Discovery of Biomarkers for Amyotrophic Lateral Sclerosis and Frontotemporal Lobar Degeneration From Human Cerebrospinal Fluid Using Mass Spectrometry-Based Proteomics

Sungtaek Oh
Johns Hopkins University
Yura Jang
Johns Hopkins University
Lawren VandeVrede
UCSF Memory and Aging Center: University of California San Francisco Memory and Aging Center
Adam L Boxer
UCSF Memory and Aging Center: University of California San Francisco Memory and Aging Center
Shanthini Sockanathan
Johns Hopkins University
Chan-Hyun Na (chanhyun@jhmi.edu)
Johns Hopkins University https://orcid.org/0000-0002-3622-2938

Research

Keywords: Amyotrophic lateral sclerosis, frontotemporal lobar degeneration, cerebrospinal fluid, mass spectrometry, parallel reaction monitoring

Posted Date: January 19th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-148356/v1

Discovery of biomarkers for amyotrophic lateral sclerosis and frontotemporal lobar degeneration from human cerebrospinal fluid using mass spectrometry-based proteomics

Sungtaek Oh[1]*, Yura Jang[1]*, Lawren VandeVrede[3], Adam L Boxer[3], Shanthini Sockanathan[2#], Chan Hyun Na[1#]

[1]Department of Neurology, Institute for Cell Engineering, Johns Hopkins University School of Medicine, 733 N. Broadway, MRB 753, Baltimore, MD 21205, USA.
[2]The Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, 725 N. Wolfe Street, PCTB 1004, Baltimore, MD 21205, USA.

[3]Memory and Aging Center, Department of Neurology, University of California, San Francisco, 675 Nelson Rising Lane, Suite 190, San Francisco, CA 94158, USA

*These authors contributed equally.

[#]Correspondence should be addressed to: Chan Hyun Na (chanhyun@jhmi.edu) and Shanthini Sockanathan (ssockan1@jhmi.edu.)

ABSTRACT

Background: Amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) are progressive neurodegenerative diseases that share clinical and neuropathologic features. Critical to the mission of developing effective therapies for ALS and FTLD is the discovery of biomarkers that can illuminate shared mechanisms of neurodegeneration, which can then be evaluated for diagnostic, prognostic or pharmacodynamic value across the disease spectrums.

Methods: Here, we merged unbiased discovery-based approaches and targeted quantitative comparative analyses between ALS and FTLD cerebrospinal fluid (CSF) to identify proteins that are altered in ALS and FTLD.

Results: Discovery mass spectrometry (MS)-based proteomic approaches combined with tandem mass tags (TMT) quantification methods from 40 CSF samples comprising 20 patients with ALS and 20 healthy control (HC) individuals identified 19 differentially expressed candidate biomarker proteins after CSF fractionation. Notably, these candidate biomarkers included novel and previously identified proteins, thus validating our approach. Candidate biomarkers were subsequently examined using parallel reaction monitoring (PRM) MS methods on 80 unfractionated CSF samples comprising 30 patients with ALS, 19 patients with FTLD, and 31 HC individuals. Two candidate biomarkers (CNTNAP2 and CLSTN1) were downregulated in both ALS and FTLD compared to healthy controls, and 11 further candidate proteins were significantly downregulated in FTLD compared to HC.

Conclusions: Taken together, this study identifies multiple novel proteins that are altered in ALS and FTLD, which provides the foundation for their evaluation and development as biomarkers for these diseases.

KEYWORDS

Amyotrophic lateral sclerosis, frontotemporal lobar degeneration, cerebrospinal fluid, mass spectrometry, parallel reaction monitoring

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disease characterized by the relentless loss of upper and lower motor neurons, eventually leading to death[1-5]. An average of 5,600 people are annually diagnosed with ALS in the United States, and ~2 per 100,000 people die of this disease annually in Europe [6-8]. In addition to motor dysfunction, a significant proportion of ALS patients also develop cognitive and behavioral symptoms seen in frontotemporal lobar degeneration (FTLD) related clinical syndromes, due to accumulation of transactive response (TAR) DNA binding protein (TDP), which accounts for approximately half of all FTLD cases as well [9, 10]. Approximately half of FTLD cases have underlying tau pathology instead of TDP. Recent studies have identified that an expanded hexanucleotide (GGGGCC) repeat insertion into the non-coding region of *C9orf72* is the most common genetic cause of both ALS and FTLD [11], and these discoveries have led to the widely accepted view that ALS and FTLD share neurodegenerative pathways and lie on a disease spectrum [12]. However, the shared pathogenic mechanisms of ALS and FTLD remain elusive, and, in the absence of an identified gene mutation, the diagnostic approach is largely based on clinical criteria [13, 14]. A key challenge for designing clinical trials for many sporadic FTLD syndromes is identifying whether the underlying pathology is due to FTLD-TDP or FTLD-tau during life since clinical syndrome often does not predict underlying pathology. In order to develop effective treatments for ALS and FTLD, there is a pressing need to improve early diagnostic capabilities to facilitate earlier therapeutic interventions prior to the manifestation of clinical symptoms and to develop prognostic measurements of disease progression.

Recently, there have been major efforts dedicated to identifying biomarkers for ALS and FTLD, and chief among these are quantitative methods to measure proteins in patient biofluids such as blood, urine, saliva and cerebrospinal fluid (CSF). Mass spectrometry (MS)-based proteomics have been considered the gold standard for protein discovery and there have been multiple attempts to develop protein biomarkers for ALS using proteomics technology [5, 6, 15-19]. So far, multiple proteins, including CHIT1, CHI3L1, CHI3L2, UCHL1, MAP2, GPNMB and neurofilament proteins, have been reported as candidate ALS biomarkers [20-23]. Many of these proteins await rigorous validation, but a few such as neurofilament are also altered in FTLD and have extremely promising clinical applications [24]. Clearly, the discovery of additional novel biomarkers is needed, but one major barrier is that many key regulatory proteins are of relatively low abundance [25], and therefore changes in protein levels are easily masked. Accordingly, the identification of useful biomarkers for ALS and FTLD requires the use of innovative, highly sensitive, and accurate methodologies.

Herein, we report findings from an initial unbiased proteomics discovery study on fractionated CSF from 20 patients with ALS and 20 healthy controls (HC). Tandem-mass-tag (TMT) technology was

used for the accurate and sensitive quantification of CSF proteins, with subsequent analysis via state-of-the-art Orbitrap Fusion Lumos ETD mass spectrometry. Several proteins from the discovery phase confirmed previously identified biomarkers for ALS, validating our approach, and many new candidates were identified. Using proteins identified from the discovery phase, a subsequent targeted, quantitative, and highly sensitive analysis was conducted using parallel reaction monitoring (PRM) targeted mass spectrometry in 80 unfractionated CSF samples from 30 patients with ALS, 19 patients with either known or predicted FTLD-tau pathology and 31 HC individuals. Several novel candidate biomarkers were identified that might be of use in better understanding shared disease mechanisms and potentially differentiating ALS and FTLD from healthy controls. Further, this work highlights that our innovative mass spectrometry approach has iterative potential in other neurodegenerative diseases.

METHODS

Collection of CSF samples

CSF samples from 11 HC individuals and 30 patients with ALS were obtained from the Northeast Amyotrophic Lateral Sclerosis (NEALS) consortium. 19 additional CSF samples from HC individuals and 19 patients with a diverse range of FTLD-related clinical syndromes were provided by the Advancing Research and Treatment for Frontotemporal Lobar Degeneration Research Consortium (ARTFL) at the University of California San Francisco (UCSF). FTLD samples included the following clinical phenotypes: behavioral variant frontotemporal dementia (bvFTD, n = 6), progressive supranuclear palsy, Richardson's syndrome (PSP-RS, n = 8), corticobasal syndrome (CBS, n = 3), and the non-fluent variant of primary progressive aphasia (nfvPPA, n = 2). Eighteen participants had either FTLD confirmed on autopsy or a known FTLD-causing genetic mutation (i.e. microtubule associated protein tau, *MAPT*), leading to high confidence in FTLD as causative in these cases, with the exception of one case with bvFTD where the clinical suspicion for FTLD was high. Lumbar punctures were performed using the atraumatic technique and collected in a polypropylene tube before transferring to a 50 ml conical polypropylene tube at room temperature (RT), which was mixed gently by inverting 3–4 times. Within 15 min of collection, CSF was centrifuged at 2,000 x g for 10 min at RT and aliquoted directly into pre-cooled polypropylene cryovials. Within 60 min of CSF collection, aliquots were frozen on dry ice and then stored at –80°C, until further analysis. For samples collected prior to 2015, the protocol used is described by Scherling CS *et al*. [26]. Study participants provided written informed consent, and all procedures were approved by the UCSF Institutional Review Board (IRB). The demographics of CSF samples used in this study are provided in Table 1.

Sample preparation and trypsin digestion for discovery experiments

Four experimental sets of 10 samples were examined, with each set including a master pool (MP) sample for normalization between sets. The MP was prepared by combining an equal volume from all 40 CSF samples including HC and ALS (Figure 1A). The CSF samples were mixed with a urea buffer, which was composed of 10 M urea/20 mM tris (2-Carboxyethyl) phosphine hydrochloride (TCEP)/80 mM chloroacetamide (CAA) in 100 mM triethylammonium bicarbonate (TEAB), at a one-to-one ratio. The samples were then incubated for 1 h at RT for reduction and alkylation. Protein digestion was carried out using LysC (lysyl endopeptidase mass spectrometry grade, Fujifilm Wako Pure Chemical Industries Co., Ltd., Osaka, Japan) at a one-to-fifty ratio for 3 h at 37°C and subsequently with trypsin digestion (sequencing grade modified trypsin, Promega, Fitchburg, WI, USA) at a one-to-fifty ratio at 37°C overnight after diluting the concentration of urea from 5 M to 2 M by adding 50 mM TEAB. Peptides were desalted using C_{18} StageTips (3M Empore™; 3M, St. Paul, MN, USA) after acidifying with 1% trifluoroacetic acid (TFA) to the final concentration. The eluted solution containing peptides was dried with a Savant SPD121P SpeedVac concentrator (Thermo Scientific) and then stored at –80°C before use.

To perform TMT -based quantitative mass spectrometry, the digested peptides from CSF samples were labeled using 11-plex TMT reagents following the manufacturer's instructions (Thermo Fisher Scientific). The MP sample was labeled with 131C, and CSFs from ALS and HC individuals were labeled with the rest of the TMT tags. The labeling reaction was performed for 1 h at RT after mixing each peptide in 100 mM TEAB with TMT reagent in acetonitrile (ACN, HPLC grade), and then quenched by adding with 1/10 volume of 1 M Tris-HCl (pH 8.0). The TMT labeled peptides were pooled, resuspended with 10 mM TEAB, and then subjected to bRPLC (basic pH reversed-phase liquid chromatography) fractionation to generate fractions on an Agilent 1260 offline HPLC system (Agilent Technologies, Santa Clara, CA, USA), which includes a binary pump, variable wavelength detector, an autosampler, and an automatic fraction collector. The pooled samples were reconstituted in solvent A (10 mM TEAB, pH 8.5) and loaded onto Agilent 300 Extend-C_{18} column (5 μm, 4.6 mm × 250 mm, Agilent Technologies). Peptides were resolved using a gradient of solvent B (10 mM TEAB in 90% ACN, pH 8.5) at a flow rate of 0.3 mL/min over 90 minutes, collecting 96 fractions. Subsequently, the fractions were concatenated into 24 fractions followed by vacuum drying using a SpeedVac (Thermo Fisher Scientific, San Jose, CA, USA). The dried peptides were suspended in 0.5% formic acid (FA), and 30% of each fraction was injected for mass spectrometry analysis.

LC-MS/MS analysis for the discovery of biomarker candidates

The peptide samples were analyzed on an Orbitrap Fusion Lumos Tribrid mass spectrometer interfaced

with an Ultimate 3000 RS Autosampler nanoflow liquid chromatography system (Thermo Fisher Scientific). The dried 24 fractionated peptides were reconstituted in 0.5% FA and then loaded onto a trap column (Acclaim™ PepMap™ 100 LC C_{18}, 5 μm, 100 μm × 2 cm, Thermo Fisher Scientific) at a flow rate of 8 μL/min. Peptides were separated on an analytical column (Easy-Spray™ PepMap™ RSLC C_{18}, 2 μm, 75 μm × 50 cm, Thermo Fisher Scientific) at a flow rate of 0.3 μL/min using a linear gradient with mobile phases consisted of 0.1% FA in water and in ACN. The total run time was 120 min. The mass spectrometer was operated in a data-dependent acquisition mode. The MS1 (precursor mass) scan range for a full survey scan was acquired from 300 to 1,800 *m/z* (mass-to-charge ratio) in the "top speed" setting with a resolution of 120,000 at an *m/z* of 200. The AGC target for MS1 was set as 1×10^6 and the maximum injection time was 50 ms. The most intense ions with charge states of 2 to 5 were isolated in a 3-sec cycle, fragmented using higher-energy collisional dissociation (HCD) fragmentation with 35% normalized collision energy, and detected at a mass resolution of 50,000 at an *m/z* of 200. The AGC target for MS/MS (MS2, fragment mass) was set as 5×10^4 and the ion filling time was 100 ms. The precursor isolation window was set to 1.6 *m/z* with a 0.4 *m/z* offset. The dynamic exclusion was set for 30 sec, and singly charged ions were rejected. Internal calibration was carried out using the lock mass option (*m/z* 445.1200025) from ambient air.

Database searches for peptide and protein identification

The acquired tandem mass spectrometry data were searched against the human UniProt database (released in May 2018, containing protein entries with common contaminants) using the SEQUEST search algorithm through the Thermo Proteome Discoverer platform (version 2.2.0.388, Thermo Fisher Scientific) for quantitation and identification. During MS/MS preprocessing, the top 10 peaks in each window of 100 *m/z* were selected for database searches. The search parameters included two maximum missed-cleavage sites by trypsin as a proteolytic enzyme. Carbamidomethyl (+57.02146 Da) at cysteine and TMT reagents (+229.162932 Da) modification at N-terminus of peptide and lysine residues were set as fixed modifications while oxidation (+15.99492 Da) of methionine was a variable modification. For MS data, MS1 error tolerance was set to 10 ppm and the MS/MS error tolerance to 0.02 Da. The minimum peptide length was set to 6 amino acids, and proteins identified by one peptide were filtered out. Both peptides and proteins were filtered at a 1% false discovery rate. The protein quantification was performed with the following parameters and methods. The most confident centroid option was used for the integration mode while the reporter ion tolerance was set to 20 ppm. MS order was set to MS2. The activation type was set to HCD. The quantification value correction was disabled. Both unique and razor peptides were used for peptide quantification. Protein groups were considered for peptide uniqueness. Missing intensity values were replaced with the minimum value. Reporter ion

abundance was computed based on the signal-to-noise ratio. Quantification value corrections for isobaric tags were disabled. The co-isolation threshold was set to 50%. The average reporter signal-to-noise threshold was set to 50. Data normalization was disabled. Protein grouping was performed by applying strict parsimony principle as following; 1) all proteins that share the same set or subset of identified peptides were grouped, 2) protein groups that have no unique peptides among the considered peptides were filtered out, 3) Proteome Discoverer iterated through all spectra and selected which peptide-spectrum match (PSM) to use in ambiguous cases to make a protein group with the highest number of unambiguous and unique peptides, and 4) final protein groups were generated. The Proteome Discoverer summed all the reporter ion abundances of PSMs for the corresponding proteins in a TMT run.

Statistical and bioinformatic analyses of the results from discovery proteomics

Bioinformatics analysis was processed with the Perseus software package (version 1.6.0.7). Each set was divided by the MP included in each set followed by dividing each column by a column median. After $\log_2$-transformation of all the values, each column was z-score-transformed. Proteins with q-value of < 0.05 in the volcano plot for graphic representation were considered significantly different in the ALS group compared with the HC group. Protein-protein association network analysis and KEGG pathway analysis were conducted using the STRING database [27].

Preparation of CSF samples for the PRM experiments

The isotopically labeled peptides were synthesized with ^{13}C and ^{15}N labeled lysine and arginine (SpikeTides L, JPT Peptide Technologies GmbH, Berlin, Germany). The synthetic peptides with known quantities were used for the PRM analysis. The CSF samples were prepared by mixing with a urea buffer mixture, which was composed of 8 M urea/10 mM TCEP/40 mM CAA in 100 mM TEAB, at a one-to-one ratio. The concentration of materials in the samples was 4 M urea/10 mM TCEP/40 mM CAA in 50 mM TEAB. After adding the buffer mixture, they were incubated for 1 h at RT for reduction and alkylation. Protein digestion was carried out using LysC (Fujifilm Wako Pure Chemical Industries Co., Ltd., Osaka, Japan) at a one-to-fifty ratio for 3 h at 37°C and then using trypsin (sequencing grade modified trypsin, Promega, Fitchburg, WI, USA) at a one-to-fifty ratio at 37°C overnight after diluting the concentration of urea from 4 M to 2 M by adding 50 mM TEAB. Peptides were purified using C_{18} StageTips (3M EmporeTM; 3M, St. Paul, MN, USA) after acidifying with TFA. The eluted solution containing peptides was dried with a SpeedVac (Thermo Fisher Scientific, San Jose, CA, USA) and then stored at –80°C before use.

LC-MS/MS analysis for validation experiments using the PRM method

The peptide samples were analyzed on an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific, Bremen, Germany) interfaced with an Ultimate 3000 RS Autosampler nanoflow liquid chromatography system (Thermo Scientific). The peptides reconstituted in 15 µl of 0.1% FA were loaded on Acclaim PepMap100 Nano-Trap column (100 µm × 2 cm, Thermo Fisher Scientific) packed with 5 µm C_{18} particles at a flow rate of 5 µl per min. The flow rate employed was 250 nl/min using a linear gradient of 6% to 28% solvent B (0.1% FA in 95% ACN) over 55 min on an EASY-Spray column (50 cm × 75 µm, Thermo Fisher Scientific) packed 2 µm C_{18} particles (Thermo Fisher Scientific), which was fitted with an EASY-Spray ion source operated at a voltage of 2.0 kV. Mass spectrometry analysis was conducted in the targeted MS2 mode. The MS1 scan range for a full survey scan was acquired from 300 to 1,600 *m/z* with a resolution of 120,000 at an *m/z* of 200. The mass resolution for MS2 was set to 30,000 at an *m/z* of 200. Automatic gain control was set to 500,000 and 100,000 ions for MS1 and MS2, respectively. The maximum ion injection time was set to 50 ms and 100 ms for MS1 and MS2, respectively. HCD fragmentation energy was set to 32%.

Statistical analyses of PRM experiments

Skyline software was used for the analysis of PRM data [28]. The quantification was based on the sum of the area under the curves (AUC) of three to last fragment ions of the selected peptides and the ratio between peptide fragments of each patient sample. Statistical analysis was performed using GraphPad Prism (GraphPad Software, version 9, Inc., CA, US). A Student unpaired, two-tailed *t*-test was used for comparison of HC with ALS and HC with FTLD. A probability value of $P < 0.05$ was considered statistically significant (Figure 5 and S3; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$, ns: not significant).

RESULTS

Quantitative proteome analysis of CSF samples

We first implemented an unbiased discovery-based approach to identify proteins that are differentially expressed in the CSF of patients with ALS compared with HC individuals. We performed a quantitative proteome analysis of 40 CSF samples comprising 20 patients with ALS and 20 HC individuals using TMT labeling-based mass spectrometry (Figure 1A and Table 1). The 40 samples were split into 4 batches of 10 and were labeled with an 11-plex TMT reagent. The MP, which is a pooled reference

sample of equal volumes from all 40 CSF samples, was placed at the 11th channel of each 11-plex TMT experimental set for the purposes of normalization between batches. After enzyme digestion and TMT labeling, the peptides were pre-fractionated using bRPLC fractionation. The fractions were analyzed on LC-MS/MS. In total, 3,828,053 MS/MS spectra were acquired, and 402,387 MS/MS spectra were assigned to peptides leading to the identification of 26,726 peptides and 3,038 proteins. Out of 3,038 proteins, 1,936 proteins were identified in all four TMT experimental batches (Figure 1B and Supplemental Table S1).

Selection of candidate biomarkers for ALS

To quantify protein abundances measured by the four TMT experimental sets, we normalized the intensity values of each protein to the intensities of the MP samples in each TMT experiment. The normalized data were subsequently subjected to statistical analyses to identify proteins that were differentially expressed between the two groups. Nineteen proteins were found to have differential expression between ALS and HC based on q-value < 0.05 (Table 2 and Figure 2A and B). Within this group, several proteins were previously identified as known candidate biomarker proteins for ALS. These include the neurofilament proteins (NEFL, NEFM and NEFH) [22], which were robustly elevated in ALS, and CHIT1 [22, 23, 29], which showed a >16-fold increase in the ALS samples compared to HC samples. CHI3L2 and UCHL1 [20, 21], which showed increases of ~6 fold and ~2 fold respectively and APOB [30], which was elevated by ~4 fold, were also identified in CSF from ALS patients. In addition, we identified several proteins that were downregulated in ALS, the most prominent of which included L1CAM1 and NPTX2, which was previously identified as downregulated in FTLD [31]. Protein-protein association network analysis showed that NEFL and APP were the most connected proteins within the panel of proteins identified (Figure 2C). Taken together, this discovery-based approach validates previously reported candidate biomarker proteins and identifies several new candidates for further validation experiments as biomarkers for ALS.

Detectability of candidate biomarker proteins in targeted PRM analysis

To further test the potential of the proteins identified in the discovery experiments as biomarkers for ALS, we opted to use PRM analysis of unfractionated CSF. PRM is a targeted mass spectrometry analysis that does not utilize chemical modification of peptides. Instead, known quantities of standard isotope labeled (SIL) peptides corresponding to the proteins of interest are spiked into the sample for accurate quantitation of the relevant endogenous target peptides. We first determined whether peptides of the target proteins were detectable by PRM analysis in unfractionated CSF samples that were

subjected to trypsin-digestion. CSF samples were pooled from 20 patients with ALS and 20 HC individuals, and 200 fmol of SIL peptides corresponding to the endogenous target peptides were added prior to trypsin-digestion to enable subsequent monitoring of relevant peptides. We evaluated 52 proteins with q-value < 0.1 (Table 2), and 16 out of the 52 proteins were detectable by PRM (Table 2). 11 of the 16 proteins were detected by 2 or more peptides, while 5 proteins were detected by 1 peptide (Supplemental Table S2). Interestingly, most of the 16 proteins that were detectable by PRM were proteins that were downregulated in ALS in the initial discovery experiments. Of note, several previously identified protein biomarkers such as neurofilament proteins, CHIT1, CHI3L2, UCHL1 and APOB were not detectable using this method.

Response curve test of candidate biomarker proteins in PRM analysis

Prior to beginning the PRM analysis, we examined the response curve of the target peptides to determine the proper range for accurate quantification of the endogenous 16 target proteins. We added varying amounts of SIL peptides (0.01, 0.1, 1, 10, 100, 1,000 and 10,000 fmol) corresponding to the targets of interest to the trypsin-digested pooled CSF samples and the SIL peptides were quantified by PRM analysis. The majority of the peptides showed linear response curves from 0.1 fmol to 10 pmol (Figure 3). APP (VESLEQEAANER), CHI3L1 (FPLTNAIK and ILGQQVPYATK), CNTNAP2 (FSFSTTK), CSPG5 (EAGSAVEAEELVK), LINGO1 (HLNINAIR), MEGF8 (LFPLPGR), MFAP4 (GFYYSLK), NCAN (ELGGEVFYVGPAR, LSSAIIAAPR and QDLPILVAK), NELL2 (AFLFQDTPR and FTGSSWIK), NPTX1 (LENLEQYSR and LPFVINDGK), NPTX2 (AAVLQLR and TESTLNALLQR), RTN4RL2 (SLEPDTFQGLER) and SERPINA3 (LINDYVK)) showed linear response curves from 0.1 fmol to 10 pmol (Figure 3A–C). APP (THPHFVIPYR and WYFDVTEGK), CHI3L1 (QLLLSAALSAGK), CLSTN1 (YISNEFK), MEGF8 (HVWTTLK), MFAP4 (ADGEYWLGLQNMHLLTLK), NELL2 (ASTATAEQFFQK and SALAYVDGK), RTN4RL2 (LFLQNNLIR), SERPINA1 (LSSWVLLMK and SVLGQLGITK) and TPI1 (QSLGELIGTLNAAK) showed linear response curves from 1 fmol to 10 pmol (Figure 3D–E). SERPINA1 (LYHSEAFTVNFGDTEEAK) and SERPINA3 (AVLDVFEEGTEASAATAVK) showed linear response curves from 100 fmol to 10 pmol (Figure 3F). In summary, almost all peptides corresponding to the 16 candidate biomarkers showed a linear response curve in the detectable range; accordingly, they were considered reliable for quantitation of endogenous peptides in CSF.

Evaluation of target peptide repeatability in PRM analysis

We next determined the repeatability of the PRM analysis by examining if we could see consistent

results when the same samples were measured repeatedly over several days. We added 30, 500 and 1,000 fmol of the heavy SIL peptides to the trypsin-digested CSF samples and performed PRM analysis over 5 days in triplicate, measuring the heavy to light (endogenous) ratio of the target peptides. All the peptides showed lower than 15% of the coefficient of variation (CV). While the peptides at 30 and 500 fmols showed lower than 10% of CV, the peptides at 1,000 fmol showed a relatively higher CV (Figure 4). We attribute these variations to the concentration differences between endogenous and SIL peptides. These observations indicate that the amount of SIL peptides added in the PRM analysis should ideally approximate the concentration of the endogenous peptides. Notably, the CV values obtained from this repeatability experiment agree well with previous reports using the same instrumentation [32], providing confidence in the PRM analysis in this study.

Targeted quantification of candidate ALS biomarker peptides in CSF from ALS, FTLD and HC individuals

To validate whether the selected candidate peptides show differences in patient samples without TMT labeling, we quantified the target peptides using PRM in unfractionated CSF of 20 patients with ALS and 20 HC patients that were our original cohort used in the discovery experiments. We also included an additional 10 CSF samples from patients with ALS and 11 HC samples so that a total of 30 patients with ALS and 31 HC individuals were surveyed (Table 1). Because ALS and FTLD are considered to lie on a disease spectrum [33], we further quantified the candidate peptides in CSF of 19 patients with FTLD (Table 1). To accurately quantify the target peptides using PRM in patients with ALS, FTLD, and HC, SIL peptides were added to the CSF samples at the beginning of the experiment as above. We added similar concentrations of the SIL peptides to the endogenous peptides, but when the concentrations of the endogenous peptides were lower than 5 pmol/ml, we added 5 pmol/ml of SIL peptides to avoid ambiguous detection. After the digestion of the CSF proteins with trypsin, the endogenous peptides and SIL peptides were monitored by PRM analysis (Figure 5A). Two proteins, CLSTN1 (P = 0.0341), and CNTNAP2 (P = 0.0281) showed statistically significant differences between HC and ALS and HC and FTLD (Figure 5B), while 11 proteins showed statistically significant differences between FTLD and HC (Figure 6). Eight peptides from 3 proteins showed no changes between HC, ALS, and FTLD (Supplemental Figure S1). For APP, NCAN and NELL2, 3, 3 and 4 peptides, respectively, showed statistically significant differences between FTLD and HC (Figure 6A). For MEGF8, MFAP4, NPTX1, NPTX2 and RTN4RL2, two peptides showed statistically significant differences between FTLD and HC (Figure 6B). For the rest of the proteins, 1 peptide showed statistically significant differences between FTLD and HC (Figure 6C). Most of the peptides that showed differential levels in FTLD did not achieve statistical significance in ALS; nevertheless, all

peptides downregulated in FTLD showed a trend of reduced concentration in ALS. These results suggest that the majority of the candidate biomarkers identified are more applicable for use in FTLD, with a select number that can be utilized for both ALS and FTLD.

DISCUSSION

The objective of this study was to use an unbiased quantitative discovery approach to identify candidate proteins that are differentially expressed in the CSF of patients with ALS compared to healthy controls, and then subsequently use these protein candidates in a targeted quantitative approach to identify novel biomarkers of ALS and FTLD. Towards this end, we performed discovery-based studies on pre-fractionated CSF from 20 patients with ALS and 20 HC individuals using an 11-plex TMT platform and discovered 19 candidate ALS biomarker proteins ($q < 0.05$). Included in this panel of proteins were several proteins that had been identified previously as being differentially expressed in ALS, and these include neurofilament proteins (NEFL, NEFM, NEFH), CHIT1, CHI3L1, CHI3L2, APOB and UCHL1. Neurofilament proteins are intermediate filament proteins that are elevated in response to neuronal damage [22, 34], while CHIT1, CHI3L1 and CHI3L2 are members of the human chitinase family that are implicated in inflammation [21]. UCHL1 is a ubiquitin-protein hydrolase important for protein homeostasis that is expressed in neurons and neuroendocrine cells, whereas APOB is associated with metabolic changes in ALS [30, 35]. Consistent with previous studies, we found that all of these candidate proteins were elevated in CSF samples from patients with ALS [17, 22, 23, 36-39]. This observation independently strengthens the potential of these proteins as reliable biomarkers for ALS, and importantly, validates our approach as a method to identify novel candidate biomarkers.

After the discovery phase, we utilized PRM analysis of unfractionated CSF without TMT labeling to validate the candidates that we had identified in the discovery phase of our study. This approach involves stringent quantitative measurement of candidate proteins. In addition, the detection and quantitation of candidates in unfractionated CSF is an important consideration for the potential development of diagnostic or prognostic biomarkers. Notably, 69% of candidates that were found to be differentially expressed in the discovery experiments with q-value < 0.1 were not detectable in the PRM analysis, presumably due to the use of trypsin-digested CSF that was not pre-fractionated. Surprisingly, proteins that showed the largest differential expression in our discovery studies such as neurofilament proteins, CHIT1, CHI3L2 and APOB were not detectable by PRM, suggesting that these proteins are expressed at relatively lower levels in CSF. Further validation of these candidates is needed to determine their utility as biomarkers for ALS.

Our PRM analysis successfully detected 16 proteins from our discovery experiments in unfractionated CSF from patients with ALS and FTLD and healthy controls. Notably, two proteins, CNTNAP2 and

CLSTN1, showed a significant reduction in CSF of patients with ALS compared with HC and were further downregulated in FTLD. CNTNAP2 is a member of the transmembrane neurexin superfamily, and mutations in CNTNAP2 are associated with autism spectrum disorders [40, 41]. The absence of CNTNAP2 in mouse models results in hyperactivity, epileptic seizures, abnormal neuronal network activity and cellular changes that include neuronal migration abnormalities, juxtaparanodal changes in myelinated axons, and reduced interneuron numbers [42, 43]. CLSTN1 or Calsyntenin-1is a type 1 transmembrane protein that is implicated in vesicular trafficking through its interaction with kinesin light chains [44]. CLSTN1 amounts are reduced in Alzheimer's disease (AD) postmortem brain, and loss of CLSTN1 leads to defective axonal trafficking of APP and increased Aβ production [45]. The reduction of CNTNAP2 and CLSTN1 in ALS and FTLD CSF samples by PRM suggests that these proteins are strongly implicated in ALS and FTLD pathogenic mechanisms. However, because changes in CNTNAP2 and CLSTN1 are also implicated in autism and AD, respectively [40, 41, 45, 46], further large PRM-based proteomics studies are needed to determine how CSF levels of these proteins differ between multiple neurologic and neurodegenerative diseases.

In addition to CNTNAP2 and CLSTN1, we found eleven other proteins with significant reductions between HC and FTLD, and a trend towards lowered concentrations in ALS. Included among this panel of proteins is NPTX2, a member of the neuronal pentraxin family that is involved in developmental and adult synaptic plasticity and is reduced in CSF of patients with FTLD and other diseases of cognition such as AD [31, 47-49]. Confirming earlier studies, we also detect reductions in NPTX1, another neuronal pentraxin that is involved in synaptic homeostasis, in the CSF of patients with FTLD [31, 50]; however, changes in NPTX1 in CSF of patients with ALS did not reach significance. It is tempting to speculate that these proteins may underlie the cognitive and behavioral changes specific to FTLD that are also seen in a subset of patients with ALS. In support of this notion, APP, CSPG5, Ngr2, and TPI1 are implicated in various aspects of synaptic biology that include synaptic development, spine morphology, synapse maturation and synaptic function [51-54]. Other proteins such as LINGO1 and NELL2 are known to have roles in neuronal survival and may reflect broader neurodegenerative changes [55, 56]. We note that the FTLD cases included in this study primarily have FTLD-tau neuropathology; therefore, further studies that include ALS-FTD and FTLD-TDP are needed to further validate these results.

In summary, our mass spectrometry-based approach has identified two proteins that are significantly reduced in the CSF of patients with ALS and patients with FTLD and has found 11 additional proteins that are downregulated in the CSF of patients with FTLD. Further validation across larger patient cohorts will be needed to test their utility as biomarkers with discriminatory power across the ALS and FTLD disease spectrum.

ETHICS DECLARATIONS

Ethics approval and consent to participate

Samples were obtained from the NEALS Biofluid Repository and the UCSF Memory and Aging Center. Informed consent was obtained by study participants and procedures were approved by IRBs of participating sites.

Consent for publication

Not applicable

Availability of data and materials

All mass spectrometry data and search results have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD023573 and project name 'Discovery of biomarkers for amyotrophic lateral sclerosis and frontotemporal lobar degeneration from human cerebrospinal fluid using mass spectrometry-based proteomics' [57]. Reviewers can access the dataset by using 'reviewer_pxd023573@ebi.ac.uk' as ID and 'OBtleZ5o' as a password.

Competing interests

A.L.B. receives research support from NIH, the Tau Research Consortium, the Association for Frontotemporal Degeneration, Bluefield Project to Cure Frontotemporal Dementia, Corticobasal Degeneration Solutions, the Alzheimer's Drug Discovery Foundation and the Alzheimer's Association. He has served as a consultant for Aeton, Abbvie, Alector, AGTC, Amgen, Arkuda, Arvinas, Asceneuron, Ionis, Lundbeck, Novartis, Passage BIO, Sangamo, Samumed, Third Rock, Toyama and UCB, and received research support from Avid, Biogen, BMS, C2N, Cortice, Eli Lilly, Forum, Genentech, Janssen, Novartis, Pfizer, Roche and TauRx. The other authors report no conflicts of interest.

Authors' contributions

CHN and SS designed research; SS obtained the samples from relevant bodies; A.L.B and L.V.V curated samples; SO. and YJ performed mass spectrometry analysis; SO, YJ and CHN analyzed the data; SO, YJ, CHN and SS wrote the manuscript with edits from A.L.B and L.V.V.; and C. H. N. and S.S. supervised research.

Funding

This work was supported by a grant awarded to S.S by alsFINDINGaCure. Data collection and dissemination of the data presented in this manuscript was supported by the NIH through the 4R-tau Neuroimaging Initiative (4RTNI, R01AG038791), the ALLFTD Consortium (U19AG063911), the former ARTFL & LEFFTDS Consortia (ARTFL: U54NS092089; LEFFTDS: U01AG045390), and the UCSF Alzheimer's Disease Research Center and Program Project Grant (ADRC P50AG023501, PPG P01AG019724), as well as the Rainwater Charitable Foundation.

Acknowledgments

We acknowledge the support of the shared instrumentation grant (S10OD021844). The authors acknowledge the invaluable contributions of the study participants and families as well as the assistance of the support staffs at each of the participating sites.

List of abbreviations

ALS: Amyotrophic lateral sclerosis

FTD: frontotemporal dementia

CSF: Cerebrospinal fluids

MS: Mass spectrometry

PRM: Parallel reaction monitoring

TMT: Tandem mass tag

REFERENCES

1. Rowland, L.P. and N.A. Shneider, *Amyotrophic lateral sclerosis.* N Engl J Med, 2001. **344**(22): p. 1688-700.
2. Mitchell, J.D. and G.D. Borasio, *Amyotrophic lateral sclerosis.* Lancet, 2007. **369**(9578): p. 2031-2041.
3. Wijesekera, L.C. and P.N. Leigh, *Amyotrophic lateral sclerosis.* Orphanet J Rare Dis, 2009. **4**: p. 3.
4. Kiernan, M.C., et al., *Amyotrophic lateral sclerosis.* Lancet, 2011. **377**(9769): p. 942-55.
5. Vu, L.T. and R. Bowser, *Fluid-Based Biomarkers for Amyotrophic Lateral Sclerosis.* Neurotherapeutics, 2017. **14**(1): p. 119-134.
6. Collins, M.A., et al., *Label-Free LC-MS/MS Proteomic Analysis of Cerebrospinal Fluid Identifies Protein/Pathway Alterations and Candidate Biomarkers for Amyotrophic Lateral Sclerosis.* J Proteome Res, 2015. **14**(11): p. 4486-501.
7. Cruz, M.P., *Edaravone (Radicava): A novel neuroprotective agent for the treatment of amyotrophic lateral sclerosis.* P t, 2018. **43**(1): p. 25-28.
8. Mehta, P., et al., *Prevalence of Amyotrophic Lateral Sclerosis - United States, 2015.* Mmwr-Morbidity and Mortality Weekly Report, 2018. **67**(46): p. 1285-1289.
9. Rabinovici, G.D. and B.L. Miller, *Frontotemporal lobar degeneration: epidemiology, pathophysiology, diagnosis and management.* CNS Drugs, 2010. **24**(5): p. 375-98.
10. Mackenzie, I.R. and R. Rademakers, *The role of transactive response DNA-binding protein-43 in amyotrophic lateral sclerosis and frontotemporal dementia.* Curr Opin Neurol, 2008. **21**(6): p. 693-700.
11. Ash, P.E., et al., *Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS.* Neuron, 2013. **77**(4): p. 639-46.
12. Starr, A. and R. Sattler, *Synaptic dysfunction and altered excitability in C9ORF72 ALS/FTD.* Brain Res, 2018. **1693**(Pt A): p. 98-108.
13. Strong, M.J., et al., *Amyotrophic lateral sclerosis - frontotemporal spectrum disorder (ALS-FTSD): Revised diagnostic criteria.* Amyotroph Lateral Scler Frontotemporal Degener, 2017. **18**(3-4): p. 153-174.
14. Brooks, B.R., et al., *El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis.* Amyotroph Lateral Scler Other Motor Neuron Disord, 2000. **1**(5): p. 293-9.

15. Barschke, P., et al., *Proteomic studies in the discovery of cerebrospinal fluid biomarkers for amyotrophic lateral sclerosis.* Expert Rev Proteomics, 2017. **14**(9): p. 769-777.
16. Chen, Y., et al., *Proteomic analysis of cerebrospinal fluid in amyotrophic lateral sclerosis.* Exp Ther Med, 2016. **11**(6): p. 2095-2106.
17. Thompson, A.G., et al., *Cerebrospinal fluid macrophage biomarkers in amyotrophic lateral sclerosis.* Ann Neurol, 2018. **83**(2): p. 258-268.
18. Kruger, T., et al., *Proteome analysis of body fluids for amyotrophic lateral sclerosis biomarker discovery.* Proteomics Clin Appl, 2013. **7**(1-2): p. 123-35.
19. Bowser, R., M.R. Turner, and J. Shefner, *Biomarkers in amyotrophic lateral sclerosis: opportunities and limitations.* Nat Rev Neurol, 2011. **7**(11): p. 631-8.
20. Oeckl, P., et al., *Proteomics in cerebrospinal fluid and spinal cord suggests UCHL1, MAP2 and GPNMB as biomarkers and underpins importance of transcriptional pathways in amyotrophic lateral sclerosis.* Acta Neuropathol, 2020. **139**(1): p. 119-134.
21. Gaur, N., et al., *The Chitinases as Biomarkers for Amyotrophic Lateral Sclerosis: Signals From the CNS and Beyond.* Front Neurol, 2020. **11**: p. 377.
22. Gagliardi, D., et al., *Diagnostic and Prognostic Role of Blood and Cerebrospinal Fluid and Blood Neurofilaments in Amyotrophic Lateral Sclerosis: A Review of the Literature.* Int J Mol Sci, 2019. **20**(17).
23. Varghese, A.M., et al., *Chitotriosidase - a putative biomarker for sporadic amyotrophic lateral sclerosis.* Clin Proteomics, 2013. **10**(1): p. 19.
24. Meeter, L.H.H., et al., *Clinical value of neurofilament and phospho-tau/tau ratio in the frontotemporal dementia spectrum.* Neurology, 2018. **90**(14): p. e1231-e1239.
25. Smaczniak, C., et al., *Proteomics-based identification of low-abundance signaling and regulatory protein complexes in native plant tissues.* Nat Protoc, 2012. **7**(12): p. 2144-58.
26. Scherling, C.S., et al., *Cerebrospinal fluid neurofilament concentration reflects disease severity in frontotemporal degeneration.* Ann Neurol, 2014. **75**(1): p. 116-26.
27. Szklarczyk, D., et al., *STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets.* Nucleic Acids Res, 2019. **47**(D1): p. D607-D613.
28. MacLean, B., et al., *Skyline: an open source document editor for creating and analyzing targeted proteomics experiments.* Bioinformatics, 2010. **26**(7): p. 966-8.
29. Gaiani, A., et al., *Diagnostic and Prognostic Biomarkers in Amyotrophic Lateral Sclerosis: Neurofilament Light Chain Levels in Definite Subtypes of Disease.* JAMA Neurol, 2017. **74**(5): p. 525-532.
30. Mariosa, D., et al., *Blood biomarkers of carbohydrate, lipid, and apolipoprotein metabolisms and risk of amyotrophic lateral sclerosis: A more than 20-year follow-up of the Swedish*

AMORIS cohort. Ann Neurol, 2017. **81**(5): p. 718-728.
31. van der Ende, E.L., et al., *Neuronal pentraxin 2: a synapse-derived CSF biomarker in genetic frontotemporal dementia.* J Neurol Neurosurg Psychiatry, 2020. **91**(6): p. 612-621.
32. Piehowski, P.D., et al., *Sources of technical variability in quantitative LC-MS proteomics: human brain tissue sample analysis.* J Proteome Res, 2013. **12**(5): p. 2128-37.
33. Gao, F.B., S. Almeida, and R. Lopez-Gonzalez, *Dysregulated molecular pathways in amyotrophic lateral sclerosis-frontotemporal dementia spectrum disorder.* Embo j, 2017. **36**(20): p. 2931-2950.
34. Yuan, A. and R.A. Nixon, *Specialized roles of neurofilament proteins in synapses: Relevance to neuropsychiatric disorders.* Brain Res Bull, 2016. **126**(Pt 3): p. 334-346.
35. Day, I.N. and R.J. Thompson, *UCHL1 (PGP 9.5): neuronal biomarker and ubiquitin system protein.* Prog Neurobiol, 2010. **90**(3): p. 327-62.
36. Zhu, S., et al., *Targeted Multiple Reaction Monitoring Analysis of CSF Identifies UCHL1 and GPNMB as Candidate Biomarkers for ALS.* J Mol Neurosci, 2019. **69**(4): p. 643-657.
37. Gille, B., et al., *Inflammatory markers in cerebrospinal fluid: independent prognostic biomarkers in amyotrophic lateral sclerosis?* J Neurol Neurosurg Psychiatry, 2019. **90**(12): p. 1338-1346.
38. Thompson, A.G., et al., *CSF chitinase proteins in amyotrophic lateral sclerosis.* J Neurol Neurosurg Psychiatry, 2019. **90**(11): p. 1215-1220.
39. Steinacker, P., et al., *Chitotriosidase (CHIT1) is increased in microglia and macrophages in spinal cord of amyotrophic lateral sclerosis and cerebrospinal fluid levels correlate with disease severity and progression.* J Neurol Neurosurg Psychiatry, 2018. **89**(3): p. 239-247.
40. Rodenas-Cuadrado, P., J. Ho, and S.C. Vernes, *Shining a light on CNTNAP2: complex functions to complex disorders.* Eur J Hum Genet, 2014. **22**(2): p. 171-8.
41. Alarcon, M., et al., *Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene.* Am J Hum Genet, 2008. **82**(1): p. 150-9.
42. Penagarikano, O., et al., *Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits.* Cell, 2011. **147**(1): p. 235-46.
43. Poliak, S., et al., *Juxtaparanodal clustering of Shaker-like K+ channels in myelinated axons depends on Caspr2 and TAG-1.* J Cell Biol, 2003. **162**(6): p. 1149-60.
44. Konecna, A., et al., *Calsyntenin-1 docks vesicular cargo to kinesin-1.* Mol Biol Cell, 2006. **17**(8): p. 3651-63.
45. Vagnoni, A., et al., *Calsyntenin-1 mediates axonal transport of the amyloid precursor protein and regulates Abeta production.* Hum Mol Genet, 2012. **21**(13): p. 2845-54.
46. Lleo, A., et al., *Changes in Synaptic Proteins Precede Neurodegeneration Markers in Preclinical Alzheimer's Disease Cerebrospinal Fluid.* Mol Cell Proteomics, 2019. **18**(3): p.

546-560.
47. Xiao, M.F., et al., *NPTX2 and cognitive dysfunction in Alzheimer's Disease.* Elife, 2017. **6**.
48. Pelkey, K.A., et al., *Pentraxins coordinate excitatory synapse maturation and circuit integration of parvalbumin interneurons.* Neuron, 2015. **85**(6): p. 1257-72.
49. Xu, D., et al., *Narp and NP1 form heterocomplexes that function in developmental and activity-dependent synaptic plasticity.* Neuron, 2003. **39**(3): p. 513-28.
50. Tsui, C.C., et al., *Narp, a novel member of the pentraxin family, promotes neurite outgrowth and is dynamically regulated by neuronal activity.* J Neurosci, 1996. **16**(8): p. 2463-78.
51. Ludewig, S. and M. Korte, *Novel Insights into the Physiological Function of the APP (Gene) Family and Its Proteolytic Fragments in Synaptic Plasticity.* Front Mol Neurosci, 2016. **9**: p. 161.
52. Roland, B.P., et al., *Structural and Genetic Studies Demonstrate Neurologic Dysfunction in Triosephosphate Isomerase Deficiency Is Associated with Impaired Synaptic Vesicle Dynamics.* PLoS Genet, 2016. **12**(3): p. e1005941.
53. Juttner, R., et al., *Impaired presynaptic function and elimination of synapses at premature stages during postnatal development of the cerebellum in the absence of CALEB (CSPG5/neuroglycan C).* Eur J Neurosci, 2013. **38**(9): p. 3270-80.
54. Wills, Z.P., et al., *The nogo receptor family restricts synapse number in the developing hippocampus.* Neuron, 2012. **73**(3): p. 466-81.
55. Zhang, Z., et al., *LINGO-1 receptor promotes neuronal apoptosis by inhibiting WNK3 kinase activity.* J Biol Chem, 2013. **288**(17): p. 12152-60.
56. Munemasa, Y., et al., *The neuronal EGF-related gene Nell2 interacts with Macf1 and supports survival of retinal ganglion cells after optic nerve injury.* PLoS One, 2012. **7**(4): p. e34810.
57. Vizcaíno, J.A., et al., *2016 update of the PRIDE database and its related tools.* Nucleic Acids Res, 2016. **44**(D1): p. D447-56.

TABLES AND FIGURES

Table 1. The demographics of CSF samples used in this study.

bvFTD: behavioral variant; CBS: corticobasal syndrome; PSP-RS: progressive supranuclear palsy (Richardson syndrome); nFV-PPA: nonfluent variant of primary progressive aphasia. ND: Not determined.

Table 2. Differential proteins between ALS and control identified in the discovery experiment.

Figure 1. Experimental strategy and identified proteins in discovery studies of ALS biomarker candidates.

Human CSF samples from 20 patients with ALS and 20 HC individuals were analyzed with four 11-plex TMT experimental batches. MP was prepared by pooling all 40 CSF samples and was added to each TMT experiment to normalize between batches. CSF proteins were digested with trypsin followed by pre-fractionation into 24 fractions prior to mass spectrometry analysis. Proteins were subsequently identified by conducting a database search of the acquired mass spectra (A)**.** The numbers of identified proteins from 4 batches of TMT-based proteomics experiments are shown in the Venn diagram (B).

Figure 2. Statistical and bioinformatic analyses of the identified CSF proteins.

The quantified CSF proteins from 20 patients with ALS and 20 HC individuals were plotted on a volcano plot. The curved lines are the boundaries for q-value of 0.05. The proteins with q-value < 0.05 are colored in red (A). The differentially expressed proteins (q-value < 0.05) are shown on a heatmap (B). Protein-protein association network analysis (C) for differentially expressed proteins (q-value < 0.05) using STRING.

Figure 3. The response curve of candidate biomarker peptides in the PRM analysis.

The SIL standard peptides ranging from 0.01 fmol to 10 pmol were added to the CSF samples and quantified by the PRM analysis to monitor the dynamic range of each target peptide. The peptides with various response curves are shown; 0.1 to 10,000 fmol (A, B and C), 1 to 10,000 fmol (D and E) and 100 to 10,000 fmol (F).

Figure 4. The repeatability of the target peptides detectable in the PRM analysis.

The repeatability experiments of 37 detectable endogenous peptides in PRM analysis were conducted. Three different amounts of target SIL peptides (30, 500, and 1,000 fmol) were added to the trypsin-digested CSF samples immediately before the PRM analysis. The PRM analyses were performed over 5 days in triplicate, measuring the heavy to light ratio of the target peptides.

Figure 5. Differentially expressed proteins between ALS and HC validated by the PRM analysis.

The selected candidate ALS biomarker proteins were validated by PRM analysis of CSF samples from 30 ALS, 19 FTD and 31 HC individuals. SIL peptides were added to the CSF samples for the accurate quantification of the target peptides, followed by trypsin-digestion and the PRM analysis. The peptide abundance was calculated by extracting the area under the curve of light endogenous and heavy SIL peptides (A). Graphs showing reductions in CLSTN1 and CNTNAP2 in ALS and FTD. All PRM-MS analyses were performed in 3 technical replicates. Student's unpaired, two-tailed t-test was conducted for statistical analysis between groups (*P<0.05; **P<0.01; ****P<0.0001) (B). Four different subtypes of FTLD are marked; red for FTLD PSP-RS, blue for FTLD bv FTD, green for FTLD CBS and purple for FTLD nfvppA. Mean $\pm$ SEM

Figure 6. Differentially expressed proteins between FTD and HC validated by the PRM analysis.

The selected candidate ALS biomarker proteins were validated by PRM analysis of CSF samples from 30 ALS, 19 FTD and 31 HC individuals. The peptides with statistical significance between FTD and HC are shown. The peptide abundance was calculated by extracting the area under the curve of light endogenous and heavy SIL peptides. All PRM-MS analyses were performed in 3 technical replicates. Proteins detected by 3-4 peptides (A), 2 peptides (B) and 1 peptide (C). All PRM-MS analyses were performed in 3 technical replicates. Student's unpaired, two-tailed *t*-test was conducted for statistical analysis between groups (*P<0.05; **P<0.01; ****P<0.0001; ns: not significant). Four different subtypes of FTLD are marked; red for FTLD PSP-RS, blue for FTLD bv FTD, green for FTLD CBS and purple for FTLD nfvppA. Mean ± SEM

Supplemental Table S1. Complete list of proteins identified by the discovery experiment

Supplemental Table S2. List of peptides detectable in the PRM analysis.

Supplemental Figure S1. Quantification of candidate ALS biomarker proteins in the PRM analysis

PRM analysis of candidate ALS biomarker peptides in CSF samples from 30 ALS, 19 FTD, and 31 HC individuals. All PRM-MS analyses were performed in 3 technical replicates. The beeswarm plots depict the peptide quantification in CSF by measuring intensities of the light and SIL heavy standard peptides. Student's unpaired, two-tailed t-test was conducted for the statistical analysis between groups. (ns: not significant). Four different subtypes of FTLD are marked; red for FTLD PSP-RS, blue for FTLD bv FTD, green for FTLD CBS and purple for FTLD nfvppA. Mean ± SEM

Figures

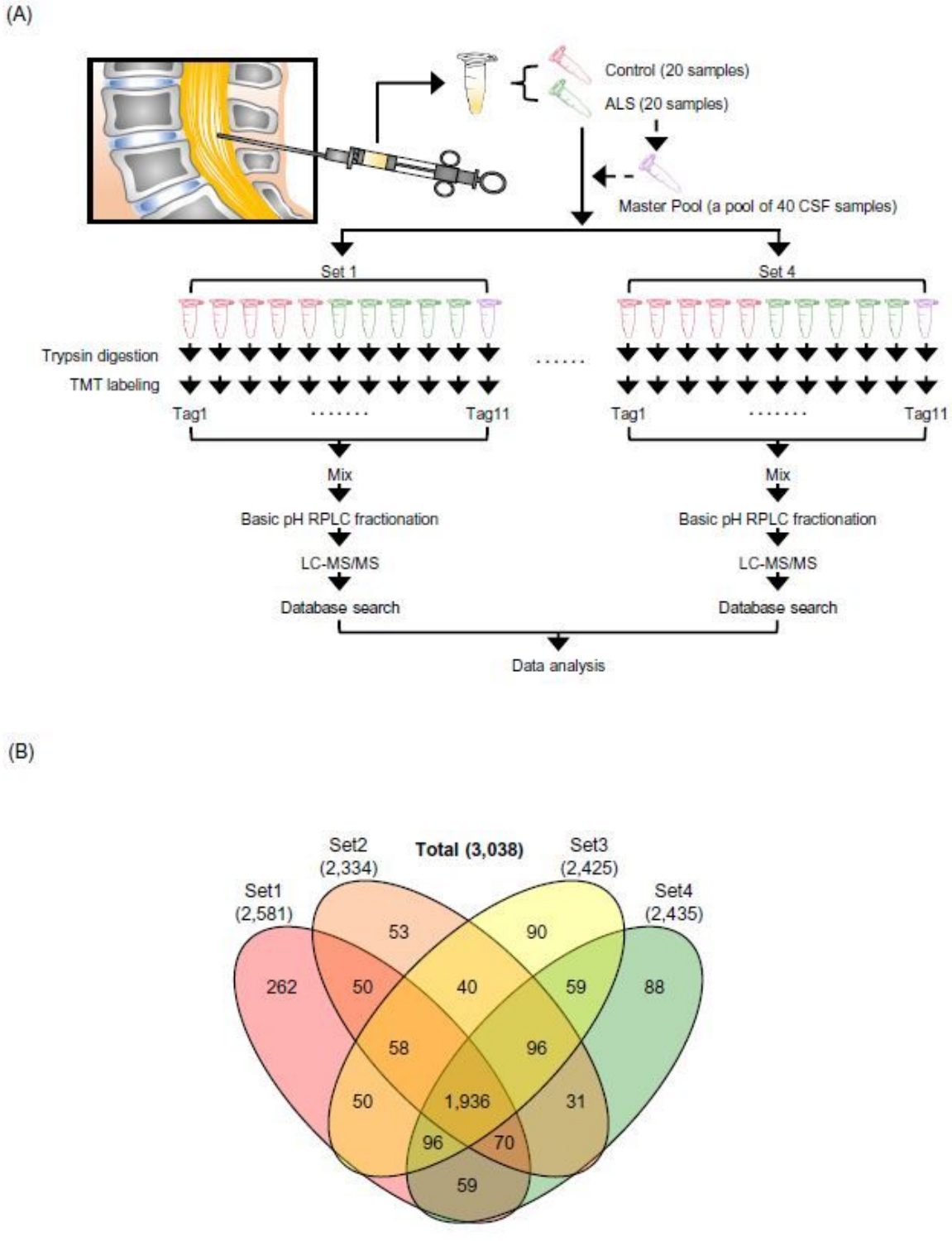

Figure 1

Experimental strategy and identified proteins in discovery studies of ALS biomarker candidates. Human CSF samples from 20 patients with ALS and 20 HC individuals were analyzed with four 11-plex TMT experimental batches. MP was prepared by pooling all 40 CSF samples and was added to each TMT

experiment to normalize between batches. CSF proteins were digested with trypsin followed by pre-fractionation into 24 fractions prior to mass spectrometry analysis. Proteins were subsequently identified by conducting a database search of the acquired mass spectra (A). The numbers of identified proteins from 4 batches of TMT-based proteomics experiments are shown in the Venn diagram (B).

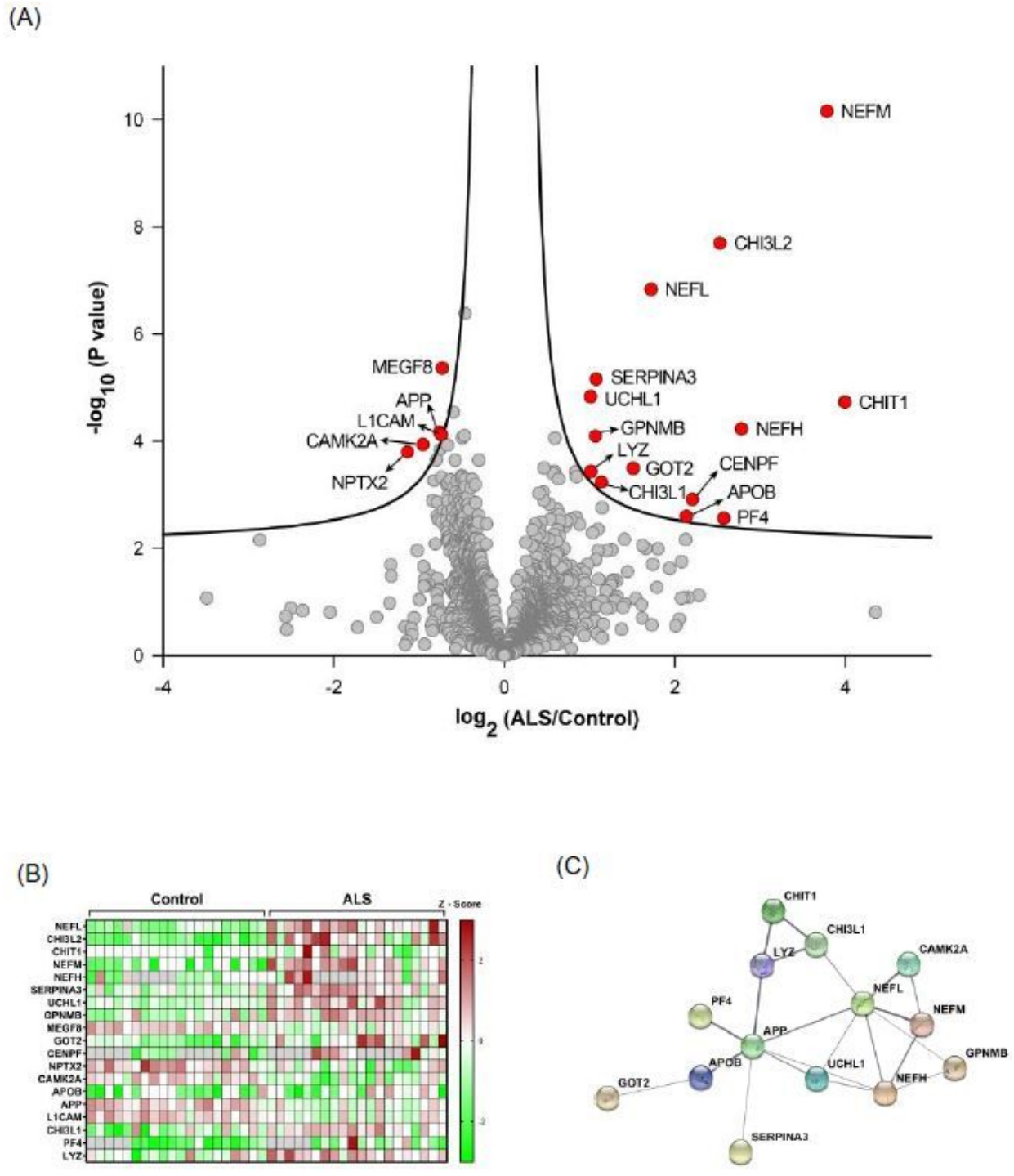

Figure 2

Statistical and bioinformatic analyses of the identified CSF proteins. The quantified CSF proteins from 20 patients with ALS and 20 HC individuals were plotted on a volcano plot. The curved lines are the boundaries for q-value of 0.05. The proteins with q-value < 0.05 are colored in red (A). The differentially expressed proteins (q-value < 0.05) are shown on a heatmap (B). Protein-protein association network analysis (C) for differentially expressed proteins (q-value < 0.05) using STRING.

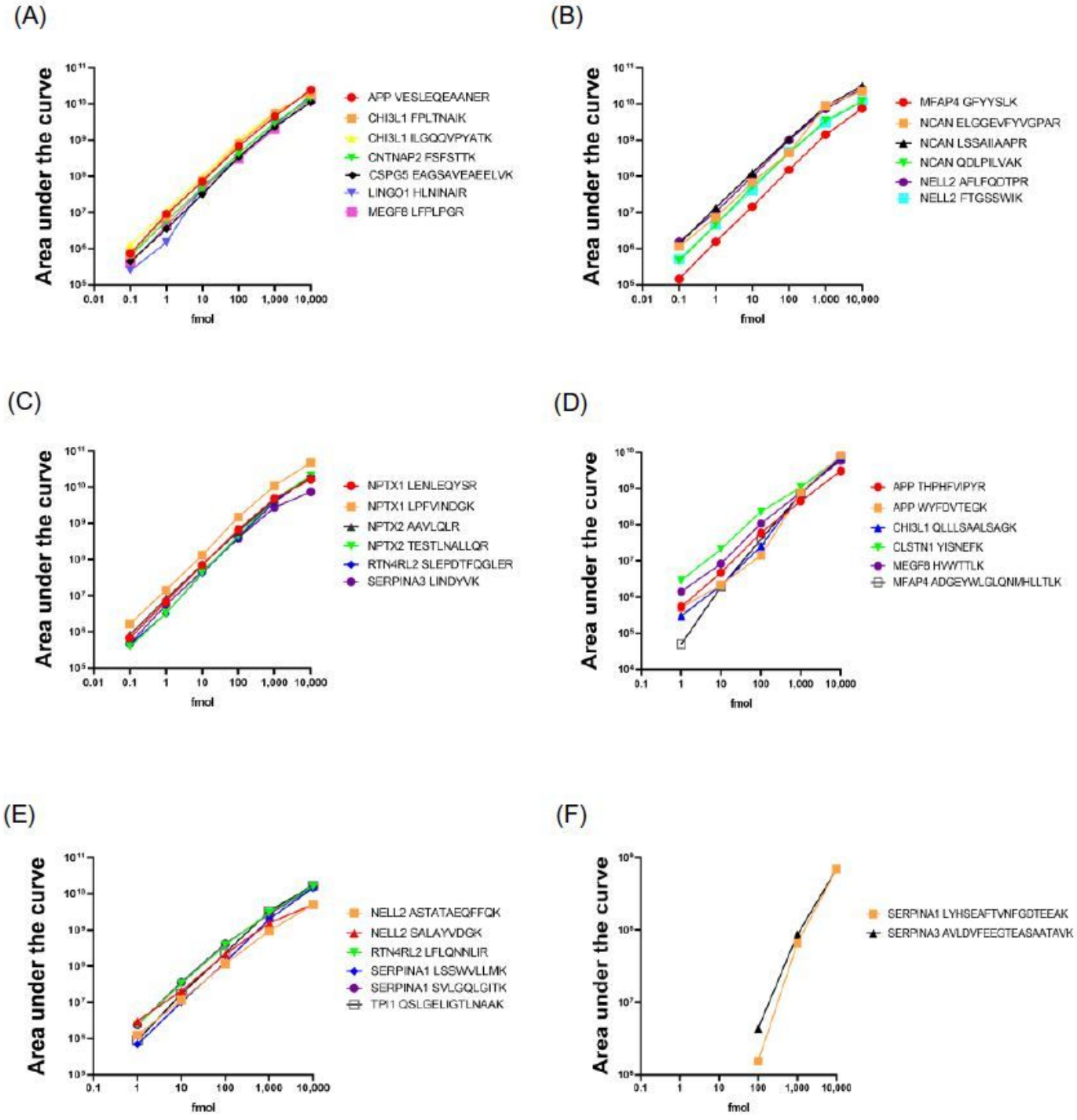

Figure 3

The response curve of candidate biomarker peptides in the PRM analysis. The SIL standard peptides ranging from 0.01 fmol to 10 pmol were added to the CSF samples and quantified by the PRM analysis to monitor the dynamic range of each target peptide. The peptides with various response curves are shown; 0.1 to 10,000 fmol (A, B and C), 1 to 10,000 fmol (D and E) and 100 to 10,000 fmol (F).

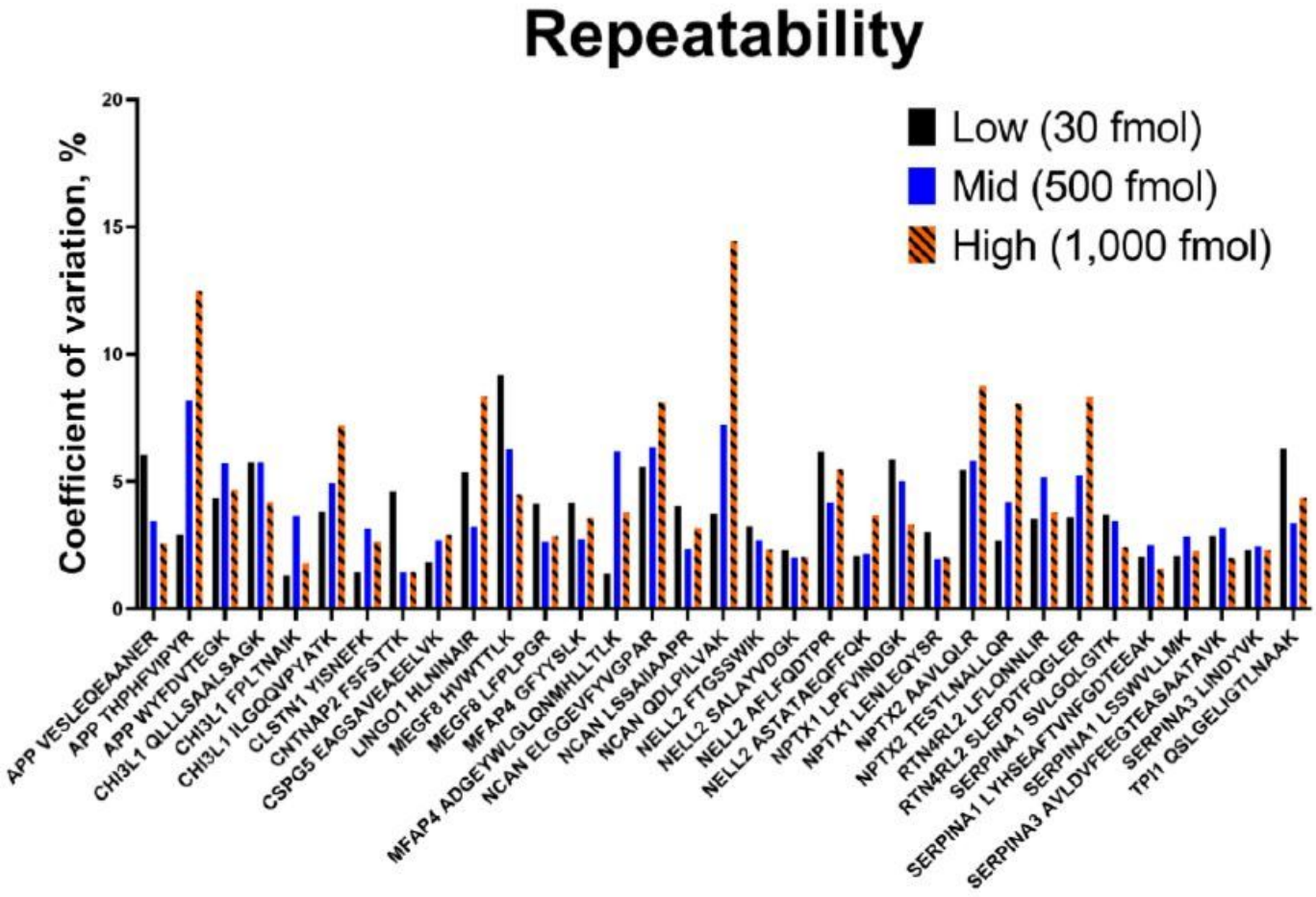

Figure 4

The repeatability of the target peptides detectable in the PRM analysis. The repeatability experiments of 37 detectable endogenous peptides in PRM analysis were conducted. Three different amounts of target SIL peptides (30, 500, and 1,000 fmol) were added to the trypsin-digested CSF samples immediately before the PRM analysis. The PRM analyses were performed over 5 days in triplicate, measuring the heavy to light ratio of the target peptides.

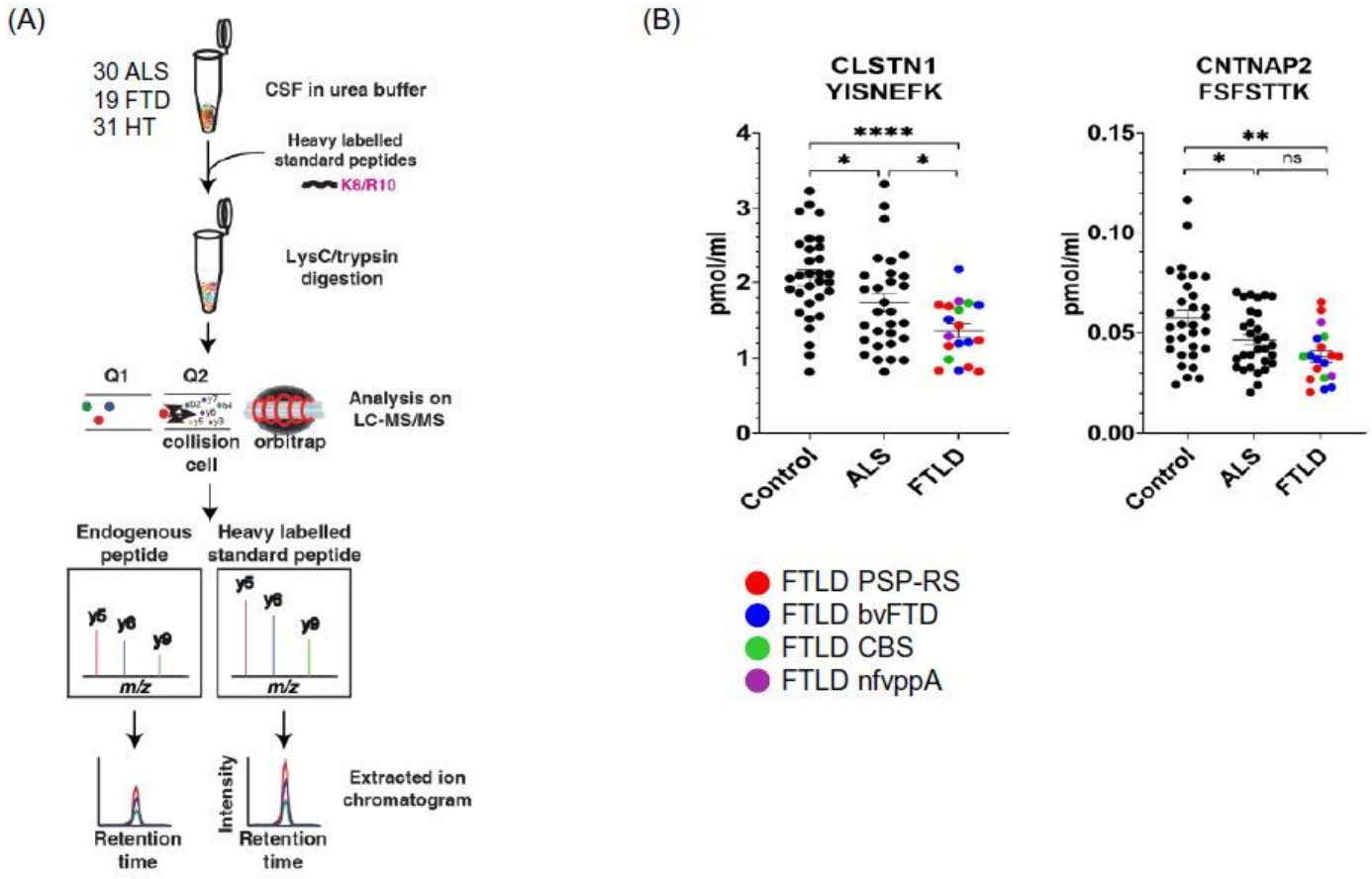

Figure 5

Differentially expressed proteins between ALS and HC validated by the PRM analysis. The selected candidate ALS biomarker proteins were validated by PRM analysis of CSF samples from 31 30 ALS, 19 FTD and HC individuals. SIL peptides were added to the CSF samples for the accurate quantification of the target peptides, followed by trypsin-digestion and the PRM analysis. The peptide abundance was calculated by extracting the area under the curve of light endogenous and heavy SIL peptides (A). Graphs showing reductions in CLSTN1 and CNTNAP2 in ALS and FTD. All PRM-MS analyses were performed in 3 technical replicates. Student's unpaired, two-tailed t-test was conducted for statistical analysis between groups (*P<0.05; **P<0.01; ****P<0.0001) (B). Four different subtypes of FTLD are marked; red for FTLD PSP-RS, blue for FTLD bv FTD, green for FTLD CBS and purple for FTLD nfvppA. Mean + SEM

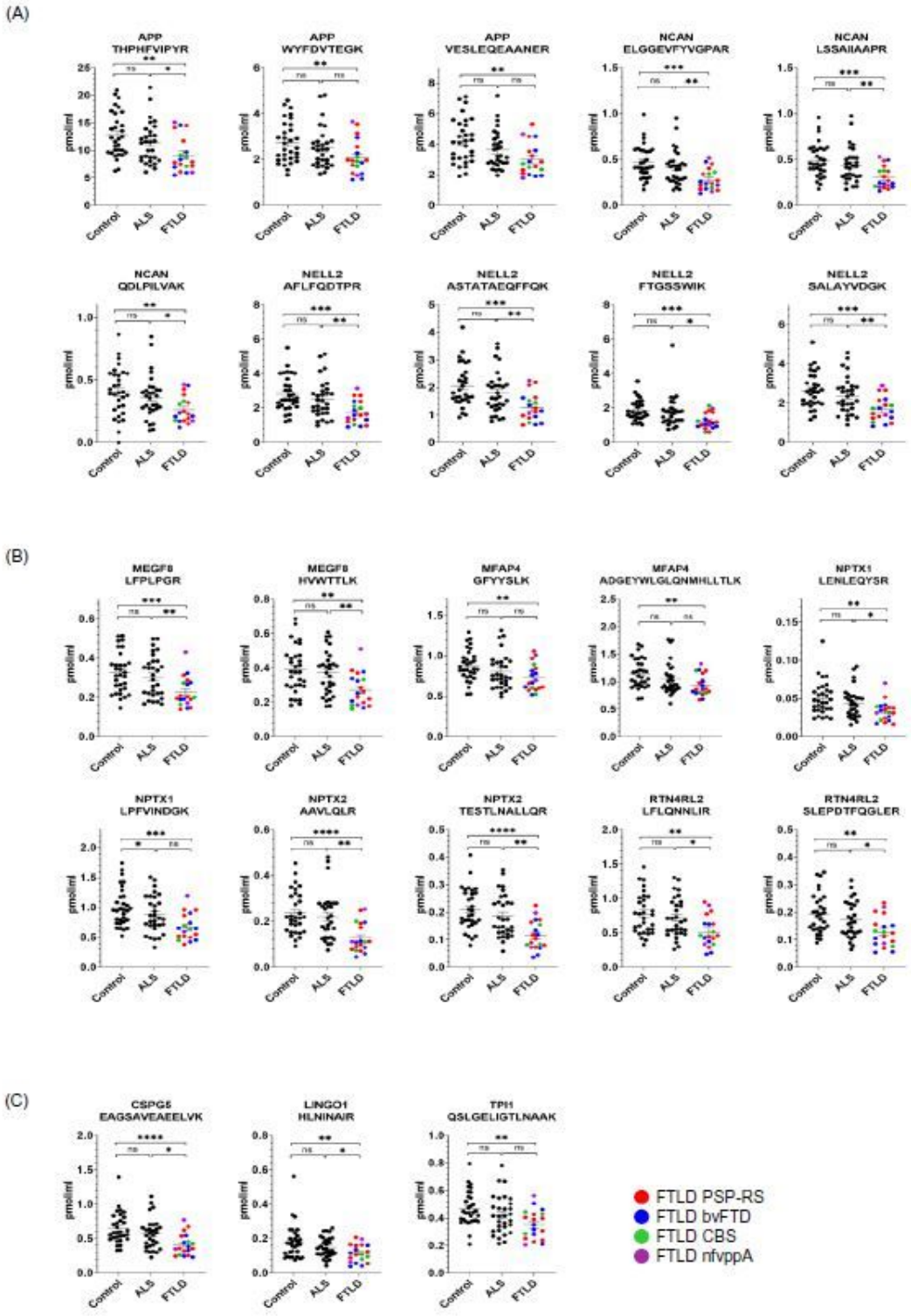

Figure 6

Differentially expressed proteins between FTD and HC validated by the PRM analysis. The selected candidate ALS biomarker proteins were validated by PRM analysis of CSF samples from 30 ALS, 19 FTD and 31 HC individuals. The peptides with statistical significance between FTD and HC are shown. The peptide abundance was calculated by extracting the area under the curve of light endogenous and heavy SIL peptides. All PRM-MS analyses were performed in 3 technical replicates. Proteins detected by 3-4

peptides (A), 2 peptides (B) and 1 peptide (C). All PRM-MS analyses were performed in 3 technical replicates. Student's unpaired, two-tailed t-test was conducted for statistical analysis between groups (*P<0.05; **P<0.01; ****P<0.0001; ns: not significant). Four different subtypes of FTLD are marked; red for FTLD PSP-RS, blue for FTLD bv FTD, green for FTLD CBS and purple for FTLD nfvppA. Mean + SEM

Supplementary Files

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

- suppfigure1.jpg
- SupplementalTableS1DiscoveryDatav02.xlsx