University of Nevada, Reno

Molecular and physiological roles of long 3′ UTR mRNA isoforms in neurons

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Cellular and Molecular Biology

By

Bongmin Bae

Dr. Pedro Miura / Dissertation Advisor

December 2021

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

BONGMIN BAE

entitled

Molecular and physiological roles of long 3' UTR mRNA isoforms in neurons

be accepted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

Pedro Miura, Ph.D.
Advisor

Alexander van der Linden, Ph.D.
Committee Member

Yumei Feng Earley, M.D., Ph.D.
Committee Member

Jung Hwan Kim, Ph.D.
Committee Member

Ruben Dagda, Ph.D
Graduate School Representative

David W. Zeh, Ph.D., Dean
Graduate School

December, 2021

Abstract

The brain is an organ where the greatest proportion of genes are expressed compared to any other part of the body. To add even more complexity, gene expression in the brain is subject to various layers of regulation through RNA processing mechanisms including alternative splicing (AS) and alternative cleavage and polyadenylation (APA). These RNA processing mechanisms contribute to increased transcriptome diversity in the brain. APA often induces the synthesis of mRNA isoforms that harbor the same protein-coding sequence but different length 3′ untranslated regions (3′ UTRs) from a single gene. Alternative 3′ UTRs regulate gene expression post-transcriptionally by modulating transcript stability, translation efficiency, or subcellular localization. In Chapter 1, we reviewed all of the reported functions of 3′ UTRs in the nervous system.

Despite the fact 3′ UTR is highly regarded in gene regulation, evidence of impacts of long 3′ UTR loss on in vivo animal is scarce. To study the physiological relevance of long 3′ UTR mRNA isoforms, we have driven our attention to the *Calm1* gene. *Calm1* is one of the three genes that encode Calmodulin which is required for proper neural development and function. In Chapter 2, we found that the expression of the long 3′ UTR mRNA isoform of *Calm1* was necessary for mouse nervous system development and function. Disruption of the *Calm1* long 3′ UTR isoform impaired dorsal root ganglion axon development in mouse embryos and neuronal activation upon novel environment exposure in young adult mice. Our results presented direct evidence for the physiological importance of the *Calm1* long 3′ UTR mRNA isoform in vivo.

To screen molecular and cellular functions of long 3′ UTRs in a fast and efficient manner, establishing an in vitro cell system is warranted. In Chapter 3, we presented mouse embryonic stem cell (mESC)-derived neurons as a suitable cell-culture system.

The transcriptomic profile of the mESC-derived neurons closely resembled the profile in the mouse cerebral cortex, showing the suitability of using this system for studying long 3′ UTRs. The mESC system is amenable to genetic manipulation via CRISPR-Cas9, thus providing as good avenue for fast generation of long 3′ UTR isoform knockout lines. As a proof of principle, a workflow for the generation of Myosin phosphatase Rho interacting protein (*Mprip)* long 3′ UTR isoform knockout cell lines, differentiation into glutamatergic neurons, and confirmation of the long 3′ UTR expression abolishment is presented.

Taking advantage of the convenient culture cell system we have established, we next aimed to explore more functions of long 3′ UTRs. A recent discovery in our lab suggested that APA and AS are closely linked RNA processing mechanisms in which long 3′ UTRs modulate upstream AS. In Chapter 4, we explored the coupling events between AS and APA in mouse neurons using Pull-a-Long-Seq (PL-Seq) pipeline, which presents a particular utility in quantifying the coordination of tandem 3′ UTR APA events with upstream cassette exon AS. PL-Seq performed on the Endonuclease V (*Endov*) gene reveals that expression of its long 3′ UTR in neurons is preferentially associated with an exon skipping event located far upstream of the terminal exon.

Acknowledgments

I would like to deeply acknowledge my advisor, Dr. Pedro Miura, for all his patience, support, and guidance throughout my training. I am tremendously fortunate to receive all the academic and scientific advice that has made me a better person and scientist.

I would like to express my gratitude to my committee members, Dr. Alexander Van der linden, Dr. Ruben Dagda, Dr. Yumei Feng Early, and Dr. Jung Hwan Kim, for their thoughtful suggestions and discussion along the way.

I also would like to thank my colleagues and all the past and current Miura lab members who were open to discuss and troubleshoot all the details of experiments and project directions.

Lastly, without the support of my friends and family, all this wouldn't have been possible.

Table of Contents

List of Figures

Chapter I: Introduction

This chapter was published in International Journal of Molecular Sciences. 2020, 21(10), 3413. Authors: Bongmin Bae and Pedro Miura.

I.1. Roles of 3′ UTRs in Neuronal Gene Regulation

The 3′ untranslated regions (3′ UTRs) of mRNAs serve as hubs for post-transcriptional control as the targets of microRNAs (miRNAs) and RNA-binding proteins (RBPs). Sequences in 3′ UTRs confer alterations in mRNA stability, direct mRNA localization to subcellular regions, and impart translational control. Thousands of mRNAs are localized to subcellular compartments in neurons—including axons, dendrites, and synapses—where they are thought to undergo local translation. Despite an established role for 3′ UTR sequences in imparting mRNA localization in neurons, the specific RNA sequences and structural features at play remain poorly understood. The nervous system selectively expresses longer 3′ UTR isoforms via alternative polyadenylation (APA). The regulation of APA in neurons and the neuronal functions of longer 3′ UTR mRNA isoforms are starting to be uncovered. Surprising roles for 3′ UTRs are emerging beyond the regulation of protein synthesis and include roles as RBP delivery scaffolds and regulators of alternative splicing. Evidence is also emerging that 3′ UTRs can be cleaved, leading to stable, isolated 3′ UTR fragments which are of unknown function. Mutations in 3′ UTRs are implicated in several neurological disorders—more studies are needed to uncover how these mutations impact gene regulation and what is their relationship to disease severity.

1. Introduction

3′ untranslated regions (UTRs) enable post-transcriptional control to provide spatiotemporal regulation of gene expression within a cell. Recognition of cis-elements by RNA-binding proteins (RBPs) drives tight modulation of gene expression by altering mRNA degradation rates, driving subcellular localization of mRNA, and regulating translation efficiency. These changes ultimately impact the spatiotemporal dynamics of protein synthesis. In this review, the roles that 3′ UTRs play in controlling mRNA dynamics in the nervous system are discussed, along with new emerging roles for 3′ UTRs that go beyond their roles in mRNA localization and translation.

2. Cis-Acting Roles for 3′ UTRs in Neuronal Gene Regulation

2.1. miRNA Regulation

The 3′ UTR is particularly well understood as the target region for microRNA (miRNA) regulation (**Figure 1A**). Several miRNAs are specifically expressed in the nervous system and have roles in neural development and maintenance. During the early stages of neural specification and differentiation, gene regulatory networks that establish neuronal identity are regulated by miRNAs (Conaco et al. 2006; Lim et al. 2005; Rajman and Schratt 2017). miRNAs also regulate axon outgrowth/pathfinding and dendritogenesis (Chiu et al. 2014; Zampa et al. 2018; Rajman and Schratt 2017). This is achieved in part by miRNA suppression of mRNAs encoding a wide variety of cytoskeletal and signaling proteins (Dajas-Bailador et al. 2012; Franke et al. 2012; Baudet et al. 2012; Magill et al. 2010). In fully mature neurons, miRNAs are involved in synaptic plasticity (Wang et al. 2012; Kiltschewskij and Cairns 2019; Hu and Li 2017) and the regulation of circadian rhythms (Xue and Zhang 2018). Given these important roles for miRNAs in neurons, it is not surprising that alterations in miRNA expression is implicated in neurodegenerative

disorders, such as Alzheimer's disease (AD) and Parkinson's disease (Hébert and De Strooper 2009; Leggio et al. 2017; Goh et al. 2019), and neuropsychiatric problems, including depression, anxiety, schizophrenia, and autism spectrum disorder (Im and Kenny 2012; Issler and Chen 2015; Wu et al. 2016). Along the same lines, mutations in 3′ UTRs encoding miRNA seed sites have also been associated with several disorders (Dini Modigliani et al. 2014; Delay et al. 2011; Vaishnavi et al. 2014).

2.2. RNA Methylation

Methylation of DNA is well understood to regulate chromatin dynamics and transcription. In recent years, it has emerged that methylation of RNA also impacts gene regulation (**Figure 1B**). While there are many types of RNA modifications, N6-methyladenosine (m6A) is one of the most abundant and the best characterized to date (Li et al. 2019). Transcriptome-wide analysis revealed that m6A modifications are

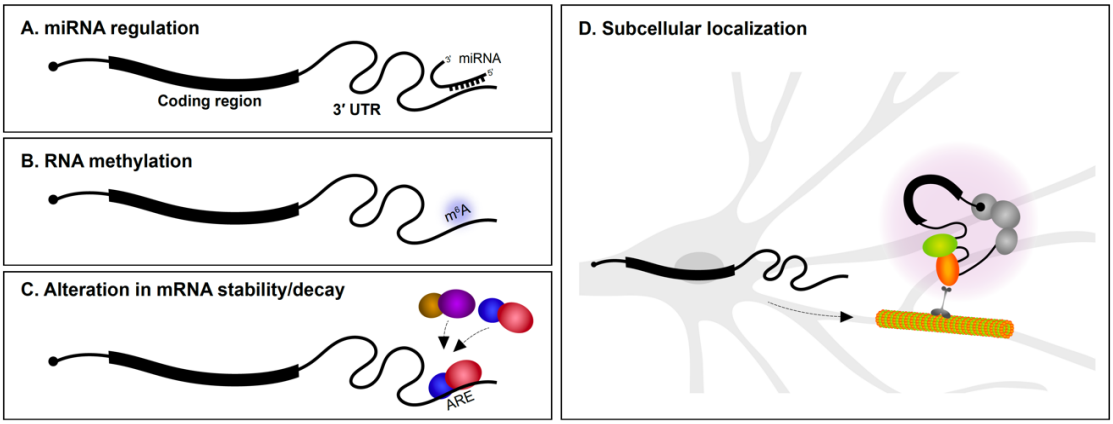


Figure 1. Post-transcriptional regulation via 3′ untranslated regions (3′ UTRs).
3′ UTRs mediate post-transcriptional gene regulation via (A) miRNA interactions, (B) RNA methylation, (C) regulating mRNA stability/decay by interaction with RNA-binding proteins (RBPs; illustrated as colored balls), and (D) imparting subcellular localization within neurons to regions such as dendrites and axons, where they can undergo local translation. ARE, AU-rich element.

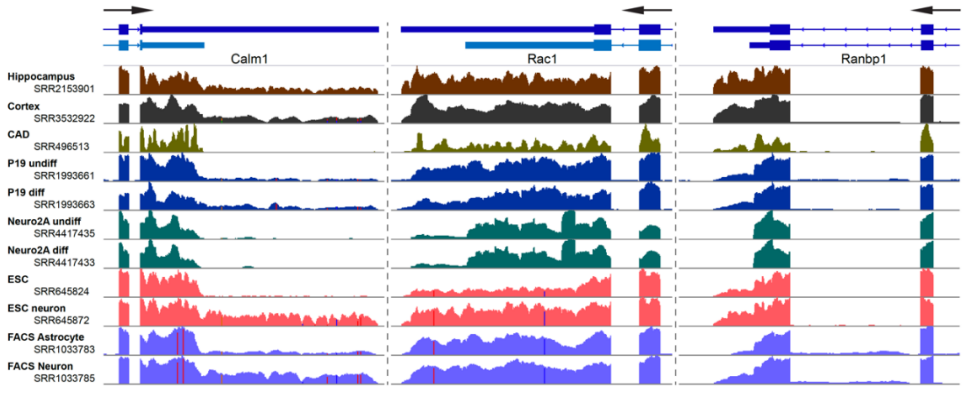
particularly enriched in 3′ UTRs (Meyer et al. 2012; Ke et al. 2015). m6A modifications play roles in axon guidance, neurogenesis, neural survival, and synaptic function (Widagdo and Anggono 2018; Du et al. 2019). m6A modifications appear to be particularly important for control of mRNA translation. YTHDF1 (YTH N6-methyladenosine RNA binding protein 1) is a m6A reader protein. By binding to m6A marked regions, predominantly in 3′ UTRs, YTHDF1 regulates protein synthesis in response to neuronal activity. Mice lacking YTHDF1 display impaired hippocampal synaptic transmission and defects in learning and memory, demonstrating the key role for RNA methylation in the nervous system (Shi et al. 2018).

2.3. mRNA Stability and Translational Control

The rate of mRNA decay is influenced by several determinants, including poly(A) tail length and 3′ UTR sequence content. RBP interactions with 3′ UTRs are important determinants of mRNA stability (**Figure 1C**). AU-rich elements (AREs) located in 3′ UTRs were first reported to mediate mRNA decay of cytokine transcripts (Caput et al. 1986). AU-rich binding proteins bind AREs and can stabilize or promote degradation of the mRNA. For instance, the KH-type splicing regulatory protein (KSRP or KHSRP) and HuD (ELAV-like protein 4) are RBPs that bind to an ARE in the 3′ UTR of the GAP-43 (Gap43) mRNA. Although both proteins target the same region, they employ antagonistic regulation of GAP-43 mRNA—KSRP enhances its turnover during axonal outgrowth of hippocampal neurons (Bird et al. 2013), whereas HuD stabilizes it in select neuronal populations (Bolognani et al. 2006). mRNA stability influences not only the number of times an mRNA can be translated, but also impacts mRNA localization since the transcript must resist degradation in order to be subcellularly localized.

Box 1. APA trends in cultured neuronal-like cells versus neurons

Selection of the proper biological system is a key for studying the dynamics of alternative 3′ UTR usage in neurons. Most previously reported studies have employed either (1) immortalized neuron-like cell lines, (2) isolated neurons, (3) stem cell/induced pluripotent stem cell derived neurons, or (4) tissues. A key benefit of using cell lines are their malleability. For instance, the properties of 3′ UTRs and localizing potentials of Tau and GAP-43 have been largely studied in P19 and PC12 cells (Aronov et al. 2001; Atlas R et al. 2004; Smith et al. 2004). Much of the research focusing on miRNA binding to 3′ UTRs has been performed in Neuro2A cells (Boissonneault et al. 2009; Lee et al. 2012) and the localization of alternative last exon 3′ UTR isoforms in neurites was studied using differentiated CAD and Neuro2A cells (Taliaferro et al. 2016). Expression of alternative 3′ UTR isoforms can vary widely between mouse tissues and cell lines. Several examples are shown in this figure. RNA-Seq reads were aligned using HISAT2 (Kim et al. 2015), processed using SAMtools (Li et al. 2009), and tracks visualized at the last two exons of the *Calm1*, *Rac1*, and *Ranbp1* genes using Integrated Genomics Viewer (Thorvaldsdóttir et al. 2013). Note the changes in read coverage pertaining to the alternative long 3′ UTRs. Gene models in light blue represent un-annotated transcript isoforms. SRA accession numbers are noted.



3. 3′ UTR-Mediated Subcellular Localization of mRNAs within Neurons

Ever since the first findings of asymmetric localization of mRNAs in ascidian eggs (Jeffery et al. 1983), subcellular localization of transcripts in polarized cells has been of great interest. In neurons, mRNAs are found in axons, dendrites, and synapses (Garner et al. 1988; Burgin et al. 1990; Bassell et al. 1998; Lyford et al. 1995) (**Figure 1D**). The number of subcellularly localized transcripts identified in neurons continues to grow as a result of advances in transcriptomics and methods to isolate dendrites and axons from soma. Most mRNA localization transcriptome studies have been performed in cultured primary explants/neurons (Taylor et al. 2009; Gumy et al. 2011; Minis et al. 2014; Zivraj et al. 2010) and immortalized neuronal-like cells such as Neuro2A and CAD cells (Taliaferro et al. 2016) (see Box 1). Methods to physically isolate long axons/processes from cultured tissues/cells include the use of compartmentalized chambers and membrane inserts. These allow axons to be separated from soma for downstream RNA analysis. We are now aware that hundreds to thousands of mRNAs are found in axons or dendrites of both the peripheral and central nervous systems (Minis et al. 2014; Tushev et al. 2018; Middleton et al. 2019), and that the transcriptome of each neuronal subcompartment is unique. It is more challenging to identify dendrite/axon localized mRNAs using sequencing approaches in vivo. Laser microdissection of dendrite-enriched regions followed by RNA-Seq has been successfully performed in various systems including rat brain slices (Tushev et al. 2018). It is clear from such studies that many mRNAs are found in dendrites and axons.

3.1. RNA Localizing Cis-Elements

The study of mRNA subcellular localization determinants has focused on *cis*-acting sequence elements in mRNAs and the RBPs that bind them. The subcellular localization of mRNAs in neurons has often been attributed to 3′ UTR sequences (Willis et al. 2011;

Andreassi and Riccio 2009). Since the identification of a localizing *cis*-element in the β-actin 3′ UTR, known as zipcode, and its *trans*-acting RBP partner, zipcode-binding protein (ZBP; as known as IMP1) (Kislauskis et al. 1994; Ross et al. 1997; Tiruchinapalli et al. 2003), considerable attention has been focused on determining 3′ UTR *cis*-elements and RBP *trans*-factors influencing mRNA localization (Minis et al. 2014; Vuppalanchi et al. 2010; Middleton et al. 2019). Although some motif enrichment has been observed, efforts to find a universal "zipcode" sequence governing mRNA localization has been unsuccessful. Sequences identified as localization elements are largely unique to each gene. Thus, mRNA localization elements cannot currently be identified by a motif-searching bioinformatics approach (for examples of localizing *cis*- and *trans*- elements, refer to (Andreassi and Riccio 2009)).

Assessing the role of 3′ UTRs in mRNA localization has heavily relied on the use of 3′ UTR reporter assays. For this type of experiment, the 3′ UTR sequence of interest is subcloned downstream of a reporter, such as a fluorescent protein (FP), or adjacent to mRNA reporter motif, such as MS2 bacteriophage coat protein-binding sequences, to observe the effect of the 3′ UTR on localization (For more experimental approaches, see Box 2). Then, a minimal localizing element can be identified by testing the effect that 3′ UTR deletions or mutations have on reporter mRNA axonal or dendritic localization. Using reporter systems, 3′ UTR sequence determinants for localization have been identified for CaMKIIα (*Camk2a*), β-actin (*Actb*), GAP-43, Importinβ1 (*Kpnb1*), and *Rgs4* mRNAs (Mayford et al. 1996; Perry et al. 2012; Yoo et al. 2013; Zhang et al. 2001; Bauer et al. 2019).

Only a few studies have uncovered the impact of 3′ UTR sequences on localization using loss of function approaches in vivo. The localizing role of the 3′ UTR of CaMKIIα in mice was confirmed by inserting a heterologous poly(A) site into the endogenous locus to

prevent full length 3′ UTR generation. This approach successfully prevented mRNAs from being localized in dendrites (Miller et al. 2002). 3′ UTR-mediated localization of β-actin mRNAs in vivo was identified using a heterologous reporter construct harboring different 3′ UTR sequences (Willis et al. 2011). The heterologous reporter transgenic approach showed that the 3′ UTR of β-actin directed expression of the reporter gene to axons (Willis et al. 2011). Localization of *Bdnf* alternative 3′ UTR mRNAs have been investigated in vivo through similar approaches as well (see section 4.2. Neural functions of long 3′ UTR mRNA isoforms). Recent technical advances in genome editing have facilitated 3′ UTR deletions with increased precision and speed in animal models. For example, the CRISPR-Cas9 (Clustered regularly interspaced short palindromic repeats-Cas9) system was used to delete part of the mTOR (*Mtor*) 3′ UTR which was found to impair local translation in dorsal root ganglion (DRG) neurons (Terenzio et al. 2018).

All these results suggest that 3′ UTRs impact the localization of transcripts in neurons in vitro and in vivo. There is plenty of evidence for certain sequences to be sufficient for mRNA localization in neurons, but there is still scant evidence that they are necessary. For example, although the 3′ UTR of β-actin was found to be sufficient to drive expression of the reporter gene in axons (Willis et al. 2011), there are still no published studies, to our knowledge, showing the impact of deleting zipcode sequences from the genome on endogenous β-actin mRNA localization. Precise genome-editing techniques such as CRISPR/Cas are now commonplace, so in the coming years we expect new studies that characterize transgenic animals with deletions in single putative localization elements.

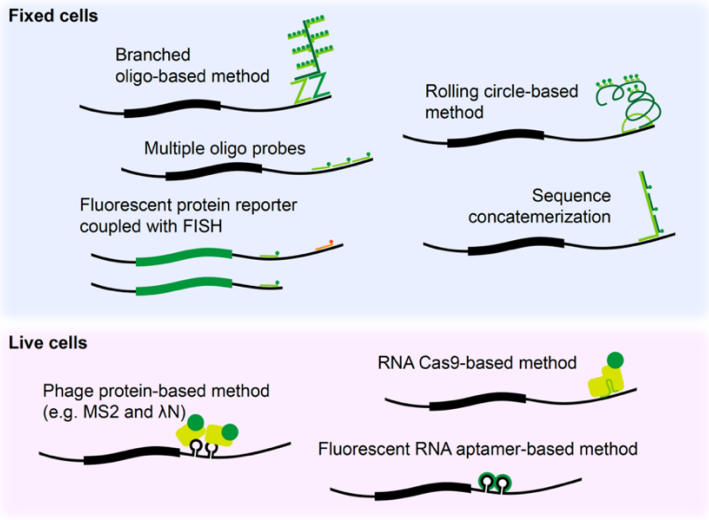
mRNA localizing elements can be more complicated than simply a primary nucleic acid sequence. For instance, the spatial arrangement of localization elements within the 3′ UTR can impact localization (Patel et al. 2012). Localizing elements can also be

comprised of structural features, such as G-quadruplex motifs (Subramanian et al. 2011). Although identification of RNA structural motifs is challenging, it is becoming more feasible with the advent of new RNA structure probing methods that allow for identification of RNA secondary structures in cells (Bevilacqua et al. 2016). Perhaps the continued search for single localization elements is misguided? It is emerging that multiple elements with redundant functions within 3′ UTRs might impact localization - 3′ end sequencing analysis of axons and cell bodies found that multiple elements spanning across 3′ UTRs influence subcellular localization as opposed to single regions (Tushev et al. 2018). It should also not be forgotten that localization elements can also be found in 5′ UTRs (Brigidi et al. 2019; Merianda et al. 2013; Meer et al. 2012). Perhaps many localized mRNAs rely on the presence of multiple, redundant localization elements in both 5′ UTRs and 3′ UTRs.

Box 2. Experimental approaches to study localization of mRNAs

Fluorophore-labeled probe-based methods, such as fluorescence in situ hybridization (FISH), have improved in situ detection of mRNAs in fixed neurons in term of resolution and sensitivity compared to previous in situ hybridization methods, allowing single molecule RNA detection. Largely, two types of RNA FISH methods are available. The first method is based on usage of multiple oligo probes each harboring a fluorophore that target a same single RNA molecule (e.g., Stellaris®) (Raj et al. 2008). The other type of method is based on amplification of fluorescence signal by in situ biochemical reactions, such as rolling-circle based method (e.g., OligoMix®) (Larsson et al. 2010), branched DNA method (e.g., RNAscope®) (Wang et al. 2012), and primer-exchange reaction based method (e.g., SABER-FISH) (Kishi et al. 2019). Advanced techniques, such as MCP-FP (MS2 bacteriophage coat protein-FP), λN-FP (N protein of bacteriophage λ-FP), RCas9-FP (dead RNA Cas9-FP), and fluorescent RNA aptamer

system, have allowed visualization of RNA trafficking in live cells. Co-expression of a MCP-FP or λN-FP protein construct and a reporter construct containing phage protein binding motif sequence upstream of 3′ UTR of interest allowed tracking of mRNA localization in live cells (Daigle et al. 2007; Fusco et al. 2003). Simultaneous delivery of RCas9-FP and target-specific single guide RNA allowed binding of the Cas9 to the mRNA of interest and visual tracking of endogenous mRNAs in live cells (Nelles et al. 2016). Use of fluorescent RNA aptamers, such as Peppers, has overcome dimensional limitations of FP tethering techniques and enhances signal-to-noise ratios allowing improved in vivo tracking (Chen et al. 2019).



3.2. Role of RBPs in mRNA Localization

How do 3′ UTR sequences in mRNAs drive localization to axons and dendrites? An attractive hypothesis is that the mRNAs are actively transported. The trafficking of RBP–mRNA complexes, referred as messenger ribonucleoproteins (mRNPs), has been associated with cytoskeleton-based transport systems (Litman et al. 1994; Gagnon and Mowry 2011; Holt and Bullock 2009). Motor proteins usually lack RNA-binding domains, thus, it is widely considered that the localizing transcripts associate with RBPs and

accessory proteins to indirectly interact with motor proteins (Gagnon and Mowry 2011). ZBP1 (which interacts with the β-actin 3′ UTR) has been reported to interact with the motor protein KIF5A (kinesin heavy chain isoform 5a) and Myosin Va (Nalavadi et al. 2012; Urbanska et al. 2017). ZBP1-containing mRNPs mediate dynamic movement and dendritic localization of β-actin mRNA (Tiruchinapalli et al. 2003).

Alternatively, an mRNA might be found in a particular subcellular compartment because of passive diffusion and long half-life conferred by 3′ UTR–RBP interactions (Medioni et al. 2012). A transcriptome-wide assessment of mRNA stability in somata and neuropil compartments of rat Cornu Ammonis (CA1) region suggested that neuropil localized 3′ UTR isoforms have longer half-life compared to non-localized 3′ UTR mRNA isoforms (Tushev et al. 2018). This might be, at least partially, due to increased GC-content and predicted secondary structure elements in neuropil localized RNAs which confer structural stability (Tushev et al. 2018). The same study also has revealed that neuropil localized 3′ UTR mRNA isoforms are enriched for neuronal RBP motifs that have been associated with mRNA localization (Tushev et al. 2018). RBP-mediated selective transport and mRNA stability-dependent localization mechanisms need not be mutually exclusive since many RBPs are multi-functional and could mediate both active localization and stability.

3.3. RNA Granules

There are emerging roles for RNA granules in neuronal mRNA localization and translation control (Krichevsky and Kosik 2001). RNA granules are membraneless intracellular structures that possess liquid-like properties to support their dynamic organization and serve as a means of compartmentalizing and accelerating biochemical reactions (Hyman et al. 2014; Hyman and Brangwynne 2011; Weber and Brangwynne 2012). These granule structures are self-assembled through liquid–liquid phase

separation. One of the components that facilitates liquid–liquid phase transition and formation of granules are proteins that contain low complexity domains. These domains induce liquid–liquid phase transition through self-aggregation (recently reviewed in (Franzmann and Alberti 2019; Martin and Mittag 2018)). Thus, it has been proposed that RBPs with low-complexity domains might control both selective recruitment of RNAs and the formation of RNA granules through phase transition (Han et al. 2012; Kato et al. 2012). For instance, FMRP (fragile X mental retardation protein) and FUS (fused in sarcoma) proteins induce phase transition and translational control in mRNPs (Murakami et al. 2015; Tsang et al. 2019). An interesting characteristic of mRNA components of these granules is that these mRNAs contain significantly longer 3′ UTR sequences and are enriched in binding motifs for the granule RBPs (Han et al. 2012), suggesting 3′ UTR contents impact RNA granule dynamics.

A recent study showed that protein constituent of RNA granules also mediates the interaction between RNA granules and moving membranous organelles to drive microtubule-dependent long-distance transport of RNAs. Annexin A11 is a protein containing a low-complexity domain and membrane binding domain. Through these domains, Annexin A11 mediates association of RNA granules with lysosomes to control RNA localization in neurons (Liao et al. 2019). In vivo functional roles in neurons for phase transition have recently been uncovered. TIAR-2 (TIA-1/TIAL RNA binding protein homolog), an RNA-binding-domain-containing TIA family protein found in *Caenorhabditis elegans*, was found to mediate phase transitioning and formation of liquid-like granules. Mutating prion-like domains in TIAR-2 mitigates granule formation and impaired axon regeneration (Andrusiak et al. 2019). How RNA granules regulate gene expression and impact nervous system functions will certainly be an area of intense investigation for years to come.

4. Alternative Polyadenylation Generates Long 3′ UTRs in Neurons

Alternative polyadenylation (APA) is an RNA processing mechanism that allows the generation of alternative mRNAs with distinct 3′ ends from a single gene through the selective usage of alternative polyadenylation (poly(A)) sites. In mammals, at least 50%–70% of protein-coding genes undergo APA, and many of these alternative 3′ UTRs show tissue-specificity of expression (Hoque et al. 2013; Lianoglou et al. 2013a; Tian et al. 2005). APA can be largely classified into two categories (**Figure 2**). One produces alternative protein isoforms through internal/intronic polyadenylation or through alternative last exon splicing choices. The other type of APA occurs within the 3′ UTR via tandem poly(A) sites and generates transcript isoforms that harbor the same coding sequences (CDS) and thus only differing in their 3′ UTR content.

Early microarray studies and work on individual genes revealed a bias among brain tissues for expressing higher levels of long versus short 3′ UTR mRNA isoforms for genes expressed in multiple tissues (Pelka et al. 2005; Costessi et al. 2006; Zhang et al. 2005). Later studies using RNA-Seq revealed the transcriptome-wide catalog of these neural-specific 3′ UTR lengthening events (see Box 3 for bioinformatics methods to quantify alternative 3′ UTR usage). In *Drosophila*, nearly 400 genes were found to express previously unannotated long 3′ UTR isoforms in head samples and late stage embryos (Smibert et al. 2012). Some of these transcripts had 3′ UTRs of staggering length, dwarfing the length of the protein-coding region of the mRNA. For example, the long 3′ UTR of *mei-P26* was found to express an 18.5 kb long 3′ UTR (Smibert et al. 2012). Investigation of long 3′ UTRs in neural tissues of mouse and human yielded similar findings of neural-specific enhancement of long 3′ UTRs with thousands of previously unannotated long 3′ UTR isoforms being identified (Miura et al. 2013).

A bias for longer 3′ UTRs in the nervous system appears to be attributed to expression in neurons as opposed to other neuronal cell types such as astrocytes, microglia, and oligodendrocytes (Ha et al. 2018; Shepard et al. 2011; Guvenek and Tian 2018). This raises two questions: (1) What is the mechanism that leads to longer 3′ UTRs in neurons? (2) Do these longer 3′ UTR mRNAs have specific neural functions?

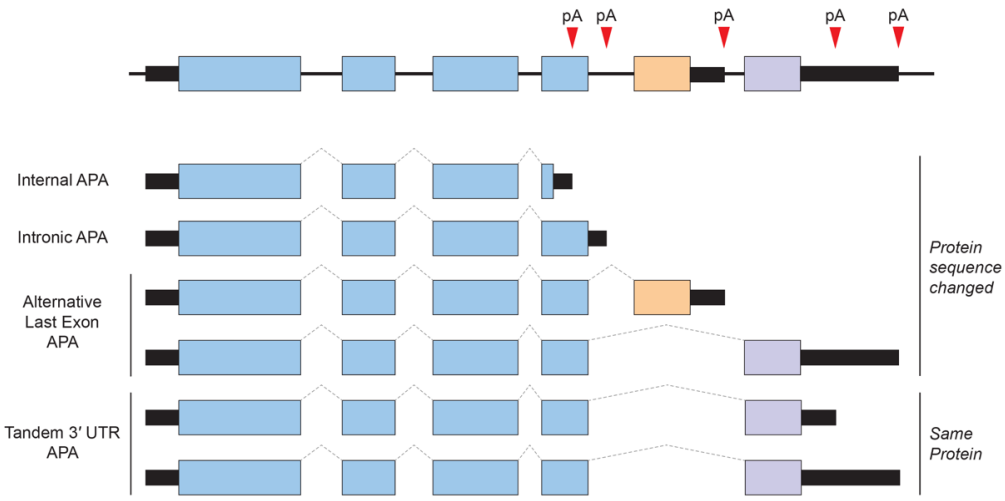


Figure 2. Types of alternative polyadenylation (APA).
APA generates alternative 3′ ends for a given gene. Tandem 3′ UTR APA is the most common and does not cause changes in protein-coding regions, whereas other APA events that are sometimes linked to regulated alternative splicing can result in alternative last exons that have different coding and 3′ UTR content. pA, poly(A) site.

4.1. APA Mechanisms Underlying Long 3′ UTR Expression in Neurons.

The mRNA 3′ end processing machinery includes cleavage and polyadenylation specificity factors (CPSF), cleavage stimulation factors (CstF), cleavage factors (CF I and II), poly(A)-binding protein (PABP), poly(A) polymerase (PAP), and RNA polymerase II (Pol II), among other proteins (Chan et al. 2011). One of the mechanisms that mediates

APA selection involves these 3′ end processing factors. APA patterns have been associated with the expression levels of CstF64 (Takagaki et al. 1996; Yao et al. 2013; Shankarling and MacDonald 2013) and CFI (Martin et al. 2012; Alcott et al. 2020). These studies suggested that the concentration of the specific processing factors might define the usage of poly(A) signals leading to the cell/tissue-specific 3′ UTR expression patterns. The speed or pausing of Pol II can also influence APA, by providing sufficient time to recruit 3′ end processing factors and to process the poly(A) site (Fusby et al. 2015). A slower RNA Pol II mutant showed preferential selection of proximal poly(A) sites in *Drosophila* (Pinto et al. 2011); however, this effect was limited to non-neural tissues, suggesting a different regulatory mechanism at play in neurons (Liu et al. 2017).

RBPs can influence APA by competing with the cleavage and polyadenylation machinery for access to the poly(A) site. One of the most striking examples of RBP regulating 3′ UTR extension is the neural-specific RBP, Elav (embryonic lethal abnormal visual protein) in *Drosophila*. *Drosophila* embryos lacking Elav were found to lack long 3′

Box 3. Quantification of alternative 3′ UTR mRNA isoforms using standard RNA-Seq data

Although RNA-Seq has become a routine procedure, the identification and quantification of alternative 3′ UTR isoforms using RNA-Seq data presents many challenges. Primarily, two types of detection algorithms are currently used: (1) de novo detection of APA isoforms based on the read density changes and (2) reliance on annotated or reported 3′ ends. De novo detection-based methods do not rely on 3′-end sequencing data or previously reported 3′ ends, thus providing unique advantages. Change-Point is a 3′ UTR APA detection software that identifies APA events between

two conditions based on read density changes (Wang et al. 2014). It compares the ratio of mapped reads in the common 3′ UTR region and the ratio of the extended 3′ UTR region between two samples and the identification of exact APA site is based on the location where ratio change is the maximum. Dynamic analyses of alternative polyadenylation from RNA-Seq (DaPars) is a de novo 3′ UTR APA detection software that uses similar read density-based approach but quantifies dynamic APA events between two or more experimental conditions and allows identification of novel long 3′ UTRs (Xia et al. 2014). In contrast to the two previous methods, APAtrap allows identification of multiple APA sites. First, novel 3′ UTR ends are identified by assessing the coverage of RNA-Seq reads through the annotated 3′ end and further downstream regions. Once the most 3′ end is defined as the distal poly(A) site, then potential proximal poly(A) sites are identified using the read density changes (Ye et al. 2018). Given the expansion of 3′ end databases, APA events can also be reliably detected using methods based on annotated Poly(A) sites. Quantification of APA (QAPA) estimates the expression of alternative 3′ UTR isoforms using annotated poly(A) sites. First, it builds a reference library comprising all the reported 3′ UTR sequences. Then, RNA-Seq reads are mapped to the 3′ UTR library using alignment-free algorithm. APA usage is quantified by the ratio of isoform expression to the sum of the expression of all detected 3′ UTR isoforms from its cognate gene. This method identifies multiple APA events and is able to process multiple datasets at a time (Ha et al. 2018). Significance analysis of alternative polyadenylation using RNA-Seq (SAAP-RS) is another method based on Poly(A) database. The 3′ UTR regions are split into upstream (UP) and downstream (DN) regions. RNA-Seq read distribution is assessed in the UP and DN region to determine the relative expression difference (Guvenek et al. 2018).

UTR isoforms for several genes, and the ectopic expression of Elav in non-neural tissues resulted in mRNAs with longer 3′ UTRs (Hilgers et al. 2012). A putative mechanism for this regulation is through Elav binding to the proximal poly(A) site, thus competing with 3′ end processing, and causing the selection of a downstream poly(A) site (Hilgers et al. 2012) (**Figure 3**). In addition to directly binding RNA, an additional role for Elav binding to DNA promoter sequences has been demonstrated, and this might also be required for Elav regulation of APA in *Drosophila* neurons (Oktaba et al. 2015). In mammalian neural tissues or neuron-like cells, the Elav homolog ELAVL3 (ELAV-like protein 3), along with NOVA2 and FUS have been shown to impact 3′ UTR APA (Grassi et al. 2019; Licatalosi et al. 2008; Masuda et al. 2015).

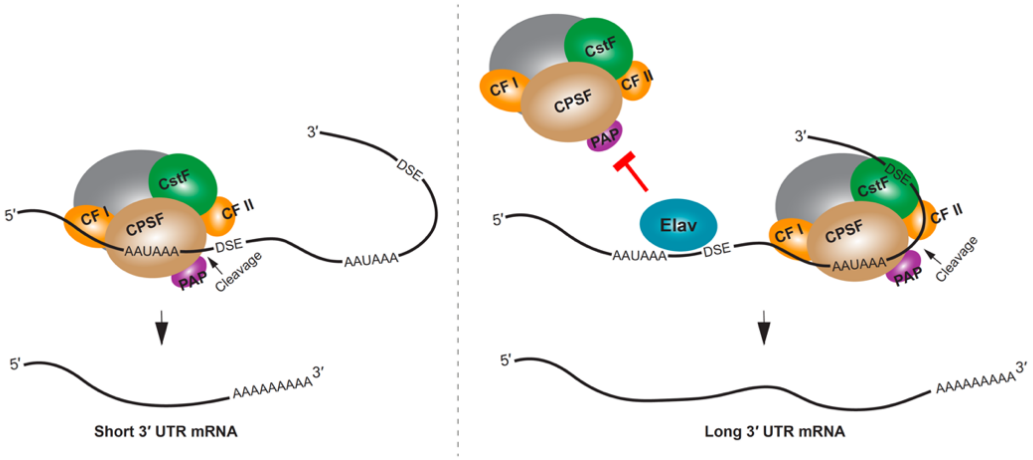


Figure 3. Mechanisms of neural APA in Drosophila.
The neuronal RBP Elav regulates APA by binding to target mRNA proximal poly(A) site region to block the 3′ end processing machinery. This promotes the usage of downstream distal poly(A) sites. Consequently, neuronal tissues of Drosophila selectively express APA isoforms with longer 3′ UTRs. DSE, downstream element; AAUAAA is shown as a representative polyA signal.

4.2. Neural Functions of Long 3′ UTR mRNA Isoforms

Given the relevance of 3′ UTRs in controlling the subcellular localization of mRNAs, one hypothesis of long 3′ UTR function has been that it localizes due to the extra 3′ UTR sequences not found in the short 3′ UTR counterpart. This phenomenon was first characterized for the long 3′ UTR mRNA isoform of *Bdnf*. Loss of the long 3′ UTR mRNA isoform was found to impair dendritic localization of the mRNA and cause defects in spine morphology and synaptic plasticity (An et al. 2008).

Based on the findings for *Bdnf*, groups have speculated that a function of the alternative long 3′ UTRs might be to impart subcellular localization. However, transcriptomic studies have failed to identify a strong bias for long 3′ UTR isoforms to be preferentially localized compared to short 3′ UTR counterparts (Tushev et al. 2018; Middleton et al. 2019). Although many 3′ UTR isoforms showed subcellular localization bias, a significant number of the alternative 3′ UTR transcripts reside in the same soma or neuropil subcompartment (Tushev et al. 2018). Another study actually found evidence that shorter 3′ UTR isoforms are more abundant than long 3′ UTR isoforms in neurites of mouse embryonic stem cell (ESC)-derived neurons (Ciolli Mattioli et al. 2019). These results suggest that long 3′ UTRs may not generally confer axon/dendrite localization.

5. Functional Relevance of Local Translation in Neurons

Regardless of the roles for alternative 3′ UTR isoforms, the notion of localized mRNAs undergoing local translation in both dendrites and axons is now well appreciated (Perry and Fainzilber 2014; Biever et al. 2020). Local translation allows precise spatiotemporal regulation of protein expression. The specific role of localized mRNA translation has been explored largely in three contexts: (1) neuron development and

neurite outgrowth, (2) synaptic function in mature neurons, and (3) neurite maintenance and regeneration.

Localization and translation of mRNAs in growing neurons has been explored especially for cytoskeleton-related genes to support neurite pathfinding and asymmetrical organization. Tau, an axonal microtubule-associated protein encoded by the *Mapt* gene, is one mRNA that localizes mainly into axons, and its role in establishing neuronal polarity might rely on axonal translation (Litman et al. 1993; Morita and Sobuě 2009; Zempel and Mandelkow 2019). *Rhoa* mRNAs axonally localize in embryonic DRG neurons in a 3′ UTR-dependent manner. During cytoskeletal rearrangement of the growth cone, 3′ UTR-dependent local translation of *RhoA* is necessary to induce growth cone collapse in response to guidance cue Sema3A (Semaphorin 3A) (Wu et al. 2005).

The ability of axonally or dendritically localized mRNAs to undergo local translation is critical for synaptic functions. Synaptic plasticity, including homeostatic scaling, is a process that is dependent on translation of axonal and dendritic proteins (Kang and Schuman 1996; Huber et al. 2000; Scarnati et al. 2018; Schanzenbächer et al. 2016). Synaptic concentration of ion channels, such as Kv1.1 (*Kcna1*), is regulated through 3′ UTR-dependent localization and translation in hippocampal neurons (Raab-Graham et al. 2006). Synaptic rearrangement is another cellular process that takes advantage of local translation. Matrix Metalloproteinase 9 (*Mmp9*), an endopeptidase that regulates the pericellular environment, is localized in a 3′ UTR-dependent manner and undergoes translation when hippocampal neurons are activated (Dziembowska et al. 2012). Neurogranin (*Nrgn*) undergoes local translation, during which FMRP exerts regulation through its 3′ UTR, and plays a role in synaptic plasticity and memory encoding (Jones et al. 2018). All these examples lead to an open conclusion that synaptic function requires local translation of specific transcripts.

Local translation is relevant for neuron maintenance and regeneration—it is implicated in retrograde signaling, mitochondrial function, and transcription factor activation related to the axon maintenance and regeneration. Several proteins involved in injury-induced retrograde singling pathways have been shown to be locally translated. Some examples include 3′ UTR mRNA isoforms of *Ranbp1* (Ran-specific binding protein 1) and Importin β, which are involved in nuclear export and import, respectively (Yudin et al. 2008; Perry et al. 2012). Mitochondria-related proteins also undergo local translation to support axon maintenance (Hillefors et al. 2007). Nerve growth factor-induced local translation of myo-inositol monophosphatase-1 (*Impa1*) is critical for nuclear CREB (cyclic AMP-responsive element-binding protein) activation and prevention of axonal degeneration (Andreassi et al. 2010). Local protein synthesis and turnover during the axon regeneration is considered to be regulated by various pathways dependent on mTOR, p38 mitogen-activated protein kinase, and caspase (Verma et al. 2005). mTOR regulates local translation of other mRNAs through its own local translation in injured axons and regulates other retrograde injury signaling molecules (Terenzio et al. 2018).

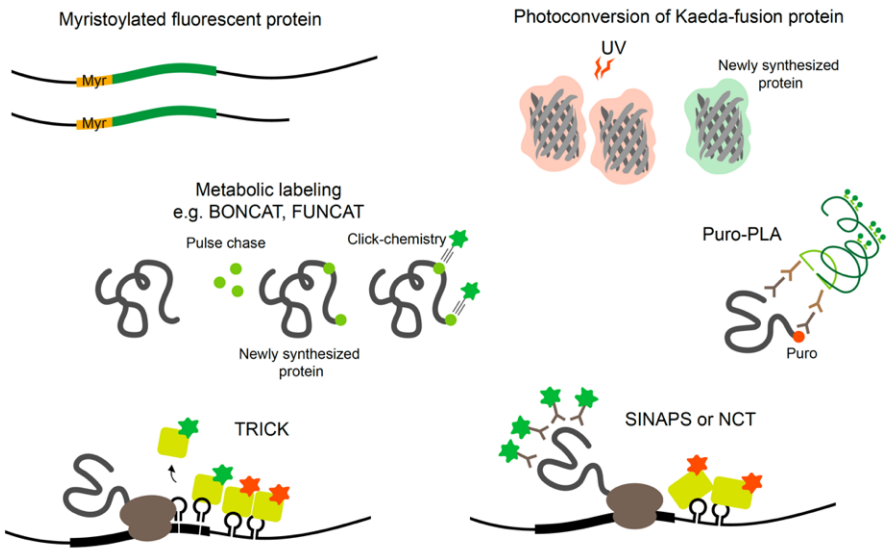
Multiple techniques have been developed in order to study local translation. Some of these techniques rely on reporter or fusion protein system and some on metabolic or genetic labeling (see Box 4). Studying local translation in vivo in animals, however, is more challenging. Some metabolic labeling techniques, such as FUNCAT (fluorescent noncanonical amino acid tagging), BONCAT (bio-orthogonal non-canonical amino acid tagging), and OP-Puro (O-propargyl-puromycin) labeling, enable global tagging of the newly synthesizing proteins in vivo through specific diet or injection of metabolic analogs (Dieterich et al. 2010, 2006; Hidalgo San Jose and Signer 2019). These techniques, however, have not been applied to in vivo nervous system and do not allow transcript (3′

UTR isoform)- or gene-specific analysis but rather confirm robustness of translation upon specific physiological contexts.

Box 4. Studying local translation

Investigating the mRNA sequences regulating mRNA localization and local translation has largely relied on the use of fluorescent protein (FP) reporters. One of the caveats of using FP to study localized translation is the ability to freely diffuse within cells. Myristoylation of FP (Myr-FP) anchors the protein with membrane limiting its diffusion thus providing spatial information on the translating protein (Aakalu et al. 2001). Technical advances have allowed visualization of local translation at increased temporal resolution. The photoconvertible FP, Kaede, fusion techniques have been used to study local translation. Kaede emits green fluorescence until it is cleaved by UV-induction then it emits red fluorescence. Newly synthesized proteins can be detected by green fluorescence after UV treatment (Raab-Graham et al. 2006). One can label newly synthesizing proteins by pulse-chase application of amino acid orthologs and click-chemistry through BONCAT or FUNCAT methods (Dieterich et al. 2006; Dieterich et al. 2010). These methods allow both purification of newly synthesized proteins and in situ visualization. Puromycin labeling-based methods incorporate Puromycin, an analog of aa-tRNA, into newly synthesized proteins to inhibit amino acid polymerization. Thus, puromycin conjugates can be used to visualize de novo protein synthesis (Starck et al. 2004). Similarly, puromycylation of a specific protein can be detected in situ by Puro-PLA (proximity ligation assay) through in situ coincidental detection of anti-Puromycin and anti-POI antibodies (Tom Dieck et al. 2015). Some techniques provide detection power both for mRNAs and newly synthesized proteins allowing bonafide detection of localized translation, although they still require multiple tagging strategies. Translating

RNA Imaging by Coat protein Knock-off (TRICK) methods takes advantage of phage coat proteins and its binding sequences, i.e., PP7 and MS2, to visualize the displacement of PP7-GFP by translating ribosomes and co-detection of its cognate mRNA by MS2-RFP (Halstead et al. 2015). Single-molecule Imaging of NAscent PeptideS (SINAPS) uses reporter constructs containing multiple SunTag epitope for visualization of translation and MCP binding sequence for mRNA detection (Wu et al. 2016). Similarly, nascent chain tracking (NCT) uses multiple FLAG epitope tags and antibody-based detection of translating proteins and MCP system to visualize mRNAs (Morisaki et al. 2016).



6. Distinct Functions for Locally Synthesized Proteins

A straightforward way to envision the role of mRNA subcellular localization is that it can provide rapid translation to generate a protein required in a spatial region of interest at a precise time. For instance, translation of β-actin mRNAs localized into axonal growth cones increases the protein levels in the same growth cone region to support axon guidance (Yao et al. 2006). An emerging theme, however, is that protein synthesis in

specific subcellular regions exposes the newly synthesized proteins to distinct proteomic repertories. In the case of the *Npas4* gene, NPAS4 (neuronal PAS domain protein 4) proteins synthesized from the dendrite-localized and soma-localized transcripts interact with distinct protein partners (**Figure 4A**). Soma-localized *Npas4* transcripts undergo de novo translation upon neuron activation through L-type voltage-gated calcium channels (L-VGCCs) in the Stratum pyramidale layer (soma region) of hippocampus. These soma-originated NPAS4 proteins associate with ARNT2 (aryl hydrocarbon receptor nuclear translocator 2) and act as nuclear transcription factors. This mechanism is distinct from the heterodimerization of NPAS4 triggered by excitatory postsynaptic potential (EPSP) induction. In the latter case, dendritically localized *Npas4* and *Arnt1* mRNAs undergo local translation, form heterodimers, and are translocated together to the nucleus to bind transcription promoter or enhancer regions (Brigidi et al. 2019). This example not only distinguishes differential means of inducing NPAS4 expression upon distinct neuronal signals but also emphasizes the importance of compartmentalization of mRNAs and its translation to modulate the protein–protein interactome.

Alternative 3′ UTRs can recruit RBPs to increase the likelihood that an RBP interacts with the nascent protein as it is translated, thus impacting protein–protein interactions. For instance, the RBP HuR (ELAV-like protein 1) interacts with the long 3′ UTR of CD47 (leukocyte surface antigen CD47) to act as an intermediate scaffold mediating SET-dependent localization of CD47 proteins to the plasma membrane (Berkovits and Mayr 2015). Another example is the long 3′ UTR isoform of BIRC3 (baculoviral IAP repeat containing 3) that, unlike the short isoform, recruits a unique set of proteins. The long 3′ UTR of BIRC3 recruits a protein complex composed by IQGAP1 (IQ motif containing GTPase-activating protein 1), RALA (Ras-related protein), STAU1 (Staufen 1), and HuR, to promote their interaction with nascent BIRC3 proteins (Lee and

Mayr 2019). Interestingly, the 3′ UTR isoforms do not differently drive mRNA or protein localization, suggesting that the long 3′ UTR-dependent function of BIRC3 is not localization or local translation. Instead, the long 3′ UTR of BIRC3 serves as a scaffold RNA to enhance protein–protein interactions (**Figure 4B**). These studies were performed in non-neuronal cell lines—in the future it will be interesting to see if analogous examples emerge from studies of alternative 3′ UTR isoforms expressed in neurons.

7. Beyond Localization and Translational Control

In addition to the conventional roles for 3′ UTRs in regulating mRNA stability, localization, and translational control, several new functions of 3′ UTRs have emerged in recent years.

7.1. Coding-Independent Functions of 3′ UTRs

Evidence is emerging that alternative 3′ UTR isoforms have functions in *trans*, independent from influencing the protein-coding functions of the gene they are transcribed from. *Ube3a1* is an essential gene implicated in Angelman syndrome and Autism spectrum disorders. An activity induced *Ube3a1* transcript is alternatively polyadenylated at intron 11 to encode an isoform lacking the catalytic domain. This alternative 3′ UTR isoform is synaptodendritically localized. Although knockdown of *Ube3a1* impaired dendritic complexity and spine morphology, this was not due to the protein-coding capacity of the transcript. *Ube3a1* RNA containing a frameshift mutation or *Ube3a1* 3′ UTR alone was able to fully rescue the dendritic phenotypes acting as endogenous miRNA sponge (Valluy et al. 2015; **Figure 4C**).

Tp53inp2 is a transcript highly enriched in axons of sympathetic neurons that has a protein coding-independent function. *Tp53inp2* is able to translate into proteins in other tissues, but it is translationally suppressed in neurons due to its long 3′ UTR, suggesting

its unique role as a non-coding mRNA in sympathetic neurons. Despite the lack of translation, *Tp53inp2* mRNA is involved in TrkA (tropomyosin-related kinase receptor A) receptor internalization and downstream signaling, and phenotypes can be rescued by non-translatable *Tp53inp2* transcript (Crerar et al. 2019; **Figure 4D**).

The precise molecular mechanisms underlying non-coding roles for alternative 3′ UTR isoforms remain to be illuminated. More examples should continue to emerge given the new feasibility of isoform specific deletions using CRISPR-Cas9 technologies, and isoform specific knockdown approaches such as short hairpin RNAs (shRNAs), antisense oligonucleotides, and CRISPR-Cas13 (Cox et al. 2017).

7.2. Coordination of Alternative Splicing and APA

Potential crosstalk between alternative splicing and APA has been suggested in the past since several RBPs have documented roles in regulating both of these pre-mRNA processing events (Hilgers et al. 2012; Licatalosi et al. 2008; Soller and White 2003). An example of a 3′ UTR isoform being associated with a particular alternative splicing event of the same gene has been recently discovered. In *Drosophila*, the expression of the long 3′ UTR isoform of the *Dscam1* gene was found to be coupled to the skipping of an upstream exon. The neuronal RBP Elav was found to induce both the expression of the long 3′ UTR and skipping of an upstream exon. Long-read sequencing on the Oxford Nanopore MinION platform demonstrated the tight connection between the alternative splicing and APA events—long 3′ UTR mRNAs in adult heads were found to always skip the upstream exon. Moreover, shRNA-mediated knock-down of exon skipping transcripts abolished the long 3′ UTR isoform of *Dscam1* and vice-versa. This connectivity of an exon skipping event and an alternative long 3′ UTR could be explained simply by the co-regulation by Elav. However, it was found that Elav could only exert its influence on the

exon skipping event when the *Dscam1* long 3′ UTR was present (Zhang et al. 2019; **Figure 4E**).

What is the mechanism of 3′ UTR-mediated regulation of exon skipping? It is possible that the 3′ UTR acts as a splicing factor delivery system. In such a model, intramolecular interactions between upstream introns and the 3′ UTR might occur, allowing RBPs bound to the 3′UTR to interact with upstream splice sites. Or perhaps 3′ UTR sequences bind to upstream intronic sequences that are important for splicing, and block association of spliceosome components? Future work is needed to determine if coupling of 3′ UTR choice to upstream alternative splicing events is widespread and whether it occurs in human neurons. To date, widespread coupling of 3′ UTR selection to alternative exons have not been widely identified because of the short (<150 nt) read length of conventional RNA-Seq approaches. The emergence of new long-read sequencing platforms and library preparation strategies could provide the opportunity to uncover global coordination of APA and alternative splicing (Volden et al. 2018; Tilgner et al. 2018).

7.3. Isolated 3′ UTR Fragments

The presence of 3′ UTR fragments that are separated from the protein-coding regions of the mRNA has been identified in a handful of studies over the past decade (**Figure 4F**). Capped analysis of gene expression (CAGE)-seq and full-length cDNA-seq identified several 3′ UTR-derived RNAs which are not directly originated from transcription by RNA Pol II promoters (Mercer et al. 2011). Additionally, in situ imaging approaches have shown a divergence in 3′ UTR to CDS expression ratios across tissues suggesting cell- and tissue-specific expression of isolated 3′ UTRs (Mercer et al. 2011; Kocabas et al. 2015). A potential role in regulating host gene translation has been suggested based on

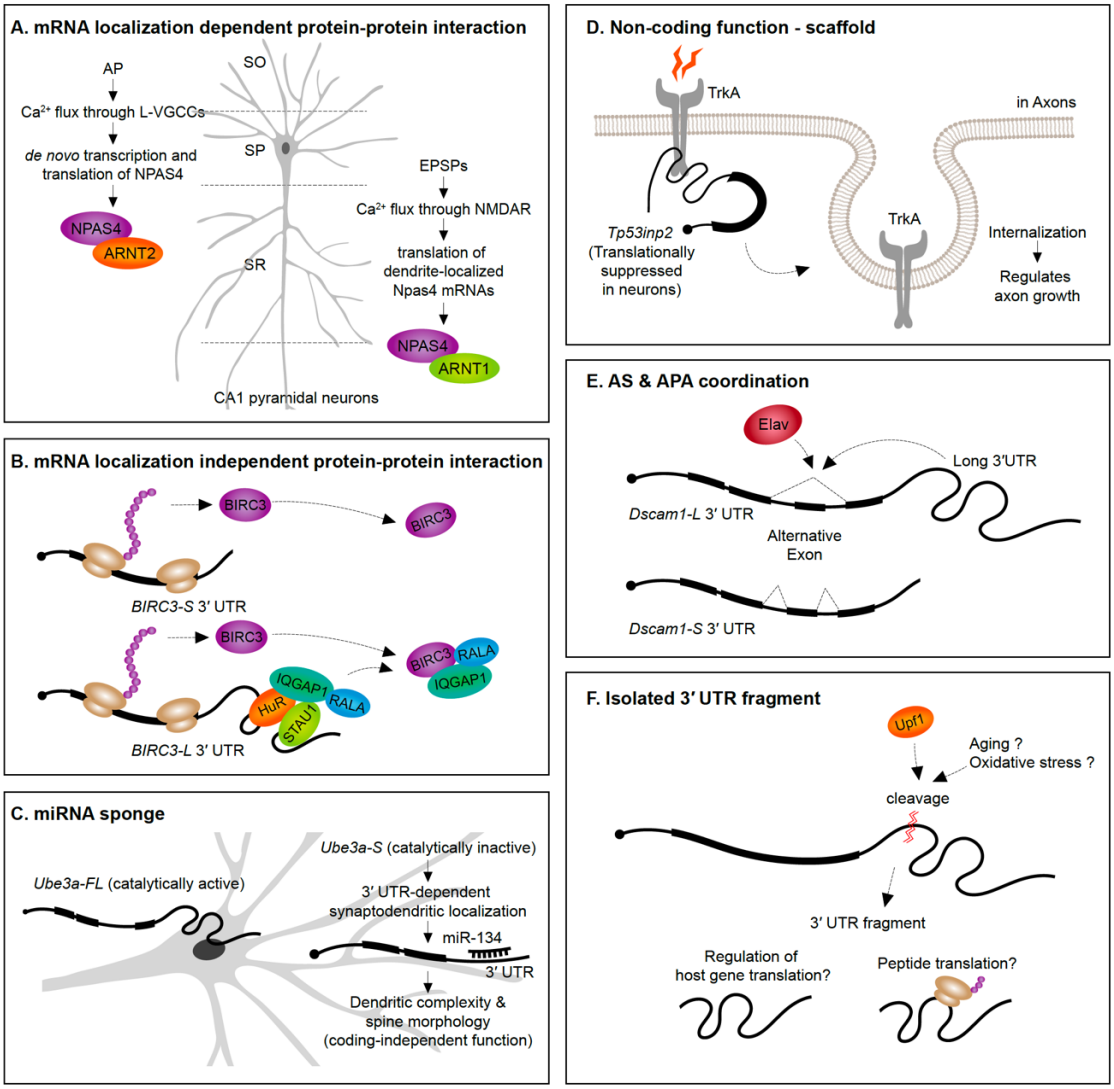


Figure 4. Newly emerging roles for 3′ UTRs

(A,B) mRNA 3′ UTRs can serve as scaffolds for the interactome of their host proteins. (C) 3′ UTRs can serve a non-coding role as a sponge for miRNAs. (D) Translationally inactive 3′ UTRs mRNA isoforms can impact internalization of receptor proteins and influence axon growth. (E) 3′ UTR sequences can act as mediators of alternative splicing on the same pre-mRNA. (F) Isolated 3′ UTR fragments cleaved from mRNAs might regulate host gene translation and/or encode peptides. AP, action potential; L-VGCC, L-type voltage-gated calcium channel; EPSP, excitatory postsynaptic potential; NMDAR, NMDA receptor; SO, stratum oriens; SP, stratum pyramidale; SR, stratum radiatum; CA1, cornus ammonis 1.

the abundance of 3′ UTR fragments being inversely correlated with protein levels from the same gene (Kocabas et al. 2015).

How are isolated 3′ UTR fragments generated? It has been suggested that some are cleaved post-transcriptionally in the cytoplasm (Malka et al. 2017). A recent study suggested that post-transcriptional cleavage of *Impa1* transcript occurs at an internal poly(A) site to generate 3′ UTR fragments in the cytoplasm (Andreassi et al. 2019). The generation of the 3′ UTR fragments appeared to occur in axons of rat cervical ganglion neurons, but not in cell bodies (Andreassi et al. 2019). In this case, 3′ UTR cleavage possibly involves the RNA surveillance machinery, as evidenced by its reliance on UPF1 (Andreassi et al. 2019). Some 3′ UTR fragments are associated with ribosomes (Kocabas et al. 2015; Sudmant et al. 2018) and some have the potential to encode peptides (Sudmant et al. 2018). Clearly, further work is needed to identify the precise mechanism(s) responsible for the generation of isolated 3′ UTR fragments. What are the molecular features of the 5′ and 3′ ends of cleaved 3′ UTRs? New library preparation strategies are needed to specifically sequence these isolated fragments. The functions of isolated fragments in the nervous system remain completely unknown, but with insight into their biogenesis and sequence features, genetic approaches to investigate their functions will surely emerge.

8. 3′ UTR-Associated Neurological Disorders and Behavior

Polymorphisms in 3′ UTRs are associated with neurological disorders and behavior. For example, the variable number of tandem repeat (VNTR) polymorphism in the 3′ UTR of the dopamine transporter (*DAT1*) has been associated with eating and substance use disorders, and symptom severity in Tourette syndrome (Shinohara et al. 2004; Tarnok et al. 2007). A single nucleotide polymorphism (SNP; rs2304297) in the 3′

UTR of *CHRNA6* gene, which encodes for a nicotinic acetylcholine receptor (nAChR) subunit, has been associated with adolescent cigarette smoking, smoking cessation, and striatum volume (Fletcher 2012; Lotfipour et al. 2010). A SNP in the 3′ UTR of the *SNCA* gene, which encodes α-synuclein, has also been associated with Parkinson's disease. Heightened levels of α-synuclein is pathogenic, thus whether the polymorphisms in the 3′ UTR of *SNCA* is implicated in regulation of *SNCA* expression is of great interest (Barrie et al. 2018; Marchese et al. 2017; Toffoli et al. 2017; Sotiriou et al. 2009). Moreover, there might be a role for 3′ UTRs variants in AD. In this regard, 3′ UTR polymorphisms in the *APP* gene have been shown to target miR-147 and miR-20a (Delay et al. 2011). Although the exact molecular mechanism by which this SNP has an effect on these behaviors is largely unknown, it is possible that these polymorphisms impact the expression of the host gene by disrupting regulatory elements.

Impaired RNA granule assembly is associated with long repeat expansion disorders such as myotonic dystrophy, amyotrophic lateral sclerosis, and spinocerebellar ataxia (Shukla and Parker 2016; Jain and Vale 2017). Repeat expansion in 3′ UTRs can sequester RBPs into granules affecting their normal functions. For instance, in the debilitating neuromuscular disorder Myotonic dystrophy type I, a repeat expansion in the 3′ UTR of the *DMPK* gene sequesters the splicing factor MBNL1 into nuclear RNA foci to prevent from functioning in normal pre-mRNA processing (Miller 2000; Mankodi 2001).

In Huntington's disease (HD), expression of many alternative 3′ UTR isoforms, among others the huntingtin (*HTT*) gene, are altered. Profiling of alternative 3′ UTR isoforms has shown that *HTT* long 3′ UTR isoform specifically is altered in multiple tissues of HD patients (Romo et al. 2017). The exact mechanism whereby the altered abundance of the long 3′ UTR isoform of *HTT* contributes to HD pathology remains elusive.

Strong evidence for the contribution of 3′ UTR mutations in neurological disease susceptibility and pathogenesis is still limited. Increased whole genome sequencing depth and sensitivity, increased numbers of sequenced individuals, and improved 3′ UTR annotations might lead to new discoveries of medically relevant 3′ UTR SNPs. The use of massively parallel reporter assays to assess the role of 3′ UTR SNPs on translational control could be a useful approach to uncover whether particular 3′ UTR SNPs have functional relevance, such as has been performed to identify human genetic variants that influence splicing (Cheung et al. 2019).

9. Conclusions

The 3′ UTR imparts post-transcriptional regulation by impacting mRNA stability, subcellular localization, and translational control. APA controls the susceptibility of transcripts to this regulation, and the functional roles for alternative 3′ UTR isoforms have been discovered for only a handful of genes. New emerging mechanisms suggest that 3′ UTRs have roles in modulating the protein interactome of newly translated proteins and coordinating alternative splicing. Given the limitations of some neuronal-like immortalized cultured systems, we call for an increased emphasis on performing 3′ UTR functional studies in primary neuronal culture and whole organisms.

10. References

1. Conaco, C.; Otto, S.; Han, J.J.; Mandel, G. Reciprocal actions of REST and a microRNA promote neuronal identity. *Proc. Natl. Acad. Sci.* **2006**, *103*, 2422–2427.
2. Lim, L.P.; Lau, N.C.; Garrett-Engele, P.; Grimson, A.; Schelter, J.M.; Castle, J.; Bartel, D.P.; Linsley, P.S.; Johnson, J.M. Microarray analysis shows that some microRNAs downregulate large numbers of-target mRNAs. *Nature* **2005**, *433*, 769–773.
3. Rajman, M.; Schratt, G. MicroRNAs in neural development: from master regulators to fine-tuners. *Development* **2017**, *144*, 2310–2322.
4. Chiu, H.; Alqadah, A.; Chang, C. The role of microRNAs in regulating neuronal connectivity. *Front. Cell. Neurosci.* **2014**, *7*, 283.
5. Zampa, F.; Bicker, S.; Schratt, G. Activity-dependent pre-miR-134 dendritic localization is required for hippocampal neuron dendritogenesis. *Front. Mol. Neurosci.* **2018**, *11*, 171.
6. Dajas-Bailador, F.; Bonev, B.; Garcez, P.; Stanley, P.; Guillemot, F.; Papalopulu, N. microRNA-9 regulates axon extension and branching by targeting Map1b in mouse cortical neurons. *Nat. Neurosci.* **2012**, *15*, 697–699.
7. Franke, K.; Otto, W.; Johannes, S.; Baumgart, J.; Nitsch, R.; Schumacher, S. miR-124-regulated RhoG reduces neuronal process complexity via ELMO/Dock180/Rac1 and Cdc42 signalling. *EMBO J.* **2012**, *31*, 2908–2921.
8. Baudet, M.; Zivraj, K.H.; Abreu-Goodger, C.; Muldal, A.; Armisen, J.; Blenkiron, C.; Goldstein, L.D.; Miska, E.A.; Holt, C.E. miR-124 acts through CoREST to control onset of Sema3A sensitivity in navigating retinal growth cones. *Nat. Neurosci.* **2012**, *15*, 29–38.
9. Magill, S.T.; Cambronne, X.A.; Luikart, B.W.; Lioy, D.T.; Leighton, B.H.; Westbrook, G.L.; Mandel, G.; Goodman, R.H. microRNA-132 regulates dendritic growth and arborization of newborn neurons in the adult hippocampus. *Proc. Natl. Acad. Sci.* **2010**, *107*, 20382–7.
10. Wang, W.; Kwon, E.J.; Tsai, L.H. MicroRNAs in learning, memory, and neurological diseases. *Learn. Mem.* 2012, *19*, 359–368.
11. Kiltschewskij, D.; Cairns, M.J. Temporospatial guidance of activity-dependent gene expression by microRNA: Mechanisms and functional implications for neural plasticity. *Nucleic Acids Res.* **2019**, *47*, 533–545.

12. Hu, Z.; Li, Z. miRNAs in synapse development and synaptic plasticity. *Curr. Opin. Neurobiol.* 2017, *45*, 24–31.
13. Xue, Y.; Zhang, Y. Emerging roles for microRNA in the regulation of Drosophila circadian clock. *BMC Neurosci.* **2018**, *19*, 1.
14. Hébert, S.S.; De Strooper, B. Alterations of the microRNA network cause neurodegenerative disease. *Trends Neurosci.* **2009**, *32*, 199–206.
15. Leggio, L.; Vivarelli, S.; L'Episcopo, F.; Tirolo, C.; Caniglia, S.; Testa, N.; Marchetti, B.; Iraci, N. microRNAs in Parkinson's Disease: From Pathogenesis to Novel Diagnostic and Therapeutic Approaches. *Int. J. Mol. Sci.* **2017**, *18*, 2698.
16. Goh, S.Y.; Chao, Y.X.; Dheen, S.T.; Tan, E.-K.; Tay, S.S.-W. Role of MicroRNAs in Parkinson's Disease. *Int. J. Mol. Sci.* **2019**, *20*, 5649.
17. Im, H.I.; Kenny, P.J. MicroRNAs in neuronal function and dysfunction. *Trends Neurosci.* 2012, *35*, 325–334.
18. Issler, O.; Chen, A. Determining the role of microRNAs in psychiatric disorders. *Nat. Rev. Neurosci.* **2015**, *16*, 201–212.
19. Wu, Y.E.; Parikshak, N.N.; Belgard, T.G.; Geschwind, D.H. Genome-wide, integrative analysis implicates microRNA dysregulation in autism spectrum disorder. *Nat. Neurosci.* **2016**, *19*, 1463–1476.
20. Dini Modigliani, S.; Morlando, M.; Errichelli, L.; Sabatelli, M.; Bozzoni, I. An ALS-associated mutation in the FUS 3′-UTR disrupts a microRNA-FUS regulatory circuitry. *Nat. Commun.* **2014**, *5*, 1–7.
21. Delay, C.; Calon, F.; Mathews, P.; Hebert, S.S. Alzheimer-specific variants in the 3′UTR of Amyloid precursor protein affect microRNA function. *Mol Neurodegener* **2011**, *6*, 70.
22. Vaishnavi, V.; Manikandan, M.; Munirajan, A.K. Mining the 3′UTR of autism-implicated genes for SNPs perturbing microRNA regulation. *Genomics Proteomics Bioinforma.* **2014**, *12*, 92–104.
23. Li, J.; Yang, X.; Qi, Z.; Sang, Y.; Liu, Y.; Xu, B.; Liu, W.; Xu, Z.; Deng, Y. The role of mRNA m6A methylation in the nervous system. *Cell Biosci.* 2019, *9*, 66.
24. Meyer, K.D.; Saletore, Y.; Zumbo, P.; Elemento, O.; Mason, C.E.; Jaffrey, S.R. Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. *Cell* **2012**, *149*, 1635–1646.

25. Ke, S.; Alemu, E.A.; Mertens, C.; Gantman, E.C.; Fak, J.J.; Mele, A.; Haripal, B.; Zucker-Scharff, I.; Moore, M.J.; Park, C.Y.; et al. A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev.* **2015**, *29*, 2037–2053.
26. Widagdo, J.; Anggono, V. The m6A-epitranscriptomic signature in neurobiology: from neurodevelopment to brain plasticity. *J. Neurochem.* **2018**, *147*, 137–152.
27. Du, K.; Zhang, L.; Lee, T.; Sun, T. m6A RNA Methylation Controls Neural Development and Is Involved in Human Diseases. *Mol. Neurobiol.* 2019, *56*, 1596–1606.
28. Shi, H.; Zhang, X.; Weng, Y.L.; Lu, Z.; Liu, Y.; Lu, Z.; Li, J.; Hao, P.; Zhang, Y.; Zhang, F.; et al. m6A facilitates hippocampus-dependent learning and memory through YTHDF1. *Nature* **2018**, *563*, 249–253.
29. Caput, D.; Beutler, B.; Hartog, K.; Thayer, R.; Brown-Shimer, S.; Cerami, A. Identification of a common nucleotide sequence in the 3′-untranslated region of mRNA molecules specifying inflammatory mediators. *Proc. Natl. Acad. Sci.* **1986**, *83*, 1670–1674.
30. Bird, C.W.; Gardiner, A.S.; Bolognani, F.; Tanner, D.C.; Chen, C.Y.; Lin, W.J.; Yoo, S.; Twiss, J.L.; Perrone-Bizzozero, N. KSRP modulation of GAP-43 mRNA stability restricts axonal outgrowth in embryonic hippocampal neurons. *PLoS One* **2013**, *8*, e79255.
31. Bolognani, F.; Tanner, D.C.; Merhege, M.; Deschênes-Furry, J.; Jasmin, B.; Perrone-Bizzozero, N.I. In vivo post-transcriptional regulation of GAP-43 mRNA by overexpression of the RNA-binding protein HuD. *J. Neurochem.* **2006**, *96*, 790–801.
32. Jeffery, W.R.; Tomlinson, C.R.; Brodeur, R.D. Localization of actin messenger RNA during early ascidian development. *Dev. Biol.* **1983**, *99*, 408–417.
33. Garner, C.C.; Tucker, R.P.; Matus, A. Selective localization of messenger RNA for cytoskeletal protein MAP2 in dendrites. *Nature* **1988**, *336*, 674–677.
34. Burgin, K.E.; Neal Waxham, M.; Rickling, S.; Westgate, S.A.; Mobley, W.C.; Kelly, P.T. In situ hybridization histochemistry of Ca2+/calmodulin-dependent protein kinase in developing rat brain. *J. Neurosci.* **1990**, *10*, 1788–1798.
35. Bassell, G.J.; Zhang, H.; Byrd, A.L.; Femino, A.M.; Singer, R.H.; Taneja, K.L.; Lifshitz, L.M.; Herman, I.M.; Kosik, K.S. Sorting of β-actin mRNA and protein to neurites and growth cones in culture. *J. Neurosci.* **1998**, *18*, 251–265.

36. Lyford, G.L.; Yamagata, K.; Kaufmann, W.E.; Barnes, C.A.; Sanders, L.K.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Lanahan, A.A.; Worley, P.F. Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-associated protein that is enriched in neuronal dendrites. *Neuron* **1995**, *14*, 433–445.
37. Taylor, A.M.; Berchtold, N.C.; Perreau, V.M.; Tu, C.H.; Li Jeon, N.; Cotman, C.W. Axonal mRNA in uninjured and regenerating cortical mammalian axons. *J. Neurosci.* **2009**, *29*, 4697–4707.
38. Gumy, L.F.; Yeo, G.S.H.; Tung, Y.-C.L.; Zivraj, K.H.; Willis, D.; Coppola, G.; Lam, B.Y.H.; Twiss, J.L.; Holt, C.E.; Fawcett, J.W. Transcriptome analysis of embryonic and adult sensory axons reveals changes in mRNA repertoire localization. *RNA* **2011**, *17*, 85–98.
39. 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. *Dev. Neurobiol.* **2014**, *74*, 365–381.
40. Zivraj, K.H.; Tung, Y.C.; Piper, M.; Gumy, L.; Fawcett, J.W.; Yeo, G.S.; Holt, C.E. Subcellular profiling reveals distinct and developmentally regulated repertoire of growth cone mRNAs. *J Neurosci* **2010**, *30*, 15464–15478.
41. Taliaferro, J.M.; Vidaki, M.; Oliveira, R.; Olson, S.; Zhan, L.; Saxena, T.; Wang, E.T.; Graveley, B.R.; Gertler, F.B.; Swanson, M.S.; et al. Distal Alternative Last Exons Localize mRNAs to Neural Projections. *Mol. Cell* **2016**, *61*, 821–833.
42. Tushev, G.; Glock, C.; Heumüller, M.; Biever, A.; Jovanovic, M.; Schuman, E.M. Alternative 3′ UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments. *Neuron* **2018**, *98*, 495–511.
43. Middleton, S.A.; Eberwine, J.; Kim, J. Comprehensive catalog of dendritically localized mRNA isoforms from sub-cellular sequencing of single mouse neurons. *BMC Biol.* **2019**, *17*, 5.
44. Aronov, S.; Aranda, G.; Behar, L.; Ginzburg, I. Axonal tau mRNA localization coincides with tau protein in living neuronal cells and depends on axonal targeting signal. *J. Neurosci.* **2001**, *21*, 6577–6587.
45. Atlas, R.; Behar, L.; Elliott, E.; Ginzburg, I. The insulin-like growth factor mRNA binding-protein IMP-1 and the Ras-regulatory protein G3BP associate with tau mRNA and HuD protein in differentiated P19 neuronal cells. *J. Neurochem.* **2004**, *89*, 613–626.

46. Smith, C.L.; Afroz, R.; Bassell, G.J.; Furneaux, H.M.; Perrone-Bizzozero, N.I.; Burry, R.W. GAP-43 mRNA in growth cones is associated with HuD and ribosomes. *J. Neurobiol.* **2004**, *61*, 222–235.
47. Boissonneault, V.; Plante, I.; Rivest, S.; Provost, P. MicroRNA-298 and microRNA-328 regulate expression of mouse β-amyloid precursor protein-converting enzyme 1. *J. Biol. Chem.* **2009**, *284*, 1971–1981.
48. Lee, S.T.; Chu, K.; Jung, K.H.; Kim, J.H.; Huh, J.Y.; Yoon, H.; Park, D.K.; Lim, J.Y.; Kim, J.M.; Jeon, D.; et al. MiR-206 regulates brain-derived neurotrophic factor in Alzheimer disease model. *Ann. Neurol.* **2012**, *72*, 269–277.
49. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360.
50. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079.
51. Thorvaldsdóttir, H.; Robinson, J.T.; Mesirov, J.P. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **2013**, *14*, 178–192.
52. Willis, D.E.; Xu, M.; Donnelly, C.J.; Tep, C.; Kendall, M.; Erenstheyn, M.; English, A.W.; Schanen, N.C.; Kirn-Safran, C.B.; Yoon, S.O.; et al. Axonal Localization of Transgene mRNA in Mature PNS and CNS Neurons. *J. Neurosci.* **2011**, *31*, 14481–14487.
53. Andreassi, C.; Riccio, A. To localize or not to localize: mRNA fate is in 3′UTR ends. *Trends Cell Biol.* **2009**, *19*, 465–474.
54. Kislauskis, E.H.; Zhu, X.; Singer, R.H. Sequences responsible for intracellular localization of β-actin messenger RNA also affect cell phenotype. *J. Cell Biol.* **1994**, *127*, 441–451.
55. Ross, A.F.; Oleynikov, Y.; Kislauskis, E.H.; Taneja, K.L.; Singer, R.H. Characterization of a beta-actin mRNA zipcode-binding protein. *Mol. Cell. Biol.* **1997**, *17*, 2158–2165.
56. Tiruchinapalli, D.M.; Oleynikov, Y.; Kelič, S.; Shenoy, S.M.; Hartley, A.; Stanton, P.K.; Singer, R.H.; Bassell, G.J. Activity-dependent trafficking and dynamic localization of zipcode binding protein 1 and β-actin mRNA in dendrites and spines of hippocampal neurons. *J. Neurosci.* **2003**, *23*, 3251–3261.

57. Vuppalanchi, D.; Coleman, J.; Yoo, S.; Merianda, T.T.; Yadhati, A.G.; Hossain, J.; Blesch, A.; Willis, D.E.; Twiss, J.L. Conserved 3′-untranslated region sequences direct subcellular localization of chaperone protein mRNAs in neurons. *J. Biol. Chem.* **2010**, *285*, 18025–18038.
58. Mayford, M.; Baranes, D.; Podsypanina, K.; Kandel, E.R. The 3′-untranslated region of CaMKIIα is a cis-acting signal for the localization and translation of mRNA in dendrites. *Proc. Natl. Acad. Sci.* **1996**, *93*, 13250–13255.
59. Perry, R.B.T.; Doron-Mandel, E.; Iavnilovitch, E.; Rishal, I.; Dagan, S.Y.; Tsoory, M.; Coppola, G.; McDonald, M.K.; Gomes, C.; Geschwind, D.H.; et al. Subcellular Knockout of Importin β1 Perturbs Axonal Retrograde Signaling. *Neuron* **2012**, *75*, 294–305.
60. Yoo, S.; Kim, H.H.; Kim, P.; Donnelly, C.J.; Kalinski, A.L.; Vuppalanchi, D.; Park, M.; Lee, S.J.; Merianda, T.T.; Perrone-Bizzozero, N.I.; et al. A HuD-ZBP1 ribonucleoprotein complex localizes GAP-43 mRNA into axons through its 3′ untranslated region AU-rich regulatory element. *J. Neurochem.* **2013**, *126*, 792–804.
61. Zhang, H.L.; Eom, T.; Oleynikov, Y.; Shenoy, S.M.; Liebelt, D.A.; Dictenberg, J.B.; Singer, R.H.; Bassell, G.J. Neurotrophin-Induced Transport of a β-Actin mRNP Complex Increases β-Actin Levels and Stimulates Growth Cone Motility. *Neuron* **2001**, *31*, 261–275.
62. Bauer, K.E.; Segura, I.; Gaspar, I.; Scheuss, V.; Illig, C.; Ammer, G.; Hutten, S.; Basyuk, E.; Fernández-Moya, S.M.; Ehses, J.; et al. Live cell imaging reveals 3′-UTR dependent mRNA sorting to synapses. *Nat. Commun.* **2019**, *10*, 3178.
63. Miller, S.; Yasuda, M.; Coats, J.K.; Jones, Y.; Martone, M.E.; Mayford, M. Disruption of Dendritic Translation of CaMKIIα Impairs Stabilization of Synaptic Plasticity and Memory Consolidation. *Neuron* **2002**, *36*, 507–519.
64. Terenzio, M.; Koley, S.; Samra, N.; Rishal, I.; Zhao, Q.; Sahoo, P.K.; Urisman, A.; Marvaldi, L.; Oses-Prieto, J.A.; Forester, C.; et al. Locally translated mTOR controls axonal local translation in nerve injury. *Science (80-.).* **2018**, *359*, 1416–1421.
65. Patel, V.L.; Mitra, S.; Harris, R.; Buxbaum, A.R.; Lionnet, T.; Brenowitz, M.; Girvin, M.; Levy, M.; Almo, S.C.; Singer, R.H.; et al. Spatial arrangement of an RNA zipcode identifies mRNAs under post-transcriptional control. *Genes Dev.* **2012**, *26*, 43–53.

66. Subramanian, M.; Rage, F.; Tabet, R.; Flatter, E.; Mandel, J.L.; Moine, H. G-quadruplex RNA structure as a signal for neurite mRNA targeting. *EMBO Rep.* **2011**, *12*, 697–704.
67. Bevilacqua, P.C.; Ritchey, L.E.; Su, Z.; Assmann, S.M. Genome-Wide Analysis of RNA Secondary Structure. *Annu. Rev. Genet.* **2016**, *50*, 235–266.
68. Brigidi, G.S.; Hayes, M.G.B.; Delos Santos, N.P.; Hartzell, A.L.; Texari, L.; Lin, P.A.; Bartlett, A.; Ecker, J.R.; Benner, C.; Heinz, S.; et al. Genomic Decoding of Neuronal Depolarization by Stimulus-Specific NPAS4 Heterodimers. *Cell* **2019**, *179*, 373-391.e27.
69. Merianda, T.T.; Gomes, C.; Yoo, S.; Vuppalanchi, D.; Twiss, J.L. Axonal Localization of Neuritin/CPG15 mRNA in Neuronal Populations through Distinct 5′ and 3′ UTR Elements. *J. Neurosci.* **2013**, *33*, 13735–13742.
70. Meer, E.J.; Wang, D.O.; Kim, S.; Barr, I.; Guo, F.; Martin, K.C. Identification of a cis-acting element that localizes mRNA to synapses. *Proc. Natl. Acad. Sci.* **2012**, *109*, 4639–4644.
71. Raj, A.; van den Bogaard, P.; Rifkin, S.A.; van Oudenaarden, A.; Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **2008**, *5*, 877–879.
72. Larsson, C.; Grundberg, I.; Söderberg, O.; Nilsson, M. In situ detection and genotyping of individual mRNA molecules. *Nat. Methods* **2010**, *7*, 395–397.
73. Wang, F.; Flanagan, J.; Su, N.; Wang, L.C.; Bui, S.; Nielson, A.; Wu, X.; Vo, H.T.; Ma, X.J.; Luo, Y. RNAscope: A novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J. Mol. Diagnostics* **2012**, *14*, 22–29.
74. Kishi, J.Y.; Lapan, S.W.; Beliveau, B.J.; West, E.R.; Zhu, A.; Sasaki, H.M.; Saka, S.K.; Wang, Y.; Cepko, C.L.; Yin, P. SABER amplifies FISH: enhanced multiplexed imaging of RNA and DNA in cells and tissues. *Nat. Methods* **2019**, *16*, 533–544.
75. Daigle, N.; Ellenberg, J. λN-GFP: An RNA reporter system for live-cell imaging. *Nat. Methods* **2007**, *4*, 633–636.
76. Fusco, D.; Accornero, N.; Lavoie, B.; Shenoy, S.M.; Blanchard, J.M.; Singer, R.H.; Bertrand, E. Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr. Biol.* **2003**, *13*, 161–167.

77. Nelles, D.A.; Fang, M.Y.; O'Connell, M.R.; Xu, J.L.; Markmiller, S.J.; Doudna, J.A.; Yeo, G.W. Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell* **2016**, *165*, 488–496.
78. Chen, X.; Zhang, D.; Su, N.; Bao, B.; Xie, X.; Zuo, F.; Yang, L.; Wang, H.; Jiang, L.; Lin, Q.; et al. Visualizing RNA dynamics in live cells with bright and stable fluorescent RNAs. *Nat. Biotechnol.* 2019, *37*, 1287–1293.
79. Litman, P.; Barg, J.; Ginzburg, I. Microtubules are involved in the localization of tau mRNA in primary neuronal cell cultures. *Neuron* **1994**, *13*, 1463–1474.
80. Gagnon, J.A.; Mowry, K.L. Molecular motors: Directing traffic during RNA localization. *Crit. Rev. Biochem. Mol. Biol.* **2011**, *46*, 229–239.
81. Holt, C.E.; Bullock, S.L. Subcellular mRNA localization in animal cells and why it matters. *Science (80-.).* 2009, *326*, 1212–1216.
82. Nalavadi, V.C.; Griffin, L.E.; Picard-Fraser, P.; Swanson, A.M.; Takumi, T.; Bassell, G.J. Regulation of zipcode binding protein 1 transport dynamics in axons by myosin va. *J. Neurosci.* **2012**, *32*, 15133–15141.
83. Urbanska, A.S.; Janusz-Kaminska, A.; Switon, K.; Hawthorne, A.L.; Perycz, M.; Urbanska, M.; Bassell, G.J.; Jaworski, J. ZBP1 phosphorylation at serine 181 regulates its dendritic transport and the development of dendritic trees of hippocampal neurons. *Sci. Rep.* **2017**, *7*, 1876.
84. Medioni, C.; Mowry, K.; Besse, F. Principles and roles of mRNA localization in animal development. *Dev.* 2012, *139*, 3263–3276.
85. Krichevsky, A.M.; Kosik, K.S. Neuronal RNA granules: A link between RNA localization and stimulation-dependent translation. *Neuron* **2001**, *32*, 683–696.
86. Hyman, A.A.; Weber, C.A.; Jülicher, F. Liquid-Liquid Phase Separation in Biology. *Annu. Rev. Cell Dev. Biol.* **2014**, *30*, 39–58.
87. Hyman, A.A.; Brangwynne, C.P. Beyond Stereospecificity: Liquids and Mesoscale Organization of Cytoplasm. *Dev. Cell.* 2011, *21*, 14–16.
88. Weber, S.C.; Brangwynne, C.P. Getting RNA and protein in phase. *Cell* 2012, *149*, 1188–1191.
89. Franzmann, T.M.; Alberti, S. Prion-like low-complexity sequences: Key regulators of protein solubility and phase behavior. *J. Biol. Chem.* 2019, *294*, 7128–7136.
90. Martin, E.W.; Mittag, T. Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions. *Biochemistry* 2018, *57*, 2478–2487.

91. Han, T.W.; Kato, M.; Xie, S.; Wu, L.C.; Mirzaei, H.; Pei, J.; Chen, M.; Xie, Y.; Allen, J.; Xiao, G.; et al. Cell-free formation of RNA granules: Bound RNAs identify features and components of cellular assemblies. *Cell* **2012**, *149*, 768–779.
92. Kato, M.; Han, T.W.; Xie, S.; Shi, K.; Du, X.; Wu, L.C.; Mirzaei, H.; Goldsmith, E.J.; Longgood, J.; Pei, J.; et al. Cell-free formation of RNA granules: Low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **2012**, *149*, 753–767.
93. Murakami, T.; Qamar, S.; Lin, J.Q.; Schierle, G.S.K.; Rees, E.; Miyashita, A.; Costa, A.R.; Dodd, R.B.; Chan, F.T.S.; Michel, C.H.; et al. ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron* **2015**, *88*, 678–690.
94. Tsang, B.; Arsenault, J.; Vernon, R.M.; Lin, H.; Sonenberg, N.; Wang, L.Y.; Bah, A.; Forman-Kay, J.D. Phosphoregulated FMRP phase separation models activity-dependent translation through bidirectional control of mRNA granule formation. *Proc. Natl. Acad. Sci.* **2019**, *116*, 4218–4227.
95. Liao, Y.C.; Fernandopulle, M.S.; Wang, G.; Choi, H.; Hao, L.; Drerup, C.M.; Patel, R.; Qamar, S.; Nixon-Abell, J.; Shen, Y.; et al. RNA Granules Hitchhike on Lysosomes for Long-Distance Transport, Using Annexin A11 as a Molecular Tether. *Cell* **2019**, *179*, 147-164.e20.
96. Andrusiak, M.G.; Sharifnia, P.; Lyu, X.; Wang, Z.; Dickey, A.M.; Wu, Z.; Chisholm, A.D.; Jin, Y. Inhibition of Axon Regeneration by Liquid-like TIAR-2 Granules. *Neuron* **2019**, *104*, 290-304.e8.
97. Hoque, M.; Ji, Z.; Zheng, D.; Luo, W.; Li, W.; You, B.; Park, J.Y.; Yehia, G.; Tian, B. Analysis of alternative cleavage and polyadenylation by 3′ region extraction and deep sequencing. *Nat. Methods* **2013**, *10*, 133–139.
98. Lianoglou, S.; Garg, V.; Yang, J.L.; Leslie, C.S.; Mayr, C. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev.* **2013**, *27*, 2380–2396.
99. Tian, B.; Hu, J.; Zhang, H.; Lutz, C.S. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.* **2005**, *33*, 201–212.
100. Pelka, G.J.; Watson, C.M.; Christodoulou, J.; Tam, P.P. Distinct expression profiles of Mecp2 transcripts with different lengths of 3′UTR in the brain and visceral organs during mouse development. *Genomics* **2005**, *85*, 441–452.

101. Costessi, L.; Devescovi, G.; Baralle, F.E.; Muro, A.F. Brain-specific promoter and polyadenylation sites of the β-adducin pre-mRNA generate an unusually long 3′-UTR. *Nucleic Acids Res.* **2006**, *34*, 243–253.
102. Zhang, H.; Lee, J.Y.; Tian, B. Biased alternative polyadenylation in human tissues. *Genome Biol* **2005**, *6*, R100.
103. Smibert, P.; Miura, P.; Westholm, J.O.; Shenker, S.; May, G.; Duff, M.O.; Zhang, D.; Eads, B.D.; Carlson, J.; Brown, J.B.; et al. Global Patterns of Tissue-Specific Alternative Polyadenylation in Drosophila. *Cell Rep.* **2012**, *1*, 277–289.
104. Miura, P.; Shenker, S.; Andreu-Agullo, C.; Westholm, J.O.; Lai, E.C. Widespread and extensive lengthening of 3′UTRs in the mammalian brain. *Genome Res.* **2013**, *23*, 812–825.
105. Ha, K.C.H.; Blencowe, B.J.; Morris, Q. QAPA: a new method for the systematic analysis of alternative polyadenylation from RNA-seq data. *Genome Biol.* **2018**, *19*, 45.
106. Shepard, P.J.; Choi, E.A.; Lu, J.; Flanagan, L.A.; Hertel, K.J.; Shi, Y. Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. *RNA* **2011**, *17*, 761–772.
107. Guvenek, A.; Tian, B. Analysis of alternative cleavage and polyadenylation in mature and differentiating neurons using RNA-seq data. *Quant. Biol.* **2018**, *6*, 253–266.
108. Wang, W.; Wei, Z.; Li, H. A change-point model for identifying 3′UTR switching by next-generation RNA sequencing. *Bioinformatics* **2014**, *30*, 2162–2170.
109. Xia, Z.; Donehower, L.A.; Cooper, T.A.; Neilson, J.R.; Wheeler, D.A.; Wagner, E.J.; Li, W. Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3′-UTR landscape across seven tumour types. *Nat. Commun.* **2014**, *5*, 5274.
110. Ye, C.; Long, Y.; Ji, G.; Li, Q.Q.; Wu, X. APAtrap: Identification and quantification of alternative polyadenylation sites from RNA-seq data. *Bioinformatics* **2018**, *34*, 1841–1849.
111. Chan, S.; Choi, E.A.; Shi, Y. Pre-mRNA 3′-end processing complex assembly and function. *Wiley Interdiscip. Rev. RNA* 2011, *2*, 321–335.
112. Takagaki, Y.; Seipelt, R.L.; Peterson, M.L.; Manley, J.L. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell* **1996**, *87*, 941–952.

113. Yao, C.; Choi, E.A.; Weng, L.; Xie, X.; Wan, J.I.; Xing, Y.I.; Moresco, J.J.; Tu, P.G.; Yates, J.R.; Shi, Y. Overlapping and distinct functions of CstF64 and CstF64τ in mammalian mRNA 3′ processing. *RNA* **2013**, *19*, 1781–1790.
114. Shankarling, G.S.; MacDonald, C.C. Polyadenylation site-specific differences in the activity of the neuronal βCstF-64 protein in PC-12 cells. *Gene* **2013**, *529*, 220–227.
115. Martin, G.; Gruber, A.R.; Keller, W.; Zavolan, M. Genome-wide Analysis of Pre-mRNA 3′ End Processing Reveals a Decisive Role of Human Cleavage Factor I in the Regulation of 3′ UTR Length. *Cell Rep.* **2012**, *1*, 753–763.
116. Alcott, C.E.; Yalamanchili, H.K.; Ji, P.; van der Heijden, M.E.; Saltzman, A.; Elrod, N.; Lin, A.; Leng, M.; Bhatt, B.; Hao, S.; et al. Partial loss of CFIm25 causes learning deficits and aberrant neuronal alternative polyadenylation. *Elife* **2020**, *9*, e50895.
117. Fusby, B.; Kim, S.; Erickson, B.; Kim, H.; Peterson, M.L.; Bentley, D.L. Coordination of RNA Polymerase II Pausing and 3′ end processing factor recruitment with alternative polyadenylation. *Mol. Cell. Biol.* **2015**, *36*, MCB.00898-15.
118. Pinto, P.A.B.; Henriques, T.; Freitas, M.O.; Martins, T.; Domingues, R.G.; Wyrzykowska, P.S.; Coelho, P.A.; Carmo, A.M.; Sunkel, C.E.; Proudfoot, N.J.; et al. RNA polymerase II kinetics in polo polyadenylation signal selection. *EMBO J.* **2011**, *30*, 2431–2444.
119. Liu, X.; Freitas, J.; Zheng, D.; Oliveira, M.S.; Hoque, M.; Martins, T.; Henriques, T.; Tian, B.; Moreira, A. Transcription elongation rate has a tissue-specific impact on alternative cleavage and polyadenylation in Drosophila melanogaster. *RNA* **2017**, *23*, 1807–1816.
120. Hilgers, V.; Lemke, S.B.; Levine, M. ELAV mediates 3′ UTR extension in the Drosophila nervous system. *Genes Dev* **2012**, *26*, 2259–2264.
121. Oktaba, K.; Zhang, W.; Lotz, T.S.; Jun, D.J.; Lemke, S.B.; Ng, S.P.; Esposito, E.; Levine, M.; Hilgers, V. ELAV links paused pol II to alternative polyadenylation in the drosophila nervous system. *Mol. Cell* **2015**, *57*, 341–348.
122. Grassi, E.; Santoro, R.; Umbach, A.; Grosso, A.; Oliviero, S.; Neri, F.; Conti, L.; Ala, U.; Provero, P.; Dicunto, F.; et al. Choice of alternative polyadenylation sites, mediated by the rna-binding protein Elavl3, plays a role in differentiation of inhibitory neuronal progenitors. *Front. Cell. Neurosci.* **2019**, *12*, 518.

123. Licatalosi, D.D.; Mele, A.; Fak, J.J.; Ule, J.; Kayikci, M.; Chi, S.W.; Clark, T.A.; Schweitzer, A.C.; Blume, J.E.; Wang, X.; et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **2008**, *456*, 464–469.
124. Masuda, A.; Takeda, J.I.; Okuno, T.; Okamoto, T.; Ohkawara, B.; Ito, M.; Ishigaki, S.; Sobue, G.; Ohno, K. Position-specific binding of FUS to nascent RNA regulates mRNA length. *Genes Dev.* **2015**, *29*, 1045–1057.
125. An, J.J.; Gharami, K.; Liao, G.-Y.; Woo, N.H.; Lau, A.G.; Vanevski, F.; Torre, E.R.; Jones, K.R.; 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.
126. Ciolli Mattioli, C.; Rom, A.; Franke, V.; Imami, K.; Arrey, G.; Terne, M.; Woehler, A.; Akalin, A.; Ulitsky, I.; Chekulaeva, M. Alternative 3′ UTRs direct localization of functionally diverse protein isoforms in neuronal compartments. *Nucleic Acids Res.* **2019**, *47*, 2560–2573.
127. Perry, R.B.; Fainzilber, M. Local translation in neuronal processes-in vivo tests of a “heretical hypothesis.” *Dev. Neurobiol.* **2014**, *74*, 210–217.
128. Biever, A.; Glock, C.; Tushev, G.; Ciirdaeva, E.; Dalmay, T.; Langer, J.D.; Schuman, E.M. Monosomes actively translate synaptic mRNAs in neuronal processes. *Science (80-.).* **2020**, *367*, 687475.
129. Litman, P.; Barg, J.; Rindzoonski, L.; Ginzburg, I. Subcellular localization of tau mRNA in differentiating neuronal cell culture: Implications for neuronal polarity. *Neuron* **1993**, *10*, 627–638.
130. Morita, T.; Sobuě, K. Specification of neuronal polarity regulated by local translation of CRMP2 and tau via the mTOR-p70S6K pathway. *J. Biol. Chem.* **2009**, *284*, 27734–27745.
131. Zempel, H.; Mandelkow, E. Mechanisms of Axonal Sorting of Tau and Influence of the Axon Initial Segment on Tau Cell Polarity. In; Springer, Singapore, 2019; pp. 69–77.
132. Wu, K.Y.; Hengst, U.; Cox, L.J.; Macosko, E.Z.; Jeromin, A.; Urquhart, E.R.; Jaffrey, S.R. Local translation of RhoA regulates growth cone collapse. *Nature* **2005**, *436*, 1020–1024.
133. Kang, H.; Schuman, E.M. A requirement for local protein synthesis in neurotrophin-induced hippocampal synaptic plasticity. *Science* **1996**, *273*, 1402–6.

134. Huber, K.M.; Kayser, M.S.; Bear, M.F. Role for rapid dendritic protein synthesis in hippocampal mGluR- dependent long-term depression. *Science (80-.).* **2000**, *288*, 1254–1256.
135. Scarnati, M.S.; Kataria, R.; Biswas, M.; Paradiso, K.G. Active presynaptic ribosomes in the mammalian brain, and altered transmitter release after protein synthesis inhibition. *Elife* **2018**, *7*, e36697.
136. Schanzenbächer, C.T.; Sambandan, S.; Langer, J.D.; Schuman, E.M. Nascent Proteome Remodeling following Homeostatic Scaling at Hippocampal Synapses. *Neuron* **2016**, *92*, 358–371.
137. Raab-Graham, K.F.; Haddick, P.C.G.; Jan, Y.N.; Jan, L.Y. Activity- and mTOR-dependent suppression of Kv1.1 channel mRNA translation in dendrites. *Science (80-.).* **2006**, *314*, 144–148.
138. Dziembowska, M.; Milek, J.; Janusz, A.; Rejmak, E.; Romanowska, E.; Gorkiewicz, T.; Tiron, A.; Bramham, C.R.; Kaczmarek, L. Activity-dependent local translation of matrix metalloproteinase-9. *J. Neurosci.* **2012**, *32*, 14538–14547.
139. Jones, K.J.; Templet, S.; Zemoura, K.; Kuzniewska, B.; Pena, F.X.; Hwang, H.; Lei, D.J.; Haensgen, H.; Nguyen, S.; Saenz, C.; et al. Rapid, experience-dependent translation of neurogranin enables memory encoding. *Proc. Natl. Acad. Sci.* **2018**, *115*, E5805–E5814.
140. Yudin, D.; Hanz, S.; Yoo, S.; Iavnilovitch, E.; Willis, D.; Gradus, T.; Vuppalanchi, D.; Segal-Ruder, Y.; Ben-Yaakov, K.; Hieda, M.; et al. Localized Regulation of Axonal RanGTPase Controls Retrograde Injury Signaling in Peripheral Nerve. *Neuron* **2008**, *59*, 241–252.
141. Hillefors, M.; Gioio, A.E.; Mameza, M.G.; Kaplan, B.B. Axon viability and mitochondrial function are dependent on local protein synthesis in sympathetic neurons. *Cell. Mol. Neurobiol.* **2007**, *27*, 701–716.
142. Andreassi, C.; Zimmermann, C.; Mitter, R.; Fusco, S.; Devita, S.; Saiardi, A.; Riccio, A. An NGF-responsive element targets myo-inositol monophosphatase-1 mRNA to sympathetic neuron axons. *Nat. Neurosci.* **2010**, *13*, 291–301.
143. Verma, P.; Chierzi, S.; Codd, A.M.; Campbell, D.S.; Meyer, R.L.; Holt, C.E.; Fawcett, J.W. Axonal protein synthesis and degradation are necessary for efficient growth cone regeneration. *J. Neurosci.* **2005**, *25*, 331–342.

144. Dieterich, D.C.; Hodas, J.J.L.; Gouzer, G.; Shadrin, I.Y.; Ngo, J.T.; Triller, A.; Tirrell, D.A.; Schuman, E.M. In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nat. Neurosci.* **2010**, *13*, 897–905.
145. Dieterich, D.C.; Link, A.J.; Graumann, J.; Tirrell, D.A.; Schuman, E.M. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc. Natl. Acad. Sci.* **2006**, *103*, 9482–9487.
146. Hidalgo San Jose, L.; Signer, R.A.J. Cell-type-specific quantification of protein synthesis in vivo. *Nat. Protoc.* **2019**, *14*, 441–460.
147. Aakalu, G.; Smith, W.B.; Nguyen, N.; Jiang, C.; Schuman, E.M. Dynamic Visualization of Local Protein Synthesis in Hippocampal Neurons. *Neuron* **2001**, *30*, 489–502.
148. Starck, S.R.; Green, H.M.; Alberola-Ila, J.; Roberts, R.W. A general approach to detect protein expression in vivo using fluorescent puromycin conjugates. *Chem. Biol.* **2004**, *11*, 999–1008.
149. Tom Dieck, S.; Kochen, L.; Hanus, C.; Heumüller, M.; Bartnik, I.; Nassim-Assir, B.; Merk, K.; Mosler, T.; Garg, S.; Bunse, S.; et al. Direct visualization of newly synthesized target proteins in situ. *Nat. Methods* **2015**, *12*, 411–414.
150. Halstead, J.M.; Lionnet, T.; Wilbertz, J.H.; Wippich, F.; Ephrussi, A.; Singer, R.H.; Chao, J.A. An RNA biosensor for imaging the first round of translation from single cells to living animals. *Science (80-.).* **2015**, *347*, 1367–1370.
151. Wu, B.; Eliscovich, C.; Yoon, Y.J.; Singer, R.H. Translation dynamics of single mRNAs in live cells and neurons. *Science (80-.).* **2016**, *352*, 1430–1435.
152. Morisaki, T.; Lyon, K.; DeLuca, K.F.; DeLuca, J.G.; English, B.P.; Zhang, Z.; Lavis, L.D.; Grimm, J.B.; Viswanathan, S.; Looger, L.L.; et al. Real-time quantification of single RNA translation dynamics in living cells. *Science (80-.).* **2016**, *352*, 1425–1429.
153. Yao, J.; Sasaki, Y.; Wen, Z.; Bassell, G.J.; Zheng, J.Q. An essential role for β-actin mRNA localization and translation in Ca2+-dependent growth cone guidance. *Nat. Neurosci.* **2006**, *9*, 1265–1273.
154. Berkovits, B.D.; Mayr, C. Alternative 3′ UTRs act as scaffolds to regulate membrane protein localization. *Nature* **2015**, *522*, 363–367.

155. Lee, S.H.; Mayr, C. Gain of Additional BIRC3 Protein Functions through 3'-UTR-Mediated Protein Complex Formation. *Mol. Cell* **2019**, *74*, 701-712.e9.
156. Valluy, J.; Bicker, S.; Aksoy-Aksel, A.; Lackinger, M.; Sumer, S.; Fiore, R.; Wüst, T.; Seffer, D.; Metge, F.; Dieterich, C.; et al. A coding-independent function of an alternative Ube3a transcript during neuronal development. *Nat. Neurosci.* **2015**, *18*, 666–673.
157. Crerar, H.; Scott-Solomon, E.; Bodkin-Clarke, C.; Andreassi, C.; Hazbon, M.; Logie, E.; Cano-Jaimez, M.; Gaspari, M.; Kuruvilla, R.; Riccio, A. Regulation of NGF Signaling by an Axonal Untranslated mRNA. *Neuron* **2019**, *102*, 553-563.e8.
158. Cox, D.B.T.; Gootenberg, J.S.; Abudayyeh, O.O.; Franklin, B.; Kellner, M.J.; Joung, J.; Zhang, F. RNA editing with CRISPR-Cas13. *Science (80-.).* **2017**, *358*, 1019–1027.
159. Soller, M.; White, K. ELAV inhibits 3′-end processing to promote neural splicing of ewg pre-mRNA. *Genes Dev.* **2003**, *17*, 2526–2538.
160. Zhang, Z.; So, K.; Peterson, R.; Bauer, M.; Ng, H.; Zhang, Y.; Kim, J.H.; Kidd, T.; Miura, P. Elav-Mediated Exon Skipping and Alternative Polyadenylation of the Dscam1 Gene Are Required for Axon Outgrowth. *Cell Rep.* **2019**, *27*, 3808-3817.e7.
161. Volden, R.; Palmer, T.; Byrne, A.; Cole, C.; Schmitz, R.J.; Green, R.E.; Vollmers, C. Improving nanopore read accuracy with the R2C2 method enables the sequencing of highly multiplexed full-length single-cell cDNA. *Proc. Natl. Acad. Sci.* **2018**, *115*, 9726–9731.
162. Tilgner, H.; Jahanbani, F.; Gupta, I.; Collier, P.; Wei, E.; Rasmussen, M.; Snyder, M. Microfluidic isoform sequencing shows widespread splicing coordination in the human transcriptome. *Genome Res.* **2018**, *28*, 231–242.
163. Mercer, T.R.; Wilhelm, D.; Dinger, M.E.; Soldà, G.; Korbie, D.J.; Glazov, E.A.; Truong, V.; Schwenke, M.; Simons, C.; Matthaei, K.I.; et al. Expression of distinct RNAs from 3′ untranslated regions. *Nucleic Acids Res.* **2011**, *39*, 2393–2403.
164. Kocabas, A.; Duarte, T.; Kumar, S.; Hynes, M.A. Widespread Differential Expression of Coding Region and 3′ UTR Sequences in Neurons and Other Tissues. *Neuron* **2015**, *88*, 1149–1156.
165. Malka, Y.; Steiman-Shimony, A.; Rosenthal, E.; Argaman, L.; Cohen-Daniel, L.; Arbib, E.; Margalit, H.; Kaplan, T.; Berger, M. Post-transcriptional 3´-UTR cleavage

of mRNA transcripts generates thousands of stable uncapped autonomous RNA fragments. *Nat. Commun.* **2017**, *8*, 2029.

166. Andreassi, C.; Luisier, R.; Crerar, H.; Blokzijl-Franke, S.; Luscombe, N.M.; Cuda, G.; Gaspari, M.; Riccio, A. 3′UTR cleavage of transcripts localized in axons of sympathetic neurons. *bioRxiv* **2019**, 170100.
167. Sudmant, P.H.; Lee, H.; Dominguez, D.; Heiman, M.; Burge, C.B. Widespread Accumulation of Ribosome-Associated Isolated 3′ UTRs in Neuronal Cell Populations of the Aging Brain. *Cell Rep.* **2018**, *25*, 2447-2456.e4.
168. Shinohara, M.; Mizushima, H.; Hirano, M.; Shioe, K.; Nakazawa, M.; Hiejima, Y.; Ono, Y.; Kanba, S. Eating disorders with binge-eating behaviour are associated with the s allele of the 3′-UTR VNTR polymorphism of the dopamine transporter gene. *J. Psychiatry Neurosci.* **2004**, *29*, 134–137.
169. Tarnok, Z.; Ronai, Z.; Gervai, J.; Kereszturi, E.; Gadoros, J.; Sasvari-Szekely, M.; Nemoda, Z. Dopaminergic candidate genes in Tourette syndrome: Association between tic severity and 3′ UTR polymorphism of the dopamine transporter gene. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2007**, *144*, 900–905.
170. Fletcher, J.M. Why Have Tobacco Control Policies Stalled? Using Genetic Moderation to Examine Policy Impacts. *PLoS One* **2012**, *7*, e50576.
171. Lotfipour, S.; Leonard, G.; Perron, M.; Pike, B.; Richer, L.; Séguin, J.R.; Toro, R.; Veillette, S.; Pausova, Z.; Paus, T. Prenatal exposure to maternal cigarette smoking interacts with a polymorphism in the α6 nicotinic acetylcholine receptor gene to influence drug use and striatum volume in adolescence. *Mol. Psychiatry* 2010, *15*, 6–8.
172. Barrie, E.S.; Lee, S.H.; Frater, J.T.; Kataki, M.; Scharre, D.W.; Sadee, W. Alpha-synuclein mRNA isoform formation and translation affected by polymorphism in the human SNCA 3'UTR. *Mol. Genet. Genomic Med.* **2018**, *6*, 565–574.
173. Marchese, D.; Botta-Orfila, T.; Cirillo, D.; Rodriguez, J.A.; Livi, C.M.; Fernández-Santiago, R.; Ezquerra, M.; Martí, M.J.; Bechara, E.; Tartaglia, G.G. Discovering the 3′ UTR-mediated regulation of alpha-synuclein. *Nucleic Acids Res.* **2017**, *45*, 12888–12903.
174. Toffoli, M.; Dreussi, E.; Cecchin, E.; Valente, M.; Sanvilli, N.; Montico, M.; Gagno, S.; Garziera, M.; Polano, M.; Savarese, M.; et al. SNCA 3′UTR genetic variants in

patients with Parkinson's disease and REM sleep behavior disorder. *Neurol. Sci.* **2017**, *38*, 1233–1240.

175. Sotiriou, S.; Gibney, G.; Baxevanis, A.D.; Nussbaum, R.L. A single nucleotide polymorphism in the 3′UTR of the SNCA gene encoding alpha-synuclein is a new potential susceptibility locus for Parkinson disease. *Neurosci. Lett.* **2009**, *461*, 196–201.
176. Shukla, S.; Parker, R. Hypo- and Hyper-Assembly Diseases of RNA–Protein Complexes. *Trends Mol. Med.* 2016, *22*, 615–628.
177. Jain, A.; Vale, R.D. RNA phase transitions in repeat expansion disorders. *Nature* **2017**, *546*, 243–247.
178. Miller, J.W. Recruitment of human muscleblind proteins to (CUG)n expansions associated with myotonic dystrophy. *EMBO J.* **2000**, *19*, 4439–4448.
179. Mankodi, A. Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum. Mol. Genet.* **2001**, *10*, 2165–2170.
180. Romo, L.; Ashar-Patel, A.; Pfister, E.; Aronin, N. Alterations in mRNA 3′ UTR Isoform Abundance Accompany Gene Expression Changes in Human Huntington's Disease Brains. *Cell Rep.* **2017**, *20*, 3057–3070.
181. Cheung, R.; Insigne, K.D.; Yao, D.; Burghard, C.P.; Wang, J.; Hsiao, Y.-H.E.; Jones, E.M.; Goodman, D.B.; Xiao, X.; Kosuri, S. A Multiplexed Assay for Exon Recognition Reveals that an Unappreciated Fraction of Rare Genetic Variants Cause Large-Effect Splicing Disruptions. *Mol. Cell* **2019**, *73*, 183-194.e8.

Chapter II: Elimination of *Calm1* long 3′ UTR mRNA isoform by CRISPR-Cas9 gene editing impairs dorsal root ganglion development and hippocampal neuron activation in mice

This chapter was published in RNA 2020, 26:1414-1430. Authors: Bongmin Bae*, Hannah N. Gruner*, Maebh Lynch, Ting Feng, Kevin So, Daniel Oliver, Grant S. Mastick, Wei Yan, Simon Pieraut, and Pedro Miura (*co-first authors)

The majority of mouse and human genes are subject to alternative cleavage and polyadenylation (APA), which most often leads to the expression of two or more alternative length 3′ untranslated region (3′ UTR) mRNA isoforms. In neural tissues, there is enhanced expression of APA isoforms with longer 3′ UTRs on a global scale, but the physiological relevance of these alternative 3′ UTR isoforms is poorly understood. *Calmodulin 1* (*Calm1)* is a key integrator of calcium signaling that generates short (*Calm1-S*) and long (*Calm1-L*) 3′ UTR mRNA isoforms via APA. We found *Calm1-L* expression to be largely restricted to neural tissues in mice including the dorsal root ganglion (DRG) and hippocampus, whereas *Calm1-S* was more broadly expressed. smFISH revealed that both *Calm1-S* and *Calm1-L* were subcellularly localized to neural processes of primary hippocampal neurons. In contrast, cultured DRG showed restriction of *Calm1-L* to soma. To investigate the *in vivo* functions of *Calm1-L*, we implemented a CRISPR-Cas9 gene editing strategy to delete a small region encompassing the *Calm1* distal polyA site. This eliminated *Calm1-L* expression while maintaining expression of *Calm1-S*. Mice lacking *Calm1-L* (*Calm1*$^{\Delta L/\Delta L}$) exhibited disorganized DRG migration in embryos, and reduced experience-induced neuronal activation in the adult hippocampus. These data indicate that *Calm1-L* plays functional roles in the central and peripheral nervous systems.

1. Introduction

Alternative cleavage and polyadenylation (APA) is the process by which a pre-mRNA is cleaved and polyadenylated at two or more polyadenylation (polyA) sites. This commonly results in two (or more) mRNAs with the same protein coding sequence but different length 3′ UTRs. APA is pervasive, occurring in ~51-79% of mammalian genes (Hoque et al. 2013; Lianoglou et al. 2013). The 3′ UTR is a major target for post-transcriptional regulation via microRNAs (miRNAs) and RNA binding proteins (RBPs); thus, harboring a longer 3′ UTR can theoretically confer additional regulatory opportunities for transcripts (Sandberg et al. 2008; Mayr and Bartel 2009). Long 3′ UTRs impact translation in a cell context-specific manner (Floor and Doudna 2016; Blair et al. 2017), and elements located in alternative 3′ UTRs can influence mRNA localization in neurons (Tushev et al. 2018). Thousands of genes in mouse and human express long 3′ UTR mRNA isoforms in brain tissues (Miura et al. 2013). For many genes, short 3′ UTR isoforms were found to be expressed across various tissues, whereas the alternative long 3′ UTR isoforms were abundant only in brain. Our understanding of the in vivo functions of neural-enriched long 3′ UTR isoforms is limited despite the large number of genes shown to be affected in *Drosophila*, Zebrafish, mice, and humans (Smibert et al. 2012; Ulitsky et al. 2012; Miura et al. 2013).

Multiple studies have examined the functions of 3′ UTRs in vivo by generating mice that either genetically delete parts of 3′ UTRs or introduce foreign polyA sites to cause production of mRNAs with truncated 3′ UTRs (Miller et al. 2002; Perry et al. 2012). For example, a genetic approach was implemented in mice to abolish the BDNF long 3′ UTR isoform while maintaining expression of its short mRNA counterpart via the insertion of tandem SV40 polyA sites downstream of the proximal polyadenylation signal. These mice displayed synaptic defects and hyperphagic obesity, ostensibly due to compromised

mRNA localization to dendrites and impaired translational control (An et al. 2008; Liao et al. 2012). The recent advent of CRISPR-Cas9 gene editing has revolutionized the speed and efficiency of generating deletion mouse strains (Wang et al. 2013). This presents an exciting new opportunity for rapidly generating 3′ UTR isoform-specific knockout mice. CRISPR-Cas9 has been implemented to successfully generate mice with a deletion that removed most of the mTOR 3′ UTR (Terenzio et al. 2018). However, successful generation of an isoform-specific, long 3′ UTR knockout mouse using CRISPR-Cas9 has not been reported to date.

Calmodulin (CaM) is the primary calcium sensor in the cell (Means and Dedman 1980; Yamniuk and Vogel 2004; Sorensen et al. 2013). CaM is expressed ubiquitously but is particularly abundant in the nervous system (Kakiuchi et al. 1982; Ikeshima et al. 1993). There are three Calmodulin genes in mammals– *Calm1*, *Calm2*, and *Calm3*. These share an identical amino acid coding sequence, but possess unique 5′ and 3′ UTRs (SenGupta et al. 1987; Fischer et al. 1988), suggesting differences in their regulation might be conferred at the post-transcriptional level. The existence of alternative 3′ UTR mRNA isoforms generated by APA for the Calmodulin 1 (*Calm1*) gene have been known for several decades (SenGupta et al. 1987; Nojima 1989; Ni et al. 1992; Ikeshima et al. 1993). These include mRNAs with a short 0.9 kb 3′ UTR (*Calm1-S*) and one with a long 3.4 kb 3′ UTR (*Calm1-L*). The functional significance of these alternative 3′ UTR isoforms is entirely unexplored.

Previous work has shown that altering CaM levels disrupts neuronal development in *Drosophila* and rodents (Vanberkum and Goodman 1995; Fritz and VanBerkum 2000; Kim et al. 2001; Kobayashi et al. 2015; Wang et al. 2015). CaM plays a role in guiding axon projections to create connections with other cells (Vanberkum and Goodman 1995; Fritz and VanBerkum 2000; Kim et al. 2001; Kobayashi et al. 2015). Additionally, CaM

interplays with a variety of synaptic proteins to play a fundamental role in regulating signaling pathways critical in synaptic plasticity (Xia and Storm 2005). Functions in the nervous system specifically for *Calm1* have been described. Notably, targeted knockdown of *Calm1*, but not *Calm2* or *Calm3*, was found to cause major migration defects in developing hindbrain neurons (Kobayashi et al. 2015). *Calm1* mRNA has been detected in cultured rat embryonic dorsal root ganglion (DRG) axons, without distinguishing *Calm1-S* or *Calm1-L* isoforms, where it undergoes local translation to promote axon outgrowth in vitro (Wang et al. 2015). *Calm1* mRNAs, especially *Calm1-L*, were found to be expressed in dendrites of rat hippocampal neurons (Tushev et al. 2018). However, whether *Calm1-L* has a specific function in the embryonic or adult nervous system has not been addressed. Here, we utilized CRISPR-Cas9 to generate the first mouse lines lacking expression of a long 3′ UTR mRNA isoform while maintaining normal expression of its short 3′ UTR APA counterpart. We uncovered that *Calm1-L* levels were high in the DRG and hippocampus, and its expression was largely restricted to neurons. In vivo phenotypic analysis of mutant embryos revealed disorganized DRG axon and cell body positioning, whereas adults displayed a reduction in hippocampal neuronal activation after enriched environment exposure. This novel deletion methodology has thus uncovered in vivo neuronal impairments attributed to the loss of a single long 3′ UTR mRNA isoform in mice.

2. Results

Calm1-L expression is enriched in neurons

To determine the relative expression of short and long *Calm1* 3′ UTR isoforms among mouse tissues, we performed Northern blot analysis– an approach uniquely suited to report on the expression of alternative 3′ UTRs (Miura et al. 2014). Northern blots performed using a universal probe (uni) that detects both short and long *Calm1* 3′ UTR

isoforms revealed the presence of both *Calm1-L* and *Calm1-S* across an array of adult tissues. The cerebral cortex showed greater enrichment of *Calm1-L* versus *Calm1-S* compared to non-brain tissues (**Fig. 5a,b**), as had been previously observed (Ikeshima et al. 1993; Miura et al. 2013). The ratio of long isoform to the sum of long and short isoforms (total *Calm1*) was found to be 0.46 in cortex, whereas in the rest of the tissues examined (gastrocnemius skeletal muscle, heart, testes, kidney, and liver) this value ranged between 0.12 – 0.28. Recent studies have shown that some 3′ UTRs can be cleaved to generate stable 3′ UTR transcripts that are separated from the protein coding portion of the mRNA (Kocabas et al. 2015; Malka et al. 2017; Sudmant et al. 2018). To test whether this might be the case for *Calm1*, and to confirm the connectivity of the long 3′ UTR region to the rest of the *Calm1* mRNA, the blot was stripped and re-probed with a probe specific for *Calm1-L* (ext). One band consistent with the size of *Calm1-L* was observed (**Fig. 5c**). Thus, the long 3′ UTR region is connected to the full-length mRNA, with no evidence of a cleaved 3′ UTR observed in the tissues examined.

We next determined the spatial expression pattern of *Calm1* 3′ UTR isoforms in mouse tissues. In situ hybridization (ISH) using Digoxigenin (DIG)-labeled probes was performed in embryonic day 13.5 (E13.5) mice and 8-weeks-old brains. Using a *Calm1* universal probe (uni; **Fig. 5a**), strong signal was observed in the brain (including forebrain (FB), midbrain (MB), and hindbrain (HB)), spinal cord (SC), and dorsal root ganglion (DRG) in parasagittal sections of embryos. Uni signal was also strong in whole coronal sections of the adult brain including amygdala, striatum, hypothalamus, cortex, and hippocampus (Hpc) (**Fig. 5d**). Expression was also observed in the nasal epithelium (N), lung (L), and intestinal epithelium (I). In contrast, ISH performed with probe against the long 3′ UTR (ext; **Fig. 5a**) showed staining largely restricted to the brain, spinal cord, and DRG in parasagittal sections of embryo and stronger signal in the adult Hpc compared to other

brain regions (**Fig. 5e**). These experiments demonstrate that *Calm1-L* is enriched in neural tissues.

The global bias for enhanced expression of longer 3′ UTR isoforms in the nervous system has generally been attributed to expression in neurons (Guvenek and Tian 2018). To test if this is the case for *Calm1*, we analyzed the relative expression of the *Calm1* 3′ UTR isoforms in different brain cell types from a previously published RNA-seq dataset (Zhang et al. 2014). We performed re-analysis of these data using QAPA, which allows for an estimation of relative polyA site usage (PAU) from RNA-seq reads (Ha et al. 2018). We analyzed the usage of *Calm1-S* and *Calm1-L* among purified neurons, microglia, astrocytes, newly formed oligodendrocytes, and endothelial cells. The coverage tracks of these data showed the highest enrichment of reads pertaining to the long 3′ UTR in neurons compared to the other cell types (**Fig. 5f**). Quantification using QAPA revealed neurons had the greatest usage of *Calm1-L* (32 %) (**Fig. 5g**).

To determine the cellular expression of *Calm1* 3′ UTR isoforms in our tissues of interest, we performed branched-oligo based single molecule Fluorescence In Situ Hybridization (RNAscope® smFISH) on primary cultures from DRG and hippocampus (**Fig. 5h-l**). Co-staining with anti-β3Tubulin (Tubb3) was used to distinguish neurons from other cell types. In cultured DRG, most cells showed robust uni signal, but ext signal was only prominent in Tubb3-positive cells (**Fig. 5i**). Counting of uni and ext puncta indicated that *Calm1-L* is mostly enriched in Tubb3-positive neurons in DRG (**Fig. 5j).** Analysis of primary hippocampal cultures similarly revealed that the *Calm1* ext signal was significantly enriched in Tubb3-positive neurons (**Fig. 5k,l**). Together with the RNA-seq analysis of isolated brain cells, these data show that *Calm1-L* expression is enriched in neurons.

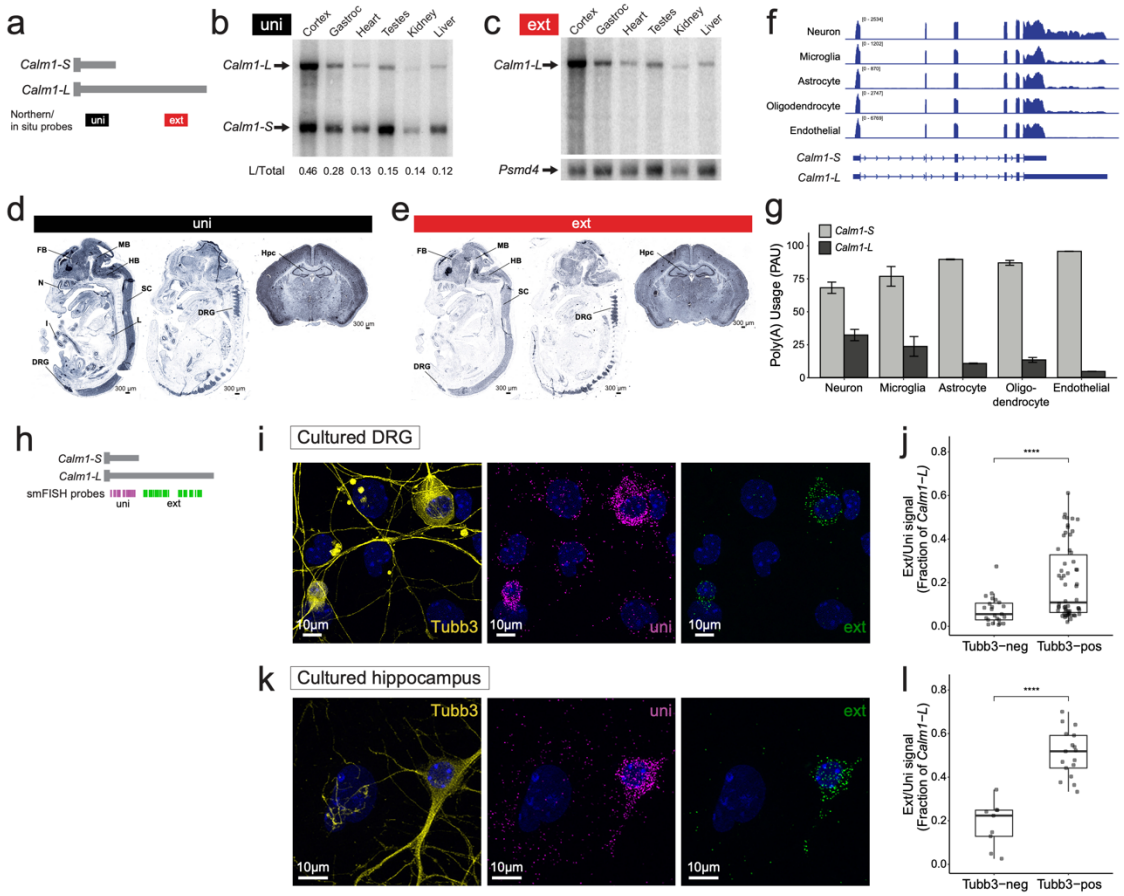


Figure 5. *Calm1-L* expression is enriched in neurons

(a) Diagram showing Northern blot and DIG in situ probes to detect either both isoforms (uni) or specifically the long 3′ UTR isoform (ext). (b) Northern blot of an array of tissues collected from adult mice probed with uni probes showing *Calm1-S* (bottom arrow) and *Calm1-L* (top arrow). The ratio of *Calm1-L* normalized to total *Calm1* (sum of Calm1-S band intensity and *Calm1-L*) is shown. (c) Northern blot performed with ext probe. The same blot was stripped and re-probed for the housekeeping gene *Psmd4* as a loading control. (d) DIG in situ hybridization of E13.5 embryo and 8w brain using the uni probe showed universal expression of *Calm1* showing strong signal in the nervous system including forebrain (FB), midbrain (MB), hindbrain (HB), spinal cord (SC), dorsal root ganglion (DRG), cortex, and hippocampus (Hpc). (e) DIG in situ performed with the ext probe showed neural tissue-specific expression pattern of *Calm1-L* with particularly strong signals in the DRG and Hpc. (f) RNA-seq tracks from purified neurons, microglia, astrocytes, newly formed oligodendrocytes, and endothelial cells are visualized. *Calm1-L* gene model is shown as annotated and *Calm1-S* gene model was illustrated based on the sequencing read coverages. Read coverage in the long 3′ UTR is particularly high in neurons. (g) QAPA was used to estimate the fraction of *Calm1-S* and *Calm1-L* found in each RNA-seq dataset. The Poly(A) usage indicates that *Calm1-L* is indeed

enriched in neurons. 2 replicates for each cell type. (h) Diagram showing RNAscope smFISH probe locations. (i-j) smFISH performed for total *Calm1* (uni) or *Calm1-L* (ext) transcripts in primary DRG culture co-labeled with β3 Tubulin (Tubb3) neuronal marker. Most cells showed robust uni signals while ext signals were restricted in Tubb3-positive cells. Significance was determined using a t-test, ****: p <= 0.0001. n = 58 Tubb3-pos cells, n = 23 Tubb3-neg cells. (k-l) smFISH performed in primary hippocampal culture showing the same trend. Significance was determined using a t-test, ****: p <= 0.0001. n = 17 Tubb3-pos cells, n = 9 Tubb3-neg cells.

Subcellular localization of *Calm1-L*

Calm1 mRNA, without distinguishing between *Calm1-S* or *-L,* has been shown to be localized to axons and dendrites in multiple studies (Zivraj et al. 2010; Gumy et al. 2011; Kobayashi et al. 2015; Wang et al. 2015; Preitner et al. 2016; Zappulo et al. 2017; Tushev et al. 2018). Given earlier work showing that elements present in long 3′ UTRs of importin β1 and Impa1 are involved in mRNA localization to axons (Perry et al. 2012; Andreassi et al. 2010), we hypothesized that *Calm1-L* might be specifically localized to axons. To test this hypothesis, we cultured dissociated embryonic DRG neurons in compartmentalized chambers or on glass coverslips and performed smFISH. DRG neurons are unipolar cells with a single axon stem bifurcating into a peripheral and a central branch, and use of embryonic DRG allows a high efficiency of isolating neurons (Melli and Höke 2009). Previous studies have shown total *Calm1* expression in sensory neuron axons by multiple approaches, including FISH, microarray, and RNA-Seq analysis of axonal transcriptomes (Zivraj et al. 2010; Gumy et al. 2011; Wang et al. 2015; Preitner et al. 2016). In agreement with these studies, we found that the probes targeting both *Calm1* isoforms (uni) revealed signal in soma and axons (**Fig. 6a,b**). *Calm1-L* expression is indicated by white puncta in these experiments due to co-localization of uni (cyan) and ext (magenta) signals (**Fig. 6a**). We found that *Calm1-L* showed little to no signal in DRG axons (**Fig. 6b′, b″**). To confirm

this expression pattern, we re-analyzed public RNA-seq data obtained from isolated axons of DRG/spinal cord (Minis et al. 2014). QAPA analysis revealed the relative expression of the long 3′ UTR to be 23% in whole DRG. In isolated axons, there was less relative expression of the long 3′ UTR (11%) (**Supplemental Fig. 1a**). This result is consistent with our FISH data showing less enrichment of *Calm1-L* in axons (**Fig. 6c**).

A recent study that employed global analysis of 3′ end sequencing found *Calm1-L* to be enriched over *Calm1-S* expression in the rat hippocampal neuropil (Tushev et al. 2018), which is a region that comprises axons, dendrites, synapses, interneurons, and glia (Cajigas et al. 2012). In order to assess subcellular localization patterns of *Calm1* 3′ UTR isoforms in mouse hippocampal neurons, we cultured postnatal day 0 (P0) mouse hippocampal neurons and performed smFISH. β3Tubulin marker was co-labeled to identify neurons. We could not simultaneously monitor short and long isoforms along with multiple cellular markers (for example, both β3Tubulin and Tau to distinguish axons and dendrites) by fluorescence microscopy, thus we refer to all neuronal extensions as "processes" for this analysis. The short 3′ UTR isoform was found in both soma and processes (cyan signals in **Fig. 6d″**). In contrast to the results observed in DRG neurons, *Calm1-L* was found to be localized in hippocampal neuron processes (**Fig. 6d′, d″** white signals depicted with arrowheads). Quantification of the number of localized puncta per neuron and maximum distance of travel for each isoform revealed no significant difference between the two *Calm1* isoforms (**Fig. 6e**, **Supplemental Fig. 1b**). However, the distance of travel from the center of the soma out to processes was found to be significantly different between *Calm1-S* and *Calm1-L*, largely due to enriched *Calm1-S* distribution in the soma (**Fig. 6f**). Collectively, these results show that *Calm1-L* is not expressed outside of the soma in DRG, but is expressed in both soma and processes in hippocampal neurons.

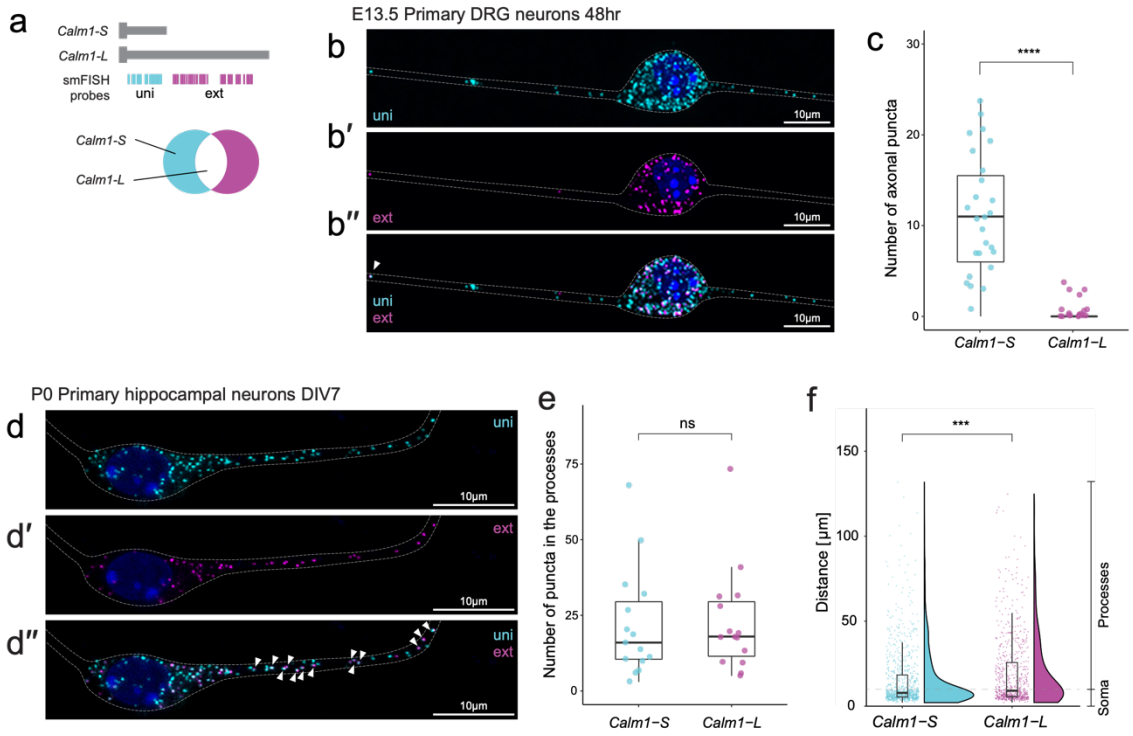


Figure 6. Subcellular localization of *Calm1-L*.

(a) Diagram showing RNAscope smFISH probe locations. When uni and ext images are merged, *Calm1-L* isoforms are shown as colocalized (white) puncta. (b) smFISH showed robust axonal localization of *Calm1-S* (uni) but *Calm1-L* (ext/magenta in b′ or white punta in b″) was observed in DRG axons. (c) When the number of axonal uni or ext puncta were counted, uni signals were robustly found in the axons of DRG neurons but little to no ext signals were found in the same regions. Significance was determined using a t-test, ****: p <= 0.0001. n = 30 neurons. (d) In primary hippocampal neurons, both *Calm1-S* and *Calm1-L* were observed in soma and neuronal processes. (e) The number of puncta corresponding to *Calm1-S* and *Calm1-L* transcript in the hippocampal processes was not found to be significantly different. Significance was determined using a t-test, ns: p > 0.05. n = 15 neurons. (f) Analysis of the distance of travel for all the *Calm1-S* and *Calm1-L* signals showed the overall distribution is different between two mRNA isoforms. Analysis done in n = 15 neurons (*Calm1-L* 684 puncta, *Calm1-S* 907 puncta). Significance was determined using a Wilcoxon test, ***: p <= 0.001.

Generation of *Calm1* long 3′ UTR deletion mice using CRISPR-Cas9

To investigate the functional role of neuron-enriched *Calm1-L in vivo*, we employed CRISPR-Cas9 gene editing to generate a series of mouse lines that lacked the expression of *Calm1-L*. We aimed to generate a mutant line which did not harbor any foreign DNA sequences to promote usage of the proximal polyA site over the distal polyA site. We designed 6 single guide RNAs (gRNAs) targeting the *Calm1* locus that we anticipated would result in three different deletions (**Fig. 7a**).

Deletion mice strains were generated by injection of all 6 gRNAs along with mRNA encoding Cas9 endonuclease. From 11 founder mice we isolated 3 lines each harboring different deletions that we confirmed using Sanger sequencing. Deletion 1 removed sequence encompassing the long 3′ UTR, including the distal polyA site. Deletion 2 removed the majority of the sequence comprising the long 3′ UTR, save for the distal polyA site. This strategy brought the proximal and distal polyA sites adjacent to one another, which we surmised might ensure cleavage at this region and prevent selection of cryptic polyA sites further downstream. Deletion 3 was designed to remove the distal polyA site, in anticipation that although the long 3′ UTR would be transcribed, it would not be cleaved and polyadenylated, thus preventing biogenesis of mature *Calm1-L* mRNA.

The effectiveness in preventing *Calm1-L* biogenesis in these three lines was determined by Northern blot analysis using the uni probe (**Fig. 5a**). We chose to analyze cortex samples instead of smaller tissues because a great deal of starting RNA is required to perform these mRNA Northern blots. Remarkably, each deletion strategy had the desired outcome. We found that Deletions 1 and 3 exhibited complete loss of *Calm-L* transcripts (**Fig. 7b**). As anticipated, Northern analysis of deletion 2 line showed a band

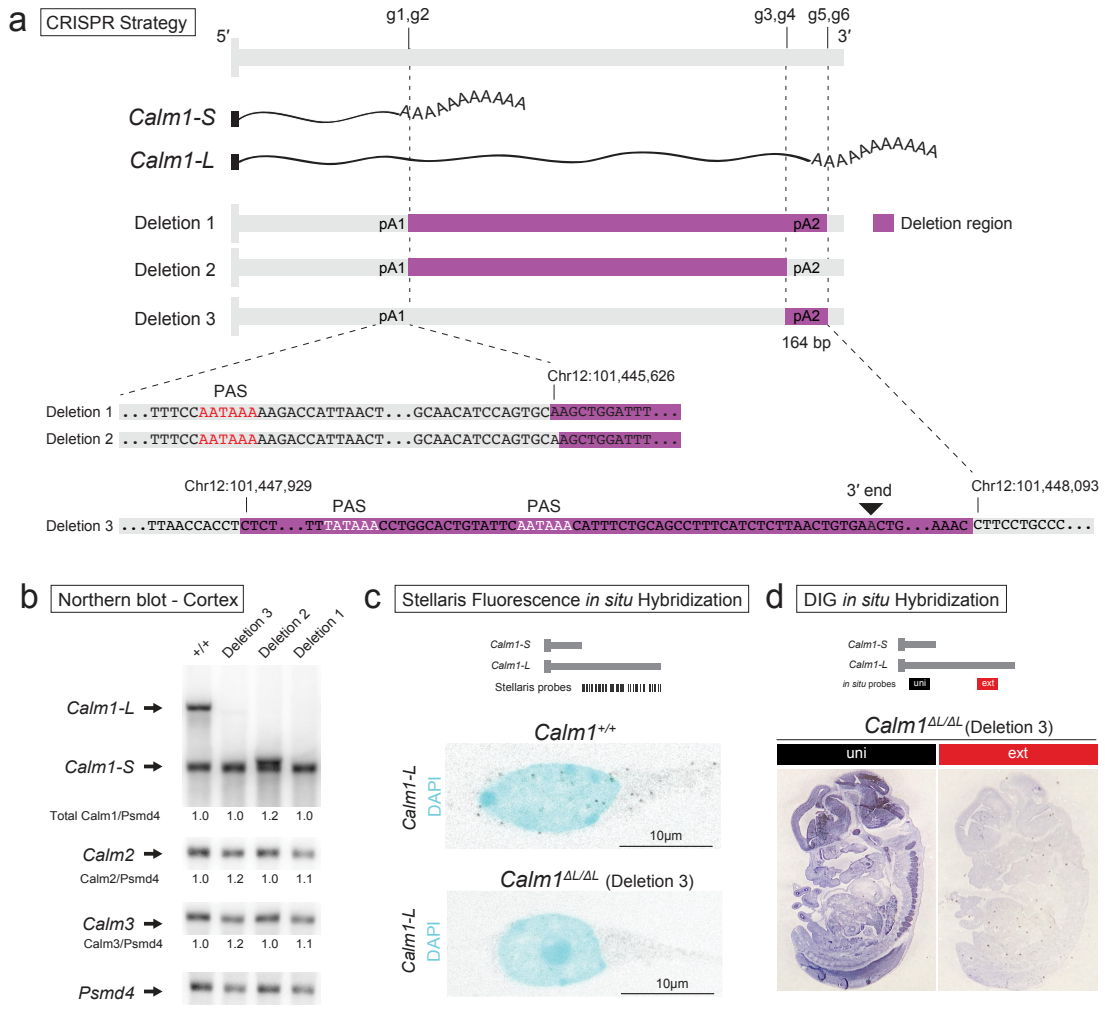


Figure 7. Generation of *Calm1* long 3′ UTR deletion mice using CRISPR-Cas9.

(a) Diagram of the strategies used to eliminate the production of mature long *Calm1* 3′ UTR transcripts. Six gRNAs (g1-g6) were injected simultaneously to generate a variety of deletions (Deletion 1 - 3). (b) *Calm1* Northern blot of adult cortex from three deletion strains and control littermate demonstrating successful deletion of the long *Calm1* 3′ UTR isoform. Note the Deletion 2 line generates a new isoform with a truncated long 3′ UTR due to the preservation of the distal PAS. Transcripts from the *Calm2* and *Calm3* genes were found to be unaltered. *Psmd4* used as a loading control. (c) Complete loss of *Calm1-L* in the *Calm1*$^{\Delta L/\Delta L}$ (Deletion 3) hippocampal neurons has been confirmed by lacked smFISH signals representing *Calm1-L*. (d) DIG *in situ* performed in the *Calm1*$^{\Delta L/\Delta L}$ (Deletion 3) embryo also revealed no expression of *Calm1-L*.

migrating slightly higher than the *Calm1-S* band, indicating the biogenesis of an ectopic transcript slightly longer than *Calm1-S* due to the second PAS (**Fig. 7b**). Transcripts from the *Calm2* and *Calm3* genes were also monitored by Northern blot and were found to be unaltered (**Fig. 7b**; for full blots, see **Supplemental Fig. 2**).

In deciding on which line to carry out phenotypic analysis, we surmised that deletion 2 was the most confounding because an ectopic transcript was generated. The deletion 1 and 3 strategies both had the desired effect on *Calm1–L* loss without altering total *Calm1* levels. We proceeded with phenotypic analysis on the deletion 3 allele because it had the smallest amount of genomic sequence deleted (164 bp) and thus reasoned that it was the least likely to have alterations in genomic elements such as unknown enhancer elements that could be present in the long 3′ UTR-encoding region. This mouse deletion line 3 was re-named *Calm1*$^{\Delta L/\Delta L}$.

We further verified the complete loss of *Calm1-L* in the *Calm1*$^{\Delta L/\Delta L}$ hippocampus using Stellaris® smFISH and in the DRG by ISH. Cultured hippocampal neurons from the *Calm1*$^{\Delta L/\Delta L}$ mice completely lacked signal representing *Calm1-L* (**Fig. 7c**). ISH for *Calm1-L* performed in E13.5 embryo sections also revealed no expression (**Fig. 7d**). We confirmed by Sanger sequencing that the *Calm1*$^{\Delta L/\Delta L}$ line did not harbor unintended mutations at predicted gRNA off target sites. (**Supplemental Fig. 3**).

Calm1$^{\Delta L/\Delta L}$ mice exhibit DRG axon development defects

Given the pronounced expression of *Calm1-L* in DRG (**Fig. 5e**) we examined *Calm1*$^{\Delta L/\Delta L}$ embryos for DRG developmental defects. The DRG is a series of ganglia in the peripheral nervous system derived from neural crest cells, which by early embryonic stage E10.5 have already begun to form distinct ganglia adjacent to the spinal neural tube caudal to the hindbrain (Le Douarin and Smith 1988; Marmigere and Ernfors 2007) (**Fig.**

8a). Cell bodies of the DRG send out dorsolateral axon projections concurrent with neurogenesis during this stage in development (Marmigere and Ernfors 2007). Unlike DRG that form adult structures, the first cervical (C1) DRG is a temporary embryonic population that loses its dorsolateral axonal projections and progressively undergoes programmed cell death from E10.5 to ~E12.5 (Fanarraga et al. 1997; van den Akker et al. 1999). The axons and cell bodies of the C1 DRG in *Calm1*$^{\Delta L/\Delta L}$ embryos at E10.5 were found to be severely disorganized relative to *Calm1*$^{+/+}$ embryos (**Fig. 8b,c**). Large groups of cell bodies of the C1 DRG in mutant embryos translocated rostral into the hindbrain adjacent to the accessory nerve (n.xi) tracks (**Fig. 8c-c″**). At the same location in *Calm1*$^{+/+}$ embryos, there were fewer translocating C1 DRG bodies that far rostral and disorganized (**Fig. 8b-b″**). The C1 DRG cell bodies in *Calm1*$^{\Delta L/\Delta L}$ grouped together in smaller ganglia (**Fig. 8c-c″**). Axons branching off these cell bodies projected aberrantly and were tightly fasciculated (**Fig. 8c′**, see arrows). Some *Calm1*$^{\Delta L/\Delta L}$ axons projected longitudinally, which was in contrast to the dorsolateral projections seen in non-C1 DRG (see Methods). Significantly more C1 DRG cell bodies rostrally migrated in the mutants compared to controls (**Fig. 8d,e**). Lastly, we quantified the number of fascicles projecting off the C1 ganglia and found there were significantly higher numbers emanating from *Calm1*$^{\Delta L/\Delta L}$ DRG compared to *Calm1*$^{+/+}$ (**Fig. 8f**). Together these data show that loss of *Calm1-L* impairs C1 DRG axon development and restricts rostral cell migration. We reasoned that the impairment of DRG development in *Calm1*$^{\Delta L/\Delta L}$ was a result of altered translation of *Calm1* resulting from loss of the long 3′ UTR isoform. We thus performed Western analysis for CaM in wild type and mutant dissected E13.5 DRG. Note that three genes encode identical CaM protein, thus this analysis cannot uniquely report on CaM generated from *Calm1*. Western analysis was performed using capillary immunoassay due to the low

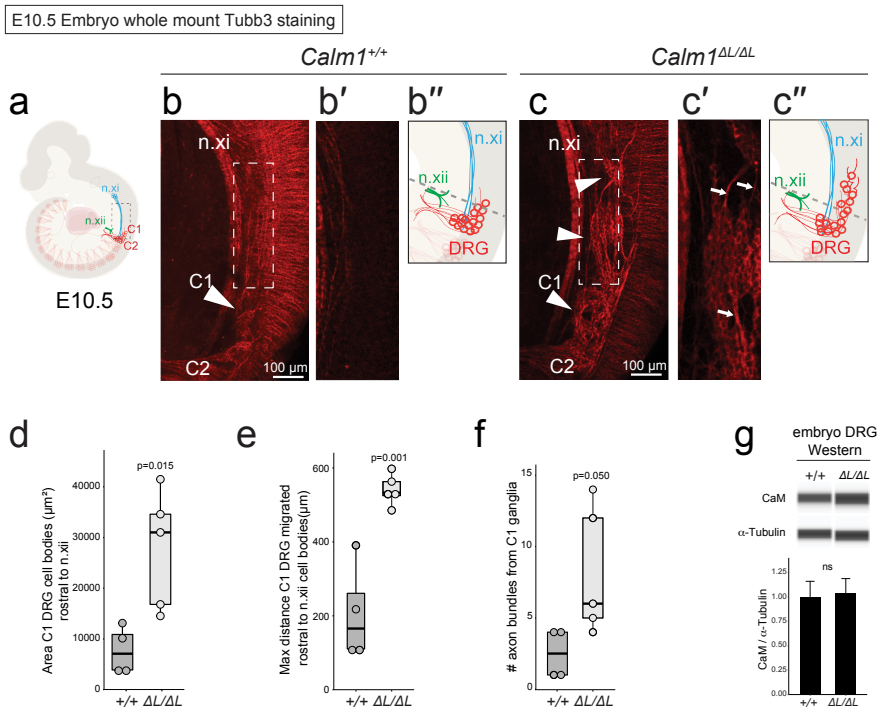


Figure 8. *Calm1*$^{\Delta L/\Delta L}$ mice exhibit DRG axon development defects
Developing C1 DRG exhibit axonal and cell body migration disorganization in *Calm1*$^{\Delta L/\Delta L}$ E10.5 embryos. (a) Schematic of *Calm1*$^{+/+}$ E10.5 embryo highlighting the morphology of the C1 and C2 DRG axons and cell bodies. (b-c) Whole mount *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ DRG morphology visualized by anti-Tubb3 labeling. (b) The cell bodies of the C1 DRG (arrowhead) in *Calm1*$^{+/+}$ embryos are bundled together to form a distinct ganglion. The neurites of the C2 DRG can be seen projecting ventrally in an organized bundle. (c) The C1 DRG in *Calm1*$^{\Delta L/\Delta L}$ animals is disorganized and consists of clusters of cell bodies (arrowheads) that extend bundles of axons (arrows). The C1 DRG cells of deletion animals migrate rostral relative to *Calm1*$^{+/+}$ and are adjacent to the n.xi tract (b′, c′). Enlarged view of single optical section of same Z-stack from (b) and (c) focusing on disorganized mutant cell body morphology of ganglion that aberrantly migrated more rostral relative to control. (d-f) In order to start measurements in the same relative location between embryos, n.xii was used as an anatomical landmark to set a beginning of measurements, indicated by the grey dashed lines in (b″, c″). (d) Quantification of area of ectopic cells bodies observed for developing DRG. (e) Distance of aberrantly clustered cells bodies in *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$. (f) Quantification of the number of axon bundles projecting off C1 ganglia. Significance determined by a t-test; n=4 C1 DRG *Calm1*$^{+/+}$; n=5 C1 DRG *Calm1*$^{\Delta L/\Delta L}$. n.xi= accessory nerve, n.xii=hypoglossal nerve. (g) Capillary western analysis of *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ embryonic DRG showing no change in overall protein levels. Significance was determined using a t-test, p = 0.802; n=3.

amounts of DRG material recovered from dissection. Western analysis failed to identify a significant change in CaM levels between *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ embryonic DRGs (**Fig. 8g**).

Subcellular localization of *Calm1* and CaM levels are unaltered in *Calm1*$^{\Delta L/\Delta L}$ hippocampus.

Next, we examined whether *Calm1*$^{\Delta L/\Delta L}$ mutants experience neurological defects in the central nervous system. We directed our attention to the hippocampus given the high expression of *Calm1-L* in this tissue (**Fig.1e**). First, we characterized whether the loss of *Calm1-L* transcripts altered the total amount of total *Calm1* transcripts or CaM protein in the hippocampus. qRT-PCR and Western blot analysis revealed no changes in total *Calm1* RNA or CaM protein levels between *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ (**Fig. 9a,b**). Additionally, we assessed changes in *Calm1* mRNA localization into the processes of hippocampal neurons. smFISH analysis in both *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ hippocampal neurons showed that the deletion of *Calm1-L* did not significantly impair localization of *Calm1* mRNAs in hippocampal neurons (**Fig. 9c,d**). Thus, our results suggested that the overall *Calm1* mRNA levels, subcellular localization of *Calm1* mRNAs, or CaM protein levels are not significantly altered in *Calm1*$^{\Delta L/\Delta L}$ hippocampus compared to *Calm1*$^{+/+}$.

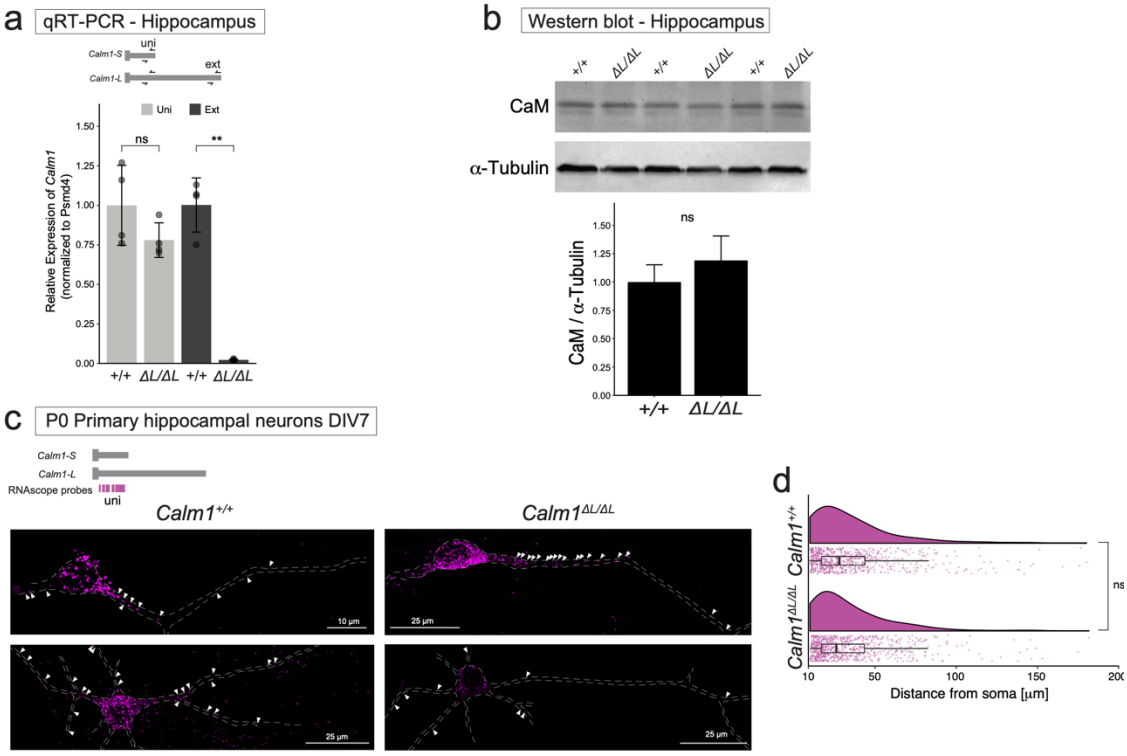


Figure 9. Characterization of *Calm1* isoforms in the hippocampus

(a) qRT-PCR analysis revealed no changes in total *Calm1* (uni) RNA levels between *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ while the *Calm1-L* expression is abolished in *Calm1*$^{\Delta L/\Delta L}$ hippocampus. Significance was determined using a t-test, ns: p > 0.05, **: p <= 0.01, n = 4 mice. (b) Western blot analysis of *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ adult hippocampus showing no change in overall protein levels. Significance was determined using a t-test, p = 0.287; n = 3 mice. (c-d) smFISH analysis in both *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ hippocampal neurons showed that the deletion of *Calm1-L* did not significantly impair localizing potential of *Calm1* mRNAs in hippocampal neurons. Significance was determined using a Wilcoxon test, ns: p > 0.05; n = 19 *Calm1*$^{+/+}$ neurons (739 puncta), n = 26 *Calm1*$^{\Delta L/\Delta L}$ neurons (802 puncta).

$Calm1^{\Delta L/\Delta L}$ mice exhibit reduced hippocampal IEG expression in response to enriched environment exposure.

CaM is an integral part of CaMK, Calcineurin, and MAPK signaling pathways which are involved in synaptic transmission and plasticity (Xia and Storm 2005; Sethna et al. 2016; Pang et al. 2010; Hagenston and Baing 2011). Ca2+ signaling mediates synaptic plasticity via post-translational alterations and initiating signaling cascades resulting in de novo transcription (Greer and Greenberg 2008; Hagenston and Bading 2011). With this in mind, we hypothesized the *Calm1-L* might impact CaM-dependent synaptic transmission and/or plasticity. As a proof of principle, we used an enriched environment (EE) paradigm to induce experience-driven expression of immediate early genes (IEGs) in $Calm1^{\Delta L/\Delta L}$ hippocampi and compared it to $Calm1^{+/+}$. IEG expression, especially the canonical example of cFos induction, is a marker of neuronal activation and is considered to be required for learning and synaptic plasticity (Yap and Greenberg 2018).

Exposure to environmental novelty leads to a selective increase of cFos in the rat Cornu Ammonis 1 (CA1) region (VanElzakker et al. 2008) and novel object and place induces cFos expression in the rat CA1 (Ito and Schuman 2012). Immediately after exposing the mice to EE for 1.5 hours, animals were sacrificed, and brains were collected. EE-induced expression of cFos was compared in $Calm1^{+/+}$ and $Calm1^{\Delta L/\Delta L}$ CA1 (**Fig.10a-c**). In home-cage control groups, less than 2% of cells in CA1 were cFos-positive. This is consistent with previous reports that showed cFos expression in the home cage condition is extremely low, with numerous sections containing no positive neurons (Drew et al. 2011). As expected, EE induced cFos expression in CA1 to 31.9 ± 12.5% in $Calm1^{+/+}$ mice (**Fig. 10b,c**). The EE induction of cFos expression in $Calm1^{\Delta L/\Delta L}$ mice was significantly reduced compared to wild type controls (19.8 ± 9.1%) (**Fig. 10b,c**,

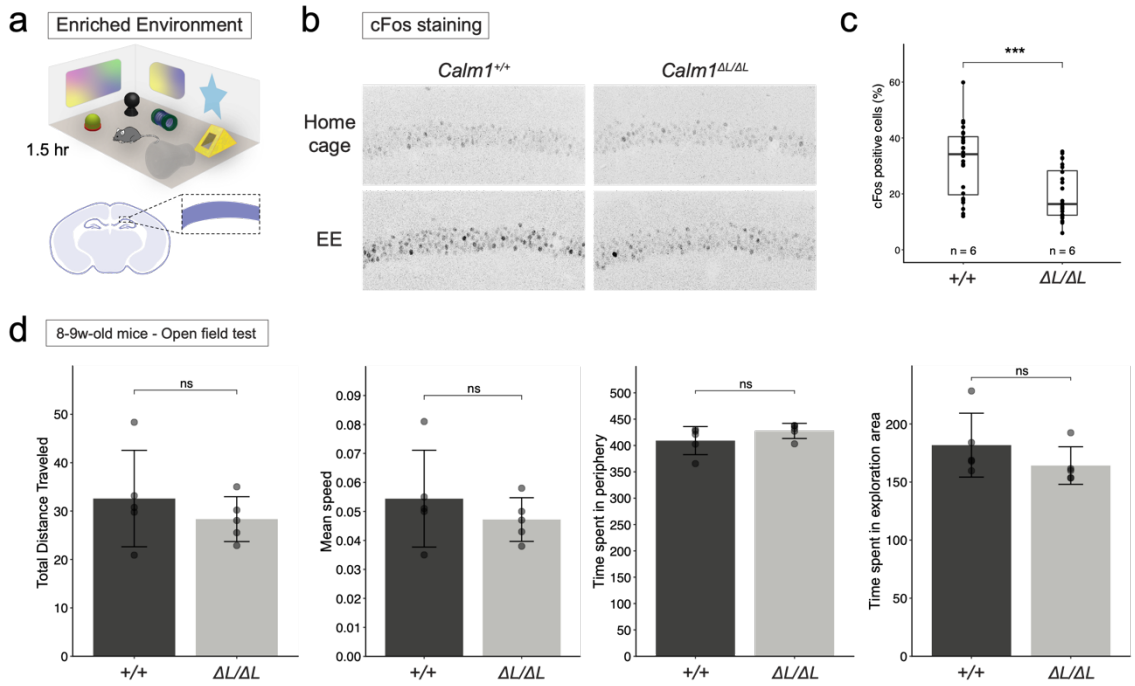


Figure 10. *Calm1-L* loss reduces hippocampal IEG expression in response to enriched environment exposure.
(a) Graphic illustration of enriched environment used in this study and the CA1 region shown in (b). (b) Representative images showing EE-induced expressions of cFos in *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ CA1. (c) Percentages of cFos-positive cells in whole CA1 region was quantified in FIJI/ImageJ. Significance was determined using a t-test, ***: p <= 0.001. n = 6 mice (2 hemispheres in 2 brain sections for each mouse). (d) Levels of total distance traveled, mean speed, time spent in periphery and exploration areas of *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ mice were compared using open field test. Significance was determined using a t-test in n = 4 mice; ns: p > 0.05.

Supplemental Fig. 4a). Analysis of Npas4, another IEG that is also activated in response to neuronal activity (Sun and Lin 2016), consistently showed reduced expression in the *Calm1*$^{\Delta L/\Delta L}$ hippocampus (**Supplemental Fig. 4b**).

The EE paradigm is influenced by multiple factors that mediate neural responses including social interaction, sensory stimulation, learning, and motor activity (Alwis and Rajan 2014). In order to exclude the possibility that the impaired locomotor function of

Calm1$^{\Delta L/\Delta L}$ mice might have impacted EE induced IEG expression, we performed an open field test. Both *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ genotypes showed similar levels of total ambulatory distance, mean speed, time spent in periphery, and time spent in exploration area (**Fig. 10d**). Thus, locomotor function and exploratory behavior were not altered by loss of Calm1-L under these conditions.

We also performed CaM western blot of hippocampal lysates from *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ mice exposed to the enriched environment (**Supplemental Fig. 4c**), and found total CaM protein levels to be unchanged. This result was similar to what was observed in Jones et al., where total CaM level in the hippocampus was unchanged in response to novel-context exposure (Jones et al. 2018). Again, our results show that total CaM level was not altered by the deletion of *Calm1-L* or by experience-dependent neuronal activation.

Calm1-L is less stable than *Calm1-S*

Unable to pinpoint the exact molecular defect responsible for the phenotypes in DRG and hippocampus resulting from *Calm1-L* loss, we turned our attention to the mRNA stability of *Calm1* 3′ UTR isoforms. A commonly used technique to estimate the half-life of mRNAs is tracing transcripts by qRT-PCR during a timeframe after blocking transcription. Due to *Calm1-S* and *Calm1-L* isoforms sharing a common short 3′ UTR region, we cannot quantify *Calm1-S* exclusively using qRT-PCR approaches in wild type tissue. To overcome this, we used *Calm1*$^{\Delta L/\Delta L}$ mice to measure the stability of *Calm1-S*. Extension qRT-PCR primers (ext) were used to detect *Calm1-L* from *Calm1*$^{+/+}$ samples, and universal qRT-PCR primers (uni) were used to uniquely detect *Calm1-S* from *Calm1*$^{\Delta L/\Delta L}$ samples (**Supplemental Fig. 5**). Hprt was quantified as a normalizing gene (La Fata et al. 2014). Primary cortical neurons were treated with actinomycin D and samples collected

immediately and at 3, 6, and 8 h post-treatment. An exponential regression equation was fitted to the relative abundance of each isoform across the timepoints as relative amount = e-kdecay*time and the half-life was estimated for each transcript. *Calm1-S* was found to have a longer half-life (t1/2= 5.9 ± 1.4 h) compared to *Calm1-L* (t1/2= 2.6 ± 0.6 h) (**Supplemental Fig. 5a**). When *Calm1* transcript levels were normalized to the stable *Hprt* transcripts, *Calm1-S* was consistently found to be more stable than *Calm1-L* (**Supplemental Fig. 5b**). To discard the possibility that $Calm1^{\Delta L/\Delta L}$ line might have overall RNA decay/stability pathways affected due to the genomic manipulation, we also measured the total *Calm1* levels using uni primers in the $Calm1^{+/+}$ samples. This again showed that the total *Calm1* in $Calm1^{+/+}$ samples has higher stability compared to the *Calm1-L* (**Supplemental Fig. 5c**). Thus, we conclude that *Calm1-S* isoform has a higher half-life compared to *Calm1-L* isoform.

Since enhanced RNA secondary structure affects mRNA half-life in mammalian systems (Sun et al. 2019), we examined predicted secondary structures in the long versus short 3′ UTRs (Bellaousov et al. 2013). We calculated the length-normalized minimum thermodynamic free energy (-ΔG/nt) of predicted RNA structures (Fischer et al. 2020) and estimated the structural complexity of the 3′ UTRs in *Calm1-S* and *Calm1-L* (**Supplemental Fig. 6**). This analysis suggested more complex RNA structures are found in the *Calm1-L* 3′ UTR (-ΔG/nt = - 0.297) compared to the *Calm1-S* 3′ UTR (-ΔG/nt = - 0.248). Thus, differential RNA structure could account for mRNA stability differences between *Calm1-S* and *Calm1-L.*

3. Discussion

Here, we successfully implemented CRISPR-Cas9 gene editing in mice to eliminate expression of a long 3′ UTR isoform while not altering the expression of its

corresponding short 3′ UTR isoform. To our knowledge, this is the first successful implementation of such an approach specifically for a long 3′ UTR isoform in vivo. We found that elimination of the *Calm1* long 3′ UTR isoform impaired both development of the DRG in embryos and activation of adult hippocampal neurons, thus establishing a functional role for *Calm1-L* in the peripheral and central nervous systems. Our method for long 3′ UTR deletion using CRISPR-Cas9 adds an important new tool for the characterization of APA-generated long 3′ UTR transcript isoforms.

Despite the prevalence of alternative length 3′ UTR isoforms in metazoan genomes, few physiological roles for these transcripts using loss-of-function genetic approaches have been identified. In a previous study, an in vivo neurological function for the *Bdnf* long 3′ UTR isoform was identified by generating a transgenic mouse that had tandem SV40 polyA sites inserted downstream of the proximal polyA signal to prevent biogenesis of the long 3′ UTR (An et al. 2008). Our strategy for generating loss of long 3′ UTR isoforms is less confounding because artificial regulatory sequences are not inserted into the genome. Including these strong regulatory sequences generates chimeric short 3′ UTR transcripts, which can affect cleavage and polyadenylation dynamics in unexpected ways.

We employed multiple gRNAs for CRISPR-Cas9 gene editing because it was not clear which, if any, of our deletion strategies would effectively prevent *Calm1* long 3′ UTR biogenesis. One concern was that removal of the distal polyA site might result in activation of downstream cryptic polyA sites. Remarkably, a single injection of this gRNA cocktail led to three different deletions which all prevented *Calm1-L* expression. We used mice generated by deletion strategy #3 for our analysis because this completely eliminated *Calm1-L* expression and only removed 164 bp of sequence encompassing the distal polyA site. However, the other two deletion strategies might be worth considering for removing long 3′ UTR isoforms for other genes. For example, one advantage of deletion #1 strategy

(which eliminates most of the long 3′ UTR and distal polyA site) is that it leaves no possibility for generating the long 3′ UTR isoform. However, implementing CRISPR for eliminating 3′ UTRs of extreme lengths (> 10 kb) is likely to be of very low efficiency. Deletion strategy #2 removes the sequence in between the proximal and distal polyA sites. The major confounding factor for this strategy is that an ectopic, truncated mRNA slightly longer than the short isoform is produced due to the distal polyA site remaining intact. Still, this approach could be useful because bringing two adjacent polyA sites in proximity is likely to prevent usage of downstream cryptic polyA sites.

Genome-wide analyses consistently have found *Calm1* to be subcellularly localized in neurons (Gumy et al. 2011; Tushev et al. 2018). In accordance with these data, our smFISH experiments show *Calm1* mRNA subcellular localization to processes in both DRG and hippocampus. A well characterized example of a gene with different subcellular localizations of 3′ UTR isoforms is *Bdnf*. *Bdnf-L* is dramatically enriched in dendrites compared to the soma restricted Bdnf-S, and removal of the long 3′ UTR isoform was found to prevent mRNA localization into dendrites (An et al. 2008). We had hypothesized that the long 3′ UTR of *Calm1* might serve a similar role– however, smFISH experiments revealed that both the short and long 3′ UTR isoforms were expressed in the processes of hippocampal neurons, and only the short 3′ UTR isoform was found in DRG axons (**Fig. 6**). Thus, cis-elements uniquely found in the long 3′ UTR of *Calm1* do not drive the axon/dendrite localization of *Calm1* mRNAs. Along these lines, recent genome-wide analysis of transcript isoform localization in somata versus neuropil similarly did not uncover a bias for longer 3′ UTR APA isoforms to be particularly localized to neuropil (Tushev et al. 2018).

Neural development requires execution of precise gene regulatory programs in response to extracellular cues, which results in cytoskeletal rearrangements allowing for

the formation of neural circuits (Tessier-Lavigne and Goodman 1996). In recent years, gene regulation at the post-transcriptional level has been emerging as an important aspect of axon growth and guidance (Holt and Schuman 2013; Zhang et al. 2019b, 2019a). In particular, there are many examples of local translation occurring in the developing growth cone in response to extracellular cues (Campbell and Holt 2001; Lin and Holt 2007; Shigeoka et al. 2013; Zivraj et al. 2010). Our data suggests that *Calm1-L* is important for C1 DRG cell body migration and axon extensions during development. This was evident from in vivo experiments that displayed dramatically disorganized cell bodies and abnormal axonal extensions in the C1 population of DRG in *Calm1*$^{\Delta L/\Delta L}$ embryos (**Fig. 8**). Future studies are needed to determine how *Calm1-L* is involved in DRG cell body positioning and axon outgrowth given that we did not observe a role for altered CaM protein nor altered axonal localization of *Calm1-L*.

EE-induced IEG expression was impaired in the hippocampus of our *Calm1*$^{\Delta L/\Delta L}$ mice. IEGs play key roles in synaptic plasticity, social and cognitive functions (Chung 2015; Coutellier et al. 2012). It has been shown that blocking CaM activity inhibits long-term potentiation (Malenka et al. 1989), supporting a critical role of CaM in synaptic plasticity. It is thus possible that loss of *Calm1-L* in *Calm1*$^{\Delta L/\Delta L}$ mice impairs synaptic plasticity and/or cognitive functions. Despite the impairment in IEG activation, we found that CaM levels were unchanged in the hippocampus of *Calm1*$^{\Delta L/\Delta L}$ mice. It is possible that *Calm1* is not regulated by activity-induced local translation. Alternatively, CaM translational changes resulting from *Calm1-L* loss might be restricted to certain subcellular regions, such as the activated synapse. Indeed, the activity-induced translation of *Calm1* has been supported by synaptoneurosome polysome fractionation experiments which showed a very small but significant enrichment of *Calm1* mRNAs in the heavy-polysome fraction after NMDAR-activation via NMDA and glutamate (Jones et al. 2018). Further investigation into the role

of *Calm1-L* in activity-induced synaptic translation of *Calm1* would be enabled by generating a mouse harboring an epitope tag fused to the *Calm1* coding sequence in the $Calm1^{+/+}$ and $Calm1^{\Delta L/\Delta L}$ backgrounds. Combined with spatiotemporal analysis of translation using a method such as Puro-PLA (Tom Dieck et al. 2015), such a transgenic line could be used to determine how *Calm1-L* 3′ UTR loss impacts activity-induced CaM translation specifically from the *Calm1* locus in vivo.

Finally, it is worth considering that the long 3′ UTR of *Calm1-L* might not be particularly important for cis-regulation of *Calm1* translation. Instead of impacting translation of their host genes, some long 3′ UTR isoforms act as scaffolds for assembling proteins that later form complexes with the protein being translated (Lee and Mayr 2019; Brigidi et al. 2019). Long 3′ UTRs have non-coding functions that are completely removed from the functions of the encoded protein. For example, long 3′ UTR isoforms of *Ube3a1* functions as a miRNA sponge in the synaptodendritic regions independent of its protein-coding ability, and a non-translatable long 3′ UTR isoform of *Tp53inp2* is involved in TrkA receptor internalization in axons (Valluy et al. 2015; Crerar et al. 2019).

The mechanism of how *Calm1* long 3′ UTR isoform loss impairs neuronal development and function remains unclear. Nonetheless, our approach has uncovered neural phenotypes both in the PNS and CNS resulting from the loss of a long 3′ UTR isoform in mice. We have demonstrated an effective strategy for eliminating long 3′ UTR isoforms via CRISPR/Cas9 gene editing without affecting short 3′ UTR expression. Thousands of alternative long 3′ UTR isoforms documented in mice are of unknown physiological relevance (Miura et al. 2013). The CRISPR approach described here can be used to generate deletion strains on a gene-by-gene basis to determine which long 3′ UTR isoforms have in vivo functional relevance.

4. Materials and Methods

Animal Use and Tissue Collection

All mice were housed in an environmentally controlled facility under the supervision of trained laboratory support personnel. Animal protocols were approved by the University of Nevada, Reno Institutional Animal Care and Use Committee (IACUC) and in accordance with the standards of the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

For embryo collection, crossed female mice were monitored daily for vaginal plugs, with positive identification being counted as E0.5 at noon that day. Pregnant mice were euthanized at noon using CO_2 asphyxiation, and then cervical dislocation was performed. E10.5 and E13.5 embryos were extracted in PBS. Embryos were immediately used for fresh tissue collection or fixed in 4% PFA in PBS by immersion. Postnatal day 0 (P0) – P1 pups were euthanized by decapitation and tissues were collected in cold PBS or HBSS. For adult tissue RNA or protein extraction, dissected tissue was flash frozen in liquid nitrogen and stored at – 80 °C or immediately used. To obtain fixed tissues, 7-9 weeks-old animals were euthanized by overdose of isoflurane inhalation followed by transcardial perfusion with abundant PBS and 4% PFA in PBS.

RT-qPCR and Northern Analysis

Flash frozen tissue was pulverized using a Cellcrusher™ tissue pulverizer. RNA was then extracted using the RNeasy Plus Universal Mini Kit (Qiagen) or Trizol (Qiagen) method and quantified using a NanoDrop spectrophotometer. For RT-qPCR, 1 or 2 µg of RNA was reverse transcribed using SuperScript III Reverse Transcriptase (Invitrogen) or Maxima Reverse Transcriptase (Invitrogen). The cDNA reaction was diluted 5-fold in ultrapure water for use in RT-qPCR. RT-qPCR was performed using SYBR™ Select

Master Mix for CFX (Applied Biosystems). The BioRad CFX96 real time PCR machine was used to carry out real time PCR and results were analyzed using the delta-delta CT method. For Northern analysis, PolyA+ RNA was extracted from total RNA using NucleoTrap mRNA kit (Machery-Nagel). Northern Blot analysis was performed as previously described (Miura et al. 2013). Briefly, polyA+ RNA samples (2 µg) was denatured in glyoxal and run in BPTE gels prior to downward transfer followed by Northern blotting using 32-P dCTP labeled DNA probes (sequences of primers used to generate probes are found in Supplemental Table 1-1). Blots were exposed overnight until desired intensity of signals was detected using Typhoon FLA7000 phosphoimager (GE).

Digoxigenin in situ hybridization

Riboprobes were generated via in vitro transcription using DIG RNA labeling mix (Roche) for the same probe regions as in Northern blot analysis. Sucrose cryoprotected E13.5 embryos or adult brains were embedded in O.C.T compound and cryosectioned at 16 -18 μm. Sections were treated in antigen retrieval solution for 5 minutes in 95°C water bath and washed in water twice. To aid permeabilization, slides were immersed in ice-cold 20% (v/v) acetic acid for 20 sec. Upon dehydration in sequential washes in 70 – 100% ethanol, slides were stored at -20°C until use. Approximately 30 ng riboprobes were used in each hybridization at 65°C overnight. Stringent washes were performed in SSC buffer (50% formamide in 2x SSC and 0.5x SSC). Anti-DIG-AP Fab antibody (Roche) incubation was performed in 2% BSA in MABT solution for 1h at room temperature. Color development was carried out using NBT and BCIP (Roche) at room temperature until desired coloring was observed (between 16 to 20 hours). Leica DM IL LED microscope or Keyence BZ-X710 all in one fluorescence microscope was used for imaging and stitched automatically.

RNA-seq analysis for 3′ UTR alternative polyadenylation identification

RNA-seq datasets for neuron, microglia, astrocyte, oligodendrocyte, and endothelial cells were downloaded from GEO accession series GSE52564, and datasets for DRG and isolated axons from GSE51572. Datasets were aligned and processed using HISAT2, Samtools, and Bamtools. Bam or bedgraph files were loaded into IGV for track visualization. For 3′ UTRs alternative polyadenylation analysis, 3′ UTRs reference was built according to QAPA algorithm with the exception of slight modification for *Calm1* 3′ UTRs. Original QAPA reference contained 6 *Calm1* 3′ UTR isoforms, but unconfirmed shorter isoforms were omitted for accurate quantification of only *Calm1-S* and *Calm1-L* expressions. Isoform abundance was quantified using Sailfish and QAPA quant algorithms. PAU values from the output file were used for comparison of 3′ UTR expression in each cell types.

Primary neuron culture

DRGs from E13.5 mouse embryos were dissected in Neurobasal medium. Cells were dissociated with 0.25% trypsin and triturated with a gel-loading tip. Dissociated cells were plated on PDL and laminin coated compartmentalized chamber (Xona microfluidics, XC450) or coverslips in culture medium (DMEM supplemented with 1X Glutamax, 10% FBS, 25 ng/ml NGF, and 1X pen/strep). In particular for the compartmentalized chamber, the somal compartment was coated only with Poly-D-lysine (PDL) and supplemented with medium containing 10 ng/ml NGF instead of 25 ng/ml to promote the growing of axons through the microgroove and toward the axonal compartment. The cultures were maintained until DIV2 at 37 ºC with 5% CO_2.

Cortices or hippocampi from P0 – P1 mouse pups were dissociated with 0.25% trypsin and plated in plating media (MEM supplemented with 0.5% w/v glucose, 0.2 mg/ml $NaHCO_3$, 0.1 mg/ml transferrin, 10% FBS, 2 mM L-glutamine, and 0.025 mg/ml Insulin) onto PDL coated 6-well plates. After one day in vitro (DIV), the media was replaced with growth media (hippocampal- MEM supplemented with 0.5% w/v glucose, 0.2 mg/ml $NaHCO_3$, 0.1 mg/ml transferrin, 5% FBS, 0.5 mM L-glutamine, and 2% B-27 supplement; cortical- Neurobasal media supplemented with 2% B27 supplement and 1X GlutaMax). 4 µM AraC (Sigma-Aldrich) was added to the growth media at DIV1 - 2 to suppress the growth of non-neuronal cells and enrich post-mitotic neurons. The cultures were maintained until DIV7 at 37 ºC with 5% CO_2.

Single molecule fluorescence in situ hybridization

Single molecule FISH (smFISH) was performed in primary neurons using Stellaris (LGC Biosearch Technologies) or RNAscope multiplex fluorescent assay according to manufacturer's instruction. Briefly, Stellaris smFISH was carried out according to manufacturer's protocol for adherent cells except that Wash buffer A was replaced with 2X SSC and 10% formamide in ultrapure water, Hybridization buffer with 2X SSC, 10% formamide, 10X dextran-sulfate, and 250 nM probe. Hybridization was carried out for ~18-24 hrs in the dark at 37 °C. For RNAscope smFISH, cells were pre-treated with permeabilization solution and protease solution, and incubated with 1X probe solution (551281-C3 or C1 for extension probe and 556541-C2 for universal probe diluted in probe diluent) for 2 hours at 40ºC. Subsequently, Amp1-FL to Amp4-FL were hybridized to amplify FISH signals. To combine FISH and immunofluorescence, immunostaining procedure was performed after the Amp4-FL hybridization step. Anti-Tubb3 primary and secondary antibodies were incubated for 1 h at 37 ºC and RT, respectively. Samples were

counterstaining with DAPI and mounted in anti-fade buffer (10 mM Tris pH 8.0, 2X SSC, and 0.4% glucose in water) or ProLong diamond antifade mountant (Invitrogen). During image acquisition, control neuron slide that had been incubated in probe diluent without probes was used to set laser intensity to ensure no background signal was counted in experimental conditions. Laser intensity and detector range were kept constant among different conditions. FISH-quant (Mueller et al. 2013) was used for puncta counting of FISH signals and FIJI/ImageJ was used to measure the distance of travel of each FISH signal from the center of the nucleus.

RNA Stability Assays and Structure Predictions

Cortical neurons were treated with 1 μg/ml Actinomycin D at DIV6. Total RNA was collected in Trizol at 0, 3, 6, and 8 hours post-treatment. 1μg of RNA was reverse transcribed using Maxima Reverse Transcriptase (Invitrogen). RT-qPCR was performed as described above. To compare the relative stability of *Calm1-S* and *Calm1-L* isoforms, expression levels of each transcript for each time point were calculated relative to 0 hours using BioRad CFX Manager software and normalized to Hprt (a stable transcript control). To estimate the half-life of each transcript, expression level for each time point was calculated relative to 0 hours without normalizing to Hprt. Exponential regression equations were fitted for each degradation plot. Half-life of each transcript was calculated using Goal-seek function in Excel to find the time point where the relative amount of transcript was reduced to 50%. The half-life of each biological replicate was used to determine mean and standard deviation for each transcript.

RNA structure prediction was performed in RNAstructure Web server (Bellaousov et al. 2013) with default setting and -ΔG/nt was calculated using the energy value given from the analysis and the length of the 3′ UTR (Fischer et al. 2020).

Western analysis

For conventional western blot analysis, total protein was extracted in RIPA buffer supplemented with protease inhibitor tablet (Pierce). Protein samples were separated in 15% discontinuous SDS-PAGE gel and transferred onto 0.2 µm PVDF membrane (Trans-Blot Turbo, Biorad). Membranes were blocked in 5% skim milk followed by overnight incubation with primary antibodies at 4 °C. Anti-CaM (Abcam 45689) and anti-α-tubulin (Sigma T9026) antibodies were used at 1:500-1,000 and 1:2,000, respectively. HRP conjugated secondary antibody incubation was performed at room temperature for 1 hour. HRP signal was detected using ProSignal Femto reagent (Prometheus) and imaged using a ChemiDoc Touch (Biorad).

For low-input western blot analysis, capillary western system (Wes™, Protein simple) was used. 500 ng of total protein lysate was loaded per capillary to visualize proteins of interest. Anti-CaM and anti-α-tubulin were used at 1:50 dilution. Quantification of band intensity was automatically estimated using Compass for SW software.

CRISPR-Cas9 gene editing

The MIT CRISPR Design Tool (http://www.genome-engineering.org/crispr) was used to design gRNAs targeted to various regions of the *Calm1* long 3′ UTR (Supplemental Table 1). Sequences were cloned into the BbsI site of pX330-U6 chimeric BB-CBh-hSpCas9 plasmid (42230; Addgene). HiScribe T7 mRNA synthesis kit (New England BioLabs) was used to in vitro transcribe the gRNAs, and RNAs were subsequently purified using RNA Clean & Concentrator™-5 (Zymo Research, Cat. R1016) before assessment on an Agilent Bioanalyzer as described (Han et al. 2015).

Super-ovulating 4-6-week-old FVB/NJ female mice were mated with C57BL/6J males. After fertilization, the eggs were collected from the oviducts. Mouse zygotes were microinjected into the cytoplasm with Cas9 mRNA (100 ng/μL) and the gRNA (100 ng/μL each) as described previously (Halim et al. 2017; Han et al. 2015; Oliver et al. 2015). After injection, zygotes were cultured in KSOM+AA medium (Millipore) for 1 h at 5% CO_2 before transfer into the 7–10-week-old female CD1 foster mothers.

Genomic DNA was isolated from tail-snips of founder mice by overnight proteinase K digestion. Two sets of primers flanking the deletion regions were used for PCR genotyping to detect the deletion using Taq Polymerase (NEB). Sibling founder mice were crossed together to generate the F1 generation. An F1 male harboring both the Deletion 2 and Deletion 3 alleles were crossed to C57BL/6 female mice to establish separate Deletion 2 and Deletion 3 lines. The F1 Deletion 1 animals were first crossed to heterozygous littermates (Del1/WT x Del1/WT). An F2 Del1 heterozygote was crossed to C57BL/6 female mice to establish the Deletion 1 line. All PCR products were resolved in 1% agarose gels. To confirm the deletions and check for any insertion events Sanger sequencing was of gel purified PCR products was performed (Nevada Genomics Center, University of Nevada, Reno). For the experiments in this paper, Deletion 3 line was backcrossed with C57BL/6 at least three times to stabilize the background and heterozygote pairs were interbreed to obtain *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ unless otherwise stated.

Immunofluorescence

For whole-mount analysis, small tears were made in embryos in the forebrain and roof plate of the hindbrain to facilitate the penetration of PFA into the neural tube. Embryos were fixed by overnight incubation in 4% PFA at 4˚C. Embryos were blocked and

permeabilized for antibody labeling by performing 3 x 5 min, 3 x 30 min, and then overnight washes of PBST (10% FBS and 1% Triton X-100 in 1X PBS). Primary antibodies (Biolegend cat#801201 anti-β3Tubulin; 1:1000) were then added and incubated for 2 nights rocking at 4 ˚C. Embryos were again washed for 3 x 5 min, 3 x 30 min, and then overnight in fresh PBST. Secondary antibodies were added at a 1:200 dilution for 2 nights at 4˚C rocking in the dark. Secondary antibodies were washed off again 3 x 5 min, 3 x 30 min, and then overnight in fresh PBST. Embryos were then immersed in 80% glycerol/PBS for a minimum of 2 hours, were mounted, and then imaged whole mount. A Leica SP8 TCS confocal microscope was used for imaging. For whole mount samples, images were taken at 1048 X 1048px, 200 hz scanning speed, 3-line averaging, 8 bit, and with the use of linear-z compensation.

For free-floating brain section staining, PFA perfused brains were post-fixed in 1% PFA for 2 nights and sectioned using vibratome. 80 µm thick sections were immediately used or stored at -20 C in antifreeze solution (30% w/v sucrose, 1% w/v PVP-40, 30% v/v ethylene glycol in PBS) until use. Sections were blocked in blocking solution for 2 hours at RT, primary antibodies (cFos SYSY 226003; Npas4 Activity Signaling AS-AB18A) were incubated overnight at 4˚C rocking in the dark. 10 min x 3 washes were performed in PBS followed by fluorophore labeled secondary antibody incubation at RT for 2 hrs. During 10 min x 3 washes, the nuclei were counterstained with DAPI. Sections were mounted in ProLong diamond antifade mountant (Invitrogen). Sections were cured for at RT for least 3 hours and stored at 4˚C until imaging. A Leica SP8 TCS confocal microscope was used for imaging at 1024 X 1024 px, 20x magnification, 3-line average and images were stitched automatically.

DRG migration in vivo analysis

Whole mount E10.5 embryos were labeled with anti-β3Tubulin (Biolegend Cat# 801201 1:1000) and were subsequently imaged by confocal microscopy. Images were analyzed using FIJI/ImageJ (NIH). The cell bodies to be quantified was defined by drawing a line parallel to the last branch of the hypoglossal nerve extending across the length of the image. The hypoglossal nerve was chosen as an anatomical landmark because its morphology was unaffected in mutant embryos and served as an independent reference point so measurements were consistent across embryos. The area and distance migrated of cell bodies and axons rostral to the hypoglossal branch were measured using a freehand selection tool in FIJI/ImageJ. The multipoint cell counting tool was used on FIJI/ImageJ to count the number of axon bundles branching off the C1 DRG. A two-sided student's t-test was performed to test for significance.

Enriched environment-induced neural activation assay

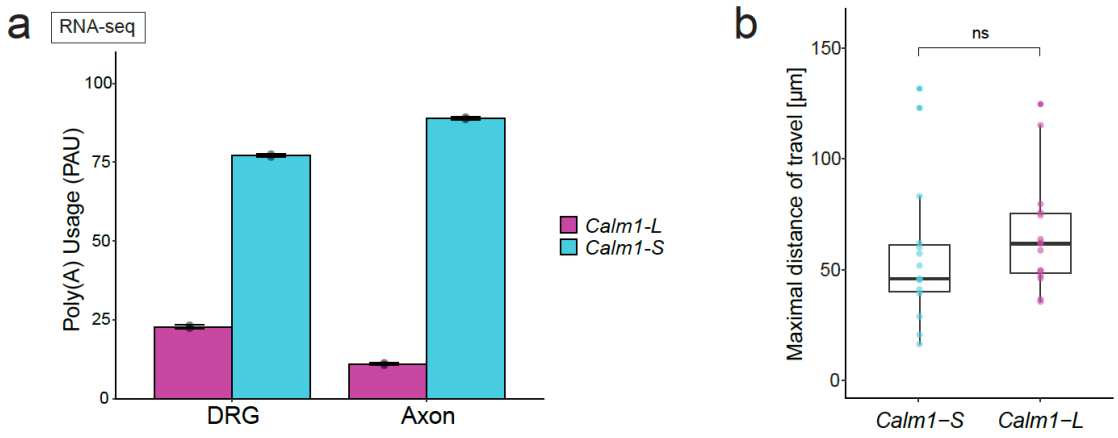
Enriched environment (EE) was created in a 83.8 cm x 51.1 cm x H 34.3 cm box with sufficient bedding material at the bottom. Different types of toys were placed in the box with the sides of the box covered with colorful papers as illustrated in **Fig 6.a**. Age matched *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ naïve animals (not used for other tests) were kept in their home cage undisturbed for at least 3 days prior the experiment and transferred to the EE box on the test day. Mice were exposed to the EE for 1.5 hours before euthanasia and tissue fixation. Brain sections were analyzed for IEG expression induction using free-floating immunofluorescence staining as described above. IEG positive cells were counted in the entire CA1 regions using FIJI/ImageJ. Images were adjusted for levels and color balance automatically and pixels were smoothened. Then, image thresholding was performed using Huang's method. After applying watershed function to separate closely

located cells, particle analysis was performed in CA1 area to count cells with pixel units above 40. The parameters were set through repetitive rounds of visual inspection that allowed the most accurate detection of IEG positive cells and applied equally across all the images. The number of IEG positive cells were then compared to the total DAPI counts in the same CA1 region to estimate the percentage of activated neurons.

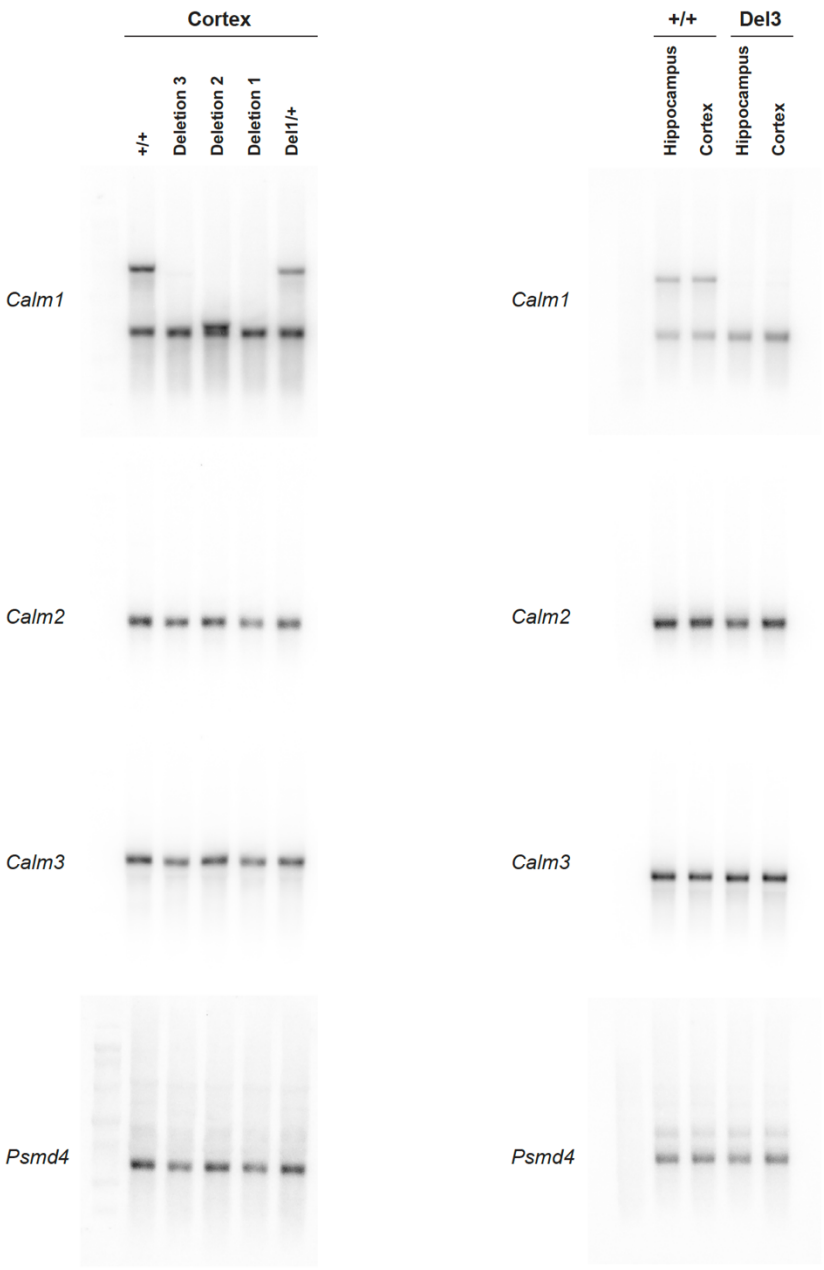
Open-field assay

To assess locomotor function and exploratory behavior of *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ mice, open field assay was performed using ANY-maze video tracking system. Day before the assay, age-matched naïve animals were placed in the behavior room in their home cage for at least 30 minutes and gently handled for a couple of minutes to minimize distress of animals during the test. The animals were returned to the housing room. On the test day, the animals were brought in their home cages into the behavior room 30 min prior to starting the test. Overhead lightning set-up was used to prevent shadows. Single animal was placed in the periphery of the arena facing the wall, and its behavior was recorded for 10 minutes. The apparatus was cleaned with 70% ethanol and dried for 5 min between each testing session. Total distance traveled by each mouse, mean speed, time spent in periphery and exploration area were obtained to compare the behavior between *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ mice using ANY-maze software.

5. Supplementary Data



Supplemental Figure 1. Subcellular localization of *Calm1-L*.
(a) re-analysis of RNA-seq data obtained from isolated axons of DRG/spinal cord (Minis et al. 2014). QAPA analysis revealed the relative expression of the long 3′ UTR to be 23% in whole DRG. In isolated axons, there was less relative expression of the long 3′ UTR (11%). (b) Quantification of the maximum distance of travel for each of the *Calm1* isoforms in DRG. Significance was determined using t-test, ns: $p > 0.05$. n = 15 neurons. Related to Figure 2.



Supplemental Figure 2. Northern blots.
Each deletion strategy had the desired outcome in both mouse cortex and hippocampus. Transcripts from the *Calm2* and *Calm3* genes were found to be unaltered in both tissues. The same blots used for *Calm1* were stripped and re-probed for each gene. Related to Figure 3b (left side) and Figure 5a (right side).

a

	crRNA	DNA	Chromosome	Position	Direction	#Bulge type	Mismatches	Bulge Size	Offtarget gene
1	TTTGTCTCGTATACCGCAAANGG	TTTGTCTCGTATACCGCAAAGGG	chr12	100209858	-	X	0	0	
2	ACGCAACATCCAGTGCAAGCNGG	ACGCAACATCCAGTGCAAGCTGG	chr12	100207399	+	X	0	0	
3	GAGTTGGTCCCCCGTCAGAGNGG	GAGTTGGTCCCCCGTCAGAGAGG	chr12	100209715	-	X	0	0	
4	TTTGTCTCGTATACCGCAAANGG	TTTGTCTCGTATACCGC-AAAGG	chr12	100209859	-	RNA	0	1	
5	TTTGTCTCGTATACCGCAAANGG	TTTGTCTCGTATACCGCA-AAGG	chr12	100209859	-	RNA	0	1	
6	TTTGTCTCGTATACCGCAAANGG	T-TGTCTCGTATACCGCAAAGGG	chr12	100209858	-	RNA	0	1	
7	TTTGTCTCGTATACCGCAAANGG	TT-GTCTCGTATACCGCAAAGGG	chr12	100209858	-	RNA	0	1	
8	AC--GCAACATCCAGTGCAAGCNGG	ACACGCAACATCCAGTGCAAGCTGG	chr12	100207397	+	DNA	0	2	
9	A--CGCAACATCCAGTGCAAGCNGG	ACACGCAACATCCAGTGCAAGCTGG	chr12	100207397	+	DNA	0	2	
10	GAGTTGGTCCCCCGTCAGAGNGG	G--TTGGTCCCCCGTCAGAGAGG	chr12	100209715	-	RNA	0	2	
11	TTTGTCTCGTATACCGCAAANGG	T--GTCTCGTATACCGCAAAGGG	chr12	100209858	-	RNA	0	2	
12	ACGCAACATCCAGTGCAAGCNGG	c-GCAACATCCAGTGCAAGCTGG	chr12	100207400	+	RNA	1	1	
13	GAGTTGGTCCCCCGTCAGAGNGG	a-GTTGGTCCCCCGTCAGAGAGG	chr12	100209715	-	RNA	1	1	
14	TT-TGTCTCGTATACCGCAAANGG	gTTTGTCTCGTATACCGCAAAGGG	chr12	100209858	-	DNA	1	1	
15	T-TTGTCTCGTATACCGCAAANGG	gTTTGTCTCGTATACCGCAAAGGG	chr12	100209858	-	DNA	1	1	
16	TTT-GTCTCGTATACCGCAAANGG	gTTTGTCTCGTATACCGCAAAGGG	chr12	100209858	-	DNA	1	1	
17	A-CGCAACATCCAGTGCAAGCNGG	cACGCAACATCCAGTGCAAGCTGG	chr12	100207398	+	DNA	1	1	
18	TTTGTCTCGTATACCGCAAANGG	TTTGTCTCGTATACCG-cAAAGG	chr12	100209859	-	RNA	1	1	
19	TTTGTCTCGTATACCGCAAANGG	TTg-TCTCGTATACCGCAAAGGG	chr12	100209858	-	RNA	1	1	
20	G-AGTTGGTCCCCCGTCAGAGNGG	aGAGTTGGTCCCCCGTCAGAGAGG	chr12	100209715	-	DNA	1	1	
21	GCTTGCACTGGATGTTGCGTNGG	GCTTGCACTGGgTGTT--GTGGG	chr8	39376664	+	RNA	1	2	Intergenic
22	GCTTGCACTGGATGTTGCGTNGG	GCTTGCACTGGgTGTTG--TGGG	chr8	39376664	+	RNA	1	2	Intergenic
23	T--TTGTCTCGTATACCGCAAANGG	gGTTTGTCTCGTATACCGCAAAGGG	chr12	100209858	-	DNA	1	2	
24	G--AGTTGGTCCCCCGTCAGAGNGG	tAGAGTTGGTCCCCCGTCAGAGAGG	chr12	100209715	-	DNA	1	2	
25	ACGCAACATCCAGTGCAAGCNGG	g--CAACATCCAGTGCAAGCTGG	chr12	100207401	+	RNA	1	2	
26	ACGCAACATCCAGTGCAAGCNGG	gC--AACATCCAGTGCAAGCTGG	chr12	100207401	+	RNA	1	2	
27	GA--GTTGGTCCCCCGTCAGAGNGG	tAGAGTTGGTCCCCCGTCAGAGAGG	chr12	100209715	-	DNA	1	2	
28	ACG--CAACATCCAGTGCAAGCNGG	ACaCGCAACATCCAGTGCAAGCTGG	chr12	100207397	+	DNA	1	2	
29	CTCTGACGGGGGACCAA--CTCNGG	CTCTGACGGGGGACCAACTCTaTGG	chr12	100209718	+	DNA	1	2	
30	GAGTTGGTCCCCCGTCAGAGNGG	Gt--TGGTCCCCCGTCAGAGAGG	chr12	100209715	-	RNA	1	2	
31	ACGC--AACATCCAGTGCAAGCNGG	ACaCGCAACATCCAGTGCAAGCTGG	chr12	100207397	+	DNA	1	2	
32	TTTGTCTCGTATACCGCAAANGG	Tg--TCTCGTATACCGCAAAGGG	chr12	100209858	-	RNA	1	2	
33	TTTGTCTCGTATACCGCAAANGG	TgT--CTCGTATACCGCAAAGGG	chr12	100209858	-	RNA	1	2	
34	CTCTGACGGGGGACCAAC--TCNGG	CTCTGACGGGGGACCAACTCTaTGG	chr12	100209718	+	DNA	1	2	
35	GAG--TTGGTCCCCCGTCAGAGNGG	tAGAGTTGGTCCCCCGTCAGAGAGG	chr12	100209715	-	DNA	1	2	
36	CTCTGACGGGGGACCAACT--CNGG	CTCTGACGGGGGACCAACTCTaTGG	chr12	100209718	+	DNA	1	2	
37	CTCTGACGGGGGACCAACTCNGG	CTCTGA--GGGGtCCAACTCAGG	chr10	62679358	-	RNA	1	2	Stox1 intron
38	TTTGTCTCGTATACCGCAAANGG	TT--TCTCGTATACaGCAAAAGG	chrX	134330263	+	RNA	1	2	Cenpi intron
39	TTTGTCTCGTATACCGCAAANGG	TTT--CTCGTATACaGCAAAAGG	chrX	134330263	+	RNA	1	2	Cenpi intron
40	GAGTTGGTCCCCCGTCAGAGNGG	cAGTTGGTCCCCCG-aAGAGTGG	chr8	121767423	+	RNA	2	1	Jph3 intron
41	GCTTGCACTGGATGTTGCGTNGG	GCTTGCgCTtGATGTTGC-TCGG	chr8	71465346	-	RNA	2	1	Mrpl34 exon
42	GAGTTGGTCCCCCGTCAGAGNGG	GAG-TGaTCCtCCGTCAGAGGGG	chr8	89063789	+	RNA	2	1	Intergenic
43	GAGTTGGTCCCCCGTCAGAGNGG	GAGT-GaTCCtCCGTCAGAGGGG	chr8	89063789	+	RNA	2	1	Intergenic
44	ACGCAACATCCAGTGCAAGCNGG	AgaCAACATCCAGTG-AAGCTGG	chr8	76542388	-	RNA	2	1	Intergenic
45	ACGCAACATCCAGTGCAAGCNGG	AgGCAACATaCAGTG-AAGCAGG	chr8	124262803	+	RNA	2	1	Galnt2 intron
46	GAGTTGGTCCCCCGTCAGAGNGG	GAGTcGGTCCtCCGT-AGAGTGG	chr8	71545799	-	RNA	2	1	Mvb12a exon
47	GAGTTGGTCCCCCGTCAGAGNGG	cAGTTGGTCCCCCGa-AGAGTGG	chr8	121767423	+	RNA	2	1	Jph3 intron
48	ACGCAACATCCAGTGCAAGCNGG	ACtCAAgATCCAGT-CAAGCTGG	chr8	104115432	+	RNA	2	1	Cdh5 intron
49	TTTGTCTCGTATACCGCAAANGG	TTgt-CTCGTATACCGCAAAGGG	chr12	100209858	-	RNA	2	1	
50	GAGTTGGTCCCCCGTCAGAGNGG	ag-TTGGTCCCCCGTCAGAGAGG	chr12	100209715	-	RNA	2	1	
51	ACGCAACATCCAGTGCAAGCNGG	cg-CAACATCCAGTGCAAGCTGG	chr12	100207400	+	RNA	2	1	
52	TTTG-TCTCGTATACCGCAAANGG	gTTtGTCTCGTATACCGCAAAGGG	chr12	100209858	-	DNA	2	1	
53	GA-GTTGGTCCCCCGTCAGAGNGG	agAGTTGGTCCCCCGTCAGAGAGG	chr12	100209715	-	DNA	2	1	
54	TTTGTCTCGTATACCGCAAANGG	TTTGTCTCGTATACC-gcAAAGG	chr12	100209859	-	RNA	2	1	
55	CTCTGACGGGGGACCAACTCNGG	CTtTGAcTGGGG-CCAACTCAGG	chr12	76602542	+	RNA	2	1	Sptb intron
56	AC-GCAACATCCAGTGCAAGCNGG	caCGCAACATCCAGTGCAAGCTGG	chr12	100207398	+	DNA	2	1	
57	ACGCAACATCCAGTGCAAGCNGG	AaGCAACATCCAGT-CAAGaTGG	chr3	126820828	+	RNA	2	1	Camk2d intron
58	TTTGCGGTATACGAGACAAANGG	TTTGaGGTATAC-AGACAAtTGG	chr3	72454474	+	RNA	2	1	Intergenic
59	CTCTGACGGGGGACCAACTCNGG	CTCTGAC-GGGGACgAgCTCGGG	chr3	146121860	-	RNA	2	1	Mcoln3 exon
60	CTCTGACGGGGGACCAACTCNGG	CTCTGACG-GGGACgAgCTCGGG	chr3	146121860	-	RNA	2	1	Mcoln3 exon

b

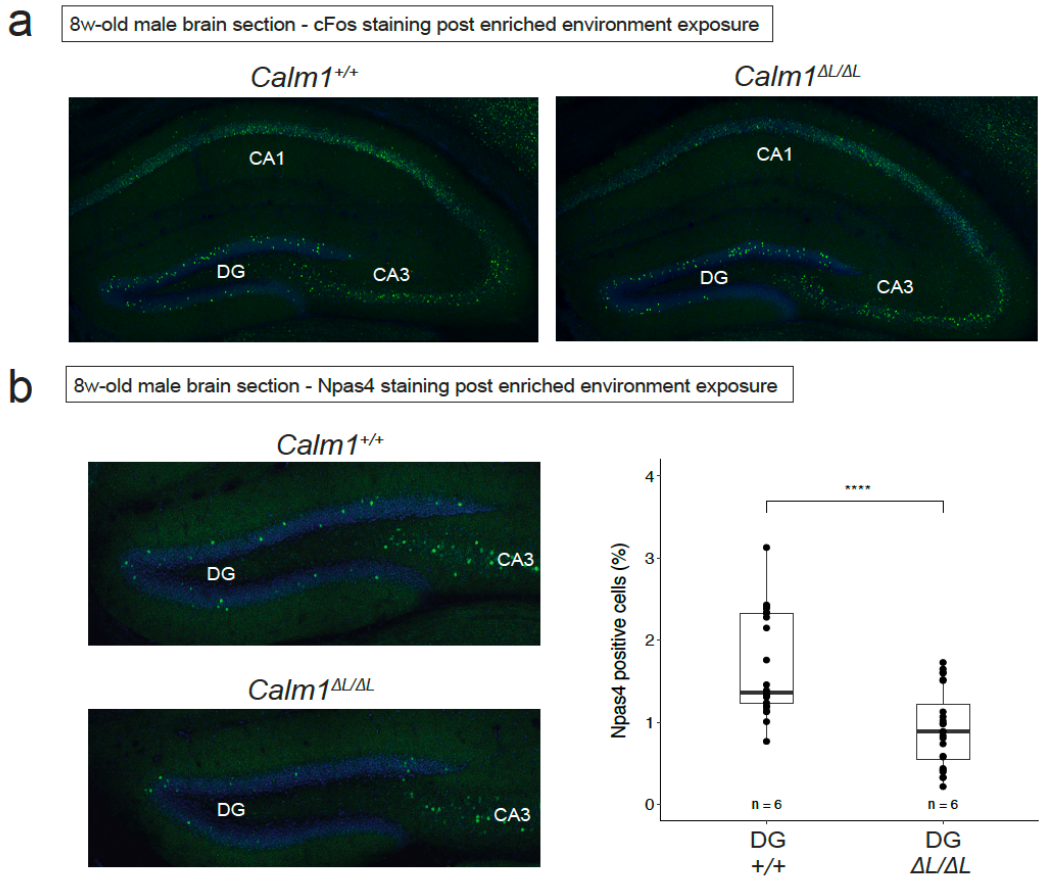
```
                chr8:124,291,290(mm9)                                  G          TC
                |                                  gRNA#4  AGTTG GTCCCCCG A GAGNGG
>BL6_Jph3       ACTGCTGCTC CCCTTCCTGC CCATGTCTGG CTATCAGTTG GTCCCCCGAA GAGTGGATCA ACTTGAACCT TCATCCCCAG CGTAGCATTC
>WT#839_Jph3    ACTGCTGCTC CCCTTCCTGC CCATGTCTGG CTATCAGTTG GTCCCCCGAA GAGTGGATCA ACTTGAACCT TCATCCCCAG CGTAGCATTC
>WT#521_Jph3    ACTGCTGCTC CCCTTCCTGC CCATGTCTGG CTATCAGTTG GTCCCCCGAA GAGTGGATCA ACTTGAACCT TCATCCCCAG CGTAGCATTC
>ΔL#838_Jph3    ACTGCTGCTC CCCTTCCTGC CCATGTCTGG CTATCAGTTG GTCCCCCGAA GAGTGGATCA ACTTGAACCT TCATCCCCAG CGTAGCATTC
>ΔL#519_Jph3    ACTGCTGCTC CCCTTCCTGC CCATGTCTGG CTATCAGTTG GTCCCCCGAA GAGTGGATCA ACTTGAACCT TCATCCCCAG CGTAGCATTC

                chr10:62,142,151(mm9)                                C     GA
(Rev Comp seq)  |                          gRNA#3 CTCTGA GGGG CCAAC TCNGG
>BL6_Stox1      GTGGTTCTCC TTTCCACCAT GTGGCTCTGA GGGGTCCAAC TCAGGTTTTC AGGCTTGGTG GCAGTTCCCT TTACCTGCTG GGTCATCTTG
>WT#839_Stox1   GTGGTTCTCC TTTCCACCAT GTGGCTCTGA GGGGTCCAAC TCAGGTTTTC AGGCTTGGTG GCAGTTCCCT TTACCTGCTG GGTCATCTTG
>WT#521_Stox1   GTGGTTCTCC TTTCCACCAT GTGGCTCTGA GGGGTCCAAC TCAGGTTTTC AGGCTTGGTG GCAGTTCCCT TTACCTGCTG GGTCATCTTG
>ΔL#838_Stox1   GTGGTTCTCC TTTCCACCAT GTGGCTCTGA GGGGTCCAAC TCAGGTTTTC AGGCTTGGTG GCAGTTCCCT TTACCTGCTG GGTCATCTTG
>ΔL#519_Stox1   GTGGTTCTCC TTTCCACCAT GTGGCTCTGA GGGGTCCAAC TCAGGTTTTC AGGCTTGGTG GCAGTTCCCT TTACCTGCTG GGTCATCTTG

                chr3:126,523,711(mm9)                                  C  G    C
                |                                  gRNA#1 A GC AACATCCAGT CAAG NGG
>BL6_Camk2d     CATGTGAGGG GTAGGAGACG GCCGGGGAGC AAGGGGAAGC AACATCCAGT CAAGATGGCC AAGGCAGACT TGGTCTGTTT TTGTTGCAGG
>WT#521_Camk2d  CATGTGAGGG GTAGGAGACG GCCGGGGAGC AAGGGGAAGC AACATCCAGT CAAGATGGCC AAGGCAGACT TGGTCTGTTT TTGTTGCAGG
>ΔL#519_Camk2d  CATGTGAGGG GTAGGAGACG GCCGGGGAGC AAGGGGAAGC AACATCCAGT CAAGATGGCC AAGGCAGACT TGGTCTGTTT TTGTTGCAGG
```

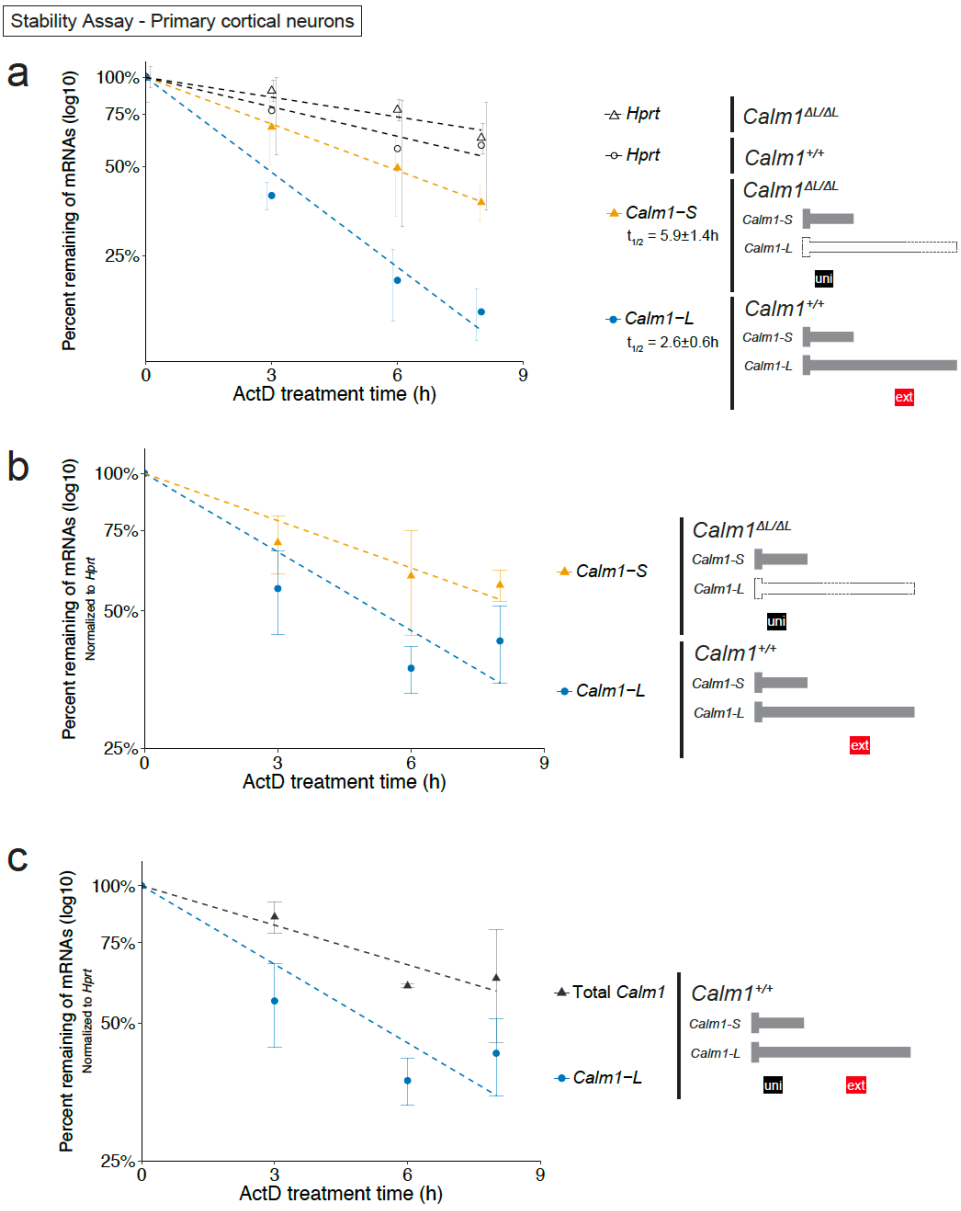
Supplemental Figure 3. CRISPR gRNA off-target analysis.

(a) CRISPR RGEN Tools was used to analyze Cas9 off-target sites using all the 6 gRNAs used in this study as input sequences. Original output file was sorted by mismatch numbers and bulge size. Top 60 positions shown here. Given that most of the top targets of gRNAs belonged to the intended *Calm1* locus (shaded in grey), these were exempted from Sanger sequencing analysis. In order to perform more genome-wide analysis, several genomic locations from the top of the list were selected (Red) (b) PCR amplification of the genomic loci followed by Sanger sequencing. Shaded pink regions are potential off-target sites which showed no alteration in *Calm1*$^{\Delta L/\Delta L}$. Numbers next to the genotypes (either WT or ΔL) indicate mouse identification number. Related to Figure 3.

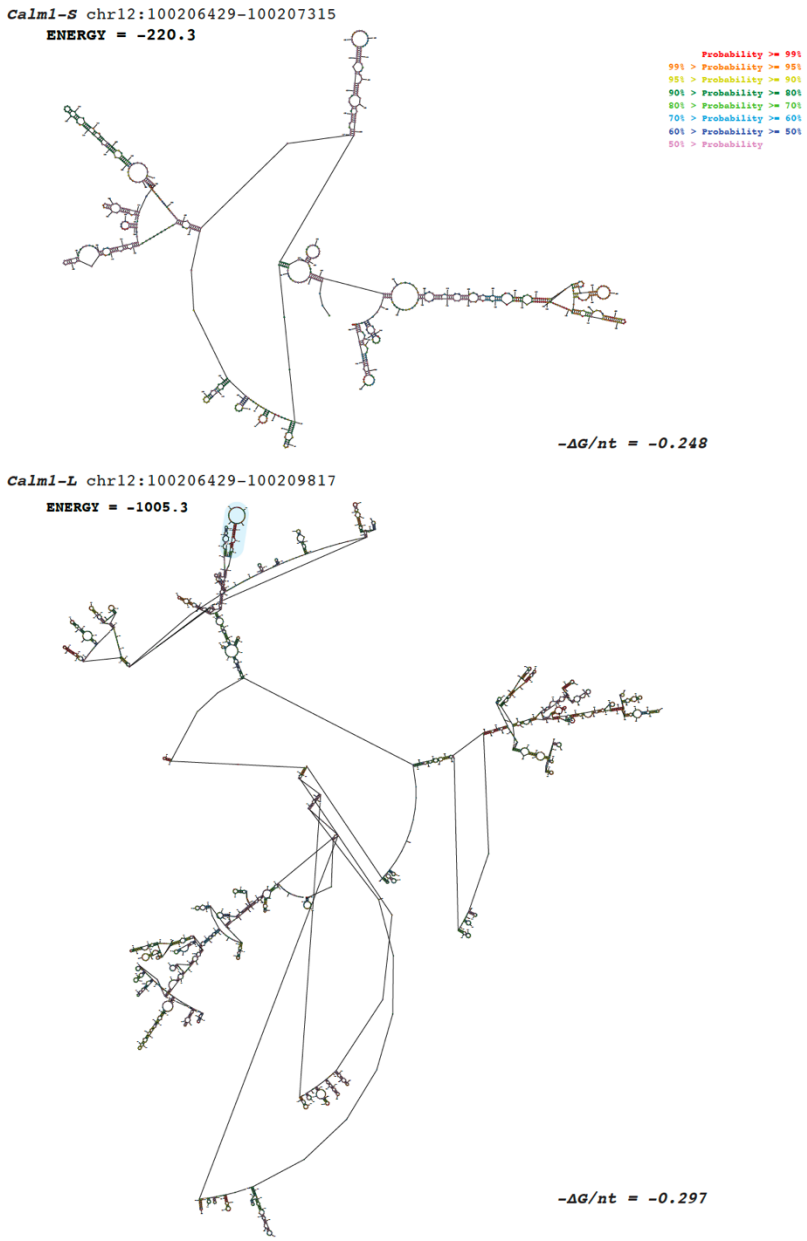


Supplemental Figure 4. *Calm1-L* loss reduces hippocampal IEG expression in response to enriched environment exposure.

(a) Full stitched images showing cFos expression in *Calm1Δ*$^{L/\Delta L}$ mice was reduced compared to *Calm1*$^{+/+}$ controls. Blue – DAPI. Green – cFos. (b) Reduced *Npas4* expression post EE in hippocampus of *Calm1*$^{\Delta L/\Delta L}$ compared to *Calm1*$^{+/+}$. Blue – DAPI. Green – *Npas4*. Right side shows % of *Npas4* positive cells. Significance was determined using t-test, ****: p <= 0.0001. n = 6 mice (2 hemispheres in 2 brain sections for each mouse). (c) Western blot analysis of CaM in hippocampus shows no difference between *Calm1*$^{+/+}$ and *Calm1*$^{\Delta L/\Delta L}$ in home cage or when exposed to enriched environment. Related to Figure 6.



Supplemental Figure 5. *Calm1-L* is less stable than *Calm1-S*.
Extension qRT-PCR primers (ext) were used to detect *Calm1-L* from *Calm1*$^{+/+}$ samples, and universal qRT-PCR primers (uni) were used to uniquely detect *Calm1-S* from *Calm1*$^{\Delta L/\Delta L}$ samples. (a) Without normalization, *Calm1-S* was found to have a longer half-life (t1/2= 5.9 ± 1.4 h) compared to *Calm1-L* (t1/2 = 2.6 ± 0.6 h) (b) When *Calm1* transcript levels were normalized to the stable *Hprt* transcripts, *Calm1-S* was consistently found to be more stable than *Calm1-L*. (c) total *Calm1* levels were measured using uni primers in the *Calm1*$^{+/+}$ samples to ensure that the shorter half-life of *Calm1-L* is not attributed to genetic background. This showed that total *Calm1* had higher stability compared to the *Calm1-L*. n = 3 biological replicates.



Supplemental Figure 6. Secondary structure analysis of *Calm1-S* and *Calm1-L*. Prediction of secondary structure using the RNAstructure software (Bellaousov et al. 2013) estimated the free energy of folding for *Calm1-S* and *Calm1-L*, -220.3 and -1005.3 kcal/mol, respectively. The length-normalized minimum thermodynamic free energy (-ΔG/nt) of predicted structures estimates the structural complexity of the 3′ UTRs in *Calm1-S* and *Calm1-L*. The value for *Calm1-L* 3′ UTR (-ΔG/nt = - 0.297) compared to the *Calm1-S* 3′ UTR (-ΔG/nt = - 0.248) suggested *Calm1-L* is more structured. Color reflects base-pairing probability. Blue shaded region is predicted as an AU-rich element uniquely found in the *Calm1-L*.

Supplemental Table 1: Probes, primers, gRNA sequences, and deletions

Northern blot/in situ probes					
Transcript Detected	**5' primer**	**3' primer**	**probe coordinates**		
Calm1 short and long 3' UTR mRNAs	N.529_Calm1_uni;CCTGCCCCTAAAGTCAAGC	N.530_Calm1_uni; TTGATGGTGTGCTCAAGTCC	chr12:101,444,821-101,445,428		
Calm1 long 3' UTR	N.559_calm1_ext;TGTACTGTTGCCAGGGAAGG	N.560_calm1_ext; TCAAACAGCCACAGTGAAGG	chr12:101,447,277-101,447,864		
Psmd4	ML_mg_9_Psmd4mRNA;CAGCCGATTCTTGGCTAAAC	ML_mg_XX_Psmd4mRNA;CCACACTGTCCAACCCAAAG	chr3:94,837,448-94,839,789		
Calm2	BB_NB_200_MmCalm2_F1;TCGCCATGTGATGACAAACC	BB_NB_201_MmCalm2_R1;CCACGCAGAGTTACAGCTCC	chr17:87,834,519-87,832,996		
Calm3	BB_NB_202_MmCalm3_F1;GGGACTGCATAGAGTGGACG	BB_NB_203_MmCalm3_R1; GGCTAGAGTCAGAAGCAAGGG	chr7:17,501,649-17,501,102		
RT-qPCR primers					
5' primer	**3' primer**	**5' primer chromosome location**	**3' primer chromosome location**		
N.529_calm1_uni;CCTGCCCCTAAAGTCAAGC	mSy.26_calm1_prox_R;GCAAATGGAACTTTAAGGAGTCC	chr12:101,444,821-101,444,840	chr12:101,444,897-101,444,919		
N.559_calm1_ext;TGTACTGTTGCCAGGGAAGG	mSy.28_calm1_ext_R;TGTTTAAGTTGCTCAAAGACACC	chr12:101,447,277-101,447,296	chr12:101,447,337-101,447,359		
ML_mg_9_Psmd4mRNA;CAGCCGATTCTTGGCTAAAC	ML_mg_10_Psmd4mRNA;TCCAGACCTAAGCAGCATGA	chr3:94,837,448-94,837,467	chr3:94,837,606-94,837,625		
BB_qP_MmHprt_F1;CTTCCTCCTCAGACCGCTTT	BB_qP_MmHprt_R1;TCATCGCTAATCACGACGCT	chrX:50,341,356-50,341,375	chrX:50,341,419-50,352,278		
Guide RNAs					
Name	**Sequence**	**Chromosome**	**Start**	**End**	**Strand**
Guide 1	ACGCAACATCCAGTGCAAGC	chr12	101445610	101445629	+
Guide 2	GCTTGCACTGGATGTTGCGT	chr12	101445610	101445629	-
Guide 3	CTCTGACGGGGGACCAACTC	chr12	101447929	101447948	+
Guide 4	GAGTTGGTCCCCCGTCAGAG	chr12	101447929	101447948	-
Guide 5	TTTGCGGTATACGAGACAAA	chr12	101448072	101448091	+
Guide 6	TTTGTCTCGTATACCGCAAA	chr12	101448072	101448091	-
Resulting Deletions					
Deletion name	**Deletion coordinates**	**Combination of Guides**	**Sanger Sequenced?**		
Large	chr12:101,445,626-101,448,077	Guide 1, Guide 2, Guide 5, Guide 6	Y		
Medium	chr12:101,445,627-101,447,931	Guide 1, Guide 2, Guide 3, Guide 4	Y		
Small	chr12:101,447,929-101,448,092	Guide 3, Guide 4, Guide 5, Guide 6	Y		
Genotyping primers					
Deletion Detected	**5' primer**	**3' primer**	**5' primer chromosome location**	**3' primer chromosome location**	
Large/Medium	m.g.2.Calm1 proxpA Fwd;CGTCCATTCCAAGTTGTAAATGC	m.g.1.Calm1 dispA Rev;CTATCTCCAGCTCCACACTCTCA	chr12:101,445,456-101,445,478	chr12:101,448,139-101,448,161	
Small	hj_m.g.7_distalPAS:TCAGCATCCTGGGAACTTCT	hj_m.g.8_distalPAS:GCGTCATCACATCCAGCTTA	chr12:101,447,500-101,447,519	chr12:101,448,491-101,448,510	

Supplemental Table 2: Stellaris smFISH probe sequences.

Stellaris probes						
Probe #	**Isoform detected**	**Sequence**	**Chromosome**	**Start**	**End**	**Dye**
1	Calm1 Long 3' UTR	caggagttgtggctagagaa	chr12	101445646	101445666	Quasar570
2	Calm1 Long 3' UTR	aagaactccgagtctgtctg	chr12	101445670	101445690	Quasar570
3	Calm1 Long 3' UTR	acttacgtggtcatagacgg	chr12	101445692	101445712	Quasar570
4	Calm1 Long 3' UTR	ctgaacatcacctgactgtg	chr12	101445748	101445768	Quasar570
5	Calm1 Long 3' UTR	ctaactgtacgctggagacg	chr12	101445791	101445811	Quasar570
6	Calm1 Long 3' UTR	agatggcctcatgtagttac	chr12	101445842	101445862	Quasar570
7	Calm1 Long 3' UTR	cttttgccatacacagtgag	chr12	101445900	101445920	Quasar570
8	Calm1 Long 3' UTR	aaggtagtgtctttgtctgg	chr12	101445924	101445944	Quasar570
9	Calm1 Long 3' UTR	agctacagtgttatgcacta	chr12	101446039	101446059	Quasar570
10	Calm1 Long 3' UTR	attgtttcgttttttgctc	chr12	101446062	101446082	Quasar570
11	Calm1 Long 3' UTR	tctacgttttggttacaggt	chr12	101446113	101446133	Quasar570
12	Calm1 Long 3' UTR	tttttaagtagaccagtgcc	chr12	101446157	101446177	Quasar570
13	Calm1 Long 3' UTR	tgcaacacacttctaagggg	chr12	101446180	101446200	Quasar570
14	Calm1 Long 3' UTR	atacaaccagttaacaccct	chr12	101446230	101446250	Quasar570
15	Calm1 Long 3' UTR	cgaacttggaagccaaagct	chr12	101446263	101446283	Quasar570
16	Calm1 Long 3' UTR	tctcattgccttgctagaag	chr12	101446302	101446322	Quasar570
17	Calm1 Long 3' UTR	aaccagggatctagaggtta	chr12	101446355	101446375	Quasar570
18	Calm1 Long 3' UTR	caatgtcttcctctctactg	chr12	101446463	101446483	Quasar570
19	Calm1 Long 3' UTR	cagtccaaggatgagcagag	chr12	101446492	101446512	Quasar570
20	Calm1 Long 3' UTR	aagtcaaagtcatcagcctg	chr12	101446523	101446543	Quasar570
21	Calm1 Long 3' UTR	accaggagggaatacctaac	chr12	101446556	101446576	Quasar570
22	Calm1 Long 3' UTR	tacaagacatatccactgcc	chr12	101446621	101446641	Quasar570
23	Calm1 Long 3' UTR	ggtcatcaactgggtgattt	chr12	101446649	101446669	Quasar570
24	Calm1 Long 3' UTR	aggttacagacaagcctgag	chr12	101446673	101446693	Quasar570
25	Calm1 Long 3' UTR	aaaccccttttgaaatcacca	chr12	101446695	101446715	Quasar570
26	Calm1 Long 3' UTR	cagcttctacagacttggta	chr12	101446750	101446770	Quasar570
27	Calm1 Long 3' UTR	tgggctgttaaagcatcatc	chr12	101446774	101446794	Quasar570
28	Calm1 Long 3' UTR	ggaggtaagcttttaggaca	chr12	101446799	101446819	Quasar570
29	Calm1 Long 3' UTR	cctagttacacaagaatcct	chr12	101446824	101446844	Quasar570
30	Calm1 Long 3' UTR	tagggtttctggttctcatg	chr12	101446851	101446871	Quasar570
31	Calm1 Long 3' UTR	tctcacgaactcaaactgcc	chr12	101446985	101447005	Quasar570
32	Calm1 Long 3' UTR	agtgagggacagtttacagg	chr12	101447055	101447075	Quasar570
33	Calm1 Long 3' UTR	ctaaagcaggatgccagtta	chr12	101447137	101447157	Quasar570
34	Calm1 Long 3' UTR	ctatgtagaggtgtgccaag	chr12	101447188	101447208	Quasar570
35	Calm1 Long 3' UTR	cccccaacagacagaaaaca	chr12	101447210	101447230	Quasar570
36	Calm1 Long 3' UTR	tcaagatgtttgccttgttc	chr12	101447238	101447258	Quasar570
37	Calm1 Long 3' UTR	cagtacattaaagcccatct	chr12	101447264	101447284	Quasar570
38	Calm1 Long 3' UTR	agctcatcttcaatgtgttt	chr12	101447355	101447375	Quasar570
39	Calm1 Long 3' UTR	gccattccttacaatatagt	chr12	101447423	101447443	Quasar570
40	Calm1 Long 3' UTR	gtgacataacactaggccta	chr12	101447450	101447470	Quasar570
41	Calm1 Long 3' UTR	tgatagatgctattcctctt	chr12	101447472	101447492	Quasar570
42	Calm1 Long 3' UTR	ggtcgctgccaaagagaaat	chr12	101447608	101447628	Quasar570
43	Calm1 Long 3' UTR	agttgggacacatcagtgta	chr12	101447633	101447653	Quasar570
44	Calm1 Long 3' UTR	tgacttctaccggacaactt	chr12	101447702	101447722	Quasar570
45	Calm1 Long 3' UTR	atccctgatttcttactcaa	chr12	101447743	101447763	Quasar570
46	Calm1 Long 3' UTR	acatcgtcgatataggggag	chr12	101447771	101447791	Quasar570
47	Calm1 Long 3' UTR	acgtccaagtacaactcttt	chr12	101447866	101447886	Quasar570
48	Calm1 Long 3' UTR	gagaggtggttaaatctcca	chr12	101447912	101447932	Quasar570

6. References

1. Alwis DS, Rajan R. 2014. Environmental enrichment and the sensory brain: the role of enrichment in remediating brain injury. *Front Syst Neurosci* **8**: 156.
2. An JJ, Gharami K, Liao G-Y, Woo NH, Lau AG, Vanevski F, Torre ER, Jones KR, Feng Y, Lu B, et al. 2008. Distinct Role of Long 3′ UTR BDNF mRNA in Spine Morphology and Synaptic Plasticity in Hippocampal Neurons. *Cell* **134**: 175–187.
3. Andreassi C, Zimmermann C, Mitter R, Fusco S, Devita S, Saiardi A, Riccio A. 2010. An NGF-responsive element targets myo-inositol monophosphatase-1 mRNA to sympathetic neuron axons. *Nat Neurosci* **13**: 291–301.
4. Bellaousov S, Reuter JS, Seetin MG, Mathews DH. 2013. RNAstructure: web servers for RNA secondary structure prediction and analysis. *Nucleic Acids Res* **41**: W471–W474.
5. Blair JD, Hockemeyer D, Doudna JA, Bateup HS, Floor SN. 2017. Widespread translational remodeling during human neuronal differentiation.
6. Brigidi GS, Hayes MGB, Delos Santos NP, Hartzell AL, Texari L, Lin PA, Bartlett A, Ecker JR, Benner C, Heinz S, et al. 2019. Genomic Decoding of Neuronal Depolarization by Stimulus-Specific NPAS4 Heterodimers. *Cell* **179**: 373-391.e27.
7. Cajigas IJ, Tushev G, Will TJ, Tom Dieck S, Fuerst N, Schuman EM. 2012. The Local Transcriptome in the Synaptic Neuropil Revealed by Deep Sequencing and High-Resolution Imaging. *Neuron* **74**: 453–466.
8. Campbell DS, Holt CE. 2001. Chemotropic Responses of Retinal Growth Cones Mediated by Rapid Local Protein Synthesis and Degradation. *Neuron* **32**: 1013–1026.
9. Chung L. 2015. A Brief Introduction to the Transduction of Neural Activity into Fos Signal. *Dev Reprod* **19**: 61–7.

10. Coutellier L, Beraki S, Ardestani PM, Saw NL, Shamloo M. 2012. Npas4: A Neuronal Transcription Factor with a Key Role in Social and Cognitive Functions Relevant to Developmental Disorders ed. K. Hashimoto. *PLoS One* **7**: e46604.
11. Crerar H, Scott-solomon E, Bodkin-clarke C, Gaspari M, Kuruvilla R, Crerar H, Scott-solomon E, Bodkin-clarke C, Andreassi C, Hazbon M, et al. 2019. Regulation of NGF Signaling by an Axonal Untranslated mRNA Report Regulation of NGF Signaling by an Axonal Untranslated mRNA. 1–11.
12. Drew LJ, Stark KL, Fénelon K, Karayiorgou M, Macdermott AB, Gogos JA. 2011. Evidence for altered hippocampal function in a mouse model of the human 22q11.2 microdeletion. *Mol Cell Neurosci* **47**: 293–305.
13. Fanarraga ML, Charite J, Hage WJ, De Graaff W, Deschamps J. 1997. Hoxb-8 gain-of-function transgenic mice exhibit alterations in the peripheral nervous system. *J Neurosci Methods* **71**: 11–18.
14. Fischer JW, Busa VF, Shao Y, Leung AKL. 2020. Structure-Mediated RNA Decay by UPF1 and G3BP1. *Mol Cell* **78**: 70-84.e6.
15. Fischer R, Koller M, Flura M, Mathews S, Strehler-Page MA, Krebs J, Penniston JT, Carafoli E, Strehler EE. 1988. Multiple divergent mRNAs code for a single human calmodulin. *J Biol Chem* **263**: 17055–17062.
16. Floor SN, Doudna JA. 2016. Tunable protein synthesis by transcript isoforms in human cells. *Elife* **5**: 1–25.
17. Fritz JL, VanBerkum MF. 2000. Calmodulin and son of sevenless dependent signaling pathways regulate midline crossing of axons in the Drosophila CNS. *Development* **127**: 1991–2000.

18. Greer PL, Greenberg ME. 2008. From Synapse to Nucleus: Calcium-Dependent Gene Transcription in the Control of Synapse Development and Function. *Neuron* **59**: 846–860.
19. Gumy LF, Yeo GS, Tung YC, Zivraj KH, Willis D, Coppola G, Lam BY, Twiss JL, Holt CE, Fawcett JW. 2011. Transcriptome analysis of embryonic and adult sensory axons reveals changes in mRNA repertoire localization. *RNA* **17**: 85–98.
20. Guvenek A, Tian B. 2018. Analysis of alternative cleavage and polyadenylation in mature and differentiating neurons using RNA-seq data. *Quant Biol* **6**: 253–266.
21. Ha KCH, Blencowe BJ, Morris Q. 2018. QAPA: a new method for the systematic analysis of alternative polyadenylation from RNA-seq data. *Genome Biol* **19**: 45.
22. Hagenston AM, Bading H. 2011. Calcium signaling in synapse-to-nucleus communication. *Cold Spring Harb Perspect Biol* **3**: a004564.
23. Halim D, Wilson MP, Oliver D, Brosens E, Verheij JB, Han Y, Nanda V, Lyu Q, Doukas M, Stoop H, et al. 2017. Loss of LMOD1 impairs smooth muscle cytocontractility and causes megacystis microcolon intestinal hypoperistalsis syndrome in humans and mice. *Proc Natl Acad Sci U S A* **114**: E2739–E2747.
24. Han Y, Slivano OJ, Christie CK, Cheng AW, Miano JM. 2015. CRISPR-Cas9 genome editing of a single regulatory element nearly abolishes target gene expression in mice--brief report. *Arter Thromb Vasc Biol* **35**: 312–315.
25. Holt CE, Schuman EM. 2013. The central dogma decentralized: new perspectives on RNA function and local translation in neurons. *Neuron* **80**: 648–657.
26. Hoque M, Ji Z, Zheng D, Luo W, Li W, You B, Park JY, Yehia G, Tian B. 2013. Analysis of alternative cleavage and polyadenylation by 3′ region extraction and deep sequencing. *Nat Methods* **10**: 133–139.

27. Ikeshima H, Yuasa S, Matsuo K, Kawamura K, Hata J, Takano T. 1993. Expression of three nonallelic genes coding calmodulin exhibits similar localization on the central nervous system of adult rats. *J Neurosci Res* **36**: 111–119.

28. Ito HT, Schuman EM. 2012. Functional division of hippocampal area CA1 via modulatory gating of entorhinal cortical inputs. *Hippocampus* **22**: 372–87.

29. Jones KJ, Templet S, Zemoura K, Kuzniewska B, Pena FX, Hwang H, Lei DJ, Haensgen H, Nguyen S, Saenz C, et al. 2018. Rapid, experience-dependent translation of neurogranin enables memory encoding. *Proc Natl Acad Sci* **115**: E5805–E5814.

30. Kakiuchi S, Yasuda S, Yamazaki R, Teshima Y, Kanda K, Kakiuchi R, Sobue K. 1982. Quantitative determinations of calmodulin in the supernatant and particulate fractions of mammalian tissues. *J Biochem* **92**: 1041–1048.

31. Kim YS, Furman S, Sink H, VanBerkum MF. 2001. Calmodulin and profilin coregulate axon outgrowth in Drosophila. *J Neurobiol* **47**: 26–38.

32. Kobayashi H, Saragai S, Naito A, Ichio K, Kawauchi D, Murakami F. 2015. Calm1 signaling pathway is essential for the migration of mouse precerebellar neurons. *Development* **142**: 375–384.

33. Kocabas A, Duarte T, Kumar S, Hynes MA. 2015. Widespread Differential Expression of Coding Region and 3′ UTR Sequences in Neurons and Other Tissues. *Neuron* **88**: 1149–1156.

34. La Fata G, Gärtner A, Domínguez-Iturza N, Dresselaers T, Dawitz J, Poorthuis RB, Averna M, Himmelreich U, Meredith RM, Achsel T, et al. 2014. FMRP regulates multipolar to bipolar transition affecting neuronal migration and cortical circuitry. *Nat Neurosci* **17**: 1693–1700.

35. Le Douarin NM, Smith J. 1988. Development of the peripheral nervous system from the neural crest. *Annu Rev Cell Biol* **4**: 375–404.

36. Lee SH, Mayr C. 2019. Gain of Additional BIRC3 Protein Functions through 3'-UTR-Mediated Protein Complex Formation. *Mol Cell* **74**: 701-712.e9.

37. Lianoglou S, Garg V, Yang JL, Leslie CS, Mayr C. 2013. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev* **27**: 2380–2396.

38. Liao GY, An JJ, Gharami K, Waterhouse EG, Vanevski F, Jones KR, Xu B. 2012. Dendritically targeted Bdnf mRNA is essential for energy balance and response to leptin. *Nat Med* **18**: 564–571.

39. Lin AC, Holt CE. 2007. Local translation and directional steering in axons. *EMBO J* **26**: 3729–3736.

40. Malenka RC, Kauer JA, Perkel DJ, Mauk MD, Kelly PT, Nicoll RA, Waxham MN. 1989. An essential role for postsynaptic calmodulin and protein kinase activity in long-term potentiation. *Nature* **340**: 554–557.

41. Malka Y, Steiman-Shimony A, Rosenthal E, Argaman L, Cohen-Daniel L, Arbib E, Margalit H, Kaplan T, Berger M. 2017. Post-transcriptional 3´-UTR cleavage of mRNA transcripts generates thousands of stable uncapped autonomous RNA fragments. *Nat Commun* **8**: 2029.

42. Marmigere F, Ernfors P. 2007. Specification and connectivity of neuronal subtypes in the sensory lineage. *Nat Rev Neurosci* **8**: 114–127.

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

44. Means AR, Dedman JR. 1980. Calmodulin--an intracellular calcium receptor. *Nature* **285**: 73–77.

45. Melli G, Höke A. 2009. Dorsal root ganglia sensory neuronal cultures: A tool for drug discovery for peripheral neuropathies. *Expert Opin Drug Discov* **4**: 1035–1045.

46. Miller S, Yasuda M, Coats JK, Jones Y, Martone ME, Mayford M. 2002. Disruption of Dendritic Translation of CaMKIIα Impairs Stabilization of Synaptic Plasticity and Memory Consolidation. *Neuron* **36**: 507–519.

47. Minis A, Dahary D, Manor O, Leshkowitz D, Pilpel Y, Yaron A. 2014. Subcellular transcriptomics-Dissection of the mRNA composition in the axonal compartment of sensory neurons. *Dev Neurobiol* **74**: 365–381.

48. Miura P, Sanfilippo P, Shenker S, Lai EC. 2014. Alternative polyadenylation in the nervous system: To what lengths will 3′ UTR extensions take us? *BioEssays* **36**: 766–777.

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

50. Mueller F, Senecal A, Tantale K, Marie-Nelly H, Ly N, Collin O, Basyuk E, Bertrand E, Darzacq X, Zimmer C. 2013. FISH-quant: automatic counting of transcripts in 3D FISH images. *Nat Methods* **10**: 277–278.

51. Ni B, Rush S, Gurd JW, Brown IR. 1992. Molecular cloning of calmodulin mRNA species which are preferentially expressed in neurons in the rat brain. *Brain Res Mol Brain Res* **13**: 7–17.

52. Nojima H. 1989. Structural organization of multiple rat calmodulin genes. *J Mol Biol* **208**: 269–282.

53. Oliver D, Yuan S, McSwiggin H, Yan W. 2015. Pervasive Genotypic Mosaicism in Founder Mice Derived from Genome Editing through Pronuclear Injection. *PLoS One* **10**: e0129457.

54. Pang ZP, Cao P, Xu W, Südhof TC. 2010. Calmodulin controls synaptic strength via presynaptic activation of calmodulin kinase II. *J Neurosci* **30**: 4132–42.

55. Perry RBT, Doron-Mandel E, Iavnilovitch E, Rishal I, Dagan SY, Tsoory M, Coppola G, McDonald MK, Gomes C, Geschwind DH, et al. 2012. Subcellular Knockout of Importin β1 Perturbs Axonal Retrograde Signaling. *Neuron* **75**: 294–305.

56. Preitner N, Quan J, Li X, Nielsen FC, Flanagan JG. 2016. IMP2 axonal localization, RNA interactome, and function in the development of axon trajectories. *Development* **143**: 2753–2759.

57. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB. 2008. Proliferating cells express mRNAs with shortened 3′ untranslated regions and fewer microRNA target sites. *Science (80-)* **320**: 1643–1647.

58. SenGupta B, Friedberg F, Detera-Wadleigh SD. 1987. Molecular analysis of human and rat calmodulin complementary DNA clones. Evidence for additional active genes in these species. *J Biol Chem* **262**: 16663–16670.

59. Sethna F, Zhang M, Kaphzan H, Klann E, Autio D, Cox CL, Wang H. 2016. Calmodulin Activity Regulates Group I Metabotropic Glutamate Receptor- Mediated Signal Transduction and Synaptic Depression. *J Neurosci Res* **94**: 401–408.

60. Shigeoka T, Lu B, Holt CE. 2013. RNA-based mechanisms underlying axon guidance. *J Cell Biol* **202**: 991–999.

61. Smibert P, Miura P, Westholm JO, Shenker S, May G, Duff MO, Zhang D, Eads BD, Carlson J, Brown JB, et al. 2012. Global Patterns of Tissue-Specific Alternative Polyadenylation in Drosophila. *Cell Rep* **1**: 277–289.

62. Sorensen AB, Sondergaard MT, Overgaard MT. 2013. Calmodulin in a heartbeat. *Febs J* **280**: 5511–5532.

63. Sudmant PH, Lee H, Dominguez D, Heiman M, Burge CB. 2018. Widespread Accumulation of Ribosome-Associated Isolated 3′ UTRs in Neuronal Cell Populations of the Aging Brain. *Cell Rep* **25**: 2447-2456.e4.

64. Sun L, Fazal FM, Li P, Broughton JP, Lee B, Tang L, Huang W, Kool ET, Chang HY, Zhang QC. 2019. RNA structure maps across mammalian cellular compartments. *Nat Struct Mol Biol* **26**: 322–330.

65. Sun X, Lin Y. 2016. Npas4: Linking Neuronal Activity to Memory. *Trends Neurosci* **39**: 264–275.

66. Terenzio M, Koley S, Samra N, Rishal I, Zhao Q, Sahoo PK, Urisman A, Marvaldi L, Oses-Prieto JA, Forester C, et al. 2018. Locally translated mTOR controls axonal local translation in nerve injury. *Science (80-)* **359**: 1416–1421.

67. Tessier-Lavigne M, Goodman CS. 1996. The molecular biology of axon guidance. *Science (80-)* **274**: 1123–1133.

68. Tom Dieck S, Kochen L, Hanus C, Heumüller M, Bartnik I, Nassim-Assir B, Merk K, Mosler T, Garg S, Bunse S, et al. 2015. Direct visualization of newly synthesized target proteins in situ. *Nat Methods* **12**: 411–414.

69. Tushev G, Glock C, Heumuller M, Biever A, Jovanovic M, Schuman EM. 2018. Alternative 3′ UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments. *Neuron* **98**: 495–511.

70. Ulitsky I, Shkumatava A, Jan CH, Subtelny AO, Koppstein D, Bell GW, Sive H, Bartel DP. 2012. Extensive alternative polyadenylation during zebrafish development. *Genome Res* **22**: 2054–66.

71. Valluy J, Bicker S, Aksoy-Aksel A, Lackinger M, Sumer S, Fiore R, Wüst T, Seffer D, Metge F, Dieterich C, et al. 2015. A coding-independent function of an alternative Ube3a transcript during neuronal development. *Nat Neurosci* **18**: 666–673.

72. van den Akker E, Reijnen M, Korving J, Brouwer A, Meijlink F, Deschamps J. 1999. Targeted inactivation of Hoxb8 affects survival of a spinal ganglion and causes aberrant limb reflexes. *Mech Dev* **89**: 103–114.

73. Vanberkum MFA, Goodman CS. 1995. Targeted Disruption of Ca2+-Calmodulin Signaling in Drosophila Growth Cones Leads to Stalls in Axon Extension and Errors in Axon Guidance. *Neuron* **14**: 43–56.

74. VanElzakker M, Fevurly RD, Breindel T, Spencer RL. 2008. Environmental novelty is associated with a selective increase in Fos expression in the output elements of the hippocampal formation and the perirhinal cortex. *Learn Mem* **15**: 899–908.

75. Wang B, Pan L, Wei M, Wang Q, Liu W-W, Wang N, Jiang X-Y, Zhang X, Bao L. 2015. FMRP-Mediated Axonal Delivery of miR-181d Regulates Axon Elongation by Locally Targeting Map1b and Calm1. *Cell Rep* **13**: 2794–2807.

76. Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/cas-mediated genome engineering. *Cell* **153**: 910–918.

77. Xia Z, Storm DR. 2005. The role of calmodulin as a signal integrator for synaptic plasticity. *Nat Rev Neurosci* **6**: 267–276.

78. Yamniuk AP, Vogel HJ. 2004. Calmodulin's flexibility allows for promiscuity in its interactions with target proteins and peptides. *Mol Biotechnol* **27**: 33–57.

79. Yap EL, Greenberg ME. 2018. Activity-Regulated Transcription: Bridging the Gap between Neural Activity and Behavior. *Neuron* **100**: 330–348.

80. Zappulo A, van den Bruck D, Ciolli Mattioli C, Franke V, Imami K, McShane E, Moreno-Estelles M, Calviello L, Filipchyk A, Peguero-Sanchez E, et al. 2017. RNA localization is a key determinant of neurite-enriched proteome. *Nat Commun* **8**: 583.

81. Zhang M, Ergin V, Lin L, Stork C, Chen L, Zheng S. 2019a. Axonogenesis Is Coordinated by Neuron-Specific Alternative Splicing Programming and Splicing Regulator PTBP2. *Neuron* **101**: 690-706.e10.

82. Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keeffe S, Phatnani HP, Guarnieri P, Caneda C, Ruderisch N, et al. 2014. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci* **34**: 11929–11947.

83. Zhang Z, So K, Peterson R, Bauer M, Ng H, Zhang Y, Kim JH, Kidd T, Miura P. 2019b. Elav-Mediated Exon Skipping and Alternative Polyadenylation of the Dscam1 Gene Are Required for Axon Outgrowth. *Cell Rep* **27**: 3808-3817.e7.

84. Zivraj KH, Tung YCL, Piper M, Gumy L, Fawcett JW, Yeo GSH, Holt CE. 2010. Subcellular Profiling Reveals Distinct and Developmentally Regulated Repertoire of Growth Cone mRNAs. *J Neurosci* **30**: 15464–15478.

Chapter III: CRISPR-mediated knockout of long 3′ UTR mRNA isoforms in mESC-derived neurons

Alternative cleavage and polyadenylation (APA) is pervasive, occurring for more than 70% of human and mouse genes. Distal poly(A) site selection to generate longer 3′ UTR mRNA isoforms is prevalent in the nervous system, affecting thousands of genes. Here, we establish mouse embryonic stem cell (mESC)-derived neurons as a suitable system to study long 3′ UTR isoforms. RNA-seq analysis revealed that mESCs-derived neurons show widespread 3′ UTR lengthening that closely resembles APA patterns found in mouse cortex. mESCs are highly amenable to genetic manipulation. We present a method to eliminate long 3′ UTR isoform expression using CRISPR/Cas9 editing. This approach can lead to clones with the desired deletion within several weeks. We demonstrate this strategy on the *Mprip* gene as a proof-of-principle. To confirm loss of long 3′ UTR expression and the absence of cryptic poly(A) site usage stemming from the CRISPR deletion, we present a simple and cost-efficient targeted long-read RNA-sequencing strategy using the Oxford Nanopore Technologies platform. Using this method, we confirmed specific loss of the *Mprip* long 3′ UTR isoform. CRISPR gene editing of mESCs thus serves as a highly relevant platform for studying the molecular and cellular functions of long 3′ UTR mRNA isoforms.

1. Introduction

Most mouse and human genes are subject to alternative cleavage and polyadenylation (APA) resulting in the expression of mRNA isoforms with different 3′ ends (Hoque et al. 2012; Lianoglou et al. 2013). Most commonly, APA results in alternative length 3′ untranslated regions (UTRs) but does not change protein-coding sequence (Sandberg et al. 2008; Chen and Shyu 1995). The usage of distal poly(A) sites is dramatically enhanced in the nervous system and results in mRNA isoforms with longer 3′ UTRs. This lengthening of the 3′ UTR in the nervous system has been observed across multiple organisms including *Drosophila*, zebrafish, mice, and humans (Ulitsky et al. 2012; Smibert et al. 2012; Miura et al. 2013).

Alternative long 3′ UTR mRNA isoforms have been found to be important in vivo. For instance, the long 3′ UTR isoform of *Bdnf* is necessary for transcript localization to dendrites, and in turn, for dendritic spine morphology and synaptic plasticity in mice (An et al. 2008). The long 3′ UTR isoform of *Calm1* is required for hippocampal neuron activation and proper C1 dorsal root ganglion (DRG) development in mouse (Bae et al. 2020). In *Drosophila*, loss of the *Dscam1* long 3′ UTR was found to severely impair axon outgrowth (Zhang et al. 2019b). These studies make it clear that the longer APA isoforms of at least some genes are necessary for nervous system development and function. However, it is less clear what molecular and cellular roles can be attributed specifically to longer, neural-enriched 3′ UTR isoforms. Long 3′ UTRs would generally be predicted to harbor regulatory *cis*-elements such as microRNA target sites, leading to regulation of translational control. However, new roles are emerging for 3′ UTRs (Bae and Miura 2020), including regulation of subcellular localization (Tushev et al. 2018), regulation of upstream alternative splicing events on the same gene (Zhang et al. 2019b), and a scaffolding role in assembly of protein complexes as they are translated (Berkovits and Mayr 2015).

CRISPR-mediated deletion of the distal poly(A) site has been established as an effective method to abolish long 3′ UTR isoform expression in both *Drosophila* and mice (Bae et al. 2020; Zhang et al. 2019b). However, to interrogate the roles of long 3′ UTR isoforms in relation to their impact on translational control, alternative splicing, and mRNA localization, a cultured cell system would have many advantages. What has been lacking to date is identification of a neuronal cell culture system that both is amenable to rapid gene-editing and also expresses long 3′ UTR isoforms at the levels observed in brain tissues.

Here, we show that mouse embryonic stem cell (mESC)-derived glutamatergic neurons exhibit robust expression of long 3′ UTR isoforms that closely resemble to the expression pattern in the mouse cortex on a transcriptome-wide level. We present a strategy for achieving CRISPR-mediated deletion of distal poly(A) sites in mESCs and demonstrate its effectiveness for deleting the long 3′ UTR isoform of the *Mprip* gene. CRISPR-mediated genetic manipulation in mESCs is fast and efficient, thus permitting the rapid exploration of the cellular and molecular functions of long 3′ UTR transcript isoforms.

2. Material and methods

Short-read RNA-seq analysis

Public RNA-seq datasets from mESC (SRR645824 SRR645826 SRR645828) and mESC-derived neurons (DIV7 for days in vitro 7; SRR645846 SRR645849 SRR645851) were used for APA analysis. Fastq files were mapped to the mouse genome using STAR 2.7. Mouse GRCm38.102 genome was used as reference. Binary bam files were visualized in IGV 2.4.17. For 3′ UTR lengthening analysis, QAPA 1.3.1 was used (Ha et al. 2018). Pre-built 3′ UTR annotation 1.3.0 was obtained from the QAPA github (github.com/morrislab/ qapa). 3′ UTR fasta sequences were extracted using the

annotation bed file and mouse mm10 genome fasta sequences. 3′ UTR isoform usage was quantified using Salmon 1.4.0 and the Salmon index generated using the 3′ UTR fasta file. PAU, which stands for Poly(A) site Usage, values were calculated from the Salmon quantification. Subsequent data processing was performed in R 3.6.0. From the QAPA output file, we selected the most distal Poly(A) site transcript by filtering for the rows with the maximum length of 3′ UTR for each of the genes to obtain the absolute distal PAU (dPAU) values. Non-APA genes with a single poly(A) site were filtered out. An expression threshold of 1 TPM across all the samples was used. T-test was performed, and FDR adjusted P-values were obtained. The dPAU fold-change between mESCs and mESC-derived neurons was calculated for each gene and a volcano plot was generated using ggplot geom_point.

For correlation analysis, first, absolute dPAU values were obtained using QAPA. For this purpose, QAPA analysis was performed independently using mESC vs mESC-derived neurons dataset or embryonic day 14 (E14) vs postnatal day 30 (P30) cortex dataset (SRR1805814 SRR1805815 SRR1805824 SRR1805825). Output files from each QAPA analysis contained dPAU values for mESC-derived neurons or P30 cortex. Then, the dPAU values were arranged in a way that the mean dPAU for mESC-derived neurons and the mean dPAU for cortex were merged into the same dataframe by gene name. Then a scatter plot was generated using ggplot geom_point, and correlation efficiency was calculated using stat_cor in R.

Neuro2A cell culture

Neuro2A (N2A) cells were obtained from ATCC (ATCC CCL-131). Cells were maintained in DMEM (Thermo Scientific 11965092) supplemented with 10% FBS (R&D systems S11150). For differentiation, 5x105 cells were plated onto a 100 mm dish with

DMEM supplemented with 2% FBS and 20 μM retinoic acid (Sigma Aldrich R2625). Media was changed daily until day 7 when neuronal-like morphology was confirmed by light microscopy and RNA was collected.

Feeder-free mESC culture

Immediately after thawing, mESCs (ES-E14TG2a; ATCC CRL-1821) were kept on MEF feeder cells for 2 passages. Cells were maintained on 0.1% gelatin (Sigma EmbryoMax ES006B) coated cell culture dishes in mESC medium (high glucose DMEM with Glutamax (Thermo Scientific 10566016) supplemented with 15% FBS (Sigma F2442; tested for mESC culture), 55 μM b-mercaptoethanol (Thermo Scientific 21985023), 1X MEM non-essential amino acids (Thermo Scientific 11140050), 1 mM sodium pyruvate (Thermo Scientific 11360070), and 1000 U/mL LIF (Sigma ESG1107)) at 37°C and 5% CO2.

Differentiation to glutamatergic neurons

mESCs were differentiated to neural progenitor cells (NPC) and then to glutamatergic neurons (Hubbard et al. 2013; Bibel et al. 2007). Briefly, $3.5x10^6$ cells were plated onto 90 mm bacteriological dishes in 15 mL NPC medium (DMEM with L-glutamine (Thermo Scientific 11965092) supplemented with 10% FBS, 1X non-essential amino acids, and 550 μM b-mercaptoethanol). On day 2, cell aggregates were collected and transferred into a 50 mL conical tube and left at room temperature for 3-5 min until cell aggregates settled down. Media was carefully discarded and replaced with new NPC media. The cell aggregates were gently mixed by pipetting up and down and returned to the bacteriological dishes. On day 4, media change was performed as describes, and 5 μM of retinoic acid was added. Media was changed again on day 6 and day 8 and retinoic acid was

supplemented. On day 10, NPC aggregates were collected and dissociated with 1 mL of TrypLE (Thermo Scientific 12604013) at 37°C for 5-7 min. To halt the reaction, 8 mL of Trypsin inhibitor (Thermo Scientific R007100) was added. The NPC aggregates were gently dissociated by pipetting up and down and filtered through 40 µm cell strainer. Cell suspension was diluted in N2 media (Neurobasal (Thermo Scientific 21103049) supplemented with 1X N2 (Thermo Scientific 17502048) and 2 mM glutamine (Thermo Scientific 25030081)) at $3x10^5$ cells/mL. 10 mL of cells were plated onto PDL (Sigma P7280) coated 100 mm cell culture dishes. Complete media change was performed at 4h (day 10) and 24h (day 11) with N2 media. At this point, short neurite extensions were visible. On day 12 (equivalent of neurons DIV 2) and 14 (DIV 4), media was replaced with B27 media (Neurobasal supplemented with 1X B27 (Thermo Scientific 17504044) and 2 mM glutamine). Cells were maintained until day 17 (DIV 7).

RNA extraction and RT-qPCR

RNA was extracted from cultured cells using the Trizol (Thermo Scientific 15596018) method. RNA was quantified using Nanodrop spectrometer. Five µg of total RNA was treated with Turbo DNase (Thermo Scientific AM1907). Then, 1 µg of DNase-treated total RNA was reverse transcribed using Maxima reverse transcriptase (Thermo Scientific EP0742). The first strand cDNA reaction was diluted 1:5 before performing RT-qPCR. RT-qPCR was performed using SYBR Select Master Mix for CFX (Thermo Scientific 4472954). BioRad CFX96 real time PCR machine was used and results were analyzed using the $\Delta\Delta$Ct method. Primers used for RT-qPCR are found in **Supplementary Table 1**.

sgRNA design

Guide RNAs (sgRNAs) were designed using CRISPick GPP sgRNA designer from the Broad Institute (https://portals.broadinstitute.org/gppx/crispick/public) by providing DNA sequence for the region spanning the distal poly(A) site. In the case of *Mprip*, each output sgRNA was inspected assuring that one sgRNA targets upstream of the dPAS and one downstream, giving preference to higher “pick order” and “combined ranked sequences”. sgRNA oligos used for *Mprip* long 3′ UTR isoform (*Mprip-L*) knockout are found in **Supplementary Table 1**.

For bulk sgRNA analysis, the same CRISPick GPP sgRNA designer tool was used in bulk mode. Bed file for the 150 bp upstream of the most distal 3′ UTR end was generated using the QAPA dPAU file as a reference. Bed file for the 150 bp downstream region was generated in a similar way for each of the genes. Then DNA sequences were extracted in fasta format using bedtools getfasta and GRCm38 mouse genome. Upstream and downstream fasta sequences were separately inputted for sgRNA design 500 sequences at a time. Only the top rank sgRNA was used for further analysis. Upstream and downstream sgRNA cut position information was used to estimate the expected deletion size for each gene. Data processing and plotting was performed in R.

Construction of plasmids

sgRNAs oligonucleotides were synthesized with CACCG at the 5′ end followed by the sgRNA sequence. The complementary oligonucleotide contained 5′ AAAC and 3′ C sequence at each end. Oligonucleotides were phosphorylated (10 μM oligonucleotides 1, 10 μM oligonucleotides 2, 1X T4 ligation buffer, and 0.5 μL T4 PNK in 10 μL reaction) at 37°C for 30 min. Then the oligos were annealed by heating at 95°C for 15 min and gradually cooling down to room temperature in a thermocycler. The pX333 plasmid

(Addgene #64073), which contains duplex U6 promoter and sgRNA scaffolds, was used for dual sgRNA cloning. The first sgRNA (for instance, the sgRNA targeting the upstream of the dPAS) was cloned into BbsI digested pX333 by standard ligation (50 ng digested pX333, 1μL of 1:250 diluted oligo duplex, 1X Quick ligation buffer, and 1μL Quick ligase (NEB) in 10 μL reaction). Ligation reaction was transformed in DH5α competent cells. Successful cloning was confirmed by Sanger sequencing. Subsequently, the second sgRNA was cloned into BsaI site following the same steps. Final sgRNA plasmids were confirmed by Sanger sequencing.

Donor plasmid harboring the antibiotic resistant cassette was constructed using pUC19 plasmid backbone. The entire backbone was PCR amplified using KOD Xtreme Hot Start DNA polymerase (Millipore 71975). 750 bp upstream to the left Cas9 cut site and 750 bp downstream to the right Cas9 cut site were PCR amplified from mESC genomic DNA and used as the homology arms. The upstream/forward primer for the left homology arm and the downstream/reverse primer for the right homology arm contained one of the sgRNAs and its PAM sequence overhang. The 20 – 25 nt overlap sequences necessary for HIFI assembly were added by a second round PCR using purified first round PCR product and primers with overlapping sequence overhangs. The neomycin resistance cassette was PCR amplified from pEN759 (kindly provided by Dr. Benoit Bruneau) using primers with overlap sequence overhangs. Each fragment was assembled into the pUC19 backbone using HIFI (NEB). Primer sequences are found in **Supplementary Table 1**.

mESC transfection and clonal analysis

Four hours prior to transfection, $2x10^5$ cells were plated onto a single well (35 mm well) of gelatin-coated 6 well plate in mESC media. After 4 hours or once the cells are

attached to the plate, 2.5 µg of donor plasmid and 0.5 µg of sgRNA plasmid was mixed with 9 µL of PEI (Polysciences 23966) in up to 200 µL of Opti-MEM (Thermo Scientific 31985062) followed by 10 min incubation at room temperature. Transfection solution was dispersed drop by drop onto the cells. The plate was returned to incubator until the next day. On day 1, full media change was performed. On day 2, appropriate antibiotics were supplemented with fresh media. On day 3, survived cells were lifted with 1 ml TrypLE (Thermo Scientific 12604013), diluted 1:100 – 1:200 and plated onto gelatin-coated 100 mm cell culture dishes. In other words, 100 µL is taken from the 1 mL cell suspension then plated into 10 mL mESC media to make a 1:100 dilution dish. 50 µL out of the 1 mL cell suspension was plated into a 100 mm dish to make a 1:200 dilution plate. The remaining cell suspension can be diluted into 10 mL mESC media and plated as a backup. On day 4, isolated cells should be observed under microscope. Media with antibiotics were replaced every day. Dishes were kept for additional 5 – 6 days until defined mESC colonies were observed. mESC colonies were manually picked using micropipettes under microscope and transferred into each wells of a gelatin coated 48 well plate. Each clone was maintained for 4 – 5 days for cell expansion. When cells were > 60% confluent, cells were dissociated with TrypLE by incubating at 37°C for 5 min. Half of the cells were used for genotyping and the other half were transferred into each well of gelatin coated 12 well plate for further cell expansion. Clones with expected genotype were gradually expanded to 6 well plates and then to 100 mm dishes.

Genomic DNA was extracted by incubating the cells in 300 µL lysis buffer with proteinase K at 65°C for 4 h - overnight. Genomic DNA was ethanol precipitated and resuspended in 200 µL of molecular grade water. Genotyping was performed by PCR. Once confirmed by PCR, a large fragment spanning the edit region was PCR amplified and Sanger sequenced. Genotyping primers are found in **Supplementary Table 1**.

Nanopore sequencing

For full-length cDNA synthesis, SMARTer PCR cDNA synthesis kit (Clontech) was used according to user manual. Total RNA was DNAse treated either on-column using PureLink DNase (Thermo Scientific 12185010) and PureLink RNA Mini kit (Thermo Scientific 12183020) or with TURBO DNase (Thermo Scientific AM1907). First strand cDNA was synthesized using 700 ng of DNase treated RNA, 3′ SMART CDS Primer II A and SMARTer II A TSO. cDNA was diluted 1:5 in water. cDNA was amplified by optimal (18 – 21) cycles of PCR using Advantage 2 PCR kit (Clontech). For enrichment of cDNA for genes of interest, in this case *Mprip*, cDNA was enriched by capturing with biotinylated xGen Lockdown probes (IDT; probe sequences found in **Supplementary Table 1**) and xGen hybridization and wash kit (IDT). Target capture probe design tool was used for the design of 120 bp probes. Probes were selected based on the location (toward the 3′ end) and GC contents (preferentially between 40 – 65°C). Captured cDNA was purified using AMPure XL beads (Beckman Coulter) then amplified using Takara LA Taq DNA polymerase Hot-Start version (Clontech).

For nanopore library preparation, captured and amplified cDNA was end-prepped (NEBNext FFPE DNA Repair Mix and NEBNext Ultra™ II End Repair Kit) and ligated (T4 DNA Ligase – NEB) with sequencing adapters (Nanopore SQK-LSK110). MinION Mk1B device and FLO-FLG001 flow cells were used for sequencing of the libraries. Minimap2 was used to map the reads by supplementing GRCm38.102 transcriptome bed file. Aligned bam files were visualized in IGV.

3. Results

mESC-derived neurons express neural long 3′ UTR mRNA isoforms

We first sought out a murine cell culture system that could recapitulate the long 3′ UTR isoforms that have been identified in mouse brain tissues (Miura et al. 2013). Mining public RNA-seq data we noticed that switching from short 3′ UTR usage to long 3′ UTR usage was evident when mESCs were differentiated into glutamatergic neurons for a handful of examples, including *Agap1* (Arf1 GTPase activating protein) and *Map4* (Microtubule-associated protein 4; **Figure 11A**). To assess whether this lengthening of 3′ UTRs during mESC neuronal differentiation occurred transcriptome-wide, we employed QAPA, a tool used to analyze APA. QAPA estimates the usage of different poly(A) sites for each gene from RNA-Seq data and known 3′ end annotations (Ha et al. 2018). We were particularly interested in usage of the most distal poly(A) site (dPAU for distal Poly(A) site Usage), since long 3′ UTRs are more abundant in neurons. QAPA analysis showed a significant increase of dPAU in mESC derived neurons compared to mESCs for 1200 genes, whereas only 45 genes showed the opposite trend (**Figure 11B**). We next wanted to determine whether these APA events were correlated with changes that occur during mouse cortical development. We performed QAPA analysis on P30 cortex RNA-Seq data to obtain absolute dPAU values. Comparison of dPAU values of all the APA genes between P30 cortex and mESC-derived neurons showed a high correlation ($R=0.84$, $p<2.2\times10^{-16}$). Thus, mESC-derived neurons closely resemble the distal poly(A) site expression patterns found *in vivo* (**Figure 11C**).

To confirm this lengthening trend with an independent experimental method, we performed mESC differentiation into glutamatergic neurons using established differentiation protocols with some adaptations (See Materials and Methods) (Hubbard et

al. 2013; Bibel et al. 2007). Feeder-free mESCs were cultured in suspension for formation of cell aggregates, during which retinoic acid was supplemented to induce formation of

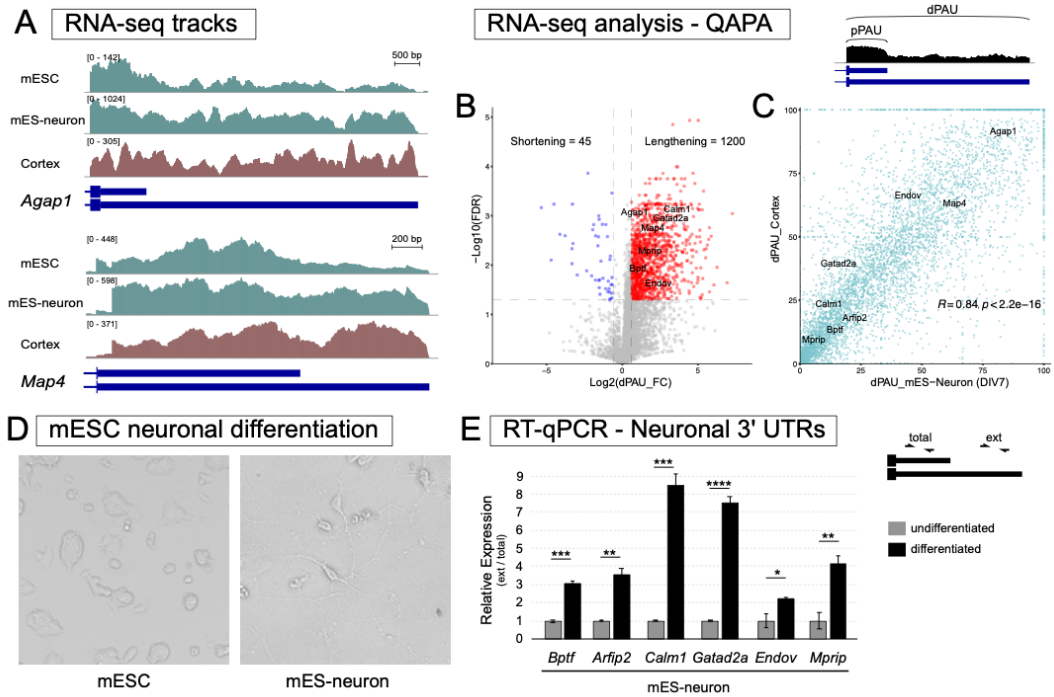


Figure 11. mESC-derived neurons express neural long 3′ UTR mRNA isoforms. (A) Visualization of RNA-seq tracks for APA genes suggested long 3′ UTR isoforms are upregulated in mESC-derived neurons compared to mESC (green tracks), and it closely resembles mouse cortical transcript profile (brown track). (B) Assessment of distal Poly(A) site usage (dPAU) values from QAPA analysis showed robust lengthening of 3′ UTRs during mESC to glutamatergic neuron differentiation. 1200 genes showed significant fold-change increase of dPAU from mESC to mESC-derived neuron stage (Red). (C) Absolute dPAU values from postnatal day 30 cortex and mESC-derived neurons were compared to assess their correlation. Comparison of dPAU values of all the APA genes between P30 cortex and mESC-derived neurons showed a high correlation (R=0.84, p<2.2x10-16) suggesting that the long 3′ UTR isoforms expression resembles the expression pattern found in vivo. (D) Picture of mESCs and differentiated glutamatergic neurons days in vitro 7 (DIV7; mESC taken at 10X objective and DIV7 neurons at 20X). Well-defines polarized morphology was observed. (E) RT-qPCR was performed for 6 genes to confirm the upregulation of long 3′ UTR isoforms in mESC-derived neurons. Long 3′ UTR isoforms expression was normalized to the total gene expression. Two-tail t-test were performed between the undifferentiated and differentiated cells; n=3, * < 0.05, ** < 0.01, *** < 0.001, **** < 0.0001.

NPCs. Retinoic acid is used to halt proliferation-associated signaling pathways and to switch from proliferation to neural differentiation stage (Janesick et al. 2015). After 10 days of NPC induction, the NPC aggregates were gently dissociated and grown for 7 additional days. Neuron differentiation media contained factors such as transferrin and insulin, which commit the cells for their final differentiation fate and assist with neuronal survival (Bottenstein and Sato 1985). After 7 days of neuronal differentiation, well-defined neuron-like polarized morphology was observed (**Figure 11D**). RT-qPCR was used to determine APA changes for 6 target genes. We confirmed a significant increase in the expression of long 3′ UTR isoforms in the mESC-derived neurons compared to mESCs (**Figure 11E**). We also investigated whether differentiation to a neuronal-like fate for Neuro2a (N2a) cells also led to 3′ UTR lengthening. However, we did not observe significant changes in APA by RT-qPCR for the same genes (**Supplementary Figure 7**). Thus, mESC-derived neurons appear to be a good system for studying the regulation and function of neuron-enriched long 3′ UTR isoforms.

CRISPR-Cas9 strategy for long 3′ UTR isoform-specific deletion

Previously, we established that CRISPR-mediated deletion of the distal poly(A) site is an effective way to abolish long 3′ UTR isoform expression in mice (Bae et al. 2020). We thus reasoned that a similar strategy could be applied in mESCs. sgRNAs are chosen that flank the distal poly(A) site to ensure that both the poly(A) signal and downstream U-rich elements (DUE) are deleted (**Figure 12A**). In most cases, a 100 – 200 bp deletion encompassing the distal poly(A) site can be designed. It is possible to delete larger regions of the long 3′ UTR sequence; however, distal poly(A) site deletion has been found to be sufficient in our experience and reduces the potential for disrupting potential DNA

enhancers or silencers in the locus. In order to allow for efficient clonal selection, homology directed repair is used to insert an antibiotic resistance cassette at the deletion site.

Two plasmids are used for generating the deletion clones, a Cas9-sgRNA construct and a donor construct. The donor construct uses pUC19 as a backbone and includes: 1) sgRNA1 target site and PAM (protospacer adjacent motif) sequence; 2) 750 bp left homology arm; 3) antibiotic resistance cassette (neomycin), flanked by FRT flippase recombination sequences; 4) 750 bp right homology arm; and 5) sgRNA2 target site and PAM sequence (**Figure 12A**). Donor constructs are flanked by sgRNA target sites and PAM sequences since the linearization of donor plasmid by sgRNA incorporation has been shown to improve HDR in 293T and iPSCs (Zhang et al. 2017). The FRT sequences allow for future removal of the antibiotics resistant cassette which can be achieved by transiently expressing the flippase enzyme in CRISPR-edited cells (Gronostajski and Sadowski 1985).

The Cas9-sgRNA plasmid and donor plasmid are easily transfected into mESCs using conventional transfection reagents such as polyethylenimine (PEI) which is a cost-effective cationic polymer that assists DNA transfection (Longo et al. 2013). Transfection of a single 35 mm well is sufficient to obtain multiple knockout clones. Antibiotic selection is initiated at day 2 post-transfection with a complete media change. On day 3, dilution of the cells is performed at a ratio of 1:100 or 1:200 into 100 mm dishes. Cells are maintained with constant antibiotic push. Roughly 8-9 days post-transfection, large mESC clusters are formed. The clusters are manually transferred into each well of a 48-well plate using a 200 µL micropipette. Within 6-7 days, sufficient cell growth is expected. Cells are gently dissociated, and half of the cells are transferred into each well of a new 12-well plate for expansion. The remaining half is used for PCR genotyping. Genotyping of 8-12 clones is

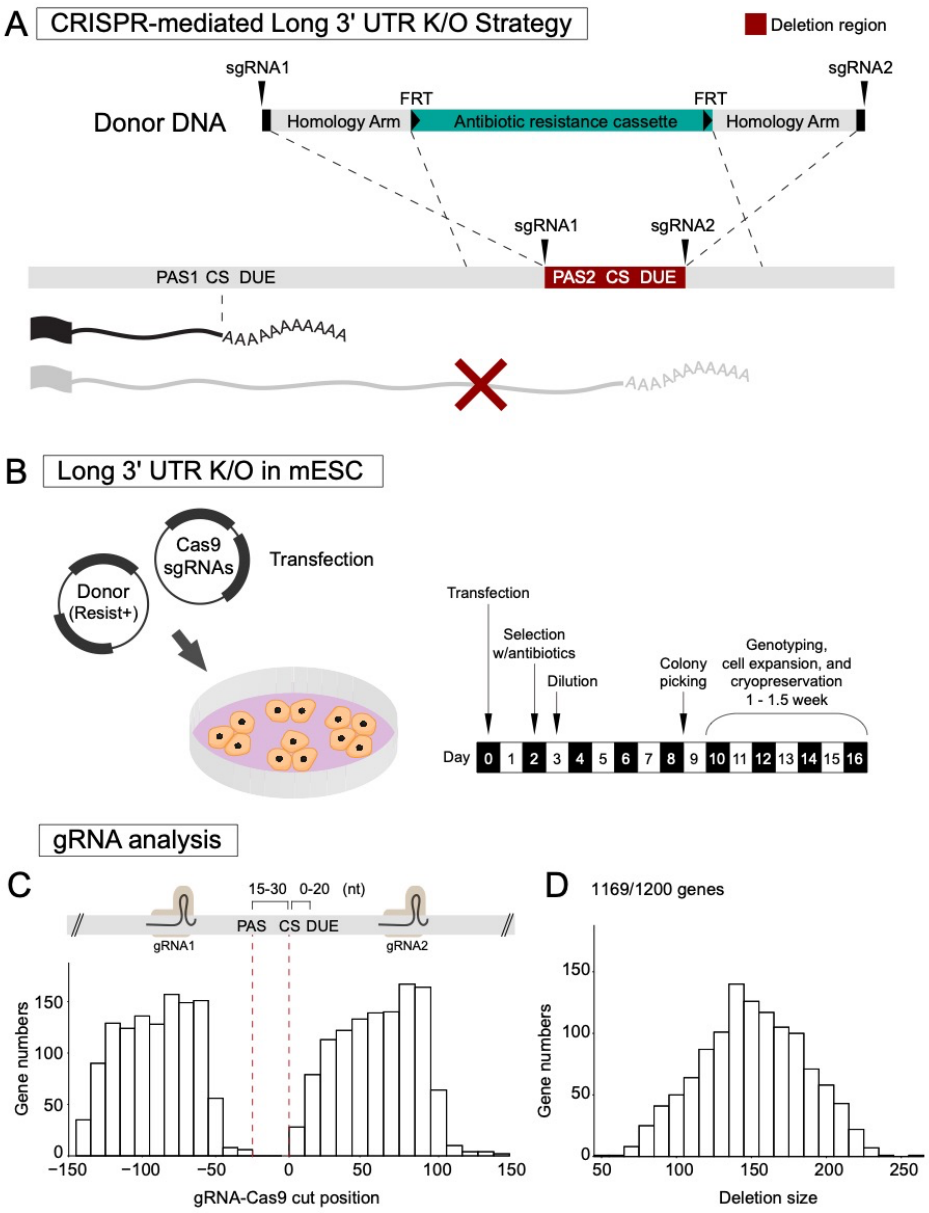


Figure 12. CRISPR-Cas9 strategy for long 3′ UTR isoform-specific knockout.
(A) CRISPR-Cas9 strategy for long 3′ UTR isoforms involves usage of a pair of sgRNAs to direct Cas9 and homology-directed repair (HDR) using donor plasmid harboring an antibiotics resistance cassette. The sgRNA pair targets regions flanking the distal poly(A) site. The donor construct for HDR includes sgRNA1 target site, protospacer adjacent motif (PAM) sequence, left homology arm, antibiotic resistance cassette, right homology arm, sgRNA2 target site, and PAM sequence. (B) The Cas9-sgRNAs plasmid and donor plasmid are transfected into mESCs using conventional transfection methods. Antibiotic selection is initiated at day 2 post-transfection. On the day 3, dilution of the cells is performed. Roughly 8-9 days post-transfection, mESC clusters are

manually selected for further genotyping and cell expansion. (C - D) Analysis for designing dual sgRNAs at 300 nt genomic sequences flanking the annotated distal 3′ end of all the 1200 genes (PAS = poly(A) signal, CS = cleavage site, DUE = downstream U-rich element). Our analysis showed it is feasible to design dual sgRNAs for 1169 genes (97.4%). On average, sgRNA pairs were located 90 ± 25 nt upstream and 62 ± 25 nt downstream of the 3′ end, introducing a deletion size of 151 ± 35 nt spanning the distal poly(A) site.

usually enough to obtain homozygous knock-in lines (**Figure 12B**). Sanger sequencing is employed to confirm the precise deletion.

To determine how many of the 1200 3′ UTR lengthening genes would be amenable to this distal poly(A) site deletion strategy, we extracted 150 bp DNA sequences upstream and downstream of each long 3′ UTR end and ran the sgRNA design tool to find the top rank sgRNAs pair for each gene. Our analysis showed it is feasible to design dual sgRNAs for 1169 out of the 1200 genes (97.4%). By plotting the sgRNA target sites relative to the distal 3′ end location, we found that that sgRNA pairs were located 90 ± 25 nt upstream and 62 ± 25 nt downstream of the 3′ end (**Figure 12C**). The expected deletion region size was 151 ± 35 nt spanning the distal poly(A) site (**Figure 12D**). This result suggests the robust applicability of the CRISPR strategy to generate long 3′ UTR deletions for nearly all genes of interest.

Generation of long 3′ UTR isoform deletion lines for the *Mprip* gene.

We applied our strategy to one of the 3′ UTR lengthening genes as a proof of principle. *Mprip* (Myosin Phosphatase Rho-Interacting Protein), encodes for a F-actin binding protein (Mulder et al. 2004). *Mprip* generates a 650 bp short 3′ UTR and a 5.2 kb long 3′ UTR (**Figure 13A**). QAPA analysis revealed that *Mprip* long 3′ UTR isoform (*Mprip*-

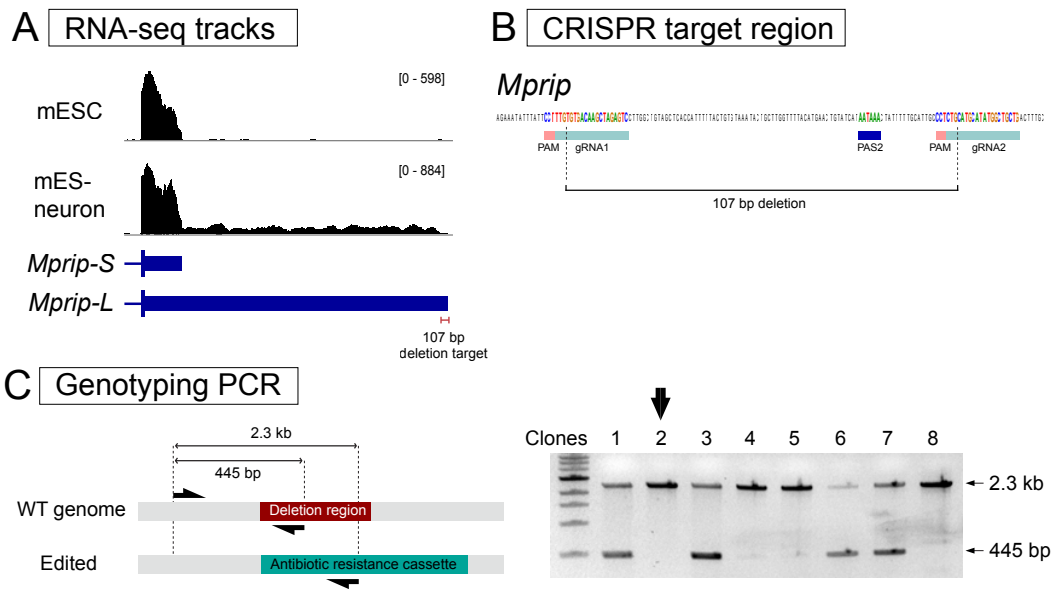


Figure 13. CRISPR-mediated generation of *Mprip* long 3′ UTR isoform knockout cells.
(A) *Mprip* generates a 650 bp short 3′ UTR isoform and a 5.2 kb long 3′ UTR isoform. Long 3′ UTR isoforms expression is increased during neuronal differentiation. (B) A pair of sgRNAs were chosen that flanked the distal poly(A) site. The strategy was designed to generate a precise 107 bp deletion. (C) PCR genotyping was performed using three primers set – a common forward primer, a reverse primer specific for the target deletion region, and a reverse primer specific for the antibiotics cassette. Homozygous deletion clones were selected for further characterization (black arrow).

L) expression is increased during neuronal differentiation by 5-fold (**Figure 13A**). RT-qPCR analysis confirmed this expression trend (**Figure 11E**). A pair of sgRNAs were chosen that flanked the distal poly(A) site. The 750 bp Homology arm regions were defined in accordance with the Cas9 cut site, which is 3 nt upstream of the PAM sequence. The strategy was designed to generate a precise 107 bp deletion (**Figure 13B**). mESCs were transfected and stable cell lines were generated as described above. PCR genotyping was used to select the clonal cells carrying the neomycin cassette (**Figure 13C**).

Homozygous deletion clones were selected (**Figure 13C**; shown with black arrow) and propagated. The 107 bp deletion was confirmed by Sanger sequencing.

Confirmation of long 3′ UTR mRNA isoform loss using Long-read sequencing

To confirm that the genomic deletion abolishes long 3′ UTR expression, differentiation into neurons is required. The preferred tool to confirm loss of a long 3′ UTR isoform is northern blot because both the expression level and length of the isoforms is revealed (An et al. 2008; Bae et al. 2020). In addition, northern analysis could also detect if spurious transcripts using cryptic poly(A) sites are activated by the genomic deletion. However, northern requires a large amount of starting material and thus is not convenient to apply on low expressed genes in mESC-derived neurons. Thus, we present targeted long-read cDNA sequencing as an alternative approach.

The targeted long-read cDNA sequencing approach relies on synthesis of template switching cDNAs and probe-based enrichment of the target prior to library preparation (**Figure 14A**). The template switching cDNA approach improves the synthesis of full-length cDNAs (Zhu et al. 2018). This cDNA synthesis relies on an oligo(dT)-based primer for detecting polyadenylated RNAs, thus allowing the recognition of 3′ ends. One of the limitations of conventional transcriptome-wide long-read cDNA sequencing is the difficulty in obtaining full length reads for low abundance genes (Byrne et al. 2017). Probe capture-based enrichment results in a library enriched for the gene of interest and is sequenced using the affordable Nanopore MinION platform.

To confirm loss of *Mprip-L* expression, WT control and CRISPR deletion cells were differentiated in glutamatergic neurons. Total RNA was extracted and full-length cDNA was synthesized. cDNA for *Mprip* gene was enriched using xGen Lockdown probe pulldown. The library was prepared using Nanopore ligation sequencing kit and

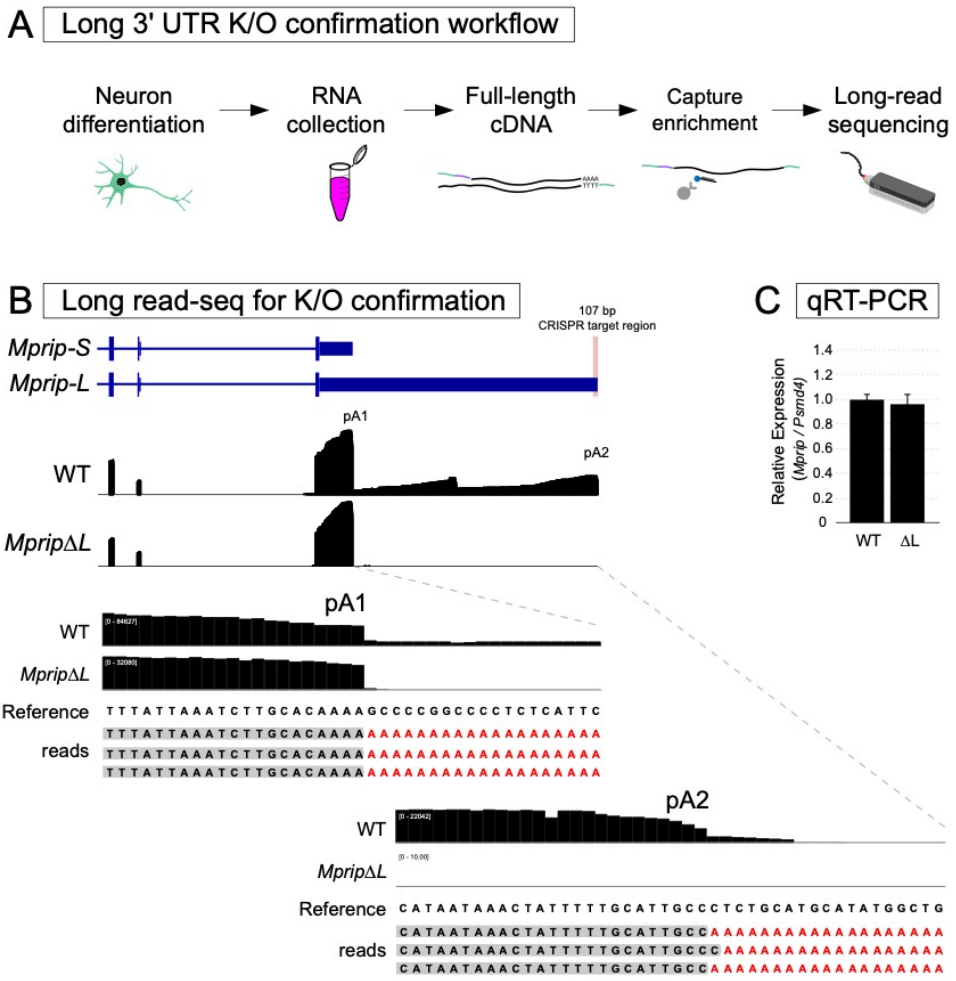


Figure 14. Confirmation of neuronal long 3′ UTR variant knockout using Long-read sequencing.

(A) Workflow for assessing lack of the long 3′ UTR isoform expression. (B) Long-read sequencing data visualized in IGV. In the WT track, the two expected 3′ ends of *Mprip* and an internal mis-priming artifact were observed. In contrast to WT neurons, the *Mprip-L* deletion neurons (*Mprip*ΔL) lacked long 3′ UTR transcripts. No evidence of cryptic poly(A) site usage was found as evidenced by lack of long reads aligning downstream of the deleted region. (C) RT-qPCR we performed in order to obtain an accurate quantification of the total levels of *Mprip* in WT vs ΔL. The deletion did not alter the overall *Mprip mRNA* levels; n=3.

sequencing was performed using flongle flow cells on the MinION. Reads were mapped to the genome using Minimap2 (Li 2018).

In the WT neuron long read sequencing libraries, the two expected 3′ ends of *Mprip* were clearly visible. A drop off of read coverage was also evident at a genomically encoded A stretch in the middle of the long 3′ UTR. This is a typical artifact of sequencing methods that rely on oligo(dT) for reverse transcription (**Figure 14B**; WT track). We confirmed that the reads did not contain untemplated poly(A) sequences at this internally primed false 3′ end. In contrast to WT neurons, the long 3′ UTR deletion neurons (*MpripΔL*) lacked long 3′ UTR transcripts (**Figure 14B**; *MpripΔL* track). There were no long reads aligning downstream of the deleted region which shows that cryptic poly(A) site usage was not activated by the distal poly(A) site deletion. This demonstrates that the deletion was effective in preventing *Mprip-L* mRNA expression. In order to obtain an accurate quantification of the total levels of *Mprip* in WT vs ΔL, RT-qPCR we performed. This showed that the deletion did not alter the overall *Mprip mRNA* levels (**Figure 14C**). Targeted Nanopore long-read sequencing thus represents a rapid and accessible method for confirming long 3′ UTR mRNA isoform loss.

4. Discussion

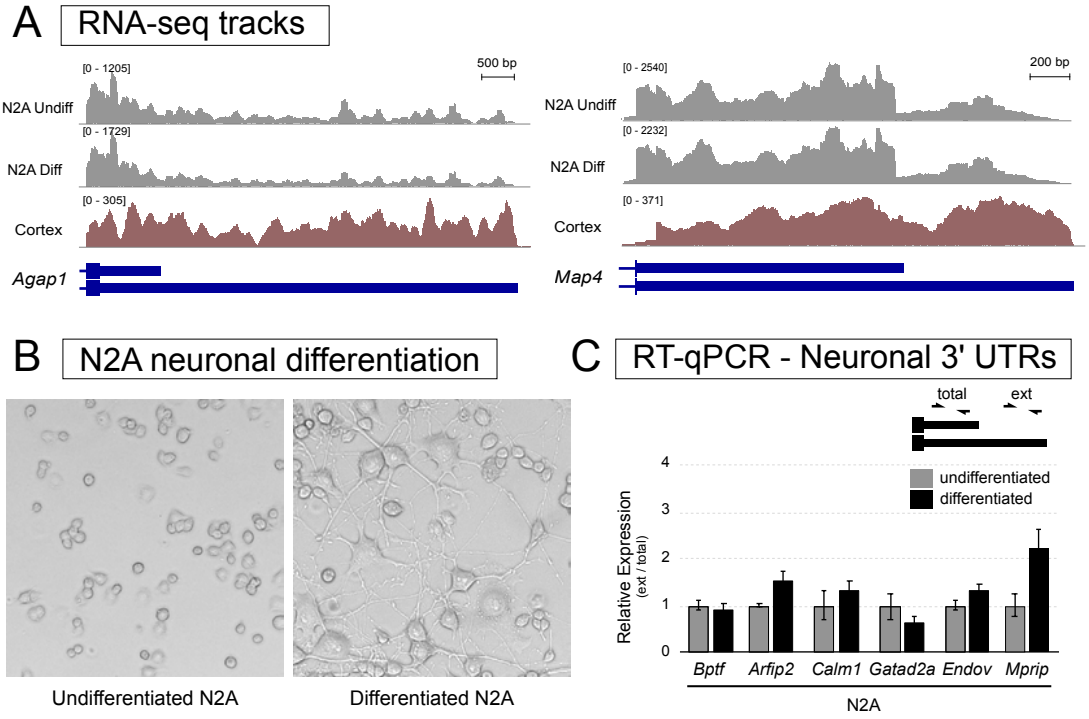
Despite the robust expression of long 3′ UTR isoforms in the nervous system, their functional relevance is underexplored. Although we have previously generated a long 3′ UTR isoform-specific knockout mouse for the *Calm1* gene (Bae et al. 2020), this approach is time consuming and expensive. An in vitro system that recapitulates in vivo long 3′ UTR isoform expression might speed up the screening of long 3′ UTR isoforms for molecular and cellular functions. Here, we establish mESC-derived glutamatergic neurons as a suitable in vitro model for studying long 3′ UTRs. In comparison to the generation of a

mouse line, abolishment of the long 3′ UTR isoform expression in mESC-derived neurons was incredibly fast, thus, allowing generation of extensive CRISPR-edited cell lines in a short period of time. mESCs can be differentiated using alternative methods such as neurogenin expression (Reyes et al. 2008) and also can be differentiated into other neuron types, including motor neurons (Wichterle and Peljto 2008) and dopaminergic neurons (Friling et al. 2009). Thus, this system can be expanded to a cell-type specific analysis of long 3′ UTR isoform function.

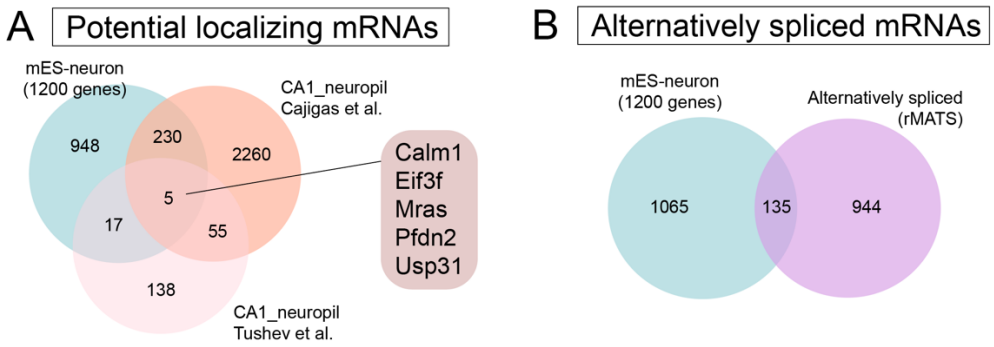
mRNA localization to subcellular regions of neurons can be influenced by 3′ UTR sequence content (Garner et al. 1988; Bassell et al. 1998; Moretti et al. 2015). Our understanding of 3′ UTR sequences that direct subcellular localization comes mostly from reporter based-experiments and genome-wide expression trends (Bauer et al. 2019; Kislauskis et al. 1994; Tushev et al. 2018). mESC-derived neurons exhibit defined neuronal processes (**Figure 11D**), thus it should be readily amenable to transcript isoform detection and quantification using fluorescence in situ hybridization. Out of the 1200 genes that showed robust lengthening of 3′ UTR in mES-neurons, 252 genes are found in previously reported hippocampal neuropil localized mRNA lists (Cajigas et al. 2012; Tushev et al. 2018) (**Supplementary Figure 8A**). Whether 3′ UTR sequences play a role in localizing these transcripts is intriguing. The roles of 3′ UTR sequences in subcellular localization for single poly(A) site genes has also been studied using CRISPR deletion approaches in mice (Terenzio et al. 2018). Our method is amenable to doing the same in mESCs. This can be achieved by targeting sgRNA1 downstream of the stop codon and sgRNA2 downstream of the poly(A) site. A donor DNA cassette can be used to re-introduce the poly(A) site along with an antibiotic resistance cassette. Our approach permits a broad investigation into the importance of 3′ UTRs, in their endogenous genomic context, and in controlling mRNA subcellular localization.

Other known 3′ UTR functions that have been shown for only a few genes, such as the RNA scaffolding role (Lee and Mayr 2019) and the role in regulating alternative splicing (Zhang et al. 2019b), can be expanded to many genes using our technique. Alternative splicing analysis using rMATS (Shen et al. 2014) suggested that at least 135 genes that present lengthening of 3′ UTR during neuronal differentiation are also subject to alternative splicing (**Supplementary Figure 8B**). Thus, providing a short list of genes that can be screened for coordinated regulation of alternative splicing and alternative polyadenylation. Culturing of cells could easily be upscaled to obtain sufficient material necessary for molecular techniques such as ribosome profiling for studying translational regulation. Although mESC-derived neurons present limitations for modeling actual neurons in animals, they can be used to identify regulators of neuronal differentiation (Linares et al. 2015; Zhang et al. 2019a), and to measure synaptic transmission properties (Whipple et al. 2020). In summary, our CRISPR-Cas9 strategy for generating long 3′ UTR isoform-specific knockouts provides a new avenue for the rapid exploration of long 3′ UTR functions at molecular and cellular levels.

5. Supplementary data



Supplementary Figure 7. Differentiation of Neuro2a cells. (A) Visualization of RNA-seq tracks for APA genes suggested long 3′ UTR isoforms are not strongly upregulated in differentiated N2a cells (grey tracks), and it shows splicing and 3′ expression discrepancy when compared to mouse cortical transcript profile (brown track). (B) Pictures of undifferentiated and differentiated N2a cells for 7 days. Morphological changes suggested successful differentiation protocol. (C) RT-qPCR performed for the same 6 genes that showed strong upregulation of long 3′ UTR isoforms in mESC-derived neurons. Long 3′ UTR isoforms expression was not significantly induced in differentiated N2a cells. Two-tail t-test were performed between the undifferentiated and differentiated cells.



Supplementary Figure 8. Potential roles of neuronal long 3′ UTR mRNA isoforms. (A) The list of 1200 genes that presented lengthening of 3′ UTR during the neuronal differentiation (Figure 1B) was compared with previously reported lists of genes localized to hippocampal neuropil. 252 genes showed overlap with at least one list and 5 were found in both localizing lists. (B) Alterative splicing analysis was performed using rMATS using the same RNA-seq dataset (splicing change ΔPSI > 0.2; FDR < 0.05). A list of genes that presented both alternative polyadenylation and alternative splicing was obtained by overlapping the rMATS list and the 1200 lengthening genes (Figure 11B). This analysis suggested at least 135 genes are subject to both alternative splicing and alternative polyadenylation.

Supplementary Table

Supplementary Table 3. List of synthetic oligonucleotides

RT-qPCR primers	qP3201_Bptf_uni_F;GCTGCCCTCATTTGTCAGCTA qP3202_Bptf_uni_R;GGTCGCCTTGTACTTCACCTA qP3203_Bptf_ext_F;CCATGTTAGCCCTGTCCTGC qP3204_Bptf_ext_R;ATAGCAGGGCGACACCAAAT
	qP3197_Arfip2_uni_F;AGTTCAACATCAAGCTGCGG qP3198_Arfip2_uni_R;CGCCCAGTGTCCTTCTTGAT qP3199_Arfip2_ext_F;ACCTGCACTGCTCTTGTTGA qP3200_Arfip2_ext_R;TTGGACCTGTAGTTGACCGC
	N.529_calm1_uni;CCTGCCCCTAAAGTCAAGC mSy.26_calm1_prox_R;GCAAATGGAACTTTAAGGAGTCC N.559_calm1_ext;TGTACTGTTGCCAGGGAAGG mSy.28_calm1_ext_R;TGTTTAAGTTGCTCAAAGACACC
	qP3179_Gatad2a_uni_F;CAGAGCTGGTTGCTGGGTTA qP3180_Gatad2a_uni_R;TTGGCAGGCATGGGAATCAT qP3181_Gatad2a_ext_F;GGTCTGACCAAGAGGGTTCG qP3182_Gatad2a_ext_R;AACTGGGCCTACAAGGGTTG
	qP3104_Endov_uni_F;GGCAACAGTGCATGGAACTC qP3105_Endov_uni_R;GCATGGTGACCACACTGACTA qP3106_Endov_ext_F;CAGGGGGAGGCAACCTAAAC qP3107_Endov_ext_R;CCTGGCAGTCCACCATTCTT
	BBqP335_MpripL_F1;GCCTTCCCACTCATGCCATA BBqP336_MpripL_R1;GAGTCTGGGGTTTGGGTGTT BBqP337_MpripU_F1;AGGCACAGCAGCCTGTTAAT BBqP338_MpripU_R1;TCTGACCATCAAGGGTGGTG
	ML_mg_9_Psmd4;CAGCCGATTCTTGGCTAAAC ML_mg_10_Psmd4;TCCAGACCTAAGCAGCATGA
gRNA oligos	BB811_Mprip;caccgGACTCTAGCTTGTCACACAA BB812_Mprip;aaacTTGTGTGACAAGCTAGAGTCc BB813_Mprip;caccgCAGCAGCCATATGCATGCAG BB814_Mprip;aaacCTGCATGCATATGGCTGCTGc
PCR primers	BB4018_pUC19_F;GAATTCACTGGCCGTCGTTTTAC BB4019_pUC19_R;AAGCTTGGCGTAATCATGG BB4020_leftHA_1st_F;CCTTTGTGTGACAAGCTAGAGTCAAGAGATACCACTGTTAAGAATC BB4021_leftHA_2nd_F;accatgattacgccaagcttCCTTTGTGTGACAAGCTAGAGT BB4022_leftHA_R;AACCGGAAATCTCCACAAAGGAATAAATATTTCTACAAATTAACA BB4023_Neo_F;ATATTTATTCCTTTGTGGAGATTTCCGGTTAAGAATTC BB4024_Neo_R;GCAGCCATATGCATGAGTGGATCCATTATGTACCTG BB4025_rightHA_F;CATAATGGATCCACTCATGCATATGGCTGCTGACTTT BB4026_rightHA_1st_R;CAGCAGCCATATGCATGCAGAGGAGACAGGATTTTGCTGTGTAAC BB4027_rightHA_2nd_R;AAACGACGGCCAGTGAATTCCAGCAGCCATATGCATGCAG
genotyping primers	BB4035_geno_F;GGCAGGGCCTCTTATATCTCG BB4036_geno_R;TGTAGAGAGAGCAGCAGGTTC BB4111_geno_R;CCAGCCTCTGAGCCCAGAAA
Lockdown capture probes	MPRIP_22_1;TCTTTTTGTTCCTTCAGACATAATGAAATCTAAAAGCAATCCTGACTTCCTGAAGAAAGACAGATCCTGTG TTACCCGGCAACTCAGAAACATCAGGTCCAAGGTGAGTGAGGTCCAGCT MPRIP_23_1;CTTATTTTTTGCTCTTTCCTCTGACAGTCCGTAATTGAGCAGGTCTCATGGGATAACTGAGTCGCGCCCG CTTCCAAGTCCCCTGTCGTATAGGTAAACCACTTCTCACTCTGCCCACTG MPRIP_24.1_1;TCTCCACCAGAGTCTGAAGGAAGGCCTGACGGTGCAGGAGCGCTTGAAGCTCTTTGAATCCAGGGAC TTGAAGAAAGATTAGCTGCATCCATTCCACTGACACATCCTTCCCAGCTGGTG MPRIP_24.1_2;GCAGCACAAGCCAAGCGTGGTTTTCATCTTCACCTGTGCATCTCATTGTTGCTGTCATTAACTGTGGGG ATGGATGCATGAGAGACTGCTGTGGGGTGCTTGCTCCAGTCCCTCCCTGTC MPRIP_24.1_3;AATTCCAACTTCCTGTTGTCTTTATCACAGCTTTTTTCCCATGGTATTCTCAGATTAGTCACGCGCAAGC AGTAGAGTGGGACGGTTTCCCCGCTGATACCTGAGCAGCCCCCTTGTCCT MPRIP_24.1_4;CCCAGCACACTCACTGTCAGTATTAAGGCACAGCAGCCTGTTAATAAGCTAGCCCTGCCTCAGAGTGC ACATGTGGGGACCTGCCAGTGGCTCTGCTCTGCCGCACACAGCGGAGCACCA MPRIP_24.1_5;CCCTTGATGGTCAGAGCAGAGCCACCGAATGTTGTAGGCTGGCCTGATTTCTGAGGGGCCTAGGAAA GTGTCCAGCCTGGCCCACTTCCTTTTGATGTGGGAGGGGGACAGCATTTTGTA MPRIP_24.1_7;CCGGCCCCTCTCATTCCTTCCTCCGCTCGTGCCTCCCTTGCTGTGCTAATCCTGGTCTTCTGGTGGCC CGGAAGAGTGACTTGCAGCCTGAAGCTTCAGACTTGTCAAGTGTGATGGGTC MPRIP_24.1_8;CTCAAAGTCCTCTGGAGGGGAACATTGCTGGGGTTTAATCTGTCTGAAAAGAAGCTGCCATGCTGCTT GGCAGGGTGAGGAGAAGCCCGTGTGCAAAGACAGACTGTCAGGACTGGTGAT MPRIP_24.1_9;AGCAGCAGTGTGTGCTCAGGACCCTCGTTAGTCTGGGCTCTGGAAGGGTCCCTTGAATTGTCTGCTTC TCATGGAGCCATTGACAGGGAGCCTGTGCTAACTCTGCTTTTCTAGCTGAAC MPRIP_24.1_11;CCCATGTCAGATGGGAGAGGCAATAGCCTCATAGCTAGATATGCGATCTGGCTTCTTAGTCCTGGCT TGTTAGTGTTAAGTGGGTCAGATGGATGGTCTCTGAGCGTACAAATTCAACTC

	MPRIP_24.1_13;AGGCAAGCCCTGGGCTCCCCTTTTCACACGTTTTTAAGTTCCACTCTGGCACATGCTCAGTGGTCATT TATGTTTTCAGCATTATGGGCAAGTCAGACCTGTTAAAACAAGCTTTGAGCT MPRIP_24.1_15;CTGCCTTGTGCTTCCTGGGATGCTACAAAGAGCCTGCGGAAGATGCCAAGCCTACCATTCAAAAGGA CACTTGATCTTTTGCTCTTCACCTCAGAATCCTAGGATATGAACTTTGCTGAA MPRIP_24.1_17;CTAGGGCTTGTTTCACGTGTGGGTATCTGGTATGTTGGCCAAATGCTGAACTGGCAGACATTCCTGG ATCATCCCCCACTGTCTACGCCAGTCTGAGCAGTAACATGTGCAGGATCAGGG MPRIP_24.1_19;TCTGTCCATTCCCTTCCCTCACTCCCTGAGCCAGGTTCACTTTTGCCTGACCCCTCTACAGGCTTCCC ATTCCCTGGAAAGAAGTTGAAGTTGGTGTGTTCAGAGAAGAGGTATGGAATG MPRIP_24.1_21;ACATTTGTGGTCTACCAGCATCTACAAGCAAACTACCATTACAGGGAACATACCATTTGACAGTGAAT GTAAATATATGCTTAAATTACATCTTGCAAGCCTGGCATGGTGGTGCACGC MPRIP_24.2_1;AAAAAAACAAAAAAAAAAAAACTTCATCTTTGCACGTGTGATGCCGAGCACATCTGCCACACAGGGCTAC TGGCTGCACCTTTTCCAGTATGGTTGTGGCCACAATGCCACAATAATAGGA MPRIP_24.2_2;GGCTGCACCTTTTCCAGTATGGTTGTGGCCACAATGCCACAATAATAGGAGGCCAGAGCCCCATGAAG TAGGGAGTCCCCAGAGGGAAGAACGCTAGGAATGTGTGCATTCACTGAGTGT MPRIP_24.3_2;TTTGTCCAAAGGGAGCCTGAGAGATTTGACTAGGAAACAAATGGGTATGAAAGATATTTGAAGGAATGT GTACTACTGTTAACCTCTCTACAGGGGGAGATGGTGCGCTGACTTCTTTCT MPRIP_24.3_4;ATCAGCTTAGTTAGATTCCTGCAGGCCTGCCCCACACACCTGACCTCCAACTAAGCCTGGCTCTCAAA GACAGGCTGGATAAGCACCCAGAGGCCCTATGCAGCCTTCCCACTCATGCCA MPRIP_24.3_6;GTTACCTCTTGGCCAGCGGAGGGTTAGCTGTTCATGCTATGCTGTAAAGATGACATCACCTAGTAAGC TACTTAGAGGCTATCAGTGTAGGGACTCATACAGCCTTTCTCATTAGAGCTC MPRIP_24.3_8;TTTGTGCAAAGTTTGTGTGAGTTTTAAAATACCCCTTTTGAAGGTATGTAAAGGCAGTGTTCAGTGCAAA CCCCTCTCAAATGTCAAGACCATTAGTTGGCGAGGCATGGTAGCACACAC MPRIP_24.4_3;CATACGCTGCTTCAGGTCCAGGAGCCTAGCTCTAACAGCGGGGCTGGACTTGAGTGCAGAGCAGCCC AGCTACCCATATCTCTAGCGACCTCACTGTATATAGCATGAGACTCCAGGTAG MPRIP_24.4_5;TGGAGCAGCTTCCATTGTGGGGCAGGACCTCATTGCTGGCATTTGCCATAAAGAGATACCACTGTTAA GAATCCTAGCCACACAAGCCCACTCCAGCTTGCCACAGCTGACATGCCATTG MPRIP_24.4_8;TAAGAACCGCAGTCACAGTGGCTCCTTGACAATCAGAACTCTAGGTGACCTTCCATTACTGCTGCTTGA CTGGGGGTGGGGTGCGTGTTTGTGAACAGGGCGCCTGGCAGGGCCTCTTAT MPRIP_24.4_10;CCAGAGAGGCTTCAAGCACAGCTTCTCAGAGCCAGTTCAGGCCTCTGTGTGTATGCACAGGGTGTGG CTTCCTGCAGACAACATGTCCCCCAACTCCTCATTCCAAGATCTTATGGTTCA¬¬¬

6. References

1. An JJ, Gharami K, Liao GY, Woo NH, Lau AG, Vanevski F, Torre ER, Jones KR, Feng Y, Lu B, et al. 2008. Distinct Role of Long 3′ UTR BDNF mRNA in Spine Morphology and Synaptic Plasticity in Hippocampal Neurons. *Cell* **134**: 175–187.
2. Bae B, Miura P. 2020. Emerging Roles for 3′ UTRs in Neurons. *International Journal of Molecular Sciences 2020, Vol 21, Page 3413* **21**: 3413.
3. Bae Bongmin, Gruner HN, Lynch M, Feng T, Kevin SO, Oliver D, Mastick GS, Wei YAN, Pieraut S, Miura P. 2020. Elimination of Calm1 long 3′-UTR mRNA isoform by CRISPR-Cas9 gene editing impairs dorsal root ganglion development and hippocampal neuron activation in mice. *RNA* **26**: 1414–1430.
4. Bassell GJ, Zhang H, Byrd AL, Femino AM, Singer RH, Taneja KL, Lifshitz LM, Herman IM, Kosik KS. 1998. Sorting of β-Actin mRNA and Protein to Neurites and Growth Cones in Culture. *Journal of Neuroscience* **18**: 251–265.
5. Bauer KE, Segura I, Gaspar I, Scheuss V, Illig C, Ammer G, Hutten S, Basyuk E, Fernández-Moya SM, Ehses J, et al. 2019. Live cell imaging reveals 3′-UTR dependent mRNA sorting to synapses. *Nature Communications 2019 10:1* **10**: 1–13. https://www.nature.com/articles/s41467-019-11123-x (Accessed September 16, 2021).
6. Berkovits BD, Mayr C. 2015. Alternative 3′ UTRs act as scaffolds to regulate membrane protein localization. *Nature 2015 522:7556* **522**: 363–367.
7. Bibel M, Richter J, Lacroix E, Barde Y-A. 2007. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. *Nature Protocols 2007 2:5* **2**: 1034–1043.
8. Bottenstein JE, Sato G. 1985. Cell Culture in the Neurosciences. *Cell Culture in the Neurosciences*.

9. Byrne A, Beaudin AE, Olsen HE, Jain M, Cole C, Palmer T, DuBois RM, Forsberg EC, Akeson M, Vollmers C. 2017. Nanopore long-read RNAseq reveals widespread transcriptional variation among the surface receptors of individual B cells. *Nature Communications 2017 8:1* **8**: 1–11.
10. Cajigas IJ, Tushev G, Will TJ, Tom Dieck S, Fuerst N, Schuman EM. 2012. The Local Transcriptome in the Synaptic Neuropil Revealed by Deep Sequencing and High-Resolution Imaging. *Neuron* **74**: 453–466.
11. Chen CYA, Shyu A Bin. 1995. AU-rich elements: characterization and importance in mRNA degradation. *Trends in Biochemical Sciences* **20**: 465–470.
12. Friling S, Andersson E, Thompson LH, Jönsson ME, Hebsgaard JB, Nanou E, Alekseenko Z, Marklund U, Kjellander S, Volakakis N, et al. 2009. Efficient production of mesencephalic dopamine neurons by Lmx1a expression in embryonic stem cells. *Proceedings of the National Academy of Sciences* **106**: 7613–7618.
13. Garner CC, Tucker RP, Matus A. 1988. Selective localization of messenger RNA for cytoskeletal protein MAP2 in dendrites. *Nature 1988 336:6200* **336**: 674–677.
14. Gronostajski RM, Sadowski PD. 1985. The FLP recombinase of the Saccharomyces cerevisiae 2 microns plasmid attaches covalently to DNA via a phosphotyrosyl linkage. *Molecular and Cellular Biology* **5**: 3274–3279.
15. Ha KCH, Blencowe BJ, Morris Q. 2018. QAPA: A new method for the systematic analysis of alternative polyadenylation from RNA-seq data. *Genome Biology* **19**: 1–18.
16. Hoque M, Ji Z, Zheng D, Luo W, Li W, You B, Park JY, Yehia G, Tian B. 2012. Analysis of alternative cleavage and polyadenylation by 3′ region extraction and deep sequencing. *Nature Methods 2012 10:2* **10**: 133–139.

17. Hubbard KS, Gut IM, Lyman ME, McNutt PM. 2013. Longitudinal RNA sequencing of the deep transcriptome during neurogenesis of cortical glutamatergic neurons from murine ESCs. *F1000Research* **2**: 35.
18. Janesick A, Wu SC, Blumberg B. 2015. Retinoic acid signaling and neuronal differentiation. *Cellular and Molecular Life Sciences 2015 72:8* **72**: 1559–1576.
19. Kislauskis EH, Zhu X, Singer RH. 1994. Sequences responsible for intracellular localization of beta-actin messenger RNA also affect cell phenotype. *Journal of Cell Biology* **127**: 441–451.
20. Lee SH, Mayr C. 2019. Gain of Additional BIRC3 Protein Functions through 3'-UTR-Mediated Protein Complex Formation. *Molecular Cell* **74**: 701-712.e9.
21. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**: 3094–3100.
22. Lianoglou S, Garg V, Yang JL, Leslie CS, Mayr C. 2013. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes & Development* **27**: 2380–2396.
23. Linares AJ, Lin CH, Damianov A, Adams KL, Novitch BG, Black DL. 2015. The splicing regulator PTBP1 controls the activity of the transcription factor Pbx1 during neuronal differentiation. *eLife* **4**.
24. Longo PA, Kavran JM, Kim MS, Leahy DJ. 2013. Transient Mammalian Cell Transfection with Polyethylenimine (PEI). *Methods in Enzymology* **529**: 227–240.
25. Miura P, Shenker S, Andreu-Agullo C, Westholm JO, Lai EC. 2013. Widespread and extensive lengthening of 3′ UTRs in the mammalian brain. *Genome Research* **23**: 812.
26. Moretti F, Rolando C, Winker M, Ivanek R, Rodriguez J, Von Kriegsheim A, Taylor V, Bustin M, Pertz O. 2015. Growth Cone Localization of the mRNA Encoding the

Chromatin Regulator HMGN5 Modulates Neurite Outgrowth. *Molecular and Cellular Biology* **35**: 2035–2050.

27. Mulder J, Ariaens A, Boomen D van den, Moolenaar WH. 2004. p116Rip Targets Myosin Phosphatase to the Actin Cytoskeleton and Is Essential for RhoA/ROCK-regulated Neuritogenesis. *https://doi.org/101091/mbc.e04-04-0275* **15**: 5516–5527.
28. Reyes JH, O′Shea KS, Wys NL, Velkey JM, Prieskorn DM, Wesolowski K, Miller JM, Altschuler RA. 2008. Glutamatergic Neuronal Differentiation of Mouse Embryonic Stem Cells after Transient Expression of Neurogenin 1 and Treatment with BDNF and GDNF: In Vitro and In Vivo Studies. *Journal of Neuroscience* **28**: 12622–12631.
29. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB. 2008. Proliferating Cells Express mRNAs with Shortened 3′ Untranslated Regions and Fewer MicroRNA Target Sites. *Science* **320**: 1643–1647.
30. Shen S, Park JW, Lu ZX, Lin L, Henry MD, Wu YN, Zhou Q, Xing Y. 2014. rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proceedings of the National Academy of Sciences of the United States of America* **111**: E5593–E5601. https://www.pnas.org/content/111/51/E5593 (Accessed January 5, 2021).
31. Smibert P, Miura P, Westholm JO, Shenker S, May G, Duff MO, Zhang D, Eads BD, Carlson J, Brown JB, et al. 2012. Global Patterns of Tissue-Specific Alternative Polyadenylation in Drosophila. *Cell Reports* **1**: 277–289.
32. Terenzio M, Koley S, Samra N, Rishal I, Zhao Q, Sahoo PK, Urisman A, Marvaldi L, Oses-Prieto JA, Forester C, et al. 2018. Locally translated mTOR controls axonal local translation in nerve injury. *Science* **359**: 1416–1421.

33. Tushev G, Glock C, Heumüller M, Biever A, Jovanovic M, Schuman EM. 2018. Alternative 3′ UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments. *Neuron* **98**: 495-511.e6.

34. Ulitsky I, Shkumatava A, Jan CH, Subtelny AO, Koppstein D, Bell GW, Sive H, Bartel DP. 2012. Extensive alternative polyadenylation during zebrafish development. *Genome Research* **22**: 2054–2066.

35. Whipple AJ, Breton-Provencher V, Jacobs HN, Chitta UK, Sur M, Sharp PA. 2020. Imprinted Maternally Expressed microRNAs Antagonize Paternally Driven Gene Programs in Neurons. *Molecular Cell* **78**: 85-95.e8.

36. Wichterle H, Peljto M. 2008. Differentiation of Mouse Embryonic Stem Cells to Spinal Motor Neurons. *Current Protocols in Stem Cell Biology* **5**: 1H.1.1-1H.1.9.

37. Zhang JP, Li XL, Li GH, Chen W, Arakaki C, Botimer GD, Baylink D, Zhang L, Wen W, Fu YW, et al. 2017. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biology* **18**: 1–18.

38. Zhang M, Ergin V, Lin L, Stork C, Chen L, Zheng S. 2019a. Axonogenesis Is Coordinated by Neuron-Specific Alternative Splicing Programming and Splicing Regulator PTBP2. *Neuron* **101**: 690-706.e10.

39. Zhang Z, K S, R P, M B, H N, Y Z, JH K, T K, P M. 2019b. Elav-Mediated Exon Skipping and Alternative Polyadenylation of the Dscam1 Gene Are Required for Axon Outgrowth. *Cell reports* **27**: 3808-3817.e7.

40. Zhu YY, Machleder EM, Chenchik A, Li R, Siebert PD. 2018. Reverse Transcriptase Template Switching: A SMART™ Approach for Full-Length cDNA Library Construction. *https://doi.org/102144/01304pf02* **30**: 892–897.

Chapter IV: Long-read sequencing reveals coordination of alternative splicing and alternative polyadenylation in mouse embryonic stem cell-derived neurons

Alternative splicing (AS) and alternative polyadenylation (APA) are RNA processing events that are responsible for generating multiple mRNA isoforms with different protein-coding exons and 3′ UTRs from a single gene. AS and APA are particularly pervasive during neuronal differentiation. The coupling of AS and APA has been speculated for decades given that multiple RNA-binding proteins can regulate both events. However, evidence showing the direct physical coupling between the AS and APA has been lacking until recently. A direct connection between AS and APA was demonstrated for the *Dscam1* gene in *Drosophila melanogaster* (Zhang et al., 2019). To extend the search for these coordinated events, we have developed a long-read sequencing approach based on the Oxford Nanopore Technology platform called Pull-a-Long-Seq (PL-Seq). PL-Seq has a particular utility in quantifying the coordination of tandem 3′ UTR APA events to upstream cassette exon AS. Here, we apply this approach to mouse embryonic stem cell (mESC)-derived neurons. PL-Seq performed on the *Endov* gene reveals that expression of its long 3′ UTR in neurons is preferentially associated with an exon skipping event located far upstream of the terminal exon.

1. Introduction

Alternative splicing (AS) diversifies the transcriptome from a limited pool of genes. AS is particularly pervasive in the nervous system and AS events in the nervous system are more conserved between vertebrate species than the splicing events in other organs (Barbosa-Morais et al. 2012). Alternative cleavage and polyadenylation (APA), specifically the usage of the distal polyadenylation signal, leads to the production of mRNAs with longer 3′ UTR sequences in the nervous system (Miura et al. 2013). Nervous system specific AS and APA impart transcriptomic diversity which contributes to the functional complexity of the nervous system. Although the functional relevance of each of individual neural-specific AS and APA events has been shown in various cases (Baj et al. 2011; Jensen et al. 2000; Gehman et al. 2011; Linares et al. 2015), whether these RNA processing mechanisms are coordinated is less clear.

In the past decades, the coupling of AS and APA has been speculated given the versatility of protein factors impacting both events (Licatalosi et al. 2008a; Masuda et al. 2015; Lagier-Tourenne et al. 2012; Movassat et al. 2016). That is, some splicing factors have been shown to influence APA or, conversely, 3′ end processing factors have shown to alter upstream splicing patterns. These observations suggested that even a single factor might be able to regulate the processing of mRNAs at multiple levels. *Elav* in *Drosophila* has shown to play a role in both AS and APA (Soller and White 2003; Lee et al. 2021b; Lisbin et al. 2001; Koushika et al. 1996, 2000). However, whether *Elav* impacts the AS and APA independently, thus coincidentally causing both transcriptomic changes or whether AS and APA have some feedback mechanism to selectively couple a particular splicing pattern and 3′ UTR isoform is yet elusive. Recently, more plausible evidence showing the direct co-regulation between the AS and APA has been uncovered. This showed a role of neuronal *Elav* in AS of *Dscam1* gene which was dependent on the long

3′ UTR isoform expression (Zhang et al. 2019). This begs the question of how pervasive the co-regulation of AS and APA is and whether AS and APA are tightly linked in the nervous system.

Conventional short-read RNA-seq techniques rely on mapping 50 – 200 bp reads to a reference genome followed by estimation of transcript isoform abundance (Li and Dewey 2011; Roberts and Pachter 2013) or estimation of transcript expression based on transcriptome k-mers counting (Patro et al. 2014). However, short-read RNA-seq usually falls short of directly connecting a particular pattern of AS to 3′ UTR isoforms. With the advent of long-read sequencing, full-length transcript isoform sequencing is now possible (Sharon et al. 2013; Oikonomopoulos et al. 2016). Current platforms of long-read sequencing have been expanded to a degree where extremely high accuracy long-read sequencing (Travers et al. 2010), detection of epigenetic modification (Flusberg et al. 2010; Rand et al. 2017), and direct RNA sequencing (Garalde et al. 2018) are all feasible.

Our group has recently developed a new long-read sequencing pipeline based on the Oxford Nanopore Technology platform, Pull-a-Long-Seq (PL-Seq; Zhang et al. *in preparation*). Although transcriptome-wide long-read sequencing provides its unique advantages in displaying assembled full-length transcript structure, its biggest limitation is the low depth of long reads covering individual genes (Byrne et al. 2017; Sharon et al. 2013; Soneson et al. 2019; Dong et al. 2021). This is due to sequencing bias toward highly expressed genes and the over-representation of reads that do not span entire transcript isoforms. Our PL-Seq pipeline overcomes this caveat by enriching the sequencing library for genes of interest using a probe-based capture of cDNA. PL-Seq has special utility in quantifying the coordination of tandem 3′ UTR APA events to upstream cassette exon AS events.

PL-Seq in various developmental stages of *Drosophila* has demonstrated direct connectivity between AS and APA events of multiple genes (Zhang et al., *in preparation*). We applied this approach to mouse embryonic stem cell (mESC) derived neurons (mES-neurons). During mESC differentiation to mES-neurons, extensive lengthening of 3′ UTRs and widespread differential cassette exon AS was quantified by analysis of short read RNA-Seq. We performed PL-Seq on one of the 27 genes that showed regulation of both AS and APA during neuronal differentiation – Endonuclease V (*Endov*). PL-Seq performed on *Endov* revealed that expression of its long 3′ UTR in neurons is preferentially associated with exon 4 skipping at far upstream of the terminal exon.

2. Results

Quantification of alternative splicing and alternative polyadenylation during neuronal differentiation of mESCs.

We first decided to generate a list of potential genes that are subject to both AS and APA during neuronal differentiation. For this purpose, we performed mESC differentiation protocol into glutamatergic neurons (Bibel et al. 2007; Hubbard et al. 2013). The expression of mESC or glutamatergic neuron marker genes, *Utf1* and *Grin2b*, respectively, were measured by RT-qPCR for confirmation of each developmental stage. RT-qPCR result showed exclusive expression of each marker in designated cell type (**Figure 15A**). Using these samples, we performed short read total RNA-seq. QAPA analysis (Ha et al. 2018) was used to identify APA genes from the short read data. Distal poly(A) site usage (dPAU) values were calculated for each gene and a fold change increase of at least 25% between the mESC and mES-neuron data was considered as lengthening events (p-value adjusted by false discovery rate, FDR < 0.05). This result demonstrated 233 genes with upregulation of longer 3′ UTR isoforms during neuronal

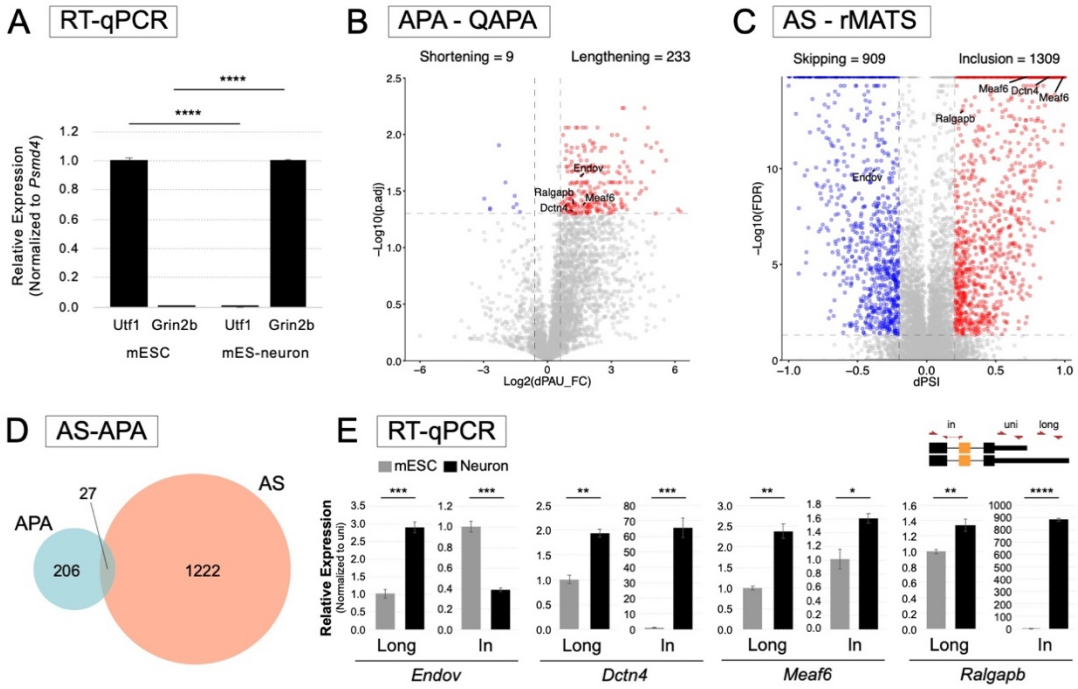


Figure 15. Alternative splicing and alternative polyadenylation during mESC neuronal differentiation.

(A) The expression of mESC or glutamatergic neuron marker genes, *Utf1* and *Grin2b*, respectively, were measured by RT-qPCR for confirmation of each developmental stage. RT-qPCR result showed exclusive expression of each marker in designated cell type. (B) QAPA analysis was performed to identify APA genes. The distal poly(A) site usage (dPAU) values of each gene were considered as lengthening if presented a fold change increase of at least 25% and FDR < 0.05. This result demonstrated 233 genes showing upregulation of longer 3′ UTR isoform during the neuronal differentiation. (C) Alternative splicing events were assessed using rMATS. All the alternative splicing events detected by rMATS were filtered by FDR < 0.05 and the inclusion level difference (dPSI > 0.2). Analysis on our RNA-seq data displayed 2,218 AS events in 1,249 genes. (D) The overlap between these two lists suggested 27 genes are subject to both AS and APA. (E) The expression of the long 3′ UTR, exon inclusion or exon skipping were validated by RT-qPCR. t-test was performed to compare RT-qPCR results. n= 3, * < 0.05; ** < 0.01; *** < 0.001; **** < 0.0001.

differentiation while only 9 showed downregulation (**Figure 15B**). When comparing these 233 genes to a list of 3′ UTR lengthening genes estimated from previously published dataset (Bae and Miura, *submitted – see Chapter 3*), we found 79% of genes showed reliable lengthening provided 185 genes overlapped between two lists (**Supplementary Figure 9A**).

Alternative splicing events were assessed for the same RNA-seq data using rMATS (Shen et al. 2014). All the alternative splicing events detected by rMATS were filtered by FDR < 0.05 and the inclusion level difference (dPSI > 0.2). Analysis on our RNA-seq data displayed 2,218 AS events in 1,249 genes (**Figure 15C**). Among them, 771 genes were detected to be alternatively spliced in both the public data and our data (**Supplementary Figure 9B**). The extend of potential co-regulation of AS and APA was assessed by selecting genes that are found in both lists of 233 3′ UTR lengthening genes and 1,249 alternatively spliced genes. There were 27 genes overlapping, suggesting that they might be subject to both AS and APA (**Figure 15D** and **Supplementary Figure 9C**).

Next, we validated some of the 27 genes for its 3′ UTR lengthening and splicing events during the mESC-derived neurons development. Split RT-qPCR primers at the exon junctions were designed either for forward or reverse primer in a way that the pair specifically detects exon inclusion events. The expression of the long 3′ UTR and exon inclusion were normalized to the expression of the total gene of interest by primers targeting constitutive exons (**Figure 15E**). All of the four targets we tested (*Endov*, *Dctn4*, *Meaf6*, *Ralgapb*) showed the expected trend according to the RNA-seq dataset. Overall, our QAPA and rMATS analysis reliably generated a list of genes subject to both AS and APA processing during the mouse neuronal development.

We performed gene ontology (GO) analysis using TopGO (Alexa et al., 2021). This analysis showed the genes with both AS and 3′ UTR lengthening are enriched in a specific

biological process. When all the 233 lengthening genes were considered as background, 5 out of the 27 genes were significantly enriched in GO term “anatomical structure formation involved in morphogenesis” (**Supplementary Figure 9D**). This suggests that AS-APA regulated genes might play a role in establishing morphological changes during neuronal differentiation.

Pull-a-Long Seq (PL-Seq) identified Endonuclease V (*Endov*) exon 4 connectivity with the 3′ UTR isoforms.

Bioinformatics analysis of RNA-seq data and RT-qPCR validation showed that neuronal differentiation regulates both AS and APA. These analyses, however, fall short in determining whether these two concurrent RNA processing events are tightly coordinated and connected for individual genes. Two scenarios would be possible (**Figure 16**). The first scenario is that the AS is equally impacting all the transcripts regardless of their 3′ UTR expression pattern. In this case, all the 3′ UTR isoforms will have the same coding power while different 3′ UTR sequences will provide post-transcriptional gene regulation. Alternatively, AS pattern could be biased toward a specific 3′ UTR isoform. Under this circumstance, each 3′ UTR isoform could possess differential coding potential translating proteins with different domains or would be subjected to differential RNA metabolism/decay mechanism due to coding frame shift.

To address this hypothesis, we focused on endonuclease 5 (*Endov*) gene. *Endov* is a highly conversed proteins found in all kingdoms which is involved in DNA repair or RNA cleavage (Sebastian Vik et al. 2013; Dalhus et al. 2015). During neuronal differentiation, exon 4 of *Endov* transcript showed increased skipping which was concomitant with expression of the long 3′ UTR (**Figure 17A**). To assess the physical connectivity between AS and APA in *Endov* transcripts, we performed PL-Seq. Briefly,

reverse transcription is performed using SMRT (Switching Mechanism at 5′ end of RNA Template) technology. *Endov* cDNA was then enriched using the probe capture technique (for probe sequences see **Supplementary Table 4**), followed by PCR amplification. Enriched and amplified cDNA was subjected to sequencing library preparation. Sequencing was then performed on the Oxford Nanopore MinION platform.

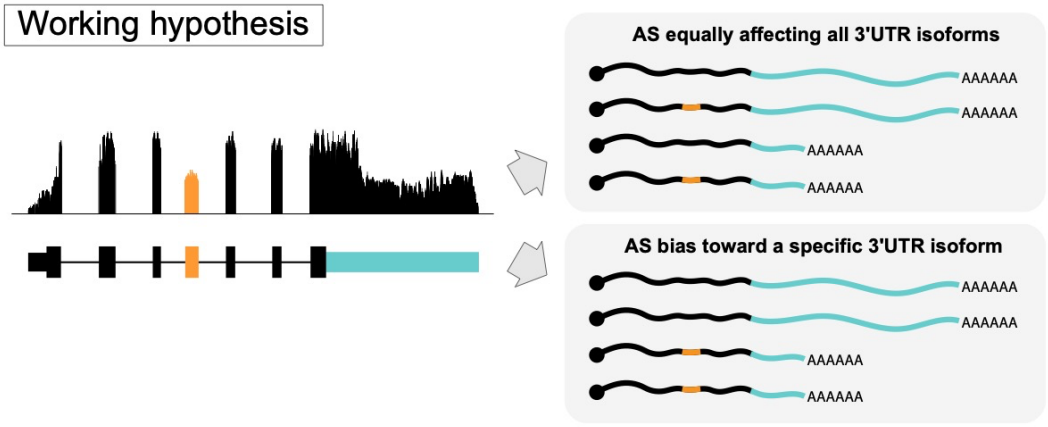


Figure 16. Working hypothesis.
In the first scenario, AS is equally impacting all the transcripts regardless of their 3′ UTR expression pattern. In this case, all the transcripts will have the same coding power while different 3′ UTR sequences will provide post-transcriptional gene regulation. Alternatively, AS pattern could be biased toward a specific 3′ UTR isoform. Under this circumstance, each 3′ UTR isoform would translate different proteins or be subjected to differential RNA metabolism/decay due to coding frame shift.

Raw fastq files were mapped to the reference genome using Minimap2 (Li 2018). From the short-read RNA-seq tracks (**Figure 17A**), the presence of a medium size 3′ UTR isoform in mES-neuron was not obvious. However, PL-Seq tracks reliably showed presence of three distinct 3′ ends for the *Endov* gene (**Supplementary Figure 10**). The

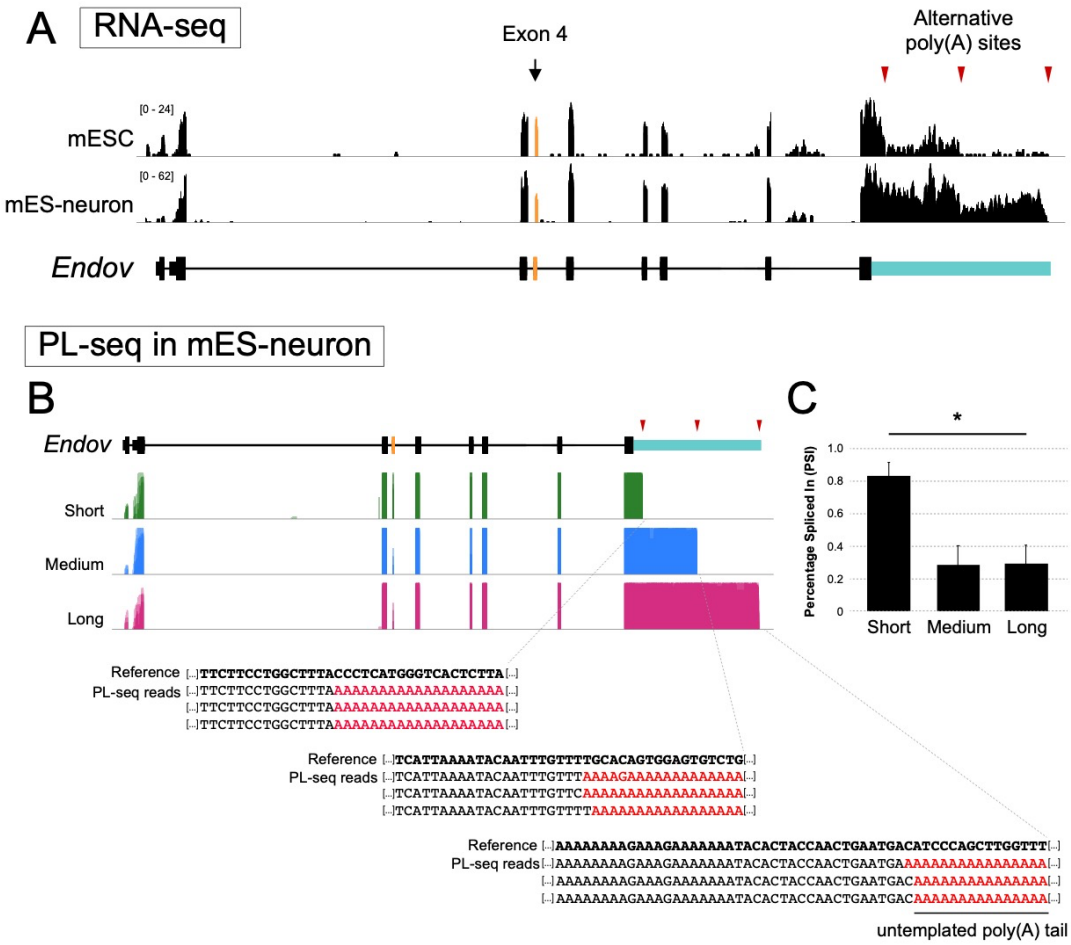


Figure 17. PL-Seq identified *Endov* exon 4 connectivity with 3′ UTR isoforms.
(A) Visualization of short-read RNA-seq supported exon 4 AS and 3′ UTR lengthening during mES-neuron differentiation. (B) PL-Seq was performed for *Endov* to assess the physical connectivity between AS and APA. Each isoform was supported by the presence of untemplated poly(A) tail at the end of the reads. All the coverage tracks from replicates were overlaid for visualization purpose. (C) Calculation of PSI was performed by counting reads that included the exon 4 divided by the total post-filtering reads. The PSI in the *Endov* short 3′ UTR isoform (83.1%) was significantly different from the value in the long 3′ UTR isoform (29.3%). n = 4 biological replicates. Paired t-test was performed. * < 0.05.

three 3′ UTR isoforms were separated using samtools. First, the longest isoform was parsed from the total bam file by filtering reads that overlapped with the long 3′ UTR exclusive region. Subsequently, the rest of the reads were used to parse medium and short 3′ UTR isoforms in a similar way. At this point, not all the reads that matched to the 3′ UTR would provide AS information. Thus, each isoform file was subjected to additional filtering. At least 80% of the upstream exon 3 and downstream exon 5 coverage was required. This step was necessary in order to obtain informative reads that would allow accurate calculation of percentage-spliced-in (PSI) values. Subsequently, at least 80% coverage of each 3′ UTR sequences was also required. These filtering steps largely discarded any reverse transcription artefacts due to random template switching in highly structured region (Schulz et al. 2021) and successfully separated all three 3′ UTR isoforms, each supported by the presence of untemplated poly(A) tail (**Figure 17B** and **Supplementary Figure 11**). Calculation of PSI was performed by counting reads that included the exon 4, then, dividing by the total reads that passes the previous filtering steps. The PSI in the *Endov* short 3′ UTR isoform (83.1%) was significantly different from the value in the long 3′ UTR isoform (29.3%; **Figure 17C**). This result showed that the exon 4 alternative splicing pattern is biased depending on which 3′ UTR isoform is selected.

Independent experimental confirmation of *Endov* AS-APA coupling.

In order to validate the PL-Seq result, alternative experimental confirmations were performed. First, we performed a nested RT-PCR experiment. We amplified mESC or mES-neuron cDNAs from a common region (**Figure 18A**; Fwd to uni) or long 3′ UTR specific region (**Figure 18A**; Fwd to ext). PCR products were gel purified and quantified. Equimolar amount of the first PCR product from each sample was used in a second PCR

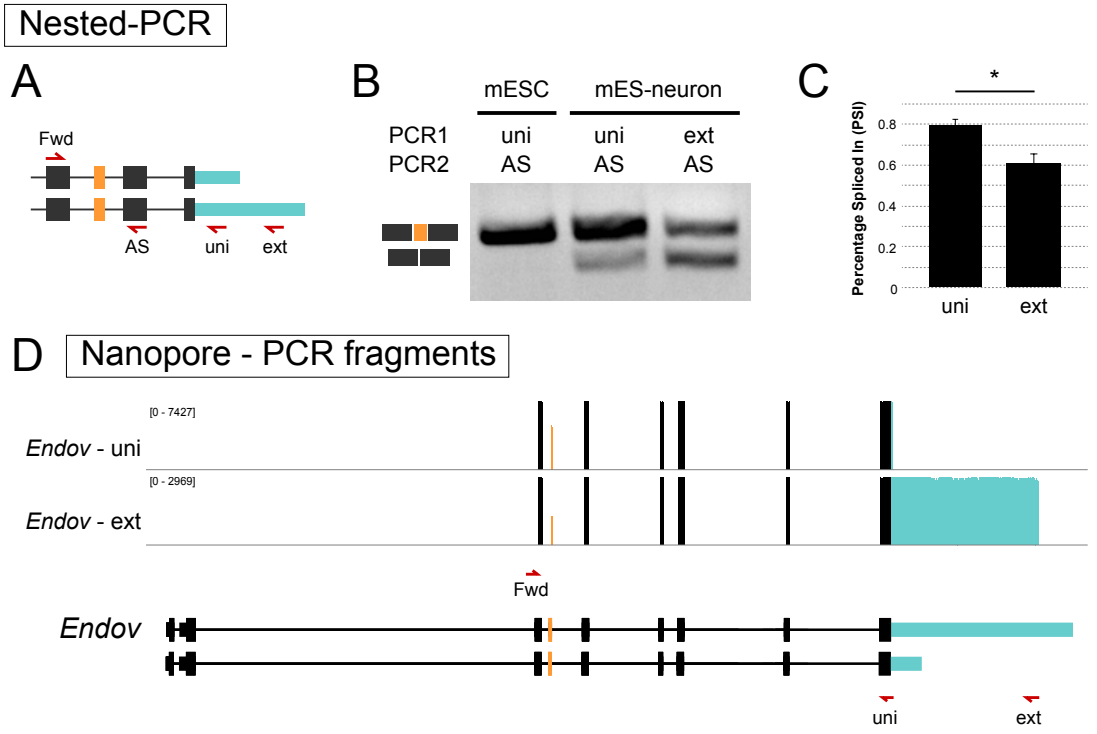


Figure 18. Validation of *Endov* AS-APA coupling.
(A-B) In the first PCR, mESC or mES-neuron cDNAs from a common region (Fwd to uni) or long 3′ UTR specific region (Fwd to ext) were amplified. Equimolar amount of the first PCR product from each sample was used in a second PCR to detect the alternative splicing patterns (Fwd to AS). Nested PCR confirmed that exon 4 is exclusively included in mESC while the extent of exon 4 inclusion varied between the common region PCR product (Fwd to uni) and the long 3′ UTR-specific PCR product (Fwd to ext). (C) Quantification of band intensities confirmed that the PSI in all the *Endov* transcripts (78.3%) significantly differed from the PSI in the long 3′ UTR transcript (61.2%). n = 3 biological replicates. t-test * < 0.05 (D) Long-read library was prepared from the upstream constitutive exon 3 to either uni or ext region using PCR. PCR-based long-read sequencing result also showed similar exon usage bias, inclusion level of the target exon in all the transcripts being higher than in the long 3′ UTR-specific PCR product. n = 1.

to detect the alternative splicing patterns (**Figure 18A**; Fwd to AS). Then, the PCR products were resolved in agarose gel. Nested PCR approach confirmed that exon 4 is exclusively included in mESC while alternatively spliced in mES-neuron (**Figure 18B**). The extent of exon 4 inclusion varied between the common region PCR product (Fwd to uni) and the long 3′ UTR-specific PCR product (Fwd to ext; **Figure 18B**). Quantification of band intensities confirmed that the PSI in all the *Endov* transcripts (78.3%) significantly differed from the PSI in the long 3′ UTR transcript (61.2%; **Figure 18C**).

For a second validation, we performed long-read sequencing of specific RT-PCR products from the *Endov* gene. Long-read library was prepared from the upstream constitutive exon 3 to either uni or ext region using PCR. PCR-based long-read sequencing result also showed similar exon usage bias. The inclusion level of the target exon (**Figure 18D** in orange) in all the transcripts (**Figure 18D**; *Endov*-uni track) was higher than the inclusion level in the long 3′ UTR-specific PCR product (**Figure 18D**; *Endov*-ext track). All our analyses suggested that skipping of exon 4 of *Endov* is differentially regulated in the short and the long 3′ UTR mRNA isoforms.

3. Discussion

AS and APA are two types of co- or post-transcriptional RNA processing mechanisms that contribute to transcriptome diversity. The significance of both AS and APA has been extensively studied to substantiate their key roles in cellular functions. However, whether AS and APA are mechanistically coordinated has been investigated to a less extent. The advent of long-read sequencing has uniquely opened up new avenues for testing the connectivity between different RNA processing events. The PL-Seq approach can specifically assess the coordination of tandem 3′ UTR APA events to upstream cassette exon AS events (Zhang et al. *in preparation*).

The nervous system exhibits an extensive expression of long 3′ UTR transcript isoforms. PL-seq performed in the mouse cortex showed some genes to be consistently and significantly coupled (**Supplementary Figure 12**). However, whether the coupling event is occurring within the same cell type or the detected isoforms are cell type-specific transcripts is ambiguous. In this context, the mES-neuron system is beneficial in addressing the AS-APA coupling hypothesis given the homogeneous cell type with strong APA changes during the differentiation. In the same cell type, neurons, both short and long 3′ UTR transcript isoforms are often expressed, thus it allows a 3′ UTR isoform-specific pair-wise analysis of AS pattern in the same sample.

PL-Seq performed on the *Endov* gene revealed that expression of its long 3′ UTR in mES-neuron is preferentially associated with the exon 4 skipping event (**Figure 17B-C**). Previous attempts on exploring the extent of AS and APA coupling have presented that the AS-APA was limited to the last intron or exon (Movassat et al. 2016; Lee et al. 2021b). However, our example on the *Endov* suggests that PL-Seq could broaden the scope to further upstream exon splicing. Implementing PL-Seq would not only confirm the coordination between APA and the terminal intron/exon splicing events but also further upstream AS events in mouse neurons.

The alternatively spliced exon in the *Endov* transcript is 40 bp in length. Thus, skipping of this exon in mES-neurons would cause a frameshift that is predicted to introduce a premature termination codon (PTC) in its exon 6 (**Supplementary Figure 13**). The introduction of PTC is often associated with nonsense-mediated mRNA decay (NMD) which serves as a surveillance mechanism or post-transcriptional gene regulation mechanism (Kurosaki et al. 2019; Lee et al. 2021a). During splicing of pre-mRNAs, exon junction complex (EJC) is deposited upstream of each exon-exon junction. These EJCs are removed from mRNAs during the first round of translation by ribosomes. However, if

a PTC is introduced and not all the EJCs are displaced, the EJC-dependent NMD mechanism is triggered. The presence of EJC could also impact the translation efficiency by hindering the interaction between the 5′ and 3′ end of mRNAs. Thus, it is intriguing to know whether the long 3′ UTR isoform associated with the skipping of exon 4 is a target of NMD to either regulate mRNA decay or translation efficiency.

A compelling question is which RBPs are involved in the coupling of these two RNA processing mechanisms in mouse neurons. In *Drosophila*, *Elav* was found to induce both neuronal AS and APA of *Dscam1* (Zhang et al. 2019; Lee et al. 2021b). Interestingly, the exon 4 AS retention for the short 3′ UTR isoform of *Endov* in mES-neurons resembled the pattern in mESCs. The neuronal-like exon 4 skipping was prominent for the long 3′ UTR isoform (**Figure 17A-B**). Thus, a single protein factor may act on a transcribing pre-mRNA to simultaneously but independently regulate exon 4 AS and APA of 3′ UTR. Are any of the RBPs reported to be regulators of AS and APA, i.e. Nova, Fus, Hu (Licatalosi et al. 2008b; Masuda et al. 2016; Lee et al. 2021b), responsible for this phenomenon? In the context of AS and APA coupling, does one of these events precede the other one? Whether splicing is required for the APA decision or long 3′ UTR transcription is necessary for AS is still elusive. Studying the transcription kinetics of the long 3′ UTR region and alternative splicing would provide good clues to address this question (Reimer et al. 2021).

Implementation of PL-Seq in the *Endov* gene served as a practical approach in looking into the AS and APA coordination in mammalian neurons. More extensive screening for AS and APA genes is warranted. In future work it will be interesting to use PL-Seq to quantify AS and APA coordination under different cellular contexts, such as neuronal activation (Denkena et al. 2020) and in disease conditions.

4. Materials and Methods

mESC and glutamatergic neuron differentiation

mESC (E14TG2a) cells between passages number 5 – 15 were routinely maintained in mESC medium (DMEM supplemented with 1x Glutamax, FBS, β-mercaptoethanol. MEM non-essential aminoacids, sodium pyruvate, and LIF) on MEF feeder or on gelatin-coated tissue culture dishes. mESC media was replaced daily and split every other day. mESC differentiation protocol has been described in detail in previous chapter. In short, $3.5x10^6$ cells were plated onto 90 mm bacteriological dishes in 15 mL NPC medium (DMEM with L-glutamine (Thermo Scientific 11965092) supplemented with 10% FBS, 1X non-essential amino acids, and 550 µM β-mercaptoethanol). Media change was performed on day 2. On day 4, 6, and 8, media change was performed and 5 µM of retinoic acid was added. On day 10, NPC aggregates were collected and dissociated with 1 mL of TrypLE (Thermo Scientific 12604013) at 37°C for 5-7 min. To halt the reaction, 8 mL of Trypsin inhibitor (Thermo Scientific R007100) was added. The NPC aggregates were gently dissociated by pipetting up and down and filtered through 40 µm cell strainer. Cell suspension was diluted in N2 media (Neurobasal (Thermo Scientific 21103049) supplemented with 1X N2 (Thermo Scientific 17502048) and 2 mM glutamine (Thermo Scientific 25030081)) at $3x10^5$ cells/mL. 10 mL of cells were plated onto PDL (Sigma P7280) coated 100 mm cell culture dishes. Complete media change was performed at 4h (day 10) and 24h (day 11) with N2 media. On day 12 (equivalent of neurons DIV 2) and 14 (DIV 4), media was replaced with B27 media (Neurobasal supplemented with 1X B27 (Thermo Scientific 17504044) and 2 mM glutamine). Cells were maintained until day 17 (DIV 7).

Alternative splicing and alternative polyadenylation analysis

For transcriptome wide AS and APA analyses, we performed mESC differentiation into glutamatergic neurons. To compare our result, we also analyzed publicly available RNA-seq data under BioProject accession PRJNA185305 - mESC (SRR645824 SRR645826 SRR645828) and mES-neurons (SRR645846 SRR645849 SRR645851). For alternative polyadenylation analysis, QAPA 1.3.1 was used with pre-compiled mouse mm10 library provided with this version of software as previous described (Chapter 3 or methods paper). Briefly, only the genes with gene-level TPM values greater than 1 were used for the following analysis. Distal poly(A) usage (dPAU) values were selected by filtering the maximum length of the 3′ UTR sequence used in the analysis. Fold change of dPAU values between mESC and DIV7 neurons was calculated and t-test was performed followed by FDR correction. For alternative splicing analysis, RNA-seq read were aligned to *Mus musculus* GRCm38.102 genome using STAR 2.7. Sorted output bam files were then fed into rMATS 4.0.2 to identify alternatively spliced cassette exons. Output file was filtered by FDR < 0.05 and |IncLevelDiff| > 0.2 to create a high-confidence spliced gene list. TopGO was used for the assessment of gene enrichment for the biological function. This gene ontology analysis was performed by using all the APA genes identified by QAPA analysis as background and genes identified to have both AS and APA as a test group. Data arrangement, statistical analysis, and graph generation was performed in R 4.0.3.

cDNA synthesis

RNA was extracted from cultured cells using Trizol (Ambion) method. Briefly, samples were triturated and lysed in Trizol for 10 min in room temperature. Phase separation was performed by adding 1/5 volume of chloroform and centrifuging at 20,000xg for 15-20 min at 4°C. Upper aqueous phase was precipitated with isopropanol

for 15 min at room temperature or 2 hours at -20°C and centrifuged at 20,000xg for 15-20 min at 4°C. RNA pellet was washed with 70% ethanol followed by centrifugation at 20,000xg for 10 min at 4°C. RNA pellet was resuspended at desired volume of water. RNA was quantified using Nanodrop spectrometer. For cDNA synthesis, 1 µg of DNase (Turbo DNase – Thermo Scientific) treated total RNA was reverse transcribed using Maxima reverse transcriptase (Thermo Scientific).

RT-qPCR

The first strand cDNA reaction was diluted 1:5 before performing RT-qPCR. RT-qPCR was performed using SYBR Select Master Mix for CFX (Thermo Scientific 4472954). BioRad CFX96 real time PCR machine was used and results were analyzed using the ΔΔCt method. Primers used for RT-qPCR are found in **Supplementary Table 5**.

SMARTer cDNA synthesis

For full-length cDNA synthesis, SMARTer PCR cDNA synthesis kit (Clontech) was used according to user manual. Total RNA was DNAse treated either on-column using PureLink DNase (Thermo Scientific 12185010) and PureLink RNA Mini kit (Thermo Scientific 12183020) or with TURBO DNase (Thermo Scientific AM1907). First strand cDNA was synthesized using 700 ng of DNase treated RNA, 3′ SMART CDS Primer II A and SMARTer II A TSO. cDNA was diluted 1:5 in water. cDNA was amplified by optimal (18 – 21) cycles of PCR using Advantage 2 PCR kit (Clontech). For enrichment of cDNA for genes of interest, cDNA was enriched by capturing with biotinylated xGen Lockdown probes (IDT; probe sequences found in **Supplementary Table 4**) and xGen hybridization and wash kit (IDT). Target capture probe design tool was used for the design of 120 bp probes. Probes were selected based on the location (toward the 3′ end) and GC contents

(preferentially between 40 – 65°C). Captured cDNA was purified using AMPure XL beads (Beckman Coulter) then amplified using Takara LA Taq DNA polymerase Hot-Start version (Clontech).

Nested RT-PCR

First PCR was performed using LongAmp Taq 2X master mix (NEB) and 2 - 3 µL of cDNA (1:5 diluted after reverse transcription) in 20 µL for 30 cycles. 2 * 20 µL PCR reactions were pooled together to obtain sufficient material. PCR product was then gel purified using QIAquick gel extraction kit (Qiagen) and eluted in 50 µL of water. Eluted DNA was quantified using Nanodrop spectrometer. For the second PCR, Taq DNA polymerase with standard buffer (NEB) was used with 25 cycles. Equimolar amount of DNA was used for the second PCR (e.g. 5 ng of DNA for 1 kb product or 15 ng of DNA for 3 kb product) according to the Nanodrop quantification. PCR products were resolved in 3% agarose gel and imaged using Gel Doc EZ (BioRad). Exposure time was adjusted to ensure band intensities were not saturated. PSI values were estimated using gel analyzer tool in Fiji.

Nanopore sequencing

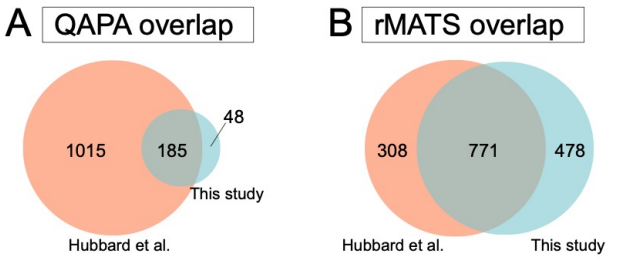
For PCR based nanopore library preparation, the same first PCR product used in nested RT-PCR was ligated with barcoding adapters or new PCR amplicon was obtained using primers with barcoding adapter sequence overhangs. Each sample was barcoded by PCR using Nanopore PCR barcoding kit (EXP-PBC001). Barcoded samples were pooled at the equimolar concentration, end-prepped (NEBNext FFPE DNA Repair Mix and NEBNext Ultra™ II End Repair Kit), and nanopore adapter was ligated (T4 DNA Ligase –

NEB) using Nanopore ligation sequencing kit (SQK-LSK109). For capture-enriched cDNA library preparation, captured and amplified cDNA was end-prepped (NEBNext FFPE DNA Repair Mix and NEBNext Ultra™ II End Repair Kit) and ligated (T4 DNA Ligase – NEB) with sequencing adapters (Nanopore SQK-LSK110). MinION Mk1B device and FLO-FLG001 flow cells were used for sequencing of the libraries.

Nanopore data analysis

Detailed analysis pipeline for PL-Seq is described in the original paper (Zhang et al. *In preparation*). Briefly, Minimap2 was used to map the reads to mouse genome and GRCm38.102 transcriptome bed file was supplemented. Samtools were used to generate and index bam files. Aligned bam files were first visualized in IGV. Different 3′ UTR isoforms were separated using samtools. Genomic coordinates were inputted as a bed file to sequentially split 3′ UTR isoforms starting with the longest 3′ UTR isoforms. Each isoform files were subjected to additional filtering. At least 80% of the upstream and downstream exon coverage was required. Subsequently, at least 80% coverage of each 3′ UTR sequences was also required. Calculation of PSI was performed by counting reads that included target exon, then, dividing by the total reads that passes the previous filtering steps.

5. Supplementary Data



C AS-APA genes

Agap1*	**Gsk3b***	**Pum2***
Bptf*	**Hp1bp3***	**Ralgapb***
Brd8	**Map3k7***	Sh3pxd2a
Dctn4*	**Map4***	Syt7
Endov*	**Meaf6***	Ube2h
Erc1*	Myo5a	Usp13
Eya3	**Necap2***	Wdr4
Fam49b*	**Nfasc***	Zfp142
Gatad2a	Pip4k2a	Zscan22

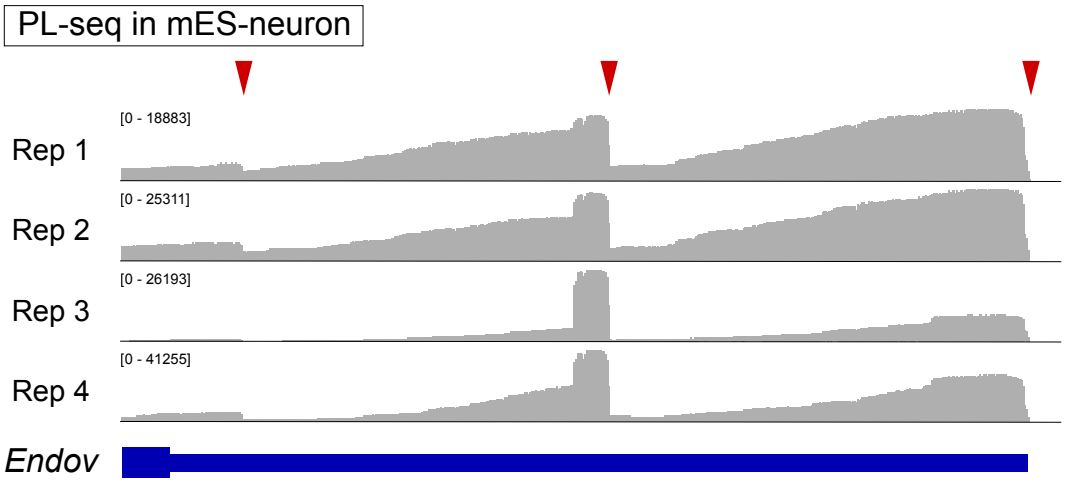
*genes also found in Hubbard et al. RNA-seq data analysis

D GO analysis

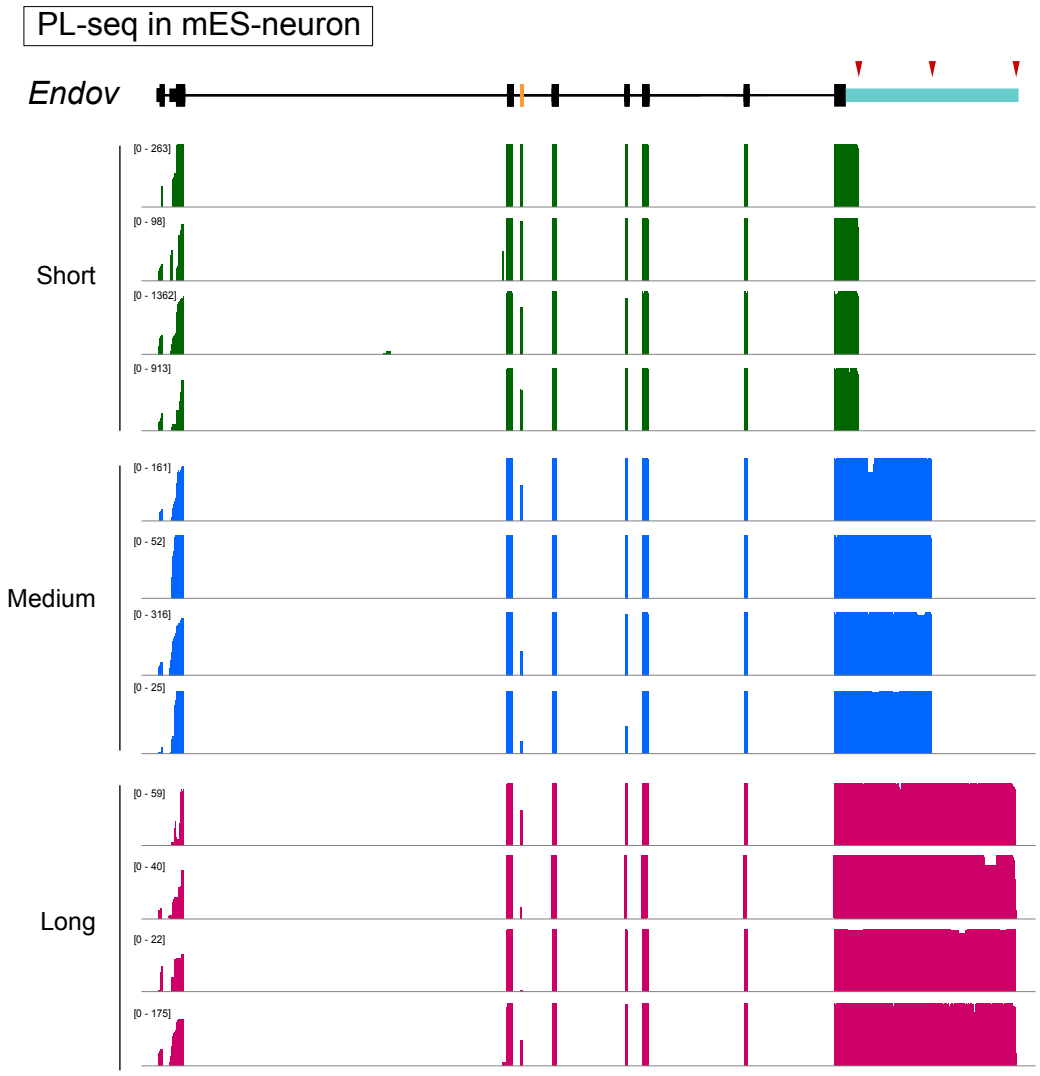
GO.ID	Term	Annotated	Significant	Expected	
GO:0048646	Anatomical structure formation involved in morphogenesis	14	5	1.65	0.0073

Ensembl_gene_id	External_gene_name	Description
ENSMUSG00000022812	Gsk3b	glycogen synthase kinase 3 beta [Source:MGI Symbol;Acc:MGI:1861437]
ENSMUSG00000036180	Gatad2a	GATA zinc finger domain containing 2A [Source:MGI Symbol;Acc:MGI:2384585]
ENSMUSG00000053617	Sh3pxd2a	SH3 and PX domains 2A [Source:MGI Symbol;Acc:MGI:1298393]
ENSMUSG00000026442	Nfasc	neurofascin [Source:MGI Symbol;Acc:MGI:104753]
ENSMUSG00000028284	Map3k7	mitogen-activated protein kinase kinase kinase 7 [Source:MGI Symbol;Acc:MGI:1346877]

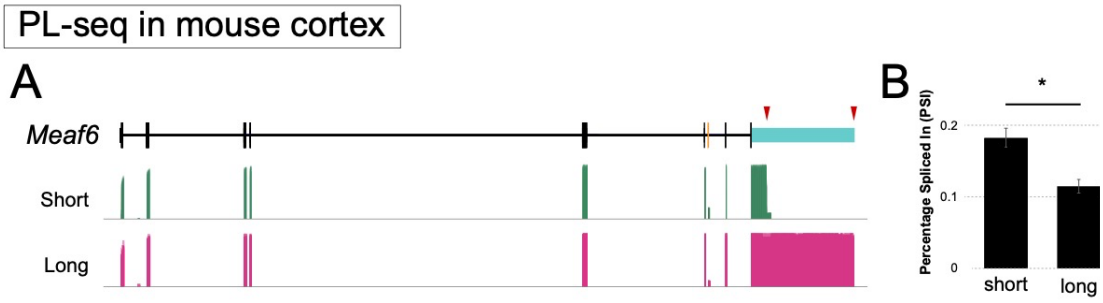
Supplementary Figure 9. Alternative splicing and alternative polyadenylation during mESC neuronal differentiation.
(A) Number of genes with 3′ UTR lengthening was estimated by using published dataset where 1,200 genes showed upregulation. Comparing this with our 233 genes, 185 genes (79%) were found in both lists. (B) Similarly comparing two datasets for AS, 771 genes were detected to be alternatively spliced in both the public and our data. (C) The extend of potential co-regulation of AS and APA was assessed by selecting genes that are found in both lists of 233 3′ UTR lengthening genes and 1,249 alternatively spliced genes. 27 genes that are subject to both AS and APA are shown here. (D) When all the 233 lengthening genes were considered as background, 5 out of the 27 genes were significantly enriched in GO term “anatomical structure formation involved in morphogenesis”.



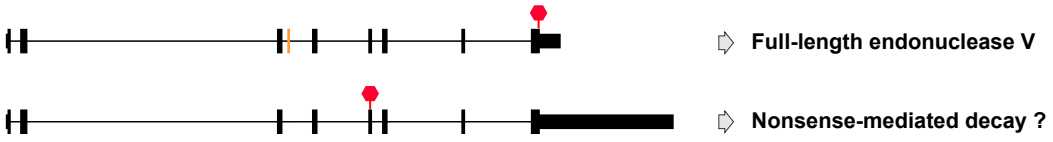
Supplementary Figure 10. Raw read information generated by PL-Seq on *Endov* gene.
Raw fastq files were mapped to the reference genome and visualized in IGV. The presence of a medium size 3′ UTR isoform in mES-neuron was not very evident in short RNA-seq. However, PL-Seq tracks reliably showed presence of three distinct 3′ of transcripts.



Supplementary Figure 11. PL-Seq on *Endov* gene.
PL-Seq pipeline supported successful detection of all three 3′ UTR isoforms. Related to Figure 3B, read coverage tracks are shown here without overlay.



Supplementary Figure 12. Pull-a-long seq in mouse cerebral cortex identified additional AS-APA linkage candidates.
(A) PL-Seq was performed in 11w mouse cortex. Additional gene was confirmed for its potential AS-APA coupling. (B) Quantification of PSI values in the short and long isoform showed significant difference of splicing pattern.



Supplementary Figure 13. *Endov* isoform as a potential NMD target.
Skipping of exon 4 introduces a PTC in *Endov* transcripts. Since PTC is one of the key features of NMD, it is intriguing whether this isoform is a target of NMD.

Supplementary Table 4. List of Lockdown capture probes used in this study.

ENDOV(338371)_3_1;GAACAGGTGGTTGCAGCCAGCATCTGACCTGTCCCTGCTGCCTCGCTGTCAGGTGGTGTATGAGGACAGCCG
CATGGTTGGCCTGAAGGCCCCCTATGTGTCAGGCTTCCTGGCCTTCCG
ENDOV(338371)_3_2;AGAGGTCCCTTTCCTGGTGGAGTTGGTACAGCGGCTGCAAGAGAAGGAACCAGATCTCATGCCCCAGGTAG
GTGCCTCCTAAGGTGAGGAGAACCCCTCTGCTTGTGCTTAGGGAGGGAA
ENDOV(338371)_5_1;GCAGGCTTCGGGGTGGCCTGCCACCTTGGTGTCCTTACAGAGCTGCCATGCATCGGGGTGGCCAAGAAGCT
CCTGCAGGTGGATGGACTGGAGAACAATGCTCTGCACAAGGAGAAGGTG
ENDOV(338371)_6.1_1;TTTACAGTTTTTTTCTGTTTCCCAGATTGTGCTCCTGCAGGCCGGAGGAGACACATTTCCTCTGATAGGCAG
CTCTGGGACTGTCCTGGGAATGGTGAGTGGCCAGAAAGCTGTGCCTTG
ENDOV(338371)_7.1_1;ACATCCTCACAGCCTCAGTCCTGAGGCTGGTCCTGCTCCACAGCAACGTTGCCCATCCATGCTTGTCCTCCT
TCCACAGGCCCTGAGGAGCCATGACCACAGCACCAAGCCCCTCTATGT
ENDOV(338371)_7.1_2;CTCTGTGGGCCACAGAATAAGCCTGGAGGTCGCTGTGCGCCTCACCCACCACTGCTGTAGGTTCCGGATCC
CAGAACCTATACGCCAGGTATGGGGCTGTGGTGGGAGTGGGAGAGTTG
ENDOV(338371)_9.1_1;TCGGCCTTCCCTCTAGGCTGGGCCCTTCATCAGGCCTCAGAGACCCCTGGAGAACAAGGGGACTTCTGACA
CACCCTGGCTTTTTATGTTTCCTGTTTCTTGGTGTCATTACCAGGAGCC
ENDOV(338371)_9.1_3;GCTTCCAGGAGCAGAAGGACCAGCAGTTGGAGGGAACCGGGCATCAGGAAGACTCGGACCTCTGGCCTC
CTTCTCCAGCCTGGGTACAGTCACCACCCTGAGAAGATGATCTCTAGGAGC
ENDOV(338371)_9.1_5;AGCCTCCTCCATCACTGGCTAGTCAGTGTGGTCACCATGCCCCAGCAGGACCATATTCCAGACCAGGAAGC
ACACACAAGCACCAAGAGTACATGAAGGACTTCAGGCCTCGGGTGCCCA
ENDOV(338371)_9.1_7;AGACTGTCATCCGTGTCATAGCCAGGCTGCAGCCCCCAAGGTGCCTGTCCTCAAGCTTGTTTGCCGCTGAG
CTTGAGGGACCTGCAGTTGGGAAAGCCTCAATCCTGGAGCAGGCTCAAG
ENDOV(338371)_9.1_9;GCGTCAAAATCCCATCCCACTAAGAAGGCGGGTCTGTTCTAGCCTGGTGGCTTCTTGTCTGGGTGGCCTTG
AAGGGGCTGCTGAGGCAGGCCAACAGCAGGGCTGCAGTGCCCTCCCTGT
ENDOV(338371)_9.1_11;AGATGGTGGGGAAGCTAGCTACTAAGTCCTGTGGATACCCACAGGGGGTGGTCAGGATGGGATGACGG
TCGGTGATGGGGAAGGACAGACGGGTGGTTCTCCTGTGATACCCAGGCCCAT
ENDOV(338371)_9.1_13;TTTCAGGTTTTGGATATAAGGAAGATAGATATGTGGTTCTGTCACCCAGTAAATGAAGTGTGGGGCAGG
AGGCAGCCTCCCCACCACACACCTGGTATGTATAGGCCAGCCTCCATGGGC
ENDOV(338371)_9.1_14;AAGAACCCCACGGCCACCCAGGGTGTCCGAGCAGGACATTCCTGCAGTGTCCACATCCACCCACTGGCCC
TCCGCAGTCTCTCACAAACATTCTGTGCACACACACACAACCACACATA
ENDOV(338371)_9.2_1;CACTCTCTCTCTCTCTCTTCATATTCCCTGACACTCTTGTGGAAGCCAGAGCTGGAGGGCCTACAGCTGG
CCCCACCATCTGCCTTCCCAACTTCCAGACCTCTCCCTTCTCAAGCCC
ENDOV(338371)_9.2_3;CAGCCGAGCTTGAGGGCCAGCACACAGCATCCTTGTTTGAAGGGAAGCTGTTTAGAGAAGTAGGGCCCCA
GAGGCCAAGGCCATCAGTCCTTAAGACCATCCTCACAGATGTGGAGGAAA
ENDOV(338371)_9.2_5;TCAACTCTTAACAGAAGACCTTCCCCAAGACACAGAAGTCAGGCCTGATGCTGCTCAGCTTGGCACCTTCC
AACACTCAGACTCTCCTCAGCCTGGGCAGAGTAGCTCTCCAGGCTGCTG
ENDOV(338371)_9.2_6;GCTGTACAGGCAGACACTGCCAGGGGCCCACTCCCCACATACTGAGGAGCAGGGCAGAACCTGAACGTG
GCCTTCTAGACAACTTTTATGTGATTTGTTCTGAAAGAAAATCCTTTCAGT
ENDOV(338371)_9.2_9;CAAGGGCCTGTCAGCCCAGCTGCCCCGCTGGGTGTGTGTGGCAGCCCAGATTTTACCTGGCTGAGCTGCTA
CAGGGAGCTAAGATAGTGAGGTGCCAGGGCCAGGGGGATGCAAACTGAA
ENDOV(338371)_9.2_10;TGAAGGAGCAGGGTGGATAGAACATAGAGTGTGGGCAAGTCTGGGGTGGGGCGTTCAGACACCCCAGA
CACCGGTGCTTGAGGGCAGGGGGCCCTAAGTGCAGTGGCGTCTCAGCTCCTT
ENDOV(338371)_9.2_12;ATTTTCATATTCATCTGCAGTTTCTGTGCCCCTTCCCAGCATCCCTACCCCTGTCCGGCAGCTGGGAGAAGT
TGGCTTCCAGGGGGAGGCAACCTAAACTAGCTGCTGTGTCTGTCACTC
ENDOV(338371)_9.2_14;AAGCCATCATGGTTCCTACATACCAGACAGGCCAGGCTAGGGAACCTTGACCCTTCCTCAACACCTCAAG
AGAGCCAATTGCTGTCTCCTGGCTTCTGGGGAGCCCATATGCTCCTTAGA
ENDOV(338371)_9.2_15;TAGACCCCAATGGGAGCATACTGTCCTGCCTCCTCAGAAGAGTCCTTGAAGGGCTAGGTCTCGGGTTGAG
ACCTGGGGATTGCATCTGTACAGGGCTCCAGCCCCAGGTCTCTG
ENDOV(338371)_9.3_1;GGTGTAGTGTCTCCAGAGTCAGCAGGATATCAGCCTGGACCACCTTCCCCTCCCGAGAGTGCCAGGGTGA
ACACCCCCCCCCCAAAAACCCACTTGTCAGCCAAAGATCTGGTATCGAAG
ENDOV(338371)_9.3_3;ATCCCCACTGCACTGCCCTAACCAATTGATTAATTAAAAGCAGAGGCACAGAACAGCGGCACGGGACCACT
GACTTGCAGATGAATGGCCCAAGAAGTAGTCAGGTGGGTGTCGCCGTGG
ENDOV(338371)_9.3_5;AATGGGAAGATAAAGTGCAGGAGCCGGGACTCCAGGCAGAGGCACCTGGATCAGGTCTTGATGGAGAG
GGTGTAGTTTGGGCTTGGGCTGGGATTTTGTCTTGTTTTGTTTTAACAGGGT

Supplementary Table 5. List of primers used in this study.

qP3081_Agap1_uni_F;CAGGACGGAACTCTTTCCCC
qP3082_Agap1_uni_R;CCGTTGTCGTGTGCTATCCT
qP3083_Agap1_ext_F;CCAGTTTTTGCCGTGCTGAA
qP3084_Agap1_ext_R;AAACCCTGGAGGCTCCGATA
qP3085_Agap1_AS_F;ACAGAAAGAAGCACCGGAGG
qP3086_Agap1_in_R;CTTTGCGTTTGGCTTCAGCA
qP3104_Endov_uni_F;GGCAACAGTGCATGGAACTC
qP3105_Endov_uni_R;GCATGGTGACCACACTGACTA
qP3106_Endov_ext_F;CAGGGGGAGGCAACCTAAAC
qP3107_Endov_ext_R;CCTGGCAGTCCACCATTCTT
qP3108_Endov_AS_F;TGAAGGCCCCCTATGTGTCAG
qP3109_Endov_in_R;CAAGAACGACCTGGGGCATGA
qP3221_Dctn4_uni_F;CATCAAAGTCACCCCACAGC
qP3222_Dctn4_uni_R;TGGTCACCTTCTTCCACTGG
qP3223_Dctn4_ext_F;CTCAGTATCTTTTGGCCGGG
qP3224_Dctn4_ext_R;aaGGCAGGCAGACAGGTAAC
qP3225_Dctn4_in_F;CTCACACACAACGGATGAACAAG
qP3226_Dctn4_in_R;CACTACATGAATAGTGTGTTGCG
qP3227_Meaf6_uni_F;TTTTGGGTCATTCGCTGCTG
qP3228_Meaf6_uni_R;AAATGCACAGTTAGCCACCC
qP3229_Meaf6_ext_F;GGGGAAGGGAGAGAGTGTTT
qP3230_Meaf6_ext_R;ATGGGAGGGTTTCAACGTGT
qP3231_Meaf6_in_F;ACGTGATGGAGGATTGATCTGA
qP3232_Meaf6_in_R;GGACATTTGTGGCTGTGAGG
qP3233_Ralgapb_uni_F;CGGAATAAGCAGTTGGAGCC
qP3234_Ralgapb_uni_R;GGAGTTCACTGACCTGGACA
qP3235_Ralgapb_ext_F;CCCCTTCATTTGGCTCTAATTTC
qP3236_Ralgapb_ext_R;TGCATGCCCCAGATTTCAAA
qP3237_Ralgapb_in_F;GACAACAGCATGACTGCATCTTTTATCC
qP3238_Ralgapb_in_R;CCATGGTGTCTGTTGACATTTGG

6. References

1. Alexa A, Rahnenfuhrer J (2021). topGO: Enrichment Analysis for Gene Ontology. R package version 2.44.0.
2. Baj G, Leone E, Chao M V., Tongiorgi E. 2011. Spatial segregation of BDNF transcripts enables BDNF to differentially shape distinct dendritic compartments. *Proc Natl Acad Sci* **108**: 16813–16818.
3. Barbosa-Morais NL, Irimia M, Pan Q, Xiong HY, Gueroussov S, Lee LJ, Slobodeniuc V, Kutter C, Watt S, Çolak R, et al. 2012. The evolutionary landscape of alternative splicing in vertebrate species. *Science (80-)* **338**: 1587–1593.
4. Bibel M, Richter J, Lacroix E, Barde YA. 2007. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. *Nat Protoc* **2**: 1034–1043.
5. Byrne A, Beaudin AE, Olsen HE, Jain M, Cole C, Palmer T, DuBois RM, Forsberg EC, Akeson M, Vollmers C. 2017. Nanopore long-read RNAseq reveals widespread transcriptional variation among the surface receptors of individual B cells. *Nat Commun 2017 81* **8**: 1–11.
6. Dalhus B, Alseth I, Bjørås M. 2015. Structural basis for incision at deaminated adenines in DNA and RNA by endonuclease V. *Prog Biophys Mol Biol* **117**: 134–142.
7. Denkena J, Zaisser A, Merz B, Klinger B, Kuhl D, Blüthgen N, Hermey G. 2020. Neuronal activity regulates alternative exon usage. *Mol Brain 2020 131* **13**: 1–24.
8. Dong X, Tian L, Gouil Q, Kariyawasam H, Su S, De Paoli-Iseppi R, Prawer YDJ, Clark MB, Breslin K, Iminitoff M, et al. 2021. The long and the short of it: unlocking nanopore long-read RNA sequencing data with short-read differential expression analysis tools. *NAR Genomics Bioinforma* **3**.

9. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, Korlach J, Turner SW. 2010. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods 2010 76* **7**: 461–465.
10. Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, Pantic N, Admassu T, James P, Warland A, et al. 2018. Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods 2018 153* **15**: 201–206.
11. Gehman LT, Stoilov P, Maguire J, Damianov A, Lin C-H, Shiue L, Ares M, Mody I, Black DL. 2011. The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. *Nat Genet 2011 437* **43**: 706–711.
12. Ha KCH, Blencowe BJ, Morris Q. 2018. QAPA: A new method for the systematic analysis of alternative polyadenylation from RNA-seq data. *Genome Biol* **19**: 45.
13. Hubbard KS, Gut IM, Lyman ME, McNutt PM. 2013. Longitudinal RNA sequencing of the deep transcriptome during neurogenesis of cortical glutamatergic neurons from murine ESCs. *F1000Research* **2**: 35.
14. Jensen KB, Dredge BK, Stefani G, Zhong R, Buckanovich RJ, Okano HJ, Yang YYL, Darnell RB. 2000. Nova-1 Regulates Neuron-Specific Alternative Splicing and Is Essential for Neuronal Viability. *Neuron* **25**: 359–371.
15. Koushika SP, Lisbin MJ, White K. 1996. ELAV, a Drosophila neuron-specific protein, mediates the generation of an alternatively spliced neural protein isoform. *Curr Biol* **6**: 1634–1641.
16. Koushika SP, Soller M, White K. 2000. The Neuron-Enriched Splicing Pattern of Drosophila erect wing Is Dependent on the Presence of ELAV Protein . *Mol Cell Biol* **20**: 1836–1845.

17. Kurosaki T, Popp MW, Maquat LE. 2019. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat Rev Mol Cell Biol 2019 207* **20**: 406–420.

18. Lagier-Tourenne C, Polymenidou M, Hutt KR, Vu AQ, Baughn M, Huelga SC, Clutario KM, Ling SC, Liang TY, Mazur C, et al. 2012. Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat Neurosci* **15**: 1488–1497.

19. Lee PJ, Yang S, Sun Y, Guo JU. 2021a. Regulation of nonsense-mediated mRNA decay in neural development and disease. *J Mol Cell Biol* **13**: 269–281.

20. Lee S, Wei L, Zhang B, Goering R, Majumdar S, Wen J, Taliaferro JM, Lai EC. 2021b. ELAV/Hu RNA binding proteins determine multiple programs of neural alternative splicing. *PLOS Genet* **17**: e1009439.

21. Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinforma 2011 121* **12**: 1–16.

22. Li H. 2018. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**: 3094–3100.

23. Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, et al. 2008a. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**: 464–469.

24. Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, et al. 2008b. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**: 464–469.

25. Linares AJ, Lin CH, Damianov A, Adams KL, Novitch BG, Black DL. 2015. The splicing regulator PTBP1 controls the activity of the transcription factor Pbx1 during neuronal differentiation. *Elife* **4**.

26. Lisbin MJ, Qiu J, White K. 2001. The neuron-specific RNA-binding protein ELAV regulates neuroglian alternative splicing in neurons and binds directly to its pre-mRNA. *Genes Dev* **15**: 2546–2561.
27. Masuda A, Takeda J ichi, Ohno K. 2016. FUS-mediated regulation of alternative RNA processing in neurons: Insights from global transcriptome analysis. *Wiley Interdiscip Rev RNA* **7**: 330–340.
28. Masuda A, Takeda JI, Okuno T, Okamoto T, Ohkawara B, Ito M, Ishigaki S, Sobue G, Ohno K. 2015. Position-specific binding of FUS to nascent RNA regulates mRNA length. *Genes Dev* **29**: 1045–1057.
29. Miura P, Shenker S, Andreu-Agullo C, Westholm JO, Lai EC. 2013. Widespread and extensive lengthening of 39 UTRs in the mammalian brain. *Genome Res* **23**: 812–825.
30. Movassat M, Crabb TL, Busch A, Yao C, Reynolds DJ, Shi Y, Hertel KJ. 2016. Coupling between alternative polyadenylation and alternative splicing is limited to terminal introns. *RNA Biol* **13**: 646–655.
31. Oikonomopoulos S, Wang YC, Djambazian H, Badescu D, Ragoussis J. 2016. Benchmarking of the Oxford Nanopore MinION sequencing for quantitative and qualitative assessment of cDNA populations. *Sci Reports 2016 61* **6**: 1–13.
32. Patro R, Mount SM, Kingsford C. 2014. Sailfish enables alignment-free isoform quantification from RNA-seq reads using lightweight algorithms. *Nat Biotechnol 2014 325* **32**: 462–464.
33. Rand AC, Jain M, Eizenga JM, Musselman-Brown A, Olsen HE, Akeson M, Paten B. 2017. Mapping DNA methylation with high-throughput nanopore sequencing. *Nat Methods 2017 144* **14**: 411–413.

34. Reimer KA, Mimoso CA, Adelman K, Neugebauer KM. 2021. Co-transcriptional splicing regulates 3′ end cleavage during mammalian erythropoiesis. *Mol Cell* **81**: 998-1012.e7.

35. Roberts A, Pachter L. 2013. Streaming fragment assignment for real-time analysis of sequencing experiments. *Nat Methods 2012 101* **10**: 71–73.

36. Schulz L, Torres-Diz M, Cortés-López M, Hayer KE, Asnani M, Tasian SK, Barash Y, Sotillo E, Zarnack K, König J, et al. 2021. Direct long-read RNA sequencing identifies a subset of questionable exitrons likely arising from reverse transcription artifacts. *Genome Biol 2021 221* **22**: 1–12.

37. Sebastian Vik E, Sameen Nawaz M, Strøm Andersen P, Fladeby C, Bjørås M, Dalhus B, Alseth I. 2013. Endonuclease V cleaves at inosines in RNA. *Nat Commun 2013 41* **4**: 1–7.

38. Sharon D, Tilgner H, Grubert F, Snyder M. 2013. A single-molecule long-read survey of the human transcriptome. *Nat Biotechnol 2013 3111* **31**: 1009–1014.

39. Shen S, Park JW, Lu ZX, Lin L, Henry MD, Wu YN, Zhou Q, Xing Y. 2014. rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci U S A* **111**: E5593–E5601.

40. Soller M, White K. 2003. ELAV inhibits 3′-end processing to promote neural splicing of ewg pre-mRNA. *Genes Dev* **17**: 2526–2538.

41. Soneson C, Yao Y, Bratus-Neuenschwander A, Patrignani A, Robinson MD, Hussain S. 2019. A comprehensive examination of Nanopore native RNA sequencing for characterization of complex transcriptomes. *Nat Commun 2019 101* **10**: 1–14.

42. Travers KJ, Chin C-S, Rank DR, Eid JS, Turner SW. 2010. A flexible and efficient template format for circular consensus sequencing and SNP detection. *Nucleic Acids Res* **38**: e159–e159.

43. Zhang Z, So K, Peterson R, Bauer M, Ng H, Zhang Y, Kim JH, Kidd T, Miura P. 2019. Elav-Mediated Exon Skipping and Alternative Polyadenylation of the Dscam1 Gene Are Required for Axon Outgrowth. *Cell Rep* **27**: 3808-3817.e7.

General discussion

Gene regulation involves multi-level mechanisms including structural and epigenetic changes at the genome level, transcriptional activation/repression, protein translation regulation and post-translational modifications. Post-transcriptional regulation is another layer of gene regulation that influences RNA metabolism once transcription is completed. Post-transcriptional regulation mediates gene expression via stabilization or degradation, transportation to subcellular compartment, or translation efficiency of mRNAs. In post-transcriptional regulation, the 3′ UTR is a region of particular relevance. Sequence or structure motifs found in 3′ UTRs recruit *trans*-factors, such as miRNAs and RBPs, to modulate the fate of transcripts (**Figure 1**). In this sense, APA has a significant impact on post-transcriptional regulation. This RNA processing event can give rise to two or more transcript molecules with distinct 3′ UTR contexts; thus, these transcripts isoforms are potentially subject to differential post-transcriptional regulation.

Deep transcriptomic analyses presented strong evidence that APA is pervasive, promoting longer 3′ UTR transcript isoform expression in the nervous system (Smibert et al., 2012; Miura et al., 2013). Transcriptome-wide findings that mRNAs are asymmetrically distributed within the human or rodent CNS and sensory neurons (Gumy et al., 2011; Minis et al., 2014; Bigler et al., 2017) and evidence that 3′ UTR harbors localizing motifs have strengthened the hypothesis that the 3′ UTR is involved in subcellular localization (Tushev et al., 2018; Ciolli Mattioli et al., 2019). Differential transcript localization coupled with localized translation confers asymmetrical protein distribution and local concentration. Additionally, it implies that translating proteins are exposed to different cellular contexts where distantly localized proteins can interact with different protein partners to regulate cellular functions (**Figure 4A, C**). On some occasions, transcript isoforms localized to the same subcellular compartment yet can mediate gene expression distinctly. For instance,

long 3′ UTR can serve as a scaffold to recruit diverse proteins to enhance their interaction with translating proteins (**Figure 4B, D**).

The contribution of the 3′ UTR in gene regulation becomes clearer as more cellular and physiological roles of 3′ UTR isoforms are being uncovered. The work presented in Chapter 2 showed differential post-transcriptional regulation of the *Calm1* 3′ UTR isoforms, *Calm1-S* and *Calm1-L*, in CNS and PNS neurons. Subcellular localization pattern varies for each isoform between different neuron types (**Figure 6**). More importantly, *Calm1-L* plays a critical role in both neuron types by mediating axon development in DRG neurons and neuronal activation in the hippocampus (**Figure 8,10**). In the future, we can reinforce this finding by performing a rescue experiment. In the context of *Calm1*$^{\Delta L/\Delta L}$ mice or primary neurons, we could transduce *Calm1-S* or *Calm1-L* expression construct using a viral vector. If the phenotypes observed above are due to an exclusive role of *Calm1-L*, and not *Calm1-S*, only the viral delivery of *Calm1-L* will restore the axon fasciculation and neuronal activation defects.

In our analysis, the steady-state *Calm1* mRNA and CaM proteins levels do not appear to vary in the WT and *Calm1*$^{\Delta L/\Delta L}$ tissues (**Figure 8g, 9a-b**). The overall unchanged mRNA levels upon the distal poly(A) site deletion, which was also observed independently in the *Mprip* gene (**Figure 14C**), suggested to us that in the absence of the alternative polyadenylation options, i.e. no distal poly(A) signal, the 3′ end processing machinery acts on the proximal poly(A) site to compensate the total *Calm1* expression by generating *Calm1-S*. The unchanged steady-state protein level would also support this assertion. Having in mind that the role of the *Calm1* long 3′ UTR might not be directly associated with translational control, what is still lacking is the exact mechanism by which *Calm1-L* loss causes the identified phenotypes. Independent of 3′ UTR-mediated translational regulation, 3′ UTRs can play a role as protein scaffolds. A plausible speculation is that

Calm1-L is necessary for the regulation of a particular signaling pathway downstream of CaM. In both PNS axonal phenotype and CNS neuron activation phenotype, CaM-dependent activation of CREB is likely involved. Mice lacking CREB displayed impaired axonal growth and projection in DRG (Lonze et al., 2002). Neuronal activation increased CREB phosphorylation and subsequence expression of cFos (Moore et al., 1996). Since nuclear translocation of CaM influences CREB activation (Deisseroth et al., 1998; Mermelstein et al., 2001; Cohen et al., 2018), it would be intriguing to determine whether *Calm1*$^{\Delta L/\Delta L}$ can directly impair this signaling pathway by reducing CaM nuclear translocation. Does *Calm1-L*, and not *Calm1-S*, recruit specific CaM binding partners, such as γCaMKII (Ma et al., 2014), via the long 3′ UTR scaffold to mediate its nuclear translocation?

Impaired cFos induction upon enrichment environment exposure has more implications in synaptic plasticity. Rapid induction of IEG expression leads to plasticity-associated gene reprogramming which is necessary during diverse learning paradigms (Minatohara et al. 2016; Yap and Greenberg, 2018). Thus, it would be interesting to expand on the impact of *Calm1-L* loss on postsynaptic plasticity. It is intriguing whether the hippocampal activation phenotype in *Calm1*$^{\Delta L/\Delta L}$ can further impact spine density or spine morphology changes in *Calm1*$^{\Delta L/\Delta L}$ hippocampal neurons. Assessing for deficits in synaptic transmission and plasticity, such as long-term potentiation and long-term depression, could also potentialize the role of *Calm1-L* in synaptic plasticity.

Despite clear interest in the molecular and physiological roles of 3′ UTRs, only a very small fraction of APA genes is known for their 3′ UTR function. Why are the functional studies so limited to a couple of genes? Generating a knockout mouse would take a great deal of time in terms of establishing the line. Having the knockout strategy setup in a cultured system might substantially improve the pace of functional screening. Given that

the long 3′ UTRs expression is particularly extensive in neurons, a cultured neuron system would be ideal. However, neuroblast cell lines, like N2A, are not an optimal system in mimicking long 3′ UTRs expression patterns found in brain tissues. Primary neurons, on the other hand, do exhibit the same APA patterns as brain tissues (data not shown), but unfortunately are not amenable for CRISPR genetic manipulation. Thus, we reasoned mES-neurons would be favorable for studying long 3′ UTR isoforms. As shown in Chapter 3, the combination of CRISPR and mES-neurons enables endogenous locus study. Given these neurons' polarized morphology, localization of 3′ UTR isoforms can be easily assessed. Then, a loss-of-function experiment needs to be performed to confirm that long 3′ UTR is indeed necessary to localize transcripts into the neuronal processes.

A transcriptome-wide screen of all APA genes would be still complex due to the dual-sgRNA nature of our strategy where a pair of sgRNA need to be cloned into a single construct and HDR donor need to be prepared for each of the individual target genes. Precise optimization at each step might be required. For instance, massive cloning of dual sgRNAs can be facilitated by using cloning strategies based on type IIS endonucleases (Adikusuma et al., 2017). A shorter HDR homology arm can be tested so that shorter DNA oligos can be synthesized instead of PCR amplification from the genomic DNA. Alternatively, the HDR can be completely omitted by relying on the clonal selection using the Cas9-GFP tag and flow cytometry. In this case, the cell recovery strategy needs to be carefully established so that it minimized cellular stress and maximizes individual cell growth in isolation.

Having an easily manageable system expands our ability to explore more regulatory aspects of 3′ UTRs. Based on a previous report showing a coregulation of AS and APA in the *Dscam1* gene in *Drosophila melanogaster* (Zhang et al., 2019; also describes in Chapter 1 **Figure 4E**), we extended the search for AS and APA coupling

genes in mES-neurons. In Chapter 4, our findings on AS pattern bias between *Endov* 3′ UTR isoforms are described. How these alternative AS and APA isoforms post-transcriptionally modulate gene expression is an important question that remains elusive. The long 3′ UTR isoform associated with neuronal exon 4 splicing originates one of the key features of NMD, a premature termination codon. In addition, harboring longer 3′ UTRs has been regarded as an attribute that distinguishes a transcript as a likely NMD target (Hogg and Goff, 2010). Thus, it is intriguing if *Endov-L* is subject to NMD-mediated post-transcriptional regulation.

RNAs undergo extensive structural changes during neuronal differentiation (Wang et al., 2021), which might be associated, at least in part, with the emergence of longer 3′ UTRs during differentiation. RNA structure changes are associated with differential RBP binding dynamics and changes in RNA processing, such as alternative splicing (Taliaferro et al., 2016; Wang et al., 2021). Thus, it is intriguing whether the association of pre-mRNA or mRNA molecules with longer 3′ UTRs induces changes in their intermolecular interaction and RNA structure and could lead to splicing changes. Investigating whether upstream splicing of exon 4 in *Endov* transcripts is affected by the structural changes caused by long 3′ UTR would be interesting.

The role of 3′ UTRs in regulating gene expression has been well regarded. Our work on *Calm1* further supported the physiological importance of the 3′ UTR-mediated post-transcriptional regulation. Some neurological disease states, such as Huntington's and Parkinson's disease, have been associated with altered 3′ UTR isoform ratios of *HTT* and *SNCA* genes. Multiple SNPs have been found in 3′ UTRs of genes associated with Tourette syndrome, Parkinson's, and Alzheimer's disease (Chapter 1 section 8). Investigating the impact of 3′ UTR isoform ratios and SNPs on gene regulation is necessary to understand whether these factors influence etiopathogenesis. Probing the

implication of 3′ UTR SNPs, such as the ones found in the neuronal genes, *DAT1*, *CHNRNA6*, and *SNCA*, or in the disease genes, *APP* and *HTT*, in disease susceptibility or pathogenesis would be a goal. Along the same lines, validation of genome-wide association study results, e.g. association of 3′ APA QTLs and Alzheimer's disease (Li et al., 2021), would corroborate the role of 3′ UTR-mediated gene regulation in disease. SNPs identified in patients can be simulated in mES-neurons and mouse models, or perhaps in human induced pluripotent cells, using gene-editing tools to systematically examine the molecular and cellular roles of the 3′ UTRs. The causal implication of these genetic features will offer a better understanding of how 3′ UTRs contribute to disease pathology.

References

1. Adikusuma, F., Pfitzner, C., and Thomas, P. Q. (2017). Versatile single-step-assembly CRISPR/Cas9 vectors for dual gRNA expression. PLoS One 12, e0187236. doi:10.1371/JOURNAL.PONE.0187236.
2. Bigler, R. L., Kamande, J. W., Dumitru, R., Niedringhaus, M., and Taylor, A. M. (2017). Messenger RNAs localized to distal projections of human stem cell derived neurons. Sci. Rep. 7, 611. doi:10.1038/s41598-017-00676-w.
3. Ciolli Mattioli, C., Rom, A., Franke, V., Imami, K., Arrey, G., Terne, M., et al. (2019). Alternative 3′ UTRs direct localization of functionally diverse protein isoforms in neuronal compartments. Nucleic Acids Res. 47, 2560–2573. doi:10.1093/nar/gky1270.
4. Cohen, S. M., Suutari, B., He, X., Wang, Y., Sanchez, S., Tirko, N. N., et al. (2018). Calmodulin shuttling mediates cytonuclear signaling to trigger experience-dependent transcription and memory. Nat. Commun. 9, 2451. doi:10.1038/s41467-018-04705-8.
5. Deisseroth, K., Heist, E. K., and Tsien, R. W. (1998). Translocation of calmodulin to the nucleus supports CREB phosphorylation in hippocampal neurons. Nat. 1998 3926672 392, 198–202. doi:10.1038/32448.
6. Gumy, L. F., Yeo, G. S. H., Tung, Y. C. L., Zivraj, K. H., Willis, D., Coppola, G., et al. (2011). Transcriptome analysis of embryonic and adult sensory axons reveals changes in mRNA repertoire localization. RNA 17, 85–98. doi:10.1261/rna.2386111.
7. Hogg, J. R., and Goff, S. P. (2010). Upf1 Senses 3′UTR Length to Potentiate mRNA Decay. Cell 143, 379–389. doi:10.1016/J.CELL.2010.10.005.
8. Li, L., Huang, K.-L., Gao, Y., Cui, Y., Wang, G., Elrod, N. D., et al. (2021). An atlas of alternative polyadenylation quantitative trait loci contributing to complex trait and disease heritability. Nat. Genet. 2021 537 53, 994–1005. doi:10.1038/s41588-021-00864-5.

9. Lonze, B. E., Riccio, A., Cohen, S., and Ginty, D. D. (2002). Apoptosis, Axonal Growth Defects, and Degeneration of Peripheral Neurons in Mice Lacking CREB. Neuron 34, 371–385. doi:10.1016/S0896-6273(02)00686-4.

10. Ma, H., Groth, R. D., Cohen, S. M., Emery, J. F., Li, B., Hoedt, E., et al. (2014). γCaMKII Shuttles Ca2+/CaM to the Nucleus to Trigger CREB Phosphorylation and Gene Expression. Cell 159, 281–294. doi:10.1016/J.CELL.2014.09.019.

11. Mermelstein, P. G., Deisseroth, K., Dasgupta, N., Isaksen, A. L., and Tsien, R. W. (2001). Calmodulin priming: Nuclear translocation of a calmodulin complex and the memory of prior neuronal activity. Proc. Natl. Acad. Sci. 98, 15342–15347. doi:10.1073/pnas.211563998.

12. Minatohara, K., Akiyoshi, M., and Okuno, H. (2016). Role of immediate-early genes in synaptic plasticity and neuronal ensembles underlying the memory trace. Front. Mol. Neurosci. 8, 78. doi: 10.3389/fnmol.2015.00078.

13. Minis, A., Dahary, D., Manor, O., Leshkowitz, D., Pilpel, Y., and Yaron, A. (2014). Subcellular transcriptomics-Dissection of the mRNA composition in the axonal compartment of sensory neurons. Dev. Neurobiol. 74, 365–381. doi:10.1002/dneu.22140.

14. Miura, P., Shenker, S., Andreu-Agullo, C., Westholm, J. O., and Lai, E. C. (2013). Widespread and extensive lengthening of 3′ UTRs in the mammalian brain. Genome Res. 23, 812. doi:10.1101/GR.146886.112.

15. Moore, A. N., Neal Waxham, M., and Dash, P. K. (1996). Neuronal Activity Increases the Phosphorylation of the Transcription Factor cAMP Response Element-binding Protein (CREB) in Rat Hippocampus and Cortex. J. Biol. Chem. 271, 14214–14220. doi:10.1074/JBC.271.24.14214.

16. Smibert, P., Miura, P., Westholm, J. O., Shenker, S., May, G., Duff, M. O., et al. (2012). Global Patterns of Tissue-Specific Alternative Polyadenylation in Drosophila. Cell Rep. 1, 277–289. doi:10.1016/j.celrep.2012.01.001.

17. Taliaferro, J.M., Lambert, N.J., Sudmant, P.H., Dominguez, D., Merkin, J.J., Alexis, M.S., Bazile, C.A., and Burge, C.B. (2016). RNA sequence context effects measured in vitro predict in vivo protein binding and regulation. Mol. Cell. 64, 294–306. doi: 10.1016/j.molcel.2016.08.035.

18. Tushev, G., Glock, C., Heumuller, M., Biever, A., Jovanovic, M., and Schuman, E. M. (2018). Alternative 3' UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments. Neuron 98, 495–511. doi:10.1016/j.neuron.2018.03.030.

19. Wang J., Zhang, T., Yu, Z., Tan, W.T., Wen, M., Shen, Y., Lambert, F.R.P., Huber, R.G., and Wan, Y. (2021). Genome-wide RNA structure changes during human neurogenesis modulate gene regulatory networks. Mol. Cell. 81, 1-12. doi:10.1016/j.molcel.2021.09.027ll.

20. Yap, E.L., Greenberg M.E. (2018). Activity-Regulated Transcription: Bridging the Gap between Neural Activity and Behavior. Neuron 100, 330–348. doi: 10.1016/j.neuron.2018.10.013.

21. Zhang, Z., So, K., Peterson, R., Bauer, M., Ng, H., Zhang, Y., et al. (2019). Elav-Mediated Exon Skipping and Alternative Polyadenylation of the Dscam1 Gene Are Required for Axon Outgrowth. Cell Rep. 27, 3808-3817.e7. doi:10.1016/j.celrep.2019.05.083.