1 Host genomics of the HIV-1 reservoir size and its decay rate during

2 suppressive antiretroviral treatment

3

Christian W. Thorball^{1,*}, Alessandro Borghesi^{1,2,*}, Nadine Bachmann^{3,4}, Chantal von 4 Siebenthal^{3,4}, Valentina Vongrad^{3,4}, Teja Turk^{3,4}, Kathrin Neumann^{3,4}, Niko Beerenwinkel^{5,6}, 5 Jasmina Bogojeska⁷, Volker Roth⁸, Yik Lim Kok^{3,4}, Sonali Parbhoo^{8,9}, Mario Wieser⁸, Jürg 6 Böni⁴, Matthieu Perreau¹⁰, Thomas Klimkait¹¹, Sabine Yerly¹², Manuel Battegay¹³, Andri 7 8 Rauch¹⁴, Patrick Schmid¹⁵, Enos Bernasconi¹⁶, Matthias Cavassini¹⁷, Roger D. Kouvos^{3,4}, Huldrych F. Günthard^{3,4}, Karin J. Metzner^{3,4}, Jacques Fellav^{1,18,§} and the Swiss HIV Cohort 9 10 Study 11 1 School of Life Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland 12 13 2 Neonatal Intensive Care Unit, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy. 14 3 Department of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, 15 Zurich, Switzerland. 16 4 Institute of Medical Virology, University of Zurich, Zurich, Switzerland. 17 5 Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland. 6 SIB Swiss Institute of Bioinformatics, Basel, Switzerland. 18 7 IBM Research - Zurich, Rüschlikon, Switzerland. 19 20 8 Department of Mathematics and Computer Science, University of Basel, Basel, Switzerland. 21 9 Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, 22 Cambridge, MA, USA. 10 Division of Immunology and Allergy, Centre Hospitalier Universitaire Vaudois, University 23 24 of Lausanne, Lausanne, Switzerland.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

- 25 11 Division Infection Diagnostics, Department Biomedicine—Petersplatz, University of Basel,
- 26 Basel, Switzerland.
- 27 12 Division of Infectious Diseases and Laboratory of Virology, University Hospital Geneva,
- 28 University of Geneva, Geneva, Switzerland.
- 29 13 Department of Infectious Diseases and Hospital Epidemiology, University Hospital Basel,
- 30 Basel, Switzerland.
- 31 14 Department of Infectious Diseases, University Hospital Bern, Bern, Switzerland.
- 32 15 Division of Infectious Diseases, Cantonal Hospital of St. Gallen, St. Gallen, Switzerland.
- 33 16 Infectious Diseases Service, Regional Hospital of Lugano, Lugano, Switzerland.
- 34 17 Division of Infectious Diseases, Centre Hospitalier Universitaire Vaudois, University of
- 35 Lausanne, Lausanne, Switzerland.
- 36 18 Precision Medicine Unit, Lausanne University Hospital (CHUV) and University of
- 37 Lausanne, Lausanne, Switzerland
- 38 * Equal contribution
- 39
- 40 § Corresponding author:
- 41 Jacques Fellay
- 42 School of Life Sciences
- 43 École Polytechnique Fédérale de Lausanne
- 44 1015 Lausanne, Switzerland
- 45 E-mail: jacques.fellay@epfl.ch
- 46 Phone number: +41216931849
- 47
- 48
- 49 Keywords: HIV, genetics, ART, latent reservoir, GWAS, exome sequencing

50 ABSTRACT

51 Introduction

A major hurdle to HIV-1 eradication is the establishment of a latent viral reservoir early after primary infection. Several factors are known to influence the HIV-1 reservoir size and decay rate on suppressive antiretroviral treatment (ART), but little is known about the role of human genetic variation.

56 Methods

We measured the reservoir size at three time points over a median of 5.4 years, and searched for associations between human genetic variation and two phenotypic readouts: the reservoir size at the first time point and its decay rate over the study period. We assessed the contribution of common genetic variants using genome-wide genotyping data from 797 patients with European ancestry enrolled in the Swiss HIV Cohort Study and searched for a potential impact of rare variants and exonic copy number variants using exome sequencing data generated in a subset of 194 study participants.

64 **Results**

Genome- and exome-wide analyses did not reveal any significant association with the size ofthe HIV-1 reservoir or its decay rate on suppressive ART.

67 **Conclusions**

Our results point to a limited influence of human genetics on the size of the HIV-1 reservoirand its long-term dynamics in successfully treated individuals.

70

- 71
- 72
- 73

74 INTRODUCTION

Combination antiretroviral treatment (ART) has turned the previously lethal infection by human immunodeficiency virus type 1 (HIV-1) into a chronic disease. Despite this significant achievement, HIV-1 as retrovirus, self-integrating its genome into the host chromosome, persists indefinitely in infected individuals during treatment [1–4], and life-long ART is required to control the infection.

80 A major hurdle to HIV-1 eradication is the establishment, already during primary infection, of 81 a latent viral reservoir of HIV-1 DNA persisting as provirus in resting memory CD4⁺ T cells 82 [1,2,5–8]. At the molecular level, chromatin remodeling, epigenetic modifications, 83 transcriptional interference, and availability of transcription factors have been considered as 84 possible mechanisms contributing to HIV-1 latency [9]. The viral reservoir is measurable 85 through different methods, including viral outgrowth assay and intracellular HIV-1 DNA 86 quantification [10,11]. Currently, there is no consensus on the best HIV-1 reservoir biomarker. 87 Total cell-associated HIV-1 DNA, easy to measure in different cell and tissue samples and 88 applicable to large populations, has been shown to be a good proxy for the reservoir size [12]. 89 Indeed, while HIV-1 DNA measurement is able to detect both integrated and nonintegrated 90 viral genomes coding for intact or defective viruses [13], total HIV-1 DNA levels have been 91 shown to correlate with viral outgrowth [14], and to predict the time to viral rebound at 92 treatment interruption [15]. Moreover, the substantial loss of nonintegrated HIV-1 DNA 93 genomes following ART initiation suggests that total HIV-1 DNA after prolonged suppression 94 is largely accounted for by integrated viral genomes [16].

95 After an initial rapid decay following ART initiation, changes of the viral reservoir size over 96 time display wide inter-individual variability. By limiting dilution culture assay, the half-life of 97 the viral reservoir was first estimated to be 44 months (95% confidence interval 27.4-114.5) in 98 individuals with undetectable viremia [4]. A more recent study showed a slow decline of total 99 HIV-1 DNA with a half-life of 13 years after the first four years of suppressive ART [17]. 100 Generally, different studies show a broad variability of the average decay rate, from 2.5 months 101 to no measurable decay [18–26]. One study even reported an increase in the viral reservoir size 102 in as much as 31% of patients in the 4-7 years following ART initiation [27], and recent data 103 from our group confirm this observation, reporting an increase in the reservoir size in 26.8% of 104 individuals in the 1.5-5.5 years after ART initiation [28].

Several factors are known to influence the decay rate of the viral reservoir: initiation of ART during acute HIV-1 infection substantially accelerates the decay rate, while viral blips and lowlevel viremia during ART slow it down, as shown in previous studies [22] and in recent data from our cohort [28]. Conversely, treatment intensification, i.e. treating with additional drugs, does not appear to influence the decay rate, suggesting that residual replication is not the main driver of the viral reservoir [29] or that it may happen in sanctuary sites.

111 Human genetic variants have been shown to influence the outcome of various infections, 112 including HIV. Previous genome-wide association studies (GWAS) addressed the role of 113 common genetic polymorphisms in several HIV-related phenotypes, including plasma viral 114 load (HIV-1 RNA) at set point, exceptional capacity to control viral replication, pace of CD4+ 115 T lymphocyte decline, time to clinical AIDS, rapid progressor status or long-term non-116 progressor status (LTNP) [30–37], and, in one single study, the amount of intracellular HIV-1 117 DNA, measured at a single time point during the chronic phase of infection [38]. Rare genetic 118 variants that are detectable through DNA sequencing technologies have been investigated far 119 less. However, a large exome sequencing study did not reveal any convincing association of 120 such variants with the natural history of HIV disease [39].

121 To date, no studies have addressed the role of human genetic variation in determining the initial 122 viral reservoir size and the reservoir decay rate over time. In the current study, we searched for

- 123 host genetic factors associated with the HIV-1 reservoir size and its long-term dynamics in a
- 124 cohort of 797 HIV-1 positive individuals on suppressive ART for at least five years.
- 125
- 126

127 METHODS

128 Ethics statement

- 129 Participants in the Swiss HIV Cohort Study (SHCS) consented to the cohort study and genetic
- 130 analyses, as approved by the corresponding local Ethics Committees.
- 131

132 Study participants

The SHCS is an ongoing, nation-wide cohort study of HIV-positive individuals, including more than 70% of all persons living with HIV in Switzerland. Clinical and laboratory information has been prospectively recorded at follow-up visits every 3-6 months since 1988 [40]. The general enrolment criteria have been described previously [28]. Additionally, availability of genome-wide genotyping data from previous studies or of a DNA sample for genotyping was required for inclusion in this study (Figure 1).

139

140 Quantification of total HIV-1 DNA

The collection of longitudinal cryopreserved peripheral blood mononuclear cells (PBMCs) from eligible participants and the quantification of total HIV-1 DNA by droplet digital PCR has been described previously along with the calculation of the reservoir decay rate [28]. Briefly, this study utilized total HIV-1 DNA quantifications from three time points at a median of ~1.5 years, ~3.5 years, and ~5.4 years after initiation of ART.

146

147 Genotyping and genome-wide association analyses

Genome-wide genotyping data were obtained from previous GWAS that used various microarrays, including the HumanCore-12, HumanHap550, Human610, Human1M and Infinium CoreExome-24 BeadChips (Illumina Inc., San Diego, CA, USA), or generated from DNA extracted from peripheral blood mononuclear cells using the HumanOmniExpress-24 BeadChip (Illumina Inc., San Diego, CA, USA).

153 Genotypes from each genotyping array were filtered and imputed separately, with variants first 154 flipped to the correct strand with BCFTOOLS (v1.8) according to the human GRCh37 reference 155 genome and filtered based on a less than 20% deviation from the 1000 genomes phase 3 EUR 156 reference panel. Genotypes were phased, and missing genotypes were imputed with EAGLE2 157 [41] and PBWT [42] respectively, using the 1000 Genomes Project Phase 3 reference panel on 158 the Sanger Imputation Service [43]. Study participants were filtered based on European 159 ancestry as determined by principal component analysis (PCA) using EIGENSTRAT (v6.1.4) 160 [44] and the HapMap project [45] as reference populations (Figure S1A). Imputed variants were 161 filtered by minor allele frequency (MAF) < 5%, missingness > 10%, deviation from Hardy-162 Weinberg equilibrium ($P_{HWE} < 1e-6$) and imputation quality score (INFO < 0.8). The remaining 163 genotypes were then combined using PLINK (v1.90b5) [46] prior to analyses.

164 To carry out the GWASs, genome-wide genotypes were tested for association with each of the 165 two study phenotypes (reservoir size or reservoir decay rate) in two separate genome-wide 166 association analyses. Statistical significance was set to the standard genome-wide significance 167 threshold of P < 5e-8 to correct for multiple testing. The associations were computed using 168 linear mixed models with genetic relationship matrixes calculated between pairs of individuals 169 according to the leave-one-chromosome-out method as implemented in GCTA mlma-loco 170 (v1.91.4beta) [47,48], only including age and sex as covariates, to avoid masking of true 171 associations by confounders. To further assess the contribution of variables previously shown 172 to be associated with either reservoir size or decay rate, we ran multiple genome-wide association analyses, each including age, sex, and one single covariate, for each of the two study
phenotypes. Finally, we conducted a GWAS including all the covariates except those showing
mutual correlations. These covariates included time on ART, time to viral suppression,
infection stage (acute or chronic), HIV-1 RNA pre-ART, last CD4+ T cell count pre-ART,
HIV-1 subtype, transmission group, and occurrence of blips or low-level viremia during
treatment.

179 Classical HLA alleles at the four-digit level and variable amino acids within HLA proteins were 180 imputed using SNP2HLA (v1.03) with the T1DGC reference panel consisting of 5,225 181 individuals of European ancestry [49]. Association analyses with the imputed HLA alleles and 182 multi-allelic amino acids was performed using linear regressions in PLINK and multivariate 183 omnibus tests, respectively. For all HLA analyses age, sex and the first principal component 184 was included as covariates.

Genotypes at specific loci, i.e. the CCR5 Δ 32 deletion (rs333) and the HLA-B*57:01 allele, known to influence the setpoint viral load (spVL) [50,51], available from genome-wide genotyping data, were tested for association with the reservoir size and its decay rate in 797 patients. High quality genotyping information on the CCR5 Δ 32 deletion was available for most individuals (N = 687), while all had available HLA information.

190

191 Exome sequencing and analysis

All coding exons were captured using either the Illumina Truseq 65 Mb enrichment kit
(Illumina Inc., San Diego, CA, USA) or the IDT xGen Exome Research Panel v1.0 (Integrated
DNA Technologies Inc., Coralville, IA, USA) and sequenced on the Illumina HiSeq4000.
Sequence reads were aligned to the human reference genome (GRCh37) including decoys with
BWA-MEM (v0.7.10) [52]. PCR duplicates were flagged using Picard tools (v2.18.14) and
variant calling performed using GATK (v3.7) [53].

To ensure a high-quality variant set across capture kit batches, all samples were merged and variants filtered based on sequencing depth ($DP \ge 20$) and genotype quality ($GQ \ge 30$) using BCFTOOLS (v1.8). Furthermore, individual genotypes were set as missing in cases of low depth (DP < 10) or low quality (GQ < 20). The effect of the included variants was annotated with SnpEff (v4.3T) [54].

For single variant association analysis, the VCF file was converted to PLINK format using BCFTOOLS and PLINK. Only variants with a MAF above 5%, missingness per variant below 5% and absence of severe deviation from Hardy-Weinberg equilibrium ($P_{HWE} > 1e-6$) were retained for the subsequent association analyses using PLINK. Sex, age and the first principal component were included as covariates. Only individuals of European descent were retained for the analyses, as determined by PCA (Figure S1B).

209 The combined effect of rare protein-altering variants (MAF < 5%), defined as either missense, 210 stop-gain, frameshift, essential splice variant or an indel by SnpEff, on the reservoir size and 211 decay rate was analyzed using optimal sequence kernel association tests (SKAT-O) [55]. For 212 the decay rate, individuals were split into two groups due to the non-normal distribution; one 213 exhibiting a very high decay over time ($< -0.03 - \log 10$ (DNA)) and another with a stable 214 reservoir size (\geq -0.03 and \leq 0.03 -log10(DNA)). For this case-control analysis we used the 215 SKATbinary function with linear weighted variants as implemented in the SKAT R package. 216 In both cases, the analyses were adjusted for age, sex, and the first principal component.

Classical HLA class I and II alleles at the four-digit level were imputed from the exome sequencing data using HLA*LA [56]. All reads mapping to the MHC region or marked as unmapped were extracted using Samtools (v1.8) and used as input into HLA*LA. For association analyses, the 4-digit HLA alleles were extracted and analyzed using PyHLA [57] assuming an additive model, a minimum frequency of 5% and including age, sex and the first principal component as covariates.

2	2	2
L	L	3

224 Copy number variation

225 Copy number variations (CNVs) were called from exome sequencing data using CLAMMS 226 [58]. CNVs were called for all samples in batches according to the exome capture kit used. 227 Within batches, samples were normalized based on coverage and potential intra-batch effects 228 adjusted for through the use of recommended mapping metrics extracted with Picard tools 229 (v2.18.14). After CNV calling, samples with the number of CNVs two times above the median 230 were excluded (N = 2). CNV association analyses were performed for duplications and 231 deletions separately for common CNVs (frequency > 5 %) with PLINK adjusting for age and 232 sex. Potential rare CNVs (frequency < 5%) impacting immune related genes were examined by 233 overlapping called CNVs with curated immune-related genes from Immport [59] which were 234 also listed as protein coding in GENCODE (v25).

235

236 Statistical analyses

All statistical analyses were performed using the R statistical software (v3.5.2), unlessotherwise specified.

239

240

241 **RESULTS**

242 Host genetic determinants of the reservoir size and long-term dynamics

To investigate the effects of host genetic variation on the size of the HIV-1 reservoir 1.5 years after ART initiation and its long-term dynamics under ART over a median duration of 5.4 years, we performed a GWAS, including 797 well-characterized HIV-1 positive individuals. All study participants were enrolled in the SHCS and were of European ancestry with longitudinal total HIV-1 DNA measurements available (Table 1). The median HIV-1 reservoir size was 2.76

(IQR: 2.48-3.03) log10 total HIV-1 DNA copies/1 million genomic equivalents measured ~1.5
years after initiation of ART (Figure S2A). The median decay rate between 1.5-5.4 years after
initiation of ART was -0.06 (IQR: -0.12-0.00) log10 total HIV-1 DNA copies/1 million
genomic equivalents per year (Figure S2B). With our sample size we had 80% power to detect
variants with a MAF of 10% explaining at least 5% of the variance in HIV-1 reservoir size or
decay rate [60].

254 First, we performed GWAS using age and sex as covariates. We did not observe any genome-255 wide significant variant (P < 5e-8) associated with either HIV-1 reservoir size or long-term 256 dynamics (Figure 2, S3). However, as we have previously determined, multiple factors are 257 associated with the HIV-1 reservoir size and its decay rate [28], some of which are correlated 258 (Figure S4). Thus, we performed additional analyses iteratively including these factors to test 259 whether they could mask genetic associations. We ran multiple GWAS each adjusting for age, 260 sex, plus one of the associated covariates, as well as all of the covariates together. The addition 261 of the covariates did not have any significant effect on the results nor the genome-wide inflation 262 factor (lambda) (Table S1).

Genetic variation in the HLA region has previously been associated with multiple HIV-related outcomes, including spVL [51]. To test whether specific HLA variants were associated with reservoir size or long-term dynamics, we imputed the HLA alleles and amino acids for all 797 individuals from the genotyping data. In line with the previous results, we did not observe any genome-wide significant associations with any HLA allele or amino acid.

268

269 Impact of protein-coding and rare variants

To assess the impact of rare variants as well as protein-coding variants missed by genotyping arrays on the HIV-1 reservoir size and long-term dynamics, we performed exome sequencing in 194 of the 797 study participants. Patients were selected at the two extremes of the observed

273 reservoir decay rate: either very rapid, or absent (no change in reservoir size over \sim 5.4 years), 274 while individuals with increasing HIV-1 reservoir sizes were excluded (N=12). Thus, the long-275 term dynamics phenotype was binarized for subsequent analyses of the decay rate, while the 276 HIV-1 reservoir size phenotype remained normally distributed (Figure S5).

To ensure that no common variants, missed by the genotype chips, were associated with the HIV-1 reservoir size or long-term dynamics, we performed a GWAS for common variants using age and sex as covariates. As with the genotyping data, we observed no genome-wide significant variants for either phenotype (Figure S6).

We then examined the potential role of rare variants (MAF < 5%) with a functional impact defined as either missense, frameshift, stop gained, splice acceptor or donor. Since HIV-1 primarily infects CD4+ T cells, we only included variants within genes expressed in these cells as determined by Gutierrez-Arcelus *et al.* [61]. The significance threshold after correcting for the number of tests performed was P = 1.21e-5. We did not observe any significant associations for either the HIV-1 reservoir or the decay rate. The *AMBRA1* gene showed the strongest association with HIV-1 reservoir size (P = 4.15e-5, not significant) (Figure S7).

To confirm the lack of HLA association seen with the genotyping data, we imputed the HLA haplotypes from the exome data using HLA*LA. Again, we did not observe any significant HLA association with the study outcomes.

291

292 Copy number variations

To examine the role of large exonic CNVs not captured by standard genotyping and exome pipelines, we called CNVs from the mapped sequencing reads of the exome samples using the software CLAMMS. The contribution of common CNVs to HIV-1 reservoir size and long-term dynamics was analyzed by association analyses including age, sex and the first principal component as covariates. No significant association was observed after Bonferroni correction (Figure S8). We also searched for rare CNVs in curated immune-related genes from Immport[59] but did not discover any suggestive immune-related CNVs .

300

301 Influence of HLA-B*57:01 and the CCR5∆32 deletion on reservoir size and long-term 302 dynamics

We have previously shown that pre-ART RNA viral load levels are associated with the HIV-1 reservoir size and the occurrence of blips [28]. The HLA-B*57:01 allele and the CCR5 Δ 32 deletion are well known genetic variants influencing HIV-1 spVL [50,51], and could thus also be associated with the with the HIV-1 reservoir size or its decay rate. However, we did not observe any nominal association (all P > 0.05) with either reservoir size or its long-term dynamics for HLA-B*57:01 and CCR5 Δ 32 (Figure S9).

309

310

311 **DISCUSSION**

312 We used a combination of genomic technologies to assess the potential role of human genetic 313 factors in determining both the HIV-1 reservoir size and its long-term dynamics in a well-314 characterized, population-based cohort. We studied 797 HIV-1 infected individuals of 315 European origin under suppressive ART over a median of 5.4 years, for whom extensive clinical 316 data are available, allowing detailed characterization and correction for potential confounders 317 [28]. We measured the HIV-1 reservoir size at three time points and selected two phenotypes 318 for our genomic study: the reservoir size at ~1.5 years after ART initiation and the slope of the 319 reservoir decay rate calculated over the three time points. Previous HIV host genetic studies 320 mostly focused on phenotypes reflecting the natural history of HIV-1 infection, prior to ART 321 initiation, including spontaneous viral control and disease progression [30–37]. A single study 322 specifically tested for associations between common genetic variants and the amount of

intracellular HIV-1 DNA, measured at a single time point during the chronic phase of infection
prior to initiation of any antiretroviral therapy [38]. Here, in contrast, we longitudinally assessed
samples collected from patients under suppressive ART to search for human genetic
determinants of the long-term dynamics of the HIV-1 reservoir during treatment.

327 We first conducted a GWAS on 797 individuals to test for association between common genetic 328 variants and the phenotypes. Given the small proportion of non-European subjects in the initial 329 study cohort, we only included patients of European ancestry to avoid any false positive 330 associations or masking of true positive associations due to different allele frequencies in small 331 proportions of individuals belonging to different subpopulations (Figure 1) [62]. Regardless of 332 including or not independent covariates other than the standard ones (i.e., sex and age), no 333 genetic variant reached the genome-wide significance threshold for association with any of the 334 two phenotypes. This may reflect a small effect size of genetic variants on the HIV-1 reservoir 335 size and decay rate. We acknowledge that a larger sample size and thus increased statistical 336 power may allow detecting genetic variants with a smaller effect size associated with the 337 phenotypes. However, it should be noted that this study is by far the largest today that has 338 investigated the size and decay of the HIV-reservoir in well characterized and well suppressed 339 HIV-infected individuals over a longer time period. Alternatively, the control of the HIV-1 340 reservoir size and its long-term dynamics may be under the control of viral or host factors other 341 than the individual genetic background. A previous report from our group had shown a 342 correlation between viral blips during the first 1.5 years of suppressive ART and the HIV-1 343 reservoir size 1.5 years after ART initiation, and between viral blips after 1.5-5.4 years of 344 suppressive ART or low-level viremia and a slower decay rate [28]. Importantly, viral blips are 345 generally thought to reflect transient increases in viral replication, probably occur under 346 multifactorial influence from viral and host factors [63-70], with these latter possibly including, 347 but not being limited to, germline genetic variation. The biological relations between viral

reservoir, decay rate, viral blips, and the contribution of the individual genetic background stillneed full elucidation.

350 Standard GWAS is designed to detect associations with common genetic variants (i.e., with a 351 MAF of at least 0.05), with little power to investigate the role of rare variants. Thus, to further 352 assess the contribution of rare variants in individuals at the extreme of the decay rate 353 distribution, we used exome sequencing in a selected subset of 194 study participants with very 354 high decay rate, or conversely, a stable reservoir size over time (Figure S5). Here again, our 355 analyses did not detect significant association with the phenotypes. Although not reaching 356 statistical significance, a rare genetic variant with potential functional impact in AMBRA1 had 357 a p-value for association just below the corrected threshold. The expression of AMBRA1, a core 358 component of the autophagy machinery, has previously been associated with long-term viral 359 control in HIV-1 non-progressors [71]. Future studies may further elucidate whether genetic 360 variation in AMBRA1 may account for inter-individual differences in the long-term dynamics 361 of the HIV-1 reservoir.

Large deletions or duplications of genomic material may be implicated in human phenotypes, with CNVs impacting the exonic regions being more likely to have a functional role. Thus, we further investigated whether any common or rare CNV spanning exonic regions was associated with the phenotype. Again, no CNV was statistically associated with the phenotypes both in the exome-wide analyses and in analyses focused on immune-related genes.

An inherent limitation of our exome-based association analyses was their inability to detect rare variants outside the coding or splice-site regions. The exonic regions account for approximately 1-2% of the whole human genome. Because many regulatory sequences are located in extragenic sites, our analysis did not fully investigate the role of highly conserved, non-coding genetic regions in influencing the phenotypes linked to HIV-1 latency.

372 Additionally, we focused on specific genetic variants, i.e., the HLA haplotypes and the 373 CCR5 Δ 32 deletion, previously demonstrated to have a role in HIV-1 related phenotypes 374 [30,51]. Indeed, previous studies unraveled a robust association between variation in the HLA 375 region and the HIV-1 spVL [30]. Likewise, heterozygosity for the CCR5 Δ 32 deletion has been 376 shown to influence spontaneous HIV-1 control [51]. Thus, we imputed HLA genotypes from 377 genotyping and exome data, and studied the CCR5 Δ 32 deletion, without, however, detecting 378 any significant associations with the phenotypes or the covariates (Figure S9). Specifically, we 379 found no correlation between HLA genotypes and HIV-1 RNA plasma levels prior to ART 380 initiation, apparently contrasting with the previous findings of an association between HLA-381 B*57:01 haplotype and spVL. This probably reflects historical changes in the therapeutic 382 approach following a diagnosis of HIV-1 infection, given that ART is currently initiated soon 383 after clinical diagnosis, before most patients reach a stable plateau of plasma viral load.

384 In our study, the quantification of the reservoir size at different time points may have been 385 influenced by factors as, for example, blips and low-level viremia, which may have reduced 386 our ability to detect significant genetic effects. It is also possible that, in the future, novel 387 methods to assess the viral reservoir will allow the detection of significant contributions of 388 genetic factors [72]. So far, it remains unanswered whether the initial response to acute 389 infection, the containment of ongoing replication, and the control of latently infected cells are 390 under the influence of the same or different molecular networks. It needs to be noted that in 391 previous work we have shown that host genetic factors as defined by GWAS did not explain 392 the severity of symptoms during acute HIV-infection, although severity of symptoms correlated 393 well with viral load and CD4 cell counts [73]. In contrast to host genetic factors, we recently 394 identified a significant impact of the viral genome on the size of the latent reservoir 1.5 years 395 after treatment initiation. However, what parts of the virus genome were responsible for this "viral effect" needs to be determined further. 396

In conclusion, our study suggests that human individual germline genetic variation has little, if any, influence on the control of the HIV-1 viral reservoir size and its long-term dynamics. Complex, likely multifactorial biological processes govern HIV-1 viral persistence. Larger studies will possibly clarify the role of common or rare genetic variants explaining small proportions of the variability of the phenotypes related to viral latency.

402

403 **Data availability**

The datasets generated during and/or analyzed during the current study are not publicly available due to privacy reasons, the sensitivities associated with HIV infections, and the representativeness of the dataset, but is available on request.

407

408 **Conflicts of interest statement**

409 H.F.G. has received unrestricted research grants from Gilead Sciences and Roche; fees for data
410 and safety monitoring board membership from Merck; consulting/advisory board membership

411 fees from Gilead Sciences, ViiV, Merck, Sandoz and Mepha.

412 T.K. has received consulting/advisory board membership fees from Gilead Sciences and from

413 ViiV Healthcare for work that has no connection to the work presented here.

414 K.J.M. has received travel grants and honoraria from Gilead Sciences, Roche Diagnostics,

415 GlaxoSmithKline, Merck Sharp & Dohme, Bristol-Myers Squibb, ViiV and Abbott; and the

416 University of Zurich received research grants from Gilead Science, Roche, and Merck Sharp &

417 Dohme for studies that Dr. Metzner serves as principal investigator, and advisory board

418 honoraria from Gilead Sciences.

419 All other authors declare no competing financial interests.

420

421 Author contributions

422	N.B., J.B., V.R., R.D.K., H.F.G., K.J.M., and J.F. contributed to the conception and design of
423	the study. C.v.S., V.V., K.N., Y.I.K., and K.J.M. contributed to the acquisition of data. A.B.,
424	C.W.T., N.B., T.T., S.P., M.W., R.D.K., and J.F. contributed to the analysis and interpretation
425	of data. J.B., M.P., T.K., S.Y., M.B., A.R., P.S., E.B., M.C., H.F.G., and the members of the
426	Swiss HIV Cohort Study (SHCS) conceived and managed the cohort, collected and contributed
427	patient samples and clinical data. A.B., C.W.T., and J.F. contributed to the drafting the article.
428	All authors read and approved the final manuscript.

429

430 Acknowledgements

431 We thank the patients for participating in the SHCS, the study nurses and physicians for 432 excellent patient care, A. Scherrer, A. Traytel, and S. Wild for excellent data management and 433 D. Perraudin and M. Amstutz for administrative assistance.

434

This work was funded within the framework of the Swiss HIV Cohort Study (SNF grant# 435 436 33CS30 177499 to H.F.G.). The data were gathered by the Five Swiss University Hospitals, 437 two Cantonal Hospitals, 15 affiliated hospitals and 36 private physicians (listed in 438 http://www.shcs.ch/180-health-care-providers). The work was furthermore supported by the 439 Systems.X grant # 51MRP0 158328 (to N.B., J.B., J.F., V.R., R.D.K., M.W., H.F.G. and K.J.M.), by SNF grant 324730B 179571 (to H.F.G.), SNF grant SNF 310030 141067/1 (to 440 441 H.F.G. and K.J.M.), SNF grants no. PZ00P3-142411 and BSSGI0 155851 to R.D.K., the 442 Yvonne-Jacob Foundation (to H.F.G.), the University of Zurich's Clinical Research Priority 443 Program viral infectious disease, ZPHI (to H.F.G) and the Vontobel Foundation (to H.F.G. and 444 K.J.M.). M.W. is partially supported by the NCCR MARVEL, funded by the Swiss National Science Foundation. 445

446

447 Members of the Swiss HIV Cohort Study:

- 448 Anagnostopoulos A, Battegay M, Bernasconi E, Böni J, Braun DL, Bucher HC, Calmy A,
- 449 Cavassini M, Ciuffi A, Dollenmaier G, Egger M, Elzi L, Fehr J, Fellay J, Furrer H, Fux CA,
- 450 Günthard HF (President of the SHCS), Haerry D (deputy of "Positive Council"), Hasse B,
- 451 Hirsch HH, Hoffmann M, Hösli I, Huber M, Kahlert CR (Chairman of the Mother & Child
- 452 Substudy), Kaiser L, Keiser O, Klimkait T, Kouyos RD, Kovari H, Ledergerber B, Martinetti
- 453 G, Martinez de Tejada B, Marzolini C, Metzner KJ, Müller N, Nicca D, Paioni P, Pantaleo G,
- 454 Perreau M, Rauch A (Chairman of the Scientific Board), Rudin C, Scherrer AU (Head of Data
- 455 Centre), Schmid P, Speck R, Stöckle M (Chairman of the Clinical and Laboratory
- 456 Committee), Tarr P, Trkola A, Vernazza P, Wandeler G, Weber R, Yerly S.
- 457
- 458

459 **REFERENCES**

- Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, Chaisson RE, et al.
 Identification of a Reservoir for HIV-1 in Patients on Highly Active Antiretroviral
 Therapy. Science. 1997 Nov 14;278(5341):1295–300.
- 463 2. Wong JK, Hezareh M, Günthard HF, Havlir DV, Ignacio CC, Spina CA, et al. Recovery of
 464 Replication-Competent HIV Despite Prolonged Suppression of Plasma Viremia.
 465 Science. 1997 Nov 14;278(5341):1291–5.
- 3. Chun T-W, Stuyver L, Mizell SB, Ehler LA, Mican JAM, Baseler M, et al. Presence of an
 inducible HIV-1 latent reservoir during highly active antiretroviral therapy. Proc Natl
 Acad Sci. 1997 Nov 25;94(24):13193–7.
- 469 4. Siliciano JD, Kajdas J, Finzi D, Quinn TC, Chadwick K, Margolick JB, et al. Long-term
 470 follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4⁺
 471 T cells. Nat Med. 2003 Jun;9(6):727–8.
- 5. Chun TW, Carruth L, Finzi D, Shen X, DiGiuseppe JA, Taylor H, et al. Quantification of
 latent tissue reservoirs and total body viral load in HIV-1 infection. Nature. 1997 May
 8;387(6629):183–8.
- 6. Chun TW, Engel D, Berrey MM, Shea T, Corey L, Fauci AS. Early establishment of a pool of latently infected, resting CD4(+) T cells during primary HIV-1 infection. Proc Natl Acad Sci U S A. 1998 Jul 21;95(15):8869–73.

- 478 7. Smith MZ, Wightman F, Lewin SR. HIV reservoirs and strategies for eradication. Curr
 479 HIV/AIDS Rep. 2012 Mar;9(1):5–15.
- 480 8. Siliciano RF, Greene WC. HIV latency. Cold Spring Harb Perspect Med. 2011
 481 Sep;1(1):a007096.
- 482 9. Ruelas DS, Greene WC. An integrated overview of HIV-1 latency. Cell. 2013 Oct
 483 24;155(3):519–29.
- 484 10. Han Y, Wind-Rotolo M, Yang H-C, Siliciano JD, Siliciano RF. Experimental
 485 approaches to the study of HIV-1 latency. Nat Rev Microbiol. 2007 Feb;5(2):95–106.
- 486 11. Hodel F, Patxot M, Snäkä T, Ciuffi A. HIV-1 latent reservoir: size matters. Future Virol.
 487 2016 Dec;11(12):785–94.
- 488 12. Avettand-Fènoël V, Hocqueloux L, Ghosn J, Cheret A, Frange P, Melard A, et al. Total
 489 HIV-1 DNA, a Marker of Viral Reservoir Dynamics with Clinical Implications. Clin
 490 Microbiol Rev. 2016 Oct;29(4):859–80.
- 491 13. Bruner KM, Wang Z, Simonetti FR, Bender AM, Kwon KJ, Sengupta S, et al. A
 492 quantitative approach for measuring the reservoir of latent HIV-1 proviruses. Nature.
 493 2019;566(7742):120–5.
- 494 14. Kiselinova M, De Spiegelaere W, Buzon MJ, Malatinkova E, Lichterfeld M,
 495 Vandekerckhove L. Integrated and Total HIV-1 DNA Predict Ex Vivo Viral Outgrowth.
 496 PLoS Pathog. 2016 Mar;12(3):e1005472.
- 497 15. Williams JP, Hurst J, Stöhr W, Robinson N, Brown H, Fisher M, et al. HIV-1 DNA
 498 predicts disease progression and post-treatment virological control. eLife. 2014 Sep
 499 12;3:e03821.
- 500 16. Koelsch KK, Liu L, Haubrich R, May S, Havlir D, Günthard HF, et al. Dynamics of
 501 total, linear nonintegrated, and integrated HIV-1 DNA in vivo and in vitro. J Infect Dis.
 502 2008 Feb 1;197(3):411–9.
- 503 17. Gandhi RT, McMahon DK, Bosch RJ, Lalama CM, Cyktor JC, Macatangay BJ, et al.
 504 Levels of HIV-1 persistence on antiretroviral therapy are not associated with markers of
 505 inflammation or activation. PLOS Pathog. 2017 Apr 20;13(4):e1006285.
- I8. Zanchetta M, Walker S, Burighel N, Bellanova D, Rampon O, Giaquinto C, et al. Longterm decay of the HIV-1 reservoir in HIV-1-infected children treated with highly active
 antiretroviral therapy. J Infect Dis. 2006 Jun 15;193(12):1718–27.
- 509 19. Chun T-W, Justement JS, Moir S, Hallahan CW, Maenza J, Mullins JI, et al. Decay of
 510 the HIV reservoir in patients receiving antiretroviral therapy for extended periods:
 511 implications for eradication of virus. J Infect Dis. 2007 Jun 15;195(12):1762–4.
- 512 20. Strain MC, Günthard HF, Havlir DV, Ignacio CC, Smith DM, Leigh-Brown AJ, et al.
 513 Heterogeneous clearance rates of long-lived lymphocytes infected with HIV: intrinsic
 514 stability predicts lifelong persistence. Proc Natl Acad Sci U S A. 2003 Apr
 515 15;100(8):4819–24.

- 516 21. Izopet J, Salama G, Pasquier C, Sandres K, Marchou B, Massip P, et al. Decay of HIV-1
 517 DNA in patients receiving suppressive antiretroviral therapy. J Acquir Immune Defic
 518 Syndr Hum Retrovirology Off Publ Int Retrovirology Assoc. 1998 Dec 15;19(5):478–
 519 83.
- Ramratnam B, Mittler JE, Zhang L, Boden D, Hurley A, Fang F, et al. The decay of the
 latent reservoir of replication-competent HIV-1 is inversely correlated with the extent of
 residual viral replication during prolonged anti-retroviral therapy. Nat Med. 2000
 Jan;6(1):82–5.
- van Rij RP, van Praag RME, Prins JM, Rientsma R, Jurriaans S, Lange JMA, et al.
 Persistence of viral HLA-DR- CD4 T-cell reservoir during prolonged treatment of HIV1 infection with a five-drug regimen. Antivir Ther. 2002 Mar;7(1):37–41.
- 527 24. Pires A, Hardy G, Gazzard B, Gotch F, Imami N. Initiation of antiretroviral therapy
 528 during recent HIV-1 infection results in lower residual viral reservoirs. J Acquir Immune
 529 Defic Syndr 1999. 2004 Jul 1;36(3):783–90.
- 530 25. Strain MC, Little SJ, Daar ES, Havlir DV, Gunthard HF, Lam RY, et al. Effect of
 531 treatment, during primary infection, on establishment and clearance of cellular reservoirs
 532 of HIV-1. J Infect Dis. 2005 May 1;191(9):1410–8.
- 533 26. Fischer M, Joos B, Niederöst B, Kaiser P, Hafner R, von Wyl V, et al. Biphasic decay
 534 kinetics suggest progressive slowing in turnover of latently HIV-1 infected cells during
 535 antiretroviral therapy. Retrovirology. 2008 Nov 26;5:107.
- 536 27. Besson GJ, Lalama CM, Bosch RJ, Gandhi RT, Bedison MA, Aga E, et al. HIV-1 DNA
 537 decay dynamics in blood during more than a decade of suppressive antiretroviral
 538 therapy. Clin Infect Dis Off Publ Infect Dis Soc Am. 2014 Nov;59(9):1312–21.
- Bachmann N, Siebenthal C von, 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 Jul 19;10(1):1–11.
- 542 29. International AIDS Society Scientific Working Group on HIV Cure, Deeks SG, Autran
 543 B, Berkhout B, Benkirane M, Cairns S, et al. Towards an HIV cure: a global scientific
 544 strategy. Nat Rev Immunol. 2012 20;12(8):607–14.
- 545 30. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, et al. A whole546 genome association study of major determinants for host control of HIV-1. Science.
 547 2007 Aug 17;317(5840):944–7.
- 548 31. Fellay J, Ge D, Shianna KV, Colombo S, Ledergerber B, Cirulli ET, et al. Common
 549 genetic variation and the control of HIV-1 in humans. PLoS Genet. 2009
 550 Dec;5(12):e1000791.
- 32. 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 Apr
 15;201(8):1141–9.

- 33. International HIV Controllers Study, Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker
 PIW, et al. The major genetic determinants of HIV-1 control affect HLA class I peptide
 presentation. Science. 2010 Dec 10;330(6010):1551–7.
- 34. 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 Feb 15;201(4):618–26.
- 560 35. Le Clerc S, Limou S, Coulonges C, Carpentier W, Dina C, Taing L, et al. Genomewide
 561 association study of a rapid progression cohort identifies new susceptibility alleles for
 562 AIDS (ANRS Genomewide Association Study 03). J Infect Dis. 2009 Oct
 563 15;200(8):1194–201.
- 564 36. Limou S, Le Clerc S, Coulonges C, Carpentier W, Dina C, Delaneau O, et al.
 565 Genomewide association study of an AIDS-nonprogression cohort emphasizes the role
 566 played by HLA genes (ANRS Genomewide Association Study 02). J Infect Dis. 2009
 567 Feb 1;199(3):419–26.
- 568 37. Limou S, Coulonges C, Herbeck JT, van Manen D, An P, Le Clerc S, et al. Multiple569 cohort genetic association study reveals CXCR6 as a new chemokine receptor involved
 570 in long-term nonprogression to AIDS. J Infect Dis. 2010 Sep 15;202(6):908–15.
- 38. Dalmasso C, Carpentier W, Meyer L, Rouzioux C, Goujard C, Chaix M-L, 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(12):e3907.
- McLaren PJ, Pulit SL, Gurdasani D, Bartha I, Shea PR, Pomilla C, et al. Evaluating the
 Impact of Functional Genetic Variation on HIV-1 Control. J Infect Dis. 2017
 27;216(9):1063–9.
- 577 40. Schoeni-Affolter F, Ledergerber B, Rickenbach M, Rudin C, Günthard HF, Telenti A, et
 578 al. Cohort Profile: The Swiss HIV Cohort Study. Int J Epidemiol. 2010 Oct
 579 1;39(5):1179–89.
- 580 41. Loh P-R, Danecek P, Palamara PF, Fuchsberger C, Reshef YA, Finucane HK, et al.
 581 Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet.
 582 2016 Nov;48(11):1443–8.
- 583 42. Durbin R. Efficient haplotype matching and storage using the positional Burrows–
 584 Wheeler transform (PBWT). Bioinformatics. 2014 May 1;30(9):1266–72.
- 43. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. A
 reference panel of 64,976 haplotypes for genotype imputation. Nat Genet [Internet].
 2016 Aug 22 [cited 2016 Aug 31];advance online publication. Available from: http://www.nature.com/ng/journal/vaop/ncurrent/full/ng.3643.html
- 589 44. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal
 590 components analysis corrects for stratification in genome-wide association studies. Nat
 591 Genet. 2006 Aug;38(8):904–9.
- 592 45. The International HapMap 3 Consortium. Integrating common and rare genetic variation
 593 in diverse human populations. Nature. 2010 Sep;467(7311):52–8.

- 594 46. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation
 595 PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015 Dec
 596 1;4(1):1–16.
- 47. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A Tool for Genome-wide Complex
 598 Trait Analysis. Am J Hum Genet. 2011 Jan 7;88(1):76–82.
- 48. Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. Advantages and pitfalls in
 the application of mixed-model association methods. Nat Genet. 2014 Feb;46(2):100–6.
- 49. Jia X, Han B, Onengut-Gumuscu S, Chen W-M, Concannon PJ, Rich SS, et al. Imputing
 Amino Acid Polymorphisms in Human Leukocyte Antigens. PLOS ONE. 2013 Jun
 603 6;8(6):e64683.
- 50. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, et al. A WholeGenome Association Study of Major Determinants for Host Control of HIV-1. Science.
 2007 Aug 17;317(5840):944–7.
- McLaren PJ, Coulonges C, Bartha I, Lenz TL, Deutsch AJ, Bashirova A, 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. 2015 Nov 24;112(47):14658–63.
- 610 52. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler
 611 transform. Bioinforma Oxf Engl. 2009 Jul 15;25(14):1754–60.
- 53. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The
 Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation
 DNA sequencing data. Genome Res. 2010 Sep 1;20(9):1297–303.
- 615 54. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for
 616 annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs
 617 in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012
 618 Jun;6(2):80–92.
- 55. Lee S, Emond MJ, Bamshad MJ, Barnes KC, Rieder MJ, Nickerson DA, et al. Optimal
 unified approach for rare-variant association testing with application to small-sample
 case-control whole-exome sequencing studies. Am J Hum Genet. 2012 Aug
 10;91(2):224–37.
- 56. Dilthey AT, Gourraud P-A, Mentzer AJ, Cereb N, Iqbal Z, McVean G. High-Accuracy
 HLA Type Inference from Whole-Genome Sequencing Data Using Population
 Reference Graphs. PLOS Comput Biol. 2016 Oct 28;12(10):e1005151.
- 626 57. Fan Y, Song Y-Q. PyHLA: tests for the association between HLA alleles and diseases.
 627 BMC Bioinformatics. 2017 Feb 6;18(1):90.
- 58. Packer JS, Maxwell EK, O'Dushlaine C, Lopez AE, Dewey FE, Chernomorsky R, et al.
 629 CLAMMS: a scalable algorithm for calling common and rare copy number variants from
 630 exome sequencing data. Bioinformatics. 2016 Jan 1;32(1):133–5.

- 631 59. Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, et al. ImmPort,
 632 toward repurposing of open access immunological assay data for translational and
 633 clinical research. Sci Data. 2018 27;5:180015.
- 634 60. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and
 635 association genetic mapping studies of complex traits. Bioinformatics. 2003 Jan
 636 1;19(1):149–50.
- 637 61. Gutierrez-Arcelus M, Teslovich N, Mola AR, Polidoro RB, Nathan A, Kim H, et al.
 638 Lymphocyte innateness defined by transcriptional states reflects a balance between
 639 proliferation and effector functions. Nat Commun. 2019 Feb 8;10(1):1–15.
- 640 62. Marees AT, de Kluiver H, Stringer S, Vorspan F, Curis E, Marie-Claire C, et al. A
 641 tutorial on conducting genome-wide association studies: Quality control and statistical
 642 analysis. Int J Methods Psychiatr Res. 2018;27(2):e1608.
- 643 63. Jones LE, Perelson AS. Transient viremia, plasma viral load, and reservoir
 644 replenishment in HIV-infected patients on antiretroviral therapy. J Acquir Immune Defic
 645 Syndr 1999. 2007 Aug 15;45(5):483–93.
- 646 64. Fletcher CV, Staskus K, Wietgrefe SW, Rothenberger M, Reilly C, Chipman JG, et al.
 647 Persistent HIV-1 replication is associated with lower antiretroviral drug concentrations
 648 in lymphatic tissues. Proc Natl Acad Sci U S A. 2014 Feb 11;111(6):2307–12.
- 649 65. Lorenzo-Redondo R, Fryer HR, Bedford T, Kim E-Y, Archer J, Pond SLK, et al.
 650 Persistent HIV-1 replication maintains the tissue reservoir during therapy. Nature. 2016
 651 Feb 4;530(7588):51–6.
- 66. Podsadecki TJ, Vrijens BC, Tousset EP, Rode RA, Hanna GJ. Decreased adherence to
 antiretroviral therapy observed prior to transient human immunodeficiency virus type 1
 viremia. J Infect Dis. 2007 Dec 15;196(12):1773–8.
- 655 67. Young J, Rickenbach M, Calmy A, Bernasconi E, Staehelin C, Schmid P, et al. Transient
 656 detectable viremia and the risk of viral rebound in patients from the Swiss HIV Cohort
 657 Study. BMC Infect Dis. 2015 Sep 21;15:382.
- 658 68. Simonetti FR, Sobolewski MD, Fyne E, Shao W, Spindler J, Hattori J, et al. Clonally
 659 expanded CD4+ T cells can produce infectious HIV-1 in vivo. Proc Natl Acad Sci U S
 660 A. 2016 Feb 16;113(7):1883–8.
- 661 69. Wang Z, Gurule EE, Brennan TP, Gerold JM, Kwon KJ, Hosmane NN, et al. Expanded
 662 cellular clones carrying replication-competent HIV-1 persist, wax, and wane. Proc Natl
 663 Acad Sci U S A. 2018 13;115(11):E2575–84.
- 664 70. Lee GQ, Orlova-Fink N, Einkauf K, Chowdhury FZ, Sun X, Harrington S, et al. Clonal
 665 expansion of genome-intact HIV-1 in functionally polarized Th1 CD4+ T cells. J Clin
 666 Invest. 2017 Jun 30;127(7):2689–96.
- 667 71. Nardacci R, Amendola A, Ciccosanti F, Corazzari M, Esposito V, Vlassi C, et al.
 668 Autophagy plays an important role in the containment of HIV-1 in nonprogressor669 infected patients. Autophagy. 2014 Jul;10(7):1167–78.

670	72.	Gaebler C, Lorenzi JCC, Oliveira TY, Nogueira L, Ramos V, Lu C-L, et al.
671		Combination of quadruplex qPCR and next-generation sequencing for qualitative and
672		quantitative analysis of the HIV-1 latent reservoir. J Exp Med. 2019 Oct
673		7;216(10):2253–64.

674 73. Braun DL, Kouyos R, Oberle C, Grube C, Joos B, Fellay J, et al. A novel Acute
675 Retroviral Syndrome Severity Score predicts the key surrogate markers for HIV-1
676 disease progression. PloS One. 2014;9(12):e114111.

677

678

679 FIGURE LEGENDS

- 680 Figure 1. Patient selection flowchart. Specific inclusion and exclusion criteria are listed for
- each selection step. ART (antiretroviral therapy); PBMCs (peripheral blood mononuclear cells);
- 682 PI (protease inhibitor); PCA (principal component analysis).

683

- 684 Figure 2. Association results with HIV-1 reservoir size. Manhattan plot with association p-
- values (-log₁₀(P)) 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
- 687 significant.

688

- 689
- 690
- 691
- 692
- 693
- 694
- 695
- 696
- 697

698 **TABLES**

Table 1. Patient characteristics	
Total number of individuals	
Genotyped	797
Genotyped + exome sequenced	194
Age at first HIV-1 DNA sample in ye	ars
median (IQR)	44 (38, 50)
Sex	
Female	123 (15.4%)
Male	674 (84.6%)
Transmission group	
HET	241 (30.2%)
IDU	77 (9.7%)
MSM	448 (56.2%)
Other	31 (3.9%)
HIV-1 subtype	
В	550 (69.0%)
Non-B	128 (16.1%)
Unknown	119 (14.9%)
Occurrence of blips or low-level vire	mia
Blips	200 (25.1%)
Low-level viremia	68 (8.5%)
None	529 (66.4%)
Time on ART	
median (IQR)	1.50 (1.28, 1.69)
Infection stage	
Acute	140 (17.6%)
Chronic	657 (82.4%)
Time to viral suppression	
median (IQR)	0.34 (0.23, 0.51)
ا Log10 HIV-1 plasma RNA pre-ART	per mL
median (IQR)	480 (248, 684)
CD4+ cell count pre-ART cells/µL bl	lood
median (IQR)	186 (90, 270)
HIV-1 reservoir size	
median (IQR)	2.76 (2.48, 3.03)
HIV-1 reservoir decay rate	
median (IOR)	-0.06 (-0.12, -0.00)

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 \geq

- 702 50 HIV-1 RNA copies/mL plasma within a 30-day window. Individuals with consecutive
- 703 measurements of \geq 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. HIV-1 reservoir size
- 706 was measured in log₁₀ total HIV-1 DNA/1 million genomic equivalents ~1.5 years after
- 707 initiating ART. The HIV-1 reservoir decay rate was based on the three measurements of total
- 708 HIV-1 DNA levels taken at the median of 1.5, 3.5 and 5.4 years after initiation of ART.

18,688 individuals (Swiss HIV Cohort Study,	19/12/2014)			
Inclusion criteria: • HIV-1 infection • Receiving ART for ≥ 5 years • No treatment interruption of > 7 days • No virological failure defined as two comeasurements > 200 HIV-1 RNA copie • PBMCs available for three time points: 1.5 +/- 0.5 years, 3.5 +/- 0.5 years, and 5.5 +/- 1 year, after initiating first-line A	onsecutive ss/ml plasma d RT			
1932 individuals				
Exclusion criteria: • Start on less potent ART regimens, i.e. therapy, less potent/unboosted PI (NF\	Exclusion criteria: • Start on less potent ART regimens, i.e., mono/ or dual therapy, less potent/unboosted PI (NFV, SQV etc.)			
1382 individuals				
Received ≥ 3 PBMC samples (mandated)	Received ≥ 3 PBMC samples (mandatory 1 st - 3 rd time point)			
1166 individuals				
Successful total HIV-1 DNA quantification in ≥ 3 PBMC samples (mandatory 1 st - 3 rd time point)				
1057 individuals				
Successfully genotyped				
995 individuals				
European ancestry (by PCA)				
797 individuals + Exome sequencing 1	194 individuals			

