Comprehensive phenotyping in multiple sclerosis: Discovery based proteomics and the current understanding of putative biomarkers

Kevin C. O’Connora, Sushmita Mimi Royb, Christopher H. Beckerb, David A. Haflera and Aaron B. Kantorb

aDepartment of Neurology and the Center for Neurologic Disease, Brigham and Women’s Hospital, Laboratory of Molecular Immunology, Department of Neurology, Harvard Medical School 77 Avenue Louis Pasteur, Boston, MA 02115, USA
bBiomarker Discovery Sciences, PDD, 1505 O’Brien Drive, Menlo Park, CA 94025, USA

Abstract. Currently, there is no single test for multiple sclerosis (MS). Diagnosis is confirmed through clinical evaluation, abnormalities revealed by magnetic resonance imaging (MRI), and analysis of cerebrospinal fluid (CSF) chemistry. The early and accurate diagnosis of the disease, monitoring of progression, and gauging of therapeutic intervention are important but elusive elements of patient care. Moreover, a deeper understanding of the disease pathology is needed, including discovery of accurate biomarkers for MS. Herein we review putative biomarkers of MS relating to neurodegeneration and contributions to neuropathology, with particular focus on autoimmunity. In addition, novel assessments of biomarkers not driven by hypotheses are discussed, featuring our application of advanced proteomics and metabolomics for comprehensive phenotyping of CSF and blood. This strategy allows comparison of component expression levels in CSF and serum between MS and control groups. Examination of these preliminary data suggests that several CSF proteins in MS are differentially expressed, and thus, represent putative biomarkers deserving of further evaluation.

1. Introduction

This article will briefly review our current knowledge of biomarkers in multiple sclerosis (MS), with particular focus on neuroimmunology, and especially antibodies. It is not intended as an exhaustive review, as such work has recently been completed in several excellent and thorough publications [1–3]. The later part of the manuscript will focus on the appraisal of groundwork biomarker discovery in MS using our own novel strategy, which allows detection, identification, and quantification of a large number of putative biomarkers in MS, leading to a “molecular signature” of proteomic patterns free of pre facto bias.

2. Multiple sclerosis

MS, the most common autoimmune disease involving the nervous system (CNS), is a chronic illness affecting CNS pathways and leading to progressive neurological dysfunction [4,5]. The disease affects twice as many women as it does men. Approximately 400,000 Americans acknowledge having MS, and every week about 200 people are diagnosed. Worldwide, MS may affect 2.5 million individuals. The high economic burden is reflected in the total estimated annual cost for all people with MS in the United States: more than $9 billion dollars [6]. Both genetic and environmental factors contribute to MS susceptibility. High-risk re-
regions include the northern United States and the Scandinavian countries, while Japan and Africa are areas of lower risk. Individuals that migrate from high-risk to low-risk regions carry their native risk for contracting MS, suggesting that exposure to an environmental factor contributes to susceptibility. Genetic studies have shown that the risk of developing MS is elevated 10- to 20-fold in first-degree relatives of individuals with MS and that the concordance rate among monozygotic twins is 30%–35%, but only 2%–5% in dizygotic twins.

Clinically, the disease can be broadly divided into a relapsing-remitting form characterized by a series of exacerbations that result in varying degrees of disability from which the patient recovers, and two progressive forms (secondary progressive, SP and primary progressive, PP) in which the patient does not experience exacerbations, but instead experiences a gradual decline. Though a relapsing-remitting onset is observed in 85–90% of all patients, the course of the disease in about 40% of them ultimately becomes progressive [5]. Thus, the current prevalence of the subtypes is approximately 45% RR, 35% SP, and 20% PP. The pathological hallmark of MS is the demyelinated plaque found throughout the brain and spinal cord, with CNS perivascular infiltration of inflammatory cells, demyelination, astrogliosis, and axonal injury (reviewed in references [7–10]). Symptoms are believed to result primarily from axonal demyelination that inhibits or blocks conduction. Amelioration of symptoms has been attributed to both partial remyelination and resolution of inflammation.

It is widely accepted that MS is a complex syndrome and may not represent a single disease entity. Based on accumulating data from immunological studies of patients with MS and a wealth of animal model data, autoimmune dysregulation has been viewed as the major contributor to tissue damage (reviewed in references [9–11]). There is widespread support for the view that MS is an immune-mediated disease, based largely on: 1) the genetic linkage to immune-related molecules, primarily MHC class II; 2) the presence of CNS inflammatory infiltrates including CD4+ T cells, CD8+ T cells, B cells, plasma cells, macrophages; 3) magnetic resonance imaging (MRI) indicative of the loss of integrity of the blood – brain barrier (BBB); 4) the response to immune-modulating therapies; 5) similarities to the animal model of experimental autoimmune encephalomyelitis (EAE); and 6) the presence of oligoclonal bands and elevated immunoglobulin levels in the cerebrospinal fluid (CSF).

In accordance with this evidence, a widely accepted model of MS immunopathology [12] suggests that activated autoreactive T cells recognize myelin antigens in the periphery, cross the CNS endothelium, leading to a cascade of events that culminates in white matter inflammation and tissue destruction. Inside the CNS, release of local cytokines, chemokines, and matrix metallo-proteins support the recruitment of subsequent waves of infiltrating effector cells (including monocytes and B cells). Mechanisms of myelin destruction and axonal damage are likely to be multiple and include direct effects of pro-inflammatory cytokines, complement-fixing antibodies, antigen-specific and non-specific cytotoxicity, and apoptosis.

While it is broadly appreciated that a sizeable percentage of MS cases involve immune-mediated pathogenesis, neurodegenerative processes not necessarily associated with inflammation appear to be principal in others. Theories that overlook inflammation as a primary event hold that MS is at least partially a genetically determined disorder characterized by metabolic neurodegeneration. Studies of the mechanisms of axonal damage and neurodegeneration in MS are still in their infancy, though recognition of axonal damage as a prominent pathological feature in MS lesions is increasing (reviewed in references [7,13–15]).

It is likely that both neurodegeneration and inflammation contribute to various extents, both among and within individuals with MS. In a seminal work underscoring the significance of both inflammatory and non-inflammatory mechanisms in MS pathology, Lassman and colleagues [16] catalogued lesions and demonstrated the existence of four fundamentally different patterns of demyelination in humans. While two patterns (I and II) showed close similarities to T-cell-mediated or T-cell plus antibody-mediated autoimmune encephalomyelitis, the other patterns (III and IV) were highly suggestive of a primary oligodendrocyte dystrophy, reminiscent of demyelination induced by a virus or toxin rather than one involving the immune system.

3. Biomarkers

Biomarkers are anatomic, physiologic, biochemical, or molecular parameters associated with the presence and severity of specific disease states. Biomarkers are measurable by a variety of methods including physical examination, laboratory assays, and medical imaging. The hallmark of a useful biomarker is that the state of health or disease of an individual can be accurately characterized by appraisal of the marker [17]. Thus, biomarkers can indicate either nor-
mal biological/pathogenic processes or pharmacological responses to a therapeutic intervention. Simple examples include CD4+ T-cell count in HIV-infected patients and HDL/LDL ratios for patients with hypercholesterolemia. Some biomarkers are not specific; prostate-specific antigen (PSA) may be increased not only in men who have prostate cancer, but also in those with benign prostatic hyperplasia or a prostate infection. More pertinent to the present report, white matter abnormalities, which can be detected by MRI, serve as biomarkers for MS.

4. Biomarkers in MS inflammation

The current inventory of prospective MS biomarkers by and large includes those that reflect the changeable inflammatory processes in the majority of cases. Many such biomarkers are found in the periphery, namely in the serum. However, analysis of CSF can add greater pathologic specificity, as it is most closely associated with the disease process. Thus, considerable efforts have been made to identify prognostic and diagnostic markers in the CSF from patients with MS.

The biomarkers reflecting the role of inflammation include cytokines, chemokines, adhesion molecules, and each of their respective receptors, complement, and antibodies. Specifically, the pro-inflammatory cytokines associated with T cells; interferon (IFN)-γ, tumor necrosis factor (TGF)-α and IL-12 may be expressed at elevated levels. Correspondingly, the anti-inflammatory cytokines such as IL-4, IL-10, and TGF-β may be overexpressed during periods of remission. In each of these cases, more evidence is required.

Breakdown of the blood-brain barrier (BBB) is necessary in some cases before immune effector cells can extravasate into the CNS. In other cases, CSF cells can extravasate through an intact BBB. Disruption of the BBB involves proteolytic activity of enzymes that degrade the extracellular matrix (ECM), chief among which are the matrix metalloproteinases (MMPs). Accordingly, elevated expression levels of MMP family members in patients with MS have been reported.

In addition to larger proteins, small-molecule biomarkers have also been investigated. For instance, it is appreciated that nitric oxide (NO) is formed in inflammatory disorders. Consequently, levels of NO, its oxidation products (NOx), and the iNOS enzyme are altered. Several studies have examined particular parts of this response. CSF nitrite levels appear to be correlated with disease exacerbation and concurrent inflammation in the CNS of patients with MS [18].

It is well-accepted that autoantibodies toward insulin and DNA are established biomarkers in the autoimmune diseases type 1 diabetes (T1D) and systemic lupus erythematosus, respectively. However, these autoantibodies have yet to be implicated directly in disease pathogenesis. In contrast, myasthenia gravis autoantibodies directed toward components of the neuromuscular junction are both diagnostic biomarkers and participants in disease pathology [19]. Neuromyelitis optica, (NMO) also termed Devic’s syndrome, is an inflammatory demyelinating disease that affects the optic nerves and spinal cord, but unlike MS, does not affect the brain. Early in its course, Devic’s is difficult to distinguish from MS, yet the distinction remains important, as treatment and prognosis of these two diseases are dissimilar. The presence of IgG serves as a specific serologic biomarker for early diagnosis of NMO and distinction from MS [20]. The NMO serum IgG reveals a characteristic pattern upon immunohistochemical staining of mouse CNS tissue. Similarly, Vincent and colleagues [21] found that the serum Ig of patients with MS recognized antigens on the surface of CNS-derived cells, whereas serum Ig from controls did not. In this case, as with NMO, the appropriate next step in biomarker refinement would be identification of the target(s).

As autoantibodies have been shown to be useful markers in some autoimmune diseases, these molecules have also been examined for their suitability as biomarkers in MS. Other approaches have focused on individual candidate antigens. Consequently, there is an abundance of reports concerning autoantibodies in MS. Although much of this work has focused on the role of autoantibodies in the immunopathogenesis of MS, most are still only putative biomarkers, as a direct role in pathology has not been reported. That said, no study has clearly demonstrated that autoantibodies in MS are either reliable biomarkers or contributors to the disease pathology. However, when considered collectively, the reports on EAE and MS certainly indicate that suspicion of such a role for autoantibodies is likely well-founded. The evidence for a consequential humoral response within patients with MS includes the presence of CSF oligoclonal bands [22] and intrathecal immunoglobulin synthesis [23], both hallmarks of MS and diagnostic biomarkers used in conjunction with clinical presentation and MRI. This combination of several biomarkers is necessary, because CSF oligoclonal bands and evidence of intrathecal in-
munoglobulin synthesis are found in conditions other than MS, such as subacute sclerosing panencephalitis (SSPE). Again, this underscores the dictum that no single biomarker is likely to be conclusive, but rather an arrangement of at least several biomarkers, forming a profile or fingerprint, will be needed to provide valuable information.

The possibility that CSF Ig in patients with MS is generated as a response to myelin self-antigens has also been investigated in detail. Antibodies specific for MBP have been identified in the CSF [24–26]. Autoantibodies reactive with myelin-associated glycoprotein (MAG) [27], the enzyme transaldolase (TAL) [28], MOG [29], and oligodendrocyte-specific protein (OSP) [30] have also been detected in the CSF of a subset of patients with MS. One series of reports [31] demonstrated that individuals with MS who have intrathecal synthesis of IgM were likely to progress from RRMS to a more severe progressive form. The presence of intrathecal IgM correlates with progression from initial stages to clinically definite MS [32]; higher EDSS scores [33] and a fraction of this IgM reactive with myelin lipids suggests a more aggressive course [34].

Serum anti-MBP antibodies have been reported [35, 36], though not in all studies [37–39]. As with MBP, reports are conflicting regarding the detection of anti-MOG antibodies in MS serum [29,40]. There are reports detecting serum autoantibodies against recombinant TAL [28], and B cells secreting autoantibodies directed against PLP have been identified in the peripheral blood of patients with MS [41]. Autoantibodies to other putative MS antigens, such as MOBP, CNPase, oβ-crystalline, and S100B have not been analyzed in sera [42].

Autoantibodies directed against non-myelin antigens have also been detected in the sera of some patients with MS. Anti-nuclear antibodies (ANA) [43], anti-cardiolipin antibodies [44], and antibodies against beta-2-glycoprotein I [45] are detected in MS sera more frequently than in normal subjects. Recently, elevated levels of antibodies towards a large panel of organ- and non-organ-specific antigens were reported in patients with MS relative to controls [46]. Currently, no B-cell antigen specific for MS or identifying a subtype of MS has been established as a true biomarker for MS.

Our own studies have indicated that the antibodies to chief candidate myelin antigens, MBP and MOG, in the serum and CSF of patients with MS, are not readily detectable [47,48] and therefore fail as useful biomarkers. Our work has shown that patients with acute disseminated encephalomyelitis (ADEM), an induced encephalomyelitis, have these autoantibodies. Moreover, reports concerning their detection in ADEM are consistent [47,49], which cannot be said of those concerning MS. Our view is that the inconsistency in the reported data of autoantibodies in MS is caused by false positives generated by acute-phase reactants that result in “sticky” antibodies. Alternatively, the difficulty in consistently detecting autoantibodies in the periphery of patients with MS is that they may be produced locally within the CNS at the site of injury. The evidence for a significant humoral response within the CNS of patients with MS includes the presence of plasma cells and IgG in CNS lesions [16,50]. That an antigen-driven response is at work is indicated in analysis of the immunoglobulin VH and VL domains from CNS samples, which reveals that these gene segments have accumulated replacement mutations in the complementary determining regions (CDRs) [51] and a skewed variable region repertoire [51,52]. These findings strongly suggest an antigen-driven response, rather than polyclonal B cell activation [53–55]. Moreover, autoantibodies to MOG were found to be bound to disintegrating myelin segments in MS lesions [56] and in IgG eluates from MS lesions [48], supporting the contention that myelin autoantibodies may mediate CNS tissue damage. IgG isolated from CNS tissue binds to MBP in solid-phase assays [57]. Collectively, these results support the hypothesis that pathologically significant antibodies are produced within the brain and may constitute specific biomarkers. Measurement of these potentially useful biomarkers is clearly limited by our current technology. Application of novel non-invasive technology or more sensitive assays for measurement in the periphery may provide access to these prospective biomarkers.

5. Biomarkers in MS neurodegeneration

The neuropathology of MS largely concerns the CNS white matter, although lesions may sometimes be found in the gray matter as well. The key cellular components of the white matter are the myelin-coated axons responsible for conducting nerve impulses. The remaining fundamental CNS support cells are categorized as glial cells and include: oligodendrocytes, which provide and maintain the axonal myelin sheaths; astrocytes, which provide mechanical and metabolic support; and microglia, which perform phagocytosis, essentially becoming activated to clean up after cell damage or death.
The gray matter is largely comprised of neurons that collect, integrate, and transmit information.

Certain traditional concepts regarding MS pathogenesis have recently come into question. For instance, it is now evident that the types of cells and molecules involved in MS are more diverse than previously thought. In particular, some CNS lesions can include either extensive oligodendrocyte apoptosis, microglial activation, axonal damage, or all of these features, while containing few or no lymphocytes [58]. While these observations do not preclude the role of inflammatory demyelination in MS pathogenesis, this form of CNS tissue damage may predate the inflammatory lesions. Thus, independent axonal pathology may contribute to the primary pathobiology of the disease.

Thus, there has been a renewed interest in the neurodegenerative mechanisms of MS. The most reported tissue injury is oligodendrocyte apoptosis and axonal degeneration and transection. Thus, irreversible neurologic disability in MS is related, in part, to axonal damage. All these pathological findings can be detected on MRI as atrophy and hypointense T1-weighted lesions, but pathogenic specificity is lacking.

Evaluation of glial cell responses in CSF and lesion tissue [59] revealed that S100B is a worthy marker for the relapsing phase of the disease and can be used along with ferritin to discriminate among the types of MS. Glial-fibrillary acidic protein (GFAP) correlated with disability scales. In another bona fide demonstration of the importance of combinations of biomarkers, the S100B: ferritin ratio discriminated patients with RR multiple sclerosis from SP, PP, or control patients and may therefore provide a marker for irreversible damage.

Markers of tissue destruction are useful for gauging neurodegenerative processes in MS. Isoprostanes, formed by free radical peroxidation, are emerging as a new class of sensitive, specific, and reliable markers of free radical damage and lipid peroxidation. CSF levels of isoprostane are elevated in MS patients [60]; however, in this context the elevation is related to degenerative rather than inflammatory phenomena [61]. Magnetic resonance spectroscopy can be used to detect levels of whole brain N-acetylaspartate (NAA); the decreases seen in MS suggest progressive neuronal cell loss. Though whole-brain NAA measurement offers a quick, highly reproducible measure of disease progression, its use is dependent upon access to MRI instrumentation.

Some putative markers that deserve further investigation and validation include cytoskeleton structural proteins like actin, tubulin, L-neurofilaments and tau protein, axon-specific antibodies, other neuronal or glial proteins like S-100, 14-3-3 and glial fibrillary acidic protein, neuronal specific enolase, and nitric oxide and its metabolites (reviewed in reference [62]). However, none currently fulfills the criterion of applicability in clinical practice.

Transected or injured axons release several molecules from their proximal extremity into the intercellular space. Both the neurofilaments contained in the cytoplasm of neurons and the autoantibodies directed toward them have been viewed as potential biomarkers in MS and other neurological disorders [63]. Investigators have searched for correlations between levels of these markers and clinical or MRI measures with success, albeit limited.

### 6. Biomarkers in MS: Proteomics, the next generation technology

Despite tremendous efforts in MS research, the specific targets of the immune response, the precise mechanism for neuronal loss, and the events leading to disease etiology are not yet clear. Better biological markers that represent the biological activity of the disease process and response to therapy are needed. Although numerous putative MS-specific biomarkers, representing different mechanisms of pathogenesis and steps along the inflammatory cascade have been proposed., none has been fully validated. As one can deduce from the descriptions of biomarkers above, most of this work has been limited to investigation of one to several markers at a time. Although many advances have been made, we are still faced with exploring complex systems with studies designed to illuminate only one or a few of these multi-level biological problems. Comprehensive proteomic phenotyping fills the unmet need for rapid and methodical examination of molecules not only in the immune system but also from the targeted organs. To comprehend a disease as enigmatic and heterogeneous as MS, a high throughput technology capable of profiling multiple changes is needed.

We have taken a high-throughput multiplexed proteomics and metabolomics approach to biomarker identification, using both discovery-based and hypothesis-based strategies [64–66]. Biological samples are analyzed by mass spectrometry because of its high sensitivity, information content, and throughput. Our approach does not require any a priori assumptions about the specific differences between sample groups. Moreover, recently developed software and bioinformatics
proteins. This approach enhances reproducibility and
identification. The digested sample is then applied to
on-line reverse-phase liquid chromatography, directly
coupled to electrospray ionization mass spectrometry
with time-of-flight (TOF) detection. TOF detection is
ideal for distinguishing individual components in com-
plex mixtures, since it provides high resolution and
high mass accuracy, to an optimum of 2 mDa.

An important technical advance for the mass spec-
trometry platform was the development of software to
pick peaks and make comparisons across a large num-
ber of different samples [67,71,72]. The data process-
ing includes 1) removing and smoothing of baseline
noise and spikes from the raw data, 2) dynamic thresh-
olding to select peaks in different regions of the spec-
trum, 3) considering both the m/z and chromatographic
dimensions to eliminate chemical noise and identify
peaks pertaining to true analyte molecular ions, 4) ad-
justing for small deviations in the elution time with
dynamic time warping algorithms, and 5) comparing
each file to a reference file to generate a single constant
based on the median of the ratios of intensities for all
components between the reference and test files and
normalizing the complete data set. De-isotoped peaks
at a certain frequency (e.g. one out of four samples in
a two-group cross-sectional comparison) are retained
in a merged peak list for statistical evaluation. Data
on system performance, including reproducibility and
linearity of the response, can be found in Wang et al.
and Roy et al. [67,68]. One-dimensional chromatog-
raphy with a hydrophobic reverse-phase column cur-
cently tracks more than 45,000 molecular components
in the serum proteome. More molecular components
can be monitored if the sample is separated using two
orthogonal chromatography methods. As outlined in
Fig. 1 on the right, we use strong cation exchange,
with collection of up to eight fractions, followed by the
same reverse-phase chromatography used in the one-
dimensional method. This enables tracking of about
30,000 molecular species [69].

Separating the differential profiling steps from pep-
tide identification is an extremely useful strategy for im-
proving throughput and reproducibility. Peptide iden-
tification is made using established tandem mass spec-
trometry approaches (MS/MS). Collisional activation
for tandem mass spectrometry provides fragments (usu-
ally broken at the peptide bond) that can be assigned
using a match to large public databases via the SE-
QUEST and MASCOT computer programs [73]. For
new studies, links are first made to our library of iden-
tified peptides based on retention time and mass. Addi-
tional significant differences between groups (such as

7. Biomarkers in MS: Proteomics, results

Our method for differential profiling of proteins and
low molecular weight molecules from biological sam-
ple has three main technical components: 1) sam-
ple preparation, 2) liquid chromatography-mass spec-
trometry instrumentation, and 3) data processing infor-
matics [65,67,69]. The overall separation and analy-
sis scheme is shown in Fig. 1. Initial sample prepa-
ration, prior to chromatography, varies with the bio-
ological sample, such as serum, CSF, urine, or cells;
whereas the instrumentation and software generally apply
across sample types. The sample is divided into a
fraction with high molecular weight (proteome) and
another with low molecular weight (metabolome, in-
cluding free-floating peptides), since these require dif-
ferent methods of preparation, mass spectrometry and
identification. Here we focus on the proteome. High-
abundance proteins, e.g., albumin and immunoglobu-
lin, which account for about 90% of the serum pro-
teins, are removed to enable better monitoring of lower-
abundance proteins. The proteome is denatured, re-
duced, alkylated, and trypsin-digested to generate pep-
tides, which are easier to resolve than large intact
proteins. This approach enhances reproducibility and
identification. The digested sample is then applied to
on-line reverse-phase liquid chromatography, directly
coupled to electrospray ionization mass spectrometry
with time-of-flight (TOF) detection. TOF detection is
ideal for distinguishing individual components in com-
plex mixtures, since it provides high resolution and
high mass accuracy, to an optimum of 2 mDa.

An important technical advance for the mass spec-
trometry platform was the development of software to
pick peaks and make comparisons across a large num-
ber of different samples [67,71,72]. The data process-
ing includes 1) removing and smoothing of baseline
noise and spikes from the raw data, 2) dynamic thresh-
olding to select peaks in different regions of the spec-
trum, 3) considering both the m/z and chromatographic
dimensions to eliminate chemical noise and identify
peaks pertaining to true analyte molecular ions, 4) ad-
justing for small deviations in the elution time with
dynamic time warping algorithms, and 5) comparing
each file to a reference file to generate a single constant
based on the median of the ratios of intensities for all
components between the reference and test files and
normalizing the complete data set. De-isotoped peaks
at a certain frequency (e.g. one out of four samples in
a two-group cross-sectional comparison) are retained
in a merged peak list for statistical evaluation. Data
on system performance, including reproducibility and
linearity of the response, can be found in Wang et al.
and Roy et al. [67,68]. One-dimensional chromatog-
raphy with a hydrophobic reverse-phase column cur-
cently tracks more than 45,000 molecular components
in the serum proteome. More molecular components
can be monitored if the sample is separated using two
orthogonal chromatography methods. As outlined in
Fig. 1 on the right, we use strong cation exchange,
with collection of up to eight fractions, followed by the
same reverse-phase chromatography used in the one-
dimensional method. This enables tracking of about
30,000 molecular species [69].

Separating the differential profiling steps from pep-
tide identification is an extremely useful strategy for im-
proving throughput and reproducibility. Peptide iden-
tification is made using established tandem mass spec-
trometry approaches (MS/MS). Collisional activation
for tandem mass spectrometry provides fragments (usu-
ally broken at the peptide bond) that can be assigned
using a match to large public databases via the SE-
QUEST and MASCOT computer programs [73]. For
new studies, links are first made to our library of iden-
tified peptides based on retention time and mass. Addi-
tional significant differences between groups (such as
cases of MS and healthy subjects) are considered based on p-values and magnitudes of differences and then identified with directed tandem mass spectrometry.

To address our first goal of serum profiling, we collected serum from 20 patients with relapsing-remitting MS and 20 normal controls. A diagnosis of MS was confirmed based on the criteria of McDonald [74]. Serum was fractioned into the metabolome and proteome and processed as described above with one-dimensional chromatographic separation. We also extended the analysis to the examination of CSF. In MS, it is expected that CSF more closely reflects events in the CNS than does peripheral blood and hence is the preferred compartment for initiation of biomarker discovery efforts, despite being more difficult to obtain. Although most proteins (\(\sim 80\%\)) in the normal CSF originate from the blood, they are generally reduced 100- to 1,000-fold. Blood proteins passively diffuse across capillary walls into the brain, extracellular fluid, and CSF. Larger molecules exchange more slowly and have a larger concentration gradient from serum to CSF than smaller molecules. Some relative concentrations of CSF to serum are IgM 0.0003, IgG 0.002, and albumin 0.005) [75]. Brain-derived proteins fall into three categories: 1) Proteins derived from neurons and glial cells like tau, S-100, and neuron-specific enolase, which are released into the ventricular and cisternal CSF, where they have concentrations of 10-, 18-, and 1-fold higher than found in serum, respectively. 2) Proteins derived primarily from the leptomeninges surrounding the brain and spinal cord that are released into the CSF, like beta trace protein (prostaglandin-D-synthase) and cystatin C, which are found at levels 30 and 5 times higher in CSF than serum, respectively; and 3) brain-derived proteins that also have a blood-derived fraction in the CSF, like transthyretin, angiotensin-converting enzyme, and s-ICAM, which typically have lower relative concentrations in the CSF compared to serum (0.05, 0.01, and 0.005, respectively). In neurological diseases with blood-CSF barrier impairment, all blood proteins are elevated in the CSF.

We collected CSF from 13 patients with MS and 13 control patients, all scheduled to undergo lumbar puncture for medical reasons. A diagnosis of MS was confirmed based on the criteria of McDonald [74]. The control CSF samples were from subjects with a variety of conditions; only subjects with a negative CSF diagnosis were chosen for this study. CSF was processed in a method similar to that described above; however, about 100 times more CSF volume (0.5 mL) was used than serum in order to load the same amount of protein (20 \(\mu\)g) into the liquid chromatography-mass spectrometry instrument.
Proteomic differences between MS and control subjects were found in both the CSF and serum. Summary statistics on the proteomic results are provided in Table 1. The serum samples tracked 1.8 times more molecular components than the CSF samples (7000 vs 4000 components). The median CVs were lower for the serum samples (28% for MS and 27% for controls) than the CSF samples (35% for MS and 40% for the controls). The larger CVs for the CSF are likely to reflect greater variation in sample collection and greater variation in the clinical status of the control CSF group.

The number of significant changes for the control vs MS group comparisons are presented for each sample type and ranked by p-value. There are multiple comparisons in this study, and more significance should be attached to differences with lower p-values. There were about twice as many differences between the MS and control groups at p-value levels < 0.01 for CSF than serum. This difference is most significant given both the smaller number of samples and the smaller number of components tracked for CSF. The result strongly supports the hypothesis that CSF is a better source than serum for initial identification of potential biomarkers for MS.

Differential profiling for CSF is presented as a heat map in Fig. 2. The MS and control groups can be readily distinguished by the pattern of expression. The plot shows the relative intensity for components (rows) with p-values < 0.05 for every subject (columns). Results are sorted by effect size, which is the mean difference between the two groups, divided by the weighted standard deviation of the groups. In addition to between-group differences, individual profiles (columns) can be identified in both the MS and control groups, suggesting biological heterogeneity among the subjects within the groups.

As expected, many proteins are found in abundance in serum, and some of these are differentially expressed in MS subjects. Identified proteins come from a variety of biological processes, including inflammation, the complement cascade, lipoproteins and acute phase reactants. This last group of molecules is of particular interest. The acute phase response is a physiologic reaction occurring soon after the onset of infection, trauma, or inflammatory processes. The most prominent change is a dramatic increase of acute phase proteins, which are any of the non-antibody proteins found in increased amounts in serum during this reaction; examples include C-reactive protein, serum amyloid A protein, fibrinogen and 1-acid glycoprotein. The concentrations of many of the proteins within this collection have been reported to be altered in MS.

At this point, about 15% of the components in the data set have been identified. Some of the specific differences between MS and control subjects are presented in Fig. 3. The x-axis plots the effect size. Most of the proteins present in Fig. 3 are tracked with more than one peptide, as indicated in the parentheses. This consistency increases our confidence in individual protein results. An example of this is seen in the consistent results for leucine-rich alpha-2-glycoprotein in serum, with three different peptides shown. Several of the components presented, such as A2GP1 and haptoglobin, are acute phase reactants that could be as-

<table>
<thead>
<tr>
<th>Table 1 Summary Statistics for Proteomic profiling of Serum and CSF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Samples</td>
</tr>
<tr>
<td>Components</td>
</tr>
<tr>
<td>Median CV</td>
</tr>
<tr>
<td>Number of significant changes</td>
</tr>
<tr>
<td>P &lt; 0.001</td>
</tr>
<tr>
<td>P &lt; 0.01</td>
</tr>
<tr>
<td>P &lt; 0.05</td>
</tr>
</tbody>
</table>

Fig. 2. Proteome Difference Map for CSF comparison. Differences with a p-value < 0.5 are shown for the MS vs. Control comparison. Each column represents one subject and each row represents one molecular component. Relative value is a normalized z score (Xi - Xave/SD), where Xi is the intensity for sample i, Xave is the average intensity, and SD is the weighted standard deviation. Values are sorted by effect size, the mean difference between the two groups, divided by the weighted standard deviation of the groups. Here, a positive effect size indicates higher means for the MS group.
associated with inflammatory conditions. Some other acute phase reactants, like alpha-1-antitrypsin, were not significantly elevated in serum.

The effect sizes shown in Fig. 3 are quite substantial, with several greater than one. For reference, in another neurological disorder, Tau and Aβ peptide have an effect size of about one in Alzheimer’s disease vs either age-matched healthy controls or age-matched dementia controls. Our calculation is based on data from Mulder et al and references therein [76]. We have also completed a proteomic analysis of serum samples from rheumatoid arthritis subjects with active disease (tender/swollen joints, elevated erythrocyte sedimentation rate, etc) and controls [65]. A broader constellation of markers consistent with a chronic inflammatory condition was observed in the rheumatoid arthritis study. Many have effect sizes greater than one. The relapsing-remitting MS subjects in the current study do not necessarily have active clinical events.

Many of the proteins identified in serum are also detected in the CSF. These include alpha-1-antitrypsin, alpha-1-acid glycoprotein, haptoglobin, transferrin, and transthyretin. Interestingly, studies that have associated these molecules with MS have been reported: Proteases and their inhibitors have been implicated in the pathogenesis of MS. Levels of the protease inhibitor, alpha-1-antitrypsin, in the CSF have been associated with disturbances in the BBB [77]. The acute phase protein, alpha-1-acid glycoprotein (orosomucoid) levels are altered in the CSF of patients with MS [78]. Although the function of alpha-1-acid glycoprotein is unknown, it has protective effects in several in vivo models of inflammation [79]. Haptoglobin irreversibly binds free hemoglobin to facilitate its removal and increased synthesis of haptoglobin is observed in conditions of extensive tissue damage and necrosis. Its altered levels have been observed in inflammatory neurological diseases [80]. Studies indicate that free radicals participate in the pathogenesis of MS, and iron has been implicated as the catalyst leading to their formation. Thus, proteins involved with iron metabolism, such as transferrin, have been proposed as putative biomarkers for demyelination [81]. Finally, the level of transthyretin, a choroid plexus-specific transport protein, has been reported to be elevated, decreased and unaffected in the CSF of patients with MS [82]. Our finding that it is elevated relative to the non-MS group underscores our belief that many of these potential markers warrant re-evaluation with new and more sensitive assays.

We also identified several proteins produced in the central nervous system. These include beta trace protein (prostaglandin-D-synthase) and cystatin C (gamma trace), which are produced by the cells of the cerebral cortex and choroid plexus, and transthyretin, which is synthesized both by the choroid plexus and by the liver. Acute phase proteins e.g. alpha-1-antitrypsin, alpha-1-acid glycoprotein, haptoglobin, and transferrin show stronger and more consistent effects in the CSF than serum. Often more consistent peptides per protein

Fig. 3. Proteomic differences between MS subjects and controls. The effect size, the mean difference between the two groups, divided by the weighted standard deviation of the groups, is shown for a subset of identified components with p < 0.05. The direction of the triangle symbols are (▲) higher in controls, (▼) lower in controls. Numbers in parentheses are the consistent components for that protein with p < 0.05.
are different in the CSF. Among proteins produced in the brain, gamma trace (cystatin C), neurotrimin, chromogranin B, and transthyretin are significantly different between the MS and control groups. Distributions for representative peptides from three of these proteins identified in CSF are presented as box and whisker plots in Fig. 4. All have effect sizes greater than one, which corresponds to the boxes of the MS and control groups being almost separated. The displayed proteins are of neurological interest. Chromogranin B (secretin 1), a neuroendocrine secretory granule protein produced in the brain and pancreas, is more abundant in the CSF of MS subjects. The single study we found, reported that the CSF levels of both chromogranin A and B are similar among MS and controls groups [83], emphasizing again that further investigation is required. Neurotrimin, a neural cell adhesion molecule produced in the brain as well as muscle, liver, and other tissues, is found at lower levels in the CSF in MS subjects. Interestingly, this protein has been implicated, in part, in regulating the development of neuronal projections and axonal fasciculation [84,85]. Transthyretin (prealbumin), which probably transports thyroxine from the bloodstream to the brain, is elevated in the CSF of MS subjects. All three proteins have potential connections to MS disease physiology and require more extensive evaluation.

8. Summary and conclusions

Exploration of the etiology and pathogenesis of MS has yielded a handful of useful biomarkers. Markers of both autoimmune and neurodegeneration processes have been reported. Still, the need for markers for accurate diagnosis, course of severity, likelihood of onset, and determinants of etiology and pathogenesis are desired. The existence of informative single MS biomarkers in such a complex system is not likely. A pattern of many single molecules that are not useful individually may hold the key. Our strategy of comprehensive phenotyping based on discovery rather than hypothesis holds promise for elucidation of such markers. In addition to these potential benefits, this strategy offers the researcher an extensive library of new candidate molecules for support of current theories or development of new ideas in the pathophysiology of the disease. Finally, it must be considered that the molecular fingerprint of a patient with MS is likely to evolve over time relative to the underlying pathology. Recognition and characterization of this alteration will also become a rich source of information regarding the disease process and may yield predictors of various disease parameters in multiple sclerosis.

Acknowledgements

KCO is a National Multiple Sclerosis Society Career Transition Award recipient (TA3000A2). DAH is an NIH, Javits Investigator Award recipient. This work was funded, in part, through a grant from The Autoimmune Prevention Center awarded to DAH, ABK, and KCO and by NIH NINDS Grant R43 NS046107 to ABK and by the NIH grants: U01DK6192601, RO1NS2424710, PO1AI39671, and PO1NS38037; and grants from the National Multiple Sclerosis Society: RG2172C9 and RG3308A10 to DAH. We thank...
Jeffrey Satkofsky and Hua Lin for contributions to the mass spectrometry work, and Praveen Kumar and Markus Anderle for contributions in the areas of data processing and statistics and Paul Guttry for editorial assistance.

References

oligoclonal IgM against myelin lipids predicts an aggressive disease course in MS, J Clin Invest 115(1) (2005), 187–194.


