Rbfox2 controls autoregulation in RNA-binding protein networks

Mohini Jangi,[1,2] Paul L. Boutz,[1] Prakriti Paul,[3] and Phillip A. Sharp[1,2,4]

[1]David H. Koch Institute for Integrative Cancer Research, [2]Department of Biology, [3]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The tight regulation of splicing networks is critical for organismal development. To maintain robust splicing patterns, many splicing factors autoregulate their expression through alternative splicing-coupled nonsense-mediated decay (AS-NMD). However, as negative autoregulation results in a self-limiting window of splicing factor expression, it is unknown how variations in steady-state protein levels can arise in different physiological contexts. Here, we demonstrate that Rbfox2 cross-regulates AS-NMD events within RNA-binding proteins to alter their expression. Using individual nucleotide-resolution cross-linking immunoprecipitation coupled to high-throughput sequencing (iCLIP) and mRNA sequencing, we identified >200 AS-NMD splicing events that are bound by Rbfox2 in mouse embryonic stem cells. These "silent" events are characterized by minimal apparent splicing changes but appreciable changes in gene expression upon Rbfox2 knockdown due to degradation of the NMD-inducing isoform. Nearly 70 of these AS-NMD events fall within genes encoding RNA-binding proteins, many of which are autoregulated. As with the coding splicing events that we found to be regulated by Rbfox2, silent splicing events are evolutionarily conserved and frequently contain the Rbfox2 consensus UGCAUG. Our findings uncover an unexpectedly broad and multilayer regulatory network controlled by Rbfox2 and offer an explanation for how autoregulatory splicing networks are tuned.

[*Keywords*: RNA-binding protein; Rbfox; alternative splicing; embryonic stem cell; iCLIP; nonsense-mediated decay]

Supplemental material is available for this article.

Received December 2, 2013; revised version accepted January 30, 2014.

Alternative pre-mRNA splicing is essential for the context-specific diversity of metazoan transcriptomes (Nilsen and Graveley 2010). Frequently, alternative exon usage leads to variations in protein-coding sequence and protein function. Alternative splicing can also impact gene expression post-transcriptionally through the nonsense-mediated mRNA decay (NMD) pathway (Lewis et al. 2003; Weischenfeldt et al. 2012). In a given cellular condition, splicing patterns are dictated by the precise balance of expressed *trans*-acting splicing factors. Most of these splicing factors recognize short sequence motifs, allowing them to regulate large, interconnected splicing networks (Irimia and Blencowe 2012). In spite of this apparent redundancy, altering the levels of an individual splicing factor can drastically reprogram splicing patterns in both normal development and disease (Yoshida et al. 2011; Han et al. 2013). It remains unclear how RNA-binding protein (RBP) expression is tightly regulated to direct these splicing programs in some cases and to prevent their aberrant activation in others.

Alternative splicing is often coupled to NMD, a post-transcriptional regulatory pathway that prevents the production of truncated proteins from transcripts containing premature termination codons (PTCs) (Kervestin and Jacobson 2012). Primarily during the pioneering round of translation, the presence of a termination codon located >50 nucleotides (nt) upstream of the last exon–exon junction triggers transcript degradation by the NMD machinery (Kervestin and Jacobson 2012). In addition to working as a quality control mechanism to avert errors in splicing and transcription, NMD can regulate gene expression when coupled to alternative splicing (AS-NMD) (Lewis et al. 2003). This can occur, for example, through the regulated inclusion of a "poison" alternative exon containing a PTC. Intriguingly, AS-NMD events are enriched among members of the SR and hnRNP families of splicing regulators, where they occur within regions of high phylogenetic conservation (Lareau et al. 2007; Ni et al. 2007). Many of these proteins and

[4]Corresponding author
E-mail sharppa@mit.edu
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.235770.113.

other core spliceosomal components bind their own transcripts and catalyze splicing of the nonproductive NMD isoform, resulting in an autoregulatory negative feedback loop that establishes a self-limiting range of protein expression (Lareau et al. 2007; Ni et al. 2007; Saltzman et al. 2011). However, how such autoregulation is modulated to allow for different steady-state protein concentrations is unknown. In addition, as global detection of AS-NMD events is hindered by the inherent instability of PTC-containing transcripts, the extent to which AS-NMD regulates gene expression has not been determined.

The Rbfox RBPs are important tissue-specific and signal-responsive alternative splicing regulators that function through binding (U)GCAUG motifs (Kuroyanagi 2009). Unlike its paralogs, Rbfox1 (Fox-1 or A2BP1) and Rbfox3 (Fox-3, HRNBP3, or NeuN), which are predominantly expressed in muscle and neuronal tissues, Rbfox2 (Fox-2 or RBM9) is the sole expressed Rbfox family member in embryonic stem cells (ESCs) and a number of epithelial and mesenchymal cancer cell lines (Jin et al. 2003; Nakahata and Kawamoto 2005; Lee et al. 2009; Venables et al. 2009; Warzecha et al. 2009). Rbfox2 has been implicated in coordinately regulating a mesenchymal-specific splicing network in cell culture models of tumorigenesis and mesodermal differentiation (Venables et al. 2009; Shapiro et al. 2011; Venables et al. 2013), is required for viability of human ESCs (Yeo et al. 2009), and is essential for normal cerebellar development in mice (Gehman et al. 2012). These observations suggest a broad role for Rbfox proteins and, in particular, the more ubiquitously expressed Rbfox2 throughout many stages of development.

Bioinformatic analyses using the high-information consensus motif recognized by the Rbfox proteins have predicted many target splicing events (Auweter et al. 2006; Wang et al. 2008; Zhang et al. 2008). Rbfox motifs are enriched upstream of tissue-specific repressed exons and downstream from enhanced exons (Wang et al. 2008; Zhang et al. 2008; Barash et al. 2010; Merkin et al. 2012; Ray et al. 2013), and RBFOX2 cross-linking immunoprecipitation (CLIP) sequencing (CLIP-seq) in human ESCs supports this context-dependent regulation (Yeo et al. 2009). While Rbfox2 recognizes its motif with high fidelity in vitro (Auweter et al. 2006), only one-third of in vivo binding events have been attributed to recognition of the minimal GCAUG motif (Yeo et al. 2009). Equally puzzling, many UGCAUG elements in transcribed regions show no evidence of binding. Despite these computational and experimental investigations, the explanation for why only a subset of motifs appears to be selected to regulate splicing in a given cellular context remains unknown. The lack of comprehensive Rbfox2-binding sites and Rbfox2-dependent gene expression and splicing changes within the same cell type has precluded a thorough examination of these questions.

Here, we generated a global analysis of Rbfox2 splicing regulation combined with a highly specific, single-nucleotide-resolution Rbfox2 RNA-binding map in mouse ESCs (mESCs), where Rbfox2 is the only expressed Rbfox paralog. We found that Rbfox2 regulates the splicing and expression of many previously unknown targets and particularly a number of RBPs by modulating AS-NMD. This demonstrates a novel advantage of CLIP-seq data in uncovering "silent" alternative splicing events, for which the instability of NMD target transcripts dampens true changes in isoform ratios. Based on our observations of RBP–Rbfox2 coregulation with a polarity predicted by Rbfox2 binding, we propose a model in which Rbfox2 tunes autoregulatory splicing events to control RBP expression levels and in turn alter their respective splicing networks. Our results position Rbfox2 at a critical node in a broader splicing regulatory network, with roles in both normal development and disease settings.

Results

Rbfox2 individual nucleotide-resolution CLIP (iCLIP) is enriched in introns and alternatively spliced regions

To determine the binding sites of Rbfox2 across the mESC transcriptome, we performed a modified iCLIP protocol optimized to provide nucleotide-level resolution and high specificity. We stably transduced V6.5 mESCs with a lentiviral vector carrying Flag-HA-tagged or untagged human RBFOX2 (FHFOX2 or UntagFOX2) expressed from a doxycycline-inducible promoter. Titration of doxycycline allowed for expression of the transgene at levels similar to endogenous Rbfox2, and both constructs were functional in regulating the splicing of candidate Rbfox2 targets (Supplemental Fig. S1A,B). After generating iCLIP libraries by sequential Flag and HA immunoprecipitation from cells expressing either FHFOX2 or UntagFOX2 as a background control (Supplemental Fig. S1C,D), reads were mapped to the genome, and single-nucleotide cross-link sites were inferred. From this, we obtained 5.64 million tagged and 1.21 million untagged cross-link sites (Supplemental Fig. S2A,B). Nearly 80% of FHFOX2 cross-links mapped to the sense strand of genes and 3% mapped to the antisense strand, suggesting that our protocol maintained a high degree of strand specificity (Supplemental Fig. S3A). Notably, FHFOX2 cross-link sites were significantly enriched in introns over UntagFOX2 cross-link sites, consistent with the known role of Rbfox2 in pre-mRNA splicing (Fig. 1A).

Statistically significant genic clusters of FHFOX2 binding were determined next using cross-link sites extended on either side by 12 nt as input into the CLIPper algorithm (available at https://github.com/YeoLab/clipper) (Lovci et al. 2013). Allowing for the calling of clusters within pre-mRNA regions, we identified 40,243 FHFOX2 clusters and only 2961 UntagFOX2 clusters (gene-level false discovery rate [FDR] ≤ 0.05; $P \leq 0.05$). We additionally required that each FHFOX2 cluster be enriched over UntagFOX2 cross-links within the bounds of the cluster using a binomial test with a Q-value threshold of 0.05, reasoning that the cross-link site distribution in the UntagFOX2 iCLIP represented background signal (Supplemental Fig. S2A). This process resulted in 35,639 high-confidence clusters. Similar to the distribution of cross-link sites, FHFOX2 iCLIP clusters were heavily weighted toward intronic regions (Supplemental Fig. S3A).

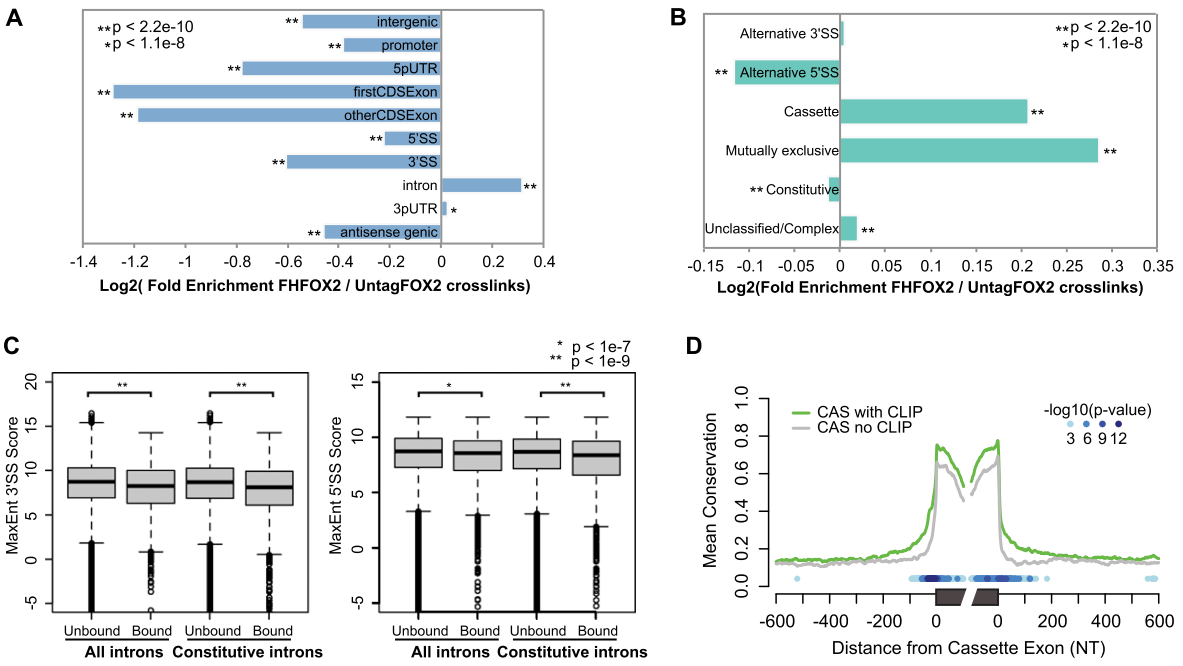

Figure 1. FHFOX2 iCLIP is enriched in conserved, alternatively spliced loci. (*A*,*B*) Fold enrichment of FHFOX2 cross-links over UntagFOX2 cross-links in genomic regions (*A*) or intronic categories (*B*). *P*-values were calculated using χ^2 test. (*C*) Box plots of splice site strength calculated using the MaxEnt algorithm for splice sites within 200 nt of FHFOX2 iCLIP clusters and unbound introns for both all introns and constitutive introns. *P*-value were calculated using Wilcoxon test. (*D*) Average Phastcons conservation score across placental mammals around expressed annotated cassette exons for cassette loci with iCLIP clusters and expression-matched unbound loci. Per-nucleotide *P*-value was calculated using Wilcoxon test.

Due to this intronic enrichment and Rbfox2's known role in regulating alternative splicing through recognition of intronic sequence elements, we next asked whether the FHFOX2 iCLIP signal was enriched around alternatively spliced regions using an intron centric approach. Correcting for the difference in library size, FHFOX2 cross-link sites were dramatically enriched compared with UntagFOX2 in introns flanking cassette exons and mutually exclusive exons, while the number of cross-link sites in constitutively spliced introns was slightly but significantly depleted (Fig. 1B). All introns were next subdivided into those without FHFOX2 iCLIP clusters and those with clusters <200 nt from either the 5′ or 3′ splice site. For both splice sites, the average splice site strength was weaker for introns marked by a splice site-proximal iCLIP cluster compared with unbound introns. The subset of introns flanking cassette exons, which as a class had significantly weaker splice sites than constitutive introns, did not have significantly different splice site strengths between FHFOX2-bound and unbound events (data not shown). We observed that the subset of FHFOX2-bound introns flanking constitutive exons had significantly weaker splice sites than unbound constitutive introns (Fig. 1C). Subdivision of constitutive splice sites according to strength revealed that weak 5′ splice sites were, on average, more strongly bound just downstream from the splice site than stronger sites (Supplemental Fig. S4). Rbfox2 may thus function to increase the efficiency of a subset of weaker constitutive splicing events in addition to its known role in influencing the choice between alternative splice sites.

Regions flanking the splice sites of cassette exons with nearby FHFOX2 clusters were more highly conserved across placental mammals. Exonic conservation for this subset was also significantly higher, indicative of preserved regulation and perhaps function of genes encoding Rbfox2-bound splicing events (Fig. 1D). These data together confirmed that, globally, FHFOX2 iCLIP signal was specific and bore hallmarks of the binding patterns of a bona fide regulator of conserved splicing events.

Rbfox2 iCLIP is strongly enriched around (U)GCAUG motifs

One defining characteristic of the Rbfox family of splicing regulators that distinguishes it from most other splicing factors is its recognition of a highly specific motif, (U)GCAUG (Jin et al. 2003; Underwood et al. 2005). While several individual Rbfox2-dependent splicing events have been specifically demonstrated to depend on this motif (Kuroyanagi 2009), a previous RBFOX2 CLIP study in humans suggested that only one-third of binding sites contained the minimal pentameric motif (Yeo et al. 2009). Due to the enhanced specificity and spatial resolution of our iCLIP protocol compared with the earlier RBFOX2 CLIP-seq study, which relied on a polyclonal antibody to the endogenous protein and lacked a negative control, we found several lines of evidence arguing that most Rbfox2 binding was driven by the canonical motif. First, 51% of intronic FHFOX2 iCLIP clusters contained at least one GCAUG occurrence, and 6.1% contained a conserved

GCAUG (Phastcons placental mammals >0.4). We also performed an unbiased motif analysis on the top 5000 clusters ranked by cross-link frequency, and, as expected, UGCAUG was the top-scoring hexamer (Fig. 2A). We found that 71% of clusters contained a motif significantly similar to the UGCAUG position-specific probability matrix generated from this analysis (FDR < 0.1; $P < 0.001$). Conversely, 16% of all conserved exact UGCAUG matches in expressed introns were within 100 nt of a cluster. To determine the fraction of motifs bound at high resolution, distance plots of FHFOX2 cross-link sites to conserved UGCAUG motifs in all introns were generated. Twenty-eight percent of UGCAUG occurrences were within 5 nt of a cross-link site, and 53% were within 300 nt, arguing that a subset of conserved UGCAUG motifs are specifically bound and detectable at high resolution (Fig. 2B, left panel). Correspondingly, cross-links peaked sharply at positions 2 and 6 in the motif, with an approximately fivefold drop in signal beyond positions 1 and 7 (Fig. 2B, right panel). This is likely due to direct interaction and cross-linking at the two uridine residues, as UV-induced cross-links occur more frequently at uridine residues (Sugimoto et al. 2012), resulting in reverse transcription terminating 1 nt upstream. Finally, in support of specific iCLIP signal reflecting functional binding, intronic UGCAUG motifs within 100 nt of a FHFOX2 cluster were significantly more conserved than an expression-matched set of unbound UGCAUG motifs (Fig. 2C). Taken together, Rbfox2 binds to a sizeable fraction of conserved UGCAUG elements across the mESC transcriptome and is largely dependent on the motif for binding. The strong preference for UGCAUG-containing binding sites in vivo is also in agreement with a recent investigation of in vitro binding preferences of Rbfox2 (NJ Lambert, A Robertson, M Jangi, S McGeary, PA Sharp, CB Burge, in prep.).

Two other motifs reached significance in the hexamer motif analysis, [CAG]C[AU]CAC (found in 6.8% of all clusters; FDR < 0.3; $P < 0.0002$) and UGUGUG (found in 18% of all clusters; FDR < 0.1; $P < 0.0003$), although the iCLIP signal around these motifs was less pronounced than around UGCAUG motifs (Fig. 2D–G). These motifs resemble the consensus motifs for hnRNPL and CELF/Cugbp1, respectively (Teplova et al. 2010; Ray et al. 2013). Motif analysis of clusters lacking the UGCAUG-related motif failed to produce significant or high-information content motifs. Consistent with this observation, 39% of CCUCAC and 72% of UGUGUG motif matches co-occurred with a UGCAUG motif within 10 nt flanking

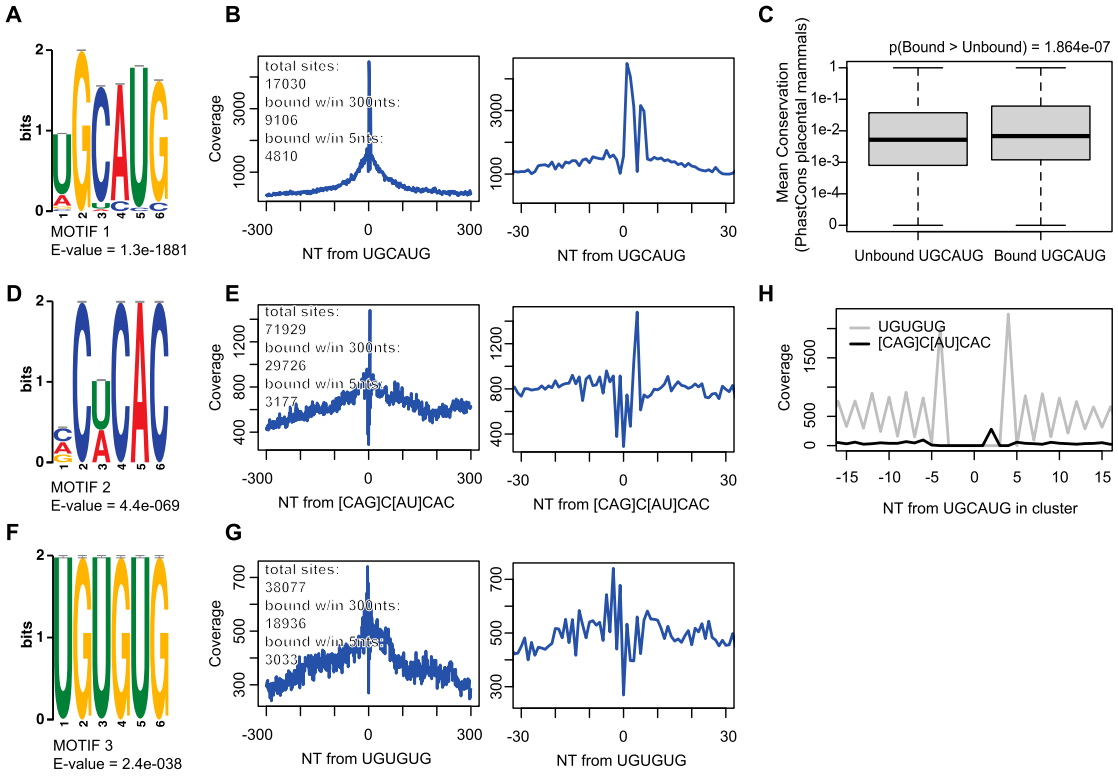

Figure 2. The majority of FHFOX2 individual nucleotide-resolution cross-linking immunoprecipitated regions contain consensus Rbfox2 motifs. (*A*,*D*,*F*) Top three enriched hexamers generated by MEME (Bailey et al. 2009) for the 20 nt flanking the summits of the top 5000 iCLIP clusters. (*B*,*E*,*G*) Total FHFOX2 cross-link coverage across intronic occurrences of the top three enriched hexamers conserved across placental mammals (average phastCons >0.4) within 300 nt (*left* panels) and 30 nt (*right* panels) windows. Position 0 is the 5′-most nucleotide of the hexamer. (*C*) Box plot of mean conservation of intronic bound (<100 nt from an iCLIP cluster) and unbound UGCAUG motifs. (*H*) Distribution of secondary motif matches around UGCAUG matches within clusters. Matches were determined using position-specific probability matrices from Figure 3A. Coverage represents the 5′-most nucleotide of each motif occurrence, with position 0 being the first nucleotide of the UGCAUG position-specific probability matrix match.

iCLIP cluster summits. To specifically examine their co-occurrence within clusters at high resolution, the distance of each secondary motif to the top motif was plotted (Fig. 2H). In accordance with the cross-link coverage plots, the UGUGUG motif was present on either side of the UGCAUG motif, with a periodicity implying a longer stretch of UG dinucleotides. The CCUCAC motif was similarly enriched upstream of and downstream from the consensus motif, although with lower overall frequency. Motifs recognized by other RBPs can thus occur in the proximity of Rbfox2-binding events but are unlikely to lead to Rbfox2 binding in the absence of a UGCAUG-related motif. Interestingly, splicing analysis of cardiac differentiation revealed enrichments of Rbfox, hnRNPL, and CELF motifs around differentially spliced exons (Kalsotra et al. 2008). This raises the possibility that Rbfox2 may compete with other splicing regulators recognizing proximal motifs in different cellular contexts.

Hundreds of splicing and gene expression changes in response to Rbfox2 loss

Our iCLIP data suggested that many Rbfox2-binding events were candidates for regulation of splicing due to, for example, enrichment within alternatively spliced regions. To determine global patterns of Rbfox2-dependent splicing changes in mESCs, we performed deep sequencing of mRNA (RNA-seq) upon shRNA-mediated depletion of Rbfox2. Two control hairpins targeting GFP and luciferase (shLuc and shGFP) and two treatment hairpins targeting the 3′ untranslated region (UTR) of Rbfox2 (shFox2-1 and shFox2-2) were lentivirally transduced into V6.5 mESCs in biological duplicates, and poly(A)-selected mRNA was prepared for sequencing on the Illumina platform. Analysis of annotated splicing changes between control and Rbfox2 knockdown using the MISO algorithm (Katz et al. 2010) revealed hundreds of putative Rbfox2-regulated splicing events, of which 18 of 21 cassette exons tested were successfully validated by RT–PCR (Supplemental Fig. S5A,B). Splicing events were quantified using percent spliced in (or ψ) to determine the percent of total isoforms represented by one of two isoforms. High-confidence Rbfox2-dependent alternative splicing changes in each splicing category were defined as those with $\Delta\psi \geq |5\%|$ and a Bayes′ factor exceeding 5 in either of two comparisons: shLuc versus shFox2-1 or shGFP versus shFox2-2. We further required events that were significant in a given comparison to change in the same direction in the other comparison and termed these "concordant" events. In this manner, we detected concordant splicing changes in all annotated categories ranging from 4% to 13% of expressed events (Supplemental Fig. S5A,C).

Rbfox2 RNA map reveals position-dependent regulation of cassette exons

Meta-exon analyses of CLIP for a number of factors, including Nova, PTB, and Mbnl, have drawn correlations between the location of the binding event with respect to the cassette exon and the direction of splicing regulation (Ule et al. 2006; Xue et al. 2009; Wang et al. 2012). Previously, motif analyses and minigene assays using Rbfox2 target cassette exons argued that binding within the ~200 nt downstream from a cassette exon resulted in Rbfox2-dependent exon inclusion, whereas binding upstream of the cassette exon correlated with repression. To globally correlate Rbfox2 binding and regulation, we overlaid the high-resolution FHFOX2 iCLIP data with the concordant events defined above. Cross-link sites were mapped in the introns flanking exons activated by Rbfox2 ($\Delta\psi \geq 0.05$; Bayes′ factor ≥ 5) and exons repressed by Rbfox2 ($\Delta\psi \leq -0.05$; Bayes′ factor ≥ 5) (Fig. 3A). The resulting distribution showed that enriched cross-linking within the ~200 nt downstream from the cassette exon as well as at more distant sites proximal to the downstream constitutive 3′ splice site correlated strongly with exon activation, whereas cross-linking within the ~300–400 nt upstream of the cassette exon correlated weakly with exon repression. The iCLIP signal around nonregulated and expressed cassette exons ($|\Delta\psi| \leq 0.02$; Bayes′ factor ≥ 3) did not show enrichment in either the upstream or downstream introns.

We next asked whether the distribution of both conserved and all canonical Rbfox2 motifs explained the cross-linking pattern that we observed. Motifs were indeed enriched within the first 200 nt downstream from Rbfox2-enhanced cassette exons, with weaker motif enrichment upstream of exons repressed by Rbfox2 (Fig. 3B). Interestingly, a subset of enhanced exons also contained UGCAUG motifs proximal to the downstream constitutive 3′ splice site, perhaps implicating distant motifs in Rbfox2-dependent exon inclusion (Lim and Sharp 1998; Lovci et al. 2013). Most of the motif occurrences near regulated exons were well conserved. This is consistent with stronger evolutionary selection maintaining robust and presumably functional splicing changes.

The motif and cross-link distributions defined a critical window ±300 nt around regulated cassette exons within which Rbfox2 binding was most potent. We further examined whether the strength of Rbfox2 binding within these critical regions impacted the degree of regulation. To this end, all detected cassette exons with Bayes′ factor >3 were separated into those with a twofold or greater difference in expression-normalized cross-link density between the upstream or downstream 300 nt. Exons with primarily upstream or downstream binding were further subdivided into quartiles of normalized cross-link density, and average $\Delta\psi$ was plotted within each of these quartiles. For both upstream and downstream bound cassette exons, higher cross-link density correlated with stronger repression and activation, respectively. Additionally, cassette exon repression by Rbfox2 tended to be less potent than exon activation (Fig. 3C). This trend persisted upon decreasing the threshold of normalized upstream versus downstream cross-link density to 1.5-fold, suggesting that small relative differences in upstream versus downstream binding were sufficient to confer exon repression or activation.

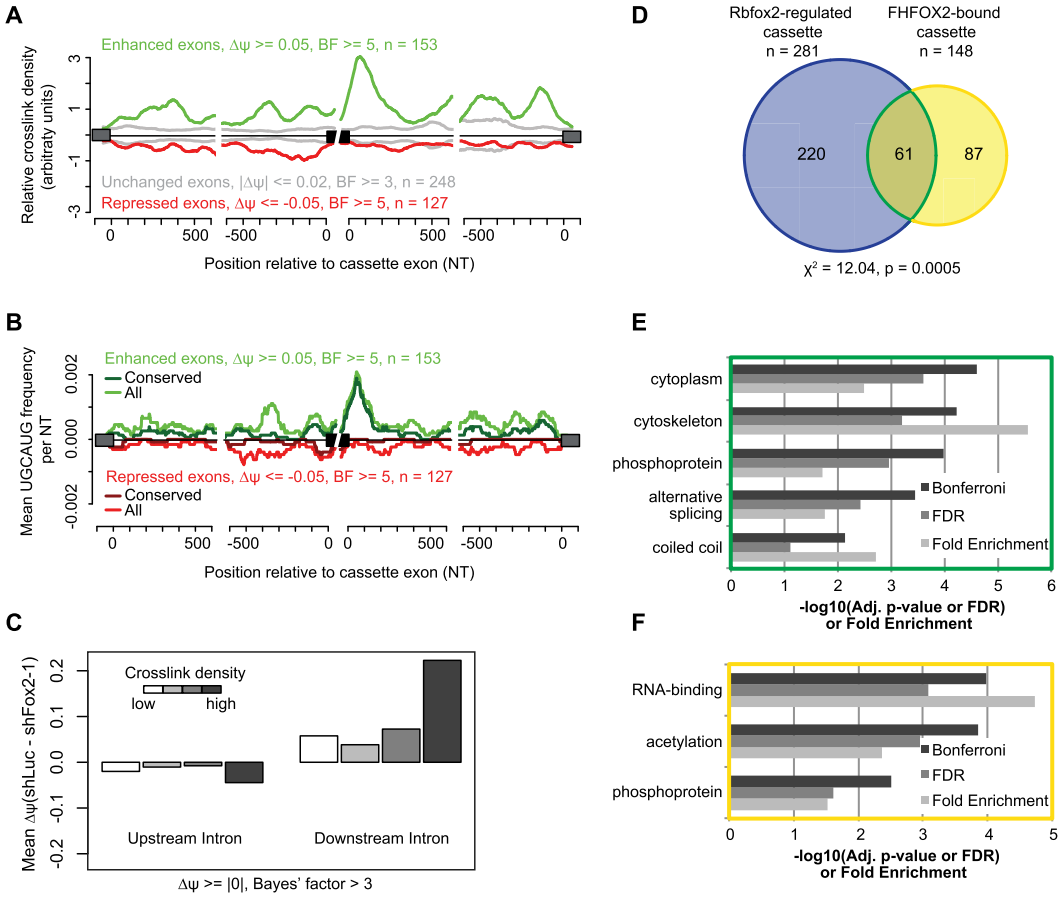

Figure 3. RNA map of Rbfox2 regulation. (*A*,*B*) Distribution of average FHFOX2 cross-link sites mapped by iCLIP (*A*) or motif frequency (*B*) around exons enhanced or repressed by Rbfox2 in mESCs as determined by RNA-seq upon Rbfox2 knockdown. (*C*) Average Δψ of cassette exons within each quartile of cross-link density in a 300-nt window upstream of or downstream from the cassette exon. (*D*) Intersection of Rbfox2-regulated exons defined by MISO ($|\Delta\psi| \geq 0.05$, Bayes' factor ≥ 5) (Katz et al. 2010) and exons within 300 nt of a FHFOX2 iCLIP cluster. *P*-value determined by χ^2 test. Unregulated, bound exons defined by $|\Delta\psi| \leq 0.02$, Bayes' factor ≥ 5. (*E*,*F*) Gene ontology analysis using DAVID (Huang et al. 2009) of bound and regulated (corresponding to Venn diagram overlap outlined in green) (*E*) or bound and unregulated (corresponding to Venn diagram region outlined in yellow) (*F*) cassette exons over the background of all cassette exons expressed in mESCs.

Silent FHFOX2-binding events are associated with NMD

The high-resolution RNA map defined above strongly correlated context-dependent binding and regulation for direct targets of Rbfox2. We subsequently asked whether binding within the critical 300 nt of an intron boundary was necessary or sufficient to confer splicing regulation by intersecting bound and regulated splicing events. Focusing first on direct targets of Rbfox2, we observed a large and statistically significant overlap between Rbfox2-regulated concordant events (regulated cassettes) and events containing a FHFOX2 iCLIP cluster within 300 nt of the cassette exon (bound cassettes) (Fig. 3D). Gene ontology analysis using the functional annotation tool DAVID (Huang et al. 2009) revealed that genes containing the 61 splicing events bound and regulated by Rbfox2 were enriched for cytoskeletal and cytoplasmic terms after correction for multiple hypothesis testing (FDR ≤ 0.1) (Fig. 3E). Many of the detected targets in this category were among those shown to be differentially regulated by Rbfox2 during epithelial–mesenchymal transitions and cell migration (Venables et al. 2009; Warzecha et al. 2009; Shapiro et al. 2011). In support of direct targeting by Rbfox2, hexamer analysis using MEME identified the UGCAUG motif as the only significantly enriched motif in the 300 nt flanking these exons (61 motifs/61 sites, $E = 0.0012$). The 220 splicing changes that were not bound by FHFOX2 did not show significant enrichment in any particular gene ontology category and were not significantly enriched for any hexamers, most likely owing to the heterogeneity of indirect effects caused by Rbfox2 loss.

Although the subset of events both bound and regulated by Rbfox2 in our study was significant, the question remained as to why some cassette exon-proximal FHFOX2-binding events were not coupled with a splicing change. By RT–PCR, only four of 13 such events tested showed any

change in splicing, suggesting that these were unlikely to be MISO false negatives (data not shown). While it could not be excluded that these were countered by the action of other splicing factors or were simply not in the correct context to elicit a splicing change, we considered the possibility that these 87 splicing events were in fact Rbfox2-regulated, but the changes were rendered silent through another mechanism. In contrast to the bound and regulated set, the splicing events that were Rbfox2-bound but significantly unchanged upon knockdown were most strongly enriched for the RBP gene ontology term (Fig. 3F). Upon closer inspection of the bound and unchanging loci, we were surprised to find that in 37 out of 87 instances, one of the two isoforms appeared to be a substrate for degradation by NMD, defined by the introduction of a PTC >50 nt upstream of the last splice junction. Isoforms that harbor a PTC arising from exon inclusion or skipping would not be detected in RNA-seq data due to the instability of the transcript, and, consequently, the change in splicing due to Rbfox2 knockdown would be underestimated for these variants. AS-NMD has been documented as a regulatory mechanism for gene expression. In particular, the expression of several RBPs is autoregulated or cross-regulated through AS-NMD-dependent mechanisms (Wollerton et al. 2004; Boutz et al. 2007; Lareau et al. 2007; Ni et al. 2007). The striking observation that a large fraction of bound and apparently unregulated events were likely to generate NMD substrates motivated the hypothesis that Rbfox2 was controlling gene expression levels by regulating AS-NMD.

Rbfox2 modulates AS-NMD events in RBPs

Due to the documented autoregulatory capacity of many RBPs, we first verified whether Rbfox2 was regulating AS-NMD within RBPs. Manual curation of the cassette exons in RBPs with proximal FHFOX2 iCLIP clusters revealed that 45 were putative NMD substrates by generating a PTC upon exon inclusion (NMD-INC), and 20 were putative NMD substrates by generating a PTC upon exon skipping (NMD-SK). iCLIP signal was distributed both upstream of and downstream from these exons, suggesting that Rbfox2 could either promote or inhibit the NMD isoform depending on the context. Based on the positional regulation criteria that we determined for cassette exons, we hypothesized that NMD-INC exons with greater downstream iCLIP density and NMD-SK exons with greater upstream iCLIP density would represent genes for which NMD is enhanced by Rbfox2 (NMD enhanced). Conversely, NMD would be suppressed by Rbfox2 for genes in which NMD-INC exons are predominantly bound upstream or NMD-SK exons are predominantly bound downstream (NMD suppressed) (Fig. 4A).

To specifically assay whether the bound exons within RBPs were in fact AS-NMD events regulated by Rbfox2, we chose six candidate splicing events within iCLIP

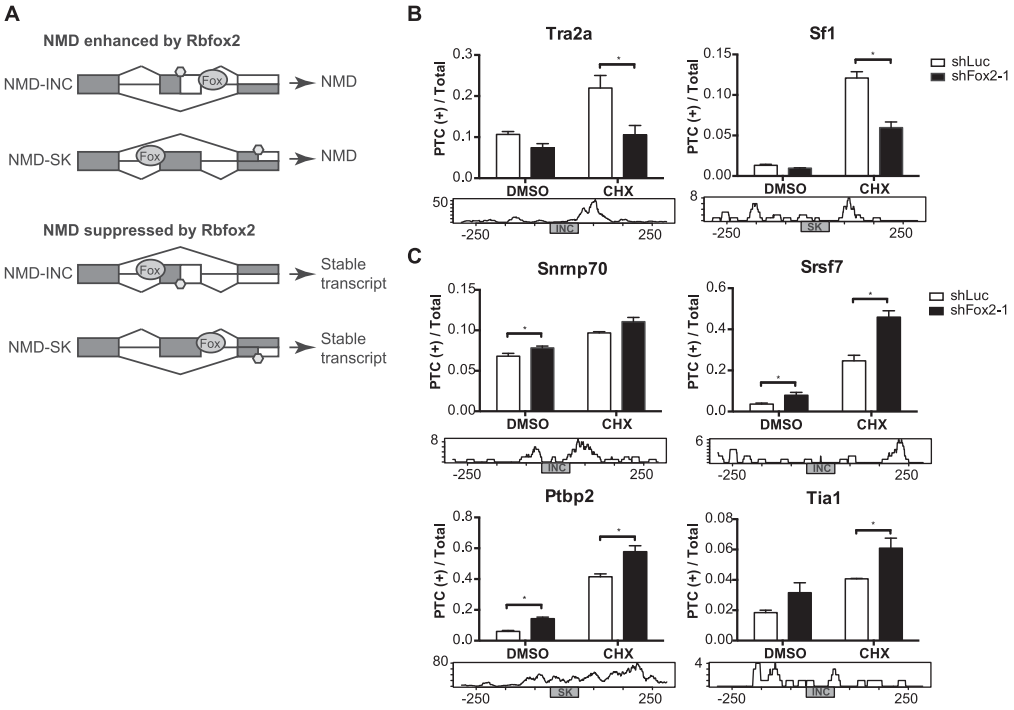

Figure 4. Rbfox2 regulates AS-NMD events in RBPs. (*A*) Schematic of modes of AS-NMD regulation by Rbfox2. Gray boxes represent coding exonic sequence, white boxes are untranslated exonic sequence, and hexagons represent PTCs. (*B*,*C*) Quantification of fraction of transcripts encoding the PTC-containing isoform in Rbfox2-destabilized (*B*) and Rbfox2-stabilized (*C*) genes after Rbfox2 knockdown and cycloheximide treatment (6 h, 10 μg/mL) from qRT–PCR using primers spanning exon junctions. Error bars represent standard error across biological triplicates; the asterisk represents $P < 0.05$, paired Student's *t*-test. Density plots *below* qRT–PCR plots depict FHFOX2 iCLIP cross-link density ±300 nt of NMD-INC (INC) or NMD-SK (SK) exons.

cross-linked transcripts encoding RBPs to examine more closely. The six genes, Snrnp70 (U170K), Srsf7, Tra2a, Sf1, Ptbp2, and Tia1, contained silent NMD splicing events with 0%–5% splicing change upon Rbfox2 knockdown and distinct predicted effects of Rbfox2 binding. Cells stably expressing hairpins against Luciferase or Rbfox2 were treated with cycloheximide for 6 h to block translation, which is essential for NMD. The type of splicing event putatively responsible for NMD (exon inclusion or exon skipping) was determined, and total transcript levels and relative levels of the NMD isoform were assayed by quantitative RT–PCR (qRT–PCR) for each condition. All six predicted NMD isoforms increased in abundance upon cycloheximide treatment, suggesting that these transcripts were indeed normally degraded by NMD (Fig. 4B,C). Two of the candidates, Tra2a and Sf1, showed significantly decreased levels of the NMD isoform upon Rbfox2 knockdown and cycloheximide treatment (Fig. 4B) and significantly increased total transcript levels (Supplemental Fig. S6A), indicating that Rbfox2 promoted the NMD isoform of these genes. The NMD isoforms of Snrnp70, Srsf7, Ptbp2, and Tia1 were up-regulated upon Rbfox2 knockdown in the absence of NMD (Fig. 4C), suggesting that for these genes, Rbfox2 suppressed the NMD isoform. Accordingly, three of these four genes showed significant decreases in total transcript levels upon Rbfox2 knockdown (Supplemental Fig. S6B).

Rbfox2 repressed exon inclusion for Sf1, Srsf7, and Snrnp70, all of which harbored iCLIP signal both upstream of and downstream from the cassette exons, suggesting that binding in both regions may lead to repression. As the iCLIP signal was predominantly downstream from the NMD-SK exon in Ptbp2 and upstream of the NMD-INC exon in Tia1, the observed increases in NMD isoforms by qRT–PCR were consistent with antagonism of the NMD isoform by Rbfox2. Strong binding downstream from the NMD-INC exon in Tra2a, indicating enhancement of the NMD isoform by Rbfox2, was similarly consistent with the relative decrease in the NMD isoform by qRT–PCR upon Rbfox2 knockdown. This argued for Rbfox2-dependent modulation of AS-NMD events within RBPs in a manner predicted by the pattern of Rbfox2 binding. Notably, 50% of the Rbfox2-dependent splicing changes in this set of exons were undetectable in the absence of NMD inhibition, indicating that genome-wide, a large number of Rbfox2-dependent splicing events go undetected under standard experimental conditions.

Autoregulation thresholds are partially set by Rbfox2

Negative autoregulation feedback loops have been proposed to maintain cellular homeostasis by buffering gene expression against stochastic variation. It has been previously demonstrated that RBPs are frequently negatively autoregulated by AS-NMD (Lareau et al. 2007; Ni et al. 2007), resulting, for example, in the case of SFRS10/Tra2b, in low degrees of cell-to-cell variability of expression (Sigal et al. 2006). The corollary to this model is that another mechanism must exist to allow for changes in expression of these autoregulated genes; for example, between cell types. We hypothesized that for a subset of RBPs, the threshold of negative autoregulation and, ultimately, the steady-state level of gene expression are established by cross-regulation of the AS-NMD event by Rbfox2. First, we verified whether the Rbfox2-regulated NMD events were themselves autoregulated. We focused on Ptbp2 and Tia1, both of which had strong motif-dependent FHFOX2 iCLIP signal in well-conserved flanking introns (Fig. 5A) and showed 1.3-fold and 1.9-fold down-regulation, respectively, in the Rbfox2 knockdown RNA-seq accompanied by down-regulation at the protein level (Fig. 5B). These have been shown or suggested in the literature to be candidates for autoregulation (Le Guiner et al. 2001; Boutz et al. 2007). Since our previous observations indicated that Rbfox2 inhibited the NMD isoform, it was possible that higher steady-state RBP expression could be achieved in the presence of Rbfox2 due to, effectively, a higher threshold for autoregulation. To test this model, epitope-tagged Ptbp2 or Tia1 were expressed from an exogenous plasmid in cells stably expressing shLuc or shFox2-1 (Fig. 5C). Upon inhibition of NMD by transient knockdown of Upf1, the relative fraction of NMD isoform increased for both Ptbp2 and Tia1 when each of these factors was overexpressed, consistent with negative autoregulation (Fig. 5D). The fraction of NMD isoform was further increased upon Rbfox2 knockdown, supporting the hypothesis that Rbfox2 was modulating the autoregulation of these RBPs. Most strikingly in the case of Tia1, autoregulation by AS-NMD under steady-state conditions was not significant and only became apparent upon Tia1 overexpression and Rbfox2 knockdown. This argues strongly for tight regulation of Tia1 NMD only above a specific threshold of gene expression.

Widespread regulation of AS-NMD by Rbfox2 controls gene expression

Our results suggest that a subset of Rbfox2-dependent binding events is rendered silent through AS-NMD, in particular within autoregulated RBPs. We next tested the prediction that this mode of regulation impacts the steady-state gene expression of these RBPs. The 65 Rbfox2-bound AS-NMD splicing events identified above were filtered for those with greater than twofold differential iCLIP density between the 300 nt upstream of and downstream from the regulated exon. These were next separated into predicted NMD-suppressed or NMD-enhanced categories (24 and 20 genes, respectively), and the gene expression upon Rbfox2 knockdown was plotted for bins of increasing normalized iCLIP density. Consistent with our predictions, RBP genes containing strongly bound NMD-suppressed events decreased significantly in expression upon Rbfox2 loss, while genes with strongly bound NMD-enhanced events increased significantly (Fig. 6A). These results lend compelling support to the model that Rbfox2 cross-regulation of AS-NMD events within RBPs impacts their gene expression patterns.

To expand our observations beyond RBPs, we next sought to define and characterize Rbfox2-bound AS-NMD events

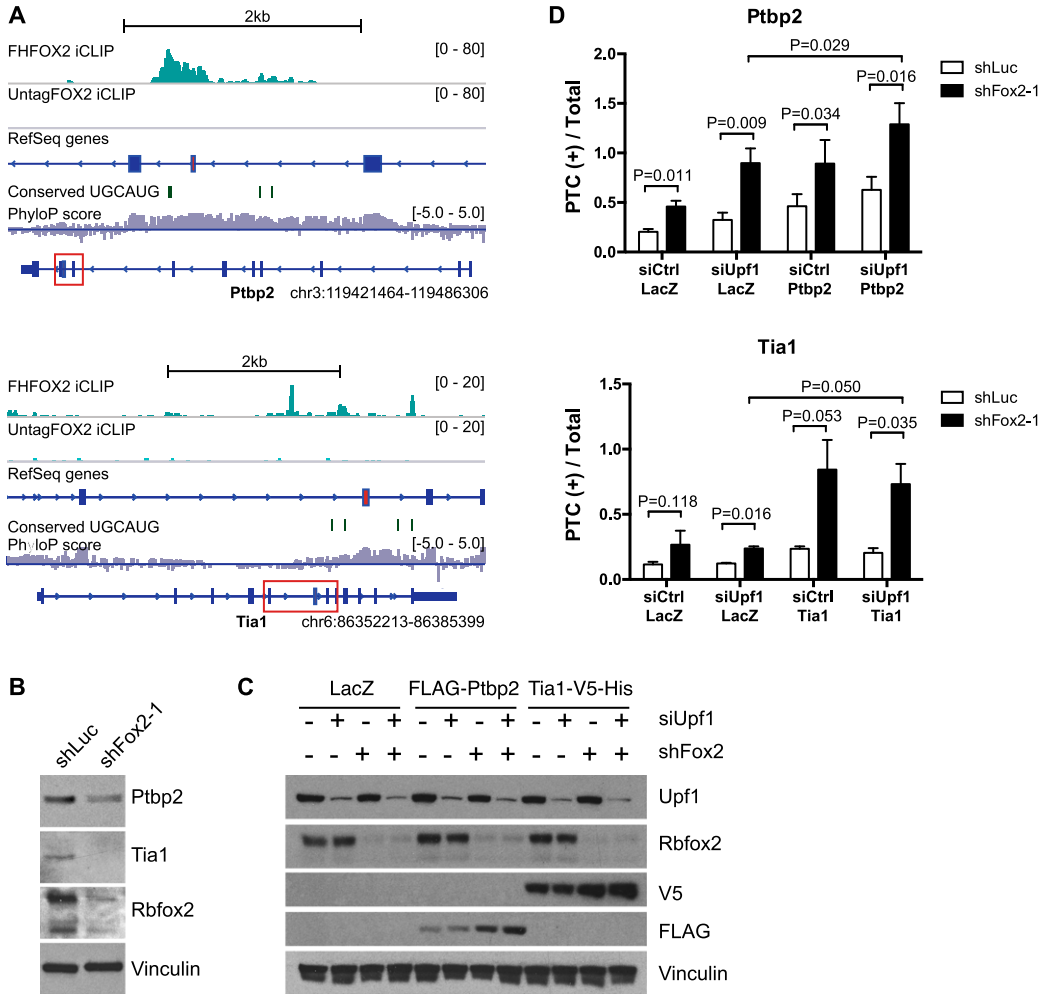

Figure 5. Rbfox2 inhibits autoregulation of FHFOX2-bound splicing factors. (*A*) Integrative Genomics Viewer genome browser image of iCLIP signal around an NMD-SK event in Ptbp2 and an NMD-INC event in Tia1. (*B*) Western blot showing decreased expression of Ptbp2 and Tia1 upon Rbfox2 knockdown in mESCs. (*C*) Western blot showing Ctrl or Upf1 and Luciferase or Rbfox2 knockdown and Ptbp2 and Tia1 exogenous expression. (*D*) Quantification of the percentage of transcripts encoding the PTC-containing isoform as in Figure 4, B and C. Error bars represent standard error across biological triplicates; *P*-values were determined using paired Student's *t*-test.

genome-wide. We developed a computational pipeline to assemble transcripts containing known and novel cassette exons using RNA-seq reads from control and Rbfox2 knockdown as well as control and Upf-1 knockdown and control and cycloheximide-treated mESCs (data from Hurt et al. 2013). Each of these transcripts was then translated in silico, assigning reading frames using Ensembl start codon annotations. If the putative alternative splicing event resulted in a PTC >50 nt upstream of the last exon–exon junction, the event was categorized as either NMD-INC or NMD-SK, depending on whether the NMD isoform resulted from exon inclusion or skipping, respectively. Isoform pairs that were coding and not NMD substrates regardless of whether the alternative exon was included or skipped were categorized as alternative coding sequences (AS-CDSs).

To validate that the algorithm was identifying true AS-NMD events, we measured splicing changes in NMD-INC, NMD-SK, and AS-CDS categories upon Upf1 knockdown using MISO. $\Delta\psi$ values (shGFP-shUpf1) were significantly negative for the 69 NMD-INC events and significantly positive for the 30 NMD-SK events with a Bayes' factor >3, supporting the stabilization of AS-NMD isoforms upon inhibition of NMD (Fig. 6B). This was further validated at the gene level using RNA-seq by measuring fold change in expression upon Upf1 knockdown. Genes containing NMD splicing events were separated into non-NMD or NMD isoforms, and fold changes in expression were compared with isoform expression from genes containing AS-CDS events. Unstable NMD isoforms were significantly increased upon Upf1 knockdown compared with AS-CDS isoforms (Fig. 6C). Stable non-NMD isoforms were significantly decreased; similar observations that an increase in NMD isoforms was coupled with a decrease in non-NMD isoforms have been reported in HeLa cells upon Upf1 knockdown (Pan et al. 2006).

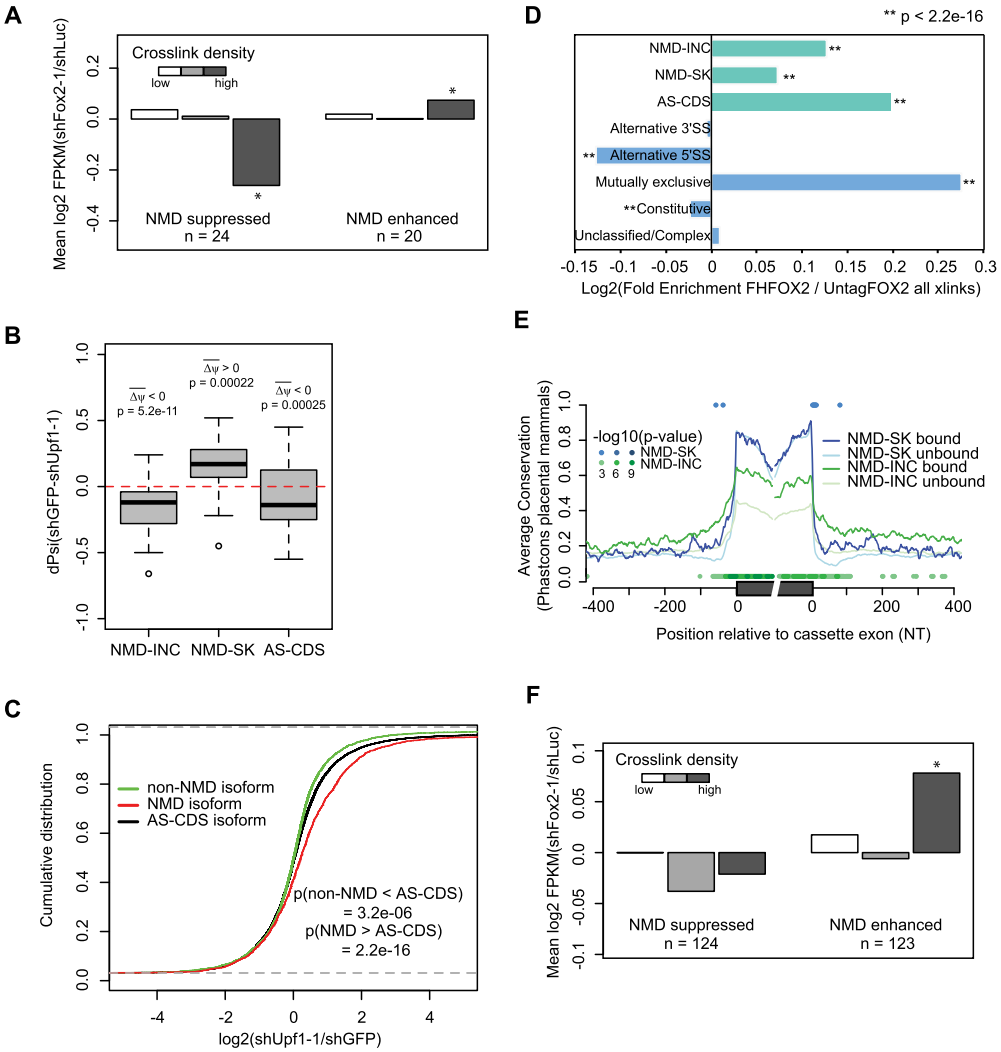

Figure 6. Rbfox2 regulation of AS-NMD has widespread effects on gene expression. (*A*) Gene expression changes upon Rbfox2 knockdown of RBPs with increasing amounts of iCLIP density around AS-NMD exons for NMD-suppressed and NMD-enhanced genes. Significance was calculated using one-sided one-sample Student's *t*-test. (*) $P < 0.05$. (*B*) Box plots of MISO splicing changes for predicted NMD-INC, NMD-SK, and AS-CDS events. *P*-values represent one-tailed one-sample Student's *t*-test. (*C*) Cdf plot of gene expression changes for each class of isoform in Upf1 knockdown versus control. (*D*) FHFOX2 iCLIP cross-links are enriched in introns flanking NMD-INC and NMD-SK splicing events over UntagFOX2 cross-links. Significance was calculated using χ^2 test. (*E*) Average Phastcons conservation score across placental mammals around expressed NMD-INC and NMD-SK cassette exons for loci with iCLIP clusters and unbound loci. (*F*) As in *A* but using all predicted AS-NMD events with greater than twofold differential binding between upstream and downstream introns.

We applied these NMD predictions to the Rbfox2-binding and regulation data to identify evidence of global AS-NMD regulation by Rbfox2. Applying the bucketing approach described previously, the FHFOX2 iCLIP signal was significantly enriched in NMD-INC and NMD-SK loci in addition to being enriched in AS-CDS loci (Fig. 6D). This argued that Rbfox2 regulation of AS-NMD is likely to be a more widespread phenomenon than would be predicted based solely on transcriptome data from Rbfox2 depletion. As seen previously for all cassette exons, the average conservation of intronic regions flanking AS-CDS, NMD-INC, and NMD-SK exons was significantly higher for those events containing a FHFOX2 iCLIP cluster within 300 nt of the exon (Fig. 6E). In particular, bound NMD-INC events showed high levels of intronic and exonic conservation; this may be partially attributable to the presence of ultraconserved elements (Fig. 6E). The specific evolutionary maintenance of Rbfox2-regulated AS-NMD cassettes argues for a widespread functional role of this mode of regulation in development.

The effect of Rbfox2-dependent splicing regulation on the expression of all genes containing bound AS-NMD events was considered next. As for RBPs, genes with greater than twofold differential iCLIP density around AS-NMD exons were separated into NMD-suppressed

(124 genes) and NMD-enhanced (123 genes) categories based on the location of Rbfox2 binding. In support of pervasive AS-NMD regulation by Rbfox2, NMD-suppressed genes with higher iCLIP density tended to decrease in expression upon Rbfox2 knockdown, although not significantly. As predicted, NMD-enhanced genes with strong Rbfox2 binding were significantly up-regulated upon Rbfox2 knockdown (Fig. 6F). This trend persisted when RBPs were removed from these gene sets. Our analysis provides strong evidence that Rbfox2 regulation of AS-NMD impacts steady-state gene expression patterns of RBPs as well as a host of other potentially autoregulated genes in mESCs.

Discussion

Here, we performed a genome-wide analysis of Rbfox2 activity in mESCs by mapping high-resolution Rbfox2-binding sites to transcriptome changes conferred by Rbfox2 loss. The use of a negative control iCLIP library, which was recently demonstrated to be critical for filtering non-specific signal inherent to CLIP experiments (Friedersdorf and Keene 2014), allowed us to detect high-fidelity binding to conserved loci containing the consensus UGCAUG motif. More than 70% of binding sites contained a similar motif, a much larger percentage than was reported in the previous RBFOX2 RNA map from human ESCs (Yeo et al. 2009). Investigation of a large subset of "silent" Rbfox2-binding events around cassette exons that were not coupled with measurable Rbfox2-dependent splicing regulation revealed an unexpected enrichment in genes regulated by AS-NMD, in particular RBPs. We demonstrated that Rbfox2 determines a threshold for the ratio of NMD to non-NMD isoforms for several of these factors. Globally, regulation of AS-NMD by Rbfox2 resulted in changes in gene expression consistent with the polarity of Rbfox2 binding and regulation, thus greatly expanding the breadth of the Rbfox2 regulatory network beyond the previously understood canonical splicing regulation. In fact, this novel general model in which one splicing regulator controls the autoregulatory splicing of another protein likely explains how levels of autoregulated proteins vary between cell states.

The observation that hundreds of silent splicing events bound by Rbfox2 are putative AS-NMD cassette exons suggests that a functional splicing activity can be attributed to the majority of Rbfox2-binding events when in the correct context; i.e., within 300 nt of an exon. This provides an interpretation for the extensive Rbfox2 binding in this region around constitutive exons. Although no evidence of alternative splicing was detected upon Rbfox2 loss or NMD inhibition for the large number of constitutive exons with proximal iCLIP clusters, our results suggest that Rbfox2 actively reinforces these constitutive exons with weak splice sites whose missplicing would only be evident upon simultaneous NMD inhibition and Rbfox2 knockdown. Conversely, a large number of splicing changes within genes of heterogeneous function observed upon Rbfox2 knockdown cannot be explained by detectable direct binding. As many targets of Rbfox2 are in fact RBPs that change in expression upon Rbfox2 knockdown, we posit that these are the products of indirect splicing changes arising from differential expression of components of the splicing machinery.

The NMD pathway is required for organismal development as well as in several tissue- and condition-specific contexts, suggesting a regulatory role extending beyond a general surveillance mechanism (Hwang and Maquat 2011). Several recent studies have highlighted the specific contribution of AS-NMD to the regulation of gene expression. Between 15% and 30% of genes within specific mouse tissues were differentially expressed due to loss of AS-NMD upon genetic ablation of Upf2 (Weischenfeldt et al. 2012). During developmental transitions, gene expression changes of individual factors, such as the PTB/nPTB switch in neuronal development, or entire networks of factors, as with myelocyte-specific proteins in granulopoiesis, also depend on splicing-coupled NMD (Boutz et al. 2007; Wong et al. 2013). These observations suggest that gene expression regulation by AS-NMD may be particularly relevant in a physiological context such as stress or differentiation. Furthermore, our results imply that the activation of entire AS-NMD networks could be initiated by altering the expression of specific splicing factors within such networks.

AS-NMD is frequently associated with negative autoregulatory feedback loops, which generate tightly controlled systems of gene expression. Best characterized at the level of transcription, negative feedback has been shown to dampen noise in transcriptional output (Becskei and Serrano 2000). In the case of splicing, deviations in the precise levels of splicing components could be expected to have far-reaching consequences in maintaining cellular homeostasis. As a result, the direct negative feedback of RBPs on the splicing of their own transcripts may have evolved to safeguard against fluctuations in expression. Several examples of conserved cross-regulation of an autoregulatory AS-NMD event within one RBP by another RBP exist (Le Guiner et al. 2001; Izquierdo and Valcarcel 2007; Spellman et al. 2007; Rossbach et al. 2009; Saltzman et al. 2011; Anko et al. 2012). Our results hint that this may be a more general characteristic of autoregulatory splicing designed to tune gene expression while maintaining the noise buffering inherent to negative autoregulation. We propose that Rbfox2 is capable of functioning in this role in a positive or negative manner, depending on the context. For the set of genes for which Rbfox2 enhances the NMD isoform, as in the case of Tra2a, Rbfox2 potentiates negative autoregulation by the RBP, thus lowering the threshold amount of RBP needed for autoregulation and decreasing its steady-state levels. Here, decreased expression of Rbfox2 results in a concomitant increase in the RBP. Conversely, for those genes for which Rbfox2 represses the NMD isoform, such as Ptbp2 or Tia1, Rbfox2 antagonizes the negative autoregulation, and higher expression of the RBP is necessary to compete out Rbfox2 and trigger splicing of the NMD isoform. Decreased expression of Rbfox2 correlates with decreased expression of the RBP in this scenario. In this manner, changes in Rbfox2 expression during normal physiological

processes would be expected to alter splicing networks not only due to changes in direct targeting by Rbfox2 but also through expression shifts of other splicing regulators. This effect of autoregulation tuning on gene expression is diagrammed in Figure 7.

The tuning of autoregulation is not necessarily limited to RBPs and other splicing factors. While it is simplest to envision direct competition or cooperativity between Rbfox2 and a splicing factor at an autoregulated splicing event, several provocative examples indicate that many AS-NMD cassettes in genes with no reported RNA-binding activity may autoregulate through noncanonical mechanisms. In particular, one NMD-INC event predicted to be enhanced by Rbfox2 occurs in the gene Sumo3, a small ubiquitin-related modifier. Sumo modification of RBPs is suggested to be relevant during stress responses, such as heat shock (Pelisch et al. 2010). This raises the possibility that down-regulation of Sumo3 by Rbfox2 may feed back on its own splicing during stress through decreased sumoylation of hnRNPs. In a similar vein, phosphorylation of SR proteins by the Clk kinases, of which NMD events in Clk1, Clk2, and Clk4 are bound by Rbfox2, may lead to SR protein-mediated regulation of Clk RNA splicing accomplished through the kinase activity of the encoded proteins (Colwill et al. 1996; Duncan et al. 1997).

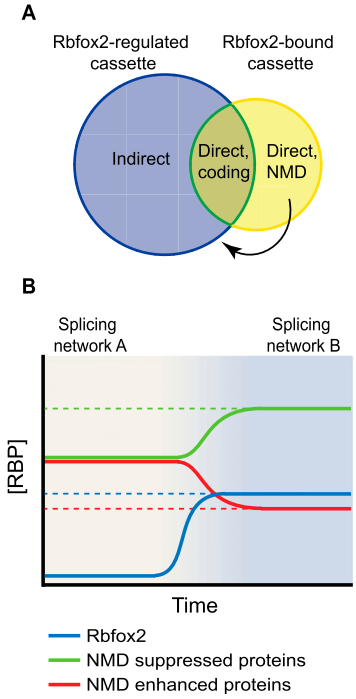

Figure 7. Rbfox2 regulates silent NMD events to tune RBP networks. (*A*) Schematic Venn diagram of Rbfox2-regulated and Rbfox2-bound cassette exons shows how the direct regulation of NMD splicing events within RBPs by Rbfox2 can mediate the indirect splicing changes seen upon Rbfox2 loss. (*B*) Model of Rbfox2-mediated tuning of RBP expression patterns. As Rbfox2 expression increases, steady-state expression of RBPs for which Rbfox2 suppresses NMD also increases. Simultaneously, steady-state expression of RBPs for which Rbfox2 enhances NMD decreases. This results in a shift from splicing network A to splicing network B, defined by both direct and indirect targets of Rbfox2.

The model that we propose has the potential to alter splicing regulation on both short and long timescales. Rbfox2 and other Rbfox family members undergo changes in expression across various tissues during development (Kuroyanagi 2009). Sustained increases in Rbfox2 protein levels would thus shift expression of the network of Rbfox2 AS-NMD targets by resetting steady-state expression (Fig. 7). However, this may also be regulated transiently; for example, in response to extracellular cues or stresses. In neuronal cells, Rbfox1 undergoes a depolarization-induced splicing switch within its own transcript, resulting in an alternate C-terminal reading frame that relocates Rbfox1 from the cytoplasm to the nucleus (Lee et al. 2009). Rbfox2 contains an analogous splicing event, raising the possibility that it could undergo similar signal-dependent relocalization to transiently activate or repress the Rbfox2 splicing network.

Our integrative genome-wide analysis has uncovered a novel network of RBPs and other autoregulated factors whose expression is modulated through AS-NMD by Rbfox2. Beyond confirming the context- and motif-dependent effects of Rbfox2 binding on splicing regulation, our results bring to light the simultaneous tunability and robustness of splicing regulatory networks arising from the cross-regulation of negative feedback loops. Importantly, this is likely to be generalizable to other splicing regulators in various physiological contexts.

Materials and methods

iCLIP library preparation

V6.5 mESCs expressing doxycycline-inducible untagged or Flag-HA-tagged human RBFOX2 (UntagFOX2 and FHFOX2) were induced for 24 h with 1 μg/mL doxycycline and cross-linked at 150 mJ/cm^2, collected, and flash-frozen. Cell pellets were used for iCLIP library preparation as described in Konig et al. (2010). Modifications to the protocol are detailed in the Supplemental Material.

iCLIP mapping and clustering

Adapter-trimmed reads from FHFOX2 and UntagFOX2 libraries were uniquely mapped to the University of California at Santa Cruz mm9 build and collapsed on unique 5′ and 3′ ends. Cross-link sites were determined from 5′-terminal nucleotides of contiguously mapping reads and deletion sites of reads mapping with deletions. FHFOX2 cross-links extended by 12 nt on either side were clustered using the Clipper algorithm (available at https://github.com/YeoLab/clipper) (Lovci et al. 2013). FHFOX2 clusters significantly enriched over UntagFOX2 read coverage were considered for further analysis. Details of mapping and clustering are described in the Supplemental Material.

Rbfox2 knockdown and RNA-seq

V6.5 mESCs were infected in biological duplicate with lentivirus expressing control (shLuciferase and shGFP) or Rbfox2 (shFox2-1

and shFox2-2) shRNAs. Cells were harvested after 3 d of selection in 1 μg/mL puromycin followed by 24 h in medium containing no puromycin and RNA was collected using Trizol. Poly(A)$^+$ RNA was prepared for sequencing on the Illumina platform.

RNA-seq analysis

Paired-end reads were aligned to mm9 using TopHat 2.0 (Trapnell et al. 2012). For gene expression analysis, mapped reads were aligned using Cufflinks 2.1.1 (Trapnell et al. 2012) to a custom annotation file generated from annotated and novel junctions called by TopHat from the RNA-seq performed in this study as well as that from Hurt et al. (2013). For splicing analysis, reads were aligned to annotated mouse splicing events using the MISO algorithm (Katz et al. 2010). Significant splicing events were determined by requiring a Bayes' factor >5 and $\Delta\psi$ >0.05 for at least one of the two comparisons of shLuc versus shFox2-1 and shGFP versus shFox2-2, with the direction of splicing change being consistent in the two comparisons. Each event was required to pass the default MISO minimum read coverage thresholds.

Metagene analysis

BedTools (Quinlan and Hall 2010) and custom Python scripts were used to quantify cross-link coverage or conservation with respect to regulated cassette exons. For conservation analysis, mean Phastcons scores across placental mammals were plotted per nucleotide per cassette exon. *P*-values represent rank sum test per nucleotide between bound and expression-matched unbound cassette exon loci. Average cross-link coverage or motif coverage was computed across a 100-nt sliding window and normalized to the total number of exons in each category.

PTC prediction

Custom gene annotations were generated using the RNA-seq libraries generated in this study and in Hurt et al. (2013). Reading frames were annotated by intersecting the custom annotation with Ensembl start codons. Internal exons were categorized as cassette exons based on the presence of annotated transcripts supporting both exon inclusion and exon skipping. Inclusion and exclusion transcript pairs were translated in silico and characterized as always coding (AS-CDS) or NMD upon skipping (NMD-SK) or NMD upon inclusion (NMD-INC) if a premature stop codon was introduced >50 nt upstream of the last exon–exon junction. NMD events were validated by further filtering by gene expression >0.1 FPKM in both control and shUpf1 conditions. MISO annotations were generated for each cassette exon and analyzed using a Bayes' factor >3 and default minimum read coverage for either control or Upf1 knockdown.

siRNA transfection and RBP overexpression

Twenty-five nanomolar AllStars Control siRNA 1 (siCtrl) or siRent1-6 (siUpf1) (Qiagen) were transfected into mESCs expressing shLuc or shFox2-1 48 h prior to harvesting, followed by cell expansion and repeat transfection in suspension with LacZ, Ptbp2, or Tia1 expression vectors 24 prior to harvesting protein and RNA.

AS-NMD gene expression analysis

Predicted AS-NMD splicing events within RBPs or within all genes were filtered for those with >10 reads and greater than twofold differential iCLIP cross-link density between the 300 nt upstream of and downstream from the alternative exon. From this point, only the intron with greater iCLIP density was considered. If the intron with greater density was upstream of an NMD-INC exon or downstream from an NMD-SK exon, the event was categorized as NMD suppressed, while if the intron with greater iCLIP density was downstream from an NMD-INC exon or upstream of an NMD-SK exon, the event was categorized as NMD enhanced. Log_2 fold change in gene expression after Rbfox2 knockdown was plotted for bins of increasing FPKM-normalized iCLIP density.

Accession numbers

RNA-seq and iCLIP data are available under Gene Expression Omnibus accession number GSE54794.

Acknowledgments

We thank members of the Sharp laboratory for insightful discussion and experimental assistance, in particular Andrew Bosson for help with the iCLIP protocol and analysis. We thank Michael Lovci and Gene Yeo for assistance with the Clipper algorithm. This work was supported by United States Public Health Service RO1-GM34277 and Integrative Cancer Biology Program grant U54 CA112967 from the National Institutes of Health to P.A.S., and partially by Koch Institute Support (core) grant P30-CA14051 from the National Cancer Institute. M.J. was supported by the David H. Koch Graduate Fellowship (2011–2012). We acknowledge the service to the Massachusetts Institute of Technology community of the late Sean Collier.

References

Anko ML, Muller-McNicoll M, Brandl H, Curk T, Gorup C, Henry I, Ule J, Neugebauer KM. 2012. The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes. *Genome Biol* **13:** R17.

Auweter SD, Fasan R, Reymond L, Underwood JG, Black DL, Pitsch S, Allain FH. 2006. Molecular basis of RNA recognition by the human alternative splicing factor Fox-1. *EMBO J* **25:** 163–173.

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME suite: Tools for motif discovery and searching. *Nucleic Acids Res* **37:** W202–W208.

Barash Y, Calarco JA, Gao W, Pan Q, Wang X, Shai O, Blencowe BJ, Frey BJ. 2010. Deciphering the splicing code. *Nature* **465:** 53–59.

Becskei A, Serrano L. 2000. Engineering stability in gene networks by autoregulation. *Nature* **405:** 590–593.

Boutz PL, Stoilov P, Li Q, Lin CH, Chawla G, Ostrow K, Shiue L, Ares M Jr, Black DL. 2007. A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes Dev* **21:** 1636–1652.

Colwill K, Pawson T, Andrews B, Prasad J, Manley JL, Bell JC, Duncan PI. 1996. The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. *EMBO J* **15:** 265–275.

Duncan PI, Stojdl DF, Marius RM, Bell JC. 1997. In vivo regulation of alternative pre-mRNA splicing by the Clk1 protein kinase. *Mol Cell Biol* **17:** 5996–6001.

Friedersdorf MB, Keene JD. 2014. Advancing the functional utility of PAR-CLIP by quantifying background binding to mRNAs and lncRNAs. *Genome Biol* **15:** R2.

Gehman LT, Meera P, Stoilov P, Shiue L, O'Brien JE, Meisler MH, Ares M Jr, Otis TS, Black DL. 2012. The splicing regulator Rbfox2 is required for both cerebellar development and mature motor function. *Genes Dev* **26:** 445–460.

Han H, Irimia M, Ross PJ, Sung HK, Alipanahi B, David L, Golipour A, Gabut M, Michael IP, Nachman EN, et al. 2013. MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. *Nature* **498:** 241–245.

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4:** 44–57.

Hurt JA, Robertson AD, Burge CB. 2013. Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. *Genome Res* **23:** 1636–1650.

Hwang J, Maquat LE. 2011. Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: To die or not to die, that is the question. *Curr Opin Genet Dev* **21:** 422–430.

Irimia M, Blencowe BJ. 2012. Alternative splicing: Decoding an expansive regulatory layer. *Curr Opin Cell Biol* **24:** 323–332.

Izquierdo JM, Valcarcel J. 2007. Two isoforms of the T-cell intracellular antigen 1 (TIA-1) splicing factor display distinct splicing regulation activities. Control of TIA-1 isoform ratio by TIA-1-related protein. *J Biol Chem* **282:** 19410–19417.

Jin Y, Suzuki H, Maegawa S, Endo H, Sugano S, Hashimoto K, Yasuda K, Inoue K. 2003. A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG. *EMBO J* **22:** 905–912.

Kalsotra A, Xiao X, Ward AJ, Castle JC, Johnson JM, Burge CB, Cooper TA. 2008. A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. *Proc Natl Acad Sci* **105:** 20333–20338.

Katz Y, Wang ET, Airoldi EM, Burge CB. 2010. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods* **7:** 1009–1015.

Kervestin S, Jacobson A. 2012. NMD: A multifaceted response to premature translational termination. *Nat Rev Mol Cell Biol* **13:** 700–712.

Konig J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J. 2010. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat Struct Mol Biol* **17:** 909–915.

Kuroyanagi H. 2009. Fox-1 family of RNA-binding proteins. *Cell Mol Life Sci* **66:** 3895–3907.

Lareau LF, Brooks AN, Soergel DA, Meng Q, Brenner SE. 2007. The coupling of alternative splicing and nonsense-mediated mRNA decay. *Adv Exp Med Biol* **623:** 190–211.

Lee JA, Tang ZZ, Black DL. 2009. An inducible change in Fox-1/A2BP1 splicing modulates the alternative splicing of downstream neuronal target exons. *Genes Dev* **23:** 2284–2293.

Le Guiner C, Lejeune F, Galiana D, Kister L, Breathnach R, Stevenin J, Del Gatto-Konczak F. 2001. TIA-1 and TIAR activate splicing of alternative exons with weak 5′ splice sites followed by a U-rich stretch on their own pre-mRNAs. *J Biol Chem* **276:** 40638–40646.

Lewis BP, Green RE, Brenner SE. 2003. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci* **100:** 189–192.

Lim LP, Sharp PA. 1998. Alternative splicing of the fibronectin EIIIB exon depends on specific TGCATG repeats. *Mol Cell Biol* **18:** 3900–3906.

Lovci MT, Ghanem D, Marr H, Arnold J, Gee S, Parra M, Liang TY, Stark TJ, Gehman LT, Hoon S et al. 2013. Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. *Nat Struct Mol Biol* **20:** 1434–1442.

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

Nakahata S, Kawamoto S. 2005. Tissue-dependent isoforms of mammalian Fox-1 homologs are associated with tissue-specific splicing activities. *Nucleic Acids Res* **33:** 2078–2089.

Ni JZ, Grate L, Donohue JP, Preston C, Nobida N, O'Brien G, Shiue L, Clark TA, Blume JE, Ares M Jr. 2007. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev* **21:** 708–718.

Nilsen TW, Graveley BR. 2010. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463:** 457–463.

Pan Q, Saltzman AL, Kim YK, Misquitta C, Shai O, Maquat LE, Frey BJ, Blencowe BJ. 2006. Quantitative microarray profiling provides evidence against widespread coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression. *Genes Dev* **20:** 153–158.

Pelisch F, Gerez J, Druker J, Schor IE, Munoz MJ, Risso G, Petrillo E, Westman BJ, Lamond AI, Arzt E, et al. 2010. The serine/arginine-rich protein SF2/ASF regulates protein sumoylation. *Proc Natl Acad Sci* **107:** 16119–16124.

Quinlan AR, Hall IM. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26:** 841–842.

Ray D, Kazan H, Cook KB, Weirauch MT, Najafabadi HS, Li X, Gueroussov S, Albu M, Zheng H, Yang A, et al. 2013. A compendium of RNA-binding motifs for decoding gene regulation. *Nature* **499:** 172–177.

Rossbach O, Hung LH, Schreiner S, Grishina I, Heiner M, Hui J, Bindereif A. 2009. Auto- and cross-regulation of the hnRNP L proteins by alternative splicing. *Mol Cell Biol* **29:** 1442–1451.

Saltzman AL, Pan Q, Blencowe BJ. 2011. Regulation of alternative splicing by the core spliceosomal machinery. *Genes Dev* **25:** 373–384.

Shapiro IM, Cheng AW, Flytzanis NC, Balsamo M, Condeelis JS, Oktay MH, Burge CB, Gertler FB. 2011. An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype. *PLoS Genet* **7:** e1002218.

Sigal A, Milo R, Cohen A, Geva-Zatorsky N, Klein Y, Liron Y, Rosenfeld N, Danon T, Perzov N, Alon U. 2006. Variability and memory of protein levels in human cells. *Nature* **444:** 643–646.

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

Sugimoto Y, Konig J, Hussain S, Zupan B, Curk T, Frye M, Ule J. 2012. Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions. *Genome Biol* **13:** R67.

Teplova M, Song J, Gaw HY, Teplov A, Patel DJ. 2010. Structural insights into RNA recognition by the alternate-splicing regulator CUG-binding protein 1. *Structure* **18:** 1364–1377.

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

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

Underwood JG, Boutz PL, Dougherty JD, Stoilov P, Black DL. 2005. Homologues of the *Caenorhabditis elegans* Fox-1 protein are neuronal splicing regulators in mammals. *Mol Cell Biol* **25:** 10005–10016.

Venables JP, Klinck R, Koh C, Gervais-Bird J, Bramard A, Inkel L, Durand M, Couture S, Froehlich U, Lapointe E, et al. 2009. Cancer-associated regulation of alternative splicing. *Nat Struct Mol Biol* **16:** 670–676.

Venables JP, Brosseau JP, Gadea G, Klinck R, Prinos P, Beaulieu JF, Lapointe E, Durand M, Thibault P, Tremblay K, et al. 2013. RBFOX2 is an important regulator of mesenchymal tissue-specific splicing in both normal and cancer tissues. *Mol Cell Biol* **33:** 396–405.

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456:** 470–476.

Wang ET, Cody NA, Jog S, Biancolella M, Wang TT, Treacy DJ, Luo S, Schroth GP, Housman DE, Reddy S, et al. 2012. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* **150:** 710–724.

Warzecha CC, Sato TK, Nabet B, Hogenesch JB, Carstens RP. 2009. ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing. *Mol Cell* **33:** 591–601.

Weischenfeldt J, Waage J, Tian G, Zhao J, Damgaard I, Jakobsen JS, Kristiansen K, Krogh A, Wang J, Porse BT. 2012. Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns. *Genome Biol* **13:** R35.

Wollerton MC, Gooding C, Wagner EJ, Garcia-Blanco MA, Smith CW. 2004. Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. *Mol Cell* **13:** 91–100.

Wong JJ, Ritchie W, Ebner OA, Selbach M, Wong JW, Huang Y, Gao D, Pinello N, Gonzalez M, Baidya K, et al. 2013. Orchestrated intron retention regulates normal granulocyte differentiation. *Cell* **154:** 583–595.

Xue Y, Zhou Y, Wu T, Zhu T, Ji X, Kwon YS, Zhang C, Yeo G, Black DL, Sun H, et al. 2009. Genome-wide analysis of PTB–RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. *Mol Cell* **36:** 996–1006.

Yeo GW, Coufal NG, Liang TY, Peng GE, Fu XD, Gage FH. 2009. An RNA code for the FOX2 splicing regulator revealed by mapping RNA–protein interactions in stem cells. *Nat Struct Mol Biol* **16:** 130–137.

Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, Sato Y, Sato-Otsubo A, Kon A, Nagasaki M, et al. 2011. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478:** 64–69.

Zhang C, Zhang Z, Castle J, Sun S, Johnson J, Krainer AR, Zhang MQ. 2008. Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2. *Genes Dev* **22:** 2550–2563.