A Candidate Gene Study of the Association of TLR4 Polymorphisms rs4986790 and rs4986791 with HIV Disease Progression and Response to HAART Treatment in Men

by
Yuan He

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Health Science

Baltimore, Maryland
April, 2015
Abstract

**Background.** Host genetic factors have been shown to be associated with the diversity in the HIV/AIDS disease progression. In addition, lipopolysaccharides (LPSs) are recognized by Toll-Like Receptor 4 (TLR4), and are associated with the HIV viral load in macrophages. Furthermore, TLR4 polymorphisms rs4986790 and rs4986791 have been shown to be associated with the HIV viral setpoint. Together, these findings suggest that these TLR4 polymorphisms are associated with HIV disease progression.

**Objective.** We studied the association of the TLR4 polymorphisms rs4986790 and rs4986791 with HIV viral setpoint, and the virologic and immunological responses to HAART treatment.

**Methods.** We designed both cross-sectional and prospective components of this cohort study to utilize men enrolled in the Multicenter AIDS Cohort Study (MACS) who had been genotyped for single nucleotide polymorphisms (SNPs) in the TLR4 region. We applied linear regression to model the HIV viral setpoint, and log-binomial regression as well as Poisson regression with robust estimation of variance to model the virologic and immunological responses to HAART.
**Results.** The results suggest that the two TLR4 SNPs have protective effects in terms of decreased HIV setpoint in a recessive model. The mean log10 HIV viral setpoint among men with rs4986791-TT was 0.11 lower than that among men with at least one C allele, and the mean log10 HIV setpoint among men with rs4986790-GG was 0.30 lower than that among men with at least one A allele. For the virologic response to HAART, we found that the probability of achieving an undetectable HIV RNA level within two years of HAART initiation was 47% higher among men with rs4986791-TT compared to men with at least one C at rs4986791 adjusting for cofactors, while the probability of achieving an undetectable HIV RNA level within two years of HAART initiation was 40% higher among men with rs4986790-GG compared to men with at least one A at rs4986790 adjusting for cofactors. In contrast, the two TLR4 SNPs were not associated with the immunological response of achieving 100 cells/μl increase in CD4+ T cells within two years following HAART treatment.

**Conclusions.** The results from this study are in accordance with previous findings that rs4986790 and rs4986791 were associated with HIV viral load, and they extended that
observation by suggesting that TLR4 genotype might also be associated with achieving undetectable viral load after HAART treatment. This novel observation needs to be confirmed in other cohorts.
Acknowledgement

I would like to express my sincere thanks to Dr. Eric Seaberg, my thesis advisor and academic advisor. During the thesis, Dr. Seaberg helped me through the whole process of obtaining data, carrying out analysis, writing and checking this manuscript. I would like to thank Dr. Seaberg for spending the time discussing this thesis with me and very kindly gave me many helpful suggestions. He showed me the correct way of doing research, which are thorough investigation, deeper thinking as well as extra attention to details. Besides, Dr. Seaberg has been very supportive since I was in the general epidemiology track. He recommended good courses to me and helped me organize the class schedule to fulfill requirement for my degree.

I would also like to thank Dr. Keri Althoff who is the second reader of my thesis and gave me very helpful comments, and Dr. Priya Duggal who gave me very instructive and useful information about the imputation part in the thesis. I am extremely grateful for your assistance and suggestions throughout my thesis.

Thank you to the staff at the department of epidemiology and school of public health, for always giving the students care and consideration. I am very proud to be studying here at Bloomberg School of Public Health.

Finally, I cannot thank enough to my parents and my family in China as well as my fiancé in Philadelphia. Thank you for always being there encouraging and supporting me. Thanks!
# Table of contents

Abstract .................................................................................................................................................................... ii

Acknowledgement .................................................................................................................................................. v

Background ............................................................................................................................................................. 1

  *Human immunodeficiency virus infection and AIDS* ......................................................................................... 1
    *Human immunodeficiency virus type 1 (HIV-1)* ................................................................................................. 1
    *HIV/AIDS progression* ......................................................................................................................................... 4
    *Diversity in progression to AIDS* ....................................................................................................................... 6
    *HAART* ............................................................................................................................................................... 8

  *Toll-like receptor 4 (TLR4)* ................................................................................................................................. 9
    *Toll-like receptor 4 and its polymorphisms* ........................................................................................................ 10
    *rs4986791 and rs4986790 in TLR4* ................................................................................................................ 11

Aim and Hypotheses ............................................................................................................................................. 16

Methods .................................................................................................................................................................. 18

  *Study Population: The MACS cohort* .................................................................................................................... 18
  *Study design* ........................................................................................................................................................ 19
  *Assessment of the outcomes* ............................................................................................................................. 19
  *Primary exposure: TLR4 SNPs* ........................................................................................................................... 21
  *SNP Imputation* ................................................................................................................................................. 22
  *Potential confounding factors* .......................................................................................................................... 23
  *Statistical Analysis* .......................................................................................................................................... 26

Results ..................................................................................................................................................................... 29

  *Aim one: HIV viral setpoint* ............................................................................................................................. 30
    *Characteristics of the study population* ........................................................................................................... 30
    *Univariate analysis* .......................................................................................................................................... 31
    *Multivariate analysis* ...................................................................................................................................... 32

  *Aim two: Virologic response to HAART* ........................................................................................................ 34
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics of the study population at baseline</td>
<td>34</td>
</tr>
<tr>
<td>Univariate analysis</td>
<td>35</td>
</tr>
<tr>
<td>Multivariate analysis</td>
<td>37</td>
</tr>
<tr>
<td>Aim three: Immunologic response to HAART</td>
<td>39</td>
</tr>
<tr>
<td>Univariate analysis</td>
<td>39</td>
</tr>
<tr>
<td>Multivariate analysis</td>
<td>40</td>
</tr>
<tr>
<td>Discussion</td>
<td>41</td>
</tr>
<tr>
<td>List of Figures</td>
<td>49</td>
</tr>
<tr>
<td>List of Tables</td>
<td>50</td>
</tr>
<tr>
<td>References</td>
<td>67</td>
</tr>
<tr>
<td>Curriculum Vitae</td>
<td>75</td>
</tr>
</tbody>
</table>
Background

*Human immunodeficiency virus infection and AIDS*

Acquired immune deficiency syndrome (AIDS) is caused by infection of human immunodeficiency virus, often strain 1 (HIV-1), and is mainly transmitted through three routes: unprotected sex with HIV-infected person; transfusion of contaminated blood; and birth from an infected mother. The hidden risk of transmission has made the disease silently spread out in the population. In 2012, the Joint United Nations Program on HIV/AIDS (UNAIDS) reported an estimation of 35.3 (32.2–38.8) million people living with HIV globally. Infected individuals gradually lose the functional immune defense system without treatment and will ultimately develop AIDS and die.

*Human immunodeficiency virus type 1 (HIV-1)*

HIV-1 is an enveloped retrovirus, which uses RNA as its genetic material and undergoes reverse transcription and DNA integration in host cells to replicate. The
envelope protein complex of HIV-1 is synthesized as a polyprotein (gp160) that is
cleaved intracellularly to a heterodimer of surface subunit gp120 and trans-membrane
subunit gp41, which two are non-covalently linked. Careful sequencing of the HIV
genome revealed seven genomic structural elements containing ten genes.

Intensive research has revealed how HIV-1 enters cells. First, the viral surface
glycoprotein gp120 of HIV-1 interacts with glycoprotein on CD4 cell surface. This
interaction induces exposure and/or formation of binding cite for specific chemokine
receptors on the CD4 cell surface, which are then bound to the HIV-1 surface protein
and "opened a door" for the virus to enter the cells. Several chemokine receptors involved
in this process have been recognized, including CCR5 used by macrophage-tropic
(M-tropic) HIV viruses and CXCR4 used by T cell line-tropic (T-tropic) HIV viruses.
This interaction between HIV-1 surface protein and cell surface glycoprotein as well as
co-receptors leads conformational changes in gp41, which results in fusion of the cell and
viral membranes. The entry process of HIV-1 determines the major target cells for HIV-1 being primarily T cells, macrophages and probably dendritic cells (DCs).

After membrane fusion, the viral capsid containing two strands of RNA and enzymes including reverse transcriptase (RT) and virus-encoded integrase (IN) protein is released into the CD4 T cells, followed by two steps to complete the virus replication. First, the RNA undergoes reverse transcription by RT to form the double-stranded DNA copy. Second, the IN protein initiates the integration of the HIV DNA into the host cell. The HIV integration in CD4 cell genome is strongly favored in active transcription regions, which may lead to active gene expression⁸.

HIV infection in the host cells causes hyper-activation of the immune system that gradually consumes T cell supplies, directly killing the CD4 T cells and destroying the architecture of immune system⁹. After sensing the presence of the viral DNA, infected CD4 T cells undergo apoptosis to sacrifice themselves¹⁰. At the same time, extracellular and cell-surface components on the HIV infected cell can conduct indirect killing of the
bystander cells that are not infected\textsuperscript{11, 12}. Failure to completely control the virus replication leads to gradual loss and ultimately depletion of CD4 T-cells in lymphoid tissues and dysfunction of the immune system\textsuperscript{13}.

\textit{HIV/AIDS progression}

HIV infection generally passes through several phrases\textsuperscript{14, 15}, and progression can be measured by plasma RNA viral load and CD4 T cell count as well as by HIV associated infections such as fever and thrush\textsuperscript{16}.

First there is a period of about 10 days called the eclipse phase following the virus transmission. During this period, viral RNA is not detectable in the plasma. Studies have shown that innate immune and inflammatory responses are activated in the eclipse phase of SIV infection\textsuperscript{17}.

Second is the acute infection phase where the virus concentration grows rapidly through replication, resulting in exponentially increased plasma virus levels\textsuperscript{18}. This usually happens about two weeks after infection, and this period is called viremia. HIV is
detectable in the form of HIV p24 antigen or HIV RNA. During the viremia period, generally minor symptoms arise including genital or oral ulcers. About one to three weeks after the period of viremia, antibodies towards HIV can be detected, and this process is called seroconversion. After seroconversion, HIV RNA normally becomes undetectable with the existence of early antibody-antigen complex\textsuperscript{19}.

The acute infection phase is followed by a clinically latent stage, where viral presence and replication is low to absent in PBMC, but still very active in lymphoid organs. Some period after infection, stable viraemia was established when the plasma viral load settles to a stable level. This is called HIV viral setpoint. The length of the period between the onset of HIV infection and reaching HIV viral setpoint varies from 3-6 month\textsuperscript{20}, 21-119 days\textsuperscript{21} and so on by different standards. Since the viral set point is associated with the speed of the disease progression as well as the risk of transmission, it is sometimes treated as a surrogate for the disease progression\textsuperscript{22}. 

5
AIDS develops when the immune system is severely damaged and one of the several events including opportunistic infections (OIs) occurs. OIs are the most common cause of death among people with HIV/AIDS among those untreated. When the CD4 count falls below 200 cells/mm\(^3\), some devastating OIs may happen, including Tuberculosis, Toxoplasmosis of brain and multiple forms of lymphoma\(^{23}\). Since the effective antiretroviral therapies were introduced in the mid 1990’s, the cause of death among people with HIV/AIDS has changed with approximately 75% of the deaths now being attributed to diseases other than AIDS, with the majority being cancer\(^{24}\).

**Diversity in progression to AIDS**

Without treatment, people infected with HIV-1 have variable rates of progression to AIDS that generally fall into one of three categories. First, there is rapid progression, where AIDS develops within three years of infection. Second, intermediate progression includes the people in whom AIDS develops about three to ten years after seroconversion. Third, long-term non-progression is the label used to describe HIV-infected individuals
who maintain a high CD4 T cell level and do not develop AIDS\textsuperscript{25}. Host characteristics including genetic factors and virology may also contribute to these differences in response to HIV.

HIV disease progression can be influenced by a variety of factors, including genetic factors and age\textsuperscript{26, 27}. Considering genetic factors, three main categories have been identified. First are genes encoding chemokine receptors on the cell surface, including some genes that encode the co-receptors during HIV entry. One of the most dramatic findings is a homozygous defect of 32-nucleotide deletion in beta-chemokine receptor 5 (CCR5) gene. CCR5 on CD4 T cell surface is the co-receptor for macrophage-tropic HIV-1. The CCR5 delta-32 deletion has been proved to convey resistance to HIV-1, so that it can prevent transmission\textsuperscript{28} and delay disease progression\textsuperscript{29}. Mutations in the promoter region of CCR5 are also associated with HIV/AIDS progression due to changed CCR5 expression on the cell surface\textsuperscript{30}. The CCR2B-64I mutation has been found to delay HIV/AIDS progression\textsuperscript{31}, although this has not been found to be universally true\textsuperscript{32}. 

Polymorphism of SDF1 (the principle ligand for CXCR4), SDF1-3'A, may also convey protection and slow the progression\textsuperscript{33}. Other polymorphisms that affect HIV disease progression include genes encoding human leukocyte antigens (HLA) include A*24, B*37, B*56\textsuperscript{34}, and mutations in other genes that control innate immune and adaptive immune systems, such as TLR4.

**HAART**

Highly active antiretroviral therapy (HAART) combines inhibitors of HIV-1 protease and reverse transcriptase. HAART can suppress HIV virus replication, restore CD4 T cell reactivity, slow down the HIV disease progression and reduce the risk of opportunistic infections by many studies\textsuperscript{35, 36}.

Though HAART is effective in most HIV infected patients, some other factors can impact the effect of HAART. For example, one study in HIV infected children showed that younger children have more rapid growth reconstitution following HAART treatment than older children\textsuperscript{37}. Weight may also have an effect on response to HAART. A
cross-sectional study showed that obesity and overweight was more common than wasting among the HIV infected subjects in Philadelphia\textsuperscript{38}. However, because of the complex impact of obesity on immune system\textsuperscript{39}, it is not clear if obesity is related to response of HAART. Smoking has also been shown to impact people's health in many ways, and it is possible that it has an impact on the immune effect. Studies have found poorer viral responses and immunological to HAART for the smokers among women\textsuperscript{40}, as well as higher risk of opportunistic respiratory infections for smokers\textsuperscript{41}.

\textbf{Toll-like receptor 4 (TLR4)}

Toll-like receptors (TLRs) are one kind of germline-encoded pattern-recognition receptors (PRRs) that play a critical role in innate immunity as well as in adaptive immunity. The most important feature about TLRs is that they can recognize pathogen-associated molecular patterns (PAMPs) from foreign pathogens\textsuperscript{42}. They are composed of a single intracellular Toll/interleukin-1 receptor (TIR) and multiple extracellular leucine-rich repeats (LRRs)\textsuperscript{43}. Different LRRs can recognize different sets
of ligands and then trigger corresponding downstream cellular signal with the help of
MyD88. Activation of TLRs initiates innate immune response by producing NF-kapaB
and inducing inflammatory cytokines such as NF-kapaB, TNF-alpha, IL-1, IL-6 and
others\textsuperscript{44}. TLRs can also mediate adaptive immune responses by controlling multiple
dendritic cell populations\textsuperscript{45}.

\textbf{Toll-like receptor 4 and its polymorphisms}

TLR4 is a very important member of the TLR family. Human TLR4 is located in
chromosome 9 from 120,466,610-120,479,149 on the forward strand on the latest
genome assembly GRCh37\textsuperscript{46}. It has four exons and four transcript mRNA variants, each
with the length between 5000bp to 6000bp\textsuperscript{47}.

TLR4 is expressed in a variety of immune cells, including dendritic cells (DCs) and
macrophages, and also CD4 T cells\textsuperscript{48}. It recognizes lipopolysaccharides (LPSs), which
are the major outer surface membrane components of Gram-negative bacteria\textsuperscript{49}. In
addition to recognizing LPS, TLR4 is also involved in recognizing chlamydial heat shock
protein (Hsp) 60, fibrinogen, virus envelope proteins such as respiratory syncytial virus (RSB) and mouse mammary tumor virus (MMTV), etc. However, the molecular structures for the latter effect have not been realized.

Among the many studies the examined the association between phylogenetic diversity in TLR4 and diseases, significant diversity in TLR4-mediated infectious disease outcome has been demonstrated. This is because that mutations in either the extracellular part of TLR4 or in the down-stream signaling pathway may change the immune responses to the foreign microbe, thus causing different outcomes. Also, there are growing studies including population studies and molecular studies identifying association between TLR4 sequence variation and different kinds of cancer.

**rs4986791 and rs4986790 in TLR4**

RS4986791 (C1196T) is a non-synonymous SNP in TLR4 gene with the minor allele frequency being 0.0407 among the 1088 worldwide individuals in the 1000Genome phase 1 genotype data. The minor allele of T is observed 204 times in the sample
The mutation from C to T leads to amino acid substitution of Thr399Ile. This missense mutation lies in the fourth exon of TLR4, and changes its extracellular domain.

RS4986790 (A896G) is also a non-synonymous SNP in TLR4. The minor allele frequency of G is 0.0599 among the 1088 worldwide individuals in the 1000Genome phase 1 genotype data. The minor allele of G is observed 300 times in the sample population. The A896G mutation leads to substitution of a conserved acid residue at amino acid 299 from aspartic to glycine (Asp299Gly). It also changes the extracellular domain.

These two mutations (rs4986791 and rs4986790) in TLR4 are predicted to affect the ligand-binding part in the extracellular domain, and have been shown to be associated with some human diseases. It has been proved that co-segregating polymorphisms of rs4986791 and rs4986790 in TLR4 are associated with endotoxin hypo-responsiveness and more susceptibility to Gram-negative bacteria in human. A cohort study found that
these two mutations are associated with lower HIV RNA levels in ARV-naïve HIV seroconverters. A meta-analysis of association between polymorphisms in TLR4 and susceptibility to cancer identified the two SNPs (rs4986791 and rs4986790) in TLR4 associated with increased risks. Both rs4986791 and rs4986790 mutations are found to exist more frequent in the mild asthma group with atopy, and the TLR4-A896G SNP is proved to be associated with decreased risk of post-traumatic sepsis using a cohort study. The diversity and large number of diseases that are found to be associated with TLR4 indicate its wide effect in immune responses.

At the same time, there are studies showing a lack of association between the two TLR4 mutations and phenotypes or diseases. For example, case-control study has found no association between rs4986790 and ankylosing spondylitis (AS), an autoimmune disorder. Another case-control study has shown that there was no significant association between rs4986791 and rs4986790 mutations and necrotizing enterocolitis (NEC), which was hypothesized to be caused by inadequate response of the innate immune system to
bacterial antigens, in very low birth weight (VLBW) infants. These results might be due to the small sample size, or because of the complex nature of these phenotypes that is controlled and mediated by many variations, in which situation the effect of two mutations only have very weak influence alone that is hard to be separated.

The variability in HIV/AIDS disease progression among HIV infected individuals indicates that disease progression might be controlled by several key factors, including host genetic factors. TLR4, being able to recognize LPSs and initiate innate immune responses, may change the risk of HIV infection and HIV/AIDS progression with its mutations. This was based on the knowledge that LPS has an inhibitory effect on HIV-1 infection in human macrophages, which are one of the main targets of HIV virus infection. Thus, macrophages with stronger ability of recognizing LPS might be less likely to get infected by HIV-1. In a prior MACS study, it was demonstrated that two TLR4 polymorphisms rs4986790 and rs4986791 being associated with slower HIV
progression. In this thesis we extended that study to examine the association of these two
SNPs with other measures of HIV progression and the response to HAART.
Aim and Hypotheses

The goal of this study was to investigate the association of two SNPs (rs4986791 and rs4986790) in TLR4 gene with markers of HIV disease severity and the immunologic and virologic responses to HAART among HIV-infected men. Since HIV viral load at the setpoint is an important indicator of the HIV disease severity and progression, we first examined the association between the two TLR4 SNPs and the HIV viral setpoint. We then examined the association of two TLR4 SNPs on the virologic response and immunologic response to initial HAART treatment.

- To achieve these goals, we designed this study to address the following three aims by testing the corresponding hypotheses: Aim 1. To examine the association of TLR4 genotype with HIV viral load, which is a marker of HIV disease progression.

Hypothesis 1: TLR4 polymorphisms at rs4986791 and rs4986790 are associated with lower HIV viral setpoint.
• Aim 2. To examine the association of TLR4 genotype with the virologic response to HAART.

Hypothesis 2: TLR4 polymorphisms at rs4986791 and rs4986790 are associated with a higher probability of achieving undetectable HIV RNA within two years following HAART initiation.

• Aim 3. To examine the association of TLR4 genotype with the immunologic response to HAART.

Hypothesis 3: TLR4 polymorphisms at rs4986791 and rs4986790 are associated with a higher probability of a 100 cell/uL increase in CD4+ T cells within two years following HAART initiation.
Methods

Study Population: The MACS cohort

The Multicenter AIDS Cohort Study (MACS), started in 1983, is an ongoing HIV study at four sites including the Johns Hopkins School of Public Health, the University of Pittsburgh Graduate School of Public Health, Northwestern University School of Medicine, and the UCLA School of Public Health. The overall aim for MACS was to investigate the disease of HIV/AIDS within the population of HIV infected men. MACS has enrolled 6,972 men who have sex with men (MSM) for semiannual interview, and has collected 1,490,995 specimens and 86,883 person-years of information through May 2011. The wide range of information collected has made this study very valuable and instructive for population-based studies of HIV/AIDS. For our study that examined the associations of the two SNPs with different HIV disease measures, we included the subgroup of MACS participants who had the TLR4 SNPs genotyped in a larger MACS-wide GWAS study.
**Study design**

We tested hypothesis one using a cross-sectional study design among 1446 MACS participants who were genotyped for rs4986791 and rs4986790. For hypothesis two and hypothesis three we used a prospective study design that included the 561 HAART initiators who also were included in the MACS GWAS study.

**Assessment of the outcomes**

After some period of HIV infection, the HIV viral load reaches a plateau that was called the HIV viral setpoint, and then the viral load will fluctuate around this setpoint value. The HIV viral load at setpoint was obtained between one and two years after HIV infection while the individual remained treatment free. For the MACS population, the HIV RNA load was measured by one of the four kits: Chiron 2nd generation ($\geq 500$ copies/ml), Roche 2nd generation ($\geq 400$ copies/ml), Roche ultra-sensitive 2nd generation ($\geq 50$ copies/ml) and Roche Cobas TaqMan ($\geq 20$ copies/ml). Among the population of 1169, more than 99% participants have detectable virus load without
treatment. As a result, the HIV-1 viral load can be treated as a continuous variable or as a categorical variable.

The viological response to HAART was defined as a binary variable indicating whether or not HIV RNA load became undetectable within two years following HAART treatment. Since the majority of the population of 561 people under this study was tested using the Roche 2nd generation kit, the virologic event of undetectable RNA virus load was set to be less than 400 copies/ml. There was only one person measured by the Chiron 2nd generation kit, whose RNA virus load was larger than 10000 copies/ml and thus can be detected using any one of the kits. All RNA virus load measurements below 400 copies/ml were treated as undetectable virus load. Eligible participants had detectable HIV RNA at baseline, and those who achieved the event of undetectable HIV RNA load within two years following HAART initiation were defined to have had a virologic response.

Immunological response was measured by CD4+ T cell count, another important indicator for the disease progression. Achieving a 100 cells/µl increase in CD4+ T cell
count compared to the last measurement within 6 months before HAART initiation is clinically important\textsuperscript{72}, and thus was considered the immunological event in the analysis.

The immunological response was also treated as binary variable.

\textbf{Primary exposure: TLR4 SNPs}

The genotypes for the two candidate TLR4 SNPs rs4986791 (chr9: 120475302 on GRCh37 genome assembly) and rs4986790 (chr9: 120475602 on GRCh37 genome assembly) were assayed by three GWAS assays, including Illumina 1MDuo, Illumina 1Mv1 and Illumina HH550.

For each hypothesis, these SNPs were coded four different ways to test different genetic models. First we fit a categorical variable that allows for there to be a different effect for each genotype. The second represents dominant effect of the minor allele with a binary variable to indicate the presence of at least one minor allele. The third model tests the recessive effect of the minor allele with a binary variable to indicate the presence of two copies of the minor allele. The fourth was an ordinal variable to evaluate the effect of
having an increasing number of the minor allele.

**SNP Imputation**

Imputation was performed to estimate and validate genotype information for rs4986791 and rs4986790 using the program IMPUTE2\(^{73, 74}\) and information about the other SNPs included on the gwas chips that were near (+/− 10kb) the locations of rs4986791 and rs4986790.

Because the majority of the participants (87.5%) belong to white non-Hispanic ethnic, we set the reference panel to the HapMap CEU population that includes Utah residents with northern and western European ancestry\(^{75}\). The -Ne parameter in IMPUTE2 controls the effective size of the population and is used to choose the model of linkage disequilibrium pattern. According to the manual of IMPUTE2, Ne was set to 11418 to represent the LD pattern in the CEU population\(^{76}\). The 1000 Genomes Phase 3 genome assembly\(^{77}\) that corresponds to the NCBI build 37 (hg19) genome assembly was used as the reference panel for imputation. Since the NCBI build 37 (hg19) genome assembly is
the same as the UCSC GRCh37 genome assembly\textsuperscript{78}, we set the imputation boundaries for rs4986791 and rs4986790 at "120465302 120485602" which is +/-10 kb from the target SNPs using the –int parameter in IMPUTE2. The -call_thresh parameter was set to 0.9 by default, indicating that genotypes from the input that have maximum probability lower than 0.9 would be treated as missing data.

\textbf{Potential confounding factors}

The potential confounders that we included for hypothesis one were race, age, BMI, smoking status and cumulative pack-years of cigarettes. For hypothesis two and hypothesis three, additional confounders were tested, including whether or not the participant had developed AIDS prior to receiving HAART treatment, the HIV viral load and CD4+ T cell count at the last measurement within two years before HAART treatment.

Since this analysis aims to study the potential effect of genotype, there is the possibility of population stratification that interferes with the effect of the interested
genotype. So race was studied and adjusted for in all the three hypotheses. Age, BMI and smoking degree were potentially related to the functionality of immune system and the individuals' health status, and thus were possible to be related to the HIV setpoint viral load as well as the viological and immunological responses to HAART.

To quantify the degree of smoking, two variables were integrated including the smoking status at HAART initiation and cumulative pack-years of cigarettes. The categorical smoking status includes never smoked, previously smoked, and currently smoking at the time of HAART intimation. However, it was not a thorough description of the effect of smoking cigarettes. For example, the sick people might quit smoking at the time of HAART initiation. Cumulative pack-years of cigarettes provided us the record of smoking history, and the heavy smokers in the previous period may be left with the impact of cigarettes after they quit. So these two variables were integrated to provide a more thorough measure of smoking.
In an HIV infected population, the mode of HIV infection might influence the HIV disease progression as well as the treatment efficacy, so we also adjusted for a history of injection drug use.

The outcome variables for the immunologic and virologic responses to HAART initiations were defined according to changes of CD4 count and HIV RNA level, respectively. For outcomes defined in this manner, the probability of having a sufficiently large change and being classified as having experienced the outcome is likely to be associated with the baseline levels of CD4 and HIV RNA. Therefore, we adjusted for the pre-HAART initiation values of the HIV disease markers. Finally, we also adjusted for whether or not the participants had previously been diagnosed with AIDS to account for the possibility that the response to HAART might differ between those with vs. without a prior AIDS diagnosis.
Statistical Analysis

Standard descriptive statistics were used to characterize the study cohorts, and then univariate analysis using linear regression or log-binomial regression was carried out to choose the appropriate set of cofactors. Then we conducted multivariate regression analyses for the three hypotheses.

For the first hypothesis, we initially defined the outcome using a binary variable defined as whether or not the HIV setpoint was higher than the median value of the population. To avoid the problem of highly skewed distribution of HIV viral load, log 10 transformed HIV viral load was used. In this way, the median was 4.3 log10 copies/ml. However, none of the men who were homozygous for the minor allele had the outcome, which made it impossible to fit the log-binomial models. Therefore, we decided to fit the model using the continuous response variable of HIV virus setpoint. The lm() function in R provided ways to fit the linear regression model.
To characterize and compare the virologic and immunologic response to HAART between people with different SNP genotypes, log-binomial regression was used to estimate the relative risk (RR) when the probability of the outcome which was >10%. To overcome the problem of non-convergence in some circumstances, we fit the models using Poisson regression with robust estimation of the variance approximation of the log-binomial regression model. The sandwich package from R provided the function `vcovHC` to perform the robust estimation of the variance.

Four models were constructed for each of the hypotheses, corresponding to four possible ways of having an effect for the SNPs. They included categorical model that used the mutation genotype as a categorical variable, dominant effect model that assumed effect of the mutation with the presence of at least one major allele, recessive effect model that assumed effect of the mutation with two copies of the minor allele, and an ordinal model to evaluate the effect of having an increasing number of the minor allele.
Thus, we used linear regression to model the HIV viral setpoint outcome, and Poisson regression with robust estimation of variance to fit the viological and immunological outcomes. All analysis was performed using R 3.0.2. Statistical significance was defined as a two-sided p-value less than 0.05.
Results

There were 1446 MACS participants who had genotype information for rs4986791 and/or rs4986790 measured by the three arrays; 263 men measured by the 1MDuo array for both SNPs, 806 men measured by the 1Mv1 for both SNPs, and 498 men measured by the HH550 array for rs4986791. Because rs4986790 was not included on the HH550 array, we use imputation to determine the most likely rs4986790 genotype for the men tested using this array.

There were 16, 15 and 6 SNPs within +/- 10kb of the rs4986791 and rs4986790 measured by 1MDuo array, 1Mv1 array and HH550 array respectively. The statistics about imputation quality are shown in Table 1. The values in the info column and the certainly column are very close to or equal to 1, indicating that the three imputations all have achieved high certainty. After imputation, the input genotypes at the SNP were masked and imputed by the other genotypes. The concordance between the input genotypes and the imputed genotypes for each SNP was stored in the concord column. The concordance values were all above 0.8 (except rs4986791 by HH550), implying that
the imputed genotypes match well with the original genotypes. Since rs4986791 was not measured by HH550 assay, the masking experiment could not be done, and thus the statistics were marked as NA. The imputation results were shown in Table 2. One individual measured by HH550 array has a "00" genotype in rs4986790, indicating poor permutation quality, and was considered as missing data for all analyses. Because of this and the fact that the two SNPs are in extremely high linkage disequilibrium (LD is ~1), we focused mainly on rs4986791 in all analyses.

**Aim one: HIV viral setpoint**

**Characteristics of the study population**

Among the 1446 MACS participants with TLR4 genotype data, the 1169 men who had HIV viral setpoint recorded in the MACS database were included in this analysis of the HIV viral setpoint. The distribution of the participants' characteristics by rs4986791 genotype is shown in Table 3. None of the characteristics evaluated in this analysis were significantly associated with having at least one copy of the minor allele T.
**Univariate analysis**

The probability of having an HIV viral setpoint greater than the median of 19953 copies/ml is shown in table 4 stratified by TLR4 genotype. The probability of having an elevated setpoint was lowest among the men who were homozygous for the minor allele for both SNPs. Specifically, only 1/9 (11%) of those with rs4986791-TT and 0/7 (0%) of those with rs4986790-GG had an elevated setpoint. Unfortunately, the small numbers, and particularly the “0” cell for rs4986790, made it impossible to fit both the log-binomial and exact logistic regression models.

Given the analytical difficulty mentioned above, we decided to model setpoint as a continuous variable. Figure 1 displays the distribution of the log10 HIV viral setpoint data. The bump between 2 and 3 reflects those who had a setpoint below 400 copies/ml that is the lower limit of detection of the most commonly used HIV RNA assay. More specifically, this corresponds to 28 individuals whose HIV viral setpoint was undetectable. Since only 1.6% (28 out of 1169) of the participants were recorded as
undetectable, we decided to include them in the analysis with the HIV setpoint equal to 400 copies/ml. Other than that little bump, the distribution of HIV viral load was approximately a normal distribution, and thus could be tested using linear regression.

The univariate associations of the TLR4 SNPs and potential confounders with the log10 HIV viral setpoint are shown in Table 5. Although the differences were not statistically significant, men with rs4986791- TT (median=4.11, IQR: 3.80 - 4.20) had lower HIV viral setpoint compared to men with CT or CC. The same was observed for men with rs4986790-GG compared to those with AG or AA. For age, the older men tended to have a higher HIV viral setpoint compared to younger men. Other covariates didn't show an apparent association with the HIV setpoint.

**Multivariate analysis**

Table 6 shows the results of multivariate regression analysis assuming a recessive gene model. The adjusted mean log10 HIV viral setpoint for the rs4986791-TT group was 0.11 (95% CI: -0.42 to 0.68) lower than among the men with at least one C allele.
Among men who were rs4986790-GG, the age and race-adjusted mean HIV viral setpoint was 0.30 (95% CI: -0.88 to 0.28) lower than that among men with at least one A allele. These results indicate the possibility of the recessive model being associated with lower HIV setpoint.

In addition to the recessive model presented above, there are three other genetic models that can be explored. They include a) different effects for each of the three genotypes, b) dominant effect of the minor allele, and c) ordinal effect corresponding to the number of minor alleles. We explored these alternate gene effect models, and the adjusted results of all four approaches are summarized in Table 7. These results indicate that the largest effect of TLR4 genotype on HIV viral setpoint is through the recessive model.
Aim two: Virologic response to HAART

Characteristics of the study population at baseline

For the aim of measuring virologic response to HAART, the population with HAART initiation records was used, which included 561 individuals. To test the virologic response, 45 individuals didn't have record for HIV RNA load at the last measurement before HAART initiation and were dropped. Then 49 individuals who had undetectable HIV RNA load at the baseline measurement were not eligible to achieve the event of becoming undetectable HIV RNA load. 19 individuals didn't have the records for HIV RNA load within two years of HAART initiation, that is, didn't have the record of the outcome of viological response. After these three filtering, 448 individuals were left eligible for the analysis. The distribution of the participants' characteristics by rs4986791 genotypes is shown in Table 8. None of the characteristics evaluated in this analysis were significantly associated with having at least one copy of the minor allele T (all p-values were larger than 0.05). However, men with at least one copy of the
rs4986791-T allele were more likely to have a history of smoking, a difference that was borderline significant (p=0.053).

**Univariate analysis**

Log binomial regression for the event of developing a viological response within two year of HAART initiation was conducted with each one of the covariates to obtain the univariate results shown in Table 9. For the TLR4 genotypes, we used Poisson regression with robust estimation of the variance instead of the log-binomial model to avoid the problem of incorrect standard error estimation due to small sample size in the TT/GG cell.

The two primary exposures showed significant protective effect in a recessive model. Men with rs4986791-TT had the risk ratio of 1.36 (95%CI: 1.28 to 1.44) compared to men with CC, while rs4986791-CT didn't show significant effect. At rs4986790, men with GG had the risk ratio of 1.35 (95%CI: 1.28 to 1.43) compared to men with AA, while men with AG didn't show significant effect.
Age and CD4+ T cell count at the baseline showed a significant positive effect in the univariate tests. As shown in table 9, men >50 years old were 1.29 (95% CI: 1.11 to 1.50) more likely to achieve an undetectable HIV RNA level compared to men <40 years old. Furthermore, men between 40 and 50 years old were 1.16 (95% CI: 1.01 to 1.36) times more likely to suppress HIV compared to men <40 years old. Higher CD4+ T cell count at the last measurement before HAART initiation showed a protective effect. Taking the participants in the category of CD4+ T cell count of <200 cells/μl as the reference group, the participants in the category with highest level of CD4+ T cell count of >500 cells/μl have the risk ratio of 1.38 (95% CI: 1.57 to 1.62) and the middle category has the risk ratio of 1.35 (95% CI: 1.18 to 1.57).

AIDS development prior to HAART initiation and HIV RNA load at baseline showed deleterious effect on the virological response, as demonstrated the findings that participants with prior AIDS had the risk ratio of 0.66 (95% CI: 0.52 to 0.80) compared to those who didn't develop AIDS, and the population with log10 of HIV RNA load at the baseline > 5 has the risk ratio of 0.82 (95% CI: 0.70 to 0.96) compared to those who had
less than 4 at the baseline.

Although there was no statistically significant difference found between the three ethnic groups, race should still be included in the model to account for potential impact of genomic background differences. However, neither BMI nor smoking associated with the viological event, so they were not included in the multivariate analysis.

**Multivariate analysis**

Table 10 shows the results of multivariate regression analysis assuming a recessive gene model using Poisson regression with robust estimation of the variance. The risk ratio of developing the viological event for the rs4986791-TT group was 1.47 (95% CI: 1.10 to 1.98) compared to the men with at least one C allele, adjusting for race, age at the HIV positive day, development of AIDS at HAART initiation, CD4+ T cell count and Log10 of HIV RNA load at the last measurement before HAART initiation. Among men who were rs4986790-GG, the adjusted risk ratio was 1.40 (95% CI: 0.99 to 1.97) compared to the men with at least one C allele.
In addition, in the model involving rs4986791, age was found to have significant protective effect (for the 40-50 population: RR=1.17, 95% CI: 1.01 to 1.34; for the >50 population: RR=1.24, 95% CI: 1.07 to 1.43) adjusting for all other variables. AIDS development prior to HAART initiation showed a negative effect (RR=0.74, 95% CI: 0.60 to 0.92). Higher CD4+ T cell count at the last measurement before HAART initiation was also found to have a positive effect on developing the viological event (for the 200-500 population: RR=1.24, 95% CI: 1.08 to 1.43; for the >500 population: RR=1.25, 95% CI: 1.05 to 1.48).

The four models testing for different effects for each of the three genotypes, dominant effect of the minor allele, recessive effect of the minor allele and ordinal effect corresponding the number of the minor allele were shown in Table 11. These results indicate that the significant protective effect of rs4986791 on viological response to HAART is through the categorical model and the recessive model.
Aim three: Immunologic response to HAART

The same subgroup of 561 MACS participants who initiated HAART and had TLR4 genotype determined from the MACS GWAS were eligible for testing the immunological response to HAART. There were 26 individuals who didn't have record for CD4+ T cell count at the last measurement before HAART initiation, and were discarded. Then 22 people didn't have any record for the outcome of CD4+ T cell count within two years following HAART initiation, leaving 513 individuals eligible for the analysis.

Univariate analysis

The probabilities of developing the immunological event were shown in table 12 stratified by the primary exposures and cofactors. Neither of the TLR4 SNPs was found to be associated with CD4 increase following HAART initiation. Among the factors included in this analysis, only BMI>30 and pre-HAART log10 HIV RNA >4 were significantly associated with a higher probability of having a strong immunologic response to HAART.
**Multivariate analysis**

The results for the multivariate analysis of immunological response were shown in Table 13 and Table 14. In the recessive models and other three models, the two SNPs didn't show a significant effect. For example, in the multivariate results for the recessive model shown in Table 13, for the effect for the rs4986791-TT genotype was RR=1.05 (95% CI is 0.60-1.86, and the effect for the rs4986790-GG genotype was RR=1.17 (95% CI is 0.65-2.11). Finally, the only factor that was found to be significantly associated with a successful immunologic response to HAART was having a pre-HAART log10 HIV RNA level greater than 4.
Discussion

Toll-like receptor 4 plays important roles in innate immune systems by recognizing lipopolysaccharides (LPS) and triggering the downstream innate immune response pathways, and TLR4 mutations that result in protein alterations might alter the way that the immune system responds to certain agents. Two non-synonymous SNP mutations in the TLR4 gene, rs4986791 and rs4986790, have been shown to be associated with phenotypes of some human diseases. In this study, we examined the association of these mutations with HIV viral setpoint, a measure of HIV disease progression, and the change in HIV RNA level and CD4 count following HAART initiation which are, respectively, virologic and immunologic responses to therapy. Our findings show that both HIV viral setpoint and the virologic response to HAART are better in men who were homozygous with the minor allele at these SNPs while there is no evidence of an association with the immunologic response to HAART and these SNPs.

Pine et al. previously showed a strong association between the major alleles of these two SNPs and the high peak viral load phenotype, which suggests that the minor alleles
might convey protection to the HIV-1 infected individuals. We explored this relationship by examining the association of these two SNPs with HIV viral setpoint. In a linear regression analysis for a recessive gene model and adjusted for age and race, we found that the mean log10 HIV viral setpoint among men with rs4986791-TT was 0.11 lower than that among men with at least one C allele. Similarly, the mean log10 HIV viral setpoint among men with rs4986790-GG was 0.30 lower than that among men with at least one A allele. While our results are in agreement with those from the Pine study, the differences were not statistically significant in our study which was likely due to the very small number of men who were homozygous for the minor alleles of these two SNPs.

We analyzed HIV setpoint as a continuous variable in this analysis, which assumes that each one unit change on the log10 scale, or each 10-fold change on the linear scale, carries the same meaning across the entire distribution. It also assumes that the mean difference between those homozygous for the minor alleles and those with at least one major allele is the same at both low and high values of the HIV setpoint. This, however,
might not be true if the difference exists only among those with, for example, a high setpoint. To examine this question, we treated HIV setpoint as a categorical variable by dividing all men according to whether or not the HIV setpoint was greater than the overall meaning setpoint. Statistically, we attempted to model the data using logistic regression, but the R statistical package we were using could not accurately account for the fact that none of the men who were rs4986790-GG had HIV setpoint greater than the overall median. Future analyses using more powerful software are required to perform this particular analysis.

The HIV setpoint analysis demonstrated that the two SNPs might have some protective effects in terms of decreased HIV viral load at the setpoint in a recessive model. Our next step was to determine whether the two SNPs had any influence on the effectiveness of HAART treatment. Specifically, we examined the association of these TLR4 SNPs with the virologic and immunological response to HAART where the virologic event was defined as achieving undetectable HIV RNA load (<400 copies/ml) within two years following HAART treatment, and the immunological event was defined
as reaching 100 cells/μl increase in CD4+ T cells within two years following HAART treatment.

For the virologic response, we found a significant protective effect of the minor allele of the rs4986791 SNP in a recessive model after adjusting for race, age, prior AIDS development before HAART initiation, CD4+ T cell count and HIV viral load at the last measurement within two years before HAART treatment. Specifically, we observed that the probability of achieving an undetectable HIV RNA level within two years of HAART initiation was 47% higher among men with rs4986791-TT compared to men with rs4986791-CC or rs4986791-CT. This suggests the possibility that TT at rs4986791 didn't only help control the HIV viral load in natural disease conditions without HAART treatment, but also assisted in controlling the HIV viral load with HAART treatment.

The mutations in rs4986790 also showed a protective effect in a recessive model after adjusting for the same cofactors. Specifically, we observed that the probability of achieving an undetectable HIV RNA level within two years of HAART initiation was 40% higher among men with rs4986790-GG compared to men with rs4986790-AG or
rs4986790-AA. This suggests the possibility that GG at rs4986790 didn't only help control the HIV viral load in natural disease conditions without HAART treatment, but also assisted in controlling the HIV viral load with HAART treatment. However, the effect of rs4986790 was not significant, which was probably due to the small sample size of the group with GG genotype at rs4986790 (n=3). While the current study included only MACS participants, this study will be extended to include WIHS participants, and this will more than triple the total sample size, and is expected to greatly increase the power for this analysis.

In contrast to the promising but inconclusive results for HIV setpoint and the virologic response to HAART, our analysis of the TLR4 association with the immunological response to HAART showed no effect for either rs4986791 or rs4986790. The immunological response, a 100 cells/$\mu l$ increase in CD4+ T cell within two year of HAART initiation, involves complex changes in immune system in itself. Although HIV virus mainly targets CD4+ T cells, there exist other complex pathways that link HIV virus load and the CD4+ T cell count. It is possible that the hyper-activation status of the
immune system caused by HIV infection lead to a lagged CD4+ T cell recovery after the
virus level decreases. There has been one study that proved the substantial delay in
restoring CD4+ T cell in gut lymphoid tissues after HAART treatment\textsuperscript{84}. It is important
to note that we were not able to find any literature to suggest that TLR4 might be
associated with this measure of immunologic response. As such, the results of this aim
can be interpreted as a “negative control” where the exposures, TLR4 SNPs, are expected
to not be associated with the outcome.

Overall, the findings from this thesis provide intriguing preliminary evidence that
these TLR4 SNPs have inhibitory effects on the HIV virus in terms of both HIV setpoint
and the response to HAART. One possible explanation for these findings involves LPS
that has an inhibitory effect on HIV-1 infection in human macrophages in vitro, probably
by altering the CCR5 expression\textsuperscript{85}. If the TLR4 mutations lead to lower LPS recognition,
then the men with these mutations should maintain a comparatively higher amount of
LPS that, in turn, could further inhibit HIV, leading to slower HIV disease progression.
This hypothesis, however, will need to be verified by further studies in other populations.
The analysis has the advantage of using log-binomial regression for hypothesis two and hypothesis three which provided an unbiased estimate of the relative risk of the study outcomes. In situations of non-convergence, Poisson regression with robust estimation of the variance was used to produce the unbiased estimation of risk ratio as well as its confidence interval. This analysis also has an important limitation regarding the available sample size. The population for analysis of HAART treatment had 561 participants in total, where only four individuals were found to have the TT genotype at rs4986791 and three were proved to have GG genotype at rs4986790. This small number in the exposure group weakens the power of analysis. As a result, it would be very helpful to have larger sample population for later analysis and verification. Since the HIV viral load may fluctuate over time, the precise setpoint levels might not be as meaningful as the relatively levels (e.g., low versus high). Thus, TT is reasonable to model the HIV viral load at setpoint as a categorical variable, divided into subgroups according to the median or quarters. However, the R software for exact logistic regression cannot process these data, so we were unable to use exact statistical methods for this analysis.
In conclusion, we analyzed the associations of the TLR4 polymorphisms as rs4986791 and rs4986790 with markers of HIV/AIDS progression and the response to HAART. The results provided evidence of the protective effect of the two SNPs against the HIV viral load in a recessive model. Before HAART treatment, our data suggest that people with homozygous genotypes of minor alleles had lower HIV viral setpoint. After HAART treatment, the TLR4 SNPs were associated with a higher probability of achieving undetectable HIV RNA load (<400 copies/ml) within two years following HAART treatment, but not with an increase in the CD4+ T cell count. Further analyses with larger sample sizes and more robust statistical methods are needed to verify these results.
List of Figures

Figure 1. Distribution of HIV Viral Load at Setpoint
## List of Tables

Table 1. Statistics for Imputation Results

<table>
<thead>
<tr>
<th>Array</th>
<th>SNP ID</th>
<th>Major allele</th>
<th>Minor allele</th>
<th>Minor allele frequency</th>
<th>info¹</th>
<th>certainty²</th>
<th>concord³</th>
<th>r²⁴</th>
</tr>
</thead>
<tbody>
<tr>
<td>1MDuo</td>
<td>rs4986791</td>
<td>A</td>
<td>G</td>
<td>0.087</td>
<td>1.000</td>
<td>1.000</td>
<td>0.966</td>
<td>0.849</td>
</tr>
<tr>
<td></td>
<td>rs4986790</td>
<td>C</td>
<td>T</td>
<td>0.074</td>
<td>1.000</td>
<td>1.000</td>
<td>0.989</td>
<td>0.929</td>
</tr>
<tr>
<td>1Mv1</td>
<td>rs4986791</td>
<td>A</td>
<td>G</td>
<td>0.054</td>
<td>1.000</td>
<td>1.000</td>
<td>0.985</td>
<td>0.880</td>
</tr>
<tr>
<td></td>
<td>rs4986790</td>
<td>C</td>
<td>T</td>
<td>0.053</td>
<td>0.992</td>
<td>0.999</td>
<td>0.986</td>
<td>0.869</td>
</tr>
<tr>
<td>HH550</td>
<td>rs4986791</td>
<td>A</td>
<td>G</td>
<td>0.051</td>
<td>1.000</td>
<td>1.000</td>
<td>0.902</td>
<td>0.971</td>
</tr>
<tr>
<td></td>
<td>rs4986790</td>
<td>C</td>
<td>T</td>
<td>0.049</td>
<td>0.911</td>
<td>0.991</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

¹ Degree of certainty for imputation
² Average certainty for the imputed genotypes
³ Concordance between the input genotype and the imputed genotype for this SNP using the results
⁴ Squared correlation between the input genotype and the "dosage" genotype for this SNP after imputation using the results
Table 2. Distribution of rs4986791 and rs4986790 after incorporating the imputation results.

<table>
<thead>
<tr>
<th>GWAS Array</th>
<th>N</th>
<th>rs4986791</th>
<th>rs4986790</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CC</td>
<td>CT</td>
</tr>
<tr>
<td>1MDuo</td>
<td>263</td>
<td>227</td>
<td>33</td>
</tr>
<tr>
<td>1Mv1</td>
<td>806</td>
<td>724</td>
<td>80</td>
</tr>
<tr>
<td>HH550</td>
<td>498</td>
<td>453</td>
<td>43</td>
</tr>
</tbody>
</table>
Table 3. Characteristics of 1169 HAART naïve MACS participants by rs4986791 genotype

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N</th>
<th>%</th>
<th>rs4986791- CC</th>
<th>rs4986791- CT/TT</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1169</td>
<td>100</td>
<td>1036</td>
<td>133</td>
<td>0.39</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>1023</td>
<td>87.5</td>
<td>902</td>
<td>121</td>
<td>0.34</td>
</tr>
<tr>
<td>Black non-Hispanic</td>
<td>69</td>
<td>5.9</td>
<td>62</td>
<td>7</td>
<td>0.16</td>
</tr>
<tr>
<td>Others</td>
<td>77</td>
<td>6.6</td>
<td>72</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30</td>
<td>382</td>
<td>32.7</td>
<td>334</td>
<td>48</td>
<td>0.39</td>
</tr>
<tr>
<td>30-40</td>
<td>573</td>
<td>49.0</td>
<td>507</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>&gt;40</td>
<td>214</td>
<td>18.3</td>
<td>195</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>90</td>
<td>7.7</td>
<td>86</td>
<td>4</td>
<td>0.16</td>
</tr>
<tr>
<td>20-25</td>
<td>817</td>
<td>69.9</td>
<td>722</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>25-30</td>
<td>216</td>
<td>18.5</td>
<td>187</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>&gt;30</td>
<td>34</td>
<td>2.9</td>
<td>30</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>12</td>
<td>1.0</td>
<td>11</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Smoke status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>434</td>
<td>37.1</td>
<td>388</td>
<td>46</td>
<td>0.75</td>
</tr>
<tr>
<td>Former</td>
<td>258</td>
<td>22.1</td>
<td>227</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>463</td>
<td>39.6</td>
<td>407</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>14</td>
<td>1.2</td>
<td>14</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Cumulative Pack-years of cigarettes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never smokers</td>
<td>432</td>
<td>37.0</td>
<td>386</td>
<td>46</td>
<td>0.95</td>
</tr>
<tr>
<td>0-5</td>
<td>197</td>
<td>16.9</td>
<td>174</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>5-20</td>
<td>222</td>
<td>19.0</td>
<td>194</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>20-35</td>
<td>167</td>
<td>14.3</td>
<td>147</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>&gt;35</td>
<td>129</td>
<td>11.0</td>
<td>115</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>22</td>
<td>1.8</td>
<td>20</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>SNP</td>
<td>Genotype</td>
<td>N</td>
<td>Proportion with log10 HIV viral setpoint &gt;4.3 (%)</td>
<td>RR</td>
<td>95% CI</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>------</td>
<td>-----------------------------------------------</td>
<td>------</td>
<td>--------------</td>
</tr>
<tr>
<td>rs4986791</td>
<td>CC</td>
<td>1036</td>
<td>51</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>124</td>
<td>56</td>
<td>1.10</td>
<td>(0.93, 1.30)</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>9</td>
<td>11</td>
<td>0.22</td>
<td>(0.03, 1.40)</td>
</tr>
<tr>
<td>rs4986790</td>
<td>AA</td>
<td>1042</td>
<td>51</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>119</td>
<td>56</td>
<td>1.11</td>
<td>(0.94, 1.32)</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>7</td>
<td>0</td>
<td>0.00</td>
<td>NA</td>
</tr>
</tbody>
</table>
Table 5. Univariate comparisons of participant characteristic and log10 HIV viral Setpoint

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>N</th>
<th>Log10 HIV viral setpoint Median (IQR range)</th>
<th>Univariate Test (linear regression)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Median (IQR range)</td>
<td>Coefficient(^1)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1169</td>
<td>4.31 (3.78, 4.77)</td>
<td></td>
</tr>
<tr>
<td>rs4986791</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>1036</td>
<td>4.31 (3.78, 4.76)</td>
<td>0</td>
</tr>
<tr>
<td>CT</td>
<td>124</td>
<td>4.38 (3.70, 4.85)</td>
<td>0.05</td>
</tr>
<tr>
<td>TT</td>
<td>9</td>
<td>4.11 (3.80, 4.20)</td>
<td>-0.12</td>
</tr>
<tr>
<td><strong>rs4986790</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>1042</td>
<td>4.31 (3.80, 4.77)</td>
<td>0</td>
</tr>
<tr>
<td>AG</td>
<td>119</td>
<td>4.40 (3.66, 4.87)</td>
<td>0.17</td>
</tr>
<tr>
<td>GG</td>
<td>7</td>
<td>4.11 (3.67, 4.19)</td>
<td>-0.16</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>1023</td>
<td>4.32 (3.81, 4.77)</td>
<td>0</td>
</tr>
<tr>
<td>Black non-Hispanic</td>
<td>69</td>
<td>4.27 (3.55, 4.67)</td>
<td>-0.16</td>
</tr>
<tr>
<td>Others</td>
<td>77</td>
<td>4.30 (3.64, 4.82)</td>
<td>-0.02</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30</td>
<td>382</td>
<td>4.25 (3.70, 4.70)</td>
<td>0</td>
</tr>
<tr>
<td>30-40</td>
<td>573</td>
<td>4.30 (3.80, 4.78)</td>
<td>0.06</td>
</tr>
<tr>
<td>&gt;40</td>
<td>214</td>
<td>4.44 (3.94, 4.83)</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-heavy (&lt;25)</td>
<td>907</td>
<td>4.30 (3.81, 4.80)</td>
<td>0</td>
</tr>
<tr>
<td>Heavy (&gt;=25)</td>
<td>250</td>
<td>4.38 (3.79, 4.75)</td>
<td>-0.03</td>
</tr>
<tr>
<td><strong>Smoke degree (Two: current smoker or previous smoker with &gt;20 CUM_PKYEAR)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One</td>
<td>665</td>
<td>4.34 (3.74, 4.79)</td>
<td>0</td>
</tr>
<tr>
<td>Two</td>
<td>504</td>
<td>4.28 (3.81, 4.74)</td>
<td>0.005</td>
</tr>
</tbody>
</table>

\(^1\) This should be interpreted as the differences of log10 transformed HIV viral load at the setpoint
### Table 6. Multivariate Regression of log10 HIV viral Load at Setpoint (Recessive Model)

<table>
<thead>
<tr>
<th>Variable</th>
<th>log10 viral load difference</th>
<th>95% CI</th>
<th>Variables</th>
<th>log10 viral load difference</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs4986791</td>
<td></td>
<td></td>
<td>RS4986790</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC/CT</td>
<td>Ref</td>
<td></td>
<td>AA/AG</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>-0.11</td>
<td>(-0.42, 0.68)</td>
<td>GG</td>
<td>-0.30</td>
<td>(-0.88, 0.28)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>Ref</td>
<td></td>
<td>White non-Hispanic</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>Black non-Hispanic</td>
<td>-0.15</td>
<td>(-0.34, 0.04)</td>
<td>Black non-Hispanic</td>
<td>-0.15</td>
<td>(-0.34, 0.04)</td>
</tr>
<tr>
<td>Others</td>
<td>0.000</td>
<td>(-0.18, 0.18)</td>
<td>Others</td>
<td>-0.001</td>
<td>(-0.18, 0.18)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;40</td>
<td>Ref</td>
<td></td>
<td>&lt;40</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>40-50</td>
<td>0.06</td>
<td>(-0.04, 0.16)</td>
<td>40-50</td>
<td>0.06</td>
<td>(-0.04, 0.16)</td>
</tr>
<tr>
<td>&gt;50</td>
<td>0.12</td>
<td>(-0.01, 0.25)</td>
<td>&gt;50</td>
<td>0.12</td>
<td>(-0.01, 0.25)</td>
</tr>
</tbody>
</table>
Table 7. Summary of Multivariate regression results for log10 HIV viral setpoint using four different genetic models (categorical, dominant, recessive, and ordinal effects)

<table>
<thead>
<tr>
<th>Model</th>
<th>Genotype</th>
<th>log10 viral load difference</th>
<th>95% CI</th>
<th>Genotype</th>
<th>log10 viral load difference</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Categorical</td>
<td>CC</td>
<td>Ref</td>
<td>-0.10</td>
<td>GG</td>
<td>-0.29</td>
<td>(-0.87,0.29)</td>
</tr>
<tr>
<td>Model</td>
<td>CT</td>
<td>0.05</td>
<td>(-0.09,0.20)</td>
<td>AG</td>
<td>0.05</td>
<td>(-0.10,0.20)</td>
</tr>
<tr>
<td>Model</td>
<td>TT</td>
<td>-0.10</td>
<td>(-0.62,0.41)</td>
<td>GG</td>
<td>-0.29</td>
<td>(-0.87,0.29)</td>
</tr>
<tr>
<td>Dominant</td>
<td>CC</td>
<td>Ref</td>
<td>-0.30</td>
<td>GG</td>
<td>-0.30</td>
<td>(-0.88,0.28)</td>
</tr>
<tr>
<td>effect</td>
<td>CT/TT</td>
<td>0.03</td>
<td>(-0.12,0.17)</td>
<td>AG/GG</td>
<td>0.03</td>
<td>(-0.12,0.17)</td>
</tr>
<tr>
<td>Recessive</td>
<td>CC/CT</td>
<td>Ref</td>
<td></td>
<td>AA/AG</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>effect</td>
<td>TT</td>
<td>-0.30</td>
<td>(-0.88,0.28)</td>
<td>GG</td>
<td>-0.30</td>
<td>(-0.88,0.28)</td>
</tr>
<tr>
<td>Ordinal</td>
<td>Number of T</td>
<td>0.03</td>
<td>(-0.10,0.16)</td>
<td>Number of G</td>
<td>0.01</td>
<td>(-0.12,0.14)</td>
</tr>
</tbody>
</table>
Table 8. Characteristics of 571 HAART treated MACS participants by rs4986791 genotypes

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N</th>
<th>%</th>
<th>rs4986791- CC Number</th>
<th>Percentage(%)</th>
<th>rs4986791- CT/TT Number</th>
<th>Percentage(%)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>561</td>
<td>100</td>
<td>497</td>
<td>88.6</td>
<td>64</td>
<td>11.4</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White-nonHispanic</td>
<td>461</td>
<td>82.2</td>
<td>409</td>
<td>88.7</td>
<td>52</td>
<td>11.3</td>
<td>0.97</td>
</tr>
<tr>
<td>Black-nonHispanic</td>
<td>65</td>
<td>11.6</td>
<td>57</td>
<td>87.7</td>
<td>8</td>
<td>12.3</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>35</td>
<td>6.2</td>
<td>31</td>
<td>88.6</td>
<td>4</td>
<td>11.4</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;40</td>
<td>154</td>
<td>27.5</td>
<td>134</td>
<td>87.0</td>
<td>20</td>
<td>13.0</td>
<td>0.47</td>
</tr>
<tr>
<td>40-50</td>
<td>252</td>
<td>44.9</td>
<td>221</td>
<td>87.7</td>
<td>31</td>
<td>12.2</td>
<td></td>
</tr>
<tr>
<td>&gt;50</td>
<td>118</td>
<td>21.0</td>
<td>108</td>
<td>91.5</td>
<td>10</td>
<td>8.5</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>37</td>
<td>6.6</td>
<td>34</td>
<td>91.9</td>
<td>3</td>
<td>8.1</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>36</td>
<td>6.4</td>
<td>32</td>
<td>88.9</td>
<td>4</td>
<td>11.1</td>
<td>0.81</td>
</tr>
<tr>
<td>20-25</td>
<td>282</td>
<td>50.3</td>
<td>246</td>
<td>87.2</td>
<td>36</td>
<td>12.8</td>
<td></td>
</tr>
<tr>
<td>25-30</td>
<td>183</td>
<td>32.6</td>
<td>165</td>
<td>90.2</td>
<td>18</td>
<td>9.8</td>
<td></td>
</tr>
<tr>
<td>&gt;30</td>
<td>55</td>
<td>9.8</td>
<td>49</td>
<td>89.1</td>
<td>6</td>
<td>10.9</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>5</td>
<td>0.9</td>
<td>5</td>
<td>100.0</td>
<td>0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Smoke status at HAART initiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>153</td>
<td>27.3</td>
<td>144</td>
<td>94.1</td>
<td>9</td>
<td>5.9</td>
<td>0.053</td>
</tr>
<tr>
<td>Former</td>
<td>240</td>
<td>42.8</td>
<td>208</td>
<td>86.7</td>
<td>32</td>
<td>13.3</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>166</td>
<td>29.6</td>
<td>144</td>
<td>86.7</td>
<td>22</td>
<td>13.3</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>2</td>
<td>0.3</td>
<td>1</td>
<td>50.0</td>
<td>1</td>
<td>50.0</td>
<td></td>
</tr>
<tr>
<td>Cumulative Pack-years of cigarettes among former and current smokers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never smokers</td>
<td>152</td>
<td>27.1</td>
<td>143</td>
<td>94.1</td>
<td>9</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>0-5</td>
<td>132</td>
<td>23.5</td>
<td>117</td>
<td>88.6</td>
<td>15</td>
<td>11.4</td>
<td></td>
</tr>
<tr>
<td>5-20</td>
<td>91</td>
<td>16.2</td>
<td>77</td>
<td>84.6</td>
<td>14</td>
<td>15.4</td>
<td>0.15</td>
</tr>
<tr>
<td>20-35</td>
<td>80</td>
<td>14.3</td>
<td>69</td>
<td>86.3</td>
<td>11</td>
<td>13.7</td>
<td></td>
</tr>
<tr>
<td>&gt;35</td>
<td>93</td>
<td>16.6</td>
<td>81</td>
<td>87.1</td>
<td>12</td>
<td>12.9</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>13</td>
<td>2.3</td>
<td>10</td>
<td>76.9</td>
<td>3</td>
<td>23.1</td>
<td></td>
</tr>
<tr>
<td>With prior developed AIDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CD4+ T cell count at baseline</td>
<td>log10 (HIV RNA load at baseline)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&lt;100</td>
<td>&lt;100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>100</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17.8</td>
<td>15.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>86</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>86.0</td>
<td>88.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14.0</td>
<td>11.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.47</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>461</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>82.2</td>
<td>93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>411</td>
<td>93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>89.2</td>
<td>88.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>11.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.8</td>
<td>11.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>93</td>
<td>90-200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>84</td>
<td>200-300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>94</td>
<td>300-400</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>57</td>
<td>400-500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>118</td>
<td>&gt;500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.6</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>4.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>92.3</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>92.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.7</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>169</td>
<td>&lt;4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>94</td>
<td>4-4.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>116</td>
<td>4.5-5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>137</td>
<td>&gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>Missing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.0</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>8.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>80.0</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>80.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20.0</td>
<td>20.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 9. Univariate comparisons of participant characteristic and virologic response to HAART

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>%</th>
<th>Univariate Test</th>
<th>Risk Ratio</th>
<th>95% Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>448</td>
<td>74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4986791</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>406</td>
<td>73</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>38</td>
<td>76</td>
<td>1.04 (0.86, 1.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>4</td>
<td>100</td>
<td>1.36 (1.28, 1.44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4986790</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>409</td>
<td>74</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>36</td>
<td>72</td>
<td>1.04 (0.79, 1.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>3</td>
<td>100</td>
<td>1.35 (1.28, 1.43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>367</td>
<td>74</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black non-Hispanic</td>
<td>55</td>
<td>81</td>
<td>1.11 (0.94, 1.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>26</td>
<td>58</td>
<td>0.78 (0.52, 1.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;40</td>
<td>136</td>
<td>65</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-50</td>
<td>211</td>
<td>75</td>
<td>1.16 (1.01, 1.36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;50</td>
<td>101</td>
<td>83</td>
<td>1.29 (1.11, 1.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-obesity (&lt;30)</td>
<td>399</td>
<td>73</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity (&gt;=30)</td>
<td>46</td>
<td>83</td>
<td>1.13 (0.95, 1.28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoke degree (Two: current smoker or previous smoker with &gt;20 CUM_PKYEAR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One</td>
<td>281</td>
<td>77</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two</td>
<td>167</td>
<td>69</td>
<td>0.91 (0.80, 1.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior AIDS before HAART initiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>365</td>
<td>79</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>83</td>
<td>52</td>
<td>0.66 (0.52, 0.80)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD4+ T cell count at baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;200</td>
<td>160</td>
<td>60</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200-500</td>
<td>202</td>
<td>81</td>
<td>1.35 (1.18, 1.57)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;500</td>
<td>86</td>
<td>83</td>
<td>1.38 (1.57, 1.62)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log10 Virus RNA load at baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;4</td>
<td>117</td>
<td>79</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-5</td>
<td>201</td>
<td>76</td>
<td>0.96</td>
<td>(0.85, 1.09)</td>
<td></td>
</tr>
<tr>
<td>&gt;5</td>
<td>130</td>
<td>65</td>
<td>0.82</td>
<td>(0.70, 0.96)</td>
<td></td>
</tr>
<tr>
<td>Variable</td>
<td>Risk Ratio</td>
<td>95% CI</td>
<td>Variable</td>
<td>Risk Ratio</td>
<td>95% CI</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------------</td>
<td>------------</td>
<td>--------------------------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>rs4986791</td>
<td></td>
<td></td>
<td>rs4986790</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC/CT</td>
<td>Ref</td>
<td></td>
<td>AA/AG</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>1.47</td>
<td>(1.10, 1.98)</td>
<td>GG</td>
<td>1.40</td>
<td>(0.99, 1.97)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>Ref</td>
<td></td>
<td>White non-Hispanic</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>Black non-Hispanic</td>
<td>1.10</td>
<td>(0.96, 1.27)</td>
<td>Black non-Hispanic</td>
<td>1.10</td>
<td>(0.96, 1.27)</td>
</tr>
<tr>
<td>Others</td>
<td>0.83</td>
<td>(0.59, 1.16)</td>
<td>Others</td>
<td>0.75</td>
<td>(0.53, 1.07)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;40</td>
<td>Ref</td>
<td></td>
<td>&lt;40</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>40-50</td>
<td>1.17</td>
<td>(1.01, 1.34)</td>
<td>40-50</td>
<td>1.17</td>
<td>(1.02, 1.35)</td>
</tr>
<tr>
<td>&gt;50</td>
<td>1.24</td>
<td>(1.07, 1.43)</td>
<td>&gt;50</td>
<td>1.24</td>
<td>(1.07, 1.43)</td>
</tr>
<tr>
<td>Prior AIDS before HAART initiation</td>
<td></td>
<td></td>
<td>Prior AIDS before HAART initiation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>Ref</td>
<td></td>
<td>No</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0.74</td>
<td>(0.60, 0.92)</td>
<td>Yes</td>
<td>0.74</td>
<td>(0.60, 0.92)</td>
</tr>
<tr>
<td>CD4+ T cell count at baseline</td>
<td></td>
<td></td>
<td>CD4+ T cell count at baseline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;200</td>
<td>Ref</td>
<td></td>
<td>&lt;200</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>200-500</td>
<td>1.24</td>
<td>(1.08, 1.43)</td>
<td>200-500</td>
<td>1.24</td>
<td>(1.08, 1.43)</td>
</tr>
<tr>
<td>&gt;500</td>
<td>1.25</td>
<td>(1.05, 1.48)</td>
<td>&gt;500</td>
<td>1.24</td>
<td>(1.05, 1.48)</td>
</tr>
<tr>
<td>Log10 Virus RNA load at baseline</td>
<td></td>
<td></td>
<td>Log10 Virus RNA load at baseline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;4</td>
<td>Ref</td>
<td></td>
<td>&lt;4</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td>4-5</td>
<td>0.99</td>
<td>(0.87, 1.12)</td>
<td>4-5</td>
<td>0.99</td>
<td>(0.87, 1.12)</td>
</tr>
<tr>
<td>&gt;5</td>
<td>0.94</td>
<td>(0.80, 1.11)</td>
<td>&gt;5</td>
<td>0.94</td>
<td>(0.80, 1.10)</td>
</tr>
</tbody>
</table>
Table 11. Summary of Multivariate regression results for viological responses using four different genetic models (categorical, dominant, recessive, and ordinal effects)

<table>
<thead>
<tr>
<th>Model</th>
<th>Genotype</th>
<th>Risk Ratio</th>
<th>95% CI</th>
<th>Genotype</th>
<th>Risk Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Categorical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>CC</td>
<td>Ref</td>
<td></td>
<td>AA</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>1.04</td>
<td>(0.88,1.22)</td>
<td>AG</td>
<td>1.01</td>
<td>(0.83,1.23)</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td><strong>1.48</strong></td>
<td><strong>(1.10,1.99)</strong></td>
<td>GG</td>
<td><strong>1.40</strong></td>
<td><strong>(0.99,1.97)</strong></td>
</tr>
<tr>
<td>Dominant effect</td>
<td>CC</td>
<td>Ref</td>
<td></td>
<td>AA</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CT/TT</td>
<td>1.08</td>
<td>(0.92,1.25)</td>
<td>AG/GG</td>
<td>1.04</td>
<td>(0.86,1.25)</td>
</tr>
<tr>
<td>Recessive effect</td>
<td>CC/CT</td>
<td>Ref</td>
<td></td>
<td>AA/AG</td>
<td>Ref</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td><strong>1.47</strong></td>
<td><strong>(1.10,1.98)</strong></td>
<td>GG</td>
<td><strong>1.40</strong></td>
<td><strong>(0.99,1.97)</strong></td>
</tr>
<tr>
<td>Ordinal effect</td>
<td>Number of T</td>
<td>1.09</td>
<td>(0.97,1.24)</td>
<td>Number of G</td>
<td>1.06</td>
<td>(0.91,1.23)</td>
</tr>
<tr>
<td>Variable</td>
<td>N</td>
<td>%</td>
<td>Risk Ratio</td>
<td>95% Confidence interval</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----</td>
<td>---</td>
<td>------------</td>
<td>-------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>513</td>
<td>69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4986791</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>455</td>
<td>69</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>52</td>
<td>65</td>
<td>0.94</td>
<td>(0.77, 1.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>6</td>
<td>67</td>
<td>0.96</td>
<td>(0.54, 1.70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4986790</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>461</td>
<td>69</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>48</td>
<td>65</td>
<td>0.96</td>
<td>(0.78, 1.19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>4</td>
<td>67</td>
<td>1.08</td>
<td>(0.61, 1.91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>419</td>
<td>69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black non-Hispanic</td>
<td>62</td>
<td>73</td>
<td>1.05</td>
<td>(0.87, 1.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>32</td>
<td>59</td>
<td>0.86</td>
<td>(0.61, 1.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;40</td>
<td>151</td>
<td>67</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-50</td>
<td>245</td>
<td>70</td>
<td>1.04</td>
<td>(0.91, 1.20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;50</td>
<td>117</td>
<td>70</td>
<td>1.05</td>
<td>(0.89, 1.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-obesity (&lt;30)</td>
<td>357</td>
<td>67</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity (&gt;=30)</td>
<td>53</td>
<td>81</td>
<td>1.20</td>
<td>(1.02, 1.37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoke degree (Two: current smoker or previous smoker with &gt;20 CUM_PKYEAR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One</td>
<td>309</td>
<td>72</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two</td>
<td>204</td>
<td>64</td>
<td>0.89</td>
<td>(0.78, 1.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior AIDS before HAART initiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>415</td>
<td>71</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>98</td>
<td>61</td>
<td>0.86</td>
<td>(0.72, 1.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD4+ T cell count at baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;200</td>
<td>173</td>
<td>69</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200-500</td>
<td>223</td>
<td>74</td>
<td>1.06</td>
<td>(0.93, 1.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;500</td>
<td>117</td>
<td>60</td>
<td>0.86</td>
<td>(0.72, 1.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Virus RNA load at baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;4</td>
<td>165</td>
<td>56</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-5</td>
<td>201</td>
<td>75</td>
<td>1.34</td>
<td>(1.15, 1.58)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

63
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;5</td>
<td>131</td>
<td>78</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.19, 1.65)</td>
</tr>
</tbody>
</table>
Table 13. Multivariate Results of Immunological Response (Recessive Model)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Risk Ratio</th>
<th>95% CI</th>
<th>Variable</th>
<th>Risk Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs4986791</td>
<td></td>
<td></td>
<td>rs4986790</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC/CT Ref</td>
<td></td>
<td></td>
<td>AA/AG Ref</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>1.05</td>
<td>(0.60, 1.86)</td>
<td>GG</td>
<td>1.17</td>
<td>(0.65, 2.11)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White non-Hispanic Ref</td>
<td></td>
<td></td>
<td>White non-Hispanic Ref</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black non-Hispanic</td>
<td>1.05</td>
<td>(0.89, 1.23)</td>
<td>Black non-Hispanic</td>
<td>1.05</td>
<td>(0.89, 1.23)</td>
</tr>
<tr>
<td>Others</td>
<td>0.86</td>
<td>(0.66, 1.13)</td>
<td>Others</td>
<td>0.86</td>
<td>(0.66, 1.13)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;40 Ref</td>
<td></td>
<td></td>
<td>&lt;40 Ref</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-50</td>
<td>1.04</td>
<td>(0.90, 1.20)</td>
<td>40-50</td>
<td>1.04</td>
<td>(0.90, 1.20)</td>
</tr>
<tr>
<td>&gt;50</td>
<td>1.03</td>
<td>(0.88, 1.21)</td>
<td>&gt;50</td>
<td>1.03</td>
<td>(0.88, 1.21)</td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td>BMI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-obesity (&lt;30) Ref</td>
<td></td>
<td></td>
<td>Non-obesity (&lt;30) Ref</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity (&gt;=30)</td>
<td>1.14</td>
<td>(0.98, 1.32)</td>
<td>Obesity (&gt;=30)</td>
<td>1.14</td>
<td>(0.98, 1.32)</td>
</tr>
<tr>
<td>Prior AIDS before HAART initiation</td>
<td></td>
<td></td>
<td>Prior AIDS before HAART initiation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Ref</td>
<td></td>
<td></td>
<td>No Ref</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0.84</td>
<td>(0.70, 1.02)</td>
<td>Yes</td>
<td>0.84</td>
<td>(0.70, 1.02)</td>
</tr>
<tr>
<td>CD4+ T cell count at baseline</td>
<td></td>
<td></td>
<td>CD4+ T cell count at baseline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;200 Ref</td>
<td></td>
<td></td>
<td>&lt;200 Ref</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200-500</td>
<td>1.07</td>
<td>(0.93, 1.22)</td>
<td>200-500</td>
<td>1.07</td>
<td>(0.93, 1.22)</td>
</tr>
<tr>
<td>&gt;500</td>
<td>0.99</td>
<td>(0.81, 1.21)</td>
<td>&gt;500</td>
<td>0.99</td>
<td>(0.81, 1.21)</td>
</tr>
<tr>
<td>Log10 Virus RNA load at baseline</td>
<td></td>
<td></td>
<td>Log10 Virus RNA load at baseline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;4 Ref</td>
<td></td>
<td></td>
<td>&lt;4 Ref</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-5</td>
<td>1.33</td>
<td>(1.12, 1.57)</td>
<td>4-5</td>
<td>1.33</td>
<td>(1.12, 1.57)</td>
</tr>
<tr>
<td>&gt;5</td>
<td>1.43</td>
<td>(1.19, 1.71)</td>
<td>&gt;5</td>
<td>1.43</td>
<td>(1.19, 1.71)</td>
</tr>
</tbody>
</table>
Table 14. Summary of Multivariate regression results for Immunological Responses using four different genetic models (categorical, dominant, recessive, and ordinal effects)

<table>
<thead>
<tr>
<th>Model</th>
<th>rs4986791</th>
<th>rs4986790</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Genotype</td>
<td>Risk Ratio</td>
</tr>
<tr>
<td>Categorical</td>
<td>CC</td>
<td>Ref</td>
</tr>
<tr>
<td>Model</td>
<td>CT</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>1.04</td>
</tr>
<tr>
<td>Dominant effect</td>
<td>CC</td>
<td>Ref</td>
</tr>
<tr>
<td></td>
<td>CT/TT</td>
<td>0.93</td>
</tr>
<tr>
<td>Recessive effect</td>
<td>CC/CT</td>
<td>Ref</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>1.05</td>
</tr>
<tr>
<td>Ordinal effect</td>
<td>Number of T</td>
<td>0.95</td>
</tr>
</tbody>
</table>
References


[36] Li, Tao Sheng, et al. Long-lasting recovery in CD4 T-cell function and viral-load reduction


[70] Abstract at 14th International AIDS Conference in Barcelona, Spain


[76] https://mathgen.stats.ox.ac.uk/impute/impute_v2.html#basic_options

[77] https://mathgen.stats.ox.ac.uk/impute/1000GP20Phase20320haplotypes20620October202014.html

[78] https://genome.ucsc.edu/FAQ/FAQreleases.html


Curriculum Vitae

Yuan He

*MHS candidate in epidemiology*

110 W 39th street, Baltimore, MD, 21210

(410)-350-1306; yhe23@jhu.edu

---

**EDUCATION**

Bloomberg School of Public Health, Johns Hopkins University, USA

*Master of Health Science, expected in June 2015.*

School of Life Sciences, Peking University, China

*B.S., Biological Sciences, July 2013*

---

**ANALYTICAL TRAINING**

*Methods in Biostatistics*


*Statistics for Genomics*

Analyze big data from ChIP-seq, RNA sequencing results.

*Multilevel statistical models*

*Bayesian Methods*

---

**EPIDEMIOLOGY TRAINING**

*Epidemiology Methods*

Basics about study design and data analysis in epidemiology research.

*Principles of Genetic Epidemiology*

*Introduction to clinical trails*

*Current topics in epidemiologic research*
**RESEARCH EXPERIENCE**

**Research Assistant**
Johns Hopkins School of Public Health. USA. December 2014-Present
Supervisor: Dr. Eric C. Seaberg
Project: Do master thesis.

**Research Assistant**
National Key Laboratory of Cell Proliferation and Differentiation, Peking University, Beijing, China. Nov.2011-July.2013
Supervisor: Dr. Hongkui Deng
Project: Knock out CCR5 on human CD34+ cells by Transcription Activator Like Effector Nuclease to create a stem cell therapy towards AIDS

**TEACHING EXPERIENCE**

**Teaching Assistant**
Course: Statistical Methods in Public Health I-III
Department of Biostatistics, Johns Hopkins School of Public Health. USA. Sep 2014-Feb 2015