

UC San Diego

UC San Diego Previously Published Works

Title

Properties of LINE-1 proteins and repeat element expression in the context of amyotrophic lateral sclerosis.

Permalink

<https://escholarship.org/uc/item/9cx2h0zb>

Authors

Pereira, Gavin

Sanchez, Laura

Schaughency, Paul

et al.

Publication Date

2018

DOI

10.1186/s13100-018-0138-z


Peer reviewed

RESEARCH

Open Access



Properties of LINE-1 proteins and repeat element expression in the context of amyotrophic lateral sclerosis

Gavin C. Pereira¹, Laura Sanchez², Paul M. Schaughency³, Alejandro Rubio-Roldán², Jungbin A. Choi¹, Evarist Planet⁴, Ranjan Batra⁵, Priscilla Turelli⁴, Didier Trono⁴, Lyle W. Ostrow⁶, John Ravits⁵, Haig H. Kazazian¹, Sarah J. Wheelan³, Sara R. Heras^{2,7}, Jens Mayer⁸, Jose Luis García-Pérez^{2,9} and John L. Goodier^{1*} 

ABSTRACT

Background: Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disease involving loss of motor neurons and having no known cure and uncertain etiology. Several studies have drawn connections between altered retrotransposon expression and ALS. Certain features of the LINE-1 (L1) retrotransposon-encoded ORF1 protein (ORF1p) are analogous to those of neurodegeneration-associated RNA-binding proteins, including formation of cytoplasmic aggregates. In this study we explore these features and consider possible links between L1 expression and ALS.

Results: We first considered factors that modulate aggregation and subcellular distribution of LINE-1 ORF1p, including nuclear localization. Changes to some ORF1p amino acid residues alter both retrotransposition efficiency and protein aggregation dynamics, and we found that one such polymorphism is present in endogenous L1s abundant in the human genome. We failed, however, to identify CRM1-mediated nuclear export signals in ORF1p nor strict involvement of cell cycle in endogenous ORF1p nuclear localization in human 2102Ep germline teratocarcinoma cells. Some proteins linked with ALS bind and colocalize with L1 ORF1p ribonucleoprotein particles in cytoplasmic RNA granules. Increased expression of several ALS-associated proteins, including TAR DNA Binding Protein (TDP-43), strongly limits cell culture retrotransposition, while some disease-related mutations modify these effects. Using quantitative reverse transcription PCR (RT-qPCR) of ALS tissues and reanalysis of publicly available RNA-Seq datasets, we asked if changes in expression of retrotransposons are associated with ALS. We found minimal altered expression in sporadic ALS tissues but confirmed a previous report of differential expression of many repeat subfamilies in *C9orf72* gene-mutated ALS patients.

Conclusions: Here we extended understanding of the subcellular localization dynamics of the aggregation-prone LINE-1 ORF1p RNA-binding protein. However, we failed to find compelling evidence for misregulation of LINE-1 retrotransposons in sporadic ALS nor a clear effect of ALS-associated TDP-43 protein on L1 expression. In sum, our study reveals that the interplay of active retrotransposons and the molecular features of ALS are more complex than anticipated. Thus, the potential consequences of altered retrotransposon activity for ALS and other neurodegenerative disorders are worthy of continued investigation.

* Correspondence: jgoodier@jhmi.edu

¹McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Full list of author information is available at the end of the article



Background

With the discovery in 1950 of transposable elements (TEs) genomes began to seem far more dynamic than hitherto conceived [1]. It is now clear that TEs have been important long-term drivers of genome evolution. Year by year, more and more ways in which mobile DNA impacts gene expression and integrity, cell variability and viability, and ultimately human health are revealed. With recent discoveries that TEs are active not only in the germline but also in somatic cells, it is evident that each of us is a mosaic of different genomes that now seem dynamic indeed (reviewed by [2] and many others).

Retrotransposon TEs include long terminal repeat (LTR) and non-LTR class elements. Both retrotranspose by a “copy and paste” mechanism involving reverse transcription of an RNA intermediate and insertion of its cDNA copy at a new site in the genome. LTR-retrotransposons, including human endogenous retroviruses (HERVs), are remnants of past germ line infections by retroviruses that subsequently lost their ability to reinfect cells. While the HERV-K(HML-2) group includes some polymorphic proviral loci [3, 4], human LTR retrotransposons generally are insertionally inactive, although many remain capable of transcription. Long Interspersed Element-1 (LINE-1, L1) retrotransposons are the only active autonomous mobile DNA in humans. Alone they occupy at least 17% of our genome and have also been responsible for the insertion *in trans* of thousands of processed pseudogenes and a million non-autonomous Short Interspersed Elements (SINEs), including Alu and SVA (composite SINE/VNTR/Alu) elements [5]. The 6.0 kilobase (kb) bicistronic human L1 has a 5' untranslated region (UTR) that functions as an internal promoter, two open reading frames (ORF1 and ORF2), and a 3' UTR. A weak promoter also exists on the antisense strand of the human L1 5' UTR [6]. ORF2 encodes a 150-kilodalton (kD) protein with DNA endonuclease and reverse transcriptase (RT) activities. While the 40 kD ORF1p RNA-binding protein is essential for retrotransposition, its exact role in retrotransposition is unclear, although it possesses RNA chaperone and packaging properties [7–9]. The great majority of L1s in the genome are 5' truncated and otherwise rearranged or mutated and so incapable of autonomous transcription.

There are 145 fully intact L1s in the human genome of which cell culture retrotransposition assays suggest about 100 remain potentially mobile in any individual diploid genome [10–12]. There are also hundreds of full length L1s lacking intact ORFs but possibly capable of generating protein [13, 14]. While L1 expression is normally suppressed by a host of cellular factors, the suppression is relaxed in embryonic stem cells, the early embryo, and some cancers (reviewed in [15, 16]).

Notably, retrotransposons are also active in some brain cells [17]. Loss of piRNA pathway proteins correlates

with elevated retrotransposon expression in *Drosophila* brain [18], although in mammals this pathway seems to act primarily in the germ line to control retrotransposon activity (see [19] for review). Early studies showed L1 retrotransposition in dividing neuronal progenitor cells (NPCs), especially those of the hippocampus [20, 21], and subsequently in non-dividing neurons [22, 23]. High-throughput sequencing of single neurons confirmed endogenous L1 retrotransposition in the human brain, although frequency estimates differed significantly (reviewed in [24–27]). Thus, it has been proposed that L1 activity contributes to neuronal plasticity [28]. Elevated L1 retrotransposition has also been reported for several human neurological conditions, including ataxia telangiectasia [29], Rett syndrome [30], autism [31], schizophrenia [32, 33], and major depressive disorder [34], as well as in neuronal cell lines or brains of patients exposed to opioids [35–37], and in brains of a mouse model of Huntington's disease [38] and hippocampi of mice following novel exploration [39] or diminished maternal care [40]. However, some of these studies relied solely on DNA amplification by quantitative (q)PCR or digital droplet PCR to compare L1 insertion copy differences between test and normal states, strategies that may fail to distinguish between *bona fide* genomic L1 insertions and contaminating extrachromosomal L1-derived nucleic acids. Some studies may therefore warrant additional verification (see also [41, 42] and discussion).

Several studies have also drawn connections between altered expression of LTR retrotransposons and amyotrophic lateral sclerosis (ALS). ALS is a fatal neurodegenerative disease involving loss of upper and lower motor neurons and afflicts 2 in 100,000 people each year. Death typically follows 2 to 3 years after onset and, while about 90% of cases are sporadic, the rest have a family history of the disease. There are no current means to reverse the course of ALS, and treatment involves efforts to slow progression of symptoms [43]. ALS has overlapping clinical presentations with frontotemporal lobar degeneration (FTLD) and its most common subtype frontotemporal dementia (FTD), a neurological condition affecting the frontal and temporal lobes and marked by cognitive and behavioral impairment. About 20% of ALS patients also exhibit FTLD, and ALS and FTLD have been seen as part of a continuous disease spectrum [44].

Increased reverse transcriptase activity from an unknown source is detectable in sera and cerebrospinal fluids of non-HIV-infected ALS patients [45–48]. Douville et al. [49] correlated this RT activity with elevated expression of a few HERV-K(HML-2) loci and increased amounts of *pol* gene transcripts and RT protein in cortical neurons of some ALS patients. Hadlock et al. [50]

noted elevated immune response to HERV-K(HML-2) Gag protein in serum samples from ALS patients, and recently it was shown that overexpression of the HERV-K(HML-2) envelope protein causes motor neuron toxicity and motor dysfunction in transgenic mice [51]. However, it cannot be excluded that some of the elevated RT activity observed is also due to increased expression of LINE-1 retrotransposons. It is reasonable to presume that cellular changes that increase HERV-K(HML-2) expression in ALS patients may similarly activate other retrotransposons. Indeed, a recent study reported global increases in expression of selected families of both LTR and non-LTR retrotransposons in ALS and FTLD patients with a hexanucleotide expansion in the Chromosome 9 Open Reading Frame 72 (*C9orf72*) ALS gene but not in sporadic ALS cases or controls [52].

The accumulation of neuronal RNA and protein aggregates, including cytoplasmic stress granules (SGs), is a pathogenic hallmark of a number of neurodegenerative diseases. Pathological aggregation of RNA-binding proteins has been implicated in ALS, FTLD, Alzheimer's disease, spinocerebellar ataxia, Huntington's disease, and inclusion body myositis. In the case of ALS, there is increasing evidence that abnormal RNA processing and abnormal self-aggregation of proteins, leading to altered RNA granule formation and malfunction of protein pathways, contribute to motor neuron death [53]. A key pathological feature of ALS is the presence of cytoplasmic inclusions in degenerating motor neurons and oligodendrocytes. Inclusions are not restricted to the spinal cord and motor cortex but are present in other brain regions, such as the frontal and temporal cortices, hippocampus, and cerebellum, and are especially evident in patients with accompanying FTD [54]. What triggers protein aggregation and what it means for cell pathology and progression of the disease remain unclear.

Aggregation of TAR DNA binding protein 43 (TDP-43, product of the *TARDBP* gene) is especially interesting as a unifying pathological marker of both FTLD and ALS. Mutations in *TARDBP* are involved in about 4% of familial (fALS) and 1% of sporadic ALS (sALS) cases. However, even lacking a mutation, TDP-43 protein, while typically nuclear in healthy cells, is cleaved and hyperphosphorylated and accumulates in ubiquitinated cytoplasmic inclusions in almost all ALS and almost half of FTLD patients (reviewed in [55]). TDP-43 protein aggregation pathology also characterizes other neurodegenerative disorders, including Parkinson's [56], Alzheimer's [57] and Huntington's [58] diseases, and inclusion body myopathies [59].

Several features of LINE-1-encoded ORF1p are reminiscent of neurodegeneration-associated proteins. ORF1p is an ubiquitinated and phosphorylated RNA-binding protein prone to forming cytoplasmic aggregates, including

SGs [60–63]. Therefore, it is conceivable therefore that abnormal expression of ORF1p in neuronal cells might aggravate formation of cytoplasmic aggregates and contribute to disease pathology. Here we analyzed subcellular localization and aggregation features of LINE-1 ORF1p and ways they may be analogous to or differ from those of neurodegeneration-associated RNA-binding proteins. We show that some ALS-associated RNA-binding protein mutants closely associate with ORF1p in cytoplasmic RNA granules of tumor cell lines, and that increasing the expression of some ALS proteins, including TDP-43, inhibits L1 retrotransposition in a cell culture reporter assay. We also considered the possibility that LINE-1 retrotransposon activity may be associated with ALS disease. Reverse transcription (RT)-qPCR analyses failed to detect significantly altered expression of non-LTR Alu or L1 elements in sALS tissues. However, by reanalyzing publicly available RNA-Seq datasets, one previously examined for TE levels [52] and one hitherto untested, we confirmed misregulation of selected TE subfamilies in *C9orf72* gene-related ALS samples. While so far the evidence is not compelling for ALS, we believe the potential for altered non-LTR retrotransposon expression playing a role in neurodegenerative disorders is worthy of continued investigation.

Results

A common LINE-1 polymorphism alters formation of ORF1p cytoplasmic RNA granules

Early studies showed that endogenous ORF1p concentrates in cytoplasmic granules in cultured cells or fixed tissues [60, 64]. Overexpression or cell stress causes ORF1p to enter stress granules (Fig. 1F) [60, 61, 65, 66]. Here we used the monoclonal α -human α -4H1 ORF1 antibody (Millipore Sigma; [67]), which targets ORF1p N-terminal amino acids (aa) 35 to 44, to examine expression of endogenous ORF1p in 2102Ep cells, an embryonal germ cell teratocarcinoma line that has unchanging human embryonal stem cell (hESC) characteristics [68, 69]. By Western blotting, α -4H1-ORF1 detected a robust approximately 42 kD band of size consistent with the ORF1p monomer, a much weaker 85 kD band that likely marks ORF1p dimers, and sporadically a band of 65 kD of unknown identity (Additional file 1: Figure S1A). Single, less intense 42 kD bands were seen in human embryonic kidney HEK 293T and neuroblastoma SH-SY5Y cells. SH-SY5Y is a thrice-cloned sub-line of the bone marrow-derived neural line SK-N-SH, which shows significantly less ORF1p signal. No endogenous ORF1p was detected in human cervical adenocarcinoma HeLa cells (Additional file 1: Figure S1A).

Confirming previous results [60], constitutive aggregation of ORF1p in the cytoplasm was detected in multiple

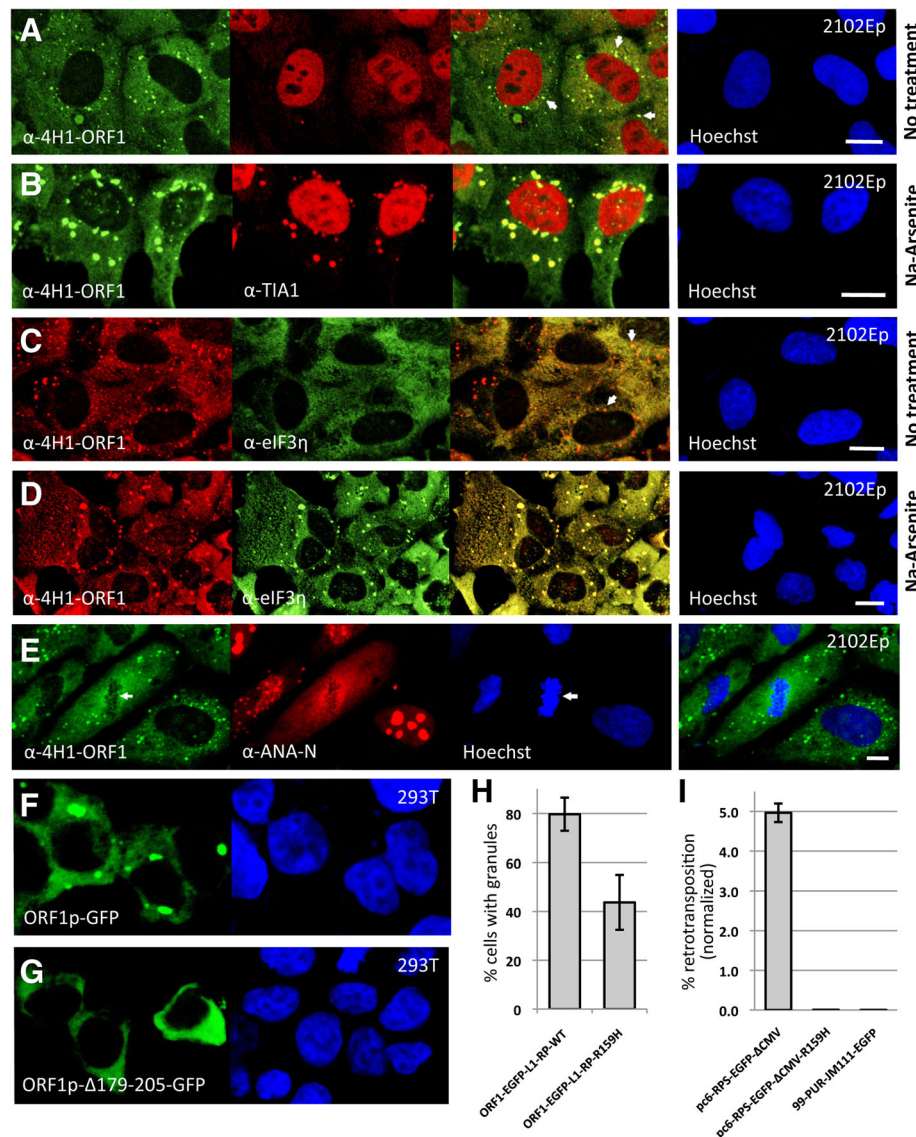


Fig. 1 Cytoplasmic localization of L1 ORF1 protein. Stress granule marker proteins TIA1 (**a, b**) and eIF3η (**c, d**) show minimal colocalization with endogenous ORF1p in cytoplasmic granules (shown by arrows) of unstressed human embryonal carcinoma 2102Ep cells (**a, c**) but colocalize in large stress granules of cells treated with Na-arsenite (0.5 mM for 1 hour) (**b, d**). Cell nuclei are stained with Hoechst. Size bars are 10 μm. **e** ORF1p cytoplasmic granules retain integrity during cellular mitosis. Patient sera-derived α-ANA-N marks nucleoli [60]. Endogenous ORF1p is mostly excluded from metaphase chromatin plates (arrow), as shown by Hoechst staining (see arrow). **f** Ectopically expressed EGFP-tagged human ORF1p induces prominent cytoplasmic granules, but (**g**) deletion of its Q-N-rich region abolishes granule formation. **h** The ORF1p R159H point mutation reduces cytoplasmic granule formation by 50%. Approximately 500 cells were examined. **i** The R159H mutation abolishes cell culture LINE-1 retrotransposition. pc6-RPS-EGFP-ΔCMV wild-type or R159H mutant retrotransposition reporter constructs were transfected in HEK 293T cells and 5 days later the percentages of EGFP-positive cells were determined by flow cytometry. The construct 99-PUR-JM111-EGFP served as a negative control for retrotransposition [84]. Each construct was tested in quadruplicate wells with results for one biological replicate shown

cell lines by multiple α-ORF1p antibodies (Additional file 1: Figure S1B-E). However, the pattern and degree of granule formation by ORF1p can vary significantly with cell type. In human neuroblastoma SH-SY5Y cells, for example, ORF1p granules are rare in the main cytoplasm but evident in neurite outgrowths (Additional file 1: Figure S1C). Notably, endogenous ORF1p cytoplasmic aggregates differ from SGs in certain ways. In the absence

of external stress, in 2102Ep cells small ORF1p aggregates are numerous but only faintly and rarely marked by SG proteins such as cytotoxic granule associated RNA binding protein (TIA1) and elongation initiation factor 3 (eIF3η) (Fig. 1A, C). Furthermore, unlike SGs [70], endogenous ORF1p granules do not obviously disassemble during cell mitosis (Fig. 1E). As previously reported, when exposed to sodium arsenite, an inducer of

oxidative stress, ORF1p redistributes to larger-sized aggregates that now mostly colocalize with SG proteins (Fig. 1B, D). Thus ORF1p aggregates form constitutively but do not chronically induce a cellular stress state that is marked by redistribution of SG proteins.

Previously, it was shown that overexpressed ORF2p and L1 RNA also colocalize with ORF1p in cytoplasmic aggregates [60, 61, 63, 71]. Interestingly, we noted [72], and others confirmed [73], that ORF2p is visible in only a minor percentage of ORF1p-positive cells when the two proteins are coexpressed from an L1 construct. The reason for this is unknown but may relate to an unconventional translation mechanism of LINE-1 ORF2 [74]. Unfortunately, although α -ORF2p antibodies exist [75–79], they are not widely available or are ineffective in detecting endogenous ORF2p, and so we did not examine ORF2p localization in this study.

Many RNA-binding proteins that form SGs have intrinsically disordered prion-like domains rich in glutamine and asparagine (Q-N) residues. Aggregation of prion-like domain proteins is characteristic of various neurodegenerative disorders including ALS (reviewed in [80]). No prion-like domain is predicted in ORF1p using the PrionW [81] or PLAAC [82, 83] algorithms. However, human (but not mouse) L1 ORF1p contains a Q-N-rich internal region (36% Q or N between residues 179 and 205, numbering according to accession number AF148856.1). Deletion of this region abolishes granule formation (Fig. 1E, G), indicating it is critical for human L1 ORF1p aggregation properties.

Previously, we and others [60, 61] showed that mutations in the N-terminus leucine zipper domain or the C-terminal domain double-point mutation R261/262A (the so-called JM111 mutation that abrogates cell culture retrotransposition; [84]) also alter ORF1p cytoplasmic aggregation. We also reported that a non-conservative mutation, R159G, inhibits ORF1p granule formation. This residue was subsequently shown to be important for RNA-binding and is within the RNP2 sequence of the ORF1p RRM (RNA recognition motif) [85]. In the present study, to ascertain the prevalence of L1s in the human genome with R159 polymorphisms, we queried the L1Base2 database [13]. L1Base2 is subdivided into 3 categories: L1s with intact ORF1 and ORF2 (FLI-L1s), L1s with intact ORF2 but disrupted ORF1 (ORF2-L1s), and non-intact L1s >4500 nucleotides in length (FLnI-L1s). Although the R159G variant was detected at only very low frequency (0.47% of 6346 alignable FLnI-L1 sequences), many other R159 polymorphisms were found, with R159H being most common. In all, we identified R159 changes to histidine, cysteine or proline residues in 4.8% of FLI-L1s, 11.5% of ORF2-L1s, and 40.3% of FLnI-L1s (Additional file 1: Figure S2A). Thus, sequence variation in the aggregation-control

R159 codon of human L1 ORF1p is common in endogenous L1s.

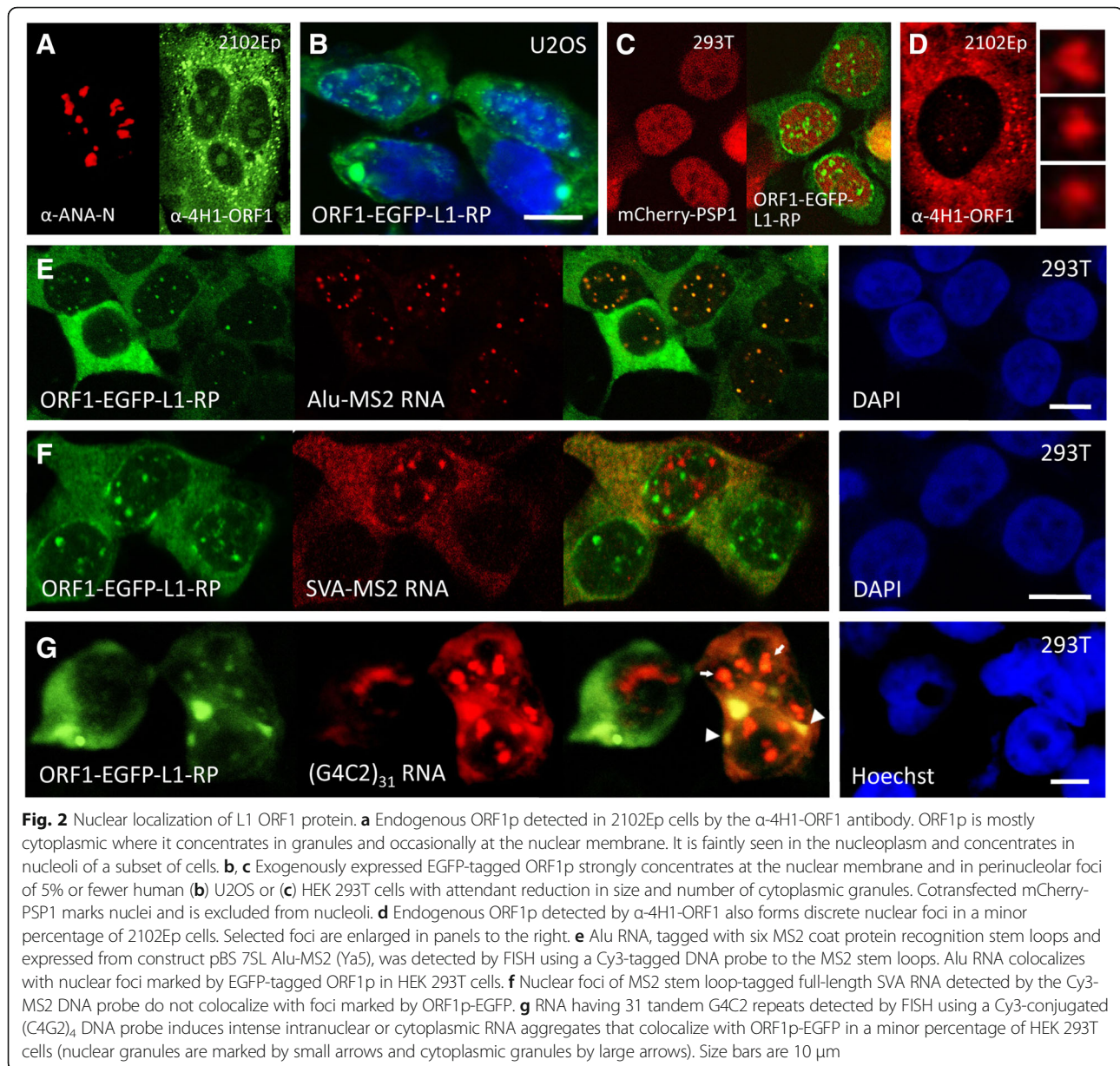
We introduced the R159H change into ORF1-EGFP-L1-RP, a construct with CMV promoter and ORF1 C-terminally tagged with EGFP followed by intact downstream L1 sequence, and as expected observed a 30% decline in the number of HEK 293T cells with ORF1p cytoplasmic granules (Fig. 1H). We next tested the effect of the R159H polymorphism in a cell culture retrotransposition assay. In this assay, an enhanced green fluorescent protein (EGFP) reporter gene reporter cassette, interrupted by a backwards γ -globin intron, is inserted in opposite transcriptional orientation into the 3' UTR of L1-RP (a highly active human L1 [86]). The EGFP reporter gene can be expressed from its own promoter only after the L1 is transcribed, the γ -globin intron is removed by splicing, the L1-reporter cassette hybrid transcript is reverse-transcribed, and its cDNA inserted in the genome [84, 87]. The R159H mutation abolished cell culture retrotransposition to levels similar to that observed for an L1 containing the ORF1p JM111 mutation that cannot form a functional L1 ribonucleoprotein (RNP) complex [88] (Fig. 1I).

Finally, we considered the possibility that the abundance of R159 polymorphisms might be due to a CpG dinucleotide methylation hotspot. Following genome bisulfite conversion, PCR amplification, cloning of the amplicons and Sanger sequencing, we queried the methylation status of nine CpGs within a 436-nt stretch (1169-1604) of ORF1 surrounding the R159 codon. Although CpGs were methylated (16 to 64%), we observed no preference for methylation at the ORF1 R159 codon (Additional file 1: Figure S2B).

Thus, L1 ORF1 polymorphisms can alter not only retrotransposition efficiency but also ORF1p aggregation dynamics for a subset of L1s abundant in the human genome.

LINE-1 ORF1 protein concentrates in nuclear aggregates

Previously we showed that both overexpressed and endogenous L1 ORF1p are not only cytoplasmic but colocalize with nucleoli of a subset of cells (Fig. 2A). Overexpressed ORF2p also enters nucleoli [76]. Here we report that exogenously expressed GFP-tagged human ORF1p also strongly concentrates at the nuclear membrane and forms small discrete perinucleolar foci in 5% or fewer of human osteosarcoma U2OS or HEK 293T cells. These cells show an attendant reduction in the size and number of cytoplasmic granules (Fig. 2B,C). Consistently, endogenous ORF1p nuclear foci are also seen in a small fraction of 2102Ep cells (Fig. 2D); presence of the foci in nuclei was confirmed by z-series confocal imaging. Recently, De Luca et al. [79] also showed in human



melanoma cells both endogenous ORF1p and ORF2p in nuclear puncta that partially colocalized.

Using the MS2-NLS-GFP detection protocol [89], we previously reported that overexpressed Alu SINE RNA forms small distinct nuclear foci that partially associate with coiled (Cajal) body marker proteins [72]. Coiled bodies are nuclear non-membrane RNP suborganelles involved in the processing of non-coding RNAs and have been linked with the rare motor neuron disease spinal muscular atrophy (SMA) [90]. In our present experiments, we show that in the minor percentage of HEK 293T cells that form ORF1p nuclear foci, these foci closely colocalize with coexpressed MS2-tagged Alu RNA (Fig. 2E, detected here by fluorescent in situ

hybridization (FISH)). Thus, ORF1 protein and Alu RNA may directly interact in the nucleus. SVA SINE RNA expressed from plasmid pcDNA SVA_{SPTA1}-MS2 is mainly cytoplasmic but also forms nuclear foci [72]. Interestingly, these foci do not colocalize with ORF1-EGFP foci, despite the fact that both Alu and SVA RNAs depend upon L1 for their retrotransposition and insert in the genome by a common mechanism (Fig. 2F). As previously reported, L1 RNA failed to form nuclear foci in our experiments [61, 72].

Certain neurodegenerative conditions, including myotonic dystrophy, fragile X-associated tremor ataxia syndrome and spinocerebellar ataxias, are associated with genes that undergo long simple repeat expansion

mutations. RNAs transcribed from these mutant genes accumulate in nuclear foci [91]. A pathogenic GGGCC (G4C2) hexanucleotide expansion in intron 1 of the *C9orf72* gene is the most common mutation associated with both ALS and FTD [92, 93] and is implicated in Huntington's disease [94]. Mutant *C9orf72* gene transcripts form toxic RNA foci in affected neuronal cells and are associated with the disease pathology [92, 95]. We transiently coexpressed in HEK 293T cells ORF1p-EGFP together with *C9orf72* RNA having 31 tandem G4C2 repeats, the latter detected by RNA FISH using a Cy3-conjugated (C4G2)₄ probe [96]. As with Alu SINE RNAs, ORF1p-EGFP granules directly overlapped or juxtaposed with G4C2₃₁ RNA granules in nuclei and cytoplasm of some cells (Fig. 2G).

Thus, being a promiscuous RNA-binding protein, it is possible that L1 ORF1p is able to bind and sequester many cellular RNAs in granules present in both the cytoplasm and nucleus.

The control of LINE-1 ORF1p nuclear localization in 2102Ep cells

Several studies have reported that cell division facilitates efficient retrotransposition, citing a failure of L1s to retrotranspose in cultured primary and tumor cells blocked at G₀ phase but disagreeing on the extent of retrotransposition loss in G₁/S-arrested cells (reporting a 3-fold to 10-fold decline; [97–99]). Mita et al. [100] recently reported that cell culture retrotransposition occurs preferentially in S-phase. On the other hand, we previously showed significant retrotransposition in non-dividing neuronal cells differentiated from hESCs [23], and similar data was earlier observed in transformed cell lines [97]. Since ORF1 protein is essential for active retrotransposition [84], we chose to examine two mechanisms postulated to control ORF1p subcellular localization, cell cycle and active nuclear export.

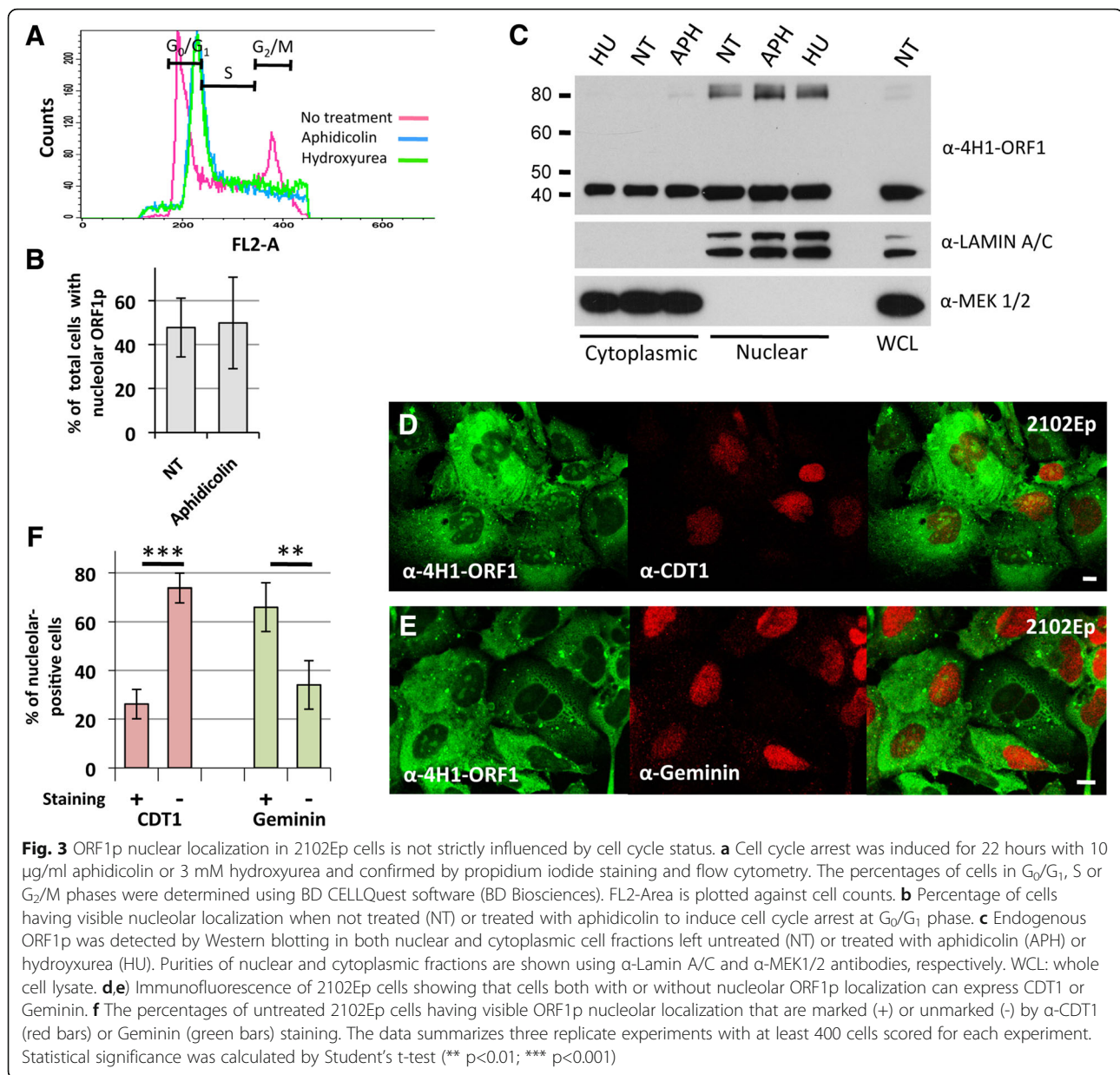
2102Ep cells are nullipotent, manifesting a stable phenotype; indeed, they are used as a reference to characterize newly derived hESC lines [69]. However, we noticed considerable variation in the percentage of 2102Ep cells showing obvious nucleolar localization of endogenous ORF1p when examining clusters of cells across a single slide using immunofluorescence (IF) and the α -4H1-ORF1 antibody (between 0.6% and 36% of total cells randomly examined in three separate experiments). Less densely clustered cells more frequently showed ORF1p nucleolar concentration. Germline tumor cells, including 2102Ep cells, are altered for cell cycling by cell-to-cell contact inhibition or serum depletion [101], and we therefore wondered if concentration of ORF1p in the nucleus might relate to cell cycle status.

Accordingly, we seeded 2102Ep cells at low densities and blocked G₁/S phase transition using aphidicolin, an

inhibitor of DNA polymerase α , or hydroxyurea, which inhibits ribonucleotide reductase causing a loss of deoxyribonucleotides [102, 103], and examined effects on endogenous ORF1p nucleolar localization. Cell cycle blockage was confirmed by propidium iodide staining followed by flow cytometry (Fig. 3A). Blocking the cell cycle at G₁/S had no significant effect on the average percentage of cells with endogenous ORF1p nucleolar localization (Fig. 3B), or on nuclear-cytoplasmic levels following cell fractionation and Western blotting (Fig. 3C). For both treated and non-treated 2102Ep cells, Western blotting showed a major amount of endogenous ORF1p in the nuclear fraction, in agreement with Sokolowski et al. [104] who reported nuclear fraction concentration of plasmid-expressed ORF1p in human HeLa and mouse NIH3T3 cells. However, it is likely that some of the ORF1p we detect in the nuclear fraction is due to copurification of insoluble ORF1p cytoplasmic aggregates [105]. Significantly, however, the amounts of ORF1p detected by Western blotting in both cytoplasmic and nuclear fractions remained unaltered by cell cycle blockage.

We next stained untreated 2102Ep cells with antibodies to chromatin licensing and DNA replication factor 1 (CDT1), a G₁ phase nuclear protein lost after initiation of S phase [106], or Geminin (GMNN), a protein expressed only in S/G₂/M phases [107], and then examined cells for L1 ORF1p nucleolar concentration. Immunocytochemistry showed that both CDT1 and Geminin marked 2102Ep cells with or without endogenous ORF1p visible in nucleoli (Fig. 3D, E). However, a majority of cells showing nucleolar ORF1p failed to stain with CDT1, while the opposite was true for Geminin (Fig. 3F). This suggests partial nucleolar exclusion of ORF1p during G₁ phase but without stringent cell cycle control. Our results in part contradict a recently published observation [100] that overexpressed LINE-1 ORF1p is nuclear in HeLa cells expressing CDT1 (G₁ phase) and almost completely cytoplasmic in cells expressing Geminin (see Discussion).

Next, we considered if LINE-1 ORF1p shuttles between the nucleus and cytoplasm, as is the case with TDP-43 and some other prion-domain RNA-binding proteins associated with neurodegenerative diseases [108]. Shuttling proteins often contain nuclear export signals (NESS), consisting of a short stretch of hydrophobic leucine-rich residues [109]. We previously reported that subcellular localization of overexpressed GFP-tagged ORF1p in HEK 293T cells was unaltered by leptomycin B (LMB), a chemical inhibitor of the chromosomal region maintenance 1 (XPO1/CRM1) nuclear export pathway [76]. We now observed that treatment of 2102Ep cells with 55 nM LMB for 18 hours also had no obvious effect on endogenous ORF1p localization



(Additional file 1: Figure S3A). On the other hand, controls revealed that LMB efficiently inhibited cytoplasmic export of endogenous cyclin B1, which contains an NES responsive to CRM1 (Additional file 1: Figure S3B) [110], as well as a GFP-tagged phosphorylation mimetic mutant of MAPKAP kinase 2 (MK2-mut T205/317E) that remains in the cytoplasm once exported from the nucleus (Additional file 1: Figure S3C) [111].

Previously, we fused a suspected ORF1p leucine-rich NES (aa 87-93, LKELMEL) and linker to the C-terminus of EGFP. While a functional NES should cause EGFP, which is normally both cytoplasmic and nuclear, to become more cytoplasmic [112], we failed to observe increased concentration of EGFP-LKELMEL in the cytoplasm [76]. For

the present study, we used the NetNES 1.1 Server [113] to predict a second NES site at the C-terminus of ORF1p (ORF1 aa 313-321, LKELLKEAL). Fusing this sequence to the N-terminus of EGFP also failed to alter distribution of EGFP (Additional file 1: Figure S3D). Moreover, altering the sequence to encode LKEAAAAAL in construct ORF1-EGFP-L1-RP failed to visibly affect its ORF1p localization (Additional file 1: Figure S3E).

In contrast to our results, Mita et al. [100] reported a 20 to 35% increase in nuclear retention of exogenous ORF1p overexpressed in HeLa cells treated with LMB. While we failed to detect NES sequences in ORF1p or obvious sensitivity to the CRM1 export pathway, we cannot exclude the possibility that LMB causes nuclear

retention of minor amounts of endogenous ORF1p not visibly obvious in our system. Also, despite previous reports of attenuated cell culture retrotransposition following G₁/S phase arrest, our results suggest this is not due to failure of ORF1p RNPs to enter nuclei, at least in 2102Ep cells, which are known to accommodate cell culture retrotransposition [114]. Moreover, despite the previous suggestion that nuclear membrane breakdown is required for nuclear entry of L1 ORF1p [100] this does not appear to be the case for this cell line.

ALS-related protein mutants colocalize with ORF1p in cytoplasmic granules

To date at least 25 genes have been linked to ALS [43, 53]. The first ALS gene discovered, superoxide dismutase (*SOD1*) [115], is mutated in about 20% of familial cases. *C9orf72* is by far the most frequent gene accounting for about 35% of fALS, 25% of fFTD, and 6% of sALS cases [92, 93, 116]. RNA-binding protein FUS (*FUS*) and *TARDBP* mutations each account for about 4% of fALS cases. Other ALS-associated genes, including alsin (*ALS2*), angiogenin (*ANG*), heterogeneous nuclear ribonucleoprotein A1 (*HNRNPA1*), optineurin (*OPTN*), sequestosome-1 (*SQSTM1/P62*), ubiquilin 2 (*UBQLN2*), TANK binding kinase 1 (*TBK1*), valosin-containing protein (*VCP*) and VAMP-associated protein B and C (*VAPB*) among others, account for only a small percentage of cases so confounding treatment strategies. ALS animal models of neurodegeneration have mostly examined the toxic effects of overexpressing disease-related aggregation-prone proteins. Mutants of several ALS-associated RNA-binding proteins are known to shift localization from the nucleus to the cytoplasm and form RNA foci in the disease state [117, 118].

Previously, in a yeast two-hybrid screen we identified FUS protein as an ORF1p interaction partner, and we confirmed that the two wild-type proteins colocalized in cytoplasmic granules of a minor percentage of stressed human nTera-2 embryonal carcinoma cells [60]. Here we generated in-house or obtained from other sources tagged constructs for selected ALS disease-associated mutants and transfected these in HEK 293T or 2102Ep cells in the presence of EGFP-tagged ORF1p or endogenous ORF1p alone. Various ALS-associated mutant but not wild-type FUS, TDP-43, and SOD1 proteins strongly colocalized with ORF1p in granules of unstressed cells (Fig. 4A–C). When cellular oxidative stress was induced by application of sodium arsenite, endogenous TDP-43 formed numerous cytoplasmic granules that strongly colocalized with endogenous ORF1p in most cells (Fig. 4D). Wild-type TDP-43 is known to form SGs in ALS neurons in response to cellular stress [119–121].

HNRNPA1 and HNRNPA2B1 are prion domain proteins that bind TDP-43 and have been linked with some

ALS cases [122, 123]. Both proteins bind the L1 RNP, and HNRNPA1 colocalizes with ORF1 in SGs, as previously reported ([60, 124]; Fig. 4E). Wild-type TIA1, recently found mutated in cases of ALS and FTD [125], also strongly colocalizes with ORF1p in stressed cells as noted above (Fig. 1B). However, some ALS-related proteins, including OPTN and ANG (Fig. 4F, G), fail to colocalize in the same granules with ORF1p.

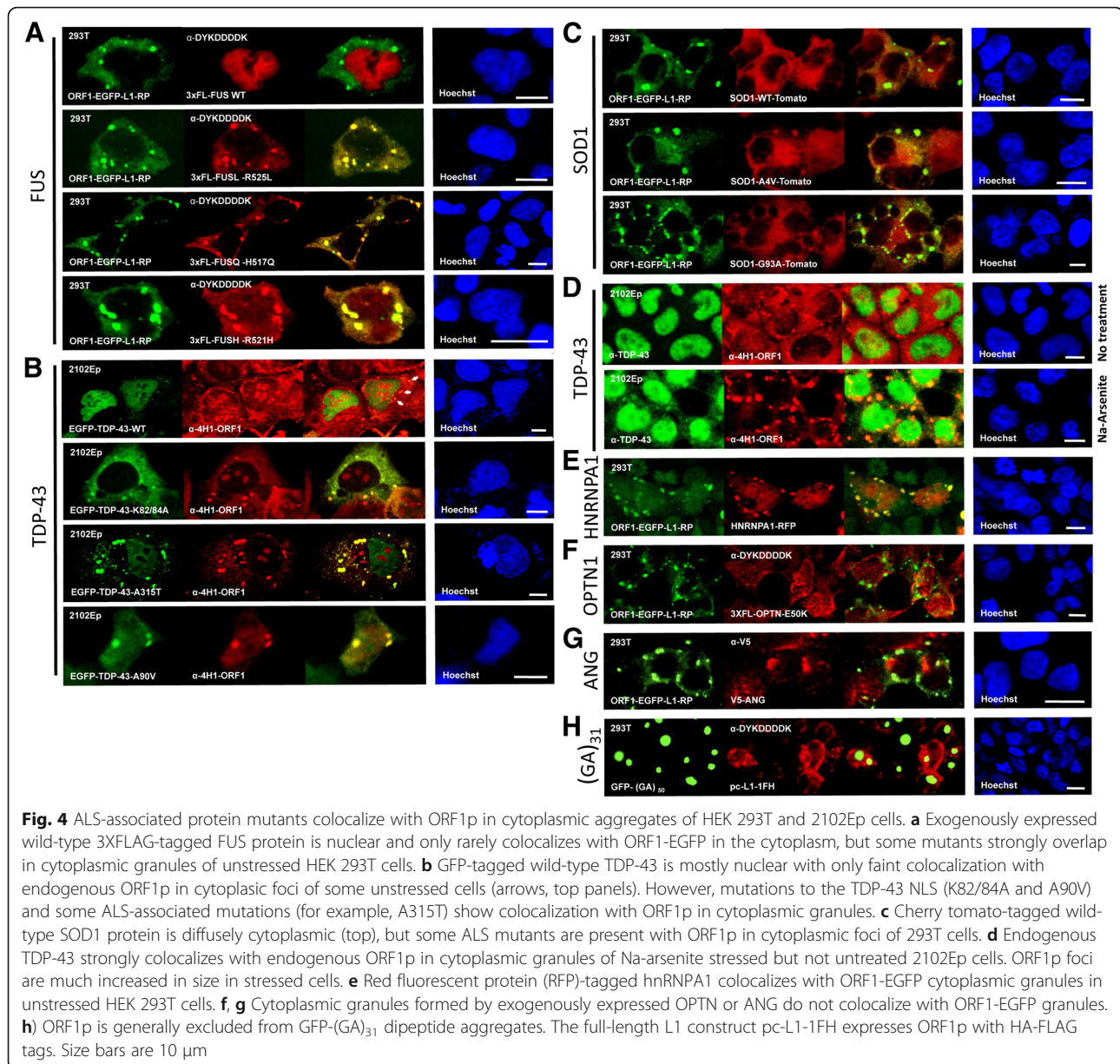
Expanded hexanucleotide repeats within transcripts of the *C9orf72* ALS gene can undergo non-conventional repeat-associated non-ATG (RAN) translation and generate dipeptide repeats that form inclusions in cerebellum, neocortex, and hippocampal neurons of C9 patients and toxic cytoplasmic aggregates in cultured neuronal cells or *Drosophila* models ([126–129], reviewed in [130]). To determine if these aggregates also colocalize with those of L1 ORF1p, we coexpressed in cultured cells a *C9orf72* RAN translation product consisting of 50 GA repeats tagged with EGFP [131] and full length L1 with FLAG-HA-tagged ORF1. However, while overexpressed dipeptide proteins formed one to three large cytoplasmic aggregates in each cell, these did not colocalize with and generally excluded ORF1p (Fig. 4H).

Thus, a subset of RNA-binding proteins mutated in ALS bind and colocalize with L1 ORF1p RNP in cytoplasmic RNA granules.

Overexpression of some ALS-associated proteins inhibit cell culture retrotransposition

We previously showed that retrotransposition occurs in non-dividing mature neurons [23]. Here we extended these analyses and tested retrotransposition in two cell lines often used to study neurodegeneration, human SH-SY5Y neuroblastoma and mouse NSC-34 neuronal cells, the latter a hybrid line generated by fusing spinal cord motor neurons with neuroblastoma cells (Fig. 5A). Due to inefficient plasmid transfection, we infected cells with the adenovirus-retrotransposon hybrid virus, A/RT-pgk-L1RP-EGFP (Ad-L1) [23, 97]. This viral construct contains L1-RP tagged with the EGFP retrotransposition reporter cassette. Retrotransposition detected by flow cytometry was 1.1% and 0.6% of gated SH-SY5Y and NSC-34 cells, respectively. Thus, both primary and transformed neuronal cell lines are competent for retrotransposition.

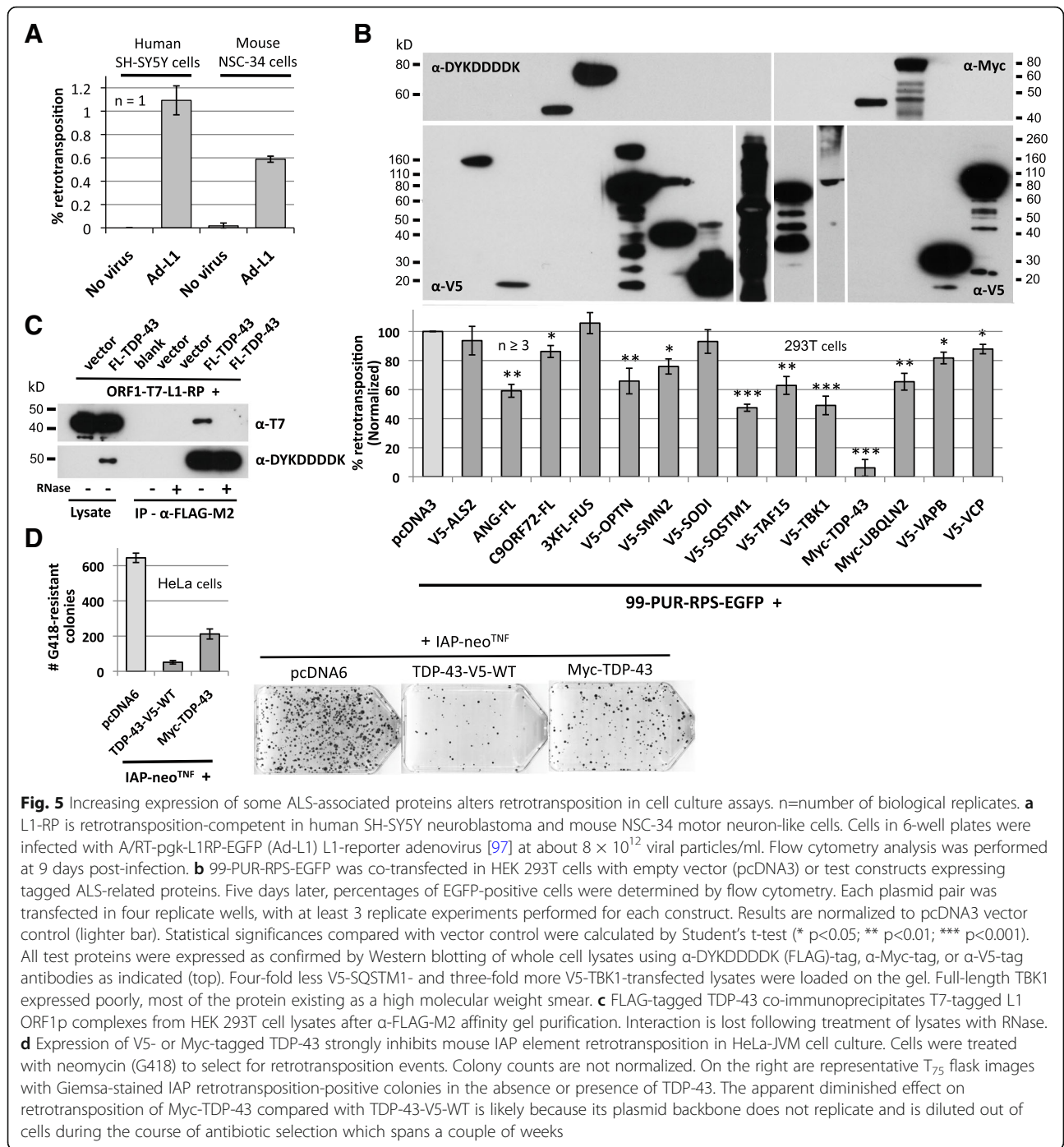
We next asked if ALS-related proteins alter L1 retrotransposition in the cell culture assay described above [84, 87]. Briefly, we transfected HEK 293T cells with the retrotransposition reporter construct 99-PUR-RPS-EGFP together with constructs expressing tagged ALS-related proteins. 99-PUR-RPS-EGFP includes full-length L1-RP with the EGFP reporter cassette in its 3' UTR cloned in a modified version of pCEP4 vector (Invitrogen) lacking a cytomegalovirus (CMV) promoter. All constructs were



expressed in HEK 293T cells (Fig. 5B, top). At least 3 biological replicates were performed. Test proteins did not cause significant cell death during the course of the experiment as determined by trypan blue exclusion staining (Additional file 1: Figure S4A). Three out of 14 proteins tested, including SQSTM1, TDP-43, and TBK1 kinase, reduced cell culture retrotransposition 50% or more when compared with cells transfected with empty vector only as control (Fig. 5B, bottom). SQSTM1/P62 is an autophagy receptor that targets bound proteins for selective degradation. Autophagy has previously been linked with retrotransposon restriction, and it was shown that SQSTM1 colocalizes with L1 RNA in stress granules, and that its knockdown causes increased

accumulation of L1 and Alu RNAs and genomic insertions in cultured cells [71]. Autophagy misregulation has also been linked with numerous neurodegenerative disorders, including ALS.

TDP-43 is a multifunctional RNA-binding protein with roles in mRNA transcription, translation, transport, splicing, and stability [132–134]. Studies in model organisms have shown that overexpression of wild-type TDP-43 mimics loss-of-function phenotypes of neurodegeneration and motor dysfunction [135, 136]. Several other studies have considered how endogenous TDP-43 levels affect expression of TEs but with inconsistent results ([49, 51, 52, 137–139]; see discussion). In the HEK 293T cell culture retrotransposition assay, ectopic expression of



TDP-43 with an N-terminal Myc-tag inhibits L1 retrotransposition over 90% (Fig. 5B). As this was the ALS-related gene that most altered retrotransposition levels, we next sought to characterize TDP-43 effect on L1 activity in more detail. To determine if ORF1p and TDP-43 interact, we co-expressed a construct containing L1-RP with T7-tagged ORF1 (ORF1-T7-L1RP) and TDP-43 with a C-terminal FLAG-tag; the two proteins co-immunoprecipitated on α -FLAG agarose (Fig. 5C). This association was

RNA-dependent and was lost upon treatment with RNase, similar to almost all other proteins previously identified within the L1 ORF1p RNP [73, 124, 140, 141].

Over-expression of TDP-43 is toxic to neurons and cell toxicity has been associated with increased cytoplasmic mislocalization of some TDP-43 mutant proteins [142, 143]. We therefore thoroughly tested for TDP-43-induced toxicity of HeLa or HEK 293T cells by three methods: 1) comparison of the effect of TDP-43

overexpression on constitutive expression of antibiotic resistance in HeLa cells (Additional file 1: Figure S4B, C), 2) trypan blue staining for cell viability in HEK 293T cells (Additional file 1: Figure S4D), and 3) MultiTox-Fluor Multiplex Cytotoxicity Assay kit (Promega) analysis in HEK 293T cells (Additional file 1: Figure S4E). Overexpression of TDP-43 had no significant effect on cell viability during the time course of our assays, indicating the drop in retrotransposition efficiency is not a reflection of cellular toxicity.

We next tested if overexpression of TDP-43 might also inhibit the mobilization of LTR TEs. Human endogenous retroviruses are thought to be incapable of replication due to the presence of inactivating mutations in their ORFs [4]. However, mouse intracisternal A particle (IAP) LTR retrotransposons actively replicate and cause new mutations by insertional mutagenesis. Using an established cell culture assay [144], we found that in HeLa cells overexpression of C-terminal V5- or N-terminal Myc-tagged TDP-43 strongly restricted retrotransposition of an IAP element tagged with a neomycin phosphotransferase reporter cassette (Fig. 5D).

In a reciprocal assay, we next asked whether loss of endogenous TDP-43 affects L1 cell culture retrotransposition. We confirmed by Western blotting that two different siRNAs efficiently repressed endogenous TDP-43 protein when transfected in HEK 293T cells. However, TDP-43 depletion had no obvious effect on L1 retrotransposition, at least using the EGFP-based retrotransposition assay (Additional file 1: Figure S5A). We note that an inherent limitation of these assays is the transient nature of the siRNA-mediated protein depletion.

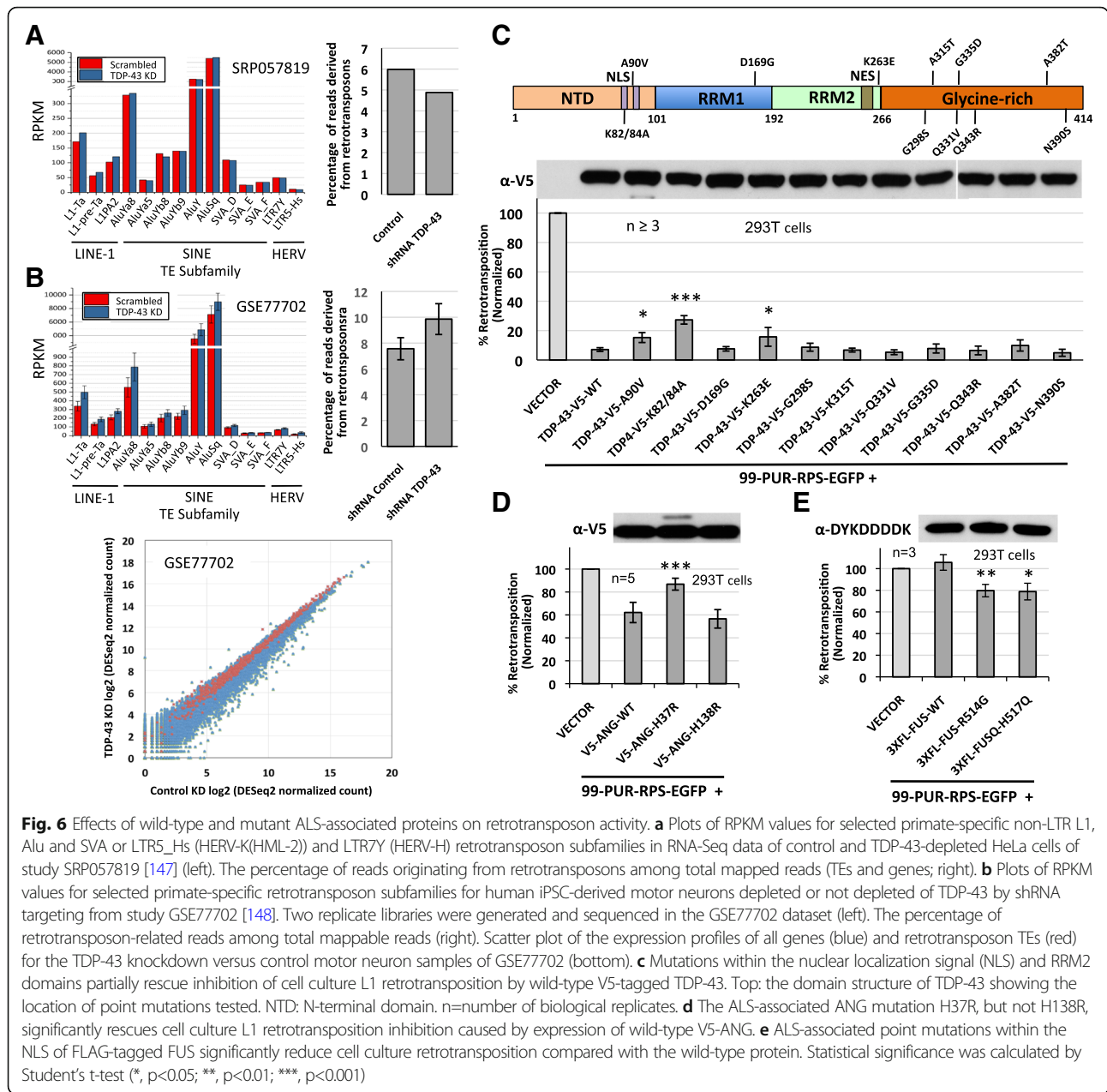
We also wondered if TDP-43 expression might affect the methylation status of the CpG island within the L1 5' UTR promoter [145]. We performed bisulfite conversion of genomic DNA from HEK 293T cells in which TDP-43 was either overexpressed (1 experiment; Additional file 1: Figure S6A) or depleted (2 independent experiments; Additional file 1: Figure S6B, C). PCR-amplified fragments containing the CpG island were cloned and at least 15 amplicons were sequenced for each sample [146]. Unexpectedly, when compared with controls, the 17 CpG residues of this region showed a significant overall increase in methylation in all experiments, although fully unmethylated sequences were found in all conditions. Therefore, one might speculate that perturbing steady-state TDP-43 protein levels alters DNA methylation status, a function for TDP-43 not to our knowledge previously reported. However, changes in L1 promoter methylation associated with TDP-43 expression were not accompanied by significant change in activity of either the L1 sense or antisense promoter in luciferase assays, at least using a plasmid-based assay (Additional file 1: Figure S6D). Moreover, TDP-43 overexpression failed to alter

levels of endogenous or ectopically expressed ORF1p in cell culture, as determined by Western blotting (Additional file 1: Figure S5B, C), nor consistently affected levels of endogenous L1 RNA in HEK 293T cells as detected by RT-PCR (Additional file 1: Figure S5D).

To complement these analyses, we next re-analyzed two available RNA-Seq datasets not previously examined for TE expression (Fig. 6A, B). The first study (SRA SRP057819) generated single-replicate paired-end 100-bp RNA-Seq data of control and TDP-43-depleted HeLa cells (using the same esiTARDBP siRNA shown in Additional file 1: Figure S5A) [147]. The second dataset (GEO GSE77702) includes single-end 50-bp RNA-Seq data (two replicates each) of wild-type human iPSC-derived motor neurons depleted by shRNAs of TDP-43, TAF15, FUS, or combined TAF15-FUS [148]. FUS and TAF15 are both members of the FET family of RNA-binding proteins, which are linked with several neurodegenerative disorders [149]. In the HEK 293T cell culture retrotransposition assay, overexpressing FUS had no effect, while TAF15 reduced retrotransposition by 38% (Fig. 5B).

To detect changes in TE expression, we plotted RPKM (Reads Per Kilobase of transcripts per Million mapped reads) values for a subset of mostly evolutionarily young primate-specific non-LTR retrotransposons, including L1, Alu and SVA, and LTR5_Hs (HERV-K(HML-2)) and LTR7Y (HERV-H) subfamilies. Study SRP057819 showed an increase (approximately 15%) in RPKM values for the TDP-43 knockdown (KD) versus control HeLa cell lines for L1s only (including L1PA2 and human-specific L1-Ta and L1-pre-Ta subfamilies) (Fig. 6A, left). However, there was a slight overall decrease in the percentage of retrotransposon (LINE, SINE, SVA, and LTR)-related RNA-Seq reads among the total number of mappable (gene and TE) reads in TDP-43 KD cells (Fig. 6A, right). For TDP-43-depleted motor neurons of study GSE77702, there was a modest but consistent increase in RPKM values for all retrotransposon subfamilies when compared with scrambled shRNA control (Fig. 6B, left). In addition, there was a modest increase from 7.6 to 9.8% in the percentage of retrotransposon RNA-Seq reads among total mapped reads (TEs and genes) for TDP-43 KD versus control motor neurons (Fig. 6B, right). A scatter plot, however, showed minimal change in the expression profile of all mapped TE subfamilies versus genes for the GSE77702 dataset (Fig. 6B, bottom).

We then analyzed the GSE77702 dataset with TETranscripts [150], a software package that uses short-read alignment files to identify differentially expressed (DE) TE subfamilies listed in RepBase, a database of representative repeat sequences in eukaryote genomes [151, 152]. A total of 192 retrotransposon subfamilies were expressed at significantly different levels in TDP-43 KD cells at an adjusted



P-value (padj) <0.05 (Additional file 2: Table S1). However, only 55 retrotransposon TEs were significantly DE in the TDP-43 KD but in neither the TAF15 nor FUS KD datasets. Notably, only three DE retrotransposon TEs (HERVK3-int, MamGyp-int, and MER51D) were unique to TDP-43 KD cells and absent in the TAF15, FUS, and combined TAF15-FUS KD groups.

Therefore, and in contrast to some reports (see discussion), our analyses do not indicate a clear TDP-43-specific link with elevated activity of TEs, particularly LINE-1 retrotransposons. In fact, overexpression of wild-type TDP-43 strongly inhibits cell culture retrotransposition of both human L1 and mouse IAP elements.

Mutation of some ALS-associated proteins alters cell culture retrotransposition

TDP-43 contains a NLS and NES, two RNA-recognition motifs (RRM1 and RRM2) that bind nucleic acids, and a C-terminal glycine-rich region that mediates protein interactions (Fig. 6C, top) [133]. A review in 2009 identified 70 pathogenic mutations in TDP-43, a majority in the glycine-rich domain [153]. We wished to determine if ALS-associated TDP-43 mutations might restore inhibition of retrotransposition by the wild-type protein, and so we tested the effects on L1 cell culture retrotransposition of a subset of mutant constructs. While all constructs expressed at levels similar to the wild-type,

most mutations had no significant effect (Fig. 6C, bottom). However, a non-ALS TDP-43 double-point mutation in the N-terminal bipartite NLS (K82/84A) restored retrotransposition 3.5-fold ($p < 0.001$). Similarly, ALS-associated NLS domain mutation A90V and RRM2 domain mutation K263E [154, 155] each rescued retrotransposition 2-fold ($p < 0.05$).

We also considered the effect of angiogenin mutations on cell culture L1 retrotransposition. ANG is a member of the pancreatic RNase A superfamily and a potent mediator of neovascularization, as well as being a host defense factor against some microorganisms [156] and an enhancer of motor neuron survival [157, 158]. To date 33 ANG mutations have been implicated in ALS and Parkinson's disease [159]. Overexpression of V5-tagged ANG protein reduced cell culture retrotransposition to 62% of empty vector control without obvious cytotoxicity (Fig. 5B, 6D, S4C). We then introduced two disease-associated mutations known to abolish ANG RNase activity [159–161]. Notably, mutation H138R (H114R in the mature protein after signal peptide cleavage) had no effect, while H37R (H13R) restored retrotransposition to 87% when compared with vector-only control (Fig. 6D).

Similarly, we examined the effect of mutations in FUS protein. Exogenous expression of wild-type FUS had no effect on cell culture retrotransposition nor obvious cytotoxicity, but ALS-related mutations in its C-terminal NLS (R514G and H517Q) inhibited retrotransposition over 20% (Fig. 5B, 6E S4D, E). Finally we tested the effect of mutations in TBK1 on L1 retrotransposition. TBK1 is a member of the IKK kinase family, and an important player in innate immune signaling. Mutations in TBK1 also impair autophagy [153]. Two mutants of TBK1, the kinase-dead mutant S172A and the ALS-associated mutant E696K [152] were tested, but neither showed any change from the 50% reduction of cell culture L1 retrotransposition caused by overexpression of the wild-type protein (Additional file 1: Figure S4F).

In sum, increased expression of some neurodegeneration-related proteins may decrease retrotransposon activity, while some disease-related mutations can modify these effects.

Retrotransposon expression in tissues of ALS patients and controls

To further determine if changes in expression of non-LTR class retrotransposons are associated with ALS, we performed RT-qPCR analyses of 108 bulk spinal cord and brain tissue samples of 38 ALS patients and 27 non-affected controls (Additional file 3: Table S2) according to methods described in [146]. We assayed 30 thoracic or cervical spinal cord samples (15 ALS, 15 controls), 16 cerebellum (9 ALS, 7 controls), 35 motor cortex samples (23 ALS, 12 controls), 19 occipital cortex

samples (14 ALS, 5 controls), and 8 hippocampal samples, all of the latter from ALS patients. Most samples were from sALS patients or patients of unknown etiology; only 5 patients had a known gene mutation. RT-qPCR primer pairs targeted the ORF1 and ORF2 regions of the young human-specific and retrotranspositionally active L1Hs subfamily (Additional file 1: Figure S7A, B) and two Alu subfamilies, AluS and AluY (Additional file 1: Figure S7C, D). Younger than AluJ elements, the AluS subfamily arose about 40 million years ago and may include some retrotransposition-competent elements [162, 163]. AluY, the youngest lineage, has the most retrotranspositionally active elements, and many genetic disorders in humans have been generated by AluY insertions [162, 164, 165]. Only L1Hs (L1P1)-type L1s are known to be retrotransposition-competent in the human genome [12, 166]. For the purposes of comparison, transcript levels were also determined for H9-hESCs [167], human embryonic fibroblasts (HEFs), and HeLa cells.

Transcript levels were determined in duplicate for each brain and spinal cord sample, normalized to *GAPDH* internal control, and averaged. We considered all measurements of sample-specific transcript levels as real and did not omit possible outliers from analyses. Averaged RT-qPCR reactions within each experiment were normalized to expression of H9-hESCs as these cells strongly express endogenous L1 RNAs [168]; means and standard deviations are shown in Additional file 1: Figure S7. As previously observed, H9-hESCs expressed 5 to 25 times more L1 RNA than differentiated cultured cells such as HEFs or HeLa [146]. Expression levels of Alu and L1 element-related sequences detected in tissue samples were as high or higher than in H9 cells. Average expression in cerebellum was 2- to >3-fold higher than in other tissues for both Alu subfamilies and for L1s (except for ORF1 in occipital cortex); however, transcript levels in cerebellum and for L1 ORF1 in all brain tissue regions varied considerably between samples. Comparing expression of Alu and L1 elements in ALS versus unaffected controls, only expression of AluS elements in occipital cortex was significantly elevated for the 14 ALS versus 5 control samples ($p = 0.02$) (Additional file 1: Figure S7C).

We next examined ORF1p expression by Western blotting of 60 brain and spinal cord tissue lysates (Additional file 1: Figure S8, Additional file 3: Table S2). There are very few studies of endogenous L1 protein expression in the brain. Baudin de The et al. [22] detected L1 ORF1p in ventral mid-brain tissues of mouse. Using a commercial antibody, Moszcynska et al. [169] showed by immunocytochemistry putative ORF2p expression in several rat brain regions, although antibody specificity was not assessed. Sur et al. [170] detected ORF1p by immunohistochemistry of various brain regions, and antibody

detection by Western blotting was confirmed for a single frontal cortex sample. Here, using Western blotting and the α -4H1-ORF1 antibody, we were, surprisingly, unable to detect an ORF1p band of appropriate size in frontal cortex, cerebellum, or hippocampal brain tissue samples, and only very faintly in some motor cortex samples, even when 50 μ g of whole cell lysate was loaded in a well (Additional file 1: Figure S8A-D) and despite the detection of L1-related RNAs expressed in these tissue types by RT-qPCR (Additional file 1: Figure S7). In contrast, we could detect a very robust full-length ORF1p signal from an equal amount of 2102Ep cell protein lysate (Additional file 1: Figure S8). Distinct bands consistent in size with full-length ORF1p were observed in some spinal cord samples; bands of smaller size were also seen, including a robust 38 kD signal of unknown origin (Additional file 1: Figure S8E). However, no overall differences in expression of ORF1p were evident in ALS compared with control spinal cord samples. Testing two different antibodies showed that failure to detect ORF1p signal in the brain was not limited to the α -4H1-ORF1 antibody (Additional file 1: Figure S8F, G).

In general, interpretation of TE expression from RT-qPCR data may be influenced by the presence of exonized TE-derived sequences in genes, the possible presence in the cell of non-integrated TE-derived cDNAs (see discussion below), and the cellular heterogeneity of the tissues analyzed. TE activation may occur in only a subset of cells within bulk tissue samples, so limiting sensitivity of detection, and in the case of motor neurons these cells may be progressively eliminated in the disease state. In sum, however, no major differences in TE expression were detected in ALS patients when compared with controls.

Retrotransposon expression in ALS RNA-Seq datasets

Prudencio et al [171] generated a paired-end total RNA-Seq data set (GSE67196) from cerebellum and frontal cortex samples of 9 healthy, 8 *C9orf72*-associated ALS (C9ALS), and 10 sALS individuals and analyzed these for differentially expressed genes. A subsequent re-analysis of the same datasets using the HOMER analyzeRepeats program revealed significantly increased global expression of repetitive element types in frontal cortex but not cerebellum of C9ALS compared with sALS patients and healthy controls [52]. Setting FDR<0.1, the authors reported 300 DE TE subfamilies in the C9ALS samples: LTR class elements predominated (46%), followed by DNA elements (19%) and LINES (18%). Notably, 91% of significant C9ALS DE repetitive elements had increased expression.

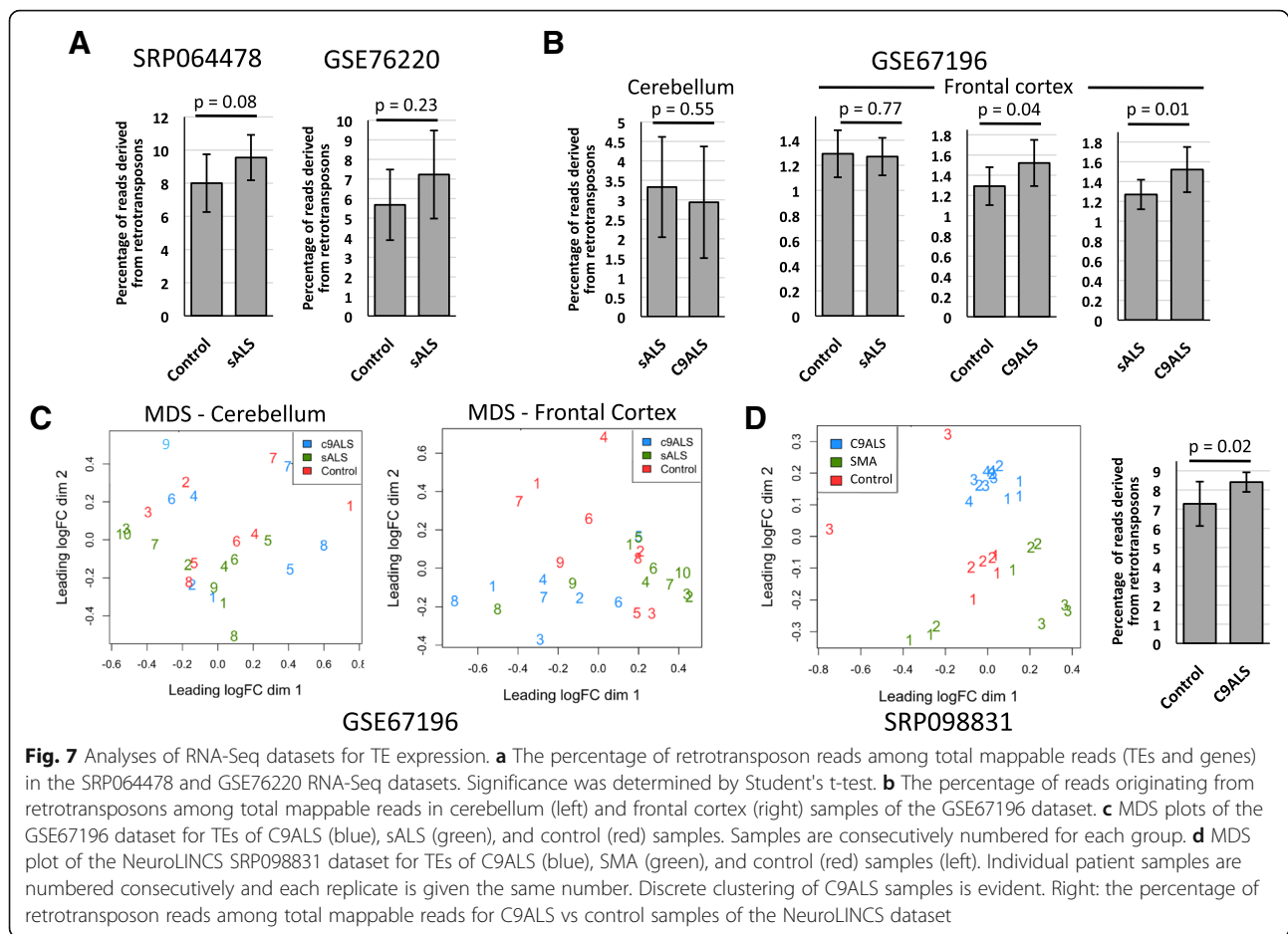
We sought to assess further the degree to which TEs are differentially expressed in sALS-associated tissues by

using Tetrascripts [150] to analyze two RNA-Seq datasets, SRP064478 and GSE76220, both publicly available in sequence read archives and neither previously examined for repeat expression (see also the Methods section). SRP064478 includes paired-end 150-bp sequence derived from total RNA of 7 sALS and 8 healthy control post-mortem cervical spinal cord samples. No significant difference in the percentage of averaged retrotransposon-derived reads among the total number of mappable reads (TEs and genes) was detected for ALS vs control samples; 1.72 to 3.41 million sample reads mapped to retrotransposon subfamilies (Fig. 7A, left; Additional file 2: Table S1). Only one significant DE retrotransposon subfamily was detected for SRP064478 ($p_{adj}<0.05$; Additional file 2: Table S1). However, overall mapping efficiency was low: about 25 million reads were alignable to the genome and 30% of these mapped to mitochondrial genes.

The GSE76220 dataset consists of single-end 50-bp sequence of total RNA isolated from laser-capture microdissected lumbar spinal cord sections of 13 sALS patients and 8 control individuals [172]. There was no significant change in the percentage of retrotransposon reads among total mappable reads in the sALS vs control samples (Fig. 7A, right; Additional file 2: Table S1). Between 0.27 and 1.26 million sample reads mapped to retrotransposons. Tetrascripts detected only four significant DE retrotransposon subfamilies (Additional file 2: Table S1).

We then reanalyzed the GSE67196 RNA-Seq dataset of Prudencio et al. [52, 171] using Tetrascripts. Significant increases in retrotransposon reads as a percentage of total mapped reads were seen for frontal cortex C9ALS vs control ($p=0.04$) and C9ALS vs sALS ($p=0.01$) samples (between 0.24 and 0.61 million sample reads mapped to retrotransposons) (Fig. 7B). As expected, multidimensional scaling (MDS) plots showed weak clustering of C9ALS samples in frontal cortex but not cerebellum samples (Fig. 7C). Tetrascripts detected no significant DE TEs in cerebellum samples of the GSE67196 dataset ($p_{adj}<0.05$). In the case of the frontal cortex samples however, and supporting Prudencio et al. [52], there were 3 DE TEs (DNAs, LTRs, LINES, SINEs, and SVAs) in sALS vs controls, 10 DE TEs in C9ALS vs controls, and 133 DE TE subfamilies in C9ALS vs sALS samples, all increased in expression and including 36% LTR, 32% DNA, 15% LINE, and 17% SINE elements (Additional file 2: Table S1).

We also analyzed for the first time TE expression for the NeuroLINCS dbGaP Study phs001231 (SRP098831). This dataset consists of poly(A)⁺ non-stranded mRNA of iPSC-derived motor neurons from 4 C9ALS and 3 SMA patients (3 sequencing replicates each) and 3 unaffected controls (2 or 3 replicates each). Transcripts of



some TE types are not polyadenylated and so are likely underrepresented in this dataset following poly(A)⁺ selection. However, although Alu elements are transcribed by RNA polymerase III and not polyadenylated, they contain both internal and 3'-end poly(A) stretches guaranteeing capture of their transcripts. An MDS plot showed C9ALS samples clustered away from SMA and control samples, while clustering of SMA from control samples was less evident (Fig. 7D, left). There was a significant increase ($p=0.02$) in the percentage of retrotransposon reads among total mappable reads in the C9ALS vs control dataset (Fig. 7D, right). The increase was also significant for C9ALS vs SMA ($p=0.002$) but not significant for SMA vs control samples ($p=0.46$) (not shown). Tetrascripts analysis showed that at $\text{padj}<0.05$, significant DE TE subfamilies (DNAs, LTRs, LINEs, SINEs, and SVAs) numbered 536 for C9ALS vs controls, 232 for C9ALS vs SMA, and 304 for SMA vs controls, most TEs being increased in expression (Additional file 2: Table S1). Three SVA and 30 Alu TEs were upregulated for C9ALS vs controls, including 6 AluY subfamilies. (The human-specific L1Hs/L1P1 subfamily was not detected). Interestingly, a recent literature review noted

at least 37 neurological and neurodegenerative disorders linked with misregulated Alu retrotransposon activity [173]. A caveat of this dataset analysis is that sample numbers were small.

Algorithms such as Homer and Tetrascripts map sequencing reads to TE consensus sequences only and locus-specific information is lost. The ability to map individual transcribed retrotransposons to their source loci can reveal (i) the particular loci that contribute to repeat family transcription differences between diseased and healthy states, (ii) the coding capacity of transcribed repeat loci of possible relevance for a specific disease, and (iii) potentially variant retrotransposon proteins and RNAs that should be considered when studying disease relevance. We therefore applied a recently developed locus-specific TE mapping pipeline (PT, EP, DT, unpublished data), as described in the Methods section, to re-analyze the GSE67196 data set [171]. The numbers of mapped loci are summarized in Additional file 1: Figure S9A. Principal component analysis (PCA) and heatmap plots again showed cerebellum TE expression to be as variable within C9ALS, sALS, and control groups as between groups and without significant clustering

(Additional file 1: Figure S9B, C, left). However, as expected, clustering of frontal cortex C9ALS samples distinct from control and sALS samples was evident (Additional file 1: Figure S9B, C right). Supporting Prudencio et al. [52], the greatest number of DE TE loci (determined as having $\text{padj} < 0.05$ and greater than 2-fold differential transcription) were identified for C9ALS vs sALS (3963 loci), followed by C9ALS vs controls (652 loci), and sALS vs controls (109 loci). However, these DE TE loci comprised only 1.8%, 0.3%, and 0.06%, respectively, of a total of 2.12×10^5 TE loci mapped (Additional file 4: Table S3).

Caveats of this type of analysis should be noted. For example, most of the significant DE L1 loci were likely not transcribed from their own promoters, since 1) almost 95% of those mapped were less than 5600 bp in length and so lacked much of their 5' UTRs, and 2) 70% of DE L1s were within genes, and so may be transcribed as part of a longer gene transcript. Moreover, only 2 younger primate-specific L1P1 and 4 L1PA2 L1 loci were differentially expressed (among a total of 164 L1P1 and 135 L1PA2 individual elements mapped). As for DE Alu loci, 81% were within genes (98% of all Alu loci being upregulated). Furthermore, most mapped Alu loci were older elements, with only 5% of them AluY subfamily members, a bias likely due to the inability of currently available algorithms to confidently map short sequence reads to young highly similar TEs. In general, designing RNA-Seq analysis pipelines that efficiently map short sequence reads of young highly similar TEs to their source loci has to date been difficult for reasons discussed below.

In summary, RNA-Seq analysis of the SRP098831 NeuroLINC dataset suggests widespread upregulation of TE sequences from numerous subfamilies in *C9orf72* ALS patients, as previously reported for the GSE67196 dataset [52]. However, additional locus-specific analysis of the GSE67196 dataset suggests that many loci contributing mappable reads were not autonomously transcribed from their own promoters and were likely part of longer gene transcripts. More detailed transcription analyses targeting a selected cohort of full-length intact intergenic TE loci are needed to validate misregulation of retrotransposon expression in *C9orf72*-associated ALS disease.

Discussion

Self-aggregation of RNA-binding proteins is a leitmotif of neurodegenerative diseases, including amyotrophic lateral sclerosis. The ORF1 protein of LINE-1 retrotransposons is also an aggregation-prone RNA-binding protein. Of the approximately 500,000 L1s in the human genome, about 5000 are full-length, or about one percent of DNA [10, 174]. Many of these L1s have the potential to be transcribed and translated, although different tissues may express different L1s [175]. We

speculate that misregulation of even a small number of these, leading perhaps to mislocalization and augmented aggregation of ORF1p, could have negative effects on some cells, including neurons. In this study, to increase our understanding of the role of ORF1p in disease, we extended previous investigations of its subcellular localization and aggregation properties. We then considered potential interactions of amyotrophic lateral sclerosis-related proteins and the ORF1p RNP and the possibility of misregulated L1 activity in the ALS state.

Analogous to gene products associated with certain neurodegenerative diseases, L1 ORF1p RNPs are prone to forming cytoplasmic RNA granules. In unstressed cells of some lines, ORF1p constitutively forms cytoplasmic granules that are only faintly and partially marked by canonical stress granule proteins. As shown in Fig. 1, stress to the cell increases both the size of ORF1p cytoplasmic aggregates and colocalization with SG proteins, and deleting a Q-N-rich region of human ORF1p abolishes aggregation. Furthermore, we showed here that L1 elements with a variant ORF1 R159 codon, a residue that controls both retrotransposition and the ability of ORF1p to seed cytoplasmic RNA aggregates, are common in the human genome. Thus, cell stress promotes and certain sequence polymorphisms alter cytoplasmic aggregation of L1 ORF1p.

A functional role for L1-associated cytoplasmic RNA granules in retrotransposition remains unknown. This begs the question, what are these constitutively expressed ORF1p aggregates? ORF1p fails to associate with Golgi, lysosome, or endoplasmic reticulum marker proteins [60, 176, 177]. Endogenous ORF1p aggregates in unstressed cells occasionally abut P-bodies but generally do not overlap ([60]; Additional file 1: Figure S1F). Guo et al. [71] found that exogenous and endogenous ORF1p colocalized with autophagosome marker LC3 protein in HEK 293T cells, an association that increased with inhibition of autophagy. We here confirmed that endogenous ORF1p granules of unstressed 2012Ep cells are also partially marked by red fluorescent protein (RFP)-tagged LC3 (Additional file 1: Figure S1G), but we failed to detect their colocalization with endogenous autophagy marker proteins ATG12 or ATG16L1 (Additional file 1: Figure S1H, I). It has also been reported that ORF1p co-IPs and colocalizes in some cytoplasmic granules with IGF2BP1/IMP1 [124], part of a multi-protein complex found in granules of neuronal axons [178–180]. IMP1 granules have been reported as distinct from P-bodies and stress granules [181]. It is therefore possible that ORF1p aggregates in more than one type of cytoplasmic structure.

Endogenous ORF1p may concentrate as perinuclear, nuclear, or nucleolar. We showed in Fig. 2 that in some cells ORF1p also forms small nuclear foci distinct from

nucleoli; ectopically expressed Alu and likely other RNAs colocalize with these foci. Other studies have also reported endogenous ORF1p nuclear localization, specifically in some human cancer cell lines and tissues [60, 76, 77, 79, 182], and murine germline, chloroleukemia, and cardiomyocyte cells [183–187]. Why ORF1p is cytoplasmic in some cells and nuclear in others is unclear but suggests a dynamic aspect of ORF1p biology that is starting to be appreciated in the retrotransposon field [100].

We therefore examined the cell cycle as one possible mechanism controlling ORF1p subcellular localization. Blocking 2102Ep cell cycling at G₁/S phase transition did not obviously alter ORF1p nuclear localization. Non-blocked cells showed significant concentration of ORF1p in nucleoli, whether the cells were in G₁ phase or not. Our results in part contradict a recent study proposing a strong cell cycle bias for ORF1p accumulation in the nucleus during mitosis where it remains during G₁ phase [100]. This discrepancy may in part be due to the fact that we queried strictly nucleolar accumulation as the most obvious feature of ORF1p nuclear localization in 2102Ep cells; failure to observe ORF1p in nucleoli does not necessarily preclude its diffuse presence in the nucleoplasm. Also, the Mita et al. study [100] tested a different cell line (HeLa) and ORF1p overexpressed from plasmids, while here we examined endogenous ORF1p localization as more biologically relevant. Indeed, 2102Ep cells mimic early human embryogenesis, where heritable L1 insertions accumulate [69, 168], while HeLa cancer cells mimic L1 activity in human cancers. While both cellular niches are known to support L1 retrotransposition, it stands to reason that differences may exist with respect to L1 regulation. Thus, the mechanisms that control L1 ORF1p nuclear localization clearly require further investigation in a range of cell lines and with care paid to their growth conditions (for as noted above, the frequency of ORF1p nucleolar localization varies with 2102Ep cell density).

Previously, we showed that ORF1p point mutations and C-terminal and N-terminal deletions increased nuclear accumulation (see text and Supplemental data of [60]). Thus, maintaining the integrity of ORF1p structure seems to be important for cytoplasmic retention and aggregation. Furthermore, concentration of ORF1p at the nuclear membrane of some cells (Fig. 2), its reported RNA-dependent association with karyopherin subunit alpha 2 (KPNA2; [124]), and the detection by mass spectrometry of importin 7 (IPO7) within an ORF2p complex [141], suggest that L1 RNPs interact with the nuclear import machinery. Indeed, it was recently shown that loss of transportin 1 (TNPO1), the beta subunit of the karyopherin receptor complex, reduces nuclear localization of epitope-tagged ORF1p [188]. It was also recently proposed that ORF1p

expression is required for nuclear ORF2p localization [141]. However, this is not supported by our earlier findings that ORF2p overexpressed alone efficiently enters nucleoli of human osteosarcoma 143B TK- cells; at that time we also mapped a functional nuclear localization signal to the N-terminus of ORF2p [76].

Might there be cellular consequences for misregulated expression or mislocalization of aggregation-prone LINE-1 proteins? L1 ORF1p is a promiscuous RNA-binding protein able to capture many cellular RNAs. Co-IP experiments with tagged L1 RNPs have identified numerous bound RNAs, including SVA and Alu SINEs and small non-coding RNAs of importance for the cell [124, 189]. Direct co-IP experiments also confirmed over 60 proteins that associate with tagged L1 ORF1p RNPs, mostly in an RNA-dependent manner [60, 73, 124, 190]. Among these were several RNA-binding proteins associated with ALS and FTLN, including FUS, HNRNPA1, HNRNPA2B1, and TDP-43. As we have shown here, pathogenic mutants of ALS proteins FUS, SOD, and TDP-43 also colocalize with ORF1p in cytoplasmic RNP aggregates. As with certain neurodegeneration-associated proteins, increased expression or mislocalization of ORF1p, whether through mutation or loss of L1 suppression, could seed protein aggregation, co-sequester other cellular proteins or RNAs, disrupt normal patterns of protein degradation or RNA processing, and trigger cytotoxicity. Retrotransposon-encoded proteins can also induce cellular stress responses. Gasior et al. [191] showed that overexpression of L1 ORF2p causes double-strand chromosome breaks. These results are consistent with observations that L1 overexpression can induce apoptosis and senescence or potentially an immune response in some cell lines [192–195]. Perhaps these are reasons for the evolution of so many cellular factors that restrict L1 activity [42].

Previous studies have considered links between TDP-43 and retrotransposon expression. TDP-43 was first identified as a transcriptional repressor that binds the RNA regulatory element TAR of HIV-1 proviruses to inhibit their expression [196]. However, a role for TDP-43 in regulating HIV or HERV transcription is not clear [197]. Douville et al. [49] found expression of HERV-K(HML2) *pol* and *TARDBP* genes to be strongly and positively correlated, and their encoded proteins colocalized in ALS neurons. Douville and Nath [198] also linked TDP-43 with altered HERV-K(HML-2) RT expression in brain tissues. Data-mining rodent and human interaction experiments, Li et al. [137] found that TDP-43 protein targets and binds LTR and non-LTR TE transcripts and that this association is reduced in cortical tissues of FTLN patients. Furthermore, reanalysis of RNA-Seq datasets of human TDP-43 overexpressed in transgenic mice [199] and endogenous TDP-43

depleted in mouse striatum [200] showed a general increase in expression of LTR, non-LTR and DNA TEs under both conditions, with concordance between TE transcripts upregulated and those bound by TDP-43 protein [137]. While it was reported [51] that TDP-43 protein bound the HERV-K LTR with an attendant increase in HERV-K(HML-2) transcription and RT activity, Manghera et al. [138] found wild-type TDP-43 bound the HERV-K(HML-2) promoter without activating its transcription, while overexpressed ALS-associated TDP-43 mutants promoted HERV-K(HML-2) protein aggregation and clearance from astrocytes (but not neurons) by stress granule formation and autophagy. Overexpression of a human TDP-43 transgene in *Drosophila* was accompanied by motor problems and derepression of retrotransposons in general and glial cell-specific upregulation of gypsy elements in particular, along with an increase in programmed cell death induced by DNA-damage [139].

We found that, while TDP-43 binds and colocalizes with the L1 ORF1p RNP, its increased expression strongly represses rather than derepresses human L1 and mouse IAP cell culture retrotransposition (without attendant cytotoxicity). On the other hand, inhibition of endogenous TDP-43 had no effect on L1 retrotransposition in HEK 293T cells. Although altered levels of TDP-43 were associated with modestly increased methylation of endogenous L1 promoters, this was not accompanied by a change in exogenous ORF1p expression or promoter effects in a luciferase assay. Moreover, a recent study by Prudencio et al. [52] found no significant association between levels of TDP-43 RNA or protein and TE expression in frontal cortex samples of a large cohort of ALS/FTLD patients. Our reanalysis of two RNA-Seq datasets [147, 148] also failed to detect strong TDP-43-specific changes in expression of retroelement subfamilies in cell lines depleted of endogenous TDP-43 protein. Therefore, a role for TDP-43 protein in aberrant retroelement activity begs further investigation.

We also reanalyzed for TE subfamily expression two RNA-Seq datasets of sALS tissue samples not previously examined for TE expression (GSE76220 and SRP064478) and one previously tested dataset of both *C9orf72* and sporadic ALS tissue samples (GSE67196). In all three datasets, we failed to find significant misregulation of TE subfamilies in sALS vs controls, consistent with the previous analysis of GSE67196 [52] and with our RT-qPCR and Western blot analyses of ALS and control brain and spinal cord tissues. However, our analysis of a NeuroLINC dataset (SRP098831) found both SMA and C9ALS vs non-ALS patient-derived iPSC cell lines differentiated to motor neurons to have significant numbers of DE TEs, including young SINE Alu subfamilies: this was in line with the previous findings of Prudencio et al. [52] that TE expression is misregulated in C9ALS vs

sALS samples of the GSE67196 dataset. However, our locus-specific analysis of the GSE67196 dataset suggested that many of the reads contributing to the retrotransposon subfamily analyses did not originate from TE sequences transcribed from their own endogenous promoters but rather from sequences contained within longer transcripts.

Several pitfalls exist for RNA-Seq analyses of differential TE expression: conclusions should be drawn with care. High copy number, close sequence similarity, and especially the frequent embedment of TE sequences in longer gene transcripts (i.e., exonization) can lead to misinterpretation. While expression of a TE subfamily may appear misregulated, a change in expression observed may in fact be due to altered expression of a gene in which a member of that TE subfamily resides. In their analysis of RNA-Seq data from HEK 293T cells, for example, Deininger et al. [174] mapped greater than 99 percent of L1-derived sequence reads within other RNAs unrelated to retrotransposition. Moreover, *bona fide* L1 transcripts originating from L1 5' UTR promoters were limited to only a small number of highly expressed full-length L1 loci.

Furthermore, we have speculated that cell conditions that induce elevated expression of L1s or HERVs, and therefore their encoded reverse transcriptases, could induce promiscuous reverse transcription of cytoplasmic RNAs ([42], see also [201]). Indeed, a recent report has demonstrated the accumulation of cytoplasmic L1-related ssDNAs in neurons derived from hESCs lacking the exonuclease TREX1, a gene mutated in Aicardi-Goutières syndrome patients [195, 202]. The ectopic cytoplasmic cDNAs so generated would be amenable to amplification during RNA-Seq or RT-qPCR protocols and so bias upwards estimates of expression from their source loci. Although as yet an unverified concern, recent studies have reported elevated levels of TE-derived cytoplasmic cDNAs or hybrid RNA/DNA molecules in cancer and other disease states [195, 201, 203]. The retrotransposon field is working to control such possible sources of error when interpreting TE expression data [174, 175]. In general, further improvements in transcriptomics, and especially single-cell based approaches, will eventually clarify the degree of deregulation of TE expression in ALS and other neurodegenerative disorders.

Conclusions

In considering links between retrotransposon expression and neurodegenerative conditions, we expanded previous knowledge of the aggregation properties of the LINE-1-encoded ORF1 protein and factors that control its accumulation. We also presented data that the cell cycle does not strongly alter nuclear localization of endogenous L1 ORF1p in nullipotent embryonal carcinoma cells. We showed that some ALS-associated protein

mutants associate with ORF1p in cytoplasmic aggregates and that increased expression of some ALS-linked proteins limit LINE-1 retrotransposition. We emphasized especially TDP-43, a protein that accumulates in the cytoplasm of a majority of ALS patients, but failed to find consistent evidence in cell culture for an effect on retrotransposon activity, in contrast to some previous reports.

By means of RT-qPCR and Western blotting of ALS tissues and reanalysis of available RNA-Seq datasets, we also sought a link between sporadic ALS and retrotransposon misregulation. In sum, clear-cut evidence is so far lacking for involvement of non-LTR retrotransposon expression in sALS. Using the same tissue samples as in the present study, we also recently profiled transcription of HERV-K(HML-2) and HERV-W LTR retrotransposons by direct Sanger sequencing of cloned cDNAs and RT-qPCR and Western blot analyses, but failed to find significant differences when comparing ALS and controls [204]. It is conceivable that previous observations of differential TE expression levels may relate to altered global DNA methylation status and other epigenetic changes observed in some ALS patients [205–208], which could in consequence cause selected TE loci to be differentially transcribed. At least, analyses of additional *C9orf72*-mutated ALS RNA-Seq datasets seem warranted. We believe examining neurodegenerative disease-affected tissues for perturbations in the aggregation dynamics of L1-encoded proteins could also prove informative. It is also reasonable to continue to apply improving methods of next-generation sequencing analysis to examine neurodegenerative and other brain diseases for misregulated activity of TEs in general and the L1 in particular, a mobile element with hundreds of thousands of copies and which through long evolution has been directly responsible for generating over a quarter of the DNA in the human genome [209].

Methods

Plasmid and RNAi constructs

Plasmid constructs were kindly provided by the following researchers: 3xFL-FUS-WT, 3xFL-FUSL-R525L, 3xFL-FUSH-R514G, 3xFL-FUSQ-H517Q, and 3xFL-FUSQ-R521H (J. Manley, Columbia University; [210]); pcDNA5 FRT/TO (G4C2)₃₁ (M. Cozzolino, University of Rome "Tor Vergata"; [96]); GFP-(GA)₅₀ (L. Petrucelli, Mayo Clinic, Florida; [131]); IAP-neo^{TNF} (M. Dewannieux, Institut Gustave Roussy; [144]); pmRFP-LC4 (T. Yoshimori, Addgene 21075); pEGFP-C1-MK2-mut T205/317E (M. Gaestel, Medical School Hannover, Germany; [111]); pcDNA3 3XFL-OPTN-E50K (H. Kawakami, Hiroshima University; [211]), mCherry-PSP1 (D. Spector, Cold Spring Harbor Laboratory; [212]); EGFP-TDP-43-WT,

EGFP-TDP-43-A315T, EGFP-TDP-43-K82/84A (B. Wolozin, Boston University; [213]); pRK5-Myc-TDP-43 (J. Wang, Johns Hopkins University, [214]); SOD1 WT-Tomato, SOD1 A4V-Tomato, SOD1 G93A-Tomato (J. Yerbury, University of Wollongong; [215]), pCS2 (+)MT UBQLN2 WT2 (Myc-UBQLN2), and pCS2 (+)MT UBQLN2 P497H (D. Ito, Keio University School of Medicine; [216]). The following plasmids have been previously described: 99-PUR-RPS-EGFP, 99-PUR-JM111-EGFP [87], ORF1-T7-LIRP, ORF1p-EGFP, HNRNPA1-RFP [60], ORF1-EGFP-L1-RP, pBS 7SL Alu-MS2 (Ya5), pcDNA SVA_{SPTA1}-MS2, ORF1-WT-GFP [72], and pc-L1-1FH [124].

Ultimate ORF cDNA clones (Invitrogen) were cloned with V5-epitope tags and tobacco etch virus (TEV) protease cleavage sites on their N-termini by shuttling them from pENTR221 vector into pcDNA3.1/nV5-DEST vector using Gateway Technology (Invitrogen). Ultimate ORF Clone ID numbers were V5-ALS2 (IOH 62502), V5-ANG (IOH29453), V5-OPTN (IOH57143), V5-SMN2 (IOH10903), V5-SOD1 (IOH4089), V5-SQS TM1 (IOH5103), V5-TAF15 (IOH40855), V5-TBK1 (IOH21006), V5-VAPB (IOH4934), and V5-VCP (IOH 52832). FL-TDP-43 was generated using Ultimate ORF Clone IOH45677 and Gateway vector pEZYflag (Y.-Z. Zhang, Addgene 18700). TDP-43-V5-WT and TDP-43-FL were generated by PCR-amplification of TDP-43 from pRK5-Myc-TDP-43 with AAG linker and C-terminal V5- and FLAG-tags, respectively, and cloned in pcDNA6/myc-His B (pcDNA6, Invitrogen). C9ORF72-FL with AAG linker and C-terminal FLAG tag was amplified from Ultimate ORF Clone IOH45695 and cloned in pcDNA6/myc-His B.

The helper-dependent adenovirus for construct A/RT-pgk-L1RP-EGFP (Ad-L1) [97, 217] was prepared as described [218]. N1-EGFP was from Clontech. The pc6-RPS-EGFP-ΔCMV retrotransposition reporter construct included full-length L1-RP and the EGFP reporter cassette in a modified version of pcDNA6/myc-His B vector lacking a cytomegalovirus (CMV) promoter. PCR and QuickChange mutagenesis methods were used to generate ORF1-EGFP-L1-RP-R159H (from ORF1-EGFP-L1-RP), ORF1-EGFP-L1-RP-LKEA AAAAL, ORF1p-Δ179-205-GFP (from ORF1p-EGFP), LKELLKEAL-EGFP (from N1-EGFP), V5-ANG-H37R and V5-ANG-H138R (from V5-ANG), V5-TBK1-S172A and V5-TBK1-E696K (from V5-TBK1), and pc6-RPS-EGFP-ΔCMV-R159H (from pc6-RPS-EGFP-ΔCMV).

siRNAs (Additional file 1: Figure S5A) were synthesized or purchased from Sigma-Aldrich:

siCNT2: CUCCCGUGAAUUGGAAUCC[dT][dT]

siCNT3: AUGUAUUGGCCUGUAUUAG[dT][dT]

siCNT4: UAAGGCUAUGAAGAGAUAC[dT][dT].

2siTDP43: AUAUCCAUAUGCACCACC[dT][dT]

esiTARDBP: EHU109221 MISSION esiRNA.

Cell culture and tissues

Human 2102Ep embryonal carcinoma cells (a gift from P.K. Andrews, University of Sheffield), human cervical cancer HeLa-JVM cells [219], human embryonic kidney (HEK) 293T cells (ATCC), human embryonic fibroblasts (HEFs, ATCC), and mouse hybrid motor neuron NSC-34 cells (a gift from D. Griffen, JHU) were grown in Dulbecco's modified Eagle's medium (DMEM). SH-SY5Y cells (ATCC CRL-2266) were grown in DMEM/F12 (Ham) medium (Gibco) and human SK-N-SH cells (a gift from D. Valle, Johns Hopkins University) were grown in Eagle's Minimum Essential Medium. Medium was supplemented with 10% FBS (Hyclone or Sigma), GlutaMax, and Pen-Strep (Invitrogen). Plasmid and siRNA transfections used FuGENE HD (Promega) or Lipofectamine 3000 (Thermo Fisher Scientific) reagents. H9 human ESCs [167] were obtained from Wicell (RRID: CVCL_9773) and cultured and passaged as previously described [168].

Post-mortem brain and spinal cord frozen tissues were obtained from the University of Maryland Brain and Tissue Bank of the NIH NeuroBioBank, the Target ALS Multicenter Postmortem Tissue Core at Johns Hopkins University, and the Department of Neurosciences of the University of California San Diego School of Medicine, as indicated in Additional file 3: Table S2. All tissues were obtained following approval of the Institutional Review Boards of the UCSD School of Medicine (to JR) and the JHU School of Medicine (IRB00066246 to JLG).

Immunofluorescence and microscopy

Commercial antibodies included rabbit (rb) α -ATG12 (D88H11), rb α -ATG16L1 (D605), rb α -CDT1 (D10F11), rb α -cyclin B1, mouse (ms) α -DYKDDDDK (FLAG)-tag (9A3) (all Cell Signaling Technology), goat (gt) α -eIF3 η (N-20, Santa Cruz), rb α -Geminin (Cell Signaling Technology or ab195047, Abcam), ms α -p70 S6 kinase (which recognizes HEDLS/EDC4; [220]) (H-9, Santa Cruz), rb α -Myc-tag (71D10, Cell Signaling Technology), ms α -TDP-43 (10782-2.AP, Proteintech), gt α -TIA1 (C-20, Santa Cruz), ms α -T7-Tag (Novagen), ms α -TLS (FUS) (BD Transduction Laboratories), rb α - β -tubulin-2 (Pierce), and ms α -V5-tag (Invitrogen). Purified ORF1p antibodies included rb polyclonal α -ORF1p-AH40.1 (4292) (a gift from M. Singer, Carnegie Institute for Science; [68]), rb monoclonal α -JH73-ORF1 (from J. Han, Tulane University and K. Burns, Johns Hopkins University; [177]), rb polyclonal α -V14-ORF1 (a gift from C. Harris, The Verto Institute; [182]) and ms α -4H1-ORF1 (from K. Burns, and Millipore MABC1152; [67]). Human α -ANA-N was obtained from a patient with autoimmune disease [76]. Donkey Cy3-, DyLight 488-, DyLight 549-, Alexa Fluor 594-, and

peroxidase-conjugated secondary antibodies were from Jackson ImmunoResearch Laboratories.

Western blotting, IF, and RNA FISH were performed as described [60, 72]. All Western blots were run on NuPAGE 4-12% Bis-Tris gels (ThermoFisher). Cells were examined using a Nikon Eclipse Ti-A1 confocal microscope with NIS-Elements AR software.

Whole-cell protein and RNA extraction

For protein extracts, tissues or cells were lysed in RIPA buffer (Sigma) with Mammalian Protease Inhibitor Cocktail and phenylmethanesulfonyl fluoride (Sigma) and homogenized with a Diagenode Bioruptor. In the case of tissues, 2 mm zirconium silicate beads (Next Advance) were added to the tubes. Samples were centrifuged at 11K at 4°C for 15 minutes to recover supernatant and resuspended in 3X SDS loading buffer. Isolation of HEK 293T cell nuclear and cytoplasmic extracts utilized the NE-PER kit (Thermo Scientific).

For RNA extracts, all brain tissue and some spinal cord tissues were disrupted and homogenized in 500 μ l of Trizol (Invitrogen) using the TissueLyser LT (Qiagen). Briefly, 30 mg of sample were transferred to a 2 ml tube containing 250 μ l of Trizol and one 5 mm stainless steel bead. The TissueLyser LT program used was 50Hz for 1 min. After a spin, the supernatant was collected and another 250 μ l were added to the sample to repeat the same procedure. Finally, both fractions were combined and RNA purification with Trizol followed the manufacturer's instructions. Some spinal cord samples were homogenized in 500 μ l of Trizol and zirconium silicate beads using a Benchmark BeadBlaster24. Following centrifugation, the supernatant was further purified using an RNeasy Mini Kit with On-column DNase digestion with RNase-Free DNase Set (Qiagen).

Next, the RNA was treated with RQ1 RNase-free DNase (Promega) for 30 min, purified with ultrapure phenol:chloroform:isoamyl alcohol mixed at 25:24:1 (v/v/v) (Ambion) and precipitated with 3 volumes of ice cold 100% ethanol and 0.1 volume 3M sodium acetate. To assure absence of cross-contaminating genomic DNA, 1 μ g of total RNA was treated again with another round of RNase-free DNase I (Invitrogen) for 15 min.

RNA Integrity numbers (RINs) are shown in Additional file 3: Table S2 (range: 2.1-10; median 6.6). RNA integrity numbers (RINs) were determined using an Agilent BioAnalyzer and Agilent RNA 6000 Nano Kit following the manufacturer's recommendations. We attribute low RIN numbers in some samples to long post-mortem intervals affecting tissue quality and to the rigorous DNase-treatments of RNA that were required to remove residual contaminating genomic DNA, a strategy necessary for our sensitive PCR amplification of multi-copy repeat cDNAs. To assess effects of RNA

quality on our analyses we plotted RIN values versus RT-qPCR Ct-values of *GAPDH* and could detect no significant effect of RIN when the various tissue types were considered separately. However, a mild effect ($R^2=0.38$) of RNA quality on Ct-levels is acknowledged when combining RIN and Ct values from all samples. Importantly, omission of samples with lower RINs did not affect our conclusions.

Retrotransposition assay

The EGFP L1 cell culture retrotransposition assay was conducted as previously described [87, 221, 222]. The IAP retrotransposition assay was carried out essentially as described in [144]. One μg of IAP-neo^{TNF} element reporter plasmid was cotransfected with 0.5 μg of empty vector or test plasmid in HeLa-JVM cells. At eighteen hours post-transfection, the cells were expanded from six-well plates to T₇₅ flasks, and two days later selection for retrotransposition events with 500 $\mu\text{g}/\text{ml}$ of G418 was begun. After 15 days of selection, cells were fixed, stained with Giemsa, and colonies were counted.

SH-SY5Y and NSC-34 cells were infected with Ad-L1 at $\sim 8 \times 10^{12}$ viral particles/mL [97, 217].

Bisulfite experiments

Bisulfite analysis was performed as described [21, 146] using the EZ DNA Methylation Gold Kit (Zymo Research). The analyses of Additional file 1: Figure S2 queried by PCR the methylation status of 9 CpGs within a 436-nt stretch (1169-1604) of L1 ORF1 surrounding the R159 polymorphism and used primers 1ORF1-R159-BisulfFOR (AGGAGTTGATGGAGTTGAAAATTAAG) and 2ORF1-R159-BisulfREV

(GACCTTCTCTCTAACTACCCTTAAC). PCR amplicons analyzed for Additional file 1: Figure S6 spanned 363 nts of the L1 5' UTR containing 17 CpG dinucleotides; primers were For (AAGGGGTTAGGGAGTTTTTTT) and Rev (TATCTATACCCTACCCCCAAA). PCR products were subcloned (TOPO TA Cloning Kit, Invitrogen), sequenced, and analyzed with the QUantification tool for Methylation Analysis, QUMA (quma.cdb.riken.jp; [223]). The significance of methylation differences was examined with Fisher's Exact Test statistics generated by QUMA.

Assessment of toxicity

To test potential protein toxicity (Additional file 1: Figure S4), we co-transfected in HeLa-JVM cells pcDNA6/myc-His B, a blasticidin S-resistance gene (bsr)-containing vector, together with empty vector (pcDNA3) or test expression constructs. On day 2, cells were expanded to T₇₅ flasks and selection with 5 $\mu\text{g}/\text{ml}$ blasticidin was begun. After 12 days, cells were fixed, stained with Giemsa and colonies were counted. Similarly, we co-transfected in HeLa cells pcDNA3, a

neomycin (neo)-resistant vector, together with either empty vector (pcDNA6/myc-His B) or test expression constructs, followed by selection of cells with 500 $\mu\text{g}/\text{ml}$ Geneticin (G418, Thermo Fisher).

Trypan Blue exclusion assays were performed in HEK 293T cells. Following staining, live and dead cells were counted using a Countess II Automated Cell Counter (Thermo Fisher Scientific). Use of the MultiTox-Fluor Multiplex Cytotoxicity Assay kit (Promega) followed manufacturer's instructions. This assay simultaneously measures cell viability and cytotoxicity in a single-reagent reaction, permitting ratios of live to dead cell readings to be calculated.

RT-qPCR

RT-qPCRs were conducted as previously described [146, 224]. A High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) was used to generate cDNA. RT-negative controls were run in parallel for all qPCR reactions. Duplicate samples were analyzed in a StepOne Real-Time PCR system (Applied Biosystems) using GoTaq qPCR Master Mix (Promega) and PCR primers at 200 nM each. We used two sets of primers to analyze endogenous L1 expression directed against L1Hs ORF1 (N-51-Fwd: GAATGATTTTGACGAGCTGAGAGAA; N-51-Rev: GTCCTCCCGTAGCTCAGAGTAATT) or L1Hs ORF2 sequence (N-22 Fwd: CAAACACCG CATATTCTCACTCA; N-22 Rev: CTTCCTGTG TCCATGTGATCTC). We also analyzed expression of AluS (AluS-Fwd: GCCGAGGCGGGCGGATCACC; AluS-Rev: GCCTCCCGAGTAGCTGGGAT) and AluY (AluY-Fwd: AGATCGAGACCATCCTGGCT; AluY-Rev: CCGCCTCCCGGGTTCACGCC). In all the cases, *GAPDH* was used as an internal normalization control (primers: *GAPDH*-Fwd: TGCACCACCAACTGCTTAG C; *GAPDH*-Rev: GGCATGGACTGTGGTCATGAG).

qPCR cycling parameters were as follows: 10 min at 95°C, 40 cycles of 15 sec at 95°C, followed by 60 sec at 60°C. Melting curve analysis was performed to confirm the identity of the amplified product. We employed the $\Delta\Delta\text{Ct}$ method [225] to determine relative differences in transcript levels. L1 and Alu transcript levels were plotted as "Fold change in transcript level" with respect to the transcript level in H9-hESCs (=1). Standard deviations were calculated based on 4 data points per sample derived from duplicate measurements and a technical replicate for each sample.

RT-PCR reactions used GoTaq Green Master Mix (Promega) and primer pairs 1) L1 5'UTR forward (ACGG AATCTCGCTGATTGCTA) and L1 5'UTR reverse (AAGCAAGCCTGGGCAATG) [226], which amplify a 98-bp fragment of L1 5' UTR, 2) ORF1-fwd (AGGA AATACAGAGAACGCCACAA) and ORF1-rev (GCTG GATATGAAATTCTGGGTTGA), which amplify a

259-bp fragment of L1 ORF1, and 3) GACTBPAIR2-FOR (TTCCAGCCTTCCTTCCTG) and HACTBPAIR2-REV (AATGATCTTGATCTTCATTGTGC), which amplify a 207-bp fragment of actin beta. PCR conditions were 2 min at 95°C, 30 sec at 95°C, 1 min at 58°C, 30 sec at 72°C for 35 cycles, followed by 1 min at 72°C.

Bioinformatic analysis

Occurrence of R159 polymorphisms in human L1 elements

L1Base 2 [13] was used for counting R159 polymorphism-containing human L1 elements. In brief, chromosome coordinates for human Full-Length, Intact LINE-1 elements (FLI-L1), human ORF2 Intact LINE-1 elements (ORF2-L1), and human Full-Length >4500nt LINE-1 elements (FLnI-L1) (Ens84.38) were obtained from L1Base 2. Corresponding sequences of L1 elements were retrieved using UCSC Table Browser [227]. Sequences of each subset were multiply aligned using MAFFT online [228] or as implemented in Geneious (Biomatters Ltd.; <https://www.geneious.com>; [229]). Occurrences within L1 ORF1 of a codon for R159 and its non-synonymous variants, including the most frequently observed codons for histidine (H), cysteine (C) or proline (P), were counted and their respective percentages calculated. Only a subset of the 13,671 L1 elements in the FLnI-L1 dataset were multiply aligned due to limitations of both the local and online versions of MAFFT. Also for the FLnI-L1 dataset, which included a greater number of evolutionarily older L1 sequences, a minority of aligned L1 sequences displayed structural rearrangements and/or higher sequence divergence in the R159 codon region resulting in unreliable prediction of sequence at the R159 codon position: these were excluded. In all, 6346 FLnI-L elements were included in the analysis.

RNA-Seq datasets

Publicly available RNA-Seq datasets were analyzed by TETranscripts software package [150]. TDP-43-related datasets SRP057819 and GSE77702 have been previously described [147, 148]. Dataset SRP064478, submitted by the Bennett Lab at Virginia Commonwealth University, consists of RNA-Seq data for total stranded RNA with >50 million 2x150 bp sequencing reads from 15 post-mortem cervical spinal cord sections (7 ALS and 8 healthy controls). GSE76220 includes 20-30 million mappable 1x50 bp reads from total stranded RNA isolated from laser capture microdissected motor neurons from post-mortem lumbar spinal cords [172]. GSE67196 consists of on average 83 million 1X100 bp reads per sample (91.5 million for cerebellum and 73.6 million for frontal cortex), as described by [52, 171].

The Library of Integrated Network-Based Cellular Signatures (LINCS)-NeuroLINCS dGAP dataset (accession number phs001231.v1.p1, SRP098831) includes RNA-Seq of iPSC-derived motor neurons from 4 C9ALS and 3 SMA

patients (3 sequencing replicates each), and 3 unaffected healthy controls (2 or 3 replicates each). It has been reported that L1 activity in iPSCs can vary with cell passage, increasing during reprogramming but subsequently subsiding [224, 226, 230]. However, passage numbers of the NeuroLINCS cell lines fall within similar ranges, from 25 to 27 for ALS and healthy control and 21 to 30 for SMA samples.

Generating RPKM plots

To generate RPKM plots, raw data was aligned to the consensus sequences for a selected group of younger retrotransposons present in RepBase [152], including L1-Ta, L1PA2, L1-pre-Ta, AluYa5, AluYa8, AluYb8, AluYb9, AluY, AluSq, SVA_D, SVA_E, SVA_F, LTR5_Hs, and LTR7Y subfamilies. Alignments were made using Bowtie 2 [231] with end-to-end sensitive parameters: -D 15 -R 2 -N 0 -L 22 -i S, 1, 1.15. We adapted the RPKM formula to provide a normalized measure of the number of reads that align with each consensus sequence based on their size. Data were plotted using OriginPro (OriginLab Corp.) with standard deviation error bars for replicates.

MDS plots were generated in R script using the edgeR package [232].

Use of TETranscripts

TETranscripts is a software package that estimates both gene and TE transcript abundances in RNA-Seq data and conducts differential expression analysis on the resultant count tables [150]. Sequences were aligned to human genome assembly GRCh38 using STAR [233]. Alignment parameters were `outFilterMultimapNmax100` and `winAnchorMultimapNmax 200`, which allow up to 100 alignments per read. TE annotation files were downloaded from <http://labshare.cshl.edu/shares/mhammelllab/www-data/TEToolkit/> (including 1181 TE types). Following the generation of a count table for gene and TE transcripts, the differential expression analysis closely followed the DESeq2 package [234] for modeling the counts data with a negative binomial distribution and computing adjusted P-values. In addition to the standard transcript abundance normalization approach used by the DESeq2 package, TETranscripts offers two additional options, reads per mapped million (RPM) and quantile normalization. All other procedures exactly followed the DESeq2 method. TETranscripts runs the DESeq2 method with a default set of general parameters. When there were no (or very few) replicates, we used the blind method for variance estimation and fit-only for SharingMode. Otherwise, we used pooled or per-condition methods and maximum SharingMode, as suggested by the DESeq2 package.

Locus-specific mapping of TEs

The pipeline to map TEs to individual genomic loci used the alignment algorithm HISAT2 [235] to map sequence reads to the human genome. Reads that mapped to more

than one genomic position were discarded. Counts per TE integrant (genomic loci) were generated using the multi-BamCov tool from the BEDtools software [236]. Normalisation for sequencing depth was performed using voom [237], with total number of reads on genes as size factors. RepeatMasker 4.0.5 (Library 20140131), a newer version than RepeatMasker 4.0 used by Prudencio et al. [52], was used to generate a list of TE subfamilies. In the case of HERVs, we re-assembled fragmented internal and LTR sequences to generate full-length HERV integrants: this step avoids bias in counts due to the highly fragmented nature of the annotated HERVs. We removed from our analyses very small and abundant repeats (low complexity and simple repeats). Any TEs with a low number of reads across all samples or which overlapped exons were also omitted from our analyses. Differential expression was performed as implemented in the voom library of Bioconductor [238]. A TE locus was considered to be differentially expressed if its fold change was greater than 2 and FDR smaller than 0.05. The Benjamini-Hochberg procedure was used to compute the FDR. Hierarchical clustering of the heatmap was performed with Pearson correlation as distance and complete agglomeration method for both, rows and columns. Any raw data files will be provided upon request to the authors.

Additional files

Additional file 1: Figure S1. Patterns of L1 ORF1p expression in various cell lines using multiple antibodies for detection. **Figure S2.** Frequency of polymorphisms and DNA methylation at the L1 ORF1p R159 residue. **Figure S3.** ORF1p does not contain a CRM1-dependent nuclear export signal. **Figure S4.** Cell culture toxicity assays. **Figure S5.** Overexpression or knockdown of TDP-43 protein does not alter L1 expression. **Figure S6.** Methylation analyses of the CpG island of the 5' UTR promoter of endogenous L1 elements show effects of altered levels of TDP-43. **Figure S7.** Expression levels of L1 (and Alu TEs determined by RT-qPCR. **Figure S8.** Representative gels showing expression of ORF1p in normal and ALS-associated tissues determined by Western blotting. **Figure S9.** TE locus-specific analyses of the GSE67196 RNA-Seq dataset [171]. (PDF 25088 kb)

Additional file 2: Table S1. Summary of significant DE TE subfamilies determined by TETRanscripts RNA-Seq datasets. (XLSX 326 kb)

Additional file 3: Table S2. Tissue samples used in this study. (PDF 65 kb)

Additional file 4: Table S3. Summary of significant individual DE TE loci in the GSE67196 RNA-Seq dataset. (XLSX 774 kb)

Abbreviations

aa: Amino acids; ALS: Amyotrophic lateral sclerosis; bp: Base pair; CTRL: Control; DE: Differentially expressed; EGFP: Enhanced green fluorescent protein; fALS: Familial ALS; FISH: Fluorescent *in situ* hybridization; FDR: False discovery rate; FTD: Frontotemporal dementia; FTLD: Frontotemporal lobar degeneration; HEF: Human embryonic fibroblast; HERV: Human endogenous retrovirus; hESC: Human embryonic stem cell; IF: Immunofluorescence; iPSC: Induced pluripotent stem cell; kb: Kilobase; kD: Kilodalton; LINE-1/L1: Long interspersed element 1; LTR: Long terminal repeat; MDS: Multidimensional scaling; NES: Nuclear export signal; NLS: Nuclear localization signal; nt: Nucleotide; ORF: Open reading frame; padj: Adjusted P-value; PCA: Principal component analysis; PCR: Polymerase chain reaction; RNP: Ribonucleoprotein; RT-qPCR: Reverse transcription quantitative PCR; RNA-Seq: RNA sequencing; RT: Reverse transcriptase; sALS: Sporadic ALS; SG: Stress granule; SMA: Spinal muscular atrophy; SVA: SINE/VNTR/Alu; TE: Transposable element

Acknowledgements

The authors appreciate reagents provided by researchers listed in the Methods section. Special thanks to the Target ALS Multicenter Postmortem Tissue Core, ChemCore, Synthesis & Sequencing Core, Ross Flow Cytometry Core, Flow Cytometry Core Analytic Laboratories, and Dr. O. Pletnikova of the Neuropathology Core of Johns Hopkins University School of Medicine, and John Cottrell of the University of Maryland Brain and Tissue Bank of the NIH NeuroBioBank. Han Wan and Alisha Soares provided technical assistance. We appreciate advice/reagents from Drs. K. Burns, L. Florea, J. Ling, N. Maragakis, J.-P. Richard, J. Rothstein, J. Wang, and J. Wong of Johns Hopkins School of Medicine, A. Ewing of the Mater Research Institute-University of Queensland, and M. Hammell of Cold Spring Harbor Laboratory.

Funding

JLG was supported by grants from the NIH National Institute of Neurological Disorders and Stroke (1R03NS087290-01), the NIH Eunice Kennedy Shriver National Institute of Child Health and Human Development (R21HD083915-01A1), and the ALS Therapy Alliance (2013-F-067). The JLG-P lab is supported by CICE-FEDER-P12-CTS-2256, Plan Nacional de I+D+I 2008-2011 and 2013-2016 (FIS-FEDER-PI14/02152), PCIN-2014-115-ERA-NET NEURON II, the European Research Council (ERC-Consolidator ERC-STG-2012-233764), by an International Early Career Scientist grant from the Howard Hughes Medical Institute (IECS-55007420), by The Wellcome Trust-University of Edinburgh Institutional Strategic Support Fund (ISFF2), and by a private donation by Ms. Francisca Serrano (Trading y Bolsa para Torpes, Granada, Spain). SRH is funded by the Government of Spain (MINECO, RYC-2016-21395 and SAF2015-71589-P). HHK is funded by the NIH National Institute of General Medical Sciences (1R01GM099875-05). SJW was supported by the NIH National Institute of General Medical Sciences (P50GM107632-01). As users of NIH NeuroLINC datasets, we acknowledge the NIH LINC grant (NS091046), sequencing support from the National Institutes of Health (P30-ES002109) through the MIT BioMicro Center, Massachusetts, Huntington's Disease Center without Walls (NS01234) and the PREDICT-HD study of the Huntington Study Group (NS01234).

Availability of data and materials

All sample information and RNA-Seq analysis summary results are available as part of the Additional files.

Authors' contributions

LS, SRH, and JLG isolated RNA and performed RT-qPCR analyses. ARR, PMS, and SJW performed TE subfamily RNA-Seq analyses, and EP, PT, and DT performed locus-specific TE RNA-Seq analyses. JAC conducted bisulfite methylation analyses. GCP and JLG performed all other experiments. GCP, JLG, JM, and JLG analyzed data. LWO, JR, and RB provided tissue samples and project advice. HHK provided reagents, advice, and support. JLG wrote the manuscript and designed experiments. GCP, JLG, and JM contributed to manuscript preparation or review.

Ethics approval and consent to participate

All post-mortem tissues were obtained following approval of the Institutional Review Boards of the UCSD School of Medicine (to JR) and the JHU School of Medicine (IRB00066246 to JLG).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

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

Author details

¹McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. ²GENYO. Centre for Genomics and Oncological Research: Pfizer, University of Granada, Andalusian Regional Government, Granada, Spain. ³Oncology Center-Cancer Biology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. ⁴School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015

Lausanne, Switzerland. ⁵Department of Neurosciences, School of Medicine, University of California at San Diego, San Diego, California, USA.

⁶Neuromuscular Division, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. ⁷Department of Biochemistry and Molecular Biology II, Faculty of Pharmacy, University of Granada, Granada, Spain.

⁸Department of Human Genetics, Medical Faculty, University of Saarland, Homburg/Saar, Germany. ⁹MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine (IGMM), University of Edinburgh, Western General Hospital, Edinburgh, UK.

Received: 20 June 2018 Accepted: 15 November 2018

Published online: 15 December 2018

REFERENCES

- McClintock B. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci U S A*. 1950;36(6):344–55.
- Kazazian HH Jr, Moran JV. Mobile DNA in Health and Disease. *N Engl J Med*. 2017;377(4):361–70.
- Wildschutte JH, Williams ZH, Montesin M, Subramanian RP, Kidd JM, Coffin JM. Discovery of unfixed endogenous retrovirus insertions in diverse human populations. *Proc Natl Acad Sci U S A*. 2016;113(16):E2326–34.
- Mager DL, Stoye JP. Mammalian Endogenous Retroviruses, Mobile DNA III: American Society for Microbiology; 2015. p. 1079–100.
- Richardson SR, Moran JV, Kopera HC, Doucet AJ, Moldovan JB, Garcia-Perez JL. The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes, Mobile DNA III: American Society for Microbiology; 2015. p. 1165–208.
- Speak M. Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol Cell Biol*. 2001;21(6):1973–85.
- Martin SL, Branciforte D, Keller D, Bain DL. Trimeric structure for an essential protein in L1 retrotransposition. *Proc Natl Acad Sci U S A*. 2003;100(24):13815–20.
- Martin SL, Cruceanu M, Branciforte D, Wai-Lun Li P, Kwok SC, Hodges RS, Williams MC. LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 protein. *J Mol Biol*. 2005;348(3):549–61.
- Khazina E, Truffaut V, Buttner R, Schmidt S, Coles M, Weichenrieder O. Trimeric structure and flexibility of the L1ORF1 protein in human L1 retrotransposition. *Nat Struct Mol Biol*. 2011;18(9):1006–14.
- Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH. Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A*. 2003;100(9):5280–5.
- Mills RE, Bennett EA, Iskow RC, Devine SE. Which transposable elements are active in the human genome? *Trends Genet*. 2007;23(4):183–91.
- Beck CR, Collier P, Macfarlane C, Malig M, Kidd JM, Eichler EE, Badge RM, Moran JV. LINE-1 retrotransposition activity in human genomes. *Cell*. 2010;141(7):1159–70.
- Penzkofer T, Jager M, Figlerowicz M, Badge R, Mundlos S, Robinson PN, Zemojtel T. L1Base 2: more retrotransposition-active LINE-1s, more mammalian genomes. *Nucleic Acids Res*. 2017;45(D1):D68–d73.
- Sokolowski M, Chynces M, deHaro D, Christian CM, Belancio VP. Truncated ORF1 proteins can suppress LINE-1 retrotransposition in trans. *Nucleic Acids Res*. 2017;45(9):5294–308.
- Scott EC, Devine SE. The Role of Somatic L1 Retrotransposition in Human Cancers. *Viruses*. 2017;9(6).
- Faulkner GJ, Garcia-Perez JL. L1 Mosaicism in Mammals: Extent, Effects, and Evolution. *Trends Genet*. 2017;33(11):802–16.
- Suarez NA, Macia A, Muotri AR. LINE-1 retrotransposons in healthy and diseased human brain. *Dev Neurobiol*. 2017;78(5):434–55.
- Perrat PN, DasGupta S, Wang J, Theurkauf W, Weng Z, Rosbash M, Waddell S. Transposition-driven genomic heterogeneity in the *Drosophila* brain. *Science*. 2013;340(6128):91–5.
- Yang F, Wang PJ. Multiple LINEs of retrotransposon silencing mechanisms in the mammalian germline. *Semin Cell Dev Biol*. 2016;59:118–25.
- Muotri AR, Chu VT, Marchetto MCN, Deng W, Moran JV, Gage FH. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature*. 2005;435(7044):903–10.
- Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, Morell M, O'Shea KS, Moran JV, Gage FH. L1 retrotransposition in human neural progenitor cells. *Nature*. 2009;460(7259):1127–31.
- Blaudin de Th FX, Reika H, Peze-Heidsieck E, Massiani-Beaudoin O, Joshi RL, Fuchs J, Prochiantz A. Engrailed homeoprotein blocks degeneration in adult dopaminergic neurons through LINE-1 repression. *EMBO J*. 2018; 37(15):e97374.
- Macia A, Widmann TJ, Heras SR, Ayllon V, Sanchez L, Benkaddour-Boumzaouad M, Munoz-Lopez M, Rubio A, Amador-Cubero S, Blanco-Jimenez E, Garcia-Castro J, Menendez P, Ng P, Muotri AR, Goodier JL, Garcia-Perez JL. Engineered LINE-1 retrotransposition in nondividing human neurons. *Genome Res*. 2017;27(3):335–48.
- Thomas CA, Paquola AC, Muotri AR. LINE-1 retrotransposition in the nervous system. *Annu Rev Cell Dev Biol*. 2012;28:555–73.
- Reilly MT, Faulkner GJ, Dubnau J, Ponomarev I, Gage FH. The role of transposable elements in health and diseases of the central nervous system. *J Neurosci*. 2013;33(45):17577–86.
- Erwin JA, Marchetto MC, Gage FH. Mobile DNA elements in the generation of diversity and complexity in the brain. *Nature Rev Neurosci*. 2014;15(8):497–506.
- Richardson SR, Morell S, Faulkner GJ. L1 retrotransposons and somatic mosaicism in the brain. *Annu Rev Genet*. 2014;48:1–27.
- Singer T, McConnell MJ, Marchetto MC, Coufal NG, Gage FH. LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? *Trends Neurosci*. 2010;33(8):345–54.
- Coufal NG, Garcia-Perez JL, Peng GE, Marchetto MCN, Muotri AR, Mu Y, Carson CT, Macia A, Moran JV, Gage FH. Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. *Proc Natl Acad Sci U S A*. 2011;108(51):20382–7.
- Muotri AR, Marchetto MCN, Coufal NG, Oefner R, Yeo G, Nakashima K, Gage FH. L1 retrotransposition in neurons is modulated by MeCP2. *Nature*. 2010; 468(7322):443–6.
- Jacob-Hirsch J, Eyal E, Knisbacher BA, Roth J, Cesarka K, Dor C, Farage-Barhom S, Kunik V, Simon AJ, Gal M, Yalon M, Moshitch-Moshkovitz S, Tearle R, Constantini S, Levanon EY, Amariglio N, Rechavi G. Whole-genome sequencing reveals principles of brain retrotransposition in neurodevelopmental disorders. *Cell Res*. 2018;28(2):187–203.
- Bundo M, Toyoshima M, Okada Y, Akamatsu W, Ueda J, Nemoto-Miyauchi T, Sunaga F, Toritsuka M, Ikawa D, Kakita A, Kato M, Kasai K, Kishimoto T, Nawa H, Okano H, Yoshikawa T, Kato T, Iwamoto K. Increased L1 retrotransposition in the neuronal genome in schizophrenia. *Neuron*. 2014;81(2):306–13.
- Doyle GA, Crist RC, Karatas ET, Hammond MJ, Ewing AD, Ferraro TN, Hahn CG, Berrettini WH. Analysis of LINE-1 Elements in DNA from Postmortem Brains of Individuals with Schizophrenia. *Neuropsychopharmacology*. 2017; 42(13):2602–11.
- Liu S, Du T, Liu Z, Shen Y, Xiu J, Xu Q. Inverse changes in L1 retrotransposons between blood and brain in major depressive disorder. *Sci Rep*. 2016;6:37530.
- Okudaira N, Ishizaka Y, Nishio H. Retrotransposition of long interspersed element 1 induced by methamphetamine or cocaine. *J Biol Chem*. 2014; 289(37):25476–85.
- Okudaira N, Ishizaka Y, Nishio H, Sakagami H. Morphine and Fentanyl Citrate Induce Retrotransposition of Long Interspersed Element-1. *In Vivo*. 2016; 30(2):113–8.
- Doyle GA, Doucet-O'Hare TT, Hammond MJ, Crist RC, Ewing AD, Ferraro TN, Mash DC, Kazazian HH Jr, Berrettini WH. Reading LINES within the cocaine addicted brain. *Brain Behav*. 2017;7(5):e00678.
- Tan H, Wu C, Jin L. A Possible Role for Long Interspersed Nuclear Elements-1 (LINE-1) in Huntington's Disease Progression. *Med Sci Monit*. 2018;24:3644–52.
- Bachiller S, Del-Pozo-Martin Y, Carrion AM. L1 retrotransposition alters the hippocampal genomic landscape enabling memory formation. *Brain Behav Immun*. 2017;64:65–70.
- Bedrosian TA, Quayle C, Novaresi N, Gage FH. Early life experience drives structural variation of neural genomes in mice. *Science*. 2018;359(6382):1395–9.
- Goodier JL. Retrotransposition in tumors and brains. *Mobile DNA*. 2014;5(1):11.
- Goodier JL. Restricting retrotransposons: a review. *Mobile DNA*. 2016;7:16.
- Brown RH, Al-Chalabi A. Amyotrophic Lateral Sclerosis. *N Engl J Med*. 2017; 377(2):162–72.
- Burrell JR, Halliday GM, Kril JJ, Ittner LM, Gotz J, Kiernan MC, Hodges JR. The frontotemporal dementia-motor neuron disease continuum. *Lancet*. 2016; 388(10047):919–31.
- Steele AJ, Al-Chalabi A, Ferrante K, Cudkowicz ME, Brown RH Jr, Garson JA. Detection of serum reverse transcriptase activity in patients with ALS and unaffected blood relatives. *Neurology*. 2005;64(3):454–8.
- MacGowan DJ, Scelsa SN, Imperato TE, Liu KN, Baron P, Polsky B. A controlled study of reverse transcriptase in serum and CSF of HIV-negative patients with ALS. *Neurology*. 2007;68(22):1944–6.

47. McCormick AL, Brown RH Jr, Cudkowicz ME, Al-Chalabi A, Garson JA. Quantification of reverse transcriptase in ALS and elimination of a novel retroviral candidate. *Neurology*. 2008;70(4):278–83.
48. Alfahad T, Nath A. Retroviruses and amyotrophic lateral sclerosis. *Antiviral Res*. 2013;99(2):180–7.
49. Douville R, Liu J, Rothstein J, Nath A. Identification of active loci of a human endogenous retrovirus in neurons of patients with amyotrophic lateral sclerosis. *Ann Neurol*. 2011;69(1):141–51.
50. Hadlock K, Miller R, Jin X, Yu S, Reis J, Mass J, Gelinis D, McGrath M. Elevated rates of antibody reactivity to HML-2/Herv-K but not other endogenous retroviruses in ALS. *Amyotroph Lateral Scler Other Motor Neuron Disord*. 2004;5(2):63.
51. Li W, Lee MH, Henderson L, Tyagi R, Bachani M, Steiner J, Campanac E, Hoffman DA, von Geldern G, Johnson K, Maric D, Morris HD, Lentz M, Pak K, Mammen A, Ostrow L, Rothstein J, Nath A. Human endogenous retrovirus-K contributes to motor neuron disease. *Sci Transl Med*. 2015;7(307):307ra153–307ra153.
52. Prudencio M, Gonzales PK, Cook CN, Gendron TF, Daugherty LM, Song Y, Ebbert MTW, van Blitterswijk M, Zhang YJ, Jansen-West K, Baker MC, DeTure M, Rademakers R, Boylan KB, Dickson DW, Petrucelli L, Link CD. Repetitive element transcripts are elevated in the brain of C9orf72 ALS/FTLD patients. *Hum Mol Genet*. 2017;26(17):3421–31.
53. Li HF, Wu ZY. Genotype-phenotype correlations of amyotrophic lateral sclerosis. *Transl Neurodegener*. 2016;5:3.
54. Al-Chalabi A, Jones A, Troakes C, King A, Al-Sarraj S, van den Berg LH. The genetics and neuropathology of amyotrophic lateral sclerosis. *Acta Neuropathol*. 2012;124(3):339–52.
55. Baralle M, Buratti E, Baralle FE. The role of TDP-43 in the pathogenesis of ALS and FTLD. *Biochem Soc Trans*. 2013;41(6):1536–40.
56. Nakashima-Yasuda H, Uryu K, Reisins J, Xie SX, Hurtig H, Duda JE, Arnold SE, Siderowf A, Grossman M, Leverenz JB, Woltjer R, Lopez OL, Hamilton R, Tsuang DW, Galasko D, Masliah E, Kaye J, Clark CM, Montine TJ, Lee VM, Trojanowski JQ. Co-morbidity of TDP-43 proteinopathy in Lewy body related diseases. *Acta Neuropathol*. 2007;114(3):221–9.
57. Amador-Ortiz C, Lin WL, Ahmed Z, Personett D, Davies P, Duara R, Graff-Radford NR, Hutton ML, Dickson DW. TDP-43 immunoreactivity in hippocampal sclerosis and Alzheimer's disease. *Ann Neurol*. 2007;61(5):435–45.
58. Schwab C, Arai T, Hasegawa M, Yu S, McGeer PL. Colocalization of transactivation-responsive DNA-binding protein 43 and huntingtin in inclusions of Huntington disease. *J Neuropathol Exp Neurol*. 2008;67(12):1159–65.
59. Weihl CC, Temiz P, Miller SE, Watts G, Smith C, Forman M, Hanson PI, Kimonis V, Pestronk A. TDP-43 accumulation in inclusion body myopathy muscle suggests a common pathogenic mechanism with frontotemporal dementia. *J Neurol Neurosurg Psychiatry*. 2008;79(10):1186–9.
60. Goodier JL, Zhang L, Vetter MR, Kazazian HH. LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex. *Mol Cell Biol*. 2007;27(18):6469–83.
61. Doucet AJ, Hulme AE, Sahinovic E, Kulpa DA, Moldovan JB, Kopera HC, Athanikar JN, Hasnaoui M, Bucheton A, Moran JV, Gilbert N. Characterization of LINE-1 ribonucleoprotein particles. *PLoS Genet*. 2010;6(10):e1001150.
62. Cook PR, Jones CE, Furano AV. Phosphorylation of ORF1p is required for L1 retrotransposition. *Proc Natl Acad Sci U S A*. 2015;112(14):4298–303.
63. MacLennan M, Garcia-Canadas M, Reichmann J, Khazina E, Wagner G, Playfoot CJ, Salvador-Palomeque C, Mann AR, Peressini P, Sanchez L, Dobie K, Read D, Hung CC, Eskeland R, Meehan RR, Weichenrieder O, Garcia-Perez JL, Adams IR. Mobilization of LINE-1 retrotransposons is restricted by Tex19.1 in mouse embryonic stem cells. *Elife*. 2017;6:e26152.
64. Martin SL, Branciforte D. Synchronous expression of LINE-1 RNA and protein in mouse embryonal carcinoma cells. *Mol Cell Biol*. 1993;13(9):5383–92.
65. Horn AV, Klawitter S, Held U, Berger A, Vasudevan AAJ, Bock A, Hofmann H, Hanschmann K-MO, Trösemeyer J-H, Flory E. Human LINE-1 restriction by APOBEC3C is deaminase independent and mediated by an ORF1p interaction that affects LINE reverse transcriptase activity. *Nucleic Acids Res*. 2014;42(1):396–416.
66. Hu S, Li J, Xu F, Mei S, Le Duff Y, Yin L, Pang X, Cen S, Jin Q, Liang C, Guo F. SAMHD1 Inhibits LINE-1 Retrotransposition by Promoting Stress Granule Formation. *PLoS Genet*. 2015;11(7):e1005367.
67. Rodić N, Sharma R, Sharma R, Zampella J, Dai L, Taylor MS, Hruban RH, Iacobuzio-Donahue CA, Maitra A, Torbenson MS, Goggins M, Shih I-M, Duffield AS, Montgomery EA, Gabrielson E, Netto GJ, Lotan TL, De Marzo AM, Westra W, Binder ZA, Orr BA, Gallia GL, Eberhart CG, Boeke JD, Harris CR, Burns KH. Long interspersed element-1 protein expression is a hallmark of many human cancers. *The American Journal of Pathology*. 2014;184(5):1280–6.
68. Leibold DM, Swergold GD, Singer MF, Thayer RE, Dombroski BA, Fanning TG. Translation of LINE-1 DNA elements in vitro and in human cells. *Proc Natl Acad Sci U S A*. 1990;87(18):6990–4.
69. Josephson R, Ording CJ, Liu Y, Shin S, Lakshminpathy U, Toumadje A, Love B, Chesnut JD, Andrews PW, Rao MS, Auerbach JM. Qualification of embryonal carcinoma 2102Ep as a reference for human embryonic stem cell research. *Stem Cells*. 2007;25(2):437–46.
70. Sivan G, Kedersha N, Elroy-Stein O. Ribosomal slowdown mediates translational arrest during cellular division. *Mol Cell Biol*. 2007;27(19):6639–46.
71. Guo H, Chitiprolu M, Gagnon D, Meng L, Perez-Iratxeta C, Lagace D, Gibbins D. Autophagy supports genomic stability by degrading retrotransposon RNA. *Nature Commun*. 2014;5:5276.
72. Goodier JL, Mandal PK, Zhang L, Kazazian HH. Discrete subcellular partitioning of human retrotransposon RNAs despite a common mechanism of genome insertion. *Hum Mol Genet*. 2010;19(9):1712–25.
73. Taylor MS, LaCava J, Mita P, Molloy K, Huang CR, Li D, Adney Emily M, Jiang H, Burns Kathleen H, Chait Brian T, Rout Michael P, Boeke Jef D, Dai L. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. *Cell*. 2013;155(5):1034–48.
74. Alisch RS, Garcia-Perez JL, Muotri AR, Gage FH, Moran JV. Unconventional translation of mammalian LINE-1 retrotransposons. *Genes Dev*. 2006;20(2):210–24.
75. Ergun S, Buschmann C, Heukeshoven J, Dammann K, Schnieders F, Lauke H, Chalajour F, Kilic N, Stratling WH, Schumann GG. Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. *J Biol Chem*. 2004;279(26):27753–63.
76. Goodier JL, Ostertag EM, Engleka K, Seleme M, Kazazian HH. A potential role for the nucleolus in L1 retrotransposition. *Hum Mol Genet*. 2004;13(10):1041–8.
77. Chen L, Dahlstrom JE, Chandra A, Board P, Rangasamy D. Prognostic value of LINE-1 retrotransposon expression and its subcellular localization in breast cancer. *Breast Cancer Res Treat*. 2012;136(1):129–42.
78. Sokolowski M, DeFreece CB, Servant G, Kines KJ, de Haro DL, Belancio VP. Development of a monoclonal antibody specific to the endonuclease domain of the human LINE-1 ORF2 protein. *Mobile DNA*. 2014;5(1):29.
79. De Luca C, Guadagni F, Sinibaldi-Vallebona P, Sentinelli S, Gallucci M, Hoffmann A, Schumann GG, Spadafora C, Sciamanna I. Enhanced expression of LINE-1-encoded ORF2 protein in early stages of colon and prostate transformation. *Oncotarget*. 2016;7(4):4048–61.
80. Harrison AF, Shorter J. RNA-binding proteins with prion-like domains in health and disease. *Biochem J*. 2017;474(8):1417–38.
81. Zambrano R, Conchillo-Sole O, Iglesias V, Illa R, Rousseau F, Schymkowitz J, Sabate R, Daura X, Ventura S. PrionW: a server to identify proteins containing glutamine/asparagine rich prion-like domains and their amyloid cores. *Nucleic Acids Res*. 2015;43(W1):W331–7.
82. Alberti S, Halfmann R, King O, Kapila A, Lindquist S. A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell*. 2009;137(1):146–58.
83. Couthouis J, Hart MP, Shorter J, De Jesus-Hernandez M, Erion R, Oristano R, Liu AX, Ramos D, Jethava N, Hosangadi D, Epstein J, Chiang A, Diaz Z, Nakaya T, Ibrahim F, Kim HJ, Solski JA, Williams KL, Mojsilovic-Petrovic J, Ingre C, Boylan K, Graff-Radford NR, Dickson DW, Clay-Falcone D, Elman L, McCluskey L, Greene R, Kalb RG, Lee VM, Trojanowski JQ, Ludolph A, Robberecht W, Andersen PM, Nicholson GA, Blair IP, King OD, Bonini NM, Van Deerlin V, Rademakers R, Mourelatos Z, Gitler AD. A yeast functional screen predicts new candidate ALS disease genes. *Proc Natl Acad Sci U S A*. 2011;108(52):20881–90.
84. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH. High frequency retrotransposition in cultured mammalian cells. *Cell*. 1996;87(5):917–27.
85. Khazina E, Weichenrieder O. Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. *Proc Natl Acad Sci U S A*. 2009;106(3):731–6.
86. Kimberland ML, Divoky V, Prchal J, Schwahn U, Berger W, Kazazian HH Jr. Full-length human L1 insertions retain the capacity for high frequency retrotransposition in cultured cells. *Hum Mol Genet*. 1999;8(8):1557–60.
87. Ostertag EM, Prak E, DeBerardinis R, Moran JV, Kazazian HH. Determination of L1 retrotransposition kinetics in cultured cells. *Nucleic Acids Res*. 2000;28(6):1418–23.

88. Kulpa DA, Moran JV. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat Struct Mol Biol.* 2006;13(7):655–60.
89. Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM. Localization of ASH1 mRNA particles in living yeast. *Mol Cell.* 1998;2(4):437–45.
90. Liu Q, Dreyfuss G. A novel nuclear structure containing the survival of motor neurons protein. *EMBO J.* 1996;15(14):3555–65.
91. Wojciechowska M, Krzyzosiak WJ. Cellular toxicity of expanded RNA repeats: focus on RNA foci. *Hum Mol Genet.* 2011;20(19):3811–21.
92. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung GY, Karydas A, Seeley WW, Josephs KA, Coppola G, Geschwind DH, Wszolek ZK, Feldman H, Knopman DS, Petersen RC, Miller BL, Dickson DW, Boylan KB, Graff-Radford NR, Rademakers R. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron.* 2011;72(2):245–56.
93. Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, Kalimo H, Paetau A, Abramzon Y, Remes AM, Kaganovich A, Scholz SW, Duckworth J, Ding J, Harmer DW, Hernandez DG, Johnson JO, Mok K, Ryten M, Trabzuni D, Guerreiro RJ, Orrell RW, Neal J, Murray A, Pearson J, Jansen IE, Sondervan D, Seelaar H, Blake D, Young K, Halliwell N, Callister JB, Toulson G, Richardson A, Gerhard A, Snowden J, Mann D, Neary D, Nalls MA, Peuralinna T, Jansson L, Isovliita VM, Kaivorinne AL, Holtta-Vuori M, Ikonen E, Sulkava R, Benatar M, Wu J, Chio A, Restagno G, Borghero G, Sabatelli M, Heckerman D, Rogaeve A, Zinman L, Rothstein JD, Sendtner M, Drepper C, Eichler EE, Alkan C, Abdullaev Z, Pack SD, Dutra A, Pak E, Hardy J, Singleton A, Williams NM, Heutink P, Pickering-Brown S, Morris HR, Tienari PJ, Traynor BJ. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron.* 2011;72(2):257–68.
94. Hensman Moss DJ, Poulter M, Beck J, Hehir J, Polke JM, Campbell T, Adamson G, Mudanohwo E, McColgan P, Haworth A, Wild EJ, Sweeney MG, Houlden H, Mead S, Tabrizi SJ. C9orf72 expansions are the most common genetic cause of Huntington disease phenocopies. *Neurology.* 2014;82(4):292–9.
95. Donnelly CJ, Zhang PW, Pham JT, Haeusler AR, Mistry NA, Vidsensky S, Daley EL, Poth EM, Hoover B, Fines DM, Maragakis N, Tienari PJ, Petrucelli L, Traynor BJ, Wang J, Rigo F, Bennett CF, Blackshaw S, Sattler R, Rothstein JD. RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron.* 2013;80(2):415–28.
96. Rossi S, Serrano A, Gerbino V, Giorgi A, Di Francesco L, Nencini M, Bozzo F, Schinina ME, Bagni C, Cestra G, Carri MT, Achsel T, Cozzolino M. Nuclear accumulation of mRNAs underlies G4C2-repeat-induced translational repression in a cellular model of C9orf72 ALS. *J Cell Sci.* 2015;128(9):1787–99.
97. Kubo S, Seleme MC, Soifer HS, Perez JL, Moran JV, Kazazian HH Jr, Kasahara N. L1 retrotransposition in nondividing and primary human somatic cells. *Proc Natl Acad Sci U S A.* 2006;103(21):8036–41.
98. Shi X, Seluanov A, Gorbunova V. Cell divisions are required for L1 retrotransposition. *Mol Cell Biol.* 2007;27(4):1264–70.
99. Xie Y, Mates L, Ivics Z, Izsvak Z, Martin SL, An W. Cell division promotes efficient retrotransposition in a stable L1 reporter cell line. *Mobile DNA.* 2013;4(1):10.
100. Mita P, Wudzinska A, Sun X, Andrade J, Nayak S, Kahler DJ, Badri S, LaCava J, Ueberheide B, Yun CY, Fenyo D, Boeke JD. LINE-1 protein localization and functional dynamics during the cell cycle. *Elife.* 2018;7.
101. Mueller S, Schittenhelm M, Honecker F, Malenke E, Lauber K, Wesselborg S, Hartmann JT, Bokemeyer C, Mayer F. Cell-cycle progression and response of germ cell tumors to cisplatin in vitro. *Int J Oncol.* 2006;29(2):471–9.
102. Ikegami S, Taguchi T, Ohashi M, Oguro M, Nagano H, Mano Y. Aphidicolin prevents mitotic cell division by interfering with the activity of DNA polymerase- α . *Nature.* 1978;275(5679):458–60.
103. Pratt WM, Ruddon RW, Ensminger WD, Maybaum J. The anticancer drugs. New York: Oxford University Press; 1994.
104. Sokolowski M, de Haro D, Christian CM, Kines KJ, Belancio VP. Characterization of L1 ORF1p self-interaction and cellular localization using a mammalian two-hybrid system. *PLoS One.* 2013;8(12):e82021.
105. Nauffer MN, Furano AV, Williams MC. Protein-nucleic acid interactions of LINE-1 ORF1p. *Semin Cell Dev Biol.* 2018;S1084-9521(17):30451–2.
106. Nishitani H, Taraviras S, Lygerou Z, Nishimoto T. The human licensing factor for DNA replication Cdt1 accumulates in G1 and is destabilized after initiation of S-phase. *J Biol Chem.* 2001;276(48):44905–11.
107. McGarry TJ, Kirschner MW. Geminin, an inhibitor of DNA replication, is degraded during mitosis. *Cell.* 1998;93(6):1043–53.
108. Ayala YM, Zago P, D'Ambrogio A, Xu YF, Petrucelli L, Buratti E, Baralle FE. Structural determinants of the cellular localization and shuttling of TDP-43. *J Cell Sci.* 2008;121(Pt 22):3778–85.
109. la Cour T, Gupta R, Rapacki K, Skriver K, Poulsen FM, Brunak S. NESbase version 1.0: a database of nuclear export signals. *Nucleic Acids Res.* 2003;31(1):393–6.
110. Yang J, Bardes ES, Moore JD, Brennan J, Powers MA, Kornbluth S. Control of cyclin B1 localization through regulated binding of the nuclear export factor CRM1. *Genes Dev.* 1998;12(14):2131–43.
111. Engel K, Kotlyarov A, Gaestel M. Leptomycin B-sensitive nuclear export of MAPKAP kinase 2 is regulated by phosphorylation. *EMBO J.* 1998;17(12):3363–71.
112. Kanwal C, Li H, Lim CS. Model system to study classical nuclear export signals. *AAPS PharmSci.* 2002;4(3):E18.
113. la Cour T, Kiemer L, Molgaard A, Gupta R, Skriver K, Brunak S. Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng Des Sel.* 2004;17(6):527–36.
114. Garcia-Perez JL, Morell M, Scheys JO, Kulpa DA, Morell S, Carter CC, Hammer GD, Collins KL, O'Shea KS, Menendez P, Moran JV. Epigenetic silencing of engineered L1 retrotransposition events in human embryonic carcinoma cells. *Nature.* 2010;466(7307):769–73.
115. Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, Donaldson D, Goto J, O'Regan JP, Deng HX, et al. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature.* 1993;362(6415):59–62.
116. Majounie E, Renton AE, Mok K, Dopper EG, Waite A, Rollinson S, Chio A, Restagno G, Nicolaou N, Simon-Sanchez J, van Swieten JC, Abramzon Y, Johnson JO, Sendtner M, Pampfllett R, Orrell RW, Mead S, Sidle KC, Houlden H, Rohrer JD, Morrison KE, Pall H, Talbot K, Ansorge O, Hernandez DG, Arepalli S, Sabatelli M, Mora G, Corbo M, Giannini F, Calvo A, Englund E, Borghero G, Floris GL, Remes AM, Laaksovirta H, McCluskey L, Trojanowski JQ, Van Deerlin VM, Schellenberg GD, Nalls MA, Drory VE, Lu CS, Yeh TH, Ishiura H, Takahashi Y, Tsuji S, Le Ber I, Brice A, Drepper C, Williams N, Kirby J, Shaw P, Hardy J, Tienari PJ, Heutink P, Morris HR, Pickering-Brown S, Traynor BJ. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* 2012;11(4):323–30.
117. Boeynaems S, Bogaert E, Van Damme P, Van Den Bosch L. Inside out: the role of nucleocytoplasmic transport in ALS and FTL. *Acta Neuropathol.* 2016;132(2):159–73.
118. Ederle H, Dormann D. TDP-43 and FUS en route from the nucleus to the cytoplasm. *FEBS Lett.* 2017;591(11):1489–507.
119. Colombrita C, Zennaro E, Fallini C, Weber M, Sommacal A, Buratti E, Silani V, Ratti A. TDP-43 is recruited to stress granules in conditions of oxidative insult. *J Neurochem.* 2009;111(4):1051–61.
120. Liu-Yesucevitz L, Bilgutay A, Zhang YJ, Vanderweyde T, Citro A, Mehta T, Zaurur N, McKee A, Bowser R, Sherman M, Petrucelli L, Wolozin B. TAR DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PLoS One.* 2010;5(10):e13250.
121. Dewey CM, Cenik B, Sephton CF, Dries DR, Mayer P 3rd, Good SK, Johnson BA, Herz J, Yu G. TDP-43 is directed to stress granules by sorbitol, a novel physiological osmotic and oxidative stressor. *Mol Cell Biol.* 2011;31(5):1098–108.
122. Buratti E, Brindisi A, Giombi M, Tisminetzky S, Ayala YM, Baralle FE. TDP-43 binds heterogeneous nuclear ribonucleoprotein A/B through its C-terminal tail: an important region for the inhibition of cystic fibrosis transmembrane conductance regulator exon 9 splicing. *J Biol Chem.* 2005;280(45):37572–84.
123. Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z, MacLac KS, Freibaum B, Li S, Mollie X, Kanagaraj AP, Carter R, Boylan KB, Wojtas AM, Rademakers R, Pinkus JL, Greenberg SA, Trojanowski JQ, Traynor BJ, Smith BN, Topp S, Gkazi AS, Miller J, Shaw CE, Kottlors M, Kirschner J, Pestronk A, Li YR, Ford AF, Gitler AD, Benatar M, King OD, Kimonis VE, Ross ED, Weihl CC, Shorter J, Taylor JP. Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature.* 2013;495(7442):467–73.
124. Goodier JL, Cheung LE, Kazazian HH. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. *Nucleic Acids Res.* 2013;41(15):7401–19.
125. Mackenzie IR, Nicholson AM, Sarkar M, Messing J, Purice MD, Pottier C, Annu K, Baker M, Perkerson RB, Kurti A, Matchett BJ, Mittag T, Temirov J, Hsiung GR, Krieger C, Murray ME, Kato M, Fryer JD, Petrucelli L, Zinman L, Weintraub S, Mesulam M, Keith J, Zivkovic SA, Hirsch-Reinshagen V, Roos RP,

- Zuchner S, Graff-Radford NR, Petersen RC, Caselli RJ, Wszolek ZK, Finger E, Lippa C, Lacomis D, Stewart H, Dickson DW, Kim HJ, Rogava E, Bigio E, Boylan KB, Taylor JP, Rademakers R. TIA1 Mutations in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Promote Phase Separation and Alter Stress Granule Dynamics. *Neuron*. 2017;95(4):808–16 e9.
126. Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin WL, DeJesus-Hernandez M, van Blitterswijk MM, Jansen-West K, Paul JW 3rd, Rademakers R, Boylan KB, Dickson DW, Petrucelli L. Unconventional translation of C9orf72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron*. 2013;77(4):639–46.
127. Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, Schmid B, Kretzschmar HA, Cruts M, Van Broeckhoven C, Haass C, Edbauer D. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTD/ALS. *Science*. 2013;339(6125):1335–8.
128. Yamakawa M, Ito D, Honda T, Kubo K, Noda M, Nakajima K, Suzuki N. Characterization of the dipeptide repeat protein in the molecular pathogenesis of c9FTD/ALS. *Hum Mol Genet*. 2015;24(6):1630–45.
129. Mizielinska S, Gronke S, Niccoli T, Ridler CE, Clayton EL, Devoy A, Moens T, Norona FE, Woollacott IOC, Pietrzyk J, Cleverley K, Nicoll AJ, Pickering-Brown S, Dols J, Cabecinha M, Hendrich O, Fratta P, Fisher EMC, Partridge L, Isaacs AM. C9orf72 repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science*. 2014;345(6201):1192–4.
130. Ji AL, Zhang X, Chen WW, Huang WJ. Genetics insight into the amyotrophic lateral sclerosis/frontotemporal dementia spectrum. *J Med Genet*. 2017;54(3):145–54.
131. Zhang YJ, Gendron TF, Grima JC, Sasaguri H, Jansen-West K, Xu YF, Katzman RB, Gass J, Murray ME, Shinohara M, Lin WL, Gass A, Stankowski JN, Daugherty L, Tong J, Perkerson EA, Yue M, Chew J, Castanedes-Casey M, Kurti A, Wang ZS, Liesinger AM, Baker JD, Jiang J, Lagier-Tourenne C, Edbauer D, Cleveland DW, Rademakers R, Boylan KB, Bu G, Link CD, Dickey CA, Rothstein JD, Dickson DW, Fryer JD, Petrucelli L. C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nat Neurosci*. 2016;19(5):668–77.
132. Sendtner M. TDP-43: multiple targets, multiple disease mechanisms? *Nat Neurosci*. 2011;14(4):403–5.
133. Lagier-Tourenne C, Polymenidou M, Cleveland DW. TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. *Hum Mol Genet*. 2010;19(11):R46–64.
134. Ratti A, Buratti E. Physiological functions and pathobiology of TDP-43 and FUS/TLS proteins. *J Neurochem*. 2016;138(Suppl 1):95–111.
135. Estes PS, Boehringer A, Zwick R, Tang JE, Grigsby B, Zarnescu DC. Wild-type and A315T mutant TDP-43 exert differential neurotoxicity in a *Drosophila* model of ALS. *Hum Mol Genet*. 2011;20(12):2308–21.
136. Xu Z, Yang C. TDP-43-The key to understanding amyotrophic lateral sclerosis. *Rare Dis*. 2014;2(1):e944443.
137. Li W, Jin Y, Prazak L, Hammell M, Dubnau J. Transposable elements in TDP-43-mediated neurodegenerative disorders. *PLoS One*. 2012;7(9):e44099.
138. Manghera M, Ferguson-Parry J, Douville RN. TDP-43 regulates endogenous retrovirus-K viral protein accumulation. *Neurobiol Dis*. 2016;94:226–36.
139. Krug L, Chatterjee N, Borges-Monroy R, Hearn S, Liao WW, Morrill K, Prazak L, Rozhkov N, Theodorou D, Hammell M, Dubnau J. Retrotransposon activation contributes to neurodegeneration in a *Drosophila* TDP-43 model of ALS. *PLoS Genet*. 2017;13(3):e1006635.
140. Moldovan JB. Identification of cellular host factors that associate with LINE-1 ORF1p and the effect of the Zinc Finger Antiviral Protein ZAP on LINE-1 retrotransposition [Ph.D. Thesis]: University of Michigan; 2015.
141. Taylor MS, Altukhov I, Molloy KR, Mita P, Jiang H, Adney EM, Wudzinska A, Badri S, Ischenko D, Eng G, Burns KH, Fenyo D, Chait BT, Alexeev D, Rout MP, Boeke JD, LaCava J. Dissection of affinity captured LINE-1 macromolecular complexes. *Elife*. 2018;77:e30094.
142. Gao J, Wang L, Huntley ML, Perry G, Wang X. Pathomechanisms of TDP-43 in neurodegeneration. *J Neurochem*. 2018; <https://doi.org/10.1111/jnc.14327>.
143. Vogt MA, Ehsaei Z, Knuckles P, Higginbottom A, Helmbrecht MS, Kunath T, Eggan K, Williams LA, Shaw PJ, Wurst W, Floss T, Huber AB, Taylor V. TDP-43 induces p53-mediated cell death of cortical progenitors and immature neurons. *Sci Rep*. 2018;8(1):8097.
144. Dewannieux M, Dupressoir A, Harper F, Pierron G, Heidmann T. Identification of autonomous IAP LTR retrotransposons mobile in mammalian cells. *Nat Genet*. 2004;36(5):534–9.
145. Crowther P, Doherty J, Linsenmeyer M, Williamson M, Woodcock D. Revised genomic consensus for the hypermethylated CpG island region of the human L1 transposon and integration sites of full length L1 elements from recombinant clones made using methylation-tolerant host strains. *Nucleic Acids Res*. 1991;19(9):2395–401.
146. Munoz-Lopez M, Garcia-Canadas M, Macia A, Morell S, Garcia-Perez JL. Analysis of LINE-1 expression in human pluripotent cells. *Methods Mol Biol*. 2012;873:113–25.
147. Ling JP, Pletnikova O, Troncoso JC, Wong PC. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science*. 2015; 349(6248):650–5.
148. Kapeli K, Pratt GA, Vu AQ, Hutt KR, Martinez FJ, Sundararaman B, Batra R, Freese P, Lambert NJ, Huelga SC, Chun SJ, Liang TY, Chang J, Donohue JP, Shiue L, Zhang J, Zhu H, Cambi F, Kasarskis E, Hoon S, Ares M Jr, Burge CB, Ravits J, Rigo F, Yeo GW. Distinct and shared functions of ALS-associated proteins TDP-43, FUS and TAF15 revealed by multisystem analyses. *Nat Commun*. 2016;7:12143.
149. Svetoni F, Frisone P, Paronetto MP. Role of FET proteins in neurodegenerative disorders. *RNA Biol*. 2016;13(11):1089–102.
150. Jin Y, Tam OH, Paniagua E, Hammell M. Tetrascripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics*. 2015;31(22):3593–9.
151. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res*. 2005;110(1-4):462–7.
152. Bao W, Kojima KK, Kohany O. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA*. 2015;6:11.
153. Pesiridis GS, Lee VM, Trojanowski JQ. Mutations in TDP-43 link glycine-rich domain functions to amyotrophic lateral sclerosis. *Hum Mol Genet*. 2009; 18(2):R156–62.
154. Winton MJ, Van Deerlin VM, Kwong LK, Yuan W, Wood EM, Yu CE, Schellenberg GD, Rademakers R, Caselli R, Karydas A, Trojanowski JQ, Miller BL, Lee VM. A90V TDP-43 variant results in the aberrant localization of TDP-43 in vitro. *FEBS Lett*. 2008;582(15):2252–6.
155. Kovacs GG, Murrell JR, Horvath S, Haraszti L, Majtenyi K, Molnar MJ, Budka H, Ghetti B, Spina S. TARDBP variation associated with frontotemporal dementia, supranuclear gaze palsy, and chorea. *Mov Disord*. 2009;24(12):1843–7.
156. Hooper LV, Stappenbeck TS, Hong CV, Gordon JL. Angiogenins: a new class of microbicidal proteins involved in innate immunity. *Nat Immunol*. 2003; 4(3):269–73.
157. Kieran D, Sebastia J, Greenway MJ, King MA, Connaughton D, Concannon CG, Fenner B, Hardiman O, Prehn JH. Control of motoneuron survival by angiogenin. *J Neurosci*. 2008;28(52):14056–61.
158. Li S, Chen Y, Sun D, Bai R, Gao X, Yang Y, Sheng J, Xu Z. Angiogenin Prevents Progranulin A9D Mutation-Induced Neuronal-Like Cell Apoptosis Through Cleaving tRNAs into tiRNAs. *Mol Neurobiol*. 2017;55(2):1338–51.
159. Bradshaw WJ, Rehman S, Pham TT, Thiagarajan N, Lee RL, Subramanian V, Acharya KR. Structural insights into human angiogenin variants implicated in Parkinson's disease and Amyotrophic Lateral Sclerosis. *Sci Rep*. 2017;7:41996.
160. Gellera C, Colombrita C, Ticozzi N, Castellotti B, Bragato C, Ratti A, Taroni F, Silani V. Identification of new ANG gene mutations in a large cohort of Italian patients with amyotrophic lateral sclerosis. *Neurogenetics*. 2008;9(1):33–40.
161. van Es MA, Schelhaas HJ, van Vught PW, Ticozzi N, Andersen PM, Groen EJ, Schulte C, Blauw HM, Koppers M, Diekstra FP, Fumoto K, Al LC, Keagle P, Bloem BR, Scheffer H, van Nuenen BF, van Blitterswijk M, van Rheeunen W, Wills AM, Lowe PP, Hu GF, Yu W, Kishikawa H, Wu D, Folkerth RD, Mariani C, Goldwurm S, Pezzoli G, Van Damme P, Lemmens R, Dahlberg C, Birve A, Fernandez-Santiago R, Waibel S, Klein C, Weber M, van der Kooij AJ, de Visser M, Verbaan D, van Hilten JJ, Heutink P, Hennekam EA, Cuppen E, Berg D, Brown RH Jr, Silani V, Gasser T, Ludolph AC, Robberecht W, Ophoff RA, Veldink JH, Pasterkamp RJ, de Bakker PI, Landers JE, van de Warrenburg BP, van den Berg LH. Angiogenin variants in Parkinson disease and amyotrophic lateral sclerosis. *Ann Neurol*. 2011;70(6):964–73.
162. Bennett EA, Keller H, Mills RE, Schmidt S, Moran JV, Weichenrieder O, Devine SE. Active Alu retrotransposons in the human genome. *Genome Res*. 2008; 18(12):1875–83.
163. Kryatova MS, Steranka JP, Burns KH, Payer LM. Insertion and deletion polymorphisms of the ancient Alu family in the human genome. *Mobile DNA*. 2017;8:6.
164. Batzer MA, Deininger PL. Alu repeats and human genomic diversity. *Nat Rev Genet*. 2002;3(5):370–9.
165. Hancks DC, Kazazian HH. Roles for retrotransposon insertions in human disease. *Mobile DNA*. 2016;7(1):9.

166. Seleme MC, Vetter MR, Cordaux R, Bastone L, Batzer MA, Kazazian HH Jr. Extensive individual variation in L1 retrotransposition capability contributes to human genetic diversity. *Proc Natl Acad Sci U S A*. 2006;103(17):6611–6.
167. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Embryonic stem cell lines derived from human blastocysts. *Science*. 1998;282(5391):1145–7.
168. Garcia-Perez JL, Marchetto MC, Muotri AR, Coufal NG, Gage FH, O'Shea KS, Moran JV. LINE-1 retrotransposition in human embryonic stem cells. *Hum Mol Genet*. 2007;16(13):1569–77.
169. Moszczynska A, Flack A, Qiu P, Muotri AR, Killinger BA. Neurotoxic Methamphetamine Doses Increase LINE-1 Expression in the Neurogenic Zones of the Adult Rat Brain. *Sci Rep*. 2015;5:14356.
170. Sur D, Kustwar RK, Budania S, Mahadevan A, Hancks DC, Yadav V, Shankar SK, Mandal PK. Detection of the LINE-1 retrotransposon RNA-binding protein ORF1p in different anatomical regions of the human brain. *Mobile DNA*. 2017;8:17.
171. Prudencio M, Belzil VV, Batra R, Ross CA, Gendron TF, Pregent LJ, Murray ME, Overstreet KK, Piazza-Johnston AE, Desaro P, Bieniek KF, DeTure M, Lee WC, Biendarra SM, Davis MD, Baker MC, Perkinson RB, van Blitterswijk M, Stetler CT, Rademakers R, Link CD, Dickson DW, Boylan KB, Li H, Petrucelli L. Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat Neurosci*. 2015;18(8):1175–82.
172. Batra R, Hutt K, Vu A, Rabin SJ, Baughn MW, Libby RT, Hoon S, Ravits J, Yeo GW. Gene Expression Signatures of Sporadic ALS Motor Neuron Populations. *bioRxiv*. 2016. <https://doi.org/10.1101/038448>.
173. Larsen PA, Hunnicutt KE, Larsen RJ, Yoder AD, Saunders AM. Warning SINEs: Alu elements, evolution of the human brain, and the spectrum of neurological disease. *Chromosome Res*. 2018;26(1-2):93–111.
174. Deininger P, Morales ME, White TB, Baddoo M, Hedges DJ, Servant G, Srivastava S, Smither ME, Concha M, DeHaro DL, Flemington EK, Belancio VP. A comprehensive approach to expression of L1 loci. *Nucleic Acids Res*. 2017;45(5):e31.
175. Philippe C, Vargas-Landin DB, Doucet AJ, van Essen D, Vera-Otarola J, Kuciak M, Corbin A, Nigumann P, Cristofari G. Activation of individual L1 retrotransposon instances is restricted to cell-type dependent permissive loci. *Elife*. 2016;5:e13926.
176. Branciforte D, Martin SL. Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. *Mol Cell Biol*. 1994;14(4):2584–92.
177. Horn AV, Celic I, Dong C, Martirosyan I, Han JS. A conserved role for the ESCRT membrane budding complex in LINE retrotransposition. *PLoS Genet*. 2017;13(6):e1006837.
178. Tiruchinapalli DM, Oleynikov Y, Kelic S, Shenoy SM, Hartley A, Stanton PK, Singer RH, Bassell GJ. Activity-dependent trafficking and dynamic localization of zipcode binding protein 1 and beta-actin mRNA in dendrites and spines of hippocampal neurons. *J Neurosci*. 2003;23(8):3251–61.
179. Donnelly CJ, Willis DE, Xu M, Tep C, Jiang C, Yoo S, Schanen NC, Kim-Safran CB, van Minnen J, English A, Yoon SO, Bassell GJ, Twiss JL. Limited availability of ZBP1 restricts axonal mRNA localization and nerve regeneration capacity. *EMBO J*. 2011;30(22):4665–77.
180. Fallini C, Rouanet JP, Donlin-Asp PG, Guo P, Zhang H, Singer RH, Rossoll W, Bassell GJ. Dynamics of survival of motor neuron (SMN) protein interaction with the mRNA-binding protein IMP1 facilitates its trafficking into motor neuron axons. *Dev Neurobiol*. 2014;74(3):319–32.
181. Jonson L, Vikesaa J, Krogh A, Nielsen LK, Hansen T, Borup R, Johnsen AH, Christiansen J, Nielsen FC. Molecular composition of IMP1 ribonucleoprotein granules. *Mol Cell Proteomics*. 2007;6(5):798–811.
182. Harris CR, Normart R, Yang Q, Stevenson E, Haffty BG, Ganesan S, Cordon-Cardo C, Levine AJ, Tang LH. Association of nuclear localization of a Long Interspersed Nuclear Element-1 Protein in breast tumors with poor prognostic outcomes. *Genes Cancer*. 2010;1(2):115–24.
183. Lucchinetti E, Feng J, Silva R, Tolstonog GV, Schaub MC, Schumann GG, Zaugg M. Inhibition of LINE-1 expression in the heart decreases ischemic damage by activation of Akt/PKB signaling. *Physiol Genomics*. 2006;25(2):314–24.
184. Kirilyuk A, Tolstonog GV, Damert A, Held U, Hahn S, Lower R, Buschmann C, Horn AV, Traub P, Schumann GG. Functional endogenous LINE-1 retrotransposons are expressed and mobilized in rat chloroleukemia cells. *Nucleic Acids Res*. 2008;36(2):648–65.
185. Soper SF, van der Heijden GW, Hardiman TC, Goodheart M, Martin SL, de Boer P, Bortvin A. Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. *Dev Cell*. 2008;15(2):285–97.
186. Malki S, van der Heijden GW, O'Donnell Kathryn A, Martin Sandra L, Bortvin A. A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. *Dev Cell*. 2014;29(5):521–33.
187. Luo YB, Zhang L, Lin ZL, Ma JY, Jia J, Namgoong S, Sun QY. Distinct subcellular localization and potential role of LINE1-ORF1P in meiotic oocytes. *Histochem Cell Biol*. 2016;145(1):93–104.
188. Idica A, Sevrioukov EA, Zisoulis DG, Hamdorf M, Daugaard I, Kadandale P, Pedersen IM. MicroRNA miR-128 represses LINE-1 (L1) retrotransposition by down-regulating the nuclear import factor TNPO1. *J Biol Chem*. 2017; 292(50):20494–508.
189. Mandal PK, Ewing AD, Hancks DC, Kazazian HH Jr. Enrichment of processed pseudogene transcripts in L1-ribonucleoprotein particles. *Hum Mol Genet*. 2013;22(18):3730–48.
190. Moldovan JB, Moran JV. The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition. *PLoS Genet*. 2015;11(5):e1005121.
191. Gasior SL, Wakeman TP, Xu B, Deininger PL. The human LINE-1 retrotransposon creates DNA double-strand breaks. *J Mol Biol*. 2006;357(5): 1383–93.
192. Belgnaoui SM, Gosden RG, Semmes OJ, Haoudi A. Human LINE-1 retrotransposon induces DNA damage and apoptosis in cancer cells. *Cancer Cell Int*. 2006;6(1):1.
193. Wallace NA, Belancio VP, Deininger PL. L1 mobile element expression causes multiple types of toxicity. *Gene*. 2008;419(1-2):75–81.
194. Yu Q, Carbone CJ, Katlinskaya YV, Zheng H, Zheng K, Luo M, Wang PJ, Greenberg RA, Fuchs SY. Type I Interferon controls propagation of Long Interspersed Element-1. *J Biol Chem*. 2015;290(16):10191–9.
195. Thomas CA, Tejwani L, Trujillo CA, Negraes PD, Herai RH, Mesci P, Macia A, Crow YJ, Muotri AR. Modeling of TREG1-Dependent Autoimmune Disease using Human Stem Cells Highlights L1 Accumulation as a Source of Neuroinflammation. *Cell Stem Cell*. 2017;21(3):319–31 e8.
196. Ou SH, Wu F, Harrich D, Garcia-Martinez LF, Gaynor RB. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J Virol*. 1995;69(6): 3584–96.
197. Nehls J, Koppensteiner H, Brack-Werner R, Floss T, Schindler M. HIV-1 replication in human immune cells is independent of TAR DNA binding protein 43 (TDP-43) expression. *PLoS One*. 2014;9(8):e105478.
198. Douville RN, Nath A. Human Endogenous Retrovirus-K and TDP-43 Expression Bridges ALS and HIV Neuropathology. *Front Microbiol*. 2017;8:1986.
199. Polyimenidou M, Lagier-Tourenne C, Hutt KR, Huelga SC, Moran J, Liang TY, Ling SC, Sun E, Wancewicz E, Mazur C, Kordasiewicz H, Sedaghat Y, Donohue JP, Shiue L, Bennett CF, Yeo GW, Cleveland DW. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci*. 2011;14(4):459–68.
200. Shan X, Chiang PM, Price DL, Wong PC. Altered distributions of Gemini of coiled bodies and mitochondria in motor neurons of TDP-43 transgenic mice. *Proc Natl Acad Sci U S A*. 2010;107(37):16325–30.
201. Dhellin O, Maestre J, Heidmann T. Functional differences between the human LINE retrotransposon and retroviral reverse transcriptases for invivo mRNA reverse transcription. *EMBO J*. 1997;16(21):6590–602.
202. Garcia Perez JL, Alarcon-Riquelme ME. The TREG1 Dinosaur Bites the Brain through the LINE. *Cell Stem Cell*. 2017;21(3):287–8.
203. Sciamanna I, Gualtieri A, Cossetti C, Osimo EF, Ferracin M, Macchia G, Arico E, Prosseda G, Vitullo P, Misteli T, Spadafora C. A tumor-promoting mechanism mediated by retrotransposon-encoded reverse transcriptase is active in human transformed cell lines. *Oncotarget*. 2013;4(12):2271–87.
204. Mayer J, Harz C, Sanchez L, Pereira GC, Maldener E, Heras SR, Ostrow LW, Ravits J, Batra R, Meese E, Garcia-Perez JL, Goodier JL. Transcriptional profiling of HERV-K(HML-2) in amyotrophic lateral sclerosis and potential implications for expression of HML-2 proteins. *Mol Neurodegener*. 2018; 13(1):39.
205. Chestnut BA, Chang Q, Price A, Lesuisse C, Wong M, Martin LJ. Epigenetic regulation of motor neuron cell death through DNA methylation. *J Neurosci*. 2011;31(46):16619–36.
206. Martin LJ, Wong M. Aberrant regulation of DNA methylation in amyotrophic lateral sclerosis: a new target of disease mechanisms. *Neurotherapeutics*. 2013;10(4):722–33.
207. Coppede F, Stoccoro A, Mosca L, Gallo R, Tarlarini C, Lunetta C, Marocchi A, Migliore L, Penco S. Increase in DNA methylation in patients with amyotrophic lateral sclerosis carriers of not fully penetrant SOD1 mutations. *Amyotroph Lateral Scler Frontotemporal Degener*. 2018;19(1-2):93–101.

208. Hamzeiy H, Savas D, Tunca C, Sen NE, Gundogdu Eken A, Sahbaz I, Calini D, Tiloca C, Ticozzi N, Ratti A, Silani V, Basak AN. Elevated Global DNA Methylation Is Not Exclusive to Amyotrophic Lateral Sclerosis and Is Also Observed in Spinocerebellar Ataxia Types 1 and 2. *Neurodegener Dis.* 2018;18(1):38–48.
209. de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet.* 2011; 7(12):e1002384.
210. Coady TH, Manley JL. ALS mutations in TLS/FUS disrupt target gene expression. *Genes Dev.* 2015;29(16):1696–706.
211. Maruyama H, Morino H, Ito H, Izumi Y, Kato H, Watanabe Y, Kinoshita Y, Kamada M, Nodera H, Suzuki H, Komure O, Matsuura S, Kobatake K, Morimoto N, Abe K, Suzuki N, Aoki M, Kawata A, Hirai T, Kato T, Ogasawara K, Hirano A, Takumi T, Kusaka H, Hagiwara K, Kaji R, Kawakami H. Mutations of optineurin in amyotrophic lateral sclerosis. *Nature.* 2010;465(7295):223–6.
212. Mao YS, Sunwoo H, Zhang B, Spector DL. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat Cell Biol.* 2011;13(1):95–101.
213. Liu-Yesucevitz L, Lin AY, Ebata A, Boon JY, Reid W, Xu YF, Kobrin K, Murphy GJ, Petrucelli L, Wolozin B. ALS-linked mutations enlarge TDP-43-enriched neuronal RNA granules in the dendritic arbor. *J Neurosci.* 2014;34(12):4167–74.
214. Zhang T, Baldie G, Periz G, Wang J. RNA-processing protein TDP-43 regulates FOXO-dependent protein quality control in stress response. *PLoS Genet.* 2014;10(10):e1004693.
215. Farrarwell NE, Lambert-Smith IA, Warraich ST, Blair IP, Saunders DN, Hatters DM, Yerbury JJ. Distinct partitioning of ALS associated TDP-43, FUS and SOD1 mutants into cellular inclusions. *Sci Rep.* 2015;5:13416.
216. Osaka M, Ito D, Yagi T, Nihei Y, Suzuki N. Evidence of a link between ubiquilin 2 and optineurin in amyotrophic lateral sclerosis. *Hum Mol Genet.* 2015;24(6):1617–29.
217. Soifer H, Higo C, Kazazian HH Jr, Moran JV, Mitani K, Kasahara N. Stable integration of transgenes delivered by a retrotransposon-adenovirus hybrid vector. *Hum Gene Ther.* 2001;12(11):1417–28.
218. Palmer DJ, Ng P. Methods for the production of helper-dependent adenoviral vectors. *Methods Mol Biol.* 2008;433:33–53.
219. Hulme AE, Bogerd HP, Cullen BR, Moran JV. Selective inhibition of Alu retrotransposition by APOBEC3G. *Gene.* 2007;390(1):199–205.
220. Stoecklin G, Mayo T, Anderson P. ARE-mRNA degradation requires the 5'-3' decay pathway. *EMBO Rep.* 2006;7(1):72–7.
221. Goodier JL, Cheung LE, Kazazian HH. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. *PLoS Genet.* 2012;8(10):e1002941.
222. Goodier JL, Pereira GC, Cheung LE, Rose RJ, Kazazian HH. The Broad-Spectrum Antiviral Protein ZAP Restricts Human Retrotransposition. *PLoS Genet.* 2015;11(5):e1005252.
223. Kumaki Y, Oda M, Okano M. QUMA: quantification tool for methylation analysis. *Nucleic Acids Res.* 2008;36(Web Server issue):W170–5.
224. Klawitter S, Fuchs NV, Upton KR, Muñoz-Lopez M, Shukla R, Wang J, Garcia-Cañadas M, Lopez-Ruiz C, Gerhardt DJ, Sebe A, Grabundzija I, Merkert S, Gerdes P, Pulgarin JA, Bock A, Held U, Witthuhn A, Haase A, Sarkadi B, Löwer J, Wolvetang EJ, Martin U, Ivics Z, Izsvák Z, Garcia-Perez JL, Faulkner GJ, Schumann GG. Reprogramming triggers endogenous L1 and Alu retrotransposition in human induced pluripotent stem cells. *Nature Commun.* 2016;7:10286.
225. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 2001;25(4):402–8.
226. Wissing S, Muñoz-Lopez M, Macia A, Yang Z, Montano M, Collins W, Garcia-Perez JL, Moran JV, Greene WC. Reprogramming somatic cells into iPS cells activates LINE-1 retroelement mobility. *Hum Mol Genet.* 2012;21(1):208–18.
227. Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, Kent WJ. The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* 2004; 32(Database issue):D493–6.
228. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.* 2017;<https://doi.org/10.1093/bib/bbx108>.
229. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thieler T, Ashton B, Meintjes P, Drummond A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 2012;28(12):1647–9.
230. Arokium H, Kamata M, Kim S, Kim N, Liang M, Presson AP, Chen IS. Deep sequencing reveals low incidence of endogenous LINE-1 retrotransposition in human induced pluripotent stem cells. *PLoS One.* 2014;9(10):e108682.
231. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357–9.
232. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26(1):139–40.
233. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15–21.
234. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 2010;11(10):R106.
235. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;12(4):357–60.
236. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010;26(6):841–2.
237. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 2014;15(2):R29.
238. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

