## Title

Host variation in type I interferon signaling genes (MX1), C-C chemokine receptor type 5 gene, and major histocompatibility complex class I alleles in treated HIV+ noncontrollers predict viral reservoir size.



## Journal

AIDS (London, England), 37(3)

## ISSN

0269-9370

## Authors

Siegel, David A
Thanh, Cassandra
Wan, Eunice
et al.

## Publication Date

2023-03-01

## DOI

10.1097/qad.0000000000003428

Peer reviewed

OPEN

# Host variation in type I interferon signaling genes (*MX1*), C–C chemokine receptor type 5 gene, and major histocompatibility complex class I alleles in treated HIV+ noncontrollers predict viral reservoir size

David A. Siegel[a], Cassandra Thanh[b], Eunice Wan[c], Rebecca Hoh[a], Kristen Hobbs[b], Tony Pan[b], Erica A. Gibson[b], Deanna L. Kroetz[d], Jeffrey Martin[e], Frederick Hecht[a], Christopher Pilcher[a], Maureen Martin[f], Mary Carrington[f,g], Satish Pillai[h], Michael P. Busch[h], Mars Stone[h], Claire N. Levy[i], Meei-Li Huang[j,k], Pavitra Roychoudhury[j,k], Florian Hladik[i,k], Keith R. Jerome[j,k], Hans-Peter Kiem[j,k], Timothy J. Henrich[d], Steven G. Deeks[a] and Sulggi A. Lee[a]

**Objective:** Prior genomewide association studies have identified variation in major histocompatibility complex (MHC) class I alleles and C–C chemokine receptor type 5 gene (*CCR5Δ32*) as genetic predictors of viral control, especially in 'elite' controllers, individuals who remain virally suppressed in the absence of therapy.

**Design:** Cross-sectional genomewide association study.

**Methods:** We analyzed custom whole exome sequencing and direct human leukocyte antigen (HLA) typing from 202 antiretroviral therapy (ART)-suppressed HIV+ noncontrollers in relation to four measures of the peripheral $CD4^+$ T-cell reservoir: HIV intact DNA, total (t)DNA, unspliced (us)RNA, and RNA/DNA. Linear mixed models were adjusted for potential covariates including age, sex, nadir $CD4^+$ T-cell count, pre-ART HIV RNA, timing of ART initiation, and duration of ART suppression.

**Results:** Previously reported 'protective' host genetic mutations related to viral setpoint (e.g. among elite controllers) were found to predict smaller HIV reservoir size. The HLA 'protective' B*57:01 was associated with significantly lower HIV usRNA ($q = 3.3 \times 10^{-3}$), and among the largest subgroup, European ancestry individuals, the *CCR5Δ32*

[a]Department of Medicine, Division of HIV, Infectious Diseases & Global Medicine, [b]Department of Medicine, Division of Experimental Medicine, [c]Institute of Human Genetics, [d]Department of Bioengineering and Therapeutic Sciences, [e]Department of Biostatistics & Epidemiology, University of California San Francisco, California, [f]Basic Science Program, Frederick National Laboratory for Cancer Research, National Cancer Institute, Frederick, and Laboratory of Integrative Cancer Immunology, Center for Cancer Research, National Cancer Institute, Bethesda, Maryland, [g]Ragon Institute of MGH, MIT and Harvard, Cambridge, Massachusetts, [h]Vitalant Blood Bank, San Francisco, California, [i]Department of Obstetrics and Gynecology, [j]Department of Laboratory Medicine and Pathology, University of Washington, and [k]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA.

Correspondence to Sulggi A. Lee, University of California San Francisco, San Francisco, CA, USA.

E-mail: sulggi.lee@ucsf.edu

Received: 23 June 2022; revised: 26 October 2022; accepted: 28 October 2022.

DOI:10.1097/QAD.0000000000003428

deletion was associated with smaller HIV tDNA ($P = 4.3 \times 10^{-3}$) and usRNA ($P = 8.7 \times 10^{-3}$). In addition, genomewide analysis identified several single nucleotide polymorphisms in *MX1* (an interferon stimulated gene) that were significantly associated with HIV tDNA ($q = 0.02$), and the direction of these associations paralleled *MX1* gene eQTL expression.

**Conclusions:** We observed a significant association between previously reported 'protective' MHC class I alleles and *CCR5Δ32* with the HIV reservoir size in noncontrollers. We also found a novel association between *MX1* and HIV total DNA (in addition to other interferon signaling relevant genes, *PPP1CB*, *DDX3X*). These findings warrant further investigation in future validation studies.



*AIDS* 2023, **37**:477–488

**Keywords: C–C chemokine receptor type 5 gene, HIV reservoir, host genetics, major histocompatibility complex class I, type I interferon**

## Introduction

Although antiretroviral therapy (ART) prolongs life, it does not fully restore health. Persistent HIV may contribute to ongoing inflammation, immune activation, and increased morbidity and mortality [1–4]. Identifying host genetic predictors of HIV persistence in ART-suppressed individuals may shed light on novel targets to reduce residual virus and HIV-associated immune dysfunction.

Most prior host genetic HIV studies have focused on identifying variants associated with viral setpoint, for example, among 'elite controllers', HIV+ individuals able to maintain viral suppression in the absence of therapy [5–14]. These studies identified several key single nucleotide polymorphisms (SNPs) in the human major histocompatibility complex (MHC), or human leukocyte antigen (HLA)-B and -C regions as well as deletions in the C–C chemokine receptor type 5 gene (*CCR5Δ32*) [15–18] and a SNP in the HLA complex 5 (*HCP5*) gene [6]. However, whether *residual* viral control during *treated* HIV disease – that is, 'the HIV reservoir' – is influenced by the same genetic variants is unknown. We performed custom whole exome sequencing among HIV noncontrollers in relation to four measures of the peripheral $CD4^+$ T-cell HIV reservoir: cell-associated 'intact' DNA [19], total DNA, unspliced RNA, and RNA/DNA (Figure S1, Supplemental Digital Content, http://links.lww.com/QAD/C713). We found that previously reported 'protective' HLA-B*57:01 [6,13] and *CCR5Δ32* [15–17] genetic variation were associated with smaller HIV reservoir size. Genomewide analyses demonstrated several novel associations with SNPs in interferon signaling-associated genes (*MX1*, *PPP1CB*, *DDX3X*) and total HIV DNA reservoir size, and intact DNA was significantly associated with gene sets representing interferon signaling pathways in a subgroup of participants (Europeans).

## Methods

### Study participants

HIV+ noncontrollers who initiated ART during chronic (>2 years) or early (<6 months) HIV infection were sampled from the UCSF SCOPE and Options cohorts (Table S1, Supplemental Digital Content, http://links.lww.com/QAD/C713). Inclusion criteria were laboratory-confirmed HIV-1 infection, availability of cryopreserved peripheral blood mononuclear cells (PBMCs), and plasma HIV RNA <40 copies/ml for ≥24 months at the time of biospecimen collection. HIV 'controllers' [20–22] were excluded (individuals with a undetectable viral load in the absence of therapy for ≥1 year). The estimated date of detected infection (EDDI) was calculated for each study participant to determine recency of infection in relation to ART initiation using the Infection Dating Tool (https://tools.incidence-estimation.org/idt/) [23]. Additional exclusion criteria were potential factors that might influence HIV reservoir quantification: recent hospitalization, infection requiring antibiotics, vaccination, or exposure to immunomodulatory drugs <6 months prior to sampling timepoint. The research was approved by the UCSF Committee on Human Research (CHR), and all participants provided written informed consent.

### Custom whole exome host DNA sequencing

Genomic DNA was extracted (AllPrep Universal Kit; Qiagen, Hilden, Germany) from negatively selected $CD4^+$ T cells from cryopreserved PBMCs (StemCell, Vancouver, Canada). Targeted exome capture was performed with custom addition of 50 Mb regulatory regions (Roche NimbleGen; Wilmington, Massachusetts, USA), sequencing libraries were generated and then run on the Illumina HiSeq 2000 system (Illumina, San Diego, California, USA). The custom regions included 50 kb upstream and 50 kb downstream of 442 candidate

genes related to cell cycle regulation, HIV host restriction factors, and HIV-host integration, which were selected based on Gene Ontology (GO) Consortium experimental evidence codes (EXP, IDA, IPI, IMP, IGI, IEP) (Table S2, Supplemental Digital Content, http://links.lww.com/QAD/C713).

## Human leukocyte antigen typing

Direct HLA typing was performed from extracted DNA following the PCR-SSOP (sequence-specific oligonucleotide probing) typing and PCR-SBT (sequence-based typing) protocols recommended by the 13th International Histocompatibility Workshop [24,25]. Locus-specific primers were used to amplify a total of 25 polymorphic exons of HLA-A & B (exons 1–4), C (exons 1–5), E (exon 3), DPA1 (exon 2), DPB1 (exons 2–4), DQA1 (exon 1–3), DQB1 (exons 2–3), DRB1 (exons 2–3), and DRB3, 4, 5 (exon 2) genes with Fluidigm Access Array (Fluidigm, Singapore) and sequenced on an Illumina MiSeq sequencer (Illumina). HLA alleles and genotypes are called using the Omixon HLA Explore (version 2.0.0) software (Omixon, Budapest, Hungary).

## HIV reservoir quantification from peripheral CD4$^+$ T cells

There is currently no 'gold standard' for measuring the HIV reservoir. Moreover, the HIV reservoir largely consists of 'defective' virus that harbors mutations prohibiting the production of infectious virus [26,27]. Here, we extracted DNA and RNA from negatively selected CD4$^+$ T cells from cryopreserved PBMCs, using the AllPrep Universal Kit (Qiagen). The frequency of 'intact' (i.e. potentially replication-competent) HIV DNA was quantified using a multiplex ddPCR assay [19]. The number of HIV positive droplets were counted using three targets per assay, and then normalized (using RPP30 and TRD genes for all cells and all non-T cells, respectively) to obtain HIV copy numbers/$10^6$ CD4$^+$ T cells. Finally, the estimated number of intact HIV genomes was corrected for shearing using a DNA shearing index (DSI). HIV total (defective + intact) DNA and unspliced (full-length) RNA was also measured for all participants. These two measures were quantified using an HIV-1 LTR-specific qPCR TaqMan assay [28]. For each measure, HIV copies were quantified in triplicate reaction wells using a seven-point standard curve (1–10 000 copies/s).

## Data processing and quality control

The bcbio bioinformatics pipeline [29] was used to perform DNA alignment, which included the Burroughs-Welcome Aligner (BWA) tool [30] and the GenomeAnalysisToolkit (GATK) HaplotypeCaller joint variant calling method [31]. Reads were initially mapped to reference genome b37, then transposed to human genome assembly GRCh38 using Picard tools [32]. SNPs and insertions or deletions (indels) were then filtered by variant quality score recalibration (VQSR) using GATK [33]. The whole genome data analysis toolset, PLINK [34], was then used to validate the chromosomal sex of each individual, filter out individuals with excessive heterozygosity, and SNPS violating Hardy-Weinberg equilibrium (HWE) at a *P*-value threshold of $1 \times 10^{-8}$. The VCFtools suite of functions were then used to summarize data, run calculations, convert data, and filter out data, and convert data, and filter out relevant SNPs [35].

The GENESIS analysis pipeline [36] was used to analyze the relatedness and ancestries of the individuals in the study. All individuals were determined to be unrelated (kinship estimates <0.05) aside from one pair of siblings, so one sibling was randomly removed from the study. The remaining 199 unrelated individuals had diverse and mixed ancestries (Fig. 1). We accounted for population stratification in the total population by including a genetic effects term with a genetic relatedness matrix (GRM), by

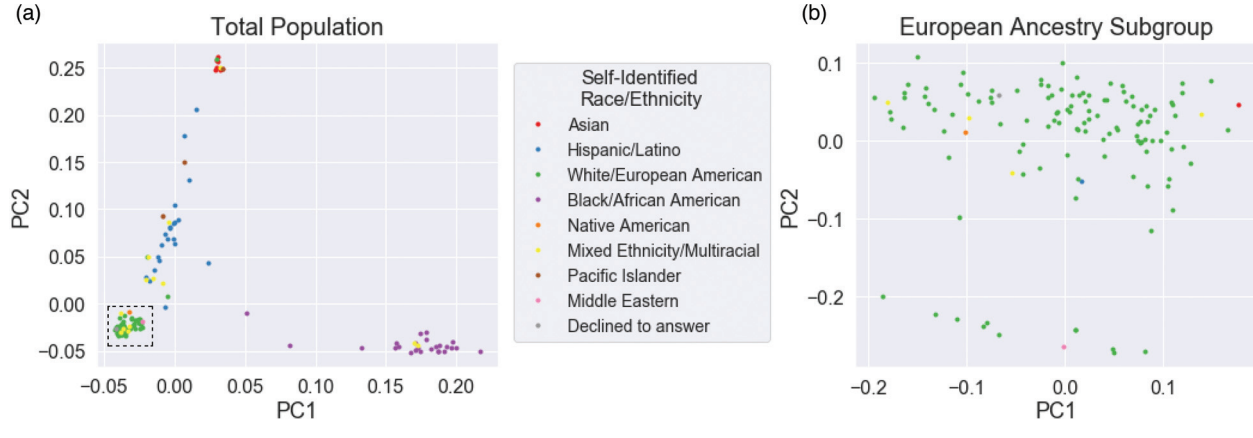


**Fig. 1. Genetic principal component analysis (PCA) plots of genotypes from the total study population (with legend with self-identified race/ethnicity) (a) and from the European ancestry subgroup European ancestry (b).** Recalculated European ancestry PCA plot is shown in panel b, from lower left dashed box in panel a.

including the first five PCs as covariates in the multivariate models, and by performing sensitivity analyses among the largest subgroup, Europeans.

### Single nucleotide polymorphism common variant analyses

Individual SNP associations were calculated with GENESIS 'assocTestSingle'. For HIV tDNA, usRNA, and intact DNA, each outcome variable was analyzed as $\log_{10}$ (copies/$10^6$ $CD4^+$ T cells + offset); RNA/DNA was analyzed as $\log_{10}$ ((RNA copies/$10^6$ $CD4^+$ T cells + offset)/(DNA copies/$10^6$ $CD4^+$ T cells + offset)). The offsets counts were given by the smallest nonzero measured values, to avoid divergences in the logarithm. Final covariates in multivariate models were sex, timing of ART initiation (Figure S2, Supplemental Digital Content, http://links.lww.com/QAD/C713), nadir $CD4^+$ T-cell count (Figure S3, Supplemental Digital Content, http://links.lww.com/QAD/C713), and the first five PCs. Pre-ART viral load (Figure S4, Supplemental Digital Content, http://links.lww.com/QAD/C713) and duration of ART suppression (Figure S5, Supplemental Digital Content, http://links.lww.com/QAD/C713) were not associated with HIV reservoir size nor improved the fit of the final models. A Gaussian link function was used, and a GRM was included with results filtered for SNPs with MAF $\geq$5%. SNP annotations were then obtained using Annovar [37].

### Gene-based rare variant analyses

Gene level multi-SNP associations were calculated with the GENESIS software package 'assocTestAggregate' function implementing the variant Set Mixed Model Association Test (SMMAT) [38] for alleles with MAF<5% with weights following the beta distribution parameters of $a_1 = 1$ and $a_2 = 25$ [39]. The same covariates, GRM, and regression family were used as for the individual SNP associations. Outcomes were quantile-normalized to follow a normal distribution. Gene regions were defined according to UCSC hg38 assembly [40].

Gene set enrichment analyses (GSEA) were performed using the Molecular Signatures Database (MSigDB) [41,42]. For all gene set analyses, introns and flanking regions of $\pm$50 kb were included in the SMMAT *P*-value calculations for each gene to account for potential regulatory regions and SNPs with smaller effects. GSEAPreranked was run with default parameters on the SMMAT gene-level $-\log_{10}(P)$.

### Human leukocyte antigen analysis

Multivariate regression models were fit using the python statsmodels OLS function [43] with covariates for sex, timing of ART initiation, nadir $CD4^+$ T-cell count, and three genetic PCs.

## Results

### Study population

A total of 202 HIV-infected ART-suppressed individuals from the UCSF SCOPE and Options HIV+ cohorts were included in the study with baseline characteristics consistent with our San Francisco-based HIV patient population (Table S1, Supplemental Digital Content, http://links.lww.com/QAD/C713). The Infection Data Tool [23] was used to calculate the estimated date of HIV infection for 147 participants with available clinical HIV test results. For the remaining 55 participants with self-report data only on date of HIV acquisition, we mean-imputed values assuming ART initiation started 2+ years from infection. This estimation is supported by prior data demonstrating that the HIV reservoir size is relatively stable after 2 years of infection [44–48] and the fact that all 55 participants had initiated ART during an era when guidelines were not based on initiating ART immediately at the time of HIV infection [49]. Sensitivity analyses excluding these individuals did not alter the overall study results.

### Earlier antiretroviral therapy initiation and lower nadir $CD4^+$ T-cell count were associated with smaller HIV reservoir, and HIV reservoir measures were correlated with each other

Consistent with prior work [27,50,51], earlier ART initiation was associated with significantly smaller HIV reservoirs (tDNA, usRNA, intact DNA) (Figure S2, Supplemental Digital Content, http://links.lww.com/QAD/C713), whereas lower nadir $CD4^+$ T-cell count was associated with larger HIV reservoirs (tDNA, usRNA, intact DNA, RNA/DNA) (Figure S3, Supplemental Digital Content, http://links.lww.com/QAD/C713). Pre-ART viral load (Figure S4, Supplemental Digital Content, http://links.lww.com/QAD/C713) and duration of ART suppression (Figure S5, Supplemental Digital Content, http://links.lww.com/QAD/C713) were not associated with HIV reservoir size. Although usRNA was correlated with both tDNA intact DNA (Figure S6a, b, Supplemental Digital Content, http://links.lww.com/QAD/C713), tDNA (by qPCR) was not associated with intact DNA (by ddPCR) (Figure S6c, http://links.lww.com/QAD/C713).

### HLA 'protective' B*57:01 and 'risk' C*07 alleles were associated with smaller and larger HIV reservoir sizes, respectively

Using a Benjamini–Hochberg [52] false discovery rate (FDR) adjusted $q$ <0.05, we examined 'protective' (B*57:01, B*27:05, B*14, C*08:02, B*52, and A*25) and 'risk' (B*35 and C*07) alleles, as previously reported for viral control among untreated HIV+ controllers vs. noncontrollers [13]. Among our treated HIV+ non-controllers, we observed a 'protective' association with HLA-B*57:01 and usRNA ($\beta = -1.5$, $q = 3.3 \times 10^{-3}$), with a similar trend for tDNA ($\beta = -1.6$, $q = 0.13$).

Similarly, previously reported HLA-C*07 'risk' allele also demonstrated a 'risk' trend (larger reservoir size) in our European subgroup (tDNA: $\beta = 0.76$, $q = 0.072$; usRNA: $\beta = 0.41$, $q = 0.10$). Further analyses employing a composite HLA variable did not identify statistically significant associations (Tables S3–S6, Supplemental Digital Content, http://links.lww.com/QAD/C713).

### *CCR5Δ32* was associated with smaller HIV reservoir size

Deletions in the C–C chemokine receptor type 5 gene (*CCR5Δ32*) have previously been associated with HIV control in the absence of therapy [15–17]. Among individuals of European ancestry (where *CCR5Δ32* is more commonly observed), *CCR5Δ32* was associated with smaller HIV reservoirs (tDNA: $\beta = -1.3$, $P = 4.3 \times 10^{-3}$; usRNA: $\beta = -0.78$, $P = 8.7 \times 10^{-3}$), with a similar trend in the total population (tDNA: $\beta = -0.86$, $P = 0.045$; usRNA: $\beta = -0.41$, $P = 0.12$), In addition, the previously reported long noncoding RNA variant which regulates differential CCR5 expression (rs1015164) [18], was significantly associated with smaller HIV reservoir size in Europeans (usRNA: $\beta = -0.39$, $P = 0.027$), which also reached near-statistical significance in the total population (usRNA: $\beta = -0.30$, $P = 0.051$),

### SNP-based analysis identified several SNPs associated with HIV tDNA, including *MX1*, for which HIV reservoir associations paralleled predicted *MX1* gene expression

A total of 1 279 156 variants from 23 733 genes were included from 199 participants for whom sequencing data passed quality control (Figure S7, Supplemental Digital Content, http://links.lww.com/QAD/C713). Final models demonstrated lambda genomic inflation factor [53] values close to 1, reflecting adequate adjustment for possible population stratification bias (Fig. 2).

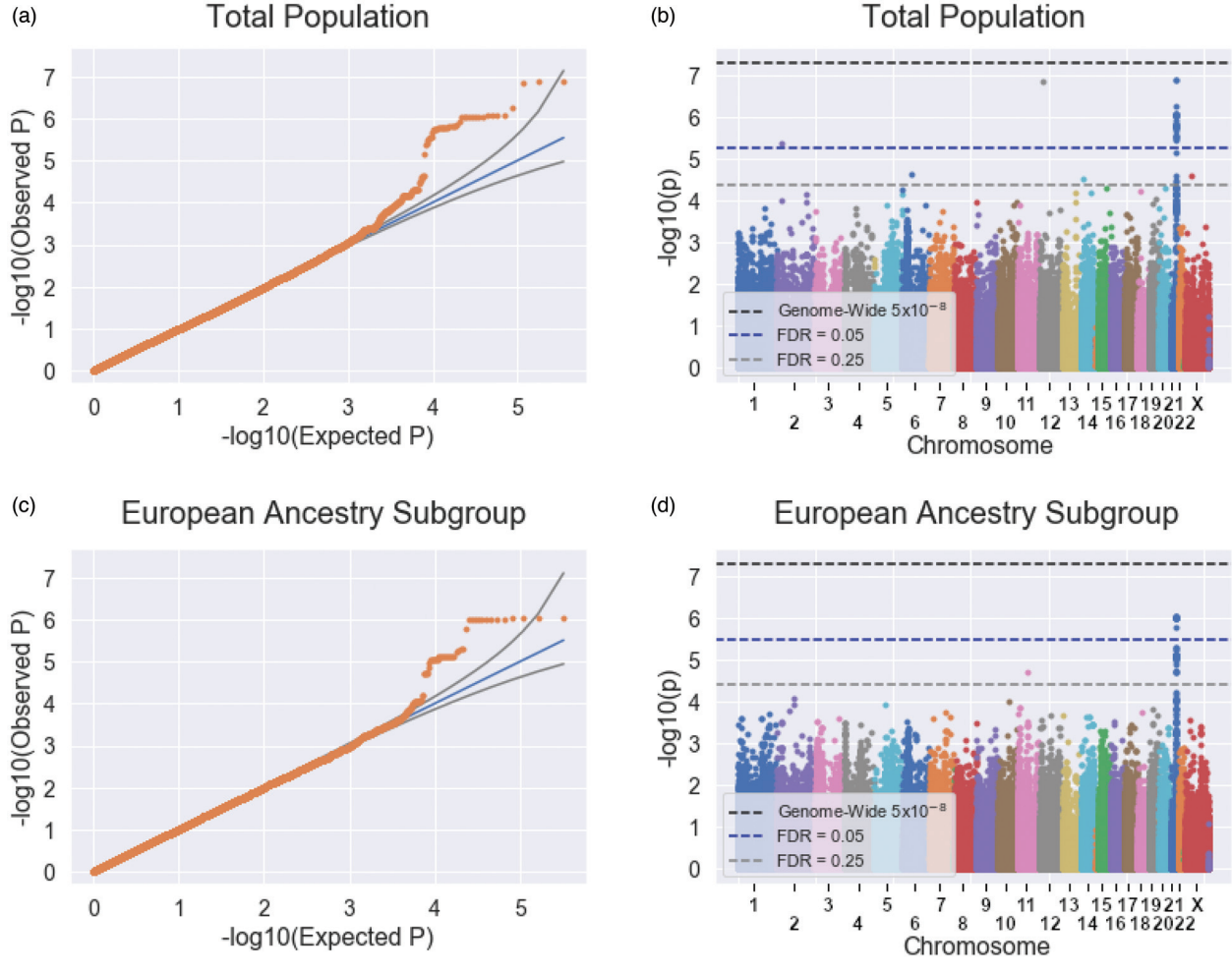


**Fig. 2. Quantile–quantile (QQ) plots (a, c) and Manhattan plots (b, d) of the total study population (a, b) and of European ancestry the European ancestry subgroup (c, d).** QQ plots: the blue line represents the expected $-\log_{10}$ *P*-values while the black lines denote the expected error bars. Manhattan plots: the horizontal black line delineates a traditional conservative genome-wide significance of *P*-value of $5 \times 10^{-8}$, while less conservative Benjamini–Hochberg false discovery rate (FDR) statistical significance of $q = 0.05$ is shown as the horizontal blue line ($q = 0.25$ is shown in grey).

The strongest genomewide associations were observed with HIV tDNA (Tables 1 and S7, Supplemental Digital Content, http://links.lww.com/QAD/C713). In particular, 44 SNPs in linkage disequilibrium (LD) in the interferon-inducible myxovirus resistance 1 (*MX1* or *MXA*) gene [54,55] were significantly associated with tDNA (all $q < 0.03$). *MX1* is a paralog of *MX2* (*MXB*), which encodes for a potent HIV-1 host restriction factor [56–58]. We then compared the directionality of the SNP hits with previously reported whole blood eQTL data at these loci [59–61] and found that the *MX1* SNPs that were associated with larger HIV tDNA were in expression quantitative trait loci (eQTL) regions predicting increased *MX1* expression while *MX1* SNPs associated with smaller HIV tDNA were predicted to decrease *MX1* expression (Table S7, Supplemental Digital Content, http://links.lww.com/QAD/C713). We also observed two additional SNPs associated with HIV tDNA: *PPP1CB* (protein phosphatase 1 catalytic subunit beta which reduces MX2's anti-HIV potency [58] and regulates HIV-1 transcription [62], $q = 0.03$) and *IRAG2* (inositol 1,4,5-triphosphate receptor associated 2, which regulates delivery of peptides to major MHC class I molecules [63] and was differentially expressed during ART initiation and interruption in HIV+ individuals [64,65], $q = 0.03$) (Table 1). HIV tDNA also showed trends with SNPs in *DDX3X* (DEAD-box helicase 3 X-linked, regulates production of type I interferons [66] and HIV-1 replication [67], $q = 0.17$) and *AKAP6* (A-Kinase Anchoring Protein 6, which binds to protein kinase A, a signaling pathway associated with HIV latency reversal and T-cell proliferation [68,69], $q = 0.20$). Among Europeans, HIV tDNA also showed a trend with a SNP in *OSBP* (oxysterol-binding protein, required for the replication of viruses such as hepatitis C and Zika [70] and has been associated with HIV-1 infection of monocyte-derived macrophages from highly-exposed seronegative individuals [71], $q = 0.14$).

For the other HIV reservoir measures, no SNPs met statistical significance in association with HIV intact DNA or RNA/DNA ratio. For HIV usRNA, a SNP <30 kilobases upstream of *BST2* which encodes for the host restriction factor tetherin [72] demonstrated a trend with usRNA ($q = 0.21$; Table 1); this SNP is intronic in *PLAVP* (protein regulating lymphocyte migration into lymph nodes [73]).

### Gene-based analyses identified several interferon signaling pathway genes that were associated with intact HIV DNA

We then performed multi-SNP gene set enrichment analyses (GSEA), aggregating SNPs by gene region to identify genes associated with HIV reservoir size. HIV intact DNA was statistically significantly associated with several gene sets, notably two gene sets highlighting STAT signaling/interferon signaling (Table S8, Supplemental Digital Content, http://links.lww.com/QAD/C713), as well as a highly significant gene set seen in Europeans, reflecting viral entry into host cell (Fig. 3, Table S8, Supplemental Digital Content, http://links.lww.com/QAD/C713). For the other HIV reservoir measures, HIV tDNA demonstrated a trend with a gene set involving glycan processing while HIV usRNA showed a trend with a gene set reflecting retroviral transcription.

## Discussion

HIV eradication remains a critical goal in reducing long-term morbidity and mortality. ART suppression does not fully restore health; HIV+ individuals have higher levels of immune activation [74] and increased mortality [3] compared to HIV-uninfected individuals. ART is also expensive and a challenge to provide sustainable access for a global population [75]. HIV cure trials to date have yielded disappointing results [76–81]. Novel approaches are needed to better target immunologic pathways that may drive HIV persistence.

Our study is the first genomic study to evaluate several measures of the HIV reservoir in HIV+ noncontrollers, including HIV intact DNA, which estimates the putative 'replication-competent' reservoir [19,82,83]. We also performed direct HLA typing of 25 polymorphic exons and included detailed clinical data on timing of ART initiation, one of the strongest clinical predictors of HIV reservoir size. Unlike prior genomic studies focusing on 'elite controllers,' we studied treated HIV+ noncontrollers, the majority of people living with HIV (PLWH). Previously reported 'protective' HLA and *CCR5Δ32* mutations during untreated disease [6,15–18] predicted smaller HIV reservoirs in our study. These findings suggest that the same immunologic pathways controlling viral setpoint might also influence HIV reservoir size during ART, and/or that variable 'penetrance' [84] of these mutations might manifest as differential pheontypes among HIV 'elite controllers' and noncontrollers.

The most striking finding from our SNP-based analysis was the identification of several SNPs in *MX1*, which encodes for a potent antiviral factor that inhibits replication of several RNA viruses, including influenza A and measles, and DNA viruses, such as hepatitis B [85], and is part of the antiviral response induced by type I (interferon gamma [IFN-α, IFN-β) and type III (IFN-λ) interferons [86]. *MX1* expression has also been shown to be upregulated in HIV+ individuals [87], especially those with higher viremia [88], as well as in latently HIV-infected cell lines [89]. Although our DNA-based genomewide results cannot infer directionality of gene function without further functional studies, for *MX1*, we were able to use previously reported whole blood expression quantitative trait loci (eQTL) data for the top hit SNP loci [59–61] to determine whether they also predict increased *MX1* expression. Indeed, SNPs

**Table 1. Single nucleotide polymorphisms (SNPs) associated with HIV total DNA (upper table) and unspliced RNA (lower table) in the total study population. Two additional SNPs were significantly associated with total HIV DNA in the subpopulation of European ancestry (middle table). For genes in which there were several SNPs in linkage disequilibrium (LD), the top SNP for each gene is shown, with the full list of SNPs shown in Table S7, Supplemental Digital Content, http://links.lww.com/QAD/C713.**

| SNP | Chrom | Position | Nearest Gene | Gene Location | MAF[a] | Beta[b] | PVE[c] | $P$[d] | $q$[e] | Description |
|---|---|---|---|---|---|---|---|---|---|---|
| HIV total DNA | | | | | | | | | | |
| Total population | | | | | | | | | | |
| rs10670165 | chr21 | 41421873 | *MX1* | Intron/exon/ 5′UTR[f] | 0.45 | −1.2 | 0.15 | $1.3 \times 10^{-7}$ | 0.02 | *MX1* is an interferon stimulated gene and encodes for MX1 (MX Dynamin Like GTPase 1). It is upregulated in HIV+ vs. uninfected individuals [87], associated with higher viremia among HIV+ individuals [88], and linked to HIV-1 latency in cell lines and HIV+ PBMC samples [89]. A closely related paralog gene (*MX2*), encodes for MX2, a potent antiviral host restriction factor which acts against HIV-1 virus [56,57] |
| rs74867009 | chr12 | 25063777 | *IRAG2; LRMP* | 5′UTR[f] | 0.06 | −2.5 | 0.15 | $1.5 \times 10^{-7}$ | 0.02 | Inositol 1, 4,5-triphosphate receptor associated 2 gene (*IRAG2*), also known as lymphoid-restricted membrane protein (*LRMP*), delivers peptides to major MHC class I molecules [63], and is differentially expressed in lymphatic tissue and PBMCs of HIV+ individuals [64], and in response to ART initiation and cessation during treatment interruption [65] |
| rs751660317 | chr2 | 28786774 | *PPP1CB* | Intronic | 0.07 | −1.8 | 0.11 | $4.1 \times 10^{-6}$ | 0.03 | *PPP1CB* encodes for protein phosphatase 1 catalytic subunit beta, which reduces the antiviral potency of MX2 against HIV [58]. It encodes for a subunit of PP1; PP1 is involved in transcription of HIV-1; inhibition of PP1 inhibits HIV-1 transcription [62]. |
| rs776025235 | chr6 | 51638799 | *PKHD1* | Intronic | 0.09 | −1.7 | 0.09 | $2.5 \times 10^{-5}$ | 0.17 | Polycystic Kidney and Hepatic Disease 1 gene (*PKHD1*) is predicted to have a transmembrane-spanning domain and an immunoglobulin-like plexin-transcription-factor domain. No direct relationship with HIV found in the literature |
| N/A | chrX | 41382082 | *DDX3X; NYX* | Intergenic | 0.18 | −1.4 | 0.09 | $2.5 \times 10^{-5}$ | 0.17 | *DDX3X* regulates the production of type I interferons [66] and encodes for DEAD-Box Helicase 3 X-Linked, a protein that shuttles HIV-1 RNA from the nucleus to the cytoplasm. *DDX3X* is upregulated in HIV-infected cells; knockdown of *DDX3X* suppresses HIV-1 replication [67], and plays a key role in innate antimicrobial immunity [66]. |
| rs17506750 | chr14 | 32599402 | *AKAP6* | Intronic | 0.07 | −1.8 | 0.09 | $3.0 \times 10^{-5}$ | 0.20 | *AKAP6* encodes for Protein Kinase A Anchoring Protein 6, which binds to regulatory subunits of protein kinase A (PKA) and anchors them to the nuclear membrane. PKA activation has been associated with HIV-1 infection, T-cell proliferation, and dysfunction [68,69] |
| European Ancestry Subgroup | | | | | | | | | | |
| rs469390 | chr21 | 41446003 | *MX1* | Intron/exon/ 5′UTR[f] | 0.54 | 1.3 | 0.2 | $1.0 \times 10^{-6}$ | 0.03 | *MX1* is an interferon stimulated gene and encodes for MX1 (MX Dynamin Like GTPase 1). It is upregulated in HIV+ vs. uninfected individuals [87], associated with higher viremia among HIV+ individuals [88], and linked to HIV-1 latency in cell lines and HIV+ PBMC samples [89]. A closely related paralog gene (*MX2*), encodes for MX2, a potent antiviral host restriction factor which acts against HIV-1 virus [56,57] |
| N/A | chr11 | 59574427 | *OSBP* | 3′UTR[f] | 0.15 | −1.8 | 0.15 | $1.9 \times 10^{-5}$ | 0.14 | *OSBP* encodes for oxysterol binding protein, which is involved in intracellular lipid transport. It is associated with in-vitro HIV-1 infection of monocyte-derived macrophages (MDMs) from highly (HIV)-exposed seronegatives (HESNs) [71], and is required for the replication of several human viruses such as hepatitis C (HCV), encephalomyocarditis (EMCV), Zika, etc. [70]. |
| HIV UNSPLICED RNA | | | | | | | | | | |
| Total population | | | | | | | | | | |
| N/A | chr19 | 17374631 | *PLVAP; BST2* | Intronic | 0.1 | −1.2 | 0.14 | $6.9 \times 10^{-7}$ | 0.21 | *PLVAP* encodes for Plasmalemma Vesicle Associated Protein, an endothelial membrane protein that controls the entry of lymphocytes and antigens into lymph nodes [73]. The gene is located within 30 kb of *BST2*, an interferon stimulated gene encoding for the host restriction factor tetherin, which is known to inhibit HIV-1 release by directly tethering virions to cells [72] |

[a]MAF = minor allele frequency in the study population.
[b]Beta = estimate from multivariate linear mixed model adjusted for age, sex, nadir $CD4^+$ T-cell count, timing of ART initiation, and ancestry.
[c]PVE = proportion of phenotype variance explained.
[d]$P$ = two-sided $P$-value.
[e]$q$ = two-sided false discovery rate (FDR) Benjamini–Hochberg $q$-value.
[f]UTR = untranslated region.

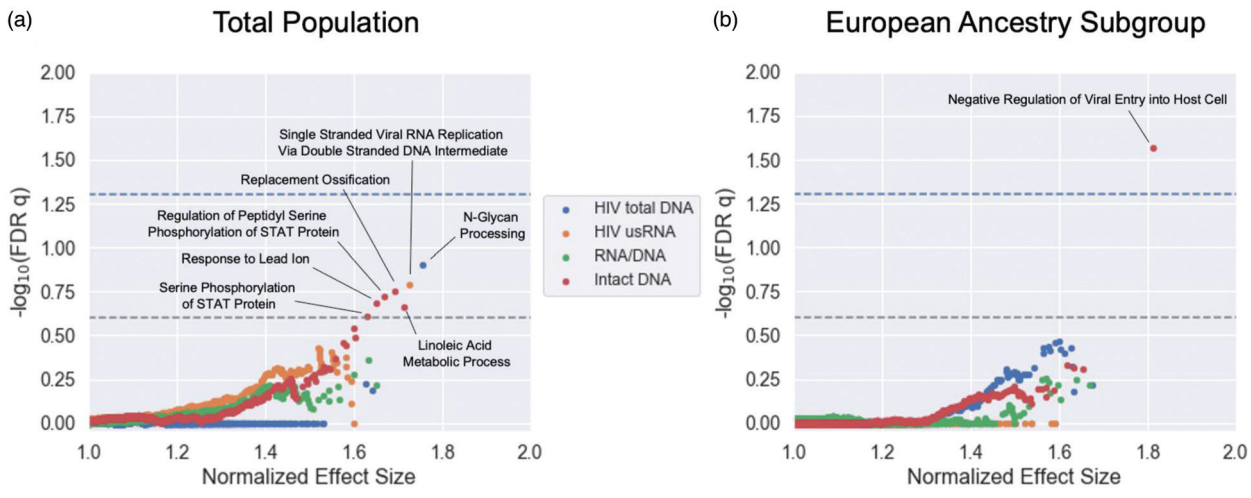


**Fig. 3. Gene set enrichment analysis (GSEA) was used to identify associations between gene pathways and HIV reservoir measures (HIV total DNA, unspliced RNA, RNA/DNA, and intact DNA) in the total study population (a) and the European ancestry subrgoup (b).** GSEA performed using rank-ordered q-values from multi-SNP rare variant analyses (minor allele frequency, MAF, <5%) using the Gene Ontology Biological Process dataset. Horizontal dashed lines represent the GSEA Benjamini–Hochberg false discovery rate (FDR) statistical significance level of $q = 0.05$ (blue) and $q = 0.025$ (grey), respectively.

associated with higher tDNA predicted increased *MX1* expression, while SNPs associated with lower tDNA predicted decreased *MX1* expression. We then performed an additional in-silico check, called 'colocalization analysis,' to integrate genomewide and eQTL results to calculate the probability of *causal* SNP effects on HIV tDNA [90]. However, this only predicted a 1% probability that the top *MX1* SNPs were causally linked to both gene expression and HIV reservoir size. For the other SNP hits, in-silico analyses were not possible since these SNPs did not map to eQTL regions to be able to query any potential directionality of the associations. Additional functional genomics studies, for example, using CRISPR–Cas9 editing [91], may further clarify a potential role for *MX1* (and other interferon pathway genes) in HIV persistence.

There are limitations to our study that deserve mention. Although the HIV reservoir has been shown to be relatively stable over time [51,92,93], our cross-sectional design provides only a 'snapshot' after a median of 5.1 years of ART suppression and may not reflect genetic associations with reservoir decay. Second, as is characteristic of many U.S.-based HIV+ cohorts, our San Francisco-based population consisted mostly of males of European ancestry. Population stratification is a critically important potential bias in any multiethnic genomic study. We approached this potential bias in at least three ways [36,94]: first by calculating principal components and including these as covariates in the final models, second by including a genetic relatedness matrix (GRM) in the models, and finally by performing stratified analyses, focusing on the largest homogenous subpopulation (European ancestry). Overall, the findings observed in the European group did not overlap with the non-Europeans in our small study (Table S9, Supplemental Digital Content, http://links.lww.com/QAD/C713). Third, the majority of the HIV reservoir persists in lymphoid tissues [95,96]. Although recent data suggests that the tissue compartment largely reflects (and is the likely source of) the peripheral compartment [44,96,97], our study may not be generalizable to the tissue HIV reservoir. Fourth, intact HIV DNA represents the potential replication-competent reservoir. Although we observed several genes that were significantly associated with intact HIV DNA in the GSEA, individual genes did not meet statistical significance. The majority of the HIV reservoir consists of defective HIV; thus, quantifying the frequency of 'intact' (e.g. replication-competent) reservoir is challenging [83,98,99]. For example, HIV intact DNA was undetectable in over half of our measured samples while total DNA was measurable in 95% of samples (Figure S6, Supplemental Digital Content, http://links.lww.com/QAD/C713). Hence, the statistical power to detect SNP associations is much lower for this assay compared to the other reservoir measures. By performing GSEA (a method that aggregates several rare variants into immunologically relevant 'gene sets' to test for an association with HIV reservoir size), we may have enhanced the ability to detect real associations with HIV intact DNA (Fig. 3, Table S8, Supplemental Digital Content, http://links.lww.com/QAD/C713).

Our findings are in contrast to two recent genomewide HIV reservoir studies, which did not identify an association with *MX1*, HLA-B*57:01, or *CCR5Δ32*, nor reported similar findings to each other [100,101]. The first study performed GWAS microarray genotyping and whole exome sequencing, measured HIV tDNA

from PBMCs (not $CD4^+$ T cells), and imputed HLA alleles (from genotypes), and did not report significant associations with HLA alleles, *CCR5Δ32*, or SNPs [100]. The second study performed GWAS microarray (no HLA or *CCR5Δ32* typing) and measured HIV tDNA and usRNA from $CD4^+$ T cells. They reported a significant association between tDNA and a SNP in *PTDSS2* (phosphatidylserine synthase 2) at genomewide $p < 5 \times 10^{-8}$ [101], which was not statistically significant in our analysis. Given the polygenic nature of the host immune response, comparing genomic studies can be challenging, let alone for studies of the HIV reservoir (often measured in different ways, from different sample types). Furthermore, how the genetic code translates from DNA to RNA to protein varies by cell type and tissue [102]. Indeed, prior genomewide association studies identifying strong genetic predictors such as HLA and *CCR5Δ32* are estimated to explain only ~13% of the variability in viral load [103]. Using a tool for genome-wide complex trait analysis (GCTA) [104], we performed a similar calculation for our HIV tDNA phenotype. The heritability of our results were estimated to be up to 0.78, but the error bars were large (±0.94), suggesting that the contribution of host genetics in determining the HIV reservoir size likely varies widely (especially when comparing small genomic studies).

Our findings support a surprising association between the host innate immune response (e.g. interferon signaling) and the HIV reservoir, and support recent data suggesting a key role for interferon signaling in determining viral rebound after ART interruption [105]. The 'transcriptionally active' reservoir may be a major source of the 'rebound-competent' reservoir" during long-term ART [106]; perhaps HIV+ individuals with enhanced innate immune responses (e.g. interferon signaling) are better able to restrict a residual 'active' HIV reservoir during ART. Our study also suggests that immune pathways determining viral control during untreated HIV disease (e.g. HLA-B*57:01, *CCR5Δ32*) may also influence HIV persistence during treated disease. Additional studies are needed to validate these findings, including replication in additional cohorts (especially female and non-European populations) and functional genomic studies (e.g. CRISPR−cas9 editing studies).

## Acknowledgements

The authors wish to acknowledge the participation of all the study participants who contributed to this work as well as the clinical research staff of the SCOPE and Options who made this research possible. All funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. All authors provided critical feedback in finalizing the report. S.A.L. conceived and designed the study with critical feedback from S.G.D., T.H., D.L.K., K.R.J., and S.P. S.G. D., J.M., F.H., C.P., R.H., and S.A.L. coordinated the collection, management, and quality control processes for the cohort clinical data and S.G.D., J.M., F.H., C.P., M.P. B., M.S. provided biospecimens. S.A.L. and E.W. performed the whole exome sequencing assays. S.A.L., C.T., K.B., T.P., E.A.G. performed participant sample processing, S.A.L. and E.W. performed the whole exome sequencing, and D.A.S. and S.A.L. performed quality control analyses and the genomic association analyses for the study. S.A.L., C.T., and K.H. performed the qPCR HIV reservoir assays (total DNA, unspliced RNA) in the lab of T.H. C.N.L. and M.L.H. performed the ddPCR HIV reservoir assay (intact DNA) in the labs of F.H., K.R. J., and H.P.K. P.R., D.A.S., T.J.H., and S.A.L. analyzed these HIV reservoir data in relation to host genomic and clinical phenotype data. M.M. and M.C. performed the HLA typing to determine HLA alleles for the analyses. D. A.S. and S.A.L. wrote the report. All authors provided critical feedback in finalizing the manuscript.

Funding: This work was supported in part by the National Institutes of Health: K23GM112526 (S.A.L.), the DARE Collaboratory (U19 AI096109; SGD), the Division of Intramural Research of the National Institutes (M.C.), UM1 AI126623 (H.P.K., K.R.J.), and NIH/NIAID R01A141003 (T.J.H.). This project has been funded in whole or in part with federal funds from the Frederick National Laboratory for Cancer Research, under Contract No. HHSN261200800001E (M.C.). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This Research was supported in part by the Intramural Research Program of the NIH, Frederick National Lab, Center for Cancer Research (M.C.). This work was also supported by the amfAR Research Consortium on HIV Eradication a.k.a. ARCHE (108072-50-RGRL; SGD) and a Collaboration for AIDS Vaccine Discovery (CAVD) grant from the Bill & Melinda Gates Foundation (INV-008500), the Reservoir Assay Validation and Evaluation Network (RAVEN) Study Group (M.B.). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Conflicts of interest

There are no conflicts of interest.

## References

1. Antiretroviral Therapy Cohort Collaboration. **Causes of death in HIV-1-infected patients treated with antiretroviral therapy, 1996–2006: collaborative analysis of 13 HIV cohort studies.** *Clin Infect Dis* 2010; **50**:1387–1396.
2. Eisele E, Siliciano RF. **Redefining the viral reservoirs that prevent HIV-1 eradication.** *Immunity* 2012; **37**:377–388.

3. Lohse N, Hansen AB, Pedersen G, Kronborg G, Gerstoft J, Sorensen HT, *et al*. **Survival of persons with and without HIV infection in Denmark, 1995–2005.** *Ann Intern Med* 2007; **146**:87–95.
4. Volberding PA, Deeks SG. **Antiretroviral therapy and management of HIV infection.** *Lancet* 2010; **376**:49–62.
5. Dalmasso C, Carpentier W, Meyer L, Rouzioux C, Goujard C, Chaix ML, *et al*. **Distinct genetic loci control plasma HIV-RNA and cellular HIV-DNA levels in HIV-1 infection: the ANRS Genome Wide Association 01 study.** *PLoS One* 2008; **3**:e3907.
6. 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.
7. Fellay J, Frahm N, Shianna KV, Cirulli ET, Casimiro DR, Robertson MN, *et al*. **Host genetic determinants of T cell responses to the MRKAd5 HIV-1 gag/pol/nef vaccine in the step trial.** *J Infect Dis* 2011; **203**:773–779.
8. Fellay J, Ge D, Shianna KV, Colombo S, Ledergerber B, Cirulli ET, *et al*. **Common genetic variation and the control of HIV-1 in humans.** *PLoS Genet* 2009; **5**:e1000791.
9. Herbeck JT, Gottlieb GS, Winkler CA, Nelson GW, An P, Maust BS, *et al*. **Multistage genomewide association study identifies a locus at 1q41 associated with rate of HIV-1 disease progression to clinical AIDS.** *J Infect Dis* 2010; **201**:618–626.
10. Le Clerc S, Limou S, Coulonges C, Carpentier W, Dina C, Taing L, *et al*. **Genomewide association study of a rapid progression cohort identifies new susceptibility alleles for AIDS (ANRS Genomewide Association Study 03).** *J Infect Dis* 2009; **200**:1194–1201.
11. Limou S, Le Clerc S, Coulonges C, Carpentier W, Dina C, Delaneau O, *et al*. **Genomewide association study of an AIDS-nonprogression cohort emphasizes the role played by HLA genes (ANRS Genomewide Association Study 02).** *J Infect Dis* 2009; **199**:419–426.
12. Pelak K, Goldstein DB, Walley NM, Fellay J, Ge D, Shianna KV, *et al*. **Host determinants of HIV-1 control in African Americans.** *J Infect Dis* 2010; **201**:1141–1149.
13. International HIVCS. Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PI, *et al*. **The major genetic determinants of HIV-1 control affect HLA class I peptide presentation.** *Science* 2010; **330**:1551–1557.
14. Petrovski S, Fellay J, Shianna KV, Carpenetti N, Kumwenda J, Kamanga G, *et al*. **Common human genetic variants and HIV-1 susceptibility: a genome-wide survey in a homogeneous African population.** *AIDS* 2011; **25**:513–518.
15. de Roda Husman AM, Koot M, Cornelissen M, Keet IP, Brouwer M, Broersen SM, *et al*. **Association between CCR5 genotype and the clinical course of HIV-1 infection.** *Ann Intern Med* 1997; **127**:882–890.
16. Dean M, Carrington M, Winkler C, Huttley GA, Smith MW, Allikmets R, *et al*. **Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study.** *Science* 1996; **273**:1856–1862.
17. Rappaport J, Cho YY, Hendel H, Schwartz EJ, Schachter F, Zagury JF. **32b CCR-5 gene deletion and resistance to fast progression in HIV-1 infected heterozygotes.** *Lancet* 1997; **349**:922–923.
18. Kulkarni S, Lied A, Kulkarni V, Rucevic M, Martin MP, Walker-Sperling V, *et al*. **CCR5AS lncRNA variation differentially regulates CCR5, influencing HIV disease outcome.** *Nat Immunol* 2019; **20**:824–834.
19. Levy CHS, Roychoudhury P, Reeves DB, Amstuz C, Zhu H, *et al*. **A highly multiplexed droplet digital PCR assay to measure the intact HIV-1 proviral reservoir.** *Cell Rep Med* 2021; **2**:100243.
20. Ramirez CM, Sinclair E, Epling L, Lee SA, Jain V, Hsue PY, *et al*. **Immunologic profiles distinguish aviremic HIV-infected adults.** *AIDS* 2016; **30**:1553–1562.
21. Boufassa F, Lechenadec J, Meyer L, Costagliola D, Hunt PW, Pereyra F, *et al*. **Blunted response to combination antiretroviral therapy in HIV elite controllers: an international HIV controller collaboration.** *PLoS One* 2014; **9**:e85516.
22. Emu B, Sinclair E, Hatano H, Ferre A, Shacklett B, Martin JN, *et al*. **HLA class I-restricted T-cell responses may contribute to the control of human immunodeficiency virus infection, but such responses are not always necessary for long-term virus control.** *J Virol* 2008; **82**:5398–5407.
23. Grebe E, Facente SN, Bingham J, Pilcher CD, Powrie A, Gerber J, *et al*. **Interpreting HIV diagnostic histories into infection time estimates: analytical framework and online tool.** *BMC Infect Dis* 2019; **19**:894.
24. Martin MP, Gao X, Lee JH, Nelson GW, Detels R, Goedert JJ, *et al*. **Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS.** *Nat Genet* 2002; **31**:429–434.
25. Martin MP, Qi Y, Gao X, Yamada E, Martin JN, Pereyra F, *et al*. **Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1.** *Nat Genet* 2007; **39**:733–740.
26. Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DI, *et al*. **Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure.** *Cell* 2013; **155**:540–551.
27. Bruner KM, Murray AJ, Pollack RA, Soliman MG, Laskey SB, Capoferri AA, *et al*. **Defective proviruses rapidly accumulate during acute HIV-1 infection.** *Nat Med* 2016; **22**:1043–1049.
28. Malnati MS, Scarlatti G, Gatto F, Salvatori F, Cassina G, Rutigliano T, *et al*. **A universal real-time PCR assay for the quantification of group-M HIV-1 proviral load.** *Nat Protoc* 2008; **3**:1240–1248.
29. Chapman B, Kirchner R, Pantano L, Naumenko S, De Smet M, Beltrame L, *et al*. bcbio/bcbio-nextgen. In. v.1.2.8. ed: Zenodo. Available at https://doi.org/10.5281/zenodo.5781867.
30. Li H. **Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.** *arXiv* 20131303.3997v2 [q-bio.GN]. Available at https://doi.org/10.48550/arXiv.1303.3997.
31. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, *et al*. **Scaling accurate genetic variant discovery to tens of thousands of samples.** *bioRxiv* 2017:201178. Available at https://www.biorxiv.org/content/10.1101/201178v3.
32. Picard toolkit. 2019.
33. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, *et al*. **A framework for variation discovery and genotyping using next-generation DNA sequencing data.** *Nat Genet* 2011; **43**:491–498.
34. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al*. **PLINK: a tool set for whole-genome association and population-based linkage analyses.** *Am J Hum Genet* 2007; **81**:559–575.
35. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, *et al*. **The variant call format and VCFtools.** *Bioinformatics* 2011; **27**:2156–2158.
36. Gogarten SM, Sofer T, Chen H, Yu C, Brody JA, Thornton TA, *et al*. **Genetic association testing using the GENESIS R/Bioconductor package.** *Bioinformatics* 2019; **35**:5346–5348.
37. Wang K, Li M, Hakonarson H. **ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data.** *Nucleic Acids Res* 2010; **38**:e164.
38. Chen H, Huffman JE, Brody JA, Wang C, Lee S, Li Z, *et al*. **Efficient variant set mixed model association tests for continuous and binary traits in large-scale whole-genome sequencing studies.** *Am J Hum Genet* 2019; **104**:260–274.
39. Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. **Rare-variant association testing for sequencing data with the sequence kernel association test.** *Am J Hum Genet* 2011; **89**:82–93.
40. Bioconductor Core Team BPMMBO. Txdb.hsapiens.ucsc.hg38.knowngene. 2017.
41. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, *et al*. **PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes.** *Nat Genet* 2003; **34**:267–273.
42. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, *et al*. **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.** *Proc Natl Acad Sci USA* 2005; **102**:15545–15550.
43. Seabold; Skipper; Perktold J. **Statsmodels: econometric and statistical modeling with python**. *Proceedings of the 9th Python in Science Conference*; 2010.
44. Josefsson L, von Stockenstrom S, Faria NR, Sinclair E, Bacchetti P, Killian M, *et al*. **The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time.** *Proc Natl Acad Sci USA* 2013; **110**:E4987–E4996.

45. Buzon MJ, Martin-Gayo E, Pereyra F, Ouyang Z, Sun H, Li JZ, *et al*. **Long-term antiretroviral treatment initiated at primary HIV-1 infection affects the size, composition, and decay kinetics of the reservoir of HIV-1-infected CD4 T cells.** *J Virol* 2014; **88**:10056–10065.
46. Archin NM, Vaidya NK, Kuruc JD, Liberty AL, Wiegand A, Kearney MF, *et al*. **Immediate antiviral therapy appears to restrict resting CD4+ cell HIV-1 infection without accelerating the decay of latent infection.** *Proc Natl Acad Sci USA* 2012; **109**:9523–9528.
47. Hocqueloux L, Avettand-Fenoel V, Jacquot S, Prazuck T, Legac E, Melard A, *et al*. **Long-term antiretroviral therapy initiated during primary HIV-1 infection is key to achieving both low HIV reservoirs and normal T cell counts.** *J Antimicrob Chemother* 2013; **68**:1169–1178.
48. Jain V, Hartogensis W, Bacchetti P, Hunt PW, Hatano H, Sinclair E, *et al*. **Antiretroviral therapy initiated within 6 months of HIV infection is associated with lower T-cell activation and smaller HIV reservoir size.** *J Infect Dis* 2013; **208**:1202–1211.
49. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in adults and adolescents with HIV. Department of Health and Human Services. Available at https://clinicalinfo.hiv.gov/sites/default/files/guidelines/documents/AdultandAdolescentGL.pdf [Accessed 18 December 2019]
50. Eriksson S, Graf EH, Dahl V, Strain MC, Yukl SA, Lysenko ES, *et al*. **Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies.** *PLoS Pathogens* 2013; **9**:e1003174.
51. Bachmann N, von Siebenthal C, Vongrad V, Turk T, Neumann K, Beerenwinkel N, *et al*. **Determinants of HIV-1 reservoir size and long-term dynamics during suppressive ART.** *Nat Commun* 2019; **10**:3193.
52. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. **Controlling the false discovery rate in behavior genetics research.** *Behav Brain Res* 2001; **125**:279–284.
53. Yang J, Weedon MN, Purcell S, Lettre G, Estrada K, Willer CJ, *et al*. **Genomic inflation factors under polygenic inheritance.** *Eur J Hum Genet* 2011; **19**:807–812.
54. Staeheli P, Haller O. **Human MX2/MxB: a potent interferon-induced postentry inhibitor of herpesviruses and HIV-1.** *J Virol* 2018; **92**:.
55. Haller O, Kochs G. **Human MxA protein: an interferon-induced dynamin-like GTPase with broad antiviral activity.** *J Interferon Cytokine Res* 2011; **31**:79–87.
56. Goujon C, Moncorge O, Bauby H, Doyle T, Ward CC, Schaller T, *et al*. **Human MX2 is an interferon-induced postentry inhibitor of HIV-1 infection.** *Nature* 2013; **502**:559–562.
57. Liu Z, Pan Q, Ding S, Qian J, Xu F, Zhou J, *et al*. **The interferon-inducible MxB protein inhibits HIV-1 infection.** *Cell Host Microbe* 2013; **14**:398–410.
58. Betancor G, Jimenez-Guardeno JM, Lynham S, Antrobus R, Khan H, Sobala A, *et al*. **MX2-mediated innate immunity against HIV-1 is regulated by serine phosphorylation.** *Nat Microbiol* 2021; **6**:1031–1042.
59. Consortium GT, Laboratory DA, Coordinating Center-Analysis Working Group, Statistical Methods groups-Analysis Working Group, Enhancing Gg, Fund NIHC, *et al*.. **Genetic effects on gene expression across human tissues.** *Nature* 2017; **550**:204–213.
60. Szabo PA, Levitin HM, Miron M, Snyder ME, Senda T, Yuan J, *et al*. **Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease.** *Nat Commun* 2019; **10**:4706.
61. Wang S, Zhang Q, Hui H, Agrawal K, Karris MAY, Rana TM. **An atlas of immune cell exhaustion in HIV-infected individuals revealed by single-cell transcriptomics.** *Emerg Microbes Infect* 2020; **9**:2333–2347.
62. Ammosova T, Jerebtsova M, Beullens M, Voloshin Y, Ray PE, Kumar A, *et al*. **Nuclear protein phosphatase-1 regulates HIV-1 transcription.** *J Biol Chem* 2003; **278**:32189–32194.
63. Collins KL. **How HIV evades CTL recognition.** *Curr HIV Res* 2003; **1**:31–40.
64. Li Q, Schacker T, Carlis J, Beilman G, Nguyen P, Haase AT. **Functional genomic analysis of the response of HIV-1-infected lymphatic tissue to antiretroviral therapy.** *J Infect Dis* 2004; **189**:572–582.
65. Vahey MT, Wang Z, Su Z, Nau ME, Krambrink A, Skiest DJ, *et al*. **CD4+ T-cell decline after the interruption of antiretroviral therapy in ACTG A5170 is predicted by differential expression of genes in the ras signaling pathway.** *AIDS Res Hum Retroviruses* 2008; **24**:1047–1066.
66. Szappanos D, Tschismarov R, Perlot T, Westermayer S, Fischer K, Platanitis E, *et al*. **The RNA helicase DDX3X is an essential mediator of innate antimicrobial immunity.** *PLoS Pathog* 2018; **14**:e1007397.
67. Yedavalli VS, Neuveut C, Chi YH, Kleiman L, Jeang KT. **Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function.** *Cell* 2004; **119**:381–392.
68. Perfettini JL, Roumier T, Castedo M, Larochette N, Boya P, Raynal B, *et al*. **NF-kappaB and p53 are the dominant apoptosis-inducing transcription factors elicited by the HIV-1 envelope.** *J Exp Med* 2004; **199**:629–640.
69. Troger J, Moutty MC, Skroblin P, Klussmann E. **A-kinase anchoring proteins as potential drug targets.** *Br J Pharmacol* 2012; **166**:420–433.
70. Roberts BL, Severance ZC, Bensen RC, Le-McClain AT, Malinky CA, Mettenbrink EM, *et al*. **Differing activities of oxysterol-binding protein (OSBP) targeting antiviral compounds.** *Antiviral Res* 2019; **170**:104548.
71. Saulle I, Ibba SV, Vittori C, Fenizia C, Mercurio V, Vichi F, *et al*. **Sterol metabolism modulates susceptibility to HIV-1 Infection.** *AIDS* 2020; **34**:1593–1602.
72. Perez-Caballero D, Zang T, Ebrahimi A, McNatt MW, Gregory DA, Johnson MC, *et al*. **Tetherin inhibits HIV-1 release by directly tethering virions to cells.** *Cell* 2009; **139**:499–511.
73. Rantakari P, Auvinen K, Jappinen N, Kapraali M, Valtonen J, Karikoski M, *et al*. **The endothelial protein PLVAP in lymphatics controls the entry of lymphocytes and antigens into lymph nodes.** *Nat Immunol* 2015; **16**:386–396.
74. Hunt PW, Martin JN, Sinclair E, Bredt B, Hagos E, Lampiris H, *et al*. **T cell activation is associated with lower CD4+ T cell gains in human immunodeficiency virus-infected patients with sustained viral suppression during antiretroviral therapy.** *J Infect Dis* 2003; **187**:1534–1543.
75. Deeks SG, Archin N, Cannon P, Collins S, Jones RB, de Jong M, *et al*. **Research priorities for an HIV cure: International AIDS Society Global Scientific Strategy 2021.** *Nat Med* 2021; **27**:2085–2098.
76. Gay CL, Kuruc JD, Falcinelli SD, Warren JA, Reifeis SA, Kirchherr JL, *et al*. **Assessing the impact of AGS-004, a dendritic cell-based immunotherapy, and vorinostat on persistent HIV-1 Infection.** *Sci Rep* 2020; **10**:5134.
77. Fidler S, Stohr W, Pace M, Dorrell L, Lever A, Pett S, *et al*. **Antiretroviral therapy alone versus antiretroviral therapy with a kick and kill approach, on measures of the HIV reservoir in participants with recent HIV infection (the RIVER trial): a phase 2, randomised trial.** *Lancet* 2020; **395**:888–898.
78. Gutierrez C, Serrano-Villar S, Madrid-Elena N, Perez-Elias MJ, Martin ME, Barbas C, *et al*. **Bryostatin-1 for latent virus reactivation in HIV-infected patients on antiretroviral therapy.** *AIDS* 2016; **30**:1385–1392.
79. Vibholm L, Schleimann MH, Hojen JF, Benfield T, Offersen R, Rasmussen K, *et al*. **Short-course toll-like receptor 9 agonist treatment impacts innate immunity and plasma viremia in individuals with human immunodeficiency virus infection.** *Clin Infect Dis* 2017; **64**:1686–1695.
80. Riddler SA, Para M, Benson CA, Mills A, Ramgopal M, DeJesus E, *et al*. **Vesatolimod, a toll-like receptor 7 agonist, induces immune activation in virally suppressed adults living with human immunodeficiency virus-1.** *Clin Infect Dis* 2021; **72**:e815–e824.
81. Elliott JH, McMahon JH, Chang CC, Lee SA, Hartogensis W, Bumpus N, *et al*. **Short-term administration of disulfiram for reversal of latent HIV infection: a phase 2 dose-escalation study.** *Lancet HIV* 2015; **2**:e520–529.
82. Bruner KM, Wang Z, Simonetti FR, Bender AM, Kwon KJ, Sengupta S, *et al*. **A quantitative approach for measuring the reservoir of latent HIV-1 proviruses.** *Nature* 2019; **566**:120–125.
83. Levy CN, Hughes SM, Roychoudhury P, Reeves DB, Amstuz C, Zhu H, *et al*. **A highly multiplexed droplet digital PCR assay to measure the intact HIV-1 proviral reservoir.** *Cell Rep Med* 2021; **2**:100243.
84. Cooper DN, Krawczak M, Polychronakos C, Tyler-Smith C, Kehrer-Sawatzki H. **Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease.** *Hum Genet* 2013; **132**:1077–1130.

85. Verhelst J, Hulpiau P, Saelens X. **Mx proteins: antiviral gatekeepers that restrain the uninvited.** *Microbiol Mol Biol Rev* 2013; **77**:551–566.
86. Verhelst J, Parthoens E, Schepens B, Fiers W, Saelens X. **Interferon-inducible protein Mx1 inhibits influenza virus by interfering with functional viral ribonucleoprotein complex assembly.** *J Virol* 2012; **86**:13445–13455.
87. Zhao G, Liu L, Su B, Zhang T, Chen P, Li W, *et al*. **The dynamic changes of interferon lambdas related genes and proteins in JAK/STAT pathway in both acute and chronic HIV-1 infected patients.** *AIDS Res Ther* 2017; **14**:31.
88. Van Hecke C, Trypsteen W, Malatinkova E, De Spiegelaere W, Vervisch K, Rutsaert S, *et al*. **Early treated HIV-1 positive individuals demonstrate similar restriction factor expression profile as long-term nonprogressors.** *EBioMedicine* 2019; **41**:443–454.
89. Lee SY, Choi BS, Yoon CH, Kang C, Kim K, Kim KC. **Selection of biomarkers for HIV-1 latency by integrated analysis.** *Genomics* 2019; **111**:327–333.
90. Hormozdiari F, van de Bunt M, Segre AV, Li X, Joo JWJ, Bilow M, *et al*. **Colocalization of GWAS and eQTL signals detects target genes.** *Am J Hum Genet* 2016; **99**:1245–1260.
91. Shifrut E, Carnevale J, Tobin V, Roth TL, Woo JM, Bui CT, *et al*. **Genome-wide CRISPR screens in primary human T cells reveal key regulators of immune function.** *Cell* 2018; **175**:1958–1971e1915.
92. Peluso MJ, Bacchetti P, Ritter KD, Beg S, Lai J, Martin JN, *et al*. **Differential decay of intact and defective proviral DNA in HIV-1-infected individuals on suppressive antiretroviral therapy.** *JCI Insight* 2020; **5**:e132997.
93. Falcinelli SD, Kilpatrick KW, Read J, Murtagh R, Allard B, Ghofrani S, *et al*. **Longitudinal dynamics of intact HIV proviral DNA and outgrowth virus frequencies in a cohort of individuals receiving antiretroviral therapy.** *J Infect Dis* 2021; **224**:92–100.
94. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. **Principal components analysis corrects for stratification in genome-wide association studies.** *Nat Genet* 2006; **38**:904–909.
95. Pantaleo G, Graziosi C, Butini L, Pizzo PA, Schnittman SM, Kotler DP, *et al*. **Lymphoid organs function as major reservoirs for human immunodeficiency virus.** *Proc Natl Acad Sci USA* 1991; **88**:9838–9842.
96. Evering TH, Mehandru S, Racz P, Tenner-Racz K, Poles MA, Figueroa A, *et al*. **Absence of HIV-1 evolution in the gut-associated lymphoid tissue from patients on combination antiviral therapy initiated during primary infection.** *PLoS Pathog* 2012; **8**:e1002506.
97. Imamichi H, Degray G, Dewar RL, Mannon P, Yao M, Chairez C, *et al*. **Lack of compartmentalization of HIV-1 quasispecies between the gut and peripheral blood compartments.** *J Infect Dis* 2011; **204**:309–314.
98. Procopio FA, Fromentin R, Kulpa DA, Brehm JH, Bebin AG, Strain MC, *et al*. **A novel assay to measure the magnitude of the inducible viral reservoir in HIV-infected individuals.** *EBioMedicine* 2015; **2**:872–881.
99. Eriksson S, Graf EH, Dahl V, Strain MC, Yukl SA, Lysenko ES, *et al*. **Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies.** *PLoS Pathog* 2013; **9**:e1003174.
100. Thorball CW, Borghesi A, Bachmann N, Von Siebenthal C, Vongrad V, Turk T, *et al*. **Host genomics of the HIV-1 reservoir size and its decay rate during suppressive antiretroviral treatment.** *J Acquir Immune Defic Syndr* 2020; **85**:517–524.
101. Zhang Z, Trypsteen W, Blaauw M, Chu X, Rutsaert S, Vandekerckhove L, *et al*. **IRF7 and RNH1 are modifying factors of HIV-1 reservoirs: a genome-wide association analysis.** *BMC Med* 2021; **19**:282.
102. Dolcemascolo R, Goiriz L, Montagud-Martinez R, Rodrigo G. **Gene regulation by a protein translation factor at the single-cell level.** *PLoS Comput Biol* 2022; **18**:e1010087.
103. McLaren PJ, Carrington M. **The impact of host genetic variation on infection with HIV-1.** *Nat Immunol* 2015; **16**:577–583.
104. Yang J, Lee SH, Goddard ME, Visscher PM. **GCTA: a tool for genome-wide complex trait analysis.** *Am J Hum Genet* 2011; **88**:76–82.
105. Gondim MVP, Sherrill-Mix S, Bibollet-Ruche F, Russell RM, Trimboli S, Smith AG, *et al*. **Heightened resistance to host type 1 interferons characterizes HIV-1 at transmission and after antiretroviral therapy interruption.** *Sci Transl Med* 2021; **13**: eabd8179. doi: 10.1126/scitranslmed.abd8179.
106. Einkauf KB, Osborn MR, Gao C, Sun W, Sun X, Lian X, *et al*. **Parallel analysis of transcription, integration, and sequence of single HIV-1 proviruses.** *Cell* 2022; **185**:66–282e215.