Multi-ancestry GWAS Identifies Novel Variants Associated with HIV-1 Viral Load Set-point

Eric O. Johnson¹, Dana B. Hancock¹, Nathan C. Gaddis¹, Christina A. Markunas¹, Yuelong Guo¹, Laura J. Bierut², Leigh Pearce³; Elizabeth Golub⁴; Howard Minkoff⁵; Chenglong Liu⁶; Mardge Cohen⁷; Mark Kuniholm⁸; Kathy Anastos⁹; Ruth M. Greenblatt¹⁰; Alex H. Kral¹¹, Bradley E. Aouizerat¹²

¹RTI International, Research Triangle Park, NC, USA; ²Washington University, St. Louis, MO; ³University of Southern California, Los Angeles, CA, USA; ⁴The Johns Hopkins University, Baltimore, MD, USA; ⁵SUNY Down State Medical Center, Brooklyn, NY, USA; ⁶Georgetown University, Washington DC, USA; ⁷Hektoen Institute for Medical Research, Chicago, IL, USA; ⁸Albert Einstein College of Medicine, Bronx, NY, USA; ⁹Montefiore Medical Center, Bronx, NY, USA; ¹⁰University of California, San Francisco, CA, USA; ¹¹RTI International, San Francisco, CA, USA; ¹²Bluestone Center for Clinical Research, New York University, New York, NY, USA.

Due to substantial reductions in HIV and AIDS incidence, HIV is now as a chronic disease in developed countries. HIV viral load (VL) is a predictor of time until HIV progression, a key measure of risk and treatment response, and a critical focal point for research. This is particularly the case among people who inject drugs (PWIDs), for whom evidence suggests HIV progression may be accelerated. Prior genome-wide association studies (GWAS) of the quantitative trait VL set-point (VLSP) in European-descent samples found replicable variant associations in the HLA gene region. We conducted the first multi-ancestry GWAS of VLSP, followed by RNA expression quantitative trait loci (eQTL) analyses to assess putative function of nominated novel variants. Discovery analyses used ~20 million 1000 Genomes imputed SNPs and indels among 705 African Americans (AAs), 215 European Americans (EAs), and 110 Hispanic HIV+ participants from the Women’s Interagency HIV Study (WIHS). Replication was tested using the Urban Health Study (UHS) HIV+ sample of 531 AAs and 258 EAs. We assessed potential function of novel variants as cis-eQTLs using data from the Genotype-Tissue Expression project (GTEx; release v6), with replication tested in the GEUVADIS consortium data. One peak was identified at p<5.0×10⁻⁸ on chromosome 6 within the HLA-B gene. The 44 genome-wide significant follow-up variants constituted 14 independent tests: multiple testing p-value <0.0036. Eighteen variant associations were replicated: rs146647111 was the top replication variant (WIHS discovery P = 4.7×10⁻¹⁶; UHS replication P = 5.3×10⁻⁵). Rs146647111 remained associated with VLSP after adjusting for all 8 known VL associated SNPs and the two independently associated classical HLA-B and HLA-A alleles, with only a modest reduction in effect size of the minor allele (AC): before adjusting β = -0.53, P=2.4×10⁻¹⁸; after adjusting β = -0.51, P=5.0×10⁻⁶. We tested rs146647111 (an intronic indel) as an eQTL for all genes within 1 Mb. The most significant association was with MICB (MHC class I polypeptide-related sequence B) (P=9.9×10⁻¹⁸), and replicated in the
independent sample (P=9.5×10^{-7}). We observed that MICB gene expression decreased in the presence of the minor, VLSP-protective AC allele. MICB encodes for a natural killer (NK) group 2 D (NKG2D) cell surface receptor ligand, which may be involved in the escape of HIV from NK cell-mediated cell death. The rs146647111–VLSP association observed across multiple ancestries is novel and independent of known variants associated with VL phenotypes. The observed effect of the rs146647111-AC allele on MICB expression suggests a biologically plausible pathway for this association: lower expression of MICB reducing capacity for HIV escape from NK cell mediated cell death.