의학박사 학위논문

Proteomic analysis of the alteration of nuclear cytoplasmic distribution of intracellular proteins in motoneuron cell lines expressing mutant SOD1 G93A

G93A SOD1 돌연변이 운동신경세포에서 핵-원형질간 세포단백의 분포 변화에 대한 proteome 분석

2015 년 08 월

서울대학교 대학원

의학과 뇌신경과학 전공

김 지 은

A thesis of the Degree of Doctor of Philosophy

G93A SOD1 돌연변이 운동신경세포에서 핵-원형질간 세포단백의 분포 변화에 대한 proteome 분석

Proteomic analysis of the alteration of nuclear cytoplasmic distribution of intracellular proteins in motoneuron cell lines expressing mutant SOD1 G93A

August 2015

The Department of Neuroscience

Seoul National University

College of Medicine

Jee Eun Kim

G93A SOD1 돌연변이 운동신경세포에서 핵-원형질간 세포단백의 분포 변화에 대한 proteome 분석

지도교수 이 광 우

이 논문을 의학박사 학위논문으로 제출함

2015 년 04 월

서울대학교 대학원

의학과 뇌신경과학 전공

김 지 은

김지은의 의학박사 학위논문을 인준함

2015 년 06 월

위 원 장 (인)

부위원장 (인)

위 원 (인)

위 원 (인)

위 원 (인)

Proteomic analysis of the alteration of nuclear cytoplasmic distribution of intracellular proteins in motoneuron cell lines expressing mutant SOD1 G93A

by

Jee Eun Kim

A thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Medicine (Neuroscience) at Seoul National University College of Medicine

June 2015

Approved by Thesis Committee:

Professor _____Chairman

Professor _____Vice chairman

Professor _____

Professor _____

Professor _____

ABSTRACT

Proteome analysis of the alteration of nuclear cytoplasmic distribution of intracellular proteins in motoneuron cell lines expressing mutant SOD1 G93A

Jee Eun Kim
Department of Neuroscience
College of Medicine
The Graduate School
Seoul National University

Background: The appropriate nuclear cytoplasmic localization of proteins is essential for eukaryotic cellular function and biological processes such as RNA processing, translation, protein-protein interaction and post-translational modification. Aberrant localization of proteins has been implicated in many human diseases, like cancer and neurodegenerative conditions. Evidence suggests that the cytoplasmic mislocalization of nuclear proteins including transactive response DNA-binding protein 43kDa (TDP-43) and fused in sarcoma (FUS) in amyotrophic lateral sclerosis and frontotemporal lobar

degeneration is associated with neurotoxicity. In this study, we investigated the proteome-wide alteration of nuclear cytoplasmic distribution in the motor neuron cell lines expressing mutant human SOD1 (G93A).

Methods: The motor neuron-like cell line NSC34 transfected with wild or mutant human SOD1 (G93A) underwent for subcellular fractionation. The proteome in the nuclear and cytoplasm fractions were analyzed using liquid chromatography-tandem mass spectrometry. Abundance difference of proteins for genotype, subcellular fraction and their interaction were analyzed. Bioinformatics analysis using TargetMine and DAVID were used to understand the function of the identified proteins. Finally, immunoblots of the significant proteins that changed the subcellular distribution in mutant cells were performed for validation.

Results: In total, 11,216 peptides and 1,925 proteins were identified, with 23% of the peptides and 32% of the proteins being found in both nuclear and cytoplasmic fractions. Using the intersection data set common to both fractions, we found that a considerable number of proteins (79%) were differentially distributed in the nuclear versus cytoplasmic compartments in the wild type. A total of 37 proteins showed a significant alteration in the nuclear cytoplasmic distribution in the mutant cells ($p<0.05$). The subcellular distribution in the mutant cells

was shifted from the cytoplasm to the nucleus for the proteins of RNA transport and processing (Dhx9, Fmr1, Srsf3, Srsf6, Tra2b), whereas the opposite was the case for the pathways of protein folding(Cct5, Cct7, Cct8), the aminoacyl-tRNA biosynthesis (Farsb, Nars, Txnrd1), the synaptic vesicle cycle (Cltc, Nsf), the Wnt signaling (Cltc, Plcb3, Plec, Psmd3, Ruvbl1) and the Hippo signaling (Camk2d, Plcb3, Ruvbl1).

Conclusions: This study allows a comprehensive understanding of the nuclear cytoplasmic distribution of intracellular proteins in motor neuron cell lines, and their alteration by the mutant SOD1 expression. Further studies are warranted, in order to elucidate the pathomechanistic implications of the aberrant localizations of the candidate proteins.

Keywords: amyotrophic lateral sclerosis, nuclear cytoplasmic mislocalization, proteome, mass spectrometry

Student number: 2011-31140

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATION

ALS: Amyotrophic lateral sclerosis

CID-HCD: Collision induced dissociation-high energy collision dissociation

FDR: False detection rate

FPKM: Fragments Per Kilobase of exon per Million fragments mapped

FUS: Fused in sarcoma

GO: Gene ontology

LC-MS/MS: Liquid chromatography tandem mass spectrometry

LTQ: linear trap quadrupole

MDS: Multidimensional scaling

MT: Mutant type

NC ratio: quantity ratio of nucleus/cytoplasmic proteins

NCE: Normalized collision energy

NGS: Next-generation transcriptome sequencing

NSC 34: Mouse motor neuron-neuroblastoma hybrid cell line

qPCR: Quantitative real time polymerase chain reaction

RNA: Ribonucleic acid

SOD1: Superoxide dismutatse 1

TCEP: Tris(2-carboxyethyl)phosphine

TDP-43: Transactive response DNA binding protein 43 kDa

TEAB: Triethylammonium bicarbonate

TMT: Tandem mass tag

tRNA: Transfer ribonucleic acid

WT: Wild type

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a catastrophic neurodegenerative disease that results from the continuing injury of motoneurons in the spinal cord, brainstem and motor cortex. Clinical pictures of ALS show a progressive spreading of motor weakness; the average survival rate of individuals with this type of condition is typically expected to be 2–3 years from onset (Boillée et al. 2006). Gaining an understanding of the pathomechanism of ALS has been the goal of research for decades, but little has been confirmed. Glutamate toxicity, apoptosis, neuroinflammation, and oxidative stresses have all been proposed as possible mechanism (Kanekura et al. 2009).

Eukaryotic cells are composed of membrane-covered compartments, like the nucleus, the cytoplasm, the mitochondria and the endoplasmic reticulum; each organelle has different protein components and resultant biological functions. The proper subcellular location of proteins is important for maintaining a variety of protein functions by influencing the protein synthesis, integrating the protein network, regulating the protein-protein interaction and post-translational modification. Recently protein mislocalization has been observed in many human diseases, such as cancer,

cardiovascular, metabolic and neurodegenerative diseases and it is regarded as the main pathogenic cause for these diseases (Hung and Wolfang, 2011).

Recently, the discovery of cytoplasmic inclusions in the motoneurons of ALS emphasized that protein aggregation and their ‘gain of toxicity’ may have the significant roles in the degeneration. These ubiquitin-positive inclusions turn out to be mostly composed of transactive response DNA binding protein 43 (TDP-43) or, less frequently, fused in sarcoma (FUS) (Dormann and Haass 2011; Neumann et al. 2006; Ito et al. 2011). Even though only 10% of ALS diagnosis are known to be familial in that it linked to a mutation in the superoxide dismutase 1 (SOD1), TDP-43, FUS and others, the protein aggregation with TDP-43 or FUS was also commonly seen in sporadic ALS cases (Dormann and Haass 2011). In normal cells, TDP-43 or FUS is predominantly located in the nucleus. However, TDP-43 or FUS, in motoneurons of the spinal cord and brain of ALS sufferers, has emerged in cytoplasmic inclusion. Interestingly, inclusion carrying cells merely show an aberrant aggregate of TDP-43 or FUS in the cytoplasm while but also showing a proportional reduction of TDP-43 or FUS in the nucleus (Arai et al. 2006; Vance et al. 2009). This proposes that a ‘loss of function’ in specific proteins in the nucleus could be another possible pathomechanism of ALS. Abnormal

cytoplasmic deposits in ALS might be the result of a defect in the nucleocytoplasmic shuttling machinery, the promotion of protein aggregation from aberrant RNA processing, stress granules and also from reduced protein degradation. The discovery of a reduced transportin or importin family in a postmortem ALS brain, and the reproducibility of the cytoplasmic insoluble aggregation by selective peptide inhibitors that inhibit the nuclear import pathway of TDP-43 or FUS supports one of those hypotheses (Cansizoglu et al. 2007; Kosugi et al. 2008).

Recently, high-throughput methods to quantify protein, as highly sensitive mass spectrometry (MS) have been applied in biological systems. This approach enables comprehensive understanding of the biological process in protein on the whole genome level, with less bias than previous analyses. It can suggest novel candidate proteins in specific diseases, including their deregulated function. Since ALS is now strongly regarded as a proteinopathy, this disease could be a good model to study with regard to this promising new quantitative technique. There have only been few reports mentioning the effect of the mutant SOD1 on protein mislocalization in the motor neuron cell; however, neither of these considered both the nuclear and cytoplasmic protein changes coincidentally, and neither used an

unbiased and comprehensive proteome analysis method in their approach.

In this study, we aimed to discover the mislocalized protein between the nucleus and the cytoplasm in the in vitro ALS model. For this purpose, we used hybrid mouse motoneuron/neuroblastoma cell lines (NSC34 cells), transfected with the mutant human SOD1 (G93A), and then compared them with the wild type in each subcellular fraction. Their fractionated cytosol and nucleus was analyzed using tandem mass tag (TMT)-labeling, linear trap quadrupole (LTQ)-Orbitrap Velos mass spectrometry, validated and looked for their biological function.

MATERIALS AND METHODS

Cell culture and subcellular fractionation

A mouse motoneuron/neuroblastoma hybrid cell line (NSC34) (Cellutions Biosystems, Ontario, Canada) was maintained in Dulbecco' s modified Eagle medium (JBI, Korea), with 10% heat inactivated fetal bovine serum (Gibco, Grand Island, NY, USA), and 1% penicillin – streptomycin (Gibco, Grand Island, NY, USA). The cells were kept in a humidified atmosphere of 5% CO_2 and 95% O_2, in an incubator, at 37° C. With a calcium phosphate method, the NSC-34 cells were transfected with the pCI-neo expression vector and the pCI-neo expression vector holding the wild type SOD1 or the mutant SOD1 (G93A), as previously described (Cashman et al. 1992). Geneticin was used for the clone selection. An immunoblot analysis confirmed the stable expression of the wild and mutant SOD1 (G93A).

For the subcellular fractionation, we used the commercially available NE-PER Nuclear and cytoplasmic extraction reagent (Thermo Scientific, USA). Briefly, the wild and mutant NSC34 cells from the subconfluent cultures were scrubbed and put into a phosphate-buffered saline (PBS) and separated by centrifugation. The supernatant was discarded,

and CER I was added to the pallet and suspended at 16,000 x g for 15 seconds. Next, CER II was mixed in, and the tube was vortexed for 10 seconds, and then centrifuged for 4 minutes at 16,000 x g The supernatant was transferred as a cytoplasmic extraction. The insoluble fraction that contained the nuclear component was suspended in ice-cold NER and vortexed for 15 seconds, every 10 minutes for a total duration of 40 minutes, until there was complete dissolution. After being centrifuged, the supernatant was prepared for nuclear fraction. The extraction was stored at -80° C, until use.

Materials for tandem mass tag

The following materials were used for TMT labeling: formic acid, urea, tris(2-carboxyethyl)phosphine (TCEP), iodoacetamide (IAA) (Sigma-Aldrich, St. Louis, MO, USA), a Labesix Plex reagents kit (Thermo Scientific, No. 90064), an HPLC-grade acetonitrile (Burdick and Jackson, (Muskegon, MI, USA). A Milli Q system (Millipore, Molsheim, France) was used for water purification.

Sample preparation for proteome analysis

100μg of the protein extract achieved from each cell fraction was dissolved in 45μL of 200mM triethylammonium bicarbonate

(TEAB) buffer (pH 8), containing 8M urea, and 500mM TCEP was mixed in, at room temperature, for 60 min. Then, the mixture was alkylated for 60 min, with 500mM IAA, in a darkened environment, at room temperature. The samples were desalted using a membrane filter of 10 KMW, and then were dissolved in 200mM TEAB (pH 8) buffer to make a final protein concentration of 1μg/μL. Each concentration of proteins was calculated using a bicinchoninic acid (BCA) assay (Thermo Scientific), as described in the manufacturer's protocol. The sequencing grade trypsin (Promega, Madison, WI, USA) and the proteins in the TEAB buffer were mixed in a 1:20 (wt/wt) ratio, and incubated overnight at 37° C (Byun et al. 2012). The nucleus and cytoplasm fractions from both the wild type SOD1 transfected and the mutant SOD1 (G93A) transfected NSC34 cells were individually labeled, using TMT-126, 128, 130 (the nucleus or cytoplasm of the wild type cells) and TMT-127, 129, and 131 (the nucleus or cytoplasm of the mutant type cells), following the manufacturer's protocol. An aqueous hydroxylamine solution (5% w/v) was blended to finish the reaction. Lastly, proteins from six samples were pooled, dried with a speed-vacuum, and melted in 0.1% formic acid, including 50μL of water, for the liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.

2D-LC-MS/MS

The 2D-LC-MS/MS system, made with a nanoACQUITY UltraPerformance LC System (Waters, USA) and an LTQ Orbitrap Elite mass spectrometer (Thermo Scientific, USA) with a nano-electrospray source were used to analyze the TMT-labeled samples (Washburn et al. 2001). A strong intensity cation exchange (5 μm, 3cm) column was located shortly before the C_{18} trap column (id 180 μm, length 20mm, particle size 5 μm; Waters). For each run, peptide solutions were laden in 5μL aliquots. Then, peptides were deranged by a salt gradient, introduced over an autosampler loop, from the strong intensity cation exchange phase into the C_{18} phase, and desalted at a flow rate of 4 μL/min for 10min. Next, the trapped peptides were detached on a 200mm homemade microcapillary column made up of C_{18} (Aqua; particle size 3 μm), filled into a 100 μm silica tube with a 5 μm orifice id.

A ten-step salt gradient was applied, using 3μL of 0, 25, 50, 100, 250, and 500mM ammonium acetate (0.1% formic acid in 5% acetonitrile), as well as 4, 5, 9 and a further 9μL of 500mM ammonium acetate (0.1% formic acid in 30% acetonitrile). The mobile phase A was composed of 0% acetonitrile and 0.1% formic acid, and B was composed of 100% acetonitrile and 0.1% formic acid. The LC gradient started with 5% B, for 1 min, and was raised to 20% B over 5 min, 50% B

over 90 min, 95% B over 1 min, and maintained at 95% B over 3 min and 5% B for a further 5 min (Hong et al. 2013; Kwon et al. 2014). Before the next run, the column was re-balanced, with 5% B, for 15 min. 2.0kV voltage was applied, in order to make an electrospray. The LTQ Orbitrap Elite was regulated with a data-dependent approach during the chromatographic separation. The MS data were obtained within the following parameters: five data-dependent collision induced dissociation-high energy collision dissociation (CID-HCD) dual MS/MS scans for each full scan, HCD scans and full scans were obtained in Orbitrap at resolutions of 60,000 and 15,000 by two-microscan averaging, CID scans were obtained in the LTQ by two-microscan averaging, 35% of normalized collision energy (NCE) in CID and 45% of NCE in HCD, ±1 Da isolation window. Fragmented ions were expelled for 60 sec. Each parent ion was initially fragmented using a CID and followed by an HCD in a CID-HCD dual scan (Byun et al. 2012).

Protein identification and quantification

An IPI mouse database (IPI.MOUSE. 7.26.2012) was used for the MS/MS spectra analysis, following the software analysis protocols. To calculate the false discovery rate (FDR), conversed sequences of all proteins were attached to the database. ProLucid was employed to classify the peptides with

the following parameter as a precursor: a mass error of 25 ppm, and a mass error distribution of fragment ion in 600 ppm (Rosen et al. 1993). Trypsin was chosen as the enzyme, with 2 potential missed cleavage sites. Lysine and N-terminus TMT modification and cysteine carbamidomethylation were selected as static modifications. Methionine oxidation was selected as a variable modification. To ensure better peptide identification and quantification, the tandem CID and HCD MS spectra from the identical precursor ions were co-analyzed using software (Li et al. 2012). Homemade software, which reporter ions from the HCD spectrum could introduce into the CID spectrum with identical precursor ions in the earlier scan, was used. Reporter ions were pulled from small windows (±20 ppm) near their anticipated m/z in the HCD spectrum. DTASelect (The Scripps Research Institute, USA) was used to filter and sort the output data in order to build the protein list. Two or more peptides were entered to identify the proteins, and 'less than 0.01' was set for a false positive rate (Tabb et al. 2002). To quantify the detected proteins, a Census was used in the IP2 pipeline (Integrated Proteomics, USA). The signal intensity of the reporter ion channel in each protein was estimated by calculating mean reporter ion's intensities from all ingredient peptides in the detected protein (Raso et al. 2012; Byun et al. 2012).

Transcriptome sequencing

To check the coverage of our proteome data and understand the detected proteome characteristics, we analyzed the corresponding transcriptome from 3 replicates of the nucleus and cytoplasm fractions in the wild and mutant NSC34 cell lines. The transcriptome NGS (next-generation transcriptome sequencing) was performed by Macrogen Inc. (Seoul, Korea), as described previously. Briefly, 1ug of the total RNA was analyzed using the TruSeq RNA library kit to construct the cDNA libraries. The protocol was made with polyA-selected RNA extraction, RNA fragmentation, random hexamer primed reverse transcription and 100nt paired-end sequencing, by an Illumina HiSeq2000 (Illumina, San Diego, CA, USA). The libraries were quantified using a quantitative real time polymerase chain reaction (qPCR), following the qPCR Quantification Protocol Guide. An Agilent Technologies 2100 Bioanalyzer was used for the qualification.

To discover alternative spliced transcripts and to measure expression levels, TopHat (which is able to report split-read alignments along the splice junctions) was used to organize the RNA-Seq reads into the genome of *homo sapiens*, which were determined by using the default selection on the Cufflinks software.

The reference genome sequence of *homo sapiens*, and the annotation data, were taken from the UCSC website (http://genome.uscs.edu). The transcript counts in the isoform and gene levels were measured, and the relative transcript abundances were measured in FPKM (Fragments Per Kilobase of exon per Million fragments mapped), using Cufflinks.

Biostatistics

All data analysis and visualization was conducted using R 3.0.0 (www.r-project.org). A proteome statistical analysis was performed as follows: The measured intensity ratios of the proteins were log2 transformed to achieve a normal distribution. The protein abundance between the two groups was analyzed by filtering, normalizing and averaging peptide intensities using the R-Rollup function of the DanteR package (Polpitiya et al. 2008). Only unique peptides that represented single proteins were selected for an evaluation of their abundance. Protein data with greater than 50% of the peptides missing from the mass spectrometry were excluded for analysis. For each of the detected proteins, we conducted a two-way ANOVA with a cellular component (nucleus, cytosol) and the genotype (mutant, wild) as factors, and compared pairs of the cellular components and genotypes, respectively, using the Tukey HSD post-hoc test. The identified proteins were regarded as being

significantly changed if it fit the cutoff value of $p<0.05$, and the fold change of abundance became 1.5 or greater.

Bioinformatics annotation

An Integrated pathway clusters analysis of the identified proteins, which exhibit significant alteration in their total amount or the nuclear cytoplasmic distribution in the mutant cells, was performed using TargetMine (http://targetmine.mizuguchilab.org/). TargetMine is an integrated data warehouse system, for target prioritization, to understand the biological meanings behind large lists of significant data. Furthermore, the identified proteins were also categorized by universal gene ontology (GO) terms, using the DAVID tool (http://david.abcc.ncifcrf.gov/). The adjusted p value, <0.05, was defined as the threshold.

Validation of proteome

The identified significantly changed proteins were validated using a Western blot. Three replicates of whole cells or the fractionated nucleus and cytoplasm in the NSC34 cell lines expressing the wild type and mutant human SOD1 (G93A) were lysed and preceded for the Western blot. The primary antibodies were: Hist1h1z (Proteintech, Chicago, IL); Atp5d,

Hist1h1e (Abcam, Cambridge, MA), Atp5b, Cand1, Cct5, Cct7, Cct8, Camk2d, Farsb, Hcfc1, Hist1h1b, Nars, pontin 52 (Santacruz Biotechnology, Santa Cruz, CA), Tardbp (Invitrogen, Carslbad, CA). The secondary antibodies were goat antimouse IgG (Santacruz Biotechnology, Santa Cruz, CA).

RESULTS

Cell fractionation

To find out if there were significant protein changes in the ALS pathology, we planned to assay for the abundance of proteins in the overall NSC34 cell lines expressing the wild and mutant human SOD1 (G93A), and also for the distribution between the nucleus and the cytoplasm. For this purpose we fractionated the nucleus and cytoplasm component, as described earlier. A Western blot analysis of subcellular preparation was checked to confirm the qualification of fractionation (Figure 1). α-tubulin, which is a marker protein of cytoplasm, was exclusively seen in the cytoplasmic fraction. The nuclear marker Lamin B was predominantly shown in the nuclear fraction, while it was absent in the cytoplasmic fraction. The quality of fractionation was reinforced by the findings from the principle component analysis of the quantified proteins from the wild and mutant cells and also from the results of the z-score representing the relative abundance of cytoplasmic and nuclear marker proteins in 12 fractionated samples (Supplemental figure S1).

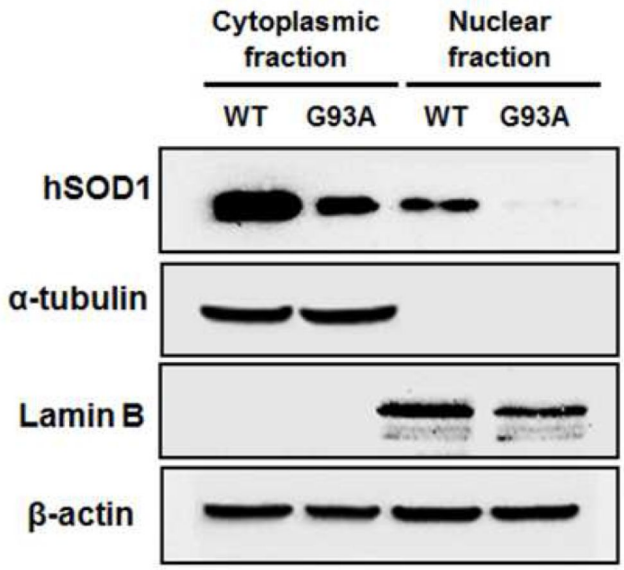


Figure 1. Western blots of marker proteins from the nuclear and cytoplasmic fractions of the NSC34 cells, expressing the wild-type (WT) and mutant human SOD1 (G93A).
α-tubulin, a cytoplasmic protein marker, was exclusively expressed in the cytoplasmic fraction; lamin B, a nuclear protein marker, was only expressed in the nuclear fraction. Finally, the quality of the subcellular fractionation was confirmed.

Protein identification and quantification

The subcellular fraction and also the total cellular extracts of the NSC34 stable cell lines expressing the human wild-type SOD1 or the mutant SOD1 (G93A) are predisposed to an LC-MS/MS analysis after the TMT tagging as previously described.

From the LC-MS/MS analysis, the MS spectra numbers detected in the cytoplasm were 31893 and 24381 in the nucleus.

The resultant MS spectra were applied in ProLucid for the peptide identification; 7543 peptides (93.8% unique peptides) in the cytoplasm, and 6341 peptides (92.7% unique peptides) in the nucleus were identified. All peptides underwent quantile normalization, and only unique peptides were selected for R-Rollup and further protein quantification. Corresponding proteins accounted for 1359 for the cytoplasm and 1200 for the nucleus. The distribution of each peptide or protein, and their overlap, was presented in Figures 2A and 2B. Peptides that were only detected in one side of the fraction were regarded as having very low abundance in that location or being undetected due to the technical limitations of mass spectrometry. Accordingly, the proteins that were detected in only one fraction were discarded, so as not to include any unnecessary bias from the imputation methods for the missing value. Finally, 634 nucleo-cytoplasmic common proteins (2403 peptides), that shared for the both cytoplasm and the nucleus underwent further evaluation.

First we investigated whether each protein level from the total cellular extraction showed any difference between the wild and mutant cells. Then, the protein quantity ratio of subcellular component, [nucleus/cytoplasm (NC) protein ratio], was compared between the wild and mutant cells to reflect the change in the subcellular distribution, by the SOD1 (G93A)

mutation. We discovered 42 up-regulated and 29 down-regulated proteins in the mutant cell compare to the wild type from the whole cell lysis (p value <0.05) (Tables 1 and 2). The fold change of the detected proteins and the corresponding significance were presented as a volcano plot. If we consider both the raw p-value 0.05 and the 1.5 fold change to be a cutoff threshold, then none of these proteins were significant altered over the threshold between the wild and the mutant cells (Figures 3A and 3C). The nucleus fraction showed 212 increased and 291 decreased proteins compared to the cytoplasmic fraction (p value <0.05) (Figure 3B, 3C). By applying the same thresholds as in the above volcano plot, the 136 down-regulated protein and the 103 up-regulated proteins were significant. A two-way statistical analysis revealed 37 proteins with significant interaction, due to the genotype (mutant versus wild type) and the subcellular fraction (nucleus versus cytoplasm). Among the proteins that changed their distribution in the presence of the SOD1 (G93A) mutation , 24 proteins changed to being predominantly located in the cytoplasm and 13 proteins changed to being predominantly located in the nucleus (p value <0.05) (Table 3). A hierarchical clustering analysis was performed to display the expression patterns of the differentially expressed proteins in the cellular component and genotype pairs (Figure 4A). As

seen in the volcano plot in Figure 4, none of these proteins showed significant increments more than 1.5 fold in the cytoplasm fraction of the mutant cell. However, 5 proteins were identified to be elevated more than 1.5 fold in the nuclear fraction of the mutant cell.

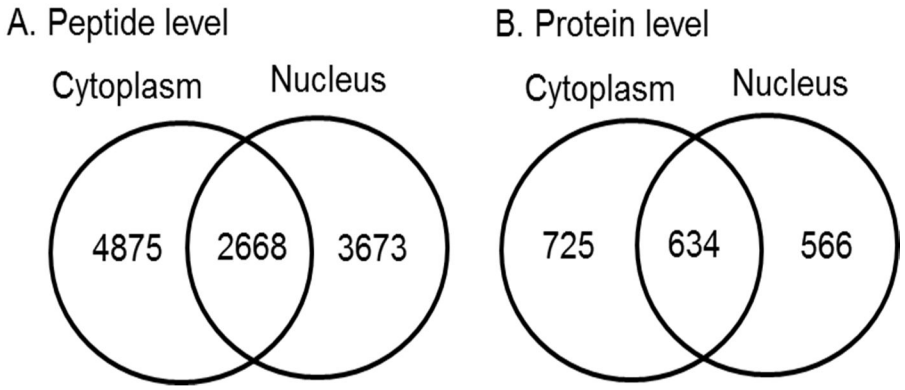


Figure 2. Venn diagrams of the total quantified peptides and proteins from the cytoplasmic and nuclear fractions of the wild type cells.
The TMT-label Orbitrap LC-MS/MS identified and quantified 1359 (7543) and 1200 (6341) proteins (peptides) from the fraction of the cytoplasm and nucleus, respectively, in the NSC34 stable cell lines expressing the human wild-type or mutant SOD1 (G93A). Only the peptides detected in both fractions were used for the relative quantification of protein abundance.

Table 1. Upregulated proteins in the NSC34 cells expressing the mutant human SOD1 (G93A)

IPI	Gene symbol	Log[2] (MT/WT)	IPI	Gene symbol	Log[2] (MT/WT)
IPI00758024.1	Prdx6	0.49	IPI00321734.7	Glo1	0.19
IPI00649135.3	Gstm1	0.49	IPI00134353.3	Nol3	0.19
IPI00115650.4	Cacybp	0.49	IPI00649406.1	Park7	0.19
IPI00331704.7	Eno2	0.43	IPI00227392.5	Ywhah	0.19
IPI00222759.3	Vat1l	0.39	IPI00230707.6	Ywhag	0.19
IPI00317309.5	Anxa5	0.35	IPI00626994.3	Ipo5	0.19
IPI00417165.3	Enah	0.34	IPI00230429.4	Kpna3	0.18
IPI00121427.1	S100a6	0.31	IPI00116498.1	Ywhaz	0.18
IPI00885558.1	Pdia3	0.29	IPI00132575.3	Cotl1	0.18
IPI00990246.1	Nme1	0.28	IPI00331556.5	Hspa4	0.17
IPI00411075.2	Pcbp3	0.27	IPI00760000.1	Ywhab	0.17
IPI00990529.1	Gstp1	0.27	IPI00660514.1	Dnajb6	0.17
IPI00461281.2	Nudcd2	0.25	IPI00131224.1	Tceb2	0.17
IPI00757109.3	Pcmt1	0.24	IPI00123342.4	Hyou1	0.16
IPI00762774.2	Eif3d	0.23	IPI00798527.1	Tnpo1	0.16
IPI00269662.1	Hnrnpa3	0.23	IPI00776252.1	Txnrd1	0.14
IPI00153728.1	Ddx19b	0.22	IPI00314153.4	Yars	0.14
IPI00339916.10	Eprs	0.21	IPI00111181.1	Vps35	0.14
IPI00116254.1	Prdx4	0.21	IPI00323357.3	Hspa8	0.11
IPI00759940.3	Fh1	0.2	IPI00116308.1	St13	0.1
IPI00122743.2	Dars	0.2	IPI00119057.1	Eif4e	0.06

Two-way ANOVA, cell-type effect, p-value < 0.05

Table 2. Downregulated proteins in the NSC34 cells expressing the mutant human SOD1 (G93A)

IPI	Gene symbol	Log[2](MT/WT)	IPI	Gene symbol	Log[2](MT/WT)
IPI00130589.8	Sod1	−0.57	IPI00123624.8	2610301G19Rik	−0.24
IPI00474974.1	Dnmt1	−0.54	IPI00230133.5	Hist1h1b	−0.23
IPI00169870.6	Glt25d1	−0.4	IPI00223371.3	Rbm39	−0.23
IPI00228616.5	Hist1h1a	−0.39	IPI00337844.5	Ranbp2	−0.22
IPI00132352.2	2610029G23Rik	−0.39	IPI00515398.1	Myh10	−0.22
IPI00109813.1	Hnrnpa0	−0.39	IPI00318725.4	Rrs1	−0.22
IPI00223714.5	Hist1h1e	−0.36	IPI00330289.4	Epb4.1l2	−0.21
IPI00113141.1	Cs	−0.36	IPI00754963.2	Mest	−0.19
IPI00229535.2	Gtf2i	−0.35	IPI00312128.3	Trim28	−0.17
IPI00331361.2	Mybbp1a	−0.34	IPI00828543.3	Hcfc1	−0.17
IPI00331597.6	Hist1h1d	−0.33	IPI00133985.1	Ruvbl1	−0.15
IPI00673465.2	Cnot1	−0.31	IPI00281011.7	Marcksl1	−0.15
IPI00154054.1	Acat1	−0.28	IPI00622811.2	Ap2m1	−0.15
IPI00515654.2	Eef1d	−0.27	IPI00881287.1	Fkbp8	−0.12
IPI00226882.7	Sec61a1	−0.25			

Two−way ANOVA, cell−type effect, p−value < 0.05

Table 3. Proteins with their nuclear−cytoplasmic distribution being significantly altered in the NSC34 cells expressing the mutant human SOD1 (G93A)

IPI	Gene symbol	$Log_2(N/C)_{WT}$	$Log_2(N/C)_{MT}$	IPI	Gene symbol	$Log_2(N/C)_{WT}$	$Log_2(N/C)_{MT}$
IPI00407130.4	Pkm2	2.17	−0.63	IPI00469268.5	Cct8	−0.32	−0.69
IPI00133985.1	Ruvbl1	2.14	−0.28	IPI00114375.2	Dpysl2	−0.35	−0.49
IPI00988949.1	Erh	1.86	0.49	IPI00896727.1	Cand1	−0.4	−0.56
IPI00311203.2	Plcb3	1.44	−0.57	IPI00918997.1	Nars	−0.42	−0.5
IPI00330289.4	Epb4.1l2	1.41	0.7	IPI00322828.2	Farsb	−0.45	−0.95
IPI00280967.3	Tardbp	1.39	−0.27	IPI00116254.1	Prdx4	−0.65	0.25
IPI00970572.1	Tra2b	1.25	0.56	IPI00310880.4	Srsf6	−0.68	0.75
IPI00116279.3	Cct5	0.9	−0.43	IPI00227392.5	Ywhah	−0.73	0.13
IPI00318841.4	Eef1g	0.89	−0.49	IPI00230707.6	Ywhag	−0.73	0.13
IPI00406790.9	Camk2d	0.62	−0.75	IPI00828543.3	Hcfc1	−0.84	−0.27
IPI00130589.8	Sod1	0.46	1.81	IPI00116498.1	Ywhaz	−0.87	0.17
IPI00387337.1	Bzw2	0.31	−0.47	IPI00227013.2	Fmr1	−0.87	−0.23
IPI00468481.2	Atp5b	0.26	1.24	IPI00648173.1	Cltc	−0.89	−0.57
IPI00322869.3	Abce1	0.14	−0.55	IPI00314439.4	Psmd3	−0.93	−0.32
IPI00331174.5	Cct7	0.1	−0.6	IPI00462453.5	Gm5619	−1.07	−0.56
IPI00776252.1	Txnrd1	0.05	−0.48	IPI00453777.2	Atp5d	−1.36	1.47
IPI00230061.3	Plec	−0.08	−0.33	IPI00339468.4	Dhx9	−1.72	0.72
IPI00454008.1	Shmt2	−0.17	−0.41	IPI00221826.1	Srsf3	−1.85	1.68
IPI00169463.1	Tubb2c	−0.29	−0.51				

Two−way ANOVA, cell type × fraction interaction effect, p−value < 0.05

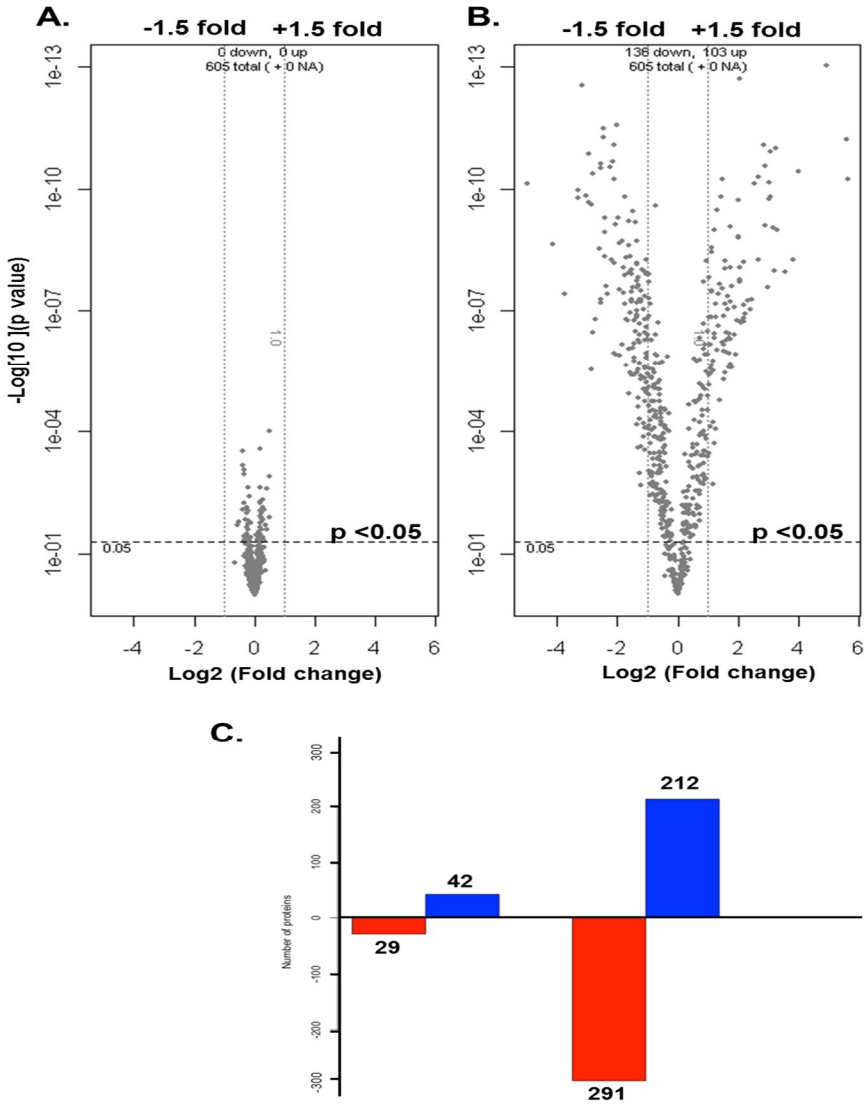


Figure 3. Volcano plots of the protein from the wild and mutant cells quantified with TMT-labeling, LTQ-Orbitrap Velos Mass Spectrometry.

The negative $\log_{10}$-transformed raw p-values of the two-way ANOVA test are plotted against the estimated log ratios ($\log_2$ fold change) of the protein abundance in the mutant versus wild-type cells (A), and the nuclear versus cytoplasmic fraction in the wild type cells (B). The numbers of the proteins that are significantly down-/up-regulated in the mutant cells and the nuclear fraction compared to the wild-type and cytoplasmic fraction, respectively, are shown in bar graphs (C).

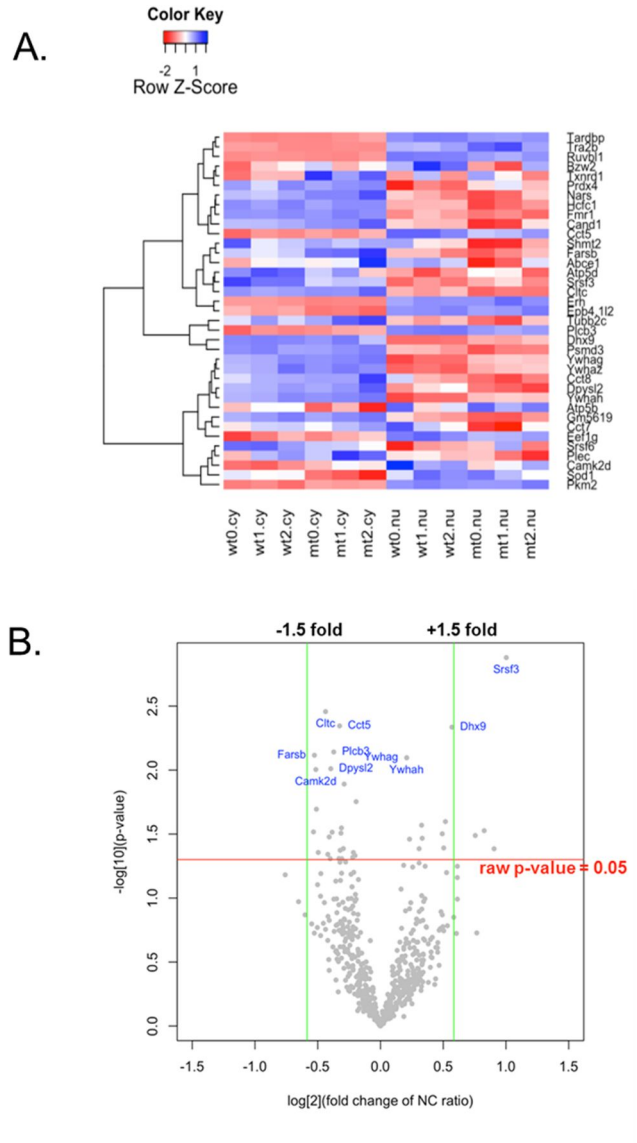


Figure 4. The alteration of the proteome nuclear–cytoplasmic distribution in the mutant cells.

(A) The heat map presentation of a hierarchical cluster of the 37 proteins that show significant alteration ($p<0.05$) in the mutant cells (versus wild-type) of the relative abundance between two subcellular fractions is shown. The red color represents "low abundance" and blue color represents "high abundance" levels. (B) The volcano plot shows the estimated $\log_2$ fold change of the nuclear–cytoplasmic (NC) ratio in the mutant versus wild-type cells in x-axis, and negative $\log_{10}$-transformed raw p-values from the cell × fraction interaction effect in the 2-way ANOVA test in the y-axis.

Functional annotation enrichment

To identify the disrupted biological function in the mutant SOD1 motoneuron from our quantified proteins, we performed an integrated pathway cluster analysis using TargetMine. The resultant integrated pathway cluster, enriched for all quantified proteins, was related to the RNA transport, the metabolism of proteins, Wnt signaling, protein processing in the endoplasmic reticulum, the cell cycle, Huntington' s disease/Alzheimer' s disease, carbon metabolism, the ribosome, aminoacyl-tRNA biosynthesis, the transcription, glycine, serine and threonine metabolism, the RNA degradation, organelle biogenesis and maintenance, glycogen storage disease, and amino acids and the derivative metabolism (Figure 5). In particular, the 37 proteins that showed significant alteration of the nuclear-cytoplasmic distribution in the mutant cells were enriched for protein folding, aminoacyl-tRNA biosynthesis, RNA transport, Wnt signaling, Huntington' s disease/Alzheimer' s disease, Synaptic vesicle cycle and Hippo signaling pathway (Table 4). Integrated pathway cluster proteins, exhibiting significant alteration of nuclear-cytoplasmic distribution in the mutant cells were plotted in the graph showing the relative protein fold changes in the mutant cells in each nuclear and cytoplasm fraction. Proteins that shift from the nucleus to the cytoplasm in the mutant cells were related to the pathway of protein folding,

aminoacyl-tRNA biosynthesis, Wnt signaling, synaptic vesicle cycle and Hippo signaling. However, the proteins enriched for RNA transport/processing were reversely moved from the cytoplasm to the nucleus in the mutant cells. For proteins corresponding to Huntington' s disease/Alzheimer' s disease, the translocation direction was not consistent in our data.

The detected proteomes were also categorized in GO terms (Table 5). Proteins with a decreased NC ratio (the quantity ratio of nucleus/cytoplasmic proteins) in the mutant cell, which means the cytoplasmic mislocalization of that protein is suspected, were functionally related to the protein folding, cell cycle and regulation of the protein translation and transcription. Alternatively, the proteins with increased NC ratios in the mutant cells are associated with ATP synthesis and phosphorylation. Significantly down-regulated proteins from the total extract of mutant cells are found to participate in the nucleosome assembly and the phosphate metabolic process.

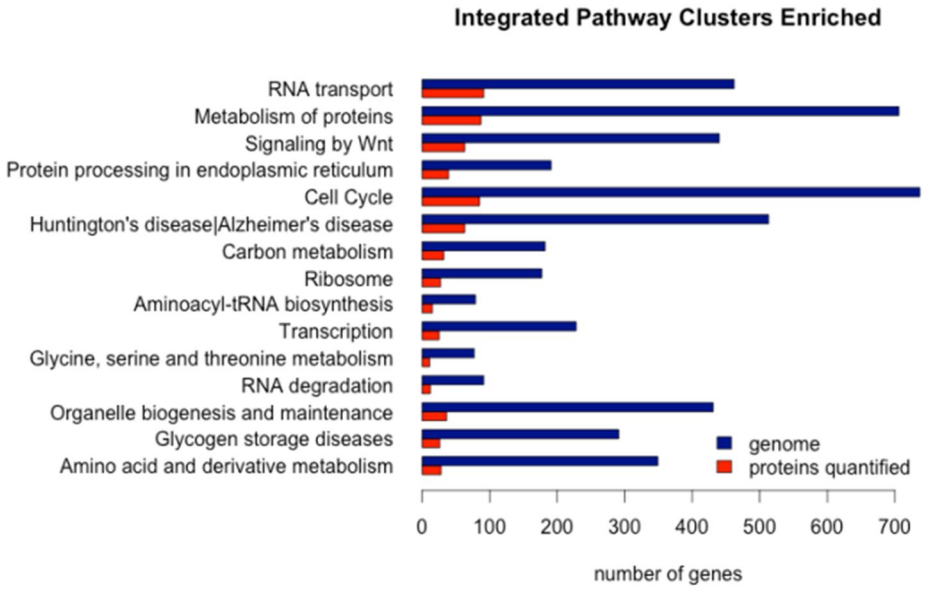


Figure 5. The integrated pathway clusters enriched in quantified proteins from the wild and mutant cells.
The identified proteins underwent integrated pathway cluster analysis using TargetMine to evaluate their biological function. The integrated pathway clusters enriched for all detected proteins were shown. Adjusted p-value <0.05

Table 4. The result of the pathway enrichment analysis

Integrated pathway clusters	Genes	P-value
Protein folding	Cct5, Cct7, Cct8	0.00055
Aminoacyl-tRNA biosynthesis	Farsb, Nars, Txnrd1	0.0018
RNA transport/ Processing of Capped Intron-Containing Pre-mRNA	Dhx9, Fmr1, Srsf3, Srsf6, Tra2b	0.0068
Signaling by Wnt	Cltc, Plcb3, Plec, Psmd3, Ruvbl1	0.011
Huntington's disease/ Alzheimer's disease	Atp5b, Atp5d, Cltc, Plcb3, Sod1	0.018
Synaptic vesicle cycle	Cltc, Nsf	0.018
Hippo signaling pathway	Camk2d, Plcb3, Ruvbl1	0.049

The integrated pathway clusters and corresponding genes, encoding the proteins and exhibiting significant alterations of the nuclear-cytoplasmic distribution in the mutant cells are shown.

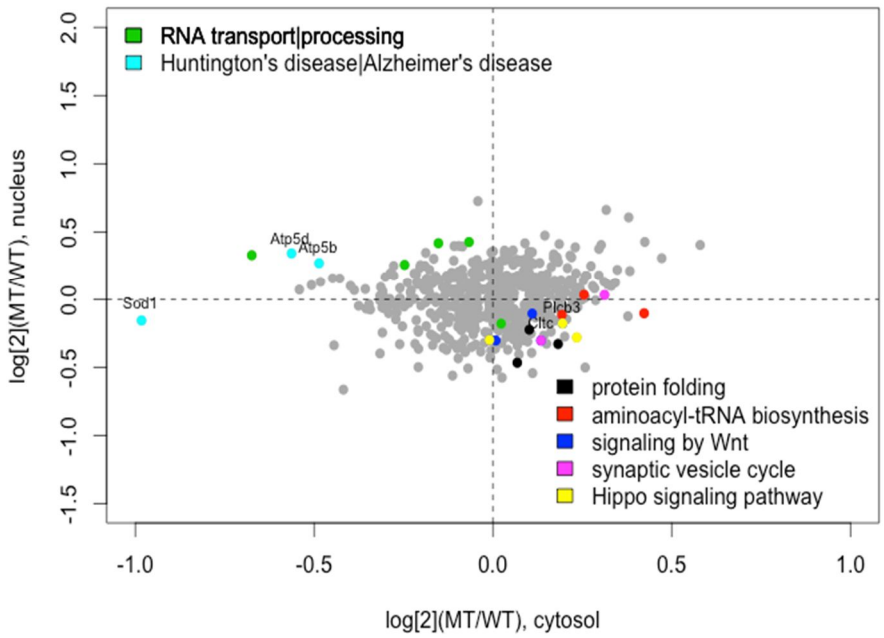


Figure 6. A scatter plot showing the log2 fold change between the mutant and wild-type cells of protein abundance in the cytoplasmic (x-axis) and nuclear (y-axis) fractions.

The color-coded dots are the integrated pathway cluster proteins exhibiting significant alterations of nuclear-cytoplasmic distribution in the mutant cells. The subcellular distributions of proteins in the mutant cells are shifted from the nucleus to the cytoplasm for the pathways of protein folding, aminoacyl-tRNA biosynthesis, Wnt signaling, synaptic vesicle cycle and Hippo signaling (right lower quadrant), and the changes are the reverse for RNA transport/processing. For proteins corresponding to Huntington's disease/Alzheimer's disease (the gene symbols annotated in the plot), there was no consistent pattern.

Table 5. The result of the functional annotation enrichment of identified proteins from the mass spectrometry

GO	Gene	P.value
Decrease of NC ratio in mutant cells		
protein folding	Cct5, Cct7, Cct8	1.20E−02
tRNA aminoacylation for protein translation	Nars, Farsb	5.90E−02
cell cycle	Ruvbl1, Camk2d, Hcfc1	1.90E−01
regulation of transcription	Ruvbl1, Cand1, Tardbp	8.20E−01
Increase of NC ratio in mutant cells		
ATP biosynthetic process	Atp5b, Atp5d	4.10E−02
phosphorylation	Atp5b, Atp5d, Sod1	4.90E−02
Downregulation in mutant cells		
nucleosome assembly	Hist1h1a, Hist1h1b, Hist1h1e	1.00E−03
phosphate metabolic process	Atp5b, Sod1, Timm50	1.10E−01
Upregulation in mutant cells		
none		

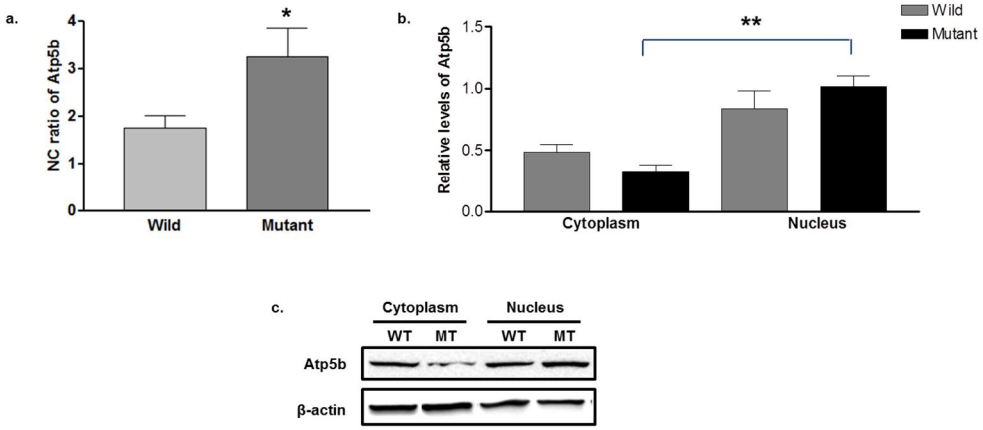
Abbreviation: GO, gene ontology; RNA, ribonucleic acid; ATP, adenosine triphosphate; NC ratio, quantity ratio of nucleus/cytoplasmic protein

Validation of proteomic data

Lastly, we preceded with the validation of the detected proteins from the mass spectrometry data. For this purpose, we prepared whole and fractionated lysates from three wild and mutant human SOD1 (G93A) transfected NSC34 cells, and conducted a Western blot. The evaluated proteins were followings: Cct5, Cct7, Cct8, Nars, Farsb, Ruvbl1, Camk2d, Hcfc1, Ruvbl1, Cand1, Tardbp, Atp5b, Atp5d, sod1, Hist1h1a, Hist1h1b and Hist1h1e. Among these candidate pathologic proteins, Atp5b (which related to ATP biosynthetic process) and Cct8 (that associated with protein folding) were confirmed to significantly change their nucleocytoplasmic distribution in the mutant cells (Figures 7A and 7B). Cct5's expression only showed a mild tendency to redistribute in the cytoplasm in the mutant cells. Using a two-way ANOVA, a significant interaction ($F_{int.}=10.11$, $p<0.05$) of the genotype (wild;mutant – $F=0.05$) and the cellular fraction (nucleus;cytoplasm – $F=139.46$) was only observed for Cct8. Furthermore, Hist1h1a, Hist1h1b and Hist1h1e, which were related to the nucleus assembly, were confirmed to be downregulated in the mutant cells by the Western blot (Figure 7C).

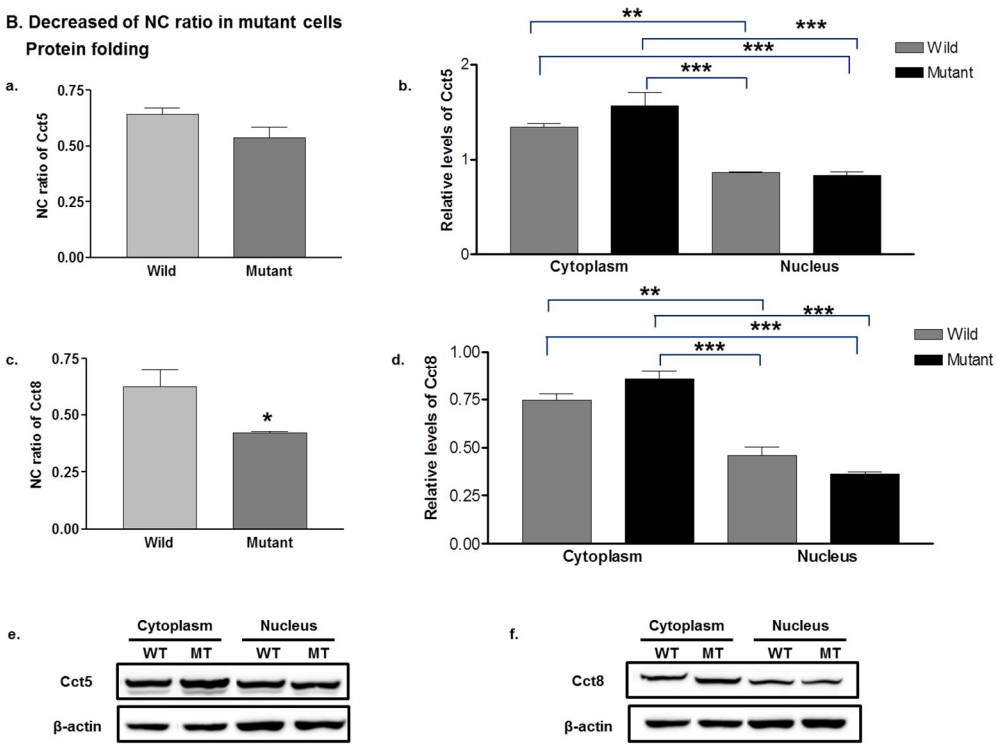
A. Increase of NC ratio in mutant cells

ATP biosynthetic process



B. Decreased of NC ratio in mutant cells

Protein folding



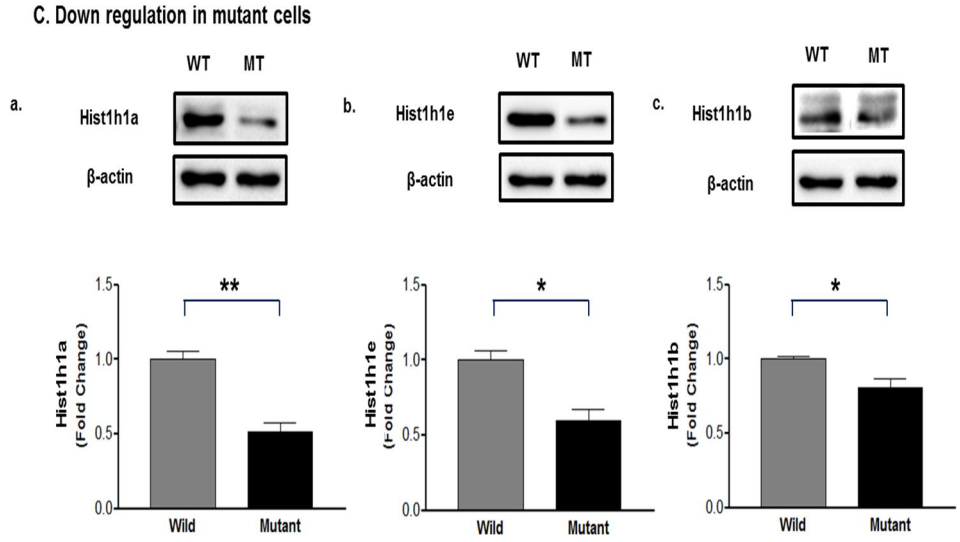


Figure 7. Validation of proteome data.

Proteins that were detected by mass spectrometry underwent a Western blot. Atp5b (which was related to the ATP biosynthetic process), Cct5, and Cct8 (that was associated with protein folding) were confirmed to significantly change their nucleocytoplasmic distribution in the mutant cells, by a Western blot. (A-c, B-e, f) Western blot results of Atp5b, Cct5 and Cct8, and (A-b, B-b, d) their quantification data for the expression level and (A-a, B-a,c)NC ratio are shown. A two way ANOVA only found a significant interaction ($F_{int.}=10.11$, $p<0.05$) between the genotype (wild; mutant – $F=0.05$) and the cellular fraction (nucleus; cytoplasm – $F=139.46$) for Cct8. The expression level of Atp5b, Cct5, and Cct8 were calculated as a relative ratio to β-actin. (C, a-c) A Western blot and the quantification data of Histone H1.1, Histone H1.4 and Histone H1.5, in the whole cellular extract. Histone H1.1, Histone H1.4 and Histone H1.5 were significantly decreased in the mutant cells. The data were reported as mean ± SE. The statistics were created via t-test, for the NC ratio and the two-way ANOVA and Tukey test, for each expression level comparison (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, $n=3$ per group)

DISCUSSION

SOD1 mutations are the second most common causative genes for familial ALS (Rosen et al. 1993). In order to identify proteome-wide alteration by the presence of the mutant SOD1, we used mutant human SOD1 (G93A) transfected NSC34 cells, which contain motoneuron characteristics. These cells lines were a good culture model for ALS, in order to investigate the independent effect of mutant SOD1 toxicity on motoneuron. In particular we focused that cytoplasmic aberrant aggregation, as TDP-43 and FUS, as hallmark pathological findings for both familial and sporadic ALS. We explored possibly mislocalized proteins, in the relation of the nucleus and the cytoplasm, in the presence of the mutant SOD1 at the proteome level. For this purpose, we applied the latest high-throughput methods (TMT-label Orbitrap LC-MS/MS) to quantify proteomes in each subcellular location.

Using a proteome analysis, we quantified 1359 and 1200 proteins from the fraction of the cytoplasm and nucleus, respectively, in NSC34 stable cell lines expressing the wild-type or mutant human SOD1 (G93A). The analysis of the 634 proteins identified in both the cytoplasmic and nuclear fractions revealed 503 proteins with significantly different nucleo-

cytoplasmic abundance, 71 proteins with a significant cell−line effect, and 37 proteins with a significant cell−line* fraction interaction. We found that a considerable proportion of the peptides and proteins (23% of the total peptides, 32% of the total proteins in the cell) in NSC34 cells were shared between the nucleus and the cytoplasm. It is a noteworthy fact that we revealed statistically significant aberrant cytoplasmic abundant proteins (even though they did not show large difference) in the mutant cells. The pathway cluster analysis of the revealed proteins, which are abnormally increased in the cytoplasmic fraction in the mutant cells, are related to protein folding (Cct5, Cct7, Cct8), aminoacyl−tRNA biosynthesis (Farsb, Nars, Txnrd1), Wnt signaling (Cltc, Plcb3, Plec, Psmd3, Ruvbl1), the synaptic vesicle cycle (Cltc, Nsf), and Hippo signaling (Camk2d, Plcb3, Ruvbl1). However, the proteins associated with the RNA transport and processing (Dhx9, Fmr1, Srsf3, Srsf6, Tra2b) pathway were reversely increased in the nucleus by the mutant SOD1. Our findings support the notion that abnormal protein aggregation and an RNA transport dysfunction might be a strong mechanism for motoneuron degeneration in ALS.

Our investigation began on the basis that we successfully fractionated the subcellular organelle. To confirm the quality of the fractionation, we repeatedly evaluated this issue using different approaches. We first checked the nuclear and

cytoplasmic markers using a Western blot, and expressed the distribution of marker proteins from the samples in the z-score, and performed the principal component analysis. Furthermore, we evaluated the organelle distribution of the marker proteins that are identified exclusively in either the cytoplasm or the nuclear fraction (Supplemental figure S4). These diverse methods strongly supported the accuracy of our fractionated data.

The SOD1 mutations, themselves, did not repress the dismutase activity or the catalytic activity to reverse the superoxide or hydrogen peroxide that was retained in the motoneuron cells (Borchelt et al. 1994). From this finding, the neurotoxicity of the SOD1 mutation can be understood in the concept of 'gain of toxicity' rather than 'loss of toxicity'. The SOD1 mutation is now known to alter free radical processing, as well as copper-zinc binding, and to aggravate the aberrant protein aggregation (Beckman et al. 2001; Cleveland and Rothstein. 2001). The human familial ALS cell culture model, and the SOD1 transgenic mice, verified the anti-SOD1 reactive cytoplasmic aggregations in both the motoneuron and the surrounding astrocytes (Bruijn et al. 1998; Durham et al. 1997). These abnormal inclusions were suggested to change protein folding, alter free radical machinery, and further protein degradation. Finally, the normal protein functions, which are

essential for motoneuron viability, might be deregulated. However, the proteins that contribute to the neurodegeneration in the SOD1 mutation were not fully evaluated in the proteomic level, and were not analyzes before in regard to subcellular mislocalization.

Among the proteins detected using mass spectrometry, Cct5, Cct8 and Atp5b were validated to regulate their nuclear cytoplasmic distribution by SOD1. Cct5 and Cct8 are members of the chaperonin, containing the TCP-1 complex related to the protein synthesis and transport. Molecular chaperones bind to the amino acid chain, and stabilize and promote protein folding in order to give proteins their proper function. Incorrect protein folding is known to lead to protein aggregation and acerbate disease. Chaperonin, containing the TCP-1 complex, was reported to suppress the huntingtin aggregation. Chaperonin, containing the TCP-1 complex (especially the subunit of Cct1) bind directly to the huntingtin protein in mutant huntingtin transgenic mice. The suppression of Cct1 resulted in an increased huntingtin protein aggregation (Shen and Frydman 2013; Shahmoradian et al. 2013). ALS and Huntington's disease share a common pathology as cytoplasmic aggregations. The Cct5 and Cct8 changes observed in our study might indicate a compensatory response to removing the aggregate result from the SOD1 mutation. Our study is the first to mention

the association of Chaperonin (containing TCP-1 complex proteins), Cct5 and Cct8 with the ALS pathology.

Atp5b, a subunit of the mitochondrial ATP synthase, is normally placed in the inner membrane of the mitochondria. In our experimental design, the mitochondria fraction was also mixed in with the cytoplasmic extracts. Relative increments of Atp5b in the nucleus might indirectly present the depletion of Atp5b in the mitochondria in the mutant cells. A lower expression of the ATP synthase, and its dysfunction, has been described in many neurodegenerative disorders, including Alzheimer's disease (Kim et al. 2000). Dysfunction in the mitochondria seemed to change the amyloid precursor protein metabolism, and enhance the amyloid β-peptides aggregation in the cytoplasm in Alzheimer's disease (Busciglo et al. 2002). Previous reports of other neurodegenerative diseases reinforced our discovered proteins as being true candidates for the ALS pathology. Validation from the total cell lysates of the mutant cells showed the decreased proteins associated with the nucleosome assembly and the phosphate metabolic process in our study. This suggests smaller amounts of DNA replication and genomic instability, which might further suggest a 'loss of function' of specific genes in the mutant SOD1 cells.

The functional pathways of other proteins that changed their distribution in the mutant cell using mass spectrometry,

but were unable to be validated are related to the tRNA aminoacylation, Wnt signaling, the synaptic vesicle cycle and Hippo signaling. Defects in the tRNA synthetase cause intracellular misfolded protein accumulation, induce the upregulation of the cytoplasmic protein chaperone, and provoke the unfolded protein response, resulting in neurodegeneration (Lee et al. 2006). Increased tRNA aminoacylation related proteins in the cytoplasm might also be a compensatory response for the mutant SOD1 toxicity. The alteration of the tRNA aminoacylation, in ALS, was suggested in several studies. Lysyl-tRNA synthetase is known to interact with the SOD1 aggregation only in mutant $SOD1^{G93A}$ mice (Kunst et al. 1997). The mitochondrial form of Lysyl-tRNA synthetase was vulnerable to aggregation before being imported to the mitochondria in the SOD1 mutation. In the end, the SOD1 mutation resulted in a protein accumulation and mitochondrial dysfunction (Kawamata et al. 2008). In the mutant $SOD1^{G93A}$ mice, a glycyl-tRNA synthetase mutation delayed the disease onset and increased the life span (Banks et al. 2009). Wnt signaling is associated with the stabilization of β-catenin, which is related to the nuclear import of the proteins. The increased Wnt signaling observed in our study supports the compensatory mechanism in the mutant cells. Additionally, the effect of the Hippo signaling pathway change, especially MST1,

was also reported in the earlier ALS mouse model (Lee et al. 2013). Interestingly, the proteins affiliated with the RNA transport were increased in the nuclei of the mutant cells, which might imply the breakdown of the RNA transport machinery, or the presence of immature RNA.

Additional functional information about the detected proteins, using GO terms, was associated with the cell cycle. The cell cycle protein change was described in ALS. Abnormally accumulated E2F-1 in the cytoplasm, and the hyperphosphorylated retinoblastoma gene product (pRb) in the nucleus was found in the motoneurons of ALS patients (Ranganathan and Bowser 2003). These increments indirectly supported the re-entry of the motoneuron cells into the G1 phase, and contribute to the cell death. The cell cycle was also altered in the mutant SOD1 cell model, and it was thought to be associated with cyclins regulator p27 (Cova et al. 2010).

A recent study explored the cytosolic proteome change in ALS in mutant human SOD1 NSC34 cell. Proteins that have the function of the RNA process, antioxidant defense, nitric oxide metabolism and protein degradation were found to alter (Allen et al. 2003). Their findings similarly point to the protein synthesis impairment in the ALS model, as our findings do. However, the concrete proteins discovered were mostly different; this discrepancy might result from the technical

difference in the proteome quantification, and also from the alteration criteria difference that we considered to be both the changes in the nucleus and the cytoplasm location (NC ratio)－they only compared for the change in cytoplasmic abundance. Another cell culture model for ALS, with TDP－43 loss, was analyzed for its proteome change in the nuclear－cytoplasmic location (Stalekar et al. 2015). The detected proteins in their research were functionally annotated for RNA processing and nuclear shuttling.

There are several limitations to our study. Even though we used the high－throughput novel technique of proteomics, the issues of proteome coverage range, the overall dynamic extent of the detection and the risk of false positive identification still remained. Furthermore, our proteome coverage was very low, when we compared it to the transcriptome data. To check the proteome detection rate using our mass spectrometry methods, we additionally performed a transcriptome analysis for the same triplicate samples of each fraction (Supplemental figure S2). The proteome coverage of the mRNA only reached 5.6% (597 proteins among 106972 mRNA were detected). This low coverage of proteome was thought to be mostly from our study design, in that we only included proteins and mRNA that co－expressed in the nucleus and cytoplasm fraction. Considering the gene expression

process, proteins that only resided in the cytoplasmic portion were hypothetically excluded for analysis. Beyond this technical limitation, the physicochemical properties of the proteins (molecular weight, hydrophobicity, coding sequence length, isoelectric points, etc.), and the biological process as posttranscriptional regulation might have influenced the detection rate. We found higher proteome coverage for more abundant transcripts (Supplemental figure S3). The proteins we detected in the mass spectrometry were suspected to be mostly abundant proteins. The small sample size and the technical limitation of mass spectrometry, such as the low ability to detect less abundant or insoluble proteins, may influence the results.

As far as we know, our results are the first data that present a comparative analysis of the proteome of a familial ALS cell culture model that considered the nucleus and cytoplasm change differently. This methodological approach, and its results, is critical because the main pathological mechanism for ALS is now known to be associated with mislocalization, not only by the absolute amount change in specific proteins or transcripts. We discovered non-referred pathological candidates (Cct5, Cct8 and Atp5b) for the nuclear cytoplasmic mislocalization; their function needs to be investigated. Our integrated omics data will provide significant

information for quantitative biology, and may suggest a clear regulatory mechanism in the familial ALS model.

REFERENCES

Allen S, Heath PR, Kirby J, et al. Analysis of the cytosolic proteome in a cell culture model of familial amyotrophic lateral sclerosis reveals alterations to the proteasome, antioxidant defenses, and nitric oxide synthetic pathways. J Biol Chem 2003;278(8):6371-6383.

Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, et al. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem Biophys Res Commun 2006;351(3):602-611.

Boillee S, Vande Velde C, Cleveland DW. ALS: a disease ofmotor neurons and their nonneuronal neighbors. Neuron 2006;52(1):39-59.

Banks GT, Bros-Facer V, Williams HP, Chia R, Achilli F, Bryson JB, Greensmith L, Fisher EM. Mutant glycyl-tRNA synthetase (Gars) ameliorates SOD1(G93A) motor neuron degeneration phenotype but has little affect on Loa dynein heavy chain mutant mice. PLoS One 2009;4(7):e6218.

Beckman JS, Estévez AG, Crow JP, Barbeito L. Superoxide dismutase and the death of motoneurons in ALS. Trends Neurosci 2001;24(11 Suppl):S15-20.

Borchelt DR, Lee MK, Slunt HS, Guarnierei M, Xu ZS, Wong PC, Rown RH, PRoce DL, Sisodia SS, Cleveland DW. Superoxide dismutase 1 with mutations linked to familial amyotrophic lateral sclerosis possesses significant activity. Biochem Biophys Res Commun 2012;420(2):397-403.

Burgess RJ, Zhang Z. Histone chaperones in nucleosome assembly and human disease. Nat Struct Mol Biol 2013;20(1):14-22.

Bruijn LI, Houseweart MK, Kato S, Anderson KL, Anderson SD, Ohama E, Reaume AG, Scott RW, Cleveland DW. Aggregation and motor neuron toxicity of an ALS-linked SOD1 mutant

independent from wild-type SOD1. Science 1998; 281(5384):1851-1854.

Busciglio J, Pelsman A, Wong C, Pigino G, Yuan M, Mori H, YanknerBA Altered metabolism of the amyloid beta precursor protein is associated with mitochondrial dysfunction in Down's syndrome. Neuron 2002;33(5):677-688.

Byun K, Young Kim J, Bayarsaikhan E, Kim D, Jeong GB, Yun KN, Kyeong Min H, Kim SU, Yoo JS, Lee B. Quantitative proteomic analysis reveals that lipopolysaccharide induces mitogen-activated protein kinase-dependent activation in human microglial cells. Electrophoresis 2012;33(24):3756-3763.

Cansizoglu AE, Lee BJ, Zhang ZC, Fontoura BM, Chook YM. Structure-based design of a pathway-specific nuclear import inhibitor. Nat Struct Mol Biol 2007;14(15):452-454.

Carvalho PC, Xu T, Han X, Cociorva D, Barbosa VC, Yates JR, 3rd. YADA: a tool for taking the most out of high-resolution spectra. Bioinformatics 2009;25(20):2734-2736.

Cashman NR, Durham HD, Blusztajn JK, Oda K, Tabira T, Shaw IT, et al. Neuroblastoma x spinal cord (NSC) hybrid cell lines resemble developing motor neurons. Dev Dyn 1992;194(3):209-221.

Chen Y-A, Tripathi LP, Mizuguchi K TargetMine, an Integrated Data Warehouse for Candidate Gene Prioritisation and Target Discovery. PLoS One 2011;6(3):e17844.

Cleveland DW, Rothstein JD. From Charcot to Lou Gehrig: deciphering selective motor neuron death in ALS. Nat Rev Neurosci 2001;2(11):806-819.

Cova E, Ghiroldi A, Guareschi S, Mazzini G, Gagliardi S, Davin A, Bianchi M, Ceroni M, Cereda C. G93A SOD1 alters cell cycle in a cellular model of Amyotrophic Lateral Sclerosis. Cell Signal 2010;22(10):1477-1484.

Dormann D, Haass C. TDP-43 and FUS: a nuclear affair. Trends Neurosci 2011;34(7):339-348.

Durham HD, Roy J, Dong L, Figlewicz DA. Aggregation of mutant Cu/Zn superoxide dismutase proteins in a culture model of ALS. J Neuropathol Exp Neurol 1997;56(5):523-530.

Gatto L, Breckels LM, Wieczorek S, Burger T, Lilley KS. Mass-spectrometry-based spatial proteomics data analysis using pRoloc and pRolocdata. Bioinformatics 2014;30(9):1322-1324.

Hong I, Kang T, Yun KN, Yoo Y, Park S, Kim J, An B, Song S, Lee S, Kim J, Song B, Kwon KH, Kim JY, Park YM, Choi S. Quantitative proteomics of auditory fear conditioning. Biochem Biophys Res Commun 2013;434(1):87-94.

Hung MC and Link W. Protein localization in disease and therapy. J Cell Sci 2011;124(Pt 20):3381-3392.

Ito D, Seki M, Tsunoda Y, Uchiyama H, Suzuki N. Nuclear transport impairment of amyotrophic lateral sclerosis-linked mutations in FUS/TLS. Ann Neurol 2011;69(1):152-162.

Kanekura K, Suzuki H, Aiso S, Matsuoka M. ER stress and unfolded protein response in amyotrophic lateral sclerosis. Mol Neurobiol 2009;39(2):81-89.

Kawamata H, Magrane J, Kunst C, King MP, Manfredi G. Lysyl-tRNA Synthetase Is a Target for Mutant SOD1 Toxicity in Mitochondria. J Biol Chem 2008;283(42):28321-28328.

Kim SH, Vlkolinsky R, Cairns N, Lubec G (2000) Decreased levels of complex III core protein 1 and complex V beta chain in brains from patients with Alzheimer' s disease and Down syndrome. Cell Mol Life Sci 2000;57(12):1810-1816.

Kosugi S, Hasebe M, Entani T, Takayama S, Tomita M, Yanagawa H. Design of peptide inhibitors for the importin alpha/beta nuclear import pathway by activity-based profiling. Chem Biol 2008;15(9):940-949.

Kunst CB , Mezey E, Brownstein MJ, Patterson D. Mutations in SOD1 associated with amyotrophic lateral sclerosis cause novel protein interactions. Nat Genet 1997;15(1):91-94.

Kwon OK, Sim J, Yun KN, Kim JY, Lee S. Global phosphoproteomic analysis of Daphnia pulex reveals evolutionary conservation of Ser/Thr/Tyr phosphorylation. J Proteome Res 2014;13(3):1327-1335.

Lee JK, Shin JH, Hwang SG, Gwag BJ, McKee AC, Lee J, Kowall NW, Ryu H, Lim DS, Choi EJ: MST1 functions as a key modulator of neurodegeneration in a mouse model of ALS. Proc Natl Acad Sci U S A 2013;110(29):12066-12071.

Lee JW, Beebe K, Nangle LA, Jang J, Longo-Guess CM, Cook SA, Davisson MT, Sundberg JP, Schimmel P, Ackerman SL. Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. Nature 2006;443(7107):50-55.

Li Z, Adams RM, Chourey K, Hurst GB, Hettich RL, Pan C. Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ Orbitrap Velos. J Proteome Res 2012;11(3):1582-1590.

Polpitiya AD, Qian WJ, Jaitly N, Petyuk VA, Adkins JN, Camp DG, 2nd, et al. DAnTE: a statistical tool for quantitative analysis of -omics data. Bioinformatics 2008;24(13):1556-1558.

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 2006;314(5796):130-133.

Ranganathan S, Bowser R. Alterations in G(1) to S phase cell-cycle regulators during amyotrophic lateral sclerosis. Am J Pathol 2003;162(3):823-835.

Raso C, Cosentino C, Gaspari M, Malara N, Han X, McClatchy D, et al. Characterization of breast cancer interstitial fluids by TmT labeling, LTQ-Orbitrap Velos mass spectrometry, and pathway analysis. J Proteome Res 2012;11(6):3199-3210.

Rosen DR, Siddique T, Patterson D, et al. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature 1993;362(6415):59-62.

Shen K, Frydman J. The interplay between the chaperonin TRiC and N-terminal region of Huntingtin mediates Huntington' s Disease aggregation and pathogenesis. Protein Quality Control in Neurodegenerative Diseases. Berlin Heidelberg:Springer, 2013;121-132.

Shahmoradian SH, Galaz-Montoya JG, Schmid MF, Cong Y, Ma B, Spiess C, Frydman J, Ludtke SJ, Chiu W. TRiC's tricks inhibit huntingtin aggregation. Elife 2013;9(2):e00710.

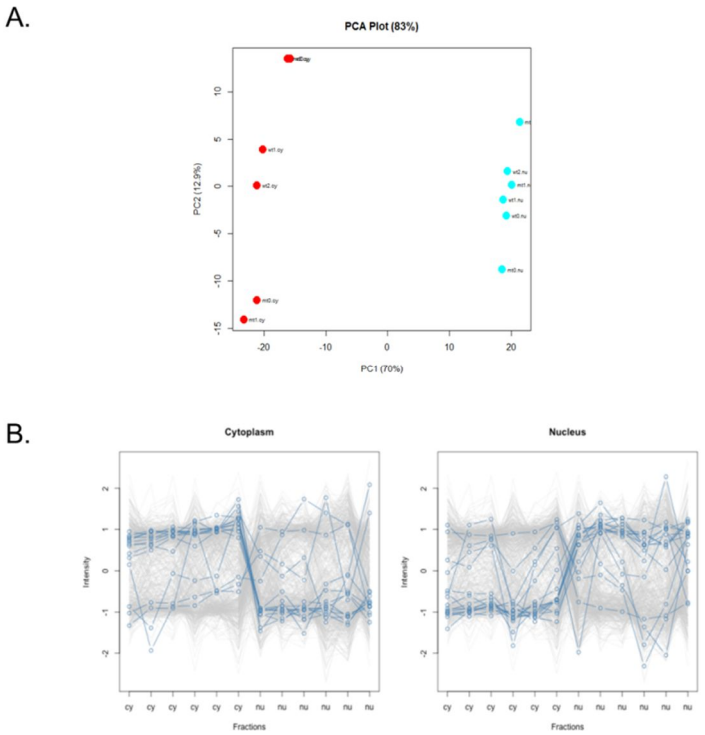
Stalekar M, Yin X, Rebolj K, et al. Proteomic analyses reveal that loss of TDP-43 affects RNA processing and intracellular transport. Neuroscience 2015;293:157-170.

Tabb DL, McDonald WH, Yates JR, 3rd. DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J Proteome Res 2002;1(1):21-26.

Vance C, Rogelj B, Hortobagyi T, De Vos KJ, Nishimura AL, Sreedharan J, et al. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science 2009(5918);323:1208-1211.

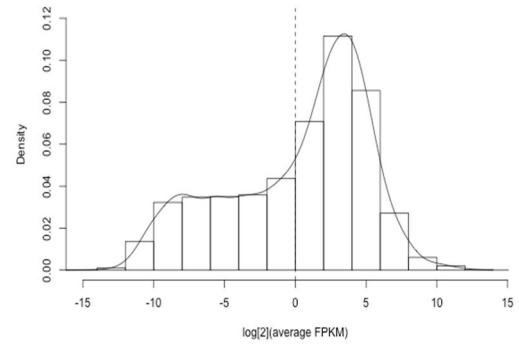
Washburn MP, Wolters D, Yates JR, 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 2001;19(3):242-247.

Supplemental Data



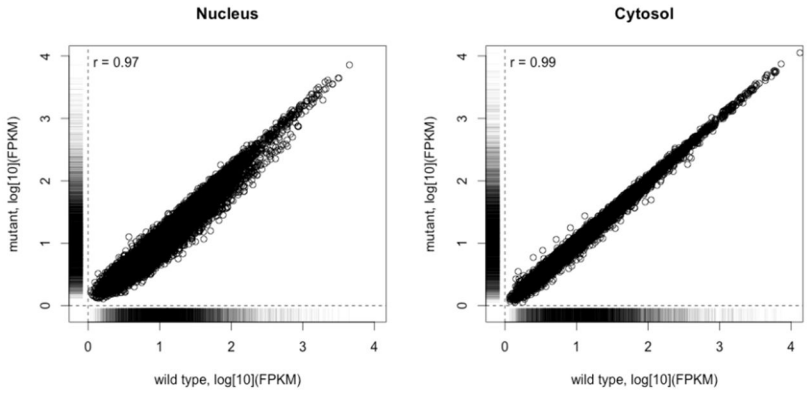
Supplemental figure S1. A quality check of the nuclear–cytoplasmic fractionation.
(A) The principal component analysis of the quantified proteins from the wild-type (wt) and mutant (mt) cells are shown. Samples from different fractions are color coded (red for cytoplasm, and cyan for nucleus). The X-axis, Y-axis were labeled with the first principal component and the second principal component, accounting for 70% and 12.9% of the total variation, respectively. (B) The distribution of cytoplasmic (left) and nuclear (right) marker proteins across 12 samples (x-axis) are presented, in their relative abundance, expressed as the z-score (y-axis). The light blue represents the marker proteins corresponding to the subcellular fraction; the light grey is for the other proteins. cy, cytoplasmic fraction; nu, nuclear fraction. The marker proteins were obtained from pRoloc' s organelle markers (Gatto L et al. 2014).

A.



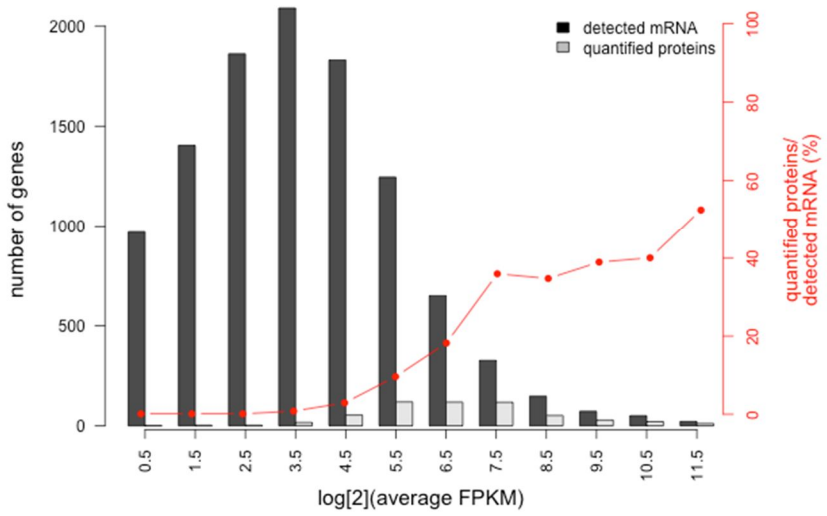
B.

C.



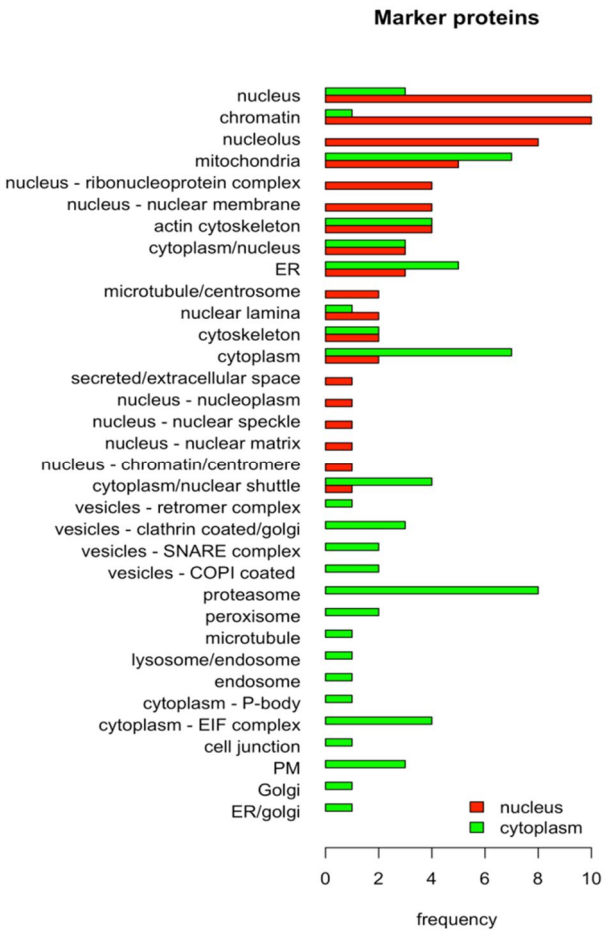
Supplemental figure S2. The distribution of the raw FPKM data and the correlation between the wild and mutant cells.

(A) The total RNA samples from the wild and mutant SOD1 (G93A) NSC34 cells were analyzed in triplicate. The average RNA expression levels (measured as fragments per kilobase of transcript per million mapped fragments, FPKM) were log-transformed. Most genes showed normal distribution, and a few genes formed a 'shoulder' to left of the distribution. We excluded the very low abundant transcripts, with average FPKM < 1 (n=7231) (presumably non-functional), from our subsequent analysis. (B, C) The absolute transcript levels from the wild and mutant cells correlate well.



Supplemental figure S3. The proteome coverage rate, dependent on the transcripts abundance.

The protein number that was detected using mass spectrometry showed an increase in the more abundant mRNA.



Supplemental figure S4. The organelle distributions of the marker proteins.

The quality of fractionation was confirmed by evaluating the organelle distribution marker proteins that are identified exclusively in the cytoplasm or nuclear fraction. The exceptional data were as following: Pds5b, Rfc4, Tcea1 were nucleus marker proteins, but were identified exclusively in the cytoplasm fraction in the data. Reversely, cytoplasm marker proteins (Aldoa, Tacc3) were exclusively shown in the nucleus fraction.

국문 초록

서론: 세포핵 및 원형질 내 단백의 정상적인 분포는 진핵세포가 정상적인 세포 기능을 유지하고 RNA 프로세싱, 번역, 단백 상호작용 및 번역 후 변형 등의 주요 생물학적 과정을 수행하는데 필수적이다. 세포 내 단백의 세포핵 및 원형질 내의 비정상적인 분포는 종양과 퇴행성질환 등의 여러 질환에서 발견되어 이들 질환의 중요한 병인으로 알려졌다. 루게릭병 및 전두측두엽치매에서도 최근 transactive response DNA-binding protein 43kDa (TDP-43) 및 Fused-in-sarcoma (FUS)로 밝혀진 단백의 비정상적인 축적이 원형질 내 관찰되며, 이러한 단백의 세포내의 이상분포가 주요 신경독성 원인으로 주목되고 있다. 본 연구에서는 루게릭병 운동신경세포모델의 세포단백질 전체를 분석하여 핵-원형질 내 분포 이상이 나타내는 단백과 그 기능을 확인하고자 한다.

방법: 정상 SOD1 또는 G93A 돌연변이 SOD1 을 가진 NSC34 세포의 원형질 및 핵을 분리하였다. 분리된 시료에 대해 액체크로마토그래피 시스템 질량분석을 진행하여 핵-원형질 내의 단백 전체에서 세포구획간, 야생형-돌연변이형 간 변화를 보이는 단백을 확인하였다. 돌연변이 SOD1 에 의해 핵-원형질 내간 분포 변화를 보이는 단백에 대해 TargetMine 및 DAVID 데이터베이스를 이용하여 그 기능을 확인하였으며, 웨스턴블롯을 통해 해당 단백의 세포 구획 내 발현량의 차이를 확인하였다.

결과: 측정된 11,216 개의 펩티드와 1,925 개의 단백 가운데 23%의 펩티드 및 32%의 단백이 세포핵 및 원형질에 동시에 분포하였다. 세포핵 및 원형질에 동시 분포하는 634 개 단백 중 503 개의 단백이 유의한 세포핵-원형질간 분포 불균형을 보이며, 야생형 및 돌연변이 세포간 유의한 차이를 보이는 단백은 71 개였다. 이 가운데 37 개의 단백이 야생형과 비교해서 돌연변이 세포에서 유의하게 세포핵-원형질간 분포차를 보였다 ($p<0.05$). 돌연변이 세포에서 세포핵에서 원형질로 단백의 이동이 관찰되는 단백(Dhx9, Fmr1, Srsf3, Srsf6, Tra2b)은 RNA 이동이나 프로세싱에 관여하였다. 반대로 돌연변이에 의해 핵에서 원형질로 이동하는 단백은 단백질접힘(Cct5, Cct7, Cct8), aminoacyl-tRNA 생합성 (Farsb, Nars, Txnrd1), Wnt 신호전달경로 (Cltc, Plcb3, Plec, Psmd3, Ruvbl1), Hippo 신호 전달 경로 (Camk2d, Plcb3, Ruvbl1) 및 시냅스소포주기(Cltc, Nsf) 에 관여하였다.

결론: 본 연구는 정상 및 돌연변이 SOD1 유전자를 발현하는 운동신경세포에서 세포핵 원형질간의 단백 분포 양상 및 변화를 전체 단백질체에서 평가하였다. 발견된 루게릭병 세포 모델에서 핵-원형질간의 분포 이상을 보이는 단백들의 병리적 작용기전을 확인하기 위해서는 향후 추가적인 연구가 필요하겠다.

주요어 : 루게릭병, 세포핵-원형질간의 분포불균형, 단백질체, 질량분석법

학 번 : 2011-31140