Published in final edited form as:
Pediatr Infect Dis J. 2014 September ; 33(9): 946–952. doi:10.1097/INF.0000000000000330.

Associations of Host Genetic Variants on CD4+ Lymphocyte Count and Plasma HIV-1 RNA in Antiretroviral Naïve Children

Min Qin, PhD[1], **Sean Brummel, PhD**[1], **Kumud K. Singh, PhD**[2], **Terry Fenton, EdD**[1], and **Stephen A. Spector, MD**[2]

[1]Center for Biostatistics in AIDS Research, Harvard School of Public Health, Boston, MA

[2]University of California, San Diego, La Jolla, CA and Rady Children's Hospital, San Diego, CA

Abstract

Background—CD4+ T-lymphocyte (CD4) counts and HIV plasma RNA concentration (RNA) are two key HIV disease markers. The complex interplay between virus and host genetics may contribute to differences in viral set point and CD4 status. Determining the effects of host genetic variation on HIV disease markers is often complicated by the use of antiretroviral therapy. In this study, the association between genetic variants and baseline HIV RNA and CD4 counts was examined in a large cohort of antiretroviral naïve children.

Methods—Specimens from 1053 HIV-infected children were screened for single nucleotide polymorphisms (SNPs) in 78 regions from 17 genes. Linear regression with a robust variance estimator was used to test the association between genetic markers with HIV RNA and CD4count, controlling for age, race/ethnicity and study. False discovery rate (FDR) controlling was used to adjust for multiple testing.

Results—The study population was 60% black, 26% Hispanic and 13% white; median age 2.35 years; 55% female. Baseline median CD4 count was 780/mm^3; median log10 HIV RNA was 5.17 copies/mL. For analyses of the associations of genetic makers with baseline CD4+, 6 HLA and 4 additional markers exhibited p-values <0.05, but none met the criteria for statistical significance with FDR controlled at 0.05. For baseline HIV RNA, HLA DRB1*15, DRB1*10, B-27/57, B-14, Cw-8, B57 were statistically significant with FDR controlled at 0.05.

Conclusions—These results provide strong evidence that HLA DRB1*15, DRB1*10, B-27/57, B-14, Cw-8, B57 are associated with HIV RNA, and play a role in HIV pathogenesis in infected children.

Keywords

HIV; host genetics; single nucleotide polymorphism; false discovery rate

CD4+ T-lymphocyte cell (CD4) count and HIV plasma RNA concentration (RNA) are the two most commonly used prognostic markers of disease status and clinical progression of

Correspondence: Name: Stephen A. Spector, M.D., 9500 Gilman Drive, Stein Clinical Research Building, MC 0672, La Jolla, CA, 92093-0672., saspector@ucsd.edu, Phone: 858 534 7170, Fax: 858 534 7411.

Conflicts of Interest: The authors have no conflicts of interest or funding to disclose.

human immunodeficiency virus type-1 (HIV) infection. Due to the complex interplay between the virus, its human host and the environment, individuals infected with HIV differ with respect to viral set point, extent of CD4 count decline and response to treatment. Host genetic variants may account for some of this variation. However, due to the successful use of antiretroviral therapy (ART) in current clinical practice, it is difficult to examine the direct relationship between host genetic variation and HIV disease markers because: 1) HIV-infected patients starting ART typically experience a rapid decline in HIV RNA levels and a steady increase in CD4 count, decreasing the amount of HIV RNA/CD4 count variability that may be explained by other factors, including host genetics; and 2) the relationship between host genetics and HIV RNA/CD4 count may be mediated by specific forms of therapy, if host genetics are related to the pharmacokinetics of antiretroviral drugs used for treatment. Thus, studying populations that are ART naïve can be helpful in understanding the effects of host genetics on the pathogenesis of HIV infection as measured by HIV RNA and CD4 count (1).

In HIV-infected adults, host genetic factors impact on the HIV viral set point before the initiation of antiretroviral therapy. In 131 antiretroviral naïve HIV-infected adults of Caucasian origin, different Fas or FasL polymorphisms were associated with an increase in CD4 count and/or a decrease in HIV RNA (2). In a study of the effect of HLA class II alleles on HIV disease outcome and HIV-specific T cell responses in a cohort of 426 antiretroviral therapy-naive, HIV-1 clade C-infected, predominantly female black South Africans, the presence of HLA class II DRB1*1303 was associated with reduced viral loads in chronic HIV-1 clade C and B infection (3). In another study of 110 antiretroviral-naive Ugandans, slow progressors with known protective HLA B*allele-associated multi-clade gag T-cell recognition had marginally significantly lower mean plasma viral loads than those who lacked those B alleles (4). Additionally, the presence of CCR5-wt/d32 heterozygous variant was associated with a lower viral load early in infection (5, 6, 7).

In the research presented here, we have examined the baseline, pre-treatment relationship between HIV RNA/CD4 count and host genetic markers that have previously been identified or hypothesized to alter the progression to HIV-related clinical endpoints in a unique antiretroviral naïve cohort of HIV-infected children. The goal was to estimate the strength of association between specific host genetic variants and pre-treatment CD4 count and HIV RNA.

SUBJECTS AND METHODS

Subjects

1053 children with symptomatic HIV infection from Pediatric AIDS Clinical Trial Group (PACTG) protocols P152 (8) and P300 (9) were included in the analyses. P152 and P300 were multicenter, prospective, randomized, double blind, placebo controlled protocols that assessed the efficacy of single or combination nucleoside reverse transcriptase inhibitor (NRTI) treatment regimens in symptomatic HIV-infected children in the United States, prior to the availability of highly active antiretroviral therapy (HAART). Important eligibility criteria included an age range of 3 months-18 years with symptomatic HIV-1 infection for P152 (8), an age range of 42 days-15 years with symptomatic HIV-1 infection for P300 (9),

and meeting the requirements that the original Centers for Disease Control (CDC) classification system had established to diagnose HIV-1 infection in children at the time of these studies (10, 11). In these two protocols, CD4 count and percentage and HIV RNA were measured at entry, prior to initiation of therapy. The current study uses these baseline CD4 and HIV RNA data as dependent variables and tests whether their values are associated with host genetic variants. HIV RNA measurements were dependent on the technology available at the time of the studies; for P152 the NASBA HIV-1 RNA QT Amplification System was used (12) and for P300 the Roche Amplicor quantitative RNA PCR assay was used (9). Of the 1053 subjects, 1045 had baseline CD4 count and 871 had baseline HIV RNA data. The summary statistics of baseline CD4 count and HIV RNA, as well as other baseline characteristics of these antiretroviral naïve children, are described in Table 1. This study followed the human experimentation guidelines of the US Department of Health and Human Services and the University of California San Diego Institutional Review Board.

Genotyping

Stored DNA samples from the 1053 children were screened for the presence of polymorphisms in host genes hypothesized to play a role in HIV-associated disease progression. Total genomic DNA was extracted from peripheral blood mononuclear cells (PBMCs) using mini blood DNA kits (Qiagen, Carlsbad, CA). Whole Genome Amplification (WGA) of DNA was performed using Qiagen WGA kits (13). Although there is potential for non-uniformity of DNA amplification with WGA, we have found 100% correlation when SNP and HLA data are compared between genomic DNA and WGA DNA. Seventy-eight single nucleotide polymorphisms (SNPs) from 17 genes including HLA, APOBEC3G, APOE, CCR5, CCR2, CX_3CR1, TNF, SDF1, RANTES, PRMT6, MCP1-alpha, IL10, IL4, IRGM, MBL2, MCP1 and CCL3L1 copy number were detected, using real-time PCR as previously described (14, 15, 16, 17, 18).

Statistical Analyses

In order to correct for heterogeneity of variance in the data, linear regression with a robust variance estimator (19) was used to test the association between the genetic markers and two outcome measures, HIV RNA and CD4 count. For each genetic marker, we used a multivariate test of the co-dominant model to determine whether there was an overall association across all categories of the variants and each of the outcome variables (HIV RNA and CD4 count) (20). The minor allele homozygotes with a frequency less than 2.5% were combined with the heterozygotes to avoid sparsely populated groups. Although genetic markers are invariant with age, because children ≤ 2 years old have more rapid disease progression than those >2 years old and have immature immune systems, it was hypothesized that the effects of host genetics on HIV disease may be different between younger and older children. Therefore, the interaction between each genetic marker and age group (age ≤2 years and >2 years) was investigated. For the genetic markers with at least a marginally significant genotype by age interaction ($p<0.1$), linear regression models were fitted to each age group separately. In the adjusted analyses, potential confounders, including age, race/ethnicity and study (P152 vs. P300) were used as adjusting covariates, which also removes the bias of the difference in CD4 count with respect to age.

In the context of performing exploratory analyses on a large number of genetic markers, setting the experiment-wise type one error rate to an exact, pre-specified level, using a procedure such as the Bonferroni correction (21), can be overly conservative, such that potentially important associations, which should be studied further, are not discovered. Conversely, without an appropriate correction, the probability of discovering false positive findings due to chance can be quite high. Thus, in exploratory research, where the goal is to generate hypotheses for further analysis while avoiding spurious findings that are due to chance, it is important to control the false discovery rate (FDR). In this paper, we apply the FDR controlling procedure developed by Benjamini and Hochberg (22) and evaluate the results with the FDR value set at 0.1, as well as 0.05, in order to discard findings likely due to chance, while identifying marginally significant, as well as highly significant, findings for further study. All genetic associations with a p-value <0.05 were included in the summary tables with 95% confidence intervals (CI). However, only P-values that correspond to a set of predictors with an FDR controlling at 0.05 were considered to be statistically significant, while p-values that correspond to an FDR between 0.1 and 0.05 were considered to be marginally significant.

RESULTS

Of the 1053 subjects, 60% were non-Hispanic black, 26% were Hispanic, 13% were non-Hispanic white, and 1% was of other or unknown race/ethnicity. Ages ranged from 44 days to 18 years (median age 2.35 years); 55% were female; 56% were from protocol P300 and 44% from protocol P152.

Baseline CD4 count, CD4% and $\log_{10}$ HIV RNA, stratified by the covariates age, race/ethnicity, and study are presented in Table 1. Of the 1053 subjects, 1045 had baseline CD4 measurements, with a median of 780 cells/mm^3 and a median CD4% of 24%; 871 had baseline HIV RNA measurements, with a median $\log_{10}$ HIV RNA value of 5.11. Age was categorized as ≤2 years and >2 years, providing a roughly balanced number of children in each age group. The older subjects, as would be expected, had significantly lower baseline CD4 count and $\log_{10}$ HIV RNA values, but also lower CD4%, compared to the younger subjects. Both baseline CD4 count and $\log_{10}$ HIV RNA are higher in subjects from P152 compared to the subjects in P300. Although the baseline CD4 count and $\log_{10}$ HIV RNA did not differ significantly across race/ethnicity and sex, these variables were included as covariates in the adjusted analyses presented below, because they had been specified a-priori as potential confounders.

Table 2 presents the results of multivariate analyses regressing baseline CD4 count on genetic markers, controlling for the covariates. For genetic markers whose effects on these outcome variables interacted at least marginally with age ($P < 0.1$), the regressions are presented separately for subjects ≤ 2 years and >2 years of age. A listing of all SNPs which were studied, along with their minor allele frequencies, which includes data from all of the 1053 subjects for whom the genetic allele data are available, is available upon request. Minor allele homozygotes with frequencies of less than 2.5% were marked with '*'.

With respect to the association of the genetic variants with baseline CD4 count, 6 HLA markers and 4 additional markers exhibited p-values <0.05 in the multivariate tests (Table 2). The presence of two of these HLA markers, A-24 and DRB1*7, exhibited no interaction with age, and their results are pooled across age groups. The effects of the 4 other HLA markers with p-values <0.05 did vary as a function of age, with the effects of 3 (DRB1*3, B-27/57, A-68) confined to the ≤ 2 year old age group and the effects of 1 (Cw-1) only apparent in the >2 year old group [Note that B-27/57 indicates that B27 and/or B57 were present]. All of the 4 non-HLA markers with p-values <0.05 had significant interactions with age, with the effects of 3 (APOBEC3G_197193-T/C, CCR5_wt/Δ32, IRGM_313-C/T) only evident in the ≤ 2 year old group and those of MBL2_A/O confined to the >2 year old group. The SNPs in Table 2 are presented from the largest (least significant) to the smallest in the multivariate Wald test p-values, along with the corresponding p-values at which the effects of each successive marker would be considered significant, when an FDR of 0.1 is applied. However, although these SNPs had exhibited nominal p-values < 0.05 for their associations with CD4 count, none met the criteria for statistical significance with the FDR controlled at 0.1.

With respect to the association of the genetic variants with baseline HIV RNA values, 10 HLA markers and 5 additional markers exhibited nominal p-values <0.05 in the multivariate tests (Table 3). Eight of these HLA markers exhibited no interaction with age (HLA_C homozygote/heterozygote, A-3, DRB1*15, DRB1*10, B-14, Cw-8, B57, B-27/57), and their results are pooled across age groups. The effects of HLA_A homozygote/heterozygote were only evident in the ≤2 year age group, while those of HLA_Cw-6 were confined to the >2 year age group. Four of the 5 non-HLA markers with p-values <0.05 had no significant interactions with age (MBL2_X/Y, CX3CR1_280-T/M, CCL3L1_Copy_Number, CCR5wt/Δ32), while effects of RANTES_28-C/G were confined to the younger group. The SNPs in Table 3 are presented from the largest (least significant) to the smallest in the multivariate Wald test p-values, along with the corresponding p-values at which the effects of each successive marker would be considered significant when an FDR of 0.1 is applied. Seven of the HLA markers, including A-3, DRB1*15, DRB1*10, B-27/57, B-14, Cw-8, B-57, along with CCR5_wt/Δ32 were marginally significant (FDR set at 0.1), while HLA markers DRB1*15, DRB1*10, B-27/57, B-14, Cw-8, B57 were statistically significant with the FDR set at 0.05.

DISCUSSION

Analyses examining the associations between genetic variants and baseline CD4 count and HIV RNA values in antiretroviral naïve HIV-infected children provide a unique opportunity to identify genetic markers that are related to HIV disease progression before antiretroviral therapy takes effect. The findings presented above allow the most promising genetic markers to be classified with respect to their potential as predictors of HIV disease status, measured by CD4 count or HIV RNA, as: 1) Strong evidence of association with FDR=0.05; 2) Promising evidence of association with nominal p <0.05 and FDR=0.1; and 3) Possible association, but considerable probability of chance findings, with nominal p <0.05 but FDR >0.1. The remaining markers exhibited less evidence of an association with HIV disease status with p ≥0.05 and FDR >0.1. Biological plausibility and consistency with previously

reported results must be taken into consideration in evaluating these findings and their implications for the role of the genetic variants that were studied.

The results presented in Table 3 suggest that HLA plays a key role in the control of HIV disease as measured by HIV RNA values. The findings which revealed the strongest evidence of a relationship between genomic markers and HIV severity were 6 HLA SNPs that were significantly associated with HIV RNA. For these associations, with FDR controlled at 0.05, the presence of HLA B-57, B27/57, B14, Cw-8 and DRB1*10 was associated with significantly lower baseline $\log_{10}$ HIV RNA values, while HLA DRB1*15 was associated with significantly higher baseline $\log_{10}$ HIV RNA. HLA A-3, which was marginally significant with FDR=0.1, also had the effect of increasing baseline $\log_{10}$ HIV RNA. Additionally, several other HLA effects were also observed in the RNA analyses where the results failed to meet the FDR criteria for statistical significance, but did show nominal p-values <0.05, e.g. HLA Cw-6, HLA_A homozygote/heterozygote, and HLA_C homozygote/heterozygote. The single non-HLA marker which exhibited at least a marginally significant relationship (FDR=0.1) with baseline $\log_{10}$ HIV RNA was CCR5, where children with the wt/Δ32 genotype had lower HIV RNA values than did those with CCR5 wt/wt homozygote. Each of these genetic variants has a strong immunologic basis for its role in the control of HIV infection (23, 24, 25). These findings suggest that HLA mediated adaptive immune responses play a critical role in the control of HIV infection and in establishing the viral set point in a given individual. The overwhelming proportion of HIV infections in the U.S. is clade B; thus, it is likely that these findings are representative of HIV-infected children in the U.S.

Recently, in a study of 572 HIV-infected children from our study population who were receiving NRTI treatment, it was shown that the presence of HLA B*57 and B*27/57 alleles was associated with delayed disease progression as measured by time to treatment failure (17). In a study of chronic HIV-1 infection in the absence of therapy, immunological and virologic outcomes for 121 African-American adolescents exhibited consistently favorable associations with HLA-B*57 (mostly B*5703) (26). Our current results, showing that B57 and B27/57 were associated with lower baseline HIV RNA, also suggest that the effects of these markers are not mediated by response to treatment, but represent direct immunologic control of HIV replication. With respect to the CCR5 gene (Δ32) effect, our results are consistent with previously reported findings (5, 6, 7, 14), and indicate that this effect is present in the absence of any antiretroviral therapy. This effect is likely mediated by the effect of the CCR5-wt/Δ32 on HIV entry into susceptible target cells.

It should be noted that the HIV RNA viral load level may depend upon whether a set point has been reached. In most HIV-1 infected individuals, certain host genetic factors, such as having a HLA-B57 or HLA-B27 type, have been associated with a lower viral load set point (27, 28, 29). A higher viral load set point over time has been associated with an accelerating clinical course of infection (30). The majority of children in our sample were perinatally infected, and infants usually reach their HIV RNA set-point within 6 months (31). Thus, it is possible that some of the youngest infants that we studied (ages 3–6 months) had not reached their viral set-points. Of the subjects whose source of HIV infection may have been other than perinatal acquisition and who may have been infected too recently to have

reached a set point, it was thought that the majority would be older children. Thus, a sensitivity analysis excluding children <6 months or >6 years, was conducted to determine whether the remaining subjects, who were most likely to have been at a HIV RNA set point, had findings consistent with those of the larger sample. The results indicated that this subgroup exhibited a similar pattern of associations as did the total sample, suggesting that the findings are not confounded by the fact that some subjects may not have been at a viral set point.

For the CD4 count analyses, no findings met the FDR criteria for statistical significance, but some HLA types had nominal p-values <0.05. This provides a further degree of evidence for the key role of HLA in the control of HIV. It is also worth noting that the presence of B*57/27 and CCR5 Δ32 heterozygosity were associated with large CD4 count effects of 386 and 492 cells/mm^3, respectively, in the younger age group ($p<0.05$; FDR >0.1), findings which are consistent with their effects on baseline HIV RNA. With sample sizes of 1045 and 871 subjects for the CD4 count and HIV RNA analyses, respectively, the analyses had considerable precision. Thus, the absence of a significant association between a given marker and CD4 count and/or HIV RNA suggests that the marker does not have a strong, direct effect on HIV when the 95% confidence interval excludes the strong effects (available upon request). However, our findings do not exclude the possibility that these genetic variants may have an important role in the innate or adaptive immunologic control of HIV.

One potential limitation of this study is that there may have been a selection bias, due to the fact that children who died of rapid HIV disease progression associated with a genetic factor prior to potential accrual into P152/P300 would not be represented in the study sample. However, under the assumption that the probability of surviving depends on the genetic marker, but not the outcomes (CD4 count/HIV RNA) after adjusting for age, the regression coefficients are not subject to bias if the analyses adjust for age (32). Survivor bias is not unique to our dataset and commonly exists in observational studies (33). To examine whether survivor bias may have caused certain genetic variants to be overrepresented among the older children in the sample, analyses assessing the association between age and genetic variants are performed, where strong associations may indicate the presence of a survivor effect. The table is not included in the publication of this article, but available upon request.

The current research was aimed at identifying genomic variants which may ultimately play key roles in understanding the genomic factors which determine HIV disease status. All of the genetic variants examined in this study were chosen because there was reason to believe that they might have an effect on HIV disease progression. A comprehensive list of the markers that were studied, which includes those where there was no significant evidence of an association with CD4 count or HIV RNA is available upon request. These markers had biologically plausible potential to affect HIV disease, but failed to exhibit an association with baseline CD4 count and HIV RNA. However, it is possible that they may predict HIV status as measured by other variables, or may combine with other genetic factors in more complex ways to influence CD4 count and HIV RNA.

In summary, accounting for multiple comparisons, for CD4 count no findings met the FDR criteria for statistical significance. In contrast, the presence of HLA B-57, B27/57, B14,

Cw-8 and DRB1*10 was associated with significantly lower baseline $\log_{10}$ HIV RNA values, while HLA DRB1*15 was associated with significantly higher baseline $\log_{10}$ HIV RNA. These findings support a strong association between specific HLA types and HIV replication in children.

Acknowledgments

Sources of Support: This research was supported in part by the Pediatric AIDS Clinical Trial Group (PACTG) [U01 AI041089] and International Maternal Perinatal Adolescent AIDS Clinical Trials (IMPAACT) Network. Overall support for the International Maternal Pediatric Adolescent AIDS Clinical Trials Group (IMPAACT) was provided by the National Institute of Allergy and Infectious Diseases (NIAID) [U01 AI068632], the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) (contract number N01-DK-9-001/HHSN267200800001C), and the National Institute of Mental Health (NIMH) [AI068632], 1R01NS077874 (SAS) and 5R01MH085608 (KKS).

The authors acknowledge and thank the P152/P300 protocol teams, participants and their families as well as site personnel for their contributions to the studies. The authors also acknowledge the Pediatric AIDS Clinical Trials Group (PACTG) for the support.

References

1. Fellay J. Review Host genetics influences on HIV type-1 disease. Antiviral Therapy. 2009:2040–2058.
2. Nasi M, Pinti M, Bugarini R, Troiano L, Lugli E, Bellodi C, Mussini C, Borghi V, Trenti T, Balli F, Esposito R, Cossarizza A. Genetic polymorphisms of Fas (CD95) and Fas ligand (CD178) influence the rise in CD4+ T cell count after antiretroviral therapy in drug-naïve HIV-positive patients. Immunogenetics. 2005 Oct; 57(9):628–35. [PubMed: 16158329]
3. Julg B, Moodley ES, Qi Y, Ramduth D, Reddy S, Mncube Z, Gao X, Goulder PJ, Detels R, Ndung'u T, Walker BD, Carrington M. Possession of HLA class II DRB1*1303 associates with reduced viral loads in chronic HIV-1 clade C and B infection. J Infect Dis. 2011 Mar 15; 203(6): 803–9. [PubMed: 21257739]
4. Serwanga J, Shafer LA, Pimego E, Auma B, Watera C, Rowland S, Yirrell D, Pala P, Grosskurth H, Whitworth J, Gotch F, Kaleebu P. Host HLA B*allele-associated multi-clade Gag T-cell recognition correlates with slow HIV-1 disease progression in antiretroviral therapy-naïve Ugandans. PLoS One. 2009; 4 (1):e4188. [PubMed: 19142234]
5. De Roda Husman AM, Koot M, Cornelissen M, Brouwer M, Broersen SM, Bakker M, et al. Association between CCR5 genotype and the clinical course of HIV-1 infection. Ann Intern Med. 1997; 127:882–890. [PubMed: 9382366]
6. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, Allikmets R, et al. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Science. 1996; 273:1856–1862. [PubMed: 8791590]
7. Meyer L, Magierowska M, Hubert J-B, Rouzioux C, Deveau C, Sanson F, et al. Early protective effect of CCR5 D32 heterozygosity on HIV-1 disease progression: relationship with viral load. AIDS. 1997; 11:F73–F78. [PubMed: 9302436]
8. Englund JA, Baker CJ, Raskino C, McKinney RE, Petrie B, Fowler MG, et al. Zidovudine, didanosine, or both as the initial treatment for symptomatic HIV-infected children. AIDS Clinical Trials Group (ACTG) Study 152 Team. N Engl J Med. 1997; 336:1704–1712. [PubMed: 9182213]
9. McKinney RE Jr, Johnson GM, Stanley K, Yong FH, Keller A, O'Donnell KJ, et al. A randomized study of combined zidovudine-lamivudine versus didanosine monotherapy in children with symptomatic therapy- naive HIV-1 infection. The Pediatric AIDS Clinical Trials Group Protocol 300 Study Team. J Pediatr. 1998; 133:500–508. [PubMed: 9787687]
10. Centers for Disease Control. Classification system for human immunodeficiency virus (HIV) infection in children under 13 years of age. MMWR Morb Mortal Wkly Rep. 1987; 36:225–36. [PubMed: 3031443]

11. Centers for Disease Control. Revised classification system for human immunodeficiency virus (HIV) infection in children under 13 years of age. MMWR Morb Mortal Wkly Rep. 1994; 43(RR-12):1–10.
12. Palumbo PE, Raskino C, Fiscus S, et al. Predictive value of quantitative plasma HIV RNA and CD4+ lymphocyte count in HIV-infected infants and children. JAMA. 1998; 279:756–61. [PubMed: 9508151]
13. Singh KK, Spector SA. Fidelity of whole-genome amplification of blood spot DNA for HLA typing and SNP analyses. Clin Genet. 2007; 72:156–159. [PubMed: 17661821]
14. Singh KK, Barroga CF, Hughes MD, Chen J, Raskino C, McKinney RE, Spector SA. Genetic influence of CCR5, CCR2, and SDF1 variants on human immunodeficiency virus 1 (HIV-1)-related disease progression and neurological impairment, in children with symptomatic HIV-1 infection. J Infect Dis. 2003 Nov 15; 188(10):1461–72. [PubMed: 14624371]
15. Singh KK, Hughes MD, Chen J, Spector SA. Genetic polymorphisms in CX3CR1 predict HIV-1 disease progression in children independently of CD4+ lymphocyte count and HIV-1 RNA load. J Infect Dis. 2005 Jun 1; 191(11):1971–80. [PubMed: 15871132]
16. Singh KK, Lieser A, Ruan PK, Fenton T, Spector SA. An age-dependent association of mannose-binding lectin-2 genetic variants on HIV-1-related disease in children. J Allergy Clin Immunol. 2008 Jul; 122(1):173–80. 180.e1–2. [PubMed: 18602571]
17. Singh KK, Gray PK, Wang Y, Fenton T, Trout RN, Spector SA. HLA alleles are associated with altered risk for disease progression and central nervous system impairment of HIV-infected children. J Acquir Immune Defic Syndr. 2011 May 1; 57(1):32–9. [PubMed: 21283014]
18. Singh KK, Wang Y, Gray KP, Farhad M, Brummel S, Fenton T, Trout R, Spector SA. Genetic variants in the host restriction factor APOBEC3G are associated with HIV-1-related disease progression and central nervous system impairment in children. J Acquir Immune Defic Syndr. 2013 Feb 1; 62(2):197–203. [PubMed: 23138837]
19. Pan W. On the Robust Variance Estimator in Generalized Estimating Equations. Biometrika. 2001; 88(3):901–906.
20. Lettre G, Lange C, Hirschhorn JN. Genetic Model Testing and Statistical Power in Population-Based Association Studies of Quantitative Traits. Gen Epidemiol. 2007; 31:358–362.
21. Bonferroni, CE. Volume in Onore di Ricarrdo dalla Volta. Universita di Firenza; 1937. Teoria statistica delle classi e calcolo delle probabilita; p. 1-62.
22. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society, B. 1995; 57:289–300.
23. Blackwell JM, Jamieson SE, Burgner D. HLA and infectious diseases. Clin Microbiol Rev. 2009; 22:370–385. [PubMed: 19366919]
24. Gandhi RT, Walker BD. Immunologic control of HIV-1. Annu Rev Med. 2002; 53:149–172. [PubMed: 11818468]
25. Carrington M, O'Brien SJ. The influence of HLA genotype on AIDS. Annu Rev Med. 2003; 54:535–551. [PubMed: 12525683]
26. Shrestha S, Aissani B, Song W, Wilson CM, Kaslow RA, Tang J. Host genetics and HIV-1 viral load set-point in African-Americans. AIDS. 2009 Mar 27; 23(6):673–7. [PubMed: 19276793]
27. Kaslow RA, Carrington M, Apple R, Park L, Munoz A, Saah AJ, et al. Influence of combinations of human major histocompatibility complex genes in the course of HIV-1 infection. Nature Med. 1996; 2:405–411. [PubMed: 8597949]
28. Migueles SA, Sabbaghian MS, Shupert WL, Bettinotti MP, Marincola FM, Martino L, et al. HLA BM5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors. Proc Natl Acad Sci U S A. 2000; 97:2709–2714. [PubMed: 10694578]
29. Navis M, Schellens I, van Baarle D, Borghans J, van Swieten P, Miedema F, et al. Viral replication capacity as a correlate of HLA B57/B5801-associated nonprogressive HIV-1 infection. J Immunol. 2007; 179:3133–3143. [PubMed: 17709528]
30. Gras L, Jurriaans S, Bakker M, van Sighem A, Bezemer D, Fraser C, et al. Viral load levels measured at set-point have risen over the last decade of the HIV epidemic in the Netherlands. PLoS One. 2009; 4:e7365. [PubMed: 19809510]

31. Richardson BA, Mbori-Ngacha D, Lavreys L, et al. Comparison of Human Immunodeficiency Virus Type 1 Viral Loads in Kenyan Women, Men, and Infants during Primary and Early. Infection J Virol. 2003 Jun; 77(12):7120–7123.

32. Glynn, RJ.; Laird, NM. Technical Report. Harvard School of Public Health, Department of biostatistics; 1986. Regression estimates and missing data: complete-case analysis.

33. Glesby MJ, Hoover DR. Survivor Treatment Selection Bias in Observational Studies: Examples from the AIDS Literature. Ann Intern Med. 1996; 124(11):999–1005. [PubMed: 8624068]

Table 1

Summary Table of Baseline CD4 Count, CD4% and HIV RNA

Outcome	Variable	Category	N (%)	Mean (95% CI)	P-value
Baseline CD4 Count	All Subjects		1045	987(937,1038)	
	Age	(0,2]	478(45.7%)	1406(1316,1496)	
		(2,18)	567(54.3%)	635(599,670)	<0.001
	Race	Black	621(59.4%)	976(910,1042)	
		Hispanic	270(25.8%)	973(873,1072)	0.960
		Other	17(1.6%)	1107(763,1450)	0.526
		White	137(13.1%)	1054(909,1200)	0.321
	Sex	Female	574(54.9%)	986(917,1056)	
		Male	471(45.1%)	989(914,1064)	0.963
	Study	152	453(43.3%)	1059(978,1140)	
		300	592(56.7%)	933(868,998)	0.016
Baseline CD4%	All Subjects		1045	24(23,25)	
	Age	(0,2]	478(45.7%)	27(26,28)	
		(2,18)	567(54.3%)	22(21,22)	<0.001
	Race	Black	621(59.4%)	24(23,25)	
		Hispanic	270(25.8%)	24(22,25)	0.841
		Other	17(1.6%)	26(22,30)	0.479
		White	137(13.1%)	26(24,28)	0.081
	Sex	Female	574(54.9%)	25(24,26)	
		Male	471(45.1%)	23(22,24)	0.019
	Study	152	453(43.3%)	25(23,26)	
		300	592(56.7%)	24(23,24)	0.152
Baseline Log_{10} HIV RNA	All Subjects		871	5.1(5.0,5.2)	
	Age	(0,2]	398(45.7%)	5.6(5.5,5.7)	
		(2,18)	473(54.3%)	4.7(4.6,4.7)	<0.001
	Race	Black	526(60.4%)	5.1(5.0,5.2)	

Outcome	Variable	Category	N (%)	Mean (95% CI)	P-value
		Hispanic	216(24.8%)	5.0(4.9,5.1)	0.164
		Other	11(1.3%)	4.9(3.9,5.9)	0.517
		White	118(13.5%)	5.3(5.1,5.4)	0.093
	Sex	Female	484(55.6%)	5.0(5.0,5.1)	
		Male	387(44.4%)	5.2(5.1,5.2)	0.072
	Study	152	320(36.7%)	5.2(5.1,5.3)	
		300	551(63.3%)	5.0(4.9,5.1)	0.005

Table 2

Analysis of Baseline CD4 Count Conditional on Genetic Markers, Adjusted by Age, Race, Study

Gene	SNP	Age X Genotype Interaction	Age Group	Level	Count	Adjusted Analysis			P-value cut off under FDR control at 0.1	Significant Controlling for FDR at 0.1
						Slope (LCI, UCI)	P-value	Multivariate Wald Test P-value		
HLA	HLA_A-24 alleles	0.3208	All	Absence	919	1419(1267,1571)		0.043	0.010	-
				Presence	106	−130(−256,−4)	0.043			
HLA	HLA_A-68 alleles	0.0496	(0, 2]	Absence	408	1436(1190,1681)		0.036	0.009	-
				Presence	64	−224(−434,−14)	0.036			
HLA	HLA_B-27/57 alleles	0.0152	(0, 2]	Absence	432	1344(1081,1608)		0.025	0.008	-
				Presence	39	380(48,712)	0.025			
HLA	HLA_DRB1*7 alleles	0.2431	All	Absence	829	1362(1207,1517)		0.020	0.007	-
				Presence	189	152(24,280)	0.020			
IRGM	IRGM_313-C/T	0.0267	(0, 2]	C/C	218	1403(1156,1649)		0.017	0.006	-
				C/T	190	89(−98,276)	0.353			
				T/T	70	470(148,792)	0.004			
CCR5	CCR5_wt/delta32	0.0829	(0, 2]	wt/wt	448	1355(1097,1612)		0.016	0.005	-
				wt/D32	23	482(89,876)	0.016			
APOBEC3G	APOBEC3G_197193-T/C	0.0533	(0, 2]	T/T	416	1434(1189,1678)		0.016	0.004	-
				C/C+T/C	61	−270(−490,−51)	0.016			
HLA	HLA_Cw-1 alleles	0.0436	(2, 18]	Absence	523	657(538,776)		0.014	0.003	-
				Presence	29	−162(−291,−32)	0.014			
MBL2	MBL2_A/O	0.0250	(2, 18]	A/A	308	646(527,765)		0.007	0.002	-
				A/O	220	−32(−103,40)	0.384			
				O/O	36	229(69,389)	0.005			
HLA	HLA_DRB1*3 alleles	0.0243	(0, 2]	Absence	368	1483(1229,1736)		0.007	0.001	-
				Presence	103	−269(−463,−75)	0.007			

Note: This table only includes the SNPs with multivariate Wald test nominal p-value <0.05. P-values are compared to the corresponding cut off values and p-values that are less than the cut off value are considered as significant under the FDR control at the pre-specified level.

-: Not significant at FDR=0.1

Table 3

Analysis of Baseline HIV Log_{10}RNA Conditional on Genetic Markers, Adjusted by Age, Race, Study

Gene	SNP	Age X Genotype Interaction	Age Group	Level	Count	Adjusted Analysis			P-value cut off under FDR control at 0.1	Significant Controlling for FDR
						Slope (LCI, UCI)	P-value	Multivariate Wald Test P-value		
MBL2	MBL2_X/Y	0.6832	All	Y/Y	643	5.56(5.40,5.72)		0.041	0.017	-
				X/X	22	0.36(0.08,0.64)	0.012			
				X/Y	203	0.04(−0.09,0.17)	0.520			
HLA	HLA_C alleles	0.3946	All	Homozygote	100	5.75(5.53,5.97)		0.031	0.016	-
				Heterozygote	757	−0.19(−0.36,−0.02)	0.031			
HLA	HLA_A alleles	0.0096	(0, 2]	Homozygote	41	5.97(5.69,6.24)		0.026	0.014	-
				Heterozygote	352	−0.27(−0.50,−0.03)	0.026			
RANTES	RANTES_28-C/G	0.0186	(0, 2]	C/C	391	5.74(5.57,5.92)		0.020	0.013	-
				G/G+C/G	7	0.42(0.07,0.78)	0.020			
HLA	HLA_Cw-6 alleles	0.0982	(2, 18]	Absence	391	4.60(4.40,4.79)		0.019	0.012	-
				Presence	72	0.22(0.03,0.40)	0.019			
CX3CR1	CX3CR1_280-T/M	0.8325	All	C/C	763	5.52(5.36,5.69)		0.019	0.011	-
				T/T+C/T	105	0.19(0.03,0.35)	0.019			
CCL3L1	CCL3L1_Copy Number	0.5775	All	[0, 2]	494	5.58(5.42,5.74)		0.016	0.010	-
				>=7	50	0.26(0.05,0.48)	0.015			
				[3, 4]	235	−0.05(−0.18,0.09)	0.495			
				[5, 6]	89	0.17(−0.02,0.36)	0.074			
HLA	HLA_A-3 alleles	0.4840	All	Absence	718	5.53(5.36,5.69)		0.008	0.009	*
				Presence	138	0.19(0.05,0.33)	0.008			
CCR5	CCR5_wt/delta32	0.1969	All	wt/wt	815	5.63(5.46,5.80)		0.005	0.008	*
				wt/D32	47	−0.35(−0.59,−0.11)	0.005			
HLA	HLA_DRB1*15 alleles	0.3394	All	Absence	610	5.52(5.35,5.68)		0.002	0.007	**
				Presence	241	0.18(0.07,0.30)	0.002			
HLA	HLA_DRB1*10 alleles	0.4121	All	Absence	821	5.59(5.43,5.75)		0.002	0.006	**

Gene	SNP	Age X Genotype Interaction	Age Group	Level	Count	Adjusted Analysis			P-value cut off under FDR control at 0.1	Significant Controlling for FDR
						Slope (LCI, UCI)	P-value	Multivariate Wald Test P-value		
				Presence	30	−0.45(−0.73,−0.17)	0.002			
HLA	HLA_B-27/57 alleles	0.6551	All	Absence	769	5.65(5.48,5.81)		<0.001	0.004	**
				Presence	84	−0.36(−0.57,−0.16)	<0.001			
HLA	HLA_B-14 alleles	0.6788	All	Absence	793	5.59(5.44,5.75)		<0.001	0.003	**
				Presence	60	−0.39(−0.60,−0.18)	<0.001			
HLA	HLA_Cw-8 alleles	0.7843	All	Absence	765	5.60(5.44,5.76)		<0.001	0.002	**
				Presence	92	−0.46(−0.65,−0.26)	<0.001			
HLA	HLA_B-57 alleles	0.9304	All	Absence	798	5.63(5.47,5.79)		<0.001	0.001	**
				Presence	55	−0.54(−0.77,−0.31)	<0.001			

Note: This table only includes the SNPs with multivariate Wald test nominal p-value <0.05. P-values are compared to the corresponding cut off values and p-values that are less than the cut off value are considered as significant under the FDR control at the pre-specified level.

-: Not significant at FDR=0.1

*: Significant at FDR=0.1

**: Significant at FDR=0.1 and FDR=0.05