2017

Deconvoluting Cell-Type Specific 3'UTR Isoform Expression in the Adult and Developing Cerebellum

Saša Jereb

DECONVOLUTING CELL-TYPE SPECIFIC 3'UTR ISOFORM EXPRESSION IN THE ADULT AND DEVELOPING CEREBELLUM

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Saša Jereb

June 2017

DECONVOLUTING CELL-TYPE SPECIFIC 3'UTR ISOFORM EXPRESSION IN THE ADULT AND DEVELOPING CEREBELLUM

Saša Jereb, Ph.D.

The Rockefeller University 2017

Alternative polyadenylation has been implicated in the regulation of mRNA translation and stability, as well as mRNA and protein localization. However, it is unclear to what extent alternative polyadenylation regulates these processes uniquely in specific cell types. Using a newly developed technique, termed conditionally-tagged poly(A) binding protein-mediated mRNA 3' end retrieval by crosslinking immunoprecipitation (cTag-PAPERCLIP), we discovered significant differences in alternative polyadenylation between granule and Purkinje cells in the mouse cerebellum, as well as between proliferating and adult granule cells. Interestingly, among transcripts that differed in alternative polyadenylation in both these comparisons, many were involved in key neuronal functions. We further characterize one such transcript, *Memo1*, which changed poly(A) site usage during granule cell development, and was concomitantly downregulated. We show that miR-124 specifically targets the long isoform of *Memo1*, which provides a mechanism for its downregulation during development. Our findings give insight into the roles for alternative polyadenylation in specific cell types and developmental stages, and help establish a platform for further functional studies.

ACKNOWLEDGMENTS

First, I would like to express my gratitude to my advisor, Dr. Robert Darnell, for his unwavering support, patience, and encouragement throughout my graduate studies.

I am also indebted to Dr. Mary E. Hatten for her guidance and generosity. I am grateful to Dr. Alexander Tarakhovsky for helpful conversations throughout the years and Dr. David Allis for his advice and reassuring optimism. Also, I am thankful to Dr. Bin Tian for accepting a position on my committee as the external examiner.

I am grateful to the members of the Darnell and Hatten labs for advice and help with reagents. They all provided many informative discussions and have been a source of support throughout the years. I am especially grateful to Dr. Hun-Way Hwang for sharing with me the mouse model he created and the protocol he developed that was critical for my thesis work. I am grateful to Dr. Xiaodong Zhu and Dr. Eve-Ellen Govek for teaching me how to prepare cerebellar slices, do microscopy and many reagents they shared with me. I am thankful to Yin Fang who taught me how to purify granule cells. Dr. Chris Park answered many of my data analysis related questions, for which I am extremely grateful. John Fak helped me with preparing RNA-sequencing library that was critical to my project. Debra Mandli provided help and administrative support. I am also grateful to Dr. Jennifer Darnell for helpful discussions over the years.

I am thankful to Dean Dr. Sid Strickland and Associate Dean Dr. Emily Harms for all their support throughout the years. I also thank Kristen Cullen, Marta Delgado, Cristian Rosario and Stephanie Fernandez for their commitment to graduate students.

Finally, I would like to thank my friends and family who have supported me throughout the years. Jonathan Green, my better half, has been a constant source of encouragement throughout my graduate studies, for which I am truly grateful.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 | Introduction

Alternative polyadenylation is a process where different ends to an mRNA transcript are determined. These alternative mRNA isoforms differ in regulatory 3′ untranslated regions (3'UTRs) and occasionally coding sequences. Alternative polyadenylation has been observed in all eukaryotes studied to date and is a very common process; for example, at least 70% of mammalian genes undergo alternative polyadenylation. Changes in alternative polyadenylation have been implicated in the regulation of subcellular protein and mRNA localization, translational regulation and mRNA stability (Tian and Manley, 2016).

Moreover, global changes in alternative polyadenylation have been observed during various physiological processes (Flavell et al., 2008; Mayr and Bartel, 2009; Sandberg et al., 2008; Shepard et al., 2011). In addition, different tissues and cell types have been shown to differ in their expression of 3'UTR isoforms (Boutet et al., 2012; L. Wang et al., 2013a).

Interestingly, evidence has shown that mammalian and fly brains express particularly long 3'UTR isoforms compared to other tissues (Haibo Zhang et al., 2005; Miura et al., 2013; Smibert et al., 2012). Neurons are highly diverse in their morphology and physiology and dynamically regulate their gene expression in response to physiological activity, all features that might call for posttranscriptional gene regulation.

However, the diversity of alternative polyadenylation within different cell types in the brain has not yet been explored. Studying the mRNA content of specific neuronal types presents a significant challenge given that single neuronal types are difficult to purify. To address this issue, the Darnell lab has developed a new technique termed conditionally-tagged poly(A) binding protein-mediated mRNA 3' end retrieval by crosslinking immunoprecipitation (cTag-PAPERCLIP) to look at mRNA expression and 3'UTR ends in specific cell types (HW Hwang et. al., in review). The technique is based on High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP), also developed by the Darnell lab, and is discussed in detail at the end of this Chapter.

In this study, we use the new technique to address whether different cell types in the cerebellum express different 3'UTR isoforms. We found large differences in 3'UTR isoform expression between Purkinje and granule cells, two main cerebellar cell types. Additionally, we looked at changes in alternative polyadenylation during cerebellar granule cell development when a cell transitions from a mitotic precursor to a post-mitotic, synaptically wired adult neuron. In these neurons, we found that transcripts that undergo changes in alternative polyadenylation are enriched for genes involved in cell proliferation and in determining cell's morphology. We address the function of one such change, in alternative polyadenylation of the *Memo1* transcript, and

demonstrate that a miRNA binding site specific to the long isoform plays a role in *Memo1* downregulation.

I will first review the current state of research on alternative polyadenylation, followed by an introduction to the cerebellum and its different cell types. Finally, I will discuss the current state of cell type-specific transcriptome profiling techniques as well as cTag-PAPERCLIP, the new technique for cell-type specific transcriptome profiling and mapping of 3'UTR ends developed by the Darnell lab.

mRNA cleavage and polyadenylation

In the early 1970s, it was discovered that mRNA labeled with radioactive adenosine contains an RNase resistant segment of labeled adenosine-rich RNA (Darnell, 2011). Later, it was found that most eukaryotic RNA transcripts transcribed by RNA polymerase II (Pol II) are polyadenylated, except for replication-dependent histone transcripts in metazoans. The genes transcribed by Pol II include protein-coding genes as well as some noncoding RNAs, including spliceosomal small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), microRNA (miRNA) precursors, and cryptic unstable transcripts (CUTs) (Richard and Manley, 2009). Other Pol II transcripts are also polyadenylated, like long non-coding RNAs (lncRNAs; non-coding transcripts of ~200 nucleotides or longer) (Tian and Manley, 2016).

The poly(A) tail of an mRNA transcript is, along with splicing, required for transporting an mRNA message from the nucleus to the cytoplasm (Huang

and Carmichael, 1996). In the cytoplasm, poly(A) tail stabilizes mRNAs by preventing mRNAs from being degraded (Ford et al., 1997) and plays a role in translation (Subtelny et al., 2014).

mRNA transcripts acquire their poly(A) tail through a two-step reaction: first, the transcript is cleaved at a specific site, and then it is polyadenylated. The site at which an RNA transcript is cleaved and polyadenylated is defined by the surrounding RNA sequence. In mammals, transcripts are cleaved and polyadenylated 10–30 nucleotides downstream from a conserved hexanucleotide – most commonly AAUAAA or AUUAAA – and 30 nucleotides upstream of a less conserved U- or GU-rich region. Additional elements upstream of the hexamer sequence described above (typically AAUAAA or AUUAAA) have been identified, namely UGUA and U-rich elements (Tian and Manley, 2016). In addition, the cleavage site is often located immediately downstream from a CA sequence. The sequences next to the polyadenylation site are divergent, and many polyadenylation sites lack one or more of the above-mentioned *cis* elements (Shi and Manley, 2015).

mRNA cleavage and polyadenylation can be reconstituted *in vitro*. The following six components were found to be essential for this *in vitro* reaction: Cleavage and polyadenylation specificity factor (CPSF), Cleavage stimulation factor (CstF), Cleavage factor I (CFI), Cleavage factor II (CFII), Poly(A) polymerase (PAP) and Poly(A)-binding protein nuclear 1 (PABPN1) (Wahle and Rüegsegger, 1999). The first four components are multi-protein

complexes and the molecular weight of the entire core machinery exceeds 1 MDa (Shi et al., 2009). The AAUAAA signal is recognized by two subunits of the CPSF complex: CPSF30 and WDR33 (Chan et al., 2014; Schönemann et al., 2014). The U/GU-rich element downstream of the cleavage site is recognized by the CstF-64 subunit within the CstF complex (C. C. MacDonald et al., 1994). The UGUA motif upstream of AAUAAA hexamer is recognized by the CFI complex (Brown and Gilmartin, 2003; G. Martin et al., 2012a). The CFI and CstF complex are both essential for cleavage, but not polyadenylation. The endonucleolytic cleavage that precedes polyadenylation is carried out by the CPSF-73 subunit of the CPSF complex (Mandel et al., 2006). The fourth complex required for cleavage is CFII (de Vries et al., 2000). After cleavage, the poly(A) tail is synthesized by the PAP enzyme. By itself, PAP is almost inactive, but its activity is enhanced by the CPSF complex and PABPN1 (Kerwitz et al., 2003). Additional components that contribute to polyadenylation have been identified, for example symplekin and C-terminal domain of Pol II, which are believed to function as scaffolds (Tian and Manley, 2016).

A functional polyadenylation signal is required for Pol II to terminate transcription, whereas a cleavage signal is not required (Connelly and Manley, 1988). *In vitro* transcription experiments have shown that the cleavage and polyadenylation machinery can pause Pol II as it is transcribing the polyadenylation signal, which may result in the gradual release of Pol II

from the DNA template (Huimin Zhang et al., 2015). In addition to this Pol II pause, the nuclear 5′-3′ exonuclease Xrn2 begins to degrade the leftover transcript downstream of the cleavage site (Huimin Zhang et al., 2015). Xrn2 degrades the downstream transcript from the 5' end, and eventually catches up with Pol II, facilitating the release of Pol II from the DNA template (Proudfoot, 2016; West et al., 2004).

Alternative polyadenylation

Alternative polyadenylation is a process where different ends to an mRNA transcript are determined by cleavage and polyadenylation at alternative sites. There are several types of alternative polyadenylation (Figure 1-1). The simplest type, called *3'UTR-alternative polyadenylation* (also 'tandem' 3'UTR-alternative polyadenylation), involves multiple poly(A) sites within the same terminal exon. This type generates multiple isoforms that only differ in their 3'UTR length, without affecting the coding sequence. In addition, other types of alternative polyadenylation affect the coding sequence of the protein and thus produce proteins with different functions. I will refer to these types as *upstream alternative polyadenylation*. There are several types of upstream alternative polyadenylation, depending on the configuration of splicing relative to the poly(A) site: alternative terminal exon alternative polyadenylation, in which alternative splicing generates isoforms that differ in their last exon; intronic alternative polyadenylation, which involves cleaving at the cryptic intronic poly(A) signal, extending an internal exon and making it the terminal

one; and internal exon alternative polyadenylation, which involves premature polyadenylation within the coding region. The last one is rare and the transcripts are likely to be degraded rapidly through the non-stop decay pathway (Elkon et al., 2013; Tian and Manley, 2016).

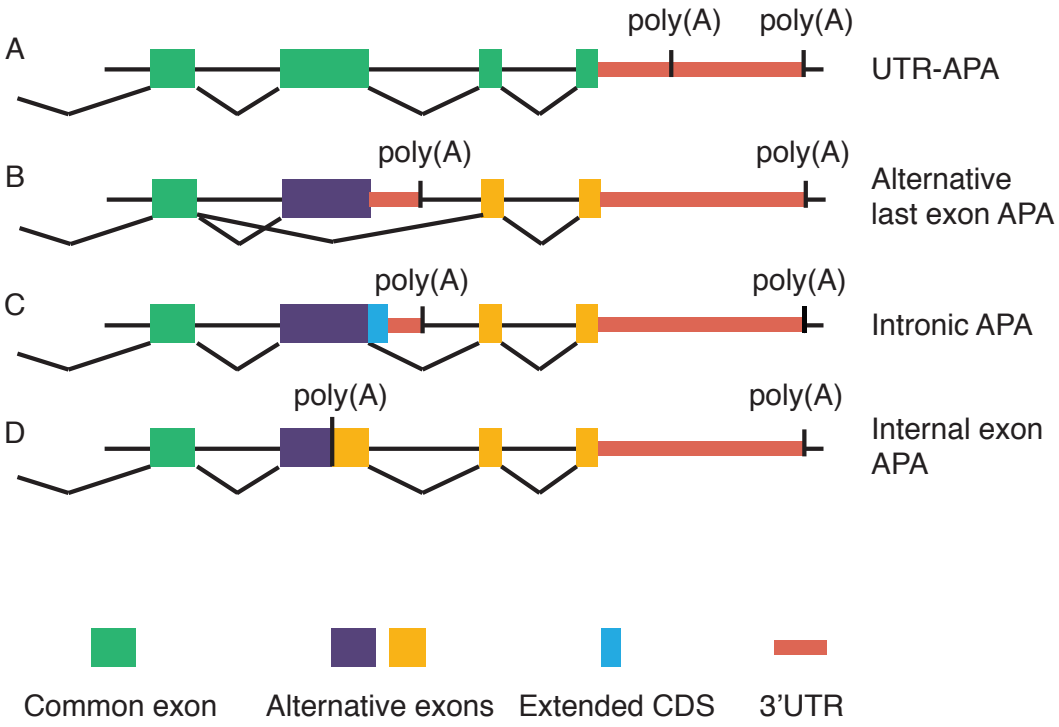


Figure 1-1 Types of alternative polyadenylation (APA)
(A) Most common type of alternative polyadenylation is 3'UTR-alternative polyadenylation, which involves alternative poly(A) sites within the same terminal exon. The 3'UTR isoforms generated by this type of alternative polyadenylation only differ in their 3′UTR length, the protein translated from the two isoforms has the same coding sequence. **(B)** Upstream alternative polyadenylation involves alternative terminal exon - alternative splicing generates isoforms that differ in their last exon. **(C)** Intronic alternative polyadenylation involves cleavage at the intronic poly(A) signal. The internal exon is extended and becomes the terminal one. **(D)** Internal exon alternative polyadenylation involves premature cleavage of a coding exon.

Alternative polyadenylation changes dynamically during various physiological processes. Global changes in alternative polyadenylation have been observed during neuronal activation (Flavell et al., 2008), T-cell activation (Sandberg et al., 2008), oncogenesis (Mayr and Bartel, 2009) and development (Ji et al., 2009; Shepard et al., 2011; L. Wang et al., 2013a).

Alternative polyadenylation patterns are also different across tissues, as assessed from expressed sequence tags (ESTs) (Beaudoing and Gautheret, 2001; Haibo Zhang et al., 2005) and high-throughput sequencing (Derti et al., 2012; Lianoglou et al., 2013). Moreover, some tissues show a global bias towards longer or shorter isoforms. In both the mammalian (Miura et al., 2013) and *Drosophila* brain (Smibert et al., 2012), 3'UTRs tend to be longer. Testes, on the other hand, tend to express very short 3'UTR isoforms, both in mouse (Liu et al., 2006) and *Drosophila* (Smibert et al., 2012).

A comparison of which poly(A) sites are used in five mammals showed that their usage is very similar in orthologous tissues, particularly in brain, kidney, and liver (Derti et al., 2012). However, another study that compared tandem poly(A) site usage in resting versus activated T-cells in human and mice did not find a conservation of poly(A) site usage, despite both species showing a trend towards using shorter 3'UTR isoforms in activated T-cells (Gruber et al., 2014).

Whereas tissue-restricted genes tend to express a single 3' UTR isoform, most ubiquitously-transcribed genes express multiple 3' UTR

isoforms. During transformation and differentiation, single 3' UTR genes tend to change their mRNA abundance levels, whereas multi 3' UTR genes mostly change 3' UTR isoform ratios to achieve tissue specificity (Lianoglou et al., 2013). This finding raises the possibility that alternative polyadenylation has an important role in the tissue-specific post-transcriptional regulation of these genes by 3'UTRs.

Consequences of alternative polyadenylation

Effects on mRNA stability

Most poly(A) sites are located in 3′ UTRs (Tian and Manley, 2016), which contain a variety of regulatory elements. For example, 3'UTR extensions have been shown to contain miRNA target sites (Boutet et al., 2012; Hwang et al., 2016; Miura et al., 2013; Nam et al., 2014). Several other motifs that are typically present in 3'UTRs can affect RNA stability, such as AU-rich elements (AREs), GU-rich elements (GREs) and PUF protein-binding elements (Tian and Manley, 2016). For example, a genetic polymorphism in the poly(A) site in human IFN-regulatory factor 5 (IRF5) is associated with a risk of developing systemic lupus erythematosus. This polymorphism leads to the expression of a long 3'UTR isoform instead of a short one. This extended 3'UTR in turn harbors an ARE element, making the long isoform less stable (Graham et al., 2007). However, it is unlikely that long isoforms are generally less stable than shorter ones. In a recent genome-wide study, long isoforms

were found to be only slightly less stable than short isoforms (measured by blocking transcription using actinomycin D) (Spies et al., 2013).

Effects on mRNA translation

The alternative use of different 3'UTRs can also influence mRNA translation – for example, as in the case of the *BDNF* transcript. The *BDNF* gene expresses two 3'UTR isoforms that have distinct neuronal functions. The long *BDNF* isoform is localized to dendrites and is only translated upon neuronal activity. The short isoform is localized to the cell body and constitutively translated. Mice that lack the long 3'UTR of *BDNF* exhibit altered dendritic spine morphology and decreased plasticity of dendritic synapses (An et al., 2008; Lau et al., 2010). Another example is the alternative polyadenylation of *polo* kinase in *Drosophila*. The long 3'UTR isoform of *polo* is expressed in abdominal epidermis precursor cells during development and is translated with much higher efficiency than the short 3'UTR isoform. Because high levels of Polo protein are required for the proliferation of epidermis precursor cells during development, deleting the *Polo* distal polyadenylation signal leads to death during development (Pinto et al., 2011).

However, whether any generalizations can be made regarding the effects of expressing long versus short 3'UTR isoforms remains unclear. One study in mouse cells reported that long isoforms had slightly greater translational efficiency than shorter isoforms (Spies et al., 2013). Conversely,

a study in human cells found just the opposite – longer 3'UTR isoforms were associated with significantly less polysomes (Floor and Doudna, 2016). It may be that the effects of expressing different 3'UTRs will be best understood by examining each isoform in each gene on a case-by-case basis, since each gene may harbor different sequence elements that could lead to a wide variety of effects.

Effects on mRNA localization

3'UTRs can also regulate the subcellular localization of mRNAs. For example, the long isoform of *BDNF* regulates its localization to dendritic processes (An et al., 2008). Several other reports of 3'UTRs regulating mRNA localization in neurons have also been described (Andreassi and Riccio, 2009; Harrison et al., 2013; Yudin et al., 2008) that suggested that the longer 3'UTR isoform (produced by 3'UTR alternative polyadenylation, see Figure 1-1) is localized to neurites. However, a genome-wide study showed that there is no systematic trend: long 3'UTR isoforms are not on average more likely to be localized to neuronal processes. Interestingly, this genome-wide study did observe that the distal alternative last exon isoforms (produced by upstream alternative polyadenylation) were more likely to localize to neurites (Taliaferro et al., 2016).

Another question is whether there is a difference in the localization of 3'UTR isoforms between the cell nucleus and cytoplasm. Two studies found

that isoforms with longer 3'UTRs tend to be overrepresented in the nucleus (Djebali et al., 2012; Neve et al., 2016). Specifically, Neve et al. found that ~10% of detected alternative 3'UTRs were more abundant in either the nucleus or the cytoplasm.

Effects on protein localization

An intriguing recent study showed that 3'UTRs could act as recruiters for other proteins that might help in the subcellular localization of the protein translated from the mRNA (Berkovits and Mayr, 2015). Specifically, the alternative 3'UTR of the mRNA encoding the transmembrane protein CD47 was found to act as a scaffold for a protein complex containing the RNA-binding protein HuR. HuR recruits the phosphatase 2A inhibitor, SET. SET in turn interacts with the newly translated cytoplasmic domains of CD47 and mediates the translocation of CD47 to the plasma membrane. On the other hand, the short isoform of CD47 is primarily localized to the endoplasmic reticulum, because it lacks the RNA sequence to bind the HuR-SET complex. An interesting possibility is that 3'UTRs could, in addition to subcellular localization, determine alternative protein functions, for example through directing post-translational modifications on the nascent protein (Mayr, 2016).

Effects on protein function

As mentioned previously, alternative polyadenylation can also affect coding sequences, in which case it is called *upstream alternative polyadenylation*. The most studied example of upstream alternative polyadenylation is the transcript encoding the IgM heavy chain. During B cell activation, the IgM gene switches from using the distal poly(A) site in the 3′-most exon to using a proximal poly(A) site in the terminal exon. This switch in polyadenylation causes the cell to produce the secreted form of the antibody instead of its membrane-bound form (Alt et al., 1980). Moreover, a genome-wide analysis has identified at least 376 mouse genes that might also use such a mechanism for regulating membrane anchoring (Davis et al., 2006).

Upstream alternative polyadenylation sometimes leads to the expression of truncated, and likely non-functional (dominant negative) proteins. For example, the expression of the mammalian polyadenylation factor CSTF77 (cleavage stimulation factor 77 kDa subunit) is controlled through a negative feedback loop that involves an intronic poly(A) site, which, when used, truncates the protein such that is non-functional (Pan et al., 2006). Because the use of this poly(A) site is regulated by CSFT77 itself, higher levels of CSTF77 feedback to produce more of the truncated protein, thus lowering the levels of the functional protein (Luo et al., 2013).

Regulation of alternative polyadenylation

One important mechanism for regulating alternative polyadenylation involves modulating the levels of core polyadenylation machinery components. For example, CSTF64 (the RNA-binding subunit of the CSTF complex) regulates the alternative polyadenylation of the IgM heavy chain transcript (described above). CSTF64 is repressed in mouse primary B cells and upregulated upon their activation. Higher levels of CSTF64 lead to increased usage of the proximal poly(A) site which leads to a switch in the expression of the IgM heavy chain from the membrane-bound isoform to the secreted isoform (Takagaki et al., 1996). This finding was recently corroborated using high-throughput sequencing (Hwang et al., 2016; Xia et al., 2014; Yao et al., 2012).

Interestingly, during somatic cell reprogramming, 3'UTRs tend to shorten, whereas the expression of mRNAs encoding for cleavage and polyadenylation factors is upregulated (Ji and Tian, 2009). The same inverse correlation between 3'UTR length and cleavage/polyadenylation complex levels was observed when comparing human fibroblasts during high and low levels of proliferation (low levels of proliferation were due to the confluence of cells in the dish). The proliferating cells also had shorter 3'UTRs and higher levels of mRNAs encoding for cleavage and polyadenylation factors (Elkon et al., 2012). These examples suggest that changes in the levels of cleavage

and polyadenylation factors could influence 3'UTR length across many genes during various physiological processes.

How might elevated levels of core polyadenylation factors lead to shorter 3'UTRs? One model is that proximal polyadenylation sites are less favored than downstream sites, and that these proximal sites are used when the availability of the core polyadenylation machinery is not limiting (Tian and Manley, 2016). Surprisingly, however, knocking down certain components of the CFI complex causes 3'UTRs to shorten (Li et al., 2015; G. Martin et al., 2012b), suggesting that different components of the cleavage/polyadenylation complex may contribute in different ways to the overall levels of long vs. short 3'UTRs.

Another protein with a role in cleavage and polyadenylation is PABPN1. PABPN1 controls the length of the poly(A) tail, and may also regulate which poly(A) site is chosen, given that knocking down PABPN1 leads to shorter 3'UTRs (de Klerk et al., 2012; Jenal et al., 2012). Moreover, a disease called autosomal-dominant oculopharyngeal muscular dystrophy (OPMD) is caused by a trinucleotide-repeat expansion in the PABPN1 gene. Expressing the mutant PABPN1 in human cells causes extensive 3'UTR shortening as well (Jenal et al., 2012). Curiously, the knockdown of its (mostly) cytoplasmic counterpart, PABPC1, was also found to modulate 3′ UTR length . (Li et al., 2015).

In addition to these changes in the levels of cleavage and polyadenylation factors, several RNA-binding proteins have also been shown to play a role in determining 3'UTR usage: Nova (Licatalosi et al., 2008), Elav (Hilgers et al., 2012; Oktaba et al., 2015), SR proteins (Müller-McNicoll et al., 2016) and Fus (Masuda et al., 2015). Elav is particularly interesting, given that it has been shown to inhibit proximal poly(A) usage during the development of the *Drosophila* nervous system and that it may be at least partially responsible for the exceptionally long 3'UTR isoforms expressed in the brain (Hilgers et al., 2012; Oktaba et al., 2015). Interestingly, a recent study from the Darnell lab showed that the m6A mRNA modification regulates alternative polyadenylation. m6A is enriched in the last exons in a gene. Furthermore, reducing m6A levels changed which 3'UTRs are expressed, and tended to cause 3′ UTRs to shorten (Ke et al., 2015).

It is also likely that polyadenylation is in competition with splicing in the case of intronic alternative polyadenylation. One line of evidence supporting this hypothesis is that intronic poly(A) sites are associated with weak 5' splice sites and large intron sizes. This presumably increases the time to splice out an intron and creates a time window for cleavage and polyadenylation of the intron (Tian et al., 2007).

U1 SNRP also plays a role in regulating alternative polyadenylation. U1 snRNP is a spliceosome component that plays an essential role in defining the 5' splice site by base pairing via U1 snRNAs. Interestingly,

knocking down U1 snRNA results not only in less splicing but also in the premature termination of most mRNAs from intronic cryptic poly(A) signals that are typically located close to the transcription start site. This suggests that U1 snRNP represses the use of cryptic intronic poly(A) sites (Berg et al., 2012).

Finally, the rate of Pol II elongation can control which poly(A) site is used. To study this possibility in *Drosophila*, Pinto et al. used a fly harboring a mutation in Pol II that slows down its rate of elongation. These mutant flies differed in alternative polyadenylation of *polo* compared to wildtype flies (Pinto et al., 2011). This finding is consistent with earlier *in vitro* evidence that sequences that cause Pol II pausing facilitate usage of an upstream poly(A) site (Yonaha and Proudfoot, 1999). Factors associated with Pol II can also regulate polyadenylation. For example, the PAF complex, a transcription elongation factor, interacts with the CPSF and CSTF cleavage/polyadenylation complexes (Rozenblatt-Rosen et al., 2009). Knocking down CDC73 or another PAF subunit, PAF1 in mouse cells leads to an increase in the use of proximal intronic poly(A) sites, which could be due to a change in Pol II kinetics or a more active role for these factors, given that CDC73 directly interacts with the cleavage and polyadenylation machinery (Yang et al., 2016).

The function and anatomy of the cerebellum

The cerebellum is a structure in the vertebrate brain that is involved in sensory-motor processing (Buckner, 2013). Animals and humans with cerebellar dysfunction have problems with motor control (e.g. inability to stand and walk, lack of motor coordination, loss of automatic, unconscious nature of most movements) (Kandel et. al., 2012). More recent studies have also uncovered a role for the cerebellum in cognition. For example, patients with cerebellar damage are unable to learn certain word-association tasks, suggesting that the cerebellum is also responsible for cognitive computations (Buckner, 2013). In addition, the cerebellum has emerged as one of the key brain regions affected in autism (Wang et al., 2013).

We chose to study alternative polyadenylation in the cerebellum because it is composed of well described cell types that are genetically accessible through Cre-diver lines (Barski et al., 2000; Matei et al., 2005). The cerebellum consists of a cortex that surrounds white matter (axon fibers) and several deep nuclei embedded in the white matter. The cortex has a relatively simple structure that is composed of two major types of neurons: granule and Purkinje cells. Granule cells are one of the smallest cell types in the brain and emit only four to five dendritic branches. These cells use glutamate as a neurotransmitter and are excitatory. Purkinje cells have a large cell body and have elaborate dendritic trees. These cells use gamma-aminobutyric acid (GABA) as their neurotransmitter and are inhibitory.

Granule cell bodies form a dense layer below a monolayer of Purkinje cell bodies. A more superficial molecular layer contains the granule cell axons and the elaborate Purkinje cell dendritic trees (Figure 1-2). Granule cell axons project from the granule layer to the molecular layer, where they bifurcate and run along an array of Purkinje cell dendritic trees; because granule cell axons run along the molecular layer of the cerebellar cortex, they are called parallel fibers (Figure 1-2). Granule cells receive inputs from outside the cerebellum and synapse onto Purkinje cells through these parallel fibers. The majority of Purkinje cells, receiving inputs from granule cells, then project to the cerebellar nuclei embedded in the white matter (Butts et al., 2014). The cerebellar circuit also contains several other cell types that are important for its physiological function – however, only granule and Purkinje cells are considered in this thesis for the purpose of comparing alternative polyadenylation in two well defined cell types.

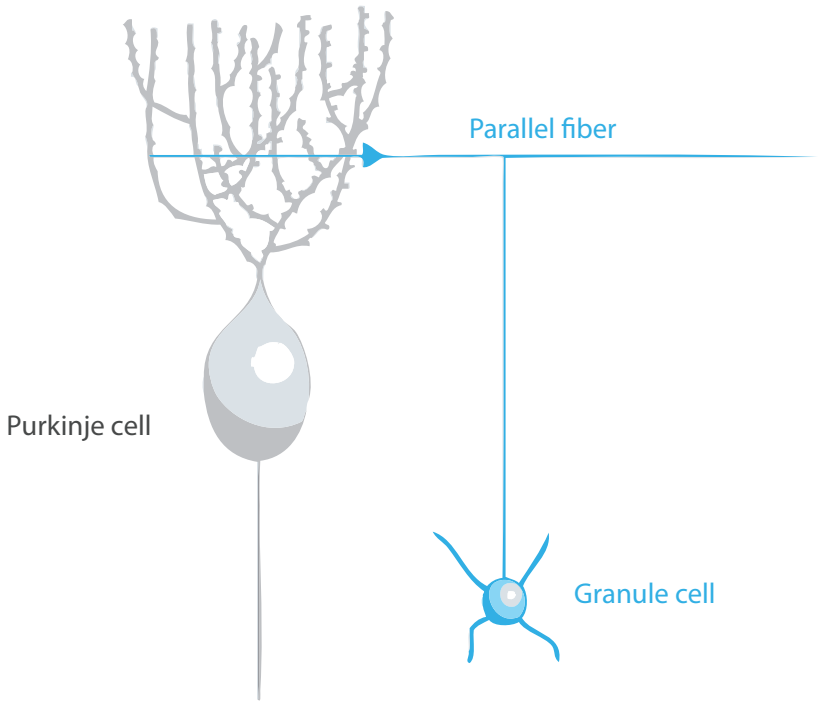


Figure 1-2 Granule and Purkinje cell circuit

Schematic of cerebellar granule cell synaptic inputs onto Purkinje cells. Modified from Tsai et. al., 2012.

Cerebellum development

This thesis examines the changes in alternative polyadenylation during granule cell development. Granule neurons are derived from granule neuron precursors that arise from an embryonic structure called the rhombic lip. In mice, rhombic lip-derived neuronal precursors migrate across the dorsal surface of the embryonic region destined to become the cerebellum (also called 'anlagen') between embryonic day 12.5 and 16, to form the external granular layer. After birth, this thin external granule layer expands to a layer that is six to eight cells deep. Granule cell precursors continue to proliferate until postnatal day 15 (Hatten in Leto et al., 2015). This proliferation is mainly driven by Sonic hedgehog (SHH) (Altaba, 1999; Wallace, 1999), which is produced by neighboring Purkinje cells. Other pathways, such as Notch signaling (Solecki et al., 2001), also contribute to the proliferation of granule cell precursors. Granule cell precursors express the transcription factor Atoh1 (also called Math1) (Machold and Fishell, 2005; V. Y. Wang et al., 2005), which is required for their specification and proliferation (Ben-Arie et al., 1997).

Following this period of proliferation in the external granule layer, granule cell precursors exit the cell cycle, downregulating Math1 and upregulating a transcription factor called NeuroD1, which is required for their differentiation (Miyata et al., 1999). After granule cells exit the cell cycle, granule cells migrate inwards, by binding to fibers of Bergmann glia, to form

the internal granule layer that is present in the fully developed, adult cerebellum. This entire process of granule cell precursor proliferation, differentiation, and inward migration starts at birth and lasts until 20 days after birth. After reaching the nascent granular layer, granule cell T-shaped axons, the parallel fibers, form synaptic connections with the extending Purkinje cell dendrites. In addition, granule cells extend dendrites that form synapses with ingrowing afferent axons, the mossy fibers (Hatten and Heintz, 1995; Leto et al., 2015).

Granule cell precursors are also known to be the cell population from which medulloblastoma, a childhood cancer, originates (Hatten and Roussel, 2011).

Cell type specific transcriptome profiling

The mammalian brain consists of many types of neurons with distinct physiological properties and morphologies. This complexity presents a serious challenge for mechanistic studies of brain development, function, and disease. Given that these differences between neuronal types must stem from different gene expression profiles (Hobert et al., 2010), an important goal in neuroscience is to describe the complement of mRNA molecules expressed in each cell type.

Several methods have been developed to study the genome-wide expression profiles of cells in complex tissues. These methods include FACS of dissociated neurons, manual sorting of fluorescent cells, immunopanning of dissociated cells, laser-capture microdissection and aspiration of patched cells after electrophysiological recordings. The main disadvantage of these methods is the introduction of experimental noise caused by tissue fixation, tissue dissociation, tissue incubation *ex vivo* or high amplification of mRNA from single-cells. In addition, these methods tend to measure only the mRNA present in the cell soma, while excluding mRNA present in dendritic and axonal processes (Kulicke et al., 2014).

Recently, several methods have been developed that enable us to purify RNA from the entirety of a cell, while preserving cell-type specificity. Two of these methods purify ribosome-associated RNA from specific cell types. In the first method, called Translating ribosome affinity purification

(TRAP), the L10a ribosomal protein tagged with EGFP is expressed in specific cell populations. To achieve cell type specific expression, the authors made transgenic mice that harbor bacterial artificial chromosome (BAC) that expresses EGFP-L10a under the control of various loci with cell type restricted expression. Tagged polysomes (multiple ribosomes bound to a single mRNA) are then immunoprecipitated from whole brains using anti-EGFP antibodies (Doyle et al., 2008; Heiman et al., 2008). Another technique uses the more popular Cre/lox recombination instead of BAC transgenic mice, which avoids concerns about BAC transgenes (e.g. the random insertion of BACs in the genome as well as the additional genes carried on BACs may produce a phenotype). The genetically modified mouse, called RiboTag mouse, carries an *Rpl22* allele with a floxed wild-type C-terminal exon followed by an identical C-terminal exon that has three hemagglutinin (HA) epitopes inserted before the stop codon. When the RiboTag mouse is crossed to a mouse expressing Cre recombinase in a cell-type specific manner, the Cre recombinase activates the expression of the epitope-tagged ribosomal protein RPL22, which is then incorporated into actively translating polysomes. Tagged polysomes, along with their associated mRNA, are then immunoprecipitated with a monoclonal antibody against HA (Sanz et al., 2009).

To identify miRNAs and their mRNA targets expressed in specific neurons, one can immunoprecipitate epitope-tagged Argounate 2 (AGO2)

that is expressed in specific cells using Cre/lox recombination. One study to apply this strategy used HITS-CLIP (described below) to purify epitope-tagged AGO2-associated mRNA expressed in specific neurons (Tan et al., 2013). A second study immunoprecipitated epitope-tagged AGO2 (bound to miRNAs) without crosslinking (He et al., 2012).

HITS-CLIP, PAPERCLIP and cTag-PAPERCLIP

In this thesis, I describe several applications of a novel method for cell-type specific transcriptome profiling termed cTag-PAPERCLIP. The approach enables mapping of the 3' ends of polyadenylated RNA expressed in specific cell types by expressing GFP-tagged Poly(A) binding protein cytoplasmic 1 (PABPC1) in selected cell populations using the Cre/lox system and subsequently purifying PABPC1-bound RNA using an anti-GFP antibody (HW Hwang et. al., in review).

cTag-PAPERCLIP is based on HITS-CLIP, a method developed by the Darnell lab that allows stringent immunopurification of RNA targets of RNA-binding proteins to generate precise genome-wide interaction maps (Licatalosi et al., 2008). The central feature of the HITS-CLIP protocol, which effectively eliminates non-specific interactions, is the formation of covalent cross-links between protein and a directly bound (within ~1 Å) RNA by UV irradiation of intact cells or tissues. In this way, endogenous protein-RNA interactions are 'frozen' *in vivo* (Moore et al., 2014), which eliminates re-association of protein–RNA complexes *in vitro* after cell lysis (Mili and Steitz,

2004) and enables stringent washes after immunoprecipitation. Following cross-linking and cell lysis, RNA is partially hydrolyzed to reduce bound RNA fragments to 'footprint' sizes (typically about 60 nucleotides long) that can then be reverse transcribed and PCR amplified. PCR products are then sequenced and mapped to the genome, which yields transcriptome-wide RNA-protein binding maps (Moore et al., 2014).

cTag-PAPERCLIP is a variation of a recently published method, termed PAPERCLIP, which is using an anti-PABPC1 antibody to purify PABPC1-bound mRNA (Hwang et al., 2016) to map 3'UTR ends. The PAPERCLIP approach was shown to be superior to other methods used to map 3'UTR ends, because it minimizes internal priming during reverse transcription, a common problem present in other methods for 3' end mapping (Shi, 2012). The reason for reduced internal priming in PAPERCLIP is that only mRNA fragments containing the poly(A) tail crosslinked to PABPC1 are retained after immunoprecipitation with anti-PABPC1 antibodies (Hwang et al., 2016).

cTag-PAPERCLIP protocol is similar to the PAPERCLIP protocol, except that anti-GFP antibodies are used instead of anti-PABPC1 antibody. The protocol and the mouse model are described in more detail in Chapter 3.

cTag-PAPERCLIP is comparable to TRAP, RiboTag and epitope-tagged AGO2 approaches, because it does not require the brain to be fixed or dissociated, or that cell bodies be extracted, to capture cell type–specific mRNA. However, unlike TRAP and RiboTag, cTag-PAPERCLIP does not

isolate only ribosome-bound mRNAs, but presumably isolates all polyadenylated RNAs present in a cell. Furthermore, cTag-PAPERCLIP, as I describe in Chapter 3, allows one to precisely map 3'UTR ends with exceptional selectivity.

Chapter 2 I Experimental procedures

Mice

Animals were maintained in an AAALAC-approved animal facility and all procedures were performed in accordance with IACUC guidelines. The *Math1-Cre* and *NeuroD1-Cre* mice were a gift from Mary E. Hatten. *Pcp2/L7-Cre* mice were obtained from the Jackson labs (catalog number: 004146).

Immunohistochemistry

Mouse brains were perfused with PBS and 4% paraformaldehyde (PFA) and fixed in 4% PFA/PBS at 4°C overnight. Fixed cerebella were then incubated in 15% sucrose/PBS followed by 30% sucrose/PBS at 4°C and finally embedded in OCT compound. Frozen brains were sliced into 30 µm thick sections on a cryostat (CM3050S, LEICA). Brain sections were then subjected to immunohistochemistry. The sections were first washed three times with PBS at room temperature (RT), incubated with 0.2% Triton X-100/PBS at RT, blocked with 1.5% normal donkey serum (NDS)/PBS at RT, and then incubated overnight at 4°C with primary antibodies in 1.5% NDS/PBS followed by incubation with Alexa 488 and 555 conjugated donkey secondary antibodies (1:1000) in 1.5% NDS/PBS. Images of immunostained specimens were taken using BZ-X700 (KEYENCE) microscope. We used the following primary antibodies for immunohistochemistry: anti-Calb1 (Santa

Cruz, catalog number: sc-7691, 1:250 dilution) anti-GFP (AVES, catalog number: GFP-1020, 1:1000 dilution).

cTag-PAPERCLIP

Bead preparation for immunoprecipitation

Protein G beads were resuspended and washed three times in Antibody binding buffer (1x PBS, 0.02% Tween-20, pH 7.4). 300 uL of beads were used per group. A group is defined as one biological replicate, or a control, e.g. overdigested control or Cre- control. Beads were then pre-coated with mouse monoclonal antibodies against GFP (19F7 and 19C8) (Kulicke et al., 2014). 20 ug of each antibody was added to 1200 mL of Antibody binding buffer per group. The solution was then added to the beads and rotated at room temperature for 30 minutes.

Brain lysate preparation

For cTag-PAPERCLIP of adult granule and Purkinje cells (cTag-PABPC1 driven by *NeuroD1* and *Pcp2* promoters) cerebella were dissected in ice cold PBS. A cell suspension was made by passing tissue through a syringe with a 18G needle 8 times (on ice). The cell suspension was then UV crosslinked in a Petri dish on a bed of ice three times at 200 mJ/cm2 (254 nm UV light) using a Stratalinker 2400 (Stratagene). For cTag-PAPERCLIP experiments on developing granule cells (cTag-PABPC1 driven by *Math1* promoter) cerebella were dissected in ice cold PBS. A tissue chopper (McIlwain) was used to make 300 um thick saggital sections of the cerebellum. Sections were

transferred to ice-cold PBS and immediately crosslinked as described above. After crosslinking, deep cerebellar nuclei were micro-dissected using a fluorescent microscope in ice-cold PBS. After dissection, pieces of cerebellar cortex ware transferred to an Eppendorf tube on ice filled with PBS and centrifuged to remove PBS. Crosslinked cell suspension and tissue sections were centrifuged to remove buffer and frozen at -80°C. When enough mice were collected, cTag-PAPERCLIP experiments were performed. Cerebella were thawed by adding 5X weight of PXL lysis buffer (1X PBS [tissue culture grade; no Mg2+, no Ca2+, GIBCO], 0.1% SDS, 0.5% Na-DOC, 0.5% NP-40 with 1 tablet/10 mL cOmplete protease inhibitor cocktail (Roche)) and resuspended by pipetting ten times. Cerebellar lysates were then pooled and incubated on ice for 8 minutes. The lysates were DNase treated with RQ1 RNase-free DNase (Promega, 25 uL per group) for 5 minutes shaking in the thermomixer at 37°C. RNase A (Affymetrix) solution was prepared (1:100 (high – for overdigested control) or 1:100,000 (low – for all other samples)) and added to each sample (10 uL dilution per 1 mL lysate for 1:100,000 dilution and 20 uL per 1 mL lysate for 1:100 dilution) and the lysate was incubated in the thermomixer for 5 minutes shaking at 37°C. The lysate was spun at 14,000 rpm for 10 minutes at 4°Cto clear the lysate and the supernatant was harvested to a new tube for immunoprecipitation. Samples were then rotated with primary antibody-loaded beads (that were previously washed with 1 X PXL 3 times) at 4°C for 2 hours. Following

immunoprecipitation, the beads were washed sequentially (1 mL washes) twice with 1X PXL lysis buffer, twice with 5X PXL (5X PBS [tissue culture grade; no Mg2+, no Ca2+, GIBCO], 0.1% SDS, 0.5% Na-DOC, 0.5% NP-40), twice with High salt buffer (15 mM Tris-HCl, pH 7.5, 5 mM EDTA, 2.5 mM EGTA, 1% Triton X- 100, 1% Na-DOC, 0.1% SDS, 1 M NaCl), twice with Low salt buffer (15 mM Tris- HCl, pH 7.5, 5 mM EDTA) twice with High stringency buffer (15 mM Tris- HCl, pH 7.5, 5 mM EDTA, 2.5 mM EGTA, 1% Triton X-100, 1% Na-deoxycholate (DOC), 0.1% SDS, 120 mM NaCl, 25 mM KCl), and two washes with PNK buffer (50mM Tris-HCl, pH 7.5, 10 mM MgCl2, 0.5% NP-40). After each capture on the magnet, the beads were resuspended by end-over-end rotation. Throughout the CLIP protocol nonstick, RNAse free microcentrifuge tubes (Ambion) were used to reduce RNA binding to tubes.

Dephosphorylation and labeling of RNA fragments

Immunoprecipitated RNA fragments were treated with Fast AP alkaline phosphatase (Thermo Scientific). The samples were resuspended in 60 µL of 1X dephosphorylation buffer containing 2.3 uL of Fast AP and incubated in the thermomixer at 37°C for 20 min, programmed to shake at 1000 rpm for 15 seconds every 2 min. This was followed 1 mL washes, once with PNK Buffer, twice with PNK buffer with EGTA (10 mM Tris pH 7.5, 20 mM EGTA, 0.5% NP-40) and twice again with with PNK buffer (no EGTA). Following dephosphorylation, the RNA fragments were labelled with radioactive

phosphate using 32P-gamma-ATP. The beads were resuspended in the following mix: 6 uL 10X T4 PNK buffer (NEB), 3 uL T4 PNK enzyme (NEB), 1.5 uL 32P-gamma-ATP (Perkin Elmer), 49.5 uL nuclease free waster, total volume 60 uL). The beads were then incubated at 37°C in the thermomixer for 20 min at 1000 rpm shaking for 15 seconds every 4 min. Beads were then washed three times with 1 mL PNK buffer and stored at 4°C overnight in 13.5 uL PNK buffer.

SDS-PAGE separation of protein-RNA complexes

After last PNK buffer wash, each tube of beads was resuspended in 30 ul NuPAGE loading buffer (LDS, Invitrogen) brought to 1x concentration with PNK buffer and containing reducing agent for Ago CLIP. The beads were incubated in the thermomixer at 70°C for 10 min shaking at 1000 rpm. Supernatants were taken off the beads and run on Novex NuPAGE 8% Bis-Tris gels (Invitrogen) in MOPS running buffer (Invitrogen) for 2.5 hours at 200V. Rainbow protein marker (Full-range, GE Healthcare) was used and electrophoresis was stopped when the purple (52 Kd) marker reached the end of the gel. Protein-RNA complexes were then transferred to Protran BA85 nitrocellulose membrane (Whatman) using a Novex wet transfer apparatus (Invitrogen). After transfer, the nitrocellulose was quickly rinsed with RNase-free PBS, blotted with Kimwipes, wrapped in plastic wrap and exposed to Biomax MR film (Kodak). The film was exposed for around 4 hours at -80°C, depending on the strength of the signal.

Purification of RNA fragments from nitrocellulose membrane

Nitrocellulose membranes were aligned with the exposed film. A piece of membrane was excised from the low RNase immunoprecipitation lanes at a size directly above the overdigested band with a scalpel. Each band of nitrocellulose membrane was further cut into smaller pieces and proteinase K treated (0.2 mL of a 1:5 dilution of proteinase K (4 mg/mL, Roche) in PK buffer (100 mM Tris-Cl pH 7.5, 50 mM NaCl, 10 mM EDTA)) in the thermomixer at 37°C, shaking at 1100 rpm for 20 minutes. Then 0.2 mL of PK buffer with 7M urea solution was added and incubated for another 20 minutes at 37°C shaking at 1100 rpm. Finally, 0.4 mL of RNA phenol (pH 6.8, Applied Biosystems/Ambion) and 0.13 mL of 49:1 CHCl3 - isoamyl alcohol were added and incubated at 37°C, shaking at 1100 rpm for additional 20 min. Tubes were spun at 20,000g in a desktop microcentrifuge at room temperature. Glycogen (0.8 ul, Applied Biosystems/Ambion), 50 uL 3M NaOAc pH 5.2, and 1 mL of 1:1 ethanol isopropanol mixture was added to the aqueous phase in a fresh tube and RNA was precipitated overnight at -20°C.

Reverse transcription and PCR amplification

Protein G beads were washed 3 times with Antibody binding buffer and blocked in 5X Denhardt's Solution (Thermo Fisher) for 45 minutes at room temperature. After that, a solution of anti-BrdU antibody (per group: 40 uL Antibody binding buffer, 5 uL 50X Denhardt's solution (Thermo Fisher) and 5 uL anti-BrdU antibody (Abcam) was added to the beads and the beads were

rotated at room temperature for another 45 minutes. Then the beads were washed with 1X IP buffer (0.3X SSPE, 1mM EDTA, 0.05% Tween-20) and stored at 4°C before use. In the meantime, RNA was spun down at 14000 rpm for 20 min at 4°C, washed 2X with 1 mL cold 75% EtOH, air dried and resuspended in 9 uL of water (Ambion) by tapping. The RNA fragments were denatured at 65°C for 5 minutes in a thermomixer (without mixing) and reverse transcribed using Superscript III reverse transcriptase (Thermo Fisher) following manufacturer's instructions with some modifications. 8 uL of the following mix was added to 8 uL of RNA fragments: 1 uL of reverse transcription primer (sequence described below), 4 uL of 5X RT buffer, 3 uL of dATP, dCTP, dGTP mix (8.2 Mm each), 1 uL of Br-dUTP (8.2 Mm). The mixture was incubated at 75°C for 3 minutes in a PCR machine (BioRad) and then held at 48°C while the second mixture (pre-warmed to 48°C) was added to it: 1 uL of 0.1 M DTT, 1 uL of RNAsin and 1 uL of Superscript III enzyme (water for no RT control). The samples were incubated for 45 min at 48°C, 15 min at 55°C, 5 min at 85°C and held at 4°C. Following the RT reaction, 1 uL of RNAse H was added to each PCR tube and the samples were incubated for 20 minutes at 37°C. Finally, 10 uL of water was added per reaction and the mixture was spun through a G-25 column to remove free Br-dUTP as per manufacturer's instructions. The following was added to the tubes after collection of the samples: 10 uL of 50X Denhardt's and 50 uL of 2X IP buffer (0.6X SSPE, 2mM EDTA, 0.1% Tween-20). The samples were then heated to

70°C for 5 minutes, equilibrated at room temperature and mixed with 25 uL of anti-BrdU coated beads. The rest of the beads were stored at 4°C. The samples were then rotated for 45 minutes at RT and washed one time with 1x IP Buffer (with 5X Denhardt's), twice with Low Salt (with 1x Denhardt's), twice with High stringency buffer (with 1x Denhardt's) and twice with 1x IP Buffer (all washes were for 3 minutes on a rotator at RT). The samples were then eluted from beads by adding Elution buffer (50 uL 2X IP buffer and 40 uL nuclease-free water per sample) followed by a 1 minute incubation at 98°C in thermomixer (1200 rpm), spun down and placed on a magnet. The eluate was collected and 10 uL of 50X Denhardt's was added to each tube for a total of 100 uL. The samples were then stored at 4°C overnight. The next day, the samples were incubated for 5 minutes at 70°C and equilibrated at RT for a minute. 25 uL anti-BrdU beads were added per tube. The beads were rotated for 45 minutes at RT. After the incubation, the beads were washed once with 1 mL 1x IP Buffer (with 5X Denhardt's), twice with Low salt buffer (with 1x Denhardt's), twice with High stringency buffer (with 1x Denhardt's) and twice with CircLigase Wash Buffer (33mM Tris-Acetate, 66mM KCl (pH 7.8)). After the last wash, CircLigase reaction mix was added to the samples (per reaction: 2 uL CircLigase buffer 10X, 4 uL 5M Betaine, 1 uL 50 Mm MnCl2, 1 uL CircLigase Enzyme, 12 uL water. The samples were then incubated for 1 hour at 60°C thermomixer (shaked at 1300 rpm every 30'' for 15''). After circular ligation, the samples were washed twice with Low salt buffer (with 1x

Denhardt's), twice with High stringency buffer (with 1x Denhardt's) and twice with Phusion Wash Buffer (50mM Tris pH 8.0)). After the last wash, care was taken to remove every bit of liquid. Phusion Mix 1 was added to the samples (per reaction: 10 uL 5X Phusion HF Buffer, 1 uL10 mM dNTPs, 37 uL nuclease free water. The samples were incubated at 98°C for 1 minute at 1200 rpm to elute cDNA, placed on magnet and eluate was collected into PCR tubes with optically clear caps. 1.5 uL of Mix 2 (per sample: 0.5 uL 20 uM P5 primer, 0.5 uL 20 uM P3 primer, 0.5 uL Phusion DNA polymerase (NEB) and 0.5 of 50X SYBR Green (Thermo Fisher) was added (diluted in water). Samples were mixed and placed in a real-time PCR machine. The following program was used for PCR: 98°C for 30'', 40 times: 98°C for 10'', 60°C for 15'', 72°C for 20''. The samples were taken out of the PCR machine after the fluorescence reached 500 RFU. After PCR, the products were run on 2% MetaPhor agarose gel and the product (a smear around ~120 bp long) was cut out and purified using Mini Elute Gel Extraction kit (Qiagen). The purpose of this step is to remove a shorter primer-dimer product which appears on the gel as a clear line, unlike the amplified cDNA fragments which appear as a smear.

RT primers

RT-1(CGAT):

/5Phos/DDDATCGNNNNNNNAGATCGGAAGAGCGTCGT/iSp18/CACTCA/i

Sp18/CAA

GCAGAAGACGGCATACGAGATTTTTTTTTTTTTTTTTTVN

RT-2(TAGC):

/5Phos/DDDGCTANNNNNNNAGATCGGAAGAGCGTCGT/iSp18/CACTCA/i

Sp18/CAA

GCAGAAGACGGCATACGAGATTTTTTTTTTTTTTTTTTVN

RT-3(CTAG):

/5Phos/DDDGACTNNNNNNNAGATCGGAAGAGCGTCGT/iSp18/CACTCA/i

Sp18/CAA

GCAGAAGACGGCATACGAGATTTTTTTTTTTTTTTTTTVN

RT-4(GATC):

/5Phos/DDDAGTCNNNNNNNAGATCGGAAGAGCGTCGT/iSp18/CACTCA/i

Sp18/CAA

GCAGAAGACGGCATACGAGATTTTTTTTTTTTTTTTTTVN

PCR primers

DP5-PE: (Allowing standard Read1 sequencing primer)

5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC

GCTCTTCCGATCT

DP3-PAT:5'-CAAGCAGAAGACGGCATA

Sequencing primers: Illumina Standard Read 1 sequencing primer:

5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCT

Individual cTag-PAPERCLIP libraries were multiplexed and sequenced on MiSeq (Illumina) to obtain 75-nt (MiSeq) single-end reads.

Analysis of cTag-PAPERCLIP data

The processing of raw reads was performed using the CIMS software package as previously described (Moore et al., 2014). Raw reads were filtered based on quality score. Filtered reads with the exact sequence were collapsed into one. Poly(A) sequence at the 3′ end was then trimmed using CutAdapt (M. Martin, 2011). Only reads that are at least 25-nt in length were mapped to mm10 reference genome. Mapping was performed using Novoalign (Novocraft) without trimming. A minimum of 25-nt match to the genome sequence was required and only those reads mapped unambiguously to the genome (single hits) were kept for downstream analysis. Reads mapping to the same genomic positions without distinct

barcodes were further collapsed into a single tag as previously described (Moore et al., 2014). All reads went through the entire process are referred to as “unique reads.” Unique reads from both replicates were merged. CIMS software package was used to cluster overlapping reads and to determine the number of reads in each cluster. To define a list of high confidence clusters we used the following criteria: clusters had to contain unique reads from all replicates in each condition, the number of reads in a cluster had to be at least 10% of gene total, clusters had to be located within 20 kb from 3′ ends annotated by RefSeq, clusters that were located upstream of a stretch of 6 or more adenines were excluded due to potential internal priming of reverse transcription primer. To identify differences in alternative polyadenylation between different conditions, the number of tags in clusters was used as an input to EdgeR software (Robinson et al., 2010). The output p-values were adjusted for multiple hypotheses testing using the qvalue package (Storey and Bass, 2003).

Gene ontology analysis

Gene ontology analysis was performed on genes showing significant changes in alternative polyadenylation using Ingenuity Pathway Analysis software (Qiagen). Background was determined as all genes with cTag-PAPERCLIP clusters expressed in Purkinje and granule cells or P0 and P21 granule cells. To do this we used the Core Analysis tool. ‘User Dataset’ was chosen as ‘Reference Set.’ Genes with significant differences in alternative

polyadenylation were given ‘Expr Other’ value of 1, all other genes were given ‘Expr Other’ value of 0. Cutoff was chosen as 1. We focused on functions related to ‘Molecular and Cellular Functions’ and ‘Physiological System Development and Function’ that are overrepresented in the set of genes that show significant changes in alternative polyadenylation between Purkinje and granule cells and during development of granule cells. We filtered the resulting list to remove functional categories that contained less than 10 genes enriched among genes that significantly change in alternative polyadenylation.

FACS sorting and RNA sequencing

GFP-positive granule cells from P0 and P21 *Math1-Cre*; *Pabpc1*^{cTag} were sorted using FACSAria II (BD). RNA was isolated using Roche High Pure RNA isolation kit as per manufacturer’s instructions. RNA sequencing library was prepared using TruSeq kit from Illumina, following manufacturer’s instructions.

Granule cell purification and culture

Granule cells were prepared as described (Hatten, 1985). Briefly, cerebella were dissected away from the brains of P6 mice. After the pial layer was peeled away, the tissue was treated with trypsin and triturated into a single-cell suspension using fine bore Pasteur pipettes. The suspension was layered unto a discontinuous Percoll gradient and separated by centrifugation. The small cell fraction was isolated, and granule cells were further enriched by

panning on tissue culture treated plastic dishes. The resulting cultures routinely contain greater than 95% of cells of the granule cell lineage (Hatten, 1985). Granule cells were plated in serum-free granule cell medium on Poly-D lysine coated 6-well plates and transfected with 50 nM miR-124-3p antagomir (Exiqon) or Memo1 target blocker (Exiqon, custom miRCURY LNA Power Inhibitor, final concentration 50 nM for both) and appropriate negative control using Lipofectamine RNAiMAX (Thermo Fisher). Cells were then allowed to differentiate (4 days) and RNA was isolated using TRIzol (Thermo Fisher) followed by High Pure RNA Isolation Kit (Roche) as per manufacturer's instructions.

Reverse transcription and quantitative PCR

This section refers to qPCR analysis of *Memo1* abundance in experiments described in Chapter 6. Reverse transcription was performed using SuperScript III reverse transcriptase (Thermo Fisher) followed by quantitative PCR using FastStart SYBR Green Master mix (Roche) in duplicates. The cycling parameters are: 95ºC for 10 min. followed by 40 cycles of 95ºC for 15 sec., 58ºC for 30 sec., 72ºC for 20 sec. Fold changes were calculated using Bio-Rad CFX Manager Software.

Primers used

Memo1 – CDS

Forward: CATCCTCGTGCACCATAGGC

Reverse: TGAGGTCCTGAGGCGGTGTA

Memo1 – extended 3'UTR

Forward: TCCCCTACTCTCCCAGTAGTC

Reverse: AGGACTTGTTGCTTCCAGGAC

Polysome fractionation

Supernatants from adult mouse brain (N=2) were loaded onto a 20%–50% w/w linear density gradient of sucrose in gradient buffer (20 mM HEPES-KOH pH 7.4, 150 mM NaCl, and 5 mM MgCl2) prepared using a Gradient Master 107 (BioComp), in 14x3x89 mm polyallomer ultracentrifuge tubes (Beckman). Gradients were centrifuged at 40,000 rpm for 2 hours at 4°C in a Beckman SW41 rotor and sixteen fractions of 0.72 ml volume were collected with continuous monitoring at 254 nm using an ISCO UA-6 UV detector. RNA from polyribosome fractions was extracted using Trizol LS Reagent according to manufacturer's instructions (Invitrogen) and then back-extracted in chloroform:isoamyl alcohol (49:1) and spun at 10,000 3 g for 15 min. The aqueous phase was precipitated with 1 uL GlycoBlue (Ambion) and 1 volume isopropanol. RNA was pelleted at 20,000 x g for 10 min at 4°C, washed twice with cold 75% ethanol and resuspended in water. Samples were digested with RQ1 DNase (Promega) for 1 hour at 37°C, extracted with phenol-chloroform, back-extracted with chloroform:isoamyl alcohol, precipitated with GlycoBlue in 2 volumes of 1:1 ethanol:isopropanol and washed as before. cDNA was prepared with equal volumes of RNA from each fraction, using random hexamers (Roche) and Superscript III Reverse Transcriptase

(Invitrogen) according to manufacturer's instructions. cDNAs were PCR amplified using FastStart Universal SYBR Green Master Mix (Roche). Primers are listed in Reverse transcription and quantitative PCR section below. Relative RNA levels were calculated using the formula relative RNA amount = 2^(40-Ct). For each gene, polyribosome fractions were plotted as percentages of the total amount of the gene on the gradient (obtained by summing values of individual fractions). Error bars refer to biological replicates of qPCR.

Single molecule RNA FISH

Granule cells were purified as described above and cultured for 14 days in granule cell medium on poly-D-lysine coated cover slips in 24-well plates. Medium was changed every 3 days (1 volume of the medium was added to the well and then 1 volume was removed). In situ hybridization was performed using the QuantiGene ViewRNA kit (Affymetrix) using probe sets aligning to *Memo1* coding sequence and extended 3'UTR (catalog numbers: VB1-20807 (coding sequence) and VB6-20266 (extended 3'UTR)). The Proteinase K treatment was omitted to preserve the integrity of the dendrites. After completion of in situ hybridizations, neurons were processed for immunofluorescence using anti-GFP antibody (described above).

Chapter 3 | Mapping of 3'UTR ends in specific cells *in vivo* with cTag-PAPERCLIP

Introduction

As mentioned in Chapter 1, we recently developed a new approach to map 3'UTR ends of polyadenylated transcripts through their interaction with poly-A binding protein cytoplasmic 1 (PABPC1) (Hwang et al., 2016) that we called poly(A) binding protein-mediated mRNA 3' end retrieval by crosslinking immunoprecipitation (PAPERCLIP). PAPERCLIP was shown to be superior to the existing techniques in accurately pinpointing the ends of 3'UTRs (Hwang et al., 2016). In this Chapter, I describe an adaptation of the PAPERCLIP technique that enables transcriptome profiling and 3'UTR mapping of selected cell populations from intact tissues, without having to first physically separate individual cell types from the tissue. To profile individual cell-types from an intact tissue, RNA-seq and other methods for high-throughput mapping of 3'UTR ends require that the cell population of interest is first physically isolated from the rest of the tissue. This can result in cellular stress, which may alter gene expression (Okaty et al., 2011). In our new technique, we genetically engineered the endogenous PABPC1 gene to be conditionally fused with GFP after Cre recombination. In this way, GFP-tagged PABP can be conditionally expressed *in vivo* in selected cell populations that express Cre recombinase. We called this mouse "cTag-PABP" since PABP is conditionally ("c") tagged with GFP. We called the overall protocol "cTag-

PAPERCLIP", since here we apply the PAPERCLIP technique to the cTag-PABP mouse. In the original PAPERCLIP protocol, anti-PABP antibodies were used to purify PABPC1-bound mRNAs. In this new protocol, a mixture of high-affinity GFP antibodies is used to purify GFP-tagged PABPC1 from Cre-expressing cells. The mouse and the protocol were developed by HW Hwang, who used the approach to profile the 3' ends of RNAs in various cells in the brain, including the study of alternative polyadenylation in microglia upon a lipopolysaccharide (LPS) challenge (HW Hwang, in review). I, in parallel, applied this technique to study alternative polyadenylation in cerebellar Purkinje and granule cells (described in this Chapter and Chapter 4) as well as granule cell development (described in Chapter 5).

Results

cTag-PAPERCLIP maps 3'UTR ends in specific neurons *in-vivo*

The genetically modified mouse used in the cTag-PAPERCLIP protocol, termed cTag PABP, expresses GFP-tagged PABPC1 in specific cell types in a Cre-dependent way (HW Hwang et. al., in review). The mouse carries a *Pabpc1* allele where two wild-type terminal exons are floxed, followed by a terminal exon that has GFP coding sequence inserted before the stop codon. After Cre-recombination, the two floxed terminal exons are removed and the exon with GFP coding sequence is inserted in the endogenous *Pabpc1* gene, which leads to the expression of PABPC1-GFP fusion protein (Figure 3-1). GFP antibody can then be used to

immunoprecipitate all polyadenylated transcripts present in a cell expressing GFP-tagged PABPC1. A schematic of the cTag-PAPERCLIP protocol is shown in Figure 3-2. In brief, crosslinked tissue is lysed, RNA partially digested and RNA-PABPC1-GFP complexes immunoprecipitated using anti-GFP antibodies. The immunoprecipitated RNA is then radiolabeled, RNA-PABPC1-GFP complexes are run on the SDS-PAGE and transferred to nitrocellulose membrane to visualize. RNA-protein complexes are then cut from the membrane (this step provides size selection), RNA is recovered from the membrane, reverse transcribed using anchored oligo (dT) primer, circularized, PCR amplified and sequenced. The sequencing reads obtained with cTag-PAPERCLIP consist of stretches As which are removed, and sequences corresponding to the ends of 3'UTRs, which are mapped to the genome and clustered to identify 3'UTR ends.

HW Hwang has shown that there is a high correlation between the total number of reads per gene obtained from regular PAPERCLIP versus cTag-PAPERCLIP (R=0.91), which suggests that the addition of GFP does not affect the binding of PABPC1 to mRNA (HW Hwang et. al., in review).

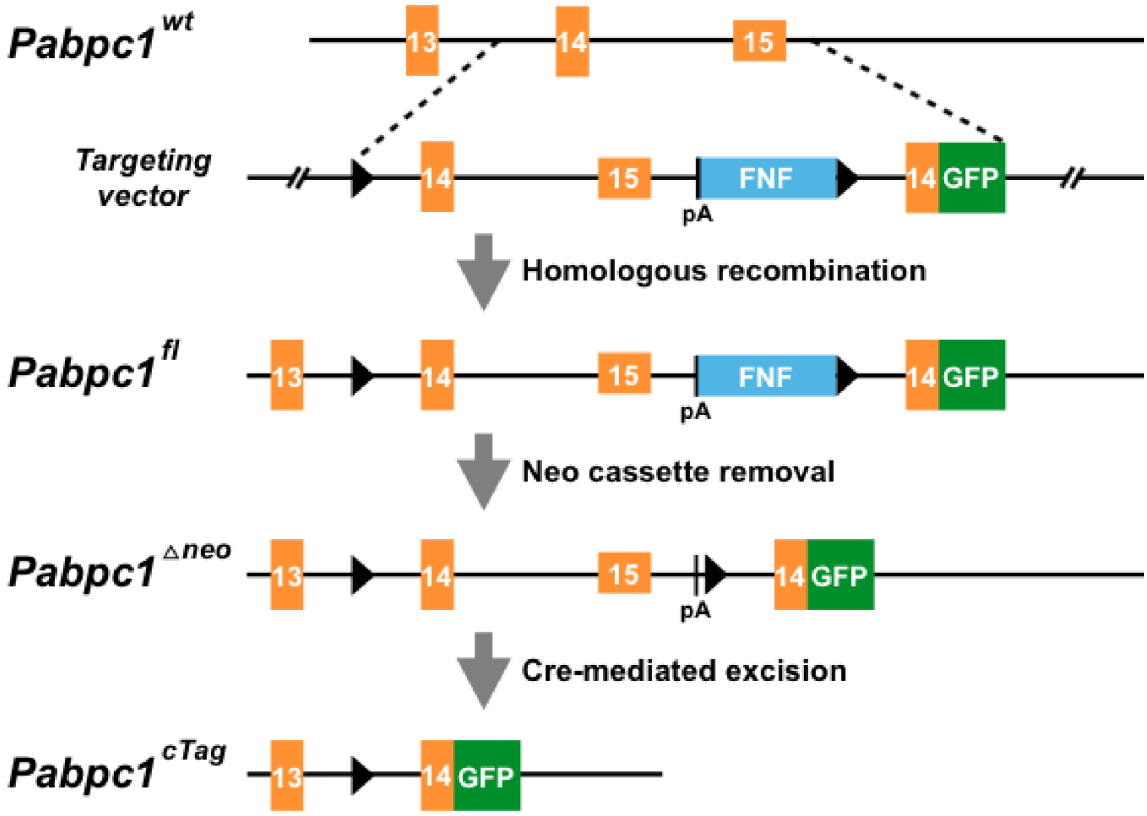


Figure 3-1 Targeting strategy for generating the knock-in cTag-PABP mouse

The targeting vector contained the endogenous *Pabpc1* exons 14 and 15, a synthetic poly(A) site (labeled pA), a *FRT-Neo-FRT* (labelled FNF) cassette and the last coding exon of *Pabpc1* fused to AcGFP sequence**.** Black arrowheads represent loxP sites (Figure courtesy of HW Hwang).

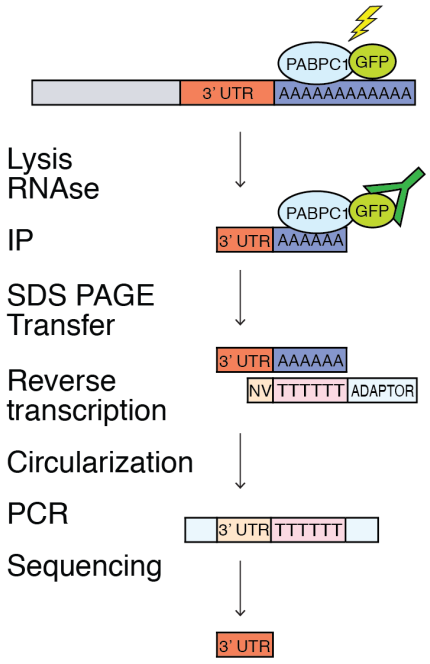


Figure 3-2 Schematic of the cTag-PAPERCLIP protocol
Tissue from mice expressing Pabpc1-GFP in specific cell types in first crosslinked and lysed. RNA is partially digested and RNA-PABPC1-GFP complexes are immunoprecipitated using anti-GFP antibodies. The immunoprecipitated RNA is then radiolabeled, RNA-PABPC1-GFP complexes are run on the SDS-PAGE and transferred to nitrocellulose membrane to visualize. RNA-protein complexes are then cut from the membrane, RNA is recovered from the membrane, reverse transcribed using anchored oligo (dT) primer, circularized, PCR amplified and sequenced.

To purify polyadenlyated RNAs in a cell-type specific manner, we bred cTag-PABP mice to *Pcp2-Cre* or *NeuroD1-Cre* mice that express Cre recombinase specifically in cerebellar Purkinje or granule cells, respectively. We first confirmed the cellular specificity of PABPC1-GFP expression by performing immunofluorescence staining of cerebellar sections using anti-GFP antibody. We observed GFP expression only in the cells of interest (Figure 3-3).

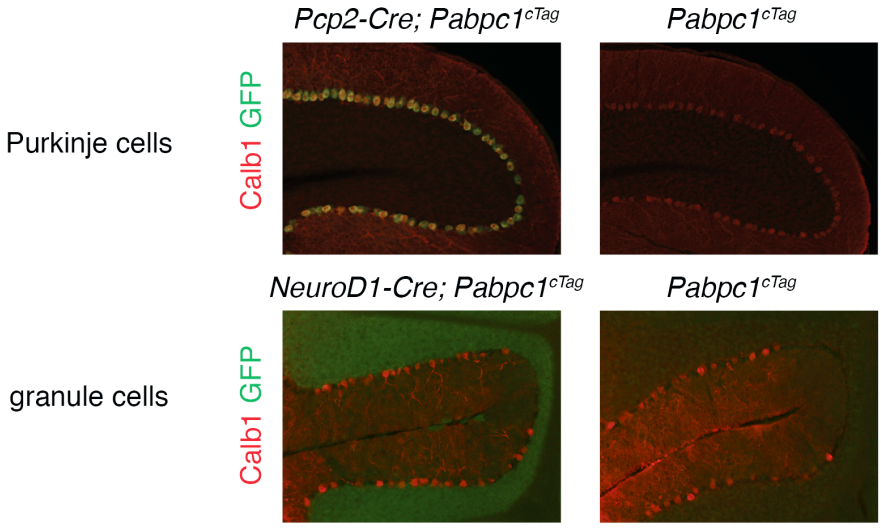


Figure 3-3 Validation of PABPC1-GFP expression in the adult cerebellar neurons

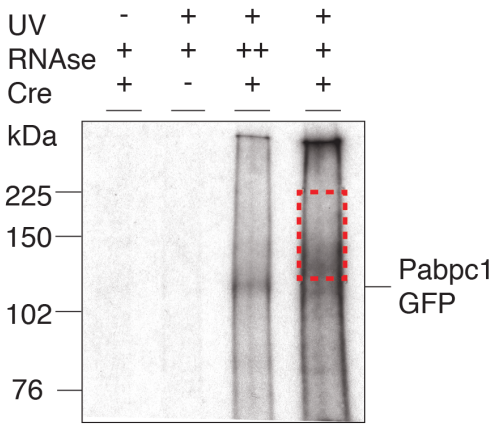
Immunofluorescence staining of cerebella from mice expressing PABPC1-GFP in Purkinje (top row) and granule cells (bottom row). Calb1: Purkinje cell marker**.** In the top row, Purkinje cells are double labelled with GFP and Calb1 and appear yellow. In the bottom row, granule cells are labelled only with GFP, whereas Purkinje cells are in a layer adjacent to granule cell layer and are labelled with Calb1.

To perform cTag-PAPERCLIP we used cerebella from 5-7 P56 *Pcp2-Cre; Pabpc1*^{cTag} and 2 P56 *NeuroD1-Cre; Pabpc1*^{cTag} mice per replicate that we crosslinked immediately after dissection. After RNase overdigestion of crosslinked samples and ^{32}P-labeling of residual RNA co-immunoprecipiated with GFP antibody, we observed a single band on the RNA autoradiogram corresponding to the size of PABPC1 protein fused to GFP (Figure 3-4A and Figure 3-5A).

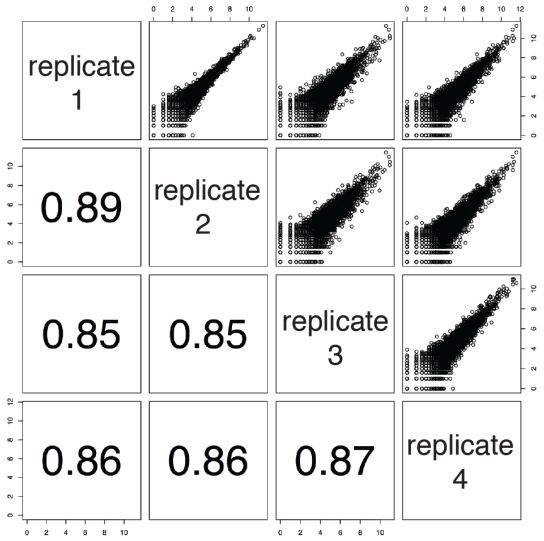
Figure 3-4 cTag-PAPERCLIP on granule cells: autoradiogram, reproducibility and genomic distribution of CLIP tags

(A) Autoradiogram of radiolabelled RNA cross-linked to PABPC1-GFP. Lanes from left to right: No-crosslinking control, *Cre*- control, overdigested control, sample used for sequencing. Red box represents the area that was cut out of the nitrocellulose membrane and processed for sequencing. **(B)** Correlation between the total number of cTag-PAPERCLIP reads per gene in three biological replicates from granule cells. R: Pearson correlation coefficient. **(C)** Genomic distribution of cTag-PAPERCLIP reads in four replicates (replicates 1-4) from granule cells. Genomic distribution of cTag-PAPERCLIP clusters that contain reads from at least three biological replicates (BC3 clusters) from granule cells.

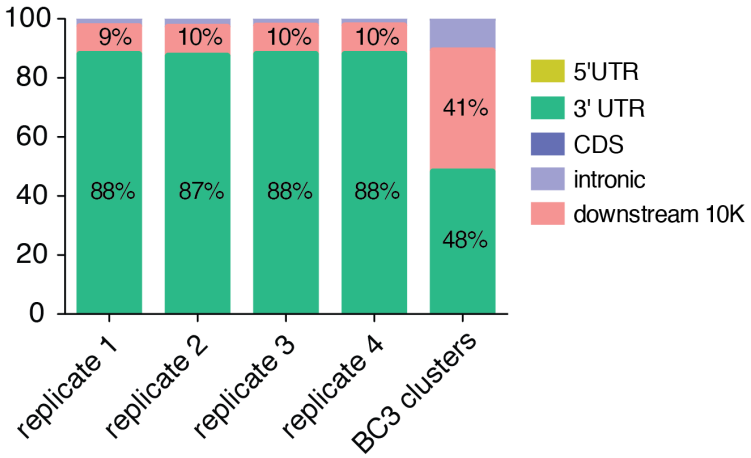
A


UV - + + +
RNAse + + ++ +
Cre + - + +
kDa
225
150
102
76
Pabpc1
GFP

B


replicate 1
0.89
replicate 2
0.85
0.85
replicate 3
0.86
0.86
0.87
replicate 4

C


100
80
60
40
20
0
9%
10%
10%
10%
88%
87%
88%
88%
41%
48%
5'UTR
3' UTR
CDS
intronic
downstream 10K
replicate 1
replicate 2
replicate 3
replicate 4
BC3 clusters

Sequencing of immunoprecipitated partially digested RNA fragments demonstrated that the technique was highly reproducible between multiple biologic replicates (Figure 3-4B and Figure 3-5B). Close to 90% cTag-PAPERCLIP reads mapped to the ends of annotated 3'UTRs of RefSeq-annotated genes (Figure 3-4C and Figure 3-5C).

We also calculated the correlation coefficient between the sum of all cTag-PAPERCLIP reads per transcript (representing the total abundance of 3'UTR isoforms) and RNA abundance (of all 3'UTR isoforms) measured by RNA sequencing of FACS sorted neurons and found that the two measurements are well correlated. This experiment is discussed in Chapter 5.

In addition, using the sum of cTag-PAPERCLIP reads per gene as a measure of gene expression, we examined the expression of marker genes for Purkinje cells. We found that known markers were among the most highly expressed genes in the cTag-PAPERCLIP dataset (Figure 3-6).

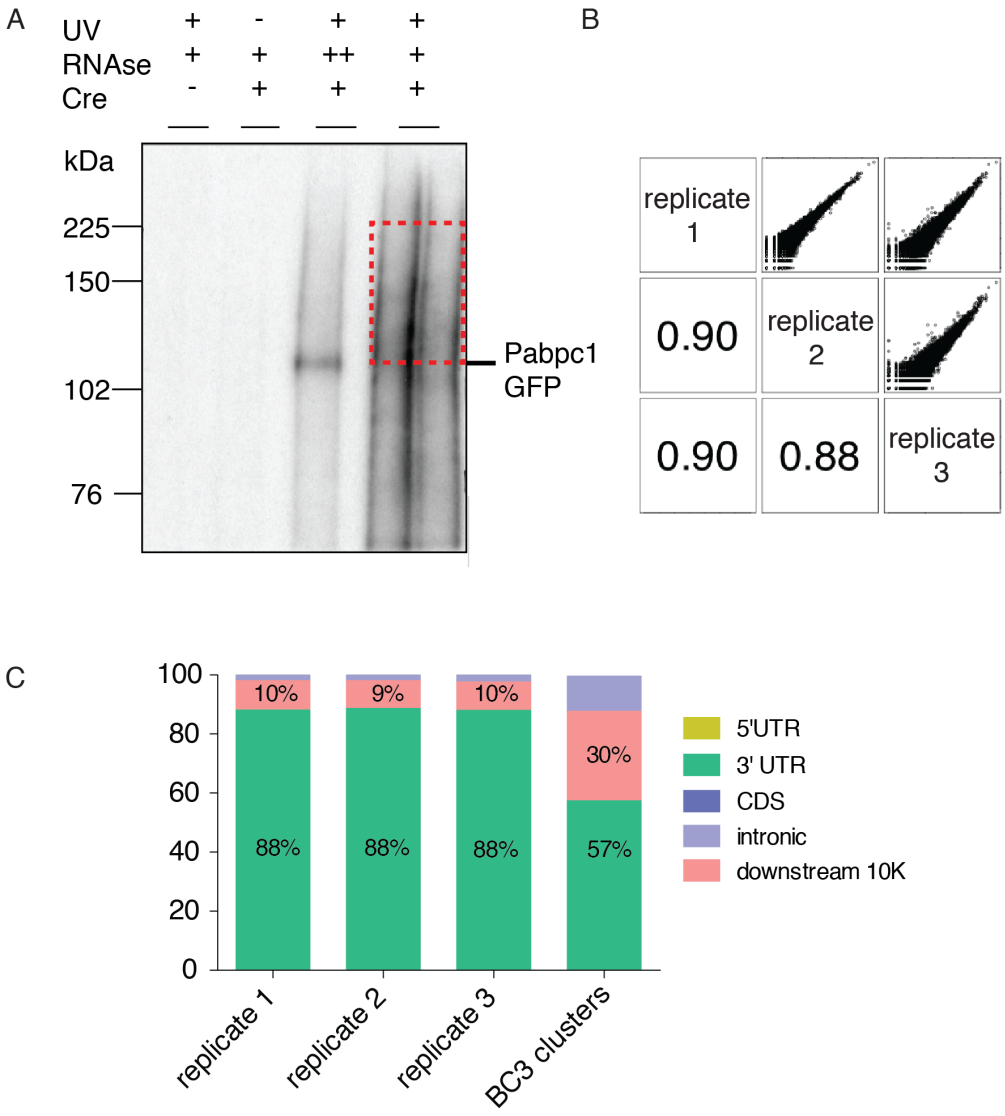


Figure 3-5 cTag-PAPERCLIP on Purkinje cells: autoradiogram, reproducibility and genomic distribution of CLIP tags

(A) Autoradiogram of radiolabelled RNA cross-linked to PABPC1-GFP. Lanes from left to right: *Cre*- control, no-crosslinking control, overdigested control, sample used for sequencing. Red box represents the area that was cut out of the nitrocellulose membrane and processed for sequencing. **(B)** Correlation between the total number of cTag-PAPERCLIP reads per gene in three biological replicates from Purkinje cells. R: Pearson correlation coefficient. **(C)** Genomic distribution of cTag-PAPERCLIP reads in three replicates (replicates 1-3) from Purkinje cells. Genomic distribution of cTag-PAPERCLIP clusters that contain reads from all three biological replicates (BC3 clusters) from Purkinje cells.

cTag-PAPERCLIP is selective at purifying mRNA from specific neurons

In addition, we ranked marker genes from non-target cell types (granule cells and glia) by expression in cTag-PAPERCLIP data and compared the them to the ranks of the same markers in Purkinje cell expression data obtained by TRAP-sequencing, an alternative method for profiling gene expression in specific cell types (Mellén et al., 2012). The analysis showed that marker genes from non-target cell types were ranked significantly lower in cTag-PAPERCLIP data (Figure 3-7), suggesting that cTag-PAPERCLIP is particularly selective at isolating mRNA from individual cell types.

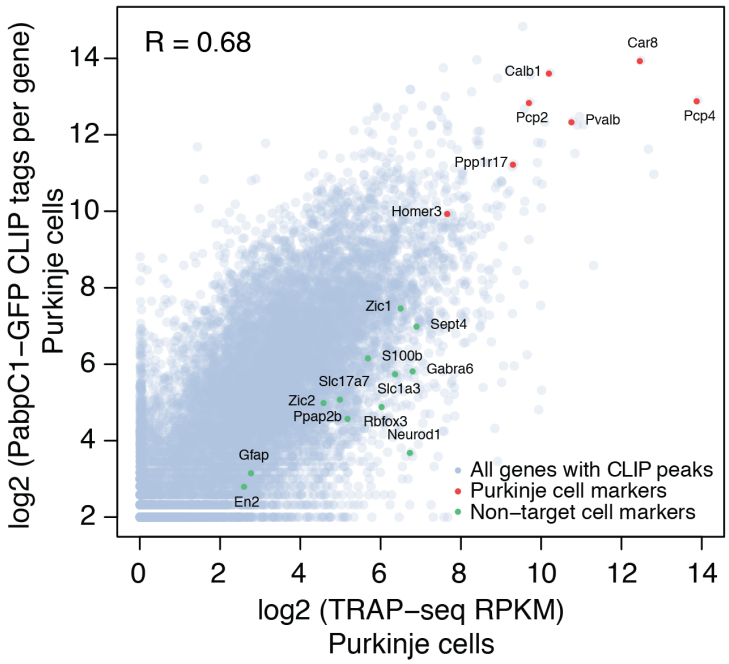


Figure 3-6 Comparison between cTag-PAPERCLIP and TRAP-seq
The x axis represents TRAP-seq Reads Per Kilobase per Million mapped reads (RPKM) per gene and the y axis represents the total unique cTag-PAPERCLIP reads per gene from Purkinje cells. Purkinje cell markers are highlighted in red and non-target cell markers (glial and granule cell markers) are highlighted in green.

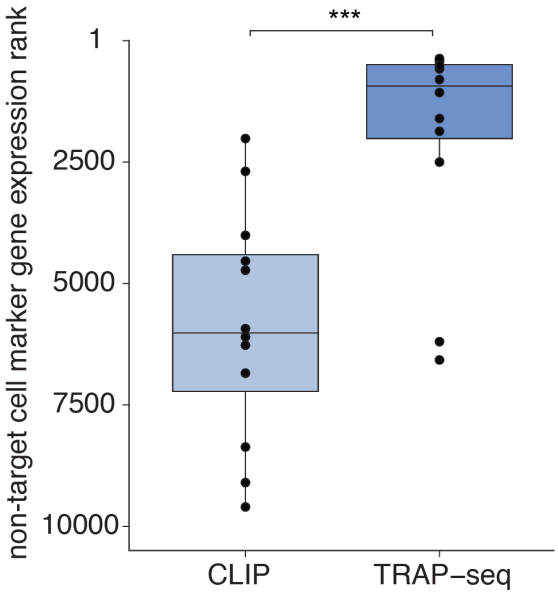


Figure 3-7 Expression of marker genes from non-target cell types in Purkinje cells

Genes in each dataset (Purkinje cell cTag-PAPERCLIP and Purkinje cell TRAP-seq) were ranked by expression. The highest expressed gene in each dataset has a rank of 1. The distribution of gene ranks for markers of non-target cells (glia and granule cells – genes highlighted in green in Figure 3-6) is shown. $P < 0.001$, Mann-Whitney test.

Discussion

A major question in neuroscience, and in biology more generally, is to understand the function of individual cell types within complex tissues. Identifying the complement of mRNA molecules expressed in specific cell types is crucial to achieve this goal. In this Chapter, I described cTag-PAPERCLIP, a new method that enables identification of 3'UTR isoforms expressed in selected cell populations *in vivo*.

Importantly, this is the first method that enables precise mapping of 3'UTR ends in specific cell types, which cannot be done with existing RNA-seq-based methods such as TRAP-seq, given that they lack the resolution to identify exact mRNA 3' ends (Miura et al., 2014).

In addition, another major advantage of this method is that it is very selective in purifying transcripts from specific cell types, and is superior in this regard compared to TRAP-sequencing, which has been previously shown to suffer from contamination from transcripts from non-target cell types (Okaty et al., 2011), possibly due to reassociation of RNA from non-tagged ribosomes with tagged ribosomes in the cell lysate (Mili and Steitz, 2004) or non-specific binding due to incomplete washing. These issues can be particularly serious when purifying RNA from rare cell types, such as Purkinje cells in the cerebellum.

In addition, cTag-PAPERCLIP is also compatible with Cre/Lox technology, allowing it to be used to study biological questions in any cell-type for which a Cre-driver is available.

In the next chapter, I describe how I used the cTag-PAPERCLIP data to analyze the alternative polyadenylation patterns in granule and Purkinje cells.

Chapter 4 | Comparison of alternative polyadenylation between Purkinje and granule cells

Introduction

We know that different tissues differ in their use of 3'UTR isoforms. Tissue-specific alternative polyadenylation preferences were first documented by analyzing datasets of expressed sequence tags (ESTs) (Haibo Zhang et al., 2005) and more recently, by analyzing high-throughput sequencing data (Lianoglou et al., 2013; Miura et al., 2013; Smibert et al., 2012). But what if we look within a tissue, and compare alternative polyadenylation between related cell types?

A recent study looked at alternative polyadenylation of the *Pax3* transcript, a key regulator of myogenesis, between two closely related subsets of muscle stem cells that differ in their anatomical location (limb muscles vs. diaphragm) and the expression of the Pax3 protein. The authors of the study found that the difference in Pax3 levels stems from the fact that the two cell populations differ in their expression of the 3'UTR isoforms of *Pax3*. The long isoform is expressed in limb muscle stem cells, which is sensitive to miR-206 downregulation. In diaphragm stem cells, the short isoform is expressed, which is insensitive to miR-206 regulation (miR-206 is expressed in high levels in both cell types). This results in markedly different levels of Pax3 protein levels between the two types of stem cells. Treating the two cell populations with a miR-206 inhibitor results in significant differences

in the proliferation of stem cells upon activation (Boutet et al., 2012).The study illustrates how differences in alternative polyadenylation between two related types can have a marked effect on protein expression of a transcript and in turn determine the identity of a cell population.

The differences in alternative polyadenylation in neurons, which are especially interesting due to their unique function, morphology and exceptional 3'UTR lengths (Miura et al., 2013) have not yet been explored. In this Chapter, I address whether neurons exhibit differences in polyadenylation and if so, whether these differences could be partially responsible for what makes these cells different from each other. Specifically, we compared 3'UTR isoform expression in Purkinje cells and cerebellar granule cells. We picked these two neuronal cell types because they are functionally and morphologically distinct; granule cells are small, excitatory and have only 4-5 dendritic branches, whereas Purkinje cells are large, inhibitory and have an exceptionally large dendritic arbor (Kandel et. al., 2012).

Results

Purkinje and granule cells differ in expression of 3'UTR isoforms

To compare cTag-PAPERCLIP data from Purkinje and granule cells we used EdgeR software, originally developed for analysis of RNA-sequencing data (Robinson et al., 2010). This analysis revealed extensive isoform diversity between the two cell types. 551 genes displayed significant differences in 3'UTR isoform expression between the two cell types (using FDR < 0.05 as a significance cut-off) (Table 1). Notably, Purkinje cells tended to express fewer long 3'UTR isoforms (produced by 3'UTR alternative polyadenylation) compared to granule cells (Figure 4-1) as well as fewer isoforms with distal alternative last exons (produced by upstream alternative polyadenylation) (Figure 4-2).

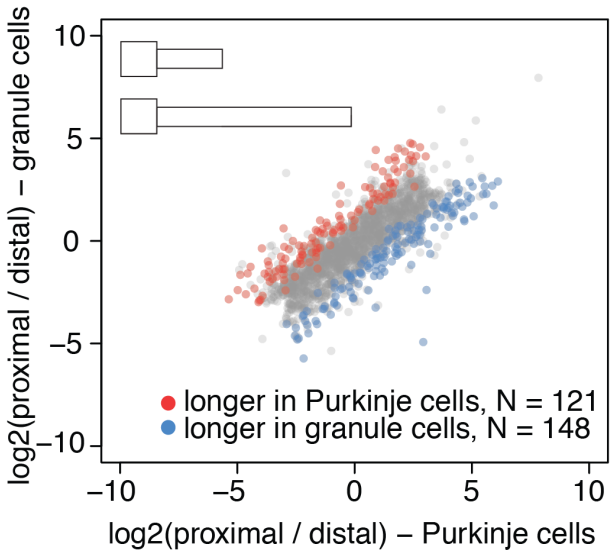


Figure 4-1 Comparison between 3'UTR alternative polyadenylation in Purkinje and granule cells

Scatterplot representing the ratio between the number of cTag-PAPERCLIP reads at the end of the proximal 3'UTR isoform and the number of cTag-PAPERCLIP reads at the end of the distal 3'UTR isoform in Purkinje vs. granule cells. Only genes subjected to 3'UTR alternative polyadenylation, with two isoforms are shown. Genes showing significantly different ratios (FDR < 0.05) are highlighted in red or blue.

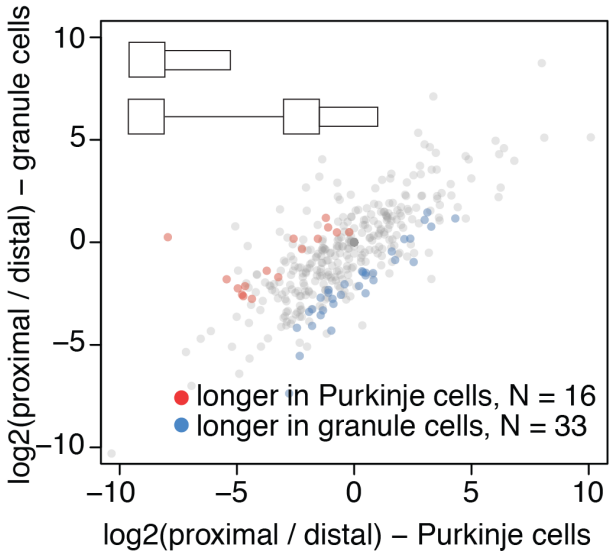


Figure 4-2 Comparison between upstream alternative polyadenylation in Purkinje and granule cells

Scatterplot representing the ratio between the number of cTag-PAPERCLIP reads at the end of the proximal 3'UTR isoform and the number of cTag-PAPERCLIP reads at the end of the distal 3'UTR isoform in Purkinje vs. granule cells. Only genes subjected to upstream alternative polyadenylation, with two isoforms are shown. Genes showing significantly different ratios (FDR < 0.05) are highlighted in red or blue.

Gene ontology analysis of differentially polyadenylated transcripts

Interestingly, using gene ontology analysis (Ingenuity software) we found that genes involved in the regulation of cell morphology and neurotransmission were significantly enriched among genes expressing different 3'UTR isoforms in the two cell types, with genes involved in branching of neurites being the most highly enriched (Figure 4-3 and Table 2).

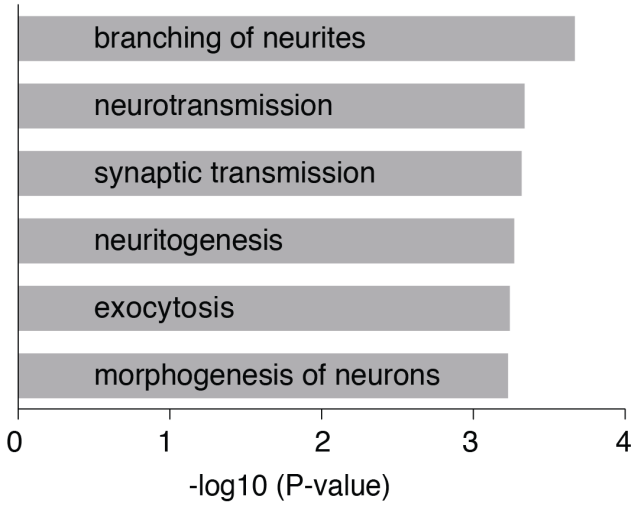


Figure 4-3 Gene ontology analysis of genes showing significant differences in alternative polyadenylation between Purkinje and granule cells.

Top 6 gene ontology categories are shown.

For example, neural cell adhesion molecule (*NCAM1)* gene, which is involved in neurite outgrowth (Fields and Itoh, 1996), expressed a different version of mRNA in Purkinje cells compared to granule cells (Figure 4-4). This data suggests that the striking morphological differences between granule and Purkinje cells may be mediated in part by fine tuning of gene expression by different 3'UTRs.

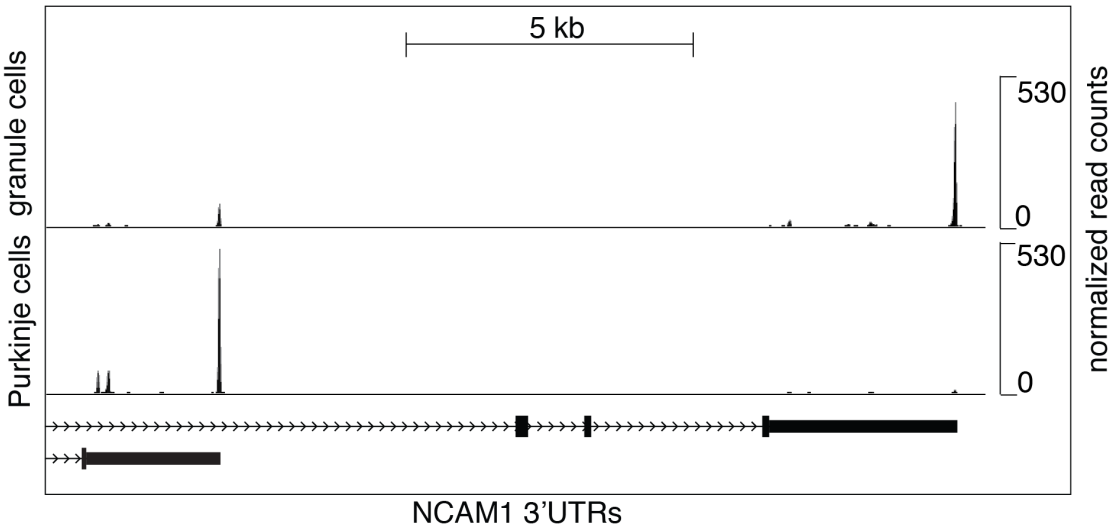


Figure 4-4 Alternative polyadenylation of the *NCAM1* transcript
cTag-PAPERCLIP clusters on the *NCAM1* 3'UTRs. The height of the clusters represents the average number of cTag-PAPERCLIP tags in a cluster (from multiple biological replicates) normalized to the total number of mapped reads. The gene structure corresponds to the RefSeq annotation.

Discussion

In this Chapter, I described large differences in the profile of alternative polyadenylation between two functionally and morphologically distinct cell types. Given that neurons are exceptionally diverse (Macosko et al., 2015), I hypothesize that alternative polyadenylation represents one mechanism that could contribute to their specific, and varied characteristics. Indeed, genes that showed differences in alternative polyadenylation between Purkinje and granule cells are involved in cellular functions like neurite branching and synaptic transmission, which are consistent with the morphological and physiological differences between these two neuronal types. This is, to our knowledge, the first demonstration of genome wide differences in alternative polyadenylation patterns at the individual neuronal-type level in an adult mouse brain *in vivo*.

In the future, it will be necessary to study the function of these neural 3'UTR isoforms *in vivo*, by deleting the 3'UTR regulatory regions in the same way that was done for the *BDNF* transcript, where the phenotype of a mouse that lacked the *BDNF* 3'UTR extension was studied (An et al., 2008). The advances in CRISPR/Cas9-mediated genome engineering may speed up such efforts (Mali et al., 2013). However, given the differences in alternative polyadenylation between neuronal types described in this thesis, it will be important to consider which neuron type is the most appropriate to study each of the 3'UTR isoforms.

Chapter 5 | Changes in alternative polyadenylation during granule cell development

Introduction

In the previous Chapter, I showed that two different types of neurons differ in alternative polyadenylation. This shows that it is important to study any process involving changes in alternative polyadenylation on the level of a specific cell type, instead of the whole tissue, given that different neurons will likely exhibit their own set of cell-type specific changes in alternative polyadenylation.

An important question in the field of alternative polyadenylation is what are the functional consequences of changes in alternative polyadenylation during embryonic development. To study this question in a specific cell type, *in vivo*, we took advantage of cTag-PAPERCLIP. We chose to study the development of granule cells in the cerebellum because they are relatively abundant as precursors, and are very abundant in the adult cerebellum, which simplified the cTag-PAPERCLIP procedure. Moreover, it is relatively easy to perform various functional studies on them, because they are easy to purify (Hatten, 1985); in addition, functional studies can also be done in *ex vivo* cerebellar explants, which provide a faithful model for the *in vivo* development of granule cells (X. Zhu et al., 2016). This makes cerebellar granule cells an ideal model system to study neuronal development.

We chose two time points for our developmental analysis: P0 (the day of birth), when granule cell precursors are proliferating in the external granule layer, and P21 (21 days after birth), when granule cells have completed their migration to the internal granule layer, and have developed dendritic processes and formed synaptic connections with Purkinje cells and ingrowing mossy fibers.

Results

Adult granule cells express longer 3'UTR isoforms compared to granule cell precursors

To purify granule cells at these two time points, we bred mice that expressed Cre under the control of the *Math1* promoter, enabling the expression of GFP-tagged PABPC1 in granule cell precursors at P0 and adult granule cells at P21. To only purify polyadenylated mRNA from granule cell precursors and adult granule cells, we performed cTag-PAPERCLIP on dissected cerebellar cortices and discarded deep cerebellar nuclei, where *Math1* is expressed as well (Figure 5-1).

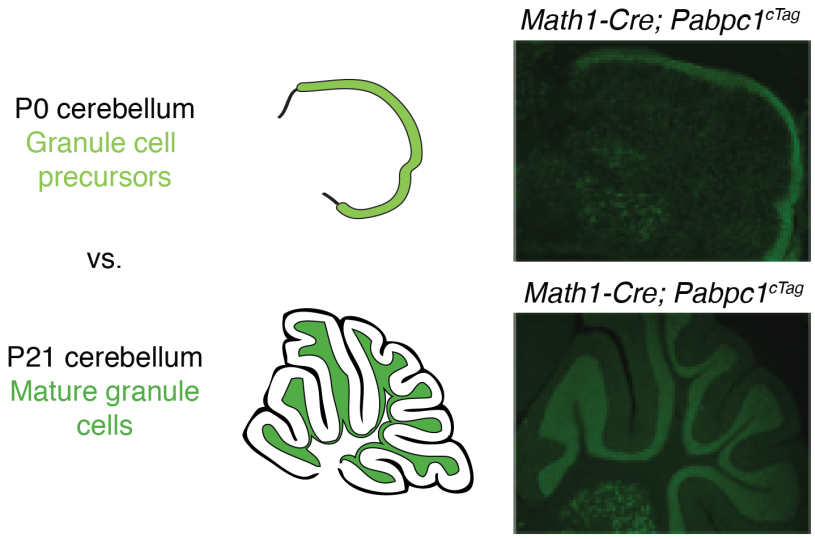


Figure 5-1 Validation of PABPC1-GFP expression in the developing and adult granule cells

Immunofluorescence staining showing cell type specific expression of conditionally tagged Pabpc1 in P0 and P21 granule cells.

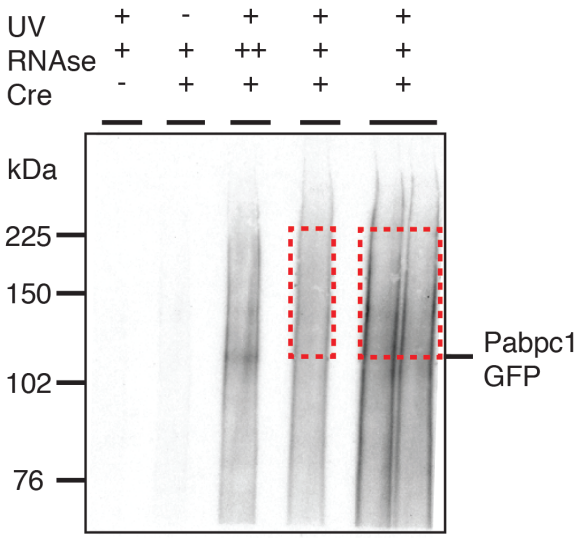
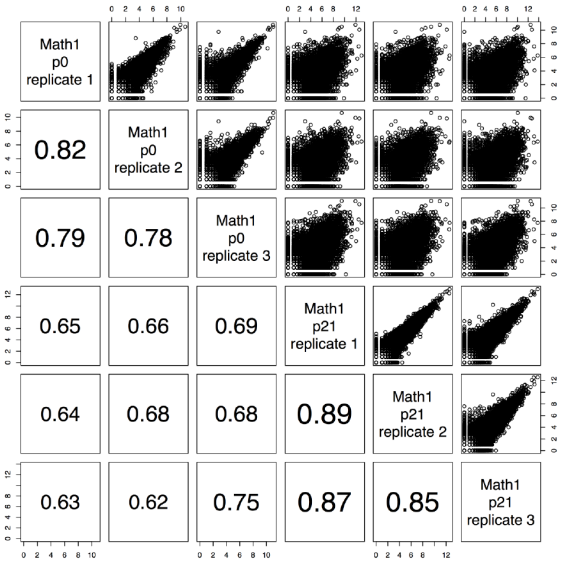


Figure 5-2 cTag-PAPERCLIP on granule cell precursors and adult granule cells: Autoradiogram

Autoradiogram of radiolabelled RNA cross-linked to PABPC1-GFP. Lanes from left to right: *Cre*- control, no-crosslinking control, overdigested control, P0 sample used for sequencing, P21 sample used for sequencing. Red box represents the area that was cut out of the nitrocellulose membrane and processed for sequencing.

A



B

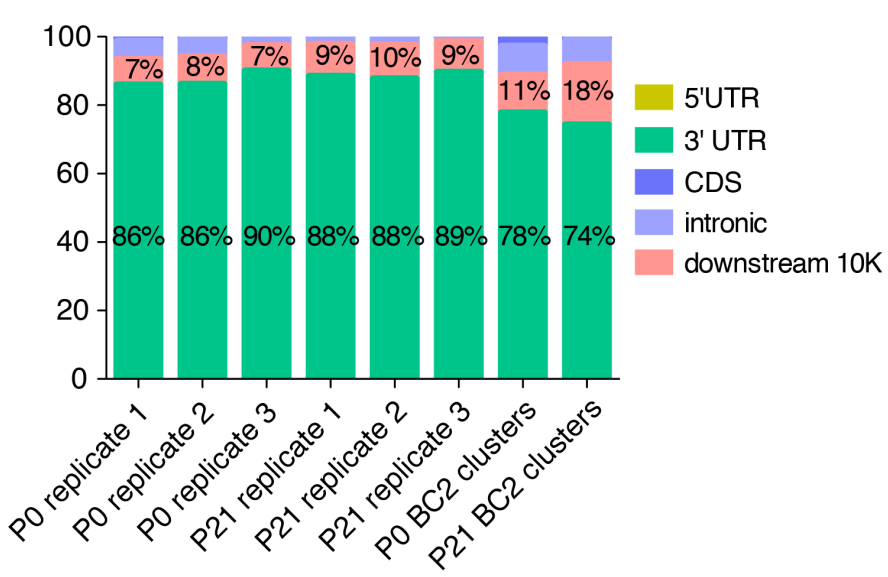


Figure 5-3 cTag-PAPERCLIP on granule cell precursors and adult granule cells: reproducibility and genomic distribution of CLIP tags

(A) Correlation between the total number of cTag-PAPERCLIP reads per gene in three biological replicates from granule cells. R: Pearson correlation coefficient. **(B)** Genomic distribution of cTag-PAPERCLIP reads in three replicates (replicates 1-3) from P0 and P21 granule cells. Genomic distribution of cTag-PAPERCLIP clusters that contain reads from at least two biological replicates (BC2 clusters) from P0 and P21 granule cells.

Analysis of cTag-PAPERCLIP data from the two time points showed that 599 genes exhibited significant changes in 3'UTR isoform expression between granule cell progenitors and mature granule cells (FDR < 0.05) (Table 3). 3'UTRs tended to lengthen during granule cell development (Figure 5-4 and Figure 5-5), consistent with data from whole developing mouse brains (Ji et al., 2009).

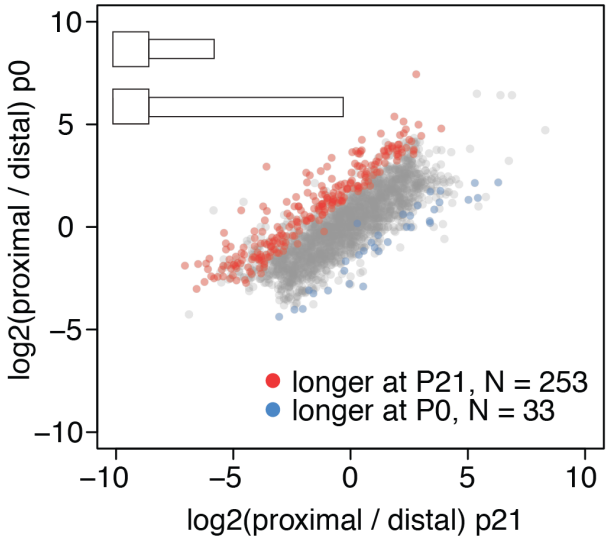


Figure 5-4 Comparison between 3'UTR-alternative polyadenylation in granule cell precursors and adult granule cells

Scatterplot representing the ratio between the number of cTag-PAPERCLIP reads at the end of the proximal 3'UTR isoform and the number of cTag-PAPERCLIP reads at the end of the distal 3'UTR isoform in P0 vs. P21 granule cells. Only genes subjected to 3'UTR-alternative polyadenylation, with two isoforms are shown. Genes showing significantly different ratios (FDR < 0.05) are highlighted in red or blue.

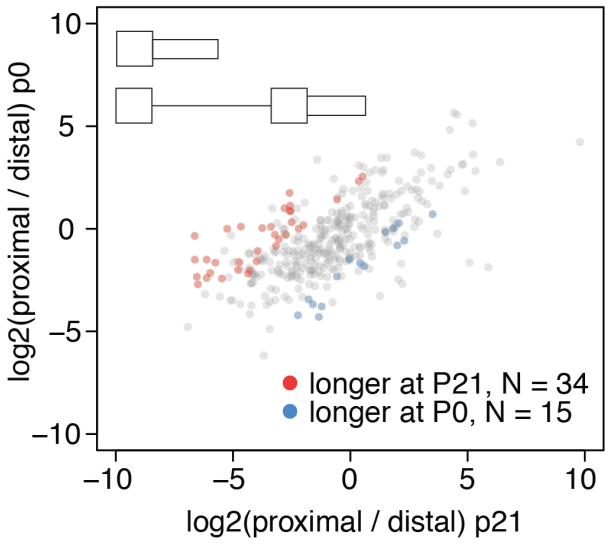


Figure 5-5 Comparison between upstream-alternative polyadenylation in granule cell precursors and adult granule cells

Scatterplot representing the ratio between the number of cTag-PAPERCLIP reads at the end of the proximal 3'UTR isoform and the number of cTag-PAPERCLIP reads at the end of the distal 3'UTR isoform in P0 vs. P21 granule cells. Only genes subjected to upstream-alternative polyadenylation, with two isoforms are shown. Genes showing significantly different ratios (FDR < 0.05) are highlighted in red or blue.

Correlation between cTag-PAPERCLIP and RNA-seq data

To independently validate the cTag-PAPERCLIP data we performed FACS sorting of GFP-positive granule cells from *Math1-Cre*; *PABP*^{cTag} mice. We then isolated RNA from the sorted cells and performed RNA-sequencing. There was a good correlation between the total number of cTag-PAPERCLIP reads per gene (i.e. a measure of total abundance of 3'UTR isoforms) and RNA-seq Reads Per Kilobase per Million mapped reads (RPKM, averaged over the whole length of a gene, including all isoforms). The correlation coefficient was 0.56. We found an even better correlation between

the ratio of RPKM per gene at P0 and P21 and the ratio of cTag-PAPERCLIP reads per gene at P0 and P21. In this case, the correlation coefficient was 0.63 (Figure 5-6). We also compared the fold change in the ratio between the short and long 3'UTR isoforms at P0 vs. P21 calculated from RNA-seq with the fold change in the same ratio calculated from cTag-PAPERCLIP. 79% of all genes that significantly change in 3'UTR isoform expression during granule cell development (as defined by cTag-PAPERCLIP data) showed the fold change in the ratio in RNA-seq data in the same direction as in cTag-PAPERCLIP data. These data indicate that cTag-PAPERCLIP quantitatively reflects 3'UTR isoform abundance.

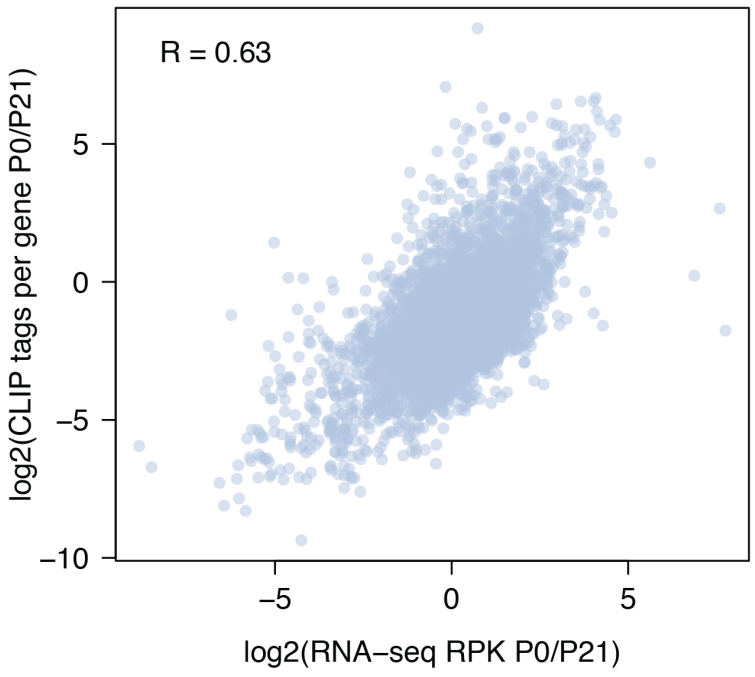


Figure 5-6 Comparison between cTag-PAPERCLIP and RNA-seq
The x axis represents the ratio between the reads per kilobase (RPK) for all expressed genes in granule cells at P0 and P21. The y axis represents the ratio between the total number of unique CLIP reads per gene (i.e. sum of all 3'UTR isoforms) in granule cells at P0 and P21. R: Pearson correlation coefficient.

Gene ontology analysis of differentially polyadenylated transcripts

Using gene ontology analysis (Ingenuity software), we found that the genes that showed different 3'UTR usage in progenitor compared to mature granule cells were significantly enriched for genes involved in the regulation of cell morphology and cell proliferation, two main features that are different between neuronal precursors and their mature counterparts (Figure 5-7 and Table 4).

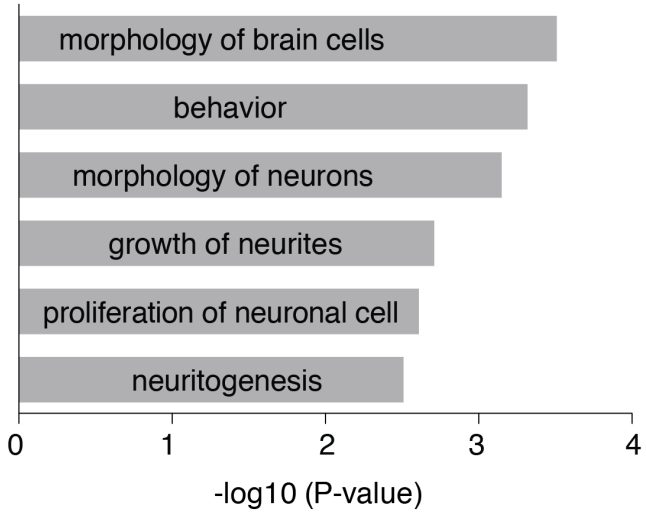


Figure 5-7 Gene ontology analysis of genes showing significant differences in alternative polyadenylation between granule cell precursors and adult granule cells.
Top 6 gene ontology categories are shown.

For example, *Sin3B,* a transcript coding for a protein that plays a role in cell cycle exit (David et al., 2008), displayed a large shift in 3'UTR isoform usage between the mitotic granule cell precursors and adult granule cells (Figure 5-8).

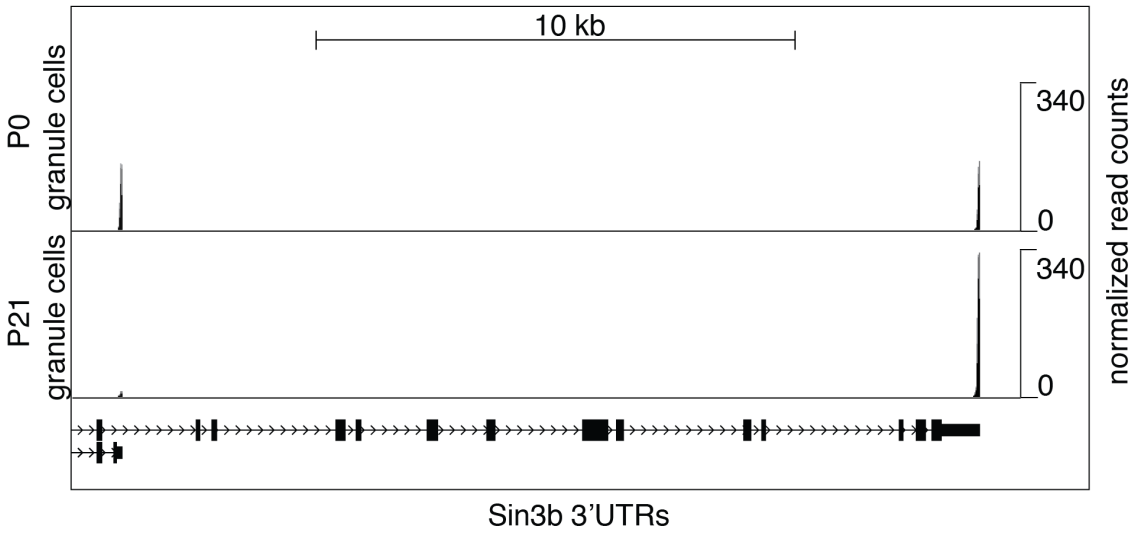


Figure 5-8 Alternative polyadenylation of the *Sin3b* transcript
cTag-PAPERCLIP clusters on the *Sin3b* 3'UTRs. The height of the clusters represents the average number of cTag-PAPERCLIP tags in a cluster (from multiple biological replicates) normalized to the total number of mapped reads. The gene structure corresponds to the RefSeq annotation.

The effect of changes in 3'UTR usage on ribosome-associated mRNA abundance

Next, we analyzed whether changes in 3'UTR isoform expression influence total gene expression during granule cell development. We used a recently published TRAP microarray data from developing granule cells (X. Zhu et al., 2016). We calculated log2 fold changes in ribosome-associated mRNA abundance between P0 and P21 granule cells. We then compared the distribution of fold changes in ribosome-associated mRNA abundance of transcripts that exhibit changes in 3'UTR usage (i.e. 3'UTRs get significantly longer or shorter) between P0 and P21 granule cells with genes that do not

significantly change in 3'UTR usage during the two time points. We did not observe any difference between the groups (Figure 5-9).

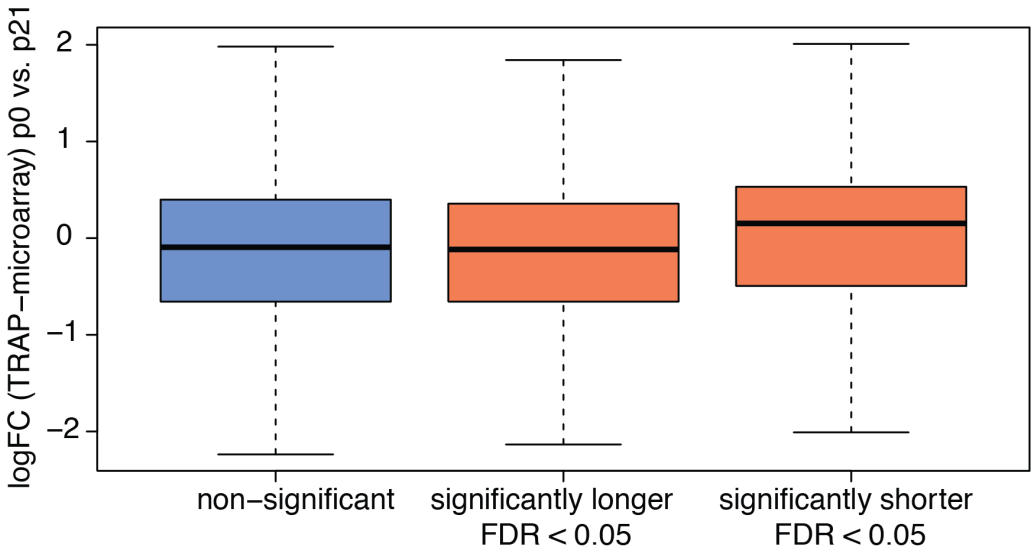


Figure 5-9 Influence of changes in alternative polyadenylation on ribosome-associated mRNA abundance between P0 and P21 granule cells.
Distribution of log2 fold-changes in ribosome-associated mRNA abundance (log2 FC) between P0 and P21 for genes that do not show significant changes in alternative polyadenylation between the two time points (blue box) and those that do (red boxes). Mann-Whitney test. P-value = NS for both comparisons (non-significant vs. significantly longer and non-significant vs. significantly shorter). Log2 FC in ribosome-associated mRNA abundance was determined using TRAP microarray data from developing cerebellar granule cells (X. Zhu et al., 2016).

The absence of a global correlation between changes in alternative 3'UTR isoform expression and total gene expression has been observed before (Gruber et al., 2014; Gupta et al., 2014). In other words, acquiring a longer 3'UTR during development is correlated to a decrease in mRNA abundance only in some cases; in other cases, it is correlated with an increase or no change in mRNA abundance.

Regulation of alternative polyadenylation during granule cell development

Next, we asked which protein factors could be involved in the lengthening of 3'UTRs during granule cell development. The expression level of polyadenylation factors as a whole has been shown to decrease during cell differentiation when 3'UTRs get longer and increase during dedifferentiation when they get shorter (Ji and Tian, 2009). Consistent with this, genes encoding polyadenylation factors tend to have binding elements for transcription factors that promote cell cycle progression in their promoters (Elkon et al., 2012; Ji and Tian, 2009).

We were interested if polyadenylation factors change in expression during granule cell development. We compiled a list of core polyadenylation factors from the literature (Shi et al., 2009). Indeed, the polyadenylation factors as a whole were slightly downregulated (median log2 FC = - 0.26), which is consistent with the above-mentioned literature. This suggests that 3'UTR lengthening during development of granule cells may at least partially be regulated by downregulation of the core polyadenylation machinery.

In addition, Elav protein has been proposed to regulate 3'UTR length during neuronal development in *Drosophila*, where the expression of Elav increases during neuronal development (Hilgers et al., 2012). We did not observe an increase in expression of mammalian homologs of Elav – Hu proteins – during granule cell development. On the contrary, we saw a

significant decrease in expression of one Hu family member, HuB, which was consistent with findings about HuB expression using *in-situ* hybridization (Okano and Darnell, 1997).

Nevertheless, since we had a Hu HITS-CLIP dataset from whole adult mouse brain available in the lab (C. Scheckel, unpublished), we decided to examine the average profile of Hu CLIP tags surrounding the proximal and distal poly(A) sites that change their relative usage during development and compare the profile to the average profile of tags around poly(A) sites that do not change their relative usage during development. We saw an increase in the density of Hu CLIP tags in proximal and distal 3'UTRs that changed their relative expression during granule cell development compared to the density of Hu CLIP tags in all 3'UTRs.

We also performed the same analysis using whole adult mouse brain Nova HITS-CLIP data (C. Marney, unpublished). Nova has been implicated in regulating alternative polyadenylation as well (Licatalosi et. al., 2008), but we saw a much smaller increase in density of Nova CLIP tags in 3'UTRs that changed their relative expression during granule cell development compared to the density of Nova CLIP tags in all 3'UTRs.

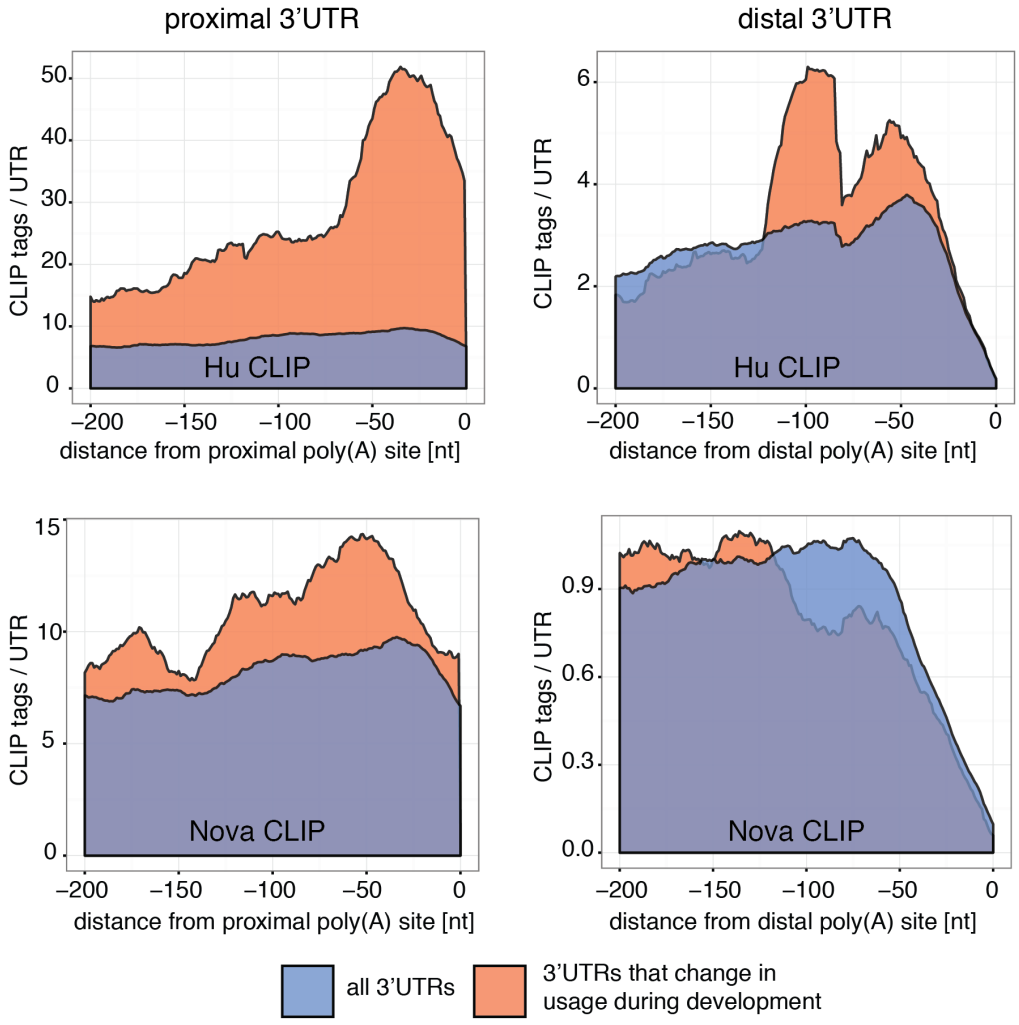


Figure 5-10 Hu and Nova binding near proximal and distal poly(A) signals

The graphs represent the average density of Hu CLIP (upper panel) and Nova CLIP (lower panel) reads per 3'UTR near poly(A) sites of proximal (short) and distal (long) 3'UTR isoforms (normalized to the abundance of each 3'UTR as determined by cTag-PAPERCLIP). The blue area represents the density of CLIP tags in all expressed 3'UTRs, and the orange area represents the density of CLIP reads in 3'UTR that show significant changes in expression during granule cell development.

Discussion

We found 599 transcripts that showed differences in alternative polyadenylation between proliferating granule cell precursors and adult granule cells, their immediate progeny. In addition, we found that transcripts that change in alternative polyadenylation during granule cell development code for proteins involved in neuron-specific processes, such as neurite outgrowth. We speculate that alternative polyadenylation is involved in establishing an adult neuron-specific gene expression program.

The finding that alternative polyadenylation particularly affects neuron-specific genes during the differentiation of granule cells is consistent with an *in-vitro* study, where the authors performed a gene ontology analysis of genes whose transcripts displayed significantly different alternative polyadenylation profiles between cultured embryonic stem cells and cultured cortical neurons. They found that genes with neuronal functions are highly enriched among the set of genes that change their alternative polyadenylation profile (e.g. 'neuron differentiation' and 'dendrite development' (Shepard et al., 2011). Our data are consistent with this study, we also found that genes with neuronal functions are enriched among genes that change in alternative polyadenylation during *in vivo* development of granule cells.

Another study profiled mRNA polyadenylation in skin stem cell lineages (L. Wang et al., 2013b). The authors of the study found 829 genes that show differential 3′-end usage between skin stem cells and their

immediate daughter cells. We also profiled cells that are closely related and found changes in alternative polyadenylation on a similar order of magnitude. However, Wang et al. did not find any enrichment for a specific functional group of genes among transcripts that changed in alternative polyadenylation. They did, however, find that components of the miRNA biogenesis pathway change in alternative polyadenylation during the development of skin cells. We checked if there was a change in alternative polyadenylation of the miRNA processing machinery during the development of granule cells. The only transcript that showed changes was *Dicer* (data not shown). Granule cell precursors express both the short and long 3'UTR isoform of *Dicer*, whereas adult granule cells express mostly the long isoform. This is inconsistent with the Wang et. al study, where they observed exactly the opposite: that the *Dicer* 3'UTR gets dramatically shorter during skin cell development. This suggests that 3'UTR profiles are not necessary regulated in the same way during the development of all cell types and underscores the need to study each cell type separately.

We have also shown the potential of Hu proteins to regulate the transition from short to long 3'UTRs during granule cell development. The increase of Hu CLIP tags from adult brain close to the proximal poly(A) site, (which used less in adult brain) is consistent with the finding from *in vitro* studies that Hu proteins can block the cleavage and polyadenylation of mRNAs when they bind in the vicinity of a poly(A) signal (H. Zhu et al., 2007).

However, it remains to be explained how this can be achieved without a change in Hu expression during granule cell development. One possibility is that there are additional factors involved.

Chapter 6 I Studying the function of *Memo1* alternative polyadenylation

Introduction

In the previous Chapter, I described the changes in alternative polyadenylation that occur during the development of granule cells. Next, we were interested in the functional consequences of these changes in 3'UTR isoform expression during granule cell development.

A switch in alternative polyadenylation of a specific gene could happen either because one isoform is required for the proper function of the adult neuron, whereas the other is required for the proper function of the neuronal precursor cell. Alternatively, the switch could regulate the proper and timely development of granule cells.

One gene that changed in alternative polyadenylation during granule cell development and which might regulate granule cell development was *Memo1. Memo1* was discovered as an intracellular binding partner of Her2/ErbB2 (Marone et al., 2004). While *Memo1* has no known function in neuronal development, it has been implicated in cell motility (G. MacDonald et al., 2014) and proliferation (Sorokin and Chen, 2012) of breast cancer cells. Given that cell proliferation and motility are also two processes that occur during the span of granule cell development we decided to study Memo1 in more detail.

Results

Memo1 changes 3'UTR isoform expression during granule cell development

Memo1 was among the top 20% of genes by fold change in 3'UTR isoform expression between granule cell precursors and adult granule neurons. Granule cell precursors express only a short 3'UTR isoform of *Memo1*, while adult granule cells express both the short and a longer isoform (Figure 6-1). The longer isoform is not annotated by RefSeq; however, we confirmed its presence by analyzing published TRAP-seq data (Mellén et al., 2012), which showed continuous read coverage between the annotated short 3'UTR end and the distal cTag-PAPERCLIP cluster 14 kb downstream (Figure 6-2).

Memo1 has no known function in neuronal development, but it has been previously implicated in cell motility (G. MacDonald et al., 2014) and proliferation (Sorokin and Chen, 2012), two processes that occur during the time points spanning our CLIP study. The expression of *Memo1* was downregulated during granule cell development (Figure 6-3) and its distal and proximal poly(A) signals are highly conserved. By analyzing published RNA sequencing data from human brains (Jaffe et al., 2015), we confirmed that both *Memo1* 3'UTR isoforms are expressed in the human brain.

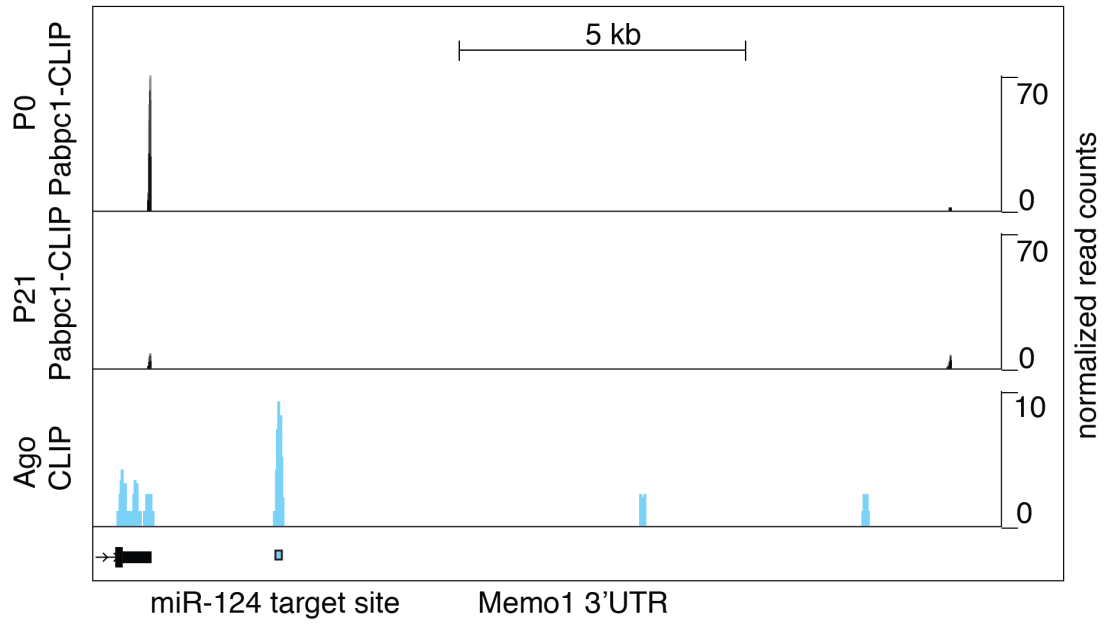


Figure 6-1 Alternative polyadenylation of the *Memo1* transcript

cTag-PAPERCLIP and Ago CLIP reads on *Memo1* 3'UTR in P0 and P21 granule cells. The location of miR-124 target site is indicated by the blue bar. RefSeq annotation of the *Memo1* 3'UTR is indicated in the bottom left corner.

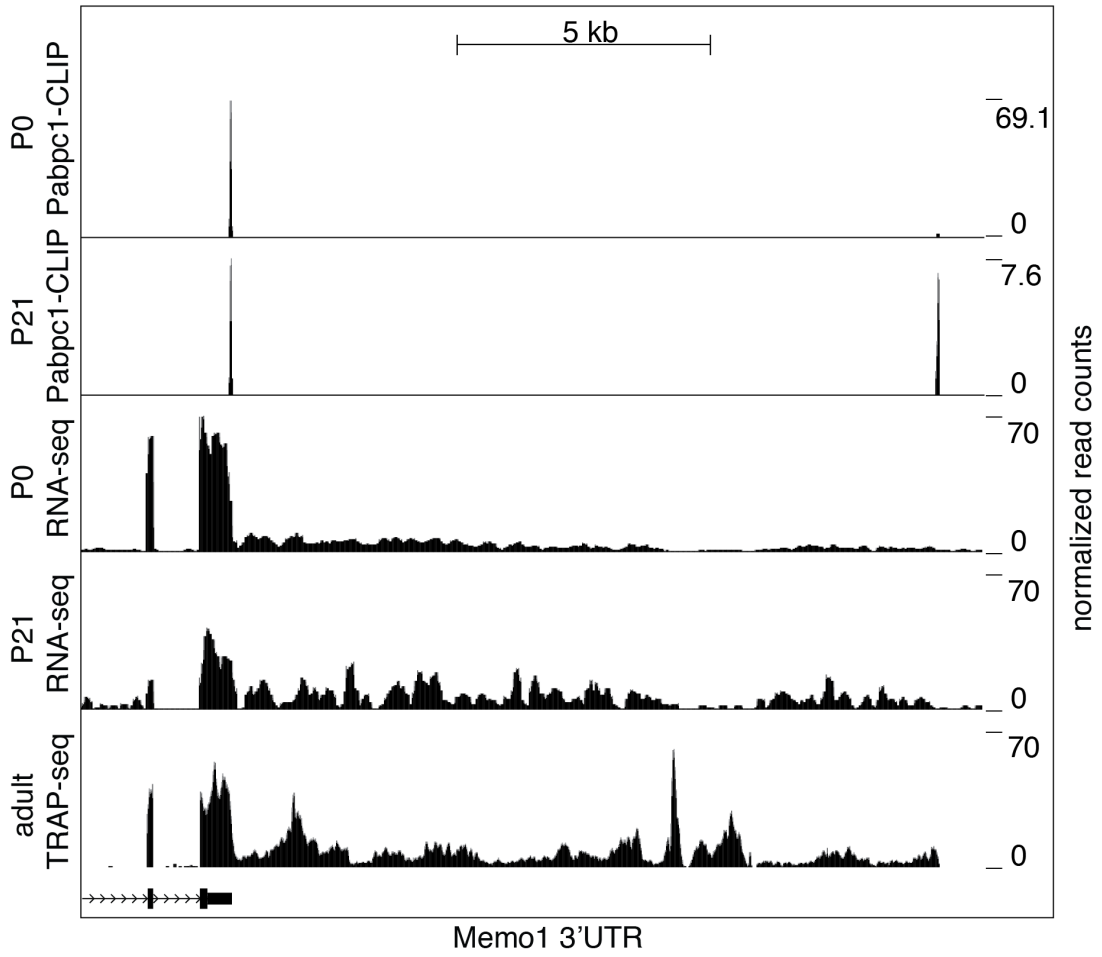


Figure 6-2 Alternative polyadenylation of the *Memo1* transcript – RNA-seq validation

The first two tracks from above represent cTag-PAPERCLIP data from P0 and P21 granule cells. The tracks below cTag-PAPERCLIP represent RNA-sequencing data from FACS sorted P0 and P21 granule cells and TRAP-sequencing data from adult granule cells (Mellén et al., 2012). In the lower left corner, RefSeq annotation for *Memo1* is shown.

miR-124 regulates *Memo1* expression during granule cell development

To explore the molecular function of the different Memo1 3'UTR isoforms in more detail, we analyzed the sequence of the two *Memo1* 3'UTRs. Using RegRNA software (Chang et al., 2013), we observed a highly conserved miR-124 target site on the long 3'UTR isoform that was absent in the short isoform (Figure 6-1). To confirm that this site was indeed bound by miR-124 in mouse brain, we examined Argonaute CLIP data from mouse brain previously generated by our laboratory (Moore et al., 2015). This analysis revealed that Argonaute is strongly bound to the miR-124 site on the long *Memo1* 3'UTR isoform in cortical neurons (Figure 6-1).

miR-124, a neuron-specific miRNA, has been shown to be upregulated during differentiation of neural stem cells in the subventricular zone and implicated in the regulation of neuronal differentiation (Cheng et al., 2009; Makeyev et al., 2007). Thus, we examined whether miR-124 expression changes during granule cell differentiation. We found a significant increase in miR-124 expression between granule cell precursors purified from P7 mice and granule cells that were differentiated *in vitro* for 2 weeks, when they morphologically resemble adult granule cells (Figure 6-4).

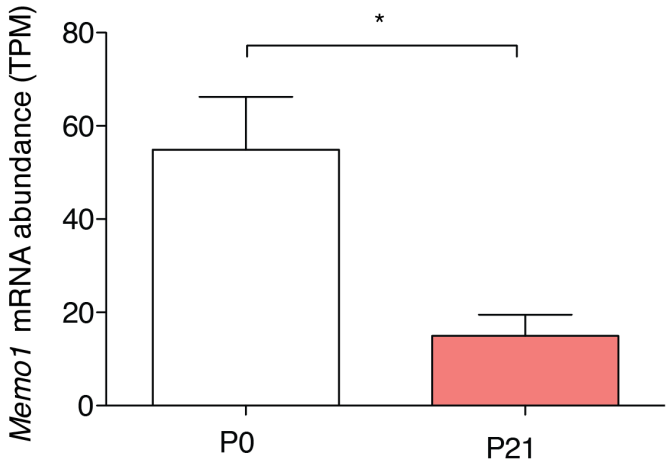


Figure 6-3 Memo1 mRNA abundance in granule cells changes during their development

Abundance of *Memo1* transcript as determined by cTag-PAPERCLIP. The y-axis represents the number of total unique reads per gene normalized to sequencing depth (RPM – reads per million). t-test. P-value < 0.05 Error bars: standard error.

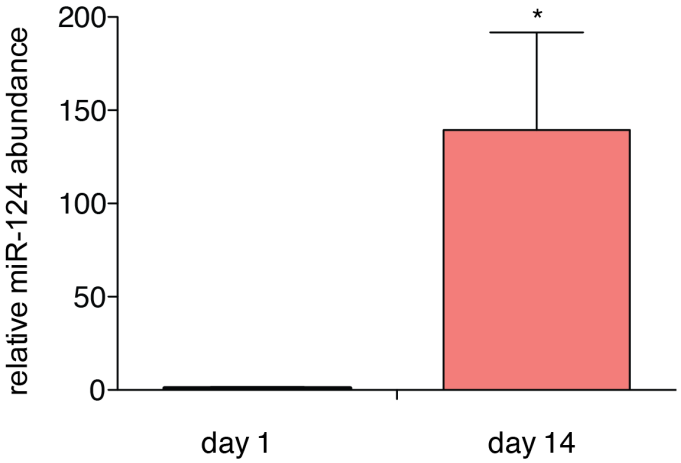


Figure 6-4 miR-124 abundance in granule cell culture

Relative abundance of miR-124 in purified granule cell precursors (day 1), after 14 days in culture (day 14, t-test, P-value < 0.05). Error bars: standard error.

We also compared miR-124 expression in purified granule cell precursors with miR-124 expression in the adult cerebellum, which is composed mostly of granule cells. Again, we found a significant increase in miR-124 expression (Figure 6-5). Thus, just as the expression of *Memo1* decreases, miR-124 expression is upregulated – pointing to a potential role for miR-124 in downregulating *Memo1* expression during granule cell development.

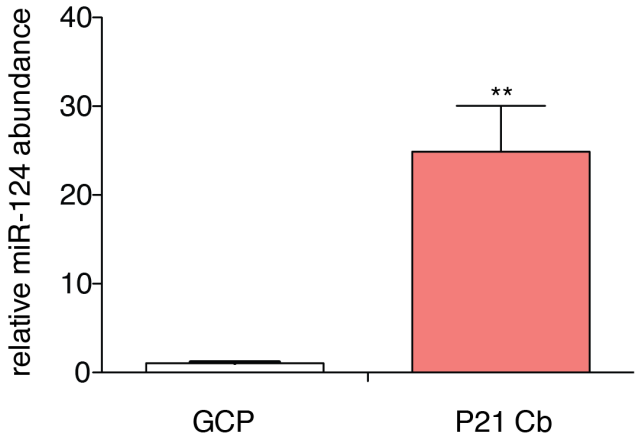


Figure 6-5 miR-124 abundance in granule cell precursors vs. P21 whole cerebellum

Relative abundance of miR-124 in granule cell precursors (GCP) purified from P6 mice and in the adult cerebellum (P21 Cb, t-test, P-value < 0.05). Error bars: standard error.

To investigate the effect of miR-124 on the abundance and stability of *Memo1* 3'UTR long isoform, we prevented the action of miR-124 using an antisense oligonucleotide (antagomir) against miR-124 in cultured primary granule cells. The *Memo1* long isoform was significantly upregulated after

transfection of miR-124 antagomir compared to scrambled control (Figure 6-6).

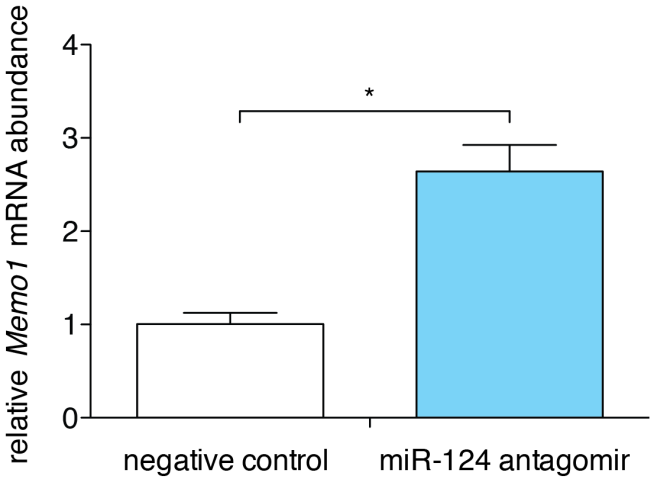


Figure 6-6 miR-124 inhibition affects *Memo1* abundance
Relative abundance of long *Memo1* isoform in cultured primary granule cells after treatment with scrambled control and miR-124 antagomir. t-test. P-value < 0.05. Error bars: standard error.

To further validate the regulation of *Memo1* expression by miR-124, we blocked the interaction between miR-124 and *Memo1* using an antisense oligonucleotide binding to the miR-124 target site on *Memo1* (target site blocker). As expected, the *Memo1* long isoform was upregulated after transfection of this target site blocker compared to a scrambled control (Figure 6-7).

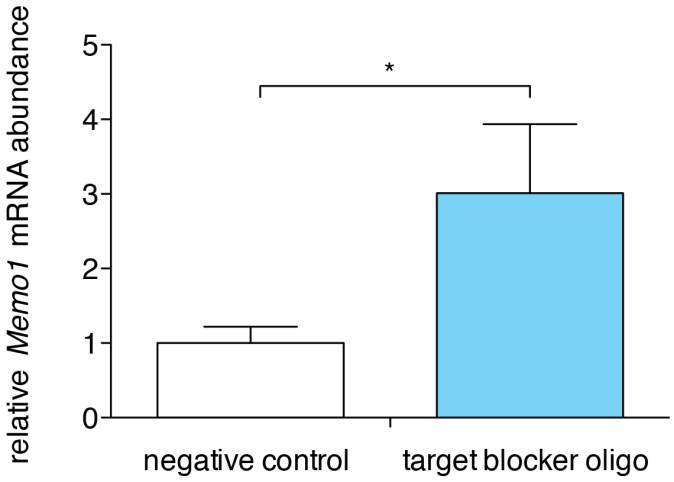


Figure 6-7 Inhibition of miR-124 target site on *Memo1* affects *Memo1* abundance

Relative abundance of long *Memo1* isoform in cultured primary granule cells after treatment with scrambled control and *Memo1* miR-124 target site blocker oligonucleotide. t-test. P-value < 0.05. Error bars: standard error.

Memo1 translation

To explore whether the two *Memo1* 3'UTR isoforms are translated differently, we fractionated mRNA from whole mouse cerebella (which are mainly composed of granule cells) using sucrose density centrifugation, which separates mRNAs based on the number of ribosomes associated with them. We then performed qPCR using a primer pair that annealed to the coding sequence of both *Memo1* isoforms and a primer pair that annealed only to the long 3'UTR isoform. We found that both total *Memo1* mRNA and the long 3'UTR isoform exhibited similar distribution on the sucrose gradient, suggesting that both isoforms are translated at a similar rate (Figure 6-8).

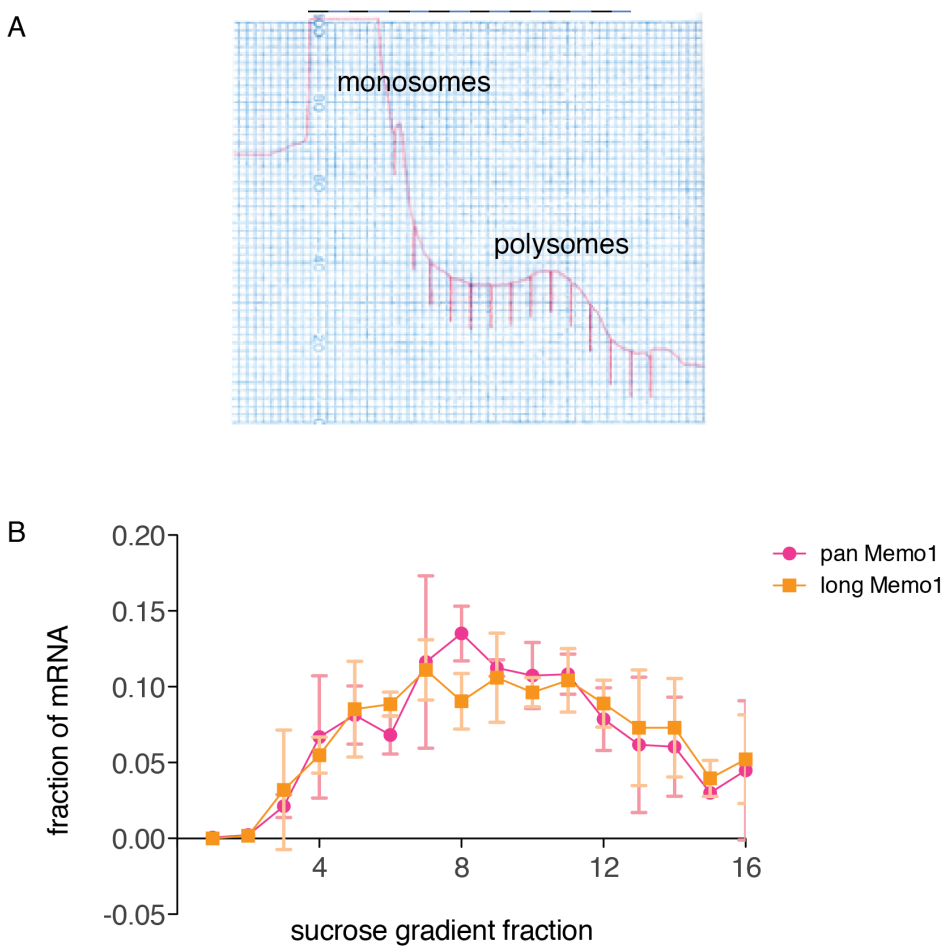


Figure 6-8 Ribosomal association of *Memo1* isoforms

(A) Polysome profile used to separate mRNAs into 16 fractions based on the number of bound ribosomes. The location of monosomes and polysomes on the gradient is indicated. **(B)** Fraction of the total *Memo1* mRNA abundance in each ribosome fraction**.** Pink dots represent abundance measured by qPCR with primers that align to an exon common to both short and long *Memo1* isoform. Orange dots represent abundance of *Memo1* measured by qPCR with primers that align only to the long isoform of *Memo1.* Error bars: Standard error.

Memo1 localization

To see if the two *Memo1* isoforms are localized to different parts of the cell, we performed single molecule RNA FISH on cultured primary granule

cells using a probe set that bound to coding sequence of both isoforms and a probe set that bound only to the long isoform. We observed that both probe sets were localized in the cell body, suggesting that the long and short *Memo1* isoform are localized to the same part of the cell (Figure 6-9).

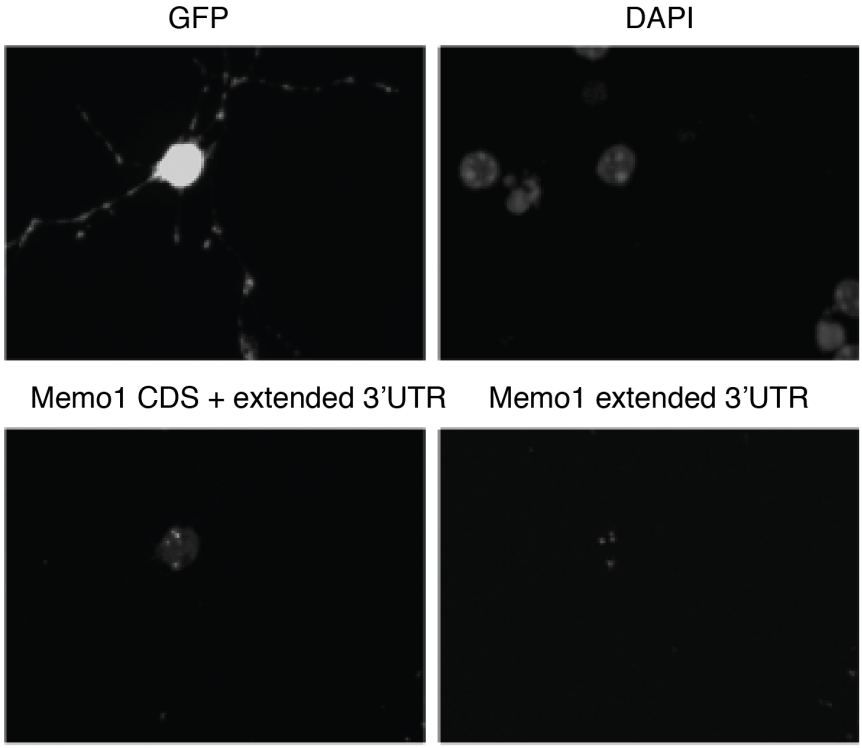


Figure 6-9 *Memo1* localization

Memo1 localization in cultured primary granule cells determined by single-molecule in-situ hybridization. Two in-situ hybridization probe sets were used (bottom left: probe sets binding to a region common to both isoforms, bottom right: probe set binding only to the extended 3'UTR region). Upper left (GFP) and upper right (DAPI) pictures show the location of the cell body and processes and the nucleus, respectively.

Discussion

Taken together, our data indicates that the change in *Memo1* 3'UTR usage during granule cell development confers a difference in the stability of the *Memo1* mRNA; however, *Memo1* translation and its subcellular localization remain the same. This change in mRNA stability provides a mechanism for *Memo1* downregulation during development.

Several reports have found that extended 3'UTR isoforms contain more miRNA target sites (Hwang et al., 2016; Ji et al., 2009; Miura et al., 2013; Nam et al., 2014; Sandberg et al., 2008), which can lead to differential regulation between the short and long 3'UTR isoform. However, to our knowledge, our study is the first to show direct evidence for an interplay between a specific miRNA and alternative polyadenylation in a specific cell type during a physiological process.

The functional consequences of alternative polyadenylation-mediated *Memo1* downregulation will need to be further investigated. For this purpose, we will use the miR-124 target protector that specifically prevents the miRNA from binding the *Memo1* transcript. We have shown in *in vitro* experiments that the relative abundance of the long 3'UTR isoform is upregulated upon target protector transfection. Upregulation of *Memo1* expression might regulate cell proliferation; it has been shown that overexpression of *Memo1* affects the proliferation of breast cells *in vitro* (Sorokin and Chen, 2012)

Another important question concerns the phenotype of granule cells upon *Memo1* knockdown. I have some preliminary data showing that knocking down *Memo1* in primary granule cells leads to cell death (data not shown).

Another intriguing question is why there is a switch in the polyadenylation of *Memo1* at all. Namely, *Memo1* downregulation could be achieved even if the long 3'UTR isoform was constitutively expressed, given that miR-124 is upregulated during development. One possibility is that the precise timing of miR-124 upregulation is not the same as the timing of the change in alternative polyadenylation. Alternative polyadenylation might provide an opportunity to fine tune the timing of *Memo1* expression; instead of being dependent on one variable (i.e. the abundance of miR-124), *Memo1* expression would be dependent on two variables (i.e. miR-124 concentration and the ratio between the short and long *Memo1* isoforms).

Chapter 7 | General discussion

Summary

In this study, we investigated 3'UTR isoform expression in specific cell types. Using cTag-PAPERCLIP, a novel method we developed, we built an atlas of 3'UTR ends in Purkinje and granule cells in the mouse cerebellum and during granule cell differentiation *in vivo*.

Previous studies have shown that alternative polyadenylation can differ between tissues, both in humans (Lianoglou et al., 2013; Haibo Zhang et al., 2005), mice (Miura et al., 2013) and *Drosophila* (Smibert et al., 2012). Furthermore, in mice, muscle stem cells in different regions of the body express different 3'UTR isoforms of the *Pax3* gene (Boutet et al., 2012), raising an intriguing question about the extent of the differences in alternative polyadenylation among related cell types within a tissue. We addressed this question in the nervous system and found that, indeed, there is a marked diversity of alternative polyadenylation among Purkinje and granule cells, two related but functionally and morphologically distinct cells. Given that neurons are exceptionally diverse (Macosko et al., 2015), we hypothesized that alternative polyadenylation is one of the mechanisms that plays a role in determining their identity. Indeed, genes that showed differences in alternative polyadenylation between Purkinje and granule cells are involved in cellular functions like 'branching of neurites' and 'synaptic transmission' which

is in line with morphological and physiological differences between the two neuronal types.

It has been reported that the lengthening of 3' UTRs is associated with development (Ji et al., 2009; Shepard et al., 2011; Ulitsky et al., 2012; L. Wang et al., 2013a). In our study, we investigated the changes in 3'UTR usage during the *in vivo* differentiation of a specific neuronal type – cerebellar granule cell. Indeed, we observed lengthening of 3'UTRs during its development. Interestingly, we found an enrichment of genes involved in neuronal development among genes that showed differences in 3'UTR usage between granule cell precursors and differentiated granule cells.

The changes in 3'UTR usage could affect mRNA abundance, translation, as well as localization of protein and mRNA (Mayr, 2016). When analyzed as a whole, we did not observe a global correlation between changes in alternative 3'UTR isoform expression and changes the abundance of ribosome-associated mRNA. This confirmed the previous observation that, globally, there is no correlation (Gruber et al., 2014; Gupta et al., 2014) or only a weak correlation (Masamha et al., 2014; Spies et al., 2013) between mRNA stability and 3'UTR length. However, the lack of the global correlation does not mean that 3'UTR changes we observed are not functionally relevant. Indeed, we found that changes in alternative polyadenylation of the *Memo1* transcript were associated with miR-124-mediated downregulation.

Future applications of cTag-PAPERCLIP

We demonstrated the utility of cTag-PAPERCLIP for transcriptome profiling and 3'UTR end mapping in specific cell types. In the future, it would be especially interesting to profile the transcriptome and alternative polyadenylation in various mouse models of diseases that affect specific cells. For example, most neurodegenerative diseases do not affect all cells in the nervous system, but rather affect specific groups of cells, whose death or dysfunction leads to the phenotypic symptoms that define the disease. For example, amyotrophic lateral sclerosis (ALS) mostly affects upper and lower motor neurons (Boillée et al., 2006), Spinocerebellar ataxia type 1 (SCA1) affects Purkinje neurons in the cerebellum (Orr, 2000) and Huntington disease affects medium spiny neurons in the striatum (Han et al., 2010). Using cTag-PAPERCLIP to compare transcriptomes of the disease-affected cells between normal mice, and mice that mimic the disease could uncover disease-specific changes in transcription in specific cells.

In the case of ALS, it would not only be interesting to profile transcriptome changes during disease progression but also to look at changes in alternative polyadenylation, given that ALS mutations have been implicated in changes in alternative polyadenylation.

For example, one of the genes mutated in ALS is Fused in sarcoma (Fus). Fus binding was shown to be frequently clustered around an alternative polyadenylation site of nascent RNA. Moreover, it was shown that Fus stalls

Pol II and prematurely terminates transcription. When an alternative polyadenylation site was located upstream of an Fus binding cluster, Fus enhanced polyadenylation by recruiting CPSF160 and upregulated the alternative short transcript. In contrast, when an alternative polyadenylation site was located downstream from an Fus binding cluster, polyadenylation was not activated and the Pol II - suppressing effect of Fus led to down-regulation of the shorter transcript (Masuda et al., 2015).

Authors of another study performed RNA-seq on brains from patients who died from sporadic ALS with no known genetic cause and ALS caused by an expansion of the hexanucleotide GGGGCC in the gene *C9ORF72*, which is the most frequent known cause of both familial and sporadic ALS. When they compared the data from both groups of ALS patients with healthy controls, they observed numerous changes in alternative polyadenylation. Most notably, they found a global shortening of 3'UTR isoforms in the cerebellum of patients with a hexanucleotide expansion in *C9ORF72* compared to healthy controls (Prudencio et al., 2015).

Additional open questions about changes in alternative polyadenylation during granule cell development

We uncovered many transcripts that change alternative polyadenylation during the development of granule cells. Interestingly, when we looked closely at the changes in alternative polyadenylation during development, we observed that many genes switch from expressing both a short and long 3'UTR isoform in the neuronal precursors to expressing just the long 3'UTR isoform in the adult neurons. The median number of (sequencing depth normalized) CLIP reads per proximal (short) 3'UTR isoform decreases from 21.7 at P0 to 7.8 at P21, while the median number of CLIP tags per distal (long) isoform increases from 24.3 to 37.7 from P0 to P21. This suggests that the lengthening of 3'UTRs arises both from a decrease in the expression of the short isoform as well as an increase in the expression of the long isoform. In other words, an average transcript goes from expressing both 3'UTR isoforms in the granule cell progenitors to expressing mostly the long 3'UTR isoform in the adult neuron.

It is interesting to speculate how the observed patterns might arise. At P0, multiple transcript variants can be expressed at the same time in each cell in the population of granule cell progenitors. Alternatively, the observed pattern could arise because half of the progenitor cell population expresses the long isoform whereas the other half expresses the short isoform. One potential source of heterogeneity is that different cells within our population of

granule cell progenitors are likely in different phases of the cell cycle. Indeed, a recent publication identified around 1300 genes with cell-cycle dependent alternative splicing changes (Dominguez et al., 2016).

The mechanism of change in 3' UTR length in granule cell population during development will differ based on which of the above models is true. Mueller et. al. call the first model, in which all progenitor cells are capable of expressing both transcripts, the 'conversion' model, and the result of the developmental transition is that each cell expresses a higher proportion of the long transcript. In the second model, which they call the 'selection' model, each cell is capable of expressing only one isoform and the global increase in 3'UTR length in the population occur when there is a change in the relative abundance of those cells expressing the long vs. the short 3'UTR isoforms (Mueller et al., 2013).

With recent advances in single-cell RNA-sequencing (Shalek et al., 2013)., it should be possible to explore this question in the future. Since we have already mapped the 3'UTR ends of polyadenylated RNAs in granule cells, RNA-sequencing could be used to quantify the expression of each isoform in single cells given this "map" of possible 3' ends.

Future experiments to elucidate the role of *Memo1* in granule cell development

The role of *Memo1* in the development of cerebellar granule cells remains unknown. One possibility is that the MEMO1 protein mediates granule cell proliferation through Insulin-like growth factor 1 (IGF-1). For example, MEMO1 has been shown to bind Insulin receptor substrate (IRS), an IGF-1 receptor binding protein, and to positively regulate breast cell proliferation (Sorokin and Chen, 2012). In addition, IGF-1 has been shown to positively regulate the proliferation of granule cell precursors (Gao et al., 1991). We propose that the MEMO1 protein is a mediator of IGF-1 signaling in granule cell precursors (Figure 7-1). Knocking it down should therefore affect the proliferation of granule cell precursors. In addition, we propose that the switch from the short to the long isoform of the *Memo1* transcript may play a role in helping granule cell precursors to exit the cell cycle, given that IGF-1 signaling has been shown to promote cell cycle progression from G1 to S phase (Mairet-Coello et al., 2009). Thus, a decrease in the levels of the MEMO1 protein through IGF-1 signaling could potentially increase the probability of cell cycle exit, and help in transitioning granule cell precursors to non-dividing mature granule cells.

We found that the long 3'UTR isoform of *Memo1* is regulated by miR-124. To specifically address the role for miR-124-mediated downregulation of *Memo1* in the regulation of granule cell proliferation, we plan to block the

interaction between miR-124 and the long 3'UTR of *Memo1* using a miR-124 target protector in primary cultured granule cells and in cerebellar slices, and then measure cell proliferation. We hypothesize that an increase in *Memo1* levels will lead to a change in the proliferation of granule cell precursors.

In addition, it would be interesting to measure the abundance and subcellular localization of the MEMO1 protein in developing granule cells using immunofluorescence. However, we were not able to perform these measurements so far, because we could not find antibodies that worked well for this purpose. Future studies with better antibodies will help to test whether *Memo1* protein levels decrease during granule cell differentiation, in lock-step with the decrease that we observe in *Memo1* mRNA levels.

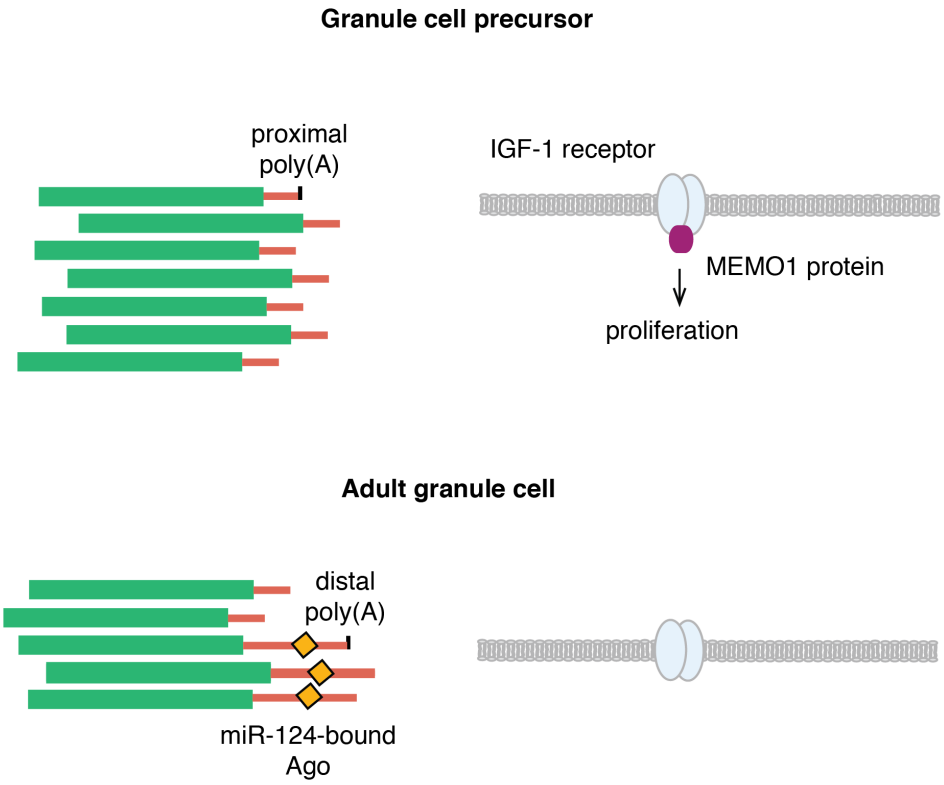


Figure 7-1 A model for the regulation of *Memo1* expression during granule cell development

In granule cell progenitors, only the short 3'UTR isoform of *Memo1* is expressed, which leads to high protein expression and binding to IRS, an IGF-1 receptor binding protein (in the figure, both IRS and MEMO1 are shown as a purple oval shape). In adult granule neurons, both long and short 3'UTR isoform are expressed. The long 3'UTR isoform is downregulated via miR-124, which results in downregulation of protein expression. Consequently, I hypothesize that *Memo1* doesn't mediate IGF-1 signaling in the adult neurons.

Appendix I Supplementary tables

Table 1 I Transcripts that show significant differences in alternative polyadenylation between granule and Purkinje cells

	Gene name	Number of 3'UTR isoforms	q-value
1	Gabbr1	2	1.35E-93
2	Trio	3	9.98E-26
3	Asph	3	7.45E-19
4	Phf14	2	1.24E-17
5	Capn15	2	2.97E-13
6	Gls	2	3.26E-13
7	Il16	3	3.54E-13
8	Tm2d1	2	1.52E-12
9	Ncam1	4	1.82E-12
10	Pld3	3	3.85E-12
11	Fbxw2	3	5.34E-12
12	Amph	2	6.51E-12
13	Robo2	2	1.39E-11
14	Cmya5	2	2.86E-11
15	Dach1	2	9.57E-11
16	Vps26b	3	1.20E-10
17	Cacna2d3	2	1.38E-10
18	Arhgap15	2	2.22E-10
19	Papola	3	3.42E-10
20	Chd9	2	4.14E-10
21	Ktn1	2	4.35E-10
22	Luc7l3	2	5.02E-10
23	Fam168a	2	8.51E-10
24	Vezt	2	1.01E-09
25	Fhod3	2	2.01E-09
26	Camta1	2	2.07E-09
27	Fut2	2	2.43E-09
28	Cdk16	2	2.97E-09
29	Gigyf2	3	3.64E-09
30	Lrrk2	3	8.28E-09
31	Wdr70	2	8.88E-09
32	Ttc14	4	1.59E-08
33	Vps13a	2	1.95E-08
34	Khsrp	2	2.41E-08
35	Kidins220	2	4.25E-08
36	Slc48a1	2	5.30E-08
37	Ptpn2	2	8.89E-08

38	Jarid2	3	9.48E-08
39	Malat1	2	1.63E-07
40	Gm37013	5	2.66E-07
41	Sap30l	2	1.01E-06
42	Cpsf2	3	1.04E-06
43	Slc39a10	2	1.13E-06
44	Sdc1	3	1.13E-06
45	6030419C18Rik	2	3.05E-06
46	Stxbp5l	2	3.19E-06
47	Ap5m1	5	3.32E-06
48	H2-D1	2	3.33E-06
49	Atp2a2	2	4.22E-06
50	Copg1	3	4.22E-06
51	Gnal	2	4.47E-06
52	Slc22a22	2	4.47E-06
53	Actr10	2	4.49E-06
54	Ctage5	2	4.82E-06
55	Ago2	2	5.22E-06
56	Tmem57	2	5.22E-06
57	Zfp609	2	5.73E-06
58	Sept11	3	5.73E-06
59	Map4	3	6.16E-06
60	Vmp1	2	6.59E-06
61	Shank1	2	6.92E-06
62	Ccdc92	2	1.04E-05
63	Slc11a2	3	1.08E-05
64	Slc39a9	2	1.12E-05
65	Katnal1	3	1.12E-05
66	Jade1	2	1.25E-05
67	Sos1	3	1.45E-05
68	Acsf3	3	1.48E-05
69	Gtpbp3	2	1.51E-05
70	Ppm1a	3	2.29E-05
71	Gtf2h1	2	2.40E-05
72	Mta3	2	2.53E-05
73	Napa	2	2.53E-05
74	Olfm1	2	3.46E-05
75	Mdh2	2	3.47E-05
76	Tardbp	3	3.96E-05
77	Zdhhc8	2	4.29E-05
78	Ccni	4	4.39E-05
79	Hs3st1	2	4.52E-05
80	Chid1	3	5.16E-05

81	Cep350	2	5.79E-05
82	Ralgapb	3	5.79E-05
83	Tmx3	2	7.65E-05
84	Tln1	2	7.78E-05
85	Robo1	2	8.42E-05
86	Wdr45b	2	8.42E-05
87	Dync1li2	2	9.38E-05
88	Slc25a46	3	9.94E-05
89	Rbbp6	3	0.000101847
90	2610507B11Rik	2	0.000115185
91	Cib2	2	0.000117262
92	Rsrc2	2	0.000118273
93	Mbd2	2	0.000120064
94	Pclo	2	0.000120064
95	Tmem248	2	0.000129519
96	Ttc39b	3	0.000131917
97	Ube2o	2	0.000133429
98	Mapre2	2	0.000136239
99	Narf	2	0.000150551
100	Orai2	2	0.000150551
101	Eif1ad	2	0.000163675
102	Pmm1	2	0.000168768
103	Map6	2	0.000175253
104	Ywhae	2	0.000175253
105	Ercc4	2	0.000178719
106	Sirpa	2	0.000179148
107	Rnf115	3	0.00019
108	3632451O06Rik	4	0.000194795
109	Dcun1d2	2	0.000194829
110	Fkbp4	2	0.000194829
111	Tmf1	5	0.000194829
112	Rbm33	3	0.00019784
113	Afg3l2	2	0.000206052
114	Clstn2	2	0.000206052
115	Rundc3a	2	0.000206052
116	Tmem38a	3	0.000210533
117	Sema6a	2	0.000217807
118	5031439G07Rik	2	0.000225306
119	Itm2b	2	0.000225306
120	March5	4	0.000234212
121	Ptar1	2	0.000255406
122	Tfdp2	3	0.000260572
123	Ube3a	4	0.000265077

124	Cct4	2	0.000309272
125	Nfic	2	0.000312149
126	Rufy2	2	0.000312149
127	Mapk1	3	0.000337012
128	Zfp438	2	0.000339047
129	Ccdc85b	2	0.000349688
130	Matr3	4	0.000388223
131	Ntrk3	2	0.000389744
132	Tbpl1	2	0.000389744
133	Ttc19	4	0.000409948
134	Sprn	2	0.000440793
135	Pik3ca	2	0.000450729
136	2410131K14Rik	2	0.000476495
137	Rfx3	2	0.000479401
138	Trak1	2	0.000501857
139	Ppp1r7	3	0.000503364
140	Nudcd1	2	0.000510349
141	Tuba4a	2	0.000510349
142	Il6st	2	0.000512091
143	Mrps35	2	0.00052822
144	Unc79	2	0.00054019
145	9330182L06Rik	4	0.000561511
146	Zbtb1	2	0.00061132
147	Plcb1	3	0.000627632
148	Agpat3	3	0.000644818
149	Pnisr	2	0.000684443
150	Rab21	5	0.000684443
151	Zfp938	3	0.000698721
152	Azi2	4	0.000736718
153	Zbtb46	2	0.000760067
154	Ccser1	3	0.000809825
155	Wdr18	3	0.000809825
156	Fads6	2	0.000862335
157	Kif26b	3	0.00086325
158	Appl1	3	0.000888787
159	Dnajc5	3	0.00092992
160	Atp13a3	3	0.000961157
161	Nipa1	2	0.000973212
162	B4galnt1	2	0.000976795
163	Opn3	2	0.001006734
164	Lamp2	3	0.001035417
165	Slc44a1	3	0.00105719
166	Nemf	2	0.001058816

167	Paqr9	2	0.001087563
168	Tax1bp1	2	0.001126312
169	Ttl	2	0.001177731
170	Ttll11	2	0.001177731
171	Eif4g2	2	0.001187375
172	Mbp	2	0.001187375
173	B3gat2	2	0.001202399
174	Mllt10	2	0.001202399
175	Pcif1	3	0.001202399
176	Sri	3	0.001202399
177	Cdc73	5	0.001202399
178	Ywhaq	2	0.001249591
179	Smim10l1	2	0.00143761
180	Ppp1r12b	3	0.001458741
181	Akt3	3	0.001483576
182	Inpp4a	2	0.001499824
183	Slc9a6	2	0.001499824
184	Tmem175	3	0.001548598
185	Rab11fip3	2	0.001554367
186	Zkscan3	2	0.001716173
187	Enpp4	2	0.001727638
188	Gnao1	2	0.001728235
189	Orc3	5	0.001728235
190	Ppp1r3f	4	0.001785304
191	Ndufs4	2	0.001794473
192	Ttc7	2	0.001823203
193	Dnajb2	2	0.001836333
194	Irf2bpl	2	0.001946469
195	Pigk	2	0.002107084
196	Nudcd3	2	0.002199492
197	Arfgef2	2	0.002276896
198	Dram2	3	0.002294422
199	Sik2	3	0.002354839
200	1110008P14Rik	2	0.00237085
201	3110043O21Rik	2	0.00237085
202	Cadm1	2	0.002414412
203	Fgfr1	2	0.002414412
204	Cog5	3	0.002414412
205	Pvrl3	2	0.002512659
206	Nup155	3	0.002513774
207	Phf20l1	3	0.002537994
208	Ttbk1	3	0.00255401
209	Fam120b	2	0.002587935

210	Psma2	2	0.002681732
211	Ccar1	3	0.002681732
212	Kcnj11	2	0.002712832
213	Trpc1	2	0.002731807
214	Sfpq	2	0.002734898
215	Emc3	2	0.002791769
216	Luc7l	2	0.002891239
217	Mkrn2	2	0.002899503
218	Tmem70	2	0.002899503
219	Vwa8	3	0.003003419
220	Ctnnbl1	3	0.003004645
221	Pbx2	3	0.003059269
222	Rab22a	2	0.003315287
223	Smarcc1	2	0.003452781
224	Calm1	2	0.003493118
225	Pvalb	2	0.003567524
226	Ppp2r3d	3	0.003605724
227	Usp15	2	0.003788459
228	Clcn3	2	0.003857266
229	Dtnb	2	0.003857266
230	Tceb3	2	0.003857266
231	Ythdf2	2	0.003959581
232	Dcc	2	0.004089991
233	Sh3gl2	2	0.004099157
234	Alg9	3	0.004121685
235	Acap3	2	0.004213891
236	Snx12	3	0.0042922
237	Raly	2	0.004295659
238	Synj2bp-cox16	3	0.004437844
239	Uqcrc2	2	0.004490743
240	Cdk13	3	0.004490743
241	Iffo1	2	0.004510938
242	Tpgs2	3	0.004518413
243	Wdr37	3	0.004518413
244	Pfkfb2	2	0.004540735
245	Ndufa9	2	0.004636918
246	Dohh	2	0.004700873
247	Eprs	2	0.004821828
248	Mfap1b	2	0.004877797
249	Slc25a16	2	0.004877797
250	Vps4a	2	0.004877929
251	Cryzl1	3	0.004877929
252	Pmpca	3	0.004877929

253	Pkib	3	0.004990859
254	Ivns1abp	3	0.005092309
255	Cln8	2	0.005167219
256	Dnajc30	2	0.005284213
257	Rims4	2	0.005345601
258	Prep	2	0.005379478
259	Slc7a6	3	0.005432843
260	Usp46	3	0.005605134
261	Slc25a39	2	0.005622362
262	Ccpg1	2	0.005708122
263	Inpp1	2	0.006006695
264	Tbc1d19	3	0.006027225
265	Ddx46	3	0.006155164
266	Pcmt1	3	0.006155164
267	Vrk1	2	0.006200393
268	Cox16	3	0.006235395
269	Appbp2	2	0.006611883
270	Gdf11	4	0.006738337
271	Timm9	2	0.006804201
272	Derl1	2	0.006841229
273	Dennd5b	2	0.006866227
274	Rbck1	2	0.006908499
275	Synrg	2	0.007075024
276	Trim9	2	0.007092295
277	Plxdc2	5	0.007194523
278	Snx29	2	0.007429402
279	Slc22a15	5	0.007442102
280	Pde4dip	3	0.007444259
281	Os9	2	0.007456501
282	Brix1	2	0.00766226
283	Ube2g1	2	0.00766226
284	Nrg2	2	0.008007353
285	Las1l	2	0.008026221
286	Macf1	2	0.008158816
287	Mb21d2	2	0.00821831
288	Cct8	2	0.008326338
289	Tia1	3	0.008414148
290	Zfp27	2	0.008435187
291	Mark3	2	0.008859494
292	Mars	2	0.008940398
293	Cacng2	3	0.00896325
294	Trim37	2	0.009012414
295	Pakap	2	0.009142597

296	Ubqln1	2	0.009142597
297	Wnk1	2	0.009355528
298	Lrch1	2	0.009462454
299	Vps37a	2	0.009462454
300	Ssh1	2	0.009617599
301	Pde3a	3	0.009643209
302	Ush2a	2	0.009873048
303	Camk4	2	0.009978559
304	Rexo1	2	0.009978559
305	Fancc	2	0.010183432
306	Smc5	4	0.010183432
307	Dhdds	2	0.010355442
308	Med29	2	0.010355442
309	Nat8l	2	0.010355442
310	Shprh	2	0.010355442
311	Ncoa3	3	0.010355442
312	Dicer1	2	0.010369881
313	Elk1	2	0.010644661
314	Khdrbs2	2	0.010672403
315	Cisd1	2	0.010752202
316	Fgf14	4	0.010954679
317	Asphd2	2	0.010983485
318	Hectd2	2	0.010983485
319	Sestd1	2	0.010983485
320	Mettl10	3	0.010983485
321	Zfp329	3	0.010983485
322	Atmin	2	0.011372636
323	Eefsec	2	0.011389006
324	Plgrkt	2	0.011485071
325	Faah	2	0.011509286
326	Epc1	2	0.011586339
327	Uggt2	3	0.011672245
328	Slc35b2	2	0.01171612
329	Rnf214	3	0.011786133
330	Vkorc1l1	2	0.011800356
331	Vps13c	2	0.011829575
332	Ntrk2	3	0.012056059
333	Homer1	6	0.01220821
334	Tpm1	2	0.012255069
335	Med13	3	0.012255069
336	Usp11	3	0.012255069
337	Emc7	2	0.01267831
338	Phka2	2	0.012692293

339	Dph3	2	0.012822623
340	Nop56	2	0.012822623
341	Sez6l2	2	0.013036901
342	Hinfp	2	0.013115101
343	Garnl3	2	0.013148497
344	Scn8a	2	0.013148497
345	Chmp4b	2	0.013540753
346	Fam57a	2	0.013540753
347	Mars2	5	0.013540753
348	Jmjd4	4	0.01371442
349	Creb1	5	0.013729811
350	Faf1	2	0.013732954
351	Emc6	2	0.0141242
352	Nol7	2	0.0141242
353	Fip1l1	3	0.015215669
354	As3mt	2	0.015371907
355	Hnrnpul2	2	0.015371907
356	Larp6	2	0.015371907
357	Anapc7	2	0.015448139
358	Cltb	3	0.015448139
359	Mef2c	2	0.01597896
360	Eif1ax	4	0.0160541
361	Cstf3	3	0.016127614
362	Sh2b3	2	0.016321983
363	Zfp941	2	0.016374243
364	Mical3	2	0.016696363
365	Arf3	2	0.017030226
366	Zbtb25	2	0.017030226
367	Atp6v1c1	2	0.017141976
368	Atp6ap2	2	0.0174271
369	Hnrnpll	2	0.0174271
370	Slc25a25	2	0.017472736
371	Far1	3	0.017551668
372	Iqce	2	0.017562494
373	Polr3c	2	0.017562494
374	Trp53bp1	2	0.017597532
375	Slc43a2	2	0.017712403
376	Ppm1h	2	0.017789161
377	Cd320	2	0.017826926
378	Dennd6a	2	0.017826926
379	Pdhb	2	0.017826926
380	Carkd	2	0.018126349
381	Abhd6	3	0.018135417

382	Nsg1	2	0.018299953
383	Alyref	2	0.01845238
384	Ago3	2	0.018732051
385	Arl6ip6	2	0.018737868
386	Golt1b	3	0.018737868
387	Scyl2	3	0.018754723
388	Aebp2	3	0.018760368
389	Anapc2	2	0.018839642
390	Pten	3	0.018839642
391	Bcan	2	0.01886929
392	Rsrp1	2	0.018983115
393	Nell2	2	0.019229586
394	Txn2	2	0.019828416
395	Ppfia2	3	0.019828416
396	Copg2	3	0.019834533
397	Chrnb2	2	0.019852604
398	Rap1b	2	0.020032705
399	Camk2a	2	0.020065901
400	4930429B21Rik	2	0.020068854
401	AI414108	2	0.020215305
402	Scrn1	2	0.020215305
403	Rps3	2	0.020276192
404	St18	2	0.02165778
405	Atxn1l	2	0.021686816
406	Smim14	2	0.021686816
407	Fkbp1a	2	0.02180567
408	Gnas	2	0.02186681
409	Txndc11	2	0.022000593
410	Cherp	2	0.022052573
411	Snx5	3	0.022129433
412	Zfp868	2	0.022236636
413	Dhx33	2	0.022412929
414	Fam71e1	2	0.022412929
415	Asb3	3	0.022866472
416	Mkks	3	0.022995931
417	Poln	2	0.023674833
418	Sike1	3	0.02375503
419	Cacna1d	3	0.024564361
420	Ints8	3	0.024564361
421	Cdh8	4	0.024783374
422	Slc35e1	3	0.024869598
423	Lrpap1	2	0.025347857
424	Ubxn2b	2	0.025771043

425	Tpp2	2	0.0258129
426	Car15	2	0.026004548
427	Lsm1	2	0.026004548
428	Ppp3ca	2	0.026004548
429	Pcf11	4	0.026004548
430	Sephs1	3	0.026057847
431	Thap1	2	0.026162221
432	Rab3gap2	3	0.026265659
433	Synj2	3	0.026265659
434	Anks1b	5	0.027157773
435	Fitm2	2	0.027186068
436	Safb2	2	0.027494032
437	Cfl2	2	0.027796283
438	Tmem64	2	0.027802026
439	Ccdc88a	4	0.028023882
440	Tmod3	2	0.028085165
441	Stam	3	0.028193661
442	Plbd2	2	0.028418796
443	Srgap2	2	0.028772948
444	Atp5b	2	0.028887746
445	Shfm1	2	0.028887746
446	Vps51	2	0.028887746
447	Atp2b3	2	0.029197137
448	Bhlhe41	2	0.029280551
449	Timm13	2	0.029280551
450	Cirbp	2	0.029288234
451	Hif1a	3	0.029548645
452	Actr1b	2	0.029877589
453	Rchy1	2	0.030136882
454	Mapk8ip2	2	0.030249373
455	Ptchd4	2	0.030249373
456	Azin1	2	0.03026784
457	Gid8	2	0.03026784
458	Rnaseh2c	2	0.03026784
459	Agfg1	3	0.03026784
460	Tmem206	3	0.03026784
461	Ralgapa2	2	0.030548715
462	Tm9sf2	2	0.030548715
463	Rnf138	2	0.031003775
464	Ahcyl2	2	0.031734993
465	Med19	2	0.031962408
466	Prpf39	2	0.031994572
467	Frmd5	2	0.032632491

468	Rap1a	2	0.032632491
469	Ap3b1	2	0.032963728
470	Dennd4a	3	0.032963728
471	Extl2	2	0.033023029
472	Mfhas1	2	0.033023029
473	Elmod2	2	0.033244118
474	Nsg2	2	0.033292184
475	Ado	2	0.033334788
476	Ncor1	4	0.033536557
477	Pcnx	3	0.033836739
478	1600012H06Rik	2	0.033939271
479	Alg11	2	0.03395052
480	Cox19	2	0.03395052
481	Nub1	3	0.03395052
482	Kif5b	3	0.034021579
483	Gemin5	2	0.034543625
484	Elmo1	2	0.034594588
485	Iqsec1	2	0.03516875
486	Lnx2	3	0.03516875
487	Tdrd3	2	0.035288255
488	Cand1	3	0.035288255
489	Vdac2	3	0.03538264
490	Cdk5	2	0.0360979
491	Gtf2h5	3	0.036404318
492	Zfp622	2	0.036461258
493	Kcnc3	2	0.036524875
494	Surf4	2	0.036686937
495	Anapc5	2	0.037357956
496	Ctnnd1	2	0.037532769
497	Srek1	3	0.038780802
498	Il1rapl1	2	0.039041715
499	Cul9	2	0.039045608
500	Dnajb9	2	0.039045608
501	Pias2	4	0.039045608
502	Ulk2	4	0.039045608
503	Mob4	3	0.039234498
504	Ip6k2	2	0.039345271
505	Srprb	2	0.039733833
506	Ptpn9	2	0.040891885
507	Zkscan1	2	0.040891885
508	Atp5g1	2	0.041066959
509	Erlec1	4	0.041066959
510	Pex14	2	0.041114721

511	Ing3	2	0.041480605
512	Nipbl	3	0.041612847
513	Bdh1	2	0.041715998
514	Pitpna	3	0.041859583
515	Prox1	2	0.042450846
516	Mrpl3	2	0.042629397
517	Sc5d	3	0.042928729
518	Hnrnpd	3	0.042936546
519	Glo1	2	0.043004897
520	Ubxn6	3	0.043004897
521	Kcnt1	3	0.04353307
522	Csmd2	2	0.043547594
523	Bzw1	2	0.043992609
524	Cbx8	2	0.044074061
525	Cep126	2	0.044074061
526	Pbx1	3	0.044471109
527	Gfm1	2	0.044866664
528	Tle1	3	0.044866664
529	Nipal2	2	0.044990305
530	Fam118b	2	0.045017095
531	Fundc2	4	0.045017095
532	Pex1	2	0.045207598
533	Ube2w	4	0.045708362
534	Myo9a	2	0.045714631
535	Ptch1	2	0.045855801
536	Alad	2	0.046104116
537	Il11	2	0.046104116
538	Ltk	2	0.046104116
539	Arfrp1	3	0.046104116
540	Plaa	3	0.046104116
541	Ints6	4	0.046132136
542	Dennd6b	2	0.046508593
543	Rwdd4a	2	0.046666369
544	Prkrir	2	0.047025999
545	Ankle2	2	0.04786654
546	Tcf4	4	0.047954947
547	Arhgef17	2	0.048492266
548	Adarb1	2	0.04864503
549	Creb3	2	0.049287719
550	Ivd	2	0.049552847
551	Arl4a	2	0.049824214

Table 2 | Gene ontology analysis of transcripts that show significant differences in alternative polyadenylation between granule and Purkinje cells

Functions Annotation	p-Value	Molecules
branching of neurites	2.12E-04	APP,ATF6,CAMK2A,CAMK2B,CDK5,CFL1, CHRNB2,CNTNAP2,EIF4G3,ELK1,FNBP1, HTT,KIDINS220,MAP1B,NCAM1,PDIA3, PIK3CA,PKP4,PTEN,RIMS4,ROBO1,ROBO2, RPS6KA3,RYR2,SHANK3,SLC12A5,TRIM9, TRIO,TRPC1,UGCG,ULK2,VEZT,ZDHHC8
ataxia	4.38E-04	AFG3L2,APP,ATP2B2,B4GALNT1,CACNG2, CACNG7,CAMK2B,CAMTA1,CANX,CHRNB2, CTSA,DAB1,DNAJC5,GNAS,HTT,MAP1B, MAPK8IP2,NCAM1,NDUFS4,PITPNA,PMPCA, SCN2B,SPRN,TMEM240,TRIO,TTL,VAMP1
neurotrans-mission of brain cells	4.56E-04	APP,CACNG2,CDK5,CNTNAP2,HIF1A, HOMER1,HTT,PTEN,PTPRS,SH3GL2, SHANK3
synaptic transmission	4.77E-04	AMPH,APP,CACNA2D3,CACNG2,CAMK2A, CDK5,CHRNB2,CLSTN2,CNTNAP2,DLGAP1, DNAJC5,DPP6,GLS,HIF1A,HLA-A,HOMER1, HTT,MAPK1,MAPK8IP2,MBP,NAPA,NCS1, NPTX1,PNKD,PTEN,PTPRS,RIC3,SCN2B, SH3GL2,SHANK3,SLC12A5,SSH1,STX1B
neuritogenesis	5.36E-04	ACVR2A,AFG3L2,AGTPBP1,ANAPC2,APC, APP,ATF6,CAMK2A,CAMK2B,CD47,CDK16, CDK5,CFL1,CHRNB2,CNTNAP2,DAB1, EIF4G3,ELK1,FNBP1,FUT8,HTT,KIDINS220, LRRK2,MAP1B,MAP6,MAPK8IP2,MBP, NCAM1,NTNG1,PDIA3,PIK3CA,PITPNA, PKP4,PTEN,RAB21,RAB35,RAP1B,RERE, RIMS4,ROBO1,ROBO2,RPS6KA3,RYR2,SHA NK3,SIRPA,SLC12A5,SLITRK1,SRGAP2, TRIM9,TRIO,TRPC1,TTL,UGCG,ULK2,VEZT, ZDHHC8
exocytosis	5.72E-04	ADGRL1,AMPH,ARFGEF2,CAMK2A,CDK16, CDK5,DNAJC5,HTT,LNPEP,NAPA,NCAM1, NCS1,PCLO,PITPNA,PNKD,RAB21,STX1B, STXBP5L,TRIM9,VAMP1,VAMP2
morphogenesis of neurons	5.87E-04	ANAPC2,APC,APP,ATF6,CAMK2A,CAMK2B, CDK16,CDK5,CFL1,CHRNB2,CNTNAP2, EIF4G3,ELK1,FNBP1,HTT,KIDINS220,LRRK2, MAP1B,MAP6,MAPK8IP2,MEF2C,NCAM1,

		PDIA3,PIK3CA,PKP4,PTEN,RAB21,RAB35, RERE,RIMS4,ROBO1,ROBO2,RPS6KA3, RYR2,SHANK3,SLC12A5,SRGAP2,TRIM9, TRIO,TRPC1,UGCG,ULK2,VEZT,ZDHHC8
synaptic transmission of nervous tissue	7.00E-04	APP,CACNA2D3,CHRNB2,CNTNAP2,HIF1A, HOMER1,HTT,PNKD,PTEN,PTPRS,RIC3, SH3GL2,SHANK3,SSH1
morphogenesis of neurites	7.42E-04	ANAPC2,APC,APP,ATF6,CAMK2A,CAMK2B, CDK16,CDK5,CFL1,CHRNB2,CNTNAP2, EIF4G3,ELK1,FNBP1,HTT,KIDINS220,LRRK2, MAP1B,MAP6,MAPK8IP2,NCAM1,PDIA3, PIK3CA,PKP4,PTEN,RAB21,RAB35,RERE, RIMS4,ROBO1,ROBO2,RPS6KA3,RYR2, SHANK3,SLC12A5,SRGAP2,TRIM9,TRIO, TRPC1,UGCG,ULK2,VEZT,ZDHHC8
folding of protein	7.78E-04	ATF6,CANX,CCT4,DNAJB2,ERP44,FKBP4, HSP90AB1,HSPA4L,HTT,LRPAP1,PDIA3,RIC3
transport of synaptic vesicles	7.81E-04	ADGRL1,AMPH,CANX,CDK5,DNAJC5,HTT, PCLO,PNKD,SH3GL2,VAMP1,VAMP2
neurotransmission	9.11E-04	AMPH,APP,CACNA2D3,CACNG2,CAMK2A, CD47,CDK5,CHRNB2,CLSTN2,CNTNAP2, DLGAP1,DNAJC5,DPP6,GLS,HIF1A,HLA-A, HOMER1,HTT,KCNJ11,MAPK1, MAPK8IP2,MBP,NAPA,NCS1,NPTX1,PNKD, PTEN,PTPRS,RIC3,RPS6KA3,SCN2B, SH3GL2,SHANK3,SLC12A5,SRI,SSH1,STX1B
abnormal morphology of telencephalon	1.15E-03	APP,ATP1B2,ATXN1L,CD81,CNTNAP2,DAB1, HTT,MAP1B,MAP3K12,NCAM1,PTEN,PTPRS, ROBO1,ROBO2,SHANK3,SMARCC1,TRIO, TTL,VPS13A
transport of metal	1.24E-03	AKT3,AP3B1,APP,ATP1B2,ATP2B2, CACNA2D3,CAMK2A,CAMK2B,CFL1, CHRNB2,CISD1,DPP6,EGLN1,FKBP4,GNAS, HIF1A,HLA-A,HTT,KCNJ11,KIF5B,MAPK1, MLLT6,NIPA1,RYR2,SLC12A5,SLC39A10,SRI, TRPC1,WNK1,YWHAE
morphology of telencephalon	1.30E-03	AKT3,APP,ATP1B2,ATXN1L,CD81,CNTNAP2, DAB1,HTT,MAP1B,MAP3K12,NCAM1,PTEN, PTPRS,ROBO1,ROBO2,SHANK3,SMARCC1, TRIO,TTL,VPS13A

Table 3 l Transcripts that show significant differences in alternative polyadenylation between granule cell precursors and adult granule cells

	Gene name	Number of 3'UTR isoforms	q-value
1	Sin3b	2	5.41E-15
2	Add2	2	1.01E-11
3	Tppp	2	7.39E-11
4	Actr10	3	2.99E-10
5	Dclk1	3	2.92E-09
6	Map6	2	3.38E-09
7	1500011B03Rik	2	7.38E-09
8	Clybl	2	1.42E-08
9	Kcnc1	2	1.42E-08
10	Osbpl1a	2	3.53E-08
11	Mbd2	3	1.24E-07
12	Mbp	2	5.29E-07
13	Sugt1	3	7.78E-07
14	Itsn1	4	1.11E-06
15	Rps6ka3	2	1.20E-06
16	Hnrnpa3	3	1.28E-06
17	Mcl1	2	3.42E-06
18	Dpp8	2	3.65E-06
19	Klc1	2	3.71E-06
20	Apba2	2	5.57E-06
21	Csk	2	5.57E-06
22	Pld3	3	5.57E-06
23	Hnrnpr	2	6.31E-06
24	Rims1	3	7.49E-06
25	Asf1a	2	7.59E-06
26	Sap30l	2	7.93E-06
27	Tubb5	2	7.93E-06
28	Sptb	3	1.36E-05
29	Rab22a	2	1.54E-05
30	Calm3	2	1.68E-05
31	Mtpn	2	1.68E-05
32	Ogfrl1	2	1.76E-05
33	Fbxo9	2	2.66E-05
34	Rps15a	2	2.92E-05
35	Tbc1d9b	2	2.92E-05
36	Sp3os	2	3.05E-05
37	Map4k4	2	3.29E-05
38	Abcd2	2	3.71E-05
39	Atp2a2	2	3.71E-05
40	Kpnb1	2	3.71E-05

41	Mtf1	3	3.71E-05
42	Stt3a	3	3.71E-05
43	Mrpl57	2	3.75E-05
44	Fam111a	2	4.26E-05
45	Klhl36	2	4.26E-05
46	Rnf214	3	4.26E-05
47	Sephs1	4	4.26E-05
48	Elavl1	3	4.62E-05
49	Ip6k2	3	5.78E-05
50	Ncl	2	7.04E-05
51	Strbp	3	7.04E-05
52	Cirbp	2	7.37E-05
53	Blvra	2	9.08E-05
54	Pip5k1b	2	9.35E-05
55	Adam23	2	9.51E-05
56	Rdx	2	9.66E-05
57	Sept6	2	9.66E-05
58	Anp32e	2	0.000102437
59	Gprc5b	2	0.000119749
60	Mapk8	2	0.000119749
61	Nsf	2	0.00012171
62	Hspa4	2	0.000132256
63	Pak1	2	0.000142139
64	Camta1	2	0.000154428
65	Dixdc1	3	0.000161794
66	Dlgap4	2	0.00016524
67	Dut	2	0.00016524
68	Kidins220	2	0.00016524
69	Sec14l1	2	0.000178799
70	Rbmx	2	0.000184336
71	Pnkd	2	0.000219509
72	Zcchc11	2	0.000219832
73	Emc2	2	0.000225603
74	Erich3	2	0.000225603
75	Trim9	3	0.000246315
76	Anks1b	4	0.000259569
77	Adrbk1	2	0.000280499
78	Zfp609	2	0.000317082
79	Caln1	2	0.000322482
80	Chd9	2	0.000353197
81	Zfp467	2	0.000418154
82	Il16	4	0.000429808
83	Mars	2	0.0004368

84	Homer2	4	0.000445619
85	Prrc2b	2	0.000450266
86	Sc5d	2	0.000465871
87	Sfrp2	2	0.000544495
88	Map2k4	3	0.000544935
89	Cd200	2	0.000559727
90	Pdgfra	2	0.000602419
91	Fem1b	2	0.000643584
92	Ralgps2	2	0.000643584
93	Kcnk10	2	0.000655004
94	AA388235	2	0.000689465
95	Tomm20	2	0.000767512
96	Agap3	2	0.000778986
97	Raly	2	0.000778986
98	Rpl18a	2	0.00079852
99	Rab11fip3	3	0.000834767
100	Emc7	2	0.000848191
101	Brd4	3	0.000848191
102	Mapre2	2	0.000868441
103	Gnao1	3	0.000871373
104	Socs4	3	0.000871373
105	Etv1	3	0.000871897
106	Bri3bp	2	0.000926462
107	Pofut2	2	0.001002791
108	Ndst3	2	0.00100886
109	Tmem143	2	0.001034179
110	Rbm33	2	0.001071397
111	Hnrnpa1	2	0.001072004
112	Ddx25	2	0.001110145
113	Memo1	2	0.001110145
114	Araf	2	0.00111328
115	Tial1	2	0.001147993
116	Orc3	3	0.001165395
117	Azi2	5	0.001170578
118	Ralgapa1	2	0.001177389
119	Asph	3	0.00121011
120	Usp46	3	0.00128673
121	Phf12	2	0.001312835
122	Malat1	2	0.001318544
123	Pgm2l1	3	0.00132271
124	Ppig	3	0.001435283
125	Ngly1	3	0.001436742
126	Mtor	2	0.001480633

127	Cep97	2	0.00154127
128	Znrf1	2	0.00154127
129	Ccnl2	2	0.00158532
130	Calm1	3	0.001790617
131	Ptprd	3	0.002163067
132	Zbtb44	3	0.002203655
133	Gspt1	4	0.002230696
134	Dmap1	2	0.002350217
135	Ubxn2a	3	0.002360357
136	Dicer1	2	0.002367966
137	Tmem57	2	0.002444281
138	Paip1	2	0.00246883
139	Cdk7	3	0.002475575
140	Clcn3	3	0.002475575
141	Svip	4	0.002475575
142	Zfp846	2	0.002537334
143	Gm3086	2	0.00261611
144	Nbea	3	0.002677104
145	Gpr137c	2	0.002717921
146	Dcun1d5	2	0.002951938
147	Tubgcp2	2	0.003005321
148	Huwe1	3	0.003048398
149	Sms	2	0.003189904
150	Ispd	2	0.003217237
151	Knop1	2	0.003217237
152	Phf21a	2	0.003217237
153	Luc7l3	2	0.003291179
154	Cds2	2	0.003300095
155	Eif3m	4	0.00332846
156	Mapk1	3	0.003448682
157	Pnrc1	2	0.00345669
158	St8sia3	3	0.003507847
159	Kctd20	2	0.00358774
160	Bcl2l11	2	0.003593984
161	Camk4	2	0.003618378
162	Nmt2	2	0.003734226
163	Kcnt1	3	0.003846626
164	Whsc1l1	5	0.003889797
165	Pdlim7	2	0.003917828
166	Grik2	4	0.003972854
167	Manbal	2	0.004084629
168	Nedd4l	2	0.004084629
169	Rcc2	2	0.004084629

170	Rsc1a1	2	0.004084629
171	Samm50	3	0.004084629
172	Blcap	2	0.00419491
173	Ilf3	2	0.004244719
174	Fam168a	2	0.004255678
175	Shank2	3	0.004255678
176	Naa15	3	0.004592365
177	Cnksr2	5	0.004697961
178	Dhx36	2	0.004700211
179	Tbl1xr1	2	0.004730654
180	4933431E20Rik	2	0.004762761
181	Ptprg	2	0.004762761
182	Apc	3	0.004762761
183	Tesk1	2	0.004775847
184	Zfp882	3	0.004775847
185	Suds3	2	0.004776466
186	Vapb	2	0.00479394
187	Mta3	3	0.004893745
188	Ahcyl1	2	0.004911514
189	Ttll11	3	0.004977995
190	Ddx19a	2	0.005123907
191	Plxdc2	3	0.005137503
192	Myo5a	2	0.005254844
193	Cstf1	2	0.005338395
194	Eps15	3	0.005338395
195	Smu1	3	0.005395669
196	Sin3a	2	0.005397934
197	Scrn1	2	0.005407743
198	Gmps	2	0.005434272
199	Mier1	2	0.005478155
200	Pcdh9	3	0.005478155
201	Nfic	2	0.005737054
202	Nrcam	2	0.005737054
203	Polr3a	2	0.005737054
204	Rngtt	2	0.005737054
205	Ppp2r2c	2	0.005848851
206	Zfp955b	2	0.005848851
207	Ppp1r7	3	0.005904873
208	Klhdc10	2	0.005957682
209	Zfp521	2	0.005957682
210	Pgm3	2	0.005966666
211	Ints12	2	0.005977399
212	Ctbp2	3	0.005977399

213	Synpo	2	0.005984325
214	Scn2a1	2	0.006066676
215	Glo1	2	0.006194603
216	Supt6	2	0.006210637
217	Pias1	4	0.006262137
218	Ak4	2	0.006289833
219	Sec63	3	0.006475643
220	Gnb1l	3	0.006531076
221	Phf14	2	0.006547165
222	Zbtb46	2	0.006562463
223	Commd2	3	0.006875959
224	Snrpd3	2	0.006915115
225	Yy1	3	0.006915115
226	Hnrnpu	3	0.00692368
227	Pask	2	0.00723964
228	Svep1	4	0.007296182
229	Fam120b	3	0.007447874
230	Man1a2	2	0.007466155
231	Rab3gap1	2	0.007466155
232	Vldlr	2	0.007595355
233	Rfx7	3	0.007595355
234	Pcmt1	4	0.007644846
235	Cpeb3	2	0.007767226
236	Ccdc127	3	0.007879846
237	Bsn	2	0.008160152
238	Aak1	2	0.008165968
239	Ccnd1	2	0.00821023
240	Sgsm1	2	0.008328194
241	Yars2	3	0.008356
242	B4galnt1	2	0.008369935
243	Neurod1	2	0.008369935
244	Vrk3	2	0.008517263
245	Fam49b	2	0.008521065
246	Snx12	2	0.008613402
247	Tceb1	2	0.008613402
248	Srr	2	0.008704522
249	Sept8	3	0.008787286
250	Atg2b	3	0.008822251
251	Lsm11	3	0.008851863
252	Pitpna	3	0.008851863
253	Gatad1	5	0.008851863
254	Mdga1	2	0.008974078
255	Ube2e3	2	0.009111356

256	Dnaja1	2	0.009138948
257	Rab10os	2	0.009346867
258	Nap1l1	3	0.009346867
259	Arl8b	3	0.009691447
260	Far1	3	0.009691447
261	Mcur1	2	0.009717358
262	Stat1	2	0.009909764
263	Eif1ax	3	0.009909764
264	Dlgap1	2	0.010075898
265	Stxbp5	3	0.010075898
266	A830080D01Rik	2	0.010154878
267	Map2k6	2	0.010158097
268	Itgb3bp	3	0.010610012
269	Ube2j1	2	0.01067413
270	Cflar	2	0.010888575
271	Neo1	2	0.011072266
272	Dync1li2	2	0.011139774
273	Ssr1	5	0.011287944
274	Insr	2	0.011403151
275	Gls	3	0.011403151
276	Tanc1	2	0.011456838
277	Klhl24	2	0.011602897
278	Fgf14	3	0.011602897
279	Slc38a9	2	0.011893192
280	Ube2g1	2	0.011893192
281	Edc3	2	0.012160575
282	Ext1	2	0.012160575
283	Commd1	2	0.012170776
284	Rpf2	2	0.012171326
285	Sema3a	2	0.012171326
286	Ppp2r5e	4	0.012171326
287	Tmem201	2	0.012181725
288	Fh1	2	0.012288246
289	Tmem106b	3	0.01234699
290	Zbtb37	2	0.012352493
291	Arhgef2	4	0.012404765
292	Pdzd11	2	0.012600123
293	Pold3	2	0.012600123
294	Zc3h14	3	0.012700249
295	Ring1	2	0.013024273
296	Mbip	2	0.013127687
297	Nipa1	2	0.013127687
298	Nsd1	3	0.013247251

299	St18	2	0.013557055
300	Med24	2	0.013571127
301	Map3k10	3	0.013897614
302	Rexo2	2	0.014000443
303	Vps29	3	0.014045314
304	Gorasp2	3	0.014199631
305	Timm10b	3	0.014199631
306	Gtpbp3	2	0.014250816
307	Nfyb	3	0.014250816
308	1500009L16Rik	2	0.014320718
309	Ttc9c	2	0.014320718
310	Ghitm	2	0.014512035
311	Unc5c	2	0.014522785
312	Wdr73	2	0.014529096
313	Ascc3	2	0.014539976
314	Pafah1b1	3	0.014539976
315	Mmd	4	0.014539976
316	Colgalt1	2	0.014607712
317	Fam210a	5	0.014815462
318	Srsf7	2	0.015080366
319	Cntn6	2	0.015128756
320	Rpl11	2	0.015128756
321	Lamp2	3	0.015128756
322	Ifnar2	2	0.015518369
323	2610100L16Rik	3	0.015518369
324	Ebpl	2	0.015608271
325	Cnot4	5	0.015963762
326	Limch1	2	0.016404862
327	Rnaseh2c	2	0.016404862
328	Faim2	4	0.016404862
329	Zfp637	2	0.016642645
330	Phactr2	2	0.017008364
331	Mpv17l	2	0.017175412
332	Wdr12	3	0.017175412
333	Otud5	3	0.017452936
334	Mrpl47	2	0.017534168
335	Nsg2	2	0.0177162
336	Fzd1	2	0.018093217
337	Psmb6	2	0.018093217
338	Ndufa12	2	0.018611396
339	Auh	2	0.019149367
340	Nus1	3	0.01924283
341	Dnajb12	2	0.019266279

342	Rab3ip	3	0.019266279
343	Ppip5k2	2	0.019723947
344	Dpf3	3	0.019782783
345	Oxct1	2	0.019884654
346	Cr1l	2	0.020162361
347	Ephb2	3	0.02017221
348	Keap1	2	0.02034753
349	2310010J17Rik	2	0.02037251
350	Map4	3	0.020986456
351	Clip1	4	0.021111635
352	Cycs	2	0.02115666
353	Negr1	3	0.02115666
354	Tmx3	3	0.02115666
355	Mrrf	4	0.02115666
356	Cd47	2	0.021360736
357	Gabbr1	2	0.021360736
358	Zfp560	3	0.021412634
359	Eefsec	3	0.021824819
360	Baiap2	2	0.021891643
361	Mok	2	0.021891643
362	Nol7	2	0.021891643
363	Evi5l	2	0.021899164
364	Aggf1	2	0.021921536
365	Nktr	2	0.021921536
366	Ssu72	2	0.022308327
367	Btbd7	2	0.022325471
368	Irf2bp1	2	0.022391708
369	E130308A19Rik	2	0.022466455
370	Spast	2	0.022522957
371	Rpl27a	2	0.022591203
372	Tmem208	2	0.022591203
373	Ap3s2	4	0.022591203
374	Dlg5	2	0.022635355
375	Efnb2	2	0.022635355
376	Impa1	2	0.022635355
377	Rps6	2	0.022635355
378	Snx13	2	0.022635355
379	Uxs1	2	0.022635355
380	Cep295	3	0.022635355
381	Safb2	3	0.022635355
382	Lsm6	4	0.022635355
383	Me2	3	0.022652502
384	Amigo1	2	0.022698575

385	Anp32a	2	0.022790069
386	1700025G04Rik	3	0.022856013
387	2510009E07Rik	2	0.023102377
388	Fbxl17	4	0.023153945
389	Tnks2	2	0.023356353
390	Atf2	4	0.023747356
391	Cep57	2	0.023936695
392	Soga1	2	0.023936695
393	Gpr89	3	0.024130546
394	Sbno1	2	0.024209931
395	Zxda	2	0.024209931
396	Ppm1h	3	0.024209931
397	Smim14	2	0.02425201
398	Gnb1	2	0.024520164
399	Pcyox1	2	0.024520164
400	Hdhd2	3	0.024520164
401	Sdf4	2	0.024531997
402	Bmp1	2	0.024561156
403	Eif4e2	2	0.02511587
404	Srf	2	0.02511587
405	Ccdc71l	2	0.025291855
406	Cep95	2	0.025291855
407	Pmm2	2	0.025291855
408	Kif5b	3	0.025304273
409	Cdc123	2	0.025384208
410	Ets2	2	0.025384208
411	G2e3	2	0.025384208
412	Raph1	3	0.025384208
413	Zfp39	2	0.025579878
414	Nck1	2	0.025594328
415	Cpne1	3	0.025739271
416	Rbx1	2	0.025755981
417	Rab30	2	0.025828905
418	Dnajc5	2	0.025840058
419	Sec62	2	0.025901379
420	Tpm1	4	0.026039078
421	Cacna1b	2	0.026168797
422	Asphd2	2	0.026177676
423	Otub1	2	0.026198168
424	Ktn1	2	0.026655688
425	Supt20	2	0.026655688
426	Alyref	3	0.02667205
427	Dcp2	2	0.027192204

428	Pitpnc1	2	0.027192204
429	Gpatch11	2	0.027234253
430	Ncoa2	2	0.027641591
431	Srpk2	2	0.027641591
432	Tfb1m	2	0.027641591
433	Kdm2a	3	0.027641591
434	Socs6	3	0.027641591
435	Hgsnat	3	0.027854147
436	Stk16	3	0.028684647
437	Dhdds	2	0.028798777
438	Ube2d3	4	0.028798777
439	Ncor1	4	0.028970534
440	Med11	2	0.029142232
441	Prep	3	0.029142232
442	Pcsk2	2	0.029333489
443	Wdr4	2	0.029333489
444	Fbxo3	3	0.029333489
445	Luc7l2	2	0.029346711
446	Zmym3	2	0.02936815
447	Reep5	2	0.029725178
448	Sgsh	2	0.029756037
449	Zfp780b	2	0.030116969
450	Rala	3	0.030216474
451	Ncs1	2	0.030432347
452	Satb1	4	0.030481108
453	Ap1b1	2	0.030718522
454	Kpna6	3	0.030718522
455	Zfyve27	3	0.030747374
456	Mga	2	0.030803732
457	Rprd1a	2	0.030803732
458	Sipa1l1	2	0.030803732
459	Unc119	2	0.030803732
460	Ggps1	3	0.030803732
461	Arpp21	2	0.030849514
462	Sucla2	2	0.030849514
463	Mrpl17	5	0.030849514
464	St3gal2	2	0.031002069
465	Mzt1	3	0.031034708
466	Hinfp	3	0.031481853
467	Ikzf2	2	0.031685636
468	Sqstm1	2	0.031729771
469	Mef2c	2	0.03186886
470	Slc30a5	3	0.031925197

471	Pak7	3	0.032125623
472	Jmy	2	0.032455702
473	Samhd1	2	0.032455702
474	Trak1	2	0.032455702
475	Ears2	2	0.032543237
476	Rnf103	2	0.032543237
477	Ppp2r2a	4	0.032543237
478	Dlat	4	0.032623764
479	Tmem135	3	0.032758061
480	Ubl7	2	0.034096251
481	Ubap1	2	0.034188966
482	AI593442	3	0.034243864
483	Recql	2	0.034281827
484	Cep170	2	0.034283084
485	Colec12	2	0.034283084
486	Slc15a4	2	0.034283084
487	Enah	3	0.034395307
488	Ccdc85b	2	0.034500631
489	Lyst	2	0.034500631
490	Gm15706	2	0.034550142
491	Ncaph2	2	0.034550142
492	Shf	2	0.034550142
493	Zdhhc18	2	0.034550142
494	Lamtor1	2	0.034771666
495	Rps6ka5	3	0.034802423
496	Adora1	2	0.034920804
497	Akap8	3	0.034939843
498	Borcs8	2	0.034967025
499	BC003965	4	0.034967025
500	S1pr1	2	0.035035163
501	Rgl1	2	0.03542942
502	Rad18	3	0.035465491
503	Nup160	3	0.035690302
504	Srprb	2	0.03605877
505	Ythdc1	3	0.036220232
506	Rprd2	5	0.036220232
507	1810013L24Rik	3	0.036335479
508	Alkbh8	3	0.036626563
509	Slc48a1	2	0.036663127
510	Gtf2f2	2	0.036749952
511	Cacnb2	3	0.036749952
512	Nupl1	4	0.036775767
513	Alad	3	0.036893493

514	Zmym4	4	0.037360843
515	Dpm1	3	0.037603232
516	Atp1b1	2	0.037730186
517	Azin2	2	0.037798101
518	Ube4b	2	0.037869201
519	Agpat5	6	0.038098622
520	Gng12	3	0.03826348
521	Eps15l1	2	0.038559661
522	Unc80	2	0.038833963
523	Senp1	2	0.038932055
524	Nr2c1	3	0.038983507
525	Smim4	2	0.039266933
526	Snrnp27	2	0.039326394
527	Mrps6	2	0.039507749
528	Sec61g	2	0.039553703
529	4930453N24Rik	3	0.039669908
530	Kif3a	3	0.039669908
531	H3f3a	2	0.039983514
532	Slc25a32	2	0.039983514
533	Tbl2	3	0.039983514
534	Pds5b	4	0.040007085
535	Kcnk9	2	0.040250076
536	Xpo6	2	0.040463683
537	Ppp1r2	4	0.040487847
538	Mkln1	3	0.040755575
539	Tmem39a	3	0.040755575
540	Cdk16	2	0.040999515
541	Srsf1	3	0.040999515
542	Ppp2r5c	3	0.041205392
543	Slc9a3r2	2	0.041259447
544	Gnas	2	0.042449483
545	Sec23ip	2	0.042449483
546	Magi1	3	0.042459739
547	Atl3	2	0.042950761
548	Mir3098	2	0.042950761
549	Ccdc186	3	0.04360625
550	Pex6	2	0.043722046
551	Ccser1	2	0.04378467
552	Ndufs4	2	0.04378467
553	Sclt1	2	0.044111195
554	Gskip	3	0.045242242
555	Prpf31	3	0.04548784
556	Opa1	2	0.045696687

557	Dcaf12	2	0.045875127
558	Caprin1	4	0.045968817
559	Map3k2	2	0.045969559
560	Rnpepl1	2	0.046168215
561	Robo2	3	0.046168215
562	2810021J22Rik	2	0.046206339
563	Stag1	2	0.046206339
564	Tpmt	2	0.046206339
565	Rab10	2	0.046884346
566	Arl3	2	0.046909255
567	Mcu	2	0.046909255
568	Srcin1	2	0.046909255
569	Tmeff2	2	0.046909255
570	Capn15	2	0.046932432
571	Dtnb	2	0.046932432
572	Dyrk1a	2	0.046932432
573	Nelfe	2	0.046932432
574	Pcca	2	0.046932432
575	Zfp651	2	0.046932432
576	Zfp945	2	0.046932432
577	Maoa	4	0.046932432
578	Nup98	4	0.046932432
579	Twistnb	2	0.047213475
580	Ank2	2	0.047289831
581	Hnrnpul2	3	0.047289831
582	Mob1b	2	0.04746583
583	Napg	2	0.04746583
584	Rock2	2	0.04746583
585	Smc5	3	0.047732834
586	Cpsf6	3	0.048168777
587	Plbd2	3	0.048168777
588	Aifm2	2	0.048709496
589	Foxn2	2	0.048709496
590	Synj2bp-cox16	3	0.048709496
591	Asxl2	2	0.04923985
592	Fra10ac1	2	0.04923985
593	Mios	2	0.049314565
594	Rbm8a	2	0.049314565
595	Ercc6	3	0.049429213
596	Cnih4	2	0.049732575
597	Fgf12	4	0.049763651
598	Gpat4	3	0.049867891
599	Gpr75	2	0.049892004

Table 4 l Gene ontology analysis of transcripts that show significant differences in alternative polyadenylation between granule cell precursors and adult granule cells

Functions Annotation	p-Value	Molecules
morphology of brain cells	3.12E-04	B4GALNT1,CHL1,CLCN3,CNTN6,CPEB3, FAIM2,GRID2,IREB2,NEUROD1,PAFAH1B1, PITPNA,UBE4B,ZNF521
behavior	4.80E-04	ADD2,APBA2,APC,ARAF,B4GALNT1, CAMTA1,CCND1,CD47,CHD7,CHL1,CLCN3, CPEB3,CSK,DICER1,DIXDC1,EPHB2,FZR1, GLS,GMFB,GNAO1,GNB1,GRID2,HOMER2, IDS,ILF3,IMPA1,INSR,IREB2,ITM2B,KCNT1, LNPEP,MAP6,MAPK1,MBD2,MTOR,NCOA2, NEGR1,NEUROD1,NFIC,NRCAM,PAFAH1B1, PAK5,PIAS1,PTPRG,RIMS1,RNF103,ROBO2, RSC1A1,SHANK2,SLC12A2,SQSTM1,SRF, SRR,STAT1,TSN,USP46,VLDLR,ZNF521
morphology of neurons	7.12E-04	B4GALNT1,BAIAP2,BCL2L11,CD47,CHL1, CLCN3,CNTN6,CPEB3,DICER1,DIXDC1, EPHB2,FAIM2,GRID2,Hnrnpa1,IDS,IREB2, ITSN1,KLC1,MAPK1,MBP,MTOR,NEUROD1, NRCAM,PAFAH1B1,PAK5,PITPNA,PTPRS, RDX,RIMS1,ROBO2,RPS6KA3,SHANK2, SIPA1L1,SLC12A2,SORT1,SPAST,UBE4B, ZNF521
growth of neurites	1.97E-03	Anp32a,APBA2,BAIAP2,CD47,CHL1,CSNK1D, DIXDC1,EPHB2,EXT1,FZR1,GNAO1,GPM6A, IMPA1,KIF1BP,MAPK1,MTOR,NCK1,NEDD4L, NRCAM,PAFAH1B1,PAK5,PTPRS,RAB22A, ROBO2,SLC12A2,SNX12,SNX3,SPAST,SRF, SYNGR1,TESK1,UBE4B,VLDLR,ZFYVE27
proliferation of neuronal cells	2.43E-03	Anp32a,APBA2,BAIAP2,CCND1,CD47,CHL1, CSNK1D,DIXDC1,EPHB2,EXT1,FZR1,GNAO1, GPM6A,IMPA1,KDM2A,KIF1BP,MAPK1, MTOR,NCK1,NEDD4L,NEGR1,NRCAM, PAFAH1B1,PAK5,PTPRS,RAB22A,ROBO2, SLC12A2,SNX12,SNX3,SPAST,SRF,SYNGR1, TESK1,UBE4B,VLDLR,ZFYVE27
neuritogenesis	3.07E-03	APC,BAIAP2,BSN,CD47,CHL1,CPEB3,CSNK2 A2,DICER1,DIXDC1,DLG5,EIF4G3,EPHB2, GPM6A,GRID2,ITSN1,JMY,MAP2K4,MAP6, MBP,MTOR,NEUROD1,NRCAM,PDGFRA, PIP5K1B,PITPNA,PPP2CA,RALA,RBFOX2,

		RDX,RIMS1,ROBO2,RPS6KA3,SHANK2, SIPA1L1,SPAST,SRF,STXBP5,TMEM106B, TRIM9,UBE4B,VLDLR,ZFYVE27
metabolism of protein	4.28E-03	AGAP3,APC,BMP1,CCND1,CIRBP,CNOT4, CPEB3,CPNE1,DDX25,DICER1,EIF3M, EIF4E2,EIF4G3,FLOT2,FZR1,HNRNPD, HOMER2,ILF3,INSR,IREB2,LNPEP,MAN1A2, MAP2K4,MAPK1,MARS,METAP2,MRPL57, MTOR,NAA15,NCK1,NCOA2,NEDD4L,NRDC, NSF,Otud5,PASK,PCYOX1,PDCD4,PPP2CA, PREP,Rbx1,RPL27A,RPS15A,SENP6,SNX12, SNX3,SQSTM1,SRF,SUGT1,UBA3,UBE2D3, UBE4B,UBXN2A
morphology of rhombencepha-lon	4.65E-03	B4GALNT1,CCND1,FAIM2,GMFB,GRID2, IREB2,NEUROD1,NRCAM,PAFAH1B1, PITPNA,ROBO2,UBE4B,VLDLR,ZNF521
abnormal morphology of neurons	5.17E-03	B4GALNT1,BCL2L11,CD47,CHL1,CLCN3, CNTN6,CPEB3,EPHB2,GRID2,Hnrnpa1,IDS, IREB2,NEUROD1,NRCAM,PAFAH1B1,PAK5, PITPNA,PTPRS,ROBO2,RPS6KA3,SHANK2, SLC12A2,SORT1,SPAST,UBE4B,ZNF521
morphology of cerebellum	5.70E-03	B4GALNT1,CCND1,FAIM2,GRID2,IREB2, NEUROD1,NRCAM,PAFAH1B1,PITPNA, ROBO2,UBE4B,VLDLR,ZNF521
anxiety	6.20E-03	APBA2,CHL1,CPEB3,CSK,DICER1,DIXDC1, IDS,INSR,MAPK1,NDST3,RNF103,SHANK2, TSN,ZNF521
coordination	6.60E-03	B4GALNT1,CCND1,CLCN3,CNTN6,GNAO1, GRID2,IREB2,KIF1B,LNPEP,NRCAM, PAFAH1B1,PTPRG,RNF103,RPS6KA3, SPAST,UBE4B,VAPB
abnormal morphology of metencephalon	7.54E-03	B4GALNT1,CCND1,GMFB,GRID2,IREB2, NEUROD1,NRCAM,PITPNA,ROBO2,UBE4B, VLDLR,ZNF521
autosomal dominant disease	8.69E-03	APC,BMP1,BNIP3L,CAMTA1,CCND1,CELF2, CHD7,CSNK1D,CYCS,DICER1,ETS2,EXT1, FAM111A,FH,GARS,GNAO1,GNB1,HOMER2, INSR,ITM2B,KCNT1,KIF1B,MARS,MTOR, NEUROD1,NIPA1,PDGFRA,RIMS1,SAMHD1, SEC63,SPAST,SPTB,SQSTM1,STAT1, TBL1XR1,TUBB,VAPB,VLDLR,ZFYVE27
transport of protein	9.12E-03	AGAP3,AP1B1,AP3S2,CEP57,COMMD1, CSNK1D,EPS15,KIF1B,KLC1,KPNA6,KPNB1, MBP,NSF,RAB2A,SNX12,SNX13,SNX3,

		SORT1,SQSTM1,TOMM70,TRAK1,VPS33A

References

Alt, F.W., Bothwell, A., Knapp, M., Siden, E., Mather, E., 1980. Synthesis of secreted and membrane-bound immunoglobulin mu heavy chains is directed by mRNAs that differ at their 3′ ends. Cell 20, 293–301. doi:10.1016/0092-8674(80)90615-7

Altaba, N.D.A.A.R.I., 1999. Sonic hedgehog regulates the growth and patterning of the cerebellum 1–12.

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., Xu, B., 2008. Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. Cell 134, 175–187. doi:10.1016/j.cell.2008.05.045

Andreassi, C., Riccio, A., 2009. To localize or not to localize: mRNA fate is in 3′ UTR ends. Trends in Cell Biology.

Barski, J.J., Dethleffsen, K., Meyer, M., 2000. Cre recombinase expression in cerebellar Purkinje cells. genesis 28, 93–98.

Beaudoing, E., Gautheret, D., 2001. Identification of alternate polyadenylation sites and analysis of their tissue distribution using EST data. Genome Research 11, 1520–1526. doi:10.1101/gr.190501

Ben-Arie, N., Bellen, H.J., Armstrong, D.L., McCall, A.E., Gordadze, P.R., Guo, Q., Matzuk, M.M., Zoghbi, H.Y., 1997. Math1 is essential for genesis of cerebellar granule neurons. Nature 390, 169–172. doi:10.1038/36579

Berg, M.G., Singh, L.N., Younis, I., Liu, Q., Pinto, A.M., Kaida, D., Zhang, Z., Cho, S., Sherrill-Mix, S., Wan, L., Dreyfuss, G., 2012. U1 snRNP Determines mRNA Length and Regulates Isoform Expression. Cell 150, 53–64. doi:10.1016/j.cell.2012.05.029

Berkovits, B.D., Mayr, C., 2015. Alternative 3' UTRs act as scaffolds to regulate membrane protein localization. Nature 522, 363–367. doi:10.1038/nature14321

Boillée, S., Vande Velde, C., Cleveland, D.W., 2006. ALS: A Disease of Motor Neurons and Their Nonneuronal Neighbors. Neuron 52, 39–59. doi:10.1016/j.neuron.2006.09.018

Boutet, S.C., Cheung, T.H., Quach, N.L., Liu, L., Prescott, S.L., Edalati, A., Iori, K., Rando, T.A., 2012. Alternative Polyadenylation Mediates MicroRNA Regulation of Muscle Stem Cell Function. Stem Cell 10, 327–336. doi:10.1016/j.stem.2012.01.017

Brown, K.M., Gilmartin, G.M., 2003. A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. Molecular Cell 12, 1467–1476.

Buckner, R.L., 2013. The Cerebellum and Cognitive Function: 25 Years of Insight from Anatomy and Neuroimaging. Neuron 80, 807–815. doi:10.1016/j.neuron.2013.10.044

Butts, T., Green, M.J., Wingate, R.J.T., 2014. Development of the cerebellum: simple steps to make a 'little brain'. Development 141, 4031–4041.

doi:10.1242/dev.106559
Chan, S.L., Huppertz, I., Yao, C., Weng, L., Moresco, J.J., Yates, J.R., III, Ule, J., Manley, J.L., Shi, Y., 2014. CPSF30 and Wdr33 directly bind to AAUAAA in mammalian mRNA 3′ processing. Genes & Development 28, 2370–2380. doi:10.1101/gad.250993.114
Chang, T.-H., Huang, H.-Y., Hsu, J.B.-K., Weng, S.-L., Horng, J.-T., Huang, H.-D., 2013. An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs. BMC Bioinformatics 14 Suppl 2, S4. doi:10.1186/1471-2105-14-S2-S4
Cheng, L.-C., Pastrana, E., Tavazoie, M., Doetsch, F., 2009. miR-124 regulates adult neurogenesis in the subventricular zone stem cell niche. Nat Neurosci 12, 399–408. doi:10.1038/nn.2294
Connelly, S., Manley, J.L., 1988. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. Genes & Development 2, 440–452.
Darnell, J.E., 2011. RNA. Life's Indispensable Molecule. Cold Spring Harbor Laboratory Press.
David, G., Grandinetti, K.B., Finnerty, P.M., Simpson, N., Chu, G.C., Depinho, R.A., 2008. Specific requirement of the chromatin modifier mSin3B in cell cycle exit and cellular differentiation. Proceedings of the National Academy of Sciences 105, 4168–4172. doi:10.1073/pnas.0710285105
Davis, M.J., Hanson, K.A., Clark, F., Fink, J.L., Zhang, F., Kasukawa, T., Kai, C., Kawai, J., Carninci, P., Hayashizaki, Y., Teasdale, R.D., 2006. Differential use of signal peptides and membrane domains is a common occurrence in the protein output of transcriptional units. PLoS Genet 2, e46. doi:10.1371/journal.pgen.0020046
de Klerk, E., Venema, A., Anvar, S.Y., Goeman, J.J., Hu, O., Trollet, C., Dickson, G., Dunnen, den, J.T., van der Maarel, S.M., Raz, V., 't Hoen, P.A.C., 2012. Poly(A) binding protein nuclear 1 levels affect alternative polyadenylation. Nucleic Acids Research 40, 9089–9101. doi:10.1093/nar/gks655
de Vries, H., Rüegsegger, U., Hübner, W., Friedlein, A., Langen, H., Keller, W., 2000. Human pre-mRNA cleavage factor II(m) contains homologs of yeast proteins and bridges two other cleavage factors. The EMBO Journal 19, 5895–5904. doi:10.1093/emboj/19.21.5895
Derti, A., Garrett-Engele, P., MacIsaac, K.D., Stevens, R.C., Sriram, S., Chen, R., Rohl, C.A., Johnson, J.M., Babak, T., 2012. A quantitative atlas of polyadenylation in five mammals. Genome Research 22, 1173–1183. doi:10.1101/gr.132563.111
Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G.K., Khatun, J., Williams, B.A., Zaleski, C., Rozowsky, J., Röder, M., Kokocinski, F., Abdelhamid, R.F., Alioto, T., Antoshechkin, I., Baer, M.T., Bar, N.S., Batut, P., Bell, K., Bell, I., Chakrabortty, S., Chen, X., Chrast,

J., Curado, J., Derrien, T., Drenkow, J., Dumais, E., Dumais, J., Duttagupta, R., Falconnet, E., Fastuca, M., Fejes-Toth, K., Ferreira, P., Foissac, S., Fullwood, M.J., Gao, H., Gonzalez, D., Gordon, A., Gunawardena, H., Howald, C., Jha, S., Johnson, R., Kapranov, P., King, B., Kingswood, C., Luo, O.J., Park, E., Persaud, K., Preall, J.B., Ribeca, P., Risk, B., Robyr, D., Sammeth, M., Schaffer, L., See, L.-H., Shahab, A., Skancke, J., Suzuki, A.M., Takahashi, H., Tilgner, H., Trout, D., Walters, N., Wang, H., Wrobel, J., Yu, Y., Ruan, X., Hayashizaki, Y., Harrow, J., Gerstein, M., Hubbard, T., Reymond, A., Antonarakis, S.E., Hannon, G., Giddings, M.C., Ruan, Y., Wold, B., Carninci, P., Guigó, R., Gingeras, T.R., 2012. Landscape of transcription in human cells. Nature 488, 101–108. doi:10.1038/nature11233

Dominguez, D., Tsai, Y.H., Weatheritt, R., Wang, Y., 2016. An extensive program of periodic alternative splicing linked to cell cycle progression. eLife Sciences. doi:10.7554/eLife.10288.001

Doyle, J.P., Dougherty, J.D., Heiman, M., Schmidt, E.F., Stevens, T.R., Ma, G., Bupp, S., Shrestha, P., Shah, R.D., Doughty, M.L., Gong, S., Greengard, P., Heintz, N., 2008. Application of a Translational Profiling Approach for the Comparative Analysis of CNS Cell Types. Cell 135, 749–762. doi:10.1016/j.cell.2008.10.029

Elkon, R., Drost, J., van Haaften, G., Jenal, M., Schrier, M., Vrielink, J.A.O., Agami, R., 2012. E2F mediates enhanced alternative polyadenylation in proliferation. Genome Biol 13, R59. doi:10.1186/gb-2012-13-7-r59

Elkon, R., Ugalde, A.P., Agami, R., 2013. Alternative cleavage and polyadenylation: extent, regulation and function. Nat Rev Genet 14, 496–506. doi:10.1038/nrg3482

Fields, R.D., Itoh, K., 1996. Neural cell adhesion molecules in activity-dependent development and synaptic plasticity. Trends in Neurosciences 19, 473–480. doi:10.1016/S0166-2236(96)30013-1

Flavell, S.W., Kim, T.-K., Gray, J.M., Harmin, D.A., Hemberg, M., Hong, E.J., Markenscoff-Papadimitriou, E., Bear, D.M., Greenberg, M.E., 2008. Genome-wide analysis of MEF2 transcriptional program reveals synaptic target genes and neuronal activity-dependent polyadenylation site selection. Neuron 60, 1022–1038. doi:10.1016/j.neuron.2008.11.029

Floor, S.N., Doudna, J.A., 2016. Tunable protein synthesis by transcript isoforms in human cells. eLife Sciences. doi:10.7554/eLife.10921.001

Ford, L.P., Bagga, P.S., Wilusz, J., 1997. The poly(A) tail inhibits the assembly of a 3“-to-5” exonuclease in an in vitro RNA stability system. Molecular and Cellular Biology 17, 398–406.

Gao, W.O., Heintz, N., Hatten, M.E., 1991. Cerebellar granule cell neurogenesis is regulated by cell-cell interactions in vitro. Neuron 6, 705–715.

Graham, R.R., Kyogoku, C., Sigurdsson, S., Vlasova, I.A., Davies, L.R.L., Baechler, E.C., Plenge, R.M., Koeuth, T., Ortmann, W.A., Hom, G.,

Bauer, J.W., Gillett, C., Burtt, N., Cunninghame Graham, D.S., Onofrio, R., Petri, M., Gunnarsson, I., Svenungsson, E., Rönnblom, L., Nordmark, G., Gregersen, P.K., Moser, K., Gaffney, P.M., Criswell, L.A., Vyse, T.J., Syvänen, A.-C., Bohjanen, P.R., Daly, M.J., Behrens, T.W., Altshuler, D., 2007. Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. Proc. Natl. Acad. Sci. U.S.A. 104, 6758–6763. doi:10.1073/pnas.0701266104

Gruber, A.R., Martin, G., Müller, P., Schmidt, A., Gruber, A.J., Gumienny, R., Mittal, N., Jayachandran, R., Pieters, J., Keller, W., van Nimwegen, E., Zavolan, M., 2014. Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells. Nat Commun 5, 5465. doi:10.1038/ncomms6465

Gupta, I., Clauder-Münster, S., Klaus, B., Järvelin, A.I., Aiyar, R.S., Benes, V., Wilkening, S., Huber, W., Pelechano, V., Steinmetz, L.M., 2014. Alternative polyadenylation diversifies post-transcriptional regulation by selective RNA-protein interactions. Mol. Syst. Biol. 10, 719. doi:10.1002/msb.135068

Han, I., You, Y., Kordower, J.H., Brady, S.T., Morfini, G.A., 2010. Differential vulnerability of neurons in Huntington's disease: the role of cell type-specific features. Journal of Neurochemistry 4, 137–19. doi:10.1111/j.1471-4159.2010.06672.x

Harrison, B.J., Flight, R.M., Gomes, C., Venkat, G., Ellis, S.R., Sankar, U., Twiss, J.L., Rouchka, E.C., Petruska, J.C., 2013. IB4-binding sensory neurons in the adult rat express a novel 3′ UTR-extended isoform of CaMK4that is associated with its localization to axons. J. Comp. Neurol. 522, 308–336. doi:10.1002/cne.23398

Hatten, M.E., 1985. Neuronal regulation of astroglial morphology and proliferation in vitro. The Journal of Cell Biology 100, 384–396.

Hatten, M.E., Heintz, N., 1995. Mechanisms of neural patterning and specification in the developing cerebellum. Annu. Rev. Neurosci. 18, 385–408. doi:10.1146/annurev.ne.18.030195.002125

Hatten, M.E., Roussel, M.F., 2011. Development and cancer of the cerebellum. Trends in Neurosciences 34, 134–142. doi:10.1016/j.tins.2011.01.002

He, M., Liu, Y., Wang, X., Zhang, M.Q., Hannon, G.J., Huang, Z.J., 2012. Cell-type-based analysis of microRNA profiles in the mouse brain. Neuron 73, 35–48. doi:10.1016/j.neuron.2011.11.010

Heiman, M., Schaefer, A., Gong, S., Peterson, J.D., Day, M., Ramsey, K.E., Suárez-Fariñas, M., Schwarz, C., Stephan, D.A., Surmeier, D.J., Greengard, P., Heintz, N., 2008. A Translational Profiling Approach for the Molecular Characterization of CNS Cell Types. Cell 135, 738–748. doi:10.1016/j.cell.2008.10.028

Hilgers, V., Lemke, S.B., Levine, M., 2012. ELAV mediates 3' UTR extension in the Drosophila nervous system. Genes & Development 26, 2259–2264.

doi:10.1101/gad.199653.112
Hobert, O., Carrera, I., Stefanakis, N., 2010. The molecular and gene regulatory signature of a neuron. Trends in Neurosciences 33, 435–445. doi:10.1016/j.tins.2010.05.006
Huang, Y., Carmichael, G.G., 1996. Role of polyadenylation in nucleocytoplasmic transport of mRNA. Molecular and Cellular Biology 16, 1534–1542.
Hwang, H.-W., Park, C.Y., Goodarzi, H., Fak, J.J., Mele, A., Moore, M.J., Saito, Y., Darnell, R.B., 2016. PAPERCLIP Identifies MicroRNA Targets and a Role of CstF64/64tau in Promoting Non-canonical poly(A) Site Usage. CellReports 15, 423–435. doi:10.1016/j.celrep.2016.03.023
Jaffe, A.E., Shin, J., Collado-Torres, L., Leek, J.T., Tao, R., Li, C., Gao, Y., Jia, Y., Maher, B.J., Hyde, T.M., Kleinman, J.E., Weinberger, D.R., 2015. Developmental regulation of human cortex transcription and its clinical relevance at single base resolution. Nat Neurosci 18, 154–161. doi:10.1038/nn.3898
Jenal, M., Elkon, R., Loayza-Puch, F., van Haaften, G., Kühn, U., Menzies, F.M., Vrielink, J.A.F.O., Bos, A.J., Drost, J., Rooijers, K., Rubinsztein, D.C., Agami, R., 2012. The Poly(A)-Binding Protein Nuclear 1 Suppresses Alternative Cleavage and Polyadenylation Sites. Cell 149, 538–553. doi:10.1016/j.cell.2012.03.022
Ji, Z., Lee, J.Y., Pan, Z., Jiang, B., Tian, B., 2009. Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proceedings of the National Academy of Sciences 106, 7028–7033. doi:10.1073/pnas.0900028106
Ji, Z., Tian, B., 2009. Reprogramming of 3′ Untranslated Regions of mRNAs by Alternative Polyadenylation in Generation of Pluripotent Stem Cells from Different Cell Types. PLoS ONE 4, e8419–13. doi:10.1371/journal.pone.0008419
Kandel, E.R., Schwartz, J.H., Jessell, T.M., Siegelbaum, S.A., Hudspeth, A.J., 2012. Principles of Neural Science, Fifth Edition. McGraw-Hill.
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., Vågbø, C.B., Kusśnierczyk, A., Klungland, A., Darnell, J.E., Jr., Darnell, R.B., 2015. A majority of m 6A residues are in the last exons, allowing the potential for 3′ UTR regulation. Genes & Development 29, 2037–2053. doi:10.1101/gad.269415.115
Kerwitz, Y., Kühn, U., Lilie, H., Knoth, A., Scheuermann, T., Friedrich, H., Schwarz, E., Wahle, E., 2003. Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. The EMBO Journal 22, 3705–3714. doi:10.1093/emboj/cdg347
Kulicke, R., Fenster, R.J., Greengard, P., Heintz, N., Heiman, M., 2014. Cell type–specific mRNA purification by translating ribosome affinity

purification (TRAP). Nature Protocols 9, 1282–1291. doi:10.1038/nprot.2014.085
Lau, A.G., Irier, H.A., Gu, J., Tian, D., Ku, L., Liu, G., Xia, M., Fritsch, B., Zheng, J.Q., Dingledine, R., Xu, B., Lu, B., Feng, Y., 2010. Distinct 3'UTRs differentially regulate activity-dependent translation of brain-derived neurotrophic factor (BDNF). Proceedings of the National Academy of Sciences 107, 15945–15950. doi:10.1073/pnas.1002929107
Leto, K., Arancillo, M., Becker, E.B.E., Buffo, A., Chiang, C., Ding, B., Dobyns, W.B., Dusart, I., Haldipur, P., Hatten, M.E., Hoshino, M., Joyner, A.L., Kano, M., Kilpatrick, D.L., Koibuchi, N., Marino, S., Martinez, S., Millen, K.J., Millner, T.O., Miyata, T., Parmigiani, E., Schilling, K., Sekerková, G., Sillitoe, R.V., Sotelo, C., Uesaka, N., Wefers, A., Wingate, R.J.T., Hawkes, R., 2015. Consensus Paper: Cerebellar Development. Cerebellum 1–40. doi:10.1007/s12311-015-0724-2
Li, W., You, B., Hoque, M., Zheng, D., Luo, W., Ji, Z., Park, J.Y., Gunderson, S.I., Kalsotra, A., Manley, J.L., Tian, B., 2015. Systematic profiling of poly(a)+ transcripts modulated by core 3' end processing and splicing factors reveals regulatory rules of alternative cleavage and polyadenylation. PLoS Genet 11, e1005166. doi:10.1371/journal.pgen.1005166
Lianoglou, S., Garg, V., Yang, J.L., Leslie, C.S., Mayr, C., 2013. Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. Genes & Development 27, 2380–2396. doi:10.1101/gad.229328.113
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., Darnell, J.C., Darnell, R.B., 2008. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456, 464–469. doi:10.1038/nature07488
Liu, D., Brockman, J.M., Dass, B., Hutchins, L.N., Singh, P., McCarrey, J.R., MacDonald, C.C., Graber, J.H., 2006. Systematic variation in mRNA 3'-processing signals during mouse spermatogenesis. Nucleic Acids Research 35, 234–246. doi:10.1093/nar/gkl919
Luo, W., Ji, Z., Pan, Z., You, B., Hoque, M., Li, W., Gunderson, S.I., Tian, B., 2013. The Conserved Intronic Cleavage and Polyadenylation Site of CstF-77 Gene Imparts Control of 3′ End Processing Activity through Feedback Autoregulation and by U1 snRNP. PLoS Genet 9, e1003613–14. doi:10.1371/journal.pgen.1003613
MacDonald, C.C., Wilusz, J., Shenk, T., 1994. The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. Molecular and Cellular Biology 14, 6647–6654.
MacDonald, G., Nalvarte, I., Smirnova, T., Vecchi, M., Aceto, N., Dolemeyer, A., Frei, A., Lienhard, S., Wyckoff, J., Hess, D., Seebacher, J., Keusch, J.J., Gut, H., Salaun, D., Mazzarol, G., Disalvatore, D., Bentires-Alj, M., Di

Fiore, P.P., Badache, A., Hynes, N.E., 2014. Memo is a copper-dependent redox protein with an essential role in migration and metastasis. Sci. Signal. 7, ra56. doi:10.1126/scisignal.2004870
Machold, R., Fishell, G., 2005. Math1 Is Expressed in Temporally Discrete Pools of Cerebellar Rhombic-Lip Neural Progenitors. Neuron 48, 17–24. doi:10.1016/j.neuron.2005.08.028
Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., Trombetta, J.J., Weitz, D.A., Sanes, J.R., Shalek, A.K., Regev, A., McCarroll, S.A., 2015. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 161, 1202–1214. doi:10.1016/j.cell.2015.05.002
Mairet-Coello, G., Tury, A., DiCicco-Bloom, E., 2009. Insulin-Like Growth Factor-1 Promotes G1/S Cell Cycle Progression through Bidirectional Regulation of Cyclins and Cyclin-Dependent Kinase Inhibitors via the Phosphatidylinositol 3-Kinase/Akt Pathway in Developing Rat Cerebral Cortex. J. Neurosci. 29, 775–788. doi:10.1523/JNEUROSCI.1700-08.2009
Makeyev, E.V., Zhang, J., Carrasco, M.A., Maniatis, T., 2007. The MicroRNA miR-124 Promotes Neuronal Differentiation by Triggering Brain-Specific Alternative Pre-mRNA Splicing. Molecular Cell 27, 435–448. doi:10.1016/j.molcel.2007.07.015
Mali, P., Esvelt, K.M., Church, G.M., 2013. Cas9 as a versatile tool for engineering biology. Nat Meth 10, 957–963. doi:10.1038/nmeth.2649
Mandel, C.R., Kaneko, S., Zhang, H., Gebauer, D., Vethantham, V., Manley, J.L., Tong, L., 2006. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. Nature 444, 953–956. doi:10.1038/nature05363
Marone, R., Hess, D., Dankort, D., Muller, W.J., Hynes, N.E., Badache, A., 2004. Memo mediates ErbB2-driven cell motility. Nature Cell Biology 6, 515–522. doi:10.1038/ncb1134
Martin, G., Gruber, A.R., Keller, W., Zavolan, M., 2012a. 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. CellReports 1, 753–763. doi:10.1016/j.celrep.2012.05.003
Martin, G., Gruber, A.R., Keller, W., Zavolan, M., 2012b. 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. CellReports 1, 753–763. doi:10.1016/j.celrep.2012.05.003
Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17, pp. 10–12. doi:10.14806/ej.17.1.200
Masamha, C.P., Xia, Z., Yang, J., Albrecht, T.R., Li, M., Shyu, A.-B., Li, W., Wagner, E.J., 2014. CFIm25 links alternative polyadenylation to

glioblastoma tumour suppression. Nature 510, 412–416. doi:10.1038/nature13261
Masuda, A., Takeda, J.-I., 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 & Development 29, 1045–1057. doi:10.1101/gad.255737.114
Matei, V., Pauley, S., Kaing, S., Rowitch, D., Beisel, K.W., Morris, K., Feng, F., Jones, K., Lee, J., Fritzsch, B., 2005. Smaller inner ear sensory epithelia in Neurog 1 null mice are related to earlier hair cell cycle exit. Dev. Dyn. 234, 633–650. doi:10.1002/dvdy.20551
Mayr, C., 2016. Evolution and Biological Roles of Alternative 3′UTRs. Trends in Cell Biology 26, 227–237. doi:10.1016/j.tcb.2015.10.012
Mayr, C., Bartel, D.P., 2009. Widespread Shortening of 3′UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells. Cell 138, 673–684. doi:10.1016/j.cell.2009.06.016
Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S., Heintz, N., 2012. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell 151, 1417–1430. doi:10.1016/j.cell.2012.11.022
Mili, S., Steitz, J.A., 2004. Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. RNA 10, 1692–1694. doi:10.1261/rna.7151404
Miura, P., Sanfilippo, P., Shenker, S., Lai, E.C., 2014. Alternative polyadenylation in the nervous system: To what lengths will 3′ UTR extensions take us? BioEssays 36, 766–777. doi:10.1002/bies.201300174
Miura, P., Shenker, S., Andreu-Agullo, C., Westholm, J.O., Lai, E.C., 2013. Widespread and extensive lengthening of 3' UTRs in the mammalian brain. Genome Research 23, 812–825. doi:10.1101/gr.146886.112
Miyata, T., Maeda, T., Lee, J.E., 1999. NeuroD is required for differentiation of the granule cells in the cerebellum and hippocampus. Genes & Development 13, 1647–1652.
Moore, M.J., Scheel, T.K.H., Luna, J.M., Park, C.Y., Fak, J.J., Nishiuchi, E., Rice, C.M., Darnell, R.B., 2015. miRNA-target chimeras reveal miRNA 3'-end pairing as a major determinant of Argonaute target specificity. Nat Commun 6, 8864. doi:10.1038/ncomms9864
Moore, M.J., Zhang, C., Gantman, E.C., Mele, A., Darnell, J.C., Darnell, R.B., 2014. Mapping Argonaute and conventional RNA-binding protein interactions with RNA at single-nucleotide resolution using HITS-CLIP and CIMS analysis. Nature Protocols 9, 263–293. doi:doi:10.1038/nprot.2014.012
Mueller, A.A., Cheung, T.H., Rando, T.A., 2013. All's well that ends well: alternative polyadenylation and its implications for stem cell biology. Current Opinion in Cell Biology 25, 222–232. doi:10.1016/j.ceb.2012.12.008
Müller-McNicoll, M., Botti, V., de Jesus Domingues, A.M., Brandl, H.,

Schwich, O.D., Steiner, M.C., Curk, T., Poser, I., Zarnack, K., Neugebauer, K.M., 2016. SR proteins are NXF1 adaptors that link alternative RNA processing to mRNA export. Genes & Development 30, 553–566. doi:10.1101/gad.276477.115

Nam, J.-W., Rissland, O.S., Koppstein, D., Abreu-Goodger, C., Jan, C.H., Agarwal, V., Yildirim, M.A., Rodriguez, A., Bartel, D.P., 2014. Global Analyses of the Effect of Different Cellular Contexts on MicroRNA Targeting. Molecular Cell 53, 1031–1043. doi:10.1016/j.molcel.2014.02.013

Neve, J., Burger, K., Li, W., Hoque, M., Patel, R., Tian, B., Gullerova, M., Furger, A., 2016. Subcellular RNA profiling links splicing and nuclear DICER1 to alternative cleavage and polyadenylation. Genome Research 26, 24–35. doi:10.1101/gr.193995.115

Okano, H.J., Darnell, R.B., 1997. A hierarchy of Hu RNA binding proteins in developing and adult neurons. J. Neurosci. 17, 3024–3037.

Okaty, B.W., Sugino, K., Nelson, S.B., 2011. A Quantitative Comparison of Cell-Type-Specific Microarray Gene Expression Profiling Methods in the Mouse Brain. PLoS ONE 6, e16493. doi:10.1371/journal.pone.0016493.s015

Oktaba, K., Zhang, W., Lotz, T.S., Jun, D.J., Lemke, S.B., Ng, S.P., Esposito, E., Levine, M., Hilgers, V., 2015. ELAV Links Paused Pol II to Alternative Polyadenylation in the Drosophila Nervous System. Molecular Cell 57, 341–348. doi:10.1016/j.molcel.2014.11.024

Orr, H.T., 2000. The Ins and Outs of a Polyglutamine Neurodegenerative Disease: Spinocerebellar Ataxia Type 1 (SCA1). Neurobiology of Disease 7, 129–134. doi:10.1006/nbdi.2000.0305

Ozsolak, F., Kapranov, P., Foissac, S., Kim, S.W., Fishilevich, E., Monaghan, A.P., John, B., Milos, P.M., 2010. Comprehensive Polyadenylation Site Maps in Yeast and Human Reveal Pervasive Alternative Polyadenylation. Cell 143, 1018–1029. doi:10.1016/j.cell.2010.11.020

Pan, Z., Zhang, H., Hague, L.K., Lee, J.Y., Lutz, C.S., Tian, B., 2006. An intronic polyadenylation site in human and mouse CstF-77 genes suggests an evolutionarily conserved regulatory mechanism. Gene 366, 325–334. doi:10.1016/j.gene.2005.09.024

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., Moreira, A., 2011. RNA polymerase II kinetics in polo polyadenylation signal selection. The EMBO Journal 30, 2431–2444. doi:10.1038/emboj.2011.156

Proudfoot, N.J., 2016. Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. Science 352, aad9926–aad9926. doi:10.1126/science.aad9926

Prudencio, M., Belzil, V.V., Batra, R., Ross, C.A., Gendron, T.F., Pregent, L.J., Murray, M.E., Overstreet, K.K., Piazza-Johnston, A.E., Desaro, P.,

Bieniek, K.F., DeTure, M., Lee, W.C., Biendarra, S.M., Davis, M.D., Baker, M.C., Perkerson, R.B., van Blitterswijk, M., Stetler, C.T., Rademakers, R., Link, C.D., Dickson, D.W., Boylan, K.B., Li, H., Petrucelli, L., 2015. Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. Nat Neurosci 18, 1175–1182. doi:10.1038/nn.4065

Richard, P., Manley, J.L., 2009. Transcription termination by nuclear RNA polymerases. Genes & Development 23, 1247–1269. doi:10.1101/gad.1792809

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. doi:10.1093/bioinformatics/btp616

Rozenblatt-Rosen, O., Nagaike, T., Francis, J.M., Kaneko, S., Glatt, K.A., Hughes, C.M., LaFramboise, T., Manley, J.L., Meyerson, M., 2009. The tumor suppressor Cdc73 functionally associates with CPSF and CstF 3' mRNA processing factors. Proceedings of the National Academy of Sciences 106, 755–760. doi:10.1073/pnas.0812023106

Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A., Burge, C.B., 2008. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science 320, 1643–1647. doi:10.1126/science.1155390

Sanz, E., Yang, L., Su, T., Morris, D.R., McKnight, G.S., Amieux, P.S., 2009. Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. Proceedings of the National Academy of Sciences 106, 13939–13944. doi:10.1073/pnas.0907143106

Schönemann, L., Kühn, U., Martin, G., Schäfer, P., Gruber, A.R., Keller, W., Zavolan, M., Wahle, E., 2014. Reconstitution of CPSF active in polyadenylation: recognition of the polyadenylation signal by WDR33. Genes & Development 28, 2381–2393. doi:10.1101/gad.250985.114

Shalek, A.K., Satija, R., Adiconis, X., Gertner, R.S., Gaublomme, J.T., Raychowdhury, R., Schwartz, S., Yosef, N., Malboeuf, C., Lu, D., Trombetta, J.J., Gennert, D., Gnirke, A., Goren, A., Hacohen, N., Levin, J.Z., Park, H., Regev, A., 2013. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 498, 236–240. doi:10.1038/nature12172

Shepard, P.J., Choi, E.-A., Lu, J., Flanagan, L.A., Hertel, K.J., Shi, Y., n.d. Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq.

Shepard, P.J., Choi, E.A., Lu, J., Flanagan, L.A., Hertel, K.J., Shi, Y., 2011. Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. RNA 17, 761–772. doi:10.1261/rna.2581711

Shi, Y., 2012. Alternative polyadenylation: new insights from global analyses. RNA 18, 2105–2117. doi:10.1261/rna.035899.112

Shi, Y., Di Giammartino, D.C., Taylor, D., Sarkeshik, A., Rice, W.J., Yates,

J.R., III, Frank, J., Manley, J.L., 2009. Molecular Architecture of the Human Pre-mRNA 3′ Processing Complex. Molecular Cell 33, 365–376. doi:10.1016/j.molcel.2008.12.028
Shi, Y., Manley, J.L., 2015. The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. Genes & Development 29, 889–897. doi:10.1101/gad.261974.115
Smibert, P., Miura, P., Westholm, J.O., Shenker, S., May, G., Duff, M.O., Zhang, D., Eads, B.D., Carlson, J., Brown, J.B., Eisman, R.C., Andrews, J., Kaufman, T., Cherbas, P., Celniker, S.E., Graveley, B.R., Lai, E.C., 2012. Global Patterns of Tissue-Specific Alternative Polyadenylation in Drosophila. CellReports 1, 277–289. doi:10.1016/j.celrep.2012.01.001
Solecki, D.J., Liu, X.L., Tomoda, T., Fang, Y., Hatten, M.E., 2001. Activated Notch2 signaling inhibits differentiation of cerebellar granule neuron precursors by maintaining proliferation. Neuron 31, 557–568.
Sorokin, A.V., Chen, J., 2012. MEMO1, a new IRS1-interacting protein, induces epithelial–mesenchymal transition in mammary epithelial cells 32, 3130–3138. doi:10.1038/onc.2012.327
Spies, N., Burge, C.B., Bartel, D.P., 2013. 3' UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. Genome Research 23, 2078–2090. doi:10.1101/gr.156919.113
Storey, J.D., Bass, A.J., 2003. Bioconductor's qvalue package Version 1.99. Annals of Statistics.
Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., Bartel, D.P., 2014. Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature 508, 66–71. doi:10.1038/nature13007
Takagaki, Y., Seipelt, R.L., Peterson, M.L., Manley, J.L., 1996. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. Cell 87, 941–952.
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., Burge, C.B., 2016. Distal Alternative Last Exons Localize mRNAs to Neural Projections. Molecular Cell. doi:10.1016/j.molcel.2016.01.020
Tan, C.L., Plotkin, J.L., Venø, M.T., Schimmelmann, von, M., Feinberg, P., Mann, S., Handler, A., Kjems, J., Surmeier, D.J., O'Carroll, D., Greengard, P., Schaefer, A., 2013. MicroRNA-128 governs neuronal excitability and motor behavior in mice. Science 342, 1254–1258. doi:10.1126/science.1244193
Tian, B., Manley, J.L., 2016. Alternative polyadenylation of mRNA precursors. Nat Rev Mol Cell Biol 18, 18–30. doi:10.1038/nrm.2016.116
Tian, B., Pan, Z., Lee, J.Y., 2007. Widespread mRNA polyadenylation events in introns indicate dynamic interplay between polyadenylation and splicing. Genome Research 17, 156–165. doi:10.1101/gr.5532707
Tsai, P.T., Hull, C., Chu, Y., Greene-Colozzi, E., Sadowski, A.R., Leech, J.M.,

Steinberg, J., Crawley, J.N., Regehr, W.G., Sahin, M., 2012. Autistic-like behaviour and cerebellar dysfunction in Purkinje cell Tsc1 mutant mice. Nature 488, 647–651. doi:10.1038/nature11310

Ulitsky, I., Shkumatava, A., Jan, C.H., Subtelny, A.O., Koppstein, D., Bell, G.W., Sive, H., Bartel, D.P., 2012. Extensive alternative polyadenylation during zebrafish development. Genome Research 22, 2054–2066. doi:10.1101/gr.139733.112

Wahle, E., Rüegsegger, U., 1999. 3'-End processing of pre-mRNA in eukaryotes. FEMS Microbiol. Rev. 23, 277–295.

Wallace, V.A., 1999. Purkinje-cell-derived Sonic hedgehog regulates granule neuron precursor cell proliferation in the developing mouse cerebellum. Current Biology 9, 445–448. doi:10.1016/S0960-9822(99)80195-X

Wang, L., Dowell, R.D., Yi, R., 2013a. Genome-wide maps of polyadenylation reveal dynamic mRNA 3'-end formation in mammalian cell lineages. RNA 19, 413–425. doi:10.1261/rna.035360.112

Wang, L., Dowell, R.D., Yi, R., 2013b. Genome-wide maps of polyadenylation reveal dynamic mRNA 3'-end formation in mammalian cell lineages. RNA 19, 413–425. doi:10.1261/rna.035360.112

Wang, V.Y., Rose, M.F., Zoghbi, H.Y., 2005. Math1 Expression Redefines the Rhombic Lip Derivatives and Reveals Novel Lineages within the Brainstem and Cerebellum. Neuron 48, 31–43. doi:10.1016/j.neuron.2005.08.024

West, S., Gromak, N., Proudfoot, N.J., 2004. Human 5′ → 3′ exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature 432, 522–525. doi:10.1038/nature03035

Xia, Z., Donehower, L.A., Cooper, T.A., Neilson, J.R., Wheeler, D.A., Wagner, E.J., Li, W., 2014. Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3′-UTR landscape across seven tumour types. Nat Commun 5, 1–13. doi:10.1038/ncomms6274

Yang, Y., Li, W., Hoque, M., Hou, L., Shen, S., Tian, B., Dynlacht, B.D., 2016. PAF Complex Plays Novel Subunit-Specific Roles in Alternative Cleavage and Polyadenylation. PLoS Genet 12, e1005794–28. doi:10.1371/journal.pgen.1005794

Yao, C., Biesinger, J., Wan, J., Weng, L., Xing, Y., Xie, X., Shi, Y., 2012. Transcriptome-wide analyses of CstF64-RNA interactions in global regulation of mRNA alternative polyadenylation. Proceedings of the National Academy of Sciences 109, 18773–18778. doi:10.1073/pnas.1211101109

Yonaha, M., Proudfoot, N.J., 1999. Specific transcriptional pausing activates polyadenylation in a coupled in vitro system. Molecular Cell 3, 593–600.

Yudin, D., Hanz, S., Yoo, S., Iavnilovitch, E., Willis, D., Gradus, T., Vuppalanchi, D., Segal-Ruder, Y., Ben-Yaakov, K., Hieda, M., Yoneda, Y., Twiss, J.L., Fainzilber, M., 2008. Localized Regulation of Axonal RanGTPase Controls Retrograde Injury Signaling in Peripheral Nerve.

Neuron 59, 241–252. doi:10.1016/j.neuron.2008.05.029
Zhang, Haibo, Lee, J.Y., Tian, B., 2005. Biased alternative polyadenylation in human tissues. Genome Biol 6, R100–13. doi:10.1186/gb-2005-6-12-r100
Zhang, Huimin, Rigo, F., Martinson, H.G., 2015. Poly(A) Signal-Dependent Transcription Termination Occurs through a Conformational Change Mechanism that Does Not Require Cleavage at the Poly(A) Site. Molecular Cell 59, 437–448. doi:10.1016/j.molcel.2015.06.008
Zhu, H., Zhou, H.-L., Hasman, R.A., Lou, H., 2007. Hu proteins regulate polyadenylation by blocking sites containing U-rich sequences. J. Biol. Chem. 282, 2203–2210. doi:10.1074/jbc.M609349200
Zhu, X., Girardo, D., Govek, E.-E., John, K., Mellén, M., Tamayo, P., Mesirov, J.P., Hatten, M.E., 2016. Role of Tet1/3 Genes and Chromatin Remodeling Genes in Cerebellar Circuit Formation. Neuron 89, 100–112. doi:10.1016/j.neuron.2015.11.030