

Title: An Alternative Splicing Event Amplifies Evolutionary Differences Between Vertebrates

Authors: Serge Gueroussov^{1,2}, Thomas Gonatopoulos-Pournatzis¹, Manuel Irimia^{1,3}, Bushra Raj^{1,2}, Zhen-Yuan Lin⁴, Anne-Claude Gingras^{2,4}, Benjamin J. Blencowe^{1,2*}

Affiliations:

¹Donnelly Centre, University of Toronto, Toronto, ON, M5S 3E1, Canada.

²Department of Molecular Genetics, University of Toronto, Toronto, ON, M5S 1A8, Canada.

³EMBL/CRG Research Unit in Systems Biology, Centre for Genomic Regulation (CRG), Barcelona, Spain.

⁴Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto M5G 1X5, Canada.

*Correspondence to:

Benjamin J Blencowe, PhD

Donnelly Centre, University of Toronto

Toronto, ON, Canada

Office 416-978-3016

Lab 416-978-7150

Fax 416-946-5545

Email: b.blencowe@utoronto.ca

One Sentence Summary:

A mammalian-specific alternative splicing event in an RNA binding protein results in major differences between vertebrate transcriptomes.

Abstract:

Alternative splicing (AS) generates extensive transcriptomic and proteomic complexity. However, the functions of species- and lineage-specific splice variants are largely unknown. Here, we show that mammalian-specific skipping of exon 9 of PTBP1 alters its splicing regulatory activities and affects the inclusion levels of numerous exons. During neurogenesis, skipping of exon 9 reduces PTBP1 repressive activity so as to facilitate activation of a brain-specific AS program. Engineered skipping of the orthologous exon in chicken cells induces a large number of mammalian-like AS changes in PTBP1 target exons. These results thus reveal that a single exon skipping event in an RNA binding regulator directs numerous AS changes between species. The results further suggest that these changes contributed to evolutionary differences in the formation of vertebrate nervous systems.

Main Text:

A major challenge in evolutionary biology is to determine which gene regulatory changes contributed to species-specific phenotypes (1-3). Comparative transcriptomic analyses revealed that vertebrate organ alternative splicing (AS) patterns diverged more rapidly than gene expression differences (4-6). These AS differences were largely attributed to changes in the utilization of conserved cis-regulatory elements (4, 5). However, a small number of lineage- and species-dependent AS changes were detected in nucleic acid binding proteins, suggesting that splicing differences in trans-acting regulators also contributed to the extensive diversity found among vertebrates (4). One such change involves mammalian-specific skipping of an exon (exon 9 in human, exon 8 in mouse) in the polypyrimidine tract binding protein 1 (PTBP1/HNRNPI) (4). PTBP1 controls cell and tissue-dependent AS, as well as other steps in gene expression (7). Exon 9 encodes an intrinsically disordered and conserved linker between RNA Recognition Motif 2 (RRM2) and RRM3, two of four RRMs in PTBP1 (Fig. 1A; Fig. S1A).

To investigate the role of mammalian-specific exon 9 skipping, RNA-Seq was used to profile AS in 293 cells engineered to express 3xFLAG-tagged PTBP1 transgenes with exon 9 included (PTBP1+Ex9) or deleted (PTBP1ΔEx9) (Fig. 1B; Fig S1B). In these cells, endogenous PTBP1 and PTBP2 (a paralog with partially redundant activity (8)) were depleted using siRNAs. The percentage of transcripts with a sequence spliced in (Percent Spliced In, PSI) was estimated for cassette exons (including 3-27 nucleotide microexons (9)), alternative 5'/3' splice sites, and more complex AS events. The percentage of transcripts with an intron retained (PIR) was estimated for intron retention (IR) events (10). AS events displaying a Δ PSI or Δ PIR of $\geq 15\%$ upon knockdown of PTBP1 and PTBP2, and a 'rescue' change $\geq 50\%$ towards levels observed in the control siRNA treatment were analyzed further (Methods).

Approximately 1,500 AS events showed differential regulation upon knockdown and rescue of PTBP1 (Fig. S1C). Comparable proportions of most classes of AS events displayed increased and decreased PSI changes. In contrast, 67% and 92% of retained introns and microexons, respectively, showed increased inclusion levels upon PTBP1 knockdown ($p < 1e-8$; Hypergeometric test), consistent with recent findings (11). Comparing the distributions of cassette exon Δ PSI values following induction of each PTBP1 isoform reveals that exon 9 inclusion results in significantly greater restoration of PSI values for both repressed (Fig. 1C, top panel; $p < 1e-12$, two-sided Mann-Whitney U-test) and stimulated exons (Fig. 1C, bottom panel; $p < 2.2e-16$, two-sided Mann-Whitney U-test). RT-PCR assays validated 23/23 (100%) of the analyzed RNA-Seq predictions (Figs. 1D-E, Fig. S2A-C). Moreover, titration experiments confirmed that PTBP1+Ex9 and PTBP1ΔEx9 impart their distinct regulatory activities due to the presence and absence of exon 9, rather than differences in relative expression (Fig. S3).

Recombinant PTBP1+Ex9 and PTBP1ΔEx9 proteins displayed similar binding affinities for a variety of RNA substrates in gel mobility shift assays (Fig. S4). Moreover, *in vivo* crosslinking followed by immunoprecipitation (CLIP-Seq) analysis revealed that these isoforms have near identical binding profiles surrounding PTBP1-regulated exons *in vivo* (Fig. S5). Sequences overlapping RRM2 and exon 9 in PTBP1 define a minimal splicing repressor domain (12). Confirming this, deletion of this region markedly reduces PTBP1 splicing regulatory activity (Fig. S6A-C). However, the presence of exon 9 without RRM2 partially restores activity

for a subset of analyzed target exons (Fig. S6D). Collectively, these results indicate that exon 9 possesses splicing regulatory activity that is partially separable from the repressive activity conferred by RRM2, and that skipping of exon 9 reduces the negative and positive regulatory activities of PTBP1 without significantly impacting RNA binding activity.

To investigate whether exon 9 inclusion influences splicing of PTBP1 target exons in other cell types, we asked if its PSI correlates significantly with PSI values of isoform-dependent and -independent target exons (Fig. S7A) across 64 diverse human cell and tissue types (Fig. S7B). As a control, we examined correlations between exon 9 PSI and PSI values of alternative exons that are not regulated by PTBP1. We observe a significantly higher correlation between exon 9 PSI and PSI of isoform-dependent target exons, than with isoform-independent target exons, or exons not regulated by PTBP1 (Fig. 2A, left panel; $p < 1e-7$, two-sided Mann-Whitney U-test). These differences in correlation are not a consequence of isoform-dependent exons having a greater sensitivity to overall PTBP1 levels, since both isoform-dependent and -independent exons display similar correlation with total PTBP1 mRNA levels (Fig. 2A, right panel $p = 0.40$, two-sided Mann-Whitney U-test). Moreover, exon 9 PSI values do not correlate significantly with PTBP1 mRNA levels (Fig. S7C; $p = 0.66$, Pearson's product-moment correlation). Exons whose inclusion levels significantly correlate with exon 9 PSI ($p < 0.05$, Pearson's product-moment correlation) are enriched in genes functionally associated with the cytoskeleton (e.g. FLNB and MYO18A) and nervous system development (e.g. APP and APBB2) (Fig. 2B, Fig. S7D, Table S1). Thus, mammalian-specific skipping of exon 9 modulates the splicing levels of a large number of functionally coherent PTBP1 target exons across diverse cell and tissue types.

PTBP1 represses a network of neural-specific alternative exons in non-neural cells and tissues (8, 13, 14). Activation of this network is required for neuronal differentiation and depends on PTBP1 silencing by miR-124 (8, 13, 14). Many PTBP1-repressed exons are activated by the neuronal-specific SR-related protein of 100 kDa (nSR100/SRRM4) (15, 16). Given that skipping of PTBP1 exon 9 reduces its repressive activity, we considered that this change might have evolved to further silence PTBP1 and render target exons permissive to activation by nSR100. To investigate this, we employed an *in vitro* model in which mouse embryonic stem (mES) cells are differentiated into cortical glutamatergic neurons (Fig. S8A-B). During differentiation, we observed the expected reduction in Ptbp1 and increase in nSR100 expression (16) (Fig. S8C).

Concomitant with reduction of Ptbp1 mRNA levels, skipping of Ptbp1 exon 8 (orthologous to human exon 9) progressively increases during neuronal differentiation of two independent mES lines (CGR8 and R1; Fig. 3A-B; Fig S8D). Transition from skipping to inclusion of exons that are negatively regulated by Ptbp1 and positively regulated by nSR100 begins between neuroepithelial stem cells and radial glia (16). This stage is associated with appreciable Ptbp1 mRNA expression, but also corresponds to when Ptbp1 Δ Ex8 supplants Ptbp1+Ex8 (Fig. 3A-B; Fig S8D). These results suggest that preferential expression of Ptbp1 Δ Ex8 further de-represses the Ptbp1 and nSR100 co-regulated neural AS network. Consistent with this proposal, Ptbp1 isoform-dependent exons, relative to Ptbp1 isoform-independent exons, display significantly higher inclusion early during differentiation (Fig. 3C, Fig. S8E; $p = 0.010$, two-sided Mann-Whitney U-test), but significantly lower inclusion at later stages (i.e. between days 1 and 7 (T4); Fig. S8D; $p = 8.3e-5$, two-sided Mann-Whitney U-test).

To confirm whether Ptbp1 Δ Ex8 is a weaker repressor than Ptbp1+Ex8, we used *in vitro* splicing assays to compare the ability of each isoform to compete with nSR100. Splicing reporters containing neural-specific exons from the *PTBP2* and *Daam1* genes (15) were incubated with comparable amounts of recombinant nSR100, PTBP1+Ex9 and PTBP1 Δ Ex9 proteins (Fig. 3D; Fig. S8F-G). PTBP1+Ex9 has a greater ability to outcompete nSR100 and promote skipping of each alternative exon than PTBP1 Δ Ex9.

To confirm whether Ptbp1 exon 8 skipping affects the kinetics of activation of neural exons *in vivo*, the CRISPR/Cas9 system was used to create mES cell lines that constitutively and uniquely express Ptbp1+Ex8 or Ptbp1 Δ Ex8 (Fig. 3E; Fig. S9A). Ptbp1 target exon levels were then monitored during differentiation of the lines into glutamatergic neurons. Expression of Ptbp1 Δ Ex8 resulted in earlier activation of target exons relative to that observed during differentiation of wild-type (WT) cells (Fig. 3F; Fig. S9B). In contrast, constitutive expression of Ptbp1+Ex8 significantly delayed activation of the same targets (Fig. 3F; Fig. S9B). Skipping of Ptbp1 exon 8 is therefore important for the activation of target exons during neuronal differentiation. Moreover, Ptbp1 Δ Ex8 expression led to earlier activation of key neural markers, whereas Ptbp1+Ex8 expression delayed their activation (Fig. 3G; Fig. S10). These results demonstrate that a change in inclusion of a single alternative exon significantly impacts the kinetics of regulatory transitions during mammalian neuronal differentiation.

Having established that skipping of exon 9/exon 8 reduces the regulatory activity of PTBP1/Ptbp1 in mammalian cells, we asked whether introducing this mammalian-specific AS event in a non-mammalian context is sufficient to alter regulation of PTBP1 target exons. One or both copies of orthologous exon 8 were deleted in the chicken DT40 B cell line to generate cells that co-express PTBP1+Ex8 and PTBP1 Δ Ex8, or that exclusively express PTBP1 Δ Ex8 (Fig. 4A). RNA-Seq analysis of the mutant and WT cell lines was used to assess the impact of exon deletion on the chicken transcriptome. Homozygous deletion of exon 8 resulted in splicing changes ($\text{PSI} \geq 15\%$) for 58 cassette exons, whereas heterozygous deletion resulted in intermediate changes in the same direction (Fig. 4B). Similar to the results in human cells, affected genes are enriched in GO terms associated with the cytoskeleton (Table S2). This set included exons in the TPM1, MAGI3, and OSBPL9 genes, whose human orthologs displayed isoform-dependent regulation in 293 cells. RT-PCR assays validated 24 of 31 (77%) of analyzed events (Fig. 4C; Fig. S11A), and further detected changes (e.g. in FLNB and PTBP2; Fig. S11B) that are isoform-dependent in 293 cells (Fig. 1E; Fig. S2A), but had smaller predicted changes ($\text{PSI} < 15\%$) in DT40 cells, possibly due to cell-type specific contextual differences. In contrast, we did not detect significant changes in 12/12 exons not predicted to be affected by PTBP1 exon 8 deletion (Fig. S12). Affected exons likely comprise direct PTBP1 targets, since their flanking intronic sequences are significantly enriched in PTBP1 binding motifs (17), and the position of the enriched motifs is consistent with the expected direction of PTBP1-dependent regulation based on *in vivo* PTBP1 binding map data (Fig. S5) (18, 19) (Fig. 4D). Thus, expression of the PTBP1 Δ Ex8 isoform in chicken cells results in a weaker splicing regulator that leads to increased inclusion of PTBP1-repressed exons, and reduced inclusion of PTBP1-stimulated exons. These results support the conclusion that an evolutionary change in PTBP1 exon 8/9 inclusion is sufficient to alter the regulation of many exons, including those with important developmental functions.

Rapid evolutionary change in AS patterns has been largely driven by the gain and loss of cis-regulatory elements (4, 5). It has been proposed that this change is in part a consequence of relaxed selection pressure acting on newly evolving splice isoforms, while the simultaneous expression of ancestral splice isoforms maintain core gene functions (20). The results of this study however provide evidence that rapid change in AS can also be driven by splicing changes in trans-acting regulators. The mammalian-specific skipping of PTBP1 exon 9/Ptbp1 exon 8 characterized here simultaneously impacts the levels of many additional AS events. Given that the affected target exons are concentrated in genes associated with cytoskeletal and neurobiological functions, and that their regulatory kinetics are significantly impacted by exon 8 inclusion levels during neuronal differentiation, it is interesting to consider that this event evolved to modulate the timing of transitions in the production of neural progenitors and mature neurons so as to affect brain morphology and complexity. In conclusion, a single AS event has served to ‘amplify’ the rate of evolutionary change in developmentally-associated AS patterns.

References

1. S. B. Carroll, Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25-36 (2008).
2. C. P. Ponting, The functional repertoires of metazoan genomes. *Nature reviews. Genetics* **9**, 689-698 (2008).
3. A. Necsulea, H. Kaessmann, Evolutionary dynamics of coding and non-coding transcriptomes. *Nature reviews. Genetics* **15**, 734-748 (2014).
4. N. L. Barbosa-Morais, M. Irimia, Q. Pan, H. Y. Xiong, S. Gueroussov, et al., The evolutionary landscape of alternative splicing in vertebrate species. *Science* **338**, 1587-1593 (2012).
5. J. Merkin, C. Russell, P. Chen, C. B. Burge, Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. *Science* **338**, 1593-1599 (2012).
6. A. Reyes, S. Anders, R. J. Weatheritt, T. J. Gibson, L. M. Steinmetz, W. Huber, Drift and conservation of differential exon usage across tissues in primate species. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 15377-15382 (2013).
7. P. Kafasla, I. Mickleburgh, M. Llorian, M. Coelho, C. Gooding, et al., Defining the roles and interactions of PTB. *Biochemical Society transactions* **40**, 815-820 (2012).
8. P. L. Boutz, P. Stoilov, Q. Li, C. H. Lin, G. Chawla, K. Ostrow, L. Shiue, M. Ares, Jr., D. L. Black, A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes Dev* **21**, 1636-1652 (2007).
9. M. Irimia, R. J. Weatheritt, J. D. Ellis, N. N. Parikshak, T. Gonatopoulos-Pournatzis, et al., A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell* **159**, 1511-1523 (2014).
10. U. Braunschweig, N. L. Barbosa-Morais, Q. Pan, E. N. Nachman, B. Alipanahi, et al., Widespread intron retention in mammals functionally tunes transcriptomes. *Genome research* **24**, 1774-1786 (2014).

11. Y. I. Li, L. Sanchez-Pulido, W. Haerty, C. P. Ponting, RBFOX and PTBP1 proteins regulate the alternative splicing of micro-exons in human brain transcripts. *Genome research* **25**, 1-13 (2015).
12. F. Robinson, C. W. Smith, A splicing repressor domain in polypyrimidine tract-binding protein. *The Journal of biological chemistry* **281**, 800-806 (2006).
13. E. V. Makeyev, J. Zhang, M. A. Carrasco, T. Maniatis, The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol Cell* **27**, 435-448 (2007).
14. Y. Xue, K. Ouyang, J. Huang, Y. Zhou, H. Ouyang, et al., Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. *Cell* **152**, 82-96 (2013).
15. J. A. Calarco, S. Superina, D. O'Hanlon, M. Gabut, B. Raj, et al., Regulation of vertebrate nervous system alternative splicing and development by an SR-related protein. *Cell* **138**, 898-910 (2009).
16. B. Raj, M. Irimia, U. Braunschweig, T. Sterne-Weiler, D. O'Hanlon, et al., A global regulatory mechanism for activating an exon network required for neurogenesis. *Mol Cell* **56**, 90-103 (2014).
17. D. Ray, H. Kazan, K. B. Cook, M. T. Weirauch, H. S. Najafabadi, et al., A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499**, 172-177 (2013).
18. M. Llorian, S. Schwartz, T. A. Clark, D. Hollander, L. Y. Tan, et al., Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. *Nature structural & molecular biology* **17**, 1114-1123 (2010).
19. A. Han, P. Stoilov, A. J. Linares, Y. Zhou, X. D. Fu, D. L. Black, De novo prediction of PTBP1 binding and splicing targets reveals unexpected features of its RNA recognition and function. *PLoS computational biology* **10**, e1003442 (2014).
20. Y. Xing, C. Lee, Alternative splicing and RNA selection pressure--evolutionary consequences for eukaryotic genomes. *Nature reviews. Genetics* **7**, 499-509 (2006).
21. J. P. Lambert, G. Ivosev, A. L. Couzens, B. Larsen, M. Taipale, et al., Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. *Nature methods* **10**, 1239-1245 (2013).
22. S. Falcon, R. Gentleman, Using GOstats to test gene lists for GO term association. *Bioinformatics* **23**, 257-258 (2007).
23. K. S. Hubbard, I. M. Gut, M. E. Lyman, P. M. McNutt, Longitudinal RNA sequencing of the deep transcriptome during neurogenesis of cortical glutamatergic neurons from murine ESCs. *Fl1000Research* **2**, 35 (2013).
24. F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, F. Zhang, Genome engineering using the CRISPR-Cas9 system. *Nature protocols* **8**, 2281-2308 (2013).
25. K. P. Orłowska, K. Klosowska, R. J. Szczesny, D. Cysewski, P. S. Krawczyk, A. Dziembowski, A new strategy for gene targeting and functional proteomics using the DT40 cell line. *Nucleic acids research* **41**, e167 (2013).

Supplementary Materials

www.sciencemag.org

Materials and Methods

Figs. S1 – S12

Tables S1 – S2
References (21-25)

Acknowledgments:

The authors thank Dr. Andrzej Dziembowski for providing DT40-Cre1 cells, Ulrich Braunschweig for assistance with CLIP-Seq analyses, and Nuno Barbosa-Morais and Tim Sterne-Weiler for advice on statistical testing. SG is supported by an NSERC Alexander Graham Bell Studentship, TGP is supported by fellowships from EMBO and OSCI, and MI was supported by a LTF from the Human Frontiers Science Program Organization. This research was supported by grants from the CIHR to BJB and A-C.G. BJB holds the Banbury Chair of Medical Research at the University of Toronto. Data presented in this manuscript is archived in the gene Gene Expression Omnibus (GEO) under accession number GSE69656.

Figure 1. A mammalian-specific alternative splicing event alters the global splicing regulatory activity of PTBP1

(a) Diagram showing position of human PTBP1 exon 9. (b) Strategy for assessing the role of exon 9 in splicing regulation. Endogenous proteins were depleted with siRNAs. Knockdown was rescued with stably integrated, siRNA-resistant PTBP1+Ex9 or PTBP1ΔEx9. (c) Scatterplots showing changes in Percent Spliced In (PSI) relative to control for all cassette exons upon expression of PTBP1+Ex9 (x-axis) or PTBP1ΔEx9 (y-axis). (d-e) Representative RT-PCR validations of AS events detected by RNA-Seq whose inclusion is: (d) preferentially repressed by PTBP1+Ex9, (e) preferentially enhanced by PTBP1+Ex9. Bar plots show mean PSIs from three independent RT-PCR experiments. Error bars correspond to standard error. (*) indicates $p < 0.05$, two-tailed t-test.

Figure 2. PTBP1 exon 9 inclusion correlates with target exon inclusion across diverse human cells and tissues

(a) Analysis of Pearson correlation coefficients (PCCs) derived from RNA-Seq comparisons of PTBP1 exon 9 PSI and PSIs of PTBP1 isoform-dependent target exons identified in Figure 1. PCCs were computed from pairwise comparisons across 64 diverse human cell and tissue types (see Supplementary Information). Plots show distributions of the absolute values of PCCs for different groups of target exons (dep: PTBP1-regulated and isoform-dependent; indep: PTBP1-regulated and isoform-independent; rand: not regulated by PTBP1). P values (two-sided Mann-Whitney U-test) from comparing the indicated distributions are shown. (b) Scatterplots showing isoform-dependent PTBP1 targets whose PSI correlates significantly (FDR multiple testing correction; $p < 0.05$, Pearson's product-moment correlation) with PTBP1 exon 9 PSI across human cells and tissues.

Figure 3. Regulation and function of Ptpb1 isoforms during neuronal differentiation

(a) RT-PCR and (b) western blot analysis of exon 8 inclusion during differentiation. (c) Plot of median Ptpb1 target exon PSIs during differentiation. Lighter shading indicates first and third quartiles. Red, mouse orthologs of human isoform-dependent events (dep); blue, isoform-independent events (indep). (d) In vitro splicing of PTBP2 exon 10 with varying combinations of purified nSR100, PTBP1+Ex9, and PTBP1ΔEx9. (e) Validation of CRISPR/Cas9-mediated deletion and knock-in of exon 8 at the level of genomic DNA, RNA, and protein by PCR, RT-PCR, and western blot, respectively. (f) Alternative splicing of Ptpb1-regulated exons was monitored by RT-PCR during differentiation in wild-type cells, cells with exon 8 deleted, or exon 8 constitutively included. Quantification represents three independent differentiation experiments. Error bars show standard error. (*) indicates $p < 0.05$, two-tailed t-test. (g) Expression of NeuN during differentiation was monitored by qRT-PCR in the same lines. Quantification as in (g).

Figure 4. Consequences of expressing a mammalian-like PTBP1 isoform in chicken cells

(a) Validation of heterozygous and homozygous deletions of PTBP1 exon 8 at the DNA, RNA, and protein levels, by PCR, RT-PCR, and western blot, respectively. (b) Heat map of average ΔPSI of exons affected by PTBP1 exon 8 deletion. Analyses were performed using two independent cell lines for each genotype. (c) Representative RT-PCR validations of AS events detected by RNA-Seq. Bar plots show mean PSIs from six independent assays (three for each independent cell line). Error bars show standard error. (*) indicates $p < 0.05$, two-tailed t-test. (d)

Number of PTBP1 motifs per 300 nucleotides of intronic sequence upstream (left panel) or downstream (right panel) of queried exons. (*) indicates $p < 0.05$, Exact binomial test.

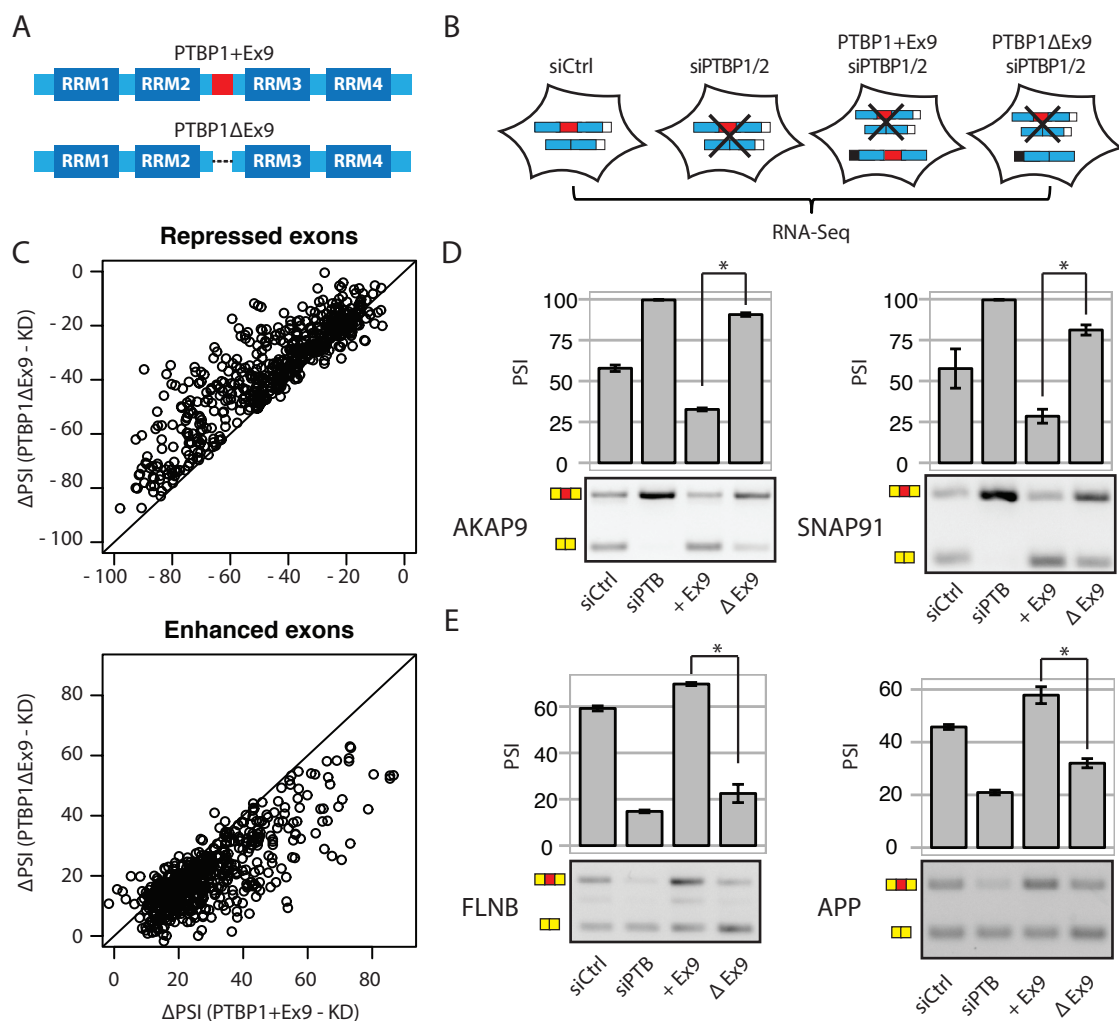


Figure 1

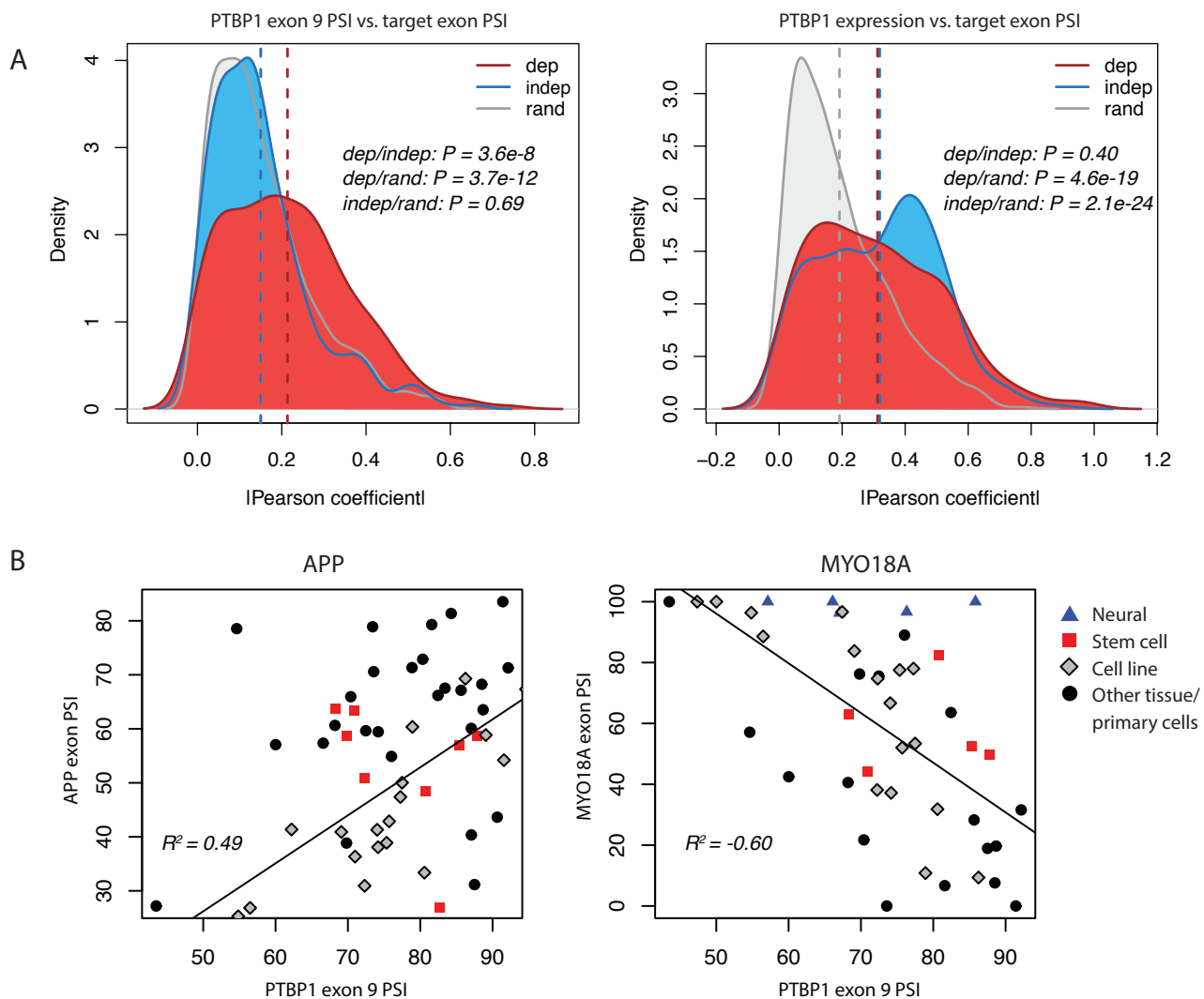


Figure 2

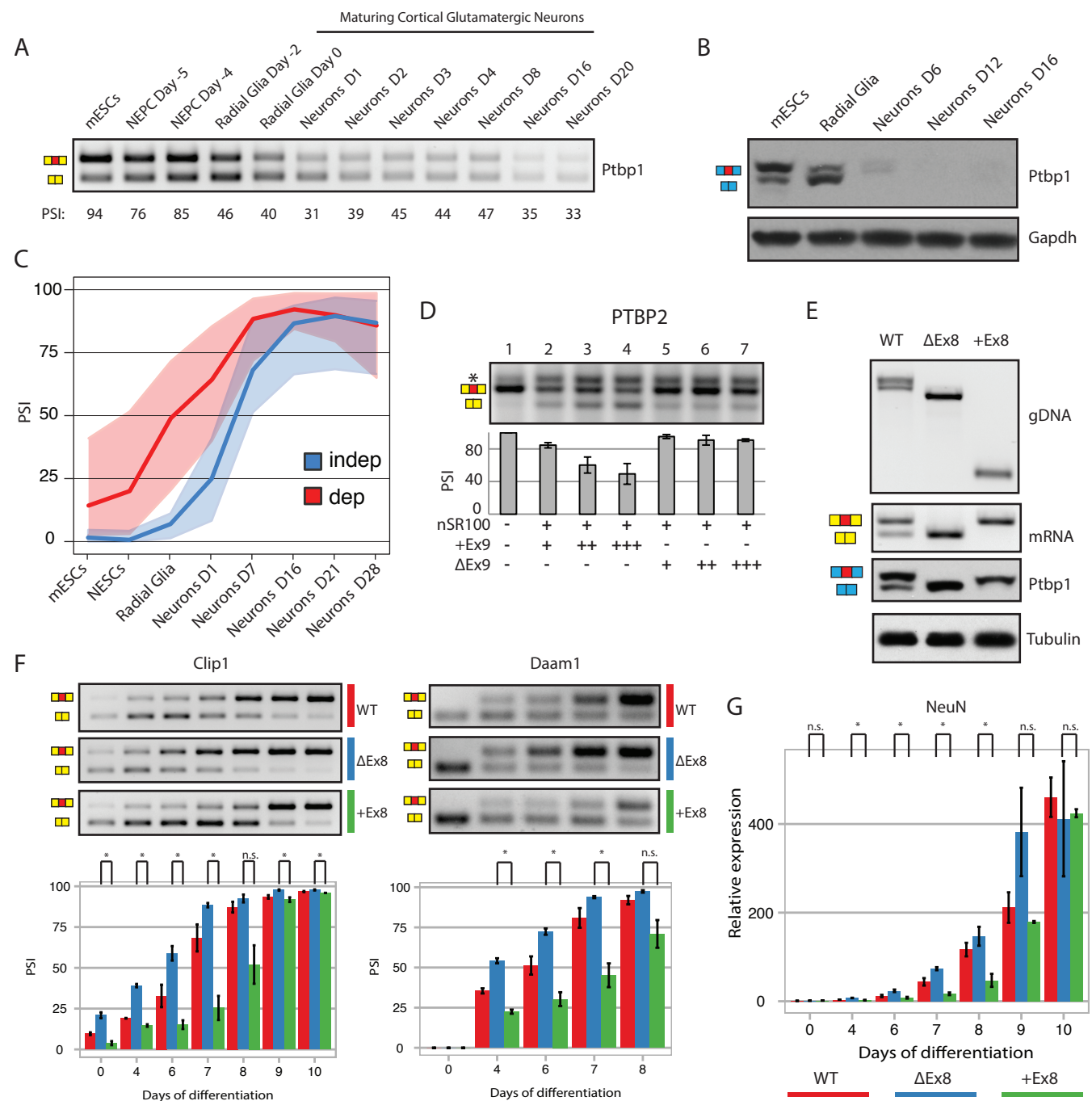


Figure 3

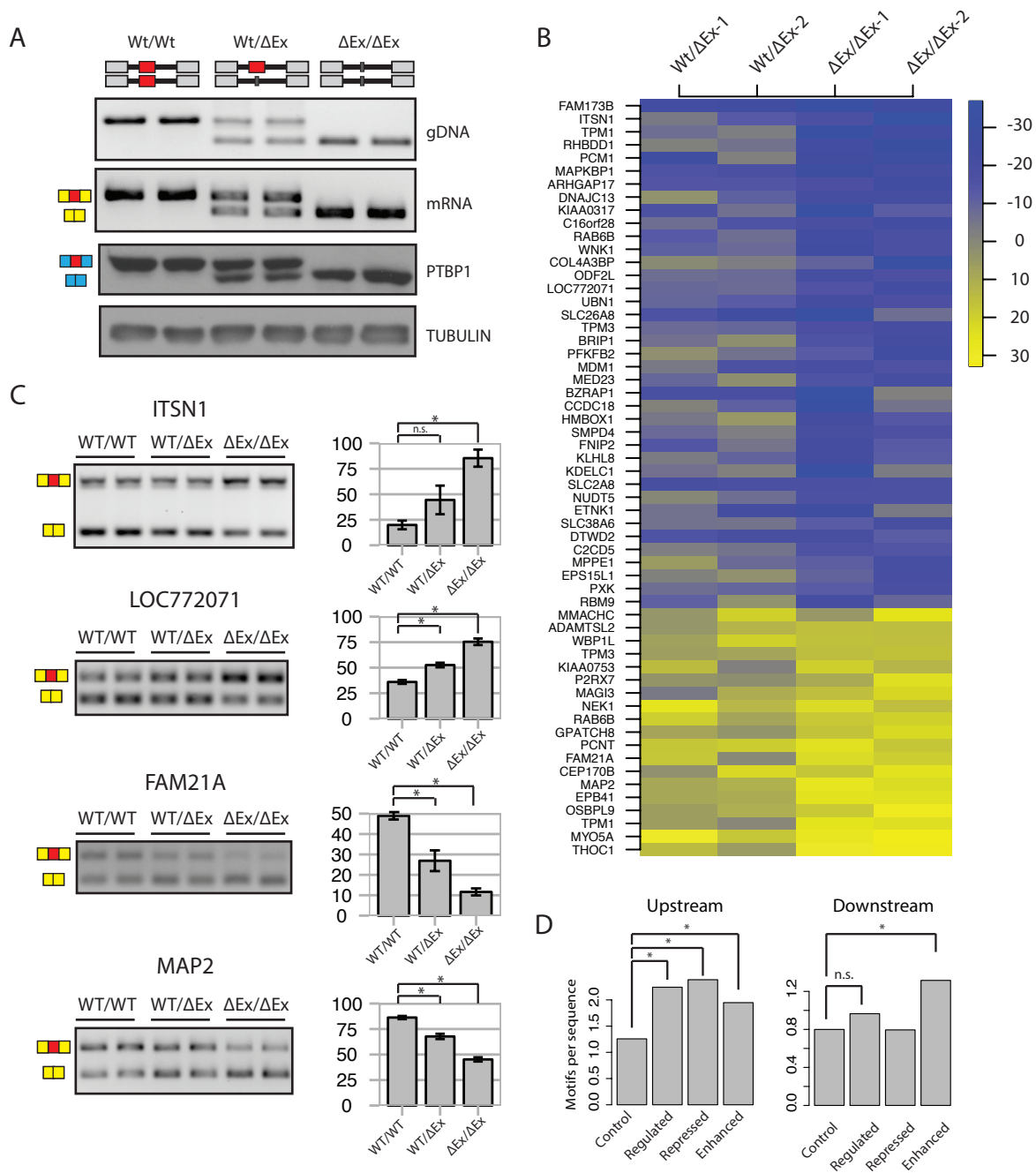


Figure 4