

# Immunopathogenic CSF TCR repertoire signatures in virus-associated neurologic disease

Satoshi Nozuma,<sup>1</sup> Yoshimi Enose-Akahata,<sup>1</sup> Kory R. Johnson,<sup>2</sup> Maria Chiara Monaco,<sup>1</sup> Nyater Ngouth,<sup>1</sup> Abdel Elkahloun,<sup>3</sup> Joan Ohayon,<sup>4</sup> Jun Zhu,<sup>5</sup> and Steven Jacobson<sup>1</sup>

<sup>1</sup>Viral Immunology Section, Neuroimmunology Branch and <sup>2</sup>Bioinformatics Section, National Institute of Neurological Disorders and Stroke, NIH, Bethesda, Maryland, USA. <sup>3</sup>Comparative Genomics and Cancer Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, Maryland, USA. <sup>4</sup>Neuroimmunology Clinic, National Institute of Neurological Disorders and Stroke, NIH, Bethesda, Maryland, USA. <sup>5</sup>Mokobio Biotechnology R&D Center, Rockville, Maryland, USA.

In this study, we examined and characterized disease-specific TCR signatures in cerebrospinal fluid (CSF) of patients with HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP). TCR  $\beta$  libraries using unique molecular identifier-based methodologies were sequenced in paired peripheral blood mononuclear cells (PBMCs) and CSF cells from HAM/TSP patients and normal healthy donors (NDs). The sequence analysis demonstrated that TCR  $\beta$  repertoires in CSF of HAM/TSP patients were highly expanded and contained both TCR clonotypes shared with PBMCs and uniquely enriched within the CSF. In addition, we analyzed TCR  $\beta$  repertoires of highly expanded and potentially immunopathogenic HTLV-1 Tax11-19-specific CD8<sup>+</sup> T cells from PBMCs of HLA-A\*0201<sup>+</sup> HAM/TSP and identified a conserved motif (PGLAG) in the CDR3 region. Importantly, TCR  $\beta$  clonotypes of expanded clones in HTLV-1 Tax11-19-specific CD8<sup>+</sup> T cells were also expanded and enriched in the CSF of the same patient. These results suggest that exploring TCR repertoires of CSF and antigen-specific T cells may provide a TCR repertoire signature in virus-associated neurologic disorders.

## Introduction

Human T-lymphotropic type 1 (HTLV-1) infects 10 million–20 million people worldwide and is the etiologic agent of adult T cell leukemia/lymphoma (1) and an inflammatory neurologic disease, HTLV-1-associated myelopathy/tropical spastic paraparesis (HAM/TSP) (2). HAM/TSP is a neuroinflammatory disease of the spinal cord with progressive spastic paraparesis and bladder dysfunction. HTLV-1 predominantly infects CD4<sup>+</sup> T cells and causes the activation and proliferation of infected cells, which express viral proteins that may lead to the activation and expansion of HTLV-1-specific CD8<sup>+</sup> cytotoxic T lymphocytes (CTLs) (3, 4). The HTLV-1 regulatory protein Tax promotes the proliferation of HTLV-1-infected lymphocytes and is an immunodominant antigen recognized by HTLV-1-specific CTLs (5). The frequency of these Tax-specific CTLs is extraordinarily high in peripheral blood and even higher in cerebrospinal fluid (CSF) (6), in which virus-specific CTLs demonstrate degranulation activity and produce proinflammatory cytokines (3, 4). It has been suggested that HTLV-1-infected CD4<sup>+</sup> T cells and HTLV-1-specific CD8<sup>+</sup> cytotoxic T cells enter the CNS and release various proinflammatory cytokines, resulting in the neural damage and degeneration in HAM/TSP patients (7).

T cell-mediated antigen recognition depends on the interaction of the T cell receptor (TCR) with antigen-major histocompatibility complex (antigen-MHC) molecules (8). Although the actual diversity of human TCR repertoires remains unknown, recent estimates of the potential number of TCR repertoires have been reported to be highly diverse in the range from  $1 \times 10^{13}$  to  $1 \times 10^{20}$  in humans (9–11). TCR diversity is generated through random rearrangement of variable regions (V), diversity (D) of the  $\beta$  chain, and joining (J) segments of TCR genes (VDJ rearrangements) and is a central component of the adaptive immune system (10). T cells become activated and clonally expanded after they encounter their cognate antigen (12). Importantly, an individual's TCR repertoire can be altered in the context of

**Conflict of interest:** The authors have declared that no conflict of interest exists.

**Copyright:** © 2021, Nozuma et al. This is an open access article published under the terms of the Creative Commons Attribution 4.0 International License.

**Submitted:** October 2, 2020

**Accepted:** January 13, 2021

**Published:** February 22, 2021

**Reference information:** *JCI Insight*. 2021;6(4):e144869.  
<https://doi.org/10.1172/jci.insight.144869>.

infections, malignancies, or immunological disorders. Analysis of the TCR repertoire can, thus, provide a better understanding of immune-mediated responses in neuroinflammatory diseases. Recent technological advances based on high-throughput sequencing (HTS), bioinformatics software, and unbiased gene amplification allow for the analysis of millions of TCR sequences (13). We have recently reported on an unbiased molecular technique that combines the use of unique molecular identifier (UMI), which is barcoded cDNA TCR transcript, with a previously published UMI-based TCR  $\beta$  data sequence analysis pipeline (14, 15). This robust UMI-based approach has been used successfully to assess the clonal expansion and diversity of the TCR repertoire in peripheral blood of HAM/TSP, multiple sclerosis (MS), and normal healthy donors (NDs), which demonstrated higher clonal T cell expansions in HAM/TSP compared with MS and NDs (14). Disease-associated TCR signatures could, thus, be useful biomarkers for immune-mediated disease diagnosis, prognosis, and response to treatment.

While we have demonstrated higher TCR clonal expansions in the peripheral blood of HAM/TSP compared with controls, no disease-specific TCR signatures have been identified (14). This could be due, in part, to observations that, in patients with neuroinflammatory neurologic disease, immune responses in the periphery may not reflect events in the CNS (16). In the CNS, robust immune responses to various viral infections and autoantigens have been demonstrated to be unique from the immune response observed in peripheral tissues. For example, comparison of TCR repertoires between blood and CSF in patients with MS showed an overlap of expanded clones, but with an intrathecal enrichment of Epstein-Barr virus-reactive (EBV-reactive) CD8<sup>+</sup> T cells (17, 18), suggesting a role for EBV in the pathogenesis of MS. While the CSF is not the CNS, it is in close proximity to the brain, accessible by lumbar puncture and considered to reflect inflammation within the CNS better than in peripheral blood (19). Therefore, comparing TCR repertoires between the periphery and the CSF might be useful for exploring local TCR signatures in viral-mediated neurologic disease such as HAM/TSP. In addition, previous reports on the TCR repertoire analysis using PCR-based techniques showed higher clonal expansions, especially in CD8<sup>+</sup> T cell subsets of HAM/TSP patients by complementarity determining region 3 (CDR3) spectratyping (20, 21) and some conserved motifs in the CDR3 region of Tax-specific CTL clones generated from HAM/TSP patients with HLA-A\*0201 by cloning and sequencing (22, 23). Since chronically activated HTLV-1-specific CTLs are thought to have key roles in the pathogenesis of HAM/TSP (7), an unbiased, comprehensive analysis based on HTS to investigate TCR repertoire signatures of HTLV-1-specific CTLs would be useful to determine if such sequences can be demonstrated in lymphocytes present in peripheral blood and CSF of HAM/TSP patients.

In this study, we examined TCR  $\beta$  repertoires in paired peripheral blood mononuclear cells (PBMCs) and lymphocytes present in CSF from patients with HAM/TSP compared with NDs. Highly clonally expanded TCR clones were detected in CSF of HAM/TSP patients compared with the CSF of NDs. In addition, we analyzed TCR  $\beta$  repertoires of HTLV-1 Tax11-19-specific CD8<sup>+</sup> T cells from PBMCs of HLA-A\*0201<sup>+</sup> HAM/TSP patients and identified that the expanded TCR clones in HTLV-1 Tax11-19-specific CD8<sup>+</sup> T cells were also expanded and enriched in the CSF of HAM/TSP.

## Results

*T cell clones are highly expanded both in blood and CSF of HAM/TSP patients compared with NDs.* A total of 20 subjects consisting of 15 HAM/TSP patients and 5 NDs were examined for sequencing of TCR  $\beta$  chain repertoires in the study. The demographic characteristics of the study population are summarized in Table 1. To study whether the TCR repertoire skewing is present in peripheral blood or CSF of HAM/TSP patients, we compared the TCR repertoire of PBMCs from 9 HAM/TSP patients and 5 NDs in which paired CSF were also obtained. In the 14 subjects, mean age of NDs and HAM/TSP patients was 50.8 and 55.6 years, respectively (no statistically significant differences). CSF cell concentration was increased in the CSF of HAM/TSP patients compared with NDs, consistent with the known CSF pleocytosis in HAM/TSP (24) (Table 1). In HAM/TSP patients, the median HTLV-1 proviral load (PVL) was 10.20% (range, 4.90%–47.17%) in PBMCs and, as reported (25), was significantly higher in CSF (36.27% [range, 19.58%–53.33%];  $P = 0.0039$ ) (Table 1).

The TCR  $\beta$  sequence data, including the numbers of total HTS sequences, total UMIs, total unique TCR  $\beta$  clonotypes, and Shannon diversity from HAM/TSP patients and NDs are summarized in Table 2. Total numbers of sequence reads generated from HTS were not statistically different between HAM/TSP patients and NDs in either PBMCs or CSF (Supplemental Figure 1A; supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.144869DS1>). On average, 631,318 and 360,179 reads

**Table 1. Demographics, clinical, and biological characteristics of HAM/TSP patients and NDs**

Subject	Sex	Age (years)	Disease duration (years) <sup>a</sup>	Race	HLA-A type	Proviral load, %		CSF cell concentration (cells/ $\mu$ L)	Tax-specific CD8 <sup>+</sup> T cells frequency, % <sup>b</sup>
						PBMCs	CSF		
HAM-1	F	50	6	African American	29,74	22.73	36.27	38	-
HAM-2	F	60	2	African American	02,30	7.45	19.58	19.1	-
HAM-3	F	44	2	African American	01,68	6.23	26.97	22	-
HAM-4	F	62	10	Caucasian	32	10.20	36.79	9.3	-
HAM-5	F	58	29	African American	03,33	47.17	48.18	11.7	-
HAM-6	M	74	1	Caucasian	0201	4.90	21.90	32.5	10.60
HAM-7	F	50	2	African American	23,33	21.50	53.33	17	-
HAM-8	F	57	20	Caucasian	11, 31	20.44	44.41	4.4	-
HAM-9	F	45	5	African American	02,34	9.22	34.09	9.1	-
HAM-10	M	69	12	Caucasian	0201	10.90	NA	NA	12.70
HAM-11	F	64	8	African American	0201	25.00	NA	NA	2.18
HAM-12	M	78	33	Caucasian	0201	16.73	NA	NA	4.40
HAM-13	M	63	16	African American	0201	14.02	NA	NA	1.21
HAM-14	M	45	2	Caucasian	0201	16.50	NA	NA	2.16
HAM-15	F	56	28	African American	0201	12.70	NA	NA	2.53
ND-1	F	62	-	African American	11, 33	-	-	3.0	-
ND-2	F	56	-	Caucasian	NA	-	-	4.5	-
ND-3	M	55	-	Caucasian	02, 69	-	-	1.6	-
ND-4	M	50	-	African American	23	-	-	0.8	-
ND-5	M	31	-	Caucasian	NA	-	-	1.3	-

<sup>a</sup>The disease duration is calculated from the time point of the first disease symptom. <sup>b</sup>The Tax11-19-specific CD8<sup>+</sup> T cell frequency is expressed as a percentage of HTLV-1 Tax11-19/HLA-A\*0201 tetramer-specific CD8<sup>+</sup> T cells in the total amount of CD8<sup>+</sup> T cells. HAM, HTLV-1-associated myelopathy/tropical spastic paraparesis; ND, normal healthy donor; PBMCs, peripheral blood mononuclear cells; CSF, cerebrospinal fluid; NA, not available.

were sequenced from the PBMCs and the CSF, respectively (Supplemental Figure 1A and Table 2). Comparison of the number of total unique TCR  $\beta$  clonotypes in PBMCs showed no significant differences between HAM/TSP patients and NDs, while in CSF, unique TCR  $\beta$  clonotypes were significantly higher in HAM/TSP patients compared with NDs (Supplemental Figure 1B and Table 2). TCR clonal size was classified into 3 groups based on the number of UMIs: (a) singletons (1 UMI); (b) clones with  $2 \leq$  UMIs  $< 8$ , and (c) expanded clones with  $\geq 8$  UMIs as previously described (14) (See Methods section for classification criteria).

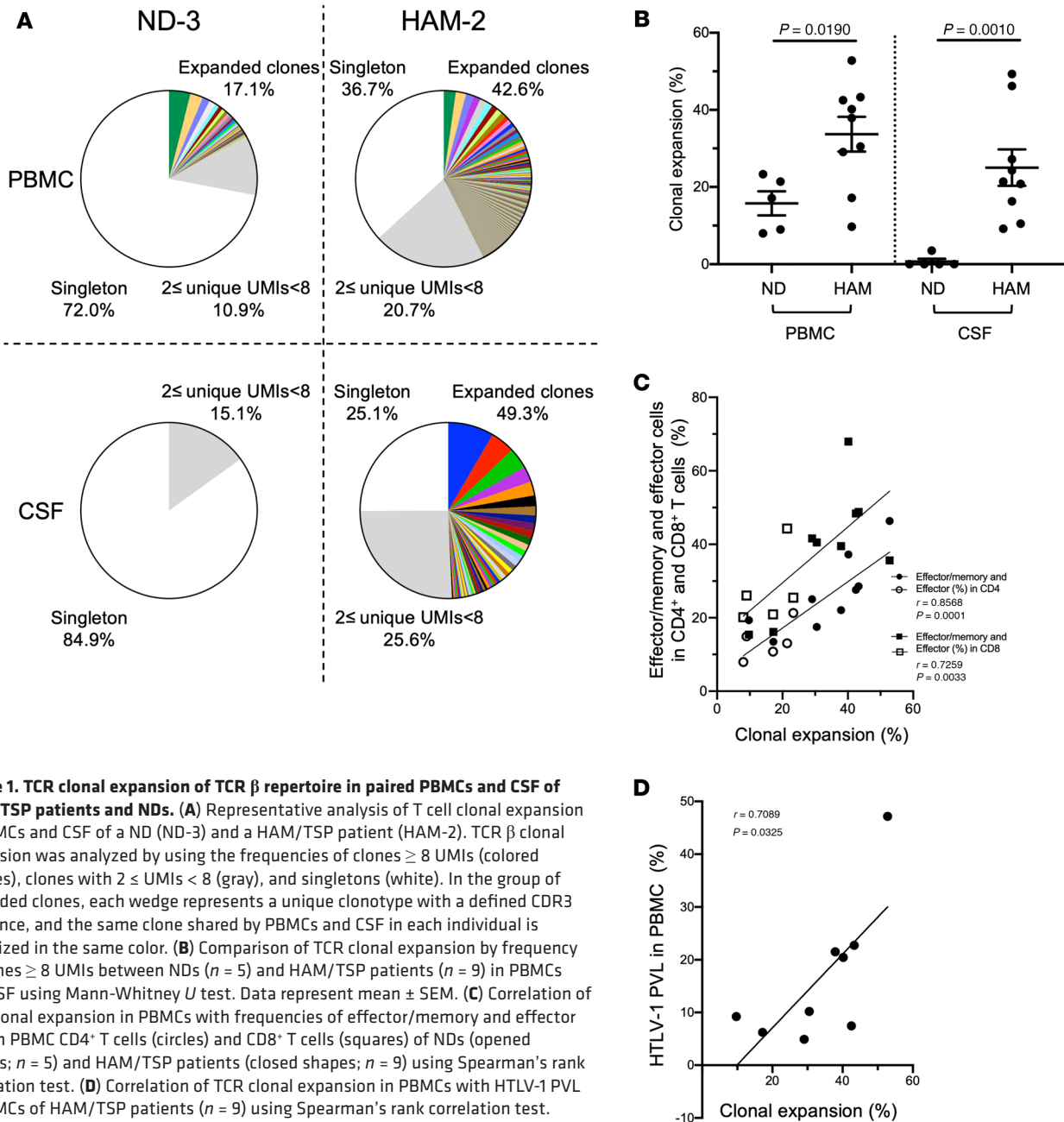
The frequencies of expanded TCR clones from a representative HAM/TSP patient (HAM-2) and a ND (ND-3) are shown in Figure 1A. In this pie chart based on our definition of TCR clonal size, expanded clones with  $\geq 8$  UMIs are shown as colored wedges, while clones with  $2 \leq$  UMIs  $< 8$  and singletons are visualized in the gray area and in the white area, respectively. Each colored wedge represents a unique clonotype with a defined CDR3 sequence. The same clone shared by PBMCs and CSF in each individual is visualized by the same color. Individual variations of clonal expansion were observed in PBMCs for both the ND and the HAM/TSP patient. By contrast in CSF, individual variations of clonal expansion were detected in the HAM/TSP patient but not in the ND (Figure 1A). Group analysis demonstrated that all HAM/TSP patients showed individual variations of clonal expansion in the CSF, while ND did not, except for the detection of only 1 expanded clone in 1 ND (ND-2; Supplemental Figure 2). These results are quantitated in Figure 1B, where each dot represents the frequency of expanded clones in PBMCs and CSF of HAM/TSP patients and NDs. HAM/TSP patients showed a higher degree of TCR clonal expansion in both PBMCs ( $P = 0.0190$ ) and CSF ( $P = 0.0010$ ), compared with NDs (Figure 1B). When the TCR clone diversity was estimated using the Shannon diversity index, a significantly lower diversity of TCR clones was observed in the peripheral blood of HAM/TSP patients compared with NDs ( $P = 0.029$ ; Supplemental Figure 1C). In CSF, there was no significant difference in TCR diversity between HAM/TSP patients and NDs (Supplemental Figure 1C).

Since HAM/TSP patients have been reported to demonstrate high frequencies of effector/memory and effector CD8<sup>+</sup> T cells (26), which suggests that they are driven by chronic antigen (viral) stimulation, we asked if this differentiation of T cells may also be associated with TCR repertoire diversity, clonal

Table 2. Summary of TCR  $\beta$  sequence data from HAM/TSP patients and NDs

Subject	Cell type	No. of total HTS sequences	No. of total UMIs	No. of unique TCR $\beta$ clonotypes	% Clonal expansion	Shannon diversity <sup>A</sup>	Shannon diversity <sup>B</sup>
HAM-1	PBMC	265,192	73,990	30,970	43.3%	5043.9	-
HAM-2	PBMC	260,965	96,946	44,155	42.5%	3403.7	-
HAM-3	PBMC	212,457	71,626	48,660	17.2%	10,295.3	-
HAM-4	PBMC	140,631	36,502	18,570	30.5%	5065.9	-
HAM-5	PBMC	63,340	31,398	10,045	52.8%	1983.6	-
HAM-6	PBMC	580,798	24,609	13,227	29.0%	4116.3	212.3
HAM-7	PBMC	1,011,365	49,740	23,579	37.9%	4121.2	-
HAM-8	PBMC	1,284,110	60,361	28,287	40.2%	5672.7	-
HAM-9	PBMC	1,053,855	31,452	24,378	9.7%	10701.2	-
HAM-10	PBMC	493,176	55,822	20,317	49.2%	-	206.7
HAM-11	PBMC	450,157	61,941	17,694	64.1%	-	175.8
HAM-12	PBMC	642,358	38,215	15,876	45.7%	-	138.9
HAM-13	PBMC	688,024	30,744	16,718	31.8%	-	206.2
HAM-14	PBMC	1,218,617	62,233	30,185	35.6%	-	269.1
HAM-15	PBMC	873,806	29,005	17,707	26.9%	-	274.6
ND-1	PBMC	710,161	40,924	25,755	23.3%	5894.8	-
ND-2	PBMC	938,524	54,162	35,970	21.4%	6945.5	-
ND-3	PBMC	479,442	16,331	12,555	17.1%	6244.4	-
ND-4	PBMC	103,0356	43,370	33,496	8.0%	11890.5	-
ND-5	PBMC	913,055	39,205	30,746	9.0%	10,979.3	-
HAM-1	CSF	173,303	7541	2854	46.2%	41.7	-
HAM-2	CSF	153,587	2989	1041	49.3%	32.4	-
HAM-3	CSF	102,783	1924	1069	24.3%	45.8	-
HAM-4	CSF	148,767	2524	1332	27.2%	36.4	-
HAM-5	CSF	168,806	1020	665	16.3%	40.4	-
HAM-6	CSF	475,004	561	320	21.4%	39.3	-
HAM-7	CSF	445,394	190	150	10.5%	38.5	-
HAM-8	CSF	438833	654	406	20.8%	37.0	-
HAM-9	CSF	328,986	337	255	9.2%	40.5	-
ND-1	CSF	386,513	147	115	0.0%	42.7	-
ND-2	CSF	416,505	255	192	3.5%	41.9	-
ND-3	CSF	432,986	332	303	0.0%	47.1	-
ND-4	CSF	655,202	65	49	0.0%	37.6	-
ND-5	CSF	610,040	48	39	0.0%	36.6	-
HAM-6	CD8 <sup>+</sup>	542,034	21,316	2202	84.7%	-	46.7
HAM-10	CD8 <sup>+</sup>	598,499	109,300	18,670	76.1%	-	108.2
HAM-11	CD8 <sup>+</sup>	415,375	85,443	10,745	84.7%	-	49.2
HAM-12	CD8 <sup>+</sup>	678,966	44,890	3552	89.1%	-	25.7
HAM-13	CD8 <sup>+</sup>	498,223	26,721	4030	81.1%	-	49.4
HAM-14	CD8 <sup>+</sup>	899,146	56,552	18,737	61.4%	-	184.0
HAM-15	CD8 <sup>+</sup>	745,549	69,655	35,989	40.6%	-	210.8
HAM-6	Tax <sup>+</sup>	366,076	541	39	90.0%	-	6.4
HAM-10	Tax <sup>+</sup>	249,796	31,234	2066	91.4%	-	38.5
HAM-11	Tax <sup>+</sup>	385,450	1313	36	96.6%	-	4.7
HAM-12	Tax <sup>+</sup>	597,171	9130	332	93.9%	-	44.2
HAM-13	Tax <sup>+</sup>	319,807	316	47	78.5%	-	21.3
HAM-14	Tax <sup>+</sup>	369,392	718	87	79.0%	-	13.4
HAM-15	Tax <sup>+</sup>	542,443	473	33	88.6%	-	6.4

<sup>A</sup>Shannon diversity is estimated by normalizing the number of unique T cell clonotypes via resampling to the size of the smallest data set of PBMCs or CSF cells. <sup>B</sup>Shannon diversity is estimated by normalizing the number of unique T cell clonotypes via resampling to the size of the smallest data set of sorted Tax11-19-specific CD8<sup>+</sup> T cells. HAM, HTLV-1-associated myelopathy/tropical spastic paraparesis; ND, normal healthy donor; HTS, high-throughput sequencing; UMIs, unique molecular identifiers; PBMC, peripheral blood mononuclear cell; CSF, cerebrospinal fluid; CD8<sup>+</sup>, enriched CD8<sup>+</sup> T cells; Tax<sup>+</sup>, sorted Tax-specific CD8<sup>+</sup> T cells.



**Figure 1. TCR clonal expansion of TCR  $\beta$  repertoire in paired PBMCs and CSF of HAM/TSP patients and NDs.** (A) Representative analysis of T cell clonal expansion in PBMCs and CSF of a ND (ND-3) and a HAM/TSP patient (HAM-2). TCR  $\beta$  clonal expansion was analyzed by using the frequencies of clones  $\geq 8$  UMIs (colored wedges), clones with  $2 \leq$  UMIs  $< 8$  (gray), and singletons (white). In the group of expanded clones, each wedge represents a unique clonotype with a defined CDR3 sequence, and the same clone shared by PBMCs and CSF in each individual is visualized in the same color. (B) Comparison of TCR clonal expansion by frequency of clones  $\geq 8$  UMIs between NDs ( $n = 5$ ) and HAM/TSP patients ( $n = 9$ ) in PBMCs and CSF using Mann-Whitney  $U$  test. Data represent mean  $\pm$  SEM. (C) Correlation of TCR clonal expansion in PBMCs with frequencies of effector/memory and effector cells in PBMC CD4<sup>+</sup> T cells (circles) and CD8<sup>+</sup> T cells (squares) of NDs (opened shapes;  $n = 5$ ) and HAM/TSP patients (closed shapes;  $n = 9$ ) using Spearman's rank correlation test. (D) Correlation of TCR clonal expansion in PBMCs with HTLV-1 PVL in PBMCs of HAM/TSP patients ( $n = 9$ ) using Spearman's rank correlation test.

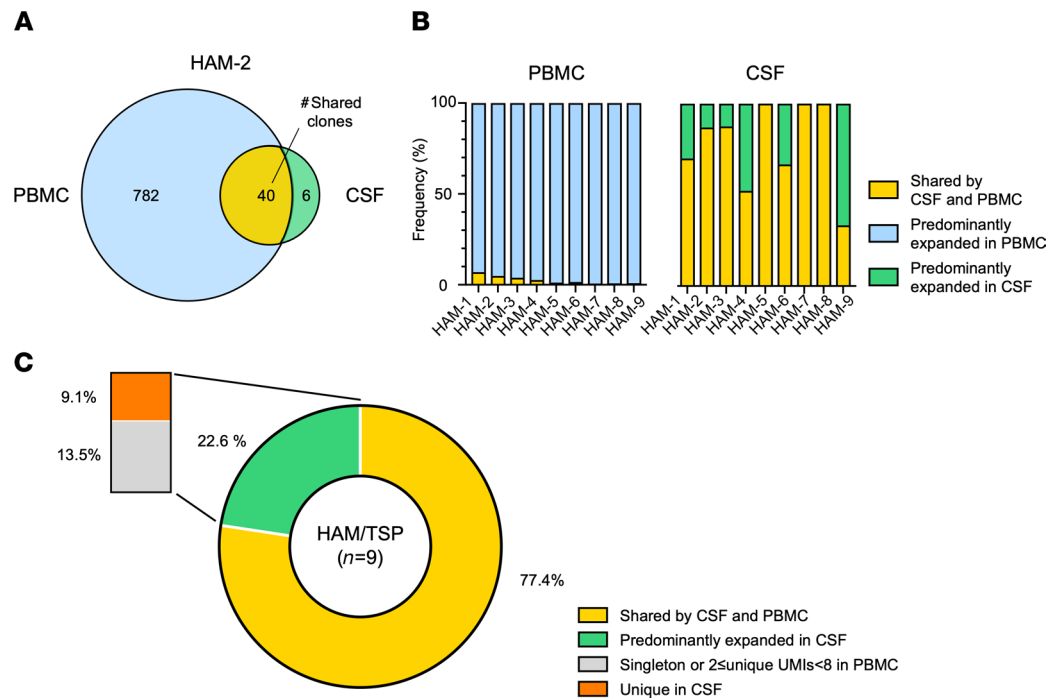
expansion, and HTLV-1 PVL. As shown in Figure 1C, the frequency of clonal expansion in PBMCs significantly correlated with T cell phenotype of effector/memory and effector cells both in CD4<sup>+</sup> and CD8<sup>+</sup> T cells. In addition, the frequency of TCR clonal expansion in PBMCs also showed significant correlations with HTLV-1 PVL in PBMCs (Figure 1D), but it did not show any correlation with disease duration or severity (data not shown). Collectively, these results demonstrate higher clonally expanded T cell clones in both PBMCs and CSF of HAM/TSP patients compared with NDs, with the degree of TCR clonal expansion particularly in the peripheral blood correlating with differentiated T cell phenotypes and HTLV-1 PVL.

*Overlap of TCR  $\beta$  repertoires of expanded clones between blood and CSF of HAM/TSP patients.* To evaluate whether there are any expanded TCR  $\beta$  clonotypes shared between PBMCs and CSF within each HAM/TSP patient, we compared the CDR3 sequences in the expanded clones (defined as unique clones with  $\geq 8$  UMIs) between PBMCs and CSF of HAM/TSP patients in order to determine if any of the CDR3 sequences were identical between PBMCs and CSF. A representative analysis of the overlap of the expanded TCR  $\beta$  clonotypes between paired PBMCs and CSF of a HAM/TSP patient (HAM-2) is shown in Figure 2A. Of the 822

expanded clonotypes in the PBMCs of HAM-2, only 40 clonotypes (4.9%) were shared with those clonotypes expanded in the CSF (Figure 2A; yellow area). Of the 46 TCR  $\beta$  expanded clonotypes in the CSF of HAM-2, 87.0% (40 of 46) were also found in the expanded clonotypes in the PBMCs (Figure 2A). Six expanded clonotypes were found to be unique in the CSF (Figure 2A; green area) — i.e., no sequences were found in the 822 expanded clonotypes in the PBMCs. Figure 2B shows the frequencies of the shared and unique TCR  $\beta$  expanded clonotypes detected in PBMCs (Figure 2B, left) and CSF (Figure 2B, right) in 9 HAM/TSP patients. In PBMCs of HAM/TSP patients, the expanded clonotypes shared with the CSF were detected at low frequencies (0.31%–6.99%; Figure 2B, left, bars in yellow), while almost all the expanded clonotypes were predominantly detected in the PBMCs (Figure 2B, left, bars in light blue). In the CSF, the expanded clonotypes shared with the PBMCs were detected at higher frequencies (33.33%–100.00%; Figure 2B, right, bars in yellow). Interestingly, of the 9 HAM/TSP patients, TCR  $\beta$  clonotypes predominantly expanded in the CSF were found in 6 patients (HAM-1, -2, -3, -4, -6, and -9; Figure 2B, right, green bars). Overall, 77.4% of expanded clonotypes in the CSF were shared by the expanded clonotypes ( $\geq 8$  UMIs) in PBMCs (Figure 2C, yellow area), and 22.6% were predominantly expanded in CSF of HAM/TSP patients (Figure 2C, green area). A subset of these CSF expanded clonotypes (13.5%) may also have been expanded in PBMCs but did not reach our formal definition of expansion in PBMCs, since they had singleton or  $2 \leq$  UMIs  $< 8$  (Figure 2C, gray area). Importantly, of the TCR  $\beta$  clonotypes predominantly expanded in the CSF, 9.1% were uniquely expanded in this compartment (Figure 2C, orange area). These results demonstrate that only a small subset of TCR  $\beta$  clonotypes in PBMCs was shared in the CSF, suggesting that most expanded CSF clonotypes were derived from expanded clonotypes in PBMCs in which a small but distinct fraction of these TCR-expanded clonotypes were intrathecally enriched.

*CDR3 relatedness of expanded T cell clonotypes in CSF of HAM/TSP patients.* Since disease-specific TCR clonotypes were not previously identified in PBMCs of HAM/TSP patients (14), we investigated whether there are any specific clonotypes shared across CSF compartments. We first compared the profile of TCR repertoires using identical CDR3 amino acid sequences in paired PBMCs and CSF across HAM/TSP patients. A heatmap analysis shows the frequency of the expanded clonotypes ( $\geq 8$  UMIs) having identical CDR3 amino acid sequences shared across HAM/TSP patients in both PBMCs and CSF (Figure 3A). While shared TCR clonotypes between CSF and PBMCs were observed within each patient, there were no shared identical TCR  $\beta$  clonotypes across both CSF and PBMCs of patients with HAM/TSP (Figure 3A). These results indicate that TCR repertoire signatures can be identified that are unique to individual patients in CSF as well as PBMCs.

Since TCR clonal repertoires in CSF were not shared across patients, we next analyzed the TCR repertoire relatedness in the CSF of the 9 HAM/TSP patients using phylogenetic tree cluster analysis with respect to CDR3 amino acid sequence similarities. Figure 3B shows the phylogenetic tree analysis with 2 layers using expanded TCR  $\beta$  clonotypes with  $\geq 8$  UMIs (total of 240 clones) in CSF of all HAM/TSP patients. Sequence data for generating the tree are provided in Supplemental Table 1. The phylogenetic tree contains 2 layers to identify HAM/TSP patients (inner layer) and expanded clone groups in CSF of HAM/TSP patients (outer layer). This approach performed an all-versus-all pairwise comparison of clonotypes detected for this group of patients or a group of expanded clones in CSF and computed a set of clonotype frequencies, the diversity of overlap, and the geometric mean for the overlap. These metrics were then used to create hierarchical clustering from samples, producing a sample-level and expanded clone group-level dendrogram (Figure 3B). Upon visualization of this dendrogram, branch lengths described the distance between repertoires (longer branch length indicates greater dissimilarity per repertoire relatedness by CDR3 amino acid sequence) and node size increase in accordance with read counts. The phylogenetic tree demonstrated that a total of 39 clusters that shared common CDR3 amino acid motifs by sequence similarity were identified in the CSF of these 9 HAM/TSP patients (Figure 3B), suggesting the existence of a TCR  $\beta$  repertoire-level signature in HAM/TSP CSF. The expanded TCR  $\beta$  clonotypes in CSF of HAM/TSP patients were distributed widely in the phylogenetic tree, but some clusters — such as clusters 8, 12, and 34 — contained the expanded TCR  $\beta$  clonotypes detected from 5 HAM/TSP patients (inner layer in Figure 3B). These results suggest that there might be relatedness of CDR3 sequences among CSF of some HAM/TSP patients. In addition, the outer layer of the dendrogram represents the expanded clonotypes of CSF classified into 3 groups based on the overlap of expanded TCR  $\beta$  clonotypes between CSF and PBMCs: (a) expanded clones with  $\geq 8$  UMIs shared by CSF and PBMCs (Figure 3B, yellow colors in the outer layer); (b) clones expanded in CSF but detected at singleton or  $2 \leq$  UMIs  $< 8$  in PBMCs (Figure 3B, gray colors in the outer layer); and (c) expanded clones unique in CSF (Figure 3B, orange colors in the outer layer). We also examined whether expanded

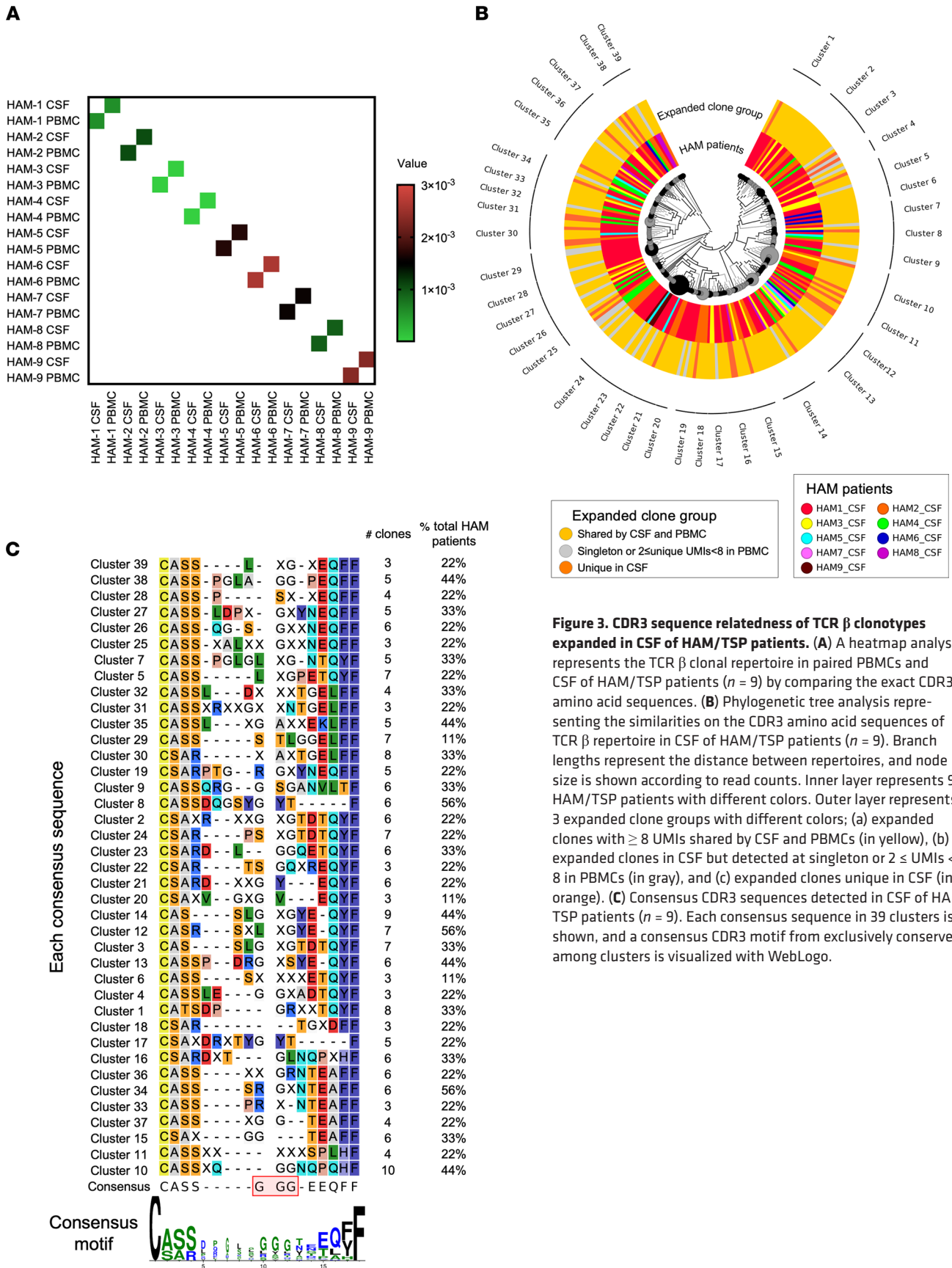


**Figure 2. Overlap of TCR  $\beta$  repertoire of expanded clones between paired blood and CSF in patients with HAM/TSP.** (A) Relationship of expanded TCR  $\beta$  clonotypes detected in PBMCs and CSF in a HAM/TSP patient (HAM-2) using a Venn diagram. The corresponding numbers of unique TCR  $\beta$  clonotype ( $\geq 8$  UMIs) are given within the diagram. (B) Proportion of shared expanded clonotypes and predominantly expanded clonotypes in PBMCs and CSF in each patient with HAM/TSP ( $n = 9$ ). (C) An average of the frequencies of shared clonotypes by CSF and PBMCs (yellow area) and predominantly expanded clonotypes in CSF (green area) in the whole TCR  $\beta$  expanded clonotypes of the CSF of HAM/TSP patients ( $n = 9$ ). In predominantly expanded clonotypes in CSF, there are 2 subgroups: clonotypes detected at singleton or  $2 \leq$  UMIs < 8 in PBMCs (gray area) and unique clonotype in CSF (undetectable in PBMCs; orange area).

clones shared by CSF and PBMCs or unique clones in CSF form a cluster; however, the expanded clone groups appeared distributed in the dendrogram, and no predominant clusters were detected (outer layer in Figure 3B). In addition, TCR clonotypes expanded in CSF were not distinguishable from the expanded TCR clonotypes shared by PBMCs and CSF by the analysis.

Based on the cluster of the dendrogram, the consensus sequences of the expanded TCR  $\beta$  clonotypes in CSF were aligned and summarized with the number of detected clones and the percentage of detected HAM/TSP patients (Figure 3C). The alignment of CDR3 sequences demonstrated that there were some similarities of CDR3 sequences between some clusters, while various CDR3 amino acid motifs were observed in the consensus sequences of each cluster. In 56% of HAM/TSP patients, the expanded TCR  $\beta$  clonotypes in CSF were detected in cluster 8, 12, or 34 (Figure 3C). Interestingly, cluster 7 and cluster 38 had a same motif PGL at positions 5–7, which were detected in 33% and 44% of HAM/TSP patients, respectively (Figure 3C). Cluster 3 and cluster 14 also had the same SLG motif at positions 8–10 in the center of the CDR3 region, and this similarity was detected in 33% and 44% of HAM/TSP patients, respectively (Figure 3C). When all clusters were compared for relatedness, a highly conserved amino acid glycine (G) at positions 10–12 of the CDR3 region was observed across the CSF repertoires of patients with HAM/TSP, in which 5 clones from 3 patients (33%) were identified. Collectively, these results demonstrate that TCR  $\beta$  clonotypes of expanded clones in CSF were unique to each patient; however, clusters with similar CDR3 motifs across HAM/TSP patients could be identified.

*Comparison of TCR repertoires among Tax11-19-specific CD8<sup>+</sup> T cells, enriched CD8<sup>+</sup> T cells, and peripheral blood in HAM/TSP patients.* Since Tax protein has been shown to be an immunodominant HTLV-1 antigen recognized by HTLV-1-specific CTLs (5), we wished to determine if public TCR clonotypes in Tax-specific CTLs could be defined. To identify TCR clonotypes in HTLV-1 Tax11-19-specific CD8<sup>+</sup> T cells, we examined and compared TCR repertoires in HAM/TSP PBMCs, in enriched CD8<sup>+</sup> T cells, and sorted Tax11-19-specific CD8<sup>+</sup> T cells from 7 HLA-A\*0201<sup>+</sup> HAM/TSP patients. Tax11-19-specific CD8<sup>+</sup> T cells were detected, with Tax11-19/HLA-A\*0201-specific tetramers in PBMCs of the 7 HAM/TSP patients at



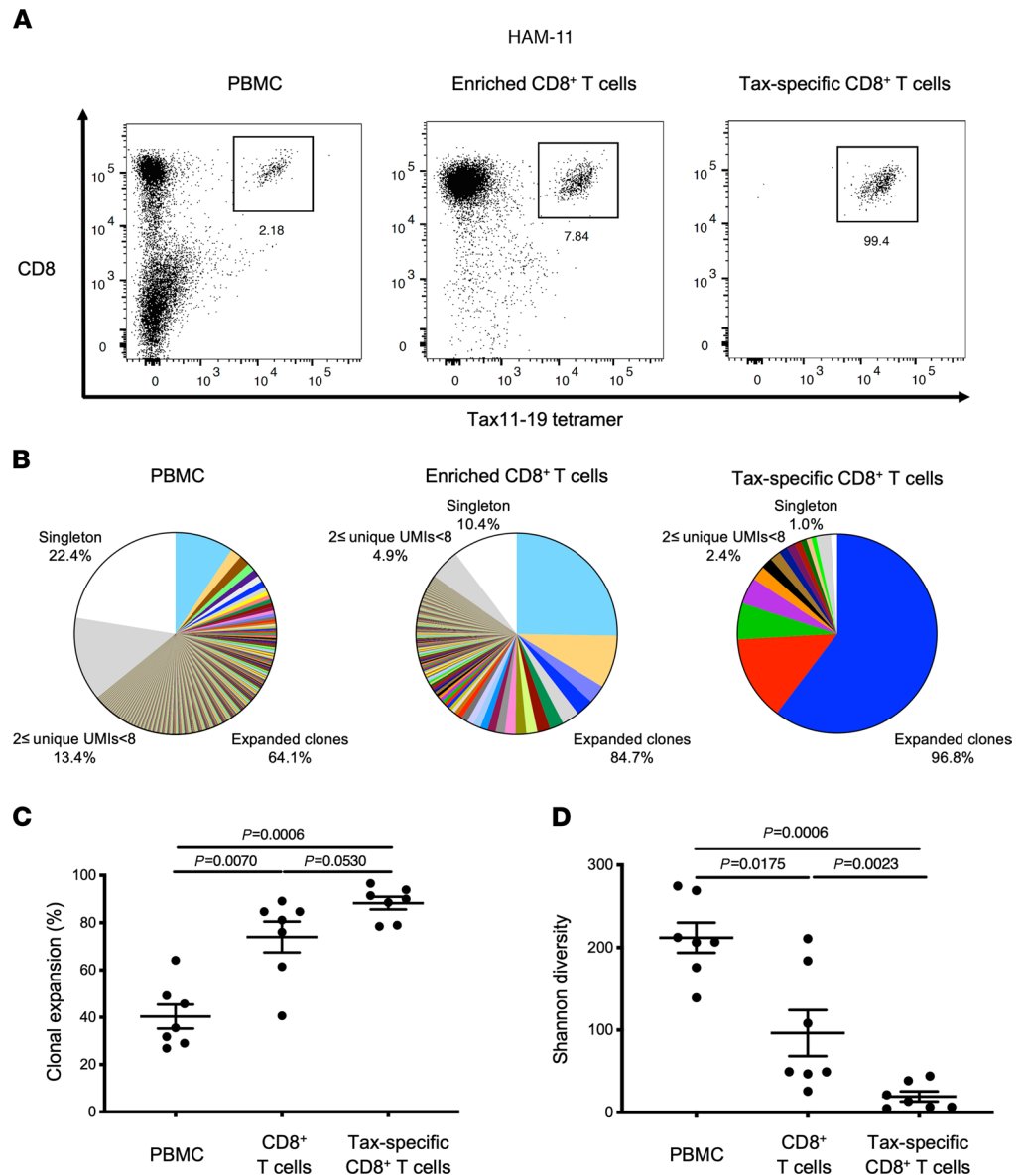
**Figure 3. CDR3 sequence relatedness of TCR β clonotypes expanded in CSF of HAM/TSP patients.** (A) A heatmap analysis represents the TCR β clonal repertoire in paired PBMCs and CSF of HAM/TSP patients ( $n = 9$ ) by comparing the exact CDR3 amino acid sequences. (B) Phylogenetic tree analysis representing the similarities on the CDR3 amino acid sequences of TCR β repertoire in CSF of HAM/TSP patients ( $n = 9$ ). Branch lengths represent the distance between repertoires, and node size is shown according to read counts. Inner layer represents 9 HAM/TSP patients with different colors. Outer layer represents 3 expanded clone groups with different colors; (a) expanded clones with  $\geq 8$  UMIs shared by CSF and PBMCs (in yellow), (b) expanded clones in CSF but detected at singleton or  $2 \leq$  UMIs  $< 8$  in PBMCs (in gray), and (c) expanded clones unique in CSF (in orange). (C) Consensus CDR3 sequences detected in CSF of HAM/TSP patients ( $n = 9$ ). Each consensus sequence in 39 clusters is shown, and a consensus CDR3 motif from exclusively conserved among clusters is visualized with WebLogo.



frequencies that ranged from 1.21% to 12.70% in CD8<sup>+</sup> T cells (Table 1). Figure 4A shows representative dot plots of Tax11-19-specific CD8<sup>+</sup> T cells of HAM-11 PBMCs, enriched CD8<sup>+</sup> T cells, and sorted Tax11-19-specific CD8<sup>+</sup> T cells. After sorting, Tax11-19-specific CD8<sup>+</sup> T cells were effectively enriched from 2.18% in PBMCs to 7.84% and 99.4% in the enriched CD8<sup>+</sup> T cells and sorted Tax11-19-specific CD8<sup>+</sup> T cells, respectively (Figure 4A). TCR repertoire analysis from these 3 populations was performed. Figure 4B shows a pie chart representation of the frequency of expanded TCR clonotypes from all 3 populations of HAM-11. A comparison of the distribution of TCR clonotypes in each group revealed that the most frequently expanded clonotypes in Tax11-19-specific CD8<sup>+</sup> T cells were not the highest-ranking expanded clonotypes in either PBMCs or purified CD8<sup>+</sup> T cells (Figure 4B). For example, the unique TCR  $\beta$  expanded clonotype depicted as blue in highly purified Tax11-19-specific CD8<sup>+</sup> T cells (representing 60.32% of all clonotypes) could only be found as the seventh most expanded clonotype in PBMCs (0.99% of all PBMC clonotypes) and the fourth most expanded clonotype in purified CD8<sup>+</sup> T cells (2.81% of all CD8<sup>+</sup> T cell clonotypes) (Figure 4B and Supplemental Table 2). Likewise, the second most expanded clonotype in Tax11-19-specific CD8<sup>+</sup> T cells (depicted in red, 13.79% of all clones) was rarely detected in PBMCs or purified CD8<sup>+</sup> T cells at 17th rank (0.37% of all PBMC clonotypes) and at 17th rank (0.80% of all CD8<sup>+</sup> T cell clonotypes), respectively (Figure 4B). Conversely, the most frequently detected clonotype in PBMCs (9.38% of all PBMC clonotypes; Figure 4B, depicted in light blue) was detected at a higher frequency in purified CD8<sup>+</sup> T cells (25.27% of all CD8<sup>+</sup> T cell clonotypes) but was only at the 20th rank of most abundant expanded clonotype in Tax11-19-specific CD8<sup>+</sup> T cells (0.15% of all Tax11-19-specific CD8<sup>+</sup> T cell clonotypes; Figure 4B and Supplemental Table 2). The complete analysis of all 7 HLA-A\*0201<sup>+</sup> HAM/TSP patients demonstrated that the most frequently expanded clonotypes in Tax11-19-specific CD8<sup>+</sup> T cells were not always detected at the higher ranks in PBMCs and CD8<sup>+</sup> T cells of HAM/TSP patients (Supplemental Figure 3). Figure 4C shows the frequency of expanded T cell clonotypes in PBMCs, enriched CD8<sup>+</sup> T cells, and sorted Tax11-19-specific CD8<sup>+</sup> T cells of HAM/TSP patients. As expected, sorted Tax11-19-specific CD8<sup>+</sup> T cells showed the highest degree of TCR clonal expansion compared with PBMCs ( $P = 0.0006$ ) and CD8<sup>+</sup> T cells ( $P = 0.0530$ ) (Figure 4C). Applying the Shannon diversity, a significantly lower diversity within the TCR repertoire was observed in sorted Tax11-19-specific CD8<sup>+</sup> T cells compared with PBMCs ( $P = 0.0006$ ) and CD8<sup>+</sup> T cells ( $P = 0.0023$ ) (Figure 4D). These results demonstrate that antigen-specific T cells have the most expanded TCR clonal repertoire compared with PBMCs or enriched CD8<sup>+</sup> T cells, suggesting that multiple clones can recognize HTLV-1 Tax11-19 in association with HLA-A\*0201. Importantly, since expanded TCR clonotypes in Tax11-19-specific CD8<sup>+</sup> T cells were not always detected at the highest rank in PBMCs or CD8<sup>+</sup> T cells, analysis of antigen-specific TCR clonotypes detectable in peripheral blood of HAM/TSP patients may not always reflect antigen-specific, immunodominant responses.

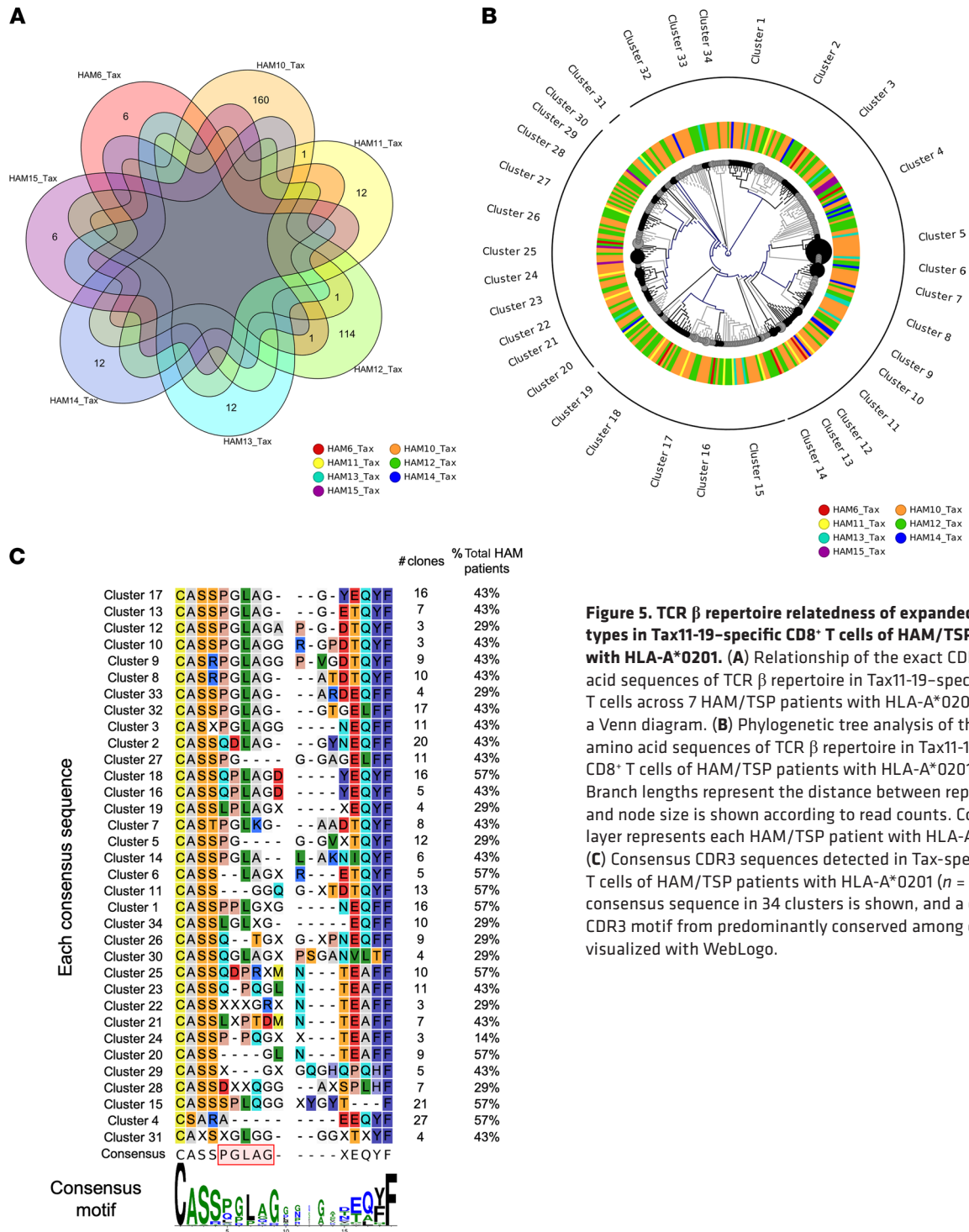
*TCR  $\beta$  clonotypes of expanded clones in HLA-A\*0201-restricted Tax11-19-specific CD8<sup>+</sup> T cells are unique to each individual but have conserved CDR3 motifs.* To determine whether there are public TCR clonotypes in HLA-A\*0201-restricted Tax-specific CTLs of HAM/TSP patients, we compared the expanded clones with at least 8 UMIs in Tax11-19-specific CD8<sup>+</sup> T cells by identical CDR3 amino acid sequences. In Figure 5A, a Venn diagram depicts the overlap between the expanded CDR3 sequences detected by each HAM/TSP patient and demonstrates no common TCR  $\beta$  clonotypes detected in all 7 patients comparing exact CDR3 amino acid sequences. Most TCR  $\beta$  clonotypes were unique to each individual patient, while a few identical CDR3 sequences were shared only between 3 pairs of HAM/TSP patients (HAM-6 and HAM-12, HAM-10 and HAM-11, and HAM-10 and HAM-12; Figure 5A). These results indicate that HAM/TSP patients known to recognize the same immunodominant HTLV-1 Tax11-19 peptide in association with the HLA-A\*0201 allele use a private TCR repertoire in which no identical shared TCR  $\beta$  clonotypes could be detected.

We next examined whether there is any CDR3 sequence relatedness in the Tax-specific CTLs. Phylogenetic analysis was performed using expanded TCR  $\beta$  clonotypes with more than 8 UMIs (total of 336 clones) in Tax11-19-specific CD8<sup>+</sup> T cells of 7 HLA-A\*0201<sup>+</sup> HAM/TSP patients (Figure 5B). Sequence data for constructing the tree are provided in Supplemental Table 3. As described above, this approach performed an all-versus-all pairwise comparison of clonotypes detected for a list of patients with HAM/TSP in Tax11-19-specific CD8<sup>+</sup> T cells, and this comparison was then used to cluster samples hierarchically, generating a sample-level dendrogram (Figure 5B). The phylogenetic analysis showed the similarities in the TCR  $\beta$  repertoires and the identification of 34 clusters that shared common CDR3 motifs by sequence similarity in Tax11-19-specific CD8<sup>+</sup> T cells of HAM/TSP patients (Figure 5B). Interestingly, some clusters, such as clusters 1, 4, 6, 11, 15, 18, 20, and 25, contained the expanded TCR  $\beta$  clonotypes



**Figure 4. TCR clonal expansion of TCR  $\beta$  repertoire in PBMCs, enriched CD8<sup>+</sup> T cells, and sorted Tax11-19-specific CD8<sup>+</sup> T cells of HAM/TSP patients with HLA-A\*0201. (A)** Tax tetramer staining in CD3<sup>+</sup> T cells of PBMCs (left), enriched CD8<sup>+</sup> T cells (center), and sorted Tax11-19-specific CD8<sup>+</sup> T cells (right) in a HAM/TSP patient (HAM-11). **(B)** Representative pie charts of T cell clonal expansion in PBMCs, enriched CD8<sup>+</sup> T cells, and sorted Tax11-19-specific CD8<sup>+</sup> T cells of a HAM/TSP patient (HAM-11). TCR  $\beta$  clonal expansion was analyzed by using the frequency of clones  $\geq 8$  UMIs (colored wedges), clones with  $2 \leq$  UMIs  $< 8$  (gray), and singletons (white). In the group of expanded clones, each wedge represents a unique clonotype with a defined CDR3 sequence, and the same clone shared among PBMCs, CD8<sup>+</sup> T cells, and Tax11-19-specific CD8<sup>+</sup> T cells is visualized in the same color. **(C)** Comparison of TCR clonal expansion by frequency of clones  $\geq 8$  UMIs among PBMCs, enriched CD8<sup>+</sup> T cells, and sorted Tax11-19-specific CD8<sup>+</sup> T cells in HAM/TSP patients with HLA-A\*0201 ( $n = 7$ ) using the Friedman test's post hoc test. Data represent mean  $\pm$  SEM. **(D)** Comparison of Shannon diversity among PBMCs, enriched CD8<sup>+</sup> T cells, and sorted Tax11-19-specific CD8<sup>+</sup> T cells in HAM/TSP patients ( $n = 7$ ) using the Friedman test's post hoc test. Data represent mean  $\pm$  SEM.

detected from 4 of 7 HAM/TSP patients (Figure 5B), suggesting that TCR  $\beta$  repertoire-level signatures and similarities could be identified in Tax-specific CTLs of HAM/TSP patients. The consensus sequences of the expanded TCR  $\beta$  clonotypes in Tax11-19-specific CD8<sup>+</sup> T cells were aligned and compared among all of the clusters (Figure 5C). The alignment of TCR  $\beta$  CDR3 sequences in Tax11-19-specific CD8<sup>+</sup> T cells from all of the clusters showed a highly conserved PGLAG amino acid sequence motif at the position 4–8 of the CDR3 region, which was detected in 46 clones (14%) and 4 HAM/TSP patients (57%).



**Figure 5. TCR  $\beta$  repertoire relatedness of expanded clonotypes in Tax11-19-specific CD8<sup>+</sup> T cells of HAM/TSP patients with HLA-A\*0201.** (A) Relationship of the exact CDR3 amino acid sequences of TCR  $\beta$  repertoire in Tax11-19-specific CD8<sup>+</sup> T cells across 7 HAM/TSP patients with HLA-A\*0201 using a Venn diagram. (B) Phylogenetic tree analysis of the CDR3 amino acid sequences of TCR  $\beta$  repertoire in Tax11-19-specific CD8<sup>+</sup> T cells of HAM/TSP patients with HLA-A\*0201 (n = 7). Branch lengths represent the distance between repertoires, and node size is shown according to read counts. Color-coded layer represents each HAM/TSP patient with HLA-A\*0201. (C) Consensus CDR3 sequences detected in Tax-specific CD8<sup>+</sup> T cells of HAM/TSP patients with HLA-A\*0201 (n = 7). Each consensus sequence in 34 clusters is shown, and a consensus CDR3 motif from predominantly conserved among clusters is visualized with WebLogo.

Our results demonstrate that patients having the same HLA-A type used private TCR repertoires to recognize HTLV-1 Tax11-19 peptide; however, there was a conserved CDR3 motif (PGLAG) in HTLV-1 Tax-specific CTLs across HLA-A\*0201<sup>+</sup> HAM/TSP patients.

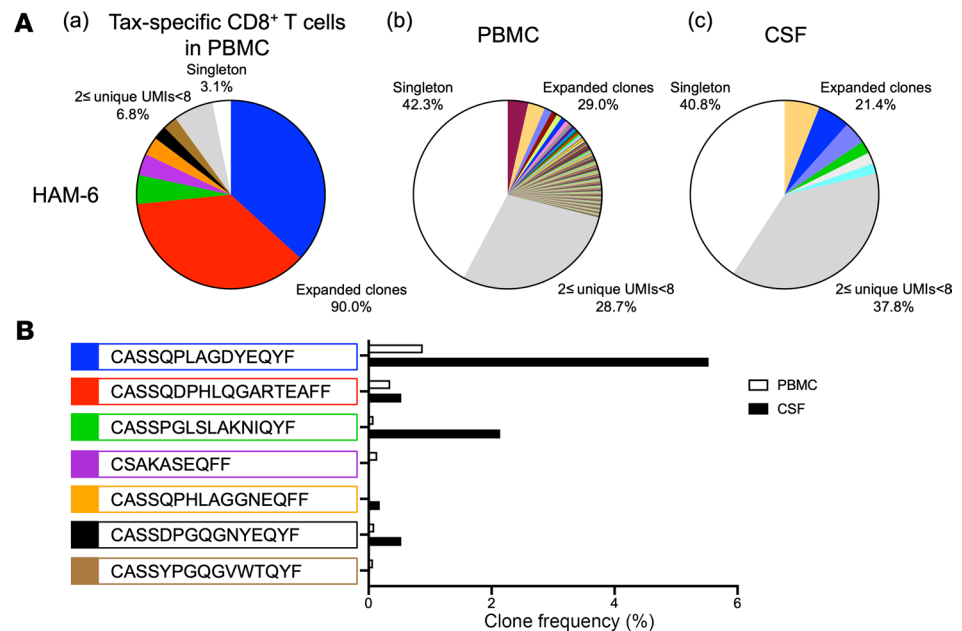
*TCR  $\beta$  expanded clonotypes of HTLV-1-specific CD8<sup>+</sup> T cells are enriched in CSF of a HAM/TSP patient.* Since it has been reported that Tax-specific CTLs are elevated both in PBMCs and CSF of HAM/TSP patients (6), expanded CTL clones in peripheral blood are thought to cross the blood-brain barrier and cause CNS inflammation (7). However, it remains unclear whether the expanded CTL clones in CSF are derived from the peripheral blood or are expanded compartmentally in the CNS. We had the opportunity to obtain the CSF from 1

HLA-A\*0201<sup>+</sup> HAM/TSP patient (HAM-6) from which we could compare the TCR  $\beta$  clonally expanded repertoires of paired samples of Tax11-19-specific CD8<sup>+</sup> T cells in PBMCs, total PBMCs, and total CSF cells (Figure 6A; left, middle, and right, respectively). In Tax11-19-specific CD8<sup>+</sup> T cells isolated from peripheral blood, 7 TCR expanded clonotypes were detected (Figure 6A, left, and Table 3). Two TCR  $\beta$  expanded clonotypes identified at the first (depicted in blue) and third rank (depicted in green) in Tax11-19-specific CD8<sup>+</sup> T cells (CASSQPLAGDYEQYF and CASSPGLSLAKNIQYF) were also detected at the second and the fourth rank in the expanded clonotypes of CSF, respectively (Figure 6A, right, and Table 3). The second most expanded TCR clonotype (depicted in red) from Tax11-19-specific CD8<sup>+</sup> T cells (CASSQDPHLQGARTEAFF) was present ( $2 \leq \text{UMIs} < 8$ ) at the 23rd rank in CSF of HAM-6 (0.53% of all CSF clones) but did not meet our definition of expanded clones  $\geq 8$  UMIs (Figure 6A, right, and Table 3). The 2 TCR  $\beta$  sequences of Tax11-19-specific CD8<sup>+</sup> T cells in PBMCs (CASSQPLAGDYEQYF and CASSPGLSLAKNIQYF) were also detected at the sixth (0.88% of all PBMC clones) and the 69th (0.08% of all PBMC clones) rank, respectively, in the expanded clonotypes of PBMCs (Figure 6A, middle, and Table 3). Importantly, when comparing the frequency of the 7 TCR  $\beta$  sequences detected in the PBMCs and the CSF of HAM-6, these 2 TCR  $\beta$  sequences of Tax11-19-specific CD8<sup>+</sup> T cells (CASSQPLAGDYEQYF and CASSPGLSLAKNIQYF) were detected more in the CSF compared with the peripheral blood (Figure 6B). Collectively, our results suggest that a subset of HTLV-1 Tax-specific TCR  $\beta$  clonotypes was clonally expanded in peripheral blood and was subsequently infiltrated, becoming highly enriched in the CSF compartment.

## Discussion

The TCR repertoire is shaped by antigen engagement and altered in the context of the disease (10). A more complete characterization of TCR repertoires may provide disease-specific signatures, including specific TCR clonotypes or oligoclonal skewed TCR repertoires in immune-mediated disorders, such as chronic viral infection and autoimmune disease. Furthermore, exploring the TCR repertoire signatures focused on expanded clones particularly in the CNS may clarify local antigen-driven T cell activation and pathologically related TCR sequences in neuroinflammatory disorders. CSF contains predominantly T cells (approximately 80%) and has unique immunoregulatory roles including control of T cell entry and migration within the CNS, and it may be associated with both protective and pathological immune responses (16, 27). To explore disease-related TCR repertoire signatures in a neurologic disorder, we investigated the TCR repertoire from CSF cells in patients with HAM/TSP, a chronic inflammatory neurologic disease associated with HTLV-1 infection. TCR  $\beta$  repertoires of HAM/TSP patients in both PBMCs and CSF were more clonally expanded compared with those of NDs. Moreover, an overlap of TCR  $\beta$  clonotypes of expanded clones was identified between peripheral blood and CSF in HAM/TSP patients. By comparison, we also had the opportunity to obtain CSF and matched PBMCs from healthy, HTLV-1-seronegative control individuals in which expanded TCR clonotypes were much less or undetectable in CSF, although — as in HAM/TSP — individual variations of clonal expansion were also observed in PBMCs of NDs. The absence of clonally expanded TCR  $\beta$  clonotypes in NDs might be associated with low number of activated T cells found in CSF of NDs, compared with HAM/TSP patients (28). The degree of TCR  $\beta$  clonal expansions in PBMCs was correlated with effector T cells and HTLV-1 PVL in the periphery of HAM/TSP patients. HAM/TSP patients are known to have a higher PVL compared with asymptomatic carriers, and this comparison is associated with increased phenotypically defined effector CD8<sup>+</sup> T cell subsets and HTLV-1 Tax-specific CD8<sup>+</sup> T cells (26, 29) where chronic viral antigen stimulations have been shown to drive clonal expansion in both infected CD4<sup>+</sup> T cells and HTLV-1-specific CD8<sup>+</sup> T cells (30). These previous reports support results in this study that demonstrate higher TCR  $\beta$  clonal expansions in peripheral blood, as well as in CSF in HAM/TSP compared with NDs. Collectively, these results strongly suggest that TCR repertoires in CSF of HAM/TSP patients may provide a local disease-specific T cell immune signature.

Although our results show that there was no identical CDR3 sequence in CSF across all HAM/TSP patients, TCR  $\beta$  repertoire clusters were detected using phylogenetic tree cluster analysis. Approximately 80% of the expanded TCR clones in CSF of HAM/TSP patients were derived from expanded clones in PBMCs, while 9% of clones appeared to be uniquely expanded in CSF of these patients. The expanded T cell clones shared by CSF and PBMCs compared with those expanded TCR clones unique to CSF may reflect a variation in the antigens recognized in each compartment and the capacity to clonally expand at each site. Indeed, TCR repertoire analyses in MS patients have shown that expanded T cell clones shared by the CNS and periphery are predominantly observed in the peripheral CD8<sup>+</sup> T cell compartment (18, 31), while expanded T cell clones



**Figure 6. Enrichment of Tax-specific TCR  $\beta$  expanded clonotypes in CSF of a HAM/TSP patient.** (A) T cell clonal expansion of sorted Tax11-19-specific CD8<sup>+</sup> T cells from PBMCs (left), PBMCs (middle), and CSF (right) of a HAM/TSP patient (HAM-6). TCR  $\beta$  clonal expansion was analyzed by using the frequency of clones  $\geq 8$  UMIs (colored wedges), clones with  $2 \leq$  UMIs  $< 8$  (gray), and singletons (white). In the group of expanded clones, each wedge represents a unique clonotype with a defined CDR3 sequence, and the same clone shared among Tax11-19-specific CD8<sup>+</sup> T cells, PBMCs, and CSF cells is visualized in the same color. (B) Frequencies of 7 TCR  $\beta$  expanded clonotypes of Tax11-19-specific CD8<sup>+</sup> T cells in PBMCs and CSF of a HAM-TSP patients (HAM-6). Each clonotype is visualized in the same color with the wedge of the pie chart in A.

unique to the CNS are found in peripheral CD4<sup>+</sup> T cells (31). In our study, no predominant cluster shared by CSF and PBMCs or unique to CSF was identified in HAM/TSP patients using phylogenetic tree analysis. Since T cell clones exclusively expanded in the CNS are considered to reflect local inflammation (18, 31), particularly in an immunopathologically mediated disease such as HAM/TSP, identification of TCR repertoire signatures that are disease and/or pathogen specific would be important.

Since T cells recognize foreign peptides through engagement with TCR and peptide-MHC (8), defining a pathogen-reactive TCR repertoire signature requires donor HLA haplotype information to filter the majority of TCRs that would not recognize a pathogen in the context of a particular HLA (10, 32). However, public TCR sequences, which are shared among multiple individuals in response to the same antigenic epitopes, have been extensively observed to a variety of viruses, including EBV, cytomegalovirus, herpes simplex virus, and HIV, and have been associated with favorable biological outcomes (33). However, more recently, TCR analyses based on HTS have demonstrated that there can be diversity of TCR sequences that recognize even the same epitopes (34–36). Consistent with these studies, our results show that HLA-A\*0210-restricted Tax11-19-specific CD8<sup>+</sup> T cells from HLA-A\*0201<sup>+</sup> HAM/TSP patients used private TCR  $\beta$  sequences to recognize the same immunodominant HTLV-1 Tax antigen. Public TCR  $\beta$  sequences (as defined by identical CD3 sequences from expanded TCR clonotypes) were rarely identified in Tax11-19-specific CD8<sup>+</sup> T cells in 7 HAM/TSP patients (Figure 5A). However, phylogenetic tree cluster analysis showed that TCR  $\beta$  repertoires of these Tax11-19-specific CD8<sup>+</sup> T cells had similar sequences and conserved motifs, such as PGLAG, in the CDR3 region. These observations are consistent with previous reports on the TCR analysis of the CDR3 region of Tax11-19-specific CD8<sup>+</sup> T cells in HLA-A\*0201<sup>+</sup> HAM/TSP patients that showed conserved motifs of 3 amino acid sequence (PG-G) (23) or 4 sequence (P/G-L-A/R-G) (22). Moreover, crystal structure analysis revealed a corner on the apex of the TCR  $\beta$  CDR3 loop (GLAG) that was shown to insert into a hydrophobic pocket formed by the  $\alpha$ -1 helix of HLA-A2 and the side chain of the Tyr residue at position 8 in the Tax11-19 peptide (22). These observations support that the motif of PGLAG is important for recognition of HTLV-1 Tax protein by antigen-specific CD8<sup>+</sup> T cells of HAM/TSP patients. Since Tax-specific CTLs are considered to play a key role in the pathogenesis of HAM/TSP (7), further research is needed to examine whether CD8<sup>+</sup> T cells with these specific motifs have protective or deleterious effects against HTLV-1 infection.

**Table 3. TCR  $\beta$  clonotypes of expanded clones shared by 3 compartments**

CDR3 sequence	Tax11-19-specific CD8 <sup>+</sup> T cells			PBMCs			CSF		
	Rank	Count	%	Rank	Count	%	Rank	Count	%
CASSQPLAGDYEQYF	1	199	36.78	6	217	0.88	2	31	5.53
CASSQPHLQGARTAEFF	2	198	36.60	12	86	0.35	23	3	0.53
CASSPGLSLAKNIQYF	3	27	4.99	69	19	0.08	4	12	2.14
CSAKASEQFF	4	20	3.70	28	35	0.14	–	–	–
CASSQPHLAGGNEQFF	5	17	3.14	1077	3	0.01	205	1	0.18
CASSDPGQGNYEQYF	6	13	2.40	49	22	0.09	20	3	0.53
CASSYPGQGVVTQYF	7	13	2.40	94	16	0.07	–	–	–

PBMCs, peripheral blood mononuclear cells; CSF, cerebrospinal fluid; CDR3, complementary determining region 3.

In the present study, we also had the opportunity to obtain CSF cells from an HLA-A\*0201<sup>+</sup> HAM/TSP patient from which we identified HTLV-1-specific TCR  $\beta$  repertoires in this compartment. We demonstrated that TCR  $\beta$  clonotypes of Tax11-19-specific CD8<sup>+</sup> T cells in the periphery were also clonally expanded and enriched in the CSF. This result is a direct demonstration that activated and clonally expanded HTLV-1-specific CD8<sup>+</sup> T cells in peripheral blood infiltrate and become highly enriched in the CNS of HAM/TSP. Interestingly, the amino acid sequences of the 2 CSF expanded TCR  $\beta$  clonotypes (CASSQPLAGDYEQYF at positions 6–9 and CASSPGLSLAKNIQYF at positions 5–7; Figure 6B and Table 3) were similar to the PLAG motif of the consensus sequence identified from HLA-A\*0201-restricted Tax11-19-specific CD8<sup>+</sup> T cells in the peripheral blood of HLA-A\*0201<sup>+</sup> HAM/TSP patients (Figure 5C). Collectively, these results suggest that these TCR  $\beta$  clonotypes may be associated with the HTLV-1- or disease-specific TCR repertoire signatures. Analysis of TCR repertoires in other diseases associated with antigen-specific T cell responses has also shown enrichment of antigen-specific CD8<sup>+</sup> T cells in sites of inflammation that were present in peripheral blood. These include enrichment of EBV-reactive CD8<sup>+</sup> T cells in CSF of MS patients (17), influenza-specific CD8<sup>+</sup> T cells in human lungs (37), and islet-reactive CD8<sup>+</sup> T cells in the pancreas of type 1 diabetes patients (38). Analysis of antigen-specific TCR repertoire signatures in the periphery could therefore be used to track and identify pathogens in sites of local inflammation. Recently, several TCR databases have been developed to annotate individual TCR repertoires to TCR sequences associated with known antigens, pathogens, and pathologies based on published literature (32). We annotated TCR  $\beta$  repertoires of expanded clones obtained from CSF cells and Tax11-19-specific CD8<sup>+</sup> T cells of HAM/TSP patients to representative TCR databases using VDJdb (36) and McPAS-TCR (39), although identical TCR  $\beta$  clonotypes were rarely identified (4 clonotypes in VDJdb and 0 in McPAS-TCR). The TCR database is presumed to need larger numbers of TCR sequences for clinical applications.

In conclusion, we have shown that TCR  $\beta$  repertoires of HAM/TSP patients are highly expanded in CSF and contain both TCR clonotypes shared with PBMCs and uniquely enriched clonotypes within the CSF. Analysis of the TCR  $\beta$  repertoire of HLA-A\*0201-restricted Tax11-19-specific CD8<sup>+</sup> T cells demonstrated the use of private TCR  $\beta$  sequences for the recognition of antigen and identified conserved motifs in the CDR3 region. Moreover, TCR  $\beta$  clonotypes of expanded clones in HTLV-1-specific CD8<sup>+</sup> T cells in the periphery were also expanded and enriched in the CSF. Exploring the TCR repertoire of CSF and antigen-specific T cells may provide a TCR repertoire signature in virus-associated neurologic disorders. Further analysis and technologies, such as the paired analysis of the 2 TCR chains ( $\alpha\beta$ ,  $\gamma\delta$ ) and single-cell isolation technologies, can serve to highlight rare and functionally important populations in the immune system of the CNS that may lead to a better understanding of disease pathogenesis.

## Methods

**Subjects.** Twenty subjects consisting of 15 HAM/TSP patients and 5 NDs were included for sequencing of TCR  $\beta$  chain repertoires by HTS. HAM/TSP patients were defined by World Health Organization criteria, and HTLV-1-uninfected healthy volunteers screened at the NIH Clinical Center (Bethesda, Maryland, USA) were evaluated as NDs. Subsets of HAM/TSP patients were used for specific studies. For analysis of paired PBMCs and CSF cells, peripheral blood and CSF samples were collected on the same day from 9 HAM/TSP patients and 5 NDs. For analysis of HTLV-1-specific CD8<sup>+</sup> T cells, peripheral blood was

obtained from 7 HLA-A\*0201<sup>+</sup> HAM/TSP patients. Subject characteristics including sex, age, ethnicity, disease status, and HLA-A type are listed in Table 1. PBMCs were isolated from the peripheral blood according to the Ficoll-Hypaque density gradient centrifugation method according to the manufacturer's instructions and then cryopreserved in liquid nitrogen until the date of use. CSF samples were obtained by nontraumatic lumbar puncture at volumes greater than 20 mL; subsequently, cells were collected by centrifugation at 400g for 10 minutes at 4°C. The CSF cells were cryopreserved in RNAlater (Invitrogen, Thermo Fisher Scientific) solution for TCR sequencing.

**HTLV-1 PVL.** HTLV-1 PVL from HAM/TSP patients was quantified based on the percentage of HTLV-1 *tax* gene detection in PBMCs by using the droplet digital PCR technique (Bio-Rad), as previously reported (25). DNA was extracted from the PBMCs and CSF cell pellets using a DNeasy Blood and Tissue kit (Qiagen) according to the manufacturer's instructions. Primers and probes specific for HTLV-1 *tax* and human ribonuclease P protein subunit 30 (*RPP30*) were used (25). All samples were tested in duplicate, and PVL was reported as the average of the 2 measurements.

**Flow cytometry and sorting.** For flow cytometric analysis, EDTA-treated whole blood from NDs and HAM/TSP patients were stained with CD3, CD4, CD8, CD27, CD45, and CD45RA (all from BD Biosciences), as well as T cell subtypes, were defined as effector/memory (CD27<sup>-</sup>CD45RA<sup>-</sup>) and effector (CD27<sup>+</sup>CD45RA<sup>+</sup>) subsets in CD4<sup>+</sup> or CD8<sup>+</sup> T cells as previously described (28, 40). All flow cytometric analyses were performed using a LSR II (BD Biosciences). For staining of HTLV-1 Tax-specific CD8<sup>+</sup> T cells in PBMCs, Tax11-19/HLA-A\*0201 tetramer (provided by NIH Tetramer Core Facility, Atlanta, Georgia, USA) was used. For sorting of Tax11-19-specific CD8<sup>+</sup> T cells, after CD8<sup>+</sup> T cells were enriched from 7 HAM/TSP patients with HLA-A\*0201 using human CD8<sup>+</sup> T cell isolation kit (Miltenyi Biotec), Tax11-19-specific CD8<sup>+</sup> T cells were sorted from the enriched CD8<sup>+</sup> T cells using FACSaria II (BD Biosciences). The purity of enriched CD8<sup>+</sup> T cells and Tax11-19-specific CD8<sup>+</sup> T cells was 82.0%–96.3% and 93.0%–100.0%, respectively. All the data were analyzed using FlowJo 10.6 software. All antibodies used for flow cytometry are provided in Supplemental Table 4.

**TCR  $\beta$  library preparation.** HTS for TCR  $\beta$  chain was performed on PBMCs, enriched CD8<sup>+</sup> T cells, sorted Tax11-19-specific CD8<sup>+</sup> T cells and CSF cells. The TCR  $\beta$  chain was amplified from  $2 \times 10^6$  cells of PBMCs and enriched CD8<sup>+</sup> T cells. The sorted Tax11-19-specific CD8<sup>+</sup> T cells and CSF cells were used at  $4 \times 10^4$  to  $5 \times 10^5$  cells and  $1 \times 10^4$  to  $2.5 \times 10^5$  cells, respectively, for the amplification of TCR  $\beta$  chain. Total RNA was extracted from PBMCs and enriched CD8<sup>+</sup> T cells using a miRNA Mini Kit (Qiagen), and RNA from sorted Tax11-19-specific CD8<sup>+</sup> T cells and CSF cells was extracted using a RNeasy Micro kit (Qiagen) following the manufacturer's instructions. The rearranged TCR  $\beta$  repertoires were amplified using previously described methods (14). Briefly, cDNA synthesis was performed using the anchored switch 5' rapid amplification of cDNA ends (5' RACE) PCR-based primer combined with a unique molecular barcode that consists of 10 random nucleotide sequences. Nested PCR was performed using a pair of primers specific for the TCR constant region and 5' RACE region, which contains a sample index and Illumina adapter sequences. Final PCR product was purified, and the quality of TCR library was analyzed using a Bioanalyzer (Agilent Technologies).

**TCR  $\beta$  repertoire sequencing and bioinformatics analysis.** TCR library were sequenced on the MiSeq platform with a 150 bp paired-end run (Illumina). After quality filtering, raw sequence reads were then processed using the MiGEC software pipeline with default parameter (15) for CDR3 extraction and VDJ gene segment alignment to the human TCR  $\beta$  germline sequences based on IMGT database. We conducted further analysis using VDJtools (41) and tcR (42), which allowed applying postanalysis of repertoire sequencing data and generated basic sample statistics. Read counts, number of clonotypes, diversity estimations, and repertoire overlap analysis within each sample or across samples were evaluated and visualized using these software tools. To evaluate the TCR  $\beta$  clonal expansion, TCR  $\beta$  clones in each subject were classified into one of the 3 groups: singletons (1 UMI), clones ( $2 \leq$  UMIs  $< 8$ ), and clones  $\geq 8$  UMIs, as previously described based on coefficients of variation analyses for each clonotype by the average number of UMIs observed across technical triplicates (14). Those clones having at least 8 UMIs were used to perform phylogenetic tree cluster analyses with respect to CDR3 amino acid sequence similarities across samples to determine clonal relatedness using CLCbio workbench (Qiagen). Consensus sequence data from the cluster analysis were then visualized with WebLogo (43).

**Statistics.** The Mann-Whitney *U* test was used to compare clonal expansion both in PBMCs and CSF between HAM/TSP patients and NDs. The Friedman test's post hoc test was used to compare clonal expansion and Shannon diversity index among PBMCs, enriched CD8<sup>+</sup> T cells, and Tax11-19-specific

CD8<sup>+</sup> T cells in HAM/TSP patients with HLA-A\*0201. Wilcoxon's signed rank test was used to compare HTLV-1 PVL between PBMCs and CSF in HAM/TSP patients. Spearman's rank correlation test was used to compare effector/memory and effector cells both in CD4<sup>+</sup> and CD8<sup>+</sup> T cells or HTLV-1 PVL in PBMCs with clonal expansion in PBMCs. All statistical analysis was performed using Prism 8 (GraphPad software) or R software (R version 3.5.1).

*Study approval.* All samples from HAM/TSP patients and NDs were collected under protocol (National Institute of Neurological Disorders and Stroke protocol nos. 98-N-0047 and 13-N-0149). All subjects gave written informed consent before inclusion in accordance with the Declaration of Helsinki, and the study was reviewed and approved by the NIH IRB.

### Author contributions

SN designed the study, performed the experiments, analyzed the data, and wrote the manuscript. YEA performed the experiments, analyzed the data, and wrote the manuscript. KRJ performed the bioinformatic analysis and analyzed the data. MCM and NN performed the experiments. AE and JZ provided experimental and high-throughput sequencing support. JO provided clinical support. SJ supervised the project, interpreted the data, and wrote the manuscript.

### Acknowledgments

This research was supported by the Intramural Research Program of the NIH, NINDS. SN was supported by JSPS Research Fellowship for Japanese Biomedical and Behavioral Research at NIH.

Address correspondence to: Steven Jacobson, Viral Immunology Section, National Institute of Neurological Disorders and Stroke, NIH, 10 Center Drive, Room 5C103, Bethesda, Maryland 20892, USA. Phone: 301.496.0519; Email: jacobsons@ninds.nih.gov.

1. Uchiyama T, et al. Adult T-cell leukemia: clinical and hematologic features of 16 cases. *Blood*. 1977;50(3):481–492.
2. Osame M, et al. HTLV-I associated myelopathy, a new clinical entity. *Lancet*. 1986;1(8488):1031–1032.
3. Kubota R, et al. Demonstration of human T lymphotropic virus type I (HTLV-I) tax-specific CD8<sup>+</sup> lymphocytes directly in peripheral blood of HTLV-I-associated myelopathy/tropical spastic paraparesis patients by intracellular cytokine detection. *J Immunol*. 1998;161(1):482–488.
4. Yamano Y, et al. Correlation of human T-cell lymphotropic virus type 1 (HTLV-1) mRNA with proviral DNA load, virus-specific CD8(+) T cells, and disease severity in HTLV-I-associated myelopathy (HAM/TSP). *Blood*. 2002;99(1):88–94.
5. Jacobson S, et al. Circulating CD8<sup>+</sup> cytotoxic T lymphocytes specific for HTLV-I pX in patients with HTLV-I associated neurological disease. *Nature*. 1990;348(6298):245–248.
6. Nagai M, et al. Increased HTLV-I proviral load and preferential expansion of HTLV-I Tax-specific CD8<sup>+</sup> T cells in cerebrospinal fluid from patients with HAM/TSP. *Ann Neurol*. 2001;50(6):807–812.
7. Jacobson S. Immunopathogenesis of human T cell lymphotropic virus type I-associated neurologic disease. *J Infect Dis*. 2002;186(Suppl 2):S187–S192.
8. Davis MM, Bjorkman PJ. T-cell antigen receptor genes and T-cell recognition. *Nature*. 1988;334(6181):395–402.
9. Laydon DJ, et al. Estimating T-cell repertoire diversity: limitations of classical estimators and a new approach. *Philos Trans R Soc Lond B Biol Sci*. 2015;370(1675):20140291.
10. Attaf M, et al.  $\alpha\beta$  T cell receptors as predictors of health and disease. *Cell Mol Immunol*. 2015;12(4):391–399.
11. Bianconi E, et al. An estimation of the number of cells in the human body. *Ann Hum Biol*. 2013;40(6):463–471.
12. Krogsgaard M, Davis MM. How T cells 'see' antigen. *Nat Immunol*. 2005;6(3):239–245.
13. Rosati E, et al. Overview of methodologies for T-cell receptor repertoire analysis. *BMC Biotechnol*. 2017;17(1):61.
14. Alves Sousa AP, et al. Comprehensive analysis of TCR-beta repertoire in patients with neurological immune-mediated disorders. *Sci Rep*. 2019;9(1):344.
15. Shugay M, et al. Towards error-free profiling of immune repertoires. *Nat Methods*. 2014;11(6):653–655.
16. Wilson EH, et al. Trafficking of immune cells in the central nervous system. *J Clin Invest*. 2010;120(5):1368–1379.
17. Lossius A, et al. High-throughput sequencing of TCR repertoires in multiple sclerosis reveals intrathecal enrichment of EBV-reactive CD8<sup>+</sup> T cells. *Eur J Immunol*. 2014;44(11):3439–3452.
18. Salou M, et al. Expanded CD8 T-cell sharing between periphery and CNS in multiple sclerosis. *Ann Clin Transl Neurol*. 2015;2(6):609–622.
19. Stangel M, et al. The utility of cerebrospinal fluid analysis in patients with multiple sclerosis. *Nat Rev Neurol*. 2013;9(5):267–276.
20. Eiraku N, et al. Clonal expansion within CD4<sup>+</sup> and CD8<sup>+</sup> T cell subsets in human T lymphotropic virus type I-infected individuals. *J Immunol*. 1998;161(12):6674–6680.
21. Ureta-Vidal A, et al. Human T cell leukemia virus type I (HTLV-I) infection induces greater expansions of CD8 T lymphocytes in persons with HTLV-I-associated myelopathy/tropical spastic paraparesis than in asymptomatic carriers. *J Infect Dis*. 2001;183(6):857–864.
22. Saito M, et al. In vivo selection of T-cell receptor junctional region sequences by HLA-A2 human T-cell lymphotropic virus



- type 1 Tax11-19 peptide complexes. *J Virol.* 2001;75(2):1065–1071.
23. Bourcier KD, et al. 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/TCR crystal structure. *J Virol.* 2001;75(20):9836–9843.
24. Nakagawa M, et al. HTLV-I-associated myelopathy: analysis of 213 patients based on clinical features and laboratory findings. *J Neurovirol.* 1995;1(1):50–61.
25. Brunetto GS, et al. Digital droplet PCR (ddPCR) for the precise quantification of human T-lymphotropic virus 1 proviral loads in peripheral blood and cerebrospinal fluid of HAM/TSP patients and identification of viral mutations. *J Neurovirol.* 2014;20(4):341–351.
26. Nagai M, et al. Increased activated human T cell lymphotropic virus type I (HTLV-I) Tax11-19-specific memory and effector CD8+ cells in patients with HTLV-I-associated myelopathy/tropical spastic paraparesis: correlation with HTLV-I provirus load. *J Infect Dis.* 2001;183(2):197–205.
27. Ousman SS, Kubes P. Immune surveillance in the central nervous system. *Nat Neurosci.* 2012;15(8):1096–1101.
28. Enose-Akahata Y, et al. Immunophenotypic characterization of CSF B cells in virus-associated neuroinflammatory diseases. *PLoS Pathog.* 2018;14(4):1007042.
29. Nagai M, et al. Analysis of HTLV-I proviral load in 202 HAM/TSP patients and 243 asymptomatic HTLV-I carriers: high proviral load strongly predisposes to HAM/TSP. *J Neurovirol.* 1998;4(6):586–593.
30. Enose-Akahata Y, et al. Role of HTLV-I Tax and HBZ in the Pathogenesis of HAM/TSP. *Front Microbiol.* 2017;8:2563.
31. Planas R, et al. Detailed characterization of T cell receptor repertoires in multiple sclerosis brain lesions. *Front Immunol.* 2018;9:509.
32. Zvyagin IV, et al. An overview of immunoinformatics approaches and databases linking T cell receptor repertoires to their antigen specificity. *Immunogenetics.* 2019;72(1–2):77–84.
33. Li H, et al. Determinants of public T cell responses. *Cell Res.* 2012;22(1):33–42.
34. Qi Q, et al. Diversification of the antigen-specific T cell receptor repertoire after varicella zoster vaccination. *Sci Transl Med.* 2016;8(332):332ra46.
35. Glanville J, et al. Identifying specificity groups in the T cell receptor repertoire. *Nature.* 2017;547(7661):94–98.
36. Shugay M, et al. VDJdb: a curated database of T-cell receptor sequences with known antigen specificity. *Nucleic Acids Res.* 2018;46(D1):D419–D427.
37. Pizzolla A, et al. Influenza-specific lung-resident memory T cells are proliferative and polyfunctional and maintain diverse TCR profiles. *J Clin Invest.* 2018;128(2):721–733.
38. Culina S, et al. Islet-reactive CD8+ T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. *Sci Immunol.* 2018;3(20):eaao4013.
39. Tickotsky N, et al. McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences. *Bioinformatics.* 2017;33(18):2924–2929.
40. Schindler MK, et al. Haploinsufficiency of immune checkpoint receptor CTLA4 induces a distinct neuroinflammatory disorder. *J Clin Invest.* 2020;130(10):5551–5561.
41. Shugay M, et al. VDJtools: unifying post-analysis of T cell receptor repertoires. *PLoS Comput Biol.* 2015;11(11):1004503.
42. Nazarov VI, et al. tcR: an R package for T cell receptor repertoire advanced data analysis. *BMC Bioinformatics.* 2015;16:175.
43. Crooks GE, et al. WebLogo: a sequence logo generator. *Genome Res.* 2004;14(6):1188–1190.