Host Genetic Regulation of HIV Set-Point Viral Load in Individuals of African Ancestry

Riley H. Tough[1,2] , David M. Tang[1], Shanelle Gingras[1,2], Paul J. McLaren[1,2]

[1]National HIV and Retrovirology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba

[2]Department of Medical Microbiology and Infectious Diseases, University of Manitoba, Canada

Purpose

To investigate how human genetic variability in highly HIV affected populations modifies viral load and disease progression. From a better understanding of the HIV host-pathogen interaction, we aim to help guide the development of host-targeted HIV therapeutics.

Acknowledgements

JC WILT INFECTIOUS DISEASES RESEARCH CENTRE

Public Health Agency of Canada
Agence de la santé publique du Canada

Conflict of Interest Disclosure: I have no conflicts of interest

Contact: Riley Tough; toughr@myumanitoba.ca

Background

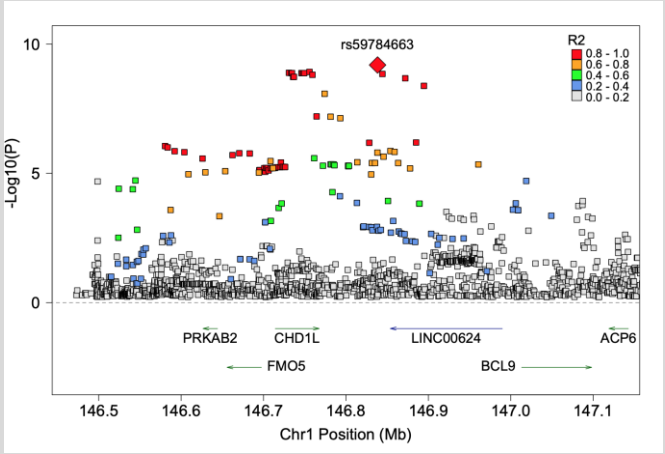


Figure 1. Enhanced view of the chromosome 1 region that is associated with a reduction of ~0.3 log10 copies/ml HIV spVL.

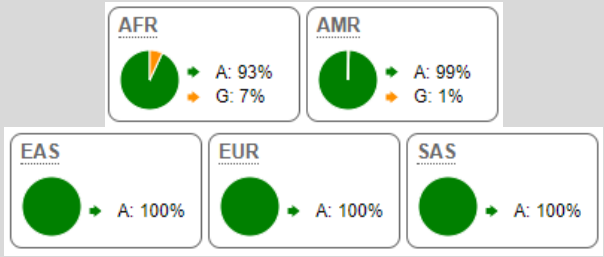


Figure 2. Population allele frequencies of rs59784663 from 1000 Genomes Phase 3 data.

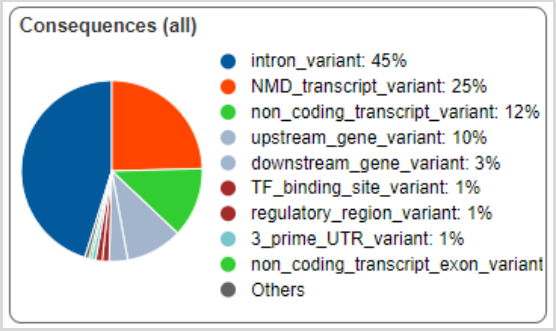


Figure 3. Functional annotation of variants in the chromosome 1 region using VEP GRCh38.

A Genome-Wide Association Study was conducted with 3,879 individuals of African Ancestry to determine if host genetic variation is associated with HIV-set point viral load (spVL).

A closer look at the chromosome 1 region shows a pattern of variants in high linkage disequilibrium, shown in orange and red, to the top associated SNP rs5978466. This pattern overlaps three coding genes: *PRKAB2, FMO5,* and *CHD1L*.

CHD1L is involved in chromatin relaxation and DNA repair[1]. *PRKAB2* is a regulatory scaffold for *AMPK,* a master regulator kinase for low-energy states[2]. *FMO5* is part of the flavin-monooxygenase family of genes that metabolise drugs, however *FMO5* appears to lack the ability to metabolise drugs[3]. All genes will be assessed for impact on HIV infection.

The top associated variant, rs59784663, is expressed at relevant levels (minor allele frequency > 1%) only in African populations (range 4-12% by geographic region) and is not present in European or Asian populations. This provides support why this region has not been detected in similar studies with individuals of European or Asian ancestry.

Functional annotation of the chromosome 1 variants show no changes to the coding sequence and a high proportion of non-coding variants. This suggests that decreased viral load may be caused by changes to gene regulation. Further work is required to determine how changes to host gene regulation impact HIV spVL.

Hypotheses

1. Individuals, carrying the variant rs59784663, will have differential gene and protein expression of *CHD1L, FMO5,* and/or *PRKAB2*.

2. Causal variants are in LD to 59784663 (R^2 ≥0.2) and will be associated with differential expression of *CHD1L, FMO5,* and/or *PRKAB2*, when used as RNA-seq selection criteria in HIV relevant cell types.

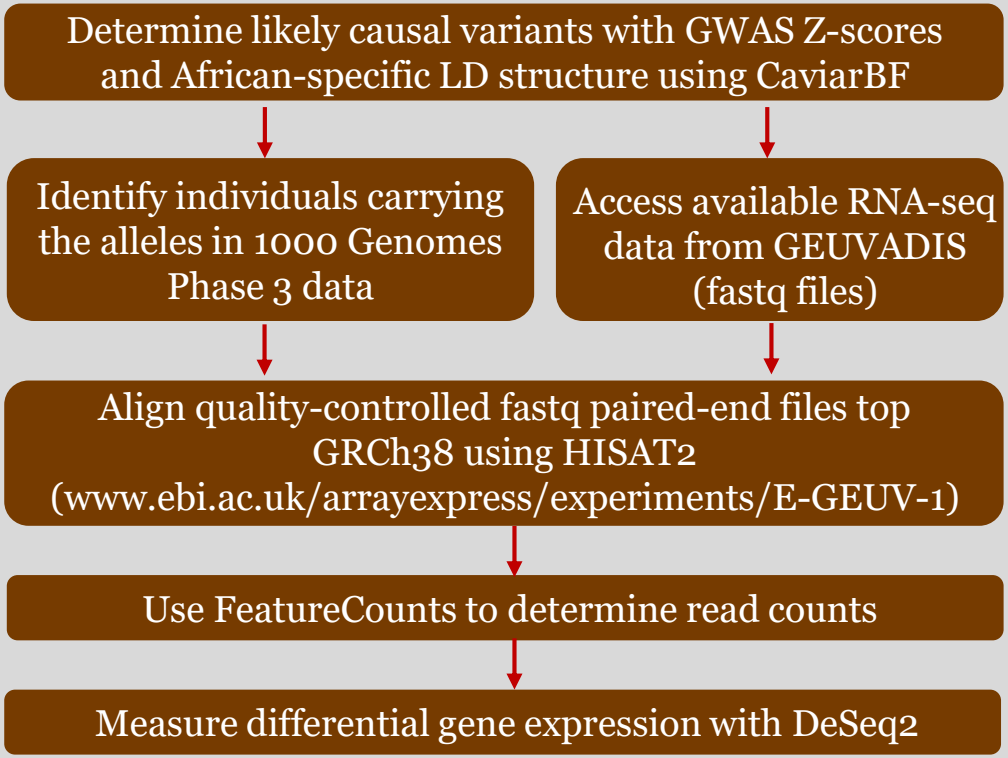
Objectives

1.Determine potentially causal variants using statistical fine-mapping and functional annotations

2. Genotype individuals of African ancestry to determine variant carriers at likely causal variants from 1000 Genomes Phase 3 data

3. Test for differential expression of CHD1L, PRKAB2, and FMO5 in EBV-transformed lymphoblastoid cells between genotype groups

Methods



Results

Table 1. Variants identified as likely causal (95% confidence set) using GWAS summary statistics and AFR-specific LD structure with a CaviarBF framework.

RSID	GWAS pval	GWAS Effect	Associated Traits*
rs59784663	6.37E-10	0.30	n/a
rs2353984	1.45E-04	-0.10	*CHD1L* abundance
rs7525622	6.36E-09	0.21	n/a
rs7520661	5.22E-08	0.20	*CHD1L* abundance
rs7520841	7.60E-05	0.11	n/a
rs6675942	5.83E-03	-0.07	n/a
rs10900350	9.80E-04	-0.08	n/a
rs7551766	9.65E-07	0.16	*CHD1L* abundance
rs66501488	9.80E-07	0.16	n/a
rs4314933	1.07E-03	-0.08	n/a
rs59987487	1.39E-04	0.13	n/a
rs2883319	1.25E-06	0.16	n/a
rs7417503	1.42E-04	0.13	n/a
rs2353975	8.93E-04	-0.08	*CHD1L* abundance
rs11239997	1.14E-03	-0.08	n/a
rs72694706	2.96E-06	0.15	n/a

*Associated traits are pulled from rVarBase

This table is a 95% confidence set, meaning that we are 95% confident that this table contains the causal variant(s).

The variants with known traits correspond to *CHD1L* abundance which suggests it may be the causal gene. However, the rVarBase database is European-centered and therefore more work is required to determine how these variants may impact African-specific populations.

How *CHD1L* would regulate HIV infection remains unclear. We decided to access the GEUVADIS RNA-seq database to determine gene expression patterns that are associated with the presence of rs59784663 because causal variants will be in linkage.

Table 2. Differentially expressed genes from Yoruban EBV-transformed lymphoblastoid cells based on rs59784663 genotype group with 7 female variant carriers and 11 female reference carriers.

GEUVADIS YRI EBV-transformed lymphoblastoid cells DESeq2 output			
HGNC	Mean Count	Log2FC	p-value
PRKAB2	401.89	-0.23	0.21
FMO5	32.30	0.10	0.65
CHD1L	1890.55	-0.10	0.40
BCL9	862.09	0.20	0.21
ACP6	190.33	-0.23	0.23
GJA5	0.07	0.00	0.97

A positive fold change represents genes upregulated in individuals who carry the rs59784663 variant allele. From Table 2, there are no genes in the chromosome 1 region associated with rs59784663 presence in Yoruban EBV-transformed lymphoblastoid cells.

The lack of association may be caused by the absence of cell-specific enhancers[4]. Cell types more relevant to HIV infection, such as CD4+ T cells or monocytes, may be required to see association signals.

Host Genetic Regulation of HIV Set-Point Viral Load in Individuals of African Ancestry

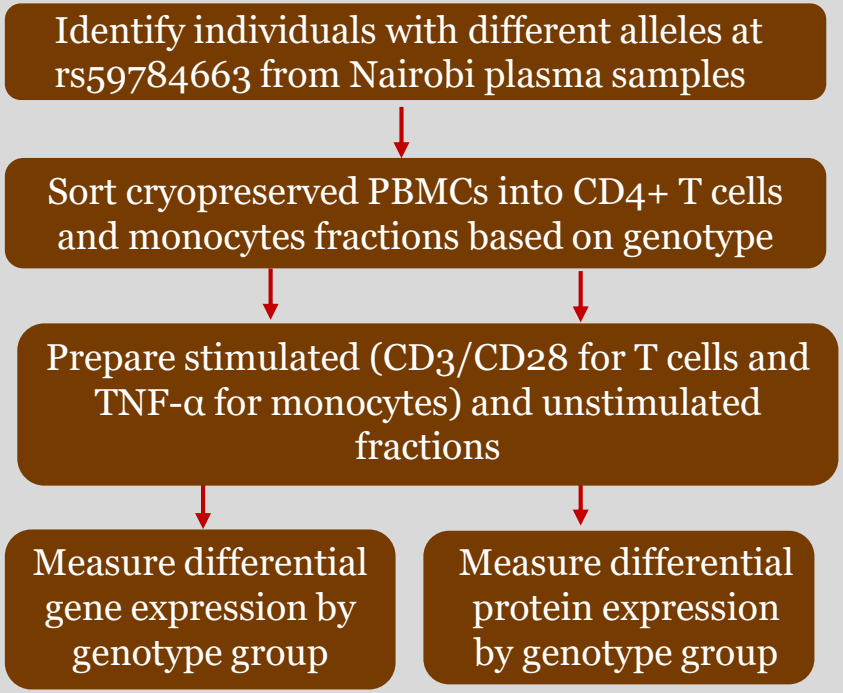
Summary

1. From fine-mapping analysis, the top associated variants are likely to impact *CHD1L* expression
2. In EBV-transformed lymphoblastoid cells, presence of rs59784663 has no significant effect on *CHD1L*, *PRKAB2*, or *FMO5* expression
3. Further work is required to determine how the presence of rs59784663 and likely causal variants impact gene expression in a more relevant cell type (primary CD4+ T cells and monocytes)
4. Further work is required to determine how these variants impact gene expression (i.e. histone modification or TF-binding domains)

References

1. Ahel et al. *Science*. 2009.
2. Thornton et al. *The Journal of Biological Chemistry*. 1998.
3. Scott et al. *Drug Metabolism and Disposition*. 2017.
4. Ernst et al. *Nature*. 2011.

Future Directions



Contact: toughr@myumanitoba.ca