2012

# C-C-motif chemokine receptor/ligand gene variants and their associations with HIV-1 transmission and pathogenesis

Liangyuan Hu
*University of Alabama at Birmingham*

# C-C-MOTIF CHEMKINE RECEPTOR/LIGAND GENE VARIANTS AND THEIR ASSOCIATIONS WITH HIV-1 TRANSMISSION AND PATHOGENESIS

By

LIANGYUAN HU

RICHARD A. KASLOW, ADVISOR
JEFFERY C EDBERG, COMMITTEE CHAIR
S LOUIS BRIDGES JR
SADEEP SHRESTHA
JIANMING TANG

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2012

C-C-MOTIF CHEMKINE RECEPTOR/LIGAND GENE VARIANTS AND THEIR ASSOCIATIONS WITH HIV-1 TRANSMISSION AND PATHOGENESIS

LIANGYUAN HU

DOCTOR OF PHILOSOPHY IN GENETICS AND GENOMIC SCIENCES

ABSTRACT

Host factors including genes and their variants are important to HIV-1 acquisition, transmission and disease progression. In particular, chemokine (C-C motif) receptors 2 and 5 genes (*CCR2* and *CCR5*) have multiple variants of interest. We first investigated the impact of *CCR2*-*CCR5* haplotypes on several outcomes among 567 HIV-1 discordant Zambian couples. HHF*2 homozygosity was associated with significantly lower VL in seroconverters (mean β=-0.58 $\log_{10}$ *P*=0.027) and the HHD/HHE diplotype was associated with significantly higher VL in the seroconverters (mean β=0.54, $\log_{10}$ *P*=0.014) adjusted for age and gender in multivariable model. HHD/HHE was associated with more rapid acquisition of infection by the HIV-1 exposed seronegatives (HESN) (HR=2.0, 95% CI=1.20-3.43, *P*=0.008), after adjustments for index partner VL and the presence of genital ulcer or inflammation in either partner in Cox multivariable models.

CC-motif chemokine ligands (CCLs) can block HIV-1 binding sites on CCR5 and inhibit viral entry. We studied single nucleotide polymorphisms (SNPs) in genes encoding three CCR5 ligands [CCL3 (MIP-1α), CCL4 (MIP-1β), and CCL5 (RANTES)] along with an adjacent gene encoding a CCR2 ligand [CCL2 (MCP-1)] on HIV-1 VL and

heterosexual transmission in Zambia cohort. We found that rs5029410 C allele (in CCL3 intron 2) was associated with lower VL in seroconverters, adjusted for gender and age (regression β=-0.57 $\log_{10}$, *P*=4x10-6). In addition, rs34171309 A allele in CCL3 exon 3 was associated with increased risk of HIV-1 acquisition in HESNs (hazard ratio=1.52, *P*=0.006) when adjusted for donor VL and genital ulcer/inflammation.

We further screened variants in *CCR5* gene and untranslated regions (UTRs) of *CCR2* gene in order to rule out the potential confounding by population-specific variants not reported in the literature. Among 27 SNPs found in 109 Zambian samples with representative CCR2-CCR5 diplotypes, 5 are within the *CCR5* coding region. Genotyping for the entire Zambian cohort (567 couples) revealed that allele T for *CCR5* SNP rs1800944 (encoding alanine to valine change) was associated with slightly higher viral load in donor partners, after adjusting for age and gender (β= 0.24 $\log_{10}$, *P*=0.026). Analyses for other outcomes and other SNPs were not conclusive.

Overall, our data favor the hypothesis that host genetic variants within several C-C-motif chemokine ligand and receptor genes can mediate HIV-1 acquisition and control of infection in our study population. Future investigation may need to focus on a) functional relevance of CCL and CCR variants; b) the potential interaction of CCL and CCR gene variants and other HIV-1-related host factors; and c) viral evolution attributable to CCR and CCL gene variants. Such efforts will be essential to producing a clear picture about the innate and adaptive immunity to HIV-1 infection.

ACKNOWLEDGMENTS

First and foremost, I have sincerest gratitude to my supervisor, Dr Richard Kaslow, who has supported me throughout my thesis with his patience and knowledge. I would like to acknowledge my committee chair Dr. Jeffery Edberg and committee members Drs. Jianming (James) Tang, Sadeep Shrestha and Louis Bridges for all their contribution to this research. I am very grateful to all investigators, staff, and participants of the Zambia-Emory HIV-1 Research Project, without whom this research would never be possible.

This study was supported primarily by the U.S. National Institute of Allergy and Infectious Diseases, through grants AI041951 and AI071906 to R.A.K., AI040951 to S.A., AI064060 to E.H., and AI076123 to J.T.

Finally, I thank my family and friends for their supporting me throughout all my studies.

TABLE OF CONTENTS

*Page*

TABLE OF CONTENTS (Continued)

*Page*

LIST OF FIGURES

*Figure* *Page*

LIST OF TABLES

*Table* *Page*

LIST OF TABLES (Continued)

*Table* *Page*

# INTRODUCTION

## Epidemiology of HIV-1 Infection

Globally, there are approximately 33.3 million people living with HIV-1 according to the "2010 Report on the Global AIDS Epidemic [1]." The annual number of new infection has been steadily declining since 1997, and the overall growth of the global acquired immunodeficiency syndrome (AIDS) epidemic caused by HIV-1 appears to have stabilized. However, the number of new infections overall are still very high. In 2009 alone, 2.6 million people were newly infected, with 97% of them are living in low and middle income countries. HIV-1 and AIDS are found in all parts of the world, but sub-Saharan Africa is the worst affected region, where 22.5 million people are living with HIV-1, and the prevalence is about 5% of the adult population [2].

There is still no cure for AIDS. Although the AIDS related deaths are also declining due to the improvements in the antiretroviral therapy over the past few years, about 1.8 million people lost their lives because of AIDS in 2009 [1]. More than 5 million people are now receiving HIV-1 therapy, but this number represents only 35% of the people who need HIV-1 therapy now according to World Health Organization guidelines.

There are 3 major ways of acquiring HIV-1 infection: through unprotected sexual relation, exposure to blood from infected individuals and mother-to-child transmission. Having unprotected sex, sharing needles/syringes, having other sexually transmitted

infections, or being born to an infected mother increases risk for HIV-1 according to Centers for Disease Control and Prevention report. In Africa, heterosexual transmission is the dominant mode of HIV-1 transmission. Data from urban Zambia suggest that 60% of the newly infected persons gain HIV-1 through heterosexual transmission [1].

Better understanding of the mechanism of HIV-1 infection is needed to provide better methods to control its spread as well as more effective and affordable therapy for infected individuals.

**Host Genetic Polymorphism and HIV-1 Infection**

Host genetic polymorphisms play a substantial role in the susceptibility and course of HIV-1 infection. Previous work has led to findings that show some consistency but also considerably variation from one population to another. As a result, the precise role of the multiple variants in the implicated host genes in altering the risk of HIV-1 infection and subsequent disease is not yet well understood [3].

**C-C-motif chemokine receptor 5 (CCR5)**

HIV-1 enters macrophages via its required CD4 receptor and coreceptors on the surface of the target cell followed by fusion of the viral envelope with the cell membrane and the release of the HIV-1 RNA into the cell [4]. CCR5 is one of the major coreceptors of HIV-1. *CCR5* Δ32 deletion is the most widely studied mutation in *CCR5* gene [5, 6]. It is a deletion of the nucleotides 794 to 825 of the cDNA sequence and results in a reading frame shift after amino acid 174. It has been shown that the severely truncated protein could not be detected at the surface of cells that normally express the protein. Extensive work has demonstrated that *CCR5* Δ32 acts as a recessive restriction gene against HIV-1

infection and exerts a dominant effect in delaying progression to AIDS among infected individuals [7]. The protective effect of *CCR5* Δ32 allele (rs333) has been shown repeatedly in different populations. However, this allele is virtually absent in non-European ethnic groups, and the vast majority of people exposed to but uninfected by HIV-1 do not carry the Δ32 allele. Thus, some other factors must account for their resistance to HIV-1 infection [8]. Multiple studies have detected variants in *CCR5* and other C-C-motif chemokine receptor genes in different populations [9-15]. However, these associations have not been consistent across the population.

**C-C-motif chemokine receptor 2 (CCR2)**

*CCR5* and *CCR2* are tightly linked on 3p22-p21, separated by 9 kb. There are 21 SNPs in *CCR2* and *CCR5* gene region and 5 kb upstream, 10 kb downstream in the HapMap database version 3 release 2 (http://hapmap.ncbi.nlm.nih.gov/). There are 20 SNPs with minor allele frequency greater or equal to 0.01 in the YRI population as showed in Figure 1.1. The CCR2-64I (rs1799864) allele is in strong linkage disequilibrium (LD) with specific alleles of the *CCR5* gene. SNPs in the 5' cis regulatory region of *CCR5* are reported to be associated with HIV-1 susceptibility, indicating that this region might be a target of natural selection. Gonzalez *et al* characterized haplotypes comprising alleles at 8 biallelic loci in the *CCR2*, *CCR5* gene region, namely HHA, HHB, HHC, HHD, HHE, HHF*1, HHF*2 (with CCR2-64I allele, rs1799864), HHG*1 and HHG*2 (with *CCR5* Δ32 allele, rs333) (Table 1.1) [9]. Our research group reported that homozygosity of the *CCR2-CCR5* HHE haplotype was associated with higher HIV-1 RNA level during the initial 42 month after HIV-1 transmission in African-American population [16]. In Caucasians, the HHF*2/HHG*2 diplotype was associated with

lowest HIV-1 RNA level (viral load) [16]. Investigation of the variants in *CCR2* and *CCR5* genes in a larger African population may provide a better understanding of the influence on HIV transmission and acquisition.

**C-C-motif chemokine ligands (CCLs)**

*CCL2, CCL3, CCL3L, CCL4, CCL4L* and *CCL5* genes are located on chromosome 17q12 region. Figure 1.2 illustrated the relative position of these *CCL* genes. *CCL2, CCL3, CCL4, CCL5* are 2 copies genes in humans. However, *CCL3L1*, *CCL3L3,* and *CCL4L* have variable copy numbers at the individual and population levels [17].

CCL2 is a natural ligand for CCR2. It has been reported that CCL2 level in the cell is strongly associated with viral load in HIV+ patients [18]. Studies on white Spaniards shows that the GG genotype at nucleotide position -2518 in *CCL2* gene promoter is over represented in HIV-1 infected subjects [19]. A recent report shows that CCL2 may assist HIV-1 entering CD4 T cells by up-regulates CXCR4 on resting CD4+ T cells in a CCR2-dependent mechanism and thereby contribute to variability in HIV infection or disease progression [20].

C-C-motif chemokine ligands *CCL3* and *CCL4* are natural ligands for CCR5. They can inhibit or reduce viral entry through CD4+ T cells through binding to CCR5 or through other indirect mechanisms [21, 22]. *CCL3* and *CCL4* are both single copy genes located on human chromosome 17. However, their paralogues, *CCL3L1-CCL3L3* and *CCL4L1-CCL4L3,* are multiple copy genes that show copy number variation (CNV), which differs in the median number in different populations. Copy numbers of *CCL3L1*

and *CCL4L1* correlate strongly with each other. It has been reported that CNV of *CCL3L1* is associated with HIV-1 infection and progression in European and Hispanic population [23]. To be precise, in this early study, a lower copy number of *CCL3L1* than the median for a given ethnic group appeared to increases the risk for HIV-1 infection. However, subsequent studies have yielded inconsistent results, and the influence of dosage of *CCL3L1* in HIV infection is still under debate. A recent report described the absence of any substantial effect of *CCL3L1* CNV on HIV-1 infection, viral load or disease progression based on a study of 277 African Americans and 1493 Europeans [24].

Similar to CCL3L genes, CCL4L genes also represent a composite of two sets of genes. *CCL4L1* and *CCL4L2* have the same exonic sequences. However, *CCL4L1* has an A to G transition in the splicing site of intron 2, producing aberrantly spliced transcripts. This gene may not be functional. On the other hand, *CCL4L2* is predicted to have a classic transcript for a chemokine [17]. Our previous study investigated 184 HIV-1 seronegative control and 227 HIV-1 seropositive adolescents showed that neither CCL3L1 nor CCL4L1 gene copy number variation had impact on susceptibility to or control of HIV-1 infection [25]. However, it was subsequently recognized that the complexity of variants in the sequence and numbers of CCL3L and CCL4L copies have not been sufficiently well defined to quantify copies of each gene or to characterize their function. Therefore, we did not continue our study on the CNVs in the Zambia cohort.

CCL5 is also an important ligand for CCR5. In 2002, a study based on 4,168 individuals from five AIDS cohort reported that SNPs and haplotypes in the *CCL5* gene have a strong dominant association with rapid progression to AIDS among HIV-1 infected African Americans [26]. Although CCL3 is regarded as the most potent ligand

for CCR5, with 10 times greater affinity than CCL5 (tested by ELISA) , CCL5 generally circulates in the serum at a much higher concentration (1000 times higher) than CCL3 or CCL4 [27]. Therefore, *CCL5* is generally regarded as an important gene in the HIV-1 infection process, and making it critical to establish whether variants in *CCL5* influence HIV-1 transmission and acquisition [25].

Through our studies, we developed a model that effectively predicts the risk of HIV-1 transmission and acquisition based on the carriage of specific combinations of the ligand and receptor variants.

## Figures

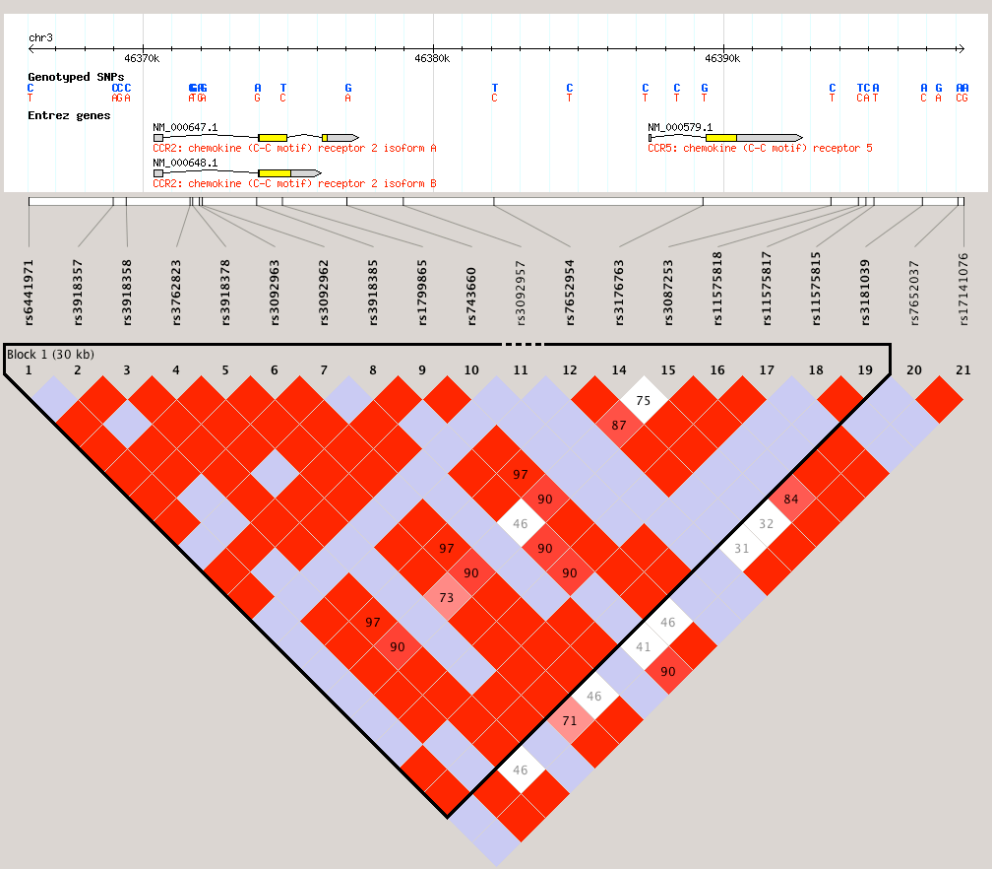



Figure 1.1 Patterns of linkage disequilibrium among 21 SNPs in *CCR2-CCR5* gene region with minor allele frequency>0.01 in the HapMap database (version 3, release 2). D' values as shown are based on HaploView version 4.2.


Figure 1.2. Figure shows the relative position of CCL gene family on the chromosome 17. The length of the genes is not drawn in scale. *CCL3L1* and *CC3L3* gene have identical exonic region. This figure was based on the NCBI reference sequences.

**Tables**

Table 1.1 Gonzalez *et al* characterized haplotypes comprising alleles at 8 biallelic loci in the *CCR2*, *CCR5* gene region, namely HHA, HHB, HHC, HHD, HHE, HHF*1, HHF*2 (with CCR2-64I allele, rs1799864), HHG*1 and HHG*2 (with *CCR5* Δ32 allele, rs333).

| Haplotypes | rs1799864 | rs2856758 | rs2734648 | rs1799987 | rs1799988 | rs41469351 | rs1800023 | rs1800024 |
|---|---|---|---|---|---|---|---|---|
| HHA | G | A | G | G | T | C | A | C |
| HHB | G | A | T | G | T | C | A | C |
| HHC | G | A | T | G | T | C | G | C |
| HHD | G | A | T | G | T | T | A | C |
| HHE | G | A | G | A | C | C | A | C |
| HHF*1 | G | A | G | A | C | C | A | T |
| HHF*2 | A | A | G | A | C | C | A | T |
| HHG*1 | G | G | G | A | C | C | A | C |
| HHG*2 | G | G | G | A | C | C | A | C |

# PRIMARY HYPOTHESES AND STUDY AIMS

**Primary Hypothesis:**

Host genetic polymorphisms and HIV-1 genomic polymorphisms are associated with HIV-1 transmission and disease control. Variants in host *CCR2* and *CCR5* and their natural ligands *CCL2*, *CCL3*, *CCL4* and *CCL5* are associated with HIV-1 acquisition and disease control in African population.

**Aim 1.** To determine and study the frequency and association of the *CCR2-CCR5* haplotypes among the HIV-1 serodiscordant couples in Zambia.

**Aim 2.** To screen the variants in *CCL2, CCL3, CCL4* and *CCL5* genes in Zambians, and to study their association with HIV-1 acquisition and disease control.

**Aim 3**. To draw a fine map of major variants in the *CCR2* 3' and 5' untranslated regions and in the *CCR5* gene, and to study their impact HIV-1-related outcomes in Zambians.

# MATERIALS AND METHODS

## Study Cohort

Zambia Emory HIV Research Project (ZEHRP) is a non-governmental organization established in 1994. The principal objectives of ZEHRP are to uncover factors related to HIV-1 risk and transmission and have a better understanding of how to prevent the spread of HIV-1 among HIV-1 infected African adults [28]. Since 1997, over 10,000 couples living in Lusaka, Zambia have been screened for HIV-1 serostatus using two rapid antibody tests and ELISA confirmation. Over 1,000 HIV-1 discordant couples have entered a prospective study. During quarterly follow-up visits, they were provided with free condoms, outpatient medical care, and appropriate counseling. Medical histories were obtained through questionnaires conducted by trained medical officers. Physical examinations following the medical interview were conducted by a physician or nurse. Hematocrit, complete blood cell counts and sedimentation rates were performed manually by trained laboratory technicians. All the clinical information is maintained by the local hospital, including age, gender, viral load, HIV-1 status, transmission type, occurrence of other sexually transmitted diseases [28].

Our research is part of ZEHRP, which broadly focuses on genetic determinants of HIV-1 infection and disease progression. We included 567 HIV-1 serodiscordant couples who are regarded as high risk of HIV-1 transmission based on their behavior and clinical exams in our study. Specifically, one partner is seropositive (Index partner); the other is

seronegative (Exposed seronegative partner). After several months follow up, some of the exposed partner became seropositive (Seroconverters). We define the seroconversion /transmission date as the mid-point between the first HIV seropositive test day and the last seronegative day. Each participant who is still seronegative at the last counseling day is designated an exposed seronegative partner. The time from the enrollment day to the last counseling day is defined as the follow-up time. We exclude the exposed seronegative partners who had a follow-up time less than 9 months in our cohort. In 2011 "workshop on HIV-exposed and resistant", it was recommended that at least 6 months of follow up time should be used to define exposed seronegatives [29]. Our definition here is stricter than the general accepted ones. All the participants were HIV-1 antiretroviral therapy naïve during the follow up time covered in this work [30].

**Study data**

Human subjects. The study was reviewed and approved by the University of Alabama at Birmingham Institutional Review Board (Appendix). The procedures for data safety included encoding the study subjects with clinical and laboratory identification numbers so that no personal information can be revealed.

# ASSOCIATION OF CHEMOKINE RECEPTOR GENE (*CCR2-CCR5*) HAPLOTYPES WITH ACQUISITION AND CONTROL OF HIV-1 INFECTION IN ZAMBIANS

by

RAKHI MALHOTRA*, LIANGYUAN HU*, WEI SONG, ILENE BRILL, JOSEPH MULENGA, SUSAN ALLEN, ERIC HUNTER, SADEEP SHRESTHA, JIANMING TANG, AND RICHARD A. KASLOW

* Equal authorship.

*Retrovirology*

## ABSTRACT

*Background:* Polymorphism in chemokine (C-C motif) receptors 2 and 5 genes (*CCR2* and *CCR5*) have been associated with HIV-1 infection and disease progression. We investigated the impact of *CCR2*-*CCR5* haplotypes on HIV-1 viral load (VL) and heterosexual transmission in an African cohort. Between 1995 and 2006, cohabiting Zambian couples discordant for HIV-1 (index seropositive and HIV-1 exposed seronegative [30]) were monitored prospectively to determine the role of host genetic factors in HIV-1 control and heterosexual transmission. Genotyping for eight *CCR2* and *CCR5* variants resolved nine previously recognized haplotypes. By regression and survival analytic techniques, controlling for non-genetic factors, we estimated the effects of these haplotypic variants on a) index partner VL, b) seroconverter VL, c) HIV-1 transmission by index partners, d) HIV-1 acquisition by HESN partners.

*Results:* Among 567 couples, 240 virologically linked transmission events had occurred through 2006. HHF*2 homozygosity was associated with significantly lower VL in seroconverters (mean beta=-0.58, $\log_{10}$ *P*=0.027) and the HHD/HHE diplotype was associated with significantly higher VL in the seroconverters (mean beta=0.54, $\log_{10}$ *P*=0.014) adjusted for age and gender in multivariable model. HHD/HHE was associated with more rapid acquisition of infection by the HESNs (HR=2.0, 95% CI=1.20-3.43, *P*=0.008), after adjustments for index partner VL and the presence of genital ulcer or inflammation in either partner in Cox multivariable models. The HHD/HHE effect was stronger in exposed females (HR=2.1, 95% CI=1.14-3.95, *P*=0.018).

C*onclusions*: Among Zambian discordant couples, HIV-1 coreceptor gene haplotypes and diplotypes appear to modulate HIV-1 VL in seroconverters and alter the

rate of HIV-1 acquisition by HESNs. These associations replicate or resemble findings reported in other African and European populations.

## BACKGROUND

Sub-Saharan Africa is home to about 10% of the world's population but bears nearly 64% of all HIV-1 infections [31], with most HIV-1 transmission occurring heterosexually. In Zambia, about one in five cohabiting couples involves an HIV-1 seropositive (index) and a seronegative (exposed) partner; these serodiscordant couples are at high risk of heterosexual transmission, with an estimated rate of eight transmission events per 100 person-years of follow-up [32].

The rate of within-couple heterosexual HIV-1 transmission is highly variable, and a number of viral, host and environmental factors may modify transmission (infectiousness), acquisition (susceptibility) or both [33]. Donor HIV-1 viral load (VL), age, sex, history of sexually transmitted infection (STI), unprotected sex, and possible HIV-1 subtype are among the major factors implicated [34, 35]. In southern Africa, unusual biological features of the predominant C subtype of HIV-1 [35] and absence of the human CC chemokine receptor 5 gene (*CCR5*) 32-bp deletion (Δ32) as a resistance factor may contribute to relatively high transmission rate.

The recognition that Caucasians who are homozygous for CCR5-Δ32 are highly resistant to HIV-1 infection was a landmark finding in research on HIV-1 transmission

[6, 8, 36, 37]. It stimulated a concerted effort to elucidate the impact of other genetic variations in *CCR5* and the adjacent gene *CCR2* on HIV-1 transmission and disease progression [16, 38, 39]. Research on the association of these variants with transmission has been largely cross-sectional or retrospective; the few prospective studies have focused on vertical (mother-to-child) transmission [40] and on HIV-1 exposed seronegatives (HESNs), in part because of the difficulty in enrolling and following HIV-1-discordant couples.

In Zambia, thousands of cohabiting and HIV-1 discordant couples have been offered voluntary counseling and testing (VCT) services since 1995 [32], and some of them have been followed for more than 10 years [41]. Despite counseling and behavioral interventions, the rate of HIV-1 transmission among these couples has remained high [42]. This circumstance permitted us to investigate the association of polymorphism in *CCR2* and *CCR5* with heterosexual transmission of phylogenetically related [43] HIV-1 within discordant partners.

The major published studies [9, 15, 16, 23, 40, 44] examining the effects of *CCR2* and *CCR5* SNPs/haplotypes/diplotypes on HIV-1 infection or disease progression have shown a wide spectrum of effects in various populations (See Additional File1; Table 2.5). We systematically tested hypotheses on these as well as other markers that occurred frequently enough in our population to permit meaningful inferences, especially in confirmation of earlier findings.

## RESULTS

### General characteristics of Zambian couples with linked HIV-1 viruses

During the study period 567 couples were eligible for analysis. Linked transmission occurred in 240 of the 567; (Table 2.1). nearly all (>95%) of the HIV-1 sequences from these transmission pairs corresponded to viral subtype C (HIV-1C) [43]. Male-to-female transmission accounted for nearly three-fifths of the incident infections (Table 2.1). Younger age of exposed women and, to a lesser extent, exposed men was associated with seroconversion. Certain non-genetic characteristics of the partners were also independently associated with increased transmission: genital ulcers or genital inflammation in any partner (HR=3.62, 95% CI: 2.65-4.93, $P<0.0001$) and high VL in the index partner (HR=1.59, 95% CI: 1.32-1.91, $P<0.0001$). These factors were retained in subsequent models that tested the impact of genetic markers.

### Distribution of *CCR2-CCR5* haplotypes in Zambian couples

Eight *CCR2-CCR5* haplotypes were observed in the frequency distribution shown in Table 2.2. Nearly 50% of all haplotypes were HHA or HHF*2. Haplotype HHB was rarely seen, and the Δ32-containing haplotype HHG*2 was not observed at all. The most common genotypes (diplotypes) were HHA/HHF*2, HHA/HHD, HHA/HHA, HHA/HHE, HHD/HHF*2, and HHE/HHF*2 (See Additional File 2; Table 2.6). The overall distribution of *CCR2-CCR5* haplotypes did not conform to HWE (Table 2.2). After stratification of the cohort into transmission and nontransmission index partners,

seroconverters, and exposed uninfected partners, the haplotype distribution deviated significantly from HWE in all three seropositive groups but not in the HESNs.

*CCR2-CCR5* determinants of VL

Although HHA and HHC had previously shown protective effects in the form of associations with lower VL in a mixed population[45], we did not observe such an effect on VL in Zambians with either haplotype overall or with any specific diplotypes containing either of them.

In prior studies HHF*2 has shown somewhat inconsistent associations with VL and disease control [9, 15, 16, 44, 46, 47]. In our Zambian study population HHF*2 showed a weak association with lower VL in both index partners (b=-0.21, $\log_{10}$ $P$=0.024*)* and seroconverters (b=-0.10, $\log_{10}$ $P$=0.089). When the index partners and seroconverters were stratified by HHF*2 genotype, a stronger association in the latter group was largely attributable to HHF*2 homozygosity (b=-0.70, $\log_{10}$ $P$=0.007) (Table 2.3).

Both HHD and HHE have been associated with higher VL in several studies [16, 23, 40, 48, 49]. In our Zambian cohort, dominant models including each haplotype plus non-genetic factors analyzed by GLM indicated that HHD was associated with higher VL (β=0.24, $\log_{10}$ $P$=0.021) in the seroconverters, but a modest effect in the opposite direction was observed in index partners. HHE showed a trend toward association with higher VL in index partners and a similar non-significant association with higher VL in

seroconverters adjusted for age and gender (Table 2.3). Because this pattern of association could be explained by combinations of haplotypes carried, we explored the effect of diplotypes further. Among all diplotypes of frequency >0.05, HHD/HHE showed the strongest association with higher VL ($\beta$=0.49, $\log_{10}$ $P$=0.02)

We next constructed a multivariable model with all the haplotypes and diplotypes that showed a trend toward association ($\log_{10}$ $P$<0.10) with higher or lower VL in either index partners or seroconverters to test their independent influences on VL (Table 2.3). In this model, by including uninformative diplotypes (HHD/HHF*2 and HHE/HHF*2) in the reference group, each diplotype implicated could be tested independently of the others. In the index partners, HHE/X shows a strong association with higher VL. In seroconverters the HHD/HHE and HHF*2/HHF*2 diplotypes remained significantly and independently associated with VL after controlling for individual haplotype effects (Table 2.3).

*CCR2-CCR5* determinants of transmission from index partners and of seroconversion in HESNs

The few studies that have attempted to assess the role of the receptor polymorphism in transmission and susceptibility have shown rather diverse associations of common *CCR5* haplotypes without any discernible pattern (See Additional File 1, Table 2.5). No SNP or haplotype carried by Zambian index partners was significantly associated with transmission (data not shown). In the survival analysis, HESNs with the HHD/HHE diplotypes showed significantly more rapid seroconversion than HESNs with

other haplotypes (Table 2.4 and Figure 2.1a) after adjustments for index partner VL and the presence of genital ulcer or inflammation in either partner. Although HHF*2 did not show statistically significant association with faster HIV-1 acquisition, we assigned it to a separate stratum in the Kaplan-Meier plot because aggregating it in the reference group would have given the appearance of a weaker HHD/HHE effect (Multivariable Cox model HR=2.0, 95% CI=1.20-3.43, *P*=0.008). Stratification by gender revealed a stronger impact of HHD/HHE on HESN women than men (Table 2.4 and Figure 2.1b) (Multivariable Cox model HR=2.1, 95% CI=1.14-3.95, *P*=0.018).

## DISCUSSION

Many investigations into genetic determinants of HIV/AIDS have evaluated the effects of Δ32, selected SNPs, and haplotypes across *CCR2*-*CCR5* on disease progression in a variety of infected populations. Studies of these markers as determinants of acquisition have usually been conducted in pairs of mothers and infants or in exposed men of European ancestry whose male sexual contacts are largely unknown [9, 40, 49-51]. Our relatively large prospective study of heterosexual discordant African couples has produced further evidence for involvement of variants in these genes in both control and occurrence of HIV-1 infection.

HHE was associated with slightly higher VL than was seen with other haplotypes, a finding consistent with observations in a number of other studies in different ethnic groups and with various modes of transmission [9, 16, 52, 53]. Further confirmation of the effect of HHE highlights its potential impact on clinical HIV-1 disease control in

diverse populations in contrast to that of the protective Δ32 variant whose distribution is confined to individuals of European ancestry. We detected an association of homozygous HHF*2 (containing CCR2-64I, rs1799864) with lower VL in recent seroconverters but less certain effects of heterozygous HHF*2. This finding is consistent with previous reports [9, 15, 16, 44, 46]. Although an early meta-analysis persuasively documented modest protection by the 64I allele against progression of HIV-1 subtype B infection [46], results in subsequent studies have been less consistent—showing association with slow progression either among Europeans but not African-Americans [15, 44] or among African-Americans but not Europeans [47, 53]. For populations with subtype C infection, however, no previous study is available as a basis for comparison.

As for the influence of *CCR2-CCR5* alleles or haplotypes on transmission and acquisition of infection, the highly significant deviation of the distribution of haplotypes from HWE among the index but not the exposed partners was strong evidence of a selective effect, and the differential deviation of the seroconverters but not the persistently seronegatives corroborated the difference. Neither chance nor systematic selection of couples into the study cohort by their *CCR2-CCR5* profile unrelated to infection seems as plausible an explanation as the direct effect on acquisition of HIV-1 infection proposed here.

No *CCR2-CCR5* variant carried by index partners was associated with an appreciable difference in transmission—not even the diplotype HHD/HHE associated with a statistically significant higher mean VL. This relative deviation in level of viremia

was apparently not equivalent to the larger deviation conferred by index partner HLA-B*57, a genetic marker associated with a significantly lower transmission rate in this population [41]. Such differential impact of the different genetic markers may reflect a threshold effect by which a deviation of VL greater than a certain level overrides any genetic influence, but the number of subjects in our cohort was insufficient to assess that possibility.

We observed a trend toward an increased rate of acquisition among the exposed partners carrying HHF*2. In another African population (Cameroon) the frequency of CCR2 64I (rs1799864, HHF*2) was higher in the HIV-1 seropositives (most likely of mixed viral subtype) than in the seronegatives [54]. However, we remain skeptical about the importance of these findings for several reasons. First, the association and its significance in Zambians diminished in the multivariable analysis. Second, previous evidence for a role of HHF*2 in occurrence of infection is sparse, and there is no other report from a prospective study. Third, considerable uncertainty remains about the functional relevance of the HHF*2 polymorphism and *CCR2* itself to HIV-1 infection [55]. Further population studies alone are unlikely to clarify more precisely the true nature of this genetic contribution.

More rapid HIV-1 acquisition among exposed seronegatives occurred in association with the HHD/HHE diplotype, and the association was stronger in exposed women than men. An association with this diplotype has not been reported before, most likely because the single SNP allele that distinguishes HHD from other haplotypes is only

frequent enough in persons of African ancestry. The relatively higher frequency (7%) of HHD/HHE in our population than in Caucasians or other smaller groups of Africans may have facilitated detection of its effect. Associations with higher risk of mother-to-child transmission have been reported for HHD in Africans [56] and with homozygous HHD in African Americans [57]. HHE has also been reported to be detrimental for HIV infection as well as disease progression, but HHD/HHE has not been studied previously as a diplotype. Although our findings do not constitute exact replication of previous work, they appear to indicate consistent effects of the two haplotypes across populations with different viral subtypes.

The effects of HHD/HHE appeared stronger in male-to-female transmission. Differences in VL among the donor groups did not explain this difference according to direction of transmission. Nor did the difference arise from any obvious difference in age or sexual exposure of the two groups. For each subgroup stratified by gender, the number of seronegative subjects carrying these genotypes (diplotypes) was relatively small. Analysis based on larger samples will be necessary to reach a reliable conclusion about such gender-specific associations.

One feature of our study worth noting is the advantage of survival analysis of time to transmission/acquisition in detecting relationships that may be weaker in the cross-sectional or case-control approach often used to assess genetic influences on HIV-1 infection. Survival methods may be more sensitive in capturing time-dependent genetic effects on infection just as they have been in the analysis of disease progression.

We did not adjust statistically for the number of genetic polymorphism tested. Rather we have emphasized those nominally significant associations with *CCR2-CCR5* variants that have previously been implicated in HIV/AIDS and de-emphasized those whose involvement was less predictable from earlier studies. The previously documented HHE association with higher VL [9, 16, 52, 53] provided ample rationale for interpreting our results as confirmatory without treating all haplotypes as equally likely to be involved. The impact of HHD/HHE on seroconversion was predicted somewhat less directly by earlier work associating HHD with a higher frequency of neonatal infection [57]. An even more important reason why these relationships cannot be readily dismissed as chance findings is that they were observed in the context of significant deviations from HWE of the haplotype distributions in each of the seropositive groups but not the seronegative group.

A consistent effect of the frequent HHE with higher VL in subtype C HIV-1-infected Africans as well as subtype B-infected Europeans and a stronger effect of HHD/HHE could have further ramifications. Since the response to antiretroviral treatment in Europeans may be modified by (Δ32) [58-60] and perhaps by other receptor variants [61, 62], investigators in African settings should consider whether similar studies of *CCR2-CCR5* polymorphism might provide epidemiologically or clinically useful prognostic information.

## CONCLUSIONS

In summary, our analysis of *CCR2-CCR5* haplotypes consisting of common combinations of SNP alleles spanning those two genes has confirmed a previously reported association of haplotype HHF*2 with favorable response to HIV-1 infection; and our longitudinal analysis of seroconversion in HESN African heterosexual partners has detected probable contributions by the HHD/HHE diplotype to acquisition of infection [9, 16, 63]. Further insight into these relationships will be gained from studies of correlation between gene variation and gene function, as well as investigation of other representative and informative populations of infected and uninfected Africans.

## METHODS

### Study population

Our study population comprised HIV-1 serodiscordant, cohabiting heterosexual couples enrolled in the Zambia-Emory HIV Research Project between 1995 and 2006. The procedures for screening, recruitment, counseling, follow-up visits and laboratory testing have been described elsewhere [42, 64]. All couples whose HESN partner acquired virologically linked HIV-1 from the index partner during follow-up were included in this study. For closer comparability to the transmitters, nontransmission couples were selected from a large number based on self-reported behavioral or clinical measures of unprotected sex. Virologically linked HIV-1 transmission was defined as identity between viruses from index and seroconverting partners, according to phylogenetic analysis of sub-genomic sequences of *gag*, *env* (gp120 and gp41), and long terminal repeat regions [43, 64]. Participant characteristics have previously been thoroughly examined as potential risk factors for transmission in this cohort [42, 43, 64,

65]. Risk factors considered here include index partner (donor) viral load (VL), age of each partner, and genital ulceration/inflammation in each partner. The study population consisted of 567 couples with: a) adequate data and biologic material for both partners, b) observation of nontransmission couples for at least nine months, c) intra-couple virologic linkage when transmission occurred, and d) none of the partners on anti-retroviral treatment.

Non-genetic factors

VL was quantified as the number of HIV-1 RNA copies per ml of plasma using Roche Amplicor 1.0 assay (Roche diagnostic Systems Inc., Branchburg, NJ) in a laboratory certified by the virology quality assurance program of the AIDS Clinical Trials Group (ACTG). The lower detection limit was 400 copies/mL of plasma. For this work, VL was transformed to $\log_{10}$ and treated as a continuous variable. Previous analyses [64] indicated that index partners with a medium number of HIV-1 RNA copies/mL ($10^4$-$10^5$, $\log_{10} = 4$-5) or a high number of copies/mL ($>10^5$, $\log_{10} >5$) were more likely to transmit the virus than those with a low number ($<10^4$, $\log_{10} <4$).

Genotyping

Genomic DNA was extracted from whole blood and buffy coats using the QIAamp blood kit and protocols recommended by the manufacturer (QIAGEN Inc., Valencia, CA). PCR-based typing differentiated the dimorphic variants at eight sites—one in *CCR2* (the SNP encoding V64I—rs1799864) and seven in *CCR5* [six SNPs in or adjacent to the cis-regulatory or promoter region (A29G—rs2856758, G303A—

rs1799987, T627C—rs1799988, C630T—rs41469351, A676G—rs1800023 and C927T—rs1800024)] and the 32-bp deletion (Δ32—rs333). *CCR5* haplotypes were typed by a combination of two methods: a PCR typing scheme and a TaqMan SNP typing scheme. The PCR typing scheme used 12 combinations of sequence-specific primers (SSP) plus four additional SSP reactions in conjunction with T627C-specific primers to define the A29G variant as described for previous work [9, 15, 16, 23, 40, 44]. Combination of variants at the eight sites form nine relatively frequent *CCR2-CCR5* haplotypes (HHA-HHE, HHF*1, HHF*2, HHG*1 and HHG*2) according to the nomenclature of the Tri-Service HIV-1 Natural History Study (TSS) [66]. HHF*2 is the only haplotype carrying the V64I mutation. A TaqMan genotyping assay was used to confirm the PCR-based SNP typing and assign *CCR5* haplotypes for 126 individuals. TaqMan assays were performed using customized TaqMan probes for 7 SNP sites; SNP alleles were assigned after real-time PCR using the ABI 7500 Fast System (Applied Biosystems) according to procedures recommended by the manufacturer.

Statistical analysis

Non-genetic factors (VL, age, gender, genital ulcer, genital inflammation, circumcision, and presence of sperm) were compared between seroconverting and non-seroconverting exposed partners using $\chi^2$ and t-tests. Hardy-Weinberg equilibrium (HWE) for each SNP and CCR haplotype distribution was assessed using SAS Genetics (see below). HWE was calculated for the entire cohort and for four separate partner groups: transmission index, nontransmission index, seroconverting, and exposed uninfected partners. Associations of frequent haplotypes/genotypes with HIV-1 VL

among the index partners and seroconverters were tested using general linear model (GLM) statistics with adjustment for age and gender.

For analysis of time-to-infection (transmission and acquisition), follow-up time for each couple was measured from the date of their enrollment into the cohort to 1) the date of HIV-1 infection (first seropositive visit) of the initially uninfected exposed partner or 2) the most recent seronegative visit prior to administrative censoring date (December 31, 2006). Time-to-infection was displayed in Kaplan-Meier plots, and comparisons between genetically distinctive groups were evaluated with Wilcoxon and log-rank tests. These plots illustrate differences in transmission associated with specific genetic markers; they do not reflect transmission rates in the entire prospectively observed discordant couple population. The overall annual HIV-1 seroincidence (7-8/100 PY) represents a one-half to two-thirds reduction in transmission following joint testing and counseling.

Statistical analysis of genetic variants of *CCR2* and *CCR5* consisted of testing hypotheses derived from earlier work on acquisition or progression of infection (See Additional File 1; Table 2.5) followed by systematic search for novel associations in our study population. Multivariable Cox proportional hazards models were used to control for non-genetic covariates. We estimated the hazard ratios (HR), its 95% confidence interval (CI), and the corresponding two-sided *P*-values. For hypotheses on genetic markers consistent with previously reported associations, statistical testing was performed without correction for multiple comparisons. All statistical analyses were done using SAS® 9.2 including SAS/Genetics™ (SAS Institute Inc., Cary, NC).

Reference:

1. UNAIDS: UNAIDS/WHO "AIDS Epidemic Update: December 2006". In *Book UNAIDS/WHO "AIDS Epidemic Update: December 2006"* (Editor ed.^eds.). City; 2006.

2. Dunkle KL, Stephenson R, Karita E, Chomba E, Kayitenkore K, Vwalika C, Greenberg L, Allen S: New heterosexually transmitted HIV infections in married or cohabiting couples in urban Zambia and Rwanda: an analysis of survey and clinical data. *Lancet* 2008, 371:2183-2191.

3. Borrow P, Shattock RJ, Vyakarnam A: Innate immunity against HIV: a priority target for HIV prevention research. *Retrovirology* 2010, 7:84.

4. Royce RA, Sena A, Cates W, Jr., Cohen MS: Sexual transmission of HIV. *N Engl J Med* 1997, 336:1072-1078.

5. Ping LH, Nelson JA, Hoffman IF, Schock J, Lamers SL, Goodman M, Vernazza P, Kazembe P, Maida M, Zimba D, et al: Characterization of V3 sequence heterogeneity in subtype C human immunodeficiency virus type 1 isolates from Malawi: underrepresentation of X4 variants. *J Virol* 1999, 73:6271-6281.

6. Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, MacDonald ME, Stuhlmann H, Koup RA, Landau NR: Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 1996, 86:367-377.

7. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, Saragosti S, Lapoumeroulie C, Cognaux J, Forceille C, et al: Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 1996, 382:722-725.

8. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, Allikmets R, Goedert JJ, Buchbinder SP, Vittinghoff E, Gomperts E, et al: 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* 1996, 273:1856-1862.

9. Zimmerman PA, Buckler-White A, Alkhatib G, Spalding T, Kubofcik J, Combadiere C, Weissman D, Cohen O, Rubbert A, Lam G, et al: Inherited resistance to HIV-1 conferred by an inactivating mutation in CC chemokine receptor 5: studies in populations with contrasting clinical phenotypes, defined racial background, and quantified risk. *Mol Med* 1997, 3:23-36.

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

11. Tang J, Shelton B, Makhatadze NJ, Zhang Y, Schaen M, Louie LG, Goedert JJ, Seaberg EC, Margolick JB, Mellors J, Kaslow RA: 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* 2002, 76:662-672.

12. Winkler CA, Hendel H, Carrington M, Smith MW, Nelson GW, O'Brien S J, Phair J, Vlahov D, Jacobson LP, Rappaport J, et al: Dominant effects of CCR2-CCR5 haplotypes in HIV-1 disease progression. *J Acquir Immune Defic Syndr* 2004, 37:1534-1538.

13. Ometto L, Bertorelle R, Mainardi M, Zanchetta M, Tognazzo S, Rampon O, Ruga E, Chieco-Bianchi L, De Rossi A: Polymorphisms in the CCR5 promoter region influence disease progression in perinatally human immunodeficiency virus type 1-infected children. *J Infect Dis* 2001, 183:814-818.

14. Tang J, Shao W, Yoo YJ, Brill I, Mulenga J, Allen S, Hunter E, Kaslow RA: Human leukocyte antigen class I genotypes in relation to heterosexual HIV type 1 transmission within discordant couples. *J Immunol* 2008, 181:2626-2635.

15. Allen S, Meinzen-Derr J, Kautzman M, Zulu I, Trask S, Fideli U, Musonda R, Kasolo F, Gao F, Haworth A: Sexual behavior of HIV discordant couples after HIV counseling and testing. *Aids* 2003, 17:733-740.

16. Trask SA, Derdeyn CA, Fideli U, Chen Y, Meleth S, Kasolo F, Musonda R, Hunter E, Gao F, Allen S, Hahn BH: Molecular epidemiology of human immunodeficiency virus type 1 transmission in a heterosexual cohort of discordant couples in Zambia. *J Virol* 2002, 76:397-405.

17. Gonzalez E, Bamshad M, Sato N, Mummidi S, Dhanda R, Catano G, Cabrera S, McBride M, Cao XH, Merrill G, et al: Race-specific HIV-1 disease-modifying effects associated with CCR5 haplotypes. *Proc Natl Acad Sci U S A* 1999, 96:12004-12009.

18. Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, Nibbs RJ, Freedman BI, Quinones MP, Bamshad MJ, et al: The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* 2005, 307:1434-1440.

19. Smith MW, Dean M, Carrington M, Winkler C, Huttley GA, Lomb DA, Goedert JJ, O'Brien TR, Jacobson LP, Kaslow R, et al: Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC), ALIVE Study. *Science* 1997, 277:959-965.

20. Easterbrook PJ, Rostron T, Ives N, Troop M, Gazzard BG, Rowland-Jones SL: Chemokine receptor polymorphisms and human immunodeficiency virus disease progression. *J Infect Dis* 1999, 180:1096-1105.

21. Kageyama S, Mimaya J, Yamada K, Kurimura T, Shiraki K: Polymorphism of CCR5 affecting HIV disease progression in the Japanese population. *AIDS Res Hum Retroviruses* 2001, 17:991-995.

22. Ioannidis JP, Rosenberg PS, Goedert JJ, Ashton LJ, Benfield TL, Buchbinder SP, Coutinho RA, Eugen-Olsen J, Gallart T, Katzenstein TL, et al: Effects of CCR5-

Delta32, CCR2-64I, and SDF-1 3'A alleles on HIV-1 disease progression: An international meta-analysis of individual-patient data. *Ann Intern Med* 2001, 135:782-795.

23. Schinkel J, Langendam MW, Coutinho RA, Krol A, Brouwer M, Schuitemaker H: No evidence for an effect of the CCR5 delta32/+ and CCR2b 64I/+ mutations on human immunodeficiency virus (HIV)-1 disease progression among HIV-1-infected injecting drug users. *J Infect Dis* 1999, 179:825-831.

24. Dorak MT, Tang J, Penman-Aguilar A, Westfall AO, Zulu I, Lobashevsky ES, Kancheya NG, Schaen MM, Allen SA, Kaslow RA: Transmission of HIV-1 and HLA-B allele-sharing within serodiscordant heterosexual Zambian couples. *Lancet* 2004, 363:2137-2139.

25. Nguyen L, Li M, Chaowanachan T, Hu DJ, Vanichseni S, Mock PA, van Griensven F, Martin M, Sangkum U, Choopanya K, et al: CCR5 promoter human haplogroups associated with HIV-1 disease progression in Thai injection drug users. *Aids* 2004, 18:1327-1333.

26. An P, Martin MP, Nelson GW, Carrington M, Smith MW, Gong K, Vlahov D, O'Brien SJ, Winkler CA: Influence of CCR5 promoter haplotypes on AIDS progression in African-Americans. *Aids* 2000, 14:2117-2122.

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

28. Martin MP, Dean M, Smith MW, Winkler C, Gerrard B, Michael NL, Lee B, Doms RW, Margolick J, Buchbinder S, et al: Genetic acceleration of AIDS progression by a promoter variant of CCR5. *Science* 1998, 282:1907-1911.

29. Mummidi S, Ahuja SS, Gonzalez E, Anderson SA, Santiago EN, Stephan KT, Craig FE, O'Connell P, Tryon V, Clark RA, et al: Genealogy of the CCR5 locus and chemokine system gene variants associated with altered rates of HIV-1 disease progression. *Nat Med* 1998, 4:786-793.

30. Ma L, Marmor M, Zhong P, Ewane L, Su B, Nyambi P: Distribution of CCR2-64I and SDF1-3'A alleles and HIV status in 7 ethnic populations of Cameroon. *J Acquir Immune Defic Syndr* 2005, 40:89-95.

31. Arenzana-Seisdedos F, Parmentier M: Genetics of resistance to HIV infection: Role of co-receptors and co-receptor ligands. *Semin Immunol* 2006, 18:387-403.

32. Joubert BR, Franceschini N, Mwapasa V, North KE, Meshnick SR: Regulation of CCR5 expression in human placenta: insights from a study of mother-to-child transmission of HIV in Malawi. *PLoS One* 2010, 5:e9212.

33. Kostrikis LG, Neumann AU, Thomson B, Korber BT, McHardy P, Karanicolas R, Deutsch L, Huang Y, Lew JF, McIntosh K, et al: A polymorphism in the regulatory region of the CC-chemokine receptor 5 gene influences perinatal transmission of human immunodeficiency virus type 1 to African-American infants. *J Virol* 1999, 73:10264-10271.

34. Valdez H, Purvis SF, Lederman MM, Fillingame M, Zimmerman PA: Association of the CCR5delta32 mutation with improved response to antiretroviral therapy. *Jama* 1999, 282:734.

35. Kasten S, Goldwich A, Schmitt M, Rascu A, Grunke M, Dechant C, Kalden JR, Harrer T: Positive influence of the Delta32CCR5 allele on response to highly active antiretroviral therapy (HAART) in HIV-1 infected patients. *Eur J Med Res* 2000, 5:323-328.

36. Barroga CF, Raskino C, Fangon MC, Palumbo PE, Baker CJ, Englund JA, Spector SA: The CCR5Delta32 allele slows disease progression of human immunodeficiency virus-1-infected children receiving antiretroviral treatment. *J Infect Dis* 2000, 182:413-419.

37. Wang YM, Wang B, Dyer WB, Lachireddy K, Peng NK, Saksena NK: Evidence for predominance of CCR5-using HIV-1 strains during highly active antiretroviral therapy. *Curr HIV Res* 2007, 5:221-234.

38. Puissant B, Roubinet F, Massip P, Sandres-Saune K, Apoil PA, Abbal M, Pasquier C, Izopet J, Blancher A: Analysis of CCR5, CCR2, CX3CR1, and SDF1 polymorphisms in HIV-positive treated patients: impact on response to HAART and on peripheral T lymphocyte counts. *AIDS Res Hum Retroviruses* 2006, 22:153-162.

39. Catano G, Chykarenko ZA, Mangano A, Anaya JM, He W, Smith A, Bologna R, Sen L, Clark RA, Lloyd A, et al: Concordance of CCR5 genotypes that influence

cell-mediated immunity and HIV-1 disease progression rates. *J Infect Dis* 2011, 203:263-272.

40. Fideli US, Allen SA, Musonda R, Trask S, Hahn BH, Weiss H, Mulenga J, Kasolo F, Vermund SH, Aldrovandi GM: Virologic and immunologic determinants of heterosexual transmission of human immunodeficiency virus type 1 in Africa. *AIDS Res Hum Retroviruses* 2001, 17:901-910.

41. Tang J, Tang S, Lobashevsky E, Zulu I, Aldrovandi G, Allen S, Kaslow RA: HLA allele sharing and HIV type 1 viremia in seroconverting Zambians with known transmitting partners. *AIDS Res Hum Retroviruses* 2004, 20:19-25.

42. Mangano A, Gonzalez E, Dhanda R, Catano G, Bamshad M, Bock A, Duggirala R, Williams K, Mummidi S, Clark RA, et al: Concordance between the CC chemokine receptor 5 genetic determinants that alter risks of transmission and disease progression in children exposed perinatally to human immunodeficiency virus. *J Infect Dis* 2001, 183:1574-1585.

43. Pedersen BR, Kamwendo D, Blood M, Mwapasa V, Molyneux M, North K, Rogerson SJ, Zimmerman P, Meshnick SR: CCR5 haplotypes and mother-to-child HIV transmission in Malawi. *PLoS One* 2007, 2:e838.

44. John GC, Bird T, Overbaugh J, Nduati R, Mbori-Ngacha D, Rostron T, Dong T, Kostrikis L, Richardson B, Rowland-Jones SL: CCR5 promoter polymorphisms in a Kenyan perinatal human immunodeficiency virus type 1 cohort: association with increased 2-year maternal mortality. *J Infect Dis* 2001, 184:89-92.

45. Ramaley PA, French N, Kaleebu P, Gilks C, Whitworth J, Hill AV: HIV in Africa (Communication arising): chemokine-receptor genes and AIDS risk. *Nature* 2002, 417:140.

FIGURE LEGENDS

Figure 2.1. Association of two *CCR2-CCR5* diplotype (HHD/HHE and HHF*2) with time to HIV-1 acquisition among initially seronegative partners of HIV-1 discordant Zambian couples. Analysis was based on 567 initially seronegative HESNs (panel a) or 295 seronegative female HESNs only (panel b). Vertical lines on each Kaplan-Meier curve represent subjects censored at the last follow up visit (before December 2006).

**Additional Files**

Additional file 1

Title: Table 2.5. Studies of associations between polymorphisms in *CCR2* and *CCR5* and acquisition or progression of HIV-1 infection.

Description: Summary of the recent publications on *CCR2-CCR5* haplotypes and association with HIV-1 acquisition or disease progression.

Additional file 2:

Title: Table 2.6. *CCR2-CCR5* haplotypes and diplotypes as observed in HIV-1 discordant Zambian couples.

Description: Frequency of *CCR2-CCR5* haplotypes and common diplotypes in overall Zambia cohort and subgroups. Rare diplotypes with count less than 12 in overall cohort are not shown.

TABLES

Table 2.1. Demographic, epidemiologic and virologic characteristics of the HIV-1 nontransmission and transmission serodiscordant Zambian couples

| | Nontransmission couples | Transmission couples | P |
|---|---|---|---|
| Characteristic | | | |
| Number of couples | 327 | 240 | |
| Male/Female (index partner) | 148/179 | 147/93 | 0.0002 |
| Age of partners (yrs) | | | |
| Index | 31.5 ± 7.9 | 30.6 ± 7.8 | 0.170 |
| Exposed | 32.0 ± 8.3 | 28.6 ± 7.3 | <0.0001 |
| Follow-up time (median [IQR], months) | 31.5 [17.0-56.1] | 17.7 [8.8-36.2] | <0.0001 |
| Male circumcised | | | |
| Index | 8.8% | 9.0% | 0.970 |
| Exposed | 19.5% | 9.9% | 0.053 |
| Genital ulcers | | | |
| Index | 12.8% | 36.6% | <0.0001 |
| Exposed | 4.1% | 26.3% | <0.0001 |
| Genital inflammation [a] | | | |
| Index | 10.3% | 26.3% | <0.0001 |
| Exposed | 6.6% | 29.1% | <0.0001 |
| Any sexually transmitted disease | | | |
| Index | 20.7% | 50.2% | <0.0001 |
| Exposed | 10.0% | 46.0% | <0.0001 |
| HIV-1 RNA level ($log_{10}$) in index partner | 4.47 ± 0.90 | 4.96± 0.70 | < 0.0001 |
| HIV-1 RNA level ($log_{10}$) in seroconverted partner | | 4.50 ± 0.80 | NA |

[a] In the 3-6 months before HIV-1 transmission (transmission couples) or latest follow-up visit (nontransmission couples).

Table 2.2. Frequencies of *CCR2-CCR5* polymorphisms among HIV-1 serodiscordant couples, index partners, and HIV-1 exposed seronegative partners

| | All | Index partners | HESN[a] partners | | |
|---|---|---|---|---|---|
| | | | All | Seroconverters | Uninfected |
| | (N=1134) | (N=567) | (N=567) | (N=240) | (N=327) |
| Haplotype | N (%) | N (%) | N (%) | N (%) | N (%) |
| HHA | 601 (26.5) | 305 (26.9) | 296 (26.1) | 133 (27.7) | 163 (24.9) |
| HHB | 44 (1.9) | 25 (2.2) | 19 (1.7) | 8 (1.7) | 11 (1.7) |
| HHC | 178 (7.9) | 69 (6.1) | 109 (9.6) | 42 (8.8) | 67 (10.2) |
| HHD | 370 (16.3) | 190 (16.8) | 180 (15.9) | 75 (15.6) | 105 (16.1) |
| HHE | 310 (13.7) | 138 (12.2) | 172 (15.2) | 61 (12.7) | 111 (17.0) |
| HHF*1 | 128 (5.6) | 76 (6.7) | 52 (4.6) | 24 (5.0) | 28 (4.3) |
| HHF*2 | 480 (21.2) | 259 (22.8) | 221 (19.5) | 99 (20.6) | 122 (18.7) |
| HHG*1 | 157 (6.9) | 72 (6.4) | 85 (7.5) | 38 (7.9) | 47 (7.2) |
| HWE:P[b] | 0.0001 | 0.0001 | 0.267 | 0.024 | 0.160 |

[a] HESN = HIV-1 exposed seronegative

[b] P values for tests of Hardy-Weinberg equilibrium in each of the patient groups.

Table 2.3. The impact of *CCR2-CCR5* haplotypes on HIV-1 viral load in Zambian index partners and seroconverters

| | Index partners (N = 567) | | | Recent seroconverters (N = 240) | | |
|---|---|---|---|---|---|---|
| Viral Load Table for CCR5 haplotype/diplotype | | | | | | |
| Haplotype/diplotype | N | β ± SE[a] | P [b] | N | β ± SE[a] | P [b] |
| HHF*2 | 232 | -0.21±0.09 | 0.024 | 89 | -0.10±0.58 | 0.089 |
| HHF*2/HHF*2 | 27 | -0.08±0.16 | 0.625 | 10 | -0.70±0.26 | 0.007 |
| HHD (All) | 164 | -0.12±0.06 | 0.052 | 69 | 0.24±0.10 | 0.021 |
| HHE (All) | 131 | 0.13±0.08 | 0.096 | 61 | 0.12±0.12 | 0.339 |
| HHD/HHE | 18 | -0.22±0.20 | 0.270 | 16 | 0.49±0.21 | 0.020 |
| HHD/X (No HHE) | 146 | -0.11±0.08 | 0.157 | 53 | 0.14±0.13 | 0.284 |
| HHE/X (No HHD) | 113 | 0.19±0.09 | 0.026 | 45 | -0.06±0.13 | 0.673 |
| Multivariable Model for Interaction[c] | | | | | | |
| Haplotype/diplotype | N | β ± SE[a] | P [b] | N | β ± SE[a] | P [b] |
| HHD/X (No HHE or HHF*2) | 109 | -0.16±0.09 | 0.560 | 37 | 0.24±0.16 | 0.118 |
| HHE/X (No HHD or HHF*2) | 69 | 0.36±0.11 | 0.002 | 31 | 0.10±0.16 | 0.556 |
| HHD/HHE | 18 | -0.16±0.20 | 0.444 | 16 | 0.54±0.22 | 0.014 |
| HHF*2/HHF*2 | 27 | -0.02±0.16 | 0.918 | 10 | -0.58±0.26 | 0.027 |
| HHF2/X (No HHD, HHE or HHF*2) | 124 | 0.05±0.09 | 0.602 | 49 | 0.10±0.14 | 0.496 |

[a] SE, standard error of the b estimate according to linear regression models.

[b] P value adjusted for the sex and age at VL for all the individuals.

[c] Individuals with HHD/HHF*2, HHE/HHF*2 diplotypes are in the reference group in the multivariable model for interaction.

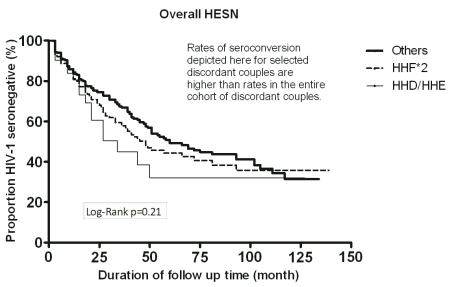
Table 2.4. Proportional hazards analysis of the effect of *CCR2-CCR5* haplotype or diplotype on HIV-1 acquisition.

| | Overall (567 couples) | | | | Male-to-Female (295 couples) | | | | Female-to-Male (272 couples) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cox model for individual CCR2-CCR5 haplotype or diplotype. | | | | | | | | | | | | |
| Haplotype/ diplotype | N | HR | 95% CI | P[a] | N | HR | 95% CI | P[a] | N | HR | 95% CI | P[a] |
| HHF*2 | 203 | 1.1 | 0.85-1.46 | 0.417 | 104 | 1.1 | 0.77-1.56 | 0.61 | 99 | 1.2 | 0.75-1.75 | 0.531 |
| HHD | 167 | 1.0 | 0.74-1.32 | 0.983 | 94 | 0.9 | 0.59-1.23 | 0.385 | 73 | 1.2 | 0.75-1.92 | 0.442 |
| HHE | 158 | 1.1 | 0.82-1.51 | 0.495 | 85 | 1.3 | 0.86-1.87 | 0.234 | 73 | 0.9 | 0.57-1.53 | 0.781 |
| HHD/HHE | 31 | 1.9 | 1.14-3.16 | 0.015 | 19 | 2.0 | 1.08-3.57 | 0.028 | 12 | 1.7 | 0.60-4.66 | 0.321 |
| Multivariable model for CCR2-CCR5 HHD/HHE diplotype and HHF*2 haplotype. | | | | | | | | | | | | |
| Genetic factors | N | HR | 95% CI | P[a] | N | HR | 95% CI | P[a] | N | HR | 95% CI | P[a] |
| HHD/HHE | 31 | 2.0 | 1.20-3.43 | 0.008 | 19 | 2.1 | 1.14-3.95 | 0.018 | 12 | 1.8 | 0.64-5.08 | 0.267 |
| HHF*2 | 203 | 1.2 | 0.90-1.57 | 0.222 | 104 | 1.2 | 0.83-1.72 | 0.337 | 99 | 1.2 | 0.77-1.82 | 0.439 |
| Any genital ulcer or | 299 | 3.6 | 2.65-4.93 | <.0001 | 162 | 3.0 | 2.04-4.51 | <.0001 | 137 | 4.6 | 2.79-7.65 | <.0001 |
| Donor VL (per 1.0 $\log_{10}$ unit) | 523 | 1.6 | 1.32-1.91 | <.0001 | 263 | 1.3 | 1.03-1.74 | 0.028 | 260 | 1.8 | 1.35-2.47 | <.0001 |

[a] All *P* values are adjusted for genital ulcer/inflammation in either partner, and index partners' VL (per $\log_{10}$).

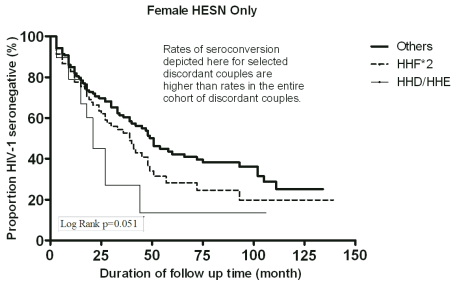
Figure 2.1

a.




| | | Number of partners at risk at various follow -up intervals (months) | | | | | |
|---|---|---|---|---|---|---|---|
| Groups | Events | 0 | 25 | 50 | 75 | 100 | 125 |
| HHF2 | 89 (44%) | 203 | 98 | 40 | 20 | 12 | 3 |
| HHD/HHE | 16 (52%) | 31 | 13 | 6 | 4 | 3 | 1 |
| Others | 135 (41%) | 333 | 178 | 81 | 49 | 29 | 9 |
| total | 240 (42%) | 567 | 289 | 127 | 73 | 44 | 13 |

b.




| | | Number of partners at risk at various follow -up intervals (months) | | | | | |
|---|---|---|---|---|---|---|---|
| Groups | Events | 0 | 25 | 50 | 75 | 100 | 125 |
| HHF2 | 54 (52%) | 104 | 44 | 14 | 8 | 5 | 2 |
| HHD/HHE | 12 (63%) | 19 | 5 | 2 | 1 | 1 | 0 |
| Others | 81 (47%) | 172 | 85 | 40 | 29 | 16 | 4 |
| total | 147 (50%) | 295 | 134 | 56 | 38 | 22 | 6 |

Supplementary Tables

Table 2.5: Studies of associations between polymorphisms in *CCR2* and *CCR5* and acquisition or progression of HIV-1 infection.

| Marker/Combination or Haplotype | Acquisition/ susceptibility | Ref† | Progression | Ref† |
|---|---|---|---|---|
| HHA | | | + | [9] |
| HHA/HHA | - | [23] | | |
| HHA/HHF*1 | - | [23] | | |
| HHB | - | [67] | | |
| HHC | - | | - | [9, 45] |
| | | | + | [9] |
| HHC/HHC, HHC/HHD, HHC/HHE, HHB/HHC | - | [23] | | |
| HHD (or 59356-T) | - | [68] | - | |
| HHD/HHE, HHD/HHG*1 | - | [23] | | |
| HHE | - | [66] | - | [40, 49] |
| HHE/HHE | | | - | [9, 16, 23] |
| HHE/HHF*2 | 0 | [66] | | |
| HHE/HHF*1, HHE/HHG*1, HHE/HHG*2 | - | [23] | | |
| HHF*1 | | | - | [9] |
| HHF*2 | 0 | [69] | 0 | [41, 44, 49, 55] |
| | + | [66] | + | [9, 15, 16, 44, 46] |
| | | | - | [54] |
| HHF*2/HHF*2, HHF*2/HHG*1 | - | [23] | | |
| P1 (HHE, HHF*1, HHF*2, HHG*1, HHG*2) | | | - | [50, 52] |
| 59353-T (HHA, HHB, HHC and HHD) | | | + | [47] |
| 59029-G (HHA, HHB, HHC and HHD) | | | + | [38, 50] |
| 59029-G/G | | | + | [38] |
| 8 SNPs in 5'UTR | 0 | [69] | | |

† References include major published population studies of associations of *CCR5* SNPs or haplotypes with HIV transmission, acquisition, or disease progression.

+ Protective effect —delayed or lower HIV progression, transmission or AIDS development.

- Risk effect —accelerated or higher risk of HIV progression, transmission or AIDS development.

0 No significant effect observed.

Supplementary Table 2.6

Table 2.6: *CCR2-CCR5* haplotypes and diplotypes as observed in HIV-1 discordant Zambian couples.

| | Overall | | Index | | Non-Index | |
|---|---|---|---|---|---|---|
| Haplotype | Count | Frequency | Count | Frequency | Count | Frequency |
| HHA | 601 | 0.265 | 305 | 0.269 | 296 | 0.261 |
| HHB | 44 | 0.019 | 25 | 0.022 | 19 | 0.017 |
| HHC | 178 | 0.079 | 69 | 0.061 | 109 | 0.096 |
| HHD | 370 | 0.163 | 190 | 0.168 | 180 | 0.159 |
| HHE | 310 | 0.137 | 138 | 0.122 | 172 | 0.152 |
| HHF*1 | 128 | 0.056 | 76 | 0.067 | 52 | 0.046 |
| HHF*2 | 480 | 0.212 | 259 | 0.228 | 221 | 0.195 |
| HHG*1 | 157 | 0.069 | 72 | 0.064 | 85 | 0.075 |
| Diplotype | | | | | | |
| HHA/HHA | 92 | 0.081 | 48 | 0.085 | 44 | 0.078 |
| HHA/HHC | 44 | 0.039 | 18 | 0.032 | 26 | 0.046 |
| HHA/HHD | 93 | 0.082 | 53 | 0.094 | 40 | 0.071 |
| HHA/HHE | 78 | 0.069 | 34 | 0.060 | 44 | 0.078 |
| HHA/HHF*1 | 26 | 0.023 | 16 | 0.028 | 10 | 0.018 |
| HHA/HHF*2 | 120 | 0.106 | 63 | 0.111 | 57 | 0.101 |
| HHA/HHG*1 | 36 | 0.032 | 15 | 0.027 | 21 | 0.037 |
| HHC/HHD | 30 | 0.027 | 11 | 0.019 | 19 | 0.034 |
| HHC/HHF*2 | 37 | 0.033 | 15 | 0.027 | 22 | 0.039 |
| HHD/HHD | 39 | 0.034 | 26 | 0.046 | 13 | 0.023 |
| HHD/HHE | 49 | 0.043 | 18 | 0.032 | 31 | 0.055 |
| HHD/HHF*2 | 78 | 0.069 | 37 | 0.065 | 41 | 0.072 |
| HHD/HHG*1 | 26 | 0.023 | 12 | 0.021 | 14 | 0.025 |
| HHE/HHF*2 | 78 | 0.069 | 44 | 0.078 | 7 | 0.012 |
| HHE/HHG*1 | 23 | 0.020 | 14 | 0.025 | 9 | 0.016 |
| HHF*1/HHF*1 | 13 | 0.012 | 10 | 0.018 | 3 | 0.005 |
| HHF*1/HHF*2 | 36 | 0.032 | 21 | 0.037 | 15 | 0.027 |
| HHF*2/HHF*2 | 45 | 0.040 | 27 | 0.048 | 18 | 0.032 |
| HHF*2/HHG*1 | 35 | 0.031 | 20 | 0.035 | 15 | 0.027 |
| Other[a] | 156 | 0.138 | 65 | 0.115 | 118 | 0.208 |

[a] Rare diplotypes observed in fewer than 12 individuals in the overall population are grouped together.

# GENETIC VARIATIONS AND HETEROSEXUAL HIV-1 INFECTION: ANALYSIS OF CLUSTERED GENES ENCODING CC-MOTIF CHEMOKINE LIGANDS

by

LIANGYUAN HU, WEI SONG, ILENE BRILL, JOSEPH MULENGA, SUSAN ALLEN, ERIC HUNTER, SADEEP SHRESTHA, JIANMING TANG, AND RICHARD A. KASLOW

*Genes and Immunity*

# ABSTRACT

Several CC-motif chemokine ligands (CCLs) can block HIV-1 binding sites on CCR5 and inhibit viral entry. We studied single nucleotide polymorphisms (SNPs) in genes encoding three CCR5 ligands [CCL3 (MIP-1α), CCL4 (MIP-1β), and CCL5 (RANTES)] along with an adjacent gene encoding a CCR2 ligand [CCL2 (MCP-1)] to identify candidate markers for HIV-1 infection and pathogenesis. Analyses of 567 HIV-1 serodiscordant Zambian couples revealed that rs5029410C (in *CCL3* intron 2) was associated with lower viral load (VL) in seroconverters, adjusted for gender and age (regression beta=-0.57 $\log_{10}$, P=$4x10^{-6}$). In addition, rs34171309A in *CCL3* exon 3 was associated with increased risk of HIV-1 acquisition in exposed seronegatives (hazard ratio=1.52, P=0.006 when adjusted for donor VL and genital ulcer/inflammation). The *CCL3* exon 3 SNP, encoding a conservative Glu-to-Asp substitution, and five neighboring SNPs in tight linkage disequilibrium all showed similar associations with HIV-1 acquisition. How these multiple *CCL3* SNPs may alter the occurrence or course of HIV-1 infection remains to be determined.

The pandemic of acquired immunodeficiency syndrome (AIDS) resulting from human immunodeficiency virus (HIV-1) infection is particularly devastating in southern Africa[31]. HIV-1 enters target cells through a two-step fusion process in which the CC-motif chemokine receptor 5 (CCR5) serve as a major coreceptor on human CD4+ T cells. Multiple studies have demonstrated that individuals who are homozygous for the *CCR5* deletion mutation (D32) lack functional CCR5 and are thus highly resistant to HIV-1 infection[6, 8, 36, 37]. CC-motif chemokine ligands CCL3 (MIP-1α), CCL4 (MIP-1β), and CCL5 (RANTES) are natural ligands for CCR5; as such they may competitively inhibit binding of the receptor by HIV-1[25, 70-73]. Variants in *CCR2,* the gene adjacent to *CCR5*, have also been reported to be independently associated with HIV-1 infection, progression and transmission[74]. CCL2, the natural ligand for CCR2, may also be an important factor for the HIV-1 transmission[19, 75, 76]. Numerous papers studied the influence of *CCR5* variants on the HIV-1 transmission, infection and disease progression. However, observations of the impact of the ligands on HIV-1 acquisition and disease control in populations of African ancestry are sparse. To complement our work on the role of the receptor polymorphisms[74], we examined the associations of single nucleotide polymorphisms (SNPs) in the CCL genes with HIV-1 heterosexual transmission and disease control in Zambia cohort.

Between 1995 and 2006 in Lusaka, Zambia, more than 10,000 couples were screened for their HIV-1 status in the Zambia-Emory HIV-1 Research Project (ZEHRP). The procedures for screening, recruitment, counseling, and follow-up visits have been described elsewhere[42, 64]. The study presented here included 567 HIV-1 serodiscordant couples who were followed for at least 9 months between 1996 and 2006.

Among these, 240 exposed seronegative (HESN) participants acquired phylogenetically linked HIV-1 [74] from their index partners after varying lengths of follow-up *(*interquartile range: 6-36 months) and 327 HESN participants remained seronegative during the study intervals. Non-genetic factors, including age, gender, donor viral load, genital ulcer or inflammation in any partner, and male circumcision were retained as covariates, as established in earlier analyses of HIV-1 transmission within the study population[74].

In the absence of selection for pre-exposure time before enrollment, the observed genetic associations with rate of HIV-1 transmission should not have been systematically affected by duration of pre-exposure. On the other hand, because non-transmitting couples had higher frequencies of non-genetic risk factors, the observed transmission rates were higher than the overall rates in the entire cohort of discordant couples. Overall, the annual HIV-1 seroincidence (7-9 events per 100 person-years) among Zambian couples was reduced by one-half to two-thirds as a result of voluntary testing and counseling[31].

We assessed recognized SNPs reported in dbSNP to map between 1 kb upstream and 500 bp downstream of the four candidate CCL genes (*CCL2, CCL3, CCL4* and *CCL5*). SNPs included in the analysis met at least one of the following criteria: 1) encodes a change in the amino acid sequence of the ligand; 2) occurs at a transcription binding site, an intron/exon boundary site, an alternative splicing site, promoter region, or 3' untranslated region; or 3) has a minor allele frequency (MAF) $\geq$0.02 in Africans/African Americans. We also included SNPs with unknown MAF to provide

additional coverage of the gene. All SNPs had to meet suitability criteria for the iPLEX SNP typing assay at the Broad Institute of MIT and Harvard[77].

Overall, 63 SNPs in 4 CCL genes passed the assay design process; 52 had a call rate of over 90%; and 35 had an MAF ≥0.01 (Table 3.2 in Supplemental Materials). SNPs within each gene tended to have strong LD as judged by D' and $r^2$ values (Figure 3.2 in Supplemental Materials). Haplotype blocks were defined by the Gabriel algorithm in Haploview 4.2 [78]. Haplotype blocks 1 and 2 are in *CCL2,* and haplotype blocks 3, 4 and 5 are in *CCL5*, *CCL3* and *CCL4*, respectively. Of the 35 SNPs, two deviated from Hardy-Weinberg equilibrium (HWE) in their distribution (Table 3.2). SNPs rs1719134 and rs13900 were out of HWE in the overall cohort but conform to HWE in the HESNs and index subgroups (q values at 0.014 and 0.175 respectively). Genotype frequencies of these 2 SNPs in the Zambian cohort were quite similar to those reported for other Africans documented in the dbSNP database (National Center for Biotechnology Information). Therefore, those few SNPs whose frequencies were out of HWE were most likely due to chance.

Of the few reports on CCL gene variants and HIV-1 disease progression[19, 79-82], none of them addressed the association between CCL variants and HIV-1 VL. By linear regression we tested the effect of the minor SNP allele on earliest available (chronic-phase) VL in index partners and the set-point VL taken at 6 months after the imputed infection date in seroconverters[74]. Carriage of one copy of the C allele of rs5029410 in *CCL3* intron 2 was strongly associated with lower VL in the seroconverters (regression $\beta$=-0.57 $\log_{10}$, $P$=$4\times10^{-6}$) with adjustment for age and gender. However, this SNP showed no association with VL in index partners ($\beta$=0.05 $\log_{10}$, $P$=0.46). Index

partners did not received antiretroviral therapy, but their average duration of infection was much longer than that of seroconverters. Thus, the SNP variant might exert its effect only early in HIV-1 infection rather than later.

The A allele of rs34171309 in *CCL3* exon 3 was associated with more rapid acquisition of HIV-1 by HESNs, as shown in an allelic proportional hazards model in the presence of non-genetic factors (HR=1.52, 95% CI 1.13-2.04, *P*=0.006) (Table 3.1). This association was also apparent in logistic regression model: the frequency of rs34171309A was higher in the SCs than HESNs (odds ratio=1.51, *P*=0.05). In Kaplan Meier plots, differences were seen in the overall cohort and in female HESNs (Figure 3.1).

The rs34171309 encodes a non-synonymous Glu-to-Asp amino acid change at position 78 in the CCL3 protein. Although this change is a conservative one, because it occurs very close to the site of CCL3 binding to the CCR5 protein, it may influence CCL3-CCR5 interaction [83]. The A allele of rs34171309 and the C allele of rs5029410 were in weak LD (D'=0.44), but rs34171309A was in strong LD with five other SNPs (rs1719130, rs1719134, rs35511254, rs1634497 and rs1634499) in *CCL3* (D'>0.8), and all these SNPs had similar associations with HIV-1 acquisition (data not shown). Three of them (rs1719130, rs35511254 and rs1719134) had minor alleles associated with fast HIV-1 disease progression in earlier studies [79], [10], consistent with our findings here. In an analysis of the initially HESNs stratified by gender we again found a stronger effect of rs34171309A in the HESN females than in the smaller number of males, but the effect for both groups is in the same direction (female: HR=1.70, 95% CI 1.17-2.47, *P*=0.006; male: HR=1.29, 95% CI 0.78-2.15, *P*=0.33). This difference in significance may be due to an interaction between this SNP and gender or simply to chance variation.

For the two SNPs highlighted in this study, genotypes for one (rs5029410) were validated by a TaqMan genotyping assay (Applied Biosystems, Inc.). Selective tests revealed over 98% concordance rate between TaqMan and iPLEX results. For rs34171309, however, variants could not be readily validated by alternative techniques. Based on alignments of homologous sequences from *CCL3* and *CCL3L1, CCL3L3* genes (Supplementary Figure 3.3), only *CCL3* was polymorphic at the nucleotide position corresponding to rs34171309. As a result, we infer that rs34171309A is most likely present in *CCL3* and not in its homologues.

At the *CCL5* locus, one SNP variant known as In1.1C (rs2280789) has been associated with higher risk of HIV-1 subtype B infection in Europeans and African Americans and with more rapid HIV-1 disease progression in African Americans[25]. In our study population, no such association could be established for In1.1C or its haplotypes. In particular, In1.1C had no association with HIV-1 acquisition (HR=1.05, 95% CI: 0.84-1.31, *P*=0.70).

We did not adjust *P* values for the number of tests implied by the number of SNPs eligible for analysis because a number of SNPs within each gene are in high LD with each other and the number of independent tests would actually be considerably lower. Only the strong association of rs5029410C (*CCL3* intron 2 variant) would have withstood even conservative Bonferroni correction. The false discovery probability (q value) for rs5029410C and rs34171309A was 0.0001 and 0.03, respectively.

Compared with data from commercially available genome wide association chip arrays or open access databases, our study provided denser coverage of the variation in

all four CCL genes studied. The 7 SNPs in *CCL3* span a 5-kb genomic region. In contrast, both the Human1M-Duo DNA Analysis BeadChip (Illumina, Inc.) and the GeneChip Human SNP 6.0 Assay (Affymetrix, Inc.) each targets a single SNP in *CCL3*. Patterns of LD, as determined for 35 CCL SNPs in our study population may help guide future research on African cohorts, because even the latest HapMap Phase II+III dataset (Release #28, August 2010) only reports 19 SNPs in the four CCL genes studied here.

In summary, we systematically screened SNPs in the genes encoding three principal natural ligands of CCR5 (HIV-1 coreceptor) and in the adjacent gene *CCR2*. Data from a large prospective cohort of HIV-1 serodiscordant couples enabled us to test SNP associations with both HIV-1 acquisition and early or chronic-phase VL in a high-risk study population. Overall, two variants in *CCL3* and none in other genes relevant to CCR5 or CCR2 function appeared to influence early events in the natural history of HIV-1 infection, either by altering the rate of HIV-1 infection in HESN partners or by modulating HIV-1 VL soon after seroconversion. Future functional studies of these SNPs may fully resolve their potential effects on HIV-1 acquisition and control.

**References:**


1. UNAIDS. UNAIDS/WHO "AIDS Epidemic Update: December 2006". In; 2006.
2. Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R*, et al.* Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 1996,**86**:367-377.
3. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM*, et al.* Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 1996,**382**:722-725.
4. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, Allikmets R*, et al.* 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* 1996,**273**:1856-1862.
5. Zimmerman PA, Buckler-White A, Alkhatib G, Spalding T, Kubofcik J, Combadiere C*, et al.* Inherited resistance to HIV-1 conferred by an inactivating mutation in CC chemokine receptor 5: studies in populations with contrasting clinical phenotypes, defined racial background, and quantified risk. *Mol Med* 1997,**3**:23-36.
6. An P, Nelson GW, Wang L, Donfield S, Goedert JJ, Phair J*, et al.* Modulating influence on HIV/AIDS by interacting RANTES gene variants. *Proc Natl Acad Sci U S A* 2002,**99**:10002-10007.
7. Rathore A, Chatterjee A, Sivarama P, Yamamoto N, Singhal PK, Dhole TN. Association of RANTES -403 G/A, -28 C/G and In1.1 T/C polymorphism with

HIV-1 transmission and progression among North Indians. *J Med Virol* 2008,**80**:1133-1141.

8. Menten P, Wuyts A, Van Damme J. Macrophage inflammatory protein-1. *Cytokine Growth Factor Rev* 2002,**13**:455-481.

9. Guan E, Wang J, Roderiquez G, Norcross MA. Natural truncation of the chemokine MIP-1 beta /CCL4 affects receptor specificity but not anti-HIV-1 activity. *J Biol Chem* 2002,**277**:32348-32352.

10. Cocchi F, DeVico AL, Garzino-Demo A, Arya SK, Gallo RC, Lusso P. Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells. *Science* 1995,**270**:1811-1815.

11. Malhotra R, Hu L, Song W, Brill I, Mulenga J, Allen S*, et al.* Association of chemokine receptor gene (CCR2-CCR5) haplotypes with acquisition and control of HIV-1 infection in Zambians. *Retrovirology* 2011,**8**:22.

12. Eugenin EA, Osiecki K, Lopez L, Goldstein H, Calderon TM, Berman JW. CCL2/monocyte chemoattractant protein-1 mediates enhanced transmigration of human immunodeficiency virus (HIV)-infected leukocytes across the blood-brain barrier: a potential mechanism of HIV-CNS invasion and NeuroAIDS. *J Neurosci* 2006,**26**:1098-1106.

13. Vilades C, Broch M, Plana M, Domingo P, Alonso-Villaverde C, Pedrol E*, et al.* Effect of genetic variants of CCR2 and CCL2 on the natural history of HIV-1 infection: CCL2-2518GG is overrepresented in a cohort of Spanish HIV-1-infected subjects. *J Acquir Immune Defic Syndr* 2007,**44**:132-138.

14. Ansari AW, Bhatnagar N, Dittrich-Breiholz O, Kracht M, Schmidt RE, Heiken H. Host chemokine (C-C motif) ligand-2 (CCL2) is differentially regulated in HIV type 1 (HIV-1)-infected individuals. *Int Immunol* 2006,**18**:1443-1451.

15. Allen S, Meinzen-Derr J, Kautzman M, Zulu I, Trask S, Fideli U*, et al.* Sexual behavior of HIV discordant couples after HIV counseling and testing. *Aids* 2003,**17**:733-740.

16. Fideli US, Allen SA, Musonda R, Trask S, Hahn BH, Weiss H*, et al.* Virologic and immunologic determinants of heterosexual transmission of human immunodeficiency virus type 1 in Africa. *AIDS Res Hum Retroviruses* 2001,**17**:901-910.

17. Gabriel S, Ziaugra L, Tabbaa D. SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Curr Protoc Hum Genet* 2009,**Chapter 2**:Unit 2 12.

18. Barrett JC. Haploview: Visualization and analysis of SNP genotype data. *Cold Spring Harb Protoc* 2009,**2009**:pdb ip71.

19. Modi WS, Lautenberger J, An P, Scott K, Goedert JJ, Kirk GD*, et al.* Genetic variation in the CCL18-CCL3-CCL4 chemokine gene cluster influences HIV Type 1 transmission and AIDS disease progression. *Am J Hum Genet* 2006,**79**:120-128.

20. Modi WS, Goedert JJ, Strathdee S, Buchbinder S, Detels R, Donfield S*, et al.* MCP-1-MCP-3-Eotaxin gene cluster influences HIV-1 transmission. *AIDS* 2003,**17**:2357-2365.

21. Gonzalez E, Rovin BH, Sen L, Cooke G, Dhanda R, Mummidi S*, et al.* HIV-1 infection and AIDS dementia are influenced by a mutant MCP-1 allele linked to increased monocyte infiltration of tissues and MCP-1 levels. *Proc Natl Acad Sci U S A* 2002,**99**:13795-13800.

22. McDermott DH, Beecroft MJ, Kleeberger CA, Al-Sharif FM, Ollier WE, Zimmerman PA*, et al.* Chemokine RANTES promoter polymorphism affects risk of both HIV infection and disease progression in the Multicenter AIDS Cohort Study. *AIDS* 2000,**14**:2671-2678.

23. Yoshiura C, Kofuku Y, Ueda T, Mase Y, Yokogawa M, Osawa M*, et al.* NMR analyses of the interaction between CCR5 and its ligand using functional reconstitution of CCR5 in lipid bilayers. *J Am Chem Soc* 2010,**132**:6768-6777.

24. Gonzalez E, Dhanda R, Bamshad M, Mummidi S, Geevarghese R, Catano G*, et al.* Global survey of genetic variation in CCR5, RANTES, and MIP-1alpha: impact on the epidemiology of the HIV-1 pandemic. *Proc Natl Acad Sci U S A* 2001,**98**:5199-5204.

**Figure Legend:**

Figure 3.1. Kaplan-Meier plots showing the relationship of rs34171309 to HIV-1 acquisition among initially HIV-1 exposed seronegative (HESN) partners of Zambia couples. (a) 567 HESNs; (b) 295 female HESNs. Genotypes AA + AC are compared with the CC genotype in order to highlight the dominant effect of the minor allele A. Rates of seroconversion depicted here for selected discordant couples are higher than rates in the entire cohort of discordant couples in the Zambia-Emory HIV-1 Research Project (ZEHRP).

Tables

Table 3.1. Association of rs34171309A allele (in *CCL3* exon 3) with acquisition of HIV-1 infection among exposed seronegative (HESN) Zambians before and after stratification by gender.

| | All HESNs (N=567) | | | Female HESNs (N=295) | | | Male HESNs (N=272) | | |
|---|---|---|---|---|---|---|---|---|---|
| Factors in model | HR[a] | 95% CI | P | HR[a] | 95% CI | P | HR[a] | 95% CI | P |
| rs34171309A[b] | 1.52 | 1.13-2.04 | 0.006 | 1.70 | 1.17-2.47 | 0.006 | 1.29 | 0.78-2.15 | 0.33 |
| GUI in both partners[c] | 8.72 | 5.72-13.28 | <0.0001 | 7.85 | 4.62-13.33 | <0.0001 | 10.39 | 5.15-20.97 | <0.0001 |
| GUI in either partner[d] | 3.04 | 2.12-4.35 | <0.0001 | 2.89 | 1.79-4.66 | <0.0001 | 3.50 | 1.99-6.18 | <0.0001 |
| Donor VL (per $\log_{10}$) | 1.37 | 1.12-1.69 | 0.003 | 1.15 | 0.86-1.53 | 0.35 | 1.66 | 1.19-2.31 | 0.003 |

[a] Hazard ratio (HR) from a Cox proportional hazards model, adjusted for all factors in the model. CI, confidence interval; GUI, genital ulcer/inflammation; VL, plasma HIV-1 viral load (RNA copies/ml).

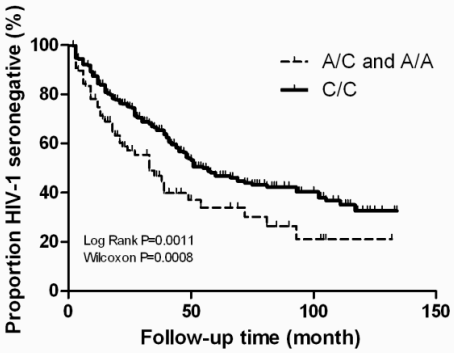
[b] Also tested in logistic regression models (see text).

[c] Genital ulcer/inflammation seen in both partners in each couple.

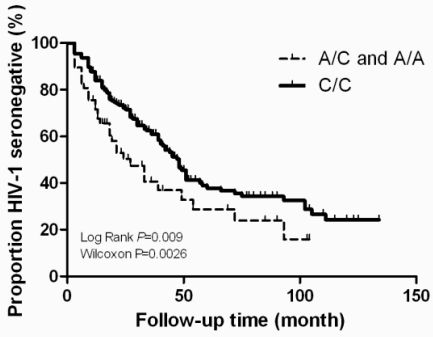
[d] Genital ulcer/inflammation seen in either partner in each couple.

**Figure 1**

**a.**


Proportion HIV-1 seronegative (%)
100
80
60
40
20
0
0
50
100
150
Follow-up time (month)
A/C and A/A
C/C
Log Rank P=0.0011
Wilcoxon P=0.0008

**b.**


Proportion HIV-1 seronegative (%)
100
80
60
40
20
0
0
50
100
150
Follow-up time (month)
A/C and A/A
C/C
Log Rank P=0.009
Wilcoxon P=0.0026

**Supplemental Materials (1 Table and 2 Figures)**

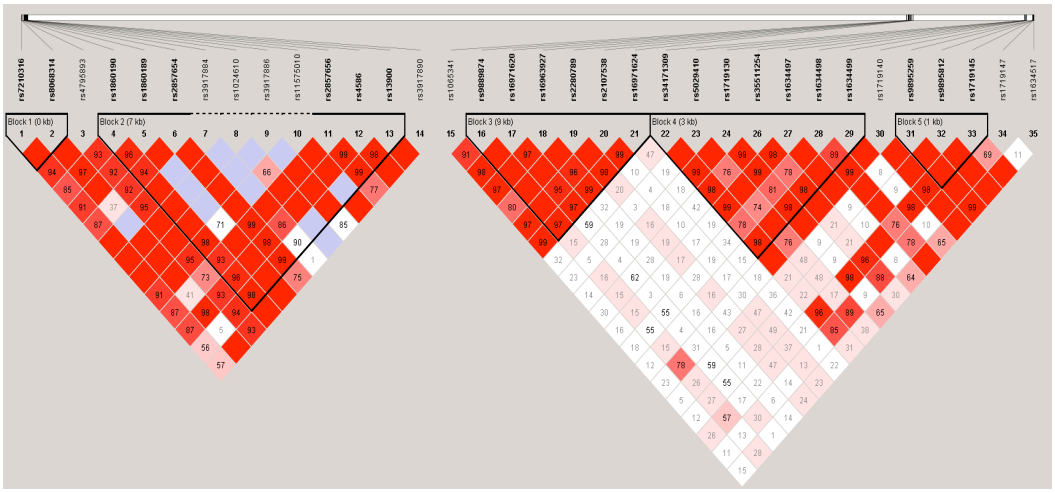
Table 3.2. List of 35 SNPs available for analysis in this study[a].

| SNP ID | Gene | Position on Chr. 17 | HWE p value | Call rate | MAF | Alleles (major:minor) |
|---|---|---|---|---|---|---|
| rs7210316 | *CCL2* | 29597620 | 0.88 | 98.4 | 0.36 | T:G |
| rs8068314 | *CCL2* | 29598394 | 0.08 | 92.3 | 0.22 | A:T |
| rs4795893 | *CCL2* | 29598561 | 0.79 | 99.6 | 0.35 | A:G |
| rs1860190 | *CCL2* | 29600686 | 0.21 | 99.3 | 0.29 | A:T |
| rs1860189 | *CCL2* | 29602471 | 0.36 | 99.8 | 0.07 | A:G |
| rs2857654 | *CCL2* | 29603644 | 0.22 | 99.6 | 0.38 | C:A |
| rs3917884 | *CCL2* | 29604093 | 0.87 | 98.4 | 0.02 | T:C |
| rs1024610 | *CCL2* | 29604344 | 0.18 | 99.7 | 0.04 | A:T |
| rs3917886 | *CCL2* | 29604755 | 0.92 | 99.2 | 0.03 | T:A |
| rs11575010 | *CCL2* | 29606052 | 0.13 | 99.5 | 0.05 | T:C |
| rs2857656 | *CCL2* | 29606120 | 0.90 | 98.4 | 0.41 | G:C |
| rs4586 | *CCL2* | 29607382 | 0.15 | 96.3 | 0.29 | C:T |
| rs13900 | *CCL2* | 29608024 | 0.01 | 99.3 | 0.19 | C:T |
| rs3917890 | *CCL2* | 29608895 | 0.10 | 99.4 | 0.06 | G:T |
| rs1065341 | *CCL5* | 31222706 | 0.55 | 98.6 | 0.29 | T:C |
| rs9889874 | *CCL5* | 31226410 | 0.34 | 99.7 | 0.19 | G:T |
| rs16971620 | *CCL5* | 31226781 | 0.60 | 99.4 | 0.07 | A:C |
| rs16963927 | *CCL5* | 31228920 | 0.68 | 98.2 | 0.28 | A:G |
| rs2280789 | *CCL5* | 31231116 | 0.32 | 99.7 | 0.21 | A:G |
| rs2107538 | *CCL5* | 31231893 | 0.90 | 99.6 | 0.49 | C:T |
| rs16971624 | *CCL5* | 31235542 | 0.63 | 98.5 | 0.24 | T:A |
| rs34171309 | *CCL3* | 31440176 | 0.49 | 99.4 | 0.10 | C:A |
| rs5029410 | *CCL3* | 31440264 | 1.00 | 98.3 | 0.14 | A:C |
| rs1719130 | *CCL3* | 31440359 | 0.44 | 99.2 | 0.19 | T:C |
| rs1719134 | *CCL3* | 31441059 | 0.00 | 99.3 | 0.11 | G:A |
| rs35511254 | *CCL3* | 31441405 | 0.34 | 98.5 | 0.10 | G:A |
| rs1634497 | *CCL3* | 31443060 | 0.63 | 97.7 | 0.19 | A:G |
| rs1634498 | *CCL3* | 31443538 | 1.00 | 99.3 | 0.10 | T:C |
| rs1634499 | *CCL3* | 31444038 | 0.53 | 99.0 | 0.19 | G:C |
| rs1719140 | *CCL4* | 31454526 | 1.00 | 99.5 | 0.08 | A:T |
| rs9895259 | *CCL4* | 31454569 | 0.08 | 99.5 | 0.12 | G:C |
| rs9895812 | *CCL4* | 31454831 | 0.16 | 99.2 | 0.11 | G:C |
| rs1719145 | *CCL4* | 31455965 | 1.00 | 97.0 | 0.29 | G:A |
| rs1719147 | *CCL4* | 31456238 | 0.12 | 98.5 | 0.34 | G:A |
| rs1634517 | *CCL4* | 31456516 | 0.82 | 96.6 | 0.28 | C:A |

[a] All had minor allele frequency (MAF) ≥0.01 in the study population.

**Supplementary Figures**

a



b

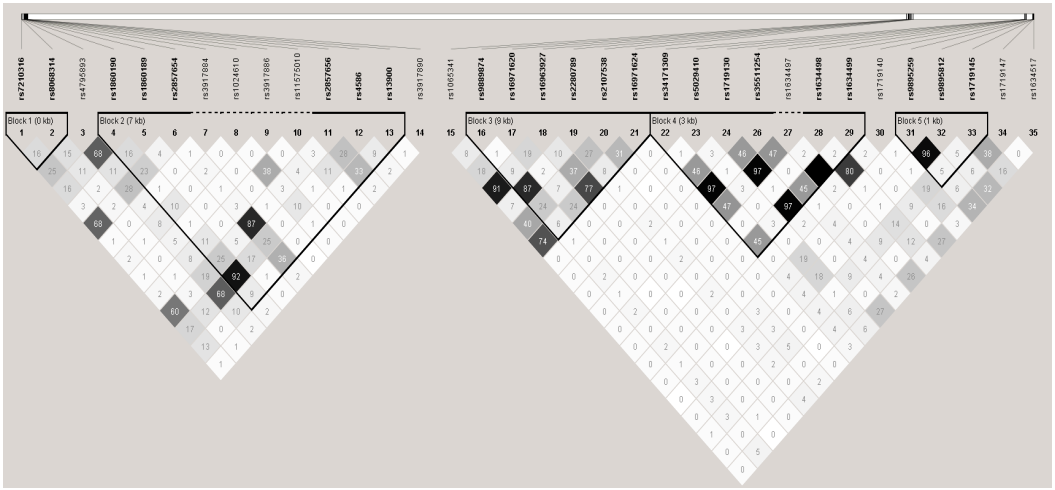


Figure 3.2. Patterns of linkage disequilibrium among 35 SNPs with minor allele frequency ≥0.01 in the study population. *CCL2* SNPs are within block 1 and block 2. SNPs in *CCL5, CCL3* and *CCL4* correspond to blocks 3, 4 and 5, respectively. The $r^2$ values (a) and D' values (b) as shown are based on HaploView version 4.2 (http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/downloads).

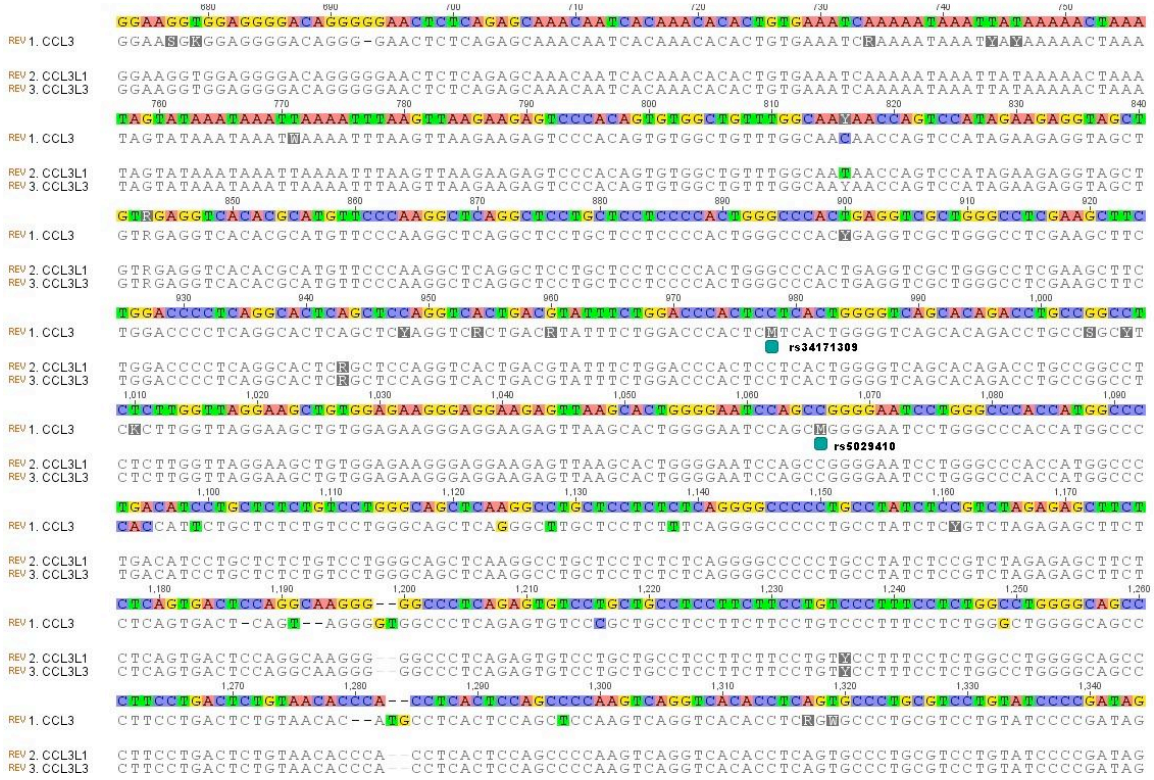



Figure 3.3. Partial alignments of *CCL3*, *CCL3L1* and *CCL3L3* gene sequences flanking the region with multiple *CCL3* SNPs, including rs34171309 and rs5029410 highlighted in this study. Known SNPs (bases shaded grey) are based on Ensembl genome browser.

# SEQUENCING OF THE *CCR2* AND *CCR5* GENES AND INVESTIGATION OF VARIANTS FOR ASSOCIATION WITH HIV-1

by

LIANGYUAN HU, ROBERT MAKOWSKY, WEI SONG, ILENE BRILL, JOSEPH MULENGA, SUSAN ALLEN, ERIC HUNTER, SADEEP SHRESTHA, JIANMING TANG, AND RICHARD A. KASLOW

## Abstract

*Background:* C-C-motif chemokine receptor 5 (CCR5) is a co-receptor for HIV-1 on human CD4+ T cells. Previous studies showed variants in *CCR2* and *CCR5* genes were associated with HIV-1 progression and acquisition.

*Results:* We sequenced the entire *CCR5* gene and the 3' and 5' untranslated regions (UTRs) of *CCR2* gene for SNP discovery in 109 Zambian individuals. We used two Caucasian individuals who are homozygous of *CCR5* Δ32 homozygous mutation (rs333) as references. Twenty-seven variants were identified in this cohort; 1 novel SNP was found in *CCR2* and 2 novel SNPs were found in *CCR5* gene. Four non-synonymous SNPs with known function on HIV-1 disease were genotyped in 567 HIV-1 serodiscordant couples in the Zambian cohort. *CCR5* coding SNP rs1800944 showed association with increased viral load in donor after adjusting for age and gender ($\beta$=0.24, $\log_{10}$ $P$=0.026). SNP rs1800944 encodes amino acid Ala-to-Val change in the CCR5 protein position 335. None of the other SNPs showed significant association with HIV-1 acquisition or viral load.

*Conclusion: CCR5* coding SNP rs1800944 may represent a previously unrecognized risk factor for HIV-1 donor viral load in Zambia cohort. Other common variants in *CCR2* and *CCR5* do not strongly influence HIV-1 acquisition or disease control in African population.

**Background:**

HIV-1 can only infect cells that express the receptor CD4 and one of several co-receptors, usually the C-C-motif chemokine receptor 5 (CCR5) [1]. Mutations producing defects in the *CCR5* gene are the best-known correlates of natural resistance to HIV infection. *CCR5* Δ32 deletion is the most widely studied mutation in *CCR5* gene [1] [2]. It is a deletion of the nucleotides 794 to 825 of the cDNA sequence and results in a reading frame shift after amino acid 174. CCR5 is truncated at codon 182. It has been shown that the severely truncated protein could not be detected at the surface of cells that normally express the protein [3]. Extensive work has demonstrated that *CCR5* Δ32 acts as a recessive restriction gene against HIV-1 infection and exerts a dominant effect in delaying progression to AIDS among infected individuals [4-5]. The protective effect of *CCR5* Δ32 allele has been proved repeatedly in different populations [6]. However, this allele is virtually absent in ethnic groups other than Caucasians, and the vast majority of people exposed to but uninfected by HIV-1 do not carry the Δ32 allele. Thus, some other factors must account for their resistance to HIV infection [7].

The purpose of our study is to sequence and screen common variants in the *CCR5* gene and *CCR2* untranslated region (UTR) in the cohort of Zambian HIV-1 discordant couples and study their association in the HIV-1 disease control and transmission. Detailed characteristics of the cohort and clinic outcome measurement are described in chapter 1 [8].

**Material and Methods:**

We randomly selected 109 individuals from 567 Zambia HIV-1 serodiscordant couples for the sequencing project. Since CCR5 Δ32 mutation (rs333) is missing in the African population, we also included 2 individuals in European American population who are homozygous of HHG*2 haplotype as references. To have a better coverage of the *CCR2-CCR5* haplotypes, we over-sampled a few individuals who are homozygous of rare haplotypes.

We sequenced the entire *CCR5* gene and untranslated region (UTR) of *CCR2* gene (PolymorphicDNA Inc., Alameda, CA). PCR and sequencing primers are listed in Table 4.3. TaqMan genotyping assay was used to type selective SNPs in the overall Zambia cohort.

We used Geneious 5.5.5 software to align the sequence and draw phylogenetic tree. Phylogenetic tree is based on rs1799864 and 7 SNPs in the *CCR5* promoter region, 1 in the coding region (rs333) and 5 SNPs in the *CCR5* 3'UTR region. Bootstrap method was used in the construction of the phylogenetic tree. We used 100 replications and the consensus support (%) was marked on the major nodes of the tree.

**Statistical Analysis**

Association of the SNPs with HIV-1 VL among the index partners and seroconverters were tested using a general linear model with adjustment for age and gender. Cox proportional hazards models and Kaplan-Meier plots were used to estimate the risk for HIV-1 acquisition. Detailed methods were described in Chapter 2 (page 27).

**Results:**

SNPs in the *CCR2-CCR5* genes in the Zambian cohort

Overall, 27 SNPs were found in the cohort (9 in *CCR2* and 18 in *CCR5*) with minor allele frequencies ranging from 0.010 to 0.482 (Table 4.1). We found 1 novel SNP in *CCR2* gene and 2 in *CCR5*. Of the 18 SNPs in *CCR5*, 5 occurred in the coding region and 4 were non-synonymous (Table 4.2).

Of the 5 SNPs in the *CCR5* coding region, only rs1800944 has minor allele frequency > 0.02. Due to the sample size, we may not have been able to detect effects of SNPs with such low minor allele frequencies in our cohort. We genotyped all 5 coding SNPs using a TaqMan assay (ABI, Foster City, CA). For the 4 rare SNPs we only tested 184 individuals (368 chromosomes). The frequencies of these SNPs in the overall cohort were similar to those seen in the selected 109 samples, with all frequencies less than 0.02. We did not perform further work on those rarer SNPs.

We tested the association of rs1800944 with HIV-1 VL in donor and seroconverters and with acquisition. This SNP was associated with higher viral load in the donor after adjusting for the effects of age and gender in a multivariable model ($\beta$=0.24, $\log_{10}$ *P*=0.026). However, the SNP showed no association with seroconverter VL or with time to acquisition after controlling for VL and genital ulcer and genital inflammation in either partner (Cox model HR=1.08, *P*=0.67).

**Phylogenetic relationship of *CCR2-CCR5* haplotypes.**

*CCR2-CCR5* haplotypes were assigned based on 8 SNPs, 1 in *CCR2* gene (rs1799864) and 7 in the CCR5 promoter region and Δ32 mutation (rs333) in the coding region of *CCR5* (Figure 4.1a) [9]. We found 5 more SNPs in the CCR5 3' UTR and

added these SNPs to the tree (Figure 4.1b) with the purpose of determining whether those 3'UTR SNPs could further subdivide the lineages from the initial trees into additional haplotypes that would show any additional association with HIV-1.

Among the SNPs in the *CCR5* 3'UTR, rs41495153 had the highest minor allele frequency (MAF) (MAF=0.266); the other 3'UTR SNPs were uncommon (MAF<0.06) and showed in linkage disequilibrium with each other. Therefore, we could reassign 3'UTR haplotypes, with the single dominant haplotype defined by rs41495153. However, inclusion of the 3'UTR portion of the sequence could not further stratify the *CCR2-CCR5* haplotypes originally defined by CCR2-64I (rs1799864), *CCR5* Δ32 mutation, and 7 *CCR5* promoter SNPs. Comparing figure 4.1a and figure 4.1b, we found that despite modest structural difference demonstrated in the 2 trees, the major branches remain unchanged. That is, haplotype HHA, HHB and HHD clustered together, and HHE, HHF1, HHF*2 and HHG*1 are on the other side of the branch.

**Conclusion and discussion**

We sequenced over 100 individuals in our Zambian cohort for variants in *CCR2* and *CCR5* genes and found 27 SNPs. 3 of the SNPs are novel. A non-synonymous SNP (rs1800944) coding for an Ala-to-Val amino acid change showed an association with higher donor VL in Zambia cohort. This SNP encodes a mutation at the C terminal tail of CCR5 (A335V). It has been reported that rs1800944 was associated with fewer binding sites for both CCL4 and CCL5, but it did not have an o effect on HIV-1 infection [10]. However, how the rs1800944 influence the CCL4 and CCL5 binding site remains unclear, and no association study on this SNP has been reported. Further research on the

function of this SNP is required to verify its role in the HIV-1 VL. None of the other SNPs were statistically associated with HIV-1 viral load or acquisition in the cohort.

**Reference:**


1. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM*, et al.* Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 1996,**382**:722-725.
2. Liang H, Baudouin C, Hamard P, Creuzot-Garcher C, Warnet JM, Brignole-Baudouin F. [Activation of TH1/TH2 pathways detected through the expression of CCR4 and CCR5 on the ocular surface of glaucomatous patients treated over the long term]. *J Fr Ophtalmol* 2006,**29**:121-126.
3. Dean M, Jacobson LP, McFarlane G, Margolick JB, Jenkins FJ, Howard OM*, et al.* Reduced risk of AIDS lymphoma in individuals heterozygous for the CCR5-delta32 mutation. *Cancer Res* 1999,**59**:3561-3564.
4. Benkirane M, Jin DY, Chun RF, Koup RA, Jeang KT. Mechanism of transdominant inhibition of CCR5-mediated HIV-1 infection by ccr5delta32. *J Biol Chem* 1997,**272**:30603-30606.
5. Cohen OJ, Paolucci S, Bende SM, Daucher M, Moriuchi H, Moriuchi M*, et al.* CXCR4 and CCR5 genetic polymorphisms in long-term nonprogressive human immunodeficiency virus infection: lack of association with mutations other than CCR5-Delta32. *J Virol* 1998,**72**:6215-6217.
6. Mangano A, Kopka J, Batalla M, Bologna R, Sen L. Protective effect of CCR2-64I and not of CCR5-delta32 and SDF1-3'A in pediatric HIV-1 infection. *J Acquir Immune Defic Syndr* 2000,**23**:52-57.
7. Zimmerman PA, Buckler-White A, Alkhatib G, Spalding T, Kubofcik J, Combadiere C*, et al.* Inherited resistance to HIV-1 conferred by an inactivating

mutation in CC chemokine receptor 5: studies in populations with contrasting clinical phenotypes, defined racial background, and quantified risk. *Mol Med* 1997,**3**:23-36.

8. Malhotra R, Hu L, Song W, Brill I, Mulenga J, Allen S*, et al.* Association of chemokine receptor gene (CCR2-CCR5) haplotypes with acquisition and control of HIV-1 infection in Zambians. *Retrovirology* 2011,**8**:22.

9. Gonzalez E, Dhanda R, Bamshad M, Mummidi S, Geevarghese R, Catano G*, et al.* Global survey of genetic variation in CCR5, RANTES, and MIP-1alpha: impact on the epidemiology of the HIV-1 pandemic. *Proc Natl Acad Sci U S A* 2001,**98**:5199-5204.

10. Dong HF, Wigmore K, Carrington MN, Dean M, Turpin JA, Howard OM. Variants of CCR5, which are permissive for HIV-1 infection, show distinct functional responses to CCL3, CCL4 and CCL5. *Genes Immun* 2005,**6**:609-619.

## Tables

Table 4.1. List of 27 SNPs in the *CCR2-CCR5* UTR region and *CCR5* coding region in the Zambia cohort.

| Name | Position | Gene | Region | MAF | Alleles |
|---|---|---|---|---|---|
| rs3918370 | 46369156 | *CCR2* | 5'UTR | 0.252 | C:G |
| rs3918358 | 46369423 | *CCR2* | 5'UTR | 0.28 | A:C |
| CCR2 Novel 1 | 46369600 | *CCR2* | 5'UTR | 0.174 | G:A |
| rs3918359 | 46369684 | *CCR2* | 5'UTR | 0.05 | T:A |
| rs3918372 | 46369744 | *CCR2* | 5'UTR | 0.174 | T:A |
| rs1799864 | 46374212 | *CCR2* | coding | 0.193 | G:A |
| rs3918388 | 46376745 | *CCR2* | 3'UTR | 0.023 | G:A |
| rs743660 | 46377022 | *CCR2* | 3'UTR | 0.046 | G:A |
| rs34138562 | 46377057 | *CCR2* | 3'UTR | 0.298 | T:G |
| rs2227010 | 46386546 | *CCR5* | 5'UTR | 0.133 | A:G |
| rs2856758 | 46386665 | *CCR5* | 5'UTR | 0.05 | A:G |
| rs2734648 | 46386844 | *CCR5* | 5'UTR | 0.284 | G:T |
| rs1799987 | 46386939 | *CCR5* | 5'UTR | 0.477 | G:A |
| rs1799988 | 46387263 | *CCR5* | 5'UTR | 0.482 | T:C |
| rs41469531 | 46387266 | *CCR5* | 5'UTR | 0.216 | C:T |
| rs1800023 | 46387312 | *CCR5* | 5'UTR | 0.055 | A:G |
| rs1800824 | 46387545 | *CCR5* | 5'UTR | 0.321 | C:T |
| rs41495153 | 46391149 | *CCR5* | 3'UTR | 0.266 | G:A |
| rs41418945 | 46391240 | *CCR5* | 3'UTR | 0.06 | G:A |
| rs41466044 | 46391243 | *CCR5* | 3'UTR | 0.06 | G:A |
| rs1800874 | 46391474 | *CCR5* | 3'UTR | 0.041 | G:T |
| rs41535253 | 46391622 | *CCR5* | 3'UTR | 0.046 | T:C |
| rs1800941 | 46414618 | *CCR5* | coding | 0.01 | C:T |
| CCR5 Novel 1 | 46414712 | *CCR5* | coding | 0.021 | C:T |
| CCR5 Novel 2 | 46415066 | *CCR5* | coding | 0.01 | C:T |
| rs1800944 | 46415397 | *CCR5* | coding | 0.094 | C:T |
| rs1800945 | 46415409 | *CCR5* | coding | 0.021 | A:G |

Table 4.2. SNPs in the *CCR5* coding region and predicted functions.

| SNP rs # | Freq in 96 chromosomes | Freq in 368 chromosomes | Function | Amino Acid change |
|---|---|---|---|---|
| rs1800941 | 0.01 | NA | Synonymous | No change |
| Novel 1 | 0.02 | <0.02 | Non-synonymous | Leu->Phe |
| Novel 2 | 0.01 | <0.02 | Non-synonymous | Arg->stop |
| rs1800944 | 0.09 | 0.10** | Non-synonymous | Ala->Val |
| rs1800945 | 0.02 | <0.02 | Non-synonymous | Tyr->Phe |

*Non-synonymous SNPs were tested for additional 184 individuals to confirm their frequency in the Zambia cohort.

**Frequency in the 567 HIV-1 serodiscordant couples.

Table 4.3. Primers used for sequencing in *CCR2* and *CCR5* 5'UTR and 3' UTR.

| Name [a] | Sequence from 5' to 3' | Gene | Specificity |
|---|---|---|---|
| CCR2-P1 | AAA GGA CAC CTG GAC TGC | *CCR2* | Forward PCR 5' UTR |
| CCR2-S1 | CCT AGA GGA TGT TAA GTG A | *CCR2* | Forward sequencing 5' UTR |
| CCR2-S2 | TAC AGG TGT GTG CCA TCA TG | *CCR2* | Reverse Sequencing 5' UTR |
| CCR2-P2 | ATT GAA AAC CCT CTC TTC TAG C | *CCR2* | Reverse PCR 5' UTR |
| CCR2-P3 | CCT TCC AGT TCC TCA TTT TTG A | *CCR2* | Forward PCR 3' UTR |
| CCR2-S3 | AAA GTT CTG CTC TGT CCC | *CCR2* | Reverse Sequencing 3' UTR |
| CCR2-P4 | CAC AAT CTT CCA CAC CAC AG | *CCR2* | Reverse PCR 3' UTR |
| | | | |
| CCR5-P1 | AGC TCT CTG CTG TCT TCT CA | *CCR5* | Forward PCR 5' UTR |
| CCR5-S1 | CAC TAA GAT CCT GGG TCC A | *CCR5* | Forward sequencing 5' UTR |
| CCR5-S2 | ACA AGA TCA CAG GGC TTT TC | *CCR5* | Reverse Sequencing 5' UTR |
| CCR5-P2 | AGG GGA ACG GAT GTC TCA | *CCR5* | Reverse PCR 5' UTR |
| CCR5-P3 | GAA GCC TCA CTG CAA GCA | *CCR5* | Forward PCR 3' UTR |
| CCR5-S3 | ATC TAG TCT CCT CCT GGA | *CCR5* | Reverse Sequencing 3' UTR |
| CCR5-P4 | TGT TCC TAG ACC TCA TAC CT | *CCR5* | Reverse PCR 3' UTR |

[a] Primers CCR2-P1, CCR2-P2, CCR2-P3 and CCR2-P4 are PCR primers; primers CCR2-S1, CCR2-S2 and CCR2-S3 are sequencing primers. The same nomenclature applies for CCR5.

Figure 4.1

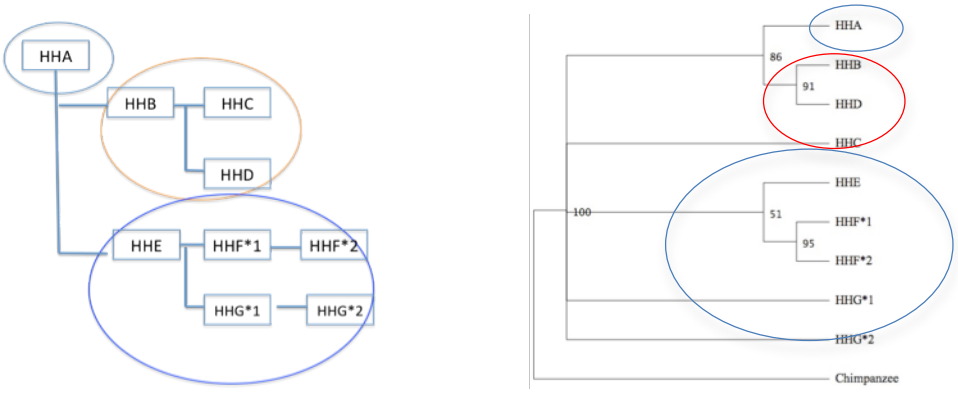



Figure 4.1 Phylogenetic trees based on *CCR2-CCR5* extended SNPs compared with reference tree on *CCR2* (rs1799864), 7 *CCR5* promoter SNPs and *CCR5* Δ32 mutation (rs333).

a. Phylogenetic tree based on rs1799864 and 7 SNPs in the *CCR5* promoter region and *CCR5* Δ32.

b. Phylogenetic tree based on rs1799864 and 7 SNPs in the *CCR5* promoter region, *CCR5* Δ32 and 5 SNPs in the *CCR5* 3'UTR region. Node labels shows the consensus support (%).

# CONCLUSIONS

We analyzed the association of variants in C-C-motif chemokine ligands and receptors and found associations of certain variants with HIV-1 transmission or VL among a population of Zambian HIV-1-discordant couples. Our findings provide further insights into the role of CCL and CCR variants associated with both HIV-1 acquisition and early or chronic-phase VL in a high-risk study population. This is the largest study so far in HIV-1 serodiscordant couples of African descent. Our findings with *CCR2-CCR5* haplotypes and *CCL3* variants at least partially corroborated those in other populations. At the same time we detected an effect of an African-specific genotype (HHD/HHE) on HIV-1 acquisition and VL. Our study is also the first one to evaluate systematically the associations between HIV-1 and SNPs within the genes for each of the principal CCR2 and CCR5 ligands: *CCL2, CCL3, CCL4* and *CCL5.*Coverage of these gene sequences was more complete than that obtained using commercially available genome wide association chip arrays or drawing upon open access databases.

HIV-1 pathogenesis is a complex and dynamic process involving many host genes. There are many genes associated with HIV-1 disease control, especially those genes involving immunity process. Each variant can only explain a small fraction on the

effect on HIV-1 related outcomes. To achieve an increasingly better understanding of the complexity of HIV-1 infection, analytic techniques would optimally incorporate the growing number of contributing factors simultaneously. For example, human leukocyteantigen (HLA) class I markers and haplotypes in killer cell immunoglobulin-like receptor (KIR) genes are reported to be associated with HIV-1 VL and transmission independently in the Zambia cohort [31-33]. In our preliminary effort to account for these effects, adding HLA and KIR variants to the multivariable regression models led to diminished significance of the effects of CCR and CCL, indicating potential correlation or interaction between these genes. Future studies of combinations of these genes may shed further light on these interrelationships.

The relatively large sample size enabled us to detect the suggestive effect of uncommon genotypes on HIV-1. However, HHD/HHE genotype is of low frequency in African population (4%) and even rarer in other population. If this genotype does predispose to infection or to poorer disease control, the effect at the population level will be modest. Another limitation of the study was the difficulty quantifying the exposure to HIV-1 of the seronegative partners. We used follow up time as a quantitative estimate of the exposure. This estimation assumes longer cohabitation time of the couple means higher exposure to the HIV-1. The actual risk behaviors were not used as cofactors for infection and controlled in our statistical models because the self reported data from this cohort is not considered sufficiently accurate. Cofactors like condom use, and frequency of unprotected sex, number of sex partners may also contribute to the risk of HIV-1 transmission, may improve modeling of genetic and non-genetic cofactors in populations with more reliable data.

In conclusion, our data favor the hypothesis that host genetic variants in CC chemokine ligands and receptors play a substantial role in HIV-1 acquisition and disease control. However, little is known about the biological function of these variants in the setting of HIV-1 infection, and in particular how the variants change the binding affinity of the CCL and CCR products or how they interact with each other in the regulation of HIV-1 infection. Further investigation must include other HIV-1 related host genes and the interaction of viral particles with CCR CCL genes.

# LIST OF ABBREVIATIONS

AIDS—acquired immunodeficiency syndrome

CCL2—C-C-motif chemokine ligand 2

CCL3—C-C-motif chemokine ligand 3

CCL4—C-C-motif chemokine ligand 4

CCL5—C-C-motif chemokine ligand 5

CCR2—C-C-motif chemokine receptor 2

CCR5—C-C-motif chemokine receptor 5

GLM—general linear model

HESN—HIV-1 exposed seronegatives

HHA, etc—human haplotype A, etc.

HIV-1—human immunodeficiency virus-1

HR—hazard ratio

HWE—Hardy-Weinberg equilibrium

SC—HIV-1 seroconverter

SNP—single nucleotide polymorphism

SSP—sequence-specific primers

UTR—untranslated region

VL—viral load

# GENERAL LIST OF REFERENCES


1. 2010 Report on the global AIDS epidemic. In; 2010.
2. Kilmarx PH. Global epidemiology of HIV. *Curr Opin HIV AIDS* 2009,**4**:240-246.
3. Kulkarni H, Agan BK, Marconi VC, O'Connell RJ, Camargo JF, He W*, et al.* CCL3L1-CCR5 genotype improves the assessment of AIDS Risk in HIV-1-infected individuals. *PLoS One* 2008,**3**:e3165.
4. Chan DC, Kim PS. HIV entry and its inhibition. *Cell* 1998,**93**:681-684.
5. Liang H, Baudouin C, Hamard P, Creuzot-Garcher C, Warnet JM, Brignole-Baudouin F. [Activation of TH1/TH2 pathways detected through the expression of CCR4 and CCR5 on the ocular surface of glaucomatous patients treated over the long term]. *J Fr Ophtalmol* 2006,**29**:121-126.
6. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM*, et al.* Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature* 1996,**382**:722-725.
7. Dean M, Jacobson LP, McFarlane G, Margolick JB, Jenkins FJ, Howard OM*, et al.* Reduced risk of AIDS lymphoma in individuals heterozygous for the CCR5-delta32 mutation. *Cancer Res* 1999,**59**:3561-3564.
8. Zimmerman PA, Buckler-White A, Alkhatib G, Spalding T, Kubofcik J, Combadiere C*, et al.* Inherited resistance to HIV-1 conferred by an inactivating mutation in CC chemokine receptor 5: studies in populations with contrasting

clinical phenotypes, defined racial background, and quantified risk. *Mol Med* 1997,**3**:23-36.

9. Gonzalez E, Bamshad M, Sato N, Mummidi S, Dhanda R, Catano G*, et al.* Race-specific HIV-1 disease-modifying effects associated with CCR5 haplotypes. *Proc Natl Acad Sci U S A* 1999,**96**:12004-12009.

10. Gonzalez E, Dhanda R, Bamshad M, Mummidi S, Geevarghese R, Catano G*, et al.* Global survey of genetic variation in CCR5, RANTES, and MIP-1alpha: impact on the epidemiology of the HIV-1 pandemic. *Proc Natl Acad Sci U S A* 2001,**98**:5199-5204.

11. Dolan MJ, Kulkarni H, Camargo JF, He W, Smith A, Anaya JM*, et al.* CCL3L1 and CCR5 influence cell-mediated immunity and affect HIV-AIDS pathogenesis via viral entry-independent mechanisms. *Nat Immunol* 2007,**8**:1324-1336.

12. Collman RG. Effect of CCR2 and CCR5 variants on HIV disease: abstract and commentary. *JAMA* 1997,**278**:2113-2114.

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

14. Wang B, Palasanthiran P, Zeigler J, Cunningham A, Saksena NK. CCR5-delta 32 gene deletion in HIV-1 infected patients. *Lancet* 1997,**350**:742.

15. Smith MW, Dean M, Carrington M, Winkler C, Huttley GA, Lomb DA*, et al.* Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study

(MHCS), San Francisco City Cohort (SFCC), ALIVE Study. *Science* 1997,**277**:959-965.

16. Tang J, Shelton B, Makhatadze NJ, Zhang Y, Schaen M, Louie LG*, et al.* 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* 2002,**76**:662-672.
17. Modi WS. CCL3L1 and CCL4L1 chemokine genes are located in a segmental duplication at chromosome 17q12. *Genomics* 2004,**83**:735-738.
18. Ansari AW, Schmidt RE, Heiken H. Prednisolone mediated suppression of HIV-1 viral load strongly correlates with C-C chemokine CCL2: In vivo and in vitro findings. *Clin Immunol* 2007,**125**:1-4.
19. Vilades C, Broch M, Plana M, Domingo P, Alonso-Villaverde C, Pedrol E*, et al.* Effect of genetic variants of CCR2 and CCL2 on the natural history of HIV-1 infection: CCL2-2518GG is overrepresented in a cohort of Spanish HIV-1-infected subjects. *J Acquir Immune Defic Syndr* 2007,**44**:132-138.
20. Campbell GR, Spector SA. CCL2 increases X4-tropic HIV-1 entry into resting CD4+ T cells. *J Biol Chem* 2008,**283**:30745-30753.
21. Berger EA, Murphy PM, Farber JM. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Annu Rev Immunol* 1999,**17**:657-700.
22. Gerard C, Rollins BJ. Chemokines and disease. *Nat Immunol* 2001,**2**:108-115.

23. Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G*, et al.* The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* 2005,**307**:1434-1440.

24. Urban TJ, Weintrob AC, Fellay J, Colombo S, Shianna KV, Gumbs C*, et al.* CCL3L1 and HIV/AIDS susceptibility. *Nat Med* 2009,**15**:1110-1112.

25. Shao W, Tang J, Song W, Wang C, Li Y, Wilson CM*, et al.* CCL3L1 and CCL4L1: variable gene copy number in adolescents with and without human immunodeficiency virus type 1 (HIV-1) infection. *Genes Immun* 2007,**8**:224-231.

26. An P, Nelson GW, Wang L, Donfield S, Goedert JJ, Phair J*, et al.* Modulating influence on HIV/AIDS by interacting RANTES gene variants. *Proc Natl Acad Sci U S A* 2002,**99**:10002-10007.

27. Townson JR, Barcellos LF, Nibbs RJ. Gene copy number regulates the production of the human chemokine CCL3-L1. *Eur J Immunol* 2002,**32**:3016-3026.

28. Modjarrad K, Zulu I, Karita E, Kancheya N, Funkhouser E, Allen S. Predictors of HIV serostatus among HIV discordant couples in Lusaka, Zambia and female antenatal clinic attendants in Kigali, Rwanda. *AIDS Res Hum Retroviruses* 2005,**21**:5-12.

29. Young JM, Turpin JA, Musib R, Sharma OK. Outcomes of a National Institute of Allergy and Infectious Diseases Workshop on understanding HIV-exposed but seronegative individuals. *AIDS Res Hum Retroviruses* 2011,**27**:737-743.

30. Chomba E, Allen S, Kanweka W, Tichacek A, Cox G, Shutes E*, et al.* Evolution of couples' voluntary counseling and testing for HIV in Lusaka, Zambia. *J Acquir Immune Defic Syndr* 2008,**47**:108-115.

31. Song W, He D, Brill I, Malhotra R, Mulenga J, Allen S*, et al.* Disparate associations of HLA class I markers with HIV-1 acquisition and control of viremia in an African population. *PLoS One* 2011,**6**:e23469.

32. Tang J, Cormier E, Gilmour J, Price MA, Prentice HA, Song W*, et al.* Human leukocyte antigen variants B*44 and B*57 are consistently favorable during two distinct phases of primary HIV-1 infection in sub-Saharan Africans with several viral subtypes. *J Virol* 2011,**85**:8894-8902.

33. Merino A, Malhotra R, Morton M, Mulenga J, Allen S, Hunter E*, et al.* Impact of a functional KIR2DS4 allele on heterosexual HIV-1 transmission among discordant Zambian couples. *J Infect Dis* 2011,**203**:487-495.

# APPENDIX

## Investigator's Progress Report

☒ Continuing Review
—OR—
☐ Final Report (When all study activities including data analysis)

—FOR—

☒ Expedited
—OR—
☐ Convened

**NOTE:** ***Not following the format shown below or ignoring or deleting questions will result in deferral of the protocol for IRB review.***

Today's Date (MM/DD/YYYY):09/23/2009

1. Name of Principal Investigator (First, Middle, Last):Richard A. Kaslow
   Email Address:rkaslow@uab.edu
   Campus Address: Department: Epidemiology Building: RPHB Room:220A UAB Zip:0022
   Name of Contact Person:Mary Shirley Title/Role:Program Coordinator
   Phone:5-8698 Fax:4-8665 Email Address:mshirley@uab.edu

2. IRB Protocol Number:X051108005 Protocol Title:CTL and HIV Polymorphisms in Heterosexual Transmission

   Study Sponsor:NIH/Emory University OGCA Tracking # or Link #:225841

3. Briefly describe the purpose of the study (2-3 sentences in non-technical, lay language).
   The purpose of this study is to understand the role that cytotoxic T lymphocyte (CTL) escape plays in HIV transmission and disease pathogenesis. Using a combination of standard and locally modified typing techniques, our laboratory will perform typing for polymorphisms in three sets of human genetic loci: class I HLA (A, B and C); class II (DRB1 and DQB1) and KIR to determine immunogenetic influences on viral evolution and escape. A better understanding of these viral host dynamics should aid in the development of a preventative vaccine in the studied population.

4. Starting Date of Project:06/30/2005 Date of Last IRB Approval: 10/31/2008

5. **Individuals Screened and Entered**

   *For projects conducted at multiple institutions, the numbers stated should be for participants enrolled by the UAB Investigator only.*

   a. Number of individuals **screened** for entry into study since the start of the project? 876
   b. Number of individuals **entered** into the study since the start of the project? 855
   c. Number of individuals **entered** into the study since the last IRB review? 131
   d. Complete the age, sex, and racial/ethnic composition grid below:

| Individuals Screened (Should reflect # in 5.a.) | | | | | Individuals Entered (Should reflect # in 5.b.) | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Racial/Ethnic Composition | Male | | Female | | Racial/Ethnic Composition | Male | | Female | |
| | Age Range | Number Screened | Age Range | Number Screened | | Age Range | Number Entered | Age Range | Number Entered |
| Caucasian | | | | | Caucasian | | | | |
| African American | | | | | African American | | | | |
| Native American | | | | | Native American | | | | |
| Asian | | | | | Asian | | | | |
| Hispanic | | | | | Hispanic | | | | |
| Other | 19-64 | 439 | 16-52 | 437 | Other | 19-64 | 429 | 16-52 | 426 |

**6. For each investigator and staff member involved in the design, conduct and reporting of the research answer the questions below:**

*The following definitions are used for Item #6:*

***Immediate family*** means spouse or a dependent of the employee. *Dependent* is any person, regardless of his or her legal residence or domicile, who receives 50% or more of his or her support from the public official or public employee or his or her spouse or who resided with the public official or public employee for more than 180 days during the reporting period.

***Financial Interest Related to the Research*** means financial interest in the sponsor, product or service being tested, or competitor of the sponsor.

Have each investigator and staff member involved in the design, conduct and reporting of the research answer the questions below:

*(Repeat this section for each individual)*

***Name:*** Richard A. Kaslow

Do you or your immediate family have any of the following? (Check all that apply)

- ☐ An ownership interest, stock options, or other equity interest related to the research of any value.
- ☐ Compensation related to the research unless it meets two tests:
  - Less than $10,000 in the past year when aggregated for the immediate family.
  - Amount will not be affected by the outcome of the research.
- ☐ Proprietary interest related to the research including, but not limited to, a patent, trademark, copyright, or licensing agreement.
- ☐ Board of executive relationship related to the research, regardless of compensation.

***Name:*** Jianming "James" Tang

Do you or your immediate family have any of the following? (Check all that apply)

- ☐ An ownership interest, stock options, or other equity interest related to the research of any value.
- ☐ Compensation related to the research unless it meets two tests:
  - Less than $10,000 in the past year when aggregated for the immediate family.
  - Amount will not be affected by the outcome of the research.
- ☐ Proprietary interest related to the research including, but not limited to, a patent, trademark, copyright, or licensing agreement.
- ☐ Board of executive relationship related to the research, regardless of compensation.

***Name:*** Ilene Brill

Do you or your immediate family have any of the following? (Check all that apply)

- ☐ An ownership interest, stock options, or other equity interest related to the research of any value.
- ☐ Compensation related to the research unless it meets two tests:
  - Less than $10,000 in the past year when aggregated for the immediate family.
  - Amount will not be affected by the outcome of the research.
- ☐ Proprietary interest related to the research including, but not limited to, a patent, trademark, copyright, or licensing agreement.
- ☐ Board of executive relationship related to the research, regardless of compensation.

***Name:*** Kui Zhang

Do you or your immediate family have any of the following? (Check all that apply)

- ☐ An ownership interest, stock options, or other equity interest related to the

research of any value.
- ☐ Compensation related to the research unless it meets two tests:
  - Less than $10,000 in the past year when aggregated for the immediate family.
  - Amount will not be affected by the outcome of the research.
- ☐ Proprietary interest related to the research including, but not limited to, a patent, trademark, copyright, or licensing agreement.
- ☐ Board of executive relationship related to the research, regardless of compensation.

***Name:*** Liangyuan Hu

Do you or your immediate family have any of the following? (Check all that apply)
- ☐ An ownership interest, stock options, or other equity interest related to the research of any value.
- ☐ Compensation related to the research unless it meets two tests:
  - Less than $10,000 in the past year when aggregated for the immediate family.
  - Amount will not be affected by the outcome of the research.
- ☐ Proprietary interest related to the research including, but not limited to, a patent, trademark, copyright, or licensing agreement.
- ☐ Board of executive relationship related to the research, regardless of compensation.

***Name:*** Rakhi Malhotra

Do you or your immediate family have any of the following? (Check all that apply)
- ☐ An ownership interest, stock options, or other equity interest related to the research of any value.
- ☐ Compensation related to the research unless it meets two tests:
  - Less than $10,000 in the past year when aggregated for the immediate family.
  - Amount will not be affected by the outcome of the research.
- ☐ Proprietary interest related to the research including, but not limited to, a patent, trademark, copyright, or licensing agreement.
- ☐ Board of executive relationship related to the research, regardless of compensation.

***Name:*** Wei Song

Do you or your immediate family have any of the following? (Check all that apply)
- ☐ An ownership interest, stock options, or other equity interest related to the research of any value.
- ☐ Compensation related to the research unless it meets two tests:
  - Less than $10,000 in the past year when aggregated for the immediate family.
  - Amount will not be affected by the outcome of the research.
- ☐ Proprietary interest related to the research including, but not limited to, a patent, trademark, copyright, or licensing agreement.
- ☐ Board of executive relationship related to the research, regardless of compensation.

***Name:*** Aimee Merino

Do you or your immediate family have any of the following? (Check all that apply)
- ☐ An ownership interest, stock options, or other equity interest related to the research of any value.
- ☐ Compensation related to the research unless it meets two tests:
  - Less than $10,000 in the past year when aggregated for the immediate family.

- Amount will not be affected by the outcome of the research.

☐ Proprietary interest related to the research including, but not limited to, a patent, trademark, copyright, or licensing agreement.

☐ Board of executive relationship related to the research, regardless of compensation.

***Name:*** **Dale Isabelle**

Do you or your immediate family have any of the following? (Check all that apply)

☐ An ownership interest, stock options, or other equity interest related to the research of any value.

☐ Compensation related to the research unless it meets two tests:

- Less than $10,000 in the past year when aggregated for the immediate family.
- Amount will not be affected by the outcome of the research.

☐ Proprietary interest related to the research including, but not limited to, a patent, trademark, copyright, or licensing agreement.

☐ Board of executive relationship related to the research, regardless of compensation.

**If you checked any of the above,** a financial interest disclosure has to be submitted to or currently on file with the CIRB and the completed CIRB Evaluation has to be available before the IRB will conduct its continuing review.

**7. Since the last IRB review,** have you received any of the following types of information?

a. Multi-center trial reports? ☐Yes ☒No
If yes, attach a copy of any multi-center trial reports not previously forwarded to the IRB, and summarize those reports here:____

b. Data and safety monitoring board reports? ☐Yes ☒No
If yes, attach a copy of any data and safety monitoring board reports not previously forwarded to the IRB, and summarize those reports here:____

c. Interim findings? ☒Yes ☐No
If yes, state both the positive and negative results received to date: **From data on HLA and KIR, we have begun to determine whether variants in those genes separately or jointly influence either acquisition of HIV or control of disease among infected individuals. In a comparison between HIV-exposed seronegatives and their seroconverting partners, no striking effects of HLA or KIR have yet been found. HLA-Bw80Ile, previously associated with favorable outcome in Europeans independently of B*57, may exert a similar effect in Zambians. Analysis is ongoing.**

d. Published literature? ☐Yes ☒No
If yes, attach a copy and summarize the published findings here: **No new findings published.**

e. Any other relevant information regarding this research, especially information about risks associated with the research? ☐Yes ☒No
If yes, attach a copy if applicable, and summarize this information here:____

f. Could any of the information described above relate to the participants' willingness to continue participating? ☐Yes ☒No
If yes, describe here whether and how this information will be provided to participants:____

**8. Since the last IRB review,** have any of the following occurred?

a. Have participants experienced any harms (expected or unexpected)? ☐Yes ☒No
If yes, attach Problem Summary Sheet, and briefly describe here the harms (serious and/or non-serious) experienced by participants:____

b. Have there been any unanticipated problems involving risks to participants or others? ☐Yes ☒No
If yes, attach Problem Report, and briefly describe here the unanticipated problems involving risks to participants or others:____

c. Have you have any problems obtaining informed consent? ☐Yes ☒No
If yes, briefly describe the problems here: **Informed consent was obtained at Emory University pursuant to an Emory IRB-approved consent process.**

d. Have any participants or others complained about the research? ☐Yes ☒No
If yes, briefly describe the number and nature of the complaints:____

e. Have any participants withdrawn from the research? ☐Yes ☒No
If yes, indicate the number of withdrawals and include the reason for each:____

f. Have any obvious, study-related benefits occurred for participants? ☐Yes ☒No
If yes, briefly describe the benefits here:____

g. Have the risks or potential benefits of this research changed? ☐Yes ☒No
If yes, briefly describe the changes here:____

**9. Protocol and/or Informed Consent Modifications**

a. Since the last IRB review, have you made modifications to the protocol or consent process/document that affect the participants?
☐Yes ☒No
If yes, have the modifications been approved by the IRB?
☐Yes ☐No
If yes, provide copies of all cover memos from all amendments approved by the IRB during the approval period.

b. At this time, are you requesting IRB review of any changes to the protocol? ☒Yes ☐No
If yes, describe the requested changes to the protocol here: **Please remove Yirong Ni, Wenshuo Shao and Lili Xie from the study protocol. Drs. Ni and Shao are no longer at UAB and Lili no longer works in Dr. Kaslow's program. Also, last year we added Aimee Merino, a graduate student in the MD/PhD program, to the protocol. We reported that she had not yet settled on a title for her dissertation. She has since chosen "HLA and LRC Gene Polymorphisms in HIV-1 Infection" as the title. Additionally, we would like to add Dale Isabelle to the list of study personnel. Dale is lab manager for the PEII and will do PCR sequencing on this study.**

c. At this time, are you requesting IRB review of any changes to the consent process and/or document? ☐Yes ☒No
If yes, describe the changes here:____
If yes *and* you are requesting review of changes to the consent document:
i. include a copy of the revised consent document with all changes highlighted
ii. include a "clean" copy for the IRB approval stamp
iii. check the following box to indicate revised consent document submitted: ☐

**10. Plans For Future Participant Enrollment**

Is the study open for enrollment? ☒Yes ☐No

If yes, provide narrative of plans for future enrollment here: **Serodiscordant couples who meet entry criteria will continue to be enrolled to the extent that an expanded sample size will benefit the analyses planned; a competitive renewal of the grant has been submitted. The decision on funding will determine how many new subjects will be enrolled and over what period of time.**

If no, complete the following items:

i. Enter date closed to enrollment:____

ii. Choose ONE of the following to describe the status of participants:

☐Participants still on therapy/receiving intervention

☐Participants off study therapy/interventions, in long-term follow-up only

☐Participants off study therapy/interventions, in data analysis only

**Principal Investigator Signature:** Richard D. Gosbee **Date:** 9/23/09

Attach to this Investigator's Progress Report a copy of the current consent form, if applicable. If your research involves children/minors under the age of 19 years; the PI must provide a memo that confirms the previously assigned CRL # or reassigns the CRL and give the reasons the CRL has changed.

**Even though some participants are 18 years of age, they are not considered minors in Zambia and Rwanda, where the age of consent is 16.**

If your study involves Gene Therapy, attach to this Investigator's Progress Report a copy of the current consent form, if applicable, and a signed copy of a memo from your Project Review Panel addressing the following questions:

1. Has the Panel's assessment of the risk-benefit ratio of this project changed? ☐Yes ☐No
   If yes, please explain.
2. Did the Panel have any recommendations regarding the protocol or the consent form? ☐Yes ☐No
   If yes, please explain.
3. If the study includes children/minors under the age of 19 years, provide a memo the previously assigned CRL # and have the Panel either confirm it or reassign it. If the panel chooses to reassign the CRL, they must give the reasons the CRL has changed.