Molecular markers in the CSF proteome differentiate neuroinflammatory diseases

Maria Louise Elkjaer
Odense Universitetshospital
Arkadiusz Nawrocki
Syddansk Universitet
Tim Kacprowski
Technische Universitat Munchen
Pernille Lassen
Syddansk Universitet
Anja Hviid Simonsen
Kobenhavns Universitet
Romain Marignier
Centre de Recherche en Neurosciences de Lyon
Tobias Sejbaek
Odense Universitetshospital
Helle Hvisted Nielsen
Odense Universitetshospital
Lene Wermuth
Odense Universitetshospital
Alyaa Rashid
Syddansk Universitet
Peter Hogh
Kobenhavns Universitet
Finn Sellebjerg
Kobenhavns Universitet
Richard Reynolds
Imperial College London
Jan Baumbach
Technische Universitat Munchen
Martin R Larssen
Syddansk Universitet
Zsolt Illes (zsolt.illes@rsyd.dk)
Syddansk Universitet https://orcid.org/0000-0001-9655-0450

Research

Keywords: CSF, proteomics, multiple sclerosis (MS), Neuromyelitis Optica Spectrum Disorder (NMOSD), molecular markers, Chitinase-3-like protein 1 (CHI3L1), metalloproteinase inhibitor 1 (TIMP1), GFAP, trypsin-1 protein, apolipoproteins, human brain lesions

Posted Date: April 6th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-19710/v1

Abstract

Background: Multiple sclerosis (MS) is characterized by different degree of inflammatory and neurodegenerative features in the early relapsing vs. progressive subtypes. By using controls with different extent of inflammation vs. neurodegeneration, we examined the CSF proteome to identify molecular markers that differentiate between subtypes of MS. Gene expression of specific proteins were explored in MS brain lesions with diverse pathological background.

Methods: (i) First, we compared the proteome by LC-MS/MS in 169 pooled CSF from MS subtypes to inflammatory/degenerative controls: AQP4-IgG-positive and AQP4-IgG-negative neuromyelitis optica spectrum disorder (NMOSD), Alzheimer's disease (AD), and healthy controls. F-test based feature selection was used to cluster diseases and MS subtypes. (ii) Next, we selected 299 molecules by comprehensive statistics, and quantified them in the individual CSF samples. (iii) We also screened the genes of MS-specific CSF proteins in transcriptomes of 73 MS brain lesions with different pathology.

Results: We identified 11 proteins that separated diseases, and 8 proteins that clustered MS subtypes. Secondary progressive (SP)MS had the most unique proteome characterized by upregulation of intrinsic pathway proteins of the coagulation pathway. SPMS also clustered far from NMOSD indicating less inflammatory pathways. Primary progressive (PP)MS was more similar to relapsing-remitting (RR)MS than SPMS. Quantification of 299 proteins in 170 individual CSF samples identified 5 molecules uniquely upregulated in MS subtypes and in AQP4-IgG-positive NMOSD, respectively. Chitinase-3-like protein 1 (CHI3L1) was upregulated in part of PPMS and remission CSF samples, and it was expressed by astrocytes in chronic active lesions. GFAP was upregulated in 70% of AQP4-IgG-positive NMOSD but only in 40% of AQP4-IgG-negative NMOSD.

Conclusions: By the combination of untargeted and targeted quantitative analysis, we identified CSF molecular markers of axonal growth inhibition, lipid binding, and protein/lipid transport that differentiated between neuroinflammatory and neurodegenerative diseases, and also MS subtypes. The majority of them were expressed in MS brain lesions suggesting their origin from the brain tissue and not from the systemic compartment. Data suggest that the CSF proteome of SPMS is different from PPMS, and astrocyte damage may not be major pathology in part of the AQP4-IgG seronegative NMOSD.

Background

Identification of specific molecular markers that reflect the pathology and disease course of multiple sclerosis (MS) is difficult because of the dynamic and complex molecular pathogeneses. Early in the course, MS is characterized by clinically active and silent phases (relapsing-remitting, RRMS). A secondary progressive phase (SPMS) evolves in a subset of patients, where a combination of neurodegenerative processes, adaptive and innate immune responses contributes to the advancing disability, and limits the efficacy of disease modifying treatments (DMTs) that target mainly systemic adaptive immune responses[1–][4]. One out of eight MS patients are diagnosed with primary progressive (PP)MS characterized by the absence of clinical relapses and gradual worsening from onset. Axonal degeneration, cortical

lesions, CNS innate immune responses, inflammatory demyelination and remyelination significantly influence the prognosis and long-term outcome of MS[1, 4, 5]. Early prediction of mechanisms that culminate in the progressive phase may provide a more individualized treatment approach and postpone the secondary phase[6].

Hypothesis-driven exploratory omics approaches are effective tools for revealing molecular pathways and quantifying differentially expressed molecules to identify multiple markers that may predict disease outcomes. Mass spectrometry is an analytical technique for the characterization of biological samples and is increasingly used in omics studies as both a nontargeted and targeted approach for discovery proteomics and quantification with high throughput abilities. Proteomics of the cerebrospinal fluid (CSF) reflects more specific changes related to CNS damage than serum, and is a powerful tool for elucidating mechanisms by networks, pathways, protein groups and individual proteins that reflect both the similar and the unique molecular events as inflammation, degeneration, reparation or oxidative stress conditions in the MS subgroups[7].

Here, we used a comprehensive two-stage approach, with an untargeted and then a quantitative targeted method to characterize the molecular landscape of the CSF in different phases of MS. We aimed to identify molecules in the CSF with a potential clinical interest in MS subtypes, and to better understand MS pathophysiology at different stages. Disease controls were selected to include conditions with strong inflammatory alterations in the CNS without major degenerative processes but with similarity to MS, i.e. AQP4-IgG^{+} and AQP4-IgG^{-} neuromyelitis optica spectrum disease (NMOSD)[8], and neurodegenerative conditions associated with innate inflammatory responses in the CNS, i.e. Alzheimer disease (AD)[9]. Based on the different protein abundances in 169 CSF samples, we: (i) clustered the diseases and MS subtypes based on similarities and differences. (ii) selected hundreds of proteins that were quantified in 170 individual CSF of the MS subgroups and controls; and (iii) compared the unique CSF proteins associated with MS stages/subtypes with MS brain lesion signatures[10].

Methods

Study design and participants

We examined the CSF proteome in a two-stage approach, with an untargeted (n = 170) and then a quantitative targeted method (n = 169, partially overlapping) (Additional file 1). CSF samples were obtained through regional, national and international collaboration (Odense, Svendborg, Esbjerg, Copenhagen, Roskilde, Lyon, Hungary) from patients with newly diagnosed, untreated RRMS (age 33.6 ± 10 years, 77% female) in relapse (n = 14) or remission MS (n = 33), untreated PPMS (n = 30, age 49 ± 8.6, 57% female), untreated SPMS (n = 26, age 45.9 ± 5.8 years, 52% female), AD (n = 22, age 72.2 ± 7.9 years, 50% females), NMOSD AQP4-IgG^{+} (n = 14, age 47.9 ± 15.3 years, 78% female), NMOSD AQP4-IgG^{-} (n = 5, age 26.8 ± 13.2, 90% female) and healthy controls (n = 33, age 37.7 ± 12.9 years, 62% female). None of the patients with MS had had received disease-modifying therapy. Relapse was verified by neurologists, and samples were taken within maximum a month after the first relapse symptoms. Patients with AQP4-IgG^{-} NMOSD were

not treated with immunosuppressive medications, while patients with AQP4-IgG$^+$ NMOSD received azathioprine or mycophenolate mofetil. NMOSD was stable in all patients.

CSF samples were obtained by lumbar puncture, collected in polypropylene tubes and gently mixed. The samples were centrifuged at 2000 × g for 10 min at 4 °C to remove cells and other insoluble materials and stored in polypropylene tubes at − 80 °C pending analysis.

The study was conducted in accordance with the approval of the Danish National Ethics Committee (S-20120066), and informed consent was obtained from each participant.

Sample Preparation for Proteomic Discovery

CSF samples of each disease group were pooled into one of three sample pools producing three technical replicates (Fig. 1A). Proteins were ethanol/acetone precipitated, re-dissolved in 2M thiourea, 20 mM dithiothreitol (DTT), and the protein amount was estimated using Qubit Protein Assay (Thermo Fisher Scientific). Following alkylation, proteins were digested with LysC (0.02 AU/mg proteins) for 4 h, and then with trypsin (50:1 ratio) overnight at 37 °C. Peptides were reverse phase (RP) purified using homemade columns of C8/R2 and C18/R3 (Applied BiosystemsTM). Purified peptides were re-dissolved in 0.1% formic acid. The peptide amount in each sample was determined by amino acid composition analysis (AAA). Subsequently, equal amounts of each sample pool were labelled with one of the iTRAQ 8plex reagent labels according to manufacturer protocol. The bulk peptide sample was fractionated using hydrophilic interaction chromatography (HILIC), and each fraction was further separated by reverse phase chromatography prior to identification by mass spectrometry (Q Exactive HF, Thermo Fisher). The three technical replicates of the sample pools were run separately (Additional file 2A).

Data Availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[11] partner repository with the dataset identifier PXD017643.

Statistical Analyses for Selection of Proteins

Proteome Discoverer software (further PD software, Thermo Scientific, v1.4) was used to process the raw mass spectrometry (MS) files, identify the proteins and generate quantitative data which was further processed by three parallel approaches.

ANOVA-based (analysis of variance). For each peptide, ANOVA was performed with the lmPerm R package to determine difference between groups. Afterwards, to determine which pairs of groups showed most differences, the Tukey's HSD (honest significant difference) test was performed as post-hoc analysis.

Limma-based (linear models). Linear regression and analysis of variance were performed with the limma R package. The ratios of a specific protein between two compared groups were $\log_2$ transformed, normalized to the median, and the 3 replicates merged into one, and proteins were significant according to q-values (FDR < 0.1). The resulting data were visualized in volcano plots and heatmaps.

Complementary analysis of the three replicates. Using the PD software, for each of the three sets the coefficient of variation CV of proteins (any subject group to healthy subjects) within the set as well as the ratio of the mean abundance between the sets were calculated. A protein was selected for further analyses, if the ratio was larger (or smaller) than 1 + 2xCV (or reciprocal). Subsequently, the PD software calculated a “global” ratio for a protein based on data from the three sets compared to healthy samples (and CV within the combined sets). Proteins were finally selected, if the protein expression was larger (or smaller) than 1 + 2xCV (or reciprocal) at least between two different conditions, and was consistently altered in a minimum of two of the three sets.

Linear discriminate analysis (LDA)

To reduce any possible batch effect, the three pools were merged after scaling them individually (per protein). An F-test based feature selection was performed, where only proteins with a FDR < 0.05 (ANOVA) were considered. Next, the set of candidate proteins were pruned for collinearity by iteratively removing the protein with the highest variance inflation factor (VIF), until only proteins with VIF < 10 remained. This resulted in 11 proteins, which were used to conduct a linear discriminant analysis (LDA). Additionally, the test was also performed only on the MS samples resulting in 8 proteins responsible for the subgroup separation according to the LDA.

Pathway Analysis

After the data were normalized to control samples, Ingenuity Pathway Analysis (IPA) was used to identify molecular pathways and perform functional analysis between different disease groups and subgroups.

Sample Preparation for Quantification

CSF from each patient was precipitated with ethanol/acetone, dissolved in urea buffer containing DTT, as described in a previous paper with parallel reaction monitoring (PRM)[12].Total protein content was estimated by AAA, and 10 ug of proteins were digested with trypsin. After digestion, Stable Isotope Standards (SIS) mix was added in equal volume to every sample (both previously prepared[12] and additional ones). Peptides in each sample were labelled with one of the TMT 11plex label. A pooled sample was prepared by mixing a small amount from approximately half of all the available samples. This pooled sample was labelled with TMT 11plex 126 label. Subsequently, this pooled sample was split equally into 17 samples, and mixed with ten other patient samples in a random manner (Additional file 2B). There were 17 TMT sets each containing at least one (if available) sample from every patient group. Samples were randomized so that each set contained a representative of each patient group, and each sample of every patient group was labelled with a different TMT label.

Peptides of every set were fractionated by HILIC and analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS). The LC method total runtime varied in length between 67 to 143 minutes depending on signal intensity of the HILIC fractions. Most of the peptides separated during the linear increase of solvent B from 10 to 35% in 38 to 120 min (corresponding to the total runtime). MS settings: Full MS: Resolution at 120000, AGC target 3e6, Maximum IT 100 ms, scan range: 325–1600 m/z. MSMS

settings: Resolution at 60000, AGC target 1e5, Maximum IT 100 ms, isolation window 1.2 m/z, NCE: 32, top 15 most intense ions of 2–4 charges (positive mode), dynamic exclusion of 15–20 seconds.

Data Processing and Statistical Analyses of Validated Proteins

The raw data was processed with the ProteomeDiscoverer software (v2.3). The samples used for analysis contained SIS standard added in the same amount to each sample and labelled with TMT along with all the other CSF peptides. Each patient group was set as one of the Categorical factors, and every patient within a patient group was set as a Biological replicate. The Pool sample was set as “Control” and every patient sample was set as “Sample”. The scaling parameter was set “On Average Control”. In this way, samples were normalized and scaled to the Pool (which is a common/identical sample across the 17 replicates). The software calculated ratios for protein abundances between any patient group and healthy controls based on proteins identified and quantified in corresponding samples from all the 17 Sets. In an alternative approach, the quantitative data from ProteomeDiscoverer were extracted and further processed in Excel (Microsoft). The constant ratio of CSF proteins to SIS were used to calculate normalization factors within each of the 17 TMT sets. Additionally, this SIS normalization could also be used for correcting the few samples that contained less than 10 ug of proteins and different amount of volume. After normalization, an average ratio for each protein (for every patient group) was calculated based on the ratios to the corresponding protein in the Pooled sample. The significance of the ratios was validated by ANOVA by using PRISM and PolySTest software [13].

Human Brain Lesion Signature

Gene names of the protein of interest were uploaded in msatlas.dk and heatmaps were produced of genes present in the human MS brain[10]. Stars were added when there was a significant difference (FDR < 0.05) between MS lesion type and control white matter from non-neurological disease brain areas.

Immunohistochemistry and RNAscope of Chronic Active Brain Lesion

Human postmortem brain tissue were supplied by the UK Multiple Sclerosis Tissue Bank (UK Multicentre Research Ethics Committee, MREC/02/2/39), funded by the Multiple Sclerosis Society of Great Britain and Northern Ireland (registered charity 207,495). Fresh-frozen blocks containing chronic active lesion from progressive MS patients were sectioned (10-µm), PFA-fixed, blocked in PBS with 10% normal horse serum (NHS) and immunostained with rabbit CHI3L1 (monoclonal antibody) 1:200 (Abcam) followed by biotinylated secondary antibody (Jackson Immunoresearch Laboratories, Cambridgeshire, UK), avidin/biotin staining (Vector Laboratories, Burlingame, CA) and 3,3'-diaminobenzidine (DAB) staining (Vector Laboratories, Burlingame, CA). The RNAscope 2.5 Duplex Assay (ACD Biosystems) was performed according to the ACD protocol for fresh-frozen tissue. Chronic active lesions were hybridized with two mRNA probes per experiment. Hs-GFAP (Cat No. 311801) was used as the astrocyte marker together with Hs-CHI3L1 (Cat No. 408121). The probes were amplified according to manufacturer’s instructions and

labeled with the following red or green color for each experiment. The Hs-CHI3L1 probe was also combined with immunohistochemistry (anti-GFAP and anti-MHCII, Abcam) as described above.

Results

Global CSF proteome landscape of MS subtypes compared to NMOSD and AD

Untargeted analysis of the CSF proteome in MS subgroups and controls

Altogether, we detected 878 proteins in the 169 CSF samples. By using F-test based feature selection, 11 proteins were able to distinguish the disease (sub)groups (Fig. 1A). These 11 proteins were used to conduct a linear discriminant analysis (LDA) including both the disease groups and healthy controls: there was no overlap between the different disease groups, and no influence of the technical batch effect (Fig. 1B). NMOSD (AQP4-IgG$^-$ and AQP4-IgG$^+$) and SPMS were the most distinct groups both from each other and from healthy controls, PPMS, RRMS (relapse, remission) and AD. The presence of genes coding these 11 proteins in the MS brain was examined by using . All were expressed in the MS brain, and 5 of them were significantly differentially expressed in different lesion types (PEBP4, CNTNAP4, NRXN1, CPQ, OLFML3) (Fig. 1C).

F-test based feature selection was also applied to the MS CSF samples separately and resulted in 8 proteins differentiating the MS subtypes (early MS in remission and relapse, SP and PPMS) (Fig. 1D). The LDA according to these 8 proteins identified also the SPMS subtype as being the most different (Fig. 1E). Seven of the 8 genes encoding for the proteins were present in the MS brain, and 3 were significantly differentially expressed: GOLM in all the lesion types (active, chronic active, inactive and remyelinating), FRZB in active and chronic active lesions, and SELENBP1 in inactive lesions (Fig. 1F).

Next, we normalized the protein levels to healthy controls, and the diseases were clustered based on the abundance in protein $\log_2$ ratios (except AQP4-IgG$^-$ NMOSD due to lack of technical replicates) (Fig. 2A). AQP4-IgG$^+$ NMOSD was the most different from the other diseases, and SPMS the most different from the other MS subtypes (Fig. 2A). Volcano maps of normalized proteins in different disease groups also indicated that AQP4-IgG$^+$ NMOSD and SPMS had the highest amount of altered proteins compared to healthy controls (FDR < 0.1) (Fig. 2B).

Functional analysis of the CSF proteome

Functional classification and molecular pathways of the proteome in the different diseases was generated by Ingenuity Pathway Analysis (IPA) (Fig. 2C, Additional file 3). The most shared pathway was “LXR/RXR Activation” by SPMS, PPMS, MS remission, and AQP4-IgG$^+$ NMOSD. “Acute Phase Response Signalling” was shared between SPMS, AQP4-IgG$^+$ NMOSD and AD. “Axonal Guidance Signalling” was shared between

MS remission, PPMS and AD. PPMS and AD shared “Intrinsic Prothrombin Activation Pathway”. AD and AQP4-IgG$^+$ NMOSD shared “Complement”. SPMS had two unique pathways: “Neuroprotective Role Of THORP1 In AD” and “Coagulation System”, while PPMS and remission had one each, “FXR/RXR Activation” and “Clathrin-mediated Endocytosis Signalling”, respectively.

While CSF in MS relapse did not share common top pathways, the top pathways were “Hematopoiesis from Pluripotent Stem Cells”, “Leucocyte Extravasation Signalling” and “Agrin Interactions at Neuromuscular Junction” (Fig. 2C). The distinct biological functional enrichment of MS relapse was also reflected by the top 5 predefined diseases or functions (Fig. 2D). The top network assigned for all the disease groups were “Metabolic Disease”, “Cellular Movement”, ”Neurological Disease” and “Psychological Disorders”, while relapse only shared “Cellular Movement”.

Unique CSF proteins in disease subtypes

Quantification of disease-specific proteins in the CSF

By combining different statistical analyses of the pooled CSF samples (ANOVA, limma, complementary analysis), we selected 299 dysregulated proteins (Fig. 3A, Additional file 4). These were quantified in 170 individual CSF samples by mass spectrometry.

CHI3L1 and metalloproteinase inhibitor 1

The two proteins, chitinase-3-like protein 1 (CHI3L1) and metalloproteinase inhibitor 1 (TIMP1) that were significantly altered in all three tests were not significantly altered in the individual samples by the quantitative proteomics (Fig. 3B). However, some of the PP and remission patients had increased levels of CHI3L1, while some of the AQP4-IgG$^+$ NMOSD patients had increased levels of metalloproteinase inhibitor 1 compared to pool (Fig. 3B). By immunohistochemistry (IHC), we also found CHI3L1 increased at the rim of chronic active lesions in progressive MS WM tissue (Fig. 3C). The morphology of cells expressing CHI3L1 in chronic active lesions was consistent with astrocytes. The astrocytic expression was confirmed by combined RNAscope and immunohistochemistry that co-localized CHI3L1 and GFAP at the chronic active rim in close proximity to MHCII expressing cells (Fig. 3D-F).

Upregulated CSF proteins in the CSF of MS

The trypsin-1 protein was the most significantly upregulated protein in RRMS in remission, PPMS and SPMS subjects compared to both the disease- and healthy controls (Fig. 4A). Apolipoprotein C-I and augurin were also upregulated in these three MS subtypes compared to healthy controls and AD patients (Fig. 4B, 4C). Receptor-type tyrosine-protein phosphatase gamma was also upregulated in these three MS subtypes compared to disease controls (Fig. 4D). Apolipoprotein A-II was significantly upregulated in SPMS compared to relapsing MS, AQP4-IgG$^+$ NMOSD and healthy controls (Fig. 4E).

AQP4-IgG$^+$ NMOSD-specific CSF proteins

GFAP, inter-alpha-trypsin inhibitor heavy chain H1, and H2, serum amyloid P-component, and actin cytoplasmic 1 protein were uniquely upregulated in AQP4-IgG^{+} NMOSD compared to all MS subtypes, AD and healthy controls (Fig. 5). Glial fibrillary acidic protein/GFAP was detected only in less than 50% of the patients with AQP4-IgG^{-} NMOSD similarly to MS and AD.

CSF proteome signatures in MS brain lesion transcriptomes

We compared the CSF proteome signatures to the recently established transcriptome signatures of different MS lesion types (). Two of the MS-specific upregulated proteins were present as transcripts in the human progressive MS brain: apolipoprotein C-I (APOC1) was significantly upregulated in active lesions, and receptor-type tyrosine-protein phosphatase gamma (PTPRG) was significantly upregulated in all WM tissue of progressive MS (NAWM and lesions) (Fig. 6A).

Three of the five altered proteins in AQP4-IgG^{+} NMOSD patients were also detected as transcripts in the MS WM brain tissue: glial fibrillary acidic protein (GFAP) was upregulated in active, inactive and remyelinating lesion types, inter-alpha-trypsin inhibitor heavy chain H2 (ITIH2) was significantly upregulated in all lesion types, while actin cytoplasmic 1 (ATCB) was not differently expressed compared to non-neurological-disease WM brain areas (Fig. 6B).

Discussion

This comprehensive two-stage proteomic study with a high number of human CSF samples (n = 169) from a spectrum of different neurological diseases provided information about the global CSF proteomic landscape in disease subgroups compared to healthy controls.

Mapping the global CSF proteome in a spectrum of neurological disorders

With F-test based feature selection, a combination of 11 proteins could separate the diseases without overlap and technical batch effect. Due to the pruning for collinearity during the F-test, these proteins have very different functions, i.e. axonal growth inhibition (RTN4R), cell-matrix interactions (CRTAC1), lipid binding and inhibition of serine proteases (PEBP4), and hydrolysis of circulating peptides (CPQ). The combination of 8 additional proteins could also separate the MS subgroups, and 4 were related to intracellular processing and transporting of synthesized proteins and lipids (GOLM1, NUCB1, NPC2, SELENBP1). LDA and differential abundance of proteins indicated that SPMS and AQP4-IgG$^{+/-}$NMOSD clustered far from each other, and SPMS differed the most from the other MS subtypes (Fig. 1A). These two disease groups also had the highest number of significantly altered proteins (FDR < 0.1) (Fig. 1B). We examined the presence of the 17 differentiating molecules among transcripts in different brain MS lesions[10], and identified 8 genes that were significantly differentially expressed. Neurorexin-1 has been related to neurodegeneration in MS[14], and NRXN1 was uniquely significantly upregulated in the chronic active lesion type associated with progressive MS. Olfactomedin-like protein 3 is a known marker of

activated ramified microglia, and OLFML3 was also significantly upregulated in chronic active lesions as well as in NAWM. Selenium-binding protein 1 is an astrocytic marker related to metabolic processes[15], and SELENBP1 was uniquely expressed in inactive lesions characterized by astrocytic scar tissue[10]. Secreted frizzled-related protein 3 is involved in axon targeting basement membrane breakdown[16], and the FRZB gene was significantly upregulated in active and chronic active lesion types. Contactin-associated protein-like 4 is involved in the formation and maintenance of myelinated axons[17], and CNTNAP4 was upregulated in the inactive lesion type. This molecular CSF profile and associated brain lesion spectrum highlights the importance of non-inflammatory mechanisms in differentiating these diseases.

We also examined pathways that were different among diseases and MS subtypes (Fig. 2). In this regard, relapse was the most distinct disease group with almost nothing in common with the other diseases. It was dominated by unique immune-related pathways, and the top predicted diseases/functions were more related to systemic than CNS-specific events. Although early MS in remission and PPMS resembled each other the most by sharing the top 5 predicted diseases/functions, they differed by specific pathways suggesting “clean-up” events in remission and regulating lipid and glucose metabolism in PPMS (FXR/RXR Activation). The unique SPMS enriched pathway was the “Coagulation” system, while PPMS and AD shared “Intrinsic Prothrombin Activation Pathway”. A previous study also found proteins involved in coagulation unique to chronic active lesion samples, suggesting dysregulation of molecules associated with coagulation in chronic active lesions[18]. Another recent study also identified higher levels of CSF proteins related to the coagulation cascade in MS patients with higher cortical lesion load[19].

Unexpectedly, in our study immune related proteins such as cytokines, chemokines, growth factors and adhesion molecules were not frequently detected. This could be because of the constrained dynamic range of mass spectrometers to truly cover the broad spectrum of lower abundance. A recent systematic review revealed 19 inflammatory proteins specifically altered in MS[20]. Not surprisingly, the majority of the upregulated MS proteins (11 of 19) were immunoglobulins. However, most of these proteins also appeared to be highly abundant in the CSF[21, 22], and low abundant proteins likely to be involved in the distinct damaging vs. repairing processes of MS remain to be discovered.

Disease-specific molecular markers

Next, 299 proteins were selected and quantified in 170 individual CSF samples (169 of these were also used for the discovery phase). This identified 12 molecules potential interest, including10 molecular markers specific to MS and AQP4-IgG^{+} NMOSD.

Molecular markers significantly altered in all three tests in the discovery analysis but not significant in quantification of individual samples

Two proteins (CHI3L1 and TIMP1) were significantly altered in all three statistical tests (ANOVA, limma, complementary analysis) in the pooled discovery CSF proteome, but were not unique to diseases in the individual quantification study. However, a subgroup of MS patients with PP and remission had increased levels of CHI3L1 (Fig. 3B). CHI3L1 (YKL-40) is a promising biomarker of inflammation in progressive

MS[23], and was originally discovered in the CSF proteome of patients with CIS converting to RRMS[16]. Immunohistochemistry and RNAscope indicated that the gene encoding CHI3L1 was primary expressed by astrocytes in the rim of chronic active lesions (Fig. 3C-F). Another recent study also found that CHI3L1 reflects disease progression, and together with the biomarker neurofilament light chain protein, it may help to discriminate MS phenotypes[24]. These data suggest that some of the emerging biomarkers in progressive MS may reflect unique molecular changes in the brain related to specific subtypes of lesions. The high expression of CHI3L1 in the CSF of patients with progressive MS[25] may be related to the increasing number of a specific subtype of chronic active lesions, and we may speculate that its level in the CSF of patients with progressive MS may even reflect the number of this lesion type in the brain. The expression of CHI3L1 by astrocytes has been recently described in neurodegenerative diseases and often appears in clusters of astrocytes[26]. Knock-out animal models indicated a protective role of CHI3L1, as traumatic brain injury and experimental autoimmune encephalomyelitis were more severe in its absence[27, 28]. CH13L1 can also influence the migratory capacity of astrocytes and reduces astrogliosis [27, 28]. It may therefore dampen the inflammation and limit astrogliosis.

TIMP-1 seemed to be highly expressed in a subset of AQP4-IgG$^+$ NMOSD patients (Fig. 3B). TIMP-1 is produced by astrocytes in both homeostasis and early/acute inflammatory events[29]. We have previously found TIMP-1 peak during acute remyelination in the cuprizone model and to be associated with reduced inflammation in the CSF of MS[12]. Induction of TIMP-1 in neurons and astrocytes was also related to early cellular events triggered by seizures and with long-lasting changes in tissue reorganization and/or neuroprotection[30]. Increased TIMP-1 levels in serum has also been proposed as a prognostic biomarker of mortality in brain trauma injury patients[31]. Therefore, increased TIMP-1 and CH13L1 in the CSF may reflect acute and chronic astrocytic responses in subgroups of MS and AQP4-IgG$^+$ NMOSD patients.

Molecular markers of MS

Two apolipoproteins were found increased in MS. These are important players in cholesterol homeostasis, and in CNS diseases for neuronal homeostasis and regeneration[32]. Apolipoprotein C-I was significantly upregulated in RRMS in remission, PPMS and SPMS, and its transcript was significantly induced in active MS lesions in SPMS brain (Fig. 6A). Apoprotein A-II was significantly altered in the CSF in SPMS compared to both AQP4-IgG$^+$ NMOSD and healthy controls (Fig. 5E). Increased levels of apoprotein A-II has been associated with fatigue in MS patients[33], and it may reflect later disease mechanisms accumulated with chronic damage. Apolipoproteins have been also linked to the genetic risk of MS: APOE genotype has been associated with disease severity and MR activity[34–36].

Trypsin-1, a protease that degrade other proteins, was also significantly upregulated in remission, PP and SPMS compared to the disease- and healthy controls. Since we were not able to detect the gene of this protein (PRSS1) expressed in the MS brain[10], its presence in the CSF may originate from the systemic compartment.

Receptor-type tyrosine-protein phosphatase gamma (PTPRG) levels were increased RRMS in remission, PP and SPMS compared to the disease controls, but only RRMS in remission was significantly upregulated compared to healthy controls (Fig. 5D). Another study also found it increased in the CSF of early MS patients compared to controls[37], suggesting that it may be induced from onset of the disease. We also found it significantly upregulated in progressive MS tissue in both NAWM and all kind of lesions (Fig. 6A).

Molecular markers in AQP4-IgG$^+$ NMOSD

We found 5 upregulated unique molecular markers in the CSF of patients with AQP4-IgG$^+$ NMOSD. Increased GFAP reflects astrocyte damage and death in AQP4-IgG$^+$ NMOSD[38, 39]. It was not increased in AQP4-IgG seronegative NMOSD indicating that at least in a subset of these patients' disease mechanisms do not primarily target astrocytes. Another study also reported higher GFAP levels in AQP4-IgG$^+$ patients compared to AQP4-IgG$^-$ NMOSD[40]. The unique elevation of serum amyloid P-component in the CSF in AQP4-IgG$^+$ NMOSD may be related to damage to the blood-brain barrier[41]. Upregulation of inter-alpha-trypsin inhibitor heavy chain H1 and H2 may represent endogenous neuroprotective immunomodulatory proteins within the CNS[42]. The ITIH2 gene was significantly upregulated in all lesion types in the MS brain (Fig. 6B), suggesting that this molecule can be an indicator of non-specific neurological inflammatory damage and control.

Conclusions

With the combination of untargeted and targeted quantitative proteomic analysis of the CSF, we identified molecular markers that differentiated between neuroinflammatory and neurodegenerative diseases, and also MS subtypes. Our data support a recent observation that the coagulation system is an important pathway in SPMS. We found that the proteome of SPMS was the most different from the other subtypes of MS including PPMS. Genes encoding proteins that clustered disease and MS subtypes could also be detected in MS brain lesions. Among specific CSF proteins in NMOSD, GFAP was present in 70% of AQP4-IgG seropositive NMOSD and significantly upregulated; however, the 40% detection level in AQP4-IgG seronegative NMOSD may suggest that astrocyte damage may not be major pathology in part of these patients. The absence of an independent cohort validation, different ages of patients in disease subgroups, and immunotherapy of NMOSD patients are limitations of the study.

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with the approval of the Danish National Ethics Committee (S-20120066), and informed consent was obtained from each participant.

Consent for publication

'Not applicable' for that section.

Availability of data and materials

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[11] partner repository with the dataset identifier PXD017643.

Competing interests

Dr. Illes has served on scientific advisory board for Biogen, Sanofi, Teva, Roche, Novartis, Merck, has received honoraria for lecturing from Biogen, Merck, Novartis, Sanofi and for clinical endpoint committees in clinical trials of NMOSD, and support for congress participation from Biogen, Genzyme, Merck, Teva, Roche. Dr. Sellebjerg has served on scientific advisory boards, been on the steering committees of clinical trials, served as a consultant, received support for congress participation, received speaker honoraria, or received research support for his laboratory from Biogen, Merck, Novartis, Roche, Sanofi Genzyme and Teva.

Funding

Lundbeckfonden R118-A11472, OTKA-K77892, Scleroseforeningen R431-A29926-B15690 and R399-A28099-B15690, Region of Southern Denmark 14/24200, Jascha Fonden 5589, Direktør Ejnar Jonasson kaldet Johnsen og hustrus mindelegat 5609, Odense University Hospital (OUH) A474 (to ZI), OTKA-NN109841, GINOP 2.3.2-15-2016-00049. JB is grateful for financial support from JB's VILLUM Young Investigator grant nr. 13154. NA received funding from JB's SDU2020 grant. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

Conception and design of the study: MLE, AN, JB, MRL, ZI

Acquisition and analysis of data: MLE, AN, TK, PL, AHS, RM, TS, HHN, LW, PH, FS, RR, JB, MRL, ZI

Drafting a significant portion of the manuscript and figures: MLE, AN, TK, JB, MRL, ZI

Acknowledgements

'Not applicable' for that section.

References

1. Bates D. Treatment effects of immunomodulatory therapies at different stages of multiple sclerosis in short-term trials. Neurology [Internet]. 2011 [cited 2019 Dec 26];76:S14–25. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21205678
2. Lassmann H, van Horssen J, Mahad D. Progressive multiple sclerosis: pathology and pathogenesis. Nat Rev Neurol [Internet]. Nature Publishing Group; 2012 [cited 2017 Jan 11];8:647–56. Available from:

http://www.nature.com/nrneurol/journal/v8/n11/full/nrneurol.2012.168.html%5Cnpapers://99cf836b-0208-4bf7-a35c-5506d5268000/Paper/p23902

3. Zaratin P, Comi G, Leppert D. 'Progressive MS – macro views': The need for novel clinical trial paradigms to enable drug development for progressive MS. Mult Scler [Internet]. 2017 [cited 2017 Oct 24];23:1649–55. Available from: https://doi.org/10.1177/1352458517729457
4. Prineas JW, Parratt JDE. Oligodendrocytes and the early multiple sclerosis lesion [Internet]. Ann. Neurol. 2012 [cited 2019 Dec 26]. p. 18–31. Available from: http://doi.wiley.com/10.1002/ana.23634
5. Kuhlmann T, Miron V, Cuo Q, Wegner C, Antel J, Brück W, et al. Differentiation block of oligodendroglial progenitor cells as a cause for remyelination failure in chronic multiple sclerosis. Brain [Internet]. 2008 [cited 2016 Dec 27];131:1749–58. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18515322
6. Derfuss T. Personalized medicine in multiple sclerosis: hope or reality? BMC Med [Internet]. 2012 [cited 2016 Feb 18];10:116. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23035757
7. Gustafsson M, Nestor CE, Zhang H, Barabási A-LL, Baranzini S, Brunak S, et al. Modules, networks and systems medicine for understanding disease and aiding diagnosis. Genome Med [Internet]. BioMed Central Ltd; 2014 [cited 2016 Jan 14];6:82. Available from: http://genomemedicine.com/content/6/10/82
8. Weinshenker BG, Wingerchuk DM. Neuromyelitis Spectrum Disorders [Internet]. Mayo Clin. Proc. 2017 [cited 2020 Jan 8]. p. 663–79. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28385199
9. Lane CA, Hardy J, Schott JM. Alzheimer's disease [Internet]. Eur. J. Neurol. 2018 [cited 2020 Jan 8]. p. 59–70. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28872215
10. Elkjaer ML, Frisch T, Reynolds R, Kacprowski T, Burton M, Kruse TA, et al. Molecular signature of different lesion types in the brain white matter of patients with progressive multiple sclerosis. Acta Neuropathol Commun [Internet]. BioMed Central; 2019 [cited 2019 Dec 28];7:205. Available from: https://actaneurocomms.biomedcentral.com/articles/10.1186/s40478-019-0855-7
11. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res [Internet]. 2019 [cited 2020 Feb 24];47:D442–50. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30395289
12. Martin NANA, Nawrocki A, Molnar V, Elkjaer MLML, Thygesen EKEK, Palkovits M, et al. Orthologous proteins of experimental de- and remyelination are differentially regulated in the CSF proteome of multiple sclerosis subtypes. PLoS One [Internet]. Public Library of Science; 2018 [cited 2018 Sep 27];13:e0202530. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30114292
13. Schwämmle V, Hagensen CE, Rogowska-Wrzesinska A, Jensen ON. PolySTest: Robust statistical testing of proteomics data with missing values improves detection of biologically relevant features. bioRxiv [Internet]. Cold Spring Harbor Laboratory; 2019 [cited 2020 Jan 21];765818. Available from: https://www.biorxiv.org/content/10.1101/765818v1
14. Kattimani Y, Veerappa AM. Dysregulation of NRXN1 by mutant MIR8485 leads to calcium overload in pre-synapses inducing neurodegeneration in Multiple sclerosis. Mult Scler Relat Disord [Internet]. 2018

[cited 2019 Dec 28];22:153–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29729524

15. Sasuclark AR, Khadka VS, Pitts MW. Cell-type specific analysis of selenium-related genes in brain. Antioxidants [Internet]. Multidisciplinary Digital Publishing Institute (MDPI); 2019 [cited 2019 Dec 28];8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31060314
16. Rich CA, Perera SN, Andratschke J, Stolt CC, Buehler DP, Southard-Smith EM, et al. Olfactory ensheathing cells abutting the embryonic olfactory bulb express Frzb, whose deletion disrupts olfactory axon targeting. Glia [Internet]. 2018 [cited 2019 Dec 28];66:2617–31. Available from: http://doi.wiley.com/10.1002/glia.23515
17. Zou Y, Zhang WF, Liu HY, Li X, Zhang X, Ma XF, et al. Structure and function of the contactin-associated protein family in myelinated axons and their relationship with nerve diseases. Neural Regen Res [Internet]. 2017 [cited 2019 Dec 28];12:1551–8. Available from: http://www.nrronline.org/text.asp?2017/12/9/1551/215268
18. Han MH, Hwang S-I, Roy DB, Lundgren DH, Price J V, Ousman SS, et al. Proteomic analysis of active multiple sclerosis lesions reveals therapeutic targets. Nature [Internet]. Nature Publishing Group; 2008 [cited 2014 May 28];451:1076–81. Available from: http://www.nature.com.proxy1-bib.sdu.dk:2048/nature/journal/v451/n7182/full/nature06559.html
19. Magliozzi R, Hametner S, Facchiano F, Marastoni D, Rossi S, Castellaro M, et al. Iron homeostasis, complement, and coagulation cascade as CSF signature of cortical lesions in early multiple sclerosis. Ann Clin Transl Neurol [Internet]. 2019 [cited 2020 Jan 24];6:2150–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31675181
20. Thygesen C, Larsen MR, Finsen B. Proteomic signatures of neuroinflammation in Alzheimer's disease, multiple sclerosis and ischemic stroke [Internet]. Expert Rev. Proteomics. 2019 [cited 2020 Feb 2]. p. 601–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31220951
21. Stoop MP, Singh V, Dekker LJ, Titulaer MK, Stingl C, Burgers PC, et al. Proteomics comparison of cerebrospinal fluid of relapsing remitting and primary progressive multiple sclerosis. Combs C, editor. PLoS One [Internet]. 2010 [cited 2020 Feb 2];5:e12442. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20805994
22. Kroksveen AC, Guldbrandsen A, Vedeler C, Myhr KM, Opsahl JA, Berven FS. Cerebrospinal fluid proteome comparison between multiple sclerosis patients and controls. Acta Neurol Scand [Internet]. John Wiley & Sons, Ltd (10.1111); 2012 [cited 2020 Jan 7];126:90–6. Available from: http://doi.wiley.com/10.1111/ane.12029
23. Burman J, Raininko R, Blennow K, Zetterberg H, Axelsson M, Malmeström C. YKL-40 is a CSF biomarker of intrathecal inflammation in secondary progressive multiple sclerosis. J Neuroimmunol [Internet]. 2016 [cited 2020 Jan 22];292:52–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26943959
24. Gil-Perotin S, Castillo-Villalba J, Cubas-Nuñez L, Gasque R, Hervas D, Gomez-Mateu J, et al. Combined Cerebrospinal Fluid Neurofilament Light Chain Protein and Chitinase-3 Like-1 Levels in Defining Disease Course and Prognosis in Multiple Sclerosis. Front Neurol [Internet]. Frontiers Media SA; 2019 [cited 2020 Jan 22];10:1008. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31608004

25. Sellebjerg F, Börnsen L, Ammitzbøll C, Nielsen JE, Vinther-Jensen T, Hjermind LE, et al. Defining active progressive multiple sclerosis. Mult Scler [Internet]. SAGE PublicationsSage UK: London, England; 2017 [cited 2018 Feb 1];23:1727–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28831853
26. Llorens F, Thüne K, Tahir W, Kanata E, Diaz-Lucena D, Xanthopoulos K, et al. YKL-40 in the brain and cerebrospinal fluid of neurodegenerative dementias. Mol Neurodegener [Internet]. 2017 [cited 2018 Jan 22];12. Available from: https://molecularneurodegeneration.biomedcentral.com/track/pdf/10.1186/s13024-017-0226-4?site=molecularneurodegeneration.biomedcentral.com
27. Wiley CA, Bonneh-Barkay D, Dixon CE, Lesniak A, Wang G, Bissel SJ, et al. Role for mammalian chitinase 3-like protein 1 in traumatic brain injury. Neuropathology [Internet]. 2015 [cited 2018 Aug 3];35:95–106. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25377763
28. Bonneh-Barkay D, Bissel SJ, Kofler J, Starkey A, Wang G, Wiley CA. Astrocyte and Macrophage Regulation of YKL-40 Expression and Cellular Response in Neuroinflammation. Brain Pathol [Internet]. 2012 [cited 2018 Aug 3];22:530–46. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22074331
29. Dhar A, Gardner J, Borgmann K, Wu L, Ghorpade A. Novel role of TGF-β in differential astrocyte-TIMP-1 regulation: Implications for HIV-1-dementia and neuroinflammation. J Neurosci Res [Internet]. NIH Public Access; 2006 [cited 2020 Jan 22];83:1271–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16496359
30. Rivera S, Tremblay E, Timsit S, Canals O, Ben-Ari Y, Khrestchatisky M. Tissue inhibitor of metalloproteinases-1 (TIMP-1) is differentially induced in neurons and astrocytes after seizures: Evidence for developmental, immediate early gene, and lesion response. J Neurosci [Internet]. Society for Neuroscience; 1997 [cited 2020 Jan 22];17:4223–35. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9151739
31. Lorente L, Martín MM, López P, Ramos L, Blanquer J, Cáceres JJ, et al. Association between serum tissue inhibitor of matrix metalloproteinase-1 levels and mortality in patients with severe brain trauma injury. PLoS One [Internet]. Public Library of Science; 2014 [cited 2020 Jan 22];9:e94370. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24728097
32. Gardner LA, Levin MC. Importance of Apolipoprotein A-I in Multiple Sclerosis. Front Pharmacol [Internet]. Frontiers; 2015 [cited 2019 Dec 29];6:278. Available from: http://journal.frontiersin.org/Article/10.3389/fphar.2015.00278/abstract
33. Browne RW, Jakimovski D, Ziliotto N, Kuhle J, Bernardi F, Weinstock-Guttman B, et al. High-density lipoprotein cholesterol is associated with multiple sclerosis fatigue: A fatigue-metabolism nexus? J Clin Lipidol [Internet]. 2019 [cited 2019 Dec 29];13:654-663.e1. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1933287419302156
34. Høgh P, Garde E, Mortensen EL, Jørgensen OS, Krabbe K, Waldemar G. The apolipoprotein E ε4-allele and antihypertensive treatment are associated with increased risk of cerebral MRI white matter hyperintensities. Acta Neurol Scand [Internet]. 2007 [cited 2020 Feb 3];115:248–53. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17376122

35. Burwick RM, Ramsay PP, Haines JL, Hauser SL, Oksenberg JR, Pericak-Vance MA, et al. APOE epsilon variation in multiple sclerosis susceptibility and disease severity: Some answers. Neurology [Internet]. 2006 [cited 2020 Feb 3];66:1373–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16682670
36. Hùgh P, Oturai A, Schreiber K, Blinkenberg M, Jùrgensen OS, Ryder L, et al. Apoliprotein E and multiple sclerosis: impact of the epsilon-4 allele on susceptibility, clinical type and progression rate. Mult Scler J [Internet]. 2000 [cited 2020 Feb 3];6:226–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10962542
37. Pavelek Z, Vyšata O, Tambor V, Pimková K, Vu DL, Kuča K, et al. Proteomic analysis of cerebrospinal fluid for relapsing-remitting multiple sclerosis and clinically isolated syndrome. Biomed Reports [Internet]. Spandidos Publications; 2016 [cited 2020 Jan 7];5:35–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27347402
38. Watanabe M, Nakamura Y, Michalak Z, Isobe N, Barro C, Leppert D, et al. Serum GFAP and neurofilament light as biomarkers of disease activity and disability in NMOSD. Neurology [Internet]. 2019 [cited 2019 Dec 31];93:e1299–311. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31471502
39. Uzawa A, Mori M, Sawai S, Masuda S, Muto M, Uchida T, et al. Cerebrospinal fluid interleukin-6 and glial fibrillary acidic protein levels are increased during initial neuromyelitis optica attacks. Clin Chim Acta [Internet]. 2013 [cited 2019 Dec 31];421:181–3. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23535508
40. Wei Y, Chang H, Li X, Wang H, Du L, Zhou H, et al. Cytokines and Tissue Damage Biomarkers in First-Onset Neuromyelitis Optica Spectrum Disorders: Significance of Interleukin-6. Neuroimmunomodulation [Internet]. 2019 [cited 2019 Dec 31];25:215–24. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30544111
41. Cummings DM, Benway TA, Ho H, Tedoldi A, Fernandes Freitas MM, Shahab L, et al. Neuronal and Peripheral Pentraxins Modify Glutamate Release and may Interact in Blood-Brain Barrier Failure. Cereb Cortex [Internet]. 2017 [cited 2019 Dec 31];27:3437–48. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28334103
42. Chen X, Rivard L, Naqvi S, Nakada S, Padbury JF, Sanchez-Esteban J, et al. Expression and localization of Inter-alpha Inhibitors in rodent brain. Neuroscience [Internet]. 2016 [cited 2019 Dec 31];324:69–81. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0306452216002323

Additional Material

Additional file 1 (.tif):Experimental overview

(A) 169 CSF samples were collected from healthy controls, MS subtypes, and disease controls with different degree of neurodegeneration and inflammation. CSF samples from each group were pooled into three technical replicates, prepared for peptide solution, labelled with iTRAQ 8plex, and distributed into three sets (a, b, c) that were analyzed by LC-MS/MS. Raw data was preprocessed, identified proteins were quantified and statistical analyses were performed. The data was examined and visualized with

bioinformatic tools. **(B)** Based on different statistical analyses (ANOVA, limma, complementary analysis- see Methods), 299 proteins were selected for quantification in individual samples. **(C)** 170 CSF samples from PPMS (n=30), relapsing MS (n=14), remission MS (n=33), SPMS (n= 26), AD (n=22), NMOSD AQP4-IgG^{+} (n=13), NMOSD AQP4-IgG^{-} (n=5) and healthy (n=27) were used to quantify the 299 proteins in each individual CSF by mass spectrometry labelled with 11TMT plex spiked with SIS, and analysed by ANOVA. Proteins were grouped based on their specific regulation in disease groups, and their gene expression was also examined in different brain WM lesion types of progressive MS by using MS-Atlas.

CSF: cerebrospinal fluid; PP/SPMS: primary/secondary progressive multiple sclerosis; AD: Alzheimer disease; NMOSD AQP4-IgG$^{+/-}$: neuromyelitis optica spectrum disorder positive/negative for immunoglobulin G antibody against aquaporin-4; LC-MS/MS: Liquid chromatography tandem mass spectrometry; iTRAQ: Isobaric tag for relative and absolute quantitation; TMT: Tandem Mass Tag; SIS: stable isotope standards; OND: other neurological diseases.

Additional file 2 (.tif): Plate design of the two mass spectrometry experiments

(A) Distribution of samples between the three iTRAQ 8plex sets.

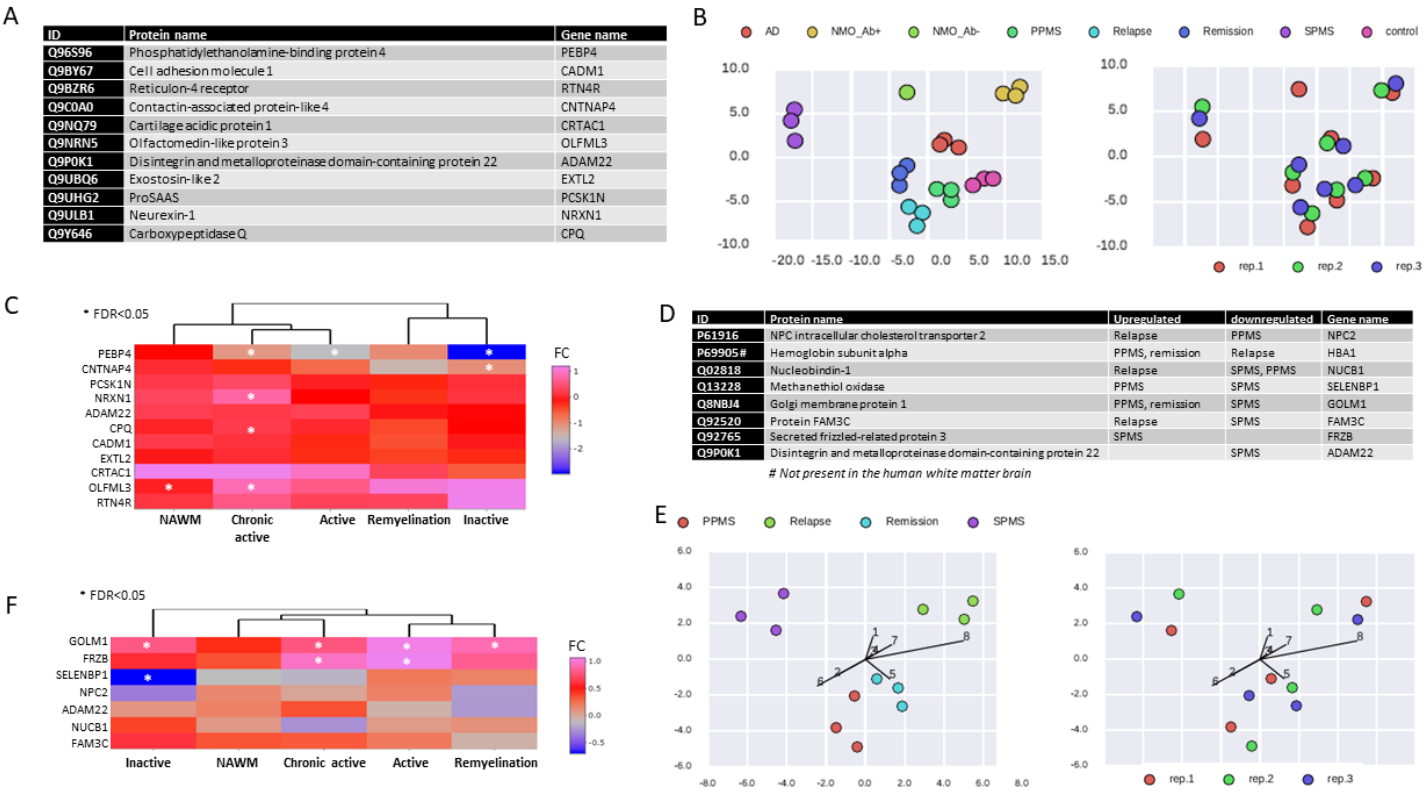
(B) Patient samples randomly distributed across 17 TMT sets. Colours represent different disease/control groups. The pooled group is a mix of all the samples.

PP/SP: primary/secondary progressive multiple sclerosis; AD: Alzheimer disease; NMOSD +/- and NMO Ab+, Ab-: neuromyelitis optica spectrum disorder serum positive/negative for immunoglobulin G antibody against aquaporin-4; iTRAQ: isobaric tag for relative and absolute quantitation; TMT: Tandem Mass Tag

Additional file 3 (.pdf): Functional analyses of MS subtypes, NMOSD and Alzheimers Disease. Comprehensive pathway analyses of the proteomic landscape in each disease subtype normalized to healthy control CSF.

Additional file 4 (.pdf): The 299 proteins selected from the three different statistical tests for quantification in individual CSF samples. The table consists three columns: the type of statistical test(s), the number of proteins and each protein ID.

Figures



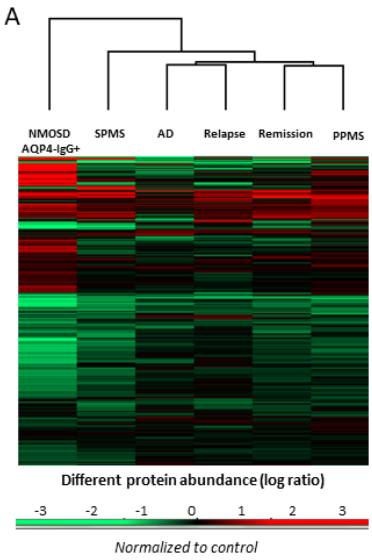
ID	Protein name	Gene name
Q96S96	Phosphatidylethanolamine-binding protein 4	PEBP4
Q9BY67	Cell adhesion molecule 1	CADM1
Q9BZR6	Reticulon-4 receptor	RTN4R
Q9C0A0	Contactin-associated protein-like 4	CNTNAP4
Q9NQ79	Cartilage acidic protein 1	CRTAC1
Q9NRN5	Olfactomedin-like protein 3	OLFML3
Q9P0K1	Disintegrin and metalloproteinase domain-containing protein 22	ADAM22
Q9UBQ6	Exostosin-like 2	EXTL2
Q9UHG2	ProSAAS	PCSK1N
Q9ULB1	Neurexin-1	NRXN1
Q9Y646	Carboxypeptidase Q	CPQ

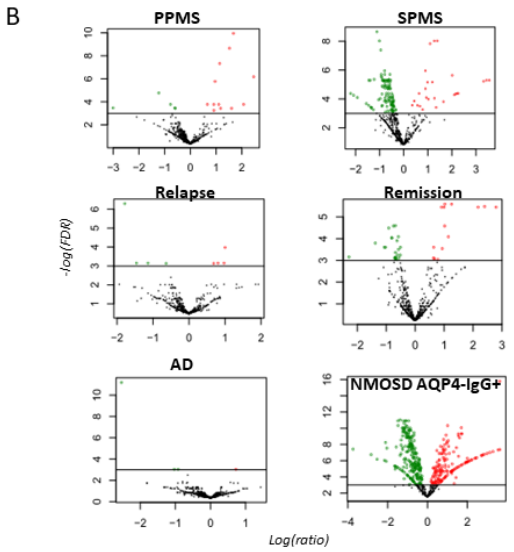
ID	Protein name	Upregulated	downregulated	Gene name
P61916	NPC intracellular cholesterol transporter 2	Relapse	PPMS	NPC2
P69905#	Hemoglobin subunit alpha	PPMS, remission	Relapse	HBA1
Q02818	Nucleobindin-1	Relapse	SPMS, PPMS	NUCB1
Q13228	Methanethiol oxidase	PPMS	SPMS	SELENBP1
Q8NBJ4	Golgi membrane protein 1	PPMS, remission	SPMS	GOLM1
Q92520	Protein FAM3C	Relapse	SPMS	FAM3C
Q92765	Secreted frizzled-related protein 3	SPMS		FRZB
Q9P0K1	Disintegrin and metalloproteinase domain-containing protein 22		SPMS	ADAM22

 # Not present in the human white matter brain

Figure 1

Protein combinations discriminating diseases and disease subtypes. (A) Combination of the 11 proteins listed were able to discriminate PPMS, MS in relapse, MS in remission, SPMS, AD, AQP4-IgG+ NMOSD, AQP4-IgG- NMOSD, and healthy controls. (B) Disease-specific discrimination by using linear discriminant analysis (LDA). (C) Hierarchical clustering of expression of genes encoding the 11 proteins in different lesions in the PMS brain white matter. Stars represent significantly differentially expressed genes (FDR<0.05) compared to non-neurological disease brains. Colour represents the log2 fold change (FC). (D) Combination of the 8 proteins listed were able to discriminate among MS subtypes. (E) LDA classifier showing discrimination between MS patients in relapse, remission, and with PPMS and SPMS. (F) Hierarchical clustering of brain lesion expression of genes encoding the 8 compound proteins that differentiate among MS subtypes. Stars represent significantly differentially expressed genes (FDR<0.05) compared to non-neurological disease brains. Colour represents the log2 fold change (FC). CSF: cerebrospinal fluid; PP/SPMS: primary/secondary progressive multiple sclerosis; AD: Alzheimer disease; NMO Ab+/-: neuromyelitis optica spectrum disorder positive/negative for immunoglobulin G antibody against aquaporin-4; FDR: False discovery rate





C

Pathway	Shared
LXR/RXR Activation	SPMS, PPMS, Remission, AQP4-IgG+ NMOSD
Acute Phase Response Signalling	SPMS, AQP4-IgG+,NMOSD AD
Axonal Guidance Signalling	Remission, PPMS, AD
Intrinsic Prothrombin Activation Pathway	PPMS, AD
Pathway	**unique**
Neuroprotective Role Of THORP1 In AD	SPMS
Coagulation System	SPMS
FXR/RXR Activation	PPMS
Clathrin-mediated Endocytosis Signalling	Remission
Hematopoiesis from Pluripotent Stem Cells	Relapse
Leucocyte Extravasation Signalling	Relapse
Agrin Interactions at Neuromuscular Junction	Relapse

D

Top 5 predicted diseases or functions

Relapse	Remission	SPMS	PPMS	NMOSD AQP4-IgG+	AD
Cellular Movement	Cardiovascular System Development and Function	Metabolic Disease	Cellular Movement	Cellular Movement	Cellular Movement
Connective Tissue Disorders	Metabolic Disease	Cellular Movement	Metabolic Disease	Metabolic Disease	Metabolic Disease
Skeletal and Muscular Disorders	Cellular Movement	Neurological Disease	Neurological Disease	Neurological Disease	Neurological Disease
Cell Morphology	Neurological Disease	Psychological Disorders	Psychological Disorders	Psychological Disorders	Psychological Disorders
Inflammatory Disease	Psychological Disorders	*Cell growth and proliferation*	Cardiovascular System Development and Function	Skeletal and Muscular Disorders	*Immune Cell Trafficking*

Figure 2

Functional analyses of the CSF proteome of disease groups normalized to healthy controls. (A) Heatmap clustering of PPMS, MS in relapse, MS in remission, SPMS, AD, and AQP4-IgG+ NMOSD versus healthy control based on the different abundance of the protein levels. Red colour represents upregulation, while green represents downregulation in disease groups compared to healthy control. (B) Volcano plots of differentially expressed proteins in each disease group compared to healthy controls. Each point represents the average value of one protein in three replicate experiments. The dark horizontal line is set when the protein expression difference is significant with FDR<0.1. (C) The top scoring canonical pathways for the different disease groups using Ingenuity Pathway Analysis (IPA) and z-score algorithms. (D) The top 5 predicted biological functions for each of the disease group compared to healthy control. Italic indicates a unique function for a disease group. CSF: cerebrospinal fluid; PP/SPMS: primary/secondary progressive multiple sclerosis; AD: Alzheimer disease; AQP4-IgG+ NMOSD: neuromyelitis optica spectrum disorder serum positive for immunoglobulin G antibody against aquaporin-4; FDR: False discovery rate;

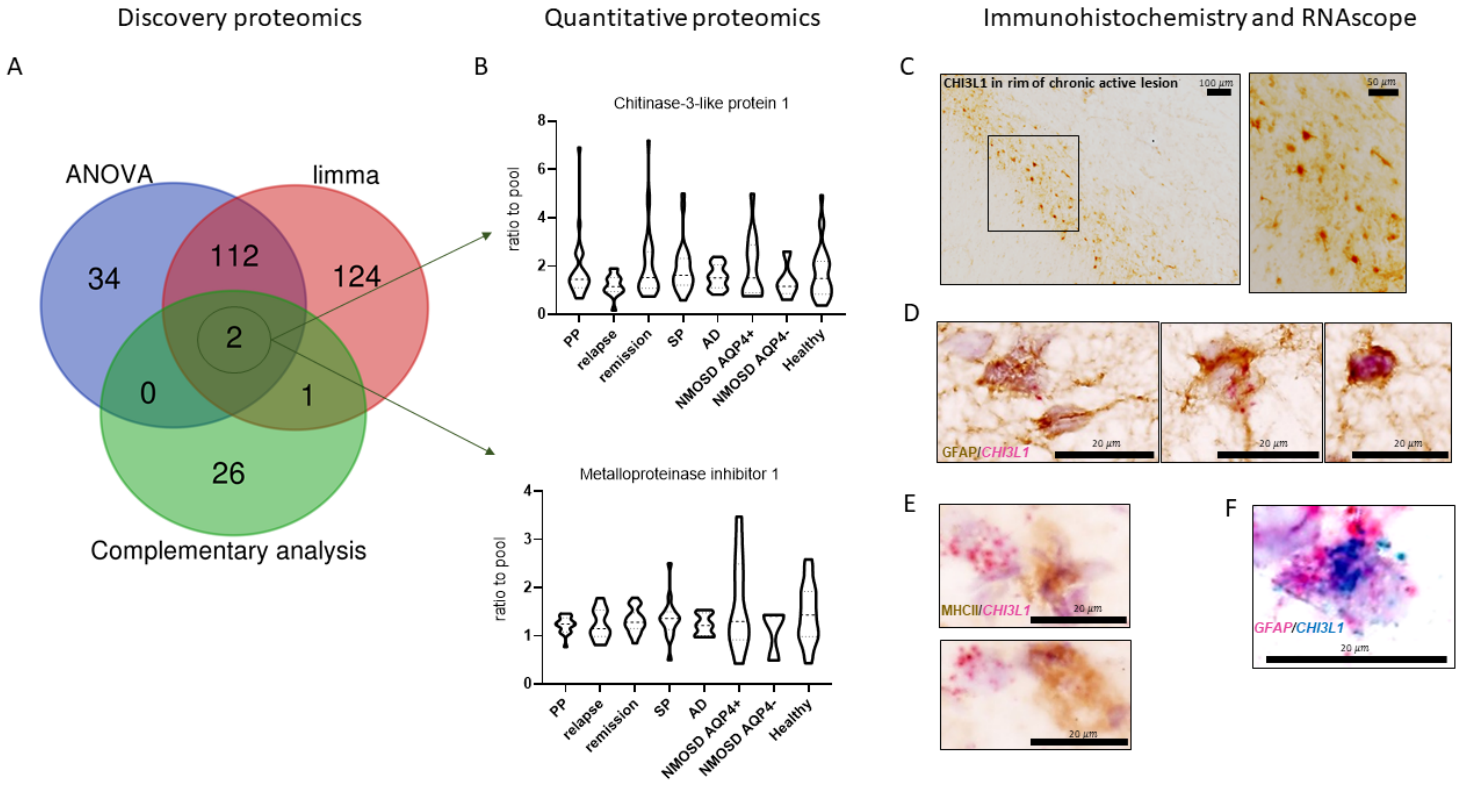


Figure 3

Chitinase-3-like protein 1 (CHI3L1) and metalloproteinase inhibitor 1 (TIMP-1). (A) Venn diagram showing the 299 proteins from discovery CSF proteomics significantly altered compared to healthy by different statistical analyses (ANOVA, limma, complementary analysis, see Methods). (B) Quantitative levels of chitinase-3-like protein 1 (CHI3L1) and metalloproteinase inhibitor 1 (TIMP-1) in individual CSF samples (detection of CHI3L1: PP=30 of 30; relapse=14 of 14; remission=33 of 33; SP=26 of 26; AD=22 of 22; AQP4+ NMOSD =13 of 13; AQP4- NMOSD =5 of 5; healthy=27 of 27) (detection of TIMP-1: PP=19 of 30; relapse=10 of 14; remission=27 of 33; SP=16 of 26; AD=13 of 22; AQP4+ NMOSD =10 of 13; AQP4- NMOSD =3 of 5; healthy=16 of 27). (C) Protein expression of CHI3L1 in the rim of a chronic active lesion of progressive MS brain. (D) Protein expression of GFAP (brown) and RNA expression of CHI3L1 (red) in the same cells (combined immunohistochemistry and RNAscope). (E) Protein expression of MHCII (brown) and RNA expression of CHI3L1 (red) in different cells close to each other at the rim of the lesion (combined immunohistochemistry and RNAscope). (F) Co-localized RNA expression of GFAP (red) with CHI3L1 (green) by RNAscope. PP/SP: primary/secondary progressive multiple sclerosis; AD: Alzheimer disease; NMOSD AQP4+/-: neuromyelitis optica spectrum disorders serum positive/negative for immunoglobulin G antibody against aquaporin-4; AD: Alzheimer disease

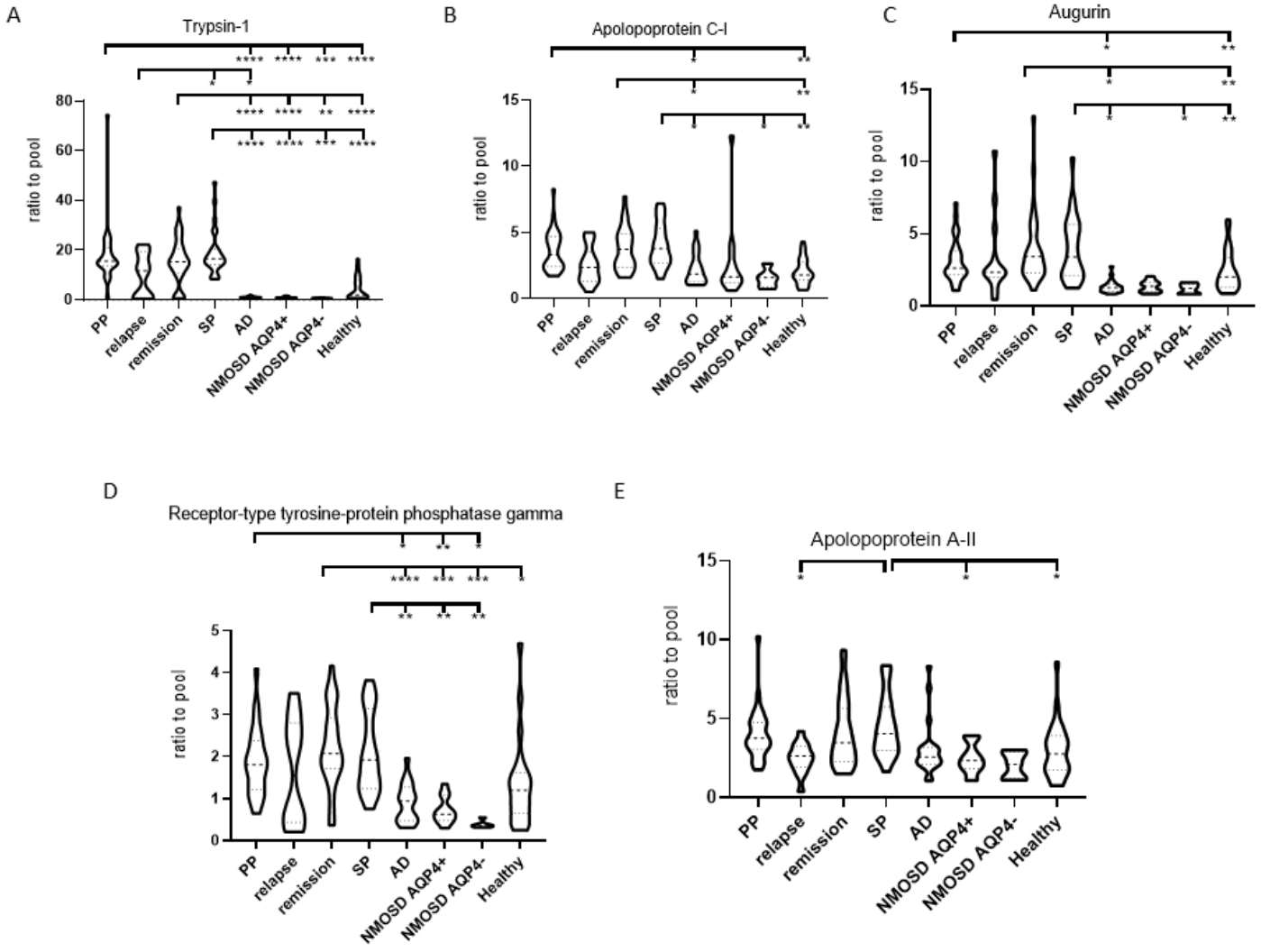


Figure 4

Molecular markers in the MS CSF. Overview of 5 proteins significantly upregulated in the CSF of MS subtypes compared to OND and healthy controls. (A) Trypsin-1 that was significantly upregulated in remission, PPMS and SPMS compared to disease- and healthy controls in individual samples (detection: PP=30 of 30; relapse=14 of 14; remission=33 of 33; SP=26 of 26; AD=22 of 22; NMOSD AQP4+=13 of 13; NMOSD AQP4- =5 of 5; healthy=27 of 27). (B) Apolipoprotein C-I and (C) augurin were significantly upregulated in remission, PPMS and SPMS compared to AD and healthy controls in individual samples. (detection: PP=30 of 30; relapse=14 of 14; remission=33 of 33; SP=26 of 26; AD=22 of 22; NMOSD AQP4+=13 of 13; NMOSD AQP4- =5 of 5; healthy=27 of 27). (D) Receptor-type tyrosine-protein phosphatase gamma was significantly upregulated in remission, PPMS and SPMS compared to disease controls (detection: PP=25 of 27; relapse=12 of 14; remission=27 of 33; SP=20 of 26; AD=18 of 22; NMOSD AQP4+=11 of 13; NMOSD AQP4- =5 of 5; healthy=22 of 27). (E) Apolipoprotein A-II was uniquely significantly upregulated in SPMS compared to MS in relapse, NMOSD AQP4+ and healthy controls (detection: PP=30 of 30; relapse=14 of 14; remission=33 of 33; SP=26 of 26; AD=22 of 22; NMOSD AQP4+=13 of 13; NMOSD AQP4- =5 of 5; healthy=27 of 27). PP/SP: primary/secondary progressive

multiple sclerosis; AD: Alzheimer disease; NMOSD AQP4+/-: neuromyelitis optica spectrum disorder positive/negative for immunoglobulin G antibody against aquaporin-4; AD: Alzheimer disease

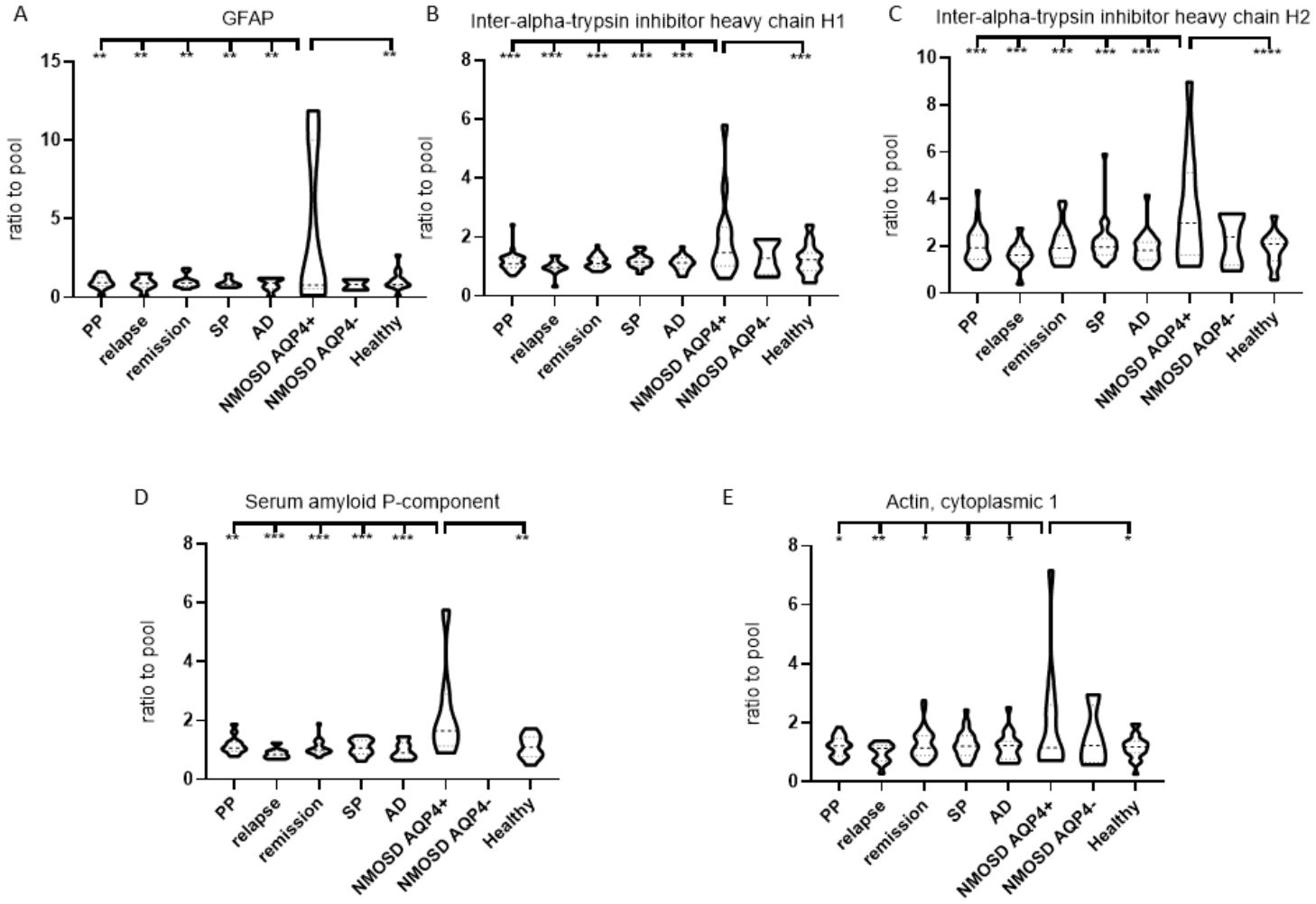


Figure 5

AQP4-IgG+ NMOSD specific molecular markers Five proteins were uniquely upregulated in the CSF of patients with AQP4-IgG+ NMOSD compared to OND and healthy controls. (A) Glial fibrillary acidic protein (GFAP) (detection: PP=12 of 30; relapse=8 of 14; remission=13 of 33; SP=10 of 26; AD=8 of 22; NMOSD AQP4-IgG+=9 of 13; NMOSD AQP4-IgG- =2 of 5; healthy=23 of 27). (B) Inter-alpha-trypsin inhibitor heavy chain H1 (detection: PP=30 of 30; relapse=14 of 14; remission=33 of 33; SP=26 of 26; AD=22 of 22; NMOSD AQP4-IgG+=13 of 13; NMOSD AQP4-IgG- =5 of 5; healthy=27 of 27). (C) Inter-alpha-trypsin inhibitor heavy chain H2 (detection: PP=30 of 30; relapse=14 of 14; remission=33 of 33; SP=26 of 26; AD=22 of 22; NMOSD AQP4-IgG+=13 of 13; NMOSD AQP4-IgG- =5 of 5; healthy=27 of 27). (D) Serum amyloid P-component (detection: PP=13 of 30; relapse=7 of 14; remission=16 of 33; SP=14 of 26; AD=10 of 22; NMOSD AQP4-IgG+=7 of 13; NMOSD AQP4-IgG- =1 of 5; healthy=12 of 27). (E) Actin, cytoplasmic 1 (detection: PP=26 of 30; relapse=13 of 14; remission=30 of 33; SP=23 of 26; AD=19 of 22; NMOSD AQP4-IgG+=12 of 13; NMOSD AQP4-IgG- =4 of 5; healthy=23 of 27). CSF: cerebrospinal fluid; NMOSD AQP4-

IgG+/-: neuromyelitis optica spectrum disorder positive for immunoglobulin G antibody against aquaporin-4; AD: Alzheimer disease

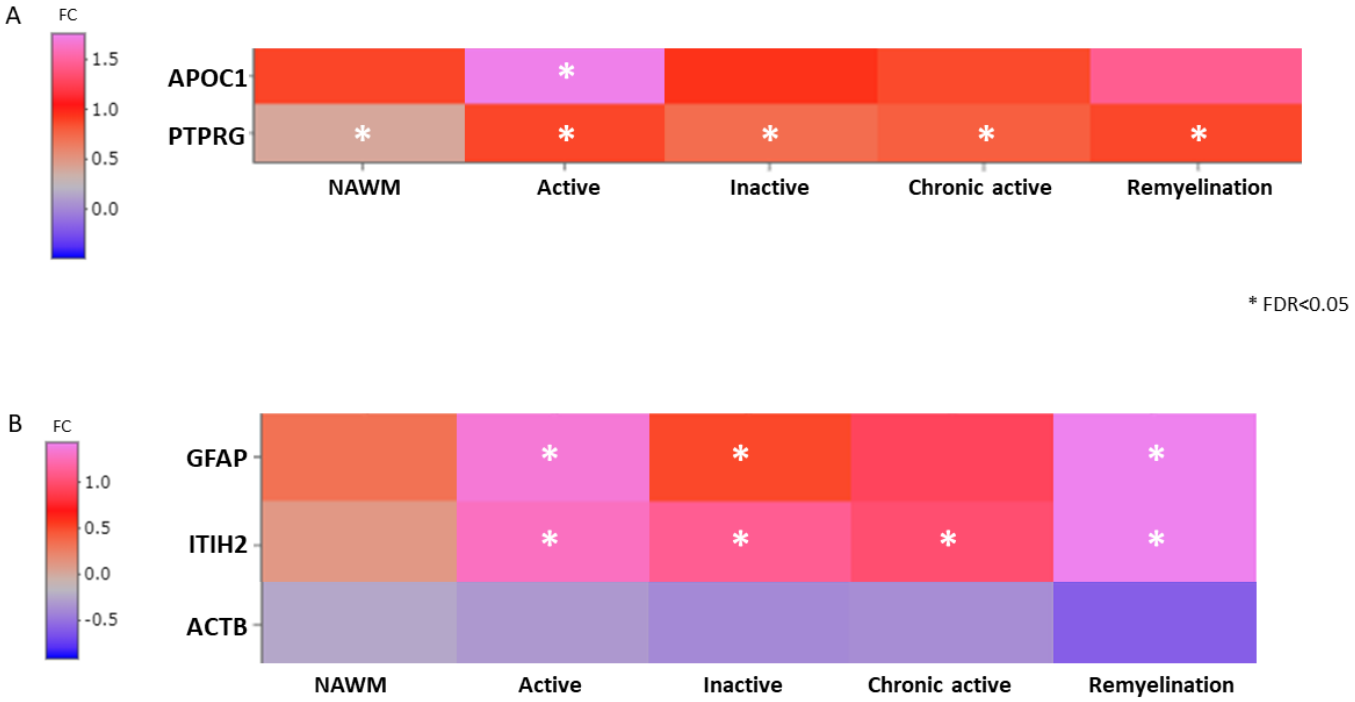


Figure 6

Expression of genes of altered disease-specific CSF proteins in transcriptomes of different MS brain lesions. The heatmaps show genes encoding the significantly altered CSF proteins that could be detected in different the brain lesions and normal-appearing white matter (NAWM) of MS. (A) The transcripts APOC1 and PTPRG encoding two MS-specific CSF proteins: apolipoprotein C-I and receptor-type tyrosine-protein phosphatase gamma. (B) The transcripts GFAP, ITH2 and ACTB encoding the NMOSD-specific CSF proteins: glial fibrillary acidic protein (GFAP), inter-alpha-trypsin inhibitor heavy chain H2, and actin cytoplasmic 1. Stars represent significantly differentially expressed genes (FDR<0.05) in the MS lesions compared to non-neurological disease brains. Colour represents the log2fold change (FC).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.tif
- Additionalfile4.pdf
- Additionalfile3.pdf
- additionalfile2.tif