The HIV-1 latent reservoir is largely sensitive to circulating T cells

Joanna A Warren[1], Shuntai Zhou[2,3], Yinyan Xu[1], Matthew J Moeser[2,3], Daniel R MacMillan[4], Olivia Council[1,2], Jennifer Kirchherr[5], Julia M Sung[5,6], Nadia R Roan[7,8], Adaora A Adimora[5], Sarah Joseph[1,6], JoAnn D Kuruc[5,6], Cynthia L Gay[5,6], David M Margolis[1,3,5,6], Nancie Archin[5,6], Zabrina L Brumme[4,9], Ronald Swanstrom[1,2,3,6], Nilu Goonetilleke[1,5,6]*

[1]Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, United States; [2]Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, United States; [3]UNC Center For AIDS Research, University of North Carolina, Chapel Hill, United States; [4]British Columbia Centre for Excellence in HIV/AIDS, Vancouver, Canada; [5]Department of Medicine, University of North Carolina, Chapel Hill, United States; [6]UNC HIV Cure Center, University of North Carolina, Chapel Hill, United States; [7]Department of Urology, University of California San Francisco, San Francisco, United States; [8]Gladstone Institute of Virology and Immunology, San Francisco, United States; [9]Faculty of Health Sciences, Simon Fraser University, Burnaby, Canada

Abstract HIV-1-specific CD8+ T cells are an important component of HIV-1 curative strategies. Viral variants in the HIV-1 reservoir may limit the capacity of T cells to detect and clear virus-infected cells. We investigated the patterns of T cell escape variants in the replication-competent reservoir of 25 persons living with HIV-1 (PLWH) durably suppressed on antiretroviral therapy (ART). We identified all reactive T cell epitopes in the HIV-1 proteome for each participant and sequenced HIV-1 outgrowth viruses from resting CD4+ T cells. All non-synonymous mutations in reactive T cell epitopes were tested for their effect on the size of the T cell response, with a ≥50% loss defined as an escape mutation. The majority (68%) of T cell epitopes harbored no detectable escape mutations. These findings suggest that circulating T cells in PLWH on ART could contribute to control of rebound and could be targeted for boosting in curative strategies.

*For correspondence:
nilu_goonetilleke@med.unc.edu

Introduction

Antiretroviral therapy (ART) has transformed human immunodeficiency virus type 1 (HIV-1) from a fatal disease to a chronic condition. However, ART must be taken life-long as interruption of therapy in people living with HIV-1 (PLWH) mostly results in viral rebound within weeks (*Davey et al., 1999*; *Crowell et al., 2019*; *Bar et al., 2016*; *Buzon et al., 2014*). This rebound results from cells harboring HIV-1 DNA that is integrated into the host genome. While >95% of proviral DNA is replication-incompetent, the remaining fraction that we define here as the 'HIV-1 reservoir', retains the capacity to produce infectious virus particles, either stochastically (*Chun et al., 1997*; *Finzi, 1997*; *Siliciano et al., 2003*; *Eisele and Siliciano, 2012*), or through interventions to induce coordinated reactivation (*Archin et al., 2012*; *Søgaard et al., 2015*; *Elliott et al., 2013*; *Elliott et al., 2015*). The largest and best characterized HIV-1 reservoir resides in circulating resting CD4+ T cells. The inducible form of the HIV-1 reservoir is present in blood cells at a low frequency (~1 in 10^6 resting CD4+ T cells), but is highly stable, with an estimated half-life of 44 months (*Crooks et al., 2015*; *Siliciano and Siliciano, 2015*; *Siliciano and Siliciano, 2004*). Natural decay of the HIV-1 reservoir

would therefore take >70 years (*Crooks et al., 2015*; *Siliciano and Siliciano, 2015*; *Siliciano and Siliciano, 2004*).

Strategies to allow PLWH to cease ART without viral rebound aim to either eliminate all HIV-1 infected cells harboring persistent, replication-competent virus (HIV-1 eradication), or to achieve a state of durable HIV-1 suppression without rebound (HIV-1 remission). Both approaches may harness HIV-1-specific CD8+ T cells to achieve reservoir reduction or elimination. T cells recognize infected cells through T cell receptors (TCR) that bind viral peptides presented on human leukocyte antigen (HLA) on the surface of the infected cell (*Townsend and McMichael, 1985*; *Townsend et al., 1984*; *Townsend et al., 1986*). On peptide recognition, CD8+ T cells can produce cytokines or release lytic molecules killing the infected cell. There is significant evidence that CD8+ T cells contribute to the control of untreated HIV-1, supported by studies of simian immunodeficiency virus (SIV) in non-human primates (*Pandrea et al., 2011*; *Barouch et al., 2002*; *Schmitz et al., 2005*; *McBrien et al., 2020*). First, certain HLA-I types, such as B*57:01/:03 and B*27:05, are overrepresented in HIV-1 elite virus controllers, defined as PLWH who exhibit sustained HIV-1 control below 50 copies/ml without ART. This suggests that specific subpopulations of CD8+ T cells are capable of controlling HIV-1 replication in the absence of therapy. Second, HIV-1-specific CD8+ T cells exert significant immune pressure on HIV-1, resulting in the accumulation of non-synonymous mutations in and around reactive T cell epitopes that can confer escape from CD8+ T cell recognition and killing (*Borrow et al., 1997*; *Goulder et al., 1997*; *Phillips et al., 1991*; *Price et al., 1997*). HIV-1 escape from CD8+ T cells, which is typically measured as a $\geq$ 50% decrease in the magnitude of the T cell response, can occur within days to weeks of infection (*Borrow et al., 1997*; *Price et al., 1997*; *Koup et al., 1994*; *Goonetilleke et al., 2009*). Both T cell immunodominance (relative strength of a T cell response) and viral entropy (sequence conservation of HIV-1 at a population level) are major determinants of the rate of virus escape in untreated infection (*Liu et al., 2013*; *Ferrari et al., 2011*; *Barton et al., 2016*). Notably, virus escape is ongoing in untreated HIV-1 infection (*Allen et al., 2005*), with new virus variants emerging over time in response to newly induced T cell responses, shifts in T cell clonotypes, or to compensate for fitness costs of earlier mutations. Lastly, when T cell selection pressure is removed, as in the case of virus transmission to recipients with HLA alleles different from the donor, reversion of mutations can occur (*Leslie et al., 2004*; *Carlson et al., 2014*).

ART is highly effective in inhibiting viral replication with multiple studies reporting that HIV-1 does not undergo directional selection during suppressive ART (*Kearney et al., 2014*; *Kearney et al., 2016*; *Rosenbloom et al., 2017*; *Van Zyl et al., 2017*). Recent studies have reported that the majority of proviruses, both total HIV-1 DNA and replication-competent viruses, that persist during ART, reflect viruses circulating during the year prior to ART initiation (*Brodin et al., 2016*; *Abrahams et al., 2019*; *Pankau et al., 2020*). These studies suggest that ART initiation is the major driver of latency formation, including the formation of the replication-competent stable HIV-1 reservoir. Proviruses that persist during ART would therefore include HIV-1 variants that have escaped from T cell selection over the entire course of untreated infection (*Phillips et al., 1991*; *Deng et al., 2015*). ART also impacts the HIV-1-specific T cell response, with reduced HIV-1 loads resulting in lower antigen drive and over time, yielding improved proliferation and a less exhausted T cell phenotype (*Rehr et al., 2008*; *Rutishauser et al., 2017*) (reviewed in *Warren et al., 2019*). We have recently shown that HIV-1-specific T cell responses are detectable ex vivo and are highly stable in PLWH on durable ART suppression (*Xu et al., 2019*), with others showing that ex vivo isolated CD8 + T cells can suppress HIV-1 superinfected CD4+ T cells as well as in vitro reactivated reservoir virus (*Sung et al., 2015*; *Shan et al., 2012*).

Collectively, these data suggest that HIV-1-specific T cell responses on ART should reflect those detected prior to the onset of ART. Likewise, the persistent HIV-1 reservoir should reflect the burden of virus escape mutants circulating pre-ART. Of note, a recent study reported that up to 98% of proviral HIV-1 sequences harbored mutations in immunoprevalent viral epitopes, though unmutated targets were also identified (*Deng et al., 2015*). A high burden of HIV-1 escape variants in the reservoir could limit approaches that aim to increase anti-HIV-1 immunity (*Ondondo et al., 2016*; *Borducchi et al., 2016*; *Mothe et al., 2015b*; *Achenbach et al., 2015*; *Hanke, 2014*).

Here, we present the first study reporting within-host HIV-1 escape frequencies in PLWH on ART. A non-selective approach was taken, specifically sequencing HIV-1 outgrowth viruses (OGVs) from the reservoir induced from resting CD4+ T cells, and mapping circulating T cell responses against

the entire HIV-1 proteome in 25 PLWH on ART. These strategies enabled us to empirically determine the extent and patterns of pre-ART escape within the stable HIV-1 reservoir. We observed that the majority, 68%, of T cell epitopes in the HIV-1 reservoir did not harbor escape variants and were recognized by circulating T cells. Furthermore, analysis of 'escape' epitopes (the remaining 32%) showed that escape was significantly lower in regions of HIV-1 that are more conserved. We conclude that the majority of circulating HIV-1-specific T cells in PLWH on ART can recognize the stable HIV-1 reservoir.

Results

Study cohort

HIV-1-specific T cell responses and the size of the HIV-1 reservoir were measured in 25 PLWH (ages 25–66, mean 46; male = 16, female = 9) durably suppressed by combination ART (viral load <50 copies/ml for 1.6–20.1 years, mean 7.2 years). Of the 25 participants, four initiated ART during acute HIV-1 infection, while the other 21 initiated ART during chronic HIV-1 infection (see Materials and Methods for definitions). Most participants (n = 24) were HLA-typed. Characteristics of the cohort are summarized in *Table 1* and full clinical details and IRB approvals are provided in Materials and Methods.

HIV-1 outgrowth viruses in PLWH on ART show no directional selection over time

The HIV-1 reservoir in resting CD4+ T cells was measured using quantitative viral outgrowth assays (QVOAs) (*Lee et al., 2017*; *Siliciano and Siliciano, 2005*; *Trumble et al., 2017*; *Ho et al., 2013*; *Massanella and Richman, 2016*) for all 25 participants (*Figure 1*). Longitudinal IUPM measurements were obtained for 10 participants (*Supplementary file 1*). From these data, we calculated the mean of IUPM measurements (mean number of IUPM = 2/participant, range 1–6) made in the 2 years prior to HIV-1 T cell epitope mapping described below. Within the cohort, measurements ranged from below the limit of detection (participants 00941 and 20296, see Material and Methods for definition of limit of detection) to 2.528 IUPM (*Table 1*, *Supplementary file 1*).

OGVs were then sequenced. Primer ID sequencing (*Lee et al., 2017*; *Emery et al., 2017*), which allows for simultaneous in-depth sequencing of multiple HIV-1 regions was performed in one participant (participant 00673, mean of 3,700 Primer ID template consensus sequences/epitope). This approach was replaced with PacBio SMRT sequencing of the p24+ wells *at endpoint dilution* (i.e. average one virus-infected cell/well) to enable nearly, full-length sequencing of the HIV-1 genome. PacBio sequencing provided long reads (two reactions spanned almost the entire length of HIV-1) which identified sites of virus diversity within and outside epitopes that were reactive in initial T cell mapping (see Materials and Methods describing 'stripes'). Note that given the initial effective population size in each well was, on average, n = 1 (*Rouzine et al., 2001*), in vitro mutations occurred stochastically, with only 4–5 mutations expected across the virus genome over the 15 day QVOA culture (*Song et al., 2012*; *Brown and Richman, 1997*). Mutations arising in culture were therefore not expected to bias results.

Phylogenetic analysis of the 22 participants for whom near full-length sequencing was available (mean = 16 sequences per participant, range 1–71) confirmed that sequences from each participant formed monophyletic clades (*Figure 2A and B*). Participants were infected with HIV-1 clade B, except for one participant (participant 00926) who was infected with HIV-1 clade G. Available HIV-1 sequences from participants treated during acute infection (n = 3) showed greater genetic similarity (maximum pairwise diversity 0.0007) than sequences from participants treated during chronic infection (n = 19) (maximum pairwise diversity 0.0088) (Mann-Whitney two tailed test, p=0.003, *Supplementary file 2*). This is consistent with reports in which most HIV-1 infections are established from a single HIV-1 variant, and therefore the early onset of ART, which halts viral evolution, results in a smaller and more homogenous HIV-1 reservoir (*Kearney et al., 2014*; *Keele et al., 2008*; *Haaland et al., 2009*; *Derdeyn et al., 2004*; *Abrahams et al., 2009*).

For seven participants, OGVs from longitudinal samples were available for sequencing (up to 4.4 years, 2–4 time points/participant) (*Figure 3*). OGV sequences from different timepoints for each participant were interspersed within-host phylogenetic trees, suggesting no major directional

Table 1. Demographic and clinical data of participants (n = 25).

Participant	Sex	Race[*]	Ethnicity[†]	Status at diagnosis[‡]	Age[§]	Years suppressed[¶]	Hiv-1 RNA[**]	Nadir CD4	Mean IUPM [n][††]	HLA type
00231	M	C	NH	A	66	9.66	1.5E+07	277	0.839 [1]	A*11:01, A*24:02, B*35:TDS[‡‡‡], B*35:02, C*04:01, C*04:01
00250	M	C	NH	A	48	7.44	6.8E+05	403	0.318 [3]	A*01:01, A*31:01, B*27:03, B*44:02, C*02:02, C*06:02
00531	M	C	NH	C	52	14.87	3.8E+05	81	1.372 [5]	A*02:01, A*02:01, B*07:02, B*44:02, C*03:04, C*07:02
00546	M	C	NH	C	56	20.09[§§]	N/A	195	0.453 [2]	A*02:01, A*32:01, B*15:01, B*40:01, C*02:02, C*03:04
00673	M	C	NH	C	49	8.05	3.2E+05	168	0.790 [4]	A*24:02, A*30:02, B*15:17, B*18:01, C*05:01, C*07:01
00674	M	C	NH	C	57	6.25	6.8E+04	338	2.064 [5]	A*02:01, A*03:01, B*07:TDVB[§§§], B*40:01, C*03:04, C*07:02
00720	M	C	NH	C	25	2.96	4.0E+05	168	0.499 [4]	A*01:01, A*02:01, B*07:02, B*18:01, C*07:02, C*12:03
00728	M	C	NH	C	30	6.00	1.8E+04	354[¶¶]	0.676 [6]	A*02:01, A*03:01, B*07:02, B*40:02, C*02:02, C*07:02
00749	M	AA	NH	C	25	4.10	N/A	402	0.365 [2]	A*03:01, A*30:02, B*15:03, B*40:01, C*02:10, C*03:04
00773	M	C	NH	A	46	2.01	2.4E+04	234	0.233 [1]	A*02:01, A*11:01, B*18:01, B*51:01, C*01:02, C*02:02
00833	M	AA	NH	C	36	8.16	7.9E+03	358[***]	0.048 [1]	A*02:02, A*74:01, B*15:16, B*82:01, C*03:02, C*14:02
00834	M	AA	NH	C	25	2.15	2.1E+04	365	0.479 [1]	A*02:05, A*30:02, B*51:01, B*58:01, C*04:01, C*16:01
00848	F	C	NH	C	44	4.78	1.4E+04	800	0.022 [1]	A*01:01, A*30:02, B*08:01, B*18:01, C*05:01, C*07:01
00870	M	C	NH	C	54	5.28	5.3E+04	154[†††]	1.400 [1]	A*11:01, A*24:02, B*35:01, B*55:01, C*03:03, C*04:01
00879	F	AA	NH	C	54	1.60	4.5E+02	612	0.093 [1]	A*03:01, A*33:01, B*14:02, B*58:02, C*06:02, C*08:02
00926	F	AA	NH	C	38	3.83	1.3E+05	508	0.262 [1]	A*02:02, A*02:05, B*:39:10, B*45:01, C*12:03, C*16:01
00929	F	AA	NH	C	61	1.77	1.9E+04	528	0.268 [1]	A*01:01, A*74:AB[****], B*15:MJMN[¶¶¶], B*81:AB[****], C*02:10, C*18:01
00930	M	C	NH	C	53	2.70	6.4E+05	174	2.084 [1]	A*01:01, A*11:01, B*08:01, B*40:01, C*03:04, C*07:01
00937	F	AA	NH	C	45	5.84	8.8E+04	260	1.276 [2]	A*03:01, A*30:01, B*35:01, B*52:01, C*04:01, C*16:01
00941	F	C	NH	C	47	14.58	6.6E+03	12	1.50E-08[‡‡][1]	N/A
01014	F	C	H	C	52	19.85	1.2E+06	16	0.338 [1]	A*23:01, A*74:01, B*07:02, B*41:02, C*15:05, C*17:03
01086	F	AA	NH	C	30	5.81	N/A	320	0.0529 [1]	A*02:05, A*03:01, B*07:06, B*53:01, C*04:01, C*07:02
01095	F	C	NH	C	54	6.56	1.1E+05	88	2.528 [1]	A*03:01, A*31:01, B*40:01, B*44:02, C*03:04, C*05:01
10430	M	C	NH	C	59	9.47	8.1E+04	78	1.367 [5]	A*02:01, A*30:01, B*13:02, B*51:01, C*06:02, C*14:02
20296	M	C	NH	A	35	6.90	2.5E+04	427	1.50E-08[‡‡] [1]	A*29:02, A*31:01, B*08:01, B*44:03, C*07:01, C*16:01
				Mean:	46	7.22	8.6E+05	293	0.775	
				Range:	25–66	1.6–20.10	4.5E+02-1.5E+07	12–800	1.50E-08–2.528	

selection over time while participants remained suppressed on ART (*Figure 3*, *Figure 3—figure supplement 1*). Note, the single long-branched sequence in participant 10430's phylogeny met quality scores for sequencing and checks for hypermutations without evidence of mislabeling and was therefore retained. Overall, these longitudinal data are consistent with observations that ART effectively halts viral replication and subsequent evolution (*Kearney et al., 2014*; *Nettles, 2005*). Given the lack of directional selection, longitudinal sequences for each participant were pooled, increasing the mean sequencing depth of OGVs from 16 to 25 sequences (range: 1–71) per participant (*Figure 3—figure supplement 2*). In general, identical OGV sequences (see definition in Materials and Methods) were identified only infrequently; however, the 5' and 3' half genome sequences from two participants (participants 00250 and 00749, *Figure 3A and E*) exhibited >20% clonality, consistent with reports that in that the HIV-1 reservoir is in part maintained by clonal expansion (*Supplementary file 3*; *Bui et al., 2017*).

The HIV-1 reservoir retains susceptibility to HLA-associated T cell responses

Bioinformatics approaches have been useful in identifying HIV-1 polymorphisms that are statistically significantly enriched among individuals expressing specific HLA class I alleles, identifying these as putative sites of viral escape (*Carlson et al., 2012*). Many of these HLA-associated polymorphisms have been experimentally verified as immune escape mutations, and lists of these HLA-associated polymorphisms are particularly comprehensive for HIV-1 subtype B (*Carlson et al., 2012*). We examined the OGV sequences recovered from 21 clade B-infected participants for whom both HIV-1 sequencing and HLA typing had been performed for the presence of known HLA-associated polymorphisms. For each HIV-1 sequence, each HLA-associated viral site (total number of HLA-associated sites n = 2,819, *Supplementary file 4*) was classified as either 'nonadapted', 'adapted', or 'possibly adapted'. 'Nonadapted' viral sites exhibited the specific HIV-1 residue predicted to be susceptible to the restricting HLA, 'adapted' sites exhibited the specific HIV-1 residue predicted to confer escape from the restricting HLA, and 'possibly adapted' sites exhibited any residue other than the 'nonadapted' form, supporting it as a possible escape variant. Minimum and maximum bounds of HLA-associated adaptation in each participant's reservoir were estimated by computing the total proportion of HLA-associated sites exhibiting 'adapted' forms (minimum bound), and those exhibiting either 'adapted' or 'possibly adapted' forms (maximum bound) (*Figure 4A*, *Supplementary file 4*; *Carlson et al., 2012*). Within participants, the proportion of HLA-associated sites exhibiting the HLA-adapted form ranged from 9–33% (adapted) and 32–66% (adapted + possibly adapted), and varied markedly by HIV-1 protein (*Carlson et al., 2012*). Importantly, our observation that, across all HIV-1 proteins, the minimum and maximum bounds for within-host escape burden ranged from 20% (adapted only) to 48% (adapted + possibly adapted) meant that between 52 and 80% of HLA-associated sites in the HIV-1 reservoir were predicted to retain susceptibility to host HLA-restricted T cell responses.
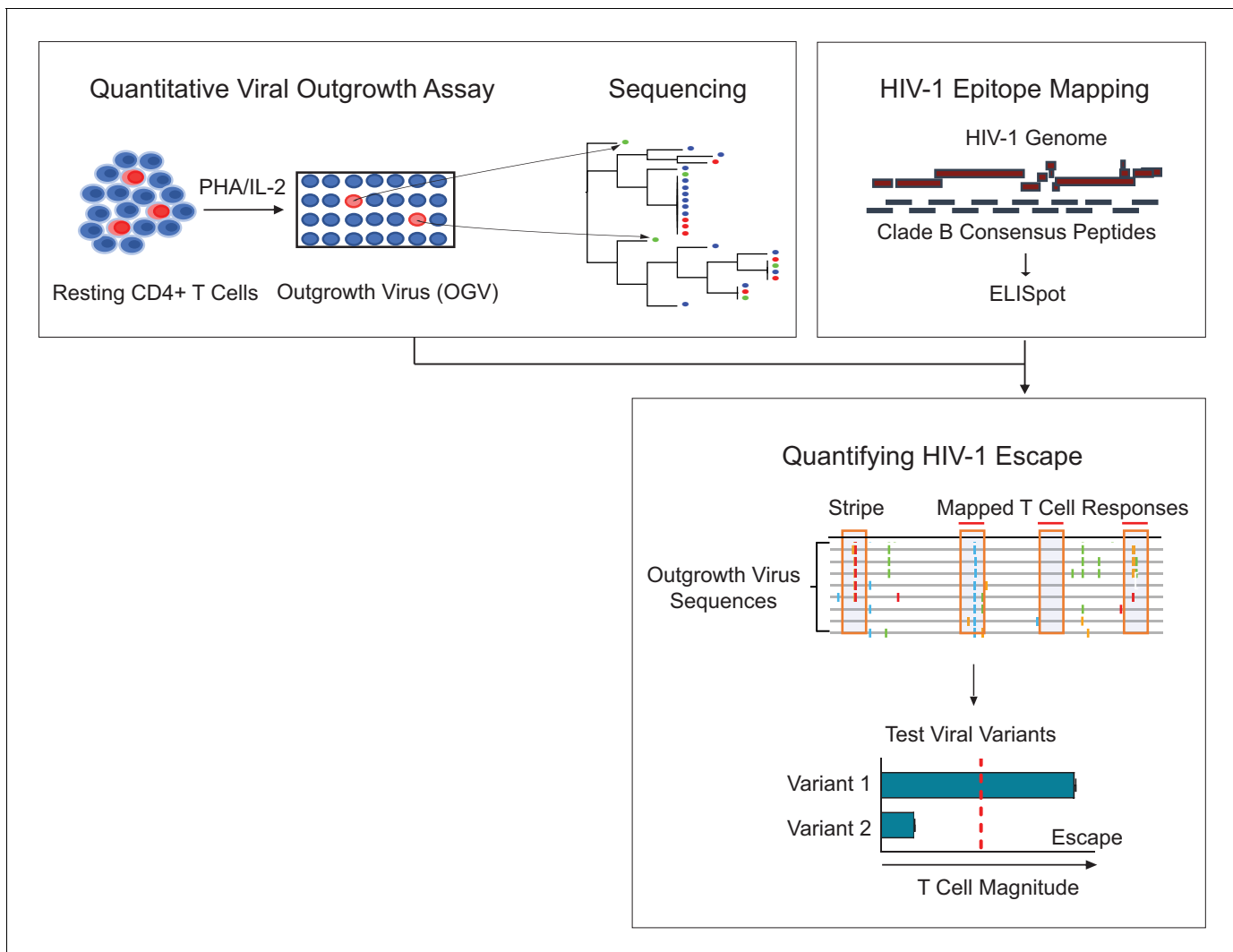
Figure 1. Schema of empirical testing of T cell escape in the HIV-1 reservoir. Top left panel. The quantitative viral outgrowth assay (QVOA) provided a minimal estimate of the size of the replication-competent reservoir virus from total resting CD4+ T cells of PLWH on ART (n = 25). Supernatants from HIV-1 p24+ wells harboring outgrowth virus (OGV) were sequenced by either PacBio (n = 22) or MiSeq (n = 1). Top right panel. Ex vivo IFN-γ ELISpot assays identified reactive T cell epitopes in clade B HIV-1 proteome in each participant. Lower panel. Sequencing and T cell data were combined to identify putative virus escape mutations by overlaying reactive T cell reactive epitopes (red horizontal line) on highlighter plots of each participant's outgrowth sequences (n = 23). Peptides spanning <u>all</u> non-synonymous changes between the clade B-defined epitope and outgrowth sequences (orange boxes) were synthesized. In addition, peptides matching OGVs in participants in which > 40% of viruses contained the same non-synonymous mutation generating a 'stripe' on the highlighter plot were also synthesized (orange box, no overhead red line). Variant peptides were examined for their impact on T cell recognition in ex vivo IFN-γ ELISpot. Virus escape was defined as a ≥ 50% difference in magnitude between variants, either Clade B sequence or OGV variants (vertical, red dotted line).

HIV-1-specific T cell responses are detectable in PLWH on ART

To empirically assess the level of T cell escape, we mapped HIV-1-specific T cell responses in the 25 PLWH on ART. T cell responses were quantified by ex vivo IFN-γ ELISpot using a matrix method containing overlapping peptides spanning the consensus sequence of the entire clade B HIV-1 proteome (*Supplementary file 5*; *Goonetilleke et al., 2009*). While HIV-1-specific T cell responses remain detectable in PLWH on ART, the overall magnitude of the T cell response is ~10 fold lower than in untreated HIV-1 infection, which may result in the underestimation in the breadth of T cell responses (*Xu et al., 2019*; *Addo et al., 2003*). To increase ELISpot assay sensitivity, we added a higher number of PBMCs, 4×10^5 cells, to each well. After the initial matrix ELISpot, individual 18 amino acid (18-mer) peptides were tested to confirm reactive T cell responses. In cases in which overlapping 18-
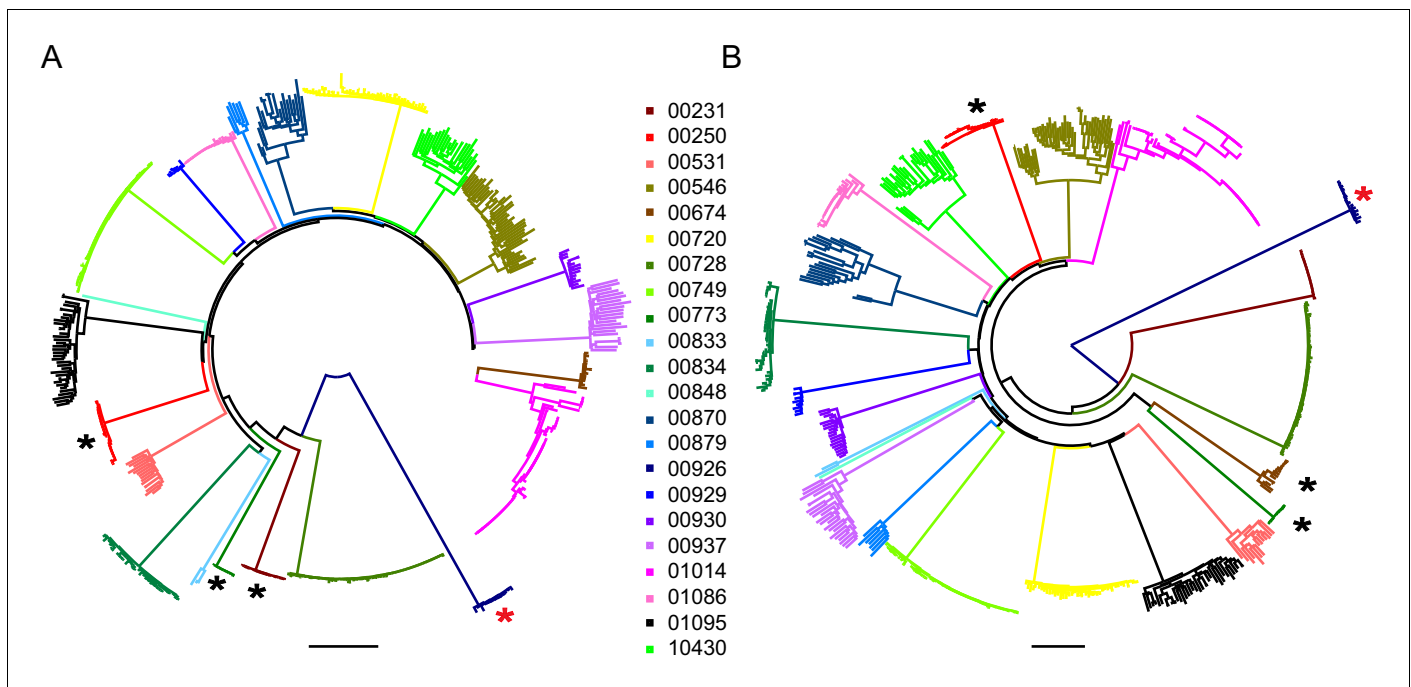
Figure 2. Neighbor-joining phylogenetic trees of sequenced outgrowth virus. The phylogenetic trees display 5' (**A**) and 3' (**B**) half genome sequences generated by PacBio sequencing of outgrowth virus from the quantitative viral outgrowth assay from n = 22 PLWH on ART. Each color indicates outgrowth virus sequences from a different PLWH on ART. Black stars denote participants who initiated ART during acute infection. All participants were infected with a clade B HIV-1 virus, except for one participant (red star), who was infected with a clade G HIV-1 virus. The genetic distance is shown by scale bar, 2 substitutions out of 100 bp. Trees are midpoint-rooted.

mer peptides were both reactive, we assumed responses were driven by a single reactive T cell epitope and only the higher magnitude response was designated reactive. Across the 25 participants, a total of 166 reactive T cell epitopes were identified (see Methods for positivity criteria). HLA typing was used to predict optimal CD8+ T cell epitopes, which were tested alongside 18-mer peptides. Optimal T cell epitopes, defined as inducing a T cell response ≥ than the corresponding 18-mer, were confirmed for 25/166 epitopes (*Supplementary files 5* and *6*).

The mean summed magnitude of HIV-1-specific T cell responses was 1,033 SFU/M (range: 156–2,856 SFU/M), and the mean T cell breadth was seven reactive epitopes (range: 1–19 epitopes) per participant (*Table 2*). Predictably, and consistent with reports in untreated HIV-1 infection, the magnitude of the mean summed HIV-1-specific T cell response correlated positively with the breadth of reactive epitopes targeted (r = 0.57, Spearman Rank two-tailed, p=0.003, 25 pairs, *Supplementary file 7*; *Addo et al., 2003*). As noted above, while the magnitude of the HIV-1-specific T cell response was weaker compared to untreated HIV-1 infection, the pattern of HIV-1 proteins targeted by T cells was similar to previous reports of untreated infection (*Figure 4B*; *Xu et al., 2019*; *Addo et al., 2003*). For example, the highest magnitude T cell response within a participant (immunodominant T cell response) most frequently targeted Gag and Nef, whereas the most frequently targeted protein across the cohort (immunoprevalent T cell response) was Gag (*Figure 4— figure supplement 1*; *Addo et al., 2003*). No differences in either T cell magnitude or breadth were observed between women and men, or between participants treated during acute versus chronic HIV-1 infection (*Supplementary file 7*). Neither T cell magnitude nor breadth exhibited any association with measured HIV-1 IUPM (*Supplementary file 8*; *Xu et al., 2019*).

Escape is not observed in the majority of T cell epitopes in PLWH on ART

To empirically quantify T cell escape in this cohort, OGV variants were tested for their effect on T cell response (*Figure 4C*). In participants for which > 1 OGV sequences were obtained (n = 23 participants and collectively 151 epitopes), peptide/s were synthesized to match each participant's OGV
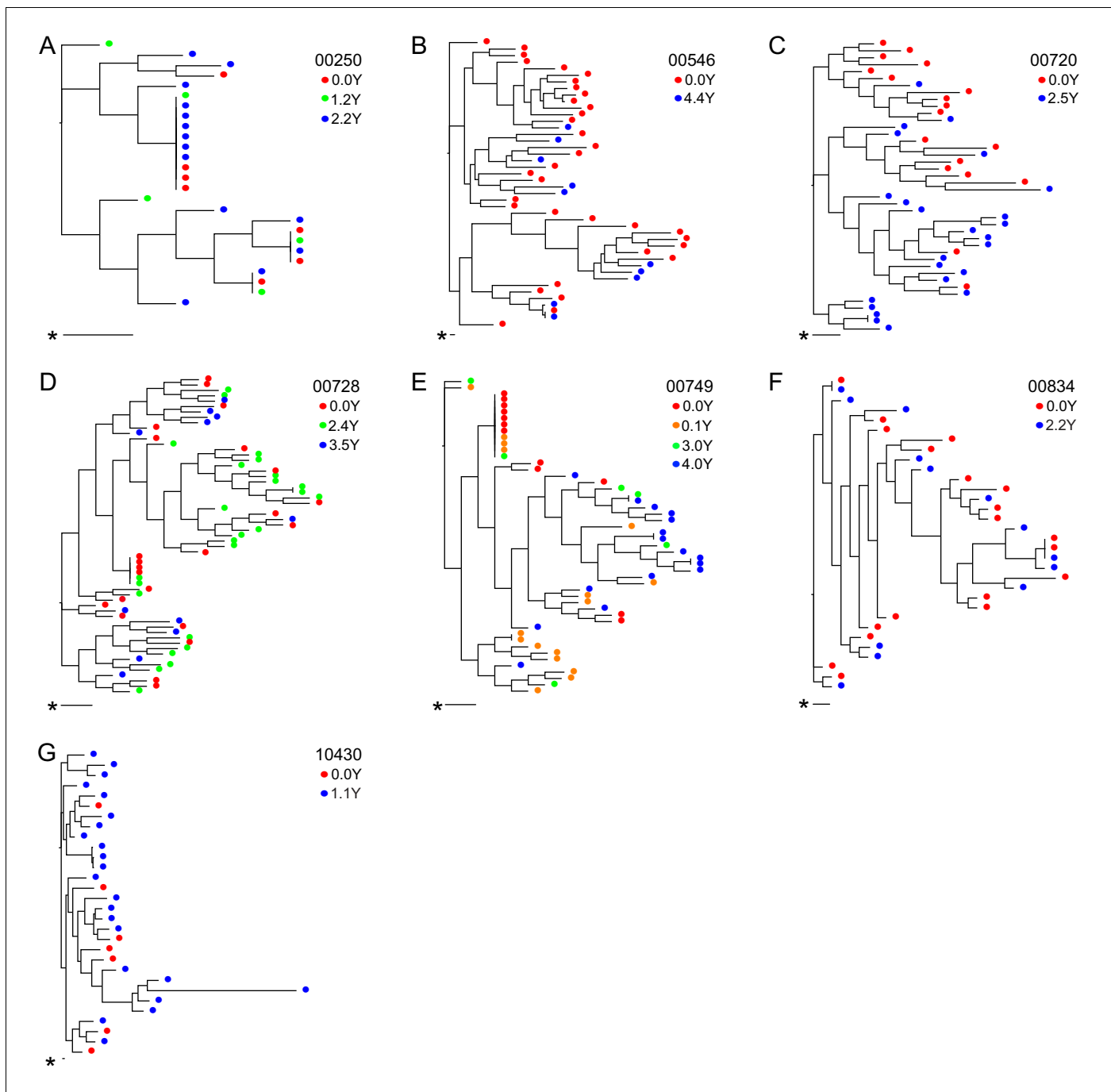
Figure 3. Neighbor-joining within-host phylogenetic trees inferred from longitudinal outgrowth virus sequences. Phylogenetic trees displaying 3' half genome sequences of longitudinal outgrowth virus sequences from seven participants (A–G). Each color represents a different time point in which outgrowth virus was sequenced for each participant. Sequences from different time points for each participant are distributed throughout the phylogeny. The genetic distance is shown by scale bar (indicated with an asterisk), 1 substitution out of 1000 bp. Trees are midpoint-rooted. (A) 00250 treated during acute infection and (E) 00749 have > 20% clonal sequences.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Longitudinal outgrowth virus sequences on neighbor-joining phylogenetic trees.

Figure supplement 2. Size of the replication-competent reservoir as measured by infectious units per million (IUPM) against the number of sequences obtained by PacBio sequencing (r = 0.755, p<0.0001, Spearman Rank, n = 24 pairs).
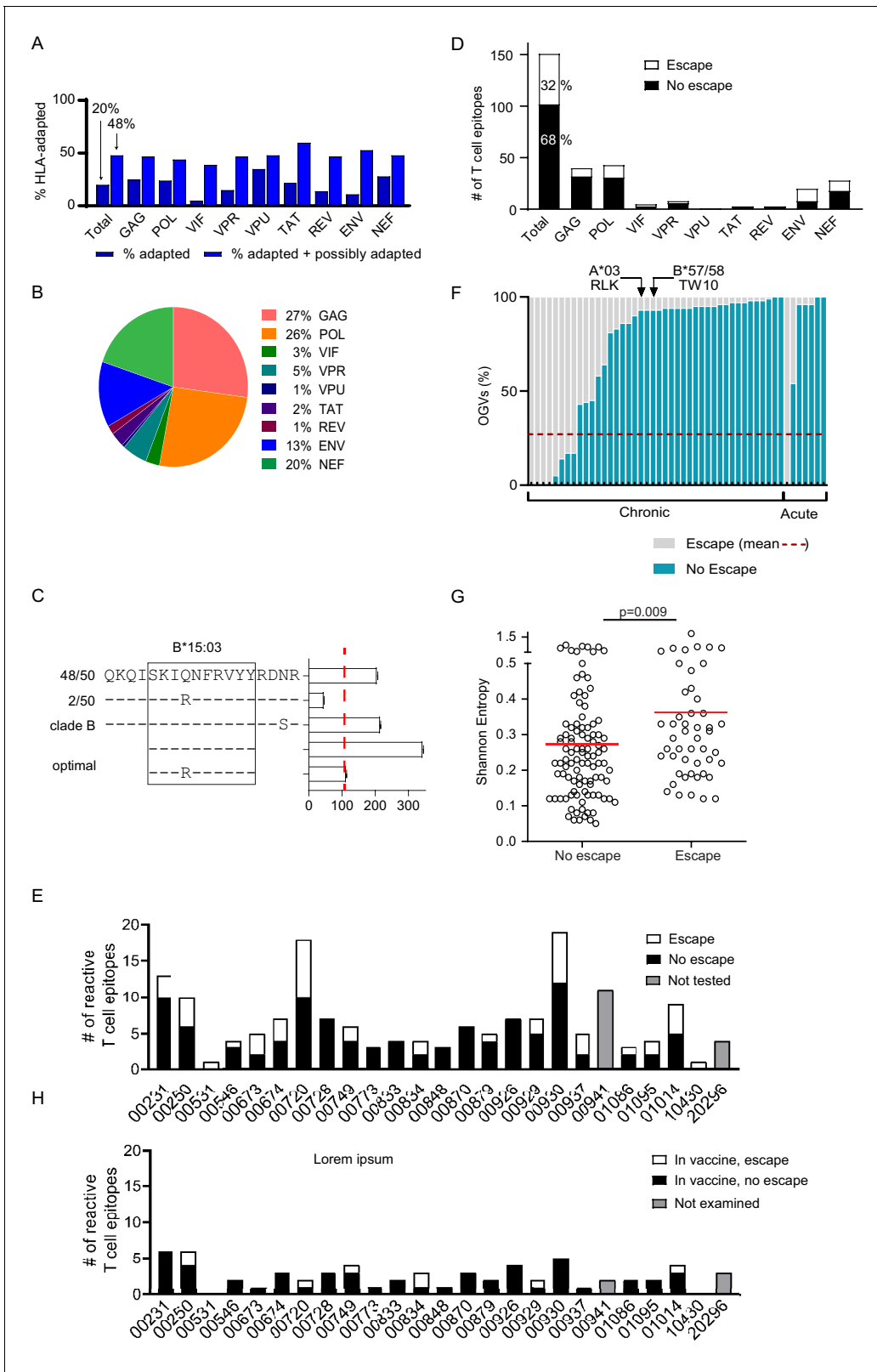
Figure 4. Quantifying T cell escape in the HIV-1 reservoir. (**A**) *HLA associations in OGVs*: The proportion of polymorphisms harboring HLA-adapted (blue bar) and HLA-adapted + possibly HLA-adapted polymorphisms (light purple) in PLWH on ART (Clade B infected with full-length sequencing, n = 22). Results are summarized as total, as well as by individual HIV-1 protein. (**B**) *T cell mapping*: Ex vivo IFN-γ ELISpot was performed to determine T cell responses (n = 166), shown by protein, against the clade B HIV-1 proteome in 25 PLWH on ART. (**C**) *Escape testing*: Variants detected in OGVs of

Figure 4 continued on next page

Figure 4 continued

participant 00749 were tested ex vivo using IFN-γ ELISpot alongside the original reactive clade B 18-mer and optimal peptide. Bars show the T cell magnitude reported as spot forming units per million (SFU/M) PBMCs. Escape was defined as a ≥ 50% decrease (red vertical dotted line) in the mean magnitude of triplicate wells compared to the peptide which elicited the greatest T cell response. The left side shows the proportion of each variant in total OGVs (n = 50). (**D**) *Cohort-level escape*: T cell escape was examined in participants for which OGV sequence data were available (n = 23 participants, n = 151 epitopes). The proportion of empirically confirmed T cell epitopes that harbored escape variant(s) (n = 49, white) or no escape variant(s) (n = 102, black) in OGV. Results are summarized in total, as well as by individual HIV-1 protein. (**E**) *Escape by participant*: The proportion of reactive T cell epitopes that harbored escape variant(s) (white) and no escape variant(s) (black) per participant (n = 23). Gray bars indicate participants (n = 2) where epitopes (n = 15) were not tested for escape due to limited OGVs. (**F**) *Within-epitope escape*: Stacked bar graph showing the proportion of individual OGVs that conferred (gray –escape) or did not confer (blue – no escape) T cell escape in each epitope for which escape was observed (n = 49). Red dotted line is the average (27%) frequency of OGVs that conferred T cell escape across all epitopes. Data are ordered in ascending order of no escape OGVs, grouped by whether the participants initiated ART in either acute or chronic HIV-1 infection. Arrows identify immunoprevalent T cell epitopes, in which the majority of OGVs do not harbor escape variants. (**G**) *Shannon entropy*: A Shannon entropy score was calculated for reactive T cell epitopes for which sequence data were available (n = 151, 23 participants). Entropy was higher in epitopes which contained T cell escape variant(s) (n = 49, median = 0.29) than epitopes with no escape variant(s) (n = 102, median = 0.23) in the OGVs (Mann Whitney two-tailed test, p=0.009). (**H**) *Escape in vaccine immunogen*: The proportion of reactive T cell epitopes in each participant (n = 25) that are targeted by the HIVconsvX immunogen (total = 59) and harbored one or more escape variants (white), or no escape variants (black) in OGVs. Gray bars indicate participants (n = 2) where epitopes in the immunogen could not be examined for escape due to limited OGVs.

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. HIV-1-specific T cell responses in PLWH on ART.

Figure supplement 2. Outgrowth virus sequences and DNA sequences on neighbor-joining phylogenetic trees.

autologous virus sequence/s. This included any viral variants that differed from the clade B peptide used to originally map the respective reactive T cell epitope. In total, 115 of 151 epitopes contained variant OGVs.

All variant peptides were tested ex vivo in an IFN-γ ELISpot assay alongside the originally tested 18mer clade B consensus peptide, and if identified, the optimal epitopes. Escape was defined as a ≥ 50% decrease in the average magnitude of triplicate wells compared to the peptide that elicited the greatest T cell response, whether autologous peptide or clade B consensus peptide (*Figure 4C*, *Supplementary file 6*). We would note that a range of escape thresholds, ranging from 50–80% (*Borrow et al., 1997*; *Goonetilleke et al., 2009*; *Radebe et al., 2015*), have been applied in previous HIV-1 studies. In this study, our goal was to maximize sensitivity to detect escape by identifying all outgrowth sequences in our dataset with the potential to compromise T cell immunity; therefore, the 50% threshold was applied.

Most variant peptides did not impact T cell magnitude. In total, virus escape was confirmed in the replication-competent HIV-1 reservoir in 49 of 151 (32%) HIV-1 T cell epitopes (*Figure 4D*) with escape occurring more frequently in epitopes in Pol, Env and Nef (25%, 23%, 21%) proteins (*Figure 4D*).

As described above, participants recognized an average of 7 T cell epitopes, ranging from 1 to 19. In each participant, one third (average = 34%) of reactive epitopes harbored virus escape mutations (*Figure 4E*, *Table 2*). In six participants, no T cell escape was observed in any reactive T cell epitope (i.e. 0% escape in 3–7 epitopes, additional DNA sequencing in *Figure 4—figure supplement 2*), whereas in the two participants who responded to only one epitope, virus escape was confirmed in both cases (i.e. 100% escape) (*Figure 4E*, *Table 2*).

At each epitope, only a minority of outgrowth viruses confer T cell escape

We next examined each of the 49 epitopes in which escape was observed to quantify the fraction of sequenced OGVs that conferred escape that is the 'depth' of escape (*Figure 4F*). On average for each epitope, 27% of OGV sequences tested produced a ≥ 50% loss in T cell response (*Figure 4F*, *Supplementary file 6*). The range of within-epitope escape was broad, starting at 1–2% to 100% of all OGVs. Note, that an escape depth of 100% (which was observed for n = 5 epitopes) occurred when all OGVs were identical but induced a lower T cell response than the originally tested consensus clade B epitope. In these cases, we assumed that the primary T cell response was induced against a consensus epitope no longer detectable in the reservoir. No differences in the depth of

Table 2. HIV-1-Specific T cell reponses (n = 25 participants).

Participant	Summed magnitude (10⁶/PMBC)	Immunodominant T cell response	Total T cell breadth	# of epitopes with escape variant(s) in outgrowth virus	Escape in immunodominant epitope (Y/N)
00231	2856	Vif153→170	13	3	Y
00250	1531	Env765→782	10	4	Y
00531	390	Gag357→374	1	1	Y
00546	895	Gag260→277	4	1	N
00673	2313	Nef73→90	5	3	N
00674	1394	Env838→855	7	3	Y
00720	1947	Pol521→538	18	8	Y
00728	439	Gag204→221	7	0	N
00749	418	Pol481→498	6	3	N
00773	1174	Pol273→290	3	0	N
00833	333	Gag164→181	4	0	N
00834	725	Pol273→290	4	2	N
00848	276	Nef89→106	3	0	N
00870	363	Nef65→82	6	0	N
00879	500	Gag292→309	5	1	N
00926	234	Gag180→197	7	0	N
00929	798	Gag180→197	7	2	N
00930	2182	Nef81→98	19	7	Y
00937	297	Pol257→274	5	3	N
00941	2216	Nef81→98	11	N/A[*]	N/A[*]
01014	1463	Nef65→82	9	4	N
01086	258	Gag17→34	3	1	Y
01095	2337	Pol313→330	4	2	N
10430	157	Nef97→114	1	1	Y
20296	331	Gag204→221	4	N/A[*]	N/A[*]
Total			166	49 (32% - 49/151)	
Mean:	1033		7	2	
Range:	157–2856		1–19	0–8	

[*]No outgrowth virus sequence data available to test for the presence of an escape variant in the sequenced outgrowth virus.

virus escape was observed between participants treated in acute or chronic infection (*Figure 4F*, Mann-Whitney two-tailed test, p=0.216). Our mapping included multiple, well-characterized immunoprevalent epitopes that also harbored a mix of escape and non-escape OGVs, but overall, mostly non-escape (*Figure 4F*, *Supplementary file 5*).

Epitope entropy but not T cell immunodominance is associated with escape in PLWH on ART

We hypothesized that escape, as observed in studies of untreated HIV-1 infection, would be associated with less conserved regions of HIV-1. Shannon entropy is used to measure the level of HIV-1 population diversity at each position along the HIV-1 proteome (*Shannon, 1948*). Here, epitope entropy was calculated as previously described (*Liu et al., 2013*), as the mean entropy of each position within the epitope. High entropy values correspond to sites with more population diversity, that is sites that are less conserved (*Allen et al., 2005*; *Shannon, 1948*). As expected, escape was more

commonly observed in regions of higher Shannon entropy (Mann-Whitney two-tailed test, p=0.009) (*Figure 4G*).

While our previous work in untreated HIV-1 infection reported that immunodominance is a major determinant in the time to viral escape, immunodominance of circulating T cell responses in PLWH on ART did not associate with the presence or absence of escape within the cognate epitope (Mann-Whitney two-tailed test, p=0.971, *Figure 4—figure supplement 1*). This may be because effective ART in this study cohort prevented directional selection and immunodominance hierarchies subsequent to the initial escape event may have shifted, whether in untreated infection or at the time of ART initiation.

Conserved HIV-1 vaccine immunogens preferentially target non-escaped T cell epitopes

Our and others' observations that T cell escape occurs more frequently in higher entropy (less conserved) epitopes (*Liu et al., 2013*; *Ferrari et al., 2011*; *Barton et al., 2016*) suggests that vaccine-induced HIV-1-specific T cells targeting low entropy (more conserved) regions may contribute to better HIV-1 control. This may be due to HIV-1-specific T cells maintaining selection pressure on regions of the virus where escape is likely to incur a fitness cost. Several conserved-element vaccine immunogens have been designed which connect the most conserved parts of the viral proteome, either directly in-frame or with linker sequences to form artificial conserved element immunogens. Vaccination with these immunogens aims to focus T cell responses against highly conserved epitopes that are generally subdominant in natural infection (*Barouch and Korber, 2010*; *Rolland et al., 2007*; *Mothe et al., 2015a*; *Borthwick et al., 2017*; *Gorse et al., 2008*). We interrogated four different conserved vaccine immunogen designs, HIVconsv (*Létourneau et al., 2007*), HIVconsvX (*Ondondo et al., 2016*), HIVACAT T cell Immunogen (HTI) (*Mothe et al., 2015a*), and the p24 conserved element (p24 CE) (*Kulkarni et al., 2013*), using our dataset to examine (1) whether the immunogen contained epitopes targeted by circulating T cell responses in PLWH on ART and (2) whether the epitopes in the immunogen contained escape variant(s) found in OGVs recovered from of our participants.

The 151 epitopes that were targeted by T cells in our study participants were differentially covered by these four immunogens. The HIVconsX immunogen contained the largest number of circulating T cell responses observed in our dataset (59/151; 39%; mean of 3 epitopes/participant, range 0–6) while the p24 CE vaccine contained the least (16/151; 11%; mean of 1 epitope/participant, range 0–4) (*Figure 4H*, *Supplementary file 9*). Nevertheless, escape frequency was on average two-fold lower in reactive T cell epitopes (12–29%) found within the immunogens relative to those excluded from immunogen designs (33–44%). For example, the overall proportion of reactive T cell epitopes within the HIVconsX immunogen that contained escape variants was 15.3% (9/59 epitopes, average 0 escape epitopes/participant, range 0–2), whereas the proportion of reactive T cell epitopes outside of the HIVconsX immunogen that contained escape variants was almost >2.5-fold higher (43.5%, average two escape epitopes/participant). This difference was statistically significant (p=0.001, Pearson's χ^2 test with Yates correction, *Supplementary file 9*). Similarly, the other three conserved immunogen vaccines also targeted T cell epitopes with a low rate of escape (*Létourneau et al., 2007*; *Mothe et al., 2015a*; *Kulkarni et al., 2013*). This suggests that current conserved immunogen designs are likely to be immunogenic in PLWH on ART and are more likely to target T cell epitopes that do not harbor pre-ART escape.

Discussion

In this study, we employed two unbiased approaches to comprehensively examine the landscape of escape in the HIV-1 reservoir in 25 PLWH durably suppressed on ART. First, we examined the frequency of published HLA-associated polymorphisms in HIV-1 subtype B OGV sequences. Overall, 20% of HLA-associated sites in the HIV-1 reservoir harbored the known HLA-adapted viral form, suggesting potential viral escape from T cell responses. This number increased to 48% if both known and putative HLA-adapted forms were included in the calculation. These estimates are consistent with substantial immune escape in the reservoir, but nevertheless support the notion that the majority of HLA-associated sites in the reservoir maintain susceptibility to host T cell responses. Next, in all participants, we mapped HIV-1-specific T cell epitopes across the full HIV-1 proteome. Across the

cohort, the mean breadth of HIV-1 epitopes targeted per person was 7. T cell targeting of HIV-1 proteins was consistent with studies of untreated infection, with T cells most often targeting Gag, Pol, and Nef (*Addo et al., 2003*). No differences were observed between treatment initiated in acute (n = 3) or chronic infection (n = 22). We then systematically examined whether non-synonymous mutations in reactive epitopes impacted T cell recognition and functional responses. Consistent with the overall HLA-adaptation estimates from the bioinformatics analysis, we experimentally confirmed that 32% (49/151) of all reactive T cell epitopes in HIV-1 OGVs contained mutations that conferred T cell escape, and that conversely, 68% of circulating HIV-specific T cell responses were able to fully recognize the autologous replication-competent-reservoir virus.

The minimal empirically defined T cell escape frequency of 32% observed in PLWH on ART in this study is consistent with studies of T cell escape in untreated HIV-1 infection. A 30% escape frequency was observed in primary untreated HIV-1 infection, in which HIV-1-specific T cell responses were comprehensively mapped against the autologous transmitted/founder virus in participants, and T cell reactivity was assessed against mutant viruses that emerged over the first 6 months of infection (*Liu et al., 2013*). In that study, similar to ours, CD8+ T cell escape occurred more frequently in epitopes of higher Shannon entropy (*Liu et al., 2013*; *Ganusov et al., 2011*; *Bansal et al., 2005*). A similar escape frequency of 33% was also observed in a large cross-sectional study of chronically infected HIV-1 Clade C untreated individuals (n >2,000), in which the T cell escape frequency was determined by examining participant HLA-associated polymorphisms in immunoprevalent HIV-1 CD8+ T cell epitopes (*Carlson et al., 2012*). In PLWH durably suppressed on ART, HIV-1-specific T cell responses are stable and maintain immunodominance hierarchies over years (*Xu et al., 2019*). Both the stability of the HIV-1-specific T cell response and the comparable level of population-level virus escape in treated and untreated individuals is consistent with studies of HIV-1 reservoir formation and kinetics. These studies showed that the HIV-1 reservoir is mostly formed at the time of ART initiation (*Brodin et al., 2016*; *Abrahams et al., 2019*; *Pankau et al., 2020*) and its subsequent rate of decay is very slow (*Crooks et al., 2015*; *Siliciano and Siliciano, 2015*; *Siliciano and Siliciano, 2004*), with only limited variation in the number of proviruses, or the types of proviral defects observed in PLWH on ART over time (*Lu et al., 2018*). More recently, longitudinal analyses of near full-length proviral sequences in PLWH on ART showed no enrichment in escaped epitopes over time suggesting that HIV-1-specific T cells do not significantly alter the provirus landscape in people durably suppressed (*Antar et al., 2020*). Furthermore, a recent study reported that brief analytical treatment interruption (ATI) (median of 5 weeks) does not significantly impact the composition of virus populations (phylogenetic analyses sequencing of HIV-1 env) or size (total DNA, cell-associated RNA, IUPM) of the latent HIV-1 reservoir (*Salantes et al., 2018*). Collectively, these studies suggest a model in which circulating T cells, as well as virus escape variants in infected cells at the time of ART initiation, are maintained throughout durable viral suppression, and during brief periods of ATI.

Our observations contrast somewhat with those of Deng and colleagues who reported a predominance of CD8+ T cell-resistant viruses in proviral HIV-1 DNA (*Deng et al., 2015*). Deng et al did not empirically investigate proteome-wide patterns of escape (i.e. escape in 32% of reactive T cell epitopes) in the replication-competent HIV-1 reservoir, as reported here. Rather, their study focused on predicting the 'proportion' of escape within epitopes in Gag. They mostly used sequencing and bioinformatics approaches to conclude that, in epitopes in which virus variation was observed, in individuals treated in chronic infection, most (>98%) proviruses contained variants previously reported to confer T cell escape. Our empirical testing, which focused on T cell escape in the persistent HIV-1 reservoir, did not yield the same high proportion of escape variants reported in that study. Indeed, the average within-epitope escape frequency in 49 empirically tested epitopes was much lower at 27% (*Supplementary file 5*). This dataset included escape variants that can emerge early after infection such as the Gag T242N escape in Gag 240–249, restricted by B*58:01 and B57:01/3 (*Leslie et al., 2004*; *Brumme et al., 2009*). Deng et al also noted a higher within-epitope escape in the DNA reservoir in individuals treated in chronic versus acute infection. While only a few participants in this study were treated in acute infection, we observed no difference in within-epitope escape frequencies between participants treated in chronic or acute HIV-1 infection. Together this suggests that the differences in the studies did not arise from the regions of HIV-1 studied, or the differences in stage of HIV-1 infection at time of ART initiation. Rather, we propose that differences in within-epitope escape frequencies between studies derived from two sources. First, our empirical approach showed clearly that HIV-1 variants often do not escape T cell recognition and therefore

prediction methods may overestimate the level of escape. Indeed, Deng et al did include examples where virus variants predicted to confer escape did not decrease T cell responses following experimental testing. Second, the HIV template, HIV DNA versus OGV RNA, used for virus sequencing largely differed between studies.

Our observations that epitopes mostly harbor a mix of escape and non-escape variants are consistent with previous studies in untreated infection reporting that T cell escape is rarely all-or-nothing, but rather occurs at different rates (days to years) with multiple variants emerging over time (*Goonetilleke et al., 2009*; *Liu et al., 2013*; *Asquith et al., 2006*). These changes reflect both ongoing compensatory changes to maintain viral fitness and selection pressure from different T cell clonotypes (*Brockman et al., 2010*; *Ladell et al., 2013*; *Ritchie et al., 2014*), and progressive dysregulation of the immune response. While the HIV-1 reservoir is mostly formed at ART initiation, almost 30% of sequences enter the HIV-1 reservoir much earlier. The mixture of escape and non-escape OGVs observed in this study likely represent OGVs at different stages of T cell escape entering the reservoir at different stages of infection.

There are limitations of this study. First, detection of escape variants may have been limited by the sequencing depth obtained from OGVs from QVOAs. In addition, different stimuli can activate different viruses, and even the negative wells from the QVOA can generate new viruses after new stimulation, contributing to an underestimation of, or a bias in the replication-competent viruses present in the culture (*Ho et al., 2013*; *Hosmane et al., 2017*). Applying a rule of three calculation, our average depth of sequencing (25 sequences/participant) provided 95% confidence that any undetected variant comprised <12% of the population. While sampling biases were minimized by our choice of sequencing methods (nearly full-length and Primer ID), which have a low probability of introducing PCR mutations (*Zhou et al., 2015*; *Salazar-Gonzalez et al., 2008*), we cannot exclude that a low-frequency variants not detected by sequencing may have conferred escape (*Abrahams et al., 2019*). It is notable that that participant OGVs sequenced over time did not exhibit directional selection suggesting that the QVOA method recovers outgrowth virus that is representative of all variants (escaped and non-escaped) in the HIV-1 reservoir. Second, our bioinformatics analysis of HLA-associated polymorphisms in HIV-1 reservoir sequences simply identifies viral variants that are enriched among individuals expressing a particular HLA allele, as defined using population-level analyses (*Carlson et al., 2012*), and thus provides only a crude estimate of adaptation burden. This is because these lists of HLA-associated polymorphisms do not capture every escape variant that occurs in vivo; moreover, not every HLA-associated viral site defined at the population level is necessarily targeted by every individual expressing the relevant HLA allele. It is notable, however, that the results of the bioinformatics approach were consistent with our empirical studies, suggesting that the estimates made correctly captured the range of escape frequencies within this cohort. Third, in our empirical studies, we cannot exclude that the use of clade B consensus peptides underestimated the total breadth of the T cell response (*Goonetilleke et al., 2009*; *Altfeld et al., 2003*). However, our approach of synthesizing autologous peptides at reactive epitopes, and experimentally testing all 'stripes' (non-synonymous changes in >40% of sequences, see Materials and Methods) in our OGVs to identify additional epitopes to study, minimizes this limitation. Fourth, our empirical approach did not detect all mechanisms of escape such as possible immune dysfunction and processing mutations. To fully assess loss of CD8+ T cell responses due to immune dysfunction, longitudinal mapping and sequencing studies pre -and post-ART are needed. While no previously documented processing mutations were detected in our dataset, we anticipate the bioinformatics analysis which enumerated all HLA-associated polymorphisms regardless of their location with respect to known (or novel) T cell epitopes, will have captured some processing mutations. Lastly, we did not confirm the CD4 or CD8 restriction of all reactive and escaped T cell epitopes in this study. However, both consistent with previous studies (*McNeil et al., 2001*; *Appay et al., 2002*; *Betts et al., 2001*; *Riou et al., 2012*) and the optimal epitope mapping performed here, it is likely that the bulk of T cell responses identified in this study were CD8-restricted.

In summary, empirical testing supported by bioinformatics analysis, showed that 68% of circulating HIV-1 specific T cells across our cohort were able to recognize the replication-competent virus from the latent reservoir. Moreover, T cells in most individuals targeted one or more epitopes with no detectable escape and within epitopes, on average, only a minority of OGVs conferred escape. Treatment interruption studies have shown that circulating HIV-1 specific T cell responses are insufficient to fully suppress virus rebound in most individuals. Recent work has shown recombinant viruses

rapidly emerge following HIV rebound (*Vibholm et al., 2019*; *Lu et al., 2018*; *Cohen et al., 2018*). Our study, supported by animal models (*Pandrea et al., 2011*) and case reports investigating viral rebound (*Smith et al., 2015*), suggest that circulating HIV-1-specific T cells very likely afford some contribution in control of virus rebound. However, T cell immunity is rapidly overwhelmed, in part due to recombination-mediated escape (*Ritchie et al., 2014*; *Streeck et al., 2008*). Altogether, this suggests that HIV-1 cure strategies that include efforts to augment T cell immunity in PLWH on ART should include immunogen designs that target existing T cell epitopes, particularly those with low entropy which, as we show here, are less likely to harbor escape in the replication-competent reservoir. Ultimately, success of these therapies will likely be dependent on the ability to exert very rapid virus control and to limit the emergence of viral recombinants.

Materials and methods

Study cohort

Participants were enrolled through either the UNC Chapel Hill HIV-1 Clinical Trials Unit, the Women's Interagency HIV-1 Study (WIHS) UNC Chapel Hill Site, or the University of California San Francisco (UCSF) site. All participants were adults (>18 years of age) with documented HIV-1 infection. Infection was confirmed by either a positive HIV test by licensed ELISA test and confirmed by western Blot (WB), multispot HIV assay, HIV-1 RNA, or two documented plasma HIV-1 RNA >1000 copies/ml.

Enrollment criteria included a stable ART regimen for $\geq$12 months prior to consent/enrollment. Stable ART was defined as $\geq$nucleoside/nucleotide reverse transcriptase inhibitors plus a non-nucleoside reverse transcriptase inhibitor, integrase inhibitors, or a protease inhibitor without interruption (missing doses for no more than 2 consecutive days or no more than 4 cumulative days) in the 24 weeks immediately prior to consent/enrollment. Other potent fully suppressive ART combinations were allowed on a case-by-case basis, after clinician review. Prior changes in or elimination of medicines for easier dosing, intolerance, or other reasons were permitted if an alternative suppressive regimen was maintained. Second, plasma HIV-1 RNA was required to be <50 copies/mL for $\geq$12 months prior to enrolment as defined by a minimum of two tests (limit of detection determined by assay employed). Third, within 60 days of consent/enrollment, plasma HIV-1 RNA was required to be <50 copies/ml and hemoglobin required to be $\geq$12 g/L. Treatment during acute HIV-1 infection was defined as treatment within 30 days of either HIV-1 PCR positive but HIV-1 antigen negative, or a HIV-1 antigen discordant test.

Class I HLA typing was performed on 24/25 participants the UNC Center for AIDS Research Core Laboratory. Study characteristics of the cohort are summarized in *Table 1*.

Study approval/IRB approvals

Participants were enrolled with the following IRB approved studies: (1) *CID 0819 - Apheresis Procedures to Obtain Leukocytes for Research Studies from HIV Positive Participants (08–1575)*, (2) *The UNC Women's Interagency HIV Study (WIHS) (12-1660)*. Review and implementation of all protocols utilized for the collection of samples for this this analysis were approved by the University of North Carolina at Chapel Hill Biomedical Institutional Review Board (IRB) *and the University of California at San Francisco IRB*. All participants provided written informed consent. All experimental protocols were approved by local Institutional Biomedical Review Boards (ethics numbers: *14–0741, 11–0228, and 13–3613, 15–1626*) and performed in accordance with the relevant guidelines. Study characteristics of the cohorts are summarized in *Table 1*.

PBMC isolation

PBMC were isolated from apheresis product by centrifugation ($\times$1200 g for 15 min at room temperature) on a Ficoll-Paque density gradient (GE Healthcare Life Sciences Ficoll-Paque Plus). Briefly, the apheresis product was diluted 1:2.5 in 2% FBS/PBS and 30 ml diluted product was overlaid onto 15 ml of Ficoll-Plaque in a 50 ml Falcon tube prior to centrifugation. PBMCs were harvested, then washed three times in 2% FBS/PBS. PBMC were counted and then frozen (*Goonetilleke et al., 2006*).

Quantitative viral outgrowth assay (QVOA) and infectious units per million (IUPM)

QVOAs were performed as previously described (*Crooks et al., 2015*). Briefly, ~5 × 10^7 million resting CD4+ T cells, defined as CD4$^+$ CD45$^+$ CD3$^+$ CD69$^-$ CD25$^-$ CD8$^-$ CD14$^-$ HLA-DR$^-$, were isolated via negative selection from PBMC (*Keedy et al., 2009*). CD4+ T cells were maximally stimulated in limiting dilutions with phytohemagglutinin (PHA) (Remel-PHA; Thermo Scientific), IL-2, and irradiated CD8-depleted PBMC from a seronegative donor, which were added twice to cultures to support virus propagation. The CD8-depleted PBMC were obtained from seronegative donors previously screened to ensure adequate CCR5 expression. Culture supernatants were collected on days 15 and 19 and assayed for p24 expression by ELISA (ABL, Rockville, MD). Only cultures that contained an equivalent or greater level of p24 antigen on day 19 compared with day 15 were scored as positive. A maximum likelihood method was used on day 15 data to estimate the frequency of latent HIV infection, reported as bias-corrected Infectious Units Per Million (IUPM) (*Supplementary file 1*; *Trumble et al., 2017*). For 10 participants, longitudinal IUPM measurements were available. IUPM values measured in the 2 years prior to initial T cell mapping were averaged (range of IUPM values: 1–6, mean number of IUPM values/participant: 2). The limit of detection was defined as ≤1 p24 positive well (range 12–36 cultures, mean 18 cultures) at the highest input of cells (2.5 million cells).

Replication-competent reservoir HIV-1 virus sequencing

Supernatants from end point dilutions of p24 positive wells from the QVOA were sequenced for replication-competent virus. Twenty-two participants had supernatants sequenced by nearly full-length sequencing, and one participant had supernatant sequenced by Primer ID. For seven participants, longitudinal p24 supernatants were available (2–4 time points over 4.4 years). In two participants, no sequencing was possible as no p24 supernatants were identified in the QVOA; in these assays IUPM was calculated as below the threshold of detection.

Primer ID sequencing for target regions

A multiplexed Primer ID library that covered partial *gag*, integrase coding domain, V1/V2 region, partial gp41 and partial *nef* region was constructed using the previous protocol (*Zhou et al., 2015*; *Dennis et al., 2018*). In brief, viral RNA was extracted from culture supernatants using QIAmp viral mini kit (Qiagen). cDNA was synthesized using a cDNA primer mixture targeting the above regions with a block of random nucleotides in each cDNA primer serving as the Primer ID, and SuperScript III reverse transcriptase (ThermoFisher). After two rounds of purification of cDNA, we amplified the cDNA and attached Illumina adaptor sequences to the PCR products. Gel-purified libraries were pooled and sequenced using MiSeq 300 bp pair-end sequencing (Illumina). The TCS pipeline (https://github.com/SwanstromLab/PID) was used for de-mutiplexing of the sequenced regions and construct Primer ID Template Consensus Sequences (TCSs) for each sequenced viral template.

Nearly full-length sequencing

Viral RNA was converted to cDNA using Superscript III reverse transcriptase and an oligo(dT) primer. The 3' and 5' half genomes were amplified in separate PCR reactions using barcoded primers, and the PCR products were pooled, and gel purified. The SMARTbell Template Prep Kit (PacBio) was used to add adaptors to amplicons, and amplicons were then submitted for PacBio sequencing (movie time of 10 hr). Sequences were demultiplexed by barcode, and then analyzed using the PacBio Long Amplicon Analysis (LAA) package. Sequences were aligned separately using MUSCLE v3.8.31 and in-house scripts were used to examine the frequencies of mutations (*Edgar, 2004a*; *Edgar, 2004b*). Neighbor-joining phylogenetic trees were built using Ninja v1.2.2 (http://wheelerlab.org/software/ninja/), and approximate maximum likelihood trees were built using FastTree 2.1 (*Price et al., 2010*).

After aligning using MUSCLE and trimming of positions with indels, we performed two-by-two comparisons of the sequences at 5HG and 3HG and calculated the percent of comparisons with 0 to 3 nucleotide differences among the total comparisons. We chose one nucleotide difference as the clonality cut-off as we estimated that the QVOA and the sequencing assay might generate two substitutions per HIV genome. Using this cut-off, the percent of clonality was defined as the percentage

of two-by-two comparisons yielding 0 to 1 nucleotide differences among all the comparisons. The Ruby module viral_seq was used to calculate the pair-wise diversity of sequences (https://rubygems.org/gems/viral_seq).

PacBio sequencing proficiency was confirmed by the strong correlation between participant IUPM and the number of full-length sequences obtained, both of which are determined by the number of p24+ ELISA wells in the QVOA assay (r = 0.755, p<0.0001), Spearman Rank, 24 pairs, (*Figure 3—figure supplement 2*).

Droplet digital polymerase chain reaction (ddPCR)

The concentration of HIV-1 DNA copies in each sample was determined using the following primers and probes spanning HXB2 coordinates 684–810 (*Malnati et al., 2008*): LTRgagF: 5'-TCTCGACG-CAGGACTCG-3', LTRgagR: 5'- TACTGACGCTCTCGCACC- 3', and LTRgagProbe: 5'/56-FAM/CTC TCTCCT/ZEN/TCTAGCCTC/31ABkFQ/- 3'. Cell concentration was determined using the following RPP30 primers and probes: RPP30-F: 5'- GATTTGGACCTGCGAGCG-3', RPP30-R: 5'-GCGGCTGTC TCCACAAGT-3', RPP30-Probe: 5'-/56-FAM/CTGACCTGA/ZEN/AGGCTCT/31ABkFQ/- 3' (*Hindson et al., 2011*). LTRgag ddPCR reactions contained the following: 10 µl ddPCR Supermix for Probes (no dUTP) (Bio-Rad), 1 µl each primer (LTRgagF and LTRgagR) at a concentration of 20 µM, 0.35 µl of LTRgag probe at a concentration of 20 µM, 6.65 µl DEPC-treated water (Thermo Fisher Scientific), and 3 µl of undiluted DNA for a total volume of 22 µl. RPP30 ddPCR reactions contained the following: 10 µl ddPCR Supermix for Probes (no dUTP) (Bio-Rad), 1 µl each primer (RPP30-F and RPP30-R) at a concentration of 20 µM, 0.35 µl of RPP30 probe at a concentration of 20 µM, 8.65 µl DEPC-treated water (Thermo Fisher Scientific), and 1 µl of DNA diluted 1:100 in DEPC-treated water. 8E5 DNA was used as a positive control and no-template controls were included in each run. Samples were run in duplicate and the values averaged.

Droplets were generated using the QX200 Automated Droplet Generator (Bio-Rad), in a final volume of 40 µl. Following droplet generation, plates were sealed with a pierceable foil seal and immediately subjected to the following PCR conditions on a C1000 Touch Thermal Cycler with 96-Deep Well Reaction Module (Bio-Rad): 95°C for 10 min, then 45 cycles of 95°C for 30 s followed by 60°C for 1 min, a final extension of 98°C for 10 min and a 4°C hold. Following thermal cycling, plates were read on the QX200 Droplet Reader (Bio-Rad) with the following set-up: Experiment = Rare Event Detection, Mix = ddPCR Supermix for Probes (no dUTP), Target 1 = FAM, Target 2 = VIC.

Analysis of ddPCR data used QuantaSoft (Bio-Rad) with thresholds for the detection of each target set using the negative controls. The same threshold was used for all reactions of a particular target (either LTRgag or RPP30) in a given run. The concentration of HIV-1 copies / µl of sample was calculated as follows: Quantasoft LTRgag concentration ×(total volume of PCR reaction/volume of DNA added). The concentration of cells/µl of sample was calculated as follows: Quantasoft RPP30 concentration x (total volume of PCR reaction/volume of DNA added) x dilution factor x 0.5.

DNA extraction

Total DNA was extracted from purified resting CD4+ T cells using the DNeasy Blood and Tissue Kit (Qiagen), following the manufacturer's instructions with the following modifications. DNA was eluted from the column using 200 µl of DEPC-treated water (Thermo Fisher Scientific). Following a one-minute centrifugation at 6000 x *g*, the eluate was re-applied to the membrane, incubated for 1 min at room temperature, and re-eluted. DNA was stored at −20°C until use.

Amplification of *gag* and *nef*

A nested PCR approach was used to generate partial *gag* and *nef* amplicons. For amplification of partial *gag* amplicons, the first-round PCR primers were Gag_F_Outer_1: '5- GGACGGCGACTGG TGAGTACG −3' (HXB2 coordinates 732–752) and Gag_R_Outer: 5' – CCAATTCCCCCTATCA TTTTTGGTTTCC −3' (HXB2 coordinates 2404–2377). Each reaction contained 14.8 µl of DEPC-treated water, 2 µl of 10 × High Fidelity Buffer, 0.8 µl of 50 µM MgSO$_4$, 0.4 µl of 10 µM dNTP mix, 0.4 µl of forward and reverse primer at 20 µM and 5U of Platinum *Taq* DNA Polymerase High Fidelity (Thermo Fisher Scientific). One copy of HIV-1 DNA was added to each reaction contained, based on ddPCR estimates. A no template control was included in each run. First round thermal cycling conditions were as follows: an initial denaturation of 2 min at 98°C, followed by 35 cycles of: 15 s at 98°C

for denaturation, 30 s at 56°C for annealing, and 2 min 30 s at 72°C for extension, followed by a final extension at 72°C for 10 min and a 4°C hold. Second round PCR primers were as follows: Gag_F_Inner_2: 5' – CTTTTGACTAGCGGAGGCTAGAAGG –3' (HXB2 coordinates: 759–783) and Gag_R_Inner_3: 5'- GAGCTTCCCTTAGCTGACCCTCTAC –3' (HXB2 coordinates: 2319–2295). Second round reactions were set up as described above, and 1 µl of first-round PCR product. Second round thermal cycling conditions were same as those described above, however, with an annealing temperature of 55°C.

A nested PCR approach was also used to generate *nef* amplicons. First round PCR primers were: U5-623F: 5' -AAATCTCTAGCAGTGGCGCCCGAACAG –3' (HXB2 coordinates 623–634) and U5-601R: 5'-TGAGGGATCTCTAGTTACCAGAGTC-3' (HXB2 coordinates 9686–9662). Reaction components were identical to those described above. One copy of HIV-1 DNA was added to each reaction and a no-template control was included in each run. First round thermal cycling conditions were as follows: an initial denaturation of 2 min at 92°C, followed by 10 cycles of 10 s at 92°C for denaturation, 30 s at 60°C for annealing, 10 min at 68°C for extension, then 20 cycles of 10 s at 92°C for denaturation, 30 s at 55°C for annealing, 10 min at 68°C for extension, followed by a final extension of 10 min at 68°C, and a 4°C hold. The second-round primers were as follows: 4653F: 5'-CCCTACAATCCCCAAAGTCAAGGAG-3' and OFM19: 5'- GCACTCAAGGCAAGCTTTATTGAGGCTTA-3'. Each reaction consisted of 34.8 µl of DEPC-treated water (Thermo Fisher Scientific), 5 µl of 10 × High Fidelity Buffer, 2 µl of 50 µM MgSO$_4$, 5 µl of 0.2 µM forward and reverse primers and 5 U of Platinum *Taq* DNA Polymerase High Fidelity (Thermo Fisher Scientific), and 2 µl of first round PCR product. Second round thermal cycling conditions were as follows: 2 min at 94°C for the initial denaturation, then 10 cycles of 15 s at 94°C for denaturation, 30 s at 55°C for annealing, and 8 min at 68°C for extension, then 25 cycles of 15 s at 94°C for denaturation, 30 s at 55°C for annealing, and 8 min at 68°C for extension with an additional 20 s of extension time added each cycle; a final 7 min at 68°C for extension, followed by a hold at 4°C. Second round PCR products were screened for positives on a 0.8% agarose gel stained with SYBR Safe DNA Gel Stain (Thermo Fisher Scientific), and visualized with a UV gel imager. Amplicons were sequenced using Sanger sequencing.

HLA-associated polymorphisms methods

HLA-associated polymorphisms have previously been identified across the HIV-1 subtype B proteome via statistical association (*Carlson et al., 2012*). Briefly, HLA-*adapted* forms denote viral residues that are enriched among individuals expressing the restricting HLA (i.e. the inferred escape form), whereas HLA-*nonadapted* forms denote viral residues that are underrepresented among individuals expressing the restricting HLA (i.e. the inferred susceptible form) (*Carlson et al., 2012*). For each HIV-1 reservoir sequence, we identified all HIV codons under selection by one or more host HLA alleles (as defined in *Carlson et al., 2012*) and classified the autologous HIV residue as 'nonadapted', 'adapted' or 'possibly adapted' (where the latter extended to any variant other than the 'nonadapted' form). For each participant, a minimum bound of HLA-associated adaptation in the reservoir was estimated by computing the total proportion of HLA-associated sites exhibiting the 'adapted' form, while a maximum escape bound was estimated by including both 'adapted' and 'possibly adapted' forms. HLA-associated polymorphisms were analyzed at the HLA subtype (4-digit) level, where ambiguous allele calls were assigned to the most common allele in the list. Note that, for this analysis, Pol residue numbering begins at position –1 of protease.

Peptides

18-mer peptides overlapping by 10 amino acids were synthesized (Sigma-Genosys, US) to match the clade B HIV-1 consensus sequence (https://www.hiv.lanl.gov/). The corresponding 18-mer autologous outgrowth reservoir virus peptide/s for all previously defined reactive 18-mers from the initial mapping ELISpot were also synthesized. In addition, optimal CD8+ T cell peptides predicted through LANL's epitope location finder ('ELF') software based on HLA-prediction software (https://www.hiv.lanl.gov/content/sequence/ELF/epitope_analyzer.html), and 18-mer peptides spanning stripes (see definition below) were synthesized. To cover variation in autologous OGV, a total of 656 peptides were synthesized.

T cell mapping against HIV-1 clade B consensus sequence

Clade B consensus peptides (n = 386) were arranged into pools in a matrix format using the Peptide Portal program (SCHARP), adapting code from the Deconvolute this! Program (*Precopio et al., 2008*). Each peptide was repeated $3 \times$ in the matrix peptide plate. Cryopreserved PBMCs were thawed and rested overnight before being placed in the ELISpot plates (Merck, Millipore) at 4×10^5 cells per well. Peptides were added to the plate with a final concentration of 4 µg/ml and incubated for 18–20 hr at 37°C, 5% CO_2. Coating, development (MabTech), and reading of ELISpot plates (AID Reader) has been described previously (*Goonetilleke et al., 2006*). Data from positive pools were then deconvoluted (*Roederer and Koup, 2003*) to identify candidate epitope peptides. Putative positive 18-mer peptide–specific T cell responses were confirmed in triplicate in a follow-up IFN-γ ELISpot with either 2- or 4×10^5 cells per well and a peptide concentration of 10 µg/ml, with 4 to 6 negative control wells (media only) and at least one positive control well (10 µg/ml PHA; Sigma-Aldrich) together. Positive T cell responses were defined as ≥12.5 SFU per million, ≥4 × background and no zero values in any replicate of antigen-stimulated wells.

Evaluating T cell escape in epitopes previously defined as reactive in clade B consensus mapping

Testing OGV variants for T cell escape (replicates, controls, positivity criteria) was the same as for the initial mapping studies with the exception that peptides were tested at 10 µM. Escape variants were defined as eliciting a 50% decrease in the mean magnitude compared to the peptide that elicited the highest magnitude T cell response in IFN-γ ELISpot. Optimally defined CD8+ epitopes (8–11 aa) were defined as equal to, or greater than the T cell magnitude elicited by the 18-mer peptide (*Carlson et al., 2012*).

Examining virus stripes and processing mutations

We observed clustered mutations in OGV sequence data outside of mapped T cell epitopes. We hypothesized that these regions may reflect pre-ART escape against T cell responses not detected using our matrix mapping approach. Regions where > 40% of the sequences contain non-synonymous mutations in the same position in the proteome were defined as 'stripes'. Twenty-four stripes were identified across participants. 18-mer peptides spanning stripes were synthesized and tested individually in ELISpot. As forecast, some T cell responses were detected. In total T cells recognized epitopes at 11/24 stripes, resulting in a total of 166 reactive T cell epitopes in 25 participants. Prediction software and HLA typing was used to screen virus sequences for the presence of antigen processing mutations (https://www.hiv.lanl.gov/content/immunology/variants/variant_search.html?db=ctl). No antigen processing mutations associated with reactive T cell epitopes were detected in the cohort.

Entropy calculations

A Shannon entropy score (*Shannon, 1948*) was calculated for all reactive epitopes in this study, 18-mer and all experimentally confirmed optimal epitopes (*Supplementary file 4*). For each reactive epitope, the LANL sequence database was used to identify submitted Clade B sequences. Between 701 and 1936 sequences were analyzed per epitope. The Shannon Entropy-Two tool from the LANL database was used to identify the variation in protein sequence alignments (reactive epitope sequence vs. all known sequences), and an entropy score was calculated by averaging the entropy score for each amino acid position (https://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy.html).

Statistics

All statistical analyses were performed using Graph Pad Prism Version 8. Data were assumed to be non-normally distributed and as appropriate, two-tailed tests applied. A p value of ≤ 0.01 was considered significant.

Acknowledgements

Additional information

Competing interests

Ronald Swanstrom: UNC is pursuing IP protection for Primer ID, and Ronald Swanstrom is listed as a coinventor and has received nominal royalties. The other authors declare that no competing interests exist.

Funding

Author contributions

Joanna A Warren, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review and editing; Shuntai Zhou, Formal analysis, Investigation, Methodology, Writing - original

draft; Yinyan Xu, Matthew J Moeser, Investigation; Daniel R MacMillan, Formal analysis; Olivia Council, Sarah Joseph, Data curation; Jennifer Kirchherr, Data curation, Investigation; Julia M Sung, Resources, Investigation, Writing - review and editing; Nadia R Roan, Resources, Writing - review and editing; Adaora A Adimora, Cynthia L Gay, Resources, Data curation; JoAnn D Kuruc, Resources, Data curation, Funding acquisition; David M Margolis, Resources, Formal analysis, Supervision, Funding acquisition, Methodology, Writing - review and editing; Nancie Archin, Resources, Data curation, Supervision, Writing - review and editing; Zabrina L Brumme, Conceptualization, Resources, Formal analysis, Supervision, Funding acquisition, Methodology, Writing - original draft, Writing - review and editing; Ronald Swanstrom, Data curation, Supervision, Writing - review and editing; Nilu Goonetilleke, Conceptualization, Resources, Data curation, Supervision, Funding acquisition, Writing - original draft, Writing - review and editing

Author ORCIDs
Joanna A Warren (iD) http://orcid.org/0000-0003-0595-0390
Matthew J Moeser (iD) http://orcid.org/0000-0002-1368-5171
Nadia R Roan (iD) https://orcid.org/0000-0002-5464-1976
Zabrina L Brumme (iD) http://orcid.org/0000-0002-8157-1037
Nilu Goonetilleke (iD) https://orcid.org/0000-0003-2278-1656

Ethics

Human subjects: Participants were enrolled with the following IRB approved studies: (1) CID 0819 - Apheresis Procedures to Obtain Leukocytes for Research Studies from HIV Positive Participants (08-1575), (2) The UNC Women's Interagency HIV Study (WIHS) (12-1660). Review and implementation of all protocols utilized for the collection of samples for this this analysis were approved by the University of North Carolina at Chapel Hill Biomedical Institutional Review Board (IRB) and the University of California at San Francisco IRB. All participants provided written informed consent. All experimental protocols were approved by local Institutional Biomedical Review Boards (ethics numbers: 14-0741, 11-0228, and 13-3613, 15-1626) and performed in accordance with the relevant guidelines.

Decision letter and Author response
Decision letter https://doi.org/10.7554/eLife.57246.sa1
Author response https://doi.org/10.7554/eLife.57246.sa2

Additional files

Supplementary files
• Supplementary file 1. Size of the replication-competent reservoir in PLWH on ART (n = 25 participants) as measured by infectious units per million (IUPM) up to 2 years prior to T cell mapping date.

• Supplementary file 2. The number of near full-length sequences obtained from p24 antigen positive wells generated in quantitative viral outgrowth assays for each participant, and the pairwise diversity of sequences obtained (5' and 3' halves).

• Supplementary file 3. Percentage of clonal sequences obtained from p24 antigen positive wells generated in quantitative viral outgrowth assays for each participant (5' and 3' halves).

• Supplementary file 4. Summary of HLA-associated polymorphism by HIV-1 protein.

• Supplementary file 5. Summary of reactive T cell epitopes in study participants.

• Supplementary file 6. Raw data from studies evaluating viral T cell escape variants in the replication-competent reservoir.

• Supplementary file 7. HIV-1-specific T cell measurements and HIV reservoir size.

• Supplementary file 8. HIV-1-specific T cell responses do not correlated with the size of the HIV reservoir. Correlation between T cell breadth and summed magnitude of T cell response to HIV-1 protein in PLWH on ART (n = 25 participants, n = 166 epitopes), and adjusted for escape variants (n = 23 participants, n = 102 epitopes excluding 49 epitopes at which escape was observed),

measured by IFN-γ ELISpot and the size of the replication-competent reservoir as measured by infectious units per million (IUPM) using Spearman Rank.

• Supplementary file 9. Reactive T cell epitopes in PLWH on ART (n = 23 participants, 151 *total* mapped epitopes including 49 *escape* epitopes) are targeted by conserved immunogen vaccines. Escape in the HIV-1 reservoir is consistently lower in T cell epitopes that fall within conserved immunogen vaccines than mapped epitopes that fall outside of the immunogens.

• Supplementary file 10. Comparison of approaches to assess pre-ART escape in the latent HIV-1 reservoir.

• Transparent reporting form

Data availability

Sequencing data have been deposited in Gen Bank under PRJNA666896, MT307344-MT308415 and MW054719-MW054856 All data generated (raw) or analyzed during this study are included in the manuscript and supporting files (supplementary files).

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Warren JA, Zhou S, Xu Y, Moeser M, MacMillan DR, Council O, Kirchherr J, Sung JM, Roan N, Adimora AA, Joseph S, Kuruc JD, Gay CL, Margolis DM, Archin NM, Brumme ZL, Swanstrom R, Goonetilleke N	2020	Primer ID sequencing for QVOA HIV sequence	https://www.ncbi.nlm.nih.gov/bioproject/666896	NCBI BioProject, PRJNA666896

References

Abrahams MR, Anderson JA, Giorgi EE, Seoighe C, Mlisana K, Ping LH, Athreya GS, Treurnicht FK, Keele BF, Wood N, Salazar-Gonzalez JF, Bhattacharya T, Chu H, Hoffman I, Galvin S, Mapanje C, Kazembe P, Thebus R, Fiscus S, Hide W, et al. 2009. Quantitating the multiplicity of infection with human immunodeficiency virus type 1 subtype C reveals a non-poisson distribution of transmitted variants. *Journal of Virology* **83**:3556–3567. DOI: https://doi.org/10.1128/JVI.02132-08, PMID: 19193811

Abrahams M-R, Joseph SB, Garrett N, Tyers L, Moeser M, Archin N, Council OD, Matten D, Zhou S, Doolabh D, Anthony C, Goonetilleke N, Karim SA, Margolis DM, Pond SK, Williamson C, Swanstrom R. 2019. The replication-competent HIV-1 latent reservoir is primarily established near the time of therapy initiation. *Science Translational Medicine* **11**:eaaw5589. DOI: https://doi.org/10.1126/scitranslmed.aaw5589

Achenbach CJ, Assoumou L, Deeks SG, Wilkin TJ, Berzins B, Casazza JP, Lambert-Niclot S, Koup RA, Costagliola D, Calvez V, Katlama C, Autran B, Murphy RL, EraMune 02 study team. 2015. Effect of therapeutic intensification followed by HIV DNA prime and rAd5 boost vaccination on HIV-specific immunity and HIV reservoir (EraMune 02): a multicentre randomised clinical trial. *The Lancet HIV* **2**:e82–e91. DOI: https://doi.org/10.1016/S2352-3018(15)00026-0, PMID: 26424549

Addo MM, Yu XG, Rathod A, Cohen D, Eldridge RL, Strick D, Johnston MN, Corcoran C, Wurcel AG, Fitzpatrick CA, Feeney ME, Rodriguez WR, Basgoz N, Draenert R, Stone DR, Brander C, Goulder PJ, Rosenberg ES, Altfeld M, Walker BD. 2003. Comprehensive epitope analysis of human immunodeficiency virus type 1 (HIV-1)-specific T-cell responses directed against the entire expressed HIV-1 genome demonstrate broadly directed responses, but no correlation to viral load. *Journal of Virology* **77**:2081–2092. DOI: https://doi.org/10.1128/JVI.77.3.2081-2092.2003, PMID: 12525643

Allen TM, Altfeld M, Geer SC, Kalife ET, Moore C, O'sullivan KM, Desouza I, Feeney ME, Eldridge RL, Maier EL, Kaufmann DE, Lahaie MP, Reyor L, Tanzi G, Johnston MN, Brander C, Draenert R, Rockstroh JK, Jessen H, Rosenberg ES, et al. 2005. Selective escape from CD8+ T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution. *Journal of Virology* **79**:13239–13249. DOI: https://doi.org/10.1128/JVI.79.21.13239-13249.2005, PMID: 16227247

Altfeld M, Addo MM, Shankarappa R, Lee PK, Allen TM, Yu XG, Rathod A, Harlow J, O'Sullivan K, Johnston MN, Goulder PJ, Mullins JI, Rosenberg ES, Brander C, Korber B, Walker BD. 2003. Enhanced detection of human immunodeficiency virus type 1-specific T-cell responses to highly variable regions by using peptides based on

autologous virus sequences. *Journal of Virology* **77**:7330–7340. DOI: https://doi.org/10.1128/JVI.77.13.7330-7340.2003, PMID: 12805431

Antar AA, Jenike KM, Jang S, Rigau DN, Reeves DB, Hoh R, Krone MR, Keruly JC, Moore RD, Schiffer JT, Nonyane BA, Hecht FM, Deeks SG, Siliciano JD, Ho YC, Siliciano RF. 2020. Longitudinal study reveals HIV-1-infected CD4+ T cell dynamics during long-term antiretroviral therapy. *Journal of Clinical Investigation* **130**: 3543–3559. DOI: https://doi.org/10.1172/JCI135953, PMID: 32191639

Appay V, Papagno L, Spina CA, Hansasuta P, King A, Jones L, Ogg GS, Little S, McMichael AJ, Richman DD, Rowland-Jones SL. 2002. Dynamics of T cell responses in HIV infection. *The Journal of Immunology* **168**:3660–3666. DOI: https://doi.org/10.4049/jimmunol.168.7.3660

Archin NM, Liberty AL, Kashuba AD, Choudhary SK, Kuruc JD, Crooks AM, Parker DC, Anderson EM, Kearney MF, Strain MC, Richman DD, Hudgens MG, Bosch RJ, Coffin JM, Eron JJ, Hazuda DJ, Margolis DM. 2012. Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. *Nature* **487**:482–485. DOI: https://doi.org/10.1038/nature11286, PMID: 22837004

Asquith B, Edwards CT, Lipsitch M, McLean AR. 2006. Inefficient cytotoxic T lymphocyte-mediated killing of HIV-1-infected cells in vivo. *PLOS Biology* **4**:e90. DOI: https://doi.org/10.1371/journal.pbio.0040090, PMID: 16515366

Bansal A, Gough E, Sabbaj S, Ritter D, Yusim K, Sfakianos G, Aldrovandi G, Kaslow RA, Wilson CM, Mulligan MJ, Kilby JM, Goepfert PA. 2005. CD8 T-cell responses in early HIV-1 infection are skewed towards high entropy peptides. *Aids* **19**:241–250. PMID: 15718834

Bar KJ, Sneller MC, Harrison LJ, Justement JS, Overton ET, Petrone ME, Salantes DB, Seamon CA, Scheinfeld B, Kwan RW, Learn GH, Proschan MA, Kreider EF, Blazkova J, Bardsley M, Refsland EW, Messer M, Clarridge KE, Tustin NB, Madden PJ, et al. 2016. Effect of HIV antibody VRC01 on viral rebound after treatment interruption. *New England Journal of Medicine* **375**:2037–2050. DOI: https://doi.org/10.1056/NEJMoa1608243, PMID: 27959728

Barouch DH, Kunstman J, Kuroda MJ, Schmitz JE, Santra S, Peyerl FW, Krivulka GR, Beaudry K, Lifton MA, Gorgone DA, Montefiori DC, Lewis MG, Wolinsky SM, Letvin NL. 2002. Eventual AIDS vaccine failure in a rhesus monkey by viral escape from cytotoxic T lymphocytes. *Nature* **415**:335–339. DOI: https://doi.org/10.1038/415335a, PMID: 11797012

Barouch DH, Korber B. 2010. HIV-1 vaccine development after STEP. *Annual Review of Medicine* **61**:153–167. DOI: https://doi.org/10.1146/annurev.med.042508.093728, PMID: 20059334

Barton JP, Goonetilleke N, Butler TC, Walker BD, McMichael AJ, Chakraborty AK. 2016. Relative rate and location of intra-host HIV evolution to evade cellular immunity are predictable. *Nature Communications* **7**:1–10. DOI: https://doi.org/10.1038/ncomms11660

Betts MR, Ambrozak DR, Douek DC, Bonhoeffer S, Brenchley JM, Casazza JP, Koup RA, Picker LJ. 2001. Analysis of total human immunodeficiency virus (HIV)-specific CD4(+) and CD8(+) T-cell responses: relationship to viral load in untreated HIV infection. *Journal of Virology* **75**:11983–11991. DOI: https://doi.org/10.1128/JVI.75.24.11983-11991.2001, PMID: 11711588

Borducchi EN, Cabral C, Stephenson KE, Liu J, Abbink P, Ng'ang'a D, Nkolola JP, Brinkman AL, Peter L, Lee BC, Jimenez J, Jetton D, Mondesir J, Mojta S, Chandrashekar A, Molloy K, Alter G, Gerold JM, Hill AL, Lewis MG, et al. 2016. Ad26/MVA therapeutic vaccination with TLR7 stimulation in SIV-infected rhesus monkeys. *Nature* **540**:284–287. DOI: https://doi.org/10.1038/nature20583, PMID: 27841870

Borrow P, Lewicki H, Wei X, Horwitz MS, Peffer N, Meyers H, Nelson JA, Gairin JE, Hahn BH, Oldstone MB, Shaw GM. 1997. Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. *Nature Medicine* **3**:205–211. DOI: https://doi.org/10.1038/nm0297-205, PMID: 9018240

Borthwick N, Lin Z, Akahoshi T, Llano A, Silva-Arrieta S, Ahmed T, Dorrell L, Brander C, Murakoshi H, Takiguchi M, Hanke T. 2017. Novel, in-natural-infection subdominant HIV-1 CD8+ T-cell epitopes revealed in human recipients of conserved-region T-cell vaccines. *PLOS ONE* **12**:e0176418. DOI: https://doi.org/10.1371/journal.pone.0176418, PMID: 28448594

Brockman MA, Brumme ZL, Brumme CJ, Miura T, Sela J, Rosato PC, Kadie CM, Carlson JM, Markle TJ, Streeck H, Kelleher AD, Markowitz M, Jessen H, Rosenberg E, Altfeld M, Harrigan PR, Heckerman D, Walker BD, Allen TM. 2010. Early selection in gag by protective HLA alleles contributes to reduced HIV-1 replication capacity that may be largely compensated for in chronic infection. *Journal of Virology* **84**:11937–11949. DOI: https://doi.org/10.1128/JVI.01086-10, PMID: 20810731

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. DOI: https://doi.org/10.7554/eLife.18889, PMID: 27855060

Brown AJ, Richman DD. 1997. HIV-1: gambling on the evolution of drug resistance? *Nature Medicine* **3**:268–271. DOI: https://doi.org/10.1038/nm0397-268, PMID: 9055846

Brumme ZL, John M, Carlson JM, Brumme CJ, Chan D, Brockman MA, Swenson LC, Tao I, Szeto S, Rosato P, Sela J, Kadie CM, Frahm N, Brander C, Haas DW, Riddler SA, Haubrich R, Walker BD, Harrigan PR, Heckerman D, et al. 2009. HLA-associated immune escape pathways in HIV-1 subtype B gag, pol and nef proteins. *PLOS ONE* **4**:e6687. DOI: https://doi.org/10.1371/journal.pone.0006687, PMID: 19690614

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 Pathogens* **13**:e1006283. DOI: https://doi.org/10.1371/journal.ppat.1006283, PMID: 28328934

Buzon MJ, Martin-Gayo E, Pereyra F, Ouyang Z, Sun H, Li JZ, Piovoso M, Shaw A, Dalmau J, Zangger N, Martinez-Picado J, Zurakowski R, Yu XG, Telenti A, Walker BD, Rosenberg ES, Lichterfeld M. 2014. Long-term antiretroviral treatment initiated at primary HIV-1 infection affects the size, composition, and decay kinetics of the reservoir of HIV-1-infected CD4 T cells. *Journal of Virology* **88**:10056–10065. DOI: https://doi.org/10.1128/JVI.01046-14, PMID: 24965451

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, et al. 2012. Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. *Journal of Virology* **86**:13202–13216. DOI: https://doi.org/10.1128/JVI.01998-12, PMID: 23055555

Carlson JM, Schaefer M, Monaco DC, Batorsky R, Claiborne DT, Prince J, Deymier MJ, Ende ZS, Klatt NR, DeZiel CE, Lin TH, Peng J, Seese AM, Shapiro R, Frater J, Ndung'u T, Tang J, Goepfert P, Gilmour J, Price MA, et al. 2014. HIV transmission. selection Bias at the heterosexual HIV-1 transmission bottleneck. *Science* **345**:1254031. DOI: https://doi.org/10.1126/science.1254031, PMID: 25013080

Chun TW, Stuyver L, Mizell SB, Ehler LA, Mican JA, Baseler M, Lloyd AL, Nowak MA, Fauci AS. 1997. Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *PNAS* **94**:13193–13197. DOI: https://doi.org/10.1073/pnas.94.24.13193, PMID: 9371822

Cohen YZ, Lorenzi JCC, Krassnig L, Barton JP, Burke L, Pai J, Lu CL, Mendoza P, Oliveira TY, Sleckman C, Millard K, Butler AL, Dizon JP, Belblidia SA, Witmer-Pack M, Shimeliovich I, Gulick RM, Seaman MS, Jankovic M, Caskey M, et al. 2018. Relationship between latent and rebound viruses in a clinical trial of anti-HIV-1 antibody 3bnc117. *Journal of Experimental Medicine* **215**:2311–2324. DOI: https://doi.org/10.1084/jem.20180936, PMID: 30072495

Crooks AM, Bateson R, Cope AB, Dahl NP, Griggs MK, Kuruc JD, Gay CL, Eron JJ, Margolis DM, Bosch RJ, Archin NM. 2015. Precise quantitation of the latent HIV-1 reservoir: implications for eradication strategies. *Journal of Infectious Diseases* **212**:1361–1365. DOI: https://doi.org/10.1093/infdis/jiv218, PMID: 25877550

Crowell TA, Colby DJ, Pinyakorn S, Sacdalan C, Pagliuzza A, Intasan J, Benjapornpong K, Tangnaree K, Chomchey N, Kroon E, de Souza MS, Tovanabutra S, Rolland M, Eller MA, Paquin-Proulx D, Bolton DL, Tokarev A, Thomas R, Takata H, Trautmann L, et al. 2019. Safety and efficacy of VRC01 broadly neutralising antibodies in adults with acutely treated HIV (RV397): a phase 2, randomised, double-blind, placebo-controlled trial. *The Lancet HIV* **6**:e297–e306. DOI: https://doi.org/10.1016/S2352-3018(19)30053-0, PMID: 31000477

Davey RT, Bhat N, Yoder C, Chun TW, Metcalf JA, Dewar R, Natarajan V, Lempicki RA, Adelsberger JW, Miller KD, Kovacs JA, Polis MA, Walker RE, Falloon J, Masur H, Gee D, Baseler M, Dimitrov DS, Fauci AS, Lane HC. 1999. HIV-1 and T cell dynamics after interruption of highly active antiretroviral therapy (HAART) in patients with a history of sustained viral suppression. *PNAS* **96**:15109–15114. DOI: https://doi.org/10.1073/pnas.96.26.15109, PMID: 10611346

Deng K, Pertea 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, et al. 2015. Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations. *Nature* **517**:381–385. DOI: https://doi.org/10.1038/nature14053, PMID: 25561180

Dennis AM, Zhou S, Sellers CJ, Learner E, Potempa M, Cohen MS, Miller WC, Eron JJ, Swanstrom R. 2018. Using Primer-ID deep sequencing to detect recent human immunodeficiency virus type 1 infection. *The Journal of Infectious Diseases* **218**:1777–1782. DOI: https://doi.org/10.1093/infdis/jiy426, PMID: 30010965

Derdeyn CA, Decker JM, Bibollet-Ruche F, Mokili JL, Muldoon M, Denham SA, Heil ML, Kasolo F, Musonda R, Hahn BH, Shaw GM, Korber BT, Allen S, Hunter E. 2004. Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. *Science* **303**:2019–2022. DOI: https://doi.org/10.1126/science.1093137, PMID: 15044802

Edgar RC. 2004a. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**:1792–1797. DOI: https://doi.org/10.1093/nar/gkh340, PMID: 15034147

Edgar RC. 2004b. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**:113. DOI: https://doi.org/10.1186/1471-2105-5-113, PMID: 15318951

Eisele E, Siliciano RF. 2012. Redefining the viral reservoirs that prevent HIV-1 eradication. *Immunity* **37**:377–388. DOI: https://doi.org/10.1016/j.immuni.2012.08.010, PMID: 22999944

Elliott J, Solomon A, Wightman F, Smith M, Palmer S, Prince M, Watson J, Hoy J, McMahon M, Lewin S. 2013. The safety and effect of multiple doses of vorinostat on HIV transcription in HIV-Infected patients receiving combination antiretroviral therapy. Conference On Retroviruses and Opportunistic Infections.

Elliott JH, McMahon JH, Chang CC, Lee SA, Hartogensis W, Bumpus N, Savic R, Roney J, Hoh R, Solomon A, Piatak M, Gorelick RJ, Lifson J, Bacchetti P, Deeks SG, Lewin SR. 2015. Short-term administration of disulfiram for reversal of latent HIV infection: a phase 2 dose-escalation study. *The Lancet HIV* **2**:e520–e529. DOI: https://doi.org/10.1016/S2352-3018(15)00226-X, PMID: 26614966

Emery A, Zhou S, Pollom E, Swanstrom R. 2017. Characterizing HIV-1 splicing by using Next-Generation sequencing. *Journal of Virology* **91**:16. DOI: https://doi.org/10.1128/JVI.02515-16

Ferrari G, Korber B, Goonetilleke N, Liu MK, Turnbull EL, Salazar-Gonzalez JF, Hawkins N, Self S, Watson S, Betts MR, Gay C, McGhee K, Pellegrino P, Williams I, Tomaras GD, Haynes BF, Gray CM, Borrow P, Roederer M, McMichael AJ, et al. 2011. Relationship between functional profile of HIV-1 specific CD8 T cells and epitope variability with the selection of escape mutants in acute HIV-1 infection. *PLOS Pathogens* **7**:e1001273. DOI: https://doi.org/10.1371/journal.ppat.1001273, PMID: 21347345

Finzi D. 1997. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science* **278**:1295–1300. DOI: https://doi.org/10.1126/science.278.5341.1295

Ganusov VV, Goonetilleke N, Liu MK, Ferrari G, Shaw GM, McMichael AJ, Borrow P, Korber BT, Perelson AS. 2011. Fitness costs and diversity of the cytotoxic T lymphocyte (CTL) response determine the rate of CTL escape during acute and chronic phases of HIV infection. *Journal of Virology* **85**:10518–10528. DOI: https://doi.org/10.1128/JVI.00655-11, PMID: 21835793

Goonetilleke N, Moore S, Dally L, Winstone N, Cebere I, Mahmoud A, Pinheiro S, Gillespie G, Brown D, Loach V, Roberts J, Guimaraes-Walker A, Hayes P, Loughran K, Smith C, De Bont J, Verlinde C, Vooijs D, Schmidt C, Boaz M, et al. 2006. Induction of multifunctional human immunodeficiency virus type 1 (HIV-1)-specific T cells capable of proliferation in healthy subjects by using a prime-boost regimen of DNA- and modified vaccinia virus Ankara-vectored vaccines expressing HIV-1 gag coupled to CD8+ T-cell epitopes. *Journal of Virology* **80**: 4717–4728. DOI: https://doi.org/10.1128/JVI.80.10.4717-4728.2006, PMID: 16641265

Goonetilleke N, Liu MK, Salazar-Gonzalez JF, Ferrari G, Giorgi E, Ganusov VV, Keele BF, Learn GH, Turnbull EL, Salazar MG, Weinhold KJ, Moore S, Letvin N, Haynes BF, Cohen MS, Hraber P, Bhattacharya T, Borrow P, Perelson AS, Hahn BH, et al. 2009. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. *Journal of Experimental Medicine* **206**:1253–1272. DOI: https://doi.org/10.1084/jem.20090365, PMID: 19487423

Gorse GJ, Baden LR, Wecker M, Newman MJ, Ferrari G, Weinhold KJ, Livingston BD, Villafana TL, Li H, Noonan E, Russell ND, HIV Vaccine Trials Network. 2008. Safety and immunogenicity of cytotoxic T-lymphocyte poly-epitope, DNA plasmid (EP HIV-1090) vaccine in healthy, human immunodeficiency virus type 1 (HIV-1)-uninfected adults. *Vaccine* **26**:215–223. DOI: https://doi.org/10.1016/j.vaccine.2007.10.061, PMID: 18055072

Goulder PJ, Phillips RE, Colbert RA, McAdam S, Ogg G, Nowak MA, Giangrande P, Luzzi G, Morgan B, Edwards A, McMichael AJ, Rowland-Jones S. 1997. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nature Medicine* **3**:212–217. DOI: https://doi.org/10.1038/nm0297-212, PMID: 9018241

Haaland RE, Hawkins PA, Salazar-Gonzalez J, Johnson A, Tichacek A, Karita E, Manigart O, Mulenga J, Keele BF, Shaw GM, Hahn BH, Allen SA, Derdeyn CA, Hunter E. 2009. Inflammatory genital infections mitigate a severe genetic bottleneck in heterosexual transmission of subtype A and C HIV-1. *PLOS Pathogens* **5**:e1000274. DOI: https://doi.org/10.1371/journal.ppat.1000274, PMID: 19165325

Hanke T. 2014. Conserved immunogens in prime-boost strategies for the next-generation HIV-1 vaccines. *Expert Opinion on Biological Therapy* **14**:601–616. DOI: https://doi.org/10.1517/14712598.2014.885946, PMID: 24490585

Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, Kitano TK, Hodel MR, Petersen JF, Wyatt PW, Steenblock ER, Shah PH, Bousse LJ, Troup CB, Mellen JC, Wittmann DK, et al. 2011. High-Throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Analytical Chemistry* **83**:8604–8610. DOI: https://doi.org/10.1021/ac202028g

Ho YC, Shan L, Hosmane NN, Wang J, Laskey SB, Rosenbloom DI, Lai J, Blankson JN, Siliciano JD, Siliciano RF. 2013. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell* **155**:540–551. DOI: https://doi.org/10.1016/j.cell.2013.09.020, PMID: 24243014

Hosmane NN, Kwon KJ, Bruner KM, Capoferri AA, Beg S, Rosenbloom DIS, Keele BF, Y-c H, Siliciano JD, Siliciano RF. 2017. Proliferation of latently infected CD4+ T cells carrying replication-competent HIV-1. *Potential Role in Latent Reservoir Dynamics* **214**:959–972. DOI: https://doi.org/10.1084/jem.20170193

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 Pathogens* **10**:e1004010. DOI: https://doi.org/10.1371/journal.ppat.1004010, PMID: 24651464

Kearney MF, Wiegand A, Shao W, Coffin JM, Mellors JW, Lederman M, Gandhi RT, Keele BF, Li JZ. 2016. Origin of rebound plasma HIV includes cells with identical proviruses that are transcriptionally active before stopping of antiretroviral therapy. *Journal of Virology* **90**:1369–1376. DOI: https://doi.org/10.1128/JVI.02139-15, PMID: 26581989

Keedy KS, Archin NM, Gates AT, Espeseth A, Hazuda DJ, Margolis DM. 2009. A limited group of class I histone deacetylases acts to repress human immunodeficiency virus type 1 expression. *Journal of Virology* **83**:4749–4756. DOI: https://doi.org/10.1128/JVI.02585-08, PMID: 19279091

Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, Sun C, Grayson T, Wang S, Li H, Wei X, Jiang C, Kirchherr JL, Gao F, Anderson JA, Ping LH, Swanstrom R, Tomaras GD, Blattner WA, Goepfert PA, et al. 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *PNAS* **105**:7552–7557. DOI: https://doi.org/10.1073/pnas.0802203105, PMID: 18490657

Koup RA, Safrit JT, Cao Y, Andrews CA, McLeod G, Borkowsky W, Farthing C, Ho DD. 1994. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *Journal of Virology* **68**:4650–4655. DOI: https://doi.org/10.1128/JVI.68.7.4650-4655.1994, PMID: 8207839

Kulkarni V, Rosati M, Valentin A, Ganneru B, Singh AK, Yan J, Rolland M, Alicea C, Beach RK, Zhang GM, Le Gall S, Broderick KE, Sardesai NY, Heckerman D, Mothe B, Brander C, Weiner DB, Mullins JI, Pavlakis GN, Felber BK. 2013. HIV-1 p24(gag) derived conserved element DNA vaccine increases the breadth of immune response in mice. *PLOS ONE* **8**:e60245. DOI: https://doi.org/10.1371/journal.pone.0060245, PMID: 23555935

Ladell K, Hashimoto M, Iglesias MC, Wilmann PG, McLaren JE, Gras S, Chikata T, Kuse N, Fastenackels S, Gostick EJI. 2013. A molecular basis for the control of preimmune escape variants by HIV-specific CD8+ T. *Cells* **38**:425–436. DOI: https://doi.org/10.1016/j.immuni.2012.11.021

Lee S-K, Zhou S, Baldoni PL, Spielvogel E, Archin NM, Hudgens MG, Margolis DM, Swanstrom R. 2017. Quantification of the latent HIV-1 reservoir using ultra deep sequencing and primer ID in a viral outgrowth assay. *JAIDS Journal of Acquired Immune Deficiency Syndromes* **74**:221–228. DOI: https://doi.org/10.1097/QAI.0000000000001187

Leslie AJ, Pfafferott KJ, Chetty P, Draenert R, Addo MM, Feeney M, Tang Y, Holmes EC, Allen T, Prado JG, Altfeld M, Brander C, Dixon C, Ramduth D, Jeena P, Thomas SA, St John A, Roach TA, Kupfer B, Luzzi G, et al. 2004. HIV evolution: ctl escape mutation and reversion after transmission. *Nature Medicine* **10**:282–289. DOI: https://doi.org/10.1038/nm992, PMID: 14770175

Létourneau S, Im EJ, Mashishi T, Brereton C, Bridgeman A, Yang H, Dorrell L, Dong T, Korber B, McMichael AJ, Hanke T. 2007. Design and pre-clinical evaluation of a universal HIV-1 vaccine. *PLOS ONE* **2**:e984. DOI: https://doi.org/10.1371/journal.pone.0000984, PMID: 17912361

Liu MK, Hawkins N, Ritchie AJ, Ganusov VV, Whale V, Brackenridge S, Li H, Pavlicek JW, Cai F, Rose-Abrahams M, Treurnicht F, Hraber P, Riou C, Gray C, Ferrari G, Tanner R, Ping LH, Anderson JA, Swanstrom R, Cohen M, et al. 2013. Vertical T cell immunodominance and epitope entropy determine HIV-1 escape. *Journal of Clinical Investigation* **123**:380–393. DOI: https://doi.org/10.1172/JCI65330, PMID: 23221345

Lu CL, Pai JA, Nogueira L, Mendoza P, Gruell H, Oliveira TY, Barton J, Lorenzi JCC, Cohen YZ, Cohn LB, Klein F, Caskey M, Nussenzweig MC, Jankovic M. 2018. Relationship between intact HIV-1 proviruses in circulating CD4 + T cells and rebound viruses emerging during treatment interruption. *PNAS* **115**:E11341–E11348. DOI: https://doi.org/10.1073/pnas.1813512115, PMID: 30420517

Malnati MS, Scarlatti G, Gatto F, Salvatori F, Cassina G, Rutigliano T, Volpi R, Lusso P. 2008. A universal real-time PCR assay for the quantification of group-M HIV-1 proviral load. *Nature Protocols* **3**:1240–1248. DOI: https://doi.org/10.1038/nprot.2008.108, PMID: 18600229

Massanella M, Richman DD. 2016. Measuring the latent reservoir in vivo. *Journal of Clinical Investigation* **126**: 464–472. DOI: https://doi.org/10.1172/JCI80567, PMID: 26829625

McBrien JB, Mavigner M, Franchitti L, Smith SA, White E, Tharp GK, Walum H, Busman-Sahay K, Aguilera-Sandoval CR, Thayer WO, Spagnuolo RA, Kovarova M, Wahl A, Cervasi B, Margolis DM, Vanderford TH, Carnathan DG, Paiardini M, Lifson JD, Lee JH, et al. 2020. Robust and persistent reactivation of SIV and HIV by N-803 and depletion of CD8$^+$ cells. *Nature* **578**:154–159. DOI: https://doi.org/10.1038/s41586-020-1946-0, PMID: 31969705

McNeil AC, Shupert WL, Iyasere CA, Hallahan CW, Mican JA, Davey RT, Connors M. 2001. High-level HIV-1 viremia suppresses viral antigen-specific CD4(+) T cell proliferation. *PNAS* **98**:13878–13883. DOI: https://doi.org/10.1073/pnas.251539598, PMID: 11717444

Mothe B, Hu X, Llano A, Rosati M, Olvera A, Kulkarni V, Valentin A, Alicea C, Pilkington GR, Sardesai NY, Rocafort M, Crespo M, Carrillo J, Marco A, Mullins JI, Dorrell L, Hanke T, Clotet B, Pavlakis GN, Felber BK, et al. 2015a. A human immune data-informed vaccine concept elicits strong and broad T-cell specificities associated with HIV-1 control in mice and macaques. *Journal of Translational Medicine* **13**:60. DOI: https://doi.org/10.1186/s12967-015-0392-5, PMID: 25879820

Mothe B, Climent N, Plana M, Rosàs M, Jiménez JL, Muñoz-Fernández MÁ, Puertas MC, Carrillo J, Gonzalez N, León A, Pich J, Arnaiz JA, Gatell JM, Clotet B, Blanco J, Alcamí J, Martinez-Picado J, Alvarez-Fernández C, Sánchez-Palomino S, Guardo AC, et al. 2015b. Safety and immunogenicity of a modified vaccinia Ankara-based HIV-1 vaccine (MVA-B) in HIV-1-infected patients alone or in combination with a drug to reactivate latent HIV-1. *Journal of Antimicrobial Chemotherapy* **70**:1833–1842. DOI: https://doi.org/10.1093/jac/dkv046, PMID: 25724985

Nettles RE. 2005. Intermittent HIV-1 viremia (Blips) and drug resistance in patients receiving HAART. *Jama* **293**: 817–829. DOI: https://doi.org/10.1001/jama.293.7.817

Ondondo B, Murakoshi H, Clutton G, Abdul-Jawad S, Wee EG, Gatanaga H, Oka S, McMichael AJ, Takiguchi M, Korber B, Hanke T. 2016. Novel Conserved-region T-cell mosaic vaccine with high global HIV-1 coverage is recognized by protective responses in untreated infection. *Molecular Therapy* **24**:832–842. DOI: https://doi.org/10.1038/mt.2016.3, PMID: 26743582

Pandrea I, Gaufin T, Gautam R, Kristoff J, Mandell D, Montefiori D, Keele BF, Ribeiro RM, Veazey RS, Apetrei C. 2011. Functional cure of SIVagm infection in rhesus macaques results in complete recovery of CD4+ T cells and is reverted by CD8+ cell depletion. *PLOS Pathogens* **7**:e1002170. DOI: https://doi.org/10.1371/journal.ppat.1002170, PMID: 21829366

Pankau MD, Reeves DB, Harkins E, Ronen K, Jaoko W, Mandaliya K, Graham SM, McClelland RS, Matsen IV FA, Schiffer J. 2020. Dynamics of HIV DNA reservoir seeding in a cohort of superinfected kenyan. *Women* **16**: e1008286. DOI: https://doi.org/10.1371/journal.ppat.1008286

Phillips RE, Rowland-Jones S, Nixon DF, Gotch FM, Edwards JP, Ogunlesi AO, Elvin JG, Rothbard JA, Bangham CR, Rizza CR. 1991. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature* **354**:453–459. DOI: https://doi.org/10.1038/354453a0, PMID: 1721107

Precopio ML, Butterfield TR, Casazza JP, Little SJ, Richman DD, Koup RA, Roederer M. 2008. Optimizing peptide matrices for identifying T-cell antigens. *Cytometry Part A* **73A**:1071–1078. DOI: https://doi.org/10.1002/cyto.a.20646

Price DA, Goulder PJR, Klenerman P, Sewell AK, Easterbrook PJ, Troop M, Bangham CRM, Phillips RE. 1997. Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. *PNAS* **94**:1890–1895. DOI: https://doi.org/10.1073/pnas.94.5.1890

Price MN, Dehal PS, Arkin AP. 2010. FastTree 2–approximately maximum-likelihood trees for large alignments. *PLOS ONE* **5**:e9490. DOI: https://doi.org/10.1371/journal.pone.0009490, PMID: 20224823

Radebe M, Gounder K, Mokgoro M, Ndhlovu ZM, Mncube Z, Mkhize L, van der Stok M, Jaggernath M, Walker BD, Ndung'u T. 2015. Broad and persistent Gag-specific CD8+ T-cell responses are associated with viral control but rarely drive viral escape during primary HIV-1 infection. *Aids* **29**:23–33. DOI: https://doi.org/10.1097/QAD.0000000000000508, PMID: 25387316

Rehr M, Cahenzli J, Haas A, Price DA, Gostick E, Huber M, Karrer U, Oxenius A. 2008. Emergence of polyfunctional CD8+ T cells after prolonged suppression of human immunodeficiency virus replication by antiretroviral therapy. *Journal of Virology* **82**:3391–3404. DOI: https://doi.org/10.1128/JVI.02383-07, PMID: 18199637

Riou C, Ganusov VV, Campion S, Mlotshwa M, Liu MK, Whale VE, Goonetilleke N, Borrow P, Ferrari G, Betts MR, Haynes BF, McMichael AJ, Gray CM. 2012. Distinct kinetics of Gag-specific CD4+ and CD8+ T cell responses during acute HIV-1 infection. *Journal of Immunology* **188**:2198–2206. DOI: https://doi.org/10.4049/jimmunol.1102813

Ritchie AJ, Cai F, Smith NM, Chen S, Song H, Brackenridge S, Abdool Karim SS, Korber BT, McMichael AJ, Gao F, Goonetilleke N. 2014. Recombination-mediated escape from primary CD8+ T cells in acute HIV-1 infection. *Retrovirology* **11**:69. DOI: https://doi.org/10.1186/s12977-014-0069-9, PMID: 25212771

Roederer M, Koup RA. 2003. Optimized determination of T cell epitope responses. *Journal of Immunological Methods* **274**:221–228. DOI: https://doi.org/10.1016/S0022-1759(02)00423-4

Rolland M, Nickle DC, Mullins JI. 2007. HIV-1 group M conserved elements vaccine. *PLOS Pathogens* **3**:e157. DOI: https://doi.org/10.1371/journal.ppat.0030157, PMID: 18052528

Rosenbloom DIS, Hill AL, Laskey SB, Siliciano RF. 2017. Re-evaluating evolution in the HIV reservoir. *Nature* **551**: E6–E9. DOI: https://doi.org/10.1038/nature24634

Rouzine IM, Rodrigo A, Coffin JM. 2001. Transition between stochastic evolution and deterministic evolution in the presence of selection: general theory and application to virology. *Microbiology and Molecular Biology Reviews* **65**:151–185. DOI: https://doi.org/10.1128/MMBR.65.1.151-185.2001, PMID: 11238990

Rutishauser RL, Hartogensis W, Deguit CD, Krone M, Hoh R, Hecht FM, Pilcher CD, Bacchetti P, Deeks SG, Hunt PW, McCune JM. 2017. Early and delayed antiretroviral therapy results in comparable reductions in CD8+ T Cell Exhaustion Marker Expression. *AIDS Research and Human Retroviruses* **33**:658–667. DOI: https://doi.org/10.1089/aid.2016.0324, PMID: 28335609

Salantes DB, Zheng Y, Mampe F, Srivastava T, Beg S, Lai J, Li JZ, Tressler RL, Koup RA, Hoxie J, Abdel-Mohsen M, Sherrill-Mix S, McCormick K, Overton ET, Bushman FD, Learn GH, Siliciano RF, Siliciano JM, Tebas P, Bar KJ. 2018. HIV-1 latent reservoir size and diversity are stable following brief treatment interruption. *Journal of Clinical Investigation* **128**:3102–3115. DOI: https://doi.org/10.1172/JCI120194, PMID: 29911997

Salazar-Gonzalez JF, Bailes E, Pham KT, Salazar MG, Guffey MB, Keele BF, Derdeyn CA, Farmer P, Hunter E, Allen S, Manigart O, Mulenga J, Anderson JA, Swanstrom R, Haynes BF, Athreya GS, Korber BT, Sharp PM, Shaw GM, Hahn BH. 2008. Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *Journal of Virology* **82**:3952–3970. DOI: https://doi.org/10.1128/JVI.02660-07, PMID: 18256145

Schmitz JE, Johnson RP, McClure HM, Manson KH, Wyand MS, Kuroda MJ, Lifton MA, Khunkhun RS, McEvers KJ, Gillis J, Piatak M, Lifson JD, Grosschupff G, Racz P, Tenner-Racz K, Rieber EP, Kuus-Reichel K, Gelman RS, Letvin NL, Montefiori DC, et al. 2005. Effect of CD8+ lymphocyte depletion on virus containment after simian immunodeficiency virus SIVmac251 challenge of live attenuated SIVmac239delta3-vaccinated rhesus macaques. *Journal of Virology* **79**:8131–8141. DOI: https://doi.org/10.1128/JVI.79.13.8131-8141.2005, PMID: 15956558

Shan L, Deng K, Shroff NS, Durand CM, Rabi SA, Yang HC, Zhang H, Margolick JB, Blankson JN, Siliciano RF. 2012. Stimulation of HIV-1-specific cytolytic T lymphocytes facilitates elimination of latent viral reservoir after virus reactivation. *Immunity* **36**:491–501. DOI: https://doi.org/10.1016/j.immuni.2012.01.014, PMID: 22406268

Shannon CE. 1948. A mathematical theory of communication. *The Bell System Technical Journal* **5**:3–55. DOI: https://doi.org/10.1002/j.1538-7305.1948.tb00917.x

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. *Nature Medicine* **9**:727–728. DOI: https://doi.org/10.1038/nm880, PMID: 12754504

Siliciano JD, Siliciano RF. 2004. A long-term latent reservoir for HIV-1: discovery and clinical implications. *Journal of Antimicrobial Chemotherapy* **54**:6–9. DOI: https://doi.org/10.1093/jac/dkh292, PMID: 15163657

Siliciano JD, Siliciano RF. 2005. Enhanced culture assay for detection and quantitation of latently infected, resting CD4+ T-cells carrying replication-competent virus in HIV-1-infected individuals. *Human Retrovirus Protocols* **304**:3–15. DOI: https://doi.org/10.1385/1-59259-907-9:003

Siliciano JM, Siliciano RF. 2015. The remarkable stability of the latent reservoir for HIV-1 in resting memory CD4+ T cells. *The Journal of Infectious Diseases* **212**:1345–1347. DOI: https://doi.org/10.1093/infdis/jiv219, PMID: 25877551

Smith NM, Mlcochova P, Watters SA, Aasa-Chapman MM, Rabin N, Moore S, Edwards SG, Garson JA, Grant PR, Ferns RB, Kashuba A, Mayor NP, Schellekens J, Marsh SG, McMichael AJ, Perelson AS, Pillay D, Goonetilleke N, Gupta RK. 2015. Proof-of-Principle for immune control of global HIV-1 reactivation in vivo. *Clinical Infectious Diseases* **61**:120–128. DOI: https://doi.org/10.1093/cid/civ219, PMID: 25778749

Søgaard OS, Graversen ME, Leth S, Olesen R, Brinkmann CR, Nissen SK, Kjaer AS, Schleimann MH, Denton PW, Hey-Cunningham WJ, Koelsch KK, Pantaleo G, Krogsgaard K, Sommerfelt M, Fromentin R, Chomont N, Rasmussen TA, Østergaard L, Tolstrup M. 2015. The depsipeptide romidepsin reverses HIV-1 latency in vivo. *PLOS Pathogens* **11**:e1005142. DOI: https://doi.org/10.1371/journal.ppat.1005142, PMID: 26379282

Song H, Pavlicek JW, Cai F, Bhattacharya T, Li H, Iyer SS, Bar KJ, Decker JM, Goonetilleke N, Liu MK, Berg A, Hora B, Drinker MS, Eudailey J, Pickeral J, Moody MA, Ferrari G, McMichael A, Perelson AS, Shaw GM, et al. 2012. Impact of immune escape mutations on HIV-1 fitness in the context of the cognate transmitted/founder genome. *Retrovirology* **9**:89. DOI: https://doi.org/10.1186/1742-4690-9-89, PMID: 23110705

Streeck H, Li B, Poon AF, Schneidewind A, Gladden AD, Power KA, Daskalakis D, Bazner S, Brander C, Rosenberg ES, Frost SD, Altfeld M, Allen TM. 2008. Immune-driven recombination and loss of control after HIV superinfection. *Journal of Experimental Medicine* **205**:1789–1796. DOI: https://doi.org/10.1084/jem.20080281, PMID: 18625749

Sung JA, Lam S, Garrido C, Archin N, Rooney CM, Bollard CM, Margolis DM. 2015. Expanded cytotoxic T-cell lymphocytes target the latent HIV reservoir. *Journal of Infectious Diseases* **212**:258–263. DOI: https://doi.org/10.1093/infdis/jiv022, PMID: 25589335

Townsend AR, McMichael AJ, Carter NP, Huddleston JA, Brownlee GG. 1984. Cytotoxic T cell recognition of the influenza nucleoprotein and hemagglutinin expressed in transfected mouse L cells. *Cell* **39**:13–25. DOI: https://doi.org/10.1016/0092-8674(84)90187-9, PMID: 6091906

Townsend AR, Rothbard J, Gotch FM, Bahadur G, Wraith D, McMichael AJ. 1986. The epitopes of influenza nucleoprotein recognized by cytotoxic T lymphocytes can be defined with short synthetic peptides. *Cell* **44**: 959–968. DOI: https://doi.org/10.1016/0092-8674(86)90019-X, PMID: 2420472

Townsend AR, McMichael AJ. 1985. Specificity of cytotoxic T lymphocytes stimulated with influenza virus studies in mice and humans. *Progress in Allergy* **36**:10–43. DOI: https://doi.org/10.1159/000409860, PMID: 3881777

Trumble IM, Allmon AG, Archin NM, Rigdon J, Francis O, Baldoni PL, Hudgens MG. 2017. SLDAssay: a software package and web tool for analyzing limiting dilution assays. *Journal of Immunological Methods* **450**:10–16. DOI: https://doi.org/10.1016/j.jim.2017.07.004, PMID: 28733216

Van Zyl GU, Katusiime MG, Wiegand A, McManus WR, Bale MJ, Halvas EK, Luke B, Boltz VF, Spindler J, Laughton B, Engelbrecht S, Coffin JM, Cotton MF, Shao W, Mellors JW, Kearney MF. 2017. No evidence of HIV replication in children on antiretroviral therapy. *Journal of Clinical Investigation* **127**:3827–3834. DOI: https://doi.org/10.1172/JCI94582

Vibholm LK, Lorenzi JCC, Pai JA, Cohen YZ, Oliveira TY, Barton JP, Garcia Noceda M, Lu CL, Ablanedo-Terrazas Y, Del Rio Estrada PM, Reyes-Teran G, Tolstrup M, Denton PW, Damsgaard T, Søgaard OS, Nussenzweig MC. 2019. Characterization of intact proviruses in blood and lymph node from HIV-Infected individuals undergoing analytical treatment interruption. *Journal of Virology* **93**:e01920-18. DOI: https://doi.org/10.1128/JVI.01920-18, PMID: 30700598

Warren JA, Clutton G, Goonetilleke N. 2019. Harnessing CD8+ T cells under HIV. *Antiretroviral Therapy* **10**:291. DOI: https://doi.org/10.3389/fimmu.2019.00291

Xu Y, Trumble IM, Warren JA, Clutton G, Abad-Fernandez M, Kirchnerr J, Adimora AA, Deeks SG, Margolis DM, Kuruc JD, Gay CL, Archin NM, Mollan KR, Hudgens M, Goonetilleke N. 2019. HIV-Specific T cell responses are highly stable on antiretroviral therapy. *Molecular Therapy - Methods & Clinical Development* **15**:9–17. DOI: https://doi.org/10.1016/j.omtm.2019.07.008, PMID: 31534983

Zhou S, Jones C, Mieczkowski P, Swanstrom R. 2015. Primer ID validates template sampling depth and greatly reduces the error rate of Next-Generation sequencing of HIV-1 genomic RNA populations. *Journal of Virology* **89**:8540–8555. DOI: https://doi.org/10.1128/JVI.00522-15, PMID: 26041299