15	2
Homozygous mutant	4	0	2	1
54A/54A	0	0	0	1
57A/57A	4	0	2	0
Compound heterozygous	2	0	0	0
54A/57A	2	0	0	0

Table 4. Characteristics of studies included in testing for heterogeneity of allele frequencies and the meta-analysis grouped according to population background.

Country	Eligible subjects		Association	Reference
	HIV+	HIV -		
Caucasian				
Denmark	96	123	Homozygosity for *MBL* variant alleles associated with increased risk of HIV-1 infection. The *MBL* variant alleles in both the heterozygous and homozygous state is also associated with shorter survival time after the diagnosis of AIDS.	Lipscombe et al., 1996
Finland	111	194	Homozygosity for *MBL* variant alleles was significantly increased in the HIV-1 infected groups.	Pastinen et al., 1998
South Africa	21	49	NONE	This study
TOTAL	**228**	**366**	**Significant association for combined *MBL* variant alleles associated with increased risk for HIV-1 infection. Specifically a significant association with homozygosity for *MBL* variants and increased risk for HIV-1 infection. Also a trend towards significance specifically for *MBL*D54G and increased risk to HIV-1 infection.**	
African				
Gabon	68	120	Homozygosity and/or compound heterozygosity for the *MBL*E57G allele is associated with an increased risk of HIV-1 infection.	Mombo et al., 2003
South Africa	114	60	NONE	This study
TOTAL	**182**	**180**	**NONE**	

HIV+, HIV-1 seropositive; HIV-, HIV-1 seronegative

Chapter 5

Discussion

Discussion

Two theories for the manner in which genetic variants influence common traits have been proposed. The first is known as the common disease/common variant hypothesis based on the idea that common susceptibility alleles for common traits arose before the differentiation of populations. Alleles at a few loci will therefore occur at high frequency across diverse ethnic groups and be useful when determining genetic associations that are identical for various populations. However, when the common disease/common variant hypothesis does not hold, it's possible that less frequent or rather population-based alleles influence susceptibility to complex traits. This could be the case in particularly the older understudied African populations who display higher levels of genetic diversity [Goldstein and Chikhi, 2002; Tishkoff and Williams, 2002; Tishkoff and Verrelli, 2003; Foster and Sharp, 2004].

The importance of host genetic factors in determining susceptibility to HIV-1 infection and progression to AIDS has been advanced by population-based association studies investigating the roles of various candidate genes. These case-control studies involve the comparison of unrelated HIV-1 seropositives to population-matched HIV-1 seronegatives and have resulted in the identification of genetic variants associated with influencing individual risks for HIV-1/AIDS. Previous studies were generally restricted to Caucasian-based populations and only more recently have been extended to African-Americans and other admixed populations [Reviewed in Carrington et al., 2001; Hogan and Hammer, 2001; Dean et al., 2002; Anastassopoulou and Kostrikis, 2003; O'Brien and Nelson, 2004; Winkler et al., 2004; Kaslow et al., 2005]. The study of host

genetic factors and HIV-1/AIDS in populations of Sub-Saharan Africa where the rampant pandemic accounts for 60% of all worldwide infections has however been largely limited.

The case-control study sample presented in this dissertation consisted of individuals belonging to an African population of Xhosa descent from South Africa (see Chapter 1.2). The study approach involved identifying and determining the distribution of previously reported and novel mutations for known candidate genes in this relatively homogenous ethnic group (see Chapter 2 to 4 and Appendix A to D). Genetic variants occurring in the Xhosa ethnic group were screened for in additional South African Caucasian individuals of European descent. It has been suggested that caution should be taken in extrapolating genetic associations from one population to another. This includes when studying various South African populations [Corfield and Brink, 2002] and when considering genetic factors influencing HIV-1/AIDS susceptibility [Ramaley et al., 2002] (see Chapters 2 to 4 and Appendix A to D).

The studies in this dissertation mainly focused on HIV-1 infection risk rather than disease progression to AIDS. This was due to insufficient information available regarding the date of seroconversion and clinical staging of the HIV-1 infected individuals. Most of these persons were tested for HIV-1 at time points after they had already presented with AIDS-related symptoms. Their inclusion in the study sample was therefore only based on their HIV-1 seropositive status and no additional information regarding possible dates of HIV-1 seroconversion was required. The population-matched HIV-1 uninfected controls are blood

donors and were recruited for the study by considering their HIV-1 seronegative status. It is unknown whether any of these HIV-1 seronegative controls have been previously exposed to HIV-1 and due to confidentiality, information regarding their health status and age was not requested.

DGGE assays were designed for the screening of both previously reported and novel genetic variants, while the TaqMan allelic discrimination assays provided high-throughput genotyping for SNPs previously implicated in HIV-1/AIDS pathogenesis (see Chapter 1.3). Both these methodologies offered reliable and reproducible mutation detection results, provided that the assays were optimally standardised. Comprehensive statistical analysis of the study sample was achieved using computational programs (see Chapter 1.2) to determine heterogeneity between the cases and controls and possible significance for establishing HIV-1/AIDS associations (see Chapters 2 to 4).

The study sample size is an important factor to consider when deciding on the validity of a genetic association. Although the sample numbers for studies in Chapters 2 to 4 are not high enough for classification as a large-scale investigation, novel African-based genetic variants were observed. The Caucasian-based *CCR5*Δ32 HIV-1/AIDS restriction mutation was completely absent in the Xhosa ethnic group. Therefore, the possibility of an African-based single variant or more likely a number of polymorphic sites that are collectively associated with susceptibility to HIV-1/AIDS needs to be explored extensively in diverse African ethnic groups.

Future prospects

A model quantifying the rate at which AIDS restriction variants delaying disease progression will undergo natural selection in South Africa has been proposed [Schliekelman et al., 2001]. South Africa has the highest HIV prevalence in the world and the model projections under HIV-1 induced selection suggest that within 100 years resistant genotype frequencies will increase from 40 to 53% and causative genotype frequencies will decrease from 20 to 10%, resulting in an increased average time to AIDS onset from 7.8 to 8.8 years. This model is however based on *CCR5* variants influencing HIV-1/AIDS susceptibility in African Americans and it is now known that the same associations cannot be assumed for African-based populations from Sub-Saharan Africa. A large study of an Ugandan population found that *CCR5* variants were not associated with determining HIV-1/AIDS risk profiles and therefore suggested that the *CCR5* gene may not be subjected to rapid evolutionary change in an African setting [Ramaley et al., 2002]. Novel associations were identified between *CCR5* promoter variants and HIV-1 infection risks for the Xhosa population from South Africa (see Chapter 2.1), but larger study samples with known disease progression for HIV-1 seropositives are required before further considering projected models for HIV-1 induced selection.

The genetic variants identified and analysed in the Xhosa ethnic group provide a strong basis for further investigation of large informative African-based study samples. Future studies should however not only consider associations with HIV-1 infection risks, but also establish possible significance with disease progression to AIDS. The latter has been proven difficult, but could eventually

be overcome by the recent improvement of HIV/AIDS awareness programs and the access to service centers that offer not only free HIV-1 testing, but also the support and care for those individuals who are HIV-1 seropositive. Future collaborations with healthcare workers who are specifically trained in the management of those persons living with HIV/AIDS, together with the HIV-1 seropositive individual's informed consent, could provide a study sample that overcomes the present limitations for AIDS progression association studies.

The studies presented in this dissertation (Chapter 2 to 4) showed association between HIV-1 infection risk and a few SNPs, including *CCR5*-2733A>G*, *CCR5*-2135C>T*, *CX3CR1*V249I (Chapter 2.1)*, *SDF1*-3'G>A (Chapter 2.2), *IL10*-819C>T, *IL10*-592C>A (Chapter 3.2) and *MBL*D54G (Chapter 4), and specific haplotypes for *IL4* and *IL10* (Chapter 3.2). The possibility exists that these SNPs and haplotypes may be in LD with other previously reported or unknown genetic variants. Functional studies based on the existing knowledge of those genetic variants associated with influencing HIV-1/AIDS susceptibility could however refine the search for the causative or resistance genetic variants. Future research would therefore be required for determining the underlying functional effects of all common genetic variants and their extended haplotypes. An example would be determining the functional consequence of the novel *IL4* haplotype (comprising *IL4*-589 C and *IL4*-33 T alleles) associated with increased risk for HIV-1 infection in the Xhosa ethnic group (see Chapter 3.2).

There is an additional approach for determining genetic associations that will be performed in the near future. It is referred to as genome-wide association

studies, which involves the genotyping of a dense set of SNPs across the entire genome to identify common genetic variation playing a role in determining common trait susceptibility. The choice of SNPs for these studies is dependent on the frequency of all common variation in the specific population studied. Many SNPs have alleles that are in strong linkage disequilibrium with other closely lying SNP alleles and therefore comprehensive genome-wide association studies can be based on a selection of evenly spaced tagSNPs representing variation across relatively large regions. A higher density of variants would need to be genotyped for comprehensive analysis of regions that show low LD. A larger number of tagSNPs are therefore required in African populations with more variation and less LD. Identifying these tagSNPs for different populations is a main goal of the HapMap project which is due for completion within the next two years [Hirschhorn and Daly, 2005; Wang et al., 2005].

The genome-wide association studies could provide a more cost-effective means to determine common genetic variation compared to the candidate gene approach. Although the candidate gene studies have had a considerable amount of success, only a limited number of genes have been investigated. The feasibility of genome-wide analysis of common SNPs will however depend on variant identification in multiple genes with distinct influences on susceptibility to complex traits. Complete genome sequencing for many case and control individuals could provide a comprehensive overview of the genetic host factors underlying common traits. This would include the screening of both coding and non-coding variants and the identification of both rare and common

variants with functional roles. Presently, this approach is not practical, particularly due to the high costs involved in automated sequence technology [Hirschhorn and Daly, 2005; Wang et al., 2005].

To date, the discovery of functional genetic variants have resulted in an improved understanding of host response to HIV-1 and raises the prospects of therapeutic intervention by targeting and disrupting complex interactions between HIV-1 and host proteins during virus exposure and post virus entry [Nolan et al., 2004; O'Brien and Nelson, 2004]. The knowledge of host genetic factors for establishing HIV-1/AIDS risk profiles specific for African populations of Sub-Saharan Africa could contribute to optimising future approaches for preventing HIV-1 infection (vaccine development) and particularly long-term HIV management (therapies administered are based on patient's genetic makeup) in pandemic-stricken countries such as South Africa.

References

- Anastassopoulou CG and Kostrikis LG. The impact of human allelic variation on HIV-1 disease. *Curr HIV Res* 2003; 1:185-203.

- Carrington M, Nelson G and O'Brien SJ. Considering genetic profiles in functional studies of immune responsiveness to HIV-1. *Immunol Lett* 2001; 79:131-40.

- Corfield VA and Brink PA. Case:control association studies of complex disease: is the search for genetic determinants a fishing expedition? *Cardiovasc J S Afr* 2002; 13:93-95.

- Dean M, Carrington M and O'Brien SJ. Balanced polymorphism selected by genetic versus infectious human disease. *Annu Rev Genomics Hum Genet* 2002; 3:263-292.

- Foster MW and Sharp RR. Beyond race: towards a whole-genome perspective on human populations and genetic variation. *Nat Rev Genet* 2004; 3:790-796.

- Goldstein DB and Chikhi L. Human migrations and population structure: What we know and why it matters. *Ann Rev Genomics Hum Genet* 2002; 3:129-152.

- Hirschhorn JN and Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005; 6:95-108.

- Hogan CM and Hammer SM. Host determinants in HIV infection and disease, Part 2: Genetic factors and implications for antiretroviral therapeutics. *Ann Intern Med* 2001; 134:978-996.

- Kaslow RA, Dorak T and Tang J. Influence of host genetic variation on susceptibility to HIV type 1 infection. *J Infect Dis* 2005; 191: S68-S77.

- Nolan D, Gaudieri S, John M and Mallal S. Impact of host genetics on HIV disease progression and treatment: new conflicts on an ancient battleground. *AIDS* 2004; 18:1231-1240.

- O'Brien SJ and Nelson GW. Human genes that limit AIDS. *Nature Genet* 2004; 36: 565-574.

- Ramaley PA, French N, Kaleebu P, et al. Chemokine-receptor genes and AIDS risk. *Nature* 2002; 417:140.

- Schliekelman P, Garner C and Slatkin M. Natural selection and resistance to HIV. *Nature* 2001; 411:545.

- Tishkoff SA and Williams SM. Genetic analysis of African populations: Human Evolution and complex disease. *Nat Rev Genet* 2002; 3:611-621.

- Tishkoff SA and Verrelli BC. Patterns of human genetic diversity: Implications for human evolutionary history and disease. *Annu Rev Genomics Hum Genet* 2003; 4:293-340.

- Winkler C, An P and O'Brien SJ. Patterns of ethnic diversity among the genes that influence AIDS. *Hum Mol Genet* 2004; 1:R9-19.

- Wang WYS, Barratt BJ, Clayton DG and Todd JA. Genome-wide association studies: Theoretical and practical concerns. *Nat Rev Genet* 2005; 6:109-118.

Appendix A

Novel mutations identified using a comprehensive

CCR5-denaturing gradient gel electrophoresis assay

AIDS 2001; 15:171-177

Novel mutations identified using a comprehensive *CCR5*-denaturing gradient gel electrophoresis assay

Desiree C. Petersen[a], Maritha J. Kotze[a], Michele D. Zeier[b], Ashraf Grimwood[c], Deon Pretorius[d], Eftyhia Vardas[e], Estrelita Janse van Rensburg[d] and Vanessa M. Hayes[d,f]

Background: Most mutations detected for the gene for CC chemokine receptor 5 (CCR5) are either relatively specific to different population groups or rarely observed in Africans.

Objectives: To develop a comprehensive mutation detection assay for the entire coding region of *CCR5* and to identify novel mutations that may play a role in genetic susceptibility to HIV-1 infection, within the diverse South African population.

Design: The study cohort consisted of 103 HIV-seropositive patients and 146 HIV-seronegative controls of predominantly African descent.

Methods: A mutation detection assay for the entire coding region of *CCR5* was designed; this included amplification of part of the coding region of *CCR2*. The assay was based on denaturing gradient gel electrophoresis (DGGE) and allowed the complete analysis of samples from 10 individuals per denaturing gel.

Results: The use of the *CCR5*-DGGE assay led to the identification of seven novel and six previously reported mutations. All novel mutations, including a common polymorphism at codon 35, occurred exclusively in non-Caucasians, indicating possible African origin.

Conclusion: A comprehensive DGGE mutation detection assay has been developed for the entire coding region of *CCR5*. Application of this assay resulted in the identification of novel *CCR5* mutations, which may have a significant effect on the normal functioning of CCR5 and thus contribute to host variability and susceptibility to HIV-1 infection and/or progression to AIDS within this population.

© 2001 Lippincott Williams & Wilkins

Introduction

Various chemokine receptors have been identified as co-receptors necessary for cellular infection by HIV [1–4]. The CC chemokine receptor 5 (CCR5) is a seven transmembrane G-coupled protein consisting of 352 amino acid residues; it binds the β-chemokines RANTES and macrophage inflammatory protein 1α and 1β [5]. It is also recognised as the principle co-receptor for the macrophage-tropic (M-tropic) strains or non-syncytium-inducing (NSI) strains of HIV-1 and facilitates fusion of the viral envelope protein with CD4+ cells during the asymptomatic phase of infection [1,2,6,7]. The expression of CCR5 at the cell surface may, therefore, have a direct influence on the individual variability and susceptibility to HIV-1 infection.

CCR5 is located at band p21 of chromosome 3 [8] and

comprises four exons and two introns, of which exon 4 contains the entire open reading frame [9]. The most widely studied *CCR5* mutation is a 32 base pair (bp) deletion (*CCR5Δ32*) within the coding region, which results in premature termination of translation and the formation of a truncated protein [8,10]. Individuals homozygous for *CCR5Δ32* have been shown to have some protection against HIV-1 infection, although this protection is not absolute. Heterozygous carriers display partial protection against HIV-1 infection, and evidence exists that a single deletion mutant in HIV-seropositive individuals may slow progression to AIDS [8,10−13]. Population surveys have shown that *CCR5Δ32* is largely confined to Caucasians (allele frequency of 0.092) and is extremely rare in Africans [10,12,14]. This suggests the presence of other possibly protective/causative mutations in the African populations and underscores the importance of comprehensive *CCR5* mutation analysis in the diverse South African populations. Several other genetic variants in the *CCR5* coding region have been described. However, their role in HIV-1 infection or progression to AIDS could not be deduced because of the low allelic frequencies of these mutations in the population groups studied [11,15−18].

In this study, we describe a comprehensive *CCR5* mutation detection assay for the entire coding region of the gene, using denaturing gradient gel electrophoresis (DGGE). This method was developed by Fischer and Lerman in 1983 [19]. It a polymerase chain reaction (PCR)-based method and is believed to be the most powerful of the pre-screening methods of mutation detection currently available. The technique involves the differential melting of double-stranded DNA molecules in a gradient with an increasing concentration of urea and formamide. The addition of a guanine and cytosine (GC)-rich fragment (GC-clamp), introduced during fragment amplification, prevents total strand dissociation and allows for the detection of single base mutations, making DGGE virtually 100% sensitive [20,21].

Using this assay, 103 HIV-seropositive patients and 146 healthy controls were screened for mutations in the coding region of *CCR5*. Our results obtained in the unique South African population are presented in this study.

Methods

Sample population
Blood samples were drawn from 103 HIV-seropositive patients (35 male and 68 female) residing in the Western Cape of South Africa (Tygerberg Hospital and Woodstock Chapel Street Community Health Clinic).

Disease progression of the majority of these individuals was unknown. Blood samples from 146 HIV-seronegative controls (56 male and 91 female) were obtained from the Western Province Blood Transfusion Service, South Africa. The study cohort consisted of Africans, predominantly Xhosa (70 HIV positive and 64 HIV negative), Coloureds (26 HIV positive and 72 HIV negative), Caucasians (seven HIV positive and two HIV negative) and Asians (eight HIV negative). An additional nine samples from seronegative 'high-risk' commercial sex workers of Zulu descent were obtained from KwaZulu-Natal, South Africa. In this study, 'African' refers to South Africans of central African descent; 'Coloured' refers to individuals of mixed ancestral descent, including San, Khoi, African Negro, Madagascan, Javanese and European origin; and 'Caucasian' refers to South Africans of European descent, mainly Dutch, French, German and British origin. Informed consent for the study was obtained from all participants. The Ethics Review Committee of the University of Stellenbosch approved the study protocol.

Primer design
DGGE PCR primers (Table 1) were designed for the entire coding region, including the donor/acceptor splice site of intron 3/exon 4, of *CCR5*, using the melt 87 computer program [22] and conditions described by Wu *et al.* [23]. The region to be analysed was divided into six overlapping PCR fragments (A−F). In order to prevent complete strand dissociation during amplification, a GC-clamp was added to the 5′ end of one of the primers in each primer set. An additional stretch of 10 GC or AT (adenine, thyamine) nucleotides were added to the 5′ end of the non-clamped primer (fragment B) or between the GC-clamp and the primer (fragments C and D), respectively, to ensure a single melting domain and thus optimal mutation detection of the fragments.

DNA amplification
Genomic DNA was extracted using conventional methods and amplified using the DGGE primer sets. Each PCR reaction mixture of 50 μl in total volume contained 100 ng genomic DNA, 0.1 mmol/l of each deoxyribonucleoside triphosphate (dNTP), 20 pmol of each primer, 2.5 mmol/l $10\times$ Mg^{2+} reaction buffer and 0.5 units DNA Taq polymerase (Boehringer Mannheim, Mannheim, Germany). Amplification was performed using the following cycling conditions; an initial denaturation at 96°C for 3 min, followed by 32 cycles of denaturation at 96°C for 45 s, annealing for 1 min, and elongation at 72°C for 1 min. The last cycle was followed by an additional extension step of 72°C for 10 min. Amplicons to be subjected to DGGE analysis required an additional heteroduplexing step, which involves denaturation at 96°C for 10 min, followed by renaturation for 45 min at the annealing

temperature of the amplification. The amplified products were checked by electrophoresis of 5 μl (10%) of each sample in a 2% agarose gel.

Denaturing gradient gel electrophoresis

The DGGE conditions were optimized using conditions previously described for broad-range DGGE analysis [24]. The six amplicons were pooled into three groups (group 1: fragments A, D and C; group 2: fragments E and B, and group 3: fragment F) and electrophoresed in a 9% polyacrylamide gel containing a 30−70% urea and formamide denaturing gradient (100% is 7 mol/l urea in 40% deionized formamide) at 59°C and 110 V overnight. Gels were stained with ethidium bromide and photographed under an ultraviolet transilluminator. The optimally designed *CCR5*-DGGE assay allows for the complete analysis of 10 patients per denaturing gel (Fig. 1).

DNA sequencing and mutation confirmation

Amplified products showing aberrant DGGE banding patterns were subjected to automated sequencing using a non-GC-clamped primer and the dye terminator sequencing kit of Applied Biosystems (www.applied-biosystems.com). Confirmation of commonly occurring polymorphisms were performed by mixing samples showing similar DGGE banding patterns, followed by a heteroduplexing step before electrophoresis on a denaturing gel [25]. Samples showing additional heteroduplex bands were subjected to sequencing.

Statistical analysis

Allele frequencies were determined by allele counting. Testing for significance of heterogeneity in mutation frequencies among HIV-seropositive and HIV-seronegative subjects was based on both the chi-square and Fisher's exact tests. Yates' corrections were applied to improve the approximation of the chi-square test.

Results

The DGGE PCR primer sets and annealing temperatures are shown in Table 1. Application of the *CCR5*-DGGE assay identified seven novel point mutations and six previously reported mutations, which are listed in Table 2 according to the codon in which they occur.

Of the seven novel mutations detected in this study, four may ultimately result in structural changes in the CCR5 protein. The first, a nonsense mutation at codon 225 (CGA-TGA), results directly in the formation of a truncated protein through conversion of the amino acid Arg to a premature stop codon. Both codon 2 (GAT-GTT) and codon 225 (CGA-CAA) mutations result in a non-conservative amino acid change (replacement of one amino acid by another with different chemical properties) from Asp to Val and from Arg to Glu, respectively. The fourth mutation at codon 107 (CTC-TTC), although resulting in a conservative amino acid change (Leu to Phe), involves the inclusion of an aromatic side chain, which may have structural and/or functional implications. All the individuals who presented with the codon 107 (CTC-TTC) mutation also presented with the codon 225 (CGA-TGA) mutation; no individual was found to have only one of these two mutations. The remaining three novel mutations were all silent mutations, occurring at codons 35, 89 and 162, respectively. A high allelic frequency for the codon 35 (CCG-CCA) polymorphism was detected in both the HIV-seropositive and HIV-seronegative in-

Table 1. *CCR5* primer sets and experimental conditions for polymerase chain reaction amplification and DGGE.

Fragment	Amplimers 5′−3′	Size (bp)	Temperature (°C)	
			Melting	Annealing
A	[40GC]TGGAGGGCAACTAAATACAT CGATTTGCTTCACATTGATT	196	67	54
B	[10GC]ATTATACATCGGAGCCCTGC [40GC]AGCATAGTGAGCCCAGAAGG	280	74	60
C	[40GC][10AT]CTGGCCATCTCTGACCTGTT GATGATTCCTGGGAGAGACG	332	73	60
D	[40GC][10AT]ACTTGGGTGGTGGCTGTGTT CATTTCGACACCGAAGCAGA	276	72	60
E	TCATGGTCATCTGCTAGTCG [40GC]GGTGTTCAGGAGAAGGACAA	192	72	58
F	[40GC]TTCTCTTCTGGGCTCCCTAC GTCACCAGCCCACTTGAGTC	390	74	60

bp, base pair.
GC-clamps used were: [40GC], CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCCG;
                                   [10GC], CGCCGCCGCG
AT-stretch used was [10AT], TATAATATTA.

Table 2. *CCR5* mutations detected in 103 HIV-seropositive patients and 146 HIV-seronegative controls, ordered according to the codon in which they occurred.

Mutation	Base change	Fragment	Africans		Coloureds		Caucasians	
			HIV+ (n = 140)	HIV− (n = 128)	HIV+ (n = 52)	HIV− (n =144)	HIV+ (n = 14)	HIV− (n = 4)
D2V[b]	GAT-GTT	A	1 (0.007)	0	0	1 (0.007)	0	0
P35[b]	CC**G**-CC**A**	B	6 (0.043)	9 (0.070)	9 (0.173)	10 (0.069)	0	0
L55Q	CTG-CAG	B	0	0	0	0	1 (0.071)	0
S75	TC**T**-TC**C**	B	1 (0.007)	0	0	0	0	0
Y89[b]	TAT-TA**C**	C	0	0	0	1 (0.007)	0	0
L107F[b,c]	**C**TC-**T**TC	C	1 (0.007)	2 (0.016)	1 (0.019)	0	0	0
P162[b]	CC**A**-CC**G**	D	1 (0.007)	0	0	0	0	0
Δ32 (185)	–	D	0	0	1 (0.019)	5 (0.035)	1 (0.071)	0
R223Q	C**G**G-C**A**G	E	0	0	0	1 (0.007)	0	0
R225X[b,c]	**C**GA-**T**GA	E	1 (0.007)	2 (0.016)	1 (0.019)	0	0	0
R225Q[b]	CG**A**-CA**A**	E	0	0	0	1 (0.007)	0	0
A335V	**G**CA-**C**TA	F	4 (0.029)	2 (0.016)	1 (0.019)	2 (0.014)	0	0
Y339F	TAC-**T**TC	F	0	1 (0.008)	0	0	0	0

[a]No mutations detected in eight Asian HIV seronegative controls.
[b]Novel mutation identified in this study.
[c]Mutations occurring together in patients.
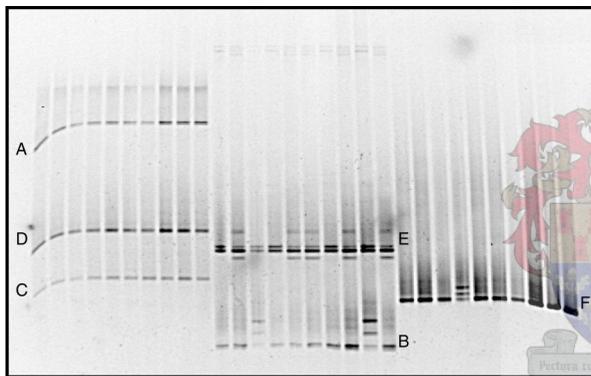
Fig. 1. Denaturing gradient gel electrophoresis banding pattern covering the entire coding region of the gene for the CC chemokine receptor 5 (CCR5) from six pooled amplicons of 10 patients; for details see the text. Lanes 1–10 contain group 1 (fragments A, D and C); lanes 11–20, group 2 (fragments E and B) and lanes 21–30, group 3 (fragment F). The multiple bands depicted for fragment E are explained in further detail in Fig. 2a. Individuals in lanes 13 and 19 are heterozygous for the codon 35 polymorphism (fragment B), and the individual in lane 24 is heterozygous for the codon 335 polymorphism (fragment F).

dividuals from African and Coloured descent (Fig. 1) and occurred in a homozygous state in a single HIV-seropositive Coloured female. This novel polymorphism was totally absent in Caucasians. Because of the low numbers of Caucasian individuals in the study cohort, further screening of 28 Caucasians for codon 35 polymorphism confirmed the absence of this mutation in this population group. All the above mentioned novel mutations were found exclusively in individuals from African or Coloured ethnic background.

The six previously reported mutations observed in this study include the most commonly studied, *CCR5Δ32* at codon 185. This deletion mutation was observed heterozygously in one HIV-seropositive Coloured and one Caucasian and in five HIV-seronegative Coloureds, while it was absent in the Africans studied. Three non-conservative mutations at codons 55 (Leu to Glu), 223 (Arg to Glu) and 339 (Tyr to Phe), all previously reported by Ansari-Lari *et al.* [15], were observed in one HIV-seropositive Caucasian, one HIV-seronegative Coloured and one HIV-seronegative African, respectively. The silent mutation at codon 75, previously reported by Carrington *et al.* [16], was found to be present in one HIV-seropositive African. The codon 335 polymorphism, involving an amino acid change from Ala to Val, has also been previously reported by Ansari-Lari *et al.* [15]. In our study, we detected this polymorphism in four HIV-seropositive and four HIV-seronegative individuals of African and Coloured descent (Fig. 1).

All HIV-seropositive and HIV-seronegative individuals presented with additional DGGE bands in fragment E, lower (L) and/or upper (U), in combination with the normal (N) *CCR5* band (Fig. 2a). Heteroduplex bands at a low percentage of urea and formamide were also noted. Excision of these aberrant bands from the gel, followed by direct sequencing, revealed 11 nucleotide variations occurring in the lower band and an additional twelfth variation was included in the upper band. Blasting the mutant sequence, using the Genbank database (www.ncbi.nlm.nih.gov), revealed that this sequence forms part of the gene for chemokine receptor 2 (*CCR2*), including codons 217–267 (Fig. 2b). The polymorphism at codon 260 (AAC to AAT)
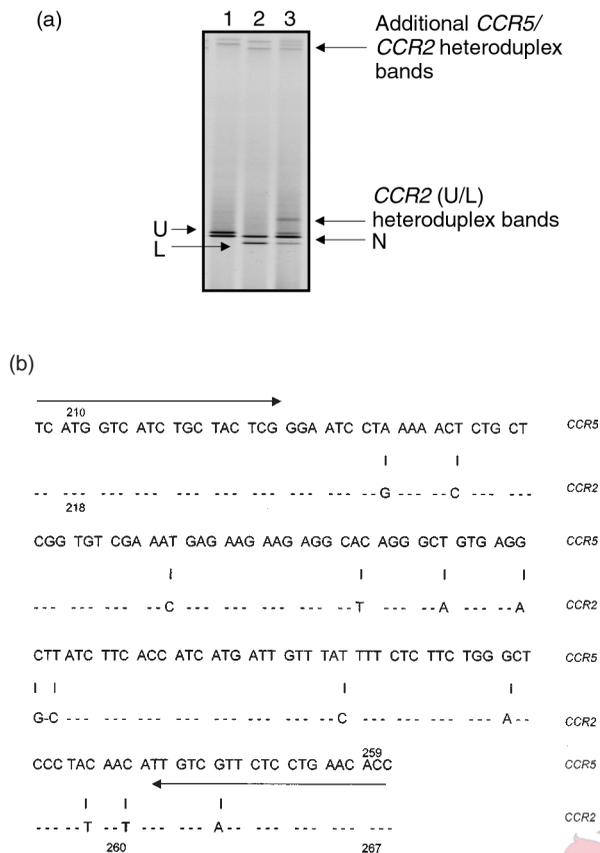
Fig. 2. Fragment E of CC chemokine receptor 5 (*CCR5*) gene. (**a**) Denaturing gradient gel electrophoresis (DGGE) banding pattern of fragment E in combination with codons 217–267 of *CCR2*. All samples (lanes 1–3) are homozygous normal (N band) for *CCR5*. Samples were either homozygous or heterozygous for the *CCR2* codon 260 polymorphism. Lane 1, homozygous 260-T (U band); lane 2, homozygous 260-C (L band); lane 3, heterozygous 260-C/T (U and L homoduplex and heteroduplex bands). Additional *CCR5/CCR2* heteroduplex bands melt at a low percentage of denaturant in the DGGE gel because of the high number of nucleotide mismatches. (**b**) *CCR5* and *CCR2* sequences amplified using DGGE fragment E primer set (arrows). The codon 260 (AAC/AAT) polymorphism of *CCR2* is in indicated in bold.

of *CCR2* showed an allelic frequency of 0.62 for T and 0.38 for C within our population group. All mutations occurring within fragment E were confirmed as being *CCR5* mutations by recognition of the DGGE band intensities and confirmation by excision of additional heteroduplex bands from the gel followed by direct sequencing.

Discussion

In this study, we describe a comprehensive and efficient mutation detection assay for the entire coding region of *CCR5*. This assay, based on DGGE with its

virtual 100% sensitivity, allows for the complete analysis of 10 patients per denaturing gel. The assay was used to screen for possible novel *CCR5* sequence variants in a predominantly African and Coloured HIV-seropositive and HIV-seronegative cohort from South Africa. Most studies to date have restricted their analysis to the *CCR5Δ32* mutation, which, although fairly common in Caucasians, is extremely rare in the African populations. Comprehensive analysis of *CCR5* is, therefore, of vital importance in the diverse South African population.

Seven novel CCR5 mutations were identified in the African and Coloured populations. No novel mutations were identified in the Caucasian or Asian populations, although numbers were small. Novel mutations at codons 107 and 225 (CGA-TGA), which occur simultaneously, and at codons 2 and 225 (CGA-CAA) may affect the functioning of CCR5 and thus provide possible protection against HIV infection and/or progression to AIDS. One cannot, however, exclude the possibility that the three novel 'silent' mutations (codons 35, 89 and 162) detected in this study affect disease progression by altering regulatory elements that affect RNA splicing [26,27]. The novel codon 35 polymorphism (CCG to CCA) occurred at an allelic frequency of 0.06 and 0.1 in the African and Coloured populations, respectively, and was absent in Caucasians, indicating that it has a definite African origin. Although numbers are small, a significantly higher allelic frequency (Fisher exact $P = 0.017$; $\chi^2 = 5.02$, 1 df, $P = 0.025$) was observed in the Coloured HIV-seropositive individuals compared with the HIV-seronegative controls, while frequency was similar in the African HIV-infected and control groups. The significance of this finding warrants further investigation. The Coloured female homozygous for the polymorphism showed normal disease progression (progression to AIDS within 8 to 10 years after HIV infection). Because of the lack of clinical information regarding disease progression of majority of the HIV-seropositive patients, the potential consequences of the different novel mutations could not be evaluated and no significant associations could be made. It is, therefore, necessary to obtain updated reports on the disease progression of all the HIV-seropositive patients.

The *CCR5Δ32* mutation, generally restricted to Caucasians, was found to be absent in the 134 Africans studied, while it occurred at an allelic frequency of 0.03 in the Coloured population. The presence of this deletion mutation in the Coloured population may be a reflection of admixture with people of Caucasian descent [14]. The presence of *CCR5Δ32* in Coloureds also provides evidence that no genetic incompatibility between ethnic groups exists for this mutation [28]. One of the two HIV-seropositive individuals who were heterozygous for *CCR5Δ32* was a Coloured male

asymptomatic 7 years after infection; the other was an asymptomatic Caucasian male with long-term non-progression (15 years since date of infection). The polymorphism at codon 335 was only observed in Africans (allelic frequency 0.02) and Coloureds (allelic frequency 0.02). This supports previous studies which suggested that the polymorphism has an African origin with an allelic frequency of approximately 0.03 and is rarely observed in Caucasian populations [15,16,18]. The mutations at codons 55, 75, 223 and 339 were found at low allelic frequencies.

Within the nine 'high-risk' seronegative commercial sex workers of Zulu descent, no possibly protective mutations were found within the coding region of *CCR5*. As this assay does not include the promoter region, the remaining 5′ and 3′ end untranslated regions and the intronic sequences, we cannot exclude the possibility that some significant mutations occurring in these regions may have been missed. Although this study cohort is small, our findings suggest that other factors (including the possibility of alternative gene involvement) may provide protection against HIV infection within this population group.

Because of the high degree of sequence homology between *CCR5* and *CCR2*, part of *CCR2* was simultaneously amplified using the *CCR5* fragment E primer set. Therefore, this assay also allows for comprehensive analysis of codons 217–267 of *CCR2*. In our study, the T allele of codon 260 (AAC-AAT) was found to occur more frequently (0.62) than the commonly reported C allele (0.38). No statistically significant differences in allelic frequencies for this polymorphism were observed and no additional *CCR2* sequence variants were detected.

The relatively high frequency of novel mutations observed in the African and Coloured patients demonstrates the effectiveness of the *CCR5*-DGGE assay and the importance of comprehensive *CCR5* gene analysis in populations where the *CCR5Δ32* mutation is rare. The recently admixed Coloured population of South Africa may, therefore, represent a valuable candidate population for the identification of genes/mutations underlying susceptibility to HIV/AIDS within the African context. Future analysis on the effect of the novel mutations on the functioning of CCR5 will result in a better understanding of this chemokine receptor and may contribute to the development of HIV therapeutic and preventative measures that focus on the interaction of HIV with the host proteins.

Acknowledgements

References

1. Deng H, Liu R, Ellmeier W, *et al.* **Identification of a major co-receptor for primary isolates of HIV-1**. *Nature* 1996, **381**: 661–666.

2. Dragic T, Litwin V, Allaway GP, *et al.* **HIV-1 entry into CD4⁺cells is mediated by the chemokine receptor CC-CKR-5.** *Nature* **1996, 381:667–673.**

3. Doranz BJ, Rucker J, Yi Y, *et al.* **A dual-tropic primary HIV-1 isolate that uses fusin and the β-chemokine receptors CKR-5, CKR-3, and CKR-2 as fusin cofactors.** *Cell* 1996, **85**:1149–1158.

4. Feng Y, Broder CC, Kennedy PE, Berger E. **HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor.** *Science* 1996, **272**:827–877.

5. Samson M, Labbe O, Mollereau C, Vassart G, Parmentier M. **Molecular cloning and functional expression of a new human CC-chemokine receptor gene.** *Biochemistry* 1996, **35**:3362–3367.

6. Alkhatib G, Combadiere C, Broder CC, *et al.* **CC CKR5: a RANTES, MIP-1α, MIP-1β receptor as a fusin cofactor for the macrophage-tropic HIV-1.** *Science* 1996, **272**:1955–1958.

7. Choe H, Farzan M, Sun Y, *et al.* **The β-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates.** *Cell* 1996, **85**:1135–1148.

8. Liu R, Paxton WA, Choe S, *et al.* **Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply exposed individuals to HIV-1 infection.** *Cell* 1996, **86**: 367–377.

9. Mummidi S, Ahuja SS, McDaniel BL, Ahuja SK. **The human CC chemokine receptor 5 (CCR5) gene.** *J Biol Chem* 1997, **272**: 30662–30671.

10. Samson M, Libert F, Doranz BJ *et al.* **Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene.** *Nature* 1996, **382**:722–725.

11. Dean M, Carrington M, Winkler C *et al.* **Genetic restriction of HIV-1 infection by a deletion allele of the *CKR5* structural gene.** *Science* 1996, **273**:1856–1862.

12. Huang Y, Paxton WA, Wolinsky SM *et al.* **The role of a mutant CCR5 allele in HIV-1 transmission and disease progression.** *Nat Med* 1996, **2**:1240–1243.

13. Zimmerman PA, Buckler-White A, Alkhatib G *et al.* **Inherited resistance to HIV-1 conferred by an inactivating mutation in the CC chemokine receptor 5: studies in populations with contrasting clinical phenotypes, defined racial background and quantified risk.** *Mol Med* 1997, **3**:23–26.

14. Martinson JJ, Chapman NH, Rees DC, Liu YT, Clegg JB. **Global distribution of the *CCR5* gene 32-basepair deletion.** *Nat Genet* 1997, **16**:100–103.

15. Ansari-Lari MA, Liu X, Metzker ML, Rut AR, Gibbs RA. **The extent of genetic variation in the *CCR5* gene.** *Nat Genet* 1997, **16**:221–222.

16. Carrington M, Kissner T, Gerrard B, Ivanov S, O'Brien SJ, Dean M. **Novel alleles of the chemokine-receptor gene *CCR5*.** *Am J Hum Genet* 1997, **61**:1261–1267.

17. Quillent C, Oberlin E, Braun J *et al.* **HIV-1-resistance phenotype conferred by combination of two separate inherited mutations of the *CCR5* gene.** *Lancet* 1998, **351**:14–18.

18. Carrington M, Dean M, Martin MP, O'Brien SJ. **Genetics of HIV-1 infection: chemokine receptor CCR5 polymorphism and its consequences.** *Hum Mol Genet* 1999, **8**:1939–1945.

19. Fischer SG, Lerman LS. **DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: correspondence with melting theory.** *Proc Natl Acad Sci USA* 1983, **80**: 1579–1583.

20. Sheffield VC, Cox DR, Lerman LS, Myers RM. **Attachment of a 40 base-pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single base changes.** *Proc Natl Acad Sci USA* 1989, **86**:232–236.

21. Abrams ES, Murdaugh SE, Lerman LS. **Comprehensive detection of single base changes in human genomic DNA using denaturing gradient gel electrophoresis and a GC-clamp.** *Genomics* 1990, **7**: 463–475.

22. Lerman LS, Silverstein K. **Computational simutation of the DNA melting and its application to denaturing gel electrophoresis.** *Meth Enzymol* 1987, **155**:482–501.

23. Wu Y, Hayes VM, Osinga J *et al.* **Improvement of fragment and primer selection for mutation detection by denaturing gradient gel electrophoresis.** *Nucl Acids Res* 1998, **26**:5432–5440.

24. Hayes VM, Wu Y, Osinga J *et al.* **Improvements in gel composition and electrophoretic conditions for broad-range mutation analysis by denaturing gradient gel electrophoresis.** *Nucl Acids Res* 1999, **27**:e29.

25. Guldberg P, Guttler F. **A simple method for identification of point mutations using denaturing gradient gel electrophoresis.** *Nucl Acids Res* 1993, **21**:2261–2262.

26. D'Souza I, Poorkaj P, Hong M *et al.* **Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements.** *Proc Natl Acad Sci USA* 1999, **90**:5598–5603.

27. Lorson CL, Hahnen E, Androphy EJ, Wirth B. **A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy.** *Proc Natl Acad Sci USA* 1999, **90**:6307–6311.

28. Kantor R, Gershoni JM. **Distribution of the CCR5 gene 32-base pair deletion in Israeli ethnic groups.** *J Acquir Immune Defic Syndr* 1999, **20**:81–84.
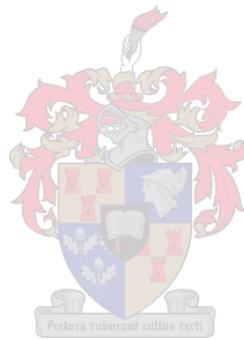
Appendix B

African-based *CCR5* single-nucleotide polymorphism

associated with HIV-1 disease progression

AIDS 2002; 16:2229 -2231

Correspondence

African-based *CCR5* single-nucleotide polymorphism associated with HIV-1 disease progression

The CC chemokine receptor 5 (CCR5), a seven transmembrane G-coupled protein, is one of the major co-receptors on CD4 T lymphocytes used for cell entry during HIV-1 infection [1–3]. A 32 base pair (bp) deletion (*CCR5Δ32*), resulting in a defective CCR5 protein, was identified as reducing the risk of HIV-1 infection in the homozygous state, while offering partial resistance to infection or reducing the rate of disease progression to AIDS and death in the hetero-zygous state [4–7]. This polymorphism has been reported to occur at allelic frequencies of 0.092 in Caucasians, and is extremely rare in populations of African ethnic origin [4,7,8]. In a previous study [9], we found *CCR5Δ32* to be absent (0/268 alleles) in our African population (predominantly Xhosa from central African descent) and present at an allele frequency of 0.03 (6/196 alleles) in our coloured population (mixed ancestral descent, including San, Khoi, African negro, Madagascan, Javanese and Eur-opean origin) from South Africa.

A number of mutations have been identified in the promoter and coding regions of *CCR5*, which may influence gene regulation or protein function. Two African-based single nucleotide polymorphisms (SNP) in the coding region of *CCR5* have been identified by ourselves and others, namely P35 [9] and A335V [10]. Although a recent study of a large Ugandan population showed no association between these polymorphic sites and either HIV-1 infection or the rate of disease progression (P35 occurred at an allelic frequency less than 0.01 within this population) [11], African-based studies have been limited to date. We used these SNPs to identify possible disease association in two African-based populations residing in the Western Cape of South Africa.

In an African setting, where clinic attendance is ham-pered by factors such as vast distances, economic burdens and a lack of patient education, it is very difficult to classify patients accurately into appropriate groups according to disease progression. Out of a study cohort of 1035 HIV-seropositive individuals, 76 (7%) could be classified into 35 normal (22 African, nine coloured, four Caucasian), 22 slow (seven African, eight coloured, seven Caucasian) and 19 fast progressors (11 African, five coloured, three Caucasian). Normal progressors were defined as individuals with a progres-sive loss in CD4 cell counts, who developed AIDS-related symptoms and co-infections within 10 years

after HIV-1 infection. All received clinical monitoring for a minimum period of 5 years. Approximately 5% of HIV-1-infected individuals remain asymptomatic for more than 10 years after seroconversion with stable CD4 cell counts, and are known as long-term non-progressors (LTNP) [12,13], whereas others, namely the fast progressors, develop AIDS within 2–5 years after seroconversion. The criteria used to determine slow and fast progression in this study is based primarily on the length of time the patient remained asympto-matic after seroconversion, as defined in Table 1. Only 17% (13/76) had received some form of antiretroviral treatment.

All individuals were screened for the P35 and A335V polymorphisms using the *CCR5*-denaturing gradient gel electrophoresis (DGGE) assay previously described by our group for fragments B and F, respectively. In addition, all slow progressors were screened for the *CCR5Δ32* using *CCR5*–DGGE fragment D primer set and conditions [9]. Manual allele counting was used for calculating allele frequencies, and statistical signifi-cance was determined using Fisher's exact test for 2×2 contingency tables (InStat version 3.0). None of the 14 Caucasians analysed were found to have either P35 or A335V and therefore, because of the African origin of the markers, were excluded from further calculations.

We first described the SNP occurring at codon 35 and resulting in a silent mutation as an African-based polymorphism, occurring at a significantly increased allelic frequency in coloured HIV-seropositive indivi-duals compared with HIV-seronegative controls, whereas similar frequencies were observed within infected versus non-infected Africans [9]. Increasing our study population we found no statistically sig-nificant differences between infected (17/164 alleles, 0.104) and non-infected (13/196 alleles, 0.066) coloured individuals ($P = 0.2511$). In addition, no statistically significant differences were found when comparing slow versus fast and slow versus normal progressors.

When comparing allelic frequencies of the codon 335 polymorphism, which results in a conservative amino acid change from an alanine to a valine, statistically significant differences were observed between both slow (4/30 alleles) versus fast (0/32 alleles) progressors ($P = 0.0491$) and slow versus normal (1/62 alleles)

Table 1. Clinical classification of the 22 slow progressors and the 19 fast progressors included in this study.

Disease progression	No. patients	No. years asymptomatic[a]	WHO clinical staging[b]				CD4 cell count (cells/mm³)[c]
			I	II	III	IV	
'True' LTNP	6	≥ 12 years (average 14)	4	1	1	0	250–800 (average 535)
'Potential' LTNP	5	10–12 years (average 10.4)	5	0	0	0	> 400 (average 602)
'Possible' LTNP	11	5–10 years (average 6.5)	11	0	0	0	> 400 (average 665)
FP	19	≤ 4 years (average 2.2)	0	5	7	7	< 400 (average 171)

FP, Fast progressor; LTNP, long-term non-progressor.
[a]Number of years the patient has remained asymptomatic after date of seroconversion.
[b]WHO, World Health Organisation clinical disease classification based on AIDS-related clinical symptoms (www.who.int).
[c]CD4 T cell counts, measured as the number of cells per cubic millimeter, only two of the 'true' LTNP had CD4 cell counts less than 500 cells/mm³ and had seroconverted 16 (CD4 cell count 280 cells/mm³) and 17 (CD4 cell count 277 cells/mm³) years previously.

progressors ($P = 0.0374$) within the African-based population groups (African and coloured). The *CCR5*-A335V SNP was found to occur homozygously in a single African woman who, remaining asymptomatic 7 years after infection, was classified as a 'possible' LTNP with a CD4 cell count of 450 cells/mm³ and recent (past year) antiretroviral treatment. Two slow progressors, one 'true' LTNP coloured man (asymptomatic 14 years after infection; CD4 cell count 610 cells/mm³) and one 'possible' LTNP coloured woman (asymptomatic 6 years after infection, CD4 cell count 914 cells/mm³), were heterozygous for this SNP, neither having received any previous antiretroviral treatment. All slow progressors were screened for the *CCR5Δ32* mutation so as to exclude its effect on disease progression. Two of the 15, one coloured woman remaining asymtomatic 10 years after infection (CD4 cell count 728 cells/mm³) and one coloured man remaining asymtomatic 8 years after infection (CD4 cell count 1171 cells/mm³), were heterozygous for the *CCR5Δ32* mutation. Neither presented with the *CCR5*-A335V SNP.

In this study, we show a statistically significant association between the *CCR5*-A335V SNP and decreased disease progression from HIV infection to AIDS within individuals of African ethnic origin residing in the Western Cape of South Africa. The variable penetrance observed for this protective association, together with previous findings that this polymorphic marker results in a functional response similar to that of wild-type CCR5 [14], suggests that A335V does not contribute directly in slowing disease progression. We therefore hypothesize that this genetic marker, in combination with other weaker genetic events, collectively acts in slowing disease progression and may ultimately contribute to providing an individual HIV susceptibility risk profile for sub-Saharan Africans.

Acknowledgements

Vanessa M. Hayes[a], Desiree C. Petersen[b], Thomas J. Scriba[b], Michele Zeier[c], Ashraf Grimwood[d] and Estrelita Janse van Rensburg[b], Departments of [a]*Urology,* [b]*Medical Virology and* [c]*Internal Medicine, University of Stellenbosch, Tygerberg School of Health Sciences, Tygerberg, South Africa; and* [d]*Chapel Street Community Health Clinic, Woodstock, South Africa.*

References

1. Alkhatib G, Combadiere C, Broder CC, *et al.* **CC CKR5: a RANTES, MIP-1α, MIP-1β receptor as a fusion cofactor for macrophage-tropic HIV-1.** *Science* 1996, **272**:1955–1958.
2. Deng H, Liu R, Ellmeier W, *et al.* **Identification of a major co-receptor for primary isolates of HIV-1.** *Nature* 1996, **381**:661–666.
3. Dragic T, Litwin V, Allaway GP, *et al.* **HIV-1 entry into CD4 cells is mediated by the chemokine receptor CC-CKR-5.** *Nature* 1996, **381**:667–673.
4. Dean M, Carrington M, Winkler C, *et al.* **Genetic restriction of HIV-1 infection by a deletion allele of the *CKR5* structural gene.** *Science* 1996, **273**:1856–1862.
5. Huang Y, Paxton WA, Wolinsky SM, *et al.* **The role of a mutant CCR5 allele in HIV-1 transmission and disease progression.** *Nat Med* 1996, **2**:1240–1243.
6. Liu R, Paxton WA WA, Choe S, *et al.* **Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply exposed individuals to HIV-1 infection.** *Cell* 1996, **86**:367–377.

7. Samson M, Libert F, Doranz BJ, *et al.* **Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene.** *Nature* 1996, **382**: 722–725.

8. Martinson JJ, Chapman NH, Rees DC, Liu YT, Clegg JB. **Global distribution of the *CCR5* gene 32-basepair deletion.** *Nat Genet* 1997, **16**:100–103.

9. Petersen DC, Kotze MJ, Zeier MD, *et al.* **Novel mutations identified using a comprehensive CCR5-denaturing gradient gel electrophoresis assay.** *AIDS* 2001, **15**:171–177.

10. Carrington M, Kissner T, Gerrard B, Ivanov S, O'Brien SJ, Dean M. **Novel alleles of the chemokine-receptor gene *CCR5*.** *Am J Hum Genet* 1997, **61**:1261–1267.

11. Ramaley P, French N, Kaleebu P, Gilks C, Whitworth J, Hill AVS. **Chemokine-receptor genes and AIDS risk.** *Nature* 2002, **417**:140.

12. Munoz A, Kirby AJ, He YD, *et al.* **Long-term survivors with HIV-1 infection: incubation period and longitudinal patterns of CD4+ lymphocytes.** *J Acquired Immune Defic Syndr Hum Retrovirol* 1995, **8**:496–505.

13. Pantaleo G, Menzo S, Vaccarezza M, *et al.* **Studies in subjects with long-term nonprogressive human immunodeficiency virus infection.** *N Engl J Med* 1995, **332**:209–216.

14. Blanpain C, Lee B, Tackoen M, *et al.* **Multiple nonfunctional alleles of CCR5 are frequent in various human populations.** *Blood* 2000, **96**:1638–1645.

Interleukin 7 production by bone marrow-derived stromal cells in HIV-1-infected patients during highly active antiretroviral therapy

Haematopoiesis is supported by a complex network of cell types, including non-haematopoietic (fibroblasts, adypocytes and endothelial cells) and haematopoietic cells [1]. Haematopoietic cells proliferate and differentiate in the interstices of this network, being in close contact with long cytoplasmic processes of stromal cells [2]. One major function of stromal cells is the production of IL-7 [3], which acts as a master regulator of T cell homeostasis, expanding naive and peripheral T cell populations [4,5]. In HIV-1-infected individuals, an inverse correlation has been observed between the levels of circulating IL-7 and the peripheral CD4 T cells [6], as the production of other cytokines diminishes [7]. It may be hypothesized that in advanced HIV-1 disease, stromal cells produce higher levels of IL-7 as a compensatory mechanism to the decline in the CD4 T cell count. At present, no data are available regarding either IL-7 production by bone marrow stromal cells in HIV-1-infected patients and the effects exerted by antiretroviral therapy on this cytokine. In this study we analysed stromal IL-7 production in bone marrow cultures before and during highly active antiretroviral therapy (HAART) in a group of HIV-1-infected patients.

Bone marrow mononuclear cells were collected from a breastbone aspirate in six HIV-1-infected patients, naive for HAART, with CD4 T cell counts between 100 and 300 cells/µl and plasma HIV-1 viral loads greater than 10 000 copies/ml. A second bone marrow evaluation was performed after 3 months of HAART. Three seronegative individuals were studied as controls. All subjects, who underwent bone marrow aspirates for peripheral blood haematological abnormalities, gave their written informed consent for the breastbone aspiration, according to the Ethical Committee procedures at our Institute. Bone marrow mononuclear cells were cultured in 24-well plates in Iscove's modified Dulbecco's medium (Gibco BRL, Life Technology Italia srl, Milan, Italy), supplemented with 10% fetal calf serum, 10% horse serum, 100 IU/ml penicillin–streptomycin and 100 IU/ml glutamine, until stromal confluence (3–4 weeks). Supernatants were collected after 24 h of culture and measurements of IL-7 were performed using enzyme-linked immunosorbent assay with an ultrasensitive kit (R&D System, Minneapolis, MN, USA). Non-parametric statistics were used (Wilcoxon test) for comparisons between the parameters analysed before and during therapy. Compared with seronegative subjects, stromal cells from HIV-1-infected individuals spontaneously produced higher levels of IL-7 before starting HAART (0.17 ± 0.10 versus 0.32 ± 0.12 pg/ml, respectively). In Fig. 1 IL-7 production is depicted for each patient before and during therapy. The levels of stromal IL-7 production significantly decreased in all patients during therapy, reaching values comparable to those observed in control subjects (0.19 ± 0.06 pg/ml, $P = 0.02$ versus baseline). This finding was associated with a significantly increased number of CD4 T cells (mean values from 188 ± 81 to 288 ± 54 cells/µl after 3 months of therapy, $P = 0.04$) and a significant decrease in HIV-1-RNA plasma levels (from 5 ± 0.6 to $1.4 \pm 0.8 \log_{10}$/ml, $P = 0.02$).

Furthermore, an inverse relationship has been observed between bone marrow stromal IL-7 production and peripheral CD4 T cell counts ($r = -0.767$; $P = 0.003$). IL-7 is well recognized as a crucial cytokine for the early development of T and B lymphocyte subpopulations. Thymic epithelial cells and bone marrow stromal cells are the primary sources of circulating IL-7 [8].

In this study we have shown that the decreased CD4 T cell count is associated with increased stromal IL-7 production. During HAART, the control of viral replication and the recovery of CD4 T cell numbers reverse bone marrow stromal IL-7 production, which reaches values comparable with those observed in control subjects. This contrasts with our observation that other cytokines (IL-2, IFN-γ, IL-4, IL-10) increase after immunological reconstitution during HAART [9].

The mechanisms responsible for the feedback between
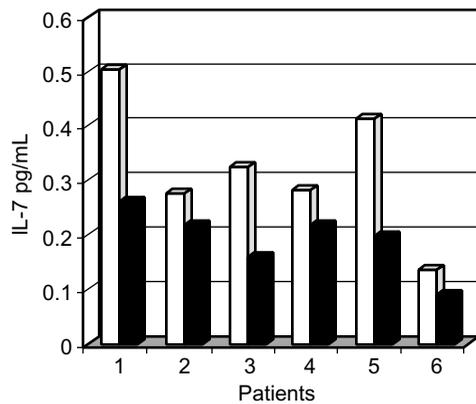
Fig. 1. IL-7 levels in supernatants of stromal cell cultures from HIV-1-infected patients before and during highly active antiretroviral therapy. □ Before highly active antiretroviral therapy (HAART); ■ during HAART.

bone marrow stromal cells and T cells remain unknown. It may be hypothesized that constitutively produced soluble factors by T cells may control the stromal cell production of IL-7 [8]. The depletion of T cells in the bone marrow, with the decreased production of such soluble factors, would lead to increased levels of IL-7, and vice versa. A possible candidate that exerts this action is transforming growth factor beta, which is known to be produced by T cells and to regulate the production of IL-7 by bone marrow stroma [10].

Antonella Isgrò[a], Fernando Aiuti[a], Ivano Mezzaroma[a], Fabrizio Franchi[b], Anna Marria Mazzone[a], Filippo Lebba[a] and Alessandro Aiuti[c], Departments of [a]Allergy and Clinical Immunology, and [b]Clinical Medicine, University of Rome 'La Sapienza', Rome, Italy; and

[c]*Telethon Institute for Gene Therapy (TIGET), Scientific Institute H.S. Raffaele, Milan, Italy.*

References

1. Dexter TM, Lajtha LG. **Proliferation of haemopoietic stem cells in vitro.** Br J Haematol 1974, **28**:525–530.
2. Jacobsen K, Osmond DG. **Microenvironmental organization and stromal cell associations of B lymphocyte precursor cells in mouse bone marrow.** Eur J Immunol 1990, **20**:2395–2404.
3. Sudo T, Ito M, Ogawa Y, et al. **Interleukin 7 production and function in stromal cell-dependent B cell development.** J Exp Med 1989, **170**:333–338.
4. Tan JT, Dudl E, LeRoy E, et al. **IL-7 is critical for homeostatic proliferation and survival of naive T cells.** Proc Natl Acad Sci U S A 2001, **96**:8732–8737.
5. Mackall CL, Fry TJ, Bare C, Morgan P, Galbraith A, Gress RE. **IL-7 increases both thymic-dependent and thymic-independent T-cell regeneration after bone-marrow transplantation.** Blood 2001, **97**:1491–1497.
6. Napolitano LA, Grant RM, Deeks SG, et al. **Increased production of IL-7 accompanies HIV-1-mediated T-cell depletion: implication for T-cell homeostasis.** Nat Med 2001, **7**:73–79.
7. Poli G, Fauci A. **Role of cytokines in the pathogenesis of human immunodeficiency virus infection.** In: *Human cytokines: their role in disease and therapy.* Aggarval BB, Puri RK (authors). Cambridge, MA: Blackwell Science; 1995. pp. 421–449.
8. Fry TJ, Mackall C. **Interleukin-7: master regulator of peripheral T-cell homeostasis.** Trends Immunol 2001, **22**:564–571.
9. Isgrò A, Aiuti A, Mezzaroma I, et al. **Improvement of IL-2 production, clonogenic capability and restoration of stromal cell function in HIV-1 patients after HAART.** Br J Haematol 2002, in press.
10. Tang J, Nuccie BL, Ritterman I, Liesveld JL, Abboud CN, Ryan DH. **TGF-β down-regulates stromal IL-7 secretion and inhibits proliferation of human B-cell precursors.** J Immunol 1997, **159**:117–125.

Nodal marginal zone lymphoma in AIDS patients: a casual association?

Four clinical categories of non-Hodgkin's lymphoma (NHL) in HIV-positive patients (HIV+) have been recognized [1]: systemic NHL, primary central nervous system lymphoma, primary effusion lymphoma, and plasmablastic lymphoma of the oral cavity.

The most frequent systemic types are diffuse large B cell and Burkitt's-like lymphoma (54% of all NHL-HIV+) [2]. However, 'low-grade lymphomas' [3] are increasingly being recognized, especially follicular and small lymphocytic lymphoma. Extranodal marginal zone B cell lymphomas have also been reported [4], but nodal marginal zone lymphomas (NMZL) have rarely been described. To our knowledge, only two previous patients have been reported [5,6]. We describe two new cases of a group of 14 NHL-HIV+ from a registry of 190 NHL patients consecutively seen between 1997 and 2001.

Case 1

An HIV-positive (bisexual risk behaviour) 53-year-old man, not receiving highly active antiretroviral therapy, was referred because of enlarged nodes and sweats. An inguinal node biopsy was compatible with NMZL (Fig. 1). His blood count was normal, $\beta_2$microglobulin was 3.9 mg/l (0.6–2.4), proteins were 95 g/l and gammaglobulins were 36 g/l (polyclonal pattern). CD4 lymphocytes were 0.261×10^9/l. A gastroscopy was normal, but intertrabecullar lymphoid infiltrates were observed in the bone marrow biopsy. A diagnosis of NMZL (stage IVB) was made and he started chemotherapy (cyclophosphamide, adriamycin, vincristine and prednisone), achieving a complete remission after six courses. Eight months later he relapsed and received salvage chemotherapy (fludarabine, cyclophosphamide and mitoxantrone) in four courses.

In-situ hybridization studies for Epstein−Barr virus (EBV) were negative and the Ki-67 proliferation rate was low (15%). No immunoglobulin heavy chain gene (IgH) (CDRIII region) rearrangement was demonstrated using the polymerase chain reaction method.

Case 2

A 40-year-old man presented with a right cervical lymphadenopathy. He was diagnosed with HIV and hepatitis C virus infections (parenteral drug abuse). He refused highly active antiretroviral therapy. He described a weight loss of 8 kg. His blood count and lactate dehydrogenase level were normal, β_2 microglobulin was 4 mg/l (0.6−2.4), and CD4 cell counts were 0.360×10^9/l. An IgM kappa monoclonal component was detected on protein electrophoresis.

The histology from the cervical lymph node revealed NMZL. Clinical staging including bone marrow biopsy, and fibergastroscopy was IB. The patient rejected therapy and was lost to follow-up.

Neoplastic monocytoid cells were positive for EBV by in-situ hybridization. The Ki-67 proliferation rate was low (25%). Clonal rearrangement of the IgH (CDRIII region) was demonstrated.

Marginal zone lymphoma is a clinicopathological entity in the WHO classification [7], and includes three subtypes: extranodal mucosa-associated lymphoid tissue lymphoma, splenic marginal zone lymphoma with or without villous lymphocytes, and NMZL with or without monocytoid B cells.

NMZL was initially reported by Sheibani et al. [8] as nodal monocytoid B cell lymphoma. This term defines cases that are primarily adenopathic and excludes cases
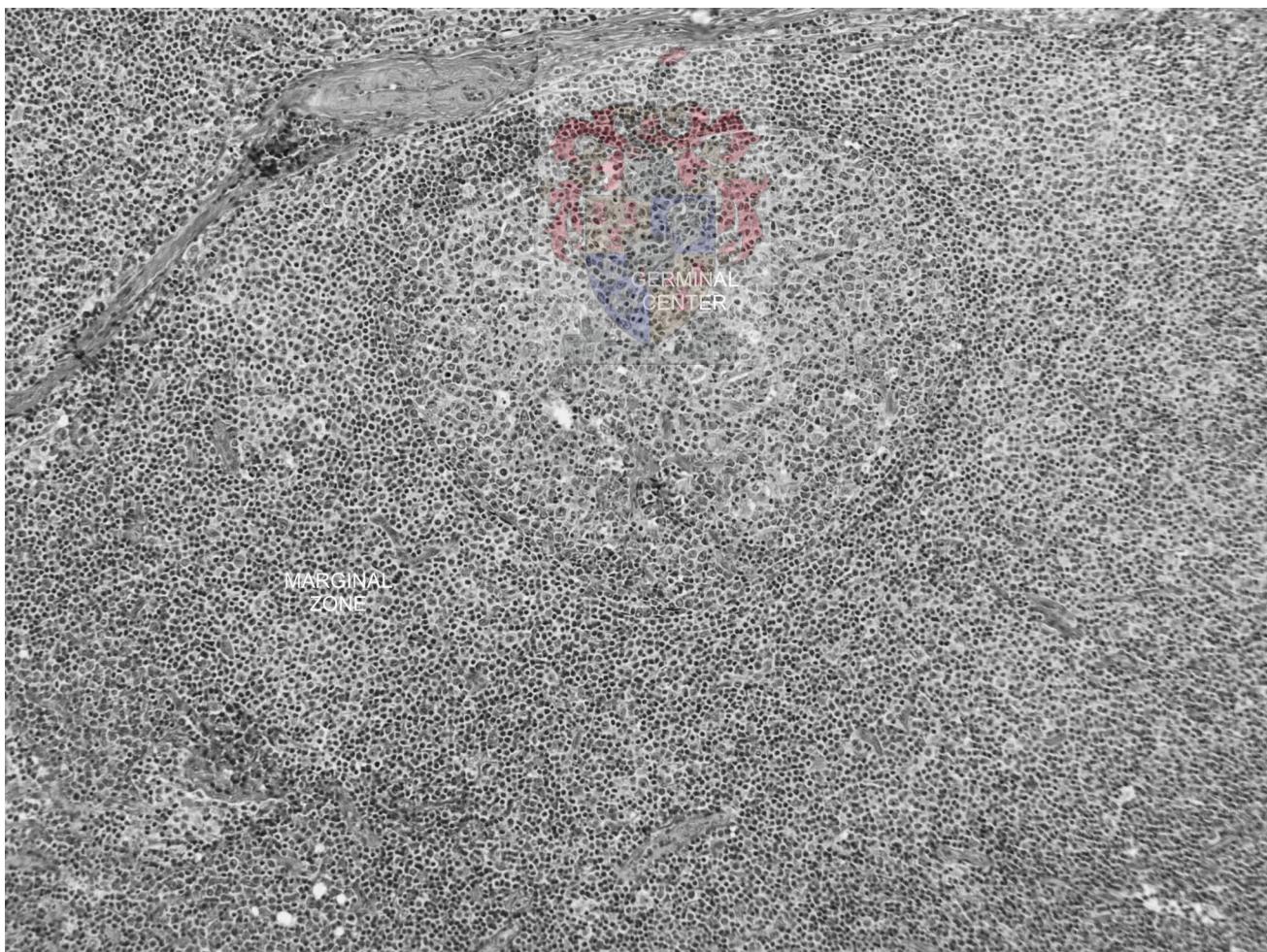

Fig. 1. Nodal marginal zone lymphoma (case 1). Pale zone corresponding to marginal zone around a germinal center. (H&E × 100).

with previous involvement of an extranodal site other than bone marrow, liver or spleen.

Clinically, NMZL mainly affects women in middle age and has an indolent course.

NMZL is extremely rare in HIV patients. According to a MEDLINE search, only two cases have previously been reported. In the first case, Sheibani *et al.* [5] presented a monocytoid B cell lymphoma in a patient who remained stable and asymptomatic during 3 years without therapy. The authors were able to demonstrate by polymerase chain reaction an HIV genome in DNA extracted from the node tissue. In the second case, Charton–Bain *et al.* [6] described an HIV patient with a monocytoid B cell lymphoma with bone marrow involvement who received monochemotherapy and was asymptomatic after a follow-up of 18 months.

NMZL must be carefully differentiated in the setting of HIV infection from secondary monocytoid B cell proliferation caused by toxoplasma, cytomegalovirus, EBV and hepatitis C infection. An IgH clonal rearrangement allows the demonstration of clonality in tumour lymphocytes.

EBV may be present in AIDS-related lymphomas, especially in primary brain lymphomas and Hodgkin's disease, in 70–80% of systemic large cell and primary effusion lymphomas, and is less frequently detected in other subtypes, such as Burkitt's lymphoma (30%) [9]. EBV was not detected in the case reported by Charton–Bain *et al.* [6].

As stated above, a careful review of case reports and clinical NMZL series has shown the rarity of the association between NMZL and HIV infection. Berger *et al.* [10] reviewed 124 patients with non–mucosa-associated lymhoid tissue marginal zone lymphoma. They failed to detect HIV infection in their patients. Other recent reviews and clinical series of NHL–HIV+ did not reveal identifiable cases of NMZL [11,12].

As a result of the small number of our NHL–HIV+ patients, we are aware that a casual association can not be excluded. However, we suggest that the true

incidence of NMZL in HIV patients might be under-estimated.

Eugenia Abella[a], Carles Besses[a], Carlos Barranco[b], Carme Pedro[a], Joan Buch[a] and Blanca Espinet[b], Departments of [a]Clinical Hematology and [b]Pathology, Hospital del Mar, IMAS, IMIM, Barcelona, Spain.

References

1. Centers for Disease Control. **Revision of the case definition of acquired immunodeficiency syndrome for national reporting – United States.** *Ann Intern Med* 1985; **103**:402–403.
2. Levine AM, Seneviratne L, Espina BM, Wohl AR, Tulpule A, Nathwani B, *et al.* **Evolving characteristics of AIDS-related lymphoma.** *Blood* 2000; **96**:4084–4090.
3. Cote TR, Biggar RJ, Rosenberg PS, Devesa SS, Percy C, Yellin FJ, *et al.* **Non-Hodgkin's lymphoma among people with AIDS: incidence, presentation, and public health burden.** *Int J Cancer* 1997; **73**:645–650.
4. Wotherspoon AC, Diss TC, Pan L, Singh N, Whelan J, Isaacson PG. **Low grade gastric B-cell lymphoma of mucosa associated lymphoid tissue in immunocompromised patients.** *Histopathology* 1996; **28**:129–134.
5. Sheibani K, Ben-Ezra J, Swartz WG, Rossi J, Kesirian J, Koo Ch, *et al.* **Monocytoid B-cell lymphoma in a patient with human immunodeficiency virus infection.** *Arch Pathol Lab Med* 1990; **114**:1264–1267.
6. Charton-Bain MC, Le Tourneau A, Weiss L, Bruneval P, Diebold J. **Rare non Hodgkin's lymphoma in the course of infection with human immunodeficiency virus. Two cases with bone marrow invasion.** *Ann Pathol* 1997; **17**:38–40.
7. Jaffe ES, Harris NL, Stein H, Vardiman JW (editors). *World Health Organization classification of tumours. Pathology and genetics of tumours of the haematopoietic and lymphoid tissues.* Lyon: IARC Press; 2001.
8. Sheibani K, Sohn CC, Burke JS, Winberg CD, Wu AM, Rappaport H. **Monocytoid B-cell lymphoma: a novel B-cell neoplasm.** *Am J Pathol* 1986; **124**:310–318.
9. Boye Hansen P, Penkowa M, Kirk O, Skinhoj P, Pedersen C, Lisse I, *et al.* **Human immunodeficiency virus-associated malignant lymphoma in eastern Denmark diagnosed from 1990 to 1996: clinical features, histopathology, and association with Epstein–Barr virus and human herpesvirus-8.** *Eur J Haematol* 2000; **64**:368–375.
10. Berger F, Felman P, Thieblemont C, Pradier Th, Baseggio L, Bryon PA, *et al.* **Non-MALT marginal zone B-cell lymphomas: a description of clinical presentation and outcome in 124 patients.** *Blood* 2000; **95**:1950–1956.
11. Dal Maso L, Rezza G, Zambon P, Tagliabue G, Crocetti E, Vercelli M, *et al.* **Non-Hodgkin lymphoma among young adults with and without AIDS in Italy.** *Int J Cancer* 2001; **93**:430–435.
12. Rabkin CS. **AIDS and cancer in the era of highly active antiretroviral therapy.** *Eur J Cancer* 2001; **37**:1316–1319.

The CD4 cell count 3 months after acute retroviral syndrome is associated with the presence of AIDS in the source individual

A source individual with AIDS was associated with a higher rate of HIV transmission [1], but the impact of the HIV stage of the source individual on the early events in the recipient individual is unclear. Whereas the CD4 cell count remains a valid prognostic marker independently of the HIV viral load [2,3], we explored

the relationship between the presence of AIDS in the source individual at the time of transmission and the CD4 cell count 3 months after a documented acute retroviral syndrome (ARS) in a recipient individual. A low level of CD4 cells might signify transmission of a virus with an increased virulence as shown with SIV [4].

A total of 19 pairs of source individuals and recipient individuals were analysed. Each of these recipient individuals was enrolled at the time of ARS in three prospective cohorts (Geneva, Switzerland, Montreal, Canada and Sydney, Australia) for studying the rate of HIV disease progression [5,6]. Data collection forms and follow-up were similar in the three centres [5,6]. At enrolment, the symptoms of ARS were recorded, and completed by biological data. The recipient individuals were invited to provide information on the most likely source individual [7]. The following anonymous information on the source individual were recorded: (i) HIV status of the source individual; (ii) if HIV positive, the presence or absence of AIDS at the time of transmission; and (iii) if the source individual was known to receive antiretroviral drugs at the time of HIV transmission to the recipient individual. A multiple linear regression was used to identify the variables associated with the CD4 cell level (in absolute numbers and in percentages) 3 months after the onset of ARS in the recipient individual.

The population consisted of 17 men (90%), and the routes of infection were homosexual intercourse for 16 patients and heterosexual intercourse for 3 patients. The CD4 cell count was measured at 3 months after onset (88.2 ± 11.0 days on average) and the average cell count was 603.7 cells/ml (± 240.2) . A total of eight recipient individuals (42%) received antiretroviral treatment at ARS on average 40 days (± 49) after onset. The results of the linear regression with CD4 cells in absolute numbers or as a percentage of total lymphocytes, as dependent variables are reported in Table 1. The presence of AIDS in the source individual was an independent predictor of the CD4 cell level 3 months after ARS in the source individual.

Carre *et al.* [8] and Ashton *et al.* [9] found a faster progression to AIDS in recipient individuals after infection by sex and by contaminated blood products when the source individuals progressed faster to AIDS. Our findings show that a difference in progression is already apparent 3 months after acute infection. Advanced HIV disease in the source individual correlates with immunodepression in the recipient individual soon after infection. That can intuitively be related to the quantity (i.e. size of inoculum or a high viral load) or the virulence (i.e. phenotype) of HIV transmitted.

In summary, a low CD4 cell count shortly after acute HIV infection through the sexual route might depend on the HIV stage of the source individual at transmission. That completes viral investigations showing that HIV strains became more virulent with time in the host [4] and then could influence early events in individuals newly infected with such HIV strains.

Philippe Vanhems[a], Emmanuelle Caillat-Vallet[a], Bernard Hirschel[b], Jean-Pierre Routy[c], Andrew Carr[d], Jeanette Vizzard[e], David A. Cooper[e], Luc Perrin[b] and The Swiss HIV Cohort Study, *[a]Laboratoire d'Epidemiologie et de Santé Publique, INSERM Unité 271, Université Claude Bernard, et Unité d'Epidémiologie, Hôpital Edouard Herriot, Lyon, France; [b]Division of Infectious Diseases, University Hospital, Geneva, Switzerland; [c]Immunodeficiency Service and Division of*

Table 1. Variables associated with the CD4 cell count in the absolute number or percentage 3 months after acute retroviral syndrome using multiple linear regression.

	Regression coefficient univariate model			Regression coefficent multivariate model			Regression coefficient univariate model			Regression coefficent multivariate model		
	CD4 cell count (absolute number)						CD4 cell count (%)					
Variables	β	Standard error	*P*	β	Standard error	*P*	β	Standard error	*P*	β	Standard error	*P*
Characteristics of the recipient individual												
Age (per 1 year increase)	−0.6	7.7	0.93	–	–		0.3	0.35	0.39	–	–	
Male sex	19.3	184.7	0.92	–	–		4.3	8.5	0.62	–	–	
Antiretroviral treatment at ARS	59.1	114.0	0.61	–	–		8.1	4.9	0.12	11.1	3.9	0.01
Characteristics of the source individual												
AIDS at transmission (n = 3)	−287.8	139.0	0.05	−319.9	141.9	0.04	−15.8	6.1	0.02	−19.1	5.2	0.002
Antiretroviral treatment at transmission (n = 6)	−33.4	121.7	0.61	–	–		3.4	5.6	0.56	–	–	

ARS, Acute retroviral syndrome.
The final model was based on a stepwise regression with variables entered into the model if the probability of its score statistics was 0.15 and variables removed from the model if the probability of its score statistics was 0.20.

Hematology, McGill University Health Center, McGill University, Montreal, Canada; ^dHIV, Immunology and Infectious Diseases Clinical Services Unit, St Vincent's Hospital, Sydney, Australia; and ^eNational Centre in HIV Epidemiology and Clinical Research, University of New South Wales, Sydney, Australia.

References

1. Quin TC, Wawer MJ, Sewankambo N, Serwada D, Chuanjun L, Wabwire-Mangen F, *et al.* **Viral load and heterosexual transmission of human immunodeficiency virus type 1.** *N Engl J Med* 2000, **342**:921–929.
2. Vlahov D, Graham N, Hoover D, Flynn C, Bartlett JG, Margolick JB, *et al.* **Prognostic indicators for AIDS and infectious disease death in HIV-infected injection drug users.** *JAMA* 1998, **279**: 35–40.
3. Saah AJ, Hoover DR, Weng S, Carrington M, Mellors J, Rinaldo CR Jr, *et al.* **Association of HLA profiles with early plasma viral load, CD4+ cell count and rate of progression to AIDS following acute HIV-1 infection. Multicenter AIDS Cohort Study.** *AIDS* 1998, **12**:2107–2113.
4. Kimata JT, Kuller L, Anderson DB, Dailey P, Overbaugh J. **Emerging cytopathic and antigenic simian immunodeficiency virus variants influence AIDS progression.** *Nat Med* 1999, **5**:535–541.
5. Vanhems P, Allard R, Cooper DA, Perrin L, Vizzard J, Hirschel B, *et al.* **Acute HIV-1 disease as a mononucleosis-like illness: is the diagnosis too restrictive?** *Clin Infect Dis* 1997, **24**:965–970.
6. Routy J-P, Vanhems P, Rouleau D, Tsoukas C, Lefebvre E, Cote P, *et al.* **Comparison of clinical features of acute HIV-1 infection in patients infected sexually or through injection drug use.** *J Acquired Immune Defic Syndr* 2000, **5**:425–432.
7. Vanhems P, Hirschel B, Phillips AN, Cooper DA, Vizzard J, Brassard J, *et al.* **Incubation time of acute HIV infection and duration of acute HIV infection are independent prognostic factors of progression to AIDS.** *J Infect Dis* 2000, **182**:334–337.
8. Carre N, Meyer L, Boufassa F, Deveau C, Persoz A, Rouzioux C, *et al.* **High risk of HIV disease progression after infection through a sexual partner with AIDS.** *AIDS* 1996, **10**:77–80.
9. Ashton LJ, Learmont J, Luo K, Wylie B, Stewart G, Kaldor JM. **HIV infection in recipients of blood products from donors with known duration of infection.** *Lancet* 1994, **344**:718–720.

Trends in hepatitis C and HIV infection among inmates entering prisons in California, 1994 versus 1999

The prevalences of hepatitis C virus (HCV) and HIV are much higher among incarcerated populations than the general public. For example, the incidence of HCV in the United States has been estimated at 1.8% [1], and more recently at 2.5% from a population-based sample of young women living in poorer neighbor-hoods in California [2]. However, 41.2% of California inmates were anti-HCV positive in 1994 [3]. In 1999, 2.1% of state and federal prison inmates were known to be HIV positive [4]. Whereas rates of HCV and HIV are higher among men within the general population, greater proportions of female inmates have been found to be infected with HCV and HIV. Among female inmates entering the California correctional system in 1994, 63.5% were anti-HCV positive compared with 39.4% of male inmates [3]. The prevalence of HIV was greater among female than male inmates (3.1 versus 2.5%) of the California prison system [3], and at nine out of 10 correctional systems across the United States [5].

Between 1995 and 2001, the incarcerated population in the United States grew an average of 4.0% annually [6]. The importance of monitoring HCV and HIV within this growing and mobile population was the reason to replicate a 1994 cross-sectional survey of inmates entering the California correctional system.

The California Department of Corrections has 13 reception centers in which male and female inmates are processed separately for entrance into the prison system. Four of the 10 male centers and two of the three female centers were selected for inclusion in the surveys. The same centers were selected in 1994 and 1999. A sample from each prison was selected based on the proportion of inmates processed at the center on a weekly basis. All incoming inmates to the California Department of Corrections receive a physical examina-tion shortly after arrival at a reception center. During the physical examination, a blood sample is obtained for syphilis serology. Inmates cannot refuse to provide a blood sample; leftover blood was used for blinded testing of HCV and HIV antibodies. Blood specimens were collected between August and September 1994 (men) and August and October 1994 (women). Sam-ples for 1999 were collected between January and March for both men and women. The same laboratory methods were used in 1994 and 1999. HCV antibodies were detected using the hepatitis C virus encoded antigen (recombinant c 100-3, HC-31 and HC-34) Abbott HCV enzyme-linked immunosorbent assay (EIA) 2.0 (Abbott Laboratories, North Chicago, IL, USA). Sera were tested for HIV antibodies using the Abbott EIA. Those specimens repeatedly reactive to EIA were confirmed by immunofluorescence assay, and any discrepancy was resolved using Western blot. Unlinked survey data were used to estimate the seroprevalence of HCV and HIV antibodies; each

Table 1. Prevalence and percentage change for HIV and hepatitis C virus by sex and race, among California inmates, 1994 and 1999.

	1994 (%)	1999 (%)	% change	% HIV positive		% change	% HCV positive		% change
				1994	1999		1994	1999	
Men	n = 4140	n = 4876							
White	29.8	23.1	−26*	2.2	1.3	−41*	49.1	36.3	−26*
African American	31.4	27.4	−17*	3.8	2.3	−39*	29.2	33.8	+16*
Latino	34.4	34.3	−4	1.5	0.6	−60*	40.2	36.0	−10*
All others	4.4	15.2	+245*	1.1	1.5	+36	24.6	27.5	+12
Total	100	100		2.4	1.4	−42*	39.4	34.2	−13*
Women	n = 624	n = 719							
White	36.4	31.3	−19	1.3	0.4	−69	58.1	26.2	−55*
African Americrican	34.3	34.9	−5	4.2	2.8	−33	37.7	29.1	−23*
Latina	23.7	25.2	−2	4.7	1.1	−77*	69.7	23.8	−66*
All others	5.6	8.6	+59	2.8	3.2	+14	80.0	11.3	−86*
Total	100	100		3.2	1.7	−47*	54.5	25.3	−54*

HCV, Hepatitis C virus.
*$P < 0.05$, one-tailed test for difference of proportions.

correctional facility provided demographic information. The California Health and Welfare Agency Committee for the Protection of Human Subjects approved the study protocols for both the 1994 and 1999 studies.

A total of 4140 male and 624 female inmates were tested in 1994, and a total of 4876 male and 719 female inmates were tested in 1999. Less than 3% of the samples in both surveys (n = 137 in 1994 and n = 135 in 1999) could not be tested, either because no blood was drawn, the quantity of the sample was too small, or the specimen was not saved.

In 1999, men entering California prisons were more likely to be infected with HCV than were women; HCV seroprevalence rates were 34.2 for male inmates and 25.3 for female inmates (Table 1). HCV antibody seroprevalence declined 13% from 1994 to 1999 among male inmates overall. However, a 16% increase was found for HCV positivity among African American men. Among female inmates, a decrease of 54% was found for HCV from 1994 to 1999.

HIV seroprevalence decreased from 1994 to 1999 by 42% for men and 47% for women. Compared with white and Latino inmates, African American male and female inmates were more likely to be infected with HIV in 1999.

The decline in HCV and HIV prevalences demonstrate a possible reduction in injection drug use or an increase in safer injecting practices within California. Whereas total admissions to publicly funded drug and alcohol treatment programs in California increased from 1995 to 1999, the number of injection drug use admissions decreased 13.4% during that time [7]. Likewise, felony drugs arrests among adults in California dropped 15.6% from 1994 to 1999; arrests for narcotic drugs declined

among men and women (21.8 and 5.5%, respectively) as did arrests for 'dangerous drugs' (including methamphetamines) during this period (men, −19.1%; women, −13.7%) [8]. Finally, perhaps changes in injection risk behaviors, decreases in needle sharing and increases in the use of syringe exchange programs, seen in New York City from 1990–1994 to 1995–1999 also took place in California during this decade [9].

Although rates of HCV and HIV among California prison inmates declined from 1994 to 1999, the approximately one in three male and one in four female inmates infected with HCV represents a serious public health concern. Control of HIV and HCV requires primary and secondary harm–reduction interventions targeted at correctional populations effectively to reduce risk behaviors during incarceration and after release. Our findings for African American inmates (i.e. the highest HIV prevalence in 1999 among both men and women; the highest HCV prevalence in 1999 among women, and the increase in HCV prevalence from 1994 to 1999 among men) strongly suggest that culturally appropriate interventions must be developed specifically for African American prisoners.

Sun from the California Department of Social Services; Ms Vanessa Baird from the DHS, Office of Multi–cultural Health; and Mr Michael Montgomery, Ms Barbara Bailey, Dr Matthew Facer and Ms Donna Zukowski from the DHS, Office of AIDS.

Juan D. Ruiz[a], Fred Molitor[b] and Julie A. Plagenhoef[c], [a]*California Department of Health Services, Office of AIDS, HIV/AIDS Epidemiology Branch, Sacramento, CA, USA;* [b]*ETR Associates, Sacramento, CA, USA; and* [c]*University of California, Davis, CA, USA.*

References

1. Alter MJ, Kruszon-Moran D, Nainan OV, *et al.* **The prevalence of hepatitis C virus infection in the United States, 1988 through 1994.** *N Engl J Med* 1999, **341**:556–562.
2. Page-Shafer KA, Cahoon-Young B, Klausner JD, *et al.* **Hepatitis C virus infection in young, low-income women: the role of sexually transmitted infection as a potential cofactor for HCV infection.** *Am J Public Health* 2002, **92**:670–676.
3. RuiRuiz JD, Molitor F, Sun RK, *et al.* **Prevalence and correlates of hepatitis C virus infection among inmates entering the California correctional system.** *West J Med* 1999, **170**:156–160.
4. US Department of Justice. *HIV in prison and jails, 1999.* Washington, DC: US Department of Justice, Bureau of Justice Statistics. July 2001, NCJ-187456.
5. Vlahov D, Brewer TF, Castro KG, *et al.* **Prevalence of antibody to HIV-1 among entrants to US correctional facilities.** *JAMA* 1991, **265**:1129–1132.
6. US Department of Justice. *Correction statistics.* Washington DC: US Department of Justice, Bureau of Justice Statistics. Available at http://www.ojp.usdoj.gov/bjs/correct.htm. Accessed 30 May, 2002.
7. State of California. *California indicators of drug and alcohol abuse.* Department of Alcohol and Drug Programs, Office of Applied Research and Analysis. Available at http://www.adp.-cahwnet.gov/RC/rc_comm.shtml. Accessed 3 June, 2002.
8. State of California. *Report on drug arrests in Calfornia from 1990 to 1999.* Office of the Attorney General, Bureau of Criminal Information and Analysis. Available at http://caag.state.ca.us/cjsc/publications/misc/drugarrests/drugs2.pdf. Accessed 30 May, 2002.
9. Maslow CB, Friedman SR, Perlis TE, Rockwell R, Des Jarlais DC. **Changes in HIV seroprevalence and related behaviors among male injection drug users who do and do not have sex with men: New York City, 1990–1999.** *Am J Public Health* 2002, **92**:382–384.
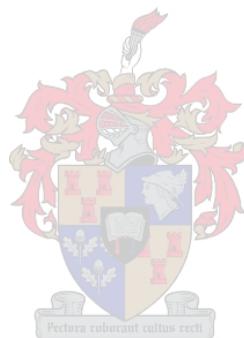
Appendix C

Novel mutations and SNPs identified in *CCR2* using a new

comprehensive denaturing gradient gel

electrophoresis assay

Hum Mut 2002; 20:253-259

RESEARCH ARTICLE

Novel Mutations and SNPs Identified in *CCR2* Using a New Comprehensive Denaturing Gradient Gel Electrophoresis Assay

Desiree C. Petersen,[1] Annette Laten,[1] Michele D. Zeier,[2] Ashraf Grimwood,[3] Estrelita Janse van Rensburg,[1] and Vanessa M. Hayes[4]*

[1]*Department of Medical Virology, University of Stellenbosch, Tygerburg, South Africa;* [2]*Department of Internal Medicine, University of Stellenbosch, Tygerburg, South Africa;* [3]*Chapel Street Community Health Clinic, Woodstock, South Africa;* [4]*Department of Urology, University of Stellenbosch, Tygerburg, South Africa*

Communicated by Georgia Chenevix-Trench

A single nucleotide polymorphism (SNP) at codon 64 in the CC chemokine receptor 2 gene (CCR2 V64I) has been associated with a dominant effect of delaying disease progression from human immunodeficiency virus-1 (HIV-1) infection to acquired immunodeficiency syndrome (AIDS). The objective of our study was to design a comprehensive mutation detection assay for the entire coding region of the CCR2A and CCR2B gene transcripts, including all relevant splice site junctions and to identify novel mutations and SNPs within our predominantly African-based population, which could influence an individual's susceptibility to HIV-1 infection and/or progression to AIDS. The mutation detection assay, based on denaturing gradient gel electrophoresis (DGGE), allowed for the complete analysis of five individuals per denaturing gel. Our study cohort consisted of 102 HIV seropositive patients and 144 HIV seronegative controls from the diverse South African population. Application of the CCR2-DGGE assay resulted in the detection of two previously reported CCR2 polymorphisms, namely CCR2 V64I and CCR2 N260N, and 11 novel mutations, including seven SNPs occurring at high allelic frequencies within specific population groups of South Africa. The large number of novel mutations/SNPs identified, using the CCR2-DGGE assay, indicates the importance for comprehensive analysis of all candidate genes in host susceptibility to HIV-1 infection, specifically in the under-studied African-based populations. Hum Mutat 20:253–259, 2002. © 2002 Wiley-Liss, Inc.

INTRODUCTION

Both isoforms of the seven transmembrane G-coupled CC-chemokine receptor 2 protein, CCR2A and CCR2B [Charo et al., 1994], bind the β-chemokine monocyte chemoattractant proteins 1 to 5 (MCP-1 to 5) [reviewed in Kalinkovich et al., 1999]. CCR2 has also been shown to act as an additional co-receptor during cellular infection of a few human immunodeficiency virus-1 (HIV-1) strains [Doranz et al., 1996]. Mutations in the *CCR2* gene (MIM# 601267), located at band p21 of chromosome 3 [Daugherty and Springer, 1997], may therefore be associated with susceptibility to HIV-1 infection and/or progression to acquired immune deficiency syndrome (AIDS).

There is a worldwide trend among geneticists away from the monogenic disorders and toward the analysis of the more common complex multifactorial diseases, where rather than a single gene defect, a combination of weaker genetic events may lead collectively to an individual's inherited disease susceptibility. This,

together with the completion of the first draft of the Human Genome Project, has resulted in an explosion in the identification of single nucleotide polymorphisms (SNPs) for use in association studies of complex diseases (NCBI's SNP database, www.ncbi.nlm.nih.gov/SNP).

A common SNP has previously been reported in the first transmembrane region of CCR2 and involves a G-A transition at codon 64, resulting in a conservative amino acid change from valine to isoleucine [Smith et al., 1997a]. Population surveys indicate that the CCR2 V64I SNP occurs at an allelic frequency of 0.10 to 0.25 within specific ethnic groups [Smith et al., 1997a; Michael et al., 1997]. In both the heterozygous and homozygous state the CCR2 V64I does not offer resistance to HIV-1 infection, but is associated with delaying progression to AIDS by two to four years [Smith et al., 1997a; Smith et al., 1997b; Michael et al., 1997; Kostrikis et al., 1998; Ioannidis et al., 2001], with its effect being more apparent in Africans than in Caucasians [Mummidi et al., 1998]. Studies indicate that CCR2 V64I does not affect CCR2 and CCR5 expression or co-receptor activity [Lee et al., 1998; Mariani et al., 1999]. Although it has been linked to a specific CCR5 promoter variant (59653-T) [Kostrikis et al., 1998], the exact mechanism by which this SNP influences disease progression still needs to be elucidated. The CCR2 V64I SNP has also been associated with insulin-dependant diabetes mellitus [Szalai et al., 1999], sarcoidosis [Hizawa et al., 1999], and coronary artery disease [Szalai et al., 2001].

This study involved the design of a comprehensive mutation detection assay for the entire coding region of both CCR2 gene transcripts (CCR2A and CCR2B), based on denaturing gradient gel electrophoresis (DGGE). DGGE is believed to be the most powerful of the polymerase chain reaction (PCR), gel-based mutation detection assays currently available. The use of the CCR2-assay led to the identification of novel mutations and SNPs in 102 HIV seropositive patients and 144 HIV seronegative controls of predominantly African ethnicity from South Africa.

METHODS
Patients

The study cohort consisted of 102 HIV seropositive patients (34 male; 68 female) of whom most are presently residents in the Western Cape of South Africa (Tygerberg Hospital and Woodstock Chapel Street Community Health Clinic). Due to lack of clinical information, precipitated by social and economical factors, the disease status was unknown for the majority of these patients. Also forming part of the study cohort, were 144 HIV seronegative healthy controls (56 male; 88 female) who are all blood donors for the Western Province Blood Transfusion Service and residing in the same geographical area as the HIV infected patients. The individuals

participating in this study included Africans, predominantly Xhosa (central African descent; 69 HIV+ and 62 HIV−), and Coloureds (mixed ancestral descent; 26 HIV+ and 72 HIV−), and to a lesser degree Caucasians (European descent; seven HIV+ and two HIV−) and Asians (eight HIV−). As described in Petersen et al. [2001], the term "African" refers to South Africans of central African descent; "Coloured" refers to individuals of mixed ancestral descent (including San, Khoi, African Negro, Madagascan, Javanese, and European origin); and "Caucasian" refers to South Africans of European descent (mainly of Dutch, French, German, and British origin). Informed consent for the study was obtained from all participants. The Ethics Review Committee of the University of Stellenbosch approved the study protocol.

Primer Design

Using the melt 87 computer program [Lerman and Silverstein, 1987] and conditions for selecting optimal PCR fragments and primers [Wu et al., 1998], DGGE primers (Table 1) were designed for the entire coding region, including the intron/exon boundaries, of both transcripts of the CCR2 gene. The coding region of CCR2B (codons 1 to 361) and most of the coding region of CCR2A (codons 1 to 313), contained in exon 2, were divided into six (A–F) overlapping amplicons. An additional amplicon "G" was designed to include the remaining coding region (exon 3) of CCR2A (codons 314 to 375). The addition of a GC-rich-fragment to the 5′ end of one of the primers in each primer set prevents complete strand dissociation during amplification. Additional stretches of GC or AT nucleotides were added to either the 5′ end of the non-clamped primer (fragments B, C, and E) or between the GC-clamp and the primer (fragment D), to ensure a single melting domain for optimal detection of all mutations (Table 1).

DNA Amplification

Genomic DNA was isolated from peripheral blood leukocytes using conventional methods and amplified using DGGE primer sets, specific for each amplicon (A–G) (Table 1). With a total volume of 50μl, each PCR reaction mixture contained 100 ng of genomic DNA, 0.1 mM of each deoxyribonucleoside triphosphate (dNTP), 20 pmol of each primer (except for fragment B, which required only 10 pmol of each primer), 2.5 mM of a $10 \times Mg^{2+}$ reaction buffer, and 0.5 units of DNA *Taq* polymerase (Boehringer Mannheim, Mannheim, Germany). Amplification was performed using a Perkin Elmer 9600 thermocycler (PE Applied Biosystems, www.appliedbiosystems.com) and the PCR cycling conditions were as follows: an initial denaturation at 96°C for 3 min, followed by 32 cycles of denaturation at 96°C for 45 sec, annealing for 1 min (annealing temperatures shown in Table 1), and elongation at 72°C for 1 min 20 sec. The last cycle was followed by an additional extension step of 72°C for 7 min. For optimal DGGE analysis, amplification was followed by a heteroduplexing step, which involves denaturation at 96°C for 10 min, followed by renaturation for 45 min at the annealing temperature of the amplification. The amplified products were checked using electrophoresis, where 5 μl (10%) of each sample was resolved on 2% agarose gel.

Denaturing Gradient Gel Electrophoresis (DGGE)

Optimized DGGE conditions were achieved by considering the conditions previously described for the improvements of broad-range DGGE analysis [Hayes et al., 1999]. DGGE was performed using the Ingeny phorU-2 system (www.ingeny.com).

TABLE 1. *CCR2* Primer Sets and Experimental Conditions for PCR Amplification and DGGE

Fragment	Amplimers, 5′-3′	Size (bp)	Melting	Annealing
A	[40GC]TGCTTATGTGGTGCCAGACT TGA ACACCAGCGAGTAGAGC	335	72	58
B	[6GC]TGATTA TGATTACGGTGCTCC [40GC]CGATTGTCAGGAGGATGATG	384	72	58
C	[40GC]GCTGTATCACATCGGTTATT [8GC]GCCACAGACATAAACAGAAT	268	73	55
D	[40GC][10AT]TGGCTGTGTTTGCTTCTGT CGAGTAGCAGATGACCATGA	220	70	55
E	[5GC]CCACACA ATA ATGAGGAACA [40GC]TGGTGCTTTCACAGTTACTC	284	73	55
F	ACCTTCCAGGAATTCTTCG [40GC]ACAATCAAACTGCTCCTCGT	346	74	55
G	TGTCTGGATCTGAGCTGGTT [40GC]TCCAAAGCAGAGATCTGTCAT	333	73	58

(Temperature column spans Melting and Annealing)

bp, base pair. GC-clamps used were as follows: [40GC], CGCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCCG; [8GC], CGCCGCCG; [6GC], CGCCGC; [5GC], CGCCG; AT-stretch used was as follows: [10AT], TATAATATTA.

The seven amplicons were electrophoresed in six lanes (fragments C and D were pooled) of a 9% polyacrylamide gel containing a 30% to 70% urea and formamide (UF) denaturing gradient (100% UF = 7 mol/L urea per 40% deionized formamide) at 59.5°C for 110 volts overnight. The gels were stained with ethidium bromide and photographed under an UV transilluminator. Thus the *CCR2*-DGGE allows for the complete analysis of five individuals per denaturing gel.

DNA Sequencing and Variant Confirmation

Automated sequencing of amplified samples showing aberrant DGGE banding patterns was performed using a non-GC-clamped primer and the Applied Biosystems (www.appliedbiosystems.com) dye terminator sequencing kit. The commonly occurring SNPs were verified by mixing samples showing similar DGGE banding patterns, followed by a heteroduplexing step before electrophoresis on a denaturing gel [Guldberg and Guttler, 1993]. Samples showing additional heteroduplex bands were subjected to sequencing for the exact determination of the sequence variants.

Statistical Analysis

Manual allele counting was used for calculating allele frequencies. Tests for significance of heterogeneity in the frequencies among HIV seropositive patients and seronegative controls for both the mutations and SNPs were performed by means of Fischer's exact test for 2 × 2 contingency tables.

RESULTS

The *CCR2* primer sets and experimental conditions for PCR amplification and DGGE analysis are shown in Table 1. Using our *CCR2*-DGGE assay, we identified two previously reported mutations and 11 novel mutations as shown in Figure 1 and listed in Table 2, according to the intron/codon in which they occur.

Previously Reported *CCR2* SNPs

The commonly occurring SNP at codon 64 (GTC–ATC) results in a conservative amino acid change from Valine to Isoleucine [Smith et al., 1997a]. The *CCR2* V64I SNP was observed homozygously in four HIV seronegative controls (one African, two Coloureds, and one Asian) and heterozygously in both the HIV seropositive patients and HIV seronegative controls of all the different South African population groups. A second *CCR2* SNP occurring at codon 260 (AAC-AAT) and resulting in a silent mutation (*CCR2* N260) has been reported in the Genbank database (www.ncbi.nlm.nih.gov). As previously reported [Petersen et al., 2001], we found the T allele to occur more frequency than the commonly reported C allele within the entire South African study cohort.

Novel Rare *CCR2* Variants

Of the 11 novel mutations, two silent mutations at codons 52 and 223 were found in two Coloureds (one HIV+ and one HIV-) and one HIV seronegative Coloured, respectively. Two non-conservative (replacement of an amino acid by another with different chemical properties) mutations at codon 233 (CGA–CAA) and codon 355 of *CCR2A* (GGA–GAA), were observed heterozygously in one HIV seronegative Coloured and in three Africans (two HIV + and one HIV−), respectively. The mutation at codon 223 involves an amino acid change from arginine to glutatmine, while the codon 355 (*CCR2A*) mutation, results in an amino acid change from glycine to glutamic acid. The disease progression for both HIV
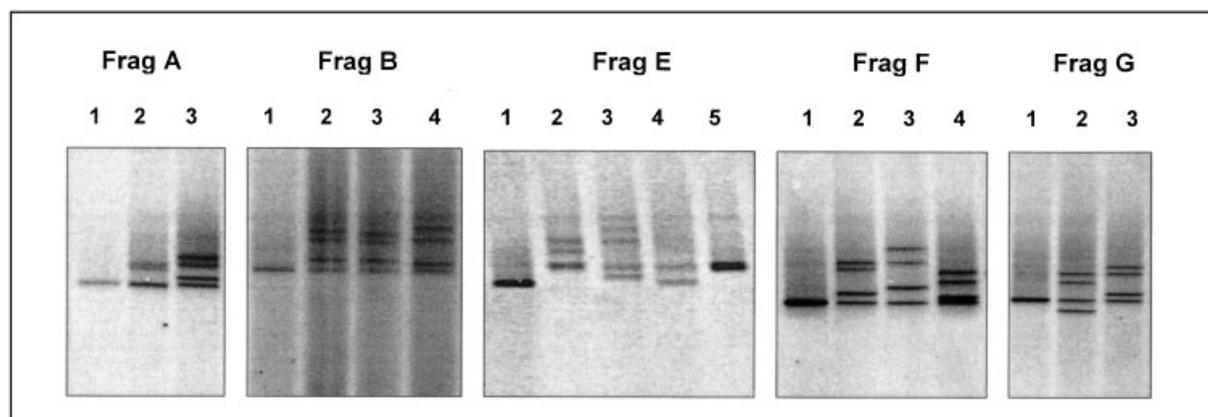
FIGURE 1. Aberrant DGGE banding patterns for all 13 *CCR2* gene mutations identified and listed in Table 2, according to the fragment (Frag) in which they were found. No mutations were identified in **Frags C** and **D**. Lane 1 of all the fragments represents the DGGE banding pattern of a normal control. Mutants are depicted as follows: **Frag A:** lanes 2, heterozygous for *CCR2* IVS1-57A>G; and lane 3, heterozygous for *CCR2* IVS1-43G>A. **Frag B:** lane 2, heterozygous for *CCR 2*V64I; lane 3, heterozygous for *CCR2* V63V; and lane 4, heterozygous for *CCR2* V52V. **Frag E:** lane 2, heterozygous for *CCR2* R233Q; lane 3, heterozygous for *CCR2* S223; lane 4, heterozygous for *CCR2* N260N; and lane 5, homozygous for *CCR2* N260N. **Frag F:** lane 2, heterozygous for *CCR2* L283L; lane 3, heterozygous for *CCR2* T287M; and lane 4, heterozygous for *CCR2* T348T. **Frag G:** lane 2, heterozygous for *CCR2* P339P; and lane 3, heterozygous for *CCR2* G355E.

TABLE 2. *CCR2* Mutations Detected in 102 HIV Seropositive Patients and 144 HIV Seronegative Controls, Ordered According to the Intron/Codon in Which They Occurred

| | | | Allele frequency | | | | | | |
| | | | Africans | | Coloureds | | Caucasians | | Asians |
Mutation	Base change	DGGE fragment	HIV+ (n=138)	HIV− (n=124)	HIV+ (n=52)	HIV− (n=144)	HIV+ (n=14)	HIV− (n=4)	HIV− (n=16)
IVS1-57A>G[a,b,c]	A–G	A	14 (0.101)	13 (0.105)	3 (0.058)	10 (0.069)	0	0	0
IVS1-43G>A[a,b,c]	G–A	A	3 (0.022)	5 (0.040)	3 (0.058)	3 (0.021)	0	0	0
V52V[a,b,c]	GTG–GTT	B	0	0	1 (0.019)	1 (0.007)	0	0	0
V63V[a,b,c]	GTC–GTT	B	1 (0.007)	1 (0.008)	5 (0.096)	5 (0.035)	2 (0.143)	0	1 (0.063)
V64I[b,c]	GTC–ATC	B	18 (0.130)	21 (0.169)	1 (0.019)	25 (0.174)	0	1 (0.250)	2 (0.125)
S223S[a,b,c]	TCG–TCA	E	0	0	0	1 (0.007)	0	0	0
R233Q[a,b,c]	CGA–CAA	E	0	0	0	2 (0.014)	0	0	0
N260N[b,c]	AAC–AAT	E	75 (0.543)	79 (0.637)	30 (0.577)	97 (0.674)	10 (0.714)	2 (0.500)	11 (0.688)
L283L[a,b,c]	CTG–CTT	F	6 (0.043)	3 (0.024)	1 (0.019)	4 (0.028)	0	0	0
T287M[a,b,c]	ACG–ATG	F	0	4 (0.032)	1 (0.019)	0	0	0	0
P339P[a,b]	CCA–CCG	G	21 (0.152)	18 (0.145)	2 (0.038)	5 (0.035)	0	0	0
T348T[a,c]	ACG–ACA	F	0	0	1 (0.019)	6 (0.042)	1 (0.071)	0	0
G355E[a,b]	GGA–GAA	G	2 (0.014)	1 (0.008)	0	0	0	0	0

[a]Novel mutation in this study.
[b]Mutation identified in *CCR2A*.
[c]Mutation identified in *CCR2B*.
n, number of alleles; HIV+, seropositive; HIV−, seronegative.

seropositive patients heterozygous for the *CCR2* G355E mutation is unknown.

Novel *CCR2* SNPs

The remaining seven novel mutations occur at an allelic frequency of greater than or equal to 0.01 and thus may be regarded as SNPs within the different South African populations studied (Table 3). Of these seven SNPs, two are found within intron 1 and involve an A to G nucleotide change at position −57 base pairs (bp) and a G to A change at position −43 bp downstream from the acceptor splice site, respectively. Four of the SNPs are silent mutations and are

observed at codons 63, 283, 339 (*CCR2A*), and 348 (*CCR2B*), while another SNP at codon 287 (ACG–ATG) involves a non-conservative amino acid change from threonine to methionine. Except for the SNP at codon 63, which was observed in all the different population groups, and the SNP at codon 348 (*CCR2B*), which was present in the Coloured and Caucasian population groups, the remaining five SNPs were found exclusively in Africans and/or Coloureds. Considering the low number of Caucasian individuals forming part of the study cohort, an additional 40 HIV seronegative Caucasians were screened for the novel SNPs identified. An absence of the IVS1-57A>G, IVS1-43G>A, L283, T287M, and P339 *CCR2* SNPs

TABLE 3. The Distribution of Novel CCR2 SNPs Within the African, Coloured, and Caucasian Population Groups

| | Total population[a] | | |
SNP	Africans (n = 262)	Coloureds (n = 262)	Caucasians (n = 98)
IVS1-57A>G[b,c]	27 (0.131)	13 (0.066)	0
IVS1-43G>A[b,c]	8 (0.031)	6 (0.031)	0
V63V[b,c]	2 (0.008)	10 (0.051)	3 (0.031)
L283L[b,c]	9 (0.034)	5 (0.026)	0
T287M[b,c]	4 (0.015)	1 (0.005)	0
P339P[b]	39 (0.149)	7 (0.036)	0
T348T[c]	0	7 (0.036)	7 (0.071)

[a]Total population includes all HIV seropositive patients and HIV seronegative controls for each population group and including the additional 40 Caucasians.
[b]Mutation identified in *CCR2A*.
[c]Mutation identified in *CCR2B*.
n, number of alleles.

and an allelic frequency of 0.075 (6/80) for the T348 *CCR2* SNP were observed and included in the total population allelic frequencies indicated in Table 3.

DISCUSSION

This study involved the design and use of a comprehensive mutation detection assay for the entire coding region of both gene transcripts of *CCR2* (*CCR2A* and *CCR2B*), for the identification of novel and previously reported mutations and SNPs in both HIV seropositive and HIV seronegative individuals in a predominantly African-based population from South Africa. The assay, based on DGGE, allows for the complete analysis of five patients per denaturing gel using a single set of experimental conditions. Previous studies have been restricted to the analysis of the *CCR2*V64I SNP in different population groups and thus comprehensive analysis of *CCR2* is important and ideal in our diverse South African study cohort.

The commonly reported *CCR2*V64I SNP, which has an allelic frequency ranging from 0.10 to 0.25 within specific populations [Smith et al., 1997a; Michael et al., 1997], was observed in all the South African population groups (non high-risk HIV seronegatives only) with allelic frequencies of 0.169 (21/124) in the Africans; 0.174 (25/144) in the Coloureds; and 0.095 (8/84) in the Caucasians. The presence of the *CCR2*V64I SNP has been dominantly associated with delaying disease progression to AIDS, provided that the patient's date of HIV-1 infection is more or less known [Smith et al., 1997a; Smith et al., 1997b; Michael et al., 1997; Kostrikis et al., 1998; Ioannidis et al., 2001]. The majority of the HIV seropositive patients forming part of our study cohort do not have known dates of HIV-1 infection. This, together with poor clinic attendance and therefore limited patient information, resulted in no associations with disease susceptibility being made. However, in the Coloured population, a statistically significant decrease in the frequency of the mutant allele was noted in the HIV seropositives, when compared to the HIV seronegative control group ($P = 0.0034$). Although analysis of the Africans does not yield a significant p value, the combination of the African and Coloured populations ($P = 0.0409$) does. The possible significance of this decrease, with regards to resistance to HIV infection within the Coloured population of mixed ancestral descent (including San, Khoi, African Negro, Madagascan, Javanese, and European origin), requires further investigation in studies that include a high-risk seronegative cohort.

Association studies with HIV-1 infection or disease progression have never been performed for the previously reported codon 260 polymorphism, which may be contributed to its lack in amino acid change. The observation that the T allele of the *CCR2*N260 SNP occurs at higher frequencies within all the population groups in this study, confirms our previous data [Petersen et al., 2001]. This is contrary to recent data, which suggests that the C allele occurs more frequently in Caucasians, Africans, and Hispanics residing in America [Clark et al., 2001]. Statistical analysis of the African-associated populations (African and Coloured) taken together, showed the mutant T allele to occur at a significantly lower frequency in the HIV seropositive group compared with the seronegative controls ($P = 0.0255$), although no statistically significant differences were noted for either the African or Coloured populations separately.

Novel mutations and SNPs were identified in all the different population groups represented in this study. The novel mutations at codons 233 and 355 (*CCR2A*) and the novel SNP at codon 287 all result in non-conservative amino acid changes, which may change the structure of the CCR2 protein and thus affect its functioning. The novel silent mutations (codons 52, 223) and SNPs (63, 283, 339 (*CCR2A*), and 348 (*CCR2B*)), as well as the two novel intronic SNPs (IVS1-57A>G and IVS1-43G>A) could all possibly influence gene expression and/or RNA splicing by altering regulatory elements

[D'Souza et al., 1999; Lorson et al., 1999]. Although CCR2 does not play a major role during HIV infection [Doranz et al., 1996], the significance of these mutations on the functioning of CCR2 is worthy of further study.

An additional 40 Caucasian seronegative control individuals were screened for all novel SNPs identified (Table 3) to determine whether these polymorphic sites have an African or Caucasian ethnic origin. SNPs occurring in the Coloured population may be either African or Caucasian in origin due to the relatively recent admixture of this group. The SNP at codon 63 was detected in all the population groups, while the SNP at codon 348 (*CCR2B*) appears to have a Caucasian origin, being found only in the Coloureds and Caucasians. The IVS1–57A > G, IVS1–43G > A, L283L, T287M, and P339P *CCR2* SNPs were found exclusively in the African and Coloured populations, suggesting an African-based origin. No significant associations with the novel SNPs and HIV-1 susceptibility and/or rates of disease progression to AIDS could be made due to insufficient information regarding the clinical status of the HIV seropositive patients.

Although expression of the CCR2A isoform is mostly restricted to the cytoplasm [Wong et al., 1997], we included the *CCR2A* transcript in our analysis. The possibility of identifying novel African-related SNPs within *CCR2A*, which although it does not alone lead to susceptibility may, in combination with other weaker genetic events, lead collectively to determining disease status. The commonly occurring African-associated SNP identified at codon 339 of the *CCR2A* transcript is currently under investigation by our group in a large cohort of patients with known disease status to determine the significance of this SNP in combination with other identified African-associated SNPs.

The large number of novel mutations identified in our South African study cohort is an indication of the efficiency of the described *CCR2*-DGGE assay and also emphasizes the importance of total gene screening for yet unidentified SNPs particularly in the under studied African ethnic populations. Further studies are required to determine the underlying mechanisms of the *CCR2* novel mutations and SNPs identified, so that its possible effects on influencing host susceptibility to HIV-1 infection and/or developing AIDS can be understood more clearly. More importantly, these novel SNPs may prove beneficial for inclusion in association studies of complex diseases, in particular haplotype studies of closely linked genetic markers specific to the HIV pandemic in sub-Saharan Africa.

ACKNOWLEDGMENTS

REFERENCES

Charo IF, Myers SJ, Herman A, Franci C, Connolly AJ, Coughlin SR. 1994. Molecular cloning and functional expression of two monocyte chemoattractant protein 1 receptors reveals alternative splicing of the carboxyl-terminal tails. Proc Natl Acad Sci USA 91:2752–2756.

Clark VJ, Metheny N, Dean M, Peterson RJ. 2001. Statistical estimation and pedigree analysis of CCR2-CCR5 haplotypes. Hum Genet 108:484–493.

Daugherty BL, Springer MS. 1997. The β-chemokine receptor genes CCR1 (CMKBR1), CCR2 (CMKBR2), and CCR3 (CMBRK3) cluster within 285 kb on human chromosome 3p21. Genomics 41:294–297.

Doranz BJ, Rucker J, Yi Y, Smyth RJ, Samson M, Peiper SC, Parmentier M, Collman RG, Doms RW. 1996. A dual-tropic primary HIV-1 isolate that uses fusin and the β-chemokine receptors CKR-5, CKR-3, and CKR-2 as fusin cofactors. Cell 85:1149–1158.

D'Souza I, Poorkaj P, Hong M, Nochlin D, Lee VM, Bird TD, Schellenberg GD. 1999. Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements. Proc Natl Acad Sci USA 90:5598–5603.

Guldberg P, Guttler F. 1993. A simple method for identification of point mutations using denaturing gradient gel electrophoresis. Nucleic Acids Res 21:2261–2262.

Hayes VM, Wu Y, Osinga J, Mulder IM, van der Vlies P, Elfferich P, Buys CH, Hofstra RM. 1999. Improvements in gel composition and electrophoretic conditions for broad-range mutation analysis by denaturing gradient gel electrophoresis. Nucleic Acids Res 27:e29.

Hizawa N, Yamaguchi E, Furuya K, Jinushi E, Ito A, Kawakami Y. 1999. The role of the C-C chemokine receptor gene polymorphism V64I (CCR2-64I) in sarcoidosis in a Japanese population. Am J Respir Crit Care Med 159:2021–2023.

Ioannidis JP, Rosenberg PS, Goedert JJ, Ashton LJ, Benfield TL, Buchbinder SP, Coutinho RA, Eugen-Olsen J, Gallart T, Katzenstein TL, Kostrikis LG, Kuipers H, Louie LG, Mallal SA, Margolick JB, Martinez OP, Meyer L, Michael NL, Operskalski E, Pantaleo G, Rizzardi GP, Schuitemaker H, Sheppard HW, Stewart GJ, Theodorou ID, Ullum H, Vicenzi E, Vlahov D, Wilkinson D, Workman C, Zagury JF, O'Brien TR. 2001. Effects of CCR5-Delta32, CCR2-64I, and SDF-1 3′A alleles on HIV-1 disease progression: an international meta-analysis of individual-patient data. Ann Intern Med 135:782–795.

Kalinkovich A, Weisman Z, Bentwich Z. 1999. Chemokine and chemokine receptors: role in HIV infection. Immunol Lett 68:281–287.

Kostrikis LG, Huang Y, Moore JP, Wolinsky SM, Zhang L, Guo Y, Deutsch L, Phair J, Neumann AU, Ho DD. 1998. A chemokine receptor CCR2 allele delays HIV-1 disease progression and is associated with a CCR5 promoter mutation. Nat Med 4:350–355.

Lee B, Doranz BJ, Rana S, Yi Y, Mellado M, Frade JM, Martinez-A C, O'Brien SJ, Dean M, Collman RG, Doms RW. 1998. Influence of the CCR2-V64I polymorphism on the human immunodeficiency virus type 1 coreceptor activity and on the chemokine receptor function of CCR2b, CCR3, CCR5 and CXCR4. J Virol 72:7450–7458.

Lerman LS, Silverstein K. 1987. Computational simutation of the DNA melting and its application to denaturing gel electrophoresis. Meth Enzymol 155:482–501.

Lorson CL, Hahnen E, Androphy EJ, Wirth B. 1999. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. Proc Natl Acad Sci USA 90:6307–6311.

Mariani R, Wong S, Mulder LC, Wilkinson DA, Reinhart AL, LaRosa G, Nibbs R, O'Brien TR, Michael NL, Connor RI, Macdonald M, Busch M, Koup RA, Landau NR. 1999. CCR2-64I polymorphism is not associated with altered CCR5 expression of coreceptor function. J Virol 73:2450–2459.

Michael NL, Louie LG, Rohrbaugh AL, Schultz KA, Dayhoff DE, Wang CE, Sheppard HW. 1997. The role of CCR5 and CCR2 polymorphisms in HIV-1 transmission and disease progression. Nat Med 3:1160–1162.

Mummidi S, Ahuja SS, Gonzalez E, Anderson SA, Santiago EN, Stephan KT, Craig FE, O'Connell P, Tryon V, Clark RA, Dolan MJ, Ahuja SJ. 1998. Genealogy of the CCR5 locus and chemokine system gene variants associated with altered rates of HIV-1 disease progression. Nat Med 4:786–793.

Petersen DC, Kotze MJ, Zeier MD, Grimwood A, Pretorius D, Vardas E, Janse van Rensburg E, Hayes VM. 2001. Novel mutations identified using a comprehensive CCR5-denaturing gradient gel electrophoresis assay. AIDS 15:171–177.

Smith MW, Dean M, Carrington M, Winkler C, Huttley GA, Lomb DA, Goedert JJ, O'Brien TR, Jacobson LP, Kaslow R, Buchbinder S, Vittinghoff E, Vlahov D, Hoots K, Hilgartner MW, O'Brien SJ. 1997a. Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Science 277:959–965.

Smith MW, Carrington M, Winkler C, Lomb D, Dean M, Huttley G, O'Brien SJ. 1997b. CCR2 chemokine receptor and AIDS progression. Nat Med 3:1052–1053.

Szalai C, Csaszar A, Czinner A, Szabo T, Panczel P, Madacsy L, Falus A. 1999. Chemokine receptor CCR2 and CCR5 polymorphisms in children with insulin-dependant diabetes mellitus. Pediatr Res 46:82–84.

Szalai C, Duba J, Prohaszka Z, Kalina A, Szabo T, Nagy B, Horvath L, Csaszar A. 2001. Involvement of polymorphisms in the chemokine system in susceptibility for coronary artery disease (CAD). Coincidence of elevated Lp(a) and MCP-1-2518G/G genotype in CAD patients. Atherosclerosis 158:233–239.

Wong L-M, Myers SJ, Tsou C-L, Gosling J, Arai H, Charo IF. 1997. Organization and differential expression of the human monocyte chemoattractant protein 1 receptor gene. J Biol Chem 272:1038–1045.

Wu Y, Hayes VM, Osinga J, Mulder IM, Looman MW, Buys CH, Hofstra RM. 1998. Improvement of fragment and primer selection for mutation detection by denaturing gradient gel electrophoresis. Nucleic Acids Res 26:5432–5440.
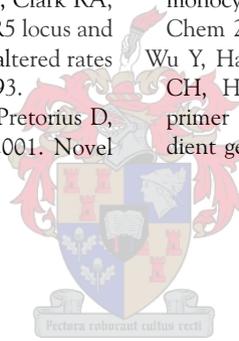
Appendix D

Functional analysis of novel *SLC11A1 (NRAMP1)*

promoter variants in susceptibility to HIV-1

J Med Genet 2004; 41:e49

Functional analysis of novel SLC11A1 (NRAMP1) promoter variants in susceptibility to HIV-1

H Donninger, T J Cashmore, T Scriba, D C Petersen, E Janse van Rensburg and V
M Hayes

Updated information and services can be found at:

http://jmg.bmjjournals.com/cgi/content/full/41/4/e49

These include:

References	This article cites 11 articles, 4 of which can be accessed free at: **http://jmg.bmjjournals.com/cgi/content/full/41/4/e49#BIBL**
	1 online articles that cite this article can be accessed at: **http://jmg.bmjjournals.com/cgi/content/full/41/4/e49#otherarticles**
Rapid responses	You can respond to this article at: **http://jmg.bmjjournals.com/cgi/eletter-submit/41/4/e49**
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article
Topic collections	Articles on similar topics can be found in the following collections Online mutation reports (137 articles)

Notes

ONLINE MUTATION REPORT

Functional analysis of novel *SLC11A1 (NRAMP1)* promoter variants in susceptibility to HIV-1

H Donninger, T J Cashmore, T Scriba, D C Petersen, E Janse van Rensburg, V M Hayes

....

The divalent cation transporter is the natural resistance associated macrophage protein 1 (formerly *NRAMP1* and now named *SLC11A1*) for solute carrier 11A1 (OMIM accession number 600266). The gene that codes for this transporter has been studied intensively for its role in conferring susceptibility to infectious diseases such as tuberculosis, leprosy, meningococcal meningitis, visceral leishmaniasis, and HIV infection, as well as to autoimmune diseases such as rheumatoid arthritis, diabetes, sarcoidosis, inflammatory bowel disease,[1] and, more recently, Kawasaki disease.[2] Most studies have investigated a functional GT repeat sequence in the promoter region of this gene[3] and have identified two commonly occurring repeat alleles and four rare repeat alleles.[4][5] The common alleles are $T(GT)_5\text{-}AC(GT)_5AC(GT)_{10}$ (allele 2) and $T(GT)_5AC(GT)_5AC(GT)_9$ (allele 3; GenBank accession number AF229163, 5768 to 5808). Allele 2, which decreases gene expression, has been associated with susceptibility to infectious diseases; the more common allele 3 enhances gene expression to protect against infectious diseases while enhancing susceptibility to autoimmune diseases. Although HIV is classified as an infectious disease, it affects the autoimmune system, which may explain why allele 3 is associated with susceptibility to HIV-1.[6][7] This study aimed to screen the promoter region of *SLC11A1* for novel sequence variations in people from sub-Saharan Africa infected with HIV-1 compared with uninfected people and to determine the effect of novel variants on normal promoter function.

MATERIALS AND METHODS
Participants
We studied 84 HIV-1 seropositive people (60 African and 24 of mixed African–European descent) and 133 HIV-1 seronegative people (64 African, 62 of mixed African–European descent, and seven Asian) who lived in the Western Cape of South Africa and who attended one of the HIV-1 clinics in Tygerberg or Langa or the blood transfusion services of the Western Cape. In this study, we use ''African'' to define people of predominantly Xhosa descent, while the mixed African–European population (known as ''Coloured'' in South Africa) consists of a well defined population with origins from an initial population mixture about 300 years ago. This study population is defined in a recent publication.[8] The ethics review committee of the University of Stellenbosch gave ethical approval, and informed consent was obtained from all study participants. We took blood from the patients and volunteers, and we extracted DNA with the Qiagen Extraction Kit (Qiagen, Valencia, CA, USA), as per the manufacturer's instructions.

Mutation detection
Two sets of denaturing gradient gel electrophoresis primers were designed to cover the upstream promoter region from the GT repeat sequence (from nucleotide −715 to nucleotide −488 from translation start site) and downstream promoter

Key points

- In contrast with previous studies, this study showed that polymorphic GT repeats in the promoter region of the *SLC11A1 (NRAMP1)* gene were not associated with altered risk for HIV-1 infection in African and African–European people from sub-Saharan Africa.

- This study aimed to identify new mutations in the promoter region of this gene that might be associated with susceptibility to infection with HIV-1 within African based populations.

- Denaturing gradient gel electrophoresis was used to assay the upstream and downstream regions of the promoter GT repeat and to screen for variants in 84 HIV-1 seropositive and 133 HIV-1 seronegative people.

- Three novel mutations and a previously reported single nucleotide polymorphism (g.332C>T) were identified. No significant associations were made between the single nucleotide polymorphism and susceptibility to HIV-1, but one of the novel mutations (g.43G>C) occurred in two HIV-1 seropositive people of African descent.

- Gene expression studies showed that the g.43G>C and g.75T>C variants enhanced promoter activity by 1.4-fold and 1.6-fold, respectively. The promoter single nucleotide polymorphism also enhanced activity 1.6-fold, but the g.561G>A variant had no effect on promoter activity.

region (from nucleotide −415 to nucleotide −132). The upstream primer pairs were forward 5'-AACAACTCTGA GAAGGGACA-3' and reverse 5'-TCTTTGATCTGGAGTT CCAA-3', and the downstream primers were forward 5'-GGGTGTGGTCATGGGGTATT and reverse 5'-TGCCCTGCC TCTTACATCAA-3'. Both forward primers were designed with a 40 base pair (bp) GC-clamp (CGCCCGCCGCGCCCC GCGCCCGGCCCGCGCCCCCGCCCG) attached to the 5'-end to prevent total strand dissociation during electrophoresis through a denaturing gradient. A 7 bp (GCCGCCG) GC stretch and a 10 bp (GCCGCCGCCG) GC stretch were added to the 5'-end of the upstream and downstream reverse primers, respectively; the latter created a single melting domain for optimal mutation detection, as previously described.[9] The repeat allele configuration was determined by direct sequencing of the amplified product created with forward primer 5'-AAGACTCGCATTAGGCCAAC-3' and reverse primer 5'-GCCTCCCAAGTTAGCTCTGA-3'.

We performed DNA amplification in a 50 μl reaction mixture that contained 100 ng genomic DNA, 0.1 mmol/l of each deoxyribonucleoside triphosphate, 20 pmol of each

primer, 2.5 mmol/l 10× magnesium ion reaction buffer, and 0.5 units DNA Taq polymerase (Boehringer Mannheim, Mannheim, Germany). We performed amplification under the following conditions: initial denaturation at 96°C for 3 minutes, 35 cycles of denaturation at 96°C for 45 seconds, annealing at 60°C for upstream amplicon and repeat region or 65°C for downstream amplicon for 1 minute, and elongation at 72°C for 1 minute 15 seconds. After cycling, we performed an additional elongation step for 10 minutes, at 72°C, followed by denaturation at 96°C for 10 minutes, and heteroduplexing for 45 minutes at the optimal annealing temperatures (the latter two steps were applicable only for the denaturing gradient gel electrophoresis primers). We electrophoresed upstream and downstream amplicons through a 9% polyacrylamide gel that contained a 40–80% urea and formamide denaturing gradient (100% urea and formamide contained 7 M urea and 40% deionised formamide) at 59°C and 110 V overnight (phorU2 system; Ingeny, Leiden, Netherlands). We subjected the repeat region amplicon and samples that showed novel denaturing gradient gel electrophoresis banding patterns to automated sequencing after product purification with a high purity polymerase chain reaction (PCR) product purification kit (Roche Diagnostics, Mannheim, Germany).

We determined allele frequencies by allele counting. We used Fisher's exact test, which included calculation of odds ratios by InStat software, to assess significance of association between HIV-1 status (seropositive v seronegative) and genotype for the previously reported repeat alleles, as well as the g.332C>T (GenBank accession number AY363243) promoter single nucleotide polymorphism.[10]

Gene expression studies

We analysed variant sequences with MatInspector (version 2.2; German Research Center for Biotechnology, Braunschweig, Germany) to establish possible loss or gain of transcription factor binding sites to determine the effect of novel promoter variants on promoter function (fig 1). To determine whether these novel mutations had any true effect on promoter activity, we cloned the various mutant promoters, together with the wild type promoter, upstream of the luciferase gene, and we transiently transfected the resulting constructs into 293 cells (ATCC: CRL-1573). We measured transactivation activity relative to Renilla luciferase in three independent experiments and determined means and standard deviations. All clones were sequenced to confirm the configuration of the repeat allele and promoter single nucleotide polymorphism.

RESULTS

Analysis of the 217 participants identified three of the previously reported repeat alleles (alleles 1, 2, and 3) and four single nucleotide variants (the previously reported g.332C>T polymorphism and three novel rare variants). All poly-

Figure 1 Schematic representation of the SLC11A1 promoter region (Genbank accession number AY363243), depicting the position of the novel mutations and known single nucleotide polymorphisms (shown below the sequence) and the various affected transcription factor binding sites (boxed regions). The core binding site for each transcription factor is highlighted in bold.

morphic alleles were in Hardy-Weinberg equilibrium in the HIV seropositive and seronegative participants for both populations. The repeat allele 3 was most common in the African and African–European admixed populations, with allele frequencies of 0.84 (101/120) and 0.73 (35/48) in participants infected with HIV and 0.81 (103/128) and 0.80 (99/124) in uninfected participants, respectively. We saw no significant associations between the presence of allele 3 in infected compared with uninfected African participants (p = 0.5073, odds ratio 1.290) and the mixed African–European population (p = 0.4124; odds ratio 0.6799) or in the homozygous states (Africans: p = 0.3406; African–European: p = 0.4715) or heterozygous states (African: p = 0.2426; African–European: p = 0.8079). Allele 2, previously associated with protection from HIV-1 infection in the homozygous state,[6] was not associated with altered risk of HIV infection in our study. Three of the seven ''slow progressors'' (three African and four African–European participants who were asymptomatic 10 years after becoming infected with HIV-1 and had not taken antiretroviral treatment) were heterozygous for allele 2, while all five ''fast progressors'' (three African and two African–European participants who progressed to AIDS within five years of infection with HIV-1) were homozygous for allele 3. Of the three slow progressors who were heterozygous for allele 2 (all African–European admixed), one was heterozygous for the CCR5-Δ32 bp deletion, which has been shown to be associated with slower progression of disease. Allele frequencies for the g.332C>T single nucleotide polymorphism in the African and African-European admixed populations infected with HIV-1 were 0.075 (9/120) and 0.02 (1/48), respectively; the frequencies for the uninfected participants were 0.05 (6/128) and 0.072 (9/124), respectively. The promoter single nucleotide polymorphism was associated with repeat allele 3 in both populations. We saw no significant associations between HIV-1 infected and non-infected African participants (p = 0.4136; odds ratio 1.706) and admixed African–European participants (p = 0.2560, odds ratio 0.2560).

The three novel sequence variants included a G to A base substitution at position 561 (g.561G>A), a T to C substitution at position 75 (g.75T>C), and a G to C transition at position 43 (g.43G>C) (GenBank accession number AY363243) at −156, −642, and −674 nucleotides from the initiator codon, respectively. The first two mutations were identified in HIV-1 seronegative participants of African and Asian descent, respectively; the g.43G>C mutation occurred in two unrelated HIV-1 seropositive patients of African descent. Promoter analysis showed that the G to C mutation at position 43 created a TCF11 transcription factor binding site. This mutation also interfered with a Nkx2–5 binding site, although it did not disrupt the site entirely. The single nucleotide polymorphism at position 332 introduced an OCT-1 binding site, and the transition at 561 altered, but did not destroy, a FoxD3 binding site. The 75 mutation did not create or alter any transcription factor binding sites.

We cloned these variants in a luciferase reporter system to determine their effect on promoter activity. Sequencing of the clones showed that all had the allele 3 repeat configuration and that all contained the C allele configuration for the g.332C>T polymorphism. Promoter analysis showed that the g.43G>C and g.75T>C mutations enhanced the activity of the SLC11A1 promoter 1.4 (SD 0.14)-fold and 1.6 (0.25)-fold, respectively, compared with the wild type sequence. The increase in activity for the g.75T>C mutation was significant (p<0.05), whereas the increase in activity for the g.43G>C mutation was not. The g.332C>T polymorphism also enhanced significantly the activity of the SLC11A1 promoter 1.6 (0.19)-fold (p<0.05), whereas the g.561G>A mutation

Figure 2 Mean (SD) transactivation activity of SLC11A1 promoter mutants in 293 cells. Activity of the wild type and mutant SLC11A1 promoters was analysed by luciferase assays. Values are expressed as luciferase activity relative to Renilla luciferase. Means of results from three independent experiments are given.

had no effect on promoter activity compared with the wild type sequence (fig 2).

DISCUSSION

The z-DNA forming polymorphic repeat in the *SLC11A1* promoter acts as a functional polymorphism and influences gene expression, with allele 3 having the strongest promoter activity.[4] Allele 3 has been associated with hyperactivation of macrophages and may be functionally linked to susceptibility to HIV-1.[11] In this study, although we did find a possible association with the presence of allele 2 and delayed disease progression, numbers were small, and this observation was not significant ($p = 0.2045$). In contrast with the literature, no associations were found with respect to the repeat polymorphism and risk of HIV infection within our population. In addition, the previously reported g.332C>T single nucleotide polymorphism was not associated with susceptibility to HIV. We used a comprehensive approach to screen the promoter region of *SLC11A1* and identified three novel sequence variants in combination with allele 3 of the polymorphic repeat. Although the g.75T>C mutation showed an increase in promoter activity, it was not associated directly with the introduction of a transcription factor binding site or with susceptibility to HIV-1; it occurred in 1/8 HIV-1 seronegative Asian participants. The lack of HIV-1 seropositive Asians in this study, however, warrants further investigation of this mutation in the Asian population. The g.43G>C novel mutation, which presented in 2/64 HIV-1 seropositive Africans but was absent in the 63 uninfected Africans, was shown to increase the promoter activity of *SLC11A1*, although this increase was not significant. The possibility that this functional mutation, which creates an additional transcription factor binding site, may lead to enhanced susceptibility to pathogen infection within the African population needs further investigation. As the exact mechanism of *SLC11A1* function in HIV-1 infection is unknown at present, the identification of novel population specific variants in the promoter region of this gene, which affects protein expression, may play a pivotal role in predicting susceptibility to HIV-1, particularly in populations from HIV stricken sub-Saharan Africa.

. . . .

Authors' affiliations

H Donninger, Departments of Internal Medicine, University of Stellenbosch, Tygerberg, South Africa
T J Cashmore, T Scriba, D C Petersen, E Janse van Rensburg, V M Hayes, Departments of Medical Virology, University of Stellenbosch, Tygerberg, South Africa
T J Cashmore, Departments of Human Genetics, University of Stellenbosch, Tygerberg, South Africa
V M Hayes, Cancer Research Program, Garvan Institute of Medical Research, St Vincent's Hospital, Sydney, Australia

Correspondence to: Dr V M Hayes, Cancer Research Program, Garvan Institute of Medical Research, 384 Victoria Street, Darlinghurst, NSW 2010, Australia; v.hayes@garvan.org.au

REFERENCES

1 **Blackwell JM**, Goswami T, Evans CAW, Sibthorpe D, Papo N, White JK, Searle S, Miller EN, Peacock CS, Mohammed H, Ibrahim M. SLC11A1 (formerly NRAMP1) and disease resistance. *Cell Microbiol* 2001;**3**:773–84.
2 **Ouchi K**, Suzuki Y, Shirakawa T, Kishi F. Polymorphism of the SC11A1 (formerly NRAMP1) gene confers susceptibility to Kawasaki disease. *JID* 2003;**187**:326–9.
3 **Blackwell JM**, Barton CH, White JK, Searle S, Baker AM, Williams H, Shaw M-A. Genomic organization and sequence of the human NRAMP gene: identification and mapping of a promoter region polymorphism. *Mol Med* 1995;**1**:194–205.
4 **Searle S**, Blackwell JM. Evidence for a functional repeat polymorphism in the promoter of the human NRAMP1 gene that correlates with autoimmune versus infectious disease susceptibility. *J Med Genet* 1999;**36**:295–9.
5 **Graham AM**, Dollinger MM, Howie SEM, Harrison DJ. Identification of novel alleles at a polymorphic microsatellite repeat region in the human NRAMP1 gene promoter: analysis of allele frequencies in primary biliary cirrhosis. *J Med Genet* 2000;**37**:150–2.
6 **Marquet S**, Sanchez FO, Arias M, Rodriguez J, Paris SC, Skamene E, Schurr E, Garcia LF. Variants of the human NRAMP1 gene and altered human immunodeficiency virus infection susceptibility. *JID* 1999;**180**:1521–5.
7 **Blackwell JM**, Searle S, Mohamed H, White JK. Divalent cation transport and susceptibility to infectious and autoimmune disease: continuation of the Ity/Lsh/Bcg/Nramp1/Slc11A1 gene story. *Immunol Lett* 2003;**85**:197–203.
8 **Hayes VM**. Genetic diversity of the alpha-1-antitrypsin gene in Africans identified using a novel genotyping assay. *Hum Mutat* 2003;**22**:59–66.
9 **Wu Y**, Hayes VM, Osinga J, Mulder IM, Looman MWG, Buys CHCM, Hofstra RMW. Improvement of fragment and primer selection for mutation detection by denaturing gradient gel electrophoresis. *Nucleic Acid Res* 1998;**26**:5432–40.
10 **Lewis LA**, Victor TC, van Helden EG, Blackwell JM, da Silva-Tatley F, Tullett S, Ethlers M, Beyers N, Donald PR, van Helden PD. Identification of C to T mutation at position −236 bp in the human NRAMP1 gene promoter. *Immunogenet* 1996;**44**:309–11.
11 **Shaw MA**, Clayton D, Atkinson SE, Williams H, Miller N, Sibthorpe D, Blackwell JM. Linkage of rheumatoid arthritis to the candidate gene NRAMP1 on 2q35. *J Med Genet* 1996;**33**:672–7.