

RESEARCH

Open Access



Alternative splicing coupled to nonsense-mediated decay coordinates downregulation of non-neuronal genes in developing mouse neurons

Anna Zhuravskaya¹ , Karen Yap¹ , Fursham Hamid^{1*} and Eugene V. Makeyev^{1*}

*Correspondence:
fursham.hamid@kcl.ac.uk;
eugene.makeyev@kcl.ac.uk

¹ Centre for Developmental
Neurobiology, King's College
London, London SE1 1UL, UK

Abstract

Background: The functional coupling between alternative pre-mRNA splicing (AS) and the mRNA quality control mechanism called nonsense-mediated decay (NMD) can modulate transcript abundance. Previous studies have identified several examples of such a regulation in developing neurons. However, the systems-level effects of AS-NMD in this context are poorly understood.

Results: We developed an R package, factR2, which offers a comprehensive suite of AS-NMD analysis functions. Using this tool, we conducted a longitudinal analysis of gene expression in pluripotent stem cells undergoing induced neuronal differentiation. Our analysis uncovers hundreds of AS-NMD events with significant potential to regulate gene expression. Notably, this regulation is significantly overrepresented in specific functional groups of developmentally downregulated genes. Particularly strong association with gene downregulation is detected for alternative cassette exons stimulating NMD upon their inclusion into mature mRNA. By combining bioinformatic analyses with CRISPR/Cas9 genome editing and other experimental approaches we show that NMD-stimulating cassette exons regulated by the RNA-binding protein PTBP1 dampen the expression of their genes in developing neurons. We also provided evidence that the inclusion of NMD-stimulating cassette exons into mature mRNAs is temporally coordinated with NMD-independent gene repression mechanisms.

Conclusions: Our study provides an accessible workflow for the discovery and prioritization of AS-NMD targets. It further argues that the AS-NMD pathway plays a widespread role in developing neurons by facilitating the downregulation of functionally related non-neuronal genes.

Keywords: Alternative splicing, Nonsense-mediated decay, Neuronal development, Downregulation of non-neuronal genes, Custom transcriptome annotation, R package



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Many mammalian transcripts can undergo alternative splicing (AS), incorporating distinct exons and, occasionally, introns into mature mRNAs depending on the circumstances [1, 2]. AS plays a crucial role in the production of tissue- and cell type-specific protein isoforms [3, 4] with numerous functionally important and evolutionarily conserved examples of this regulation identified in the nervous system [5–8].

Besides diversifying the proteome, AS can regulate gene expression by changing the ratio between productively spliced mRNAs that encode functional proteins and unproductively spliced isoforms eliminated by RNA quality control mechanisms [9–11]. mRNAs containing premature translation termination codons > 50–55 nucleotides (nt) upstream of the last exon-exon splice junction typically undergo nonsense-mediated decay (NMD) in the cytoplasm following a pioneer round of translation [12–14]. NMD is well known for its role in destabilizing aberrant transcripts produced because of splicing errors or mutations in *cis*-elements or *trans*-acting factors involved in splicing regulation [15, 16].

Additionally, the scheduled production of alternatively spliced transcripts sensitive to NMD (AS-NMD) can facilitate the normal control of gene expression. For instance, several RNA-binding proteins use AS-NMD to auto- or cross-regulate their expression levels [17–21]. These post-transcriptional circuitries often rely on two types of alternative cassette exons (CEs) that trigger NMD when included into or skipped from mature mRNAs. We will refer to the former group as NMD-stimulating (NS-CEs; also known as “poison exons”), and the latter, NMD-repressing (NR-CEs).

AS-NMD is also known to contribute to the gene expression dynamics in development. A striking example of this regulation is provided by the large-scale upregulation of retained introns (RIs) in differentiating granulocytes and possibly other cell types, triggering NMD-mediated downregulation of functionally related genes [22–25]. Notably, bioinformatics analyses of tissue- and cell type-specific transcriptomes have identified NMD-regulating CE in diverse groups of genes [26, 27]. Yet, it is currently unknown whether specific groups of CE, or other exonic events such as alternative donors and acceptors (ADs and AAs), might coordinate gene expression dynamics on a scale similar to that observed for RIs.

Several lines of evidence point to the possibility of such systems-level regulation in developing neurons. For example, the RNA-binding protein PTBP1 controls NMD-repressing CE in transcripts encoding the neuron-enriched PTBP1 paralog PTBP2 and the post-synaptic proteins GABBR1 and DLG4/PSD-95 [28–31]. These exons are predominantly skipped in embryonic stem cells (ESCs) and neural progenitor cells (NPCs)—where PTBP1 is abundant—giving rise to NMD-sensitive splice forms. The downregulation of PTBP1 in developing neurons by the neuron-enriched microRNA miR-124 facilitates the CE inclusion, thus protecting *Ptbp2*, *Gabbr1*, and *Dlg4* mRNAs from NMD [29]. PTBP1 also represses NMD-stimulating CE and ADs in several non-neuronal genes including *Bak1*, *Flna*, and *Hps1*, limiting their expression to ESCs and NPCs [32–34]. The decline in PTBP1 levels in neurons promotes the inclusion of these events, rendering the mature mRNAs sensitive to NMD.

Mutations in key NMD factors are known to lead to neurological and psychiatric defects [35–38]. NMD is also known to control neural cell identity, fine-tune the

expression of important axonal and dendritic components, and regulate the excitation/inhibition balance in mature neurons [39–43]. It is tempting to speculate that these diverse functions involve coordinated regulation of different groups of AS-NMD events in a differentiation stage-specific manner.

However, understanding the developmental functions of AS-NMD at a systems level is a challenging task due to the lack of straightforward approaches for shortlisting biologically meaningful events. Moreover, NMD targets are inherently unstable and tend to be underrepresented in reference transcriptome annotations. The quantitative aspect of AS-NMD is also poorly understood. In most cases, it remains unclear whether this pathway drives developmental changes in gene expression or simply fine-tunes the abundance of productively spliced transcripts.

To systematically identify AS-NMD events with a strong potential to regulate genes in developing neurons, we developed a user-friendly computer package factR2 that annotates custom transcriptomes and prioritizes targets with significant correlation between NMD-protective splicing patterns and gene expression levels. By combining factR2 with longitudinal analysis of neuronal differentiation and acute inhibition of the NMD pathway, we uncovered a widespread role of NMD-stimulating CEs in downregulating non-neuronal genes in developing neurons. We validate our findings and demonstrate their generality using ex vivo neural cultures, subcellular fractionation, minigene and CRISPR-Cas9 experiments, and single-cell data analyses.

Results

An inducible system for longitudinal analysis of AS-NMD in developing neurons

We began by establishing a single-dish protocol for neuronal differentiation in vitro that avoids cell dissociation and replating steps (Fig. 1A). Inducible expression of the proneuronal transcription factor neurogenin 2 (NGN2; encoded by the *Ngn2/Neurog2* gene) has been shown to promote efficient differentiation of mouse ESCs into glutamatergic neurons [44]. In that study, *Ngn2* was expressed from a constitutive promoter following the

(See figure on next page.)

Fig. 1 Systematic analysis of the AS-NMD program in developing neurons. **A** Doxycycline (Dox) inducible differentiation of mouse ESCs into neurons. *Top*, RT-qPCR analysis of differentiation stage-specific markers in Dox-induced TRE-Ngn2 samples. Data are averaged from 3 experiments with SD shown by the shaded areas. *Bottom*, Dox-induced TRE-Ngn2 cells were treated for 6 h with DMSO or CHX and analyzed by RNA-seq. **B** factR2 streamlines the discovery of functional AS-NMD events and provides additional transcriptome annotation functions. **C** Genes containing Whippet-shortlisted AS-NMD events (see the “Results” section for details) show a stronger upregulation by CHX compared to non-AS-NMD genes. This is particularly prominent for genes containing NMD-stimulating cassette exons (NS-CEs). The $-\log_{10}$ -transformed FDR values for CHX-upregulated genes are presented as box plots and compared by a two-tailed Wilcoxon rank sum test. Dotted line, FDR = 0.001 cutoff used throughout this study. **D** Genes containing Whippet-shortlisted factR2 AS-NMD events are more likely to be significantly upregulated (FDR < 0.001) by CHX than genes lacking such events. **E** Distribution of AS-NMD events responding to CHX at both the splicing and gene expression levels at different stages of neuronal differentiation. **F** Wild-type (WT) mouse ESCs (IB10) and NPCs and neurons from mouse embryonic cortex were treated with CHX or DMSO and analyzed by RNA-seq as in (A). CHX-responsive AS-NMD events detected in these samples were intersected with three groups of the TRE-Ngn2 factR2 events: not regulated by CHX, regulated by CHX at the splicing level, and regulated by CHX at both the splicing and gene expression levels. The ESC and neural samples were analyzed separately. Note particularly strong enrichment for the third TRE-Ngn2 group. **G** Nuclear and cytoplasmic fractions of untreated samples introduced in (F) were analyzed by RNA-seq, and the nuclear enrichment of putative AS-NMD events in the three TRE-Ngn2 groups mentioned in (F) was calculated by Whippet

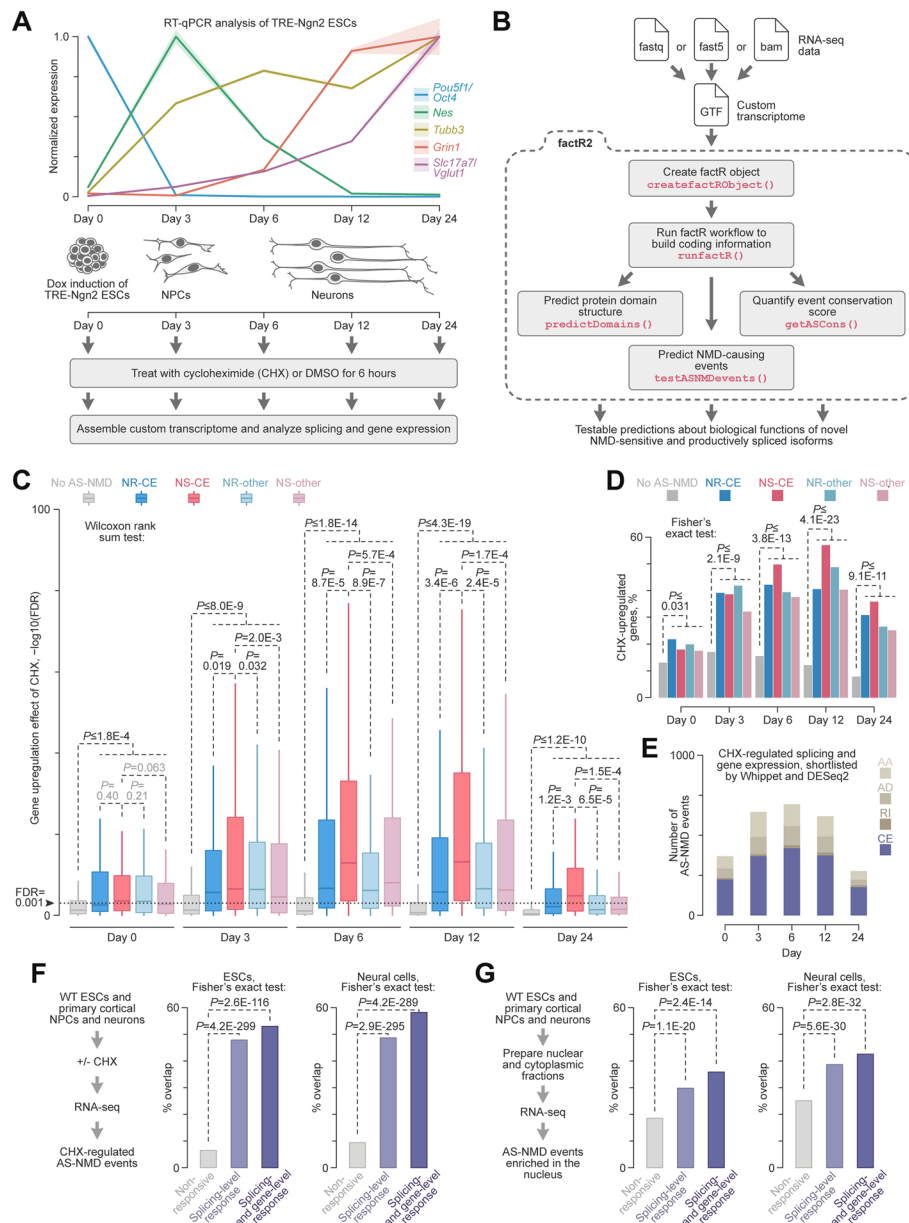


Fig. 1 (See legend on previous page.)

4-hydroxytamoxifen-induced Cre recombination [44]. Yet, in mice, *Ngn2* upregulation in NPCs is transient, followed by a decrease in its expression at later stages of differentiation [45]. To better recapitulate the endogenous regulation, we knocked in a mouse *Ngn2* transgene into the A2Lox ESC line [46] under the doxycycline (Dox) inducible promoter *TRE* (Fig. 1A).

The resultant TRE-Ngn2 line maintained its ESC properties in the 2i+LIF medium [47] without Dox, producing dome-shaped colonies positive for the POU5F1/OCT4 marker (Additional file 1: Fig. S1A). When treated with Dox (see the “Materials and methods” section for details), most cells lost OCT4, acquired spindle-shaped morphology typical for NPCs, and increased the expression of the NPC marker vimentin (VIM) by differentiation day 3 (Additional file 1: Fig. S1B). To mimic transient upregulation of

Ngn2 in neurogenesis, we progressively reduced Dox concentration beginning from day 2 (see the “Materials and methods” section). By differentiation day 6, most cells stained positive for the early neuronal marker tubulin β III (TUBB3/TUJ1) and developed characteristic neuronal morphology with round somas and readily detectable neurites (Additional file 1: Fig. S1C).

The differentiating cultures were viable for at least 24 days and showed an increased expression of the late neuronal marker MAP2 by days 12–24 (Additional file 1: Fig. S1C). By day 24, cells displayed signs of neuronal maturation, including the presence of ankyrin G (ANK3/ANKG) at the axon initial segment and punctate distribution of the postsynaptic density protein-95 (DLG4/PSD-95), as well as the presynaptic marker Synaptophysin (SYP) in neurites (Additional file 1: Fig. S1D–F). Reverse transcriptase-quantitative PCR (RT-qPCR) assays confirmed stage-specific expression of corresponding mRNA markers (Fig. 1A). Consistent with the transient Dox treatment used in our protocol, *Ngn2* expression peaked at day 3 and declined at the later stages of differentiation (Additional file 1: Fig. S1G).

To understand the role of AS-NMD in neurodevelopment, we treated differentiating TRE-*Ngn2* cultures for 6 h prior to RNA extraction with cycloheximide (CHX), to repress NMD, or DMSO, as a control (Fig. 1A). The samples collected on days 0, 3, 6, 12, and 24 were then analyzed by RNA-seq. Principal component analysis (PCA) of gene expression in this experiment showed a tight clustering of biological replicates and a clear separation of samples according to the differentiation stage and the treatment (Additional file 1: Fig. S2A). Further inspection of the DMSO samples confirmed stage-specific expression of pluripotency markers on day 0, NPC markers on day 3, and progressive upregulation of mature neuronal and glutamatergic markers between days 6 and 24 (Additional file 1: Fig. S2B). This developmental trajectory was further confirmed by deconvolving cell types in the DMSO samples using MuSiC [48] (Additional file 1: Fig. S2C, D).

We concluded that transient expression of transgenic NGN2 in mouse ESCs provides a suitable system for time-resolved analyses of AS-NMD targets in developing neurons.

A bioinformatics tool for annotating AS-NMD events in custom transcriptomes

To enable in-depth analysis of AS-NMD targets, we updated our prototype R tool factR (functional annotation of custom transcriptomes; [49]) with a suite of relevant functions. The new package, factR2 (Fig. 1B; [50]), allocates transcripts identified by RNA-seq to their genes of origin and identifies NMD-sensitive isoforms. It further detects splicing events underlying the NMD sensitivity and analyzes their regulation potential. factR2 can also deduce the domain structure of productively spliced protein-coding sequences and plot RNA and protein isoform structures for data exploration. Overall, factR2 provides a rich set of functions for the functional annotation of custom transcriptomes and AS-NMD analyses (Additional file 2: Table S1).

As a proof of principle, we searched for NMD-sensitive transcripts among the 90,159 mRNAs present in the GENCODE M26 reference transcriptome using factR2's predict-NMD function. The workflow was kept blind to the GENCODE transcript biotype annotation. This returned 13,042 putative NMD targets containing premature termination codons > 50 nt upstream of the last exon-exon junction. The set included 7161 out of

the 7201 transcripts (i.e. 99.4%) with the “nonsense_mediated_decay” biotype, as well as 5881 additional transcripts that had not been previously flagged as possible NMD targets (Additional file 1: Fig. S3A).

Transcript-level analysis of our DMSO- and CHX-treated TRE-Ngn2 RNA-seq data (Fig. 1A) by HISAT2-StringTie [51] predicted 57,535 novel mRNA isoforms absent from the GENCODE annotation. Of these, 30,228 transcripts were classified by factR2 as NMD-sensitive. The analysis of the combined NMD-sensitive transcriptome using factR2’s testASNMDevents function revealed 19,198 putative AS-NMD events. These events were classified (Additional file 1: Fig. S3B, C) as cassette exons (CEs), alternative donors (ADs), alternative acceptors (AAs), and retained introns (RIs) that either stimulate or repress NMD when included into mature mRNA (NS and NR events, respectively).

These data demonstrate the utility of factR2 and suggest that the extent of AS-NMD regulation in developing neurons is substantially greater than currently thought.

factR2 identifies AS-NMD events with a strong regulation potential

To shortlist high-quality AS-NMD targets, we first examined changes in their splicing patterns and gene expression in response to CHX. We used Whippet [52] for the splicing-level analysis and selected CEs, ADs, AAs, and RIs consistently changing their percent spliced in (PSI) values in response to CHX at least at one differentiation stage (NS events, $\Delta\text{PSI} > 0.1$; NR events, $\Delta\text{PSI} < -0.1$; probability > 0.9). We additionally selected NS events with a PSI average of > 0.9 for DMSO- and CHX-treated samples, and NR events with a PSI average of < 0.1 . This was done because these groups have a limited capacity to change splicing in response to CHX but may still be regulated at the gene expression level.

Genes with Whippet-shortlisted events were more frequently upregulated in response to CHX compared to genes lacking AS-NMD events (Fig. 1C). In many cases, the false discovery rate (FDR) of this effect, calculated by the DESeq2 Wald test [53], was < 0.001 . Cassette exons stimulating NMD upon their inclusion (NS-CEs) were associated with particularly robust regulation in neural cells (days 3–24; Fig. 1C). The fractions of Whippet-shortlisted entries passing the FDR < 0.001 gene-level regulation cutoff were significantly larger compared to genes lacking factR2-predicted AS-NMD events (Fig. 1D). Overall, we identified 1432 non-redundant AS-NMD events responding to CHX at both splicing and gene levels (Additional file 3: Table S2). The number of such events peaked at the intermediate stages of differentiation (days 3–12), with CEs dominating the distributions throughout the entire time course (Fig. 1E). Notably, 49.2% of the CHX-responsive events were not previously annotated in the GENCODE transcriptome. Similar results were obtained when we shortlisted CHX-responsive splicing events by rMATS [54] instead of Whippet (Additional file 1: Fig. S3D).

This analysis suggests that the AS-NMD pathway controls hundreds of genes expressed at different stages of neuronal differentiation.

Many AS-NMD events identified in TRE-Ngn2 cells are regulated in natural samples

To validate our approach, we performed RNA-seq analyses of wild-type (WT) mouse ESCs (IB10 line) and primary cortical NPCs and neurons treated with CHX or DMSO

(Fig. 1F and Additional file 1: Fig. S4A–D). Splicing events responding to CHX in these “natural” samples (with the Whippet cutoffs described above) were significantly over-represented among the CHX-responsive TRE-Ngn2 factR2 events compared to the non-responsive TRE-Ngn2 factR2 control (Fig. 1F).

Since NMD occurs in the cytoplasm [12–14], transcripts including NMD-stimulating and excluding NMD-repressing events are expected to be enriched in the nucleus. We reasoned that such nuclear enrichment could provide a CHX-independent validation method for AS-NMD targets. With this in mind, we sequenced nuclear and cytoplasmic transcriptomes from untreated WT ESCs and cortical NPCs and neurons and compared the two compartments using Whippet (Fig. 1G). Nucleus-enriched NR and nucleus-depleted NS events were overrepresented in the CHX-responsive TRE-Ngn2 factR2 sets (Fig. 1G). Notably, the sequence context of events responding to CHX in both the TRE-Ngn2 and natural cells was more conserved across vertebrates compared to a non-responsive control (Additional file 1: Fig. S4E, F).

These data argue that many AS-NMD events identified by factR2 are regulated in vivo. Since the targets responding to CHX at both the splicing and gene expression levels showed the largest overlap with the natural regulation (Fig. 1F, G) and the highest median conservation scores (Additional file 1: Fig. S4E, F), we prioritized this set in our subsequent analyses.

AS-NMD is widespread among genes downregulated in developing neurons

So far, we focused on the impact of AS-NMD on individual stage-specific transcriptomes. To explore the potential effects of this pathway on gene expression dynamics in development, we analyzed the correlation between splicing patterns protecting mRNA from NMD (i.e. inclusion of NR events and skipping of NS events) and gene expression levels. This was done by applying factR2’s function testGeneCorr to genes differentially expressed (DESeq2 likelihood-ratio test; FDR < 0.001) in the control-treated longitudinal TRE-Ngn2 data.

Compared to the non-responsive control, CHX-responsive AS-NMD events were enriched for positively correlated entries (Fig. 2A). The enrichment was predominantly due to CEs rather than other types of AS-NMD events (Fig. 2A, B). The largest fraction of significantly correlated events was detected for NS-CEs, with the magnitude of this effect increasing for more stringent correlation P -values and Whippet cutoffs (Fig. 2A, B and Additional file 1: Fig. S5). The enrichment of correlated events among NS-CEs ranged from ~2.4-fold ($\Delta\text{PSI} > 0.1$ and probability > 0.9 Whippet cutoffs and $P < 0.05$ correlation cutoff; Fig. 2A, B) to ~5.8-fold ($\Delta\text{PSI} > 0.25$ and probability > 0.95 Whippet cutoffs and $P < 0.001$ correlation cutoff; Additional file 1: Fig. S5). Hereafter, we will use the term “facilitating” to refer to CHX-responsive events showing significant positive correlation ($P < 0.05$; FDR < 0.1) between NMD-protective splicing patterns and gene expression.

Since the above results pointed to a possible role of AS-NMD in gene expression dynamics, we wondered if such a mechanism might preferentially feed into specific types of temporal regulation. We clustered differentially expressed genes (DESeq2 likelihood-ratio test; FDR < 0.001) according to their TRE-Ngn2 time-course trajectories using

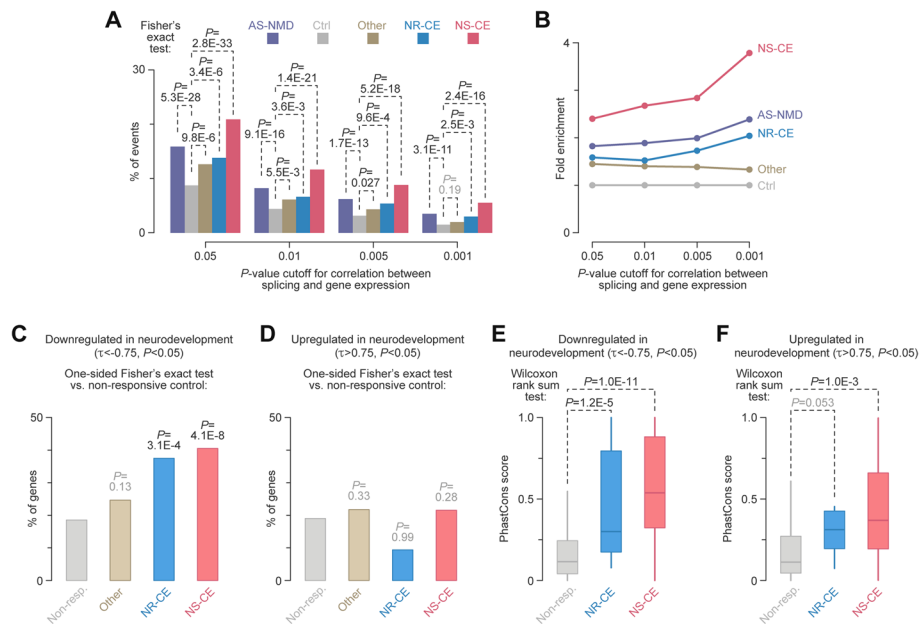


Fig. 2 Neurodevelopmentally downregulated genes are often controlled by AS-NMD. **A, B** factR2 events responding to CHX at both the splicing and gene levels show a stronger correlation with developmental changes in gene expression compared to non-responsive events. **A** Percentages of events correlating with gene expression dynamics in differentiating TRE-Ngn2 cells for different P -value cutoffs. Ctrl are factR2 events not responding to CHX. AS-NMD, Other, NR-CE, and NS-CE are, respectively, all AS-NMD events, AD, AA, and RI AS-NMD events, NMD-repressing cassette exons, and NMD-stimulating cassette exons responsive to CHX. **B** Fold enrichments of the correlated events in the shortlisted categories in **(A)** plotted as a function of the correlation P -value cutoff. NS-CEs show the strongest enrichment for correlated entries, with the fold enrichment increasing for tighter cutoffs. **C** One-sided Fisher's exact test showing that the monotonic downregulation trend (Kendall's $\tau < -0.75$, $P < 0.05$) is significantly overrepresented among genes with CHX-responsive facilitating CEs, particularly NS-CEs, correlating with gene expression dynamics ($P < 0.05$). Genes with non-responsive factR2 events are used as a control. **D** No significant difference is detected for the monotonic upregulation trend (Kendall's $\tau > 0.75$, $P < 0.05$). **E, F** PhastCons [55] analysis shows strong conservation of the intronic sequence context in NS-CEs in general and NS-CEs residing in monotonically downregulated genes in particular

DP_GP_cluster [56] or k-means methods (Additional file 1: Fig. S6, Additional file 1: Fig. S7A–H, and Additional file 4: Table S3).

Notably, facilitating AS-NMD events were often associated with neurodevelopmentally downregulated gene clusters (Additional file 1: Fig. S7C–F and Additional file 5: Table S4). The enrichment was especially robust for genes with facilitating NS-CEs (Additional file 1: Fig. S7F–H). Similar trends were detected when we analyzed the trajectories of individual genes without clustering (Additional file 1: Fig. S7I and Fig. 2C, D).

Consistent with the opposite effects of NR-CE and NS-CE inclusion on mRNA stability, Kendall's analysis of the PSI dynamics revealed a skew towards negative τ values in NR-CEs and positive τ values in NS-CEs (Additional file 1: Fig. S7J). Events with significant positive PSI trends ($P < 0.05$) were significantly overrepresented in the NS-CE group (Additional file 1: Fig. S7K). Moreover, the intronic sequence context of NS-CEs encoded in monotonically downregulated genes was often conserved across species (Fig. 2E, F).

Overall, our pipeline identified 87 facilitating NS-CEs associated with monotonic gene downregulation (Kendall's $\tau < -0.75$, $P < 0.05$; the first tab in Additional file 5: Table S4).

Metascape [57] analysis of this set revealed significant enrichment for RNA metabolism and localization as well as cell cycle regulation functions (Additional file 1: Fig. S8A and the first tab in Additional file 6: Table S5). Other examples of the overrepresented functional categories included RHO GTPase effectors and actin cytoskeleton organization (Additional file 1: Fig. S8A).

To validate our bioinformatics predictions, we examined four monotonically downregulated genes ($\tau < -0.75$, $P < 0.05$; the first tab in Additional file 5: Table S4) with evolutionarily conserved NS-CEs (Additional file 1: Fig. S9). These included three examples encoding RNA metabolism factors, *Fbl*, *Srsf9*, and *Xpo1* (the first tab in Additional file 6: Table S5), and the *Cttnal1* gene predicted to contribute to actin filament binding and RHO signaling (<https://www.ncbi.nlm.nih.gov/gene/8727>). All four NS-CEs were readily detectable in CHX-treated samples showing progressively stronger inclusion at later stages of differentiation.

Thus, AS-NMD in general and NMD-stimulating cassette exons in particular may play a role in downregulation of functionally related groups of genes in developing neurons.

Comprehensive identification of AS-NMD events regulated by PTBP1

Previous studies have identified several AS-NMD events controlled by PTBP1, an RNA-binding protein expressed at high levels in ESCs and NPCs and downregulated in neurons [28–31, 34]. We therefore used PTBP1-regulated events to benchmark the performance of the factR2-based workflow. *Ptbp1* was monotonically downregulated in differentiating TRE-Ngn2 cells, and the shortlist of facilitating AS-NMD events included previously characterized PTBP1 targets (e.g., NR-CEs maintaining open reading frames in *Ptbp2*, *Gabbr1*, and *Dlg4*; Additional file 1: Fig. S10A and Additional file 5: Table S4).

To identify PTBP1-dependent AS-NMD targets systematically, we examined factR2 events correlating with *Ptbp1* expression in DMSO-treated TRE-Ngn2 cells and changing their splicing in ESCs or NPCs in response to PTBP1 knockdown alone or together with its functionally similar paralog, PTBP2 (Fig. 3A; RNA-seq data from [58]; Whippet $|\Delta\text{PSI}| > 0.1$ and probability > 0.9 cutoffs; see the “Materials and methods” section for further details).

Notably, the PTBP1-dependent behavior was significantly overrepresented among the high-quality TRE-Ngn2 factR2 events (Fig. 3B). Of the 46 facilitating AS-NMD events regulated by PTBP1, 29 were NS-CEs (Additional file 5: Table S4). Similar to the trend observed for all facilitating NS-CEs (Additional file 1: Fig. S7E, F), the PTBP1-regulated NS-CEs were strongly enriched in monotonically downregulated ($\tau < -0.75$, $P < 0.05$) but not upregulated ($\tau > 0.75$, $P < 0.05$) genes (Additional file 1: Fig. S10B, C).

Metascape analysis showed that monotonically downregulated genes with PTBP1-controlled NS-CEs were enriched for the RHO GTPase effector and actin cytoskeleton organization functions (Additional file 1: Fig. S8B and the second tab in Additional file 6: Table S5). Another enriched category related to the programmed cell death (apoptosis modulation by HSP70; Additional file 1: Fig. S8B). Interestingly, the overrepresentation of the RHO GTPase effector and actin cytoskeleton organization terms among all downregulated genes with NS-CEs (Additional file 1: Fig. S8A) was largely due to the PTBP1-regulated targets. Indeed, these terms were missing in the Metascape output when we

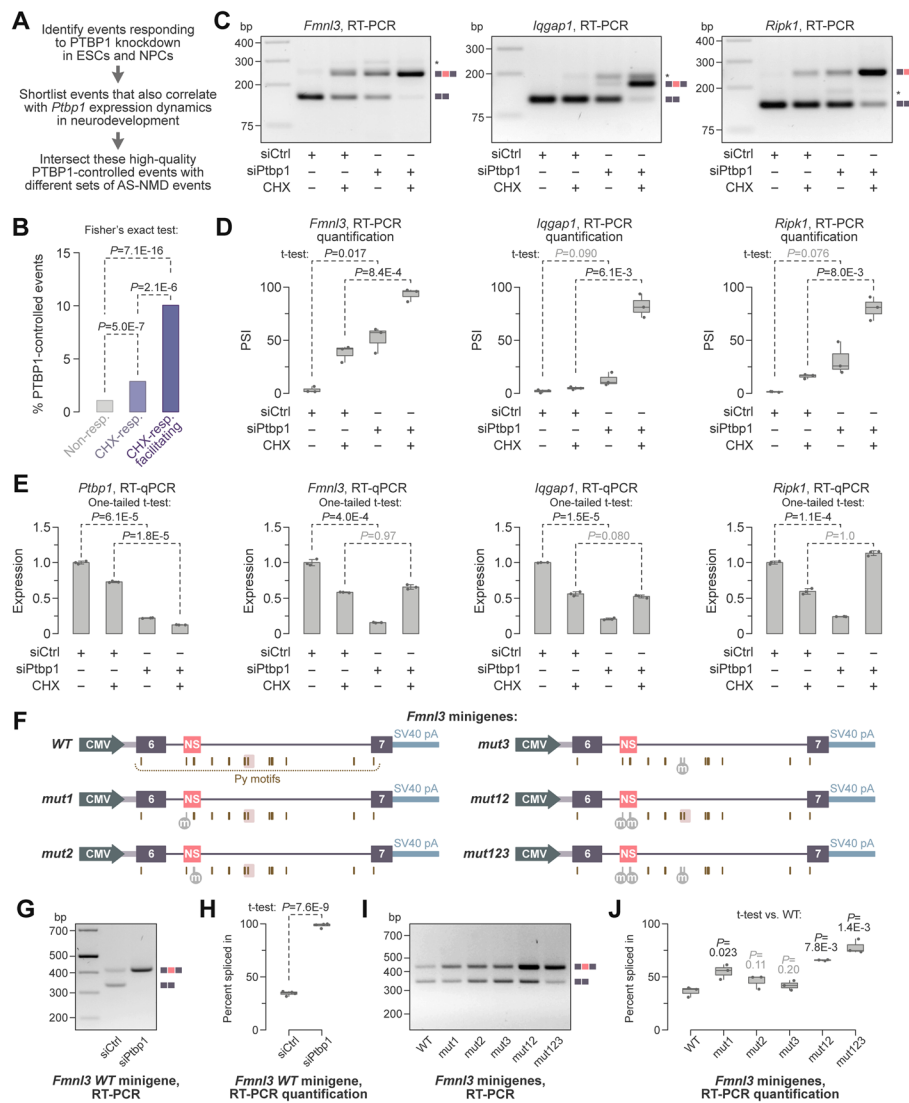


Fig. 3 Role of PTBP1 in AS-NMD regulation in developing neurons. **A** The workflow used to shortlist PTBP1-controlled events. **B** AS-NMD events with strong regulatory potential are enriched for PTBP1-controlled entries. **C–E** Uninduced TRE-Ngn2 ESCs were treated for 48 h with a PTBP1-specific (siPtbp1) or a non-targeting control siRNA (siCtrl) and post-treated with CHX (+ CHX) or DMSO (– CHX) for an additional 6 h. **C** The samples were analyzed by RT-PCR with primers flanking the *Fmn13*, *Iqgap1*, and *Ripk1* NS-CEs. PTBP1 represses NS-CE inclusion and CHX stabilizes NS-CE-containing isoforms. **D** PSI values in **(C)** were quantified from 3 experiments and shown as box plots. **E** RT-qPCR analyses with primers against constitutively spliced parts of *Fmn13*, *Iqgap1*, and *Ripk1* show that siPtbp1 downregulates these genes in the presence of DMSO but not CHX. *Ptbp1*-specific primers were used to estimate its knockdown efficiency. The data were averaged from three experiments \pm SD. **F** *Fmn13* minigenes. Pyrimidine-rich motifs (Py; YUCUYY and YYUCUY) predicted to recruit PTBP1 within the NS-CE or at intronic positions overlapping with a PTBP1 iCLIP cluster [58] (blush highlight) were mutated (“m”) in the corresponding minigenes. **G** TRE-Ngn2 ESCs pretreated with siPtbp1 or siCtrl were transfected with the wild-type (WT) *Fmn13* minigene and analyzed by RT-PCR with minigene-specific primers. The minigene recapitulates the repressive effect of PTBP1 on the *Fmn13* NS-CE. **H** Quantification of the NS-CE PSI values in **(G)**. **I** Untreated TRE-Ngn2 ESCs were transfected with the *Fmn13* minigenes introduced in **(F)** and analyzed by RT-PCR. **J** Quantification of the PSI values in **(I)** showing that the mutation of the upstream exonic Py (*mut1*) promotes the NS-CE inclusion, especially when combined with downstream exonic Py (*mut2*) and iCLIP Py (*mut3*) mutations. Box plots in **H, J** show data from 3 experiments compared by a two-tailed *t*-test assuming unequal variances

analyzed monotonically downregulated genes with PTBP1-independent NS-CEs (Additional file 1: Fig. S8C and the third tab in Additional file 6: Table S5).

These data suggest that PTBP1 controls a wider range of NMD-stimulating cassette exons than previously thought, and that many genes containing these events belong to specific functional categories.

PTBP1-mediated repression of NS-CEs promotes the expression of their genes in ESCs

We selected three neurodevelopmentally downregulated genes with PTBP1-repressed NS-CEs for further analyses: *Fmnl3* and *Iqgap1* encoding RHO GTPase effectors involved in actin cytoskeleton organization and *Ripk1*, a regulator of programmed cell death (the second tab in Additional file 6: Table S5). The three genes were strongly downregulated in developing neurons and their expression was rescued by CHX (Additional file 1: Fig. S10D-F). Their NS-CEs were evolutionarily conserved and neighbored or overlapped with PTBP1 interaction motifs (YUCUYY and YYUCUY) and PTBP1 binding sites identified by iCLIP [58] (Additional file 1: Fig. S11). PTBP1 was proposed to control *Ripk1* and *Iqgap1* NS-CE splicing [27, 59–61], but the role of these exons in the regulation of their host genes during neuronal differentiation was not addressed experimentally. To the best of our knowledge, the *Fmnl3* NS-CE has not been previously identified.

To validate the role of PTBP1 in the expression of these targets, we transfected TRE-Ngn2 ESCs with *Ptbp1*-specific or control siRNAs (si*Ptbp1* or siCtrl, respectively) and post-treated the cells with either CHX or DMSO. Subsequent RT-PCR analyses showed that siPTBP1 stimulated the inclusion of all three NS-CEs, which was further enhanced by CHX (Fig. 3C, D). PTBP1 knockdown dampened the total mRNA abundance for the three genes in the DMSO- but not CHX-treated samples (Fig. 3E). Importantly, the downregulation of *Fmnl3*, *Iqgap1*, and *Ripk1* by si*Ptbp1* was also significantly rescued by an siRNA against the key NMD factor UPF1 (Additional file 1: Fig. S12).

To test if PTBP1 might regulate NS-CE splicing directly, we cloned the *Fmnl3* NS-CE in its natural sequence context into a mammalian expression vector (Fig. 3F). Transcripts produced from this minigene lacked an in-frame translation initiation codon and were not predicted to undergo NMD. The wild-type minigene (*WT*) recapitulated the endogenous splicing regulation in TRE-Ngn2 ESCs (Fig. 3G, H). There are two PTBP1 binding sites inside the *Fmnl3* NS-CE and two such positions in the downstream intron that overlap with a PTBP1 iCLIP cluster (Additional file 1: Fig. S11A). Mutation of the upstream exonic site (*mut1*; Fig. 3F) increased the NS-CE inclusion in the presence of PTBP1 (Fig. 3I, J). Mutations of the downstream exonic site (*mut2*) or the two iCLIP positions (*mut3*) did not alter the minigene splicing pattern individually, but enhanced the NS-CE stimulation effect of *mut1* in the *mut12* and *mut123* minigenes (Fig. 3F, I–J).

We concluded that PTBP1 promotes the expression of several non-neuronal genes in ESCs by repressing NMD-stimulating cassette exons.

***Ptbp1* downregulation in developing neurons may act as a switch for NS-CE inclusion**

The TRE-Ngn2 RNA-seq data showed a substantial decrease in *Ptbp1* expression between differentiation days 3 (NPCs) and 6 (young neurons), which coincided with the expression trajectories of *Fmnl3*, *Iqgap1*, and *Ripk1* (Additional file 1: Fig. S10A). To find

out whether the *Ptbp1* dynamics could promote the NS-CE inclusion in these targets, we analyzed the TRE-Ngn2 differentiation time course collecting DMSO- or CHX-treated samples daily between days 0 and 6 (Fig. 4).

RT-PCR analysis of DMSO-treated samples (Fig. 4A) did not detect the NS-CE-containing isoforms of *Fmnl3*, *Iqgap1*, and *Ripk1* between days 0 and 2. Small amounts of these isoforms appeared on days 3–4 and accumulated at the subsequent stages of differentiation. Treating the cells with CHX tended to increase the relative abundance of the NS-CE-containing isoforms with a robust difference between the CHX and the DMSO PSI values in newly developed neurons (Fig. 4B).

Since CHX protects NS-CE-containing isoforms from NMD, we used the CHX-treated data to follow developmental changes in pre-mRNA splicing patterns (Fig. 4B). In all three cases, the time-course curves were S-shaped with little or no changes in the NS-CE inclusion between days 0 and 2, a rapid increase between days 2 and 4, and a plateau between days 4 and 6. Using the median PSI value for the three genes as a threshold (43.8%; dotted line in Fig. 4B) placed the onset of the rapid increase phase at day 2 for

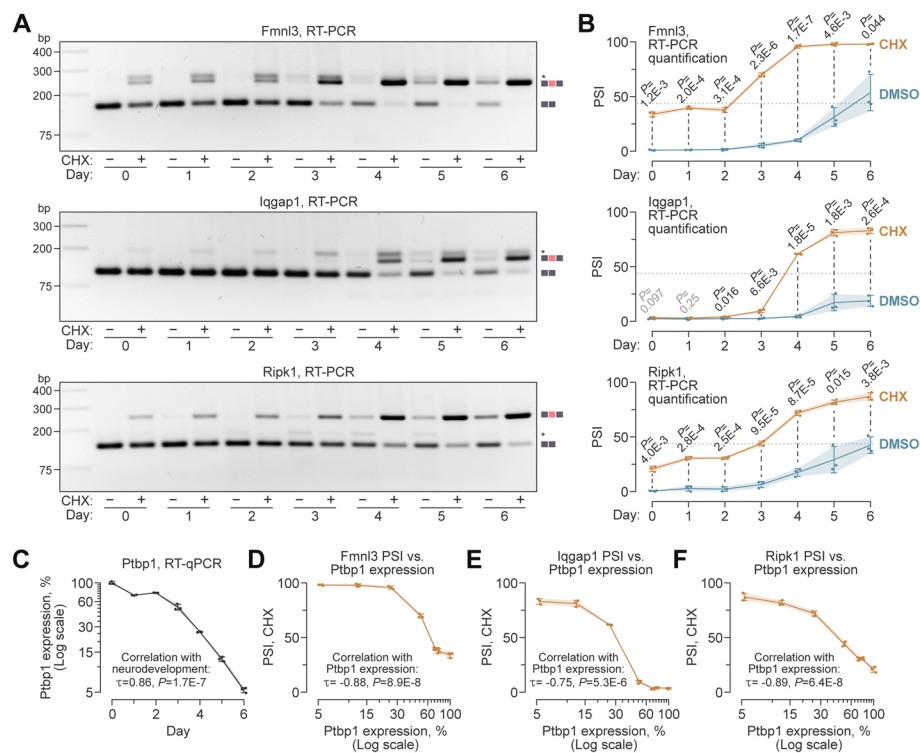


Fig. 4 PTBP1-repressed NS-CEs are regulated in a switch-like manner during neuronal differentiation. **A** RT-PCR analyses of Dox-induced TRE-Ngn2 ESCs treated with DMSO or CHX on days 0–6 show that the inclusion of NS-CEs into the *Fmnl3*, *Iqgap1*, and *Ripk1* mRNAs increases during neuronal differentiation. **B** NS-CE PSI values for the DMSO and CHX samples in **(A)** averaged from three experiments \pm SD and compared by two-tailed *t*-test assuming unequal variances. Note that the CHX curves are S-shaped, consistent with a switch-like regulation. Dotted lines, PSI median of the CHX-treated data for the three genes used to estimate the onset of the rapid increase phase. **C** RT-qPCR analysis *Ptbp1* expression dynamics during the day 0–6 differentiation period. Note that the abundance of the *Ptbp1* mRNA decreases monotonically beginning from day 2, consistent with the onset of the rapid increase phase in the CHX curves in **(B)**. **D–F** NS-CE inclusion in **(D)** *Fmnl3*, **(E)** *Iqgap1*, and **(F)** *Ripk1* shows a strong negative correlation with *Ptbp1* expression, as expected. Correlations in **(C–F)** were analyzed using the Kendall's method

Fmnl3 and day 3 for *Iqgap1* and *Ripk1*. RT-qPCR analysis of *Ptbp1* expression showed that it was relatively stable between days 0 and 2, and progressively declined beginning from day 2 (Fig. 4C). Notably, NS-CE-inclusion values showed a strong negative correlation with *Ptbp1* for all three targets (Fig. 4D–F).

Taken together, these results suggest that reduced *Ptbp1* expression may facilitate a switch towards the inclusion of NMD-stimulating cassette exons during the transition from NPCs to neurons.

NS-CEs work in coordination with NMD-independent gene downregulation mechanisms

Increased inclusion of NS-CEs may contribute to gene downregulation in developing neurons in four different ways. (1) It may act as the sole downregulation mechanism. Alternatively, it may work alongside AS-NMD-independent repression mechanisms by (2) preempting, (3) coinciding with, or (4) following them in development. We modeled these four scenarios assuming possible cooperation with transcriptional repression (Additional file 1: Fig. S13 and Materials and Methods). The key predictions from this simple model were that AS-NMD should consistently diminish gene expression later in development in all four scenarios and accelerate the onset of downregulation in scenarios (1) and (2), but not (3) and (4).

To find out which of the above mechanisms could be used in developing neurons, we targeted NS-CE sequences in *Fmnl3*, *Iqgap1*, and *Ripk1* using appropriate CRISPR gRNAs and Cas9 (Additional file 1: Fig. S14). For each gene, we selected two TRE-Ngn2 ESC clones with biallelic NS-CE disruption (Additional file 1: Fig. S14). RT-PCR analyses of undifferentiated (day-0 ESCs) and Dox-differentiated (day-6 neurons) samples treated with DMSO or CHX showed that the mutant clones lost or had a severely impaired ability to splice in NMD-stimulating exons (Additional file 1: Fig. S15).

We then analyzed the expression dynamics of the three genes in the wild-type (WT) and mutant cells undergoing neuronal differentiation (Fig. 5A–C). As expected, the three WT genes were downregulated in day-6 neurons compared to day-0 ESCs. *Iqgap1* and *Fmnl3* additionally showed transient upregulation peaks on days 1–2, which could not be detected in our day 0, 3, 6, 12, and 24 RNA-seq time series. The mutants generally followed the WT trends (Fig. 5A–C), with two key differences revealed by normalizing the data by the wild-type values (Fig. 5D–F).

First, NS-CE inactivation led to progressive upregulation of gene expression at later developmental stages (Fig. 5D–F). For all 3 genes, the mutant trajectories diverged from their WT counterparts after an initial lag period. By using the median of the normalized mutant values for all three genes as a cutoff (1.73; dotted lines in Fig. 5D–F), we estimated that the WT-mutant dichotomy point was around day 2 for *Fmnl3* and day 3 for *Iqgap1* and *Ripk1*. According to Additional file 1: Fig. S13, these points mark the onset of facilitating AS-NMD in the WT, and they indeed coincided with the corresponding parts of the NS-CE inclusion curves in CHX-treated WT samples (Fig. 4B).

Second, the temporal relationship between the AS-NMD dichotomy points in Fig. 5D–F and the monotonically downregulated parts of the mutant trajectories in Fig. 5A–C varied depending on the gene. The estimated onset of AS-NMD preceded the monotonic downregulation for the *Fmnl3* (day 2 vs. day 3), while following it for *Iqgap1* and *Ripk1* (day 3 vs. day 2). These behaviors matched models (2) and (4) in

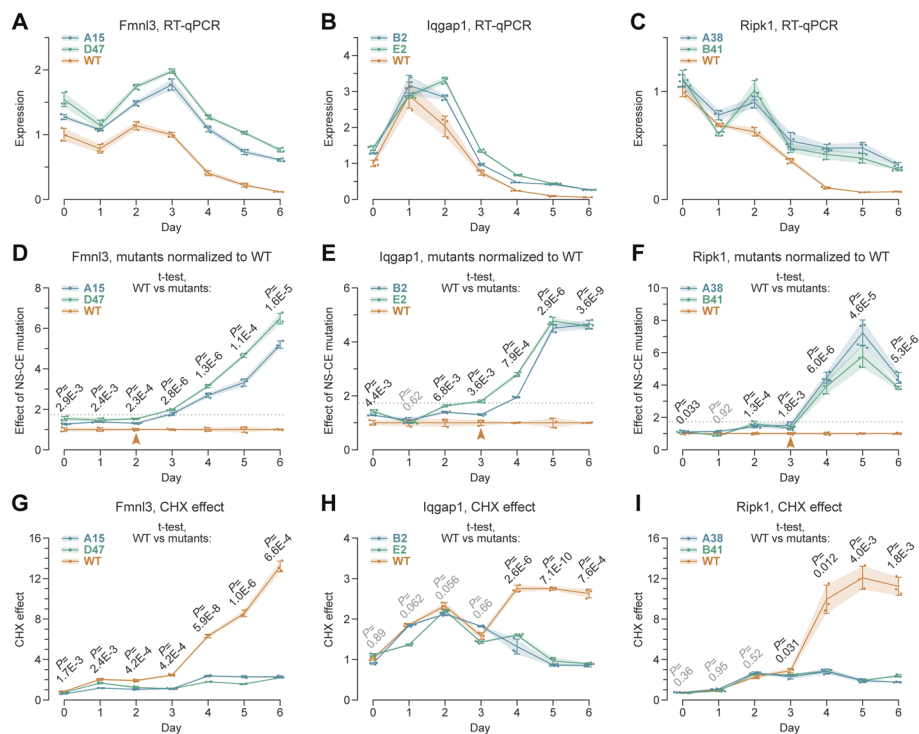


Fig. 5 NS-CEs facilitate neurodevelopmental downregulation of non-neuronal genes. **A–C** TRE-Ngn2 ESCs with wild-type (WT) or biallelically mutated NS-CEs (*Fmnl3*, clones A15, and D47; *Iqgap1*, clones B2 and E2; and *Ripk1*, clones A38 and B41) were differentiated into neurons and the gene expression dynamics was analyzed by RT-qPCR. Note that the mutant (**A**)*Fmnl3*, (**B**)*Iqgap1*, and (**C**)*Ripk1* have a higher residual expression in neurons compared to their WT counterparts. Mutant *Fmnl3* is also expressed somewhat higher than the WT at earlier stages of differentiation. **D–F** Normalizing the data in (**A–C**) to the WT values confirms that (**D**)*Fmnl3*, (**E**) *Iqgap1*, and (**F**)*Ripk1* with mutated NS-CEs have a substantially reduced ability to undergo downregulation at later stages of development. We defined WT-mutant dichotomy points as differentiation stages preceding the increase of normalized expression values averaged for the two mutants above the median of all normalized mutant values calculated for the three genes (1.73; dotted line). This places the dichotomy points at day 2 for *Fmnl3* and day 3 for *Iqgap1* and *Ripk1* (arrowheads). **G–I** Normalizing CHX-treated time-course to the corresponding DMSO-treated controls for (**G**)*Fmnl3*, (**H**)*Iqgap1*, and (**I**)*Ripk1* confirms that the AS-NMD regulation is lost or markedly diminished in the NS-CE mutants, and that days 2–3 correspond to the onset of facilitating AS-NMD in the WT

Additional file 1: Fig. S13, with an apparent 1-day difference between the onset of NMD-dependent and NMD-independent downregulation mechanisms. Confirming the role of the NS-CEs in the AS-NMD regulation, CHX strongly upregulated the WT but not the mutant genes at the corresponding differentiation stages (Fig. 5G–I).

Thus, in the three examples tested, NS-CEs are essential for full-scale downregulation of non-neuronal genes in developing neurons, working in coordination with NMD-independent repression mechanisms. Our additional bioinformatics analysis showed that genes containing functional NMD-stimulated cassette exons generally maintained declining neurodevelopmental trends in the presence of CHX, even though their downregulation was more robust in DMSO- compared to CHX-treated samples (Additional file 1: Fig. S16). This points to a wider crosstalk between AS-NMD and other gene regulation mechanisms.

Discussion

Our study offers valuable insights into the gene regulation functions of the AS-NMD pathway. Firstly, it introduces a straightforward workflow for shortlisting biologically relevant AS-NMD targets. Secondly, it reveals the extensive involvement of AS-NMD in the downregulation of non-neuronal genes in developing neurons. Thirdly, it argues that the activity of NMD-stimulating cassette exons (NS-CEs) is often coordinated with other downregulation mechanisms to ensure full-scale repression of their target genes in mature neurons (Fig. 6).

Additionally, our optimized inducible differentiation system provides a useful resource for time-resolved analyses of gene expression dynamics in developing neurons (Fig. 1A, Additional file 1: Fig. S1 and Additional file 1: Fig. S2). We also

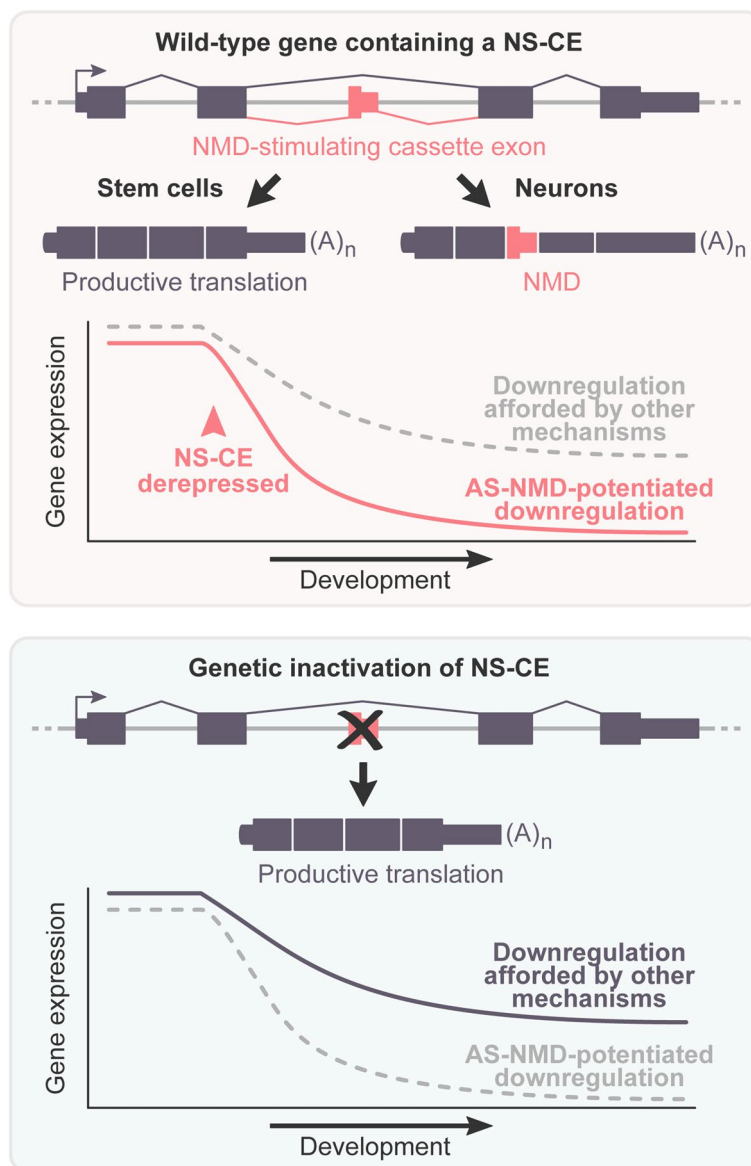


Fig. 6 Model for the role of NMD-stimulating cassette exons in gene downregulation in developing neurons

believe that users working with custom transcriptomes deduced from either short- or long-read sequencing data will appreciate the ability of factR2 (Fig. 1B and Additional file 2: Table S1) to quickly allocate sample-specific transcripts to their original genes (done when creating the factR2 object), predict open reading frames (buildCDS), identify transcripts containing NMD-promoting features (predictNMD), and pinpoint AS events underlying NMD sensitivity (testASNMDevents). Functions testGeneCorr and getAScons are helpful for assessing the regulatory potential and evolutionary conservation of splicing events, while plotTranscripts provides rich data visualization capabilities. Although not showcased in this study, factR2 can be also used to predict the domain structure of proteins encoded by productively spliced isoforms (predictDomain).

In our evaluation, factR2 provides a wider range of functions for the functional analysis of custom transcriptomes and shortlisting AS-NMD with strong regulation potential compared to similar currently available tools [62–66] (Additional file 2: Table S1). The efficiency of factR2 as an AS-NMD discovery tool is supported by its ability to identify the vast majority of the GENCODE-annotated NMD targets (99.4%). Moreover, when applied to a custom transcriptome containing both reference and novel isoforms, factR2 expanded the NMD biotype annotation by ~6-fold (43,270 vs. 7201).

factR2 analysis of TRE-Ngn2 neuronal differentiation time series identified hundreds of AS-NMD targets responding to the NMD inhibitor CHX at specific time points (Additional file 3: Table S2 and Fig. 1E). Confirming their biological relevance and authenticity, many of these events were also detected in genetically unperturbed ESCs and ex vivo neural cells (Fig. 1F, G). Interestingly, the number of CHX-responsive targets increased transiently in NPCs and developing neurons (days 3–12; Fig. 1E), suggesting that AS-NMD might be particularly important at differentiation stages associated with a large-scale rewiring of the transcriptome. The relative paucity of AS-NMD targets in mature neurons (day 24) is consistent with previously reported downregulation of the NMD factors, UPF1 and MLN51 by the neuronal microRNAs [39, 40].

Our data corroborate earlier findings that a considerable fraction of mammalian genes generate NMD-sensitive isoforms [14, 26, 63, 67, 68]. A key advancement of the present study is the identification of events with a strong potential to modulate gene expression during development. By focusing on this high-quality subset, we show that AS-NMD events, and their NS-CE subset in particular, are enriched in genes downregulated during neuronal differentiation (Fig. 2, Additional file 1: Fig. S7 and Additional file 5: Table S4). Notably, our reanalysis of an RNA-seq dataset for a functionally distinct type of neurons, dentate gyrus granule cells [69, 70], argues for a broader association between NS-CEs and gene downregulation in developing brain (Additional file 7: Supplementary Results, Additional file 1: Fig. S17, and Additional file 8: Table S6).

The enrichment of neurodevelopmentally downregulated genes is an important finding since previously characterized examples of exonic AS-NMD events have shown varied effects on their host genes [28–34]. Whether specific regulation scenarios are favored at the systems level has been an open question. In contrast, a post-transcriptional circuitry operating in developing granulocytes involves a coordinated increase in intron retention in 86 functionally related genes, dampening their expression in an NMD-dependent manner [22].

We detected 87 NS-CEs in genes monotonically downregulated in differentiating TRE-Ngn2 cells and enriched for specific biological functions (Additional file 1: Fig. S8A, Additional file 5: Table S4 and Additional file 6: Table S5). Thus, despite the obvious difference in the type of regulated splicing events, AS-NMD provides a transcriptome-wide mechanism for repressing functionally related genes in both neuronal and granulocyte development. The sequence context of the facilitating NS-CEs identified by our pipeline showed a remarkable degree of evolutionary conservation (Fig. 2E–F and Additional file 1: Fig. S17C, D). This points to the functional relevance of these events and distinguishes them, e.g., from the less conserved “cryptic” exons identified by examining transcriptome-wide effects of PTBP1/PTBP2 knockdown [61].

It should be noted that we utilized relatively lenient cutoffs (Whippet $\Delta\text{PSI} > 0.1$, probability > 0.9 , and correlation $P < 0.05$) to shortlist AS-NMD events potentially influencing gene expression in neurodevelopment (the first tab in Additional file 5: Table S4). factR2 users may consider adjusting these parameters, based on the data in Fig. 2A–B, Additional file 1: Fig. S5, and the second tab of Additional file 5: Table S4, to strike the optimal balance between capturing all regulatory events and events with the strongest regulation potential.

Our data suggest that PTBP1-controlled events are highly enriched for NS-CEs residing in neurodevelopmentally downregulated genes and that its NR-CE-containing targets upregulated in neurons (*Ptbp2*, *Gabbr1*, and *Dlg4*; [28–31]) are more of an exception rather than the rule (Additional file 1: Fig. S10). Interestingly, PTBP1-controlled NS-CEs are strongly enriched for the RHO GTPase effectors and actin cytoskeleton organization functions (Additional file 1: Fig. S8B and Additional file 6: Table S5). Another category over-represented in this set of genes relates to programmed cell death. These functions are known to play important roles in brain development and disease [71–74].

Conversely, neurodevelopmentally downregulated genes with PTBP1-independent NS-CEs are enriched for functions related to pre-mRNA splicing and other aspects of cellular RNA metabolism (Additional file 1: Fig. S8C and Additional file 6: Table S5). This is surprising, considering the previously established role of conserved NS-CEs in maintaining RNA-binding protein homeostasis through negative feedback loops [18, 20]. Understanding the mechanisms that allow such exons to downregulate RNA-associated factors in developing neurons is an intriguing question for future studies. It is possible that cross-regulation between functionally related proteins or developmental changes in protein modification or cellular localization patterns contributes to this process [19, 75].

An integral aspect of our work is the quantification of the extent to which NS-CEs modulate gene expression in developing neurons. Biallelic inactivation of the PTBP1-repressed NS-CEs in *Fmnl3*, *Iqgap1*, and *Ripk1* suggested that these *cis*-elements facilitate downregulation of their host genes during neuronal differentiation (Fig. 5). Indeed, *Fmnl3*, *Iqgap1*, and *Ripk1* lacking functional NS-CEs were expressed 4–6-fold higher compared to the wild type on differentiation days 4–6. It is also evident that NS-CEs do not function alone since some downregulation is observed even in their absence. Notably, the NMD-dependent and NMD-independent downregulation mechanisms appear to be coordinated in time with a ± 1 -day precision (Fig. 5).

Further studies will be required to understand the nature of the NMD-independent repression mechanism(s) partnered with NS-CEs. We assumed transcriptional

repression in our model (Additional file 1: Fig. S13). However, the wild-type and mutant trajectories would be expected to diverge in a similar manner if AS-NMD cooperated with NMD-independent post-transcriptional repression, e.g., dampening mRNAs abundance via microRNAs [76], RNA binding proteins [77], or nuclear retention of incompletely spliced mRNAs [78]. Exploring potential crosstalk between AS-NMD and these, as well as other processes regulating mRNA stability, nucleocytoplasmic distribution, and translational activity [79–81] in different developmental and physiological contexts may unveil novel mechanisms regulating eukaryotic gene expression.

Conclusions

We systematically elucidated the role of AS-NMD during neuronal differentiation using the newly developed bioinformatics suite factR2 and relevant experimental approaches. This revealed that NMD-stimulating cassette exons play a widespread role in promoting the downregulation of non-neuronal genes in neurons. With the increasing availability of RNA sequencing data, particularly from single-cell and long-read studies, we anticipate that the systematic use of factR2 as the transcriptome annotation tool will unveil the global impacts of AS-NMD in a variety of biological contexts.

Materials and methods

DNA constructs

Plasmids p2lox and pX330-U6-Chimeric_BB-CBh-hSpCas9 were kindly provided by Michael Kyba (Addgene plasmid #34,635; [46]) and Feng Zhang (Addgene plasmid #42,230; [82]), respectively. pEGFP-N3 was from Clontech and pGEM-T Easy was from Promega (cat# A137A). New constructs were generated as described in Additional file 9: Table S7 using routine molecular cloning techniques and enzymes from New England Biolabs. We used a Quikchange site-directed mutagenesis protocol with the KAPA HiFi DNA polymerase (Kapa Biosystems, cat# KK2101) to modify the *Fmnl3* minigene plasmid. Primers used for cloning, mutagenesis, and other purposes are listed in Additional file 10: Table S8. All constructs were verified by Sanger sequencing.

ESC culture

A2lox [46], IB10 [83], and TRE-Ngn2 (this study) mouse ESCs were cultured in a humidified incubator at 37 °C, 5% CO₂, in tissue-culture plates or dishes coated with gelatin (Millipore, cat# ES-006-B). The 2i + LIF medium [84] used for ESC maintenance contained a 1:1 mixture of Neurobasal (Thermo Fisher Scientific, cat# 21,103,049) and DMEM/F12 (Sigma, cat# D6421) media supplemented with 100 units/ml PenStrep (Thermo Fisher Scientific, cat# 15,140,122), 1 μM PD03259010 (Cambridge Bioscience, cat# SM26-2), 3 μM CHIR99021 (Cambridge Bioscience, cat# SM13-1), 0.5 mM L-Glutamine (Thermo Fisher Scientific, cat# 25,030,024), 0.1 mM β-mercaptoethanol (Sigma, cat# M3148), 1000 units/ml ESGRO LIF (Millipore, cat# ESG1107), 0.5 × B-27 supplement without vitamin A (Thermo Fisher Scientific, cat# 12,587,010) and 0.5 × N2 supplement. N2 100 × stock was prepared using DMEM/F12 medium as a base and contained 5 mg/ml BSA (Thermo Fisher Scientific, 15,260,037), 2 μg/ml progesterone (Sigma, P8783-1G), 1.6 mg/ml putrescine (Sigma, P5780), 3 μM sodium selenite solution (Sigma, S5261), 10 mg/ml apo-transferrin (Sigma, T1147), and 1 mg/ml insulin (Sigma,

I0516) and stored in single-use aliquots at -80°C . Cells were typically passaged every 2–3 days by treating the cultures with 0.05% Trypsin–EDTA (Thermo Fisher Scientific, cat# 15,400,054) for 8–10 min at 37°C . After quenching trypsin with FBS (Thermo Fisher Scientific, cat# SH30070.03E), cells were washed once with Neurobasal medium, spun down at $260 \times g$ for 4 min and plated at 1:3 to 1:6 dilution. The identity of all ESC lines was confirmed by their ability to form OCT4-positive, dome-shaped colonies in 2i + LIF medium (see the “[Immunofluorescence analyses](#)” section). A2lox cells were additionally validated by performing inducible cassette exchange reactions, and TRE-Ngn2 cells were validated by their growth in the presence of G418/geneticin (see the section below).

Production of Dox-inducible Ngn2 knock-in ESCs

A2lox cells were pre-treated overnight with 2 $\mu\text{g}/\text{ml}$ doxycycline (Dox; Sigma, cat# D9891) to activate Cre expression, trypsinized, FBS-quenched, and transfected in suspension with 1 μg of p2Lox-based plasmid pML156 (Additional file 9: Table S7) mixed with 3 μl of Lipofectamine 2000 (Thermo Fisher Scientific, cat# 11,668,019) and 100 μl of Opti-MEM I (Thermo Fisher Scientific, cat# 31,985,070). The transfection was done in 4 ml of 2i + LIF medium at $\sim 2.5 \times 10^5$ cells/ml in 6-cm bacterial dishes. Cells were pelleted 2 h post transfection, resuspended in 2 ml of fresh 2i + LIF medium, serially diluted, and replated into a gelatinized tissue-culture 6-well plate. 350 $\mu\text{g}/\text{ml}$ of G418/geneticin (Sigma, cat# 10,131,019) was added 36 h post transfection and the incubation was continued for an additional 10 days with regular medium changes to allow G418-resistant cells to form colonies. These were picked, expanded, and analyzed for Dox-inducible expression of *Ngn2* by reverse-transcriptase quantitative PCR (RT-qPCR).

Genetic inactivation of NMD-stimulating cassette exons

NS-CEs were knocked out by co-transfecting single-cell suspensions of TRE-Ngn2 cells with four plasmids encoding Cas9 and distinct gene-specific CRISPR gRNAs, as well as the pEM584 plasmid [85] encoding a blasticidin resistance gene. 1.2 μg of a mixture containing equal amounts of each of these plasmids was combined with 3 μl of Lipofectamine 2000 and 100 μl of Opti-MEM I. TRE-Ngn2 cells were transfected with the mixture in 4 ml of 2i + LIF at $\sim 2.5 \times 10^5$ cells/ml in 6-cm bacterial dishes. Cells were collected 2 h post transfection, spun down, and the cell pellets were serially diluted in fresh 2i + LIF prior to plating in 6-well format. Antibiotic selection with 8 $\mu\text{g}/\text{ml}$ blasticidin (Sigma, cat# 15,205) was started 24 h post transfection and lasted for 3 days. Cells were then cultured in 2i + LIF medium for another 6–10 days. Surviving colonies were picked, expanded for 3–4 days, and PCR-genotyped (see below). In cases where heterogeneity was suspected, the clones were dissociated to single-cell suspensions and allowed to form colonies (without blasticidin). These were picked, expanded, genotyped, and used for subsequent analyses.

Genotyping

Genomic DNA was extracted and analyzed using the PCR-BIO Rapid Extract PCR Kit (PCR Biosystems; cat# PB10.24–08) following the manufacturer’s protocol. Amplified DNA fragments were resolved by electrophoresis in 1% agarose gels alongside

GeneRuler 1 kb Plus DNA Ladder (Thermo Fisher Scientific, cat# SM1331). CRISPR/Cas9-induced mutations in AS-NMD-regulated genes were characterized with primers designed to amplify either short (~0.5 kb) or long (~2 kb) fragments centered on the targeted exon. Gel-purified PCR fragments were cloned into the pGEM-T Easy vector (Promega, cat# A137A) and sequenced using Sanger's method.

Transient transfection

Cells were counted using a hemocytometer, and $\sim 2 \times 10^5$ cells in 1 ml of 2i + LIF medium were seeded per well of a gelatinized 12-well plate and immediately transfected with 50 pmol of an appropriate siRNA (Horizon Discovery/Dharmacon; Additional file 11: Table S9) premixed with 2.5 μ l of Lipofectamine 2000 and 100 μ l of Opti-MEM I, as recommended. In experiments involving simultaneous knockdown of two genes, 100 pmol of mixed siRNAs was combined with 4 μ l of Lipofectamine 2000 and 100 μ l Opti-MEM I. Cells were exposed to siRNA-containing complexes overnight, followed by a medium change to fresh 2i + LIF next morning. RNAs were extracted 48 h post transfection. In experiments requiring CHX treatment, cells grown for 48 h following siRNA transfection were supplemented with 100 μ g/ml CHX (Sigma, cat# C4859) or an equal volume of DMSO (Sigma, cat# D2650) and incubated for an additional 6 h prior to RNA extraction. In minigene experiments, cells were transfected with 1 μ g of a corresponding plasmid mixed with 3 μ l of Lipofectamine 2000 and 100 μ l of Opti-MEM I and incubated for 24 h. Alternatively, cells were pre-treated with siRNAs for 24 h and then transfected with 1 μ g of minigene plasmid, as above, and incubated for another 24 h prior to RNA extraction.

Inducible neuronal differentiation

We have adapted previously published protocols for *Ngn2*-induced glutamatergic neuronal differentiation [44, 45] (Additional file 12: Table S10). On the day of *Ngn2* induction (day 0), cell culture plates (12-well plates were Corning™ Costar™ cat# 3513 and 6-well plates were Thermo Fisher Scientific Nunc™ cat# 140675) were coated with Geltrex™ (Gibco, cat# A1413302), diluted 1:50 in cold DMED/F12, for 2 h in a humidified incubator at 37 °C. Geltrex was aspirated and TRE-Ngn2 ESCs were immediately plated into wells, without washing, in the iN-D0 medium containing a 1:1 mixture of Neurobasal (Thermo Fisher Scientific, cat# 21103049) and DMEM/F12 (Sigma, cat# D6421) supplemented with 100 units/ml PenStrep (Thermo Fisher Scientific, cat# 15140122), 1 \times N2 (Gibco, cat# 17502048), 1 \times B27 with retinoic acid (Gibco, cat# 17504044), 1 μ g/ml laminin (Sigma, cat# L2020), 20 μ g/ml insulin (Sigma, cat# I0516), 500 μ M L-glutamine (Gibco, cat# 25030-024), 2 mM db-cAMP (Sigma, cat# D0627), 10 ng/ml NT-3 (Miltenyi Biotec, cat# 130-093-973), 100 μ M β -mercaptoethanol and 2 μ g/ml Dox. If not stated otherwise, we seeded 2×10^4 ESCs per well of a Geltrex-coated 12-well plate for subsequent RT-PCR and RT-qPCR analyses and $1-3 \times 10^3$ cells per well of a 12-well plate containing a Geltrex-coated 18-mm glass coverslip (VWR, cat# 631-1580) for immunostaining. The cells attached to the wells were then cultured in a humidified incubator at 37 °C, 5% CO₂.

On day 2, the medium was changed completely to the iN-D2 medium, similar to iN-D0 except containing 200 μ M L-ascorbic acid (Sigma, A4403) instead of β -mercaptoethanol

and 1 µg/ml instead of 2 µg/ml Dox. Starting from day 4, we replaced only ~50% of conditioned medium with fresh Dox-free medium. The iN-D4 medium used for this purpose on day 4 had the same composition as iN-D2 but lacked Dox. In short-term differentiation experiments, samples were collected on day 6 without further medium changes. In experiments where induced neurons were maintained longer than 6 days, the half-volume changes on day 6 onwards were done using the iN-D6 medium containing Neurobasal supplemented with 1 × B27 with retinoic acid, 1 × CultureOne supplement (Gibco, cat# 15674028), 500 µM L-glutamine, 1 µg/ml laminin, 100 units/ml PenStrep, 200 µM L-ascorbic acid, 10 ng/ml NT-3, 10 ng/ml BDNF (Miltenyi Biotec, cat# 130-093-811), and 2 mM db-cAMP.

RT-qPCR and RT-PCR analyses

Total RNAs were purified using the EZ-10 DNAaway RNA Miniprep Kit (BioBasic, cat# BS88136), as recommended. Reverse transcription (RT) was performed at 50 °C for 40 min using SuperScript IV reagents (Thermo Fisher Scientific, cat# 18090200), 5 µM of random decamer (N10) primers, and 2 units/µl of murine RNase inhibitor (New England Biolabs, M0314L). In minigene experiments, purified RNAs were additionally incubated with 2 units of RQ1-DNase (Promega, cat# M6101) per 1 µg of RNA at 37 °C for 30 min to eliminate plasmid DNA contamination. RQ1-DNase was inactivated by adding the Stop Solution. The RNAs were then immediately reverse-transcribed as described above.

cDNA samples were analyzed by qPCR using a Light Cycler[®]96 Real-Time PCR System (Roche) and the qPCR BIO SyGreen Master Mix (PCR Biosystems; cat# PB20.16). All RT-(q)PCR primers are listed in Additional file 10: Table S8. RT-qPCR signals were normalized to expression levels of housekeeping genes selected based on our RNA-seq data. *Cnot4* was used for longitudinal differential expression analyses. *Gars* was used to compare between CHX- and DMSO-treated samples.

To analyze NS-CE inclusion, cDNA samples were amplified using DreamTaq DNA Polymerase (Thermo Fisher Scientific cat# EP0702) or ROCHE Taq DNA polymerase (Sigma-Aldrich, cat# 11596594001). The RT-PCR products were then resolved by electrophoresis in 2% agarose gels alongside GeneRuler 1 kb Plus DNA Ladder. Band intensities were quantified using Image Studio Lite (Version 5.2) (LI-COR Biosciences). Percent spliced in (PSI) values were calculated by normalizing the intensity of the NS-CE inclusion product by combined intensity of the inclusion and skipping products and multiplying the quotient by 100.

TRE-Ngn2 samples for RNA-sequencing

For RNA-Seq, TRE-Ngn2 cells were differentiated into neurons in a 6-well plate format as described above. We plated 6–8 × 10⁴ cells per well to produce day 3, 6, 12, and 24 differentiated samples. Day 0 non-induced samples were generated by plating TRE-Ngn2 ESCs in 2i + LIF at 1 × 10⁵ per well and then lysing the culture 3 days later. All time points were lysed using TRIzol reagent (Thermo Fisher Scientific, cat# 15596026). The lysates were immediately frozen, stored at –80 °C, and processed simultaneously using the TRIzol Plus RNA Purification Kit (Thermo Fisher Scientific cat# 12183555). RNAs were eluted in nuclease-free water (Invitrogen, cat# AM9939).

Natural samples for RNA sequencing

To identify natural AS-NMD targets, cultures of wild-type (IB10) ESCs, NPC neurospheres, and primary neurons were treated with either DMSO or 100 µg/ml CHX for 6 h at 37 °C. Alternatively, we prepared nuclear and cytoplasmic fractions from untreated cells using a PARIS™ kit (Thermo Fisher Scientific, cat# AM1921).

Primary NPC cultures were obtained by dissecting embryonic day 14.5 (E14.5) mouse embryonic cortices in Hank's Balanced Salt Solution (1 × HBSS, Thermo Fisher Scientific; cat# 14025092) and dissociating them by trituration. The NPCs were then cultured as neurospheres in reconstituted NeuroCult® Proliferation Kit medium (STEMCELL Technologies, cat# 05702) supplemented with 20 ng/ml recombinant human EGF (STEMCELL Technologies, cat# 78006). For passaging, neurospheres were dissociated with NeuroCult® Chemical Dissociation Kit (STEMCELL Technologies, cat# 05707), as recommended. We also prepared adherent NPC cultures for immunofluorescence experiments by plating dissociated neurospheres to polyornithine (Sigma; cat# P3655-10MG) and fibronectin (STEMCELL Technologies, cat# 07159) coated dishes.

Primary cortical neurons were isolated from E15.5 mouse embryos and cultured as described [86]. Briefly, cortices were dissociated with 2.5% trypsin (Thermo Fisher Scientific, cat# 15090046) and plated on poly-L-lysine (Sigma, cat# P2636) coated 6-well plates or 18-mm round coverslips in Minimum Essential Media (MEM) with L-glutamine (Thermo Fisher Scientific, cat# 11095080), 0.6% glucose (Fisher Scientific, cat# 10335850) and 10% horse serum (GIBCO/Life Technologies) at 1×10^6 neurons per well and 2.5×10^4 neurons per coverslip.

Neurons in 6-well plates were cultured for 5 days without glial feeders, whereas neurons attached to coverslips were transferred to wells containing a monolayer of newborn rat astrocytes [86], and cultured in neuronal maintenance medium (MEM with L-glutamine, 0.6% glucose and 1 × Neurocult SM1 neuronal supplement (STEMCELL Technologies, cat# 05711)) for 21 days replacing half of the medium every 3–4 days.

Astrocytes for long-term neuronal cultures were prepared by dissociating newborn rat cortices with 2.5% trypsin and 1 mg/ml DNase (Merck, cat# 10104159001) and plating the cells in MEM with L-glutamine supplemented with 0.6% glucose, 10% FBS (GE Hyclone, cat# HYC85), 100 units/ml penicillin and 100 µg/ml streptomycin (Thermo Fisher Scientific, cat# 15140122) at 2×10^6 cells per 25-cm² flask [86]. The medium was changed once after 3 days, and the cultures were maintained for a total of 7 days to allow astrocytes to expand to ~90% confluence.

We isolated RNAs from the whole-cell and fractionated material using either the Purelink™ RNA mini kit (Thermo Fisher Scientific, cat# 12183025; IB10 ESCs and NPCs) or the RNeasy micro kit (Qiagen, cat# 74004; primary neurons).

RNA sequencing

Total RNA concentration was estimated using Nanodrop and measured with the Qubit RNA BR Assay Kit (Fisher Scientific, cat# Q10211). Additionally, RNA quality was evaluated using the Agilent Technologies RNA ScreenTape Assay (5067–5576 and 5067–5577).

For the TRE-Ngn2 time series, intact poly(A) RNA was purified from total RNA samples (100–500 ng) with oligo(dT) magnetic beads. Stranded mRNA sequencing libraries were then prepared as described using the Illumina TruSeq Stranded mRNA Library Prep kit (20020595) and TruSeq RNA UD Indexes (20022371). Purified libraries were qualified on an Agilent Technologies 2200 TapeStation using a D1000 ScreenTape assay (cat# 5067–5582 and 5067–5583). The molarity of adapter-modified molecules was defined by quantitative PCR using the Kapa Biosystems Kapa Library Quant Kit (cat# KK4824). Individual libraries were normalized to 1.30 nM in preparation for Illumina sequence analysis. Sequencing libraries were chemically denatured and applied to an Illumina NovaSeq flow cell using the NovaSeq XP chemistry workflow (20021664). Following the transfer of the flowcell to an Illumina NovaSeq instrument, a 2×51 cycle paired end sequence run was performed using the NovaSeq S1 reagent Kit (20027465).

For the natural samples, intact poly(A) RNA was purified from total RNA (100–500 ng) with oligo(dT) magnetic beads and stranded mRNA sequencing libraries were prepared as described using the Illumina TruSeq Stranded mRNA Library Preparation Kit (RS-122–2101, RS-122–2102). Purified libraries were qualified on an Agilent Technologies 2200 TapeStation using a D1000 ScreenTape assay (cat# 5067–5582 and cat# 5067–5583). The molarity of adapter-modified molecules was defined by quantitative PCR using the Kapa Biosystems Kapa Library Quant Kit (cat# KK4824). Individual libraries were normalized to 5 nM and equal volumes were pooled in preparation for Illumina sequence analysis. Sequencing libraries (25 pM) were chemically denatured and applied to an Illumina HiSeq v4 single-read flow cell using an Illumina cBot. Hybridized molecules were clonally amplified and annealed to sequencing primers with reagents from an Illumina HiSeq SR Cluster Kit v4-cBot (GD-401–4001). The flowcell was transferred to an Illumina HiSeq 2500 instrument (HCSv2.2.38 and RTA v1.18.61) and a 50-cycle single-read sequence run was performed using HiSeq SBS Kit v4 sequencing reagents (FC-401–4002). The library preparation and sequencing steps were performed by the High-Throughput Genomics facility at the Huntsman Cancer Institute, University of Utah, USA.

Immunofluorescence analyses

Non-induced TRE-Ngn2 ESCs were plated in 2i+LIF medium onto 18-mm gelatin-coated coverslips (VWR, cat# 631–1580) at 6×10^4 cells per well of a 12-well plate. Dox-induced TRE-Ngn2 cells were plated onto 18 mm coverslips coated with Geltrex, as described above. Cells were fixed with 4% PFA in $1 \times$ PBS for 15 min at room temperature. After 3 washes with $1 \times$ PBS, cells were permeabilized with 0.1% Triton X-100 in $1 \times$ PBS for 5 min, followed by another 3 washes with $1 \times$ PBS. The coverslips were then incubated in the blocking buffer (5% horse serum, 5% goat serum, 1% BSA in $1 \times$ PBS) for 1 h at room temperature. Blocked samples were incubated with primary antibodies (Additional file 13: Table S11) diluted in the blocking buffer at 4 °C in a humidified chamber overnight. The next day, coverslips were washed 3 times with $1 \times$ PBS and incubated with Alexa Fluor-conjugated secondary antibodies diluted 1:400 in the blocking buffer at room temperature for 1 h. Coverslips were then washed 3 times with $1 \times$ PBS and stained with 0.5 µg/ml DAPI (Thermo Fisher Scientific, cat# D1306) in $1 \times$ PBS for

5 min. Coverslips were mounted on microscope slides in a drop of ProLong Gold Anti-fade Mountant solution (Thermo Fisher Scientific, cat# P36934). Images were taken using a ZEISS Axio Observer Z1 Inverted Microscope with alpha Plan-Apochromat 100x/1.46 oil immersion objective.

Bioinformatics

RNA-seq data were quality-controlled by FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and trimmed by Trimmomatic [87]. HISAT2 [51] index and splice site files were prepared using the GRCm39 mouse genome (GRCm39.primary_assembly.genome.fa) and M26 transcriptome (gencode.vM26.primary_assembly.annotation.gtf) files from GENCODE (ftp://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_mouse/release_M26/):

```
hisat2-build -p [n_threads] GRCm39.primary_assembly.genome.fa GRCm39_genome
hisat2_extract_splice_sites.py gencode.vM26.primary_assembly.annotation.gtf > M26.ss.txt
```

We then aligned our longitudinal TRE-Ngn2 differentiation data using HISAT2:

```
hisat2 -p [n_threads] -k 50 -fr -rna-strandness RF -dta-cufflinks -no-unal \
--known-splicesite-infile M26.ss.txt \
-x GRCm39_genome \
-1 [read_1.fastq] \
-2 [read_2.fastq] \
-S [aligned.sam]
```

SAM files were converted to the BAM format and optionally concatenated, sorted, and indexed using SAMtools [88]. To produce CPM-normalized strand-specific RNA-seq coverage plots, sorted and indexed [aligned.bam] files were converted into the bedGraph format using deepTools [89]:

```
bamCoverage -p [n_threads] -bs 1 -of bedgraph \
--filterRNAstrand forward \
--effectiveGenomeSize 2654621837 --normalizeUsing CPM \
-b [aligned.bam] -o [forward.bedGraph]
```

```
bamCoverage -p [n_threads] -bs 1 -of bedgraph \
--filterRNAstrand reverse \
--effectiveGenomeSize 2654621837 --normalizeUsing CPM \
-b [aligned.bam] -o [reverse.bedGraph]
```

The resultant [forward.bedGraph] and [reverse.bedGraph] files were then visualized in IGV [90].

To identify novel transcripts enriched for AS-NMD events, we first analyzed individual stage-specific DMSO- and CHX-treated samples by StringTie [51]:


```
stringtie -p [n_threads] --rf -M 1 -u -f 0.05 \
  -G gencode.vM26.primary_assembly.annotation.gtf \
  -o [sample_specific.gtf] \
  [sample_specific.bam]
```

Different sample_specific.gtf files were then merged by cuffmerge [91]:

```
cuffmerge -p [n_threads] --min-isoform-fraction 0.05 \
  -g gencode.vM26.primary_assembly.annotation.gtf \
  [list_of_sample_specific_gtfs_to_merge.txt]
```

and novel transcripts in the resultant merged.gtf file were shortlisted by removing entries with intron chains identical to those in gencode.vM26.primary_assembly.annotation.gtf (class code "="):

```
grep -v -P "class_code \"= \";" merged.gtf > novel.gtf
```

The remaining transcripts were analyzed by GffCompare [92] to allocate them to known genes where possible:

```
gffcompare novel.gtf -r gencode.vM26.primary_assembly.annotation.gtf
```

We then used R [93] and factR2 [50] to annotate the resultant gffcmp.merged.gtf file, combine it with the gencode.vM26.primary_assembly.annotation.gtf reference and extract AS-NMD events from thus produced extended_M26.gtf:

```
library(factR2).
```

```
library(rtracklayer)
custom.gtf <- import("gffcmp.merged.gtf")
ref.gtf <- import("gencode.vM26.primary_assembly.annotation.gtf")
seqlevels(ref.gtf) <- unique(c(seqlevels(custom.gtf), seqlevels(ref.gtf)))
extended_M26.gtf <- c(ref.gtf, custom.gtf)
factRobj <- createfactRObj(extended_M26.gtf, "vM26")
factRobj <- buildCDS(factRobj)
factRobj <- predictNMD(factRobj)
factRobj <- testASNMDevents(factRobj)
factRobj <- getAScons(factRobj, type = "flanks", padding = 100)
exportGTF(factRobj, "extended_M26.gtf")
exportTable(factRobj, "extended_M26_NMD_exons.txt", "exons")
```

Transcript abundance in our TRE-Ngn2 differentiation experiment was quantified using Kallisto [94] with an extended_M26.gtf-based index:

```
kallisto quant -t [n_threads] -rf-stranded \
  -i extended_M26.index -o [out_dir] \
  [read_1.fastq] [read_2.fastq]
```

Kallisto-deduced transcript-specific counts for individual biological replicates were imported into R using the tximport package [95] and analyzed by DESeq2 [53]. Genes not represented by ≥ 5 sequencing reads in ≥ 2 samples were considered poorly expressed and excluded from subsequent analyses. DMSO and CHX data for specific differentiation time points were compared pairwise using the Wald test:

```
dds <- DESeq(dds, fitType = 'local')
```

Differentially expressed genes in DMSO time-course data were analyzed by the likelihood-ratio test:

```
dds_lrt <- DESeq(dds_lrt, test = "LRT", reduced = ~ 1)
```

Unless indicated otherwise, genes with FDR (calculated as Benjamini-Hochberg-adjusted P -value) < 0.001 were considered differentially regulated. The PCA plot in Additional file 1: Fig. S2A was produced by DESeq2's plotPCA function using VST-transformed expression values (VST_{gene}), which were calculated for all detectably expressed genes by the vst(dds, blind = FALSE) function of DESeq2. VST_{gene} values were also used for correlation and developmental trend analyses. Heatmaps in Additional file 1: Fig. S2B and Additional file 1: Fig. S10A were produced using the ComplexHeatmap R package and centered and scaled VST_{gene} values for manually selected marker genes.

Alternative splicing patterns were analyzed by Whippet with an extended_M26.gtf-based index:

```
julia whippet-quant.jl [read_1.fastq] [read_2.fastq] -biascorrect \
  -x extended_M26.whippet.graph.jls -o [out_dir]
```

Condition-specific psi.gz files produced by the above code were then compared for each differentiation stage as follows:

```
julia whippet-delta.jl -s 1 -a condition_A.psi.gz -b condition_B.psi.gz \
  -o [output_prefix]
```

The difference (Δ) between condition-specific percent spliced in values (PSI; defined in Whippet as a fraction rather than %) for CHX-treatment and nucleus-cytoplasm fractionation experiments was calculated as:

$$\Delta\text{PSI} = \text{PSI}_{\text{CHX}} - \text{PSI}_{\text{DMSO}}$$

and

$$\Delta\text{PSI} = \text{PSI}_{\text{Nucleus}} - \text{PSI}_{\text{Cytoplasm}}$$

respectively.

We used the following filters to shortlist regulated NMD-stimulating (NS) and NMD-repressing (NR) events. NS events, $\Delta\text{PSI} > 0.1$ and probability > 0.9 ; NR events, $\Delta\text{PSI} < -0.1$ and probability > 0.9 . To retain events that control mRNA sensitivity to NMD but are nearly completely included or excluded at the splicing level, we also shortlisted NS events

where PSI averaged across conditions was >0.9 , and NR events where the average was <0.1 . In some cases, the stringency of Whippet cutoffs was increased to $|\Delta\text{PSI}|>0.25$ and probability >0.95 . Additionally, regulated alternative splicing (AS) events were occasionally shortlisted by rMATS v4.1.2 [54] with $|\Delta\text{PSI}|>0.1$ and FDR <0.05 cutoffs. These instances are explicitly indicated in the text, figures, and tables. Shortlisted events from the cassette/core exon (CE), alternative donor/5' splice site (AD), alternative acceptor/3' splice site (AA), and retained intron (RI) groups were used for further analyses.

For correlation and trend analyses we used factR2's function `testGeneCorr` subjecting PSI values to arcsine-square root variance-stabilizing transformation [96] and adding the $2/\pi$ coefficient to maintain transformed PSI within the $[0, 1]$ interval:

$$\text{VST}_{\text{PSI}} = 2\arcsin(\sqrt{\text{PSI}})/\pi$$

The correlation between the AS-NMD splicing pattern and gene expression was then calculated as Pearson's product-moment correlation between VST_{gene} and VST_{PSI} for NMD-repressing events and between VST_{gene} and $1 - \text{VST}_{\text{PSI}}$ for NMD-stimulating events.

Evolutionary conservation of the intronic sequence context of AS-NMD events was analyzed using factR2's function `getASCons` and the PhastCons [55] mm39 35-vertebrate conservation track from the UCSC genome browser (<http://hgdownload.soe.ucsc.edu/goldenPath/mm39/phastCons35way/>). For each event type, we examined a 200-nt sequence, defining the context as 100-nt intronic sequences preceding and following CEs, the first and last 100 nt of RIs, 200-nt intronic sequences following ADs, and 200-nt intronic sequences preceding AAs.

PTBP1-regulated events were shortlisted by Whippet analyses of ESCs and NPCs, where PTBP1 was knocked down alone or together with its functionally similar paralog, PTBP2 ($|\Delta\text{PSI}|>0.1$; probability >0.9 cutoffs). If the PSI values changed significantly in response to multiple treatments, individual PTBP1 knockdown experiments were given priority over combined PTBP1 and PTBP2 knockdowns. Moreover, ESCs were prioritized over NPCs. We additionally requested that the event's PSI correlates with *Ptbp1* expression levels (Pearson's $P < 0.05$) in our TRE-Ngn2 RNA-seq differentiation dataset and that the sign of the correlation coefficient (Pearson's r) is opposite to the sign of ΔPSI in response to PTBP1 knockdown in ESCs or NPCs.

We estimated the proportions of different cell types in differentiating TRE-Ngn2 cultures using the R package MuSiC [48]. Raw gene expression counts from our bulk RNA-seq dataset were compared with mouse single-cell RNA-seq data for ESCs (GSM2098554), NPCs (GSE67833; [97]), and adult cortices (GSE185862; [98]). We first converted count matrices to the ExpressionSet object format:

```
library(Biobase)
metadata <- data.frame(row.names = colnames(counts), group = colnames(counts))
metalabels <- data.frame(labelDescription = c("group"))
metadata <- AnnotatedDataFrame(data = metadata, varMetadata = metalabels)
bulk.eset <- ExpressionSet(counts, phenoData = metadata)
```

and then performed cellular deconvolution using MuSiC's `music_prop` function. The output data frame was processed and visualized in `ggplot2` [99].

Genes were clustered according to their neurodevelopmental expression trajectories using DP_GP_cluster [56] as follows:

```
DP_GP_cluster.py -i [expression.txt] -o [output_prefix] \
  -fast -n 1000 -max_iters 1000 \
  -check_burnin_convergence -check_convergence \
  -cluster_uncertainty_estimate -plot -p pdf
```

where expression.txt was a table with gene IDs and averages of VST_{gene} values for each differentiation time point. The optimal_clustering.txt output file was used for further analyses.

To estimate AS-NMD contributions to gene downregulation in development, we modeled changes in mRNA abundance as a function of time ($R(t)$) depending on mRNA synthesis rate (v_{sr}) and the decay constant (k_{dr}) related to mRNA half-life ($t_{1/2}$) [100, 101] as follows:

$$k_{dr} = \frac{\ln 2}{t_{1/2}}$$

For simplicity, we assumed that the rate at which productively spliced mRNA is synthesized can decrease through the onset of transcriptional repression (r_t) or/and an increase in AS-NMD repression (r_n) at specific time points within the day 2–4 period, whereas the k_{dr} remains constant. We also assumed that NMD-sensitive isoforms are much less stable than productively spliced mRNA, and therefore, ignored their contribution to $R(t)$. This produced the following equation, which was solved using the ode45 function or the R package pracma:

$$\frac{dR(t)}{dt} = r_t r_n v_{sr} - k_{dr} R(t)$$

In the model described in Additional file 1: Fig. S13, we used published transcriptome-wide estimates for v_{sr} (1.76 mRNA molecules per cell per hour; the median rate measured for mouse NIH 3T3 cells; [101]) and $t_{1/2}$ (7.08 h; the median half life for proliferating and differentiating mouse ESCs; [102]). However, similar results were obtained when we estimated the v_{sr} and $t_{1/2}$ parameters by averaging the *Fmnl3*-, *Iqgap1*-, and *Ripk1*-specific values ($v_{sr}=2.63$ mRNA molecules per cell per hour in NIH 3T3 cells and $t_{1/2} = 5.11$ h in proliferating (LIF+) and differentiating (RA+) mouse ESCs). r_t was set to 1 in early development. To estimate r_t following the onset of repression in developing neurons, we averaged the ratios between expression levels on days 5–6 and days 0–1 for the *Fmnl3*, *Iqgap1*, and *Ripk1* mutants in Fig. 5A–C. This resulted in $r_t = 0.39$. We estimated the initial (stem cell-specific) and the final (neuron-specific) r_n values from the PSI_{CHX} data (Fig. 4B):

$$r_n = 1 - \frac{PSI_{\text{CHX}}}{100}$$

by averaging day-0–1 and day-5–6 *Fmnl3*, *Iqgap1*, and *Ripk1* PSI_{CHX} values, respectively. This yielded the initial $r_n = 0.78$ and the final $r_n = 0.12$.

Single-cell RNA-sequencing data from developing mouse dentate gyrus were downloaded from <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE95753>. A total of 2000 FASTQ files from the granule cell lineage were sampled and combined into four

groups to represent the transcriptomes of the four distinct cell subpopulations: radial glia-like cells, neuronal intermediate progenitors, neuroblasts, and granule cells. Pooled FASTQ files were aligned to the mm39 genome by HISAT2, and custom transcriptomes were constructed using StringTie2, as described above. Group-specific GTF files were then merged:

```
stringtie --merge -F 0.5 -T 0.5 -G gencode.vM26.primary_assembly.annotation.gtf -o stringtie.merged.gtf [* .gtf]
```

The `stringtie.merged.gtf` file was processed by `factR2` to annotate AS-NMD events and score the interspecies conservation of intronic regions flanking NS-CEs. Transcript abundances and splicing patterns in individual cells were analyzed by Kallisto and Whippet, respectively, with `stringtie.merged.gtf`-based indexes. In this analysis, high-quality NS-CEs were identified as events predicted by `factR2` to stimulate NMD and showing a correlation (Pearson's $r > 0.2$ and $P < 0.05$) between NMD-protective splicing and gene expression after applying appropriate variance-stabilizing transformations. Cassette exons that were not predicted to be NMD-inducing by `factR2` were used as a control group.

Statistical analyses

Unless stated otherwise, all statistical procedures were performed in *R*. We repeated experiments at least three times and showed the results as averages \pm SD. Data obtained from RT-(q)PCR quantifications were typically analyzed using a two-tailed Student's *t*-test assuming unequal variances. Other types of data were compared using a two-tailed Wilcoxon rank sum test, one-sided Kolmogorov–Smirnov (KS) test, or Fisher's exact test. Where necessary, *p*-values were adjusted for multiple testing using Benjamini–Hochberg correction (FDR). Correlation analyses were done using Pearson's product-moment and Kendall's rank correlation methods, as specified in the text. Numbers of experimental replicates, *P*-values, and the tests used are indicated in the Figures and/or Figure legends.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-024-03305-8>.

Additional file 1. Supplementary Figures.

Additional file 2. Table S1. Key functions of currently available mRNA isoform/AS-NMD analysis programs.

Additional file 3. Table S2. CHX-responsive AS-NMD events.

Additional file 4. Table S3. Ranked gene trajectory clusters.

Additional file 5. Table S4. Facilitating AS-NMD events in the TRE-Ngn2 differentiation dataset showing significant positive correlation between NMD-protective splicing patterns and gene expression with $P < 0.05$ and FDR < 0.1 .

Additional file 6. Table S5. Metascape results.

Additional file 7. Supplementary Results.

Additional file 8. Table S6. Facilitating AS-NMD events in developing dentate gyrus granule cells.

Additional file 9. Table S7. Plasmids generated in this study.

Additional file 10. Table S8. DNA oligonucleotides used in this study.

Additional file 11. Table S9. siRNAs used in this work.

Additional file 12. Table S10. Media for induced neuronal differentiation of TRE-Ngn2 cells.

Additional file 13. Table S11. Antibodies used in this work.

Additional file 14. Peer review history.

Acknowledgements

We thank Snezhka Olfiferenko for valuable discussions, Brian Dalley for the help with RNA sequencing, and Michael Kyba and Feng Zhang, for reagents.

Review history

The review history is available as Additional file 14.

Peer review information

Tim Sands was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Authors' contributions

AZ performed the TRE-Ngn2 experiments, analyzed and interpreted the data, and contributed to writing the manuscript; KY performed the experiments with wild-type ESCs and neural cortical cells and analyzed and interpreted the data; FH developed factR2 and performed bioinformatics analyses; EVM obtained funding for this project, performed bioinformatics analyses and was a major contributor to the writing of the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the Biotechnology and Biological Sciences Research Council (BB/M001199/1, BB/M007103/1, and BB/V006258/1).

Availability of data and materials

The RNA-seq data generated and analyzed in this study are available from BioStudies under the accession number E-MTAB-13134 [103]. The source code and a detailed user manual for factR2 are available on GitHub [50]. Details of the key bioinformatics procedures used in our work are also described on GitHub [104]. Uncropped versions of our gel and microscopy images have been published on Figshare (<https://doi.org/10.6084/m9.figshare.25943524.v1>).

Declarations

Ethics approval and consent to participate

All animal procedures were approved by the Institutional Animal Care and Use Committee and the Home Office (Project Licence PP5451667).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 8 July 2023 Accepted: 7 June 2024

Published online: 20 June 2024

References

1. Baralle FE, Giudice J. Alternative splicing as a regulator of development and tissue identity. *Nat Rev Mol Cell Biol.* 2017;18:437–51.
2. Ule J, Blencowe BJ. Alternative splicing regulatory networks: functions, mechanisms, and evolution. *Mol Cell.* 2019;76:329–45.
3. Marasco LE, Kornblihtt AR. The physiology of alternative splicing. *Nat Rev Mol Cell Biol.* 2023;24:242–54.
4. Wright CJ, Smith CWJ, Jiggins CD. Alternative splicing as a source of phenotypic diversity. *Nat Rev Genet.* 2022;23:697–710.
5. Raj B, Blencowe BJ. Alternative splicing in the mammalian nervous system: recent insights into mechanisms and functional roles. *Neuron.* 2015;87:14–27.
6. Vuong CK, Black DL, Zheng S. The neurogenetics of alternative splicing. *Nat Rev Neurosci.* 2016;17:265–81.
7. Mazin PV, Khaitovich P, Cardoso-Moreira M, Kaessmann H. Alternative splicing during mammalian organ development. *Nat Genet.* 2021;53:925–34.
8. Merkin J, Russell C, Chen P, Burge CB. Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. *Science.* 2012;338:1593–9.
9. Garcia-Moreno JF, Romao L. Perspective in alternative splicing coupled to nonsense-mediated mRNA decay. *Int J Mol Sci.* 2020;21:9424.
10. Han X, Wei Y, Wang H, Wang F, Ju Z, Li T. Nonsense-mediated mRNA decay: a “nonsense” pathway makes sense in stem cell biology. *Nucleic Acids Res.* 2018;46:1038–51.
11. Mockenhaupt S, Makeyev EV. Non-coding functions of alternative pre-mRNA splicing in development. *Semin Cell Dev Biol.* 2015;47–48:32–9.
12. Hug N, Longman D, Caceres JF. Mechanism and regulation of the nonsense-mediated decay pathway. *Nucleic Acids Res.* 2016;44:1483–95.
13. Kishor A, Fritz SE, Hogg JR. Nonsense-mediated mRNA decay: The challenge of telling right from wrong in a complex transcriptome. *Wiley Interdiscip Rev RNA.* 2019;10:e1548.
14. Kurosaki T, Popp MW, Maquat LE. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat Rev Mol Cell Biol.* 2019;20:406–20.

15. Lindeboom RGH, Vermeulen M, Lehner B, Supek F. The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. *Nat Genet.* 2019;51:1645–51.
16. Karousis ED, Muhlemann O. The broader sense of nonsense. *Trends Biochem Sci.* 2022;47:921–35.
17. Wollerton MC, Gooding C, Wagner EJ, Garcia-Blanco MA, Smith CW. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Mol Cell.* 2004;13:91–100.
18. Lareau LF, Inada M, Green RE, Wengrod JC, Brenner SE. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature.* 2007;446:926–9.
19. Leclair NK, Brugiolo M, Urbanski L, Lawson SC, Thakar K, Yurieva M, George J, Hinson JT, Cheng A, Graveley BR, Anczukow O. Poison Exon Splicing Regulates a Coordinated Network of SR Protein Expression during Differentiation and Tumorigenesis. *Mol Cell.* 2020;80(648–665):e649.
20. Muller-McNicoll M, Rossbach O, Hui J, Medenbach J. Auto-regulatory feedback by RNA-binding proteins. *J Mol Cell Biol.* 2019;11:930–9.
21. Pervouchine D, Popov Y, Berry A, Borsari B, Frankish A, Guigo R. Integrative transcriptomic analysis suggests new autoregulatory splicing events coupled with nonsense-mediated mRNA decay. *Nucleic Acids Res.* 2019;47:5293–306.
22. Wong JJ, Ritchie W, Ebner OA, Selbach M, Wong JW, Huang Y, Gao D, Pinello N, Gonzalez M, Baidya K, et al. Orchestrated intron retention regulates normal granulocyte differentiation. *Cell.* 2013;154:583–95.
23. Braunschweig U, Barbosa-Morais NL, Pan Q, Nachman EN, Alipanahi B, Gonatopoulos-Pournatzis T, Frey B, Irimia M, Blencowe BJ. Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res.* 2014;24:1774–86.
24. Jacob AG, Smith CWJ. Intron retention as a component of regulated gene expression programs. *Hum Genet.* 2017;136:1043–57.
25. Iannone C, Kainov Y, Zhuravskaya A, Hamid F, Nojima T, Makeyev EV. PTBP1-activated co-transcriptional splicing controls epigenetic status of pluripotent stem cells. *Mol Cell.* 2023;83(203–218):e209.
26. Yan Q, Weyn-Vanhentenryck SM, Wu J, Sloan SA, Zhang Y, Chen K, Wu JQ, Barres BA, Zhang C. Systematic discovery of regulated and conserved alternative exons in the mammalian brain reveals NMD modulating chromatin regulators. *Proc Natl Acad Sci U S A.* 2015;112:3445–50.
27. Mironov A, Petrova M, Margasyuk S, Vlasenok M, Mironov AA, Skvortsov D, Pervouchine DD. Tissue-specific regulation of gene expression via unproductive splicing. *Nucleic Acids Res.* 2023;51:3055–66.
28. Boutz PL, Stoilov P, Li Q, Lin CH, Chawla G, Ostrow K, Shiue L, Ares M Jr, Black DL. A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes Dev.* 2007;21:1636–52.
29. Makeyev EV, Zhang J, Carrasco MA, Maniatis T. The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol Cell.* 2007;27:435–48.
30. Spellman R, Llorian M, Smith CW. Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1. *Mol Cell.* 2007;27:420–34.
31. Zheng S, Gray EE, Chawla G, Porse BT, O'Dell TJ, Black DL. PSD-95 is post-transcriptionally repressed during early neural development by PTBP1 and PTBP2. *Nat Neurosci.* 2012;15(381–388):S381.
32. Lin L, Zhang M, Stoilov P, Chen L, Zheng S. Developmental Attenuation of Neuronal Apoptosis by Neural-Specific Splicing of Bak1 Microexon. *Neuron.* 2020;107(1180–1196):e1188.
33. Zhang X, Chen MH, Wu X, Kodani A, Fan J, Doan R, Ozawa M, Ma J, Yoshida N, Reiter JF, et al. Cell-Type-Specific Alternative Splicing Governs Cell Fate in the Developing Cerebral Cortex. *Cell.* 2016;166(1147–1162):e1115.
34. Hamid FM, Makeyev EV. Regulation of mRNA abundance by polypyrimidine tract-binding protein-controlled alternate 5' splice site choice. *PLoS Genet.* 2014;10:e1004771.
35. Addington AM, Gauthier J, Piton A, Hamdan FF, Raymond A, Gogtay N, Miller R, Tossell J, Bakalar J, Inoff-Germain G, et al. A novel frameshift mutation in UPF3B identified in brothers affected with childhood onset schizophrenia and autism spectrum disorders. *Mol Psychiatry.* 2011;16:238–9.
36. Laumonnier F, Shoubbridge C, Antar C, Nguyen LS, Van Esch H, Kleefstra T, Briault S, Frys JP, Hamel B, Chelly J, et al. Mutations of the UPF3B gene, which encodes a protein widely expressed in neurons, are associated with nonspecific mental retardation with or without autism. *Mol Psychiatry.* 2010;15:767–76.
37. Harekaki T, Sridharan J, Dvora S, Weinstein DC. Regulation of vertebrate embryogenesis by the exon junction complex core component Eif4a3. *Dev Dyn.* 2010;239:1977–87.
38. Wittkopp N, Huntzinger E, Weiler C, Sauliere J, Schmidt S, Sonawane M, Izaurralde E. Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. *Mol Cell Biol.* 2009;29:3517–28.
39. Bruno IG, Karam R, Huang L, Bhardwaj A, Lou CH, Shum EY, Song HW, Corbett MA, Gifford WD, Gecz J, et al. Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay. *Mol Cell.* 2011;42:500–10.
40. Lou CH, Shao A, Shum EY, Espinoza JL, Huang L, Karam R, Wilkinson MF. Posttranscriptional control of the stem cell and neurogenic programs by the nonsense-mediated RNA decay pathway. *Cell Rep.* 2014;6:748–64.
41. Colak D, Ji SJ, Porse BT, Jaffrey SR. Regulation of axon guidance by compartmentalized nonsense-mediated mRNA decay. *Cell.* 2013;153:1252–65.
42. Giorgi C, Yeo GW, Stone ME, Katz DB, Burge C, Turrigiano G, Moore MJ. The EJC factor eIF4AIII modulates synaptic strength and neuronal protein expression. *Cell.* 2007;130:179–91.
43. Eom T, Zhang C, Wang H, Lay K, Fak J, Noebels JL, Darnell RB. NOVA-dependent regulation of cryptic NMD exons controls synaptic protein levels after seizure. *Elife.* 2013;2:e00178.
44. Thoma EC, Wischmeyer E, Offen N, Maurus K, Siren AL, Scharl M, Wagner TU. Ectopic expression of neurogenin 2 alone is sufficient to induce differentiation of embryonic stem cells into mature neurons. *PLoS ONE.* 2012;7:e38651.
45. Hulme AJ, Maksour S, St-Clair Glover M, Mielliet S, Dottori M. Making neurons, made easy: the use of Neurogenin-2 in neuronal differentiation. *Stem Cell Reports.* 2022;17:14–34.
46. Iacovino M, Bosnakovski D, Fey H, Rux D, Bajwa G, Mahen E, Mitanoska A, Xu Z, Kyba M. Inducible cassette exchange: a rapid and efficient system enabling conditional gene expression in embryonic stem and primary cells. *Stem Cells.* 2011;29:1580–8.

47. Mulas C, Kalkan T, von Meyenn F, Leitch HG, Nichols J, Smith A. Defined conditions for propagation and manipulation of mouse embryonic stem cells. *Development*. 2019;146:dev173146.
48. Wang X, Park J, Susztak K, Zhang NR, Li M. Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nat Commun*. 2019;10:380.
49. Hamid F, Alasoo K, Vilo J, Makeyev E. Functional Annotation of Custom Transcriptomes. *Methods Mol Biol*. 2022;2537:149–72.
50. Hamid F. factR v.2: functional annotation of custom transcriptomes in R. GitHub; <https://github.com/f-hamidlab/factR2>; 2024.
51. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc*. 2016;11:1650–67.
52. Sterne-Weiler T, Weatheritt RJ, Best AJ, Ha KCH, Blencowe BJ. Efficient and accurate quantitative profiling of alternative splicing patterns of any complexity on a laptop. *Mol Cell*. 2018;72(187–200):e186.
53. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
54. Shen S, Park JW, Lu ZX, Lin L, Henry MD, Wu YN, Zhou Q, Xing Y. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci U S A*. 2014;111:E5593–5601.
55. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res*. 2005;15:1034–50.
56. McDowell IC, Manandhar D, Vockley CM, Schmid AK, Reddy TE, Engelhardt BE. Clustering gene expression time series data using an infinite Gaussian process mixture model. *PLoS Comput Biol*. 2018;14:e1005896.
57. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10:1523.
58. Linares AJ, Lin CH, Damianov A, Adams KL, Novitsch BG, Black DL. The splicing regulator PTBP1 controls the activity of the transcription factor Pbx1 during neuronal differentiation. *Elife*. 2015;4:e09268.
59. Callow MG, Watanabe C, Wickliffe KE, Bainer R, Kummerfield S, Weng J, Cuellar T, Janakiraman V, Chen H, Chih B, et al. CRISPR whole-genome screening identifies new necroptosis regulators and RIPK1 alternative splicing. *Cell Death Dis*. 2018;9:261.
60. Hensel JA, Nicholas SE, Kimble AL, Nagpal AS, Omar OMF, Tyburski JD, Jellison ER, Menoret A, Ozawa M, Rodriguez-Oquendo A, et al. Splice factor polypyrimidine tract-binding protein 1 (Ptbp1) primes endothelial inflammation in atherogenic disturbed flow conditions. *Proc Natl Acad Sci U S A*. 2022;119:e2122227119.
61. Ling JP, Chhabra R, Merran JD, Schaugency PM, Wheelan SJ, Corden JL, Wong PC. PTBP1 and PTBP2 repress nonconserved cryptic exons. *Cell Rep*. 2016;17:104–13.
62. Vitting-Seerup K, Porse BT, Sandelin A, Waage J. spliceR: an R package for classification of alternative splicing and prediction of coding potential from RNA-seq data. *BMC Bioinformatics*. 2014;15:81.
63. Weischenfeldt J, Waage J, Tian G, Zhao J, Damgaard I, Jakobsen JS, Kristiansen K, Krogh A, Wang J, Porse BT. Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns. *Genome Biol*. 2012;13:R35.
64. Hsu MK, Lin HY, Chen FC. NMD Classifier: a reliable and systematic classification tool for nonsense-mediated decay events. *PLoS ONE*. 2017;12:e0174798.
65. Vitting-Seerup K, Sandelin A. IsoformSwitchAnalyzeR: analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. *Bioinformatics*. 2019;35:4469–71.
66. Entizne JC, Guo W, Calixto CPG, Spensley M, Tzioutziou N, Zhang R, Brown JWS. TranSuite: a software suite for accurate translation and characterization of transcripts. *bioRxiv* 2020:2020.2012.2015.422989.
67. Lewis BP, Green RE, Brenner SE. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci U S A*. 2003;100:189–92.
68. Hurt JA, Robertson AD, Burge CB. Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. *Genome Res*. 2013;23:1636–50.
69. Hochgerner H, Zeisel A, Lonnerberg P, Linnarsson S. Conserved properties of dentate gyrus neurogenesis across postnatal development revealed by single-cell RNA sequencing. *Nat Neurosci*. 2018;21:290–9.
70. Munster-Wandowski A, Gomez-Lira G, Gutierrez R. Mixed neurotransmission in the hippocampal mossy fibers. *Front Cell Neurosci*. 2013;7:210.
71. Colgan LA, Yasuda R. Plasticity of dendritic spines: subcompartmentalization of signaling. *Annu Rev Physiol*. 2014;76:365–85.
72. Govek EE, Newey SE, Van Aelst L. The role of the Rho GTPases in neuronal development. *Genes Dev*. 2005;19:1–49.
73. Yamaguchi Y, Miura M. Programmed cell death in neurodevelopment. *Dev Cell*. 2015;32:478–90.
74. Yuan J, Amin P, Ofengeim D. Necroptosis and RIPK1-mediated neuroinflammation in CNS diseases. *Nat Rev Neurosci*. 2019;20:19–33.
75. Zhou Z, Fu XD. Regulation of splicing by SR proteins and SR protein-specific kinases. *Chromosoma*. 2013;122:191–207.
76. Makeyev EV, Maniatis T. Multilevel regulation of gene expression by microRNAs. *Science*. 2008;319:1789–90.
77. Dai W, Li W, Hoque M, Li Z, Tian B, Makeyev EV. A post-transcriptional mechanism pacing expression of neural genes with precursor cell differentiation status. *Nat Commun*. 2015;6:7576.
78. Yap K, Lim ZQ, Khandelia P, Friedman B, Makeyev EV. Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. *Genes Dev*. 2012;26:1209–23.
79. Bresson S, Tollervey D. Surveillance-ready transcription: nuclear RNA decay as a default fate. *Open Biol*. 2018;8:170270.
80. Khong A, Parker R. The landscape of eukaryotic mRNPs. *RNA*. 2020;26:229–39.
81. Labno A, Tomecki R, Dziembowski A. Cytoplasmic RNA decay pathways - Enzymes and mechanisms. *Biochim Biophys Acta*. 2016;1863:3125–47.
82. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339:819–23.
83. Robanus-Maandag E, Dekker M, van der Valk M, Carozza ML, Jeanny JC, Dannenberg JH, Berns A, te Riele H. p107 is a suppressor of retinoblastoma development in pRb-deficient mice. *Genes Dev*. 1998;12:1599–609.

84. Martello G, Smith A. The nature of embryonic stem cells. *Annu Rev Cell Dev Biol.* 2014;30:647–75.
85. Khandelia P, Yap K, Makeyev EV. Streamlined platform for short hairpin RNA interference and transgenesis in cultured mammalian cells. *Proc Natl Acad Sci U S A.* 2011;108:12799–804.
86. Kaech S, Banker G. Culturing hippocampal neurons. *Nat Protoc.* 2006;1:2406–15.
87. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114–20.
88. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25:2078–9.
89. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 2016;44:W160–165.
90. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nat Biotechnol.* 2011;29:24–6.
91. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* 2010;28:511–5.
92. Pertea G, Pertea M. GFF Utilities: GffRead and GffCompare. *F1000Res.* 2020;9:ISCB Comm J-304.
93. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2021.
94. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol.* 2016;34:525–7.
95. Sonesson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* 2015;4:1521.
96. Faigenbloom L, Rubinstein ND, Kloog Y, Mayrose I, Pupko T, Stein R. Regulation of alternative splicing at the single-cell level. *Mol Syst Biol.* 2015;11:845.
97. Llorens-Bobadilla E, Zhao S, Baser A, Saiz-Castro G, Zwadlo K, Martin-Villalba A. Single-cell transcriptomics reveals a population of dormant neural stem cells that become activated upon brain injury. *Cell Stem Cell.* 2015;17:329–40.
98. Yao Z, van Velthoven CTJ, Nguyen TN, Goldy J, Sedeno-Cortes AE, Baftizadeh F, Bertagnolli D, Casper T, Chiang M, Crichton K, et al. A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell.* 2021;184(3222–3241):e3226.
99. Wickham H. *ggplot2: elegant graphics for data analysis.* New York: Springer-Verlag; 2016.
100. Ben-Tabou de-Leon S, Davidson EH. Modeling the dynamics of transcriptional gene regulatory networks for animal development. *Dev Biol.* 2009;325:317–28.
101. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. Global quantification of mammalian gene expression control. *Nature.* 2011;473:337–42.
102. Sharova LV, Sharov AA, Nedorezov T, Piao Y, Shaik N, Ko MS. Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. *DNA Res.* 2009;16:45–58.
103. Makeyev EV: RNA-seq analysis of AS-NMD targets in developing mouse neurons. *BioStudies*; <https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-13134>; 2024.
104. Hamid F. Scripts for analysis of alternative splicing and nonsense-mediated decay in developing mouse neurons. *GitHub*; <https://github.com/f-hamidlab/Zhuravskaya-ASNMD>; 2024.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.