HIV

# Sex differences in HIV-1 reservoir cell selection are linked to altered innate immune profiles

**Toong Seng Tan[1,2], Ce Gao[1,2], Alexander S. Hochroth[1,2], Liliana Vela[1,2], Leah Carrere[1,2], Sruthi Kalavacherla[1,2], Seohyun Hong[1], Melanie Lancien[1], Chloe M. Naasz[1,2], Aischa Niesar[1], Samantha K. Marzi[1,2], Isabelle C. Roseto[1,2], Benjamin Bone[1,2], Xiaodong Lian[1,2], Yuko Yuki[3], Mathias Viard[3], Rebecca Hoh[4], Deborah K. McMahon[5], Ronald J. Bosch[6], Seble G. Kassaye[7], Rajesh T. Gandhi[8], Mary Carrington[1,3], Steven G. Deeks[4], Phyllis C. Tien[9], Michael J. Peluso[4], Jeffrey M. Jacobson[10], Mathias Lichterfeld[1,2], Xu G. Yu[1,2]***



**HIV-1 persistence despite suppressive antiretroviral therapy (ART) is primarily because of infected memory CD4 T cells, so-called viral reservoir cells, that harbor chromosomally integrated viral DNA as a "provirus" and resist clearance by the human immune system. Biological sex affects host immune responses and may influence selection and evolution of HIV-1 reservoir cells during long-term ART for HIV infection. We assessed more than 4073 individual proviruses through single-molecule amplification from 30 females and 35 males living with HIV-1 and treated with ART for a median of 20 years. We observed that the HIV-1 reservoir profile in females was characterized by lower proviral phylogenetic complexity, an increased proportion of clonally expanded intact proviruses, and a higher proportion of intact proviruses integrated into repressive heterochromatin locations of the human genome. The evolution of this distinct viral reservoir profile in females was associated with an improved signature of innate immune responses, specifically those of NK cells. On the contrary, signs of viral sequence adaptation to adaptive T cell immune responses were more pronounced in intact HIV-1 proviruses from males. Collectively, these data suggest a stronger ability of the female immune system to drive immune selection of HIV-1 reservoir cells during ART, putatively because of improved innate immune function.**

## INTRODUCTION

There are currently no curative treatment options for HIV-1 infection; lifelong pharmacological suppression therapy therefore remains the only available treatment modality for the more than 39 million individuals living with HIV-1 worldwide (*1*). Indefinite viral persistence despite suppressive antiretroviral therapy (ART) is primarily due to infected memory CD4 T cells, frequently termed "viral reservoir cells," that harbor chromosomally integrated viral DNA as a "provirus" and resist human immune clearance. The ability of these cells to evade immune-mediated elimination has been attributed to low or absent viral gene transcription that can mask these cells from host immune recognition. This condition, typically referred to as "viral latency" or proviral "transcriptional silence," may be linked to the lack of specific host transcription factors required for effective viral gene expression in memory CD4 T cells (*2*). Nevertheless, recent studies have demonstrated that ongoing viral gene expression is detectable in a marked number of infected cells from individuals receiving suppressive ART (*3*) consistent with the fact that HIV-1 uses complex molecular trafficking mechanisms (*4*) to preferentially integrate into highly expressed genes in accessible euchromatin regions that can support viral transcription. After primary infection of CD4 T cells, ~80% of viral DNA sequences are integrated in such highly expressed genes that promote access to epigenetic host factors facilitating viral gene expression and that generally represent a chromatin environment conducive to viral transcriptional activity (*5*).

Rather than being inconsequential bystander effects of a "leaky reservoir" (*6*), increasing evidence suggests that ongoing viral transcriptional activity during ART, even when limited to small, abortive viral transcripts, can be functionally relevant for influencing the fate and persistence of viral reservoir cells. This view is most definitely supported by studies highlighting the selective long-term persistence of intact proviruses in heterochromatin and repressive chromatin locations, where residual viral gene expression is ineffective or actively suppressed (*2*, *7*, *8*). Proviruses with these specific integration site features, sometimes described as being "blocked and locked" or in "deep latency" (*9*), seem to have a longitudinal selection advantage and tend to accumulate and dominate the viral integration site landscape in persons with natural control of HIV (*7*, *8*). The preferential persistence of intact proviruses in heterochromatin regions can be most readily interpreted as consequences of immune-mediated selection effects during which specific components of the human immune system recognize and target HIV-1 reservoir cells with ongoing viral transcription, presumably because viral transcripts can be sensed by immune recognition mechanisms (*10*, *11*). In addition to proviral integration locations, such immune selection processes also seem to depend on proviral sequence characteristics. Many proviral sequences exhibit large deletions that result from error-prone reverse transcription (*12*); proviruses with such deletions or

[1]Ragon Institute of MGH, MIT and Harvard, Cambridge, MA 02139, USA. [2]Infectious Disease Division, Brigham and Women's Hospital, Boston, MA 02115, USA. [3]Basic Science Program, Frederick National Laboratory for Cancer Research, National Cancer Institute, Frederick, MD and Laboratory of Integrative Cancer Immunology, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892, USA. [4]Division of HIV, Infectious Diseases and Global Medicine, University of California San Francisco, San Francisco, CA 94143, USA. [5]University of Pittsburgh, Pittsburgh, PA 15260, USA. [6]Center for Biostatistics in AIDS Research, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA. [7]Department of Medicine, Georgetown University Medical Center, Washington, DC 20057, USA. [8]Infectious Disease Division, Massachusetts General Hospital, Boston, MA 02114, USA. [9]Department of Medicine, University of California San Francisco and Department of Veterans Affairs Medical Center, San Francisco, CA 94121, USA. [10]School of Medicine, Case Western Reserve University, Cleveland, OH 44106, USA.
*Corresponding author. Email: xyu@mgh.harvard.edu

other defects seem to be subject to very limited, if any, immune selection pressure, arguably because effective host immune recognition may require the presence of functionally and structurally complete viral transcripts (*8*, *10*). Together, these results suggest that host immune mechanisms can target and eliminate at least some of the viral reservoir cells in people living with HIV-1 and that the pool of reservoir cells surviving long-term during ART is specifically adjusted to evade host immune responses through distinct chromosomal locations in transcriptionally repressive chromatin (*13*) or through viral sequence deletions precluding viral immune recognition. The type of host immune responses that are most effective in recognizing, attacking, and selecting HIV-1 reservoir cells are now unknown but represent an active area of investigation.

Although females represent more than half of all people living with HIV-1 worldwide and, in certain geographic regions, strongly exceed the frequencies of male counterparts (*14*), relatively little is now known about immune effects on HIV-1 reservoir cells in females. Studies evaluating the selection and evolution of the HIV-1 reservoir cell pool involved, in the majority of cases, men who have sex with men; moreover, prospective clinical trials designed to target and reduce HIV reservoir cells also mostly recruited males (*15*, *16*). However, a series of prior studies clearly demonstrated that antiviral immunity is frequently influenced by sex and that immune mechanisms, specifically those involving innate immune recognition and function, are heavily affected by sex differences (*15*, *17*–*19*). Here, we evaluated HIV-1 reservoir profiles in a cohort of individuals comprising 30 females and 35 males who remained on long-term suppressive ART (LT-ART) for a median of 20 years. Our results suggest that immune selection of viral reservoir cells in heterochromatin locations is more effective in females, implying that sex-dependent immune mechanisms may play a critical role for recognizing and targeting HIV-1–infected cells during suppressive ART.

## RESULTS

### Proviral reservoir size and clonality in people living with HIV-1 on LT-ART

Previous studies demonstrated that viral reservoir cells can dynamically evolve over time in a process that seems to be at least partially influenced by host immune selection pressure (*13*). To understand this phenomenon on a larger population level and to dissect potential sex-based differences of this likely immune-driven viral reservoir transformation, we recruited a study cohort of 65 participants (30 females and 35 males, all cis-gender) who had been on LT-ART for a median of 20 (range, 15 to 24) consecutive years with no more than two recorded plasma viremia blips (<100 copies/ml) (table S1). We first quantified their HIV-1 reservoir cell pool using near full-length individual proviral sequencing (FLIP-seq). In total, we analyzed 4073 individual proviral genomes from LT-ART participants ($n = 191$ intact and $n = 3882$ defective proviruses). The viral reservoir profiles of people living with HIV-1 on ART for moderate durations (median of 8 years) (mART) (table S1), partially described in our previous work (*8*, *13*), were analyzed as comparison cohorts for the LT-ART study group. These investigations demonstrated that the frequencies of total ($P = 0.0451$) and intact ($P = 0.0062$) HIV-1 proviruses were lower in LT-ART persons compared with mART individuals (Fig. 1, A and B); however, there were no differences between the mART and LT-ART groups when comparing defective HIV-1 DNA sequences (Fig. 1C). The relative proportion of intact proviruses among all analyzed proviruses was significantly lower ($P < 0.0001$) in the LT-ART (4.69%) compared with the mART (7.38%) group (Fig. 1D and fig. S3C); these observations support the hypothesis that genome-intact proviruses are more susceptible to host immune effects and decline faster over time compared with defective proviruses (*20*). Given that longitudinal dynamics of HIV-1 proviruses can be heavily influenced by clonal proliferation of virally infected cells (*21*), we analyzed clusters of sequence-identical proviruses that result from clonal turnover of reservoir cells. The proportion of sequence-identical intact proviruses detected was higher in the LT-ART compared with the mART cohort (Fig. 1E), consistent with prior reports showing increasing proportions of clonally expanded proviruses over extended periods of viral suppression (*13*, *22*). Together, these results indicate that genome-intact proviruses can readily survive for more than 20 years of continuous antiviral therapy, partially because of clonal proliferation of infected cells; however, lower frequencies of intact proviruses after long-term ART, relative to moderate durations of ART, suggest continuous decay of genome-intact HIV-1 DNA during extended periods of ART.

### Sex differences in the HIV-1 reservoir landscape

Given that sex may influence immune selection of HIV-1 reservoir cells during suppressive LT-ART, we subsequently compared viral reservoir cells in male versus female study participants. Overall, there were no significant differences in terms of age, clinical characteristics, years on ART, years of documented viral suppression, distribution of protective human leukocyte antigen (HLA) alleles, or quantities of cells subjected to viral sequencing assays between female and male participants in our LT-ART cohort (table S2). Whereas the size of total, intact, and defective proviruses did not differ profoundly between females and males in each of the analyzed study cohorts (Fig. 2, A to C), we observed a notable number of LT-ART participants in whom, within the available number of peripheral blood mononuclear cell (PBMC) samples, we failed to detect intact proviruses (Figs. 1B and 2B), suggesting that in these study persons, frequencies of intact proviruses were low. Within a total of $n = 2573$ ($n = 94$ intact and $n = 2479$ defective) individual HIV-1 genomes detected in male LT-ART participants and a total of $n = 1500$ ($n = 97$ intact and $n = 1403$ defective) proviral sequences from female LT-ART study persons, we noted that the proportion of intact proviruses among all analyzed HIV-1 genomes was 1.8-fold higher in female (6.45%) compared with male participants (3.59%) (Fig. 2D), although this did not reach statistical significance. The frequencies (Fig. 2E) and the proportions (Fig. 2F) of intact clonal proviruses were higher in LT-ART females than in males, suggesting a more important role of clonal proliferation for maintaining the viral reservoir in females compared with males. Corresponding to this observation, we noted that the landscape of intact proviruses in females was more profoundly dominated by large clonal populations of intact proviruses, resulting in a more limited phylogenetic complexity within the pool of intact proviruses (Fig. 2, G and H). Overall, although failing to identify quantitative differences in frequencies of proviral sequences in reservoir cells between females and males, our data suggest sex-based qualitative differences in the structure and composition of the viral reservoir cell pool.

### Integration site profiles of intact proviruses

Cells harboring large clones of intact proviruses that persist and expand during LT-ART may proliferate because of selection advantages that protect them from immune recognition (*2*, *13*); therefore,

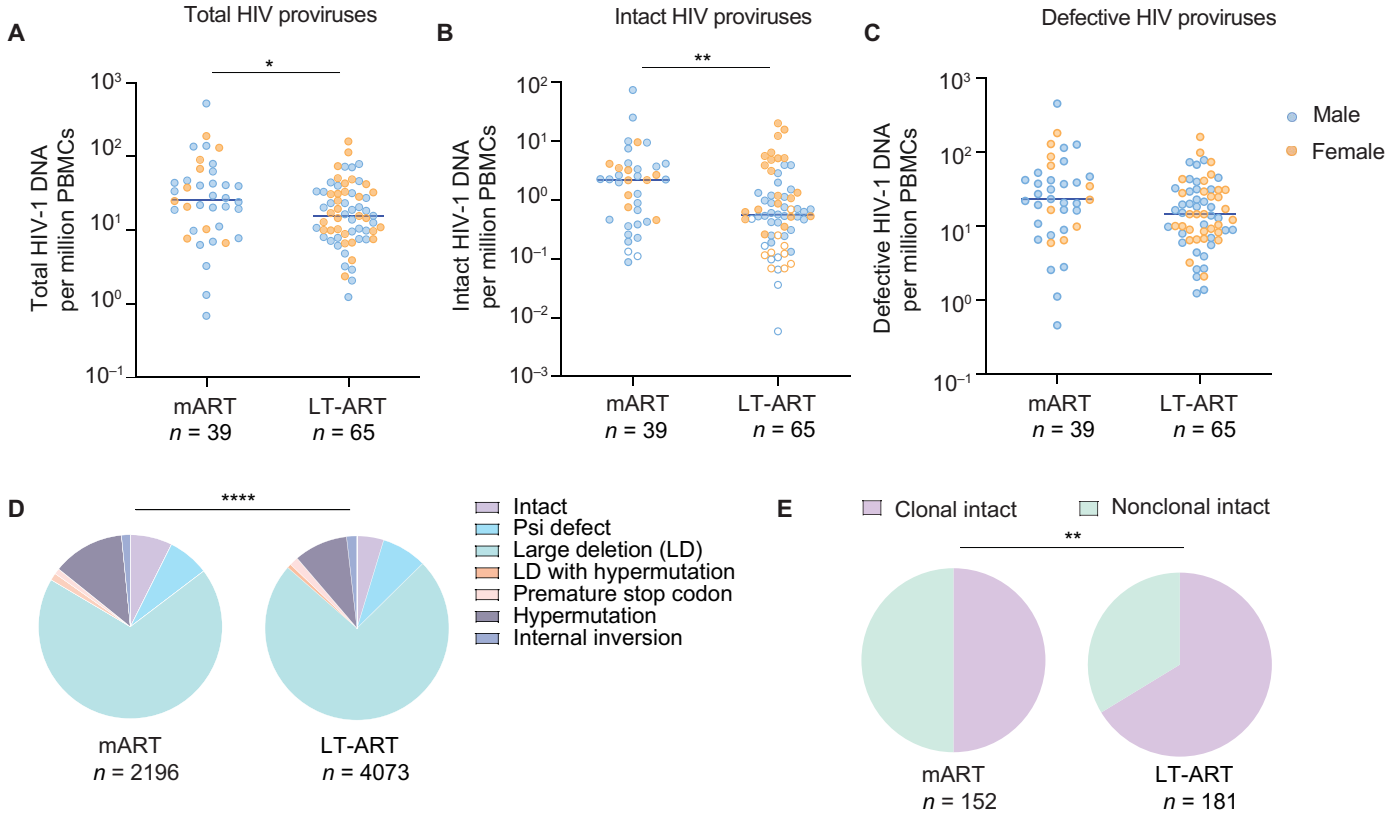


**Fig. 1. Proviral reservoir size and clonality in people living with HIV-1 on LT-ART.** (**A** to **C**) The frequencies of total (A), intact (B), and defective (C) proviruses in individuals on mART and individuals on LT-ART. Open circles represent data at the limit of detection. Horizontal lines reflect the median. (**D**) Proportions of intact or defective proviruses among all proviral genomes. Psi, packaging signal. (**E**) Proportions of clonal and nonclonal intact proviruses. Participants with at least two detectable genome-intact proviruses were included. Mann-Whitney *U* nonparametric tests were used in (A) to (C); two-tailed $\chi^2$ tests were used in (D) and (E). ****$P < 0.0001$, **$P < 0.01$, and *$P < 0.05$. The number of participants in (A) to (C) is represented by *n*; the number of viral sequences in (D) and (E) is represented by *n*.

these clones can be interpreted as a biomarker of enhanced immune selection pressure that HIV-1 reservoir cells are exposed to. To investigate the clonal structure of the reservoir cell pool in more detail, we profiled the chromosomal locations of intact proviruses, using matched integration site and proviral sequencing (MIP-seq). We assessed a total of 115 unique integration sites (50 from intact proviral DNA and 65 from defective proviral DNA) in our LT-ART cohort (Figs. 3 and 4). This detailed analysis demonstrated that in a large proportion (8 of 14, 57.1%) of female LT-ART study participants, integration sites of intact proviruses were strongly biased toward nongenic heterochromatin regions (Fig. 3). This was particularly evident in study participants LT-F02, LT-F03, LT-F06, and LT-F08, in all of whom we observed large clones of intact proviruses integrated in nongenic centromeric satellite or peri-centromeric microsatellite DNA, which typically consists of condensed heterochromatin with repressive epigenetic chromatin features (*23*). In three additional female study participants (LT-F04, LT-F05, and LT-F07), large clones of intact proviruses were integrated in *ZNF* genes located on chromosomes 4, 12, and 20; these *ZNF* genes are packed with inhibitory epigenetic histone modifications (*24*), including those deposited by the human silencing hub (HUSH) complex (*25*). One more female study participant (LT-F01) demonstrated a large genome-intact clone in a nongenic region on chromosome 12. In three female study participants (LT-F09, LT-F10, and LT-F11), we noted a more diverse integration site profile for intact proviruses: In LT-F09, a large intact provirus was detected within intronic microsatellite DNA of the *DLEC1* gene; in LT-F11, we detected a large clone integrated into the *RPS6KA3* gene (associated with possible cell-autonomous growth) (*26*) and a single provirus integrated in the *FMN1* gene. Similarly, in LT-F12, we found a large clone integrated into the *SH3KBP1* gene. In LT-F10, we observed two proviruses in nongenic regions and one in a genic location. In LT-F13, we failed to capture any clonally expanded intact proviruses despite analyzing up to 19 million PBMCs. In LT-F14, we captured a small intact proviral clone; however, because of limited sample availability, we were not able to identify any integration sites of intact proviruses using the MIP-seq assay (Fig. 3).

In our subsequent analysis of the 17 male participants from whom a detailed integration site profile was obtained (Fig. 4), we observed one participant (LT-M03) with a large genome-intact proviral clone integrated in centromeric satellite DNA. In two additional study participants (LT-M01 and LT-M05), intact clonal proviruses were integrated in nongenic regions far away from host transcriptional start sites (TSSs); moreover, there were two participants (LT-M02 and LT-M04) with intact proviruses in *ZNF* genes. In 12 of the 17 male participants (70.6%), we failed to detect dominant intact proviruses in nongenic regions, nongenic centromeric/peri-centromeric DNA, or *ZNF* genes; instead, we noted dominant intact proviruses integrated in typical genic locations, including the *CCDC91* gene

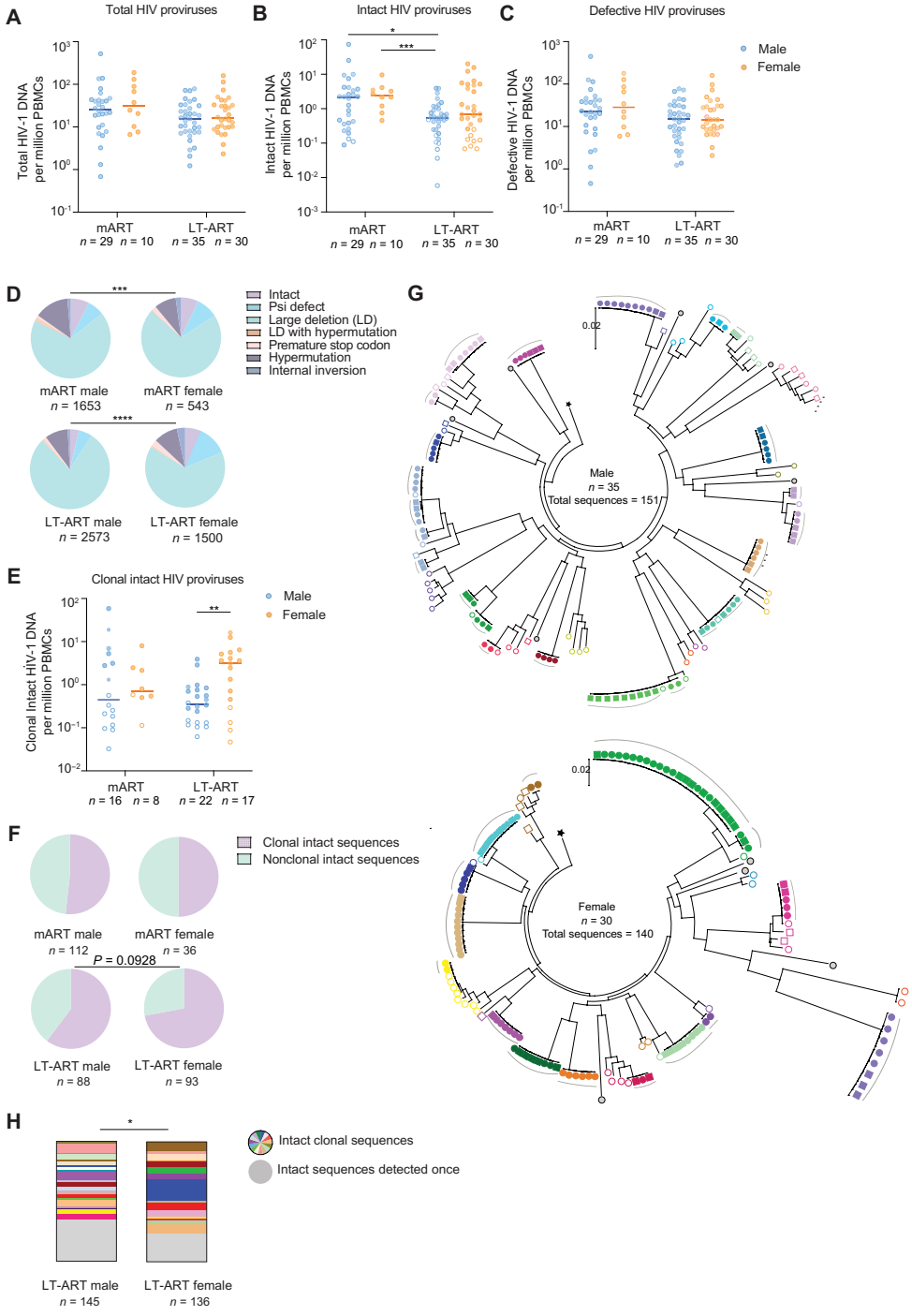


**Fig. 2. Sex-based differences in reservoir size and clonality in people living with HIV-1 on LT-ART.** (**A** to **C**) Frequencies of total (A), intact (B), and defective (C) proviruses in individuals on mART or LT-ART, segregated on the basis of sex. Open circles represent data at the limit of detection. (**D**) Proportions of intact or defective proviruses. (**E** and **F**) Frequencies (E) and proportions (F) of clonally expanded intact proviruses in indicated cohorts stratified on the basis of sex. Participants with at least two detectable intact proviruses were included. (**G**) Circular maximum-likelihood phylogenetic trees for all genome-intact proviral sequences from LT-ART males and females. HXB2, reference HIV-1 sequence. Symbols indicate sequences generated by FLIP-seq (circle) or MIP-seq (square). Genome-intact proviral sequences from a given study participant are depicted in the same color. Gray circles reflect detection of a sole single genome-intact sequence in a given study participant. Clonal sequences, defined by complete sequence identity, are indicated by gray arches. Open symbols represent nonclonal genome-intact sequences in participants with at least two genome-intact sequences detected. The maximum-likelihood phylogenetic trees were inferenced using PhyML. (**H**) Stacked bar charts reflect the clonal composition of HIV-1 reservoir cells in LT-ART males and females. Each color represents an individual clone of sequence-identical genome-intact proviruses, whereas gray-colored sub-bars represent proviruses detected only once. Sequences derived from participants with at least two genome-intact sequences are included. Horizontal lines reflect the median [(A) to (C) and (E)]. Mann-Whitney *U* nonparametric tests were used in (A) to (C) and (E); two-tailed $\chi^2$ tests were used in (D), (F), and (H). ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, and *$P < 0.05$. *n* indicates the number of participants in (A) to (C), (E), and (G) or the number of viral sequences in (D), (F), and (H).

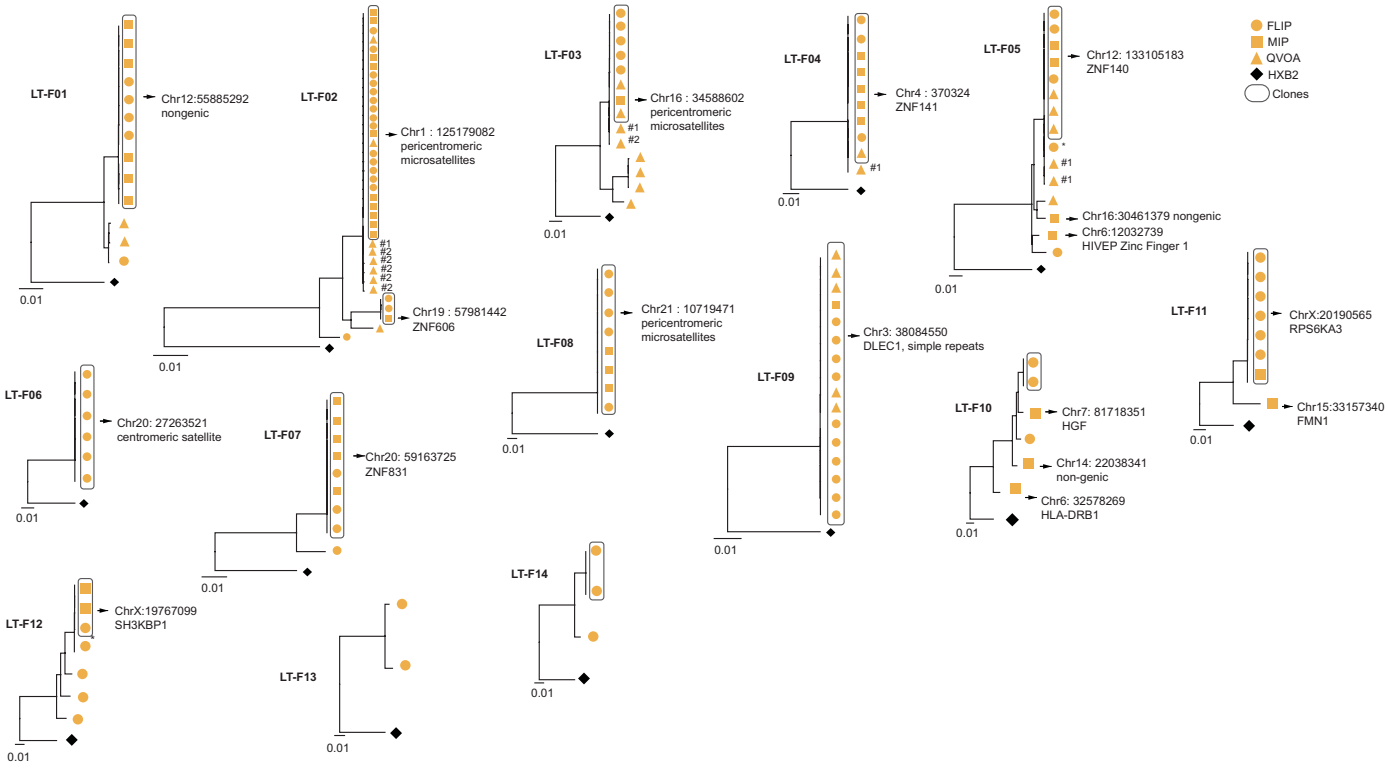


**Fig. 3. Chromosomal positioning of intact HIV-1 proviruses in females on LT-ART.** Maximum-likelihood phylogenetic trees of intact proviruses from females on LT-ART (LT-F01 to LT-F014). Coordinates of chromosomal integration sites and corresponding gene names (if applicable) are indicated. Symbols indicate sequences generated by FLIP-seq, by MIP-seq, or from qVOAs. "*" indicates sequences generated by FLIP-seq or MIP-seq that may be part of adjacent clonal clusters but differ by 1 to 2 bp, possibly because of PCR amplification or sequencing errors. "#*n*" indicates sequences derived from replication-competent proviral genomes retrieved in qVOAs that differ by *n* number of base pairs from adjacent clonal sequences; these sequence variations likely reflect mutations originating from viral sequence diversification during in vitro culture in qVOA assays.

(LT-M06), the *PRPSAP1* gene (LT-M07), the *DGKD* gene (LT-M08), the *FAM230I* gene (LT-M09), the *ATXN7* gene (LT-M10), the *OR4M2-OT1* gene (LT-M11), the *GLB1* gene (LT-M16), and the *HSPA5* gene (LT-M17). For LT-M12, LT-M13, LT-M14, and LT-M15, we noticed a phylogenetically more diverse portfolio of intact proviruses with corresponding detectable integration sites being almost exclusively located in genic regions; in LT-M16, a dominant intact proviral clone located in the *GLB1* gene was paired with two subdominant proviruses located in peri-centromeric regions. Together, these results demonstrate a higher frequency of female participants with evidence for selection of intact proviruses in heterochromatin regions (Fig. 5, A and B). No major selection effects were noted among defective proviruses from our study participants of either sex (Fig. 5B).

Whereas integration of proviruses into heterochromatin regions is associated with transcriptional repression in direct ex vivo assessments (*10*) and in tissue culture experiments (*23*, *27*), our prior work showed that such integration is not associated with irreversible latency; in particular, in vitro stimulation of T cells with nonphysiological latency-reversing agents can reactivate proviruses integrated in heterochromatin, as evidenced by effective detection of replication-competent proviruses in quantitative viral outgrowth assays (qVOAs) from elite controllers, who control HIV-1 replication in the absence of ART (*7*). We performed qVOAs on a total of 13 LT-ART participants with adequate PBMC sample availability (female $n = 8$; male $n = 5$) (tables S3 and S4). Our findings showed no differences in proviral inducibility between elite controllers and LT-ART participants (fig. S3F) (*7*) and no sex differences in proviral inducibility among LT-ART participants (fig. S3G).

## Genomic and epigenetic integration site features

To further explore integration site features of intact proviruses from individuals on long-term ART, we aligned integration sites of intact proviruses to high-resolution Hi-C data capturing intra- and inter-chromosomal three-dimensional (3D) chromatin interactions, allowing us to segregate chromatin compartments with permissive or repressive characteristics (*28*). These data demonstrated that a considerable proportion (30.6%) of intact proviruses from LT-ART participants were integrated in compartments B1, B2, and B4, which are generally associated with repressive heterochromatin features (*28*); this overrepresentation in compartment B was significantly more pronounced ($P < 0.0001$) in female than in male study participants (Fig. 5, C and D). No enrichment of defective proviruses in the B compartments was noted, and alignment of defective proviruses to 3D Hi-C data failed to demonstrate differences between female and male study participants. We also compared the relative positioning of the intact proviruses to host TSSs, as determined by RNA sequencing (RNA-seq) in primary CD4 T cells. We observed a significantly ($P = 0.0477$) elevated chromosomal distance of intact proviruses to the most proximal TSSs in females on long-term ART compared with males (Fig. 5, E and F). As an additional analysis step, we evaluated

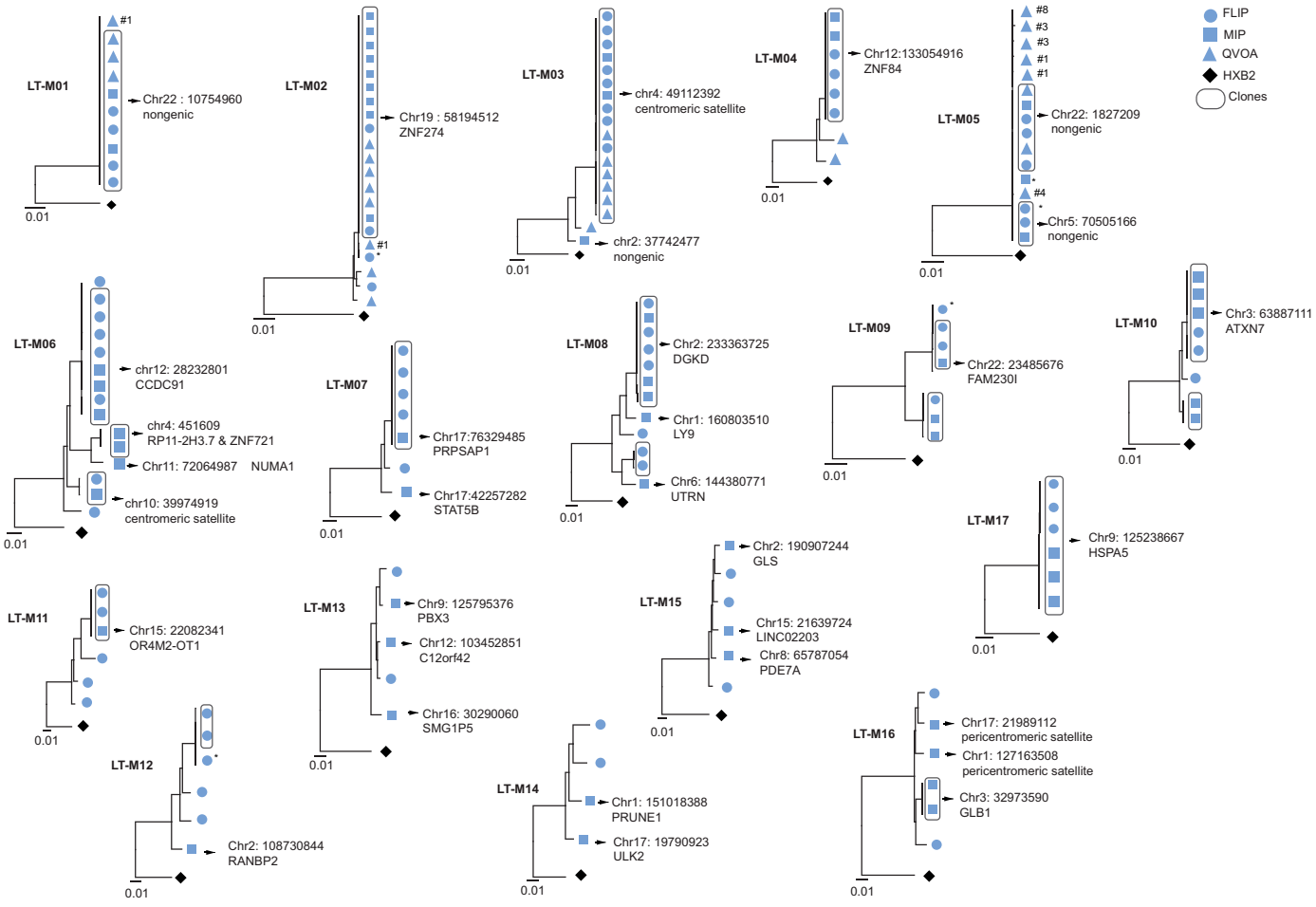


**Fig. 4. Chromosomal positioning of intact HIV-1 proviruses in males on LT-ART.** Maximum-likelihood phylogenetic trees of intact proviruses from males on LT-ART (LT-M01 to LT-M017). Coordinates of chromosomal integration sites and corresponding gene names (if applicable) are indicated. Symbols indicate sequences generated by FLIP-seq, by MIP-seq, or from qVOAs. "*" indicates sequences generated by FLIP-seq or MIP-seq that may be part of adjacent clonal clusters but differ by 1 to 2 bp, possibly because of PCR amplification or sequencing errors. "#*n*" indicates sequences derived from replication-competent proviral genomes retrieved in qVOAs that differ by *n* number of base pairs from adjacent clonal sequences; these sequence variations likely reflect mutations originating from viral sequence diversification during in vitro culture in qVOA assays.

the chromosomal location of intact proviruses relative to the activating histone modifications H3K4me1 and H3K27ac, using genome-wide chromatin immunoprecipitation sequencing (ChIP-seq) data from primary CD4 T cells deposited by the Roadmap Consortium (*29*). We observed that the intact proviruses from LT-ART participants displayed significant de-enrichment ($P < 0.0001$) of both H3K27ac and H3K4me1 marks in proximity to their integration sites, compared with corresponding data from mART study persons. In contrast, no similar changes were observed for defective proviruses from both study cohorts (Fig. 5G). Reductions of activating histone modifications in proximity to intact proviruses were significantly ($P = 0.0450$) more notable in female than in male LT-ART study participants (Fig. 5H).

## Characteristics of viral escape from T and B cell immune responses

We next investigated viral sequence variation in HIV-1 proviruses obtained from our study participants and evaluated evidence of sequence adaptation to HIV-1–specific T cells or B cells (Fig. 6). First, pair-wise comparisons of all distinct intact proviruses within a given study person demonstrated that sequences from LT-ART participants had lower intraindividual viral diversity compared with sequences from the mART group (Fig. 6A), suggesting that longer durations of ART may limit the persistence of distinct viral quasispecies and restrict the phylogenetic complexity of the viral reservoir pool. Within the LT-ART group, the intraindividual phylogenetic distances among intact proviruses were lower in females compared with males (Fig. 6B), although these differences did not reach statistical significance ($P = 0.0675$). Overall, we noted that the frequencies of sequence variations associated with resistance to broadly neutralizing antibodies (bNAbs) or cytotoxic T lymphocytes (CTLs), determined using previously described algorithms (*8*), were not different between individuals from the mART and LT-ART groups (Fig. 6, C, E, and G). However, genome-intact sequences in females from the LT-ART group displayed significantly ($P = 0.0193$) fewer CTL mutations compared with males (Fig. 6, F and H). No such

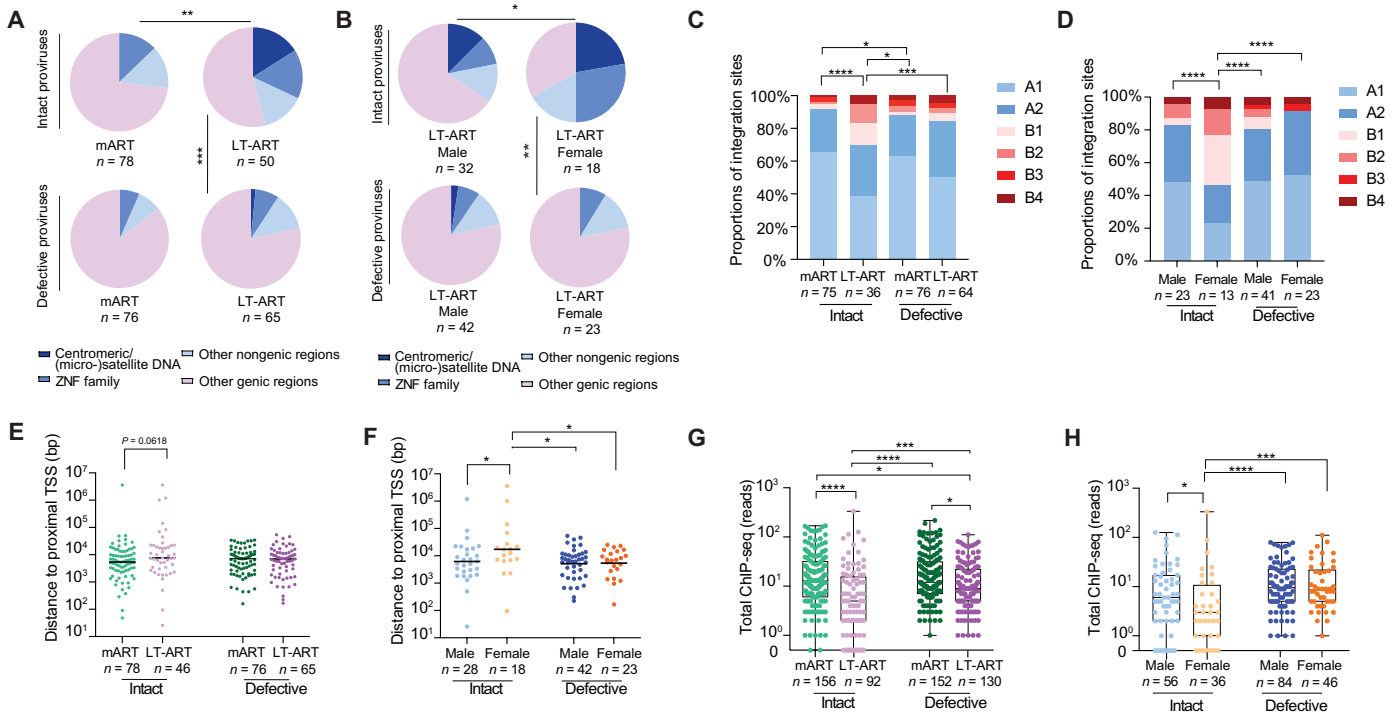


**Fig. 5. Integration site profile of intact and defective HIV-1 proviruses in people living with HIV-1 on LT-ART.** (**A** and **B**) Proportions of intact and defective proviruses in defined genomic regions in indicated study cohorts (A) and in LT-ART males and females (B). (**C** and **D**) Proportions of proviral integration sites located in Hi-C genomic compartments A and B (and associated subcompartments) in mART and LT-ART individuals (C) and in LT-ART males and females (D). (**E** and **F**) Chromosomal distance of proviral integration sites to most proximal host TSSs, as determined by RNA-seq in CD4 T cells from reference datasets in indicated study cohorts (E) or in LT-ART males and females (F). Horizontal lines reflect the median. (**G** and **H**) Box and whisker plots reflect activating histone protein modifications in proximity to proviral integration sites inferred from reference datasets of the ROADMAP consortium (*68*). The median, the 25% and 75% percentiles, and the minimum and maximum of H3K27ac and H3K4me1ChIP-seq reads within ±2500 bp of the proviral integration site are shown in indicated study cohorts (G) and in LT-ART males and females (H). Integration sites not covered in the reference dataset were excluded in (C) to (H). Kruskal-Wallis nonparametric tests were used in (E) to (H), and two-tailed $\chi^2$ tests were used in (A) to (D). Nominal *P* values are shown in (E) to (H). ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, and *$P < 0.05$. Clonal integration sites are counted once. *n* indicates the number of proviral integration sites.

sex-specific differences were observed for bNAb-associated resistance mutations (Fig. 6D).

## Distinct innate immune signatures in females on LT-ART

The distinct viral reservoir and integration site profile of intact proviruses in females on long-term ART may reflect a more efficient immune-mediated selection process of HIV-1 reservoir cells within females. To evaluate this, we characterized phenotypic signatures of innate immune cells in selected LT-ART study participants ($n = 9$ females and $n = 11$ males) from whom sufficient cells were available, using a multidimensional 27-color spectral flow cytometry assay. Within these study persons, five females (LT-F01 to LT-F05) and five males (LT-M01 to LT-M05) had dominant intact proviruses in heterochromatin locations ($HTC^{high}$); four females (LT-F09 and LT-F11 to LTF13) and six males (LT-M06, LT-M11, and LT-M13 to LT-M16) had no or weak evidence of a selection for intact proviruses in heterochromatin ($HTC^{low/neg}$). As an initial analysis step, we used a linear discriminant analysis to evaluate the global phenotypic profile of innate immune cells [natural killer (NK) cells, monocytes, conventional dendritic cells, and plasmacytoid dendritic cells]; this analysis demonstrated a distinct representation of innate immune cells in $HTC^{high}$ females relative to the other study subgroups (Fig. 7A). We observed that proportions of $CD64^+PD\text{-}L1^+$ myeloid dendritic cells (mDCs), previously shown to be enriched in elite controllers (*30*), were more frequent in females, specifically in those with intact proviruses enriched in heterochromatin locations (Fig. 7B). Moreover, the proportions of $CD64^+PD\text{-}L1^+$ mDCs were positively correlated with the frequencies of clonally expanded intact proviruses and with the proportions of intact proviruses in heterochromatin chromosomal locations (Fig. 7C). No such associations were observed for plasmacytoid dendritic cells (fig. S7E).

In our subsequent analysis, we focused on the phenotypic profiles of NK cells in our study participants. We noted that a global linear discriminant analysis of NK cell immunophenotypes produced notable distinctions between the study subgroups (Fig. 7D). In particular, NK cells from females displayed a phenotypic profile characterized by a more limited ensemble expression of inhibitory killer cell immunoglobulin-like receptors (KIRs) and of the inhibitory receptor T cell immunoreceptor with Ig and ITIM domains (TIGIT) (Fig. 7, E to G); in contrast, the highest proportions of NK cells expressing combinations of these inhibitory receptors were found in $HTC^{low/neg}$ males (Fig. 7, E to G). Frequencies of clonal intact proviruses and proportions of intact proviruses in heterochromatin regions were inversely correlated with proportions of NK cells expressing these inhibitory receptors (Fig. 7H). The expression of the activating or cytotoxicity-related receptors CD161, Nkp30, and NKG2D and of NKG2A, which, despite its role

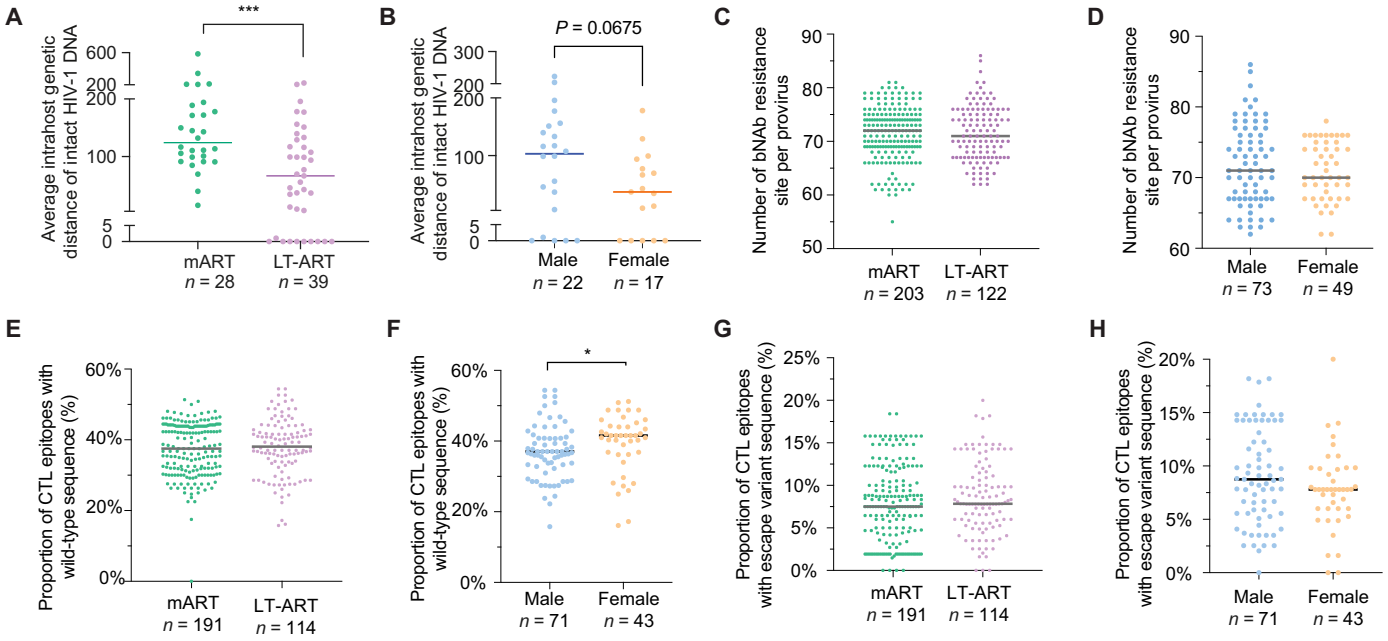


**Fig. 6. Footprints of viral immune escape from host T and B cell immune responses.** (**A** and **B**) Average genetic distances, determined by pairwise comparisons of nucleotide differences, between all intact proviruses from a given study participant, in indicated study cohorts (A) and in males and females on LT-ART (B). Participants with at least two intact proviruses detected were included. (**C** and **D**) Numbers of signature sites associated with resistance to bNAbs in intact proviral sequences of indicated study cohorts (C) and in males and females on LT-ART (D). Signature sites associated with resistance to four classes of bNAbs were determined as described before (*66*). (**E** to **H**) Proportions of CTL epitopes (restricted by autologous HLA class I alleles) within intact proviruses that harbor the clade B consensus wild-type sequences [(E) and (F)] or previously described CTL escape variants [(G) and (H)]. Defined escape mutations listed in the LANL HIV Immunology Database (www.hiv.lanl.gov) were considered. Horizontal lines reflect the median. Mann-Whitney *U* nonparametric tests were used. ***$P < 0.001$ and *$P < 0.05$. Clonal sequences are counted once. *n* indicates the number of participants in (A) and (B) or the number of viral sequences in (C) to (H).

as an inhibitory receptor, has been associated with superior cytotoxic and noncytotoxic NK cell activities in HIV-1 infection (*31*, *32*), was elevated in NK cells from females while being reduced in males without evidence of selection for intact proviruses in heterochromatin locations (Fig. 7, I to K). In addition, statistical correlations supported a trend for a positive association between expression of these NK cell markers, the frequencies of clonal intact proviruses, and the proportions of intact proviruses in heterochromatin locations (Fig. 7L). Furthermore, the ratio of NKG2A-expressing NK cells relative to NKG2C-expressing NK cells, associated with better control of HIV-1 viremia in some studies (*33*, *34*), was lowest in HTC$^{low/neg}$ males and positively associated with clonality and heterochromatin locations of intact proviruses (Fig. 7, M and N). The expression of activating NK cell markers was associated with the proportions of CD64$^{+}$PD-L1$^{+}$ mDCs (fig. S9), supporting a form of immunological cross-talk as described previously (*35*). Together, these results suggest a role of innate immune cells in driving selection of intact HIV-1 proviruses in heterochromatin regions during long-term ART.

### HIV-1–specific T cell responses in people living with HIV-1 on LT-ART

To complement our immunophenotypic analysis, we also conducted multiparametric spectral flow cytometry assays to evaluate adaptive HIV-1–specific CD4 and CD8 T cell responses based on the secretion of cytokines or effector molecules [interferon-γ (IFN-γ), interleukin-2 (IL-2), and tumor necrosis factor–α (TNF-α)] and up-regulation of defined surface markers (CD25, OX40, CD69, and CD40L) in response to stimulation with HIV-1 Gag peptides (Fig. 8). Whereas HIV-1–specific T cell responses were readily detectable in our study participants despite more than 20 years of continuous pharmacological suppression of plasma viremia, a global linear discriminant analysis demonstrated only relatively weak separations between the phenotypic and functional profile of HIV-1–specific CD4 (Fig. 8A) and CD8 (Fig. 8J) T cells among the study subgroups. We noted that cytokine-secreting HIV-1–specific CD4 and CD8 T cells in HTC$^{high}$ males exceeded those of most of the other study subgroups; however, this was not true when antigen-specific T cells were measured on the basis of antigen-specific up-regulation of surface markers (Fig. 8, B, D, K, and M). Moreover, a correlative statistical analysis failed to demonstrate associations between HIV-1–specific CD4 and CD8 T cell responses and the clonality or heterochromatin locations of intact proviruses (Fig. 8, C, E, L, and N). Nevertheless, relative to females, male study participants were characterized by higher frequencies of polyfunctional HIV-1–specific CD4 (Fig. 8, F to H) and CD8 (Fig. 8, O to Q) T cells, which were particularly enriched in HTC$^{high}$ males; however, these polyfunctional T cell responses were not correlated with proviral reservoir characteristics (Fig. 8, I and R). Together, these results suggest a limited, if any, influence of adaptive T cell responses on the sex-specific selection and evolution of HIV-1 reservoir cells.

## DISCUSSION

Sex differences can have a marked impact on host immune responses, as evidenced in a number of viral infections (*17*). In the context of HIV-1 infection, female sex has been associated with lower set point viremia (*36*, *37*), higher life expectancy (*38*), delayed viral

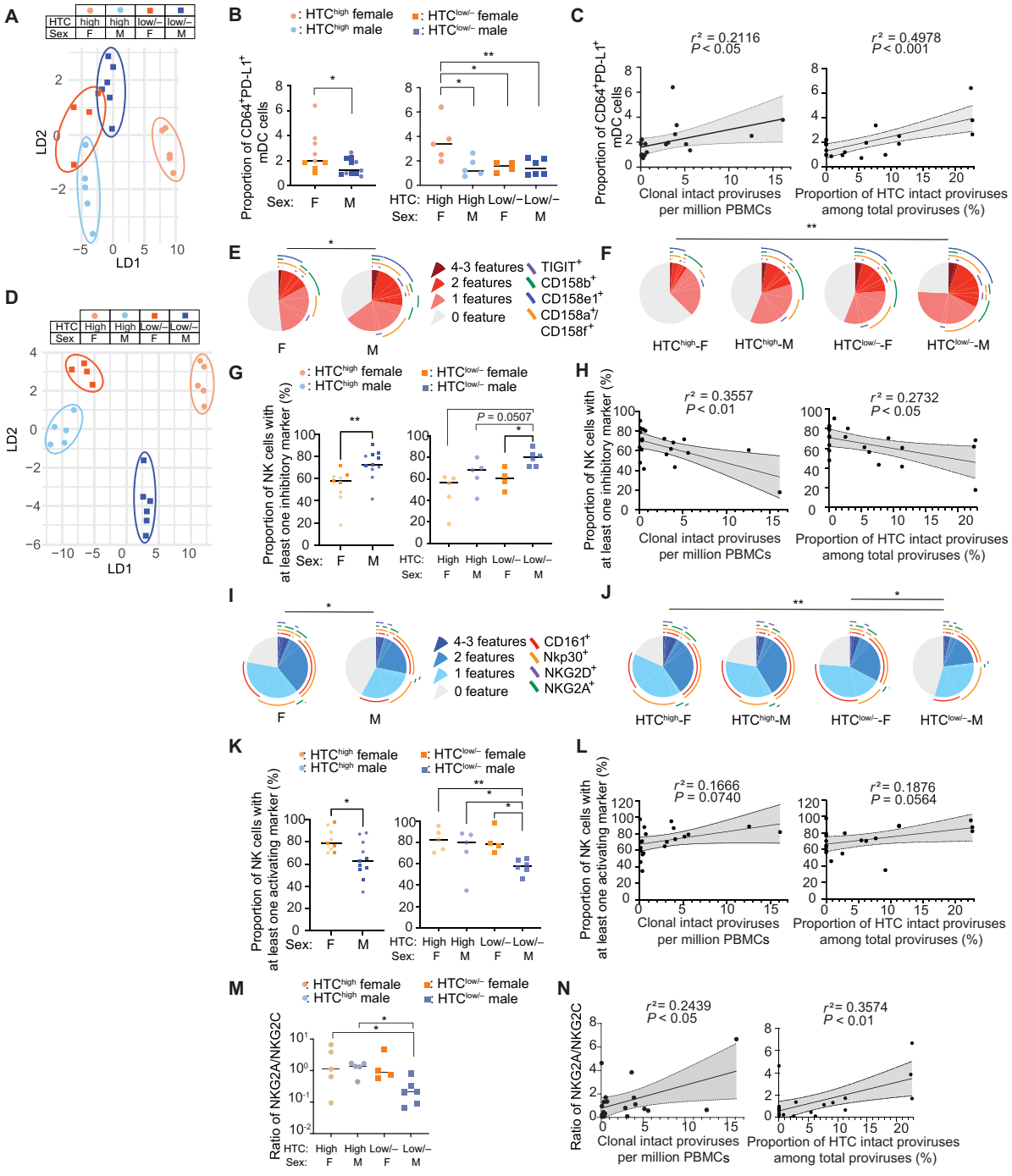


**Fig. 7. Innate immune responses in people living with HIV-1 on LT-ART.** (**A** and **D**) Linear discriminant analysis of total innate immune cells (A) and NK cells (D) in indicated study subgroups of LT-ART participants is shown. The phenotypic profiles of total innate immune cells and NK cells in PBMC samples were characterized by the expression of 25 (A) or 15 (D) relevant surface markers. (**B** and **C**) Proportions of $CD64^+PD\text{-}L1^+$ mDCs in indicated study subgroups (B) and their associations with frequencies of clonal intact proviruses and proportions of intact proviruses in heterochromatin regions among total proviruses (C). (**E** to **L**) SPICE diagrams highlighting proportions of NK cells with zero, one, or more inhibitory [(E) to (H)] or activating [(I) to (L)] phenotypic markers. Pie chart colors reflect the proportions of cells expressing the indicated numbers of markers; arcs reflect the proportions of cells expressing a given marker. Proportions of NK cells with at least one feature [(G) and (K)] and their associations with frequencies of clonal intact proviruses and proportions of intact proviruses in heterochromatin regions among total proviruses [(H) and (L)]. (**M** and **N**) Ratios of NKG2A-expressing versus NKG2C-expressing NK cells in indicated study subgroups (M) and their associations with frequencies of clonal intact proviruses and proportions of intact proviruses in heterochromatin regions among total proviruses (N). The study subgroups are segregated on the basis of sex (F, female; M, male) and the proportion of intact HIV-1 reservoir in heterochromatin locations ($HTC^{high}$ or $HTC^{low/-}$). PD-L1, program cell death ligand 1. Horizontal bars reflect the median. Mann-Whitney *U* tests or Kruskal-Wallis nonparametric tests were used in (B), (G), (K), and (M), permutation tests were used in (E), (F), (I), and (J), and linear regressions were used in (C), (H), (L), and (N). Linear regression coefficients ($r^2$) and nominal *P* values are shown. ***$P < 0.001$, **$P < 0.01$, and *$P < 0.05$.

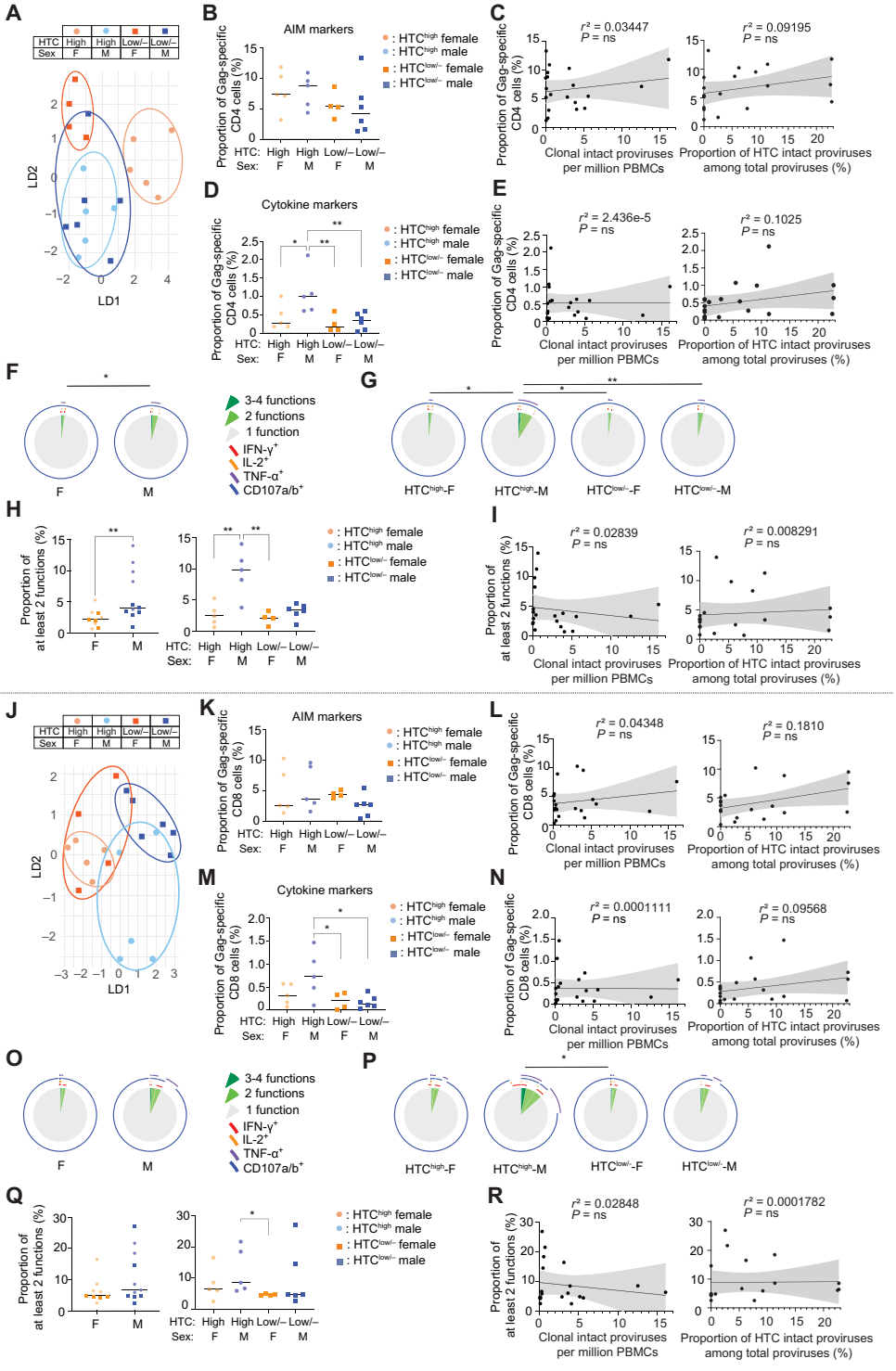


**Fig. 8. T cell immune responses in people living with HIV-1 on LT-ART.** (**A** and **J**) Linear discriminant analysis of CD4 (A) and CD8 (J) T cells in indicated study subgroups in participants on LT-ART. The phenotypic and functional profile was characterized using the expression of 10 shared surface and intracellular markers after stimulation with HIV-1 Gag peptide pool. (**B** to **E** and **K** to **N**) Proportions of HIV-1–Gag–specific CD4 [(B) and (D)] and CD8 [(K) and (M)] T cells in total CD4 or CD8 T cells, identified on the basis of up-regulation of any of the activation-induced markers (CD25, OX40, CD69, and CD40L) [(B) and (K)] or secretion of cytokines (IFN-γ, IL-2, and TNF-α) [(D) and (M)] in indicated study subgroups. Associations with frequencies of clonal intact proviruses and proportions of intact proviruses in heterochromatin regions among total proviruses are indicated [(C), (E), (L), and (N)]. (**F**, **G**, **O**, and **P**) SPICE diagrams highlighting proportions of HIV-1–specific CD4 [(F) and (G)] and CD8 [(O) and (P)] T cells with one or more of the indicated functions. Pie chart colors reflect the number of effector functions, whereas proportions of cells expressing a specific effector function are indicated by an arc. (**H**, **I**, **Q**, and **R**) Proportions of HIV-1–specific CD4 (H) and CD8 (Q) T cells with at least two functions and their associations with frequencies of clonal intact proviruses and proportions of intact proviruses in heterochromatin regions among total proviruses [(I) and (R)]. The study subgroups are segregated on the basis of sex and the proportion of intact HIV-1 reservoir in heterochromatin locations ($HTC^{high}$ or $HTC^{low/-}$). Horizontal bars in dot plots represent the median. Mann-Whitney *U* tests or Kruskal-Wallis nonparametric tests were used in (B), (D), (H), (K), (M), and (Q), permutation tests were used in (F), (G), (O), and (P), and linear regressions were used in (C), (E), (I), (L), (N), and (R). Linear regression coefficients ($r^2$) and nominal *P* values are shown. $^{**}P < 0.01$ and $^{*}P < 0.05$. ns, not significant.

rebound after treatment interruption (*39*), and an up to fourfold higher probability of developing an elite controller or posttreatment controller phenotype (*40*, *41*). Moreover, two elite controllers who maintained undetectable intact proviruses and may have reached a spontaneous, immune-mediated virological cure of HIV-1 infection were both females (*7*, *42*). To what extent sex differences may influence viral reservoir cells, the major barrier to a cure of HIV-1 infection, remains an area of active investigation. Prior studies relying on quantitative measurements of the reservoir cell pool have largely failed to detect any differences between females and males (*16*, *43*, *44*). Using a more complex qualitative analysis approach that involved single-genome amplification of more than 4073 individual proviruses, we here demonstrate profound sex-specific differences in viral reservoir cell structure and composition, most dominantly characterized by the increased frequencies of genome-intact proviruses integrated in heterochromatin locations in females compared with

males. We propose that these structural differences in the viral reservoir cell profile in females result from sex-specific immune selection forces, implying that the female immune system is more effective in driving immune selection pressure within the viral reservoir cell pool.

Previous studies demonstrated at least two distinct features of viral reservoir cells in females: Scully *et al.* and Gianella *et al.* (*16*, *43*) both reported a reduction in multiply spliced cell-associated viral RNA and reduced residual plasma viremia in females on suppressive ART. Both observations are consistent with, and can possibly be explained by, a more effective selection of proviruses in heterochromatin regions where viral transcription is suppressed. A considerable number of studies have now demonstrated that integration of proviruses in heterochromatin is associated with lower or absent transcriptional activity; this was true in in vitro models of viral latency (*27*) and when participant-derived cells were assayed directly ex vivo using single-cell analytic technologies (*10*). Preferential selection of intact proviruses in heterochromatin in females does not exclude the possibility that proviral transcription in females is also inhibited by additional factors, including female sex hormones (*45*), or sex-related epigenetic differences in regulation of proviral gene transcription (*46*). Prior work also suggests that the inducibility of intact proviruses in in vitro viral outgrowth assays is reduced in females (*44*); however, our study failed to demonstrate sex-specific differences in viral reactivation in vitro. We argue that the highly stochastic nature of the in vitro viral outgrowth assay (*47*), the finding that only approximately 1 in 100 intact proviruses can be reactivated in this assay system, and the high probability that viral reservoir cells die during viral reactivation in in vitro culture (*48*) reduce the usefulness of the in vitro viral outgrowth assay for assessing HIV-1 reservoir cells.

Using a comprehensive immunophenotyping approach, we noted that the selection of large clones of intact proviruses integrated in transcriptionally repressive chromatin regions in females was most closely related to NK cell responses, specifically to subsets of NK cells characterized by phenotypic features of enhanced functionality that were expanded in females. Supporting this association, recent data suggest that female sex can be associated with improved NK cell functionality and with increased proportions of NK cells with enhanced effector functions, specifically during advanced stages of the age continuum (*49*), such as the study participants analyzed in our work. To some extent, such improved NK cell responses in females may be genetically determined, although microanatomical changes in cytokine milieu (*50*), altered cross-talk with alternative immune cells (*51*), differential expression of the X-chromosomally encoded epigenetic regulator ubiquitously transcribed tetratricopeptide repeat, X chromosome (UTX) (*52*), and hormonal changes during peri- and postmenopause (*49*), which were likely present in most of our study persons, may also have contributed. The true effect of NK cell responses against viral reservoir cells will critically depend on the surface expression of the ligands for activating and inhibitory NK cell markers on virally infected cells, including nonclassical HLA class I molecules that serve as ligands for KIRs and for NKG2A. Recent technology advances now permit direct assessment of the surface phenotype of viral reservoir cells encoding for intact proviruses (*11*) and will be highly informative as next steps for evaluating how NK cells can interact with and target the viral reservoir cell pool.

Our study failed to identify an important association between HIV-1–specific T cells and qualitative or quantitative features of viral reservoir cells; in particular, there was no association between HIV-1–specific T cells and the evolution of intact proviruses in heterochromatin regions. This finding, supported by other recent observations in humans (*53*) and nonhuman primates (*54*), casts doubt on the ability of virus-specific T cells to effectively engage viral reservoir cells during suppressive ART. We observed more limited CTL resistance mutations in intact proviruses from females; this might suggest a more effective immune control at early stages of infection before ART initiation in females, resulting in reduced viral turnover with fewer opportunities for viral sequence adaptation to CTL-mediated immune pressure (*55*). Such a hypothesis is supported by prior studies indicating lower set point viremia in untreated females compared with age-matched males (*56*). Alternatively, reduced frequencies of CTL resistance mutations in females could reflect a more potent immune clearance of intact proviruses during long durations of ART suppression in females (*2*, *13*) or generally reveal weaker exposure to CTL-mediated immune pressure in females.

Our work has several limitations. First, our study only analyzes HIV-1 reservoir cells in peripheral blood, which might not entirely reflect the reservoir characteristics in lymphoid tissues, the likely site from which viruses rebound after treatment interruptions (*57*). Nevertheless, prior studies show that clonal clusters of intact proviral sequences are frequently shared between viral reservoir cells in blood and tissues (*58*). Moreover, we only conducted a cross-sectional analysis of immune profiles and viral reservoir structures in our study participants, because longitudinal samples collected from our study participants over 2 decades of treatment were not available. Characterizing the interplay between host immune responses and viral reservoir cells during prolonged periods of ART in females and males will therefore require additional longitudinal studies. In addition, almost all our study participants were of advanced ages by the time of sample collection, and it is possible that sex-dependent immune effects on the viral reservoir cell pool might materialize differently during infancy, adolescence, and young adulthood.

In conclusion, by using a high-resolution molecular approach for viral reservoir cell analysis, our study demonstrates marked sex-specific differences in the HIV-1 reservoir profile and suggests more efficient selection of intact proviral sequences in heterochromatin in females, most likely as a result of innate immune activity. Our study reinforces the importance of considering sex in the design and implementation of cure-directed medical interventions and suggests that females may be better candidates to explore innate immunity-dependent strategies for targeting HIV-1 reservoir cells.

## MATERIALS AND METHODS

### Study design

We conducted a cross-sectional, noninterventional study to evaluate sex differences in HIV-1 reservoir cell profiles. Study participants living with HIV-1 were recruited at Massachusetts General Hospital, Brigham and Women's Hospital, the University of California, San Francisco, Case Western Reserve University, AIDS Clinical Trials Group, and the Multicenter AIDS Cohort Study (MACS) and Women's Interagency HIV Study (WIHS) Combined Cohort Study (MWCCS). No formal sample size calculations, randomizations, or investigator blinding operations were performed. All study participants provided written consent in accordance with the Declaration of Helsinki before enrollment. PBMC samples were collected according to protocols approved by the respective Institutional Review Boards. Clinical and demographic characteristics of study participants are summarized in tables S1 and S2.

### Intact proviral DNA assay

PBMCs were subjected to DNA extraction using DNeasy Blood and Tissue kits (QIAGEN, 69504). The Intact Proviral DNA Assay (IPDA) was performed using primers and probes described previously (*59*).

### Whole-genome amplification

DNA was diluted to single viral genomes based on Poisson distribution. Subsequently, DNA in each well was subjected to multiple displacement amplification with phi29 polymerase (QIAGEN REPLI-g Single Cell Kit, 150345), per the manufacturer's protocol.

### HIV-1 near–full-genome sequencing

DNA resulting from full-genome amplification reactions was subjected to HIV-1 near–full-genome amplification using a nonmultiplexed multiamplicons approach as described previously (*22*, *60*). All near–full-length polymerase chain reaction (PCR) products were subjected to Illumina MiSeq sequencing at the Massachusetts General Hospital (MGH) DNA Core facility. Resulting short reads were de novo assembled using Ultracycler v1.0 and aligned to HXB2 to identify large deleterious deletions (<8000 bp of the amplicon aligned to HXB2), out-of-frame indels, premature/lethal stop codons, internal inversions, or packaging signal deletions (≥15-bp insertions and/or deletions relative to HXB2), using an automated in-house pipeline written in Python programming language (https://github.com/BWH-Lichterfeld-Lab/Intactness-Pipeline) (*61*). The presence/absence of APOBEC-3G/3F–associated hypermutations was determined using the Los Alamos HIV Sequence Database Hypermut 2.0 program (*62*). Viral sequences that lacked all mutations listed above were classified as "genome-intact." Multiple sequence alignments were performed using Multiple Alignment using Fast Fourier Transform (MAFFT) (*63*) in Geneious Prime software v2023.1 (https://geneious.com). Phylogenetic distances between sequences were examined using a PhyML tree (*64*) in Geneious Prime software. Viral sequences were considered clonal if they had completely identical consensus sequences; single-nucleotide variations in primer binding sites were not considered for clonality analysis. The proportions of optimal CTL epitopes (restricted by autologous HLA class I alleles) matching the clade B consensus sequence and CTL escape variants restricted by autologous HLA class I alleles described in the LANL HIV Immunology Database (www.hiv.lanl.gov) were determined. The number of sequence mutations in clade B proviruses associated with HLA class I–mediated pressure (*65*) and the sensitivity of proviral species to bnAbs (*66*) were calculated as previously described.

### Integration site analysis

Integration sites associated with each viral sequence were obtained using integration site loop amplification, as previously described (*67*); DNA produced by whole-genome amplification was used as template. Biocomputational identification of integration sites was performed as previously described (*29*, *67*). The final list of integration sites and corresponding chromosomal annotations was obtained using Ensembl (v86, www.ensembl.org), the UCSC Genome Browser (www.genome.ucsc.edu), and GENCODE (v39, www.gencodegenes.org). Repetitive genomic sequences harboring HIV-1 integration sites were identified using RepeatMasker (www.repeatmasker.org). All HIV-1 integration sites of study participants are listed in data file S1.

### RNA-seq and Hi-C sequencing data

RNA-seq data generated from primary CD4 T cells deposited in the National Center for Biotechnology Information Gene Expression Omnibus database (GSE144334) were used, as described in a previous publication (*7*). Hi-C sequencing data were described by Rao *et al.* (*28*). ChIP-seq data are derived from primary human memory CD4 T cells listed in the Roadmap Consortium (*68*).

### Viral outgrowth assays

Viral outgrowth assays were performed as described previously (*13*). Demographic information, number of cells assayed, and infectious units per million (IUPM) of selected participants are summarized in tables S3 and S4.

### Flow cytometry

For immunophenotyping of selected study participants (table S5), PBMCs were thawed, stained with LIVE/DEAD Blue Viability Dye (Invitrogen) for 15 min, and subsequently blocked with 2 μl of FcR blocking reagent (Miltenyi Biotec) for 10 min. Next, cells were incubated for 25 min with different combinations of appropriately titrated antibodies as listed in table S6. Afterward, the cells were fixed in 2% paraformaldehyde in phosphate-buffered saline and acquired on an Aurora spectral cytometer (Cytek Biosciences) at the Ragon Institute Imaging Core Facility at MGH. To analyze HIV-1–specific T cell responses, PBMCs were stimulated with HIV-1 clade B Gag peptide pool (mix of 123 overlapping 15-mer peptides at 2 μg/ml for each peptide, NIH AIDS Reagent Program, ARP-12425) or staphylococcal enterotoxin B (SEB) at 0.4 μg/ml (Sigma-Aldrich) in the presence of anti-CD28/anti-CD49d costimulatory antibodies at 1 μg/ml (BD Biosciences) and CD107a/b antibodies for an hour before adding brefeldin A (5 μg/ml) and monensin (1 μg/ml). After overnight stimulation, cells were stained with LIVE/DEAD Blue Viability Dye and with surface antibodies as listed in table S7 for 25 min. After staining, cells were washed, fixed, and permeabilized using the FoxP3 transcription factor staining buffer set (eBioscience) for 30 min at 4°C. Intracellular cytokine staining was performed using antibodies as listed in table S7 for 30 min at 4°C. Cells were acquired on an Aurora spectral cytometer (Cytek Biosciences). Fluorescence minus one (FMO) controls were prepared as part of the gating strategy. Unstimulated negative controls were subtracted as background. Data were analyzed using FlowJo v.10.8.1 software (Tree Star LLC) and using the SPICE software (version 6.0). Demographic information of study subgroups is summarized in table S5. Gating strategies are summarized in figs. S12 and S13.

### Statistical analyses

Data are presented as pie charts, bar charts, scatterplots, and box-and-whisker plots overlaid with individual values. All statistical analyses were performed using GraphPad Prism 8.2.1 and SPICE v6.0 software using tests as indicated in the figure legends. *P* values less than 0.05 were considered significant. Individual-level data are presented in data file S2.

## Supplementary Materials

**The PDF file includes:**

Figs. S1 to S13
Tables S1 to S7

**Other Supplementary Material for this manuscript includes the following:**
Data files S1 and S2
MDAR Reproducibility Checklist

## REFERENCES AND NOTES


1. UNAIDS, AIDSinfo (2023); https://aidsinfo.unaids.org/.
2. M. Lichterfeld, C. Gao, X. G. Yu, An ordeal that does not heal: Understanding barriers to a cure for HIV-1 infection. *Trends Immunol.* **43**, 608–616 (2022).
3. K. G. Lassen, J. R. Bailey, R. F. Siliciano, Analysis of human immunodeficiency virus type 1 transcriptional elongation in resting $CD4^+$ T cells in vivo. *J. Virol.* **78**, 9105–9114 (2004).
4. A. Ciuffi, M. Llano, E. Poeschla, C. Hoffmann, J. Leipzig, P. Shinn, J. R. Ecker, F. Bushman, A role for LEDGF/p75 in targeting HIV DNA integration. *Nat. Med.* **11**, 1287–1289 (2005).
5. R. Craigie, F. D. Bushman, HIV DNA integration. *Cold Spring Harb. Perspect. Med.* **2**, a006890 (2012).
6. H. Takata, J. L. Mitchell, J. Pacheco, A. Pagliuzza, S. Pinyakorn, S. Buranapraditkun, C. Sacdalan, L. Leyre, S. Nathanson, J. C. Kakazu, J. Intasan, P. Prueksakaew, N. Chomchey, N. Phanuphak, M. de Souza, E. K. Haddad, M. Rolland, S. Tovanabutra, S. Vasan, D. C. Hsu, N. Chomont, L. Trautmann, An active HIV reservoir during ART is associated with maintenance of HIV-specific $CD8^+$ T cell magnitude and short-lived differentiation status. *Cell Host Microbe* **31**, 1494–1506.e4 (2023).
7. C. Jiang, X. Lian, C. Gao, X. Sun, K. B. Einkauf, J. M. Chevalier, S. M. Y. Chen, S. Hua, B. Rhee, K. Chang, J. E. Blackmer, M. Osborn, M. J. Peluso, R. Hoh, M. Somsouk, J. Milush, L. N. Bertagnolli, S. E. Sweet, J. A. Varriale, P. D. Burbelo, T. W. Chun, G. M. Laird, E. Serrao, A. N. Engelman, M. Carrington, R. F. Siliciano, J. M. Siliciano, S. G. Deeks, B. D. Walker, M. Lichterfeld, X. G. Yu, Distinct viral reservoirs in individuals with spontaneous control of HIV-1. *Nature* **585**, 261–267 (2020).
8. X. Lian, C. Gao, X. Sun, C. Jiang, K. B. Einkauf, K. W. Seiger, J. M. Chevalier, Y. Yuki, M. Martin, R. Hoh, M. J. Peluso, M. Carrington, E. Ruiz-Mateos, S. G. Deeks, E. S. Rosenberg, B. D. Walker, M. Lichterfeld, X. G. Yu, Signatures of immune selection in intact and defective proviruses distinguish HIV-1 elite controllers. *Sci. Transl. Med.* **13**, eabl4097 (2021).
9. G. Vansant, A. Bruggemans, J. Janssens, Z. Debyser, Block-and-lock strategies to cure HIV infection. *Viruses* **12**, 84 (2020).
10. K. B. Einkauf, M. R. Osborn, C. Gao, W. Sun, X. Sun, X. Lian, E. M. Parsons, G. T. Gladkov, K. W. Seiger, J. E. Blackmer, C. Jiang, S. A. Yukl, E. S. Rosenberg, X. G. Yu, M. Lichterfeld, Parallel analysis of transcription, integration, and sequence of single HIV-1 proviruses. *Cell* **185**, 266–282.e15 (2022).
11. W. Sun, C. Gao, C. A. Hartana, M. R. Osborn, K. B. Einkauf, X. Lian, B. Bone, N. Bonheur, T.-W. Chun, E. S. Rosenberg, B. D. Walker, X. G. Yu, M. Lichterfeld, Phenotypic signatures of immune selection in HIV-1 reservoir cells. *Nature* **614**, 309–317 (2023).
12. J. Kuniholm, C. Coote, A. J. Henderson, Defective HIV-1 genomes and their potential impact on HIV pathogenesis. *Retrovirology* **19**, article no. 13 (2022).
13. X. Lian, K. W. Seiger, E. M. Parsons, C. Gao, W. Sun, G. T. Gladkov, I. C. Roseto, K. B. Einkauf, M. R. Osborn, J. M. Chevalier, C. Jiang, J. Blackmer, M. Carrington, E. S. Rosenberg, M. M. Lederman, D. K. McMahon, R. J. Bosch, J. M. Jacobson, R. T. Gandhi, M. J. Peluso, T. W. Chun, S. G. Deeks, X. G. Yu, M. Lichterfeld, Progressive transformation of the HIV-1 reservoir cell profile over two decades of antiviral therapy. *Cell Host Microbe* **31**, 83–96.e5 (2023).
14. M. Griesbeck, E. Scully, M. Altfeld, Sex and gender differences in HIV-1 infection. *Clin. Sci.* **130**, 1435–1451 (2016).
15. S. Rao, Sex differences in HIV-1 persistence and the implications for a cure. *Front. Glob. Womens Health* **3**, 942345 (2022).
16. E. P. Scully, M. Gandhi, R. Johnston, R. Hoh, A. Lockhart, C. Dobrowolski, A. Pagliuzza, J. M. Milush, C. A. Baker, V. Girling, A. Ellefson, R. Gorelick, J. Lifson, M. Altfeld, G. Alter, M. Cedars, A. Solomon, S. R. Lewin, J. Karn, N. Chomont, P. Bacchetti, S. G. Deeks, Sex-based differences in human immunodeficiency virus type 1 reservoir activity and residual immune activation. *J. Infect. Dis.* **219**, 1084–1094 (2019).
17. S. L. Klein, K. L. Flanagan, Sex differences in immune responses. *Nat. Rev. Immunol.* **16**, 626–638 (2016).
18. M. M. Addo, M. Altfeld, Sex-based differences in HIV type 1 pathogenesis. *J. Infect. Dis.* **209** (suppl. 3), S86–S92 (2014).
19. V. Taneja, Sex hormones determine immune response. *Front. Immunol.* **9**, 1931 (2018).
20. M. J. Peluso, P. Bacchetti, K. D. Ritter, S. Beg, J. Lai, J. N. Martin, P. W. Hunt, T. J. Henrich, J. D. Siliciano, R. F. Siliciano, G. M. Laird, S. G. Deeks, Differential decay of intact and defective proviral DNA in HIV-1-infected individuals on suppressive antiretroviral therapy. *JCI Insight* **5**, e132997 (2020).
21. N. N. Hosmane, K. J. Kwon, K. M. Bruner, A. A. Capoferri, S. Beg, D. I. Rosenbloom, B. F. Keele, Y. C. Ho, J. D. Siliciano, R. F. Siliciano, Proliferation of latently infected $CD4^+$ T cells carrying replication-competent HIV-1: Potential role in latent reservoir dynamics. *J. Exp. Med.* **214**, 959–972 (2017).
22. K. B. Einkauf, G. Q. Lee, C. Gao, R. Sharaf, X. Sun, S. Hua, S. M. Y. Chen, C. Jiang, X. Lian, F. Z. Chowdhury, E. S. Rosenberg, T.-W. Chun, J. Z. Li, X. G. Yu, M. Lichterfeld, Intact HIV-1 proviruses accumulate at distinct chromosomal positions during prolonged antiretroviral therapy. *J. Clin. Invest.* **129**, 988–998 (2019).
23. M. K. Lewinski, D. Bisgrove, P. Shinn, H. Chen, C. Hoffmann, S. Hannenhalli, E. Verdin, C. C. Berry, J. R. Ecker, F. D. Bushman, Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. *J. Virol.* **79**, 6610–6619 (2005).
24. M. J. Vogel, L. Guelen, E. de Wit, D. Peric-Hupkes, M. Lodén, W. Talhout, M. Feenstra, B. Abbas, A. K. Classen, B. van Steensel, Human heterochromatin proteins form large domains containing KRAB-ZNF genes. *Genome Res.* **16**, 1493–1504 (2006).
25. I. A. Tchasovnikarova, R. T. Timms, N. J. Matheson, K. Wals, R. Antrobus, B. Göttgens, G. Dougan, M. A. Dawson, P. J. Lehner, Epigenetic silencing by the HUSH complex mediates position-effect variegation in human cells. *Science* **348**, 1481–1485 (2015).
26. P. M. Pereira, A. Schneider, S. Pannetier, D. Heron, A. Hanauer, Coffin-Lowry syndrome. *Eur. J. Hum. Genet.* **18**, 627–633 (2010).
27. G. Vansant, H. C. Chen, E. Zorita, K. Trejbalová, D. Miklík, G. Filion, Z. Debyser, The chromatin landscape at the HIV-1 provirus integration site determines viral expression. *Nucleic Acids Res.* **48**, 7801–7817 (2020).
28. S. S. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, E. L. Aiden, A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
29. E. Serrao, P. Cherepanov, A. N. Engelman, Amplification, next-generation sequencing, and genomic DNA mapping of retroviral integration sites. *J. Vis. Exp.* **2016**, e53840 (2016).
30. E. Martin-Gayo, M. B. Cole, K. E. Kolb, Z. Ouyang, J. Cronin, S. W. Kazer, J. Ordovas-Montanes, M. Lichterfeld, B. D. Walker, N. Yosef, A. K. Shalek, X. G. Yu, A reproducibility-based computational framework identifies an inducible, enhanced antiviral state in dendritic cells from HIV-1 elite controllers. *Genome Biol.* **19**, 10 (2018).
31. I. Lisovsky, G. Isitman, R. Song, S. DaFonseca, A. Tremblay-McLean, B. Lebouché, J. P. Routy, J. Bruneau, N. F. Bernard, A Higher frequency of NKG2A+ than of NKG2A- NK cells responds to autologous HIV-infected CD4 cells irrespective of whether or not they coexpress KIR3DL1. *J. Virol.* **89**, 9909–9919 (2015).
32. Z. B. Davis, A. Cogswell, H. Scott, A. Mertsching, J. Boucau, D. Wambua, S. Le Gall, V. Planelles, K. S. Campbell, E. Barker, A conserved HIV-1-derived peptide presented by HLA-E renders infected T-cells highly susceptible to attack by NKG2A/CD94-bearing natural killer cells. *PLOS Pathog.* **12**, e1005421 (2016).
33. E. Brunetta, M. Fogli, S. Varchetta, L. Bozzo, K. L. Hudspeth, E. Marcenaro, A. Moretta, D. Mavilio, Chronic HIV-1 viremia reverses NKG2A/NKG2C ratio on natural killer cells in patients with human cytomegalovirus co-infection. *AIDS* **24**, 27–34 (2010).
34. L. Romero-Martín, C. Duran-Castells, M. Olivella, M. Rosás-Umbert, M. Ruiz-Riol, J. Sanchez, D. Hartigan-O'Connor, B. Mothe, À. Olvera, C. Brander, Disruption of the HLA-E/NKG2X axis is associated with uncontrolled HIV infections. *Front. Immunol.* **13**, 1027855 (2022).
35. R. Thomas, X. Yang, NK-DC crosstalk in immunity to microbial infection. *J. Immunol. Res.* **2016**, 6374379 (2016).
36. M. Gandhi, P. Bacchetti, P. Miotti, T. C. Quinn, F. Veronese, R. M. Greenblatt, Does patient sex affect human immunodeficiency virus levels? *Clin. Infect. Dis.* **35**, 313–322 (2002).
37. S. Napravnik, C. Poole, J. C. Thomas, J. J. Eron Jr., Gender difference in HIV RNA levels: A meta-analysis of published studies. *J. Acquir. Immune Defic. Syndr.* **31**, 11–19 (2002).
38. A. Trickey, C. A. Sabin, G. Burkholder, H. Crane, A. d'Arminio Monforte, M. Egger, M. J. Gill, S. Grabar, J. L. Guest, I. Jarrin, F. C. Lampe, N. Obel, J. M. Reyes, C. Stephan, T. R. Sterling, R. Teira, G. Touloumi, J. C. Wasmuth, F. Wit, L. Wittkop, R. Zangerle, M. J. Silverberg, A. Justice, J. A. C. Sterne, Life expectancy after 2015 of adults with HIV on long-term antiretroviral therapy in Europe and North America: A collaborative analysis of cohort studies. *Lancet HIV* **10**, e295–e307 (2023).
39. C. N. Le, P. Britto, S. S. Brummel, R. M. Hoffman, J. Z. Li, P. M. Flynn, T. E. Taha, A. Coletti, M. G. Fowler, R. J. Bosch, R. T. Gandhi, K. L. Klingman, J. A. McIntyre, J. S. Currier, Time to viral rebound and safety after antiretroviral treatment interruption in postpartum women compared with men. *AIDS* **33**, 2149–2156 (2019).
40. S. S. D. de Azevedo, D. G. Caetano, F. H. Côrtes, S. L. M. Teixeira, K. dos Santos Silva, B. Hoagland, B. Grinsztejn, V. G. Veloso, M. G. Morgado, G. Bello, Highly divergent patterns of genetic diversity and evolution in proviral quasispecies from HIV controllers. *Retrovirology* **14**, 29 (2017).
41. C. Goujard, I. Girault, C. Rouzioux, C. Lécuroux, C. Deveau, M. L. Chaix, C. Jacomet, A. Talamali, J. F. Delfraissy, A. Venet, L. Meyer, M. Sinet, HIV-1 control after transient antiretroviral treatment initiated in primary infection: Role of patient characteristics and effect of therapy. *Antivir. Ther.* **17**, 1001–1009 (2012).
42. G. Turk, K. Seiger, X. Lian, W. Sun, E. M. Parsons, C. Gao, Y. Rassadkina, M. L. Polo, A. Czernikier, Y. Ghiglione, A. Vellicce, J. Varriale, J. Lai, Y. Yuki, M. Martin, A. Rhodes, S. R. Lewin, B. D. Walker, M. Carrington, R. Siliciano, J. Siliciano, M. Lichterfeld, N. Laufer, X. G. Yu, A possible sterilizing cure of HIV-1 infection without stem cell transplantation. *Ann. Intern. Med.* **175**, 95–100 (2022).
43. S. Gianella, S. M. Tran, S. Morris, M. Vargas, M. Porrachia, M. F. Oliveira, S. Lada, M. Zhao, G. B. Ellsworth, J. S. Mathad, T. Wilkin, Sex differences in CMV replication and HIV persistence during suppressive ART. *Open Forum Infect. Dis.* **7**, ofaa289 (2020).

44. J. L. Prodger, A. A. Capoferri, K. Yu, J. Lai, S. J. Reynolds, J. Kasule, T. Kityamuweesi, P. Buule, D. Serwadda, K. J. Kwon, K. Schlusser, C. Martens, E. Scully, Y. H. Choi, A. D. Redd, T. C. Quinn, Reduced HIV-1 latent reservoir outgrowth and distinct immune correlates among women in Rakai, Uganda. *JCI Insight* **5**, e139287 (2020).
45. E. L. Szotek, S. D. Narasipura, L. Al-Harthi, 17β-Estradiol inhibits HIV-1 by inducing a complex formation between β-catenin and estrogen receptor α on the HIV promoter to suppress HIV transcription. *Virology* **443**, 375–383 (2013).
46. V. S. Ratnu, M. R. Emami, T. W. Bredy, Genetic and epigenetic factors underlying sex differences in the regulation of gene expression in the brain. *J. Neurosci. Res.* **95**, 301–310 (2017).
47. Y. C. Ho, L. Shan, N. N. Hosmane, J. Wang, S. B. Laskey, D. I. Rosenbloom, J. Lai, J. N. Blankson, J. D. Siliciano, R. F. Siliciano, Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell* **155**, 540–551 (2013).
48. M. M. K. Hansen, B. Martin, L. S. Weinberger, HIV latency: Stochastic across multiple scales. *Cell Host Microbe* **26**, 703–705 (2019).
49. A. Al-Attar, S. R. Presnell, C. A. Peterson, D. T. Thomas, C. T. Lutz, The effect of sex on immune cells in healthy aging: Elderly women have more robust natural killer lymphocytes than do elderly men. *Mech. Ageing Dev.* **156**, 25–33 (2016).
50. M. Pujantell, N. T. Skenteris, J. M. Claussen, B. Grünhagel, R. J. Thiele, M. Altfeld, Sex-dependent differences in type I IFN-induced natural killer cell activation. *Front. Immunol.* **14**, 1277967 (2023).
51. L. Mensching, A. Hoelzemer, N. K. Cells, Monocytes and macrophages in HIV-1 control: Impact of innate immune responses. *Front. Immunol.* **13**, 883728 (2022).
52. M. I. Cheng, J. H. Li, L. Riggan, B. Chen, R. Y. Tafti, S. Chin, F. Ma, M. Pellegrini, H. Hrncir, A. P. Arnold, T. E. O'Sullivan, M. A. Su, The X-linked epigenetic regulator UTX controls NK cell-intrinsic sex differences. *Nat. Immunol.* **24**, 780–791 (2023).
53. W. Sun, C. Gao, G. T. Gladkov, I. Roseto, L. Carrere, E. M. Parsons, C. Gasca-Capote, J. Frater, S. Fidler, X. G. Yu, M. Lichterfeld, RIVER Trial Study Group, Footprints of innate immune activity during HIV-1 reservoir cell evolution in early-treated infection. *J. Exp. Med.* **221**, e20241091 (2024).
54. M. Statzu, W. Jin, E. J. Fray, A. K. H. Wong, M. R. Kumar, E. Ferrer, S. S. Docken, M. Pinkevych, J. B. McBrien, C. M. Fennessey, B. F. Keele, S. Liang, J. L. Harper, S. Mutascio, L. Franchitti, H. Wang, D. Cicetti, S. E. Bosinger, D. G. Carnathan, T. H. Vanderford, D. M. Margolis, J. V. Garcia-Martinez, A. Chahroudi, M. Paiardini, J. Siliciano, M. P. Davenport, D. A. Kulpa, R. S. Siliciano, G. Silvestri, $CD8^+$ lymphocytes do not impact SIV reservoir establishment under ART. *Nat. Microbiol.* **8**, 299–308 (2023).
55. B. Vrancken, G. Baele, A. M. Vandamme, K. van Laethem, M. A. Suchard, P. Lemey, Disentangling the impact of within-host evolution and transmission dynamics on the tempo of HIV-1 evolution. *AIDS* **29**, 1549–1556 (2015).
56. T. R. Sterling, D. Vlahov, J. Astemborski, D. R. Hoover, J. B. Margolick, T. C. Quinn, Initial plasma HIV-1 RNA levels and progression to AIDS in women and men. *N. Engl. J. Med.* **344**, 720–725 (2001).
57. M. A. De Scheerder, B. Vrancken, S. Dellicour, T. Schlub, E. Lee, W. Shao, S. Rutsaert, C. Verhofstede, T. Kerre, T. Malfait, D. Hemelsoet, M. Coppens, A. Dhondt, D. De Looze, F. Vermassen, P. Lemey, S. Palmer, L. Vandekerckhove, HIV rebound is predominantly fueled by genetically identical viral expansions from diverse reservoirs. *Cell Host Microbe* **26**, 347–358.e7 (2019).
58. W. Sun, Y. Rassadkina, C. Gao, S. I. Collens, X. Lian, I. H. Solomon, S. S. Mukerji, X. G. Yu, M. Lichterfeld, Persistence of intact HIV-1 proviruses in the brain during antiretroviral therapy. *eLife* **12**, RP89837 (2023).
59. K. M. Bruner, Z. Wang, F. R. Simonetti, A. M. Bender, K. J. Kwon, S. Sengupta, E. J. Fray, S. A. Beg, A. A. R. Antar, K. M. Jenike, L. N. Bertagnolli, A. A. Capoferri, J. T. Kufera, A. Timmons, C. Nobles, J. Gregg, N. Wada, Y. C. Ho, H. Zhang, J. B. Margolick, J. N. Blankson, S. G. Deeks, F. D. Bushman, J. D. Siliciano, G. M. Laird, R. F. Siliciano, A quantitative approach for measuring the reservoir of latent HIV-1 proviruses. *Nature* **566**, 120–125 (2019).
60. B. Cole, L. Lambrechts, P. Gantner, Y. Noppe, N. Bonine, W. Witkowski, L. Chen, S. Palmer, J. I. Mullins, N. Chomont, M. Pardons, L. Vandekerckhove, In-depth single-cell analysis of translation-competent HIV-1 reservoirs identifies cellular sources of plasma viremia. *Nat. Commun.* **12**, 3727 (2021).
61. G. Q. Lee, K. Reddy, K. B. Einkauf, K. Gounder, J. M. Chevalier, K. L. Dong, B. D. Walker, X. G. Yu, T. Ndung'u, M. Lichterfeld, HIV-1 DNA sequence diversity and evolution during acute subtype C infection. *Nat. Commun.* **10**, 2737 (2019).
62. P. P. Rose, B. T. Korber, Detecting hypermutations in viral sequences with an emphasis on G → A hypermutation. *Bioinformatics* **16**, 400–401 (2000).
63. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
64. S. Guindon, J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel, New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
65. J. M. Carlson, C. J. Brumme, E. Martin, J. Listgarten, M. A. Brockman, A. Q. Le, C. K. S. Chui, L. A. Cotton, D. J. H. F. Knapp, S. A. Riddler, R. Haubrich, G. Nelson, N. Pfeifer, C. E. DeZiel, D. Heckerman, R. Apps, M. Carrington, S. Mallal, P. R. Harrigan, M. John, Z. L. Brumme, International HIV Adaptation Collaborative, Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. *J. Virol.* **86**, 13202–13216 (2012).
66. C. A. Bricault, K. Yusim, M. S. Seaman, H. Yoon, J. Theiler, E. E. Giorgi, K. Wagh, M. Theiler, P. Hraber, J. P. Macke, E. F. Kreider, G. H. Learn, B. H. Hahn, J. F. Scheid, J. M. Kovacs, J. L. Shields, C. L. Lavine, F. Ghantous, M. Rist, M. G. Bayne, G. H. Neubauer, K. McMahan, H. Peng, C. Chéneau, J. J. Jones, J. Zeng, C. Ochsenbauer, J. P. Nkolola, K. E. Stephenson, B. Chen, S. Gnanakaran, M. Bonsignori, L. D. Williams, B. F. Haynes, N. Doria-Rose, J. R. Mascola, D. C. Montefiori, D. H. Barouch, B. Korber, HIV-1 neutralizing antibody signatures and application to epitope-targeted vaccine design. *Cell Host Microbe* **26**, 296 (2019).
67. T. A. Wagner, S. McLaughlin, K. Garg, C. Y. K. Cheung, B. B. Larsen, S. Styrchak, H. C. Huang, P. T. Edlefsen, J. I. Mullins, L. M. Frenkel, Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science* **345**, 570–573 (2014).
68. A. Kundaje, W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M. J. Ziller, V. Amin, J. W. Whitaker, M. D. Schultz, L. D. Ward, A. Sarkar, G. Quon, R. S. Sandstrom, M. L. Eaton, Y. C. Wu, A. R. Pfenning, X. Wang, M. Claussnitzer, Y. Liu, C. Coarfa, R. A. Harris, N. Shoresh, C. B. Epstein, E. Gjoneska, D. Leung, W. Xie, R. D. Hawkins, R. Lister, C. Hong, P. Gascard, A. J. Mungall, R. Moore, E. Chuah, A. Tam, T. K. Canfield, R. S. Hansen, R. Kaul, P. J. Sabo, M. S. Bansal, A. Carles, J. R. Dixon, K. H. Farh, S. Feizi, R. Karlic, A. R. Kim, A. Kulkarni, D. Li, R. Lowdon, G. Elliott, T. R. Mercer, S. J. Neph, V. Onuchic, P. Polak, N. Rajagopal, P. Ray, R. C. Sallari, K. T. Siebenthall, N. A. Sinnott-Armstrong, M. Stevens, R. E. Thurman, J. Wu, B. Zhang, X. Zhou, A. E. Beaudet, L. A. Boyer, P. L. De Jager, P. J. Farnham, S. J. Fisher, D. Haussler, S. J. Jones, W. Li, M. A. Marra, M. T. McManus, S. Sunyaev, J. A. Thomson, T. D. Tlsty, L. H. Tsai, W. Wang, R. A. Waterland, M. Q. Zhang, L. H. Chadwick, B. E. Bernstein, J. F. Costello, J. R. Ecker, M. Hirst, A. Meissner, A. Milosavljevic, B. Ren, J. A. Stamatoyannopoulos, T. Wang, M. Kellis, Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).


**Acknowledgments:** We gratefully acknowledge all study participants, the MGH DNA core facility, and the dedication of the staff at the MWCCS (MACS/WIHS Combined Cohort Study) sites. **Funding:** X.G.Y. is supported by the American Foundation for AIDS Research (amfAR, no. 110393-72-RPRL), by NIH grants DA047034, MH134823, and AI155171, and the Bill and Melinda Gates Foundation (INV-002703). M. Lichterfeld is supported by NIH grants AI184094, AI152979, AI176579, and AI155233. X.G.Y. and M. Lichterfeld are members of the DARE, ERASE, PAVE, and BEAT-HIV Martin Delaney Collaboratory (UM1 AI164560, AI164562, AI164566, and AI164570). M.J.P. is supported by NIH grant no. K23AI157875. This project has been funded in whole or in part with federal funds from the Frederick National Laboratory for Cancer Research, under contract no. 75N91024F00011. This research was supported in part by the Intramural Research Program of the NIH, Frederick National Lab, Center for Cancer Research. Data in this manuscript were collected by the Women's Interagency HIV Study (WIHS), now the MACS/WIHS Combined Cohort Study (MWCCS). P.C.T. was supported by the Northern California site and by U01-HL146242. S.K. was supported by the Metropolitan Washington site and by U01-HL146205 with additional support from the Data Analysis and Coordination Center of U01-HL146193. The contents of this publication are solely the responsibility of the authors and do not represent the official views of the NIH and the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government. The MWCCS is funded primarily by the National Heart, Lung, and Blood Institute (NHLBI), with additional cofunding from the National Institute of Child Health and Human Development (NICHD), National Institute on Aging (NIA), National Institute of Dental and Craniofacial Research (NIDCR), National Institute of Allergy and Infectious Diseases (NIAID), National Institute of Neurological Disorders and Stroke (NINDS), National Institute of Mental Health (NIMH), National Institute on Drug Abuse (NIDA), National Institute of Nursing Research (NINR), National Cancer Institute (NCI), National Institute on Alcohol Abuse and Alcoholism (NIAAA), National Institute on Deafness and Other Communication Disorders (NIDCD), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute on Minority Health and Health Disparities (NIMHD), and in coordination and alignment with the research priorities of the NIH, Office of AIDS Research (OAR). Data collection from the MWCCS Northern California and Metropolitan Washington site was also supported by UL1-TR000004 (UCSF CTSA), UL1-TR001409 (DC CTSA), KL2-TR001432 (DC CTSA), and TL1-TR001431 (DC CTSA). This work was supported by the National Institute of Allergy and Infectious Diseases (UM1AI068634 to R.J.B. and UM1AI068636 and UM1AI106701 for the contributing ACTG studies). **Author contributions:** M. Lichterfeld and X.G.Y. conceptualized the study. T.S.T., A.S.H., L.V., L.C., S.K., S.H., and M. Lancien performed the whole-genome amplification and HIV-1 sequencing experiments. T.S.T., L.C., C.M.N., S.K.M., and I.C.R. carried out the integration site analysis of in vivo–infected cells. T.S.T., L.V., and B.B. performed the in vitro qVOAs. T.S.T. designed and performed the immunological assays. A.N. assisted in the flow cytometry experiment and related analyses. M.C., M.V., and Y.Y. produced the HLA class I and KIR genotyping data. R.H., S.G.D., D.K.M., R.J.B., S.G.K., R.T.G., P.C.T., M.J.P., J.M.J., and M. Lichterfeld contributed source clinical PBMC samples. C.G. conducted bioinformatic analysis. T.S.T., X.L., C.G., M. Lichterfeld, and X.G.Y. provided input for study

design, data analysis, and presentation. M. Lichterfeld and X.G.Y. supervised the study. T.S.T., M. Lichterfeld, and X.G.Y. wrote the original draft. All authors reviewed and edited the manuscript. **Competing interests:** M.J.P. has served on a Data Safety Monitoring Board for American Gene Technologies. M. Lichterfeld has acted as a paid consultant for Merck, ViiV, and MPM BioImpact. X.G.Y. has acted as a paid consultant for MPM BioImpact. The other authors declare that they have no competing interests. **Data and materials availability:** All data associated with this study are present in the paper or the Supplementary Materials. Study participants' viral sequencing data can be made available to academic investigators via a data sharing agreement. Access to individual-level data from the MWCCS may be obtained upon review and approval of a MWCCS concept sheet (https://statepi.jhsph.edu/mwccs/work-with-us/).

Submitted 18 November 2024
Resubmitted 10 February 2025
Accepted 27 August 2025
Published 17 September 2025
10.1126/scitranslmed.adu7154