

## Supplementary information

---

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

---

In the format provided by the authors and unedited

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

**Authors:** Tobias V. Lanz<sup>1,2,3,4</sup>, R. Camille Brewer<sup>1,4</sup>, Peggy P. Ho<sup>5</sup>, Jae-Seung Moon<sup>1,4</sup>, Kevin M. Jude<sup>6</sup>, Daniel Fernandez<sup>7</sup>, Ricardo A. Fernandes<sup>6</sup>, Alejandro M. Gomez<sup>1,4</sup>, Gabriel-Stefan Nadj<sup>1,4</sup>, Christopher M. Bartley<sup>8</sup>, Ryan D. Schubert<sup>9</sup>, Isobel A. Hawes<sup>9</sup>, Sara E. Vazquez<sup>10</sup>, Manasi Iyer<sup>11</sup>, J. Bradley Zuchero<sup>11</sup>, Bianca Teegen<sup>12</sup>, Jeffrey E. Dunn<sup>13</sup>, Christopher B. Lock<sup>13</sup>, Lucas B. Kipp<sup>13</sup>, Victoria C. Cotham<sup>14,15</sup>, Beatrix M. Ueberheide<sup>14,15</sup>, Blake T. Aftab<sup>16</sup>, Mark S. Anderson<sup>17</sup>, Joseph L. DeRisi<sup>10,18</sup>, Michael R. Wilson<sup>9</sup>, Rachael J. M. Bashford-Rogers<sup>19</sup>, Michael Platten<sup>2,3,20</sup>, K. Christopher Garcia<sup>6</sup>, Lawrence Steinman<sup>5</sup> & William H. Robinson<sup>1,4,\*</sup>

## Affiliations:

<sup>1</sup>Division of Immunology and Rheumatology, Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA.

<sup>2</sup>Department of Neurology, Mannheim Center for Translational Neurosciences (MCTN), Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany.

<sup>3</sup>Department of Neurology and National Center for Tumor Diseases, University Hospital Heidelberg, Heidelberg, Germany.

<sup>4</sup>The Geriatric Research, Education, and Clinical Center (GRECC), VA Palo Alto Health Care System, Palo Alto, CA, USA.

<sup>5</sup>Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Beckman Center for Molecular Medicine, Stanford, CA, USA.

<sup>6</sup>Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Beckman Center for Molecular Medicine, Stanford, CA, USA.

<sup>7</sup>Macromolecular Structure Knowledge Center, Stanford ChEM-H Institute, Stanford, CA, USA.

<sup>8</sup>Weill Institute for Neurosciences, Department of Psychiatry and Behavioral Sciences, University of California San Francisco, San Francisco, CA, USA.

<sup>9</sup>Weill Institute for Neurosciences, Department of Neurology, University of California San Francisco, San Francisco, CA, USA.

<sup>10</sup>Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, CA, USA.

<sup>11</sup>Department of Neurosurgery, Stanford University School of Medicine, Stanford, CA, USA.

<sup>12</sup>Institute of Experimental Immunology, Euroimmun AG, Lubeck, Germany.

<sup>13</sup>Division of Neuroimmunology, Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA, USA.

<sup>14</sup>Department of Biochemistry and Molecular Pharmacology, NYU Perlmutter Cancer Center, NYU School of Medicine, New York, NY, USA.

<sup>15</sup>NYU Langone Health Proteomics Laboratory, Division of Advanced Research Technologies, NYU School of Medicine, New York, NY, USA.

<sup>16</sup>Preclinical Science and Translational Medicine, Atara Biotherapeutics, South San Francisco, CA, USA.

<sup>17</sup>Department of Medicine, Diabetes Center, University of California San Francisco, San Francisco, CA, USA.

<sup>18</sup>Chan Zuckerberg Biohub, University of California San Francisco, San Francisco, CA, USA.

<sup>19</sup>Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK.

<sup>20</sup>DKTK Clinical Cooperation Unit Neuroimmunology and Brain Tumor Immunology, German Cancer Research Center (DKFZ), Heidelberg, Germany.

\*e-mail: wrobins@stanford.edu

## **Supplementary Information**

<b>Table of Content</b>	<b>Page</b>
Supplementary Table 1	4
Supplementary Table 2	5
Supplementary Table 3	6
Supplementary Table 4	7
Supplementary Table 5	8
Supplementary Table 6	10
Supplementary Figure 1	11
Supplementary Figure 2	12
Supplementary Discussion	13
References	14

**Supplementary Table 1**

	HLA-A 1	HLA-A 2	HLA-B 1	HLA-B 2	HLA-C 1	HLA-C 2
<b>MS9</b>	02:05	68:01	15:03	50:01	02:10	06:02
<b>MS12</b>	31:01	32:01	40:01	40:01	03:04	03:04
<b>MS20</b>	02:01	02:01	18:01	51:01	07:01	14:02
<b>MS21</b>	01:01	03:01	07:02	58:01	03:02	07:02
<b>MS28</b>	01:01	03:01	07:02	07:02	07:02	07:02
<b>MS30</b>	01:01	24:02	07:02	08:01	07:01	07:02
<b>MS31</b>	01:01	02:01	08:01	13:02	06:02	07:02
<b>MS37</b>	03:01	68:02	15:10	35:01	03:04	04:01
<b>MS39</b>	01:01	03:01	08:01	51:01	07:01	15:02

	HLA-DRB1_1	HLA-DRB1_2	HLA-DRB345_1	HLA-DRB345_2	HLA-DQA1_1	HLA-DQA1_2	HLA-DQB1_1	HLA-DQB1_2	HLA-DPA1_1	HLA-DPA1_2	HLA-DPB1_1	HLA-DPB1_2
<b>MS9</b>	07:01	15:01	4*01:01	5*01:01	01:02	02:01	02:02	06:02	01:03	01:03	03:01	04:01
<b>MS12</b>	01:01	04:04	4*01:03	null	01:01	03:01	03:02	05:01	01:03	02:01	01:01	02:01
<b>MS20</b>	08:01	11:01	3*02:02	null	04:02	05:05	04:02	03:01	01:03	01:03	04:01	04:02
<b>MS21</b>	15:01	13:02	3*01:01	5*01:01	01:02	01:02	06:02	06:09	01:03	01:03	04:01	104:01
<b>MS28</b>	07:01	15:01	4*01:01	5*01:01	01:02	02:01	02:02	06:02	02:01	02:01	01:01	11:01
<b>MS30</b>	03:01	15:01	3*01:01	5*01:01	01:02	05:01	02:01	06:02	01:03	02:01	01:01	15:01
<b>MS31</b>	03:01	15:01	3*01:01	5*01:01	01:02	05:01	02:01	06:02	01:03	01:03	03:01	03:01
<b>MS37</b>	01:01	08:04	null	null	01:01	05:01	03:01	05:01	01:03	01:03	02:01	04:02
<b>MS39</b>	03:01	07:01	3*01:01	4*01:03	02:01	05:01	02:01	03:03	01:03	01:03	03:01	06:01

**Supplementary Table 1.** Genotype of full HLA locus. HLA-DRB1\*15:01 (dark red) is the main risk allele for MS, HLA-DRB1\*03:01 (bright red) is a secondary risk allele for MS.





## Supplementary Table 4

Protein Antigen	Name in Floures	Name in Suopj. Figure	Product Name	Company	catalogue #	Lot #	residues	source	Usulj
GlialCAM ECD	GlialCAM ECD	GlialCAM ECD A	HEPACAM Protein, Human, Recombinant (Ecd, His Tag)	SINO Biological	R6047-H08H	ECH011-2806	AA34-240	HEK293	0.10
GlialCAM ECD	GlialCAM ECD	GlialCAM ECD B	Recombinant Human HEPACAM protein	Abcam	ab151670	GR3329367-1	AA34-240	HEK293	0.10
GlialCAM ECD	GlialCAM ECD	GlialCAM ECD C	Recombinant Human HEPACAM	Sino Biological	R6047-H08H	LC10F0E2808	AA1-240	HEK293	0.10
GlialCAM ECD	GlialCAM ECD	GlialCAM ECD D	Recombinant Human HEPACAM protein	Abcam	ab151670	GR3241129-1	AA34-240	HEK293	0.10
GlialCAM ICD	GlialCAM ICD	GlialCAM ICD A	Hepracam intracellular region, self expression	self		batch 1	AA262-416	E. Coli	0.10
GlialCAM ICD	GlialCAM ICD	GlialCAM ICD C	Hepracam intracellular region, self expression	self		batch 2	AA262-416	E. Coli	0.10
Hepracam full-length	Hepracam full-length	GlialCAM Full Length	HEPACAM (NM_152722) Human Recombinant Protein	OrfGene	TP323693	WX02C25F	AA1-416	HEK293T	0.10
Peptide Antigen	Name in Suopj. Figure	Name in Suopj. Figure	AA Sequence	Company	Peptide Length	Usulj			
GlialCAM p315-334		GlialCAM p315-334	IKNSPTEENRPAEPSPGSAE	Sigma Aldrich	20	0.10			
GlialCAM p320-339		GlialCAM p320-339	TEENRPAEPSPGSAE	Sigma Aldrich	20	0.10			
GlialCAM p321-340		GlialCAM p321-340	TEENRPAEPSPGSAE	Sigma Aldrich	20	0.10			
GlialCAM p327-346		GlialCAM p327-346 A	PEPRSAEPSPGPGYVSPAV	Sigma Aldrich	20	0.10			
GlialCAM p327-346		GlialCAM p327-346 B	PEPRSAEPSPGPGYVSPAV	Sigma Aldrich	20	0.10			
GlialCAM p333-352		GlialCAM p333-352	TEPGPGYVSPAVPGRSPG	Sigma Aldrich	20	0.10			
GlialCAM p334-353		GlialCAM p334-353 A	EPGPGYVSPAVPGRSPGL	Sigma Aldrich	20	0.10			
GlialCAM p334-353		GlialCAM p334-353 B	EPGPGYVSPAVPGRSPGL	Sigma Aldrich	20	0.10			
GlialCAM p340-359		GlialCAM p340-359	VYSPVAVPGRSPGLPGRAR	Sigma Aldrich	20	0.10			
GlialCAM p341-360	GlialCAM p341-360	GlialCAM p341-360	VYSPVAVPGRSPGLPGRAR	Sigma Aldrich	20	0.10			
GlialCAM p346-365		GlialCAM p346-365	VPRSPGLTIRARVYRSP	Sigma Aldrich	20	0.10			
GlialCAM p348-367		GlialCAM p348-367	GRSPGLTIRARVYRSP	Sigma Aldrich	20	0.10			
GlialCAM p352-371		GlialCAM p352-371	GLPISARVYRSP	Sigma Aldrich	20	0.10			
GlialCAM p355-374		GlialCAM p355-374	IRSAARVYRSPSPATGRT	Sigma Aldrich	20	0.10			
GlialCAM p358-377		GlialCAM p358-377	ARVYRSPSPATGRT	Sigma Aldrich	20	0.10			
GlialCAM p362-381		GlialCAM p362-381	SPSPSPATGRT	Sigma Aldrich	20	0.10			
GlialCAM p364-383		GlialCAM p364-383	SPSPSPATGRT	Sigma Aldrich	20	0.10			
GlialCAM p369-388	GlialCAM p369-388	GlialCAM p369-388 A	PATGRTSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p369-388	GlialCAM p369-388	GlialCAM p369-388 B	ATGRTSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p370-389		GlialCAM p370-389 A	ATGRTSPSPSPSPGSR	GenScript	20	0.10			
GlialCAM p370-389		GlialCAM p370-389 B	ATGRTSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p370-389		GlialCAM p370-389 C	ATGRTSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p376-395		GlialCAM p376-395 A	SSPPRSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p376-395		GlialCAM p376-395 B	SSPPRSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p378-389		GlialCAM p378-389	PPRSPSPGSR	GenScript	12	0.10			
GlialCAM p382-401	GlialCAM p382-401	GlialCAM p382-401	PSPPRSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p383-402		GlialCAM p383-402	SSPPRSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p369-388 Ctr380		GlialCAM p369-388 Ctr380	ATGRTSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p369-388 Ctr387		GlialCAM p369-388 Ctr387	ATGRTSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p369-388 Ctr373		GlialCAM p369-388 Ctr373	ATGRTSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p369-388 Ctr373/380		GlialCAM p369-388 Ctr373/380	ATGRTSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p369-388 Ctr373/380/387		GlialCAM p369-388 Ctr373/380/387	ATGRTSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p370-389 pSer383/384		GlialCAM p370-389 pSer383/384	ATGRTSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p370-389 pSer376	GlialCAM p370-389 pSer376	GlialCAM p370-389 pSer376	ATGRTSPSPSPSPGSR	GenScript	20	0.10			
GlialCAM p370-389 pSer376/377	GlialCAM p370-389 pSer376/377	GlialCAM p370-389 pSer376/377 A	ATGRTSPSPSPSPGSR	Sigma Aldrich	20	0.10			
GlialCAM p370-389 pSer376/377	GlialCAM p370-389 pSer376/377	GlialCAM p370-389 pSer376/377 B	ATGRTSPSPSPSPGSR	GenScript	20	0.10			
GlialCAM p370-389 pSer377	GlialCAM p370-389 pSer377	GlialCAM p370-389 pSer377	ATGRTSPSPSPSPGSR	GenScript	20	0.10			
GlialCAM p370-389 pSer377/383	GlialCAM p370-389 pSer377/383	GlialCAM p370-389 pSer377/383	ATGRTSPSPSPSPGSR	GenScript	20	0.10			

Supplementary Table 4. GlialCAM proteins and peptides.



## Supplementary Table 5

### Crystallographic data collection and refinement statistics

EBNA1 peptide 386-405/MS39p2w174 Fab	
<b>Data collection</b>	
Beamline	SSRL BL12-2
Wavelength (Å)	0.97946
Space group	I222
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	119.66, 137.56, 179.00
$\alpha$ , $\beta$ , $\gamma$ (°)	90.00, 90.00, 90.00
Matthews coefficient (Å <sup>3</sup> /Da) <sup>a</sup>	3.7
Solvent content (%)	66.8
Anisotropy	0.72
Resolution (Å) <sup>b</sup>	45.14(2.19)
<i>R</i> <sub>merge</sub> <sup>c</sup>	0.16(5.92)
<i>I</i> / $\sigma$ <i>I</i> ratio <sup>d</sup>	5.5(0.7)
Uncorrected completeness (%) <sup>e,f</sup>	96.1(78.0)
Spherical completeness (%) <sup>g</sup>	54.7(13.1)
Ellipsoidal completeness (%) <sup>h</sup>	92.5(72.4)
Reflections (total/unique)	51,233(3,637)
Redundancy <sup>j</sup>	3.2(2.0)
<b>Refinement</b>	
Resolution (Å)	45.14-2.50
No. reflections/test set	36,574/1,912
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub> <sup>k</sup>	21.1/25.2
Mean B value (Å <sup>2</sup> )	56.4
<i>F</i> <sub>obs</sub> - <i>F</i> <sub>calc</sub> correlation <sup>l</sup>	0.90
No. atoms	
Protein	6,778
Ligand/ion	4 glycerol/1 chlorine
Water	24
<i>B</i> -factors	
Protein	46 (overall)
Ligand/ion	73 (glycerol)/45 (chlorine)
Water	52
Ramachandran statistics <sup>m</sup>	
Most favored/allowed regions (%)	100
Disallowed regions (%)	0
R.m.s. deviations	
Bond lengths (Å)	0.33
Bond angles (°)	0.55

<sup>a</sup>Ratio of the volume of the asymmetric unit to the molecular weight of all protein in the asymmetric unit

<sup>b</sup>Value in parentheses is for the highest-resolution shell: 2.24 – 2.19 Å.

<sup>c</sup>Reliability factor for symmetry-related reflections calculated as:  $R_{\text{merge}} = \frac{\sum_{hkl} \sum_{j=1}^N |I_{hkl} - I_{hkl}(j)|}{\sum_{hkl} \sum_{j=1}^N I_{hkl}(j)}$ , where *N* is the redundancy of the data. In parentheses, the cumulative value at the highest-resolution shell

<sup>d</sup>Ratio of mean intensity to the mean standard deviation of the intensity over the entire resolution range

<sup>e</sup>Fraction of measured reflections to possible observations at the resolution range

<sup>f</sup>Completeness of the uncorrected data set

<sup>g</sup>Completeness within the sphere that contains the ellipsoid, after removing reflections outside of the ellipsoid

<sup>h</sup>The amount of expected reflections within the ellipsoid that are actually measured

<sup>j</sup>Number of measurements of individual, symmetry unique reflections

<sup>k</sup>Average deviation between the observed and calculated structure factors calculated as:  $R_{\text{work}} = \sum_{\text{hkl}} \frac{||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum_{\text{hkl}} |F_{\text{obs}}|}$ , where the  $F_{\text{obs}}$  and  $F_{\text{calc}}$  are the observed and calculated structure factor amplitudes of reflection hkl.  $R_{\text{free}}$  is equal to  $R_{\text{factor}}$  but for a randomly selected 5.0 % subset of the total reflections that were held aside throughout refinement for cross-validation

<sup>l</sup>Correlation coefficient between observed and calculated structure factor amplitudes

<sup>m</sup>According to Molprobit for non-proline and non-glycine residues

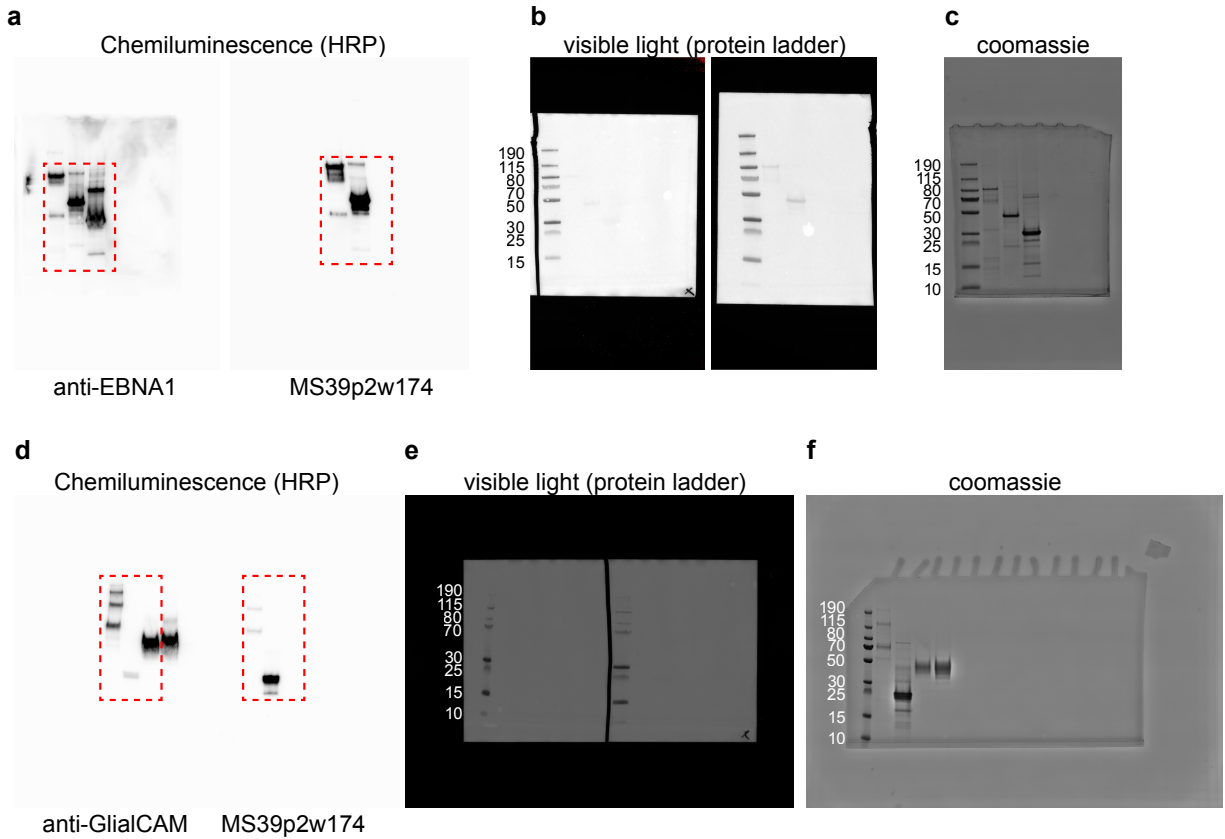
**Supplementary Table 5.** Crystallographic data and refinement statistics.

## Supplementary Table 6

antibody	ligand	KD (M)	KD Error	kon (M <sup>-1</sup> s <sup>-1</sup> )	kon Error	koff (s <sup>-1</sup> )	koff Error	n
MS39p2w174	EBNA1 protein	1.99E-09	6.25E-10	4.07E+04	1.61E+04	7.32E-05	1.06E-05	4
MS39p2w174 germline	EBNA1 protein	4.19E-09	7.62E-10	2.94E+04	1.65E+04	1.14E-04	4.53E-05	4
MS39p2w174	Glia1CAM protein	1.91E-10	1.73E-11	2.61E+05	5.01E+04	5.00E-05	1.21E-05	3
MS39p2w174 germline	Glia1CAM protein	1.05E-08	4.12E-09	4.28E+04	1.39E+04	4.12E-04	4.79E-05	3
MS39p2w174	EBNA1 AA386-405	2.67E-09	7.83E-11	2.40E+05	5.28E+03	6.40E-04	1.24E-05	4
MS39p2w174	Glia1CAM AA370-389	3.02E-07	3.10E-08	7.15E+04	6.39E+03	2.16E-02	1.09E-03	4
MS39p2w174	Glia1CAM AA370-389 pSer376	6.10E-09	2.69E-10	2.80E+05	1.10E+04	1.71E-03	3.40E-05	4
MS39p2w174	Glia1CAM AA370-389 pSer376/377	3.73E-09	1.50E-10	3.23E+05	1.14E+04	1.20E-03	2.33E-05	4

**Supplementary Table 6.**  $K_D$ ,  $K_{on}$ , and  $K_{off}$  values for antibody-ligand pairs. Values correspond to **Fig. 2 k,l** (MS39p2w174 and Germline with EBNA1 protein), **Fig. 3 e,f** (MS39p2w174 and Germline with Glia1CAM protein), and **Fig. 3 k,l** (MS39p2w174 with EBNA1 AA386-405, Glia1CAM AA370-389, and phosphorylated Glia1CAM AA370-389). Measurements are representative of at least 3 independent experiments. Values are averages of n replicate measurements in serial dilutions.

## Supplementary Figure 1



**Supplementary Figure 1.** Raw western blot images. **a-c**, recombinant EBNA1 proteins (full-length, AA328-641, and AA408-641), corresponding to (**Fig. 2c**). **d-f**, recombinant GlialCAM proteins (full-length, ICD, and ECD), corresponding to (**Fig. 3c**). Recombinant proteins were run on three separate blots, stained with commercial anti-EBNA1 or anti-GlialCAM (**a,d**, left), MS39p2w174 (**a,d**, right), and coomassie (**c,f**). Red frames: cropped images used for figures.



## Supplementary Discussion

### Patient collective and MS diagnosis

As mentioned in the main manuscript, all MS patients included in this study cohort were selected for CSF pleocytoses above 10 cells /  $\mu\text{l}$ . This selection was made due to technical considerations, as single-cell B cell sorting and sequencing became increasingly challenging below a threshold of 5 cells /  $\mu\text{l}$  and yielded significantly more BCR sequences above 10 cells /  $\mu\text{l}$ . CSF volumes ranged from 5-15 mL per patient. When selecting patients with high pleocytosis, we were aware of the possible misdiagnosis of neuromyelitis optica spectrum disorders (NMOSD). Special care was taken to follow the current McDonald criteria for the diagnosis of MS<sup>1,2</sup>. None of the patients met the diagnostic criteria for NMOSD, in particular spinal lesions spanning  $\geq 3$  segments<sup>3</sup>. Patients were tested for antibodies against aquaporin-4 and MOG and showed negative results. Patient MS37, who showed the highest pleocytosis of 57 cells /  $\mu\text{l}$ , had a spinal lesion spanning one segment in addition to several small subventricular contrast enhancing lesions. In order to avoid a misdiagnosis, her serum (and CSF) were tested repeatedly and in different laboratories for anti-aquaporin-4 and anti-MOG antibodies and always tested negative. Considering all symptoms and laboratory results, the patient did not meet criteria of NMOSD and the MS diagnosis remained. Of note, patient MS39, from whom antibody MS39p2w174 was selected, possessed 18 cells /  $\mu\text{l}$ , which was in the middle of the range of CSF cell counts (Extended Data Table 1).

### HLA phenotype and HLA binding prediction

While our manuscript focuses on B cell biology and antibody-mediated molecular mimicry, we briefly discuss our data on human and mouse T cell reactivity against EBNA1 and GlialCAM (Extended Data Fig. 10). Patients' HLA genotypes could profoundly influence T cell reactivity to either antigen. We should point out that patient MS39, from whom antibody MS39p2w174 originated, is not a carrier of the most significant MS risk allele HLA-DRB1\*15:01 (Supplementary Table 1). The patient carries HLA-DRB1\*03:01, a second HLA class II risk allele for MS with lower significance. The patient did not carry any of the protective HLA class I alleles (HLA-A\*02:01, HLA-B\*44:02, HLA-B\*38:01 and HLA-B\*55:01)<sup>4,5</sup>. We used the prediction tools SYFPEITHI, NetMHC, and IEDB to predict binding of EBNA1 and GlialCAM peptides to the patient's molecular HLA setup, and included the main MS risk allele HLA-DRB1\*15:01 (Supplementary Fig. 2). We listed the 10 highest binding peptides of EBNA1 and GlialCAM for each HLA allele as well as the highest prediction for a peptide containing the central binding motifs of antibody MS39p2w174 (EBNA1<sub>AA394-400</sub> and GlialCAM<sub>AA377-383</sub>). These motifs were not predicted to be presented particularly well on either HLA class I or HLA class II and were never amongst the top 10 predicted binders. However, multiple other regions of both proteins are predicted to be presented well on several alleles, with GlialCAM peptides binding slightly better to HLA class I alleles (Supplementary Fig. 2a) and EBNA1 peptides binding slightly better to HLA class II alleles (Supplementary Fig. 2b). Peptides of both proteins can be presented well on HLA-DRB1\*15:01, but only EBNA1 binds tightly to HLA-DRB1\*03:01. Presentation of EBNA1 on HLA-DRB1\*15:01 and HLA-DRB1\*03:01 might be part of the explanation how these alleles contribute to MS risk.

## References

1. Thompson, A. J. *et al.* Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* **17**, 162–173 (2018).
2. Polman, C. H. *et al.* Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann. Neurol.* **69**, 292–302 (2011).
3. Wingerchuk, D. M. *et al.* International consensus diagnostic criteria for neuromyelitis optica spectrum disorders. *Neurology* **85**, 177–189 (2015).
4. Moutsianas, L. *et al.* Class II HLA interactions modulate genetic risk for multiple sclerosis. *Nat. Genet.* **47**, 1107–1113 (2015).
5. de la Concha, E. G. *et al.* DRB1\*03:01 haplotypes: differential contribution to multiple sclerosis risk and specific association with the presence of intrathecal IgM bands. *PLoS One* **7**, e31018 (2012).