Potential of a unique antibody gene signature to predict conversion to clinically definite multiple sclerosis

Elizabeth M. Cameron, Sade Spencer, Jonathan Lazarini, Christopher T. Harp, E. Sally Ward, Mark Burgoon, Gregory P. Owens, Michael K. Racke, Jeffrey L. Bennett, Elliot M. Frohman, Nancy L. Monson

Article history:
Received 13 March 2009
Received in revised form 19 May 2009
Accepted 29 May 2009

Keywords:
Multiple sclerosis
B lymphocytes
Antibodies
Gene rearrangement
Mutational signature

1. Introduction

B cells have historically been implicated in the pathogenesis of multiple sclerosis (MS) since elevated central nervous system (CNS) gamma immunoglobulins were first described in MS patients in the 1940s (Kabat et al., 1948). Additional evidence of B cell involvement in MS pathogenesis is extensive (reviewed in Antel and Bar-Or, 2006; Duddy and Bar-Or, 2006; Franciotta et al., 2008; Frohman et al., 2006; Martin Mdel and Monson, 2007; McLaughlin and Wucherpfennig, 2008; Owens et al., 2006) and has been recently substantiated by the efficacy of rituximab (Rituxan), a B cell depleting antibody, in a cohort of patients at risk to develop MS predicted conversion to MS with 91% accuracy in a small cohort of clinically isolated syndrome patients. If confirmed, signature prevalence would be a novel genetic diagnostic tool candidate for patients with early demyelinating disease of the central nervous system.

We identified a unique antibody gene mutation pattern (i.e. “signature”) in cerebrospinal fluid (CSF) B cells from multiple sclerosis (MS) patients not present in control populations. Prevalence of the signature in CSF B cells of patients at risk to develop MS predicted conversion to MS with 91% accuracy in a small cohort of clinically isolated syndrome patients. If confirmed, signature prevalence would be a novel genetic diagnostic tool candidate for patients with early demyelinating disease of the central nervous system.

© 2009 Elsevier B.V. All rights reserved.
cells of MS patients would harbor antibody gene mutation patterns that would be distinct from VH4 expressing peripheral B cells derived from healthy controls. To address this contention, we characterized antibody gene mutations in a VH4 subdatabase extracted from the parent heavy chain antibody database consisting of 373 CSF-derived B cells from 11 patients with definite MS. Our analysis revealed a unique pattern of antibody gene replacement mutations in CSF B cells from MS patients that was not prevalent in antibody gene repertoires from CSF B cells of OND patients. Furthermore, prevalence of this conspicuous signature in B cell antibody repertoires from patients with a first inflammatory demyelinating episode (a clinically isolated syndrome; CIS) can predict conversion to clinically definite MS (CDMS) within 3–18 months after initial sampling.

2. Materials and methods

2.1. Patient description

CSF was collected from 10 RRMS patients, one PPMS patient (M484), three patients with other neurological diseases (OND341, ataxia; OND758, headache, and OND116, chronic inflammatory demyelinating polyneuropathy), and two patients with one demyelinating event suggestive of MS (i.e. Clinically Isolated Syndrome (CIS)) at UT Southwestern Medical Center (UTSWMC) (Harp et al., 2007; Monson et al., 2005) in accordance with the UTSWMC Institutional Review Board (IRB). CSF was collected from nine patients with CIS at University of Colorado Denver (UCD) as previously described (Bennett et al., 2006) in accordance with the UCD IRB. The CIS patients had a single episode of demyelination (optic neuritis, brainstem or spinal cord syndrome), and the majority had multiple lesions on MRI satisfying the dissemination in space criterion of the McDonald criteria. None of the patients had received immunomodulatory agents for at least 1 month prior to lumbar puncture. A second relapse of sample acquisition, thus not fulfilling the dissemination in time criterion (McDonald et al., 2001; Polman et al., 2005). Subsequent definition of definite MS was made using the revised McDonald criteria (Polman et al., 2005). Conversion to definite MS was not revealed to the antibody sequence analysis team until after signature score predictions had been calculated.

2.2. MS and CIS B cell antibody database generation

At UTSWMC, antibody repertoires were generated from CD19+ CSF B cells using single cell PCR as previously described (Harp et al., 2007; Monson et al., 2005). The MSCSF database consists of antibody rearrangements from 373 CD19+ CSF B cells from 10 RRMS and 1 PPMS patient recruited at UTSWMC. The CISC SF database consists of antibody rearrangements from 302 CD19+ CSF B Cells from 10 CIS patients (ON4-8 did not have a CD19+ B cell antibody repertoire) and 226 CD138+ CSF plasma cells from 7 CIS patients (CIS132, CIS429, ON4-10 and ON3-4 did not have CD138+ CSF plasma cell antibody repertoires). To clarify, antibody repertoires from CIS patients at UCD were generated from both single CD19+ CSF B Cells and single CD138+ CSF plasma cells (Bennett et al., 2008), while antibody repertoires from CIS patients at UTSWMC were generated from single CD19+ CSF B cells only. Since the resultant databases (CIS CD19+ CSF from UTSWMC, CIS CD19+ and CIS CD138+ CSF from UCD) were similar in mutational frequency, variable heavy chain (VH) gene family usage, and heavy chain joining segment (JH) usage, the two databases were combined for analysis (Supplemental Table 1).

2.3. Control B cell antibody database generation

The healthy control peripheral blood (HCPB) antibody database has been used in multiple studies (Brezinschek et al., 1997, 1998; Dorner et al., 1997, 1998a,b,c; Farner et al., 1999; Hansen et al., 2000; Harp et al., 2007; Monson et al., 2005, 2000) and consists of 348 CD19+ or CD19+IgM+ peripheral B cells from two healthy control donors. The memory HCPB antibody database (mHCPB) consists of 205 sequences from the HCPB antibody database that contain 4 or more mutations (less than 98% homology to the germline sequence, n = 123) combined with sequences from a HCPB antibody database generated from class-switched IgD CD27+ memory B cells (n = 82) (Tian et al., 2007) (Genbank 535266–535274, 535324–535368, 535381–535408, and 535416–535418). As expected, the class-switched IgD CD27+ memory B cell database had a higher percentage of mutated codons that resulted in a replacement than the mHCPB database (compare 64.7% vs 70.3%, p = 0.002 by χ² test). The OND CD19+ CSF antibody database consists of 65 sequences. UCD and UTSWMC cell isolation and IgM amplification was performed similarly. All sequences were re-confirmed by our
laboratory using IgBLAST (those obtained from UCD and from GenBank (http://www.ncbi.nlm.nih.gov/igblast/), and only codons 24–93 were considered in the analysis.

2.4. Mutation analyses

Frequency of replacement mutations (RF) was calculated as the number of replacement mutations at each codon position divided by the total number of replacement mutations in each VH4 subdatabase and displayed as a percentage. The MSCSF database contains 373 sequences with 2475 replacement mutations, and the CISCSF database contains 302 CD19+ and 226 CD138+ sequences with 4081 replacement mutations (2052 in CD19+ and 2029 in CD138+). The HCPB database contains 348 sequences with 1086 replacement mutations, and the mHCPB database contains 205 sequences with 1857 replacement mutations. The ONDCSF database contains 65 sequences with 482 replacement mutations, and the MSPB database contains 156 sequences with 1373 replacement mutations. In total, 1675 sequences and 10,373 replacement mutations were analyzed in this manner. Table 1 contains VH4 sequence numbers and Table 2 legend contains number of VH4 replacement mutations. Codon domains and numbers were defined by Kabat (Kabat et al., 1983), and Tomlinson in V-base (http://www.ncbi.nlm.nih.gov/igblast/), respectively.

2.5. Statistical strategy for signature identification

Codons included in the signature were identified using three criteria. First, we identified codons that had statistically different RF values in the MSCSFVH4 database compared to HCPBVH4 by Goodness of Fit test where the expected frequency is the RF calculated in HCPBVH4. Twenty-four codons passed this criterion. Next, codon positions that had an RF in both the MSCSFVH4 and HCPBVH4 databases that was less than the average +2 S.D. of the memory HCPBVH4 subdatabase were excluded. Thus, since the average ± S.D. RF of the memory HCPBVH4 database was 0.68 ± 0.59, any individual codon RF less than 1.86 in both databases was excluded. Fourteen codons passed this additional criterion. Eight of these 14 codons (31B, 32, 40, 56, 57, 60, 81, and 89) were defined as “hot” since the RF at that codon position within the MSCSFVH4 database was statistically higher compared to the HCPBVH4 database. Six of these 14 codons (30, 43, 52, 77, 82 and 82a) were defined as “cold” since the RF at that codon position within the MSCSFVH4 database was statistically less compared to the HCPBVH4 database. Two of the 6 “cold” codons (52 and 82a) were excluded because the RF value in the MSCSFVH4 database at that codon position was significantly higher than 1.86 (the average +2 S.D. of the memory HCPBVH4 subdatabase). The overall signature consistently consisted of codons 30, 31B, 32, 40, 43, 56, 57, 60, 77, 81, 82, and 89. This analysis was not biased by differences in the prevalence of particular codons (31B in particular), as individual VH4 gene frequencies in MSCSF were similar to HCPB by $\chi^2$ test using a Bonferroni corrected p-value of 0.004 (data not shown).

2.6. Statistical computation of the signature score

Signature scores were generated by calculating Z-scores for the RF values at the 6 codons within the signature (31B, 40, 56, 57, 81 and 89) that had the most significant difference in RF compared to HCPBVH4 at each codon position. The Z-score formula is: (RF at codon X minus the average RF in HCPBVH4) divided by the standard deviation of the average RF in HCPBVH4. For example, the average RF in HCPBVH4 within the 6 signature codons was 1.6 ± 0.9 and so an RF of 4.4 at codon 31B would be assigned a score of 3.1 ((Z-score = (4.4−1.6)/0.9)). Individual Z-scores at each of the 6 codon positions were then added to generate the composite signature Z-score. The average composite signature score in the MSCSFVH4 database was 10.9 ± 2.0 and so any signature score of an individual CIS patient above 6.8 (average −2 S.D.) was predicted to convert to CDMS. Of note, both the ONDCSFVH4 signature score (at 4.5), and the MSPB signature score (at 2.0) were below the threshold for MS conversion. CD19+ CSF B cell and CD138+ CSF plasma cell mutation positions both contributed to each CIS patient’s signature score, while the MSCSFVH4 signature scores were only composed of CD19+ CSF B cells.

2.7. VH4 structure

A human VH4–30.4 antibody structure described in (Guddat et al., 1993) was obtained from the Protein Data Bank (http://www.rcsb.org) under the identification moniker 1MCO, and adapted using the

Table 2

<table>
<thead>
<tr>
<th>Codon</th>
<th>Location</th>
<th>MSCSFVH4 RF</th>
<th>HCPBVH4 RF</th>
<th>Fold increase</th>
<th>p-value</th>
<th>mHCPBVH4 RF</th>
<th>Fold increase</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>31B</td>
<td>CDR1</td>
<td>3.5</td>
<td>0.5</td>
<td>7.0</td>
<td>0.001</td>
<td>0.8</td>
<td>4.4</td>
<td>0.001</td>
</tr>
<tr>
<td>32</td>
<td>CDR1</td>
<td>2.3</td>
<td>1.5</td>
<td>1.5</td>
<td>0.05</td>
<td>2.1</td>
<td>1.1</td>
<td>NS</td>
</tr>
<tr>
<td>40</td>
<td>FR2</td>
<td>2.7</td>
<td>1.0</td>
<td>2.7</td>
<td>0.001</td>
<td>1.1</td>
<td>2.5</td>
<td>0.001</td>
</tr>
<tr>
<td>56</td>
<td>CD2</td>
<td>5.5</td>
<td>3.0</td>
<td>1.8</td>
<td>0.001</td>
<td>3.2</td>
<td>1.7</td>
<td>0.001</td>
</tr>
<tr>
<td>57</td>
<td>CD2</td>
<td>2.0</td>
<td>1.0</td>
<td>2.0</td>
<td>0.005</td>
<td>0.5</td>
<td>3.7</td>
<td>0.001</td>
</tr>
<tr>
<td>60</td>
<td>CD2</td>
<td>2.4</td>
<td>1.5</td>
<td>1.6</td>
<td>0.05</td>
<td>3.7</td>
<td>1.3</td>
<td>NS</td>
</tr>
<tr>
<td>81</td>
<td>FR3</td>
<td>4.7</td>
<td>3.0</td>
<td>1.5</td>
<td>0.005</td>
<td>1.3</td>
<td>1.5</td>
<td>NS</td>
</tr>
<tr>
<td>89</td>
<td>FR3</td>
<td>2.0</td>
<td>1.0</td>
<td>2.0</td>
<td>0.005</td>
<td>1.3</td>
<td>1.5</td>
<td>NS</td>
</tr>
<tr>
<td>Hotspot total</td>
<td></td>
<td>25.0</td>
<td>12.6</td>
<td>2.0</td>
<td>0.001</td>
<td>13.8</td>
<td>1.8</td>
<td>0.001</td>
</tr>
<tr>
<td>30</td>
<td>FR1</td>
<td>2.0</td>
<td>4.0</td>
<td>0.5</td>
<td>0.005</td>
<td>2.9</td>
<td>0.7</td>
<td>NS</td>
</tr>
<tr>
<td>43</td>
<td>FR2</td>
<td>0.9</td>
<td>2.0</td>
<td>0.5</td>
<td>0.025</td>
<td>1.3</td>
<td>0.7</td>
<td>NS</td>
</tr>
<tr>
<td>77</td>
<td>FR3</td>
<td>1.5</td>
<td>2.5</td>
<td>0.6</td>
<td>0.05</td>
<td>1.6</td>
<td>0.9</td>
<td>NS</td>
</tr>
<tr>
<td>82</td>
<td>FR3</td>
<td>0.7</td>
<td>2.5</td>
<td>0.3</td>
<td>0.001</td>
<td>1.6</td>
<td>0.5</td>
<td>0.05</td>
</tr>
<tr>
<td>Coldspot total</td>
<td></td>
<td>5.1</td>
<td>8.5</td>
<td>0.5</td>
<td>0.001</td>
<td>7.4</td>
<td>0.7</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Abbreviations in table: CDR, complementary determining region; FR, framework; RF, replacement frequency; MSCSF, multiple sclerosis cerebrospinal fluid; HCPB, healthy control peripheral blood; mHCPB, memory HCPB; NS, not significant.

a As defined by Kabat (Kabat et al., 1983).

b Comparing HCPBVH4 or mHCPBVH4 to MSCSFVH4 RFs at each codon position using $\chi^2$ goodness-of-fit where expected frequency is the RF calculated in HCPBVH4 and mHCPBVH4 respectively.

c Codon used in the calculation of signature score.

d Previously published replacement hotspot (Dorner et al., 1997; Dorner et al., 1998a).

e “Hotspot total” is the total RF within codons 31B, 32, 40, 56, 57, 60, 81 and 89. “Coldspot total” is the total RF within codons 30, 43, 77 and 82. 199, 337 and 965 replacements respectively were included in this analysis for HCPBVH4, mHCPBVH4, and MSCSFVH4.
RasMol program (mc2.cchem.berkeley.edu/Rasmol/) to highlight codons within the designated signature of the heavy chain variable region.

3. Results

The 51 antibody heavy chain variable genes are subdivided into 7 different families (Cook and Tomlinson, 1995) (http://www.ncbi.nlm.nih.gov/igblast/), and it has been well-established that peripheral blood B cells from healthy donors utilize VH antibody genes most often from the VH3 family (“HCPB” in Table 1 and (Brezinschek et al., 1995; Brezinschek et al., 1997; Huang et al., 1992; Kraj et al., 1997; Wardemann et al., 2003; Yurasov et al., 2005)). In contrast, it has been reported by us and others that B cells in the CSF of MS patients often utilize VH4 antibody genes more frequently than those in the VH3 family (“MSCSF” in Table 1 and (Baranzini et al., 1999; Colombo et al., 2000; Harp et al., 2007; Monson et al., 2005; Owens et al., 1998, 2003, 2007; Qin et al., 1998; Ritchie et al., 2004)). The CISCSF antibody database consisting of CD19+ B cells only had a similar frequency of B cells that utilize VH4 family genes in comparison to HCPB (26.2% vs. 21.8%, \( p = 0.20 \) by \( \chi^2 \) test) (Table 1); in contrast, when CD138+ plasma cells were included, the CISCSF had a higher frequency of B cells that utilize VH4 family genes in comparison to HCPB (35.0% vs. 21.8%, \( p = 0.0001 \) by \( \chi^2 \) test). Some individual CIS patient CSF B cell antibody repertoires were enriched for B cells utilizing VH4 family genes in comparison to the random expected frequency, as reported previously (Bennett et al., 2008). CSF-derived B cell antibody repertoires from patients with Other Neurological Diseases (OND) were not enriched for VH4-expressing CSF B cells in comparison to HCPB (23.1% vs. 21.8%, \( p = 0.85 \) by \( \chi^2 \) test) or mHCPB (23.1% vs. 16.6%, \( p = 0.24 \) by \( \chi^2 \) test), indicating that VH4 over-expression in the CSF of MS patients was not due to bias in the ability of VH4 expressing B cells to enter the CNS.

3.1. Identification of codons within MSCSF that are enriched for replacement mutations

Since VH4 expressing B cells are enriched in the CSF of MS patients, we hypothesized that mutational analysis would reveal a pattern (i.e. “signature”) of antibody gene replacement mutations that is unique to VH4 expressing B cells from the CNS of MS patients in comparison to HCPB. In order to test this hypothesis, the percentage of replacement mutations (RF) at each codon within the VH4 subdatabase extracted from the parent database (MSCSFVH4) was determined and compared to the RF at each codon position within the VH4 subdatabase extracted from the parent HCPB database (HCPBVH4). Replacement frequencies were used so that only those mutations resulting in an amino acid change would be considered. In addition, codon amino acid replacement can result from 1, 2, or 3 nucleotide changes within the codon, and so replacement frequencies limit bias based on the number of nucleotides in a codon that are mutated to generate a replacement. Hot spots were defined as those codon positions within MSCSFVH4 with a statistically higher RF at a particular codon position in comparison to HCPBVH4 (Table 2). Using this approach, 8 codon positions (31B, 32, 40, 43, 56, 57, 60, and 89) were identified that have a total RF value in MSCSFVH4 (25.0%) that was statistically higher than in HCPBVH4 (12.6%) (\( p = 0.001 \) by \( \chi^2 \) test). Cold spots were defined as those codon positions within MSCSFVH4 with a statistically lower RF at a particular codon position in comparison to HCPBVH4 (Table 2). Four codons (30, 43, 77 and 82) were identified as cold spots that have a total RF value in MSCSFVH4 (5.1%) that was statistically less than in HCPBVH4 (8.5%) (\( p = 0.001 \) by \( \chi^2 \) test).

Individual MS patient RFs within the 8 hot spot codons of the signature ranged from 22.5 to 34.1% (data not shown), indicating that some individual patient MSCSF repertoires had a greater enrichment of replacements at these 8 codon positions than others. Also, the variability of RF values within the 8 hot spot codons of the signature in individual VH4 genes in MSCSFVH4 ranged from 14.5 to 36% (data not shown).
genes that contain codon 31B, there was a 3.1-fold increase in RF of the

When the analysis was restricted to those B cells expressing

oligoclonal bands; CDMS, clinically definite MS. Current criteria for diagnosis of MS requires dissemination of lesions both in time and space (Barkhof et al., 1997; Polman et al., 2005; Tintore et al., 2000). When MRI lesions alone are not sufficient to confirm diagnosis, CSF abnormalities can be used to meet the criteria of dissemination in space (Polman et al., 2005; Siritho and Freedman, 2009). Risk of conversion to clinically definite MS in patients who have had a single demyelinating event is 50–90% if the patient has an abnormal MRI (Beck et al., 2003; Brex et al., 2002; Cole et al., 1998; O’Riordan et al., 1998; Soderstrom et al., 1998), but of those patients with a normal MRI, up to 29% had oligoclonal bands and converted to CDMS (Cole et al., 1998).

In order to test whether signature prevalence could predict conversion to MS, we generated CSF B cell repertoires from patients who had one demyelinating event that placed them “at risk” to develop MS. Such patients are typically diagnosed with CIS. CD19+ B cell and CD138+ plasma cell repertoires from the CSF of two CIS patients at UT-SWMC and nine patients at UCHSC were generated and analyzed for RF values within the 6 codons of the signature defined in the MSCPBFVH4 database that had the most significant difference in RF compared to HCPBFVH4 at each codon position (codons 31B, 40, 56, 81 and 89). RF values were combined using a signature score that accounts for RF variance as described in Materials and methods. The average signature score in the MSCPBFVH4 database was 10.9±2.0 (range 7.6–11.9), and so any individual CIS patient score that was 6.8 (average signature score – 2 S.D.) or higher was predicted to develop MS (Fig. 2 and Table 3). Notably, the signature score from a pool of VH4 expressing CSF B cells of 3 OND patients was 4.5, and the signature score from a pool of VH4 expressing peripheral blood B cells of 3 CDMS patients was 2.0, and thus did not reach the 6.8 signature score threshold. Also, signature scores based on CD19+ B cell sequences only (in the patients where this was possible) did not change predictions based on signature score. This was expected since there is significant overlap in the antibody gene repertoires of CD19+

Table 3
CIS patient summary and signature score predictions.

<table>
<thead>
<tr>
<th>Subject no.</th>
<th>Time to LPb</th>
<th>MRI brain lesions</th>
<th>OCB</th>
<th>CD19/CD138c</th>
<th>CD19 VH4 biasd</th>
<th>CD138 VH4 biasd</th>
<th>Signature scoree</th>
<th>Prediction based on signature score</th>
<th>Definite MS</th>
<th>Time to MS diagnosisf</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS132</td>
<td>1</td>
<td>GD+</td>
<td>Yes</td>
<td>19/NA</td>
<td>No</td>
<td>NA</td>
<td>12.1</td>
<td>CDMS</td>
<td>Clinical</td>
<td>18</td>
</tr>
<tr>
<td>CS429</td>
<td>1</td>
<td>GD+</td>
<td>Yes</td>
<td>56/NA</td>
<td>No</td>
<td>NA</td>
<td>15.0</td>
<td>CDMS</td>
<td>Clinical</td>
<td>3</td>
</tr>
<tr>
<td>CS13</td>
<td>2</td>
<td>WML</td>
<td>Yes</td>
<td>24/76</td>
<td>Yes</td>
<td>Yes</td>
<td>15.5</td>
<td>CDMS</td>
<td>MRI</td>
<td>3</td>
</tr>
<tr>
<td>ON3-3</td>
<td>4</td>
<td>GD+</td>
<td>Yes</td>
<td>39/13</td>
<td>No</td>
<td>No</td>
<td>11.3</td>
<td>CDMS</td>
<td>Clinical</td>
<td>3</td>
</tr>
<tr>
<td>ON3-5</td>
<td>1.75</td>
<td>GD+</td>
<td>Yes</td>
<td>35/44</td>
<td>Yes</td>
<td>Yes</td>
<td>12.8</td>
<td>CDMS</td>
<td>Clinical</td>
<td>2</td>
</tr>
<tr>
<td>ON4-7</td>
<td>1</td>
<td>Noneb</td>
<td>Yes</td>
<td>17/20</td>
<td>No</td>
<td>Yes</td>
<td>10.2</td>
<td>CDMS</td>
<td>Clinical</td>
<td>5</td>
</tr>
<tr>
<td>ON4-8</td>
<td>1.5</td>
<td>WML</td>
<td>Yes</td>
<td>NA/18</td>
<td>NA</td>
<td>Yes</td>
<td>9.6</td>
<td>CDMS</td>
<td>Clinical</td>
<td>5</td>
</tr>
<tr>
<td>ON5-2</td>
<td>1</td>
<td>GD+</td>
<td>Yes</td>
<td>29/12</td>
<td>Yes</td>
<td>Yes</td>
<td>7.9</td>
<td>CDMS</td>
<td>Clinical</td>
<td>3</td>
</tr>
<tr>
<td>ON3-1</td>
<td>10</td>
<td>WML</td>
<td>Yes</td>
<td>23/45</td>
<td>No</td>
<td>No</td>
<td>6.4</td>
<td>–</td>
<td>NA</td>
<td>–</td>
</tr>
<tr>
<td>ON4-10</td>
<td>1.25</td>
<td>WML</td>
<td>No</td>
<td>NA/18</td>
<td>No</td>
<td>No</td>
<td>2.2</td>
<td>No</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ON3-4</td>
<td>1.5</td>
<td>Noneb</td>
<td>No</td>
<td>28/NA</td>
<td>No</td>
<td>NA</td>
<td>11.3</td>
<td>CDMS</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Abbreviations in table: CIS, clinically isolated syndrome; ON, optic neuritis; LP, lumbar puncture; GD+, gadolinium enhancing lesion positive; WML, white matter lesions by T2; OCB, oligoclonal bands; CDMS, clinically definite MS; NA, not applicable.

b Months from first demyelinating event to LP.

c Values given are number of unique sequences in CD19 repertoire/CD138 repertoire; family usage can be found in (Bennett et al., 2008; Harp et al., 2007); CS429 was a 62 y.o. male first presenting with optic neuritis and the repertoire had 2% VH1 usage, 0% VH2, 77% VH3, and 19% VH4.

d Bias was considered significantly different from random frequency (Cook and Tomlinson, 1995) or expected frequency in HCPB (Brezinschek et al., 1997, 1998; Dorner et al., 1997, 1998a,b; Farner et al., 1999; Hansen et al., 2000; Monson et al., 2000). VH2 bias was also observed in CD19 repertoires from ON4-7, and CD138 repertoires from ON3-3, ON3-5 and ON4-7.

e Signature score was calculated as outlined in Materials and methods, and uses both CD19 and CD138 sequences. The average score among MS patients is 10.9±2.0.

f Months from first demyelinating event to MS diagnosis.

One spinal cord lesion was observed by T2 weighted MRI.

Fig. 2. Signature score in individual MS and CIS patients. Signature scores were generated by calculating Z-scores for the RF values at the 6 codons within the signature (31B, 40, 56, 57, 81 and 89). Individual Z-scores at each of the codon positions were compiled to generate the composite signature Z-score. MS patient signature scores are shown as black circles (●), CIS patient signature scores that resulted in prediction of CDMS are black squares (■), and CIS patient signature scores that resulted in prediction of unlikely to convert to definite MS are in open squares (□). The average composite signature score in the MSCPBFVH4 database was 10.9±2.0 (black line) and so any signature score of an individual CIS patient above 6.8 (average – 2 S.D.) was predicted to convert to CDMS. For reference, ONDCSFVH4 group signature score was 4.5, and MSPBHSV signature score was 2.0.
B cells and CD138^+ plasma cells from the CSF of the same patients (Martin Mdel and Monson, 2007; Ritchie et al., 2004), suggesting that the memory B cell pool present in the CSF is the reservoir for differentiation of plasma cells in the CSF.

As indicated in Table 3, prediction of conversion to CDMS using the antibody gene signature score was accurate in 8 of 8 CIS patients that converted to CDMS. Lack of signature prevalence also accurately predicted that 2 of 2 patients who had recently experienced a first demyelinating event (ON3-1 and ON4-10) would not develop CDMS, and indeed, have not developed CDMS up to 2 years after initial sampling. One additional patient who had recently experienced a first demyelinating event (ON3-4) had a high signature score (11.3), but had not converted to CDMS at the 2-year follow-up. The antibody gene signature yielded a sensitivity of 100%, specificity of 67%, positive predictive value of 89%, negative predictive value of 100%, and accuracy of 91%, as defined by others applying the McDonald Criteria to identify CIS patients that would convert to MS (Dalton et al., 2002). Most patients in this cohort converted to CDMS within 3–6 months of repertoire sampling, although in the case of CIS132, conversion to CDMS was not confirmed until 17 months after antibody repertoire sampling (Table 3). MRI, OCB and VH4/VH2 bias are also useful in assessing probability of MS conversion (Bennett et al., 2008; Freedman et al., 2005; Frohman et al., 2003; Korteweg et al., 2006; Paolino et al., 1996; Soderstrom et al., 1998), but were not considered in calculating the signature score.

4. Discussion

The intense somatic hypermutation accumulation in MSCSF514 enabled us to identify a unique antibody gene signature enriched for replacement accumulation at codons 31B, 32, 40, 56, 57, 60, 81 and 89 that was not observed in HCPB514 or ONDCSF514. Of note, any residual effect of naive B cells on the RF calculation was minimized by tabulating only those sequences with mutations resulting in amino acid replacements. This approach minimized bias in the signature that may have reflected enrichment of mutation accumulation in CNS derived B cells (which are mostly memory and thus have high mutation rates) compared to peripheral B cells (which are mostly naive and thus have low mutation rates). In addition, 5 of the 8 hot codons of the signature retain higher RF values in the MSCSF database compared to the memory HCPB database. Finally, if signature score reflected enrichment of mutation accumulation due to the repertoire’s high memory representation, then all signature scores from the CISCSF antibody repertoires should have been high since they were all heavily enriched for memory B cells. This was not the case, since CISCSF repertoires ON3-1 and ON4-10, despite being heavily enriched for memory B cells (with mutation frequencies of 5.2% and 6.7%, respectively), had signature scores below the threshold of 6.8 (ON3-1 score = 6.4, ON4-10 score = 2.2).

It was compelling to investigate whether the antibody gene signature may be of value to identify CIS patients who would subsequently develop MS, since early and accurate diagnosis of MS is of tantamount importance in clinical care (Stuve et al., 2008). Signature prevalence could be used to identify patients who would be diagnosed with CDMS within 3–18 months of experiencing their first demyelinating event. Of note, patient ON3-4 had a signature score that indicated this patient would convert to CDMS (score = 11.3), but did not demonstrate a lesion load by MRI, banding by OCB, or VH4/VH2 bias, and had not developed CDMS up to 2 years after sampling was performed (Table 3). It will be interesting to determine whether this patient is diagnosed with CDMS over time. It is also important to note that the majority of patients in this cohort already had evidence of MS risk as indicated by positive MRI and OCB. Patient ON4-7, however, did not present with brain lesions by MRI, but had a signature score that indicated this patient would be diagnosed with CDMS (score = 10.2). Indeed, this patient did convert to CDMS within 5 months of CSF B cell antibody repertoire sampling, and provides a reasonable example of how signature prevalence may predict CDMS diagnosis in patients that either do not present with brain lesions by MRI, or who are not evaluated by MRI at this stage of their disease. It will be interesting to determine whether the combination of MRI and signature prevalence would be useful in predicting MS conversion. Signature prevalence may also provide an evaluation mechanism to identify the most appropriate patient candidates to receive B cell depletion therapies, for example. Certainly this is a priority since a recent investigation demonstrated significant efficacy of Rituxan in RRMS patients (Hauser et al., 2008).

Given the urgency for the early identification of MS and the rapid initiation of disease modifying therapy, presentation of a molecular signature in the CSF B cells of CIS patients who develop MS may provide a unique tool for identifying at risk individuals. However, wide implementation of the current form of this approach would be problematic since the AGS scores presented here were generated using a specialized technique that is labor intensive (single cell PCR) and requires fresh CSF for sampling. Developing other approaches to generate AGS data that maintains accuracy, but does not require a specialized laboratory to perform, is attainable and of paramount importance.

In addition, many early MS patients have atypical clinical presentations or unremarkable MRI scans, and patients with alternative inflammatory conditions may mimic idiopathic demyelinating disease. In these circumstances, the advent of a molecular diagnostic signature would increase diagnostic sensitivity and specificity. Investigating the utility of the antibody gene signature in such patients is ongoing in our laboratory.

The presence of a mutational signature among clonally expanded VH4 germline antibodies in MSCSF may be helpful in understanding disease pathogenesis. For example, the VH4 germline mutational signature may be the direct result of antigen targeting in the humoral immune response. Therefore, determining the antigen specificity of signature-enriched antibodies from CSF B cells of patients with definite MS and CIS is one of the first steps towards dissecting whether signature-enriched B cells have the potential to participate in MS pathogenesis. Of note, 5 of the 8 signature codons (31B, 32, 56, 57 and 60) we identified as having a unique accumulation of amino acid replacements in MSCSF514 are predicted to have direct antigen contact since they reside in CDRs (Supplemental Fig. 1). Dissecting the relative contribution of replacement mutations at each of these signature codons as well as those outside of the antigen binding region will address the impact of both codon classifications (direct and indirect antigen binding capacity) on antigen binding affinity.

In summary, a unique signature of antibody gene replacement mutations was identified in the MSCSF514 database that is not observed in healthy control peripheral blood or CSF-derived VH4-expressing B cells from patients with other neurological diseases. Prevalence of the signature was accurate in identifying CIS patients that would convert to CDMS, but needs to be tested on a larger cohort of patients at both high and low risk to develop MS. Identifying the antigen specificity of signature-enriched CSF B cells from these patients may also reveal a unique group of antigens that are central to initiation of humoral autoimmunity in the CNS. It is likely that the MS-specific VH4 antibody gene signature provides both a new focus of investigation to further elucidate the role of B cells and their antibody products in MS and a new candidate molecular diagnostic tool for MS.

Acknowledgements

The authors thank the patients who consented to sampling for this study. Special thanks to Dr. Olaf Stüve for critical review of this manuscript, Bonnie Darnell, Angie Mobley and Elizabeth Curry for their technical expertise in cell sorting, and Tanya Hendricks, Elayna Tillman and Timothy Ahearne for repertoire analysis support.


