Methods to investigate intrathecal adaptive immunity in neurodegeneration

Hamilton Oh¹, Olivia Leventhal¹, Tony Wyss-Coray¹, Benoit Lehallier¹ and David Gate¹,*

¹Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, California, USA
*Corresponding author: dgate@stanford.edu

AUTHORS’ EMAIL ADDRESSES
Hamilton Oh: hoh3@stanford.edu
Olivia Leventhal: olivialeventhal@stanford.edu
Tony Wyss-Coray: twc@stanford.edu
Benoit Lehallier: lehallib@stanford.edu
David Gate: dgate@stanford.edu

ABSTRACT

Background: Cerebrospinal fluid (CSF) provides basic mechanical and immunological protection to the brain. Historically, analysis of CSF has focused on protein changes, yet recent studies have shed light on cellular alterations. Evidence now exists for involvement of intrathecal T cells in the pathobiology of neurodegenerative diseases. However, a standardized method for long-term preservation of CSF immune cells is lacking. Further, the functional role of CSF T cells and their cognate antigens in neurodegenerative diseases are largely unknown.

Results: We present a method for long-term cryopreservation of CSF immune cells for downstream single cell RNA and T cell receptor sequencing (scRNA-TCRseq) analysis. We observe preservation of CSF immune cells, consisting primarily of memory CD4⁺ and CD8⁺ T cells. We then utilize unbiased bioinformatics approaches to quantify TCR sequence similarity within and between disease groups. By this method, we identify clusters of disease-associated, antigen-specific TCRs from clonally expanded CSF T cells of patients with neurodegenerative diseases.

Conclusions: Here, we provide a standardized approach for long-term storage of CSF immune cells. Additionally, we present unbiased bioinformatic approaches that will facilitate the discovery of target antigens of clonally expanded T cells in neurodegenerative diseases. These novel methods will help improve our understanding of adaptive immunity in the central nervous system.

Keywords: cerebrospinal fluid cells, intrathecal cells, neurodegeneration, adaptive immunity, T cells, T cell receptor (TCR), antigen
BACKGROUND

The cerebrospinal fluid (CSF) provides insight into brain physiology of living individuals. Biochemical analysis of CSF is routinely utilized as a diagnostic tool in neurodegeneration (1-3). For example, in Alzheimer’s disease (AD), changes in protein levels of total tau, phosphorylated tau and amyloid-β are indicative of disease pathology. However, while CSF biomarkers guide diagnosis of patients with neurodegenerative diseases, the cells patrolling the interstitial fluid are often centrifuged and discarded. Only in cases of extreme central nervous system inflammation, such as meningitis or encephalitis, are CSF cells utilized as a diagnostic. Blood cells, on the other hand, are routinely used to assess health and disease. We surveyed the literature to gain insight into CSF cell composition and summarized the main studies (Table 1). The CSF is primarily composed of antigen-experienced memory CD4+ and CD8+ T cells, plasmacytoid dendritic cells, and CD56high natural killer (NK) cells. Naïve T cells, monocytes, granulocytes, myeloid dendritic cells, basophils, and B cells are not as abundant in the CSF as they are in peripheral blood (2). Moreover, the cellular density of CSF (1,000-3000 cells per mL) is drastically diluted compared to blood (millions of cells per mL) (1, 4, 5).

To our knowledge, only the multiple sclerosis field has rigorously evaluated live CSF cells to assess disease (1, 5-14). However, recent studies provide evidence of CSF T cells in the pathobiology of other neurodegenerative diseases. These include the discovery of hypocretin-specific CD8+ T cells in narcolepsy (15) and T cell clones shared between Rammusen encephalitis patients (16). Moreover, Sulzer et al. report the presence of α-synuclein specific T cells in Parkinson’s disease (PD), though they studied the peripheral blood, not CSF (17). Finally, our group reported the presence of α-synuclein specific T cells in the parenchyma of AD brains (18). Cumulatively, these advances provide evidence of T cell involvement in the pathobiology of several neurodegenerative diseases and emphasize the value of CSF cells to study the role of adaptive immunity in neurodegenerative disease.

Most of the aforementioned studies have relied on freshly isolated CSF cells (1, 5, 6, 9-20) or extraction of genomic DNA or RNA from frozen cells (7, 8). Analyzing fresh CSF cells provides the highest number of viable cells and the closest approximation of their endogenous physiology, but introduces batch effects and time restrictions. Conversely, cryopreservation allows for parallel analysis of multiple samples acquired longitudinally, minimizing batch effects and time restrictions. Increased evidence of CSF cell involvement in neurodegenerative disease warrants a standardized approach for the long-term preservation of CSF immune cells and methods to assess the role of CSF immune cells in pathobiology. Here, we report a method for the long-term storage and subsequent analysis of CSF cells by scRNA-TCRseq. Moreover, we present unbiased bioinformatics approaches to identify disease-associated TCRs. These methods will enable researchers to discover novel disease mechanisms in neurodegenerative diseases, particularly in the context of adaptive immunity.

RESULTS

We developed a standardized workflow for the isolation and cryopreservation of CSF immune cells for scRNA-TCRseq (Figure 1a). Following extraction of CSF by lumbar puncture, cells were pelleted by centrifugation, checked for blood contamination (Supplementary Figure 1a), and cryopreserved. Samples without blood contamination were thawed at 37°C and live cells were sorted by fluorescence activated cell sorting (FACS; Figure 1b), resulting in an average of 4,961 live cells per sample (Figure 1c) and 71% viability of singlets (Figure 1d). Length of storage had no impact on cell viability (Figure 1e).

Here, we sequenced CSF cells from a total of 24 subjects – 8 healthy controls, 4 patients with clinical AD, 5 with mild cognitive impairment (MCI), and 7 with PD (Table 2). After thawing and sorting live cells, we prepared single cell libraries then amplified the global transcriptome and TCRαβ genes from each sample. This resulted in an average of 1,323 sequenced cells per sample (Figure 1f). Clustering of all samples and marker gene expression analysis revealed separation of CD8+ and CD4+ T cells, NK cells, monocytes, dendritic cells (DCs), B cells, and plasma cells (Figure 2a-b). Cells with sequenced TCRs (Figure 2c) overlapped with cells that expressed the pan-T cell marker CD3D (Figure 2d), confirming T cell identity and productive TCR sequencing.
Importantly, we did not detect platelet genes, such as PPBP, among our CSF cells, confirming that our samples were not contaminated with blood. Additionally, clusters were not enriched for particular subjects, diagnosis, processing batch or sequencing batch (Supplementary Figure 2a-d). After clustering and annotating cells, we calculated the average proportion of each cell type in our CSF samples (Figure 2e). CD4+ T cells were the most abundant cell type (64%) followed by CD8+ T Cells (30%), myeloid cells (3.5%), NK cells (2%), plasma cells (0.5%) and B cells (0.3%) (Figure 2e). There was no significant difference in the cell type composition between disease groups (Supplementary Figure 2e). These results indicate that our method of long-term preservation of CSF cells retains several immune cell populations and is suitable for downstream scRNA-TCRseq analysis.

We next sought to develop methods to facilitate the discovery of TCR-antigen specificities relevant to neurodegenerative diseases. We first assessed whether clonally expanded T cells were present in the CSF, as determined by identification of cells with identical TCR sequences. Indeed, we identified numerous clonally expanded T cells (Figure 2f) and observed that larger clones (of size > 5) were enriched in CD8+ T cells compared to CD4+ T cells (Figure 2g-i).

Given the abundance of clonally expanded T cells in the CSF of patients with neurodegenerative disease, we reasoned that examining TCR similarity within and between disease groups could provide insight into disease-relevant TCR-antigen specificities. We first used a Levenshtein similarity (L-sim) score to quantify TCR similarity. This score is based on the Levenshtein distance algorithm, which calculates the number of edits, deletions, or insertions required to make two strings identical. L-sim includes a string length normalization and transformation of the final value to be between 0 and 1, with 1 representing identical TCR sequences. To examine TCR similarity within and between neurodegenerative diseases, we calculated the L-sim score for each possible pair of TCRs in our dataset. To increase the confidence of this analysis and to focus on disease-relevant TCRs, we filtered our dataset to only include complete, unambiguous TCRαβ sequences from clonally expanded T cells (Figure 3a). Using this unbiased method, we identified several clusters of similar TCR sequences (L-sim > 0.8) (Figure 3b). Interestingly, the largest clusters contained TCR sequences exclusively from patients with neurodegenerative diseases (Figure 3c) and were expressed by CD8+ T cells (Figure 3d). Moreover, three of the TCRβ sequences in Cluster 1 were identical (CASSLGQAYEQYF) and have been shown to be specific for Epstein-Barr virus (EBV) epitope EBNA3A (18).

To more broadly understand TCR similarity within and between disease groups, we developed a TCR node network that displays connections between similar TCRs (L-sim > 0.8) organized by samples and disease. This analysis revealed that AD and MCI patients have numerous similar TCR sequences (Figure 3e). Visualizing TCR similarity using a stricter L-sim cutoff of 0.9 (Figure 3f) or 1 (Supplementary Figure 3a) showed that highly similar TCRs in our dataset were almost exclusively shared within AD and MCI patients. This suggests that AD and MCI patients may have similar T cell antigen specificities in the CSF, indicative of a disease-specific adaptive immune response.

In addition to Levenshtein similarity, we show that motif analysis can be used to identify TCR sequences that share antigen specificity within and between disease groups. Using a sliding window approach, we quantified the frequency of every motif of every size from our pool of TCRβ sequences (Figure 3g). We then inspected antigen specificity of large, frequent motifs by searching the motifs in the McPAS-TCR specificity public database (21). This search uncovered motifs which were highly correlated with specific antigens, including the GATNEKL motif which was found only in patients with neurodegenerative disease (Figure 3h). This motif has been reported to be expressed by CD8+ T cells with specificity against cytomegalovirus (CMV; Figure 3i). In summary, we show that TCR analysis using Levenshtein similarity and motif analysis can be used to identify disease-associated TCRs as well as their antigen targets.

**DISCUSSION**

Most studies performed on human cells rely on peripheral blood mononuclear cells. However, utilizing peripheral cells as a read-out of immune changes in neurodegeneration limits our ability to understand central
immunity. CSF cells, on the other hand, provide a way to directly study immunology in the central nervous system. Indeed, recent studies of intrathecal immunity by us and others have shown this understudied immune compartment to be relevant to the pathobiology of neurodegenerative diseases (1, 5-20, 22). Yet, CSF cells have been widely understudied because of the invasive method of CSF extraction (typically lumbar puncture), and because cells are often discarded for proteomic analysis. Cells not discarded are typically analyzed fresh—generally by flow cytometry—which limits the scope of analysis.

Our simple and quick method allows for the preservation of an average 4,961 live cells after thawing, with no effect of storage length on cell viability. We find 64% of intrathecal cells are CD4⁺ T cells and 30% are CD8⁺ T cells, compared to 60-83% and 11-20%, respectively, reported by others (1, 20). We find that monocytes, dendritic cells, and NK cells combined make up 6% of CSF cells, whereas other studies report that monocytes make up 5-12%, dendritic cells less than 4%, and NK cells around 5% of fresh CSF cells (1). We also detected small proportions of plasma cells (0.3%) and B cells (0.5%), which aligns with the less than 1% of B cells previously reported in fresh CSF (22). The differences in CSF composition between this study and previous studies may be explained by the increased susceptibility to cell loss of myeloid cells compared to lymphocytes from freshly isolated CSF samples, (23) combined with the added stress of the freeze and thaw process. Overall, while we find a slight reduction in the number of recovered myeloid cells and NK cells, our method preserves T cells, the most abundant cell type in fresh CSF. Preservation of these cells enables the molecular study of adaptive immunity in neurodegenerative disease while minimizing batch effects and time restrictions.

Beyond introducing a method to store human CSF cells, we show that scRNA-TCRseq analysis can provide insight into the TCR repertoire and antigen specificities of clonally expanded T cells within and between neurodegenerative disease groups. We provide an easy-to-use and easily modifiable R script that can act as a base template for future studies. By quantifying the similarity of each TCR to every other TCR using the L-sim score, we show that TCRs from patients with neurodegenerative diseases, namely AD and MCI, cluster together. Additionally, we show that motif analysis can be used to identify TCRs that may have shared antigen specificity. Interestingly, both analyses independently revealed clusters of TCRs in patients with neurodegenerative diseases that were specific to viruses (i.e. EBV and CMV). However, these clusters represent only a small fraction of the TCR repertoire in the CSF. Moreover, our current knowledge of TCR antigen-specificity is largely limited to studies of infectious disease, biasing results to viral epitopes.

It will be important to continue to expand our knowledge of TCR antigen specificity in neurodegenerative diseases, especially in diseases where infection is not thought to be the primary cause. This is especially important given the growing evidence that T cells negatively contribute to the pathobiology of several neurodegenerative diseases, including AD and PD (17, 18). Using these methods to first identify candidate disease-specific TCRs, followed by the identification of their antigens, may reveal novel self/non-self antigen targets of clonally expanded T cells in neurodegenerative diseases.

CONCLUSIONS

In conclusion, we provide a detailed method for long-term preservation of CSF cells and subsequent analysis by scRNA-TCRseq to study clonally expanded T cells in neurodegenerative disease. Importantly, we uncover several disease-associated TCR clusters of unknown specificity. Although these TCRs are highly similar, additional studies will need to confirm their shared specificity for antigens. Future studies utilizing high throughput peptide screens, such as yeast surface display platforms, could be employed to identify novel antigens. Theoretically, these novel antigens could serve as novel therapeutic targets or biomarkers for neurodegenerative diseases. With increasing innovation and throughput in technologies that aid in the identification of TCR-antigen specificities, the mapping of clonally expanded TCRs will improve our understanding of the role of T cells in the pathobiology of neurodegenerative diseases. Ultimately, we expect these methods to aid in the development of novel adaptive immune-focused treatments for neurodegenerative disease.
METHODS

Study participants
Samples were acquired through the Stanford Brain Rejuvenation Program, the NIA funded Stanford Alzheimer’s Disease Research Center (ADRC), the University of California at San Francisco ADRC and the University of California at San Diego ADRC. Collection of CSF was approved by the Institutional Review Board of each university; written consent was obtained from all subjects. A total of 34 living subjects were used in this study, 24 of which were used for scRNA-TCRseq. The 24 subjects included 8 healthy controls, 4 patients with AD, 5 with MCI, and 7 with PD.

Cryopreservation of CSF cells
CSF was collected by lumbar puncture, then centrifuged at 300 rcf for 10 minutes at 4 °C to pellet immune cells. Importantly, CSF samples were checked for blood contamination by examining the pellet for the presence of red blood cells by eye. An example of a CSF sample contaminated with blood is shown in Supplemental Figure 1a. Note that cells should remain at 4 °C until they are further processed, but it is best to freeze the cells as quickly as possible to limit cell death. The supernatant (cell free CSF) was aliquoted, carefully leaving behind 100 μl of CSF with the pelleted cells. 100 μl of CSF was left so that cells were concentrated enough for counting and viability measurements. The pelleted cells were then gently resuspended in the 100 μl CSF and 10 μl of resuspended cells were then removed for counting. Importantly, cells were gently resuspended by first tapping the bottom of the tube and then gently triturating 10 times, making sure not to touch the pipette tip to the edge of the tube. Then 10 μl CSF was removed and mixed with 10 μl trypan blue to assess red blood cell content and viability. Cells were then visualized on a TC20 automated cell counter (BioRad) and cell number, viability and the presence or absence of red blood cells was recorded. CSF samples contaminated with blood were discarded. The resuspended cells were then mixed with 900 μl Recovery Cell Culture Freezing Medium (Thermo Fisher Scientific). This medium is an optimized version of the typical freezing medium, containing high-glucose Dulbecco's Modified Eagle Medium with 10% serum and 10% dimethyl sulfoxide. We utilized this medium because it is quality tested for pH, osmolality, sterility, and endotoxin and each lot is quality tested on CHO-K1 cells. The freezing medium was first thawed at 37 °C, aliquoted, and stored at -20 °C. Before use, the medium was thawed at 37 °C and kept on ice. After each aliquot was thawed, the freezing medium was stored at 4 °C for up to one month. All samples were frozen overnight at −80 °C in a Mr. Frosty freezing container (Thermo Fisher Scientific) and transferred the following day to liquid nitrogen for storage. CSF cells were stored in liquid nitrogen on average 266 days.

Preparation of frozen CSF cells for analysis
CSF cells were thawed at 37 °C in a water bath with the media submerged and the top of the tube out of the water. Cells were kept in the water bath for as little as possible and removed when the media was nearly completely thawed. The cells were then removed and gently pipetted into a 5 mL flow cytometry tube containing 3 mL pre-warmed (37 °C) sorting buffer (PBS with 0.04% bovine serum albumin (BSA)). The tube was then rinsed once with the sorting buffer and placed into the same flow cytometry tube. Cells were then centrifuged at 350 rcf for ten minutes. The supernatant was removed, and cells were resuspended in 500 μl sorting buffer. ½ μl of Sytox Red (Thermo Fisher Scientific) was added to the sample immediately before sorting by flow cytometry. Live cells were sorted into 1.5 mL Eppendorf tubes containing 750 μL sorting buffer. Once all the samples were sorted, cells were spun at 350 rcf at 4 °C in a spinning bucket centrifuge for 7 minutes. The supernatant was then removed, leaving behind 10 μl. 10 μl was left behind to resuspend the CSF cells and load the entire volume for droplet scRNA-seq.

Drop-seq of CSF cells
Chromium Single Cell 5’ Library & Gel Bead kit, Chromium Single Cell 5’ Library construction kit, Chromium Single Cell A Chip Kit, and Chromium i7 Multiplex kit (10X Genomics) were used for scRNAseq of CSF cells. We followed 10x Genomic’s User Guide for library construction. The only change we made to their protocol was in Step 1, GEM Generation & Barcoding. The user guide recommends loading a certain volume of cell suspension
stock depending on the concentration of the cell suspension and the user’s desired cell recovery. However, because CSF contains such low cell numbers, we loaded all the cells that were resuspended in 10 μl of sorting buffer. We then added the 10 μl of cell suspension and 21.7 μl of nuclease free water, which results in the same total volume of cell suspension/water that the user guide recommends. After library construction, libraries were sequenced by Novogone on a Novoseq S4 sequencer and FASTQ files were generated by Novogone. Cell Ranger v.3.0.2 was used to generate gene-expression matrices for CSF cells. Reads from the 10X v.2 5’ paired library were mapped to the human genome build GRCh38 3.0.0. The 5’ gene-expression libraries were then analyzed with the Cell Ranger count pipeline and the resulting expression matrix was used for further analysis in the Seurat package v.3.0.

**Seurat clustering of CSF cells**

Individual sample expression matrices were loaded into R using the function Read10x under the Matrix package v.1.2-15. The expression matrix for each sample was merged into one Seurat object using the CreateSeuratObject and MergeSeurat functions. The Seurat package v.3.0 was used for filtering, variable gene selection, normalization, scaling, dimensionality reduction, clustering and visualization. Genes were excluded if they were expressed in fewer than 10 cells and cells were excluded if they expressed fewer than 200 genes. Cells that expressed more than 1,600 genes, more than 6,000 UMIs and more than 10% mitochondrial genes were excluded from the analysis. The sctransform normalization method was used to normalize, scale, select variable genes and regress out sequencing and experimental batch, mitochondrial mapping percentage, number of UMIs, and number of genes. After filtering and normalization, there were 26,797 cells and 14,953 genes. Following PCA, 5 principle components were selected for clustering tSNE dimensionality reduction.

**Calculation of Levenshtein similarities**

Levenshtein similarity (L-sim) scores were calculated using the levenshteinSim function in the RecordLinkage package for R (24). L-sim calculation incorporates the Levenshtein distance algorithm, which quantifies the number of edits, deletions, or insertions required to make two strings identical. L-sim includes a string length normalization and transformation of the final value to be between 0 and 1, with 1 representing identical TCR sequences:

\[
\text{Levenshtein Similarity} = 1 - \frac{\text{lev}_{a,b}(i,j)}{\max(A,B)}
\]

\[
\text{Levenshtein Distance} = \begin{cases} 
\max(i,j) & \text{if } \min(i,j) = 0, \\
\min(\text{lev}_{a,b}(i-1,j) + 1) & \text{otherwise.}
\end{cases}
\]

**TCR network analysis**

TCR networks that show connections between similar TCRs organized by patient IDs and diagnosis groups were generated using the qgraph function in the qgraph package for R (25). Analysis from Figure 3b-d include only clonal TCRs with unambiguous alpha and beta chains. Supplementary Figure 3a includes only clonal TCRs with unambiguous beta chains. Unweighted networks were generated with all subjects and split per diagnosis group.

**TCR Motif analysis**

A custom script was used to identify motifs in our TCR beta chains. Only clonal TCRs with unambiguous beta chains were included for analysis. Identified motifs were searched in the McPAS-TCR database, a manually curated database of TCR sequences found to be associated with pathological conditions in mice and humans (21).
LIST OF ABBREVIATIONS
T cell receptor = TCR; CSF = cerebrospinal fluid; scRNA-TCRseq = single cell RNA-TCR sequencing; Alzheimer’s disease = AD; Mild cognitive impairment = MCI; Parkinson’s disease = PD; Healthy control = HC; Natural killer = NK; Dendritic cells = DCs; Levenshtein similarity = L-sim; Epstein-Barr virus = EBV; Cytomegalovirus = CMV.

DECLARATIONS

Ethics approval and consent to participate
Collection of CSF was approved by the Institutional Review Board of each university, and written consent was obtained from all subjects

Consent for publication
Not applicable

Availability of data and materials
The scRNAseq datasets analyzed during the current study are available in the Gene Expression Omnibus (GEO) repository under accession number GSE134578. The fresh cell counts and flow cytometry data generated in the current study are available from the corresponding author upon request. The R scripts for TCR similarity analysis are available on github: https://github.com/hamiltonoh/TCR_similarity_analysis.git

Competing interests
The authors declare that they have no competing interests.

Funding
This work was supported by a National Institute of Neurologic Disease and Stroke K99/R00 Pathway to Independence Award (NS112458-01A1) (D.G.), an Irene Diamond Fund/AFPAR Postdoctoral Transition Award in Aging (D.G.), a National Institutes of Health National Institute on Aging (NIA) F32 Fellowship (AG055255-01A1) (D.G.), and the NIA funded Stanford Alzheimer’s Disease Research Center P50AG047366.

Author’s contributions
D.G. and O.L. designed experiments. H.O., O.L. and D.G. prepared the manuscript. D.G. conducted flow cytometry analysis and sorting. H.O. and O.L. analyzed scRNAseq data. B.L. designed and performed bioinformatics analysis. All authors edited, read and approved the final manuscript.

Acknowledgements
We thank V. Henderson and the entire Stanford Alzheimer's disease Research Center team; G. Kerchner and S. Sha for CSF collection and D. Channappa for CSF processing.
REFERENCES

1. Han S, Lin YC, Wu T, Salgado AD, Mexhitaj I, Wuest SC, et al. Comprehensive immunophenotyping of cerebrospinal fluid cells in patients with neuroimmunological diseases. J Immunol. 2014;192(6):2551-63.

2. Perrin RJ, Fagan AM, Holtzman DM. Multimodal techniques for diagnosis and prognosis of Alzheimer's disease. Nature. 2009;461(7266):916-22.

3. Blennow K, Zetterberg H. The Past and the Future of Alzheimer's Disease Fluid Biomarkers. J Alzheimers Dis. 2018;62(3):1125-40.

4. Ransohoff RM, Engelhardt B. The anatomical and cellular basis of immune surveillance in the central nervous system. Nat Rev Immunol. 2012;12(9):623-35.

5. Schafflick D, Xu CA, Hartlehnert M, Cole M, Schulte-Mecklenbeck A, Lautwein T, et al. Integrated single cell analysis of blood and cerebrospinal fluid leukocytes in multiple sclerosis. Nat Commun. 2020;11(1):247.

6. Aly L, Yousef S, Schippling S, Jelcic I, Breiden P, Matschke J, et al. Central role of JC virus-specific CD4+ lymphocytes in progressive multi-focal leucoencephalopathy-immune reconstitution inflammatory syndrome. Brain. 2011;134(Pt 9):2687-702.

7. Lossius A, Johansen JN, Vartdal F, Robins H, Jurate Saltyte B, Holmoy T, et al. High-throughput sequencing of TCR repertoires in multiple sclerosis reveals intrathecal enrichment of EBV-reactive CD8+ T cells. Eur J Immunol. 2014;44(11):3439-52.

8. Salou M, Garcia A, Michel L, Gainche-Salmon A, Loussouarn D, Nicol B, et al. Expanded CD8 T-cell sharing between periphery and CNS in multiple sclerosis. Ann Clin Transl Neurol. 2015;2(6):609-22.

9. Johansen JN, Vartdal F, Desmarais C, Tutturen AE, de Souza GA, Lossius A, et al. Intrathecal BCR transcriptome in multiple sclerosis versus other neuroinflammation: Equally diverse and compartmentalized, but more mutated, biased and overlapping with the proteome. Clin Immunol. 2015;160(2):211-25.

10. Planas R, Metz I, Ortiz Y, Vilarrasa N, Jelcic I, Salinas-Riester G, et al. Central role of Th2/Tc2 lymphocytes in pattern II multiple sclerosis lesions. Ann Clin Transl Neurol. 2015;2(9):875-93.

11. Hoglund RA, Lossius A, Johansen JN, Homan J, Benth JS, Robins H, et al. In Silico Prediction Analysis of Idiotope-Driven T-B Cell Collaboration in Multiple Sclerosis. Front Immunol. 2017;8:1255.

12. Rathbone E, Durant L, Kinsella J, Parker AR, Hassan-Smith G, Douglas MR, et al. Cerebrospinal fluid immunoglobulin light chain ratios predict disease progression in multiple sclerosis. J Neurol Neurosurg Psychiatry. 2018;89(10):1044-9.

13. Beltran E, Gerdes LA, Hansen J, Flierl-Hecht A, Krebs S, Blum H, et al. Early adaptive immune activation detected in monozygotic twins with prodromal multiple sclerosis. J Clin Invest. 2019;129(11):4758-68.

14. Tomescu-Baci u A, Johansen JN, Holmoy T, Greiff V, Stensland M, de Souza GA, et al. Persistence of intrathecal oligoclonal B cells and IgG in multiple sclerosis. J Neuroimmunol. 2019;333:576966.

15. Latorre D, Kallweit U, Armentani E, Foglierini M, Mele F, Cassotta A, et al. T cells in patients with narcolepsy target self-antigens of hypocretin neurons. Nature. 2018;562(7725):63-8.

16. Schneider-Hohendorf T, Mohan H, Bien CG, Breuer J, Becker A, Gorlich D, et al. CD8(+) T-cell pathogenicity in Rasmussen encephalitis elucidated by large-scale T-cell receptor sequencing. Nat Commun. 2016;7:11153.

17. Sulzer D, Alcalay RN, Garretti F, Cote L, Kanter E, Agin-Liebes J, et al. T cells from patients with Parkinson’s disease recognize α-synuclein peptides. Nature. 2017;546(7660):656-61.

18. Gate D, Saligrama N, Leventhal O, Yang AC, Unger MS, Middeldorp J, et al. Clonally expanded CD8 T cells patrol the cerebrospinal fluid in Alzheimer's disease. Nature. 2020;577(7790):399-404.

19. Albert ML, Austin LM, Darnell RB. Detection and treatment of activated T cells in the cerebrospinal fluid of patients with paraneoplastic cerebellar degeneration. Ann Neurol. 2000;47(1):9-17.

20. Hummert MW, Alvermann S, Gingele S, Gross CC, Wiendl H, Mirenska A, et al. Immunophenotyping of cerebrospinal fluid cells by Chipcytometry. J Neuroinflammation. 2018;15(1):160.

21. Tickot sny N, Sagiv T, Prilusky J, Shifrut E, Friedman N. McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequencing. Bioinformatics. 2017;33(18):2924-9.
22. Kivisakk P, Mahad DJ, Callahan MK, Trebst C, Tucky B, Wei T, et al. Human cerebrospinal fluid central memory CD4+ T cells: evidence for trafficking through choroid plexus and meninges via P-selectin. Proc Natl Acad Sci U S A. 2003;100(14):8389-94.

23. de Graaf MT, de Jongste AH, Kraan J, Boonstra JG, Sillevis Smitt PA, Gratama JW. Flow cytometric characterization of cerebrospinal fluid cells. Cytometry B Clin Cytom. 2011;80(5):271-81.

24. Sariyar M, Borg A. The RecordLinkage Package: Detecting Errors in Data. The R Journal. 2010;2(2).

25. Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D. qgraph: Network Visualizations of Relationships in Psychometric Data. Journal of Statistical Software; Vol 1, Issue 4 (2012). 2012.

**FIGURE LEGENDS**

**Figure 1. Cryopreservation of human cerebrospinal fluid cells retains cellular viability.**

a) Workflow for preserving CSF cells for scRNA-TCRseq. b) Gating strategy for sorting live cryopreserved cells by flow cytometry. c) Quantification of live cell count (live singlets) from flow cytometry sorting. Mean ± s.e.m. d) Quantification of viability (percent of live singlets out of all singlets) from sorting by flow cytometry. Mean ± s.e.m. e) Linear regression of cell viability versus length of storage. f) Quantification of the number of CSF cells captured for sequencing.

**Figure 2. scRNA-TCRseq of human cerebrospinal fluid reveals clonally expanded CD8+ T cells.**

a) Seurat clustering of 24 CSF samples displayed on tSNE. b) Heatmap of cluster marker genes. c) Cells with detected TCR displayed on tSNE. d) CD3D expression displayed on tSNE. e) Quantification of average cell type distribution based on Seurat clustering. f) Clones – cells with TCR sequences shared with other cells – displayed on tSNE. Only cells with detected TCRs are included. g) Clones of different sizes displayed on tSNE. h) Quantification of number of T cell types per clone size. Only cells with detected TCR included. GNLY+ CD8 T Cells and CD8 T cells were combined as CD8+ T cells. i) Quantification of % T cell types per clone size.

**Figure 3. Analysis of T cell receptor sequence similarity within and between neurodegenerative disease groups.**

a) Workflow for quantifying TCRαβ similarity. Clonal TCRs with unambiguous CDR3a and CDR3b sequences were retained for analysis b) At left: All versus all heatmap highlighting TCRs with a L-sim score > 0.8. At right: Inset of heatmap to show three clusters of similar TCRs. Bright yellow indicates high L-sim score. c) Metadata of clustered TCRs shown in b). d) Three clusters of similar TCRs identified in c) highlighted on tSNE. Clusters of TCRs overlap with CD8 T cells. e) TCR network displaying connections between samples with similar sequences (L-sim > 0.8) identified in b). Each node is a unique patient sample with each small circle sprouting off a node representing a unique clonal TCR. Network is organized by disease group. f) TCR network (L-sim > 0.9) g) Quantification of shared motifs present in TCRβ sequences. Clonal TCRs with unambiguous CDR3b sequences were retained for analysis. h) Table of motifs of length 7. GATNEKL motif highlighted and metadata table for motif shown. i) Table of pathogens specific to GATNEKL motif based on search in public McPAS-TCR database. Table of GATNEKL motif containing T cell types based on search in public McPAS-TCR database.

**Supplementary Figure 1. Identification of blood contamination in CSF**

a) Representative CSF pellet with visible blood contamination. Samples contaminated with blood should be discarded for molecular analysis of CSF cells.

**Supplementary Figure 2. Visualization of sample metadata.**

a) tSNE colored by sample. b) tSNE colored by diagnosis. c) tSNE colored by processing batch. d) tSNE colored by sequencing batch. e) Quantification of average cell type distribution per diagnosis based on Seurat clustering.

**Supplementary Figure 3. Networks of all samples showing connections between patients with similar TCRs.**
a) Network displaying connections between samples with similar TCRαβ sequences (Levenshtein Similarity > 0.9). b) Network displaying connections between samples with identical TCRβ sequences (Levenshtein Similarity = 1.0). Network includes clonal TCRs with unambiguous CDR3b sequences.
Figure 1

Isolate CSF by lumbar puncture
Pellet CSF cells and count
Cryopreserve cells
Thaw and sort live cells
10X Single cell mRNA and TCR barcode
Sequence

Gating Strategy for cryopreserved CSF cells

% Live Singlets vs Length of Storage

Number of Cells Sequenced
Blood Contamination
a

Identical TCRβ
(Levenshtein Similarity = 1)
Table 1

| Author et al. | Year | Journal | Title | Cell preservation method | Method of analysis | Disorder studied |
|--------------|------|---------|------|--------------------------|--------------------|-----------------|
| Lueg et al.  | 2015 | Neurobiology of Aging | Clinical relevance of specific T-cell activation in the blood and cerebrospinal fluid of patients with mild Alzheimer’s disease | CSF cells processed within 20 minutes | Flow cytometry | AD & MCI |
| Tomescu-Baciu et al. | 2019 | Journal of Neuroimmunology | Persistence of intrathecal oligoclonal B cells and IgG in multiple sclerosis | CSF cells processed immediately | Quantitative mass spectrometry of oligoclonal bands and immunoglobulin heavy-chain from cDNA sequencing | B cells and IgG in MS patients at different time points |
| Giunti et al. | 2003 | Journal of Leukocyte Biology | Phenotypic and functional analysis of T cells homing into the CSF of subjects with inflammatory diseases of the CNS | Unclear | Flow cytometry & culture of T cell clones from CSF (maintained by PGA stimulation) | Inflammatory neurological diseases (including MS) |
| Hauser et al. | 1983 | Neurology | CSF cells in multiple sclerosis Monoclonal antibody analysis and relationship to peripheral blood T-cell subsets | Fresh CSF cells processed within 4 hours of lumbar puncture | Fluorescence microscopy | MS |
| Calebreti et al. | 1998 | Journal of Neuroimmunology | Cytokine gene expression in cells derived from CSF of multiple sclerosis patients | CSF cells processed immediately | RNA extraction + RT-PCR | MS |
| Aly et al. | 2011 | Brain: A Journal of Neurology | Central role of JC virus-specific CD4+ lymphocytes in progressive multi-focal leukoencephalopathy-immune reconstitution inflammatory syndrome | PHA expanded CSF cells | Proliferation assay (3H-thymidine incorporation) of cells stimulated with peptides | MS |
| Lossius et al. | 2014 | European Journal of Immunology | High-throughput sequencing of TCR repertoires in multiple sclerosis reveals intrathelial enrichment of EBV-reactive CD8+ T cells | Pellets sampled in liquid nitrogen & genomic DNA extracted | TCRβ sequencing | MS |
| Salou et al. | 2015 | Annals of Clinical and Translational Neurology | Expanded CUB T-cell sharing between periphery and CNS in multiple sclerosis | CSF cells stored frozen in Tissue & total RNA extracted (CSF obtained from oisterna magna) | TCRβ sequencing | MS |
| Johansen et al. | 2015 | Clinical Immunology | Intrathelial BCR transcriptome in multiple sclerosis versus other neuroinflammation: Equally diverse and compartmentalized, but more mutated, biased and overlapping with the proteome | CSF cells processed within 20 minutes | Sequencing of IGHV and mass spectrometry | MS |
| Planas et al. | 2015 | Annals of Clinical and Translational Neurology | Central role of Th2/Tc2 lymphocytes in patients with multiple sclerosis lesions | CSF cells expanded in culture for 2 weeks | CSF cells expanded in vitro and then FACs sorted based on V β | MS |
| Hoglund et al. | 2017 | Frontiers in Immunology | In Silico Prediction Analysis of Idiotope-Driven T-B Cell Collaboration in Multiple Sclerosis | CSF cells processed immediately | Sequencing IGHV-VDJ region with Adaptive | MS |
| Rathbone et al. | 2018 | Journal of Neuroimmunology | Cerebrospinal fluid immunoglobulin light chain ratios predict disease progression in multiple sclerosis | CSF cells processed immediately | Flow cytometry & sequencing of IGHV and IGKV with adaptive | MS |
| Schafflick et al. | 2020 | Nature Communications | Integrated single-cell analysis of blood and cerebrospinal fluid leukocytes in multiple sclerosis | Processed and kept at 4 °C | Single-cell RNA sequencing (10x) | MS |
| Hummert et al. | 2018 | Journal of Neuroimmunology | Immunophenotyping of cerebrospinal fluid cells by Chipcytometry | Fresh CSF cells fixed with 4% phosphate-buffered formaldehyde | Chipcytometry | MS/CIS, OND (infectious & autoimmune), tumor, NIND, NND, unspeciﬁed MS-discordant monozygotic twin pairs |
| Beltran et al. | 2019 | The Journal of Clinical Investigation | Early adaptive immune activation detected in mononuclear cells with prodromal multiple sclerosis | Fresh CSF and blood samples processed within 1 hour after collection | Single-cell RNA sequencing (Smart-Seq2) | MS |
| Latrre et al. | 2016 | Nature | T cells in patients with narcolepsy target self-antigens of hypocretin neurons | CSF cells expanded in culture within a few hours of sampling (15 days in culture) | Polyclonal expansion of CSF cells followed by TCR Vβ sequencing | Narcolepsy |
| Kivstak et al. | 2003 | NNAS | Human cerebrospinal fluid central memory CD4+ T cells: Evidence for trafficking through choroid plexus and meninges via P-selectin | Fresh CSF cells processed within 20 minutes of sampling | Flow cytometry | Noninflammatory neurological diseases |
| Albert et al. | 2001 | Annals of Neurology | Detection and treatment of activated T cells in the cerebrospinal fluid of patients with paranaoplastic cerebellar degeneration | Unclear | Flow cytometry | Paraneoplastic cerebellar degeneration |
| Schneider-Hohenfert et al. | 2016 | Nature Communications | CDB+ T-cell pathogenicity in Rasmussen encephalitis elucidated by large-scale T-cell receptor sequencing | Unclear | TCR sequencing with Adaptive | Rasmussen encephalitis |
| Han et al. | 2014 | Journal of Immunology | Comprehensive Immunophenotyping of Cerebrospinal Fluid Cells in Patients with Neuroinmunomologous Diseases | CSF cells processed immediately | Flow cytometry | Neuroinmunomolgical disorders |
| Lanz et al. | 2019 | Frontiers in Immunology | Single-Cell High-Throughput Technologies in Cerebrospinal Fluid Research and Diagnosis | Review | NA | NA |
| Glanz et al. | 1998 | Cancer | Cerebrospinal fluid cytology in patients with cancer: minimizing false- negative results | Fresh cells or kept at 4°C for 48 hours | Flow cytometry | Cancer |
| Pittman et al. | 2013 | Arch Pathol Lab Med | Utility of flow cytometry of cerebrospinal fluid as a screening tool in the diagnosis of central nervous system lymphoma | Unclear | Flow cytometry | CNS lymphoma |
| Svenningsson | 1995 | Neuroimmunology | Lymphocyte phenotype and subset distribution in normal cerebrospinal fluid. | Cells analyzed fresh and fixed before flow analysis | Flow cytometry | Healthy |
| de Graaf et al. | 2011 | Cytometry B Clin Cytom | Central memory CD4+ T cells dominate the normal cerebrospinal fluid. | Fresh CSF cells (within 6 hours of lumbar puncture) | Flow cytometry | Neurological diseases |
| Enose-Akahata | 2018 | PLoS Pathog | Immunophenotypic characterization of CSF B cells in virus-associated neuroinflammatory diseases | Unclear | Flow cytometry | Virus-associated neuroinflammatory disease |
| Gross et al. | 2016 | Proc Natl Acad Sci USA | Impaired NK-mediated regulation of T-cell activity in multiple sclerosis is reconstituted by IL-2 receptor modulation | Fresh CSF cells processed within 20 minutes of sampling | Flow cytometry | MS |
| Author et al. | Year | Journal | Methodology | Tissue | Notes |
|--------------|------|---------|-------------|--------|-------|
| Kraan et al. | 2008 | Curr Protoc Cytom | Flow cytometric immunophenotyping of cerebrospinal fluid. | fresh CSF cells | Flow cytometry |
| Cepok et al. | 2005 | Brain | Short-lived plasma blasts are the main B cell effector subset during the course of multiple sclerosis. | fresh CSF cells | Flow cytometry |
| Schirmer et al. | 2015 | Open Forum infectious disease | Extensive recruitment of plasma blasts to the cerebrospinal fluid in toscana virus encephalitis. | fresh CSF cells | Flow cytometry, microscopic cell-counting and May-Grünwald cell stain |
| Oreja-Guevara | 1998 | Acta Neurol Scand | Analysis of lymphocyte subpopulations in cerebrospinal fluid and peripheral blood in patients with multiple sclerosis and inflammatory diseases of the nervous system. | fresh CSF cells | Flow cytometry, MS & nervous system inflammatory diseases |
| Kowarik et al. | 2014 | J Neurology | Immune cell subtyping in the cerebrospinal fluid of patients with neurological diseases. | fresh CSF cells | Flow cytometry, neurological diseases |
| Brucklacher-Waldert | 2009 | Brain | Phenotypical and functional characterization of T helper 17 cells in multiple sclerosis. | freshly used and expanded in culture by PHA stimulation | TCRβ sequencing of in vitro clones & flow cytometry |
| Huang et al. | 2009 | Annals of Neurology | Characterization of natural killer cells in paired CSF and blood samples during neuroinflammation. | freshly used and put into culture | Flow cytometry |
| Schirmer et al. | 2015 | Open Forum Infectious disease | Intrathecal T-cell clonal expansions in patients with multiple sclerosis. | freshly used and expanded in culture by PHA stimulation | Flow cytometry, inflammatory and non-inflammatory neurological diseases |
| Cameron et al. | 2019 | J Neuroimmunology | Immunoglobulin class-switched B cells form an active immune axis between CNS and periphery in multiple sclerosis. | fresh CSF cells | Flow cytometry |
| Pranzatelli et al. | 2018 | Clin Exp Immunol | Cerebrospinal fluid gammadelta T cells in children with inflammatory and non-inflammatory neurological disorders. | fresh CSF cells | Flow cytometry |
| Harmer et al. | 2015 | Clin Exp Immunol | High interindividual variability in the CD4/CD8 T cell ratio and natalizumab concentration levels in the cerebrospinal fluid of patients with multiple sclerosis. | fresh CSF cells | Flow cytometry |
| Schulte-Meckenbeck | 2019 | Eur J Neurol | Immunophenotyping of cerebrospinal fluid cells in ischaemic stroke. | CSF cells processed within 20 minutes | Flow cytometry, ischemic stroke |
| Ligocki | 2015 | ASN Neuro | A distinct class of antibodies may be an indicator of gray matter autoimmunity in early and established relapsing remitting multiple sclerosis patients. | no mention of storing cells so most likely analyzed fresh | single-cell sorting of CD19+ B cells and subsequent VH & VL sequencing |
| Kowarik et al. | 2014 | J Neurology | Immune cell subtyping in the cerebrospinal fluid of patients with neurological diseases. | fresh CSF cells | Flow cytometry, neurological diseases |
| Cepok et al. | 2007 | Brain | Patterns of cerebrospinal fluid pathology correlate with disease progression in multiple sclerosis. | fresh CSF cells | Flow cytometry |
| Hammerschmidt | 2015 | Sci Trans Med | Enriched CD16 high CCR4+ gammadelta T cells in the cerebrospinal fluid of patients with multiple sclerosis. | fresh CSF cells | Flow cytometry |
| Palanichamy | 2014 | Sci Trans Med | Immunoglobulin class-switched B cells form an active immune axis between CNS and periphery in multiple sclerosis. | analyzed cells fresh (lysed them immediately to extract RNA) | Ig-VH sequencing from isolated RNA |
| Stern et al. | 2014 | Sci Trans Med | B cells populating the multiple sclerosis brain mature in the draining cervical lymph nodes. | stored at -80°C | Immunoglobulin (Ig) variable region heavy (VH) and light chain (VL) sequencing from extracted RNA |
| Lovato et al. | 2011 | Brain | Related B cell clones populate the meninges and parenchyma of patients with multiple sclerosis. | stored at -80°C | Ig variable region heavy chain sequencing from extracted RNA |
| Obermeier et al. | 2008 | Nature Medicine | Matching of oligodendrocytoma immunoglobulin transcripts and proteomes of cerebrospinal fluid in multiple sclerosis. | unclear, but only extracted RNA from cells | Ig-G, Ig-A and Ig-X sequencing from extracted RNA |
| Cameron et al. | 2009 | J Neuroimmunology | Phenotypical and functional characterization of T helper 17 cells in multiple sclerosis. | fresh CSF cells | Flow cytometry |
| von Budingen | 2012 | J Clin Invest | B cell exchange across the blood-brain barrier in multiple sclerosis. | stored at -80°C | Deep repertoire sequencing of IgG heavy chain variable region genes (IgG-VH) from extracted RNA |
| de Paula Alves Sousa | 2016 | Annals of Clinical and Translational Neurology | Intrathecal T-cell clonal expansions in patients with multiple sclerosis. | stored at -80°C | TCRβ sequencing from extracted RNA |
| Planas et al. | 2018 | Frontiers in Immunology | Detailed characterization of T cell receptor repertoires in multiple sclerosis brain lesions. | stored at -80°C | T cell receptor β-chain variable gene (TRBV) sequencing (5′-3′) of genomic (gDNA) with PCR amplification and next generation deep immune repertoire sequencing of immunoglobulin (Ig) heavy chain variable regions (VH) |
| Greenfield | 2019 | JCI Insight | The active intrathecal B cell response in LGI1 antibody encephalitis. | not described | PLCR amplification and next generation deep immune repertoire sequencing of immunoglobulin (Ig) heavy chain variable regions (VH) |
| Rans et al. | 2015 | Lancet | Longitudinally persistent cerebrospinal fluid B cells can resist treatment in multiple sclerosis. | stored 7 of 20 CSF samples and 1 of 20 PBMC samples at -80°C as unsorted cell pellets and the remaining were analyzed fresh | Ig heavy chain variable region repertoire sequencing on B cells (flow sorted & bulk) |

**Legend:**
- **CSF:** Cerebrospinal fluid
- **PBMC:** Peripheral blood mononuclear cells
- **RNA:** RNA extraction
- **gDNA:** Genomic DNA
- **PHO:** Phytohemagglutinin
- **IgG:** Immunoglobulin G
- **VH:** Variable heavy chain
- **VL:** Variable light chain
| Kreye et al. 2016 | Brain | Human cerebrospinal fluid monoclonal N-methyl-D-aspartate receptor autoantibodies are sufficient for encephalitis pathogenesis. | Cells resuspended in 500 µl freezing medium [45% RPMI, 45% foetal calf serum, 10% dimethylsulphoxide (DMSO)] and stored at −80°C | Flow cytometric single cell sorting → sequencing of IGH, IHL, IGK | Encephalitis |
|-----------------|-------|---------------------------------------------------------------|-----------------------------------------------------------------|----------------------------------------------------------------|----------------|
| Farhadian 2018 | UCI Insight | Single-cell RNA sequencing reveals microglia-like cells in cerebrospinal fluid during virologically suppressed HIV | Fresh CSF and blood | scRNA-seq | HIV |
| Jordao et al. 2019 | Science | Single-cell profiling identifies myeloid cell subsets with distinct fates during neuroinflammation. | Did not isolate CSF Cells | scRNA-seq | EAE (mouse) |
| SAMPLE | AGE | SEX | CONDITION | PROCESSING BATCH | SEQUENCING BATCH |
|--------|-----|-----|-----------|------------------|------------------|
| CSF01  | 88  | M   | HC        | 1                | 1                |
| CSF02  | 77  | M   | AD        | 1                | 1                |
| CSF03  | 71  | F   | HC        | 1                | 1                |
| CSF04  | 74  | M   | HC        | 1                | 1                |
| CSF05  | 72  | M   | MCI       | 1                | 1                |
| CSF06  | 64  | F   | AD        | 1                | 1                |
| CSF07  | 57  | M   | PD        | 1                | 1                |
| CSF08  | 74  | M   | HC        | 1                | 1                |
| CSF09  | 83  | F   | AD        | 2                | 2                |
| CSF10  | 66  | F   | PD        | 2                | 2                |
| CSF11  | 66  | M   | MCI       | 2                | 2                |
| CSF12  | 73  | F   | HC        | 2                | 2                |
| CSF13  | 76  | F   | HC        | 2                | 2                |
| CSF14  | 65  | F   | PD        | 2                | 2                |
| CSF15  | 48  | M   | MCI       | 2                | 2                |
| CSF16  | 67  | F   | HC        | 2                | 2                |
| CSF17  | 78  | M   | MCI       | 3                | 2                |
| CSF18  | 58  | F   | AD        | 3                | 2                |
| CSF19  | 83  | F   | MCI       | 3                | 2                |
| CSF20  | 66  | M   | HC        | 3                | 2                |
| CSF21  | 63  | M   | PD        | 3                | 2                |
| CSF22  | 77  | M   | PD        | 3                | 2                |
| CSF23  | 69  | M   | HC        | 3                | 2                |
| CSF24  | 62  | F   | PD        | 3                | 2                |