Distal Alternative Last Exons Localize mRNAs to Neural Projections

Citation: Taliaferro, J. Matthew et al. "Distal Alternative Last Exons Localize mRNAs to Neural Projections." Molecular Cell 61, 6 (March 2016): 821–833 © 2016 Elsevier Inc

As Published: http://dx.doi.org/10.1016/J.MOLCEL.2016.01.020

Publisher: Elsevier BV

Persistent URL: http://hdl.handle.net/1721.1/116334

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

HHS Public Access

Author manuscript

Mol Cell. Author manuscript; available in PMC 2017 March 17.

Published in final edited form as:

Mol Cell. 2016 March 17; 61(6): 821–833. doi:10.1016/j.molcel.2016.01.020.

Distal Alternative Last Exons Localize mRNAs to Neural Projections

J. Matthew Taliaferro[1], **Marina Vidaki**[2], **Ruan Oliveira**[3], **Sara Olson**[4], **Lijun Zhan**[4], **Tanvi Saxena**[1,2], **Eric T. Wang**[3], **Brenton R. Graveley**[4], **Frank B. Gertler**[1,2], **Maurice S. Swanson**[3], and **Christopher B. Burge**[1,*]

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142

[2]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02142

[3]Department of Molecular Genetics and Microbiology, University of Florida College of Medicine, Gainesville, FL 32610

[4]Department of Genetics and Genome Sciences, Institute for Systems Genomics, University of Connecticut Health Center, Farmington, CT 06030

Summary

Spatial restriction of mRNA to distinct subcellular locations enables local regulation and synthesis of proteins. However, the organizing principles of mRNA localization remain poorly understood. Here, we analyzed subcellular transcriptomes of neural projections and soma of primary mouse cortical neurons and two neuronal cell lines and found that alternative last exons (ALEs) often confer isoform-specific localization. Surprisingly, gene-distal ALE isoforms were 4 times more often localized to neurites than gene-proximal isoforms. Localized isoforms were induced during neuronal differentiation and enriched for motifs associated with muscleblind-like (Mbnl) family RNA-binding proteins. Depletion of *Mbnl1* and/or *Mbnl2* reduced localization of hundreds of transcripts, implicating Mbnls in localization of mRNAs to neurites. We provide evidence supporting a model in which the linkage between genomic position of ALEs and subcellular localization enables coordinated induction of localization-competent mRNA isoforms through a post-transcriptional regulatory program that is induced during differentiation and reversed in cellular reprogramming and cancer.

*address correspondence to: cburge@mit.edu.

Author Contributions

The experiments were conceived, designed and analyzed by JMT and CBB and executed by JMT. Cortical neurons were dissected by MV. Mouse husbandry, breeding, and genotyping was done by RO. Knockdowns of RBPs in K562 cells were performed by SO and LZ. RNA-seq experiments from mouse fibroblasts were performed by TS and ETW. The manuscript was written by JMT and CBB and edited by BRG, FBG and MSS.

Accession Numbers

RNA-seq data has been deposited to the Gene Expression Omnibus (GEO) under accession number GSE67828.

INTRODUCTION

Patterns of protein localization contribute to the specialized functions of cellular compartments and are often driven by localization of the corresponding mRNA. RNA localization is quite widespread, with up to 70% of mRNAs nonuniformly localized in *Drosophila* embryos, and similar localizations observed for the encoded proteins (Lécuyer et al., 2007). In mammalian cells, mRNAs encoding proteins of different functional classes have distinct patterns of localization (Wang et al., 2012). Proper germ cell formation in the fly embryo relies in part on the high concentration of *Oskar* protein at the anterior end, which is achieved through the localization of *oskar* mRNA (Ephrussi et al., 1991). Mammalian fibroblasts use enrichment of beta-actin mRNA at the leading edges of lamellipodia for directed cell motility (Mili et al., 2008), and many neuronal messages are enriched in neurites (axons, dendrites, and their precursors), including mRNAs important for proper response to stimuli (Leung et al., 2006).

Polarized cells often receive different stimuli from the apical and basal surfaces and must direct their responses to the appropriate cellular location. In some cases, signaling to up- or down-regulate translation of specific mRNAs in the vicinity of the stimulus may produce a rapid and robust response (Buxbaum et al., 2014). Mis-regulation of RNA localization in neurons is associated with many neurological diseases, including spinomuscular atrophy (SMA) and amyotrophic lateral sclerosis (ALS) (Paushkin et al., 2002; Tolino et al., 2012).

In a handful of well-studied cases, localization is known to involve specific RNA binding proteins (RBPs) that associate with mRNAs, motor proteins that transport mRNA along the cytoskeleton, and adapter proteins that link RBP to motor protein (Martin and Ephrussi, 2009). The RBPs that target messages for localization usually associate with RNA based on the presence of linear sequence motifs or RNA secondary structures (Ghosh et al., 2012; Ross et al., 1997). However, the extent to which they regulate localization transcriptome-wide is often unknown. Known RNA elements associated with RNA localization are often found in the 3′ untranslated regions (UTRs) of messages (Andreassi and Riccio, 2009).

Thousands of mammalian genes generate mRNA isoforms differing in their 3′ UTRs. Alternative 3′ UTR isoforms are highly conserved between human and mouse, contain many regulatory elements, and have been implicated in a variety of cellular processes (Miura et al., 2013). Generally, expression of isoforms with shorter 3′ UTRs is associated with rapidly proliferating cells (Mayr and Bartel, 2009; Sandberg et al., 2008), while expression of longer 3′ UTR isoforms increases during development, with brain and muscle tending to express messages with the longest 3′ UTRs (Ji et al., 2009; Ramsköld et al., 2009; Ulitsky et al., 2012). While specific alternative 3′ UTRs can have large effects on transcript stability and protein production (Mayr and Bartel, 2009; Sandberg et al., 2008; Yang et al., 2003), a recent genome-wide analysis found that most alternative 3′ UTRs have little or no effect on either mRNA stability or translation (Spies et al., 2013), raising questions about why alternative 3′ UTRs are so abundant and conserved.

Because of the importance of mRNA localization in neurons and the large physical distances involved, we chose to study mRNA localization in neuronal cells. Although hundreds of

mRNAs are known to be enriched in neurites (Cajigas et al., 2012; Gumy et al., 2011; Minis et al., 2013; Taylor et al., 2009), the RNA features required for localization remain largely unknown. We sought to determine RNA sequences and associated *trans*-factors that regulate mRNA localization in neuronal systems. We used differential enrichment of mRNA isoforms in the transcriptomes of soma and neurite to identify RNA regions and motifs associated with localization. Our results implicate proteins of the muscleblind (Mbnl) family in localization of hundreds of mRNA isoforms in neurons, and uncover a surprising connection between relative genomic position and subcellular localization for a major class of alternative 3′ UTR isoforms.

RESULTS

Similar global patterns of mRNA localization in CAD, N2A and primary neuronal cells

To identify RNA transcripts enriched in neuronal projections, we mechanically fractionated mouse neuronal cells using porous membranes that allow projection growth through the membrane (Poon et al., 2006). To increase the robustness of our results, we used two different cell lines: N2A, a brain-derived neuroblastoma line; and CAD, a brain-derived catecholaminergic neuronal line. In addition, primary cortical neurons from E18.5 mouse embryos were analyzed using the same approach (Fig. 1A–B, Fig. S1A–C). RNA from both fractions was isolated and subjected to strand-specific polyA-selected paired-end RNA-seq analysis. For each gene, a “localization ratio” (LR) was defined as the ratio of expression (measured by fragments per kilobase of mRNA per million mapped reads, FPKM) in the neurite fraction divided by expression in the soma fraction. Thus, genes with log(LR) > 0 are enriched in neurites and those with log(LR) < 0 are enriched in soma. The LR values of genes were highly concordant between the two cell lines ($R_{Spearman} = 0.89$, $p < 2.2 \times 10^{-16}$) (Fig. 1C, S1D). The LR values in these lines were also correlated with those in primary cortical neurons ($R_{Spearman} = 0.38$, $p < 2.2 \times 10^{-16}$) (Fig. S1E) and with those from a similar fractionation of primary mouse dorsal root ganglia (DRG) ($R_{Spearman} = 0.38$, $p < 2.2 \times 10^{-16}$) (Fig. S1F) (Minis et al., 2013), suggesting that these cell lines capture general features of the neuronal RNA localization program.

The localized genes identified here were consistent with previous studies. Several genes with known projection-enriched RNA localization patterns, including beta-actin *(Actb)*, neurogranin (*Nrgn*), and *Ranbp1*, were identified as neurite-enriched in both CAD and N2A cells (Fig. S1G and references in figure legend). Overall, of our stringent group of 778 genes enriched in neurites of both CAD and N2A cells, 537 overlapped with a set of ~4000 genes identified as enriched in the peripheral axons of mouse DRG ($P = 2 \times 10^{-44}$, binomial test) (Minis et al., 2013), and 86 were shared with a set of ~300 genes enriched in the axons of rat cortical neurons ($P = 8 \times 10^{-15}$, binomial test) (Taylor et al., 2009). Messages localized to projections of both cell lines and primary cortical neurons preferentially encoded ribosomal and mitochondrial proteins, consistent with previous reports (Gumy et al., 2011; Moccia et al., 2003) (Fig. S1H, I). Conversely, transcripts localized to the soma fraction were enriched for genes with nuclear functions (Fig. S1H).

Distal alternative last exons are strongly associated with RNA localization

We reasoned that if a pair of transcript isoforms differ in their extent of localization to neurites, the RNA elements driving this difference should be located in the segment(s) that differ between the isoforms (Fig. 1D). We therefore sought to identify pairs of alternative mRNA isoforms that differed in their localization, using the MISO software for statistical analysis of RNA-seq data (Supplementary Experimental Procedures). We focused on four of the most common types of alternative isoforms in mammals: alternative first exons (AFE), alternative last exons (ALE), skipped exons (SE) and consecutive polyadenylation sites (PAS) or "tandem 3′ UTRs" (tandem UTR) (other types of isoforms are shown in Fig. S2). To assess isoform-specific localization, we compared percent spliced in (PSI or ψ) values between compartments. PSI is defined as the fraction of a gene's transcripts that contain the longer ("inclusion") isoform for SEs and tandem UTRs, and as the fraction of transcripts that contain the gene-distal alternative exon for AFEs and ALEs (Fig. 1E).

Differential localization was assessed based on differences in PSI between neurite and soma, defined as $\Delta\psi = \psi_{neurite} - \psi_{soma}$, for over 40,000 alternative isoform pairs derived from a previous RNA-seq analysis of mouse tissues (Merkin et al., 2012). Thus, enrichment of the distal or inclusion isoform (blue) in neurites yields a positive Δψ while enrichment of the proximal or exclusion isoform (red) yields a negative Δψ. By focusing on relative abundance of isoforms, this approach controls for gene-level contributions to localization. As seen for LRs, we observed good agreement between Δψ values measured in CAD and N2A cells, both in the identities of genes and isoforms exhibiting differential isoform enrichment ($P < 2.2 \times 10^{-16}$, binomial test, Fig. S1J) and also in the relative magnitude of enrichment ($R_{Spearman} = 0.74$, Fig. S1K, Table S1). Additionally, we observed reasonable agreement in Δψ values between the cultured cell lines and primary cortical neurons ($R_{Spearman} = 0.35$, Fig. S1L, Table S1) for isoform pairs expressed in both cell lines and primary neurons, further supporting the utility of these cell lines as a model for neuronal RNA localization.

Previously, differential localization of alternative mRNA isoforms has been observed in a few cases (An et al., 2008; Buckley et al., 2011; Harrison et al., 2014; Whittaker et al., 1999). Here, we observed hundreds of isoform pairs with significant differences in ψ between projection and soma, using statistical criteria similar to those used previously for comparisons between cell states or types (Experimental Procedures, Fig. 1E). Thus, differential localization of alternative mRNA isoforms is a relatively common phenomenon. Using more stringent criteria, we identified a confident set of 195 localized ALEs and 96 localized tandem UTRs (Fig. S1O). Of the isoform types analyzed, ALE and tandem 3′ UTRs had the highest absolute and relative numbers of differentially localized pairs in both cell types (Fig. 1E, Fig. S1O). This trend persisted even when controlling for the increased statistical power for analysis of ALE and tandem UTR isoforms that results from the relatively large sizes of 3′ UTRs (Fig. S1M, N). The individual ALE pairs that were localized were highly similar between CAD and N2A cells ($R_{Spearman} = 0.74$) and moderately similar between CAD and primary neurons ($R_{Spearman} = 0.35$) (Fig. S1K, L).

The biases toward ribosomal and mitochondrial functions observed when analyzing gene-level LRs were not observed among genes containing neurite localized distal ALEs, perhaps because ribosomal and mitochondrial genes rarely contain ALEs. Inspection of the list of

genes with neurite distal ALEs revealed several genes encoding neurotransmitter receptors, ion channels and trafficking proteins but no strong gene ontology biases, indicating that this set of genes is quite diverse.

Previously, localization determinants have been identified in both UTRs and coding regions, but more commonly in 3′ UTRs (Andreassi and Riccio, 2009). Since ALEs and tandem UTRs predominantly alter 3′ UTRs, this analysis provides evidence for a predominant and general role of 3′ UTRs in determining message localization. This observation may help to explain the widespread presence and conservation of alternative 3′ UTR isoforms, despite recent evidence that these isoforms rarely impact mRNA stability or translation (Spies et al., 2013). For this reason, we chose to focus here on the role of 3′ UTRs in neurite localization.

In two previous examples of differential localization of alternative 3′ UTR isoforms, the longer tandem UTR isoform was localized to neurites (An et al., 2008; Harrison et al., 2014). Here, we observed similar numbers of tandem UTR pairs having the shorter, proximal PAS isoform localized to neurites as of pairs having the longer isoform localized (Fig. 1F). However, when analyzing ALEs in CAD cells, we observed a dramatic bias: in 80% of pairs with significant differential localization, the distal ALE isoform was localized to neural projections ($p < 2.2 \times 10^{-16}$, chi-square test, Fig. 1F). A trend of similar magnitude in the same direction was observed in N2A cells and in primary cortical neurons (Fig. 1F, Fig. S2A–C), and also in mouse DRG (Minis et al., 2013), indicating that this phenomenon occurs in peripheral neurons as well (Fig. S2D). Thus, isoform-level analysis of four different neuronal localization systems revealed an unexpected connection between the subcellular localization of ALE isoforms and the relative genomic position (distal versus proximal) of the alternative exons.

Alternative 3′ UTRs confer neurite or soma localization

We hypothesized that the 3′ UTR portion of differentially localized ALE isoforms confers mRNA localization. To test this hypothesis, we fused UTRs from differentially localized distal and proximal ALEs to reporter genes and expressed them in CAD cells. RNA localization was monitored by RNA fluorescence in situ hybridization (RNA FISH), using reporters that encoded a fluorescent protein, providing a control for transfection and expression efficiency (Fig. S2E). For all three of the tested reporters we observed robust localization of the distal ALE UTR reporter to projections, with much less localized RNA detected for the corresponding proximal ALE reporters (Fig. 2A, B, Fig. S2F). Similarly, when analyzing the relative abundance of these constructs between soma and neurite fractions using qRT-PCR, we found that RNA from the distal ALE construct was consistently enriched in neurites relative to that from the proximal ALE-containing construct (Fig. 2C). Therefore we conclude that the 3′ UTR portion of distal ALE isoforms is often sufficient to confer localization of mRNA to projections.

We also considered whether differential stability of isoforms in different cellular compartments might explain observed differences in mRNA abundance between compartments. We monitored changes in mRNA abundance following either inhibition of transcription by treatment with actinomycin D (Fig. S2G) or physical separation of neurite from soma (Fig. S2H), but did not observe differences in decay rates that could explain

differential abundance. Instead, differential abundance presumably results from active trafficking of mRNAs to projections or anchoring of messages by projection-specific proteins.

Localized distal ALEs possess distinctive properties

To help understand the determinants of localization, we examined properties of localized distal ALE UTRs as a class. We identified 421 distal ALE isoforms that were preferentially localized to neurites, with criteria including $\Delta\psi \geq 0.1$ in both CAD and N2A cells (Table S2). The median ratio of $\psi_{neurite} / \psi_{soma}$ for this set was 1.3, with a range from 1.1 to more than 16. These isoforms form a set of "neurite distal" UTRs, with the corresponding proximal ALE isoforms of these genes forming a set of "soma proximal" UTRs. Nonlocalized distal and proximal UTRs were defined from ALE isoform pairs that did not differ significantly in localization by the above criterion. We similarly defined classes of ALEs in cortical neurons and DRG using $\Delta\psi$ values from the respective cell types.

We considered a variety of mRNA features that might impact localization. Neurite distal UTRs identified in cultured and primary cortical neurons had median sizes of 461 and 429 bases, respectively, substantially shorter than the other classes of UTRs (Fig. 3A left, Fig. S3A), but the opposite trend held in DRG (Fig. S3B), suggesting that there is not a simple relationship between 3′ UTR length and localization. Some known localization elements involve RNA secondary structure (Martin and Ephrussi, 2009), and we observed that neurite distal UTRs contained more conserved secondary structure on average, based on folding of homologous sets of UTRs from mouse, human, rat, dog, and cow using the RNAalifold algorithm (Bernhart et al., 2008) (Fig. 3A middle, S3C). *Cis*-acting regulatory elements involved in mRNA localization are likely to be conserved. Localized distal UTRs from the cell lines had higher average conservation, based on PhastCons score (Siepel and Haussler, 2005), throughout their length (Fig. 3A right), consistent with increased abundance of conserved regulatory elements in these UTRs. We also considered whether the protein-coding capacity of ALEs might contribute to mRNA localization. We observed no difference in the abundance of mitochondrial and secretory pathway targeting peptides (Emanuelsson et al., 2007) among the four classes of isoforms defined above (Fig. S3D), providing no evidence that these peptide motifs contribute to neurite localization. This observation is consistent with our reporter assays showing that the 3′ UTR is often sufficient to confer localization.

We next sought to understand the interaction between gene and isoform expression and neuronal differentiation, since the requirement for localization of mRNAs is expected to increase as neurites grow during neuronal differentiation. By RNA-seq analysis of CAD cells before and after inducing differentiation by withdrawal of serum, we observed increased expression of distal ALE isoforms. Dividing ALE isoforms based on their localization properties as in Figure 3A, we observed preferential expression of neurite distal ALE isoforms upon differentiation, with no trend observed for nonlocalized distal ALEs (Fig. 3B). The average expression level of genes containing neurite distal ALEs did not change (Fig. S3E). Thus, our data support a model in which preferential expression of localized mRNA isoforms during differentiation results primarily from shifts in relative

isoform abundance (changes in PSI, mediated by post-transcriptional mechanisms) rather than gene expression changes.

We also examined RNA-seq data from the differentiation of human neural precursor cells (NPCs) to neurons (Sauvageau et al., 2013). We observed a significant trend for expression of distal ALE isoforms, particularly early in the time course (Fig. S3F, G), suggesting that preferential expression of distal ALE isoforms during neuronal differentiation is conserved across species.

Muscleblind proteins promote localization of mRNAs to neurites

To determine candidate RNA-binding factors involved in localization of mRNA isoforms, we searched for sequence motifs that were both enriched in the UTRs of neurite distal ALEs compared to soma proximal ALEs and conserved between mouse and human (Fig. 4A). We observed strong enrichment (~1.4-fold to > 2-fold) and strong sequence conservation of several 6mers matching consensus binding motifs of the muscleblind-like (Mbnl) family of RNA binding proteins (RBPs), and of a few other 6mers. Furthermore, motifs containing the Mbnl family motif YGCU (Y = C or U) were enriched in neurite distal UTRs from both cell lines, from primary cortical neurons, and from primary DRG cells (Fig. 4B, Fig. S4A, B). Neurite distal UTRs identified in the cell lines and primary DRG were also enriched for *in vivo* MBNL1 binding sites as identified in mouse brain by crosslinking/immunoprecipitation-sequencing (CLIP-seq) (Fig. 4C, Fig. S4C) (Wang et al., 2012). The mouse genome expresses three Mbnl genes, of which *Mbnl1* and *Mbnl2* are expressed in neurons, as well as other cell and tissue types (Charizanis et al., 2012; Suenaga et al., 2012). Mbnl family proteins are well established as regulators of alternative splicing and have also been implicated in the localization of integrin alpha3 mRNA to adhesion plaques in cancer cell lines (Adereth et al., 2005), and in localizing hundreds of mRNAs to membrane locations in mouse myoblasts (Wang et al., 2012).

Because of the strong enrichment of associated motifs and binding sites, and previous reports implicating Mbnl proteins in mRNA localization, we hypothesized that Mbnl proteins play a major role in localization of mRNAs to neural projections. To test this hypothesis, we depleted *Mbnl1* and *Mbnl2* simultaneously in CAD cells by RNAi, and we also dissected cortical neurons from E18.5 *Mbnl1* knockout (KO) and *Mbnl2* KO mouse embryos. Expression of *Mbnl1* and *Mbnl2* were reduced by approximately 70% in CAD cells treated with siRNAs relative to controls (Fig. S4D). CAD cells or primary cortical neurons were fractionated into soma and projection as before, and both fractions were subjected to RNA-seq, from which the LR values of genes were measured. To assess the change in localization we defined the “Difference in log Localization Ratios”, DLR = $\log_2(LR_{kd}) - \log_2(LR_{control})$. Thus, positive values of DLR indicate increased neurite localization following knockdown and negative values indicate decreased neurite localization. Comparing DLR values we found that a large subset of mRNAs that were neurite localized in control cells became less neurite localized following Mbnl knockdown (Fig. 4D),, suggesting that Mbnl proteins contribute to neurite localization of many genes. By comparison, the DLR values of nonlocalized genes were centered around zero, indicating no systematic change in localization (Fig. 4D). Similarly, LR values for localized genes

were significantly decreased in cortical neurons from *Mbnl1* and *Mbnl2* KO embryos, while nonlocalized genes were not systematically affected (Fig. S4E, F).

Based on the RNA yields following fractionation of the two cell lines, we estimate that approximately 99% of the total RNA within these cells is contained within the soma fraction while ~1% is contained within projections (Fig. S4G). The loss of neurite localization of a transcript should therefore cause a large reduction in neurite expression, but only a small increase in soma expression relative to the larger pool of somal transcripts. Consistent with this expectation, genes whose LR decreased following *Mbnl1/Mbnl2* knockdown had substantially reduced expression in projections (Fig. S4I), but only slight increases in somal expression compared to nonlocalized genes (Fig. S4H). These data are consistent with Mbnl depletion exerting a primary affect on mRNA levels in projections rather than affecting overall mRNA expression levels.

We also analyzed the relationship between presence of Mbnl motifs and change in RNA localization following Mbnl depletion. To assess Mbnl-dependent changes in isoform localization in WT versus Mbnl-depleted cells we use $\Delta\Delta\psi$, defined as an isoform pair's $\Delta\psi$ value in Mbnl-depleted cells minus its $\Delta\psi$ in WT cells. An isoform pair where the distal isoform becomes less localized following Mbnl depletion will exhibit a reduced $\Delta\psi$ in Mbnl-depleted versus WT cells, and therefore have $\Delta\Delta\psi < 0$. Thus, $\Delta\Delta\psi$ represents the change in differential localization following Mbnl depletion. For example, consider the ALE pair in the gene *Hsd17b4* in WT and *Mbnl1*$^{-/-}$;*Mbnl2*$^{-/-}$ double knockout (DKO) primary neurons. The PSI values of this ALE pair in WT neurite and soma were 0.87 and 0.24, respectively, yielding a $\Delta\psi$ of 0.63, indicating preferential enrichment in neurites. Its PSI values in DKO neurite and soma were 0.29 and 0.21, respectively, yielding a $\Delta\psi$ of 0.08. The strength of neurite enrichment of the distal isoform has therefore decreased dramatically, as reflected in its $\Delta\Delta\psi$ value of $0.08 - 0.63 = -0.55$. The majority of ALE isoforms in CAD cells had negative $\Delta\Delta\psi$ values upon Mbnl depletion, consistent with a role for Mbnls in promoting neurite localization of distal ALE isoforms (Fig. S4K). Furthermore, we observed a significant correlation of ALE $\Delta\Delta\psi$ values between *Mbnl1* KO and *Mbnl2* KO cells, indicating that MBNL1 and MBNL2 may influence the localization of overlapping sets of mRNAs (and may partially compensate for each other's absence) (Goodwin et al., 2015) (Fig. S4L).

Defining "Mbnl-sensitive" ALEs as those with a $\Delta\Delta\psi$ value at least one standard deviation below the mean, we observed that the expression of genes containing Mbnl-sensitive ALE isoforms was unchanged following Mbnl depletion (Fig. S4M), but that distal ALE isoforms had large decreases in expression in projections and small increases in the soma, paralleling the changes in LR and soma and projection expression at the gene level (Fig. S4N). Distal ALE isoforms that increased in inclusion during differentiation of human NPCs to neurons (Fig. S4K) were also enriched for Mbnl motifs in their 3′ UTRs (Fig. S4J), suggesting that Mbnl proteins may play a role in RNA localization in human neurons. Although Mbnl proteins were recently observed to impact cleavage and polyadenylation (CPA) (Batra et al., 2014), this activity would not affect the $\Delta\psi$ defined here, which reflects differences in localization of isoforms rather than absolute isoform abundance. Furthermore, although

Mbnl is a known splicing factor, we did not observe any splicing changes in known localization factors.

Under our hypothesis that Mbnl proteins localize specific mRNAs to neurites, there should be a relationship between the number of Mbnl binding motifs present in a transcript's 3′ UTR and changes in localization following Mbnl depletion. Because RNAi yields only a partial reduction in *Mbnl1*/*Mbnl2* levels and *Mbnl1* KO mice still express *Mbnl2* (and vice versa), these systems achieved ~50–70% reduction in total levels of MBNL1 + MBNL2 together. Therefore, we expected that transcripts with modest numbers of MBNL sites and weaker binding would be more susceptible to mis-localization after Mbnl depletion than those with many sites, which might more effectively compete for the reduced pool of Mbnl proteins. UTRs with low Mbnl motif densities had the greatest decrease in localization following *Mbnl1/Mbnl2* knockdown or KO, more so than UTRs lacking Mbnl motifs or UTRs with higher Mbnl motif densities, which were less sensitive to depletion (Figs. S4O–V). Moreover, this relationship between motif density and mis-localization was true only for Mbnl motifs and not for motifs of 20 other RBPs used as controls (Fig. S4T–V, gray lines). When Mbnls were completely depleted from cells through the use of DKO cells, the UTRs that were most sensitive to Mbnl depletion were most enriched for Mbnl motifs (Fig. 4E). Together, these observations support a direct role for Mbnl proteins in promoting mRNA localization to neurites.

ALE and tandem UTR isoforms are coordinately regulated in diverse cellular contexts

Independent of the precise mechanism by which Mbnl proteins direct mRNA localization, which remains to be worked out, the larger puzzle presented by this study is why there should be a relationship between the relative genomic position of ALEs and the subcellular localization of the resulting mRNA isoform. A mechanistic link seems unlikely, since the localization of mRNAs to neurites presumably occurs after nuclear RNA processing and/or export. Instead, we hypothesize that this relationship reflects a regulatory strategy, in which differentiating neurons alter the RNA processing machinery to preferentially produce distal ALE isoforms in order to coordinately induce expression of many neurite-localized mRNAs. This hypothesis makes specific predictions, including: 1) that distal ALE isoforms are systematically induced during cellular differentiation/development generally (as seen in Fig. 3B for neurite distal ALE isoforms in CAD cells); and 2) that there are post-transcriptional regulatory programs or factors that preferentially promote expression of distal versus proximal ALE isoforms (or of proximal versus distal), enabling coordinated regulation in various contexts.

To test the first of these predictions, we analyzed RNA-seq data from five available developmental, differentiation, or reprogramming systems (Fig. 5A). We observed strong biases toward expression of distal ALE and distal tandem UTR isoforms generally during neuronal differentiation of human NPCs and of mouse CAD cells *in vitro* (Fig. 5A). We also observed similar trends during cardiac differentiation of mESCs *in vitro*, and a strong bias toward distal ALEs but not tandem UTRs during mouse cardiac ventricle development *in vivo*. Previously, we have shown that Mbnl proteins contribute to mRNA localization in myoblasts (Wang et al., 2012). In the other direction, we observed a bias toward proximal

ALE isoform expression following the reprogramming of human fibroblasts to iPSCs (Fig. 5A) (Gallego Romero et al., 2014). Thus our first prediction, that distal ALE expression should generally increase during cellular differentiation, was borne out.

A trend toward expression of distal tandem UTR isoforms has been observed previously in mouse myoblast differentiation and during mouse embryonic development (Ji et al., 2009), and a reverse trend toward expression of proximal tandem UTRs has been observed in association with cellular proliferation, oncogenic transformation and reprogramming of iPSCs (Ji et al., 2009; Mayr and Bartel, 2009; Sandberg et al., 2008). However, general trends in ALE isoform expression have not been as well studied. To assess whether ALE isoform expression changes are altered during the general de-differentiation that occurs in cancer, we used available RNA-seq data to examine changes in isoform expression in comparisons of liver cancer to matched normal liver controls, and also in comparisons of lung cancer to matched normal lung. These comparisons predominantly showed a trend toward increased expression of proximal ALE and tandem UTR isoforms in tumors relative to controls, in the great majority of tumors of both types (Fig. 5B). Together, these observations about isoform abundance suggest the existence of a general association between differentiation and distal ALEs that is reversed in cancer and other cases of de-differentiation.

Previous studies have suggested that a weakening of the activity of intrinsic CPA machinery may underlie the shift toward distal tandem UTR expression in differentiation (Ji et al., 2009) as might be expected if CPA is controlled by kinetic competition between PASs. Furthermore, induction of the expression of core CPA factor *Cstf2* (aka *Cstf-64*) promotes expression of the proximal ALE isoform of IgM in a B cell line (Takagaki and Manley, 1998), and recent studies indicate that reduction in levels of U1 snRNP promotes proximal PAS usage (Berg et al., 2012). More generally, ALE isoform regulation could involve various post-transcriptional mechanisms, since ALE choice entails use of different 3′ splice sites and different PAS (Di Giammartino et al., 2011).

To test the second prediction of our hypothesis – that there are factors that preferentially promote (or inhibit) expression of distal versus proximal ALE isoforms in bulk – we analyzed changes in ALE and tandem UTR expression using RNA-seq data following RNAi of dozens of RBPs, including a number of splicing and CPA factors, which was conducted in human K562 erythroleukemia cells as part of an ENCODE Phase 3 project, and related data. This analysis identified candidate factors whose activity could contribute to systematic shifts toward proximal or distal PAS in mammalian cells. The strongest effect on tandem UTRs was observed for depletion of CPA factor *CFIM25*, whose depletion resulted in a predominant shift toward proximal PAS isoforms (as observed following knockdown in glioblastoma cells (Masamha et al., 2014)), and also toward proximal ALE isoforms. In the other direction, depletion of CPA factor *CSTF2T* (a paralog of *CSTF2* (Di Giammartino et al., 2011)) resulted in a shift toward distal PAS isoforms for both tandem UTRs and ALEs, and knockdowns of certain other RBPs also triggered systematic shifts in one direction or the other. Therefore, our prediction that there are factors that preferentially promote distal or proximal PAS isoforms of ALEs and tandem UTRs in mammalian cells was confirmed,

supporting our hypothesis that the relative genomic position of ALEs that direct localization enables their coordinate regulation.

Comparing across all of the RBPs, we observed a high correlation between the effects on distal ALEs and the effects on distal tandem UTRs ($r = 0.54$, $P = 8.63 \times 10^{-6}$), providing evidence of co-regulation of these two classes of isoforms (Fig. 5C) (Li et al., 2015). Consistent with this idea, the proportion of genes with increased distal versus proximal ALE expression was strongly positively correlated with the corresponding proportion for tandem UTRs across the samples analyzed in Fig. 5A, B ($r = 0.74$), a tighter correlation than was observed for comparisons of other types of alternative isoforms (Fig. 5D), suggesting co-regulation of these two classes of 3′ UTR isoforms.

DISCUSSION

Several previous studies have analyzed RNA localization at the gene level, without regard for individual isoforms (Cajigas et al., 2012; Gumy et al., 2011; Minis et al., 2013; Taylor et al., 2009). Here, we have assayed RNA localization at the isoform level, identifying hundreds of alternative 3′ UTRs associated with mRNA localization to neurites. This approach enabled us to hone in on relevant transcript regions and to identify motifs associated with localization, revealing that distal 3′ UTR isoforms are preferentially neurite localized. We also identified Mbnl proteins, which are central to pathology of myotonic dystrophy (DM) (Lee and Cooper, 2009), as regulators of RNA localization in neurons. Inhibiting Mbnls alters localization of mRNAs encoding proteins of neurological importance (Table S3), raising the possibility that localization defects contribute to the various neurological symptoms observed in DM.

Functions of 3′ UTRs

States of higher cell proliferation and oncogenic transformation are associated with increased expression of transcripts from upstream tandem PAS (Mayr and Bartel, 2009; Sandberg et al., 2008). Conversely, a trend toward higher expression of transcripts from distal tandem polyadenylation sites has been observed during cellular differentiation and development in a number of systems (Ji et al., 2009; Miura et al., 2013). However, the functional consequences of these shifts in 3′ UTR isoforms have remained largely unclear, and a recent global assessment found that most alternative 3′ UTRs have little or no effect on either translation efficiency or mRNA stability (Spies et al., 2013).

Regulation and function of alternative 3′ UTRs is likely to impact a variety of cell types and states, including disease states. For example, depletion of a specific factor, CfIm25, leads to a pronounced shift toward proximal 3′ UTR isoform expression and an increase in cell proliferation and tumorigenicity in glioblastoma cells (Masamha et al., 2014). Our data suggest that the differences between alternative 3′ UTR isoforms more often involve altered mRNA localization. Beyond neuronal differentiation, mRNA localization is important in diverse mammalian cell types (Mili et al., 2008; Wang et al., 2012). In some cases, alternative 3′ UTRs may impact protein localization independent of mRNA localization (Berkovits and Mayr, 2015).

Genomic organization of localization-inducing sequences

It has been recognized that the 5′ and 3′ ends of genes are often variable (Davuluri et al., 2008; Proudfoot, 2011). In particular, many mammalian gene families such as protocadherins, cytochrome p450s, and various receptor families express two or more alternative first exons (AFEs) from alternative promoters (Wu and Maniatis, 1999). In general, the literature has supported regulation of individual alternative promoters by specific transcription factors rather than coordinated directional shifts toward proximal or distal promoter use. Nor have general trends in the functions of proximal versus distal AFEs been observed. This situation contrasts with our findings on ALEs, where we observe wholesale shifts toward distal ALE isoform expression during differentiation, as well as a pattern in which distal ALE isoforms are preferentially neurite-localized. These trends suggest the existence of mechanisms or factors that shift ALE choice in a coordinated, directional fashion during differentiation or disease processes (Fig. 5E). Regulating ALE isoforms in bulk, e.g. via changes in cleavage and polyadenylation activity, may allow cellular control of the localization properties of hundreds or thousands of transcripts by altering the activity of a few post-transcriptional regulatory factors.

EXPERIMENTAL PROCEDURES

Cell culture and fractionation

N2A cells were grown in standard DMEM (Gibco) supplemented with 10% FBS. CAD cells were grown in DMEM/F12 (Gibco) supplemented with 10% FBS. Primary cortical neurons were grown in Neurobasal medium (Gibco) with B-27 supplements (Gibco). To fractionate, polyethylene terephthalate membranes with 1 μm pores (Millipore PIRP30R48) were treated on their underside with 0.2% matrigel in DMEM for 30 min at 37° C. 4 mL of media were then placed in each well of a 6 well plate and the membranes were placed in the plate. 2 mL of confluent cells (~1×10^6 cells) were then plated on the top of the membrane and allowed to attach for 1 h. For N2A and CAD cells, the media below the membrane and on top of the cells was then replaced with media lacking serum. The cells were incubated at 37° C for 24 h (or 48 h for primary neurons) until fractionation.

The media was then removed and both sides of the membrane were rinsed with PBS. 1 mL of PBS was placed on the top of the membrane. Cell bodies were scraped in the PBS from the top of the membrane using a cell scraper. The membrane, still containing projections, was then cut out of its plastic housing and incubated with RLT lysis buffer (Qiagen) at 4°C for 15 min. Six membranes in a six well plate were combined and used as a single prep. RNA was then purified from both fractions using a Qiagen RNeasy Micro Kit. Typically, between 500 and 1000 ng of total RNA was collected from projection fractions in a single prep.

RNA-seq

For the N2A and CAD cell fractionation, strand-specific, poly-A selected libraries were constructed using the dUTP incorporation method and sequenced on an Illumina HiSeq sequencer with paired end 60 bp reads. Each sample was fractionated, prepared, and sequenced in triplicate, yielding approximately 35–50 million read pairs per replicate. For

library preparation and sequencing methods for the primary neuron samples, see Supplemental Experimental Procedures. The ENCODE shRNA knockdown RNA-seq experiment are available at www.encodeproject.org under accession number ENCSR089EOA

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Myriam Heiman, Phillip Sharp and members of the Burge lab for helpful comments on this manuscript, and Athma Pai for assistance with preparation of supplemental materials. This work was supported by an NIH Fellowship (J. M. T.), by an NSF equipment grant (no. 0821391), ENCODE project funding (U54HG007005) (B. G. and C. B. B.), the Ludwig Center at MIT (F. G.), and by grants from the NIH (B. G., F. G., M. S. and C. B. B.).

References

Adereth Y, Dammai V, Kose N, Li R, Hsu T. RNA-dependent integrin α3 protein localization regulated by the Muscleblind-like protein MLP1. Nat Cell Biol. 2005; 7:1240–1247. [PubMed: 16273094]

An JJ, Gharami K, Liao GY, Woo NH, Lau AG, Vanevski F, Torre ER, Jones KR, Feng Y, Lu B, et al. Distinct role of long 3′ UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. Cell. 2008; 134:175–187. [PubMed: 18614020]

Andreassi C, Riccio A. To localize or not to localize: mRNA fate is in 3′UTR ends. Trends Cell Biol. 2009; 19:465–474. [PubMed: 19716303]

Batra R, Charizanis K, Manchanda M, Mohan A, Li M, Finn DJ, Goodwin M, Zhang C, Sobczak K, Thornton CA, et al. Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. Mol Cell. 2014; 56:311–322. [PubMed: 25263597]

Berg MG, Singh LN, Younis I, Liu Q, Pinto AM, Kaida D, Zhang Z, Cho S, Sherrill-Mix S, Wan L, et al. U1 snRNP determines mRNA length and regulates isoform expression. Cell. 2012; 150:53–64. [PubMed: 22770214]

Berkovits BD, Mayr C. Alternative 3′ UTRs act as scaffolds to regulate membrane protein localization. Nature. 2015; 522:363–367. [PubMed: 25896326]

Bernhart SH, Hofacker IL, Will S, Gruber AR, Stadler PF. RNAalifold: improved consensus structure prediction for RNA alignments. BMC Bioinformatics. 2008; 9:474. [PubMed: 19014431]

Buckley PT, Lee MT, Sul JY, Miyashiro KY, Bell TJ, Fisher SA, Kim J, Eberwine J. Cytoplasmic intron sequence-retaining transcripts can be dendritically targeted via ID element retrotransposons. Neuron. 2011; 69:877–884. [PubMed: 21382548]

Buxbaum AR, Wu B, Singer RH. Single -Actin mRNA Detection in Neurons Reveals a Mechanism for Regulating Its Translatability. Science. 2014; 343:419–422. [PubMed: 24458642]

Cajigas IJ, Tushev G, Will TJ, tom Dieck S, Fuerst N, Schuman EM. The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. Neuron. 2012; 74:453–466. [PubMed: 22578497]

Charizanis K, Lee KY, Batra R, Goodwin M, Zhang C, Yuan Y, Shiue L, Cline M, Scotti MM, Xia G, et al. Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. Neuron. 2012; 75:437–450. [PubMed: 22884328]

Davuluri RV, Suzuki Y, Sugano S, Plass C. The functional consequences of alternative promoter use in mammalian genomes. Trends in Genetics. 2008

Di Giammartino DC, Nishida K, Manley JL. Mechanisms and consequences of alternative polyadenylation. Mol Cell. 2011; 43:853–866. [PubMed: 21925375]

Emanuelsson O, Brunak S, von Heijne G, Nielsen H. Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc. 2007; 2:953–971. [PubMed: 17446895]

Ephrussi A, Dickinson LK, Lehmann R. Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. Cell. 1991; 66:37–50. [PubMed: 2070417]

Gallego Romero I, Pavlovic BJ, Hernando-Herraez I, Banovich NE, Kagan CL, Burnett JE, Huang CH, Mitrano A, Chavarria CI, Ben-Nun IF, et al. Generation of a Panel of Induced Pluripotent Stem Cells From Chimpanzees: a Resource for Comparative Functional Genomics (bioarxiv). 2014

Ghosh S, Marchand V, Gáspár I, Ephrussi A. Control of RNP motility and localization by a splicing-dependent structure in oskar mRNA. Nature Publishing Group. 2012; 19:441–449.

Goodwin M, Mohan A, Batra R, Lee KY, Charizanis K, Gómez FJF, Eddarkaoui S, Sergeant N, Buée L, Kimura T, et al. MBNL Sequestration by Toxic RNAs and RNA Misprocessing in the Myotonic Dystrophy Brain. CellReports. 2015; 12:1159–1168.

Gumy LF, Yeo GSH, Tung YCL, Zivraj KH, Willis D, Coppola G, Lam BYH, Twiss JL, Holt CE, Fawcett JW. Transcriptome analysis of embryonic and adult sensory axons reveals changes in mRNA repertoire localization. Rna. 2011; 17:85–98. [PubMed: 21098654]

Harrison BJ, Flight RM, Gomes C, Venkat G, Ellis SR, Sankar U, Twiss JL, Rouchka EC, Petruska JC. IB4-binding sensory neurons in the adult rat express a novel 3′ UTR-extended isoform of CaMK4 that is associated with its localization to axons. J Comp Neurol. 2014; 522:308–336. [PubMed: 23817991]

Ji Z, Lee JY, Pan Z, Jiang B, Tian B. Progressive lengthening of 3′ untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proc Natl Acad Sci USA. 2009; 106:7028–7033. [PubMed: 19372383]

Lee JE, Cooper TA. Pathogenic mechanisms of myotonic dystrophy. Biochem Soc Trans. 2009; 37:1281–1286. [PubMed: 19909263]

Leung KM, van Horck FPG, Lin AC, Allison R, Standart N, Holt CE. Asymmetrical beta-actin mRNA translation in growth cones mediates attractive turning to netrin-1. Nat Neurosci. 2006; 9:1247–1256. [PubMed: 16980963]

Lécuyer E, Yoshida H, Parthasarathy N, Alm C, Babak T, Cerovina T, Hughes TR, Tomancak P, Krause HM. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. Cell. 2007; 131:174–187. [PubMed: 17923096]

Li W, You B, Hoque M, Zheng D, Luo W, Ji Z, Park JY, Gunderson SI, Kalsotra A, Manley JL, et al. Systematic profiling of poly(a)+ transcripts modulated by core 3′ end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. PLoS Genet. 2015; 11:e1005166. [PubMed: 25906188]

Martin KC, Ephrussi A. mRNA localization: gene expression in the spatial dimension. Cell. 2009; 136:719–730. [PubMed: 19239891]

Masamha CP, Xia Z, Yang J, Albrecht TR, Li M, Shyu AB, Li W, Wagner EJ. CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. Nature. 2014; 510:412–416. [PubMed: 24814343]

Mayr C, Bartel DP. Widespread shortening of 3′UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell. 2009; 138:673–684. [PubMed: 19703394]

Merkin J, Russell C, Chen P, Burge CB. Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. Science. 2012; 338:1593–1599. [PubMed: 23258891]

Mili S, Moissoglu K, Macara IG. Genome-wide screen reveals APC-associated RNAs enriched in cell protrusions. Nature. 2008; 453:115–119. [PubMed: 18451862]

Minis A, Dahary D, Manor O, Leshkowitz D, Pilpel Y, Yaron A. Subcellular transcriptomics-Dissection of the mRNA composition in the axonal compartment of sensory neurons. Devel Neurobio. 2013 n/a–n/a.

Miura P, Shenker S, Andreu-Agullo C, Westholm JO, Lai EC. Widespread and extensive lengthening of 3′ UTRs in the mammalian brain. Genome Res. 2013; 23:812–825. [PubMed: 23520388]

Moccia R, Chen D, Lyles V, Kapuya E, EY, Kalachikov S, Spahn CMT, Frank J, Kandel ER, Barad M, et al. An Unbiased cDNA Library Prepared from Isolated Aplysia Sensory Neuron Processes Is Enriched for Cytoskeletal and Translational mRNAs. The Journal of …. 2003

Paushkin S, Gubitz AK, Massenet S, Dreyfuss G. The SMN complex, an assemblyosome of ribonucleoproteins. Current Opinion in Cell Biology. 2002; 14:305–312. [PubMed: 12067652]

Poon MM, Choi SH, Jamieson CAM, Geschwind DH, Martin KC. Identification of process-localized mRNAs from cultured rodent hippocampal neurons. J Neurosci. 2006; 26:13390–13399. [PubMed: 17182790]

Proudfoot NJ. Ending the message: poly(A) signals then and now. Genes Dev. 2011

Ramsköld D, Wang ET, Burge CB, Sandberg R. An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. PLoS Comp Biol. 2009; 5:e1000598.

Ross AF, Oleynikov Y, Kislauskis EH, Taneja KL, Singer RH. Characterization of a beta-actin mRNA zipcode-binding protein. Mol Cell Biol. 1997; 17:2158–2165. [PubMed: 9121465]

Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB. Proliferating cells express mRNAs with shortened 3′ untranslated regions and fewer microRNA target sites. Science. 2008; 320:1643–1647. [PubMed: 18566288]

Sauvageau M, Goff LA, Lodato S, Bonev B, Groff AF, Gerhardinger C, Sanchez-Gomez DB, Hacisuleyman E, Li E, Spence M, et al. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. eLife. 2013; 2:e01749. [PubMed: 24381249]

Siepel A, Haussler D. Phylogenetic hidden Markov models. Statistical Methods in Molecular Evolution. 2005

Spies N, Burge CB, Bartel DP. 3′ UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. Genome Res. 2013; 23:2078–2090. [PubMed: 24072873]

Suenaga K, Lee KY, Nakamori M, Tatsumi Y, Takahashi MP, Fujimura H, Jinnai K, Yoshikawa H, Du H, Ares M, et al. Muscleblind-like 1 knockout mice reveal novel splicing defects in the myotonic dystrophy brain. PLoS ONE. 2012; 7:e33218. [PubMed: 22427994]

Takagaki Y, Manley JL. Levels of polyadenylation factor CstF-64 control IgM heavy chain mRNA accumulation and other events associated with B cell differentiation. Mol Cell. 1998; 2:761–771. [PubMed: 9885564]

Taylor AM, Berchtold NC, Perreau VM, Tu CH, Jeon NL, Cotman CW. Axonal mRNA in uninjured and regenerating cortical mammalian axons. J Neurosci. 2009; 29:4697–4707. [PubMed: 19369540]

Tolino M, Köhrmann M, Kiebler MA. RNA-binding proteins involved in RNA localization and their implications in neuronal diseases. European Journal of Neuroscience. 2012; 35:1818–1836. [PubMed: 22708593]

Ulitsky I, Shkumatava A, Jan CH, Subtelny AO, Koppstein D, Bell GW, Sive H, Bartel DP. Extensive alternative polyadenylation during zebrafish development. Genome Res. 2012; 22:2054–2066. [PubMed: 22722342]

Wang ET, Cody NAL, Jog S, Biancolella M, Wang TT, Treacy DJ, Luo S, Schroth GP, Housman DE, Reddy S, et al. Transcriptome-wide Regulation of Pre-mRNA Splicing and mRNA Localization by Muscleblind Proteins. Cell. 2012; 150:710–724. [PubMed: 22901804]

Whittaker KL, Ding D, Fisher WW, Lipshitz HD. Different 3′ untranslated regions target alternatively processed hu-li tai shao (hts) transcripts to distinct cytoplasmic locations during Drosophila oogenesis. Journal of Cell …. 1999

Wu Q, Maniatis T. A Striking Organization of a Large Family of Human Neural Cadherin-like Cell Adhesion Genes. Cell. 1999; 97:779–790. [PubMed: 10380929]

Yang E, van Nimwegen E, Zavolan M, Rajewsky N, Schroeder M, Magnasco M, Darnell JE. Decay rates of human mRNAs: correlation with functional characteristics and sequence attributes. Genome Res. 2003; 13:1863–1872. [PubMed: 12902380]

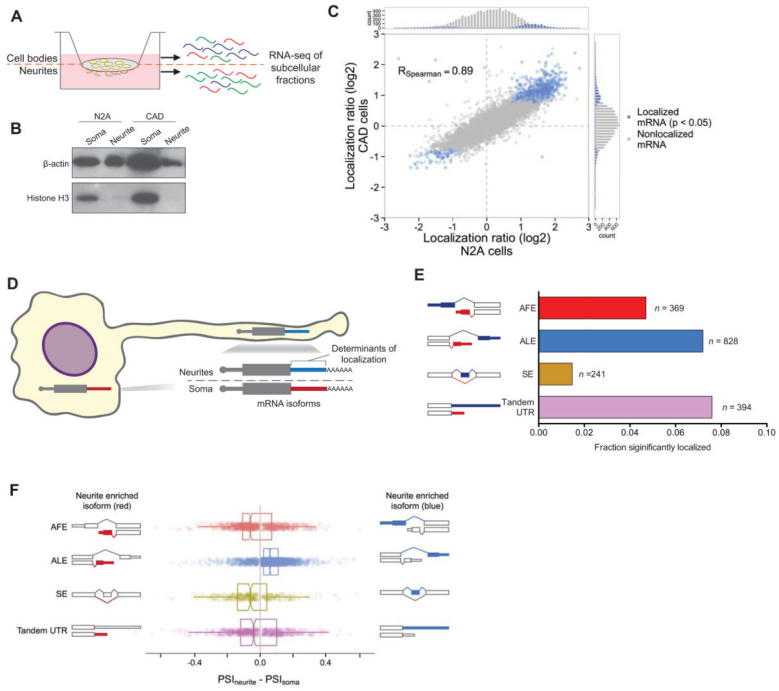

Figure 1. Cellular fractionation and sequencing reveals mRNA isoforms associated with neurite localization

A) Cells are grown on top of porous membranes, allowing growth of neurites through the pores, enabling fractionation. B) Soma and neurite lysates from primary cortical neurons were immunoblotted for beta-actin, a marker of both soma and neurite, and histone H3, a marker of soma. C) LRs in two cell lines. Differentially enriched genes in both cell lines are shown in blue. D) Schematic showing differential isoform enrichment. E) The fraction of the expressed alternative isoform pairs that were significantly differentially enriched between soma and neurite fractions for different classes of alternative isoforms. At left, the inclusion isoform is pictured in blue, and the exclusion isoform is pictured in red. F) Distribution of $\Delta\psi$ values of different isoform classes. Boxes indicate 25th and 75th percentiles, lines indicate 5^{th} and 95^{th} percentiles. See also Figure S1, Table S1.

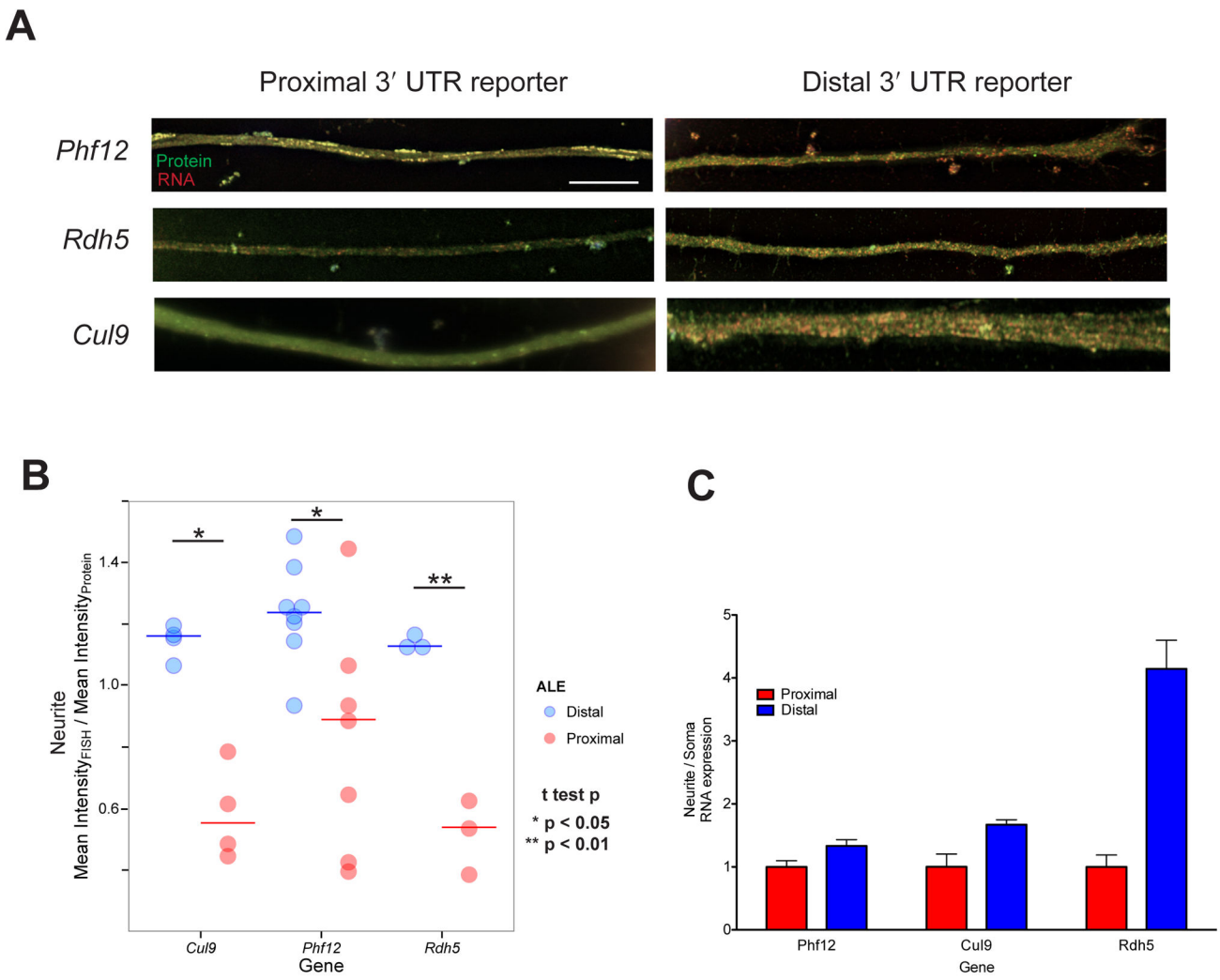

Figure 2. 3′ UTRs of neurite distal ALE isoforms confer neurite localization

A) The subcellular localization of RNA from a reporter gene (Fig. S2E) containing the proximal (left) or distal (right) alternative last exon from the indicated gene was monitored using RNA FISH. The fluorescent protein product of the reporter is colored in green while probes against the RNA are shown in red. B) Quantification of FISH results. Values are the mean intensity across the projection in the red channel divided by the mean intensity in the green channel. * $p < 0.05$; ** $p < 0.01$. C) qRT-PCR analysis of neurite versus soma expression of proximal and distal reporter genes (mean and SD of 6 replicates). See also Figure S2.

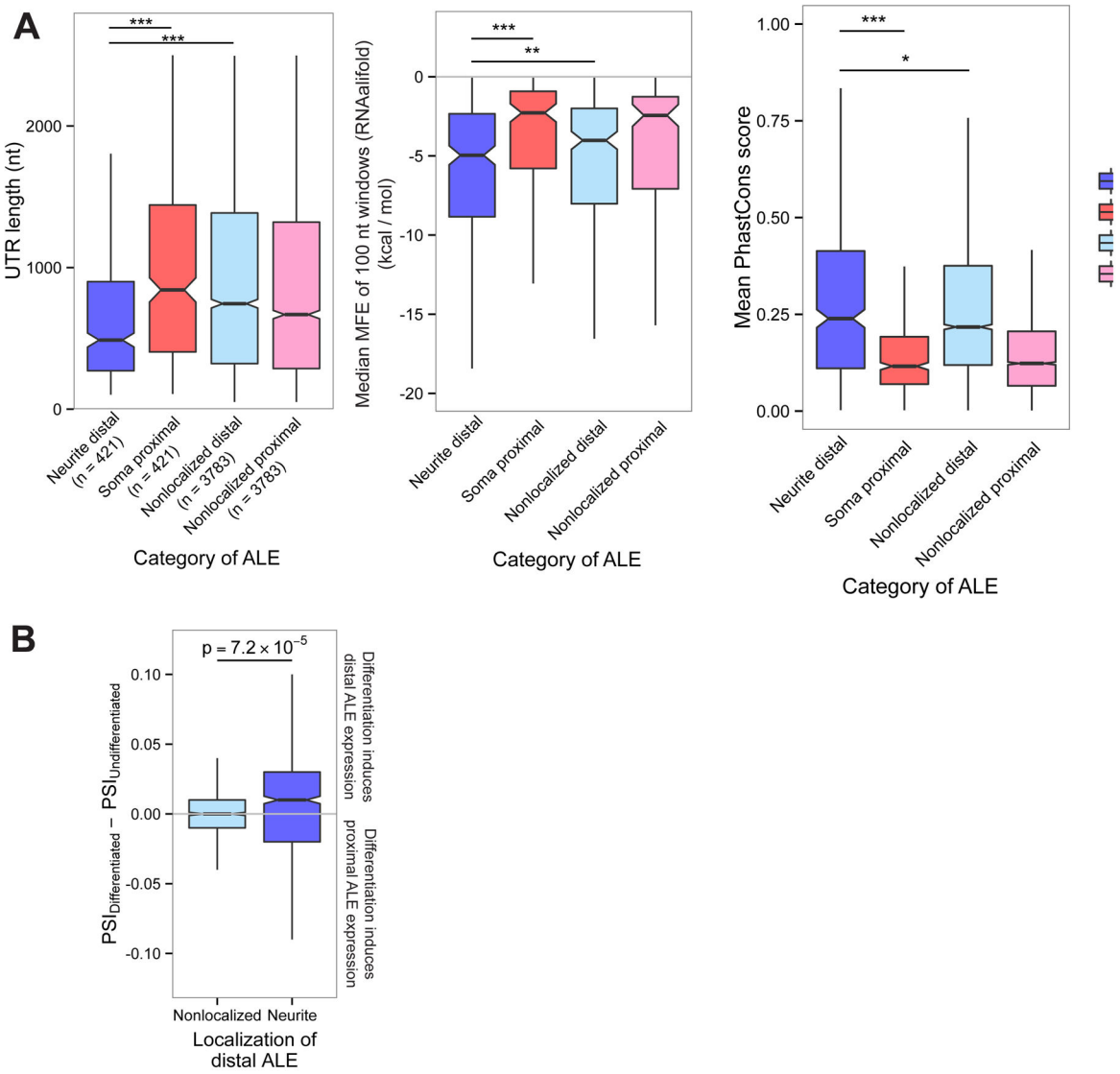

Figure 3. Distinctive properties of 3′ UTRs of neurite-localized distal ALEs

A) Left: Lengths of UTRs of neurite-localized distal ALEs identified in N2A and CAD cells, proximal ALEs of the same genes, and distal and proximal ALEs not associated with localization. Middle: UTRs from the indicated regions were aligned with homologous regions from human, rat, dog, and cow. RNA secondary structure minimum free energies (MFE) were then calculated for successive 100 nt windows of the alignment using RNAalifold. For each alignment, the median MFE was recorded. Right: PhastCons scores of 30-way alignments of UTRs from the indicated classes of ALEs. The score for each UTR was defined as the mean PhastCons score for all basepairs within the UTR. B) Increased PSI values following differentiation of CAD cells indicate preferential accumulation of neurite distal ALE isoforms but not of nonlocalized distal ALE isoforms. See also Figure S3, Table S2.

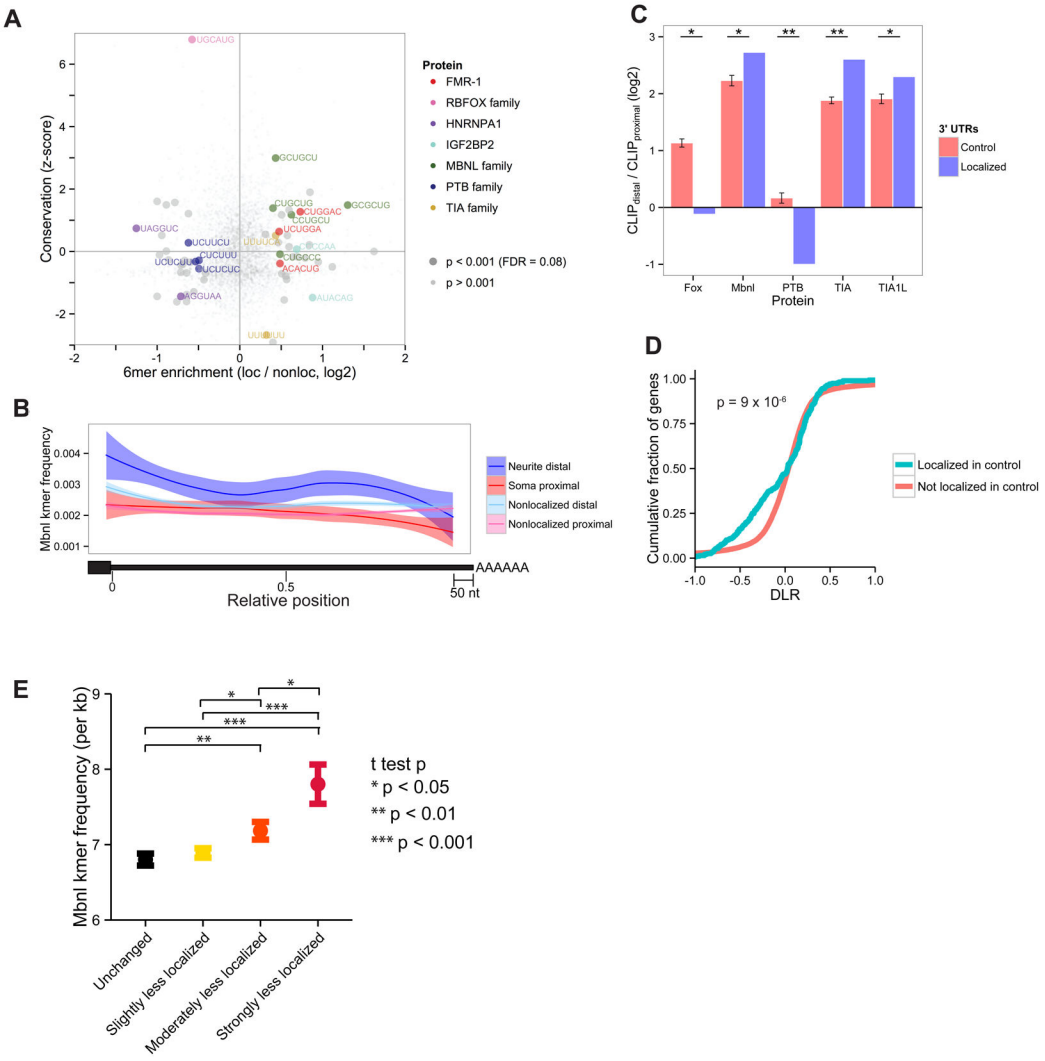

Figure 4. Mbnl motifs are enriched and conserved within localized distal ALEs, and Mbnl promotes RNA localization to projections

A) Enrichment of 6mers (hexanucleotides) between neurite distal and soma proximal UTRs and conservation of 6mers between mouse and human. Conservation is measured by a z-score representing the number of SD above the mean conservation of 50 control 6mers matched for CpG and C+G% content, in neurite distal UTRs. B) Metagene analysis of Mbnl motif frequency across UTRs from indicated classes (excluding the last 50 nt to exclude PAS motifs). These classes correspond to those defined in Figure 2. C) Relative CLIP-seq cluster densities in the UTRs of distal and proximal ALEs. Control UTRs consist of randomly sampled UTRs from all ALE events that were not differentially localized. Error bars are the standard error of random samplings of controls. D) Change in LR upon Mbnl knockdown for genes that were (blue) or were not (pink) localized in the control sample. E) Mbnl motif frequency across 3′ UTRs of ALEs as a function of the change in localization of that ALE in cortical neurons from *Mbnl1* / *Mbnl2* DKO mice. ALEs were classified by their ΔΔψ values as described in Supplemental Methods. Error bars represent +/− SEM. See also Figure S4, Table S3.

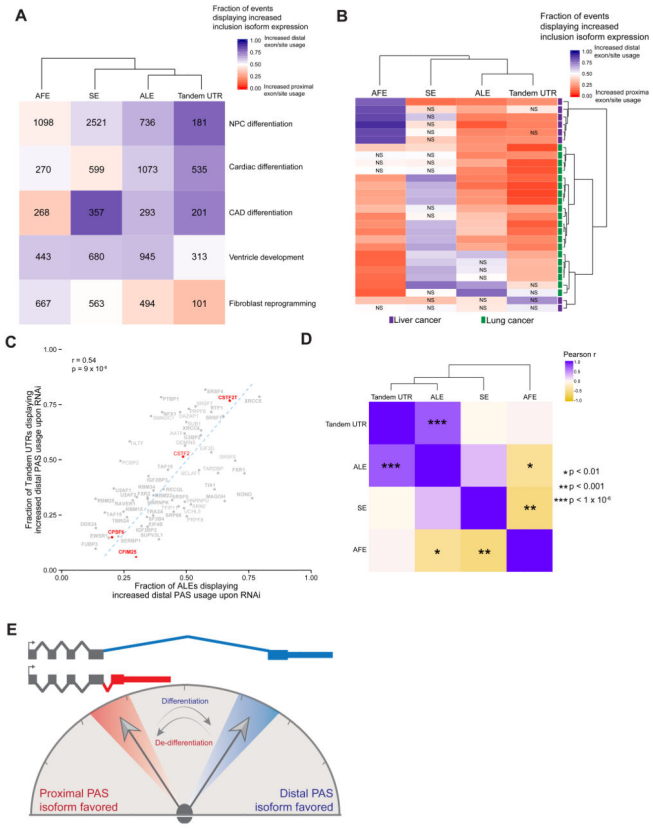

Figure 5. ALE and tandem UTR isoforms are generally coordinately regulated

A) For each row, the fraction of alternative isoform events that displayed an increase in relative abundance of the inclusion isoform in the differentiated or reprogrammed sample relative to its corresponding control was calculated. For each class of isoforms red indicates a shift in the differentiated/reprogrammed sample towards the proximal AFE, ALE or tandem UTR isoform or toward exon skipping, while blue corresponds to a shift toward the distal AFE, ALE or tandem UTR isoform or exon inclusion. The number inside the boxes corresponds to the number of significantly changing alternative isoforms in each sample. B) As in A, but comparing cancer samples to matched non-tumor controls. All samples significantly biased toward distal or proximal by chi-square test ($P < 0.05$) except those marked NS. C) The fraction of tandem UTR and ALE events displaying shifts towards distal PAS following the knockdown of RNA binding proteins in K562 cells. Gene names with significant shifts toward distal or proximal shown in bold. D) Correlation and clustering of isoform types indicated in A and B. E) Generally, development and differentiation result in a shift toward the inclusion of more distal ALEs. Conversely, becoming cancerous results in a shift toward more proximal ALEs.