Title
Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GlialCAM

Journal
Nature, 603(7900)

ISSN
0028-0836

Authors
Lanz, Tobias V
Brewer, R Camille
Ho, Peggy P
et al.

Publication Date
2022-03-10

DOI
10.1038/s41586-022-04432-7

Peer reviewed

HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2022 September 01.

Published in final edited form as:

Nature. 2022 March ; 603(7900): 321–327. doi:10.1038/s41586-022-04432-7.

Clonally Expanded B Cells in Multiple Sclerosis Bind EBV EBNA1 and GlialCAM

Tobias V. Lanz[1,2,3], **R. Camille Brewer**[1], **Peggy P. Ho**[4], **Jae-Seung Moon**[1], **Kevin M. Jude**[5], **Daniel Fernandez**[6], **Ricardo A. Fernandes**[5], **Alejandro M. Gomez**[1], **Gabriel-Stefan Nadj**[1], **Christopher M. Bartley**[7,8], **Ryan D. Schubert**[9], **Isobel A. Hawes**[9], **Sara E. Vazquez**[10], **Manasi Iyer**[11], **J. Bradley Zuchero**[11], **Bianca Teegen**[12], **Jeffrey E. Dunn**[13], **Christopher B. Lock**[13], **Lucas B. Kipp**[13], **Victoria C. Cotham**[14], **Beatrix M. Ueberheide**[14], **Blake T. Aftab**[15], **Mark S. Anderson**[16], **Joseph L. DeRisi**[10,17], **Michael R. Wilson**[9], **Rachael J.M. Bashford-Rogers**[18], **Michael Platten**[2,3,19], **K. Christopher Garcia**[5], **Lawrence Steinman**[4], **William H. Robinson**[1,*]

[1]Division of Immunology and Rheumatology, Department of Medicine, Stanford University School of Medicine, 269 Campus Drive, Stanford, CA 94305, United States, and the Geriatric Research, Education, and Clinical Centers (GRECC), VA Palo Alto Health Care System, 3801 Miranda Ave, Palo Alto, CA 94304, United States

[2]Department of Neurology, Mannheim Center for Translational Neurosciences (MCTN), Medical Faculty Mannheim, University of Heidelberg, Theodor-Kutzer-Ufer 1-3, 68167 Mannheim, Germany

[3]Department of Neurology and National Center for Tumor Diseases, University Hospital Heidelberg, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany

[4]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Beckman Center for Molecular Medicine, 279 Campus Drive, Stanford, CA 94305, United States

[5]Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Beckman Center for Molecular Medicine, 279 Campus Drive, Stanford, CA 94305, United States

[6]Stanford ChEM-H Institute, Macromolecular Structure Knowledge Center, 290 Jane Stanford Way, Stanford, CA 94305, United States

[7]Hanna H. Gray Fellow, Howard Hughes Medical Institute, 4000 Jones Bridge Rd, Chevy Chase, MD 20815, United States

[*]Corresponding Author: William H. Robinson, Division of Immunology and Rheumatology, Department of Medicine, Stanford University School of Medicine, 269 Campus Drive, Stanford, CA 94305, United States, wrobins@stanford.edu.

Author contributions: Conceptualization, T.V.L., L.S., W.H.R.; Methodology, T.V.L., R.C.B., P.P.H., K.M.J., D.F., R.A.F., A.M.G., R.D.S., B.T., V.C.C., B.M.U., J.-S.M., M.I., J.B.Z.; Software, T.V.L., R.C.B., D.F., K.M.J., V.C.C.; Validation, T.V.L., R.C.B., K.M.J., B.M.U., R.J.M.B.-R., R.A.S., K.C.G., L.S., W.H.R., Formal Analysis, T.V.L., R.C.B., K.M.J., D.F., A.M.G., C.M.B., V.C.C., B.M.U.; Investigation, T.V.L., R.C.B., P.P.H., D.F., G.-S.N., C.M.B., R.D.S., I.A.H., S.E.V., B.T., V.C.C., J.-S.M., M.I.; Resources, T.V.L., P.P.H., D.F., B.T., J.E.D., C.B.L., L.B.K., B.M.U., M.R.W., M.S.A., J.L.D., M.P., R.A.S., K.C.G., L.S., W.H.R.; Data Curation, T.V.L., R.C.B., K.M.J., D.F., V.C.C., B.M.U.; Writing – Original Draft, T.V.L.; Writing – Review & Editing, T.V.L., R.C.B., P.P.H., L.S., W.H.R.; Visualization, T.V.L., R.C.B., K.M.J.; Supervision, T.V.L., W.H.R.; Project Administration, T.V.L., W.H.R.; Funding Acquisition, T.V.L., P.P.H., B.T.A., M.P., L.S., W.H.R.

Competing interests: W.H.R. owns equity in, serves as a consultant to, and is a member of the Board of Directors of Atreca, Inc. L.S. owns equity in and serves as a consultant to Atreca, Inc. T.V.L. and W.H.R. are coinventors on a patent application filed by Stanford University that includes antibodies to EBV. The remaining authors declare no competing interests.

[8]Weill Institute for Neurosciences, Department of Psychiatry and Behavioral Sciences, University of California San Francisco, 675 Nelson Rising Ln San Francisco, CA 94158, San Francisco, United States

[9]Weill Institute for Neurosciences, Department of Neurology, University of California San Francisco, 675 Nelson Rising Ln San Francisco, CA 94158, San Francisco, United States

[10]Department of Biochemistry and Biophysics, University of California San Francisco, 1700 4th Street, San Francisco, CA 94158, United States

[11]Department of Neurosurgery, Stanford University School of Medicine, 1201 Welsh Road, Stanford, CA, United States

[12]Institute of Experimental Immunology, Euroimmun AG, Seekamp 31, 23560 Lübeck, Germany

[13]Division of Neuroimmunology, Department of Neurology and Neurological Sciences, Stanford University School of Medicine, 213 Quarry Road, Stanford, CA, United States

[14]Department of Biochemistry and Molecular Pharmacology, NYU Perlmutter Cancer Center, and NYU Langone Health Proteomics Laboratory, Division of Advanced Research Technologies, NYU School of Medicine, 430 East 29th St, New York, NY, 10016, United States

[15]Preclinical Science and Translational Medicine, Atara Biotherapeutics, 611 Gateway Blvd South San Francisco, CA 94080, United States

[16]Department of Medicine, Diabetes Center, University of California San Francisco, 513 Parnassus Ave, San Francisco, CA 94143, United States

[17]Chan Zuckerberg Biohub, University of California San Francisco, 499 Illinois Street, San Francisco, CA 94158, United States

[18]Wellcome Centre for Human Genetics, University of Oxford, Roosevelt Dr, Headington, Oxford OX3 7BN, United Kingdom

[19]DKTK Clinical Cooperation Unit Neuroimmunology and Brain Tumor Immunology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

Abstract

Multiple sclerosis (MS) is a heterogenous autoimmune disease in which autoreactive lymphocytes attack the myelin sheath of the central nervous system (CNS). B lymphocytes in the cerebrospinal fluid (CSF) of MS patients contribute to inflammation and secrete oligoclonal immunoglobulins[1,2]. Epstein-Barr virus (EBV) infection has been linked to MS epidemiologically, but its pathological role remains unclear[3]. Here we demonstrate high-affinity molecular mimicry between the EBV transcription factor EBNA1 and the CNS protein GlialCAM, and provide structural and in-vivo functional evidence for its relevance. A cross-reactive CSF-derived antibody was initially identified by single-cell sequencing of the paired-chain B cell repertoire of MS blood and CSF, followed by protein microarray-based testing of recombinantly expressed CSF-derived antibodies against MS-associated viruses. Sequence analysis, affinity measurements, and the crystal structure of the EBNA1-peptide epitope in complex with the autoreactive Fab fragment allowed for tracking the development of the naïve EBNA1-restricted antibody to a mature EBNA1/GlialCAM cross-reactive antibody. Molecular mimicry is facilitated by a post-

translational modification of GlialCAM. EBNA1 immunization exacerbates the mouse model of MS and anti-EBNA1/GlialCAM antibodies are prevalent in MS patients. Our results provide a mechanistic link for the association between MS and EBV, and could guide the development of novel MS therapies.

The presence of oligoclonal bands (OCB) in cerebrospinal fluid (CSF) and the efficacy of B cell depleting therapies emphasize the importance of B cells in the pathobiology of multiple sclerosis (MS)[2]. Anti-viral antibodies against mumps, measles, varicella-zoster, and Epstein-Barr Virus (EBV) are often present in MS[4,5], but their relevance is unclear. Anti-EBV antibody titers in over 99% of MS patients provide evidence for an epidemiological link between MS and EBV[6]. Symptomatic infectious mononucleosis during EBV infection increases risk for MS[7]. Molecular mimicry between virus and self-antigens is a potential mechanism that might explain this association[8]. Antibodies against certain EBV nuclear antigen 1 (EBNA1) regions have been found in MS patients, including the region AA365–426[5,9–12], which we describe here in our identification of molecular mimicry between EBNA1 and the glial cellular adhesion molecule GlialCAM. The potential significance of this mimicry in the pathophysiology of MS is described in detail.

The B cell repertoire in MS CSF plasmablasts is highly clonal

CSF and blood samples were obtained from MS patients during the onset of disease (clinically isolated syndrome, n=5) or an acute episode of relapsing-remitting MS (n=4). Patients with a CSF pleocytosis of >10 cells/μl were selected (Extended Data Table 1, Supplementary Discussion). Single B cells were sorted by flow cytometry (Extended Data Fig. 1a,b). Characteristic phenotypic differences of B cells in blood and CSF were observed[13,14], including (i) high plasmablast (PB) counts in CSF vs. blood (Extended Data Fig. 1c,d), (ii) differing expression levels of integrin-α4 and HLA-DR in PB but not non-PB B cells (Extended Data Fig. 1e–j, Supplementary Table 1), and (iii) high abundance of immunoglobulin G (IgG) in CSF PB (Fig. 1a, Extended Data Fig. 1k,l).

We sorted PB from blood and B cells from matched CSF by flow cytometry, and sequenced their full-length paired heavy-chain (HC) and light-chain (LC) VDJ regions[15]. 13,578 paired sequences from blood PB and 1,689 from CSF B cells passed filter thresholds. The CSF repertoire is highly clonal (Fig. 1a, Extended Data Fig. 2a,3), suggesting antigen-specific proliferation of select clones within the CSF. Notably, CSF repertoires in three non-MS control patients with neuroinflammatory conditions (Extended Data Table 1) did not exhibit such extensive clonality and high IgG usage (Fig. 1a,b, Extended Data Fig. 2b–d). While the amount of somatic hypermutation (SHM) in immunoglobulin HC-V genes (IGHV) and LC-V genes (IGLV) does not differ between PB in blood and CSF (Extended Data Fig. 2e), HC-complementarity determining region 3 (CDR3) lengths are longer in CSF PB than in blood, suggesting ongoing intrathecal SHM (Extended Data Fig. 2f). Consistent with previous reports[16–19], repertoires in the CSF are skewed towards usage of five IGHV genes (Fig. 1c, Extended Data Fig. 2g–i). This suggests that select autoantigens drive PB survival and proliferation in CSF.

Clonal PB are the predominant source of oligoclonal immunoglobulins in MS CSF

Immunoglobulin levels are elevated in CSF of MS patients, and the presence of a few highly abundant OCB is a hallmark of MS. To determine if clonal PB are the source of intrathecal immunoglobulins, we isolated immunoglobulins from CSF, characterized them by mass spectrometry, and compared the spectra with their corresponding antibody repertoire sequence datasets. For 87% of clonal sequences, peptides matching variable chain sequences unique to the patient were identified, whereas this occurred only for 40% of singleton sequences (Fig. 1d). Highly abundant immunoglobulins, defined by detection of ≥ten peptide-spectral matches (PSM ≥10), aligned almost exclusively to clonally expanded B cells (Extended Data Fig. 2j) and to PB, which display more clonality than non-PB B cells (Extended Data Fig. 2k,l), suggesting that clonally expanded PB are the predominant source of CSF OCB.

CSF-derived monoclonal antibodies bind EBV antigens

A total of 148 BCR sequences from MS CSF, each representative of a clonal expansion, were expressed as monoclonal antibodies (mAbs) (Extended Data Fig. 3). MS CSF mAbs showed low polyreactivity (polyreactive mAbs MS: 5.4%; controls: 15.4%; Extended Data Fig. 4, 5b). To test for anti-viral reactivities, mAbs were probed on protein microarrays containing 2 EBV viral lysates, 23 latent and lytic EBV proteins, 240 peptides spanning four prominent EBV proteins, and lysates of 7 other MS-associated viruses, including measles, rubella, and varicella-zoster virus (VZV)[4] (Fig. 2a,b, Extended Data Fig. 5a, 6, 7a, Supplementary Tables 2,3). One-third of the expressed mAbs bind to EBV proteins and peptides and ~20% to other viruses, in particular to VZV and cytomegalovirus (CMV) (Fig. 2a). Notably, half of the VZV-reactive antibodies cross-react to CMV and EBV, indicative of conserved epitopes in herpes viruses.

Interestingly, we found mAbs from 6 out of 9 MS patients that bind to the EBV transcription factor EBNA1 (Fig. 2a, Extended Data Fig. 5a), and mAbs binding to EBNA1 peptides were identified in 8 of 9 patients (Fig. 2b, Extended Data Fig. 6). Anti-EBNA1-reactivity has been implicated in MS epidemiology[20], and the region AA365–425 is the target of stronger antibody responses in MS than in healthy individuals[5,9–12]. Protein and peptide microarrays revealed that mAb MS39p2w174 binds EBNA1 within this MS-associated region at AA386–405 (Fig. 2b, Extended Data Fig. 6). This epitope was confirmed by western blot analysis using full-length and truncated EBNA1 proteins (Fig. 2c), and by ELISA-based EBNA1 peptide scans (Fig. 2d). Alanine scanning identified the proline/arginine-rich region AA394–399 as the central epitope for MS39p2w174 (Fig. 2e, Extended Data Fig. 7b).

Crystal structure reveals key residues of MS39p2w174-Fab / EBNA1$_{AA386–405}$ interaction

While the presence of antibodies against the broader EBNA1 region AA365–425 has been described in MS[5,9–12], their relevance remains elusive. We solved the crystal structure of EBNA1$_{AA386–405}$ in complex with the MS39p2w174-Fab at a resolution of 2.5 Å (Fig. 2f–j, Extended Data Fig. 7c–e, PDB ID: 7K7R[21]). We demonstrate close interactions of the EBNA1 residues P394-P398 with all CDRs except for LC-CDR2. Residues Tyr31 and Tyr38 on LC-CDR1 together with Trp38 on HC-CDR1 and Pro108, Pro109, and Tyr114 on HC-CDR3 create a hydrophobic cage for the peptide's two N-terminal prolines Pro394 and Pro395 and the proximal side chain of Arg396 (Fig. 2h,i). The C-terminal end of the antibody binding groove is wider, and Pro398 is carried by a large aromatic tryptophan residue (Trp114 in HC-CDR1)(Fig. 2g,h). The central arginines Arg395 and Arg396 hydrogen bond with residues on HC-CDR2, HC-CDR3, and HC framework region 2.

The IGHV encoding gene of MS39p2w174 is IGHV3–7, one of the IGHV chains over-represented in CSF (Fig. 1c). Interestingly, all but two of the residues that directly interact with EBNA1 are unmutated germline residues (Extended Data Fig. 7f,g). We hypothesized that the unmutated ancestor of MS39p2w174 (germline) might have an innate propensity to bind EBNA1$_{AA386–405}$. Indeed, germline was significantly more polyreactive than MS39p2w174 (Extended Data Fig. 4), and bio-layer interferometry revealed that germline binds EBNA1 with only slightly lower affinity than MS39p2w174 (K_D MS39p2w174: 1.99 nM, SD: 0.63 nM; germline: 4.19 nM, SD: 0.76 nM) (Fig. 2k,l), indicating that SHM is not required for effective EBNA1 binding and that naïve B cells have EBNA1-specificity.

Molecular mimicry between EBNA1$_{AA386–405}$ and GlialCAM

We probed mAb MS39p2w174 on HuProt microarrays, which contained >16,000 proteins spanning >80% of the human proteome[22], and discovered that MS39p2w174 binds glial cell adhesion molecule (GlialCAM) (Fig. 3a). GlialCAM is an immunoglobulin superfamily cell adhesion molecule expressed in the CNS (Extended Data Fig. 8a, www.proteinatlas.org[23]) by astrocytes and oligodendrocytes[24–26]. Earlier proteomics studies on MS brain lesions by our laboratory revealed expression of GlialCAM in chronic-active plaques[27]. MS39p2w174 was also probed using human proteome-wide phage display immunoprecipitation and sequencing (PhIP-Seq) library[28]. MS39p2w174 did not highly enrich for any single peptide, suggesting low affinity for multiple native peptides (Extended Data Table 2). Peptide motif analysis identified a common Pro/Arg-rich motif that closely resembles the central epitope in EBNA1 (AA395–399, Fig. 2e, Extended Data Fig. 8b). Comparing PhIP-Seq results with Huprot results yielded two overlapping targets: the ubiquitously expressed actin filament associated protein 1 (AFAP1, without the Pro/Arg-rich motif) and the CNS protein GlialCAM (=HEPACAM). Binding of MS39p2w174 to the intracellular domain (ICD, AA262–416) of GlialCAM was confirmed by ELISA (Fig. 3b) and western blot (Fig. 3c). Several regions in mouse brain distinctly stain for MS39p2w174 with predominantly glial patterns, including radial Bergmann glia in cerebellum, glia limitans, and likely

perivascular glial cells in the hippocampus and brainstem (Fig. 3d, Extended Data Fig. 8c). MS39p2w174 stained rat oligodendrocytes and transgenic GlialCAM expressing K562 cells, where GlialCAM generates tight junctions (Extended Data Fig. 8d–f).

While we demonstrated that MS39p2w174 and its unmutated germline ancestor bound EBNA1 with similar affinity (Fig. 2k,l), their binding affinities to GlialCAM differ significantly: affinity of MS39p2w174 to GlialCAM is ≥10-fold increased (K_D EBNA1: 1.99 nM, SD: 0.63 nM; vs. K_D GlialCAM: 190 pM, SD: 17pM), while germline binds GlialCAM with lower affinity (K_D EBNA1: 4.19 nM, SD: 0.76 nM vs. K_D GlialCAM: 10.46 nM, SD: 4.12 nM) (Fig. 3e,f). Evidently, while germline harbors a propensity to bind to EBNA1, SHM of MS39p2w174 increased its affinity to the CNS mimic GlialCAM by an order of magnitude.

Phosphorylation of GlialCAM$_{AA370–389}$ at pSer376 enables MS CSF-derived MS39p2w174 binding

The EBNA1 epitope AA386–405 is located between the protein's long N-terminal Gly-Ala-rich low-complexity region (AA90–380) and its highly structured DNA-binding domain (AA461–607, PDB: 1B3T[29]). The GlialCAM region AA337–385, identified by phage display (Supplementary Table 4), is located at the C-terminal end of the ICD and contains a proline-rich region that resembles the central epitope of EBNA1 (Fig. 3g). Intracellular B cell autoantigens have been described in several autoimmune diseases[30–32]. MS39p2w174 reacts with both proteins on western blots under denaturing conditions (Fig. 2c, Fig. 3c), suggesting linear epitopes for both targets. This is consistent with predictions that both epitopes are located in intrinsically disordered regions of their respective proteins (Fig 3h,i).

Nevertheless, while MS39p2w174 binds with similar affinity to both the EBNA1 protein and EBNA1$_{AA386–405}$ peptide (K_D protein: 1.99 nM, SD: 0.63 nM; K_D peptide: 2.67 nM, SD: 0.078 nM), its binding affinity for the native peptide GlialCAM$_{AA370–389}$ is three orders of magnitude lower than for GlialCAM protein (K_D protein: 190 pM, SD: 17pM; K_D peptide: 302 nM, SD: 31 nM). Part of this discrepancy could be due to multimerization of GlialCAM[33]. Additionally, the intracellular domain of GlialCAM is heavily phosphorylated[33–35] (Extended Data Fig. 8g, phosphosite.org[36]), and post-translational modifications (PTMs) often determine antibody-antigen interactions[37]. We determined if phosphorylation at one of the four serine residues surrounding the central epitope region (residues Ser376, 377, 383, and 384) could increase the binding affinity of MS39p2w174 to GlialCAM$_{AA370–389}$. Indeed, phosphorylation at Ser376 increased MS39p2w174 binding by ~50-fold (K_D native peptide: 302 nM, SD: 0.078 nM; K_D pSer376 peptide: 6.1 nM, SD: 0.27 nM), and additional phosphorylation of Ser377 further enhanced binding affinity (K_D pSer376/pSer377 peptide: 3.73 nM, SD: 0.15 nM) (Fig. 3j–l). In contrast, citrullination of arginine residues Arg373, 380, and 387 did not alter binding to GlialCAM$_{AA370–389}$ (Extended Data Fig. 8h). The important residue Arg397 in EBNA1$_{AA386–405}$, which engages in two hydrogen-bonds with Glu64 at HC-CDR2 (Fig. 2h,j), is replaced with alanine in GlialCAM$_{AA370–389}$ (Ala381) (Fig. 3g). This likely explains the differential binding affinity to EBNA1 and GlialCAM peptides.

Phosphorylation at Ser376 likely promotes binding by providing new polar interactions to the proximal LC, possibly with Arg36, a positively charged residue that is mutated from asparagine in germline (Fig. 2h, Extended Data Fig. 7g). Together, our results show that post-translational phosphorylation enables cross-reactivity of anti-EBNA1 MS39p2w174 to GlialCAM.

Anti-GlialCAM IgG reactivity is elevated in MS patients

To determine if the observed anti-GlialCAM reactivity of MS39p2w174 represents a broader phenomenon in MS, we tested the remaining 147 clonal MS mAbs for reactivity against GlialCAM protein and peptides spanning GlialCAM$_{AA315–395}$. We found ten additional mAbs from seven patients that bound the ICD and seven more from four patients that bound the extracellular domain (ECD) (Fig. 3m), demonstrating that MS PBs generate antibodies against multiple GlialCAM epitopes. We then tested plasma samples from MS patients (*n=36*) and healthy individuals (*n=20*) for reactivity to EBNA1 and GlialCAM. As expected, all MS patients and the majority of healthy individuals exhibited plasma reactivity to EBNA1 protein (Fig. 3n). Reactivity to EBNA1$_{AA386–405}$ and to GlialCAM was significantly increased in MS patients (Fig. 3o,p). Increased plasma IgG levels against EBNA1$_{AA386–405}$, GlialCAM$_{AA370–389}$, and phosphorylated GlialCAM$_{AA370–389\ pSer376}$ were confirmed in a second independent patient cohort (MS, *n=71*; controls, *n=50*; Extended Data Fig. 9a). Similarly, we detected increased plasma IgG levels against GlialCAM, GlialCAM$_{AA370–389}$, and phosphorylated GlialCAM$_{AA370–389\ pSer376}$ in a third cohort (MS, *n=67*; controls, *n=31*; Extended Data Fig. 9b). To address if reactivity against EBNA1$_{AA386–405}$ could be blocked by GlialCAM, we showed inhibition of anti-EBNA1$_{AA386–405}$ reactivity by GlialCAM$_{AA370–389\ pSer376}$ in nine high-reactivity samples (Fig. 3q, Extended Data Fig. 9c), indicating that this molecular mimicry is prevalent in a subset of MS patients.

Immunization with EBNA1$_{AA386–405}$ aggravates EAE

To assess the effect of an anti-EBNA1$_{AA386–405}$ immune response on neuroinflammation, we utilized experimental autoimmune encephalomyelitis (EAE), the mouse model of MS. SJL/J mice were immunized with scrambled control peptide or EBNA1$_{AA386–405}$. Three weeks later, EAE was induced by a second immunization of the respective peptides mixed with PLP$_{AA139–151}$. Mice in the EBNA1$_{AA386–405}$ group generated robust antibody responses to both EBNA1$_{AA386–405}$ (Fig. 4a) and GlialCAM ICD (Fig. 4b). The antibody response to PLP$_{AA139–151}$ was unaltered in both groups (Extended Data Fig. 10a). In addition, EBNA1$_{AA386–405}$ immunization induced a strong CD4+ T cell response while the T cell response against PLP$_{AA139–151}$ remained stable in both groups (Extended data Fig. 10b). EBNA1$_{AA386–405}$ stimulated the secretion of B cell stimulatory Th1 cytokines IFN-γ, TNF, and IL-12, as well as IL-6 and IL-10 (Extended Data Fig. 10c–g), but suppressed the key Th17 cytokine IL-17 (Extended Data Fig. 10h). Clinically, the EBNA1 group exhibited more severe paresis (Fig. 4c), with enhanced CNS immune cell infiltration (Fig. 4d,e), and demyelination (Extended Data Fig. 10i,j).

To further assess T cell reactivity in humans, PBMCs of MS patients (*n=7*) with elevated anti-EBNA1 and anti-GlialCAM titers were stimulated with EBNA1 and GlialCAM proteins and peptides and compared to PBMCs from healthy individuals. The CD4+ T cells trended towards more IFN-γ expression upon stimulation with $EBNA1_{AA386–405}$ and GlialCAM ECD and ICD proteins, which resembled CD4+ T cells in mice (Extended Data Fig. 10k,l). In CD8+ T cells, high expression of IFN-γ and granzyme-B indicated a robust CD8+ T cell response against EBNA1 in both groups, while only CD8+ T cells from MS patients responded to GlialCAM ICD and GlialCAM ECD (Fig. 4f). One patient (MS16, Fig. 3o,p) showed extraordinarily high counts of IFN-γ+ granzyme-B+ CD8+ T cells upon stimulation with EBNA1, GlialCAM ICD, and $GlialCAM_{AA370–389}$ (Extended Data Fig. 10m).

Together, we showed anti-GlialCAM antibodies are generated in response to $EBNA1_{AA386–405}$ immunization. Further, co-immunization of mice with $EBNA1_{AA386–405}$ and $PLP_{AA139–151}$ enhanced CNS immune cell infiltration and demyelination, two prominent features of human MS pathology. In addition to anti-GlialCAM antibody titers, our human T cell data suggest an important role for CD8+ T cells in the immune response against GlialCAM.

Discussion

Viral triggers of MS and other autoimmune diseases have long been the subject of intense investigation, but evidence for their functional relevance is scarce[8]. By utilizing paired-chain BCR repertoire sequencing, rational selection of clonal antibody sequences, and three independent high-throughput proteomic platforms, we identified a monoclonal antibody from CSF of an MS patient that binds the MS-associated epitope $EBNA1_{AA386–405}$ with high affinity, and cross-reacts with GlialCAM. We demonstrate the presence of cross-reactive EBV EBNA1 and GlialCAM antibodies in approximately 20 – 25% of MS patients, and show that co-immunization of mice with this EBNA1 epitope exacerbates autoimmune demyelination. Our findings demonstrate a mechanistic link between EBV infection and the pathobiology of MS.

Our analysis of the single-cell paired-chain BCR repertoire from blood and CSF of MS patients revealed features of intrathecal oligoclonal expansion with ongoing SHM, strikingly different from CSF BCR repertoires in other neuroinflammatory diseases. PB in the CSF of MS patients are highly clonal with longer CDR3 regions and skew towards use of five IGHV genes[16–19,38,39]. IGHV3–7 has been previously observed in MS CSF and OCB[17,19,39–41]. The discovery that the IGHV3–7-based mAb MS39p2w174 cross-binds EBV EBNA1 and GlialCAM provides a direct link between IGHV3–7 and MS pathobiology.

Structural and binding data of MS39p2w174 from the CSF of an MS patient, and its unmutated germline ancestor show that germline only binds $EBNA1_{AA486–405}$ with high affinity, and that SHM is required for high affinity cross-reactivity with GlialCAM. Due to the CNS-restricted expression of GlialCAM, it is likely that immature IGHV3–7 expressing B cells encoding a germline precursor of MS39p2w174 entered the CNS/CSF space, encountered GlialCAM antigen, and then underwent affinity maturation that generated clones encoding high affinity anti-GlialCAM antibodies. This is concordant with our BCR

repertoire analysis, which revealed a highly clonal and skewed repertoire in CSF with low polyreactivity, indicative of a select group of B cells undergoing SHM, likely driven by perpetual activation by CNS GlialCAM.

The crystallographic structure revealed that Arg397 in $EBNA1_{AA386–405}$ engages in two hydrogen-bonds with HC-CDR2 and is likely responsible for a significant part of the high binding affinity between MS39p2w174 and EBNA1. We show that phosphorylation of GlialCAM at Ser376 restores high affinity. Epitopes altered by PTMs have been described as targets for autoimmune responses, including citrullinated proteins in rheumatoid arthritis and phosphorylated Ro/La in systemic lupus erythematosus[37,42]. Tissue-specific differences of PTMs could contribute to a lack of central tolerance. Several kinases have been described as risk genes for MS, including MERTK, MAPK1 and 3, TYK2, potentially contributing to alternative phosphorylation patterns in the CNS[43].

In addition, we show that co-immunization of mice with $EBNA1_{AA386–405}$ generates a robust B cell response against GlialCAM and aggravates EAE. Our finding of activated PB in the CSF expressing exceptionally high levels of HLA-DR, suggest that these B cells present antigens and exchange inflammatory signals with Tfh.

In summary, we present a detailed picture of the B cell repertoire in MS CSF and blood, demonstrating activated intrathecal IgG+ PBs undergoing affinity maturation. Analysis of a CSF-derived mAb targeting $EBNA1_{AA386–405}$ revealed molecular mimicry to GlialCAM. These findings provide a mechanistic link between EBV infection and the pathobiology of MS.

Methods

Study design and human Subjects.

Patient samples were collected at Stanford University CA, USA, and the University of Heidelberg, Germany. Relapsing remitting MS (RRMS) was diagnosed according to the current McDonald criteria[44,45]. None of the patients met the diagnostic criteria for NMOSD, in particular spinal lesions spanning ≥3 segments[46]. Patients were tested for antibodies against aquaporin-4 and MOG and showed negative results. All included patients had elevated CSF white blood cell counts ≥10 cells/μl, and blood contaminated CSF samples were excluded by visual and microscopic inspection. Paired peripheral blood and CSF samples were obtained at the time of clinical onset (clinically isolated syndrome, CIS) or during an acute relapse. All but one patient had not received any MS-specific treatment prior to sample collection (Extended Data Table 1). 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). Written informed consent was obtained from each patient.

Cell preparation, antibody staining, and flow cytometric cell sorting.

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 antibodies against the following cell surface markers: CD20 (clone L27, dilution 1:10), CD38 (clone HB7, dilution 1:30), IgD (clone IA6–2, dilution 1:20) (all BD Biosciences), CD3 (clone OKT3, dilution 1:60), CD19 (clone HIB19, dilution 1:20), CD27 (clone O323, dilution 1:20), IgM (clone MHM-88, dilution 1:40), HLA-DR (clone L243, dilution 1:100), α4 integrin (clone 9F10, dilution 1:100) (all BioLegend), IgA (clone IS11–8E10, dilution 1:20) (Miltenyi Biotec), and Sytox blue (ThermoFisher Scientific, dilution 1:500). Single cells were sorted with a FACSAria II cell sorter (BD Biosciences) using FACSDiva (v8.0, BD Biosciences) into 96-well PCR plates (BioRad). For single-cell repertoire sequencing, plasmablasts were sorted from PBMC (plasmablast gate (4.12%) in panel 5 in the representative flow cytometry plots shown in Extended Data Fig. 1a). All B cells were sorted from CSF (B cell gate (73.9%) in panel 4 in the representative flow cytometry plots shown in Extended Data Fig. 1b). FlowJo Version 10.7.1 (BD) and R version 3.6.1 was used to evaluate flow cytometry data.

Single-cell BCR repertoire sequencing.

BCR repertoire sequencing was carried out using our in house developed plate-bound single-cell sequencing technology as described previously[15,47,48]. Briefly, reverse transcription with oligo-dT was carried out in separate wells, attaching unique well-ID barcodes by template switching activity of Maxima Reverse Transcriptase (ThermoFisher Scientific). Barcoded cDNA from each plate were pooled and amplified in 3 consecutive PCRs, including attaching plate-specific barcodes and sequencing adapters. PCRs were carried out separately for HC of IgG, IgA, and IgM, as well as for κ for λ LC, and separate libraries were generated from each, gel-purified, cleaned with Ampure XP beads (Beckman Coulter), and sequenced on an Illumina MiSeq (Illumina) with 2 × 330 paired-end reads.

Sequence analysis.

The MiSeq FASTQ workflow was used for Fastq generation and plate demultiplexing. R version 3.6.1 was used for custom analyses. Paired reads of sequences that passed quality thresholds were stitched and separated by plate and well IDs. Similar reads sharing the same plate and well IDs were clustered into operational taxonomic units (OTUs)[49]. Consensus sequences were aligned to germline variable-chain immunoglobulin sequences with IMGT HighV-QUEST v1.3.1[50], which reports V, D, and J germline genes, HC and LC CDR3-lengths, and non-silent mutation counts and locations. Clonal expansions were defined based on sharing the same HC and LC V and J genes and exhibiting >70% amino acid identity within the HC and LC CDR3s. Percent clonality represents the percent of all sequences that fulfill these criteria. To calculate IGHV, IGLV, IGHJ, and IGLJ gene usages, percent abundance of each particular gene was calculated in blood and CSF PB of each patient and means were calculated across patients. Genes that were present in less than three CSF samples were excluded from this analysis. While our sequencing method preferentially captures PB sequences due to higher amounts of immunoglobulin mRNA (Extended Data Fig. 2m), enough non-PB B cell sequences passed filter thresholds to compare the non-PB repertoire to the PB repertoire in 7 patients (Extended Data Fig. 2a,b). For patient samples MS12 and C6 only PB were captured (while gating on all B cells), and for MS39 only PB were sorted. For phylogenetic analysis, sequences were binned according to their HC V-gene

family and V-gene. Concatenated LC and HC were then aligned with Muscle[51] and clustered with PhyML[52] using maximum-likelihood clustering. Each tree-partition was rooted by their HC V-gene. Phylogenetic trees were drawn in Python using the ETE 3 toolkit[53].

Peptide identification with mass spectrometry.

Immunoglobulins were purified from 1.5ml of CSF samples with Protein A (ThermoFisher Scientific). The purified IgGs were reduced with 0.02 M dithiothreitol at 57°C for 1 hour, alkylated with 0.05 M iodoacetamide at room temperature (RT) in the dark, and digested with trypsin overnight at RT. Peptides were extracted and desalted as previously described[54]. An aliquot of the peptide mixtures was loaded onto an Acclaim PepMap 100 precolumn (75μm × 2cm, C18, 3μm, 100Å) in-line with an EASY-Spray, PepMap column (75μm × 50cm, C18, 2μm, 100Å) with a 5μm emitter using the autosampler of an EASY-nLC 1000 (ThermoFisher Scientific). The peptides were gradient eluted into a Lumos Fusion Tribrid (ThermoFisher Scientific) mass spectrometer using a 120min gradient from 5% to 35% solvent B (90% acetonitrile, 0.5% acetic acid), followed by 10 minutes from 35% to 45% solvent B and 10 min from 45 to 100% B. High resolution full MS spectra were acquired with a resolution of 120,000, an AGC target of 4×10^5, a maximum ion time of 50 ms, and scan range of 400 to 1800 m/z. Following each full MS scan as many data-dependent HCD MS/MS spectra were acquired in the orbitrap as possible in a 3 second cycle time. Monoisotopic precursor selection (MIPS) was set to peptide, precursors with a charge state of 2 – 7 and minimum intensity of 5×10^4 were selected for MS/MS. Dynamic exclusion was set to 60 seconds after a single selection. All MS/MS spectra were collected using the following instrument parameters: resolution of 30,000, an AGC target of 10^5, maximum ion time of 120 ms, two microscans, 1.6 m/z isolation window, and Normalized Collision Energy (NCE) of 32.

The MS/MS spectra were searched against the respective peptide specific database including common contaminant proteins using the search engine Byonic[54]. The search parameters were set to trypsin allowing two missed cleavages, fixed modification of carbamidomethyl on cysteine, variable modification of oxidation on methionine and deamidation on glutamine and asparagine. Peptides mapping to variable regions of IgG were manually verified. In order to include only sequence-specific peptides, peptides that aligned to non-immunoglobulin or constant-region sequences were excluded from the analysis, as were peptides that aligned to the repertoire of multiple patients. Included were peptides that aligned to one variable sequence in a single patient. Peptides that aligned to more than one variable sequence in a single patient were included if all matching sequences were exact matches or clonally related, in which case the peptide was counted as representative for all matches. Counts of identical or non-identical peptide spectral matches (PSM) per sequence were tallied for each sequence. Sequences that had >1 or >10 matching peptides were presented as percentage of all sequences (Fig. 1j,k, Extended Data Fig. 2g,h). The mass spectrometry files are accessible at MassIVE (massive.ucsd.edu) under accession number MSV000086829.

Selection and recombinant expression of mAbs.

Representative antibodies from the largest clonal B cell expansions in the CSF of each patient were selected for recombinant expression. In patients with more than 10 large clonal expansions, sequences were preferentially chosen based on their usage of one of the 11 most abundant IGHV genes in the CSF (Extended Data Fig. 3). HC and LC variable sequences were custom generated (IDT), and cloned into pFuse vectors (Invivogen), containing human IgG constant region or kappa or lambda constant regions, respectively. Fab HC were expressed in in-house plasmids, containing the constant-region C1 up to Cys103. Plasmids were transfected into Expi293T cells using Expifectamine (ThermoFisher Scientific). Culture medium was harvested after 4 and 7 days post transfection. mAbs and Fabs were purified with protein A and protein G resins, respectively (ThermoFisher Scientific). Antibody concentrations were measured with a nanodrop spectrophotometer (ThermoFisher Scientific) and hIgG quantitation ELISAs (Bethyl Laboratories) and checked for purity on SDS protein gels with Coomassie staining.

Protein expression and purification.

EBNA1 proteins and peptides were obtained from: full-length AA1–641 (Abcam), AA328–641 (Virion Serion), AA408–641 (ProspecBio). GlialCAM proteins and peptides: full-length AA34–416 (OriGene), ECD AA34–234 (Novoprotein), and ICD AA262–416 with N-terminal His-Tag was cloned into a pet30(+) vector, expressed in BL21 chemically competent E. coli (Sigma Aldrich) to an OD of 600nm, and induced with IPTG (Sigma Aldrich) for 3h at 37°C. Cell pellets were disrupted by sonication and proteins were purified with cOmplete His-Tag Purification Resins (Roche Life Science), followed by size-exclusion purification (Cytiva). For all other used peptides and proteins see Supplementary Tables 2–4.

Planar protein microarrays.

Protein microarrays were generated as described previously [55,56] (https://web.stanford.edu/group/antigenarrays/). In brief, peptides, recombinant proteins, and lysates were diluted at the indicated concentrations in a 1:1 solution of PBS/water and protein printing buffer (ArrayIt) (Supplementary Tables 2–4), aliquoted on 384-well plates, and printed on SuperEpoxy Slides using a NanoPrint LM210 system (ArrayIt). Two independent quadruplicates of each analyte were spotted, and some proteins were used in several versions / preparations from different sources (Supplementary Table 2). Ready-made HuProt Arrays version 3.1 were obtained from CDI labs. Arrays were circumscribed with a hydrophobic marker, blocked overnight at 4 °C in PBS containing 3% FCS and 0.1% Tween-20, and incubated with individual mAbs at a concentration of 1 μg/ml for 1h at 4°C, then washed twice for 20 min in blocking buffer on a rotating shaker. Arrays were then incubated with Cy-3-conjugated secondary goat anti-human IgG (0.8 μg/mL) (Jackson ImmunoResearch) for 1h at 4 °C, then washed twice for 30 min in blocking buffer, twice for 30 min in PBS, and twice for 15 s in water. Arrays were spun dry and scanned with a GenePix 4000B scanner (Molecular Devices). Median pixel intensities for each fluorescent spot were determined with GenePix Pro-3.0 software (Molecular Devices). Z-scores for each row of antigens were calculated for viral antigens, raw intensities were analyzed for

GlialCAM arrays. Heatmaps were generated with Morpheus software (The Broad Institute; https://software.broadinstitute.org/morpheus).

PhIP-Seq.

PhIP-Seq was performed using a human proteome-wide library expressing overlapping 49-amino-acid peptides with a 24-AA sliding window approach starting at the N-terminus. Briefly, 2uL (1ug/mL) of substrate antibody was diluted 1:100 in blocking buffer for two sequential rounds of immunoprecipitation. After the second round of immunoprecipitation and amplification in E. coli, next generation sequencing libraries were prepared for paired-end 150 base next generation DNA sequencing on the Illumina Hi-Seq platform as previously described[28,57]. After alignment of the reads to the reference peptide sequences, quality control was performed and only reads present at an abundance of at least 10 reads per hundred thousand were carried forward. The number of reads mapping to each peptide were then counted and individually scored as a percentage of the total.

ELISA.

Cytokine ELISA kits were used according to manufacturers' instructions: mouse IL-6, IL-10, IL-12, IFN-γ, and TNF (BD Biosciences), and IL-17A (ThermoFisher Scientific). For protein and peptide ELISAs, MaxiSorp 384-well plates (ThermoFisher Scientific) were coated with 1 μg/ml peptide or protein in carbonate-bicarbonate buffer at 4°C overnight, then washed 6x with PBST (PBS + 0.05 % Tween20), blocked with blocking buffer (PBS + 1% BSA) for 1h, and mAbs were applied at 1 μg/ml in blocking buffer. Human and mouse plasma samples were diluted 1:100 and T cell supernatants 1:4 in blocking buffer. After overnight incubation at 4°C, plates were washed again 6x with PBST, secondary antibody HRP-conjugated goat anti-human IgG (Jackson ImmunoResearch), was applied for 1h at RT, and after 6 additional washes with PBST, plates were developed with TMB substrate (ThermoFisher Scientific), stopped with 1N sulfuric acid, and read on a SpectraMax Paradigm Microplate Reader (Molecular Devices). For plasma ELISAs with blocking of plasma IgG, MaxiSorp 384-well plates were coated with 2μg/ml recombinant Protein G (Acro Biosystems) at 4°C over night, then washed 6x with PBST, and incubated with 1:100 diluted plasma at 4°C over night. Plates were again washed 6x with PBST, and then incubated with the respective blocking peptides at 10μg/ml for 2h at RT. Then, biotinylated $EBNA1_{AA385–405}$ was added at 1μg/ml and incubated for 1h at RT. Plates were washed again 6x with PBST, incubated with HRP-conjugated streptavidin (BioLegend) for 1h at RT, and developed with TMB substrate as described above.

Western blotting.

Western blots were run according to standard protocols. Briefly, purified proteins were boiled in Laemmli-buffer with 10% beta-mercaptoethanol for 5 minutes, run on 4–12% Criterion XT Bis-Tris protein Gels (Bio-Rad), then transferred onto a nitrocellulose membrane using a Trans-Blot Turbo semi-dry transfer system (Bio-Rad) and stained with MS39p2w174 at 10μg/ml or with mouse anti-EBNA1 antibody (Biorbyt) or mouse anti-GlialCAM antibody (R&D Systems) followed by secondary HRP-conjugated goat anti-human IgG and anti-mouse IgG (Jackson ImmunoResearch). Coomassie gels were run concomitantly, fixed with 10% methanol and 7% acetic acid, and stained according

to standard protocols. Uncropped western blot and Coomassie images are available in Supplementary Figure 1.

Fluorescent Immunohistochemistry on mouse brain slices and Immunofluorescence on primary cultured rat oligodendrocytes.

An adult mouse (F1 generation of FVB × C57blk6 cross) was transcardially perfused with 4% paraformaldehyde and post-fixed in 4% PFA overnight at 4C. After sucrose equilibration, the brain was blocked in OCT and sectioned at 12μm on a standard cryostat. Sections were permeabilized and blocked in PBS containing 10% lamb serum and 0.1% triton x-100. Sections were immunostained with concomitantly expressed control mAb anti-DSG3 (Acc#: HQ338093.1 and HQ338094.1) (18μg/mL), MS39p2w174 (18μg/mL), or PBS in blocking buffer overnight at 4C. Sections were washed five times with PBS over 1 hour and counterstained with anti-human IgG for 1 hour at room temperature (Alexa Fluor 488, Jackson ImmunoResearch) (2μg/mL). Nuclei were stained with DAPI at 1:2000 for 5 minutes at room temperature. Rat oligodendrocyte precursor cells were prepared from rat embryos followed by panning and in-vitro differentiation into primary rat oligodendrocytes[58]. Cells on cover slips were permeabilized with ice-cold 100% methanol for 10min, blocked with 10% donkey serum for 1h at RT, and then stained with isotype control (anti-DSG3) or MS39p2w174 at 10μg/ml in 1% donkey serum for 1h at RT, before incubation with secondary Alexa Fluor 647 donkey anti-human IgG antibody (Jackson ImmunoResearch) for 1h at RT. Confocal images were taken with a Zeiss LSM 880 confocal microscope using ZEN software (Zeiss).

Bio-Layer Interferometry.

Association and dissociation constants of mAbs to proteins and peptides were measured with bio-layer interferometry on an Octet QK device (Fortebio / Sartorius, Fremont, CA) according to standard protocols. For peptide kinetics, biotinylated peptides were bound to high precision streptavidin (SAX) biosensors (peptide concentration in solution: 100 nM) and mAbs MS39p2w174 and germline were probed as analytes in concentrations ranging from 10 – 270 nM. For protein kinetics, mAbs were bound to anti-hIgG Fc Capture (AHC) biosensors (mAb concentration in solution: 20 nM), and proteins were probed as analytes in concentrations ranging from 1.56 – 125 nM. Data was analyzed with BLI analysis software (Fortebio / Sartorius, version 7.1) and with GraphPad prism (version 8.4). Buffer controls were subtracted, and curves were fitted globally for each group consisting of all concentrations of the same ligand. K_D values ± SD as well as association / dissociation curves were reported and plotted with GraphPad prism (version 8.4). K_D, K_{on}, and K_{off} values are shown in Supplementary Table 6.

Prediction of Protein Disorder.

Order and disorder along the amino acid sequences of EBNA1 and GlialCAM were analyzed with PONDR (Predictor of Natural Disordered Regions, WSU Research Foundation)[59], using the VSL2 algorithm.

Crystallization of antibody-antigen complexes.

EBNA1$_{AA386–405}$ 20mer peptides (>98% purity) (Sigma Aldrich) were mixed with MS39p2w174-Fab (15 mg/ml) in a 1:7.5 molar ratio and incubated overnight. Crystals for MS39p2w174-Fab + EBNA1$_{AA386–405}$ grew in 0.48M Sodium Citrate, 0.72M Sodium/Potassium Phosphate, and 3% MPD (v/v) in 0.1M HEPES, pH 6.9 (Extended Data Fig. 6c). Crystals were harvested, cryo-protected with a quick dip in a cryo-solution containing the well solution with 25 % glycerol, and flash cooled in liquid nitrogen. Data were collected at beamline SSRL 12–2 at the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, and processed and scaled using XDS/aimless and Staraniso[60,61]. Crystals belonged to space group I222 (a = 119.66 Å, b = 137.56 Å, c = 179.00 Å, $\alpha = \beta = \gamma = 90°$) and contained two Fab-peptide complexes per asymmetric unit (Extended Data Fig. 6d). Phaser was used for molecular replacement[62] with the model structure 4LRI (PDB)[63], stripped of all CDR loops. Loops were re-constructed with Coot[64] and structures were refined with phenix.refine[65,66] in iterations with Coot. Measurements and figure design was done with Pymol v2.1[67]. The structure was deposited in the protein data bank (PDB, www.rcsb.org [26]) with PDB ID: 7K7R.

Mouse Immunization, EAE, and histology.

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. 8-week-old female SJL/J and FVB × C57BL/6 mice were purchased from The Jackson Laboratory. The mice were housed in recyclable individually ventilated (IVC) cages, with a 12-light/hour dark cycle, at a temperature of 70 degrees F, and with 50% humidity. Mice were immunized s.c. with 200 μg/mouse of EBNA1$_{AA386–405}$ (peptide sequence: SQSSSSGSPPRRPPPGRRPF) or scrambled control peptide (peptide sequence: SPSRPGRSRSRGSPFPQPSP) (10 mice per group), mixed with 100 μg/mouse of CpG (ODN1826, Invivogen, San Diego, CA) in 100 μl/mouse incomplete Freund's adjuvant (BD Difco, Franklin Lakes, NJ). 3 weeks later EAE was induced by s.c. immunization with 200 μg/mouse of PLP$_{AA139–151}$, mixed with the same peptides used in the first immunization, in 100μl of incomplete Freund's adjuvant, supplemented with 200 μg/mouse of mycobacterium tuberculosis (strain H37 RA, BD Difco). Serum samples were obtained by retro-orbital blood draws 3 days before the 1st and 2nd immunizations (day −24 and day −3), and during termination of the experiment (day 50). Mice were weighed daily, and disease severity was assessed according to a 5-point standard scoring system: 0, no clinical signs; 1, loss of tail tone; 2, hind limb weakness; 3, complete hind limb paralysis; 4, hind limb and forelimb paralysis; 5, moribund or dead. Mice were euthanized on day 50 post induction of EAE by deep anesthesia with i.p. injections of 0.01 ml/g body weight 7.2% Xylazine (Bayer Healthcare, Leverkusen, Germany) and 10.8% Ketamine (Pfizer, New York City, NY). Lymph nodes and spleens were extracted, and mice were then perfused with 10 ml PBS and 20 ml 4% paraformaldehyde (PFA) (Electron Microscopy Sciences, Hatfield, PA). Brains and spinal cords were extracted, stored in 4% PFA overnight, followed by 30% sucrose in PBS. Tissues were embedded in paraffin, sectioned, and stained for H&E as well as Luxol fast blue according to standard protocols. Infiltration of inflammatory cells into the spinal cord on H&E slides was assessed by a blinded investigator by counting lesions of infiltrating cells per slide, taking lesion size into account: 1: small infiltrate (<10 cells), 2:

medium infiltrate (<100 cells), 3: large infiltrate (>100 cells). Demyelination was assessed by a blinded investigator according to a histological score[68]: 0.5, single demyelinated spot; 1, several spots; 2, confluent sites of demyelination; 3, extensive demyelination, less than half of a spinal cord; 4, demyelination of more than half of the spinal cord; and 5, extensive demyelination affecting >85 % of the total white matter of the spinal cord.

In vitro human T cell stimulation with proteins and peptides.

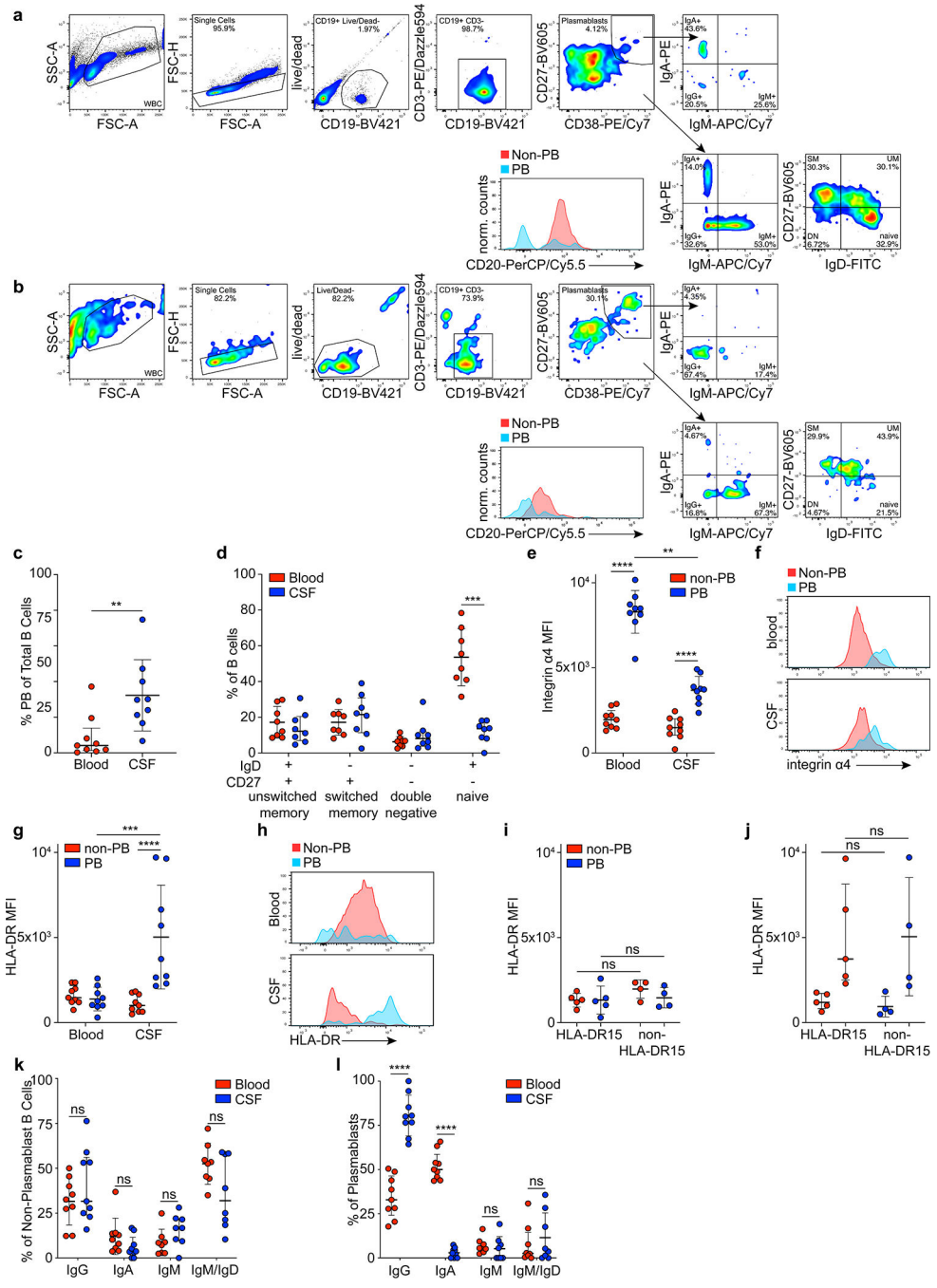
Cryopreserved PBMCs were thawed and stabilized overnight at 37°C. The cells were pre-incubated for 30 minutes at 37°C, 5% CO_2 with Polymyxin B (Sigma Aldrich) at a concentration of 10 μg/mL. The cells were then incubated for 16 hours with 100 μM of each recombinant protein or peptide in the presence of 2 μg/ml of anti-human CD28 (clone CD28.2, BD Biosciences) and anti-CD49d (clone 9F10, BioLegend) antibodies and IL-2 (50 IU/ml, Peprotech, Cranbury, NJ) and IL-7 (5 ng/ml, Peprotech). To detect intracellular staining, eBioscience™ Protein Transport Inhibitor Cocktail (500X, ThermoFisher Scientific) was added during the final 5 hours of culture. After 16 hours, the cells were labeled with Fixable Viability Stain 510 (BD Biosciences) for live cell staining and fluorophore conjugated anti-CD3 (clone SK7, BD Biosciences), CD4 (clone RPA-T4, BD Biosciences), CD8 (clone RPA-T8, BD Biosciences), Granzyme B (clone GB11, BD Biosciences), IFNγ (clone B27, BD Biosciences) and IL-17A (clone BL168, BD Biosciences) antibodies and detected using a BD LSR Fortessa.

Data analysis and statistics.

The publicly available dataset from Han et al.[27] was searched for abundance of GlialCAM in MS lesions. The old accession number Q8N7I3 was found, which was annotated in 2008 as unknown hypothetical protein and has since been replaced by accession number Q14CZ8. GlialCAM was identified with 2.5 mean spectral counts (MSC) in control tissue, 1.3 MSC in chronic plaques, 1.8 MSC in acute plaques, and 8 MSC in chronic-active plaques.

GraphPad Prism version 8.4.1 and R version 3.6.1 were used for statistical analyses. Statistical tests used are indicated in the respective methods section or in the figure legends

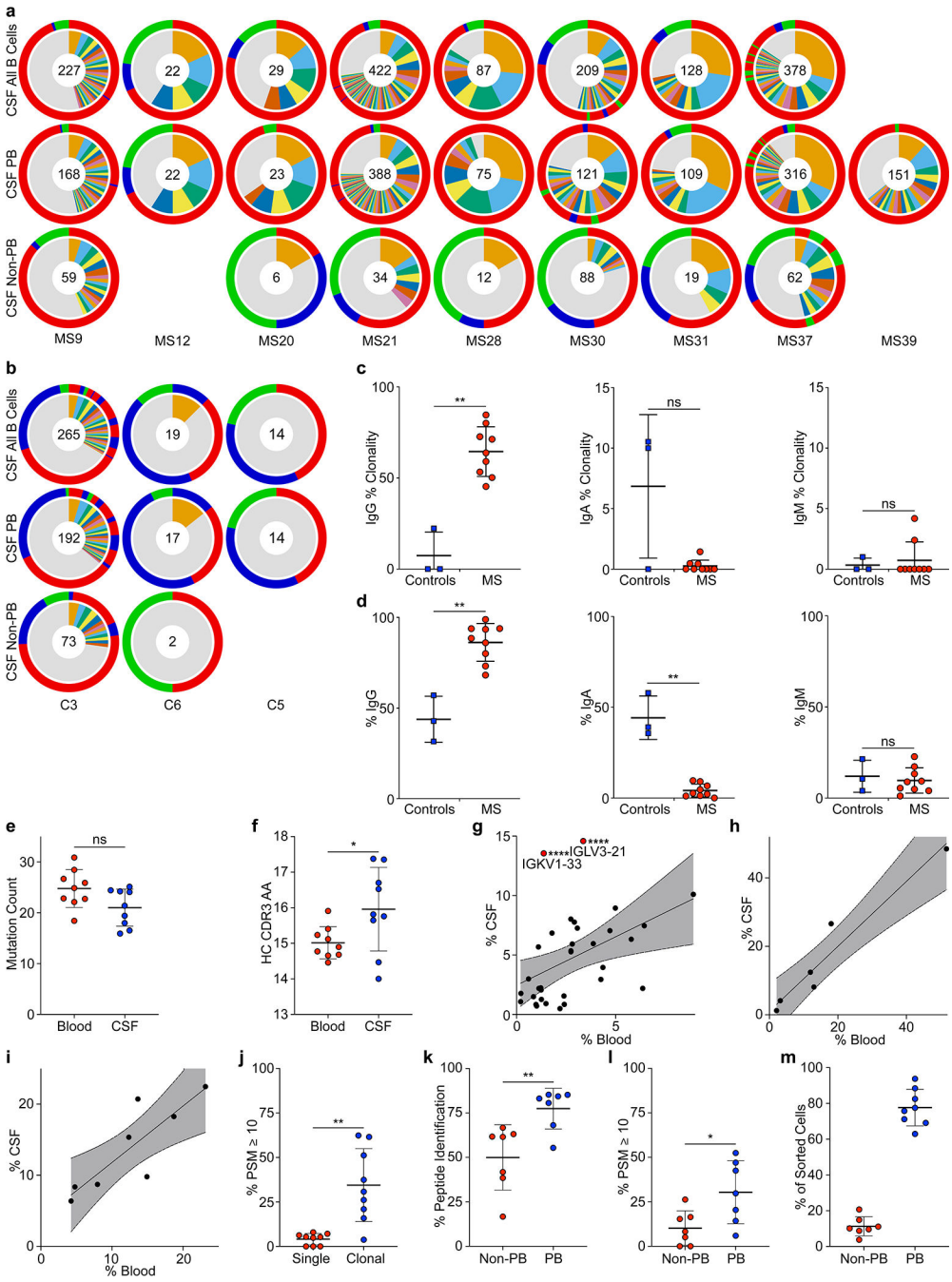
Extended Data



Extended Data Figure 1: Analysis of B cell phenotypes in MS blood and CSF.

a-l, Flow cytometry data, **a**,**b**, representative flow cytometry plots are shown for **a**, blood and **b**, CSF. **c**, plasmablasts as percent of all B cells in MS blood and CSF, means ± SD of *n=9* patient samples, ** $P = 0.004$, two-tailed Mann-Whitney test. **d**, non-plasmablast B cell subsets as percent of all B cells in blood (red) and CSF (blue), means ± SD of *n=8* patient samples, *** $P = 0.0006$, two-tailed Mann-Whitney test, Holm-Sidak corrected

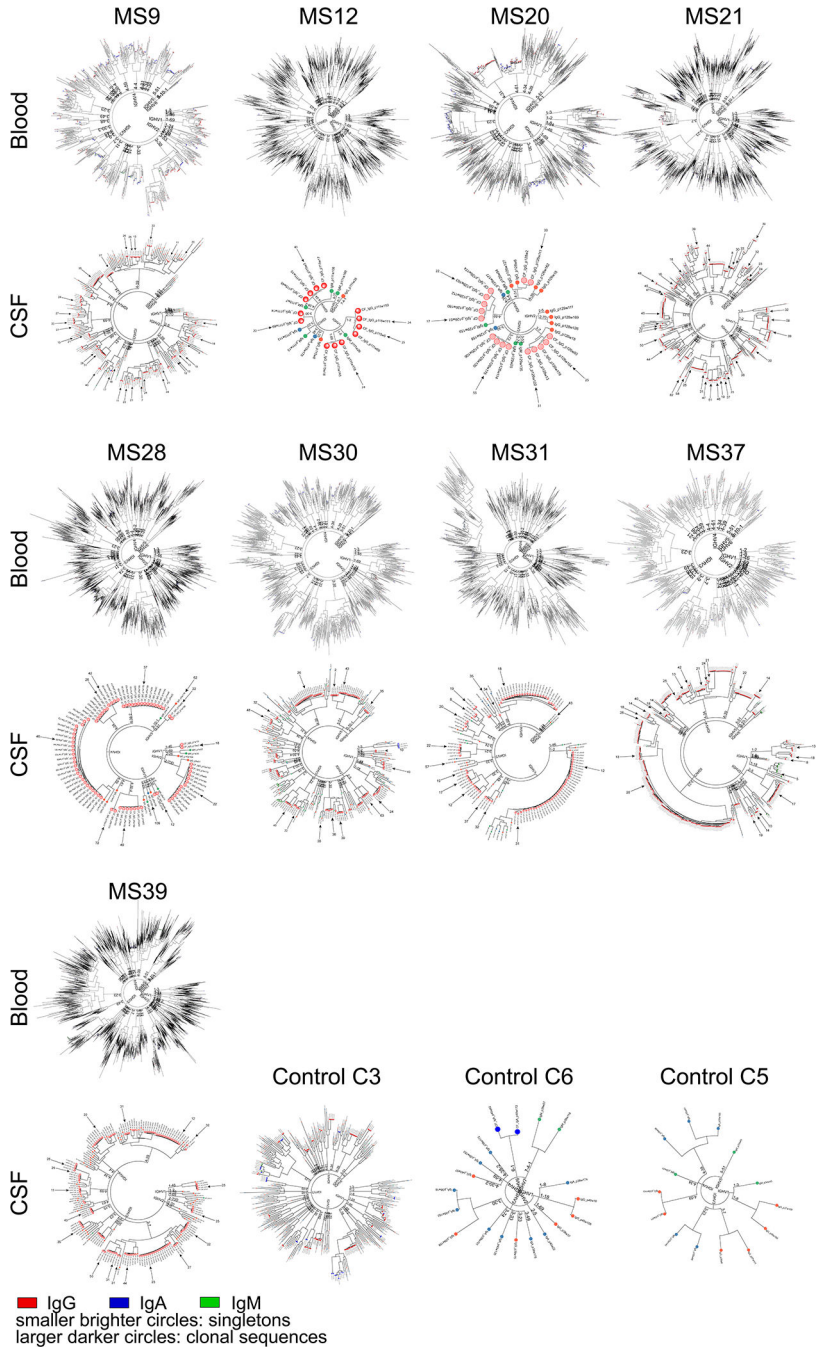
for multiple comparisons. **e**, Integrin alpha-4 expression in non-plasmablast B cells (red) and plasmablasts (blue), mean MFI ± SD of *n=9* patient samples, ****$P < 0.0001$, **$P = 0.0013$, two-way ANOVA, Tukey adjusted for multiple comparisons, **f**, representative histogram showing integrin alpha-4 expression in non-plasmablast B cells (red) and plasmablasts (blue) in blood (top panel) and CSF (lower panel), **g**, HLA-DR expression in non-plasmablast B cells (red) and plasmablasts (blue) in blood and CSF, mean MFI ± SD of *n=9* patient samples, ****$P < 0.0001$, ***$P = 0.0002$, two-way ANOVA, Tukey adjusted for multiple comparisons, and **h**, representative histogram showing HLA-DR expression in non-plasmablast B cells (red) and plasmablasts (blue) in blood (top panel), and CSF (lower panel). **i**, HLA-DR expression in patients carrying HLA-DRB1*15:01 (HLA-DR15, *n=5*) vs. other HLA-genotypes (non-HLA-DR15, *n=4*) in **i**, blood, and **j**, CSF, mean MFI ± SD, significance levels calculated with two-way ANOVA, **k**,**l**, Immunoglobulin classes in **k**, non-plasmablast B cells and **l**, plasmablasts in blood (red) and CSF (blue), mean MFI ± SD of *n=9* patient samples, ****$P < 0.0001$, two-way ANOVA, Holm-Sidak adjusted for multiple comparisons. Plasmablasts, PB; unswitched memory B cells, UM; switched memory B cells, SM; double negative B cells, DN.



Extended Data Figure 2: Extended BCR repertoire data.

a-i, Single-cell BCR repertoire sequencing data, **a**, individual repertoires from all CSF B cells (top row) and subdivided into CSF plasmablasts (middle row) and non-plasmablast B cells (bottom row) of *n=9* MS patients. **b**, Individual repertoires of all CSF B cells (top row) and subdivided into CSF plasmablasts (middle row) and non-plasmablast B cells (bottom row) of *n=3* control patients. Numbers indicate number of sequences, inner circle: colored wedges represent clonal expansions and grey area represents singleton antibody sequences, outer circle: immunoglobulin classes, red: IgG, blue: IgA, green: IgM, sequence

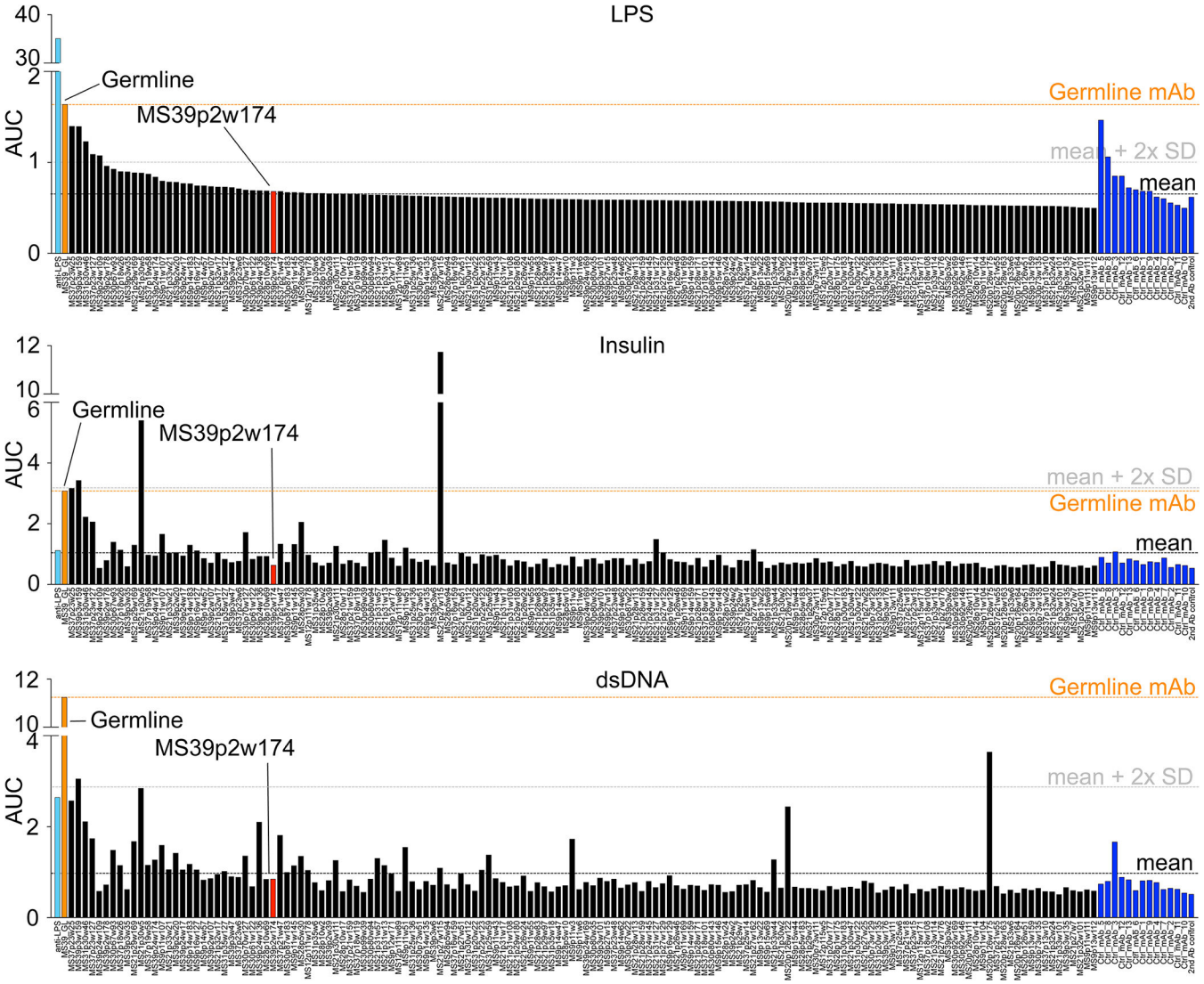
locations in outer circle correspond to inner circle. No non-plasmablast B cells were sorted for MS12 and C5. Only plasmablasts were sorted for MS39. **c**, Clonality, percent of clonal sequences in CSF B cells are shown, comparing BCR repertoires of control patients (*n=3*) to MS patients (*n=9*). Data corresponds to data shown in (Fig. 1b) and is separated into immunoglobulin classes IgG (left), IgA (center), and IgM (right). Means ± SD of individuals' repertoires are shown. **d**, Immunoglobulin class distribution, percent of IgG (left), IgA (center), and IgM (right) of all CSF B cells are shown for *n=3* control patients and *n=9* MS patients. Means ± SD of individuals' repertoires are shown. **e**, IGHV and IGLV cumulated mutation count in plasmablasts in blood (red) vs. CSF (blue), means ± SD of *n=9* patients samples. **f**, Mean HC CDR3 lengths (amino acid sequences) of plasmablasts in blood (red) vs. CSF (blue), means ± SD of *n=9* patient samples. **g-i**, Immunoglobulin gene distribution in blood vs. CSF plasmablasts for **g**, IGLV, IGKV1–33, ****P < 10^{-6}, IGLV3–21, ****$P = 3\times10^{-6}$ according to unpaired two-tailed Student's t tests, Holm-Sidak adjusted for multiple comparisons, **h**, IGHJ, and **i**, IGLJ. Each dot represents the usage of one gene across *n=9* MS patient repertoires in the respective compartments. Linear regression lines and 95% confidence interval are shown. **j**, Mass spectrometry data of purified CSF immunoglobulins, showing variable chain sequences that could be uniquely identified in singleton BCR sequences vs. plasmablast sequences, peptide-spectral matches (PSM) cutoff ≥10, means ± SD of *n=9* MS patients, **$P = 0.0012$. **k**,**l**, Same mass spectrometry data set as in (**j**), showing variable chain sequences that could be uniquely identified in non-plasmablast BCR sequences vs. plasmablast sequences, means ± SD of *n=7* MS patients, **k**, PSM cutoff ≥1, **$P = 0.007$, **l**, PSM cutoff ≥10, *$P = 0.037$. **l**, **m**, Single-cell sequencing efficacy in non-plasmablast B cells (red) vs. plasmablasts (blue) in CSF. Fraction of sequences that passed filter thresholds are shown as percentages of the number of sorted cells in the respective group, means ± SD of *n=8* patient samples (no non-PB value for MS39). **c,d,j-l**, *P* according to unpaired two-tailed Mann-Whitney test. **d-i**, *P* according to unpaired two-sided Student's t-test. Immunoglobulin heavy V gene, IGHV; Immunoglobulin heavy J gene, IGHJ; Immunoglobulin light V gene, IGLV; Immunoglobulin light J gene, IGLJ; peptide-spectral matches, PSM.



Extended Data Figure 3: Phylogenetic trees of B cells from MS blood and CSF.

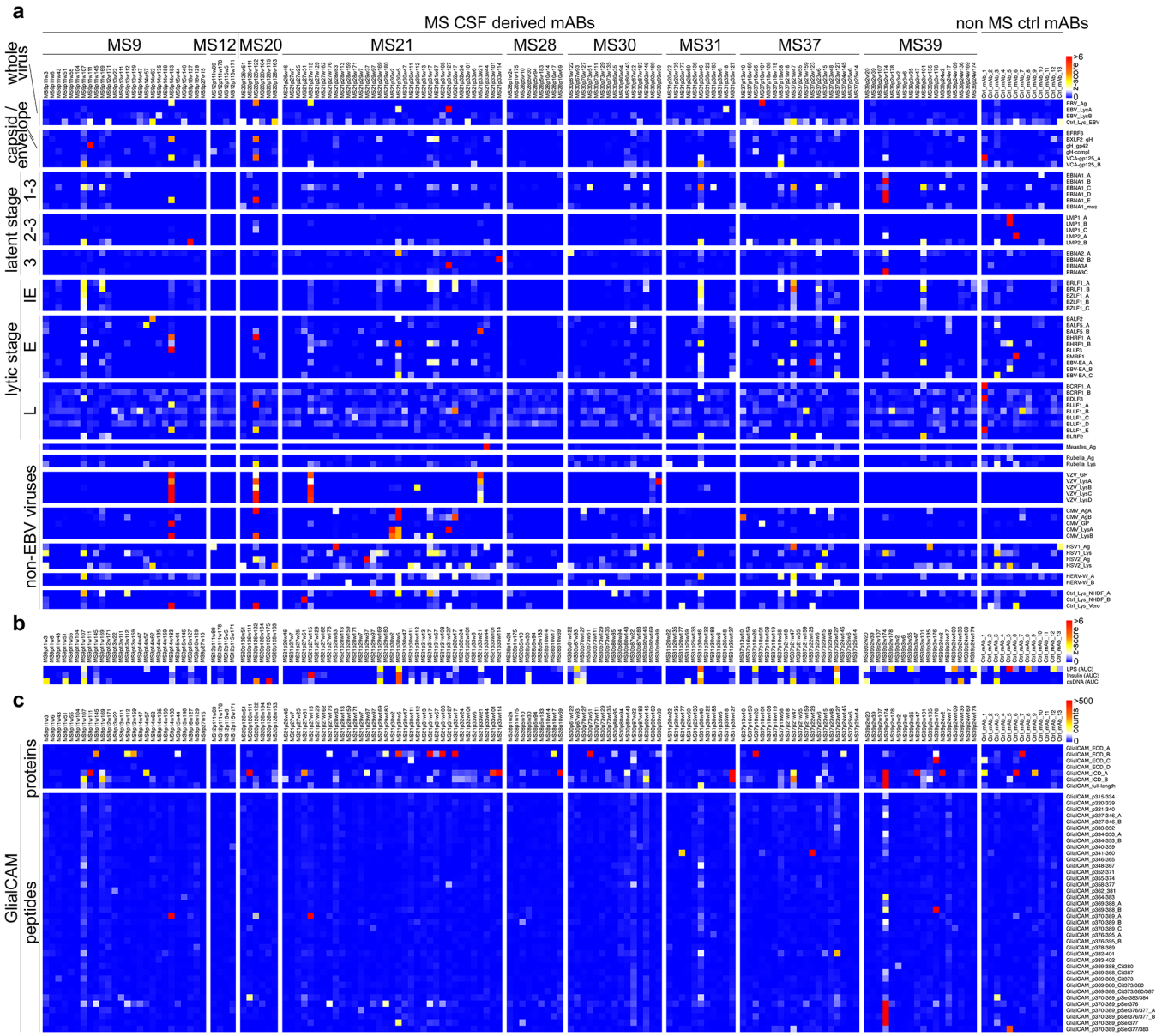
Blood plasmablasts (top rows) and CSF B cells (bottom rows) of *n=9* MS patients and CSF B cells of *n=3* control patients are shown. Each node represents the full-length heavy chain and light chain sequence of a single B cell. Trees are binned according to their IGHV families and genes, then the concatenated heavy chain and light chain sequences are clustered. IgG (red), IgA (blue), IgM (green). Smaller brighter circles indicate singleton B cells, larger darker circles indicate clonal expansions. Arrows indicate sequences that were

expressed as mAbs, numbers indicate V-gene mutation loads in heavy and light chains. Immunoglobulin heavy V gene, IGHV.



Extended Data Figure 4: Polyreactivity of recombinantly expressed antibodies.

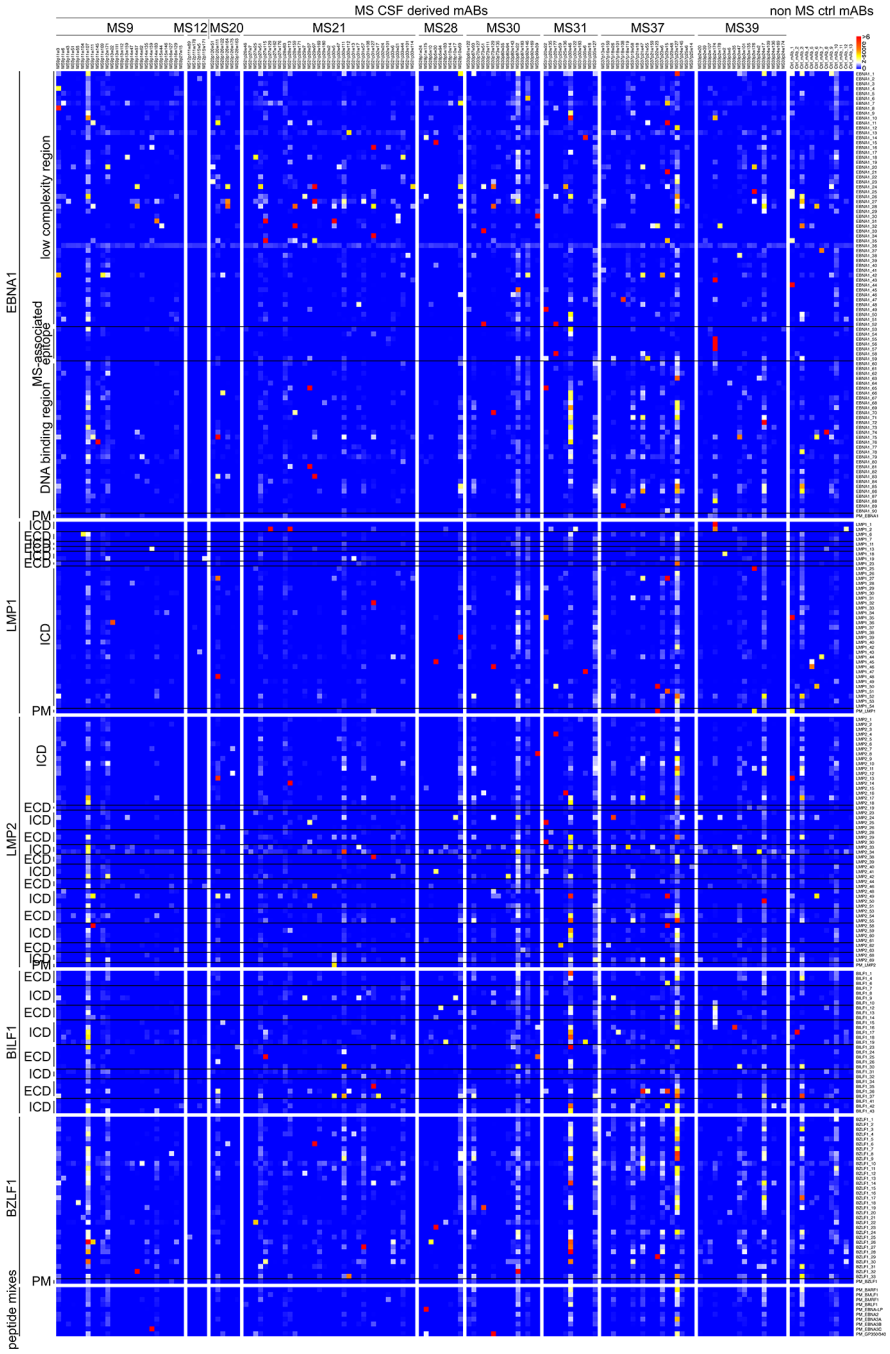
a, ELISA data showing reactivity of recombinant mAbs against LPS (top), human insulin (middle), and dsDNA (bottom). Reactivity is represented in the order of decreasing reactivity to LPS in MS mAbs and control mAbs, respectively. Measurements were carried out in duplicates at 0.1, 1, and 10 μg/ml mAb concentrations and the area under the curve (AUC) for each mAb is shown from one experiment. Commercial anti-LPS antibody (cyan), MS39p2w174 (red), germline (orange), control mAbs (blue).



Extended Data Figure 5: MS CSF mAb reactivity to EBV and GlialCAM antigens.

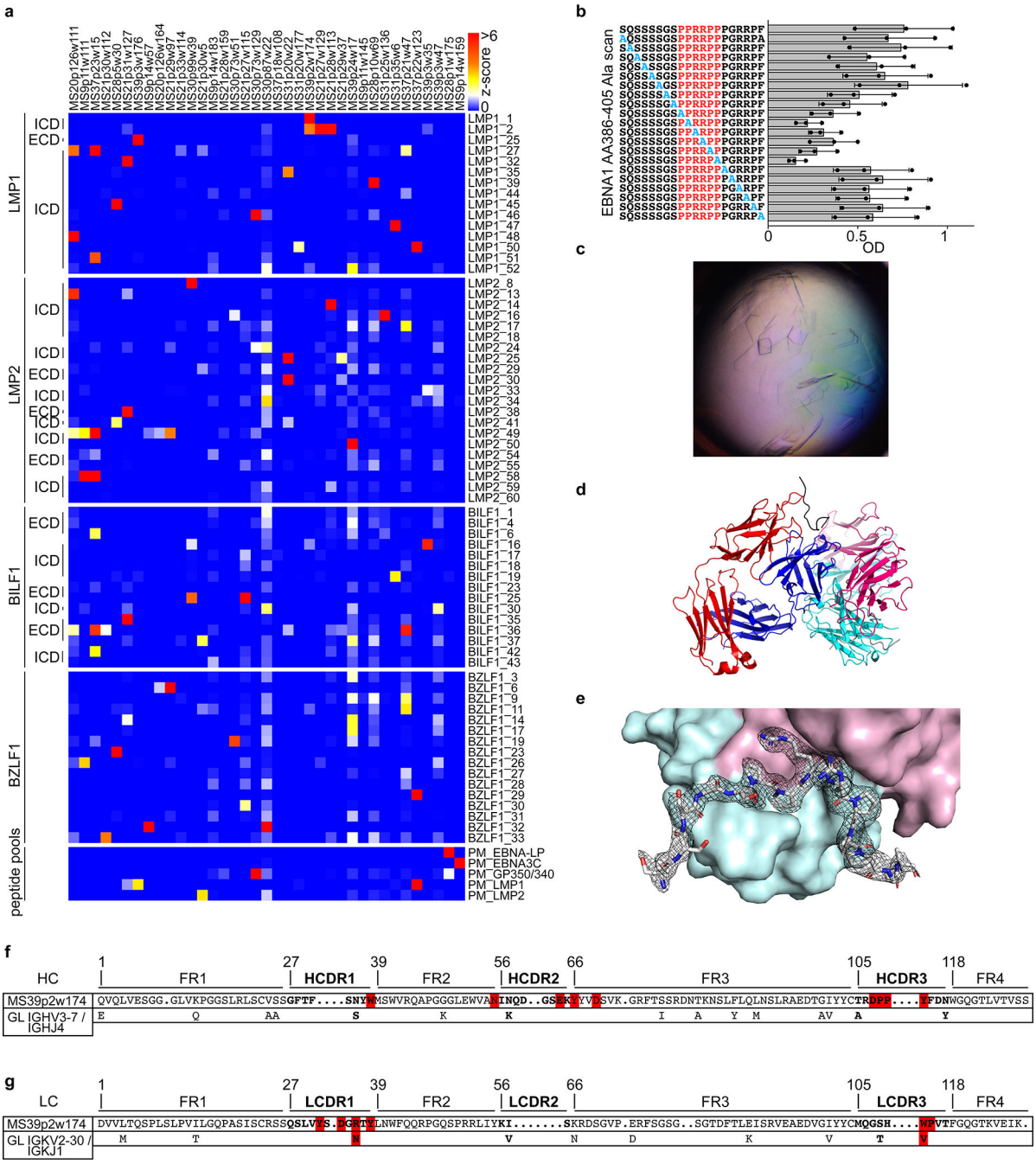
a, mAb reactivities to EBV virus lysates and recombinant EBV proteins as well as to other virus lysates. Z-scores for each antigen are shown, measurement of one microarray experiment, measured in 8 technical replicates. **b**, mAb reactivities to LPS, Insulin, and dsDNA to assess polyreactivity. Z-scores of area under the curve (AUC) of ELISA measurements at antibody concentrations of 0.1, 1, and 10 μg/ml are shown, each measurement was carried out in duplicates. **c**, mAb reactivities to GlialCAM proteins, peptides, and phosphorylated or citrullinated peptides. Mean reactivities (mean fluorescence intensity counts) are shown from one microarray experiment, measured in 8 technical replicates. Immediate early latency stage protein, IE; early, E; late, L; intracellular domain, ICD; extracellular domain, ECD; phosphorylated Serine, pSer; citrulline residue, Cit; _B

- _E: duplicate probes of same / similar lysates and proteins (different preparations or batches).



Extended Data Figure 6: MS CSF mAb reactivity to EBV peptides.

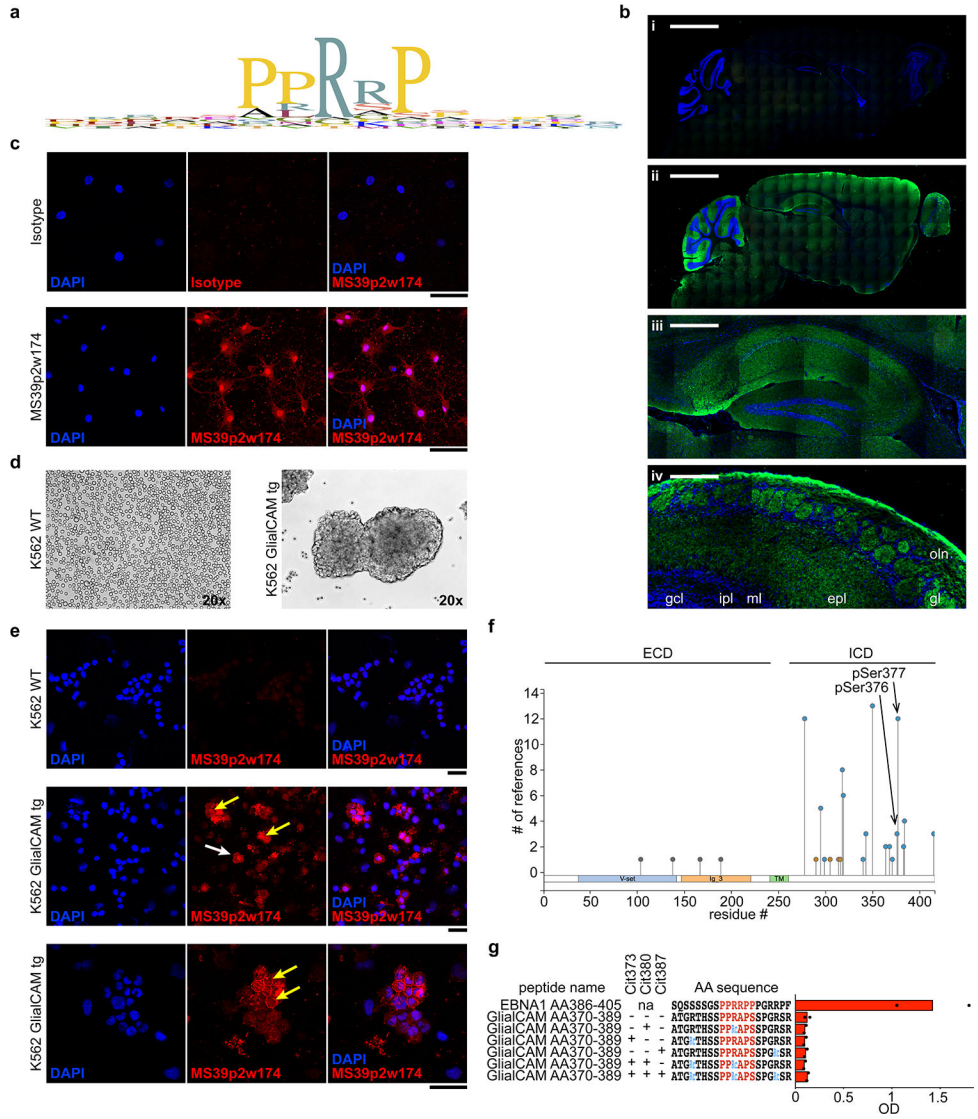
a, mAb reactivities to EBV peptides. Z-scores for each antigen are shown, measurement of one microarray experiment, measured in 8 technical replicates. Intracellular domain, ICD; extracellular domain, ECD; peptide mix, PM.



Extended Data Figure 7: mAb reactivity to EBV peptides and extended structural data for the EBNA1$_{AA386–405}$ / MS39p2w174-Fab complex.

a, mAb reactivities of selected reactive mAbs against the selected reactive peptide antigens. Z-scores for each antigen are shown, measurement of one microarray experiment, measured in 8 technical replicates. **b**, ELISA-based alanine-scan on EBNA1$_{AA386–405}$, corresponding to (Fig. 2e). Mean OD (450 nm) ± SD from three independent experiments, each carried out in triplicates. **c**, 20x image of protein crystals in hanging drop. **d**, Asymmetric unit containing two peptide-Fab complexes in a diagonal orientation, heavy chain (red/pink), light chain (blue/cyan), peptide (black/gray). **e**, EBNA1$_{AA386–405}$ peptide and its

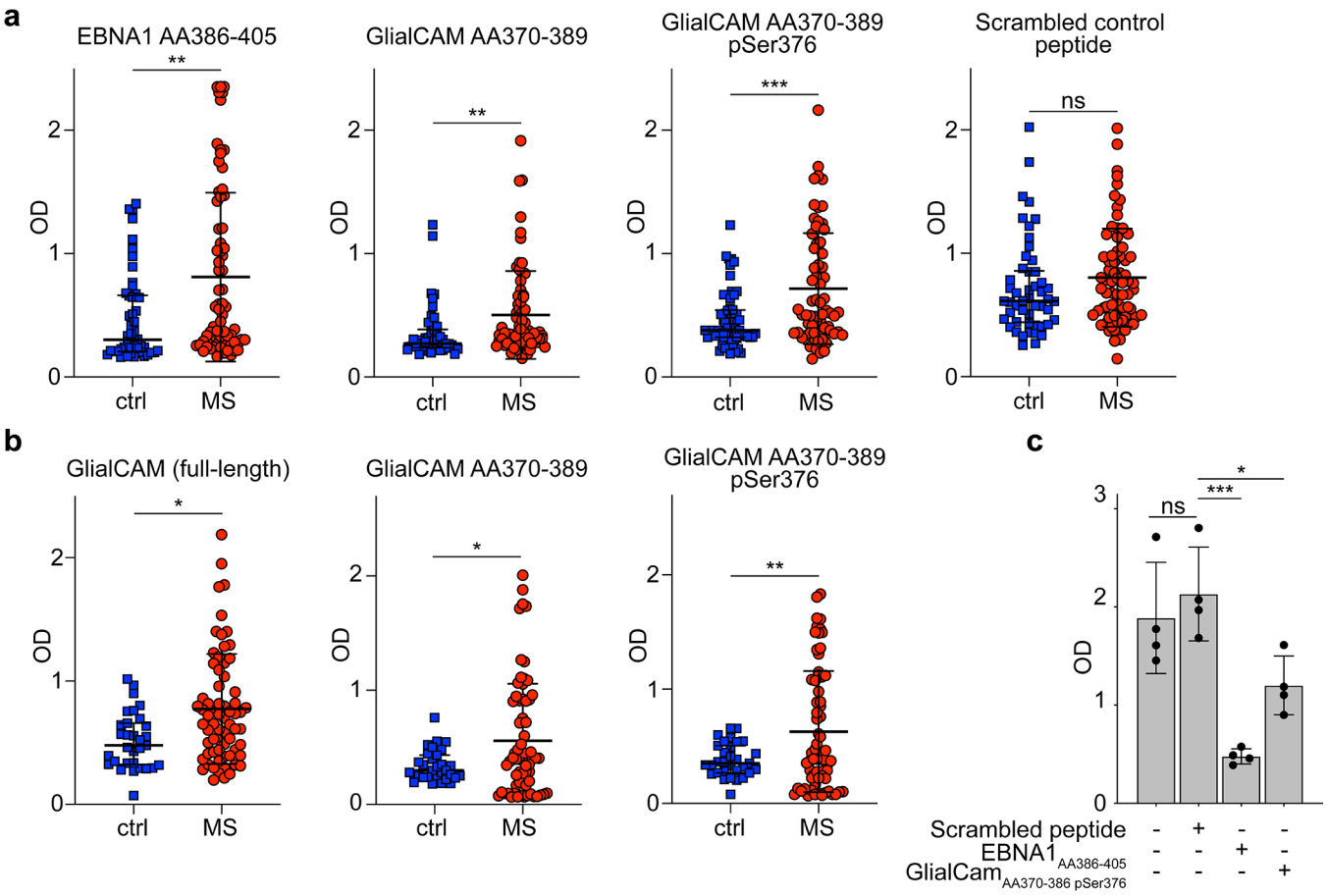
2mFoDFc map (contoured at 1σ) are shown, depicted on heavy chain (cyan) and light chain (pink) in surface representation. **f**,**g**, Amino acid sequences of variable regions of **f**, mAb MS39p2w174 heavy chain and **g**, light chain. Bold font: CDR, regular font framework regions. Of the germline variable genes (bottom rows), only residues that differ from MS39p2w174 sequence are shown, red: residues that closely interact with $EBNA1_{AA386–405}$ according to crystal structure. dots: gaps introduced during IMGT GapAlign for alignment and numbering purposes, numbers: residue numbers according to IMGT unique numbering. Intracellular domain, ICD; extracellular domain, ECD; heavy chain, HC; light chain, LC; complementarity determining region, CDR; framework region, FR; germline, GL.



Extended Data Figure 8: Extended characteristics of $GlialCAM_{AA370–389}$ and immunofluorescence stainings with MS39p2w174.

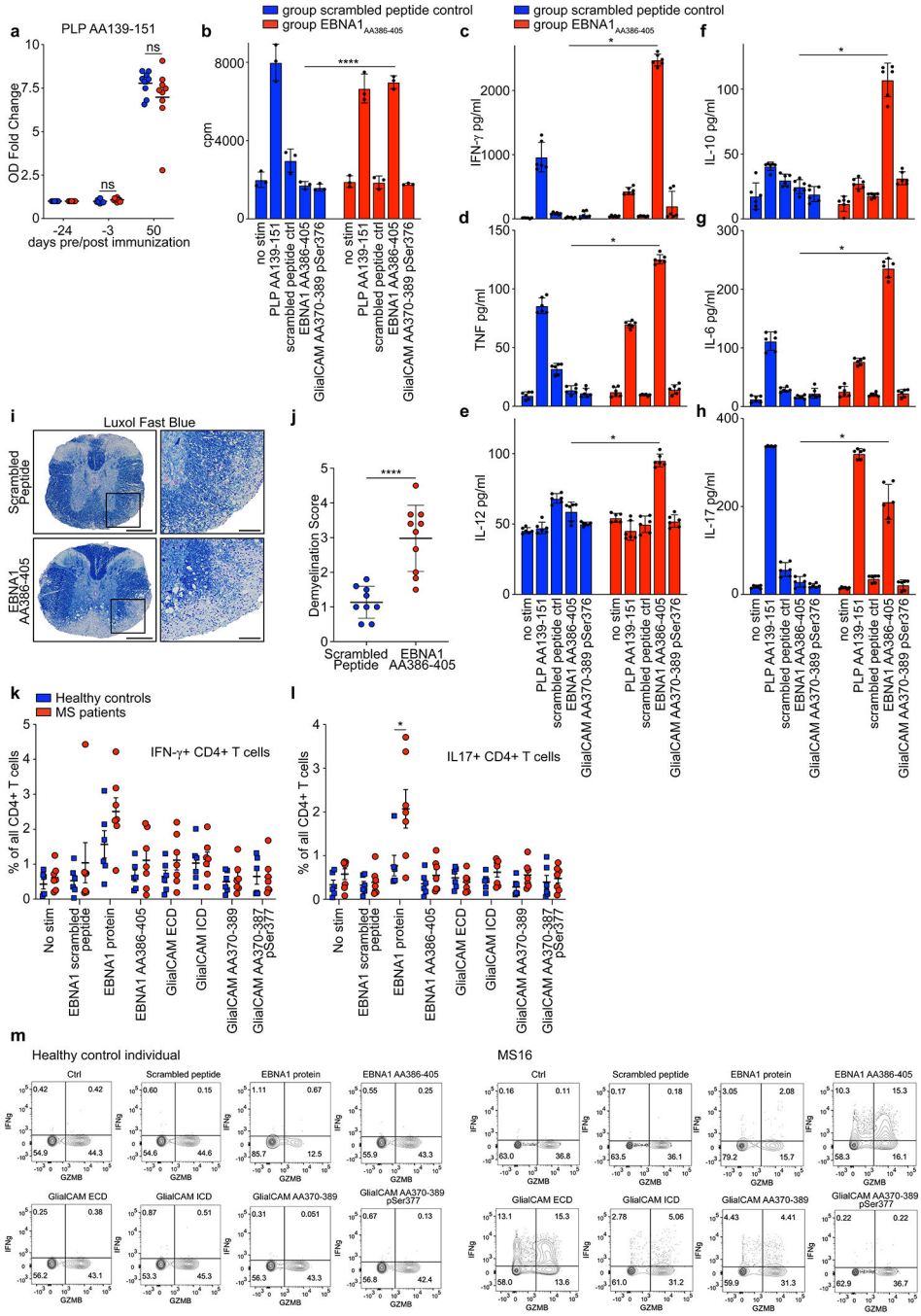
a, Expression pattern of GlialCAM in human tissues (proteinatlas.org). **b**, Phage display PhiP-Seq data, showing alignment of Pro/Arg-rich region and adjacent residues of all phage display peptides enriched above 100 / 10^5 reads. **c**, Immunofluorescence of mouse

brain slices stained with (i) control antibody, and (ii-iv) MS39p2w174 (green) and DAPI (blue). (i,ii) full brain, scale bars: 2000 μm, (iii) magnification of hippocampus with prominent MS39p2w174 staining, scale bar: 400 μm, and (iv) olfactory bulb with prominent MS39p2w174 staining in the olfactory nerve (oln), glomerular (gl), and external plexiform layers (epl), but not the mitral (ml), internal plexiform (ipl), or granule cell (gcl) layers, scale bar: 100 μm. **d**, Immunofluorescence staining of primary rat oligodendrocytes with isotype control antibody (top panel) and MS39p2w174 (bottom panel). **e**, K562 cells in culture, wildtype (left) and transduced with full-length GlialCAM (right). **f**, Immunofluorescence with MS39p2w174 on WT K562 cells (top) and GlialCAM-tg K562 cells (center and bottom). White arrow: single K562 cell, orange arrow: high intensity MS39p2w174 staining on the cell boarder between transgenic K562 cells in bulks. **d**,**f**, Scale bars: 40 μm. **c-f**, representative micrographs of at least two experiments. **g**, Overview of phosphorylated residues in GlialCAM, identified by mass spectrometry (phosphoSite.org). The two phosphorylated serine residues of interest are indicated with arrows. **h**, ELISA, measuring binding of MS39p2w174 to native and citrullinated GlialCAM$_{AA370–389}$ peptides, means of *n=2* independent experiments, each carried out in triplicates. Wildtype, WT; extracellular domain, ECD; intracellular domain, ICD; phosphorylated serine, pSer; citrulline residue, Cit.



Extended Data Figure 9: Plasma reactivity against EBNA1 and GlialCAM proteins and peptides in healthy control individuals and MS patients.

a, ELISA measurement of antigen-specific IgG reactivity against peptides $EBNA1_{AA386–405}$, $GlialCAM_{AA370–389}$, phosphorylated $GlialCAM_{AA370–389\ pSer376}$, 2x-phosphorylated $GlialCAM_{AA370–389\ pSer376\ pSer377}$, and scrambled peptide control in plasma samples of healthy control individuals (*n=50*) and MS patients (*n=71*). Means ± SD in each patient group is shown. Representative OD (450 nm) measurements of two independent experiments, each carried out in duplicates. ** $P < 0.01$, *** $P < 0.001$ according to two-tailed Mann-Whitney test, Tukey corrected for multiple comparisons. **b**, ELISA measurements of antigen-specific IgG reactivity against GlialCAM full-length protein, $GlialCAM_{AA370–389}$, and phosphorylated $GlialCAM_{AA370–389\ pSer376}$ in plasma samples of a separate cohort of healthy control individuals (*n=31*) and MS patients (*n=67*). Means ± SD across patient groups are shown. Representative OD (450 nm) measurements of two independent experiments, each carried out in duplicates. * $P < 0.05$, ** $P < 0.01$ according to two-tailed Mann-Whitney test, Tukey corrected for multiple comparisons. **c**, ELISA measurements of mAB MS39p2w174 binding to $EBNA1_{AA386–405}$, without interference as well as blocked with scrambled peptide control, $EBNA1_{AA386–405}$, and $GlialCAM_{AA370–389\ pSer376}$, as a positive control to (Fig. 3q). Mean OD (450 nm) ± SD of quadruplicate measurements from *n=1* experiment are shown. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ according to one-way ANOVA, Tukey corrected for multiple comparisons.



Extended Data Figure 10: T cell response against EBNA$_{AA386–405}$.

a, ELISA data showing mouse plasma IgG responses against PLP$_{AA139–151}$ at the indicated timepoints pre and post EAE induction, for scrambled peptide immunized mice (blue, *n=10*) and EBNA1$_{AA386–405}$ immunized mice (red, *n=10*). Mean OD (450 nm) fold change ± SD, significance levels according to unpaired two-tailed Mann-Whitney test. Means ± SD, representative of three independent experiments, each carried out as triplicate measurements. **b**, T cell proliferation measurement by ^{3}H-thymidine incorporation in splenocytes and lymph node cells of mice immunized with scrambled peptide (blue) and EBNA1$_{AA386–405}$

(red). Cells from *n=10* mice per group were pooled and mean counts per minute (cpm) ± SD of triplicate measurements are shown. **c-h**, ELISA measurements of cytokines in cell culture supernatant of mouse splenocytes and lymph node cells of mice immunized with scrambled peptide (blue) or $EBNA1_{AA386–405}$ (red) and re-stimulated with the indicated peptides. Cells from *n=10* mice per group were pooled and mean cpm ± SD of six replicate measurements are shown. ****P = 8.9×10^{-5}, unpaired two-tailed Student's t test, Holm-Sidak corrected for multiple comparisons. **c**, IFN-γ, **d**, TNF, **e**, IL-12, **f**, IL-10, **g**, IL-6, **h**, IL-17, *$P <$ 0.05, significance levels according to unpaired two-tailed Mann-Whitney test, Holm-Sidak corrected for multiple comparisons. **i**, Representative Luxol Fast Blue stained spinal cords from scrambled peptide group (top panel) and $EBNA1_{AA386–405}$ group (bottom panel). Scale bars left images: 200 μm, right images: 50 μm**. j**, Statistical evaluation of Luxol Fast Blue scores, means of at least 4 coronal spinal cord sections per mouse and means ± SD for each group (*n=9*) are shown. ****$P < 0.0001$, unpaired two-tailed Mann-Whitney test. **k**,**l**, Flow cytometry data of PBMC from healthy control individuals (*n=6*, blue) and MS patients (*n=7*, red), showing percent of **k**, IFN-γ+ and **l**, IL-17+ CD4+ T cells in all CD4+ T cells. Mean MFI ± SEM are shown for the respective groups. Significance levels were assessed by two-way ANOVA, followed by FDR calculation using the two-stage step-up method of Benjamini, Krieger and Yekutieli, * FDR < 0.1. **m**, Flow cytometry data, representative dot plots are shown for two individuals from the data set presented in Fig. 4f. Healthy control individual (left) and MS patient MS16 (right). Expression levels of Granzyme-B (GZMB) and IFN-γ are presented under the indicated stimulations.

Extended Data Table 1:

Patient collective.

	patient	sex	age range	Origin	Diagnosis	LP indication	MS Treatment	CSF cells/ ul	OCB
1	MS9	female	41–45	Stanford	RRMS	diagnostic	none	12	+
2	MS12	female	16–20	Heidelberg	CIS	diagnostic	none	11	+
3	MS20	male	21–25	Heidelberg	CIS	diagnostic	none	10	+
4	MS21	female	36–40	Heidelberg	RRMS	diagnostic	Fingolimod (paused for 4 weeks)	15	+
5	MS28	female	26–30	Heidelberg	CIS	diagnostic	none	17	+
6	MS30	female	21–25	Heidelberg	RRMS	diagnostic	none	48	+
7	MS31	female	21–25	Heidelberg	CIS	diagnostic	none	10	+
8	MS37	female	16–20	Heidelberg	CIS	diagnostic	none	57	+
9	MS39	female	26–30	Stanford	RRMS	diagnostic	none	18	+
10	C3	male	60–65	Heidelberg	neuroborreliosis	diagnostic	none	825	+
11	C5	female	45–50	Heidelberg	neuro-Behçet's	diagnostic	none	56	+
12	C6	male	75–80	Heidelberg	viral encephalitis	diagnostic	none	48	+

Extended Data Table 2:

Phage display.

	gene	RefSeq ID	peptide sequence	percent enrichment
1	CCDC88C	NP_001073883.2	RGSPHRGSLDRTDASTDLAMRSWPSELGSRTCSTSATTTAPSNSTPIAR	2.31%
2	PTPRU	NP_005695.3	RPGDGGTG**RPGPP**LISRTKCAEPMRAPKGLAFAEIQARQLTLQWEPLGY	2.01%
3	WNT4	XP_011539901.1	RSGERTQRRKEGGSLGQCWPEAIVGTQD**PMRSPP**YLWGRNQGGTHTGSA	1.10%
4	YLPM1	NP_062535.2	RERGLGRSDFGRDRGPFRPEPGDGGEKMYPYHRDE**PPRAP**WNHGEERGH	0.91%
5	ATG16L1	XP_011526107.1	RRRSVSSFPVPQDNVDTHPGSGKEVRVPATALCVFDAHDGEVNAVQFSP	0.87%
6	MAST1	XP_011526107.1	RFSALLEPSRFSAPQEDEDEARL**RRPPRP**SSDPAGSLDARAPKEETQGE	0.72%
7	AFAP1	XP_006713972.1	ALRNRLAQLRKERKDLRAAIEVNAGRKPQAILEEKLKQLEEECRQKEAE	0.64%
8	PMS2	XP_011513729.1	**PRRSP**LGQKRGMLSSSTSGAISDKGVLRPQKEAVSSSHGPSDPTDRAEV	0.61%
9	EIF3A	NP_003741.1	DRPSWRNTDDD**RPPRR**IADEDRGNWRHADDD**RPPRR**GLDEDRGSWRTAD	0.56%
10	CDC42BPG	NP_059995.2	PEEKGRVARGSG**PQRP**HSFSEAL**RRP**ASMGSEGLGGDADPMKRKPWTSL	0.54%
11	GRIN3B	XP_011525971.1	SGTAHVTGLGACFRPE**RPPRRP**HHTLADRGARNEAAAGPFGLKTHTQRS	0.53%
12	FAM21A	XP_006717894.2	GK**RRP**QTRAARRLAAQESSETEDMSV**PRGP**IAQWADGAISPNGHRPQLR	0.47%
13	LOC102723743	XP_006715874.1	RGNPQGIPETGLWSNPQGIPETALRGNPQGKAANSAGARRCGVT**PRGPR**	0.47%
14	WDR62	XP_011525141.1	AQ**PLRRP**SSVGELASLGQELQAITTATTPSLDSEGQEPALRSWGNHEAR	0.44%
15	PLEKHA6	XP_006711276.2	ASYGRQDATVWIPSPSRQPVYYDELDAASSSLRRLSLQPRSHSV**PRSP**S	0.43%
16	RBMXL1	NP_062556.2	GLVRSSSGMGGRAPLSRGRDSYGG**PPRREP**LPSRRDVYLSPRDDGYSTK	0.38%
17	GLIALCAM	NP_689935.2	PPGYSVSPAVPGRSPGLPIRSARRYPRSPARSPATGRTHSS**PPRAP**SSP	0.36%
18	MAEL	NP_001273306.1	RGLPVARVADAIPYCSSDWAKPVFT**PLRRP**GMLVPKQNVSPPDMSALSL	0.35%
19	PPP1R26	NP_055626.3	**RGPP**VLKSCLSKSKRDSGEGPGKKPPSVFGSTAERMRQEGAASQDAALA	0.34%
20	MAPK8IP2	XP_011528981.1	E**PPRRP**AFLPVGPDDTNSEYESGSESEPDLSEDADSPWLLSNLVSRMIS	0.33%
21	JPH4	NP_115828.2	AKLIAQDLQPMLEA**PGRRPR**QDSEGSDTEPLDEDSPGVYENGLTPSEGS	0.31%
22	ZAN	NP_775082.2	CESPCLQNPCQNDGQCREQGATFTCECEVGYGGGLCMEPRDA**PPPRKP**A	0.29%
23	CDH18	XP_005248285.1	MAGQVGGLSGSTTVNITLTDVNDN**PPRFP**QKHYQLYVPESAQVGSAVGK	0.29%
24	ZAN	XP_011514857.1	PCQNDGQCREQGATFTCECEVGYGGGLCMEPRDA**PPPRKP**EASNLVGVL	0.28%
25	ZNF324B	XP_011525261.1	PREKTFTEYRVPGRQ**PRTP**ERQKPCAQEVPGRAFGNASDLKAASGGRDR	0.27%
26	ZC3H12D	NP_997243.2	LQLQPRGEHRPRDLHGDLLS**PRRPP**DDPWA**RPPR**SDRFPGRSVWAEPAW	0.26%
27	KIAA0226	XP_006713890.1	GSERRSTSFPLSG**PPRKP**QESRGHVSPAEDQTIQAPPVSVSALARDSPL	0.26%
28	MYO7A	XP_006718622.1	EEEGNRALIKHRDSSLGLWERAMALDTCPGEG**RRAAP**SALSQLSVLAGV	0.24%
29	CHTF18	XP_005255528.1	NPVL**RRPP**ILEDYVHVTSTEGVRAYLVLRADPMAPGVQGSLLHVPWRGG	0.24%
30	CDC42BPG	XP_011543457.1	VARGSGPQRPHSFSEAL**RRP**ASMGSEGLGGDADPMKRKPWTSLSSESVS	0.23%
31	ABL1	NP_005148.2	**PLRR**QVTVAPASGLPHKEEAGKGSALGTPAAAEPVTPTSKAGSGAPGGT	0.21%
32	TCF20	XP_006724376.1	YRGNASPGAATHDSLSDYGPQDSRPT**PMRRVP**GRVGGREGMRGRSPSQY	0.20%
33	C9orf50	XP_011516963.1	**PPRSPP**GRPHGAQVPRLKAALTHNPSGEGSRPCRQRCPFRVRFADETLQ	0.20%
34	YLPM1	NP_062535.2	GSRERI**PPRR**AGSRERGPPRGPGSRERGLGRSDFGRDRGPFRPEPGDGG	0.18%
35	PHF8	XP_005262057.1	PATSSLQAWWTGGQDRSSGSSSSGLGTVSNSPASQRTPGKR**PIKRP**AYW	0.18%
36	WDR62	XP_005258866.1	EGPIVATLAQ**PLRRP**SSVGELASLGQELQAITTATTPSLDSEGQEPALR	0.18%

	gene	RefSeq ID	peptide sequence	percent enrichment
37	LPPR3	NP_001257295.1	RPVAREKTSLGSLKRASVDVDLLA**PRSP**MAKENMVTFSHTLPRASAPSL	0.17%
38	EGLN1	XP_005273224.1	GKEE**PPAR**SSLFQEKANLYPPSNTPGDALSPGGGLRPNGQTKPLPALKL	0.17%
39	CHD4	XP_006719022.1	ASEEGDEDFDERSEA**PRRP**SRKGLRNDKDK**PLPP**LLARVGGNIEVLGFN	0.17%
40	SCNN1B	XP_011544215.1	RAQASYAG**PPP**TVAELVEAHTNFGFQPDTA**PRSP**NTGPYPSEQALPIPG	0.17%
41	PTPRM	NP_001098714.1	GGTGSPGPALRTRTKCADPMRGPRKLEVVEVKSRQITIRWEPFGYNVTR	0.16%
42	EIF3A	NP_003741.1	GMDDDRG**PRRGP**EEDRFSRRGADDDRPSWRNTDDD**RPPRR**IADEDRGNW	0.16%
43	MORN1	XP_005244855.1	MAAAGEGTPSSRGPRRD**PPRRPP**RNGHGKLLFKDGSYYEGAFVDGEITG	0.16%
44	SFMBT2	NP_001025051.1	**PVRRPPP**ERTRRGRGAPAASSAEEGEKCPPTKPEGTEDTKQEEEERLVL	0.16%
45	TBC1D1	XP_011511971.1	MKPQREHADRGPWEPHHIEPHCGSPSLACTEDYSELGEL**PPRSP**LEPVC	0.16%
46	FAM114A1	XP_005262729.1	VGHGLTAVKEKAGATLRIHGVNSGSSEGAQPNTENGVPEITDAATDQGP	0.16%
47	TRIP10	XP_006723003.1	QPMNRAPSDSSLGTPSDGRPELRGPGRSRTKRWPFGKKNK**PRPPP**LSPL	0.16%
48	IRF3	XP_006723260.1	WPVTLPDPGMSLTDRGVMSYVRHVLSCLGGGLALWRAGQWLWAQRLGHC	0.16%
49	REEP4	NP_079508.2	DSDTEDECWSDTEAV**PRAPARPR**EKPLIRSQSLRVVKRKPPVREGTSRS	0.15%
50	WDR62	XP_011525141.1	ASSRARISRSISLGDSEGPIVATLAQ**PLRRP**SSVGELASLGQELQAITT	0.15%
51	PLEKHA4	XP_011525459.1	QTLS**RPPTPRR**GPPSEAGGGK**PPRSP**QHWSQEPRTQVSREAHSGSPTYL	0.15%
52	NOM1	NP_612409.1	RLQRTAGPEQGPGLGGRSGAEEASGHRQDTEERARPAPSRD**PSPPR**KPR	0.14%
53	UPB1	XP_011528524.1	REDFELQGYAFEAAEEQL**RRPR**IVHVGLVQNRIPLPANAPVAEQVSALH	0.14%
54	PKD1	XP_011520832.1	GSGVVYTWSLEEGLSWETSEPFTTHSFPTPGLHLVTMTAGNPLGSANAT	0.14%
55	MAP3K1	XP_011541708.1	GNSPSGRTVKSESPGVRRKRVSPVPFQSGRIT**PPRRAP**SPDGFSPYSPE	0.14%
56	WNT4	XP_011539901.1	RTQRRKEGGSLGQCWPEAIVGTQDPM**RSPP**YLWGRNQGGTHTGSASEGH	0.13%
57	YLPM1	NP_062535.2	GEKMYPYHRDE**PPRAP**WNHGEERGHEEFPLDGRNAPMERERLDDWDRER	0.13%
58	RBL2	NP_005602.3	PPASTTRRRLFVENDSPSDGGT**PGRMPPQP**LVNAVPVQNVSGETVSVTP	0.13%
59	LOC105376336	XP_011547586.1	VGETSRESASSPQDGGRHHRDS**PRRAP**SPTRKTWGDGAGRAKRAGVGCG	0.13%
60	IQCA1L	XP_011514509.1	YDSETLREEKRKEVELEIRIQVDELMRQELRKLRLAVDKEEER**PLRAP**K	0.12%
61	ATG16L1	XP_011547155.1	DNVDTHPGSGKEVRVPATALCVFDAHDGEVNAVQFSPGSRLLATGGMDR	0.12%
62	ZNF536	XP_011525856.1	GSGSDQESQSVSRSTTPGSSNVTEESGVGGGLSQTGSAQEDSPHPSSPS	0.12%
63	TCP10L2	NP_001138593.1	SLKTERINSGKTPPQEDREKS**PPGRR**QDRSPA**PTGRP**TPGAERREVSED	0.12%
64	CDH6	XP_011512223.1	MGGQMGGLSGTTTVNITLTDVNDN**PPRFP**QSTYQFKTPESSPPGTPIGR	0.12%
65	MAPK8IP2	XP_011528982.1	SPASE**PEPPREPPRRP**AFLPVGPDDTNSEYESGSESEPDLSEDADSPWL	0.12%
66	BRD1	NP_001291738.1	IGLEEASGMHLPERPAAA**PRRP**FSWEDVDRLLDPANRAHLGLEEQLREL	0.12%
67	ABCC6	NP_001162.4	TEPGTSTKDPRGTSAG**RRP**ELRRERSIKSVPEKDRTTSEAQTEVPLDDP	0.11%
68	FAM109B	XP_005261430.1	QRRSSWKSVASRCKPQAPNHRAAGLENGHCLSKDSSPVGLVEEAGSRSA	0.11%
69	MAPK8IP3	XP_011520731.1	TEMIQTYVEHIERSKMQQVGGNSQTESSL**PGRSPR**QSWRKSRKERPTSL	0.11%
70	TJP1	XP_006725617.1	NVPDLSDSIHSANASERDDISEIQSLASDHSGRSHD**RPPRRSR**SRSPDQ	0.11%
71	OPALIN	NP_001035192.1	SIEAMEESDRPCEISEIDDNPKISEN**PRRSP**THEKNTMGAQEAHIYVKT	0.11%
72	COL4A6	XP_011529155.1	GNPGPVGI**PSPRRP**MSNLWLKGDKGSQGSAGSNGFPGPRGDKGEAGRPG	0.11%
73	FMNL1	XP_006722125.1	TAMGPS**RRPPEP**EKAPPA**APTRP**SALELKVEELEEKGLIRILRGPGDAV	0.11%
74	EIF3A	NP_003741.1	WRHADDD**RPPRR**GLDEDRGSWRTADEDRGPRRGMDDDRGPRRGGADDER	0.10%

	gene	RefSeq ID	peptide sequence	percent enrichment
75	RIMBP2	XP_011536404.1	VPDDVEVYLSDAPSHYSQDTPMRSKAK**RVPP**EGSGTARRAPSPTVHLHS	0.10%
76	LSR	NP_001247419.1	GPPSGRVERAMSEVTSLHEDDWRS**RPSRGP**ALTPIRDEEWGGHS**PRSPR**	0.10%

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank Wolfgang Wick for support with sample collection, Ben Bell for advice on crystallization, George Harauz for insightful discussions on poly-proline motifs, and the staff at SSRL for assistance with data collection.

Funding:

This work was supported by NIH R01 AR063676 and U19 AI110491 to W.H.R., the Juvenile Diabetes Research Foundation and Lupus Research Alliance Funding to W.H.R. and T.V.L., the German Research Foundation (DFG, LA3657/1) to T.V.L., Atara to L.S. and P.P.H., and the German Research Foundation to M.P. (DFG, project 406052676; PL-315/5–1). The mass spectrometric experiments are in part supported by the NYU Grossman School of Medicine and a shared instrumentation grant NIH 1S10OD010582–01A1. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02–76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by NIH GIGMS P30GM133894. The contents of this publication are solely the responsibility of the authors and do not represent the official views of NIGMS or NIH.

Data availability statement:

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.

Main Text References

1. Cencioni MT, Mattoscio M, Magliozzi R, Bar-Or A & Muraro PA B cells in multiple sclerosis - from targeted depletion to immune reconstitution therapies. Nat. Rev. Neurol. 17, 399–414 (2021). [PubMed: 34075251]
2. Hauser SL et al. Ocrelizumab versus Interferon Beta-1a in Relapsing Multiple Sclerosis. N. Engl. J. Med. 376, 221–234 (2017). [PubMed: 28002679]
3. Bar-Or A et al. Epstein–Barr Virus in Multiple Sclerosis: Theory and Emerging Immunotherapies. Trends Mol. Med. 26, 296–310 (2020). [PubMed: 31862243]
4. Jarius S et al. The MRZ reaction as a highly specific marker of multiple sclerosis: re-evaluation and structured review of the literature. J. Neurol. 264, 453–466 (2017). [PubMed: 28005176]
5. Wang Z et al. Antibodies from Multiple Sclerosis Brain Identified Epstein-Barr Virus Nuclear Antigen 1 & 2 Epitopes which Are Recognized by Oligoclonal Bands. J. Neuroimmune Pharmacol 16, 567–580 (2021). [PubMed: 32808238]
6. Thacker EL, Mirzaei F & Ascherio A Infectious mononucleosis and risk for multiple sclerosis: a meta-analysis. Ann. Neurol. 59, 499–503 (2006). [PubMed: 16502434]
7. Nielsen TR et al. Effects of infectious mononucleosis and HLA-DRB1*15 in multiple sclerosis. Multiple Sclerosis Journal vol. 15 431–436 (2009). [PubMed: 19153174]

8. Tengvall K et al. Molecular mimicry between Anoctamin 2 and Epstein-Barr virus nuclear antigen 1 associates with multiple sclerosis risk. Proc. Natl. Acad. Sci. U. S. A. 116, 16955–16960 (2019). [PubMed: 31375628]
9. Ruprecht K et al. Multiple sclerosis: The elevated antibody response to Epstein–Barr virus primarily targets, but is not confined to, the glycine–alanine repeat of Epstein–Barr nuclear antigen-1. Journal of Neuroimmunology vol. 272 56–61 (2014). [PubMed: 24798244]
10. Jafari N et al. No evidence for intrathecal IgG synthesis to Epstein Barr virus nuclear antigen-1 in multiple sclerosis. J. Clin. Virol. 49, 26–31 (2010). [PubMed: 20638898]
11. Salzer J et al. Epstein-Barr virus antibodies and vitamin D in prospective multiple sclerosis biobank samples. Mult. Scler. 19, 1587–1591 (2013). [PubMed: 23549431]
12. Sundqvist E et al. Epstein-Barr virus and multiple sclerosis: interaction with HLA. Genes Immun. 13, 14–20 (2012). [PubMed: 21776012]
13. Ramesh A et al. A pathogenic and clonally expanded B cell transcriptome in active multiple sclerosis. Proc. Natl. Acad. Sci. U. S. A. 117, 22932–22943 (2020). [PubMed: 32859762]
14. Kowarik MC et al. Immune cell subtyping in the cerebrospinal fluid of patients with neurological diseases. J. Neurol. 261, 130–143 (2014). [PubMed: 24162037]
15. Tan Y-C et al. Barcode-enabled sequencing of plasmablast antibody repertoires in rheumatoid arthritis. Arthritis rheumatol. 66, 2706–2715 (2014). [PubMed: 24965753]
16. Palanichamy A et al. Immunoglobulin class-switched B cells form an active immune axis between CNS and periphery in multiple sclerosis. Sci. Transl. Med. 6, 248ra106 (2014).
17. von Büdingen H-C et al. B cell exchange across the blood-brain barrier in multiple sclerosis. J. Clin. Invest. 122, 4533–4543 (2012). [PubMed: 23160197]
18. Rounds WH et al. MSPrecise: A molecular diagnostic test for multiple sclerosis using next generation sequencing. Gene 572, 191–197 (2015). [PubMed: 26172868]
19. Bankoti J et al. In multiple sclerosis, oligoclonal bands connect to peripheral B-cell responses. Ann. Neurol. 75, 266–276 (2014). [PubMed: 24375699]
20. Mescheriakova JY, van Nierop GP, van der Eijk AA, Kreft KL & Hintzen RQ EBNA-1 titer gradient in families with multiple sclerosis indicates a genetic contribution. Neurol Neuroimmunol Neuroinflamm 7, (2020).
21. Berman HM et al. The Protein Data Bank. Acta Crystallogr. D Biol. Crystallogr 58, 899–907 (2002). [PubMed: 12037327]
22. Jeong JS et al. Rapid Identification of Monospecific Monoclonal Antibodies Using a Human Proteome Microarray *. Mol. Cell. Proteomics 11, (2012).
23. Uhlén M et al. Proteomics. Tissue-based map of the human proteome. Science 347, 1260419 (2015). [PubMed: 25613900]
24. Gilbert A, Vidal XE, Estevez R, Cohen-Salmon M & Boulay A-C Postnatal development of the astrocyte perivascular MLC1/GlialCAM complex defines a temporal window for the gliovascular unit maturation. Brain Struct. Funct 224, 1267–1278 (2019). [PubMed: 30684007]
25. Favre-Kontula L et al. GlialCAM, an immunoglobulin-like cell adhesion molecule is expressed in glial cells of the central nervous system. Glia 56, 633–645 (2008). [PubMed: 18293412]
26. López-Hernández T et al. Mutant GlialCAM causes megalencephalic leukoencephalopathy with subcortical cysts, benign familial macrocephaly, and macrocephaly with retardation and autism. Am. J. Hum. Genet. 88, 422–432 (2011). [PubMed: 21419380]
27. Han MH et al. Proteomic analysis of active multiple sclerosis lesions reveals therapeutic targets. Nature 451, 1076–1081 (2008). [PubMed: 18278032]
28. O'Donovan B et al. High-resolution epitope mapping of anti-Hu and anti-Yo autoimmunity by programmable phage display. Brain Commun 2, fcaa059 (2020). [PubMed: 32954318]
29. Bochkarev A, Bochkareva E, Frappier L & Edwards AM The 2.2 A structure of a permanganate-sensitive DNA site bound by the Epstein-Barr virus origin binding protein, EBNA1. J. Mol. Biol. 284, 1273–1278 (1998). [PubMed: 9878348]
30. Casciola-Rosen LA, Anhalt G & Rosen A Autoantigens targeted in systemic lupus erythematosus are clustered in two populations of surface structures on apoptotic keratinocytes. Journal of Experimental Medicine vol. 179 1317–1330 (1994). [PubMed: 7511686]

31. Pisetsky DS & Lipsky PE New insights into the role of antinuclear antibodies in systemic lupus erythematosus. Nat. Rev. Rheumatol. 16, 565–579 (2020). [PubMed: 32884126]
32. Schellekens GA et al. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. Arthritis Rheum. 43, 155–163 (2000). [PubMed: 10643712]
33. Moh MC, Zhang C, Luo C, Lee LH & Shen S Structural and functional analyses of a novel ig-like cell adhesion molecule, hepaCAM, in the human breast carcinoma MCF7 cells. J. Biol. Chem. 280, 27366–27374 (2005). [PubMed: 15917256]
34. Britton D et al. Quantification of pancreatic cancer proteome and phosphorylome: indicates molecular events likely contributing to cancer and activity of drug targets. PLoS One 9, e90948 (2014). [PubMed: 24670416]
35. Herskowitz JH et al. Phosphoproteomic analysis reveals site-specific changes in GFAP and NDRG2 phosphorylation in frontotemporal lobar degeneration. J. Proteome Res. 9, 6368–6379 (2010). [PubMed: 20886841]
36. Hornbeck PV, Chabra I, Kornhauser JM, Skrzypek E & Zhang B PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. Proteomics 4, 1551–1561 (2004). [PubMed: 15174125]
37. Zavala-Cerna MG et al. The clinical significance of posttranslational modification of autoantigens. Clin. Rev. Allergy Immunol. 47, 73–90 (2014). [PubMed: 24840362]
38. Bashford-Rogers RJM, Smith KGC & Thomas DC Antibody repertoire analysis in polygenic autoimmune diseases. Immunology 155, 3–17 (2018). [PubMed: 29574826]
39. Obermeier B et al. Matching of oligoclonal immunoglobulin transcriptomes and proteomes of cerebrospinal fluid in multiple sclerosis. Nat. Med. 14, 688–693 (2008). [PubMed: 18488038]
40. Singh V et al. Cerebrospinal-fluid-derived Immunoglobulin G of Different Multiple Sclerosis Patients Shares Mutated Sequences in Complementarity Determining Regions *. Mol. Cell. Proteomics 12, 3924–3934 (2013). [PubMed: 23970564]
41. Colombo M et al. Accumulation of clonally related B lymphocytes in the cerebrospinal fluid of multiple sclerosis patients. J. Immunol. 164, 2782–2789 (2000). [PubMed: 10679121]
42. Terzoglou AG, Routsias JG, Avrameas S, Moutsopoulos HM & Tzioufas AG Preferential recognition of the phosphorylated major linear B-cell epitope of La/SSB 349–368 aa by anti-La/SSB autoantibodies from patients with systemic autoimmune diseases. Clin. Exp. Immunol. 144, 432–439 (2006). [PubMed: 16734612]
43. International Multiple Sclerosis Genetics Consortium et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476, 214–219 (2011). [PubMed: 21833088]
44. Thompson AJ et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol. 17, 162–173 (2018). [PubMed: 29275977]
45. Polman CH et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann. Neurol. 69, 292–302 (2011). [PubMed: 21387374]
46. Wingerchuk DM et al. International consensus diagnostic criteria for neuromyelitis optica spectrum disorders. Neurology 85, 177–189 (2015). [PubMed: 26092914]
47. Tan Y-C et al. High-throughput sequencing of natively paired antibody chains provides evidence for original antigenic sin shaping the antibody response to influenza vaccination. Clin. Immunol. 151, 55–65 (2014). [PubMed: 24525048]
48. Blum LK et al. Circulating plasmablasts are elevated and produce pathogenic anti-endothelial cell autoantibodies in idiopathic pulmonary arterial hypertension. Eur. J. Immunol. 48, 874–884 (2018). [PubMed: 29369345]
49. Edgar RC UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods 10, 996–998 (2013). [PubMed: 23955772]
50. Alamyar E, Duroux P, Lefranc M-P & Giudicelli V IMGT(®) tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. Methods Mol. Biol. 882, 569–604 (2012). [PubMed: 22665256]
51. Edgar RC MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797 (2004). [PubMed: 15034147]

52. Guindon S et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321 (2010). [PubMed: 20525638]
53. Huerta-Cepas J, Serra F & Bork P ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. Mol. Biol. Evol. 33, 1635–1638 (2016). [PubMed: 26921390]
54. Bern M, Kil YJ & Becker C Byonic: advanced peptide and protein identification software. Curr. Protoc. Bioinformatics Chapter 13, Unit13.20 (2012).
55. Robinson WH et al. Autoantigen microarrays for multiplex characterization of autoantibody responses. Nat. Med. 8, 295–301 (2002). [PubMed: 11875502]
56. Kuerten S Autoantibodies against central nervous system antigens in a subset of B cell–dominant multiple sclerosis patients. Proc. Natl. Acad. Sci 202011249, 2011249117 (2020).
57. Schubert RD et al. Pan-viral serology implicates enteroviruses in acute flaccid myelitis. Nat. Med. 25, 1748–1752 (2019). [PubMed: 31636453]
58. Emery B & Dugas JC Purification of oligodendrocyte lineage cells from mouse cortices by immunopanning. Cold Spring Harb. Protoc. 2013, 854–868 (2013). [PubMed: 24003195]
59. Obradovic Z et al. Predicting intrinsic disorder from amino acid sequence. Proteins 53 Suppl 6, 566–572 (2003). [PubMed: 14579347]
60. Evans PR & Murshudov GN How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr 69, 1204–1214 (2013). [PubMed: 23793146]
61. Tickle I STARANISO: use of a WebGL-based 3D interactive graphical display to represent and visualise data quality metrics for anisotropic macromolecular diffraction data. Acta Crystallogr. A Found. Adv. 75, e162–e162 (2019).
62. McCoy AJ et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007). [PubMed: 19461840]
63. Fouts AE et al. Mechanism for neutralizing activity by the anti-CMV gH/gL monoclonal antibody MSL-109. Proc. Natl. Acad. Sci. U. S. A. 111, 8209–8214 (2014). [PubMed: 24843144]
64. Emsley P, Lohkamp B, Scott WG & Cowtan K Features and development of coot. Acta Crystallogr. D Biol. Crystallogr 66, 486–501 (2010). [PubMed: 20383002]
65. Liebschner D et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D Struct. Biol. 75, 861–877 (2019). [PubMed: 31588918]
66. Afonine PV et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr 68, 352–367 (2012). [PubMed: 22505256]
67. The PyMOL Molecular Graphics System, Version 21, Schrödinger, LLC.
68. Lanz TV et al. Tryptophan-2,3-Dioxygenase (TDO) deficiency is associated with subclinical neuroprotection in a mouse model of multiple sclerosis. Sci. Rep. 7, 41271 (2017). [PubMed: 28117398]

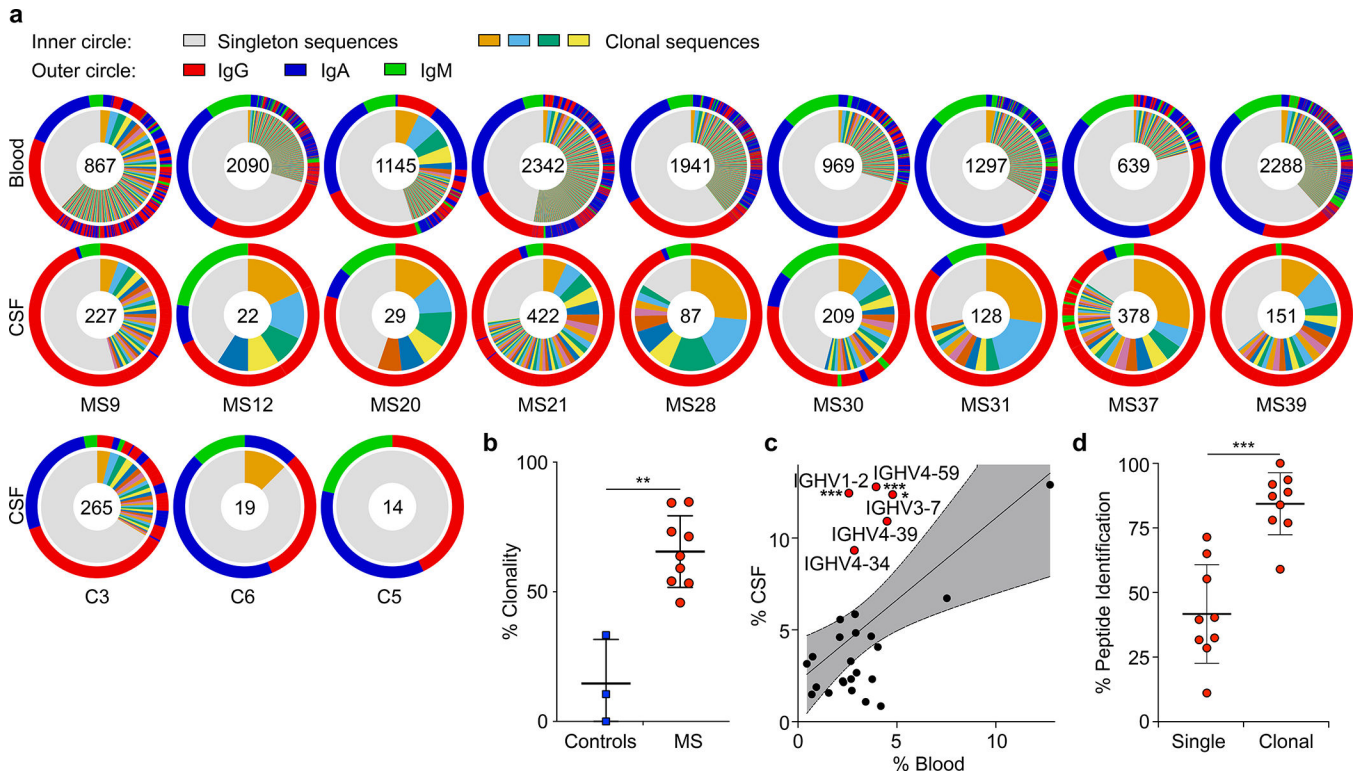


Figure 1: B cell repertoires in MS blood and CSF.

a-c, Single-cell BCR repertoire sequencing data, **a**, individual repertoires from paired blood plasmablasts (top row) and CSF B cells (second row) of MS patients and CSF B cells of control patients (bottom row) (*n=9* MS patients, *n=3* control patients), numbers indicate number of sequences, inner circle: colored wedges represent clonal expansions and grey area represents singleton antibody sequences, outer circle: immunoglobulin classes, red: IgG, blue: IgA, green: IgM, sequence locations in outer circle correspond to inner circle. **b**, Clonality analysis, percent of clonal sequences in CSF B cells are shown, comparing BCR repertoires of control patients (*n=3*) to MS patients (*n=9*). Means ± SD of each group are shown. $^{**}P = 0.0091$, two-tailed Mann-Whitney test. **c**, IGHV gene distribution in blood vs. CSF plasmablasts, each dot represents the usage of one IGHV gene across *n=9* MS patient repertoires in the respective compartments. Linear regression line and 95% confidence interval is shown, IGHV1–2, $^{***}P = 5.6\times10^{-4}$; IGHV4–59, $^{***}P = 9.2\times10^{-4}$; IGHV3–7, $^{*}P = 0.025$, unpaired two-tailed Student's t tests, Holm-Sidak corrected for multiple comparisons. **d**, Mass spectrometry data of purified CSF immunoglobulins, showing variable chain sequences that were uniquely identified with mass spectrometry in the singleton BCR sequences vs. clonal sequences (peptide-spectral matches (PSM) cutoff ≥1), means ± SD of *n=9* MS patients. $^{***}P = 0.0002$, two-tailed Mann-Whitney test.

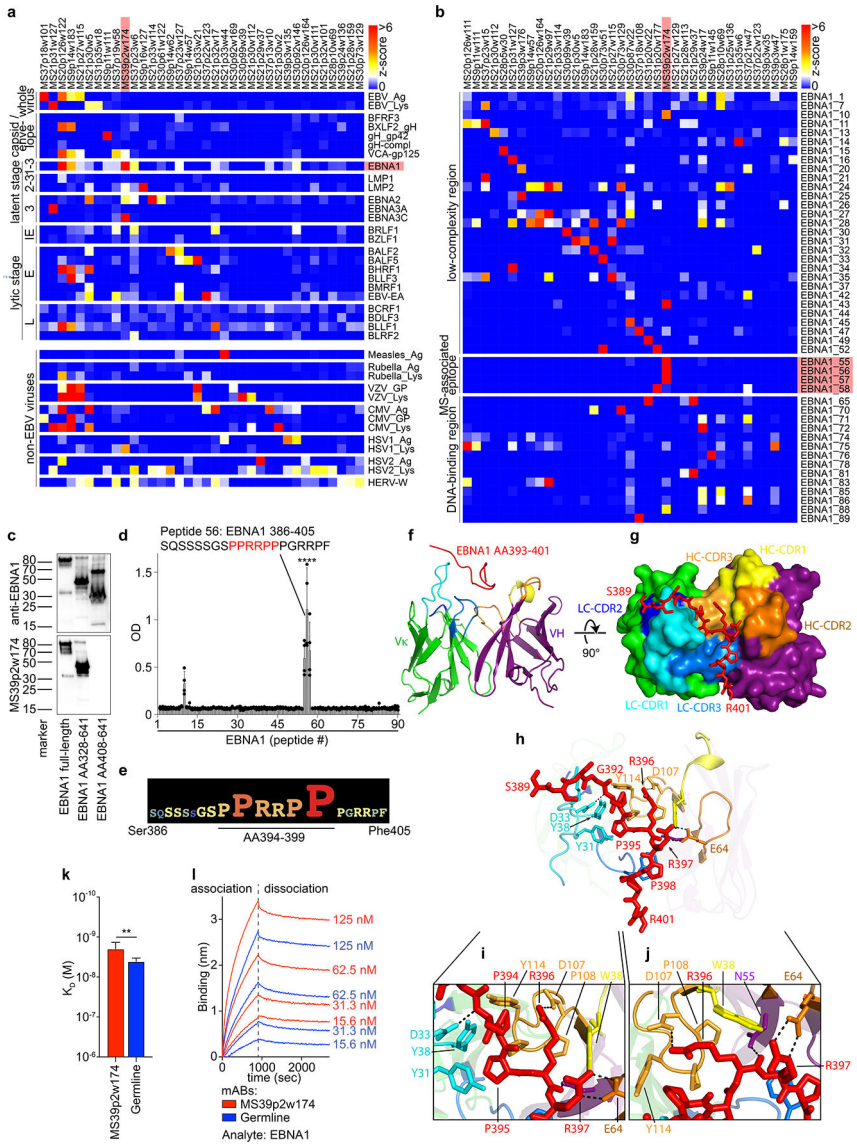


Figure 2: MS CSF B cell mAb reactivity to EBV proteins and interaction of MS39p2w174 with EBNA1$_{AA386–405}$·

a, Protein microarray data showing MS CSF mAb reactivities (z-scores) to viral lysates and EBV proteins, **b**, MS CSF mAB reactivities to EBNA1 peptides. Selected mAbs with highest reactivities are shown (n=36 (**a**) and n=37 (**b**) out of n=148). Data from one experiment carried out in 8 technical replicates. IE: immediate early, E: early, and L: late lytic stage, red: mAb MS39p2w174 and antigen EBNA1 / MS-associated region. **c**, Western blot analysis of recombinant EBNA1 (full-length and truncated proteins), stained on separate blots with commercial anti-C-terminal EBNAI antibody (top panel) and MS39p2w174 (bottom panel), molecular weight marker in kDa. **d**,**e**, ELISA data, **d**, overlapping peptide scan of MS39p2w174 binding EBNA1 peptides (20mers, 13AA overlap), means ± SD of *n=4* independent experiments, each measured in duplicates. **e**, Alanine-scan, EBNA1$_{AA386–405}$ logo representation showing the contribution of each residue to binding of MS39p2w174. **f-j**, Crystal structure of MS39p2w174 in complex

with EBNA1 AA386–405. **f**, Cartoon representation, showing $EBNA1_{AA393–401}$ in the binding groove. Peptide (red), HC (green/blue), LC (purple/yellow/orange), CDR loop colors correspond to annotations in (**g**). **g**, View of the binding groove from the top. Surface representation of the Fab with $EBNA1_{AA386–402}$ in stick representation. **h**, Cartoon and stick representation outlining close interactions. Major H-bond forming residues are represented as sticks. H-bonds < 3.2 Å are represented as black dashed lines. **i**, Magnification of peptide in the hydrophobic cage, **j**, magnification of region around Arg396 to emphasize polar contacts of HC residues with Arg396 and Arg397. **k**,**l**, Bio-layer interferometry measurement of MS39p2w174 (red) and germline (blue) affinity to EBNA1 full-length protein. **k**, K_D in M, mean K_D ± SD, calculated from four serial dilutions from one representative experiment out of three independent experiments, $^{**}P = 0.0043$, unpaired two-tailed Student's t test. **l**, association and dissociation curves corresponding to (**k**).

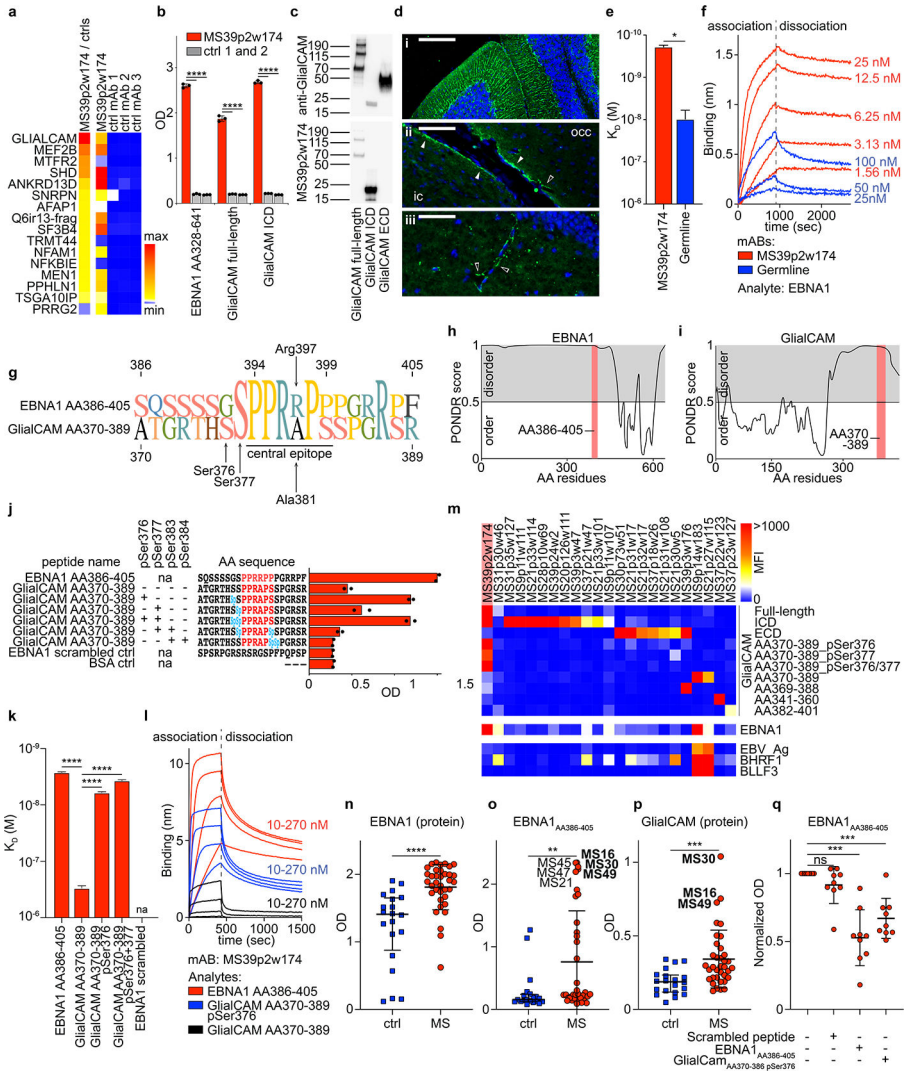


Figure 3: Molecular mimicry between EBNA1 and GlialCAM.

a, Protein microarray data showing top 16 results of HuProt array for MS39p2w174, compared to 3 control mAbs, sorted from top to bottom by the ratio of MS39p2w174 / average of controls reactivity (left column, min: 89, max: 911). Raw counts are shown in the four columns on the right (min: 1, max: 36450). **b**, ELISA showing binding of MS39p2w174 and two control mAbs to recombinant proteins $EBNA1_{AA328–641}$ as well as GlialCAM full-length (AA34–416) and ICD (AA262–416). Means ± SD of *n=3* independent experiments are shown, each carried out in triplicates. **c**, Western blot analysis of recombinant GlialCAM (full-length, ICD, and ECD), stained on separate blots with commercial anti-GlialCAM antibody (top panel) and MS39p2w174 (bottom panel), molecular weight marker in kDa. **d**, Fluorescent immunohistochemistry of mouse brain regions, stained with MS39p2w174, (i) cerebellum, scale bar: 150 µm, (ii) inferior colliculus (ic) and occipital cortex (occ) with perivascular glia (unfilled arrows) and glia limitans staining (filled arrows), scale bar: 60 µm, (iii) dentate gyrus with perivascular staining (unfilled arrows), scale bar: 40 µm. **e**,**f**, Bio-layer interferometry measurement of MS39p2w174 (red) and germline (blue) affinity to EBNA1 full-length protein. **k**, K_D

in M, mean K_D ± SD, calculated from three serial dilutions from one representative experiment out of three independent experiments, $*P = 0.0124$, unpaired two-tailed Student's t test, **f**, association and dissociation curves corresponding to (**e**). **g**, Alignment of AA sequences of $EBNA1_{AA386–405}$ and $GlialCAM_{AA370–389}$ and pointing out the central epitope region. **h**,**i**, Prediction of disorder with PONDR for **h**, EBNA1 and **i**, GlialCAM. High scores indicate disorder, red areas: epitope regions. **j**, ELISA data showing binding of MS39p2w174 to $EBNA1_{AA386–405}$ and $GlialCAM_{AA370–389}$ non-phosphorylated and phosphorylated at the indicated serine residues, means from *n=2* independent experiments are shown, representative for 5 experiments, each carried out in duplicates. pSer / Sp: phosphorylated serine. **k**,**l**, Bio-layer interferometry measurement of MS39p2w174 affinity to GlialCAM 20mer peptides. K_D in M, mean K_D ± SD, calculated from four serial dilutions from one representative experiment out of three independent experiments, $****P < 0.0001$, unpaired two-tailed ANOVA, Tukey corrected for multiple comparisons. pSer: phosphorylated serine residues. **l**, association and dissociation curves corresponding to (**k**). **m**, Protein microarray data showing mAb reactivities (MFI) to GlialCAM proteins, peptides, and phosphorylated peptides, as well as cross-reactivities to EBNA1 and other EBV proteins. Data from one experiment carried out in 8 technical replicates. ICD: intracellular domain, ECD: extracellular domain, pSer: phosphorylated serine residues. **n-p**, ELISA data showing human plasma reactivities against **n**, EBNA1 protein AA328–641, $****P < 0.0001$, unpaired two-tailed Mann-Whitney test, **o**, $EBNA1_{AA386–405}$, $**P < 0.0044$, unpaired two-tailed Mann-Whitney test, and **p**, GlialCAM protein, $***P < 0.0002$, unpaired two-tailed Mann-Whitney test, means ± SD from *n=20* control individuals and *n=36* MS patient samples. **q**, ELISA measurements of plasma antibody reactivity against $EBNA1_{AA386–405}$, without interference as well as blocked with scrambled peptide control, $EBNA1_{AA386–405}$, and $GlialCAM_{AA370–389\ pSer376}$. Mean OD (450 nm), normalized to unblocked sample, ± SD of *n=9* patient samples, each measured in quadruplicates. **e**,**k**,**n-p**, *P* values according to unpaired two-sided Student's t-test, **q**, one-way ANOVA, Dunnett corrected for multiple comparisons. $*P < 0.05$, $**P < 0.01$, $***P < 10^{-3}$, $****P < 10^{-4}$.

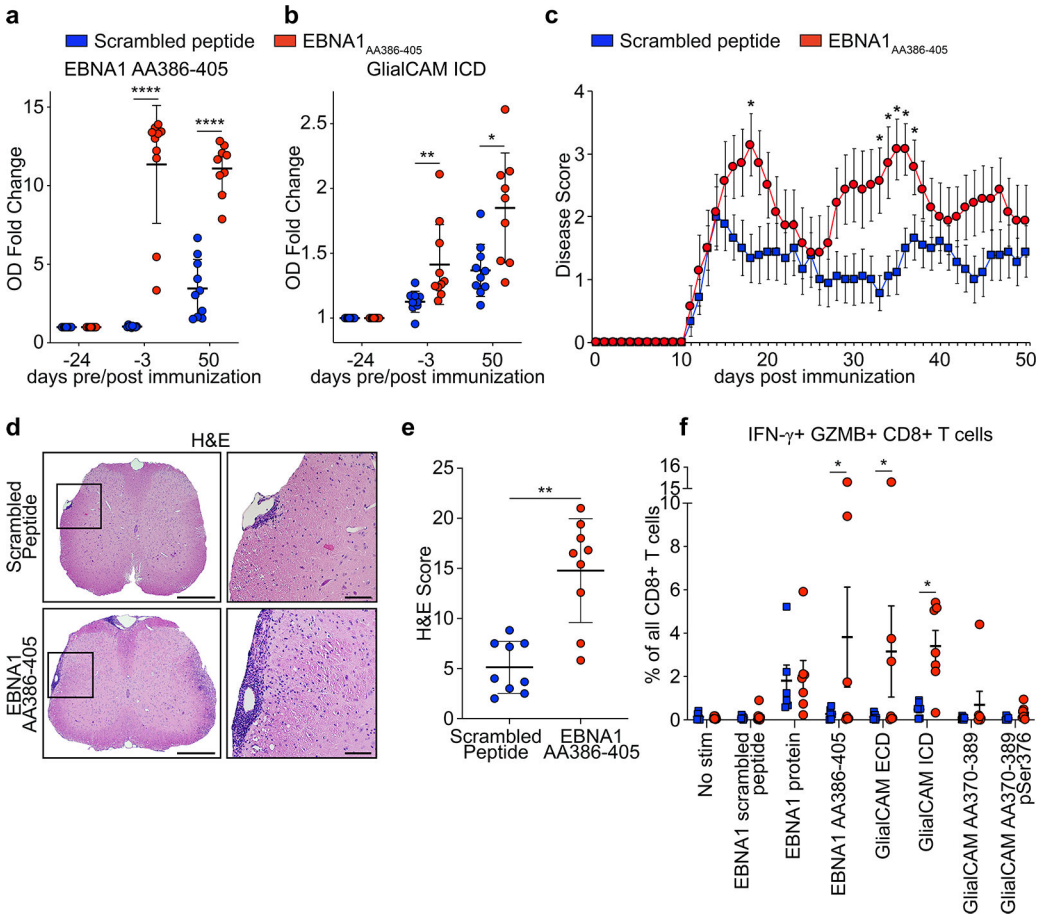


Figure 4: Anti-EBNA1$_{AA386–405}$ immunization exacerbates autoimmune-mediated demyelination *in vivo*.

a,**b**, ELISA data showing mouse plasma IgG response at the indicated timepoints pre and post EAE induction, for scrambled peptide immunized mice (blue, *n=10*) and EBNA1$_{AA386–405}$ immunized mice (red, *n=10*) against **a**, EBNA1$_{AA386–405}$, mean OD (450 nm) fold change ± SD, ****$P < 0.0001$, unpaired two-tailed Mann-Whitney test, Holm-Sidak corrected for multiple comparisons, and **b**, GlialCAM ICD, mean OD (450 nm) fold change ± SD, **$P = 0.0022$, *$P = 0.0152$, unpaired two-tailed Mann-Whitney test, Holm-Sidak corrected for multiple comparisons. Means ± SD, representative of three independent experiments, each carried out as triplicate measurements. **c**, EAE scores of mice immunized with scrambled peptide (blue, *n=9*) and EBNA1$_{AA386–405}$ (red, *n=7*), 3 weeks prior and on the day of EAE immunization (day 0), means of clinical scores ± SEM, *$P<0.05$ unpaired two-tailed Mann Whitney test. **d**,**e**, Spinal cord histology, **d**, representative H&E-stained spinal cords from scrambled peptide group (top panel) and EBNA1$_{AA386–405}$ group (bottom panel). Scale bars left images: 200 μm, right images: 50 μm**. e**, Statistical evaluation of H&E score, means of at least 4 coronal spinal cord sections per mouse and means ± SD for each group (*n=9*), **$P = 0.0012$, unpaired two-tailed Mann-Whitney test. **f**, Flow cytometry data of PBMC from healthy control individuals (*n=6*, blue) and MS patients (*n=7*, red), showing percent of IFN-γ+ GZMB+ CD8+ T cells in all CD8+ T cells. Means ± SEM are shown for the respective groups. Significance levels were assessed by two-way

ANOVA, followed by FDR calculation using the two-stage step-up method of Benjamini, Krieger and Yekutieli, * FDR < 0.1.