Research paper

Identification of two highly antigenic epitope markers predicting multiple sclerosis in optic neuritis patients

Helle Sadam[a,b], Arno Pihlak[a], Mariliis Jaago[a,b], Nadežda Pupina[a], Annika Rähni[a,b], Maarja Toots[a], Antti Vaheri[c], Janne K. Nieminen[d], Mika Siuko[e], Pentti J. Tienari[d], Kaia Palm[a,b,*]

[a] Protobios Llc, Mäealuse 4, Tallinn 12618, Estonia
[b] Department of Chemistry and Biotechnology, Tallinn University of Technology, Akadeemia tee 15, Tallinn 12618, Estonia
[c] Department of Virology, University of Helsinki, Medicum, Haartmaninkatu 3, 00014, Finland
[d] Neurology, Neurocentre, Helsinki University Hospital and Translational Immunology Research Program, Faculty of Medicine, University of Helsinki, Haartmaninkatu 8, Helsinki 00014 Finland
[e] Department of Ophthalmology, Helsinki University Hospital and University of Helsinki, Helsinki, Finland

A R T I C L E   I N F O

A B S T R A C T

Background: Optic neuritis (ON) can occur as an isolated episode or will develop to multiple sclerosis (MS) a chronic autoimmune disease. What predicts ON progression to MS remains poorly understood.
Methods: We characterised the antibody epitope repertoire in three independent clinical cohorts (discovery (*n* = 62), validation (*n* = 20) and external cohort (*n* = 421)) using mimotope variation analysis (MVA), a next generation phage display technology to identify epitopes that associate with prognosis of ON.
Findings: We observed distinct epitope profiles for ON, MS and the controls, whereas epitope repertoires of sera and CSF were highly similar. Two unique and highly immunogenic epitopes A and B were detected in subjects with ON progressing to MS. These epitopes A and B were strongly associated with herpesviral antigens (VCA p18 of Epstein-Barr virus (EBV); gB of Cytomegalovirus (CMV)). ROC addressed 75% of MS subjects with ON onset correctly (at 75% sensitivity and 74.22% specificity) based on the two-epitope biomarker analysis.
Interpretation: This is the first report on epitope diagnostics for MS employing the unbiased strategy of MVA for identification of novel immunological features of disease.
Funding: The Estonian Ministry of Education, The Estonian Research Council (PRG573, PRG805 and PSG691), H2020-MSCA-RISE-2016 (SZTEST), H2020-NMBP-2017 (PANBIORA), Helsinki University Hospital, Mary and Georg C. Ehrnrooth, Finnish Eye, Sigrid Jusélius and Magnus Ehrnrooth Foundations.

1. Introduction

Multiple sclerosis (MS) is the most prevalent chronic inflammatory disease of the central nervous system (CNS). Most patients present with relapsing MS, with episodes of relapse and remission phases, whereas ~10% have a progressive course from the onset.

The causes of MS are not known. Associations with HLA-DRB1×15:01 (HLA class II) and HLA-A*02:01 (class I) along with more than 200 genetic variants have been reported to influence the risk of MS [1]. Female gender, Scandinavian/Celtic descent, low childhood vitamin D status, adolescent obesity, smoking, and infectious background − all include major risk factors of MS (reviewed in [2]). MS is considered an autoimmune disease, but the antigens of CNS targeted by T and B cells are currently largely unknown (reviewed in [3]). About 90% of patients with MS have IgG oligoclonal bands (OCBs) in the cerebrospinal fluid (CSF) [4]. OCBs are also common in other types of CNS inflammation [4,5]. Reactivity of OCBs and intrathecal antibodies against bacterial (e.g., *Chlamydia pneumoniae* [6-8]), human herpesvirus (Epstein-Barr virus (EBV), Human Herpesvirus 6 (HHV-6)) and other viral (measles, rubella, and zoster ("MRZ") [9-14]) antigens have been reported in patients with MS, although underlying mechanisms remain unclear.

The optic nerve is one of the major targets in MS. About 20% of MS patients present with optic neuritis (ON) as one of the first symptoms, whereas during the course of the disease ON may occup in 50% of patients. In a follow-up of a large cohort of ON patients the risk of developing MS after ON was 30% at 5 years, close to 40% at 10 years, and 50% after 15 years [15].

* Corresponding author.
E-mail address: kaia@protobios.com (K. Palm).

Research in Context

Evidence before the study

Multiple sclerosis (MS) is the most common neurological disease in young adults, affecting more than 2 million individuals worldwide. About 20% of MS patients experience optic neuritis (ON) as the presenting symptom, but not all ON patients develop MS. Thus, it is important to identify prognostic biomarkers of the development of MS after ON. B cells are now recognised to play a central role in MS. The MS community has traditionally focused on well-defined candidate autoantigens, whereas novel techniques can expand the number of target antigens in an unbiased manner.

Added value of this study

Here, we present a general methodology for antibody epitope repertoire analysis to rapidly discover the immunodominant epitopes, and to develop high performance epitope-specific diagnostic tests. Using next generation phage display technology (MVA) we provide a broad, high-resolution view on humoral immunity associated with MS and report on the prognostic value of viral antibody epitopes as novel biomarkers for the risk of MS after the first episode of ON. These epitopes with cross-reactivity to antigens of common viral pathogens (EBV, CMV) potentially reflect pathogenic mechanisms in the development of MS. In addition, we show a high clinical potential of these epitopes as blood biomarkers in addressing MS correctly for 76.6% of cases.

Implication of all the available evidence

Because of the significant clinical heterogeneity of MS, biomarkers that reflect pathogenic mechanisms may thus be valuable for predicting disease progression and facilitating individualised treatments. The identification of circulating biomarkers' immunological significance presents an attractive solution for early diagnosis and prognosis of MS.

Although some immunological factors have previously been associated with MS, virtually nothing is known whether different sites of onset have different predictive markers for MS. Here we employed a hypothesis-free approach of mimotope variation analysis (MVA), a next generation phage display method to analyse antibody epitope profiles of subjects with ON who did or did not progress to MS. Furthermore, we identified novel epitope biomarkers for assessing ON progression to MS from blood.

2. Methods

2.1. Study population

The discovery cohort included CSF and sera/plasma samples of 24 treatment-naïve Finnish patients who initially received the diagnosis of ON and 38 controls (**Table 1**). Fifteen of the ON patients were subsequently diagnosed with relapsing MS (ON^{MS}) during the median follow-up time of 52 months (range 38–69 months), whereas 9 patients did not develop MS (denoted ON^{ON}). Using standard isoelectric focusing and agarose gel electrophoresis OCBs were found in all CSF samples of the discovery cohort. In the ON^{ON} group the average number of CSF-specific OCBs was 12; in the ON^{MS} group the average number of CSF-specific OCBs was 15 (the numbers of CSF-specific OCBs observed are in **Figure S1a**).

The validation cohort included, sera/plasma and CSF samples collected at diagnostic phase (treatment-naïve) of 20 Finnish patients with relapsing MS, out of which 10 presented with ON (MS^{ON}) and 10 with other symptoms (MS^{other}) (**Table 1**). All patients with MS in the discovery and validation cohorts fulfilled McDonald 2005 and 2017 criteria, their clinical features on baseline EDSS, visual functional score and brain MRI findings are summarised in **Table S1** and OCBs in the CSF on **Figure S1a**.

The diagnostic model using ROC analysis was tested in an independent external cohort of Estonians including healthy individuals (n = 229, **Table 2**) and in subjects with different ICD-10 diagnosis codes, but without any notification of demyelinating disease (n = 192, **Table 2**).

2.2. Ethics statement

Patients for this study were recruited at the Department of Neurology and Department of Ophthalmology of Helsinki University Hospital, Finland. Written informed consent was obtained from all study participants. This study was pre-approved by the regional ethics committee at the Helsinki University Hospital (Dno 83/13/03/01/ 2013), Ethics Review Committee on Human Research of the University of Tartu, Estonia: 212T-24 (issued 13.03.12), 177/T-2 (issued 15.12.08), 211/M-22 (issued 23.01.12) and 281/T-5 (issued 16.04.18) and from Tallinn Medical Research Ethics Committee 1161 (issued 13.09.2007). The healthy control (HC) samples were from donors of the North Estonia Medical Blood Centre (Tallinn, Estonia).

2.3. Mimotope variation analysis (MVA)

For qualitative and quantitative characterisation of antibody epitopes from blood sera/plasma and CSF samples, we used the MVA method as described previously [16]. Analysis included discovery (sera/plasma and CSF samples; n = 62), validation (sera/plasma and CSF; n = 20) and independent external (sera/plasma; n = 421) cohorts. On average, MVA generated 3 million peptide sequences per sample,

Table 1
Description of clinical samples studied by MVA.

Characteristics	Discovery cohort ON^{ON}	ON^{MS}	Controls CTRL	CTRL	Validation cohort MS^{ON}	MS^{other}
Group size (n)	9*	15**	27	11	10	10
Diagnose on sample collection	ON	ON	-	-	MS	MS
ON development into MS	No	Yes	-	-	Yes	NA
Sample type	CSF(9), serum(5)	CSF(15),serum(10)	serum	CSF	CSF (10)/serum (10)	CSF (10)/serum (10)
Gender (female/male)	3/6	13/3	7/20	8/3	7/3	9/1
Age (average)	35	32	64	31	32	32

Serum=serum or plasma.
  * 1 sample with non-native Finnish background.
  ** 2 samples with non-native Finnish background.

Table 2
External cohorts used for ROC analysis.

Characteristics	Samples from subjects with no demyelinating disease	Samples from healthy donors
Group size (n)	192	229
Sample type	sera/plasma	sera/plasma
Gender (female/male (NA))	141/41(10)	98/131
Age (average) (NA)	43 (10)	39

out of which 350,000 peptides, on average, were with unique amino acid sequence (data structure shown on **Figure S1b**). Altogether, the size of the described antigenic repertoire of the discovery cohort encompassed about 20 million unique peptide sequences (data not shown). Although the majority of these peptides were largely individual-specific as observed by the data structure analysis of the most frequent 5000 and 20,000 peptides from each sample (top 5000 and top20000 peptides, respectively), the studied samples shared a substantial fraction of common peptide antigen characteristics across all datasets (**Figure S1c** shows the heatmap image of MVA profiles from the CSF samples of patients in the discovery cohort).

2.4. Data processing and statistical analysis

Data processing was performed as described previously [16]. Statistical analyses (ANOVA, *t*-test, correlation analyses, Receiver Operating Characteristic (ROC), distribution profiles) and corresponding visualisations were done using MedCalc Statistical Software (version 17.0.4, MedCalc Software Bvba, Belgium). For clustering and hypergeometric test SPEXS2 Software was used. In-house data analysis scripts were used to perform peptide divergence, peptide abundance, and coefficient of variation computations, as well as motif alignments and sequence annotations against user-defined reference sequence database (IEDB.org (01.08.2019)). Excel VBA (Visual Basic for Applications) scripts were used for these data visualisations.

2.5. Measurements of total IgG and protein

Total amount of IgGs in sera/plasma and CSF were measured using ELISPOT method. In brief, diluted sera/plasma samples and CSF samples were printed onto nitrocellulose film slides (Amersham Bioscience) by SpotBot® 4 Personal Microarrayer (Arrayit). Purified human IgG (Sigma, i4506) was used for standard curve preparation with rabbit anti-human IgG (HRP) (Abcam, ab6759; dilution 1:1000) as a secondary antibody. Results were scanned by using Ettan™ DigeImager (GE Healthcare Life Sciences) and quantified by ImageQuant TL (GE Healthcare Life Sciences). Total protein concentration in CSF was measured using Pierce BCA Protein Assay kit (Thermo Scientific) according to the kit protocol. Bovine serum albumin (Naxo) was used for standard curve preparation. absorbence was measured at 450 nm with SpectraMax Paradigm (Molecular Devices).

2.6. ELISPOT of peptide displaying phages

In vitro mutagenesis method was used to generate recombinant phages displaying peptides of interest (TLPMDTSPRAHW (vector for cluster B), TLPMDASPRAHW (control vector of cluster B) and DYKDDDDK (FLAG tag)) in the N-terminus of the pIII of M13. In addition, peptides of cluster A - NETIYNTTLKYGGGGDYKDDD(LYS(BIO-TIN)); control peptides of cluster A - NETIANAAAKAGGGGDYKDDD(LYS(BIOTIN)) were synthesised by Genescript (US). For ELISPOT, peptides or peptide-displaying phages printed onto nitrocellulose filter pads (Amersham Bioscience) by SpotBot® 4 (Arrayit) were exposed to human precleared sera/plasma (dilution 1:100) or CSF (dilution 1:2) for 1 h at room temperature (RT) and then incubated with rabbit anti-human IgG (HRP) (Abcam, ab6759; dilution 1:1000) as a

secondary antibody. Images were scanned using Ettan™ DIGEImager (GE Healthcare Life Sciences).

2.7. Western blot analysis

For western blot (WB) analysis, recombinant phages displaying sequences of interest (ELEKAYKTTLSY (vector of cluster A), TLPMDTSPRAHW (vector of cluster B) and DYKDDDDK (FLAG tag)) at the N-terminus of the pIII of the M13 were generated with in vitro mutagenesis. For WB, 30 μg of protein lysate of 1×10^{13} phages with the following primary antibodies: anti-Flag (Sigma-Aldrich F3165; 1:3600), precleared human sera (dilution 1:750) and precleared human CSF (dilution 1:7,5) samples were incubated with secondary rabbit anti-mouse (Abcam, ab6728; dilution 1:10 000) or rabbit anti-human IgG antibodies (Abcam, ab6759; dilution 1:10 000). The ECL Femto kit (Amersham) was used for detection of target proteins.

2.8. CMV and EBV seropositivity

Human Cytomegalovirus- (CMV) and EBV-specific IgGs were determined by the ISO 17,025 accredited enzyme-linked immunosorbent assay (ELISA). For CMV, for analysis of sera/plasma anti-CMV IgG ELISA kit (EUROIMMUN EI 2570 9601 G) was used. Anti-EBV IgG ELISA kit (EUROIMMUN EI 2731 9601) was used for measuring EBV seropositivity in sera/plasma samples. Analyses were carried out in accordance with the manufacturer's specifications. absorbence was measured at 450 nm with SpectraMax Paradigm (Molecular Devices).

2.9. Role of the funding source

We confirm that all funders played no role in study design, data collection, data analysis, interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

3. Results

3.1. Discovery study of antibody epitope profiles differentiating ON from MS

Effective biomarkers for assessment of ON prognosis, in particular those measurable in blood, are largely lacking. MS is driven by systemic immune activation of autoimmunity against CNS, thereof CSF has been the primary object of biomarker exploration. Regarding this, we performed MVA immunoprofiling of CSF and sera samples on the cohort of 24 subjects i) with isolated optic neuritis ON^ON and ii) with initial diagnosis of ON, who were later diagnosed with MS ((ON^MS), **Fig. 1a**).

A total of 20 million unique peptides from the discovery cohort were included in downstream analyses. This allowed a fine molecular description of the antigenic repertoires of ON patients (data structure analysis shown in **Figure S1b** and heatmap image of MVA profiles in the CSF samples of the patients shown in **Figure S1c**). Analysis of the most abundant fractions resulted in defined sets of peptides that were common to samples of CSF and sera/plasma from the same individual ($r > 0.7$ (Pearson); $p < 0.001$ (ANOVA); **Fig. 1b-c**). Subsets of
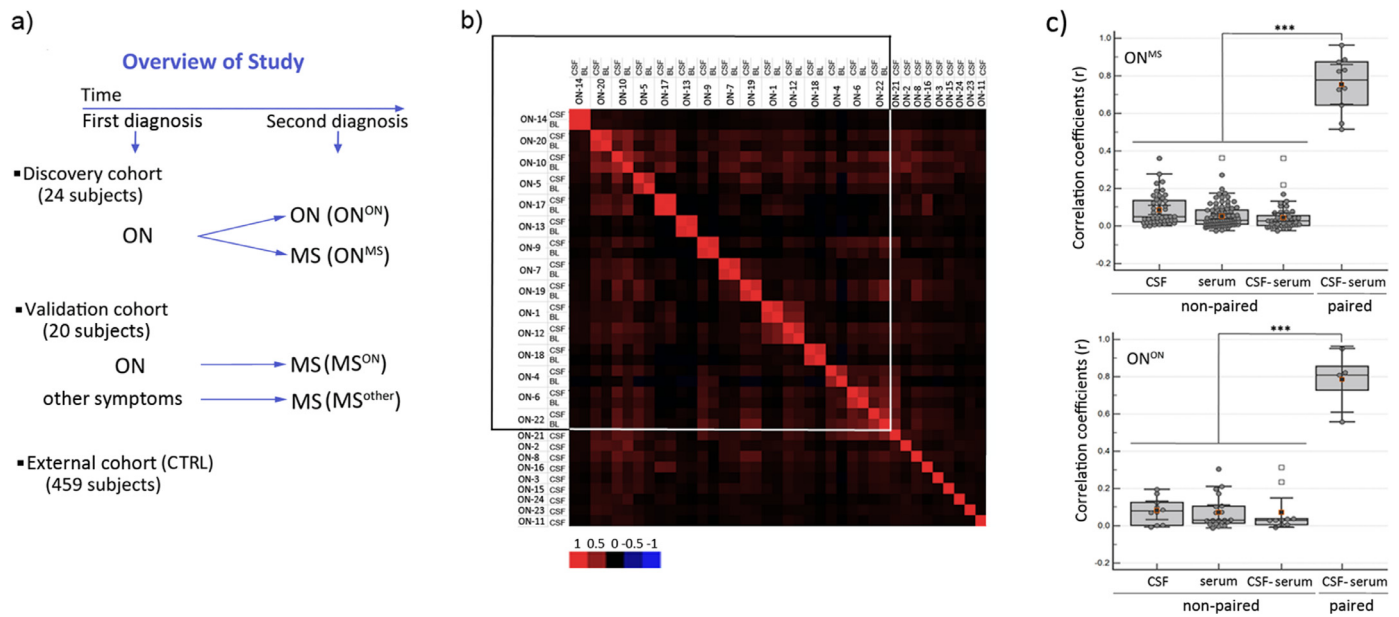
Fig. 1. Highly individual but similar immunoreactive epitope profiles of matched sera and CSF samples as observed by MVA. a) Clinical study design: for discovery cohort, CSF and sera/plasma samples were collected from 24 ON patients, of which 15 were later diagnosed with MS (ON^{MS}), whereas the remaining nine patients were not (ON^{ON}). 27 sera samples and 11 CSF (CTRL) samples from individuals without ON or MS diagnosis were used as controls (see Table 1-2). For validation cohort, independent samples from 10 MS patients with ON onset (MS^{ON}) and 10 with MS but with other symptoms onset (MS^{other}) were collected at diagnostic phase. For ROC analysis of biomarker performance, another independent external cohort including healthy individuals ($n = 229$) and subjects without any known diagnosis of the demyelinating disease ($n = 192$) was added to CTRLs. b) Pearson correlation matrix analysis of Top5000 peptides from sera/plasma and CSF samples of 24 individuals diagnosed with ON revealed high similarity of epitope repertoires in paired CSF and sera/plasma samples. The range of correlation coefficient values is shown below the figure by the colour-coded panel; *boxed samples* - paired CSF and sera/plasma samples. c) Box plots of ANOVA analysis of Pearson correlation values (r) in ON^{MS} ($n = 15$) (*top*) and ON^{ON} ($n = 9$) (*bottom*) peptide datasets are shown. Epitope profiles were dissimilar in non-paired CSF and sera/plasma samples as compared to paired samples. *y-axes* —correlation coefficients (r); *x-axes* — non-paired samples of CSF (1) and sera (2), non-paired CSF and sera/plasma (3) and paired CSF and sera/plasma samples (4)); ***- ANOVA $p < 0.001$.

peptides were also shared between individuals ($r < 0.3$ (Pearson); $p < 0.001$ (t-Test); **Fig. 1b-c).** The exchange of antibody forming cells across the blood brain barrier (BBB) in individuals with ON and MS has previously been suggested from sequencing data of B cell receptors [17] and IgG heavy chain variable regions [18]. Our analysis, by demonstrating close similarity of the immunodominant antibody epitope profiles of blood and CSF, also argues for the free exchange of antibodies across the BBB in individuals with ON and MS.

3.2. Two immunodominant epitopes discriminate between ON^{ON} and ON^{MS}

To examine the group-discriminating features, using unsupervised clustering of the most abundant (immunodominant) peptides from each study cohort (altogether 1.4 million peptides) we delineated 1669 group-specific epitopes ($p < 0.005$ (hypergeometric test)). The overall data analysis scheme is shown in **Figure S2a. Figure S2b** shows the distribution of these 1669 epitopes across study cohorts. Correlation analysis that followed revealed two major consensus motifs, clusters A and B (coefficient of variation > 1.2, **Fig. 2a** and top sequence motifs are listed in **Table S2**). Cluster A contained peptides with Y..TL.Y amino acid patterns, whereas cluster B was formed by P.. T.PR pattern-containing peptides. The heatmap analysis of median peptide abundance as shown in **Fig. 2b** further confirmed the differential distribution of peptides forming clusters A and B in different clinical cohorts.

Next, we evaluated the association of clusters A and B with clinical diagnosis using ANOVA. In blood samples, as shown in **Figs. 2c**, cluster A epitopes were most abundant in ON^{ON} and controls ($p < 0.05$ (ANOVA) as compared to ON^{MS}), whereas a trend of cluster B epitopes being more abundant in ON^{MS} and ON^{ON} as compared with controls was observed ($p = 0.077$ (ANOVA)). Furthermore, the group-discriminating features of clusters A and B were similarly detected in the CSF of ON^{MS} and ON^{ON} as compared with controls (cluster A CTRL vs.

ON^{ON} $p < 0.005$ (ANOVA); cluster B CTRL vs. ON^{MS} $p < 0.05$ (ANOVA); **Fig. 2c**).

3.3. Immunodominant epitopes A and B mimic highly antigenic epitopes of CMV and EBV

Epitopes of clusters A and B harvested the most abundant antibody immune response, suggesting their possible association with common human pathogens. Reviewed epitopes of human pathogens from IEDB database were aligned to consensus sequences of clusters A and B. Peptides of cluster A unequivocally defined the antigen domain 2 (AD2) of the neutralising epitope of CMV glycoprotein B (gB), whereas peptides of cluster B were close mimics of the C-terminal epitope of capsid antigen (VCA) p18 protein of EBV (**Fig. 3**). As shown in **Fig. 3a** and **3b** and further corroborated by western blot analysis by using recombinant phages displaying prototype peptides, epitopes derived of AD2 ($_{70}$ETI**Y**NT**TL**K**Y**$_{80}$) of CMV gB and of EBV VCA p18 ($_{161}$GGQ**P**HD**TAPR**GARKK$_{175}$) could act as true antigens mimicked by clusters, i.e., epitopes of A and B, respectively. Further validation using ELISPOT analyses confirmed the MVA detected seropositivity of epitopes A and B at accuracy of 95%. Comparative analysis of seropositivity by MVA and ELISPOT are shown in **Figure S3a**. The findings of the herpesviral antigenic background of these two epitopes were further strengthened by clinical ELISA testing. The highest anti-CMV seroresponse was observed in the cohorts of ON^{ON}, MS^{other} and CTRL, whereas all study subjects in the discovery cohort were seropositive for anti-VCA EBV (percentages of CMV and EBV seropositive and -negative subjects in study groups are shown in **Figure S3b**). Moreover, all study subjects, who showed antibodies against clusters A or B epitopes in MVA, were seropositive for CMV and EBV, accordingly (**Figure S3c and S3d**).

Interestingly, the alignment analysis of other MVA-defined consensus motifs (**Fig. 2a**) resulted in delineating epitopes associated with Epstein−Barr nuclear antigen 1 (EBNA1) and also with other
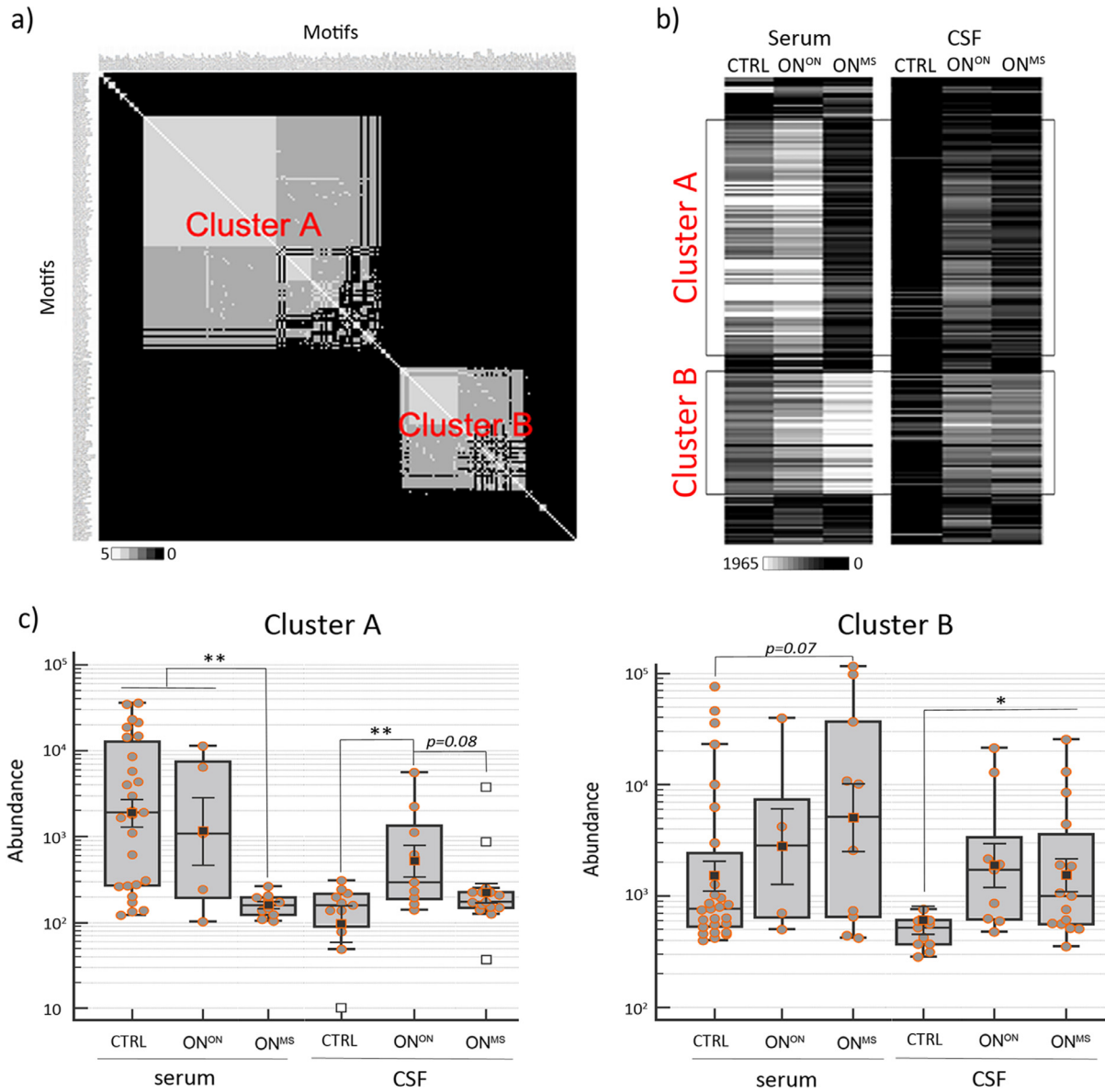
Fig. 2. Epitope clusters A and B with group-specific features discriminate between ON^{MS}, ON^{ON} and CTRL. a) Strong sequence associations in peptides forming epitope clusters A and B became evident by the amino acid positional identity count matrix analysis of 205 motifs underlying these clusters (Figure S2). *White* − high correlation; *black* − low correlation. b) The heatmap depicts distribution of the same 205 motifs as in B according to the median values of motif abundance across different study groups ($n = 64$). *White* − maximum median value (1965); *black* − minimum median value (0). Study cohort acronyms are provided on the top of the heatmap. *CSF, BL* - samples of discovery cohort; *framed boxes* - median values of clusters A and B motifs across different study cohorts. c) Box plot analysis of peptides containing clusters A and B epitopes as detected by MVA in the CSF and sera samples across different study groups ($n = 64$). *y- axes* - peptide abundance (in $\log_{10}$); *x-axes* − study groups; *p-* ANOVA on log-transformed data p value, *- $p < 0.05$; **- $p < 0.01$.

epitopes of gB of CMV (**Table S3**). However, these were detected by MVA at considerably (50 to 135 times) lower abundances as compared to epitopes of A and B. Nonetheless, the epitopes aligning to EBNA1 were clearly more frequent in ON^{MS} as compared to controls (**Table S3**).

3.4. Value of epitopes A and B as clinical biomarkers

We next tested epitopes of clusters A and B in the independent validation cohort of 20 subjects with relapsing MS and in 448 controls (**Fig. 4** and **Tables 1** and **2**). MVA analysis of the blood samples of the validation cohort showed that MS patients either with ON onset (MS^{ON}) or other MS onset (MS^{other}) had roughly similar patterns of response to peptide epitopes of A and B, although these of MS^{ON} differed slightly more from the controls (**Figure S4a**). **Fig. 4a**

shows that by combining sera findings from ON^{MS} and MS^{ON} groups, clusters A and B epitopes specifically discriminated between these diagnostic groups and controls ($p < 0.05$ (ANOVA), **Fig. 4a**). Furthermore, normalising the ELISPOT findings from CSF of MS patients to total IgG content yielded a high positive correlation between CSF and peripheral blood in the immunoreactivity to epitopes of A and B ($r = 0.95$ (Pearson), $p < 0.0001$ (t-Test); $r = 0.71$ (Pearson), $p < 0.0001$ (t-Test); **Fig. 4b).** The measurements of total amounts of proteins and IgGs in the CSF samples of study cohorts are in **Figure S4b**.

We used ROC analysis to assess the predictive value of epitopes of clusters A and B as blood-based diagnostic biomarkers for MS after ON (**Fig. 4c** and **S4c**). The prediction value of these two epitope biomarkers was high in stratifying subjects with ON^{MS} and MS^{ON} ($n = 20$) from healthy controls ($n = 256$) (**Table 1** and **2**) with balanced accuracy of 74.61 at 75% sensitivity and 74.22% specificity (AUC 0.796,
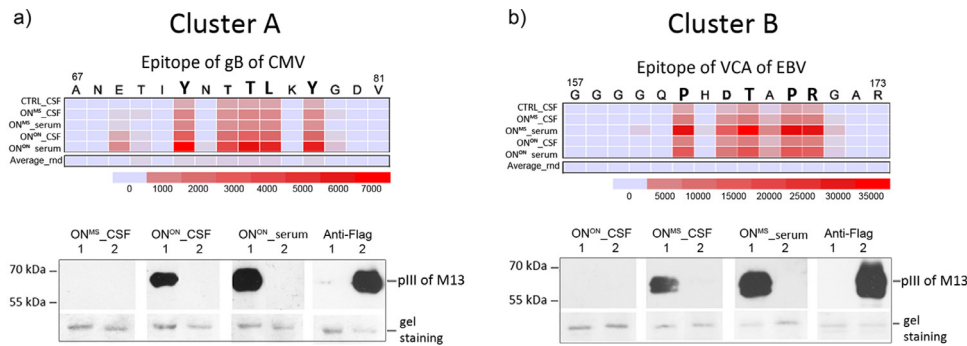
Fig. 3. Epitope clusters A and B mimic highly antigenic epitopes of CMV and EBV. a-b) Heatmap images of epitopes of clusters A and B alignments to IEDB epitopes (01.08.2019). The criteria for homology searches were set to 4 amino acid similarity matches and group median values for each epitope motif were used in alignment calculations. On top of the heat-map panels primary sequences of gB CMV (Uniprot code Q2FAM8) and VCA p18 EBV (Uniprot code P14348) are shown with amino acids defining the core epitopes of clusters A and B. Below the heatmap panels, the scale of relative alignment loads (0−35 000) is shown in colour code. The acronyms of the discovery and validation cohorts are shown in the left. *Average_rnd* − average median values of scrambled motifs derived of clusters A and B motifs. Representative images from validation studies of antigenic epitope predictions using western blot analysis are shown in the lower part of the figure (see full blots on Figure S5). Western blot analysis of recombinant phages containing the epitope gB CMV or VCA EBV was performed using recombinant phages encoding Y..TL.Y-pIII or P..T.PR-pIII (1) and flag-pIII fusion proteins (2) with primary antibodies: i) pre-cleared serum/plasma (dilution 1:750), CSF (dilution 1:7,5); ii) mouse anti-FLAG antibody (Sigma Aldrich, No. 287) and secondary antibodies: i) rabbit anti-human-HRP (Abcam), ii) rabbit anti-mouse-HRP (Abcam). Protein molecular weight markers (kDa) are shown in the left. *Gel staining* − protein loading control with Coomassie blue staining, *pIII* − pIII protein of M13. Source of primary antibodies is indicated on top of the blot.

CI 0.743 to 0.842, *p*<0.0001; **Fig. 4c**). The ROC curve analysis is shown in **Fig S4c**, and the calculated prediction values for each patient are shown in **Table S4**. Moreover, when the independent external cohort of subjects with non-demyelinating disease (*n* = 192; **Table 2**) was used as a control group, the prediction values for clusters A and B epitope biomarkers for blood-based detection of MS with ON onset were even higher with balanced accuracy of 76.62 at 60% sensitivity and 93.23% specificity (AUC 0.82, CI 0.761 to 0.869, *p*<0.0001; **Fig. 4**c and **S4c, Table S4**).

4. Discussion

In this study, we present an unbiased antibody epitope discovery strategy that resulted in delineation of two epitope biomarkers stratifying subjects with ON and those with MS after ON from healthy controls and patients without any demyelinating disease. Using hypothesis-free MVA immunoprofiling approach we observed highly individual immune response profiles in the paired peripheral blood and CSF samples of the study subjects (**Fig. 1**). Group comparative
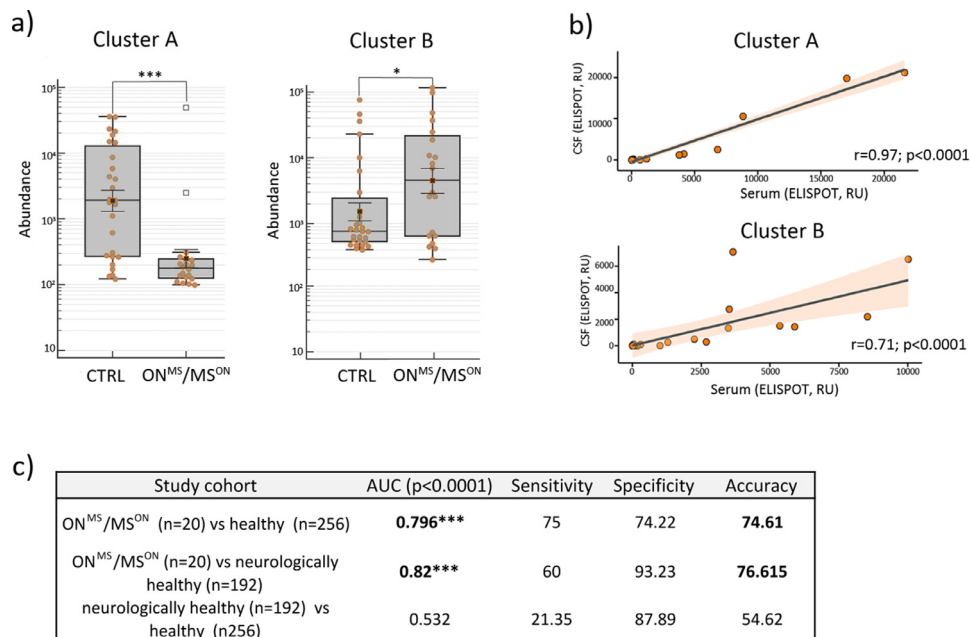

Fig. 4. Epitope clusters A and B predict MS with ON onset at high diagnostic accuracy in validation cohorts. a) Validation of epitopes of clusters A and B by MVA confirmed the diagnostic power of these as two blood-based biomarkers to detect MS with ON onset. Box plot analysis of immunoreactivity values of clusters A and B epitopes as obtained by MVA in sera samples of MS with ON onset (samples of ON^MS and MS^ON, *n* = 20) and controls. *y-axes* - peptide abundance (in log₁₀); *x-axes* −study cohorts; *p*- ANOVA on log-transformed data p value, *- *p*<0.05, ***- *p*<0.001. b) High immunoreactivity to epitopes of clusters A and B in sera was always co-detected with high immunoreactivity in CSF. Pearson correlation analysis of ELISPOT values of gB CMV (cluster A) and VCA p18 of EBV (cluster B) in the CSF samples of MS^ON and MS^other patients (*n* = 20) upon normalisation to total IgG amount. *y-axes* - ELISPOT values in CSF; *x-axes* − ELISPOT values in serum; *r* - correlation coefficient; *p* - correlation significance level (t-Test). c) ROC analysis data of the predictive value of clusters A and B epitope biomarkers in different validation groups − MS with ON onset (samples of ON^MS and MS^ON, *n* = 20; **Table 1-2**), healthy (CTRL samples from discovery cohort (*n* = 27) and samples from external cohorts of healthy individuals (*n* = 256) and subjects with non-demyelinating diseases (*n* = 192; **Table 2**). *Accuracy* − balanced accuracy of sensitivity and specificity; *area under the curve* (AUC).

clustering analysis of peptide epitope profiles identified case-discriminating antigenic repertoires forming two major clusters A and B (**Fig. 2**). Our further analysis determined these epitopes A and B as mimics of highly antigenic epitopes of CMV and EBV (**Fig. 3**). In validation analyses, we confirmed these two epitopes (A and B) as highly potential serologic biomarkers of MS (**Fig. 4**).

Antibody epitope repertoire shared similarity in the CSF and serum. Our study demonstrated the power of next generation phage display MVA method for high-throughput profiling of antibody response at the epitope precision of the clinical samples from the early phase of MS with ON onset. Despite the highly individual top immune response profiles, we observed that paired sera and CSF samples showed highly similar antigenic features (**Fig. 1**), arguing for shared robust antibody responses both in the periphery and in the CNS. Although findings of similar patterns of immunoreactivity to single antigens in the peripheral blood and CSF have been reported (29−31), our current data extended these findings to the depths of millions of peptide antigens. Furthermore, amongst these peptide antigens we describe repertoires (epitopes forming clusters A and B) that can distinguish different demyelinating inflammatory conditions (ON^{ON}, ON^{MS}, MS^{ON} and MS^{other}) and controls (**Figs. 2** and **4**). The emerging questions still remain of the origin and also of role of these loads of antigenic epitopes in the physiology and pathology of the brain. Although our data provided early answers to these inquiries, future studies would be needed to elucidate the mechanisms leading to the peptide antigen signatures described here.

Two highly antigenic epitopes mimicking herpesviruses behind pathophysiology of MS after ON. We show here that the highly antigenic epitopes of CMV gB and EBV VCA p18 (**Fig. 3**) could be true antigens underlying epitope clusters A and B, respectively. These specific epitopes have not, to our knowledge, been previously reported in the neurological disease literature. Previous studies about the associations between viral, bacterial, and helminthic infections and the prevalence of MS contributed towards creation of the "hygiene hypothesis" [19]. Despite plethora of conflicting reports including also those on herpesviruses [20-22], the mechanisms behind pathogens in the incidence of MS have remained unclear. This is the first study describing associations between the highly antigenic epitopes of gB CMV and VCA p18 EBV with MS (**Figs. 2** and **3**) suggesting direct involvement of these two herpesviruses in the pathophysiology of MS. Thus, why and how the humoral response against these two viral epitope mimics is associated with CNS antigens remains an open question.

Epitope-specific serologic biomarkers of ON prognosis and MS disease. We found that different demyelinating inflammatory conditions (ON^{ON}, ON^{MS}, MS^{ON} and MS^{other}) were associated with different antibody epitopes (**Figs. 2** and **4**). Based on these promising data, one would argue that epitopes including the two A and B clusters could detect MS with different onsets. Biomarker risk score model with 74.61% accuracy established epitopes of A and B as blood-based biomarkers of MS with ON onset. For stratifying MS from other nonmyelinating disease, the clinical accuracy of these two biomarkers reached a value as high as 76.62% (**Fig. 4c**). To date, there are no biomarkers that predict ON progression to MS . Most recent findings indicated that serum neurofilament light chain (sNfL) might serve as a biomarker from very early stages of MS (see also ref in [23]). It will be important to investigate the additional diagnostic value of epitopes A and B would carry for patients with first MS symptoms and high initial sNfL levels all combined as serologic biomarkers for the disease course prognosis.

Strengths of this study include the unbiased nature of the discovery phase to define epitopes that predictthe risk of developing MS after ON. Further, we highlight the study cohort of native Finnish (59 participants out of 62), residents of Southern Finland, providing a fair homogeneity of the studied population. In addition, this is the first study in MS clinical research to characterise antibody epitope repertoires of this magnitude from CSF and serum accenting on their potential clinical value. It should be mentioned that application of methods like MVA with high epitope resolution and links to novel antigens and biological pathways clearly facilitate the discovery of immune response specifics. Agreeably, mechanisms of these immune features in underlying pathophysiologies reflecting neurodegenerative processes associated with MS still remain to be elucidated.

Our study was a pilot-scale small sample size study and limited at variable degrees of metadata available for subjects of different cohort. Hence the findings may limit generalisable extrapolation and may include biases of factors such as genetic and environmental factors (smoking, infection history, etc.) that might contribute to distinct MS immunological signatures in the blood. Further longitudinal studies including different pre-symptomatic stages would strengthen our findings and should be able to establish the role of the specific epitopes in MS pathogenesis. Another limitation of this study is that only the two major and highly antigenic epitopes could be thoroughly studied. It also remains unclear how these associations between sera and CSF profiles of epitopes A and B relate to MS pathogenesis and whether these associations are cause or consequence of the disease. Further delineation of detected antibody epitopes is desirable as these might provide further understanding of the pathophysiological context behind MS progression.

Our findings indicate that specific antibody epitope biomarkers of MS exist in the blood of patients at different stages of disease. These epitope biomarkers can be combined into a multivariate model with high discriminatory potential. Further validation of the selected two-epitope biomarker analysis from this initial study are warranted to assess their exact value in MS development. Overall our results stress the importance of dissecting global antibody immunoreactivity patterns at the epitope level towards the personalised care of MS.

Contributors

HS - conceptualisation, data processing, formal analysis, investigation, methodology, software, supervision, validation, data visualisation, writing original draft, reviewing and editing; **AP** – data processing, formal analysis, investigation, methodology, software conceptualisation, formal analysis, visualisation; **MJ** – data processing, investigation, methodology, software conceptualisation, data visualisation, writing, reviewing and editing; **NP** – writing, reviewing and editing; **AR** – writing, reviewing and editing; **MT** – writing, reviewing and editing; **AV** – conceptualisation, writing, reviewing and editing**; JN** – clinical data analysis, writing, reviewing and editing**; MS** – funding acquisition, clinical data analysis, writing, reviewing and editing; **PT** – conceptualisation, patient recruitment, funding acquisition, clinical data analysis, writing, reviewing and editing, and **KP** – conceptualisation, formal analysis, investigation, methodology, visualisation, project administration, supervision, funding acquisition, writing original draft, reviewing and editing. All authors had full access to all underlying data. All authors read and approved the final version of the manuscript.

Data sharing statement

Data on annotation are freely available from the database of IEDB.org. Any data not published within the article or supplementary materials will be shared in anonymised format by request from any qualified investigator. If desired, please contact the corresponding author of this article.

Declaration of Competing Interests

AP and KP are inventors of the patent application (PCT Application No. US/14079626) filed by Protobios that covers the use of phage display method to manipulate and monitor humoral immunity. PT

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ebiom.2021.103211.

References

[1] International Multiple Sclerosis Genetics C, Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kemppinen A, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet 2013;45(11):1353–60.

[2] Ascherio A, Munger KL. EBV and autoimmunity. Curr Top Microbiol Immunol 2015;390(Pt 1):365–85.

[3] Lassmann H. Pathogenic mechanisms associated with different clinical courses of multiple sclerosis. Front Immunol 2018;9:3116.

[4] Stangel M, Fredrikson S, Meinl E, Petzold A, Stuve O, Tumani H. The utility of cerebrospinal fluid analysis in patients with multiple sclerosis. Nat Rev Neurol 2013;9 (5):267–76.

[5] Meinl E, Krumbholz M, Hohlfeld R. B lineage cells in the inflammatory central nervous system environment: migration, maintenance, local antibody production, and therapeutic modulation. Ann Neurol 2006;59(6):880–92.

[6] Derfuss T, Gurkov R, Then Bergh F, Goebels N, Hartmann M, Barz C, et al. Intrathecal antibody production against Chlamydia pneumoniae in multiple sclerosis is part of a polyspecific immune response. Brain 2001;124(Pt 7):1325–35.

[7] Sriram S, Stratton CW, Yao S, Tharp A, Ding L, Bannan JD, et al. Chlamydia pneumoniae infection of the central nervous system in multiple sclerosis. Ann Neurol 1999;46(1):6–14.

[8] Yao SY, Stratton CW, Mitchell WM, Sriram S. CSF oligoclonal bands in MS include antibodies against Chlamydophila antigens. Neurology 2001;56(9):1168–76.

[9] Felgenhauer K, Schadlich HJ, Nekic M, Ackermann R. Cerebrospinal fluid virus antibodies. A diagnostic indicator for multiple sclerosis? J Neurol Sci 1985;71 (2–3):291–9.

[10] Vartdal F, Vandvik B, Norrby E. Viral and bacterial antibody responses in multiple sclerosis. Ann Neurol 1980;8(3):248–55.

[11] Alenda R, Alvarez-Lafuente R, Costa-Frossard L, Arroyo R, Mirete S, Alvarez-Cermeno JC, et al. Identification of the major HHV-6 antigen recognized by cerebrospinal fluid IgG in multiple sclerosis. Eur J Neurol 2014;21(8):1096–101.

[12] Derfuss T, Hohlfeld R, Meinl E. Intrathecal antibody (IgG) production against human herpesvirus type 6 occurs in about 20% of multiple sclerosis patients and might be linked to a polyspecific B-cell response. J Neurol 2005;252(8):968–71.

[13] Leibovitch EC, Jacobson S. Evidence linking HHV-6 with multiple sclerosis: an update. Curr Opin Virol 2014;9:127–33.

[14] Virtanen JO, Pietilainen-Nicklen J, Uotila L, Farkkila M, Vaheri A, Koskiniemi M. Intrathecal human herpesvirus 6 antibodies in multiple sclerosis and other demyelinating diseases presenting as oligoclonal bands in cerebrospinal fluid. J Neuroimmunol 2011;237(1–2):93–7.

[15] Optic Neuritis Study G. Multiple sclerosis risk after optic neuritis: final optic neuritis treatment trial follow-up. Arch Neurol 2008;65(6):727–32.

[16] Sadam H, Pihlak A, Kivil A, Pihelgas S, Jaago M, Adler P, et al. Prostaglandin D2 receptor DP1 antibodies predict vaccine-induced and spontaneous narcolepsy type 1: large-scale study of antibody profiling. EBioMedicine 2018;29:47–59.

[17] von Budingen HC, Kuo TC, Sirota M, van Belle CJ, Apeltsin L, Glanville J, et al. B cell exchange across the blood-brain barrier in multiple sclerosis. J Clin Invest 2012;122(12):4533–43.

[18] Bankoti J, Apeltsin L, Hauser SL, Allen S, Albertolle ME, Witkowska HE, et al. In multiple sclerosis, oligoclonal bands connect to peripheral B-cell responses. Ann Neurol 2014;75(2):266–76.

[19] Leibowitz U, Antonovsky A, Medalie JM, Smith HA, Halpern L, Alter M. Epidemiological study of multiple sclerosis in Israel. II. Multiple sclerosis and level of sanitation. J Neurol Neurosurg Psychiatry 1966;29(1):60–8.

[20] Langer-Gould A, Wu J, Lucas R, Smith J, Gonzales E, Amezcua L, et al. Epstein-Barr virus, cytomegalovirus, and multiple sclerosis susceptibility: a multiethnic study. Neurology 2017;89(13):1330–7.

[21] Maple PAC, Tanasescu R, Gran B, Constantinescu CS. A different response to cytomegalovirus (CMV) and Epstein-Barr virus (EBV) infection in UK people with multiple sclerosis (PwMS) compared to controls. J Infect 2020;80(3):320–5.

[22] Munger KL, Hongell K, Cortese M, Aivo J, Soilu-Hanninen M, Surcel HM, et al. Epstein-barr virus and multiple sclerosis risk in the finnish maternity cohort. Ann Neurol 2019;86(3):436–42.

[23] Bittner S, Steffen F, Uphaus T, Muthuraman M, Fleischer V, Salmen A, et al. Clinical implications of serum neurofilament in newly diagnosed MS patients: a longitudinal multicentre cohort study. EBioMedicine 2020;56:102807.