Dynamic Perturbations of the T-Cell Receptor Repertoire in Chronic HIV Infection and following Antiretroviral Therapy

James M. Heather[1]*, Katharine Best[1,2], Theres Oakes[1], Eleanor R. Gray[1], Jennifer K. Roe[1], Niclas Thomas[1], Nir Friedman[3], Mahdad Noursadeghi[1] and Benjamin Chain[1]

[1] Division of Infection and Immunity, University College London, London, UK, [2] Centre for Mathematics and Physics in the Life Sciences and Experimental Biology (CoMPLEX), University College London, London, UK, [3] Department of Immunology, Weizmann Institute, Rehovot, Israel

HIV infection profoundly affects many parameters of the immune system and ultimately leads to AIDS, yet which factors are most important for determining resistance, pathology, and response to antiretroviral treatment – and how best to monitor them – remain unclear. We develop a quantitative high-throughput sequencing pipeline to characterize the TCR repertoires of HIV-infected individuals before and after antiretroviral therapy, working from small, unfractionated samples of peripheral blood. This reveals the TCR repertoires of HIV+ individuals to be highly perturbed, with considerably reduced diversity as a small proportion of sequences are highly overrepresented. HIV also causes specific qualitative changes to the repertoire including an altered distribution of V gene usage, depletion of public TCR sequences, and disruption of TCR networks. Short-term antiretroviral therapy has little impact on most of the global damage to repertoire structure, but is accompanied by rapid changes in the abundance of many individual TCR sequences, decreases in abundance of the most common sequences, and decreases in the majority of HIV-associated CDR3 sequences. Thus, high-throughput repertoire sequencing of small blood samples that are easy to take, store, and process can shed light on various aspects of the T-cell immune compartment and stands to offer insights into patient stratification and immune reconstitution.

Keywords: T-cell, TCR, repertoire, sequencing, HIV, AIDS, ART

1. INTRODUCTION

During the chronic phase of HIV infection, viral load and total T-cell numbers remain relatively stable. However, analysis of T-cell turnover reveals a more dynamic picture, with continuous virus turnover and rapid CD4 cell death offset by increased T-cell proliferation (1). HIV infection typically leads to development of AIDS, although this can be prevented with combination antiretroviral therapy (ART). ART rapidly reduces viremia to very low or undetectable levels and is associated with a long-term increase in peripheral blood CD4 T-cell numbers and reduced risk of secondary infections (2). In this study, we examine the structure of the T-cell receptor (TCR) repertoire in chronically infected treatment-naïve individuals and monitor the changes that occur immediately after viral replication is blocked by ART.

HIV infection affects many features of the immune system, and it remains unclear which changes are ultimately responsible for different aspects of the pathology. HIV infection of CD4 T-cells leads to an increased rate of cell death by both direct virological and indirect immunological mechanisms (3, 4). HIV also drives CD8[+] antigen-specific clonal expansion as well as unspecific cell activation and exhaustion (5, 6).

Many studies have investigated TCRs in HIV-infected individuals. Flow cytometry, DNA hybridization, and quantitative PCR have shown decreased expression of certain V genes (7–10). Spectratyping analysis has revealed a skewing of CDR3 length distributions in certain V families, indicating clonal expansions (11–13). A decrease in TCR diversity associated with HIV infection has also been measured using a number of experimental approaches (14–16), although such observations have typically been limited to either antigen- or subset-specific populations. The overall impact of viral cytotoxicity, viral antigenicity, and immune escape on the clonal structure of the T-cell population remain incompletely understood.

The introduction of high-throughput DNA sequencing allows analysis of the TCR repertoire at greater depth than previous techniques, with quantitative, nucleotide-level resolution (17, 18). Previous analysis of TCR repertoires from HIV patients often involved complex cell fractionation protocols, which pose a challenge in a routine clinical setting. However, we demonstrate that high-throughput analysis of TCR repertoire profiles from unseparated whole blood samples can reveal multiple levels of dysregulation in HIV patients. The diversity of the TCR repertoire is profoundly diminished during chronic infection, with expansion of large private clones driving the formation of highly idiosyncratic repertoires. HIV infection is also associated with skewing of V gene usage and disruption of the population structures that typify healthy repertoires. These perturbations remained largely unchanged after 3 months of ART, despite viral load being reduced to undetectable levels in most patients coincident with a significant increase in CD4 T-cell numbers. However, ART was accompanied by major changes at the sequence level, with many rapid, dramatic changes of frequency. Diverse immunological parameters and rapid restructuring events can therefore be captured from even small blood samples, which stands to offer new methods to monitor immune perturbation and reconstitution in patients suffering from infectious or immunodeficient conditions.

2. MATERIALS AND METHODS

2.1. Study Design

This laboratory study aimed to test whether and how deep sequencing of TCR repertoires from HIV patients might produce results of interest to clinical and immunological investigations. Adult HIV[+] ART-naïve patients due to start treatment were recruited, with the first sixteen eligible candidates being processed. Exclusion criteria included febrile or AIDS-related illness; tumors; coinfection with Hepatitis B or C; immunomodulatory therapy, recent vaccination, or breaks in treatment. 2.5 ml of peripheral blood was drawn into Tempus tubes, and RNA was extracted as per the manufacturer's instructions (Life Technologies). Residual gDNA was removed using

the TURBO-DNase kit, and globin mRNA was depleted using GLOBINclear (Life Technologies). RNA was likewise prepared from consenting adult healthy controls. Human participation in this study was approved by UK Research Ethics Committee. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

2.2. Unbiased TCR Amplification and Sequencing

Our protocol amplifies all rearrangements (within a chain) with a single primer pair, thus avoiding primer bias (Figure S1 in Supplementary Material). Addition of molecular barcodes to each cDNA molecule allows further correction for PCR amplification and errors. Primer sequences are shown in Table S1 in the Supplementary Material.

RNA was reverse transcribed using oligonucleotides (αRC2 and βRC2) directed against each constant regions' 5'. 500 ng of RNA was mixed with 1 µl of each RC2 primer (10 µM) and 1 µl of a 10-mM dNTP mix and heated at 65°C, 5 min before rapid cooling on ice. 1 µl of DTT (0.1 µM), 1 µl of RNasin (20–40 U/µl, Promega), 1 µl of SuperScript III reverse transcriptase (RT) (200 U/µl, Life Technologies) and 4 µl of 5× FS buffer were added, before incubation at 55°C, 20 min, then heat inactivated at 70°C, 15 min. RT reactions were purified and eluted into 10 µl of Tris-Cl using MinElute columns (Qiagen).

cDNA was ligated to a 5' RACE adapter, incorporating a random hexamer and SP2, one of the two sequencing primer sites, in 30 µl: 1× T4 RNA ligase buffer; 20 U T4 RNA Ligase 1 (NEB); 0.33 mM dATP; 25% PEG 8000; 1 mM hexammine chloride; 0.1 mg/ml acetylated-BSA, 0.33 µM SP2-6N ligation oligonucleotide and 5 µl of purified cDNA. Ligations were incubated for 23 h at 16°C and heat inactivated for 10 min at 65°C. Products were diluted with 70 µl of water and purified using a 1:1 ratio of AMPure XP beads (Beckman Coulter), eluting in 30 µl of water.

Ligated DNA underwent second strand synthesis, priming off SP2. Final concentrations in 50 µl were 0.2 µM dNTPs, 0.5 µM SP2, and 1 U Phusion, in 1X HF buffer. Thermal conditions were 95°C, 3 min; a slow ramp to 80°C, 10 s; another slow ramp to 58°C, 45 s; and final extension for 5 min at 72°C. Slow ramping involves descending at 0.5°C/s to encourage optimal annealing at the highest temperature. Bead-purified samples were split in two for separate chain-specific third strand synthesis. Reaction conditions were the same as the second strand reaction, except the SP2 primer was replaced with either SP1-6N-Ix-αRC1 or SP1-6N-Ix-βRC1 ("x" being a multiplexing index). Bead-purified samples then underwent "PCR1" to add the P5 and P7 elements: 0.2 mM each dNTP, 0.5 µM of each primer (P5-SP1 and P7-X-SP2), and 1 U Phusion in 1X HF buffer ("X" being another index). Initial denaturation was at 95°C for 3 min, before ramping slowly to 69°C, 15 s, and extended at 72°C, 1 min, before 3 cycles of 98°C, 10 s to 72°C, 1 min. Final extension was at 72°C for 5 min. Purified products were then subjected to "PCR2," to amplify to sufficient concentrations. Reaction conditions were the same as in PCR1, except primers used were P5s and P7. The PCR program began with initial denaturation at 95°C, 3 min, before 23 cycles of: 98°C, 10 s; 69°C, 15 s; 72°C, 40 s, then final extension at 72°C for 5 min and bead purification. Samples thus received

27 cycles of PCR in total. Final products were quantified on a Qubit (Life Technologies) and sized on a Bioanalyzer (Agilent) before dilution to 4 nM. Normalized libraries were sequenced on an Illumina MiSeq, using 2×250 PE kits.

2.3. Data Analysis

Demultiplexed FASTQ files were analyzed using a modified version of the TCR assignation software Decombinator, developed previously in our lab (19). This modified Decombinator additionally outputs barcode sequence information (comprised of the two random hexamer sequences added prior to amplification, concatenated together) in addition to the V, J, and insert information. This barcode information is then used to error-correct the TCR sequences: as the probability of any two identical TCRs acquiring the same 12-mer barcode (from the 4^{12} possibilities) is low, sequences that share a barcode and yet do not match the most frequent (within a 20% nucleotide identity threshold across the Decombinator-defined V–J spanning sequence) are assumed to be PCR or sequencing errors, and are discarded. Then, by counting the number of barcode sequences that are associated with a given rearrangement (clustering barcodes with a Levenshtein distance of 3 to allow for errors in the barcode sequence itself), we can estimate how many original cDNA molecules encoded the same TCR chain, reducing the impact of PCR duplication. Decombinator and error-correction scripts were written in Python. All remaining analyses were carried out in R, using the following packages: ineq (Gini index calculation), vegan (diversity indices), and igraph and RCytoscape (network/cluster analysis). IMGT TCR nomenclature and CDR3 definition (running from the last conserved cysteine in the V to the conserved phenylalanine in the FGXG motif in the J gene) have been used throughout (20). HIV-associated CDR3s were harvested from references (6, 21–38) and MAIT sequences from references (39, 40), while CMV-associated sequences were obtained via reference (41). Statistical significance was determined using T test (paired or unpaired as appropriate) except where indicated.

2.4. Data Availability

FASTQ data are available from the Sequence Read Archive, accession number SRP045430. Processed TCR data and published antigen-/subset-specific CDR3s are available for download (DOIs: 10.6084/m9.figshare.1153921 and 10.6084/m9.figshare.1153817). TCR analysis scripts are freely available (DOI: 10.6084/m9.figshare.1190861[1]), as is the standard version of Decombinator ((19), RRID:OMICS_00001[2]).

3. RESULTS

Two unfractionated blood samples were taken from HIV^+ patients – one immediately prior to (S1) and another 3 months after commencing ART (S2) – from which RNA was extracted. Ten uninfected donors were also sampled as controls, four of whom were sampled twice 3 months apart. RNA was reverse

transcribed, and cDNA molecules were labeled with molecular barcodes before amplification and sequencing (Figure S1 in Supplementary Material).

Patient CD4 counts showed a modest but significant increase in most individuals (14/16) after 3 months of therapy, yet remained below normal values (Figure S2A in Supplementary Material). CD8 counts showed no consistent change over this period, while the CD4:CD8 T-cell ratio slightly increased (Figures S2B,C in Supplementary Material). Viral genome counts decreased rapidly and were undetectable in 11/16 individuals by the second sample collection (Figure S2D in Supplementary Material) and in all patients after a further year of ART. Clinical data and the number of sequences obtained are summarized in Table S2 in the Supplementary Material.

3.1. HIV$^+$ TCR Repertoires Show a Profoundly Decreased Richness and Diversity Which Does Not Recover During Short-Term ART
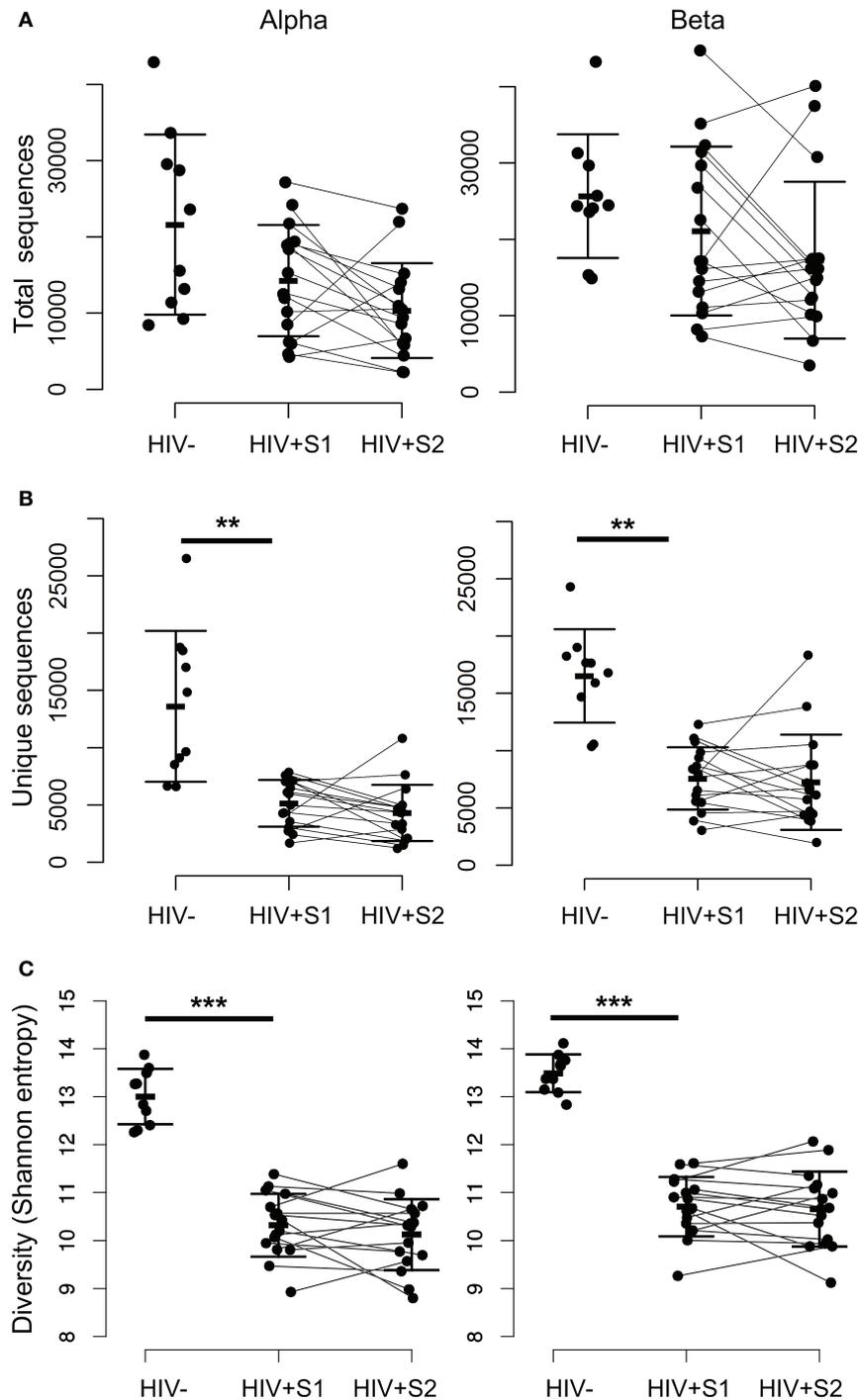
The complementarity determining region 3 (CDR3) section of a given T-cell receptor is that which is encoded at the hyper-variable recombination junction, and forms the primary contacts with the peptide-MHC antigen (42). Therefore, as the major site of biological interest, it is the primary sequence investigated in this study. There is no significant difference in the total number of TCR CDR3 sequences of either chain between HIV$^-$ and HIV$^+$ repertoires (**Figure 1A**), in agreement with the lack of total T-cell lymphocytopenia in these patients. In contrast, there is a significant decrease ($p \ll 0.01$) in the number of distinct sequences observed (**Figure 1B**), as HIV$^+$ samples contain on average threefold fewer sequences per chain. The TCR repertoire is determined both by the number of different sequences (richness) and the relative abundances of each clonotype. The combination of these properties is typically quantified as diversity, with decreased diversity likely reflecting a decreased ability of the repertoire to recognize and respond across the spectrum of potential pathogen antigens (43). One frequently used measure of diversity is the Shannon information index, or Shannon entropy, which is significantly decreased ($p < 0.01$) in untreated HIV$^+$ repertoires (**Figure 1C**) and shows no recovery after therapy. This low diversity shows a weak inverse correlation to viral load (Pearson correlations: alpha chain R $= -0.48$, $p = 0.06$; beta chain R $= -0.41$, $p = 0.12$). Random sampling of the same number of sequences to produce size-matched repertoire files shows the same significant decrease in diversity, thus it is not a feature of the lower number of sequences found in patient repertoires (Figure S2E in Supplementary Material).

3.2. Altered V Gene Usage in Chronic HIV Infection

The distribution of V gene usage for both chains, measured for distinct sequences (counting each TCR only once to avoid bias from the effects of clonal amplification), is shown in **Figures 2A,B**. As reported previously, the expression pattern of different V genes is non-uniform and conserved between individuals (44–46). Eight V alpha and eight V beta genes showed significant

---

FIGURE 1 | The repertoires of HIV-infected individuals have a smaller number of distinct sequences, and lower diversity. **(A)** The total number of alpha (left) and beta chain (right) CDR3 sequences obtained from each HIV⁻ and HIV⁺ individual, at both the pretreatment (S1) and treated (S2) bleed. The bars show mean and SD. **(B)** As for **(A)** but showing the number of unique CDR3 sequences observed within each set. The difference between the mean of the HIV⁻ and HIV⁺ S1 samples was significant, $p < 0.01$. There was no significant difference between HIV⁺ S1 and HIV⁺ S2. **(C)** As for **(A)** but showing CDR3 diversity measured by the Shannon information index (or Shannon entropy) for each sample.

differences in mean expression between patients and controls, suggesting that HIV infection might differentially modulate certain parts of the repertoire, for example by driving expansion of HIV (or coinfecting pathogen) reactive T-cells using specific genes. A recent repertoire analysis of Treg cells sorted from HIV-infected individuals (47) showed a similarly non-uniform V and J

FIGURE 2 | Altered distribution and interindividual differences in V gene usage in the repertoires of HIV+ individuals. The frequency of V alpha **(A)** and V beta **(B)** gene usage for the unique TCR repertoires of each donor sample within the two groups (HIV− in black or untreated HIV+ in red). Bars show the mean proportion for each V gene in each group. Asterisks (*) show those V genes, which differ significantly in their usage (T test, $p < 0.05$) between HIV− and HIV+. Interindividual differences in patterns of V gene usage (measured using Jensen–Shannon distances) for alpha **(C)** and beta **(D)** V genes. The plot shows the median and interquartile range of the Jensen–Shannon distances between each individual and all other individuals in turn from the same group. HIV− are shown in black; HIV+ are shown in red.

gene distribution, with many of the most frequently used V genes observed in that study being similarly over represented in this study (e.g. TRBV6-5).

V gene usage profiles between HIV$^+$ repertoires also differed by a greater amount than between healthy volunteer repertoires. This was measured by calculating Jensen–Shannon distances for each pair of V gene usage distributions; HIV$^+$ pairs showed greater distances than HIV$^-$ pairs (**Figures 2C,D**). Chronic HIV infection therefore leads to severe idiosyncratic perturbations of T-cell compartments, driving each individual's repertoire away from typical ranges into their own distinct state. In contrast, the distributions of J gene differences showed few significant differences between HIV$^+$ and HIV$^-$ repertoires (Figure S3 in Supplementary Material). This may reflect the shorter length and greater propensity for base removal during recombination in the J genes relative to the V genes, and in the case of beta chain sequences, there being far fewer J genes for rearrangements to draw upon.

Comparing the distribution of V genes of HIV$^+$ repertoires before and after therapy revealed that those relatively perturbed genes tended to return towards healthy values (Figure S4 in Supplementary Material) but changes observed were small.

3.3. Decreased TCR Sequence Sharing Between Repertoires of HIV$^+$ Individuals

Despite the stochasticity inherent in V(D)J recombination, TCR sequences can often be found to be shared between individuals, being classified as *private* when they occur in only a few individuals and *public* when found in many (45, 48–51). To see whether HIV infection impacts upon TCR sharing, we first compared the proportion of CDR3s shared between pairs of HIV$^+$ individuals to that of HIV$^-$ pairs by using the Jaccard index (a normalized measure of sharing, **Figure 3A**), which revealed a profound loss of sharing among HIV-infected patients relative to healthy controls. This property is not a feature of sample size, as size-matching data by random selection shows the same result (Figure S5A in Supplementary Material). Similarly, by counting the number of CDR3s that occur in 5 or more out of 10 HIV$^-$ or HIV$^+$ individuals, we see a decrease in the number of highly shared sequences (**Figure 3B**), again indicating a loss of TCR sharing between HIV$^+$ patients and a tendency toward idiosyncratic repertoires. Figure S5B in Supplementary Material shows the intradonor Jaccard indices (i.e., the overlap of each individual's S1 and S2 samples) for reference.

3.4. HIV Infection Is Associated with Disruption of Networks of Similar CDR3s

In order to explore the connectivity between CDR3s, we measured the distance between every pair of CDR3 amino acid sequences in an individual, using the Levenshtein distance (the minimum number of changes required to turn one sequence into another). We then constructed networks by creating an edge between unique CDR3 sequences (nodes), which differ by a distance of one (**Figure 3C**; Figure S6 in Supplementary Material). The size of the largest alpha CDR3 network (cluster size) and the maximum number of edges going to any one node (the maximal degree) for each individual are shown in **Figures 3D,E**. The maximum cluster size and the maximum network degree in the HIV$^+$ repertoires

are much smaller than the largest clusters in the uninfected repertoires, indicating landscapes of CDR3s that are less similar than those seen in controls. Healthy donor alpha repertoires are more similar, revealed by increased frequency of CDR3s, which differ by a Levenshtein distance of five or less (**Figure 3F**). No differences were seen between patients and controls for these parameters in the beta chain networks (Figure S7 in Supplementary Material), perhaps reflecting the greater diversity of the beta sequences owing to the presence of TRBD genes.
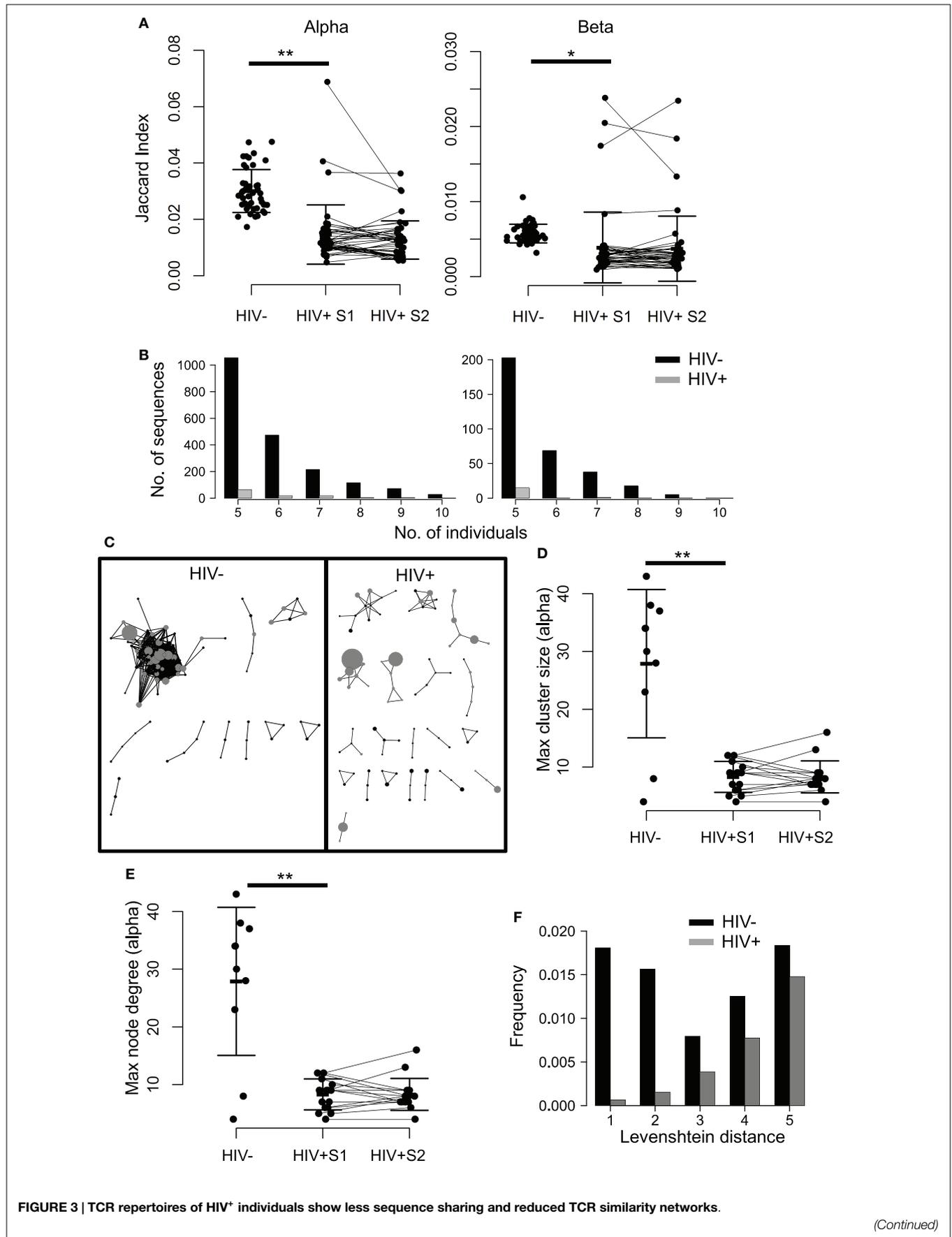
3.5. ART Is Accompanied by Large Changes in the Abundances of Individual Alpha and Beta Chain Sequences

By averaging the 100 most common sequences per donor, we observed that the most abundant CDR3s in HIV$^+$ samples are significantly more frequent than controls (**Figure 4A**; Figure S8A in Supplementary Material). While the average frequency falls slightly following therapy, it remains higher and more variable than in HIV$^-$ controls. These highly abundant sequences skew the repertoire frequency distributions of HIV$^+$ individuals, with the most common sequences in patients occupying over twice as much of the repertoire than the most common sequences found in uninfected controls (**Figure 4B**).

The Gini index is a measure of unevenness that can measure this skewness, which ranges from zero for a uniform population (all TCRs being equally present) and tending toward one (maximum inequality). The mean Gini index is significantly increased in HIV$^+$ samples compared to the healthy controls (**Figure 4C**), which could be considered reflective of increased oligoclonality. In contrast to the diversity as measured by Shannon entropy, the mean Gini index shows a significant recovery following therapy, yet remains above the mean of healthy controls. However, this observation is not replicated in size-matched data (Figure S8B in Supplementary Material). Interestingly, the Gini indices of the pre-treatment patient data significantly correlate to their CD8 counts (alpha R = 0.63, p = 0.01; beta R = 0.50, p = 0.05) but are independent of CD4 counts (alpha R = 0.21, p = 0.44; beta R = 0.06, p = 0.81). It is worth noting that the slight differences between the age and sex profiles of our control and uninfected donor cohorts fail to explain the differences observed in either measure of diversity (Figure S9 in Supplementary Material).

Because of the degeneracy of the genetic code and the similarity of germline TCR genes, identical CDR3s can be created by multiple DNA sequences (here referred to as translational convergence). By plotting the average number of different TCR rearrangements which translate to produce the 100 most abundant CDR3s in each group, we see lower translational convergence in the patient group (**Figure 4D**). Thus, despite the abundant CDR3s in HIV patients being far more frequent, they are translated from fewer DNA sequences. This suggests that abundant CDR3s in HIV patients arise from large clonal expansions.

We tracked the frequencies of the 100 most abundant sequences per chain before and after therapy (example donors shown in **Figure 5**, with all donors in Figure S10 in Supplementary Material). Frequency histograms of the ratio of sequence abundance in S2 versus S1 (**Figure 5B**) show that HIV$^+$ samples show more dispersion than controls, indicating that these sequences

FIGURE 3 | TCR repertoires of HIV+ individuals show less sequence sharing and reduced TCR similarity networks.

(Continued)

undergo larger frequency changes compared to the most common sequences in uninfected donors.

The histogram of the HIV⁺ samples shows one peak close to 1 (similar to HIV⁻ histograms) and another near 0.01, i.e., a decrease of approximately 100-fold. These changes occurred over 14 weeks which translates to a half-life on the order of 30 days, compatible with the known half-life of effector memory T-cells (52).

3.6. Mucosal-Associated Invariant T-Cell Sequences Are Depleted in HIV⁺ Repertoires and Do Not Recover Following ART

We searched our data for sequences belonging to non-MHC restricted T-cell subsets, many of which use TCRs which include "invariant" alpha chains, those which use a tightly conserved selection of V and J genes and near-identical CDR3s. Sequences associated with invariant natural killer T (iNKT) and germline-encoded mycolyl lipid-reactive (GEM) cells were extremely rare in our data, consistent with a reported very low frequency in blood. In contrast, CDR3s derived from mucosal-associated invariant T (MAIT) cells – which can constitute upwards of 10% of the healthy CD8⁺ T-cell pool (53) – were present in the context of the correct genes (TRAV1-2 with TRAJ33, TRAJ12, or TRAJ20) at high frequencies in healthy controls (**Figure 6A**). MAIT CDR3s were severely depleted in HIV⁺ repertoires and showed no evidence of recovery following therapy. V genes from beta chains known to pair with the invariant MAIT alpha chains (TRBV6-1 and TRBV20-1) (39) were among the few that were significantly decreased in the repertoires of HIV-infected donors (**Figure 2A**).

3.7. ART Induces Rapid Changes in the Numbers of HIV- and CMV-Associated CDR3s

While this study was not designed to gather alpha-beta chain pairing information, there are a number of reports in which even single chains from T-cells reactive against a particular antigen share sequence identity or features, and some TCRs can recognize the same peptide in the context of different MHC molecules (26, 54–56). Furthermore, there are several published reports of HIV-specific public TCR responses (21, 32, 33), sometimes even finding the same CDR3 in combination with different V and J genes (57). We therefore collected the CDR3 sequences of HIV-specific T-cells from the literature (92 alpha and 702 beta) and searched for matches in our data. A set of CDR3s from
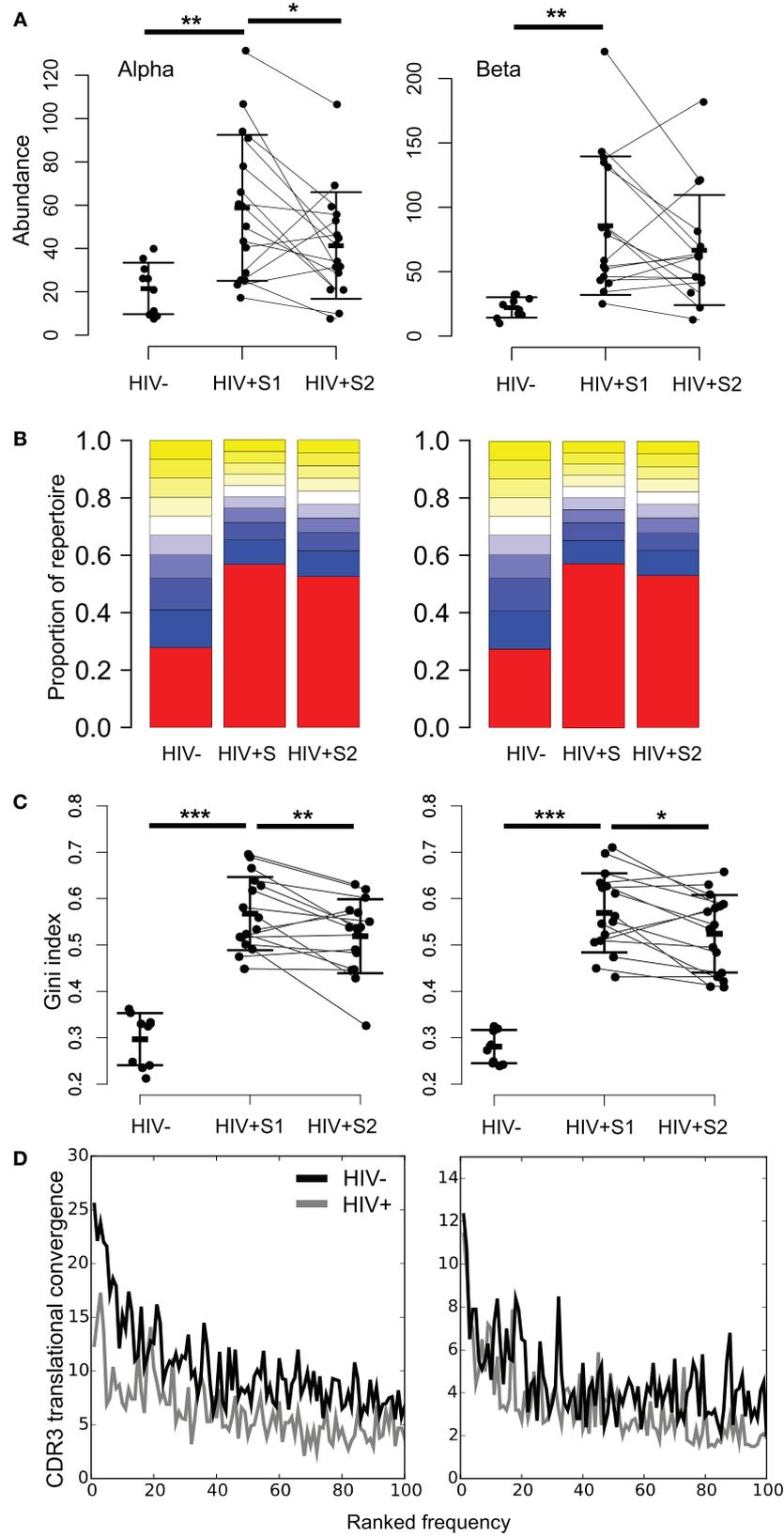
CMV-specific T-cells were also collected (13 alpha and 253 beta), as an example of another chronic virus known to drive large oligoclonal expansions (58–60). Because these sequences were typically obtained by isolating viral-specific pMHC-tetramer⁺ cells, the HLA restriction of the set was highly skewed to a small number of different genotypes.

We detected 15 examples of expanded HIV-associated CDR3s in our data (equal to or greater than five copies in at least one of a patients' bleeds), comprised of 11 different CDR3s found across seven different individuals (**Figure 6B**). Some individuals had multiple matches, perhaps reflecting a match between their HLA type (not determined here) and those used to originally isolate the sequences. We found 17 examples of expanded CMV-associated CDR3s, comprised of ten different CDR3s observed in eight individuals. There was considerable overlap between the set of patients containing HIV- and CMV-associated CDR3s, again perhaps reflecting HLA genotype. Only two virally associated CDR3s were found in this analysis in the uninfected repertoires and only at the edge of detection at a single time point. As a control, we selected ten similarly sized sets of CDR3s from an independent published set derived from patients with Juvenile Idiopathic arthritis (61). Only an average of 4.5 of these CDR3s (±SD of 0.8) were found in our patients, suggesting the viral-associated CDR3 frequencies do indeed reflect actual virus exposure. Over the course of therapy, HIV-associated CDR3s significantly decreased in frequency, while CMV-associated sequences did not significantly alter (**Figures 6C,D**), although the majority of them (12 out of 17) did become more frequent.

4. DISCUSSION

We have developed a high-throughput sequencing pipeline to measure quantitative parameters of the TCR repertoires of a cohort of HIV-infected patients, before and 3 months after the start of antiretroviral therapy. From a small unsorted blood sample, we are able to extract tens of thousands of TCR sequences per donor, providing the most comprehensive view of the global impact of HIV infection upon the TCR repertoire to date. Despite a relatively small cohort size, the results robustly identify multiple aspects of immune dysregulation and document some interesting changes that follow initiation of ART.

Dynamic changes in the TCR repertoire are a common feature of many viral infections. However, chronic HIV infection presents a number of special characteristics. During chronic HIV infection, the repertoire is severely restricted, averaging threefold fewer distinct sequences than uninfected controls. This depletion
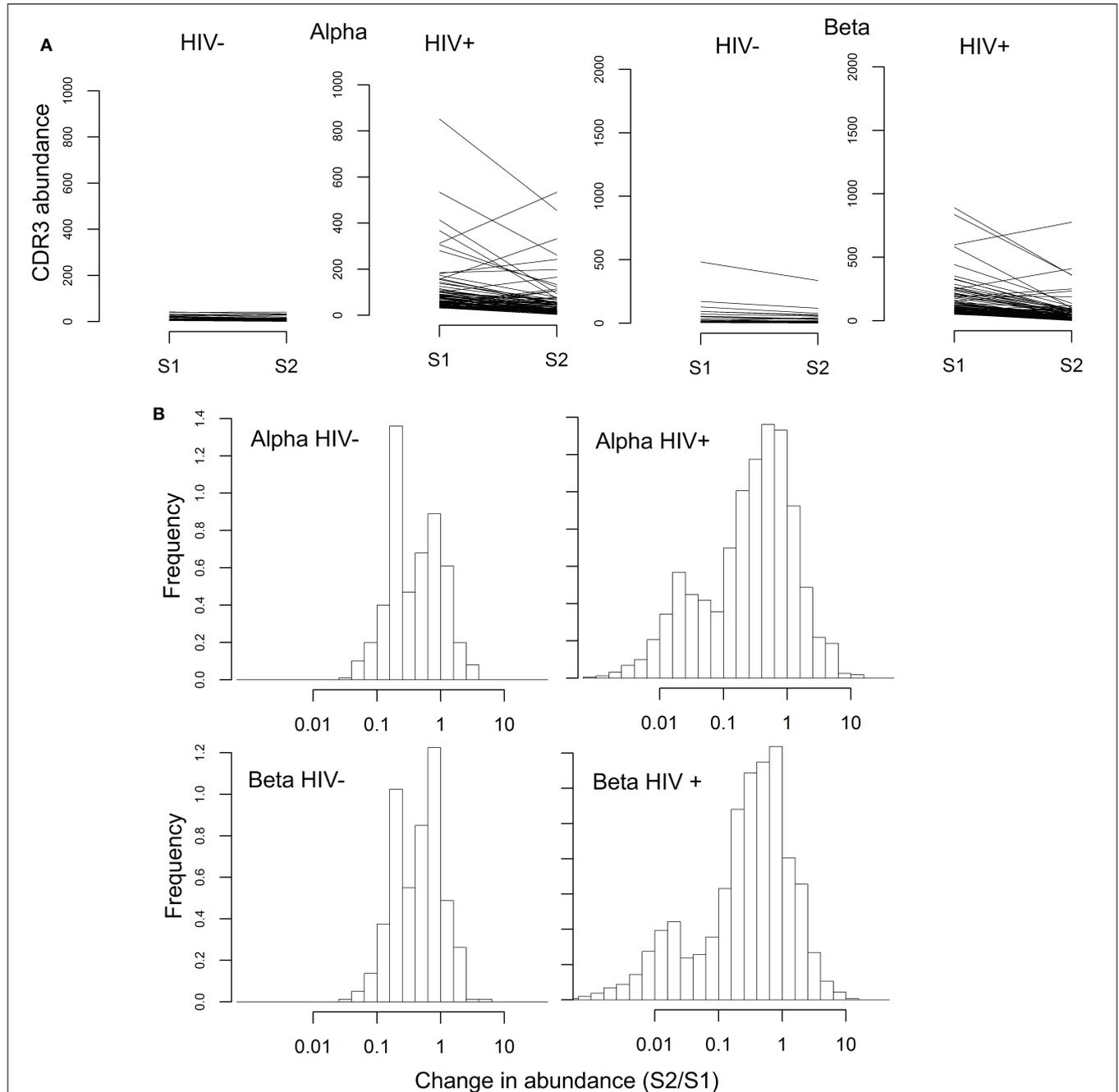
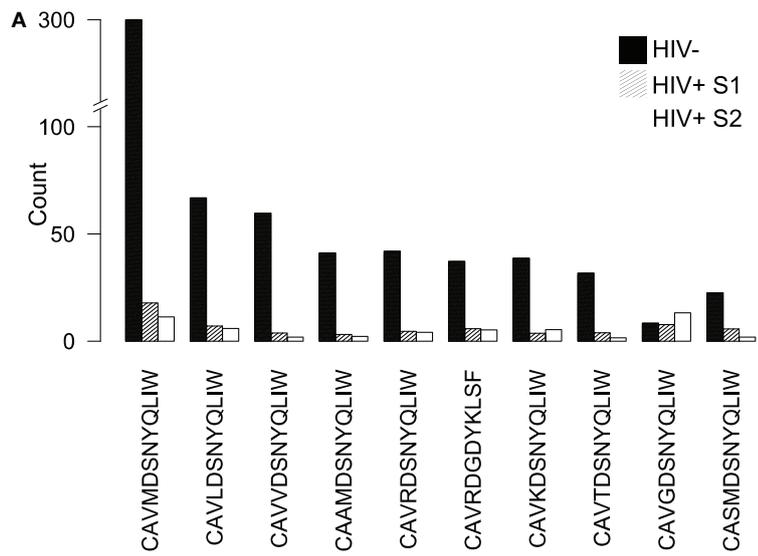**FIGURE 4 | HIV⁺ repertoires have a skewed TCR chain abundance distribution**.

(Continued)

FIGURE 4 | Continued

(A) The mean abundance of the 100 most abundant alpha (left) and beta chain (right) CDR3s in HIV$^-$ and HIV$^+$ before (S1) and after (S2) therapy. Bars show SD. T test: $**p < 0.01$, $*p < 0.05$. **(B)** Size distribution of TCR chain abundances. The plot shows the proportion of alpha or beta sequences in the HIV$^-$ and HIV$^+$ groups before (S1) and after (S2) therapy, which fall within each 10 percentile range of the size distribution. The percentile ranges are shown in decreasing order from the largest 10% (red) to the smallest 10% (yellow at top of each bar stack). **(C)** The Gini index of the TCR alpha or beta chain abundance distributions in HIV$^-$ and HIV$^+$ before (S1) and after (S2) therapy. T test: $***p < 0.001$, $**p < 0.01$, and $*p < 0.05$. **(D)** Average translational convergence of alpha and beta CDR3s in HIV$^+$ and HIV$^-$ repertoires. The HIV$^+$ and HIV$^-$ distributions differ significantly (one-way ANOVA, $p < 0.01$).
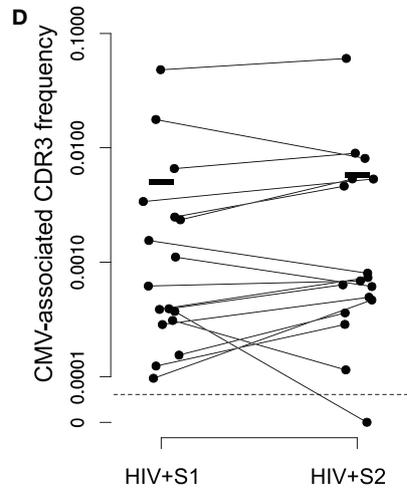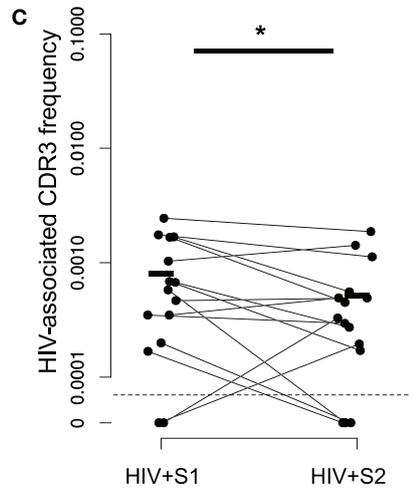

FIGURE 5 | Rapid changes in abundance of individual TCR sequences following ART. **(A)** Abundance of the 100 most common alpha and beta sequences before (S1) and after (S2) ART therapy, in a representative HIV$^-$ and HIV$^+$ individual. **(B)** Change in abundance of the 100 most common alpha and beta sequences following ART, expressed as a ratio of abundance in S2 over the abundance in S1. Sequences absent in one of the two samples were assigned a size of 1 in that sample, so that the ratio between the two samples did not become infinite.

FIGURE 6 | Changes in frequency of MAIT and virally associated CDR3 sequences in HIV repertoires and following ART. (A) Average numbers of the ten most frequent MAIT alpha chain sequences in the HIV[−] (filled bars) and HIV[+] repertoires before (dashed bars) and 3 months after (open bars) the start of ART. All sequences required an exact sequence match to a published MAIT CDR3 in the context of the correct germline V and J TCR genes. These genes are TRAV1-2 and TRAJ33, with the exception of CAVRDGDYKLSF, which instead uses the non-canonical TRAJ20 pairing. **(B)** Published CDR3 sequences from T-cells reported to be HIV or CMV specific were found in the repertoires of the HIV[+] individuals shown. CDR3 sequences were only included if they occurred at an abundance of five or greater in either of the two of a patient's samples. Underlining indicates alpha chain sequences, the rest being beta chain. **(C,D)** The proportion that each CDR3 listed in **(B)** occupies in the repertoires in which they were found, in both their pretreatment (S1) and midtreatment (S2) samples, for both HIV- and CMV-associated CDR3s (**(C,D)**, respectively). Horizontal lines indicate mean frequencies. The average clone size of the HIV CDR3s in sample 2 was significantly less than the size in S1 (*T* test, $p < 0.05$).

is consistent with the short life span of infected CD4$^+$ T-cells ([1], [62]) and significant cell death in bystander uninfected cells early in disease progression ([3]). The shrinkage of the repertoire is opposed by an overexpansion of a small proportion of TCRs. The decrease in CD4$^+$ counts and CD4:CD8 ratios combined with the positive correlation between Gini index and CD8$^+$ T-cell count suggest that repertoire skewing can be attributed both to the decline in CD4 cells and large expansions, primarily in the relatively oligoclonal CD8 compartment. This hypothesis needs to be tested further by analyzing the TCR repertoire of sorted CD4 and CD8 T-cell populations, as well as various effector and memory subsets. It is also worth noting that large clonal expansions might not necessarily be directed against HIV antigens but might recognize those belonging to other common coinfections: indeed, there is some evidence that the majority of T-cell responses in untreated patients are not directed toward HIV peptides ([63]).

HIV infection was also associated with qualitative repertoire changes. Repertoires from healthy adults are characterized by highly conserved non-uniform distributions of different V and J genes, thought to reflect intrinsic features of the recombination process ([45], [49], [64]). These distributions are significantly altered in the repertoires of the HIV-infected individuals, which may partially reflect dramatic depletion of some non-classical T-cell subsets, such as MAIT cells, which preferentially use certain V and J TCR genes. The HIV-associated decrease in MAIT cells has been reported previously ([53], [65]) and is consistent with the extensive damage to the gastrointestinal immune compartment, which is a feature of HIV infection ([66], [67]). Further studies looking at the repertoires of gamma-delta T-cells – which are enriched in gastrointestinal tissues ([68]) – and those of all T-cells from gut-associated lymphoid tissue will be interesting in this regard.

Healthy repertoires are also characterized by the presence of public CDR3s, which are shared between many different individuals ([45], [48]–[51]). We also observed that repertoires from healthy donors contain sets of CDR3s with very similar amino acid sequences and that these sets often contain some of the most abundant CDR3s observed. Here, we observed that HIV infection associates with a profound depletion both of the number of CDR3s shared between people and CDR3 similarity within individuals. This decreased sharing may, in part, reflect differences in HLA types between individuals, although we have no reason to believe the HLA diversity is greater within the HIV cohort than among the controls. It may also be driven by heterogeneity in the microbial antigens presented over the course of infection, both in terms of the HIV mutating and in the complement of other infections a given patient has. The function of public CDR3s and CDR3 networks remain unknown, although both features have been suggested to play a role in the maintenance of self-tolerance ([50]). Disruption of these features may therefore aggravate the immune dysregulation, which is a feature of chronic HIV infection.

ART is a rapid and effective means to lower viral burden and initiate restoration of immune function. Paradoxically, recovery of the many parameters of immune function (e.g. CD4 count) occurs much slower, requiring years to return to normal ranges, likely related to low adult thymic output. TCR diversity and gene usage patterns are at best only partially restored after 3 months of ART. However, at an individual sequence level, initiation of therapy is accompanied by many rapid sequence expansions and contractions, including a reduction in the frequency of HIV-associated CDR3s. The T-cell receptor repertoire is therefore highly dynamic and undergoes substantial reshaping, which accompanies the rapid fall in viral load. Further functional studies will be needed to establish the causal relationship between the individual clonal dynamics and the increased ability to control opportunistic infections, which results from ART. It will also be instructive to obtain longer follow up samples from the HIV-infected cohort to document the long-term reconstitution of the immune repertoire. In particular, it will be interesting to see if the aberrant clonal expansions observed in the HIV repertoires remain and play any role in the chronic inflammatory phenotype, which seems to contribute to long-term morbidity in HIV-infected individuals, even when viral load is controlled.

This study only processed blood from sixteen HIV$^+$ individuals, which constitutes a major limitation regarding the possible widespread applicability of our findings. Additionally, our treated S2 bleeds were taken after only approximately 3 months of antiretroviral therapy, which might be considered a relatively short time period to observe immune reconstitution in adult humans, given their reduced thymic output. However, we do observe a significant recovery in CD4 cell numbers over this period, in agreement with studies tracking CD4 cell numbers in larger cohorts of patients, which often show that the first 3 months of treatment produce the greatest absolute and relative CD4 T-cell recovery ([69], [70]). A related sampling limitation is that concerning the difference in number of different TCR sequences obtained from either HIV$^+$ individuals or uninfected controls, which makes their analysis non-trivial. One can try to address this by randomly selecting equally sized samples from both groups, yet this results in analyses effectively including a higher proportion of rare events from the HIV$^+$ data.

One of the major strengths of the approach described in this study is that all findings are derived from a fraction of RNA extracted from a small sample of peripheral blood. These samples are easily taken and stored, and using whole blood obviates the need to fractionate different cell populations, avoiding the cost, technical barriers, and health risks involved in processing potentially infectious biological material. However, this arrangement also imposes limitations on the depth of repertoire sequencing and the ability to separate out dynamics of different T-cell populations. If required – for instance to track rare T-cell clones – depth could be increased in future studies by taking more blood, processing more RNA or sequencing final libraries more deeply, yet as many of the interesting phenomena discussed here were manifest at a population structure level, this may not be required. Subpopulation information might also be inferred from correlations with routine clinical blood counts, as presented here, or might even be estimated from the properties of the TCR chains themselves, especially in the case of CD4 and CD8 T-cell sequence differentiation as their repertoires show marked divergence in several respects ([71]). This sequencing protocol could of course also be applied to sorted cell populations were such data required. Another limitation of the data presented here is one which affects the majority of the high-throughput TCR repertoire published data sets, thus far, in that we do not know which alpha chains

pair with which betas in our samples. Such data would allow the determination of actual clonotypes and would permit cloning and functional testing of sequences of interest. However, while there are a number of protocols which potentially offer this ability – involving single-cell sorting (72), cellular emulsion PCR (73), in-cell nucleic acid linking (74), or bioinformatic inference of paired chains (75) – none are yet capable of producing data from the same volume of cells as bulk sequencing without large increases in handling time or cost.

This study therefore demonstrates a means by which TCR repertoire sequencing even from small, unfractionated blood samples – collection of which could be readily incorporated into a number of clinical studies – can be used to investigate a wide variety of immune parameters. Our results confirm and build upon a number of previous studies investigating repertoire perturbation during HIV infection. We are able to simultaneously track repertoire diversity, gene usage, and specific sequences in the data produced from a single experiment. These data highlight the dynamic nature of the repertoire, which is rapidly reshaped following the ART-induced fall in HIV load. These dynamic changes may lead to improved immune function despite a persistent contraction of the repertoire. In contrast, long-term damage to the repertoire may be much more difficult to repair and may contribute to long-term morbidity. Further studies, following larger cohorts of HIV$^+$ individuals through the various phases of HIV infection and after long-term therapy, will help determine the clinical and immunological correlates of perturbation of the TCR repertoire. These data will help inform a rational approach to treatment optimization and may offer novel biomarkers for patient stratification.

Moreover, this study demonstrates the potential of TCR repertoire sequencing for monitoring patients with infectious or immunodeficient diseases.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://journal.frontiersin.org/article/10.3389/fimmu.2015.00644

REFERENCES

1. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* (1995) **373**:123–7. doi:10.1038/373123a0

2. Moore RD, Chaisson RE. Natural history of HIV infection in the era of combination antiretroviral therapy. *AIDS* (1999) **13**(14):1933–42. doi:10.1097/00002030-199910010-00017

3. Alimonti JB. Mechanisms of CD4+ T lymphocyte cell death in human immunodeficiency virus infection and AIDS. *J Gen Virol* (2003) **84**(7):1649–61. doi:10.1099/vir.0.19110-0

4. Doitsh G, Galloway NLK, Geng X, Yang Z, Monroe KM, Zepeda O, et al. Cell death by pyroptosis drives CD4 T-cell depletion in HIV-1 infection. *Nature* (2013) **505**:509–14. doi:10.1038/nature12940

5. Kuchroo VK, Anderson AC, Petrovas C. Coinhibitory receptors and CD8 T cell exhaustion in chronic infections. *Curr Opin HIV AIDS* (2014) **9**(5):439–45. doi:10.1097/COH.0000000000000088

6. Wilson JDK, Ogg GS, Allen RL, Goulder PJR, Kelleher A, Sewell AK, et al. Oligoclonal expansions of CD8+ T cells in chronic HIV infection are antigen specific. *J Exp Med* (1998) **188**(4):785–90. doi:10.1084/jem.188.4.785

7. Bahadoran P, Rieux-Laucat F, Le Deist F, Blanche S, Fischer A, de Villartay JP. Lack of selective V beta deletion in peripheral CD4+ T cells of human immunodeficiency virus-infected infants. *Eur J Immunol* (1993) **23**(8):2041–4. doi:10.1002/eji.1830230850

8. Imberti L, Sottini A, Bettinardi A, Puoti M, Primi D. Selective depletion in HIV infection of T cells that bear specific T cell receptor V beta sequences. *Science* (1991) **254**(5033):860–2. doi:10.1126/science.1948066

9. Laurence J, Hodstev AS, Posnett DN. Superantigen implicated in dependence of HIV-1 replication in T cells on TCR V beta expression. *Nature* (1992) **358**:255–9. doi:10.1038/358255a0

10. Rebai N, Pantaleo G, Demarest JF, Ciurli C, Soudeyns H, Adelsberger JW, et al. Analysis of the T-cell receptor beta-chain variable-region repertoire in monozygotic twins discordant for human immunodeficieny virus: evidence for perturbations of specific Vbeta segments in CD4+ T cells of the virus-positive twins. *Proc Natl Acad Sci U S A* (1994) **91**(February):1529–33. doi:10.1073/pnas.91.4.1529

11. Connors M, Kovacs J, Krevat S, Gea-Banacloche J, Sneller M, Flanigan M, et al. HIV infection induces changes in CD4+ T-cell phenotype and depletions within the CD4+ T-cell repertoire that are not immediately restored by antiviral or immune-based therapies. *Nat Med* (1997) **3**(5):533–40. doi:10.1038/nm0597-533

12. Gea-Banacloche JC, Martino L, Mican JM, Hallahan CW, Baseler M, Stevens R, et al. Longitudinal changes in CD4+ T cell antigen receptor diversity and naive/memory cell phenotype during 9 to 26 months of antiretroviral therapy of HIV-infected patients. *AIDS Res Hum Retroviruses* (2000) **16**(17):1877–86. doi:10.1089/08892220050195838

13. Kharbanda M, Mccloskey TW, Pahwa R, Sun M, Pahwa S. Alterations in T-cell receptor Vbeta repertoire of CD4 and CD8 T Lymphocytes in human immunodeficiency virus-infected children. *Clin Diagn Lab Immunol* (2003) **10**(1):53–8. doi:10.1128/CDLI.10.1.53-58.2003

14. Baum PD, Young JJ, Schmidt D, Zhang Q, Hoh R, Busch M, et al. Blood T cell receptor diversity decreases during the course of HIV infection but the potential for a diverse repertoire persists. *Blood* (2012) **119**:3469–77. doi:10.1182/blood-2011-11-395384

15. Conrad JA, Ramalingam RK, Duncan CB, Smith RM, Wei J, Barnett L, et al. Antiretroviral therapy reduces the magnitude and T cell receptor repertoire diversity of HIV-specific T cell responses without changing T cell clonotype dominance. *J Virol* (2012) **86**(8):4213–21. doi:10.1128/JVI.06000-11

16. Kostense S, Raaphorst FM, Notermans DW, Joling J, Hooibrink B, Pakker NG, et al. Diversity of the T-cell receptor BV repertoire in HIV-1-infected patients reflects the biphasic CD4+ T-cell repopulation kinetics during highly active antiretroviral therapy. *AIDS* (1998) **12**(18):F235–40. doi:10.1097/00002030-199818000-00001

17. Freeman JD, Warren RL, Webb JR, Nelson BH, Holt RA. Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing. *Genome Res* (2009) 19(10):1817–24. doi:10.1101/gr.092924.109

18. Robins HS, Campregher PV, Srivastava SK, Wacher A, Turtle CJ, Kahsai O, et al. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* (2009) 114(19):4099–107. doi:10.1182/blood-2009-04-217604

19. Thomas N, Heather J, Ndifon W, Shawe-Taylor J, Chain B. Decombinator: a tool for fast, efficient gene assignment in T-cell receptor sequences using a finite state machine. *Bioinformatics* (2013) 29(5):542–50. doi:10.1093/bioinformatics/btt004

20. Lefranc M-P, Pommié C, Ruiz M, Giudicelli V, Foulquier E, Truong L, et al. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev Comp Immunol* (2003) 27(1):55–77. doi:10.1016/S0145-305X(02)00039-3

21. Berger CT, Frahm N, Price DA, Mothe B, Ghebremichael M, Hartman KL, et al. High-functional-avidity cytotoxic T lymphocyte responses to HLA-B-restricted Gag-derived epitopes associated with relative HIV control. *J Virol* (2011) 85(18):9334–45. doi:10.1128/JVI.00460-11

22. Bourcier K, Lim D, Ding Y. Conserved CDR3 regions in T-cell receptor (TCR) CD8+ T cells that recognize the Tax11-19/HLA-A*0201 complex in a subject infected with human T-cell leukemia virus type 1: relationship of T-cell fine specificity and major histocompatibility complex/peptide. *J Virol* (2001) 75(20):9836–43. doi:10.1128/JVI.75.20.9836-9843.2001

23. Chen H, Ndhlovu ZM, Liu D, Porter LC, Fang JW, Darko S, et al. TCR clonotypes modulate the protective effect of HLA class I molecules in HIV-1 infection. *Nat Immunol* (2012) 13(7):691–700. doi:10.1038/ni.2342

24. Conrad JA, Ramalingam RK, Smith RM, Barnett L, Lorey SL, Wei J, et al. Dominant clonotypes within HIV-specific T cell responses are PD-1hi and CD127low reactivity and display reduced variant cross-reactivity. *J Immunol* (2012) 186(12):6871–85. doi:10.4049/jimmunol.1004234

25. Dong T, Stewart-Jones G, Chen N, Easterbrook P, Xu X, Papagno L, et al. HIV-specific cytotoxic T cells from long-term survivors select a unique T cell receptor. *J Exp Med* (2004) 200(12):1547–57. doi:10.1084/jem.20032044

26. Gillespie GMA, Stewart-Jones G, Rengasamy J, Beattie T, Bwayo JJ, Plummer FA, et al. Strong TCR conservation and altered T cell cross-reactivity characterize a B*57-restricted immune response in HIV-1 infection. *J Immunol* (2006) 177(6):3893–902. doi:10.4049/jimmunol.177.6.3893

27. Iglesias MC, Almeida JR, Fastenackels S, van Bockel DJ, Hashimoto M, Venturi V, et al. Escape from highly effective public CD8+ T-cell clonotypes by HIV. *Blood* (2011) 118(8):2138–49. doi:10.1182/blood-2011-01-328781

28. Kolowos W, Schmitt M, Herrman M, Harrer E, Löw P, Kalden JR, et al. Biased TCR repertoire in HIV-1-infected patients due to clonal expansion of HIV-1-reverse transcriptase-specific CTL clones. *J Immunol* (1999) 162(12):7525–33.

29. Kou ZC, Puhr JS, Rojas M, McCormack WT, Goodenow MM, Sleasman JW. T-cell receptor Vbeta repertoire CDR3 length diversity differs within CD45RA and CD45RO T-cell subsets in healthy and human immunodeficiency virus-infected children. *Clin Diagn Lab Immunol* (2000) 7(6):953–9. doi:10.1128/CDLI.7.6.953-959.2000

30. Lee SK, Xu Z, Lieberman J, Shankar P. The functional CD8 T cell response to HIV becomes type-specific in progressive disease. *J Clin Invest* (2002) 110(9):1339–47. doi:10.1172/JCI200216028

31. Lichterfeld M, Williams KL, Mui SK, Shah SS, Mothe BR, Sette A, et al. T cell receptor cross-recognition of an HIV-1 CD8+ T cell epitope presented by closely related alleles from the HLA-A3 superfamily. *Int Immunol* (2006) 18(7):1179–88. doi:10.1093/intimm/dxl052

32. Mendoza D, Royce C, Ruff LE, Ambrozak DR, Quigley MF, Dang T, et al. HLA B*5701-positive long-term nonprogressors/elite controllers are not distinguished from progressors by the clonal composition of HIV-specific CD8+ T cells. *J Virol* (2012) 86(7):4014–8. doi:10.1128/JVI.06982-11

33. Motozono C, Kuse N, Sun X, Rizkallah PJ, Fuller A, Oka S, et al. Molecular basis of a dominant T cell response to an HIV reverse transcriptase 8-mer epitope presented by the protective allele HLA-B*51:01. *J Immunol* (2014) 192(7):3428–34. doi:10.4049/jimmunol.1302667

34. Pantaleo G, Demarest JF, Soudeyns H, Graziosi C, Denis F, Adelsberger JW, et al. Major expansion of CD8+ T cells with a predominant Vbeta usage during the primary immune response to HIV. *Nature* (1994) 370:463–7. doi:10.1038/370463a0

35. Simons BC, VanCompernolle SE, Smith RM, Wei J, Barnett L, Lorey SL, et al. Despite biased TRBV gene usage against a dominant HLA B57-restricted epitope, TCR diversity can provide recognition of circulating epitope variants. *J Immunol* (2008) 181:5137–46. doi:10.4049/jimmunol.181.7.5137

36. Stewart-Jones GB, Simpson P, van der Merwe P, Easterbrook P, McMichael AJ, Rowland-Jones SL, et al. Structural features underlying T-cell receptor sensitivity to concealed MHC class I micropolymorphisms. *Proc Natl Acad Sci U S A* (2012) 109(50):E3483–92. doi:10.1073/pnas.1207896109

37. Sun X, Fujiwara M, Shi Y, Kuse N, Gatanaga H, Appay V, et al. Superimposed epitopes restricted by the same HLA molecule drive distinct HIV-specific CD8+ T cell repertoires. *J Immunol* (2014) 193(1):77–84. doi:10.4049/jimmunol.1400375

38. Yu XG, Lichterfeld M, Williams KL, Martinez-Picado J, Walker BD. Random T-cell receptor recruitment in human immunodeficiency virus type 1 (HIV-1)-specific CD8+ T cells from genetically identical twins infected with the same HIV-1 strain. *J Virol* (2007) 81(22):12666–9. doi:10.1128/JVI.01450-07

39. Greenaway HY, Ng B, Price DA, Douek DC, Davenport MP, Venturi V. NKT and MAIT invariant TCRα sequences can be produced efficiently by VJ gene recombination. *Immunobiology* (2013) 218(2):213–24. doi:10.1016/j.imbio.2012.04.003

40. Lepore M, Kalinichenko A, Colone A, Paleja B, Singhal A, Tschumi A, et al. Parallel T-cell cloning and deep sequencing of human MAIT cells reveal stable oligoclonal TCRβ repertoire. *Nat Commun* (2014) 5(May):3866. doi:10.1038/ncomms4866

41. Zvyagin IV, Pogorelyy MV, Ivanova ME, Komech EA, Shugay M, Bolotin DA, et al. Distinctive properties of identical twins' TCR repertoires revealed by high-throughput sequencing. *Proc Natl Acad Sci U S A* (2014) 111(16):5980–5. doi:10.1073/pnas.1319389111

42. Rudolph MG, Stanfield RL, Wilson IA. How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol* (2006) 24:419–66. doi:10.1146/annurev.immunol.23.021704.115658

43. Nikolich-Zugich J, Slifka MK, Messaoudi I. The many important facets of T-cell repertoire diversity. *Nat Rev Immunol* (2004) 4(2):123–32. doi:10.1038/nri1292

44. Miles JJ, Douek DC, Price DA. Bias in the $\alpha\beta$ T-cell repertoire: implications for disease pathogenesis and vaccination. *Immunol Cell Biol* (2011) 89(3):375–87. doi:10.1038/icb.2010.139

45. Ndifon W, Gal H, Shifrut E, Aharoni R, Yissachar N, Waysbort N, et al. Chromatin conformation governs T-cell receptor J gene segment usage. *Proc Natl Acad Sci U S A* (2012) 109(39):15865–70. doi:10.1073/pnas.1203916109

46. Warren RL, Freeman JD, Zeng T, Choe G, Munro S, Moore R, et al. Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. *Genome Res* (2011) 21(5):790–7. doi:10.1101/gr.115428.110

47. Angin M, Klarenbeek PL, King M, Sharma SM, Moodley ES, Rezai A, et al. Regulatory T cells expanded from HIV-1-infected individuals maintain phenotype, TCR repertoire and suppressive capacity. *PLoS One* (2014) 9(2):e86920. doi:10.1371/journal.pone.0086920

48. Elhanati Y, Murugan A, Callan CG, Mora T, Walczak AM. Quantifying selection in immune receptor repertoires. *Proc Natl Acad Sci U S A* (2014) 111(27):9875–80. doi:10.1073/pnas.1409572111

49. Li H, Ye C, Ji G, Han J. Determinants of public T cell responses. *Cell Res* (2012) 22(1):33–42. doi:10.1038/cr.2012.1

50. Madi A, Shifrut E, Reich-zeliger S, Gal H, Best K, Ndifon W, et al. T-cell receptor repertoires share a restricted set of public and abundant CDR3 sequences that are associated with self-related immunity. *Genome Res* (2014) 24:1603–12. doi:10.1101/gr.170753.113

51. Venturi V, Kedzierska K, Price DA, Doherty PC, Douek DC, Turner SJ, et al. Sharing of T cell receptors in antigen-specific responses is driven by convergent recombination. *Proc Natl Acad Sci U S A* (2006) 103(49):18691–6. doi:10.1073/pnas.0608907103

52. Macallan DC, Wallace D, Zhang Y, De Lara C, Worth AT, Ghattas H, et al. Rapid turnover of effector-memory CD4(+) T cells in healthy humans. *J Exp Med* (2004) 200(2):255–60. doi:10.1084/jem.20040341

53. Cosgrove C, Ussher JE, Rauch A, Gartner K, Kurioka A, Huhn MH, et al. Early and nonreversible decrease of CD161++/MAIT cells in HIV infection. *Blood* (2013) 121(6):951–61. doi:10.1182/blood-2012-06-436436

54. Threlkeld SC, Wentworth PA, Kalams SA, Wilkes BM, Ruhl DJ, Keogh E, et al. Degenerate and promiscuous recognition by CTL of peptides presented by the MHC class I A3-like superfamily. *J Immunol* (1997) **159**(4):1648–57.

55. Ueno T, Tomiyama H, Takiguchi M. Single T cell receptor-mediated recognition of an identical HIV-derived peptide presented by multiple HLA class I molecules. *J Immunol* (2002) **169**(9):4961–9. doi:10.4049/jimmunol.169.9.4961

56. Venturi V, Chin HY, Asher TE, Ladell K, Scheinberg P, Bornstein E, et al. TCR beta-chain sharing in human CD8+ T cell responses to cytomegalovirus and EBV. *J Immunol* (2008) **181**(11):7853–62. doi:10.4049/jimmunol.181.11.7853

57. Almeida JR, Price DA, Papagno L, Arkoub ZA, Sauce D, Bornstein E, et al. Superior control of HIV-1 replication by CD8+ T cells is reflected by their avidity, polyfunctionality, and clonal turnover. *J Exp Med* (2007) **204**(10):2473–85. doi:10.1084/jem.20070784

58. Khan N, Shariff N, Cobbold M, Bruton R, Ainsworth JA, Sinclair AJ, et al. Cytomegalovirus seropositivity drives the CD8 T cell repertoire toward greater clonality in healthy elderly individuals. *J Immunol* (2002) **169**(4):1984–92. doi:10.4049/jimmunol.169.4.1984

59. Klarenbeek PL, Remmerswaal EBM, Ten Berge IJM, Doorenspleet ME, van Schaik BDC, Esveldt RE, et al. Deep sequencing of antiviral T-cell responses to HCMV and EBV in humans reveals a stable repertoire that is maintained for many years. *PLoS Pathog* (2012) **8**(9):e1002889. doi:10.1371/journal.ppat.1002889

60. Ouyang Q, Wagner WM, Wikby A, Walter S, Aubert G, Dodi AI, et al. Large numbers of dysfunctional CD8+ T lymphocytes bearing receptors for a single dominant CMV epitope in the very old. *J Clin Immunol* (2003) **23**(4):247–57. doi:10.1023/A:1024580531705

61. Spreafico R, Rossetti M, van Loosdregt J, Wallace CA, Massa M, Magni-Manzoni S, et al. A circulating reservoir of pathogenic-like CD4+ T cells shares a genetic and phenotypic signature with the inflamed synovial microenvironment. *Ann Rheum Dis* (2014) 1–7. doi:10.1136/annrheumdis-2014-206226

62. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* (1996) **271**:1582–6. doi:10.1126/science.271.5255.1582

63. Betts MR, Ambrozak DR, Douek DC, Bonhoeffer S, Brenchley JM, Casazza JP, et al. Analysis of total human immunodeficiency virus (HIV)-specific CD4+ and CD8+ T-cell responses: relationship to viral load in untreated HIV infection. *J Virol* (2001) **75**(24):11983–91. doi:10.1128/JVI.75.24.11983-11991.2001

64. Murugan A, Mora T, Walczak AM, Callan CG. Statistical inference of the generation probability of T-cell receptors from sequence repertoires. *Proc Natl Acad Sci U S A* (2012) **109**:16161–6. doi:10.1073/pnas.1212755109

65. Leeansyah E, Ganesh A, Quigley F, So A, Andersson J, Hunt PW, et al. Activation, exhaustion, and persistent decline of the antimicrobial MR1-restricted MAIT-cell population in chronic HIV-1 infection. *Blood* (2013) **121**(7):1124–35. doi:10.1182/blood-2012-07-445429

66. Brenchley JM, Price DA, Schacker TW, Asher TE, Silvestri G, Rao S, et al. Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat Med* (2006) **12**(12):1365–71. doi:10.1038/nm1511

67. Mehandru S, Poles MA, Tenner-Racz K, Jean-Pierre P, Manuelli V, Lopez P, et al. Lack of mucosal immune reconstitution during prolonged treatment of acute and early HIV-1 infection. *PLoS Med* (2006) **3**(12):2335–48. doi:10.1371/journal.pmed.0030484

68. Vantourout P, Hayday A. Six-of-the-best: unique contributions of $\gamma\delta$ T cells to immunology. *Nat Rev Immunol* (2013) **13**(2):88–100. doi:10.1038/nri3384

69. Kaufmann GR, Perrin L, Pantaleo G, Opravil M, Furrer H, Telenti A, et al. CD4 T-lymphocyte recovery in individuals with advanced HIV-1 infection receiving potent antiretroviral therapy for 4 years: the Swiss HIV Cohort Study. *Arch Intern Med* (2003) **163**:2187–95. doi:10.1001/archinte.163.18.2187

70. Khanna N, Opravil M, Furrer H, Cavassini M, Vernazza P, Bernasconi E, et al. CD4+ T cell count recovery in HIV type 1-infected patients is independent of class of antiretroviral therapy. *Clin Infect Dis* (2008) **47**(8):1093–101. doi:10.1086/592113

71. Emerson R, Sherwood A, Desmarais C, Malhotra S, Phippard D, Robins H. Estimating the ratio of CD4+ to CD8+ T cells using high-throughput sequence data. *J Immunol Methods* (2013) **391**(1–2):14–21. doi:10.1016/j.jim.2013.02.002

72. Han A, Glanville J, Hansmann L, Davis MM. Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat Biotechnol* (2014) **32**(7):684–92. doi:10.1038/nbt.2938

73. Turchaninova MA, Britanova OV, Bolotin DA, Shugay M, Putintseva EV, Staroverov DB, et al. Pairing of T-cell receptor chains via emulsion PCR. *Eur J Immunol* (2013) **43**(9):2507–15. doi:10.1002/eji.201343453

74. Embleton MJ, Gorochov G, Jones PT, Winter G. In-cell PCR from mRNA: amplifying and linking the rearranged immunoglobulin heavy and light chain V-genes within single cells. *Nucleic Acids Res* (1992) **20**(15):3831–7. doi:10.1093/nar/20.15.3831

75. Howie B, Sherwood AM, Berkebile AD, Berka J, Emerson RO, Williamson DW, et al. High-throughput pairing of T cell receptor alpha and beta sequences. *Sci Transl Med* (2015) **7**(301):ra131–301. doi:10.1126/scitranslmed.aac5624