The Role of a Single Nucleotide Polymorphism of the CD4 Gene in HIV-1 Disease Progression and Mother to Child Transmission

by

Françoise C.M. Vouriot

A thesis submitted to the Faculty of Graduate Studies in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Department of Medical Microbiology

University of Manitoba

Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

The Role of Single Nucleotide Polymorphism of the CD4 Gene in HIV-1 Disease Progression and Mother to Child Transmission

BY

Françoise C. M. Vouriot

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

Master of Science

Abstract

There is a single nucleotide polymorphism (868T) in the gene encoding the primary HIV-1 receptor, CD4, which encodes an amino acid substitution of tryptophan for arginine in the third domain. The purpose of this study was to determine the effect of this polymorphism on HIV prevalence, incidence, and rate of disease progression in a cohort of highly exposed women. CD4 genotyping was performed on 364 women from an open cohort of female commercial sex workers from Nairobi, Kenya (ML cohort), 284 women from an antenatal clinic in Nairobi (MCH cohort) and on 257 of their children. Participants in the ML cohort included 144 initially HIV-1 uninfected individuals (57 of whom seroconverted during follow-up) and 220 who were HIV-1 infected at the time of enrollment. The MCH cohort participants included 76 HIV-1 uninfected individuals, 31 seroconverters and 177 HIV infected women. MCH children included 167 HIV-1 uninfected children, 19 seroconverting children (late transmission due to breast feeding) and 71 children infected with HIV-1 at the time of birth. Bi-directional DNA sequencing was performed to determine their CD4 genotype. The CD4 868T polymorphism was found to be similarly prevalent in HIV-1 infected and HIV-1 uninfected groups. Kaplan Meier survival analyses showed that 868T heterozygosity was associated with incidence of HIV infection i.e. faster time to seroconversion (Log-rank $p = 0.0046$, Wilcoxon $p = 0.0088$) in the ML cohort. When the polymorphism frequency was analyzed with respect to disease progression, seroconversion, and CD4 T-cell dynamics, ML 868T heterozygotes were shown to be 2.936 times more likely to be rapid progressors ($p = 0.0197$, CI95 1.187-7.262) and Kaplan Meier analysis showed a trend towards a more rapid CD4+ T-cell decline among ML participants (log-rank $p = 0.$ 0.0910).

Homozygosity for the polymorphism was associated with more rapid disease progression in Kaplan Meier analyses of MCH CD4+ T-cell decline to both less than 400 cells per cubic mm ($p = 0.04$) and less than 200 cells per cubic mm ($p = 0.0295$). Mothers heterozygous and homozygous for the CD4 868T polymorphism showed an increased rate of breastmilk transmission of HIV-1 to their children as compared to wild type mothers (868C/868T $p = 0.0316$ O.R. 4.522 CI95 1.142-17.902, 868T/868T $p = 0.0021$ O.R. 19.5 CI95 2.947-129.026). In addition, children were shown to have a higher gene frequency of CD4 868T than adults ($p = 0.003$). In these populations, the presence of the CD4 868T allele is not associated with HIV prevalence but is associated with HIV incidence, and once infected more rapid disease progression. It is also associated with an increased rate of mother to child transmission of HIV-1 by breastfeeding.

Dedication

I would like to dedicate this thesis to the Kenyan Mothers and their children.

Acknowledgements

I first of all want to thank my supervisor Dr. Keith Fowke for his support and encouragement and for teaching me how to ask and answer scientific questions. I would also like to thank my advisory committee, Dr. Teresa Zelinski and Dr. John Wylie. Dr. Xiaojuan Mao, Dr. Ma Luo and Sue Ramdahin were absolutely invaluable in teaching me the fine art of sequencing. Dr. Ma Luo also deserves recognition for graciously letting me use her HLA typing data. Sue was especially helpful to me in so many ways – from deciphering databases to locating long-ago catalogued samples. Thanks to Mary Cheang for all her help in statistical analysis. I also want to acknowledge Jonathan Bellas for developing the ARMS-PCR protocol. I also have to thank the department administrative support, Angela Nelson, Sharon Tardi, Jude Zieske and Robin Scott among others that kept me on track. I would also like to thank Dr. Joanne Embree for letting me forage into the MCH cohort and for helpful discussion and direction especially regarding pediatric HIV infection.

A huge thank you to Dr. Frank Plummer and his lab including Leslie Slaney and John Rutherford. Dr. Blake Ball for demonstrating that genetics and immunology can live in the same brain and Jillian Waruk, Tammy Stuart, Allison Land, Lyle McKinnon and Shezhad Iqbal for invaluable friendship and support.

Thanks is probably not the right word to acknowledge the members of my lab, Steve Wayne, Paul McLaren, Sandy Koesters, Judie Alimonti, Julius Oyugi and Sarah Myk who have been ever supportive, helpful and when all else fails, endlessly entertaining. They showed me how it's done.

Finally and most importantly, I would like to thank the women and children of the ML and MCH cohorts as well as the staff at both clinics. This study was only possible because of their cooperation and dedication.

Table of Contents **Page**

Abstract 1

Dedication 3

Acknowledgements 4

Table of Contents 6

List of Figures 8

List of Tables 10

Introduction 11

1 Overview 11

2 Discovery of HIV as the causative agent of AIDS 11

2.1 discovery of AIDS 11

2.2 discovery of HIV 12

2.3 HIV transmission 13

3 HIV virology 14

3.1 structure and taxonomy 14

3.2 HIV-1 viral life cycle 15

4 HIV-1 Replication and AIDS 17

4.1 The human immune system 17

4.2 Clinical presentation of HIV-1 disease 18

4.3 Disease progression is variable 21

5 Immunogenetics 22

6 Genetic Mechanisms affecting HIV susceptibility and disease progression 23

7 CD4 Molecule 25

7.1 CD4 protein function 25

7.2 CD4 is a signaling molecule 26

8 CD4 868T 29

9 Hypothesis and Objectives 32

Materials and Methods 33

Study design and cohorts 33

Identification of Disease Progression Groups 35

Sample collection 39

DNA Isolation 40

DNA isolation from plasma 41

ARMS-PCR Screening Assay 41

CD4 Sequence Analysis Confirmatory Assay 46

Nested PCR for Plasma Samples 49

Statistical Analysis 49

Results 53

ARMS-PCR Screening assay 53
ML COHORT 56
CD4 868T not associated with ML HIV prevalence 56
CD4 868T associated with ML HIV incidence 56
CD4 868T association with ML HIV disease progression 61
CD4 868T association with lower CD4 counts in ML cohort 61
Confounding variables 65
MCH COHORT 65
CD4 868T not associated with MCH HIV prevalence 65
CD4 868T associated with MCH rapid CD4+ T-cell decline to <400 cells/mm^3 68
CD4 868T associated with rapid MCH CD4+ T-cell decline to <200 cells/ mm^3 71
MOTHER TO CHILD TRANSMISSION IN MCH COHORT 74
CD4 868T not associated with HIV prevalence in Kenyan children 74
CD4 868T not associated with MCH children's CD4+ T-cell decline to less than 25% 76
CD4 868T not associated with MCH children's CD4+ T-cell decline to less than 20% 79
Mother's CD4 868T genotype does not affect in utero HIV transmission 79
Mother's CD4 868T genotype associated with increased breast milk HIV Transmission 82
Prevalence of 868T is higher among Kenyan children than among Kenyan adults 87

Discussion 91

Appendix 1 104
Solutions

Appendix 2 105
Abbreviations

References 108

List of Figures **Page**

1. Organization of the HIV-1 genome and the protein products of the genes. 16
2. Proposed CD4 signal transduction pathway. 27
3. CD4 genetic neighborhood. 31
4. ML cohort flow chart. 36
5. MCH mothers cohort flow chart 37
6. MCH children cohort flow chart 38
7. ARMS-PCR analysis of genetic single nucleotide polymorphism at position 868 of the nucleotide sequence of the CD4 gene. 44
8. Electropherograms of individuals a) homozygous for the wild type cytosine at nucleotide position 868 of the CD4 gene, b) heterozygous for 868T and c) homozygous for 868T. 50
9. CD4 868T prevalence in low-risk and high-risk individuals as screened by ARMS-PCR. 54
10. Kaplan Meier plot of time from ML enrollment to HIV-1 seroconversion i.e. HIV incidence. 59
11. Kaplan Meier plot of ML CD4+ T-cell decline to less than 400 cells/mm^3. 63
12. Kaplan Meier plot of MCH CD4+ T-cell decline to less than 400 cells/mm^3. 69
13. Kaplan Meier plot of MCH CD4+ T-cell decline to less than 200 cells/mm^3. 72
14. Kaplan Meier plot of MCH children's CD4+ T-cell decline to less than 25% CD4+ cells from total lymphocytes. 77
15. Kaplan Meier plot of MCH children's CD4+ T-cell decline to less than 20% CD4+ cells from total lymphocytes. 80

16. Breast milk mother to child transmission of HIV based on mother's CD4 genotype. 85

17A. Age dependent distribution of CD4 868T. A) All children and adults tested in this study. 89

17B. Age dependent distribution of CD4 868T. B) Total HIV uninfected children and adults tested in this study. 89

List of Tables **Page**

1. ARMS-PCR primer sequences. 42
2. Sequencing primers. 47
3. HIV Disease Progression and Frequency of CD4 genotype and CD4 allele in ML cohort. 57
4. Association of biographical data with CD4 868T genotype in ML cohort. 66
5. MCH HIV status and frequency of CD4 genotype and CD4 allele. 67
6. MCH children HIV status and frequency of CD4 genotype and CD4 allele. 75
7. *In utero* mother to child transmission of HIV based on mother's CD4 genotype. 83
8. Breast milk mother to child transmission of HIV based on mother's CD4 genotype. 84

1. Overview

The AIDS pandemic is arguably the worst crisis the global community has had to face in the last three decades. With between 34 and 46 million people currently living with HIV infection (1), it is more than a health crisis but also a social and political one as well. Entire generations have been orphaned in the developing world and the disease has single handedly reversed development in some countries by devastating the youth and skilled labour and greatly reducing the average life expectancy. The progress of the pandemic does not seem to be abated yet with 5 million new infections in 2003 and 3 million deaths, 500 000 of them children under the age of 15 (1).

Though treatments are available in the industrialized world, they are not widely accessible in the developing world nor are they the sole answer to slowing the pandemic. Only effective prevention strategies including a prophylactic vaccine can actually be useful tools to end the current crisis. Until such time as an effective vaccine can be instituted, efforts must be continued to better understand the virus and its pathogenesis in order to provide information for treatment and vaccine design.

2. Discovery of HIV as the causative agent of AIDS

2.1 discovery of AIDS

A group of five otherwise healthy men were diagnosed with *Pneumocystis carinii* in 1981 and reported to the Centers for Disease Control in Atlanta, USA (2). By 1982, several other cases were reported as well as other conditions that are normally associated with

severe immunodeficiency (3). This condition was termed acute immune deficiency syndrome or AIDS and was notably afflicting gay men, intravenous drug users and individuals who had received frequent blood transfusions, such as hemophiliacs. AIDS was later noted in other groups as well and the evidence from these incidents indicated a possible infectious causative agent (4, 5).

2.2 discovery of HIV

The possible causative agent was first isolated in 1983. A human retrovirus similar to human T-cell leukemia virus (HTLV) was isolated from the lymph node of a man suffering from AIDS (6). The same year, Gallo et al isolated a similar virus from another AIDS patient (7). The various groups studied their isolates and soon discovered that they were not in fact HTLV but were very similar and so named the new virus HTLV III, LAV and ARV depending on the research group involved.

The virus was recognized as novel and named Human Immunodeficiency Virus (HIV) and classified as a lentivirus (8). In 1986, a second separate subtype was identified in Western Africa (9) which is transmitted in much the same way as the initial isolates but had a longer disease progression. This subtype was named HIV-2 with the original etiological agent of the North American and European AIDS being named HIV-1.

HIV-1 has a large capacity for mutation. Since its discovery, many different strains have emerged and forced a classification system to be implemented. There are three larger

groups named M (major), N (non-M, non-O) and O (outliers). M is by far the largest and most common group and is subdivided into clades. Clades A and C are more closely related to each other than to the others and clades B and D are more closely related to each other than they are to A and C. There is some disagreement as to the other named clades and whether or not they are true clades or they are recombinants of the four larger clades due to coinfection of a patient. Both HIV-1 and HIV-2 are thought to have been transmitted to humans from monkeys and apes that harbour simian immunodeficiency viruses (SIV). The closest related SIV to HIV is that found in chimpanzees and this virus is a descendant of the red-capped mangabey SIV and that of the greater spot-nosed monkey. The SIVcpz of *Pan troglodytes troglodytes* is thought to have given rise to the M group viruses while SIV cpz of *Pan troglodytes schweinfurtheii* is associated with the N group HIV-1 viruses. HIV-2 is closely related to SIV of sooty mangabeys. Since various SIVs are associated with varying groups of HIV, it is believed that multiple zoonotic jumps were required to produce the HIV variety seen today (10).

2.3 HIV transmission

HIV-1 is transmitted through bodily fluids and is not an airborne or aerosol-acquired virus. Horizontal transmission is accomplished through sexual contact (both heterosexual and homosexual) and through the exchange of blood, blood products and other fluids encompassing both medical procedures and intravenous drug use for example (11). Mother to child transmission is also another significant route of transmission. Infants can acquire the virus from their mothers at various points during pregnancy (*in utero*), during

delivery or through breastfeeding (12). In the developed world with the advent of antiretroviral therapy (ART) administration during delivery and adequate alternatives to breastfeeding, mother to child transmission has virtually been eliminated. It is still a very significant route of transmission in the developing world, however, where ARTs are not always available and where there are no viable options to breastfeeding. Under these conditions, HIV infected pregnant women will infect their infants in 25% of the cases *in utero* or at delivery and breastfeeding their children for 2 years will increase their chances of transmitting the virus to 45% (13-15).

3. HIV virology

3.1 structure and taxonomy

HIV is a member of the *lentivirinae* (lenti – slow) which are part of the family Retroviridae (named for their DNA intermediate stage). It is an enveloped, helical virus with a diploid positive sense RNA genome of approximately 9 kb. The virion is relatively small, about 110-120 nm in diameter. The envelope is composed of host cell plasma membrane which, unlike many viruses, still contains many host cell surface proteins.

The HIV genome is composed of three genetic regions (Figure 1): firstly *gag* or group specific antigen which comprises the capsid proteins, p24, p7, p9 and the matrix protein, p17. The second, *pol* encodes the polymerase gene reverse transcriptase, the integrase gene and the protease gene. The third region, *env* encodes the envelope proteins gp160

which associate as trimers on the surface of the virion once they are cleaved to form the gp120 (attachment) and gp41 (fusion) portions of the spike. Several accessory and regulatory proteins are also encoded by the genome namely rev, nef, vif, vpr, vpu, vpx and tat. The genome is also flanked by two terminal repeats and unique non-repeated regions internal to the terminal repeats named U5 at the 5' end and U3 at the 3' end (16).

3.2 HIV-1 viral life cycle

HIV has a very narrow host range and a very specific cell tropism in that host. The virus attaches to its primary cellular receptor, the cell surface protein CD4 (17) through its gp120 envelope spike in order to mediate entry. A second coreceptor (CCR5 or CXCR4 or some other minor cell surface receptors) interaction is also required which creates a conformational change in gp 120 revealing gp 41, the fusion protein which permits the virion to fuse with the plasma membrane and allows entry.

CD4 is found primarily on CD4+ T-helper lymphocytes but is also found on macrophages and dendritic cells (18). CCR5 coreceptors are primarily found on macrophages. Viruses that use this coreceptor are termed R5 viruses. CXCR4 is primarily found on T-lymphocytes and viruses that use this coreceptor are termed X4 viruses. Once the virus enters the cell, uncoating occurs and the genome is exposed. The key identifying feature of retroviruses is their peculiar way of generating their viral mRNA's. Unlike most positive sense RNA genome viruses that use their genomes directly as message, retroviruses engage in an additional step using reverse transcriptase

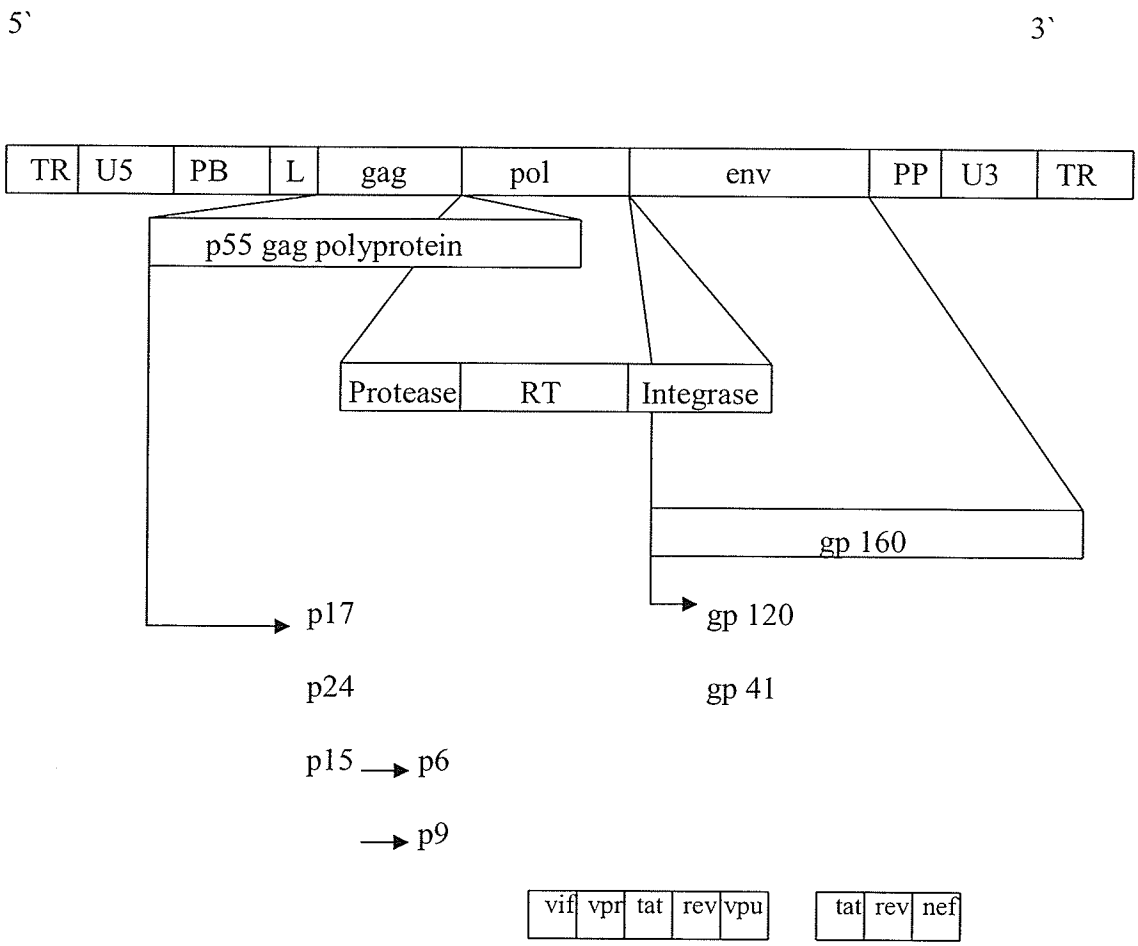


Figure 1: Organization of the HIV-1 genome and the protein products of the genes.

to generate a cDNA copy of their genomes. The generation of the cDNA is done using a lysine tRNA as a primer (19). This DNA copy (called provirus) is then imported into the nucleus, possibly by Vif and Vpr, where it is integrated into the host genome by Integrase. This integrated copy is then converted to mRNA using host machinery. The first mRNA's produced are those for Tat, Rev and Nef, all regulatory protein products from short multiply spliced mRNA. The viral proteins are then made by host translational machinery in the cytoplasm. Tat binds the 5' end of the provirus and increases transcription. Rev binds RNA structures in the *env* region to allow the nuclear export of incompletely spliced mRNA. Nef has multiple functions not the least of which is downregulation of MHC molecules and the CD4 molecule on the infected cell's surface. Vif associates with the cytoskeleton to increase the infectious potential and Vpr associates with the matrix protein to provide a nuclear localization signal. Once gp 160 is produced, it migrates to the cell membrane and is cleaved into gp120, gp 41, Vpr, Vpu and Vif during virion maturation (20).

The capsid proteins associate with progeny genome made in the nucleus and associate with matrix protein that associates with the gp160 embedded in the cellular membrane. The virion is then ready for budding (20).

4. HIV-1 Replication and AIDS

4.1 The human immune system

The immune system is composed of two branches. The innate immune system is the first line of defense and is not antigen specific but involves cells and chemical signals that produce inflammation. The second branch termed the acquired immune system is antigen specific so must be "trained" to recognize and produce the full effective response to an invader. The acquired immune system has two arms called humoral immunity and cell-mediated immunity. The effector cell of humoral immunity is the B-lymphocyte that produces immunoglobulins (antibodies) that recognize and bind antigens. These antibodies can prevent infectious agents from infecting or binding their targets in the body. They can also recruit cells of the innate immune system called natural killer cells through antibody-dependent cell-mediated cytotoxicity (21). The cells of cell-mediated immunity are CD8+ cytotoxic T-lymphocytes. These cells recognize and kill cells infected with viruses or are otherwise displaying non-self characteristics like cancerous cells. CD4+ T-lymphocytes are cells that control both arms of acquired immunity and can modulate the innate immune system as well as receive information from the innate system. CD4+ T-cells recognize foreign proteins bound to host MHC II molecules. These cells then produce cytokines that help B-cells produce antibodies or recruit and activate other effector cells. Without them, the immune system is incapable of proper function (21).

4.2 Clinical presentation of HIV-1 disease

Both the Centers for Disease Control and the Walter Reed Army Medical Center in the United States have a classification system which outlines the course of HIV infection and

the development of AIDS. The Walter Reed system has 6 stages while the CDC has 3 clinical categories which are then subdivided into 3 CD4+ T-cell count (cells per cubic mm of blood) categories.

The Walter Reed system begins at stage 0 which is the point at which infection first occurs. Since this is generally the first encounter with the virus, the immune system goes through a primary response and allows a very large viral replication burst to occur. The highest viral titres occur during this primary acute stage and it is at this time that an infected individual is most likely to transmit the virus. This stage last approximately six weeks until a full immune recognition occurs and lowers the viral titre. Flu-like symptoms may occur during this stage.

Stage 1 begins when the body starts to fight back. Antibodies (humoral response) and cell-mediated immunity begin to be raised against the invading pathogen. Specifically, the antibodies made are to prevent viral attachment to the host cell receptor and CD8+ T-cells, and the mediators of cell-mediated immunity begin to recognize and kill infected CD4+ cells. Stage 2 then begins when the infected cells begin to gather in the lymph nodes and a large daily CD4+ T-lymphocyte turnover begins. This stage lasts 3-5 years. During this time, CD4+ T-cell counts remain normal though they do slightly decline.

Stages 3 and 4 are linked. During this time the CD4+ T-cell decline goes from 700 cells/cubic mm to 400 cells/cubic mm. This stage lasts about 18 months. Once a level of 500 cells/cubic mm is reached, aggressive chemotherapy such as highly active

antiretroviral therapy (HAART) is begun in developed nations. Stage 4 begins when the CD4+ T-cell count remains steadily under 400 cells/cubic mm.

Stage 5 is when the patient has CD4+ T-cell counts below 400 cells/cubic mm but remain above 200. A patient is considered to be in stage 5 if they develop an opportunistic infection as this is a definitive sign that HIV is winning the battle.

In stage 6, the CD4+ T-cell count remains below 200 cells/cubic mm and numerous opportunistic infections manifest themselves including pneumonia caused by *Pneumocystis carinii* and Kaposi's sarcoma, a rare herpesvirus 8-mediated cancer. Patients in stage 6 are considered to have AIDS and on average survive no more than 5 years (22). The CDC ranks stage 6 patients as category C 1993.

These stages do not apply to children. It has been reported that children born infected with HIV develop clinical symptoms sooner (23) and have the same predictive factors for progression and course of infection (without treatment) in developed and developing nations (24). Disease progression is more rapid in girls than in boys and peak viremia occurs within the first 6 months of life (24). Since children's CD4 T-cell counts are variable during childhood, hard counts are not commonly used as measurements of health in HIV infected children unless the age of the child is taken into the calculation. CD4+ T-cells as a percentage of total lymphocytes are reported, with 45% being the normal state and 20% being considered AIDS stage illness (25).

When an individual becomes infected with HIV, the immune system responds in a major cascade. In fact, the period of time between infection and the development of AIDS which was once deemed a "latency" period has now been characterized as a period of time when the immune system still has the upper hand. Up to 10^{10} virions are produced and cleared each day in an asymptomatic HIV-infected person (26-28). The body produces large quantities of cytotoxic T-lymphocytes and neutralizing antibodies (immunoglobulins that can prevent viral infection) but due to the progressive loss of CD4+ T-cells, these responses cannot be properly managed and become less effective (29-32).

4.3 Disease progression is variable

The system of disease stage classification is both generalized and based on observations in the developed world. In the developing world, disease progression is more rapid (33) and there have also been several reports of great variation in the length of time between infection and the development of clinical AIDS (33-36). These findings have prompted the description of several progression classifications and fueled research into individual differences that could be responsible for this variation.

The first observed condition was that of long-term non-progression. Individuals infected with the virus were shown to live longer (some have survived 20 years) than what was predicted with CD4 T-cell counts above 400 cells/cubic mm of blood (35, 37). The other end of the spectrum was also observed as well and termed rapid progression. These are

patients whose CD4 T-cell counts drop dramatically to less than 200 cells/cubic mm usually with 3-5 years of initial infection (33).

Several factors could be involved in why one patient will progress differently from another. Access to antiretroviral therapy, initial viral load set point, age at onset of infection, geographic location and various genetic and immunological factors have been postulated as being associated or at least possibly involved in the rate of progression of infection (36).

5. Immunogenetics

A patient's genetics can affect the outcome of disease. Various genetic disorders have been identified and recently certain genetic advantages or disadvantages have been shown to affect the course of infectious disease. These changes are part of a genetic variability that affects evolution and underscores the observation that no disease or organism can boast a 100% infection rate.

Certain genetic changes can be associated with human disease states. For example, asthma-related illness has been attributed to genetic alterations in afflicted individuals. A particular single nucleotide polymorphism (SNP) haplotype mapped to the 133 kilobase region containing the orphan G protein-coupled receptor GPRA (G-protein receptor for asthma susceptibility) has been recently associated with both high levels of serum IgE and with classical asthma symptoms (38).

It has already been established that certain membrane receptor alleles can affect the progression of infectious diseases. For example, the presence of at least one copy of a particular MHC class II allele, HLA-DQB1*0503, has been associated with clinical cases of *Mycobacterium tuberculosis*-caused tuberculosis. This SNP causing an amino acid change seemed to be present more often in individuals who progressed to clinical symptomatic tuberculosis than in the infected, asymptomatic individuals in this Cambodian cohort (39). This work suggested, much like our study, that the SNP didn't affect the susceptibility to infection by the bacillus but that it did cause an altered disease progression. Similar results have been shown in studies of NRAMP1 polymorphisms and the susceptibility and progression of *Mycobacterium leprae* caused leprosy (40).

6. Genetic mechanisms affecting HIV susceptibility and disease progression

Many previous studies have shown that genetic alterations in genes important to the immune system can affect susceptibility to HIV infection and HIV disease progression. A 32 bp deletion in the gene for CCR5, an HIV coreceptor, abolishes the expression of that gene. Individuals homozygous for this mutation have been shown to be resistant to infection with R5 and dual tropic viruses (41-43). Heterozygotes express low levels of CCR5 and have been shown to have delayed disease progression though demonstrated no significant changes in susceptibility to infection (44-47). When examined for its role in mother to child transmission, the deletion showed no effect on susceptibility regardless of

homozygous vs. heterozygous state but did delay progression of illness in children with the mutation (48, 49).

Although CCR5 mutations are not lethal, CXCR4 mutations are rare because this molecule is vital to cells which express it. Mutations have been noted in its ligand, SDF-1. The SDF1-3'A mutation is a SNP in the 3' untranslated region of the transcript. This change has been shown to lower susceptibility to infection in adults but showed no effect in mother to child transmission (50). The only association noted was in one study which showed that the SDF-1 genotype of the mother affected whether or not her child became HIV-infected (51). The manner in which this polymorphism is proposed to affect susceptibility is unknown although it is speculated that the mutation increases translation of the mRNA causing an excess of the ligand that outcompetes HIV for CXCR4 (50).

An IL-4 promoter polymorphism was identified and associated with delayed disease progression although very little work has been done to explain the relationship leading to this observation (52).

An amino acid substitution in CCR2, a minor HIV coreceptor, was identified and showed to be located in one of the seven transmembrane domains of the receptor. Though a functional explanation could not be found, it was associated with delayed disease progression. It was later discovered to be in almost complete linkage disequilibrium with a CCR5 promoter polymorphism which is associated with a decreased expression of CCR5 which also delays disease progression (53, 54).

The only polymorphism in the literature to date which has been associated with accelerated disease progression is located in the gene for CX3CR1. This is a very minor HIV coreceptor but its major ligand is fractalkine, a lymphocyte chemoattractant molecule. It is believed that the structurally different CX3CR1 molecule binds its ligand less efficiently causing a general down-modulation of the immune system allowing HIV more access to immune cells without immune system interruption (55-57).

7. CD4 Molecule

7.1 CD4 protein function

Maddon et al [12] described CD4 as the primary HIV-1 receptor in 1985. It was shown that the molecule, generally found as a dimer, is a highly conserved, 55kDa single chain class I glycoprotein expressed on 60% of peripheral blood T-lymphocytes. CD4 has a cytoplasmic tail for signal transduction and 4 immunoglobulin-like extracellular domains with the most distal named D1 and the most proximal named D4. D1 has been found to contain the HIV-1 envelope gp120 protein-binding region. The critical amino acid involved in that interaction is the phenylalanine at position 43 which binds inside the gp 120 binding pocket (58). The molecule is known to be expressed on many types of cells, most importantly on macrophages and T-helper cells. Primarily, CD4 was shown to bind the conserved regions of major histocompatibility complex class II (MHC II) as its natural ligand, accounting for its lack of variability (59). This lack of variation has been exploited by HIV-1 and at least one member of the human herpesviruses (HHV-7) as a suitable receptor target (60, 61). The T-cell receptor (TCR) recognizes a specific antigen

peptide which is presented by MHC II on antigen presenting cells (B-cells, macrophages, etc.). CD4 enhances and stabilizes that interaction.

7.2 CD4 is a signaling molecule

When CD4 binds its natural ligand, a signal transduction pathway begins (Figure 2). During T-cell activation, CD4 associates with the TCR/CD3 complex. CD4's cytoplasmic domain is the vital element for signaling. The cytoplasmic tail is non-covalently associated through two cysteine residues with Lck, an Src-related protein tyrosine kinase of 56kDa. Activation of this protein is critical to intracellular CD4 signaling. Lck activation leads to the phosphorylation of the CD3 immunoreceptor tyrosine-based activation motifs (ITAMs). The phosphorylation of ITAMs in the zeta chain allows zeta associated protein 70 kDa (ZAP-70) to interact with CD3. Once this happens, ZAP-70 becomes phosphorylated. ZAP-70 activation leads to the initiation of a number of signal transduction pathways. An early event following Lck activation is the activation of Ras proteins which are the 21kDa gene products of *ras* oncogenes. These are important for normal cell division. If Ras is bound to guanine triphosphate (GTP), it is in its active form, while if it is bound to guanine diphosphate (GDP) it is in its inactive form. Ras has been shown to be dependent on Lck tyrosine kinase upstream for T-cell activation. Between CD4/Lck/ZAP-70 and Ras are a number of possible adapter proteins including Shc which becomes phosphorylated after Lck activation and seems to control the level of Ras activation (62). ZAP-70 also activates phospholipase C gamma which in turn produces IP_3 (inositol 1,4,5 triphosphate) and DAG (diacylglycerol). IP_3 activates

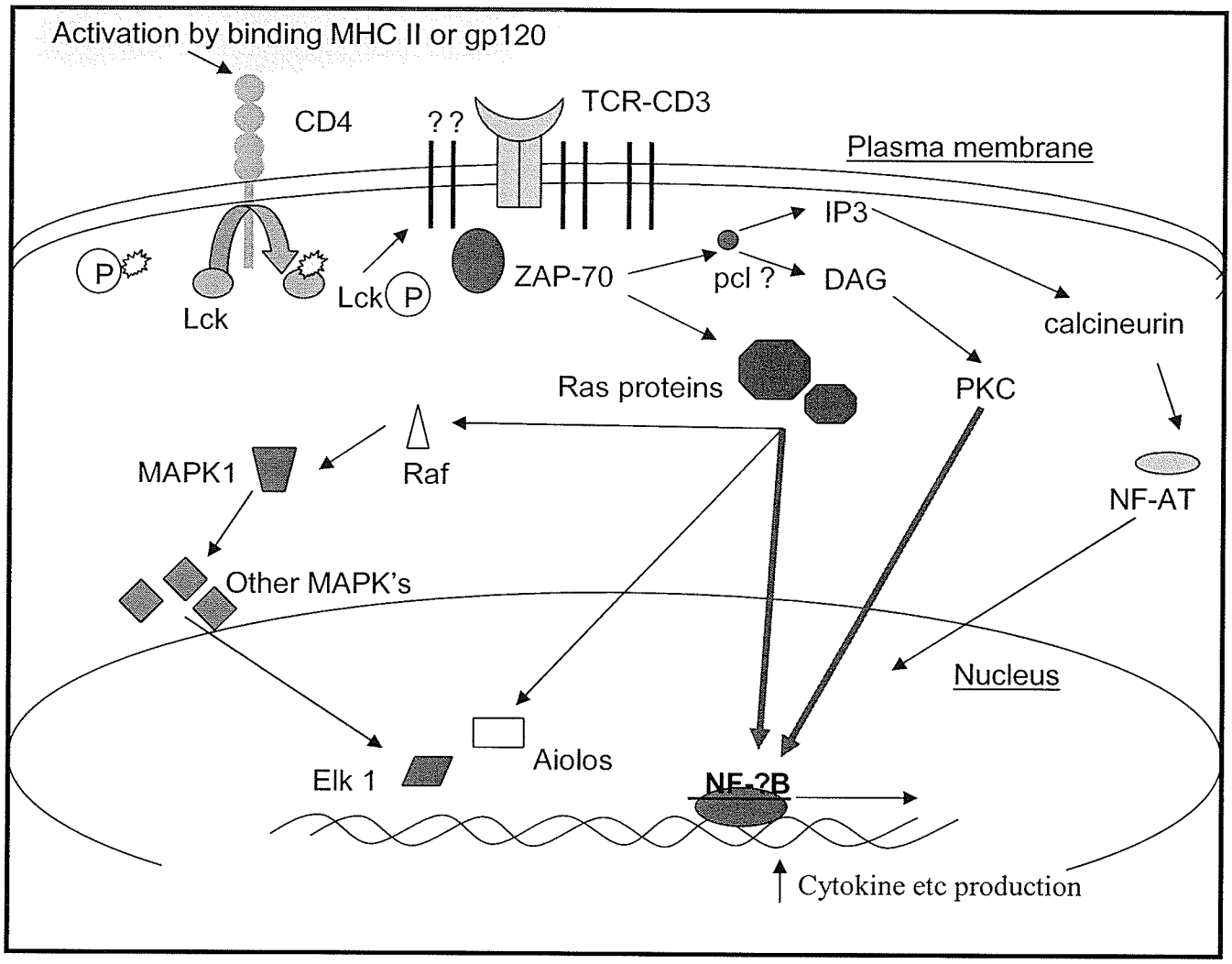


Figure 2: Proposed CD4 signal transduction pathway.

Figure 2: Proposed CD4 signal transduction pathway. CD4, once activated, triggers a signal transduction cascade resulting in the increased transcription of NF-kappa B regulated immunomodulatory genes (62, 63).

calcineurin through an increase in intracellular Ca^{2+}, which activates NF-AT by dephosphorylation (a T-cell specific nuclear factor). DAG activates PKC (protein kinase C) which also activates NF-?B.

Multiple signaling molecules for regulation of many gene types including both pro- and anti-apoptotic pathways, use the Ras-mediated signal pathway. Activated Ras can transmit a signal to many molecules, the most common being Raf. Raf activates mitogen activated protein kinase 1 (MEK1 or MAPK1) which then activates a series of MAP kinases that eventually translocate to the nucleus and activate nuclear transcription factors including Elk1. Ras can also directly associate with the transcription factor Aiolos – a mediator of Bcl-2 inhibition related apoptosis - and interleukin-2 deprivation especially enhances this interaction. The most important transcription factor activated by Ras however is nuclear factor kappa B (NF-?B) which is important in the regulation of cytokine, chemokine and growth factor genes (63). Cytokines are a vital part of CD4+ T-cell communication to the immune system. They permit the recruitment and activation of other immune cells including cytotoxic T-lymphocytes (cell-mediated immunity) and B-cells (humoral or antibody-mediated immunity).

8. CD4 868T

The CD4 gene is located on chromosome 12 on the p arm at position 12 (Figure 3). It is near the gene for B7 isoform, also called B7-2. The gene has 3084 coding base pairs spread out in 10 exons among 32 297 total base pairs (GenBank accession number

U47924). Presently, few polymorphisms in the gene for CD4 have been identified and none have been studied for their effect on HIV-1 infection and progression. Of the eight known alleles of CD4, only one polymorphism is reported to occur in the coding region (64). A single nucleotide polymorphism from a cytosine to a thymidine located at position 868 of the nucleotide sequence encodes an amino acid change from arginine to tryptophan in D3 at amino acid position 240.

The SNP is an autosomal codominant allele causing both wild type CD4 and CD4 868T to be expressed on the surface of CD4+ cells in heterozygous individuals (65). This creates the possibility that the heterozygous state may also exert an effect. CD4 868T heterozygotes have been found at a 20% frequency among African-American adults and very rarely (>1%) among Caucasians (65). It is not unusual for polymorphisms to have different prevalences depending on ethnic backgrounds; a polymorphism of human caspase-12 is found almost exclusively among populations of African descent (66). The initial identification of this SNP stemmed from the observation that some people of African descent had CD4 molecules that were unable to bind an IgG antibody named anti-OKT4 that was directed at an epitope in the third domain of CD4 (67). When sequence analysis revealed 868T, it was hypothesized that this amino acid change caused alterations in the tertiary structure of CD4's third domain which was later supported by Kyte-Doolittle hydrophobicity analysis (64, 65, 67).

Even though the polymorphism is in D3 of CD4 while the gp120 portion of HIV-1 binds D1 of CD4, the location of the polymorphism could still have a dramatic effect on the

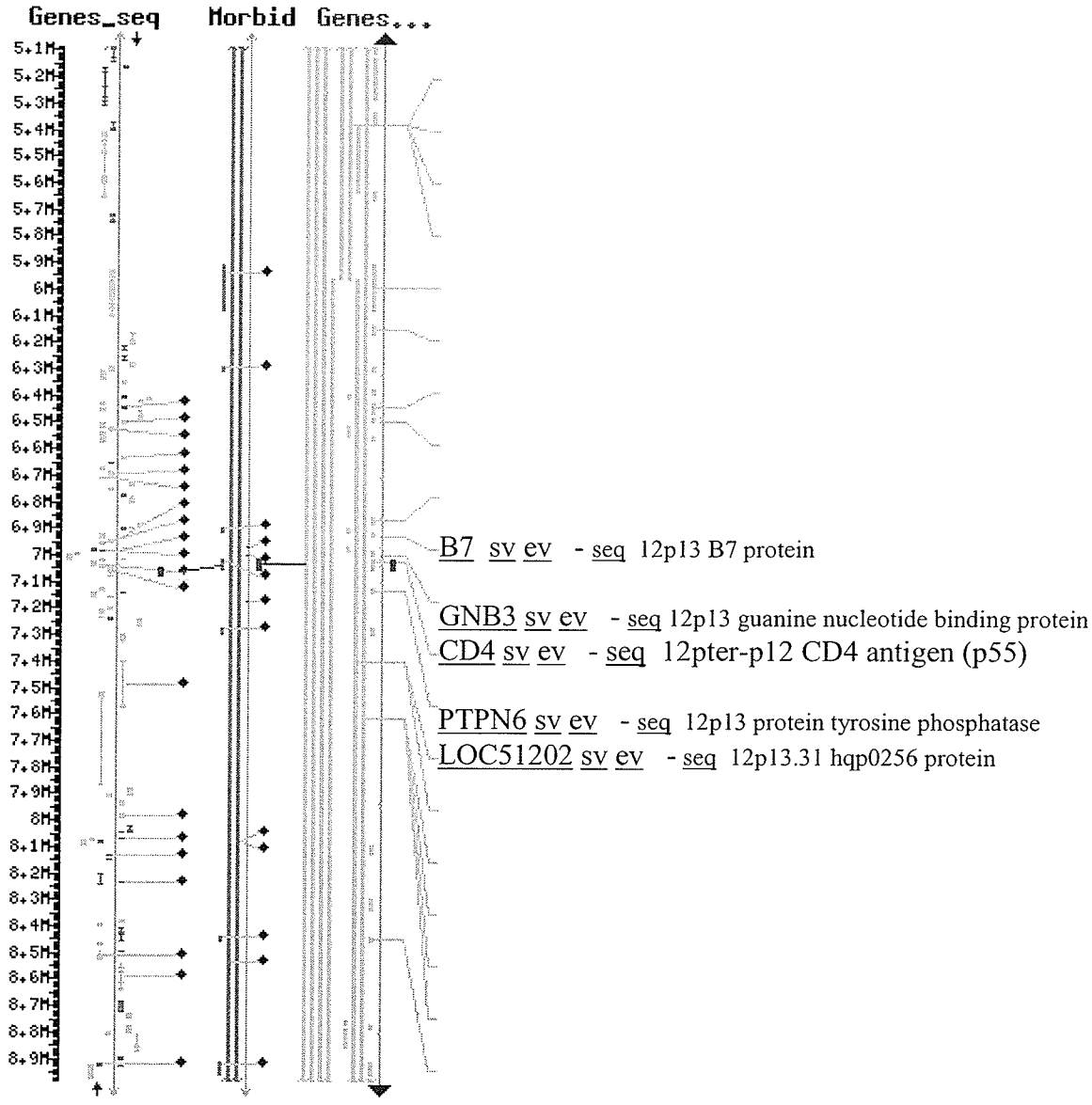


Figure 3: CD4 genetic neighborhood. Portion of the terminal end of the p arm of chromosome 12 (NCBI).

interactions between gp120, or the chemokine secondary receptors and CD4. For example, an amino acid substitution in a non-active site locus of MHC I can significantly alter the peptide repertoire and T-cell recognition of the MHC molecule (68). Specifically in the case of CD4, recent data does suggest that the redox state of the disulfide bond in D2 of CD4 greatly affects the binding of gp120 (69).

9. Hypothesis and Objectives

The hypotheses for this project are two-fold. We hypothesize that the presence of the CD4 polymorphism 868T is a genetic risk factor for seroconversion and rapid disease progression in adults and that it is also a genetic risk factor for mother to child transmission of HIV.

In order to explore this hypothesis we propose two specific objectives:

1) To determine if the CD4 868T polymorphism is correlated with susceptibility to HIV infection and disease progression in adults
2) To assess the epidemiology of CD4 868T in mother to child HIV transmission

Materials and Methods

Study design and cohorts

The Majengo Commercial Sex Workers cohort and the Mother-Child Health cohort from Nairobi, Kenya were utilized for this study. These protocols were approved by the University of Manitoba Ethics Review panel and the Kenyatta National Hospital Ethics Review Committee.

The ML study of commercial sex workers is an open cohort followed since 1985 in the Pumwani district of Nairobi, Kenya (70-73). Female sex workers attending the Majengo Sexually Transmitted Infection Clinic were enrolled. Informed consent was obtained from willing participants and they were assigned a unique study number and identification card to confirm their identity on subsequent visits to the clinic. Women were asked to return to the clinic at least twice a year for follow-up. At each visit, a physical examination was performed and demographic, sexual, medical, obstetric and contraceptive histories were obtained. Serum samples were obtained at each visit and tested for a variety of sexually transmitted infections (STI's) and for HIV serology and PCR. At the time of testing, the cohort had enrolled a total of 1927 women over the 18 years of operation, although at any one time approximately 500 were active members from whom samples may be obtained. An analysis of potential selection bias indicated an equal distribution of CD4 868T and CD4 868C among those who were lost to follow-up or died.

The MCH study of perinatal HIV-1 transmission was founded in 1986 in the Pumwani sector of Nairobi, Kenya (74). This study follows the mothers and their children. Women who presented to Pumwani Maternity Hospital, a major maternity hospital (one that delivers more that 25 000 children/year) in the region, were invited to enroll in the cohort study if they delivered their child between 7:30 am and 12:00 am Monday to Friday.

First the mothers' sera were screened for antibodies to HIV-1 and those found to be infected and some randomly selected seronegative women and their infants were enrolled. The sera were screened by ELISA and positive samples were repeated the following day. Those samples which were repeatedly positive were tested by Western blot and seropositivity determined by reaction to one core protein (p24, p17 or p15) and to one envelope protein (gp41, gp120/160). Women with negative Western blot, irrespective of ELISA results were considered uninfected. One seronegative mother was enrolled for every one to four seropositive mothers that entered the study. In the 24 hours following delivery, demographic, social, medical and obstetric histories as well as the circumstances of labor and delivery were obtained from the mothers and recorded. Mothers were physically examined for signs of HIV infection and infants were examined to determine general health status (including birth weight, length and head circumference) and gestational age. Two weeks following delivery, mothers were examined for sexually transmitted infections.

Follow-up was conducted in similar ways for all mother-infant pairs enrolled. Infant visits were at monthly intervals for the first 6 months of life and then every 3 months. Histories of intercurrent illnesses, breastfeeding and nutrition were taken and physical examinations were performed at each visit. All Kenyan required immunizations were administered (BCG, oral Polio, diphtheria, pertussis, tetanus and measles). Mothers and infants provided serum samples at birth and every 3 months.

In both cohorts, individuals enrolled were encouraged to visit their respective study clinic if they had health concerns for either themselves or for their children.

Individuals for this particular study were chosen based on HIV status, availability of extracted DNA samples and availability of samples for the children born to the women of the antenatal cohort. In order to perform the ARMS-PCR screening assays, 533 adult female individuals were chosen from both the ML and MCH cohorts (285 ML and 248 MCH individuals respectively). For the CD4 sequencing confirmatory assays, 648 adult females including 364 ML (Figure 4, at the time of testing, 1874 women were part of the cohort though few had available DNA samples) and 284 MCH individuals (Figure 5) and 276 children of the MCH cohort (Figure 6) were chosen. The 276 children were chosen from the children of HIV infected women of the MCH cohort.

Identification of Disease Progression Groups

Women from the commercial sex workers cohort (ML cohort) were defined as rapid progressors, long-term non-progressors and normal progressors based on the

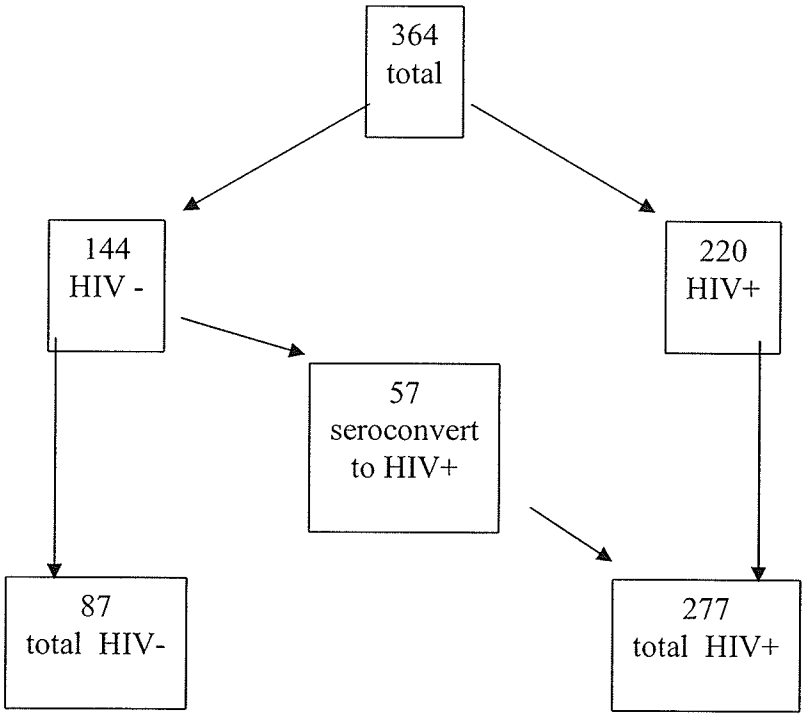


Figure 4: ML cohort flow chart

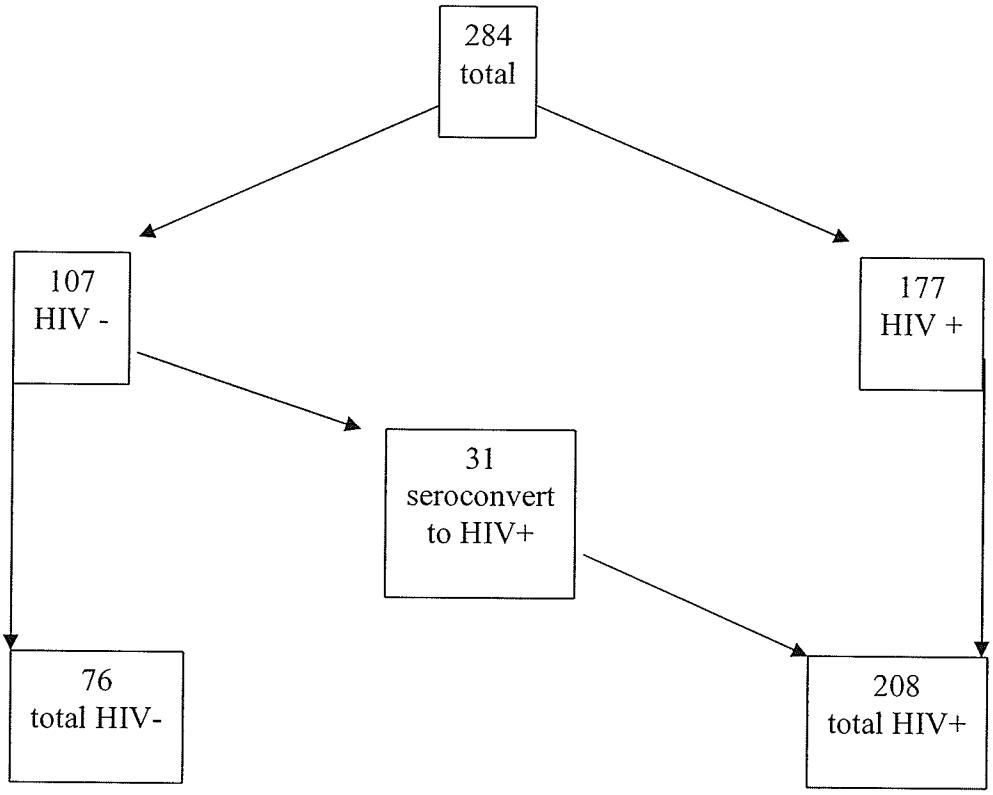


Figure 5: MCH mothers cohort flow chart

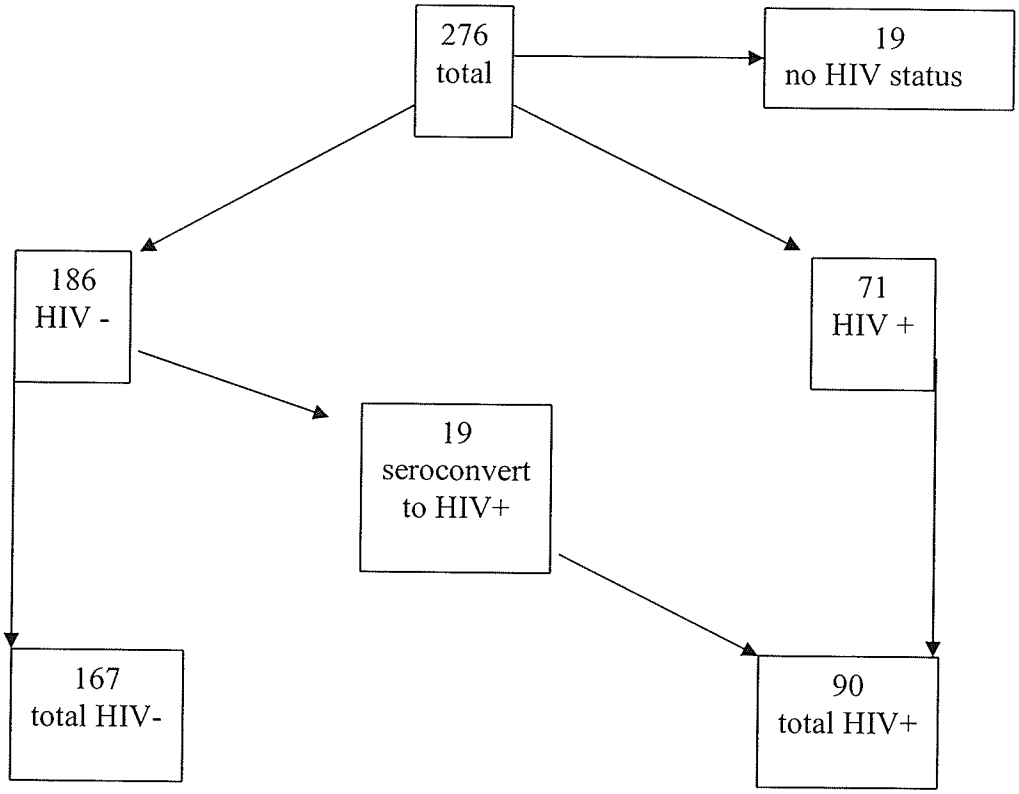


Figure 6: MCH children cohort flow chart

following definitions. Rapid progressors were defined as women whose CD4+ T-lymphocyte counts reached less than 200 cells/mm^3 within 3 years of seroconversion, long-term non-progressors were defined as women whose CD4+ T-lymphocyte counts remained above 400 cells/ mm^3 for more than 8 years after seroconversion or enrollment in the cohort. Normal progressors were defined as women who progress to AIDS (< 200 CD4+T-cells/ cubic mm) in 3-8 years following the average progression observed in sub-Saharan Africa.

Infants born to the women enrolled in the MCH cohort were tested for HIV infection by serology and PCR at birth, 6 weeks, 6 months and 9 months of age. They were then seen and tested at every 3-month appointment follow-up. If a child tested PCR positive for the virus at birth or at the first visit, they were determined to have been infected *in utero* or during delivery. If the first PCR positive test occurred at 9 months or later, the child was determined to have been infected through breast milk transmission (74, 75).

Sample collection

Samples of peripheral blood mononuclear cells (PBMC) were obtained from cohort participants in the ML and MCH cohorts. These cells were isolated from whole blood by ficol purification. Blood was drawn into sodium heparin vacutainers (Becton Dickinson) and mixed to prevent coagulation. The blood was then layered onto 1/3 volume of Ficoll-Hypaque (Sigma) and centrifuges at 500xg for 30 minutes in a table top centrifuge

(Beckman GPR) at room temperature. The plasma was then removed aseptically in a laminar flow hood for long-term cryopreservation and used for HIV-1 testing. The PBMC layer was removed in a similar way and placed in a new sterile tube. Cells were washed twice with 50 ml PBS at 400 x g for 20 minutes at room temperature. Cells were then stained with trypan blue and counted using a hemocytometer and 5×10^6/ml were cryopreserved at -70°C in freezing media (90% FCS, 10% DMSO).

DNA Isolation

DNA for PCR and sequencing procedures was isolated from between 5 and 6 x 10^6 PBMC using a QIAamp DNA Mini Kit (Qiagen). The cells were pelleted and resuspended in 200 ul of PBS (phosphate buffered saline). In order to lyse the cells, 20 ul of Proteinase K (920mg/ml) was added to the resuspended pellet and mixed. The rest of the procedure was carried out as per the manufacturer's instructions using buffers included in the kit. Each DNA sample resulting from the procedure contained a 200 ul solution of DNA eluted in the commercial buffer. The samples were stored at –20°C .

DNA isolation from plasma

Isolation of DNA from plasma was performed using the same kit as for the isolation from PBMC. 200 ul of plasma was used in place of pelleted cells. The resulting DNA was then amplified using a nested PCR strategy in order to generate enough amplicons for sequencing.

ARMS-PCR Screening Assay

Amplification refractory mutation system polymerase chain reaction (ARMS-PCR) was performed as a screening assay for the CD4 868T polymorphism using sequence specific primers for each of the two alleles. The final 3' most nucleotide position in each of these primers corresponded to position 868 in the nucleotide sequence of the CD4 gene. The allele specific primers were forward primers and both alleles used a common reverse primer to complete the PCR amplification primer pair (Table 1).

The PCR procedure was conducted by making two master mixes. One master mix contained the primer specific for the wild type allele while the other contained the primer for the 868T allele. Each mix contained 5ul of Invitrogen™ 10x PCR buffer without magnesium (Mg), 1.5ul Invitrogen™ 50mM $MgCl_2$, 8.0ul of dNTP's, 28.25ul double distilled autoclaved water, 0.25ul Invitrogen™ recombinant *Taq* DNA polymerase and 5.0ul (0.5uM per reaction) of a 5uM primer mix of both a forward and reverse primer for

Table 1: ARMS-PCR primer sequences. The sequences are listed 5` to 3`.

Allele		**Primer sequence**
Wild type 5'	nt* 923-945	5` GAACAAGGAAGTGTCTGTAAAAC 3`
868T 5'	nt 923-945	5` GAACAAGGAAGTGTCTGTAAAAT 3`
3' common	nt 1049-1071	5` TTCCTGTTTTCGCTTCAAGGGCC 3`

* nt = nucleotide

each reaction. The primer mix was made from reconstituted primers from Invitrogen™ which were set at a concentration of 50uM. The dNTP's (Invitrogen™) were made in an aliquoted mix of 25 ul of each 100mM dNTP solution in 1900 ul of double distilled autoclaved water.

Forty-eight ul of master mix was added to round-top thin wall 200ul reaction tubes followed by 2ul of desired template DNA. A negative control using deionized distilled autoclaved water as the template was added to each reaction set. As a positive control, the DNA of an individual known to be heterozygous for the 868T polymorphism, termed C1, was also included in each reaction set.

The PCR amplification was performed on an MJ Research PTC-200 thermocycler under these conditions: 94°C for 2 minutes, followed by 30 cycles of 94°C for 30 seconds, 64°C for 30 seconds, 72°C for 30 seconds with a final elongation of 72°C for 7 minutes. Each individual sample was amplified with both the wild type and 868T primers in separate reaction tubes.

These PCR products were then resolved on a 2% agarose gel by electrophoresis and scored based on whether or not bands appeared as wild type homozygous, heterozygous or homozygous for the 868T polymorphism (Figure 7). Each gel was also run using a 100bp DNA ladder (Invitrogen™) as a size standard. Each fully resolved gel was photographed under ultraviolet radiation using a BioRad Gel Doc® system.

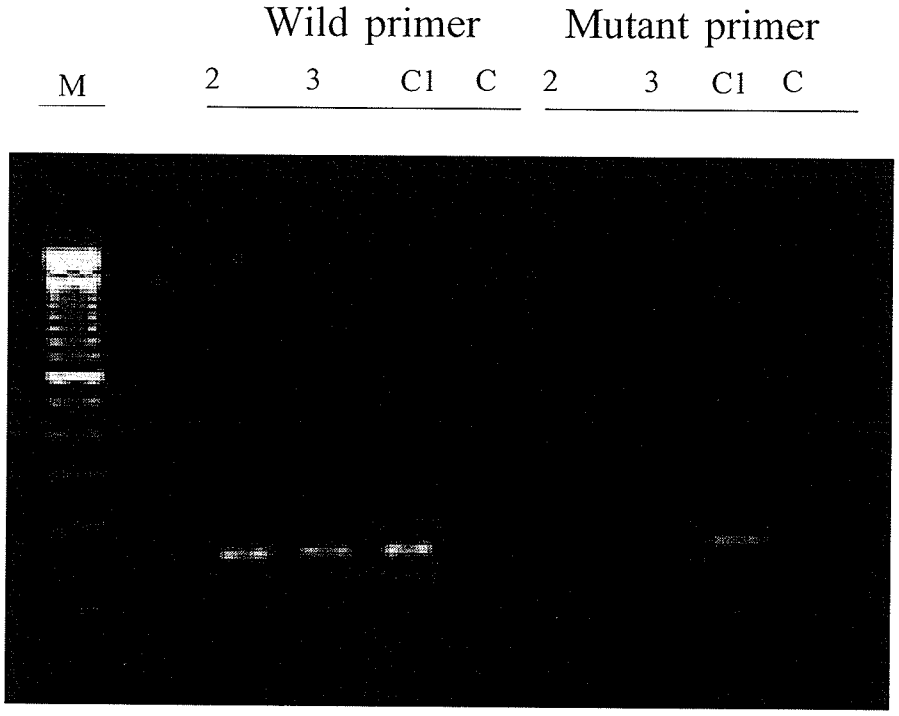


Figure 7: ARMS-PCR analysis agarose gel.

Figure 7: ARMS-PCR analysis of genetic single nucleotide polymorphism at position 868 of the nucleotide sequence of the CD4 gene. M indicates the molecular weight standard, lanes 2 and 3 are homozygous wild type samples, lane C1 is the heterozygous positive control and lane C is the negative water control.

CD4 Sequence Analysis Confirmatory Assay

Human genomic DNA samples were first PCR amplified in the region of the CD4 gene containing the nucleotide position 868. The PCR procedure was performed using a master mix of 5 ul of Invitrogen 10x Mg- PCR buffer, 1.5 ul of Invitrogen™ 50mM $MgCl_2$, 8.0 ul of dNTP, 28.25 ul of double distilled autoclaved water, 0.25ul of Invitrogen™ recombinant *Taq* DNA polymerase and 5.0 ul of a 5uM primer mix of primer 5N and 3C (listed in Table 2) per each reaction. Each reaction tube contained 48 ul of this mix and 2ul of the desired DNA template. An approximately 300 base pair region was amplified using the primers listed in Table 2. The PCR was performed on an MJ Research PTC-200 thermocycler using the same conditions as indicated for ARMS-PCR above.

The PCR products were then purified using Millipore Microcon® PCR Centrifugal Filter Devices. Each filter device was inserted into the collection tube provided in each kit. Four hundred and fifty ul of pH 8.0 TE buffer (see appendix 1 for solution protocols) was added to each filter device followed by 50 ul of PCR product. The apparatus was then centrifuged at 1000xg for 15 minutes in a table top microfuge (Fisher IEC Micromax). In order to recover the DNA, 20ul of deionized distilled, autoclaved water was added to each filter device and the device was placed upside down into a clean 1.5ml eppendorf collection tube. The whole apparatus was then centrifuged for 2 minutes at 1000g. These purified samples were then stored at –20°C.

Table 2: Sequencing primers. Sequences are listed from 5` to 3`. Locations are designated by nt = nucleotide.

Primer Name and Location	**Primer Sequence**	**Procedure**
5`outer nt 737-757	5` CAAAATAGACATCGTGGTGCT 3`	PCR amplification
5N nt 812-833	5` GTTCTCCTTCCCACTCGCCTTT 3`	forward sequence/PCR amplification
3N nt 1049-1071	5` TTCCTGTTTTCGCTTCAAGGGCC 3`	reverse sequence
3`outer nt 1072-1095	5`CCAGGTTCACTTCCTGATGCAAC 3`	PCR amplification

Cycle sequencing was then performed on these PCR products using Applied Biosystems PRISM® Big Dye™ Terminator Mix version 3.0 which contains fluorescently labeled dideoxynucleotide triphosphates, appropriate salt concentrations and high fidelity sequencing grade Ampli*Taq*® DNA polymerase. Each reaction tube contained 2 ul of Terminator mix, 1.5 ul of 3.2pmol/ul primer, 1.0 ul of double distilled autoclaved water and 1 ul of DNA template (the purified PCR amplified region of CD4). The primers used to generate both forward (5N) and reverse (3N) sequences are listed in Table 2. The cycle sequencing procedure was conducted on an MJ Research PTC-200 thermocycler under these conditions: 96°C for 3 minutes followed by 60 cycles of 96°C for 30 seconds, 50°C for 30 seconds and 60°C for 4 minutes.

The amplified samples were precipitated using a NaOAc - ethanol method. To each cycle sequenced reaction, 1 ul of 3M NaOAc pH 5.2 was added followed by 3x volume of 95% ethanol. The samples were left to precipitate overnight at room temperature. The samples were then centrifuged at maximum speed for 30 minutes. The pellet was then washed with 200 ul of 70% ethanol and centrifuged for 15 minutes. The pellet was then dried.

The pellets were then resuspended in 20 ul of high grade formamide. The samples were sequenced on an Applied Biosystems 3100 Automated Sequencer and analyzed using Applied Biosystems Sequence Analysis software. The samples were sequenced using a 20 second injection time and the BC-3100SRseqoffFTOFF analysis module.

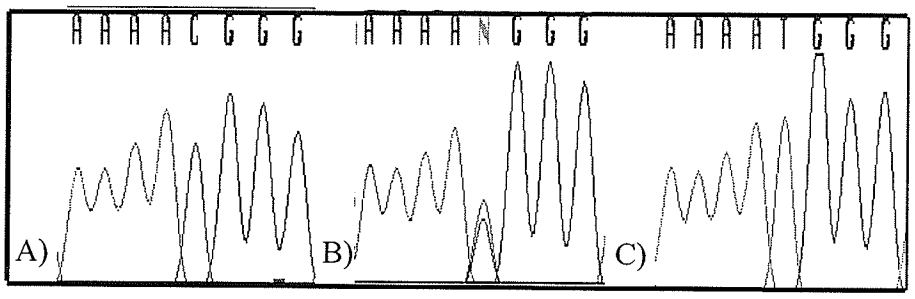
Heterozygotes were identified by the appearance of both a C and a T peak overlapping at position 868 (Figure 8). Some sequence analysis was also performed using Sequencher 4.1 software (Gene Codes Corporation, Ann Arbor, Michigan, USA).

Nested PCR for plasma samples

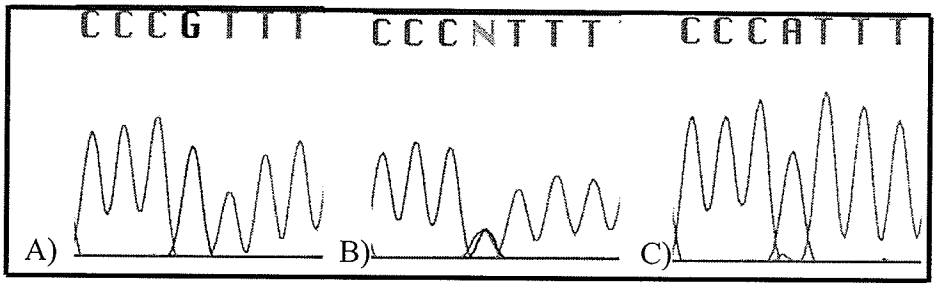
In the case of DNA samples isolated from plasma, a nested PCR strategy was employed. The first round of PCR was conducted using the outer primers listed in Table 2 and the procedure used was the same as above with the exception that 4 ul of DNA template was added to each reaction. This round was followed by column purification as described above. The second round employed the inside primers (called 5N and 3N) listed in Table 2. To prevent PCR cross contamination, this round was performed in a separate ultraviolet-decontaminated area and was conducted the same way as for round 1. These column-purified samples were then ready for cycle sequencing as described above.

Statistical Analysis

Chi square tests and Fisher's Exact Test were performed for comparison between categorical variables. Kaplan Meier survival plots were generated to assess the rate of CD4+ T-cell decline of seropositive individuals with and without the polymorphism and the length of time between enrollment of seroconverters and their seroconversion date.



Forward Primer 5N



868 **868** **868**

Reverse Primer 3N

Figure 8: Electropherograms of individuals a) homozygous for the wild type cytosine at nucleotide position 868 of the CD4 gene, b) heterozygous for 868T/C and c) homozygous for 868T.

Figure 8: Electropherograms of individuals a) homozygous for the wild type cytosine at nucleotide position 868 of the CD4 gene, b) heterozygous for 868T and c) homozygous for 868T. Images generated from Sequencher 4.1 sequence analysis software package (Gene Codes Corportation, Ann Arbor, Michigan, U.S.A.) after importing files from ABI 3100 automated sequencer (Applied Biosystems).

Log rank and Wilcoxon tests were applied. The Cox Proportional Hazard model was also used to confirm Kaplan Meier results by assessing the risk of seroconversion or CD4+ T-cell decline to end point and comparing that risk in each of the CD4 genotype groups. Logistic regression multivariate analysis (analysis of discrete rather than continuous variables) was also performed for comparisons between the various HIV serology groups in each population and their CD4 genotype by evaluating possible associations between factors (i.e., CD4 genotype, HIV status, HIV disease progression etc) and HIV transmission and progression.

Allele frequencies were determined using Python for Population Genetics (PyPop) software (University of California, Berkeley, USA) (76).

Results

ARMS-PCR Screening assay

To screen large numbers of DNA samples for the 868T polymorphism, an ARMS-PCR system was employed. Five hundred and thirty-three Kenyan women (285 ML and 248 MCH) were screened using this system. The results of this assay are summarized in Figure 9. There were no differences between the number of ML individuals with the polymorphism and the number of MCH individuals with the polymorphism. However, when the groups were split based on HIV status, there was a statistically significant result. Out of 161 HIV uninfected women, only 19 were positive for 868T, while 125 out of 372 HIV infected women were 868T positive (Chi square $p < 0.000001$, O.R. 3.78, CI95 2.18-6.63). When the group was split by HIV status and by risk group as is depicted in Figure 9, 63 of 157 MCH HIV infected women and 11 of 91 MCH HIV uninfected women were CD4 868T positive (Chi square $p < 0.0001$, O.R. 3.8, CI95 1.9-7.5). When the ML cohort members were examined, 62 of the 215 HIV infected women and 8 of the 70 HIV uninfected women had the polymorphism (Chi square $p = 0.019$ RR = 2.02 CI95 1.1-3.7). To confirm these results, select sequencing was performed on these samples. Some discrepancies were noted, prompting the sequencing of the entire sample in order to confirm. The ARMS-PCR assay was discordant with the sequence data in 27% of the cases. The entire work was then repeated using the sequence-based assay.

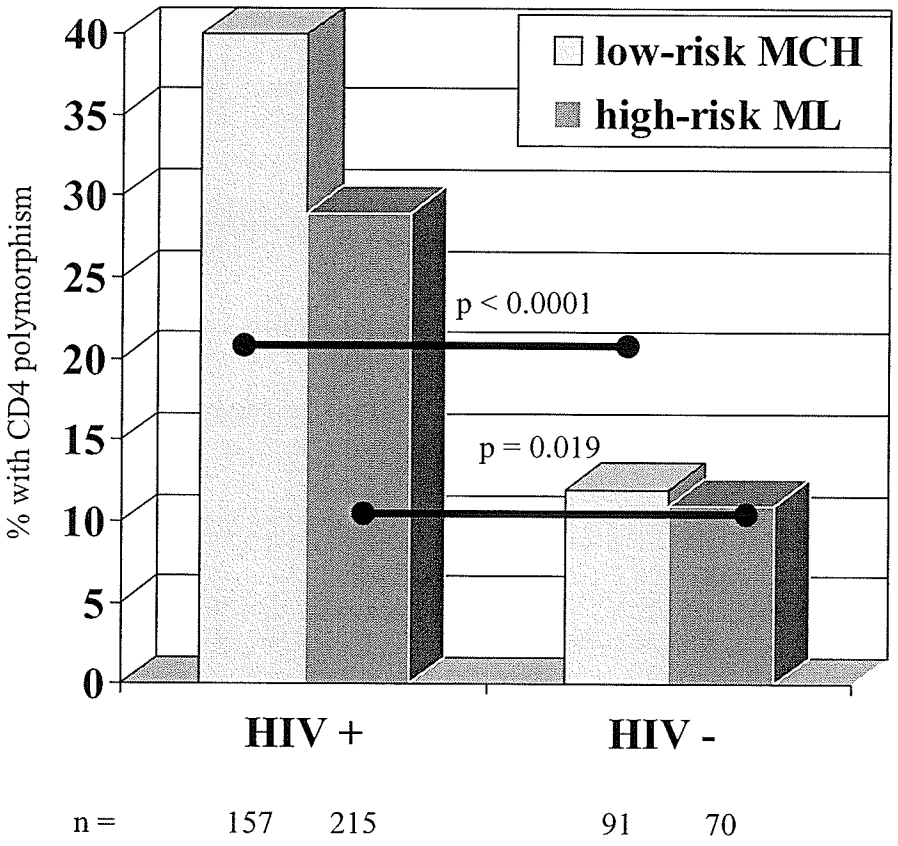


Figure 9: CD4 868T prevalence in low-risk (MCH) and high-risk (ML) individuals as screened by ARMS-PCR.

Figure 9: CD4 868T prevalence in low-risk (MCH) and high-risk (ML) individuals as screened by ARMS-PCR. The group was split by HIV status and by risk group. 63 of 157 MCH HIV infected women and 11 of 91 MCH HIV uninfected women were CD4 868T positive (Chi square $p < 0.0001$, O.R. 3.8, CI95 1.9-7.5). In the ML cohort 62 of the 215 HIV infected women and 8 of the 70 HIV uninfected women had the polymorphism (Chi square $p = 0.019$ RR = 2.02 CI95 1.1-3.7).

ML COHORT

CD4 868T not associated with ML HIV prevalence

In order to determine whether the CD4 868T polymorphism correlated with HIV prevalence in commercial sex workers, the percentage of individuals with the polymorphism in seronegative and seropositive groups was compared. The 364 individuals that were CD4-typed were separated into 3 groups based on their HIV status. The number of individuals within each group and their CD4 genotypes are shown in Table 3. In the HIV-1 negative group, 24% of individuals had at least one copy of the polymorphic CD4 allele. Twenty-nine percent of the HIV-1 infected group (individuals that were HIV-1 positive upon enrollment in the cohort) had the polymorphism. In the HIV-1 seroconverting group, 40% of the individuals possessed at least one copy of the polymorphic CD4 gene. If both the HIV-1 infected and seroconverting groups were combined, they showed that 31% of all individuals in these groups were CD4 868T positive. There were no significant differences in the prevalence of the polymorphism in the three categories studied (Chi-square $p = 0.1317$, OR = 1.380, CI95 = 0.917-2.078).

CD4 868T associated with ML HIV incidence

While the polymorphism was not associated with HIV prevalence, the higher CD4 868T frequency in the seroconverters suggested that a study of HIV incidence was warranted.

Table 3: ML HIV Disease Progression and Frequency of CD4 genotype and CD4 allele.

				HIV positive	
Genotype	HIV -	HIV+	HIVsc[a]	LTNP[b]	Rapid Progressors
CD4	(n = 87)	(n = 220)	(n = 57)	(n = 21)	(n = 19)
868C/868C	66	157	34	15	11
868C/868T	12	57	20	5	7
868T/868T	9	6	3	1	1

$?^2$ square test for trend comparing HIV- to LTNP to rapid progressors: p = 0.04 for progression of heterozygotes, logistic regression comparing LTNP and rapid progressors O.R. 2.936 (p = 0.0197, CI95 1.187-7.262)

				HIV positive		
Allele	HIV -	HIV+	HIVsc	LTNP	Rapid Progressors	Frequency
CD4	(2n = 174)	(2n = 440)	(2n = 114)	(2n = 42)	(2n = 38)	
868C	144	371	88	35	29	0.82
868T	30	69	26	7	9	0.18

[a] sc = seroconverter, [b] LTNP = long-term non-progressor

The next step undertaken was to determine if the SNP was associated with the incidence of HIV in this highly exposed ML cohort. A Kaplan Meier survival analysis (Figure 10) was performed on the seroconverting members of the cohort. The time between the date of enrollment and the calculated seroconversion date was plotted for each seroconverting individual. The comparison was made to uninfected individuals in the cohort and between genotypes. The analysis was performed on 125 HIV uninfected individuals. Eighty-nine of these individuals were homozygous for wild type allele, 25 were heterozygous and 11 were homozygous for 868T. Forty-eight individuals of the 125 HIV uninfected women seroconverted to HIV+ were then used for this analysis: 16 were heterozygous for CD4 868T, 29 were homozygous for the wild type allele and 3 were homozygous for 868T. Due to the small number of seroconverters homozygous for the 868T allele (3), only wild type and heterozygous individuals were included in the final analysis.

This analysis showed that heterozygous individuals seroconverted faster (shorter time between enrollment and seroconversion) than individuals homozygous for the wild type CD4 allele (Log-rank $p = 0.0046$, Wilcoxon $p = 0.0088$). The mean time to seroconversion for wild type individuals was 2889 days or 7.91 years as compared to heterozygotes who had a mean time to seroconversion of 1236 days or 3.38 years.

To confirm the Kaplan Meier result that the polymorphism is associated with an increased incidence of HIV infection, logistic regression analysis was performed on the

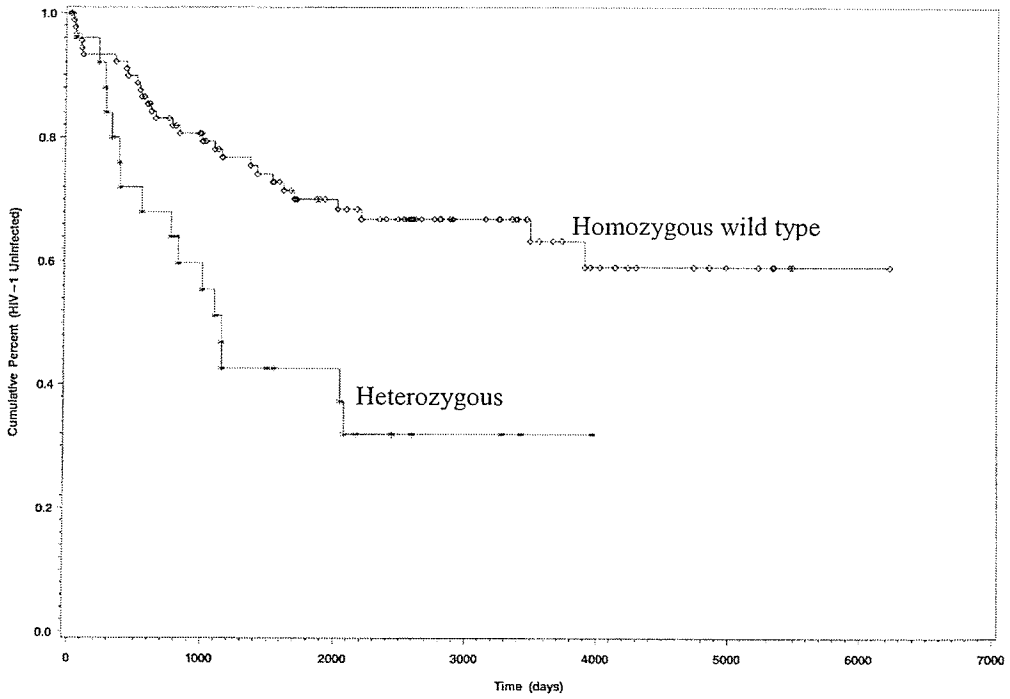


Figure 10: Kaplan Meier plot of time from enrollment to HIV-1 seroconversion i.e. HIV incidence.

Figure 10: Kaplan Meier plot of time from enrollment to HIV-1 seroconversion i.e. HIV incidence. The plot was constructed using 114 HIV uninfected individuals of which 45 seroconverted (of the 114 HIV negative individuals, 89 were homozygous for the wild type allele and 25 were heterozygous). Of these 45 seroconverters, 16 were heterozygous for CD4 868T (asterisks) and 29 were homozygous for the wild type allele (diamonds). Log rank test $p = 0.0046$ and Wilcoxon test $p = 0.0088$.

seroconverters. The same seroconverters who were part of the Kaplan Meier test in Figure 10 were analyzed and compared to uninfected controls. This showed that the heterozygous state is associated with seroconversion, supporting the Kaplan Meier findings (p = 0.0087, OR 2.488, CI95 1.260-4.913).

CD4 868T association with ML HIV disease progression

In addition to a potential effect on susceptibility to HIV infection, the CD4 868T may play a role in HIV disease progression. In order to determine if the polymorphism was associated with HIV disease progression, the cohort was divided into three progression groups based on the rate of CD4+ T-cell decline: uninfected individuals, long-term non-progressors and rapid progressors (who were a subset of the seroconverter group since their time of infection was known). The CD4 868T genotypes of these extreme ends of the HIV disease progression profile were compared to one another (Table 3). As summarized in Table 3, of the 21 long-term non-progressors, 5 were heterozygous for CD4 868T (23.81%) and of the 19 rapid progressors, 7 were heterozygous (36.84%), which is a significantly higher frequency than in the uninfected controls (heterozygote frequency was 13.8%) when compared in a chi-square test for trend. By logistic regression comparing rapid progressors to long-term non-progressors, heterozygotes were shown to be 2.936 times more likely to be rapid progressors (p = 0.0197, CI95 1.187-7.262) than wild type individuals.

CD4 868T association with lower CD4 counts in ML cohort

To determine the speed with which individuals with the polymorphism progressed to low CD4 counts as compared to wild type HIV-1 seropositive individuals, CD4+ T-cell decline was analyzed by Kaplan Meier analysis (Figure 11). Only individuals that either entered the cohort HIV positive or became infected during follow-up were included in this analysis. CD4 decline was measured as the time between the initial CD4+ T-cell count (if above 400) at the first clinic visit and the visit at which a count of below 400 cells/cubic mm was reached. A patient could only be included if a minimum of 2 visits to the clinic were recorded, one with a count above 400 and one with a count below 400. Of the 277 individuals typed in this study, 232 individuals fit the criteria. One hundred and forty-eight were homozygous for the wild type allele, 75 were heterozygous and 9 were homozygous for the polymorphic allele.

The 400 cells/cubic mm T-cell count was chosen because it represents a count that is definitively below 500 cells/cubic mm which is a clinical marker for the beginning of antiretroviral treatment.

Figure 11 shows a trend that the time between the first visit and the visit at which the CD4 count is below 400 cells/cubic mm is shorter for individuals homozygous for CD4 868T than for genotypically wild type individuals and heterozygous individuals (log-rank $p = 0.0910$, Wilcoxon $p = 0.1298$).

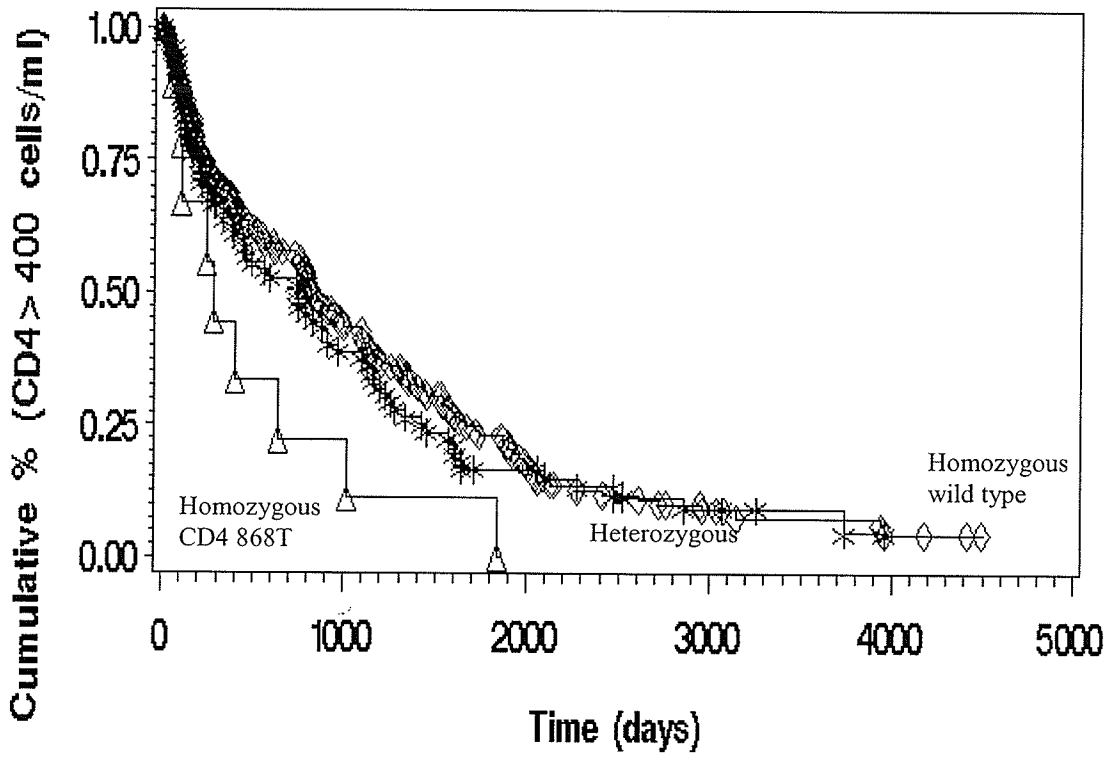


Figure 11: Kaplan Meier plot of ML CD4+ T-cell decline to less than 400 cells/cubic mm.

Figure 11: Kaplan Meier plot of ML CD4+ T-cell decline to less than 400 cells/cubic mm. Of the 196 seropositive individuals used for this analysis, 148 were homozygous for the wild type allele (diamonds), 75 were heterozygous (asterisks) and 9 were homozygous for the CD4 868T allele (triangles). Log rank test $p = 0.0910$ and Wilcoxon test $p = 0.1298$.

Confounding variables

In order to exclude the effects of possible confounding factors, dates of birth, tribal origin, and behavioural data such as time of enrollment into the cohort and start of prostitution were compared between CD4 genotype groups (Table 4). None of these possible confounding factors was shown to be associated with the effect seen by CD4 868T.

To determine if other genetic factors associated with susceptibility to HIV infection or disease progression were responsible for the CD4 868T effect, the ML CD4 genotype was compared to their human leukocyte antigen (HLA) alleles class I and class II that in published reports have been shown to be associated with accelerated or decreased disease progression (77). No significant associations were seen (data not shown).

MCH COHORT

CD4 868T not associated with MCH HIV prevalence

In order to determine if there was a correlation between the polymorphism and HIV prevalence among low-risk members of the MCH cohort, seronegative women were compared to women infected with HIV (Table 5). Of the 284 MCH women genotyped for CD4, 31.6% of HIV negative women were 868T positive, 29.3% of HIV infected women were 868T positive and 45.2% of seroconverting women were positive for 868T.

Table 4: Association of biographical data with CD4 868T genotype

Variable	CD4 Genotype	Mean	Standard Deviation	t-test
Date of birth	C/T, T/T	1960.88	6.425	
	C/C	1960.60	6.993	0.7265
Prostitution start date	C/T, T/T	1987.08	4.975	
	C/C	1986.40	5.526	0.2690
Cohort enrollment date	C/T, T/T	1991.04	3.072	
	C/C	1990.74	3.238	0.4183

	CD4 genotype		
	C/T, T/T	C/C	$?^2$
Tribal Group 1	50	106	
Tribal Group 2	24	74	
Tribal Groups 3-18	22	36	0.1891

Table 5: MCH HIV status and frequency of CD4 genotype and CD4 allele.

Genotype CD4	HIV - (n =107)	HIV+ (n =177)	HIVsc[a] (n =31)
868C/868C	52	125	17
868C/868T	19	45	12
868T/868T	5	7	2

Allele CD4	HIV - (2n =214)	HIV+ (2n =354)	HIVsc (2n =62)	Frequency
868C	123	295	46	0.82
868T	29	59	16	0.18

[a] sc = seroconverter

If all the HIV infected women are combined (those enrolled HIV+ and those who seroconverted during follow-up) 33.3% had at least one copy of the polymorphism. There were no significant differences between the seronegative and seropositive groups with respect to the number of 868T positive individuals (chi square p= 0.9806).

CD4 868T associated with MCH rapid CD4+ T-cell decline to <400 cells/cubic mm

In order to determine if the CD4 868T polymorphism had an effect on the CD4+ T-cell decline to less than 400 cells/cubic mm of blood, a Kaplan Meier analysis was performed on all cohort members who entered the cohort positive or seroconverted during follow-up (Figure 12). The individuals were followed from the time of enrollment (or seroconversion) until the first CD4+ T-cell count of less than 400 cells/ cubic mm. In order to be included in the analysis, at least one count above 400 and one count below or at 400 must have been recorded.

Of the 208 possible individuals that could have been used in this analysis, 159 had appropriate data to be used in the Kaplan Meier plot. Of the 159 included individuals, 113 were homozygous for the wild type allele, 40 were heterozygous and 6 were homozygous for 868T.

In this analysis, the individuals homozygous for 868T had a much steeper CD4+ T-cell decline than both heterozygous and homozygous wild type individuals (log rank p =

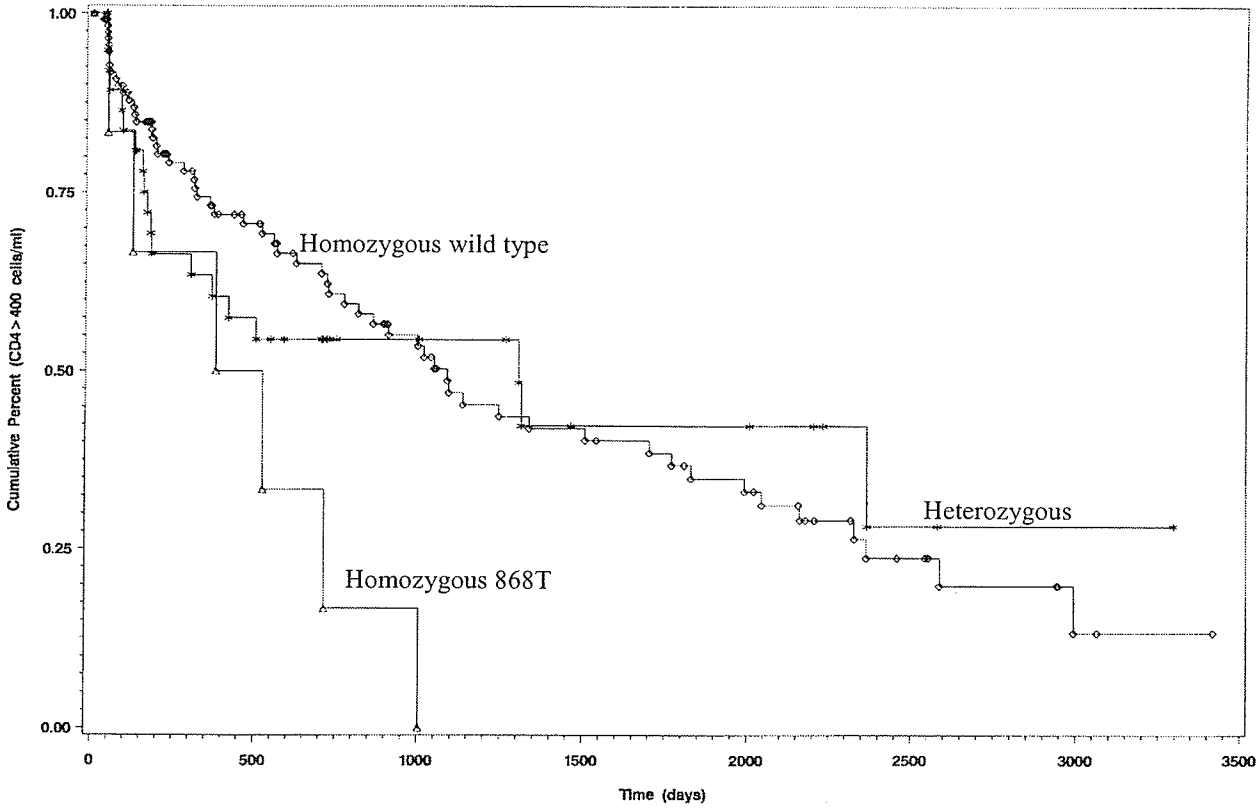


Figure 12: Kaplan Meier plot of MCH CD4+ T-cell decline to less than 400 cells/cubic mm.

Figure 12: Kaplan Meier plot of MCH CD4+ T-cell decline to less than 400 cells/cubic mm. Of the 159 seropositive individuals used for this analysis, 113 were homozygous for the wild type allele (diamonds), 40 were heterozygous (asterisks) and 6 were homozygous for 868T (triangles). Log rank test $p = 0.04$.

0.04). Patients homozygous for 868T dropped to a CD4 count below 400 in a mean of 467.33 days (1.28 years) while homozygous wild type women dropped to below 400 in 1371.88 days (3.76 years) and heterozygotes fell to below 400 in 1246.22 days (3.41 years). When Cox proportional hazard test was performed to confirm this data, 868T homozygous women were shown to have a 2.8 times faster decline to <400 cells/cubic mm ($p = 0.04$ CI95 1.220-6.599).

CD4 868T associated with rapid MCH CD4+ T-cell decline to <200 cells/cubic mm

As indicated in the introduction to this thesis, Walter Reed Stage 6 disease is considered clinical AIDS. To be considered stage 6, patients must have CD4+ T-cell counts below 200 cells/cubic mm. To determine whether or not 868T positive individuals have a faster CD4+ T-cell decline to AIDS level counts, a Kaplan Meier and Cox proportional hazard test were performed on HIV infected and HIV seroconverting individuals in the cohort providing they had at least one count above 200 and one below (Figure 13). Each patient was followed in the analysis from the date of enrollment (or seroconversion) to the first count below 200 cells/cubic mm. Of the possible 208 HIV infected tested members of the cohort, 159 patients had sufficient CD4 count information to be included in the Kaplan Meier analysis. One hundred and thirteen were homozygous for the wild type allele, 40 were heterozygous and 6 were homozygous for 868T.

The Kaplan Meier plot shows a faster CD4+ T-cell decline for 868T homozygotes than for heterozygotes and homozygous wild type individuals (log rank $p = 0.0295$). The Cox

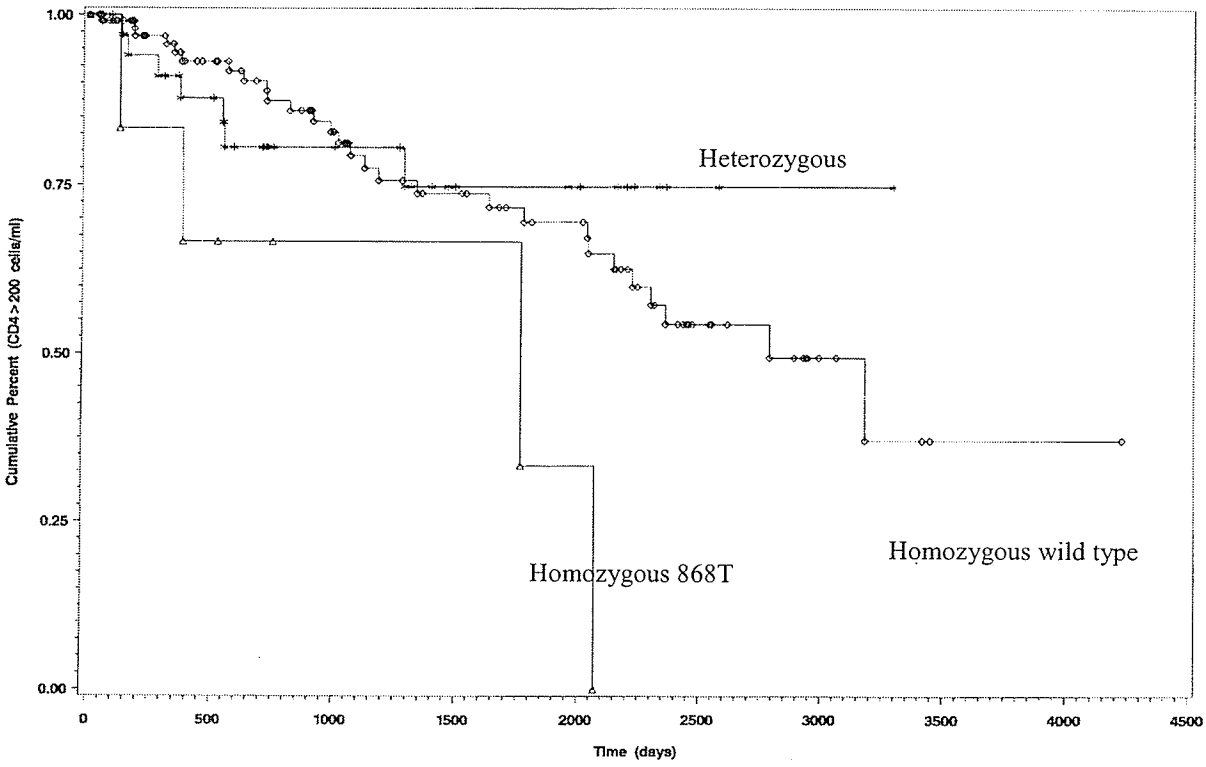


Figure 13: Kaplan Meier plot of MCH CD4 T-cell decline to less than 200 cells/cubic mm.

Figure 13: Kaplan Meier plot of MCH CD4+ T-cell decline to less than 200 cells/cubic mm.

Figure 13: Kaplan Meier plot of MCH CD4+ T-cell decline to less than 200 cells/cubic mm. Of the 159 seropositive individuals used for this analysis, 113 were homozygous for the wild type allele (diamonds), 40 were heterozygous (asterisks) and 6 were homozygous for 868T (triangles). Log rank test $p = 0.0295$.

proportional hazard test showed that 868T homozygous women had a 3.5 times faster CD4+ T-cell decline to AIDS level counts (p =0.0295 CI95 1.230-10.308).

MOTHER TO CHILD TRANSMISSION IN MCH COHORT

CD4 868T not associated with HIV prevalence in Kenyan children

The CD4 868T status of children of HIV infected MCH mothers was compared to their HIV status in order to determine if the polymorphism affects the prevalence of HIV in this population (Table 6). A total of 278 children from 178 HIV infected MCH mothers were tested for the polymorphism. Of these, 19 had indeterminate HIV status and were eliminated from further analysis. The remaining 257 children were from 160 HIV+ mothers.

There were no statistically significant differences in 868T distribution among different HIV infection groups of children. Of the included 257 children, 73 of the 167 (43.71%) HIV uninfected children had the 868T polymorphism while 31 of the 90 (34.40%) HIV infected children were 868T positive (Chi square p = 0.1487).

Table 6: MCH children HIV status and frequency of CD4 genotype and CD4 allele.

Genotype CD4	HIV - (n =167)	HIV+ (n =71)	HIVbmt[a] (n =19)
868C/868C	94	47	12
868C/868T	60	19	4
868T/868T	13	5	3

Allele CD4	HIV - (2n =334)	HIV+ (2n =142)	HIVbmt (2n =38)	Frequency
868C	248	113	28	0.76
868T	86	29	10	0.24

[a] Bmt refers to late seroconversion due to breast milk transmission.

CD4 868T not associated with MCH children's CD4+ T-cell decline to less than 25%

Disease progression in children is more rapid than in adults. To determine whether or not the 868T polymorphism had an effect on the speed of CD4+ T-cell decline in children, a Kaplan Meier analysis was performed (Figure 14) on HIV infected children. In this analysis, the children were followed from the point of entrance into the cohort (birth) until the end point. The end point was determined to be the first CD4+ T-cell count that amounted to less than 25% of total lymphocytes. This end point is roughly equivalent to a CD4+ T-cell count of approximately 400-500 cells/cubic mm in adults.

Of the possible 90 HIV infected children, 64 children met the criterion of having a minimum of two CD4+ T-cell counts with one of these above and one below the 25% of total lymphocyte end point. Of the 64 seropositive individuals used for this analysis, 40 were homozygous for the wild type allele, 21 were heterozygous and 3 were homozygous for 868T.

The Kaplan Meier plot shows that the CD4+ T-cell decline is similar between the wild type homozygotes, heterozygotes and 868T homozygotes. The log rank test value was $p = 0.42$.

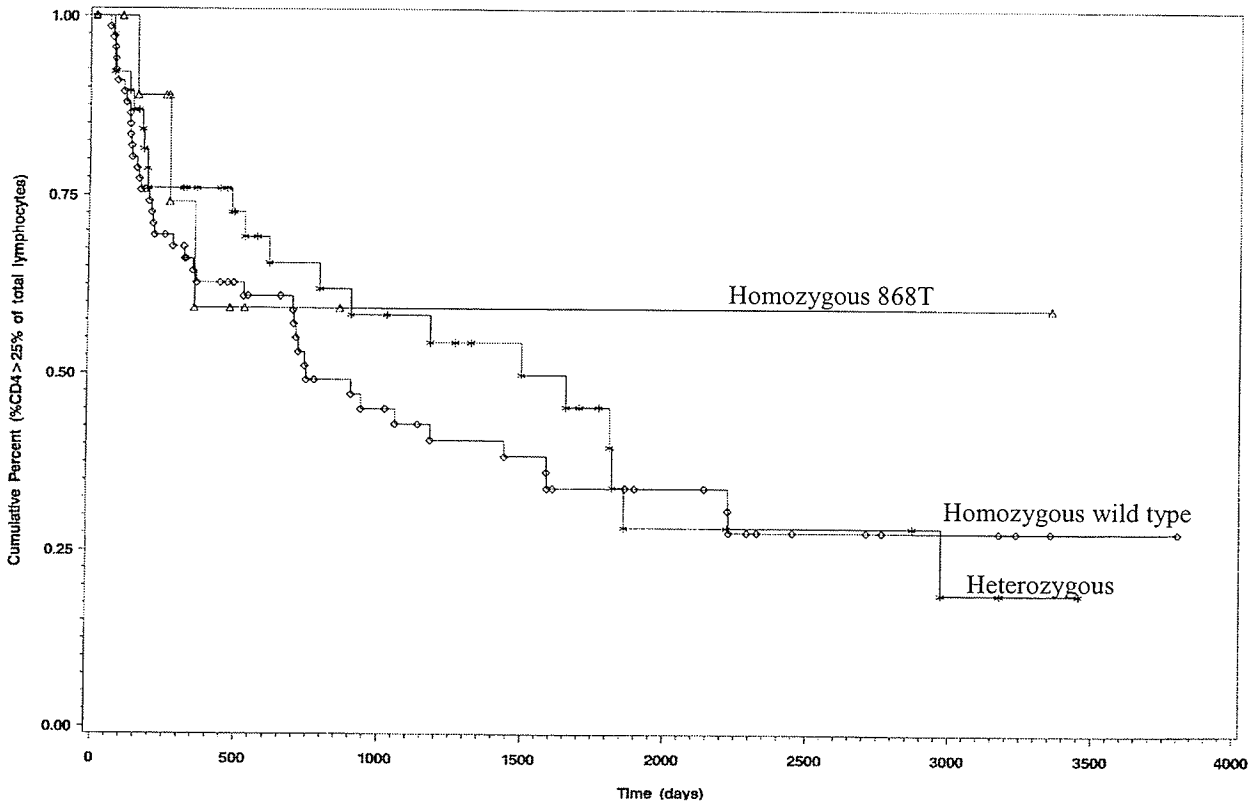


Figure 14: Kaplan Meier plot of MCH children's CD4+ T-cell decline to less than 25% CD4+ cells from total lymphocytes.

Figure 14: Kaplan Meier plot of MCH children's CD4+ T-cell decline to less than 25% CD4+ cells from total lymphocytes. Of the 64 seropositive individuals used for this analysis, 40 were homozygous for the wild type allele (diamonds), 21 were heterozygous (asterisks) and 3 were homozygous for 868T (triangles). Log rank test $p = 0.42$.

CD4 868T not associated with MCH children's CD4+ T-cell decline to less than 20%

The CD4+ T-cell count based clinical point that determines AIDS stage in children is when CD4+ T-cell counts reach 20% of total lymphocytes. In order to determine if the 868T polymorphism affected the time between enrollment (in this case birth) or seroconversion and AIDS stage illness, a Kaplan Meier analysis was performed on the HIV infected children of the cohort who developed HIV infection either *in utero* or through breast milk transmission (Figure 15).

To be included in this analysis, a minimum of two CD4+ T-cell counts including a CD4+ T-cell count above and below 20% of total lymphocytes had to have been recorded. Of the possible 90 HIV infected children tested, 34 children had enough clinical data to be included. Of the 34 seropositive individuals used for this analysis, 21 were homozygous for the wild type allele, 11 were heterozygous and 3 were homozygous for 868T.

As seen in Figure 15, the CD4+ T-cell decline in children to AIDS stage disease is similar regardless of 868T genotype. The log rank test value for this analysis is $p = 0.57$.

Mother's CD4 868T genotype does not affect *in utero* HIV transmission

Of the members of the MCH mothers and the MCH children's cohorts, 170 are matched mothers and their children. This pool was used to determine whether or not the 868T

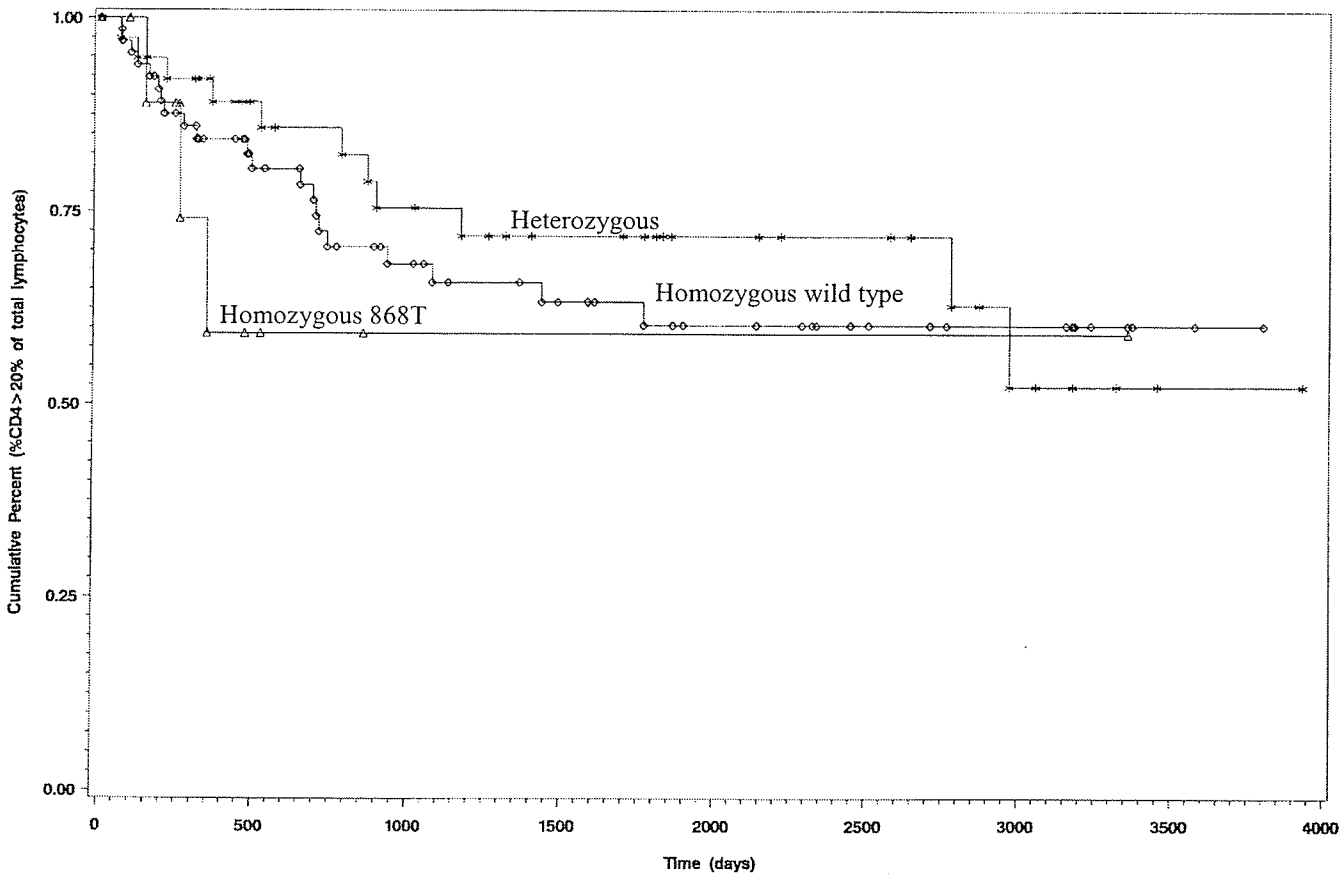


Figure 15: Kaplan Meier plot of MCH children's CD4+ T-cell decline to less than 20% CD4+ cells from total lymphocytes.

Figure 15: Kaplan Meier plot of MCH children's CD4+ T-cell decline to less than 20% CD4+ cells from total lymphocytes. Of the 34 seropositive individuals used for this analysis, 21 were homozygous for the wild type allele (diamonds), 11 were heterozygous (asterisks) and 3 were homozygous for 868T (triangles). Log rank test $p = 0.57$.

polymorphism affects the transmission of the virus to children. Every mode of transmission is different so *in utero* transmission was first analyzed (Table 7).

Chi square and logistic regression analysis was applied to the data. Of the 99 homozygous wild type women, 21 gave birth to HIV infected children, while 16 of 64 heterozygous women gave birth to HIV infected children and 3 out of 7 868T homozygous women transmitted the virus to their children. There were no statistical differences between these groups.

Mother's CD4 868T genotype associated with increased breast milk HIV transmission

Using the same pool of 170 matched mothers and their children, the effect of CD4 868T on breast milk transmission was examined (Table 8, Figure 16). Different modes of mother to child transmission involve different factors so breast milk transmission was analyzed separately.

As seen in Table 8 and Figure 16, out of the 99 homozygous wild type women, 3 had children seroconvert late due to breast milk transmission. Of the 64 heterozygous women, 8 transmitted the virus to their children by breast-feeding while 3 of the 7 homozygous 868T women had children infected with HIV by breast milk.

Table 7: *In utero* mother to child transmission of HIV based on mother's CD4 genotype.

Maternal CD4 genotype	Kids HIV status		$?^2$	Odds Ratio	CI95[b]
	HIV -	HIV+[a]			
(n = 170)	(n = 128)	(n = 40)	(p=)		
868C/868C (n = 99)	78	21			
868C/868T (n = 64)	46	16	0.5007	1.292	0.613-2.723
868C/868C (n = 7)	4	3	0.2016	2.786	0.578-13.425

[a] HIV + refers to children infected *in utero* or during delivery, [b] CI95 is the 95% confidence interval.

Table 8: Breast milk mother to child transmission of HIV based on mother's CD4 genotype.

Maternal CD4 genotype	Kids HIV status		?[2]	Odds Ratio	CI95[b]
	HIV -	HIV bmt[a]			
(n = 170)	(n = 128)	(n = 14)	(p=)		
868C/868C (n = 99)	78	3			
868C/868T (n = 64)	46	8	0.0316*	4.522	1.142-17.902
868C/868C (n = 7)	4	3	0.0021*	19.500	2.947-129.026

* statistically significant $p<0.05$

[a] HIV bmt refers to late seroconversion by breast milk transmission, [b] CI95 is the 95% confidence interval.

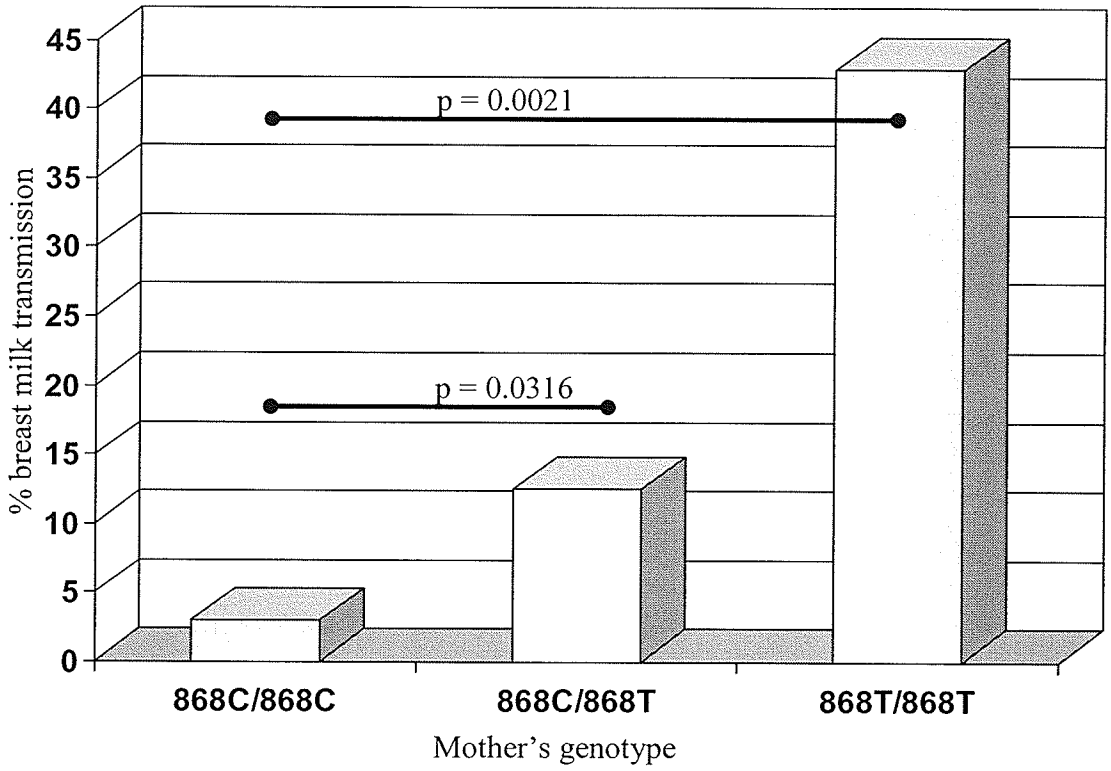


Figure 16: Breast milk mother to child transmission of HIV based on mother's CD4 genotype.

Figure 16: Breast milk mother to child transmission of HIV based on mother's CD4 genotype. Of the 99 wild type mothers, 3 transmitted HIV to their children through breast milk (3%). Of the 64 heterozygous mothers, 8 transmitted the virus to their children by breast milk (12.5%) a significantly higher amount than wild type women ($p = 0.0316$ O.R. 4.522 CI95 1.142-17.902). The 7 women homozygous for 868T transmitted HIV to their children by breast milk 3 times (42.857%), also a significantly higher rate than wild type women ($p = 0.0021$ O.R. 19.5 CI95 2.947-129.026).

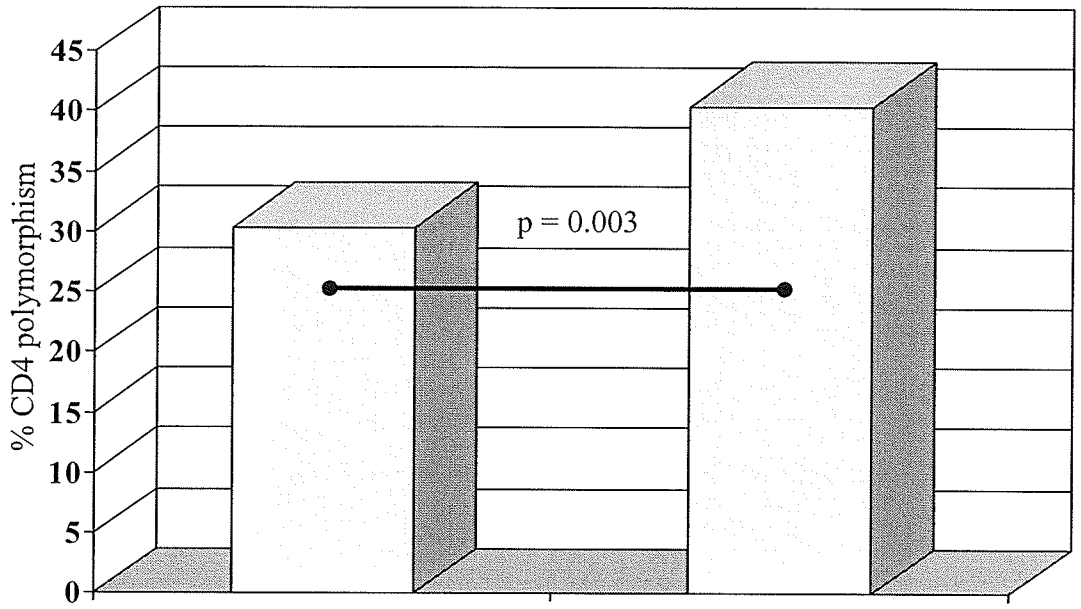
Both Chi Square and logistic regression analysis were performed comparing the amount of breast milk transmission in heterozygous and 868T homozygous women to wild type women. A statistically significant number of heterozygous women transmitted HIV to their children by breastfeeding ($p = 0.0316$, O.R. 4.522, CI95 1.142-17.902). In addition, homozygous 868T women were 19.5 times more likely to transmit HIV to their children by breastfeeding ($p = 0.0021$, O.R. 19.500, CI95 2.947-129.026) than wild type women.

Prevalence of 868T is higher among Kenyan children than among Kenyan adults

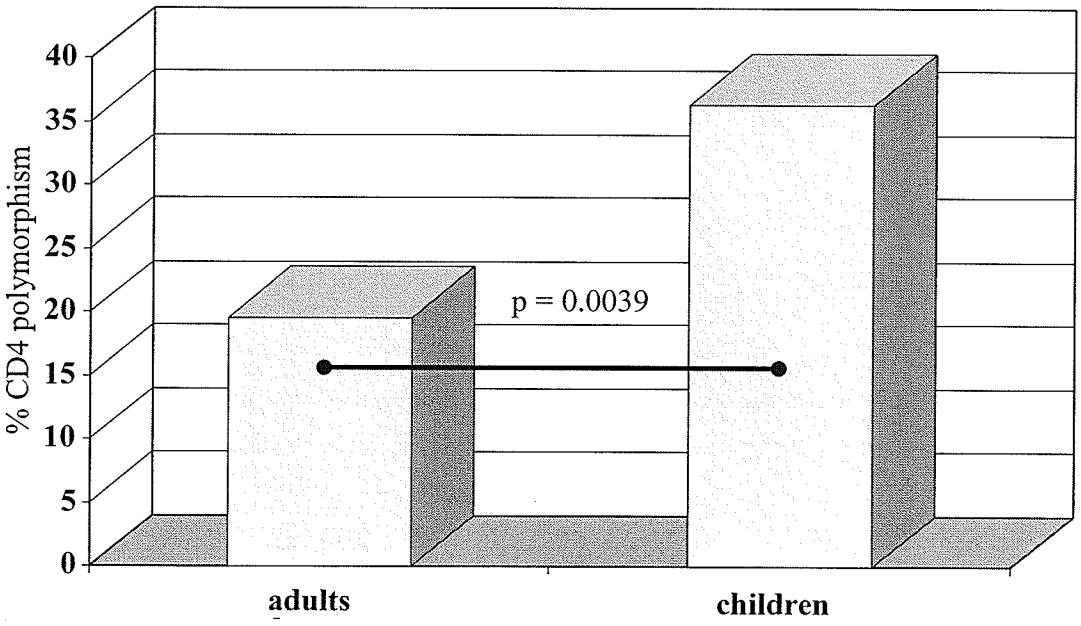
When the prevalence of the 868T polymorphism was analyzed separately in the adult and children's group, a difference was noted. In order to determine if the prevalence was statistically different between the different age groups, a chi square analysis was done on all the adults versus all the children regardless of HIV status, then the age groups were also compared based on HIV infection.

In Figure 17A, all adults (ML and MCH) were compared to all children tested for CD4 868T in this study. Of the 648 adults, 30.40% (197 individuals) had the polymorphism while 40.47% (104 individuals) of the 257 of children were 868T positive. In this analysis, there are significantly more children possessing at least one copy of 868T when compared to the adults ($p = 0.003$).

To control for the confounding effect HIV status may have on CD4 genotype only HIV uninfected individuals were compared. As depicted in Figure 17B, of the 158 HIV uninfected adults, 19.60% (31 individuals) and 36.36% (61 individuals) of the 167 children had the polymorphism. When only HIV- adults and children were compared in Figure 17B, the results remained significant ($p = 0.0039$). All the HIV infected adults were also compared to the HIV infected children (results not shown) and the distribution of CD4 868T was also determined to be statistically different ($p = 0.04$).



A)



B)

Figure 17: Age dependent gene frequency of CD4 868T. A) All children and adults tested in this study. B) HIV uninfected individuals only.

Figure 17: Age dependent gene frequency of CD4 868T. A) All children and adults tested in this study. Of the 648 adults, 30.40% (197 individuals) had the polymorphism while 40.47% (104 individuals) of the 257 of children were 868T positive (Chi square p = 0.003). B) Total HIV uninfected children and adults tested in this study. Of the 158 adults, 19.60% (31 individuals) and of the 167 children, 36.36% (61 individuals) had the polymorphism (Chi square p = 0.0039).

Discussion

From two Nairobi-based cohorts we identified three groups of individuals based on their HIV status and the stage of their disease, if infected. The ML sex workers cohort was divided into HIV uninfected women, HIV+ women who entered the cohort with a positive test for HIV infection and seroconverting women who were HIV negative upon enrollment in the cohort but became infected over the course of follow-up. The same groups were also identified in the mothers of the MCH cohort. The HIV infected women of the ML cohort were further subdivided into normal progressors and long-term non-progressors (LTNP). These LTNP were women who were HIV infected but remained at CD4+ T-cell counts above 400 cells/cubic mm for 8 years or more. LTNP were identified in both the initially infected and seroconverting women. Among the ML seroconverters, a subset of rapid progressors was identified. These were women who progressed to AIDS level CD4+ T-cell counts (<200 cells/cubic mm) within 3 years of seroconversion. The LTNP and RP definitions were developed based on the more rapid disease progression seen in sex workers as compared to other African women (78).

The children of the MCH cohort were divided in a very similar manner to the adult cohorts with HIV uninfected and HIV infected groups. Some children identified as HIV+ were determined to have been infected *in utero*. Those that would in an adult cohort be called seroconverters, since they were initially HIV PCR negative prior to becoming positive, are called bmt. These are children determined to have been infected through breast milk transmission explaining the observed late seroconversion.

Within these groups, a great variation in disease progression to AIDS exists. Many valid correlates of progression and mother to child transmission of HIV have been brought forward and this study seeks to add another factor to the list of contributors. The question posed by this work is whether or not a single nucleotide polymorphism of the gene for CD4 exerts an effect on the susceptibility to HIV infection in adults, on the disease progression of HIV infected adults and on the perinatal and postnatal mother to child transmission of HIV. **This question led to the development of the hypothesis that the 868T polymorphism in the coding region of CD4 is a genetic risk factor for seroconversion, rapid disease progression of infected adults and mother to child transmission.**

The finding that a polymorphism in an HIV-1 receptor is involved in altering susceptibility and the course of the disease is not surprising and has been shown for a number of coreceptor polymorphisms (41-44, 46, 47, 79). There are also examples of polymorphisms affecting disease progression of various infectious diseases including the presence of at least one copy of a particular MHC class II allele, HLA-DQB1*0503 which has been associated with clinical cases of *Mycobacterium tuberculosis*-caused tuberculosis. In a Cambodian study (39) this SNP, which causes an amino acid change, correlates with progression to clinically symptomatic tuberculosis. This work suggested that the SNP didn't affect the susceptibility to infection by the bacillus but it did cause an altered disease progression. Studies of NRAMP1 polymorphisms have indicated no effect on susceptibility but a marked effect on progression of *Mycobacterium leprae* infection (40). An accelerated disease progression to AIDS has also been shown with a

structural variant of the CX_3CR1 chemokine receptor (55-57). Perhaps the best described effect of a genetic polymorphism on HIV disease progression as well as susceptibility to infection is CCR5Δ32 (41-44, 46, 47, 79). However, our study is the first to demonstrate that a polymorphism in the primary HIV-1 receptor, CD4, is associated with a higher HIV incidence and accelerated disease progression in infected individuals.

This study did not find an association between the CD4 868T polymorphism and HIV prevalence among the highly exposed ML cohort sex workers (Chi-square p = 0.1317, OR = 1.380, CI95 = 0.917-2.078). Yet, we did observe that the highest prevalence of the polymorphism was among those individuals who were infected with HIV while enrolled in the cohort. Since prevalence is defined as the proportion of persons with a particular disease within a given population and incidence is the number of new cases during a specified time period in a given population, should not the presence of CD4 868T correlate with both HIV incidence and prevalence? If a disease were non-fatal, a steady incidence rate would result in an ever-increasing prevalence. However, if a particular genetic condition has a high probability of death, a prevalence study may underestimate the frequency of that state since a higher proportion of those with the condition would not survive and those with the less lethal condition would be over-represented. In our study those ML with the CD4 868T polymorphism are 2.9 times (p = 0.0197, CI95 1.187-7.262) more likely to be rapid progressors (Table 3) and have a trend (log-rank p = 0.0910, Wilcoxon p = 0.1298) towards a more rapid CD4+ T-cell decline to < 400 cells/cubic mm (Figure 11), so it is possible that they are at a greater risk of dying sooner

and, therefore, would be under-represented in a prevalence study. Since an incidence study measures the earliest indication of disease, which in our case is HIV seroconversion, incidence would more accurately reflect the association of the CD4 868T polymorphism with the condition (HIV disease in this case). Rapid progressing patients are better represented in an incidence study because their initial seroconversion is recorded which is an equivalent measure to every other patient tested. The incidence of HIV infection, as indicated by the Kaplan Meier time to seroconversion graph (Figure 10), would support the role of the polymorphism increasing susceptibility to HIV infection (Log-rank p = 0.0046, Wilcoxon p = 0.0088). However, given that these commercial sex workers are highly exposed to HIV, the protective effect of having the wild type allele may be overcome by the large number of exposures to HIV, eventually resulting in infection. This would eventually result in a similar prevalence of the polymorphism among infected and uninfected groups.

This work did not find an association of the CD4 868T polymorphism with HIV prevalence in either MCH women (chi square p= 0.9806) or their children (Chi square p = 0.1487). However, among the MCH mothers, seroconverters still had a higher frequency of the polymorphism as was seen with the ML sex workers (Table 5). In this case, as in the ML cohort, individuals with the polymorphism might be under-represented in a prevalence study if this genetic condition conferred increased mortality. The CD4 868T positive women were in fact shown to have a steeper CD4+ T-cell decline to both treatment level (<400 cells / cubic mm, log rank p = 0.04) and to AIDS level (<200 cells / cubic mm, log rank p = 0.0295) counts than wild type women (Figures 12 and 13). The

curves in these figures as compared to the Kaplan Meier plot for ML CD4+ T-cell decline (Figure 11) show a similar decline although the difference in the ML cohort is not statistically significant and can only be interpreted as a trend.

The children of the HIV infected members of the MCH cohort did not show an association of CD4 868T with prevalence of HIV (Table 6). Unlike the adult groups, children that seroconverted due to breast milk transmission (the child cohort equivalent of a seroconverter since a seroconversion date can be calculated) did not show a higher frequency of CD4 868T compared to HIV infected and HIV uninfected children. This observation was supported by the Kaplan Meier plots of CD4+ T-cell decline (Figures 14 and 15). Since each mode of transmission is so drastically different from one another in terms of entry, viral dominance (X4 vs. R5) etc., it is not entirely surprising to find that the disease progression of children does not seem to be affected by the polymorphism. The disease progression of children is very rapid (23) when compared with that of adults therefore a genetic factor might not be well-represented in a survival plot.

Mother to child transmission of the HIV virus can occur by at least 3 routes. Transmission during gestation or *in utero* indicates cross-placental transfer of HIV virions that infect the fetus. Infection can also occur during delivery due to a transfer of infected blood and fluids to the neonate. For the purposes of this study, both gestational and delivery HIV infections were deemed *in utero* transmission as it is extremely difficult to assess the infection time when a neonate tests positive for the virus. Late seroconversion or breast milk transmission is the third route of infection and occurs due

to infected breast milk being ingested and virions entering the child through the gastrointestinal system. The viral load of breast milk depends directly and primarily on the serum viral load of the mother (80). This is the only route of mother to child virus entry that depends on ingestion of the virus instead of direct blood-to-blood contact. Since the survival curves and prevalence tests seemed to indicate that the polymorphism in children did not affect their susceptibility to infection or their disease progression, the effect of the CD4 genotype of the HIV infected mother on the HIV status of her child and the means of transmission of the virus to infected children were examined. Looking at the information in Tables 7 and 8 and Figure 16, the genotype of the mother was not associated with an increased infection of her child when all HIV infected children were combined or when only children deemed to have been infected *in utero* were analyzed. However, when only children infected via breast milk transmission (Table 8 and Figure 16) were analyzed, 868T heterozygous mothers were 4.5 times more likely to transmit to their infants and 868T homozygous mothers were 19.5 times more likely to transmit to their infants than wild type mothers. This is a potentially very important finding. This is another indication of a dose effect – the probability of infection increases with the number of 868T alleles a mother possesses. It is also interesting to note that because this is a common SNP among Africans but is rare in Caucasians, it might be a contributor to the observations that African-American women transmit HIV more commonly to their infants than Caucasian women (81) and that breast milk transmission is estimated to be higher in African cohorts than in other study groups (80).

Serum HIV viral load is directly associated with the viral load in breast milk. Cunningham *et al* (81) showed that race, especially being of African descent, was independently associated with a higher viral load at delivery. If CD4 868T increases the serum viral load of a mother thereby increasing her breast milk viral load, an increase in the rate of breast milk transmission would be observed as it was in this study. Unfortunately, in this study, the samples were not well suited for viral load determination so these types of analyses were not performed.

A particularly interesting finding of this study was the observation that the population of children studied had a higher frequency of the polymorphism than adults (0.24 in children as opposed to 0.18 in adults) regardless of HIV status (Figures 17A and 17B). The HIV/AIDS epidemic has raged in sub-Saharan Africa for 20 years, which is a very short time by evolutionary standards. It is therefore unlikely that the virus itself has altered the genetic makeup of Kenyans to this extent. It is, however, worth indicating that if the altered CD4 868T causes a change to the immune system in some way, the children possessing the polymorphism might be susceptible to various other infections including, but not restricted to, HIV. Although it may seem counterproductive, this allele seems to be in equilibrium in this population. One possible explanation could be attributed to the concept of Darwinian Fitness. Darwinian Fitness assumes that if an allele can be carried in an individual to reproductive maturity, it will not completely leave a population (82). Therefore, if CD4 868T can be carried in members of this Kenyan population to adulthood (for a variety of reasons including counter-acting HIV-1 resistance factors) it will be continuously maintained in the population. There are not enough generations in

this study to properly assess whether or not HIV-directed selection pressure is at work on this allele. There is also a possibility that this allele has an unknown positive role as does sickle-cell anemia heterozygosity against malaria (83).

This study involved 3 populations derived from two cohorts. The first population involved 364 women of a Kenyan commercial sex workers cohort. Follow-up in this cohort continues from enrollment until the woman either dies or leaves the area and the patients in the study have been followed for up to 15 years. Roughly 20% of cohort participants become lost to follow-up. No selection bias was observed among our subjects that were lost to follow-up or died. From the second low-risk mother and child health cohort, 285 mothers and 257 children were examined. New enrollment in this cohort ended in 2000 but mothers and children continue to be followed, some for up to 10 years. Our study size is large enough to reveal effects generated by this particular genetic polymorphism. Cohorts of smaller and similar sizes have been used to reveal prognostic values of many polymorphisms with similar allele frequencies to CD4 868T including CCR5Δ32, CX3CR1 and MDR-1 variants, IL-4 –589T and SDF1-3'A (44, 47, 50, 57). The allele frequencies in some of these populations for example were 0.044-0.14 for CCR5Δ32 studies and 0.21 for SDF1-3'A studies (84) while CD4 868T has an allele frequency of 0.18 in our adult populations and 0.24 in our population of children.

HIV infection and disease progression are dependent on a combination of many factors, none of which alone explain all cases. There are some rapid progressors identified in this study that do not, in fact, have the CD4 868T polymorphism. These individuals may

have other clinical factors or genetic polymorphisms that have affected their clinical outcomes.

The mechanism of how the polymorphism is associated with HIV susceptibility and disease progression is not known. There are potential epidemiological explanations. It is possible that those with and without the polymorphism differed in their exposure to HIV or in their genetic background. Since the risk of HIV acquisition, made up of several factors including condom use and HIV prevalence among sex workers clients, varied as a function of time over the 20 years of the ML cohort, we sought to determine if those with the polymorphism differed in time dependent variables. As indicated in Table 4, there was no difference in their age, duration of prostitution or enrollment into the cohort, suggesting similar levels of exposure. Table 4 also indicates the CD4 polymorphism is not associated with any particular tribal group. This suggests there the CD4 polymorphism is not simply a marker of tribal genetic variation. The lack of association with tribal origin also discounts any tribal differences in sex practices or HIV exposure that could account for the association with HIV incidence and disease progressions. Although our examination is by no means exhaustive, to date we have found no demographic or behavioural link with the CD4 868T polymorphism.

To explain the association between CD4 868T and HIV incidence and disease progression, there are also several potential molecular mechanisms worth investigating. The amino acid change in CD4 that results from 868T is suspected to alter the tertiary structure of the molecule, however, the effect on the function of the molecule is not

known. The lack of OKT4 IgG binding to CD4 868T suggests this tertiary structural change (64). One possibility would be an alteration to the normal function of the CD4 molecule causing an abnormal immune response by CD4+ cells resulting in an increased sensitivity to any infection. These differences could be manifested in an altered binding or capacity to activate cells (external changes) or in the ability of the molecule to initiate its signal transduction cascade which results in the production of cytokines (internal changes). External changes to CD4 could result in increased or decreased affinity of CD4 for its natural ligand. Chronic activation of CD4+ cells results in increased activation induced cell death (85) so if CD4 868T has a higher affinity which causes increased activation, or if the polymorphism somehow caused an increase in cytokine production, it may be the cause of the differential CD4+ cell decline in 868T positive patients. HIV proteins are known to interact with CD4 in the cell membrane. For example, Nef has been shown to down-modulate CD4 and cause its targeting to the degradative pathway thus increasing viral infectivity by an as of yet undescribed mechanism (86). If CD4 868T affects the affinity of Nef for CD4, a difference in the viral load and CD4+ T-cell decline could be an observed effect. Additionally, the efficiency of gp120 cross-linking CD4 on the surface of uninfected cells, which contributes to by-standard apoptosis, may be different with the different CD4 isotypes (87, 88). If the CD4 868T isotype has an increased affinity for gp120, it is possible that it results in an increased apoptosis of bystander T-cells as was observed by Holm *et al* with gp120 mutants of HIV-1 that had increased affinity for CD4 (89). The altered CD4 isotype could affect the binding efficiency of gp120 on HIV or possibly the coreceptor CCR5 which is constitutively associated with CD4 (90, 91). A more efficient gp120 – CD4 –

CCR5 interaction may result in increased sensitivity to infection during initial exposure and, once infected with HIV, could result in more infected cells over several cycles of virus replication and affect disease progression. As viral load has been highly correlated with disease progression, the relationship between the polymorphism and viral load must be examined. Unfortunately, in this study there were too few samples of sufficient quality to perform a complete viral load analysis. It is also possible that the altered CD4 molecule encoded by the 868T SNP is not the cause of the effects shown in this study but the polymorphism is in linkage disequilibrium with the true factor much like the haplotype of CCR2b-64I and CCR5 promoter alleles (92). The CD4 gene is located on chromosome 12 on the p arm at position 12, near the gene for B7-2. Further studies are required to address the nature of the mechanism of the association between CD4 868T and HIV infection and disease progression.

The results of our study show that the CD4 868T polymorphism may have a significant impact on HIV incidence, disease progression and mother to child transmission in these two cohorts. Given the combination of high prevalences of both HIV and the 868T polymorphism in African populations, the effect on the African epidemic could be dramatic. Studies in other cohorts with differing modes of HIV transmission, ethnic background, and HIV clades would be warranted. Further study of the mechanism of the enhanced disease progression could help further elucidate the mechanism of HIV-1 entry and specifically the complex interaction between gp120, CD4 and the coreceptors. Studies on binding affinities of both wild type and 868T CD4 molecules would help assess whether or not the CD4+ T-cell decline can be attributed to bystander apoptosis.

Work involving visualization of the interaction and recruitment of molecules such as coreceptors or viral proteins to the site of CD4 –gp 120 binding could also elucidate the underlying reasons for the observations outlined in this thesis. Sedgh *et al* noted that breastfeeding increased maternal HIV viral load and decreased CD4+ T-cell numbers (93). Perhaps an interesting direction might be to look at whether or not a second or subsequent child is more likely to be infected with HIV after one sibling is breastfed and whether or not the CD4 868T polymorphism has a role in these observations.

The work detailed in this thesis shows that a SNP in the coding region of CD4, an important protein of the immune system and the primary cellular receptor for HIV-1, is associated with more rapid disease progression, an increased HIV incidence, a more rapid CD4+ T-cell decline and an increased risk of breast milk transmission. To correlate these results to a polymorphism of this frequency in a population with such a high prevalence of HIV infection is an exceedingly important finding. In addition to possibly contributing to the magnitude of the African pandemic, the involvement of CD4 868T in breast milk transmission adds another somber warning against breastfeeding by HIV infected mothers. Although this is a detrimental suggestion in Africa, it may be further encouraged in the developed world among women of African descent who have other feeding options available to them. It is also another indication that not only antiretroviral drugs but also preventative measures such as premixed infant formula should be donated to the developing world. Although this work could be viewed as discouraging, any information that could lead to better and more accurate prognoses should be pursued. This work contributes to the identification of predictive factors that may be used to

develop novel treatment approaches and could be useful in the development of new therapeutic agents.

APPENDIX 1 - Solutions

1M Tris pH 8.0 – 100ml volume
15.76g Tris

adjust to pH 8.0 with NaOH

0.5M EDTA pH 8.0 – 100ml volume

23.26g disodium EDTA-H_2O

adjust pH with glacial acetic acid

TE buffer pH 8.0 – 100ml volume
(10mM Tris, 1mM EDTA)

1ml of 1M Tris

0.2ml of 0.5M EDTA

make up volume to 100ml with double distilled water

3M NaOAc pH 5.2 – 100ml volume
40.81 g NaOAc

adjust to pH 5.2 with acetic acid

Appendix 2 – Abbreviations

Ab - antibody

AIDS – acquired immunodeficiency syndrome

ARMS-PCR – amplification refractory mutation system polymerase chain reaction

ART – antiretroviral therapy

Bmt – breast milk transmission

CCR5 – c-chemokine receptor 5

CD4 – cluster of differentiation 4

CI95 – confidence interval 95%

CTL – cytotoxic T-lymphocyte

CXCR4 – cx-chemokine receptor 4

DNA – deoxyribonucleic acid

ddNTP – dideoxynucleotide triphosphate

dNTP – deoxynucleotide triphosphate

ELISA – enzyme linked immunosorbant assay

Env – envelope

g - gram

Gag – group specific antigen

gp120 – glycoprotein 120 kDa

HAART – highly active antiretroviral therapy

HIV – human immunodeficiency virus (indicates HIV-1)

HIV+ – HIV infected

HIV- – HIV uninfected

Ig - immunoglobulin

IL – interleukin

LTNP – long-term non-progressor

M - molar

MCH – mother-child health cohort

ML – commercial sex workers cohort

mg - milligram

mL - milliliter

MTCT – mother to child transmission

NRAMP – natural resistance associated monocyte protein

NTP- nucleotide triphosphate

O.R. – odds ratio

p - probability

PBMC – peripheral blood mononuclear cells

PCR- polymerase chain reaction

Pol - polymerase

R5 virus – HIV-1 with CCR5 tropism

RP – rapid progressor

RT – reverse transcriptase

RNA- ribonucleic acid

sc – seroconverter

SIV – simian immunodeficiency virus

STD/STI – sexually transmitted disease / infection

T-cell – thymus-originating lymphocyte

µL – microlitre

$?^2$ – Chi Square test

X4 virus – HIV-1 with CXCR4 tropism

xg – times gravity

References

1. 2003. UNAIDS/WHO global AIDS statistics. *AIDS Care 15:144.*

2. 1982. Update on acquired immune deficiency syndrome (AIDS)--United States. *MMWR Morb Mortal Wkly Rep 31:507.*

3. Masur, H., M. A. Michelis, J. B. Greene, I. Onorato, R. A. Stouwe, R. S. Holzman, G. Wormser, L. Brettman, M. Lange, H. W. Murray, and S. Cunningham-Rundles. 1981. An outbreak of community-acquired Pneumocystis carinii pneumonia: initial manifestation of cellular immune dysfunction. *N Engl J Med 305:1431.*

4. Kreiss, J. K., L. W. Kitchen, H. E. Prince, C. K. Kasper, and M. Essex. 1985. Antibody to human T-lymphotropic virus type III in wives of hemophiliacs. Evidence for heterosexual transmission. *Ann Intern Med 102:623.*

5. Chamberland, M. E., K. G. Castro, H. W. Haverkos, B. I. Miller, P. A. Thomas, R. Reiss, J. Walker, T. J. Spira, H. W. Jaffe, and J. W. Curran. 1984. Acquired immunodeficiency syndrome in the United States: an analysis of cases outside high-incidence groups. *Ann Intern Med 101:617.*

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

7. Gallo, R. C., P. S. Sarin, E. P. Gelmann, M. Robert-Guroff, E. Richardson, V. S. Kalyanaraman, D. Mann, G. D. Sidhu, R. E. Stahl, S. Zolla-Pazner, J. Leibowitch,

and M. Popovic. 1983. Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science 220:865.*

8. Coffin, J., A. Haase, J. A. Levy, L. Montagnier, S. Oroszlan, N. Teich, H. Temin, K. Toyoshima, H. Varmus, P. Vogt, and et al. 1986. Human immunodeficiency viruses. *Science 232:697.*

9. Clavel, F., D. Guetard, F. Brun-Vezinet, S. Chamaret, M. A. Rey, M. O. Santos-Ferreira, A. G. Laurent, C. Dauguet, C. Katlama, C. Rouzioux, and et al. 1986. Isolation of a new human retrovirus from West African patients with AIDS. *Science 233:343.*

10. Rambaut, A., D. Posada, K. A. Crandall, and E. C. Holmes. 2004. The causes and consequences of HIV evolution. *Nat Rev Genet 5:52.*

11. Signorelli, D., and R. J. Joseph, 2nd. 1998. Epidemiology of HIV and AIDS. A retrospective look. *Clin Podiatr Med Surg 15:179.*

12. Burns, D. N., and L. M. Mofenson. 1999. Paediatric HIV-1 infection. *Lancet 354 Suppl 2:SII1.*

13. Sullivan, J. L. 2003. Prevention of mother-to-child transmission of HIV--what next? *J Acquir Immune Defic Syndr 34 Suppl 1:S67.*

14. Sullivan, J. L., and K. Luzuriaga. 2001. The changing face of pediatric HIV-1 infection. *N Engl J Med 345:1568.*

15. Kriebs, J. M. 2002. The global reach of HIV: preventing mother-to-child transmission. *J Perinat Neonatal Nurs 16:1.*

16. Levy, J. A. 1993. Pathogenesis of human immunodeficiency virus infection. *Microbiol Rev 57:183.*

17. Sattentau, Q. J., and R. A. Weiss. 1988. The CD4 antigen: physiological ligand and HIV receptor. *Cell 52:631.*

18. Rowland-Jones, S. L. 1999. HIV: The deadly passenger in dendritic cells. *Curr Biol 9:R248.*

19. Cen, S., A. Khorchid, H. Javanbakht, J. Gabor, T. Stello, K. Shiba, K. Musier-Forsyth, and L. Kleiman. 2001. Incorporation of lysyl-tRNA synthetase into human immunodeficiency virus type 1. *J Virol 75:5043.*

20. Chinen, J., and W. T. Shearer. 2002. Molecular virology and immunology of HIV infection. *J Allergy Clin Immunol 110:189.*

21. Goepfert, P. A. 2003. Making sense of the HIV immune response. *Top HIV Med 11:4.*

22. Redfield, R. R., D. C. Wright, and E. C. Tramont. 1986. The Walter Reed staging classification for HTLV-III/LAV infection. *N Engl J Med 314:131.*

23. Shearer, W. T., T. C. Quinn, P. LaRussa, J. F. Lew, L. Mofenson, S. Almy, K. Rich, E. Handelsman, C. Diaz, M. Pagano, V. Smeriglio, and L. A. Kalish. 1997. Viral load and disease progression in infants infected with human immunodeficiency virus type 1. Women and Infants Transmission Study Group. *N Engl J Med 336:1337.*

24. Rouet, F., C. Sakarovitch, P. Msellati, N. Elenga, C. Montcho, I. Viho, S. Blanche, C. Rouzioux, F. Dabis, and V. Leroy. 2003. Pediatric viral human immunodeficiency virus type 1 RNA levels, timing of infection, and disease progression in African HIV-1-infected children. *Pediatrics 112:e289.*

25. Delozier, E. P. 1995. Health Services Technology/Assessment Text (HSTAT). *Med Ref Serv Q 14:39.*

26. Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, B. H. Hahn, and et al. 1995. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature 373:117.*

27. Pantaleo, G., C. Graziosi, J. F. Demarest, L. Butini, M. Montroni, C. H. Fox, J. M. Orenstein, D. P. Kotler, and A. S. Fauci. 1993. HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease. *Nature 362:355.*

28. Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz. 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature 373:123.*

29. Fowke, K. R., N. J. Nagelkerke, J. Kimani, J. N. Simonsen, A. O. Anzala, J. J. Bwayo, K. S. MacDonald, E. N. Ngugi, and F. A. Plummer. 1996. Resistance to HIV-1 infection among persistently seronegative prostitutes in Nairobi, Kenya [see comments]. *Lancet 348:1347.*

30. Rowland-Jones, S., J. Sutton, K. Ariyoshi, T. Dong, F. Gotch, S. McAdam, D. Whitby, S. Sabally, A. Gallimore, T. Corrah, and et al. 1995. HIV-specific cytotoxic T-cells in HIV-exposed but uninfected Gambian women [published erratum appears in Nat Med 1995 Jun;1(6):598]. *Nat Med 1:59.*

31. Girard, M., B. Meignier, F. Barre-Sinoussi, M. P. Kieny, T. Matthews, E. Muchmore, P. L. Nara, Q. Wei, L. Rimsky, K. Weinhold, and et al. 1995.

Vaccine-induced protection of chimpanzees against infection by a heterologous human immunodeficiency virus type 1. *J Virol 69:6239.*

32. Berman, P. W., K. K. Murthy, T. Wrin, J. C. Vennari, E. K. Cobb, D. J. Eastman, M. Champe, G. R. Nakamura, D. Davison, M. F. Powell, J. Bussiere, D. P. Francis, T. Matthews, T. J. Gregory, and J. F. Obijeski. 1996. Protection of MN-rgp120-immunized chimpanzees from heterologous infection with a primary isolate of human immunodeficiency virus type 1. *J Infect Dis 173:52.*

33. Anzala, O. A., N. J. Nagelkerke, J. J. Bwayo, D. Holton, S. Moses, E. N. Ngugi, J. O. Ndinya-Achola, and F. A. Plummer. 1995. Rapid progression to disease in African sex workers with human immunodeficiency virus type 1 infection. *J Infect Dis 171:686.*

34. Easterbrook, P. J. 1999. Long-term non-progression in HIV infection: definitions and epidemiological issues. *J Infect 38:71.*

35. Rowland-Jones, S. 1999. Long-term non-progression in HIV infection: clinico pathological issues. *J Infect 38:67.*

36. Simonsen, J. N., K. R. Fowke, K. S. MacDonald, and F. A. Plummer. 1998. HIV pathogenesis: mechanisms of susceptibility and disease progression. *Curr Opin Microbiol 1:423.*

37. Cao, Y., L. Qin, L. Zhang, J. Safrit, and D. D. Ho. 1995. Virologic and immunologic characterization of long-term survivors of human immunodeficiency virus type 1 infection. *N Engl J Med 332:201.*

38. Laitinen, T., A. Polvi, P. Rydman, J. Vendelin, V. Pulkkinen, P. Salmikangas, S. Makela, M. Rehn, A. Pirskanen, A. Rautanen, M. Zucchelli, H. Gullsten, M.

Leino, H. Alenius, T. Petays, T. Haahtela, A. Laitinen, C. Laprise, T. J. Hudson, L. A. Laitinen, and J. Kere. 2004. Characterization of a common susceptibility locus for asthma-related traits. *Science 304:300.*

39. Goldfeld, A. E., J. C. Delgado, S. Thim, M. V. Bozon, A. M. Uglialoro, D. Turbay, C. Cohen, and E. J. Yunis. 1998. Association of an HLA-DQ allele with clinical tuberculosis. *Jama 279:226.*

40. Abel, L., F. O. Sanchez, J. Oberti, N. V. Thuc, L. V. Hoa, V. D. Lap, E. Skamene, P. H. Lagrange, and E. Schurr. 1998. Susceptibility to leprosy is linked to the human NRAMP1 gene. *J Infect Dis 177:133.*

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

42. Huang, Y., W. A. Paxton, S. M. Wolinsky, A. U. Neumann, L. Zhang, T. He, S. Kang, D. Ceradini, Z. Jin, K. Yazdanbakhsh, K. Kunstman, D. Erickson, E. Dragon, N. R. Landau, J. Phair, D. D. Ho, and R. A. Koup. 1996. The role of a mutant CCR5 allele in HIV-1 transmission and disease progression [see comments]. *Nat Med 2:1240.*

43. Dean, M., M. Carrington, C. Winkler, G. A. Huttley, M. W. Smith, R. Allikmets, J. J. Goedert, S. P. Buchbinder, E. Vittinghoff, E. Gomperts, S. Donfield, D. Vlahov, R. Kaslow, A. Saah, C. Rinaldo, R. Detels, and S. J. O'Brien. 1996. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study,

Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science 273:1856.*

44. Eugen-Olsen, J., A. K. Iversen, P. Garred, U. Koppelhus, C. Pedersen, T. L. Benfield, A. M. Sorensen, T. Katzenstein, E. Dickmeiss, J. Gerstoft, P. Skinhoj, A. Svejgaard, J. O. Nielsen, and B. Hofmann. 1997. Heterozygosity for a deletion in the CKR-5 gene leads to prolonged AIDS-free survival and slower CD4 T-cell decline in a cohort of HIV-seropositive individuals. *Aids 11:305.*

45. Mulherin, S. A., T. R. O'Brien, J. P. Ioannidis, J. J. Goedert, S. P. Buchbinder, R. A. Coutinho, B. D. Jamieson, L. Meyer, N. L. Michael, G. Pantaleo, G. P. Rizzardi, H. Schuitemaker, H. W. Sheppard, I. D. Theodorou, D. Vlahov, and P. S. Rosenberg. 2003. Effects of CCR5-Delta32 and CCR2-64I alleles on HIV-1 disease progression: the protection varies with duration of infection. *Aids 17:377.*

46. Samson, M., F. Libert, B. J. Doranz, J. Rucker, C. Liesnard, C. M. Farber, S. Saragosti, C. Lapoumeroulie, J. Cognaux, C. Forceille, G. Muyldermans, C. Verhofstede, G. Burtonboy, M. Georges, T. Imai, S. Rana, Y. Yi, R. J. Smyth, R. G. Collman, R. W. Doms, G. Vassart, and M. Parmentier. 1996. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene [see comments]. *Nature 382:722.*

47. Stewart, G. J., L. J. Ashton, R. A. Biti, R. A. Ffrench, B. H. Bennetts, N. R. Newcombe, E. M. Benson, A. Carr, D. A. Cooper, and J. M. Kaldor. 1997. Increased frequency of CCR-5 delta 32 heterozygotes among long-term non-progressors with HIV-1 infection. The Australian Long-Term Non- Progressor Study Group. *Aids 11:1833.*

48. Ioannidis, J. P., D. G. Contopoulos-Ioannidis, P. S. Rosenberg, J. J. Goedert, A. De Rossi, T. Espanol, L. Frenkel, M. J. Mayaux, M. L. Newell, S. G. Pahwa, C. Rousseau, G. Scarlatti, S. Sei, L. Sen, and T. R. O'Brien. 2003. Effects of CCR5-delta32 and CCR2-64I alleles on disease progression of perinatally HIV-1-infected children: an international meta-analysis. *Aids 17:1631.*

49. Philpott, S., H. Burger, T. Charbonneau, R. Grimson, S. H. Vermund, A. Visosky, S. Nachman, A. Kovacs, P. Tropper, H. Frey, and B. Weiser. 1999. CCR5 genotype and resistance to vertical transmission of HIV-1. *J Acquir Immune Defic Syndr 21:189.*

50. Soriano, A., C. Martinez, F. Garcia, M. Plana, E. Palou, M. Lejeune, J. I. Arostegui, E. De Lazzari, C. Rodriguez, A. Barrasa, J. I. Lorenzo, J. Alcami, J. del Romero, J. M. Miro, J. M. Gatell, and T. Gallart. 2002. 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 186:922.*

51. John, G. C., C. Rousseau, T. Dong, S. Rowland-Jones, R. Nduati, D. Mbori-Ngacha, T. Rostron, J. K. Kreiss, B. A. Richardson, and J. Overbaugh. 2000. Maternal SDF1 3'A polymorphism is associated with increased perinatal human immunodeficiency virus type 1 transmission. *J Virol 74:5736.*

52. Kwa, D., R. P. van Rij, B. Boeser-Nunnink, J. Vingerhoed, and H. Schuitemaker. 2003. Association between an interleukin-4 promoter polymorphism and the acquisition of CXCR4 using HIV-1 variants. *Aids 17:981.*

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

54. Tang, J., B. Shelton, N. J. Makhatadze, Y. Zhang, M. Schaen, L. G. Louie, J. J. Goedert, E. C. Seaberg, J. B. Margolick, J. Mellors, and R. A. Kaslow. 2002. Distribution of chemokine receptor CCR2 and CCR5 genotypes and their relative contribution to human immunodeficiency virus type 1 (HIV-1) seroconversion, early HIV-1 RNA concentration in plasma, and later disease progression. *J Virol 76:662.*

55. Faure, S., L. Meyer, D. Costagliola, C. Vaneensberghe, E. Genin, B. Autran, J. F. Delfraissy, D. H. McDermott, P. M. Murphy, P. Debre, I. Theodorou, and C. Combadiere. 2000. Rapid progression to AIDS in HIV+ individuals with a structural variant of the chemokine receptor CX3CR1. *Science 287:2274.*

56. Faure, S., L. Meyer, E. Genin, P. Pellet, P. Debre, I. Theodorou, and C. Combadiere. 2003. Deleterious genetic influence of CX3CR1 genotypes on HIV-1 disease progression. *J Acquir Immune Defic Syndr 32:335.*

57. Brumme, Z. L., W. W. Dong, K. J. Chan, R. S. Hogg, J. S. Montaner, M. V. O'Shaughnessy, and P. R. Harrigan. 2003. Influence of polymorphisms within the CX3CR1 and MDR-1 genes on initial antiretroviral therapy response. *Aids 17:201.*

58. Choe, H. R., and J. Sodroski. 1992. Contribution of charged amino acids in the CDR2 region of CD4 to HIV-1 gp120 binding. *J Acquir Immune Defic Syndr 5:204.*

59. Maddon, P. J., D. R. Littman, M. Godfrey, D. E. Maddon, L. Chess, and R. Axel. 1985. The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: a new member of the immunoglobulin gene family. *Cell 42:93.*

60. Furukawa, M., M. Yasukawa, Y. Yakushijin, and S. Fujita. 1994. Distinct effects of human herpesvirus 6 and human herpesvirus 7 on surface molecule expression and function of CD4+ T cells. *J Immunol 152:5768.*

61. Lusso, P., P. Secchiero, R. W. Crowley, A. Garzino-Demo, Z. N. Berneman, and R. C. Gallo. 1994. CD4 is a critical component of the receptor for human herpesvirus 7: interference with human immunodeficiency virus. *Proc Natl Acad Sci U S A 91:3872.*

62. Ravichandran, K. S., T. L. Collins, and S. J. Burakoff. 1996. CD4 and signal transduction. *Curr Top Microbiol Immunol 205:47.*

63. Perez-Sala, D., and A. Rebollo. 1999. Novel aspects of Ras proteins biology: regulation and implications. *Cell Death Differ 6:722.*

64. Hodge, T. W., D. R. Sasso, and J. S. McDougal. 1991. Humans with OKT4-epitope deficiency have a single nucleotide base change in the CD4 gene, resulting in substitution of TRP240 for ARG240. *Hum Immunol 30:99.*

65. Fuller, T. C., J. E. Trevithick, A. A. Fuller, R. B. Colvin, A. B. Cosimi, and P. C. Kung. 1984. Antigenic polymorphism of the T4 differentiation antigen expressed on human T helper/inducer lymphocytes. *Hum Immunol 9:89.*

66. Saleh, M., J. P. Vaillancourt, R. K. Graham, M. Huyck, S. M. Srinivasula, E. S. Alnemri, M. H. Steinberg, V. Nolan, C. T. Baldwin, R. S. Hotchkiss, T. G.

Buchman, B. A. Zehnbauer, M. R. Hayden, L. A. Farrer, S. Roy, and D. W. Nicholson. 2004. Differential modulation of endotoxin responsiveness by human caspase-12 polymorphisms. *Nature 429:75.*

67. Lederman, S., J. A. DeMartino, B. L. Daugherty, I. Foeldvari, M. J. Yellin, A. M. Cleary, N. Berkowitz, I. Lowy, N. S. Braunstein, G. E. Mark, and et al. 1991. A single amino acid substitution in a common African allele of the CD4 molecule ablates binding of the monoclonal antibody, OKT4. *Mol Immunol 28:1171.*

68. Macdonald, W. A., A. W. Purcell, N. A. Mifsud, L. K. Ely, D. S. Williams, L. Chang, J. J. Gorman, C. S. Clements, L. Kjer-Nielsen, D. M. Koelle, S. R. Burrows, B. D. Tait, R. Holdsworth, A. G. Brooks, G. O. Lovrecz, L. Lu, J. Rossjohn, and J. McCluskey. 2003. A Naturally Selected Dimorphism within the HLA-B44 Supertype Alters Class I Structure, Peptide Repertoire, and T Cell Recognition. *J Exp Med 198:679.*

69. Matthias, L. J., P. T. Yam, X. M. Jiang, N. Vandegraaff, P. Li, P. Poumbourios, N. Donoghue, and P. J. Hogg. 2002. Disulfide exchange in domain 2 of CD4 is required for entry of HIV-1. *Nat Immunol 3:727.*

70. Kreiss, J. K., D. Koech, F. A. Plummer, K. K. Holmes, M. Lightfoote, P. Piot, A. R. Ronald, J. O. Ndinya-Achola, L. J. D'Costa, P. Roberts, and et al. 1986. AIDS virus infection in Nairobi prostitutes. Spread of the epidemic to East Africa. *N Engl J Med 314:414.*

71. Plummer, F. A., and J. O. Ndinya-Achola. 1990. Sexually transmitted diseases and HIV-1: interactions in transmission and role in control programs [editorial]. *East Afr Med J 67:457.*

72. Plummer, F. A., M. A. Wainberg, P. Plourde, P. Jessamine, L. J. D'Costa, I. A. Wamola, and A. R. Ronald. 1990. Detection of human immunodeficiency virus type 1 (HIV-1) in genital ulcer exudate of HIV-1-infected men by culture and gene amplification [letter]. *J Infect Dis 161:810.*

73. Ronald, A., F. Plummer, E. Ngugi, J. O. Ndinya-Achola, P. Piot, J. Kreiss, and R. Brunham. 1991. The Nairobi STD program. An international partnership. *Infect Dis Clin North Am 5:337.*

74. Braddick, M. R., J. K. Kreiss, J. B. Embree, P. Datta, J. O. Ndinya-Achola, H. Pamba, G. Maitha, P. L. Roberts, T. C. Quinn, and K. K. Holmes. 1990. Impact of maternal HIV infection on obstetrical and early neonatal outcome. *Aids 4:1001.*

75. Embree, J. E., M. Braddick, P. Datta, J. Muriithi, C. Hoff, J. K. Kreiss, P. L. Roberts, B. J. Law, H. O. Pamba, J. O. Ndinya-Achola, and et al. 1989. Lack of correlation of maternal human immunodeficiency virus infection with neonatal malformations. *Pediatr Infect Dis J 8:700.*

76. Lancaster, A., M. P. Nelson, D. Meyer, G. Thomson, and R. M. Single. 2003. PyPop: a software framework for population genomics: analyzing large-scale multi-locus genotype data. *Pac Symp Biocomput:514.*

77. MacDonald, K. S., K. R. Fowke, J. Kimani, V. A. Dunand, N. J. Nagelkerke, T. B. Ball, J. Oyugi, E. Njagi, L. K. Gaur, R. C. Brunham, J. Wade, M. A. Luscher, P. Krausa, S. Rowland-Jones, E. Ngugi, J. J. Bwayo, and F. A. Plummer. 2000. Influence of HLA supertypes on susceptibility and resistance to human immunodeficiency virus type 1 infection. *J Infect Dis 181:1581.*

78. Anzala, O. A., N. J. Nagelkerke, J. J. Bwayo, D. Holton, S. Moses, E. N. Ngugi, J. O. Ndinya-Achola, and F. A. Plummer. 1995. Rapid progression to disease in African sex workers with human immunodeficiency virus type 1 infection [see comments]. *J Infect Dis 171:686.*

79. Paxton, W. A., and S. Kang. 1998. Chemokine receptor allelic polymorphisms: relationships to HIV resistance and disease progression. *Semin Immunol 10:187.*

80. Coutsoudis, A., F. Dabis, W. Fawzi, P. Gaillard, G. Haverkamp, D. R. Harris, J. B. Jackson, V. Leroy, N. Meda, P. Msellati, M. L. Newell, R. Nsuati, J. S. Read, and S. Wiktor. 2004. Late postnatal transmission of HIV-1 in breast-fed children: an individual patient data meta-analysis. *J Infect Dis 189:2154.*

81. Cunningham, C. K., R. Balasubramanian, I. Delke, R. Maupin, L. Mofenson, A. Dorenbaum, J. L. Sullivan, A. Gonzalez-Garcia, E. Thorpe, M. Rathore, and R. D. Gelber. 2004. The Impact of Race/Ethnicity on Mother-to-Child HIV Transmission in the United States in Pediatric AIDS Clinical Trials Group Protocol 316. *J Acquir Immune Defic Syndr 36:800.*

82. Russell, P. J. 1998. *Genetics*. Benjamin/Cummings Publishing Company, Menlo Park, CA.

83. Hebbel, R. P. 2003. Sickle hemoglobin instability: a mechanism for malarial protection. *Redox Rep 8:238.*

84. O'Brien, S. J., and J. P. Moore. 2000. The effect of genetic variation in chemokines and their receptors on HIV transmission and progression to AIDS. *Immunol Rev 177:99.*

85. Lawn, S. D., S. T. Butera, and T. M. Folks. 2001. Contribution of immune activation to the pathogenesis and transmission of human immunodeficiency virus type 1 infection. *Clin Microbiol Rev 14:753.*

86. Alexander, M., Y. C. Bor, K. S. Ravichandran, M. L. Hammarskjold, and D. Rekosh. 2004. Human immunodeficiency virus type 1 Nef associates with lipid rafts to downmodulate cell surface CD4 and class I major histocompatibility complex expression and to increase viral infectivity. *J Virol 78:1685.*

87. Bottarel, F., M. J. Feito, M. Bragardo, S. Bonissoni, D. Buonfiglio, S. DeFranco, F. Malavasi, T. Bensi, U. Ramenghi, and U. Dianzani. 1999. The cell death-inducing ability of glycoprotein 120 from different HIV strains correlates with their ability to induce CD4 lateral association with CD95 on CD4+ T cells. *AIDS Res Hum Retroviruses 15:1255.*

88. Desbarats, J., J. H. Freed, P. A. Campbell, and M. K. Newell. 1996. Fas (CD95) expression and death-mediating function are induced by CD4 cross-linking on CD4+ T cells. *Proc Natl Acad Sci U S A 93:11014.*

89. Holm, G. H., C. Zhang, P. R. Gorry, K. Peden, D. Schols, E. De Clercq, and D. Gabuzda. 2004. Apoptosis of bystander T cells induced by human immunodeficiency virus type 1 with increased envelope/receptor affinity and coreceptor binding site exposure. *J Virol 78:4541.*

90. Xiao, X., L. Wu, T. S. Stantchev, Y. R. Feng, S. Ugolini, H. Chen, Z. Shen, J. L. Riley, C. C. Broder, Q. J. Sattentau, and D. S. Dimitrov. 1999. Constitutive cell surface association between CD4 and CCR5. *Proc Natl Acad Sci U S A 96:7496.*

91. Wang, X., and R. Staudinger. 2003. Interaction of soluble CD4 with the chemokine receptor CCR5. *Biochem Biophys Res Commun 307:1066.*

92. Kostrikis, L. G., Y. Huang, J. P. Moore, S. M. Wolinsky, L. Zhang, Y. Guo, L. Deutsch, J. Phair, A. U. Neumann, and D. D. Ho. 1998. A chemokine receptor CCR2 allele delays HIV-1 disease progression and is associated with a CCR5 promoter mutation. *Nat Med 4:350.*

93. Sedgh, G., D. Spiegelman, U. Larsen, G. Msamanga, and W. W. Fawzi. 2004. Breastfeeding and maternal HIV-1 disease progression and mortality. *Aids 18:1043.*