Clinical

HIGHLIGHTS

FRONTLINE | Research Article

Stereotyped B-cell responses are linked to IgG constant region polymorphisms in multiple sclerosis

Ida Lindeman#1,2,3, Justyna Polak#2,3, Shuo-Wang Qiao1,2,3, Trygve Holmøy4,5, Rune A. Høglund4,5, Frode Vartdal2,3, Pål Berg-Hansen6, Ludvig M. Sollid1,2,3 and Andreas Lossius3,4,7

1 Department of Immunology, Oslo University Hospital, Oslo, Norway
2 Department of Immunology, Institute of Clinical Medicine, University of Oslo, Oslo, Norway
3 K.G. Jebsen Coeliac Disease Research Centre, University of Oslo, Oslo, Norway
4 Department of Neurology, Akershus University Hospital, Lørenskog, Norway
5 Department of Neurology, Institute of Clinical Medicine, University of Oslo, Oslo, Norway
6 Department of Neurology, Oslo University Hospital, Oslo, Norway
7 Department of Molecular Medicine, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway

Clonally related B cells infiltrate the brain, meninges, and cerebrospinal fluid of MS patients, but the mechanisms driving the B-cell response and shaping the immunoglobulin repertoires remain unclear. Here, we used single-cell full-length RNA-seq and BCR reconstruction to simultaneously assess the phenotypes, isotypes, constant region polymorphisms, and the paired heavy- and light-chain repertoires in intrathecal B cells. We detected extensive clonal connections between the memory B cell and antibody-secreting cell (ASC) compartments and observed clonally related cells of different isotypes including IgM/IgG1, IgG1/IgA1, IgG1/IgG2, and IgM/IgA1. There was a strong dominance of the G1m1 allotype constant region polymorphisms in ASCs, but not in memory B cells. Tightly linked to the G1m1 allotype, we found a preferential pairing of the immunoglobulin heavy-chain variable (IGHV)4 gene family with the κ variable (IGKV)1 gene family. The IGHV4-39 gene was most used and showed the highest frequency of pairing with IGKV1-5 and IGKV1(D)-33. These results link IgG constant region polymorphisms to stereotyped B-cell responses in MS and indicate that the intrathecal B-cell response in these patients could be directed against structurally similar epitopes.

Keywords: B cells • Cerebrospinal fluid • Immunoglobulins • Multiple sclerosis • Single-cell RNA sequencing

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Introduction

More than half a century ago, Kabat and colleagues discovered that MS patients have increased levels of IgG in their CSF [1]. Using more sensitive techniques, such as agarose electrophoresis and isolectric focusing, intrathecally synthesized IgG can be detected as oligoclonal bands (OCBs) in more than 90% of patients [2]. OCBs of other isotypes, such as IgM and IgA, can also be found to a variable extent and might be of prognostic value [3, 4]. A proportion of the secreted IgG proteome in the CSF matches IgG transcripts from B-lineage cells in the brain parenchyma, meninges, and the CSF of MS patients, suggesting...
that CSF IgG is secreted by B-lineage cells within these compart-
ments [5,6]. Accordingly, studies have demonstrated an intra-
thal enrichment of antibody-secreting cells (ASCs) that are clon-
ally related [7–11], have undergone somatic hypermutation [7,9–
13], and have B-cell receptors (BCRs) displaying a biased usage
of variable heavy-chain (V\textsubscript{H}) genes toward the
\textit{IGHV4} family [9,10,14]. These studies have provided important biological
insights, but they were based on cloning and sequencing of V\textsubscript{H}
genes from a limited number of single cells or bulk sequencing
that does not permit pairing of V\textsubscript{H} and variable light-chain (V\textsubscript{L})
genes. Analysis of paired V\textsubscript{H}:V\textsubscript{L} sequences from a sufficient num-
ber of single B-lineage cells is key to establish the degree of clonal
expansion, determine the mutational load, accurately trace clonal
evolution based on somatic mutations within both immunoglob-
ulin chains, and to detect convergent B-cell responses with
V\textsubscript{H}:V\textsubscript{L} pairing preferences.

The constant region of IgG1 comprises the G1m allotypes
that are polymorphic markers encoded by the \textit{IGHG1} gene. We
have recently demonstrated that MS patients heterozygous for
the \textit{IGHG1} alleles expressing the G1m1 and G1m3 allotypes of
IgG1 display a selective intrathecal enrichment of ASCs expressing
the G1m1 allotype [15]. This was unexpected, as maturing
B cells in the bone marrow undergo random allelic exclusion of
the immunoglobulin genes. Thus, the immunoglobulin alleles are
expected to be evenly distributed in the B-cell pool of heterozy-
gous carriers, which was the case in the blood of MS patients,
and in the CSF of patients with \textit{Lyme neuroborreliosis} [15] and
\textit{varicella zoster} virus encephalitis [16].

To dissect the mechanisms shaping the intrathecal
immunoglobulin repertoires in MS, we performed an in-depth
study of B-lineage cells in MS patients using a full-length single-
cell RNA-sequencing (scRNA-seq) protocol [17]. We combined
this with BraCeR [18], a bioinformatic pipeline that provides
accurate reconstruction of paired BCR sequences, clonality infer-
ence, and lineage tracing of single B cells. This approach has
recently proven useful in the study of autoimmune plasma cells
in celiac disease [19]. In the present study, it allowed us to simul-
taneously assess the detailed phenotypes, isotypes, allotypes, and
for the first time, trace the clonal evolution of CSF B cells based
on somatic mutations within full-length paired immunoglobulin
V\textsubscript{H}:V\textsubscript{L} sequences.

Results

Full-length RNA-seq of single B-lineage cells in the
CSF of MS patients

We performed flow cytometry index-sorting and scRNA-seq of B-
lineage cells collected from MS patients during diagnostic lum-
bar puncture (Fig. 1, Table 1 and Supporting information Table
S1). ASCs were sorted first, and other B-cell phenotypes were
included on every other 96-well plate if the number of cells was
sufficient (Supporting information Fig. S1). After quality control,
we analyzed the transcriptomic profiles of 1621 ASCs from 21 MS
patients and another 544 B-lineage cells from ten of the patients

Figure 1. Schematics of single B-cell sorting, sequencing library preparation, and data analysis. Fresh CSF samples were stained and analyzed by
flow cytometry. Debris and doublets were excluded (not shown). After excluding CD3\textsuperscript{+}, CD14\textsuperscript{+}, and CD16\textsuperscript{+} cells in a dump channel, CD38\textsuperscript{++}
CD27\textsuperscript{++} single antibody-secreting cells (ASCs) were sorted into 96-well plates containing lysis buffer. For every other plate, we inverted the latter gate and
collected single CD38\textsuperscript{−}/dim CD27\textsuperscript{−}/dim CD19\textsuperscript{+} B cells. From the lysed cells, single-cell RNA-sequencing (scRNA-seq) libraries were prepared following
a modified Smart-Seq2 protocol and sequenced on an Illumina NextSeq500 platform. B-cell receptor (BCR) sequences were reconstructed using
BraCeR. Some of the elements in the figure were modified from Servier Medical Art, licensed under a Creative Common Attribution 3.0 Generic
License. http://smart.servier.com.
Table 1. Patient characteristics

| ID  | Sex | Age | Diagnosis          | Initial presenting symptom | Disease duration (months) | Relapses | Months since last relapse | CSF cell count\(^a\) | OCB\(^b\) | Albumin ratio | IgG index | Allotype\(^c\) | IGHG1 Alleles\(^d\) |
|-----|-----|-----|---------------------|----------------------------|---------------------------|----------|--------------------------|----------------------|----------|---------------|-----------|----------------|------------------|
| MS1 | M   | 27  | RR-MS               | Optic neuritis             | 15                        | 1        | 15                       | 16                   | +        | 5.1           | 0.93      | G1m1          | *02/(*10?)       |
| MS2 | M   | 36  | SP-MS               | Motor symptoms (PTM)       | 120                       | 3        | 12                       | 36                   | +        | 12            | 1.7       | G1m1/G1m3     | *02/03           |
| MS3 | F   | 26  | RR-MS               | Motor symptoms (PTM)       | 36                        | 1        | 36                       | 4                    | +        | 3.8           | 1.1       | G1m3          | *03               |
| MS4 | M   | 20  | RR-MS               | Sensory symptoms (PTM)     | 48                        | 1        | 48                       | 19                   | +        | 6.2           | 0.81      | G1m3          | *03               |
| MS5 | F   | 46  | RR-MS               | Vertigo                    | 24                        | 3        | 3                        | 19                   | +        | 3.1           | 0.72      | G1m3          | *03               |
| MS6 | F   | 49  | RR-MS               | Vertigo                    | 12                        | 1        | 12                       | 4                    | +        | 3.8           | 0.69      | G1m1/G1m3     | *02/03           |
| MS7 | M   | 37  | RR-MS               | Optic neuritis             | 0                         | 1        | 0                        | 5                    | +        | 7.5           | 0.67      | G1m1/G1m3     | *02/03           |
| MS8 | F   | 31  | RR-MS               | Motor symptoms             | 1                         | 1        | 1                        | 33                   | +        | 3.2           | 2.1       | G1m1/G1m3     | *02/03           |
| MS9 | F   | 24  | RR-MS               | Motor symptoms             | 3                         | 1        | 3                        | 4                    | +        | 4.0           | 0.69      | G1m3          | *03               |
| MS10| M   | 52  | RR-MS               | Motor symptoms (PTM)       | 7                         | 1        | 7                        | 6                    | +        | 5.0           | 1.1       | G1m1/G1m3     | *02            |
| MS11| F   | 39  | RR-MS               | Optic neuritis             | 0                         | 1        | 0                        | 18                   | +        | 6.4           | 1.1       | G1m3          | *03               |
| MS12| F   | 39  | RR-MS               | Optic neuritis             | 24                        | 2        | 6                        | 10                   | +        | 3.8           | 0.59      | G1m3          | *03               |
| MS13| F   | 38  | RR-MS               | Sensory symptoms           | 1                         | 2        | 12                       | 4                    | +        | 3.8           | 0.54      | G1m3          | *03               |
| MS14| M   | 21  | RR-MS               | Sensory symptoms (PTM)     | 1                         | 1        | 1                        | 21                   | +        | 6.4           | 0.95      | G1m1/G1m3     | *02/08 or *08    |
| MS15| M   | 36  | RR-MS               | Sensory symptoms           | 7                         | 1        | 7                        | 6                    | +        | 7.2           | 0.54      | G1m3          | *03               |
| MS16| F   | 37  | RR-MS               | Optic neuritis             | 132                       | 1        | 132                      | 8                    | +        | 6.2           | 1.7       | G1m1/G1m3     | *02/03           |
| MS17| M   | 25  | RR-MS               | Sensory symptoms (PTM)     | 0                         | 1        | 0                        | 6                    | +        | 7.5           | 0.58      | G1m1/G1m3     | *02            |
| MS18| F   | 44  | RR-MS               | Sensory symptoms           | 9                         | 2        | 2                        | 4                    | +        | 5.6           | 0.66      | G1m3          | *03               |
| MS19| F   | 48  | RR-MS               | Sensory symptoms           | 72                        | 2        | 0                        | 17                   | +        | 4.3           | 1.3       | G1m1/G1m3     | *07/03           |
| MS20| M   | 35  | RR-MS               | Motor symptoms             | 12                        | 2        | 0                        | 5                    | +        | 3.3           | 0.75      | G1m1/G1m3     | *02/03           |
| MS21| F   | 39  | RR-MS               | Optic neuritis             | 48                        | 2        | 13                       | 9                    | +        | 2.8           | 1.5       | G1m1/G1m3     | *02            |

\(^a\) Mononuclear cells count per microliter of CSF.

\(^b\) Positive (+) for more than two CSF restricted oligoclonal bands on isoelectric focusing.

\(^c\) Determined in serum.

\(^d\) Inferred computationally by quantifying the transcription of each allele in IgG1-expressing cells.

\(^e\) Too few cells to determine if patient is homozygous for IGHG1*02 or also carries another allele coding for G1m1.

\(^f\) G1m3-encoding allele could not be determined computationally or is uncertain due to too few cells and/or strong skewing toward G1m1 usage.

F, female; M, male; PTM, partial transverse myelitis; RR-MS, relapsing-remitting multiple sclerosis; SP-MS, secondary progressive multiple sclerosis.
ASCs in the CSF constitute a continuity from immature plasmablasts to more differentiated phenotypes

The phenotype of ASCs in the CSF of MS patients has not been unequivocally established [15,20,22]. We found that approximately 22% of the ASCs expressed the proliferation marker MKI67 (Ki-67), suggesting they were plasmablasts rather than end-differentiated plasma cells (Fig. 2E and Supporting information Fig. S2C). Notably, the ASCs clustered based on inferred cell-cycle phase when visualized by UMAP (Fig. 2E), and MKI67 was as anticipated, expressed in cells assigned to the G2M or S phases. This held true also when excluding MKI67 from the list of genes that is used to infer cell-cycle phase (Supporting information Fig. S2C). When visualized by UMAP, ASCs assigned to the nonproliferative G1/0 phase (66.9%) displayed a seemingly gradual transition between a major population of ASCs with high expression of SDC1 (CD138) and a smaller population expressing MS4A1 (CD20) (Fig. 2D). The percentage of ASCs in the different cell-cycle phases was, however, not significantly different from the other B-cell lineages (Fig. 2E, right panel). Taken together, these data suggest that ASCs in the CSF of MS patients are proliferating plasmablasts within different stages of differentiation. However, a proportion of the ASCs in the inferred G1/0 phase may represent true end-differentiated nondividing plasma cells.

IgG, IgA, and IgM/IgD B-lineage cells show distinct transcriptional profiles and mutational load

Full-length recombinant BCR sequences, stretching into the constant region, allowed us to define the isotype of sorted B-lineage cells. As expected from earlier studies [15,23], IgG1 was by far the most prevalent isotype among ASCs, while both IgM and IgG1 isotypes were often used in the memory subset (Fig. 2F). We observed a gradual transition between naive, nonclass-switched

Table 2. Features used to phenotype B-lineage cells

| B-cell population                  | Phenotype                        | Marker                                      |
|-----------------------------------|-----------------------------------|---------------------------------------------|
| Antibody secreting cells (ASCs)   |                                   | CD27++/CD38++                               |
| Naive                             |                                   | CD27−, CD19+ IgD and/or IgM; mutation rate <0.01; mutations <5% |
| Memory                            | Non-switched CD27++               | CD27+, CD19+ IgG or IgA                     |
|                                   | Non-switched CD27−/dim            | CD27−/dim, CD19+ IgG or IgM                 |
|                                   | Switched CD27++                   | CD27++, CD19+ IgG or IgA                    |
|                                   | Switched CD27−/dim                | CD27−/dim, CD19+ IgG or IgM                 |

ASCs in the CSF constitute a continuity from immature plasmablasts to more differentiated phenotypes

The phenotype of ASCs in the CSF of MS patients has not been unequivocally established [15,20,22]. We found that approximately 22% of the ASCs expressed the proliferation marker MKI67 (Ki-67), suggesting they were plasmablasts rather than end-differentiated plasma cells (Fig. 2E and Supporting information Fig. S2C). Notably, the ASCs clustered based on inferred cell-cycle phase when visualized by UMAP (Fig. 2E), and MKI67 was as anticipated, expressed in cells assigned to the G2M or S phases. This held true also when excluding MKI67 from the list of genes that is used to infer cell-cycle phase (Supporting information Fig. S2C). When visualized by UMAP, ASCs assigned to the nonproliferative G1/0 phase (66.9%) displayed a seemingly gradual transition between a major population of ASCs with high expression of SDC1 (CD138) and a smaller population expressing MS4A1 (CD20) (Fig. 2D). The percentage of ASCs in the different cell-cycle phases was, however, not significantly different from the other B-cell lineages (Fig. 2E, right panel). Taken together, these data suggest that ASCs in the CSF of MS patients are proliferating plasmablasts within different stages of differentiation. However, a proportion of the ASCs in the inferred G1/0 phase may represent true end-differentiated nondividing plasma cells.

IgG, IgA, and IgM/IgD B-lineage cells show distinct transcriptional profiles and mutational load

Full-length recombinant BCR sequences, stretching into the constant region, allowed us to define the isotype of sorted B-lineage cells. As expected from earlier studies [15,23], IgG1 was by far the most prevalent isotype among ASCs, while both IgM and IgG1 isotypes were often used in the memory subset (Fig. 2F). We observed a gradual transition between naive, nonclass-switched

Table 2. Features used to phenotype B-lineage cells

| B-cell population                  | Phenotype                        | Marker                                      |
|-----------------------------------|-----------------------------------|---------------------------------------------|
| Antibody secreting cells (ASCs)   |                                   | CD27++/CD38++                               |
| Naive                             |                                   | CD27−, CD19+ IgD and/or IgM; mutation rate <0.01; mutations <5% |
| Memory                            | Non-switched CD27++               | CD27+, CD19+ IgG or IgA                     |
|                                   | Non-switched CD27−/dim            | CD27−/dim, CD19+ IgG or IgM                 |
|                                   | Switched CD27++                   | CD27++, CD19+ IgG or IgA                    |
|                                   | Switched CD27−/dim                | CD27−/dim, CD19+ IgG or IgM                 |

ASCs in the CSF constitute a continuity from immature plasmablasts to more differentiated phenotypes

The phenotype of ASCs in the CSF of MS patients has not been unequivocally established [15,20,22]. We found that approximately 22% of the ASCs expressed the proliferation marker MKI67 (Ki-67), suggesting they were plasmablasts rather than end-differentiated plasma cells (Fig. 2E and Supporting information Fig. S2C). Notably, the ASCs clustered based on inferred cell-cycle phase when visualized by UMAP (Fig. 2E), and MKI67 was as anticipated, expressed in cells assigned to the G2M or S phases. This held true also when excluding MKI67 from the list of genes that is used to infer cell-cycle phase (Supporting information Fig. S2C). When visualized by UMAP, ASCs assigned to the nonproliferative G1/0 phase (66.9%) displayed a seemingly gradual transition between a major population of ASCs with high expression of SDC1 (CD138) and a smaller population expressing MS4A1 (CD20) (Fig. 2D). The percentage of ASCs in the different cell-cycle phases was, however, not significantly different from the other B-cell lineages (Fig. 2E, right panel). Taken together, these data suggest that ASCs in the CSF of MS patients are proliferating plasmablasts within different stages of differentiation. However, a proportion of the ASCs in the inferred G1/0 phase may represent true end-differentiated nondividing plasma cells.

IgG, IgA, and IgM/IgD B-lineage cells show distinct transcriptional profiles and mutational load

Full-length recombinant BCR sequences, stretching into the constant region, allowed us to define the isotype of sorted B-lineage cells. As expected from earlier studies [15,23], IgG1 was by far the most prevalent isotype among ASCs, while both IgM and IgG1 isotypes were often used in the memory subset (Fig. 2F). We observed a gradual transition between naive, nonclass-switched

Table 2. Features used to phenotype B-lineage cells

| B-cell population                  | Phenotype                        | Marker                                      |
|-----------------------------------|-----------------------------------|---------------------------------------------|
| Antibody secreting cells (ASCs)   |                                   | CD27++/CD38++                               |
| Naive                             |                                   | CD27−, CD19+ IgD and/or IgM; mutation rate <0.01; mutations <5% |
| Memory                            | Non-switched CD27++               | CD27+, CD19+ IgG or IgA                     |
|                                   | Non-switched CD27−/dim            | CD27−/dim, CD19+ IgG or IgM                 |
|                                   | Switched CD27++                   | CD27++, CD19+ IgG or IgA                    |
|                                   | Switched CD27−/dim                | CD27−/dim, CD19+ IgG or IgM                 |

ASCs in the CSF constitute a continuity from immature plasmablasts to more differentiated phenotypes

The phenotype of ASCs in the CSF of MS patients has not been unequivocally established [15,20,22]. We found that approximately 22% of the ASCs expressed the proliferation marker MKI67 (Ki-67), suggesting they were plasmablasts rather than end-differentiated plasma cells (Fig. 2E and Supporting information Fig. S2C). Notably, the ASCs clustered based on inferred cell-cycle phase when visualized by UMAP (Fig. 2E), and MKI67 was as anticipated, expressed in cells assigned to the G2M or S phases. This held true also when excluding MKI67 from the list of genes that is used to infer cell-cycle phase (Supporting information Fig. S2C). When visualized by UMAP, ASCs assigned to the nonproliferative G1/0 phase (66.9%) displayed a seemingly gradual transition between a major population of ASCs with high expression of SDC1 (CD138) and a smaller population expressing MS4A1 (CD20) (Fig. 2D). The percentage of ASCs in the different cell-cycle phases was, however, not significantly different from the other B-cell lineages (Fig. 2E, right panel). Taken together, these data suggest that ASCs in the CSF of MS patients are proliferating plasmablasts within different stages of differentiation. However, a proportion of the ASCs in the inferred G1/0 phase may represent true end-differentiated nondividing plasma cells.

IgG, IgA, and IgM/IgD B-lineage cells show distinct transcriptional profiles and mutational load

Full-length recombinant BCR sequences, stretching into the constant region, allowed us to define the isotype of sorted B-lineage cells. As expected from earlier studies [15,23], IgG1 was by far the most prevalent isotype among ASCs, while both IgM and IgG1 isotypes were often used in the memory subset (Fig. 2F). We observed a gradual transition between naive, nonclass-switched
Figure 2. Transcriptional and mutational profiling of B-lineage cells from the cerebrospinal fluid of multiple sclerosis patients. (A) Frequencies of sorted CD19+ B cells being classified as naïve or memory B-cell subsets (Table 2). (B) UMAP projection of all B-lineage cells (upper left; n = 21,216 cells), only the CD19+ B cells (upper right; n = 10,544 cells), and all B-lineage cells colored according to isotype (lower left) and IgG1 (G1m) allotype (lower right). (C) Heatmap showing expression of genes of particular interest in all B-lineage cells in (B). ASC: antibody-secreting cell, HLA: human leukocyte antigen. (D) UMAP projections colored by genes of interest or median fluorescence intensity (MFI) of the cell surface markers CD19, CD27, and CD38 (bottom row) obtained during index sorting. (E) MKI67 expression and inferred cell-cycle phase in each B-lineage population. (F) Isotype distribution based on reconstructed B-cell receptor (BCR) sequences. Each data point represents a patient with more than one cell for a given cell
and class-switched memory B cells when visualized by UMAP (Fig. 2G). Genes that were found to be significantly higher expressed in memory B cells of the IgM/IgD compared to IgG and IgA isotypes included DUSP1, CD69, and RGS2 (encoding regulator of G-protein signaling 2). CTS (encoding cathepsin H) and CD86 were significantly higher expressed in IgG memory B cells, suggesting an increased importance of antigen presentation in these cells. IgG memory B cells also had the highest expression of CXCR3, which is shown to be important for recruitment of B-lineage cells in inflamed tissues [24]. In contrast to a recent report [25], we did not detect a distinct regulatory phenotype of IgA-expressing B-lineage cells in the CSF, and we did not find evidence to support that IgA-expressing B-lineage cells express higher levels of gut-homing markers than their IgM/IgD and IgG counterparts (Fig. 2G).

Next, we investigated the mutational load of each reconstructed BCR sequence. The numbers of somatic mutations varied between populations (Fig. 2H) and isotype (sub)classes (Fig. 2I and J). In line with previous reports [26, 27], CD27−/dim memory B cells displayed significantly fewer mutations than their CD27+ counterparts. However, this was not statistically significant on an individual level when excluding MS2, MS19, and MS20 from the analysis (Supporting information Table S4). Among ASCs, we observed the highest number of mutations in class-switched cells. In the memory B-cell population, IgD-only cells showed the highest mutational load, followed by class-switched memory cells (Fig. 2I and J).

### B-lineage cells in the CSF of MS patients undergo intrathecal maturation

Next, we analyzed clonal relationships between B-lineage cells using BraCeR. BraCeR, for paired heavy and light chains, assigns clonal relationships between cells, which are then depicted as clonal networks, and additionally, lineage trees for individual clones, consisting of cells with different mutational patterns (Fig. 3). In all 21 patients, clonally related cells were detected, with a median of 61.6% [39.7–83.1%] of all ASCs belonging to a clone (Fig. 3A and B, and Supporting information Table S5). For the first time, we traced the clonal evolution of B-lineage cells based on somatic mutations in the paired full-length variable heavy- and light-chain genes. We reconstructed lineage trees in 17 patients (Fig. 3C, Supporting information Fig. S3A and Table S5).

Interestingly, we commonly observed highly mutated clones consisting solely of cells with identical mutational patterns (Supporting information Table S5). The four patients with no reconstructed lineage trees were in the lower range of sorted ASCs and displayed clonally expanded ASCs with the exact same mutational patterns.

Our approach, where both heavy and light chains are involved in determining clonal relations combined with transcriptomic profiling of cells, made it possible to trace linkage between memory B cells and ASCs. Strikingly, we observed clonal sharing between memory B and ASCs in seven out of ten patients from whom both cell populations were sorted (Fig. 3D and E, and Supporting information Fig. S3B). The majority of memory cells belonging to clones containing ASCs shared identical V-regions with at least one ASC, while we also observed the acquisition of mutations. In one patient, we found clonally expanded memory B cells with no detected members of the clones among the ASCs (Fig. 3F). The lack of connections between memory B cells and ASC in three patients could possibly be attributed to low cell numbers (Fig. 3E). We also explored the clonal evolution of the ASCs in relation to their cell-cycle stage and found extensive clonal connections between cells in different stages (Supporting information Fig. S4A–D).

The most commonly found expanded sequences other than IgG1 were of the IgG2 isotype subclass (Table 3, Supporting information Fig. S4E). In eight patients, we also found expanded IgA1 and/or IgM ASCs. Additionally, we detected expanded IgG3 ASCs in two patients and expanded IgG4 ASCs in one patient. Interestingly, we found examples of clones comprising B-lineage cells of different isotypes, including IgM and IgG1, IgG1 and IgA1, IgG1 and IgG2, and IgM and IgA1 (Fig. 3D and G). This could suggest that class-switch recombination has taken place intrathecally and that at least a proportion of the IgG, IgM, and IgA B-lineage cells share a common origin and target the same antigens.

### ASCs show a more biased usage of the G1m1 allotype than memory B cells

BraCeR infers the CH1 region of the IGHC alleles of reconstructed heavy chains, allowing for detection of isotypes and distinction between the G1m1 and G1m3 allotypes on a cellular level. Due to random allelic exclusion of immunoglobulin genes, there is a 50/50 distribution of the IGHC alleles among B-lineage cells in
Figure 3. B-lineage cells in the CSF are clonally related, undergo somatic hypermutation, isotype switching, and intrathecal maturation. (A) Example of clonal relationships between antibody secreting cells (ASCs) of patient MS15 inferred by BraCeR. (B) Sizes of clonal groups found for ASCs in each patient. The colors indicate number of cells belonging to each clone group, size of boxes is proportionate to the size of clones. Only clonally expanded cells are shown. The grey columns on the top depict the total number of ASCs included in the analysis for each patient. (C) Representative ASC lineage trees generated by BraCeR for two clones, with inferred germinal sequences at the root. The number of mutations between each node is shown next to the branch. The size of each node is proportionate to the number of cells containing a given unique BCR sequence. The number
the blood of heterozygous individuals [15]. We have previously
shown, based on flow cytometry, that G1m1-expressing ASCs are
preferentially enriched in the CSF of MS patients [15]. To
address this on a transcriptomic level, we limited the analysis
of G1m1/G1m3 heterozygous patients and calculated the propor-
tion of G1m1-expressing cells among IgG1 ASCs and memory
B cells (Fig. 4A and B, Supporting information Table S6). In
agreement with our previous study, the results showed a predom-
inance of G1m1 ASCs (Fig. 4A; median 97.8% [66.7–100%]).
On an individual level, a statistically significant skewing toward
the G1m1 allotype in the ASC population was seen in 10 of
11 of the G1m1/G1m3 heterozygous patients (Fig. 4B). The
G1m1 bias was significantly higher in ASCs than memory B cells
(Fig. 4A).

ASCs expressing G1m1 show preferential pairing of
\( V_H \) and \( V_L \) genes

Previous studies have shown that the \( IGHV4 \) gene segments are
dominating the CNS and CSF heavy-chain repertoires [14,28].
We stratified the heavy-chain repertoires of ASCs according to
G1m allotype and found that a preferential use of \( IGHV4 \) gene
segments is associated with the G1m1 allotype, but not with
the G1m3 allotype (Fig. 5A and Supporting information Fig.
S5A and B). In the G1m1-expressing memory B-cell population,
on the other hand, we did not detect a similar bias of \( IGHV4 \)
gene segments usage (Supporting information Fig. S5B). To
confirm the association between \( IGHV4 \) gene segments and G1m1
in a completely independent patient cohort, we reanalyzed \( IGHV \)
sequencing data from bulk CSF B-lineage cells of 12 MS patients
previously published by us [10,29]. We used the total \( IGHV \)
pool, driving bias toward ASCs-derived sequences (Supporting
information Fig. S6A) [30], but analyzing unique sequences yielded
similar results (Supporting information Fig. S6B). Next, we inves-
tigated the repertoire of the light chains in single IgG1 ASCs. The
results showed that a strong preference for the \( \kappa \) light chain was
connected to the G1m1 allotype of the paired heavy chain, but not
the G1m3 allotype (Fig. 5B and Supporting information Fig. S7A).
No significant \( \kappa \) light chain bias was seen for G1m1-expressing
memory B cells, but this could possibly be attributed to few data
points and large variation (Supporting information Fig. S7A). We
also investigated the use of light-chain constant region genes and
alleles (Supporting information Fig. S7B and C). In agreement
with the expected high allele frequency of the Km3 allotype in
Caucasians [31], we detected use of the Km3-encoding \( IGKC*01 \)
allele in all patients with \( \kappa \)-expressing cells. Two patients were

Table 3. Clonal expansion of other isotypes than IgG1

| ID   | IgA1 | IgM | IgG2 | IgG3 | IgG4 |
|------|------|-----|------|------|------|
| MS2  | -    | -   | 1 (2) | -    | -    |
| MS3  | -    | 1 (2) | 1 (11) | -    | -    |
| MS4  | -    | 2 (3) | 1 (4) | -    | -    |
| MS5  | -    | 2 (3-5) | 1 (2) | 1 (6) | -    |
| MS6  | -    | 2 (2) | 1 (2) | -    | -    |
| MS8  | 1 (2) | -   | 2 (2–7) | -    | -    |
| MS9  | 1 (2) | -   | 1 (2) | -    | -    |
| MS10 | 2 (2) | -   | -    | -    | -    |
| MS11 | 1 (2) | 1 (3) | 2 (2–3) | 1 (3) | -    |
| MS14 | -    | -   | 2 (7–12) | -    | 1 (2) |
| MS20 | 1 (5)\(^a\) | 1 (2) | 3 (2–22)\(^b\) | -    | -    |

Numbers of clonotypes for each patient with numbers in brackets indi-
cating sizes of clonal expansion.

All cells are expanded ASCs except the shaded, which are expanded
memory B cells.

\(^a\) Clone additionally contains one IgM ASC.

\(^b\) Part of a large expanded clone containing 22 IgG2 and 19 IgG1 ASCs.

The intrathecal ASC population is skewed toward the G1m1
allotype in G1m1/G1m3 heterozygous MS patients. (A) G1m1 allotype
distribution in IgG1 memory B cells (\( n = 6 \)) and antibody-secreting cells
(ASCs; \( n = 11 \)) of G1m1/G1m3 heterozygous patients. Each dot rep-
resents the percentage of IgG1 cells that are of the G1m1 allotype for
each patient and horizontal solid lines depict median for all patients.
Lines connect ASCs and memory B cells from the same patients. The
patient with SP-MS is indicated in green, while the two patients pre-
viously treated with corticosteroids are indicated in blue. Frequencies
of G1m1 cells were compared within each population (one-sided binom-
ial test) and between paired memory and ASC populations (paired
t-test), where four patients were excluded due to absence of a memory
B-cell population (no CD38\(^{dim}\)/CD27\(^{dim}\)/CD19\(^{+}\) cells sorted) and one patient due to only one IgG1 memory B cell being present for this
patient. The horizontal dashed line at 50% represents the expected dis-
tribution of G1m1 cells due to random allelic exclusion. The indicated
p-value significance levels are: \( * p < 0.05; ** p < 0.001; ns > 0.05. \) (B)
p-values resulting from two-sided binomial tests for each population
within each patient. Solid horizontal lines indicate median values. The
horizontal dashed line represents statistical significance (\( p = 0.05. \) (A–
B) Data shown are analyses of ASCs from 11 G1m1/G1m3 heterozygous
MS patients and memory B cells from six of the patients.
heterozygous for IGKC*01 and the Km1,2-encoding IGKC*04 alleles, and exhibited no preferential usage of either alleles (Supporting information Fig. S7C). No serological allotypes are known for the \(\lambda\) light chain, which instead exhibits variation through a variable number of different IGLC genes. We observed no strong preference of any specific IGLC genes in G1m1 and G1m3 cells (Supporting information Fig. S7C).

The analysis of single B-lineage cells allowed us to gain insight into the \(\text{V}_{\text{H}}:\text{V}_{\text{L}}\) pairings, which is key to understanding antigen-driven B-cell responses and may sometimes show specific signatory combinations [32,33]. To exclude the influence of clonal expansion on the pairing frequencies, only one sequence per clonotype was considered. In ASCs expressing G1m1, we observed a preferential pairing of IGHV4 with IGKV1, and to a lesser degree,
IGKV3 gene segments (Fig. 5C). The same pairing preference was not observed for G1m3-expressing ASCs. This pattern was also evident when taking clonal expansion into account (Supporting information Fig. S8A), but it was not clear in the memory B-cell population (Supporting information Fig. S8B). Finally, we explored the pairing frequencies of genes within the IGHV4 and IGKV1 gene families in G1m1-expressing ASCs. The IGHV4-39 gene was mostly used and showed the highest frequency of pairing with IGKV1-5 (frequency of 9.9%) and IGKV1(D)-33 (8.4%) (Fig. 5D).

Discussion

In the present study, we analyzed the paired full-length rearranged immunoglobulin heavy- and light-chain gene usage and transcriptome of intrathecal B-lineage cells in MS. We achieved this by combining full-length transcriptomic profiling of single cells along with analysis using BraCeR. The results reveal a common pattern across patients with a preferential V_H:V_L pairing of the nature of the antigen, the stereotyped B-cell response with the G1m1 bias is not present to the same degree in intrathecal memory B cells. The mechanisms driving such a biased maturation are currently not known, but one important clue comes from the findings of a connection between the G1m1 allotype and particular V_H and V_L gene segments, suggesting that the ASCs target particular antigenic epitopes. Linkage disequilibrium between constant region alleles expressing the light chain [41].

Preferential pairing of IGHV4 and IGKV1 in the BCR of ASCs expressing G1m1 bears a resemblance to stereotyped B-cell responses observed in other diseases, including celiac disease [42], HIV infection [43], and influenza infection [44], that are driven by particular antigenic epitopes [45]. The target antigens of the intrathecal humoral immune response in MS, however, have not been unequivocally defined. Some early studies reported reactivity against myelin-associated antigens [46,47] and EBV [48], but none of these findings have been reproduced in independent studies [8,49]. More recent studies have shown that some antibodies expressed by CSF B cells recognize neuronal nuclei and/or astrocytes [50] and lead to demyelination in spinal cord explant cultures [51]. Finally, a study reported that a proportion of CSF IgG might target cellular debris [52]. Nevertheless, independent of the nature of the antigen, the stereotyped B-cell response with a preferential V_H:V_L pairing demonstrated here argues that these B-lineage cells target a set of epitopes that may be shared between patients.

We have previously demonstrated that the G1m1 allotype dominates the intrathecal humoral immune response in G1m1/G1m3 heterozygous MS patients, but not in controls with neuroborreliosis [15]. The present study confirms the observation in MS patients using gene expression analysis of single CSF ASCs and extends the findings demonstrating that the G1m1 bias is not present to the same degree in intrathecal memory B cells. The mechanistic driving such a biased maturation are currently not known, but one important clue comes from the findings of a connection between the G1m1 allotype and particular V_H and V_L gene segments, suggesting that the ASCs target particular antigenic epitopes. Linkage disequilibrium between constant region alleles expressing the light chain [41].

Previous studies have demonstrated that immunoglobulin heavy-chain sequences from the brain, CSF and cervical lymph nodes of MS patients are clonally expanded and have undergone somatic hypermutation, which is indicative of an antigen-driven immune response [7,9,10,12, 14,21,34-36]. IGHV4 gene segments have been portrayed as dominant in the CSF and brain of patients with MS and clinically isolated syndrome [14,28,37], and have also been shown to acquire specific mutations and displaying increased specificity toward neuroantigens [38,39]. At a closer look, however, it seems that a IGHV bias is only present in a proportion of patients, although this has not been made a matter of contention by previous studies [9,10]. The present results offer a potential explanation, which is possible to verify in previously published data if the G1m1 carrier status of the patients is determined. Indeed, reanalyzing data from our own group that were generated using bulk sequencing of CSF B cells and a multiplex PCR technique confirmed the association between the G1m1 carrier status and a biased usage of IGHV4 gene segments [10,29]. Curiously, the present data also show a connection between G1m1 expression and κ light-chain usage. While κ light-chain usage was highly dominant among ASCs of G1m1-carriers, several G1m3 homozygous patients showed a predominance of κ light chains in the ASC population. Thus, the G1m1 carrier status could have implications for the diagnostic sensitivity of free light-chain levels in the CSF, which is increasingly being recognized as a valuable diagnostic test in MS [40] and might correlate with the frequency of ASCs expressing the light chain [41].
CNS B cells [56]. Although this has not been investigated during neuroinflammation in humans, it introduces the possibility that intrathecal B cells in MS could originate within or immediately adjacent to the CNS. In line with this idea, B-cell follicles have been demonstrated in the meninges of patients with longstanding disease [57] and could also be present already at an earlier stage [58]. Expression of the activation-induced cytidine deaminase gene (AICDA), which encodes the DNA-editing deaminase involved in somatic hypermutation and class-switch recombination, has previously been detected in intrathecal B-lineage cells [13]. In the present study, we were able to confirm this observation in a small proportion of the single cells investigated. More importantly, we observed extensive connections between the intrathecal memory and ASC compartments and found evidence of isotype switching from IgM to IgG1, from IgG1 to IgA1, and from IgG1 to IgG2.

Our observations of a relatively high mutational rate in IgM ASCs in the CSF of MS patients are in line with a previous study that demonstrated a very high level of somatic mutations in IgM chains in CSF compared to blood [13]. We also detected expanded and extensively mutated IgA ASCs, which confirms early studies detecting IgA-producing cells in the CSF [59,60]. A recent study found that a proportion of IgA-producing cells in the CSF recognize gut microbiota, and the authors proposed that such cells could represent a population of regulatory cells [25]. In the present study, however, we did not detect any distinct regulatory phenotype of IgA-expressing B-lineage cells in the CSF. We also observed a few instances of clonal connections between IgA1 and IgG1/IgM isotypes, arguing against disparate origins. However, it is possible that different populations of IgA plasma cells in MS may play divergent roles in the disease process.

This is the first study to provide a detailed clonal evolution of B-lineage cells in the CSF of MS patients based on somatic mutations in paired heavy- and light-chain sequences. In contrast to previous studies utilizing different variants of bulk amplification and sequencing of heavy-chain transcripts [9–11], we found evidence of a more focused humoral immune response characterized by smaller lineage trees with fewer offspring per ancestor node. This may seem surprising, as one might have expected larger lineage trees due to the fact that we were able to detect mutations also within the light chains. A likely explanation is that high-throughput bulk sequencing introduces PCR and sequencing errors, which are inherently impossible to discern from somatic mutations [61]. Accordingly, studies using bulk sequencing of heavy-chain transcripts are losing important information of somatic mutations in the light-chain genes, while at the same time, inevitably interpreting PCR and sequencing errors as somatic mutations in heavy-chain genes.

There are limitations to our study that need to be recognized. The number of B-lineage cells in the CSF of MS patients is generally scarce, and it is important to keep in mind that the number of cells in a collected sample only represents a small part of a larger population of B-lineage cells with a plethora of different phenotypes within the CNS. Underscoring this fact, we recently demonstrated that the clonal overlap of CSF B cells at two different time points is much lower than the overlap of the secreted CSF IgG proteome [29]. Moreover, the lack of a control group of patients with other inflammatory neurological diseases does not allow us to conclude that the transcriptomic profile of CSF B cells is unique for MS. The strong dominance of the G1m1 allotype constant region polymorphisms in ASCs, on the other hand, recapitulates our previous findings on a protein level [15]. In this previous study, the G1m1 allotype bias was not present in the CSF of patients with Lyme neuroborreliosis, indicating that it is not a general consequence of neuroinflammation.

Our study represents one of the most detailed investigations of B-lineage cells in the CSF of MS patients to date. For the first time, we demonstrate a strong connection between a G1m allotype and the heavy- and light-chain variable gene repertoires. We anticipate that such interactions can point to biological effects of the G1m allotypes that might be of importance in humoral immune responses in MS.

### Materials and methods

#### Patients and sample collection

Twenty-four subjects with symptoms and brain MRI scans strongly suggestive of MS were recruited during diagnostic work-up at the Departments of Neurology at Akershus University Hospital and Oslo University Hospital. All the patients were of Caucasian ethnicity. Three patients were excluded due to too low frequency of ASCs and/or absence of OCBs (Supporting information Table S1). Based on MRI scans, clinical symptoms, and CSF findings, 20 included patients met the criteria of relapsing-remitting MS according to the 2017 McDonald revisions [62], and one patient had developed a secondary-progressive disease (Table 1). MS19 and MS20 had been previously treated with intravenous infusions of methylprednisolone at 1 gram per day for 3 days within 2 weeks of inclusion. None of the other patients had received corticosteroids or any other type of immunomodulatory treatment. The study was approved by the Committee for Research Ethics at the South-Eastern Norwegian Health Authority (2009/23, S-04143a). All patients signed a written informed consent.

CSF and serum were collected during diagnostic lumbar puncture. The first 7 mL of CSF was collected for diagnostic purposes, and 2 × 9 mL CSF was subsequently collected in 15 mL Falcon tubes. Collected CSF was centrifuged (1600 rpm, 10 min, room temperature [RT]) and cell pellets were resuspended in approximately 100 μL of supernatant. Cell count and the presence of red blood cells was checked in a Bürker counting chamber, and none of the CSF samples contained enough cells to indicate contamination of lymphocytes from blood. Serum was left at RT for at least 30 min after collection, centrifuged (2800 rpm, 10 min, RT) and stored at −80°C. It was subsequently thawed and used to determine the expression of the G1m allotypes of each subject by ELISA according to a previously published protocol [15].
Single-cell sorting by flow cytometry

For flow cytometry cell sorting, we adhered to the guidelines for the use of flow cytometry and cell sorting in immunological studies [63]. The CSF cell suspension was transferred to a V-bottom plate, centrifuged (1600 rpm, 4 min, 4°C) and the supernatant was removed. Cells were resuspended in 40 μL of Live/Dead Fixable Violet (Molecular Probes L34956) according to the manufacturer’s instructions and incubated at RT for 15 min in the dark. A total of 10 μL of 30% BSA was added, followed by a mixture of the following anti-human antibodies from BD Biosciences: CD3-BV510 (HIT3a), CD14-BV510 (MqP9), CD16-BVS10 (3G8), CD19-AP488 (HIB19), CD27-PE-CF594 (M-T271), CD38-APC (HIT2), IgD-APC-H7 (IA6-2). Cell suspensions were incubated for 30 min on ice, centrifuged (1600 rpm, 4 min, 4°C), and washed once with flow buffer (PBS, 0.5% BSA, 2 mM EDTA). Cells were resuspended in 200 μL of flow buffer and taken for single-cell index sorting on a FACS ARIA III cell sorter equipped with 408, 488, 561, and 633 nm lasers (BD Biosciences) at the flow cytometry core facility at the Oslo University Hospital. Single ASCs and B cells were sorted from the same tube into separate plates (Fig. 1). The ASCs were sorted first, and other B-cell phenotypes were included on every other 96-well plate, if the number of cells was sufficient. We performed index sorting into 96-well plates (BioRad) containing 2 μL 1:20 RNase inhibitor [Clontech] in 0.2% v/v Triton X-100 [Sigma]).

Generation of scRNA-seq libraries

Plates containing sorted single cells were spun down (2500 rpm, 5 min, 4°C) and stored at –80°C or immediately processed for cDNA synthesis. cDNA synthesis was done following the Smart-Seq2 protocol [17] with minor modifications as described below. For reverse transcription, SmartScribe RT (Clontech) and corresponding buffer was used. Further, a modified TSO primer with biotinylation was introduced: Bio-AAGCAGTTGATCAACGCAGAGTACGrGrG. In the cDNA amplification step, ASCs underwent 21 cycles of amplification and CD19+ B cells were amplified with 22 cycles. Amplified cDNA was purified with 20 μL Ampure XP (Agencourt) beads per well. Tagmentation was done using in-house made Tn5 transposase [64] dually indexed with Nextera (XT) N7xx and S5xx index primers (125 nm). Purified libraries, consisting of up to 384 cells each, were sequenced on the Illumina NextSeq500 platform at the Norwegian Sequencing Centre using 75 bp paired-end reads, resulting in approximately 1 million reads per cell.

Processing of raw sequencing data

Low-quality sequences and adapter sequences were trimmed off the raw scRNA-seq reads with Cutadapt version 1.18 [65] and Trim Galore version 0.6.1 in paired-end mode. We then quantified transcript expression using Salmon version 0.11.3 [66], building the salmon index using cDNA sequences from GRCh38.94 and a k-mer length of 25. Quantified transcripts were subsequently collapsed to gene level, and corrected for transcript length using tximport version 1.8.0 [67]. Quality control was performed in R version 3.5.3 using the scater package [68], and was based on the following measures: Number of detected genes and reads, percent mitochondrial genes, reads mapping to the reference, percent immunoglobulin genes, and successful reconstruction of at least one productively rearranged heavy-chain BCR sequence by BraCeR [18]. Library-specific threshold values are given in Supporting information Table S2.

B-cell receptor analysis

Raw reads were provided as input to BraCeR [18] in assemble mode to reconstruct full-length paired heavy- and light-BCR chains for each cell. We ran BraCeR assembly with--threshold 5000 for ASCs and--threshold 500 for B cells to filter out reconstructed BCRs likely arising from well-to-well contamination or PCR errors in indices. BraCeR yielded the full-length recombined V-region of heavy and light chains in most cells, and additionally sufficient coverage of the constant region to determine isotypes and G1m1 or G1m3 allotypes based on perfect, full-length alignment to the CH1 region of the IGHG1 alleles.

BraCeR was subsequently used in summarise mode to identify clonal relationships between cells and construct lineage trees based on paired heavy and light chains. We identified clonally related B-lineage cells separately for each patient with the following parameters: --include_multiplets --infer_lineage. Likely cell multiplets were then manually removed as previously described [69]. All heavy- and light-chain rearrangements used in the subsequent analyses were productive.

Information regarding variable (V), diversity (D), and joining (J) gene segment usage, light chain, isotype, clonality, etc. was extracted from the summary files generated by BraCeR and used for more extensive BCR repertoire analysis using Change-O and Alakazam version 0.3.0 [70]. The number of V-segment somatic mutations was identified for the most highly expressed heavy and light chain in each cell using IMG/M/HighV-QUEST [71] and normalized by the length of the V-segment of each reconstructed sequence. To remove ambiguous gene assignment of duplicated V genes, those gene segments were collapsed into newly created subgroups, for example, IGKV1-33 and IGKV1D-33 were collapsed into IGKV1D-33. A Circos plot, visualizing the IGHV4:IGKV1 gene pairing frequencies, was generated using the Circos Table Viewer [72].

IGHG1 allele inference

While the G1m allotypes expressed by each patient were determined serologically, we also computationally inferred the specific IGHG1 alleles carried by each individual using Salmon [66] and
a custom script. In short, reads from IgG1 cells were mapped to all IGHG1 reference allele sequences obtained from IMGT, and the IGHG1 allele with highest mapping rate to the CH1, CH2, and CH3 parts of IGHG1 was determined as the IGHG1 allele for each cell. Allele assignments for each cell for each patient were manually inspected, and patient-specific alleles were determined based on the most frequent allele assignments.

Gene expression analysis and phenotyping of B-lineage cells

Subsets of B-lineage cells were initially determined according to the surface expression of CD27 and CD38 measured by flow cytometry, isotype inferred by BraCeR, and somatic mutation load (Table 2). Immunoglobulin genes are frequently expressed in B-lineage cells and were discarded from the gene expression analysis before normalization to avoid masking of non-immunoglobulin-related transcriptional differences. Genes with more than three reads detected in more than five cells were retained. The gene expression matrix was subset according to cells of interest, and subsequently normalized by total reads and logarithmized as $X = \ln(X + 1)$ with scanny version 1.4.4 [73]. We then identified highly variable genes using the highly_variable_genes function of scanny ($min\ mean = 0.1$, $max\ mean = 10$, $min\ disp = 0.25$). To remove batch effects and reduce patient-to-patient variation, we regressed out number of detected genes and reads, percent mitochondrial genes, and patient-specific variation, while retaining variation explained by cell type, using NaiveDE (https://github.com/Teichlab/NaiveDE) and the highly variable genes. Last, the expression matrix was scaled using sklearn (scikit-learn v0.21.3) [74]. We then used scanny to run Principal Component Analysis of the scaled expression matrices, and visualized the populations using UMAP for dimension reduction [75]. The first 25 principal components were used for visualization of all the cells together, while only the first seven principal components were used for visualization of only the CD19+ B cells. Cell cycle phase of each cell was inferred using the score_genes_cell_cycle function of scanny using the provided regev_lab_cell_cycle_genes.txt file for specifying genes associated with the S and G2M phases.

Statistics

Statistical analysis and plots were made using R version 3.5.3 and ggpubr version 0.2.4. Somatic mutation frequencies were compared between all sequences of different isotypes or B-lineage populations using a Wilcoxon rank-sum test, while median somatic mutation frequencies for each patient were compared across B-lineage populations with a Wilcoxon signed-rank test. Frequencies of G1m1 cells were compared within ASCs and memory B cells using a binomial test. Paired t-test was used to analyze differences in G1m1 usage between ASCs and memory B cells and between G1m1 and G1m3 cells in G1m1/G1m3 heterozygous patients. Wilcoxon rank-sum test was used to test differences in Igκ usage for IgG1 ASCs expressing G1m1 or G1m3 and proportion of cells using IGHV family genes depending on the expressed G1m allotype. All tests were two-sided with a significance level of 0.05 unless otherwise stated in the figure legends. To ensure that the patient with SP-MS (MS2) and the two patients who had been previously treated with corticosteroids (MS19 and MS20) did not introduce any bias, all statistical analyses were performed with and without these three patients (Supporting information Table S4).

Acknowledgments: We would like to express our gratitude to patients who participated in this study. We thank Łukasz Wyrożemski for invaluable assistance in performing experiments. The study was funded by grants from the Norwegian Women’s Public Health Association, the Norwegian Research Council (Project No. 314376), the University of Oslo World-leading research program on human immunology (WL-IMMUNOLOGY), and by Stiftelsen KG Jebsen (Project No. SKGJ-MED-017). In addition, A.L. is the recipient of Per B. Larsen’s grant 2018, Odd Fellow’s research Grant 2019, and the Norwegian Neurological Association research prize in multiple sclerosis 2019 provided by Sanofi Genzyme.

Conflict of interest: AL and JP are recipients of unrestricted research grants provided by Sanofi Genzyme and Novartis, respectively.

Author contributions: AL, JP, LMS, FV, and TH conceived the study and designed the experiments. AL, RH, TH, and PBH recruited patients and provided patient data. AL, IL, JP, SWQ, and FV conducted experiments. IL, JP and AL conducted data analysis. AL, IL, and JP drafted the manuscript. All authors contributed to revising the manuscript.

Peer review: The peer review history for this article is available at https://publons.com/publon/10.1002/eji.202149576.

Data availability statement: The raw scRNA-seq fastq files, reconstructed BCR sequences and accompanying metadata have been deposited at the European Genome-Phenome Archive under accession number EGAS00001005745, with controlled access.

References

1. Kabat, E. A., Glusman, M. and Knaub, V., Quantitative estimation of the albumin and gamma globulin in normal and pathologic cerebrospinal fluid by immunochemical methods. Am. J. Med. 1948. 4: 653–662.
2. Stangel, M., Fredriksson, S., Meini, L., Petzold, A., Stuve, O. and Tumani, H., The utility of cerebrospinal fluid analysis in patients with multiple sclerosis. Nat. Rev. Neurol. 2013. 9: 267–276.
Clonal relationships of CSF B cells in treatment-naive patients with multiple sclerosis: equally diverse and complex. J. Immunol. 2020. S: eabc1791.

Fectau, J. F., Coté, G. and Nerón, S., A new memory CD27-IgG+ B cell population in peripheral blood expressing VH genes with low frequency of somatic mutation. J. Immunol. 2006. 177: 3728–3736.

Berkowska, M. A., Driessen, G. J., Bikov, V., Grossericht-Wagner, C., Stamatopoulos, K., Cerutti, A., He, B. et al., Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways. Blood 2011. 118: 2150–2158.

Owens, G. P., Kraus, H., Burgoon, M. P., Smith-Jensen, T., Devlin, M. E. and Gilden, D. H., Restricted use of VH4 germline segments in an acute multiple sclerosis brain. Ann. Neurol. 1998. 43: 236–243.

Tomescu-Baciu, A., Johansen, J. N., Holmøy, T., Greiff, V., Stensland, M., de Souza, G. A., Vadtal, F. et al., Persistence of intrathecal oligoclonal B cells and IgG in multiple sclerosis. J. Neuroimmunol. 2019. 333: 576966.

Chen-Bettecken, U., Wecker, T. and Schimpf, A., Transcriptional control of μ- and κ-Gene expression in resting and bacterial lipopolysaccharide-activated normal B cells. Immunobiology 1987. 174: 162–176.

Lefranc, M. P. and Lefranc, G., Human Gm, Km, and Am allotypes and their molecular characterization: a remarkable demonstration of polymorphism. Methods Mol. Biol. 2012. 882: 635–680.

DeKosky, B. J., Lungu, O. I., Park, D., Johnson, E. L., Charab, W., Chrysotomou, C., Kuroda, D. et al., Large-scale sequence and structural comparisons of human naive and antigen-experienced antibody repertoires. Proc. Natl. Acad. Sci. U. S. A. 2016. 113: E2636–2645.

Roy, B., Neumann, R. S., Snir, O., Iversen, R., Sandve, G. K., Lundin, K. E. A. and Solidi, L. M., High-throughput single-cell analysis of B cell receptor usage among autoantigen-specific plasma cells in celiac disease. J. Immunol. 2017. 199: 782–791.

Stern, J. N., Yaari, G., Vander Heiden, J. A., Church, G., Donahue, W. F., Hintzen, R. Q., Husniy, A. et al., B cells populating the multiple sclerosis brain mature in the draining cervical lymph nodes. Sci. Transl. Med. 2014. 6: 248ra107.

Owens, G. P., Ritchie, A. M., Burgoon, M. P., Williamson, R. A., Corboy, J. R. and Gilden, D. H., Single-cell repertoire analysis demonstrates that clonal expansion is a prominent feature of the B cell response in multiple sclerosis cerebrospinal fluid. J. Immunol. 2003. 171: 2725–2733.

Qin, Y., Duquette, P., Zhang, Y., Okele, M., Da, R. R., Richardson, J., Antel, J. P. et al., Intrathecal B-cell clonal expansion, an early sign of humoral immunity, in the cerebrospinal fluid of patients with clinically isolated...
syndrome suggestive of multiple sclerosis. Lab. Invest. 2003. 83: 1081–1088.

37 Bennett, J. L., Haubold, K., Ritchie, A. M., Edwards, S. J., Burgoon, M., Shearer, A. J., Gilden, D. H. et al., CSF IgG heavy-chain bias in patients at the time of a clinically isolated syndrome. J. Neuroimmunol. 2008. 199: 126–132.

38 Cameron, E. M., Spencer, S., Lazarini, J., Harp, C. T., Ward, E. S., Burgoon, M., Owens, G. P. et al., Potential of a unique antibody gene signature to predict conversion to clinically definite multiple sclerosis. J. Neuroimmunol. 2009. 213: 123–130.

39 Rivas, J. R., Ireland, S. J., Chkhaidze, R., Rounds, W. H., Lim, J., Johnson, J., Ramirez, D. M. et al., Peripheral VH4+ plasmablasts demonstrate autoreactive B cell expansion toward brain antigens in early multiple sclerosis patients. Acta Neuropathol. 2017. 133: 43–60.

40 Leurs, C. E., Twaalfhoven, H., Lissenberg-Witte, B. I., van Pesch, V., Dujmovic, I., Drulovic, J., Castellazzi, M. et al., Kappa free light chains is a valid tool in the diagnostics of MS: A large multicenter study. Mult. Scler. 2020. 26: 912–923.

41 Rathbone, E., Durant, L., Kinsella, J., Parker, A. R., Hassan-Smith, G., Douglas, M. R. and Cournos, S. J., Cerebrospinal fluid immunoglobulin light chain ratios predict disease progression in multiple sclerosis. J. Neurol. Neurosurg. Psychiatry. 2018. 89: 1044–1049.

42 Di Niro, R., Mesin, L., Zheng, N. Y., Stannaes, J., Morrissey, M., Lee, J. H., Huang, M. et al., High abundance of plasma cells secreting transglutaminase 2-specific IgA autoantibodies with limited somatic hypermutation in celiac disease intestinal lesions. Nat. Med. 2012. 18: 441–445.

43 Gorny, M. K., Wang, X. H., Williams, C., Volsky, B., Revész, K., Witover, B., Burda, S. et al., Preferred use of the VH5-1 gene segment by the human immune response to code for antibodies against the V3 domain of HIV-1. Mol. Immunol. 2009. 46: 917–926.

44 Ekiert, D. C., Babbha, G., Elsliger, M. A., Friesen, R. H., Jongeneelen, M., Throsby, M., Goudsmidt, J. et al., Antibody recognition of a highly conserved influenza virus epitope. Science 2009. 324: 246–251.

45 Henry Dunand, C. J. and Wilson, P. C., Restricted, canonical, stereotyped and convergent immunoglobulin responses. Philos. Trans. R. Soc. Lond. B Biol. Sci. 2015. 370.

46 O’Connor, K. C., Appel, H., Bregoli, L., Call, M. E., Catz, I., Chan, J. A., Moore, N. H. et al., Antibodies from inflamed central nervous system tissue recognize myelin oligodendrocyte glycoprotein. J. Immuno. 2005. 175: 1974–1982.

47 Kanter, J. L., Narayana, S., Ho, P. P., Catz, I., Warren, K. G., Sobel, R. A. Steinman, L. et al., Lipid microarrays identify key mediators of autoimmune brain inflammation. Nat. Med. 2006. 12: 138–143.

48 Cepok, S., Zhou, D., Srivastava, R., Nessler, S., Stei, S., Bussow, K., Sommer, N. et al., Identification of Epstein-Barr virus proteins as putative targets of the immune response in multiple sclerosis. J. Clin. Invest. 2005. 115: 1352–1362.

49 Otto, C., Oltmann, A., Stein, A., Frenzel, K., Schroeter, J., Habel, P., Gartner, B. et al., Intrathecal EBV antibodies are part of the polyspecific immune response in multiple sclerosis. Neurology 2011. 76: 1316–1321.

50 Ligocki, A. J., Rivas, J. R., Rounds, W. H., Guzman, A. A., Li, M., Spadaro, M., Lahey, L. et al., A distinct class of antibodies may be an indicator of gray matter autoimmunity in early and established relapsing remitting multiple sclerosis patients. ASN Neuro. 2015. 7: 1759091415609613.

51 Blauth, K., Soltys, J., Matschulat, A., Reiter, C. R., Ritchie, A., Baird, N. L., Bennett, J. L. et al., Antibodies produced by clonally expanded plasma cells in multiple sclerosis cerebrospinal fluid cause demyelination of spinal cord explants. Acta Neuropathol. 2015. 130: 765–781.

52 Brandle, S. M., Obermeier, B., Senel, M., Bruder, J., Mentele, R., Khademi, M., Olsson, T. et al., Distinct oligoclonal band antibodies in multiple sclerosis recognize ubiquitous self-proteins. Proc. Natl. Acad. Sci. U. S. A. 2016. 113: 7864–7869.

53 Torres, M. and Casadevall, A., The immunoglobulin constant region contributes to affinity and specificity. Trends Immunol. 2008. 29: 91–97.

54 Buck, D., Albrecht, E., Aslam, M., Goris, A., Hauenstein, N., Jochim, A., International Multiple Sclerosis Genetics Consortium, et al., Genetic variants in the immunoglobulin heavy chain locus are associated with the IgG index in multiple sclerosis. Ann. Neurol. 2013. 73: 86–94.

55 Delgado-Garcia, M., Matesanz, F., Alcina, A., Fedetz, M., Garcia-Sanchez, M. I., Ruiz-Pena, J. L., Fernandez, O. et al., A new risk variant for multiple sclerosis at the immunoglobulin heavy chain locus associates with intrathecal IgG, IgM index and oligoclonal bands. Mult. Scler. 2015. 21: 1104–1111.

56 Brioschi, S., Wang, W. L., Peng, W., Wang, M., Shchukina, I., Greenberg, Z. J., Bando, J. K. et al., Heterogeneity of meningeal B cells reveals a lymphopoietic niche at the CNS borders. Science 2021. 373: eaab9277.

57 Serafini, B., Rosicicrelli, B., Magliozzi, R., Stigliano, E. and Aloisi, F., Detection of ectopic B-cell follicles with germinal centers in the meninges of patients with secondary progressive multiple sclerosis. Brain Pathol. 2004. 14: 164–174.

58 Absinta, M., Vuolo, L., Rao, A., Nair, G., Sati, P., Cortese, I. C., Ohayon, J. et al., Gadolinium-based MRI characterization of leptomeningeal inflammation in multiple sclerosis. Neurology 2015. 85: 18–28.

59 Eniri, M., Immunoglobulin-containing cells in multiple-sclerosis plaques. Lancet 1977. 310: 478–480.

60 Henriksson, A., Kam-Hansen, S. and Link, H., IgM, IgA and IgG producing cells in cerebrospinal fluid and peripheral blood in multiple sclerosis. Clin. Exp. Immunol. 1985. 62: 176–184.

61 Robasky, K., Lewis, N. E. and Church, G. M., The role of replicates for error mitigation in next-generation sequencing. Nat. Rev. Genet. 2014. 15: 56–62.

62 Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., Coetzee, T., Comi, G., Correale, J. et al., Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol. 2018. 17: 162–173.

63 Cossarizza, A., Chang, H. D., Radbruch, A., Acs, A., Adam, D., Adam-Klages, S., Agace, W. W. et al., Guidelines for the use of flow cytometry and cell sorting in immunological studies (second edition). Eur. J. Immunol. 2019. 49: 1457–1973.

64 Picelli, S., Bjorklund, A. K., Reinius, B., Sagasser, S., Winberg, G. and Sandberg, R., Tn5 transposase and fragmentation procedures for massively scaled sequencing projects. Genome Res. 2014. 24: 2033–2040.

65 Martin, M., Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBO J. 2011. 17: 10–12.

66 Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. and Kingsford, C., Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 2017. 14: 417–419.

67 Soneson, C., Love, M. I. and Robinson, M. D., Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res 2015. 4: 1521.

68 McCarthy, D. J., Campbell, K. R., Lun, A. T. and Wills, Q. F., Scater: preprocessing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics 2017. 33: 1170–1186.
69 Lindeman, I. and Stubbington, M. J. T., Antigen Receptor Sequence reconstruction and clonality inference from scRNA-Seq data. Methods Mol. Biol. 2019. 1935: 223–249.

70 Gupta, N. T., Vander Heiden, J. A., Uduman, M., Gadala-Maria, D., Yaari, G. and Kleinstein, S. H., Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. Bioinformatics 2015. 31: 3356–3358.

71 Brochet, X., Lefranc, M. P. and Giudicelli, V., IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. Nucleic. Acids. Res. 2008. 36: W503–W508.

72 Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J. et al., Circos: an information aesthetic for comparative genomics. Genome Res. 2009. 19: 1639–1645.

73 Wolf, F. A., Angerer, P. and Theis, F. J., SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018. 19: 15.

74 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M. et al., Scikit-learn: machine learning in Python. J. Mach. Learn Res. 2011. 12: 2825–2830.

75 McInnes, L., Healy, J. and Melville, J., UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv:1802.03426v3 [preprint] 2018.

Abbreviations: ASC: antibody-secreting cells · BCR: B-cell receptor · OCBs: oligoclonal bands · RT: room temperature · scRNA-seq: single-cell RNA-sequencing · UMAP: Uniform Manifold Approximation and Projection

Full correspondence: Dr. Andreas Lossius, Department of Molecular Medicine, Institute of Basic Medical Sciences, University of Oslo, PO Box 1110 Blindern, 0317 Oslo, Norway.
e-mail: andreas.lossius@medisin.uio.no

Received: 12/8/2021
Revised: 19/11/2021
Accepted: 10/1/2022
Accepted article online: 29/1/2022