From the Division of Infectious Diseases and Tropical Medicine
Ludwig-Maximillians-Universität, München
Director: Prof. Dr. med. Michael Hoelscher

Phylogenetic comparison of plasma- and cell subset-derived viral sequences to identify the cellular origin of HIV plasma viremia

Dissertation
to obtain the degree
Doctor of Medicine (Dr. med.)
at the
Faculty of Medicine
Ludwig-Maximillians-Universität, München

Submitted by:
Astrid Hielscher

Leinefelde, 2020

With the approval
of the Faculty of Medicine
of the University, Munich

Reviewer:	Prof. Dr. Michael Hoelscher
Co-Reviewer:	PD Dr. Hanna-Mari Baldauf
	PD Dr. Bernd Sporer
Supervisor:	Dr. rer. nat. Christof Geldmacher
Dean:	Prof. Dr. med. dent. Reinhard Hickel
Date of oral examination:	5.03.2020

Abstract

$CD4^+$ T cells provide the major target for HIV infection initiation but productive viral infection is accompanied with CD4 receptor down-modulation. However, it is still not completely understood which cell subsets support HIV infection and replication in vivo and contribute to plasma viremia. Under the assumption that cell-associated HIV nucleotide sequences, quasi-identical to plasma-derived sequences, originate from HIV infected cells producing plasma virus, we analyzed plasma and cell-associated HIV sequences in sorted memory $CD45RA^-$ T cell populations including $CD4^{low/-}/CD8^-$ T cells, regulatory T cells (Tregs, $CD4^+CD25^+FoxP3^+$), T-follicular helper cells (Tfh, $CD4^+CXCR5^+$) and non-T-follicular helper cells (non-Tfh, $CD4^+CXCR5^-$). A nested PCR was used to amplify, clone and sequence the highly variable HIV Envelope (Env) V1V3 region (HxB6559 to 7320) from the sorted cell populations. Nucleotide sequences were aligned with respect to the predicted amino acid sequence of the reference alignment using MEGA6. Up to 4 substitutions between cell and plasma-associated sequences were counted as quasi-identical. Collectively, 621 EnvV1V3 sequences from 9 viremic HIV positive subjects were analyzed. Dual infection evidence was detected in 4 subjects and quasi-identical cell-plasma sequence pairs were detected in 5 of the 9 subjects. For patient H76 3 cell-derived sequences from the $CD4^{low/-}/CD8^-$ T cell subset were quasi-identical with 4 distinct plasma sequences. For patient H241 17 cell-derived sequences from the $CD4^+CD25^+FoxP3^+$ T cell subset were quasi-identical with 2 plasma sequences. For patient H304 1 $CD4^+CXCR5^+$ and 1 $CD4^+CXCR5^-$ cell subset was quasi-identical with 1 plasma sequence. For patient H442 1 $CD4^+CXCR5^-$ cell subset was quasi-identical with 10 plasma-derived sequences and 1 $CD4^{low/-}/CD8^-$ T cell subset was quasi-identical with 1 plasma-derived sequence. For H446 1 $CD4^+CXCR5^-$ T cell subset was quasi-identical with 1 plasma derived sequence and further 2 $CD4^+CXCR5^-$ T cell subsets were quasi-identical with 1 other plasma derived sequence. These results are consistent with the concept that the majority of infected memory CD4 T cells in peripheral blood may not contribute to plasma viremia. Furthermore, none of the cell subset-derived sequences preferentially clustered with plasma virus sequences. Rather, the data indicate that all of the studied memory T cell populations can either contribute to plasma virus production and/or have recently become infected.

Zusammenfassung

Während $CD4^+$ T-Zellen das Hauptziel einer beginnenden HIV-Infektion darstellen, geht eine produktive Virusinfektion mit einer Herabregulation der CD4-Rezeptoren einher. Bis heute ist jedoch nicht eindeutig geklärt, welche Zellpopulationen eine HIV-Infektion und Replikation *in vivo* fördern, und zur Plasmavirämie beitragen. Um dies genauer zu charakterisieren, haben wir Plasma und Zell assoziierte HIV Sequenzen aus separierten $CD45RA^-$ T-Gedächtniszellpopulationen, einschließlich $CD4^{niedrig/-}/CD8^-$ T-Zellen, regulatorischen T-Zellen (Tregs, $CD4^+CD25^+FoxP3^+$), follikulären T-Helferzellen (Tfh, $CD4^+CXCR5^+$) und nicht-follikulären T-Helferzellen (nicht-Tfh, $CD4^+CXCR5^-$) untersucht. Hierbei nahmen wir an, dass Zell assoziierte HIV-Nukleotidsequenzen, quasi-identisch zu Plasma assoziierten Sequenzen, ihren Ursprung aus HIV-infizierten, Virus produzierenden Zellen haben. Um die hochvariable HIV Envelope (Env) V1V3 Region (HxB6559 bis 7320) aus geordneten Zellpopulationen zu amplifizieren, klonieren und zu sequenzieren, wurde eine nested-PCR verwendet. Die resultierenden Nukleotidsequenzen wurden in Aminosäuresequenzen übersetzt, und mithilfe von MEGA6 auf die Referenzsequenz ausgerichtet. Bis zu 4 Substitutionen zwischen Zell- und Plasma assoziierten Sequenzen wurden als quasi-identisch gewertet. Insgesamt wurden 621 EnvV1V3-Sequenzen von 9 viremischen, HIV-positiven Probanden analysiert. Dabei wurden bei 4 Probanden Doppelinfektionen mit unterschiedlichen HIV-Stämmen nachgewiesen und quasi-identische Zell-Plasma-Sequenzpaare bei 5 der 9 Probanden entdeckt. Im Detail waren bei Patient H76 3 Sequenzen der $CD4^{niedrig/-}/CD8^-$ T-Zellpopulation quasi-identisch zu 4 Plasmasequenzen. Für Patient H241 waren 17 Sequenzen der $CD4^+CD25^+FoxP3^+$ T-Zellen quasi-identisch zu 2 Plasmasequenzen. Für Patient H304 war 1 Sequenz der $CD4^+CXCR5^+$ T-Zell Population und 1 Sequenz der $CD4^+CXCR5^-$ T-Zellpopulation quasi-identisch zu 1 Plasma assoziierten Sequenz. Bei Patient H442 wurde 1 Sequenz der $CD4^+CXCR5^-$ Zell assoziierten Population quasi-identisch mit 10 Plasma assoziierten Sequenzen, sowie 1 $CD4^{niedrig}/-/CD8^-$ T-Zell Population quasi identisch mit 1 Plasma assoziierten Sequenz gefunden. In Patient H446 wurde 1 $CD4^+CXCR5^-$ T-Zellpopulation quasi-identisch mit 1 Plasma assoziierten Population und 2 weitere $CD4^+CXCR5^-$ T-Zellpopulationen quasi-identisch mit einer anderen Plasma assoziierten Population entdeckt. Alles in allem wurde keine einzelne Zellpopulation identifiziert, die dominant Plasma assoziierte Sequenzen produziert. Diese Ergebnisse unterstützen vielmehr das Konzept, dass alle der untersuchten CD4 T-Gedächtniszellen im peripheren Blut zur Plasmavirusproduktion beitragen und/oder kürzlich infiziert wurden.

Table of Contents

1. Introduction

Human immune deficiency virus (HIV) infection induces an ongoing $CD4^+$ T cell depletion and results in a severe immune system vulnerability towards opportunistic infections, which leads in progress to the fulminant Acquired Immune Deficiency Syndrome (AIDS).
A large meta-analysis (Patel et al., 2014) revealed accurate estimates of per-act HIV transmission risks from various exposures. Parental exposure routes through blood transfusions revealed the highest risk for HIV transmission with 92.5% (9250 transmissions per 10000 exposures to an infected source). Needle sharing of drug users revealed 0.63% (63 transmissions per 10000 exposures to an infected source) and percutaneous needle stick with 0.23% (23 transmissions per 10000 exposures to an infected source). For pathways through semen or mucosal surfaces highest transmission rates for individuals with no condom or antiretroviral therapy use was detected for receptive anal intercourse with 1.38% (138 transmissions per 10000 exposures), followed by 0.11% (11 transmissions per 10000 exposures) for insertive anal intercourse, 0.08% (8 transmissions per 10000 exposures) for receptive penile-vaginal intercourse and 0.04% (4 transmissions per 10000 exposures) for insertive penile-vaginal intercourse. For receptive and insertive oral sex the transmission risk was suggested to be low but not zero, as the estimation is often confounded by complex patterns of sexual exposure such as higher-risk exposures that are accompanied with the same sexual encounter. Among 8965 receptive oral sex acts no transmission could be observed. Although HIV transmission via oral sex would be plausible no precise numeric estimate could be provided. Vertical transmission from mother to child revealed a 22.6% risk (2260 transmissions per 10000 exposures to an infected source).

The viral crossing of the mucosa is enabled via cell free virions, infected cells, or virion associated with Langerhans cells, dendritic cells (DCs) or monocytes/macrophages and effects primarily $CD4^+$ T cells. The virus can also infect natural killer (NK) cells, mucosal and cervical epithelial cells, mastocytes, astrocytes and microglia in the central nervous system, skin fibroblasts and bone marrow stem cells (Dahabieh, Battivelli and Verdin, 2015).

Mechanisms of early events of HIV transmission remain not completely understood as infected individuals are difficult to identify, caused by the frequent absence of clinical symptoms during the beginning of infection. Based on studies of the non-human primate simian immunodeficiency virus (SIV) infection model, lymph nodes (LN) and the gut-associated lymphoid tissue (GALT) play an important role. Particularly in the GALT, within the intestinal lamina propria in transition to an acute status of infection, the virus leads to rapid depletion of $CD4^+$ memory T cells, the major target for HIV replication, (Moir, Chun and Fauci, 2011).

An earlier study of Mattapallil et al. (2005) revealed that the efficient infected memory $CD4^+$ T cell depletion, in mucosal tissue as well as LN and peripheral blood peaks in means at 10-14 days after infection while at least some viral gene expression takes place. The distinct depletion comprises about 80% of the infected cells by various mechanisms. These mechanisms include $CD8^-$T cell killing, that detect infected or virus triggered cytopathic mechanisms and virus-induced cytolysis associated with cellular activation.

In the chronic stage, generally characterized by slow naive and memory $CD4^+$ T cell loss, only low frequencies of HIV infected $CD4^+$ T cells, despite high plasma viremia could be detected in earlier studies. Naive $CD4^+$ T cells were found to be nearly resistant to infection (Nishimura et al., 2004). Hence, the severe memory CD4 T cell depletion is partially masked by this unaffected naive T cell population to infection (Mattapallil et al., 2005).

While peripheral blood contains only 2-5% of all lymphocytes, the gastrointestinal mucosa presents the largest lymphoid organ of the body and is known to be highly permissive for HIV-1 entry, the dominant site for HIV-1 replication and hosts large amounts of proinflammatory, HIV-1 stimulating cytokines (Cerf-Bensussan and Guy-Grand, 1991; Smit-McBride et al., 1998; McGowan et al., 2004). Evidence of $CD4^+$ T cell depletion could be detected in both but significantly larger in mucosa mononuclear cells and less in PBMCs during acute and early HIV-1 infection. Further the mucosal $CD4^+$ T cell depletion occurs before changes in PBMCs appear. Based on investigations in both SIV infection and acute or early HIV-1 infected individuals $CD4^+$ T cell depletion primarily takes place in the effector sites, potentially through decreased cell proliferation, apoptosis or activation-induced cell death (Krammer, 2000). These mechanisms lead to degraded homing of cells to the effector sites from peripheral compartments after antigen detection through direct cytopathic effects by the HIV-1 virus, immune activation-induced cell death or both (McCune, 2001). Expression of HIV-1 RNA was detected predominantly in the lymphoid follicles or T cell zone and germinal center of the gastrointestinal mucosa inductive site, possible through the effector compartment loss at the time of biopsy sampling (Mehandru et al., 2004).

High replication of HIV in memory $CD4^+$ T cells is likely to contribute to high plasma viremia. Plasma peak viremia appears 2 to 3 weeks after exposure if the patient is not undergoing Antiretroviral Therapy (ART), then decreases rapidly and subsequently reaches a steady state or certain set point that indicates the progression of infection (Moir, Chun and Fauci, 2011). Besides the GALT the virus spreads also systemically in all lymphatic tissue (LT) and establishes stable viral reservoirs in latently infected resting and proliferating $CD4^+$ T cells, which present challenges in virus elimination. Likely associated with $CD4^+$ T cell depletion in the GALT, massive depletion of $CD4^+$ T cells in peripheral

blood occurs during the acute phase of infection. Although this depletion is balanced through increased cellular immune activation and a proliferative response, in the ongoing process the hosts immune system constantly loses the ability to restock and fails to completely control viral replication. HIV escapes both cellular and humoral immune responses through its ability to mutate rapidly.

In the chronic phase, continuing immune system activation, differentiation and exhaustion under persistent HIV viremia results in elevated cell turnover and transformations in cellular phenotype and function (Moir, Chun and Fauci et al., 2011). Eventually the infection reaches an advanced level in which the $CD4^+$ T cell depletion is fulminant in the progression of infection, which results in AIDS (Moir, Chun and Fauci, 2011). The rising depletion of CD4 T cells correlates with the increasing loss of immune mediated control against opportunistic pathogens in AIDS, although timing of opportunistic infection is not necessarily related to the degree of the CD4 T cell loss, as some opportunistic infections can occur even in the presence of high CD4 T cell counts. Hence, HIV leads to a generalized immune dysfunction and a low immune reaction to opportunistic pathogens (Lane et al., 1983; Geldmacher and Koup, 2012).

ART inhibits HIV replication to undetectable levels in peripheral blood but still cannot eradicate cellular reservoirs that had been latently infected before therapy started and thus fails to eliminate the virus completely. These latently infected reservoirs enable HIV to rebound when ART is interrupted through virus reactivation. LT is an important source of rebounding virus during treatment interruption (Rothenberger et al., 2015). Based on studies in peripheral blood, resting $CD4^+$ memory T cells can be latently infected by replication competent virus. These cells have been suggested to go over a long-living quiescent state, hosting HIV provirus transcriptionally silent and immunologically inert, once they manage to get away from both cytotoxic $CD8^+$ T cells as well as cytopathic viral effects. Therefore, they represent the fundamental origin of recrudescent infection and the main target for virus elimination (Chun et al., 1997a, b; Finzi, 1997; Wong et al., 1997).
Earlier studies indicated that reservoirs from which infection rebounds after treatment interruption appear seem to originate either from persistent low-level virus production under lower concentration of antiretroviral drugs in LT or from latently infected cell reactivation (Estes et al., 2017). Although ART suppresses viral replication and decreases plasma viremia to the extend below sensitivity of clinical laboratory assay, lower drug levels were found in LT indicating incomplete suppression of viral replication, in comparison to the completely suppressive drug levels in peripheral blood (Fletcher et al., 2014; Rothenberger et al., 2015; Lorenzo-Redondo et al., 2016). Other parameters potentially contributing to ongoing infection include tissue fibrosis, gastrointestinal-tract damage, associated inflammation and $CD4^+$ T cell depletion to a larger extend in advanced stages of infection

at the time of ART initiation (Schacker et al., 2002, 2005, 2006; Estes et al., 2008; Estes, Haase and Schacker, 2008).

In a comparative study, mechanisms for HIV persistence during natural control was investigated in a rare cohort of elite virus controllers in the absence of ART (Walker and Yu, 2013; Migueles and Connors, 2015) characterized by very small frequencies of CD4 T cells hosting HIV DNA (Julg et al., 2010) and replication-competent HIV (Blankson et al., 2007) in contrast to non-controllers. In blood of controllers, provirus emerged clonally expanded within highly differentiated cells and at times expressed virions after stimulation. LT in this population was comprising numerous virus transcripts in T-follicular helper cells and other memory $CD4^+$ T cell populations. Hence, this study revealed HIV persistence notwithstanding virological control through ongoing lymphatic cell infection, viability and recirculation of some of these cells as well as long-term remaining provirus in clonally expanded cells (Boritz et al., 2016).

Consequently, the underlying mechanisms of HIV replication and persistence are not completely understood. Since viral replication and reactivation in latently infected, resting memory T-helper cells potentially enables HIV-1 persistence (Guenthard et al., 2000; Strain et al., 2003; Chun et al., 2005), current ART is not targeting these long-living cells for complete viral extirpation (Chun et al., 1997b; Wong et al., 1997; Finzi, 1997, Finzi et al., 1999; Blankson, Persaud, and Siliciano, 2002). However, it does reconstitute and protect memory $CD4^+$ T cells from further HIV-induced depletion (Guadalupe et al., 2003; Brenchley et al., 2004a). Although ART thus reconstitutes immunity to the majority of opportunistic pathogens, its potential is still limited due to drug side effects, insufficient viral suppression and increasing drug resistance, which leads to the necessity to find solutions with long term curative approach. For these purposes an understanding of potential reservoirs as sources for viral rebound is mandatory.

$CD4^+$ T cells are essential for both supporting B-cells for efficient humoral immunity and supporting $CD8^+$ T cells for cellular immunity to infections as well as immune system regulation. Mediated through various $CD4^+$ T cell subsets these multiple helper functions are underlying a wide range of differentiation pathways (Zhu, Yamane and Paul, 2010). In general, lymphocytes are affected by HIV infection through activation increase, subpopulation differentiation, cellular debility, senescence and decreased rehabilitation potential. Intracellular viral loads of HIV DNA vary in different $CD4^+$ T cell subsets arising by their differentiation level, (Brenchley et al., 2004b; Chomont et al., 2009; Ganesan et al., 2010) functional status (Casazza et al., 2009) and pathogen specificity (Douek et al., 2002; Geldmacher et al., 2010; Geldmacher and Koup, 2012). Therefore, various CD4 T cell populations

seem to differ in their susceptibility towards HIV infection and viral production.

T cells consist of naive (antigen-naive) and memory (antigen-experienced) subsets. Memory T cells are divided into memory stem cells (TSCM), central memory T cells (TCM), transitional memory T cells (TTM) and effector memory T cells (TEM). Naive T -cells (TN) cells appear to be barely receptive to HIV infection. After activation TCM cells survive, proliferate and migrate into secondary LT, differentiate into TEM cells via T cell receptor crosslinking or infrequently in response to homeostatic cytokines. TEM cells present a major target of acute HIV infection and subsequent cell reduction in the mucosa. Functional and transcriptional features of TTM cells are in between TCM and TEM cells (Farber, Yudanin and Restifo, 2014). HIV associated immune dysfunction include impairment of T cell regeneration and differentiation, thymic function loss and LN damage. Analogues processes are ongoing in NK and B cells with resulting innate and adaptive (humoral) immune malfunction. Functional competence rather than the number of cells therefore plays a superordinate role (Moir, Chun and Fauci, 2011). In a study of Chomont et al. (2009) two main steady HIV reservoirs within memory $CD4^+$ T cells had been identified comprising of TCM cells facilitating HIV persistence during ART due to their low cellular proliferation rate and intrinsic capacity to survive for decades (Combadiere et al., 2004; Riou et al., 2007) and TTM cells especially in patients with low $CD4^+$ counts. Both TN and TEM seem to play a minor role in HIV reservoir provision during ART. An additional cellular long-term viral reservoir may be TSCM cells that are characterized by a high self-renewal potential and homeostatic proliferation and are also permissive to HIV infection (Gattinoni et al. 2011; Buzon et al., 2014).
While cell activation status is oblique for viral replication it still remains unknown if distinct memory $CD4^+$ T cell populations play a major role in active virus replication and production or if diverse cell populations may be responsible for ongoing plasma virus contribution.
$CD4^+$ T cells often downregulate CD4 as an effect of productive HIV infection. Hence, productively infected CD4 T cells may be enriched in the $CD4^-/CD8^-$ memory T cell population. Therefore, cells with CD4 low or no surface expression may represent the majority of productively infected cells expressing HIV-RNA and further the predominant cell type supporting ongoing HIV-1 production in peripheral blood or LT (Kaiser et al., 2007). Furthermore, CD4 T cell susceptibility by HIV is affected by cellular activation levels and viral entry receptor expression as well as functional characteristics, such as cytokine and chemokine synthesis, (Geldmacher and Koup, 2012).

CD45, a tyrosine protein phosphatase is regulating the src family kinases, key regulators of signal transduction. CD45RA is expressed on naive CD4 T cells and effector cells while CD45RO represents a memory marker in $CD4^+$ T cells. Furthermore, in HIV infected subjects, central and effector

memory T cells gain expression of CD45RO and lose expression of CD45RA during antigen stimulation (Pinto, Covas and Victorino, 1991).

Interleukin-2 (IL-2) is a cytokine that drives antigen-specific T cell proliferation and survival. Memory $CD4^+$ T cells expressing the interleukin 2 (IL-2) receptor alpha chain (CD25) are known to facilitate efficient HIV infection in LT. In contrast, in vitro productive HIV infection seems to be almost completely suppressed, when IL-2 signaling is being intercepted (Finberg et al., 1991; Ramilo et al., 1993; Goletti et al., 1996; Chou et al., 1997; Boyman and Sprent, 2012). Hence, $CD25^+$ $CD4^+$ T cells were suggested early on to actively produce virus (Ramilo et al., 1993; Borvak et al., 1995).

A high proportion of $CD25^+$ memory $CD4^+$ T cells co-express the transcription factor forkhead box P3 (FoxP3), designated regulatory T cells (Tregs). FoxP3 is essential for the development and function of $CD25^+$ Treg cells (Fontenot, Gavin, and Rudensky, 2003; Hori, Nomura and Sakaguchi, 2003; Khattri et al., 2003), which are known to mediate suppressive effects on T and B cells (Sakaguchi, 2004). To prevent over-reactive immunity, Tregs play a crucial role in immune response modulation. Tregs are capable to migrate into secondary LT and suppress B-cell antibody responses directly (Lim et al., 2005).
Nuclear factor of activated T cells (NFAT) are transcription factors in lymphocytes involved in the gene expression regulation of cytokines, such as IL-2. Various mechanisms had been identified in which FOXP3 inhibits NFAT responsive gene expression. This includes transactivation or repression of NFAT target genes (Zheng and Rudensky, 2007), inhibition of transcriptional activity, modulated by NFAT and nuclear factor „kappa-light-chain-enhancer" of activated B-cells (NfκB) (Bettelli, Dastrange, and Oukka, 2005). FoxP3 further suppresses transcription activity through forkhead binding site interaction (Wu et al., 2006; Marson et al., 2007).

Controversial findings in Treg numbers in HIV patients had been described in various studies with either an increase (Andersson et al., 2005; Epple et al., 2006; Nilsson et al., 2006; Estes et al., 2006) during HIV infection or decrease (Oswald-Richter et al., 2004; Kinter et al., 2004; Apoil et al., 2005; Eggena et al., 2005; Tsunemi et al., 2005) possibly depending on the stage of infection, immune activation status and viral load. However, during acute HIV infection Tregs may function as competent target cells for HIV infection whereas chronification of HIV infection is associated with a cumulative depletion of this cell subset (Holmes, Zhang and Su, 2008).

Tfh cells are a specialized CD4 T cell subset defined by cell surface expression of programed death receptor 1 (PD-1) and the homing receptor for B Cell follicles, CXC chemokine receptor 5 (CXCR5)

(Fazilleau et al., 2009; Qi, 2016). By facilitating signals to B-cells in germinal centers (Gcs), Tfh cells induce B-cell proliferation, isotype switching, somatic hypermutation and affinity maturation during immune responses and contribute to long-lived plasma cell and memory B-cell responses (Mueller, Hoepkin and Lipp, 2003; Crotty, 2014, 2015). For differentiation Tfh cells are depending on dendritic cell (DC) priming of naive antigen-specific CD4 T cells (Crotty, 2011, 2012; Webb and Lintermann, 2017). Resulting from the challenges of studying LT in humans, investigations were depending increasingly on peripheral Tfh cells, a memory $CD4^+$ T cell subset similar to LN Tfh cells. These peripheral Tfh cells express the B-cell follicular homing molecule CXCR5 but show moderate PD-1 levels (Akiba et al. 2005; Morita et al., 2011; Nicholas et al., 2016; Schultz et al., 2017). By 20-25% CXCR5 is expressed of peripheral blood human central memory $CD4^+$ T cells but its function is still not completely understood (Chevalier et al., 2011). Stable levels of peripheral Tfh levels however are considered to be essential for continuing immune system functionality. Tfh cells are infected by follicular DCs transporting infectious HIV virions into lymphatic organs and are considered to function as a predominant viral reservoir with integrated HIV genomes (Vinuesa, 2012; Perreau et al., 2013; Fukazawa et al., 2015; Boritz et al., 2016). In contrast to the GC Tfh cell expansion a remarkable peripheral Tfh decrease had been detected in the chronic course of HIV infection (Pallikkuth, Fischl and Pahwa et al., 2013; Boswell et al., 2014). Further low CD4 T cell counts during HIV infection seem to be accompanied with significantly decreased CXCR5 levels and hence B-cell immunopathology, which leads to impaired humoral immunity under disease progression (Chong et al., 2004; Cagigi et al., 2008).

PD-1 is a cell surface receptor functioning as a down-regulator of the immune system and inflammation, which contributes to self-tolerance. Hence, PD-1 leads to balanced surveillance against autoimmune processes by stimulating apoptosis in antigen specific T cells in LN and contemporaneous repressing apoptosis in regulatory T cells (anti-inflammatory, suppressive T cells; Francisco, Sage and Sharpe, 2010; Fife and Pauken, 2011). Immune cell exhaustion had been noticed in multiple persisting viral infections. PD-1 signaling is linked to virus-specific T cell exhaustion. $PD\text{-}1^+$ memory $CD4^+$ T cells during HIV were detected to have higher levels of HIV-1 DNA than other memory T cell subsets, based on studies of PBMCs (Chomont et al., 2009; Perreau et al., 2013). Amongst ART disruption the PD-1 expression on PBMCs correlates with plasma viremia relapse (Hurst et al., 2015). In contrast suppression of $PD\text{-}1^+$ $CD8^+$ initiation leads to survival and effector response increase against HIV (Moir, Chun and Fauci, 2011). CD4 T cells corresponding to Tfh cell populations, thus $CXCR5^+PD\text{-}1^+$ and $CXCR5^-PD\text{-}1^+$ cells, were found to be significantly enriched in distinct HIV specific CD4 T cells and Tfh cells. Both had been reported to contain the highest viral levels of HIV-infected cells and therefore may represent a predominant virus replication and production site (Perreau et al., 2013).

Ki67 presents a marker for in vivo cell proliferation activity and provides an indirect measurement of HIV infection related proliferation in the different CD4 T cell subsets. Increased $Ki67^{+}CD4^{+}$ T cell levels had been reported in patients suffering from HIV in both LT (Tenner-Racz et al., 1998) and PBMCs (Sachsenberg et al., 1998), in particular during primary HIV-1 infection.

It remains largely unknown which HIV-infected cell populations contribute to plasma viremia. Abundance of viral RNA in $CD4^{-}/CD8^{-}$ T cells from viremic patients observed in a previous study (Kaiser et al., 2007) supports the hypothesis that especially the $CD4^{low}$ and $CD4^{-}$ subsets are a major cell population contributing to plasma virus production. To investigate whether such a particular cell subset contributes to ongoing plasma virus production, we analyzed the HIV highly variable envelope EnvV1V3 region in 4 different memory T cell populations and in plasma from 9 viremic patients for phylogenetic comparisons. Under the assumption that quasi-identical pairs of cell-derived and plasma-derived nucleotide sequences may provide evidence that a specific cell subset recently participated in plasma virion production or became recently infected we analyzed memory T cell populations delineated into $CD4^{low/-}$ and $CD4^{+}$ subsets. $CD4^{+}$ memory T cells were further delineated into Tregs ($CD25^{+}/FoxP3^{+}$) and non-Tregs ($CD25^{-}/FoxP3^{-}$). Non-Tregs were further delineated into Tfh cells ($CXCR5^{+}$) and non-Tfh cells ($CXCR5^{-}$).

2. Material and Methods

2.1 Subjects

Study subjects belonged to the large high-risk cohort of 600 female bar workers, who had enrolled in a prospective study of HIV-1 superinfection (HIV Superinfection Study [HISIS]) in the Mbeya region of Southwest Tanzania. Blood samples were taken, after giving consent, at enrollment between September and December 2000 and every 3 months after, for a period of up to 4 years.

Subject	Infection status by the time of blood sample collection	Subtype	Dual infection evidence
H442	21-24 months	AC	Dual
H76	>36 months	C	Single
H457	>36 months	C	Single
H533	>36 months	CD	Dual
H293	>36 months	C	Single
H304	30-33 months	C	Single
H446	>36 months	C	Single
H241	>36 months	ACD	Dual
H346	24-27 months	CD	Single

Table 1 Infection status, subtype and dual infection evidence for the 9 subjects of the HISIS cohort at Follow up 12 (36 months after participant enrollment)

The HIV-1 status was determined using HIV enzyme-linked immunoassay tests (Enzygnost Anti HIV1/2 Plus, Dade Behring, Liederbach, Germany and Determine HIV ½, Abbott, Wiesbaden, Germany). Discrepancies in the results were resolved by using the Western blot assay (Genelabs Diagnostics, Geneva, Switzerland). Plasma HIV-1 RNA levels were measured with the Amplicor HIV-1 Monitor assay (Roche Diagnostics Indianapolis, IN; Geldmacher et al., 2007b).

9 antiretroviral naive, HIV positive subjects were selected for this study at Follow up 12 (36 months after enrollment). Three of these 9 subjects had seroconverted in the progress of the study. Subtype identification and dual infection evidence was determined (Table 1) by MHA (Multi hybridization assay; Hoelscher et al., 2002).

Ethical approvals were obtained from the Mbeya Regional and National Ethics Committees of the Tanzanian National Institute for Medical Research (NIMR)/Ministry of Health in Dar es Salaam, Tanzania and from the ethics committee of the University of Munich.

2.2 Cell preparation

40 ml of venous blood was drawn from each participant by using CPDA (citrate phosphate dextrose adenine) anticoagulant tubes and processed within less than 6 hours after blood draw as described previously (Geldmacher et al., 2007a). Plasma samples were collected after centrifugation. PBMCs (Peripheral blood mononuclear cells) were isolated by centrifugation on a Ficoll-Hypaque (Biocoll separating solution; Biochrom AG, Berlin, Germany) density gradient. PBMCs and plasma samples were cryopreserved in liquid nitrogen and -80°C respectively before usage in this study.

2.3 PBMC- and Plasma staining for Fluorescence activated cell sorting (FACS)

The following steps had been performed in a S3** laboratory carried out in a laminar flow safety cabinet which enabled a save and coordinated processing to all specimen and applying safety instructions for bloodborne pathogens in the S3** risk group.

2.3.1 Thawing process and cell quantification

In a first step complete media, consisting of 500 ml RPMI medium 1640 (1X; Gibco™ GlutaMAX™), 50 ml heat inactivated FCS from PAA (inactivated fetal calf serum) and 5ml Pen Strep from Sigma, was prepared.
Thawing media was prepared with 4 µl Benzonase per 20 ml complete media. Frozen PBMCs were thawed in a 37°C water bath and then dropwise diluted with 10 ml of thawing media. After centrifugation of 400xg for 5 minutes, which remained the adjustment for all centrifugation steps for FACS preparation, supernatant was discarded, the pellet resuspended in the remaining thawing media and centrifugation was repeated. Supernatant was discarded and the pellet diluted in 2 ml complete media for cell counting. An aliquot of this suspension was stained in a 1:10 dilution with trypan blue for counting by hemocytometer to gain an approximation of the cell quantity (Table 9).

2.3.2 Antibody staining panel for flow cytometric cell sorting

In order to identify different memory $CD4^+$ T cell populations, a staining panel was developed using QA/QC PBMC. The full antibody-fluorochrome panel consisted of 9 markers (Table 2). Fluorochrome labelled antibodies were titrated from 5 µl, 2.5 µl, 1.25 µl, 0.625 µl and 0.3125 µl each to achieve optimal PBMC staining results, defined by the best separation between marker positive and marker negative cell population.

Marker	Fluorochrome
CD3	FITC anti-human from BioLegend
CD4	ECD from Bechman Coulter Inc.
CD8	APC-Alexa Fluor 750 from Beckman Coulter Inc.
CD25	Brilliant Violet 605 anti-human from BioLegend
CD45RA	Brilliant Violet 421 anti-human from BioLegend
PD1 (CD279)	PC5.5 from Beckman Coulter Inc.
CXCR5 (CD185)	PECy7 from BioLegend
Ki67	PE from BioLegend
FOXP3	Alexa Fluor 647 from Beckman Coulter Inc.

Table 2 Full antibody-fluorochrome panel including 9 markers binding various fluorochromes

2.3.3 Cell staining with conjugated fluorochrome antibodies

The remaining suspension was centrifuged, supernatant discarded and the pellet stained with the first fluorochrome antibody panel regarding the optimized titration values including CD3, CD4, CD8, CD25, CD45RA, PD1 and CXCR5 (Table 3).

Fluorochrome antibody marker	Volume in µl
CD3 FITC	1.25
CD4 ECD	1
CD8 APC AF750	0.5
CD25 BV 605	1.25
CD45RA BV 421	2
PD1 PC5.5	0.625
CXCR5 PECy7	0.625

Table 3 Optimized titration values for the Fluorochrome antibody markers

During a 25 minutes incubation in darkness, Fixation/Permeabilization buffer from eBioscience was prepared from Fixation/Permeabilization concentrate and Fixation/Permeabilization diluent in a 1:4 dilution and another Permeabilization Buffer with a concentrate from eBioscience and double distilled water in a 1:10 dilution. After incubation samples were washed with FACS wash buffer (1:25 dilution of 2% FCS in RPMI medium 1640) and centrifuged. The supernatant was discarded and the pellet resuspended in Fixation/Permeabilization buffer and incubated for 30 to 60 minutes in darkness. Permeabilization buffer was then added to each sample. Centrifugation was repeated and supernatant discarded. To enhance the cell isolation, filter paper was used to absorb the residual buffer before adding the fluorochrome antibodies Ki67 and FoxP3 (Table 4) for intracellular staining. Samples were then incubated for 30 minutes in darkness.

Fluorochrome antibody marker	Volume in µl
Ki67 PE	1.25
FoxP3 AF647	2.5

Table 4 Optimized titration values for the Fluorochrome antibody markers

Another washing step with Permeabilization buffer followed. The dilution was centrifuged, supernatant discarded and the pellet diluted in Flow cytometry fixation buffer (Table 5). The Flow cytometry fixation buffer was prepared in a 1:10 dilution with distilled water, while the cells conjugated with fluorochrome antibodies.

Cell quantity	Flow cytometry fixation buffer in µl
$<1x10^6$	150
$1x10^6$ - $5x10^6$	250
$5x10^6$ - 10^7	350
10^7 - $3x10^7$	500

Table 5 Flow cytometry fixation buffer proportion to the cell quantities

2.3.4 Compensation

The full antibody panel was compensated to prevent staining artefacts from overlapping fluorochrome emission spectra. Individual antibody staining was performed on 1 drop of Versa Comp positive control beads (VersaComp Antibody Capture positive beads from Beckman

Coulter Inc.) and a negative control bead (BD CompBeads negative control from Beckman Coulter Inc.) Each of these stained beads was acquired to set up the compensation matrix before fluorescence activated cell sorting.

2.4 Fluorescence activated cell sorting (FACS)

In cooperation with the Institute of Virology, Faculty of Medicine at the Technical University of Munich, the fixed cells were sorted using a FACS Aria III cell sorter from Becton, Dickinson and Company BD Bioscience to sort the populations of interest including $CD4^{low/-}$, $CD25^{+}$/$FOXP3^{+}$, $CXCR5^{+}$ and $CXCR5^{-}$ cells. Sorted populations, constantly preserved on ice in between further processing, were centrifuged at 17900xg for 3 minutes. The entire supernatant was discarded and the pellet stored at -80°C.

For FACS analysis the gating tree for the populations of interest had been analyzed using FlowJo version 10.1r5. From lymphocytes, only singlet cells were selected leaving duplets, triplets and larger cell clusters behind. Single lymphocytes were separated by gating only $CD3^{+}$ cells and further only the $CD45RA^{-}$ cell subset of latter population. From the $CD45RA^{-}$ cell subset $CD4^{+}CD8^{-}$ and $CD4^{low/-}$ and $CD8^{-}$ cells were gated. The $CD4^{low/-}$ cell population was sorted while the $CD4^{+}$ cell subset was further separated into $CD25^{+}$/$FoxP3^{+}$ and $CD25^{-}$/$FoxP3^{-}$. The $CD25^{+}$/$FoxP3^{+}$ population was sorted and the $CD25^{-}$/$FoxP3^{-}$ population was further distinguished into $CXCR5^{+}$ and $CXCR5^{-}$ subsets, while both of them were sorted.

2.5 Amplification and phylogenetic comparison of HIV envelope sequences from plasma and sorted cell populations

2.5.1 Cell lysis

For cell lysis, the pellet was resuspended in 10mM TRIS-HCl buffer (pH 8), made from molecular biology grade pre-made 1M TRIS-HCl buffer (Trizma hydrochloride buffer solution, Sigma) and a 100 µg/ml proteinase K (Roche). For every 10^5 cells 10 µl of latter diluent was layered above the sorted pellet. If the cell number was below $3x10^5$ still a minimum of diluent of 30 µl was used, which was applied for the majority of the sorted cell populations. The suspension was carefully

vortexed and subsequently incubated for a minimum of 1 hour at 56°C. Again, vortexed and briefly centrifuged at 17900xg for a few seconds, the suspension was then incubated another 10 minutes at 95°C to inactivate the proteinase K. To relieve the cell debris from the bottom the suspension was vortexed and centrifuged for 30 seconds, with the same settings as earlier.

2.5.2 Nested Polymerase chain reaction (PCR)

In advance (in cooperation with the Institute of Virology, Faculty of Medicine at the Technical University of Munich), HIV cDNA from 3µl extracted plasma RNA was synthesized using the reverse primer ACD_Env7521R (5'ATGGGAGGGGCATAYATTGC) and the Superscript III reverse transcriptase (Life Technologies, Darmstadt, Germany) according to the manufactures instructions and as described previously (Chachage et al., 2016).

A nested PCR was used to amplify the highly variable envelope region of V1 to V3 (EnvV1V3; Hxb 6559 to 7320) from the previously synthetized cDNA of plasma RNA and DNA of the sorted cell populations. On behalf of this method non-specific binding due to amplification of uncalculated primer binding sites could be reduced. In this two-round PCR the second round amplified the target region on the first-round template. For the detection of the subtypes A, C and D PCR primer pairs had previously been optimized to amplify the EnvV1V3 region (Chachage et al., 2016). All primers for the nested-PCR were diluted in a proportion of 100pmol/µl with water molecular biology reagent and stored at -20°C before usage.

The first-round PCR was performed with 10µl template (supernatant from lyses cells, refered to 2.5.1) in a 50µl reaction mixture, including water molecular biology grade (Sigma-Aldrich), deoxynucleotide trisphosphates (dNTPs 100mM, Thermo Scientific), 10x Rxn buffer (from Invitrogen), MgCl2 50mM from Invitrogen, primer Env 6420 F (5'CATAATGTCTGGGCYACACATGC; Thermo Scientific), primer Env 7521 R (5'ATGGGAGGGGCATAYATTGC; Thermo Scientific) and Platinum Taq Polymerase (Invitrogen; Table 6).

Nested PCR reaction components	Volume in µl
Water molecular biology reagent	19
dNTPs 100mM	4
10x Rxn buffer	5
MgCl2 50mM	3.5
Primer Env 6420 F	4
Primer Env 7521 R	4
Platinum Taq Polymerase	0.5
Total	40

Table 6 First-round PCR reaction content proportions

The second-round PCR was performed with 5 µl of the first-round product in a 50 µl reaction mixture (Table 7), including the same contents as for round one and the primers Env 6559 F (5'GGGAYSAAAGCCTAAARCCATGTG; Thermo Scientific) and Env 7320 R (5'GTTGTAATTTCTRRRTCCCCTCC; Thermo Scientific).

For both rounds a Thermocycler, TProfessional Standard from Biometra was used. The polymerase was activated for 10 minutes at 95°C and amplified for 41 cycles (30 seconds at 94°C, 30 seconds at 53°C and 1 minute and 30 seconds at 72°C) followed by a final elongation step of 7 minutes at 72°C.

Nested PCR reaction components	Volume in µl
Water molecular biology reagent	28.25
dNTPs 100mM	4
10x Rxn buffer	5
MgCl2 50mM	3.5
Primer Env 6559 F	2
Primer Env 7320 R	2
Platinum Taq Polymerase	0.25
Total	45

Table 7 Second round PCR reaction content proportion

3 µl of the second PCR reaction were resuspended in 7 µl Blue Juice to screen for the amplicon of interest using gel electrophoresis by use of a 250 base pair (bp) DNA ladder from Invitrogen. The

PCR product imaging after gel electrophoresis was visualized through Grab-IT, Version 2.59.
As the deployed Taq Polymerase didn't include proof reading the positive-control template Du422 (clone 1 [SVPC5]; Li et al., 2006) was endpoint diluted equal to Chachage et al. (2016) to evaluate the error rate of the nested-PCR on behalf of a 10-fold-dilution series and amplified as described earlier. From the last detectable dilution step the EnvV1V3 product was then cloned as described earlier. The sequences of 21 clones had been analyzed and compared to the original Du422 template sequence. The variability appeared to be in a minimal extend of only 0-4 base pare errors through nested-PCR processing and therefore, quasi-identical assumption was assumed within this range at the Chachage et al. (2016) study and was transferred to this study.

2.5.3 Gel Extraction

According to the protocol by the manufacturer the Qiagen Gel extraction kit was used to extract the 760 bp EnvV1V3 amplicon from the Gel according to manufacturer instructions. Depending on the resulting DNA quantity 10 µl up to 47 µl of the second-round PCR product were resuspended in Blue Juice in a 1:10 ratio in favor for the product and transferred to gel electrophoresis. A 250 bp ladder was used to identify the region of interest. To avoid DNA contamination the entire equipment used was decontaminated with 2% hypochlorite solution (Merck). Precisely the 760 bp bands were selectively excised out with a scalpel, the gel slice weighted and suspended in QG solubilization buffer (QIAquick) in a proportion of 3 volumes buffer to 1 volume gel (100 mg gel~100 µl). This suspension was left at 50°C for 10-15 minutes until it was completely diluted. Next, 1 volume gel Isopropanol (2-Propanol molecular biology grade, AppliChem) was resuspended.
The entire volume was disposed into QIAquick spin columns, each no more than 400 mg of gel volume. For cloning 500 µl QG solubilization buffer was added to each column, the whole suspension centrifuged at 17900xg for 1 minute and the supernatant discarded.
Subsequently, 750 µl PE wash buffer was added to each QIAquick spin columns. To use the DNA for salt sensitive cloning, the columns were incubated for 2-5 minutes after buffer addition, centrifuged twice with the same setting as before and supernatant discarded to remove residual wash buffer. The columns were transferred to new microcentrifuge tubes. For DNA extraction 50 µl water molecular biology reagent (from Sigma) was added to the center of the QIAquick membrane. To increase the yield of purified DNA, the suspension was incubated for 2 minutes before centrifugation. To proof successful extraction 5 µl of the DNA suspension was tested again on gel electrophoresis.

2.5.4 Cloning reaction of purified DNA fragments

The TOPO cloning kit (Invitrogen by Life Technologies, Darmstadt, Germany) was used for cloning. 0.5 µl salt solution (1.2 M NaCl and 0.06 M $MgCl_2$) and 0.5 µl TOPO (PCR 4-TOPO vector), including the precut plasmid vector pCR™ 4-TOPO were added to 2 µl of the before extracted DNA fragment and incubated for 15 minutes at room temperature. The reaction was put on ice and 25 µl of One Shot TOP10 chemically competent *E.coli* (Invitrogen) were added, gently mixed and incubated for 20 minutes, remaining on ice. Next, the suspension was heat-shocked at 42°C for 30 seconds and replaced on ice afterwards. 250 µl of room temperature S.O.C medium was added and the entire suspension shaken for 60 minutes. 50 µl of each transformation was spread on an ampicillin plate and left in an incubator at 37°C overnight for approximately 18 hours.

After the 18 hours incubation time clones were picked and resuspended in 2 ml of LB ampicillin medium (containing 100 µg/ml ampicillin) and cultured overnight at 37°C on a shaker for approximately 18 hours again. The plates were wrapped with Parafilm, when clones were left and stored at 4°C, if further clones were needed.

2.5.5 Plasmid preparation

The PureLink Quick Plasmid MiniPrep Kit of the TOPO cloning kit (Invitrogen by Life Technologies, Darmstadt, Germany) was used for plasmid preparation. The overnight 2 ml cultures were centrifuged at 6800xg for 3 minutes, supernatant was discarded and the remaining supernatant removed with filter paper. After resuspension, 250 µl Buffer P1 (RNase A Buffer) and 250 µl buffer P2 (lysis buffer) was added and the suspension shaked 4-6 times within less than 5 minutes. Another 350 µl of buffer N3 (neutralization buffer) was added and again everything shaked 4-6 times. The whole suspension was centrifuged at 17900xg for 10 minutes. From the supernatant 800 µl were transferred to a new spin column, centrifuged at 17900xg for approximately 30-60 seconds and supernatant discarded. 500 µl of buffer PB (binding buffer) was added, centrifuged for 30-60 seconds with the same setting as latter and supernatant discarded. 750 µl of buffer PE (wash buffer) was added and again the same centrifugation and supernatant discard followed. Another centrifugation at 17900xg for 1 minute was performed to remove residual wash buffer. The columns were displaced to new tubes and 50 µl water molecular biology reagent (from Sigma) was added and everything incubated for 2 minutes before a further centrifugation at 17900xg for 1 minute.

2.5.6 Verification of cloned insert

To proof the successful cloning, the restriction enzyme EcoRI (10U/µl, Fermentas) was used to verify the insert of interest. The mastermix was prepared including 3.5 µl water molecular biology reagent (Sigma), 1 µl 10x EcoRI buffer (restriction buffer) and 0.5 µl EcoRI.
5 µl of the DNA extract was resuspended in 5 µl mastermix and incubated at 37°C for 90 minutes. From this suspension, another control gel electrophoresis was performed to proof for the amplified sequence at 760 bp.

2.5.7 Molecular Sequencing

16 to 27 EnvV1V3 containing clones per cloning reaction from up to 5 compartments (different cell subsets or plasma) were sequenced unidirectional using M13rev primers at Eurofins Genomics (Ebersberg, Germany). Collectively 621 EnvV1V3 sequences from 9 subjects were analyzed.

2.5.8 Phylogenetic comparison of HIV envelope sequences from Plasma and sorted cell subsets

Individual EnvV1V3 sequences from each clone and subset were analyzed for each individual using MEGA version 6.06. Regarding to the predicted nucleotide sequence of the reference alignment extracted from the Los Alamos HIV database (http://www.hiv.lanl.gov/content/sequence/NEWALIGN/align.html) nucleotide sequences were aligned as former described (Pollakis et al., 2015). The alignments were manually corrected by complementation and primer cut off.

In cooperation with Dr. Georgios Pollakis at the Department of Clinical Infection, Microbiology and Immunology (CIMI), the NIHR Health Protection Research Unit in Emerging and Zoonotic Infections (HPRU EZI) at the Institute of Infection and Global Health (IIGH) in Liverpool the evolutionary history was inferred on behalf of the maximum likelihood method based on the general time-reversible substitution model (GTR+G; Nei and Kumar, 2000) and was rooted in previous outbreaks. The tree with the highest log likelihood is presented for each analysis. Next to the branches the percentage is presented in which the associated taxa clustered together. For the heuristic search the initial tree(s) was/were obtained automatically by applying neighbor-joining

and BioNJ algorithms to a matrix of pairwise distances, estimated using the maximum composite likelihood (MCL) approach and further selecting the topology with superior log likelihood value (Fig. 3, Fig. 4).

The nucleotide variation of plasma to the cell population of interest (Fig. 5 and Fig. 6) was plotted using Prism Version 4.0 software (GraphPad, Inc.) including the interquartile range of 25% to 75% with the mean distance for each participant.

For identification of quasi-identical sequences, the phylogenetic maximum likelihood tree analysis in MEGA 6.06 was examined in order to find clusters of plasma-derived sequences with cell-derived sequences. Subsequent manual counting of nucleotide differences between those sequences enabled to proof for quasi-identical or close to quasi-identical clusters as defined by the Du422 cut off of 4 nucleotide substitutions as described in Chachage et al. (2016).

3. Results

3.1 $CD4^+$ T cell counts and Viral Load

Subject	$CD4^+$ count (cells/µl)	Viral Load (copies/ml)
H442	463	51200
H76	227	15800
H457	348	>750000
H533	473	5890
H293	253	Not available
H304	350	69200
H446	143	258000
H241	359	Not available
H346	251	119000

Table 8 $CD4^+$ T cell counts (cells per µl) and viral load (copies per ml) at Follow up 12 for the 9 selected subjects of the HISIS cohort

For the phylogenetic characterization of HIV EnvV1V3 sequences within the sorted memory T cell populations, PBMCs of 9 viremic HIV positive female individuals from the HISIS cohort were analyzed in this study. These patients were HIV infected between 21 to over 36 months, a median $CD4^+$ T cell count of 348 cells/µl and a median plasma viral load of 69200 copies/ml (Table 8).

3.2 Cell counting before FACS

Subject	Number of counted PBMCs / 2 ml
H442	10.9 $x10^6$
H76	6.15 $x10^6$
H457	9.75 $x10^6$
H533	13.5 $x10^6$
H293	36.3 $x10^6$
H304	18 $x10^6$
H446	9.5 $x10^6$
H241	8.3 $x10^6$
H346	7.7 $x10^6$

Table 9 Quantities of cells for each subject after PBMC thawing process before FACS

After thawing, 6 to 36 million cells with a median of $9.75x10^6$ PBMCs per 2 ml were counted to receive an approximation for the cell numbers for later cell staining and sorting (Table 9).

3.3 FACS

3.3.1 Gating tree

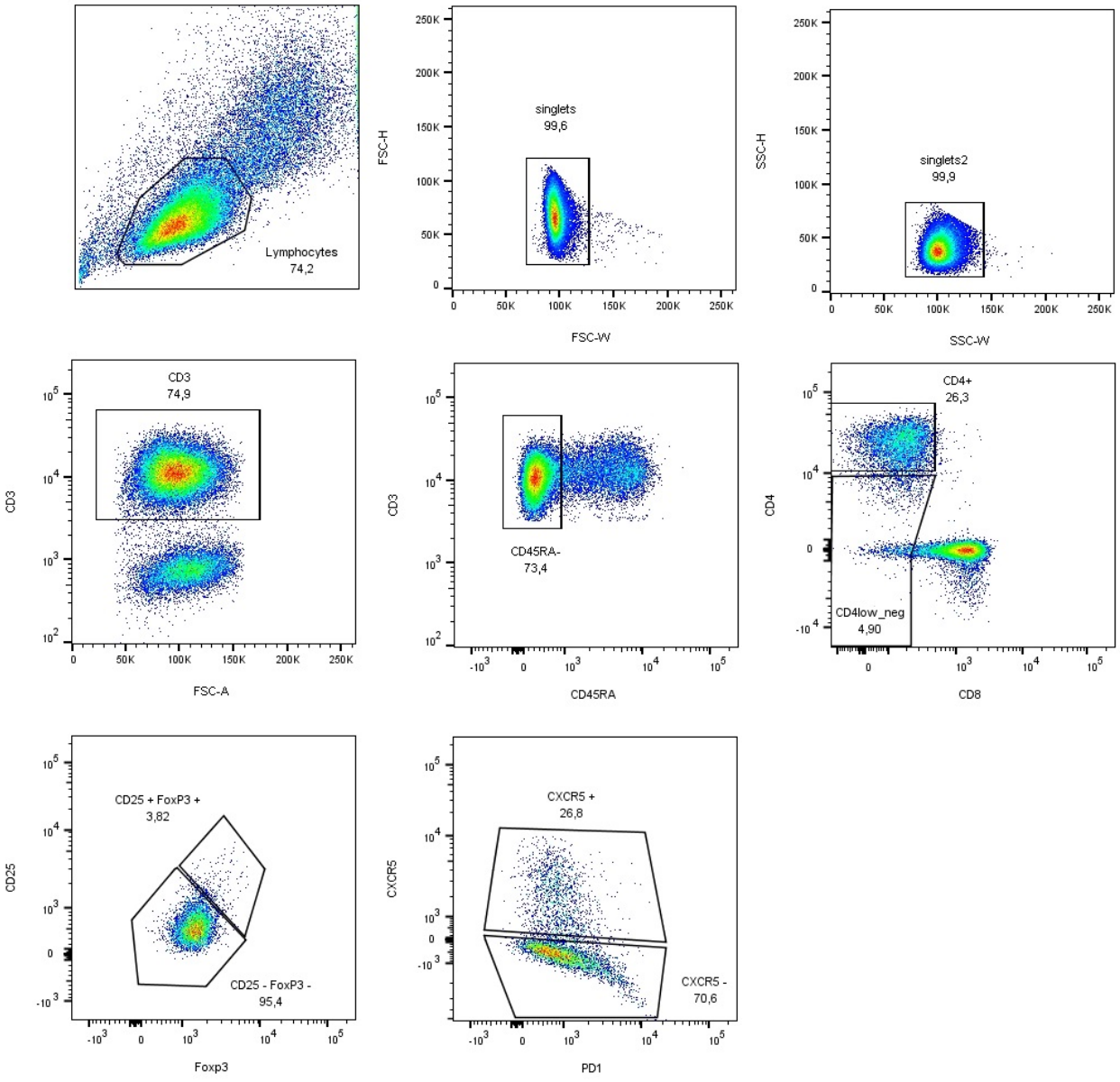


Fig. 1 Representative gating tree during cell sorting. Lymphocytes were gated from all monocyte cells as shown for subject H442 from which only singlets with an amount of about 99% were selected. About 75% of T-lymphocytes ($CD3^+$ cells) were selected from singlets and further divided into naive ($CD45RA^+$) and memory T cells ($CD45RA^-$), latter with about 73%. From the memory T cells the $CD4^{low/-}$ population with approximately 5% was sorted. From $CD4^+$ T cells, Tregs ($CD25^+/FoxP3^+$) with about 4% were sorted. The $CD25^-/FoxP3^-$ cells were then distinguished and separately sorted into Tfh containing cells with about 27% ($CXCR5^+$) and non-Tfh cells ($CXCR5^-$) with 71%

Plasma virus sequences were phylogenetically compared with cell-derived sequences of $CD4^{low/-}$, $CD4^+CD25^+/FoxP3^+$, $CD4^+CXCR5^+$ and $CD4^+CXCR5^-$ memory T cell subpopulations.

Primarily from all detected monocyte cells, selectively lymphocytes, in particular only singlets were

gated, excluding duplets, triples and larger cell clusters. The gated singlet T-lymphocytes ($CD3^+$) were separated from other cell populations, such as dendritic cells, natural killer cells and B-lymphocytes and divided into naive ($CD45RA^+$) and memory T cells ($CD45RA^-$). From the memory T cell population $CD4^{low/-}/CD8^-$ cell subsets were sorted. From the $CD4^+/CD8^-$ cell population Treg cells ($CD25^+/FOXP3^+$) were sorted and non-Treg cells were further separately sorted by dividing them into Tfh cells ($CXCR5^+$) and non-Tfh cells ($CXCR5^-$; Fig. 1).

3.3.2 Cell quantification after FACS

Subject	CD4 $^{low/-}$	$CD25^+/FoxP3^+$	$CXCR5^+$	$CXCR5^-$
H442	$2.8x10^5$	18175	$2.8x10^5$	$7.8x10^5$
H76	80000	$4.6x10^4$	$3.4x10^5$	$4.9x10^5$
H457	$1.8x10^5$	$1.9x10^4$	$1.1x10^5$	$2.8x10^5$
H533	$1.4x10^5$	10153	64769	$3.2\ x10^5$
H293	$2.9x10^5$	36246	62066	$2.2\ x10^5$
H304	81945	43645	85053	17047
H446	31940	6905	20024	92665
H241	57348	7743	2211	$1.8\ x10^5$
H346	84272	8935	63561	$2.2x10^5$

Table 10 Quantities of all sorted populations of interest for each subject in absolute numbers. Quantities of the sorted populations of interest without successful EnvV1V3 amplification are colored in black, sorted populations of interest with successful EnvV1V3 amplification are colored in blue.

From the 6 up to 36 million cells per 2 ml that had been counted after thawing (Table 9), approximately $2x10^3$ to $7.8x10^5$ cells were sorted for the 4 subpopulations of interest for each subject (Table 10). For the $CD4^{low/-}$ subset the median over all subjects was 81945, for the $CD25^+/FOXP3^+$ subject 10153, for the $CXCR5^+$ subset 62813.5 and for the $CXCR5^-$ subset 205275.5 (Table 10).

3.4 Amplification of EnvV1V3 DNA using a nested PCR

Values for the subpopulations that had not been successfully amplified for the EnvV1V3 region (black colored values) and values for the subpopulations that had been successfully amplified for the EnvV1V3 region (blue colored values) are presented in Table 10.

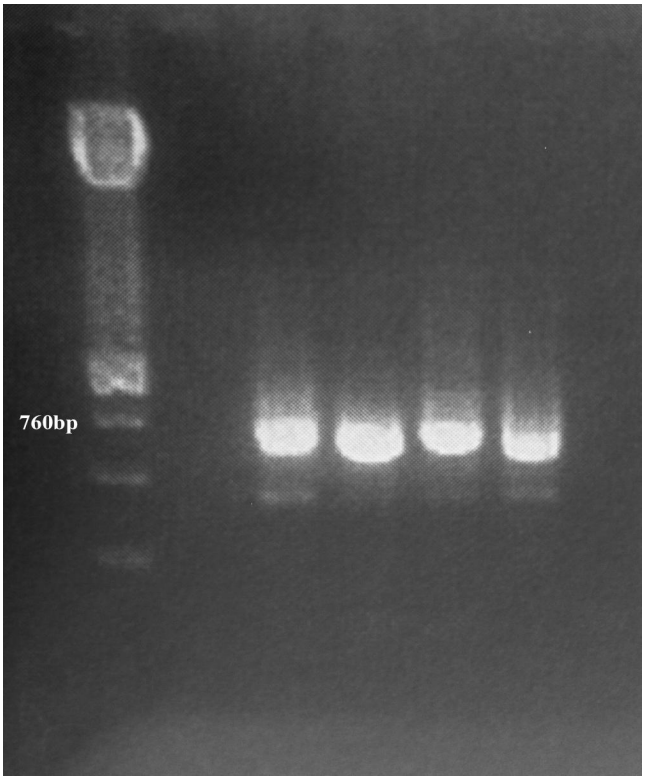


Fig. 2 The EnvV1V3 sequence, comprising approximately 760 bp, was verified in all subjects in various subpopulations, here in all 4 populations of interest for subject H442

The 760 bp target sequence of EnvV1V3 was successfully amplified on behalf of nested PCR processing within the V1 to V3 region of Hxb 6559 to 7320 for all of the 9 subjects in diverse distribution within the 4 cell populations of interest (Fig.2, Table 10). Particularly for Treg cells the EnvV1V3 could be amplified and sequenced for 7 out of 9 subjects. For the Tfh cells and non-Tfh cell populations EnvV1V3 was detected for both in 6 out of 9 subjects. An EnvV1V3 amplicon was detected for the $CD4^{low/-}$ subpopulation in only 3 subjects.

3.5 Phylogeny of EnvV1V3 sequences from plasma and cell derived subsets

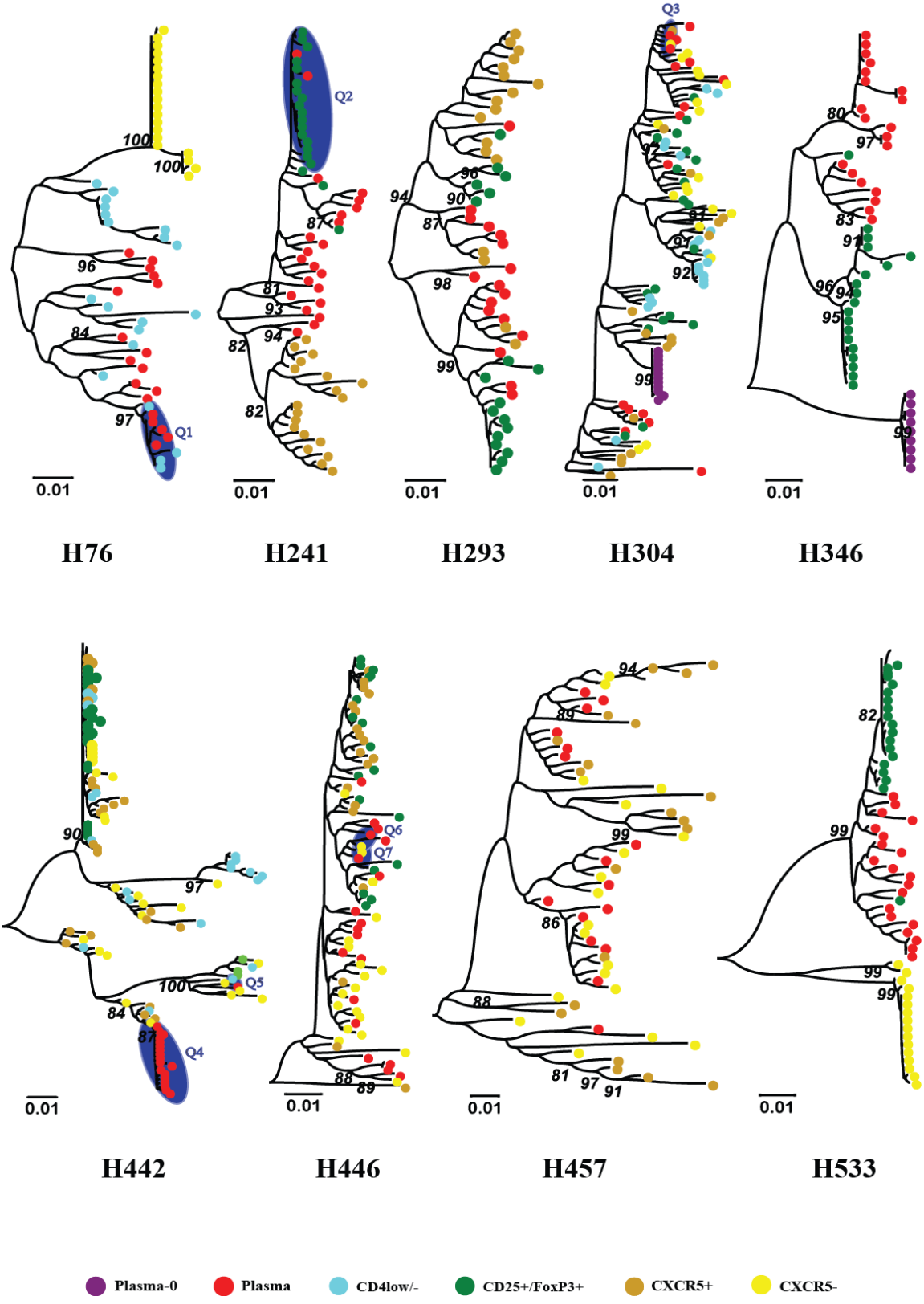


Fig. 3 Phylogenetic relationship of clonal EnvV1V3 sequences derived from plasma and from $CD4^{low/-}$, $CD4^{+}CD25^{+}/FoxP3^{+}$, $CD4^{+}CXCR5^{+}$ and $CD4^{+}CXCR5^{-}$ cell subsets. Phylogenetic distances between various subpopulations in percent are presented with individual probability estimations of the distribution marked on the phylogenetic branches. Plasma-0 represents the very early stage of infection, less than 3 months after infection with a primarily seronegative status.

The individual phylogenetic trees of each subject visualize the genetic variations of clonal EnvV1V3 sequences within plasma at the early stage of infection, specifically the first seropositive follow-up, less than 3 months after infection with a primarily seronegative status (Plasma-0; H304, H346 and H605; Fig. 3, Fig. 4), within plasma at Follow up 12, between plasma to cell and within variations of the selected cell populations. With a probability estimation between 80 up to 100 the phylogenetic distribution is represented for each subject (Fig. 3). Clusters of the $CD4^{low/-}$ -derived sequences with plasma-derived sequences were found in 2 of the 9 subjects including a cluster of 3 $CD4^{low/-}$ derived sequences quasi-identical with 4 plasma populations (H76; Q1; Fig. 3; Table 11) and 1 $CD4^{low/-}$ derived sequences quasi-identical with 1 plasma population (H442; Q5, Fig. 3; Table 11). The largest quasi-identical pair was detected for the $CD25^{+}$/$FoxP3^{+}$ subset with 17 cell-derived sequences clustering with 2 plasma-derived sequences (H241; Q2; Fig. 3; Table 11). 1 quasi-identical pair was detected for 1 Tfh ($CXCR5^{+}$) cell-derived sequence clustering with 1 plasma derived sequence (H304; Q3; Fig. 3; Table 11). Most quasi-identical pairs were detected for non-Tfh ($CXCR5^{-}$) cells with a large cluster of 1 cell-derived sequence with 10 plasma-derived sequences (H442; Q4; Fig. 3; Table 11), 2 clusters of 1 cell-derived sequence with 1 plasma-derived sequence (H304; Q3 and H446; Q6; Fig. 3; Table 11) and 1 cluster of 2 cell-derived sequences with 1 plasma-derived sequence (H446; Q7; Fig. 3; Table 11). No quasi-identical clusters could be detected for 4 subjects (H457, H533, H293 and H346; Fig. 3; Table 11).

Subject	Plasma	$CD4^{low/-}$	$CD25^{+}$/$FoxP3^{+}$	$CXCR5^{+}$	$CXCR5^{-}$	Quasi-identical pair number
H76	4	3(1-3S)	0	0	0	Q1
H241	2	0	17(0-3S)	0	0	Q2
H304	1	0	0	1(3S)	1(3S)	Q3
H442	10	0	0	0	1(4S)	Q4
	1	1(4S)	0	0	0	Q5
H446	1	0	0	0	1(3S)	Q6
	1	0	0	0	2(3S)	Q7
H457	0	0	0	0	0	
H533	0	0	0	0	0	
H293	0	0	0	0	0	
H346	0	0	0	0	0	

Table 11 Quantities of quasi-identical plasma-derived sequences and distinct cell-derived sequences selectively for the 4 investigated subpopulations. No quasi-identical clusters could be detected for the 4 grey highlighted subjects. The quantities of substitutions are presented in parenthesis (S).

With the endpoint diluted molecular clone of the subtype isolate Du422 clone 1 the PCR-related sequence background variation was controlled. The 21 Du422 sequences contained either 0, 1 or 2 nucleotide substitutions, neither insertions nor deletions in 95% of the amplicons. Hence, up to 4 substitutions between cell- and plasma-derived sequence variants were considered as quasi-identical for the present study, likewise Chachage et al. (2016). In 2 subjects (H76, H533) deletions in the V-loop were detected for both in a relatively similar region. Additionally, a frameshift was found for H76 demonstrating the region to be defective.

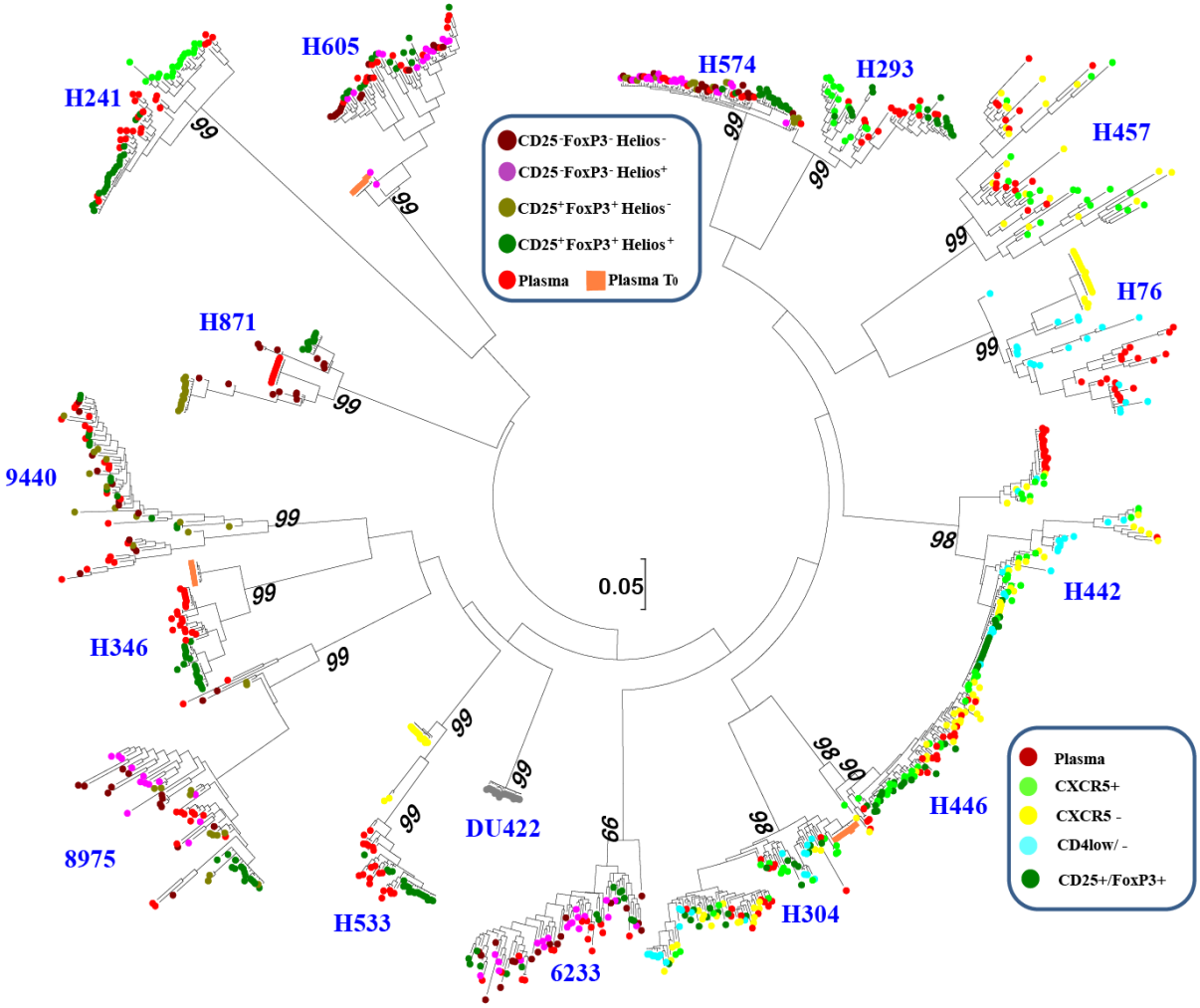


Fig. 4 Phylogenetic tree including subjects of the study of Chachage et al. (2016) with $CD25^{+}/FoxP3^{+}Helios^{+}$, $CD25^{+}/FoxP3^{+}Helios^{-}$, $CD25^{-}/FoxP3^{-}Helios^{+}$ and $CD25^{-}/FoxP3^{-}Helios^{-}$ sorted populations and the present study with $CD4^{low/-}$, $CD4^{+}CD25^{+}/FoxP3^{+}$, $CD4^{+}CXCR5^{+}$ and $CD4^{+}CXCR5^{-}$ sorted populations and plasma (Plasma) at Follow up 12 and at onset of infection (Plasma T_0, first seropositive follow-up). Phylogenetic distances between various subpopulations in percent are presented with individual probability estimations of the distribution marked on the phylogenetic branches.

The wide phylogenetic distribution of cell-derived sequences in regard to plasma-derived sequences for the investigated 9 subjects indicate various memory T cell-subsets rather than a distinct cell subset may contribute to plasma viremia as the data does not indicate a coherence of a predominant cell subset derived sequences clustering with plasma virus sequences.
The phylogenetic tree including the results of the study of Chachage et al. (2016, Fig. 4) provides an overview of the nucleotide sequence distribution. Regarding the hypothesis to find evidence for a distinct cell subset contributing to plasma viremia, comparing the phylogenetic distribution from $CD4^{low/-}$-derived, Tfh-derived, Treg-derived and non-Treg-derived nucleotide sequences respectively to the plasma-derived sequences, no distinct correlation could be detected for a certain cell population contributing to plasma viral replication.

3.6 HIV nucleotide variations between plasma- and cell-derived sequences

The nucleotide distance of plasma to cell differed with a variation between approximately 0 up to 25%, the mean values between approximately 1 to 8% and the interquartile ranges, the middle 50% between the upper and lower quartile, varied approximately between 1 to 12.5%.
Significant increased nucleotide variation in correlation to the infection duration could not be detected with continuity as the distribution of nucleotide variation in subjects with shorter infection duration showed partially larger nucleotide variation (H442; 21-24 months infected; Fig. 5) in comparison to subjects with longer infection duration (H346; 24-27 months infected; Fig. 5). Further, for all subjects, infected over 36 months, varying nucleotide distribution with significant differences were detected (H457 with large nucleotide variation in comparison with H293, H446, H241 and H76; Fig. 5). A distribution of 2 clusters with significant diverse distances from the plasma-derived nucleotide sequences was detected for H533 and earlier in the study of Chachage et al. (2016) for H8710. Although these findings may be explained by dual infection evidence for H533 (Table 1, Fig. 9), no such evidence for dual infection were detected for H8710 (Table 13).

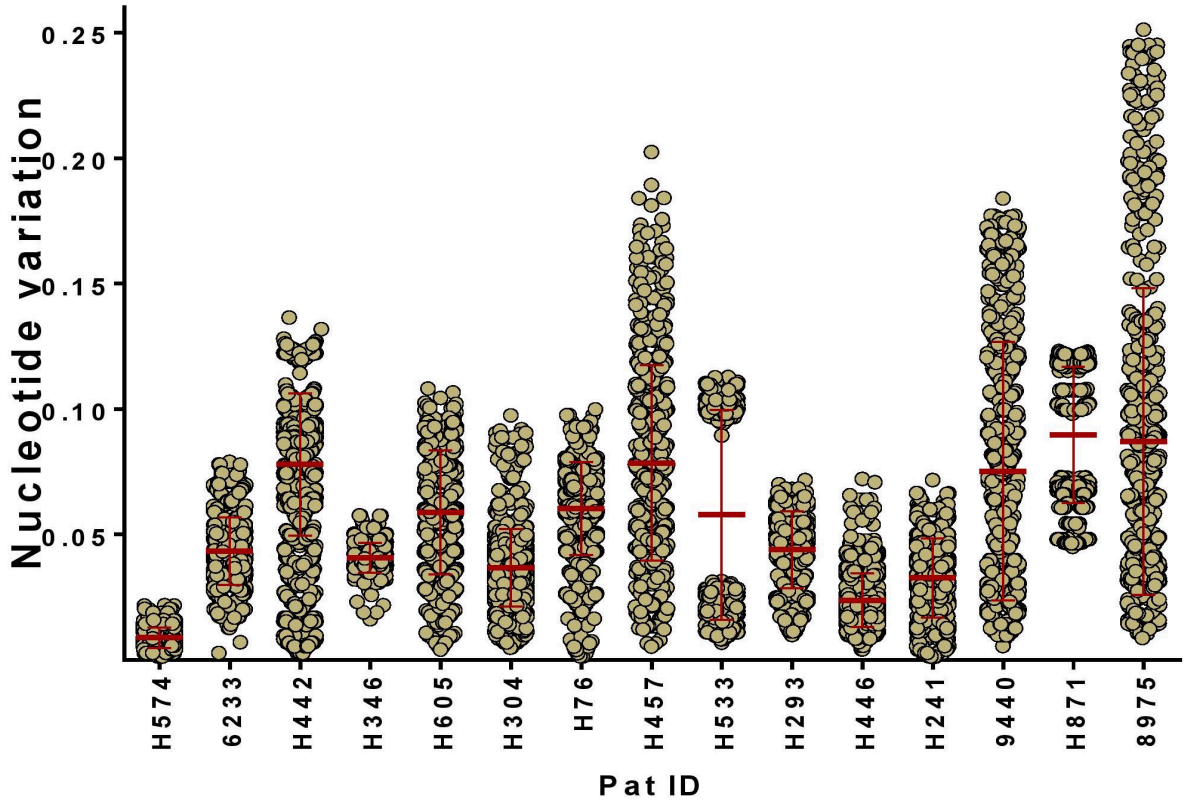


Fig. 5 Nucleotide distance variation (in %) of plasma-derived nucleotide sequences for the 9 subjects of this study and the 6 subjects of the study of Chachage et al. (2016). Each value represents a plasma-derived nucleotide sequence at its distance to a distinct cell-derived nucleotide sequence. The horizontal red bars indicate the mean value of nucleotide variation, the vertical bar represents the interquartile range of nucleotide variation.

With the shortest infection time of 21-24 months a large cluster of 1 Tfh-cell-derived nucleotide population with 10 plasma-derived nucleotide sequences (H442; Q4; Fig. 3; Table 11) was detected but also for subjects infected over 36 months large populations of 17 $CD25^{+}$/$FoxP3^{+}$ cell-derived nucleotide sequences clustering with 2 plasma-derived sequences could be detected (H241; Q2; Fig. 3; Table 11).

Dual infection evidence may be an explanation for the discontinuity of nucleotide variation increase over time. Various subjects were previously found to be dual infected by the time of this study onset (H574, H605, H442, H533, 241; Table 1) while dual infection evidence was detected during the phylogenetic analysis during this study (H8975, 9440, H457, H346; Fig. 7; Fig. 8; Fig. 9). Hence, viral sequence variation seems to underline multiple variables, which may have an impact on the genetic alteration instead of a single aspect, such as infection duration or dual infection evidence.

	H76	H241	H304	H442	H446	H457	H533	H293	H346
Plasma Nucleotide variation maximum in %	8	6	10	5	5	15	4	6	6
Plasma Mean in %	4	3	3	1	3	5	2	3	3
Plasma Interquartile range in %	2-6	2-4	2-5	0-2	2-4	3-8	1-3	2-5	1-4
$CD4^{low/-}$ Nucleotide variation maximum in %	9		10	14					
$CD4^{low/-}$ Mean in %	5		4	8					
$CD4^{low/-}$ Interquartile range in %	3-7		3-5	5-12					
$CD25^{+}/FoxP3^{+}$ Nucleotide variation maximum in %		6	10	13	6		4	7	6
$CD25^{+}/FoxP3^{+}$ Mean in %		3	3	9	3		2	4	4
$CD25^{+}/FoxP3^{+}$ Interquartile range in %		2-4	2-5	8-10	2-4		1-3	2-6	3-5
$CXCR5^{+}$ Nucleotide variation maximum in %		7	10	13	8	21		7	
$CXCR5^{+}$ Mean in %		4	3	7	3	9		5	
$CXCR5^{+}$ Interquartile range in %		3-5	2-5	4-10	2-4	5-12		3-6	
$CXCR5^{-}$ Nucleotide variation maximum in %	10		10	14	6	19	12		
$CXCR5^{-}$ Mean in %	7		3	7	3	7	10		
$CXCR5^{-}$ Interquartile range in %	6-8		2-5	4-9	2-4	4-11	9-11		

Table 12 Nucleotide variation maximum, Mean and Interquartile range for within plasma and plasma to the distinct cell population of interest for all 9 subjects (in %)

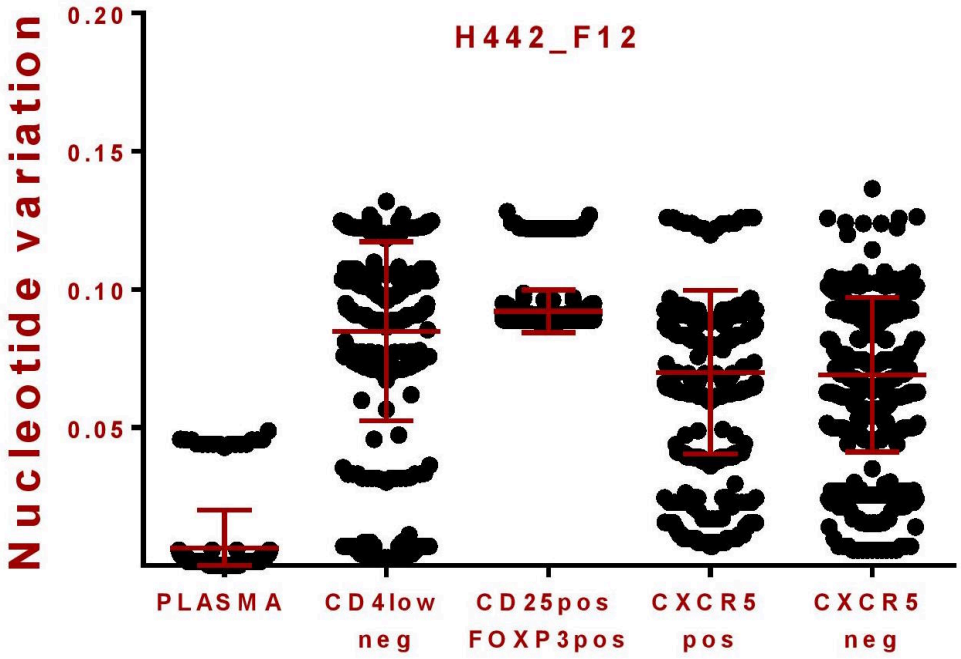


Fig. 6 Nucleotide variation within plasma and for all plasma-derived sequences to each cell population of interest (in %).

Respectively nucleotide variations within plasma and the individual cell populations to the plasma are presented for each subject (Fig. 6, Appendix: S2, S4, S6, S8, S10, S12, S14, S16). The nucleotide variation maximum, mean and interquartile range are presented for each subject (Table 12).

Nucleotide distance maxima within plasma populations with approximately 15% variation, from $CXCR5^{+}$populations to plasma with approximately 21% variation and from $CXCR5^{-}$ populations to plasma with approximately 19% variation were detected with the largest maxima for H457 (>36 infected; Appendix: S10; Table 12).

For the $CD4^{low/-}$ population with approximately 14% as well as for the $CD25^{+}/FoxP3^{+}$ population with approximately 13% largest nucleotide variation maxima were detected for H442 (21-24 months infected; Fig. 6; Table 12). For latter subject two considerably independent plasma subpopulations were detected, one with a minimal nucleotide variation maximum of approximately 1% and a second of an approximate 5% nucleotide variation maximum. These findings may be explained by dual infection.

Considering the assumption that time of infection may be accompanied with nucleotide variation increase the findings of this study are controversy. Plasma nucleotide variation maxima were lowest for H533 (>36 months infected; Appendix: S12; Table 12) with approximately 4% and highest for H457 (>36 months infected; Appendix: S10; Table 12) with approximately 15%. Within the $CD4^{low/-}$ population lowest nucleotide variation maxima with approximately 9% were detected for H76 (>36

months infected; Appendix: S2; Table 12) and largest nucleotide variation maxima with approximately 14% were detected for H442 (21-24 months infected; Fig. 6; Table 12). Latter nucleotide variation discrepancy regarding the infection duration may be explained by dual infection evidence for H442. Similar discrepancy could be detected for the $CD25^{+}/FoxP3^{+}$ population were lowest nucleotide variation maximum of approximately 4% was found for H533 (>36 months infected; Appendix: S12; Table 12) and largest nucleotide variation maximum for H442 with approximately 13% (21-24 months infected; Fig. 6; Table 12). For the $CXCR5^{+}$ populations lowest nucleotide variation maxima with approximately 7% were found for H293 (>36 months infected; Appendix: S4; Table 12) and H241 (>36 months infected; Appendix: S14; Table 12) and largest nucleotide variation maxima for H457 with approximately 21% (>36 months infected; Appendix: S10; Table 12). For the $CXCR5^{-}$ population lowest nucleotide variation maximum was found for H446 with approximately 6% (>36 months infected; Appendix: S8; Table 12) and largest nucleotide variation maximum for H457 with approximately 19% (>36 months infected; Appendix: S10; Table 12).

This controversy of nucleotide variation increase, not necessarily being accompanied with infection prolongation, may partially be explained by dual infection evidence (H442, H457, H533, H346 and H241) but additionally seem to underline the assumption, that multiple factors may be involved in nucleotide alterations. Hence, multiple cell populations rather than a single cell subset may contribute to plasma viremia.

3.7 Coinfection

During analysis, processing phylogeny indicated coinfection for distinct subjects of this study, as their phylogenetic distance could not be explained by nucleotide variation over time but rather coinfection.

Through further analysis within the viral sequence, evidence of dual infection was confirmed in the V1C2 region (HxB2 6615→7025) for H442 with a significant distance of 2 clusters of nucleotide sequences (Fig. 7).

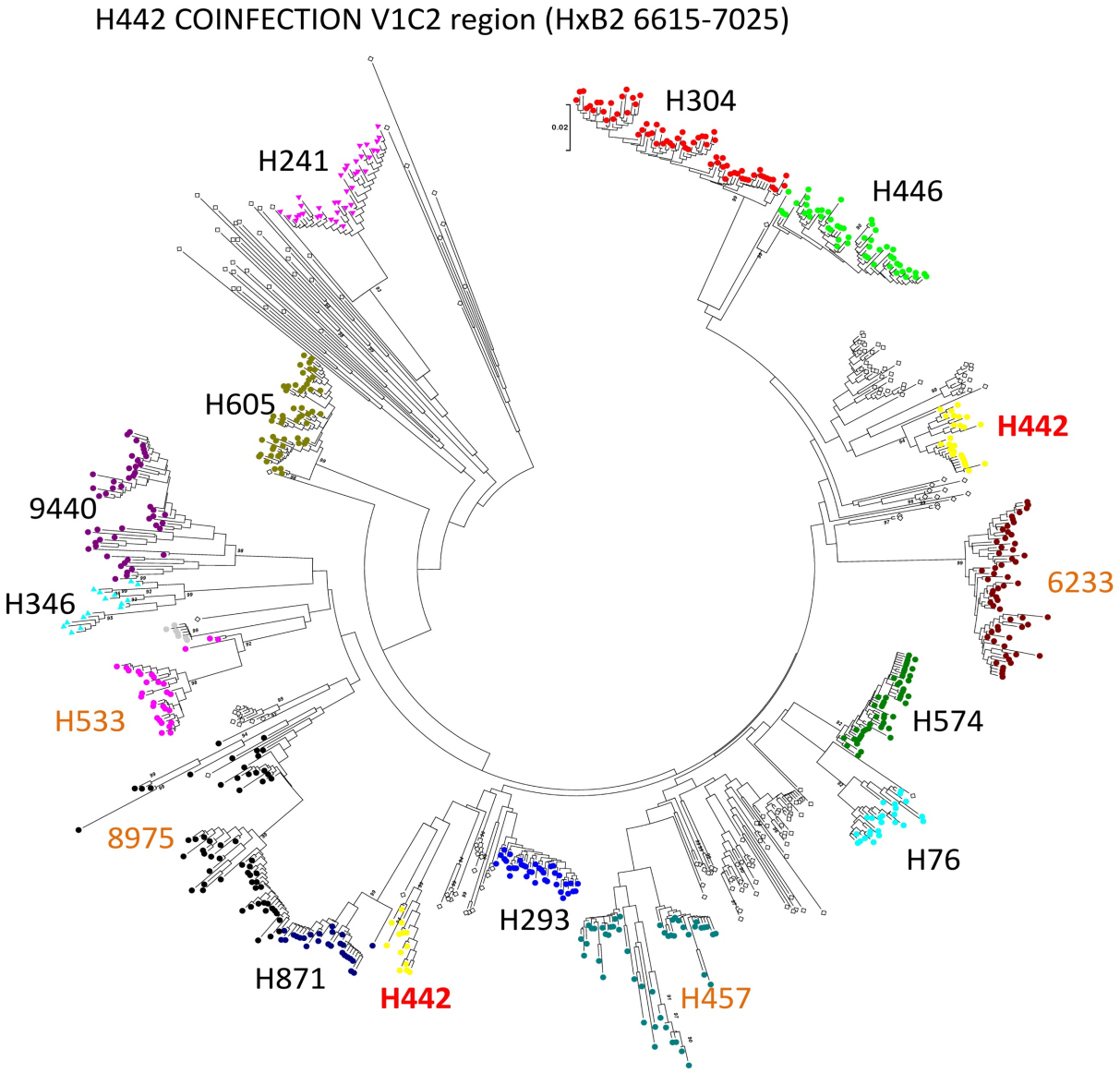


Fig. 7 Phylogenetic analysis of the V1C2 region of HxB2: 6615→7025) for dual infection. For subject H442 remarkably 2 significant different infecting strains of nucleotide sequence clusters indicating coinfection could be detected

Within the V3 region (HxB2 7026→7245), evidence for dual infection was detected for 9440 and H457 with large distances of two phylogenetic populations for each subject. In the same region evidence for coinfection was again detected for H442 with two phylogenetic populations significantly distant to each other (Fig. 8).

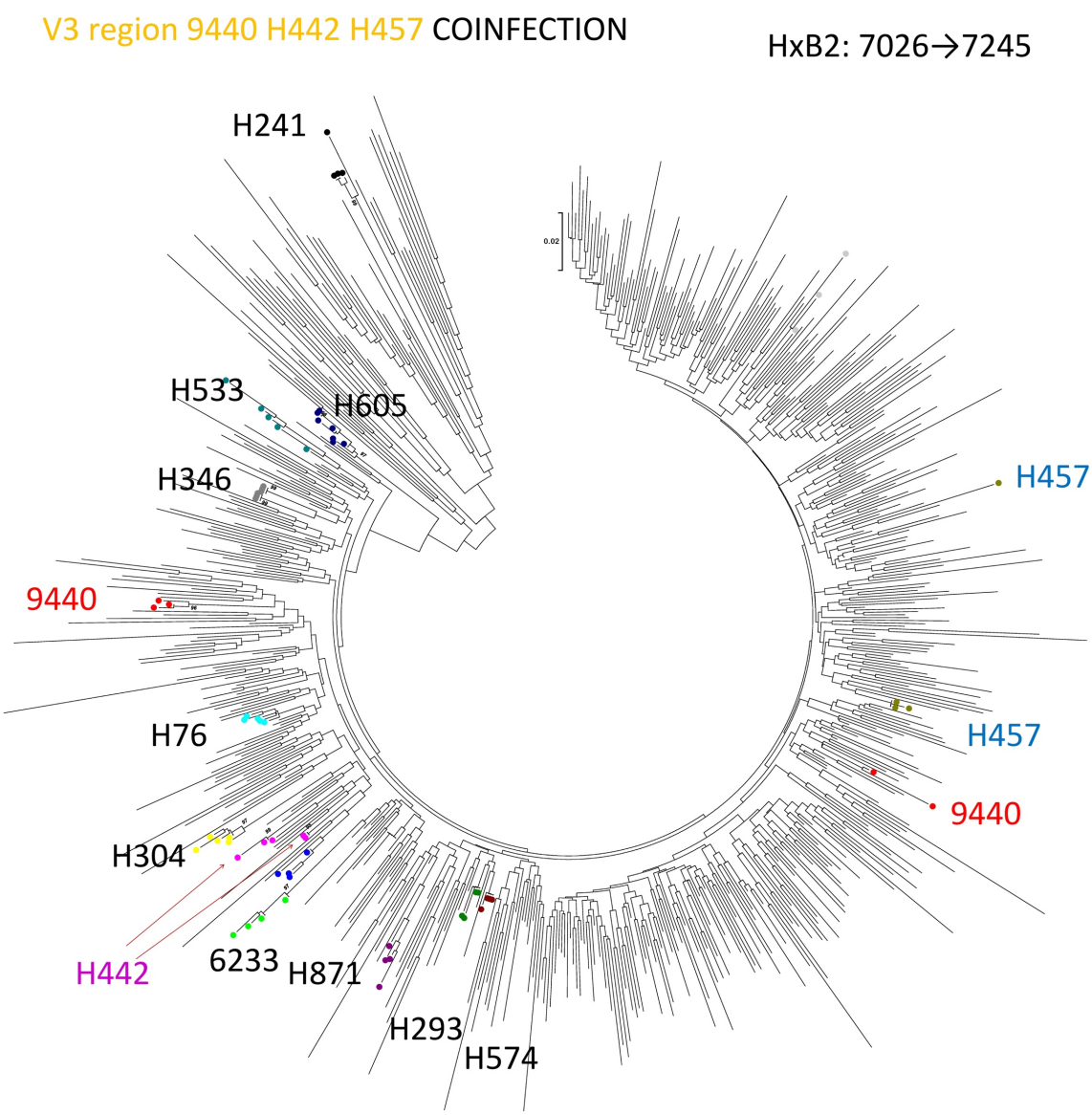


Fig. 8 Dual infection evidence for subject 9440, H457 and H442 within the V3 region of HxB2: 7026→7245. For each subject 2 significant distant clusters were detected.

For the same region, within the genome sequence of 7008→7223, evidence for coinfection was detected for 8975, H346 and H533 (Fig. 9).

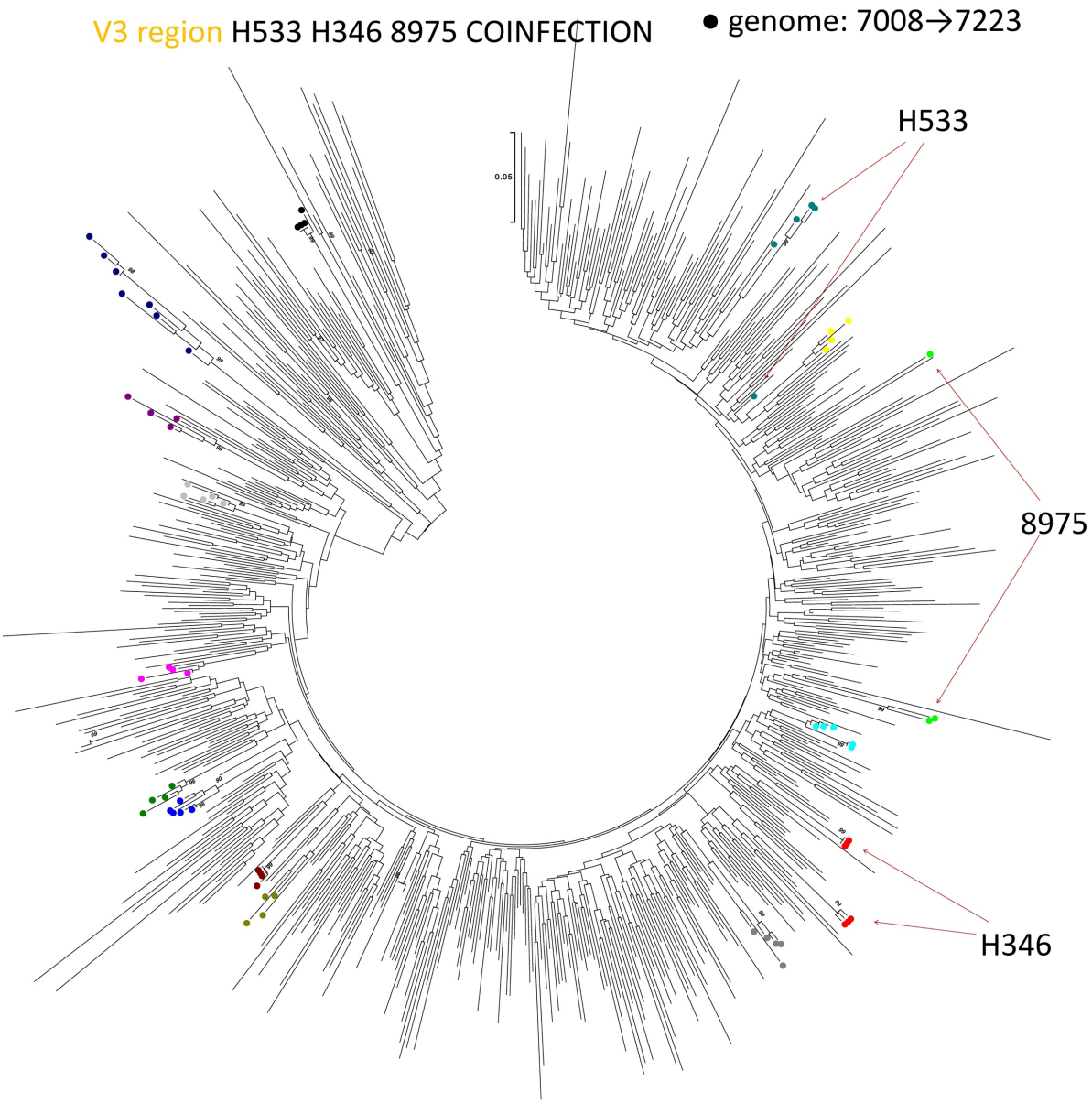


Fig. 9 Within the genome sequence of 7008→7223 coinfection for 8975, H346 and H533, latter previously known to be dual infected, was detected within the V3 with 2 significant distant clusters for each subject

H442 and H533 had former been determined as dual infected by Multi-region Hybridization assay (MHA). Dual infection status for 9440 and 8975 remained unknown (Appendix: Table 13) until this study and for H457 and H346 single infection status had been determined by MHA before this phylogenetic analysis. These results demonstrate that all these subjects were infected with more than one HIV strain.

4. Discussion

The extend of HIV disease progression is highly dependent on active HIV replication in $CD4^+$ cells and therefore, predicted by the HIV plasma viral load (Mellors et al., 1996; Ioannidis et al., 1996). While HIV infection rates vary for different CD4 T cell populations (Brenchley et al., 2004a; Brenchley et al., 2004b; Chomont et al., 2009; Ganesan et al., 2010), memory $CD4^+$ T cells seem to be the dominant substrate for virus replication throughout most of the infection duration (Veazey et al., 1998; Zhang et al., 1999; Haase, 1999; Schacker et al., 2001). It still remains unknown to what degree particular immune cell populations contribute to plasma virus production.

To assess the possible contribution of distinct memory T cells to plasma virion production, blood samples of 9 subjects with ART-naive HIV-1 infection were included in this study in order to phylogenetically analyze and compare the highly variable EnvV1V3 region sequences of plasma virus with viral sequences derived from sorted memory ($CD45RA^-$) T cell subpopulations. These included $CD4^{low/-}$, Treg cells ($CD4^+CD25^+/FoxP3^+$), Tfh cells ($CD4^+CXCR5^+$) and non-Tfh cells ($CD4^+CXCR5^-$). The infection duration at the time of blood sampling varied from 21-24 months (subject H442), 24-27 months (subject H346), 30-33 months (subject H304) and >36 months (subjects H76, H457, H533, H293, H446, H241). The EnvV1V3 region could be amplified in the $CD4^{low/-}$ memory T cell subsets for 3 subjects, in the $CD4^+CD25^+/FoxP3^+$ cells for 7 subjects, in $CD4^+CXCR5^+$ and $CD4^+CXCR5^-$ in 6 subjects. Subsequent cloning of the PCR products enabled genetic sequencing and later comparison of EnvV1V3 between cell-derived and plasma-derived populations.

The EnvV1V3 amplicon sequence variation, introduced by nested PCR amplification was assessed with an endpoint-diluted positive control plasmid (Du422 clone 1), as previously described by Chachage et al. (2016). 80% of cloned amplicon EnvV1V3 sequences differed only by ≤2 nucleotide substitutions, without insertions or deletions. However, up to 4 substitutions were considered as quasi-identical between cell- and plasma-derived sequence variants, regarding the performed nested PCR with 2 rounds.

Despite the relatively low number of sorted cells used for the nested EnvV1V3 PCR reaction, most amplicons were detected within the sorted $CD25^+/FoxP3^+$ CD4 T cells (7 out of the 9 subjects), followed by EnvV1V3 amplification in 6 out of the 9 subjects for $CXCR5^+$ and $CXCR5^-$ CD4 T cells and only 3 subjects with EnvV1V3 amplification for the $CD4^{low/-}$population (blue colored values, Table 10). These results are consistent with the results of Chachage et al. (2016) showing that HIV

DNA loads are comparatively high in the $CD25^+/FoxP3^+$ memory CD4 T cell population. However, within the Treg cell populations in the present study, out of the 7 EnvV1V3 amplicons only for 1 subject a quasi-identical cluster of 17 Treg derived sequences and 2 plasma derived sequences were detected. Consequently, within this subject, Tregs may have been recently infected or were actively contributing to plasma virus at the time of blood sampling. These data argue against a major contribution of circulating Treg cells to plasma virion production in ART naive HIV positive subjects, despite their comparatively high infection rates (Chachage et al., 2016).

As mentioned above in 6 out of 9 subjects EnvV1V3 amplicons were detected in the $CXCR5^+$ T cell population, which is consistent with a high HIV infection rate of this cell subset, considering the relatively low numbers of sorted cells. However, similar to the Treg population, only 1 quasi-identical sequence cluster was found with 1 Tfh-derived EnvV1V3 sequence clustering with 1 plasma-derived sequence (H304; Q3, Fig. 3; Table 11). Lymph node Tfh cells are considered to be a major substrate for HIV viral replication as evident by the correlation of HIV DNA load in these cells and the plasma viral load (Perreau et al., 2013). Potential explanations for the disparities between the present study and the results from Perreau et al. (2013) include primarily the investigation of sorted circulating, but not LN derived Tfh cells in the present study. Further different phenotypic markers were used to define the Tfh subset. This study used solely CXCR5 to define Tfh cells, whereas Perreau et al. (2013) used T cells that co-expressed CXCR5 and PD1. Hence it is not completely clear whether $CXCR5^+$ cells truly represent Tfh cells, as $CXCR5^+$ memory T cells include both PD1 positive and PD1 negative cell populations.

In peripheral blood $CXCR5^-$ memory CD4 T cells represent the major memory CD4 T cell population. In 6 out of 9 subjects the EnvV1V3 region could be amplified. In 3 of these 6 subjects quasi-identical sequence clusters were detected within this non-Tfh, non-Treg memory cell population: H304 (1 non-Tfh cell derived nucleotide sequence clustering with the previously mentioned quasi-identical pair of 1 Tfh and 1 plasma derived nucleotide sequence; Q3, Fig. 3; Table 11), H442 (1 non-Tfh cell derived nucleotide sequence clustering with 10 plasma-derived sequences; Q4, Fig. 3; Table 11) and H446 (1 non-Tfh cell derived nucleotide sequence clustering with 1 plasma-derived sequence; Q6, Fig. 3; Table 11 and another quasi-identical pair of 2 non-Tfh cell derived nucleotide sequences clustering with another plasma-derived sequence; Q7, Fig. 3; Table 11). Noticeably, for H304 the quasi-identical cluster of 1 Tfh ($CXCR5^+$) and 1 non-Tfh ($CXCR5^-$) population with 1 plasma-derived population can be interpreted that either both cell populations had recently become infected by the same (or highly similar) virus or that CXCR5 expression was up or downregulated in the offspring of a cell that had become HIV infected. Unfortunately, little is known about whether differentiated cells of a

specific lineage, such as Tregs or Tfh, can convert back into a less differentiated cell type.

Only in 3 out of 9 subjects the EnvV1V3 region could be amplified in the $CD4^{low/-}$ memory T cell population. Interestingly, 2 of these 3 cell-derived sequences showed quasi-identical clusters with plasma virus sequences. For H76 1 cluster contained 3 cell-derived sequences clustering with 4 plasma derived sequences (Q1, Fig. 3, Table 11) and for H442 1 cell-derived sequence clustered with 1 plasma-derived sequence (Q5, Fig. 3, Table 11). This relative accumulation of quasi-identical clusters in the $CD4^{low/-}$ memory T cell subset could be the result from productive HIV infection and thus, HIV-1 induced downregulation of the CD4 surface expression as described earlier (Kaiser et al., 2007). Hence, our results are consistent with those reported by Kaiser et al. (2007), which showed that $CD4^{low/-}$ memory T cells contained the highest proportion of unspliced (us) HIV RNA per 10^6 cells (Kaiser et. 2007) and therefore, play a crucial role in HIV plasma virion production.

Quasi-identical sequence-clusters could be detected in 5 out of the 9 subjects, whereas generally within the EnvV1V3 amplicons, a large sequence variability was detected in all investigated subjects. These findings suggest that rather than a single T cell subset, more likely various memory T cell subsets support the viral replication cycle and are involved in ongoing plasma virus production. Furthermore, individual EnvV1V3 sequences derived from plasma and cellular compartments were often very highly variable and intermingled in the absence of compartmentalization. Hence, each of these individual sequences must have originated from a different cellular source, producing plasma virus (or became recently HIV infected) at any given time point. Blood accounts for an average of 8% of the body weight. As we only analyzed a few milliliters of the average 5 to 8 liters of total human blood volume, these results highlight that a huge number of different cellular sources produce plasma virus simultaneously in viremic subjects.

In consideration of the different times since infection, Chachage et al. (2016) detected a correlation of infection duration with increasing EnvV1V3 sequence variation and decreasing rates of quasi-identical pairs between plasma and cell-associated sequences. The present study could not confirm these findings since no correlation of infection duration with decreasing frequencies of quasi-identical clusters between plasma and cell-derived EnvV1V3 sequences and thus increasing sequence variation was detected. A possible explanation could be that HIV infection had occurred more than once in some of our study subjects. Occurrence of unique recombinant forms between 2 or more different HIV strains could potentially mask the linear relationship between infection duration and increased viral diversity in an HIV positive subject off ART (McCutchan et al., 2005). In order to further investigate this assumption, all available EnvV1V3 sequences from 15 subjects, including those from

the Chachage et al. study (2016) were subjected to a detailed analysis of EnvV1V3 subregions. Indeed, various individuals must have been co-infected with different HIV strains, as evident by highly divergent variant sequences in the same EnvV1V3 subregion within the same individual. Evidence of such dual infection was detected, mainly in the V3 region at position HxB2 7026→7245 (H442, H457, 9440) and at position 7008→7223 (H533, H346, 8975). Additionally, for H442 evidence for dual infection was also confirmed in the V1C2 region at position HxB2 6615→7025. These results underline that dual HIV infections were quite common in female bar workers from the HISIS study, conducted from 2000-2005, but also in further subjects from a more recent study conducted in the Mbeya region, Tanzania (Worm-HIV infection, conducted from 2010-2013 (Arroyo et al., 2005; Chachage et al., 2014).

Only 2 out of 7 co-infected individuals had quasi-identical cluster pairs between plasma and cell-derived sequences. As dual infection may undergo "unique" possible interchange, specifically genetic recombination between viral strains with non-linearly increasing viral diversity, the detection of circulating viral species as quasi-identical clusters between plasma- and cell-derived sequences and therefore, EnvV1V3 sequences of distinct cell populations as predominant origins of plasma viremia may presumably be more unlikely to be detected.

Considering the fact that EnvV1V3 nucleotide variation did not correlate directly with infection duration or dual infection status in this study, additional variables may also be important and likely include viral load, extend of natural immune control and intrinsic properties of infected cells, such as APOBEC activity, a family of cellular polynucleotide cytidine deaminases that are potent inhibitors of HIV-1 infection (Malim et al., 2009). Hence, the frequency of quasi-identical cell-plasma sequence pairs in an individual may be influenced by multiple factors.

Earlier studies have used different approaches to identify potential cellular sources of the plasma virus. In the study of Kaiser et al. (2007) approximately 10000 copies of usRNA HIV nucleic acid copies/10^6 PBMCs were detected in activated $CD4^+$ T cells, which are considered a major substrate for HIV replication (Maier et al., 2000). The highest levels of almost 15000 usRNA HIV nucleic acid copies/10^6 PBMCs were detected in $CD4^-/CD8^-$ T cells, probably related to the HIV-1 induced downregulation of the CD4 surface expression, as an effect of productive HIV infection (Kaiser et al., 2007). The results of this study are consistent with the latter finding, as for the 3 subjects, in which EnvV1V3 was successfully amplified, quasi-identical clusters in 2 subjects could be detected between $CD4^{low/-}CD8^-$ memory T cell-derived and plasma-derived sequences.
A study of a nonhuman primate (NHP) SIV infection model used in situ HIV RNA hybridization to

identify transcriptionally active cellular SIV reservoirs in different organs, but excluding the blood compartment (Estes et al., 2017). This study showed that in the absence of ART more than 98% of detectable $vRNA^+$ and $vDNA^+$ cells can be found in LT, particularly in LN and GALT (Estes et al., 2017). LT contains high numbers of Tfh and other memory CD4 T cells that should be susceptible to productive HIV infection (Perreau et al., 2013, Banga et al., 2016). The infected cell pool in blood contains archival, quiescent proviruses, within highly differentiated cells, which appeared clonally expanded and occasionally expressed virions upon stimulation (Boritz et al., 2016). Homeostatic cellular proliferation may indeed be an important mechanism contributing to persistence of HIV in vivo by passing on HIV proviral DNA without the production of virions (Chomont et al., 2009). High homeostatic proliferation of HIV infected Tregs could potentially contribute to the high frequency of detected EnvV1V3 amplicons without a high number of quasi-identical clusters with concurrent plasma-derived sequences.

In this study, about $3{,}5x10^5$ of memory T cells were sorted from a median of about $10x10^6$ PBMCs. From this cell pool, a median of about $8\ x10^4$ cells and a median $1x10^4$ were sorted for the $CD4^{low/-}$ memory T cell subset (median 23% of memory $CD3^+$ cells) and $CD25^+/FoxP3^+$ cells (median of 2,9% $CD4^+$ memory T cells), respectively. From the remaining memory $CD4^+$ T cells, a median of $6x10^4$ and $2x10^5$ were sorted for $CXCR5^+$ (median 17%) and $CXCR5^-$ (median 57%) were sorted respectively.

After FACS, during cell lysis, a minimum of 10mM TRIS-HCl buffer of 30 µl was layered above the sorted cell pellet. From this cell-buffer dilution only 10 µl was used in the first-round nested PCR. Hence, cellular DNA corresponding to only 1 third of the sorted cells was used for amplification in the EnvV1V3 nested PCR. A median of 10 µl of plasma and a median of 15 µl from the second-round PCR product was then used for DNA gel extraction after gel electrophoresis and only 2 µl of 50 µl eluted EnvV1V3 amplicons were finally used for molecular cloning. Hence, the number of sorted cells and plasma-associated HIV sequences, that were included in our phylogenetic analyses, only represent a very tiny fraction of the HIV sequences that exist in blood in vivo.

Future studies may increase the validity of phylogenetic relationships between plasma- and certain cell subset-derived HIV sequences by studying much larger cell numbers and significantly larger rates of HIV sequences per cell subset and plasma. Furthermore, amplification of the target sequence within each cell population may increase sensitivity of EnvV1V3 PCR due to larger DNA quantities. Although investigations of the human LT may be challenging due to invasive procedures and thus ethical issues, analysis of these compartments with significantly higher levels of $vRNA^+$ and $vDNA^+$

cells will be important to understand the cellular source of plasma virus production.

In conclusion, the hypothesis of preferential plasma virion production from a specific memory T cell population could not be supported in this study. Moreover, the question arises whether once the reservoir of HIV-1 DNA infected cells is established, which parameters influence their lifespan and what influences their long-term reservoir contribution. Therefore, still little is known about the cellular source of active plasma virus production.

Innovative phylogenetic correlations and insights into the compartmentalized intra-host population dynamics will provide new perspectives in understanding the persistence of HIV-1 in its host environment. These consolidated findings will facilitate meliorating of cellular pharmacokinetics and spatial distribution of ART to entirely inhibit viral replication and sustain immune system function as a supposition to eliminate viral reservoirs and enable long-term cure for HIV-1 infection.

5. References

Akiba, H., Takeda, K., Kojima, Y., Usui, Y., Harada, N., Yamazaki, T., Ma, J., Tezuka, K., Yagita, H. and Okumura, K. (2005) 'The Role of ICOS in the CXCR5+ Follicular B Helper T Cell Maintenance In Vivo', *The Journal of Immunology*, 175(4), pp. 2340–2348. doi: 10.4049/jimmunol.175.4.2340.

Andersson, J., Boasso, A., Nilsson, J., Zhang, R., Shire, N. J., Lindback, S., Shearer, G. M. and Chougnet, C. a (2005) 'The prevalence of regulatory T cells in lymphoid tissue is correlated with viral load in HIV-infected patients.', *Journal of immunology (Baltimore, Md. : 1950)*, 174(6), pp. 3143–3147. doi: 174/6/3143 [pii].

Apoil, P. A., Puissant, B., Roubinet, F., Abbal, M., Massip, P. and Blancher, A. (2005) 'FOXP3 mRNA levels are decreased in peripheral blood CD4+ lymphocytes from HIV-positive patients.', *Journal of acquired immune deficiency syndromes (1999)*, 39(4), pp. 381–385. doi: 10.1097/01.qai.0000169662.30783.2d.

Arroyo, M. A., Hoelscher, M., Sateren, W., Samky, E., Maboko, L., Hoffmann, O., Kijak, G., Robb, M., Birx, D. L. and McCutchan, F. E. (2005) 'HIV-1 diversity and prevalence differ between urban and rural areas in the Mbeya region of Tanzania', *AIDS*. doi: 10.1097/01.aids.0000183515.14642.76.

Banga, R., Procopio, F. A., Noto, A., Pollakis, G., Cavassini, M., Ohmiti, K., Corpataux, J. M., De Leval, L., Pantaleo, G. and Perreau, M. (2016) 'PD-1+and follicular helper T cells are responsible for persistent HIV-1 transcription in treated aviremic individuals', *Nature Medicine*. doi: 10.1038/nm.4113.

Bettelli, E., Dastrange, M. and Oukka, M. (2005) 'Foxp3 interacts with nuclear factor of activated T cells and NF-kappa B to repress cytokine gene expression and effector functions of T helper cells', *Proceedings of the National Academy of Sciences of the United States of America*, 102(14), pp. 5138–5143.

Blankson, J. N., Persaud, D., Siliciano, R. F., In, P. R. and Ecycling, P. A. R. (2002) 'The Challenge of Viral Reservoirs in HIV-1 Infection', *Annual review of medicine*, 53(November), pp. 557–593. doi: 10.1093/ajae/aau087.

Blankson, J. N., Bailey, J. R., Thayil, S., Yang, H.-C. C., Lassen, K., Lai, J., Gandhi, S. K., Siliciano, J. D., Williams, T. M. and Siliciano, R. F. (2007) 'Isolation and characterization of replication-competent human immunodeficiency virus type 1 from a subset of elite suppressors', *Journal of virology*, 81(5), pp. 2508–18. doi: 10.1128/JVI.02165-06.

Boritz, E. A., Darko, S., Swaszek, L., Wolf, G., Wells, D., Wu, X., Henry, A. R., Laboune, F., Hu, J., Ambrozak, D., Hughes, M. S., Hoh, R., Casazza, J. P., Vostal, A., Bunis, D., Nganou-Makamdop, K., Lee, J. S., Migueles, S. A., Koup, R. A., Connors, M., Moir, S., Schacker, T., Maldarelli, F., Hughes, S. H., Deeks, S. G. and Douek, D. C. (2016) 'Multiple Origins of Virus Persistence during Natural Control of HIV Infection', *Cell*, 166(4), pp. 1004–1015. doi: 10.1016/j.cell.2016.06.039.

Borvak, J., Chou, C. S., Bell, K., Van Dyke, G., Zola, H., Ramilo, O. and Vitetta, E. S. (1995) 'Expression of CD25 defines peripheral blood mononuclear cells with productive versus latent HIV infection.', *Journal of immunology (Baltimore, Md. : 1950)*, 155(6), pp. 3196–3204.

Boswell, K. L., Paris, R., Boritz, E., Ambrozak, D., Yamamoto, T., Darko, S., Wloka, K., Wheatley, A., Narpala, S., McDermott, A., Roederer, M., Haubrich, R., Connors, M., Ake, J., Douek, D. C., Kim, J., Petrovas, C. and Koup, R. A. (2014) 'Loss of Circulating CD4 T Cells with B Cell Helper Function during Chronic HIV Infection', *PLoS Pathogens*, 10(1). doi: 10.1371/journal.ppat.1003853.

Boyman, O. and Sprent, J. (2012) 'The role of interleukin-2 during homeostasis and activation of the immune system', *Nature Reviews Immunology*, pp. 180–190. doi: 10.1038/nri3156.

Brenchley, J. M., Hill, B. J., Ambrozak, D. R., Price, D. A., Guenaga, F. J., Casazza, J. P., Kuruppu, J., Yazdani, J., Migueles, S. A., Connors, M., Roederer, M., Douek, D. C. and Koup, R. A. (2004a) 'T-Cell Subsets That Harbor Human Immunodeficiency Virus (HIV) In Vivo: Implications for HIV Pathogenesis', *Journal of Virology*, 78(3), pp. 1160–1168. doi: 10.1128/JVI.78.3.1160-1168.2004.

Brenchley, J. M., Schacker, T. W., Ruff, L. E., Price, D. A., Taylor, J. H., Beilman, G. J., Nguyen, P. L., Khoruts, A., Larson, M., Haase, A. T. and Douek, D. C. (2004b) 'CD4 + T Cell Depletion during all Stages of HIV Disease Occurs Predominantly in the Gastrointestinal Tract', *The Journal of Experimental Medicine*, 200(6), pp. 749–759. doi: 10.1084/jem.20040874.

Buzon, M. J., Sun, H., Li, C., Shaw, A., Seiss, K., Ouyang, Z., Martin-Gayo, E., Leng, J., Henrich, T. J., Li, J. Z., Pereyra, F., Zurakowski, R., Walker, B. D., Rosenberg, E. S., Yu, X. G. and Lichterfeld, M. (2014) 'HIV-1 persistence in CD4+ T cells with stem cell-like properties.', *Nature Medicine*, 20(2), pp. 139–42. doi: 10.1038/nm.3445.

Cagigi, A., Mowafi, F., Dang, L. V. P., Tenner-Racz, K., Atlas, A., Grutzmeier, S., Racz, P., Chiodi, F. and Nilsson, A. (2008) 'Altered expression of the receptor-ligand pair CXCR5/CXCL13 in B cells during chronic HIV-1 infection', *Blood*, 112(12), pp. 4401–4410. doi: 10.1182/blood-2008-02-140426.

Casazza, J. P., Brenchley, J. M., Hill, B. J., Ayana, R., Ambrozak, D., Roederer, M., Douek, D. C., Betts, M. R. and Koup, R. a (2009) 'Autocrine production of beta-chemokines protects CMV-Specific CD4 T cells from HIV infection.', *PLoS pathogens*, 5(10), p. e1000646. doi: 10.1371/journal.ppat.1000646.

Cerf-Bensussan, N. and Guy-Grand, D. (1991) 'Intestinal intraepithelial lymphocytes', *Gastroenterol Clin North Am*, 20(3), pp. 549–576.

Chachage, M., Podola, L., Clowes, P., Nsojo, A., Bauer, A., Mgaya, O., Kowour, D., Froeschl, G., Maboko, L., Hoelscher, M., Saathoff, E. and Geldmacher, C. (2014) 'Helminth-Associated Systemic Immune Activation and HIV Co-receptor Expression: Response to Albendazole/Praziquantel Treatment', *PLoS Neglected Tropical Diseases*. doi: 10.1371/journal.pntd.0002755.

Chachage, M., Pollakis, G., Kuffour, E. O., Haase, K., Bauer, A., Nadai, Y., Podola, L., Clowes, P., Schiemann, M., Henkel, L., Hoffmann, D., Joseph, S., Bhuju, S., Maboko, L., Sarfo, F. S., Eberhardt, K., Hoelscher, M., Feldt, T., Saathoff, E. and Geldmacher, C. (2016) 'CD25+FoxP3+ memory CD4 T cells are frequent targets of HIV infection in vivo', *Journal of Virology*, (July), p. JVI.00612-16. doi: 10.1128/JVI.00612-16.

Chevalier, N., Jarrossay, D., Ho, E., Avery, D. T., Ma, C. S., Yu, D., Sallusto, F., Tangye, S. G. and Mackay, C. R. (2011) 'CXCR5 expressing human central memory CD4 T cells and their relevance for humoral immune responses.', *Journal of immunology (Baltimore, Md. : 1950)*, 186(10), pp. 5556–68. doi: 10.4049/jimmunol.1002828.

Chomont, N., El-Far, M., Ancuta, P., Trautmann, L., Procopio, F. A., Yassine-Diab, B., Boucher, G., Boulassel, M.-R., Ghattas, G., Brenchley, J. M., Schacker, T. W., Hill, B. J., Douek, D. C., Routy, J.-P., Haddad, E. K. and Sékaly, R.-P. (2009) 'HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation.', *Nature medicine*, 15(8), pp. 893–900. doi: 10.1038/nm.1972.

Chong, Y., Nabeshima, S., Furusyo, N., Murata, M., Yamaji, K. and Hayashi, J. (2004) 'Downregulation of CXCR5 in CD27- B cells of HIV-1 infected patients', *Journal of Medical Virology*, 73(3), pp. 362–367. doi: 10.1002/jmv.20099.

Chou, C. S., Ramilo, O. and Vitetta, E. S. (1997) 'Highly purified CD25- resting T cells cannot be infected de novo with HIV-1.', *Proceedings of the National Academy of Sciences of the United States of America*, 94(4), pp. 1361–5. doi: 10.1073/pnas.94.4.1361.

Chun, T. W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J. a, Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T. C., Kuo, Y. H., Brookmeyer, R., Zeiger, M. a, Barditch-Crovo, P. and Siliciano, R. F. (1997a) 'Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection.', *Nature*, 387(6629), pp. 183–188. doi: 10.1038/387183a0.

Chun, T. W., Stuyver, L., Mizell, S. B., Ehler, L. a, Mican, J. a, Baseler, M., Lloyd, a L., Nowak, M. a and Fauci, a S. (1997b) 'Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy.', *Proceedings of the National Academy of Sciences of the United States of America*, 94(24), pp. 13193–7. doi: 10.1073/pnas.94.24.13193.

Chun, T. W., Nickle, D. C., Justement, J. S., Large, D., Semerjian, A., Curlin, M. E., O'Shea, M. A., Hallahan, C. W., Daucher, M., Ward, D. J., Moir, S., Mullins, J. I., Kovacs, C. and Fauci, A. S. (2005) 'HIV-infected individuals receiving effective antiviral therapy for extended periods of time continually replenish their viral reservoir', *Journal of Clinical Investigation*, 115(11), pp. 3250–3255. doi: 10.1172/JCI26197.

Combadiere, B., Boissonnas, A., Carcelain, G., Lefranc, E., Samri, A., Bricaire, F., Debre, P. and Autran, B. (2004) 'Distinct Time Effects of Vaccination on Long-Term Proliferative and IFN-γ–producing T Cell Memory to Smallpox in Humans', *The Journal of Experimental Medicine*, 199(11), pp. 1585–1593. doi: 10.1084/jem.20032083.

Crotty, S. (2011) 'Follicular helper CD4 T cells (TFH).', *Annual review of immunology*, 29(1), pp. 621–663. doi: 10.1146/annurev-immunol-031210-101400.

Crotty, S. (2012) 'The 1-1-1 fallacy', *Immunological Reviews*, 247(1), pp. 133–142. doi: 10.1111/j.1600-065X.2012.01117.x.

Crotty, S. (2014) 'T Follicular Helper Cell Differentiation, Function, and Roles in Disease', *Immunity*, pp. 529–542. doi: 10.1016/j.immuni.2014.10.004.

Crotty, S. (2015) 'A brief history of T cell help to B cells', *Nature Reviews Immunology*, 15(3), pp. 185–189. doi: 10.1038/nri3803.

Dahabieh, M. S., Battivelli, E. and Verdin, E. (2015) 'Understanding HIV Latency: The Road to an HIV Cure', *Annual Review of Medicine*, 66(1), pp. 407–421. doi: 10.1146/annurev-med-092112-152941.

Douek, D. C., Brenchley, J. M., Betts, M. R., Ambrozak, D. R., Hill, B. J., Okamoto, Y., Casazza, J. P., Kuruppu, J., Kunstman, K., Wolinsky, S., Grossman, Z., Dybul, M., Oxenius, A., Price, D. A., Connors, M. and Koup, R. A. (2002) 'HIV preferentially infects HIV-specific CD4+T cells', *Nature*, 417(6884), pp. 95–98. doi: 10.1038/417095a.

Eggena, M. P., Barugahare, B., Jones, N., Okello, M., Mutalya, S., Kityo, C., Mugyenyi, P. and Cao, H. (2005) 'Depletion of regulatory T cells in HIV infection is associated with immune activation.', *Journal of immunology (Baltimore, Md. : 1950)*, 174(7), pp. 4407–4414. doi: 10.4049/jimmunol.174.7.4407.

Epple, H. J., Loddenkemper, C., Kunkel, D., Tröger, H., Maul, J., Moos, V., Berg, E., Ullrich, R., Schulzke, J. D., Stein, H., Duchmann, R., Zeitz, M. and Schneider, T. (2006) 'Mucosal but not peripheral FOXP3+ regulatory T cells are highly increased in untreated HIV infection and normalize after suppressive HAART', *Blood*, 108(9), pp. 3072–3078. doi: 10.1182/blood-2006-04-016923.

Estes, J. D., Li, Q., Reynolds, M. R., Wietgrefe, S., Duan, L., Schacker, T., Picker, L. J., Watkins, D. I., Lifson, J. D., Reilly, C., Carlis, J. and Haase, A. T. (2006) 'Premature Induction of an Immunosuppressive Regulatory T Cell Response during Acute Simian Immunodeficiency Virus Infection', *The Journal of Infectious Diseases*, 193(5), pp. 703–712. doi: 10.1086/500368.

Estes, J., Baker, J. V., Brenchley, J. M., Khoruts, A., Barthold, J. L., Bantle, A., Reilly, C. S., Beilman, G. J., George, M. E., Douek, D. C., Haase, A. T. and Schacker, T. W. (2008) 'Collagen Deposition Limits Immune Reconstitution in the Gut', *The Journal of Infectious Diseases*, 198(4), pp. 456–464. doi: 10.1086/590112.

Estes, J. D., Haase, A. T. and Schacker, T. W. (2008) 'The role of collagen deposition in depleting CD4+ T cells and limiting reconstitution in HIV-1 and SIV infections through damage to the secondary lymphoid organ niche', *Seminars in Immunology*, pp. 181–186. doi: 10.1016/j.smim.2008.04.002.

Estes, J. D., Kityo, C., Ssali, F., Swainson, L., Makamdop, K. N., Del Prete, G. Q., Deeks, S. G., Luciw, P. A., Chipman, J. G., Beilman, G. J., Hoskuldsson, T., Khoruts, A., Anderson, J., Deleage, C., Jasurda, J., Schmidt, T. E., Hafertepe, M., Callisto, S. P., Pearson, H., Reimann, T., Schuster, J., Schoephoerster, J., Southern, P., Perkey, K., Shang, L., Wietgrefe, S. W., Fletcher, C. V., Lifson, J. D., Douek, D. C., McCune, J. M., Haase, A. T. and Schacker, T. W. (2017) 'Defining total-body AIDS-virus burden with implications for curative strategies', *Nature Medicine*, 23(11), pp. 1271–1276. doi: 10.1038/nm.4411.

Farber, D. L., Yudanin, N. A. and Restifo, N. P. (2014) 'Human memory T cells: Generation, compartmentalization and homeostasis', *Nature Reviews Immunology*, pp. 24–35. doi: 10.1038/nri3567.

Fazilleau, N., Mark, L., McHeyzer-Williams, L. J. and McHeyzer-Williams, M. G. (2009) 'Follicular Helper T Cells: Lineage and Location', *Immunity*, pp. 324–335. doi: 10.1016/j.immuni.2009.03.003.

Fife, B. T. and Pauken, K. E. (2011) 'The role of the PD-1 pathway in autoimmunity and peripheral tolerance', *Annals of the New York Academy of Sciences*, 1217(1), pp. 45–59. doi: 10.1111/j.1749-6632.2010.05919.x.

Finberg, R. W., Wahl, S. M., Allen, J. B., Soman, G., Strom, T. B., Murphy and Nichols, J. C. (1991) 'Selective elimination of HIV-1-infected cells with an interleukin-2 receptor-specific cytotoxin', *Science*, 252(5013), pp. 1703–1705.

Finzi, D. (1997) 'Identification of a Reservoir for HIV-1 in Patients on Highly Active Antiretroviral Therapy', *Science*, 278(5341), pp. 1295–1300. doi: 10.1126/science.278.5341.1295.

Finzi, D., Blankson, J., Siliciano, J. D., Margolick, J. B., Chadwick, K., Pierson, T., Smith, K., Lisziewicz, J., Lori, F., Flexner, C., Quinn, T. C., Chaisson, R. E., Rosenberg, E., Walker, B., Gange, S., Gallant, J. and Siliciano, R. F. (1999) 'Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy.', *Nature medicine*, 5(5), pp. 512–517. doi: 10.1038/8394.

Fletcher, C. V., Staskus, K., Wietgrefe, S. W., Rothenberger, M., Reilly, C., Chipman, J. G., Beilman, G. J., Khoruts, A., Thorkelson, A., Schmidt, T. E., Anderson, J., Perkey, K., Stevenson, M., Perelson, A. S., Douek, D. C., Haase, A. T. and Schacker, T. W. (2014) 'Persistent HIV-1 replication is associated with lower antiretroviral drug concentrations in lymphatic tissues', *Proceedings of the National Academy of Sciences*, 111(6), pp. 2307–2312. doi: 10.1073/pnas.1318249111.

Fontenot, J. D., Gavin, M. A. and Rudensky, A. Y. (2003) 'Foxp3 programs the development and function of CD4+CD25+ regulatory T cells', *Nat.Immunol.*, 4(1529–2908 (Print)), pp. 330–336. doi: 10.1038/ni904.

Francisco, L. M., Sage, P. T. and Sharpe, A. H. (2010) 'The PD-1 pathway in tolerance and autoimmunity', *Immunological Reviews*, pp. 219–242. doi: 10.1111/j.1600-065X.2010.00923.x.

Fukazawa, Y., Lum, R., Okoye, A. A., Park, H., Matsuda, K., Bae, J. Y., Hagen, S. I., Shoemaker, R., Deleage, C., Lucero, C., Morcock, D., Swanson, T., Legasse, A. W., Axthelm, M. K., Hesselgesser, J., Geleziunas, R., Hirsch, V. M., Edlefsen, P. T., Piatak, M., Estes, J. D., Lifson, J. D. and Picker, L. J. (2015) 'B cell follicle sanctuary permits persistent productive simian immunodeficiency virus infection in elite controllers', *Nature Medicine*, 21(2), pp. 132–139. doi: 10.1038/nm.3781.

Ganesan, A., Chattopadhyay, P. K., Brodie, T. M., Qin, J., Gu, W., Mascola, J. R., Michael, N. L., Follmann, D. A. and Roederer, M. (2010) 'Immunologic and Virologic Events in Early HIV Infection Predict Subsequent Rate of Progression', *The Journal of Infectious Diseases*, 201(2), pp. 272–284. doi: 10.1086/649430.

Gattinoni, L., Lugli, E., Ji, Y., Pos, Z., Paulos, C. M., Quigley, M. F., Almeida, J. R., Gostick, E., Yu, Z., Carpenito, C., Wang, E., Douek, D. C., Price, D., June, C. H., Marincola, F. M., Roederer, M. and Restifo, N. P. (2011) 'A human memory T cell subset with stem cell-like properties', *Nature medi*, 17(10), pp. 1290–1297. doi: 10.1038/nm.2446.

Geldmacher, C., Currier, J. R., Herrmann, E., Haule, A., Kuta, E., McCutchan, F., Njovu, L., Geis, S., Hoffmann, O., Maboko, L., Williamson, C., Birx, D., Meyerhans, A., Cox, J. and Hoelscher, M. (2007a) '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', *Journal of Virology*, 81(5), pp. 2440–2448. doi: 10.1128/JVI.01847-06.

Geldmacher, C., Currier, J. R., Gerhardt, M., Haule, A., Maboko, L., Birx, D., Gray, C., Meyerhans, A., Cox, J. and Hoelscher, M. (2007b) 'In a mixed subtype epidemic, the HIV-1 Gag-specific T-cell response is biased towards the infecting subtype', *AIDS*, 21(2), pp. 135–143. doi: 10.1097/01.aids.0000247589.77061.f7.

Geldmacher, C., Ngwenyama, N., Schuetz, A., Petrovas, C., Reither, K., Heeregrave, E. J., Casazza, J. P., Ambrozak, D. R., Louder, M., Ampofo, W., Pollakis, G., Hill, B., Sanga, E., Saathoff, E., Maboko, L., Roederer, M., Paxton, W. a, Hoelscher, M. and Koup, R. a (2010) 'Preferential infection and depletion of Mycobacterium tuberculosis-specific CD4 T cells after HIV-1 infection.', *The Journal of experimental medicine*, 207(13), pp. 2869–2881. doi: 10.1084/jem.20100090.

Geldmacher, C. and Koup, R. A. (2012) 'Pathogen-specific T cell depletion and reactivation of opportunistic pathogens in HIV infection', *Trends in Immunology*, 33(5), pp. 207–214. doi: 10.1016/j.it.2012.01.011.

Goletti, D., Weissman, D., Jackson, R. W., Graham, N. M., Vlahov, D., Klein, R. S., Munsiff, S. S., Ortona, L., Cauda, R. and Fauci, A. S. (1996) 'Effect of Mycobacterium tuberculosis on HIV replication. Role of immune activation.', *Journal of immunology (Baltimore, Md. : 1950)*, 157(3), pp. 1271–8. doi: 10.5402/2011/810565.

Guadalupe, M., Reay, E., Sankaran, S., Prindiville, T., Flamm, J., McNeil, A. and Dandekar, S. (2003) 'Severe CD4+ T-cell depletion in gut lymphoid tissue during primary human immunodeficiency virus type 1 infection and substantial delay in restoration following highly active antiretroviral therapy.', *Journal of virology*, 77(21), pp. 11708–17. doi: 10.1128/JVI.77.21.11708.

Guenthard, H. F., Wong, J. K., Spina, C. a, Ignacio, C., Kwok, S., Christopherson, C., Hwang, J., Haubrich, R., Havlir, D. and Richman, D. D. (2000) 'Effect of influenza vaccination on viral replication and immune response in persons infected with human immunodeficiency virus receiving potent antiretroviral therapy.', *The Journal of infectious diseases*, 181(2), pp. 522–31. doi: 10.1086/315260.

Haase, a T. (1999) 'Population biology of HIV-1 infection: viral and CD4+ T cell demographics and dynamics in lymphatic tissues.', *Annual review of immunology*, 17, pp. 625–56. doi: 10.1146/annurev.immunol.17.1.625.

Hoelscher, M., Dowling, W. E., Sanders-Buell, E., Carr, J. K., Harris, M. E., Thomschke, A., Robb, M. L., Birx, D. L. and McCutchan, F. E. (2002) 'Detection of HIV-1 subtypes, recombinants, and dual infections in east Africa by a multi-region hybridization assay', *AIDS*, 16(15), pp. 2055–2064. doi: 10.1097/00002030-200210180-00011.

Holmes, D., Jiang, Q., Zhang, L. and Su, L. (2008) 'Foxp3 and Treg cells in HIV-1 infection and immuno-pathogenesis', *Immunologic Research*, pp. 248–266. doi: 10.1007/s12026-008-8037-x.

Hori, S., Nomura, T. and Sakaguchi, S. (2003) 'Control of regulatory T cell development by the transcription factor Foxp3.', *Science*, 299(5609), pp. 1057–1061. doi: 10.1126/science.1079490.

Hurst, J., Hoffmann, M., Pace, M., Williams, J. P., Thornhill, J., Hamlyn, E., Meyerowitz, J., Willberg, C., Koelsch, K. K., Robinson, N., Brown, H., Fisher, M., Kinloch, S., Cooper, D. A., Schechter, M., Tambussi, G., Fidler, S., Babiker, A., Weber, J., Kelleher, A. D., Phillips, R. E. and Frater, J. (2015) 'Immunological biomarkers predict HIV-1 viral rebound after treatment interruption', *Nature Communications*, 6. doi: 10.1038/ncomms9495.

Ioannidis, J. P., Cappelleri, J. C., Lau, J., Sacks, H. S. and Skolnik, P. R. (1996) 'Predictive value of viral load measurements in asymptomatic untreated HIV-1 infection: a mathematical model.', *AIDS (London, England)*, 10(3), pp. 255–62.

Julg, B., Pereyra, F., Buzón, M. J., Piechocka-Trocha, A., Clark, M. J., Baker, B. M., Lian, J., Miura, T., Martinez-Picado, J., Addo, M. M. and Walker, B. D. (2010) 'Infrequent Recovery of HIV from but Robust Exogenous Infection of Activated CD4 $^{+}$ T Cells in HIV Elite Controllers', *Clinical Infectious Diseases*, 51(2), pp. 233–238. doi: 10.1086/653677.

Kaiser, P., Joos, B., Niederost, B., Weber, R., Gunthard, H. F. and Fischer, M. (2007) 'Productive human immunodeficiency virus type 1 infection in peripheral blood predominantly takes place in CD4/CD8 double-negative T lymphocytes', *J Virol*, 81(18), pp. 9693–9706. doi: 10.1128/JVI.00492-07.

Khattri, R., Cox, T., Yasayko, S.-A. and Ramsdell, F. (2003) 'An essential role for Scurfin in CD4+CD25+ T regulatory cells.', *Nature immunology*, 4(4), pp. 337–342. doi: 10.1038/ni909.

Kinter, A. L., Hennessey, M., Bell, A., Kern, S., Lin, Y., Daucher, M., Planta, M., McGlaughlin, M., Jackson, R., Ziegler, S. F. and Fauci, A. S. (2004) 'CD25(+)CD4(+) Regulatory T Cells from the Peripheral Blood of Asymptomatic HIV-infected Individuals Regulate CD4(+) and CD8(+) HIV-specific T Cell Immune Responses In Vitro and Are Associated with Favorable Clinical Markers of Disease Status', *The Journal of Experimental Medicine*. The Rockefeller University Press, 200(3), pp. 331–343. doi: 10.1084/jem.20032069.

Krammer, P. H. (2000) 'CD95's deadly mission in the immune system', *Nature*, pp. 789–795. doi: 10.1038/35037728.

Lane, H. C., Masur, H., Edgar, L. C., Whalen, G., Rook, A. H. and Fauci, A. S. (1983) 'Abnormalities of B-Cell Activation and Immunoregulation in Patients with the Acquired Immunodeficiency Syndrome', *New England Journal of Medicine*, 309(8), pp. 453–458. doi: 10.1056/NEJM198308253090803.

Li, M., Salazar-gonzalez, J. F., Derdeyn, C. A., Morris, L., Williamson, C., Robinson, J. E., Decker, M., Li, Y., Salazar, M. G., Victoria, R., Mlisana, K., Karim, S. A., Greene, K. M., Bilska, M., Zhou, J., Allen, S., Chomba, E., Mulenga, J., Gao, F., Zhang, M., Korber, B. T. M., Hahn, B. H., Montefiori, D. C., Li, M., Salazar-gonzalez, J. F., Derdeyn, C. A., Morris, L., Williamson, C., Robinson, J. E., Decker, J. M., Li, Y., Salazar, M. G., Polonis, V. R., Mlisana, K., Karim, S. A., Hong, K., Greene, K. M., Bilska, M., Zhou, J., Allen, S., Chomba, E., Mulenga, J. and Vwalika, C. (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. Virology*, 80(23), pp. 11776–11790. doi: 10.1128/JVI.01730-06.

Lim, H. W., Hillsamer, P., Banham, A. H. and Kim, C. H. (2005) 'Cutting edge: direct suppression of B cells by CD4+ CD25+ regulatory T cells.', *Journal of immunology (Baltimore, Md. : 1950)*, 175(7), pp. 4180–4183. doi: 10.4049/jimmunol.175.7.4180.

Lorenzo-Redondo, R., Fryer, H. R., Bedford, T., Kim, E. Y., Archer, J., Kosakovsky Pond, S. L., Chung, Y. S., Penugonda, S., Chipman, J. G., Fletcher, C. V., Schacker, T. W., Malim, M. H., Rambaut, A., Haase, A. T., McLean, A. R. and Wolinsky, S. M. (2016) 'Persistent HIV-1 replication maintains the tissue reservoir during therapy', *Nature*, 530(7588), pp. 51–56. doi: 10.1038/nature16933.

Maier, R., Bartolomé-Rodríguez, M. M., Moulon, C., Weltzien, H. U. and Meyerhans, a (2000) 'Kinetics of CXCR4 and CCR5 up-regulation and human immunodeficiency virus expansion after antigenic stimulation of primary CD4(+) T lymphocytes.', *Blood*.

Malim, M. H. (2009) 'APOBEC proteins and intrinsic resistance to HIV-1 infection', *Philosophical Transactions of the Royal Society B: Biological Sciences*. The Royal Society, 364(1517), pp. 675–687. doi: 10.1098/rstb.2008.0185.

Marson, A., Kretschmer, K., Frampton, G. M., Jacobsen, E. S., Polansky, J. K., MacIsaac, K. D., Levine, S. S., Fraenkel, E., von Boehmer, H. and Young, R. a (2007) 'Foxp3 occupancy and regulation of key target genes during T-cell stimulation.', *Nature*, 445(7130), pp. 931–935. doi: 10.1038/nature05478.

Mattapallil, J. J., Douek, D. C., Hill, B., Nishimura, Y., Martin, M. and Roederer, M. (2005) 'Massive infection and loss of memory CD4+T cells in multiple tissues during acute SIV infection', *Nature*, 434(7037), pp. 1093–1097. doi: 10.1038/nature03501.

McCune, J. M. (2001) 'The dynamics of CD4+ T-cell depletion in HIV disease', *Nature*, pp. 974–979. doi: 10.1038/35073648.

McCutchan, F. E., 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. and Birx, D. L. (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', *Journal of Virology*. doi: 10.1128/JVI.79.18.11693-11704.2005.

McGowan, I., Elliott, J., Fuerst, M., Taing, P., Boscardin, J., Poles, M. and Anton, P. (2004) 'Increased HIV-1 mucosal replication is associated with generalized mucosal cytokine activation', *Journal of Acquired Immune Deficiency Syndromes*, 37(2), pp. 1228–1236. doi: 10.1097/01.qai.0000131846.12453.29.

Mehandru, S., Poles, M. A., Tenner-Racz, K., Horowitz, A., Hurley, A., Hogan, C., Boden, D., Racz, P. and Markowitz, M. (2004) 'Primary HIV-1 Infection Is Associated with Preferential Depletion of CD4 $^{+}$ T Lymphocytes from Effector Sites in the Gastrointestinal Tract', *The Journal of Experimental Medicine*, 200(6), pp. 761–770. doi: 10.1084/jem.20041196.

Mellors, J. W., Rinaldo, C. R., Gupta, P., White, R. M., Todd, J. A. and Kingsley, L. A. (1996) 'Prognosis in HIV-1 Infection Predicted by the Quantity of Virus in Plasma', *Science*, 272(5265), pp. 1167–1170. doi: 10.1126/science.272.5265.1167.

Migueles, S. A. and Connors, M. (2015) 'Success and failure of the cellular immune response against HIV-1', *Nat Immunol*, 16(6), pp. 563–570. doi: 10.1038/ni.3161.

Moir, S., Chun, T. W. and Fauci, A. S. (2011) 'Pathogenic mechanisms of HIV disease', *Annu Rev Pathol*, 6, pp. 223–248. doi: 10.1146/annurev-pathol-011110-130254.

Morita, R., Schmitt, N., Bentebibel, S.-E., Ranganathan, R., Bourdery, L., Zurawski, G., Foucat, E., Dullaers, M., Oh, S., Sabzghabaei, N., Lavecchio, E. M., Punaro, M., Pascual, V., Banchereau, J. and Ueno, H. (2011) 'Human Blood CXCR5(+)CD4(+) T Cells Are Counterparts of T Follicular Cells and Contain Specific Subsets that Differentially Support Antibody Secretion', *Immunity*, 34(1), pp. 108–121. doi: 10.1016/j.immuni.2010.12.012.

Mueller, G., Höpken, U. E. and Lipp, M. (2003) 'Erratum: The impact of CCR7 and CXCR5 on lymphoid organ development and systemic immunity (Immunological Reviews (2003) 195 (117-135))', *Immunological Reviews*, p. 265. doi: 10.1046/j.1600-065X.2003.00073.x.

Nei, M. and Kumar, S. (2000) *Molecular Evolutionand Phylogenetics*, *Archives of virology*. doi: 10.1046/j.1365-2540.2001.0923a.x.

Nicholas, K. J., Flaherty, D. K., Smith, R. M., Sather, D. N. and Kalams, S. A. (2016) 'Chronic HIV-1 infection impairs superantigen-induced activation of peripheral CD4+CXCR5+PD-1+ cells, with relative preservation of recall antigen-specific responses', *JAIDS Journal of Acquired Immune Deficiency Syndromes*, p. 1. doi: 10.1097/QAI.0000000000001152.

Nilsson, J., Boasso, A., Velilla, P. A., Zhang, R., Vaccari, M., Franchini, G., Shearer, G. M., Andersson, J. and Chougnet, C. (2006) 'HIV-1-driven regulatory T-cell accumulation in lymphoid tissues is associated with disease progression in HIV/AIDS', *Blood*, 108(12), pp. 3808–3817. doi: 10.1182/blood-2006-05-021576.

Nishimura, Y., Igarashi, T., Donau, O. K., Buckler-White, A., Buckler, C., Lafont, B. A. P., Goeken, R. M., Goldstein, S., Hirsch, V. M. and Martin, M. A. (2004) 'Highly pathogenic SHIVs and SIVs target different CD4+ T cell subsets in rhesus monkeys, explaining their divergent clinical courses', *Proceedings of the National Academy of Sciences*, 101(33), pp. 12324–12329. doi: 10.1073/pnas.0404620101.

Oswald-Richter, K., Grill, S. M., Shariat, N., Leelawong, M., Sundrud, M. S., Haas, D. W. and Unutmaz, D. (2004) 'HIV infection of naturally occurring and genetically reprogrammed human regulatory T-cells', *PLoS Biology*, 2(7). doi: 10.1371/journal.pbio.0020198.

Pallikkuth, S., Fischl, M. a and Pahwa, S. (2013) 'Combination Antiretroviral Therapy With Raltegravir Leads to Rapid Immunologic Reconstitution in Treatment-Naive Patients With Chronic HIV Infection.', *The Journal of infectious diseases*, 208(10), pp. 1613–23. doi: 10.1093/infdis/jit387.

Patel, P., Borkowf, C. B., Brooks, J. T., Lasry, A., Lansky, A. and Mermin, J. (2014) 'Estimating per-act HIV transmission risk: A systematic review', *AIDS*. doi: 10.1097/QAD.0000000000000298.

Perreau, M., Savoye, A.-L., De Crignis, E., Corpataux, J.-M., Cubas, R., Haddad, E. K., De Leval, L., Graziosi, C. and Pantaleo, G. (2013) 'Follicular helper T cells serve as the major CD4 T cell compartment for HIV-1 infection, replication, and production.', *The Journal of experimental medicine*, 210(1), pp. 143–56. doi: 10.1084/jem.20121932.

Pinto, L., Covas, M. J. and Victorino, R. M. (1991) 'Loss of CD45RA and gain of CD45RO after in vitro activation of lymphocytes from HIV-infected patients.', *Immunology*, 73(2), pp. 147–50.

Pollakis, G., Baan, E., van Werkhoven, M. B., Berkhout, B., Bakker, M., Jurriaans, S. and Paxton, W. A. (2015) 'Association between gp120 envelope V1V2 and V4V5 variable loop profiles in a defined HIV-1 transmission cluster.', *AIDS (London, England)*. England, 29(10), pp. 1161–1171. doi: 10.1097/QAD.0000000000000692.

Qi, H. (2016) 'T follicular helper cells in space-time', *Nature Reviews Immunology*, pp. 612–625. doi: 10.1038/nri.2016.94.

Ramilo, O., Bell, K. D., Uhr, J. W. and Vitetta, E. S. (1993) 'Role of CD25+ and CD25-T cells in acute HIV infection in vitro', *J Immunol*, 150(11), pp. 5202–5208.

Riou, C., Yassine-Diab, B., Van grevenynghe, J., Somogyi, R., Greller, L. D., Gagnon, D., Gimmig, S., Wilkinson, P., Shi, Y., Cameron, M. J., Campos-Gonzalez, R., Balderas, R. S., Kelvin, D., Sekaly, R.-P. and Haddad, E. K. (2007) 'Convergence of TCR and cytokine signaling leads to FOXO3a phosphorylation and drives the survival of CD4 $^{+}$ central memory T cells', *The Journal of Experimental Medicine*, 204(1), pp. 79–91. doi: 10.1084/jem.20061681.

Rothenberger, M. K., Keele, B. F., Wietgrefe, S. W., Fletcher, C. V., Beilman, G. J., Chipman, J. G., Khoruts, A., Estes, J. D., Anderson, J., Callisto, S. P., Schmidt, T. E., Thorkelson, A., Reilly, C., Perkey, K., Reimann, T. G., Utay, N. S., Nganou Makamdop, K., Stevenson, M., Douek, D. C., Haase, A. T. and Schacker, T. W. (2015) 'Large number of rebounding/founder HIV variants emerge from multifocal infection in lymphatic tissues after treatment interruption', *Proceedings of the National Academy of Sciences*, 112(10), pp. E1126–E1134. doi: 10.1073/pnas.1414926112.

Sachsenberg, N., Perelson, A. S., Yerly, S., Schockmel, G. A., Leduc, D., Hirschel, B. and Perrin, L. (1998) 'Turnover of CD4+ and CD8+ T lymphocytes in HIV-1 infection as measured by Ki-67 antigen.', *The Journal of experimental medicine*, 187(8), pp. 1295–303. doi: 10.1084/jem.187.8.1295.

Sakaguchi, S. (2004) 'Naturally arising CD4+ regulatory t cells for immunologic self-tolerance and negative control of immune responses', *Annu Rev Immunol*, 22, pp. 531–562. doi: 10.1146/annurev.immunol.21.120601.141122.

Schacker, T., Little, S., Connick, E., Gebhard, K., Zhang, Z. Q., Krieger, J., Pryor, J., Havlir, D., Wong, J. K., Schooley, R. T., Richman, D., Corey, L. and Haase, A. T. (2001) 'Productive infection of T cells in lymphoid tissues during primary and early human immunodeficiency virus infection.', *The Journal of infectious diseases*, 183(4), pp. 555–62. doi: 10.1086/318524.

Schacker, T. W., Nguyen, P. L., Beilman, G. J., Wolinsky, S., Larson, M., Reilly, C. and Haase, A. T. (2002) 'Collagen deposition in HIV-1 infected lymphatic tissues and T cell homeostasis Rapid Publication', *Journal of Clinical Investigation*, 110(8), pp. 1133–1139. doi: 10.1172/JCI200216413.Introduction.

Schacker, T. W., Reilly, C., Beilman, G. J., Taylor, J., Skarda, D., Krason, D., Larson, M. and Haase, A. T. (2005) 'Amount of lymphatic tissue fibrosis in HIV infection predicts magnitude of HAART-associated change in peripheral CD4 cell count', *AIDS*. England: The Rockefeller University Press, 19(18), pp. 2169–2171. doi: 10.1097/01.aids.0000194801.51422.03.

Schacker, T. W., Brenchley, J. M., Beilman, G. J., Reilly, C., Pambuccian, S. E., Taylor, J., Skarda, D., Larson, M., Douek, D. C. and Haase, A. T. (2006) 'Lymphatic Tissue Fibrosis Is Associated with Reduced Numbers of Naive CD4+ T Cells in Human Immunodeficiency Virus Type 1 Infection', *Clinical and Vaccine Immunology*, 13(5), pp. 556–560. doi: 10.1128/CVI.13.5.556-560.2006

Schultz, B. T., Teigler, J. E., Pissani, F., Oster, A. F., Kranias, G., Alter, G., Marovich, M., Eller, M. A., Dittmer, U., Robb, M. L., Kim, J. H., Michael, N. L., Bolton, D. and Streeck, H. (2017) 'Circulating HIV-Specific Interleukin-21+CD4+ T Cells Represent Peripheral Tfh Cells with Antigen-Dependent Helper Functions', *Immunity*. Elsevier, 44(1), pp. 167–178. doi: 10.1016/j.immuni.2015.12.011.

Smit-McBride, Z., Mattapallil, J. J., McChesney, M., Ferrick, D. and Dandekar, S. (1998) 'Gastrointestinal T lymphocytes retain high potential for cytokine responses but have severe CD4(+) T-cell depletion at all stages of simian immunodeficiency virus infection compared to peripheral lymphocytes.', *Journal of virology*, 72(8), pp. 6646–56.

Strain, M. C., Günthard, H. F., Havlir, D. V, Ignacio, C. C., Smith, D. M., Leigh-Brown, a J., Macaranas, T. R., Lam, R. Y., Daly, O. a, Fischer, M., Opravil, M., Levine, H., Bacheler, L., Spina, C. a, Richman, D. D. and Wong, J. K. (2003) 'Heterogeneous clearance rates of long-lived lymphocytes infected with HIV: intrinsic stability predicts lifelong persistence.', *Proceedings of the National Academy of Sciences of the United States of America*, 100(8), pp. 4819–24. doi: 10.1073/pnas.0736332100.

Tenner-Racz, K., Stellbrink, H.-J., van Lunzen, J., Schneider, C., Jacobs, J.-P., Raschdorff, B., Großschupff, G., Steinman, R. M. and Racz, P. (1998) 'The Unenlarged Lymph Nodes of HIV-1–infected, Asymptomatic Patients with High CD4 T Cell Counts Are Sites for Virus Replication and CD4 T Cell Proliferation. The Impact of Highly Active Antiretroviral Therapy', *The Journal of Experimental Medicine*. The Rockefeller University Press, 187(6), pp. 949–959.

Tsunemi, S., Iwasaki, T., Imado, T., Higasa, S., Kakishita, E., Shirasaka, T. and Sano, H. (2005) 'Relationship of CD4+CD25+ regulatory T cells to immune status in HIV-infected patients.', *AIDS (London, England)*, 19(9), pp. 879–886. doi: 00002030-200506100-00004 [pii].

Veazey, R. S., DeMaria, M., Chalifoux, L. V, Shvetz, D. E., Pauley, D. R., Knight, H. L., Rosenzweig, M., Johnson, R. P., Desrosiers, R. C. and Lackner, A. A. (1998) 'Gastrointestinal tract as a major site of CD4+ T cell depletion and viral replication in SIV infection', *Science (New York, N.Y.)*, 280(5362), pp. 427–431. doi: 10.1126/science.280.5362.427.

Vinuesa, C. G. (2012) 'HIV and T follicular helper cells: A dangerous relationship', *Journal of Clinical Investigation*, pp. 3059–3062. doi: 10.1172/JCI65175.

Walker, B. D. and Yu, X. G. (2013) 'Unravelling the mechanisms of durable control of HIV-1', *Nature Reviews Immunology*, pp. 487–498. doi: 10.1038/nri3478.

Webb, L. M. C. and Linterman, M. A. (2017) 'Signals that drive T follicular helper cell formation', *Immunology*. doi: 10.1111/imm.12778.Wei, X., Decker, J. M., Wang, S., Hui, H., Kappes, J. C., Wu, X., Salazar-Gonzalez, J. F., Salazar, M. G., Kilby, J. M., Saag, M. S., Komarova, N. L., Nowak, M. A., Hahn, B. H., Kwong, P. D. and Shaw, G. M. (2003) 'Antibody neutralization and escape by HIV-1', *Nature*, 422(6929), pp. 307–312. doi: 10.1038/nature01470.

Wong, J. K., Hezareh, M., Günthard, H. F., Havlir, D. V, Ignacio, C. C., Spina, C. a and Richman, D. D. (1997) 'Recovery of replication-competent HIV despite prolonged suppression of plasma viremia.', *Science (New York, N.Y.)*, 278(5341), pp. 1291–1295. doi: 10.1126/science.278.5341.1291.

Wu, Y., Borde, M., Heissmeyer, V., Feuerer, M., Lapan, A. D., Stroud, J. C., Bates, D. L., Guo, L., Han, A., Ziegler, S. F., Mathis, D., Benoist, C., Chen, L. and Rao, A. (2006) 'FOXP3 Controls Regulatory T Cell Function through Cooperation with NFAT', *Cell*, 126(2), pp. 375–387. doi: 10.1016/j.cell.2006.05.042.

Zhang, Z., Schuler, T., Zupancic, M., Wietgrefe, S., Staskus, K. a, Reimann, K. a, Reinhart, T. a, Rogan, M., Cavert, W., Miller, C. J., Veazey, R. S., Notermans, D., Little, S., Danner, S. a, Richman, D. D., Havlir, D., Wong, J., Jordan, H. L., Schacker, T. W., Racz, P., Tenner-Racz, K., Letvin, N. L., Wolinsky, S. and Haase, a T. (1999) 'Sexual transmission and propagation of SIV and HIV in resting and activated CD4+ T cells.', *Science (New York, N.Y.)*, 286(1999), pp. 1353–1357. doi: 10.1126/science.286.5443.1353.

Zheng, Y. and Rudensky, A. Y. (2007) 'Foxp3 in control of the regulatory T cell lineage', *Nature Immunology*. doi: 10.1038/ni1455.

Zhu, J., Yamane, H. and Paul, W. E. (2010) 'Differentiation of Effector CD4 T Cell Populations', *Annual Review of Immunology*, 28(1), pp. 445–489. doi: 10.1146/annurev-immunol-030409-101212.

6. Appendix

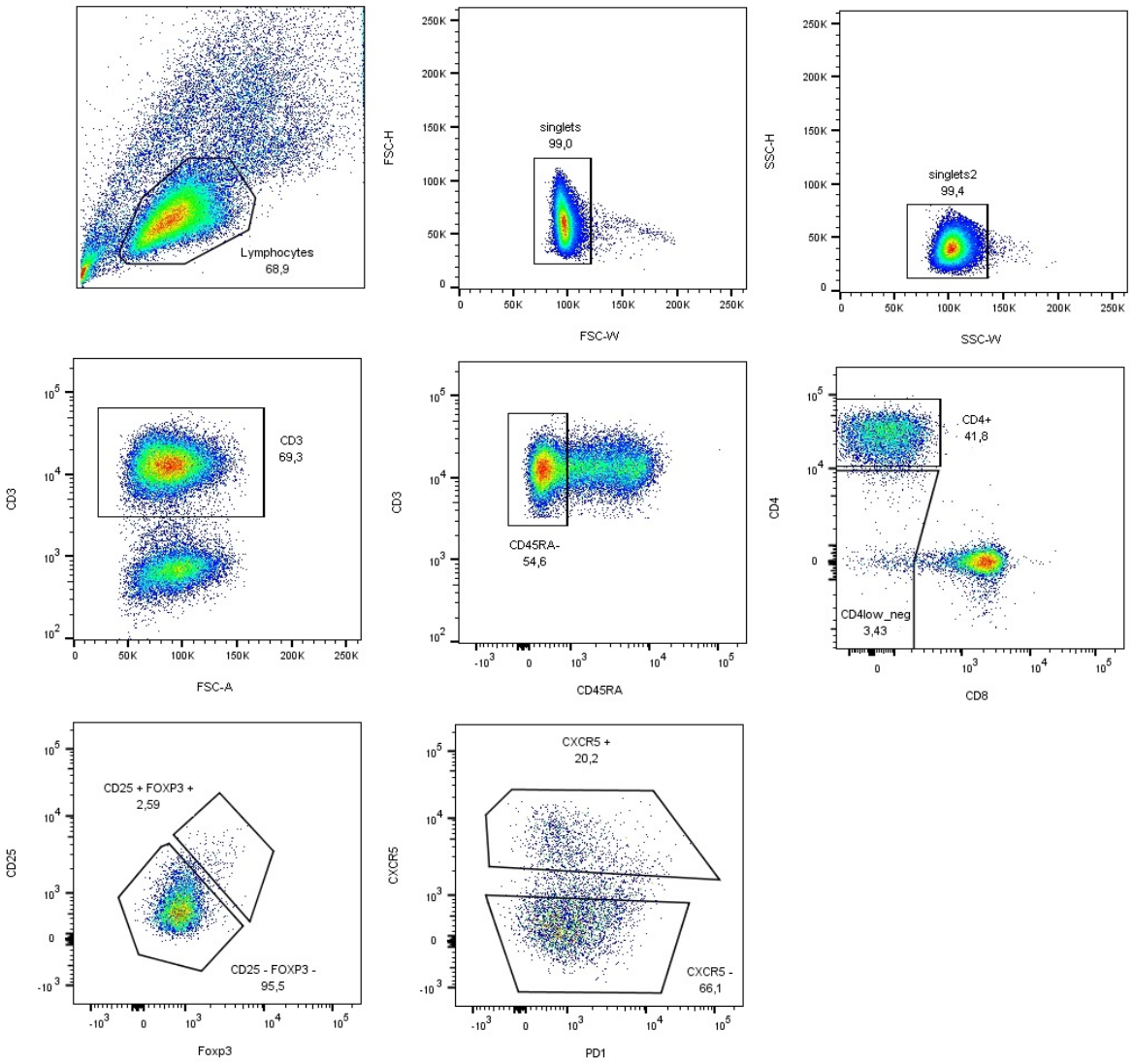
6.1 Infection status, subtype and dual infection evidence of the cohort of Chachage et al. (2016)

Subject	Infecion status by the time of blood sample collection	Subtype	Dual infection evidence
H574	9-12 months	AC	Dual
H605	27-30 months	AC	Dual
6233K12	16-38 months	C	unknown
9440A11[a]	>38 months	C	unknown
8710U11	>38 months	C	unknown
8975T11	>54 months	C	unknown

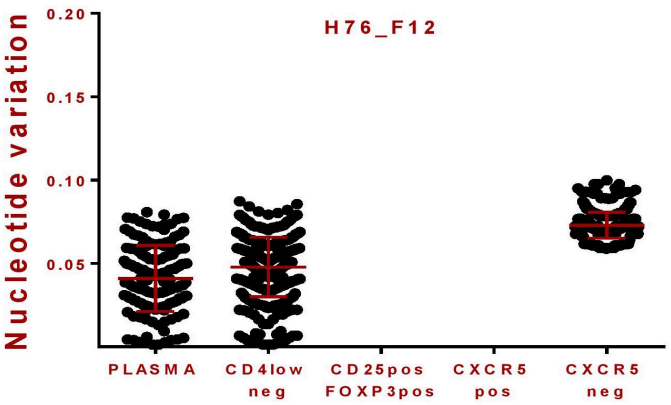
Table 13 Infection status, subtype and dual infection evidence of the cohort of the Chachage et al. (2016) study

6.2 Gating trees and HIV nucleotide variations between plasma and cell derived sequences

H76

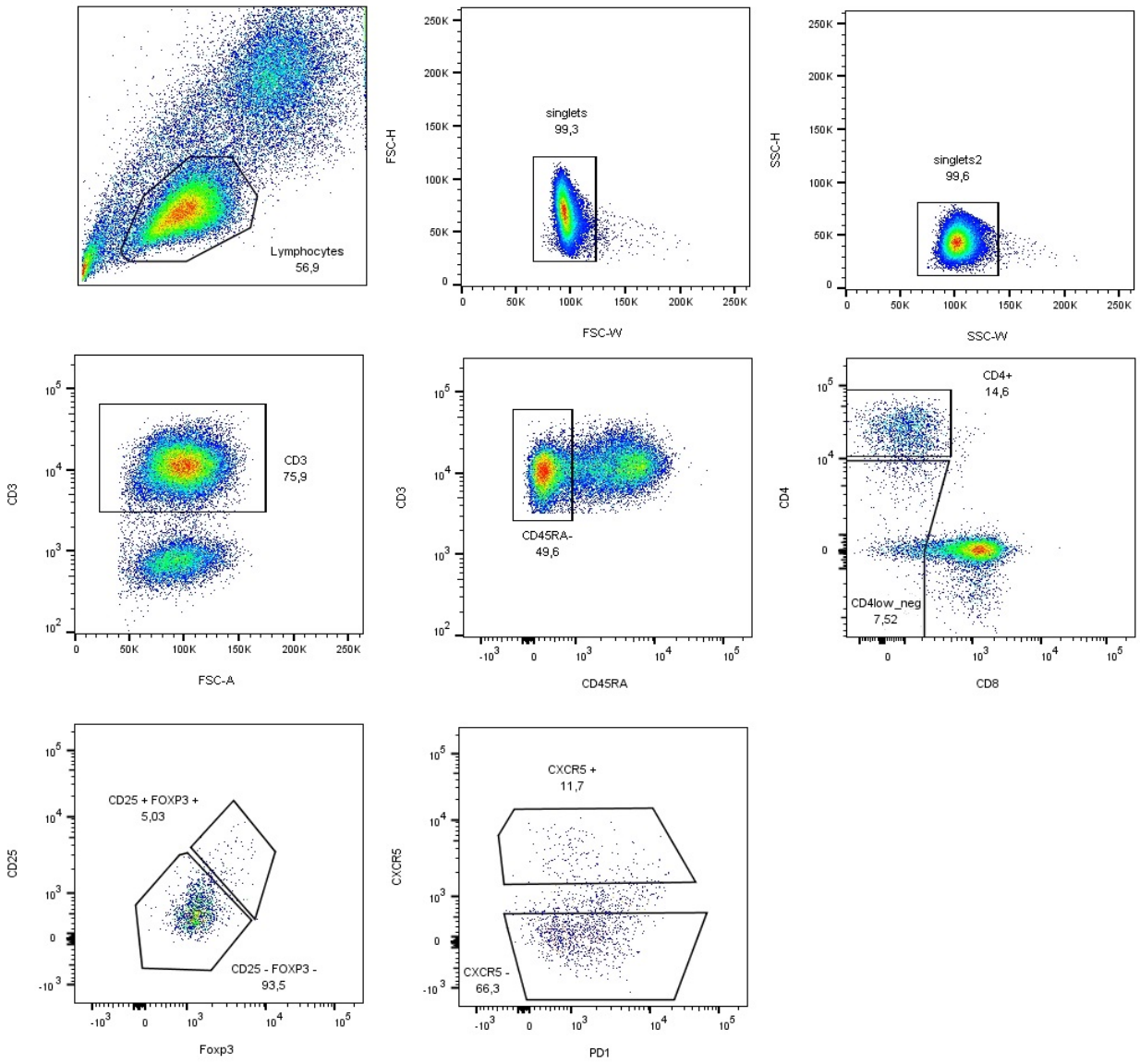


S1 Gating tree of subject H76

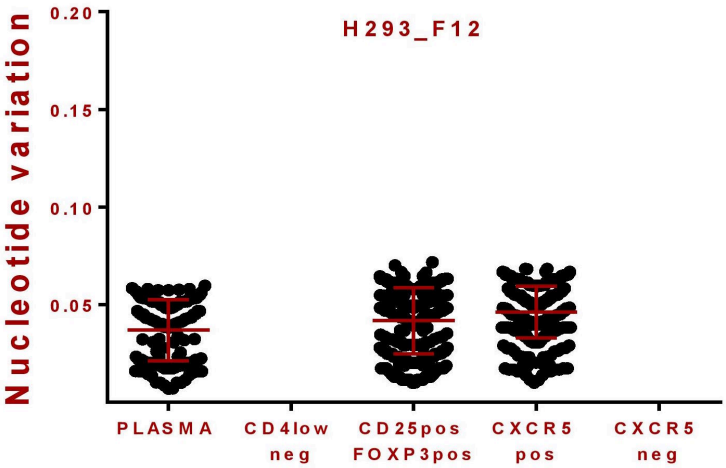


S2 Nucleotide variation between plasma and cell-derived viral nucleotide sequences for subject H76.

H293

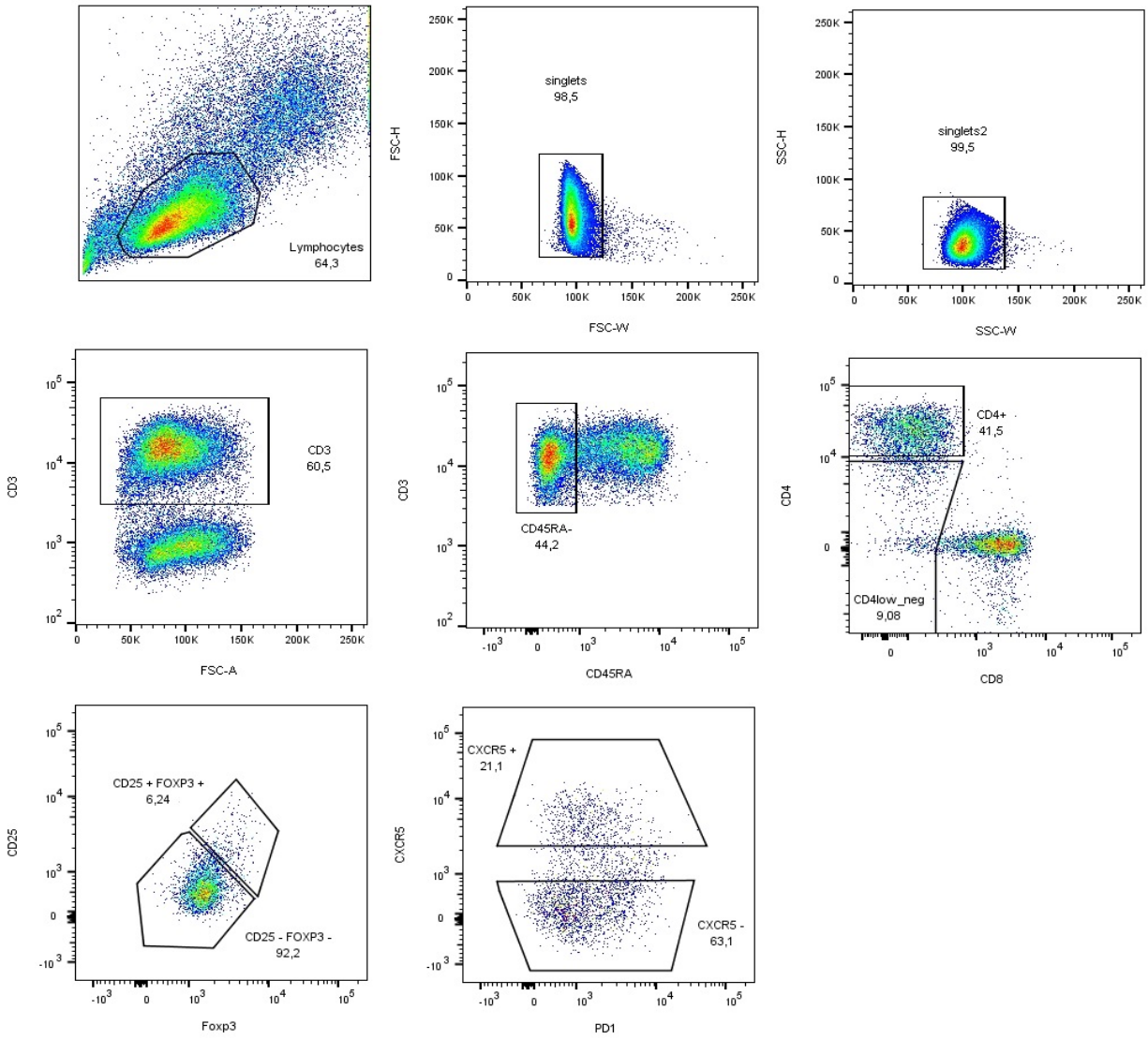


S3 Gating tree for subject H293

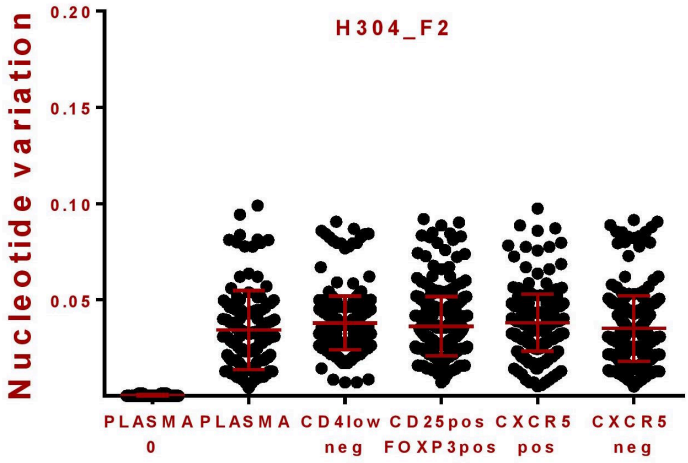


S4 Nucleotide variation between plasma and cell-derived viral nucleotide sequences for subject H293

H304

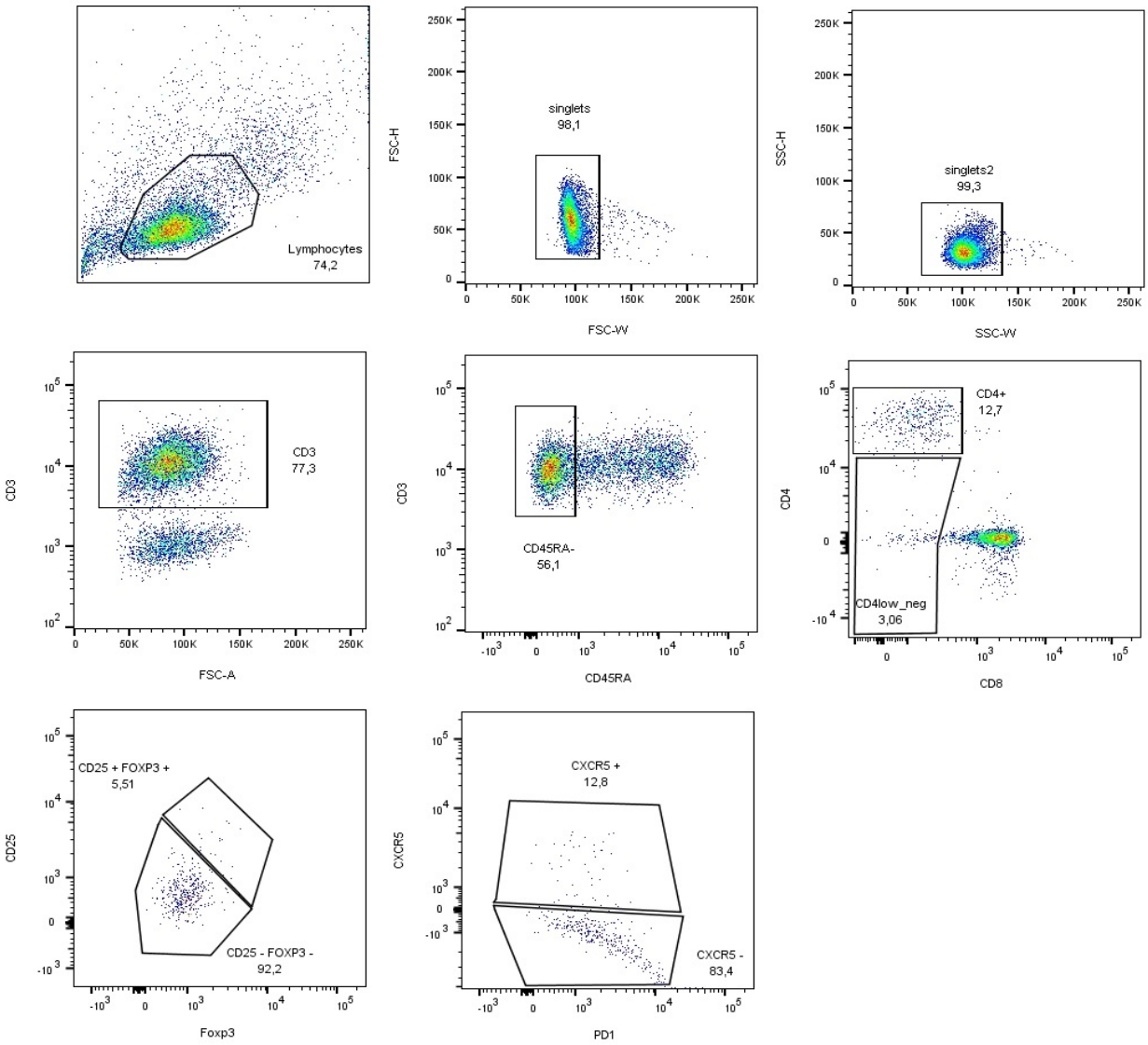


S5 Gating tree for subject H304

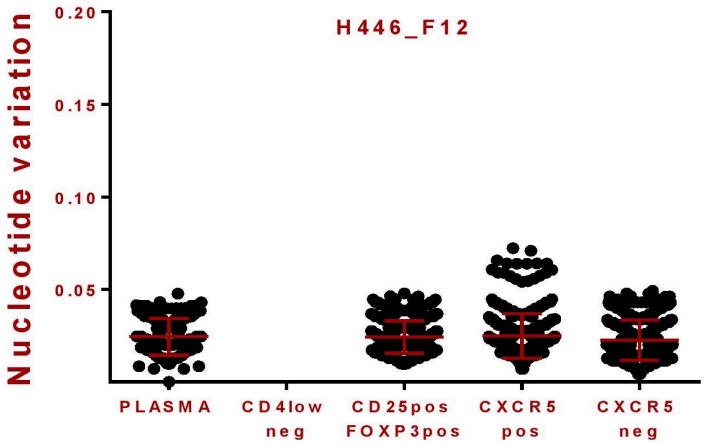


S6 Nucleotide variation between plasma and cell-derived viral nucleotide sequences for subject H304

H446

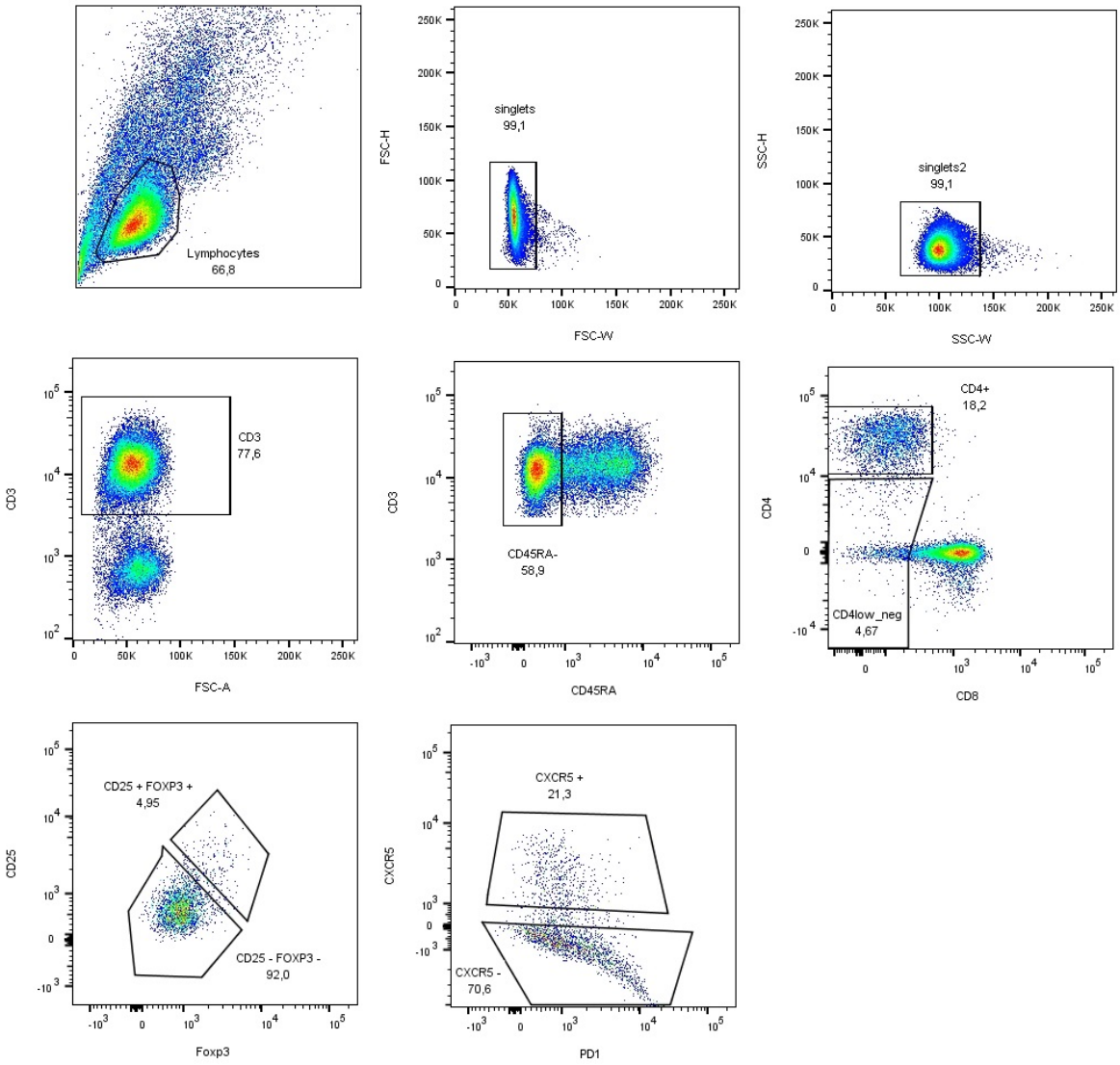


S7 Gating tree for subject H446

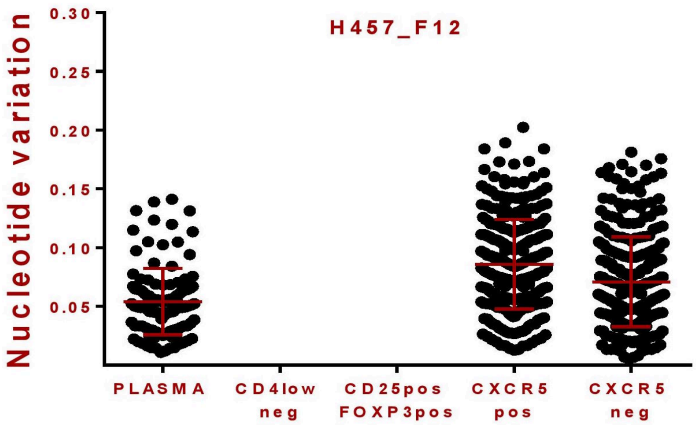


S8 Nucleotide variation between plasma and cell-derived viral nucleotide sequences for subject H446

H457

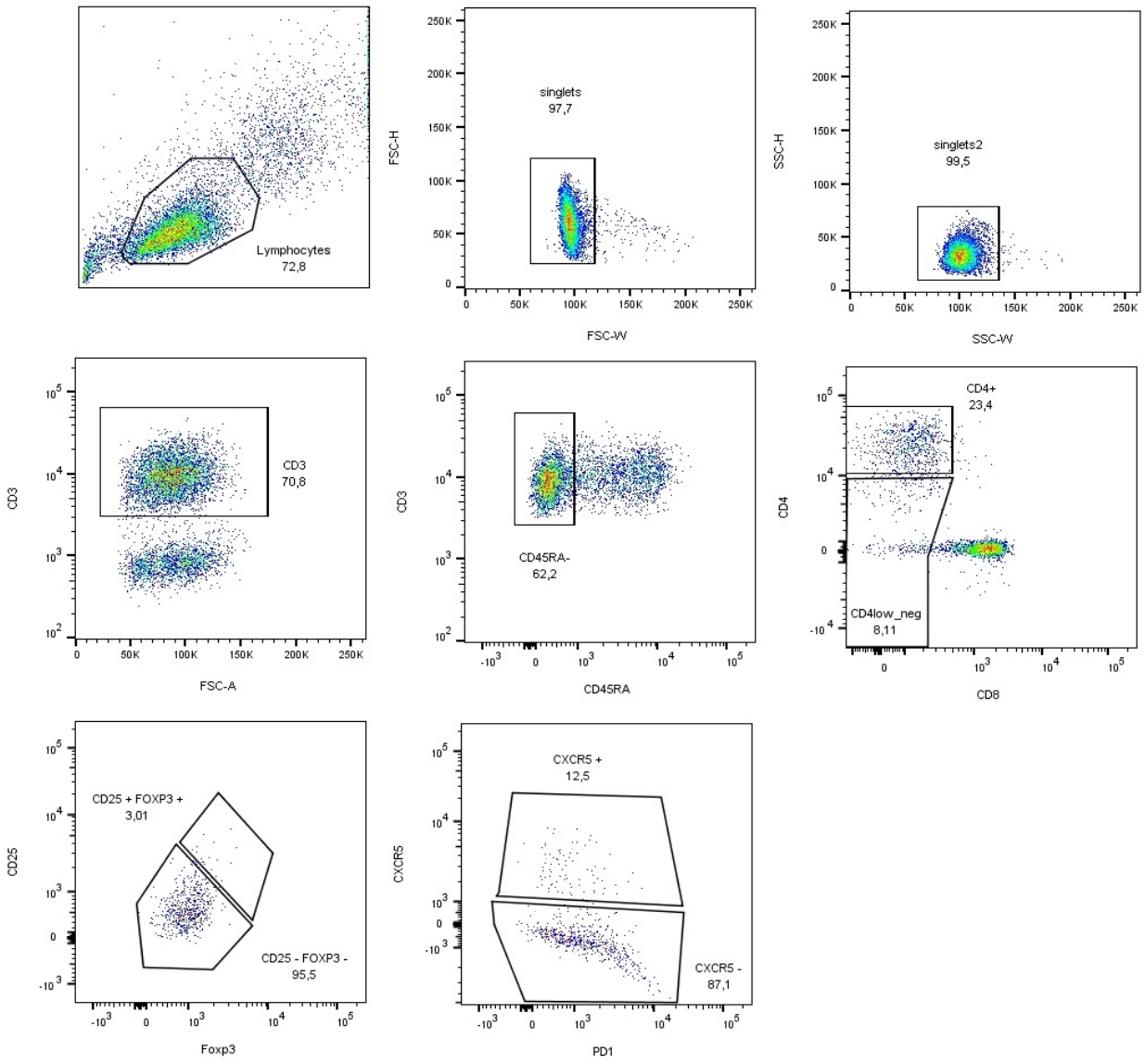


S9 Gating tree for subject H457

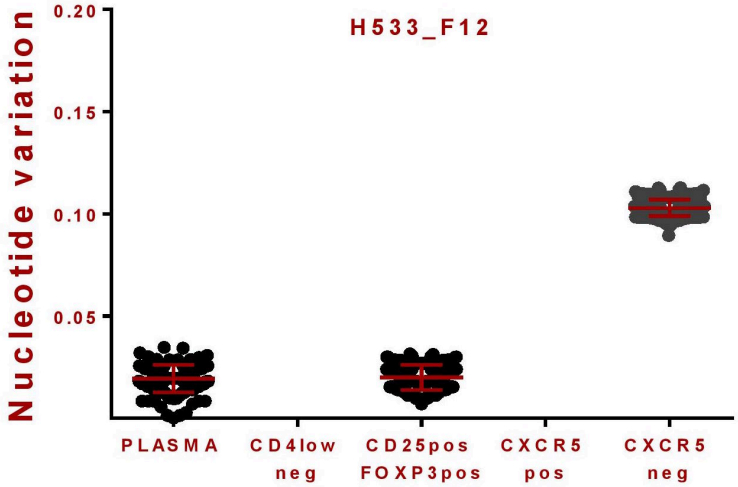


S10 Nucleotide variation between plasma and cell-derived viral nucleotide sequences for subject H457

H533

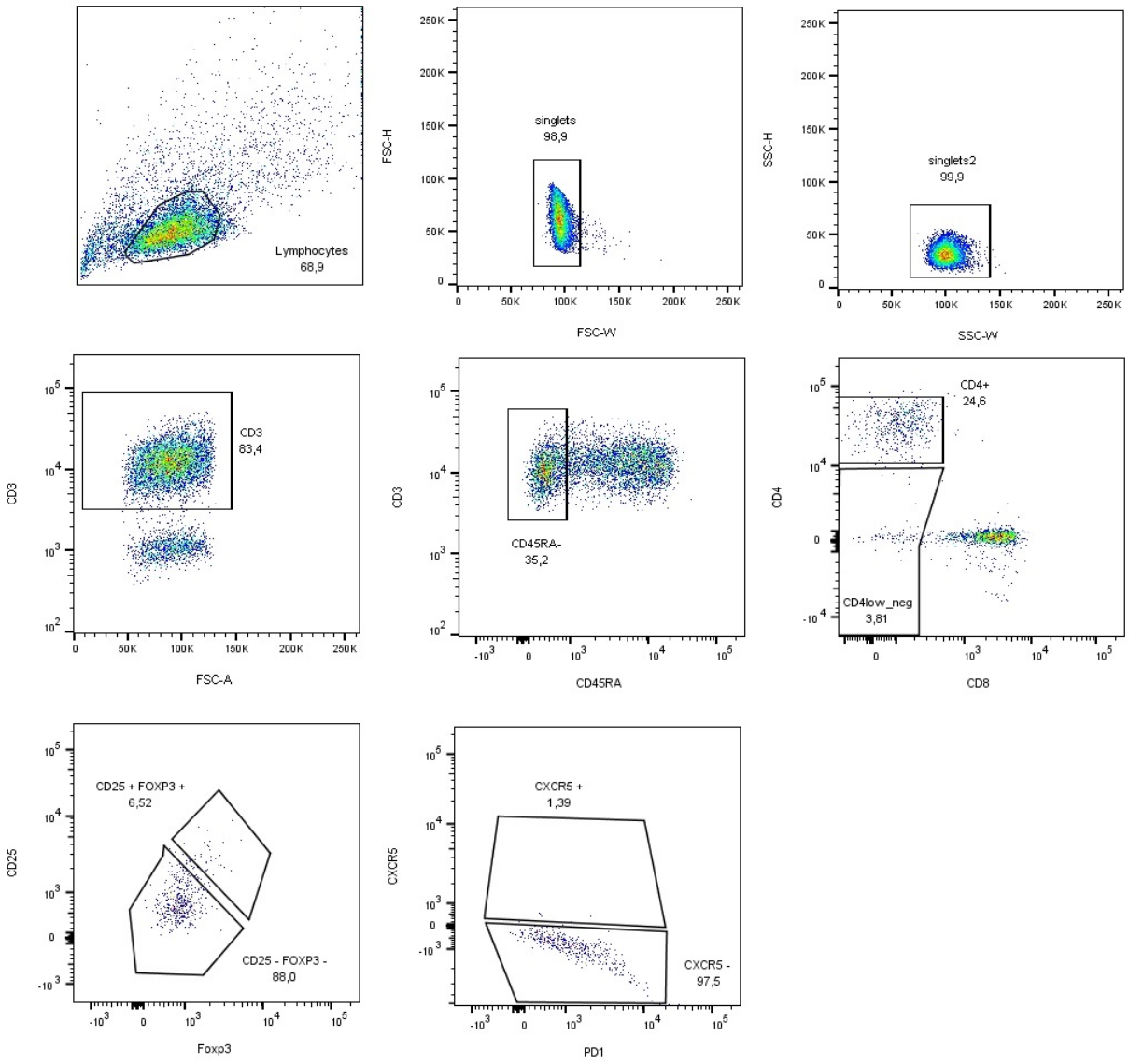


S11 Gating tree for subject H533

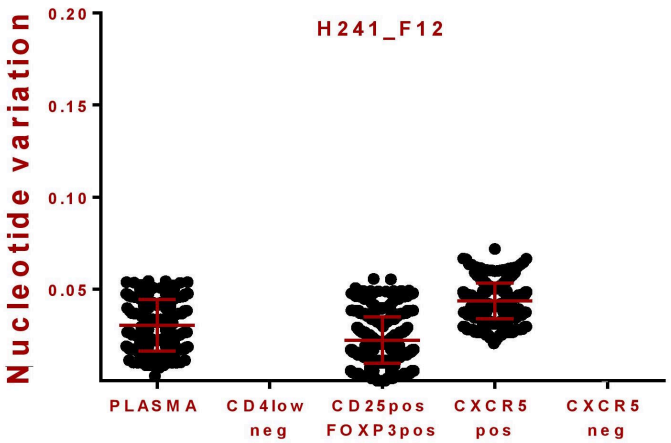


S12 Nucleotide variation between plasma and cell-derived viral nucleotide sequences for subject H533

H241

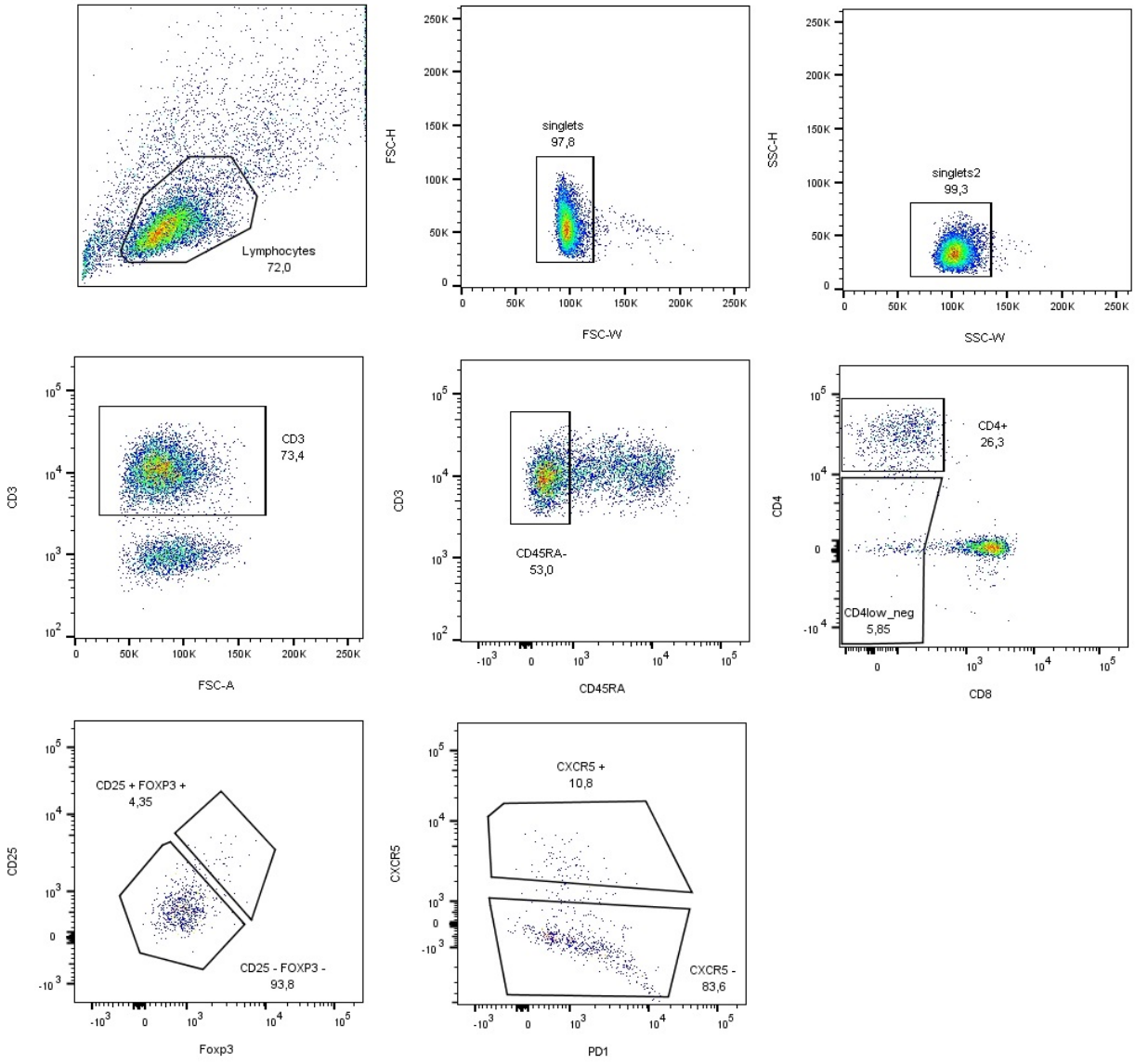


S13 Gating tree for subject H241

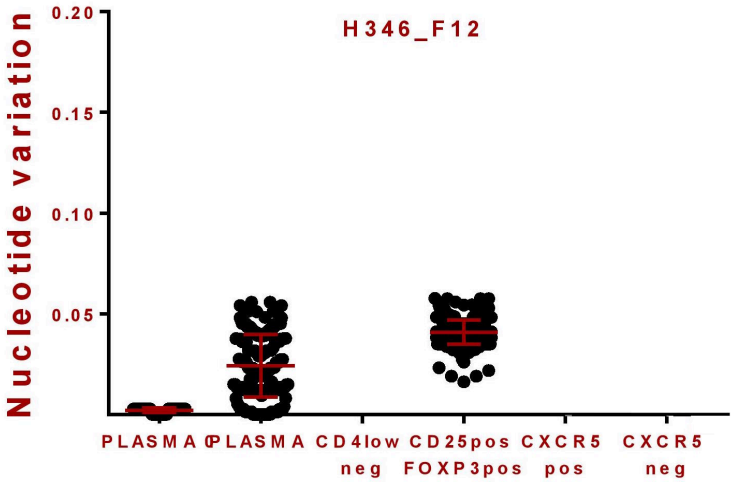


S14 Nucleotide variation between plasma and cell-derived viral nucleotide sequences for subject H241

H346



S15 Gating tree for subject H346



S16 Nucleotide variation between plasma and cell-derived viral nucleotide sequences for subject H346

7. Acknowledgements

I am heartily thankful to my supervisor, Dr. rer. nat. Christof Geldmacher, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. I would also like to express my sincere gratitude to Dr. Georgios Pollakis from the Institute of Infection and Global Health at the University of Liverpool for his support in the phylogenetic analysis and Mohammed Ahmed, from the Department of Infectious Diseases and Tropical Medicine for his support in statistical analyzing questions.

Lastly, I offer my regards and blessings to all of those who supported me in any respect during the completion of this Doctoral thesis.

Astrid Hielscher

8. Statutory declaration

I hereby declare that this doctoral thesis has been composed solely by myself, independently and without the help of others. Except where stated otherwise, by reference or acknowledgment, the work presented is entirely my own.

The work has neither been submitted to any other examining authority nor published yet.

Munich, 12.12.2018

Astrid Hielscher