TET2 genetic variation affects HIV viral load in ART-naïve persons

Daniel D. Murray
for the INSIGHT START study group and the FIRST study group

Background

- Compared to the general population, HIV-positive persons continue to be at greater risk of a variety of clinical events, even with optimum antiretroviral therapy (ART)
- Identifying factors that influence this risk is key for two reasons
    1. Understanding the underlying pathogenesis of HIV-disease
    2. Identification of clinically relevant biomarkers
- Previous studies, mostly in European populations, have identified SNPs in HLA and CCR5 that explain much of the genetic induced variation in HIV-VL[1-5]
- Other genes are clearly involved in HIV pathogenesis and variation in these genes may also influence HIV-VL

1. Fellay J, Ge D, Shianna KV, et al; NIAID Center for HIV/ AIDS Vaccine Immunology (CHAVI). Common genetic variation and the control of HIV-1 in humans. PLoS Genet 2009; 5:e1000791.
2. Pereyra F, Jia X, McLaren PJ, et al. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science 2010; 330:1551–7. 7.
3. Fellay J, Shianna KV, Ge D, et al. A whole-genome association study of major determinants for host control of HIV-1. Science 2007; 317:944–7.
4. McLaren PJ, Coulonges C, Bartha I, et al. Polymorphisms of large effect explain the majority of the host genetic contribution to variation of HIV-1 virus load. Proc Natl Acad Sci U S A 2015; 112:14658–63
5. Ekenberg, C., et al., *Single nucleotide polymorphisms in HLA alleles are associated with HIV-1 viral load in demographically diverse, ART-naive participants from the START trial.* J Infect Dis, 2019.

Our targeted approach

- Can genetic information from clinical studies compliment molecular evidence and further our understanding of HIV pathogenesis?
- Target a pathway suspected to be involved in HIV-pathogenesis and determine whether variation in this pathway affects HIV-VL

Hypothesis: TET2 is a critical regulator of HIV-replication, and that genetic variation within this pathway will impact this function

Aim: Use the START and FIRST cohorts to assess the impact of genetic variation within the TET2 pathway on HIV-VL

Why TET2?

- TET2 is a host gene involved in demethylation
- TET2 involved the regulation of endogenous retroviral elements[1,2]
- TET2 function has been linked to HTLV-1 (a retrovirus closely related to HIV) induced malignancy[3]
- In the context of HIV, one recent study has suggested that the HIV-protein VPR selectively degrades TET2, enhancing IL-6 expression and viral replication[4]
- Preliminary work in our previous GWAS[5] observed a non-genome wide significant signal in TET2 that was below the minor allele frequency (MAF) cut-off used in the final manuscript, but encouraged us to explore this region further

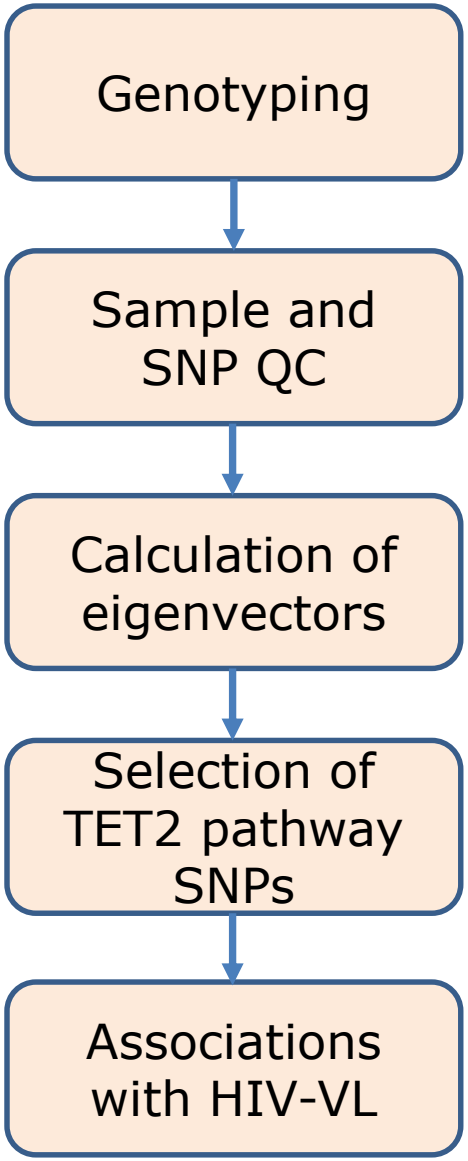
1. Guallar, D., et al., RNA-dependent chromatin targeting of TET2 for endogenous retrovirus control in pluripotent stem cells. Nat Genet, 2018. 50(3): p. 443-451.
2. Deniz, O., et al., SETDB1 prevents TET2-dependent activation of IAP retroelements in naive embryonic stem cells. Genome Biol, 2018. 19(1): p. 6.
3. Lv, L., et al., Vpr Targets TET2 for Degradation by CRL4(VprBP) E3 Ligase to Sustain IL-6 Expression and Enhance HIV-1 Replication. Mol Cell, 2018. 70(5): p. 961-970 e5.
4. Yeh, C.H., et al., Mutation of epigenetic regulators TET2 and MLL3 in patients with HTLV-I-induced acute adult T-cell leukemia. Mol Cancer, 2016. 15: p. 15.
5. Ekenberg, C., et al., *Single nucleotide polymorphisms in HLA alleles are associated with HIV-1 viral load in demographically diverse, ART-naive participants from the START trial.* J Infect Dis, 2019.

The cohorts - START and FIRST

Two ART naïve cohorts from the INSIGHT network (http://insight.ccbr.umn.edu/)

	START	FIRST
Participants (genetic consent) (n =)	2546	544
Age (years) Median (IQR) or Percent	36 (29, 45)	38 (32, 44)
Female (%)	20	20
Race (self reported)		
Asian (%)	1	<1
Black (%)	23	57
White/other (%)	76	43
CD4+ count (cells/mm^3) Median (IQR)	651 (585, 759)	220 (43, 345)
HIV RNA level (log10 copies/mL) Median (IQR)	4.17 (3.54, 4.66)	5.09 (4.53, 5.54)
Region of Residence		
U.S. (%)	18	100
Europe/Australia/Israel (%)	49	0
South America/Mexico (%)	20	0
Asia[2] (%)	0	0

Methods - overview



Custom content Affymetrix SNP CHIP – enriched for SNPs/genes involved in the immune response (including TET2)

No imputation was performed

Used in associations to control for population structure

All SNPs (n = 888) across the TET2 pathway (IDH1 and IDH2 are regulators of TET2) with MAF > 1%

Gene and SNP level association using an additive model

Association with HIV-VL at study entry

- After QC and MAF filtering we analysed 292 and 345 SNPs for START and FIRST, respectively
- Associations with HIV-VL at study entry were performed
  1. At the gene level – using SKAT-O[1] + gender and first four eigenvectors
  2. At the SNP level – using linear regression + gender and first four eigenvectors
- Benjamin-Hochberg procedure was used to control the false discovery rate to 5% (q-value < 0.05)
- Associations were performed independently in each cohort

1. Ionita-Laza, I.*, Lee, S.*, Makarov, V., Buxbaum, J. Lin, X. (2013) Sequence kernel association tests for the combined effect of rare and common variants. *American Journal of Human Genetics, 92, 841–853*

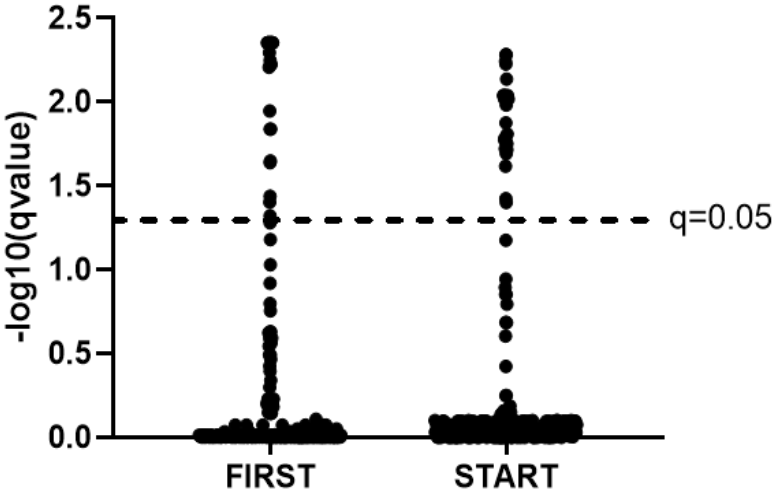
Results - Associations with HIV-VL

- Gene level associations with HIV-VL in both START and FIRST
- SNP level associations in both START and FIRST
- 36 SNPs were associated with HIV-VL ($q < 0.05$) in one of either START or FIRST
- 15 of these SNPs were associated ($q < 0.05$) with HIV-VL in both cohorts
- 35/36 SNPs associated with HIV-VL were in TET2
- No gene level associations were observed in IDH1 or IDH2

Gene level associations

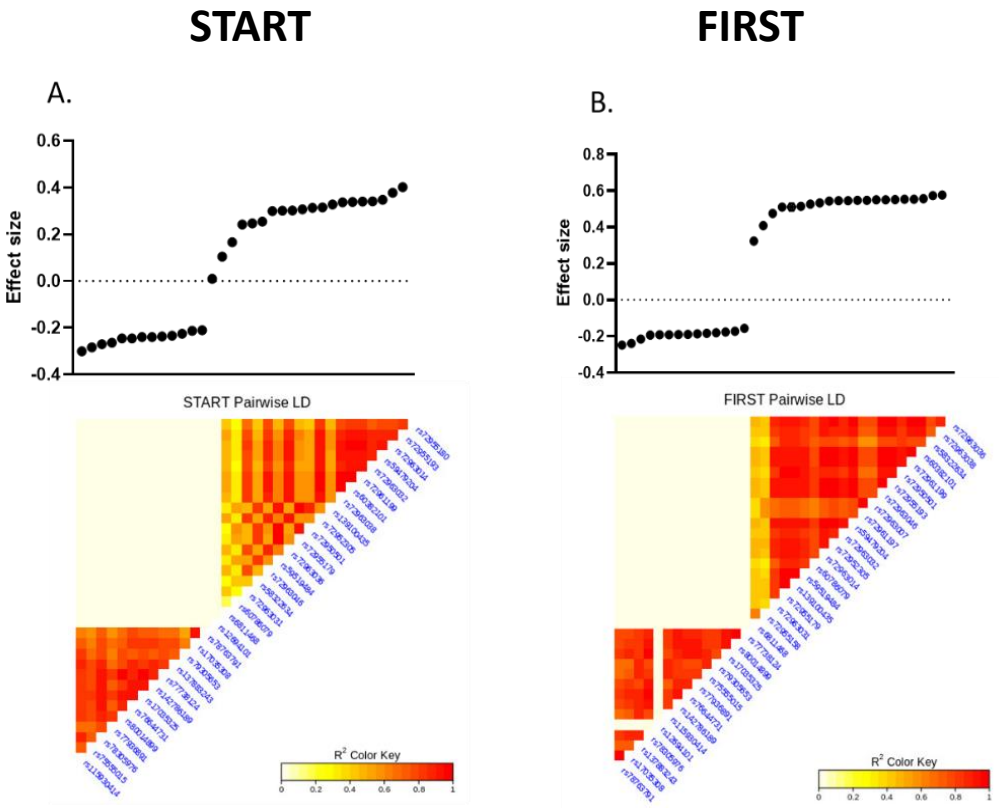
Study	SKAT-O p-value
START	0.000136
FIRST	0.000546

SNP level associations



Linkage Disequilibrium (LD) of SNPs associated with HIV-VL

- Two (maybe 3) groups of SNPs in LD
- One group is associated with higher HIV-VL
- One group is associated with lower HIV-VL
- All TET2 SNPs that associated with higher HIV-VL in START associated with higher HIV-VL in FIRST
- TET2 SNPs that associated with a lower HIV-VL in START associated with a lower HIV-VL in FIRST
- Strong LD structure makes identifying a causal SNP difficult

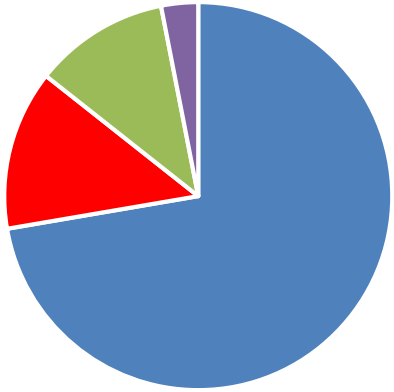


SNPs associated with HIV-VL are predominantly present in persons of Black race

START
(23% Black)

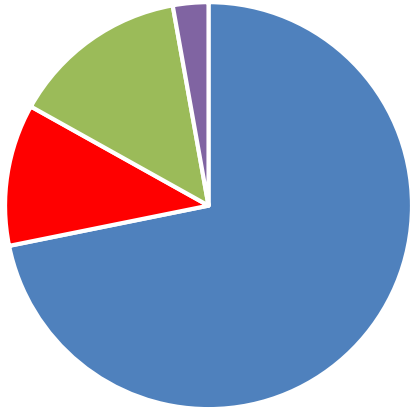
↓ HIV-VL

rs115930414 (n=195)



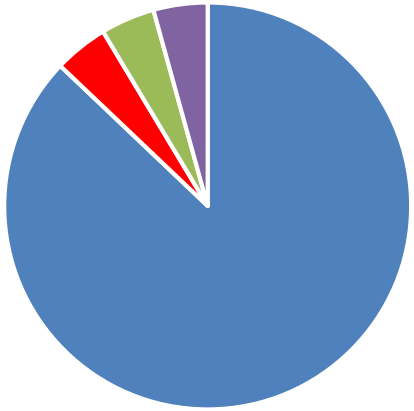
↑ HIV-VL

rs72963036 (n=71)

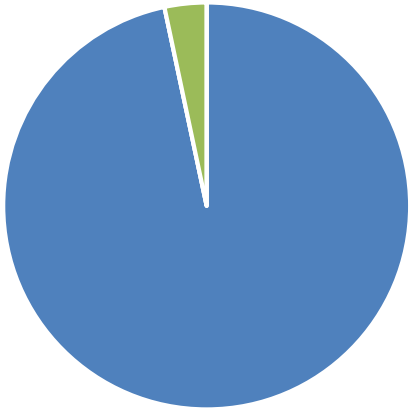


FIRST
(57% Black)

rs115930414 (n=93)



rs72963036 (n=30)



Black
White
Latin
Other

Literature associations

- Most of the SNPs were part of the enrichment of the SNP CHIP and have not been reported in the literature previously
- One SNP, rs72963007, has been reported in the literature
  - Associated with an increased risk of adult T cell leukaemia caused by HTLV-1[1]
  - This association was in persons of African descent
  - This SNP was present in 13% of adult T cell leukaemia (ATL) patients compared to 5% of an ethnically matched control population

1. Marcais, A., et al., *Adult T cell leukemia aggressivenness correlates with loss of both 5-hydroxymethylcytosine and TET2 expression.* Oncotarget, 2017. **8**(32): p. 52256-52268.

Strengths and limitations of the study

Strengths

- Validation across independent cohorts
- Enrichment of TET2 in the INSIGHT CHIP
- Diversity of the cohort – population specific signals

Limitation

- Diversity of the cohort – population structure as a confounder?
- The mechanism of action is unclear

No genetic association model is perfect and should be viewed more as a screening tool than a final result

We need additional confirmation/validation/accumulation of supporting evidence

Conclusion

- Gene and SNP level associations indicate genetic variation within the TET2 gene affects HIV-VL
- These data supports previous molecular evidence that TET2 is involved in HIV-replication
- Further work is required to validate and identify the mechanism behind this change in TET2 function
- Further work is required to identify the role of TET2 in HIV-replication

Acknowledgements

Study participants and staff involved in the START and FIRST studies

Centre of Excellence for Health, Immunity and Infections (CHIP), Copenhagen, Denmark

- Christina Ekenberg
- Adrian Zucco
- Cameron MacPherson
- Joanne Reekie
- Marie Helleberg
- Alvaro Borges
- Jens Lundgren

Rigshospitalet, Copenhagen, Denmark

- Jan Gerstoft

University of Minnesota, USA

- Birgit Grund

Kirby Institute, Australia

- Mark Polizzotto

Tulane University Medical Center, USA

- Dahlene Fusco

APHP-Hôpital Saint Louis, Paris, France

- Julien Gras

Supported by the National Institute of Allergy and Infectious Diseases (United States), Agence Nationale de Recherches sur le SIDA et les Hépatites Virales (France), National Health and Medical Research Council (Australia), National Research Foundation (Denmark), Bundes ministerium für Bildung und Forschung (Germany), European AIDS Treatment Network, Medical Research Council (United Kingdom), National Institute for Health Research, National Health Service (United Kingdom), and the University of Minnesota. Antiretroviral drugs were donated to the central drug repository by AbbVie, Bristol-Myers Squibb, Gilead Sciences, GlaxoSmithKline/ViiV Healthcare, Janssen Scientific Affairs, and Merck.

INSIGHT Array Content – supporting slides

Modules from UK Biobank Array

Module Name	# Markers
Alzheimer's Disease	803
ApoE	1,147
Autoimmune/Inflammatory	258
Blood Phenotypes	2,545
Cancer common variants	343
Cardiometabolic	377
HLA	13,519
KIR	1616
Lung function phenotypes	8,645
Common mitochondrial DNA variants	180
Neurological disease	19,791
Pharmacogenetics/ADME	2,856

Module Name	# Markers
Y chromosome markers	807
Rare variants in cancer predisposition genes	6,543
Rare variants in cardiac predisposition genes	1,710
Rare, possibly disease causing, mutations	13,729
eQTL	17,115
Fingerprint	262
NHGRI GWAS catalog	8,136
Protein truncating variants	30,581
Other rare coding variants	80,581
Genome-wide coverage for common variants	348,569
Genome-wide coverage for rare variants	37,000

Custom Content — 93,000 SNPs

Ddimer	Hematopoiesis
Bone Mineral Density	Coronary Heart Disease
COPD	Pharmacogenetics/ADME
Immune Function/Response	Others

725,000 unique markers represented on the array
Including an enrichment of TET2 as part of haematopoiesis