nature portfolio

Corresponding author(s): William H. Robinson

Last updated by author(s): Jan 5, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	BD FACSDiva (v.8.0) was used for collection of flow cytometry data. BLI analysis software (Fortebio/Sartorius, version 7.1) was used for collection of biolayer-interferometry data. GenePix Pro (ArrayIt, version 3.0) was used for collection of microarray data. SoftMax Pro (Molecular Devices) was used for collection of ELISA data. BLI analysis software (Fortebio/Sartorius, version 7.1) was used for collection of biolayer-interferometry data. Zeiss ZEN (blue edition) software was used for fluorescent immunostaining data collection.
Data analysis	FlowJo Version 10.7.1 (BD) and R version 3.6.1 was used to evaluate flow cytometry data. GraphPad Prism 9.1.0 was used for statistical analyses. BLI analysis software (Fortebio/Sartorius, version 7.1) were used for interpretation of biolayer-interferometry data. For single cell repertoire analysis, Illumina software was used to de-multiplex the raw reads, and R version 3.6.1 and Python scripts were used for downstream analyses, and IMGT HighV-QUEST v1.3.1 for sequence alignment to immunoglobulin databases. For phylogenetic analysis, we used Muscle (reference 51), and clustered the sequences with PhyML52 using maximum-likelihood clustering. Phylogenetic trees were drawn in Python using the ETE 3 toolkit (reference 53). GenePix Pro 3.0 software was used for microarray analysis (Molecular Devices). The PONDR algorithm was used for prediction of protein disorder. Software for structural analysis: XDS/aimless, Staraniso, Phenix/Phaser/Phenix.refine, Coot version 0.8.1, Pymol version 2.4.0. GraphPad Prism version 8.4.1 and R version 3.6.1 were used for statistical analyses. Morpheus software (The Broad Institute, https://software.broadinstitute.org/morpheus) was used for Heatmap generation. The mass spectroscopy data was analyzed using the search engine Byonic (reference 54). Prediction of Protein Disorder were analyzed with PONDR (Predictor of Natural Disordered Regions, WSU Research Foundation) using the VSL2 algorithm (reference 59).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The genomic datasets analyzed during the study have been uploaded to SRA, Accession #: PRJNA780931. Mass spectrometry data is available at www.massive.ucsd.edu, Accession #: MSV000086829. Structural data is available at www.rcsb.org PDB ID: 7K7R. Requests for data and materials should be addressed to tlanz@stanford.edu or wrobins@stanford.edu.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Our single-cell repertoire study includes n=9 individuals, yielding a total of 13,578 paired sequences from blood and 1,689 from CSF B cells, representing one of the largest single-cell repertoire data set from CSF B cells. A limiting factor is the availability of CSF and low abundance of B cells in CSF samples from MS patients. The study size is in line with prior repertoire studies investigating both, single-cell and bulk B cell repertoires in MS and other diseases (Palanichamy A et al., Sci Transl Med. 2014;6(248):248ra106.; Bashford-Rogers RJM, et al., Nature. 2019;574(7776):122-126.; Ramesh A et al., Proc Natl Acad Sci U S A. 2020;117(37):22932-22943).

Data exclusions

CSF samples with low white blood cell counts (<10 cells / µl) were excluded from the study, as not enough B cells could be isolated for repertoire analysis. Below a threshold of ~20 B cells generating a meaningful antibody repertoire becomes challenging as clonal families can hardly be identified with low cell numbers (Tan et al. Arthritis Rheumatol. 2014, 66(10):2706-15, PMID: 24965753; Lu et al, Clinical Imm., 2014, 152(1-2):77-89, PMID: 24589749). CSF samples with <10 cells / µl do not contain sufficient numbers of B cells to generate phylogenetic trees of the B cell repertoire that identifies representative clonal families and thus would not be of use in characterizing the antibody repertoire in MS.

Replication

All regimens of biological and technical replication are annotated in the methods section. FACS sorting of and single-cell sequencing of B cells was performed once for each patient sample, with each entire sample being exhausted by the sort. For ELISA experiments with in-house expressed recombinant monoclonal antibodies, experiments were performed at least 3 independent times. In each experiment included were at least 2 technical replicates per sample. Western Blot analyses were performed at in at least 3 independent experiments. EAE experiments were performed at least three times with 10 mice per subgroup in each experiment. For ELISA experiments on human MS samples, experiments were performed on MS samples from 3 independent cohorts of patients, with each cohort analyzed in at least 2-3 independent experiments. Custom-made microarray analysis was performed once in 8-fold technical replicates and for several important proteins multiple versions were included, each in 8-fold technical replicates. All attempts at replication were successful.

Randomization

As this study does not explore group differences, no specific randomization strategy has been applied for the selection of study participants. MS is a disease that predominantly affects females, which is reflected in our study cohort which contained paired CSF and blood samples from n = 8 female and n = 1 male MS patients.

Mice were randomly assigned a treatment group. Animals from each treatment group were housed together in the same cages.

Blinding

Participants were de-identified. As this study does not explore differences between patient subgroups, investigators were not blinded to the de-identified sample disease states for sequencing experiments, given comparisons were not being made between patient subgroups. The study does not rely on subjective measures, but analyzes sequencing data.

For the flow cytometry experiments, investigators were not blinded because all patients had MS and the primary purpose of the flow cytometry sort was to isolate B cells for single-cell antibody repertoire sequencing.

Investigators were not blinded during ELISAs, given ELISA provides a non-subjective quantitative readout.

For the mouse EAE experiments, investigators scoring EAE severity were blinded to their immunization groups. For analysis of EAE histopathology, the scoring pathologist was blinded to the immunization groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

1) anti-CD19 (clone HIB19, BioLegend #302234, Brilliant Violet 421)
2) anti-CD20 (clone L27, BD Biosciences #340955, PerCP-Cy5.5)
3) anti-CD38 (clone HB7, BD Biosciences #335790, PE-Cy7)
4) anti-CD3 (clone OKT3, BioLegend #317346, PE/Dazzle 594)
5) anti-CD27 (clone O323, BioLegend #302830, Brilliant Violet 605)
6) anti-IgM (clone MHM-88, BioLegend #314520, APC-Cy7)
7) anti-IgD (clone IA6-2, BD Biosciences #555778, FITC)
8) anti-HLA-DR (clone L243, BioLegend #307626, Alexa Fluor 700)
9) anti-α4 integrin (clone 9F10, BioLegend #304308, APC)
10) anti-IgA (clone IS11-8E10, Miltenyi Biotec #130-113-476, PE)
11) anti-EBNA1 (clone EBS-I-024, Biorbyt #orb557160)
12) anti-GlialCAM/Hepacam (clone 419305, R&D Systems #MAB4108)
13) Cy-3-conjugated secondary goat anti-human IgG antibody (Jackson Immunoresearch #109-165-088)
14) Alexa Fluor 700 conjugated anti-CD3 (clone SK7, BD Biosciences #566796)
15) In-house recombinant antibodies derived from CSF repertoires were expressed as outlined in the methods section.

Validation

Western blot antibodies were tested on western blot membranes containing recombinant target protein and unrelated proteins.
Links to data sheets containing validation data:
1) https://www.biolegend.com/en-us/products/brilliant-violet-421-anti-human-cd19-antibody-7144?GroupID=BLG5913
2) https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/clinical-discovery-research/single-color-antibodies-ruo-gmp/percp-cy-5-5-mouse-anti-human-cd20.340955
3) https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/clinical-discovery-research/single-color-antibodies-ruo-gmp/pe-cy-7-mouse-anti-human-cd38.335790
4) https://www.biolegend.com/en-us/search-results/pe-dazzle-594-anti-human-cd3-antibody-11986
5) https://www.biolegend.com/en-us/products/brilliant-violet-605-anti-human-cd27-antibody-7804
6) https://www.biolegend.com/en-us/products/apc-cyanine7-anti-human-igm-antibody-7403
7) https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/fitc-mouse-anti-human-igd.562023
8) https://www.biolegend.com/fr-ch/products/alexa-fluor-700-anti-human-hla-dr-antibody-3403?GroupID=BLG10409
9) https://www.biolegend.com/fr-ch/products/apc-anti-human-cd49d-antibody-582
10) https://www.miltenyibiotec.com/US-en/products/iga-antibody-anti-human-is11-8e10.html#pe:30-tests-in-60-ul
11) https://www.biorbyt.com/ebv-ebna1-antibody-orb557160.html
12) https://www.rndsystems.com/products/human-hepacam-antibody-419305_mab4108
13) https://www.jacksonimmuno.com/catalog/products/109-165-088
14) https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/apc-r700-mouse-anti-human-cd3.566796
15) In house recombinant antibodies. Amounts of in-house antibodies were measured by nanodrop and anti-human IgG ELISA (Bethyl Inc.), and then size-separated by western blot and stained with coomassie protein stain to assess the purity.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Expi293T (human, Thermo Fisher Scientific, A14527), BL21 chemically competent E. coli (Sigma Aldrich, St. Louis, MO)

Authentication

Expi293T cells were distributed with certificates of authentication, and used for transient transfection to express the in-house recombinant antibodies. Recombinant antibody expression was successful with the expressed antibodies confirmed to be intact as described above - thus the cell served their purpose for recombinant monoclonal antibody production, and no further validation of the cell line was performed.

Mycoplasma contamination

Expi293F were tested regularly and were found negative for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Female 8-week-old SJL/J mice were purchased from the Jackson Laboratory and used for EAE-experiments. Tissue from an adult mouse FVB x C57BL/6 was used for histology. The FVB x C57BL/6 mice used were also 8-week-old females purchased from Jackson Laboratory. The mice were housed in recyclable individually ventilated (IVC) cages, with a 12-hour light/dark cycle, at a temperature of 70 degrees F, and with 50% humidity.
Wild animals	No wild animals were used in the study.
Field-collected samples	No field collected samples were used in the study.
Ethics oversight	All animal experiments were performed in accordance with state and federal guidelines and regulations, and approved by the Stanford Institutional Animal Care and use Committee.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	n=12 patients between 18 and 76 years of age were included in the study. Nine patients with multiple sclerosis and three with non-MS neuro-inflammatory diseases. Three male and nine female patients were included, somewhat reflecting the female-dominated sex-distribution of multiple sclerosis. Patient details are listed in Extended Data Table 1.
Recruitment	Patients with a likely diagnosis of multiple sclerosis were recruited by their treating physicians on the neurological wards of the university hospitals in Stanford and Heidelberg during routine work-ups that required lumbar punctures. The diagnoses of multiple sclerosis were ultimately confirmed for all included samples. Beyond the requirement for having the diagnosis of MS, the sole exclusion criteria was a CSF cell count below 10 cells / µl CSF. Below a threshold of ~20 B cells in total, generating a meaningful antibody repertoire becomes challenging as clonal families can hardly be identified with low cell numbers (Tan et al. Arthritis Rheumatol. 2014, 66(10):2706-15, PMID: 24965753; Lu et al, Clinical Imm., 2014, 152(1-2):77-89, PMID: 24589749). In our experience, it is extremely challenging to retrieve >20 viable cells and ultimately single-cell antibody sequences from those cells, if the initial CSF sample has less than 10 cells / µl CSF. All included MS patients undergoing lumbar punctures were recruited in a sequential fashion, without any additional selection or non-selection of potential patients - as a result we do not believe there was any self-selection bias or other biases. Written consent was obtained from all patients.
Ethics oversight	All experimental protocols were approved by the institutional review board of Stanford University (IRB# 34529) and the ethics committee of the medical faculty of the University of Heidelberg (IRB# S-466/2015).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- [x] The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- [x] The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- [x] All plots are contour plots with outliers or pseudocolor plots.
- [x] A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	CSF was centrifuged immediately after lumbar puncture and cells were counted. PBMCs were isolated from heparin blood by density gradient centrifugation using Ficoll PLUS media (Cytiva). Cells were magnetically separated with anti-CD19 magnetic beads (Dynabeads CD19 Pan B cell isolation kit, Invitrogen), then stained according to standard protocols, using the antibodies detailed above.
Instrument	FACSAria II cell sorter (BD Biosciences)
Software	FACS Diva (data collection), FlowJo (data analysis).
Cell population abundance	Plasmablasts in peripheral blood: median: 4.1%, SD:12.2 of all B cells. Plasmablasts in CSF: median: 29.8%, SD:20 of all B cells.
Gating strategy	Details of the gating strategy are annotated in Extended Data Fig. 1a and b. Briefly, we applied forward and side scatter parameters (FSC-A, FSC-W, SSC) and live/dead stain (Sytox Blue) to select for live lymphocytes. Sorted plasmablasts from peripheral blood were gated on CD3-/CD19+/IgD-/CD27+/CD38+ (plasmablast gate (4.12%) in panel 5 in the representative

flow cytometry plots shown in Extended Data Fig. 1a), low expression of CD20 on plasmablasts was confirmed. Sorted B cells from CSF were gated on CD3-/CD19+ (B cell gate (73.9%) in panel 4 in the representative flow cytometry plots shown in Extended Data Fig. 1b). A plasmablast population expressing CD19+/IgD-/CD27+/CD38+/CD20low was identified in CSF and was included in the sorted population.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.