HIV Diversity and Genetic Compartmentalization in Blood and Testes during Suppressive Antiretroviral Therapy

Rachel L. Miller,^a Rosalie Ponte,^{b,c} Bradley R. Jones,^d Natalie N. Kinloch,^a Fredrick H. Omondi,^a Mohammad-Ali Jenabian,^e Franck P. Dupuy,^{b,c} Remi Fromentin,^f Pierre Brassard,^g Vikram Mehraj,^{b,c,f} Nicolas Chomont,^{f,h} Art F. Y. Poon,ⁱ Jeffrey B. Joy,^{d,j} Zabrina L. Brumme,^{a,d} Jean-Pierre Routy,^{b,c} for the ORCHID Study Group

^aFaculty of Health Sciences, Simon Fraser University, Burnaby, British Columbia, Canada

^bResearch Institute, McGill University Health Centre, Montréal, Quebec, Canada

^cChronic Viral Illness Service and Division of Hematology, McGill University Health Centre, Montréal, Quebec, Canada

^dBritish Columbia Centre for Excellence in HIV/AIDS, Vancouver, British Columbia, Canada

^eDepartment of Biological Sciences, Université du Québec à Montréal, Montréal, Quebec, Canada

^fCentre de Recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Quebec, Canada

^gThe Metropolitan Centre for Plastic Surgery, Montréal, Quebec, Canada

^hDepartment of Microbiology, Infection and Immunology, Université de Montréal, Montréal, Quebec, Canada

ⁱDepartment of Pathology & Laboratory Medicine, Western University, London, Ontario, Canada

^jDepartment of Medicine, University of British Columbia, Vancouver, British Columbia, Canada

ABSTRACT HIV's ability to persist during suppressive antiretroviral therapy is the main barrier to cure. Immune-privileged tissues, such as the testes, may constitute distinctive sites of HIV persistence, but this has been challenging to study in humans. We analyzed the proviral burden and genetics in the blood and testes of 10 individuals on suppressive therapy who underwent elective gender-affirming surgery. HIV DNA levels in matched blood and testes were quantified by quantitative PCR, and subgenomic proviral sequences (*nef* region) were characterized from single templates. HIV diversity, compartmentalization, and immune escape burden were assessed using genetic and phylogenetic approaches. Diverse proviruses were recovered from the blood (396 sequences; 354 *nef*-intact sequences) and testes (326 sequences; 309 *nef*-intact sequences) of all participants. Notably, the frequency of identical HIV sequences varied markedly between and within individuals. Nevertheless, proviral loads, within-host unique HIV sequence diversity, and the immune escape burden correlated positively between blood and testes. When all intact *nef* sequences were evaluated, 60% of participants exhibited significant blood-testis genetic compartmentalization, but none did so when the evaluation was restricted to unique sequences per site, suggesting that compartmentalization, when present, is attributable to the clonal expansion of HIV-infected cells. Our observations confirm the testes as a site of HIV persistence and suggest that individuals with larger and more diverse blood reservoirs will have larger and more diverse testis reservoirs. Furthermore, while the testis microenvironment may not be sufficiently unique to facilitate the seeding of unique viral populations therein, differential clonal expansion dynamics may be at play, which may complicate HIV eradication.

IMPORTANCE Two key questions in HIV reservoir biology are whether immune-privileged tissues, such as the testes, harbor distinctive proviral populations during suppressive therapy and, if so, by what mechanism. While our results indicated that blood-testis HIV genetic compartmentalization was reasonably common (60%), it was always attributable to differential frequencies of identical HIV sequences between sites. No blood-tissue data set retained evidence of compartmentalization when only unique HIV sequences per site were considered; moreover, HIV immune escape mutation burdens were highly concordant between sites. We conclude that the princi-

Citation Miller RL, Ponte R, Jones BR, Kinloch NN, Omondi FH, Jenabian M-A, Dupuy FP, Fromentin R, Brassard P, Mehraj V, Chomont N, Poon AFY, Joy JB, Brumme ZL, Routy J-P, for the ORCHID Study Group. 2019. HIV diversity and genetic compartmentalization in blood and testes during suppressive antiretroviral therapy. *J Virol* 93:e00755-19. <https://doi.org/10.1128/JVI.00755-19>.

Editor Frank Kirchhoff, Ulm University Medical Center

Copyright © 2019 American Society for Microbiology. All Rights Reserved.

Address correspondence to Zabrina L. Brumme, zbrumme@sfu.ca.

R.L.M. and R.P. contributed equally to this work, as did Z.L.B. and J.-P.R.

Received 6 May 2019

Accepted 8 June 2019

Accepted manuscript posted online 12 June 2019

Published 13 August 2019

pal mechanism by which blood and testis reservoirs differ is not via seeding of divergent HIV sequences therein but, rather, via differential clonal expansion of latently infected cells. Thus, while viral diversity and escape-related barriers to HIV eradication are of a broadly similar magnitude across the blood and testes, clonal expansion represents a challenge. The results support individualized analysis of within-host reservoir diversity to inform curative approaches.

KEYWORDS HIV, clonal expansion, diversity, genetic compartmentalization, reservoir, testes

The HIV reservoir, a small pool of primarily CD4⁺ T cells that harbor replication-competent virus despite long-term suppressive combination antiretroviral therapy (cART) (1), is the major barrier to cure. During untreated infection, HIV populations evolve within the host (2–5), and individual viral sequences are continually archived into the reservoir as integrated proviruses (6). There, they can persist long term within either the original infected cell or clonal descendants thereof (7–10). Much of our understanding of HIV persistence is derived from HIV sequences isolated from blood during long-term cART; these studies have revealed that the within-host HIV reservoir is genetically diverse (11–13), that it frequently contains immune escape mutations (14, 15), and that it commonly features clonally expanded cell populations harboring identical proviruses (10, 16–19).

The HIV reservoir in blood, however, may not reflect that within tissues. HIV reservoir composition in immune-privileged sites, defined as those that are protected by physical barriers that limit immune cell trafficking (20) and/or that are equipped with mechanisms to suppress immune responses locally (21, 22), may be distinct from that elsewhere, but this has been challenging to study in healthy humans. The testes represent such a site (21, 23). Nonhuman primate studies have demonstrated that the cytokine responses of testicular T cells to mitogen stimulation are lower than those of blood T cells (24, 25); moreover, human testes harbor higher frequencies of CD39⁺ T-regulatory cells (21), which are capable of suppressing HIV-specific CD8⁺ T-cell responses (26). Despite this, however, leukocyte populations, including central and effector memory CD4⁺ T cells that represent potential viral targets, are nevertheless present in the testes (24). Indeed, immunodeficiency viruses penetrate into the testes early during infection (22, 27) and subsequently persist (22, 28–30). As such, the testes have been hypothesized to represent a distinctive HIV sanctuary site (21, 25).

A key outstanding question is whether proviruses persisting in testes are genetically distinct from those in blood and, if so, how. Blood-tissue genetic compartmentalization could arise in a number of ways (31). It could happen if, during untreated HIV infection, tissue HIV populations replicated under reduced immune pressures and/or experienced limited subsequent mixing with blood; if this were the case, then provirus populations archived into blood and tissue HIV reservoirs would reflect their divergent evolution pre-cART. Genetically distinct HIV reservoirs could also arise as a result of restricted gene flow between compartments; for example, if only a small number of HIV strains initially penetrated into tissue, descendant viral populations in tissue could form nested subclades within the more diverse populations in blood. Differential proviral distributions could also arise via unequal clonal expansion and/or contraction of latently HIV-infected cells in blood and tissue either before or during cART. The latter would result in viral populations that are not different in terms of unique viral lineages but, rather, that are different in terms of their frequencies of particular identical proviruses. Investigating blood-tissue genetic compartmentalization will shed light on the extent to which tissue microenvironments shape latent within-host HIV populations and the extent to which the blood reservoir mirrors that elsewhere in the body, information that is relevant to the design of curative approaches guided by individualized analysis of within-host reservoir composition (32, 33).

While studies in nonhuman primates have compared the genetic structures of within-host simian immunodeficiency virus populations in male genital organs versus

FIG 1 Quantifying and characterizing the HIV reservoir in blood and testes. (A) Representative immunostaining of a frozen testis section (participant 5) showing cells expressing CD3 (red) or CD4 (green). Nuclei were counterstained with DAPI (blue). CD4 T cells (CD3⁺ CD4⁺) appear in yellow in the merged image. (B) HIV DNA loads, expressed in number of copies per million cells in whole-tissue lysate, in left and right testes. ND, not determined; trace, HIV DNA was detected but not quantifiable. The asterisk denotes the fact that HIV DNA loads in testes are measured as copies per million total cells (not per million CD4⁺ T-cells, as for blood). (C) Marginal positive correlation between HIV DNA loads in blood and testes (where the latter is expressed as the average measurements for the left and right testes). (D) Unique HIV sequence distribution in blood and testes, matched by participant. (E) Significant correlation between the HIV DNA load and unique HIV sequence frequency in blood. (F) Lack of a significant correlation between the HIV DNA load and unique HIV sequence frequency in testes.

those of populations in blood (28, 34), no studies, to our knowledge, have investigated this in healthy, HIV-infected humans. We thus investigated the within-host proviral burden, genetic diversity, and compartmentalization in 10 HIV-infected, cART-suppressed individuals undergoing elective bilateral orchiectomy for gender affirmation surgery (21, 23).

RESULTS

HIV characterization in blood and testes. At the time of blood and tissue sampling, all participants had maintained plasma viremia suppression on cART for ≥ 6 months. At this time, the blood HIV DNA levels in the participants ranged from 11 to 4,003 (median, 512; interquartile range [IQR], 100 to 1,111) HIV DNA copies/ 10^6 CD4⁺ T cells. The presence of CD4⁺ T cells in the testes was confirmed by immunostaining (Fig. 1A). The testis HIV reservoir size, quantified in terms of the number of viral copies per total cell equivalent (CD4⁺ T cells were not isolated from tissue), ranged from trace quantities to 198 (median, 3; IQR, 0.4 to 10) copies/ 10^6 cells in the right testis and trace quantities to 39.2 (median, 1; IQR, 0.4 to 15) copies/ 10^6 cells in the left testis (Fig. 1B).

TABLE 1 Participant HLA and HIV sequence characteristics

Participant	Allele			HIV subtype	Total no. of intact sequences (total no. of unique sequences, % unique sequences) from:	
	HLA-A	HLA-B	HLA-C		Blood	Testes
1	A*32:01-A*34:02	B*15:18-B*52:01	C*07:04-C*16:01	B	30 (21, 70)	99 (8, 8)
2	A*23:01-A*30:01	B*07:02-B*42:01	C*07:02-C*17:01	B	26 (2, 8)	9 (7, 78)
3	A*02:01-A*25:01	B*08:01-B*18:01	C*07:01-C*12:03	B	36 (10, 28)	40 (24, 60)
4	A*11:01-A*24:02	B*15:02-B*35:03	C*08:01-C*12:02	B	31 (20, 65)	15 (12, 80)
5	A*02:01-A*26:01	B*15:01-B*38:01	C*03:04-C*12:03	B	31 (20, 65)	84 (21, 25)
6	A*33:01-A*68:01	B*14:02-B*44:02	C*07:04-C*08:02	B	45 (40, 89)	16 (15, 94)
7	A*24:02-A*29:02	B*18:25-B*27:05	C*01:02-C*12:03	B	37 (16, 43)	13 (9, 69)
8	A*02:01-A*02:01	B*08:01-B*44:02	C*05:01-C*07:01	B	41 (23, 56)	8 (5, 63)
9	A*02:01-A*11:01	B*18:01-B*27:04	C*07:04-C*12:02	B	45 (25, 56)	21 (8, 38)
10	A*26:01-A*26:01	B*38:01-B*51:01	C*12:03-C*15:02	C	32 (7, 22)	4 (3, 75)
Total					354 (184, 52)	309 (112, 36)

The HIV DNA loads in the right and left testes correlated significantly between participants (Spearman's $r = 0.73$, $P = 0.03$; not shown). HIV DNA levels in blood and testes, calculated as the mean of the right and left sections, also correlated marginally between participants (Spearman's $r = 0.54$, $P = 0.1$) (Fig. 1C).

In total, 722 subgenomic HIV sequences (*nef*) were isolated by single-genome amplification from blood ($n = 396$) and at least one tissue section ($n = 326$) for all 10 participants, even those for whom testis HIV DNA loads were below the limit of quantification (participants 7 and 10). Even though different tissue sections were used for HIV DNA quantification and sequence recovery, testis HIV DNA levels correlated with the total number of HIV sequences retrieved from $\sim 100 \mu\text{g}$ tissue (Spearman's $r = 0.61$; $P = 0.07$; not shown), suggesting the absence of major location-based sampling biases. After removal of 50 defective and/or hypermutated sequences and 9 putative within-host recombinants, 663 intact *nef* sequences remained: 354 (26 to 45 per participant; median, 34; IQR, 31 to 40) from blood and 309 (4 to 99 per participant; median, 16; IQR, 10 to 35) from testes (Table 1). For eight participants (participants 1 to 5, 7, 9, and 10), intact *nef* sequences were isolated from both the right and left testes; however, for participant 6, only the right testis yielded sequences, and for participant 8, only right testis sections were available for analysis. Due to the modest number of testis *nef* sequences recovered for some participants, our primary analysis combined all within-host testis sequences together regardless of sampling location, though we tested the validity of this assumption in participants for whom sufficient data were available (see below).

All participants harbored both unique HIV *nef* sequences (observed only once) and identical *nef* sequences (observed more than once), but at markedly different frequencies (Table 1). This could be visualized by plotting each participant's total number of unique sequences as a function of the total number of sequences collected (Fig. 2). While for some participants, notably, participant 6, unique sequences continued to be recovered during sampling, for many others (e.g., participant 1), ongoing sampling largely yielded sequences that had been recovered previously. Notably, there was no significant relationship between the frequency of unique sequences recovered from blood and testes within a given participant (Wilcoxon signed-rank test, $P = 0.5$; Fig. 1D). Participants 1 and 2 represented the most extreme examples of this: whereas 70% of HIV *nef* sequences isolated from participant 1's blood were unique, the sequences isolated from testes were almost all (92%) identical, whereas for participant 2 the opposite was true (Table 1). This indicates that one cannot generalize, based on location alone, where identical HIV sequences are more likely to be observed. In blood, we observed a highly significant positive correlation between HIV DNA levels and unique HIV *nef* sequence frequency (Spearman's $r = 0.90$, $P = 0.0008$; Fig. 1E). This supports the notion that larger blood reservoirs tend to be more genetically diverse (as

FIG 2 Relationship between the total number of unique HIV sequences collected per participant as a function of the total number of sequences collected. The dashed line denotes a hypothetical data set where every sampled sequence is unique.

opposed to more highly clonally expanded). In contrast, no significant correlation between HIV DNA levels and unique *nef* sequence frequency was observed in testes; in fact, the relationship trended in the opposite direction (Spearman's $r = -0.49$, $P = 0.2$; Fig. 1F). This indicates that, in contrast to blood, the relationship between reservoir size and diversity in testes is less straightforward.

A phylogeny inferred from an alignment comprising all intact HIV *nef* sequences plus select subtype reference strains confirmed that each participant's sequences formed monophyletic clades with high ($\geq 99.7\%$) bootstrap support and revealed that participants 1 to 9 harbored HIV subtype B, while participant 10 harbored subtype C (Fig. 3). Though all within-host phylogenies featured both unique and identical HIV *nef* sequences, tree topology and diversity varied markedly between individuals (Fig. 4). Participant 2's phylogeny was remarkable, in that 25 of the 26 *nef* sequences isolated from blood were identical, but those isolated from testes were diverse (though the other lone unique blood *nef* sequence was also quite distant from the others) (Fig. 4A). Of note, the abundant blood *nef* sequence was also recovered from both the right and left testes; in addition, there was a second *nef* sequence that was independently recovered from both testes. Participant 6 harbored diverse HIV *nef* sequences within both blood and the right testis (no sequences were recovered from the left testis) and exhibited the highest within-host diversity overall (Fig. 4B). Nevertheless, identical *nef* sequences, including three examples of identical sequences isolated from both blood and the right testis, were identified. Participant 5 harbored multiple closely related monotypic clades, including one sequence that was abundant in both testes (and that was detected in independent samplings thereof) and that was also present in blood, alongside other diverse sequences in blood and tissue (Fig. 4C). Participants 1 and 9 were notable in that they both featured a large, genetically distinct monotypic clade detected solely in testes (in participant 1, this sequence was detected in both testes, while in participant 9, it was detected in the right testis only), yet both also featured other identical HIV *nef* sequences that were isolated from both blood and testes (Fig. 4D and E). Participant 8 featured six clusters of identical sequences, four of which were detected solely in blood and two of which were detected in both blood and the right testis (a left testis section was not available for this participant) (Fig. 4F). Participant 3 harbored four clusters of identical sequences: one large one that was found predominantly in blood but that was also recovered from the right testis and three others that were exclusively observed in either blood or tissue (including one sequence observed

FIG 3 Between-host HIV phylogeny. The numbers on the branches indicate bootstrap values supporting within-host monophyletic clades. The scale is the estimated number of substitutions per nucleotide site. NL4-3 and MJ4 are subtype B and C reference sequences, respectively.

in both the right and left testes) (Fig. 4G). Of all individuals studied, participants 4, 7, and 10 exhibited the most limited within-host diversity (participant 10's proviruses were the least diverse of all, suggestive of early cART initiation; of note, this individual, the only one for whom some clinical information was available, had been diagnosed with infection only 3 years earlier) (Fig. 4H to J). Nevertheless, like all others studied, participants 4, 7, and 10 also featured at least one identical HIV *nef* sequence isolated from both blood and testes.

We next investigated the extent to which the HIV DNA diversity in the testes correlated with that in blood. Because identical sequences artificially lower average pairwise HIV diversity measures (and our data already indicated that identical sequence distributions varied markedly between compartments), we collapsed identical sequences down to a single copy per compartment and computed median within-host patristic (tip-to-tip phylogenetic) distances between all pairs of sequences sampled from each site. Median within-host patristic distances were 2.25×10^{-2} (IQR, 1.83×10^{-2} to 3.70×10^{-2}) substitutions per nucleotide site in blood and 2.33×10^{-2} (IQR, 1.79×10^{-2} to 4.05×10^{-2}) substitutions per nucleotide site in testes, values that correlated significantly with one another (Spearman's $r = 0.94$, $P = 0.0002$; Fig. 5). These results indicate that the diversity of unique HIV sequences in the testes generally reflects that in blood.

HIV genetic compartmentalization. Our major objective was to assess whether proviral sequences in the blood and testes exhibit significant population structure or genetic compartmentalization. As the presence of identical sequences influences compartmentalization detection (especially if their distribution varies markedly across sites),

FIG 4 Within-host HIV reservoir diversity in blood and testes. Phylogenies, inferred from nucleotide sequence alignments (HIV *nef*), are midpoint rooted, with scales denoting the estimated number of substitutions per nucleotide site. Matched highlighter plots, made from amino acid sequence alignments, show substitutions relative to the master sequence (the top sequence in the phylogeny). Symbols denote sampling location: blood (filled circles), right testis (open squares), and left testis (open diamonds). The red A and black B markers for participants 1, 3, 4, 5, 7, 8, and 9 identify sequences collected from independent sections of the same testis (the HIV genetic compartmentalization results for these within-tissue comparisons are summarized in Table 3). Bootstrap values of between 70 and 90% are reported to the left of their respective nodes; those of >90% are marked with asterisks.

FIG 4 (Continued)

FIG 4 (Continued)

FIG 5 Significant positive correlation between blood and testes proviral diversity. Values represent average within-host patristic (tip-to-tip phylogenetic) distances between all pairs of sequences sampled from each site, after collapsing identical *nef* sequences to a single copy per compartment.

we analyzed the data in two ways: by including all sequences (overall) and by collapsing identical sequences down to a single copy per compartment (unique).

When assessing within-host data sets overall, both distance- and tree-based tests agreed that 6 of 10 participants (participants 1 to 5 and 9) displayed significant blood-testis HIV genetic compartmentalization, while participants 7 and 8 did not (Fig. 6A and Table 2). For participants 6 and 10, the distance-based test did not detect compartmentalization, but the tree-based test did; as per our predefined criteria, these participants' overall data sets were deemed not compartmentalized. Notably, all six participants whose overall data sets were deemed compartmentalized harbored at least one abundant identical HIV *nef* sequence that predominated or that was found exclusively in one compartment (Fig. 4).

Importantly, however, when the analysis was restricted to unique *nef* sequences per site, no participant retained consistent evidence of compartmentalization between blood and testes (Fig. 6B and Table 2). These observations are notable for two reasons. First, the observation that 40% of the within-host data sets lacked any evidence of significant HIV population structure between blood and testes overall indicates that the HIV proviral composition in the testes is not always distinct from that in blood. Second, the observation that no participants exhibited evidence of blood-testis HIV compartmentalization when the analysis was restricted to unique *nef* sequences per compartment strongly suggests that HIV population structure, when present, is driven by the differential distribution of identical sequences across blood and tissue and not by the presence of distinct HIV lineages within these sites.

Our primary analyses combined all sequences from a given site together. However, given that only a limited amount of biological material was analyzed, sampling biases are a potential concern. For participants 5, 6, and 7, independent peripheral blood mononuclear cell (PBMC) aliquots were received and analyzed at different times, allowing us to investigate whether independent sampling of HIV *nef* sequences from blood yielded consistent distributions. It did in all cases: proviruses isolated from independent PBMC aliquots exhibited no evidence of compartmentalization by any method (Table 3), supporting their sampling from the same overall gene pool. Moreover, for seven participants (participants 1, 3 to 5, and 7 to 9), sufficient HIV sequences were isolated from different sections of the same testis (see the tip label annotations in Fig. 4), yielding eight opportunities to investigate whether independent samplings of the same site yielded consistent HIV sequence distributions (for participant 5, independent samplings of both the right and left testes yielded sufficient sequences for evaluation). Of these eight within-testis data sets, consistent evidence of overall com-

FIG 6 Genetic compartmentalization results for each participant. “Distance” denotes the results of the Hudson, Boos, and Kaplan nonparametric test for population structure, “Tree” denotes the results of the Slatkin-Maddison test (57), and “Consensus” denotes the final result (compartmentalization was declared only if the results of both tests agreed). Blue denotes compartmentalization; gray denotes no compartmentalization. (A) Overall results when comparing all *nef* sequences from blood to all those from testes; (B) unique results with identical *nef* sequences collapsed to a single copy per compartment.

partmentalization was detected in only one: the right testis of participant 7 (Table 3). Note that consistent support for within-tissue compartmentalization in participant 7’s right testis did not remain after the analysis was restricted to unique sequences (not shown), though limited statistical power is acknowledged. Phylogenetic segregation of the HIV sequences retrieved from these two independent samplings of the right testis is indeed apparent (Fig. 4J); however, further sampling of additional tissue sections yielded HIV sequences that intermixed with these, and it is important to underscore that there was no support for compartmentalization between blood and testes within this participant (Fig. 6). Together, these observations suggest that HIV sampling from blood and testes was not majorly biased by the specific aliquot or tissue section analyzed.

For seven participants (participants 1 to 5, 7, and 9), we had sufficient HIV *nef* sequences to investigate compartmentalization between the right and left testes. When analyzing overall within-host data sets, only participant 9 consistently exhibited between-testis compartmentalization (Table 2). However, after restricting the analysis to only unique *nef* sequences per site, no consistent evidence of compartmentalization remained. Consistent with the results of our analyses of HIV sequences in blood and tissue, these results suggest that in the rare cases where population structure was observed between proviral populations in the right and left testes, this was driven by the differential distribution of identical sequences between sites.

HIV immune escape in blood and testes. Given that cellular immune responses are suppressed in the testes (21, 23), we wished to investigate whether blood and

TABLE 2 Blood-testis and right-left testis HIV genetic compartmentalization results

Comparison	Test	Statistic	Result for the following participant ^c :									
			1	2	3	4	5	6	7	8	9	10
Blood vs testes overall	Distance	K_{ST} ^a	0.61	0.53	0.09	0.06	0.17	0.01	0.002	0.02	0.26	0.11
		<i>P</i> value	**	**	**	*	**	NS	NS	NS	**	NS
	Tree	No. of migrations ^b	7	4	16	9	18	10	12	7	7	3
		<i>P</i> value	**	**	**	*	**	**	NS	NS	**	*
Blood vs testes unique	Distance	K_{ST}	0.10	0	0.04	0.01	0.004	0.01	0	0.02	0.10	0
		<i>P</i> value	NS	NS	NS	NS	NS	NS	NS	NS	*	NS
	Tree	No. of migrations	3	2	8	9	15	10	9	4	6	2
		<i>P</i> value	**	NS	NS	NS	NS	*	NS	NS	NS	NS
Right testis vs left testis overall	Distance	K_{ST}	0.14	0.02	0.03	0.05	0		0		0.65	
		<i>P</i> value	*	NS	NS	NS	NS		NS		**	
	Tree	No. of migrations	38	3	9	1	30		4		2	
		<i>P</i> value	NS	NS	NS	*	NS		NS		**	
Right testis vs left testis unique	Distance	K_{ST}	0	0.02	0	0.07	0		0		0.21	
		<i>P</i> value	NS	NS	NS	NS	NS		NS		*	
	Tree	No. of migrations	3	3	9	1	9		4		2	
		<i>P</i> value	NS	NS	NS	*	NS		NS		NS	

^a K_{ST} , statistic produced by the Hudson, Boos, and Kaplan nonparametric test for population structure, where 0 indicates no compartmentalization and 1 indicates complete compartmentalization.

^bThe number of migrations between compartments required to explain the input phylogeny, as defined by the Slatkin-Maddison test (57). The smaller the number, the stronger the support for compartmentalization.

^c*, 0.01 < *P* < 0.05; **, *P* < 0.01; NS, not significant.

testis HIV sequences differed with respect to their immune escape mutation burden. We investigated this in two ways. First, the total escape burden was estimated by identifying all HIV codons associated with one or more host HLA alleles and classifying the HIV residue at that site as adapted (inferred escaped) or susceptible based on published lists of HLA-associated polymorphisms defined at the population level in HIV subtype B (35). As such, the analysis was restricted to the 9 subtype B-infected participants. For each sequence, we calculated the percentage of HLA-associated sites exhibiting an adapted (or possibly adapted) form and computed the mean for each within-host data set (e.g., participant 7's data set is 43% adapted to the host HLA; Fig. 7A). Next, we estimated within-host escape complexity by quantifying the percentage of published or predicted optimally described HLA-restricted cytotoxic T lymphocyte (CTL) epitopes exhibiting within-host amino acid variation (e.g., 5/5 [100%] for participant 7). Both metrics correlated positively in blood and testes (Spearman's *r* = 0.98 and *P* = 0.0001 for escape burden [Fig. 7B] and Spearman's *r* = 0.93 and *P* = 0.001 for escape complexity [Fig. 7C]). In fact, these values were highly concordant (Lin's concordance coefficient for the data presented in Fig. 7B was 0.97 [95% confidence interval {CI}, 0.90 to 0.99], while that for the data presented in Fig. 7C was 0.95 [95% CI, 0.80 to 0.99]). Examples of the generally concordant nature of within-host HIV variation in key HLA-restricted epitopes in blood and testes are shown in Fig. 7D. Consistent with previous reports (15), these examples also reveal that HLA-susceptible and -adapted forms of the same CTL epitope commonly coexist in the reservoir: 6 of the 9 subtype B-infected individuals harbored at least one epitope where this occurred.

DISCUSSION

The availability of matched blood and tissue from our unique cohort (21, 23) allowed the first genetic exploration of the hypothesis that the testes constitute a distinctive HIV sanctuary site in healthy HIV-infected humans on suppressive cART (21, 25). Our study confirmed that CD4⁺ T cells are present in testicular tissue and that HIV DNA could be detected in the testes of all participants, though in two cases these were below quantification limits. Overall, the testis proviral burdens measured in the present study were lower than those previously reported in an autopsy study of cART-treated individuals; however, in the latter study, some individuals had discontinued cART near

TABLE 3 Genetic compartmentalization when sampling HIV sequences from the same site or tissue

Comparison	Test	Statistic	Result for the following participant ^c :									
			1	2	3	4	5	6	7	8	9	10
PBMC aliquot A vs aliquot B overall	Distance	K_{ST} ^a					0.02	0	0			
		<i>P</i> value					NS	NS	NS			
	Tree	No. of migrations ^b					9	11	13			
		<i>P</i> value					NS	NS	NS			
Right testis location A vs location B overall	Distance	K_{ST}			0.01		0		0.51	0.03	0.27 [–]	
		<i>P</i> value			NS		NS		**	NS	NS	
	Tree	No. of migrations			12		5		1	3	6	
		<i>P</i> value			NS		NS		*	NS	NS	
Left testis location A vs location B overall	Distance	K_{ST}	0			0.09	0					
		<i>P</i> value	NS			NS	NS					
	Tree	No. of migrations	17			3	4					
		<i>P</i> value	NS			NS	NS					

^a K_{ST} , statistic produced by the Hudson, Boos, and Kaplan nonparametric test for population structure, where 0 indicates no compartmentalization and 1 indicates complete compartmentalization.

^bThe number of migrations between compartments required to explain the input phylogeny, as defined by the Slatkin-Maddison test (57). The smaller the number, the stronger the support for compartmentalization.

^c*, 0.01 < *P* < 0.05; **, *P* < 0.01; NS, not significant.

the end of life, which could have increased proviral levels in tissue (30). The proviral burden in blood correlated with that in the testes (Fig. 1C), suggesting that individuals with larger blood reservoirs will tend to have larger tissue reservoirs (36). These observations confirm the testes as a site of HIV persistence during cART.

Our results also yielded insights into within-host proviral diversity. First, identical HIV *nef* sequences, which are suggestive of clonal expansion, were observed at markedly different frequencies within and between participants. While in some (e.g., participants 1 and 5) a particular *nef* variant dominated within the testes, in others (e.g., participant 2) a particular variant dominated in the blood, while in yet others (e.g., participant 6) identical *nef* sequences were relatively rare. These observations are consistent with the clonal expansion of latently infected cells as a driver of HIV persistence during cART in many persons (10, 16–18, 37, 38). However, the frequencies of identical *nef* sequences followed no consistent pattern between blood and testes, suggesting that one cannot generalize, based on location alone, where clonally expanded latently infected cell populations are more likely to reside. Disregarding clonal expansion, however, within-host blood and testis HIV diversity correlated significantly (Fig. 5), suggesting that individuals with diverse blood reservoirs also harbor diverse HIV sequences in the testes.

Our study also sheds some light on two outstanding questions: first, whether the within-host testicular microenvironment may be distinct enough to influence the proviral composition therein, and second, in the cases where blood-testis HIV genetic compartmentalization is detected, how this arises (31). Our observations reveal that, while blood-testis genetic compartmentalization was reasonably common if one analyzed within-host HIV *nef* sequences as a whole (two established tests consistently deemed 60% of overall within-host data sets to be compartmentalized), no participant retained consistent evidence of blood-testis compartmentalization when only unique HIV *nef* sequences per compartment were analyzed. Strictly speaking, therefore, the proviral populations in the testes of the majority of cART-suppressed individuals do differ from those in blood, but only because of their unbalanced distributions of identical sequences. This in turn suggests that the principal mechanism in which blood and testis HIV reservoirs differ from one another is not via seeding of divergent viral sequences therein pre-cART, nor is it due to restricted trafficking of viral lineages between blood and testes. Rather, our observations suggest that the extent and/or dynamics of proliferation of latently HIV-infected cells (and/or subsequent waning of descendant populations [10]) can differ substantially between blood and testes and

FIG 7 Immune escape landscape in blood and testes. (A) Example of escape burden analysis using participant 7's HIV Nef amino acid sequence alignments in blood and testes. Red, orange, and blue represent HLA-adapted, possibly adapted, and susceptible forms, respectively. For each sequence, the sum of HLA-adapted and possibly adapted sequences is divided by the total number of HLA-associated sites to yield the overall percentage of adapted sites (values shown at the end of each sequence). These values are then averaged to yield site- and participant-specific averages. For viral sites associated with more than one HLA allele (e.g., Nef codon 105), the inferred escape profile is shown for the underlined allele. (B) Correlation between escape burden in blood and testes. (C) Correlation between epitope complexity, defined here as the percentage of known or predicted HLA-restricted CTL epitopes exhibiting amino acid variation, between blood and testes. (D) Examples of escaped and susceptible forms coexisting within known or predicted CTL epitopes within the participants' reservoirs, where the height of the residue represents its frequency.

that this in turn drives the differences in the HIV population structure between these sites. While it is tempting to speculate that the site harboring the sequence in abundance is where the clonal expansion occurred and the other is where clonal descendants subsequently migrated, we cannot rule out the possibility that cells harboring identical HIV proviruses preexisted in both sites as a result of a previous expansion and migration event and that these underwent another expansion, in parallel, more recently. Our observation that the HIV population structure, when present, is due to the differential distribution of identical sequences across anatomical sites is consistent with studies of viral diversity within the female genital tract (39, 40). More broadly, our observations are consistent with the notion that major blood-tissue compartmentalization during cART, at least in terms of unique HIV lineages, may not be the norm (41–43).

At first glance, our finding that the composition of unique proviral lineages in the testes does not differ significantly from that in the blood may appear to contradict the former's status as an HIV sanctuary site (21, 44). The recovery of proviruses from the testicular tissue of all participants, however, confirms the testes as a site of HIV persistence during cART. Rather, our results suggest that the testis microenvironment may not be sufficiently unique to persistently maintain distinct HIV lineages therein, at least not over the time course of cART suppression of the participants studied. Similarly, while reduced anti-HIV immune responses likely facilitate HIV persistence in the testes in general, our observation that the HIV immune escape burden correlated strongly between blood and testes (Fig. 5) suggests that the extent of local immunosuppression may not be sufficient to markedly influence the immune escape landscape therein. However, there is an alternative explanation for the lack of genetic compartmentalization of unique HIV lineages between blood and testes: we cannot exclude the possibility that low-level HIV replication may still occur in certain tissues, including the testes, during cART, and that this could lead to the ongoing reshaping of the HIV populations persisting in blood.

Other limitations of our study and data interpretation should also be acknowledged. Though proviral sequences were obtained via single-template amplification, only a subgenomic HIV region (*nef*) was analyzed due to the challenges of recovering HIV from the testes (including from two participants for whom the testis proviral loads were below quantification limits). As such, we must acknowledge the possibility that sequences that are intact within *nef* may harbor defects elsewhere in the genome, nor can we definitively classify identical *nef* sequences as being derived from clonally expanded cell populations. As HIV sequences were isolated from PBMCs or whole tissue, the specific immune cells harboring them are unknown. This may be particularly important for the testes, which are rich in macrophages, which may contribute to HIV persistence there (30). Limited sampling must be acknowledged, particularly from the testes, which could compromise our power to detect compartmentalization. However, our independent recovery of the same sequence from separate testis sections, our lack of observation of compartmentalization within and between testes in 21/22 comparisons performed (Tables 2 and 3), our recovery of at least one identical HIV sequence from both blood and testes in each participant, and the observation that, for many participants, increased proviral sampling largely yielded sequences that had been recovered previously suggest that our results are not simply attributable to biased and/or limited sampling. Furthermore, our qualitative observation that blood and testis sequences were reasonably well intermixed in all phylogenies (that is, that no data set exhibited overt segregation by site) further supports the notion that these two sites do not differ substantively in terms of a unique HIV lineage distribution. We must also acknowledge that limited within-host diversity can contribute to uncertainty in phylogenetic reconstruction (as evidenced by low bootstrap values for some nodes in the within-host trees), which could in turn influence the results of tree-based compartmentalization tests, but the relatively high concordance between the tree- and distance-based tests (Fig. 6; Tables 2 and 3) suggests that this is not a major concern. As *nef* was sequenced, the possibility of a differential distribution of antiretroviral resistance

mutations across sites (as a result of the reduced penetration of some antiretrovirals into the testes [44]) could not be investigated. As clinical histories were unknown, we could not investigate correlates of blood-tissue HIV diversity and compartmentalization, nor can we rule out the possibility that evidence of genetic compartmentalization between blood and tissues may become apparent only after much more prolonged durations of cART suppression (e.g., as a result of latently infected cells having different half-lives in blood than in the testes as a result of reduced immune surveillance in the latter). Finally, proviral sequences could not be studied in the context of pre-cART viral populations, as these samples were not available.

In conclusion, and again noting the caveats associated with subgenomic HIV sequencing, our observation that a unique proviral sequence distribution and the immune escape mutation burden in testicular tissue are not unrepresentative of those within blood suggests that diversity and escape-related barriers to HIV eradication are likely to be of a broadly similar magnitude across these sites. However, the marked differences in the quantity and distribution of identical HIV sequences between hosts and between anatomical sites within each host underscore the potential challenges of differentially distributed clonally expanded latently infected cell populations to HIV eradication.

MATERIALS AND METHODS

Study participants and ethics statement. Blood and testicular tissues were donated by 10 HIV-infected but otherwise healthy adults aged 27 to 52 years who underwent elective bilateral orchiectomy for gender affirmation (the ORCHID cohort [21, 23]). At the time of surgery, all participants had maintained viremia suppression on cART for at least 6 months; however, full clinical histories (including infection date estimates) were unavailable. This study was approved by the Institutional Review Boards of the Research Institute of the McGill University Health Centre and Simon Fraser University. All participants provided written informed consent.

Blood and tissue collection. Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood by Ficoll density gradient centrifugation and stored at -80°C until use. Testicular tissue was processed within 1 h of surgery (21, 23). For genetic analysis, $\sim 100\text{-}\mu\text{g}$ tissue sections were snap-frozen upon reception in liquid nitrogen; for histology, 10-mm^3 sections were embedded in CryoMatrix material (Thermo Fisher) before snap-freezing.

Immunostaining. Frozen embedded tissues were cut into $5\text{-}\mu\text{m}$ -thick sections (Leica CM3050S cryostat), fixed in acetone-methanol for 5 min at 4°C , stained overnight at 4°C using anti-CD3 and anti-CD4 (Dako-Agilent), and exposed to conjugated Alexa Fluor secondary antibodies (Thermo Fisher) for 30 min at room temperature. The slides were counterstained using 4',6-diamidino-2-phenylindole (DAPI) before mounting in Fluoromount-G mounting medium (Southern Biotech). Images were acquired using a Nikon epifluorescence microscope with a Zyla sCMOS camera and analyzed using ImageJ (v1.47g) software (National Institutes of Health).

HIV quantification and sequencing. DNA was extracted from CD4^{+} T cells negatively isolated from PBMCs (StemCell Technologies) and directly from snap-frozen tissue (Qiagen). The HIV DNA burden was measured using a quantitative PCR assay capable of detecting a single viral genome per PCR, where primers are located in the HIV long terminal repeat-*gag* region and are optimized for detection of multiple HIV subtypes (21, 45). For each participant, a minimum of 5 million PBMCs plus at least four tissue sections, two each from the right and left testes, were allocated for HIV sequencing, with the exception of participant 8, for whom only two right testis sections were available. Genomic DNA was extracted from PBMCs directly (Invitrogen), whereas tissues were further sectioned into $\sim 50\text{-}\mu\text{g}$ fragments prior to genomic DNA extraction. HIV amplification was performed using single-genome amplification of a subgenomic fragment (HIV *nef*). The *nef* gene was selected based on its relatively high within-host diversity, its richness in phylogenetic signal, and its representativeness of within-host HIV evolution and diversity within the rest of the HIV genome (8). This gene also represents the one most likely to be intact within proviruses sampled during long-term cART (46). Single-genome amplification, performed using high-fidelity enzymes (Roche Expand HiFi) and oligonucleotide primers designed to amplify HIV sequences from all major subtypes, was achieved by endpoint dilution, such that ~ 25 to 30% of the resulting nested PCRs were successful (8). Negative PCR controls, included in every amplification, always remained negative. Amplicons were sequenced on an automated DNA sequencer using BigDye (v3.1) chemistry (Applied Biosystems). Chromatograms were base called using the Sequencher (v5.0) program (GeneCodes). Sequences exhibiting nucleotide mixtures, defects, or hypermutations (identified using the Hypermut program [47]) were excluded, as were sequences exhibiting evidence of within-host recombination (identified using the RDP4 program [48]).

Sequences were codon aligned using the MAFFT method (49) and edited in the AliView (v1.18) alignment viewer and editor (50). For each within-host alignment, the best-fitting substitution models for both genetic distance and phylogenetic-based tests were determined using the jModelTest (v2.1.10) program (51). Phylogenies were inferred by approximate maximum likelihood using the FastTree2 (v2.1.10) program (52) under the best available nucleotide substitution model (for participants 7 and 10, the generalized time-reversible [GTR] model [53] with no rate heterogeneity among sites was the best fit; for all others, a GTR model with rate heterogeneity using a gamma distribution was the best fit). Node

support values were derived from 1,000 bootstraps. Within-host phylogenies were midpoint rooted in FigTree (v1.4.4) software (<http://tree.bio.ed.ac.uk/software/figtree/>) and visualized with the ggtree program (54).

Genetic compartmentalization tests and other statistical analyses. As there are different ways to define genetic compartmentalization (e.g., as genetic heterogeneity, or the presence of distinct lineages, between two viral subpopulations), compartmentalization tests occasionally yield discordant results (55). As recommended, we applied more than one test and classified a data set as “compartmentalized” only when the results of both tests agreed (55). We employed one genetic distance-based test (56) and one tree-based test (57), both of which were implemented in the HyPhy (v2.22) program (58). For the distance-based test, we employed the Hudson, Boos, and Kaplan nonparametric test for population structure (56), a test that has been validated for within-host HIV data sets (59), including from the reservoir, where it is sometimes referred to as the “nonparametric test for panmixia” (59–61). This test compares the mean pairwise distances between sequences from different subpopulations versus the same subpopulation (compartmentalization is supported if the mean pairwise distances of sequences from the same subpopulation are smaller than those from different subpopulations) and computes the K_{ST} statistic, whose values range from 0 (denoting no compartmentalization) to 1 (denoting complete compartmentalization). Statistical significance, expressed as a P value, is assessed via a population structure randomization test. Using the jModelTest (v2.1.10) program (51), we identified the Tamura-Nei 1993 (TN93) nucleotide substitution model to be the best-fitting distance-based model available in the HyPhy (v2.22) program (58); use of this model is also consistent with previous applications of this test to within-host HIV data sets (55, 62). For the tree-based test, we used the Slatkin-Maddison test (57). This test determines the minimum number of between-compartment migrations to explain the distribution of compartments on the tree tips: the smaller the number, the stronger the support for compartmentalization. Statistical significance is assessed via a lineage permutation test. Importantly, because identical HIV sequences (especially those that are present in abundance and whose frequencies differ between sites) increase the likelihood of compartmentalization detection (39), we analyzed each data set in two ways: overall and with identical sequences collapsed to a single copy per compartment.

HLA class I sequence-based typing was performed to allele-level resolution (63). HLA-associated adapted (inferred HLA-escaped) and nonadapted (inferred HLA-susceptible) polymorphisms in HIV subtype B were defined using a published list derived from an analysis of linked HIV/HLA genotypes from an independent cohort of 1,888 subtype B-infected individuals (35). HLA-restricted, optimally described CTL epitopes in the participants' provirus sequences were defined using the Los Alamos HIV Molecular Immunology Database with current updates (64), where epitopes restricted by HLA alleles closely related to the participants' allele(s) were identified as putative epitopes (35). Correlations were assessed using Spearman's correlation. P values of <0.05 were considered statistically significant.

Data availability. GenBank accession numbers for intact sequences are [MK301609](#) to [MK302069](#) and [MK849402](#) to [MK849612](#).

ACKNOWLEDGMENTS

The research reported in this publication was funded in part by Canadian HIV Cure Enterprise Team grants HIG-133050 (to J.-P.R., Z.L.B., and A.F.Y.P.) and HB2-164064 (to J.-P.R.) from the Canadian Institutes for Health Research (CIHR) in partnership with CANFAR and the International AIDS Society (IAS), by CIHR team grant HB1-164063 (to Z.L.B.), by CIHR project grants PJT-159625 (to J.B.J. and Z.L.B.) and PJT-155990 (to A.F.Y.P.), by the National Institutes of Health under award number NIHR21A127029 (to J.B.J., A.F.Y.P., Z.L.B.), by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health with cofunding support from the National Institute on Drug Abuse, the National Institute of Mental Health, and the National Institute of Neurological Disorders and Stroke under award number UM1AI126617 (to Z.L.B.), and by a Simon Fraser University Next Big Question fund award (to Z.L.B.). This study was also supported by the Fonds de recherche du Québec-Santé (FRQ-S): Réseau SIDA/Maladies infectieuses and Thérapie cellulaire (to J.-P.R.). N.N.K. was supported by a CIHR Frederick Banting and Charles Best M.Sc. award. F.H.O. was supported by the Canadian Queen Elizabeth II Diamond Jubilee Scholarships program, a joint initiative of the Rideau Hall Foundation, Community Foundations of Canada, and the Association of Universities and Colleges of Canada, and also received a fellowship from the Sub-Saharan African Network for TB/HIV Research Excellence (SANTHE), a DELTAS Africa Initiative (grant number DEL-15-006). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS) Alliance for Accelerating Excellence in Science in Africa (AESA) and is supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust (grant number 107752/Z/15/Z) and the UK government. M.-A.J. holds the CIHR Canada Research Chair tier 2 in immunovirology. A.F.Y.P. is supported by a CIHR new investigator award (FRN-130609). N.C. is supported by a

research scholar career award of the Quebec Health Research Fund (FRQ-S; number 253292). Z.L.B. is supported by a Michael Smith Foundation for Health Research (MSFHR) scholar award. J.-P.R. is the holder of the Louis Lowenstein Chair in Hematology and Oncology at McGill University.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, AAS, the NEPAD Agency, the Wellcome Trust, the UK government, or the other funding agencies listed.

We thank Gabriel Stang and Jeff Knaggs for software programming and bioinformatics assistance and Mark Brockman for helpful discussions. We are grateful to the participants, who made this study possible by their donations of blood and tissue.

REFERENCES

- Finzi D, Hermankova M, Pierson T, Carruth LM, Buck C, Chaisson RE, Quinn TC, Chadwick K, Margolick J, Brookmeyer R, Gallant J, Markowitz M, Ho DD, Richman DD, Siliciano RF. 1997. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science* 278: 1295–1300. <https://doi.org/10.1126/science.278.5341.1295>.
- Shankarappa R, Margolick JB, Gange SJ, Rodrigo AG, Upchurch D, Farzadegan H, Gupta P, Rinaldo CR, Learn GH, He X, Huang XL, Mullins JL. 1999. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J Virol* 73:10489–10502.
- Herbeck JT, Rolland M, Liu Y, McLaughlin S, McNeven J, Zhao H, Wong K, Stoddard JN, Raugi D, Sorensen S, Genowati I, Birditt B, McKay A, Diem K, Maust BS, Deng W, Collier AC, Stekler JD, McElrath MJ, Mullins JL. 2011. Demographic processes affect HIV-1 evolution in primary infection before the onset of selective processes. *J Virol* 85:7523–7534. <https://doi.org/10.1128/JVI.02697-10>.
- Dapp MJ, Kober KM, Chen L, Westfall DH, Wong K, Zhao H, Hall BM, Deng W, Sibley T, Ghorai S, Kim K, Chen N, McHugh S, Au L, Cohen M, Anastos K, Mullins JL. 2017. Patterns and rates of viral evolution in HIV-1 subtype B infected females and males. *PLoS One* 12:e0182443. <https://doi.org/10.1371/journal.pone.0182443>.
- Zanini F, Brodin J, Thebo L, Lanz C, Bratt G, Albert J, Neher RA. 2015. Population genomics of intrapatient HIV-1 evolution. *Elife* 4:e11282. <https://doi.org/10.7554/eLife.11282>.
- Whitney JB, Hill AL, Sanisetty S, Penaloza-MacMaster P, Liu J, Shetty M, Parenteau L, Cabral C, Shields J, Blackmore S, Smith JY, Brinkman AL, Peter LE, Mathew SI, Smith KM, Borducchi EN, Rosenbloom DI, Lewis MG, Hattersley J, Li B, Hesselgesser J, Geleziunas R, Robb ML, Kim JH, Michael NL, Barouch DH. 2014. Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys. *Nature* 512:74–77. <https://doi.org/10.1038/nature13594>.
- Brodin J, Zanini F, Thebo L, Lanz C, Bratt G, Neher RA, Albert J. 2016. Establishment and stability of the latent HIV-1 DNA reservoir. *Elife* 5:e18889. <https://doi.org/10.7554/eLife.18889>.
- Jones BR, Kinloch NN, Horacek J, Ganase B, Harris M, Harrigan PR, Jones RB, Brockman MA, Joy JB, Poon AFY, Brumme ZL. 2018. Phylogenetic approach to recover integration dates of latent HIV sequences within-host. *Proc Natl Acad Sci U S A* 115:E8958–E8967. <https://doi.org/10.1073/pnas.1802081115>.
- Siliciano JD, Kajdas J, Finzi D, Quinn TC, Chadwick K, Margolick JB, Kovacs C, Gange SJ, Siliciano RF. 2003. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4⁺ T cells. *Nat Med* 9:727–728. <https://doi.org/10.1038/nm880>.
- Wang Z, Gurule EE, Brennan TP, Gerold JM, Kwon KJ, Hosmane NN, Kumar MR, Beg SA, Capoferri AA, Ray SC, Ho YC, Hill AL, Siliciano JD, Siliciano RF. 2018. Expanded cellular clones carrying replication-competent HIV-1 persist, wax, and wane. *Proc Natl Acad Sci U S A* 115:E2575–E2584. <https://doi.org/10.1073/pnas.1720665115>.
- Hiener B, Horsburgh BA, Eden JS, Barton K, Schlub TE, Lee E, von Stockenstrom S, Odeval L, Milush JM, Liegler T, Sinclair E, Hoh R, Boritz EA, Douek D, Fromentin R, Chomont N, Deeks SG, Hecht FM, Palmer S. 2017. Identification of genetically intact HIV-1 proviruses in specific CD4⁺ T cells from effectively treated participants. *Cell Rep* 21:813–822. <https://doi.org/10.1016/j.celrep.2017.09.081>.
- Kearney M, Spindler J, Shao W, Maldarelli F, Palmer S, Hu SL, Lifson JD, KewalRamani VN, Mellors JW, Coffin JM, Ambrose Z. 2011. Genetic diversity of simian immunodeficiency virus encoding HIV-1 reverse transcriptase persists in macaques despite antiretroviral therapy. *J Virol* 85:1067–1076. <https://doi.org/10.1128/JVI.01701-10>.
- Josefsson L, von Stockenstrom S, Faria NR, Sinclair E, Bacchetti P, Killian M, Epling L, Tan A, Ho T, Lemey P, Shao W, Hunt PW, Somsouk M, Wylie W, Douek DC, Loeb L, Custer J, Hoh R, Poole L, Deeks SG, Hecht F, Palmer S. 2013. The HIV-1 reservoir in eight patients on long-term suppressive antiretroviral therapy is stable with few genetic changes over time. *Proc Natl Acad Sci U S A* 110:E4987–E4996. <https://doi.org/10.1073/pnas.1308313110>.
- Deng K, Perteau M, Rongvaux A, Wang L, Durand CM, Ghiaur G, Lai J, McHugh HL, Hao H, Zhang H, Margolick JB, Gurer C, Murphy AJ, Valenzuela DM, Yancopoulos GD, Deeks SG, Strowig T, Kumar P, Siliciano JD, Salzberg SL, Flavell RA, Shan L, Siliciano RF. 2015. Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations. *Nature* 517:381–385. <https://doi.org/10.1038/nature14053>.
- Brumme ZL, Sudderuddin H, Ziemniak C, Luzuriaga K, Jones BR, Joy JB, Cunningham CK, Greenough T, Persaud D. 2019. Genetic complexity in the replication-competent latent HIV reservoir increases with untreated infection duration in infected youth. *AIDS* 33:211–218. <https://doi.org/10.1097/QAD.0000000000002045>.
- Bui JK, Sobolewski MD, Keele BF, Spindler J, Musick A, Wiegand A, Luke BT, Shao W, Hughes SH, Coffin JM, Kearney MF, Mellors JW. 2017. Proviruses with identical sequences comprise a large fraction of the replication-competent HIV reservoir. *PLoS Pathog* 13:e1006283. <https://doi.org/10.1371/journal.ppat.1006283>.
- Maldarelli F, Wu X, Su L, Simonetti FR, Shao W, Hill S, Spindler J, Ferris AL, Mellors JW, Kearney MF, Coffin JM, Hughes SH. 2014. HIV latency. Specific HIV integration sites are linked to clonal expansion and persistence of infected cells. *Science* 345:179–183. <https://doi.org/10.1126/science.1254194>.
- Wagner TA, McLaughlin S, Garg K, Cheung CY, Larsen BB, Styrchak S, Huang HC, Edlefsen PT, Mullins JL, Frenkel LM. 2014. HIV latency. Proliferation of cells with HIV integrated into cancer genes contributes to persistent infection. *Science* 345:570–573. <https://doi.org/10.1126/science.1256304>.
- Lee GQ, Orlova-Fink N, Einkauf K, Chowdhury FZ, Sun X, Harrington S, Kuo HH, Hua S, Chen HR, Ouyang Z, Reddy K, Dong K, Ndung'u T, Walker BD, Rosenberg ES, Yu XG, Lichterfeld M. 2017. Clonal expansion of genome-intact HIV-1 in functionally polarized Th1 CD4⁺ T cells. *J Clin Invest* 127:2689–2696. <https://doi.org/10.1172/JCI93289>.
- Mital P, Hinton BT, Dufour JM. 2011. The blood-testis and blood-epididymis barriers are more than just their tight junctions. *Biol Reprod* 84:851–858. <https://doi.org/10.1095/biolreprod.110.087452>.
- Jenabian MA, Costiniuk CT, Mehraj V, Ghazawi FM, Fromentin R, Brousseau J, Brassard P, Belanger M, Ancuta P, Bendayan R, Chomont N, Routy JP, Orchid Study Group. 2016. Immune tolerance properties of the testicular tissue as a viral sanctuary site in ART-treated HIV-infected adults. *AIDS* 30:2777–2786. <https://doi.org/10.1097/QAD.0000000000001282>.
- Le Tortorec A, Le Grand R, Denis H, Satie A-P, Mannioui K, Roques P, Maillard A, Daniels S, Jégou B, Dejucq-Rainsford N. 2008. Infection of semen-producing organs by SIV during the acute and chronic stages of the disease. *PLoS One* 3:e1792. <https://doi.org/10.1371/journal.pone.0001792>.
- Ponte R, Dupuy FP, Brimo F, Mehraj V, Brassard P, Belanger M, Yurchenko E, Jenabian M-A, Bernard NF, Routy J-P. 2018. Characterization of myeloid cell populations in human testes collected after sex reassignment

- surgery. *J Reprod Immunol* 125:16–24. <https://doi.org/10.1016/j.jri.2017.10.043>.
24. De Rose R, Fernandez CS, Hedger MP, Kent SJ, Winnall WR. 2013. Characterisation of macaque testicular leucocyte populations and T-lymphocyte immunity. *J Reprod Immunol* 100:146–156. <https://doi.org/10.1016/j.jri.2013.09.003>.
 25. Winnall WR, Lloyd SB, De Rose R, Alcantara S, Amarasekera TH, Hedger MP, Girling JE, Kent SJ. 2015. Simian immunodeficiency virus infection and immune responses in the pig-tailed macaque testis. *J Leukoc Biol* 97: 599–609. <https://doi.org/10.1189/jlb.4A0914-438R>.
 26. Nikolova M, Carriere M, Jenabian MA, Limou S, Younas M, Kok A, Hue S, Seddiki N, Hulin A, Delaneau O, Schuitemaker H, Herbeck JT, Mullins JI, Muhtarova M, Bensussan A, Zagury JF, Lelievre JD, Levy Y. 2011. CD39/adenosine pathway is involved in AIDS progression. *PLoS Pathog* 7:e1002110. <https://doi.org/10.1371/journal.ppat.1002110>.
 27. Shehu-Xhilaga M, Kent S, Batten J, Ellis S, Van der Meulen J, O'Bryan M, Cameron PU, Lewin SR, Hedger MP. 2007. The testis and epididymis are productively infected by SIV and SHIV in juvenile macaques during the post-acute stage of infection. *Retrovirology* 4:7. <https://doi.org/10.1186/1742-4690-4-7>.
 28. Matusali G, Dereuddre-Bosquet N, Le Tortorec A, Moreau M, Satie AP, Mahe D, Roumaud P, Bourry O, Sylla N, Bernard-Stoecklin S, Pruvost A, Le Grand R, Dejuicq-Rainsford N. 2015. Detection of simian immunodeficiency virus in semen, urethra, and male reproductive organs during efficient highly active antiretroviral therapy. *J Virol* 89:5772–5787. <https://doi.org/10.1128/JVI.03628-14>.
 29. North TW, Higgins J, Deere JD, Hayes TL, Villalobos A, Adamson L, Shacklett BL, Schinazi RF, Luciw PA. 2010. Viral sanctuaries during highly active antiretroviral therapy in a nonhuman primate model for AIDS. *J Virol* 84:2913–2922. <https://doi.org/10.1128/JVI.02356-09>.
 30. Lamers SL, Rose R, Maidji E, Agsald-Garcia M, Nolan DJ, Fogel GB, Salemi M, Garcia DL, Bracci P, Yong W, Commins D, Said J, Khanlou N, Hinkin CH, Sueiras MV, Mathisen G, Donovan S, Shiramizu B, Stoddart CA, McGrath MS, Singer EJ. 2016. HIV DNA is frequently present within pathologic tissues evaluated at autopsy from combined antiretroviral therapy-treated patients with undetectable viral loads. *J Virol* 90: 8968–8983. <https://doi.org/10.1128/JVI.00674-16>.
 31. Nickle DC, Jensen MA, Shriner D, Brodie SJ, Frenkel LM, Mittler JE, Mullins JI. 2003. Evolutionary indicators of human immunodeficiency virus type 1 reservoirs and compartments. *J Virol* 77:5540–5546. <https://doi.org/10.1128/jvi.77.9.5540-5546.2003>.
 32. Theiler J, Yoon H, Yusim K, Picker LJ, Fruh K, Korber B. 2016. Epigraph: a vaccine design tool applied to an HIV therapeutic vaccine and a pan-filovirus vaccine. *Sci Rep* 6:33987. <https://doi.org/10.1038/srep33987>.
 33. Tumiotto C, Riviere L, Bellecave P, Recordon-Pinson P, Vilain-Parce A, Guidicelli GL, Fleury H, Provir/Latitude 45 Collaborating Group. 2017. Sanger and next-generation sequencing data for characterization of CTL epitopes in archived HIV-1 proviral DNA. *PLoS One* 12:e0185211. <https://doi.org/10.1371/journal.pone.0185211>.
 34. Houzet L, Perez-Losada M, Matusali G, Deleage C, Dereuddre-Bosquet N, Satie AP, Aubry F, Becker E, Jegou B, Le Grand R, Keele BF, Crandall KA, Dejuicq-Rainsford N. 2018. Seminal simian immunodeficiency virus in chronically infected cynomolgus macaques is dominated by virus originating from multiple genital organs. *J Virol* 92:e00133-18. <https://doi.org/10.1128/JVI.00133-18>.
 35. Carlson JM, Brumme CJ, Martin E, Listgarten J, Brockman MA, Le AQ, Chui CK, Cotton LA, Knapp DJ, Riddler SA, Haubrich R, Nelson G, Pfeifer N, Deziel CE, Heckerman D, Apps R, Carrington M, Mallal S, Harrigan PR, John M, Brumme ZL. 2012. Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. *J Virol* 86:13202–13216. <https://doi.org/10.1128/JVI.01998-12>.
 36. Chomont N, El-Far M, Ancuta P, Trautmann L, Procopio FA, Yassine-Diab B, Boucher G, Boullassel M-R, Ghattas G, Brenchley JM, Schacker TW, Hill BJ, Douek DC, Routy J-P, Haddad EK, Sékaly R-P. 2009. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat Med* 15:893–900. <https://doi.org/10.1038/nm.1972>.
 37. Mullins JI, Frenkel LM. 2017. Clonal expansion of human immunodeficiency virus-infected cells and human immunodeficiency virus persistence during antiretroviral therapy. *J Infect Dis* 215:S119–S127. <https://doi.org/10.1093/infdis/jiw636>.
 38. Murray AJ, Kwon KJ, Farber DL, Siliciano RF. 2016. The latent reservoir for HIV-1: how immunologic memory and clonal expansion contribute to HIV-1 persistence. *J Immunol* 197:407–417. <https://doi.org/10.4049/jimmunol.1600343>.
 39. Bull M, Learn G, Genowati I, McKernan J, Hitti J, Lockhart D, Tapia K, Holte S, Dragavon J, Coombs R, Mullins J, Frenkel L. 2009. Compartmentalization of HIV-1 within the female genital tract is due to monotypic and low-diversity variants not distinct viral populations. *PLoS One* 4:e7122. <https://doi.org/10.1371/journal.pone.0007122>.
 40. Bull ME, Learn GH, McElhone S, Hitti J, Lockhart D, Holte S, Dragavon J, Coombs RW, Mullins JI, Frenkel LM. 2009. Monotypic human immunodeficiency virus type 1 genotypes across the uterine cervix and in blood suggest proliferation of cells with provirus. *J Virol* 83:6020–6028. <https://doi.org/10.1128/JVI.02664-08>.
 41. Dahl V, Gisslen M, Hagberg L, Peterson J, Shao W, Spudich S, Price RW, Palmer S. 2014. An example of genetically distinct HIV type 1 variants in cerebrospinal fluid and plasma during suppressive therapy. *J Infect Dis* 209:1618–1622. <https://doi.org/10.1093/infdis/jit805>.
 42. Chun TW, Nickle DC, Justement JS, Meyers JH, Roby G, Hallahan CW, Kottitil S, Moir S, Mican JM, Mullins JI, Ward DJ, Kovacs JA, Mannon PJ, Fauci AS. 2008. Persistence of HIV in gut-associated lymphoid tissue despite long-term antiretroviral therapy. *J Infect Dis* 197:714–720. <https://doi.org/10.1086/527324>.
 43. Kearney MF, Anderson EM, Coomer C, Smith L, Shao W, Johnson N, Kline C, Spindler J, Mellors JW, Coffin JM, Ambrose Z. 2015. Well-mixed plasma and tissue viral populations in RT-SHIV-infected macaques implies a lack of viral replication in the tissues during antiretroviral therapy. *Retrovirology* 12:93. <https://doi.org/10.1186/s12977-015-0212-2>.
 44. Huang Y, Hoque MT, Jenabian MA, Vyboh K, Whyte SK, Sheehan NL, Brassard P, Belanger M, Chomont N, Fletcher CV, Routy JP, Bendayan R. 2016. Antiretroviral drug transporters and metabolic enzymes in human testicular tissue: potential contribution to HIV-1 sanctuary site. *J Antimicrob Chemother* 71:1954–1965. <https://doi.org/10.1093/jac/dkw046>.
 45. Vandergeeten C, Fromentin R, Merlini E, Lawani MB, DaFonseca S, Bakeman W, McNulty A, Ramgopal M, Michael N, Kim JH, Ananworanich J, Chomont N. 2014. Cross-clade ultrasensitive PCR-based assays to measure HIV persistence in large-cohort studies. *J Virol* 88:12385–12396. <https://doi.org/10.1128/JVI.00609-14>.
 46. Bruner KM, Wang Z, Simonetti FR, Bender AM, Kwon KJ, Sengupta S, Fray EJ, Beg SA, Antar AAR, Jenike KM, Bertagnolli LN, Capoferri AA, Kufera JT, Timmons A, Nobles C, Gregg J, Wada N, Ho YC, Zhang H, Margolick JB, Blankson JN, Deeks SG, Bushman FD, Siliciano RF, Laird GM, Siliciano RF. 2019. A quantitative approach for measuring the reservoir of latent HIV-1 proviruses. *Nature* 566:120–125. <https://doi.org/10.1038/s41586-019-0898-8>.
 47. Rose PP, Korber BT. 2000. Detecting hypermutations in viral sequences with an emphasis on G → A hypermutation. *Bioinformatics* 16:400–401. <https://doi.org/10.1093/bioinformatics/16.4.400>.
 48. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. 2015. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol* 1:vev003. <https://doi.org/10.1093/ve/vev003>.
 49. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059–3066. <https://doi.org/10.1093/nar/gkf436>.
 50. Larsson A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30:3276–3278. <https://doi.org/10.1093/bioinformatics/btu531>.
 51. Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9:772. <https://doi.org/10.1038/nmeth.2109>.
 52. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>.
 53. Tavaré S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lect Math Life Sci* 17:57–86.
 54. Yu G, Smith D, Zhu H, Guan Y, Lam T. 2017. GGTREE: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol* 8:28–36. <https://doi.org/10.1111/2041-210X.12628>.
 55. Zarate S, Pond SL, Shapshak P, Frost SD. 2007. Comparative study of methods for detecting sequence compartmentalization in human immunodeficiency virus type 1. *J Virol* 81:6643–6651. <https://doi.org/10.1128/JVI.02268-06>.
 56. Hudson RR, Boos DD, Kaplan NL. 1992. A statistical test for detecting

- geographic subdivision. *Mol Biol Evol* 9:138–151. <https://doi.org/10.1093/oxfordjournals.molbev.a040703>.
57. Slatkin M, Maddison WP. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* 123:603–613.
 58. Pond SL, Frost SD, Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679. <https://doi.org/10.1093/bioinformatics/bti079>.
 59. Achaz G, Palmer S, Kearney M, Maldarelli F, Mellors JW, Coffin JM, Wakeley J. 2004. A robust measure of HIV-1 population turnover within chronically infected individuals. *Mol Biol Evol* 21:1902–1912. <https://doi.org/10.1093/molbev/msh196>.
 60. Chaillon A, Gianella S, Lada SM, Perez-Santiago J, Jordan P, Ignacio C, Karris M, Richman DD, Mehta SR, Little SJ, Wertheim JO, Smith DM. 2018. Size, composition, and evolution of HIV DNA populations during early antiretroviral therapy and intensification with maraviroc. *J Virol* 92: e01589-17. <https://doi.org/10.1128/JVI.01589-17>.
 61. Kearney MF, Spindler J, Shao W, Yu S, Anderson EM, O'Shea A, Rehm C, Poethke C, Kovacs N, Mellors JW, Coffin JM, Maldarelli F. 2014. Lack of detectable HIV-1 molecular evolution during suppressive antiretroviral therapy. *PLoS Pathog* 10:e1004010. <https://doi.org/10.1371/journal.ppat.1004010>.
 62. Gianella S, Kosakovsky Pond SL, Oliveira MF, Scheffler K, Strain MC, De la Torre A, Letendre S, Smith DM, Ellis RJ. 2016. Compartmentalized HIV rebound in the central nervous system after interruption of antiretroviral therapy. *Virus Evol* 2:vew020. <https://doi.org/10.1093/ve/vew020>.
 63. Cotton LA, Rahman MA, Ng C, Le AQ, Milloy MJ, Mo T, Brumme ZL. 2012. HLA class I sequence-based typing using DNA recovered from frozen plasma. *J Immunol Methods* 382:40–47. <https://doi.org/10.1016/j.jim.2012.05.003>.
 64. Frahm N, Baker B, Brander C. 2008. Identification and optimal definition of HIV-derived cytotoxic T lymphocyte (CTL) epitopes for the study of CTL escape, functional avidity and viral evolution, p 3–24. *In* Korber BTM, Brander C, Haynes BF, Koup R, Moore JP, Walker BD, Watkins DI (ed), *HIV molecular immunology 2008*. Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, NM.