Submitter Name: Kyle Sullivan
PI Name: Eric Johnson
Daniel Jacobson

Submitted Email: sullivanka@ornl.gov
PI email: ejohnson@rti.org
jacobsonda@ornl.gov

Multi-omic integration uncovers biological pathways underlying HIV viral load

Kyle A. Sullivan[1], Melyssa Minto[2], J. Izaak Miller[1], Bryan C. Quach[2], Caryn Willis[2], Michael R. Garvin[1], Alice Townsend[3], Peter Kruse[3], Matthew Lane[3], David Kainer[1], Mikaela Cashman[1,4], Ke Xu[5,6], Bradley E. Aouizerat[7], Dana B. Hancock[2], Daniel A. Jacobson[1], Eric O. Johnson[2, 8]

[1] Computational and Predictive Biology Group, Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN; [2] GenOmics and Translational Research Center, RTI International, Research Triangle Park, NC; [3] Bredesen Center for Interdisciplinary Research and Graduate Education, University of Tennessee-Knoxville, Knoxville, TN; [4] Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA; [5] Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA; [6] Veterans Affairs Connecticut Healthcare System, West Haven, CT, USA; [7] Translational Research Center, College of Dentistry, New York University, New York, NY, USA; [8] Fellow Program, RTI International, Research Triangle Park, NC

While antiretroviral therapy has significantly improved disease prognosis in people living with HIV, substance use can compromise adherence to ART and exacerbate inflammation, thereby accelerating disease progression directly and indirectly. Understanding biological mechanisms underlying HIV viral load is a necessary foundation for future studies exploring the mechanisms underlying how substance use exacerbates HIV disease progression. Here, we integrated multi-omic datasets and used two machine learning network biology tools (GRIN and MENTOR) to identify biological mechanisms associated with HIV viral load across 10 distinct cohorts: the Veterans Aging Cohort Study (VACS; n = 2,465), Swiss HIV Cohort (n = 198), and 7 cohorts of whole-genome sequencing data used to obtain an MHC locus fine mapping reference panel (n = 21,546). We integrated three genes from HIV set point viral load GWAS using fine mapping from WGS data (*HLA-A*, *HLA-B*, *HLA-C*), 258 differentially expressed genes from microarray data based on CD4+ T-cell viral load from HIV-infected, untreated individuals from the Swiss HIV Cohort, 143 genes based on differential DNA methylation status from the VACS cohort, and 8 genes known to affect the pharmacokinetics of anti-HIV drugs (*ABCC2*, *CYP2A6*, *CYP2B6*, *CYP2C9*, *CYP2C19*, *CYP3A4*, *SLCO1B1*, and *UGT1A1*). Using GRIN, we retained 194 multi-omic genes based on their high network interconnectivity. Using MENTOR, we then collaboratively interpreted subsets of the 194 multi-omic viral load genes. We identified the following biological processes implicated by multi-omics of HIV viral load: cell cycle checkpoint, histone modifications, mitochondrial function, oxidative stress, viral replication, and interferon signaling.