Low level of expression of C-terminally truncated human FUS causes extensive changes in the spinal cord transcriptome of asymptomatic transgenic mice

The running title: **Transcriptome of asymptomatic FUS transgenic mice.**

Ekaterina A. Lysikova[1,2,*], Sergei Funikov[3], Alexander P. Rezvykh[3,4], Kirill D. Chaprov[1], Michail S. Kukharsky[1,5], Aleksey Ustyugov[1], Alexey V. Deykin[6], Ilya. A. Flyamer[7], Shelagh Boyle[7], Sergey O. Bachurin[1], Natalia Ninkina[1,2] and Vladimir L. Buchman[1,2]

[1]*Institute of Physiologically Active Compounds, Russian Academy of Sciences, 1 Severnyj Proezd, Chernogolovka, Russian Federation*
[2] *School of Biosciences, Cardiff University, Sir Martin Evans Building, Museum Avenue, Cardiff, CF10 3AX, UK*
[3]*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 119991 Vavilova str., 32, Moscow, Russia*
[4]*Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Moscow Region, 141701, Russia*
[5]*Pirogov Russian National Research Medical University, Ostrovitianova str 1, Moscow, Russian Federation*
[6]*Institute of Gene Biology, Russian Academy of Sciences, Vavilova str., 34/5, Moscow, 19334, Russian Federation*
[7]*MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK*

*Corresponding author lysikovaea@ipac.ac.ru Tel. +84965242588
Institute of Physiologically Active Compounds, Russian Academy of Sciences, 1 Severnyj Proezd, Chernogolovka, Russian Federation

Abstract

A number of mutations in a gene encoding RNA-binding protein FUS have been linked to the development of a familial form of amyotrophic lateral sclerosis known as FUS-ALS. C-terminal truncations of FUS by either nonsense or frameshift mutations lead to the development of FUS-ALS with a particularly early onset and fast progression. However, even in patients bearing these highly pathogenic mutations the function of motor neurons is not noticeably compromised for at least a couple of decades, suggesting that until cytoplasmic levels of FUS lacking its C-terminal nuclear localisation signal reaches a critical threshold, motor neurons are able to tolerate its permanent production.

In order to identify how the nervous system responds to low levels of pathogenic variants of FUS we produced and characterised a mouse line, L-FUS[1-359], with a low neuronal expression level of a highly aggregation-prone and pathogenic form of C-terminally truncated FUS. In contrast to mice that express substantially higher level of the same FUS variant and develop severe early onset motor neuron pathology, L-FUS[1-359] mice do not develop any clinical or histopathological signs of motor neuron deficiency even at old age. Nevertheless, we detected substantial changes in the spinal cord transcriptome of these mice compared to their wild type littermates. We suggest that at least some of these changes reflect activation of cellular mechanisms compensating for the potentially damaging effect of pathogenic FUS production. Further studies of these mechanism might reveal effective targets for therapy of FUS-ALS and possibly, other forms of ALS.

Introduction

Mutations in the gene encoding FUS are among the most frequent causes of familial ALS (fALS) and many of these mutations lead to partial change of FUS localisation from the nucleus to the cytoplasm of motor neurons. Mislocalised FUS aggregates forming large cytoplasmic inclusions, and negatively affects various important cell functions [1-4]. Aggregation of FUS has also been observed in the cytoplasm of neurons in a subset of patients with sporadic form of ALS (sALS) and in patients with certain variants of frontotemporal lobar degeneration (FTLD-FUS) in the absence of any mutations or polymorphisms of the gene encoding FUS [5-9].

Various invertebrate and vertebrate animal models of FUS malfunction, including several transgenic FUS mouse lines, have been produced and characterised in order to shed light on molecular mechanisms of ALS-FUS and FTLD-FUS but the results and conclusions of these studies were not always consistent (reviewed in [10-12]). This might reflect a complex combination of intrinsic (e.g. differences, even subtle, in genetic background) and environmental factors (e.g. type of animal diet) that affect the manifestation of pathology in animals with genetic alterations of FUS structure or expression. However, the course of the disease development in all animal models and in ALS patients is common in one aspect: motor neurons of transgenic mice expressing pathogenic variants of human FUS, as well as motor neurons of ALS patients carrying germ line mutations of FUS, survive and function normally for a relatively long period of time. This suggests that motor neurons successfully employ certain intracellular defence systems that prevent damaging effects of the pathogenic protein when its level is still low. An ability to artificially boost the activity of these systems, either directly or via modulation of their regulatory pathways, can help counteracting pathogenic effects of FUS and potentially other ALS-linked proteins, even when they accumulate in the neuronal cytoplasm beyond a certain threshold and the intrinsic neuronal defence systems become overloaded. Therefore, it is important to identify how the nervous system responds to low levels of pathogenic variants of FUS and prevents potentially damaging consequences of its permanent production.

Previously we have described production and characterisation of a transgenic mouse line FUS 1-359 clone 19 (S-FUS[1-359]) with neurospecific expression of a C-terminally truncated form of human FUS (tr-hFUS) lacking its nuclear localisation signal (NLS) and arginine/glycine rich (RGG) domains [13]. This modification recapitulates consequences of FUS gene mutations in certain forms of familial ALS [14-17] typified histopathologically by mislocalisation of modified FUS and clinically by an aggressive disease with young onset and rapid progression [18]. Although hemizygous mice of the S-FUS[1-359] line express tr-hFUS at a lower level than the endogenous mouse FUS, accumulation of tr-hFUS in the cytoplasm of neurons leads to its aggregation and the development of FUS-proteinopathy, characterised by the presence of large inclusions, neuronal dysfunction and death. These pathological changes are most pronounced in the motor neurons of the spinal cord, causing the development of severe and fast progressing motor neuron disease in young adult S-FUS[1-359] mice and death of animals within one to three weeks from the onset of clinical signs [13].

Here we assessed changes in the spinal cord transcriptome of L-FUS[1-359] mice, another transgenic mouse line produced along with S-FUS[1-359] line using the same transgenic construct. Although these mice have substantially lower levels of tr-hFUS expression in the spinal cord and do not develop clinical or histopathological signs of motor neuron deficiency, we found a significant number of transcripts that differentially regulated in the spinal cord of L-FUS[1-359] mice when compared to their wild type littermates.

Materials and Methods

Animals

The L-FUS[1-359] mouse line was produced by pronuclear microinjection of a transgenic cassette for expression of human FUS under control of regulatory elements of the mouse Thy-1 gene as described in our previous publications [13,19]. The cassette was identical to the cassette used for production of S-FUS[1-359] mouse line and therefore, the same mouse genotyping

protocol was employed [13]. Mice were maintained under 12/12 hour dark/light cycle with free access to food and water.

Comparison of transgene copy number by quantitative PCR

Genomic DNA from ear biopsies was purified with Wizard® SV Genomic DNA Purification System (Promega) and 0.1 µg was used for real-time qPCR amplification with primers 5' – TCTTTGTGCAAGGCCTGGGT – 3' and 5' – TAATCATGGGCTGTCCCGTT – 3' targeting human FUS in the transgenic cassette. Amplification of the GAPDH gene fragment with primers 5' – CACTGAGCATCTCCCTCACA – 3' and 5' – GTGGGTGCAGCGAACTTTAT – 3' was used as a reference when the relative number of transgenic cassettes in DNA samples was assessed. The CFX96 real-time PCR detection system (Biorad) and CYBR GreqPCR protocol were employed with the cycle parameters: 10 min at 95°C followed by 40 cycles of 15 s at 95°C and 60 s at 60°C.

Motor behaviour testing

Animal motor performance was assessted on the inverted grid and accelerating rotarod as described previously [20,21].

A locomotor activity was monitored in a square 30 x 30 cm activity camera under 25 lux illumination using the TRU SCAN Activity Monitoring System (Coulbourn Instruments, Whitehall, PA, USA).

RNA extraction

Total RNA was extracted from thoracic spinal cords of 9-week old male mice using RNeasy Plus Mini Kit with genomic DNA eliminator columns (Qiagen). For quantification and the quality control of RNA samples Qubit fluorimetry and TapeStation analysis were employed.

Quantitative RT-PCR

1 µg of total RNA was reverse transcribed in the presence of random hexamer primers according to manufacturer's instructions (Eurogene, Russia) and resulting cDNA was used for qPCR reaction with GAPDH as a reference gene as previously described [22,23]. For analysis of FUS expression same primers and amplification protocol as for the genomic qPCR above were used. The following

combinations of primers were used for analysis of expression of selected transcripts identified in RNA sequencing experiment:

Mpz: 5' – TCAAGATGGACGCGACACAA – 3'

                5' – CAAAGGCCCAGAGTGTCTCA – 3'

Prx: 5' – CTCAGCTTGCAAGAAGGGGA – 3'

                5' – CAGGCGAAGTGCATCCTCAT – 3'

PPARA: 5' – CTGCCTTCCCTGTGAACTGA – 3'

                5' – ACAGAGCGCTAAGCTGTGAT – 3'

Gpr179:   5' – CATGGTTGTGGCGGAAATGC – 3'

                5' – TGTAGCGTGGTTCGTGGAAG – 3'

Fggy: 5' – GCAAAGATCCAGTGTTTGTACCA – 3'

                5' – GCTGAGCCTTCTTAATCAGATCC – 3'

Protein extraction and Western blot analysis

Total proteins were extracted from the mouse spinal cords, separated by SDS-PAGE and transferred to PVDF membrane by semidry blotting as described elsewhere [13,19]. Human FUS protein was detected using rabbit polyclonal antibody 14080 (a kind gift from Don Cleveland) specific to its N-terminal epitope and endogenous mouse protein – with a rabbit polyclonal antibody specific to C-terminal epitope absent in human FUS[1-359] protein (NSJ Bioreagents). Secondary anti-rabbit HRP-conjugated antibodies (GL Healthcare) and WesternBright TM Sirius chemiluminescent detection system (Advansta) were used for detection of specific protein bands. For loading control membranes were re-probed with mouse monoclonal antibody against beta-actin (clone AC-15, Sigma-Aldrich).

Histology and immunohistochemistry

The protocol for preparation of histological sections of the mouse spinal cord and immunodetection of human FUS with antibody 14080 was described previously [13]. The same

protocol was used to detect endogenous mouse FUS with an antibody specific to its C-terminal epitope.

Nissl staining of transverse sections through thoracic and lumbar regions of the spinal cord was performed as described elsewhere [13, 22]. The number of large cell body containing Nissl granules, clear nuclear envelope and intensely stained nucleolus was counted for every 6th of 8µm thick sections. An average number from six sections per spinal cord region per animal was used for further statistical analysis.

RNA sequencing

cDNA libraries for the dual indexed single-end sequence analysis were prepared from equal amounts (270 ng) of each total RNA sample using Illumina TruSeq Stranded Total RNA LT Sample Prep Kit. Following libraries' quality checks and normalisation, cDNA libraries were sequenced on Illumina NextSeq 500 to generate single end 75 bp reads.

Sequencing data analysis

Processing of RNA sequencing raw data included the adapter, length and quality trimming by Trimmomatic, mapping of reads to the mouse genome (release GRCm38) by STAR aligner, counting the overlap of reads with genes by featureCounts, implemented in PPLine script [24-27]. Differential gene expression analysis was performed with the edgeR package using a Fisher exact test between experimental groups [28]. We used the Benjamini-Hochberg method for multiple testing FDR correction. The genes with expression level ≥ 1 Log10 CPM (counts per million) and FDR ≤ 0.05 were taken into the account and considered as differentially expressed. Multidimensional scaling between all experimental samples were performed with limma package, pairwise distances were calculated using root-mean-square of Log2FoldChange values between experimental groups [29].

Sorting of microglial and neuronal genes of the whole spinal cord samples were performed using the published datasets of purified microglia and laser-microdissected ventral horns of mouse spinal cord [30,31]. We applied a 5-fold change cutoff of gene expression levels to discriminate neurons-

and microglia-specific genes as suggested previously [31,32]. These criteria were used to classify differentially expressing genes (DEGs) from our set as either neuron-specific or microglia-specific or shared, if the difference was less than 5 times.

The analysis of enriched gene ontology (GO) terms was performed using DAVIDWebService package for R with a P-value = 0.05 (Fisher exact test) and Q-value = 0.05 (BH-corrected p-value) as cutoffs [33]. Only those GO terms that satisfied these criteria and also showed enrichment by differentially expressed genes ≥ 4 were taken for further analysis. Then, redundant GO terms were removed using REVIGO software [34].

Differential expression analysis, data visualization, and GSEA (Gene Set Enrichment Analysis) were performed using R project for statistical computing [35]. Visualization of experimental data was made with ggplot2 and GOplot R packages [36,37].

RNA sequencing data were deposited in Gene Expression Omnibus (GEO) under the number GSE130604.

Results

Transgenic mice expressing low level of C-terminally truncated human FUS in their neurons do not develop motor neuron pathology

A founder of L-FUS[1-359] transgenic mouse line was produced by pronucleus microinjection of the same transgenic construct as was used for production of the previously described S-FUS[1-359] line, namely a linear DNA fragment containing human FUS cDNA encoding amino acids 1 to 359 (tr-hFUS) under control of regulatory elements of mouse Thy-1 gene [13]. The L-FUS[1-359] line was first established on a C57Bl6J background and then transferred to CD1 genetic background by serial backcrosses (>5) with the wild type CD1 mice. Intercrosses of hemizygous animals produced homozygous L-FUS[1-359] mice and their wild type littermates that were used in some of experiments described below. Analysis of genomic DNA revealed that L-FUS[1-359] mice carry the same number of tandemly arranged repeats of the transgenic cassette as S-FUS[1-359] mice

(Suppl. Fig. 1a). However, the localisation on different chromosomes (Suppl. Fig. 1b) and presumably, different chromatin organisation around the integration sites led to substantially different expression levels of the transgenic cassette in the neural tissues of these two mouse lines. In the spinal cord of 9-week old hemizygous L-FUS[1-359] mice weak predominantly cytoplasmic staining with antibody specifically recognising human FUS protein was detected in motor neurons (Suppl. Fig. 2a). Quantitative RT-PCR revealed that the level of tr-hFUS mRNA in the spinal cord of 9-week old L-FUS[1-359] mice is 20 times lower than in the spinal cord of presymptomatic S-FUS[1-359] of the same age and maintained at the same CD1 genetic background (Fig. 1a). Consistently, Western blot analysis with human FUS specific antibody revealed substantially lower level of tr-hFUS protein in the spinal cords of L-FUS[1-359] mice when compared to S-FUS[1-359] mice (Fig. 1b, c). Accumulation of tr-hFUS in the cytoplasm of spinal motor neurons of hemizygous S-FUS[1-359] mice leads to its aggregation with formation of large cytoplasmic inclusions at the onset of clinical signs, i.e. at the age of 13 to 20 weeks (Fig. 1e and Ref. 13). However, tr-hFUS remains diffusely distributed in the cytoplasm and is also present in the nucleus of spinal motor neurons of ageing homozygous L-FUS[1-359] mice (Fig. 1e). The latter observation is consistent with only partial cytoplasmic mislocalisation of FUS observed in neurons of mice that express low levels of human FUS variants carrying point mutations in the NLS or completely lacking this domain [38-40]. No obvious changes in localisation of endogenous mouse FUS in motor neurons (Suppl. Fig. 2b) or the level of expression of this protein in the spinal cord (Fig. 1b, d) of L-FUS[1-359] mice compared to their wild type littermates were observed (p>0.05).

The low level of tr-hFUS expression is not sufficient for triggering obvious disease phenotypes even in homozygous L-FUS[1-359] mice – no changes in performance of these animals in the accelerating rotarod and inverted grid tests, and their locomotor activity in novel non-anxiogenic environment were found (Fig. 2). The lifespan of these animals was not different from the lifespan of their littermates (Suppl. Fig. 3) and even in ageing 18-month old animals no sign of spinal motor neuron loss was noted (Fig. 3).

9

Changes of spinal cord transcriptomes caused by expression of C-terminally truncated human FUS

To assess whether a low level of expression of a pathogenic variant of FUS affects gene expression in the spinal cord of transgenic mice we employed RNA sequencing to compare spinal cord transcriptomes of four 9-week old hemizygous L-FUS[1-359] mice and four of their wild type littermates. As a result of deep-sequencing, we obtained from 13 to 17 million reads for each library produced from an individual mouse thoracic spinal cord.

Our analysis of sequencing data revealed 272 DEGs with the false discovery rate (FDR) < 0.05 that clearly distinguish gene expression profiles of L-FUS[1-359] and WT mice on multidimensional scaling (Fig. 4a, b, Supplementary table 1). To validate results of RNA sequencing analysis we carried out real-time qRT-PCR analysis of two of three most upregulated (*Mpz* and *Prx*) and two of three most downregulated (*Gpr179* and *FGGY*) transcripts, as well as one randomly selected transcript that did not show differential expression (*PPARA*). Results of qRT-PCR analysis appeared consistent with results of RNA sequencing analysis (Suppl. Fig. 4). In order to discriminate DEGs that are intrinsic for neurons and glial cells we used the published datasets of purified microglia and laser-microdissected ventral horns of mice spinal cord applying a 5-fold change cutoff of gene expression levels [30,31]. As a result, we obtained 200 genes (> 73% of DEG, FDR < 0.05) that are more specific for spinal cord neurons than for microglial cells and only 7 genes (< 3% of DEG, FDR < 0.05) that are intrinsic to microglia (Fig. 4c). Remaining 65 genes (~ 24% of DEG) were shared by both types of cells (Fig. 4c). Thus, the observed gene expression changes in the spinal cord of L-FUS[1-359] mice take place predominantly in neurons.

The gene ontology (GO) enrichment analysis of genes differentially expressed in L-FUS[1-359] spinal cord revealed several groups of genes encoding proteins involved in biological processes important for normal function of the nervous system. The largest group, in terms of percentage of involved genes to a number of total genes in a group, includes upregulated genes for claudin 19 (*Cldn19*), tenascin XB (*Tnxb*), stabilin 1 (*Stab1*), basal cell adhesion molecule (*Bcam*), collagens

(*Col14a1*, *Col15a1*, *Col18a1*), collagen-binding proteins (*Tgfbi, Nid2*), semaphorins (*SEMA3G, SEMA4B* and *SEMA5A*) and several other genes involved in cell adhesion and organisation of extracellular matrix (green and yellow groups in Fig. 4d).

A significant number of genes involved in negative regulation of neuronal differentiation is also upregulated in the spinal cord of L-FUS[1-359] mice, however genes encoding transcriptional factor SOX2 and OLIG2 that play a central role in neuronal cell differentiation and spinal motor neuron specification, respectively are significantly downregulated (red group in Fig. 4d).

The circadian rhythm related genes are highly represented between DEGs – all three paralogs of the "core" clock activator gene *period – Per1*, *Per2* and *Per3,* genes encoding the circadian transcriptional repressors CIART and BHLHE40, and transcription factor DBP are upregulated in the spinal cord neurons of L-FUS[1-359] mice, whereas genes encoding transcriptional activators ARNTL and NPAS2, and transcriptional repressor NFIL3 are significantly downregulated (blue group in Fig. 4d).

Expression of genes encoding the proteins assisting protein folding in the endoplasmic reticulum such as chaperones and chaperonins *Hspa2*, *Hsp90b1*, *Dnajb11*, *Cryab*, *Mkks*, protein disulfide isomerases *Pdia4*, *Pdia6* as well as calreticulin (*Calr*) are decreased in the spinal cord of L-FUS[1-359] mice (lilac group in Fig. 4d).

We also observed the altered expression levels of genes encoding several actin-binding proteins (*Gas2l3*, *Flna*, *Coro6*, *Diaph a*nd *Ermn)*, a myosin (*Myh11*) and members of dynein family (*Dnah5, Dync1i1, Dynll1, Dnaic1, Dnah8)* (Fig. 4d, Supplementary table 1).

Discussion

Familial forms of ALS with mutations in the *FUS* gene are characterised by a relatively early onset and a fast progression of the disease (reviewed in Refs. 4,12,18). Even so, despite a constant production of a pathogenic variant of FUS, spinal motor neurons of FUS-ALS patients function normally for decades (average age of FUS-ALS onset is ~43 years [12]). This suggest that by

activating certain compensatory mechanisms neurons are able to neutralise toxicity of pathogenic variants of FUS until their accumulation reaches a critical threshold, particularly when it is coupled with mislocalisation, or the cell defence machinery is compromised by an external or internal stress. Mouse models of FUS-ALS display the same pattern of the disease progression: animals are presymptomatic for a long period of their postnatal life (3-12 months, depending of the introduced genetic modification), but after manifestation of the first clinical signs of pathology the disease progresses rapidly, mice develop severe motor dysfunction and die within several days (reviewed in Refs. 10-12). However, this phenotype is typical for mouse lines expressing high levels or/and highly pathogenic variants of human FUS. In several recently produced models a problem of artificially high levels of endogenous FUS expression was solved by using knock-in techniques [40,41], a relevant promoter [38] or selection of mouse lines with an appropriate level of transgene expression [42]. Although the common trend for later onset and longer disease duration was observed in these studies, each model displayed distinct behavioural and histopathological characteristics.

Expression of human FUS carrying point mutations in the NLS (either FUS[R521C] or FUS[R521H]) at the level comparable with that of the endogenous mouse FUS caused progressive motor dysfunction due to the deficiency in muscle innervation, loss of spinal motor neurons and their axons in the absence of FUS mislocalisation to the cytoplasm and formation of inclusions [38]. In the spinal cord of ageing symptomatic transgenic mice lacking endogenous mouse FUS the most prominent changes revealed by RNA sequencing analysis were downregulation of mRNA encoding proteins involved in protein synthesis and synaptic function, and upregulation of mRNA encoding chaperons and proteins involved in RNA metabolism although no changes in a global splicing pattern were observed [38].

In another recent study, human FUS lacking NLS was expressed under control of Thy-1 promoter in the nervous system of transgenic mice at the level of around 80% of the endogenous mouse FUS expression [42]. These mice developed a progressive motor impairment from the age of

12 weeks and their lifespan was significantly decreased. Cytoplasmic accumulation of deltaNLS FUS and formation of large inclusions were observed in upper motor neurons before the onset of clinical signs, but in the spinal motor neurons, only at the age of one year. Neuroinflammation and the loss of upper but not lower motor neurons was prominent in one-year old mice. At the same age, cerebral cortex transcriptome analysis identified a large number of genes up- and downregulated as the consequence of deltaNLS FUS expression but again, no notable changes of splicing pattern were found [42].

To reveal changes in gene expression taking place in the nervous system of mice carrying ALS-related modifications of FUS before they develop any clinical signs of pathology, Humphrey et al. [43] compared results of transcriptomic analysis of neural tissues from homozygous late-gestation embryo (E17.5 spinal cord or E18.5 brain) of two independent FUS knock-in mouse lines expressing FUS lacking NLS [40,41,44]. These data were compared to data obtained in the same studies by RNA sequencing analysis of similar neural samples of FUS knockout mice, which allowed discrimination of changes caused by the loss of FUS function and the gain of function by FUS lacking NLS and therefore partially mislocalised to the cytoplasm. A substantial overlap of DEGs in knockout and knock-in models suggested that at the presymptomatic stage the effects of the substitution of endogenous FUS to deltaNLS FUS on gene expression is due to the loss of function, i.e. an inadequate functionality of deltaNLS FUS that remains nuclear in neurons of studied mice. Between identified DEGs, the most prominent changes common for knockout and knock-in, as well as specific for homozygous knock-in mice were upregulation for those that code for proteins involved in RNA metabolism and downregulation of those that code for proteins involved in synaptic transmission [43]. This is consistent with considerable changes of splicing pattern identified in these studies and the growing body of evidence that defects of synaptic transmission, particularly in the neuromuscular junctions, are very early events in the development of pathology in animal models and FUS-ALS.

In our previously described model, S-FUS[1-359] transgenic mouse line, expression of highly aggregation-prone and pathogenic tr-hFUS protein caused a relatively early onset (at the age of ~13-18 weeks) and severe and fast progressing motor pathology, despite its level in neural tissues at presymptomatic stage being lower than the level of endogenous FUS [13]. In the spinal cord of mice of a new L-FUS[1-359] transgenic line described here, the level of Thy-1 promoter-driven expression of the same tr-hFUS protein was even lower than in the spinal cord of S-FUS[1-359] transgenic line. Consequently, these animals did not develop motor deficiency, had the same lifespan as their wild type littermates and were phenotypically indistinguishable from them. We suggested that when the abundance of this pathogenic protein is low, neurons are able to neutralise its damaging effects by adjusting certain intracellular mechanisms and pathways. In an attempt to identify these adjustments, we analysed changes in the spinal cord transcriptomes of young adult L-FUS[1-359] mice. Importantly, in this model no pathology-driven changes in gene expression take place and therefore, any observed expression changes should be linked to adaptation to production of tr-hFUS.

Substantial changes in the gene expression profile (272 of DEGs, FDR < 0.05) were found in the spinal cord of L-FUS[1-359] mice although for most DEGs fold changes were relatively subtle (average log2 fold change for upregulated DEGs was 0.8 and for downregulated DEGs – 0.6). The majority of identified DEGs represented neuron-specific genes or genes expressed in neurons and other types of cells, suggesting that gene expression changes is direct effect of tr-hFUS expression in neurons. C-terminally truncated tr-hFUS lacks the ability to bind target RNAs and be recruited into physiological RNA granules due to the deletion of main RNA-binding motifs, namely RGG, Zn-finger and a part of RRM [45-49], and therefore cannot directly affect RNA metabolism. Therefore, it was not surprising that mRNA encoding proteins involved in RNA metabolism were barely represented in L-FUS[1-359] DEGs in contrast to high representation of these mRNA in DEGs identified in studies of mouse lines expressing low levels of FUS variants with intact RNA-binding domains [38,40,43].

Between gene clusters identified in our analysis of RNA sequencing data, the most prominent were upregulated expression of genes encoding proteins involved in cell adhesion and structural organisation of the extracellular matrix. These proteins function as regulators of cell migration, axon guidance and structuring of synaptic contacts, and therefore play important roles in the assembly and function of neuronal networks. Despite being predominantly a nuclear protein, FUS has important functions in axonal and synaptic processes [50-55]. It is feasible that the presence of tr-hFUS in these neuronal compartments might have a dominant negative effect on these functions of the endogenous FUS. Increased production of proteins that can provide additional stabilisation of neuronal networks might be a compensatory mechanism that efficiently prevents potentially damaging effects of tr-hFUS. Consistently, we also found modified expression of several genes encoding proteins involved in organisation of actin cytoskeleton and motor proteins, indicating that the remodeling of cytoskeletal structures and cellular transport take place in the spinal cord of L-FUS[1-359] mice.

Unexpectedly, expression of tr-hFUS protein did not prompt upregulation of genes encoding proteins involved in protein folding. Moreover, several chaperone or chaperonine encoding genes appeared to be downregulated in the spinal cord of L-FUS[1-359] mice, although only slightly. It is feasible that because at this level of expression tr-hFUS does not reach a concentration threshold required for its efficient aggregation, activation of protein folding mechanisms is not required.

Further detailed studies are required to explain prominent changes in expression of several genes encoding proteins involved in circadian rhythms because of different directions of these changes and complex feedback loops regulating expression of these genes.

Overall, we found that the pattern of gene expression changes in the spinal cord of asymptomatic mice expressing a low level of cytoplasmically mislocalised, aggregation-prone and deficient in RNA binding variant of human FUS, is substantially different from patterns previously reported for either symptomatic or presymptomatic mice expressing various human FUS variants at similarly low levels. Observed changes may reflect efficient adaptation of spinal motor neurons to expression

of a potentially pathogenic protein and therefore, molecular mechanisms and pathways affected by these changes might be considered as valid targets for preventive therapy of FUS-ALS.

Data availability statement

Data generated during this study are included in this published article and its Supplementary Information files and the RNA sequencing data were deposited in Gene Expression Omnibus (GEO) under the number GSE130604.

Ethical approval

All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted (The Bioethics committee of Institute of Physiologically Active Compounds, Russian Academy of Sciences approval No. 20 dated 23.06.2017). All animal work was carried out in accordance with this approval and the Rules of Good Laboratory Practice in Russian Federation (2016).

References

1. Deng HX, Zhai H, Bigio EH et al (2010) FUS-immunoreactive inclusions are a common feature in sporadic and non-SOD1 familial amyotrophic lateral sclerosis. Ann. Neurol. 67(6): 739-748. doi: 10.1002/ana.22051

2. Al-Chalabi A, Jones A, Troakes C, King A, Al-Sarraj S, van den Berg LH (2012) The genetics and neuropathology of amyotrophic lateral sclerosis. Acta Neuropathol. 124(3): 339-352. doi: 10.1007/s00401-012-1022-4

3. Polymenidou M, Lagier-Tourenne C, Hutt KR, Bennett CF, Cleveland DW, Yeo GW (2012) Misregulated RNA processing in amyotrophic lateral sclerosis. Brain Res. 1462: 3-15. doi: 10.1016/j.brainres.2012.02.059

4. Deng H, Gao K, Jankovic J (2014) The role of FUS gene variants in neurodegenerative diseases. Nat. Rev. Neurol. 10(6): 337-348. doi: 10.1038/nrneurol.2014.78

5. Mackenzie IR, Munoz DG, Kusaka H, Yokota O, Ishihara K, Roeber S, Kretzschmar HA, Cairns NJ, Neumann M (2011) Distinct pathological subtypes of FTLD-FUS. Acta Neuropathol. 121(2): 207-218. doi: 10.1007/s00401-010-0764-0

6. Mackenzie IR, Rademakers R, Neumann M (2010) TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. Lancet Neurol. 9(10): 995-1007. doi: 10.1016/S1474-4422(10)70195-2

7. Neumann M, Rademakers R, Roeber S, Baker M, Kretzschmar HA, Mackenzie IR (2009) A new subtype of frontotemporal lobar degeneration with FUS pathology. Brain 132(Pt 11): 2922-2931. doi: 10.1093/brain/awp214

8. Neumann M, Roeber S, Kretzschmar HA, Rademakers R, Baker M, Mackenzie IR (2009) Abundant FUS-immunoreactive pathology in neuronal intermediate filament inclusion disease. Acta Neuropathol. 118(5): 605-616. doi: 10.1007/s00401-009-0581-5

9. Snowden JS, Hu Q, Rollinson S (2011) The most common type of FTLD-FUS (aFTLD-U) is associated with a distinct clinical form of frontotemporal dementia but is not related to mutations in the FUS gene. Acta Neuropathol. 122(1): 99-110. doi: 10.1007/s00401-011-0816-0

10. Nolan M, Talbot K, Ansorge O (2016) Pathogenesis of FUS-associated ALS and FTD: insights from rodent models. Acta Neuropathol. Commun. 4(1): 99. doi: 10.1186/s40478-016-0358-8

11. Van Damme P, Robberecht W, Van Den Bosch L (2017) Modelling amyotrophic lateral sclerosis: progress and possibilities. Dis. Models Mech. 10(5): 537-549. doi: 10.1242/dmm.029058

12. Shang Y, Huang EJ (2016) Mechanisms of FUS mutations in familial amyotrophic lateral sclerosis. Brain Res. 1647: 65-78. doi: 10.1016/j.brainres.2016.03.036

13. Shelkovnikova TA, Peters OM, Deykin AV et al (2013) Fused in Sarcoma (FUS) protein lacking nuclear localization signal (NLS) and major RNA binding motifs triggers proteinopathy and severe motor phenotype in transgenic mice. J. Biol. Chem. 288(35): 25266-25274. doi: 10.1074/jbc.M113.492017

14. DeJesus-Hernandez M, Kocerha J, Finch N (2010) De novo truncating FUS gene mutation as a cause of sporadic amyotrophic lateral sclerosis. Hum. Mutat. 31(5): E1377-1389. doi: 10.1002/humu.21241

15. Bosco DA, Lemay N, Ko HK, Zhou H, Burke C, Kwiatkowski TJ Jr, Sapp P, McKenna-Yasek D, Brown RH Jr, Hayward LJ (2010) Mutant FUS proteins that cause amyotrophic lateral sclerosis incorporate into stress granules. Hum. Mol. Genet. 19(21): 4160-4175. doi: 10.1093/hmg/ddq335

16. Belzil VV, Langlais JS, Daoud H, Dion PA, Brais B, Rouleau GA (2012) Novel FUS deletion in a patient with juvenile amyotrophic lateral sclerosis. Arch. Neurol. 69(5): 653-656. doi: 10.1001/archneurol.2011.2499

17. Waibel S, Neumann M, Rosenbohm A, Birve A, Volk AE, Weishaupt JH, Meyer T, Müller U, Andersen PM, Ludolph AC (2013) Truncating mutations in FUS/TLS give rise to a more aggressive ALS-phenotype than missense mutations: a clinico-genetic study in Germany. Eur. J. Neurol. 20(3): 540-546. doi: 10.1111/ene.12031

18. Lattante S, Rouleau GA, Kabashi E (2013) TARDBP and FUS mutations associated with amyotrophic lateral sclerosis: Summary and update. Hum. Mutat. 34(6): 812-826. doi: 10.1002/humu.22319

19. Robinson HK, Deykin AV, Bronovitsky EV et al (2015) Early lethality and neuronal proteinopathy in mice expressing cytoplasm-targeted FUS that lacks the RNA recognition motif. Amyotroph. Lateral scler. Frontotemporal Degener. 16(5-6): 402-409. doi: 10.3109/21678421.2015.1040994

20.   Ninkina N, Peters O, Millership S, Salem H, van der Putten H, Buchman VL (2009)

Gamma-synucleinopathy: neurodegeneration associated with overexpression of the mouse

protein. Hum. Mol. Genet. 18(10): 1779-1794. doi: 10.1093/hmg/ddp090

21.   Connor-Robson N, Peters OM, Millership S, Ninkina, N, Buchman VL (2016)

Combinational losses of synucleins reveal their differential requirements for compensating

age-dependent alterations in motor behavior and dopamine metabolism. Neurobiol. Aging

46: 107-112. doi: 10.1016/j.neurobiolaging.2016.06.020

22.   Peters OM, Millership S, Shelkovnikova TA, Soto I, Keeling L, Hann A, Marsh-Armstrong

N, Buchman VL, Ninkina N (2012) Selective pattern of motor system damage in gamma-

synuclein transgenic mice mirrors the respective pathology in amyotrophic lateral sclerosis.

Neurobiol. Dis. 48(1): 124-131. doi: 10.1016/j.nbd.2012.06.016

23.   Millership S, Ninkina N, Guschina IA, Norton J, Brambilla R, Oort PJ, Adams SH, Dennis

RJ, Voshol PJ, Rochford JJ, Buchman VL (2012) Increased lipolysis and altered lipid

homeostasis protect gamma-synuclein-null mutant mice from diet-induced obesity. Proc.

Natl. Acad. Sci. U.S.A. 109(51): 20943-20948. doi: 10.1073/pnas.1210022110

24.   Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina

sequence data. Bioinformatics 30(15): 2114-2120. doi: 10.1093/bioinformatics/btu170

25.   Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M,

Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29(1): 15-

21. doi: 10.1093/bioinformatics/bts635

26.   Krasnov GS, Dmitriev AA, Kudryavtseva AV et al (2015) PPLine: An automated pipeline

for SNP, SAP, and splice variant detection in the context of proteogenomics. J. Proteome.

Res. 14(9): 3729-3737. doi: 10.1021/acs.jproteome.5b00490

27.   Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for

assigning sequence reads to genomic features. Bioinformatics 30(7): 923-930. doi:

10.1093/bioinformatics/btt656

28.   Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26(1): 139-140. doi: 10.1093/bioinformatics/btp616

29.   Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43(7): e47. doi: 10.1093/nar/gkv007

30.   Bandyopadhyay U, Cotney J, Nagy M, Oh S, Leng J, Mahajan M, Mane S, Fenton WA, Noonan JP, Horwich AL (2013) RNA-Seq profiling of spinal cord motor neurons from a presymptomatic SOD1 ALS mouse. PLoS One 8(1): e53575. doi: 10.1371/journal.pone.0053575

31.   Chiu IM, Morimoto ET, Goodarzi H, Liao JT, O'Keeffe S, Phatnani HP, Muratet M, Carroll MC, Levy S, Tavazoie S, Myers RM, Maniatis T (2013) A neurodegeneration-specific gene-expression signature of acutely isolated microglia from an amyotrophic lateral sclerosis mouse model. Cell Rep. 4(2): 385-401. doi: 10.1016/j.celrep.2013.06.018

32.   Funikov SY, Rezvykh AP, Mazin PV, Morozov AV, Maltsev AV, Chicheva MM, Vikhareva EA, Evgen'ev MB, Ustyugov AA (2018) FUS(1-359) transgenic mice as a model of ALS: pathophysiological and molecular aspects of the proteinopathy. Neurogenetics 19(3): 189-204. doi: 10.1007/s10048-018-0553-9

33.   Fresno C, Fernandez EA (2013) RDAVIDWebService: a versatile R interface to DAVID. Bioinformatics 29(21): 2810-2811. doi: 10.1093/bioinformatics/btt487

34.   Supek F, Bosnjak M, Skunca N, Smuc T (2011) REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One 6(7): e21800. doi: 10.1371/journal.pone.0021800

35.   R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org. Accessed 16 April 2019

36.   Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New
       York

37.   Walter W, Sanchez-Cabo F, Ricote M (2015) GOplot: an R package for visually combining
       expression data with functional analysis. Bioinformatics 31(17): 2912-2914. doi:
       10.1093/bioinformatics/btv300

38.   Lopez-Erauskin J, Tadokoro T, Baughn MW et al (2018) ALS/FTD-linked mutation in FUS
       suppresses intra-axonal protein synthesis and drives disease without nuclear loss-of-function
       of FUS. Neuron 100(4): 816-830. doi: 10.1016/j.neuron.2018.09.044

39.   Sharma A, Lyashchenko AK, Lu L, Nasrabady SE, Elmaleh M, Mendelsohn M, Nemes A,
       Tapia JC, Mentis GZ, Shneider NA (2016) ALS-associated mutant FUS induces selective
       motor neuron degeneration through toxic gain of function. Nat. Commun. 7: 10465. doi:
       10.1038/ncomms10465

40.   Devoy A, Kalmar B, Stewart M et al (2017) Humanized mutant FUS drives progressive
       motor neuron degeneration without aggregation in 'FUSDelta14' knockin mice. Brain
       140(11): 2797-2805. doi: 10.1093/brain/awx248

41.   Scekic-Zahirovic J, Sendscheid O, El Oussini H et al (2016) Toxic gain of function from
       mutant FUS protein is crucial to trigger cell autonomous motor neuron loss. EMBO J.
       35(10): 1077-1097. doi: 10.15252/embj.201592559

42.   Shiihashi G, Ito D, Yagi T, Nihei Y, Ebine T, Suzuki N (2016) Mislocated FUS is sufficient
       for gain-of-toxic-function amyotrophic lateral sclerosis phenotypes in mice. Brain 139(Pt
       9):2380-2394. doi: 10.1093/brain/aww161

43.   Humphrey J, Birsa N, Milioto C et al (2019) FUS ALS-causative mutations impact FUS
       autoregulation and the processing of RNA-binding proteins through intron retention.
       bioRxiv 567735; doi: https://doi.org/10.1101/567735

44. Scekic-Zahirovic J, Oussini HE, Mersmann S et al (2017) Motor neuron intrinsic and extrinsic mechanisms contribute to the pathogenesis of FUS-associated amyotrophic lateral sclerosis. Acta Neuropathol. 133(6): 887-906. doi: 10.1007/s00401-017-1687-9

45. Lerga A, Hallier M, Delva L, Orvain C, Gallais I, Marie J, Moreau-Gachelin F (2001) Identification of an RNA binding specificity for the potential splicing factor TLS. J. Biol. Chem. 276(9): 6807-6816. doi: 10.1074/jbc.M008304200

46. Bentmann E, Neumann M, Tahirovic S, Rodde R, Dormann D, Haass C (2012) Requirements for stress granule recruitment of fused in sarcoma (FUS) and TAR DNA-binding protein of 43 kDa (TDP-43). J. Biol. Chem. 287(27): 23079-23094. doi: 10.1074/jbc.M111.328757

47. Liu X, Niu C, Ren J, Zhang J, Xie X, Zhu H, Feng W, Gong W (2013) The RRM domain of human fused in sarcoma protein reveals a non-canonical nucleic acid binding site. Biochim. Biophys. Acta, 1832(2): 375-385. doi: 10.1016/j.bbadis.2012.11.012

48. Shelkovnikova TA, Robinson HK, Connor-Robson N, Buchman VL (2013) Recruitment into stress granules prevents irreversible aggregation of FUS protein mislocalized to the cytoplasm. Cell cycle 12(19): 3194-3202. doi: 10.4161/cc.26241

49. Shelkovnikova TA, Robinson HK, Southcombe JA, Ninkina N, Buchman VL (2014) Multistep process of FUS aggregation in the cell cytoplasm involves RNA-dependent and RNA-independent mechanisms. Hum. Mol. Genet. 23(19): 5211-5226. doi: 10.1093/hmg/ddu243

50. Belly A, Moreau-Gachelin F, Sadoul R, Goldberg Y (2005) Delocalization of the multifunctional RNA splicing factor TLS/FUS in hippocampal neurones: exclusion from the nucleus and accumulation in dendritic granules and spine heads. Neurosci. Lett. 379(3): 152-157. doi: 10.1016/j.neulet.2004.12.071

51. Fujii R, Okabe S, Urushido T, Inoue K, Yoshimura A, Tachibana T, Nishikawa T, Hicks GG, Takumi T (2005) The RNA binding protein TLS is translocated to dendritic spines by

mGluR5 activation and regulates spine morphology. Curr. Biol. 15(6): 587-593. doi: 10.1016/j.cub.2005.01.058

52.   Schoen M, Reichel JM, Demestre M, Putz S, Deshpande D, Proepper C, Liebau S, Schmeisser MJ, Ludolph AC, Michaelis J, Boeckers TM (2015) Super-Resolution Microscopy Reveals Presynaptic Localization of the ALS/FTD Related Protein FUS in Hippocampal Neurons. Front. Cell. Neurosci. 9: 496. doi: 10.3389/fncel.2015.00496

53.   So E, Mitchell JC, Memmi C, Chennell G, Vizcay-Barrena G, Allison L, Shaw CE, Vance C (2018) Mitochondrial abnormalities and disruption of the neuromuscular junction precede the clinical phenotype and motor neuron loss in hFUSWT transgenic mice. Hum. Mol. Genet. 27(3): 463-474. doi: 10.1093/hmg/ddx415

54.   Udagawa T, Fujioka Y, Tanaka M et al (2015) FUS regulates AMPA receptor function and FTLD/ALS-associated behaviour via GluA1 mRNA stabilization. Nat. Commun. 6: 7098. doi: 10.1038/ncomms8098

55.   Yasuda K, Zhang H, Loiselle D, Haystead T, Macara IG, Mili S (2013) The RNA-binding protein Fus directs translation of localized mRNAs in APC-RNP granules. J. Cell Biol. 203(5): 737-746. doi: 10.1083/jcb.201306058

Author contributions statement

Natalia Ninkina and Vladimir L. Buchman conceived the project, designed the study, carried out certain experiments, interpreted experimental data and wrote the manuscript. Ekaterina A. Lysikova acquired and analysed molecular biology and behaviour data, prepared figures and wrote drafts of manuscript sections. Sergei Funikov and Alexander P. Rezvykh performed bioinformatic analysis of RNAseq data, drafted results and prepared corresponding table and figures. Kirill D. Chaprov and Michail S. Kukharsky performed immunohistochemical and Western blot analyses. Aleksey Ustyugov and Alexey V. Deykin produced transgenic mice and generated survival data. Ilya A. Flyamer and Shelagh Boyle defined chromosomal localisation of transgenic cassettes in both

hFUS[1- 359] lines. Sergey O. Bachurin interpreted experimental data. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Figure legends

Figure 1. Expression of tr-hFUS in the thoracic spinal cord of L-FUS[1-359] and S-FUS[1-359] mice. a) Relative mRNA levels quantified by real-time qRT-PCR with primers specific for human FUS mRNA. Total RNA samples from four 9-week old hemizygous male animals were analysed in triplicates (*$p<0.01$, Mann-Whitney U-test). b) Western blot analysis of total protein samples extracted from spinal cords of two hemizygous male L-FUS[1-359] , two hemizygous male S-FUS[1-359] and two wild type mice (all 9-week old) with antibodies that specifically recognise human FUS (tr-hFUS panel) or full-length mouse FUS (fl-mFUS) or neurofilament-L (NF-L). Anti-beta-actin antibody was used as a loading control. c, d) Relative levels of tr-hFUS and endogenous mouse FUS proteins quantified by Western blot analysis of total spinal cord proteins of four animals per genotype with antibody that specifically recognise human FUS (c) or mouse FUS (d) as illustrated in the panel c (*$p<0.05$, Mann-Whitney U-test). e) Immunohistochemical staining of transverse section through the spinal cord of 18-month old hemizygous male L-FUS[1-359] mouse and 4-month old early symptomatic hemizygous male S-FUS[1-359] with antibody specific for human FUS. Note accumulation of tr-hFUS in large cytoplasmic and small nuclear inclusions in an anterior horn motor neuron of the S-FUS[1-359] mouse and diffuse distribution of tr-hFUS in the cytoplasm and the nucleus of the L-FUS[1-359] mouse. Scale bars, 50 µm.

Figure 2. Animal performance in motor behaviour tests and their locomotor activity in novel non-anxiogenic environment. Bar charts show means±SEM of values obtained by testing one-year

old homozygous male L-FUS[1-359] and wild type (WT) mice (10-12 per genotype). a) The latency to fall from the accelerating rotarod. No significant difference between performance of two groups of mice were found (p>0.05, Mann-Whitney U-test). b) The latency to fall from the inverted grid. Note that all animals in each group successfully completed this 60-second test. c) Total distance covered by an animal over the 5-minute test in the activity camera. No significant difference between performance of two groups of mice were found (p>0.05, Mann-Whitney U-test).

Figure 3. No loss of spinal motor neurons in ageing L-FUS[1-359] mice. Nissl stained transverse section through the spinal cord of 18-month old hemizygous male L-FUS[1-359] mouse, its wild type littermate and 5-month old symptomatic hemizygous male S-FUS[1-359] mouse. Note that while few motor neurons remained in the anterior horn of S-FUS[1-359] mice (arrowheads) are either pale stained or shrinked, the anterior horn of ageing L-FUS[1-359] mice have the same complement and morphology of motor neurons as anterior horn of wild type animals. The bar chart shows means±SEM of the number of motor neurons in anterior horns per section (averaged from 6 section separated from each other by 40 µm, per animal) separately for thoracic and lumbar spinal cord of 6 L-FUS[1-359] and 6 wild type mice (p>0.05, Mann-Whitney U-test).

Figure 4. Gene expression profile changes in the spinal cord of 9-week old hemizygous L-FUS(1-359) mice compared to wild type littermates. a) Multidimensional scaling plot of RNA sequencing data for L-FUS[1-359] mice (L, in purple) and wild type mice (WT, in green). b) Volcano plot demonstrating an amount of DEGs in the spinal cord of L-FUS[1-359] mice compared to wild type mice. Only DEGs with statistically significant (FDR < 0.05) difference between studied groups are shown. c) Distribution of DEGs between neuron-specific and microglia-specific groups of genes. The sorting of genes was carried out as described in Materials and Methods. d) Gene ontology analysis of DEGs. The most enriched GO terms of biological processes are present (p < 0.05) and

sorted by their expression level (Log2FC, FDR < 0.05) in L-FUS[1-359] mice in comparison to WT

mice.