CD25+FoxP3+ memory CD4 T cells are frequent targets of HIV infection in vivo

Mkunde Chachage,[a]# Georgios Pollakis,[b] Edmund Osei Kuffour,[c] Kerstin Haase,[d] Asli Bauer,[a,e] Yuka Nadai,[e] Lilli Podola,[a,e] Petra Clowes,[a,e] Matthias Schiemann,[f,g] Lynette Henkel,[f] Dieter Hoffmann,[h] Sarah Joseph,[j] Sabin Bhuju,[k] Leonard Maboko,[a] Fred Stephen Sarfo,[l] Kirsten Eberhardt,[m] Michael Hoelscher,[e,n] Torsten Feldt,[c] Elmar Saathoff,[e,n] Christof Geldmacher[e,n]#

NIMR Mbeya Medical Research Centre, Mbeya, Tanzania[a]; Institute of Infection and Global Health, University of Liverpool, Liverpool, UK[b]; Department of Gastroenterology, Hepatology and Infectious Diseases, University Hospital Düsseldorf, Germany[c]; Department of Genome Oriented Bioinformatics, Technische Universität München, Freising, Germany[d]; Division of Infectious Diseases and Tropical Medicine, Medical Center of the University of Munich (LMU), Munich, Germany[e]; Institute for Medical Microbiology, Immunology and Hygiene[f] & Clinical Cooperation Groups "Antigen-specific Immunotherapy" and "Immune-Monitoring" [g] & Institute of Virology[h], Helmholtz Center Munich, Technische Universität München, Munich, Germany; MRC Clinical Trials Unit at UCL, London, UK[j]; Genome Analytics, Helmholtz Centre for Infection Research, Braunschweig, Germany[k]; Kwame Nkrumah University of Science & Technology, Kumasi, Ghana[l]; Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany[m]; German Center for Infection Research (DZIF), partner site Munich, Munich, Germany[n]

22

23 Running Title: Infection of CD25+FoxP3+ memory CD4 T cells by HIV

24

25 #Address correspondence to Christof Geldmacher, geldmacher@lrz.uni-muenchen.de and

26 Mkunde Chachage, mchachage@nimr-mmrc.org.

27

28 Word count abstract: 227

29 Word count text: 6180

30

Abstract

Interleukin 2 (IL2) signaling through the IL2 receptor alpha chain+ (CD25) facilitates HIV replication in vitro and facilitates homeostatic proliferation of CD25+FoxP3+CD4+ T cells. CD25+FoxP3+CD4+ T cells may therefore constitute a suitable subset for HIV infection and plasma virion production.

CD25+FoxP3+CD4+ T cell frequencies, absolute numbers and the expression of CCR5 and cell cycle marker Ki67 were studied in peripheral blood from HIV+ and HIV- study volunteers. Different memory CD4+ T cell subsets were then sorted for quantification of cell-associated HIV-DNA and phylogenetic analyses of the highly variable EnvV1V3 region in comparison to plasma-derived virus sequences.

In HIV+ subjects, 51% (median) of CD25+FoxP3+CD4+ T cells expressed the HIV co-receptor CCR5. Very high frequencies of Ki67+ cells were detected in CD25+FoxP3+ (median, 27.6%) in comparison to memory CD25-FoxP3- memory CD4+ T cells (median, 4.1%, $p<0.0001$). HIV-DNA content was 15-fold higher in CD25+FoxP3+ compared to CD25-FoxP3- memory CD4+ T cells ($p=0.003$). EnvV1V3 sequences derived from CD25+FoxP3+ memory CD4+ T cells did not preferentially cluster with plasma-derived sequences. Quasi-identical cell-plasma-sequence pairs were rare and their proportion further decreased with the estimated HIV infection duration.

These data suggest that specific cellular characteristics of CD25+FoxP3+ memory CD4+ T cell might facilitate efficient HIV infection in vivo and passage of HIV DNA to cell progeny in the absence of active viral replication. Contribution of this cell population to plasma virion production remains unclear.

Importance:

Despite recent advances in the understanding of AIDS virus pathogenesis, it is incompletely understood, which cell subsets support HIV infection and replication *in vivo*. *In vitro*, the IL2 signaling pathway and IL2 dependent cell cycle induction are essential for HIV infection of stimulated T cells. CD25+FoxP3+ memory CD4 T cells - often referred to as regulatory CD4 T cells – depend on IL2 signaling for homeostatic proliferation *in vivo*. Our results show that CD25+FoxP3+ memory CD4+ T cells often express the HIV co-receptor CCR5, are significantly more proliferative and contain more HIV-DNA compared to CD25-FoxP3- memory CD4 T cell subsets. The specific cellular characteristics of CD25+FoxP3+ memory CD4+ T cell probably facilitate efficient HIV infection in vivo and passage of HIV DNA to cell progeny in the absence of active viral replication. However contribution of this cell subset to plasma viremia remains unclear.

Introduction

The Acquired Immunodeficiency Syndrome (AIDS) is caused by HIV infection and is characterized by the failure of the immune system to control diverse opportunistic infections facilitated by the progressive loss of CD4 T cells. The rate of CD4 T cell depletion correlates with set point levels of HIV-1 viral load in plasma (1) and is critically dependent on ongoing viral replication. Antiretroviral therapy (ART) blocks viral replication, reverses CD4 T cell depletion (2) and reconstitutes immunity to most opportunistic pathogens. Replication of HIV within CD4 T cells significantly contributes to plasma viral load and thus to HIV disease

progression (3). It is well established that intra-cellular HIV DNA load in vivo are influenced by CD4 T cell differentiation (4–6), functional properties of CD4 T cells (7) and pathogen-specificity (8–10) and that T cell activation and proliferation contribute to productive HIV infection of memory CD4 T cells (11–15). Together these results imply that, depending on their biological properties, different CD4 T cell subsets might differ in their susceptibility to HIV infection and their contribution to virion production in vivo. Perhaps the best characterized CD4 T cell subset in this regard are follicular CD4 T helper cells (Tfh), which are essential for germinal center formation and which reside in the periphery of B cell follicles within secondary lymphoid organs (reviewed in (16)). Recent data demonstrate that Tfh cells are a major reservoir for HIV replication in vivo (17, 18) and contribute to persistent SIV virion production even in elite controlling, aviremic macaques (19). In viremic macaques virion production appears to be less restricted anatomically (19) and other cell subsets are likely to contribute.

One such cell subset could be memory CD4 T cells expressing the IL2 receptor alpha chain (CD25). Interception of IL2 signaling, which is required for antigen-specific proliferation and survival of CD4 T cells (reviewed in (20)), almost completely abrogates productive HIV infection in cell cultures stimulated in vitro (13, 21–23). Moreover, expression of CD25 defines a CD4 T cell population that efficiently supports productive HIV infection in lymphoid tissue explants (10, 14). In vivo, CD25 expression is characteristic for CD4 T cells (24–26) co-expressing the transcription factor forkhead box P3 (FoxP3) often referred to as regulatory T cells (Tregs). CD25+FoxP3+ CD4 T cells can suppress the activation, proliferation and effector functions of a wide range of immune cells, including CD4 and CD8 T cells (reviewed in (27)), activities shown essential for the maintenance of self-tolerance, but which can also impede the

clearance of chronic infections (28, 29). The vast majority (>80%) of circulating CD25+FoxP3+ CD4 T cells express the memory marker CD45RO (30, 31) and high frequencies of these cells co-express the cell cycle marker Ki67 in peripheral blood (10-20%) and even more so in secondary lymphoid tissue (40-80%) (30, 32) indicating high levels of in vivo proliferation. Doubling time of memory CD25+FoxP3+ CD4 T cells in humans is only 8 days, which is 3-fold and 25-fold less than that of memory and naïve CD4 T cells, respectively (33). These specific cell characteristics and the proposed mechanism of constant IL2 dependent homeostatic replenishment of this cell subset (33, 34) support the hypothesis that CD25+FoxP3+ CD4 T cells are particularly susceptible to HIV infection in vivo and may contribute to plasma virus production in viremic HIV progressors - potentially driven by IL-2 secreted by auto-antigen-specific T cells (35).

To address this hypothesis, we analyzed peripheral blood of HIV-positive and HIV-negative individuals for CD25+FoxP3+ CD4 T cell numbers and frequencies, expression of HIV co-receptor CCR5 and the cell proliferation marker Ki67 in relation to HIV infection. We have also assessed the levels of cell associated viral DNA and the phylogenetic relationship between cell and plasma derived HIV envelope sequences relative to other memory CD4 T cell subsets. Confirming previous reports (36), our data show that high proportions of circulating CD25+FoxP3+ CD4 T cells express the HIV co-receptor CCR5. Furthermore, memory CD25+FoxP3+ CD4 T cells from HIV+ subjects contained high frequencies of Ki67+ cells, and higher levels of HIV DNA and compared to memory CD4 T cells that were CD25-FoxP3-. However, phylogenetic comparison of the highly variable HIV Env_V1V3 region between plasma and cell-derived virus sequences did not allow definite conclusions about the cellular

origin of plasma virions, because sequences from both compartments behaved similar and intermingled with no evidence of compartmentalization. Instead, we observed that the phylogenetic distance between plasma and memory cell-derived viral sequences increases with the duration of HIV infection, with simultaneous decrease in the proportion of detectable quasi-identical cell-plasma-sequence pairs.

Journal of Virology

Materials and Methods

Cohorts, Study volunteers and blood processing. WHIS cohort: 361 adult volunteers were enrolled into a prospective cohort (WHIS) that studies the interaction between HIV-1 and Helminth infection in the Mbeya region in South West Tanzania. The WHIS cohort study is described in detail elsewhere (37). HIV status was determined using HIV 1/2 STAT-PAK, (Chem-bio Diagnostics Systems) and positive results were confirmed using ELISA (Bio-Rad). Discrepancies between HIV 1/2 STAT-PAK and ELISA were resolved by Western Blot (MPD HIV Blot 2.2, MP Biomedicals). 40ml of venous blood were drawn from each participant using anticoagulant tubes (CPDA, EDTA; BD Vacutainer) Absolute CD4 T cell counts were determined in anti-coagulated whole blood using the BD Multitest IMK kit (BD) according to manufacturer instructions. Blood samples were processed within less than 6 hours of the blood draw. Frequencies of CD25+FoxP3+ CD4 T cells and surface CCR5 expression were determined in fresh, anticoagulated whole blood as described below. The absolute numbers of CD25+FoxP3+ CD4 T cells in the peripheral blood was calculated from the total CD4 T cell counts and the percentage CD25+FoxP3+ CD4 T cells. Peripheral Blood Mononuclear Cells (PBMC) were isolated using the Ficoll centrifugation method and Leucosep Tubes (Greiner Bio one) according to Standard Protocols. **HHECO and HISIS cohort:** PBMCs from 28 HIV-positive blood donors who were recruited from a previously described cohort (HHECO) at the Komfo Anokye Teaching Hospital in Kumasi, Ghana (38, 39) and PBMCs from the previously described HISIS cohort (40) were also isolated by centrifugation of heparinized venous blood on a Ficoll/Hypaque (Biocoll Seperating Solution, Biochrom AG, Berlin, Germany) density gradient, prior to cryopreservation.

Ethics Statement. Ethical approvals for the WHIS and HISIS cohorts were obtained from the Mbeya Regional and the National Ethics committee of the Tanzanian National Institute for Medical Research (NIMR)/Ministry of Health in Dar es Salaam and from the Ethics committee of the University of Munich. HHECO study was approved by the appropriate ethics committees of the Kwame Nkrumah University of Science and Technology (Ghana) and of the medical association in Hamburg (Germany) (38, 39). Signed informed consent was obtained from all participants.

Characterization of CD25+FoxP3+ CD4 T cells in fresh whole blood. Fresh anticoagulated whole blood samples from the WHIS cohort were incubated for 30 minutes using the following fluorochrome labeled monoclonal **antibodies for cell surface staining** (mABs);CD3-Pacific Blue (BD), CD4 Per-CP Cy5.5 (eBioscience), CD25 PE-Cy7 (eBioscience), and CCR5 APC-Cy7 (BD). Red blood cells in samples were then lysed by incubating and washing samples twice for 10 minutes with 1X cell lysis solution (BD). Intracellular FoxP3 was detected with FoxP3 Alexa Fluor 647 (eBioscience) according to manufacturer's instructions. Cells were finally fixed with 2% paraformaldehyde prior to acquisition. Acquisition was performed on FACS CANTO II (BD). Compensation was conducted with antibody capture beads (BD) stained separately with the individual antibodies used in the test samples. Flow cytometry data was analyzed using FlowJo (version 9.5.3; Tree Star Inc).

Characterization of memory CD25+FoxP3+ CD4 T cells. Cell surface markers of immune regulation and cell proliferation/cell turnover were stained on cryopreserved PBMCs of

individuals from the HHECO cohort using anti-CD3 PerCP, anti-CD4 Pacific Blue, anti-CD45RA Alexa Flour 700, and anti-CD25 PE-Cy7 (BD Biosciences, Germany). The stained cells were later fixated and permeabilized (FoxP3 Staining Buffer Set, eBioscience) for intracellular staining using anti-FoxP3-PE (Biolegend, Germany) and anti-Ki67-Alexa-Flour-647 (BD Biosciences, Germany). Flow cytometric data was acquired with the LSRII flow cytometer (BD Biosciences, Germany). Compensation was conducted with antibody capture beads (BD CompBeads Set Anti-Mouse Ig, κ, BD Biosciences, Germany), stained separately with the individual flourochrome conjugated monoclonal antibodies used in all samples. Flow cytometry measurements were analyzed using FlowJo® version 9.6.2 (Tree Star, San Carlos, USA).

Cell sorting. Cryopreserved PBMCs from HIV+ WHIS (n=15) and HISIS (n=6) participants were thawed and washed twice in pre-warmed (37°C) complete media (RPMI plus 10% heat inactivated Fetal Bovine Serum (GIBCO) that was supplemented with Benzonase (5U/ml, Novagen). Surface staining was performed with CD3-Pacific Blue, CD4 Per-CP Cy5.5, CD25 PeCy7 and CD45RO PE (BD) for 30 minutes in the dark at RT; intracellular staining was performed with FoxP3 Alexa Fluor 647 (eBioscience) and Helios FITC (BioLegend) according to the CD25+FoxP3+ CD4 T cells staining protocol mentioned above. Cell sorts were performed on a FACSAria cell sorter (BD) after gating on CD3+CD4+CD45RO+ cells into "regulatory T cells populations" (CD25+FoxP3+Helios+ and CD25+FoxP3+Helios-) and memory populations (CD25-FoxP3-Helios+ and CD25-FoxP3-Helios-) as shown in Fig 4A. Between 293 and 750,000 fixed CD4 T cells from each of the four different populations were collected, depending on the number of PBMCs available from each individual. Cells were collected on FACS buffer consisting of PBS mixed with 0.5% Bovine Serum Albumin (BSA, Sigma), 2mM EDTA and

0.2% Sodium Azide at pH 7.45. Median of fixed cell count number collected for each population were as follows: CD25+FoxP3+Helios+ (Median: 9017 and IQR: 3931-14412); CD25+FoxP3+Helios-: (Median: 4381 and IQR: 1579-9799); CD25-FoxP3-Helios+ (Median: 2646 and IQR: 1336-5644) and CD25-FoxP3-Helios- (Median: 185000 and IQR: 79000-315000). Sorted Cells were then centrifuged at 13000 rpm for 3 minutes and the supernatant removed. Cell pellet was stored at −80°C until further analysis.

Quantification of cell-associated HIV gag DNA. Quantification of cell associated HIV gag DNA was performed as previously described (8) with minor modifications. Sorted CD4 T cell subsets were lysed in 30 µl of 0.1 mg/ml proteinase K (Roche) containing 10mM, pH8 Tris-Cl (Sigma) for 1 hour at 56°C followed by Proteinase K inactivation step for 10 min at 95°C. Cell lysates were then used to quantify cell associated HIV DNA was quantified by qPCR as previously described with some modifications (10). Briefly, Gags primers and probe used were as follows: 783gag, forward, 5′-GAG AGA GAT GGG TGC GAG AGC GTC-3′ (Tm>60), 895gag, reverse, 5′-CTK TCC AGC TCC CTG CTT GCC CA-3′ (Tm>60); FAM-labeled probe 844gagPr, 5′-ATT HGB TTA AGG CCA GGG GGA ARG AAA MAA T-3′ and had been designed to optimally cover subtypes A, C and D prevalent in Mbeya Region (10). To quantify the cell number in each reaction mix, the human prion gene copy number was also assessed by qPCR. Prion primers and probe sequences were as follows: Prion forward: 5′TGC TGG GAA GTG CCA TGA G-3′; Prion reverse: 5′CGG TGC ATG TTT TCA CGA TAG-3′; probe 5′FAM-CAT CAT ACA TTT CGG CAG TGA CTA TGA GGA CC-TAMRA (41). 5 µl of lysate was used in a total reaction volume of 25 µl containing 0.8 µM Gag primers or 0.4 µM Prion primers, 0.4 µM probe, a 0.2 mM concentration of each deoxynucleoside triphosphate, 3.5 mM $MgCl_2$

and 0.65 U platinum *Taq* in the supplied buffer. Standard curves were generated using HIV-1 gag gene (provided by Brenna Hill, Vaccine Research Center, NIH, Bethesda) and prion gene encoding plasmids. Real time PCR was performed in a Bio-Rad cycler CFX96 (Bio-Rad): 5-min at 95°C, followed by 45 cycles (15 seconds at 95°C and 1 minute at 60°C). To assure comparability of the results, cell-associated gag DNA from the 4 different memory CD4 T cell subsets, which were sorted from one patient, were always quantified simultaneously. Cell-associated gag DNA in memory CD25+FoxP3+ CD4 T cells and CD25-/FoxP3- memory CD4 T cells independent of Helios Expression was calculated as follows: ∑Gag DNA load (Helios+)+(Helios-) divided by ∑sorted cells in 5 µl lysate (Helios+)+(Helios-).

Amplification and phylogenetic comparison of HIV Envelope sequences from plasma and sorted cell populations. A highly variable Envelope region spanning the V1 to V3 region (EnvV1V3, Hxb 6559 – 7320) was amplified using a nested PCR strategy from 10 µl of lysed sorted cells (described above) or from plasma virus cDNA. HIV RNA was extracted with Sample Preparation Systems RNA on the automatic extractor m24sp instrument (Abbott molecular, USA) following the manufacturer's instructions. The HIV cDNA was synthesized from 3 µl of extracted RNA using the reverse primer ACD_Env7521R 5`ATGGGAGGGGCATAYATTGC and the Superscript III reverse transcriptase (Life technologies, Darmstadt) according to manufacturer instructions. Newly designed PCR primer pairs were optimized for detection of subtypes A, C and D were used to amplify the EnvV1V3 region. The 1st round PCR was performed with 10 µl of template in a 50 µl reaction (0.5 µl (=5U)) Platinum Taq (Life technologies, Darmstadt), 2.0 mM primers; ACD_Env6420F

240 5'CATAATGTCTGGGCYACACATGC and ACD_Env7521R 5`ATGGGAGGGGC
241 ATAYATTGC, 3.5 mM $MgCl_2$, 4 µl of dNTPs at 95°C for 10 min followed by 45 cycles (94°C-
242 30 seconds, 55°C-30 seconds, 72°C-90 seconds) and 7 minutes at 72°C. The 2nd round PCR was
243 performed with 2 µl of first round PCR product in a 50 µl reaction (0.25ul (2.5U) AmpliTaq
244 Gold (Life technologies, Darmstadt), 2.0 mM ACD_Env6559F
245 5`GGGAYSAAAGCCTAAARCCATGTG and ACD_Env7320R GTTGTAATTTCTRRR
246 TCCCCTCC, 2.0 mM $MgCl_2$, 4 µl of dNTPs at 95°C for 10 min followed by 45 cycles (94°C-30
247 seconds, 53°C-30 seconds, 72°C-90 seconds) and 7 minutes at 72°C. The second round PCR
248 products were extracted from agarose gel and then cloned using the TOPO-TA cloning Kit for
249 sequencing (Life technologies, Darmstadt) including the pre-cut vector pCR4.1 and One Shot®
250 chemically competent E. coli according to manufacturer instructions. EnvV1V3 sequences from
251 11-23 clones/population/subject were then sequenced unidirectional using Mnrev primers at
252 Eurofins Genomics (Ebersberg, Germany). In total, 384 EnvV1V3 sequences from 6 subjects
253 were analyzed.

254 To assess the error rate of the applied nested PCR strategy, the positive control template
255 (Du422, clone 1 (SVPC5)) (42) was endpoint diluted using a 10-fold dilution series and
256 amplified as described above. The EnvV1V3 product from the last detectable dilution step was
257 then cloned as described above. Sequences from 21 clones were analyzed and compared to the
258 original Du422 template sequence.

259

260 **Phylogenetic analyses.** Nucleotide sequences were aligned with respect to the predicted amino
261 acid sequence of the reference alignment extracted from the Los Alamos HIV database

(http://www.hiv.lanl.gov/content/sequence/NEWALIGN/align.html) as previously described (43) Evolutionary analyses were conducted in MEGA6 (44). The evolutionary history is inferred by using the Maximum Likelihood method based on the General Time Reversible substitution model (GTR+G) (45) and is rooted on previous outbreaks. Upon each analysis the tree with the highest log likelihood is shown. The percentage of trees in which the associated taxa clustered together is presented next to the branches. Initial tree(s) for the heuristic search are obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value.

Next Generation Sequencing (NGS). Library preparation from EnvV1V3PCR second round products was done using TruSeq DNA PCR-Free Sample Preparation Kit (Illumina Inc., San Diego, CA, USA) with 550 bp as insert size following the manufacturer's instruction. The libraries were controlled with Agilent Bioanalyzer HS Chip (Agilent Technologies) and sequenced using MiSeq Desktop Sequencer (Illumina Inc.) using MiSeq Reagent Kits v3 (Illumina Inc.). The sequencing was done to 250 cycles in both directions. The produced reads were processed through a quality control pipeline that removed all reads containing unresolved positions or had a mean quality below 20. Furthermore, poly-A tails and low quality read ends were trimmed away. All reads that had a length below 30nt after trimming were also excluded from further analysis. An initial mapping was created for each sample, by placing the reads onto the HIV HXB2 reference sequence (GenBank identifier K03455.1 (46)) using segemehl (version 0.1.6) (47). The difference parameter was set to two in order to increase the sensitivity given the origin of the sequences being a highly variable viral genome. Using an adapted samtools (version

285 0.1.19) (48) pipeline, we created a consensus sequence for each sample from the initial mapping
286 to use as individual reference for a second round of alignments. This was necessary as the
287 official HIV reference sequence is very diverse from our set of reads, thus the initial mapping
288 was only able to place an unsatisfyingly low number of reads onto this sequence. The second
289 individual mapping was able to use a higher number of reads and create sufficient alignments
290 which were used as input for the quasispecies reconstruction tool QuasiRecomb (49). It uses an
291 expectation maximization algorithm to not only reconstruct the single sequences present in the
292 viral population, but to also assign their relative proportions.

293

294 **Statistical analysis.** Data analyses were performed using Prism version 4.0 software (GraphPad,
295 Inc.). Comparisons of two groups were performed using the Mann-Whitney test. Comparisons of
296 paired groups were performed using the Wilcoxon matched pairs test. For correlation analyses
297 the Spearman r, Pearson two-tailed statistical test or linear regression were used. Differences
298 were considered significant at P values of <0.05. Tests used for statistical analysis are described
299 in the figure legends.

300

301 **Results**

302 **Study subjects**

303 Table 1 provides an overview of the subjects included in this study. A total of 258 HIV
304 negative and 103 HIV positive adults (Mean age, 34.3 years) from the WHIS cohort (37) were
305 included in this study of which 217 (60%) of these were female. The vast majority of HIV+

306 subjects from the WHIS cohort were treatment naïve (97%) with a median $CD4^+$ T cell count of
307 396.3 cells/µl and median Log_{10} plasma viral load was 4.7 copies/ml. 28 subjects from the
308 previously described HHECO cohort were included for the in-depth characterization of memory
309 CD25+FoxP3+ CD4 T cells ((38, 39); also described in Table 1). PBMCs from 6 viremic HIV+
310 subjects from the HISIS cohort (40) were used for the characterization of HIV infection within
311 different memory T cell subsets.

312

313 **Correlation between CD4 T cells and CD25+FoxP3+ CD4 T cell counts in HIV infected**
314 **subjects**

315 We first determined and compared the frequency and absolute numbers of
316 CD25+FoxP3+ CD4+ T cells in fresh anticoagulated peripheral blood of HIV+ (treatment naïve,
317 n=100) and HIV- subjects (n=258) from the WHIS cohort. A representative dot plot and gating
318 of CD25+FoxP3+CD4+ T cell is shown in Fig 1A. In HIV+ compared to HIV-neg. individuals,
319 CD25+FoxP3+CD4+ T cell frequencies were moderately increased (Fig 1B, HIV+: median,
320 2.5%; IQR, 1.5-4.5% versus HIV-: median, 2.1%; IQR, 1.5-2.9; p= 0.03;), but absolute numbers
321 of CD25+FoxP3+CD4+ T cell were significantly decreased with median counts of 10.16 cells/µl
322 (IQR, 4.88- 18.57 cells/µl) in HIV+ subjects and 17.75 cells/µl (IQR, 11.06- 24.56 cells/µl) in
323 HIV- subjects ($p<0.0001$, Fig 1C). Within HIV+ subjects there was a positive correlation
324 between CD25+FoxP3+CD4+ T cell and CD4 T cell counts ($p<0.0001$, $r= 0.6152$, Fig 1D).
325 Confirming previous reports (50–54), our data shows that the depletion of CD25+FoxP3+CD4+
326 T cells is closely linked to the loss of CD4 T cells.

327

High frequencies of CD25+FoxP3+ CD4 T cells express HIV-co receptor CCR5 and the cell cycle marker Ki67

In order to determine, whether CD25+FoxP3+CD4+ T cells could potentially support entry of HIV, we assessed the expression of the HIV co-receptor CCR5. Fresh anticoagulated whole blood was used for improved CCR5 staining. A representative plot is shown in Fig 2A. A considerable proportion of CD25+FoxP3+CD4+ T cell expressed CCR5 (median, 53.7%), which was higher than previously observed in total memory CD4 T cells (median, 40%; data not shown). HIV infection was associated with a moderate decrease in the frequency of CCR5+ CD25+FoxP3+CD4+ T cells (Fig 2B; median, 50.9% compared to 54.5%; p= 0.01).

We next studied the cell cycle status of memory CD25+FoxP3+ and CD25-FoxP3- CD4+ T cells in HIV infected subjects and analyzed cellular Ki67 expression using cryopreserved PBMC samples (n=28 from HHECO cohort, Table 1). The representative dots plots for Ki67 staining in memory (CD45RA-) CD25+FoxP3+ and CD25-FoxP3- CD4+ T cells are shown in 3A. HIV infected study participants had very high frequencies of Ki67+ memory CD25+FoxP3+ CD4 T cells (median, 27.6%, Fig 3B) despite the majority of subjects from the HHECO cohort being on ART. Importantly, frequencies of Ki67+ cells detected were 6.7-fold higher in CD25+FoxP3+ compared to CD25-FoxP3- memory CD4+ T cells (median, 4.1%, $p<0.0001$), consistent with high in vivo proliferation of memory CD25+FoxP3+ CD4 T cells. Correlation analysis demonstrated a close association between the proportion of Ki67+CD25+FoxP3+ and Ki67+CD25-FoxP3-memory CD4 T cells (p=0.005, r=0.51, Fig 3C), linked to the level of CD4 T cell depletion in HIV+ subjects (p=0.1, r=(-)0.3, Fig 3D and p=0.02, r=(-)0.4, 3E). Memory CD25+FoxP3+ CD4 T cells could hence potentially support CCR5 mediated viral entry and subsequent steps of the viral life cycle due to their high in vivo proliferation. The correlation

between the frequency of Ki67+ memory T cells and memory CD25+FoxP3+ CD4 T cells and the fact that loss of these cell subsets is closely linked, support the proposed mechanism of constant replenishment of memory CD25+FoxP3+ CD4 T cells from the memory CD4 T cell pool (30) also during HIV infection.

Memory Helios+ and Helios- CD25+FoxP3+ CD4 T cells are frequent targets for HIV infection in vivo

To determine in vivo HIV infection rates of memory CD25+FoxP3+ CD4 T cells, we sorted four different subsets of CD45RO+ memory CD4 T cells on the basis of their Helios, CD25 and FoxP3 expression (Fig 4A) for 22 subjects (WHIS cohort, plus 6 subjects from HISIS cohort, Table 1) and quantified HIV gag DNA within the sorted subsets. Helios is an Ikaros transcriptional factor family member, which is critical for the regulatory function of CD25+FoxP3+ CD4 T cells (55–58) is a negative regulator of IL2 signaling in CD25+FoxP3+ CD4 T cells (59). A large fraction of CD25+FoxP3+ CD4 T cells expressed the memory marker CD45R0 in HIV+ subjects (median, 87.3%; IQR, 71.85%-93.55%) and most of these expressed Helios (median, 76.30%; IQR, 69.18%-84.43%; data not shown), consistent with a regulatory cell function of this subset. In contrast, only a minor fraction of CD25-FoxP3- memory CD4 T expressed Helios (median, 1.65%; IQR, 1.15%-2.75%). HIV gag DNA was detected in >80% of memory CD25+FoxP3+ and CD25-FoxP3- CD4 T cells with a 15-fold higher median gag DNA load in CD25+FoxP3+ compared to CD25-FoxP3- memory CD4 T cells ($\sum$Helios^{+}Helios-, 16072 versus 1074 copies/10^6 cells; p=0.003; Fig 4B). From 16 subjects we also determined the plasma viral load (pVL) and found correlation between log cell associated DNA gag in memory

CD25-FoxP3- memory CD4 T cells and log pVL (p=0.025, r=0.56, data not shown). No such association was detected for memory CD25+FoxP3+ (p=0.1, r=0.39, data not shown).

Fig 4C shows the levels of HIV gag DNA within these memory CD4 T cell subsets further delineated by Helios expression. Compared to the largest sorted memory CD4 T cell population in the blood (FoxP3-CD25-Helios-), which contained a median of 154.4 HIV copies/10^6 cells (IQR, 0-10241 copies/10^6 cells), levels of HIV gag DNA were substantially increased in the other subsets; the FoxP3+CD25+Helios- CD4 T cells (119-fold increased, median, 18407 copies/10^6 cells; IQR, 1556-106067 copies/10^6 cells; p= 0.007), FoxP3-CD25-Helios+ CD4 T cells (104-fold increased, median, 16096 copies/10^6 cells; IQR, 837.9–47903 copies/10^6 cells, p= 0.029) and FoxP3+CD25+Helios+ CD4 T cells (26-fold increased, median, 4106 copies/10^6 cells; IQR, 0-446m copies/10^6 cells; p=0.072). Together these data demonstrate that CD25+FoxP3+ memory CD4 T cells and in particular the small Helios- population, contain high HIV DNA levels in vivo. Likewise, the small CD25-FoxP3- Helios+ memory CD4 T cell population contained substantially increased HIV DNA copies. In comparison, the main CD25-FoxP3- Helios- memory CD4 T cell subset (>90% of memory CD4 T cells in peripheral blood) of which high cell numbers were sorted for all 22 subjects, contained few and surprisingly often undetectable gag DNA copies. Together these data suggest that CD25+FoxP3+ and also CD25-FoxP3-Helios+ memory CD4 T cells are frequent targets for HIV infection. However, the lack of correlation between plasma viral load and Gag DNA loads in CD25+FoxP3+ memory CD4 T cells is inconsistent with the hypothesis of significant plasma virus production by this cell subset.

Phylogenetic sequence analyses of the highly variable EnvV1V3 region in plasma virus and sorted memory CD4 T cell populations

To assess whether memory CD25+FoxP3+ CD4 T cells could potentially contribute to plasma virion production, we compared the highly variable Envelope V1V3 region from cell (CD25+FoxP3+ and CD25-FoxP3- memory CD4 T cell subsets) DNA and plasma virus sequences in seven viremic subjects. The estimated HIV infection duration varied from 9 months (H574), 27-30 months (H605), 1.3 – 3.3 years (6233K12), above 3.2 years for 3806A11, 8710U11 and 9440A11 and above 4.5 years for 8975T11. PCR related sequence background variation was controlled for by using an endpoint diluted molecular clone of the subtype isolate Du422 clone 1. Ten of the 21 Du422 sequences did not contain any nucleotide substitutions compared to the template sequence, seven sequences had one and three sequences had two substitutions. Hence, the PCR protocol introduced only two or less nucleotide substitutions and no insertions or deletions in 95% of the amplicons. We hence considered up to four substitutions between cell- and plasma-derived sequence variants as quasi-identical. EnvV1V3 amplicons containing clones from 6 of the 7 subjects were subjected to Sanger sequencing and clonal sequences were analyzed using Maximum likelihood method (Figs 5A and 5B). In 4 of these 6 subjects (H574, H605, 6233K12, 9440A11) we found quasi-identical cell- and plasma-derived EnvV1V3 sequence pairs (Table 2). For subject H574 (9- months HIV infected) viral sequences were closely related to each other and sequences from all four sorted cell populations were closely related to plasma virus (Fig 5B, Table 2). 11.4% (8 of 70) of cell-derived sequences were quasi-identical to plasma-derived sequence variants. For subject H605 (27 to 30 months infected) the closest sequence was derived from the “dominant” memory CD4 T cell subset (CD25-FoxP3-Helios-, 3 substitutions) and 6.8% (3 of 44) of cell-derived sequences were quasi-

417 identical to plasma-derived sequence variants. For subject 6233K12 (16 to 38 months infected)
418 only 1 of 53 EnvV1V3 sequences was quasi-identical to a plasma-derived sequence variant and
419 was derived from the CD25-FoxP3-Helios+ memory T cell subset. The three subjects (8710U11,
420 8975T11 and 9440A11) were infected for at least 3.2 years and the closest cell-derived sequence
421 to a plasma-derived sequence variant had 32, 54 and 4 substitutions respectively. Hence, we
422 detected a single "quasi-identical pair" between cell- and plasma-derived EnvV1V3 sequences
423 (derived from CD25+Helios+ memory CD4 T cells) only in one of these three subjects. In
424 summary, sequences derived from CD25+FoxP3+ memory CD4 T cells (or those derived from
425 the other sorted memory CD4 T cell subsets) were not preferentially clustering with plasma-
426 derived sequence variants. Quasi-identical cell- and plasma-derived EnvV1V3 sequence pairs
427 were generally infrequent and their proportion decreased with HIV infection duration (Fig 5C,
428 $p=0.03$, $r= (-)0.85$) as the nucleotide distances between cell- and plasma-derived sequences (Fig
429 5D, $p=0.02$, $r^2=0.84$) and also between individual plasma-derived sequences (Fig 5E, $p=0.02$,
430 $r^2=0.95$) increased. To ascertain the relatedness of the plasma sequences and the sequences
431 isolated from the cell fractions we estimated the nucleotide variation within each fraction. The
432 estimation was performed using the neighbor-joining model with the Kimura-2 parameter
433 method. The sequence diversity analyses showed that the sequence diversity in plasma was not
434 different from the estimated diversity between the plasma and the cell-derived sequences (data
435 not shown).

436 We also analyzed plasma- and cell-derived EnvV1V3 amplicons from two HIV+ subjects
437 (3806A11 and 9440A11) infected for more than 3.2 years using next generation sequencing to
438 detect "rare" quasi-identical sequence pairs we might have missed in the previous analyses.
439 Between 780 and 10000 EnvV1V3 sequences were first reconstructed using QuasiRecomb (49).

The 50 most frequent sequences/population were aligned and sequences compared (Fig 6). The closest cell-associated and plasma sequences were 6 and 14 nucleotide substitutions apart for 3806A11 and 9440A11, respectively, inconsistent with a major contribution of the sorted peripheral memory CD4+ T cell subsets to plasma virus production. Blast searching all plasma sequence variants against the 150 highest frequency cell-derived variants (per sorted cell subset) identified the closest pairs as 4 (3806A11, CD25-FoxP3-Helios+) and 10 (9440A11, CD25+Helios+) nucleotides apart.

Discussion

HIV plasma viremia predicts the rate of HIV disease progression (1, 60) and depends on active HIV viral replication in CD4+ cells. Memory CD4 T cells are most probably the primary substrate for virus replication (11, 61–63). HIV infection rates differ substantially between different CD4 T cell subsets (4–6, 64). Recent data show that follicular T Helper (Tfh) cells are a prime target for virus replication and contribute to virion production even in elite controlling rhesus macaques (19) and most probably to plasma viremia (17). To what extent other CD4+ cell subsets contribute to plasma virus production in viremic progressors is unclear. In various in vitro infection models, HIV replication is associated with IL2 signaling and CD25 expression on stimulated CD4 T cells (10, 13, 14, 21–23). Because IL2 is important for the homeostatic proliferation of the CD25+FoxP3+ CD4 T cells (35, 65), and because of high in vivo proliferation rates of this subset (32), we hypothesized that CD25+FoxP3+ CD4 T cells constitute a prime target for HIV infection and may contribute to plasma virion production in vivo.

462 Consistent with a previous report, we show that a large fraction of CD25+FoxP3+ CD4 T
463 cells, express the HIV co-receptor CCR5 (35), potentially supporting viral entry. Although
464 frequencies of CD25+FoxP3+ CD4 T cells were slightly elevated in viremic, HIV+ subjects,
465 absolute cell numbers of this subset were significantly depleted, which confirms previously
466 published data (50, 52, 67). A greater proportion of CD25+FoxP3+ memory CD4 T cells from
467 HIV+ subjects expressed Ki67+ with almost one third of these cells "cycling" at any given time.
468 This pattern – depleted cell counts despite increased fractions of Ki67+, "cycling" cells
469 demonstrates that homeostasis of CD25+FoxP3+ CD4 T cells is heavily perturbed by HIV
470 infection. Furthermore, expression of CCR5 and high proportions of cycling cells within
471 CD25+FoxP3+ CD4 T cells should support both cell entry and reverse transcription of HIV,
472 which is supported by the increased HIV DNA loads observed in memory CD25+FoxP3+ CD4 T
473 cells observed in this study (12, 36). Other reports show discrepant results regarding in vivo
474 levels of HIV DNA in "regulatory" CD4 T cells - typically defined by $CD25^{high}$ phenotype,
475 instead of the definition using co-expression of CD25 and FoxP3 that we used (66, 68, 69). Tran
476 et al. observed a higher infection rate in $CD25^{high}$ than CD25 negative CD4 T cells (70), but did
477 not exclude naïve CD4 T cells – which are not susceptible to CCR5-topic strains which
478 predominate throughout most of the infection course. Of note, high in vivo proliferation of
479 memory CD25+FoxP3+ CD4 T cells could also potentially pass on proviral HIV DNA to the cell
480 progeny in the absence of productive HIV infection during ART. Previous studies reported that
481 $CD25^{high}$ T cells (which were >99% FoxP3+) release virus upon in vitro re-stimulation and have
482 ~3-fold higher HIV infection rates compared to other CD4 T cells upon in vitro activation (36,
483 70). Together these data suggest that CD25+FoxP3+ CD4 T cells are a prime cellular target for
484 HIV infection that might serve as an important HIV reservoir during ART.

We next wanted to address whether memory CD25+FoxP3+ CD4 T cells could potentially contribute to plasma virion production. Because cell fixation complicates analyses of HIV transcription in sorted cell populations defined by intranuclear transcription factors (such as FoxP3), we decided to study the phylogenetic relationship between plasma- and cell-derived sequences within the highly variable EnvV1V3 region; if CD25+FoxP3+ memory CD4 T cells significantly contribute to plasma virion production, EnvV1V3 DNA sequences derived from this cell population should often be quasi-identical or preferentially cluster with plasma-derived sequences. A previous study had reported rapid replacement of cell- and plasma-derived HIV sequences by an incoming superinfecting HIV strain (71), implying a highly dynamic exchange between these two compartments. In our study, detection of quasi-identical sequence pairs derived from cells and plasma was rare and their fraction further decreased with infection duration, which is consistent with the broadening of the viral reservoir with time. There was no clear pattern of phylogenetic clustering of the plasma virus with any of the cell subset-derived sequences we had sorted. In fact, cell-derived sequences did not "behave differently" from plasma-derived sequences and sequences from both compartments intermingled. Our phylogenetic data therefore do not allow definite conclusions about the cellular origin of plasma virions. The high variability between individual plasma-derived sequences during chronic infection emphasizes that a huge number of infected cells must contribute to plasma virion production at any given time during chronic infection. It might hence be difficult to determine the exact cellular origins of plasma virus through phylogenetic sequence analyses. Nonetheless, in our analyses of individual sequences, we did find several quasi-identical sequence pairs between plasma and CD25+FoxP3+ CD4 T cells, indicating that they may contribute to the plasma viremia. One limitation of our study was that we used comparatively small amounts of

508 PBMC and plasma (compared to the total body amount) for phylogenetic analyses and we
509 therefore probably included insufficient numbers for detection of clusters of cell- and plasma-
510 derived sequences (76). Virus sequences from very large amounts of specimen will need to be
511 analyzed and optimally include material from secondary lymphoid tissues for more conclusive
512 answers. Secondary lymphoid tissues are thought to constitute the primary site for virion
513 production (reviewed in (63)). After ART interruption, onset of viral RNA transcription in lymph
514 nodes coincides with a rise in plasma viral load (73). CD25+FoxP3+ CD4 T cells in secondary
515 lymphoid organs contain high frequencies of Ki67+ "cycling" cells with significant capacity for
516 IL2 production and often express a CD69+ "recently activated" phenotype (74) and hence differ
517 from those in peripheral blood. A recent study detected colocalization of SIV_p27- and FoxP3
518 expression in intestinal tissues using confocal microscopy (75). We therefore consider it likely
519 that CD25+FoxP3+ CD4 T cells in lymphoid tissues are a targeted by HIV, but additional studies
520 will be needed to define the role of CD25+FoxP3+ CD4 T cells for plasma virion production in
521 vivo.

522

523 We also sorted memory CD4 T cell populations depending on their Helios expression. Helios
524 is an Ikaros transcriptional factor family member is critical for the regulatory function of
525 CD25+FoxP3+ CD4 T cells (55–57) and for the prevention of autoimmunity (58). Helios
526 modulates cell cycle progression and sustained cell survival through regulation of genes involved
527 in IL-2 signaling (58, 59). Helios expression is also linked to expression of a range of
528 suppressive T cell markers and can be induced in CD4 T cells upon in vitro activation (72, 73).
529 In vitro, dividing CD25+FoxP3+CD4 T cells co-express Helios, while non-dividing regulatory T
530 cells lose expression of FoxP3 and Helios, suggesting Helios as a marker of recently divided

531 cells. In the same set of in vitro experiments, CD25-Helios+ CD4 T cells were composed of a
532 highly activated "effector" memory cells (72). We detected higher median Gag DNA loads in
533 memory CD25+FoxP3+ in both Helios positive (26-fold increased) and negative (119-fold
534 increased) as well as CD25-FoxP3- Helios+ memory CD4 T cells (104-fold increased) compared
535 to FoxP3-CD25- Helios- memory CD4 T cells. It is remarkable that we often did not detect HIV-
536 DNA in this "dominant" memory CD4 T cell subset. A history of more frequent or recent cell
537 divisions within CD25-FoxP3- Helios+ memory CD4 T cells might have contributed to high
538 HIV susceptibility in this memory subset, whereas removal of such cells in the sorted CD25-
539 FoxP3-Helios- memory CD4 T cells, could potentially explain the low HIV infection rates
540 observed in this memory cell subset. "Non-activated", circulating memory CD4 T cells are
541 probably less susceptible and accumulate less HIV DNA over time, in comparison to other
542 memory CD4 T cell subsets with a history of in vivo proliferation. Helios deficient regulatory
543 CD4 T cells exhibit an activated phenotype, increased capacity to secrete IFN-γ and develop into
544 non-anergic cells under inflammatory conditions (58, 74). Increased responsiveness to cellular
545 activation in comparison to their Helios+ counterparts signaling could potentially explain the
546 higher HIV-DNA levels in CD25+FoxP3+ Helios- memory CD4 T cells compared to their
547 Helios+ counterparts. These data show that Helios and CD25/FoxP3 expression patterns are
548 linked to different cellular HIV infection rates, consistent with a role of the IL2 signaling
549 pathway for HIV infection in vivo.

550

551 In conclusion, we find that homeostasis of CD25+FoxP3+ CD4 T cells is heavily
552 perturbed during HIV infection. High expression of HIV coreceptor-CCR5 and in vivo
553 proliferation potentially facilitates efficient HIV infection of memory CD25+FoxP3+ CD4 T

554 cells. Furthermore, high proliferative activity of this cell subset is likely to passage of HIV DNA
555 to cell progeny in the absence of active viral replication. This subset could therefore serve as an
556 important viral reservoir during ART. Neither circulating memory CD25+FoxP3+ CD4 T cell-
557 nor any of the other memory CD4 T cell subset-derived EnvV1V3 sequences preferentially
558 clustered with plasma-derived sequences. Instead, sequences from the two compartments
559 intermingled and the genetic distance in-between and within the two compartments increased
560 with infection duration, precluding definite conclusion about the cellular origin of the plasma
561 virus in this study.

562

563 **Acknowledgements**

564 We would like to thank Brenna Hill from the Vaccine Research Center, NIH in Bethesda for
565 providing HIV_Gag DNA standard and Andreas Wieser from the Max von Pettenkofer Institute,
566 Medical Center of the University of Munich (LMU) for advice to include a PCR error control.
567 The following reagent was obtained through the NIH AIDS Reagent Program, Division of AIDS,
568 NIAID, NIH from Drs. D. Montefiori, F. Gao, C. Williamson and S. Abdool Karim: Du422,
569 clone 1 (SVPC5).

570

References

1. **Mellors JW**, **Rinaldo CR**, **Gupta P**, **White RM**, **Todd JA**, **Kingsley LA**. 1996. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. Science (80) **272**:1167–70.

2. **Li TS**, **Tubiana R**, **Katlama C**, **Calvez V**, **Ait Mohand H**, **Autran B**. 1998. Long-lasting recovery in CD4 T-cell function and viral-load reduction after highly active antiretroviral therapy in advanced HIV-1 disease. Lancet **351**:1682–6.

3. **Verhofstede C**, **Reniers S**, **Van Wanzeele F**, **Plum J**. 1994. Evaluation of proviral copy number and plasma RNA level as early indicators of progression in HIV-1 infection: correlation with virological and immunological markers of disease. AIDS **8**:1421–7.

4. **Ganesan A**, **Chattopadhyay PK**, **Brodie TM**, **Qin J**, **Mascola JR**, **Michael NL**, **Follmann DA**, **Roederer M**. 2010. Immunological and Virological Events in Early HIV Infection Predict Subsequent Rate of Progression. J Infect Dis **201**:272–284.

5. **Brenchley JM**, **Schacker TW**, **Ruff LE**, **Price DA**, **Taylor JH**, **Beilman GJ**, **Nguyen PL**, **Khoruts A**, **Larson M**, **Haase AT**, **Douek DC**. 2004. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. J Exp Med **200**:749–59.

6. **Chomont N**, **El-Far M**, **Ancuta P**, **Trautmann L**, **Procopio F a**, **Yassine-Diab B**, **Boucher G**, **Boulassel M-R**, **Ghattas G**, **Brenchley JM**, **Schacker TW**, **Hill BJ**, **Douek DC**, **Routy J-P**, **Haddad EK**, **Sékaly R-P**. 2009. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. Nat Med **15**:893–900.

7. **Casazza JP**, **Brenchley JM**, **Hill BJ**, **Ayana R**, **Ambrozak D**, **Roederer M**, **Douek DC**,

Betts MR, **Koup RA**. 2009. Autocrine production of beta-chemokines protects CMV-Specific CD4 T cells from HIV infection. PLoS Pathog **5**:e1000646.

8. **Douek DC**, **Brenchley JM**, **Betts MR**, **Ambrozak DR**, **Hill BJ**, **Okamoto Y**, **Casazza JP**, **Kuruppu J**, **Kunstman K**, **Wolinsky S**, **Grossman Z**, **Dybul M**, **Oxenius A**, **Price DA**, **Connors M**, **Koup RA**. 2002. HIV preferentially infects HIV-specific CD4+ T cells. Nature **417**:95–8.

9. **Geldmacher C**, **Koup RA**. 2012. Pathogen-specific T cell depletion and reactivation of opportunistic pathogens in HIV infection. Trends Immunol **33**:207–14.

10. **Geldmacher C**, **Ngwenyama N**, **Schuetz A**, **Petrovas C**, **Reither K**, **Heeregrave EJ**, **Casazza JP**, **Ambrozak DR**, **Louder M**, **Ampofo W**, **Pollakis G**, **Hill B**, **Sanga E**, **Saathoff E**, **Maboko L**, **Roederer M**, **Paxton WA**, **Hoelscher M**, **Koup RA**. 2010. Preferential infection and depletion of Mycobacterium tuberculosis-specific CD4 T cells after HIV-1 infection. J Exp Med **207**:2869–81.

11. **Zhang Z**, **Schuler T**, **Zupancic M**, **Wietgrefe S**, **Staskus KA**, **Reimann KA**, **Reinhart TA**, **Rogan M**, **Cavert W**, **Miller CJ**, **Veazey RS**, **Notermans D**, **Little S**, **Danner SA**, **Richman DD**, **Havlir D**, **Wong J**, **Jordan HL**, **Schacker TW**, **Racz P**, **Tenner-Racz K**, **Letvin NL**, **Wolinsky S**, **Haase AT**. 1999. Sexual transmission and propagation of SIV and HIV in resting and activated CD4+ T cells. Science **286**:1353–7.

12. **Zack JA**, **Arrigo SJ**, **Weitsman SR**, **Go AS**, **Haislip A**, **Chen IS**. 1990. HIV-1 entry into quiescent primary lymphocytes: molecular analysis reveals a labile, latent viral structure. Cell **61**:213–22.

13. **Chou CS**, **Ramilo O**, **Vitetta ES**. 1997. Highly purified CD25- resting T cells cannot be

615 infected de novo with HIV-1. Proc Natl Acad Sci U S A **94**:1361–5.

616 14. **Biancotto A**, **Iglehart SJ**, **Vanpouille C**, **Condack CE**, **Lisco A**, **Ruecker E**, **Hirsch I**,
617 **Margolis LB**, **Grivel J-C**. 2008. HIV-1 induced activation of CD4+ T cells creates new
618 targets for HIV-1 infection in human lymphoid tissue ex vivo. Blood **111**:699–704.

619 15. **Maenetje P**, **Riou C**, **Casazza JP**, **Ambrozak D**, **Hill B**, **Gray G**, **Koup RA**, **de Bruyn**
620 **G**, **Gray CM**. 2010. A steady state of CD4+ T cell memory maturation and activation is
621 established during primary subtype C HIV-1 infection. J Immunol **184**:4926–35.

622 16. **Crotty S**. 2014. T Follicular Helper Cell Differentiation, Function, and Roles in Disease.
623 Immunity **41**:529–542.

624 17. **Perreau M**, **Savoye A-L**, **De Crignis E**, **Corpataux J-M**, **Cubas R**, **Haddad EK**, **De**
625 **Leval L**, **Graziosi C**, **Pantaleo G**. 2013. Follicular helper T cells serve as the major CD4
626 T cell compartment for HIV-1 infection, replication, and production. J Exp Med **210**:143–
627 56.

628 18. **Pallikkuth S**, **Sharkey M**, **Babic DZ**, **Gupta S**, **Stone GW**, **Fischl MA**, **Stevenson M**,
629 **Pahwa S**. 2015. Peripheral T Follicular Helper Cells Are the Major HIV Reservoir Within
630 Central Memory CD4 T Cells in Peripheral Blood from chronic HIV infected individuals
631 on cART. J Virol JVI.02883–15.

632 19. **Fukazawa Y**, **Lum R**, **Okoye AA**, **Park H**, **Matsuda K**, **Bae JY**, **Hagen SI**, **Shoemaker**
633 **R**, **Deleage C**, **Lucero C**, **Morcock D**, **Swanson T**, **Legasse AW**, **Axthelm MK**,
634 **Hesselgesser J**, **Geleziunas R**, **Hirsch VM**, **Edlefsen PT**, **Piatak M**, **Estes JD**, **Lifson**
635 **JD**, **Picker LJ**. 2015. B cell follicle sanctuary permits persistent productive simian
636 immunodeficiency virus infection in elite controllers. Nat Med **21**:132–9.

637 20. **Boyman O**, **Sprent J**. 2012. The role of interleukin-2 during homeostasis and activation
638 of the immune system. Nat Rev Immunol **12**:180–90.

639 21. **Finberg RW**, **Wahl SM**, **Allen JB**, **Soman G**, **Strom TB**, **Murphy JR**, **Nichols JC**.
640 1991. Selective elimination of HIV-1-infected cells with an interleukin-2 receptor-specific
641 cytotoxin. Science **252**:1703–5.

642 22. **Ramilo O**, **Bell KD**, **Uhr JW**, **Vitetta ES**. 1993. Role of CD25+ and CD25-T cells in
643 acute HIV infection in vitro. J Immunol **150**:5202–8.

644 23. **Goletti D**, **Weissman D**, **Jackson RW**, **Graham NM**, **Vlahov D**, **Klein RS**, **Munsiff SS**,
645 **Ortona L**, **Cauda R**, **Fauci AS**. 1996. Effect of Mycobacterium tuberculosis on HIV
646 replication. Role of immune activation. J Immunol **157**:1271–8.

647 24. **Baecher-Allan C**, **Brown JA**, **Freeman GJ**, **Hafler DA**. 2003. CD4+CD25+ regulatory
648 cells from human peripheral blood express very high levels of CD25 ex vivo. Novartis
649 Found Symp **252**:67–91, 106–14.

650 25. **Seddiki N**, **Santner-Nanan B**, **Martinson J**, **Zaunders J**, **Sasson S**, **Landay A**,
651 **Solomon M**, **Selby W**, **Alexander SI**, **Nanan R**, **Kelleher A**, **Fazekas de St Groth B**.
652 2006. Expression of interleukin (IL)-2 and IL-7 receptors discriminates between human
653 regulatory and activated T cells. J Exp Med **203**:1693–700.

654 26. **Hori S**, **Takahashi T**, **Sakaguchi S**. 2003. Control of autoimmunity by naturally arising
655 regulatory CD4+ T cells. Adv Immunol **81**:331–71.

656 27. **Sakaguchi S**, **Yamaguchi T**, **Nomura T**, **Ono M**. 2008. Regulatory T cells and immune
657 tolerance. Cell **133**:775–87.

658 28. **Aandahl EM**, **Michaëlsson J**, **Moretto WJ**, **Hecht FM**, **Nixon DF**. 2004. Human CD4+

CD25+ regulatory T cells control T-cell responses to human immunodeficiency virus and cytomegalovirus antigens. J Virol **78**:2454–9.

29. **Kinter AL**, **Horak R**, **Sion M**, **Riggin L**, **McNally J**, **Lin Y**, **Jackson R**, **O'shea A**, **Roby G**, **Kovacs C**, **Connors M**, **Migueles SA**, **Fauci AS**. 2007. CD25+ regulatory T cells isolated from HIV-infected individuals suppress the cytolytic and nonlytic antiviral activity of HIV-specific CD8+ T cells in vitro. AIDS Res Hum Retroviruses **23**:438–50.

30. **Booth NJ**, **McQuaid AJ**, **Sobande T**, **Kissane S**, **Agius E**, **Jackson SE**, **Salmon M**, **Falciani F**, **Yong K**, **Rustin MH**, **Akbar AN**, **Vukmanovic-Stejic M**. 2010. Different proliferative potential and migratory characteristics of human CD4+ regulatory T cells that express either CD45RA or CD45RO. J Immunol **184**:4317–26.

31. **Antons AK**, **Wang R**, **Oswald-Richter K**, **Tseng M**, **Arendt CW**, **Kalams SA**, **Unutmaz D**. 2008. Naive precursors of human regulatory T cells require FoxP3 for suppression and are susceptible to HIV infection. J Immunol **180**:764–773.

32. **Peters JH**, **Koenen HJPM**, **Fasse E**, **Tijssen HJ**, **Ijzermans JNM**, **Groenen PJTA**, **Schaap NPM**, **Kwekkeboom J**, **Joosten I**. 2013. Human secondary lymphoid organs typically contain polyclonally-activated proliferating regulatory T cells. Blood **122**:2213–23.

33. **Vukmanovic-stejic M**, **Zhang Y**, **Cook JE**, **Fletcher JM**, **Mcquaid A**, **Masters JE**, **Rustin MHA**, **Taams LS**, **Beverley PCL**, **Macallan DC**, **Akbar AN**. 2006. Human CD4+CD25hiFoxp3+ regulatory T cells are derived by rapid turnover of memory populations in vivo. J Clin Invest **116**:2423–2433.

34. **Vukmanovic-Stejic M**, **Agius E**, **Booth N**, **Dunne PJ**, **Lacy KE**, **Reed JR**, **Sobande**

TO, **Kissane S**, **Salmon M**, **Rustin MH**, **Akbar AN**. 2008. The kinetics of CD4+Foxp3+ T cell accumulation during a human cutaneous antigen-specific memory response in vivo. J Clin Invest **118**:3639–50.

35. **Sakaguchi S**, **Ono M**, **Setoguchi R**, **Yagi H**, **Hori S**, **Fehervari Z**, **Shimizu J**, **Takahashi T**, **Nomura T**. 2006. Foxp3+ CD25+ CD4+ natural regulatory T cells in dominant self-tolerance and autoimmune disease. Immunol Rev **212**:8–27.

36. **Oswald-Richter K**, **Grill SM**, **Shariat N**, **Leelawong M**, **Sundrud MS**, **Haas DW**, **Unutmaz D**. 2004. HIV infection of naturally occurring and genetically reprogrammed human regulatory T-cells. PLoS Biol **2**:E198.

37. **Chachage M**, **Podola L**, **Clowes P**, **Nsojo A**, **Bauer A**, **Mgaya O**, **Kowour D**, **Froeschl G**, **Maboko L**, **Hoelscher M**, **Saathoff E**, **Geldmacher C**. 2014. Helminth-associated systemic immune activation and HIV co-receptor expression: Response to Albendazole/Praziquantel treatment. PLoS Negl Trop Dis **8**:e2755.

38. **Sarfo FS**, **Eberhardt KA**, **Dompreh A**, **Kuffour EO**, **Soltau M**, **Schachscheider M**, **Drexler JF**, **Eis-Hübinger AM**, **Häussinger D**, **Oteng-Seifah EE**, **Bedu-Addo G**, **Phillips RO**, **Norman B**, **Burchard G**, **Feldt T**. 2015. Helicobacter pylori Infection Is Associated with Higher CD4 T Cell Counts and Lower HIV-1 Viral Loads in ART-Naïve HIV-Positive Patients in Ghana. PLoS One **10**:e0143388.

39. **Eberhardt KA**, **Sarfo FS**, **Dompreh A**, **Kuffour EO**, **Geldmacher C**, **Soltau M**, **Schachscheider M**, **Drexler JF**, **Eis-Hubinger AM**, **Haussinger D**, **Bedu-Addo G**, **Phillips RO**, **Norman B**, **Burchard GD**, **Feldt T**. 2015. Helicobacter pylori Coinfection Is Associated with Decreased Markers of Immune Activation in ART-Naive HIV-Positive and in HIV-Negative Individuals in Ghana. Clin Infect Dis **61**:1615–1623.

40. **Geldmacher C**, **Currier JR**, **Herrmann E**, **Haule A**, **Kuta E**, **McCutchan F**, **Njovu L**, **Geis S**, **Hoffmann O**, **Maboko L**, **Williamson C**, **Birx D**, **Meyerhans A**, **Cox J**, **Hoelscher M**. 2007. CD8 T-cell recognition of multiple epitopes within specific Gag regions is associated with maintenance of a low steady-state viremia in human immunodeficiency virus type 1-seropositive patients. J Virol **81**:2440–8.

41. **Hoffmann D**, **Wolfarth B**, **Hörterer HG**, **Halle M**, **Reichhuber C**, **Nadas K**, **Tora C**, **Erfle V**, **Protzer U**, **Schätzl HM**. 2010. Elevated Epstein-Barr virus loads and lower antibody titers in competitive athletes. J Med Virol **82**:446–51.

42. **Li M**, **Salazar-Gonzalez JF**, **Derdeyn CA**, **Morris L**, **Williamson C**, **Robinson JE**, **Decker JM**, **Li Y**, **Salazar MG**, **Polonis VR**, **Mlisana K**, **Karim SA**, **Hong K**, **Greene KM**, **Bilska M**, **Zhou J**, **Allen S**, **Chomba E**, **Mulenga J**, **Vwalika C**, **Gao F**, **Zhang M**, **Korber BTM**, **Hunter E**, **Hahn BH**, **Montefiori DC**. 2006. Genetic and neutralization properties of subtype C human immunodeficiency virus type 1 molecular env clones from acute and early heterosexually acquired infections in Southern Africa. J Virol **80**:11776–90.

43. **Pollakis G**, **Baan E**, **van Werkhoven MB**, **Berkhout B**, **Bakker M**, **Jurriaans S**, **Paxton WA**. 2015. Association between gp120 envelope V1V2 and V4V5 variable loop profiles in a defined HIV-1 transmission cluster. AIDS **29**:1161–71.

44. **Tamura K**, **Stecher G**, **Peterson D**, **Filipski A**, **Kumar S**. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol **30**:2725–9.

45. **Nei M**, **Kumar S**. 2000. Molecular Evolution and Phylogenetics. Oxford University Press.

46. **Benson DA**, **Karsch-Mizrachi I**, **Lipman DJ**, **Ostell J**, **Sayers EW**. 2009. GenBank. Nucleic Acids Res **37**:D26–31.

47. **Hoffmann S**, **Otto C**, **Kurtz S**, **Sharma CM**, **Khaitovich P**, **Vogel J**, **Stadler PF**, **Hackermüller J**. 2009. Fast mapping of short sequences with mismatches, insertions and deletions using index structures. PLoS Comput Biol **5**:e1000502.

48. **Li H**, **Handsaker B**, **Wysoker A**, **Fennell T**, **Ruan J**, **Homer N**, **Marth G**, **Abecasis G**, **Durbin R**. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**:2078–9.

49. **Töpfer A**, **Zagordi O**, **Prabhakaran S**, **Roth V**, **Halperin E**, **Beerenwinkel N**. 2013. Probabilistic inference of viral quasispecies subject to recombination. J Comput Biol **20**:113–23.

50. **Angin M**, **Kwon DS**, **Streeck H**, **Wen F**, **King M**, **Rezai A**, **Law K**, **Hongo TC**, **Pyo A**, **Piechocka-Trocha A**, **Toth I**, **Pereyra F**, **Ghebremichael M**, **Rodig SJ**, **Milner DA**, **Richter JM**, **Altfeld M**, **Kaufmann DE**, **Walker BD**, **Addo MM**. 2012. Preserved function of regulatory T cells in chronic HIV-1 infection despite decreased numbers in blood and tissue. J Infect Dis **205**:1495–500.

51. **Schulze Zur Wiesch J**, **Thomssen A**, **Hartjen P**, **Tóth I**, **Lehmann C**, **Meyer-Olson D**, **Colberg K**, **Frerk S**, **Babikir D**, **Schmiedel S**, **Degen O**, **Mauss S**, **Rockstroh J**, **Staszewski S**, **Khaykin P**, **Strasak A**, **Lohse AW**, **Fätkenheuer G**, **Hauber J**, **van Lunzen J**. 2011. Comprehensive analysis of frequency and phenotype of T regulatory cells in HIV infection: CD39 expression of FoxP3+ T regulatory cells correlates with progressive disease. J Virol **85**:1287–97.

748 52. **Presicce P**, **Orsborn K**, **King E**, **Pratt J**, **Fichtenbaum CJ**, **Chougnet CA**. 2011.
749 Frequency of circulating regulatory T cells increases during chronic HIV infection and is
750 largely controlled by highly active antiretroviral therapy. PLoS One **6**:e28118.

751 53. **Montes M**, **Lewis DE**, **Sanchez C**, **Lopez de Castilla D**, **Graviss EA**, **Seas C**, **Gotuzzo**
752 **E**, **White AC**. 2006. Foxp3+ regulatory T cells in antiretroviral-naive HIV patients. AIDS
753 **20**:1669–71.

754 54. **Simonetta F**, **Lecuroux C**, **Girault I**, **Goujard C**, **Sinet M**, **Lambotte O**, **Venet A**,
755 **Bourgeois C**. 2012. Early and long-lasting alteration of effector CD45RA(-)Foxp3(high)
756 regulatory T-cell homeostasis during HIV infection. J Infect Dis **205**:1510–9.

757 55. **Sugimoto N**, **Oida T**, **Hirota K**, **Nakamura K**, **Nomura T**, **Uchiyama T**, **Sakaguchi S**.
758 2006. Foxp3-dependent and -independent molecules specific for CD25+CD4+ natural
759 regulatory T cells revealed by DNA microarray analysis. Int Immunol **18**:1197–209.

760 56. **Thornton AM**, **Korty PE**, **Tran DQ**, **Wohlfert EA**, **Murray PE**, **Belkaid Y**, **Shevach**
761 **EM**. 2010. Expression of Helios, an Ikaros transcription factor family member,
762 differentiates thymic-derived from peripherally induced Foxp3+ T regulatory cells. J
763 Immunol **184**:3433–41.

764 57. **Getnet D**, **Grosso JF**, **Goldberg M V**, **Harris TJ**, **Yen H-R**, **Bruno TC**, **Durham NM**,
765 **Hipkiss EL**, **Pyle KJ**, **Wada S**, **Pan F**, **Pardoll DM**, **Drake CG**. 2010. A role for the
766 transcription factor Helios in human CD4(+)CD25(+) regulatory T cells. Mol Immunol
767 **47**:1595–600.

768 58. **Kim H-J**, **Barnitz RA**, **Kreslavsky T**, **Brown FD**, **Moffett H**, **Lemieux ME**, **Kaygusuz**
769 **Y**, **Meissner T**, **Holderried TAW**, **Chan S**, **Kastner P**, **Haining WN**, **Cantor H**. 2015.

770 Stable inhibitory activity of regulatory T cells requires the transcription factor Helios.
771 Science (80) **350**:334–339.

772 59. **Baine I**, **Basu S**, **Ames R**, **Sellers RS**, **Macian F**. 2012. Helios Induces Epigenetic
773 Silencing of Il2 Gene Expression in Regulatory T Cells. J Immunol **190**:1008–16.

774 60. **Ioannidis JP**, **Cappelleri JC**, **Lau J**, **Sacks HS**, **Skolnik PR**. 1996. Predictive value of
775 viral load measurements in asymptomatic untreated HIV-1 infection: a mathematical
776 model. AIDS **10**:255–62.

777 61. **Haase AT**. 1999. Population biology of HIV-1 infection: viral and CD4+ T cell
778 demographics and dynamics in lymphatic tissues. Annu Rev Immunol **17**:625–56.

779 62. **Veazey RS**, **DeMaria M**, **Chalifoux L V**, **Shvetz DE**, **Pauley DR**, **Knight HL**,
780 **Rosenzweig M**, **Johnson RP**, **Desrosiers RC**, **Lackner AA**. 1998. Gastrointestinal tract
781 as a major site of CD4+ T cell depletion and viral replication in SIV infection. Science
782 **280**:427–31.

783 63. **Schacker T**, **Little S**, **Connick E**, **Gebhard K**, **Zhang ZQ**, **Krieger J**, **Pryor J**, **Havlir**
784 **D**, **Wong JK**, **Schooley RT**, **Richman D**, **Corey L**, **Haase AT**. 2001. Productive
785 infection of T cells in lymphoid tissues during primary and early human
786 immunodeficiency virus infection. J Infect Dis **183**:555–62.

787 64. **Brenchley JM**, **Hill BJ**, **Ambrozak DR**, **Price DA**, **Guenaga FJ**, **Casazza JP**, **Kuruppu**
788 **J**, **Yazdani J**, **Migueles SA**, **Connors M**, **Roederer M**, **Douek DC**, **Koup RA**. 2004. T-
789 cell subsets that harbor human immunodeficiency virus (HIV) in vivo: implications for
790 HIV pathogenesis. J Virol **78**:1160–8.

791 65. **Setoguchi R**, **Hori S**, **Takahashi T**, **Sakaguchi S**. 2005. Homeostatic maintenance of

792 natural Foxp3(+) CD25(+) CD4(+) regulatory T cells by interleukin (IL)-2 and induction
793 of autoimmune disease by IL-2 neutralization. J Exp Med **201**:723–35.

794 66. **Moreno-Fernandez ME**, **Zapata W**, **Blackard JT**, **Franchini G**, **Chougnet CA**. 2009.
795 Human regulatory T cells are targets for human immunodeficiency Virus (HIV) infection,
796 and their susceptibility differs depending on the HIV type 1 strain. J Virol **83**:12925–33.

797 67. **Nilsson J**, **Boasso A**, **Velilla PA**, **Zhang R**, **Vaccari M**, **Franchini G**, **Shearer GM**,
798 **Andersson J**, **Chougnet C**. 2006. HIV-1-driven regulatory T-cell accumulation in
799 lymphoid tissues is associated with disease progression in HIV/AIDS. Blood **108**:3808–
800 17.

801 68. **Dunham RM**, **Cervasi B**, **Brenchley JM**, **Albrecht H**, **Weintrob A**, **Sumpter B**,
802 **Gordon S**, **Klatt NR**, **Frank I**, **Donald L**, **Douek DC**, **Paiardini M**, **Silvestri G**,
803 **Engram J**, **Sodora DL**, **Sodora L**. 2008. CD127 and CD25 Expression Defines CD4+ T
804 Cell Subsets That Are Differentially Depleted during HIV Infection. J Immunol
805 **180**:5582–5592.

806 69. **Chase AJ**, **Yang H-C**, **Zhang H**, **Blankson JN**, **Siliciano RF**. 2008. Preservation of
807 FoxP3+ regulatory T cells in the peripheral blood of human immunodeficiency virus type
808 1-infected elite suppressors correlates with low CD4+ T-cell activation. J Virol **82**:8307–
809 15.

810 70. **Tran T-A**, **de Goër de Herve M-G**, **Hendel-Chavez H**, **Dembele B**, **Le Névot E**, **Abbed**
811 **K**, **Pallier C**, **Goujard C**, **Gasnault J**, **Delfraissy J-F**, **Balazuc A-M**, **Taoufik Y**. 2008.
812 Resting regulatory CD4 T cells: a site of HIV persistence in patients on long-term
813 effective antiretroviral therapy. PLoS One **3**:e3305.

71. **McCutchan FE**, **Hoelscher M**, **Tovanabutra S**, **Piyasirisilp S**, **Sanders-Buell E**, **Ramos G**, **Jagodzinski L**, **Polonis V**, **Maboko L**, **Mmbando D**, **Hoffmann O**, **Riedner G**, **von Sonnenburg F**, **Robb M**, **Birx DL**. 2005. In-depth analysis of a heterosexually acquired human immunodeficiency virus type 1 superinfection: evolution, temporal fluctuation, and intercompartment dynamics from the seronegative window period through 30 months postinfection. J Virol **79**:11693–704.

72. **Akimova T**, **Beier UH**, **Wang L**, **Levine MH**, **Hancock WW**. 2011. Helios expression is a marker of T cell activation and proliferation. PLoS One **6**:e24226.

73. **Verhagen J**, **Wraith DC**. 2010. Comment on "Expression of Helios, an Ikaros transcription factor family member, differentiates thymic-derived from peripherally induced Foxp3+ T regulatory cells". J Immunol **185**:7129.

74. **Sebastian M**, **Lopez-Ocasio M**, **Metidji A**, **Rieder SA**, **Shevach EM**, **Thornton AM**. 2015. Helios Controls a Limited Subset of Regulatory T Cell Functions. J Immunol **196**:144–55.

Figure Legends

Fig 1. Frequencies and absolute numbers of CD25+FoxP3+ CD4 T cells in the peripheral blood in relation to HIV infection. Representative dot plots and gating strategy for the detection of regulatory T cells through CD25 and FoxP3 expression on $CD3^{+}CD4^{+}$ T cells from fresh anticoagulated whole blood of WHIS subjects are shown in (**A**). CD25+Foxp3+ CD4 T cell frequencies and absolute numbers were compared between HIV- and HIV+ subjects in (**B**) and (**C**), respectively. A correlation analysis of absolute CD4 counts and CD25+Foxp3+ CD4 T cell counts is shown in (**D**). Statistical analysis was performed using Mann-Whitney test when comparing groups and Spearman r statistical test for correlation analyses.

Fig 2. Ex vivo HIV-co receptor (CCR5) expression on CD25+Foxp3+ CD4 T cells. Shown is (**A**) a histogram overlay for CCR5 expression on total CD4 T cells (grey) and CD25+Foxp3+ CD4 T cells (black). The frequencies of CCR5+ expressing CD25+Foxp3+ CD4 T cells are compared between HIV negative and positive subjects in (**B**). For maximum staining sensitivity, fresh anticoagulated whole blood of individuals from the WHIS cohort was used to determine CCR5 expression on CD4 T cells. Statistical analysis was performed using Mann-Whitney test.

Fig 3. Ki67 expression in memory CD25+FoxP3+ and CD25-FoxP3- CD4 T cells in HIV+ subjects. Representative dot plots for Ki67 staining are shown in (**A**). A comparison of the frequencies of Ki67+ cells in memory CD25+FoxP3+ and CD25-FoxP3- memory CD4 T cells in HIV+ subjects is shown in (**B**). A correlation analysis of the frequency of Ki67+ cells between CD25+FoxP3+ (Y axis) and CD25-FoxP3- (X axis) memory CD4 T cells is shown in (**C**). A correlation analysis of the frequency of Ki67+ cells among CD25+FoxP3+ and CD25-FoxP3- CD4 T cells versus CD4 T cell frequencies (% of CD3) is shown in (**D**) and (**E**) respectively.

852 The analysis was done using cryopreserved PBMC samples from HIV+ HHECO study
853 participants. Memory status of CD4 T cells was determined by CD45RA staining. Statistical
854 analysis was performed using Mann-Whitney test when comparing groups and Spearman r
855 statistical test for correlation analyses.

856 **Fig 4. Quantification of Cell associated HIV gag DNA in sorted memory CD4 T cell subsets**.
857 Gating/sorting strategy used to sort different memory CD4 T cell populations delineated by
858 Helios, CD25 and FoxP3 expression (**A**). The number of gag copies/10^6cells detected in $CD25^-$
859 /$FoxP3^-$ and $CD25^+$/$FoxP3^+$ memory CD4 T cells from 21 different subjects is shown in (**B**). The
860 number of gag copies/10^6cells detected in these memory CD4 T cell subsets further delineated by
861 Helios expression is shown in (**C**). Gag DNA within different CD4 T cell populations of the
862 same subject was quantified during the same RT-PCR run. Cryopreserved PBMC from the
863 WHIS and HISIS cohorts were used for cell sorting. The statistical analysis was performed using
864 the Wilcoxon-rank-matched pairs test.

865

866 **Fig 5. Phylogenetic relationship of HIV Envelope sequences derived from plasma and**
867 **sorted memory CD4 T cell populations.** Plasma- and cell-derived sequences of the highly
868 variable EnvV1V3 region (Hxb 6559–7320) were amplified cloned, sequenced (n=384, Sanger
869 method) and analyzed for 6 viremic subjects from the WHIS and HISIS cohorts with differing
870 HIV infection duration**.** The phylogenetic relationship was inferred by the Maximum Likelihood
871 method based on the General Time Reversible substitution model (GTR+G, **A and B**).
872 Correlation between frequency of cell-derived sequences that were quasi-identical to plasma-
873 derived sequences and the estimated infection duration is shown in (**C**). Linear regression
874 analysis (green line) between: (**D**) the distance of the EnvV1V3 sequences derived from plasma

875 to the sequences extracted from the corresponding cellular fractions and the estimated duration
876 of infection, and (**E**) plasma sequences diversity plotted against the estimated duration of
877 infection. The red line indicates a non-linear analysis performed using a second order polynomial
878 equation taking into account the best-fit values. The evolutionary distances were computed
879 using the Kimura 2-parameter method (75) and are in the units of the number of base
880 substitutions per site including both Transitions + Transversions. The rate variation among sites
881 was modeled with a gamma distribution. The analysis was conducted in MEGA6 (70). No
882 sequence diversity was observed in the 8710 plasma fraction, probably because the number of
883 viruses sampled in each PCR was very low (Table 2). We hence excluded 8710 results from the
884 linear regression analysis. P and r-values were calculated with the Pearson two-tailed statistical
885 test.
886
887 **Fig 6. Phylogenetic analyses of HIV Envelope sequences derived from plasma and sorted**
888 **memory CD4 T cell populations using a using Next Generation sequencing.**
889 Shown is the phylogenetic analyses of EnvV1V3 sequences from the 50 most frequently detected
890 sequences derived from either plasma or the different sorted memory CD4 T cell subsets for two
891 viremic subjects of the WHIS cohort. The phylogenetic relationship was inferred by the
892 Maximum Likelihood method based on the General Time Reversible substitution model
893 (GTR+G). EnvV1V3 amplicons were directly subjected to next generation sequencing. Quasi-
894 species reconstruction was performed using the software QuasiRecomb. The applied methods are
895 described in detail in the material and methods section.
896

Tables

Table 1. Characteristics of study subjects from different cohorts.

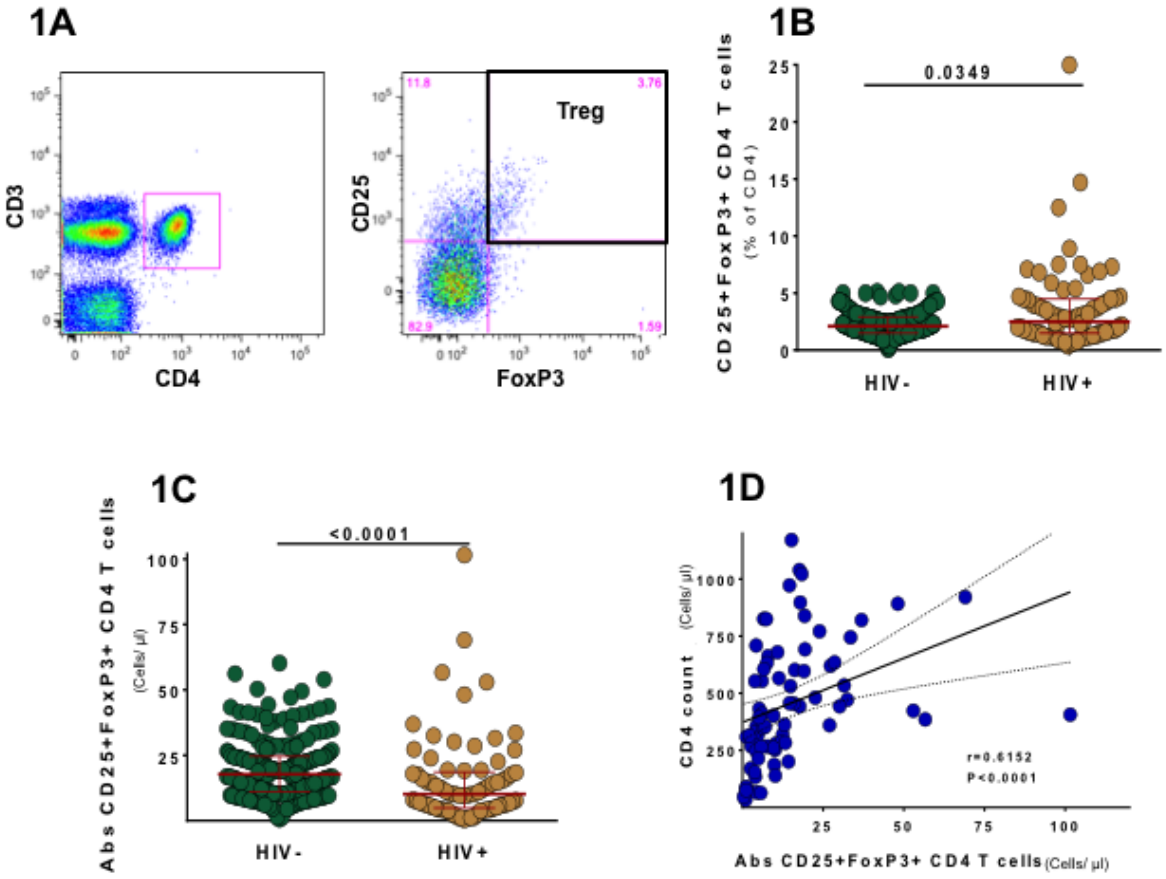
	WHIS	HHECO	HISIS
N	361	28	6
HIV pos., N	103	28	6
Females, N	217	25	6
Age, mean (SD)	34.3 (11.05)	38.8 (7.5)	28 (3.2)
Median CD4, cell/ul (IQR)*	396 (265-603)	629 (444-900)	496 (231 - 707)
Median log pVL, copy/ml (IQR)*	4.67 (3.74-5.23)	1.59 (1.59-3.82)	4.9 (4.4 - 5.5)
On ARV treatment, N (%)*	3 (0.8)	20 (71.4)	0 (0)

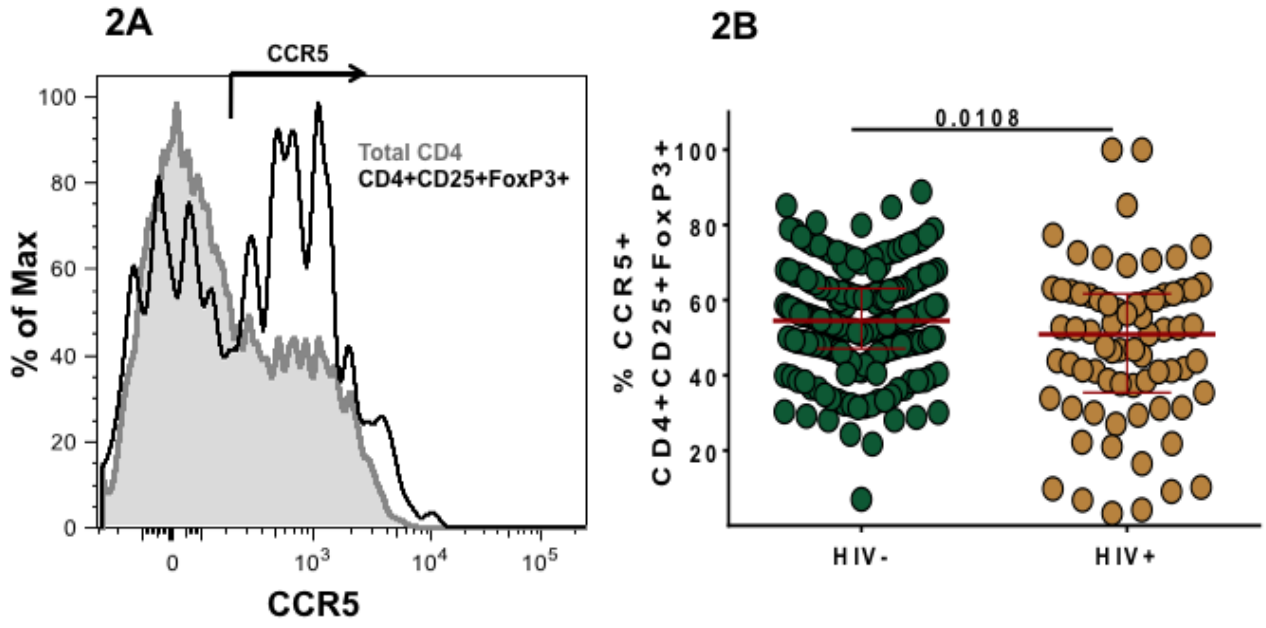
***Data shown for HIV positive subjects only**

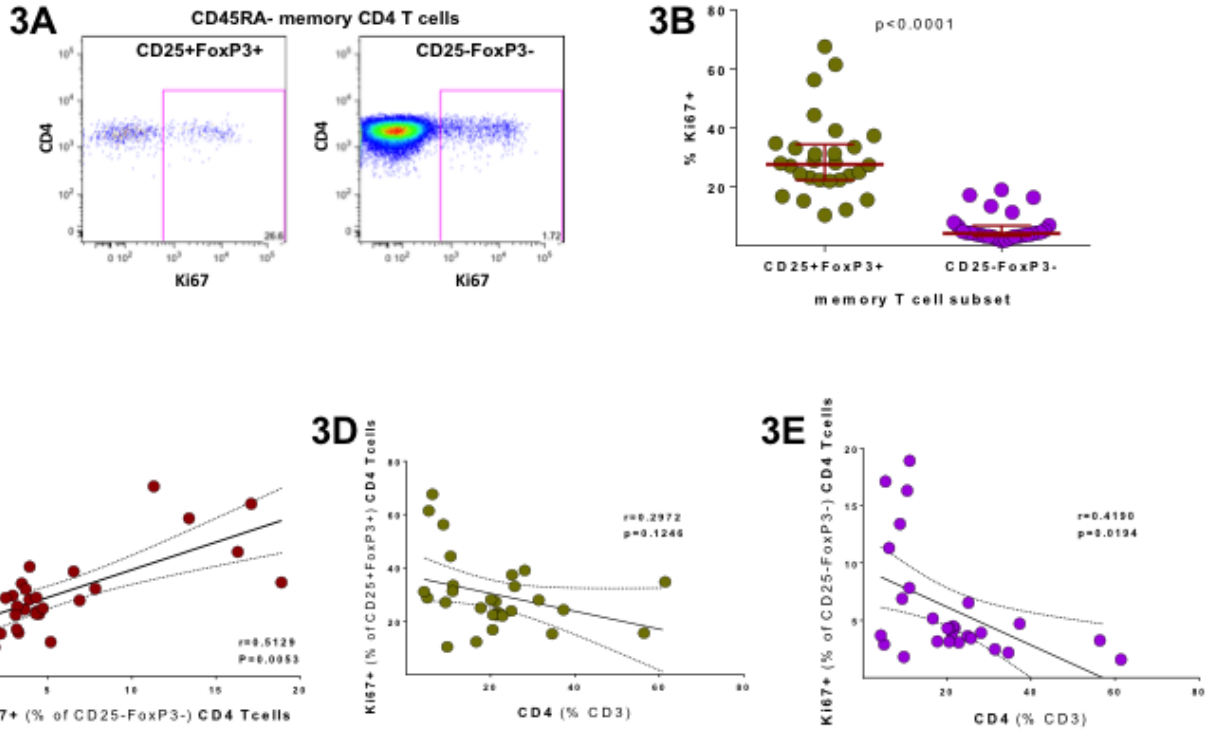
Table 2. Key data of the EnvV1V3 phylogenetic studies and HIV infection duration for 6 viremic subjects.

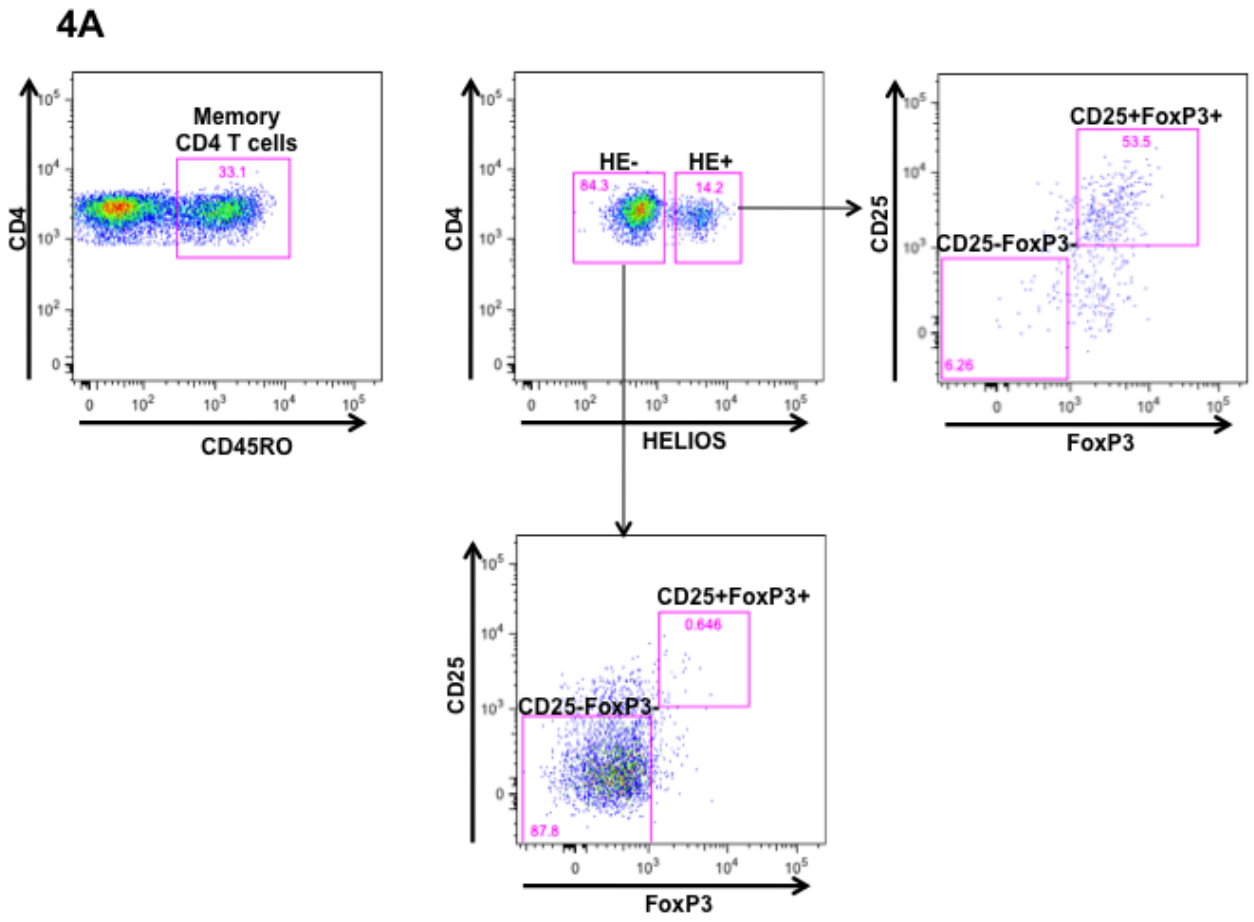
Subject ID	HIV Infection duration (months)	% of cell-derived sequences quasi-identical to plasma-derived sequences (n)	mean number of nucleotide substitutions between plasma and cell-derived sequences	cellular origin of closest sequence	Number of nucleotide substitutions	cellular origin of most distant sequence	Number of nucleotide substitutions
H574	9 to 12	11.4 (8 of 70)	6	CD25+FoxP3+Helios-	1	CD25+FoxP3+Helios+	16
				CD25+FoxP3+Helios+	1		
				CD25-FoxP3-Helios+	1		
				CD25-FoxP3-Helios-	1		
H605	27 to 30	6.8 (3 of 44)	39	CD25-Foxp3-Helios-	3	CD25+FoxP3+Helios+	32
6233K12	16 to 38	1.9 (1 of 53)	30	CD25-FoxP3-Helios+	2	CD25-FoxP3-Helios+	30
9440A11*	>38	2.6 (1 of 38)	46	CD25+Helios+	4	CD25+Helios+	76
8710U11	>38	0 (0 of 39)	57	CD25-FoxP3-Helios-	32	CD25-FoxP3-Helios-	67
8975T11	>54	0 (0 of 55)	53	CD25-FoxP3-Helios+	54	CD25-FoxP3-Helios-	86

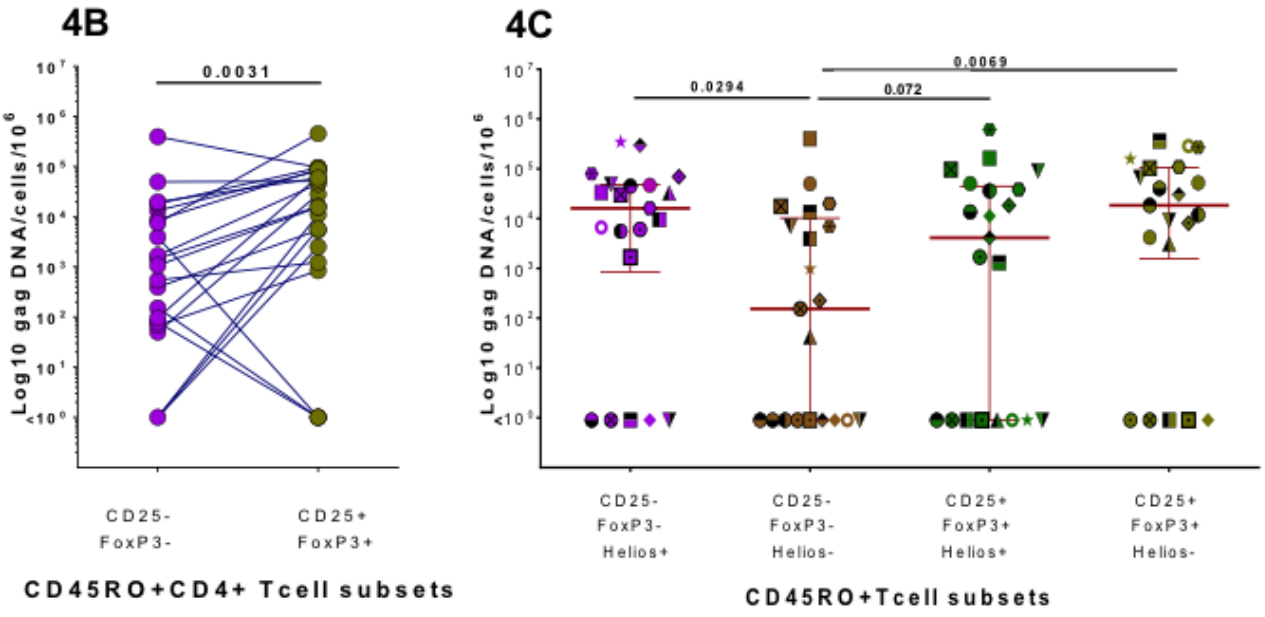
***Cells (PBMCs) from this subject were sorted into four populations only on the basis of CD25 and Helios expression on memory (CD45RO) CD4 T cells**



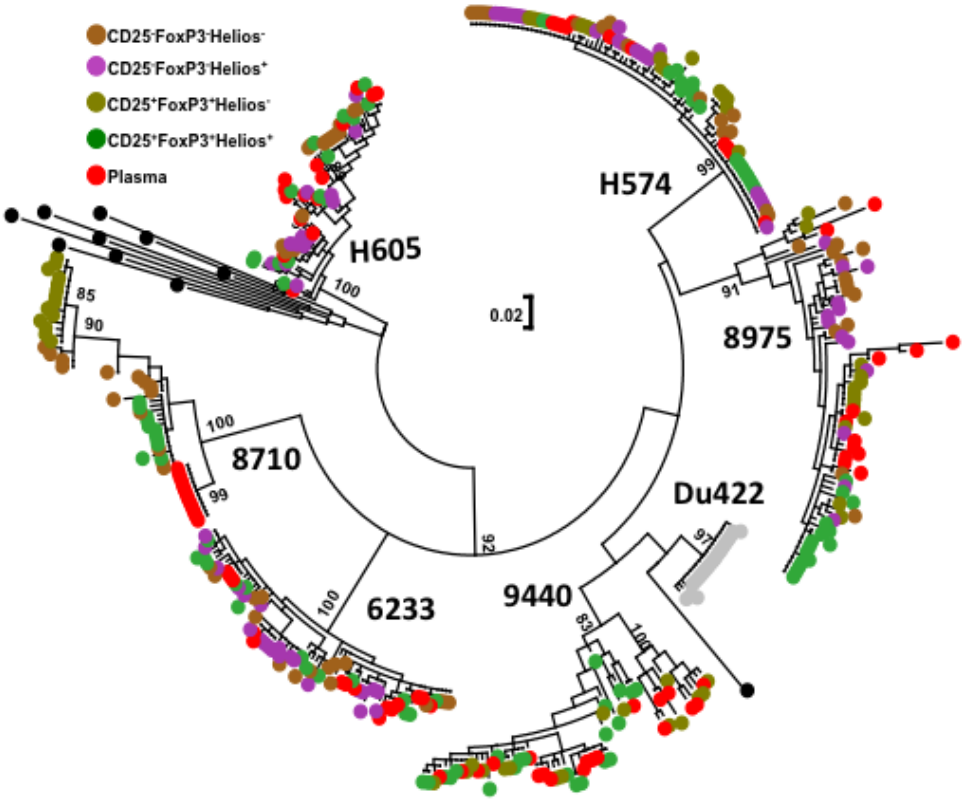




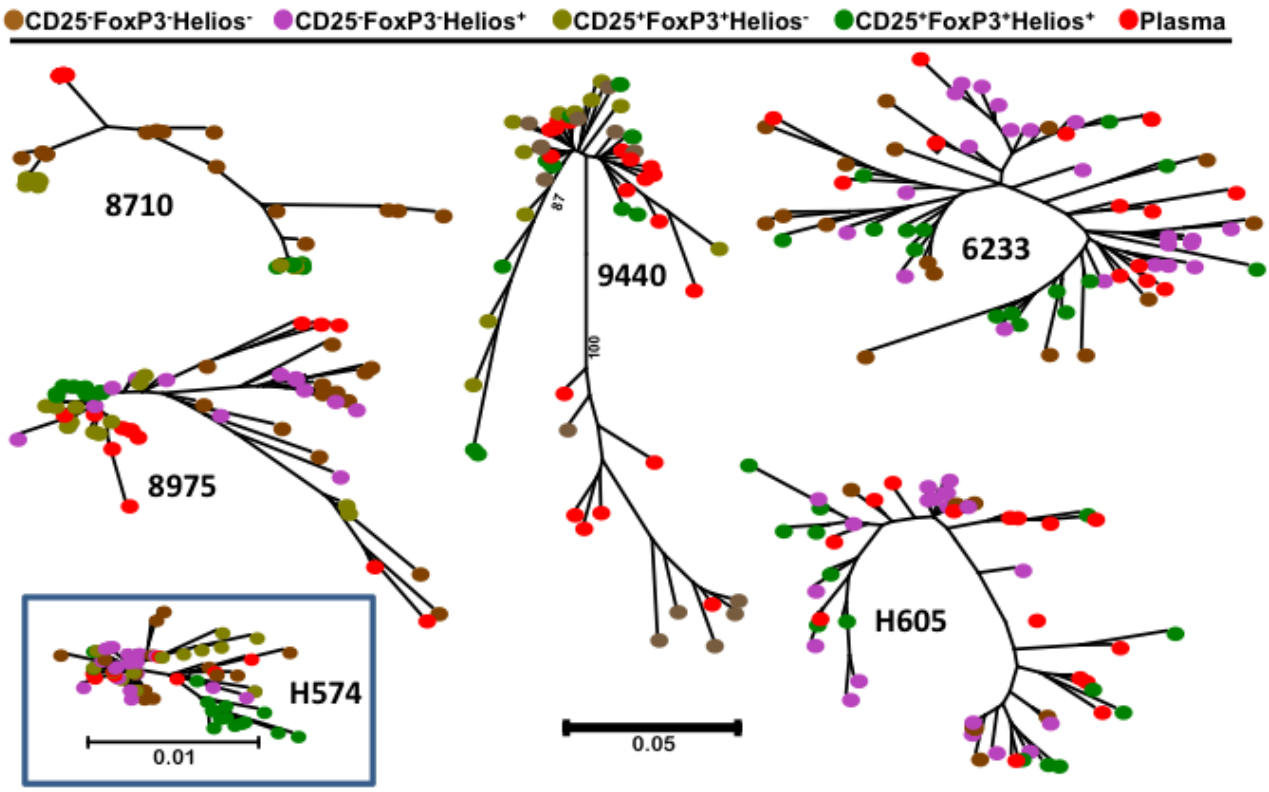




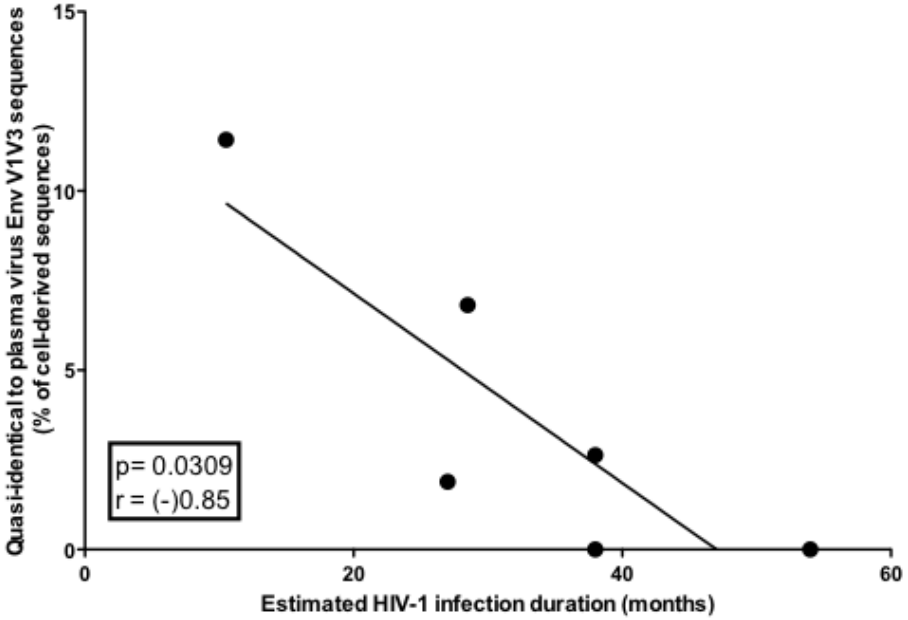
5A



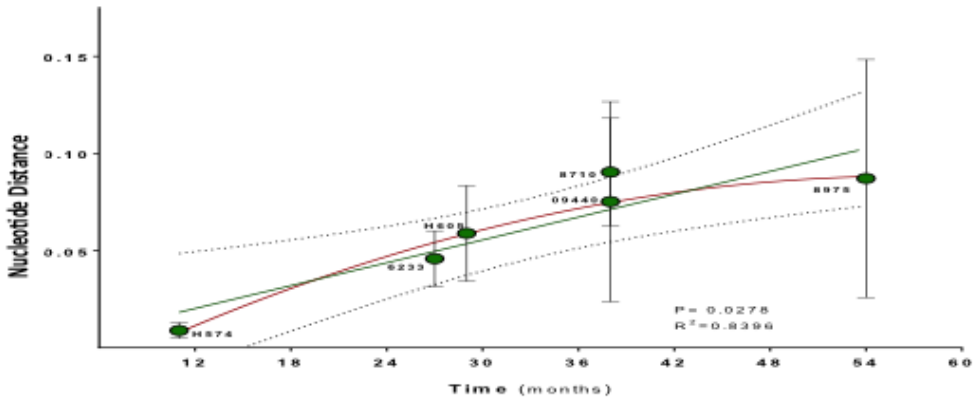
5B



5C



5D



5E