Stem Cell Research 27 (2018) 172–179

Contents lists available at ScienceDirect

Stem Cell Research

journal homepage: www.elsevier.com/locate/scr

Characterization of the lncRNA transcriptome in mESC-derived motor neurons: Implications for FUS-ALS

Silvia Biscarini [a], Davide Capauto [a], Giovanna Peruzzi [a], Lei Lu [b], Alessio Colantoni [c], Tiziana Santini [a], Neil A. Shneider [b], Elisa Caffarelli [d], Pietro Laneve [a,*,1], Irene Bozzoni [a,c,d,e,**]

[a] Center for Life Nano Science@Sapienza, Istituto Italiano di Tecnologia, Rome, Italy
[b] Department of Neurology, Center for Motor Neuron Biology and Disease, Columbia University, New York, NY, USA
[c] Department of Biology and Biotechnology, Sapienza University of Rome, Italy
[d] Institute of Molecular Biology and Pathology of CNR, Rome, Italy
[e] Institute Pasteur Fondazione Cenci-Bolognetti, Sapienza University of Rome, Italy

ARTICLE INFO

Article history:
Received 9 October 2017
Received in revised form 22 December 2017
Accepted 15 January 2018
Available online 31 January 2018

Keywords:
LncRNAs
Mouse ESCs
Motor neurons
Differentiation
FUS
ALS

ABSTRACT

Long non-coding RNAs (lncRNAs) are currently recognized as crucial players in nervous system development, function and pathology. In Amyotrophic Lateral Sclerosis (ALS), identification of causative mutations in FUS and TDP-43 or hexanucleotide repeat expansion in C9ORF72 point to the essential role of aberrant RNA metabolism in neurodegeneration. In this study, by taking advantage of an *in vitro* differentiation system generating mouse motor neurons (MNs) from embryonic stem cells, we identified and characterized the long non-coding transcriptome of MNs. Moreover, by using mutant mouse MNs carrying the equivalent of one of the most severe ALS-associated FUS alleles (P517L), we identified lncRNAs affected by this mutation. Comparative analysis with human MNs derived *in vitro* from induced pluripotent stem cells indicated that candidate lncRNAs are conserved between mouse and human. Our work provides a global view of the long non-coding transcriptome of MN, as a prerequisite toward the comprehension of the still poorly characterized non-coding side of MN physiopathology.

1. Material and methods

1.1. RNA-Seq and bioinformatics analysis

TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero treatment (Illumina) was used to obtain sequencing libraries from total RNA extracted from sorted GFP(+) MNs differentiated from three different mESC clones (C57BL/6J strain). The sequencing reaction, which produced 100 nucleotide long paired end reads, was performed on an Illumina Hiseq 2500 Sequencing system.

Trimmomatic software (Bolger et al., 2014) was used to remove adapter sequences and poor quality bases from raw reads; reads whose length after trimming was <30 nt were discarded. We also filtered out reads aligning to rRNAs, tRNAs, snRNAs, snoRNAs and other small non-coding species which resulted to be overrepresented according to FastQC software (available online at http://www.bioinformatics.babraham.ac.uk/projects/fastqc); this first alignment was performed using Bowtie 2 software (Langmead and Salzberg, 2012). Preprocessed reads were aligned to GRCm38 mouse genome and Ensembl transcriptome (Flicek et al., 2014) using TopHat2 (Kim et al., 2013) with parameters -i 50 –library-type fr-firststrand; to estimate mean and variance of inner distance distribution, which are required by TopHat2 software, we previously aligned reads to a non-redundant set of mRNA sequences derived from Ensembl 77 gene annotation using BWA software (Li and Durbin, 2010), then we calculated mean and variance of inner distance distribution from aligned read pairs whose inner distance was within interval [Q1 − 2(Q3 − Q1), Q3 + 2(Q3 − Q1)] (Q1 = first quartile, Q3 = third quartile). Reads mapping to mitochondrial genome were filtered off using Samtools software (Li et al., 2009). Cuffdiff 2 software (Differential analysis of gene regulation at transcript resolution with RNA-seq) was used to perform differential expression analysis.

Abbreviations: lncRNA, Long non-coding RNA; ALS, Amyotrophic Lateral Sclerosis; MN, Motor neuron; PCG, Protein-coding Gene.

* Correspondence to: P. Laneve, Center for Life Nano Science@Sapienza, Istituto Italiano di Tecnologia, Viale Regina Elena, 291, 00161 Rome, Italy.
** Correspondence to: I. Bozzoni, Department of Biology and Biotechnology, Sapienza University of Rome, Piazzale Aldo Moro, 5, 00185 Rome, Italy.
E-mail addresses: silvia.biscarini@uniroma1.it (S. Biscarini), davide.capauto@uniroma1.it (D. Capauto), giovanna.peruzzi@iit.it (G. Peruzzi), ll2807@columbia.edu (L. Lu), alessio.colantoni@uniroma1.it (A. Colantoni), tiziana.santini@iit.it (T. Santini), ns327@columbia.edu (N.A. Shneider), elisa.caffarelli@uniroma1.it (E. Caffarelli), pietro.laneve@iit.it, pietro.laneve@uniroma1.it (P. Laneve), irene.bozzoni@uniroma1.it (I. Bozzoni).
[1] Current address: Institute of Molecular Biology and Pathology of CNR, Rome, Italy.

https://doi.org/10.1016/j.scr.2018.01.037

Heatmap of differentially expressed lincRNAs was generated using heatmap3 R package (Zhao et al., 2014) from log2 transformed FPKM values.

1.2. Cell cultures

Spinal motor neurons (MNs) were differentiated from mESCs HB9::GFP Fus$^{+/+}$, Fus$^{-/-}$ or Fus $^{P517L/P517L}$ (C57BL/6J strain; one clone for each genetic background) as described in (Errichelli et al., 2017; Wichterle and Peljto, 2008), by culturing embryoid bodies (EBs) in ADFNK medium, complemented with B27 Supplement, Retinoic Acid (RA) and Smoothened Agonist (SAG).

1.3. Fluorescence-activated cell sorting (FACS)

The differentiated mixed population was dissociated by Papain Dissociation system (Worthington). Cells were resuspended in PBS without Ca++Mg++ complemented with 2.5% Horse Serum, 0.4% Glucose, DNAseI (50 μg/ml), containing 2% B27 Supplement and sorted for GFP expression using a FACSAriaIII (Becton Dickinson, BD Biosciences) equipped with a 488 nm laser and FACSDiva software (BD Biosciences version 6.1.3). The analysis was based on FlowJo software (Tree Star). An aliquot of each collected sample was evaluated for purity resulting >99%.

Highly purified MNs were plated on 0.01% poly-L-ornithine, 10 μg/ml natural mouse laminin (Sigma-Aldrich) coated dishes, in MN medium (Neurobasal medium, 2% horse serum, 1% B27, 1% Pen/Step, 0.25% 2-mercaptoethanol, 0.25% Glutamax, 0.025 mM L-glutamic acid) supplemented with 10 ng/ml BDNF, 10 ng/ml GDNF, 10 ng/ml CNTF, 10 ng/ml NT3 (ThermoFisher) and ROCK-inhibitor (20 μM) for the first 48 h.

1.4. RNA preparation and analysis

Total RNA from cells was extracted with the Quick RNA MiniPrep (Zymo Research) and retrotranscribed with SuperScript VILO (Life Technologies). The real-time qRT-PCR analysis was performed with SYBR Green Power-UP (Life Technologies) using the housekeeping gene Atp5o (ATP synthase, H+ transporting, mitochondrial F1 complex, O subunit) as internal control.

1.5. Cellular fractionation

At the end of differentiation, the mixed cell population was dissociated as for FACS analysis and plated on poly-L-ornithine/laminin-coated dishes. After three more days of culture, cells were washed with PBS without Ca++Mg++, 100 μl of buffer A (Tris 20 mM pH 8.0, NaCl 10 mM, MgCl2 3 mM, NP40 (IGEPAL) 0,10%, EDTA 0,2 mM, DTT 1 mM, Protease inhibitor cocktail 1×, Ribolock 1×) were added and cells were scraped, transferred to a 1,5 ml tube and incubated on ice for 5 min. Nuclei were pelleted 400 ×g 5 min at 4 °C. The supernatant, corresponding to the cytoplasmic fraction was stored on ice. The nuclear pellet was washed twice with 50 μl of buffer A. RNA was then extracted adding 600 μl of RNA Lysis buffer (QuickRNA Miniprep) to both fractions as described in "RNA preparation and analysis".

1.6. Coding potential calculation

Coding potential calculation was performed as recommended by CPC (http://cpc2.cbi.pku.edu.cn/) and CPAT (http://lilab.research.bcm.edu/cpat/index.php) online tools.

1.7. Statistical analysis

Results are expressed as means ± SD. Statistical differences were analyzed by two-tailed Student's *t*-test. A P-value < 0.05 was considered as statistically significant. *P < 0.05, **P < 0.01, ***P < 0.001.

1.8. Oligonucleotides

Oligonucleotide sequences used in this study are listed in Table S1.

2. Introduction

Long non-coding RNAs (lncRNAs) are a heterogeneous class of non proteinogenic molecules longer than 200 nt which, in the last decades, have been shown to play several roles in cellular biology both in physiological and pathological conditions (Ponting et al., 2009).

Current views describe lncRNAs as less expressed than mRNAs, but showing high tissue specificity (Derrien et al., 2012); in particular, among non-ubiquitous transcripts, 40% display brain-specific expression (Derrien et al., 2012), suggesting their relevant role in the nervous system. Despite deep sequencing technologies promoted the annotation process, uncovering an unexpected abundance of lncRNAs, very few species have been functionally characterized so far (Briggs et al., 2015). LncRNAs participate in various stages along the path from pluripotent to differentiated cells; indeed they may act in the exit from pluripotency (Guttman et al., 2011); guide neural fate choice driving transcription factor localization (Ng et al., 2013; Vance et al., 2014), regulate local translation at synapses (Zalfa et al., 2003; Zhong et al., 2009) or even influence neuronal excitability (Zhao et al., 2013). Some lncRNAs have also been implicated in neurodegeneration, such as BACE1AS, involved in amyloid plaques formation in Alzheimer disease (Faghihi et al., 2008) or NEAT1, essential for paraspeckles formation and over-expressed in Amyotrophic Lateral Sclerosis (ALS) (Nishimoto et al., 2013).

ALS is an incurable adult-onset neurodegenerative disease, which affects upper and lower motor neurons (MNs), and leads to paralysis and death in 3–5 years from diagnosis (Taylor et al., 2016). Several genetic alterations are associated with ALS (Renton et al., 2014), including causative mutations in FUS, TDP-43 and expansions in C9ORF72 point to the essential role of aberrant RNA metabolism in ALS pathogenesis (Ling et al., 2013). FUS is a multifunctional RNA/DNA binding protein accounting for 4% of familial ALS cases (Renton et al., 2014). FUS mutations frequently result in protein mis-localization from the nucleus to the cytoplasm leading to pathological mechanisms linked to its loss and/or gain-of-function in these cellular compartments (Vance et al., 2009); this is the case for aggressive human mutation FUS^{P525L} and its murine equivalent, Fus^{P517L}.

In spite of the high number of neural-specific lncRNAs characterized and the availability of several differentiation protocols for different neuronal cell types, the profiling of motor neuronal lncRNA transcriptome is still lacking; this has impaired a thorough understanding of the molecular mechanisms underlying ALS.

To address these issues we combined the efficient generation of *in vitro*-derived mouse MNs with RNA-seq analysis to reveal a MN-specific lncRNA signature. We identified several species showing MN-restricted expression, conservation in human MNs and altered expression in Fus depleted or mutated MNs.

3. Results

3.1. Identification of lncRNAs expressed in MN differentiation and maturation

As previously described, mESCs harboring a GFP reporter under the control of the MN-specific HB9 promoter can be efficiently differentiated *in vitro* and FAC-sorted as MNs (Fig. S1A–D) (Errichelli et al.,

2017; Wichterle and Peljto, 2008; Wichterle et al., 2002). RNA-seq data obtained from total RNA of GFP+ cells (GSE101097; Errichelli et al., 2017) were compared with those of publicly available RNA-seq experiments from undifferentiated mESCs (GSE43390; Williams et al., 2015). We found 9644 genes differentially expressed (average FPKM in at least one condition ≥ 1 and q-value < 0.05) between the two cell types (Table S2); out of these, 469 encoded for *bona fide* lncRNAs (Table S2 and Fig. 1A). The family of lncRNAs up-regulated in MNs derived from 270 *loci* (Table S2) and includes some species already known to play key roles in neurogenesis; among them, Miat (Aprea et al., 2013), Rmst (Ng et al., 2013), Hotairm1 (Lin et al., 2011), Meg3, Rian, and Mirg (Mo et al., 2015).

Up-regulated lncRNAs were ranked according to fold-change and filtered for FPKM > 4. The neural expression of the top 40 hits was

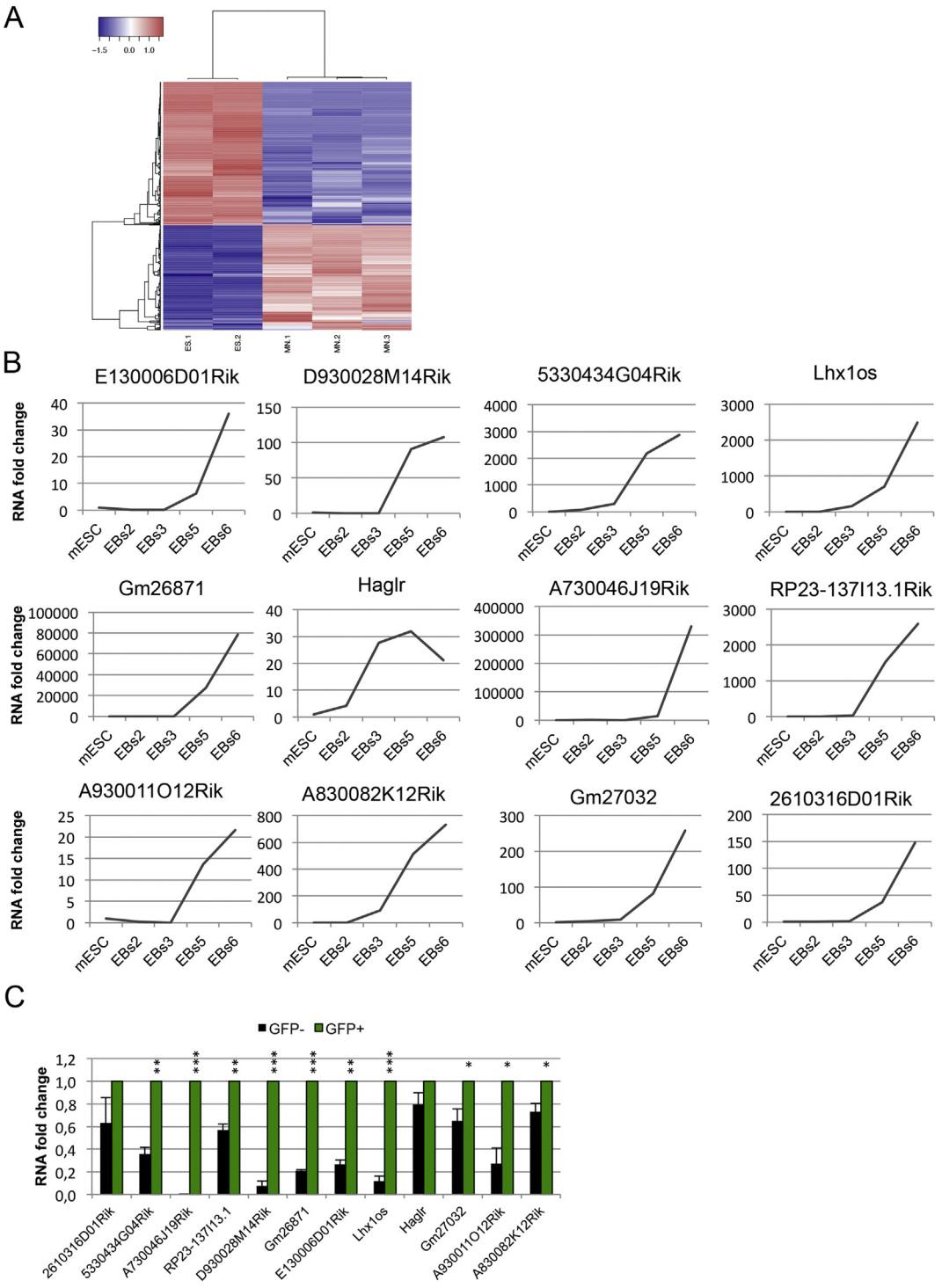

Fig. 1. Identification of lncRNAs expressed in MN differentiation. A. Heat-map of differentially expressed lncRNAs between mESCs and MNs (q-value < 0.05; average FPKM in at least one condition ≥ 1). B. qRT-PCR analysis of the 12 lncRNAs up-regulated during MN differentiation. The expression is relative to levels detected in mESCs, set as 1. Differentiation time points are indicated as days below the diagrams (EBsn: embryoid bodies day n). Transcript names are indicated above each graph. Normalization vs Atp5o expression levels. C. qRT-PCR analysis of the 12 lncRNAs in GFP+ (green bars) and GFP- (black bars) cell populations. The expression is relative to levels in GFP+ cells (MNs), set as 1, n = 3. Normalization vs Atp5o expression.

assessed by analyzing publicly available longRNA-seq data from 26 mouse tissues (Stamatoyannopoulos et al., 2012), resulting in the identification of 12 transcripts with neural-restricted expression (Fig. S2A and B). The 12 lncRNAs were strongly expressed in MNs (fold change > 10) with a FPKM value between 4 and 151 (Fig. S3A). The differential expression of the 12 lncRNAs was then validated through qRT-PCR on RNA from mESCs and GFP+ MNs immediately after sorting or after six additional days of maturation (Fig. S3B). For all the 12 lncRNAs we were able to confirm the up-regulation in MNs. Notably, three species (2610316D01Rik, 5330434G04Rik and Gm26871) maintained or even increased their expression upon further maturation, whereas other transcripts (Haglr and E130006D01Rik) showed a drastic drop in expression after six additional days of culture (Fig. S3B).

We analyzed the expression profile of the 12 lncRNAs during differentiation, in order to identify transcripts expressed when progenitor and post-mitotic MNs were specified (Fig. 1B). Most of the 12

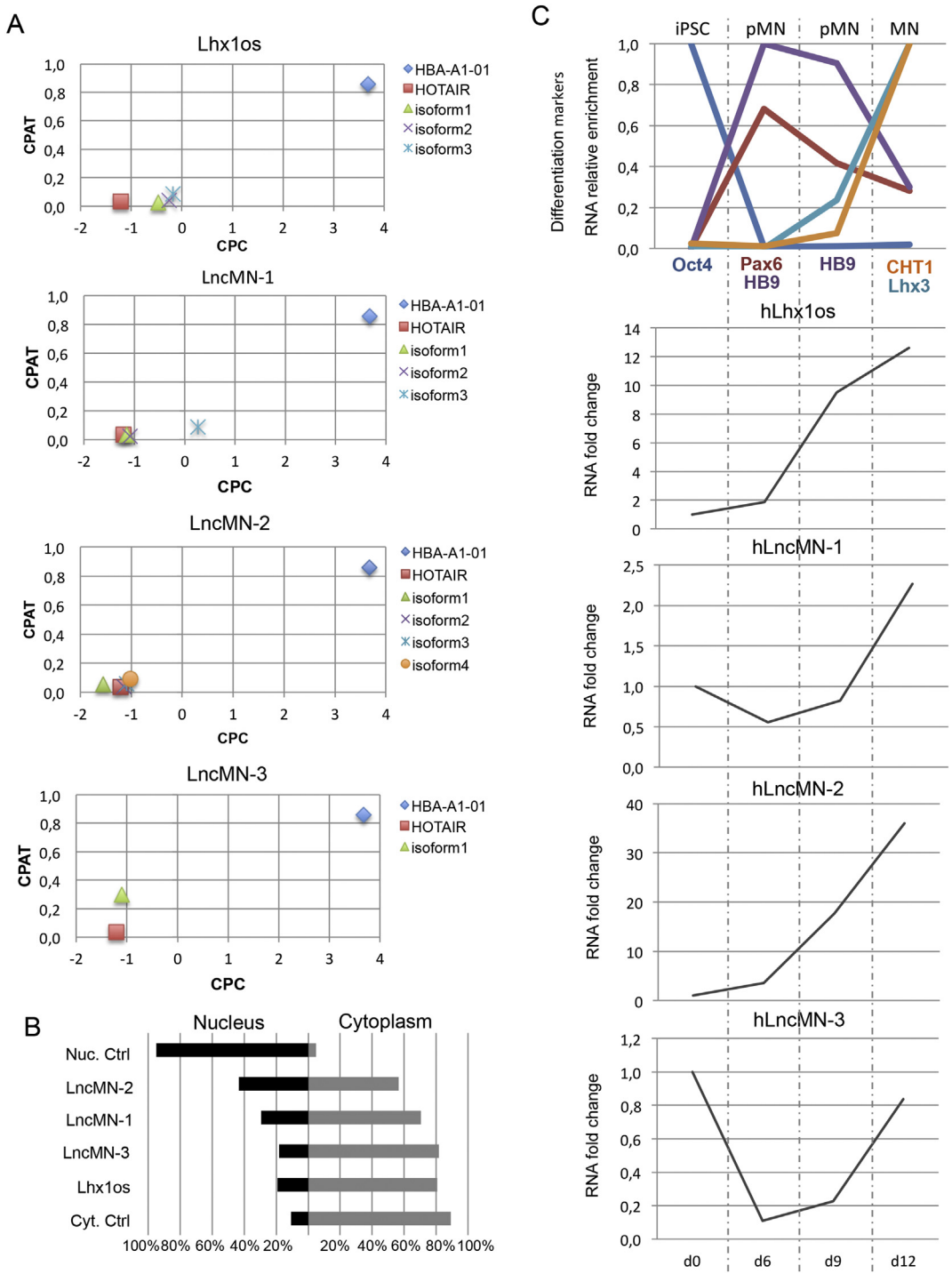

Fig. 2. Molecular characterization of selected lncRNAs. A. Codogeneity graph displaying, for each transcript variant of each gene analyzed (reported above each diagram), the coding score from Coding Potential Calculator (CPC, x-axis) and Coding Potential Assessment tool (CPAT, y-axis). Non-coding scores are <0 for CPC and <0,4 for CPAT. Hemoglobin alpha, adult chain 1 transcript 1 (HBA-A1-01) represents the coding transcript control, HOTAIR the non-coding transcript control. B. RNA sub-cellular localization analysis: transcript distribution is expressed as percentage; a nuclear control (sno55 RNA) and a cytoplasmic control (Gapdh mRNA) are also shown. C. Expression profile of selected lncRNAs during human MN differentiation from iPSCs. Top graph displays the expression of differentiation markers: Oct4 (stem cells), Pax6 (MN progenitors), HB9, CHT1 and Lhx3 (MNs). LncRNA expression is relative to expression in iPSCs, set as 1. Normalization vs Atp5o expression.

transcripts were initially expressed in the transition from neural progenitor (pN) to MN progenitor (pMN) stages, as defined by the expression of Pax6 and Olig2 or HB9 and ChAT respectively (Fig. S1C). Interestingly A730046J19Rik showed a marked up-regulation coincident with MNs specification (Fig. 1B).

MN-specificity was further tested by analyzing lncRNA expression levels in GFP+ *versus* GFP - cell populations. Several candidates showed a MN-restricted expression, as confirmed by their absence in the GFP- population (Fig. 1C), which mainly contains MN progenitors as well as V1, V2 and V3 interneurons (Fig. S1D).

We select for further analysis Lhx1os, 5330434G04Rik (renamed as LncMN-2) and A730046J19Rik (renamed as LncMN-3) for their MN-enriched expression while 2610316D01Rik (renamed as LncMN-1) because of its interesting genomic localization (see below).

3.2. Selected lncRNA molecular characterization

To characterize Lhx1os, lncMN-1, lncMN-2 and lncMN-3 we started by examining their genomic *loci* in order to determine i) the structure of the gene, ii) the number of annotated splice variants, iii) the degree of sequence conservation, iv) the presence of repetitive region in the exons and v) the function of surrounding genes (Fig. S4A–D).

Lhx1os shows three splicing variants, producing transcripts between 524 and 599 nt in length and composed of 4 or 5 exons (Fig. S4A). This gene is divergent to Lhx1, a transcript encoding a morphogenetic factor of the LIM family involved in lateral MN differentiation (Alaynick et al., 2011), head development and MN axon guidance (Hunter and Rhodes, 2005).

LncMN-1 has three splice variants giving rise to 664–924 nt and 7/9-exons transcripts. This lncRNA is divergent to the Pcdh10 gene, which codes for a protocadherin involved in motor neuronal cell adhesion (Fig. S4B) (Machado et al., 2014).

LncMN-2 is an intergenic lncRNA transcribed in 4 splicing variants composed of 4–5 exons with lengths ranging from 511 to 3545 nt (Fig. S4C).

LncMN-3 is a 3-exons, 4707 nt long, intergenic ncRNA (Fig. S4D), characterized by an exceptionally high sequence conservation in mammals (BLAT search identifies a non-continuous sequence of 1217 nt showing 87,4% base conservation; Fig. S5).

The non-coding nature of the four species was assessed through the Coding Potential Calculator (CPC) and Coding Potential Assessment Tool (CPAT). Considering Hemoglobin alpha, adult chain 1 transcript 1 (HBA-A1-01) and HOTAIR as coding and non-coding controls respectively, we could annotate the four hits as non-coding molecules (Fig. 2A). Concerning lncMN-3, despite the bioinformatics analysis points to its non-coding nature, the presence of an ORF of 267 nt may suggest this lncRNA is a member of the emerging class of bifunctional RNAs, which display independent functions for the transcript and the encoded protein (Andres-Pablo et al., 2017).

Further characterization of their sub-cellular localization indicated that, with the exception of lncMN-2, which is almost equally distributed, the three lncRNAs are 70–80% cytoplasmic (Fig. 2B).

Among our candidates, the only hit with high sequence conservation in human is lncMN-3 (Fig. S5); however, all four RNAs show synteny conservation in human (orthologue transcripts: RP11-445F12.1, RP11-9G1.3, MIR325HG, LINC00890). In order to assess the motor neuronal expression of human counterparts, we took advantage of an *in vitro* system to convert human iPSCs into MNs (De Santis et al., 2017; Errichelli et al., 2017; Hill et al., 2016). Similarly to the murine system, the human cell line contains a HB9::GFP cassette that enables the purification of MNs. Fig. 2C shows that three species increased their expression during human *in vitro* MN differentiation mimicking the situation observed in mouse. Also in this case, the expression peak paralleled the expression of markers of progenitor and post-mitotic MNs. LncMN-3 drops at MN progenitor stage and successively rises back to iPSC level (Fig. 2C).

3.3. LncRNA de-regulation in mouse FUS-ALS in vitro model

Considering that ALS is characterized by the specific loss of MNs and given that FUS mutations point to a relevant role of RNA dysmetabolism in neurodegeneration (Ling et al., 2013), we investigated whether the lncRNAs we identified showed any change in expression in an *in vitro* model of FUS-ALS. By analyzing RNA expression in differentiated $Fus^{P517L/P517L}$ MNs we found that, compared to $Fus^{+/+}$ MNs, Lhx1os was up-regulated (fold change 2,5), while lncMN-1 and lncMN-2 were down-regulated (fold change 0,5 and 0,8 respectively; Fig. 3A). No alteration was instead observed for lncMN-3 (Fig. 3A).

To discriminate whether Lhx1os, lncMN-1 and lncMN-2 altered expression in $Fus^{P517L/P517L}$ depended on Fus gain- or loss-of-function, we analyzed their levels in FUS knock-out MNs ($Fus^{-/-}$). Notably, in these cells the three lncRNAs exhibited the same alterations observed in $Fus^{P517L/P517L}$ MNs (Fig. 3B), indicating a loss-of-function mechanism. LncMN-3 expression was not analyzed in $Fus^{-/-}$ MN since it was not affected by FUS mutation.

Given that Lhx1os and lncMN-1 are divergent from two protein-coding genes (PCGs) involved in neuronal physiology, we investigated the behavior of the PCGs during differentiation and we observed a strong co-regulation of the two divergent transcripts in both mouse and

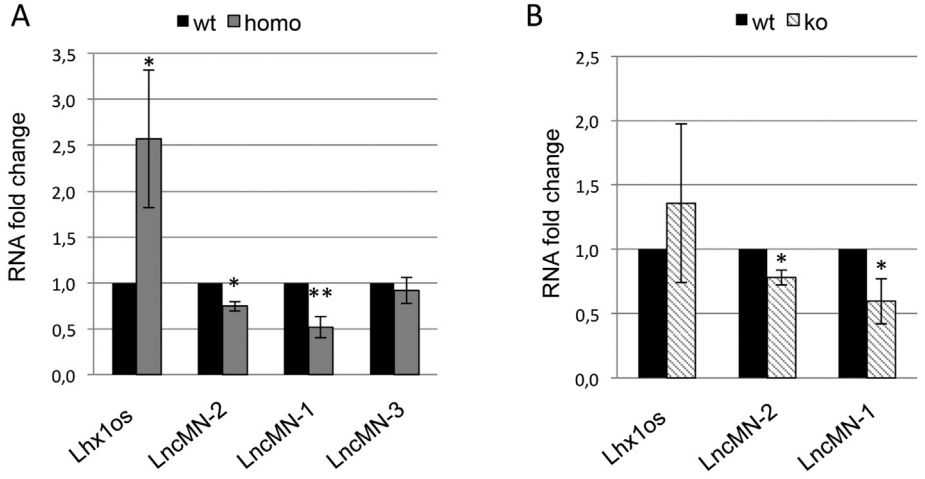

Fig. 3. LncRNA expression in FUS-ALS MNs. A. qRT-PCR analysis of selected lncRNAs in $Fus^{P517L/P517L}$ MNs (homo, gray bars), relative to $Fus^{+/+}$ MNs (wt, black bars), set as 1; n = 3. Normalization vs Atp5o expression levels. One clone for each genetic background analyzed. B. qRT-PCR analysis of specific lncRNAs in $Fus^{-/-}$ MNs (ko, striped bars), relative to $Fus^{+/+}$ MNs (wt, black bars), set as 1; n = 3. Normalization vs Atp5o expression levels. One clone for each genetic background analyzed.

human models (Fig. 4A). Notably, both PCGs showed the same trend of deregulation of the neighboring lncRNAs also upon FUS mutation (Fig. 4B), indicating a co-regulated response to Fus for these divergent transcriptional units.

4. Discussion

Transcriptome sequencing of bulk cell populations has led to the idea that lncRNAs are poorly expressed. However recent single cell analysis has revealed that lncRNAs are expressed at levels comparable to mRNAs (Liu et al., 2016). In this work, we analyzed the long non-coding transcriptome of FAC-Sorted mESC-derived MNs and identified a set of strongly induced lncRNAs. Interestingly, the expression of a subgroup of these was activated in the late stages of differentiation, when post-mitotic MNs are specified, possibly indicating a role for these molecules not only in MN differentiation but also in their mature physiology. Notably, our candidates are conserved by sequence and/or synteny in human and are expressed in MNs derived from human iPSCs. Considering the generally poor conservation of this class of transcripts, it is therefore likely that the identified species might be involved in some conserved and critical MN function. While LncMN-2 and LncMN-3 are located in intergenic regions, Lhx1os and LncMN-1 are transcribed in opposite direction from PCGs involved in MN physiology. Interestingly, we observed a strong co-regulation of the divergent transcripts in both human and mouse models. Recent work highlighted a non-random distribution of lncRNAs in mammalian genomes with 20% of them being divergent from PCGs. Moreover, a higher level of expression correlation was observed in the case of lncRNAs divergent from PCGs compared to pairs of neighboring PCGs (Luo et al., 2016). Interestingly, Gene Ontology term analysis indicated that PCGs divergent from lncRNA genes are enriched for transcription factors involved in pattern specification (Luo et al., 2016). In our case, Lhx1 (PCG of Lhx1os) is a LIM-domain containing transcription factor involved in urogenital, kidney, liver, and nervous system development but in adult tissues its expression is restricted to the kidney and the brain (Hunter and Rhodes, 2005); moreover, insertions/deletions in the chromosomal region harboring this gene have been associated with several cases of mental retardation (OMIM). With respect to the MN physiology, the expression of Lhx1 determines the specification of laterally positioned Lateral Motor Column MNs (lLMC) which are the MNs projecting to the dorsal muscles of the limbs (Tsuchida et al., 1994). Lhx1 is also known to control the expression of adhesion molecules, therefore participating in the establishment of axon trajectory (Kania and Jessell, 2003).

Another very interesting candidate for the study of MN function is the PCG antisense to LncMN-1, Pcdh10, a calcium-dependent adhesion protein belonging to the non-clustered protocadherin family (Kim et

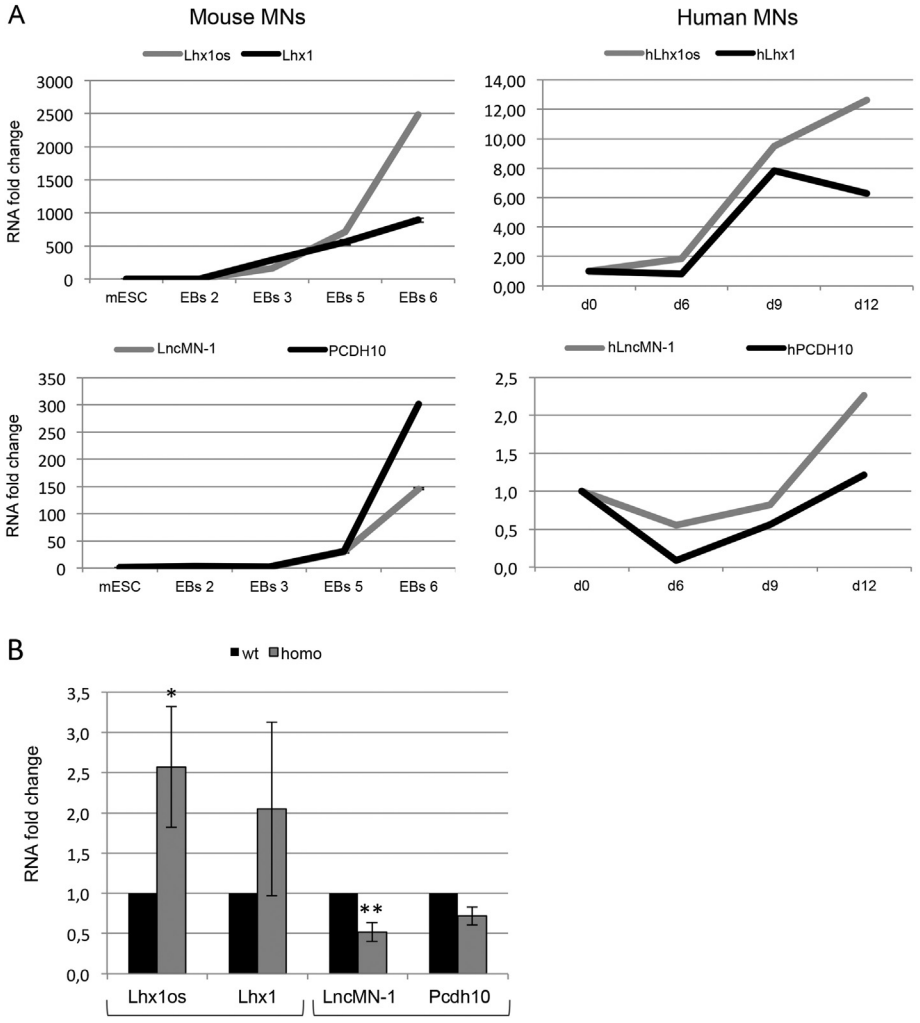

Fig. 4. Expression of divergent lncRNAs/PCG pairs. A. qRT-PCR analysis of divergent lncRNAs (Lhx1os and LncMN-1) and their neighboring genes (Lhx1 and Pcdh10) during mESC (left panels) or iPSC (right panels) differentiation to MNs. Expression in stem cells was set as 1. Normalization vs Atp5o expression levels. B. qRT-PCR analysis of divergent lncRNAs (Lhx1os and LncMN-1) and their neighboring genes (Lhx1 and Pcdh10) in $Fus^{P517L/P517L}$ MNs (homo, gray bars), relative to $Fus^{+/+}$ MNs (wt, black bars), set as 1. Normalization vs Atp5o expression levels.

al., 2011). Pcdh10 mediates homophilic adhesion and it is involved in neural cells clustering and axon pathfinding (Machado et al., 2014; Uemura et al., 2007).

In ALS, MN degeneration has been linked to several proteins involved in RNA metabolism. The molecular mechanisms linking gene mutation to MN degeneration are currently under investigation and models to discriminate the role of mutated proteins, specifically in MNs at pre-symptomatic stage, may shed light on the pathogenic mechanisms of disease onset. In our work, the use of mESC lines carrying either a null or the P517L FUS allele in homozygous conditions, revealed that three of the conserved lncRNAs (Lhx1os, lncMN-1 and lncMN-2) are affected by Fus through a loss-of-function mechanism. In the case of Lhx1os and lncMN-1, the PCGs expressed in divergent orientation also showed the same trend of deregulation in conditions of Fus depletion or mutation, indicating a co-regulated response to Fus. Since the divergent PCGs have a relevant role in neurogenesis, it will be interesting to analyze whether the non-coding transcripts work independently from the PCGs or if they are involved, along with Fus, in controlling the expression of the divergent gene.

In conclusion, through the use of an *in vitro* stem cell-based differentiation system, this work provides interesting lncRNA candidates for the study of new regulatory mechanisms involved in the control of MN differentiation or activity, which may contribute to ALS pathogenesis.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scr.2018.01.037.

Competing interests

The authors declare no conflict of interest.

Funding

This work was partially supported by grants from: ERC-2013 (AdG 340172–MUNCODD), AriSLA full grant 2014 'ARCI', MAECI Grant circALS, Telethon (GGP16213), Epigen-Epigenomics Flagship Project, Human Frontiers Science Program Award RGP0009/2014, AFM-Telethon (17835) and Fondazione Roma NCDs 2013 to IB. This work was also partially supported by NIH/NINDS grant R01 NS07377 to NAS; LL was supported by the Judith and Jean Pape Adams Foundation.

Authors' contributions

SB performed cellular/molecular experiments and wrote the paper; DC contributed to ES cell manipulation; GP performed FAC-sorting and analysis; LL established ES cell lines and performed MN differentiation; AC performed bioinformatics analysis; TS and EC contributed to data analysis; NAS supervised the generation of ES cell lines; PL performed MN differentiation/RNA collection and coordinated the experimental activity; IB conceived the project and wrote the paper.

Acknowledgements

We thank A. Rosa and R. De Santis for human iPSC RNA samples, M. Ballarino for helpful discussions and D. Zagoura for cell cultures.

References

Alaynick, W.A., Jessell, T.M., Pfaff, S.L., 2011. SnapShot: spinal cord development. Cell 146, 178 (e171).

de Andres-Pablo, A., Morillon, A., Wery, M., 2017. LncRNAs, lost in translation or licence to regulate? Curr. Genet. 63, 29–33.

Aprea, J., Prenninger, S., Dori, M., Ghosh, T., Monasor, L.S., Wessendorf, E., Zocher, S., Massalini, S., Alexopoulou, D., Lesche, M., Dahl, A., Groszer, M., Hiller, M., Calegari, F., 2013. Transcriptome sequencing during mouse brain development identifies long non-coding RNAs functionally involved in neurogenic commitment. EMBO J. 32, 3145–3160.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

Briggs, J.A., Wolvetang, E.J., Mattick, J.S., Rinn, J.L., Barry, G., 2015. Mechanisms of long non-coding RNAs in mammalian nervous system development, plasticity, disease, and evolution. Neuron 88, 861–877.

De Santis, R., Santini, L., Colantoni, A., Peruzzi, G., de Turris, V., Alfano, V., Bozzoni, I., Rosa, A., 2017. FUS mutant human motoneurons display altered transcriptome and microRNA pathways with implications for ALS pathogenesis. Stem Cell Rep. 9, 1450–1462.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., Lagarde, J., Veeravalli, L., Ruan, X., Ruan, Y., Lassmann, T., Carninci, P., Brown, J.B., Lipovich, L., Gonzalez, J.M., Thomas, M., Davis, C.A., Shiekhattar, R., Gingeras, T.R., Hubbard, T.J., Notredame, C., Harrow, J., Guigo, R., 2012. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res. 22, 1775–1789.

Errichelli, L., Dini Modigliani, S., Laneve, P., Colantoni, A., Legnini, I., Capauto, D., Rosa, A., De Santis, R., Scarfo, R., Peruzzi, G., Lu, L., Caffarelli, E., Shneider, N.A., Morlando, M., Bozzoni, I., 2017. FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons. Nat. Commun. 8, 14741.

Faghihi, M.A., Modarresi, F., Khalil, A.M., Wood, D.E., Sahagan, B.G., Morgan, T.E., Finch, C. E., St Laurent, G., 3rd, P.J. Kenny, Wahlestedt, C., 2008. Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. Nat. Med. 14, 723–730.

Flicek, P., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., Gil, L., Giron, C.G., Gordon, L., Hourlier, T., Hunt, S., Johnson, N., Juettemann, T., Kahari, A.K., Keenan, S., Kulesha, E., Martin, F.J., Maurel, T., McLaren, W.M., Murphy, D.N., Nag, R., Overduin, B., Pignatelli, M., Pritchard, B., Pritchard, E., Riat, H.S., Ruffier, M., Sheppard, D., Taylor, K., Thormann, A., Trevanion, S.J., Vullo, A., Wilder, S.P., Wilson, M., Zadissa, A., Aken, B.L., Birney, E., Cunningham, F., Harrow, J., Herrero, J., Hubbard, T.J., Kinsella, R., Muffato, M., Parker, A., Spudich, G., Yates, A., Zerbino, D.R., Searle, S.M., 2014. Ensembl 2014. Nucleic Acids Res. 42, D749–755.

Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R., Bruhn, L., Yang, X., Amit, I., Meissner, A., Regev, A., Rinn, J.L., Root, D.E., Lander, E.S., 2011. lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature 477, 295–300.

Hill, S.J., Mordes, D.A., Cameron, L.A., Neuberg, D.S., Landini, S., Eggan, K., Livingston, D.M., 2016. Two familial ALS proteins function in prevention/repair of transcription-associated DNA damage. Proc. Natl. Acad. Sci. U. S. A. 113, E7701–E7709.

Hunter, C.S., Rhodes, S.J., 2005. LIM-homeodomain genes in mammalian development and human disease. Mol. Biol. Rep. 32, 67–77.

Kania, A., Jessell, T.M., 2003. Topographic motor projections in the limb imposed by LIM homeodomain protein regulation of ephrin-A:EphA interactions. Neuron 38, 581–596.

Kim, S.Y., Yasuda, S., Tanaka, H., Yamagata, K., Kim, H., 2011. Non-clustered protocadherin. Cell Adhes. Migr. 5, 97–105.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, R36.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Li, H., Durbin, R., 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., S. Genome Project Data Processing, 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

Lin, M., Pedrosa, E., Shah, A., Hrabovsky, A., Maqbool, S., Zheng, D., Lachman, H.M., 2011. RNA-Seq of human neurons derived from iPS cells reveals candidate long non-coding RNAs involved in neurogenesis and neuropsychiatric disorders. PLoS One 6, e23356.

Ling, S.C., Polymenidou, M., Cleveland, D.W., 2013. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. Neuron 79, 416–438.

Liu, S.J., Nowakowski, T.J., Pollen, A.A., Lui, J.H., Horlbeck, M.A., Attenello, F.J., He, D., Weissman, J.S., Kriegstein, A.R., Diaz, A.A., Lim, D.A., 2016. Single-cell analysis of long non-coding RNAs in the developing human neocortex. Genome Biol. 17, 67.

Luo, S., Lu, J.Y., Liu, L., Yin, Y., Chen, C., Han, X., Wu, B., Xu, R., Liu, W., Yan, P., Shao, W., Lu, Z., Li, H., Na, J., Tang, F., Wang, J., Zhang, Y.E., Shen, X., 2016. Divergent lncRNAs regulate gene expression and lineage differentiation in pluripotent cells. Cell Stem Cell 18, 637–652.

Machado, C.B., Kanning, K.C., Kreis, P., Stevenson, D., Crossley, M., Nowak, M., Iacovino, M., Kyba, M., Chambers, D., Blanc, E., Lieberam, I., 2014. Reconstruction of phrenic neuron identity in embryonic stem cell-derived motor neurons. Development 141, 784–794.

Mo, C.F., Wu, F.C., Tai, K.Y., Chang, W.C., Chang, K.W., Kuo, H.C., Ho, H.N., Chen, H.F., Lin, S. P., 2015. Loss of non-coding RNA expression from the DLK1-DIO3 imprinted locus correlates with reduced neural differentiation potential in human embryonic stem cell lines. Stem Cell Res Ther 6 (1).

Stamatoyannopoulos, J.A., Snyder, M., Hardison, R., Ren, B., Gingeras, T., Gilbert, D.M., Groudine, M., Bender, M., Kaul, R., Canfield, T., Giste, E., Johnson, A., Zhang, M., Balasundaram, G., Byron, R., Roach, V., Sabo, P.J., Sandstrom, R., Stehling, A.S., Thurman, R.E., Weissman, S.M., Cayting, P., Hariharan, M., Lian, J., Cheng, Y., Landt, S.G., Ma, Z., Wold, B.J., Dekker, J., Crawford, G.E., Keller, C.A., Wu, W., Morrissey, C., Kumar, S.A., Mishra, T., Jain, D., Byrska-Bishop, M., Blankenberg, D., Lajoie, B.R., Jain, G., Sanyal, A., Chen, K.B., Denas, O., Taylor, J., Blobel, G.A., Weiss, M.J., Pimkin, M., Deng, W., Marinov, G.K., Williams, B.A., Fisher-Aylor, K.I., Desalvo, G., Kiralusha, A., Trout, D., Amrhein, H., Mortazavi, A., Edsall, L., McCleary, D., Kuan, S., Shen, Y., Yue, F., Ye, Z., Davis, C.A., Zaleski, C., Jha, S., Xue, C., Dobin, A., Lin, W., Fastuca, M., Wang, H., Guigo, R., Djebali, S., Lagarde, J., Ryba, T., Sasaki, T., Malladi, V.S., Cline, M.S., Kirkup, V.M., Learned, K., Rosenbloom, K.R., Kent, W.J., Feingold, E.A., Good, P.J.,

Pazin, M., Lowdon, R.F., Adams, L.B., 2012. An encyclopedia of mouse DNA elements (mouse ENCODE). Genome Biol. 13, 418.
Ng, S.Y., Bogu, G.K., Soh, B.S., Stanton, L.W., 2013. The long noncoding RNA RMST interacts with SOX2 to regulate neurogenesis. Mol. Cell 51, 349–359.
Nishimoto, Y., Nakagawa, S., Hirose, T., Okano, H.J., Takao, M., Shibata, S., Suyama, S., Kuwako, K., Imai, T., Murayama, S., Suzuki, N., Okano, H., 2013. The long non-coding RNA nuclear-enriched abundant transcript 1_2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis. Mol. Brain 6, 31.
Ponting, C.P., Oliver, P.L., Reik, W., 2009. Evolution and functions of long noncoding RNAs. Cell 136, 629–641.
Renton, A.E., Chio, A., Traynor, B.J., 2014. State of play in amyotrophic lateral sclerosis genetics. Nat. Neurosci. 17, 17–23.
Taylor, J.P., Brown Jr., R.H., Cleveland, D.W., 2016. Decoding ALS: from genes to mechanism. Nature 539, 197–206.
Tsuchida, T., Ensini, M., Morton, S.B., Baldassare, M., Edlund, T., Jessell, T.M., Pfaff, S.L., 1994. Topographic organization of embryonic motor neurons defined by expression of LIM homeobox genes. Cell 79, 957–970.
Uemura, M., Nakao, S., Suzuki, S.T., Takeichi, M., Hirano, S., 2007. OL-Protocadherin is essential for growth of striatal axons and thalamocortical projections. Nat. Neurosci. 10, 1151–1159.
Vance, C., Rogelj, B., Hortobagyi, T., De Vos, K.J., Nishimura, A.L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P., Ganesalingam, J., Williams, K.L., Tripathi, V., Al-Saraj, S., Al-Chalabi, A., Leigh, P.N., Blair, I.P., Nicholson, G., de Belleroche, J., Gallo, J. M., Miller, C.C., Shaw, C.E., 2009. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science (New York, N.Y.) 323, 1208–1211.
Vance, K.W., Sansom, S.N., Lee, S., Chalei, V., Kong, L., Cooper, S.E., Oliver, P.L., Ponting, C.P., 2014. The long non-coding RNA Paupar regulates the expression of both local and distal genes. EMBO J. 33, 296–311.
Wichterle, H., Peljto, M., 2008. Differentiation of mouse embryonic stem cells to spinal motor neurons. Curr. Protoc. Stem Cell Biol. Chapter 1. Unit 1H 1 1-1H 1 9.
Wichterle, H., Lieberam, I., Porter, J.A., Jessell, T.M., 2002. Directed differentiation of embryonic stem cells into motor neurons. Cell 110, 385–397.
Williams, L.H., Fromm, G., Gokey, N.G., Henriques, T., Muse, G.W., Burkholder, A., Fargo, D. C., Hu, G., Adelman, K., 2015. Pausing of RNA polymerase II regulates mammalian developmental potential through control of signaling networks. Mol. Cell 58, 311–322.
Zalfa, F., Giorgi, M., Primerano, B., Moro, A., Di Penta, A., Reis, S., Oostra, B., Bagni, C., 2003. The fragile X syndrome protein FMRP associates with BC1 RNA and regulates the translation of specific mRNAs at synapses. Cell 112, 317–327.
Zhao, X., Tang, Z., Zhang, H., Atianjoh, F.E., Zhao, J.Y., Liang, L., Wang, W., Guan, X., Kao, S.C., Tiwari, V., Gao, Y.J., Hoffman, P.N., Cui, H., Li, M., Dong, X., Tao, Y.X., 2013. A long noncoding RNA contributes to neuropathic pain by silencing Kcna2 in primary afferent neurons. Nat. Neurosci. 16, 1024–1031.
Zhao, S., Guo, Y., Sheng, Q., Shyr, Y., 2014. Advanced heat map and clustering analysis using heatmap3. Biomed. Res. Int. 2014, 986048.
Zhong, J., Chuang, S.C., Bianchi, R., Zhao, W., Lee, H., Fenton, A.A., Wong, R.K., Tiedge, H., 2009. BC1 regulation of metabotropic glutamate receptor-mediated neuronal excitability. J. Neurosci. 29, 9977–9986.