Published in final edited form as:

AIDS. 2009 March 27; 23(6): 673–677. doi:10.1097/QAD.0b013e328325d414.

Host Genetics and HIV-1 Viral Load Set-point in African-Americans

Sadeep Shrestha^{1,*}, Brahim Aissani¹, Wei Song¹, Craig M. Wilson^{1,2}, Richard A. Kaslow^{1,2}, and Jianming Tang²

¹Department of Epidemiology, University of Alabama at Birmingham, Birmingham, Alabama.

²Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama.

Abstract

Objective—In a recent genome-wide association study of HIV-1-infected individuals in the Euro-CHAVI cohort, viral load set-point was strongly associated with genotypes defined by two SNPs (rs9264942 and rs2395029) within the human MHC region on chromosome 6. We attempted to confirm this finding in African-Americans and assess if these SNPs are in linkage disequilibrium (LD) with HLA class I alleles that mediate innate and adaptive immunity.

Design—Our analyses relied on 121 African American adolescents with chronic HIV-1 infection and quarterly immunological and virological outcome measures in the absence of therapy.

Methods—PCR-based techniques were used to genotype two SNPs along with HLA class I alleles. Their associations with HIV-1 viral load set-point and longitudinal CD4⁺ and CD8⁺CD38⁺ T-cell counts were tested in univariate and multivariate models.

Results—The CC genotype at rs9264942 was associated with reduced viral load but not with immunological outcomes or category of disease control. Consistent associations of HLA-B*57 (mostly B*5703) with favorable virological and immunological outcomes were observed, but not rs2395029G allele at the *HCP5* locus, which is in absolute linkage disequilibrium with B*5701 (in individuals of European descent), and not B*5703.

Conclusion—While rs9264942 and B*57 (but not rs2395029G) are clearly associated with control of viral load set-point among African-Americans, fine-mapping of MHC SNPs in populations of African and European descent should help reveal the true variants and the underlying functional mechanisms.

Keywords

HIV-1; genetics; viral load; African American

Introduction

Immunological and clinical outcomes in the natural history of HIV-1 infection and progression to AIDS can vary considerably both at the individual or population level; part of such variation

Corresponding Author: Sadeep Shrestha, Department of Epidemiology, University of Alabama at Birmingham, 1665 University Blvd, RPHB Room 217L, Birmingham, Alabama 35294-0022, Phone: (205) 934 6459; Fax: (205) 934 8665, E-mail: sshresth@uab.edu.

Authors' contributions

Sadeep Shrestha supervised the study design and execution, performing the data analyses and writing the manuscript. James Tang and Richard A Kaslow provided immunogenetic expertise, Brahim Aissani provided analytic support for the study and Wei Song provided support with laboratory methods. Craig Wilson assisted with clinical data of the REACH participants. All the authors participated in the writing and the reading of the article and have seen and approved the final version of the manuscript.

is attributable to a number of host genetic factors. In particular, several allelic variants in genes encoding chemokine receptors and ligands, cytokines, human leukocyte antigen (HLA) and products mediating cellular immunity have been shown to influence HIV-1 infection, progression of immunodeficiency, or various AIDS-defining outcomes [1–5]. However, these known genes and their variants do not fully explain the highly variable course of HIV-1 infection or its pathogenic mechanisms in most populations. Indeed, host genetic contributions to HIV-1 infection and subsequent disease progression have been reported mostly in individuals of European descent, with fewer confirmations in other populations or in experimental systems. In general, genetic variation and disease association studies are known to be confounded by differences in human genetic makeup and in disease risk associated among populations of specific ethnic origin with distinctive genetic profiles. Thus replication of immunogenetic research findings in major ethnic groups should help distinguish population-specific relationships from those that are more widely applicable.

A quantifiable outcome in HIV-1 infected individuals is the level of viremia in patient plasma. After the acute phase of HIV-1 infection, viremia (plasma viral load) usually stabilizes at a steady state commonly known as the “set-point.” Set-point viral load results from the equilibrium between the HIV-1 virus and immune response and is an important predictor of HIV-1 disease progression [6]. Fellay et al recently reported that two independent single nucleotide polymorphisms (SNPs), i.e., rs9264942 (T/C) about 35 kb upstream of the *HLA-C* gene and rs2395029 (T/G) in HERV-derived HLA complex P5 gene (*HCP5*), were associated with low viral load set-point in a multi-center genome-wide association study (GWAS) of treatment-naïve Europeans [7]. We aimed to determine if the novel findings derived from the GWAS also apply to African American population with a different genetic profile.

Materials and Methods

We analyzed participants in the Reaching for Excellence in Adolescent Care and Health (REACH) study. The characteristics of the cohort, recruitment, follow-up and observational research objectives are described in detail elsewhere [8,9]. Briefly, HIV-1 seropositive and high-risk seronegative adolescents (13 to 19 years old) were recruited from 15 clinical sites in 13 US cities. Participants were followed on a quarterly basis for basic science and clinical evaluations, including documentation of demographic and risk behaviors, collection of medical history and various biological samples, along with tests for HIV-1 infection, other sexually transmitted infections, CD4⁺ cell counts, and immunological outcomes. HIV-1 RNA concentration (viral load) and immunological outcomes were measured every three months. CD4⁺ and CD8⁺CD38⁺ lymphocytes were quantified by flow cytometry in National Institute of Allergy and Infectious Disease (NIAID)-certified laboratories at the clinical sites. Plasma viral load was measured in a centralized laboratory using either nucleic acid sequence-based amplification (NASBA, lower limit = 400 copies/mL) or NucliSens assays (Organon Teknika, Durham, NC, lower limit = 80 copies/mL). Viral loads under the lower limit of detection were assigned half of log₁₀ (lower detection limit) values for each respective visits. All individuals included in the analysis had multiple measurements (≤ 4 visits) of HIV-1 RNA (transformed to log₁₀ for normalization), CD4⁺ count, and CD8⁺CD38⁺ count at treatment-free visits.

Our analyses focused on a subset of African American adolescents [8] who were free of clinical AIDS and not receiving any antiretroviral therapy (ART) at multiple 3-month follow-up intervals (total = 396 person-visits). All participants had acquired HIV-1 infection through sexual activity or injection drug use. Although these relatively young patients were HIV-1 seropositive at baseline, vertical transmission could be excluded and seroconversion was deemed recent as judged by i) self-reported sexual and other risk behaviors, ii) high CD4⁺ count and iii) relatively stable viral load that could be treated as a proxy for viral load set-point [8,10]. We estimated the viral load set-point based on observations from 2–4 sequential visits,

which is consistent with earlier reports that it is cost-effective to estimate the viral set-point based on one or two measurements obtained between 5 and 12 months after HIV-1 infection [10].

High-resolution HLA class I (*HLA-A*, *HLA-B* and *HLA-C*) typing was performed by a combination of molecular techniques, including automated DNA hybridization with sequence-specific oligonucleotide (SSO) probes, PCR with sequence-specific primers (SSP), and automated sequencing-based typing (SBT). SNP genotyping was performed by the allelic discrimination TaqMan assay (Applied Biosystems). We tested for consistency with Hardy-Weinberg equilibrium expectation using one-way goodness-of-fit chi-square. To confirm the specificity of the rs9264942 genotyping (due to several SNPs in the region), we sequenced the regions (flanking 250 bp each side) in a selected subset of African American individuals from the entire REACH cohort (19 with B*5703, 4 with B*5701, and 45 with other *HLA-B* genotypes). We assessed patterns of linkage disequilibrium (LD) of rs9264942 along with other variants in the sequenced region with *HLA-B* alleles (specifically B*5701 and B*5703) and alleles of *HLA-A* and *HLA-C*.

Among those participants with multiple clinical visits in the study, we excluded the measurements from the visits where the subsequent viral loads did not appear to be in steady state, i.e. inter-visit viral load variation exceeded 3-fold ($0.5 \log_{10}$) copies per ml of plasma. Individuals with no more than two steady state viral loads were excluded from the study. We used Proc Mixed in SAS version 9.2 (SAS Institute, Cary, NC) to fit the uneven repeated measurement ANOVA model and estimated mean viral load [11] by genetic variant (allele or genotype) after accounting for confounding covariates and multiple comparisons using Tukey-Kramer adjustment. To account for the differences in the evaluation between viral load observations which are close to each other than to those which are further apart we used autoregressive covariance structure. This method is powerful in providing the precision of the mean and the standard error of the viral set-point with such uneven viral load data over uneven period of time. We also performed alternate analysis by first estimating the viral load set point as the average of multiple viral-load measurements and used the standard linear regression model. Additionally, we used the mixed model to assess how genotype influenced CD4⁺ and CD8⁺CD38⁺ over time, post viral load set-point, as surrogates for treatment free HIV progression. We also determined the allelic and genotypic heterogeneity among 19 “controllers” (VL<1000 copies/ml and CD4 >450 × 10⁶ cell/l), 23 “non-controllers” (VL>16,000 copies/ml and CD4 <450 × 10⁶ cell/l) and 79 “intermediates”, as previously described [12]. Categorical outcomes were assessed by χ^2 and Cochran-Armitage trend tests and the continuous outcomes ($\log_{10}$ viral load and absolute CD4⁺) by t-tests and F-tests.

Results

In all, 39 out of 160 eligible patients did not meet the inclusion criteria for inter-visit viral load variation to estimate the viral set-point and were excluded from the analyses. Among 121 eligible African American patients (median age = 16.8), 102 (84%) were females and the excluded group (39 subjects) did not differ in comparison. Consistent with other female cohorts, the viral load set-point in our cohort of predominantly females is relatively lower. Using both mixed model and standard linear regression model approaches showed similar results and here we only present the results from a more conservative mixed model approach. In a univariate model, individuals with B*57 alleles (mostly B*5703) had significantly lower mean viral load than others without B*57 (Figure 1a, 2.75 ± 0.20 vs. 3.70 ± 0.07 ; $p=0.0002$). We also confirmed a strong gradient of B*57 frequency among “controllers” (67%), “intermediates” (33%) and “non-controllers” (0) ($p<0.0001$ for Cochran-Armitage trend test). Also, among individuals with B*57, absolute CD4⁺ were higher by an average of 216 ± 60 cells

(785 ± 56 vs. 569 ± 20 , $p=0.02$) during the viral set-point period, but no difference in $CD8^+CD38^+$ ($p=0.45$). The *HCP5* rs2395029 (G) allele was rare (0.16%) among the 121 study participants (close to reported (0.2%) in general African American population); it appeared to be in complete linkage disequilibrium with less frequent HLA-B*5701 ($r^2=1$, $n=4$) as previously reported [7] but absent in a total of 19 HLA-B*5703 carriers. Sequencing of the *HCP5* region also revealed a dense set of SNPs within the 500 bp region; rs17206855 (T/C, 10%), rs2255221 (G/T, 23%), rs2255223 (G/A, 3%), rs11752262 (A/G, 13%), rs2395029 (T/G, 5%), rs3130907 (A/G, 3%), rs2243621 (C/T, 21%), rs2395030 (G/T, 1%), and rs2263318 (G/A, 21%), but there was no distinct LD of either HLA B*57 or rs2395029 with any of these polymorphisms. B*57 trended to be in LD with both Cw*06 ($r^2 = 0.50$) and Cw*18 ($r^2 = 0.44$), but these *HLA-C* variants showed no clear associations by themselves.

In our study population, the minor allele (C) of rs9264942 SNP had a frequency of 32%, slightly lower than the frequency observed by Fellay et al among Europeans (41%). The rs9264942 SNP genotypes did not deviate from Hardy-Weinberg equilibrium ($p = 0.48$) and allele C was associated with reduced HIV-1 viral load in the African American adolescents (Figure 1b). In particular, individuals with the rs9264942CC genotype (frequency = 8%) had lower viral load (mean $\pm$ stderror = $2.96 \pm 0.23 \log_{10}$) than the TC heterozygotes (frequency = 47.5%) ($3.69 \pm 0.10 \log_{10}$, $p = 0.0085$) and TT homozygotes (frequency = 44.5%) ($3.64 \pm 0.10 \log_{10}$, $p = 0.004$), suggesting that the association of the C allele (38.5%) with lower viral load fit a recessive model. The favorable association of CC genotype with reduced viral load, however, did not extend to immunological outcomes ($CD4^+$, $p=0.11$ and $CD8^+CD38^+$, $p=0.34$) or disease progression (controllers vs intermediates vs non-controllers, $p=0.42$, Cochran-Armitage trend test). In multivariate models, HIV-1 viral load set-point did not differ statistically by gender (small male sample size) or age, while the genetic associations remained independent (B*57⁺ vs. B*57⁻: 2.53 ± 0.21 vs. $3.53 \pm 0.08 \log_{10}$ viral load, $p<0.0001$; and rs9264942: CC, 2.56 ± 0.23 vs. 3.20 ± 0.14 for CT and 3.33 ± 0.12 for TT, $p=0.004$).

Discussion

SNP polymorphisms implicated by the recent GWAS as influential to HIV-1 viral load set-point in infected Europeans may provide important benchmarks for the fine mapping of functionally critical genotypes that mediate HIV-1 pathogenesis. Here, we partially confirmed the effect of rs9264942C genotype in an African American population. Our results also support that B*57 and not the rs2395029G genotype in the *HCP5* region is associated with highly favorable virological and immunological outcomes of HIV-1 infection. B*5701, which is very rare (allele frequency = 0.4%) in African-Americans [13], is in perfect LD with rs2395029G in *HCP5*. In contrast, B*5703 (allele frequency = 2.4%) as the most common B*57 allele in this ethnic group [13] was not found in LD with *HCP5* SNP allele. Notably, none of the 16 predominantly African-American elite controllers studied by Han et al. had rs2395029G [14].

While B*57 and rs2395029 SNP might be surrogates for a true variant in LD, B*57 alleles have consistently been associated with better control of HIV-1 infection [5,15]. HLA-B*5701 and B*5703 differ by an aspartic acid-to-asparagine substitution at amino acid position 114 and a serine-to-tyrosine change at position 116 in the alpha 2 helix. Although both of these amino acids are located in structural proximity of the F pocket and contribute to peptide-binding groove, neither appears to significantly impact the structure of the P9 anchoring residue of antigenic peptides; indeed, both alleles are known to share similar peptide-binding preferences, with a dominant response to the same HIV-1 Gag epitope KF11 (KAFSPEVIPMF), and to induce potent CD8 response against several p24 epitopes [16].

Unlike B*57, the rs9264942CC genotype was associated only with viral load set-point, as previously observed in GWAS [7]. We thoroughly examined the pattern of LD of this SNP

and HLA-B*57 alleles with markers in *HLA-A*, *HLA-C*, and the neighboring *HCP5* gene region but found no other nearby marker that could explain the observed relationships. HLA-B*57 (and not rs2395029G) was the major factor associated with immunological and virological control of HIV-1 infection in African-American adolescents, while the rs9264942CC genotype was associated with modest impact on viral load without influencing CD4⁺ T-cells.

Major progress in genotyping technology and understanding of the human genome has led to a rapid increase of genetic association studies of complex diseases including HIV-1 infection. With the completion of Human Genome Project and International HapMap Project, use of population-specific SNPs scattered throughout the human genome is considered useful in addressing how individual genotypes or their combinations can influence disease outcomes [17]. Our observations illustrate the importance of replicating initial findings especially from GWAS in one major population to others with varying LD patterns and haplotype distributions. Such replication is critical in establishing causal relationships and providing insights into their biological mechanisms. HLA-B*57 remains a major protective genetic factor among populations of African-ancestry, where the public health burden of HIV/AIDS is substantial. Future studies with fine-mapping efforts in this MHC region should help explain whether B*57 has direct biological relevance or it indeed tags other true functional variants as a result of LD. Through continuing translational and applied research, understanding how genetic variations contribute to effective immune control of HIV-1 infection in populations with the greatest disease burden should provide insights to drug development and new interventional strategies.

Acknowledgement

We thank investigators and staff of the Adolescent Medicine HIV/AIDS Research Network and the youth who participated in the REACH project for their valuable contributions. We also thank our entire team within the Program in Epidemiology of Infection and Immunity (PEII) in the Department of Epidemiology for excellent technical assistance. There is no conflict of interest.

The parent study and this sub-study conformed to the procedures for informed consent (parental permission was obtained wherever required) approved by institutional review boards at all sponsoring organizations and to human-experimentation guidelines set forth by the United States Department of Health and Human Services. The REACH study was supported by grant U01-HD32830 from National Institute of Child Health and Human Development. This work was supported in part by a developmental award from the UAB Center for AIDS Research (5P30 AI27767-20 and by grants R01-AI41951 and R01-AI51173 from National Institute of Allergy and Infectious Diseases.

References

1. O'Brien SJ, Nelson GW. Human genes that limit AIDS. *Nature Genetics* 2004;36:565–574. [PubMed: 15167933]
2. Shrestha S, Strathdee SA, Galai N, Oleksyk T, Fallin MD, Mehta S, et al. Behavioral risk exposure and host genetics of susceptibility to HIV-1 infection. *J Infect Dis* 2006;193:16–26. [PubMed: 16323127]
3. Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, et al. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science*. 2005
4. Telenti A, Goldstein DB. Genomics meets HIV-1. *Nat Rev Microbiol* 2006;4:865–873. [PubMed: 17041633]
5. Tang J, Kaslow RA. The impact of host genetics on HIV infection and disease progression in the era of highly active antiretroviral therapy. *Aids* 2003;17:S51–S60. [PubMed: 15080180]
6. Mellors JW, Margolick JB, Phair JP, Rinaldo CR, Detels R, Jacobson LP, Munoz A. Prognostic value of HIV-1 RNA, CD4 cell count, and CD4 Cell count slope for progression to AIDS and death in untreated HIV-1 infection. *Jama* 2007;297:2349–2350. [PubMed: 17551128]
7. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, et al. A whole-genome association study of major determinants for host control of HIV-1. *Science* 2007;317:944–947. [PubMed: 17641165]

8. Rogers AS, Futterman DK, Moscicki AB, Wilson CM, Ellenberg J, Vermund SH. The REACH Project of the Adolescent Medicine HIV/AIDS Research Network: design, methods, and selected characteristics of participants. *J Adolesc Health* 1998;22:300–311. [PubMed: 9561462]
9. Wilson CM, Houser J, Partlow C, Rudy BJ, Futterman DC, Friedman LB. The REACH (Reaching for Excellence in Adolescent Care and Health) project: study design, methods, and population profile. *J Adolesc Health* 2001;29:8–18. [PubMed: 11530299]
10. Mei Y, Wang L, Holte SE. A comparison of methods for determining HIV viral set point. *Stat Med* 2008;27:121–139. [PubMed: 17787030]
11. Kelley CF, Barbour JD, Hecht FM. The relation between symptoms, viral load, and viral load set point in primary HIV infection. *J Acquir Immune Defic Syndr* 2007;45:445–448. [PubMed: 17514014]
12. Tang J, Wilson CM, Meleth S, Myracle A, Lobashevsky E, Mulligan MJ, et al. Host genetic profiles predict virological and immunological control of HIV-1 infection in adolescents. *Aids* 2002;16:2275–2284. [PubMed: 12441799]
13. Cao K, Hollenbach J, Shi X, Shi W, Chopek M, Fernandez-Vina MA. Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Hum Immunol* 2001;62:1009–1030. [PubMed: 11543903]
14. Han Y, Lai J, Barditch-Crovo P, Gallant JE, Williams TM, Siliciano RF, Blankson JN. The role of protective HCP5 and HLA-C associated polymorphisms in the control of HIV-1 replication in a subset of elite suppressors. *Aids* 2008;22:541–544. [PubMed: 18301071]
15. Altfeld M, Addo MM, Rosenberg ES, Hecht FM, Lee PK, Vogel M, et al. Influence of HLA-B57 on clinical presentation and viral control during acute HIV-1 infection. *Aids* 2003;17:2581–2591. [PubMed: 14685052]
16. Yu XG, Lichterfeld M, Chetty S, Williams KL, Mui SK, Miura T, et al. Mutually exclusive T-cell receptor induction and differential susceptibility to human immunodeficiency virus type 1 mutational escape associated with a two-amino-acid difference between HLA class I subtypes. *J Virol* 2007;81:1619–1631. [PubMed: 17121793]
17. Deloukas P, Bentley D. The HapMap project and its application to genetic studies of drug response. *Pharmacogenomics J* 2004;4:88–90. [PubMed: 14676823]

Figure 1.

Genetic variants associated with viral load set-point among 121 HIV-1-infected African-Americans. Both B*57 (a: B*57⁺ vs. B*57⁻, p<0.0001) and the rs9264942CC genotype (b; TT+TC vs. CC, p=0.004) are associated with lower viral load. Mean values and the standard error groups defined by the respective genotypes.