Acta Neuropathologica
https://doi.org/10.1007/s00401-019-02093-x

ORIGINAL PAPER

Proteomics in cerebrospinal fluid and spinal cord suggests UCHL1, MAP2 and GPNMB as biomarkers and underpins importance of transcriptional pathways in amyotrophic lateral sclerosis

Patrick Oeckl[1] · Patrick Weydt[1,2] · Dietmar R. Thal[3,4] · Jochen H. Weishaupt[1] · Albert C. Ludolph[1] · Markus Otto[1]

Received: 4 September 2019 / Revised: 11 October 2019 / Accepted: 1 November 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Amyotrophic lateral sclerosis (ALS) is a devastating neurodegenerative disease and the proteins and pathways involved in the pathophysiology are not fully understood. Even less is known about the preclinical disease phase. To uncover new ALS-related proteins and pathways, we performed a comparative proteomic analysis in cerebrospinal fluid (CSF) of asymptomatic ($n=14$) and symptomatic ($n=14$) ALS mutation carriers and sporadic ALS patients ($n=12$) as well as *post-mortem* human spinal cord tissue (controls: $n=7$, ALS, $n=8$). Using a CSF-optimized proteomic workflow, we identified novel (e.g., UCHL1, MAP2, CAPG, GPNMB, HIST1H4A, HIST1H2B) and well-described (e.g., NEFL, NEFH, NEFM, CHIT1, CHI3L1) protein level changes in CSF of sporadic and genetic ALS patients with enrichment of proteins related to transcription, cell cycle and lipoprotein remodeling (total protein IDs: 2303). No significant alteration was observed in asymptomatic ALS mutation carriers representing the prodromal disease phase. We confirmed UCHL1, MAP2, CAPG and GPNMB as novel biomarker candidates for ALS in an independent validation cohort of patients ($n=117$) using multiple reaction monitoring. In spinal cord tissue, 292 out of 6810 identified proteins were significantly changed in ALS with enrichment of proteins involved in mRNA splicing and of the neurofilament compartment. In conclusion, our proteomic data in asymptomatic ALS mutation carriers support the hypothesis of a sudden disease onset instead of a long preclinical phase. Both CSF and tissue proteomic data indicate transcriptional pathways to be amongst the most affected. UCHL1, MAP2 and GPNMB are promising ALS biomarker candidates which might provide additional value to the established neurofilaments in patient follow-up and clinical trials.

Keywords Amyotrophic lateral sclerosis · Biomarker · Cerebrospinal fluid · Proteomics · Multiple reaction monitoring · UCHL1

Introduction

Amyotrophic lateral sclerosis (ALS) is a rare and fatal neurodegenerative disease characterized by severe motor neuron loss and finally death due to respiratory failure within 3–5 years. Except for about 5% of ALS patients with a clear monogenetic background (gALS) the cause of sporadic ALS (sALS) remains elusive. The neuropathological hallmark of ALS is TDP-43-positive protein inclusions and several mechanisms have been implicated in ALS pathophysiology such as inflammation, mitochondrial dysfunction, protein and RNA homeostasis but their relative contributions to the pathogenesis are unclear [1]. Furthermore, disease-modifying treatment strategies were largely unsuccessful so far [2] implying additional yet unknown mechanisms with an important role in the disease process.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00401-019-02093-x) contains supplementary material, which is available to authorized users.

✉ Markus Otto
markus.otto@uni-ulm.de

1 Department of Neurology, Ulm University Hospital, Oberer Eselsberg 45, 89081 Ulm, Germany

2 Department of Neurodegenerative Diseases and Geriatric Psychiatry, University of Bonn, Bonn, Germany

3 Laboratory of Neuropathology, Institute of Pathology, Ulm University, Ulm, Germany

4 Department of Imaging and Pathology, KU Leuven and Department of Pathology, UZ Leuven, Louvain, Belgium

Published online: 07 November 2019

Even less is known about the preclinical disease phase and the available information is mainly derived from mutation-based animal models. Diagnosis of ALS is based on clinical symptoms [3] hampering the investigation of the presymptomatic phase in patients. The most promising ALS biomarker candidates are neurofilament levels in cerebrospinal fluid (CSF) and blood but they emerge during the symptomatic disease phase and do not seem to change during disease progression [4, 5].

Hence, there is an urgent need to identify new ALS-related pathways to close the gaps in our current understanding of ALS pathophysiology and to potentially provide new targets for disease-modifying treatment strategies. Furthermore, the establishment of new biomarker candidates for early diagnosis and disease progression is an important requirement to improve treatment of patients including drug development.

The use of unbiased proteomics is a valuable tool for the identification of disease-related pathways and new biomarker candidates for further validation. It has led to the identification of the inflammatory protein CHIT1 as a CSF marker in ALS [6] and this has been validated by several studies [7–10]. However, previous proteomic studies did not reach the sensitivity to detect proteins in the low concentration range where most biomarker candidates for neurological diseases are expected including established ones such as neurofilaments, tau or synucleins. The ongoing technical advances in the proteomics field now enable a deeper proteomic coverage in CSF samples making novel biomarkers accessible. This is especially important for the small changes that can be expected to occur during the earliest disease phases. The investigation of asymptomatic ALS mutation carriers offers the unique option to study the preclinical ALS disease phase. We and others showed that the CSF biomarkers CHIT1 [8] and neurofilaments [5, 11] increase within the time of symptom onset. Furthermore, a proteomic analysis comparing changes in CSF and spinal cord tissue is currently not available but can provide mechanistic information of the proteins.

The aim of the present study was to uncover novel biomarker candidates and disease-related pathways of ALS by comparative proteomic analysis of human CSF and spinal cord samples from symptomatic ALS patients and to identify early changes in CSF of asymptomatic ALS mutation carriers. We analyzed CSF samples from 14 asymptomatic ALS mutation carriers, 12 sALS, 14 gALS and 16 control patients using isobaric tags for relative and absolute quantification (iTRAQ). The identified biomarker candidates were further evaluated in a second validation cohort consisting of 61 sALS, 24 gALS and 32 control patients using multiple reaction monitoring (MRM). *Post-mortem* spinal cord tissue from 8 ALS and 7 control patients were analyzed by label-free proteomics.

Materials and methods

Reagents

The iTRAQ Reagents Multi-Plex Kit (4-plex) was purchased from AB Sciex (Darmstadt, Germany, #4352135), Flagellin from *B. subtilis* from InvivoGen (San Diego, USA, FLA-BS Ultrapure) and β-Lactoglobulin from Thermo (Darmstadt, Germany, #400979). Stable isotope-labeled (SIL) peptides were custom-synthesized from Thermo and QPrESTs were provided by or purchased from Atlas Antibodies AB (Bromma, Sweden).

Patients

Patients were recruited between 2009 and 2017 at the Department of Neurology, Ulm University Hospital. ALS was diagnosed according to the revised El Escorial criteria [12] and all ALS patients were tested for known ALS gene mutations. Asymptomatic ALS mutation carriers were recruited through the German Presymptomatic ALS study and where identified by genetic testing of first-degree relatives of gALS patients. Relatives without known ALS mutation and no neurodegenerative disease were assigned to the control group of the discovery cohort. Controls in the validation cohort were age and sex-matched individuals without neurodegenerative disease. Demographic characteristics of the discovery (iTRAQ) and validation cohorts (MRM) are listed in Table 1. All patients or their legal representatives gave written informed consent to be included in this study and the Ethics Committee of Ulm University approved the study.

CSF was collected by lumbar puncture during diagnostic workup or recruitment, centrifuged and stored at − 80 °C within 2 h in polypropylene tubes. The ALS functional rating scale-revised (ALSFRS-r) was used to assess disease severity.

Post-mortem spinal cord tissue (cervical/thoracal) was collected from neuropathologically confirmed ALS patients ($n = 4$ c9ALS, $n = 4$ sALS) and non-neurodegenerative controls ($n = 7$) with the following diagnoses: myocardial infarction (3), multiple organ failure, cirrhosis of the liver, aortic dissection and tricuspid insufficiency. Both groups were comparable with respect to age and *post-mortem* delay time (Online Resource 1) but with a predominance of male subjects.

Table 1 Demographic characteristics of the CSF discovery (iTRAQ) and validation (MRM) cohorts

	N (f/m)	Age (year)[a]	Mutations	ALSFRS-r[a]	Age at onset (year)[a]	Duration at LP (months)[a]
Discovery cohort						
Con	16 (11/5)	44 (29–51)	–	–	–	–
Asymptomatic ALS mutation carriers	14 (11/3)	45 (35–52)	8x *C9orf72* 5x *SOD1* 1x *TARDBP*	–	–	–
sALS	12 (5/7)	67 (58–74)	–	39 (32–44)	66 (57-68)	12 (6–21)
gALS	14 (5/9)	62 (53–75)	9x *C9orf72* 5x *SOD1*	39 (35–45)	54 (52–69)	12 (6–27)
p value[b]	< 0.001	< 0.001	–	0.8	0.29	0.82
Validation cohort						
Con	32 (13/19)	63 (57–72)	–	–	–	–
gALS	24 (9/15)	58 (48–66)	16x *C9orf72* 4x *SOD1* 2x *FUS* 1x *NEK* 1x *TBK1*	42 (37–45)	57 (46–64)	10 (7–22)
sALS	61 (21/40)	64 (55–70)	–	42 (36–46)	63 (52–69)	13 (6–27)
p value[b]	0.84	0.06	–	0.91	0.05	0.31

ALS amyotrophic lateral sclerosis, *ALSFRS-r* ALS functional rating scale-revised, *Con* controls, *f* female, *gALS* genetic ALS, *LP* lumbar puncture, *m* male, *sALS* sporadic ALS

[a]Median and interquartile range

[b]Chi-squared test for sex, Mann–Whitney test (two groups) or Kruskal–Wallis test and Dunn´s post hoc test (> 2 groups)

Sample preparation of CSF samples for iTRAQ analysis

A volume of 200 µL CSF per patient was spiked with 100 mM TEAB (triethylammonium bicarbonate), 5 mM TCEP [tris(2-carboxyethyl)phosphine hydrochloride], 10 mM CAA (2-chloroacetamide) and two non-human proteins (flagellin, β-lactoglobulin) to correct for variations during sample preparation. Samples were incubated at 60 °C for 30 min (400 rpm) and then buffer exchanged with 500 mM TEAB using 3 kDa MWCO centrifugal filters (Millipore UFC500396) and finally concentrated to a volume of 25 µL. Trypsin/LysC (Promega) was added with an enzyme-to-substrate ratio of 1:50 and incubated for 16 h at 27 °C. Digestion was stopped by addition of ethanol (75% final) followed by the addition of the respective iTRAQ labels (114, 115, 116 or 117) and incubation for 1 h at 22 °C. The labeling reaction was stopped with formic acid (10% final) and all samples from one iTRAQ 4-plex set were combined and vacuum dried. Pellets were redissolved in 200 µL 1% TFA and fractionated with strong cation exchange (SCX) STAGE Tips (Sigma 66889-U) into six fractions: 125, 160, 220, 300, 450 mM ammonium acetate in 20% acetonitrile/0.5% formic acid and 5% ammonium hydroxide in 80% acetonitrile. After vacuum drying, the fractions were dissolved in 12 µL 0.5% TFA and analyzed by LC–MS/MS.

In each iTRAQ 4-plex sample set, one CSF pool sample was included enabling the comparison of samples from different iTRAQ 4-plex sets and patient groups were equally distributed between the iTRAQ sets.

Sample preparation of human spinal cord tissue

Tissue (whole sections) was lysed in 10 × volume 6 M GndHCl, 10 mM TCEP, 40 mM CAA, 100 mM TEAB by sonication and then heated for 10 min at 95 °C. The lysates were buffer exchanged with 50 mM TEAB (4 ×) using a Microcon-30 kDa centrifugal filter (Millipore #MRCF0R030) and concentrated on the filter. Fifty microliters of trypsin/LysC (Promega #V5072) in 50 mM TEAB were added to a final protein-to-enzyme ratio of 50:1. Digestion was performed for 16 h at 37 °C and 600 rpm. The resulting peptides were centrifuged through the filter and centrifugation was repeated after addition of 50 µL 0.2% TFA to increase recovery. The flow through was pooled and then fractionated by SCX Stage Tips into three fractions (160, 300 mM ammonium acetate and 5% ammonium hydroxide) and further processed as described above.

CSF sample preparation for MRM analysis

A volume of 200 µL CSF was spiked with 0.8 mM TCEP, 4.2 mM CAA, 83 mM TEAB and a mixture of stable isotope labeled peptides and QPrESTs as internal standards (Online Resource 2). The samples were incubated for 10 min at 95 °C and then digested for 16–18 h at 37 °C by addition of 12 µL trypsin/LysC (0.1 µg/µL). Digestion was stopped with 800 µL 1.25% TFA and samples fractionated with SCX Stage tips into six fractions (75, 125, 200, 300, 450 mM ammonium acetate and 5% ammonium hydroxide) and vacuum dried. Samples were re-dissolved in 27.5 µL 6% acetonitrile/0.1% TFA and stored in the autosampler until analysis.

LC–MS/MS analysis of proteomic samples

LC–MS/MS analysis of fractionated iTRAQ-labeled CSF samples and tissue lysates was performed with an UltiMate 3000 RSLCnano system and a Q Exactive mass spectrometer (both Thermo, Dreieich, Germany). The digested peptides were concentrated on a 20 × 0.075 mm, 3 µm trap column (Thermo PepMap100 C18) and separated on a PepMap100 C18, 50 × 0.050 mm, 2 µm analytical column (Thermo) with a column temperature of 60 °C. Mobile phase of the loading pump (trap column) was 0.05% TFA/2%MeOH (flow rate: 5 µL/min) and of the nano pump (analytical column) 4% DMSO/0.1% formic acid (A) and 4%DMSO/76% acetonitrile/0.1% formic acid (B) (flow rate 150 nL/min). Peptides were eluted with a step-gradient from 1-53%B within 340 min (CSF) or 160 min (tissue) and infused into the Q Exactive in positive ion mode (2300 V, 270 °C capillary temperature, S-lens RF level 69). Data were acquired by data-dependent acquisition (CSF: top12, tissue: top15) with the following settings: full MS: resolution 70000, AGC target 3e6, max injection time 120 ms, scan range 400–1400 *m/z*, MS^2: resolution 35,000 (CSF) or 17,500 (tissue), AGC target 1e6, max injection time 120 ms (CSF) or 65 ms (tissue), isolation window 1.6 m/z, NCE 25, dynamic exclusion of 40 s and a fixed first mass of 100 m/z (CSF iTRAQ samples only).

MRM analysis

The fractionated CSF samples were analyzed using an Agilent 1260 HPLC pump, Eksigent MicroLC200 and QTRAP6500 mass spectrometer (AB Sciex, Darmstadt, Germany). Peptides were loaded on a C18 PepMap100, 5 µm, 0.3 × 5 mm trap column (Thermo Fisher Scientific) with a flow rate of 200 µL/min of 1% methanol/0.05% TFA. Separation was performed on an Eksigent HALO Fused-core C18, 2.7 µm, 0.5 × 100 mm column at 40 °C with mobile phase A: 4% DMSO/0.1% formic acid, and mobile phase B: 4% DMSO/96% acetonitrile/0.1% formic acid and a linear gradient from 1 to 30% B within 9.85 min (15 µL/min) and measured in positive ion mode with the QTRAP6500 using the following settings: 5500 V, 175 °C source temperature, curtain gas 40psi, nebulizer gas (gas 1) 20 psi, heater gas (gas 2) 30 psi, declustering potential 100 V. The measured transitions, transition-specific settings and retention times of the scheduled MRM are listed in Online Resource 2.

Proteomic and MRM data analysis

Data from the CSF iTRAQ analysis and spinal cord tissue (label-free proteomics) were analyzed using MaxQuant v1.5.2.8 [13] and Perseus v1.5.2.6 software [14].

MaxQuant settings for CSF iTRAQ samples were reporter ion MS2 with iTRAQ 4-plex and a precursor ion fraction of ≥ 0.75. For tissue samples, the MaxLFQ algorithm [15] was used for protein quantification. For all samples, N-terminal acetylation and methionine oxidation were set as variable modifications and carbamidomethylation on cysteine residues as fixed modification. Trypsin without cleavage before proline was set as the enzyme allowing up to two missed cleavages. Match between runs was selected. The human reference proteome from UniProt (downloaded 07-Jun-2016) was used for protein identification with a peptide and protein FDR of 1%. For CSF iTRAQ samples, the sequences of β-Lactoglobulin and flagellin were also included.

Reporter ion intensities of all proteins were first normalized to the mean intensity of flagellin and β-lactoglobulin (internal standards) within each iTRAQ sample to account for variations during sample preparation and these corrected intensities were then normalized to the intensities of the CSF pool sample within each iTRAQ set to enable comparison between iTRAQ sets.

Normalized reporter ion intensities (CSF iTRAQ) and LFQ intensities (tissue) were log2 transformed for further statistical analysis. Groups were compared by Student´s *t* test (two-tailed) and visualized in Volcano Plots. Correction for multiple testing was performed with permutation-based FDR (0.05). For hierarchical clustering and heatmaps, missing values were replaced by imputation from a normal distribution (width 0.3, down shift 1.8) and intensities were standardized using *z* score. Clustering was performed by Euclidean distance and average linkage. Gene ontology (GO) terms (molecular function, biological process, and cellular compartment) and Reactome pathways were downloaded from UniProt (20-Oct-2017) and enrichment analysis was performed using Fisher´s exact test and Benjamini–Hochberg FDR (0.05). Spearman's rank correlation coefficient was used for correlation analysis of log2 intensities and expected time to disease onset in asymptomatic ALS mutation carriers.

MRM data were analyzed with Skyline v4.2.0 [16] and the mean light-to-heavy (L/H) ratio of 1–3 transitions per peptide (see Online Resource 2) was calculated. For each protein, 1–2 peptides were measured.

Hemoglobin determination

Hemoglobin concentration in the CSF pool sample was determined to estimate blood contamination in the samples of the iTRAQ experiment. Measurement was performed with MRM using a stable isotope-labeled peptide as previously described [17].

Statistics

Statistical analysis of patient characteristics and MRM data was performed with GraphPad Prism v5.00. Groups were compared by Mann–Whitney test (2 groups) or Kruskal–Wallis test and Dunn's post hoc test (>2 groups). All tests were two-tailed and unpaired. Sex distribution was analyzed by Chi-squared test. Spearman´s rank correlation coefficient was used for correlation analysis of MRM data with clinical data. A p value < 0.05 was regarded significant.

Results

Proteomic comparison of CSF samples

To identify novel disease pathways and altered protein levels in the asymptomatic disease stage and in symptomatic gALS and sALS, we used quantitative proteomics to compare CSF samples from asymptomatic ALS mutation carriers ($n = 14$) with age- ($p = 0.79$) and sex- ($p = 0.15$) matched controls ($n = 16$) and sALS ($n = 12$) and gALS ($n = 14$) cases. Characteristics of the individuals are given in Table 1. We established an optimized proteomic workflow for CSF analyses allowing deep proteomic coverage on a single patient level in an applicable volume of CSF (200 µL). We identified a total of 2303 proteins excluding potential contaminants (Online Resource 3). Proteins identified with ≥ 1 unique peptide and quantified in ≥ 3 samples in at least one group (1929 proteins, Online Resource 3) were used for quantitative comparisons. Hemoglobin concentration in the CSF pool was 10.9 ng/mL and samples with blood contamination, indicated by high intensity of hemoglobin (i.e., > 20 × than CSF pool), were excluded (2 Con, 1 Car, 2 sALS, 3 gALS). This threshold was selected based on the experience with other biomarker candidates [17, 18].

We first compared the symptomatic ALS patients (sALS + gALS) with the control group to test whether we are able to detect established biomarkers as verification of our proteomic workflow (Fig. 1a). A total of 32 proteins were significantly changed in the ALS group after correcting for multiple testing (FDR = 0.05) (Table 2).

The well-established ALS markers NEFL, NEFH and CHIT1 [9] were markedly upregulated in ALS whereas another marker of neurodegeneration, tau protein, which has previously been shown to be not altered in ALS [19], was unchanged. This indicates that our proteomic workflow is able to correctly identify disease-relevant, low-abundant proteins in CSF samples. Only three proteins were significantly downregulated in ALS, all belonging to the histone protein family (Fig. 1a and Table 2). Consistent changes (FDR 0.05) were observed for the proteins NEFM, NEFL, NEFH, GPNMB, CHIT1, CHI3L1, CHI3L2, APOC2, APOC3 (increased) and HIST1H2B (decreased) when gALS and sALS were compared with controls separately. Enrichment analysis revealed an over-representation of proteins from the neurofilament, nucleosome, nuclear nucleosome and intermediate-density lipoprotein particle cellular compartments, from the biological process of nucleosome assembly and from pathways involved in transcription, cell cycle and lipoprotein remodeling (Fig. 1b). Hierarchical clustering using the 32 differentially regulated proteins identified two main clusters of patients (Fig. 1c) with cluster 1 containing 86% of ALS patients (19 out of 22) and cluster 2 containing 93% of controls (13 out of 14). The profile of three ALS patients on the left in cluster 1 seemed to differ slightly from the other ALS cases in this cluster but we observed no unique difference of these patients regarding age, sex, genotype, age at disease onset, disease duration at sample collection or ALSFRS-r score. Cluster 1 included only one control patient (specificity 95%) and cluster 2 contained three ALS patients (specificity 81%).

Validation of ALS biomarker candidates by MRM

We established an MRM method for proteins listed in Table 2 to validate their biomarker potential in a second cohort of patients (validation cohort, Table 1). The proteins NEFL, NEFH, CHIT1, CHI3L1 and CHI3L2 have already been studied extensively by our [5, 8, 9, 19, 20] and other groups [4, 7, 10, 11], which is why we did not further validate them here. From the remaining proteins, we could successfully establish an MRM for NEFM (one peptide), GPNMB (one peptide), MAP2 (one peptide), SERPINA3 (two peptides), UCHL1 (two peptides), CAPG (two peptides) and GBA (two peptides). The method showed good performance for all measured proteins/peptides with intra-assay CVs of 3.2–18.3%. Using the established MRM method in the validation cohort, we could confirm significantly increased CSF concentrations of NEFM, GPNMB, MAP2, UCHL1 and CAPG in ALS compared with age- and sex-matched control patients (Fig. 2). For SERPINA3 and GBA no significant difference between groups was observed

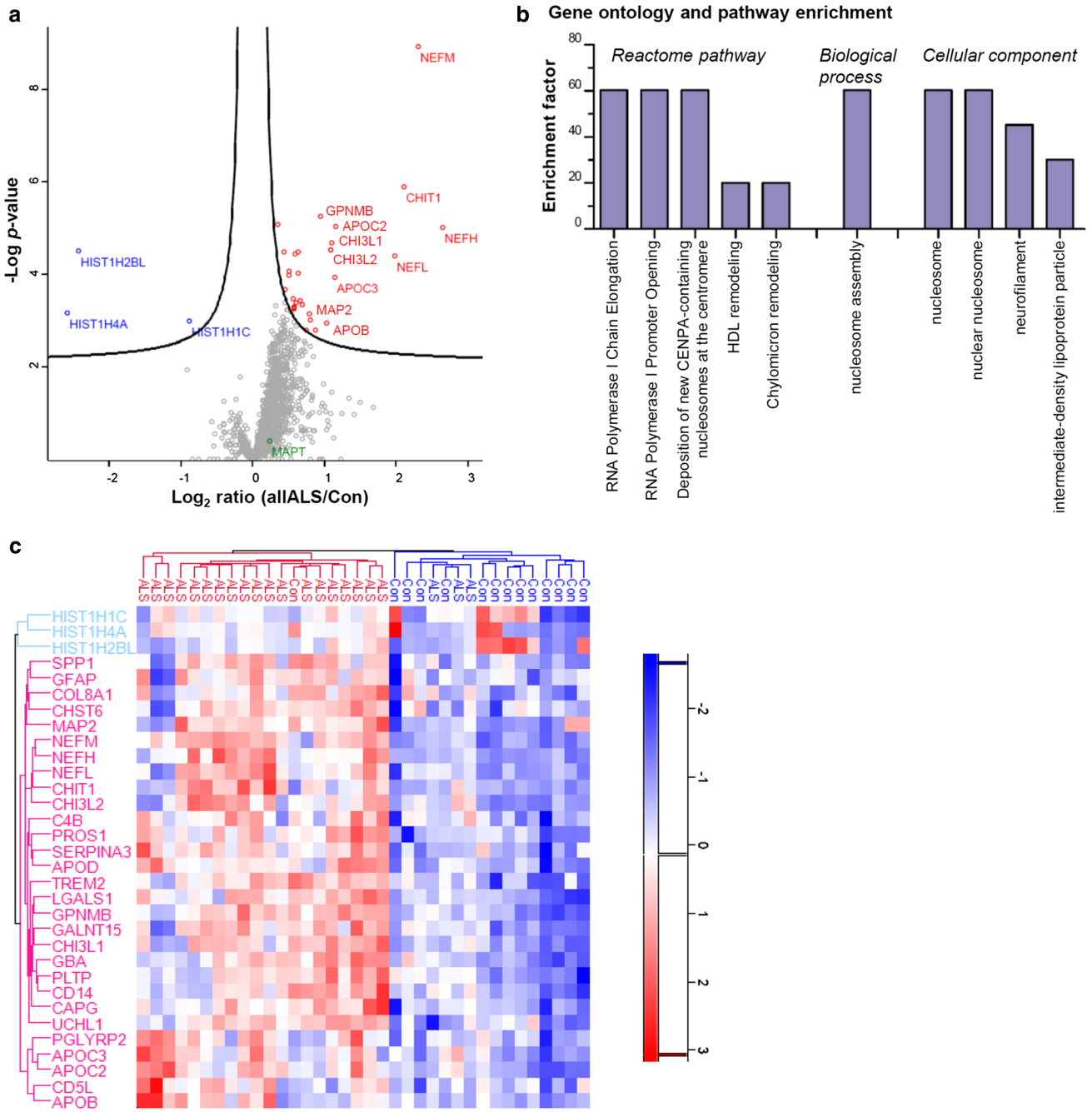

Fig. 1 iTRAQ proteomics in CSF uncovers differences in ALS patients. **a** Volcano plot comparing protein levels in CSF of ALS (allALS, $n=22$) and control patients (Con, $n=14$). Proteins above the bold line are significantly different with a false discovery rate of 0.05. Upregulated proteins in ALS are highlighted in red and downregulated in blue. Tau protein (MAPT) as a negative control is indicated in green. **b** Gene ontology and pathway enrichment analysis of the significantly different proteins in ALS CSF. **c** Heat map and hierarchical clustering of patients using the 32 significantly different proteins in ALS

in the validation cohort ($p>0.18$ for SERPINA3, $p>0.07$ for GBA, Online Resource 4).

There was a significant correlation of age at disease onset and GPNMB ($r=0.52, p<0.001$), MAP2 ($r=0.30, p=0.04$), UCHL1 ($r=0.61, <0.001$) and CAPG ($r=0.40, p<0.001$) but not NEFM ($r=0.28$, $p=0.05$). Disease duration at time of lumbar puncture (LP) showed a significant negative correlation with GPNMB ($r=-0.37, p<0.001$) and MAP2 ($r=-0.30, p=0.04$) but not with the other proteins (UCHL1: $r=-0.20$, $p=0.06$, CAPG: $r=0.01$, $p=0.95$, NEFM: $r=0.12, p=0.39$). The ALSFRS-r score correlated significantly with GPNMB ($r=-0.44, p<0.001$), UCHL1

Table 2 Significantly different proteins (FDR = 0.05) in CSF of ALS patients

Protein/protein group	Gene names	UniProt ID	Log2 ratio	− Log *p* value
Neurofilament heavy polypeptide	NEFH	P12036	2.64	5.02
Neurofilament medium polypeptide	NEFM	P07197	2.30	8.92
Chitotriosidase-1	CHIT1	Q13231	2.10	5.89
Neurofilament light polypeptide	NEFL	P07196	1.98	4.40
Apolipoprotein C-II	APOC2	P02655	1.15	5.03
Apolipoprotein C-III	APOC3	P02656	1.15	3.93
Chitinase-3-like protein 1	CHI3L1	P36222	1.10	4.69
Chitinase-3-like protein 2	CHI3L2	Q15782	1.09	4.52
Apolipoprotein B-100	APOB	P04114	1.03	2.95
Transmembrane glycoprotein NMB	GPNMB	Q14956	0.95	5.26
Collagen alpha-1(VIII) chain	COL8A1	P27658	0.87	2.80
Osteopontin	SPP1	P10451-5	0.81	3.01
Microtubule-associated protein 2	MAP2	P11137	0.79	3.14
CD5 antigen-like	CD5L	O43866	0.75	2.79
Carbohydrate sulfotransferase 6, Carbohydrate sulfotransferase 5	CHST6, CHST5	Q9GZX3, Q9GZS9	0.69	3.34
Monocyte differentiation antigen CD14	CD14	P08571	0.67	3.44
Alpha-1-antichymotrypsin	SERPINA3	P01011	0.64	4.02
Ubiquitin carboxyl-terminal hydrolase isozyme L1	UCHL1	P09936	0.63	4.48
N-acetylmuramoyl-L-alanine amidase	PGLYRP2	Q96PD5	0.61	3.39
Apolipoprotein D	APOD	P05090	0.60	4.45
Glial fibrillary acidic protein	GFAP	P14136	0.58	3.29
Complement C4-B	C4B	P0C0L5	0.58	3.29
Triggering receptor expressed on myeloid cells 2	TREM2	Q9NZC2	0.58	3.27
Polypeptide *N*-acetylgalactosaminyltransferase 15	GALNT15	Q8N3T1	0.57	3.47
Macrophage-capping protein	CAPG	P40121	0.51	3.99
Glucosylceramidase	GBA	P04062	0.50	4.06
Phospholipid transfer protein	PLTP	P55058	0.45	3.66
Galectin-1	LGALS1	P09382	0.44	4.48
Vitamin K-dependent protein S	PROS1	P07225	0.36	5.07
Histone H4	HIST1H4A	P62805	− 2.58	3.16
Histone H2B	HIST1H2B (See Table S1)	(See Table S1)	− 2.42	4.52
Histone H1	HIST1H1C, HIST1H1E, HIST1H1D	P16403, P10412, P16402	− 0.88	3.00

($r = -0.25$, $p = 0.02$) and CAPG ($r = -0.30$, $p = 0.005$) but not MAP2 ($r = -0.22$, $p = 0.13$) and NEFM ($r = 0.01$, $p = 0.97$).

To evaluate the prognostic value of the biomarkers, we correlated the CSF levels with disease duration from LP to death in a subset of patients (gALS and sALS) where time of death was known ($n = 46$). GPNMB ($r = -0.45$, $p = 0.002$, Fig. 3a) but not UCHL1 ($r = -0.28$, $p = 0.06$), CAPG ($r = -013$, $p = 0.38$), MAP2 ($r = -0.23$, $p = 0.24$) and NEFM ($r = -0.02$, $p = 0.91$) showed a significant negative correlation with disease duration until death. ALS patients with CSF GPNMB levels above the median had a significantly shorter survival time than patients with low GPNMB levels (Fig. 3b, c).

Proteomic comparison of asymptomatic ALS mutation carriers and controls

To identify pathways and biomarker candidates that are already changed in the presymptomatic phase of ALS, we compared the CSF proteome of asymptomatic ALS mutation carriers with age- ($p = 0.79$) and sex- ($p = 0.15$) matched controls (Table 1). We did not observe a significant difference between the two groups after correcting for multiple testing (FDR 0.05, Fig. 4). Using less stringent criteria (no FDR), which is common in CSF proteomic analyses to identify pathways and possible biomarker candidates for further validation, we detected 21 proteins with a *t* test *p* value < 0.05 and log2 ratio of ≥ 0.25 (3 proteins) or ≤ − 0.25

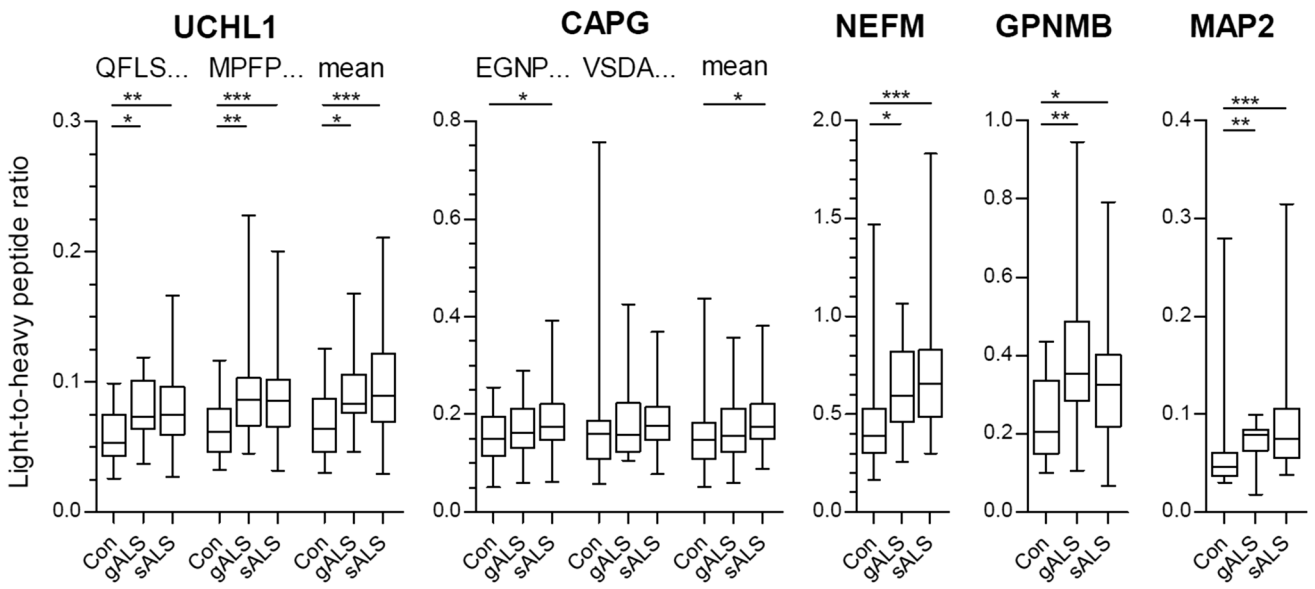

Fig. 2 MRM analysis of biomarker candidates in CSF of the validation cohort. The proteins UCHL1, CAPG, NEFM, GPNMB and MAP2 were measured in cerebrospinal fluid (CSF) by multiple reaction monitoring (MRM). The validation cohort consisted of genetic ALS (gALS, $n=24$), sporadic ALS ($n=61$) and age-matched control patients ($n=32$). The following peptides were measured by MRM: UCHL1 (QFLSETEK, MPFPVNHGASSEDTLLK, mean of both peptides), CAPG (EGNPEEDLTADK, VSDATGQMNLTK, mean of both peptides), NEFM (FEEEAR), MAP2 (DQGGAGEGLVR), GPNMB (IPDENCQINR). Groups were compared by Kruskal–Wallis test and Dunn's post hoc test. Boxes are median and interquartile range, whiskers are min and max. $*p<0.05$, $**p<0.01$, $***p<0.001$

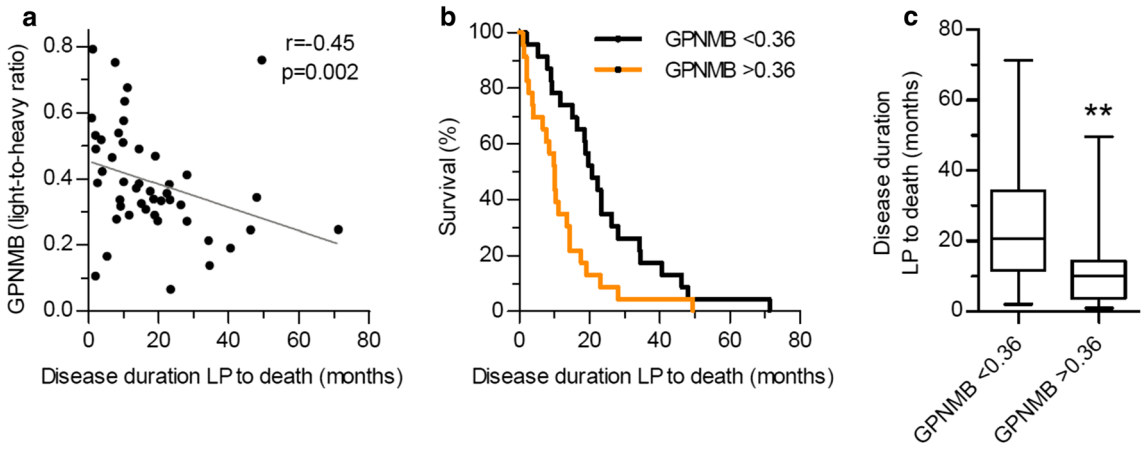

Fig. 3 Higher CSF GPNMB levels in ALS are associated with shorter survival. **a** Correlation analysis (Spearman) of CSF GPNMB levels with disease duration from LP to death in a subset of gALS and sALS patients from the validation cohort ($n=46$) where time of death was available. **b** Kaplan–Meier curve and **c** box blot of ALS patients with CSF GPNMB levels below and above the median concentration show shorter survival of ALS patients with high GPNMB levels (**b** log-rank test, $p=0.006$; **c** Mann–Whitney test, $**p=0.002$). Boxes are median and interquartile range, whiskers are min and max

(18 proteins) (Fig. 4, Online Resource 5). However, enrichment analysis yielded no significantly enriched GO term or pathway in the 21 proteins.

To identify proteins which might change gradually until disease onset, we correlated protein levels of asymptomatic ALS mutation carriers with the expected time to disease onset. This time span was estimated using the parental time of disease onset as described previously [5]. Only proteins with valid intensity values in all these patients (998 proteins) were used for correlation to increase reliability. Sixteen proteins showed a significant correlation (Spearman $r>0.5$ or <-0.5 and $p<0.05$) with the expected time to disease onset and are listed in Table 3. A separate correlation analysis of these proteins in asymptomatic *C9orf72* and *SOD1* mutation carriers is given in Online Resource 5.

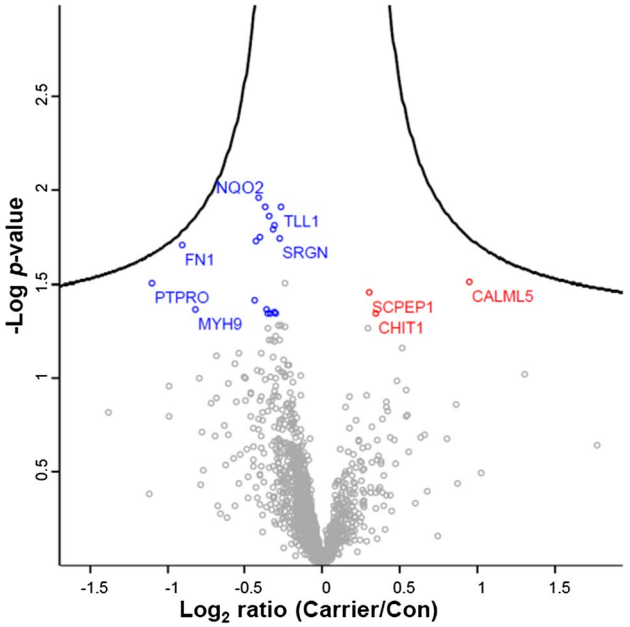

Fig. 4 No significant proteomic changes in CSF of asymptomatic ALS mutation carriers. A volcano plot comparing protein levels in CSF of asymptomatic ALS mutation carriers (Carrier, $n=13$) and age- and sex-matched healthy controls (Con, $n=14$). No significant changes reaching a false discovery rate of 0.05 (bold line) were observed. Proteins with a t test p value <0.05 and a log2 ratio of ≥ 0.25 (red) or ≤ -0.25 (blue) are indicated

Proteomics of *post-mortem* tissue

We used label-free proteomics for comparison of *post-mortem* spinal cord tissue from ALS ($n=8$) and control patients ($n=7$). Age ($p=0.49$), sex ($p=0.57$) and post-mortal delay of tissue collection ($p=0.63$) was comparable between the groups (Online Resource 1). We identified 6810 protein groups after excluding potential contaminants (Online Resource 6) and used those with ≥ 3 valid values in both groups (5115 proteins) for the quantitative comparison. A total of 292 proteins were significantly changed in ALS (139 downregulated, 153 upregulated, FDR = 0.05) as shown in Fig. 5a and Online Resource 7 including ALS-associated proteins such as APOE, hnRNPA2b1, LMNB1, NEFH, NEFL, NEFM, PRPH, SARM1 and SOD2. Table 4 lists the Top 30 proteins (based on fold change and p value). Proteins involved in mRNA splicing (R-HSA-72163, $p=0.016$) and of the catalytic step 2 spliceosome (GO:0071013, $p=0.045$) and neurofilament (GO:0005883, $p=0.030$) compartments were significantly enriched in the group of changed proteins. Based on the 292 differentially regulated proteins, hierarchical clustering identified two main clusters (Fig. 5b). Cluster 1 contained 6 out of 7 control patients (86%) and cluster 2 contained all ALS patients (100%) and one control corresponding to a specificity of 100% and 89% of cluster 1 and 2.

Table 3 List of proteins correlating with expected time to disease onset in asymptomatic ALS mutation carriers and albumin for comparison

Gene names	Spearman r	95% CI	p value	Link to neurological disease
NUCB1	– 0.70	– 0.91 to – 0.22	0.008	AD [50]
NUCB2	– 0.56	– 0.85 to 0.01	0.049	Anxiety/depression [51]
SPARC	– 0.75	– 0.92 to – 0.33	0.003	
ACHE	– 0.56	– 0.85 to 0.01	0.049	ALS [52, 53]
AGT	– 0.59	– 0.87 to – 0.04	0.035	ALS [54, 55] ALS-associated gene (http://alsod.iop.kcl.ac.uk)
CDH8	– 0.58	– 0.86 to – 0.03	0.037	Autism-Gene [56]
CFHR2	0.59	0.05 to 0.87	0.033	
FAT2	– 0.59	– 0.87 to – 0.04	0.035	SCA [57] and autism [58] gene ALS [59]
FSTL5	– 0.65	– 0.89 to – 0.15	0.015	
GALNT18	– 0.62	– 0.88 to – 0.09	0.024	
GMFB	– 0.63	– 0.88 to – 0.10	0.022	PD [60]
MCAM	– 0.59	– 0.87 to – 0.05	0.033	ALS [61, 62], MS [63]
PLD4	– 0.57	– 0.86 to – 0.00	0.044	ALS [64]
SIRPA; SIRPB1	– 0.67	– 0.90 to – 0.17	0.012	AD [65, 66]
SLPI	– 0.68	– 0.90 to – 0.19	0.010	ALS [67], FTLD [68]
UBA52; RPS27A; UBB; UBC	– 0.57	– 0.86 to – 0.00	0.044	ALS and others [69]
ALB	0.23	– 0.39 to 0.70	0.459	na

AD Alzheimer´s disease, ALS amyotrophic lateral sclerosis, *CI* confidence interval, *FTLD* frontotemporal lobar degeneration, *PD* Parkinson's disease, *SCA* spinocerebellar ataxia

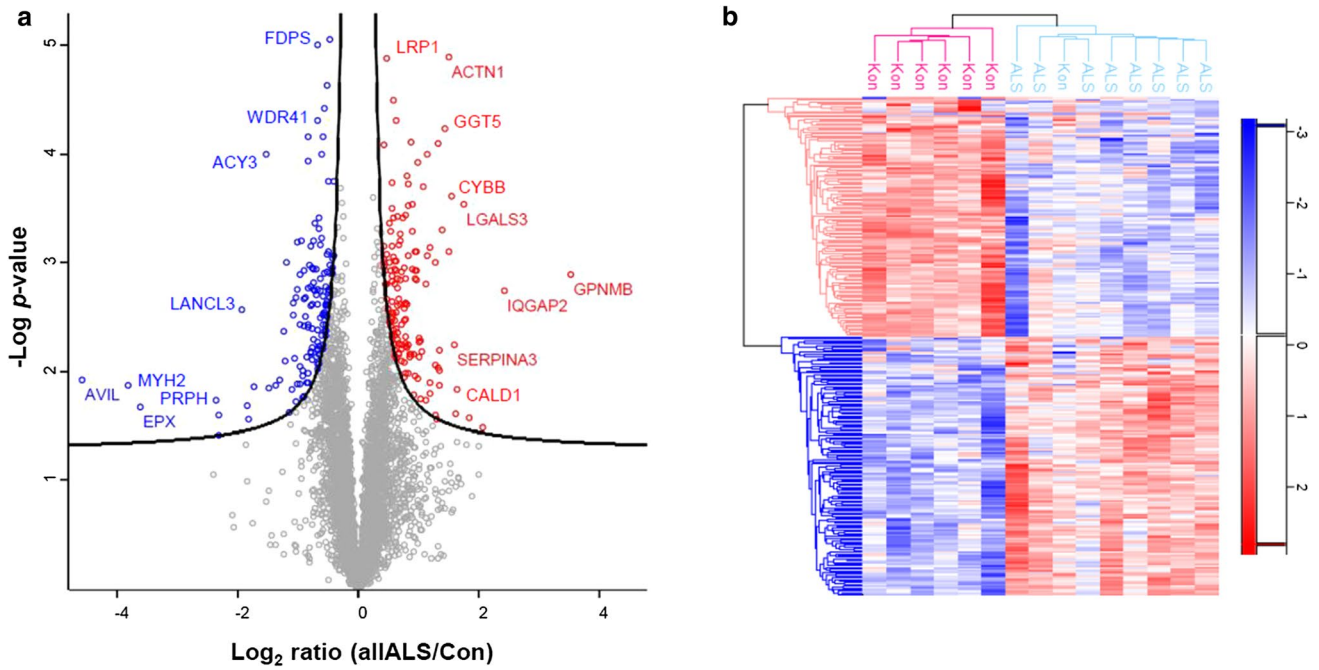

Fig. 5 Label-free proteomics in spinal cord tissue. **a** Volcano plot comparing protein levels in *post-mortem* spinal cord tissue of ALS (allALS, $n=8$) and control patients (Con, $n=7$). Proteins above the bold line are significantly different with a false discovery rate of 0.05. Upregulated proteins in ALS are highlighted in red and downregulated in blue. **b** Heat map and hierarchical clustering of patients using the 292 significantly different proteins

Table 4 Top 30 proteins changed in *post-mortem* spinal cord tissue of ALS

Gene names	− Log *p* value	Log2 ratio	Gene names	− Log *p* value	Log2 ratio
AVIL	1.92	− 4.6	SCYL1	4.63	− 0.5
GPNMB	2.90	3.5	LAMTOR3	4.49	0.6
ACTN1	4.89	1.5	MAST1	4.16	− 0.8
MYH2	1.87	− 3.8	MVD	4.42	− 0.6
GGT5	4.23	1.4	WDR41	4.31	− 0.7
FDPS	5.00	− 0.7	ITGAX	4.11	0.9
CSNK1D;CSNK1E	5.05	− 0.5	SOD2	4.30	0.6
ACY3	3.99	− 1.5	HMGB1; HMGB1P1	3.93	1.0
HLA-DRA	4.10	1.3	CAPG	3.70	1.1
LRP1	4.88	0.5	MVK	3.94	− 0.8
EPX	1.67	− 3.6	KIF3C	4.16	− 0.6
LGALS3	3.54	1.8	S100A6	3.30	1.4
IQGAP2	2.75	2.4	PDPK1; PDPK2	4.00	− 0.6
CYBB	3.61	1.6	ANXA1	3.10	1.5
NAMPT	4.00	1.2	LMNB1	3.80	0.8

Merger of CSF and tissue proteomic data

To strengthen our proteomic results in CSF and spinal cord tissue, we merged both data sets. A change in both CSF and tissue will increase the reliability of these proteins and a comparison of the direction of changes can yield hints to the related mechanisms. Quantitative data in both data sets were available for 1049 protein groups (Online Resource 8) and the changes in ALS (log2 ratios) are plotted in Fig. 6. Proteins changed in CSF and spinal cord tissue can be divided into four groups: (1) downregulated in tissue and upregulated in CSF, (2) upregulated in both, (3) upregulated in tissue, downregulated in CSF and (4) downregulated in both (Online Resource 9). The equation $y=1/6x$ was

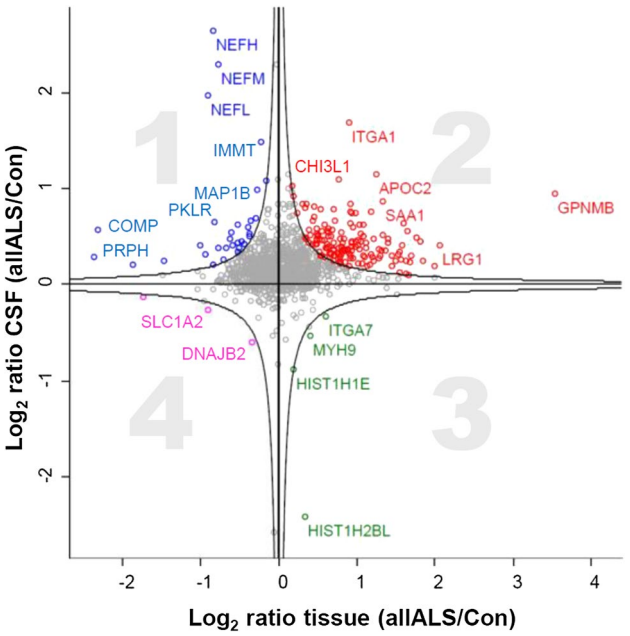

Fig. 6 Merge of CSF and spinal cord proteomic data divide proteins in four groups. The scatter plot shows 1049 proteins which were identified in both CSF and spinal cord tissue. Log2 ratios of protein levels in ALS (allALS) and controls (con) in CSF (x axis) and spinal cord tissue (y axis) are shown. Proteins can be divided into four groups: (1) downregulated in tissue and upregulated in CSF, including the neurofilament proteins (blue) (2) upregulated in both (red) (3) upregulated in tissue, downregulated in CSF (green) and (4) downregulated in both (pink). We used the equation $y = 1/6x$ to separate changed proteins from unchanged

used as threshold for group allocation (Fig. 6). Group (2) is enriched of proteins involved in the complement and coagulation cascade and ECM-receptor interaction and group (3) is enriched with proteins of the nucleosome (GO:0000786) and nuclear nucleosome (GO:0000788). Neurofilaments and other cytoskeletal proteins were present in group (1) but no significant enrichment was observed in groups (1) and (4).

Discussion

In the present study, we investigated proteomic changes in CSF and spinal cord tissue from ALS patients. Both CSF and tissue proteomic data indicate transcriptional pathways to be amongst the most severely affected pathogenic mechanisms relevant for ALS. We identified novel ALS biomarker candidates by CSF proteomics, i.e., UCHL1, CAPG, GPNMB and MAP2, which we confirmed with MRM in an independent validation cohort. Of note, no significant changes were observed in CSF from asymptomatic ALS mutation carriers supporting the hypothesis of a sudden disease onset and a short preclinical disease phase.

We used an optimized proteomic workflow for the analysis of CSF samples which enabled us to detect low abundance proteins in the pg/mL range, such as neurofilaments, synucleins and tau. This was not possible in previous proteomic studies comparing single patient samples [6, 10], as it required large volumes (> 1 mL) of pooled CSF samples [21, 22]. As a proof-of-concept for biomarker identification, we observed increased concentrations of the established markers neurofilaments (NEFL, NEFM, NEFH), CHIT1 and CHI3L1 (also known as YKL40) in CSF of symptomatic ALS patients [5, 8, 10, 19] while the neurodegeneration marker tau, known to be unchanged in ALS [19], was not altered in our proteomic analysis. In addition, we discovered four new ALS biomarker candidates, UCHL1, CAPG, GPNMB and MAP2, which we confirmed in an independent validation cohort of patients with a different method (MRM).

Ubiquitin C-terminal hydrolase-L1 (UCHL1) is highly expressed in the brain and a mutation of the UCHL1 gene was linked to Parkinson's disease [23]. It is thought to be involved in the ubiquitin–proteasome-system due to its ability to cleave ubiquitin from small peptides and its importance for maintaining a pool of free ubiquitin. Recent studies indicate that there are also other functions of UCHL1 that might be more important [24]. UCHL1 is essential for axonal health [25] and the absence of UCHL1 leads to the degeneration of corticospinal neurons [26] and motor axons [25], both of which are cardinal features of ALS [1]. Further support for a role of UCHL1 in ALS pathogenesis comes from studies showing alternative polyadenylation of UCHL1 mRNA in ALS [27] and reduced gene expression of UCHL1 in the motor cortex and spinal cord of ALS patients [28]. In agreement with Lederer et al. [28], we observed reduced UCHL1 protein levels in post-mortem spinal cord tissue of ALS patients ($p < 0.05$) although it did not reach statistical significance after correction for multiple testing. In contrast, UCHL1 levels in CSF were increased in our study in two independent cohorts of ALS patients. The elevated UCHL1 levels might reflect the massive release of UCHL1 into the extracellular space during axonal degeneration, similar to what is expected for neurofilaments, which show a similar profile to UCHL1 (i.e., increased in CSF, reduced in tissue). The reduced UCHL1 levels in tissue might reflect the reduced number of axons/neurons in ALS. Alternatively, the reduced UCHL1 could also be the result of an actively increased UCHL1 secretion from cells resulting in an intracellular UCHL1 deficit inducing axonal degeneration. Massive UCHL1 secretion would also explain the increased CSF levels. Clearly, more research is needed to elucidate the nature of the observed UCHL1 changes but our observations strongly support at least a secondary role of UCHL1 in ALS pathogenesis. Further studies must also address how UCHL1 levels in CSF change during the disease course and

whether it can be used as a biomarker for disease progression and provide an added value to the measurement of neurofilaments.

MAP2 and tau belong to the MAP2/tau family of microtubule-associated proteins (MAPs) which are predominantly located in neurons and are important for regulating the microtubule network, but also have additional functions [29]. In mature neurons, tau is mainly located to the axon [29, 30], whereas MAP2 is restricted to the somatodendritic compartment [29]. In agreement with previous studies [19, 31], we did not observe significant changes of CSF tau in ALS using proteomics, which is a robust yet incompletely understood finding in ALS since the CSF levels of other axonal proteins, i.e., neurofilaments, are markedly increased. In contrast to tau, we observed a significant increase of MAP2 levels in CSF of ALS patients. MAP2 is located to the soma and dendrites of neurons [29] and might therefore be a marker of motor neuron degeneration in ALS which is also supported by immunohistochemistry [32]. It has been shown that MAP2 can induce neurites [29] and elevated CSF levels of MAP2 could reflect a compensatory increase of MAP2 expression in the CNS to compensate axonal loss. However, we did not observe increased MAP2 expression in spinal cord tissue but we cannot rule out that upper motor neurons contribute more to the CSF MAP2 pool. The determination of MAP2 levels in CSF as a marker of motor neuron loss would complement the established axonal markers (i.e., neurofilaments) and allow for a better characterization of the disease in patients and in clinical trials and should be followed up in further studies.

CAPG (capping actin protein, gelsolin-like) and GPNMB (glycoprotein non-metastatic melanoma protein B) have been related to inflammatory processes [33, 34]. Both show significantly increased levels in CSF and spinal cord tissue of ALS patients in our study although changes of CAPG in CSF are moderate limiting its use as a biomarker. In agreement with our observation, increased levels of GPNMB in CSF and a higher number of GPNMB depositions in spinal cord of sALS patients have been described by Tanaka and colleagues [35]. In addition, our data show that gALS patients are characterized by an increase similar to sporadic cases. Increased expression of GPNMB has also been observed in other neurodegenerative diseases such as Alzheimer´s disease [36] and Parkinson´s disease [37, 38] in parallel with markers of so called "disease-associated microglia" such as TREM2 [36]. TREM2 is also increased in CSF of ALS patients in our study supporting the presence of disease-associated microglia. Several studies showed that increased GPNMB expression is neuroprotective [38–40] and might reflect a common compensatory effect to the degenerative process in neurodegenerative diseases. Thus, its suitability as a biomarker for differential diagnosis is questionable, although the extent of GPNMB elevation might be useful, but needs further validation. However, we observed a shorter survival time of ALS patients with high GPNMB levels in CSF and a correlation of GPNMB with the ALSFRS score, a clinical measure of disease severity. This indicates that the amount of GPNMB expression is related to the degree of neurodegeneration and might therefore be a potential marker for disease progression or treatment effects in clinical trials. Since neurofilaments as the most established fluid biomarker for ALS diagnosis do not change over time, a marker of disease progression is urgently needed [4]. Longitudinal studies must clarify whether our observed association of GPNMB levels in CSF with survival and disease duration is a suitable follow-up marker for ALS or if GPNMB is a prognostic marker but stays constant during the disease process similar to neurofilaments.

We performed enrichment analyses in the CSF and spinal cord tissue data sets to identify new ALS-related pathways. We used the list of all identified proteins in each data set for comparison in enrichment analyses instead of the whole human proteome to avoid sampling bias [41] and increase the robustness of the results which is why the number of significantly enriched pathways is lower compared with other studies. A consistent finding in the CSF and tissue data sets was the enrichment of pathways involved in transcription in ALS patients.

In spinal cord tissue of ALS patients, we found a significant enrichment of proteins involved in mRNA splicing and the catalytic step 2 spliceosome. This included the proteins HNRNPC, SNRPB2, MAGOH, SNRPE, LSM3, SNRPD2, SRSF1 and LSM7 which were significantly upregulated in ALS. Our data are in agreement with transcriptomic data showing alterations of mRNA splicing in ALS motor neurons [42]. Alternative splicing is one of several proposed functions of TDP-43, which forms the characteristic insoluble protein aggregates in ALS, and therefore, splicing defects have been proposed as a possible contributor to ALS pathogenesis [43]. Recent evidence indicates that TDP-43 loss-of-function in splicing repression due to cytoplasmic mislocalization is a key pathological event in ALS [44]. It leads to the decline of STMN2 which is important for axonal growth and regeneration [45]. In agreement with our observation that several proteins of the spliceosome are affected in ALS spinal cord, a recent hypothesis assumes that there is an equilibrium between several RNA-binding proteins including TDP-43 and splicing defects appear if any of these proteins becomes dysfunctional [46, 47]. This theory also provides a link between sporadic and genetic ALS since splicing defects seem to be a common mechanism [46]. Cytoplasmic mislocalization of the splicing factor SNRPB2 due to the *C9orf72*-associated dipeptide repeat proteins were accompanied with splicing defects in induced pluripotent stem cell (iPSC)-derived motor neurons of c9ALS patients [48]. SNRPB2 was also upregulated in ALS spinal cord

tissue in our study which would be in agreement with a cytoplasmic accumulation of the protein. In addition, there is evidence that histones which we found dysregulated in ALS CSF also play a role in pre-mRNA splicing [49] providing a link between our two datasets. The ALS-related CSF/tissue profile of histones in our study (reduced in CSF, increased in tissue) would also be in support of a cytoplasmic accumulation of these histones but needs confirmation.

One of the major aims of our study was to discover early changes in ALS pathogenesis by proteomic investigation of CSF from asymptomatic ALS mutation carriers. However, we did not find any significant changes in these individuals. Even when we increased sensitivity by excluding correction for multiple testing and using all proteins with a p value < 0.05, there were no significantly enriched pathways or functional annotations. This observation is in agreement with previous studies from our and other groups which observed changes of neurofilaments [5, 11] or inflammatory markers [8] shortly before or simultaneously with disease onset supporting a short preclinical phase of ALS. On the other hand, neurodegeneration might start suddenly when the straw of factors that have accumulated over years breaks the camel´s back. The asymptomatic ALS mutation carriers investigated here might be in different preclinical ALS stages and combining them in a single group might not be appropriate to identify proteins that change gradually over time. Therefore, we correlated protein levels in CSF with the expected time to disease onset in the asymptomatic ALS mutation carriers. We observed 16 proteins showing a strong and significant correlation with time to disease onset. Serum albumin was used as a negative control and did not correlate. Eleven of the 16 proteins have already been associated with neurologic diseases before and seven even with ALS (ACHE, AGT, FAT2, MCAM, PLD4, SLPI and ubiquitin). However, although many of the proteins in this list have already been linked to neurodegenerative diseases in other studies, the performance of the correlation analysis is based on several assumptions which might not necessarily be true such as the gradual change of protein levels over time, the expected time of disease onset and a 100% penetrance of gene mutations. Therefore, this list should be regarded as a selection of candidates which might be promising for further validation.

A limitation of our study is the significantly lower age of the controls compared with the ALS group in the CSF iTRAQ experiment. Since the sample number that can be used for iTRAQ proteomic studies is limited due to the required analysis time and costs, we included only one group of controls that were age- and sex-matched to the asymptomatic ALS mutation carriers to maximize the sensitivity and specificity for this comparison. We assume that the age-difference between ALS and controls is also responsible for the skewed appearance of the Volcano plot reflecting the known age-dependent increase of blood-derived proteins in CSF. However, the fact that we successfully identified already established ALS biomarkers such as neurofilaments and no change of tau as a negative control shows that the age-effect is less pronounced for CNS-derived proteins and the comparison is valid to identify new biomarker candidates. In addition, to overcome this limitation, we validated the new biomarker candidates by MRM in a separate and larger cohort of patients who were matched for age and sex and confirmed our observations from the iTRAQ experiment.

In conclusion, our proteomic data in both CSF and spinal cord of ALS patients indicate that pathways involved in transcription are attractive candidates for the pathogenesis of ALS. We successfully identified and validated UCHL1, MAP2 and GPNMB as new ALS biomarker candidates in CSF which might also be useful as follow-up markers for disease progression and clinical trials. The lack of significant proteomic changes in CSF of asymptomatic ALS mutation carriers is in support of a short preclinical phase of ALS.

Acknowledgements We are grateful to all the patients who participated in this study. We thank Stephen Meier for his excellent technical assistance. QPrESTs for MRM analysis were partly provided by Atlas Antibodies AB (Bromma, Sweden).

Author contributions Conception and design: PO, MO. Data acquisition, analysis and interpretation: PO, PW, DRT, JHW, ACL, MO. All authors substantially revised the manuscript.

Funding This study was supported by the EU Joint Programme-Neurodegenerative Diseases networks SOPHIA (01ED1202A), BiomarkAPD (01ED1203F) and PreFrontAls (01ED1512), the German Federal Ministry of Education and Research (FTLDc 01GI1007A, MND-Net 01GM1103A), the EU (NADINE 246513, FAIR-PARK II 633190), the German Research Foundation/DFG (SFB1279), the foundation of the state Baden-Württemberg (D.3830), Boehringer Ingelheim Ulm University BioCenter (D.5009) and the Thierry Latran Foundation. DRT receives grants from Fonds Wetenschappelijk Onderzoek (FWO) G0F8516 N, and C1-internal funds from KU Leuven (C14-17-107). The funding sources were not involved in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Compliance with ethical standards

Conflict of interest PO, PW, JHW, ACL and MO report no conflict of interest. DRT received consultant honorary from GE-Healthcare (UK), and Covance Laboratories (UK), speaker honorary from Novartis Pharma AG (Switzerland), travel reimbursement from GE-Healthcare (UK) and UCB (Belgium) and collaborated with Novartis Pharma AG (Switzerland), Probiodrug (Germany), GE-Healthcare (UK), and Janssen Pharmaceutical Companies (Belgium).

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee (Ethics Committee of Ulm) and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

References

1. Brown RH, Al-Chalabi A (2017) Amyotrophic lateral sclerosis. N Engl J Med 377:162–172. https://doi.org/10.1056/NEJMra1603471
2. Petrov D, Mansfield C, Moussy A, Hermine O (2017) ALS clinical trials review: 20 years of failure Are we any closer to registering a new treatment. Front Aging Neurosci 9:68. https://doi.org/10.3389/fnagi.2017.00068
3. Ludolph A, Drory V, Hardiman O, Nakano I, Ravits J, Robberecht W et al (2015) A revision of the El Escorial criteria—2015. Amyotroph Lateral Scler Frontotemporal Degener 16:291–292. https://doi.org/10.3109/21678421.2015.1049183
4. Lu C-H, Macdonald-Wallis C, Gray E, Pearce N, Petzold A, Norgren N et al (2015) Neurofilament light chain: a prognostic biomarker in amyotrophic lateral sclerosis. Neurology 84:2247–2257
5. Weydt P, Oeckl P, Huss A, Müller K, Volk AE, Kuhle J et al (2016) Neurofilament levels as biomarkers in asymptomatic and symptomatic familial amyotrophic lateral sclerosis. Ann Neurol 79:152–158. https://doi.org/10.1002/ana.24552
6. Varghese AM, Sharma A, Mishra P, Vijayalakshmi K, Harsha HC, Sathyaprabha TN et al (2013) Chitotriosidase—a putative biomarker for sporadic amyotrophic lateral sclerosis. Clin Proteom 10:19. https://doi.org/10.1186/1559-0275-10-19
7. Chen X, Chen Y, Wei Q, Ou R, Cao B, Zhao B et al (2016) Assessment of a multiple biomarker panel for diagnosis of amyotrophic lateral sclerosis. BMC Neurol 16:173. https://doi.org/10.1186/s12883-016-0689-x
8. Oeckl P, Weydt P, Steinacker P, Anderl-Straub S, Nordin F, Volk AE et al (2019) Different neuroinflammatory profile in amyotrophic lateral sclerosis and frontotemporal dementia is linked to the clinical phase. J Neurol Neurosurg Psychiatry 90:4–10. https://doi.org/10.1136/jnnp-2018-318868
9. Steinacker P, Verde F, Fang L, Feneberg E, Oeckl P, Roeber S et al (2018) 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 89:239–247
10. Thompson AG, Gray E, Thézénas M-L, Charles PD, Evetts S, Hu MT et al (2018) Cerebrospinal fluid macrophage biomarkers in amyotrophic lateral sclerosis. Ann Neurol 83:258–268. https://doi.org/10.1002/ana.25143
11. Benatar M, Wuu J, Andersen PM, Lombardi V, Malaspina A (2018) Neurofilament light: a candidate biomarker of presymptomatic amyotrophic lateral sclerosis and phenoconversion. Ann Neurol 84:130–139. https://doi.org/10.1002/ana.25276
12. Brooks BR, Miller RG, Swash M, Munsat TL (2000) El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph Lateral Scler Other Motor Neuron Disord 1:293–299
13. Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26:1367–1372. https://doi.org/10.1038/nbt.1511
14. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T et al (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods 13:731–740. https://doi.org/10.1038/nmeth.3901
15. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteom 13:2513–2526. https://doi.org/10.1074/mcp.M113.031591
16. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B et al (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 26:966–968
17. Oeckl P, Metzger F, Nagl M, von Arnim CAF, Halbgebauer S, Steinacker P et al (2016) Alpha-, Beta-, and gamma-synuclein quantification in cerebrospinal fluid by multiple reaction monitoring reveals increased concentrations in Alzheimer's and Creutzfeldt-Jakob disease but no alteration in synucleinopathies. Mol Cell Proteom 15:3126–3138. https://doi.org/10.1074/mcp.M116.059915
18. Hong Z, Shi M, Chung KA, Quinn JF, Peskind ER, Galasko D et al (2010) DJ-1 and alpha-synuclein in human cerebrospinal fluid as biomarkers of Parkinson's disease. Brain 133:713–726
19. Steinacker P, Feneberg E, Weishaupt J, Brettschneider J, Tumani H, Andersen PM et al (2016) Neurofilaments in the diagnosis of motoneuron diseases: a prospective study on 455 patients. J Neurol Neurosurg Psychiatry 87:12–20
20. Oeckl P, Jardel C, Salachas F, Lamari F, Andersen PM, Bowser R et al (2016) Multicenter validation of CSF neurofilaments as diagnostic biomarkers for ALS. Amyotroph Lateral Scler Frontotemporal Degener 17:404–413
21. Macron C, Lane L, Núñez Galindo A, Dayon L (2018) Deep dive on the proteome of human cerebrospinal fluid: a valuable data resource for biomarker discovery and missing protein identification. J Proteome Res 17:4113–4126. https://doi.org/10.1021/acs.jproteome.8b00300
22. Zhang Y, Guo Z, Zou L, Yang Y, Zhang L, Ji N et al (2015) A comprehensive map and functional annotation of the normal human cerebrospinal fluid proteome. J Proteom 119:90–99. https://doi.org/10.1016/j.jprot.2015.01.017
23. Leroy E, Boyer R, Auburger G, Leube B, Ulm G, Mezey E et al (1998) The ubiquitin pathway in Parkinson's disease. Nature 395:451–452. https://doi.org/10.1038/26652
24. Bishop P, Rocca D, Henley JM (2016) Ubiquitin C-terminal hydrolase L1 (UCH-L1): structure, distribution and roles in brain function and dysfunction. Biochem J 473:2453–2462. https://doi.org/10.1042/BCJ20160082
25. Genç B, Jara JH, Schultz MC, Manuel M, Stanford MJ, Gautam M et al (2016) Absence of UCHL 1 function leads to selective motor neuropathy. Ann Clin Transl Neurol 3:331–345. https://doi.org/10.1002/acn3.298
26. Jara JH, Genç B, Cox GA, Bohn MC, Roos RP, Macklis JD et al (2015) Corticospinal motor neurons are susceptible to increased ER stress and display profound degeneration in the absence of UCHL1 function. Cereb Cortex 25:4259–4272. https://doi.org/10.1093/cercor/bhu318
27. Patel R, Brophy C, Hickling M, Neve J, Furger A (2019) Alternative cleavage and polyadenylation of genes associated with protein turnover and mitochondrial function are deregulated in Parkinson's, Alzheimer's and ALS disease. BMC Med Genom 12:60. https://doi.org/10.1186/s12920-019-0509-4
28. Lederer CW, Torrisi A, Pantelidou M, Santama N, Cavallaro S (2007) Pathways and genes differentially expressed in the motor cortex of patients with sporadic amyotrophic lateral sclerosis. BMC Genom 8:26. https://doi.org/10.1186/1471-2164-8-26
29. Dehmelt L, Halpain S (2004) The MAP2/Tau family of microtubule-associated proteins. Genome Biol 6:204. https://doi.org/10.1186/gb-2004-6-1-204
30. Xia D, Gutmann JM, Götz J (2016) Mobility and subcellular localization of endogenous, gene-edited Tau differs from that of over-expressed human wild-type and P301L mutant Tau. Sci Rep 6:29074. https://doi.org/10.1038/srep29074
31. Grossman M, Elman L, McCluskey L, McMillan CT, Boller A, Powers J et al (2014) Phosphorylated tau as a candidate biomarker for amyotrophic lateral sclerosis. JAMA Neurol 71:442. https://doi.org/10.1001/jamaneurol.2013.6064

32. Kikuchi H, Doh-ura K, Kawashima T, Kira J, Iwaki T (1999) Immunohistochemical analysis of spinal cord lesions in amyotrophic lateral sclerosis using microtubule-associated protein 2 (MAP2) antibodies. Acta Neuropathol 97:13–21
33. van der Lienden M, Gaspar P, Boot R, Aerts J, van Eijk M (2018) Glycoprotein non-metastatic protein B: an emerging biomarker for lysosomal dysfunction in macrophages. Int J Mol Sci 20:66. https://doi.org/10.3390/ijms20010066
34. Witke W, Li W, Kwiatkowski DJ, Southwick FS (2001) Comparisons of CapG and gelsolin-null macrophages. J Cell Biol 154:775–784. https://doi.org/10.1083/jcb.200101113
35. Tanaka H, Shimazawa M, Kimura M, Takata M, Tsuruma K, Yamada M et al (2012) The potential of GPNMB as novel neuroprotective factor in amyotrophic lateral sclerosis. Sci Rep 2:573. https://doi.org/10.1038/srep00573
36. Hüttenrauch M, Ogorek I, Klafki H, Otto M, Stadelmann C, Weggen S et al (2018) Glycoprotein NMB: a novel Alzheimer's disease associated marker expressed in a subset of activated microglia. Acta Neuropathol Commun 6:108. https://doi.org/10.1186/s40478-018-0612-3
37. Moloney EB, Moskites A, Ferrari EJ, Isacson O, Hallett PJ (2018) The glycoprotein GPNMB is selectively elevated in the substantia nigra of Parkinson's disease patients and increases after lysosomal stress. Neurobiol Dis 120:1–11. https://doi.org/10.1016/j.nbd.2018.08.013
38. Neal ML, Boyle AM, Budge KM, Safadi FF, Richardson JR (2018) The glycoprotein GPNMB attenuates astrocyte inflammatory responses through the CD44 receptor. J Neuroinflamm 15:73. https://doi.org/10.1186/s12974-018-1100-1
39. Nagahara Y, Shimazawa M, Ohuchi K, Ito J, Takahashi H, Tsuruma K et al (2017) GPNMB ameliorates mutant TDP-43-induced motor neuron cell death. J Neurosci Res 95:1647–1665. https://doi.org/10.1002/jnr.23999
40. Nagahara Y, Shimazawa M, Tanaka H, Ono Y, Noda Y, Ohuchi K et al (2015) Glycoprotein nonmetastatic melanoma protein B ameliorates skeletal muscle lesions in a SOD1 ^{G93A} mouse model of amyotrophic lateral sclerosis. J Neurosci Res 93:1552–1566. https://doi.org/10.1002/jnr.23619
41. Timmons JA, Szkop KJ, Gallagher IJ (2015) Multiple sources of bias confound functional enrichment analysis of globalomics data. Genome Biol 16:186. https://doi.org/10.1186/s13059-015-0761-7
42. Highley JR, Kirby J, Jansweijer JA, Webb PS, Hewamadduma CA, Heath PR et al (2014) Loss of nuclear TDP-43 in amyotrophic lateral sclerosis (ALS) causes altered expression of splicing machinery and widespread dysregulation of RNA splicing in motor neurones. Neuropathol Appl Neurobiol 40:670–685. https://doi.org/10.1111/nan.12148
43. Gao J, Wang L, Huntley ML, Perry G, Wang X (2018) Pathomechanisms of TDP-43 in neurodegeneration. J Neurochem 146:7–20. https://doi.org/10.1111/jnc.14327
44. Donde A, Sun M, Ling JP, Braunstein KE, Pang B, Wen X et al (2019) Splicing repression is a major function of TDP-43 in motor neurons. Acta Neuropathol 138:813–826. https://doi.org/10.1007/s00401-019-02042-8
45. Klim JR, Williams LA, Limone F, Guerra San Juan I, Davis-Dusenbery BN, Mordes DA et al (2019) ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. Nat Neurosci 22:167–179. https://doi.org/10.1038/s41593-018-0300-4
46. Conlon EG, Fagegaltier D, Agius P, Davis-Porada J, Gregory J, Hubbard I et al (2018) Unexpected similarities between C9ORF72 and sporadic forms of ALS/FTD suggest a common disease mechanism. Elife 7:e37754. https://doi.org/10.7554/eLife.37754
47. Gitler AD, Fryer JD (2018) A matter of balance. Elife 7:e40034. https://doi.org/10.7554/eLife.40034
48. Yin S, Lopez-Gonzalez R, Kunz RC, Gangopadhyay J, Borufka C, Gygi SP et al (2017) Evidence that C9ORF72 dipeptide repeat proteins associate with U2 snRNP to cause mis-splicing in ALS/FTD patients. Cell Rep 19:2244–2256. https://doi.org/10.1016/j.celrep.2017.05.056
49. Neves LT, Douglass S, Spreafico R, Venkataramanan S, Kress TL, Johnson TL (2017) The histone variant H2A.Z promotes efficient cotranscriptional splicing in *S. cerevisiae*. Genes Dev 31:702–717. https://doi.org/10.1101/gad.295188.116
50. Lin P, Li F, Zhang Y, Huang H, Tong G, Farquhar MG, Xu H (2006) Calnuc binds to Alzheimer's beta-amyloid precursor protein and affects its biogenesis. J Neurochem 100:1505–1514. https://doi.org/10.1111/j.1471-4159.2006.04336.x
51. Kühne S, Schalla M, Friedrich T, Kobelt P, Goebel-Stengel M, Long M et al (2018) Nesfatin-130-59 injected intracerebroventricularly increases anxiety, depression-like behavior, and anhedonia in normal weight rats. Nutrients 10:1889. https://doi.org/10.3390/nu10121889
52. Berger ML, Veitl M, Malessa S, Sluga E, Hornykiewicz O (1992) Cholinergic markers in ALS spinal cord. J Neurol Sci 108:114–117
53. Campanari M-L, García-Ayllón M-S, Ciura S, Sáez-Valero J, Kabashi E (2016) Neuromuscular junction impairment in amyotrophic lateral sclerosis: reassessing the role of acetylcholinesterase. Front Mol Neurosci 9:160. https://doi.org/10.3389/fnmol.2016.00160
54. Kawajiri M, Mogi M, Higaki N, Tateishi T, Ohyagi Y, Horiuchi M et al (2009) Reduced angiotensin II levels in the cerebrospinal fluid of patients with amyotrophic lateral sclerosis. Acta Neurol Scand 119:341–344. https://doi.org/10.1111/j.1600-0404.2008.01099.x
55. Lin F-C, Tsai C-P, Kuang-Wu Lee J, Wu M-T, Tzu-Chi Lee C (2015) Angiotensin-Converting Enzyme Inhibitors and Amyotrophic Lateral Sclerosis Risk. JAMA Neurol 72:40. https://doi.org/10.1001/jamaneurol.2014.3367
56. Pagnamenta AT, Khan H, Walker S, Gerrelli D, Wing K, Bonaglia MC et al (2011) Rare familial 16q21 microdeletions under a linkage peak implicate cadherin 8 (CDH8) in susceptibility to autism and learning disability. J Med Genet 48:48–54. https://doi.org/10.1136/jmg.2010.079426
57. Nibbeling EAR, Duarri A, Verschuuren-Bemelmans CC, Fokkens MR, Karjalainen JM, Smeets CJLM et al (2017) Exome sequencing and network analysis identifies shared mechanisms underlying spinocerebellar ataxia. Brain 140:2860–2878. https://doi.org/10.1093/brain/awx251
58. Butler M, Rafi S, Hossain W, Stephan D, Manzardo A (2015) Whole exome sequencing in females with autism implicates novel and candidate genes. Int J Mol Sci 16:1312–1335. https://doi.org/10.3390/ijms16011312
59. Bereman MS, Beri J, Enders JR, Nash T (2018) Machine learning reveals protein signatures in CSF and plasma fluids of clinical value for ALS. Sci Rep 8:16334. https://doi.org/10.1038/s41598-018-34642-x
60. Werner CJ, Heyny-von Haussen R, Mall G, Wolf S (2008) Proteome analysis of human substantia nigra in Parkinson's disease. Proteome Sci 6:8. https://doi.org/10.1186/1477-5956-6-8
61. Cox LE, Ferraiuolo L, Goodall EF, Heath PR, Higginbottom A, Mortiboys H et al (2010) Mutations in CHMP2B in lower motor neuron predominant amyotrophic lateral sclerosis (ALS). PLoS One 5:e9872. https://doi.org/10.1371/journal.pone.0009872
62. Garbuzova-Davis S, Woods RL, Louis MK, Zesiewicz TA, Kuzmin-Nichols N, Sullivan KL et al (2010) Reduction of circulating endothelial cells in peripheral blood of ALS patients. PLoS One 5:e10614. https://doi.org/10.1371/journal.pone.0010614
63. Larochelle C, Lécuyer M-A, Alvarez JI, Charabati M, Saint-Laurent O, Ghannam S et al (2015) Melanoma cell adhesion molecule-positive CD8 T lymphocytes mediate central nervous system

inflammation. Ann Neurol 78:39–53. https://doi.org/10.1002/ana.24415
64. Figueroa-Romero C, Hur J, Bender DE, Delaney CE, Cataldo MD, Smith AL et al (2012) Identification of epigenetically altered genes in sporadic amyotrophic lateral sclerosis. PLoS ONE 7:e52672. https://doi.org/10.1371/journal.pone.0052672
65. Foveau B, Correia AS, Hébert SS, Rainone S, Potvin O, Kergoat M-J et al (2019) Stem cell-derived neurons as cellular models of sporadic Alzheimer's disease. J Alzheimer's Dis 67:893–910. https://doi.org/10.3233/jad-180833
66. Gaikwad S, Larionov S, Wang Y, Dannenberg H, Matozaki T, Monsonego A et al (2009) Signal regulatory protein-β1. Am J Pathol 175:2528–2539. https://doi.org/10.2353/ajpath.2009.090147
67. Lilo E, Wald-Altman S, Solmesky LJ, Ben Yaakov K, Gershoni-Emek N, Bulvik S et al (2013) Characterization of human sporadic ALS biomarkers in the familial ALS transgenic mSOD1G93A mouse model. Hum Mol Genet 22:4720–4725. https://doi.org/10.1093/hmg/ddt325
68. Ghidoni R, Flocco R, Paterlini A, Glionna M, Caruana L, Tonoli E et al (2014) Secretory leukocyte protease inhibitor protein regulates the penetrance of frontotemporal lobar degeneration in progranulin mutation carriers. J Alzheimers Dis 38:533–539. https://doi.org/10.3233/JAD-131163
69. Atkin G, Paulson H (2014) Ubiquitin pathways in neurodegenerative disease. Front Mol Neurosci 7:63. https://doi.org/10.3389/fnmol.2014.00063

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.