LETTER

Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations

Kai Deng[1], Mihaela Pertea[1,2], Anthony Rongvaux[3], Leyao Wang[4], Christine M. Durand[1], Gabriel Ghiaur[5], Jun Lai[1], Holly L. McHugh[1], Haiping Hao[6], Hao Zhang[7], Joseph B. Margolick[7], Cagan Gurer[8], Andrew J. Murphy[8], David M. Valenzuela[8], George D. Yancopoulos[8], Steven G. Deeks[9], Till Strowig[3], Priti Kumar[10], Janet D. Siliciano[1], Steven L. Salzberg[2,11], Richard A. Flavell[3,12], Liang Shan[3] & Robert F. Siliciano[1,13]

Despite antiretroviral therapy (ART), human immunodeficiency virus (HIV)-1 persists in a stable latent reservoir[1,2], primarily in resting memory CD4[+] T cells[3,4]. This reservoir presents a major barrier to the cure of HIV-1 infection. To purge the reservoir, pharmacological reactivation of latent HIV-1 has been proposed[5] and tested both *in vitro* and *in vivo*[6–8]. A key remaining question is whether virus-specific immune mechanisms, including cytotoxic T lymphocytes (CTLs), can clear infected cells in ART-treated patients after latency is reversed. Here we show that there is a striking all or none pattern for CTL escape mutations in HIV-1 Gag epitopes. Unless ART is started early, the vast majority (>98%) of latent viruses carry CTL escape mutations that render infected cells insensitive to CTLs directed at common epitopes. To solve this problem, we identified CTLs that could recognize epitopes from latent HIV-1 that were unmutated in every chronically infected patient tested. Upon stimulation, these CTLs eliminated target cells infected with autologous virus derived from the latent reservoir, both *in vitro* and in patient-derived humanized mice. The predominance of CTL-resistant viruses in the latent reservoir poses a major challenge to viral eradication. Our results demonstrate that chronically infected patients retain a broad-spectrum viral-specific CTL response and that appropriate boosting of this response may be required for the elimination of the latent reservoir.

HIV-1 establishes latent infection in resting CD4[+] T cells[3,4]. Recent efforts to eradicate HIV-1 infection have focused on reversing latency without global T-cell activation[5]. However, inducing HIV-1 gene expression in latently infected cells is not sufficient to cause the death of these cells if they remain in a resting state[9]. Boosting HIV-1-specific immune responses, including CTL responses, may be required for clearance of the latent reservoir[9]. CTLs have a significant role in suppressing HIV-1 replication in acute infection[10–14]. Because of this strong selective pressure, HIV-1 quickly acquires mutations to evade CTL recognition[12,13,15–18]. CTL escape has been studied primarily through the analysis of plasma virus[12,13,16,18–20], and CTL-based vaccines have been designed based on conserved epitopes[21,22]. A systematic investigation of CTL escape in the latent reservoir will be of great importance to the ongoing CTL-based virus eradication efforts, because latent HIV-1 probably represents the major source of viral rebound after treatment interruption. Earlier studies have suggested the presence of CTL escape mutations in proviral DNA[15,17], but it remains unclear to what extent the latent reservoir in resting CD4[+] T cells is affected by CTL escape, whether mutations detected in proviral DNA are representative of the very small fraction of proviruses that are replication competent, and, most importantly, whether the CTL response can recognize and clear infected cells after latency is reversed.

To investigate CTL escape variants in the latent reservoir, we deep sequenced the proviral HIV-1 DNA in resting CD4[+] T cells from 25 patients (Extended Data Table 1). Among them, 10 initiated ART during the acute phase (AP; within 3 months of infection) while the other 15 initiated ART during the chronic phase (CP) of infection. The sequencing was focused on Gag because it is an important target of the CTL response[23] and is highly conserved, which facilitates the detection of escape variants. Our data show that previously documented CTL escape variants completely dominate the viral reservoirs of nearly all CP-treated patients (Extended Data Fig. 1 and Supplementary Table 1). This trend is especially obvious for several well characterized CTL epitopes: the human leukocyte antigen (HLA)-A2-restricted epitope SLYNTVATL (SL9), the HLA-A3-restricted epitope RLRPGGKKK (RK9) and the HLA-B57/58-restricted epitope TSTLQEQIGW (TW10) (Fig. 1a and Extended Data Fig. 1, highlighted in coloured boxes.). In these epitopes, close to 100% of the sequences harboured escape mutations. Comparison of mutation frequencies between *HLA* allele-relevant and -irrelevant epitopes in CP-treated patients suggests that the CTL escape mutations identified are specific to each patient's HLA type (Fig. 1b). By contrast, except for SL9 from patient AP01 and RK9 from patient AP08, few if any CTL escape mutations were archived in AP-treated patients (Fig. 1c and Extended Data Fig. 1). The striking difference between AP- and CP-treated patients (Fig. 1c) indicates that, unless treatment is initiated within the first several months of infection, the latent reservoir becomes almost completely dominated by variants resistant to dominant CTL responses.

To confirm that variants detected at high frequency in the latent reservoir represent functional CTL escape mutants, cells from seven CP-treated subjects were tested for reactivity to synthetic peptides representing wild-type and mutant versions of the relevant epitopes. As expected, there were only minimal responses to previously documented CTL escape mutants by patient CD8[+] T cells, and no *de novo* response was detected (Fig. 1d and Extended Data Fig. 2). In contrast, all tested subjects retained a strong response to peptides representing the wild-type epitopes, suggesting that the wild-type virus was initially transmitted, with subsequent evolution of CTL escape variants. Most HIV-1 proviruses detected in patients are defective[24]. Therefore, to determine whether these CTL escape variants can be reactivated and lead to viral rebound if therapy is stopped, we isolated replication-competent viruses from the latent reservoirs of nine CP-treated patients. We found that all the dominant CTL escape mutations that had been identified in proviruses in resting CD4[+] T cells were also present in the replication-competent viruses that grew out after T-cell activation (Fig. 1e and Extended Data Fig. 3), indicating that these CTL escape variants not only dominate the population of proviruses, but can also be released and replicate once latency is reversed.

[1]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. [2]Center for Computational Biology, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. [3]Department of Immunobiology, Yale University School of Medicine, New Haven, Connecticut 06510, USA. [4]Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, Connecticut 06510, USA. [5]Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. [6]Deep Sequencing and Microarray Core, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. [7]Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205, USA. [8]Regeneron Pharmaceuticals Inc., Tarrytown, New York 10591, USA. [9]Department of Medicine, University of California, San Francisco, San Francisco, California 94110, USA. [10]Department of Medicine, Yale University School of Medicine, New Haven, Connecticut 06510, USA. [11]Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA. [12]Howard Hughes Medical Institute, New Haven, Connecticut 06510, USA. [13]Howard Hughes Medical Institute, Baltimore, Maryland 21205, USA.
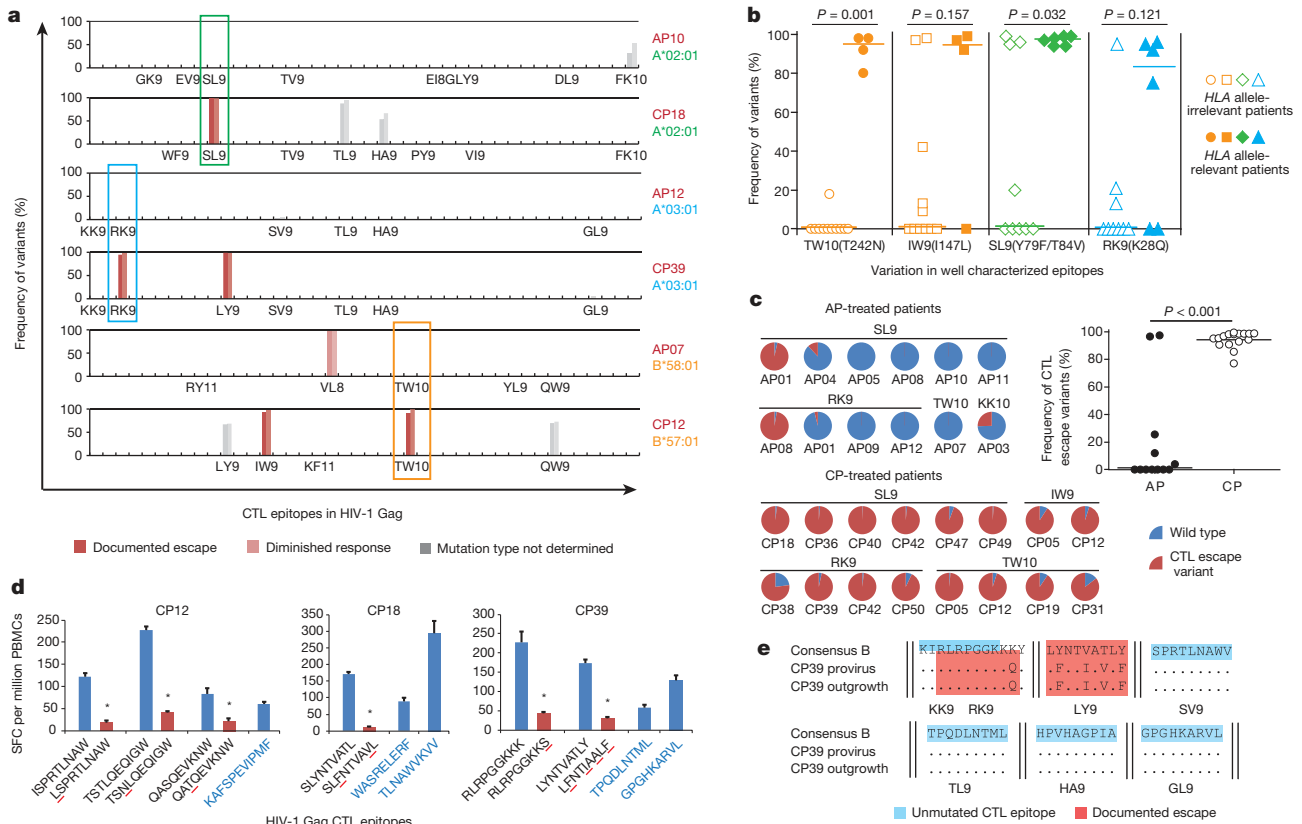
Figure 1 | CTL escape variants dominate the latent reservoir of CP-treated but not AP-treated patients. a, Frequency of variants in Gag CTL epitopes in proviruses from resting CD4[+] T cells. Representative results from six patients are shown. Only optimal CTL epitopes relevant to each patient's HLA type are listed. Results from both PacBio (left bar) and MiSeq (right bar) sequencing are shown. The effect on CTL recognition (denoted by colour) is determined from information in the Los Alamos National Laboratory (LANL) HIV Molecular Immunology Database. **b**, CTL escape variants identified by sequencing are specific to HLA type. Frequencies of documented escape-associated variants in four well characterized epitopes are shown for all 15 CP-treated patients. Medians and *P* values from Mann–Whitney test are shown. **c**, Comparison of CTL escape variant frequency in proviruses between CP- and AP-treated patients. Only well characterized epitopes are shown.

Medians and *P* values from Mann–Whitney test are shown. **d**, Characterization of CTL responses against HIV-1 Gag epitopes by interferon-γ ELISpot. The peptides tested are listed below the *x*-axis (black type, epitopes in which sequence variation was detected; blue type, no variation). The observed mutation is underlined in red, and CTL escape (defined by the absence of positive response after mutation) is denoted by an asterisk above the bar. The peptide concentration was 10 μg ml[−1]. PBMCs, peripheral blood mononuclear cells; SFC, spot-forming cell. Error bars represent standard error of the mean (s.e.m.), *n* = 3. **e**, Sequences in Gag CTL epitopes for proviral DNA and outgrowth virus from resting CD4[+] T cells in patient CP39. CTL epitopes with no observed variation are highlighted in blue. Epitopes with documented escape mutations are shaded in red.

We next asked whether the host CTL response could recognize and eliminate the cells infected with these escape variants. We infected activated CD4[+] T cells from these patients with autologous, replication-competent virus derived from the latent reservoir (Extended Data Fig. 4a). The infected cells were then co-cultured with autologous CD8[+] T cells, either unstimulated or pre-stimulated, to assess HIV-1-specific cytolytic activity. Non-specific activation of CD8[+] T cells was not observed after co-culture with phytohaemagglutinin (PHA)-activated CD4[+] T cells (Extended Data Fig. 4b). From all 13 CP-treated subjects tested, CD8[+] T cells pre-stimulated by a Gag peptide mixture efficiently killed autologous infected CD4[+] T cells (median 61% elimination), while unstimulated CD8[+] T cells from most subjects had significantly less effect (median 23% elimination) (Fig. 2a and Extended Data Fig. 4c, d). CD8[+] T cells from 7/7 healthy donors completely failed to eliminate autologous infected cells (Fig. 2a), confirming that the observed killing was HIV-1 specific. The killing effect was enhanced by increasing the effector to target ratio (Extended Data Fig. 5a), and was cell–cell contact dependent (Extended Data Fig. 5b). When the co-culture was maintained over time in the absence of ART, viral replication was significantly reduced, but not completely inhibited by pre-stimulated CD8[+] T cells (Fig. 2b). We found that peptide mixtures from other HIV-1 proteins (Nef, Tat, Rev and Env) could also boost CTL responses and facilitate the elimination of infected cells (Fig. 2c), and that CTLs pre-stimulated

by Gag peptides generally had the highest activity. Together, these results demonstrate that chronically infected patients retain CTL clones that can recognize and eliminate autologous infected CD4[+] T cells, despite the presence of CTL escape mutations in dominant epitopes. However, these clones require stimulation with antigen for optimal activity.

To characterize further which CTL population contributed to the elimination of cells infected by CTL escape variants, we compared the killing activity of two specific CTL populations: the population that targets epitopes in which escape has been identified and the one that targets unmutated epitopes (Fig. 3a). CD8[+] T cells from patients CP36 and CP39 were pre-stimulated with interleukin (IL)-2 and different synthetic peptides representing the wild-type forms of the relevant epitopes. After incubation for 6 days, each CTL population exhibited significant proliferation compared with no treatment or IL-2 alone (Fig. 3b, c). Pentamer staining for three available epitopes revealed that the number of epitope-specific CD8[+] T cells increased dramatically after stimulation with wild-type peptides (Fig. 3b). After co-culture with autologous target cells infected with latent reservoir-derived viruses, CTLs targeting unmutated epitopes clearly showed stronger cytolytic activity than the IL-2 only controls, while CTLs targeting epitopes with identified escaped mutations showed no significant killing (Fig. 3d). CTLs pre-stimulated by the Gag peptide mixture exhibited stronger killing than all single-peptide-stimulated populations (Fig. 3d).
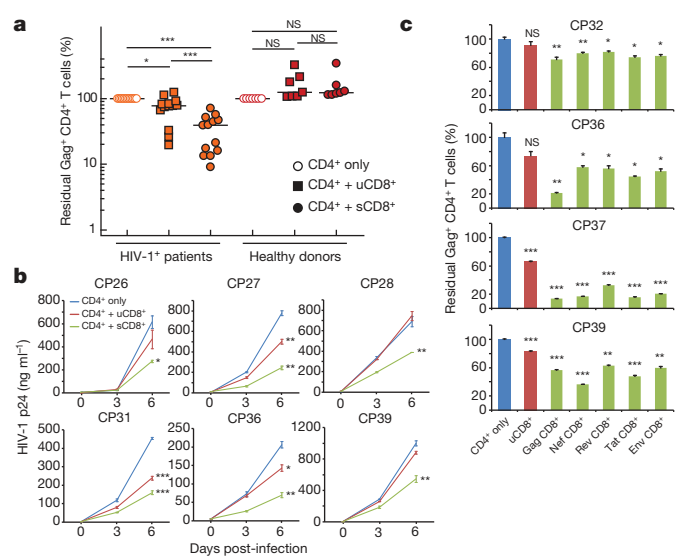
Figure 2 | CD8+ T cells pre-stimulated with a mixture of Gag peptides eliminate autologous CD4+ T cells infected with autologous HIV-1 from resting CD4+ T cells. a, Pre-stimulated CD8+ T cells (sCD8+) eliminate autologous infected CD4+ T cells more efficiently than unstimulated CD8+ T cells (uCD8+). Each symbol represents the mean of three replicates. Medians and P values from Mann–Whitney test are shown. **b,** sCD8+ cells inhibit viral growth in autologous infected CD4+ T cells with higher efficacy than uCD8+ cells. p24, HIV-1 capsid (core) protein. **c,** sCD8+ cells pre-stimulated by different viral peptides eliminate autologous CD4+ T cells infected with viruses derived from resting CD4+ T cells. **b, c,** Results were compared with CD4+ only using paired t-tests. Error bars represent s.e.m., $n = 3$. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, NS, not significant ($P > 0.05$).

reported mouse system named MISTRG[25]. Whereas the previously reported MISTRG mice bear a bacterial artificial chromosome (BAC) transgene encoding human SIRP-α, the newly generated MIS(KI)TRG mice harbour a knock-in replacement of the endogenous mouse *Sirpa* gene with a humanized version. With humanization by knock-in replacement of the *Csf1*, *Csf2*, *Il3*, *Tpo* and *Sirpa* genes in the *Rag2*[−/−] *Il2rg*[−/−] genetic background, MIS(KI)TRG mice are highly permissive for human haematopoiesis and support the reconstitution of robust human lymphoid and myelomonocytic systems. With the demonstrated development of functional T lymphocytes and monocytes/macrophages, MIS(KI)TRG mice provide a useful humanized mouse host for HIV-1 infection studies. Bone marrow biopsies were obtained from study participants and
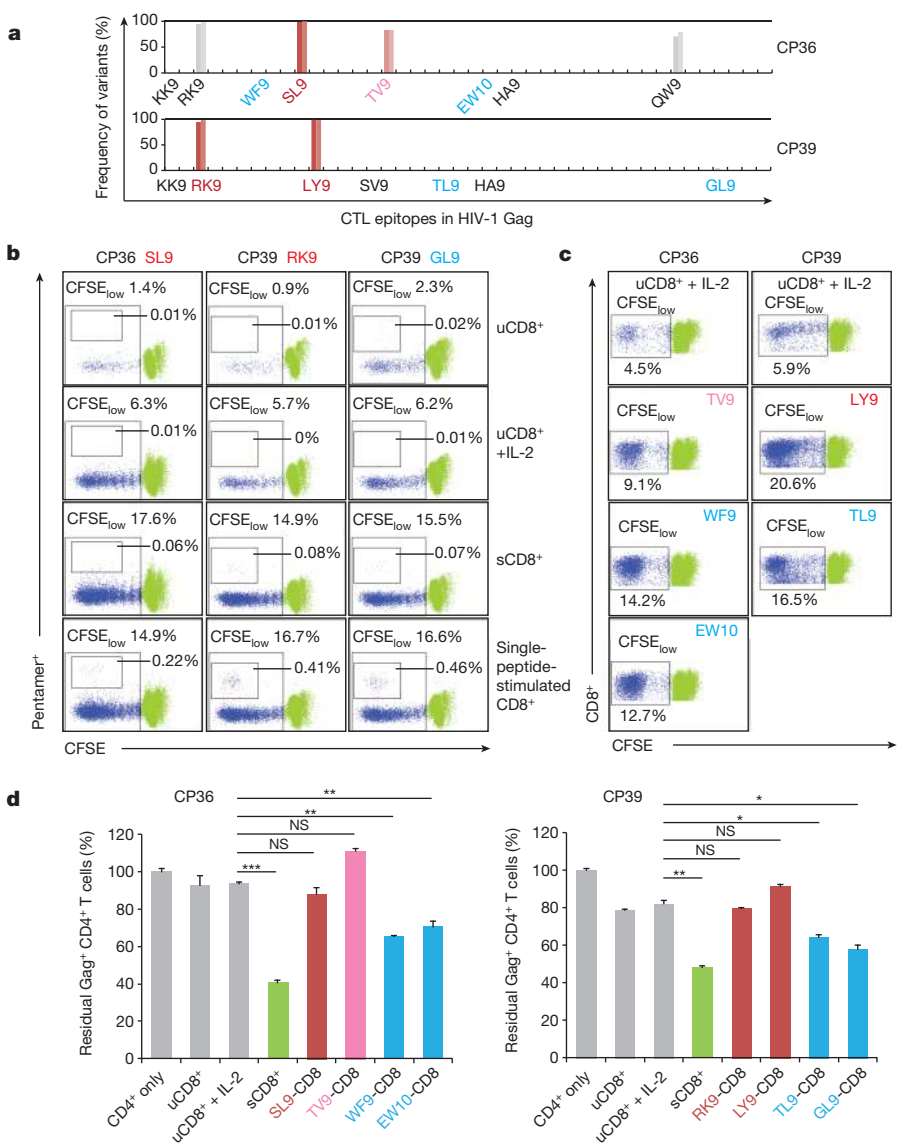
To test whether CTLs that recognize unmutated viral epitopes can inhibit HIV-1 replication and clear infected cells *in vivo*, we generated patient-derived humanized mice using an improved version of a recently

Figure 3 | CD8+ T cells targeting unmutated epitopes, not epitopes with identified escape mutations, eliminate CTL escape variants.
a, Frequency of variants in Gag CTL epitopes in proviruses from resting CD4+ T cells from patients CP36 and CP39. Epitopes tested with single-peptide stimulation are denoted in colours (red or pink, epitopes with escape observed; blue, no escape observed). **b,** Epitope-specific CD8+ T cells proliferate markedly after single-peptide stimulation. Only CD8+ T cells are shown. Percentages of carboxyfluorescein succinimidyl ester (CFSE)low, pentamer-positive cells are indicated for unstimulated cultures (uCD8+) with or without IL-2, for cultures stimulated with Gag peptide mixture (sCD8+) and IL-2, and for cultures stimulated with the indicated single peptides and IL-2. Wild-type versions of peptides were used for all single-peptide stimulations. **c,** CD8+ T-cell proliferative responses after single-peptide stimulation. Only CD8+ T cells are shown. Percentages of CFSElow cells are indicated. **d,** CD8+ T cells targeting unmutated epitopes, not epitopes with identified escaped mutations, eliminate autologous CD4+ T cells infected with CTL escape variants. Error bars represent s.e.m., $n = 3$. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, NS, not significant ($P > 0.05$), paired t-test.

purified CD34$^+$ cells were used to reconstitute the MIS$^{(KI)}$TRG mice. We infected these patient-derived humanized mice with primary HIV-1 isolates grown from resting CD4$^+$ T cells from the same patient and then evaluated the antiviral effect of autologous CD8$^+$ T cells (Fig. 4a). MIS$^{(KI)}$TRG mice engrafted with bone marrow CD34$^+$ cells from patient CP18 successfully developed human T-lymphocyte and monocyte/macrophage subsets (Fig. 4b, c), which were sufficient to support HIV-1 infection (Fig. 4d). Plasma HIV-1 RNA levels peaked 20–30 days after infection (Extended Data Fig. 6a). Depletion of CD4$^+$ T cells was clearly evident 12 days after infection in peripheral blood and spleen (Fig. 4e and Extended Data Fig. 6b, c). Cell-associated HIV-1 RNA was detected in both T cells and macrophages/monocytes (Extended Data Fig. 6d). Viral infection was also observed in various tissues in which a large number of memory CD4$^+$ T cells were detected (Extended Data Fig. 7). In control mice or mice that received autologous patient CD8$^+$ T cells pre-stimulated with a peptide representing the unmutated dominant SL9 epitope, levels of plasma HIV-1 RNA and proviral DNA in peripheral blood continued to increase from day 14 to day 29 after infection (Fig. 4f). In sharp contrast, mice that received CD8$^+$ T cells pre-stimulated with unmutated epitopes (Gag mix or WF9) had a significantly lower level of viral replication (Fig. 4g, h). Dramatic decreases in plasma HIV-1 RNA of 100- to 1,000-fold were observed in all three mice that received CD8$^+$ T cells pre-stimulated with the mixture of Gag peptides including dominant and subdominant epitopes. Two of three mice had undetectable levels of plasma HIV-1 RNA and proviral DNA in peripheral blood measured at three time points (Fig. 4g). We performed the same experiments using patient CP36-derived humanized mice and a reduction of peripheral HIV-1 RNA and DNA levels was also observed in mice that received CP36 CD8$^+$ T cells pre-stimulated with the mixture of Gag peptides (Extended Data Fig. 8). Since the post-engraftment lifespan of MIS$^{(KI)}$TRG mice is only 10–12 weeks[25], we were only able to investigate the acute phase of HIV-1 infection and demonstrate the *in vivo* functionality of patient CD8$^+$ T cells. Future developments of the MIS$^{(KI)}$TRG model will prolong the post-engraftment lifespan of these mice and allow studies of the establishment and clearance of the HIV-1 latent reservoir *in vivo*. Together, our *in vitro* and *in vivo* experiments demonstrate that only CTL clones targeting unmutated epitopes are effective against cells infected with the viral variants that are likely to represent the major source of rebound HIV-1 after reversal of latency.

The seeding of the HIV-1 latent reservoir starts just a few days after infection[26], before the development of a robust CTL response[14]. This is consistent with our finding that patients who initiated treatment early, in the acute infection stage, have few if any CTL escape variants archived in the latent reservoir. However, if treatment was initiated in chronic infection, CTL escape variants became dominant in the latent reservoir, indicating a complete replacement of the initially established 'wild-type' reservoir. The mechanism behind this replacement warrants further investigation, but probably reflects the dynamic nature of the reservoir in untreated infection. In any event, the overwhelming presence of escape variants in the latent reservoir of chronic patients certainly presents an additional barrier to eradication efforts. The striking difference between AP- and CP-treated patients presents another argument for early treatment of HIV-1 infection; early treatment not only reduces the size of the latent reservoir[27], but also alters the composition of the reservoir, as shown here, in a way that may enhance the efficacy of potential CTL-based eradication therapies.

The hierarchy of HIV-1-specific CTL response in acute infection appears to have an important role in initial viral suppression, as demonstrated by the fact that certain immunodominant CTL populations are frequently linked to lower set point viraemia later in infection[17,28]. These immunodominant responses in acute infection have been identified as the major selection force driving the development of CTL escape mutations[13,20]. Here we show that these immunodominant response-driven mutations are not only archived in the latent reservoir, but also in fact dominate the latent provirus population in CP-treated patients.
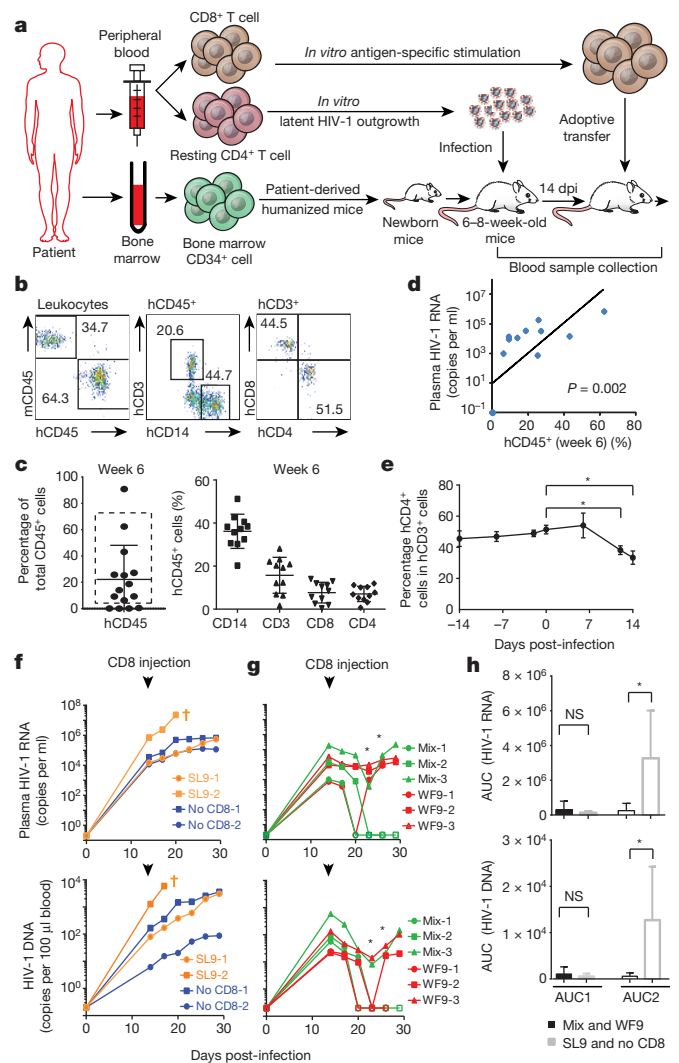

Figure 4 | Broad-spectrum CTLs suppress *in vivo* replication of HIV-1 from the latent reservoir of the same patients in patient-derived humanized mice. **a**, Experimental design. dpi, days post-infection. **b**, **c**, Efficient engraftment of patient CP18-derived haematopoietic cells in MIS$^{(KI)}$TRG mice at week 6. Representative flow cytometry analysis (**b**) and summary (**c**) of human CD45$^+$ cells, human T-lymphocyte and monocyte subsets. Eleven out of fifteen mice (enclosed in the rectangle) were used for HIV-1 infection. h, human; m, mouse. **d**, Correlation between frequency of peripheral human CD45$^+$ cells (6 weeks after engraftment) and plasma HIV-1 RNA levels (14 days after infection). **e**, Depletion of human CD4$^+$ T cells in peripheral blood after HIV-1 infection. *$P < 0.05$, paired t-test, $n = 11$. **f**, **g**, Reduction of levels of plasma HIV-1 RNA and copies of peripheral blood HIV-1 DNA after injection of viral-specific CTLs. Filled symbols, above detection limit; open symbols, below detection limit. *$P < 0.05$, unpaired t-test. **h**, Effect of CTLs on the level of viral replication *in vivo*. The area under the curve (AUC) of the viraemia versus time plot for each mouse in **f** ($n = 3$) and **g** ($n = 6$) before (AUC1) or after (AUC2) injection of CD8$^+$ T cells was calculated to represent quantitatively viral replication over time. Mouse SL9-2 was excluded from AUC analysis owing to early death. Error bars represent s.e.m. *$P < 0.05$, unpaired t-test.

Therefore, directing CTL responses to unmutated viral epitopes is essential to clear latent HIV-1. Owing to bias in antigen presentation or recognition[29], common vaccination strategies will probably re-stimulate immunodominant CTL clones that do not kill infected cells after the reversal of latency. Stimulation of CTL responses with viral peptides circumvents antigen processing and is able to elicit broad-spectrum CTL responses against unmutated regions of viral proteins. Our study suggests that latent HIV-1 can be eliminated in chronically infected patients despite the overwhelming presence of CTL escape variants.

Future directions in therapeutic vaccine design need to focus on boosting broad CTL responses, as also reported elsewhere[30], and/or manipulating immunodominance.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

1. Siliciano, J. D. et al. Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4$^+$ T cells. Nature Med. **9,** 727–728 (2003).
2. Strain, M. C. et al. Heterogeneous clearance rates of long-lived lymphocytes infected with HIV: intrinsic stability predicts lifelong persistence. Proc. Natl Acad. Sci. USA **100,** 4819–4824 (2003).
3. Finzi, D. et al. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. Science **278,** 1295–1300 (1997).
4. Chun, T. W. et al. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. Nature **387,** 183–188 (1997).
5. Richman, D. D. et al. The challenge of finding a cure for HIV infection. Science **323,** 1304–1307 (2009).
6. Archin, N. M. et al. Expression of latent HIV induced by the potent HDAC inhibitor suberoylanilide hydroxamic acid. AIDS Res. Hum. Retroviruses **25,** 207–212 (2009).
7. Contreras, X. et al. Suberoylanilide hydroxamic acid reactivates HIV from latently infected cells. J. Biol. Chem. **284,** 6782–6789 (2009).
8. Archin, N. M. et al. Administration of vorinostat disrupts HIV-1 latency in patients on antiretroviral therapy. Nature **487,** 482–485 (2012).
9. Shan, L. et al. Stimulation of HIV-1-specific cytolytic T lymphocytes facilitates elimination of latent viral reservoir after virus reactivation. Immunity **36,** 491–501 (2012).
10. Walker, B. D. et al. HIV-specific cytotoxic T lymphocytes in seropositive individuals. Nature **328,** 345–348 (1987).
11. Koup, R. A. et al. Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. J. Virol. **68,** 4650–4655 (1994).
12. Borrow, P. et al. Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. Nature Med. **3,** 205–211 (1997).
13. Goonetilleke, N. et al. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. J. Exp. Med. **206,** 1253–1272 (2009).
14. McMichael, A. J., Borrow, P., Tomaras, G. D., Goonetilleke, N. & Haynes, B. F. The immune response during acute HIV-1 infection: clues for vaccine development. Nature Rev. Immunol. **10,** 11–23 (2010).
15. Phillips, R. E. et al. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. Nature **354,** 453–459 (1991).
16. Koup, R. A. Virus escape from CTL recognition. J. Exp. Med. **180,** 779–782 (1994).
17. Goulder, P. J. et al. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. Nature Med. **3,** 212–217 (1997).
18. Goulder, P. J. & Watkins, D. I. HIV and SIV CTL escape: implications for vaccine design. Nature Rev. Immunol. **4,** 630–640 (2004).
19. Henn, M. R. et al. Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. PLoS Pathog. **8,** e1002529 (2012).
20. Liu, M. K. et al. Vertical T cell immunodominance and epitope entropy determine HIV-1 escape. J. Clin. Invest. **123,** 380–393 (2013).
21. Rolland, M., Nickle, D. C. & Mullins, J. I. HIV-1 group M conserved elements vaccine. PLoS Pathog. **3,** e157 (2007).
22. Létourneau, S. et al. Design and pre-clinical evaluation of a universal HIV-1 vaccine. PLoS ONE **2,** e984 (2007).
23. Yu, X. G. et al. Consistent patterns in the development and immunodominance of human immunodeficiency virus type 1 (HIV-1)-specific CD8$^+$ T-cell responses following acute HIV-1 infection. J. Virol. **76,** 8690–8701 (2002).
24. Ho, Y. C. et al. Replication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. Cell **155,** 540–551 (2013).
25. Rongvaux, A. et al. Development and function of human innate immune cells in a humanized mouse model. Nature Biotechnol. **32,** 364–372 (2014).
26. Whitney, J. B. et al. Rapid seeding of the viral reservoir prior to SIV viraemia in rhesus monkeys. Nature **512,** 74–77 (2014).
27. Ananworanich, J. et al. Impact of multi-targeted antiretroviral treatment on gut T cell depletion and HIV reservoir seeding during acute HIV infection. PLoS ONE **7,** e33948 (2012).
28. Goulder, P. J. et al. Evolution and transmission of stable CTL escape mutations in HIV infection. Nature **412,** 334–338 (2001).
29. Le Gall, S., Stamegna, P. & Walker, B. D. Portable flanking sequences modulate CTL epitope processing. J. Clin. Invest. **117,** 3563–3575 (2007).
30. Hansen, S. G. et al. Cytomegalovirus vectors violate CD8$^+$ T cell epitope recognition paradigms. Science **340,** 1237874 (2013).

Author Contributions K.D., L.S., R.A.F. and R.F.S. conceived and designed the research studies; K.D., J.L., H.Z., J.B.M. and L.S. performed the in vitro experiments; K.D., A.R., L.W., C.G., A.J.M., D.M.V., G.D.Y., T.S., P. K. and L.S. performed animal experiments; C.M.D., G.G., H.L.M. and S.G.D. provided patient samples; K.D., M.P., L.W., H.H., J.D.S., S.L.S., L.S. and R.F.S. analysed data; K.D., L.S. and R.F.S. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.F.S. (rsiliciano@jhmi.edu), L.S. (liang.shan@yale.edu) or R.A.F. (richard.flavell@yale.edu).

METHODS

Human subjects. Peripheral blood or bone marrow for the isolation of primary CD4$^+$, CD8$^+$ T cells or CD34$^+$ cells was obtained from 30 HIV-1-infected patients (Extended Data Table 1) and 7 healthy adult volunteers. All patients had been on ART for at least 2 years and had maintained undetectable plasma HIV-1 RNA levels (<50 copies per ml) for at least 1 year before study. Ten AP-treated patients were recruited from the OPTIONS cohort at the University of California San Francisco (UCSF). This study was approved by the Johns Hopkins Internal Review Board and by the UCSF Committee on Human Research. Written informed consent was provided by all study participants. HLA typing for each patient was performed by the Johns Hopkins University Immunogenetics Laboratory.

Sample preparation for deep sequencing. Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood by Ficoll gradient separation. CD4$^+$ T cells were purified from PBMCs by negative selection using CD4$^+$ Isolation Kit II (Miltenyi). Resting CD4$^+$ T cells were then purified from CD4$^+$ T cells by negative selection using CD25, CD69 and HLA-DR microbeads (Miltenyi). Genomic DNA was extracted from 5 million resting CD4$^+$ T cells from each patient using QIAamp DNA Mini Kit (Qiagen). The *gag* gene was amplified from genomic DNA by a two-round nested PCR using these primers: 5′ outer primer (5′-TTGACTAGCGGAG GCTAGAAGG-3′); 3′ outer primer (5′-GATAAAACCTCCAATTCCCCCTAT C-3′); 5′ inner primer (5′-GAGAGATGGGTGCGAGAGCGTC-3′); 3′ inner primer (5′-CTGCTCCTGTATCTAATAGAGC-3′). For each patient, the entire genomic DNA from 5 million resting CD4$^+$ T cells was evenly distributed as a template into 80 PCR reactions. The reactions were performed by using High Fidelity Plantinum Taq Polymerase (Life Technologies) following the manufacturer's instructions. PCR amplicons were purified by gel extraction after gel electrophoresis.

Deep sequencing. For PacBio RS single-molecule sequencing, amplicons were barcoded with a group of 10 bp indexes and then multiple samples were pooled together to generate a SMRTbell sequencing library following the Pacific Biosciences template preparation and sequencing-C2 user guide for 2 kb insert size and using the Pacific Biosciences DNA template preparation kit. For MiSeq sequencing, the pooled amplicon DNA was end repaired, adenylated, and ligated to Illumina TruSeq adaptors and PCR enriched for 10 cycles. The resulting library was then run on a bioanalyser high-sensitivity DNA chip for size and concentration determination. The library was then sequenced on MiSeq for paired-end 250 bp reads. The sequence reads from PacBio and MiSeq were demultiplexed using Fastx-Toolkit.

Data analysis for deep-sequencing results. For the paired MiSeq reads, the two reads were first merged using FLASH[31]. MiSeq and PacBio reads from each individual were then aligned to the reference HIV-1 consensus B Gag sequence using Bowtie2 (ref. 32). A custom program was written using Perl scripts to identify and compute the frequency of all sequence variants that caused non-synonymous amino acid changes in each individual's relevant optimal Gag epitopes (based on reported information in the HIV Molecular Immunology Database, Los Alamos National Laboratory (http://www.hiv.lanl.gov/content/immunology/index.html)) according to their HLA type. For each individual, variants that occurred at a frequency >3% were retained. Additionally, for PacBio reads, sequences with identified premature stop codons were eliminated from the analysed results. For each identified variation, the mutation type regarding CTL recognition was determined by matching with the information in the before-mentioned database. The five mutation types adopted in this paper are: documented escape (no CTL response when patient cells are challenged with the variant peptide); inferred escape (variant is predicted to be an escape mutant by longitudinal study or transmission study, but the reactivity of the variant is not tested experimentally); diminished response (experimental data suggest partial escape as evidenced by decreased CTL response); susceptible form (CTL response is elicited when patient cells are challenged with the variant peptide); and mutation type not determined (no experimental data on CTL recognition of this variant).

ELISpot assays. The ELISpot assays were performed using Human IFN-γ ELISpot PLUS kit (Mabtech) according to methods previously described[33] and the manufacturer's instructions. PBMCs were added at 200,000 cells per well and synthetic peptides were added in a final concentration of 0.1, 1 or 10 μg ml^{-1}. A response was considered positive if it was threefold higher than the mean background (cell only control) and greater than 55 SFC per million cells. The number of specific T cells was calculated by subtracting the mean background values.

Recovery and sequencing of patient viruses from resting CD4$^+$ T cells. Co-culture assays were performed to recover and amplify replication-competent viruses as previously described[34]. The viruses were recovered from 5–10 million resting CD4$^+$ T cells. The concentration of outgrowth viruses was determined by p24 ELISA (PerkinElmer). Total RNA of outgrowth viruses was extracted using TRIzol LS reagent (Life Technologies). Residual DNA was then removed by TURBO DNase (Life Technologies) treatment. First-strand complementary DNA was synthesized using SuperScript III Reverse Transcriptase (Life Technologies) and the *gag* gene was amplified from cDNA using the *gag* outer primer pair mentioned above. The

PCR amplicons were then purified by gel extraction and sequenced by regular Sanger sequencing.

***In vitro* HIV-1 infection.** PBMCs from HIV-1-infected patients and healthy donors were stimulated by adding 0.5 μg ml^{-1} PHA and IL-2 (100 U ml^{-1}) to basal media (RPMI with 10% heat-inactivated fetal bovine serum and antibiotics) for 3 days before isolation of CD4$^+$ T cells. Each patient's activated CD4$^+$ T cells were infected with viruses recovered from the same patient's resting CD4$^+$ T cells. Healthy donors' CD4$^+$ T cells were infected with a laboratory strain virus, BaL. The virus concentration used in infection was equivalent to the p24 concentration of 200 ng ml^{-1}. All infections were performed by centrifugation of target cells with virus at 1,200g for 2 h.

Stimulation of CD8$^+$ T cells. PBMCs from CP-treated patients were cultured in the presence of IL-2 (100 U ml^{-1}) with a mixture of consensus B Gag (or Nef, Rev, Tat, Env) peptides (800 ng ml^{-1} for each) (NIH AIDS Reagent Program), or with individual synthetic peptide (0.5 μg ml^{-1}) (Genemed Synthesis). CD8$^+$ T cells were purified after 6 days of incubation by positive selection using human CD8 microbeads (Miltenyi). To monitor CTL proliferation, PBMCs were stained with CFSE (Life Technologies) before incubation and with the relevant pentamer (Proimmune) after incubation. PBMCs were then stained with CD8-APC (Becton Dickson (BD)) and analysed by flow cytometry using FACS Canto II (BD).

Co-culture of autologous CD4$^+$ and CD8$^+$ T cells. Three hours after infection, CD4$^+$ T cells were mixed with autologous unstimulated or stimulated CD8$^+$ T cells at a 1:1 ratio in basal media at 5 million cells per ml. Two days after co-culture, enfuvirtide (T-20, Roche) was added into the culture at 10 μM to prevent further infection events except if the measurement was p24 ELISA. Three days after co-culture, cells were stained with CD8-APC (BD) first, fixed and permeabilized with Cytoperm/Cytofix (BD pharmingen), then stained for intracellular p24 Gag (PE, Coulter). Cells were analysed by flow cytometry using FACS Canto II (BD). For measurement of viral growth, 5 μl of supernatant was taken from the co-culture at days 0, 3 and 6, and subjected to p24 ELISA. For analysis of cell contact dependence, CD4$^+$ and CD8$^+$ T cells were placed in separate chambers of *trans*-well plates (0.4 μm, Costar).

Generation and infection of patient-derived humanized mice. The previously reported MISTRG mouse in the *Rag2*$^{-/-}$ *Il2rg*$^{-/-}$ 129 × Balb/c (N2) genetic background harbours knock-in replacements of the endogenous mouse *Csf1*, *Csf2*, *Il3* and *Tpo* genes with humanized versions and a BAC transgene encoding human SIRP-α[25]. We generated the Sirpa$^{(KI)}$ mouse, which harbours a knock-in replacement of the endogenous mouse *Sirpa* gene with a humanized version. The Sirpa$^{(KI)}$ mouse will be thoroughly described elsewhere (manuscript in preparation). The improved MIS$^{(KI)}$TRG mouse was generated by breeding Sirpa$^{(KI)}$ mice to MITRG mice. All animal experimentations were performed in compliance with Yale Institutional Animal Care and Use Committee protocols. MIS$^{(KI)}$TRG mice were maintained with continuous treatment with enrofloxacin in the drinking water (Baytril, 0.27 mg ml^{-1}). Patient bone marrow or fetal liver CD34$^+$ cells were isolated by CD34 microbeads selection (miltenyi). Newborn mice (within their first 3 days of life) were sublethally irradiated (X-ray irradiation with X-RAD 320 irradiator, PXi; 1 × 150 cGy) and 100,000 fetal liver or 250,000 patient CD34$^+$ cells in 20 μl of PBS were injected into the liver with a 22-gauge needle (Hamilton Company). Both male and female mice with comparable engraftment levels (percentage of hCD45$^+$, hCD3$^+$ and hCD14$^+$ cells in the blood) were separated randomly into the experimental groups 6–8 weeks after engraftment. Mice engrafted with patient CD34$^+$ cells were infected by retro-orbital injection with HIV-1 (100 ng p24), which was recovered and expanded from the resting CD4$^+$ T cells of the same patient (CD34$^+$ cell donor), as mentioned earlier. Mice engrafted with fetal liver CD34$^+$ cells were infected by intravenous injection with HIV-1 BaL (100 ng p24). Mice with less than 5% human CD45$^+$ cells in the peripheral blood were excluded from the infection study. Mice with more than 70% human CD45$^+$ cells in the peripheral blood were also excluded because they were unhealthy due to human macrophage/monocyte-caused anaemia[25]. Twenty million autologous CD8$^+$ T cells with or without pre-stimulation were injected intravenously 9 or 14 days after infection. Group allocation was blinded. Peripheral blood samples were collected by retro-orbital bleeding at different time points before and after injection of CD8$^+$ T cells. Engraftment of human CD45$^+$ cells as well as lymphoid and myeloid subsets was determined by flow cytometry. Plasma HIV-1 RNA in peripheral blood was measured by one-step reverse transcriptase (Invitrogen) real-time PCR using the following primers and probe, described previously[8]: forward (5′ → 3′) ACATCAAGCAGCCATGCAAAT, reverse (5′ → 3′) TCTGGCCTGGTGCAATAGG, and probe (5′ → 3′) VIC-CTA TCCCATTCTGCAGCTTCCTCATTGATG-TAMRA. Assay sensitivity is 200 RNA copies per ml of plasma. HIV-1 DNA in peripheral blood was also measured by real-time PCR using the same primers and probe mentioned earlier, with assay sensitivity at 5 copies per 100 μl of blood. Total viral DNA in PBMCs was determined by measuring copies of viral DNA per 100 μl blood and blood volume per mouse (80 μl blood per 1 g body weight). To quantitate total viral DNA in tissues, spleens,

livers and lungs of infected mice were collected. For the spleen, single-cell suspensions were treated with ACK lysis buffer. Liver and lung leukocytes were isolated by digesting tissues with 100 U ml^{-1} collagenase IV and 0.02 mg ml^{-1} DNase I (Sigma), followed by density gradient centrifugation.

Statistical analysis. For comparison of HIV-1 variant frequency (Fig. 1b, c) and viral infection in HIV-1 BaL-infected mice (Extended Data Figs 6 and 7), we applied Mann–Whitney tests. For comparison of the inhibitory effect of autologous CTLs (Fig. 2a), we applied a Wilcoxon matched pairs test. For comparison of viral replication in humanized mice (Fig. 4f–h), we applied an unpaired t-test. For all other comparisons, paired t-tests were applied. All tests were calculated by the GraphPad Prism 6 software, and conducted as two-tailed tests with a type I error rate of 5%. No statistical method was used to predetermine sample size.

31. Magoč, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27,** 2957–2963 (2011).
32. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9,** 357–359 (2012).
33. Streeck, H., Frahm, N. & Walker, B. D. The role of IFN-γ Elispot assay in HIV vaccine research. *Nature Protocols* **4,** 461–469 (2009).
34. Siliciano, J. D. & Siliciano, R. F. Enhanced culture assay for detection and quantitation of latently infected, resting CD4$^+$ T-cells carrying replication-competent virus in HIV-1-infected individuals. *Methods Mol. Biol.* **304,** 3–15 (2005).
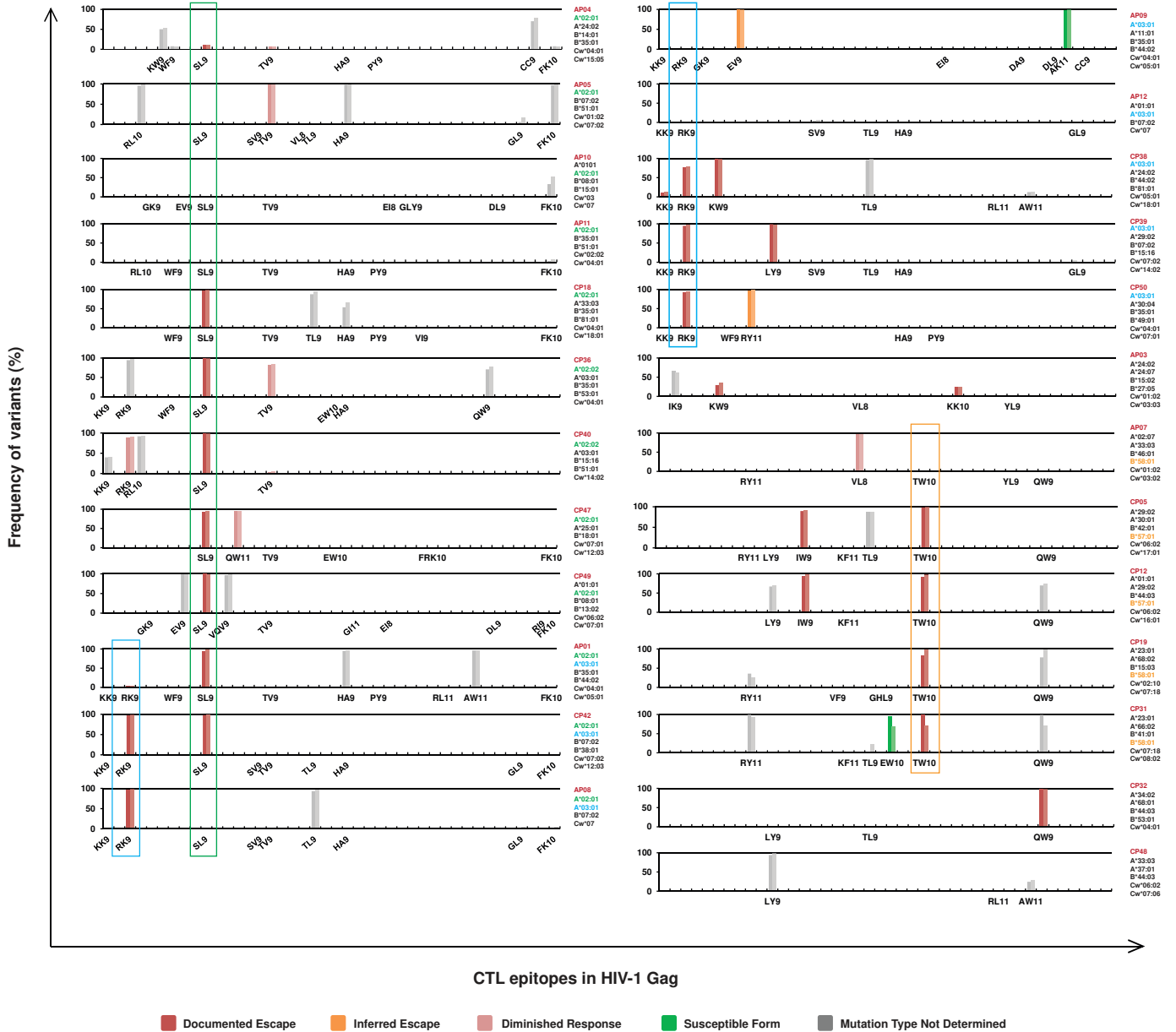
Extended Data Figure 1 | CTL escape variants dominate the latent reservoir of CP-treated HIV-1-positive individuals, but not AP-treated individuals. Frequency of variants in Gag CTL epitopes in proviruses from resting CD4[+] T cells. Results of all 25 patients tested are shown. Only optimal CTL epitopes relevant to each patient's HLA type are listed in linear positional order on the x-axis. Results from both PacBio (left bar) and MiSeq (right bar) sequencing platforms are shown for each epitope. The absence of bars above a listed epitope indicates that only wild-type sequences were detected. For each mutation in a CTL epitope, information regarding the effect of the mutation on CTL recognition from the Los Alamos National Laboratory (LANL) HIV Molecular Immunology Database or from ELISpot assays described in Methods was used to assign the mutation to one of the categories indicated at the bottom. See Methods for definitions of these categories.
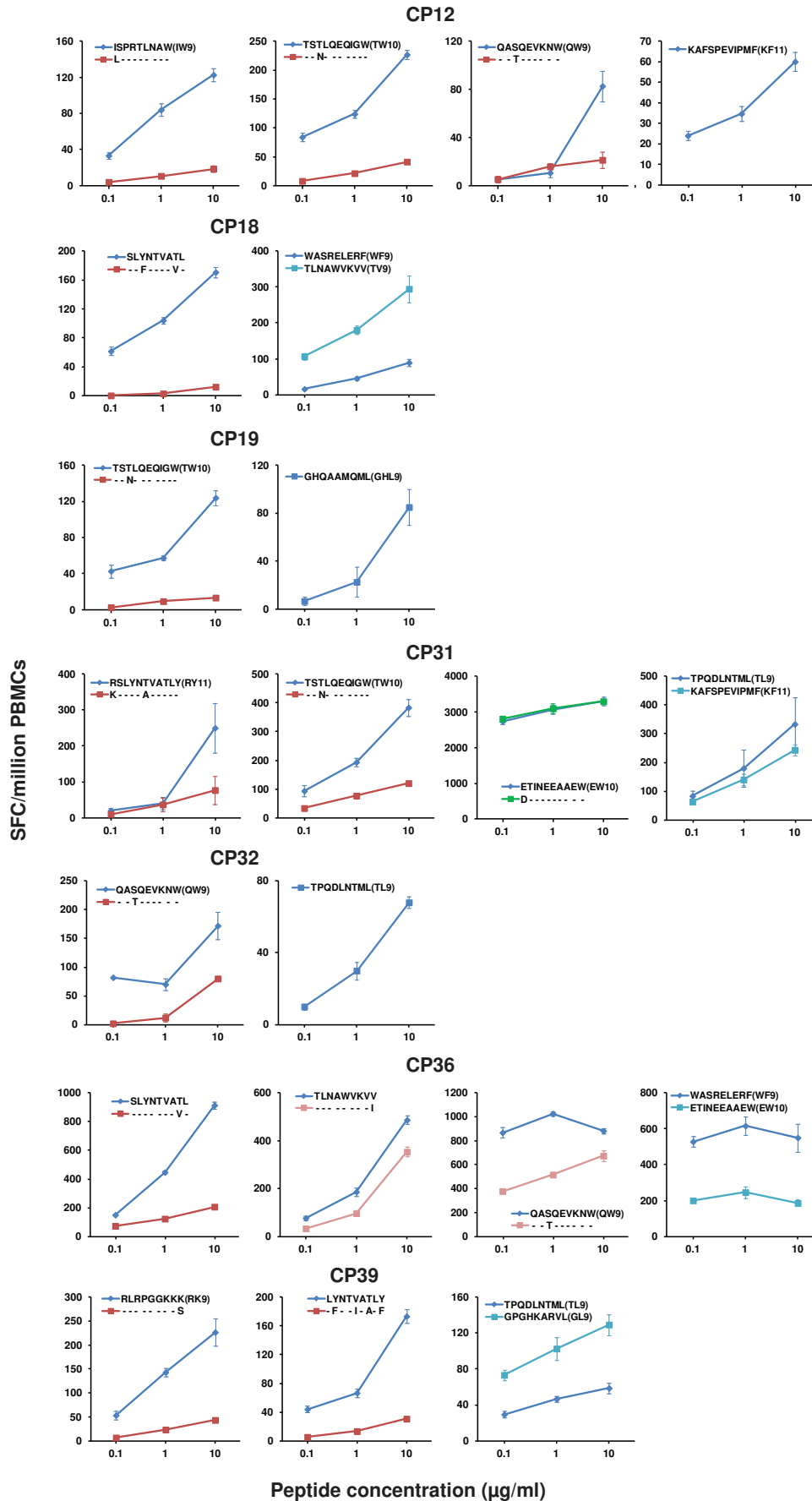
Extended Data Figure 2 | Characterization of CTL responses against HIV-1 Gag epitopes by interferon-γ ELISpot. Results of seven patients tested are shown. The peptides tested are listed for each patient in each graph. Error bars represent s.e.m., $n = 3$.

Extended Data Figure 3 | Partial Gag sequences from proviral DNA and outgrowth virus from resting CD4⁺ T cells from eight CP-treated patients. CTL epitopes with no observed variation are highlighted in blue. Documented escape mutations (red shading), inferred escape mutations (yellow shading), diminished response (pink shading), susceptible form (green shading) or undetermined variations (grey shading) in relevant optimal epitopes are indicated. See Methods for definitions of these types of mutations. This figure supplements Fig. 1e, as a total of nine CP-treated patients were tested.
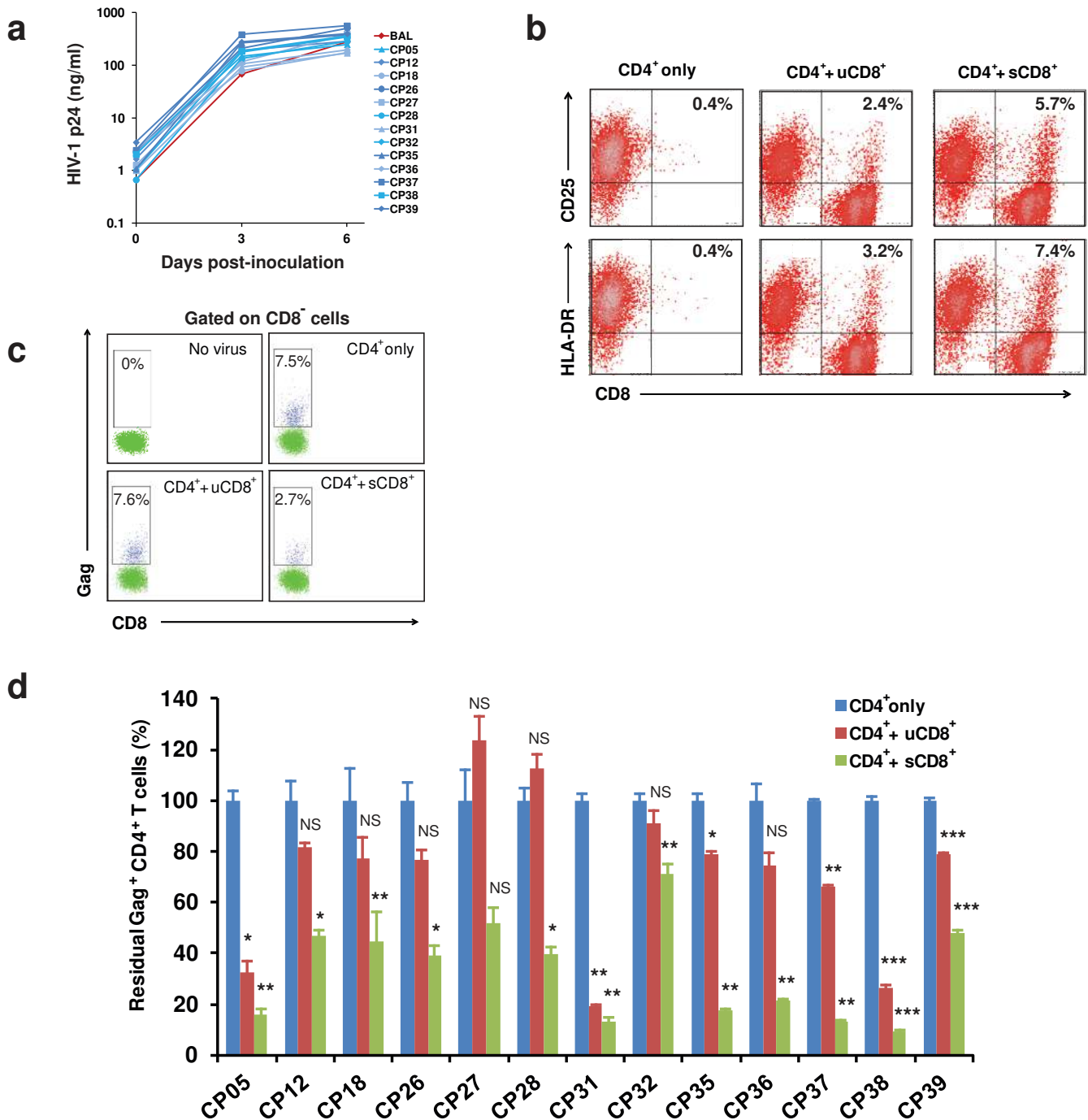
Extended Data Figure 4 | CD8+ T cells pre-stimulated with a mixture of consensus B Gag peptides eliminate autologous CD4+ T cells infected with autologous HIV-1 from resting CD4+ T cells. a, HIV-1 isolated from ART-treated individuals replicates as well as the laboratory strain virus BaL. p24 values represent mean of three replicates. Error bars represent s.e.m., $n = 3$. **b,** CD8+ T cells are not stimulated after co-culture with PHA-activated CD4+ T cells. **c,** A representative flow cytometric analysis of CTL-mediated killing after co-culture of infected CD4+ T cells with autologous CD8+ T cells. CTL activity is measured by the percentage of Gag-positive, CD8-negative cells after 3 days of co-culture relative to cultures without CD8+ T cells. **d,** Pre-stimulated CD8+ T cells eliminate autologous infected CD4+ T cells more efficiently than non-stimulated CD8+ T cells. All results were normalized to the CD4 only control group. Error bars represent s.e.m., $n = 3$. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, NS, not significant ($P > 0.05$), paired t-test.
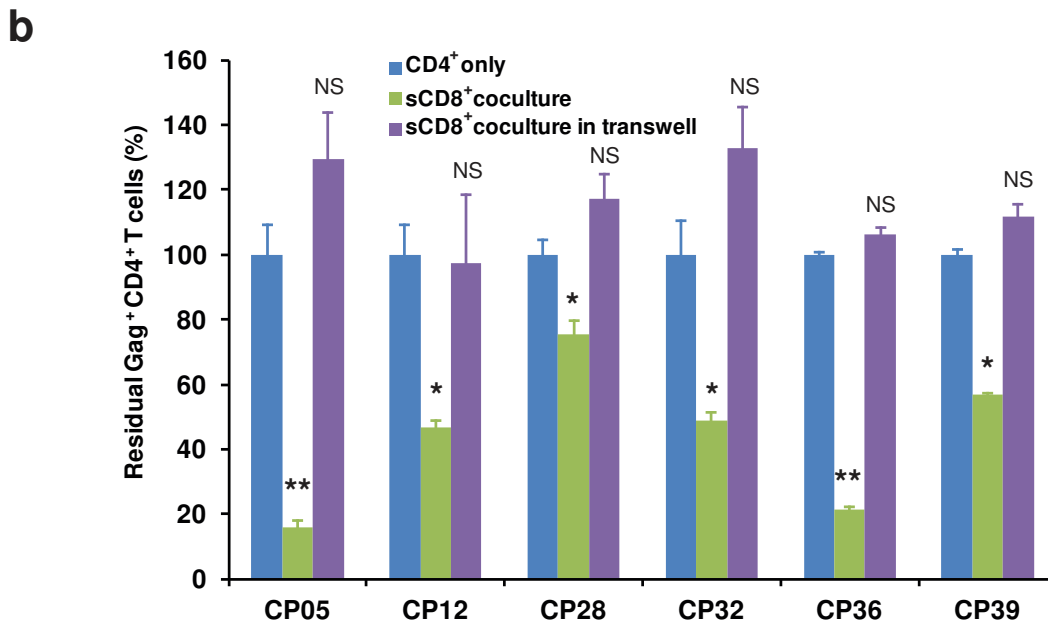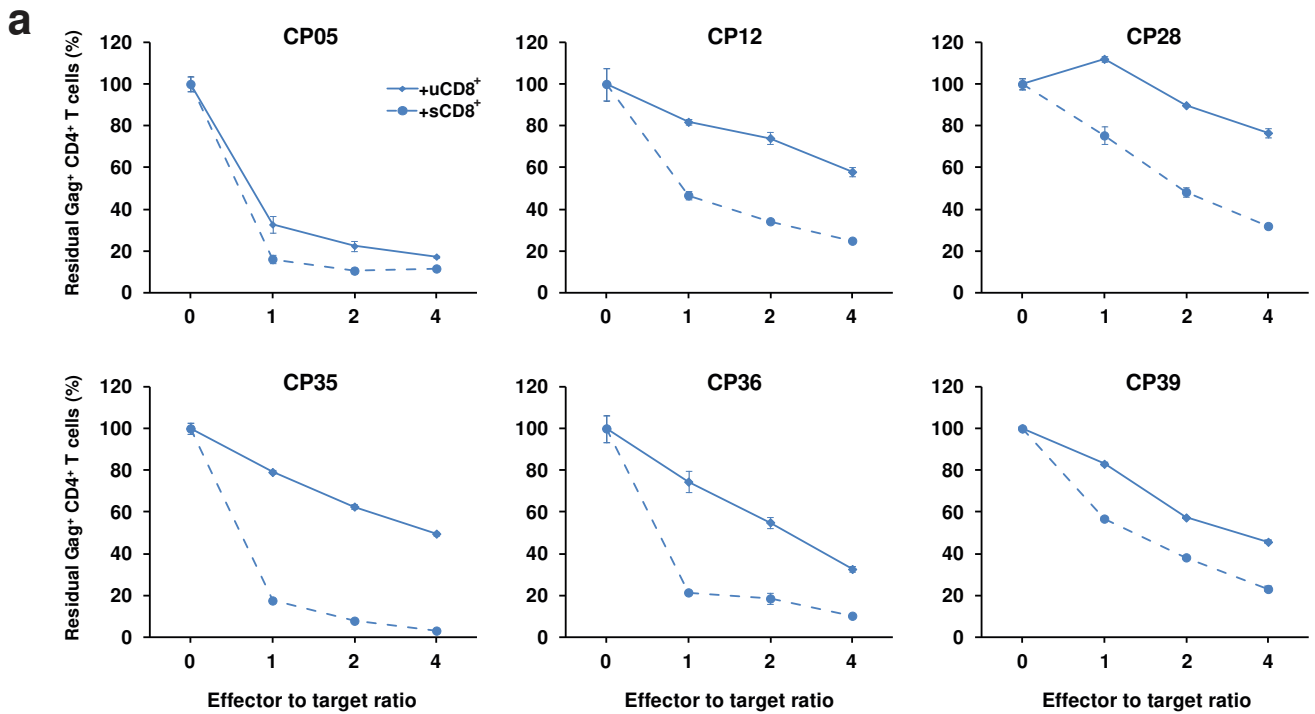
Extended Data Figure 5 | The elimination of infected CD4+ T cells is mediated by direct killing by autologous CD8+ T cells. a, Killing of infected CD4+ T cells is enhanced by increased effector to target ratios for both pre-stimulated and non-stimulated CD8+ T cells. **b,** Killing of the infected CD4+ T cells depends on direct cell–cell contact between CD4+ T cells and CTLs. All results were normalized to the CD4+ only control group. Error bars represent s.e.m., $n = 3$. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, NS, not significant ($P > 0.05$), paired t-test.
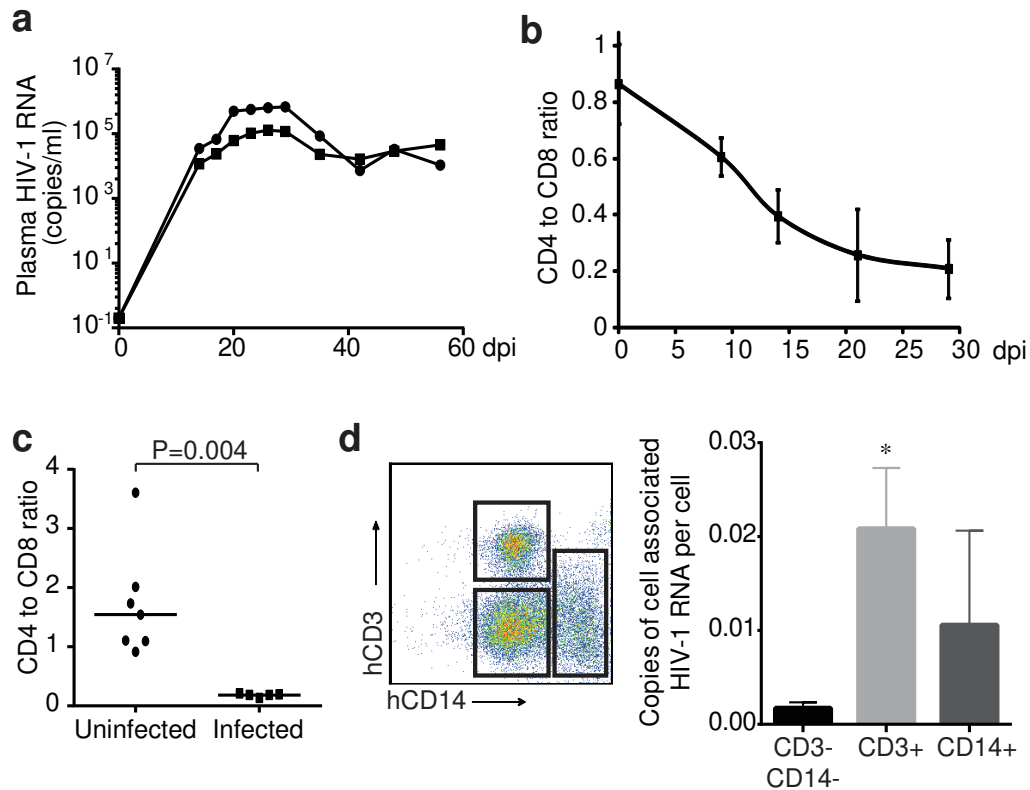
Extended Data Figure 6 | Viral dynamics and depletion of CD4[+] T cells in humanized mice. **a**, Viral dynamics in CP18-infected MIS[(KI)]TRG mice. CP18-derived MIS[(KI)]TRG mice were infected with autologous HIV-1. Plasma HIV-1 RNA levels were measured from day 0 to day 56. **b**, Depletion of CD4[+] T cells in peripheral blood of HIV-1 BaL-infected mice. MIS[(KI)]TRG mice engrafted with fetal liver CD34[+] cells were infected with HIV-1 BaL. The CD4 to CD8 ratio in peripheral blood was measured by fluorescence-activated cell sorting (FACS) from day 0 to day 29 after infection. Error bars represent s.e.m., $n = 5$. **c**, Depletion of CD4[+] T cells in spleen of HIV-1 BaL-infected mice. MIS[(KI)]TRG mice engrafted with fetal liver CD34[+] cells were infected with HIV-1 BaL. The CD4 to CD8 ratio in spleen was measured by FACS 20 days after infection. Medians and P values from Mann–Whitney test are shown. **d**, Detection of cell-associated HIV-1 RNA in T cells and macrophages/monocytes. CD3[+] and CD14[+] human cells from HIV-1-infected MIS[(KI)]TRG mice from spleen and lung were purified by FACS. CD3[−]CD14[−] cells were also collected as controls. Cell-associated HIV-1 RNA was quantified by *gag*-specific quantitative polymerase chain reaction (qPCR). Error bars represent s.e.m., $n = 3$. *$P < 0.05$, unpaired *t*-test.
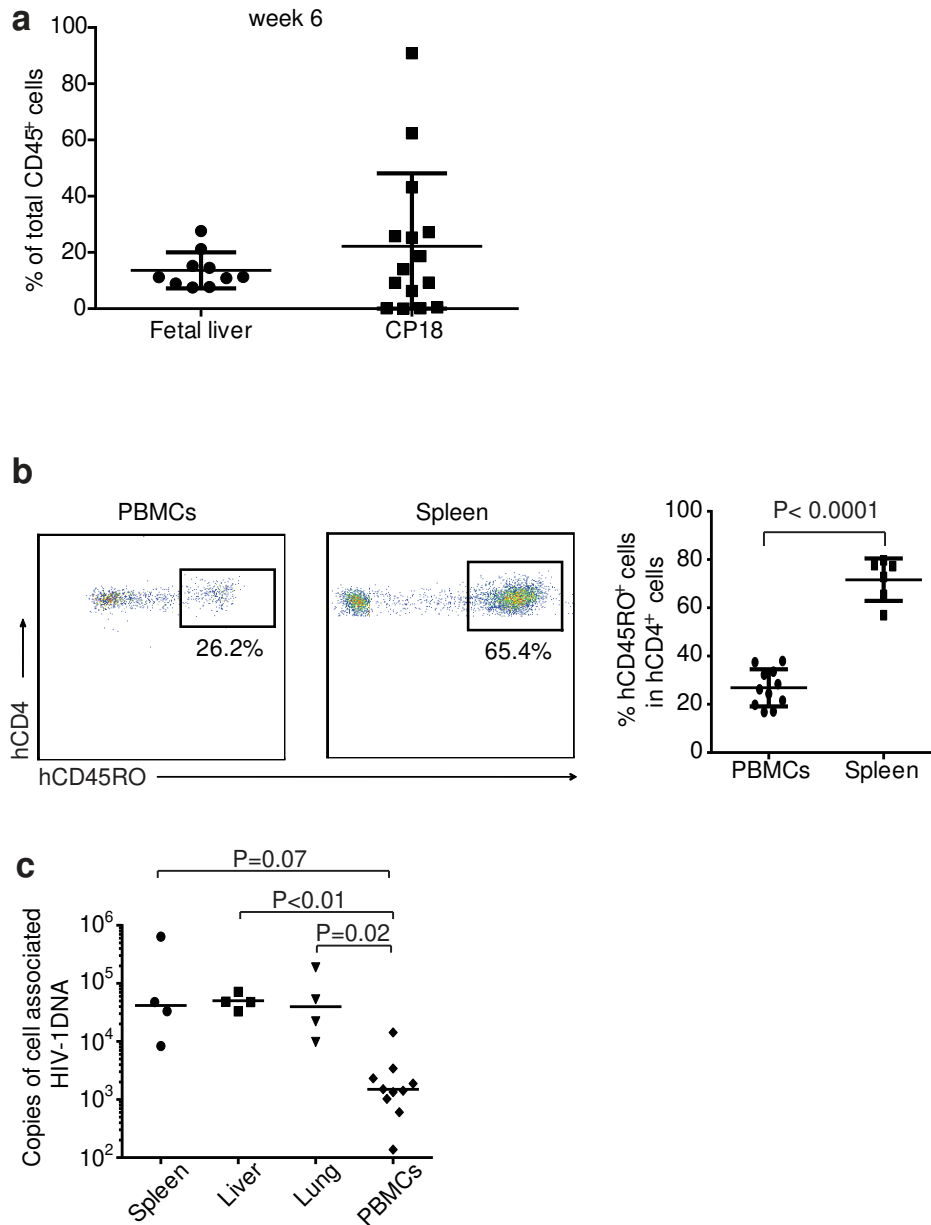
Extended Data Figure 7 | HIV-1 infection occurs in peripheral blood and tissues in humanized mice. a, Engraftment levels of MIS(KI)TRG mice with fetal liver or patient CD34+ cells. **b**, Memory CD4+ T cells are detected in MIS(KI)TRG mice after infection. MIS(KI)TRG mice were infected with HIV-1 BaL. Peripheral blood and indicated tissues from infected mice were collected at 20 days post-infection. Memory CD4+ T cells were determined by CD45RO staining. **c**, Total number of cell-associated HIV-1 DNA in blood and tissues. DNA from peripheral blood or indicated tissues was isolated for the measurement of total amount of cell-associated HIV-1 DNA by real-time PCR. **b**, **c**, Medians and *P* values from Mann–Whitney test are shown.
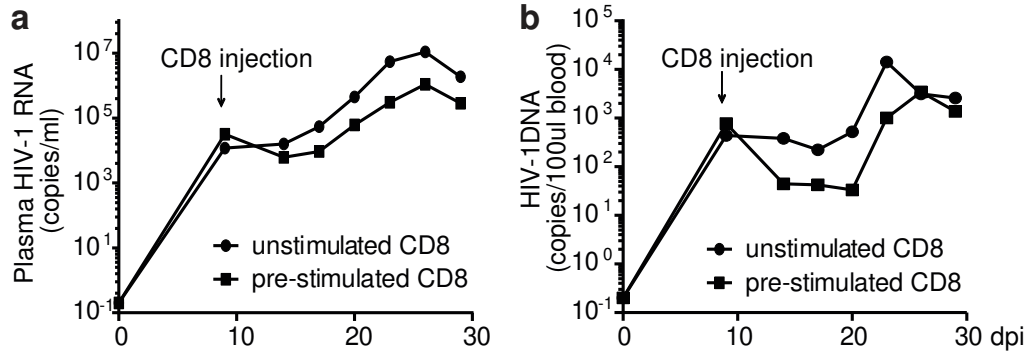
Extended Data Figure 8 | Broad-spectrum CTLs suppress *in vivo* infection of patient-derived humanized mice with autologous latent HIV-1. The generation of patient CP36-derived humanized mice is described in Fig. 4. Mice were infected with autologous viruses at 6 weeks old. CD8⁺ T cells from patient CP36 were pre-stimulated with the mixture of Gag peptides or left untreated for 6 days *in vitro*, and were injected into mice intravenously 9 days after infection. Plasma HIV-1 RNA (**a**) and HIV-1 DNA (**b**) in peripheral blood were measured by real-time PCR.

Extended Data Table 1 | Characteristics of study subjects

Patient	Year of Diagnosis	CD4 count* (cells/μl)	Plasma HIV-1 RNA† (copies/ml)	Time on ART (years)	Treatment start time after infection (days)
AP01	2006	1251	<50	7	64
AP03	2004	595	<50	9	34
AP04	1998	953	<50	15	77
AP05	2002	618	<50	11	39
AP07	2012	592	<50	1.5	67
AP08	2008	780	<50	6	28
AP09	2012	1069	<50	2	39
AP10	2006	513	<50	7	50
AP11	2007	874	<50	6	10
AP12	2007	629	<50	6	15
CP05	2001	500	<50	10	>180
CP12	1997	1074	<50	15	>180
CP18	1998	773	<50	>4	>180
CP19	2006	620	<50	6	>180
CP26	1994	640	<50	16	>180
CP27	1987	784	<50	4	>180
CP28	1998	614	<50	6	>180
CP31	2000	619	<50	10	>180
CP32	1999	780	<50	2	>180
CP35	2002	738	<50	10	>180
CP36	2003	1119	<50	4	>180
CP37	1999	730	<50	12	>180
CP38	1986	870	<50	3	>180
CP39	1996	1152	<50	10	>180
CP40	2002	544	<50	7	>180
CP42	1987	684	<50	16	>180
CP47	1986	792	<50	14	>180
CP48	1998	641	<50	8	>180
CP49	1992	864	<50	5	>180
CP50	2001	964	<50	4	>180

* Patient CD4 count was measured during this study.
† Plasma HIV-1 RNA levels for all patients were <50 copies per ml for at least 1 year before this study.