

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
12 January 2012 (12.01.2012)

PCT

(10) International Publication Number
WO 2012/004276 A2

- (51) **International Patent Classification:**
G01N 33/68 (2006.01)
- (21) **International Application Number:**
PCT/EP201 1/061344
- (22) **International Filing Date:**
5 July 2011 (05.07.2011)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/361,730 6 July 2010 (06.07.2010) US
- (71) **Applicants (for all designated States except US):** FONDAZIONE TELETHON [IT/IT]; Via Carlo Spinola, 16, 1-00154 Roma (IT). FONDAZIONE SALVATORE MAUGERI CLINICA DEL LAVORO E DELLA RIABILITAZIONE [IT/IT]; Via Salvatore Maugeri, 4, 1-27 100 Pavia (IT).
- (72) **Inventors; and**
- (75) **Inventors/Applicants (for US only):** BONETTO, Valentina [IT/IT]; c/o Istituto di Ricerche Farmacologiche "Mario Negri" DTI, Via La Masa, 19, 1-20156 Milano (IT). BENDOTTI, Caterina [IT/IT]; c/o Istituto di Ricerche Farmacologiche "Mario Negri" DTI, Via La Masa, 19, 1-20156 Milano (IT). MORA, Gabriele [IT/IT]; c/o IRCCS Fondazione S. Maugeri, Via Camaldoli, 64, 1-20138 Milano (IT). CORBO, Massimo [IT/IT]; c/o Centra Clinico Nemo, A.O. Niguarda, Via Ca' Granda, 1-201 62 Milano (IT).
- (74) **Agents:** CAPASSO, Olga et al; c/o De Simone and Partners S.p.A., Via Vincenzo Bellini, 20, 1-001 98 Rome (IT).
- (81) **Designated States (unless otherwise indicated, for every kind of national protection available):** AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) **Designated States (unless otherwise indicated, for every kind of regional protection available):** ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))
- with sequence listing part of description (Rule 5.2(a))

(54) **Title:** MULTIPROTEIN BIOMARKERS OF AMYOTROPHIC LATERAL SCLEROSIS IN PERIPHERAL BLOOD MONONUCLEAR CELLS, DIAGNOSTIC METHODS AND KITS

(57) **Abstract:** The present invention refers to a method and to a diagnostic kit to diagnose and/or to make prognostic predictions and/or to monitor the efficacy of a therapy of Amyotrophic lateral sclerosis (ALS) in a subject.



WO 2012/004276 A2

MULTIPROTEIN BIOMARKERS OF AMYOTROPHIC LATERAL SCLEROSIS IN PERIPHERAL BLOOD MONONUCLEAR CELLS, DIAGNOSTIC METHODS AND KITS

5 **Field of the invention**

Amyotrophic lateral sclerosis (ALS) is a fatal progressive motor neuron disease, for which there are still no diagnostic/prognostic test and therapy. Specific molecular biomarkers would facilitate clinical studies and speed up the development of effective treatments. Authors used a proteome-based approach to identify in peripheral blood mononuclear cells (PBMC) a panel of protein biomarkers that are closely associated with ALS and can distinguish, with high discriminatory power, between two levels of disease severity (90%), ALS patients from healthy controls (98%), and from patients with neurological disorders that may resemble ALS (91%). These biomarkers are easily measurable in large-scale immunoassays and thus suitable to develop a helpful test in clinical practice. The protein profile changes detected in PBMC of ALS patients are suggestive of possible pathogenic mechanisms such as endoplasmic reticulum stress, nitrate stress, disturbances in redox regulation and RNA processing, previously detected in spinal cord of patients and animal models. This overlap endorses the use of PBMC for *in vitro* mechanistic studies of ALS and reinforces the theory that ALS is a multi-cellular disease. Finally, this is the first study that provides a highly feasible strategy, potentially applicable to several neurological diseases, to identify and validate multiprotein biomarkers in PBMC, useful both for clinical and experimental studies.

Prior art

Amyotrophic lateral sclerosis (ALS) is an incurable neurodegenerative disorder of unknown cause arising from progressive degeneration of motor neurons and resulting in paralysis and death, usually within 2-4 years from diagnosis (1). Its incidence is between 1.5 and 2.5 per 100.000 per year (2): approximately 90% of cases are sporadic and the remaining 10% are familial. The diagnosis is mostly based on clinical assessment with a history of progression of symptoms and is thus made with a delay of about a year from symptom onset (3), quite likely beyond the therapeutic window of a disease-modifying drug (4). Moreover, the clinical course varies widely. No ALS biomarkers are currently in clinical use except for the genetic forms (10% of ALS cases). At the moment the diagnosis is based on clinical assessments together with a history of progression of symptoms and

thus made with a delay of about a year from symptom onset. ALS biomarkers are needed not only to support early diagnosis, but also to monitor disease progression, and to accelerate (by decreasing the length and the cost of clinical trials) the identification of effective treatments, that are not available at the moment.

- 5 The pathological process in ALS is now recognized as extending beyond motor neurons (5-10), so it can be regarded as a multi-cellular/multi-systemic disease. In particular, peripheral blood mononuclear cells (PBMC) display traits of the disease such as down-regulation of Bcl-2 (11, 12), increased oxidative stress (13), intracellular calcium dysregulation (7) and glutamatergic dysfunction (14).
- 10 A proteomic approach, such as 2D difference gel electrophoresis (2D DIGE), offers the possibility to reveal multiple protein alterations and is a robust system for identification of potential protein biomarkers (15). PBMC are readily accessible clinical samples and offer a series of advantages over serum/plasma and cerebrospinal fluid (CSF). Bio-fluids have wide inter-individual variability and a broad range of protein abundance, which make them
- 15 difficult to analyze by proteomic approaches (16, 17). The cellular proteome is relatively stable, less complex to analyze and gives direct information on alterations of cellular pathways, hence insights into possible pathogenic mechanisms.

We here reported a proteome-based strategy to identify and validate disease biomarkers in PBMC. By using this procedure we found a panel of proteins that are closely associated

20 with ALS and have high potential in clinical applications and translational medicine. Moreover, our results support the use of PBMC of sporadic ALS (sALS) patients for mechanistic studies.

Nardo et al. (13) mentions that PBMCs proteomic analysis may be used to identify ALS biomarkers.

- 25 Massignan et al. (22) refers to PDI and CypA as potential pharmacological targets.

The use of proteomic approaches to identify biomarkers in ALS has been correlated with the use of PBMCs also in US2009/305258 and WO2010148323. However cited markers are different from those of the instant invention. In particular, WO2010148323 differs from the present invention as it is directed to detecting presence, absence, or quantity of an

30 Xenotropic Murine Leukemia Virus-Related Virus (XMRV) antigen, immunopeptide or polynucleic acid in PBMCs.

WO2011005628 discloses methods for diagnosing ALS and determining drug efficacy based on the characterization of TDP-43 in a tissue, specifically in cerebrospinal fluid. However PBMCs are not mentioned.

Though some of the protein biomarkers of the instant invention were previously found as
5 hallmark of diseases affecting the central nervous system, no detection in PBMCs is disclosed. Studies in spinal cord tissues of sALS patients showed that PDI and ERp57 were up-regulated (Atkin JD et al. Endoplasmic reticulum stress and induction of the unfolded protein response in human sporadic amyotrophic lateral sclerosis. *Neurobiol Dis* 2008, 30:400-407), CypA and HSC70 were accumulated in the detergent-insoluble fraction
10 (Basso M et al. Characterization of detergent-insoluble proteins in ALS indicates a causal link between oxidative stress and aggregation in pathogenesis. *PLoS ONE* 2009, 4(12)), HSC70 was present in hyaline inclusions (Watanabe M et al. Histological evidence of protein aggregation in mutant SOD1 transgenic mice and in amyotrophic lateral sclerosis neural tissues. *Neurobiol Dis* 2001, 8:933-941), and TDP-43 was identified as the major
15 component of the ubiquitinated inclusions (Neumann M et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* 2006, 314:130-133).

US2008/0289964 discloses an assay for the diagnosis of patients with ALS or ALS-like disorders that manifest ALS-like symptoms based on the use of 2D gel electrophoresis to
20 separate the complex mixture of proteins found in blood serum.

Apart from disclosure that elevated act^{NT} can be detected in PBMCs of ALS patients, none of the other proteins have been disclosed as potential markers of ALS in PBMCs. Furthermore, the other proteins identified in this invention are not nitrotyrosine modified. Authors found that a combination of PBMC markers provides a reliable diagnosis of ALS.
25 The detection of the combination of markers gives significantly better results than using only act^{NT}.

Description of the invention

Authors identified by a proteome-based approach in peripheral blood mononuclear cells (PBMC) a panel of protein biomarkers that are greatly associated with ALS and can
30 distinguish, with high discriminatory power, between two levels of disease severity (90%), ALS patients from healthy controls (98%) and from patients with neurological disorders that resemble ALS (91%). These PBMC multiprotein biomarkers are easily measurable in large-scale immunoassays and thus suitable to develop a test to be used in clinical practice.

Authors found that at least three biomarkers each one belonging to a different group as listed in Table 1, or allelic variants, or protein precursor, or functional fragments, or gene transcripts thereof, are valuable for set up a diagnostic/prognostic test from peripheral blood mononuclear cells (PBMCs), or extracts thereof, isolated from a biological sample of a subject. The method may also relate to the comparison with respect to a proper control.

Object of the invention is therefore a method to diagnose and/or to make prognostic predictions and/or to monitor the efficacy of a therapy of Amyotrophic lateral sclerosis (ALS) in a subject comprising the steps of:

- a) detecting or measuring the amount of at least three biomarkers each one belonging to a different group of groups I to VII as listed in Table 1, or allelic variants, or protein precursors, or functional fragments, or gene transcripts or post-translationally modified isoforms thereof, from peripheral blood mononuclear cells (PBMCs) or extracts thereof, being isolated from a biological sample of the subject;
- b) eventually comparing such level with respect to a proper control.

In a preferred embodiment the method to diagnose and/or to make prognostic predictions and/or to monitor the efficacy of a therapy of Amyotrophic lateral sclerosis (ALS) in a subject comprises the steps of:

- a) detecting or measuring the amount of at least three biomarkers belonging to the group of: Calreticulin (CALR) (SEQ ID No. 22); TAR DNA-binding protein 43 (TDP-43) (SEQ ID No. 42); Peroxiredoxin-2 (PRDX2) (SEQ ID No. 10); Protein disulfide-isomerase (PDI) (SEQ ID No. 19); protein disulfide-isomerase A3 (ERp57) (SEQ ID No. 20); Proteasome activator complex subunit 1 (PA28a) (SEQ ID No. 23); chloride intracellular channel protein 1 (CLIC1) (SEQ ID No. 15); interleukin-1 receptor-associated kinase 4 (IRAK4) (SEQ ID No. 38); Far upstream element-binding protein 1 (FUBP1) (SEQ ID No. 39); Glutathione S-transferase omega-1 (GSTO1) (SEQ ID No. 12); Heat shock cognate 71 kDa protein (HSC70) (SEQ ID No. 16); peptidyl-prolyl cis-trans isomerase A (CypA) (SEQ ID No. 18); tyrosine nitrated actin (Actin^{NT}) (SEQ ID No. 24); heterogeneous nuclear ribonucleoprotein A2/B1 (hnRNPA2/B1) (SEQ ID No. 40), or allelic variants, or protein precursors, or functional fragments, or post-translationally modified isoforms or gene transcripts thereof, from peripheral blood mononuclear cells (PBMCs) or extracts thereof, isolated from a biological sample of the subject;
- b) eventually comparing such level with respect to a proper control.

In a more preferred embodiment the method is performed by detecting or measuring the

amount of at least: chloride intracellular channel protein 1 (CLIC1) (SEQ ID No. 15); tyrosine nitrated actin (actin^{NT}) (SEQ ID No. 24); heterogeneous nuclear ribonucleoprotein A2/B1 (hnRNPA2/B1) (SEQ ID No. 40), or allelic variants, or protein precursors, or functional fragments, or post-translationally modified isoforms or gene transcripts thereof.

5 More preferably the method should further detect or measure the amount of Interleukin-1 receptor-associated kinase 4 (IRAK4) (SEQ ID No. 38) and Peptidyl-prolyl cis-trans isomerase A (CypA) (SEQ ID No. 18), or allelic variants, or protein precursors, or functional fragments, or post-translationally modified isoforms or gene transcripts thereof. Such embodiment is preferred when a differential diagnosis is to be performed with respect to
10 other neurological diseases.

In an alternative preferred embodiment to make prognostic predictions and/or to monitor the efficacy of a therapy of Amyotrophic lateral sclerosis (ALS) in a subject, step a) is performed by further detecting or measuring the amount of at least: TDP-43 (SEQ ID No. 42); protein disulfide-isomerase A3 (ERp57) (SEQ ID No. 20); peptidyl-prolyl cis-trans
15 isomerase A (CypA) (SEQ ID No. 18), or allelic variants, or protein precursor, or functional fragments, or post-translationally modified isoforms or gene transcripts thereof.

Any techniques known by the expert in the field may be used to detect and measure biomarkers, preferred ones are immunochemical techniques or mass spectrometric techniques.

20 In the context of the invention a subject is an human or a mammalian animal, also to be used for drug tests.

Peripheral blood mononuclear cells (PBMCs) are freshly isolated from a biological sample or properly stored, as usually performed by experts in the field.

It is another object of the invention, a diagnostic kit to diagnose and/or to make prognostic
25 predictions and/or to monitor the efficacy of a therapy of Amyotrophic lateral sclerosis (ALS) in a subject comprising selective ligands for the biomarkers or allelic variants or functional fragments or post-translationally modified isoforms or gene transcripts thereof, as above disclosed. In a preferred embodiment the selective ligands are immunoglobulins or synthetic or recombinant fragments thereof.

30 Authors provide a highly feasible strategy to identify and validate multiprotein biomarkers, potentially applicable to several neurological diseases, both for diagnosis and treatment monitoring. The use of highly accessible peripheral blood cells renders the proteomic profile easily analyzed and able to underline a pathological condition of the central nervous

system with high accuracy.

ALS Markers can be identified in PBMCs by different methods known in the art. The invention relates also to diagnostic/prognostic tests. ALS Markers were identified in PBMCs by a proteomic approach (2D DIGE and MALDI TOF/TOF mass spectrometry) and validated by immunochemical assay (dot blot analysis on single samples) comparing healthy individuals, sporadic ALS patients (in two levels of ALS severity) and non-ALS neurological patients that may resemble ALS (Multiple Sclerosis, diabetic peripheral neuropathy, demyelinating polyneuropathy, alcoholic polyneuropathy, axonal peripheral neuropathy, Charcot-Marie-Tooth disease, spinocerebellar ataxia, myelopathy localized at C6-C7, transverse myelitis, polymyositis and Lambert-Eaton myasthenic syndrome).

The present invention will be described through non-limitative examples, with reference to the following figures:

Figure 1. Proteome-based strategy to identify specific biomarkers of ALS in PBMC.

(A) PBMC of healthy controls and sALS patients with two levels of disease severity, low, with a ALSFRS-R score >24 (ALS >24), or a high one, with a ALSFRS-R score ≤ 24 (ALS <24), were analyzed by 2D DIGE. Samples were pooled for analysis (n=11 for each group). The candidate biomarkers, 71 differentially expressed proteins, were identified by MALDI TOF/TOF mass spectrometry (Table 1). The ALS network (the panel of proteins and their direct interactors) was then examined by network analysis to identify correlations with ALS pathogenesis (Fig. 5). Large-scale validation was done on 14 proteins by dot blot analysis with single samples from an independent set of sALS patients, healthy and non-ALS neurological disorder controls (Table 2-3, Fig. 2,3). We obtained five protein biomarkers that significantly distinguish ALS patients from healthy and non-ALS neurological controls. The 14 PBMC proteins were also measured in G93A SOD1 rats: 5 of them are similarly regulated in the patients and animal model (translational biomarkers).

(B) Representative Sypro-Ruby stained 2D gel of the PBMC proteome. The numbered spots correspond to the differentially expressed proteins listed in Table 1 and Tables 6-7.

(C) Validation by dot blot analysis was done with commercially available antibodies except for the in-house developed anti-actin^{NT} antibody (Fig. 7A-7B). A representative dot blot with the anti-actin^{NT} antibody and quantification of the immunoreactive signals are shown. The box-plot shows relative immunoreactivity in ALS >24 (n=30) (grey box) and ALS <24 (n=30) (black box) patients as percentage of healthy controls (n=30) (white box). The bottom and top of the boxes are the lower and upper quartiles, respectively, the lines

inside the boxes indicate the median and the ends of the whiskers represent the minimum and maximum of all the data. * $p < 0.05$ versus controls as assessed by univariate logistic regression.

Figure 2. Validation of the 14 selected protein biomarkers versus healthy controls:

5 **statistical analysis.** Large-scale validation on single PBMC samples from ALS patients and healthy controls was done by dot blot analysis. Receiver operating characteristics (ROC) curves and analysis of the area under the curve (AUC) were used to find the discriminatory power for proteins showing a significant association with ALS or with disease severity. Results were expressed as odds ratios (OR) and 95% confidence intervals
10 (95% CI). A 95% CI not including the value of 1 indicates a statistically significant result. All probability values were two-sided and $p < 0.05$ was considered significant. (A) Univariate logistic regression: healthy controls (n=30) versus ALS, ALS>24 (n=30) and ALS<24 (n=30). All 14 proteins were significantly associated with ALS (p values in bold type). CLIC1, actinNT and hnRNPA2/B1 were significant in multivariate analysis
15 (highlighted in grey): a ROC curve for the combination of the three proteins and relative AUC is shown. (B) Univariate logistic regression: ALS>24 (n=30) versus ALS<24 (n=30). TDP-43, ERP57 and CypA gave a significant association with disease severity (in bold type). Multivariate analysis indicated that ERP57 was the protein most closely associated with disease severity (highlighted in grey). A ROC curve for ERP57 is shown.

20 **Figure 3. Validation of the 14 selected protein biomarkers versus neurological disorder controls: statistical analysis.** Large-scale validation on single PBMC samples from ALS>24 and neurological disorder (ND) patients was done by dot blot analysis. ROC curves and analysis of AUC were used to find the discriminatory power for proteins showing a significant association with ALS. Results were expressed as OR and 95% CI.
25 All probability values were two-sided and $p < 0.05$ was considered significant. Univariate logistic regression: ALS>24 (n=20) versus neurological disorder controls (n=23). CALR, CLIC1, IRAK4, GSTO1, CypA had significant associations with ALS (in bold type). IRAK4 and CypA were significant in multivariate analysis (highlighted in grey): a ROC curve for the combination of the two proteins and relative AUC is shown.

30 **Figure 4. Analysis of protein biomarkers in the G93A SOD1 transgenic rat model: comparison with sALS patients.** Human PBMC (left column): box-plot shows relative immunoreactivity for the indicated proteins measured by dot blot assay in ALS>24 (n=30) (grey box) and ALS<24 (n=30) (dark grey box) patients as percentage of healthy controls

(control) (n=30) (white box). *, significantly different ($p<0.05$) from control (univariate logistic regression). Rat PBMC (middle column): box-plot shows relative immunoreactivity for the indicated proteins measured by WB (except for actin^{NT}, measured by dot blot) in PBMC lysates (15 μ g) of presymptomatic (pres) (n=6) and symptomatic (symp) G93A SOD1 (n=6) rats as percentage of nonTg (control) (n=8) rats. Protein level values were normalized to actin loading control. *, significantly different ($p<0.05$) from controls (one-way ANOVA followed by Newman-Keuls multiple comparison test). Rat spinal cord (right column): box-plot shows relative immunoreactivity for the indicated proteins measured by dot blot in ventral horn tissue lysates of presymptomatic (pres) (n=6) and (symp) G93A SOD1 (n=6) rats as percentage of nonTg (control) (n=8) rats. Immunoreactivity was normalised to the actual amount of protein loaded, detected after Red Ponceau staining. *, significantly different ($p<0.05$) from control (one-way ANOVA followed by Newman-Keuls multiple comparison test).

Figure 5. Analysis of the PBMC ALS network. Enrichment of the ALS network in GO molecular function (A), GO biological processes (B) and GO cellular component (C) and distribution in the five modules of the ALS network.

Figure 6. Longitudinal analysis of protein biomarkers CypA, ERp57, TDP-43 and SOD1 as negative control. Graphs show relative immunoreactivity for the indicated proteins measured by dot blot assay in PBMC of sALS patients (n=13) over a period of six months. PBMC were collected at t=0 (T1, white column), at t=3 months (T2, grey column) and at t=6 months (T3, black column). Immunoreactivity was normalized to the actual amount of protein loaded, detected after Red Ponceau staining. Data are the means \pm SEM of relative immunoreactivity. *, significantly different ($p<0.05$) from T1 as assessed by repeated measures ANOVA followed by Tukey's multiple comparison test; #, significantly different ($p<0.05$) from T2 as assessed by repeated measures ANOVA followed by Tukey's multiple comparison test.

Figure 7. A: Anti-nitrated actin (actin^{NT}) antibody preparation and characterization by dot blot assay. Purified human actin was nitrated *in vitro* and used as antigen in rabbits for raising polyclonal antibodies. The rabbit polyclonal antiserum was tested by dot blot with actin, actin^{NX}, bovine serum albumin (BSA) and BSA^{NX}, prepared by the same procedure as actin^{NX}. Figure 7 A shows that anti-actin^{NX} does not efficiently recognize another nitrated protein and has more than three times affinity for actin^{NT} of unmodified actin. 3 μ g of actin, actin^{NX}, BSA and BSA^{NX} were loaded in each slot on the nitrocellulose

membrane. The membrane was probed overnight with the polyclonal antiserum diluted 1:7500. Immunoreactivity was normalized to the actual amount of protein loaded on the membrane as detected after Red Ponceau staining. **B: Characterization of the anti-actin^{NT} antibody.** The anti-actin^{NX} antibody was further characterized to determine the specificity of the immunoreactivity against PBMC lysates from controls and ALS patients by Western blotting. Figure 7 B shows that the antiserum recognized a band at 42 kDa, which is the expected Mw for actin^{NT}, but also other bands at higher Mw. These are probably polymerized actin forms. These high-Mw species are essentially only detected in the patients and make this antibody especially useful to distinguish ALS patients from controls, also in a dot blot assay, as shown in Figure 3. Equal amounts of PBMC lysates (30 µg) from healthy controls and ALS patients were analyzed. A representative experiment is shown. The PVDF membrane was probed first with the anti-actin antibody (mouse monoclonal, 1:1000 dilution, Chemicon), and the signal was revealed with a goat anti-mouse Qdot® 800-conjugated secondary antibody (Invitrogen, 1:1000 dilution) on a laser scanner Molecular Imager FX (Bio-Rad), then with the anti-actin^{NT} antibody, and the signal was revealed with an anti-rabbit peroxidase-conjugated secondary antibody/chemiluminescent HRP Substrate (Millipore) on a ChemiDoc XRS system (Bio-Rad).

Figure 8. Analysis of the ALS network. The list of the GO terms enriched in the giant component of the ALS network, with relative p-value and number of proteins associated, is shown. Out of 492 proteins, 245, 239 and 199 were annotated in the three branches of GO, respectively cellular compartment, biological process and molecular function.

Figure 9. Analysis of protein biomarkers in dorsal horn spinal cord of the G93A SOD1 transgenic rat. Bars show relative immunoreactivity for the indicated proteins measured by dot blot in ventral horn tissue lysates of presymptomatic (pres, grey bars) (n=6) and (symp, dark grey bars) G93A SOD1 (n=6) rats and non transgenic (control, white bars) (n=8) rats. Immunoreactivity was normalized to the actual amount of protein loaded, detected after Red Ponceau staining. None of the protein levels were significantly different from control (one-way ANOVA followed by Newman-Keuls multiple comparison test).

Results

2D DIGE proteomic analysis and validation

Figure 1 schematically shows the proteome-based strategy used to identify and validate protein biomarkers of ALS in PBMC. In the first phase, PBMC of healthy controls and sALS patients with two levels of disease severity (Table 4), low, with a functional rating scale (ALSFRS-R) score(18)>24 (ALS>24), and high, with a ALSFRS-R score≤24 (ALS<24), were analyzed by 2D DIGE (Fig. 1A). The analysis, done with 11 pooled samples for each group, detected a total of 129 differential protein spots in the three experimental groups (Table 5). From these spots we used MALDI TOF/TOF mass spectrometry to identify 71 corresponding proteins (Fig. 1B) that were the candidate protein biomarkers (Tables 6-7). We classified these proteins in different categories based on their most recognized function (Table 1). Some of the proteins were identified in multiple spots, at lower Mw or different pi than the ones expected (Table 6). These are fragments of the protein or post-translationally modified isoforms.

We then validated selected proteins by dot blot analysis on single samples of 60 sALS patients, ALS>24 (n=30) and ALS<24 (n=30), and healthy controls (n=30) (Table 8). From the panel of the candidate biomarkers we validated at least 12 proteins, for the following reasons: they are also expressed in the central nervous system (CNS), they are associated with neurodegenerative processes, some of them (PDI, ERp57, HSC70, GSTO1, CypA, PRDX2) specifically with ALS (13, 19-24), and are easily detectable by commercially available specific antibodies. In the same validation analysis we also measured, by using an in-house developed antibody (Fig. 7A-B), nitrated actin (actin^{NT}) (Fig. 1C), whose level was already found high in PBMC of a small cohort of sALS patients and G93A SOD1 transgenic rats (13) as well as in the spinal cord of the G93A SOD1 mouse model of familial ALS (25); and TDP-43, identified as the major component of ubiquitinated inclusions in brains of patients with ALS and frontotemporal lobar degeneration (26) and mutated in 1-3% of ALS cases (27, 28). We confirmed that all 14 proteins can significantly distinguish healthy controls from sALS patients (Fig. 2A, Table 2 and Tables 10-11). Multivariate logistic analysis showed that CLIC1, actin^{NT} and hnRNPA2/B1 were the proteins most associated with ALS in comparison with healthy controls, with 98% discriminatory power (AUC 0.981) (Fig. 2A). Moreover, CypA, TDP-43 and ERp57 levels significantly differed in ALS>24 and ALS<24 patients (Figure 2B and Table 2, indicating that these proteins can discriminate between patients with high and low disease severity). Multivariate logistic analysis showed that ERp57 was the most associated with more

severe ALS, with 89% discriminatory power (AUC 0.893) (Fig. 2B). In all 60 sALS patients the levels of the 14 protein biomarkers were independent of sex and age.

We validated the specificity of the 14 biomarkers with PBMC samples from patients with neurological disorders that in some cases may resemble ALS, e.g. some peripheral neuropathies (Table 9). We analyzed our multiprotein biomarkers by dot blot in ALS > 24 patients (n=20) and in controls with neurological disorders (n=23) (Table 3). Univariate logistic analysis showed that CALR, CLIC1, IRAK4, GSTO1 and CypA were associated with ALS (Fig. 3). Multivariate logistic analysis showed that IRAK4 and CypA were the proteins most associated with ALS rather than neurological disorders, with 91% discriminatory power (AUC 0.905) (Fig. 3). It is interesting to note that an equally high level of TDP-43 was also detected in patients with neurological disorders.

CypA, GSTO1, FUBP1, CLIC1 and actin^{NT} are translational biomarkers of ALS

Translational biomarkers are molecules that can be assessed in both human and animal models and be useful as disease, efficacy and toxicity biomarkers in preclinical and clinical settings accelerating the discovery and development of new treatments. To test whether our 14 PBMC proteins were translational biomarkers, their level was measured in PBMC of a transgenic SOD1^{G93A} rat model of ALS, at pre-symptomatic and symptomatic stages of the disease, in comparison with samples from non-transgenic (nonTg) rats by immunoblot. Five protein biomarkers, CypA, GSTO1, FUBP1, CLIC1 and actin^{NT}, showed similar changes in PBMC of ALS patients and SOD1^{G93A} rats (Fig. 4). It is interesting to note that FUBP1 was one of the least specific ALS biomarkers in PBMC of sporadic patients. It is possible that this protein is more associated with SOD1 mutation-induced alterations and is more informative for the mutant SOD1 fALS forms. The proteins changed significantly already at a presymptomatic stage, suggesting that they might underline early disease alterations also in patients. ERp57, PRDX2, PA28a showed an altered expression in PBMC of SOD1G93A rats compared to nonTg rats, but differently from PBMC of sALS patients. In particular, ERp57 and PA28a were respectively down- and up-regulated at a presymptomatic stage, while PRDX2 was up-regulated only at a symptomatic stage (data not shown). The remaining proteins IRAK4, PDI, HSC70, CALR, TDP-43 and hnRNPA2/B1 did not change their expression with the progression of the disease and if compared with controls.

The five translational biomarkers were also investigated in the rat spinal cord at a presymptomatic and symptomatic stages of the disease (Fig. 4). Lumbar spinal cords were

sectioned in ventral and dorsal horns to investigate whether proteins alterations were associated or not with motor neurons. In ventral horns, all of them changed similarly to PBMC already at a presymptomatic stage except for CLIC1, up-regulated only at a symptomatic stage. In dorsal horns, the proteins did not change with the progression of the disease and if compared with controls (Fig. 9), suggesting that protein alterations observed in ventral horns may be associated more specifically with motor neurons.

In conclusion, CypA, GSTO1, FUBP1, CLIC1 and actin^{NT} are translational biomarkers that might be markers of multi-cellular/multi-systemic alterations underlining pathogenic events both in the human sporadic and animal mutant SOD 1-linked disease forms.

10 *The PBMC ALS network: modularity and functional annotations*

To address the biological relevance of the PBMC protein biomarkers, we further characterized the proteomic alterations by network analysis. To this purpose, a PBMC ALS network was built by retrieving (from protein interaction databases) the known physical interactions among the proteins reported in Table 1 (core proteins) and their direct interactors (first neighbors). Out of the 43 core proteins (protein fragments and post-translationally modified isoforms were excluded from the 71 identified proteins), 41 are reported to interact with other proteins. In total, 499 proteins (41 core proteins and 458 first neighbors) share 1540 interactions within the ALS network. In addition, 492 proteins and 1536 interactions (corresponding to 98.6% and 99.7% of all the proteins and interactions of the network respectively) form a giant component, while the remaining 7 proteins form 3 small islands (of 2 or 3 proteins each). For comparison, we generated random networks, starting from sets of 41 random proteins with the same number of interactions as the ALS proteins. The resulting random networks have similar topological parameters as the ALS network, including degree exponent (1.46 *versus* 1.40, random and ALS respectively), average clustering coefficient *C* (0.171 *versus* 0.185) and average shortest path *d* (3.63 *versus* 3.35), thus indicating that the topology of the ALS network is not unusual. Interestingly, however, the average length of the shortest paths between the proteins used to generate the networks was significantly smaller in the ALS than in the random networks (3.16 and 3.80 respectively; Chi-squared test $p < 0.01$), which suggests that, on average, many proteins within the ALS network have a general propensity to establish physical and possibly functional interactions among themselves.

To identify biological features shared by the ALS proteins, we analyzed the gene ontology (GO) terms, which are significantly enriched in the giant component of the ALS network.

Out of 492 proteins, 245, 239 and 199 were annotated in the three branches of GO, i.e. respectively cellular compartment, molecular function and biological process (29). Figure 8 lists the enriched GO terms. We next examined whether the enriched GO terms were clustered in modules of the network, which are groups of densely inter-connected proteins with shared localizations and/or functions (30). After dividing the ALS network into five modules, we searched which GO annotations were enriched in the modules (Fig. 5A-C). In general, while modules 2, 3 and 5 contain proteins that are mostly localized in membranes, modules 1 and 4 are more heterogeneous and with interesting features that did not immediately emerge in the initial analysis. For instance, module 1 is enriched in proteins that are mainly localized to either the cytosol or the nucleus, and that are involved in the molecular function of nucleotide (RNA and DNA) binding. In module 4 proteins are almost equally localized to cytosol and nucleus and transcription regulation is one of the most enriched biological processes.

CypA, TDP-43 and ERp57 as biomarkers for disease progression

We next examined in a pilot longitudinal study whether CypA, TDP-43 and ERp57 changed over time. We collected three longitudinal PBMC samples from an independent set of 13 sALS patients over a period of six months and used a statistical model for repeated measures to assess the effect of time on protein levels, detected by dot blot assay. The patients had on average a disease duration of 21 ± 10 months, a 33 ± 5 ALSFRS-R score at the first draw and a 28 ± 7 ALSFRS-R score at the third draw (Table 12). The level of all the three protein biomarkers showed a significant increase within six months from the first PBMC collection and for TDP-43 already within 3 months (Fig. 6). As a negative control, in the same set of samples we measured SOD1, which did not significantly change over time (Fig. 6).

Discussion

A number of gene expression studies have demonstrated the utility of PBMC as a source of biomarkers in neurological disorders (31-33). This is the first study that describes a highly feasible strategy to identify and validate multiprotein biomarkers in PBMC, potentially applicable to several neurological/neurodegenerative diseases. Protein profiles are directly connected to changes in molecular pathways related to health and disease, therefore have the potential to accurately monitor the progression of the disease or the response to a treatment.

We have previously shown that PBMC undergo immunophenotypic changes in sALS patients (12). We have now identified a panel of protein biomarkers in these cells that are associated with sALS with high discriminatory power. These PBMC proteins are easily measurable in large-scale immunoassays aimed at developing diagnostic/prognostic tests for clinical use. The great advantages of such an *in vitro* test is low invasiveness for the patient compared to CSF tests (34, 35), the consequent greater availability of samples for large clinical studies, including longitudinal ones, and the simple laboratory procedures involved.

The ideal diagnostic marker should detect disease before clinical diagnosis, which is highly challenging for a rare and sporadic disease. ALS patients very often see the specialized neurologist only after months from the first symptoms, when they are unquestionably ill. It is therefore very difficult to test and validate the applicability of biomarkers in preclinical diagnosis. However, our PBMC protein biomarkers seems to be promising to support prompt clinical diagnosis, since all the 14 validated proteins can distinguish with high significance ALS patients with ALSFRS-R>24 from healthy controls (Table 9). Moreover, in a G93A rat model of ALS, alterations of the level of CypA, GSTO1, FUBP1, CLIC1 and actinNT are detected before disease onset. It is possible that some of these biomarkers may predict the onset of SOD 1-linked familial ALS.

There are no precise measures of ALS disease progression that allow for short-term monitoring of the disease and assessment of treatment efficacy. In clinical trials survival time is therefore used as the primary measure of outcome. This requires large number of patients followed over a long period of time making ALS clinical trials very expensive. A panel of biomarkers that can reliably assess disease progression would enable a substantial reduction of the costs of the clinical trial and accelerate therapy development in ALS. ERp57, CypA and TDP-43, that were able to discriminate between patients with high and low disease severity, were selected for a pilot longitudinal study and proved to be good candidates for such applications. Large longitudinal studies are now needed to further validate the use of these proteins in clinical practice.

Up to 10% of patients initially diagnosed as having ALS are false positive (36, 37). A similar percentage is false negative and undergoes inappropriate medical or surgical procedures (38, 39). Thus, there is a need for biomarkers that distinguish ALS with high accuracy from neurological disorders that in some cases may resemble it, e.g. some peripheral neuropathies that are treatable and do not have a fatal prognosis. We found that

there are PBMC protein biomarkers that significantly distinguish ALS from neurological disorders, such as chaperones (CALR, CypA), proteins involved in redox homeostasis (GSTO1, CLIC1) and immune responses (IRAK4). IRAK4, which has no previous association with ALS, has a central function in innate immunity (40). Its up-regulation may
5 be connected with the high cytokine levels observed in sALS patients (34).

Finally, multivariate analysis helped us in defining the most convenient combination of protein biomarkers that could be potentially used in clinics (i) to support diagnosis (CLIO, actin^{NT} and hnRNPA2/B1), (ii) to determine disease severity (ERp57), and (iii) to contribute to differential diagnosis of ALS from other neurological conditions (CypA and
10 IRAK4).

The pathogenesis of sALS is largely unknown. The concept of PBMC as a window into the CNS has been already proposed for several neurological disease states (31-33, 41). CNS and immune cells communicate through multiple mechanisms and have several similarities in receptor expression and transduction processes (42). We therefore hypothesized that
15 PBMC protein profiles could help to elucidate pathways underlying ALS etiology. Indeed, some of the protein biomarkers identified in PBMC of sALS patients were previously found as hallmarks of disease in CNS. Studies in spinal cord tissues of sALS patients showed that PDI and ERp57 were up-regulated (20), CypA and HSC70 were accumulated in the detergent-insoluble fraction (43), HSC70 was present in hyaline inclusions (24), and
20 TDP-43 was identified as the major component of the ubiquitinated inclusions (26).

The protein profile changes detected in PBMC of ALS patients are suggestive of possible pathogenic mechanisms. For instance, up-regulation of endoplasmic reticulum (ER) chaperones (PDI, ERp57, CALR) is a typical cellular response to ER stress that triggers the unfolded protein response leading eventually to cell death (44). The increased level of a nitrotyrosine-linked protein, actin^{NT}, is indicative of nitrative stress. Alterations in CypA, GSTO1, PRDX2 suggest disturbances in cellular redox regulation (45-48). It is important to note that all these pathogenic alterations were previously reported in the spinal cord of
25 sALS patients and mutant SOD1 animal models (19, 20, 25, 49-52). HnRNPA2/B1 belongs to the family of heterogeneous nuclear ribonucleoproteins that participates to several RNA-related biological processes such as transcription, pre-mRNA processing, mRNA transport to the cytoplasm and translation (53). It is also a binding partner of TDP-43 and seems to be crucial for at least one of its putative functions (54). The up-regulation of hnRNPA2/B and TDP-43 in PBMC of both ALS^{>24} and ALS^{<24} patients may
30

underline aberrant RNA processing events that are now emerging as central in neurodegeneration in ALS and other neurological disorders (55). All these data endorse the use of PBMC for further *in vitro* mechanistic studies and reinforce the theory that ALS is a multi-cellular disease. Experimental models for the sporadic form are not available and all mechanistic studies are done with transgenic cells and animals expressing one of the mutant genes linked to familial ALS. Studies with PBMC would have the advantage to consider the influence of the genetic background of the patient.

Translational biomarkers, that link responses between human and animal model, are of particular interest because their role in the pathogenesis can be investigated in detail in the animal model where they can also offer important preliminary information for clinical trials. We found that CypA, GSTO1, FUBP1, CLIC1 and actin^{NT} are translational biomarkers and, as well as PDI and ERp57, are altered in the spinal cord of SOD1 G93A rodents before disease onset (20, 22, 25), suggesting a possible involvement in pathways that trigger the disease. Further mechanistic studies in the animal models with these proteins are now warranted.

As further evidence that PBMC are a reliable source of biologically relevant biomarkers for ALS, we report that the PBMC ALS network is structurally and functionally connected. First, the probability of finding ALS proteins linked either directly or indirectly (via a common neighbor) was higher in the ALS network than in random networks. Second, the ALS network was significantly enriched in GO annotations that were also enriched in genomic/proteomic data obtained from spinal cord and motor neurons of familial ALS animal models (56-58). In addition, GO terms related to DNA/RNA binding, transcription regulation and nucleus were enriched in two highly-connected groups of nodes (modules 1 and 4) within the ALS network. Genes with similar GO annotations (TARDBP and FUS) are mutated in familial and sporadic ALS (27, 28, 59, 60) and clustered within modules of a recent gene expression network derived from peripheral blood of sALS patients (61).

In conclusion, we identified and validated a panel of highly promising protein biomarkers of ALS in PBMC that may be useful in clinical studies, helping elucidate pathogenic mechanisms and pin-pointing pathways to tackle for future therapeutic interventions. Moreover, we provided a straightforward strategy to identify and validate multiprotein biomarkers that can be easily applied to several neurological diseases.

Materials and Methods

Patients and controls

Patients with sALS, according to revised El Escorial criteria(62), were examined and blood samples were collected at the IRCCS Fondazione S. Maugeri, in Milano and Pavia, Italy. The patients were then divided into two groups according to the ALSFRS-R score, ALS>24 or ALS<24 patients. None of the sALS patients had systemic inflammatory conditions as detected by the erythrocyte sedimentation rate and total blood cell count. Blood samples of non-ALS neurological disorder controls were provided by the NEMO, Niguarda Ca' Granda Hospital, Milano, Italy and the IRCCS Fondazione S. Maugeri, in Milano and Pavia, Italy. None of the patients was receiving drugs that might interfere with total blood cell count. Tables 4 and 7-8 summarize the characteristics of ALS patients and non-ALS neurological disorder controls. Blood samples of healthy donors were provided by the Transfusion Medical Centre at the IRCCS Policlinico S. Matteo, Pavia, Italy. Informed consent was obtained from all subjects, according to the ethics committee guidelines at the different centers.

PBMCs

Samples of peripheral venous blood from patients and controls were collected in EDTA pre-coated vials (Vacuette K3E K3EDTA, Greiner bio-one). PBMC were isolated from EDTA-blood by Ficoll-Hypaque (Ficoll-Plaques™ Plus, GE Healthcare) density gradient centrifugation. Mononuclear cells were harvested from the interface and washed three times with RPMI 1640 medium (EuroClone). Platelets were eliminated by an additional wash and centrifugation at 200 x g for 10 min. Patients' and controls' PBMC were stored as pellets at -80°C. Just before analysis PBMC proteins were obtained by cell lysis in 20 mM Tris-HCl pH 7.5, 0.1% NP40 and 0.1% SDS supplemented with Protease Inhibitors (Sigma). Proteins were quantified by the BCA protein assay (Pierce).

Two-dimensional difference in gel electrophoresis (2D DIGE)

PBMC proteins were prepared for 2D DIGE analysis. Three pools of 25 µg from 11 healthy controls, 11 ALS>24 and 11 ALS<24 patients were methanol-precipitated. Proteins were then dissolved in 30 mM Tris-HCl pH 8.5, 7 M urea, 2 M thiourea, CHAPS 4% (w/v) and Cydye-labeled according to the manufacturer's instructions (GE Healthcare) with minor modifications. Briefly, 25 µg of each pool was labeled with 200 pmol of Cy3 or Cy5 dye for 30 min in ice in the dark. To exclude preferential labeling of the dyes, each sample was also reverse labeled. As an internal standard, aliquots of each pool were mixed and labeled with Cy2 dye. Labeled samples were then resuspended in Destreak Solution™ (GE Healthcare) with IPG buffer pH 3-10 NL 0.5% v/v (GE Healthcare) added and loaded into

7 cm-IPG strips pi range 3-10NL (GE Healthcare). Isoelectrofocusing was done on an IPGphor apparatus (GE Healthcare) with the following protocol: 30 V for 270 Vhrs, 200V for 200 Vhrs, 2000 V for 2000 Vhrs, a linear gradient of 3500 V for 1375 Vhrs, 3500 V for 7000 Vhrs, a linear gradient of 8000 V for 8625 Vhr, 8000 V for 32000 Vhr and forever at 30 V. SDS-PAGE was done using precast 10% polyacrylamide SDS gel (Invitrogen). Four 2D gels were run with the three experimental groups: healthy controls and ALS patients, ALS>24, and ALS<24. Each gel contained two experimental groups, one Cy3-labelled, the other Cy5-labelled plus the Cy2-labelled internal standard. Gel images were captured by the laser scanner Molecular Imager FX (Bio-Rad). Image analysis was done with Progenesis PG240 v2006 software (Nonlinear Dynamics). For each spot the normalized volume was standardized against the intra-gel standard, dividing the value for each spot normal volume by the corresponding internal standard spot normal volume within each gel. The values for each spot in each group were expressed as the mean of the Cy3- and Cy5-labelled analyses. The values for ALS patients were reported as fold change: higher (positive) or lower (negative) spot volume of the samples from ALS patients in comparison with healthy controls.

Protein identification

Protein spots were located and excised from 2D gels with the EXQuest™ spot cutter (Bio-Rad). Spots were processed and gel-digested with modified trypsin from bovine pancreas (Roche) and identified by mass spectrometry (MS), essentially as previously described(25). Digestion, desalting and concentration with ZipTip® pipette tips with C18 resin (Millipore) and MALDI target deposition were carried out on an automated protein digester DigestPro MS (Intavis AG). Peptide mass fingerprinting and tandem mass spectrometry (MS/MS) were done on a 4800 MALDI TOF/TOF mass spectrometer (Applied Biosystems). The mass spectra were internally calibrated with trypsin autolysis fragments. The five most abundant precursor ions, out of the exclusion mass list (ions from human keratin and trypsin), were selected for MS/MS analysis. The combined MS and MS/MS data were submitted by GPS Explorer v.3.6 software (Applied Biosystems) to the MASCOT database search engine (Version 2.1, Matrix Science) and searched with the following parameters: Uniprot_Swissprot 57.8 database over all *Homo sapiens* protein sequences deposited, no fixed modifications, as possible modifications carboamidomethylation of cysteine and oxidation of methionine, 1 missed trypsin cleavage, a mass tolerance of ± 0.1 Da for the peptide masses and ± 0.3 Da for the MS/MS

fragment ion masses. A protein was regarded as identified if the MASCOT protein score, based on the combined MS and MS/MS data, was above the 5% significance threshold for the database(63). Some of the proteins were identified by liquid chromatography- (LC)-MS/MS using the microfluid chip-based technology for nanoelectrospray coupled to an ion trap mass spectrometer (Agilent 1200 LC/MSD Trap XCT) as previously described(64).
5 Data were acquired sequentially in MS mode (scan range 200-2000 amu) and in data-dependent mode, automatically recording the MS/MS spectra of the four most abundant ions in every scan cycle. Data files of all MS/MS spectra in a LC run were merged, and submitted as an "mgf" file (BioWorks Rev 3.1 SRI, Thermo Scientific) to the MASCOT
10 database search engine in MS/MS Ion Search mode. A protein was regarded as identified if MASCOT individual ion scores were >36, which indicate identity or extensive homology (p<0.05), for at least three matched peptides. Search parameters were the same as above.

Anti-actin^{NT} antibody preparation

Purified human non-muscle actin (>99% pure; Cytoskeleton, Inc.) (1.5 mg/mL) was
15 incubated for 24 hours at room temperature in a nitration solution containing 20 mM sodium acetate pH 5.6, 9 μ M FeCl₃, 10 mM NaNO₂, 0.3% H₂O₂. Nitration of the protein was verified by Western blotting (WB) with the anti-nitrotyrosine antibody (clone HM.1 1; Hycult Biotechnology). Actin^{NT} was used as antigen in rabbits for raising polyclonal antibodies at Eurogentec S.A. LIEGE Science Park (Belgium). The anti-actin^{NT} antibody
20 was characterized as described in Fig. 7A-B.

Dot blot analysis

Aliquots (3 μ g) of PBMC samples were loaded on nitrocellulose membrane, Trans-Blot Transfer Medium (Bio-Rad), by vacuum deposition on the Bio-Dot SF blotting apparatus (Bio-Rad). Membranes were probed overnight with primary antibodies: rabbit polyclonal
25 anti-CALR (the immunogen is the recombinant human protein) (1:5000), rabbit polyclonal anti-PDI (the immunogen is purified bovine liver protein) (1:4000) and mouse monoclonal anti-ERp57 (the immunogen is the recombinant full-length human protein) (1:500) from StressGen, mouse monoclonal anti-HSC70 (specific for an epitope mapping between amino acids 580-601 at the C-terminus of the human protein) (1:1000), goat polyclonal
30 anti-PRDX2 (epitope mapping near the N-terminus of human protein) (1:2000), goat polyclonal anti-PA28a (epitope mapping within an internal region of the mouse protein) (1:2000), mouse monoclonal anti-CLIC1 (the immunogen is the recombinant full-length human protein) (1:2500), goat polyclonal anti-IRAK4 (the immunogen is a peptide

mapping at the C-terminus of the human protein) (1:1000) and rabbit polyclonal anti-FUBP1 (epitope corresponding to amino acids 65-106 mapping near the N-terminus of FBP1 of human origin) (1:1000) from Santa Cruz Biotechnology, mouse monoclonal anti-GSTO1 (the immunogen is partial recombinant human protein with GST tag) (1:500) and
5 mouse monoclonal anti-hnRNPA2B1 (the immunogen is the recombinant full-length human protein with GST tag) (1:2500) from Abnova, rabbit polyclonal anti-CypA (the immunogen is the recombinant full-length human protein) (1:2500) from Upstate Biotechnology, rabbit polyclonal anti-actin^{NT} (1:7500) developed in-house, and rabbit polyclonal anti-TDP-43 (the immunogen is the recombinant full-length human protein with
10 GST tag) (1:2000), kindly provided by Francisco Baralle, ICGEB, Trieste, Italy, and then with secondary anti-mouse, anti-rabbit or anti-goat peroxidase-conjugated secondary antibodies (Santa Cruz Biotechnology). Blots were developed by Immobilon Western Chemiluminescent HRP Substrate (Millipore) on the ChemiDoc XRS system (Bio-Rad). Densitometry was done with Progenesis PG240 v2006 software (Nonlinear Dynamics).
15 Immunoreactivity was normalized to the actual amount of proteins loaded on the membrane as detected after Red Ponceau staining (Fluka).

Statistical analysis

Continuous variables such as age and all data of the considered proteins were described using 'standard' statistics (mean, standard deviation), and relative and absolute frequencies
20 were used for categorical variables. Spearman correlation analysis was done to assess the associations between the 14 proteins. Univariate and multivariate logistic regression models were also built to identify the level of association between proteins and ALS patients. Receiver operating characteristics (ROC) curves and analysis of the area under the curve (AUC) were used to find the discriminatory power for proteins showing a
25 significant association with ALS or disease severity. Results were expressed as odds ratios (OR) and 95% confidence intervals (95% CI). A 95% CI not including the value of 1 indicates a significant result. Dependence on sex and age (at the time of PBMC collection) was analyzed by two-way ANOVA. All probability values were two-sided and $p < 0.05$ was considered statistically significant. Version 9.1 of the SAS statistical software (SAS
30 Institute, Inc, Cary, NC, USA) and version 5.03 of the GraphPad Prism software (GraphPad Software, Inc, La Jolla, CA, USA) were used.

Analysis of protein biomarkers in the rat model

NonTg and Tg rats expressing a high copy number of mutant human SOD1 with Gly-93-Ala substitution were bred and maintained at the Mario Negri Institute for Pharmacological Research, Milan, Italy. Animals were housed at $21\pm 1^{\circ}\text{C}$ with relative humidity $55\pm 10\%$ and 12 h of light. Food (standard pellets) and water were supplied *ad libitum*. Transgenic
5 rats were identified with PCR on DNA from tail biopsies. G93A SOD1 rats were killed at presymptomatic (15 weeks of age) and early-symptomatic (20 weeks of age) stages of disease. NonTg rats were used as controls (20 weeks of age). Procedures involving animals and their care were conducted in conformity with the institutional guidelines that are in compliance with national (D.L. No. 116, Suppl. 40, Feb. 18, 1992 Circolare No. 8, G.U.,
10 14 luglio 1994) and International laws and policies (EEC Council Directive 86/609. OJ L 358,1, Dec. 12, 1987; *NIH Guide for Care and use of Laboratory Animals*, U.S. National Research Council, 1996).

PBMC

Blood was sampled directly by intracardiac puncture from rats and collected in EDTA pre-coated vials (Vacuette K3E K3EDTA, Greiner bio-one). PBMC were isolated from blood,
15 washed and stored as pellets at -80°C in the same way as the human samples. PBMC proteins were obtained by cell lysis in 50 mM Tris-HCl, pH 7.5, 2% (w/v) CHAPS, 37.5 U benzonase (Merck), supplemented with Protease Inhibitors (Sigma). Proteins were quantified by the BCA protein assay (Pierce). Proteins were separated by SDS-PAGE on
20 precast 12% Criterion XT Bis-Tris gels (Bio-Rad) and immunoblotted, as described(43). The blots were probed with primary and secondary antibodies, as described for dot blot assay of the human samples. Blots were also probed with an antibody that recognizes actin (mouse monoclonal, 1:1000 dilution, Chemicon) for loading control and developed by Immobilon Western Chemiluminescent HRP Substrate (Millipore) on the ChemiDoc XRS
25 system (Bio-Rad). Densitometry was done with Progenesis PG240 v2006 software (Nonlinear Dynamics).

Spinal cord

Spinal cords were removed and the lumbar tract was cut into ventral and dorsal horn sections on a cryostat. Ventral and dorsal horn sections were suspended in 1% SDS (5
30 $\mu\text{L}/\text{mg}$), sonicated and boiled for 10 min. Homogenates were centrifuged at 12000 RPM and supernatants analyzed by dot blot, as described for the human samples except for the FUBP1 and PA28a analyses where the mouse monoclonal anti-FUBP1 antibody (1:500)

from Santa Cruz Biotechnology and the rabbit polyclonal anti-PA28a antibody (1:1000) from Cell Signaling were used.

Network analysis

Generation and modularity of the ALS network

5 To assemble the interaction network of the ALS proteins, the interactions involving the proteins reported in Table 1 ("core proteins") with their direct interactors ("first neighbors") were retrieved from the Human Protein Reference Database (HPRD) (65, 66). Out of the 43 core proteins, 41 are reported to interact with other proteins. Then, proteins and interactions were visualized as nodes and edges of the network using Cytoscape 2.6.1
10 (67) and yEd Graph and analyzed with the Cytoscape plug-in Network Analyzer (68). For comparison, we generated random networks, starting from sets of 41 random proteins with the same number of interactions as the ALS proteins. Like the ALS proteins, the random ones were obtained in proteomic studies (Supporting References) but, unlike the ALS proteins, they were selected randomly, without any bias for a particular experimental
15 setting. The Rives and Galitski algorithm was applied to sub-divide the ALS network into modules (69). Briefly, an adjacency matrix of all-pairs shortest path distances was subjected to average linkage hierarchical clustering (70) with uncentered correlation as similarity metric, using Cluster 3.0 (71). The resulting dendrogram, generated using Java Tree View 1.0.13(72), was parsed into modules, by cutting the tree after the 4th bifurcation.

20 *GO analysis*

BinGO software(73) was used to assess the enrichment of the test set (the ALS network or modules thereof) in GO terms (along each of the three GO branches), compared to a reference set of similar size, randomly chosen from all the human genes annotated along the same GO branch. The statistical significance of the enriched GO annotations was first
25 assessed with the hyper-geometric test, then the P-values were corrected with the Benjamini-Hochberg correction of false discovery rate (74). After adjustment for the significant terms ($P < 0.05$), terms were selected with an intermediate level of specificity along the GO hierarchy. Fine-grained terms, i.e. those not included in the GO Slim (generic GOA) dataset (75), were first removed from the BinGO output lists. Then coarse-
30 grained terms were also removed, i.e. those that are not terminal in the Slim-restricted dataset (specifically, the terms that are preceded by "is a" and/or "is part of" text strings).

Table 1| Protein biomarkers.

Group	Spot	Protein name	Uniprot	SEQ ID No.
Energy metabolism				
I	1	ATP synthase subunit beta	P06576	1
	2	Triosephosphate isomerase	P60174	2
	3-6	Phosphoglycerate kinase 1	P00558	3
	7-10	Alpha-enolase	P06733	4
	11-13	Fructose-bisphosphate aldolase A	P04075	5
	14	Glyceraldehyde-3 -phosphate dehydrogenase	P04406	6
	15	L- Lactate dehydrogenase A chain	P00338	7
	16	L-Lactate dehydrogenase B chain	P07195	8
Redox regulation				
II	17	Flavin reductase	P30043	9
	18	Peroxiredoxin-2 (PRDX2)	P32119	10
	19	Peroxiredoxin-6	P30041	11
	20	Glutathione S-transferase omega- 1 (GSTO1)	P78417	12
	21	Superoxide dismutase [Mn]	P04179	13
	22	Protein DJ- 1	Q99497	14
	23	Chloride intracellular channel protein 1 (CLIO)	000299	15
Protein folding and degradation				
III	24-25	Heat shock cognate 71 kDa protein (HSC70)	PI1142	16
	26	78 kDa glucose-regulated protein	P11021	17
	27	Peptidyl-prolyl cis-trans isomerase A (CypA)	P62937	18
	28	Protein disulfide-isomerase (PDI)	P07237	19
	29-32	Protein disulfide-isomerase A3 (ERp57)	P30101	20
	33-34	Endoplasmic reticulum protein ERp29	P30040	21
	35	Calreticulin (CALR)	P27797	22
	36	Proteasome activator complex subunit 1 (PA28a)	Q06323	23
Cytoskeleton-associated				
IV	37-43	Actin	P60709	24
	44-48	Vinculin	P18206	25
	49-51	Moesin	P26038	26
	52	Tropomyosin alpha-4 chain	P67936	27
	53	Alpha-actinin- 1	P12814	28
	54-55	Macrophage-capping protein	P40121	29
	56	F-actin capping protein subunit-alpha 1	P52907	30
	57-58	Talin- 1	Q9Y490	31
	59	Transgelin-2	P37802	32
	60	Filamin-A	P21333	33
	61	Golgin subfamily B member 1	Q14789	34
	#	Nitrated actin (actin ^{NT})	P60709	24
Inflammatory response				
V	62	Group XIIA secretory phospholipase A2	Q9BZM1	35
	63	Annexin A2	P07355	36
	64	Leukocyte elastase inhibitor	P30740	37
	65	Interleukin-1 receptor-associated kinase 4 (IRAK4)	Q9NWZ3	38
DNA/RNA binding				
VI	66-67	Far upstream element-binding protein 1 (FUBP1)	Q96AE4	39
	68	Heterogeneous nuclear ribonucleoproteins A2/B1 (hnRNPA2/B1)	P22627	40
	69	Probable ATP-dependent RNA helicase DDX41	Q9UJV9	41
	##	TAR DNA-binding protein 43 (TDP-43)	Q13148	42
Others				
VII	70	AH receptor-interacting protein	000170	43
	71	Spindle and kinetochore-associated protein 1	Q96BD8	44

#,##: All the proteins were identified in the 2D DIGE analysis except for actin^{NT} and TDP-43, identified as hallmark of the disease in previous studies (13, 26-28).

5 **Table 2| Validation of 14 selected protein biomarkers in patients and healthy controls.**

Protein	controls	ALS>24	ALS≤24	ALS>24 vs controls*	ALS≤24 vs controls*	ALS≤24 vs ALS>24*
CALR	2.8±0.7	4.9±2.8	4.4±2.3	1.8	1.6	-1.1
<i>TDP-43</i>	<i>4.1±1.6</i>	<i>5.2±2.1</i>	<i>6.8±2.6</i>	<i>1.3</i>	<i>1.7</i>	<i>1.3</i>
PRDX2	0.9±0.5	1.4±0.7	2.0±1.1	1.6	2.2	1.4
PDI	1.0±0.3	1.6±0.7	1.8±0.7	1.6	1.8	1.1
<i>ERp57</i>	<i>2.5±0.7</i>	<i>3.8±1.4</i>	<i>4.7±2.4</i>	<i>1.5</i>	<i>1.9</i>	<i>1.2</i>
PA28a	4.1±4.5	1.4±1.1	1.6±1.6	-2.9	-2.6	1.1
CLIC1	0.4±0.3	0.8±0.4	0.9±0.5	2.0	2.3	1.1
IRAK4	1.0±0.5	2.2±1.5	2.4±1.5	2.2	2.4	1.1
FUBP1	1.7±1.5	0.6±0.6	0.9±1.0	-2.8	-1.9	1.5
GSTO1	1.2±0.7	0.7±0.5	0.6±0.4	-1.7	-2.0	-1.2
HSC70	0.8±0.5	1.7±1.2	2.3±2.3	2.1	2.9	1.4
<i>CypA</i>	<i>1.4±0.7</i>	<i>3.4±2.7</i>	<i>2.0±1.2</i>	<i>2.4</i>	<i>1.4</i>	<i>-1.7</i>
Actin ^{NT}	0.7±0.3	1.5±0.8	1.6±0.6	2.1	2.3	1.1
hnRNPA2/B1	2.9±1.8	5.7±2.7	6.4±3.0	2.0	2.2	1.1

10 Large-scale validation was done by dot blot assay with the specific antibodies on single PBMC samples from ALS patients, ALS>24 (n=30) and ALS≤24 (n=30), and healthy controls (n=30); mean ± SD of the relative immunoreactivity of the specific antibody normalized to the actual amount of proteins loaded on the membrane as detected after Red Ponceau staining; *, -fold change as the mean increase (positive) and decrease (negative) in protein concentration in samples from ALS patients in comparison with healthy controls or between the two levels of disease severity. All 14 proteins were significantly associated with ALS. CLIC1, actin^{NT} and hnRNPA2/B1 (in bold), significant also in multivariate analysis, are the most associated with ALS. The analysis of these three proteins is suggested to support diagnosis. TDP-43, ERp57 and CypA (in bold and italics) are significantly associated with disease severity and their analysis is suggested to assess disease progression.

15 **Table 3| Validation of the 14 biomarkers in patients and ND controls.**

Protein	ALS>24	ND	ND vs ALS>24*
CALR	3.2±1.3	4.9±2.8	1.5
TDP-43	2.0±0.5	2.2±0.6	1.1
PRDX2	0.6±0.2	0.5±0.2	-1.2
PDI	0.8±0.4	0.9±0.4	1.1
ERp57	0.1±0.1	0.1±0.1	1.0
PA28a	0.3±0.2	0.3±0.2	1.0
CLIC1	1.5±0.8	2.9±2.1	1.9
<i>IRAK4</i>	<i>1.0±0.4</i>	<i>0.7±0.2</i>	<i>-1.4</i>
FUBP1	2.8±2.0	2.0±1.3	-1.4
GSTO1	0.5±0.3	0.8±0.6	1.6

HSC70	2.3±2.2	3.4±2.6	1.5
CypA	2.6±1.0	4.5±2.2	1.7
Actin ^{NT}	3.0±1.2	2.8±1.3	-1.1
hnRNPA2/B1	1.7±1.0	1.9±0.9	1.1

Large-scale validation was done by dot blot assay with the specific antibodies on single PBMC samples from ALS patients, ALS>24 (n=20), and ND controls (n=23); mean ± SD of the relative immunoreactivity of the specific antibody normalized to the actual amount of proteins loaded on the membrane as detected after Red Ponceau staining; *, -fold change as the mean increase (positive) and decrease (negative) in protein concentration in samples from ALS patients in comparison with ND controls. CALR, CLIC1, IRAK4, GSTO1, CypA (in bold) had significant associations with ALS. The analysis of these proteins are suggested to support differential diagnosis of ALS from other neurological conditions that may resemble it. IRAK4 and CypA (in bold and italics), significant also in multivariate analysis, is the minimal combination.

Table 4 | Main characteristics of healthy individuals and ALS patients used in the proteomic 2D DIGE analysis.

Sample	Clinical diagnosis	Age ¹	Sex	Score ²	Onset ³	Duration ⁴	Survival ⁵
1-11	Healthy	54±5	6(M), 5(F)	-	-	-	-
12-22	ALS	65±13	5(M), 6(F)	>24			
12	ALS	56	F	28/48	spinal	60	>98*
13	ALS	76	M	25/48	spinal	96	111
14	ALS	80	F	37/48	n.a	60	83
15	ALS	61	M	25/48	bulbar	36	>72*
16	ALS	39	F	39/48	spinal	n.a.	n.a.
17	ALS	66	F	27/48	bulbar	32	>56*
18	ALS	79	M	28/48	spinal	24	46
19	ALS	81	F	40/48	bulbar	24	51
20	ALS	59	M	33/48	spinal	9	30
21	ALS	52	M	45/48	bulbar	6	29
22	ALS	61	F	40/48	spinal	7	>43*
23-33	ALS	63±9	3(M), 8(F)	≤24			
23	ALS	73	F	15/48	spinal	17	24
24	ALS	52	F	24/48	spinal	60	75
25	ALS	64	M	24/48	spinal	43	70
26	ALS	63	F	22/48	n.a.	17	27
27	ALS	52	F	17/48	bulbar	8	11
28	ALS	72	F	12/48	spinal	27	29
29	ALS	61	M	12/48	spinal	28	30
30	ALS	72	F	21/48	spinal	n.a.	n.a.
31	ALS	47	M	17/48	spinal	n.a.	n.a.
32	ALS	63	F	23/48	spinal	16	31
33	ALS	72	F	24/48	spinal	24	26

¹Age at PBMC collection; ²ALSFRR-R score; ³Site of onset; ⁴Disease duration (months) from the onset of symptoms to PBMC collection; ⁵Disease duration (months) from the onset of symptoms to death; -, not applicable; *, patient still alive (feb-2010); n.a, not available.

Table 5 | Spot volume changes in 2D DIGE analysis.

Spot volumes¹	ALS>24 vs ctr	ALS<24 vs ctr	ALS>24 vs ALS<24
increased	83	55	112
decreased	36	68	17
unchanged	107	103	97

¹Normalized spot volumes. In the 2D DIGE analysis only the 226 matched spots in the three experimental groups were quantified. The spot volume was considered increased or decreased if the ratio of the volume in the patients to controls or in ALS>24 patients to ALS<24 patients was >1.4; the spots with a ratio <1.4 were considered unchanged.

5

Table 6| Candidate protein biomarkers identified by mass spectrometry

Spot	Uniprot KB ¹	Protein name	Mw _{cal}	pI _{cal}	Mw _{obs}	pI _{obs}	Sequence coverage	Matched peptides	Mascot score
1	P06576	ATP synthase subunit beta (ATPase)	56.5	5.2	60.0	4.9	46%	21	156
2	P60174	Triosephosphate isomerase	26.6	6.4	28.8	7.3	79%	16	325
3	P00558	Phosphoglycerate kinase 1	44.5	8.3	54.4	8.5	72%	36	506
4	P00558	Phosphoglycerate kinase 1*	44.5	8.3	35.3	6.2	34%	8	333
5	P00558	Phosphoglycerate kinase 1*	44.5	8.3	35.1	5.8	16%	4	147
6	P18669	Phosphoglycerate mutase 1*	28.7	6.6	26.8	6.9	50%	12	220
7	P06733	Alpha-enolase	47.1	7.0	55.4	6.3	31%	12	173
8	P06733	Alpha-enolase	47.1	7.0	55.1	6.8	34%	12	241
9	P06733	Alpha-enolase	47.1	7.0	55.1	7.0	43%	20	113
10	P06733	Alpha-enolase	47.1	6.7	55.4	6.5	38%	19	102
11-13	P04075	Fructose-bisphosphate aldolase A	39.3	8.3	45.0	7.8/8.3/8.6	35%	12	154
14	P04406	Glyceraldehyde-3-phosphate dehydrogenase	35.9	8.5	36.9	8.4	53%	16	161
15	P00338	L- Lactate dehydrogenase A chain#	36.5	8.5	32.4	8.5	11%	3	85
16	P07195	L-Lactate dehydrogenase B chain	36.5	5.7	33.4	5.7	44%	19	90
17	P30043	Flavin Reductase	22.1	7.1	24.6	7.2	42%	7	75
18	P32119	Peroxiredoxin-2 (PRDX2)#	21.7	5.6	25.8	5.0	24%	8	114
19	P30041	Peroxiredoxin-6	24.9	6.0	26.0	6.5	37%	7	181
20	P78417	Glutathione S-transferase omega-1 (GSTO1)	27.5	6.2	27.0	6.4	34 %	13	95
21	P04179	Superoxide dismutase [Mn]	24.7	8.3	24.7	7.4	34%	6	62
22	Q99497	Protein DJ-1	19.8	6.3	23.4	6.1	42%	5	70
23	O00299	Chloride intracellular channel protein 1 (CLIC1)	28.7	5.7	28.0	4.9	63%	14	130
24	P11142	Heat shock cognate 71 kDa protein (HSC70)	71.0	5.3	70.0	5.1	36%	25	161
25	P11142	HSC70*	71.0	5.3	24.6	5.3	21%	32	179
26	P11021	78 kDa glucose-regulated protein	72.2	5.0	75.0	4.9	11%	10	87
27	P62937	Peptidyl-prolyl cis-trans isomerase A (CypA)	18.0	7.7	18.7	7.8	61%	11	84
28	P07237	Protein disulfide-	57.1	4.7	62.1	4.4	28%	15	111

		isomerase (PDI)							
29	P30101	Protein disulfide-isomerase A3 (ERp57)	56.7	6.0	60.6	5.4	13%	9	125
30	P30101	ERp57	56.7	6.0	60.4	5.8	13%	9	139
31	P30101	ERp57*	56.7	6.0	50.0	6.2	32%	25	121
32	P30101	ERp57*	56.7	6.0	50.0	6.0	23%	17	94
33	P30040	Endoplasmic reticulum protein ERp29	28.9	6.7	27.2	6.4	29%	8	72
34	P30040	Endoplasmic reticulum protein ERp29	28.9	6.7	27.4	5.8	39%	14	68
35	P27797	Calreticulin (CALR)	48.1	4.2	55.8	3.9	27%	11	90
36	Q06323	Proteasome activator complex subunit 1 (PA28a)	28.7	5.7	29.1	5.4	46%	11	84
37	P60709	Actin*	42.0	5.3	38.0	5.1	42%	19	149
38	P60709	Actin *	42.0	5.3	38.0	5.3	46%	19	380
39	P60709	Actin*	42.0	5.3	26.6	4.9	22%	7	257
40	P60709	Actin*	42.0	5.3	26.5	5.3	18%	6	270
41	P60709	Actin*	42.0	5.3	27.8	5.2	30%	13	317
42	P60709	Actin *	42.0	5.3	26.5	5.1	23%	8	210
43	P60709	Actin*	42.0	5.3	26.5	4.8	62%	26	378
44	P18206	Vinculin	123.7	5.5	120	5.2	20%	19	122
45	P18206	Vinculin*	123.7	5.5	62.3	6.2	16%	16	100
46	P18206	Vinculin	123.7	5.5	126	5.7	57%	74	468
47	P18206	Vinculin	123.7	5.5	127	6.4	21%	27	164
48	P18206	Vinculin	123.7	5.5	130	6.2	12%	15	100
49	P26038	Moesin	67.7	6.0	71.0	6.3	25%	14	83
50	P26038	Moesin	67.7	6.0	71.0	6.7	36%	23	237
51	P26038	Moesin	67.7	6.0	71.0	7.0	7%	4	71
52	P67936	Tropomyosin alpha-4 chain	28.5	4.6	28.2	4.3	43%	18	135
53	P12814	Alpha-Actinin-1	102.9	5.2	125	5.0	26%	19	95
54	P40121	Macrophage-capping protein	38.4	5.8	40.0	5.9	16%	5	67
55	P40121	Macrophage-capping protein-G	38.4	5.8	40.0	6.2	32%	8	66
56	P52907	F-actin capping protein subunit-alpha 1	32.9	5.4	35.2	5.1	31%	7	76
57	Q9Y490	Talin-1*	269.7	5.7	121	5.2	2%	5	72
58	Q9Y490	Talin-1*	269.7	5.7	38.0	6.6	2%	13	83
59	P37802	Transgelin-2	22.3	8.4	21.7	8.3	41%	9	82
60	P21333	Filamin-A*#	280.6	5.7	110	5.7	2%	7	124
61	Q14789	Golgin subfamily B member 1*	375.7	4.9	110	5.0	20%	83	64
62	Q9BZM1	Group XIII secretory phospholipase A2	21.0	6.9	26.0	6.8	26%	5	65
63	P07355	Annexin A2#	38.5	7.5	38.8	7.8	26%	10	228
64	P30740	Leukocyte elastase inhibitor	42.7	5.9	36.5	6.3	17%	5	64
65	Q9NWZ3	Interleukin-1 receptor-associated	51.5	5.2	25.1	6.2	31%	16	59

		kinase 4 (IRAK4)*							
66	Q96AE4	Far upstream element-binding protein 1 (FUBP1)	67.5	7.1	71.0	6.9	29%	19	166
67	Q96AE4	FBP1	67.5	7.1	71.0	7.2	20%	12	74
68	P22627	Heterogeneous nuclear ribonucleoproteins A2/B1(hnRNPA2/B1)	37.4	8.9	38.4	8.9	29%	7	404
69	Q9UJV9	Probable ATP-dependent RNA helicase DDX41	69.7	6.4	56.0	6.8	22%	14	72
70	000170	AH receptor-interacting protein	37.6	6.0	40.0	7.0	27%	10	60
71	Q96BD8	Spindle and kinetochore-associated protein 1	28.9	6.7	27.7	5.6	25%	6	64

¹Entry from the UniProt Knowledgebase database; Mw_{calc} and pl_{calc} , calculated Mw and pi; Mw_{obs} and pl_{obs} , observed Mw and pi; *, protein with lower than expected Mw, possibly a fragment. Proteins were identified based on combined MS and MS/MS on a 4800 MALDI TOF/TOF mass spectrometer (Applied Biosystems) except those indicated by #. These were identified by LC-MS/MS using the microfluid chip-based technology for nanoelectrospray coupled to an ion trap mass spectrometer (Agilent 1200 LC/MSD Trap XCT); MASCOT score, MASCOT protein score derived from the combination of MS and MS/MS data.

Table 7 | Candidate protein biomarkers: 2D DIGE quantitative analysis.

Spot	Protein name	ALS>24 vs controls (fc) ¹	ALS≤24 vs controls (fc)
Energy metabolism			
1	ATP synthase subunit beta	1.6	-1.7
2	Triosephosphate isomerase	1.7	n.c.
3	Phosphoglycerate kinase 1	1.6	-2.0
4	Phosphoglycerate kinase 1*	3.3	4.3
5	Phosphoglycerate kinase 1*	4.5	3.0
6	Phosphoglycerate mutase 1*	1.7	1.5
7	Alpha-enolase	3.1	n.c.
8	Alpha-enolase	2.5	n.c.
9	Alpha-enolase	2.0	n.c.
10	Alpha-enolase	2.7	n.c.
11-13	Fructose-bisphosphate aldolase A	3.2	1.5
14	Glyceraldehyde-3 -phosphate dehydrogenase	1.9	n.c.
15	L- Lactate dehydrogenase A chain	n.c.	-1.4
16	L-Lactate dehydrogenase B chain	2.8	n.c.
Redox regulation			
17	Flavin Reductase	-1.7	n.c.
18	PRDX2	1.7	2.8
19	Peroxiredoxin-6	2.7	1.4
20	GSTO1	n.c.	-1.5
21	Superoxide dismutase [Mn]	n.c.	1.4
22	Protein DJ- 1	2.4	3.6
23	CLIC1	1.5	2.0
Protein folding and degradation			
24	HSC70	1.5	n.c.
25	HSC70*	2.8	2.5
26	78 kDa glucose-regulated protein	nc	-2.9
27	CypA	1.5	n.c.
28	PDI	1.4	1.6
29	ERp57	-1.4	-1.8
30	ERp57	n.c.	2.0
31	ERp57*	2.0	2.0
32	ERp57*	2.0	n.c.
33	Endoplasmic reticulum protein ERp29	1.8	1.8
34	Endoplasmic reticulum protein ERp29	1.6	1.6
35	CALR	2.5	n.c.
36	PA28a	-1.4	-1.7
Cytoskeleton-associated			
37	Actin*	n.c.	2.0
38	Actin *	3.6	n.c.
39	Actin*	2.3	3.2
40	Actin*	1.5	2.8
41	Actin*	1.8	3.5
42	Actin *	1.7	2.5
43	Actin*	2.5	3.9
44	Vinculin*	n.c.	-1.5

45	Vinculin*	1.5	-1.5
46	Vinculin	3.4	2.3
47	Vinculin	3.4	2.3
48	Vinculin	-2.4	-2.8
49	Moesin	-2.2	-1.5
50	Moesin	-1.5	-1.6
51	Moesin	-1.7	-1.9
52	Tropomyosin alpha-4 chain	3.0	n.c.
53	Alpha-Actinin-1	-1.9	-2.0
54	Macrophage-capping protein	1.4	n.c.
55	Macrophage-capping protein	2.8	2.9
56	F-acti capping protein subunit-alpha 1	1.9	1.4
57	Talin-1*	2.5	2.7
58	Talin-1*	3.4	3.0
59	Transgelin-2	2.7	3.5
60	Filamin-A	-1.5	-1.5
61	Golgin_subfamily_B member 1	-1.8	-1.8
Inflammatory response			
62	Group XIII secretory phospholipase A2	1.5	1.8
63	Annexin A2	2.7	1.4
64	Leukocyte elastase inhibitor	4.6	2.2
65	IRAK4*	2.7	n.c.
DNA/RNA binding			
66	FUBP1	-2.0	-3.5
67	FUBP1	n.c.	-1.8
68	hnRNPA2/B1	1.6	n.c.
69	Probable ATP-dependent RNA helicase DDX41	2.7	2.9
Others			
70	AH receptor-interacting protein	2.6	3.2
71	Spindle and kinetochore-associated protein 1	2.0	3.5

¹fc, -fold change: the mean of two 2D DIGE experiments (dye-swap normalization) for increase (positive) and decrease (negative) of normalized spot volumes in samples from ALS patients, ALS>24 (pool of 11) and ALS<24 (pool of 11), in comparison with healthy controls (pool of 11); n.c., no changed, -fold changes <1.4; only -fold changes ≥ 1.4 are reported. *, protein with lower than expected Mw, probably a fragment.

5

Table 8 | Main characteristics of healthy individuals and ALS patients used for the validation analysis (Figure 2A-B).

Sample	Clinical diagnosis	Age ¹	Sex	Score ²	Onset ³	Duration ⁴	Survival ⁵
1-30	Healthy	46±4	18(M), 12(F)	-	-	-	-
31-60	ALS	62±11	15(M), 15(F)	>24			
31	ALS	50	F	32/48	spinal	n.a.	n.a.
32	ALS	61	M	32/48	spinal	29	41
33	ALS	76	M	37/48	spinal	96	111
34	ALS	56	F	25/48	spinal	60	>98*
35	ALS	74	F	28/48	bulbar	17	52
36	ALS	49	M	38/48	spinal	5	32
37	ALS	63	F	32/48	bulbar	16	30

38	ALS	52	M	34/48	bulbar	28	33
39	ALS	66	F	27/48	bulbar	32	>56*
40	ALS	39	F	39/48	spinal	n.a.	n.a.
41	ALS	59	M	33/48	spinal	9	30
42	ALS	58	M	43/48	bulbar	33	59
43	ALS	64	M	29/48	spinal	38	43
44	ALS	68	F	40/48	spinal	n.a.	n.a.
45	ALS	40	F	40/48	spinal	20	>43*
46	ALS	59	M	28/48	spinal	25	>43*
47	ALS	55	M	33/48	spinal	28	57
48	ALS	72	M	35/48	spinal	13	>34*
49	ALS	69	F	26/48	bulbar	26	40
50	ALS	53	F	28/48	spinal	19	21
51	ALS	62	M	29/48	spinal	108	>13 1*
52	ALS	50	F	27/48	spinal	7	9
53	ALS	83	F	25/48	spinal	35	>62*
54	ALS	72	M	25/48	spinal	n.a.	n.a.
55	ALS	52	M	45/48	bulbar	6	29
56	ALS	79	M	28/48	spinal	24	46
57	ALS	81	F	40/48	bulbar	24	51
58	ALS	61	F	40/48	spinal	7	>43*
59	ALS	77	F	27/48	bulbar	32	>56*
60	ALS	61	M	25/48	bulbar	36	>72*
61-90	M:S	62±9	12(M), 18(F)	≤24			
61	ALS	61	M	12/48	spinal	28	30
62	ALS	60	M	21/48	bulbar	17	25
63	ALS	69	F	24/48	bulbar	22	34
64	ALS	63	M	17/48	spinal	24	40
65	ALS	62	M	19/48	spinal	6	11
66	ALS	65	M	22/48	spinal	30	66
67	ALS	64	M	24/48	spinal	43	70
68	ALS	72	F	12/48	spinal	27	29
69	ALS	55	F	13/48	spinal	51	61
70	ALS	48	F	16/48	spinal	19	>55*
71	ALS	42	F	17/48	spinal	38	66
72	ALS	49	F	24/48	spinal	24	54
73	ALS	73	F	15/48	spinal	17	24
74	ALS	56	F	10/48	bulbar	48	57
75	ALS	59	F	10/48	spinal	32	54
76	ALS	73	F	23/48	spinal	21	>43*
77	ALS	48	F	19/48	spinal	10	29
78	ALS	70	F	24/48	bulbar	60	>55*
79	ALS	63	M	11/48	spinal	42	86
80	ALS	64	F	21/48	spinal	22	30
81	ALS	52	F	24/48	spinal	60	75
82	ALS	74	M	17/48	spinal	n.a.	n.a.
83	ALS	63	F	23/48	spinal	16	31
84	ALS	69	F	22/48	bulbar	19	49
85	ALS	63	M	10/48	spinal	10	13
86	ALS	72	F	21/48	spinal	n.a.	n.a.
87	ALS	71	F	20/48	bulbar	16	20
88	ALS	71	M	24/48	spinal	24	26
89	ALS	52	F	17/48	bulbar	8	11
90	ALS	47	M	17/48	spinal	n.a.	n.a.

¹Age at PBMC collection; ²ALSFRRS-R score; ³Site of onset; ⁴Disease duration (months) from the onset of symptoms to PBMC collection; ⁵Disease duration (months) from the onset of symptoms to death; -, not applicable; *, patient still alive (feb-2010); n.a, not available.

Table 9 | Characteristics of ALS patients and ND controls used for the validation analysis (Figure 3).

Sample	Clinical diagnosis	Age ¹	Sex	Score ²	Onset ³	Duration ⁴	Survival ⁵
1-20	ALS	65±9	10(M), 10(F)	>24			
1	ALS	76	M	32/48	bulbar	16	30
2	ALS	58	M	25/48	spinal	60	61
3	ALS	71	F	30/48	spinal	27	48
4	ALS	61	M	26/48	bulbar	36	>72*
5	ALS	63	M	26/48	bulbar	25	30
6	ALS	52	M	25/48	spinal	36	58
7	ALS	67	M	39/48	spinal	25	30
8	ALS	56	F	32/48	bulbar	36	51
9	ALS	53	M	25/48	spinal	16	31
10	ALS	77	F	27/48	spinal	7	9
11	ALS	75	F	35/48	spinal	24	44
12	ALS	77	F	26/48	bulbar	26	40
13	ALS	66	F	27/48	bulbar	32	>56*
14	ALS	68	F	40/48	spinal	n.a.	n.a.
15	ALS	68	M	28/48	bulbar	16	27
16	ALS	67	F	34/48	spinal	5	11
17	ALS	70	F	25/48	spinal	2	15
18	ALS	58	M	45/48	spinal	4	15
19	ALS	72	M	30/48	bulbar	9	17
20	ALS	66	F	35/48	spinal	10	24
21-43	ND controls	64±14	13(M), 11(F)	-	-	-	-
21	diabetic peripheral neuropathy	79	M	-	-	-	-
22	diabetic peripheral neuropathy	80	F	-	-	-	-
23	diabetic peripheral neuropathy	70	F	-	-	-	-
24	CIDP ⁶	79	F	-	-	-	-
25	CIDP	61	M	-	-	-	-
26	CIDP	62	M	-	-	-	-
27	CIDP	81	M	-	-	-	-
28	alcoholic polyneuropathy	38	M	-	-	-	-
29	alcoholic polyneuropathy	49	M	-	-	-	-
30	alcoholic polyneuropathy	60	M	-	-	-	-
31	axonal peripheral neuropathy	79	M	-	-	-	-
32	axonal peripheral neuropathy	76	F	-	-	-	-
33	axonal peripheral neuropathy	69	F	-	-	-	-
34	Charcot-Marie-Tooth disease	38	F	-	-	-	-
35	spinocerebellar ataxia	43	F	-	-	-	-
36	multiple sclerosis	51	F	-	-	-	-
37	multiple sclerosis	52	F	-	-	-	-
38	multiple sclerosis	59	F	-	-	-	-
39	myelopathy localized at C6-C7	61	M	-	-	-	-

40	transverse myelitis	63	M	-	-	-	-
41	transverse myelitis	73	M	-	-	-	-
42	polymyositis	74	M	-	-	-	-
43	Lambert-Eaton myasthenic syndrome	77	M	-	-	-	-

¹Age at PBMC collection; ²ALSFRS-R score; ³Site of onset; ⁴Disease duration (months) from the onset of symptoms to PBMC collection; ⁵Disease duration (months) from the onset of symptoms to death; -, not applicable; ⁶CIDP, chronic inflammatory demyelinating polyneuropathy; *, patient still alive (feb-2010); n.a, not available.

5

Table 10| Univariate logistic regression: controls versus ALS>24

Protein	OR	95% CI		P-value	AUC
CALR	2.002	1.268	3.163	0.0029	0.711
TDP-43	1.377	1.010	1.877	0.0430	0.657
PRDX2	6.600	1.807	24.101	0.0043	0.723
PDI	48.689	4.370	542.533	0.0016	0.818
ERp57	4.819	1.940	11.971	0.0007	0.816
PA28a	0.608	0.412	0.898	0.0123	0.721
CLIC1	39.908	5.412	294.281	0.0003	0.794
IRAK4	13.379	3.207	55.810	0.0004	0.852
FUBP1	0.283	0.115	0.695	0.0059	0.806
GSTO1	0.278	0.098	0.785	0.0157	0.701
HSC70	4.167	1.664	10.437	0.0023	0.742
CypA	3.312	1.516	7.235	0.0027	0.787
Actin ^{NI}	79.972	7.635	837.669	0.0003	0.864
hnRNPA2/B1	1.865	1.305	2.664	0.0006	0.792

Results are expressed as odds ratios (OR) and 95% confidence intervals (95% CI). A 95% CI not including the value of 1 indicates a statistically significant result. All probability values were two-sided and p<0.05 was considered statistically significant (values in bold type).

10

Table 11| Univariate logistic regression: controls versus ALS<24

Protein	OR	95% CI		P-value	AUC
CALR	2.863	1.367	5.995	0.0053	0.752
TDP-43	1.859	1.320	2.620	0.0004	0.810
PRDX2	1.027	1.012	1.041	0.0003	0.836
PDI	1.043	1.018	1.068	0.0006	0.843
ERp57	1.020	1.008	1.032	0.0007	0.893
PA28a	0.707	0.527	0.949	0.0209	0.715
CLIC1	1.036	1.016	1.057	0.0004	0.794
IRAK4	7.887	2.262	27.497	0.0012	0.835
FUBP1	0.608	0.379	0.976	0.0392	0.723
GSTO1	0.061	0.010	0.354	0.0019	0.771
HSC70	6.794	2.342	19.709	0.0004	0.822
CypA	1.830	1.001	3.345	0.0497	0.620

Table 11| Univariate logistic regression: controls versus ALS<24

Protein	OR	95% CI		P-value	AUC
Actin ^{NT}	1.059	1.026	1.093	0.0003	0.937
hnRNPA2/B1	1.931	1.359	2.745	0.0002	0.832

Results are expressed as odds ratios (OR) and 95% confidence intervals (95% CI). A 95% CI not including the value of 1 indicates a statistically significant result. All probability values were two-sided and $p < 0.05$ was considered statistically significant (values in bold type).

5

Table 12| Characteristics of sALS patients used for the longitudinal study (Figure 6)

Sample	Clinical diagnosis	Age ¹	Sex	Score T1 ²	Score T2 ²	Score T3 ²	Site of Onset	Disease duration ³
1	ALS	38	M	41	34	30	spinal	5
2	ALS	64	M	30	20	19	bulbar	47
3	ALS	80	M	24	22	23	bulbar	19
4	ALS	73	M	42	37	36	spinal	15
5	ALS	69	M	37	31	30	bulbar	16
6	ALS	63	M	34	31	24	bulbar	21
7	ALS	69	M	35	27	27	spinal	10
8	ALS	64	M	28	23	12	bulbar	22
9	ALS	59	F	30	29	31	spinal	18
10	ALS	71	F	33	28	13	spinal	26
11	ALS	57	F	38	35	32	spinal	23
12	ALS	74	M	28	21	16	bulbar	35
13	ALS	60	F	31	29	30	spinal	20

¹Age at first PBMC collection; ²ALSFRS-R score at PBMC collection: at t=0 (T1), at t=3 months (T2) and at t=6 months (T3); ³Disease duration (months) from the onset of symptoms to first PBMC collection.

References

1. Chio, A., Logroscino, G., Hardiman, O., Swingler, R., Mitchell, D., Beghi, E., and Traynor, B.G. 2008. Prognostic factors in ALS: A critical review. *Amyotroph Lateral Scler.* 4.
- 5 2. Chio, A., Mora, G., Calvo, A., Mazzini, L., Bottacchi, E., and Mutani, R. 2009. Epidemiology of ALS in Italy: a 10-year prospective population-based study. *Neurology* 72:725-731.
3. Zoccolella, S., Beghi, E., Palagano, G., Fraddosio, A., Samarelli, V., Lamberti, P., Lepore, V., Serlenga, L., and Logroscino, G. 2006. Predictors of delay in the
10 diagnosis and clinical trial entry of amyotrophic lateral sclerosis patients: a population-based study. *J Neurol Sci* 250:45-49.
4. Turner, M.R., Kiernan, M.C., Leigh, P.N., and Talbot, K. 2009. Biomarkers in amyotrophic lateral sclerosis. *Lancet Neurol* 8:94-109.
5. Aguirre, T., Van Den Bosch, L., Goetschalckx, K., Tilkin, P., Mathijs, G.,
15 Cassiman, J.J., and Robberecht, W. 1998. Increased sensitivity of fibroblasts from amyotrophic lateral sclerosis patients to oxidative stress. *Ann Neurol* 43:452-457.
6. Boillee, S., Vande Velde, C , and Cleveland, D.W. 2006. ALS: a disease of motor neurons and their nonneuronal neighbors. *Neuron* 52:39-59.
7. Curti, D., Malaspina, A., Facchetti, G., Camana, C , Mazzini, L., Tosca, P., Zerbi,
20 F., and Ceroni, M. 1996. Amyotrophic lateral sclerosis: oxidative energy metabolism and calcium homeostasis in peripheral blood lymphocytes. *Neurology* 47:1060-1064.
8. Dobrowolny, G., Aucello, M., Rizzuto, E., Beccafico, S., Mammucari, C ,
25 Boncompagni, S., Belia, S., Wannenes, F., Nicoletti, C , Del Prete, Z., et al. 2008. Skeletal muscle is a primary target of SOD1G93A-mediated toxicity. *Cell Metab* 8:425-436.
9. Ono, S. 2000. The skin in amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Other Motor Neuron Disord* 1:191-199.
10. Wong, M., and Martin, L.J. 2010. Skeletal Muscle-Restricted Expression of Human
30 SOD1 Causes Motor Neuron Degeneration in Transgenic Mice. *Hum Mol Genet.*[Epub ahead of print].
11. Cova, E., Cereda, C , Galli, A., Curti, D., Finotti, C , Di Poto, C , Corato, M.,
Mazzini, G., and Ceroni, M. 2006. Modified expression of Bcl-2 and SOD1 proteins in lymphocytes from sporadic ALS patients. *Neurosci Lett* 399:186-190.
- 35 12. Mantovani, S., Garbelli, S., Pasini, A., Alimonti, D., Perotti, C , Melazzini, M., Bendotti, C , and Mora, G. 2009. Immune system alterations in sporadic amyotrophic lateral sclerosis patients suggest an ongoing neuroinflammatory process. *J Neuroimmunol* 210:73-79.
13. Nardo, G., Pozzi, S., Mantovani, S., Garbelli, S., Marinou, K., Basso, M., Mora, G.,
40 Bendotti, C , and Bonetto, V. 2009. Nitroproteomics of peripheral blood mononuclear cells from patients and a rat model of ALS. *Antioxid Redox Signal* 11:1559-1567.
14. Pouloupoulou, C , Davaki, P., Koliaraki, V., Kolovou, D., Markakis, I., and Vassilopoulos, D. 2005. Reduced expression of metabotropic glutamate receptor
45 2mRNA in T cells of ALS patients. *Ann Neurol* 58:946-949.
15. Alban, A., David, S.O., Bjorkesten, L., Andersson, C , Sloge, E., Lewis, S., and Currie, I. 2003. A novel experimental design for comparative two-dimensional gel

- analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics* 3:36-44.
16. Hu, Y., Malone, J.P., Fagan, A.M., Townsend, R.R., and Holtzman, D.M. 2005. Comparative proteomic analysis of intra- and interindividual variation in human cerebrospinal fluid. *Mol Cell Proteomics* 4:2000-2009.
- 5 17. Zhang, J. 2007. Proteomics of human cerebrospinal fluid - the good, the bad, and the ugly. *Proteomics Clin Appl* 1:805-819.
18. Cedarbaum, J.M., Stambler, N., Malta, E., Fuller, C, Hilt, D., Thurmond, B., and Nakanishi, A. 1999. The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). *J Neurol Sci* 169:13-21.
- 10 19. Atkin, J.D., Farg, M.A., Turner, B.J., Tomas, D., Lysaght, J.A., Nunan, J., Rembach, A., Nagley, P., Beart, P.M., Cheema, S.S., et al. 2006. Induction of the unfolded protein response in familial amyotrophic lateral sclerosis and association of protein-disulfide isomerase with superoxide dismutase 1. *J Biol Chem* 281:30152-30165.
- 15 20. Atkin, J.D., Farg, M.A., Walker, A.K., McLean, C, Tomas, D., and Home, M.K. 2008. Endoplasmic reticulum stress and induction of the unfolded protein response in human sporadic amyotrophic lateral sclerosis. *Neurobiol Dis* 30:400-407.
- 20 21. Kato, S., Kato, M., Abe, Y., Matsumura, T., Nishino, T., Aoki, M., Itoyama, Y., Asayama, K., Awaya, A., Hirano, A., et al. 2005. Redox system expression in the motor neurons in amyotrophic lateral sclerosis (ALS): immunohistochemical studies on sporadic ALS, superoxide dismutase 1 (SOD1)-mutated familial ALS, and SOD1-mutated ALS animal models. *Acta Neuropathol* 110:101-112.
- 25 22. Massignan, T., Casoni, F., Basso, M., Stefanazzi, P., Biasini, E., Tortarolo, M., Salmona, M., Gianazza, E., Bendotti, C, and Bonetto, V. 2007. Proteomic analysis of spinal cord of presymptomatic amyotrophic lateral sclerosis G93A SOD1 mouse. *Biochem Biophys Res Commun* 353:719-725.
- 30 23. van de Giessen, E., Fogh, I., Gopinath, S., Smith, B., Hu, X., Powell, J., Andersen, P., Nicholson, G., Al Chalabi, A., and Shaw, C.E. 2008. Association study on glutathione S-transferase omega 1 and 2 and familial ALS. *Amyotroph Lateral Scler* 9:81-84.
24. Watanabe, M., Dykes-Hoberg, M., Culotta, V.C., Price, D.L., Wong, P.C., and Rothstein, J.D. 2001. Histological evidence of protein aggregation in mutant SOD1 transgenic mice and in amyotrophic lateral sclerosis neural tissues. *Neurobiol Dis* 8:933-941.
- 35 25. Casoni, F., Basso, M., Massignan, T., Gianazza, E., Cheroni, C, Salmona, M., Bendotti, C, and Bonetto, V. 2005. Protein nitration in a mouse model of familial amyotrophic lateral sclerosis: possible multifunctional role in the pathogenesis. *J Biol Chem* 280:16295-16304.
- 40 26. Neumann, M., Sampathu, D.M., Kwong, L.K., Truax, A.C., Micsenyi, M.C., Chou, T.T., Bruce, J., Schuck, T., Grossman, M., Clark, CM., et al. 2006. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* 314:130-133.
- 45 27. Kabashi, E., Valdmanis, P.N., Dion, P., Spiegelman, D., McConkey, B.J., Vande Velde, C, Bouchard, J.P., Lacomblez, L., Pochigaeva, K., Salachas, F., et al. 2008. TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat Genet* 40:572-574.

28. Sreedharan, J., Blair, L.P., Tripathi, V.B., Hu, X., Vance, C, Rogelj, B., Ackerley, S., Durnall, J.C., Williams, K.L., Buratti, E., et al. 2008. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* 319:1668-1672.
29. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M.,
5 Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25:25-29.
30. Hartwell, L.H., Hopfield, J.J., Leibler, S., and Murray, A.W. 1999. From molecular to modular cell biology. *Nature* 402:C47-52.
- 10 31. Achiron, A., and Gurevich, M. 2006. Peripheral blood gene expression signature mirrors central nervous system disease: the model of multiple sclerosis. *Autoimmun Rev* 5:517-522.
32. Borovecki, F., Lovrecic, L., Zhou, J., Jeong, H., Then, F., Rosas, H.D., Hersch, S.M., Hogarth, P., Bouzou, B., Jensen, R.V., et al. 2005. Genome-wide expression profiling of human blood reveals biomarkers for Huntington's disease. *Proc Natl AcadSci USA* 102:1 1023-1 1028.
- 15 33. Maes, O.C., Xu, S., Yu, B., Chertkow, H.M., Wang, E., and Schipper, H.M. 2007. Transcriptional profiling of Alzheimer blood mononuclear cells by microarray. *Neurobiol Aging* 28:1795-1809.
- 20 34. Mitchell, R.M., Freeman, W.M., Randazzo, W.T., Stephens, H.E., Beard, J.L., Simmons, Z., and Connor, J.R. 2009. A CSF biomarker panel for identification of patients with amyotrophic lateral sclerosis. *Neurology* 72:14-19.
35. Pasinetti, G.M., Ungar, L.H., Lange, D.J., Yemul, S., Deng, H., Yuan, X., Brown, R.H., Cudkowicz, M.E., Newhall, K., Peskind, E., et al. 2006. Identification of
25 potential CSF biomarkers in ALS. *Neurology* 66:1218-1222.
36. Davenport, R.J., Swingler, R.J., Chancellor, A.M., and Warlow, CP. 1996. Avoiding false positive diagnoses of motor neuron disease: lessons from the Scottish Motor Neuron Disease Register. *J Neurol Neurosurg Psychiatry* 60:147-151.
- 30 37. Traynor, B.J., Codd, M.B., Corr, B., Forde, C, Frost, E., and Hardiman, O. 2000. Amyotrophic lateral sclerosis mimic syndromes: a population-based study. *Arch Neurol* 57:109-1 13.
38. Kraemer, M., Buerger, M., and Berlit, P. 2009. Diagnostic problems and delay of diagnosis in amyotrophic lateral sclerosis. *Clin Neurol Neurosurg*.
- 35 39. Srinivasan, J., Scala, S., Jones, H.R., Saleh, F., and Russell, J.A. 2006. Inappropriate surgeries resulting from misdiagnosis of early amyotrophic lateral sclerosis. *Muscle Nerve* 34:359-360.
40. Suzuki, N., and Saito, T. 2006. IRAK-4~a shared NF-kappaB activator in innate and acquired immunity. *Trends Immunol* 27:566-572.
- 40 41. Tang, Y., Lu, A., Aronow, B.J., and Sharp, F.R. 2001. Blood genomic responses differ after stroke, seizures, hypoglycemia, and hypoxia: blood genomic fingerprints of disease. *Ann Neurol* 50:699-707.
42. Gladkevich, A., Kauffman, H.F., and Korf, J. 2004. Lymphocytes as a neural probe: potential for studying psychiatric disorders. *Prog Neuropsychopharmacol Biol Psychiatry* 28:559-576.
- 45 43. Basso, M., Samengo, G., Nardo, G., Massignan, T., D'Alessandro, G., Tartari, S., Cantoni, L., Marino, M., Cheroni, C, S., D.B., et al. 2009. Characterization of detergent-insoluble proteins in ALS indicates a causal link between nitrate stress and aggregation in pathogenesis. *PLoS ONE* 4:e8130.

44. Kim, I., Xu, W., and Reed, J.C. 2008. Cell death and endoplasmic reticulum stress: disease relevance and therapeutic opportunities. *Nat Rev Drug Discov* 7:1013-1030.
45. Board, P.G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermini, L.S., Schulte, G.K., Danley, D.E., Hoth, L.R., Griffor, M.C., Kamath, A.V., et al. 2000. Identification, characterization, and crystal structure of the Omega class glutathione transferases. *J Biol Chem* 275:24798-24806.
- 5 46. Ghezzi, P., Casagrande, S., Massignan, T., Basso, M., Bellacchio, E., Mollica, L., Biasini, E., Tonelli, R., Eberini, I., Gianazza, E., et al. 2006. Redox regulation of cyclophilin A by glutathionylation. *Proteomics* 6:817-825.
- 10 47. Milton, R.H., Abeti, R., Averaimo, S., DeBiasi, S., Vitellaro, L., Jiang, L., Curmi, P.M., Breit, S.N., Duchen, M.R., and Mazzanti, M. 2008. CLIC1 function is required for beta-amyloid-induced generation of reactive oxygen species by microglia. *J Neurosci* 28:1 1488-1 1499.
48. Wood, Z.A., Schroder, E., Robin Harris, J., and Poole, L.B. 2003. Structure, mechanism and regulation of peroxiredoxins. *Trends Biochem Sci* 28:32-40.
- 15 49. Beal, M.F., Ferrante, R.J., Browne, S.E., Matthews, R.T., Kowall, N.W., and Brown, R.H., Jr. 1997. Increased 3-nitrotyrosine in both sporadic and familial amyotrophic lateral sclerosis. *Ann Neurol* 42:644-654.
50. Ferrante, R.J., Shinobu, L.A., Schulz, J.B., Matthews, R.T., Thomas, C.E., Kowall, N.W., Gurney, M.E., and Beal, M.F. 1997. Increased 3-nitrotyrosine and oxidative damage in mice with a human copper/zinc superoxide dismutase mutation. *Ann Neurol* 42:326-334.
- 20 51. Barber, S.C., Mead, R.J., and Shaw, P.J. 2006. Oxidative stress in ALS: a mechanism of neurodegeneration and a therapeutic target. *Biochim Biophys Acta* 1762:1051-1067.
- 25 52. Ilieva, E.V., Ayala, V., Jove, M., Dalfo, E., Cacabelos, D., Povedano, M., Bellmunt, M.J., Ferrer, I., Pamplona, R., and Portero-Otin, M. 2007. Oxidative and endoplasmic reticulum stress interplay in sporadic amyotrophic lateral sclerosis. *Brain* 130:31 11-3123.
- 30 53. Krecic, A.M., and Swanson, M.S. 1999. hnRNP complexes: composition, structure, and function. *Curr Opin Cell Biol* 11:363-371.
54. Buratti, E., Brindisi, A., Giombi, M., Tisminetzky, S., Ayala, Y.M., and Baralle, F.E. 2005. TDP-43 binds heterogeneous nuclear ribonucleoprotein A/B through its C-terminal tail: an important region for the inhibition of cystic fibrosis transmembrane conductance regulator exon 9 splicing. *J Biol Chem* 280:37572-37584.
- 35 55. Anthony, K., and Gallo, J.M. 2010. Aberrant RNA processing events in neurological disorders. *Brain Res* 1338C:67-77.
56. Ferraiuolo, L., Heath, P.R., Holden, H., Kasher, P., Kirby, J., and Shaw, P.J. 2007. Microarray analysis of the cellular pathways involved in the adaptation to and progression of motor neuron injury in the SOD1 G93A mouse model of familial ALS. *J Neurosci* 27:9201-9219.
- 40 57. Lukas, T.J., Luo, W.W., Mao, H., Cole, N., and Siddique, T. 2006. Informatics-assisted protein profiling in a transgenic mouse model of amyotrophic lateral sclerosis. *Mol Cell Proteomics* 5:1233-1244.
- 45 58. Malaspina, A., Jokic, N., Huang, W.L., and Priestley, J.V. 2008. Comparative analysis of the time-dependent functional and molecular changes in spinal cord degeneration induced by the G93A SOD1 gene mutation and by mechanical compression. *BMC Genomics* 9:500.

59. Kwiatkowski, T.J., Jr., Bosco, D.A., Leclerc, A.L., Tamrazian, E., Vanderburg, C.R., Russ, C, Davis, A., Gilchrist, J., Kasarskis, E.J., Munsat, T., et al. 2009. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* 323:1205-1208.
- 5 60. Vance, C, Rogelj, B., Hortobagyi, T., De Vos, K.J., Nishimura, A.L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P., et al. 2009. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 323:1208-1211.
- 10 61. Saris, C.G., Horvath, S., van Vught, P.W., van Es, M.A., Blauw, H.M., Fuller, T.F., Langfelder, P., DeYoung, J., Wokke, J.H., Veldink, J.H., et al. 2009. Weighted gene co-expression network analysis of the peripheral blood from Amyotrophic Lateral Sclerosis patients. *BMC Genomics* 10:405.
- 15 62. Brooks, B.R. 1994. El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical limits of amyotrophic lateral sclerosis" workshop contributors. *J Neurol Sci* 124 Suppl:96-107.
- 20 63. Pappin, D.J., Hojrup, P., and Bleasby, A.J. 1993. Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol* 3:327-332.
64. Carpi, D., Korkalainen, M., Airoidi, L., Fanelli, R., Hakansson, FL, Muhonen, V., Tuukkanen, J., Viluksela, M., and Pastorelli, R. 2009. Dioxin-sensitive proteins in differentiating osteoblasts: effects on bone formation in vitro. *Toxicol Sci* 108:330-343.
- 25 65. Keshava Prasad, T.S., Goel, R., Kandasamy, K., Keerthikumar, S., Kumar, S., Mathivanan, S., Telikicherla, D., Raju, R., Shafreen, B., Venugopal, A., et al. 2009. Human Protein Reference Database—2009 update. *Nucleic Acids Res* 37:D767-772.
66. Mishra, G.R., Suresh, M., Kumaran, K., Kannabiran, N., Suresh, S., Bala, P., Shivakumar, K., Anuradha, N., Reddy, R., Raghavan, T.M., et al. 2006. Human protein reference database—2006 update. *Nucleic Acids Res* 34:D411-414.
- 30 67. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498-2504.
- 35 68. Assenov, Y., Ramirez, F., Schelhorn, S.E., Lengauer, T., and Albrecht, M. 2008. Computing topological parameters of biological networks. *Bioinformatics* 24:282-284.
69. Rives, A.W., and Galitski, T. 2003. Modular organization of cellular networks. *Proc Natl Acad Sci USA* 100:1128-1133.
- 40 70. Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95:14863-14868.
71. de Hoon, M.J., Imoto, S., Nolan, J., and Miyano, S. 2004. Open source clustering software. *Bioinformatics* 20:1453-1454.
- 45 72. Saldanha, A.J. 2004. Java Treeview—extensible visualization of microarray data. *Bioinformatics* 20:3246-3248.
73. Maere, S., Heymans, K., and Kuiper, M. 2005. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* 21:3448-3449.

74. Reiner, A., Yekutieli, D., and Benjamini, Y. 2003. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* 19:368-375.
- 5 75. Martin, D., Brun, C, Remy, E., Mouren, P., Thieffry, D., and Jacq, B. 2004. GOToolBox: functional analysis of gene datasets based on Gene Ontology. *Genome Biol* 5:R101.

Claims

1. A method to diagnose and/or to make prognostic predictions and/or to monitor the efficacy of a therapy of Amyotrophic lateral sclerosis (ALS) in a subject comprising the steps of:

- 5 a) detecting or measuring the amount of at least three biomarkers each one belonging to a different group of groups I to VII as listed in Table 1, or allelic variants, or protein precursors, or functional fragments, or post-translationally modified isoforms or gene transcripts thereof, from peripheral blood mononuclear cells (PBMCs), or extracts thereof, isolated from a biological sample of the subject;
- 10 b) eventually comparing such level with respect to a proper control.

2. A method to diagnose and/or to make prognostic predictions and/or to monitor the efficacy of a therapy of Amyotrophic lateral sclerosis (ALS) in a subject according to claim 1 comprising the steps of:

- a) detecting or measuring the amount of at least three biomarkers belonging to the group
15 of: Calreticulin (CALR) (SEQ ID No. 22); TAR DNA-binding protein 43 (TDP-43) (SEQ ID No. 42); Peroxiredoxin-2 (PRDX2) (SEQ ID No. 10); Protein disulfide-isomerase (PDI) (SEQ ID No. 19); protein disulfide-isomerase A3 (ERp57) (SEQ ID No. 20); Proteasome activator complex subunit 1 (PA28a) (SEQ ID No. 23); chloride intracellular channel protein 1 (CLIC1) (SEQ ID No. 15); interleukin-1 receptor-associated kinase 4 (IRAK4)
20 (SEQ ID No. 38); Far upstream element-binding protein 1 (FUBP1) (SEQ ID No. 39); Glutathione S-transferase omega-1 (GSTO1) (SEQ ID No. 12); Heat shock cognate 71 kDa protein (HSC70) (SEQ ID No. 16); peptidyl-prolyl cis-trans isomerase A (CypA) (SEQ ID No. 18); tyrosine nitrated actin (Actin^{NT}) (SEQ ID No. 24); heterogeneous nuclear ribonucleoprotein A2/B1 (hnRNPA2/B1) (SEQ ID No. 40), or allelic variants, or protein
25 precursors, or functional fragments, or post-translationally modified isoforms or gene transcripts thereof, from peripheral blood mononuclear cells (PBMCs), or extracts thereof, isolated from a biological sample of the subject;
- b) eventually comparing such level with respect to a proper control.

3. The method to diagnose Amyotrophic lateral sclerosis (ALS) in a subject
30 according to claim 2 wherein step a) is performed by detecting or measuring the amount of at least: chloride intracellular channel protein 1 (CLIC1) (SEQ ID No. 15); tyrosine nitrated actin (actin^{NT}) (SEQ ID No. 24); heterogeneous nuclear ribonucleoprotein A2/B1 (hnRNPA2/B1) (SEQ ID No. 40), or allelic variants, or protein precursors, or functional

fragments, or post-translationally modified isoforms or gene transcripts thereof.

4. The method to diagnose Amyotrophic lateral sclerosis (ALS) in a subject according to claim 3 wherein step a) is performed by detecting or measuring the amount of also Interleukin-1 receptor-associated kinase 4 (IRAK4) (SEQ ID No. 38) and Peptidyl-prolyl cis-trans isomerase A (CypA) (SEQ ID No. 18), or allelic variants, or protein precursors, or functional fragments, or post-translationally modified isoforms or gene transcripts thereof.

5. The method to make prognostic predictions and/or to monitor the efficacy of a therapy of Amyotrophic lateral sclerosis (ALS) in a subject according to claim 2 wherein step a) is performed by detecting or measuring the amount of at least: TDP-43 (SEQ ID No. 42); protein disulfide-isomerase A3 (ERp57) (SEQ ID No. 20); peptidyl-prolyl cis-trans isomerase A (CypA) (SEQ ID No. 18), or allelic variants, or protein precursors, or functional fragments, or post-translationally modified isoforms or gene transcripts thereof.

6. The method according to any of previous claims wherein the detecting and measuring is performed by immunochemical techniques or mass spectrometric techniques.

7. The method according to any of previous claims wherein the subject is an human or a mammalian animal.

8. The method according to any of previous claims wherein peripheral blood mononuclear cells (PBMCs) isolated from a biological sample are freshly isolated or properly stored.

9. A diagnostic kit to diagnose and/or to make prognostic predictions and/or to monitor the efficacy of a therapy of Amyotrophic lateral sclerosis (ALS) in a subject comprising selective ligands for the biomarkers or allelic variants, or protein precursors, or functional fragments, or post-translationally modified isoforms or gene transcripts thereof, as described in any of claims 1 to 5.

10. The diagnostic kit according to claim 9 wherein the selective ligands are immunoglobulins or synthetic or recombinant fragments thereof.

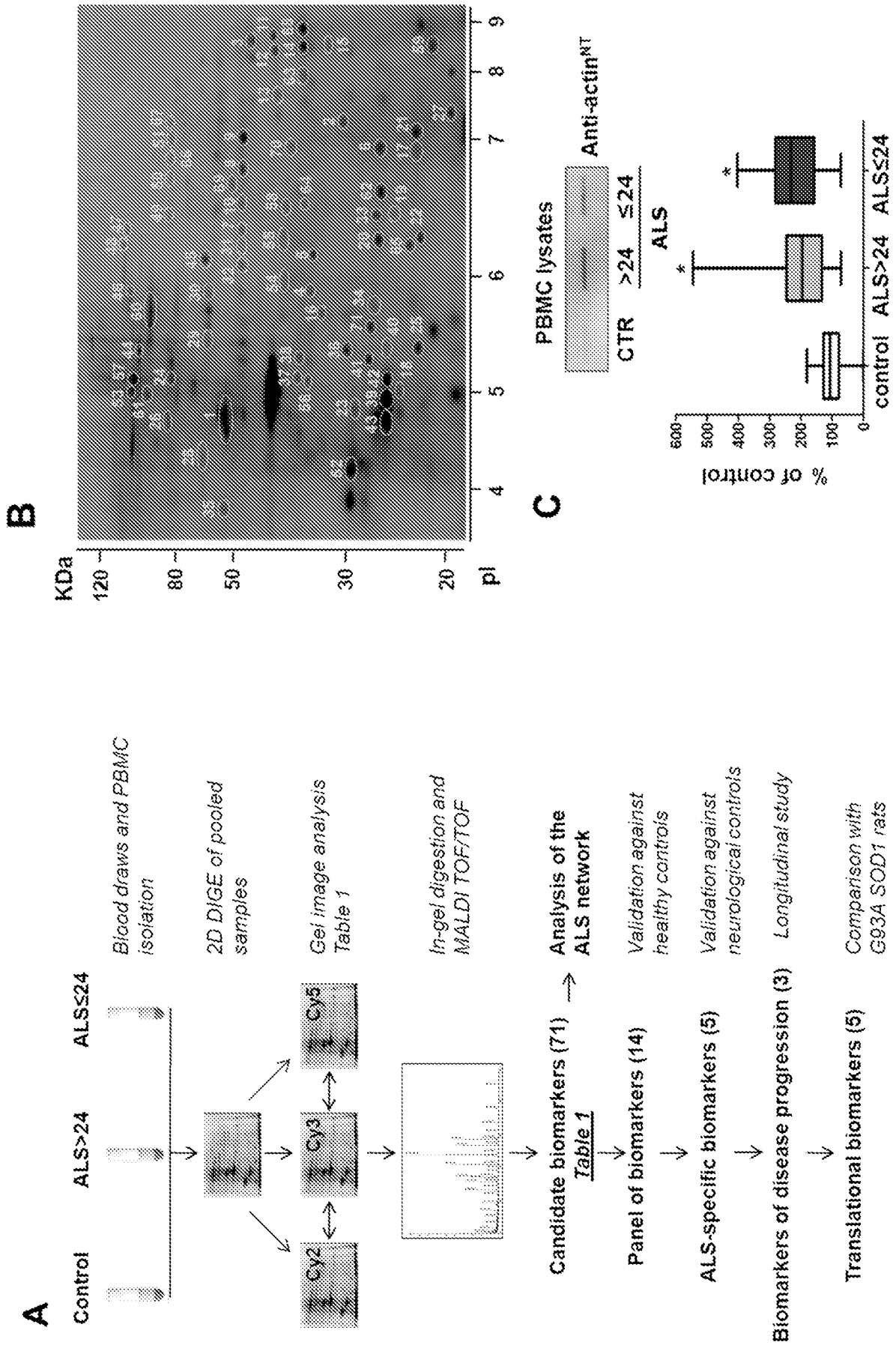
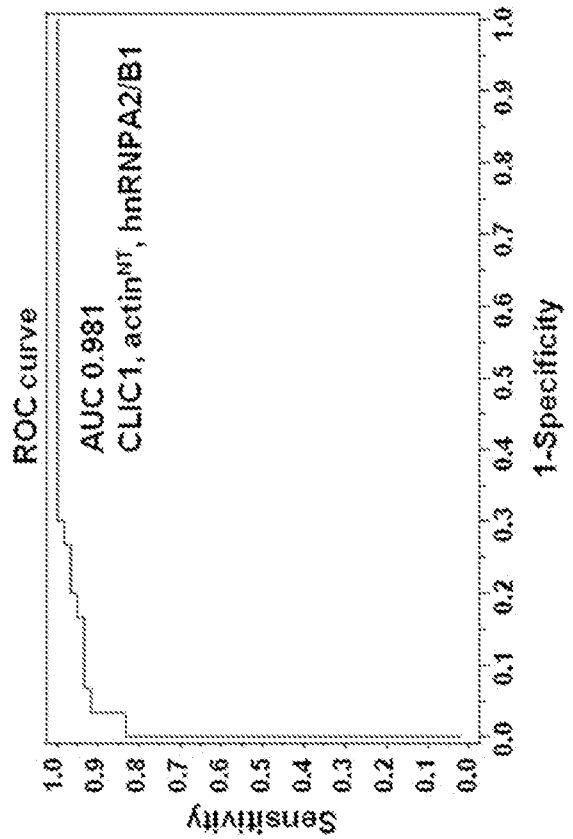


Fig. 1

A

ALS versus healthy controls					
Protein	OR	95% CI	P-value	AUC	
CALR	2.100	1.320	0.0018	0.732	
TDP-43	1.567	1.195	0.0011	0.733	
PRDX2	8.253	2.753	0.0002	0.779	
PDI	43.269	6.364	0.0001	0.831	
ERp57	2.238	1.253	0.0065	0.697	
PA28a	0.667	0.512	0.0027	0.718	
CLIC1	33.935	6.164	<0.0001	0.794	
IRAK4	10.750	3.263	<0.0001	0.843	
FUBP1	0.503	0.324	0.0021	0.764	
GSTO1	0.191	0.070	0.0011	0.736	
HSC70	5.121	2.147	0.0002	0.782	
CypA	2.267	1.278	0.0051	0.703	
Actin ^{NT}	109.248	12.469	<0.0001	0.981	
hnRNPA2/B1	1.916	1.401	<0.0001	0.812	



B

ALS>24 versus ALS<24					
Protein	OR	95% CI	P-value	AUC	
CALR	0.918	0.747	0.4168	0.525	
TDP-43	1.361	1.066	0.0134	0.687	
PRDX2	1.909	0.983	0.0584	0.665	
PDI	1.304	0.614	0.4895	0.568	
ERp57	7.349	2.327	0.0007	0.893	
PA28a	1.096	0.752	0.6341	0.494	
CLIC1	1.422	0.474	0.5301	0.533	
IRAK4	1.088	0.767	0.6369	0.522	
FUBP1	1.671	0.855	0.1328	0.608	
GSTO1	0.564	0.182	0.3213	0.547	
HSC70	1.214	0.868	0.2577	0.562	
CypA	0.660	0.456	0.0277	0.671	
Actin ^{NT}	1.315	0.607	0.4880	0.587	
hnRNPA2/B1	1.094	0.911	0.3374	0.567	

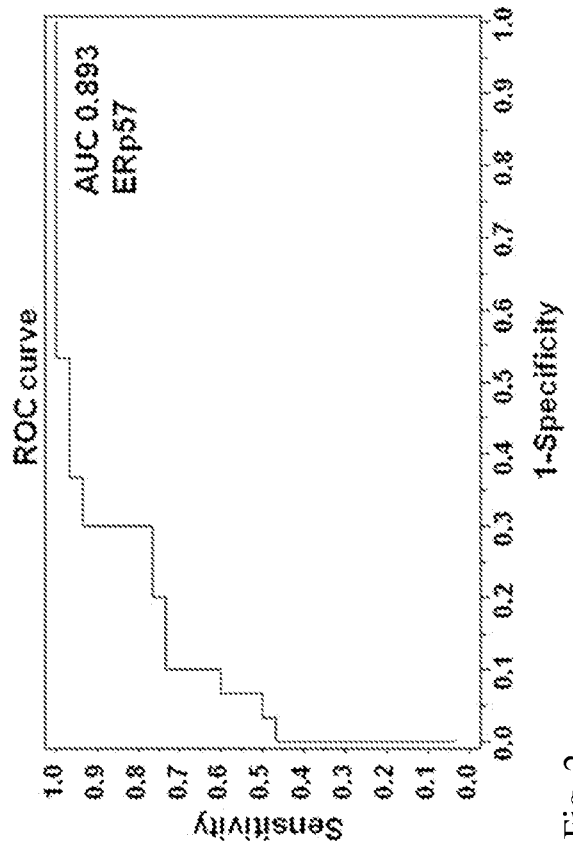


Fig. 2

ALS>24 versus neurological disorder controls					
Protein	OR	95% CI		P-value	AUC
CALR	1.487	1.044	2.116	0.0278	0.688
TDP-43	1.808	0.613	5.338	0.2833	0.624
PRDX2	0.146	0.004	5.257	0.2927	0.569
PDI	1.004	0.988	1.021	0.6273	0.574
ERp57	1.053	0.969	1.145	0.2214	0.629
PA28a	4.523	0.093	221.137	0.4470	0.571
CLIC1	1.007	1.001	1.012	0.0212	0.708
IRAK4	0.967	0.939	0.996	0.0239	0.707
FUBP1	0.746	0.501	1.111	0.1494	0.611
GSTO1	8.201	1.098	61.271	0.0403	0.693
HSC70	1.218	0.919	1.614	0.1694	0.609
CypA	2.385	1.235	4.606	0.0097	0.763
Actin ^{NT}	0.998	0.993	1.003	0.5163	0.546
hnRNPA2/B1	1.244	0.640	2.418	0.5190	0.562

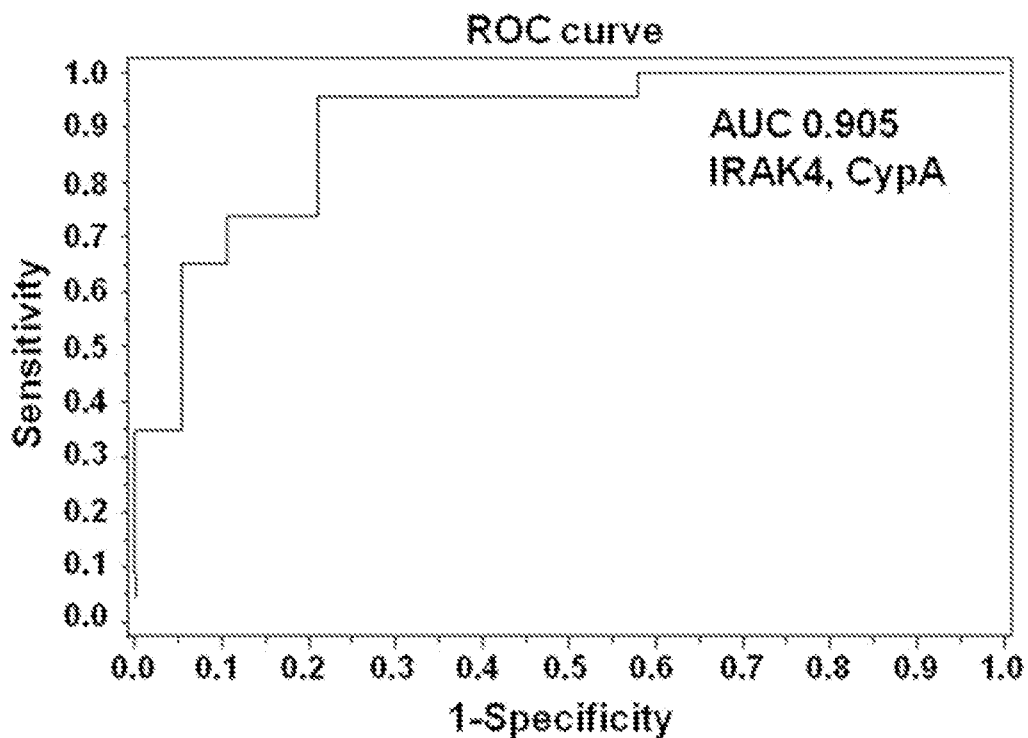


Fig. 3

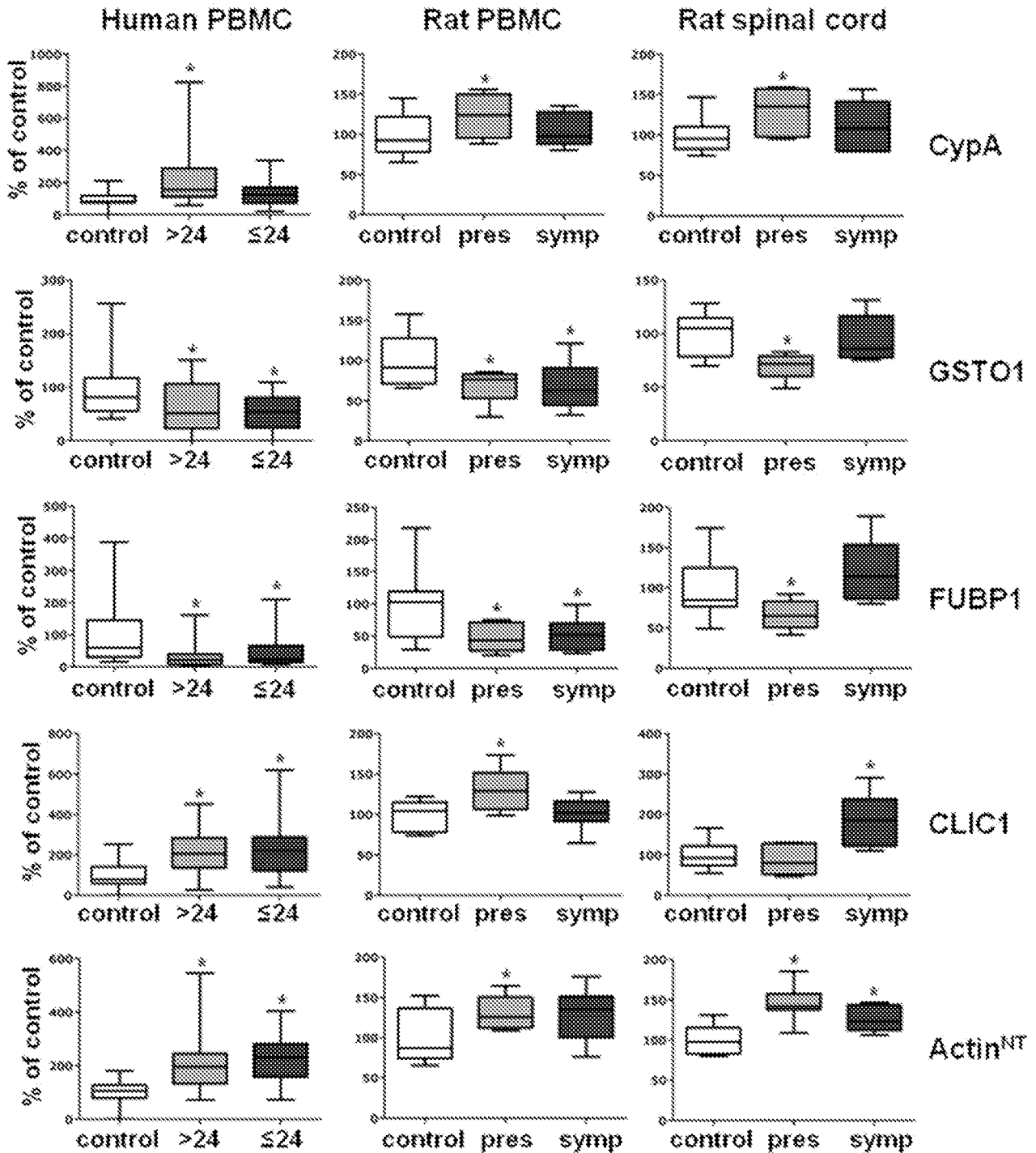


Fig. 4

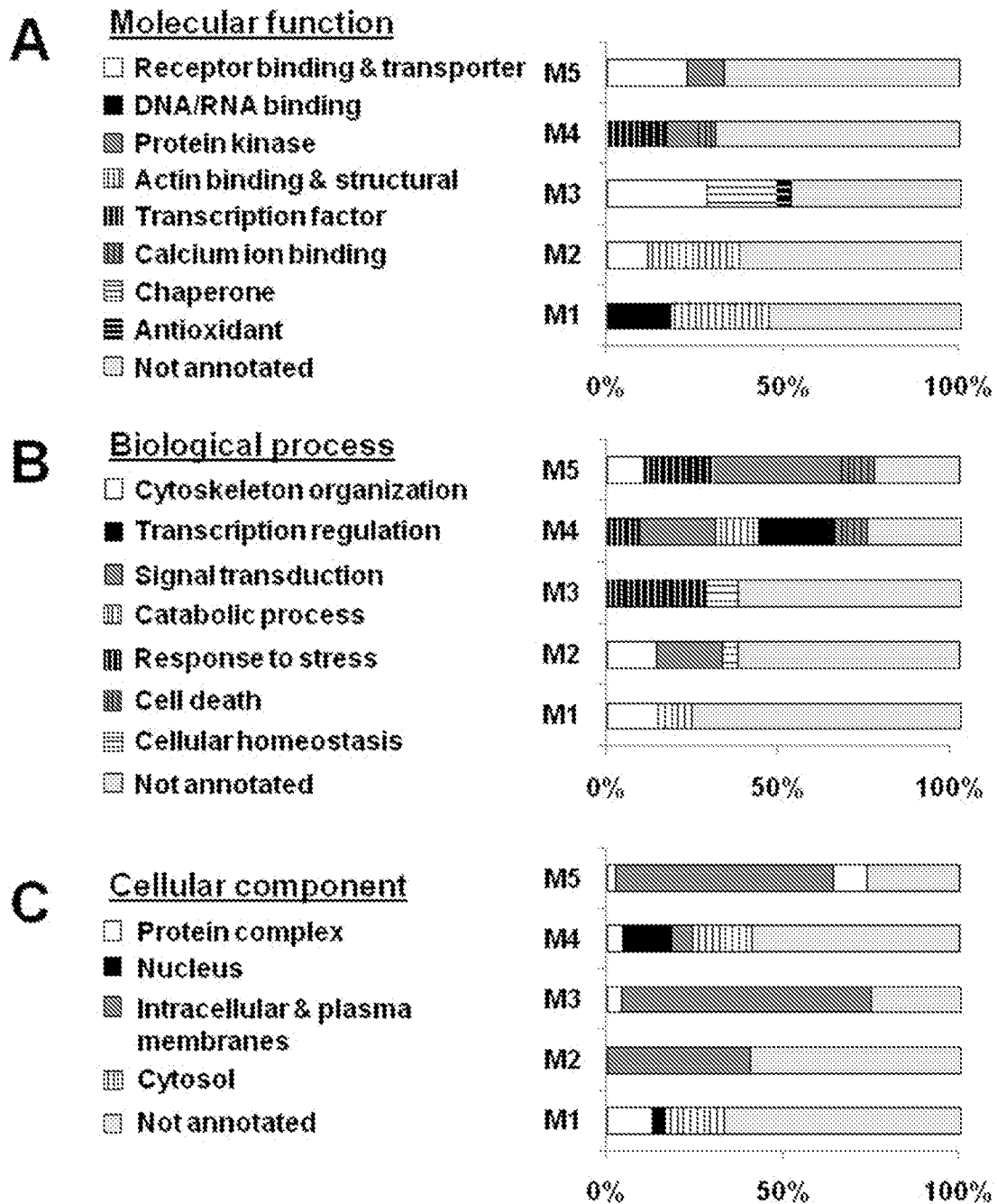


Fig. 5

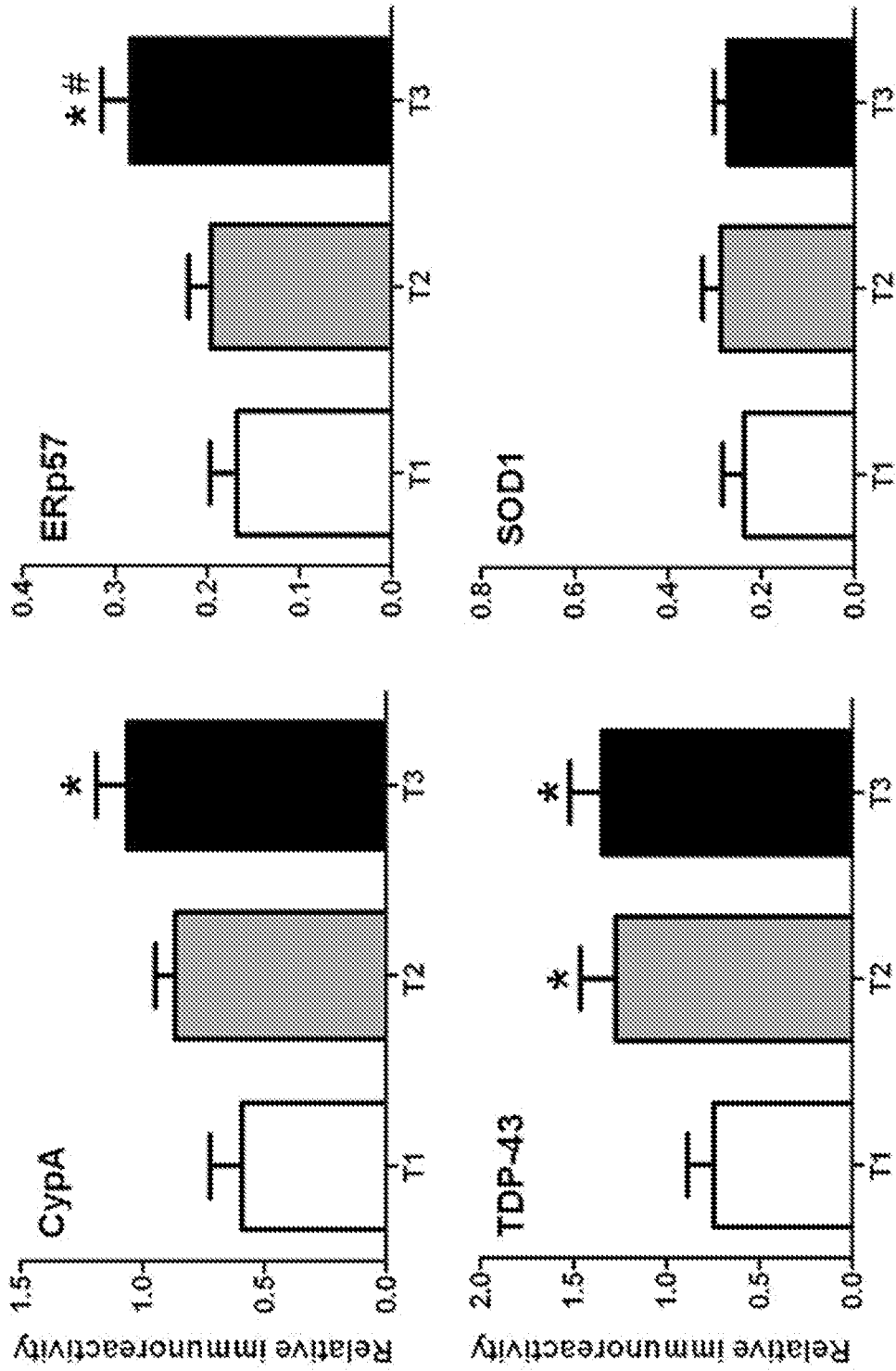


Fig. 6

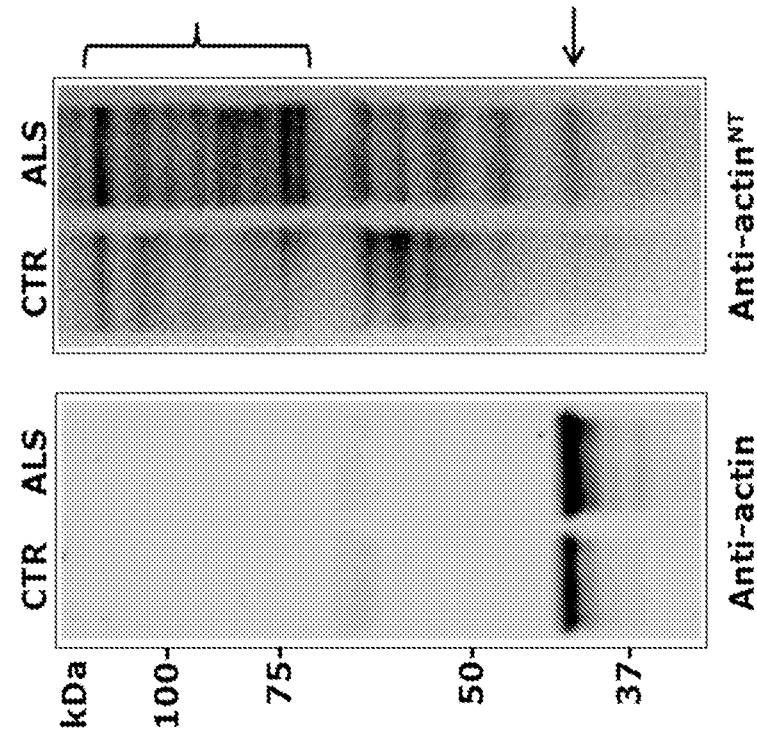


Fig. 7 B

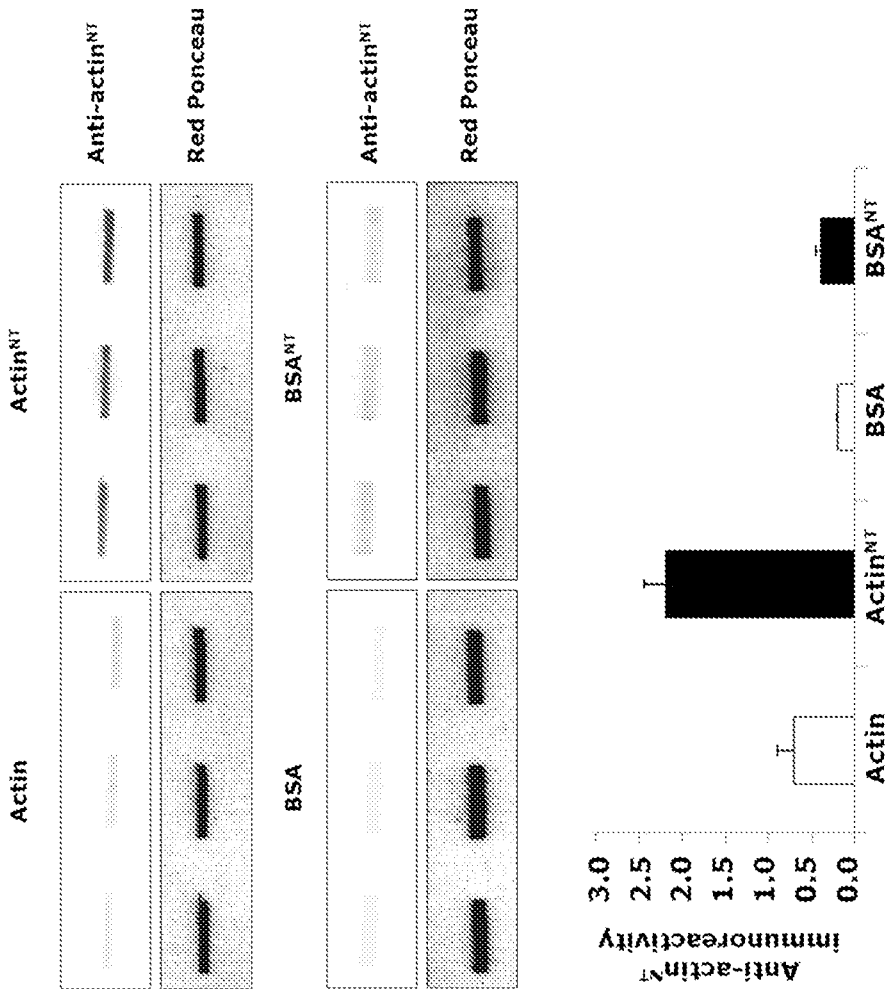
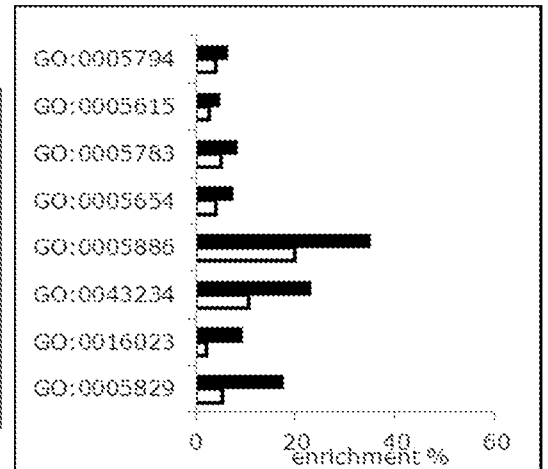
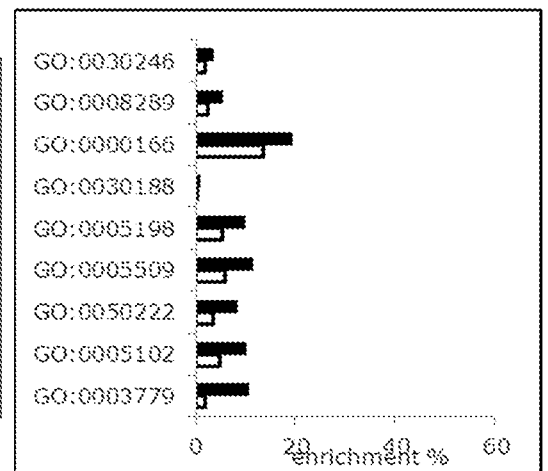


Fig. 7 A

GO:ID	Cellular Component (CC)	n	p-value
GO:0005794	Golgi apparatus	29	1.59E-02
GO:0005615	extracellular space	22	1.18E-02
GO:0005783	endoplasmic reticulum	38	2.01E-03
GO:0005654	nucleoplasm	34	3.87E-04
GO:0005886	plasma membrane	158	3.59E-14
GO:0043234	protein complex	104	8.16E-15
GO:0016023	cytoplasmic membrane-bounded vesicle	42	3.51E-15
GO:0005829	cytosol	90	1.41E-21



GO:ID	Molecular Function (MF)	N	p-value
GO:0030246	carbohydrate binding	17	6.63E-03
GO:0008289	lipid binding	25	1.94E-04
GO:0000166	nucleotide binding	92	1.49E-04
GO:0030188	chaperone regulator activity	4	4.92E-05
GO:0005198	structural molecule activity	47	2.51E-05
GO:0005509	calcium ion binding	54	2.06E-06
GO:0050222	protein kinase activity	39	1.81E-06
GO:0005102	receptor binding	48	8.30E-07
GO:0003779	actin binding	51	4.03E-24



GO:ID	Biological Process (BP)	N	p-value
GO:0007267	cell-cell signaling	27	9.67E-03
GO:0044404	symbiosis, encompassing mutualism through parasitism	4	8.14E-03
GO:0009653	anatomical structure morphogenesis	39	2.36E-03
GO:00016032	viral reproduction	8	3.67E-04
GO:0009056	catabolic process	44	5.53E-05
GO:0007610	behavior	27	8.89E-07
GO:0008219	cell death	37	5.11E-07
GO:0006464	protein modification process	75	1.76E-07
GO:00019725	cellular homeostasis	29	7.57E-08
GO:00050791	regulation of biological process	266	7.47E-08
GO:0007165	signal transduction	157	2.70E-09
GO:0009605	response to external stimulus	49	1.54E-09
GO:0006950	response to stress	85	4.83E-12
GO:0007010	cytoskeleton organization and biogenesis	52	1.37E-15

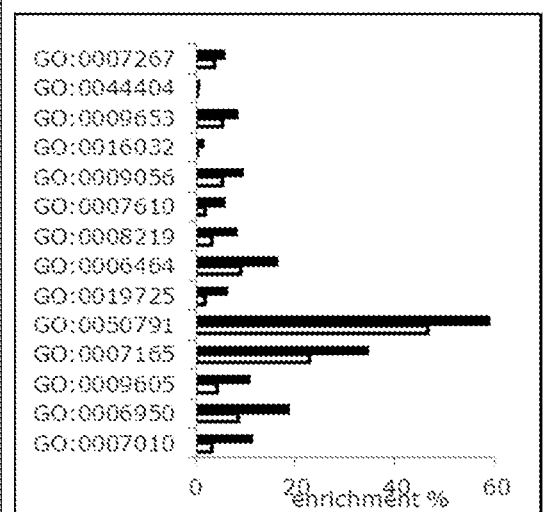


Fig. 8

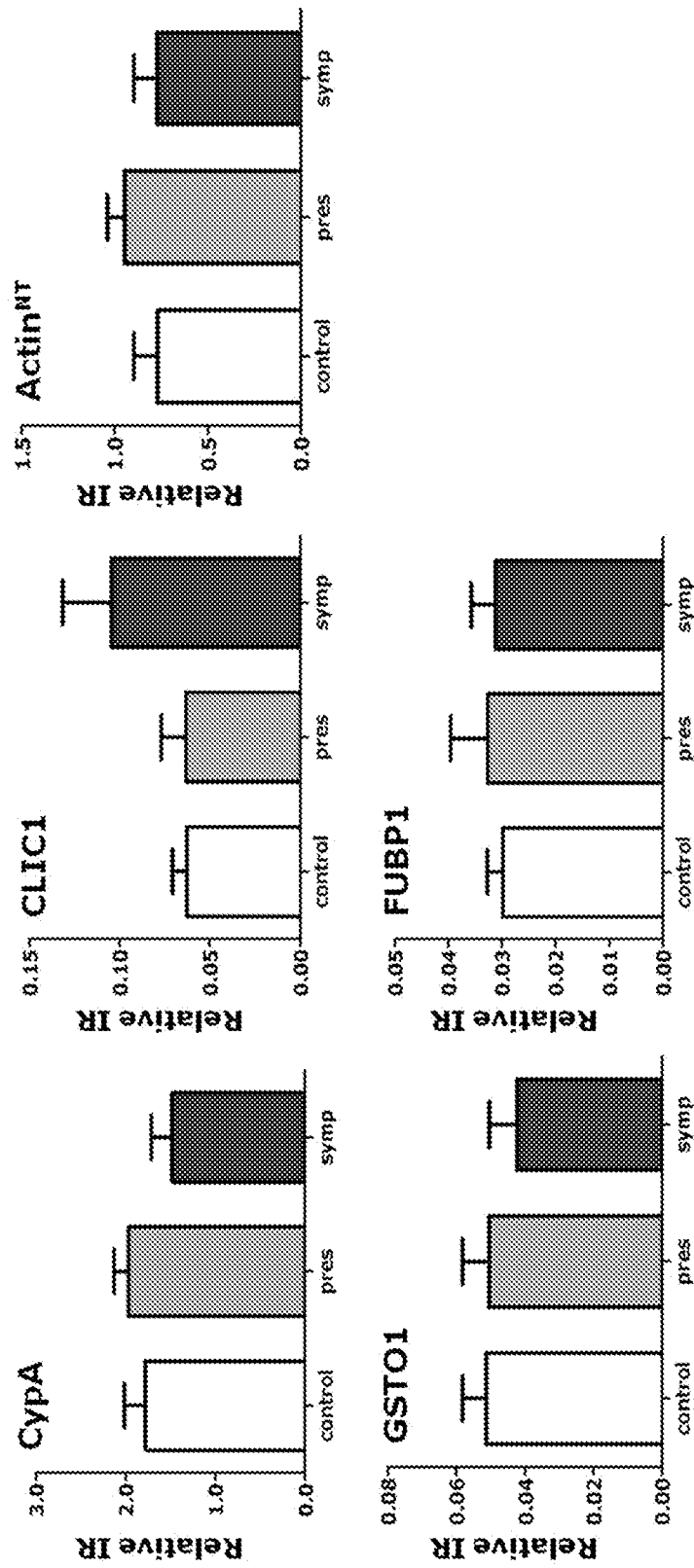


Fig. 9