Title
Identification and Characterization of PTBP1-Regulated Splicing Events During Neuronal Differentiation.

Author
Linares, Anthony

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Identification and Characterization of PTBP1-Regulated Splicing Events During Neuronal Differentiation.

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular Biology

by

Anthony James Linares

2015

ABSTRACT OF THE DISSERTATION

Identification and Characterization of PTBP1 Regulated

Splicing Events During Neuronal Differentiation.

by

Anthony James Linares

Doctor of Philosophy in Molecular Biology

University of California, Los Angeles, 2015

Professor Douglas L. Black, Chair

Alternative splicing is an important form of gene regulation during tissue development. In the mammalian nervous system, complex splicing patterns produce new protein isoforms with different structures and functions specific to neurons. These splicing patterns are regulated in a temporal and cell-specific manner by the expression of specialized pre-mRNA binding proteins (RBPs). The polypyrimidine tract binding (PTB) proteins, PTBP1 and PTBP2, are RBPs that control extensive programs of alternative splicing during neuronal differentiation and maturation. However, the neuronal splicing programs regulated by PTBP1 and PTBP2 are not fully characterized, and the cellular function of their regulated isoforms remains largely unknown. Here we define the PTBP1 splicing program during *in vitro* neuronal differentiation and show that PTBP1 guides developmental gene expression programs by regulating an alternative exon in the homeodomain transcription factor Pbx1.

We use RNA-sequencing to identify thousands of alternative exons that are differentially spliced as mouse embryonic stem cells (ESC) differentiate into neuronal progenitor cells (NPC) and then into neurons, during the transition from PTBP1 expression to PTBP2. We then define the subset of these neuronal exons controlled by PTBP1 in ESCs and NPCs by RNA-sequencing analysis after PTBP1 depletion and PTBP1 crosslinking-immunoprecipitation (iCLIP-seq). Among these targets, we find that PTBP1 represses splicing of exon 7 in the Pbx1 transcript leading to expression of the Pbx1b isoform in ESCs and NPCs. To study the consequences of this Pbx1 splicing switch, we use the CRISPR-Cas9 technology to induce or eliminate exon 7 splicing in different cell types. We delete regulatory elements in Pbx1 intron 6 to force exon 7 splicing and Pbx1a expression in ESCs. We find that the early expression of Pbx1a enhances expression of a defined group of neurogenic genes, many of which were previously identified as Pbx1 targets. Cas9-targeted deletion of Pbx1 exon 7 forces expression of Pbx1b, with the loss of Pbx1a, in differentiated motor neurons. We find that these cells differentiate normally but exhibit altered expression of Pbx1 target genes. These data demonstrate that one component of the PTBP1 regulatory network is to control the activity of the transcription factor Pbx1 and thus to alter the transcriptional program of neuronal differentiation.

The dissertation of Anthony James Linares is approved.

William Edward Lowry

Bennett G. Novitch

Douglas L. Black, Committee Chair

University of California, Los Angeles

2015

This dissertation is dedicated to my parents.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my parents for everything they have done for me. I would like to thank my girlfriend for her unwavering support. I would like to thank fellow Black lab members for their guidance. I would like to thank Douglas Black for his mentorship and patience. There are many other friends, colleagues, and mentors that I am thankful for.

Chapter 2 is a manuscript that is in preparation:
Anthony J Linares, Chia-Ho Lin, Katrina L Adams, Bennett G Novitch, Douglas L Black (2015). The splicing regulator PTBP1 controls the activity of the homeodomain transcription factor Pbx1 during neuronal differentiation. *In preparation.*

This work was supported by UCLA MSTP Training Grant, NIH T32 NS048004, and NIH T32 NS07449. D.L. Black is an investigator of the Howard Hughes Medical Institute.

VITA

EDUCATION

09/10 to present	Molecular Biology Interdepartmental Program at UCLA, Los Angeles, California PhD candidate
08/08 to present	David Geffen School of Medicine, Medical Scientist Training Program.at UCLA, Los Angeles, California. MD candidate
08/04 to 05/08	Pomona College, Claremont, CA B.A., Chemistry, 05/08

PUBLICATIONS

1) Han A, Stoilov P, Linares AJ, Zhou Y, Fu XD, Black DL. De Novo Prediction of PTBP1 Binding and Splicing Targets Reveals Unexpected Features of Its RNA Recognition and Function. PLoS Comput. Biol. 10(1):e1003442, 2014. PMCID: PMC3907290.
2) Friedland GD, Linares AJ, Smith CA, Kortemme T. A simple model of backbone flexibility improves modeling of side-chain conformational variability. J. Mol. Biol. 380:757-775, 2008. PMCID: PMC3574579.
3) Garza-López RA, Linares A, Yoo A, Evans G, Kozak, JJ. Invariance relations for random walks on simple cubic lattices. Chem. Phys. Lett. 421:287-294, 2006.

PRESENTATIONS

1) Linares AJ and Black DL. Functional characterization of PTBP1 regulated alternative splicing events during neuronal differentiation. Oral presentation at the RNA Society Annual Meeting in Quebec City, Canada June 2014.

CHAPTER 1
Introduction

Alternative pre-messenger RNA splicing

Alternative pre-messenger (pre-mRNA) splicing is an essential form of gene regulation during tissue development (Black, 2003). It is estimated that greater than 95% of multi-exon mammalian genes undergo this process, which allows exons, portions of exons, or introns to be included or excluded from the final mRNA transcript (Pan et al., 2008; Wang et al., 2008).

From a single multi-exon gene, many transcripts can be produced that contribute to transcriptional and proteomic complexity (Black, 2003; Braunschweig et al., 2013; Kalsotra and Cooper, 2011). The most common splicing pattern is the inclusion or exclusion of cassette exons. Others patterns include mutually exclusive exons, alternative 5' and 3' splice sites, alternative promoters, alternative poly-A sites, and retained introns. Often splicing will produce multiple mRNA products that encode protein isoforms with different cellular functions. However, the biological significance of alternatively spliced products is not well understood.

In addition to regulating proteomic diversity, alternative splicing can regulate mRNA stability. If a splicing change introduces a premature termination codon (PTC) more than 50 nucleotides from an exon junction complex (EJC), the mRNA is targeted to the non-sense mediated decay (NMD) pathway (Schoenberg and Maquat, 2012). Like alternative splicing induced NMD (AS-NMD), intron retention can lead to the nuclear retention and decay of the mRNA (Yap and Makeyev, 2013).

Alternative splicing regulation

Alternative pre-mRNA splicing occurs when the spliceosome changes splice site choice (Black, 2003; Chen and Manley, 2009; Matlin et al., 2005). The spliceosome, a large ribonuclearprotein (RNP) complex, catalyzes intron removal and exon ligation from a pre-mRNA. To conduct the splicing reactions, the major spliceosome, which is composed of five small nuclear ribonucleoproteins (snRNPs; U1, U2, U4, U5, and U6) and auxiliary factors, assembles on specific pre-mRNA signals at intron/exon junctions. The signals required for splicing activity include the 5' splice sites, the branch point sequence, the pyrimidine tract, and the 3' splice site. During early spliceosome assembly, the U1 snRNP binds to the 5' splice site and the auxiliary factors SF1 and U2AF bind 3' splice elements. The spliceosome then transitions through several complexes to catalyze the splicing reactions.

Specialized RNA binding proteins (RBPs) and specific RNA sequence elements influence splice site choice (Black, 2003; Chen and Manley, 2009; Fu and Ares, 2014; Matlin et al., 2005). There are many RBPs that regulate splicing such as the serine/arginine-rich (SR) and heterogeneous nuclear ribonucleoproteins (hnRNP) group of proteins. These factors recognize *cis*-regulatory elements in the pre-mRNA to interact with the splicing machinery. The *cis*-regulatory elements that recruited specific RBPs are categorized based on their location and action. They include exonic splicing enhancers (ESEs), exonic splicing silencers (ESSs), intronic splicing enhancers (ISEs), and intronic splicing silencers (ISSs). For example, the SR proteins are known to bind ESEs to activate the splicing, while the hnRNP proteins are known to bind ISSs to repress the splicing. However, SR proteins can repress splicing and hnRNP proteins can activate splicing though a variety of mechanisms. Notably, many splicing factors, including the

SR and hnRNP proteins, exhibit tissue-specific expression patterns, which contribute to the complex splicing patterns observed during development.

Polypyrimidine tract binding (PTB) proteins

Members of the hnRNP group include the polypyrimidine tract binding (PTB) proteins (Keppetipola et al., 2012). PTBP1, one of the first known splicing regulators, is widely expressed but generally absent in muscle and neurons. In mammals, there are two PTBP1 paralogs that show more restricted expression patterns: PTBP2 is present in neurons, testis, and other cell types (Markovtsov et al., 2000; Polydorides et al., 2000), whereas PTBP3 (Rod1) is expressed primarily in hematopoietic cells (Yamamoto et al., 1999).

PTBP1 was first observed in HeLa nuclear extract, where it bound the polypyrimidine tract of the adenovirus pre-mRNA (Garcia-Blanco et al., 1989). It was proposed that PTBP1 assisted with 3' splice site recognition, but later work demonstrated that PTBP1 negatively affected this process to repress spliceosome assembly (Keppetipola et al., 2012). Early studies identified PTBP1-repressed exons specific to neuronal and muscle tissues. These targets include exons in the α- and β-tropomyosin, α-actinin, c-src, $GABA_A$ receptor-γ2, and calcitonin/CGRP transcripts (Ashiya and Grabowski, 1997; Chan and Black, 1997; Gooding et al., 1998; Lou et al., 1999; Mulligan et al., 1992; Southby et al., 1999).

Biochemical and structural studies have provided insights into the mechanisms of PTBP1 splicing repression. To bind RNA, PTBP1 has four RNA recognition motifs (RRMs) that each interact with pyrimidine rich elements (Oberstrass et al., 2005). Often multiple PTBP1 binding sites are required for PTBP1 repression (Keppetipola et al., 2012). PTBP1 binding sites are located within the regulated exon and in the adjacent introns. Intronic regulatory elements are

found at the polypyrimidine tract of the 3' splice site, upstream of the branchpoint, and downstream of the 5' splice site. Studies on model exons have shown that PTBP1 can prevent early spliceosome assembly by outcompeting U2AF binding at the polypyrimidine tract of the 3' splice site (Spellman and Smith, 2006). However, PTBP1 can also prevent different stages of spliceosome assembly by binding upstream of the branchpoint sequence and near the 5' splice site (Sharma et al., 2005; 2008). The mechanisms of PTBP1 splicing repression are diverse and likely depend on the regulated exons.

Genome-wide studies have provided additional insights into PTBP1 splicing regulation. Splicing sensitive microarrays and RNA-sequencing have identified large networks of splicing targets regulated by PTBP1. Interestingly, these data demonstrated that PTBP1 positively regulates exon splicing for a set of targets (Boutz et al., 2007; Llorian et al., 2010). The mechanisms by which PTBP1 promotes exon inclusion, however, are not well understood. For other splicing regulators like the neuro-oncological ventral antigen (Nova) and RNA-binding fox (Rbfox) proteins, binding sites downstream of the regulated exon promote exon splicing, while binding sites upstream repress exon splicing (Ule et al., 2006; Yeo et al., 2009).

Genome-wide models of PTBP1 binding support the idea that PTBP1 regulates splicing through diverse mechanisms. One study identified PTBP1 bound RNA in HeLa cells with crosslinking-immunoprecipitation sequencing (CLIP-seq) (Xue et al., 2009). An RNA binding map was constructed, relating PTBP1 binding with PTBP1 splicing regulation, by integrating PTBP1 functional and binding data. Another study combined PTBP1 functional data with bioinformatics analyses to construct a different RNA binding map (Llorian et al., 2010). Both studies found that PTBP1 often binds upstream or downstream of a regulated exon to repress splicing. However, the two models disagreed about the location of PTBP1 binding for splicing

activation. These results are not mutually exclusive and likely reflect the diverse modes of PTBP1 regulation.

Splicing regulation during neuronal development

Alternative splicing contributes to the transcriptional and proteomic complexity needed for neuronal lineage commitment, differentiation, and maturation. The first example of alternative splicing in neurons was found when the pre-mRNA encoding calcitonin in the thyroid was differentially spliced to produce the neuropeptide CGRP in neurons (Amara et al., 1982). Since then thousands of alternatively spliced isoforms specific to the mammalian nervous system have been identified. These events are particularly abundant in synaptic proteins, where the neurexin genes are alternatively spliced to produce over thousands of different isoforms (Ullrich et al., 1995). More recently, high-throughput sequencing experiments have shown that the vertebrate and mammalian nervous systems exhibit complex and conserved splicing patterns (Barbosa-Morais et al., 2012; Merkin et al., 2012).

Neuronal splicing patterns are largely determined by the tissue-specific expression of splicing factors that recognize specific *cis*-regulatory sequences (Norris and Calarco, 2012; Zheng and Black, 2013). In addition to the PTB proteins, there are other RBPs that regulate splicing programs in the developing and adult brain, including members of the Nova, Rbfox, serine/arginine repetitive matrix (SRRM) and other protein families. These splicing regulators control the expression of isoforms essential for neural development and maturation. However, the complex splicing programs regulated by neuronal splicing factors are incompletely characterized, and the cellular functions of the regulated isoforms are poorly understood.

The Nova proteins were first discovered as antigens in the paraneoplastic opsoclonus-myoclonus ataxia (POMA) syndrome and were subsequently found to be exclusively expressed in the nervous system (Posner and Darnell, 1993). To bind RNA, the proteins possess three KH-type RNA-binding domains that recognize the sequence-specific motif YCAY (Ule et al., 2006). Knockout mouse models have shown that these factors are required for the survival of spinal cord and brainstem neurons (Jensen et al., 2000) and the migration of late-generated cortical and Purkinje neurons (Yano et al., 2010). Splicing-sensitive microarrays, CLIP-seq studies, and bioinformatics analyses identified direct Nova target exons involved in synaptic function and axon guidance (Ule et al., 2006; 2005; Zhang et al., 2010). Follow-up studies characterized the function of Nova-regulated exons in the Dab1 and Agrin transcripts, which are important for the neuronal migration of cortical neurons and synaptic function at neuromuscular junctions, respectively (Ruggiu et al., 2009; Yano et al., 2010).

The Rbfox family of proteins includes Rbfox1, Rbfox2, and Rbfox3. Rbfox2 is broadly expressed, whereas Rbfox1 is expressed in neurons, heart, and muscle; and Rbfox3 expression is limited to neurons. These proteins are very similar in sequence, particularly in their RNA binding domain. To regulate alternative splicing, the proteins recognize the motif (U)GCAUG in the introns flanking the regulated exon (Underwood et al., 2005). Similar to the Nova proteins, an Rbfox motif upstream of a regulated exon promotes exon skipping, while a downstream motif promotes exon inclusion (Yeo et al., 2009). Rbfox mutant mice are susceptible to seizures (Gehman et al., 2011) and display disrupted cerebellar development (Gehman et al., 2012). The observed phenotypes are in part due to Rbfox-dependent exons in proteins relevant to synaptic transmission and membrane excitability (Gehman et al., 2011; 2012). CLIP-seq experiments of

all three Rbfox proteins identified direct targets important for brain development (Weyn-Vanhentenryck et al., 2014).

SRRM4 is another splicing factor that is exclusively expressed in neuronal cells (Calarco et al., 2009). SRRM4 binds intronic UGC-containing motifs near the 3' splice site of regulated exons to activate neuronal exon inclusion (Raj et al., 2014). A recent study showed that SRRM4 knockout mice have impaired neurite outgrowth, cortical layering, and axon guidance (Quesnel-Vallières et al., 2015). RNA-sequencing identified that SRRM4 regulates the inclusion of a large network of brain-specific alternative exons (Calarco et al., 2009). One SRRM4 target is the repressor element 1 (RE-1) silencing transcription factor (REST/NRSF) (Raj et al., 2011). REST is expressed in NPCs and non-neuronal cells, where it represses neuronal gene expression programs. With neuronal differentiation, SRRM4 regulates REST splicing and promotes the expression of the inactive REST4 isoform, which removes the repression of neuronal target genes, including SRRM4.

PTB proteins during neuronal development

The polypyrimidine tract binding (PTB) proteins, PTBP1 and PTBP2, regulate a large set of splicing events during neuronal development. In neuronal progenitor cells (NPCs) and non-neuronal cells, PTBP1 expression is high while PTBP2 expression is low (Boutz et al., 2007; Li et al., 2014; Licatalosi et al., 2012). In these cells, PTBP1 represses PTBP2 expression through AS-NMD and additional post-transcriptional mechanisms (Boutz et al., 2007; Makeyev et al., 2007; Spellman et al., 2007). With neuronal differentiation, PTBP1 expression is repressed by miRNA124 allowing for the induction of PTBP2 expression (Makeyev et al., 2007). During the early transition in PTB protein expression, many splicing events occur that are essential for

neuronal maturation and survival. This early program affects a diverse set of targets, including those involved in cytoskeletal assembly, membrane excitation, and synaptic maturation (Boutz et al., 2007; Tang et al., 2011; Zheng et al., 2012). As neurons mature, PTBP2 expression is eventually reduced, which allows for the expression of adult-specific splicing patterns (Li et al., 2014; Zheng et al., 2012). Together, the transitions in PTBP1 and PTBP2 expression define three different splicing programs that guide neuronal development and maturation.

PTBP1 and PTBP2 are encoded on separate genes but share 74% peptide identity. Their peptide sequences are particularly conserved in their four RRMs, giving them similar RNA-binding properties (Markovtsov et al., 2000). Despite their similarities, PTBP1 and PTBP2 regulate overlapping but different sets of targets during neuronal differentiation. Some exons are responsive to PTBP1 but not PTBP2, and these include c-src N1 exon, $GABA_A$ receptor-γ2 exon 9, and CaV1.2 exon 8a (Ashiya and Grabowski, 1997; Chan and Black, 1997; Tang et al., 2011). Other exons are responsive to both proteins, including Dlg4 exon 18 (Zheng et al., 2012). In these cases, PTBP1 and PTBP2 act together to repress adult splicing patterns during neuronal maturation (Li et al., 2014; Zheng et al., 2012). The mechanisms that determine the differences in PTBP1 and PTBP2 activity are poorly understood.

Like other splicing factors, the PTB proteins regulate other aspects of posttranscriptional regulation in neurons. A recent study demonstrated that in non-neuronal cells, PTBP1 blocks miRNA targeting of transcripts regulating neuronal differentiation, including component of the REST complex (Xue et al., 2013). The loss of PTBP1 in mouse embryonic fibroblast (MEF) cells results in the expression of neuronal-specific transcription factors and the transdifferentiation of MEF cells into neurons. Another study showed that PTBP1 regulates the expression of several presynaptic proteins (Stx1b, Vamp2, Sv2a) through 3'-terminal intron

retention (Yap et al., 2012). In non-neuronal cells, PTBP1 promotes intron retention in these transcripts, which leads to their nuclear retention and decay. PTBP2 likely regulates similar posttranscriptional programs, but they are not well characterized.

The proper expression of PTBP1 and PTBP2 are essential for neuronal maturation. The loss of PTBP2 in the developing mouse cortex leads to its degeneration during early postnatal development (Li et al., 2014). Likewise, cultured neurons lacking PTBP2 fail to mature and die in culture. The maturation defects are consistent with the splicing changes observed in the PTBP2 knock out mice. PTBP2 targets affect genes involved in cytoskeletal assembly, neurite growth, and synaptic function, many of which are PTBP1 targets. Meanwhile, the re-expression of PTBP1 and PTBP2 in mature neuronal cultures represses synaptic activity and dendritic spine formation (Zheng et al., 2012). These changes are in part mediated by PTB protein repression of Dlg4 exon 18, a scaffolding protein important for the function of glutamatergic synapses. The skipping of Dlg4 exon 18 introduces a new termination codon (PTC) in exon 19, resulting in the NMD of the transcript. Thus one function of the PTB proteins is to prevent the expression of adult isoforms during critical periods of neuronal development.

The early splicing program regulated by PTBP1 and not by PTBP2 has been difficult to characterize in the developing mouse brain. Most of the work identifying PTBP1 targets has been carried out in non-neuronal cell lines, and may not accurately represent neuronal development (Llorian et al., 2010; Xue et al., 2009; 2013). Likewise, the cellular functions of many PTBP1-regulated isoforms have not been well characterized. Here we use embryonic stem cell (ESC) culture to define the PTBP1-specific splicing program during *in vitro* neuronal development. We are particularly interested in characterizing how the PTBP1-specific splicing program contributes to neuronal lineage commitment and differentiation.

References

Amara, S.G., Jonas, V., Rosenfeld, M.G., Ong, E.S., and Evans, R.M. (1982). Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. Nature *298*, 240–244.

Ashiya, M., and Grabowski, P.J. (1997). A neuron-specific splicing switch mediated by an array of pre-mRNA repressor sites: evidence of a regulatory role for the polypyrimidine tract binding protein and a brain-specific PTB counterpart. Rna *3*, 996–1015.

Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov, S., Lee, L.J., Slobodeniuc, V., Kutter, C., Watt, S., Colak, R., et al. (2012). The evolutionary landscape of alternative splicing in vertebrate species. Science *338*, 1587–1593.

Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. Annu. Rev. Biochem. *72*, 291–336.

Boutz, P.L., Stoilov, P., Li, Q., Lin, C.-H., Chawla, G., Ostrow, K., Shiue, L., Ares, M., and Black, D.L. (2007). A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. Genes Dev. *21*, 1636–1652.

Braunschweig, U., Gueroussov, S., Plocik, A.M., Graveley, B.R., and Blencowe, B.J. (2013). Dynamic integration of splicing within gene regulatory pathways. Cell *152*, 1252–1269.

Calarco, J.A., Superina, S., O'Hanlon, D., Gabut, M., Raj, B., Pan, Q., Skalska, U., Clarke, L., Gelinas, D., van der Kooy, D., et al. (2009). Regulation of vertebrate nervous system alternative splicing and development by an SR-related protein. Cell *138*, 898–910.

Chan, R.C., and Black, D.L. (1997). The polypyrimidine tract binding protein binds upstream of neural cell-specific c-src exon N1 to repress the splicing of the intron downstream. Mol. Cell. Biol. *17*, 4667–4676.

Chen, M., and Manley, J.L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat. Rev. Mol. Cell Biol. *10*, 741–754.

Fu, X.-D., and Ares, M. (2014). Context-dependent control of alternative splicing by RNA-binding proteins. Nat. Rev. Genet. *15*, 689–701.

Garcia-Blanco, M.A., Jamison, S.F., and Sharp, P.A. (1989). Identification and purification of a 62,000-dalton protein that binds specifically to the polypyrimidine tract of introns. Genes Dev. *3*, 1874–1886.

Gehman, L.T., Meera, P., Stoilov, P., Shiue, L., O'Brien, J.E., Meisler, M.H., Ares, M., Otis, T.S., and Black, D.L. (2012). The splicing regulator Rbfox2 is required for both cerebellar development and mature motor function. Genes Dev. *26*, 445–460.

Gehman, L.T., Stoilov, P., Maguire, J., Damianov, A., Lin, C.-H., Shiue, L., Ares, M., Mody, I., and Black, D.L. (2011). The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. Nat. Genet. *43*, 706–711.

Gooding, C., Roberts, G.C., and Smith, C.W. (1998). Role of an inhibitory pyrimidine element and polypyrimidine tract binding protein in repression of a regulated alpha-tropomyosin exon. Rna *4*, 85–100.

Jensen, K.B., Dredge, B.K., Stefani, G., Zhong, R., Buckanovich, R.J., Okano, H.J., Yang, Y.Y., and Darnell, R.B. (2000). Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability. Neuron *25*, 359–371.

Kalsotra, A., and Cooper, T.A. (2011). Functional consequences of developmentally regulated alternative splicing. Nat. Rev. Genet. *12*, 715–729.

Keppetipola, N., Sharma, S., Li, Q., and Black, D.L. (2012). Neuronal regulation of pre-mRNA splicing by polypyrimidine tract binding proteins, PTBP1 and PTBP2. Crit. Rev. Biochem. Mol. Biol. *47*, 360–378.

Li, Q., Zheng, S., Han, A., Lin, C.-H., Stoilov, P., Fu, X.-D., and Black, D.L. (2014). The splicing regulator PTBP2 controls a program of embryonic splicing required for neuronal maturation. Elife *3*, e01201–e01201.

Licatalosi, D.D., Yano, M., Fak, J.J., Mele, A., Grabinski, S.E., Zhang, C., and Darnell, R.B. (2012). Ptbp2 represses adult-specific splicing to regulate the generation of neuronal precursors in the embryonic brain. Genes Dev. *26*, 1626–1642.

Llorian, M., Schwartz, S., Clark, T.A., Hollander, D., Tan, L.-Y., Spellman, R., Gordon, A., Schweitzer, A.C., la Grange, de, P., Ast, G., et al. (2010). Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. Nat. Struct. Mol. Biol. *17*, 1114–1123.

Lou, H., Helfman, D.M., Gagel, R.F., and Berget, S.M. (1999). Polypyrimidine tract-binding protein positively regulates inclusion of an alternative 3'-terminal exon. Mol. Cell. Biol. *19*, 78–85.

Makeyev, E.V., Zhang, J., Carrasco, M.A., and Maniatis, T. (2007). The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. Mol. Cell *27*, 435–448.

Markovtsov, V., Nikolic, J.M., Goldman, J.A., Turck, C.W., Chou, M.Y., and Black, D.L. (2000). Cooperative assembly of an hnRNP complex induced by a tissue-specific homolog of polypyrimidine tract binding protein. Mol. Cell. Biol. *20*, 7463–7479.

Matlin, A.J., Clark, F., and Smith, C.W.J. (2005). Understanding alternative splicing: towards a cellular code. Nat. Rev. Mol. Cell Biol. *6*, 386–398.

Merkin, J., Russell, C., Chen, P., and Burge, C.B. (2012). Evolutionary dynamics of gene and

isoform regulation in Mammalian tissues. Science *338*, 1593–1599.

Mulligan, G.J., Guo, W., Wormsley, S., and Helfman, D.M. (1992). Polypyrimidine tract binding protein interacts with sequences involved in alternative splicing of beta-tropomyosin pre-mRNA. J. Biol. Chem. *267*, 25480–25487.

Norris, A.D., and Calarco, J.A. (2012). Emerging Roles of Alternative Pre-mRNA Splicing Regulation in Neuronal Development and Function. Front Neurosci *6*, 122.

Oberstrass, F.C., Auweter, S.D., Erat, M., Hargous, Y., Henning, A., Wenter, P., Reymond, L., Amir-Ahmady, B., Pitsch, S., Black, D.L., et al. (2005). Structure of PTB bound to RNA: specific binding and implications for splicing regulation. Science *309*, 2054–2057.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat. Genet. *40*, 1413–1415.

Polydorides, A.D., Okano, H.J., Yang, Y.Y., Stefani, G., and Darnell, R.B. (2000). A brain-enriched polypyrimidine tract-binding protein antagonizes the ability of Nova to regulate neuron-specific alternative splicing. Proc. Natl. Acad. Sci. U.S.a. *97*, 6350–6355.

Posner, J.B., and Darnell, R.B. (1993). Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. Neuron *11*, 657–672.

Quesnel-Vallières, M., Irimia, M., Cordes, S.P., and Blencowe, B.J. (2015). Essential roles for the splicing regulator nSR100/SRRM4 during nervous system development. Genes Dev. *29*, 746–759.

Raj, B., Irimia, M., Braunschweig, U., Sterne-Weiler, T., O'Hanlon, D., Lin, Z.-Y., Chen, G.I., Easton, L.E., Ule, J., Gingras, A.-C., et al. (2014). A global regulatory mechanism for activating an exon network required for neurogenesis. Mol. Cell *56*, 90–103.

Raj, B., O'Hanlon, D., Vessey, J.P., Pan, Q., Ray, D., Buckley, N.J., Miller, F.D., and Blencowe, B.J. (2011). Cross-regulation between an alternative splicing activator and a transcription repressor controls neurogenesis. Mol. Cell *43*, 843–850.

Ruggiu, M., Herbst, R., Kim, N., Jevsek, M., Fak, J.J., Mann, M.A., Fischbach, G., Burden, S.J., and Darnell, R.B. (2009). Rescuing Z+ agrin splicing in Nova null mice restores synapse formation and unmasks a physiologic defect in motor neuron firing. Proc. Natl. Acad. Sci. U.S.a. *106*, 3513–3518.

Schoenberg, D.R., and Maquat, L.E. (2012). Regulation of cytoplasmic mRNA decay. Nat. Rev. Genet. *13*, 246–259.

Sharma, S., Falick, A.M., and Black, D.L. (2005). Polypyrimidine tract binding protein blocks the 5' splice site-dependent assembly of U2AF and the prespliceosomal E complex. Mol. Cell *19*, 485–496.

Sharma, S., Kohlstaedt, L.A., Damianov, A., Rio, D.C., and Black, D.L. (2008). Polypyrimidine tract binding protein controls the transition from exon definition to an intron defined spliceosome. Nat. Struct. Mol. Biol. *15*, 183–191.

Southby, J., Gooding, C., and Smith, C.W. (1999). Polypyrimidine tract binding protein functions as a repressor to regulate alternative splicing of alpha-actinin mutally exclusive exons. Mol. Cell. Biol. *19*, 2699–2711.

Spellman, R., and Smith, C.W.J. (2006). Novel modes of splicing repression by PTB. Trends Biochem. Sci. *31*, 73–76.

Spellman, R., Llorian, M., and Smith, C.W.J. (2007). Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1. Mol. Cell *27*, 420–434.

Tang, Z.Z., Sharma, S., Zheng, S., Chawla, G., Nikolic, J., and Black, D.L. (2011). Regulation of the mutually exclusive exons 8a and 8 in the CaV1.2 calcium channel transcript by polypyrimidine tract-binding protein. J. Biol. Chem. *286*, 10007–10016.

Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., Gaasterland, T., Blencowe, B.J., and Darnell, R.B. (2006). An RNA map predicting Nova-dependent splicing regulation. Nature *444*, 580–586.

Ule, J., Ule, A., Spencer, J., Williams, A., Hu, J.-S., Cline, M., Wang, H., Clark, T., Fraser, C., Ruggiu, M., et al. (2005). Nova regulates brain-specific splicing to shape the synapse. Nat. Genet. *37*, 844–852.

Ullrich, B., Ushkaryov, Y.A., and Südhof, T.C. (1995). Cartography of neurexins: more than 1000 isoforms generated by alternative splicing and expressed in distinct subsets of neurons. Neuron *14*, 497–507.

Underwood, J.G., Boutz, P.L., Dougherty, J.D., Stoilov, P., and Black, D.L. (2005). Homologues of the Caenorhabditis elegans Fox-1 protein are neuronal splicing regulators in mammals. Mol. Cell. Biol. *25*, 10005–10016.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470–476.

Weyn-Vanhentenryck, S.M., Mele, A., Yan, Q., Sun, S., Farny, N., Zhang, Z., Xue, C., Herre, M., Silver, P.A., Zhang, M.Q., et al. (2014). HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. Cell Rep *6*, 1139–1152.

Xue, Y., Ouyang, K., Huang, J., Zhou, Y., Ouyang, H., Li, H., Wang, G., Wu, Q., Wei, C., Bi, Y., et al. (2013). Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. Cell *152*, 82–96.

Xue, Y., Zhou, Y., Wu, T., Zhu, T., Ji, X., Kwon, Y.-S., Zhang, C., Yeo, G., Black, D.L., Sun,

H., et al. (2009). Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. Mol. Cell *36*, 996–1006.

Yamamoto, H., Tsukahara, K., Kanaoka, Y., Jinno, S., and Okayama, H. (1999). Isolation of a mammalian homologue of a fission yeast differentiation regulator. Mol. Cell. Biol. *19*, 3829–3841.

Yano, M., Hayakawa-Yano, Y., Mele, A., and Darnell, R.B. (2010). Nova2 regulates neuronal migration through an RNA switch in disabled-1 signaling. Neuron *66*, 848–858.

Yap, K., and Makeyev, E.V. (2013). Regulation of gene expression in mammalian nervous system through alternative pre-mRNA splicing coupled with RNA quality control mechanisms. Mol. Cell. Neurosci. *56*, 420–428.

Yap, K., Lim, Z.Q., Khandelia, P., Friedman, B., and Makeyev, E.V. (2012). Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. Genes Dev. *26*, 1209–1223.

Yeo, G.W., Coufal, N.G., Liang, T.Y., Peng, G.E., Fu, X.-D., and Gage, F.H. (2009). An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. Nat. Struct. Mol. Biol. *16*, 130–137.

Zhang, C., Frias, M.A., Mele, A., Ruggiu, M., Eom, T., Marney, C.B., Wang, H., Licatalosi, D.D., Fak, J.J., and Darnell, R.B. (2010). Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. Science *329*, 439–443.

Zheng, S., and Black, D.L. (2013). Alternative pre-mRNA splicing in neurons: growing up and extending its reach. Trends Genet. *29*, 442–448.

Zheng, S., Gray, E.E., Chawla, G., Porse, B.T., O'Dell, T.J., and Black, D.L. (2012). PSD-95 is post-transcriptionally repressed during early neural development by PTBP1 and PTBP2. Nat. Neurosci. *15*, 381–8–S1.

CHAPTER 2

The splicing regulator PTBP1 controls the activity of the homeodomain transcription factor Pbx1 during neuronal differentiation

ABSTRACT

PTBP1 and PTBP2 control alternative splicing programs during neuronal development, but the cellular functions of most PTBP1/2-regulated isoforms remain unknown. We show that PTBP1 guides developmental gene expression by regulating the transcription factor Pbx1. We identify differentially spliced exons as mouse embryonic stem cells (ESCs) differentiate into neuronal progenitor cells (NPCs) and neurons, when cells transition from PTBP1 to PTBP2 expression. We define those exons controlled by PTBP1 in ESCs and NPCs by RNA-seq analysis after PTBP1 depletion and PTBP1 crosslinking-immunoprecipitation. We find that PTBP1 represses Pbx1 exon 7 and the expression of its neuronal isoform Pbx1a in ESC. Using CRISPR-Cas9 to delete regulatory elements for exon 7, we induce Pbx1a expression in ESCs, finding that this activates transcription of specific neuronal genes including known Pbx1 targets. Thus PTBP1 controls the activity of Pbx1 and suppresses its neuronal transcriptional program prior to differentiation.

INTRODUCTION

Alternative splicing is an important form of gene regulation during tissue development. In the mammalian nervous system, large scale changes in splice site choice produce many new mRNAs encoding protein isoforms with different structures and functions that are specific to neurons (Li et al., 2007; Licatalosi and Darnell, 2006; Norris and Calarco, 2012; Yap and Makeyev, 2013; Zheng and Black, 2013). These splicing patterns are regulated in a temporal and cell-specific manner by the expression of specialized pre-mRNA binding proteins (RBPs) (Black, 2003; Braunschweig et al., 2013; Fu and Ares, 2014; Lee and Rio, 2015). Several RBPs have been shown to control the expression of isoforms essential for neuronal development and maturation (Charizanis et al., 2012; Gehman et al., 2012; 2011; Iijima et al., 2011; Ince-Dunn et al., 2012; Jensen et al., 2000; Li et al., 2014; Licatalosi et al., 2012; Quesnel-Vallières et al., 2015; Yano et al., 2010). However, the complex programs of splicing affecting neuronal development are only beginning to be characterized, and the cellular functions of the regulated isoforms are often poorly understood.

The polypyrimidine tract binding (PTB) proteins, PTBP1 and PTBP2, regulate a large set of splicing events during neuronal differentiation (Keppetipola et al., 2012). PTBP1 is expressed in neuronal progenitor cells and many non-neuronal cells. Upon neuronal differentiation, PTBP1 expression is repressed allowing the induction of PTBP2 and the initiation of a neuronal splicing program that is essential for neuronal maturation and survival (Boutz et al., 2007; Li et al., 2014; Licatalosi et al., 2012; Makeyev et al., 2007; Tang et al., 2011; Zheng et al., 2012). In addition to splicing, the PTB proteins affect other aspects of posttranscriptional regulation in neurons, including miRNA targeting (Xue et al., 2013; Yap et al., 2012). It has been shown that depletion of PTBP1 from fibroblasts is sufficient to drive cells towards a neuronal phenotype (Xue et al.,

2013). As neurons mature, PTBP2 expression is eventually reduced, giving rise to an adult neuronal splicing program (Li et al., 2014; Licatalosi et al., 2012; Tang et al., 2011; Zheng et al., 2012).

Together, these transitions in PTBP1 and PTBP2 expression define three phases of neuronal maturation each with different splicing programs.

PTBP1 and PTBP2 are encoded on separate genes and have very similar sequences and RNA-binding properties (Markovtsov et al., 2000; Oberstrass et al., 2005). These proteins target overlapping but not identical sets of splicing events (Boutz et al., 2007; Li et al., 2014; Llorian et al., 2010; Tang et al., 2011; Zheng et al., 2012). Some exons are responsive to both proteins (Boutz et al., 2007; Li et al., 2014; Spellman et al., 2007; Zheng et al., 2012). In these cases, PTBP1 and PTBP2 serve to repress adult splicing patterns during neuronal maturation, yielding isoforms that are expressed only after PTBP2 levels have declined. In other cases, exons are more sensitive to PTBP1 and shift their splicing when PTBP1 is shut off early in neuronal differentiation even though PTBP2 is present (Boutz et al., 2007; Makeyev et al., 2007; Markovtsov et al., 2000; Tang et al., 2011). PTBP1 is known to regulate splicing in a wide variety of cell-types, and much of the work on PTBP1 targeting has used non-neuronal cell lines such as HeLa cells (Llorian et al., 2010; Spellman et al., 2007; Xue et al., 2013; 2009). How the PTBP1-specific splicing program contributes to early aspects of neuronal differentiation remains poorly understood.

Directed differentiation of embryonic stem cells (ESCs) provides a versatile model for the study of neuronal commitment, differentiation, and maturation. These systems have illuminated a wide range of genetic contributors to neuronal phenotype including epigenetic modifiers, transcription factors, miRNAs, and signaling molecules (Hirabayashi and Gotoh,

2010; Kosik, 2006; Louvi and Artavanis-Tsakonas, 2006; Ronan et al., 2013; Temple, 2001). Here we use mouse ESC culture to characterize the PTBP1 splicing program during early neuronal differentiation. We identify a diverse set of splicing events regulated by PTBP1, including an alternative exon in the homeodomain transcription factor Pbx1, whose switch in splicing induces a neuronal transcriptional program early in neuronal development.

RESULTS

The transition in PTB protein expression occurs during *in vitro* neuronal differentiation.

To examine PTBP1 and PTBP2 expression in an *in vitro* model of neuronal development, we differentiated mouse embryonic stem cells (ESCs) into motor neurons (MNs). Mouse ESCs (Day -2) were grown in aggregate culture for two days to form embryoid bodies (EBs; Day 0), which were then treated with retinoic acid (RA) and a Sonic hedgehog (Shh) pathway agonist for 5 days to induce MN formation (Adams et al., 2015; Wichterle et al., 2002). To facilitate MN identification and isolation, we used a mouse ESC line that expresses eGFP under the control of the MN specific HB9 promoter (Wichterle et al., 2002). At Day 5, the GFP+ embryoid bodies were dissociated and MNs plated onto Matrigel in the presence of neurotrophic factors to permit further MN maturation (Figure 2-1A).

Similar to observations in the brain, as ESCs differentiate into MNs, PTBP1 expression declines, while PTBP2 expression is induced (Figure 2-1B and C) (Boutz et al., 2007; Tang et al., 2011; Zheng et al., 2012). By immunofluorescence, the GFP+ cells express the postmitotic MN marker HB9 and the neuronal marker TuJ1 (Figure 2-2A, and data not shown). The presence of PTBP2 in GFP+ cells was also seen by immunofluorescence (Figure 2-2B). To monitor PTBP1 and PTBP2 protein specifically in the differentiated neurons, we used fluorescence-

activated cell sorting (FACS) to isolate GFP+ cells at Days 5 and 8. By western blot, these GFP positive cells express low amounts of PTBP1 and high levels of PTBP2 at Day 5, but by Day 8 only trace amounts of PTBP1 are found while PTBP2 expression remains high (Figure 2-2C). We did not observe the late decline in PTBP2 expression that occurs during neuronal maturation *in vivo*. It may be that this decline in PTBP2 expression takes place with longer culture times, or that the cells do not fully mature in this system. These results show that the switch from PTBP1 to PTBP2 expression occurs during MN differentiation, similar to earlier observations in the developing cortex.

We previously observed high levels of PTBP1 and low levels of PTBP2 in neuronal progenitor cells (NPCs) (Boutz et al., 2007; Li et al., 2014; Licatalosi et al., 2012). Nestin+ NPCs are observed early in EB cultures, but these cells could not be isolated in pure populations. To examine NPCs in more detail, ESCs were grown in N2B27 media and expanded in the presence of EGF and FGF, to yield cultures where greater than 80% of the cells were Nestin+ and with very few TuJ1+ neurons (Conti et al., 2005; Ying et al., 2003) (Figure 2-3A) (Figure 2-3B and C). By western blot, we found that both PTBP1 and PTBP2 proteins were expressed in these NPC cultures. As these NPCs were differentiated into neurons with retinoic acid, PTBP1 protein expression was lost while PTBP2 protein was induced (Figure 2-3D). The EB and monolayer culture systems likely give rise to different neuronal populations with differences in gene expression attributable to different neuronal types (Abranches et al., 2009). Nevertheless, the transition in PTBP1-PTBP2 protein expression is common to both models of *in vitro* neuronal development, as well as embryonic brain.

We next examined alternative splicing patterns under different states of PTBP1 and PTBP2 expression. Poly-A plus RNA was isolated from undifferentiated ESCs (high PTBP1 and low

PTBP2), Nestin+ progenitor cells (intermediate PTBP1 and PTBP2), and post-mitotic HB9-GFP+ MNs (low Ptbp1 and high PTBP2). Strand specific libraries for each cell type were subjected to 100 nt paired-end sequencing to generate 200 to 250 million mapped reads per sample. Alternative splicing events were quantified using SpliceTrap to identify cassette exons, 5' and 3' splice sites, and retained introns that differ between all sample pairs (Wu et al., 2011). Differentially spliced events were filtered for p-value below 0.05 and ranked by the change in percent spliced in value (ΔPSI). Using a ΔPSI cutoff of 15%, we identified a large set of alternative splicing events that change across neuronal differentiation (Figure 2-4A). With these filters 1201 differential splicing events were identified between ESCs and MNs, 1218 differential events between NPCs and MNs, and 750 differential events between ESC and NPC. To identify enriched functional categories, transcripts exhibiting differential splicing were compared to the total expressed transcript set by Gene Ontology analyses (Huang et al., 2009). Differential splicing events between NPCs and MNs were enriched (FDR<0.05) in biological processes and cellular compartments relevant to neuronal differentiation such as cell projection organization, axonogenesis, and cell morphogenesis (Figure 2-4B and C). It is expected that a subset of these splicing changes are the result of changes in PTBP1 and PTBP2 expression.

PTBP1 regulates a large set of neuronal exons in ESCs and NPCs.

To identify neuronal splicing events controlled by the PTB proteins and in particular PTBP1, RNAi knockdowns were performed with two independent sets of control, PTBP1, and PTBP2 siRNAs. PTBP1 siRNA treatment depleted the protein by 80% (n=3) and 85% (n=2) in ESCs and NPCs respectively (Figure 2-5A and B). As seen previously, PTBP1 depletion induced PTBP2 protein expression (Boutz et al., 2007; Makeyev et al., 2007; Spellman et al., 2007). To

assess the effect of PTBP2, we performed PTBP1/2 double knockdowns to remove 84% and 93% PTBP2 in ESCs and NPCs respectively (Figure 2-5A and B). 50 nt paired-end RNA sequencing was performed on the control and knockdown samples, and the splicing was analyzed with SpliceTrap.

Using the p-value and PSI cutoffs previously described, we identified 418 splicing events altered by PTBP1 knockdown in ESCs, including 362 cassette exons, 50 alternative 5' or 3' splice sites, and 6 retained introns (Figure 2-6B). This number was only slightly increased by double PTBP1/2 depletion. The splicing changes between the single and double knockdowns showed a strong correlation ($R^2 = 0.91$, Figure 2-6C), indicating that these events are more strongly affected by PTBP1. A majority of the exons altered by PTB protein depletion also change in the same direction during ESC to MN differentiation (exons in the upper right and lower left quadrants of the plots in Figure 2-5C). Applying a ΔPSI cutoff of 15% for both data sets, 97 neuronally activated exons are PTBP1 repressed, and 56 neuronally skipped exons are PTBP1 activated (indicated by red and blue data points respectively in Figure 2-5C, left panel). Similar numbers of PTBP1/2 repressed (101) and activated (55) neuronal exons were observed following PTBP1/2 double knockdown (Figure 2-5C, right panel). As expected, many exons that change during neuronal differentiation are not affected by PTB protein depletion, indicating likely regulation by other proteins (grey data points along the y-axes of the plots in Figure 2-5C).

PTBP1 depletion from NPCs altered 325 splicing events (264 cassette exons, 44 alternative 5' or 3' splice sites, and 17 retained introns, Figure 2-6B). Comparing this list to the splicing changes between NPCs and Day 5 GFP+ MNs identified fewer correlated exons (66 Ptbp1-repressed exons increased with differentiation; 23 Ptbp1-activated exons decreased with differentiation; Figure 2-5D, left panel). Interestingly, PTBP1/2 double knockdown had a bigger

effect in NPCs than in ESCs (116 PTBP1/2-repressed exons increased with differentiation; 45 PTBP1/2-activated exons decreased with differentiation; Figure 2-5D, middle panel). These changes in splicing after double PTBP1/2 knockdown in NPCs also tended to be larger than PTBP1 knockdown alone, yielding a lower correlation than seen in ESC ($R^2 = 0.69$, Figure 2-6D). Knockdown of PTBP2 alone in NPCs had only a small effect on the splicing of neuronal exons, indicating that PTBP2 largely regulates a subset of PTBP1 targets, with relatively few targets of its own (Figure 2-6A) (Figure 2-5D, right panel). This is in agreement with previous results that both PTBP1 and PTBP2 repress splicing of a variety of exons in the developing mouse brain, such as Dlg4 (Psd-95) exon 18 (Boutz et al., 2007; Li et al., 2014; Zheng et al., 2012). In these new data, PTBP1 and PTBP2 single knockdown each only moderately increased Dlg4 exon 18 splicing (ΔPSI = 8% and 13% respectively), whereas PTBP1/2 double knockdown increased exon 18 splicing by 42%. Other exons where PTBP2 can compensate for the loss of PTBP1 are found in the Magi1, Ap2a1, and Tnrc6a transcripts (Table 2-1). Thus, PTBP1 and PTBP1/2 depletion from NPCs identifies exons that are coregulated by both proteins. Whereas, depletion from ESCs identifies exons affected largely by PTBP1.

PTBP1 iCLIP-seq identifies direct PTBP1 splicing targets during neuronal differentiation.

To identify direct binding of PTBP1 to transcripts in these cells, we performed crosslinking-immunoprecipitation followed by sequencing (iCLIP-seq) (König et al., 2011; 2010). UV irradiated ESC and NPC cultures were immunoprecipitated with PTBP1 antibody or with Flag antibody as a negative control (Figure 2-7A). Cross-linked RNA fragments were converted into cDNA using a modification of the iCLIP protocol and subjected to high-density sequencing (König et al., 2010). After removing PCR duplicates, unique sequencing reads were mapped to

the UCSC known Genes table (Hsu et al., 2006). To identify significant crosslinking sites, the FDR for each position was calculated (König et al., 2010). Applying an FDR cutoff of 0.01, we obtained 918,725 and 914,084 significant reads of RNA crosslinked to PTBP1 in ESCs and NPCs respectively.

The PTBP1-bound sequences were analyzed for base content and location. The iCLIP protocol generally yields sequence reads that terminate one nucleotide downstream of the crosslink site (König et al., 2010). We compiled a set of binding regions extending 20 nucleotides upstream and downstream of a crosslink site. We measured the frequency of all pentamer motifs within these binding sites, relative to randomly chosen intervals from the same introns. As expected, PTBP1 binding sites in both ESC and NPC were highly enriched for CU-rich pentamers. The top 20 motifs included pentamers comprised of only C and U nucleotides, as well as some with G nucleotides, in agreement with the described binding specificity of PTBP1 (Figure 2-7B) (Amir-Ahmady et al., 2005; Han et al., 2014; Licatalosi et al., 2012; Llorian et al., 2010; Oberstrass et al., 2005). The enriched motifs were very similar between ESC and NPC samples (R^2= 0.93, Figure 2-7B). The locations of the binding sites were also consistent with previous observations, with most being found in introns and some in 3' UTRs (Figure 2-7C and D) (Xue et al., 2009).

We next examined the alternative exons sensitive to PTBP1 knockdown (ΔPSI $\geq$ 15%) for PTBP1 iCLIP clusters within 500 nt upstream or downstream ($\geq$ four significant reads per cluster). In ESCs, we identified 170 PTBP1-dependent exons that are bound by PTBP1 (Figure 2-7E). In NPCs, we identified 108 PTBP1-dependent exons with PTBP1 binding (Figure 2-7F). These include many known PTBP1 targets transcripts such as Ganab, Rod1, Tpm1, Rbm27, Fam38a, Mink1, and Ptbp2 (Boutz et al., 2007; Makeyev et al., 2007; Spellman et al., 2007; Xue

et al., 2009). Some known target exons with PTBP1 binding did not exhibit splicing changes above our cutoff, including Pkm2, Eif4g2, Dlg4, and Ptbp1 (Xue et al., 2009; Zheng et al., 2013). These exons may be controlled by additional factors in ESCs and NPCs.

Overlapping all of the datasets, we defined a minimal set of 104 direct PTBP1 targets in ESCs and NPCs, whose splicing changes during neuronal differentiation (70 Ptbp1 repressed, 34 Ptbp1 activated; Tables 2-2 and 2-3). Many other exons are expected to be direct targets but were excluded due to failure to fulfill a cutoff value in one dataset. The exons in this minimal set affect a wide range of cellular processes including GTPase signaling (Dock7), synaptic function (Gabbr1), and transcriptional regulation (Tcf20) (Figure 2-8A-D). For example, Gabbr1 is a subunit of a metabotropic gamma-aminobutyric acid (GABA) receptor. PTBP1 induced skipping of Gabbr1 exon 15 leads to premature translation termination in exon 16 and predicted nonsense mediated decay of the Gabbr1 mRNA (Makeyev et al., 2007). We observe PTBP1 binding within exon 15 and upstream, consistent with the increased exon 15 splicing observed during neuronal differentiation and after PTBP1 depletion (Figure 2-8A). This gene is similar to Dlg4, where PTBP1 regulates its overall expression during development rather than the production of a neuronal isoform. However unlike Dlg4, Gabbr1 is more sensitive to PTBP1 than PTBP2, allowing its expression to be induced earlier.

PTBP1 regulates a switch in Pbx1 isoform expression.

The above data define the PTBP1 regulatory program in a new level of detail. For early cells along the neuronal lineage, the PTBP1 program is distinct from that regulated by PTBP2. PTBP1 target transcripts influence an array of important cellular functions, and understanding how splicing alters each target's activity in differentiating cells is a challenge. We were particularly

interested in how the early PTBP1 program affects gene regulation. We identified a number of transcriptional regulators whose splicing is controlled by PTBP1 during ESC differentiation into NPCs, including Tcf20, Med23, Gatad2a, and Pbx1. These early changes in the structure of transcriptional regulators have the potential to broadly alter gene expression programs during neuronal development (Demir and Dickson, 2005; Gabut et al., 2011; Raj et al., 2011).

We focused on the homeodomain transcription factor Pbx1 because of its known interactions with the Hox family of developmental regulators (LaRonde-LeBlanc and Wolberger, 2003; Piper et al., 1999). Multiple studies implicate Pbx1 and its family members as transcriptional regulators during neuronal development (Maeda et al., 2002; Schulte and Frank, 2014; Sgadò et al., 2012; Vitobello et al., 2011). Exon 7 in the Pbx1 transcript is conserved across mammals and its inclusion leads to production of the Pbx1a isoform. Skipping exon 7 shifts the translational reading frame to introduce a new termination codon (PTC) in exon 8 (Figure 2-9A). This translation stop does not result in NMD but instead generates the shorter isoform, Pbx1b, which retains the DNA binding homeodomain but lacks the 83 amino acids at the C-terminus of Pbx1a. During development, early embryonic tissues predominantly express Pbx1b, whereas Pbx1a is found in neural tissues (Redmond et al., 1996; Schnabel et al., 2001). We found a similar isoform switch in cultured cells, where ESCs express low levels of Pbx1b. As ESCs differentiate into NPCs and HB9+ MNs, both overall Pbx1 expression and exon 7 splicing are induced, such that MNs predominantly express the Pbx1a isoform (Figure 2-9B and C). Previous work indicates that the Pbx1 splice isoforms have different transcriptional activities and cellular functions (Asahara et al., 1999; Di Rocco et al., 1997; Kamps et al., 1991). However, the role of these functional differences during neuronal differentiation is not understood.

We find that Pbx1 exon 7 is regulated by PTBP1. Exon 7 is derepressed as ESCs differentiate into NPCs and further into motor neurons (Figure 2-9B and C). Depletion of PTBP1 from ESCs or NPCs strongly induces exon 7 splicing (Figure 2-9D and E). In NPCs, a PTBP1 iCLIP cluster is found in the intron upstream of exon 7 (Figure 2-9B). These iCLIP tags are not seen in ESCs presumably due to lower Pbx1 transcript levels in these cells (Figure 2-9B and C). These data indicate that one consequence of PTBP1 depletion during neuronal differentiation is to switch Pbx1 expression from the Pbx1b isoform to Pbx1a.

Intron 6 deletions in Pbx1 induce exon 7 splicing in ESC.

PTBP1 represses the production of Pbx1a in ESCs and during early neuronal differentiation. To examine the role of Pbx1a repression, we modified the Pbx1 locus using the CRISPR-Cas9 technology to force exon 7 splicing in cells containing PTBP1 (Cong and Zhang, 2015; Doudna and Charpentier, 2014). Intron 6 upstream of exon 7 contains a PTBP1 iCLIP cluster as well as regions of mammalian sequence conservation, possibly indicative of additional splicing regulatory elements. Guide RNA pairs were designed to target the Cas9 endonuclease to sites within intron 6 (Figure 2-10A). After Cas9 catalyzes double-strand breaks at each site, non-homologous end joining (NHEJ) of the free ends will potentially generate genomic deletions within intron 6 (Figure 2-11A). The largest deletion, between guides 1 and 4, would remove 5.5 kb of intron 6 containing all of the conserved regions, while preserving the splice sites (Figure 2-10A). HB9-GFP ESCs were transiently transfected with expression plasmids for Cas9 and two guide RNAs, and clones were isolated from single transfected cells and genotyped by PCR, with the deletion event confirmed by sequencing. Using guides g1 and g4, we identified no homozygous deletions among 60 ESC clones, although heterozygotes were readily isolated

(Figure 2-11B). In total, we isolated five heterozygous deletion clones with preserved splice sites. Shorter deletions were also tested. Guides 3 and 4 generated a 1.4 kb deletion that includes the PTBP1 binding site. Again this deletion was only isolated as a heterozygote (1/6 clones). In contrast, guides 2 and 3 generated a 1.5 kb deletion from the middle of the intron that was isolated as a homozygous allele (3/10 clones).

Interestingly, these deletions had different effects on Pbx1 splicing. Clones carrying the heterozygous g1-g4 deletion (I6 +/-) exhibited dramatically increased exon 7 splicing (Figure 2-10B and C). Exon 7 was included in only 16% of the Pbx1 mRNA in clones of the parental ESC line, whereas exon 7 was included in 51% of the mRNA in the I6 +/- clones. This indicates that RNA from the mutant allele is almost entirely splicing in exon 7 and confirms that intron 6 contains splicing silencer elements for this exon.

We also tested the other deletions within intron 6 for their effects on exon 7 splicing. The deletion derived from guides 3 and 4 was similar to the long deletion using guides 1 and 4, where the heterozygous deletion exhibited strongly induced exon 7 splicing (g3-4 in Figure 2-10B and C). In contrast, the homozygous intronic deletion derived from guides 2 and 3 had a minimal effect, with exon 7 splicing similar to wild type in 3 separate clones (g2-3 in Figure 2-10B and C). The inability to isolate cells carrying homozygous mutations that increase exon 7 splicing suggests that either the gain of Pbx1a or the loss of Pbx1b is disadvantageous to ESC growth.

These experiments open up many possibilities for using Cas9-mediated mutagenesis to map splicing regulatory elements and to alter the splicing phenotype of particular genes. We next assessed the effect of forced Pbx1a expression during neuronal differentiation.

The early expression of Pbx1a promotes neuronal gene expression.

To characterize how early Pbx1a expression alters transcriptional programs, we performed transcriptional profiling with RNA-seq to compare cells with different Pbx1 genotypes. Wild type (n=3) and I6 +/- (n=5) ESC clones were differentiated into MNs. At Day 2, the I6 +/- cells predominantly expressed the Pbx1a mRNA isoform, and increased Pbx1a protein compared to wildtype cells (Figure 2-10D and E). The Day 2 samples were then subjected to 50 nt single-end RNA sequencing, with the data analyzed by the Cufflinks pipeline. We performed hierarchical clustering on the replicates, and as expected the gene expression profiles were similar across replicates, with the profiles of the wild type and Pbx1 mutant cell replicates segregating into two groups (data not shown).

Using a twofold change in FPKM as the cutoff, we identified 33 differentially expressed genes in the I6 +/- cultures (31 induced, 2 repressed). Examining these genes with the largest changes in expression confirmed that a Pbx1a neuronal regulatory program was activated in the mutant cells. Several of the top 20 induced transcripts are known to play important roles in neuronal differentiation including Phox2b, Cntn2, Ntng2, Olig1, Isl1, Nrp2, Ngfr, Nav2, and Slit1, and many of these were previously shown to be affected by Pbx1 (Igf2, Isl1, Dlk1, and Meox1; Figure 2-12B) (Jürgens et al., 2009; Kim et al., 2002; Thiaville et al., 2012). Notably, the gene exhibiting the largest change in expression was Phox2b, a marker of hindbrain visceral motor neurons and a known Pbx1a target (Pattyn et al., 2000; Samad et al., 2004).

To increase the number of differentially expressed genes, we relaxed the cutoff to a 1.5-fold change in expression and identified 196 induced and 25 repressed genes in the mutant relative to the wild type cells. Gene ontology (GO) analyses were performed on the induced gene set to assess functional enrichments relative to the total transcripts present in the D2 MN cultures. Within the biological processes ontology, several terms were significantly enriched

(FDR<0.05) including axonogenesis, pattern specification, regulation of transcription, cell adhesion, cell motion, cell fate commitment, and heart development (Figure 2-12C). Notably, the induced genes include the MN markers Olig2, Lhx3, and Mnx1, again consistent with Pbx1a activating a neuronal transcriptional program.

Overlapping the 196 induced events with a published Pbx1 ChIP data set identified 52 genes with nearby Pbx1 binding (Penkov et al., 2013). We were particularly interested in the up-regulation of Hoxa3 in the mutant cells (Figure 2-12D). The Hoxa3 locus contains a conserved enhancer with two Hox-Pbx-binding sites that are necessary for its *in vivo* activity in the mouse hindbrain (Manzanares et al., 2001). Our data indicate that the switch to Pbx1a during neuronal development contributes to the activation Hoxa3. Thus, the intronic mutation that induces Pbx1a expression in the presence of PTBP1 alters gene activity early in neuronal differentiation and increases the expression of genes affecting neuronal fate.

DISCUSSION

Neuronal development is driven by a complex set of genetic events affecting every step in gene expression from transcriptional regulation by specialized DNA binding proteins to translational repression by miRNAs. These levels of regulation are intricately connected, where one type of regulator will target effectors at different steps in the gene expression pathway. In particular, alternative splicing plays a key role in defining the neuronal proteome with thousands of spliced isoforms found specifically in neurons or particular neuronal subtypes (Barbosa-Morais et al., 2012; Grabowski and Black, 2001; Merkin et al., 2012). A variety of RNA binding proteins alter splicing patterns during development and in the adult brain, including members of

the PTBP, Rbfox, Nova, CELF, SRRM and other protein families (Norris and Calarco, 2012; Zheng and Black, 2013). The changes in isoform expression controlled by these factors determine the activity of many essential neuronal genes. However, these regulatory programs are incompletely characterized and the cellular roles of their alternatively spliced products are often not understood.

The polypyrimidine tract binding proteins, PTBP1 and PTBP2, influence neuronal differentiation through programs of posttranscriptional gene regulation. In previous work, we found that the transition from PTBP1 to PTBP2 expression during neuronal differentiation causes a large scale reprogramming of splicing patterns (Boutz et al., 2007). More recently, we identified a second transition late in neuronal maturation, where PTBP2 levels decline allowing for the expression of many adult-specific spliced isoforms (Li et al., 2014; Zheng et al., 2012). Like many splicing regulators, the PTB proteins also affect other aspects of mRNA metabolism. Notably, the depletion of PTBP1 from mouse embryonic fibroblasts (MEFs) has been shown to induce the trans-differentiation of these cells into neurons (Xue et al., 2013). This was linked to PTBP1 bound in 3' UTRs altering miRNA targeting of transcripts driving neuronal differentiation, including the REST transcriptional repressor. We have not found that PTBP1 depletion is sufficient to induce neuronal differentiation of ESCs. This could reflect a difference between ESCs and MEFs, or differences in our culture conditions. Notably, we find here that, like its 3' UTR targets, the PTBP1 splicing program also includes targets that enhance neuronal development, such as Pbx1.

Earlier studies demonstrated that while PTBP1 and PTBP2 share many targets, exons can have different sensitivities to the two proteins (Keppetipola et al., 2012). This raised questions about the mechanisms of their differential targeting and the biological roles of the early and late

neuronal splicing programs. The late program was illuminated in the PTBP2 knockout mouse where many exons that are repressed by both proteins were identified (Li et al., 2014; Licatalosi et al., 2012). However, the early program composed of exons primarily targeted by PTBP1 was harder to define in the developing brain. By examining the PTBP1 splicing program in ESC culture, we now identify a large set of targets primarily responsive to PTBP1 regulation. Interestingly in ESCs, exons that are sensitive PTBP1 depletion are largely not responsive to the concomitant PTBP2 induction. This contrasts with NPCs where many exons whose splicing shifts with PTBP1 depletion are more strongly affected by PTBP1/2 co-depletion. We previously found that in mice heterozygous for the PTBP2 knockout mutation, certain exons were spliced at 50% the level seen in the homozygous null, indicating a strong effect of PTBP2 concentration. Similarly, we find here that certain exons, such as Gabbr1 exon 15, are very sensitive to changes in PTBP1 expression, such as the reduction seen when ESCs differentiate into NPCs. Like PTBP2, the PTBP1 targets display a range of responsiveness to changes in PTBP1 concentration, and the ESC system gives us a new tool for examining the earlier PTBP1 dependent program.

Alternative exons are usually regulated by ensembles of splicing factors acting to repress or activate their splicing (Fu and Ares, 2014; Lee and Rio, 2015). A question of interest is how targeting by the PTB proteins is affected by other neuronal splicing factors. Two known PTBP1 cofactors that may also affect PTBP2 are Matrin3 and Raver1 (Coelho et al., 2015; Hüttelmaier et al., 2001; Rideau et al., 2006). Both these proteins are likely to affect the activity of PTBP1 and PTBP2 on certain targets. We find that both proteins are well expressed in ES cells, and Matrin3 is strongly upregulated with differentiation into NPCs and neurons. We recently identified several potential cofactors that alter the splicing of the PTBP1/2 target exon in Dlg4 (Zheng et al., 2013). These will also be interesting to examine in relation to additional PTBP

targets, and whether they more strongly affect exons controlled by PTBP1, PTBP2 or both proteins. A protein that can counteract PTBP repression is nSR100/SRRM4, which is induced with neuronal differentiation and whose targets include some PTBP1/2 targets (Calarco et al., 2009; Raj et al., 2014). SRRM4 expression coincides with PTBP2, and its role may be to specifically antagonize the effects of PTBP2 on certain exons in immature neurons. It will be also interesting to identify the SRRM4 target exons during NPC differentiation and to assess their sensitivity to PTBP2 compared to PTBP1.

The intent of this study was to define a set of targets that are primarily responsive to PTBP1, and thus may affect early neuronal lineage commitment and differentiation. We identify a diverse group of transcripts that are sensitive to PTBP1 depletion from ESCs and which change their splicing when ESCs differentiate into NPCs and then into early neurons, as PTBP1 is reduced. Among these targets, we focused on Pbx1, which contains an alternative exon 7 that is highly responsive to PTBP1 concentration and is implicated in neuronal transcriptional regulation.

Using Cas9 genome editing, we created Pbx1 alleles that eliminate PTBP1 repression of exon 7, causing its premature splicing during development (Cong and Zhang, 2015; Doudna and Charpentier, 2014). This early switch in Pbx1 isoforms was sufficient to induce a variety of transcripts, including multiple mRNAs implicated in neuronal lineage commitment. Thus, one function of PTBP1 is to repress the neuronal form of Pbx1 and prevent its action at target genes early in development. It will be interesting to examine how different interactions of the two Pbx1 isoforms might lead to their different activities. Pbx1b lacks 83 amino acids at the C-terminus of Pbx1a but retains domains that facilitate interactions with the Hox, Prep, and Meis transcription factors (Longobardi et al., 2014). An earlier study found that Pbx1a and Pbx1b differ in their

interactions with the transcriptional corepressors NCoR1 and NCoR2 and their ability to repress reporter gene expression (Asahara et al., 1999).Thus, the changes in Pbx1 structure will likely influence its DNA-binding specificity and/or enhancer complex assembly at specific genes.

A splicing factor will often regulate a large network of targets that may overlap with the targets of other factors. The complexity of these regulatory programs is a challenge for understanding their biological roles. Mouse mutations in splicing regulators often exhibit lethal or highly pleiotropic phenotypes. Among the many affected transcripts, few isoforms are usually sufficiently analyzed to understand their altered activity. Cas9 genome editing offers an exciting new tool for examining differential isoform activity in the developing nervous system. Here we show that it is relatively straightforward to engineer an endogenous locus to express only one splice variant. Combining genome editing with stem cell differentiation allows the characterization of isoform activity in specific cells at specific times in development. This approach opens the way for investigating how other PTB protein targets function within the larger programs of neurogenesis.

MATERIAL AND METHODS

Tissue culture

46C and HB9-GFP mouse ESCs were grown on 0.1% gelatin-coated dish with CF1 mouse embryonic fibroblasts (Applied StemCell, Inc., Menlo Park, CA) in ESC media. ESC media consisted of DMEM (Fisher Scientific, Hampton, NH) sup
lemented with 15% ESC-qualified fetal bovine serum (Life Technologies, Carlsbad, CA), 1X non-essential amino acids (Life Technologies), 1X GlutaMAX (Life Technologies), 1X ESC-qualified nucleosides (EMD Millipore, Billerica, MA), 0.1 mM 2-Mercaptoethanol (Sigma-

Aldrich, St. Louis, MO), and 1000U/ml ESGRO leukemia inhibitor factor (EMD Millipore). 46C ESCs were a gift of J. Sanford, University of California, Santa Cruz. HB9-GFP ESCs were a gift of B. Novitch, University of California, Los Angeles.

46C and HB9-GFP ESCs were differentiated into NPCs according to (Conti et al., 2005), with minor modifications. Briefly, feeder-free ESCs were differentiated on 0.1% gelatin-coated dish in N2B27 media. N2B27 media consisted of 1:1 mixture of Neurobasal media (Life Technologies) and DMEM/F12 (Fisher Scientific) supplemented with 0.5X B27 without Vitamin A (Life Technologies), 0.5X N2 Supplement (Life Technologies), 1X GlutaMAX (Life Technologies), and 0.1 mM 2-Mercaptoethanol (Sigma-Aldrich). After seven days of differentiation, cells were trypsinized and grown in aggregate culture for two days in N2B27 supplemented with 10 ng/ml recombinant human EGF (PeproTech, Rocky Hill, NJ) and 10 ng/ml recombinant human FGF-basic (PeproTech). Aggregates were plated on poly-ornithine-coated (15 ug/ml, Sigma) and fibronectin-coated (1.5 ug/ml, Sigma-Aldrich) dishes in NPC media. NPC media consisted of DMEM/F12 (Fisher Scientific) supplemented 1X B27 without Vitamin A (Life Technologies), 1X GlutaMAX (Life Technologies), 0.1 mM 2-Mercaptoethanol (Sigma-Aldrich), 10 ng/ml recombinant human EGF (PeproTech), and 10 ng/ml recombinant human FGF-basic (PeproTech). NPCs were maintained on poly-ornithine-coated (15 ug/ml, Sigma-Aldrich) and fibronectin-coated (1.5 ug/ml, Sigma-Aldrich) dishes in NPC media for one to two passages. NPCs were differentiated with the removal of EGF and FGF, along with the addition of 1 uM all-trans retinoic acid (RA, Sigma-Aldrich).

HB9-GFP ESCs were differentiated to MNs according to (Wichterle et al., 2002), with minor modifications. Briefly, feeder-free ESCs were grown as aggregate culture for two days in MN media. MN media consisted of 1:1 mixture of Neurobasal media (Life Technologies) and

DMEM/F12 (Fisher Scientific) supplemented with 10% Knockout Serum Replacement (Life Technologies), 1X GlutaMAX (Life Technologies), and 0.1 mM 2-Mercaptoethanol (Sigma-Aldrich). After two days, the MN media was supplemented with 1X N2 Supplement (Life Technologies), 1 uM all-trans retinoic acid (RA, Sigma-Aldrich), and 1 uM smoothened agonist (EMD Millipore). After five days of RA and SAG addition, EBs were plated on BD Matrigel (BD Biosciences), with the addition of 10 ng/ml BDNF (R&D Systems, Minneapolis, MN), 10 ng/ml GDNF (R&D Systems), and 10 ng/ml CNTF (R&D Systems).

Cell sorting of GFP+ MNs

EBs were collected 5 or 8 days after RA and SAG addition. EBs were dissociated in Acutase (Innovative Cell Technologies, San Diego, CA) for 10 minutes and sorted using a BD FACSAria cell sorter at the UCLA Broad Stem Cell center core facility.

siRNA transfections

46C ESCs and NPCs were transfected with Silencer Select Ptbp1 (Life Technologies; s72335, s72336) and Ptbp2 (Life Technologies; s80149, s80149) siRNAs using RNAiMax (Life Technologies) according to the manufacturer's recommendations. Silencer Negative Control siRNA #1 (Life Technologies) and siGENOME Non-Targeting siRNA Pool #1 (GE Healthcare, Pittsburg, PA) were used as negative controls. ESCs were reverse transfected with 20 nM of siRNAs, treated again 24 hours later, and collected 72 hours post-transfections. NPCs were reverse transfected with 10 nM of siRNAs and collected 48 hours post-transfections.

Western blots and immunocytochemistry

Western blots were performed on total protein from cell cultures lysed in RIPA buffer supplemented with protease inhibitors (Roche, Basel, Switzerland) and Benzonase (Sigma-Aldrich). Lysates were diluted in 4X SDS loading buffer, heated for 10 min at 95°C, and loaded onto 10% polyacrylamide Laemmli SDS PAGE gels. Gels were run under standard electrophoresis conditions. Transfers were performed on a Novex X-Cell mini-cell transfer apparatus (Life Technologies) onto Immobilon-FL PVDF membranes (EMD Millipore). The membranes were probed under standard conditions with primary antibodies overnight at 4°C. The following primary antibodies were used: rabbit anti-PTBP1 antibody PTB-NT (1:3000) (Markovtsov et al., 2000), rabbit anti-PTBP2 antibody nPTB-IS2 (1:1000) (Sharma et al., 2005), rabbit anti-Pbx1 antibody (1:500, Cell Signaling, Danvers, MA), rabbit anti-U170K antibody (1:1000) (Sharma et al., 2005), and mouse anti-GAPDH 65C antibody (1:10000, Life Technologies). For fluorescent detection, the membranes were probed with ECL Plex Cy3 and Cy5-conjugated goat anti-mouse and goat anti-rabbit secondary antibodies (1:2000; GE Healthcare). The blots were then scanned on a Typhoon Phosphorimager (GE Healthcare) and quantified using ImageQuant software (GE Healthcare). For chemiluminescent detection, the membranes were probed with Amersham ECL HRP Conjugated Antibodies (1:4000, GE Healthcare) and developed using SuperSignal West Femto Maximum Sensitivity Substrate (Life Technologies) and Kodak BioMax XAR film (Sigma-Aldrich).

Cell cultures were fixed in 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA). Adherent fixed cells were incubated in permeabilization buffer (PBS, 0.25% Triton X-100), washed with block solution (PBS, 0.1% Triton X-100, 1% bovine serum albumin), and probed with primary antibodies overnight at 4°C. Fixed EB cell cultures were cryoprotected in 30% sucrose-PBS, frozen in Tissue-Tek OCT (Electron Microscopy Sciences),

and stored at −80°C until use. 10-μm sections were prepared on a cryostat. Sections were washed with block solution, and probed with primary antibodies overnight at 4°C. The following primary antibodies were used: chicken anti-GFP antibody (1:1000, Abcam, Cambridge, United Kingdom), rabbit anti-HB9 antibody (1:8000) (Arber et al., 1999), rabbit anti-PTBP2 antibody nPTB-IS2 (1:1000), mouse anti-Nestin antibody (1:250, Developmental Studies Hybridoma Bank, Iowa City, IA), and mouse-anti TuJ1 antibody (1:1000, Covance, Princeton, NJ). Samples were washed in PBS before incubation with Alexa-conjugated secondary antibodies (Life Technologies) for 2 hours at room temperature. Samples were rinsed in PBS and mounted with Prolong Gold AntiFade containing nuclear stain DAPI (Life Technologies). Images were acquired using the Carl Zeiss Laser Scanning System LSM 510 META confocal microscope.

RNA isolation, RT-PCR, and real-time PCR

Total RNA was collected from cell cultures using Trizol (Life Technologies) according to the manufacturer's instructions. RNA was quantified (A260) using a Nanodrop-1000 spectrophotometer (Nanodrop Technologies/Thermo Fisher, San Jose, CA). Total RNA (0.5–1 μg) was used for each sample in a 10-μL reaction with 0.25 μL of SuperScript III RT (Life Technologies). PCR reactions contained 200,000–500,000 counts per minute of 32P-labeled or 0.2 μM FAM-labeled reverse primer. PCR reactions were run for 24 cycles for ESCs, 19 cycles for NPCs, and 20 cycles for MN cultures at an annealing temperature of 60°C. PCR reactions were mixed 1:2 with 95% formamide and loaded onto 8% polyacrylamide, 7.5 M urea gels. Gels were run under standard electrophoresis conditions, dried, and imaged on a Typhoon Imager (GE Healthcare). Bands were quantified with ImageQuant software (GE Healthcare). Real-time PCR was performed using SensiFAST SYBER Lo-ROX Kit (Bioline, London, United Kingdom) on a

QuantStudio 6 Real-Time PCR System (Life Technologies) according to the manufacturers' instructions. Relative mRNA levels were determined using a standard curve with beta-actin as a normalizing control.

The splicing primer sequences are Pbx1_F, 5'- GTCACAGCCACCAATGTGTC-3';

Pbx1_R, 5'-TGCGAGTCCGTCACTGTATC-3'.

The real-time primer sequences are

Pbx1-RT_F, 5'-TGGAGAAGTATGAGCAGGCA-3';

Pbx1-RT_R, 5'-CTGGATGGAGCTGAACTTGC-3';

Actb-RT_F, 5'-GGCTGTATTCCCCTCCATCG-3';

Actb-RT_R, 5'-CCAGTTGGTAACAATGCCATGT-3'.

Generation of CRISPR cell lines

Guide RNAs targeting Pbx1 intron 6 were designed using http://crispr.mit.edu/ and cloned into a modified pX330-U6-Chimeric_BB-CBh-hSpCas9 plasmid (gift of B. Stahl and J. Doudna, University of California, Berkley) according to the published protocol (Ran et al., 2013). To generate Cas9-targeted deletions, HB9-GFP ESCs were transfected with two modified pX330 constructs and pBABE-puro (Addgene, Cambridge, MA; #42230) using BioT (Bioland Scientific, Paramount, CA) according to the manufacturer's recommendations. Transfected cells were then treated with 0.5 ug/ml puromycin (InVivoGen, San Diego, CA). Following puromycin selection, clonal ESC lines were isolated and genotyped for intron 6 deletions. Some intron 6 deletions using guides 1 and 4 resulted in the loss of splice sites needed for exon 7 splicing. After genotyping to identify intron 6 deletions, ESC clones were subjected to secondary tests to confirm exon 7 splicing.

The guide RNA sequences are

Pbx1-I6_g1, 5'-GGAACCTCAATCATGTGCCC-3';

Pbx1-I6_g2, 5'-CGGCTAGATAGTCTCTGCGT-3';

Pbx1-I6_g3, 5'-CTAACAGACTGTAACTTGTC-3';

Pbx1-I6_g4, 5'-TCCTCTTGGCTGTTGGTTGC-3'.

The genotyping primer sequences are

Pbx1-I6_g1-4_del_F, 5'-CTTGCTGTCCCCATTGTGTC-3';

Pbx1-I6_g1-4_del_R, 5'- ATCCCCATTGAGTGACTGCA-3';

Pbx1-I6_g1-4_wt_F, 5'- TGTACTCTCCTGCTGTGTGA-3';

Pbx1-I6_g1-4_wt_R, 5'- ATCCCCATTGAGTGACTGCA-3';

Pbx1-I6_g2-3_F, 5'-GTTGTAAGTATGCACCCCGG-3';

Pbx1-I6_g2-3_R, 5'-TGGGGAGTTTGCATCCTCAT-3';

Pbx1-I6_g3-4_F, 5'- AGGAAGTTCAGGCCAAAGTCT-3';

Pbx1-I6_g3-4_R, 5'- ATCCCCATTGAGTGACTGCA-3'.

iCLIP sequencing and data analysis

46C ESCs and NPCs were cross-linked at 100 mJ/cm2. Cell pellets were collected and flash-frozen for iCLIP library preparation (König et al., 2010). Modifications to the protocol are described in Supplementary file 11. Flag and PTBP1 libraries were subjected to 100 bp single-end sequencing at the UCLA Broad Stem Cell center core facility (Illumina HiSeq2000). Data analyses were performed according to (König et al., 2010), with minor modifications. Briefly, PCR duplicates were removed by comparing random portions of the sequenced barcodes. Unique reads were mapped to the mouse genome (mm9/NCBI37) using Bowtie, allowing up to two

nucleotide mismatches (Langmead et al., 2009). Mapped reads were further mapped to the longest transcripts in Known Gene table (Hsu et al., 2006). Crosslink sites were defined as the nucleotide upstream of each iCLIP read, and the False Discovery Rate (FDR) calculated according to (König et al., 2010), with significant crosslink sites (FDR < 0.01) used for clustering and downstream analyses. Significant crosslinking sites were extended 20 nucleotides on either side, and overlapping sites compiled into clusters. Genomic sequences 30 nucleotides upstream and downstream from each crosslink site were used for motif enrichment analyses. Z-scores for all pentamer motifs were calculated by comparing motif frequencies relative to randomly chosen intervals from the same introns. Cassette exons were evaluated for PTBP1 clusters containing ≥ four significant reads, and defined as likely targets if a cluster was present within the cassette exon or within the intron sequence 500 nt upstream or downstream.

RNA sequencing and data analysis

To identify splicing changes during neuronal differentiation, polyA-plus RNA was isolated from HB9-GFP ESCs, NPCs, and Day 5 GFP+ MNs. Paired-end libraries were constructed using the TruSeq mRNA Library Prep Kit (Illumina, San Diego, CA) with modifications to generate strand-specific libraries (Li et al., 2014). The libraries were subjected to 100 nt paired-end sequencing at the UCLA Broad Stem Cell center core facility (Illumina HiSeq2000). Data were analyzed using the Cufflinks pipeline to generate 200-250 million mapped reads per sample (Trapnell et al., 2012). Alternative exon inclusion levels were determined by mapping to exon duo and trio databases using SpliceTrap (Wu et al., 2011). Gene Ontology analyses were performed using DAVID (Huang et al., 2009).

To identify splicing changes following PTB protein depletion, polyA-plus RNA was isolated from 46C ESCs treated with siControl, siPtbp1, or siPtbp1/2 (n=3); and 46C NPCs treated with siControl, siPtbp1, siPtbp1/2, or siPtbp2 (n=2). Paired-end, strand-specific libraries were constructed using the TruSeq Stranded mRNA Library Prep Kit (Illumina). The libraries were subjected to 50 nt paired-end sequencing (Illumina HiSeq2000) to generate 30 to 50 million mapped reads per replicate. The replicates were then pooled for SpliceTrap analyses (Wu et al., 2011).

To identify gene expression changes with Pbx1a induction, polyA-plus RNA was isolated from HB9-GFP Day 2 MNs differentiated from wild type (n=3) and I6 +/- (n=5) ESC clones. Sequencing libraries were constructed using the TruSeq Stranded mRNA Library Prep Kit (Illumina) and subjected to 50 nt single-end sequencing (Illumina HiSeq2000) to generate 25 to 40 million mapped reads per replicate. Differential gene expression was calculated using the Cufflinks pipeline (q-value < 0.05). Normalized FPKM values were calculated by Cuffnorm and were used to perform a Welch's t-test (p-value < 0.05). Gene Ontology analyses were performed using DAVID (Huang et al., 2009). Differentially expressed genes were evaluated for Pbx1 ChIP clusters (Penkov et al., 2013). Pbx1 binding was determined if a ChIP cluster was present within a window 1 kb upstream and downstream of the gene locus.

FIGURE LEGENDS

Figure 2-1: The transition in PTB protein expression occurs during *in vitro* neuronal differentiation. (**A**) HB9-GFP ESCs express eGFP under the MN-specific promoter HB9. The addition of retinoic acid (RA) and a Sonic hedgehog (Shh) agonist drives MN differentiation. (**B**) Western blot shows the loss of PTBP1 protein and gain of PTBP2 and GFP proteins as ESCs differentiate into HB9+ MNs. U170K served as a loading control. (**C**) Quantification of relative PTBP1 and PTBP2 protein expression across MN differentiation. Error bars represent standard error of the mean (SEM, n=2).

Figure 2-2: GFP+ cells express both the postmitotic MN marker HB9 and PTBP2. (**A**) Day 5 MN cultures co-express GFP and HB9. (**B**) Day 8 MN cultures co-express GFP and PTBP2. (**C**) Western blot of PTBP1 and PTBP2 protein expression in ESCs, sorted Day 5 GFP+ MNs, and sorted Day 8 GFP+ MNs. GAPDH served as a loading control.

Figure 2-3: The transition in PTB protein expression occurs as ESCs differentiate into NPCs and neurons in monolayer culture. (**A**) ESCs are differentiated into NPCs in N2B27 media and then maintained in the presence of EGF and FGF. The removal of EGF and FGF, along with the addition of RA, differentiates NPCs into TuJ1+ neurons. (**B**) NPC cultures homogenously express the progenitor marker Nestin, (**C**) while very few cells express the neuronal marker TuJ1. (**D**) Western blot of PTBP1 and PTBP2 as ESCs differentiate into NPCs and neurons.

Figure 2-4: Transitions in alternative splicing occur during neuronal differentiation. (**A**) Alternative splicing events in ESCs, NPCs, and GFP+ MNs were quantified using SpliceTrap (Wu et al., 2011). The bar graph indicates the number of cassette exons, 5' and 3' splice sites, and retained introns whose splicing differs between each sample pair ($|\Delta PSI| \geq 15\%$). (**BC**) Gene ontology analyses of genes showing the differential splice events compared to all expressed genes were performed using DAVID (FDR<0.05) (Huang et al., 2009).

Figure 2-5: PTBP1 regulates a large set of neuronal exons in ESCs and NPCs. (**AB**) Left: Western blots of ESCs and NPCs treated with siControl, siPtbp1, or both siPtbp1 and siPtbp2. Similar results were obtained with an independent set of siRNAs. Right: Bar graphs showing the relative PTBP1 and PTBP2 protein expression ± SEM following siRNA treatment in ESCs (n=3) and NPCs (n=2). Cells with the highest level of either protein are normalized to 1. (**C**) Scatter plots compare the splicing changes in individual exons resulting from PTB protein depletion in ESCs (X-axis) with splicing changes between ESCs and MNs (Y-axis). (**D**) Scatter plots compare the splicing changes resulting from PTB protein depletion in NPCs (X-axis) with splicing changes between NPCs and MNs (Y-axis). (**CD**) Data points in red correspond to neuronally spliced exons that are PTB protein repressed ($\Delta PSI \geq 15\%$). Data points in blue correspond to neuronally skipped exons that are PTB protein activated ($\Delta PSI \leq -15\%$). Data points in grey correspond to the other cassette exons measured by SpliceTrap.

Figure 2-6: **The PTB proteins regulate a large set of exons in ESCs and NPCs.** (**A**) PTBP2 protein is depleted from NPCs using two independent siRNAs. (**B**) Bar chart indicating the number of differentially spliced events between Control samples and PTB protein knockdown

samples in ESCs and NPCs (|ΔPSI| ≥ 15%). (**CD**) Scatter plots compare the splicing changes from PTBP1 depletion with splicing changes from PTBP1 and PTBP2 dual depletion in (**C**) ESCs (**D**) and NPCs.

Figure 2-7: The base content and location of PTBP1 iCLIP-sequencing reads are consistent with known PTBP1 binding properties. (**A**) Autoradiograph of the nitrocellulose membrane containing protein-RNA complexes isolated from ESCs and NPCs using either Flag or PTBP1 antibodies. Boxes indicate the region of the membrane used for iCLIP library preparation. (**B**) Scatter plot compares the calculated pentamer Z-scores from ESCs with scores from NPCs. The top 23 pentamer motifs, which had Z scores >200, are listed. (**CD**) Pie charts showing the percent of the significant PTBP1 iCLIP reads that were found in 5' UTR (blue), coding sequence (CDS) (red), intron (green), and 3' UTR (purple) regions. (**EF**) All expressed cassette exons with nearby PTBP1 iCLIP clusters (within 500 nt, blue circles) are overlapped with PTBP1-responsive cassette exons (green circles, |ΔPSI| ≥ 15%). P-values were calculated using the hypergeometric test.

Figure 2-8: Examples of neuronally spliced cassette exons that are PTBP1 repressed. (**A-D**) Genome Browser tracks show aligned RNA-seq reads from ESCs (red), NPCs (blue), and GFP+ MNs (yellow). The scale indicates the number of mapped reads in the highest peak. PSI values for the cassette exons were calculated with SpliceTrap (Wu et al., 2011). Significant PTBP1 iCLIP sequencing reads from ESCs (magenta) and NPCs (green) are overlaid on the Genome Browser tracks.

Figure 2-9: PTBP1 regulates a switch in Pbx1 isoform expression. (**A**) Exons are represented as boxes, while horizontal lines represent introns. Inclusion of exon 7 results in the longer Pbx1a protein (top). Skipping of exon 7 results in a frameshift and translation termination at a stop codon (*) in exon 8 to yield the shorter Pbx1b protein isoform (bottom). The DNA-binding homeodomain (HD) is indicated by a blue box. (**B**) Genome Browser tracks of aligned RNA-seq reads show exon 7 inclusion as ESCs (red) differentiate into NPCs (blue) and GFP+ MNs (yellow). A significant PTBP1 iCLIP cluster is present upstream of exon 7 in NPCs (green). (**C**) Both Pbx1 exon 7 (solid line, n=3) and Pbx1 mRNA expression (dashed line, n=2) are induced with MN differentiation. (**DE**) ESCs and NPCs were treated with siControl, siPtbp1, or both siPtbp1 and siPtbp2. (**D**) RT-PCR of Pbx1 exon 7 splicing following siRNA treatment in ESCs (24 PCR cycles) and NPCs (19 PCR cycles). (**E**) Bar chart of Pbx1 exon 7 splicing (Mean ± SEM, n=3). Statistical analyses were performed using paired one-tailed Student's t-test (p-value<0.01**, 0.05*).

Figure 2-10: Intron 6 deletions in Pbx1 induce exon 7 splicing. (**A**) Guide RNAs were designed to target regions along Pbx1 intron 6. The location of the PTBP1 iCLIP cluster is indicated by the black box. (**B**) RT-PCR of Pbx1 exon 7 splicing in ESC clones carrying Cas9-mediated deletions. (**C**) Bar chart of Pbx1 exon 7 splicing in ESC clones with intron 6 deletions. Error bars indicate the SEM for three independent clones, except for g3-4 where only one clone was isolated. (**D**) RT-PCR of Pbx1 exon 7 splicing and (**E**) western blot of Pbx1 in D2 MN cultures show that Pbx1a is the predominant isoform in the I6 +/- cell line at day 2.

Figure 2-11: Cas9-targeted deletions of Pbx1 intron 6 switch Pbx1 isoform expression. (**A**) Schematic of Cas9-targeted deletion of Pbx1 intron 6. (**B**) Location of primers used for intron 6 genotyping PCRs. A list of primer sequences is provided in Supplementary file 10 (**C**) Genotyping PCR of the wild type and the intron 6 deletion cell lines. Table showing the number of clones screened and number of edited cell lines isolated. (**D**) Bar graph of Pbx1 exon 7 splicing in D2 MN wild type (n=3 clones) and I6 +/- (n=5 clones) cultures. Error bars indicate SEM.

Figure 2-12: The early expression of Pbx1a promotes neuronal gene expression. (**A**) Genome Browser tracks of aligned RNA-seq reads from wild type (top) and I6 +/- (bottom) Day 2 MN cultures. The asterisk (*) indicates Pbx1 exon 7, which is induced in the mutant. (**B**) Bar charts of the top 20 induced genes in the I6 +/- cell lines compared to expression in wild type cells. Genes highlighted in bold have known roles in neuronal differentiation. Error bars indicate SEM. (**C**) Gene ontology analysis of the 196 genes induced by 1.5 fold in the I6 +/- cultures. (**D**) Bar chart of Hoxa3 expression. Statistical analyses were performed using Welch's t-test (p-value<0.01**).

Figure 2-1: The transition in PTB protein expression occurs during *in vitro* neuronal differentiation.

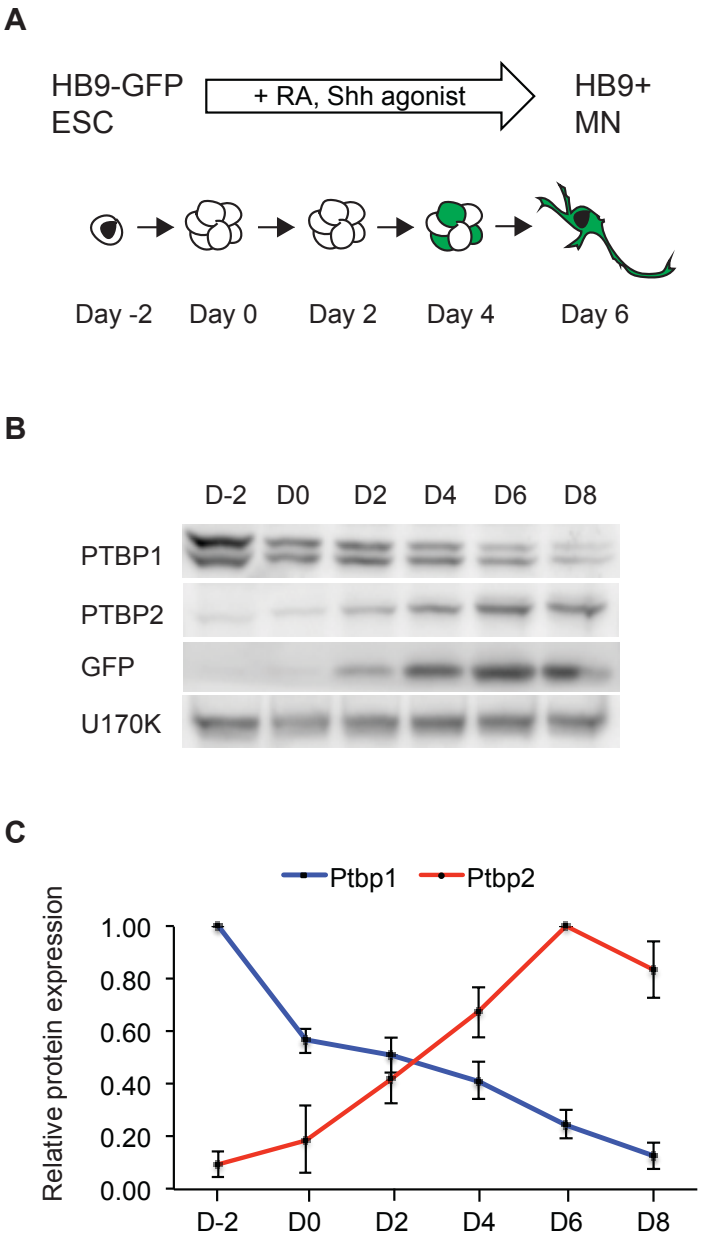


Figure 2-2: GFP+ cells express both the postmitotic MN marker HB9 and PTBP2.

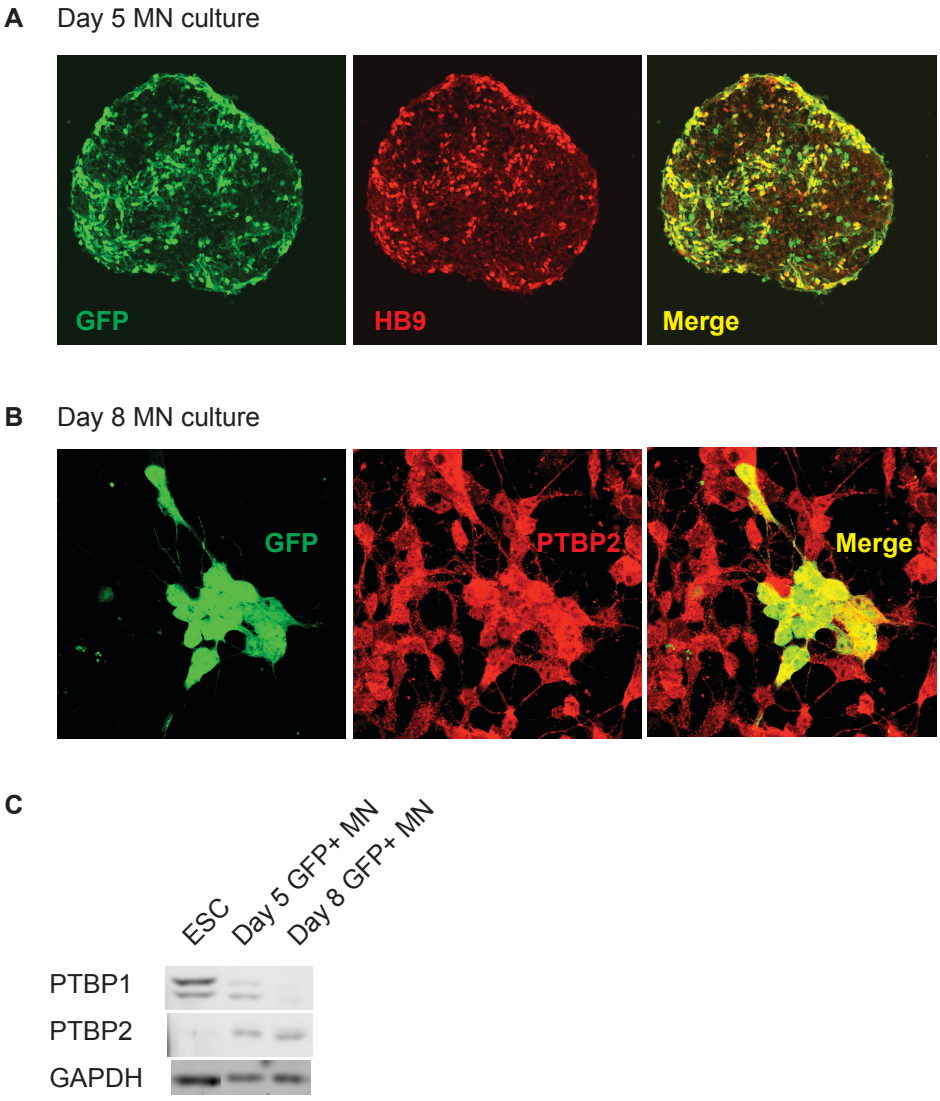


Figure 2-3: The transition in PTB protein expression occurs as ESCs differentiate into NPCs and neurons in monolayer culture.

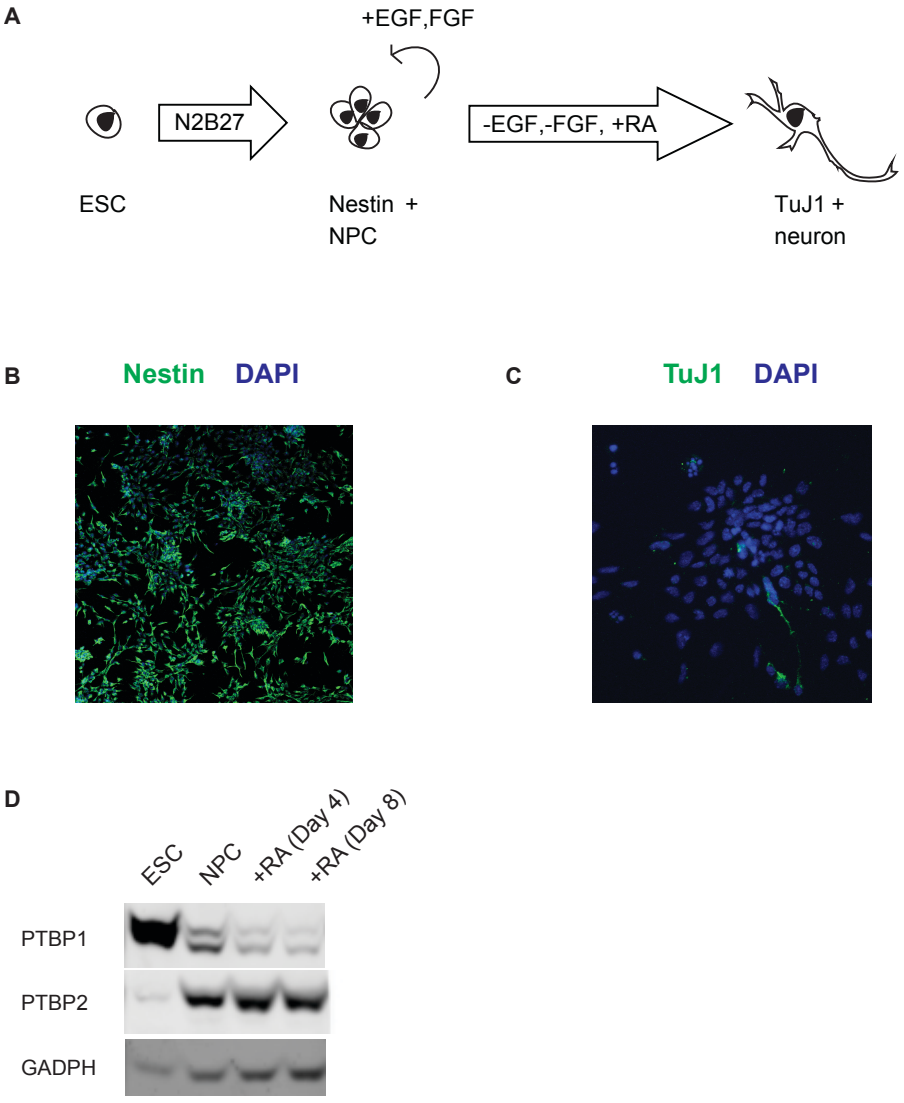


Figure 2-4: Transitions in alternative splicing occur during neuronal differentiation.

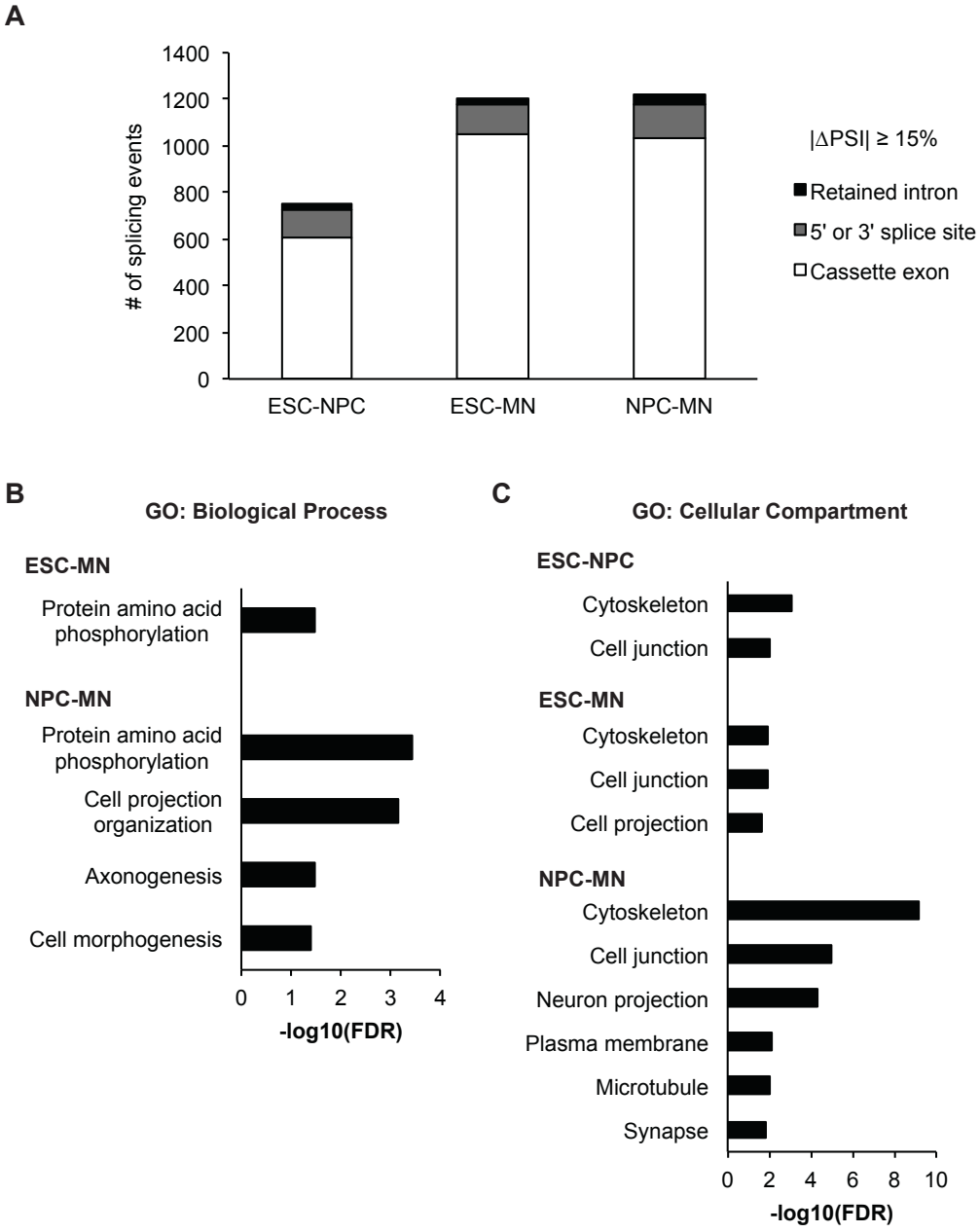


Figure 2-5: PTBP1 regulates a large set of neuronal exons in ESCs and NPCs.

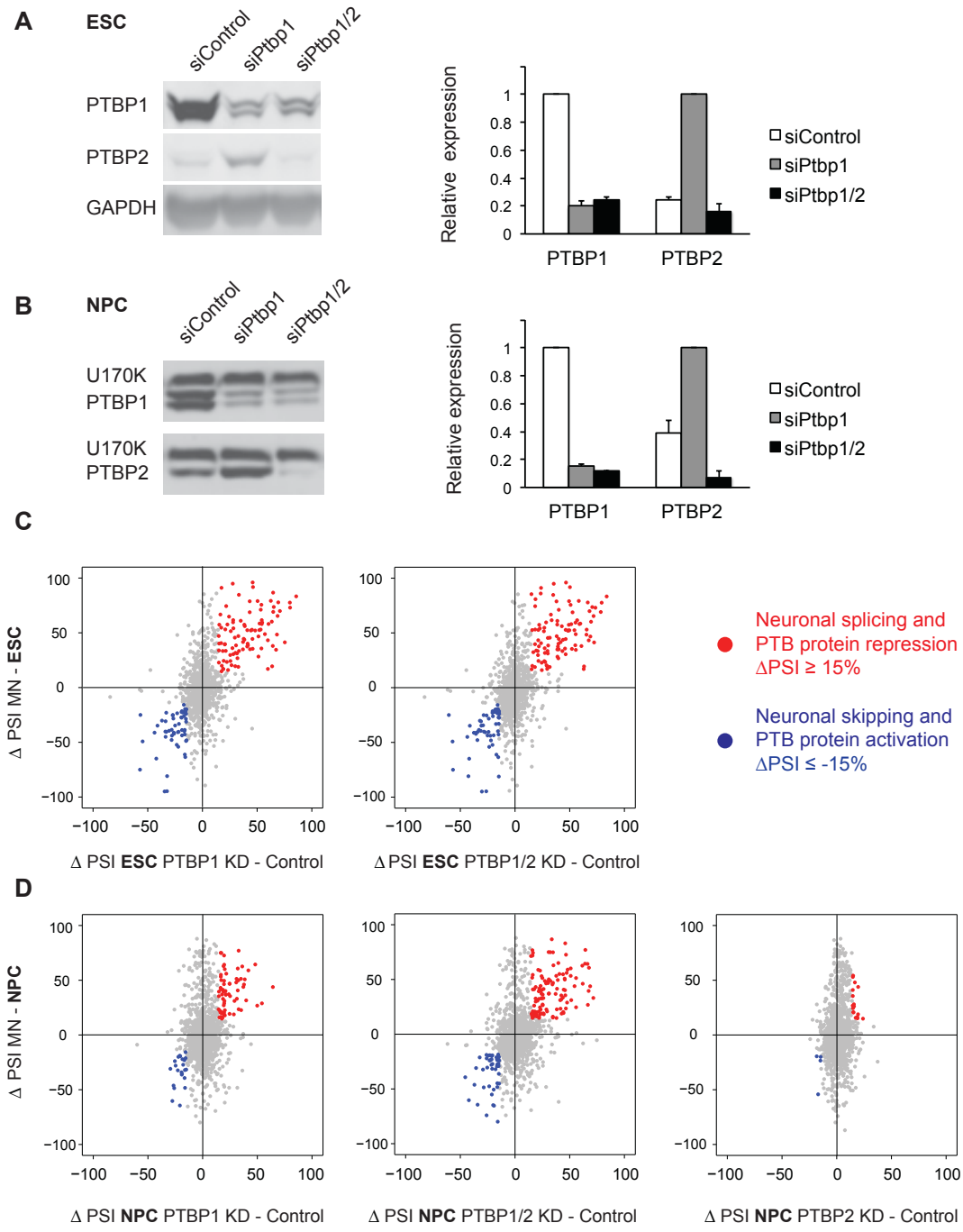


Figure 2-6: **The PTB proteins regulate a large set of exons in ESCs and NPCs.**

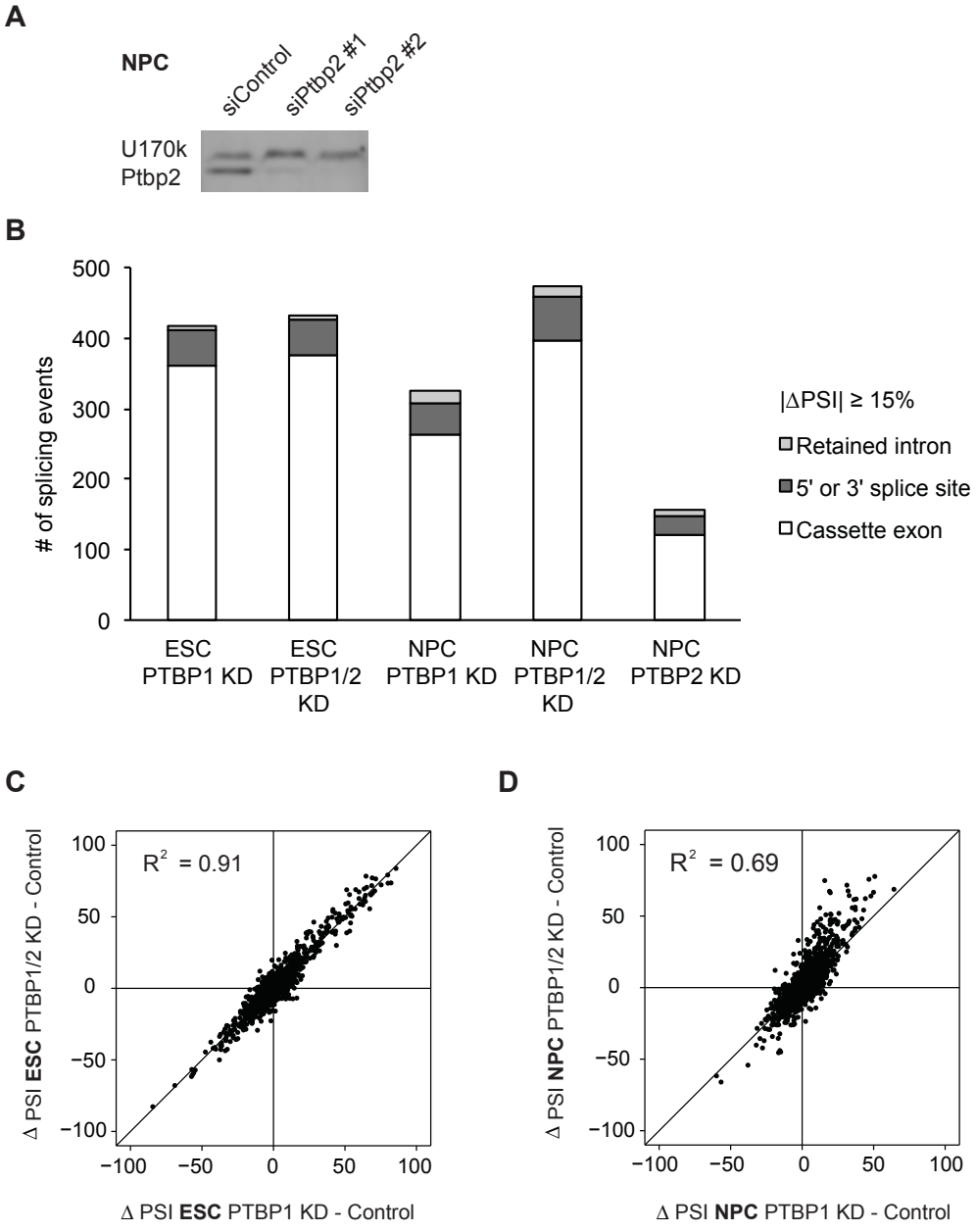
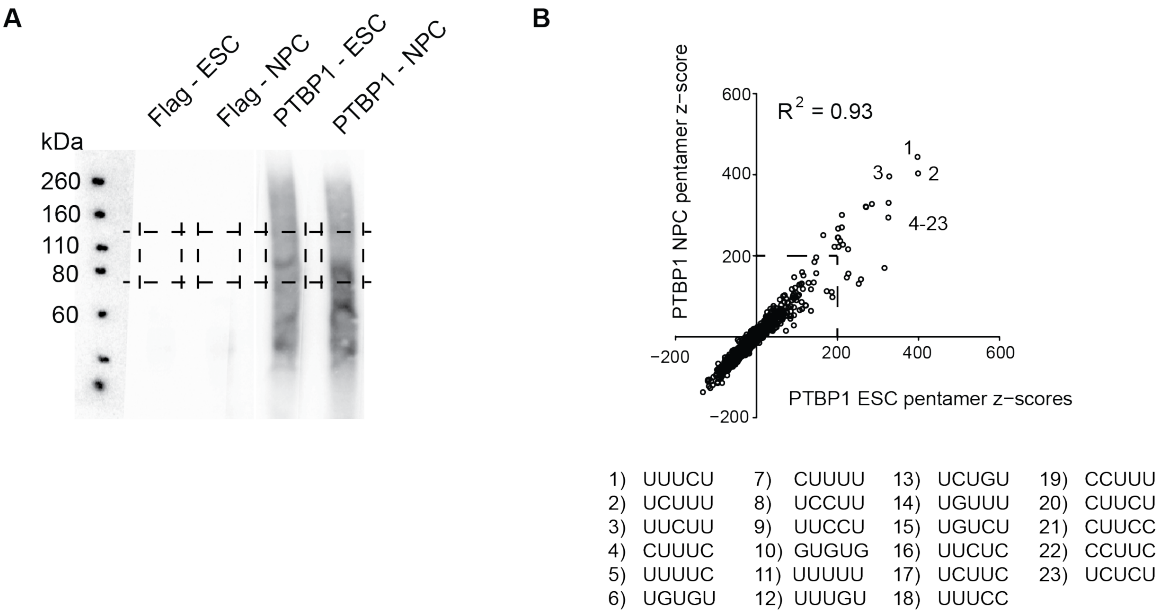
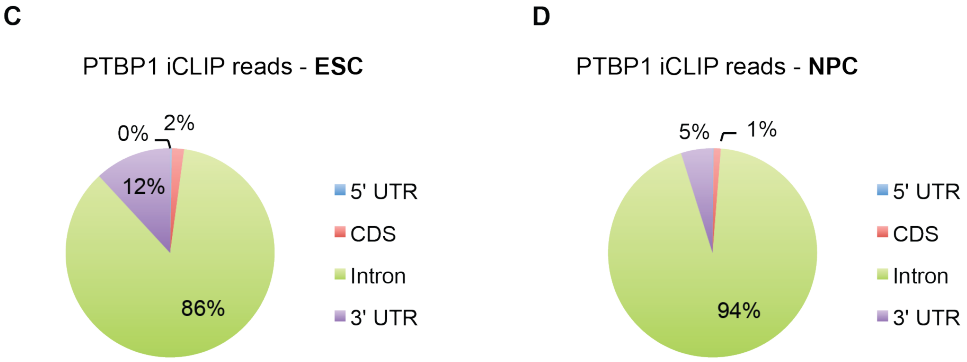


Figure 2-7: The base content and location of PTBP1 iCLIP-sequencing reads are consistent with known PTBP1 binding properties.





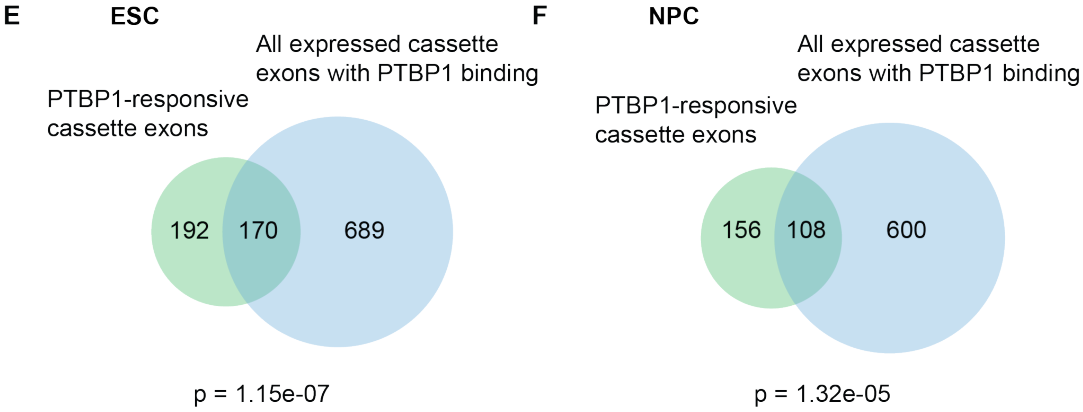


Figure 2-8: Examples of neuronally spliced cassette exons that are PTBP1 repressed.

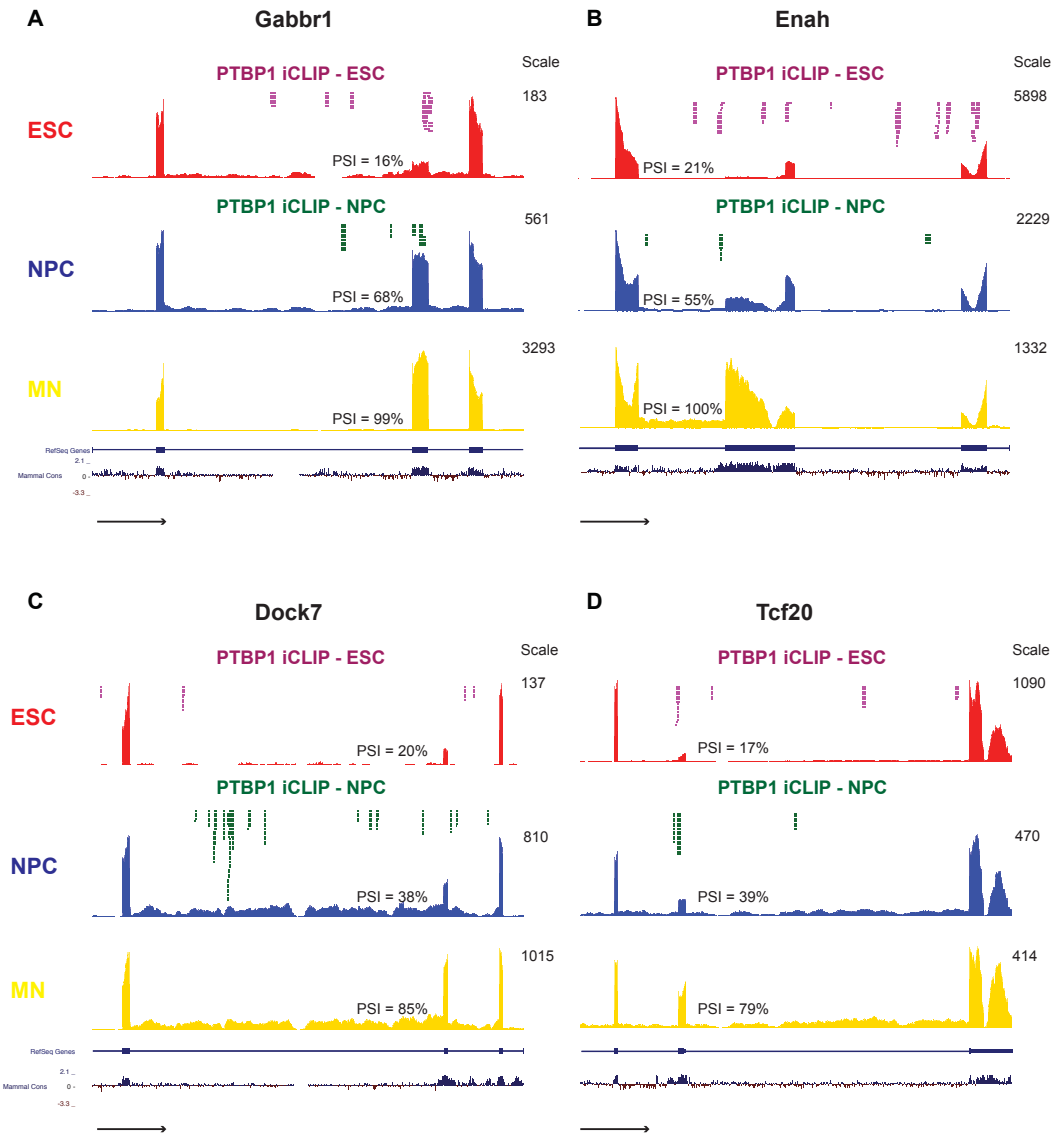


Figure 2-9: PTBP1 regulates a switch in Pbx1 isoform expression.

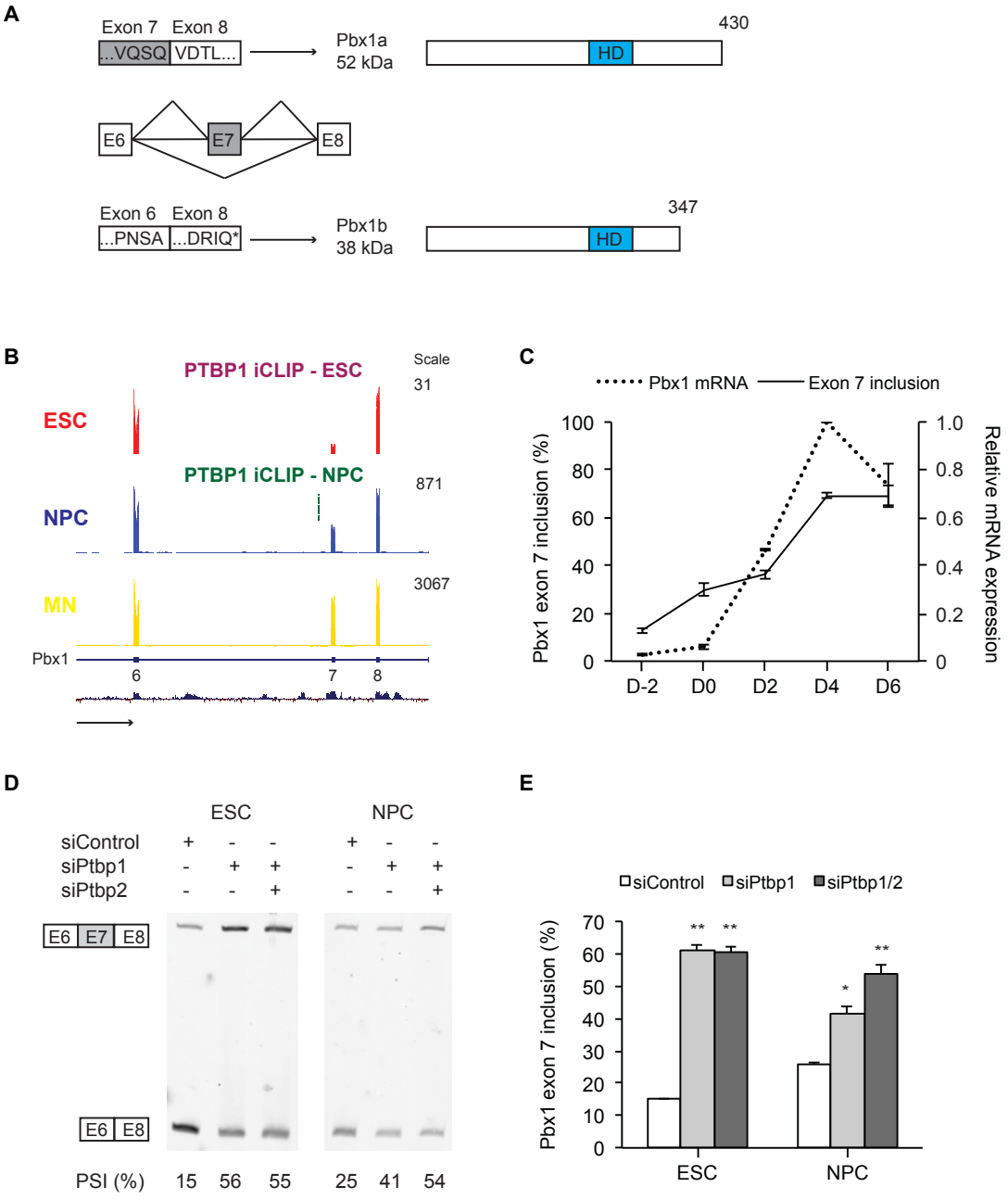


Figure 2-10: Intron 6 deletions in Pbx1 induce exon 7 splicing.

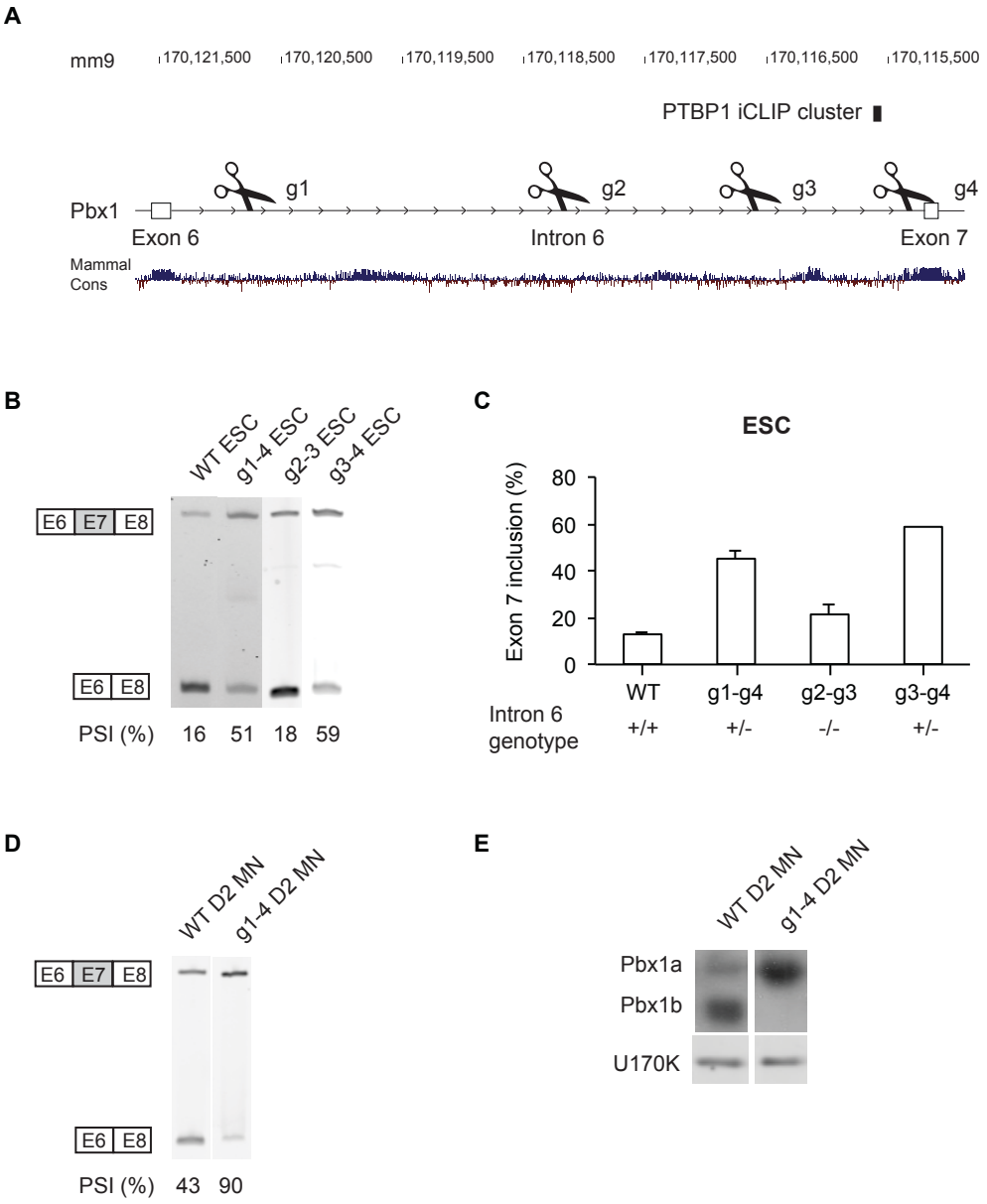


Figure 2-11: Cas9-targeted deletions of Pbx1 intron 6 switch Pbx1 isoform expression.

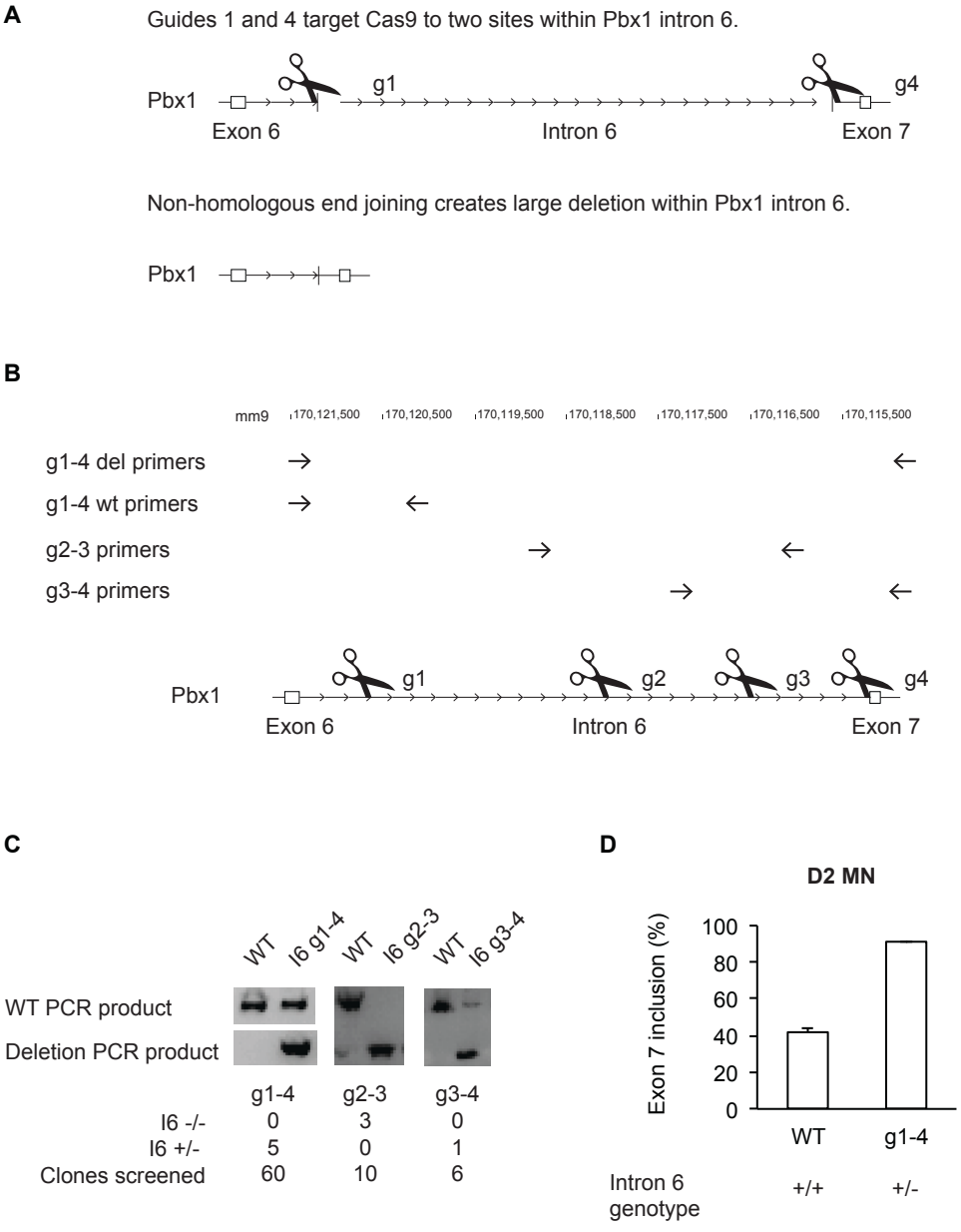


Figure 2-12: The early expression of Pbx1a promotes neuronal gene expression.

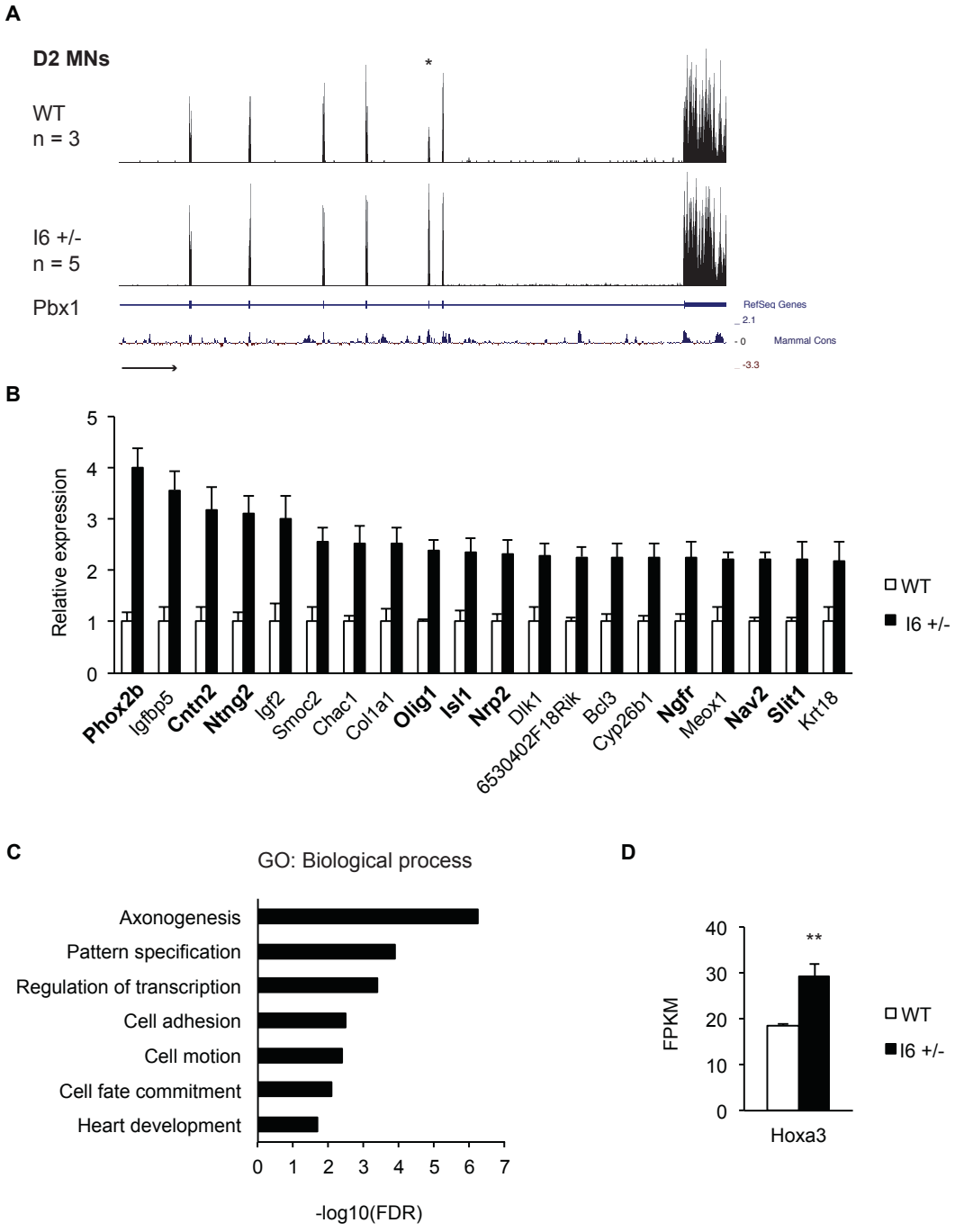


Table 2-1: Top 20 exons co-repressed by PTBP1 and PTBP2 in NPCs.

Genesymbol	Coordinates	dPSI PTBP1 KD	dPSI PTBP2 KD	dPSI PTBP1-2 KD	Description
Magi1	chr6:93630782-93630871(-)	0.16	0.03	0.75	membrane-associated guanylate kinase, WW and PDZ
Ap2a1	chr7:52158831-52158897(-)	0.19	0.04	0.67	AP-2 complex subunit alpha-1 isoform a
Tnrc6a	chr7:130323762-130323909(+)	0.20	0.07	0.66	trinucleotide repeat-containing gene 6A protein
Clip2	chr5:134980603-134980708(-)	0.12	0.06	0.56	CAP-Gly domain-containing linker protein 2
Ap2b1	chr11:83178296-83178338(+)	0.20	0.14	0.51	AP-2 complex subunit beta isoform a
Pdlim7	chr13:55609224-55609241(-)	0.17	0.04	0.51	PDZ and LIM domain protein 7 isoform a
Thoc5	chr11:4811437-4811473(+)	0.13	0.05	0.48	THO complex subunit 5 homolog
Vav2	chr2:27123670-27123757(-)	0.20	0.06	0.48	guanine nucleotide exchange factor VAV2
Sbf2	chr7:117484803-117484878(-)	0.20	0.07	0.47	SET-binding factor 2
Slit2	chr5:48582241-48582253(+)	0.11	0.00	0.46	Mus musculus neurogenic extracellular slit protein (Slit2) mRNA, partial cds.
Scrib	chr15:75879886-75879970(-)	0.20	0.09	0.44	protein scribble homolog
Gripap1	chrX:7380719-7380812(+)	0.08	0.12	0.44	SubName: Full=GRIP1 associated protein 1; Flags: Fragment;
Osbpl9	chr4:108780485-108780554(-)	0.13	0.04	0.44	oxysterol-binding protein-related protein 9
Rpn2	chr2:157148958-157149006(+)	0.15	-0.01	0.42	dolichyl-diphosphooligosaccharide--protein
Dlg4	chr11:69857159-69857269(+)	0.08	0.13	0.42	disks large homolog 4 isoform 1
Map3k4	chr17:12435515-12435671(-)	0.17	0.08	0.41	mitogen-activated protein kinase kinase kinase
Phldb1	chr9:44509028-44509169(-)	0.10	-0.03	0.40	pleckstrin homology-like domain family B member
Ptk2	chr15:73042416-73042425(-)	0.20	0.10	0.39	focal adhesion kinase 1 isoform 1
Tpm3	chr3:89894934-89895013(+)	0.16	0.03	0.39	tropomyosin alpha-3 chain
Arhgef1	chr7:25704202-25704370(+)	0.10	0.03	0.38	rho guanine nucleotide exchange factor 1 isoform

Table 2-2: Top 20 exons directly repressed by PTBP1 in ESCs.

Genesymbol	Coordinates	dPSI PTBP1 KD	Description
Gabbr1	chr17:37204158-37204309(+)	0.86	gamma-aminobutyric acid type B receptor subunit
Lass5	chr15:99567977-99568084(-)	0.80	LAG1 longevity assurance homolog 5
Dpf2	chr19:5903730-5903772(-)	0.75	zinc finger protein ubi-d4
Kiaa1310	chr1:36404061-36404139(-)	0.71	RecName: Full=Uncharacterized protein KIAA1310;
Sbf2	chr7:117484803-117484878(-)	0.69	SET-binding factor 2
Fnbp1	chr2:30900401-30900416(-)	0.69	formin-binding protein 1 isoform e
Fbxo25	chr8:13938921-13938948(+)	0.65	F-box only protein 25
Tnk2	chr16:32678754-32678799(+)	0.64	activated CDC42 kinase 1 isoform 1
Depdc5	chr5:33301929-33301995(+)	0.64	DEP domain-containing protein 5 isoform 1
Rapgef1	chr2:29565675-29565828(+)	0.63	Rap guanine nucleotide exchange factor (GEF) 1
Enah	chr1:183851788-183852631(-)	0.63	protein enabled homolog isoform 2
Osbpl9	chr4:108780485-108780554(-)	0.60	oxysterol-binding protein-related protein 9
R3hdm1	chr1:130087214-130087316(+)	0.54	R3H domain (binds single-stranded nucleic
Tcf20	chr15:82644910-82645033(-)	0.54	transcription factor 20 isoform a
Mllt4	chr17:13962949-13962994(+)	0.54	afadin
Vav2	chr2:27123670-27123757(-)	0.54	guanine nucleotide exchange factor VAV2
Ap1b1	chr11:4933244-4933265(+)	0.53	AP-1 complex subunit beta-1
ra175a	chr9:47637459-47637492(+)	0.53	SubName: Full=RA175 isoform j; Flags: Fragment;
Ctnnd1	chr2:84466176-84466294(-)	0.52	catenin delta-1 isoform 1
Arhgef1	chr7:25704202-25704370(+)	0.51	rho guanine nucleotide exchange factor 1 isoform

Table 2-3: Top 20 exons directly repressed by PTBP1 in NPCs.

Genesymbol	Coordinates	dPSI PTBP1 KD	Description
Kiaa1310	chr1:36404061-36404139(-)	0.64	RecName: Full=Uncharacterized protein KIAA1310;
Dennd1a	chr2:37660591-37660720(-)	0.54	DENN domain-containing protein 1A
Rod1	chr4:59559020-59559054(-)	0.50	regulator of differentiation 1 isoform 1
Ralgapa1	chr12:56836020-56836161(-)	0.48	ral GTPase-activating protein subunit alpha-1
Mink1	chr11:70422909-70422933(+)	0.42	misshapen-like kinase 1 isoform 1
Gabbr1	chr17:37204158-37204309(+)	0.40	gamma-aminobutyric acid type B receptor subunit
Tbc1d9b	chr11:49977663-49977714(+)	0.39	TBC1 domain family member 9B
Lass5	chr15:99567977-99568084(-)	0.36	LAG1 longevity assurance homolog 5
Rtn4	chr11:29605534-29605591(+)	0.35	reticulon-4 isoform A
Pbrm1	chr14:31836252-31836365(+)	0.35	SubName: Full=Putative uncharacterized protein Pbrm1;
Wnk1	chr6:119903788-119904250(-)	0.33	serine/threonine-protein kinase WNK1 isoform 5
Fbxo25	chr8:13938921-13938948(+)	0.33	F-box only protein 25
Exoc7	chr11:116158884-116158977(-)	0.32	exocyst complex component 7 isoform 1
Fnbp1	chr2:30900401-30900416(-)	0.32	formin-binding protein 1 isoform e
Lphn3	chr5:82092410-82092449(+)	0.31	latrophilin-3
Lphn1	chr8:86450160-86450175(+)	0.31	latrophilin-1 precursor
Tnk2	chr16:32678754-32678799(+)	0.29	activated CDC42 kinase 1 isoform 1
Rtn4	chr11:29606409-29608770(+)	0.25	reticulon-4 isoform A
Ctnnd1	chr2:84466176-84466294(-)	0.25	catenin delta-1 isoform 1
Dock7	chr4:98660337-98660427(-)	0.25	dedicator of cytokinesis protein 7

REFERENCES

Abranches, E., Silva, M., Pradier, L., Schulz, H., Hummel, O., Henrique, D., and Bekman, E. (2009). Neural differentiation of embryonic stem cells in vitro: a road map to neurogenesis in the embryo. PLoS ONE *4*, e6286.

Adams, K.L., Rousso, D.L., Umbach, J.A., and Novitch, B.G. (2015). Foxp1-mediated programming of limb-innervating motor neurons from mouse and human embryonic stem cells. Nat Commun *6*, 6778.

Amir-Ahmady, B., Boutz, P.L., Markovtsov, V., Phillips, M.L., and Black, D.L. (2005). Exon repression by polypyrimidine tract binding protein. Rna *11*, 699–716.

Arber, S., Han, B., Mendelsohn, M., Smith, M., Jessell, T.M., and Sockanathan, S. (1999). Requirement for the homeobox gene Hb9 in the consolidation of motor neuron identity. Neuron *23*, 659–674.

Asahara, H., Dutta, S., Kao, H.Y., Evans, R.M., and Montminy, M. (1999). Pbx-Hox heterodimers recruit coactivator-corepressor complexes in an isoform-specific manner. Mol. Cell. Biol. *19*, 8219–8225.

Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov, S., Lee, L.J., Slobodeniuc, V., Kutter, C., Watt, S., Colak, R., et al. (2012). The evolutionary landscape of alternative splicing in vertebrate species. Science *338*, 1587–1593.

Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. Annu. Rev. Biochem. *72*, 291–336.

Boutz, P.L., Stoilov, P., Li, Q., Lin, C.-H., Chawla, G., Ostrow, K., Shiue, L., Ares, M., and Black, D.L. (2007). A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. Genes Dev. *21*, 1636–1652.

Braunschweig, U., Gueroussov, S., Plocik, A.M., Graveley, B.R., and Blencowe, B.J. (2013). Dynamic integration of splicing within gene regulatory pathways. Cell *152*, 1252–1269.

Calarco, J.A., Superina, S., O'Hanlon, D., Gabut, M., Raj, B., Pan, Q., Skalska, U., Clarke, L., Gelinas, D., van der Kooy, D., et al. (2009). Regulation of vertebrate nervous system alternative splicing and development by an SR-related protein. Cell *138*, 898–910.

Charizanis, K., Lee, K.-Y., Batra, R., Goodwin, M., Zhang, C., Yuan, Y., Shiue, L., Cline, M., Scotti, M.M., Xia, G., et al. (2012). Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. Neuron *75*, 437–450.

Coelho, M.B., Attig, J., Bellora, N., König, J., Hallegger, M., Kayikci, M., Eyras, E., Ule, J., and Smith, C.W.J. (2015). Nuclear matrix protein Matrin3 regulates alternative splicing and forms overlapping regulatory networks with PTB. Embo J. *34*, 653–668.

Cong, L., and Zhang, F. (2015). Genome engineering using CRISPR-Cas9 system. Methods Mol. Biol. *1239*, 197–217.

Conti, L., Pollard, S.M., Gorba, T., Reitano, E., Toselli, M., Biella, G., Sun, Y., Sanzone, S., Ying, Q.-L., Cattaneo, E., et al. (2005). Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. PLoS Biol. *3*, e283.

Demir, E., and Dickson, B.J. (2005). fruitless splicing specifies male courtship behavior in Drosophila. Cell *121*, 785–794.

Di Rocco, G., Mavilio, F., and Zappavigna, V. (1997). Functional dissection of a transcriptionally active, target-specific Hox-Pbx complex. Embo J. *16*, 3644–3654.

Doudna, J.A., and Charpentier, E. (2014). Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096.

Fu, X.-D., and Ares, M. (2014). Context-dependent control of alternative splicing by RNA-binding proteins. Nat. Rev. Genet. *15*, 689–701.

Gabut, M., Samavarchi-Tehrani, P., Wang, X., Slobodeniuc, V., O'Hanlon, D., Sung, H.-K., Alvarez, M., Talukder, S., Pan, Q., Mazzoni, E.O., et al. (2011). An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. Cell *147*, 132–146.

Gehman, L.T., Meera, P., Stoilov, P., Shiue, L., O'Brien, J.E., Meisler, M.H., Ares, M., Otis, T.S., and Black, D.L. (2012). The splicing regulator Rbfox2 is required for both cerebellar development and mature motor function. Genes Dev. *26*, 445–460.

Gehman, L.T., Stoilov, P., Maguire, J., Damianov, A., Lin, C.-H., Shiue, L., Ares, M., Mody, I., and Black, D.L. (2011). The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. Nat. Genet. *43*, 706–711.

Grabowski, P.J., and Black, D.L. (2001). Alternative RNA splicing in the nervous system. Prog. Neurobiol. *65*, 289–308.

Han, A., Stoilov, P., Linares, A.J., Zhou, Y., Fu, X.-D., and Black, D.L. (2014). De novo prediction of PTBP1 binding and splicing targets reveals unexpected features of its RNA recognition and function. PLoS Comput. Biol. *10*, e1003442.

Hirabayashi, Y., and Gotoh, Y. (2010). Epigenetic control of neural precursor cell fate during development. Nat. Rev. Neurosci. *11*, 377–388.

Hsu, F., Kent, W.J., Clawson, H., Kuhn, R.M., Diekhans, M., and Haussler, D. (2006). The UCSC Known Genes. Bioinformatics *22*, 1036–1046.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44–57.

Hüttelmaier, S., Illenberger, S., Grosheva, I., Rüdiger, M., Singer, R.H., and Jockusch, B.M.

(2001). Raver1, a dual compartment protein, is a ligand for PTB/hnRNPI and microfilament attachment proteins. J. Cell Biol. *155*, 775–786.

Iijima, T., Wu, K., Witte, H., Hanno-Iijima, Y., Glatter, T., Richard, S., and Scheiffele, P. (2011). SAM68 regulates neuronal activity-dependent alternative splicing of neurexin-1. Cell *147*, 1601–1614.

Ince-Dunn, G., Okano, H.J., Jensen, K.B., Park, W.-Y., Zhong, R., Ule, J., Mele, A., Fak, J.J., Yang, C., Zhang, C., et al. (2012). Neuronal Elav-like (Hu) proteins regulate RNA splicing and abundance to control glutamate levels and neuronal excitability. Neuron *75*, 1067–1080.

Jensen, K.B., Dredge, B.K., Stefani, G., Zhong, R., Buckanovich, R.J., Okano, H.J., Yang, Y.Y., and Darnell, R.B. (2000). Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability. Neuron *25*, 359–371.

Jürgens, A.S., Kolanczyk, M., Moebest, D.C.C., Zemojtel, T., Lichtenauer, U., Duchniewicz, M., Gantert, M.P., Hecht, J., Hattenhorst, U., Burdach, S., et al. (2009). PBX1 is dispensable for neural commitment of RA-treated murine ES cells. In Vitro Cell. Dev. Biol. Anim. *45*, 252–263.

Kamps, M.P., Look, A.T., and Baltimore, D. (1991). The human t(1;19) translocation in pre-B ALL produces multiple nuclear E2A-Pbx1 fusion proteins with differing transforming potentials. Genes Dev. *5*, 358–368.

Keppetipola, N., Sharma, S., Li, Q., and Black, D.L. (2012). Neuronal regulation of pre-mRNA splicing by polypyrimidine tract binding proteins, PTBP1 and PTBP2. Crit. Rev. Biochem. Mol. Biol. *47*, 360–378.

Kim, S.K., Selleri, L., Lee, J.S., Zhang, A.Y., Gu, X., Jacobs, Y., and Cleary, M.L. (2002). Pbx1 inactivation disrupts pancreas development and in Ipf1-deficient mice promotes diabetes mellitus. Nat. Genet. *30*, 430–435.

Kosik, K.S. (2006). The neuronal microRNA system. Nat. Rev. Neurosci. *7*, 911–920.

König, J., Zarnack, K., Luscombe, N.M., and Ule, J. (2011). Protein-RNA interactions: new genomic technologies and perspectives. Nat. Rev. Genet. *13*, 77–83.

König, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat. Struct. Mol. Biol. *17*, 909–915.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

LaRonde-LeBlanc, N.A., and Wolberger, C. (2003). Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. Genes Dev. *17*, 2060–2072.

Lee, Y., and Rio, D.C. (2015). Mechanisms and Regulation of Alternative Pre-mRNA Splicing. Annu. Rev. Biochem. *84*, 291–323.

Li, Q., Lee, J.-A., and Black, D.L. (2007). Neuronal regulation of alternative pre-mRNA splicing. Nat. Rev. Neurosci. *8*, 819–831.

Li, Q., Zheng, S., Han, A., Lin, C.-H., Stoilov, P., Fu, X.-D., and Black, D.L. (2014). The splicing regulator PTBP2 controls a program of embryonic splicing required for neuronal maturation. Elife *3*, e01201–e01201.

Licatalosi, D.D., and Darnell, R.B. (2006). Splicing regulation in neurologic disease. Neuron *52*, 93–101.

Licatalosi, D.D., Yano, M., Fak, J.J., Mele, A., Grabinski, S.E., Zhang, C., and Darnell, R.B. (2012). Ptbp2 represses adult-specific splicing to regulate the generation of neuronal precursors in the embryonic brain. Genes Dev. *26*, 1626–1642.

Llorian, M., Schwartz, S., Clark, T.A., Hollander, D., Tan, L.-Y., Spellman, R., Gordon, A., Schweitzer, A.C., la Grange, de, P., Ast, G., et al. (2010). Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. Nat. Struct. Mol. Biol. *17*, 1114–1123.

Longobardi, E., Penkov, D., Mateos, D., De Florian, G., Torres, M., and Blasi, F. (2014). Biochemistry of the tale transcription factors PREP, MEIS, and PBX in vertebrates. Dev. Dyn. *243*, 59–75.

Louvi, A., and Artavanis-Tsakonas, S. (2006). Notch signalling in vertebrate neural development. Nat. Rev. Neurosci. *7*, 93–102.

Maeda, R., Ishimura, A., Mood, K., Park, E.K., Buchberg, A.M., and Daar, I.O. (2002). Xpbx1b and Xmeis1b play a collaborative role in hindbrain and neural crest gene expression in Xenopus embryos. Proc. Natl. Acad. Sci. U.S.a. *99*, 5448–5453.

Makeyev, E.V., Zhang, J., Carrasco, M.A., and Maniatis, T. (2007). The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. Mol. Cell *27*, 435–448.

Manzanares, M., Bel-Vialar, S., Ariza-McNaughton, L., Ferretti, E., Marshall, H., Maconochie, M.M., Blasi, F., and Krumlauf, R. (2001). Independent regulation of initiation and maintenance phases of Hoxa3 expression in the vertebrate hindbrain involve auto- and cross-regulatory mechanisms. Development *128*, 3595–3607.

Markovtsov, V., Nikolic, J.M., Goldman, J.A., Turck, C.W., Chou, M.Y., and Black, D.L. (2000). Cooperative assembly of an hnRNP complex induced by a tissue-specific homolog of polypyrimidine tract binding protein. Mol. Cell. Biol. *20*, 7463–7479.

Merkin, J., Russell, C., Chen, P., and Burge, C.B. (2012). Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. Science *338*, 1593–1599.

Norris, A.D., and Calarco, J.A. (2012). Emerging Roles of Alternative Pre-mRNA Splicing Regulation in Neuronal Development and Function. Front Neurosci *6*, 122.

Oberstrass, F.C., Auweter, S.D., Erat, M., Hargous, Y., Henning, A., Wenter, P., Reymond, L., Amir-Ahmady, B., Pitsch, S., Black, D.L., et al. (2005). Structure of PTB bound to RNA: specific binding and implications for splicing regulation. Science *309*, 2054–2057.

Pattyn, A., Hirsch, M., Goridis, C., and Brunet, J.F. (2000). Control of hindbrain motor neuron differentiation by the homeobox gene Phox2b. Development *127*, 1349–1358.

Penkov, D., Mateos San Martín, D., Fernandez-Díaz, L.C., Rosselló, C.A., Torroja, C., Sánchez-Cabo, F., Warnatz, H.J., Sultan, M., Yaspo, M.L., Gabrieli, A., et al. (2013). Analysis of the DNA-binding profile and function of TALE homeoproteins reveals their specialization and specific interactions with Hox genes/proteins. Cell Rep *3*, 1321–1333.

Piper, D.E., Batchelor, A.H., Chang, C.P., Cleary, M.L., and Wolberger, C. (1999). Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. Cell *96*, 587–597.

Quesnel-Vallières, M., Irimia, M., Cordes, S.P., and Blencowe, B.J. (2015). Essential roles for the splicing regulator nSR100/SRRM4 during nervous system development. Genes Dev. *29*, 746–759.

Raj, B., Irimia, M., Braunschweig, U., Sterne-Weiler, T., O'Hanlon, D., Lin, Z.-Y., Chen, G.I., Easton, L.E., Ule, J., Gingras, A.-C., et al. (2014). A global regulatory mechanism for activating an exon network required for neurogenesis. Mol. Cell *56*, 90–103.

Raj, B., O'Hanlon, D., Vessey, J.P., Pan, Q., Ray, D., Buckley, N.J., Miller, F.D., and Blencowe, B.J. (2011). Cross-regulation between an alternative splicing activator and a transcription repressor controls neurogenesis. Mol. Cell *43*, 843–850.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc *8*, 2281–2308.

Redmond, L., Hockfield, S., and Morabito, M.A. (1996). The divergent homeobox gene PBX1 is expressed in the postnatal subventricular zone and interneurons of the olfactory bulb. J. Neurosci. *16*, 2972–2982.

Rideau, A.P., Gooding, C., Simpson, P.J., Monie, T.P., Lorenz, M., Hüttelmaier, S., Singer, R.H., Matthews, S., Curry, S., and Smith, C.W.J. (2006). A peptide motif in Raver1 mediates splicing repression by interaction with the PTB RRM2 domain. Nat. Struct. Mol. Biol. *13*, 839–848.

Ronan, J.L., Wu, W., and Crabtree, G.R. (2013). From neural development to cognition: unexpected roles for chromatin. Nat. Rev. Genet. *14*, 347–359.

Samad, O.A., Geisen, M.J., Caronia, G., Varlet, I., Zappavigna, V., Ericson, J., Goridis, C., and Rijli, F.M. (2004). Integration of anteroposterior and dorsoventral regulation of Phox2b transcription in cranial motoneuron progenitors by homeodomain proteins. Development *131*, 4071–4083.

Schnabel, C.A., Selleri, L., Jacobs, Y., Warnke, R., and Cleary, M.L. (2001). Expression of Pbx1b during mammalian organogenesis. Mech. Dev. *100*, 131–135.

Schulte, D., and Frank, D. (2014). TALE transcription factors during early development of the vertebrate brain and eye. *243*, 99–116.

Sgadò, P., Ferretti, E., Grbec, D., Bozzi, Y., and Simon, H.H. (2012). The atypical homeoprotein Pbx1a participates in the axonal pathfinding of mesencephalic dopaminergic neurons. Neural Dev *7*, 24.

Sharma, S., Falick, A.M., and Black, D.L. (2005). Polypyrimidine tract binding protein blocks the 5' splice site-dependent assembly of U2AF and the prespliceosomal E complex. Mol. Cell *19*, 485–496.

Spellman, R., Llorian, M., and Smith, C.W.J. (2007). Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1. Mol. Cell *27*, 420–434.

Tang, Z.Z., Sharma, S., Zheng, S., Chawla, G., Nikolic, J., and Black, D.L. (2011). Regulation of the mutually exclusive exons 8a and 8 in the CaV1.2 calcium channel transcript by polypyrimidine tract-binding protein. J. Biol. Chem. *286*, 10007–10016.

Temple, S. (2001). The development of neural stem cells. Nature *414*, 112–117.

Thiaville, M.M., Stoeck, A., Chen, L., Wu, R.-C., Magnani, L., Oidtman, J., Shih, I.-M., Lupien, M., and Wang, T.-L. (2012). Identification of PBX1 target genes in cancer cells by global mapping of PBX1 binding sites. PLoS ONE *7*, e36054.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562–578.

Vitobello, A., Ferretti, E., Lampe, X., Vilain, N., Ducret, S., Ori, M., Spetz, J.-F., Selleri, L., and Rijli, F.M. (2011). Hox and Pbx factors control retinoic acid synthesis during hindbrain segmentation. Developmental Cell *20*, 469–482.

Wichterle, H., Lieberam, I., Porter, J.A., and Jessell, T.M. (2002). Directed differentiation of embryonic stem cells into motor neurons. Cell *110*, 385–397.

Wu, J., Akerman, M., Sun, S., McCombie, W.R., Krainer, A.R., and Zhang, M.Q. (2011). SpliceTrap: a method to quantify alternative splicing under single cellular conditions. Bioinformatics *27*, 3010–3016.

Xue, Y., Ouyang, K., Huang, J., Zhou, Y., Ouyang, H., Li, H., Wang, G., Wu, Q., Wei, C., Bi, Y., et al. (2013). Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. Cell *152*, 82–96.

Xue, Y., Zhou, Y., Wu, T., Zhu, T., Ji, X., Kwon, Y.-S., Zhang, C., Yeo, G., Black, D.L., Sun, H., et al. (2009). Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the

general splicing repressor to modulate exon inclusion or skipping. Mol. Cell *36*, 996–1006.

Yano, M., Hayakawa-Yano, Y., Mele, A., and Darnell, R.B. (2010). Nova2 regulates neuronal migration through an RNA switch in disabled-1 signaling. Neuron *66*, 848–858.

Yap, K., and Makeyev, E.V. (2013). Regulation of gene expression in mammalian nervous system through alternative pre-mRNA splicing coupled with RNA quality control mechanisms. Mol. Cell. Neurosci. *56*, 420–428.

Yap, K., Lim, Z.Q., Khandelia, P., Friedman, B., and Makeyev, E.V. (2012). Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. Genes Dev. *26*, 1209–1223.

Ying, Q.-L., Stavridis, M., Griffiths, D., Li, M., and Smith, A. (2003). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. Nat. Biotechnol. *21*, 183–186.

Zheng, S., and Black, D.L. (2013). Alternative pre-mRNA splicing in neurons: growing up and extending its reach. Trends Genet. *29*, 442–448.

Zheng, S., Damoiseaux, R., Chen, L., and Black, D.L. (2013). A broadly applicable high-throughput screening strategy identifies new regulators of Dlg4 (Psd-95) alternative splicing. Genome Res. *23*, 998–1007.

Zheng, S., Gray, E.E., Chawla, G., Porse, B.T., O'Dell, T.J., and Black, D.L. (2012). PSD-95 is post-transcriptionally repressed during early neural development by PTBP1 and PTBP2. Nat. Neurosci. *15*, 381–8–S1.

CHAPTER 3

Pbx1a is not required for motor neuron differentiation

INTRODUCTION

In Chapter 2, we identified a large set of neuronal exons controlled by PTBP1 in ESCs and NPCs. Among these targets, we focused on the homeodomain transcription factor Pbx1 because of its known interactions with the Hox family of developmental regulators (LaRonde-LeBlanc and Wolberger, 2003; Piper et al., 1999). We found that PTBP1 represses the alternative splicing of Pbx1 exon 7, leading to Pbx1b isoform expression in ESCs and NPCs. Using CRISPR-Cas9 genome editing, we deleted regulatory elements in Pbx1 intron 6 to induce exon 7 splicing and Pbx1a expression in the presence of PTBP1. We found that the early expression of Pbx1a induces the expression of specific neuronal genes, including several known Pbx1 targets. These results suggest that Pbx1a is more active than Pbx1b at neuronal target genes.

Here we further investigate the function of Pbx1a during motor neuron (MN) differentiation. We use CRISPR-Cas9 genome editing to delete exon 7 from the Pbx1 locus to force Pbx1b expression all cell types. We find that the loss of Pbx1a does not influence the efficiency of MN differentiation. Using RNA-seq, we identify altered gene expression patterns in the GFP+ MNs lacking exon 7, including several Pbx1 target genes.

RESULTS

Exon 7 deletions in the Pbx1 locus force Pbx1b expression

PTBP1 represses the expression of Pbx1a in ESCs and NPCs. To examine the role of Pbx1a during MN differentiation, we used CRISPR-Cas9 genome editing to eliminate exon 7 from the Pbx1 locus. Guide RNA pairs were designed to target the Cas9 endonuclease to sites upstream and downstream of exon 7 (Figure 3-1A). Guides A and C would remove 821 bp, including exon 7, while guides B and D would remove 860 bp. HB9-GFP ESCs were transiently transfected with plasmids encoding Cas9 and each pair of guide RNAs. Clones were then isolated from single transfected cells and genotyped by PCR (Figure 3-1B), with the deletion events confirmed by Sanger sequencing. In total, we isolated four clones homozygous for the gA-C deletion and two clones homozygous for the gB-D deletion. In the absence of exon 7 (E7 -/-), the edited cells are forced to express Pbx1b at the protein level (Figure 3-2B).

Forced Pbx1b expression in GFP+ MNs affects the transcription of Pbx1 target genes.

We differentiated the wild type and E7 -/- cell lines into MNs and monitored their differentiation capacity with FACS. At Day 5 of differentiation, we did not observe a difference in the percentage of GFP+ cells between the different cell lines (data not shown). To characterize how the loss of exon 7 splicing alters transcriptional programs, we performed RNA-seq to compare GFP+ MNs with different Pbx1 genotypes. The samples were subjected to 50 nt single-end RNA sequencing, with the data analyzed by the Cufflinks pipeline (Figure 3-2A).

Using a twofold change in FPKM as the cutoff identified 93 and 146 differentially expressed genes in E7 -/- MNs using guides gA-C and gB-D, respectively. Between the two comparisons, 23 events overlapped (17 induced and 6 repressed genes; Figure 3-3A). The small

overlap suggests that there is variation between the two E7 -/- lines, perhaps owing to Cas9 off-target effects (Pattanayak et al., 2013). Of the 23 shared genes, we observed expression changes in several transcriptional regulators, including the immediate early genes Fos, Egr1, and Jund (Figure 3-3B). Notably, a previous study found Pbx1 binding at the Fos enhancer in MCF7 breast cancer cells (Magnani et al., 2011). Comparing the 23 genes with a published Pbx1 ChIP data set identified Pbx1 binding at the transcriptional start sites of *Mafa*, *Ier5l*, and *H2afj* (Penkov et al., 2013). The expression of these targets was increased in both E7 -/- lines, suggesting that Pbx1a represses or Pbx1b enhances their expression in HB9+ MNs.

DISCUSSION

Pbx1 has been implicated as a transcriptional regulator during neuronal development (Maeda et al., 2002; Sgadò et al., 2012; Vitobello et al., 2011). However, the requirement of Pbx1a during neuronal differentiation is not well understood. A previous study found that full-length Pbx1 is not essential for the neuronal commitment and differentiation of RA-treated ESCs (Jürgens et al., 2009). We find that Pbx1a is not needed for MN differentiation. It is possible that Pbx1b or Pbx family members compensate for Pbx1a activity in developing MNs. Interestingly, the Pbx1a C-terminus shares homology with Pbx3, which undergoes a similar splicing event during development (data not shown) (Monica et al., 1991). In future studies, it will be interesting to examine whether Pbx1a function is needed for later stages of neuronal maturation. Likewise, it is worth investigating the role of Pbx1a in different neuronal populations. A previous study demonstrated that the loss of Pbx1a influences the axonal guidance of dopaminergic neurons in the mouse midbrain (Sgadò et al., 2012).

Using RNA-sequencing we find gene expression changes in the Pbx1 mutant MNs. However, there are differences between the two Pbx1 mutant cell lines. One concern is Cas9 off-target effects, and our findings should inform future experimental designs using CRISPR-Cas9 genome editing. An alternative approach includes using the Cas9 nickase to create mutations at splice sites with homology directed repair. Despite the differences between the two Pbx1 E7 -/- data sets, we identified several targets with Pbx1 ChIP binding. However, the function of *Mafa*, *Ier5l*, and *H2afj* during MN development is not well understood.

MATERIAL AND METHODS

Tissue culture

46C and HB9-GFP mouse ESCs were grown on 0.1% gelatin-coated dish with CF1 mouse embryonic fibroblasts (Applied StemCell, Inc., Menlo Park, CA) in ESC media. ESC media consisted of DMEM (Fisher Scientific, Hampton, NH) supplemented with 15% ESC-qualified fetal bovine serum (Life Technologies, Carlsbad, CA), 1X non-essential amino acids (Life Technologies), 1X GlutaMAX (Life Technologies), 1X ESC-qualified nucleosides (EMD Millipore, Billerica, MA), 0.1 mM 2-Mercaptoethanol (Sigma-Aldrich, St. Louis, MO), and 1000U/ml ESGRO leukemia inhibitor factor (EMD Millipore).

HB9-GFP ESCs were differentiated to MNs according to (Wichterle et al., 2002), with minor modifications. Briefly, feeder-free ESCs were grown as aggregate culture for two days in MN media. MN media consisted of 1:1 mixture of Neurobasal media (Life Technologies) and DMEM/F12 (Fisher Scientific) supplemented with 10% Knockout Serum Replacement (Life Technologies), 1X GlutaMAX (Life Technologies), and 0.1 mM 2-Mercaptoethanol (Sigma-Aldrich). After two days, the MN media was supplemented with 1X N2 Supplement (Life Technologies), 1 uM all-trans retinoic acid (RA, Sigma-Aldrich), and 1 uM smoothened agonist (EMD Millipore).

Cell sorting of GFP+ MNs

EBs were collected 5 days after RA and SAG addition. EBs were dissociated in Acutase (Innovative Cell Technologies, San Diego, CA) for 10 minutes and sorted using a BD FACSAria cell sorter at the UCLA Broad Stem Cell center core facility.

Generation of CRISPR cell lines

Guide RNAs targeting Pbx1 exon 7 were designed using http://crispr.mit.edu/ and cloned into a modified pX330-U6-Chimeric_BB-CBh-hSpCas9 plasmid (gift of B. Stahl and J. Doudna, University of California, Berkley) according to the published protocol (Ran et al., 2013). To generate Cas9-targeted deletions, HB9-GFP ESCs were transfected with two modified pX330 constructs and pBABE-puro (Addgene, Cambridge, MA; #42230) using BioT (Bioland Scientific, Paramount, CA) according to the manufacturer's recommendations. Transfected cells were then treated with 0.5 ug/ml puromycin (InVivoGen, San Diego, CA). Following puromycin selection, clonal ESC lines were isolated and genotyped for exon 7 deletions. After genotyping to identify exon7 deletions, ESC clones were subjected to secondary tests to confirm exon 7 splicing.

The guide RNAs sequences are

guide A, 5'-GGAGAGTATCTATACACCGC-3';

guide B, 5'-AAGAATATGGTCACCGATAC-3';

guide C, 5'-TGGTGACCATGCTTCGGCTC-3';

guide D, 5'-CTAACATCGAGAAGTCATCC-3'.

The genotyping primer sequences are

Pbx1-E7_F, 5'-GCAGGTGAAAGCCCATGTTA-3';

Pbx1-E7_R, 5'- GGCTTCCTGAGGTATCTGTCT-3'.

Western blots

Western blots were performed on total protein from cell cultures lysed in RIPA buffer supplemented with protease inhibitors (Roche, Basel, Switzerland) and Benzonase (Sigma-

Aldrich). Lysates were diluted in 4X SDS loading buffer, heated for 10 min at 95°C, and loaded onto 10% polyacrylamide Laemmli SDS PAGE gels. Gels were run under standard electrophoresis conditions. Transfers were performed on a Novex X-Cell mini-cell transfer apparatus (Life Technologies) onto Immobilon-FL PVDF membranes (EMD Millipore). The membranes were probed under standard conditions with primary antibodies overnight at 4°C. The following primary antibodies were used: rabbit anti-Pbx1 antibody (1:500, Cell Signaling, Danvers, MA) and rabbit anti-U170K antibody (1:1000) (Sharma et al., 2005). For fluorescent detection, the membranes were probed with ECL Plex Cy3 and Cy5-conjugated goat anti-mouse and goat anti-rabbit secondary antibodies (1:2000; GE Healthcare). The blots were then scanned on a Typhoon Phosphorimager (GE Healthcare)

RNA sequencing and data analysis

To identify gene expression changes with the loss of Pbx1a, polyA-plus RNA was isolated from Day 5 GFP+ MNs differentiated from wild type (n=2) and E7 -/- ESC clones (guides A-C, n=4; guides B-D, n=2). Sequencing libraries were constructed using the TruSeq Stranded mRNA Library Prep Kit (Illumina) and subjected to 50 nt single-end sequencing (Illumina HiSeq2000) to generate 25 to 40 million mapped reads per replicate. Differential gene expression was calculated using the Cufflinks pipeline (q-value < 0.05). Normalized FPKM values were calculated by Cuffnorm and were used to perform a Welch's t-test (p-value < 0.05). Differentially expressed genes were evaluated for Pbx1 ChIP clusters (Penkov et al., 2013). Pbx1 binding was determined if a ChIP cluster was present within a window 1 kb upstream and downstream of the gene locus.

FIGURE LEGENDS

Figure 3-1: Cas9-targeted deletions of Pbx1 exon 7. (**A**) Guide RNAs were designed to target regions upstream and downstream of Pbx1 exon 7. (**B**) Genotyping PCR of the wild type and the exon 7 deletion cell lines.

Figure 3-2: Pbx1 exon 7 deletions forces Pbx1b expression in HB9+ MNs. (**A**) Genome Browser tracks of aligned RNA-seq reads from wild type (top) and E7 -/- (bottom) GFP+ MN cultures. The asterisk (*) indicates Pbx1 exon 7, which is excluded in the mutants. (**B)** Western blot shows that Pbx1b is the only isoform expressed in the E7 -/- MN cultures.

Figure 3-3: The loss of Pbx1 exon 7 influences gene expression patterns in HB9+ MNs. (**A**) Using a twofold change in FPKM as the cutoff identified 23 differentially expressed genes that were shared between the two E7 -/- cell lines. (**B**) Bar charts of the 23 differentially expressed genes in the E7 -/- cell lines relative to the wild type cells. Genes with Pbx1 ChIP binding are indicated in bold.

Figure 3-1: Cas9-targeted deletions of Pbx1 exon 7.

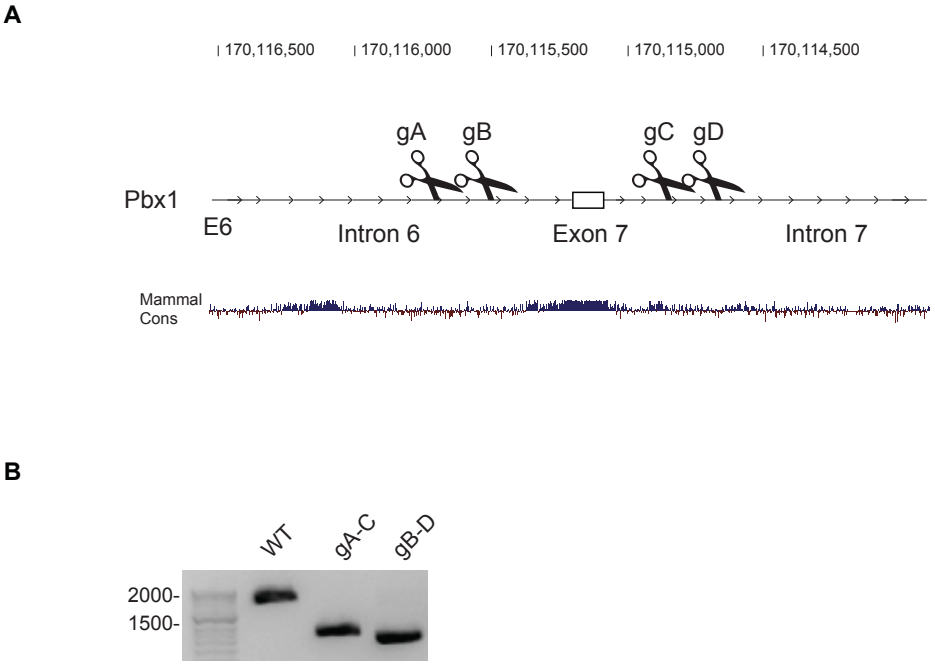


Figure 3-2: Pbx1 exon 7 deletions forces Pbx1b expression in HB9+ MNs.

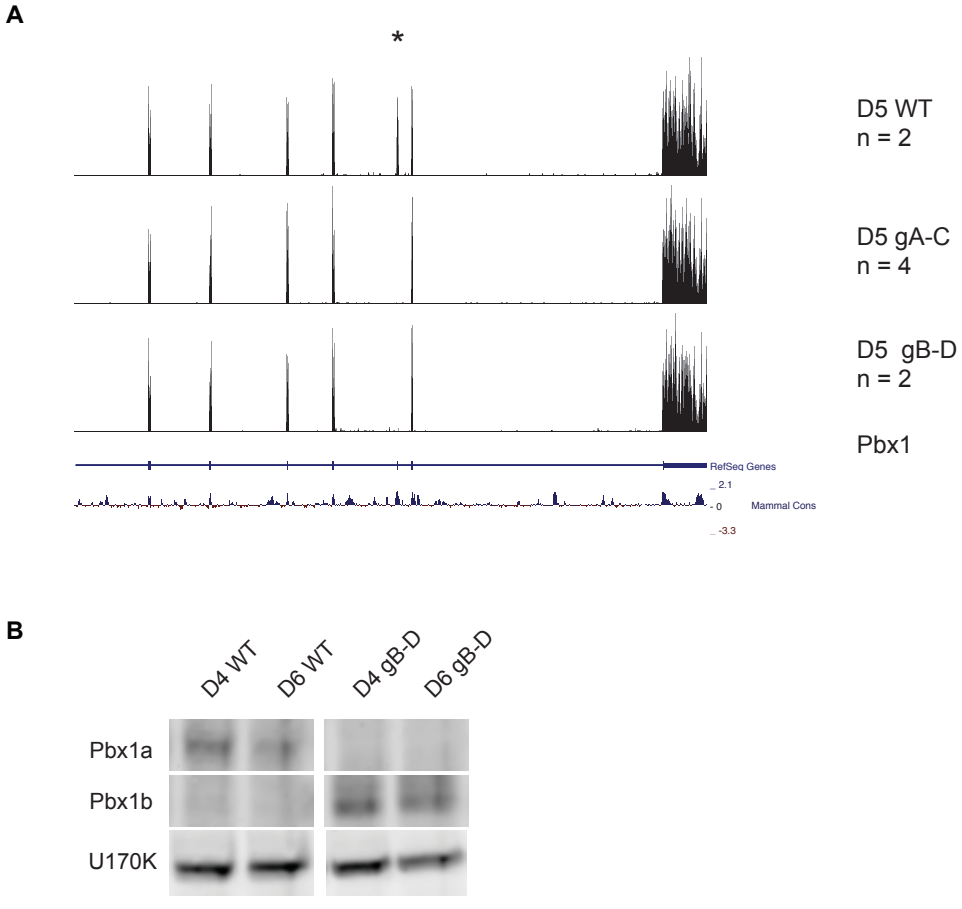
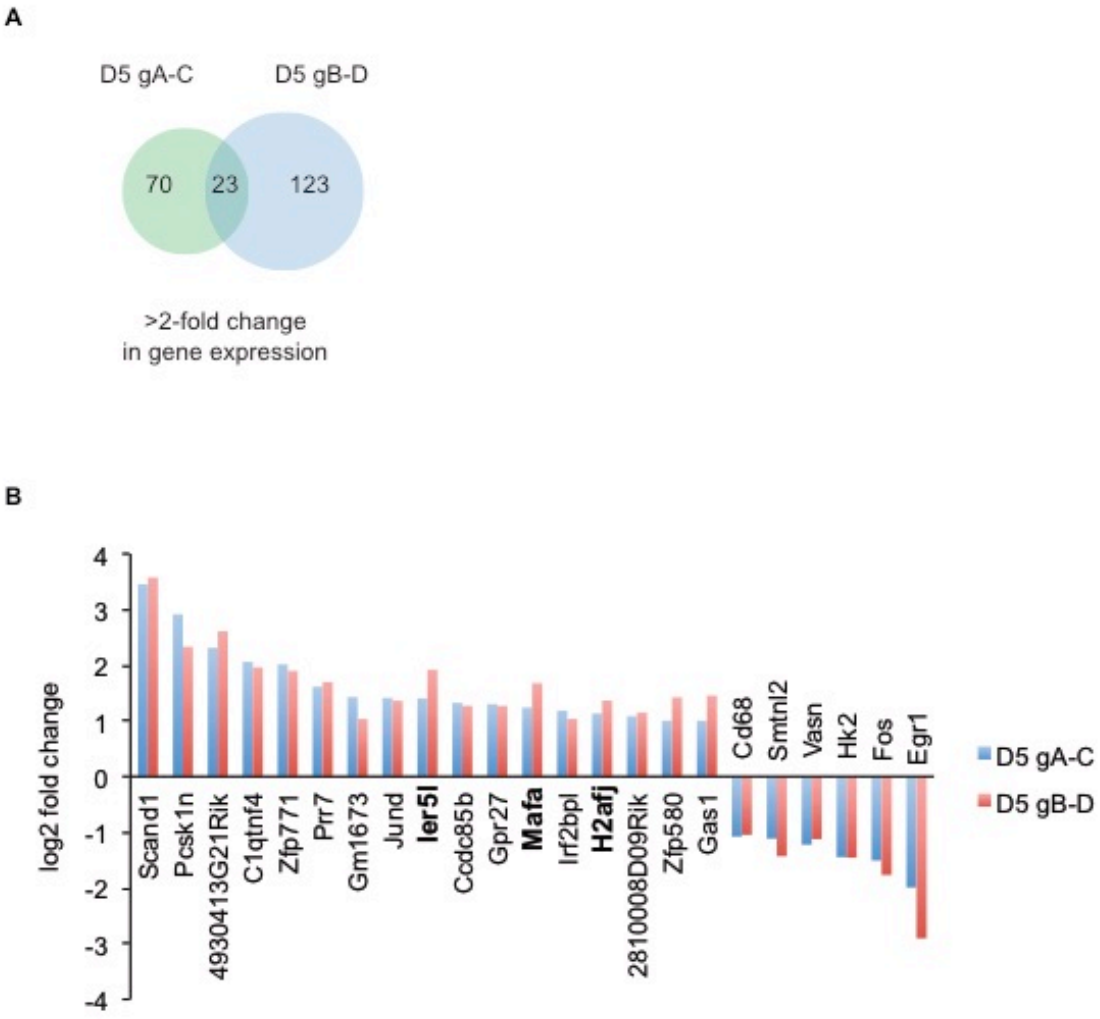


Figure 3-3: The loss of Pbx1 exon 7 influences gene expression patterns in HB9+ MNs.



REFERENCES

Jürgens, A.S., Kolanczyk, M., Moebest, D.C.C., Zemojtel, T., Lichtenauer, U., Duchniewicz, M., Gantert, M.P., Hecht, J., Hattenhorst, U., Burdach, S., et al. (2009). PBX1 is dispensable for neural commitment of RA-treated murine ES cells. In Vitro Cell. Dev. Biol. Anim. *45*, 252–263.

LaRonde-LeBlanc, N.A., and Wolberger, C. (2003). Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior. Genes Dev. *17*, 2060–2072.

Maeda, R., Ishimura, A., Mood, K., Park, E.K., Buchberg, A.M., and Daar, I.O. (2002). Xpbx1b and Xmeis1b play a collaborative role in hindbrain and neural crest gene expression in Xenopus embryos. Proc. Natl. Acad. Sci. U.S.a. *99*, 5448–5453.

Magnani, L., Ballantyne, E.B., Zhang, X., and Lupien, M. (2011). PBX1 genomic pioneer function drives ERα signaling underlying progression in breast cancer. PLoS Genet. *7*, e1002368.

Monica, K., Galili, N., Nourse, J., Saltman, D., and Cleary, M.L. (1991). PBX2 and PBX3, new homeobox genes with extensive homology to the human proto-oncogene PBX1. Mol. Cell. Biol. *11*, 6149–6157.

Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A., and Liu, D.R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat. Biotechnol. *31*, 839–843.

Penkov, D., Mateos San Martín, D., Fernandez-Díaz, L.C., Rosselló, C.A., Torroja, C., Sánchez-Cabo, F., Warnatz, H.J., Sultan, M., Yaspo, M.L., Gabrieli, A., et al. (2013). Analysis of the DNA-binding profile and function of TALE homeoproteins reveals their specialization and specific interactions with Hox genes/proteins. Cell Rep *3*, 1321–1333.

Piper, D.E., Batchelor, A.H., Chang, C.P., Cleary, M.L., and Wolberger, C. (1999). Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. Cell *96*, 587–597.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc *8*, 2281–2308.

Sgadò, P., Ferretti, E., Grbec, D., Bozzi, Y., and Simon, H.H. (2012). The atypical homeoprotein Pbx1a participates in the axonal pathfinding of mesencephalic dopaminergic neurons. Neural Dev *7*, 24.

Sharma, S., Falick, A.M., and Black, D.L. (2005). Polypyrimidine tract binding protein blocks the 5' splice site-dependent assembly of U2AF and the prespliceosomal E complex. Mol. Cell *19*, 485–496.

Vitobello, A., Ferretti, E., Lampe, X., Vilain, N., Ducret, S., Ori, M., Spetz, J.-F., Selleri, L., and Rijli, F.M. (2011). Hox and Pbx factors control retinoic acid synthesis during hindbraib segmentation. Developmental Cell *20*, 469–482.

Wichterle, H., Lieberam, I., Porter, J.A., and Jessell, T.M. (2002). Directed differentiation of embryonic stem cells into motor neurons. Cell *110*, 385–397.

CHAPTER 4

Concluding remarks

The splicing regulators PTBP1 and PTBP2 control large splicing programs during neuronal development. However, the early program regulated by PTBP1 has been difficult to characterize in the mouse brain, and most of the work characterizing the PTBP1 splicing program has been performed in non-neuronal cell lines. Therefore, the goals of this dissertation were to 1) define the PTBP1-specific regulatory network and 2) characterize the consequences of the PTBP1 splicing program during early neuronal lineage commitment and differentiation.

We find that ESCs are a suitable system to characterize the PTBP1-specific splicing program. As ESCs differentiate into NPCs and neurons, thousands of alternative splicing events change, which corresponds to the transition from PTBP1 to PTBP2 expression. Of these neuronal splicing events, we identify a large set or exons that is directly regulated by PTBP1 and not PTBP2. Interestingly, we find that alternative exons show different sensitivities to PTBP1 and PTBP2 in ESCs and NPCs. In ESCs, exons sensitive to PTBP1 depletion are largely not affected by PTBP2 induction, whereas in NPCs, many exons that are sensitive to PTBP1 depletion show larger changes in splicing with PTBP1/2 double knockdown. Similar to previous studies, we identified a subset of these events in NPCs that are co-regulated by the PTB proteins, including Dlg4. Future studies can use NPCs as a system to study why certain exons show different sensitivities to PTBP1 and PTBP2. Overall, these results demonstrate the PTBP1 splicing program changes with differentiation and likely depends on the expression of other splicing regulators.

Our work provides the most comprehensive picture of PTBP1 splicing regulation during early neuronal differentiation. We find that PTBP1 targets influence many aspects of neuronal differentiation and maturation, including cytoskeletal assembly, synaptic function, and transcriptional regulation. We were particularly interested in PTBP1 regulation of transcriptional regulators, where changes in the structure and function of these proteins have the potential to affect gene expression programs important for neurogenesis. We focused on the PTBP1 target Pbx1 because of its known interactions with the Hox family of proteins. In ESCs and NPCs, PTBP1 represses exon 7 in the Pbx1 transcript to promote Pbx1b expression. As NPCs differentiate into neurons, PTBP1 expression is reduced and Pbx1a expression is induced.

A major advance in our study is the application of the CRISPR-Cas9 technology to study the biology and mechanisms of splicing regulation during neuronal differentiation. We show that deleting regulatory elements in Pbx1 intron 6 can promote Pbx1a expression, and the early expression of Pbx1a induces the expression of specific neuronal genes, including several Pbx1 targets. Thus one component of the PTBP1 splicing program is to regulate Pbx1 activity and transcriptional programs relevant to neurogenesis. In addition, our work shows that CRISPR-Cas9 genome editing can be used to map splicing regulatory elements at the endogenous locus, a significant advance over mini-gene studies.

By combining ESC culture and genome editing, we have developed an approach to study isoform activity in specific neuronal populations. These methods will help characterize the cellular functions of the other PTBP1-specific targets identified in this study. Likewise, these approaches can be applied to study the late PTBP splicing program. Recent work has shown that late PTBP program represses adult-specific isoforms during critical periods of maturation, and

cortical neurons lacking PTBP2 fail to mature and die in culture. It will be interesting to investigate whether the loss of PTBP2 affects the development and maturation of HB9+ MNs.