

Supplementary Figures

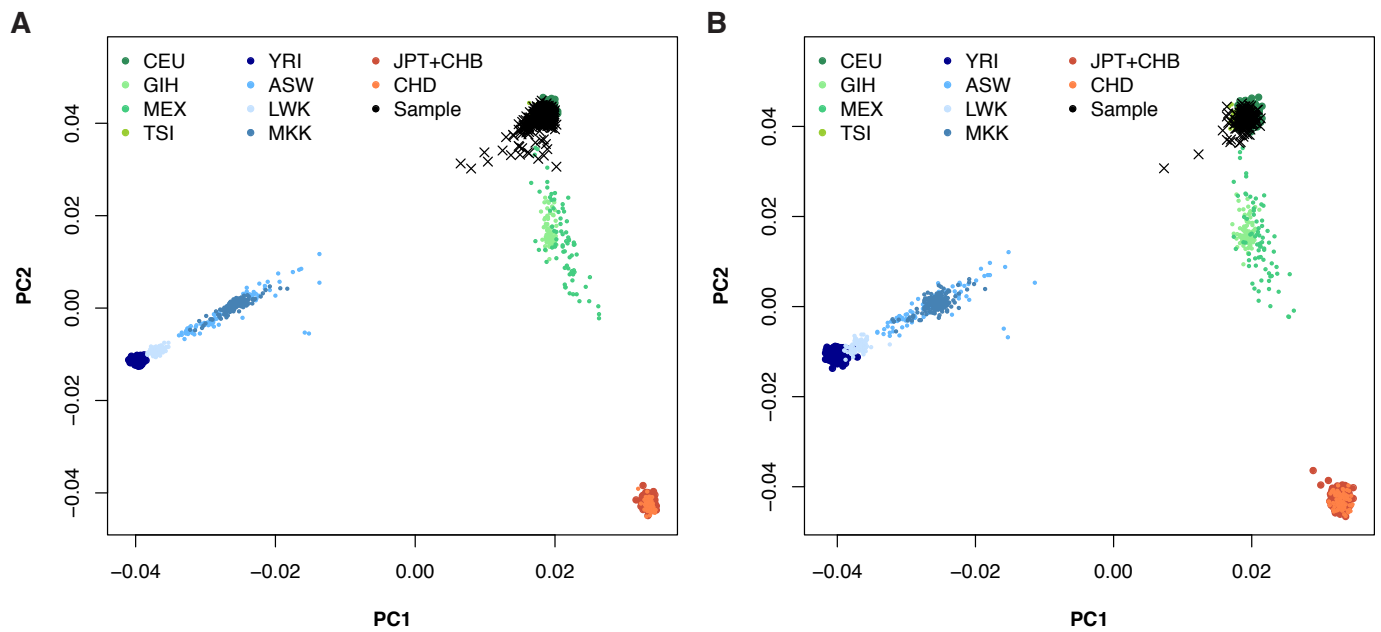


Figure S1. Principal component analyses (PCA) with population references from the HapMap project. (A) The 797 genotyped individuals (black crosses) colocalizing with the HapMap reference samples from CEU (Northern Europeans from Utah) and TSI (Tuscans from Italy). (B) The 194 individuals also exome sequenced (black crosses) cluster as expected still with CEU and TSI.

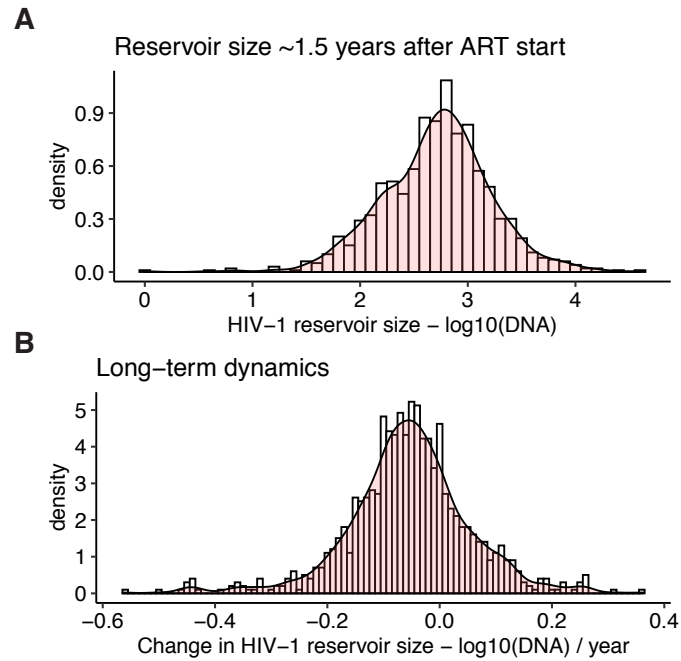


Figure S2. The HIV-1 reservoir size and decay rates for 797 genotyped individuals on ART for a median of 5.4 years. (A) Histograms of the HIV-1 reservoir size in \log_{10} total HIV-1 DNA/1 million genomic equivalents measured ~1.5 after initiating ART. (B) Histogram of the HIV-1 reservoir decay rate in \log_{10} total HIV-1 DNA/1 million genomic equivalents per year based on the three measurements of total HIV-1 DNA levels taken at the median of 1.5, 3.5 and 5.4 years after initiation of ART.

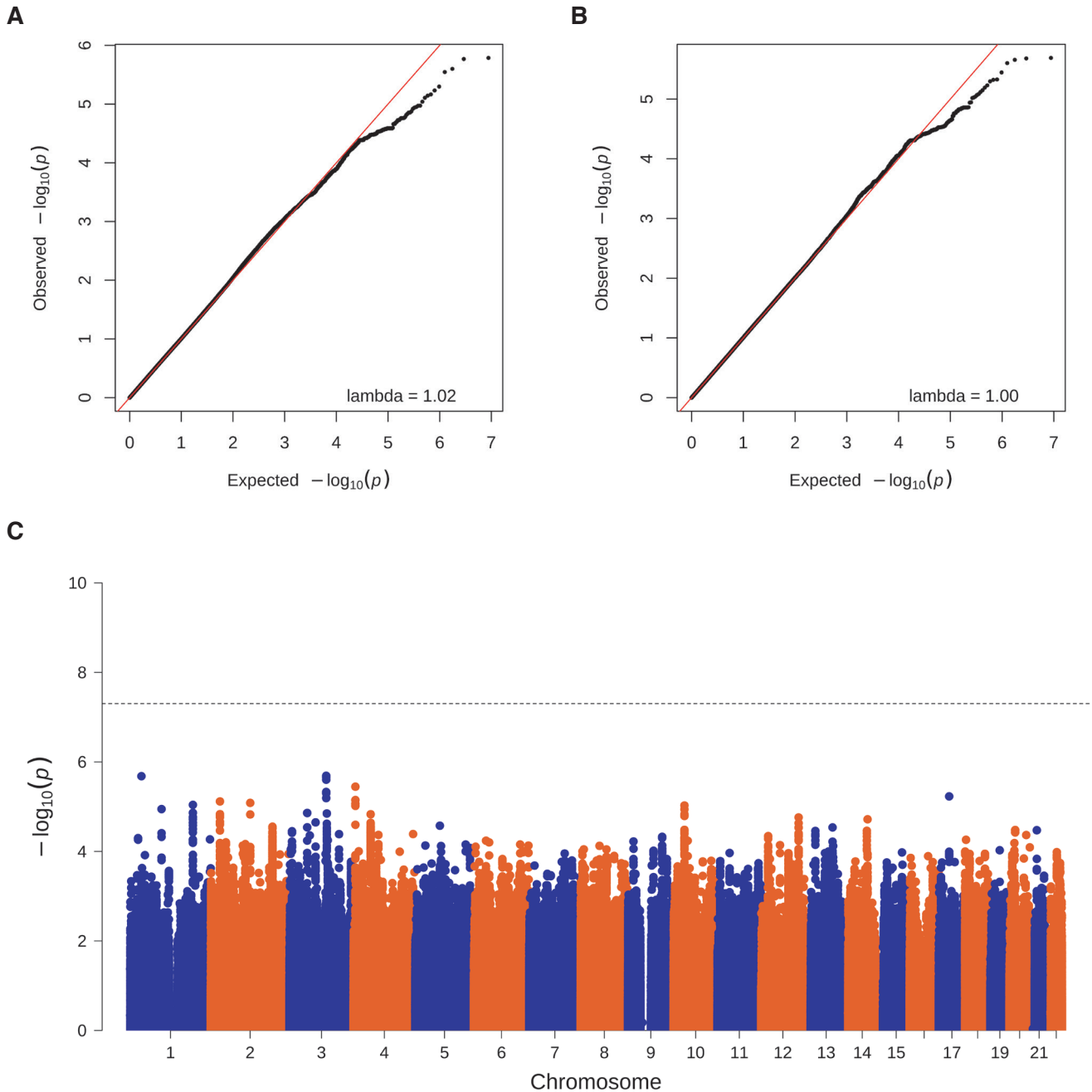


Figure S3. GWAS results for the HIV-1 reservoir size and decay rate. (A) Quantile-quantile plot for the GWAS of HIV-1 reservoir size showing the observed $-\log_{10}(P)$ (black dots, y-axis) versus expected $-\log_{10}(P)$ under the null hypothesis (red line). Lambda indicates the genomic inflation factor. Values ~ 1 indicates the lack of genomic inflation due to confounding factors (B) Quantile-quantile plot for the GWAS of the HIV-1 decay rate between 1.5 – 5.5 years after ART initiation. (C) Manhattan plot for the GWAS of the HIV-1 decay rate with association p-values per genetic variant plotted by genomic position. Dashed line indicates the threshold for genome-wide significance ($P = 5e-8$). No variants were found to be genome-wide significant.

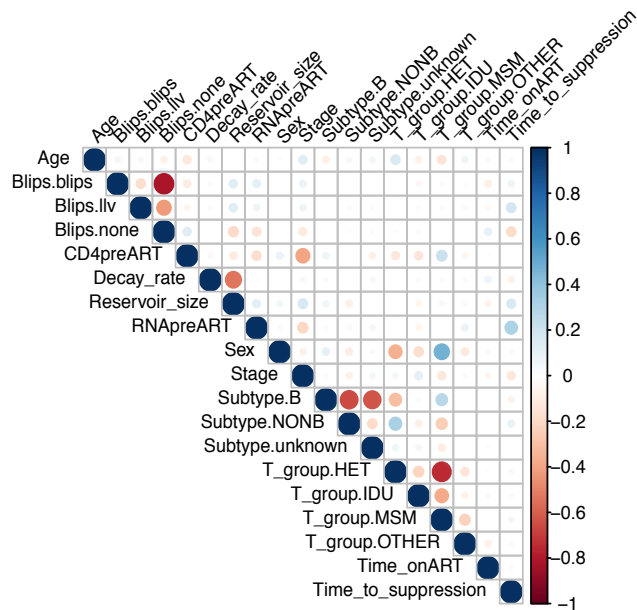


Figure S4. Spearman correlations between known determinants of the HIV-1 reservoir size and decay rate and other variables. The size and color intensity of the circles indicates the level of correlation between variables. Blips.blips refers to the occurrence of viral blips defined by measurements of ≥ 50 HIV-1 RNA copies/mL plasma within a 30-day window. Blips.llv refers to individuals exhibiting low-level viremia defined as consecutive measurements of ≥ 50 HIV-1 RNA copies/mL plasma for more than 30 days. Transmission group (T_group) indicates the self-reported route of infection (heterosexual (HET), intravenous drug usage (IDU), men who have sex with men (MSM), and other (including transfusions and unknown)). Time to viral suppression was the time from initiation of ART to the first viral load measurement below 50 copies/mL HIV-1 plasma RNA. HIV-1 reservoir size was measured in log₁₀ total HIV-1 DNA/1 million genomic equivalents ~ 1.5 years after initiating ART.

Table S1. GWAS sensitivity analysis

Phenotype	Covariates	min P	Lambda
HIV-1 reservoir size	Basic (age + sex)	1.64E-06	1.02
	RNA pre-ART	2.78E-06	1.01
	CD4 pre-ART	1.54E-06	1.02
	Time on ART	1.96E-06	1.02
	Time to suppression	1.13E-06	1.02
	Stage	1.10E-06	1.02
	T_group.HET	1.71E-06	1.02
	T_group.IDU	1.82E-06	1.02
	T_group.MSM	1.68E-06	1.02
	T_group.OTHER	1.53E-06	1.02
	Subtype - B	1.21E-06	1.02
	Subtype - Non B	1.69E-06	1.02
	Subtype - Unknown	1.52E-06	1.02
	Blips	9.99E-07	1.02
	Low-level viremia	2.76E-06	1.01
	No blips or viremia	1.43E-06	1.02
All	2.39E-06	1.01	
HIV-1 reservoir decay rate	Basic (age + sex)	2.05E-06	1.00
	RNA pre-ART	2.12E-06	1.00
	CD4 pre-ART	1.98E-06	1.00
	Time on ART	1.97E-06	1.00
	Time to suppression	1.91E-06	1.00
	Stage	2.03E-06	1.00
	T_group.HET	2.02E-06	1.00
	T_group.IDU	1.94E-06	1.00
	T_group.MSM	1.94E-06	1.00
	T_group.OTHER	2.21E-06	1.00
	Subtype - B	2.08E-06	1.00
	Subtype - Non B	2.14E-06	1.00
	Subtype - Unknown	2.11E-06	1.00
	Blips	2.21E-06	1.00
	Low-level viremia	2.24E-06	1.00
	No blips or viremia	2.07E-06	1.00
All	2.26E-06	0.99	

Covariates indicates the covariates added in the linear mixed model together with age and sex. Transmission group indicates the self-reported route of infection (heterosexual (HET), intravenous drug usage (IDU), men who have sex with men (MSM), and other (including transfusions and unknown)). The occurrence of viral blips was defined by measurements of ≥ 50 HIV-1 RNA copies/mL plasma within a 30-day window. Individuals with consecutive measurements of ≥ 50 HIV-1 RNA copies/mL plasma for longer durations were classified as exhibiting low-level viremia. Time to viral suppression was the time from initiation of ART to the first viral load measurement below 50 copies/mL HIV-1 plasma RNA. All, indicates that all independent covariates listed in the table were added to the model. Min P is the minimal observed p-value in the corresponding GWAS. Lambda indicates the genomic inflation factor. All values were ~ 1.00 , indicating an absence of genomic inflation of the test statistics due to confounding factors.

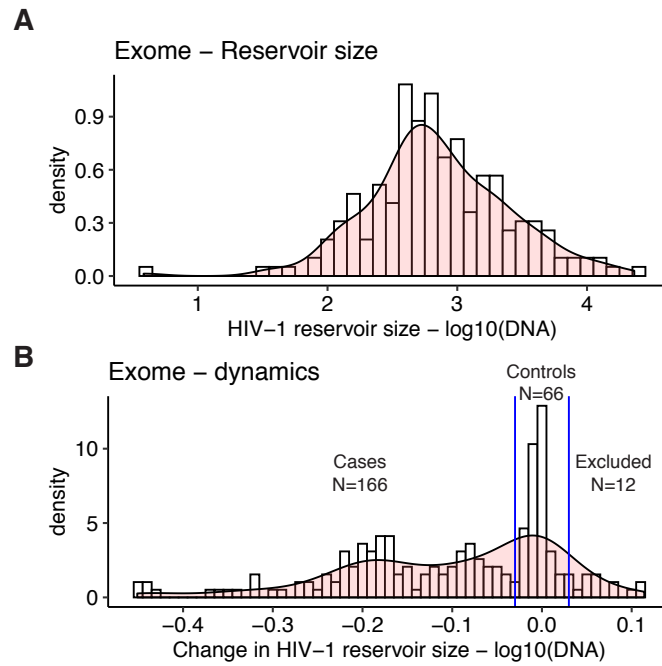


Figure S5. The HIV-1 reservoir size and decay rates for the 194 exome sequenced individuals. (A) Histogram of the HIV-1 reservoir size in \log_{10} total HIV-1 DNA/1 million genomic equivalents measured ~ 1.5 after initiating ART. (B) The non-normal distribution of the decay rate (long-term dynamics) meant that the individuals were split into cases with a decreasing HIV-1 reservoir ($< -0.03 -\log_{10}(\text{DNA})$) and controls with no change in their HIV-1 reservoir size over 5.4 years (≥ -0.03 and $\leq 0.03 -\log_{10}(\text{DNA})$). The blue lines indicate the cutoff points. Individuals with increasing HIV-1 reservoir sizes were excluded (N=12).

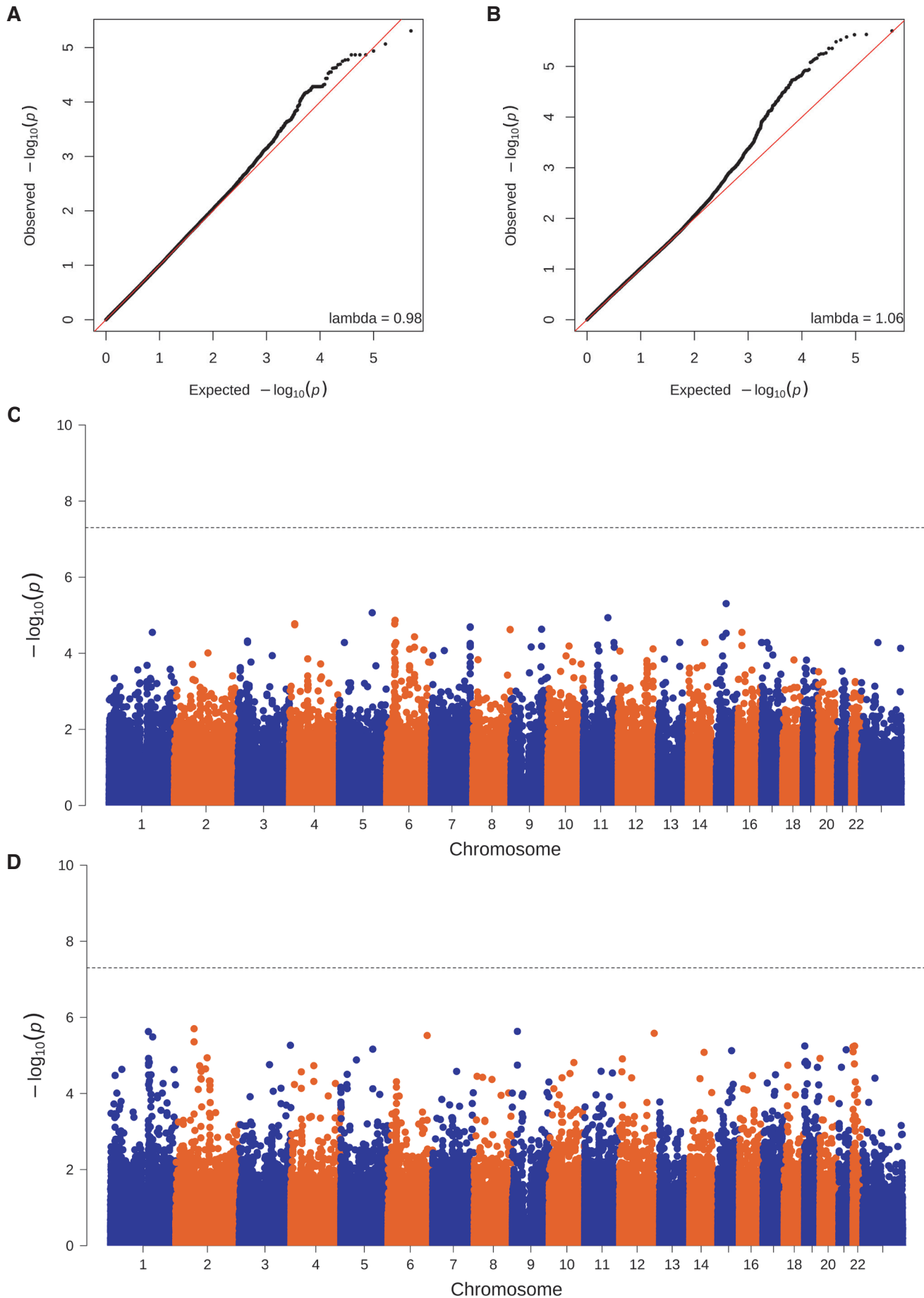


Figure S6. Associations with common variants in 194 exome sequenced individuals. (A) Quantile-quantile plot for GWAS of HIV-1 reservoir size showing the observed $-\log_{10}(P)$ (black dots, y-axis) versus expected $-\log_{10}(P)$ under the null hypothesis (red line). (B) Quantile-quantile plot for GWAS of the HIV-1 decay rate between 1.5 – 5.5 years after ART initiation. (C) Manhattan plot for GWAS of the HIV-1 reservoir size with association p-values per genetic variant plotted by genomic position. Dashed line indicates the threshold for genome-wide significance ($P = 5e-8$). No variants were found to be genome-wide significant. (D) Manhattan plot for the GWAS of the HIV-1 decay rate. No variants were found to be genome-wide significant.

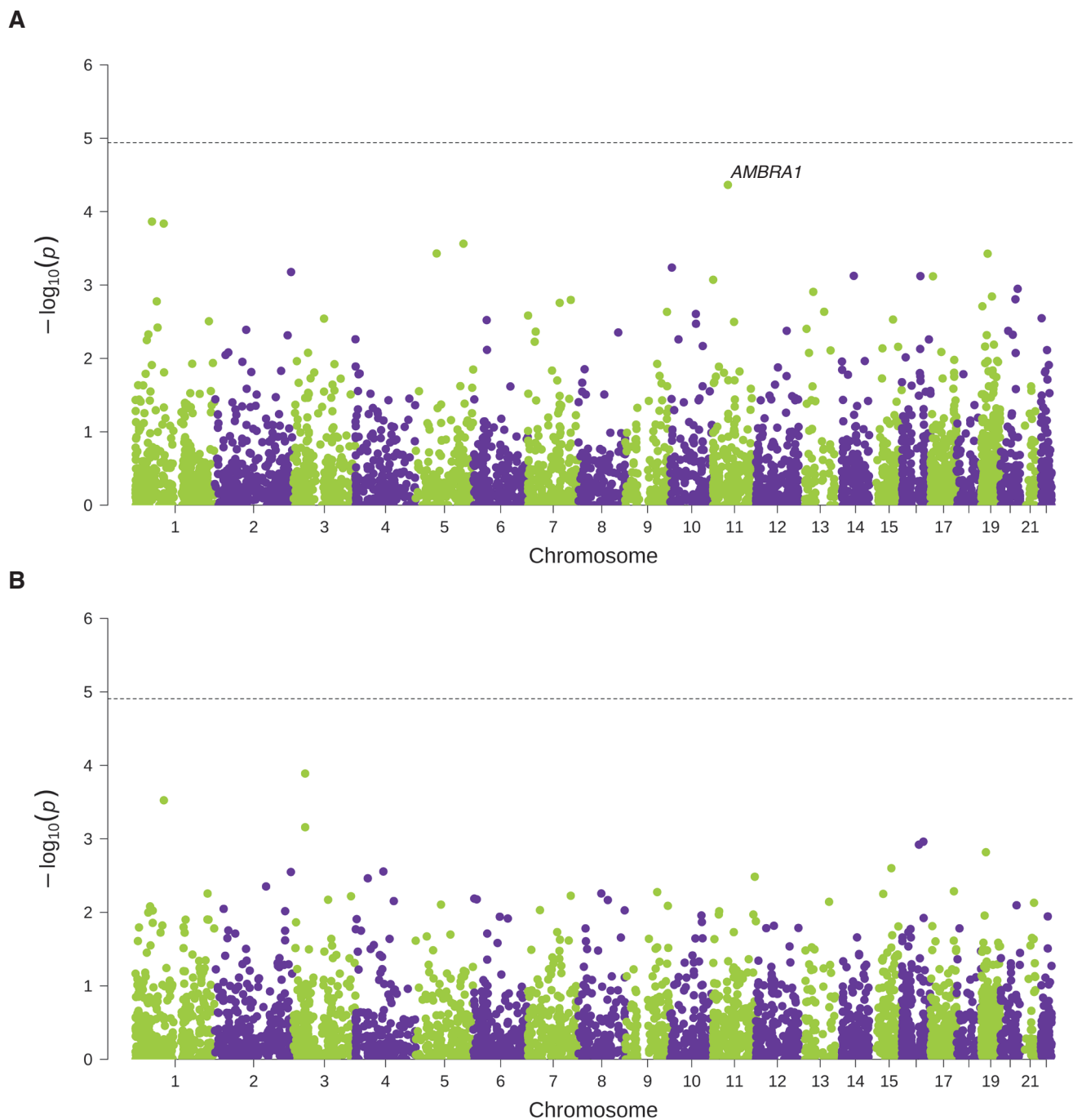


Figure S7. Association results for rare functional variants with optimal sequence kernel association tests (SKAT-O). (A) Manhattan plot for rare variant association analysis using SKAT-O of the HIV-1 reservoir size with association p-values per genetic variant plotted by genomic position. Dashed line indicates the threshold for genome-wide significance following Bonferroni correction ($P = 1.21 \times 10^{-5}$). No variants were found to be genome-wide significant. (B) Manhattan plot for the rare variant association analysis of the HIV-1 decay rate. No variants were found to be genome-wide significant.

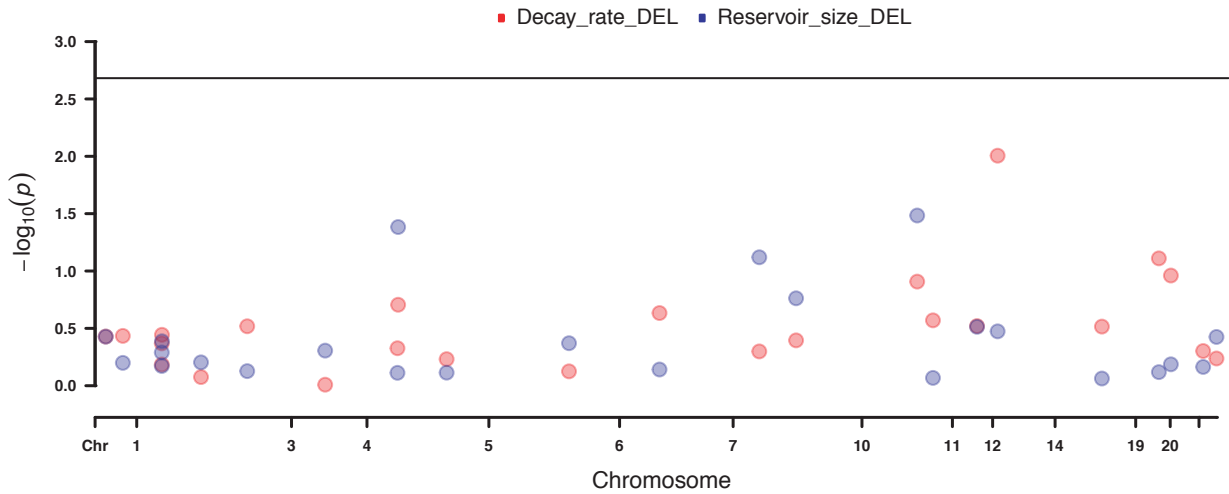
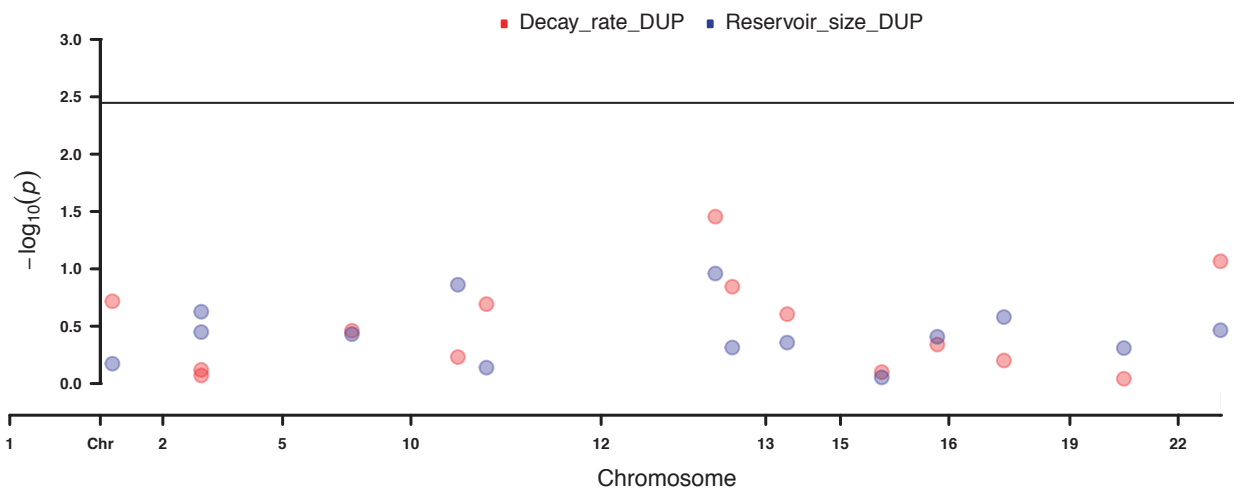
A**B**

Figure S8. Association results for common exonic copy number variations (CNVs). (A) Manhattan plot for CNV deletions (DEL) with association p-values per CNV plotted by genomic position. The association of each CNV with either reservoir size (blue dots) or decay rate (red dots) is indicated. (B) Manhattan plot for CNV duplications (DUP) with association p-values per CNV plotted by genomic position. The association of each CNV with either reservoir size (blue dots) or decay rate (red dots) is marked.

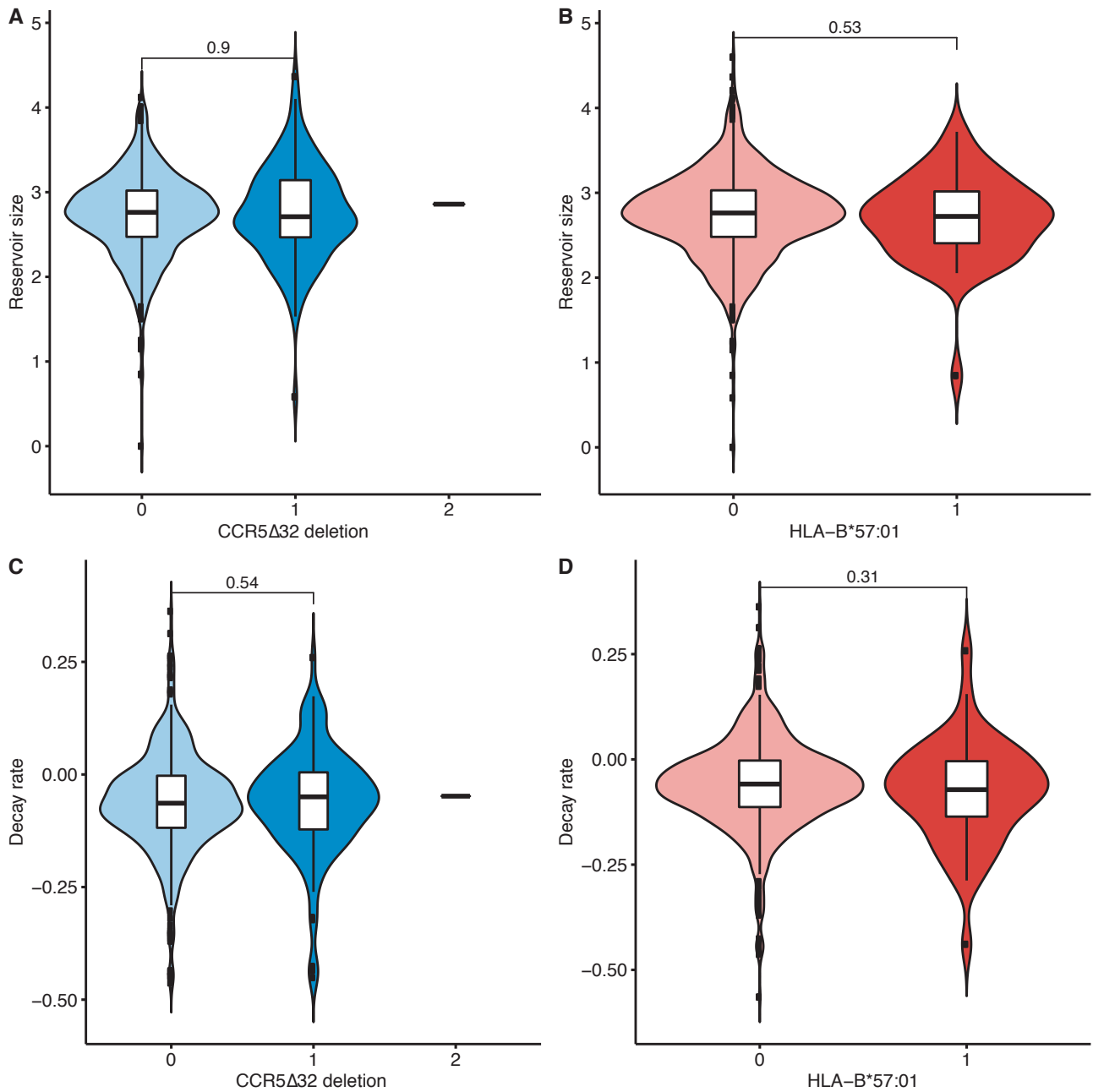


Figure S9. Effect of CCR5Δ32 deletion and the HLA-B*57:01 allele on HIV-1 reservoir size and decay rate in 797 individuals. Statistical differences between groups were determined by Wilcoxon signed-rank tests. (A) Violin plots showing the HIV-1 reservoir size in log₁₀ total HIV-1 DNA/1 million genomic equivalents of individuals carrying either none, one or two alleles of the CCR5Δ32 deletion as tagged by the rs333 SNP. (B) Violin plot with the HIV-1 reservoir size levels grouped by presence of the HLA-B*57:01 allele. (C) Violin plot showing HIV-1 reservoir decay rates of individuals with and without the CCR5Δ32 deletion. (D) Violin plot with the decay rates of individuals with and without the HLA-B*57:01 allele.