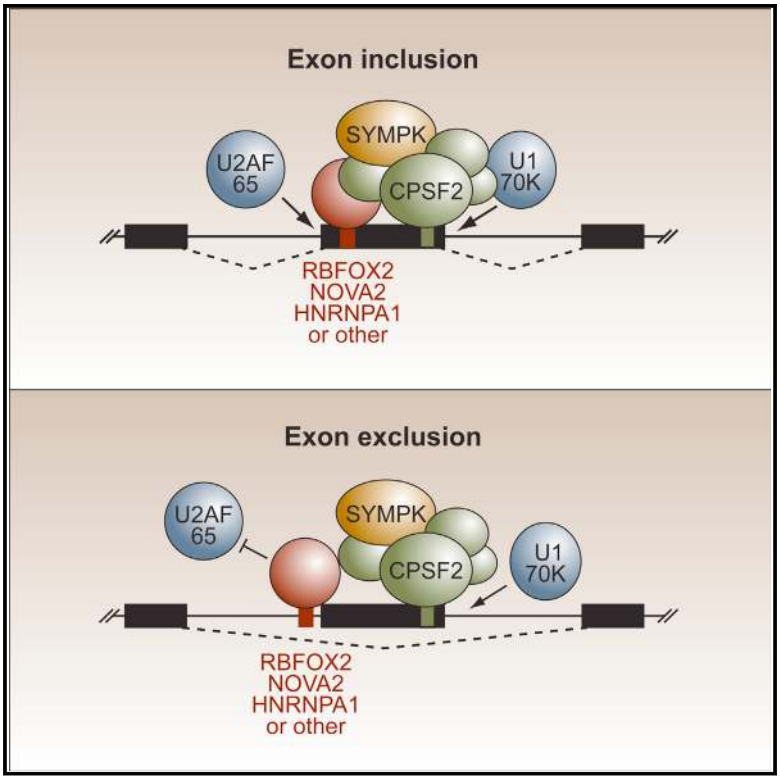
Article

Molecular Cell

Global Promotion of Alternative Internal Exon Usage by mRNA 3′ End Formation Factors

Graphical Abstract



Highlights

- mRNA 3′ end formation factors CPSF and SYMPK are cofactors for splicing regulators
- CPSF/SYMPK is recruited to the pre-mRNA via interaction with the splicing regulator
- CPSF/SYMPK is a global promoter of internal exon usage
- CPSF/SYMPK regulates binding of early intron recognition factors U2AF and U1 snRNP

Authors

Ashish Misra, Jianhong Ou, Lihua J. Zhu, Michael R. Green

Correspondence

michael.green@umassmed.edu

In Brief

The mechanisms that regulate alternative precursor mRNA splicing are largely unknown. Using a large-scale RNAi screen, combined with genome-wide RNA sequencing and crosslinking studies, Misra et al. demonstrate an unanticipated role for the mRNA 3′ end formation factors CPSF and SYMPK in global promotion of alternative splicing.

Accession Numbers

GSE60392

Misra et al., 2015, Molecular Cell *58*, 819–831
June 4, 2015
http://dx.doi.org/10.1016/j.molcel.2015.03.016

CellPress

Molecular Cell
Article

Global Promotion of Alternative Internal Exon Usage by mRNA 3′ End Formation Factors

Ashish Misra,[1,2] Jianhong Ou,[2] Lihua J. Zhu,[2,3] and Michael R. Green[1,2,*]
[1]Howard Hughes Medical Institute
[2]Department of Molecular, Cell and Cancer Biology
[3]Program in Bioinformatics and Integrative Biology
University of Massachusetts Medical School, Worcester, MA 01605, USA
*Correspondence: michael.green@umassmed.edu
http://dx.doi.org/10.1016/j.molcel.2015.03.016

SUMMARY

The mechanisms that regulate alternative precursor mRNA (pre-mRNA) splicing are largely unknown. Here, we perform an RNAi screen to identify factors required for alternative splicing regulation by RBFOX2, an RNA-binding protein that promotes either exon inclusion or exclusion. Unexpectedly, we find that two mRNA 3′ end formation factors, cleavage and polyadenylation specificity factor (CPSF) and SYMPK, are RBFOX2 cofactors for both inclusion and exclusion of internal exons. RBFOX2 interacts with CPSF/SYMPK and recruits it to the pre-mRNA. RBFOX2 and CPSF/SYMPK then function together to regulate binding of the early intron recognition factors U2AF and U1 small nuclear ribonucleoprotein particle (snRNP). Genome-wide analysis reveals that CPSF also mediates alternative splicing of many internal exons in the absence of RBFOX2. Accordingly, we show that CPSF/SYMPK is also a cofactor of NOVA2 and heterologous nuclear ribonucleoprotein A1 (HNRNPA1), RNA-binding proteins that also regulate alternative splicing. Collectively, our results reveal an unanticipated role for mRNA 3′ end formation factors in global promotion of alternative splicing.

INTRODUCTION

The vast majority of structural genes in higher eukaryotes contain intervening sequences (introns) that are removed from the precursor mRNAs (pre-mRNAs) by splicing. In ~95% of human genes, splice sites can be differentially selected to produce distinct mRNA and protein isoforms from the same pre-mRNA, a process called alternative splicing. Alternative splicing plays important roles in diversifying the proteome and controlling gene expression (Keren et al., 2010; Kornblihtt et al., 2013; Nilsen and Graveley, 2010). The importance of understanding alternative splicing regulation is underscored by its well-established roles in multiple biological processes (Irimia and Blencowe, 2012; Kalsotra and Cooper, 2011). For example, differential expression of mRNA isoforms is important for development (Kalsotra and Cooper, 2011), and alterations in alternative splicing can contribute to the initiation or progression of cancer and other diseases (Cooper et al., 2009).

The mechanisms that regulate alternative splicing, particularly on a genome-wide level, remain largely unknown. Regulation of splice site selection is thought to primarily occur at the earliest stages of the assembly pathway by RNA-binding proteins that either promote or repress the use of core splicing signals (Chen and Manley, 2009). RBFOX2 (also called FOX-2) is an RNA-binding protein that is conserved from worms to humans and specifically recognizes the RNA element UGCAUG (Underwood et al., 2005; Zhang et al., 2008). Binding of RBFOX2 within or 5′ to the alternative exon causes exclusion (or skipping), whereas binding to the 3′ intron promotes inclusion (Chen and Manley, 2009; Huh and Hynes, 1994; Venables et al., 2009; Yeo et al., 2009; Zhang et al., 2008). The basis of this differential, location-dependent regulation remains to be determined.

In addition to pre-mRNA splicing, RNA-binding proteins also function at multiple steps during gene expression, including transcriptional elongation (Lin et al., 2008), mRNA 3′ end formation (Chan et al., 2011), and mRNA nuclear export (Blanchette et al., 2004). Here, we describe a large-scale RNAi screen to identify factors required for splicing regulation by RBFOX2. Our results reveal an unanticipated role for mRNA 3′ end formation factors in global promotion of alternative splicing.

RESULTS

A Large-Scale shRNA Screen Reveals CPSF as an RBFOX2 Cofactor

To identify possible cofactors required for RBFOX2 to repress splicing, we performed a large-scale small hairpin RNA (shRNA) screen (Figure 1A) based upon a previously described three-exon mini-gene reporter for exon exclusion (Wang et al., 2004). Exons 1 and 3 form a complete mRNA encoding GFP. Exon 2 is normally included to form an mRNA that does not encode functional GFP (GFP−). However, insertion of a binding site for a splicing repressor into exon 2 causes skipping of this exon, producing an mRNA encoding functional GFP (GFP+). We inserted an RBFOX2-binding site (Lim and Sharp, 1998; Venables et al., 2009; Yeo et al., 2009; Zhang et al., 2008) into exon 2 and derived a Flp-In-293 cell line in which this reporter construct was integrated at a single site (GFP/Flp-In-293 cells).

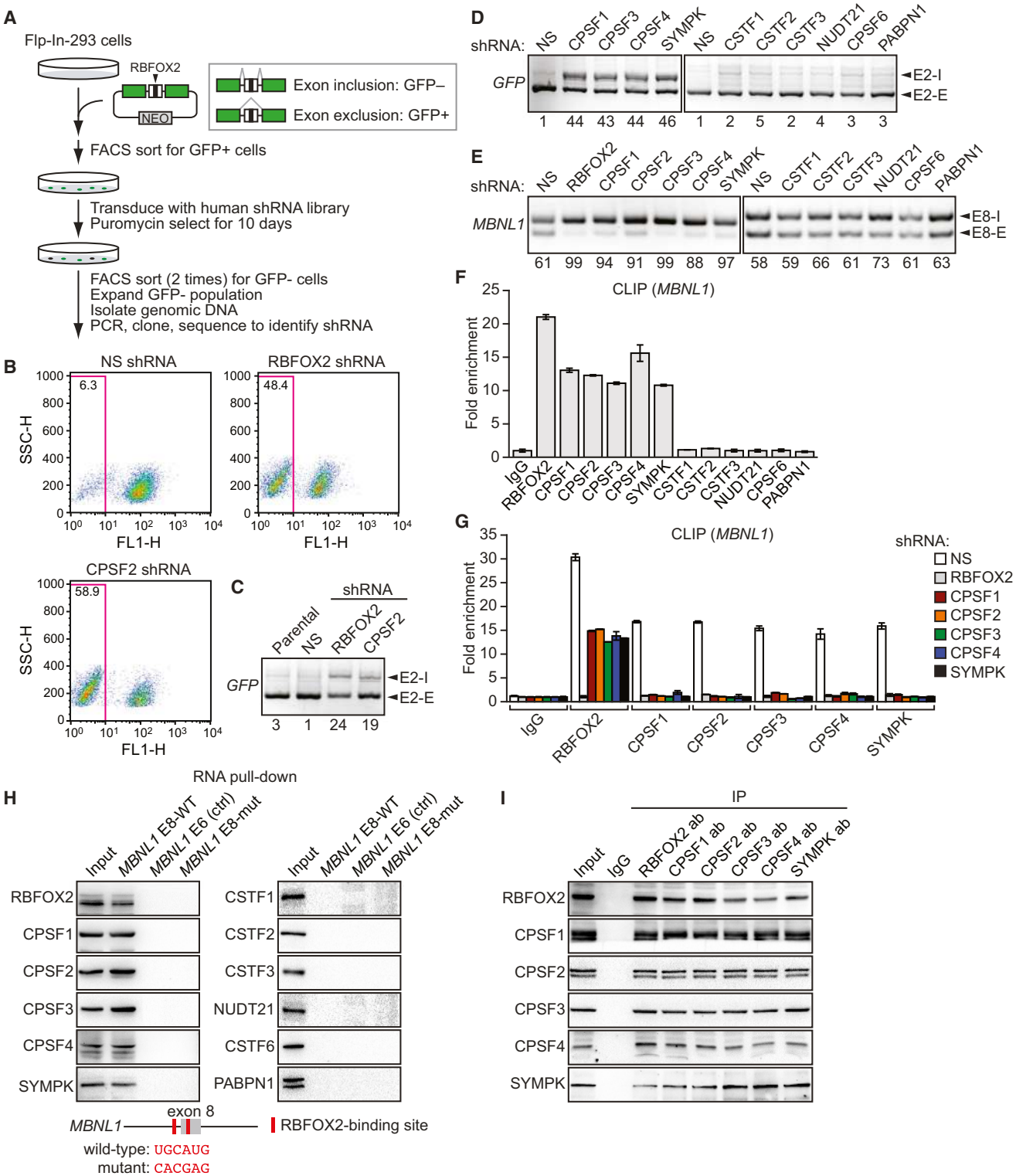


Figure 1. A Large-Scale shRNA Screen Reveals CPSF as an RBFOX2 Cofactor

(A) Schematic of the screen.

(B) FACS analysis showing the percentage of GFP− cells in GFP/Flp-In-293 cells expressing an NS, RBFOX2, or CPSF2 shRNA.

(C) RT-PCR analysis of *GFP* splicing in parental (FACS-sorted) GFP/Flp-In-293 cells, and cells expressing an NS, RBFOX2, or CPSF2 shRNA. The positions of mRNAs in which exon 2 is included (E2-I) or excluded (E2-E) are shown. The percentage of the signal corresponding to exon inclusion is shown.

(legend continued on next page)

Fluorescence-activated cell sorting (FACS) was used to derive a population of cells that was >95% GFP+ (Figure S1A).

GFP/Flp-In-293 cells were transduced with lentiviral shRNA pools from the RNAi Consortium (TRC)-Hs1.0 human shRNA library, GFP− cells were isolated by two rounds of FACS selection and expanded, and shRNA candidates were identified by DNA sequencing. To avoid indirect effects, we prioritized candidates that were implicated in RNA-related processes and nuclear localized (Table S1). Unexpectedly, one of the candidates was cleavage and polyadenylation specificity factor 2 (CPSF2) (also called CPSF100), a component of the multi-subunit CPSF complex that catalyzes the cleavage step of mRNA 3′ end formation (Chan et al., 2011; Colgan and Manley, 1997; Mandel et al., 2008). Because of the well-established role of CPSF in pre-mRNA processing, we elected to focus on CPSF2. Validation experiments confirmed that knockdown of RBFOX2 or CPSF2 in the GFP/Flp-In-293 reporter cell line resulted in a substantially increased percentage of GFP− cells (Figure 1B) and increased exon 2 inclusion (Figure 1C) relative to that obtained using a control non-silencing (NS) shRNA. Notably, knockdown of CPSF2 did not affect RBFOX2 levels and vice versa (Figures S1B and S1C).

We next tested the role of the three other CPSF subunits as well as additional mRNA 3′ end formation factors. Figure 1D shows that CPSF1 (also called CPSF160), CPSF3 (also called CPSF73), and CPSF4 (also called CPSF30) and SYMPK, a protein that can stably associate with CPSF subunits (Sullivan et al., 2009), were required for efficient RBFOX2-directed alternative splicing of the GFP reporter. By contrast, knockdown of other known mRNA 3′ end formation factors, including cleavage and stimulation factors CSTF1–3, NUDT21 (also called CPSF5), CPSF6, and the poly(A)-binding protein PABPN1, had no effect on splicing. Similar results were obtained using a second, unrelated shRNA to CPSF1–4 and SYMPK (Figure S1D). Knockdown efficiency of each shRNA was confirmed by qRT-PCR (Figure S1E) and immunoblot analyses (Figure S1F). Notably, like CPSF2, knockdown of CPSF1, 3, or 4 or SYMPK did not affect RBFOX2 levels (Figures S1C and S1F). In addition, knockdown of either CPSF2 or RBFOX2 did not affect expression of CPSF1, 3, or 4 or SYMPK (Figure S1G).

To confirm the role of CPSF and SYMPK (CPSF/SYMPK) on RBFOX2-directed splicing repression, we analyzed splicing of an endogenous pre-mRNA, *MBNL1* exon 8, whose exclusion is promoted by RBFOX2 (Venables et al., 2009; Zhang et al., 2008). Knockdown of RBFOX2, CPSF1–4, or SYMPK, but not other mRNA 3′ end formation factors, substantially increased *MBNL1* exon 8 inclusion (Figures 1E and S1H).

We next asked whether CPSF was bound to the pre-mRNA using a UV crosslinking/immunoprecipitation (CLIP) assay (Jensen and Darnell, 2008). For this CLIP experiment and those described below, following immunoprecipitation, RNA was size selected to ~200–250 base pairs (bp) followed by qRT-PCR using primers designed to amplify a region within 200 bp upstream (for an exclusion event) or downstream (for an inclusion event) of the alternatively spliced exon. In this particular experiment, primers were designed to amplify the intron upstream of *MBNL1* exon 8, which contains an RBFOX2-binding site (Venables et al., 2009; Zhang et al., 2008). Figure 1F shows that RBFOX2, CPSF1–4, and SYMPK were bound to the intron upstream of *MBNL1* exon 8. By contrast, other mRNA 3′ end formation factors were not associated with this same *MBNL1* pre-mRNA region, although, as expected, they were associated with the *MBNL1* cleavage and polyadenylation site (Figure S1I). Knockdown of RBFOX2 resulted in complete loss of CPSF1–4 and SYMPK binding (Figure 1G). Notably, binding of RBFOX2 was reduced ~50%–60% following knockdown of CPSF1–4 or SYMPK. Interestingly, depletion of any CPSF subunit or SYMPK resulted in loss of binding of all other factors, which may be explained by disruption of a CPSF/SYMPK complex.

To confirm the CLIP results, we performed an in vitro RNA pull-down assay. Briefly, a biotinylated RNA substrate containing *MBNL1* exon 8 plus ~100 bp of flanking intron sequence was bound to streptavidin beads and incubated in Flp-In-293 nuclear extract. Proteins retained after stringent washing were eluted and analyzed by immunoblotting. Figure 1H shows that the RNA substrate was bound by RBFOX2, CPSF1–4, and SYMPK, but not other mRNA 3′ end formation factors. Notably, there was no detectable binding of RBFOX2, CPSF1–4, and SYMPK to a biotinylated RNA substrate that contained constitutively spliced *MBNL1* exon 6 or an *MBNL1* exon 8 derivative in which the RBFOX2-binding site was mutated.

To delineate the region of the CPSF2 protein that was required for its RNA-binding activity, we constructed a series of CPSF2 truncation mutants and tested them for RNA binding in an in vitro RNA pull-down assay. The results show that the RNA-binding activity of CPSF2 requires the metallo-β-lactamase (MβL) and β-CASP motifs, but not the C-terminal domain (Figure S1J).

The above-mentioned results raised the possibility that CPSF/SYMPK was associated with the pre-mRNA due to a direct interaction with RBFOX2. Consistent with this possibility,

(D and E) RT-PCR analysis of *GFP* splicing in GFP/Flp-In-293 cells (D) and of endogenous *MBNL1* in Flip-In-293 cells (E) expressing a control NS shRNA, or an shRNA targeting an mRNA 3′ end formation factor.

(F) CLIP analysis monitoring binding of RBFOX2, CPSF1–4, SYMPK, and other mRNA 3′ end formation factors to *MBNL1* exon 8. Data are represented as mean ± SD.

(G) CLIP analysis monitoring binding of RBFOX2, CPSF1–4, and SYMPK to *MBNL1* exon 8 in cells expressing an NS, RBFOX2, CPSF1–4, or SYMPK shRNA. Data are represented as mean ± SD.

(H) RNA pull-down assay. A biotinylated RNA containing *MBNL1* regulated exon 8 (E8-WT), an exon 8 derivative containing a mutation in the RBFOX2-binding site (E8-mut) or, as a control, constitutively spliced exon 6 (E6 [ctrl]) as well as 100 bp of upstream and downstream intron sequences was incubated in nuclear extract and analyzed for bound proteins by immunoblotting.

(I) Co-immunoprecipitation analysis. Cell extracts were immunoprecipitated in the presence of RNase with an RBFOX2, CPSF1–4, or SYMPK antibody, or as a control with IgG, and immunoblotted for each protein.

See also Figure S1.

co-immunoprecipitation from nuclear extract shows that RBFOX2, CPSF1–4, and SYMPK are stably associated with one another (Figure 1I). In this and the co-immunoprecipitation experiments presented below, co-immunoprecipitations were performed in the presence of RNase to rule out the possibility that the apparent protein-protein interactions were mediated by RNA. We could also demonstrate co-immunoprecipitation of in vitro translated CPSF1–4 and SYMPK (Figure S1K), ruling out the possibility of a "bridging" interaction by some other nuclear protein.

Genome-wide Analysis of RBFOX2- and CPSF2-Mediated Changes in Alternative Splicing of Internal Exons

To determine the role of CPSF in alternative splicing on a genome-wide level, we performed paired end RNA sequencing (RNA-seq) to compare splicing events in control and CPSF2 knockdown cells. RBFOX2 knockdown cells were analyzed in parallel. The RNA-seq experiments were performed using two biological replicates. Each sample produced 78–110 million paired-end reads, >85% of which mapped uniquely to the human genome. RNA-seq analysis was performed using DEXSeq, a statistical method to test for differential exon usage in RNA-seq (Anders et al., 2012). The biological replicates demonstrated a high level of reproducibility for both differential gene expression, as determined by reads per kilobase of transcript per million mapped reads (RPKM) ($R^2 > 0.99$) (Figure S2A), and alternative splicing ($R^2 > 0.97$) (Figure S2B). Also, comparison of RPKM between knockdown cell lines showed comparable gene expression ($R^2 > 0.96$) (Figure S2C). We excluded terminal exons, whose splicing could be altered due to defective mRNA 3′ end formation (Kyburz et al., 2006; Martinson, 2011) following RBFOX2 or CPSF2 knockdown. We also used Cufflinks, a software program that tests for differential expression and regulation in RNA-seq samples (Trapnell et al., 2013), to eliminate genes whose expression levels were altered by RBFOX2 or CPSF2 knockdown.

The RNA-seq results show, as expected, that RBFOX2 promoted exclusion of some internal exons (5,106) and inclusion of other internal exons (754) (Figure 2A; Table S2). Likewise, CPSF2 also promoted both internal exon exclusion (9,081) and inclusion (2,326) (Figure 2A; Table S3). Notably, ~67% of the splicing events that were altered following RBFOX2 knockdown were comparably affected by CPSF2 knockdown (Figure 2A; Table S4). In addition to these RBFOX2-regulated, CPSF2-regulated events, the RNA-seq results revealed two additional categories of alternative splicing events: RBFOX2-regulated, CPSF2-independent and RBFOX2-independent, CPSF2-regulated, which are analyzed in greater detail below.

Identification of Genome-wide Binding Sites for RBFOX2 and CPSF2

To identify the direct genome-wide targets of RBFOX2 and CPSF2, we performed genome-wide individual-nucleotide resolution UV crosslinking/immunoprecipitation (iCLIP) assays (König et al., 2010) in Flp-In-293 cells. For each factor, two independent replicate iCLIP experiments were performed (Figure S2D).

Unexpectedly, only ~18% of the CPSF2 iCLIP tags are located in the 3′ UTR (Figure 2B), suggesting that CPSF2 has a major role in RNA processing events other than mRNA 3′ end formation. Consistent with this idea, analysis of the iCLIP data also revealed that CPSF2 does not have a distinct binding peak near the cleavage and polyadenylation site (Figures S2E and S2F). Notably, there was a significant correlation between our CPSF2 results and a published CPSF1 CLIP-seq dataset (Martin et al., 2012) (Figure S2G).

To identify the sequence motifs that are enriched in RBFOX2 and CPSF2 clusters, we used the motif analysis program HOMER (Heinz et al., 2010). As expected, the top motif based on p value for RBFOX2 was the well-known RBFOX2 motif UGCAUG (p = 1e-265), whereas the top motif for CPSF2 was AGGUAG (p = 1e-232) (Figure 2C). Bioinformatic analysis revealed that there was no substantial enrichment of the CPSF2 motif near cleavage and polyadenylation sites (Figure S2H).

To determine whether a CPSF subunit or SYMPK binds to the AGGUAG motif, we performed RNA pull-down assays using a biotinylated RNA substrate and in vitro translated proteins. In this assay, the biotinylated substrate was a *β-globin* exon 2 derivative containing two inserted AGGUAG motifs. The results of Figure 2D show that CPSF2, but not CPSF1, 3, or 4 or SYMPK, bound to the AGGUAG-containing RNA substrate, but not to a control biotinylated *β-globin* exon 2 RNA that lacked the motif.

Validation of the RNA-Seq Results

To validate these RNA-seq results, we analyzed by RT-PCR several representative alternatively spliced exons whose inclusion or exclusion was promoted by both RBFOX2 and CPSF2. For all RBFOX2-regulated, CPSF2-regulated exons analyzed, knockdown of RBFOX2, CPSF1–4, or SYMPK had a comparable effect on alternative splicing (Figures 3A, 3B, S3A, and S3B). By contrast, knockdown of CSTF3, an essential cleavage and polyadenylation factor that is not a CPSF/SYMPK subunit, had no effect on splicing of these pre-mRNAs (Figures S3A and S3B) or any of the other pre-mRNAs analyzed below. The corresponding iCLIP tracks revealed multiple RBFOX2- and CPSF2-binding peaks in or adjacent to the alternatively spliced exon (Figures S3A and S3B). CLIP analysis showed that binding of RBFOX2 and CPSF2 was substantially enriched within 200 bp downstream (Figure 3C) or upstream (Figure 3D) of the alternatively spliced exon. For other components of the CPSF/SYMPK complex, CLIP analysis revealed that they were also bound to the same region of the pre-mRNA (Figures 3C and 3D). Analysis of a representative RBFOX2-regulated, CPSF2-regulated exon, *MEF2A* exon 8, revealed that knockdown of RBFOX2 resulted in complete loss of CPSF1–4 and SYMPK binding (Figure 3E). Notably, binding of RBFOX2 was reduced ~50%–60% following depletion of CPSF/SYMPK.

In addition to the alternatively spliced exons that were regulated by both RBFOX2 and CPSF2, the RNA-seq results identified a large number of CPSF2-regulated exons (11,407), ~66% of which appeared to be RBFOX2 independent (see Figure 2A). RT-PCR analysis of several of these alternatively spliced exons confirmed in all cases that knockdown of CPSF1–4 or SYMPK had the expected effect on exon inclusion or exclusion, whereas RBFOX2 knockdown had no effect (Figures 4A, 4B, S4A, and S4B). Consistent with these results, CLIP analysis showed that

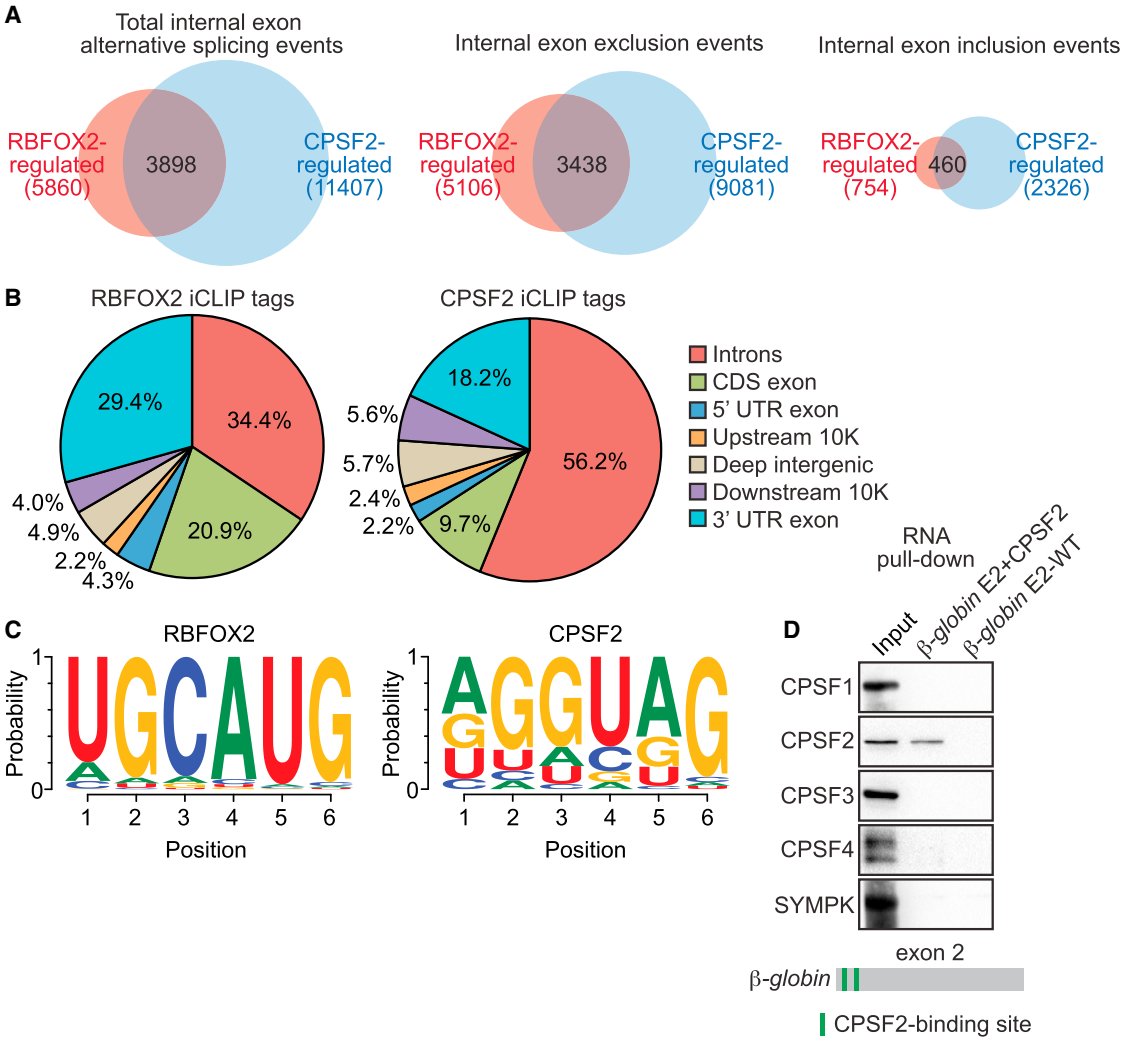


Figure 2. Genome-wide Analysis of RBFOX2- and CPSF2-Mediated Changes in Alternative Splicing of Internal Exons

(A) Venn diagrams showing the overlap between total (left), exon exclusion (middle), and exon inclusion (right) RBFOX2-regulated and CPSF2-regulated internal splicing events.

(B) Distribution of RBFOX2 and CPSF2 iCLIP tags.

(C) Top binding motifs, based on p values, from RBFOX2 and CPSF2 iCLIP analysis.

(D) RNA pull-down assay. A biotinylated RNA substrate containing *β-globin* exon 2 harboring (E2+CPSF2) or lacking (E2-WT; control) two AGGUAG motifs was incubated with in vitro translated proteins and analyzed for bound proteins by immunoblotting.

See also Figures S2 and S7.

CPSF1–4 and SYMPK were bound near the regulated exons, whereas RBFOX2 binding was not detected (Figures 4C and 4D).

Finally, the RNA-seq results also revealed that ~33% of RBFOX2-regulated exons were CPSF2 independent (see Figure 2A). Analysis of representative alternatively spliced exons revealed that in all cases, knockdown of RBFOX2 had the predicted effect on exon inclusion or exclusion, whereas knockdown of CPSF1–4 or SYMPK had no effect (Figures S4C and S4D). CLIP assays revealed that RBFOX2, but not CPSF1–4 or SYMPK, were bound adjacent to the regulated exon (Figure S4E). Furthermore, in contrast to what we found for RBFOX2-regulated, CPSF/SYMPK-regulated exons, RBFOX2 binding was unaffected by depletion of CPSF1–4 or SYMPK (Figure S4F).

RBFOX2 and CPSF/SYMPK Promote Exon Exclusion and Inclusion by Regulating Binding of U2AF and U1 snRNP

One mechanism to promote alternative splicing is by regulating binding of the early intron recognition factors U2AF, which binds to the Py tract/3′ splice site, and U1 small nuclear ribonucleoprotein particle (snRNP), which binds to the 5′ splice site (Chen and Manley, 2009). To gain insight into how CPSF functions with RBFOX2 to modulate splicing, we used a CLIP assay to measure binding of the U2AF subunit U2AF65 and the U1 snRNP subunit U1 70K in the upstream intron or the downstream intron, respectively, relative to the alternatively spliced exon. We first measured U2AF65 and U1 70K binding to exon 8 of the *MEF2A* pre-mRNA, whose inclusion is promoted by RBFOX2

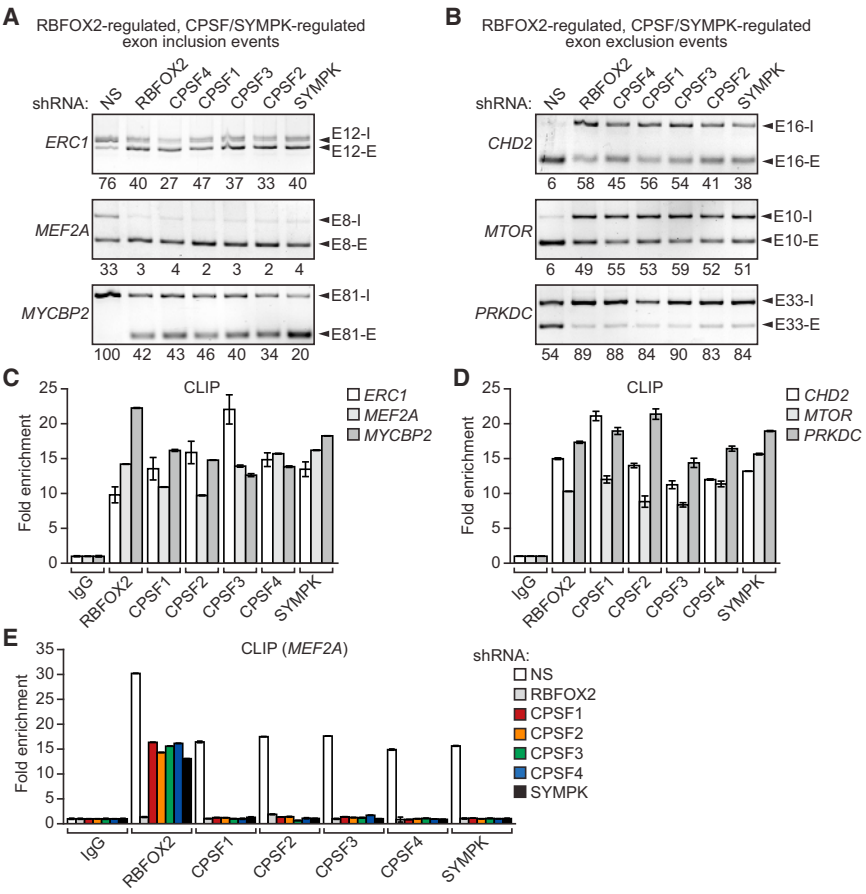


Figure 3. Validation of RNA-Seq Results for RBFOX2-Regulated, CPSF/SYMPK-Regulated Exons

(A and B) RT-PCR analysis monitoring inclusion (A) or exclusion (B) of representative exons in cells expressing an NS, RBFOX2, CPSF1–4, or SYMPK shRNA.

(C and D) CLIP analysis monitoring binding of RBFOX2, CPSF1–4, and SYMPK to the regulated exons of the indicated pre-mRNAs. Data are represented as mean ± SD.

(E) CLIP analysis monitoring binding of RBFOX2, CPSF1–4, and SYMPK to *MEF2A* exon 8 in cells expressing an NS, RBFOX2, CPSF1–4, or SYMPK shRNA. Data are represented as mean ± SD.

See also Figure S3.

and CPSF (see Figure 3A). Figure 5A shows that in control cells expressing an NS shRNA, binding of U2AF65 and U1 70K could be detected at the introns adjacent to both the regulated exon 8 and the constitutively spliced exon 5. Knockdown of either RBFOX2 or CPSF2 resulted in decreased binding of U2AF65 and U1 70K at the intron adjacent to *MEF2A* exon 8, but had no effect on binding at the intron adjacent to the constitutive exon 5. Similar results were obtained for nine other pre-mRNAs whose inclusion is promoted by RBFOX2 and CPSF (Figure S5A; Table S4).

We performed a comparable analysis on *MBNL1* exon 8, whose exclusion is promoted by RBFOX2 and CPSF (see Figure 1E). Figure 5B shows that in control NS shRNA cells, U2AF65 binding was not detected at the upstream intron adjacent to regulated *MBNL1* exon 8, but was readily detected at the intron upstream of constitutively spliced exon 6. Notably, knockdown of either RBFOX2 or CPSF2 substantially increased binding of U2AF65 to the intron upstream of regulated exon 8. By contrast, U1 70K binding was detected at both the downstream intron adjacent to regulated exon 8 and the intron upstream of constitutively spliced exon 6, and was not affected by RBFOX2 or CPSF2 knockdown. Similar results were obtained for nine other pre-mRNAs whose exclusion is promoted by RBFOX2 and CPSF (Figure S5B; Table S4).

We also analyzed U2AF65 and U1 70K binding for two RBFOX2-regulated, CPSF/SYMPK-independent pre-mRNAs. Figure S5C shows that knockdown of RBFOX2 substantially decreased binding of U2AF65 and U1 70K at *DIAPH1* exon 2, whose inclusion is promoted by RBFOX2 (see Figure S4C). By contrast, Figure S5D shows that RBFOX2 knockdown increased binding of U2AF65, but had no effect on U1 70K binding at *AGRN* exon 7, whose exclusion is promoted by RBFOX2 (see Figure S4D). As expected, for both exons, knockdown of CPSF2 had no effect (Figures S5C and S5D).

Finally, we analyzed U2AF65 and U1 70K binding for two CPSF/SYMPK-regulated, RBFOX2-independent pre-mRNAs. Figure S5E shows that knockdown of CPSF2 decreased binding of U2AF65 and U1 70K at the intron adjacent to *PDHX* exon 11, whose inclusion is promoted by CPSF/SYMPK (see Figure 4A). By contrast, Figure S5F shows that CPSF2 knockdown substantially increased binding of U2AF65 to the intron upstream of *RALGAPA1* exon 4, whose exclusion is promoted by CPSF/SYMPK (see Figure 4B). As expected, for both exons, knockdown of RBFOX2 had no effect (Figures S5E and S5F).

RBFOX2 and CPSF/SYMPK Promote Recruitment of U1 snRNP to the 5′ Splice Site, Leading to Exon Inclusion

We next performed several experiments to understand how RBFOX2 and the CPSF/SYMPK complex promote exon inclusion. The co-immunoprecipitation results show that RBFOX2 and CPSF2 are stably associated with U1 70K, but not U2AF65 (Figure 5C). To determine whether RBFOX2 and CPSF/SYMPK affected recruitment of U1 snRNP or U2AF to the pre-mRNA, a biotinylated RNA substrate containing *MEF2A* exon 8 was bound to streptavidin beads and incubated with nuclear extracts derived from RBFOX2 or CPSF2 knockdown Flp-In-293 cells. Proteins retained after stringent washing were eluted and analyzed by immunoblotting. The results show that knockdown of RBFOX2 or CPSF2 reduced association of U2AF65 and U1 70K with the RNA (Figure 5D).

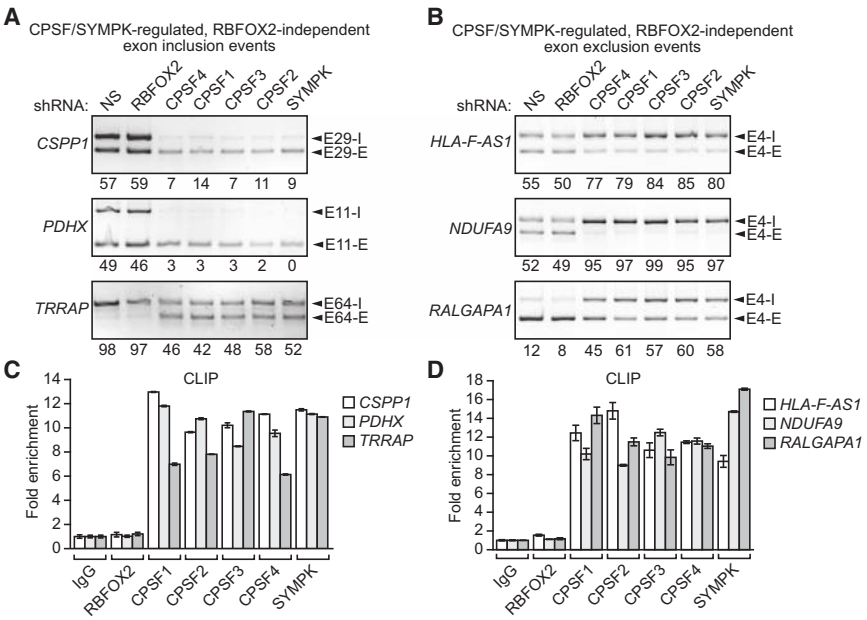


Figure 4. Validation of RNA-Seq Results for CPSF/SYMPK-Regulated, RBFOX2-Independent Exons

(A and B) RT-PCR analysis monitoring inclusion (A) or exclusion (B) of representative exons in cells expressing an NS, RBFOX2, CPSF1–4, or SYMPK shRNA.

(C and D) CLIP analysis monitoring binding of RBFOX2, CPSF1–4, and SYMPK to the regulated exons of the indicated pre-mRNAs. Data are represented as mean ± SD.

See also Figure S4.

To confirm and extend these results, we performed RNA pull-down assays on a control wild-type *MEF2A* exon 8 pre-mRNA substrate and mutant *MEF2A* exon 8 pre-mRNA substrates that lacked either the RBFOX2-binding site, the CPSF2-binding site, the 5′ splice site, or the Py tract/3′ splice site. Figure 5E shows that mutation of the RBFOX2-binding site reduced binding of RBFOX2, as expected, and also substantially decreased binding of CPSF2. Mutation of the CPSF2-binding site led to reduced binding of CPSF2 as well as RBFOX2, consistent with the results of our CLIP experiments (Figure 3E). Mutation of the RBFOX2- or CPSF2-binding site also resulted in decreased binding of U1 70K and U2AF65, consistent with our CLIP results (Figure 5A). Binding of U1 70K was lost following mutation of the 5′ splice site, whereas mutation of the Py tract/3′ splice site had little effect. Notably, however, binding of U2AF65 was abolished by mutation of either the 5′ splice site or Py tract/3′ splice site. Collectively, these results suggest a model in which RBFOX2, in conjunction with CPSF/SYMPK, mediates recruitment of U1 snRNP, which in turn promotes the recruitment of U2AF, resulting in exon inclusion.

RBFOX2 and CPSF/SYMPK Sterically Interfere with U2AF65 Binding, Leading to Exon Exclusion

We next studied the role of RBFOX2 and the CPSF/SYMPK complex in exon exclusion. The RNA pull-down assay of Figure 5D shows that in control nuclear extracts, there was no detectable binding of U2AF65 to the intron preceding excluded *MBNL1* exon 8. By contrast, in nuclear extracts derived from RBFOX2 or CPSF2 knockdown cells, binding of U2AF65 was readily detectable.

To confirm and extend these results, we performed RNA pull-down assays on a control wild-type *MBNL1* exon 8 pre-RNA substrate and mutant *MBNL1* exon 8 RNA substrates that lacked either the RBFOX2-binding site, the CPSF2-binding site, the 5′ splice site, or the Py tract/3′ splice site. Figure 5F shows that mutation of the RBFOX2-binding site eliminated binding of RBFOX2, as expected, and also reduced binding of CPSF2. Mutation of the CPSF2-binding site eliminated CPSF2 binding and also reduced binding of RBFOX2. Notably, mutation of either the RBFOX2- or CPSF2-binding site resulted in increased binding of U2AF65. Binding of U1 70K was substantially decreased by mutation of the 5′ splice site, but relatively unaffected by mutation of the 3′ splice site. However, binding of U2AF65 was abolished by mutation of either the 5′ splice site or Py tract/3′ splice site. These results suggest that RBFOX2 and the CPSF/SYMPK complex interfere with binding of U2AF to the intron preceding an excluded exon.

CPSF/SYMPK Is Also a Cofactor for the Splicing Regulators NOVA2 and HNRNPA1

We hypothesized that for CPSF/SYMPK-regulated, RBFOX2-independent exons there were other factors that functionally substituted for RBFOX2. To identify such candidate splicing factors, we used SFmap, which predicts and maps known splicing factor-binding sites on RNA sequences (Paz et al., 2010), to analyze the occurrence of motifs in CPSF2-regulated exons. This analysis identified several motifs, including those for the well-studied splicing regulators NOVA2 and heterologous nuclear ribonucleoprotein A1 (HNRNPA1), that were enriched in exons or flanking introns as compared to control constitutively spliced introns or exons (Figure S6A).

Analysis of the list of CPSF/SYMPK-regulated, RBFOX2-independent exons revealed several genes that are known targets of NOVA2 (Tollervey et al., 2011; Ule et al., 2006) or HNRNPA1 (Huelga et al., 2012). For two of the NOVA2-regulated pre-mRNAs, knockdown of NOVA2 (Figure S6B), CPSF1–4, or SYMPK affected splicing of the regulated exon comparably (Figures 6A and S6C). Knockdown of CPSF1–4 or SYMPK did not affect *NOVA2* expression (Figure S6D). CLIP analysis confirmed that NOVA2, CPSF1–4, and SYMPK were substantially enriched near the alternatively spliced exons (Figure 6B). Furthermore, recruitment of CPSF/SYMPK was dependent on NOVA2, and NOVA2 binding was reduced by depletion of CPSF/SYMPK (Figure 6C), similar to what we found for RBFOX2 (Figure 1G). Analogous results were

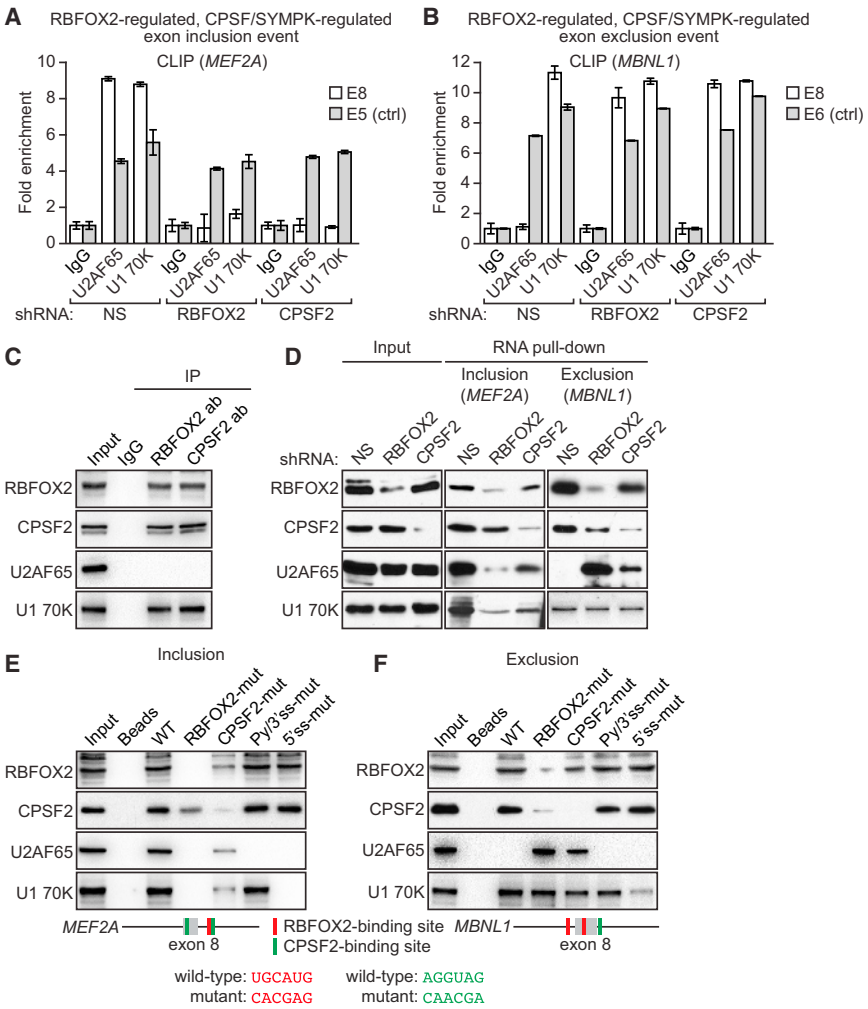


Figure 5. RBFOX2 and CPSF/SYMPK Promote Exon Exclusion and Inclusion by Regulating Binding of U2AF and U1 snRNP

(A and B) CLIP analysis monitoring binding of U2AF65 and U1 70K to regulated and constitutive exons in *MEF2A* (A) and *MBNL1* (B) in cells expressing an NS, RBFOX2, or CPSF2 shRNA. Data are represented as mean ± SD.

(C) Co-immunoprecipitation analysis. Cell extracts were immunoprecipitated in the presence of RNase with an RBFOX2 or CPSF2 antibody, or as a control with IgG, and immunoblotted for RBFOX2, CPSF2, U2AF65, or U1 70K.

(D) RNA pull-down assays monitoring binding of RBFOX2, CPSF2, U2AF65, and U1 70K to a biotinylated RNA substrate containing *MEF2A* exon 8 (inclusion) or *MBNL1* exon 8 (exclusion) in extracts prepared from Flp-In-293 cells expressing an NS, RBFOX2, or CPSF2 shRNA.

(E and F) RNA pull-down assays monitoring binding of RBFOX2, CPSF2, U2AF65, and U1 70K to a biotinylated RNA substrate containing a wild-type *MEF2A* exon 8 (E) or *MBNL1* exon 8 (F) or mutant derivative in which the RBFOX2-binding site, CPSF2-binding site, Py tract/3′ splice site, or 5′ splice site was mutated.

See also Figure S5.

obtained for two pre-mRNAs that were regulated by HNRNPA1 (Figures 6D–6F and S6E–S6G).

NOVA2 and HNRNPA1 Promote Exon Inclusion and Exclusion by a Mechanism Similar to That of RBFOX2

The finding that CPSF/SYMPK was a cofactor for RBFOX2, NOVA2, and HNRNPA1 suggested that the three proteins might promote exon exclusion by a common mechanism. In support of this possibility, the RNA pull-down experiments of Figure 7A show that, like RBFOX2 (see Figure 1I), NOVA2 and HNRNPA1 could recruit CPSF/SYMPK to a regulated exon, consistent with the CLIP results described above (see Figures 6D and 6G). Moreover, the co-immunoprecipitation results of Figures 7B and 7C show that both NOVA2 and HNRNPA1 stably associate with CPSF/SYMPK in the absence of RNA. Finally, the CLIP results of Figures 7D and 7E show that NOVA2 and HNRNPA1, in conjunction with CPSF, promote binding of U1 70K and U2AF65 on included exons, and interfere with binding of U2AF65 to excluded exons (see also Figures S6H–S6K). Collectively, these results indicate that RBFOX2, NOVA2, and HNRNPA1 regulate splicing by a common mechanism.

DISCUSSION

In this study, we find that in addition to their well-established role in mRNA 3′ end formation, CPSF and SYMPK, but not other mRNA 3′ end formation factors, are involved in alternative splicing of ~11,400 internal human exons. Functional and physical interactions between factors involved in splicing and mRNA 3′ end formation have been previously demonstrated (reviewed in Martinson, 2011). However, these interactions have been thought to be important for either splicing of terminal introns or mRNA 3′ end formation, but not for promoting alternative internal exon usage as we show here. To our knowledge, interactions between mRNA 3′ end formation factors and the splicing regulators analyzed here have not been previously reported.

Two Classes of RBFOX2-Regulated Exons

Our RNA-seq, iCLIP, and directed CLIP results have revealed two classes of RBFOX2-regulated exons that differ in their requirement for CPSF/SYMPK. At RBFOX2-regulated, CPSF/SYMPK-regulated exons, CPSF/SYMPK is required for alternative splicing and is either recruited by pre-mRNA-bound RBFOX2 or associates with the pre-mRNA as an RBFOX2/CPSF/SYMPK complex. By contrast, at RBFOX2-regulated, CPSF/SYMPK-independent exons, CPSF/SYMPK is neither required for alternative splicing nor bound nearby. These findings raise the possibility that there are two forms of RBFOX2 that are either free or associated with CPSF/SYMPK.

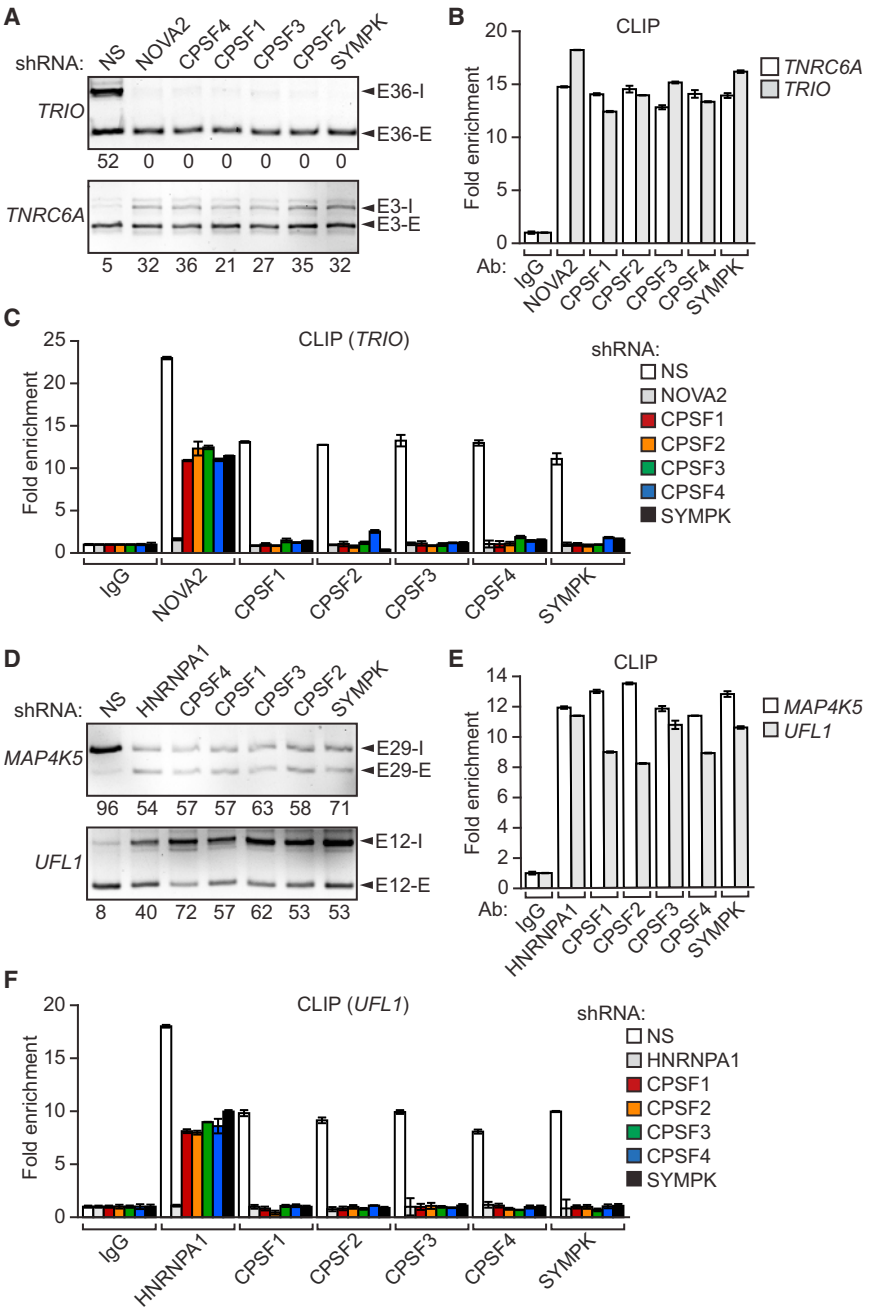


Figure 6. CPSF/SYMPK Is Also a Cofactor for the Splicing Regulators NOVA2 and HNRNPA1

(A) RT-PCR analysis monitoring exon inclusion (*TRIO*) or exclusion (*TNRC6A*) in cells expressing an NS, NOVA2, CPSF1–4, or SYMPK shRNA.

(B) CLIP analysis monitoring binding of NOVA2, CPSF1–4, and SYMPK to the regulated exons of the indicated pre-mRNAs. Data are represented as mean ± SD.

(C) CLIP analysis monitoring binding of NOVA2, CPSF1–4, and SYMPK to *TRIO* exon 36 in cells expressing an NS, NOVA2, CPSF1–4, or SYMPK shRNA. Data are represented as mean ± SD.

(D) RT-PCR analysis monitoring exon inclusion (*MAPK45*) or exclusion (*UFL1*) in cells expressing an NS, HNRNPA1, CPSF1–4, or SYMPK shRNA.

(E) CLIP analysis monitoring binding of HNRNPA1, CPSF1–4, and SYMPK to the regulated exons of the indicated pre-mRNAs. Data are represented as mean ± SD.

(F) CLIP analysis monitoring binding of HNRNPA1, CPSF1–4, and SYMPK to *UFL1* exon 12 in cells expressing an NS, HNRNPA1, CPSF1–4, or SYMPK shRNA. Data are represented as mean ± SD.

See also Figure S6.

CPSF/SYMPK Functions as a Cofactor by Regulating Binding of the Early Intron Recognition Factors U1 snRNP and U2AF

The CPSF/SYMPK complex is recruited to the alternatively spliced exon by a sequence-specific RNA-binding protein, such as RBFOX2, NOVA2, or HNRNPA1. CPSF/SYMPK then functions as a cofactor by facilitating the ability of RBFOX2, NOVA2, or HNRNPA1 to regulate binding of the early intron recognition factors U2AF and U1 snRNP. Notably, previous bioinformatic, biochemical, and genetic studies have shown that NOVA2 regulates early spliceosome assembly (Chen and Manley, 2009; Ule et al., 2006). Our findings are consistent with the general notion that frequently it is the earliest events in spliceosome assembly that are modulated to regulate alternative splicing (Chen and Manley, 2009). For RBFOX2, NOVA2, and HNRNPA1, binding upstream of the alternative exon causes exclusion, whereas binding to the downstream intron promotes inclusion (Chen and Manley, 2009; Venables et al., 2009; Yeo et al., 2009; Zhang et al., 2008). We have shown that CPSF/SYMPK can also function to promote either inclusion or exclusion, reinforcing the role of CPSF/SYMPK as a cofactor of RBFOX2, NOVA2, and HNRNPA1.

Notably, CPSF/SYMPK enhances RBFOX2 binding at CPSF/SYMPK-regulated exons, but not at CPSF/SYMPK-independent exons (see Figures 1G, 3E, and S4F). Thus, the distinction between CPSF/SYMPK-dependent and -independent exons may be related, at least in part, to differences in the intrinsic strength of the RBFOX2-binding sites between the two classes. For example, free RBFOX2 may be able to bind to strong RBFOX2-binding sites, whereas association of RBFOX2 with weak sites may require the additional binding enhancement provided by CPSF/SYMPK.

At included exons, binding of RBFOX2 and CPSF/SYMPK leads to recruitment of U1 70K and U2AF65 to the 5′ splice site

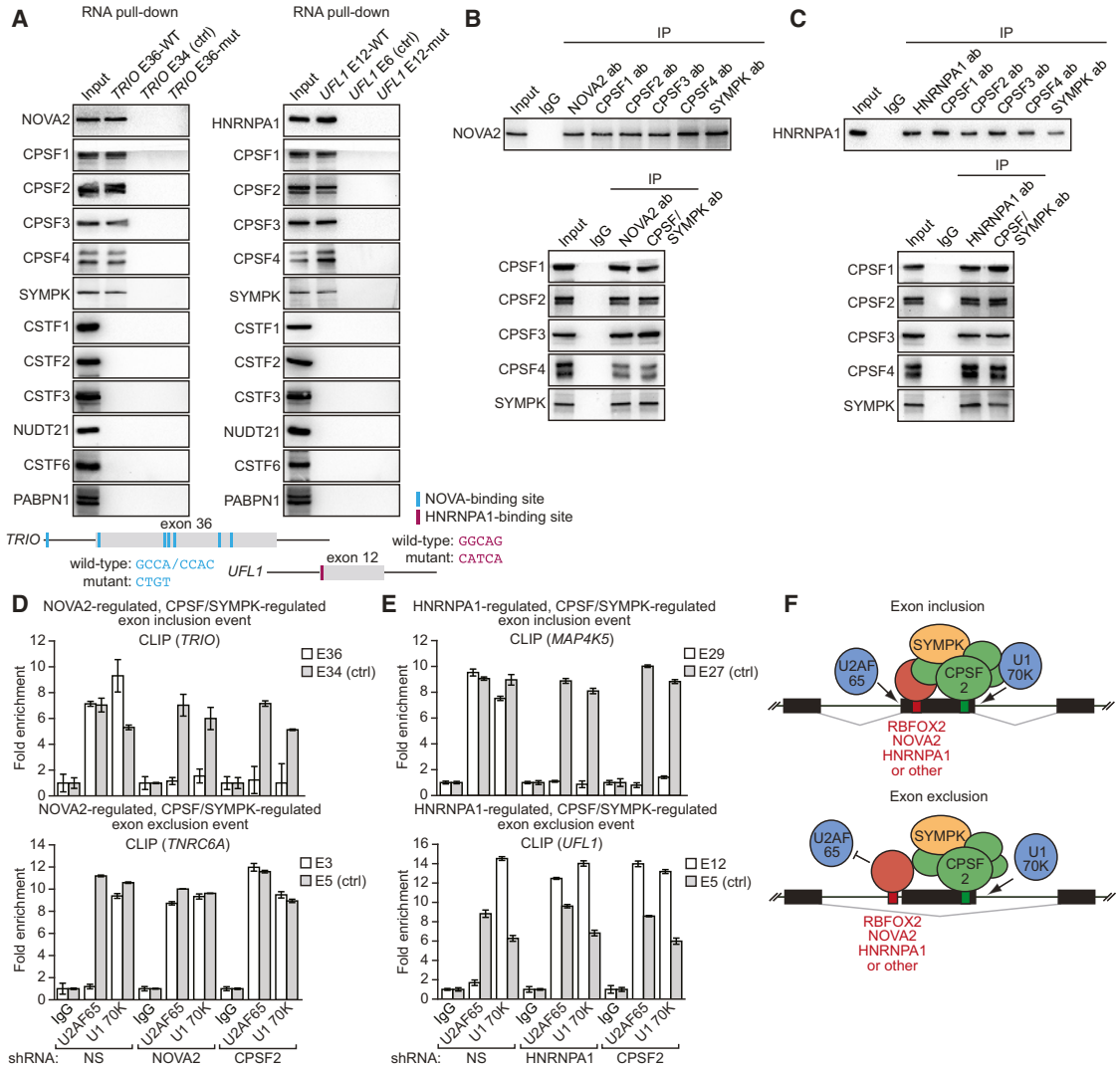


Figure 7. NOVA2 and HNRNPA1 Promote Exon Inclusion and Exclusion by a Mechanism Similar to That of RBFOX2

(A) RNA pull-down assay. A biotinylated RNA containing *TRIO* regulated exon 36 (left) or *UFL1* regulated exon 12 (right), derivatives containing a mutation in the NOVA2- or HNRNPA1-binding site or, as a control, a constitutively spliced exon (exon 34 of *TRIO* or exon 6 of *UFL1*) as well as 100 bp of upstream and downstream intron sequences was incubated in nuclear extract and analyzed for bound proteins by immunoblotting.

(B and C) Co-immunoprecipitation analysis. Cell extracts were immunoprecipitated in the presence of RNase with a NOVA2 (B) or HNRNPA1 (C) antibody, or CPSF1–4 or SYMPK antibody, or an IgG control, and immunoblotted for each protein.

(D) CLIP analysis monitoring binding of U2AF65 and U1 70K to regulated and constitutive exons in *TRIO* (top) and *TNRC6A* (bottom) in cells expressing an NS, NOVA2, or CPSF2 shRNA. Data are represented as mean ± SD.

(E) CLIP analysis monitoring binding of U2AF65 and U1 70K to regulated and constitutive exons in *MAP4K5* (top) and *UFL1* (bottom) in cells expressing an NS, HNRNPA1, or CPSF2 shRNA. Data are represented as mean ± SD.

(F) Model for CPSF/SYMPK-mediated regulation of alternative splicing of internal exons.

See also Figure S6.

and 3′ splice site, respectively (Figure 7F, top). The finding that RBFOX2, CPSF2, and U1 70K interact in an RNA-independent manner (Figure 5C) supports our conclusion that RBFOX2 and CPSF/SYMPK facilitate the recruitment of U1 snRNP to the pre-mRNA. Notably, mutational analysis of the 3′ splice site demonstrates that U2AF recruitment occurs subsequent to association of U1 snRNP with the 5′ splice site (Figure 5E). Similar results have been shown for NOVA2 and HNRNPA1, whose binding to the intron is known to promote spliceosome assembly (Black and Graveley, 2006; Ule et al., 2006).

Integration of our genome-wide iCLIP data with alternative splicing profiles reveals that at included exons, RBFOX2 and CPSF2 binding is significantly enriched inside the alternatively spliced exon and at the 5′ splice site (Figures S7A and S7B), consistent with the role of these proteins in facilitating 5′ splice site recognition. Furthermore, RBFOX2-regulated,

CPSF2-regulated exons show significant RBFOX2 enrichment at the exon-downstream intron junction (i.e., 5′ splice site), whereas CPSF2 shows significant enrichment within the exon as well as in the upstream and downstream introns (Figure S7C). It has been previously reported that the majority of alternative splicing events show combinatorial regulation, whereby the choice of splice site is governed by multiple RNA-binding proteins (Black, 2003; Matlin et al., 2005). In this regard, one reason why there are multiple CPSF2-binding peaks adjacent to the alternatively spliced exon is that CPSF2 is a cofactor for other factors such as NOVA2 and HNRNPA1, which are bound adjacent to RBFOX2-regulated, CPSF/SYMPK-regulated exons.

At excluded exons, depletion of RBFOX2 or CPSF2 results in increased U2AF65 binding to the pre-mRNA (Figure 5D), suggesting that binding of RBFOX2 and CPSF/SYMPK interferes with recruitment of U2AF to the 3′ splice site upstream of the excluded exon (Figure 7F, bottom). Consistent with this model, a previous study showed the RBFOX2 could block pre-spliceosome formation in vitro (Zhou and Lou, 2008). Interestingly, we found that RBFOX2-regulated, CPSF/SYMPK-independent exons have significantly weaker 3′ splice sites compared to exons that are modulated by both RBFOX2 and CPSF/SYMPK (Figure S7D). Perhaps both RBFOX2 and CPSF/SYMPK are required to block U2AF binding at strong 3′ splice sites, whereas at weaker 3′ splice sites RBFOX2 alone is sufficient.

Integration of our iCLIP data with alternative splicing profiles into an RNA map reveals a significant enrichment of RBFOX2 and CPSF2 immediately upstream of the 3′ splice site (Figure S7E). These results suggest that RBFOX2 and the CPSF/SYMPK complex sterically interfere with access of U2AF to the 3′ splice site. However, these findings do not rule out the possibility that at least on some excluded exons the repressive complex functions by interfering with exon definition by, for example, blocking communication between U1 snRNP and U2AF. This possibility is consistent with the observation that some U2AF binding is detectable. Such an indirect mechanism has been previously proposed for Py tract binding protein (PTB) acting through an exonic silencer in alternative splicing of *Fas* exon 6 (Izquierdo et al., 2005). In addition to RBFOX2, our results indicate that other RNA-binding proteins, such as NOVA2 and HNRNPA1, function through a similar mechanism (Figures 7D and 7E), and previous studies have also suggested that HNRNPA1 and NOVA block recruitment of U2AF and U1 snRNP, leading to exon exclusion (Del Gatto-Konczak et al., 1999; Ule et al., 2006; Zhu et al., 2001).

In contrast to included exons, where knockdown of RBFOX2 or CPSF2 results in reduced binding of both U1 70K and U2AF65, at excluded exons binding of U1 70K is unaffected by depletion of RBFOX2 or CPSF2 (Figure 5D). Thus, at excluded exons stable association of U1 snRNP with the 5′ splice is not dependent on RBFOX2 and CPSF/SYMPK, presumably because complementarity between U1 small nuclear RNA (snRNA) and the 5′ splice site is sufficient (Séraphin et al., 1988; Siliciano and Guthrie, 1988).

Collectively, our results reveal a new paradigm for CPSF/SYMPK function in promoting alternative splicing and provide mechanistic insights into communication between mRNA 3′ end formation factors and the spliceosome. Previous studies have identified a subcomplex containing CPSF2, CPSF3, and SYMPK, but not other mRNA 3′ end formation factors, that is involved in histone mRNA 3′ end formation (Sullivan et al., 2009). Notably, the CPSF/SYMPK subcomplex we identified includes CPSF3, which is the cleavage and polyadenylation site endonuclease (Mandel et al., 2006). Furthermore, although CPSF1, 3, and 4 have been previously shown to bind RNA (Barabino et al., 1997; Keller et al., 1991; Murthy and Manley, 1995), in the CPSF/SYMPK subcomplex the primary RNA-binding factor is CPSF2. It will be important to determine whether the CPSF/SYMPK subcomplex also has a more general role in mRNA 3′ end formation. Consistent with this idea, genome-wide studies have identified a large number of RBFOX- and NOVA-binding sites in 3′ UTRs, implicating these proteins as regulators of alternative polyadenylation (Licatalosi et al., 2008; Wang et al., 2008). Our results raise the possibility that the CPSF/SYMPK subcomplex may facilitate RBFOX- and NOVA-directed alternative polyadenylation.

EXPERIMENTAL PROCEDURES

Reporter Cell Line Construction

The parental GFP reporter construct (Wang et al., 2004) was kindly provided by Christopher Burge (Massachusetts Institute of Technology). Oligonucleotides containing an RBFOX2-binding site were annealed and inserted into the reporter construct using XhoI and ApaI sites. The construct was transfected into Flp-In-293 cells (Invitrogen), which were selected for 10 days with 150 μg/ml hygromycin. Individual hygromycin-resistant clones were isolated and analyzed by FACS. Cells with the maximum GFP fluorescence intensity (FL1-H) were used for the RNAi screen.

RNAi Screening

The Open Biosystems/Thermo Scientific TRC-Hs1.0 lentiviral human shRNA library was obtained through the University of Massachusetts (UMass) Medical School RNAi Core Facility. GFP/Flp-In-293cells were transduced with lentiviral pools and puromycin-selected for 2 days. GFP− cells were FACS sorted 10 days after infection, grown, and sorted again after 7 days, at which point cells were pooled, genomic DNA was isolated, and shRNAs were identified. Subsequent directed RNAi-mediated knockdown experiments used shRNAs listed in Supplemental Experimental Procedures.

FACS Analysis

Individual knockdown cells were FACS sorted, using a Becton Dickinson FACSCalibur flow cytometer, 10 days after lentiviral transduction and puromycin selection. Greater than 1×10^5 cells were analyzed for each sample. Data were processed and analyzed using FlowJo software.

Splicing Assays

Splicing assays were carried out in Flp-In-293 cells. Total RNA was extracted from knockdown cell lines using an RNeasy Kit (QIAGEN). Reverse transcription was carried out using SuperScript II reverse transcriptase (Invitrogen). Semiquantitative PCR using exon-specific primers (listed in Supplemental Experimental Procedures) was performed to obtain alternatively spliced products. Quantitative information of PCR products was obtained by staining the agarose gels using ethidium bromide and quantifying the signal using ImageJ software (NIH). Exon numbers were assigned according to Refseq nomenclature.

CLIP Assays

CLIP experiments were carried out essentially as previously described (König et al., 2010), with modifications as described in Supplemental Experimental Procedures and with gene-specific primers listed in Table S5. The qRT-PCR results were normalized to that obtained with an immunoglobulin G (IgG) isotype control, which was set to 1.

RNA Pull-Down Assays

RNA substrates were PCR amplified from Flp-In-293 cell genomic DNA using gene-specific primers listed in Supplemental Experimental Procedures. The PCR products were cloned into pGEM-T Easy (Promega), and mutations were introduced by full vector amplification. In vitro biotin-RNA synthesis was performed using an AmpliScribe T7-Flash Biotin-RNA Transcription Kit (Epicenter Biotechnologies). Synthesized RNAs were analyzed qualitatively and quantitatively by electrophoresis and UV spectrometry. RNA pull-down assays were performed as previously described (Iioka et al., 2011).

Co-immunoprecipitation Assays

Flp-In-293 cells were washed two times with cold PBS and lysed in 1 ml of immunoprecipitation (IP) lysis buffer (50 mM Tris-Cl [pH 7.4], 250 mM NaCl, 5 mM EDTA, 0.5% Triton X-100, 0.5 mM DTT, 1× complete protease inhibitor [Roche]) on ice. The lysate was cleared by centrifugation at 16,000 × *g* for 30 min at 4°C. Whole-cell lysate (2 mg per sample) was incubated with various antibodies (listed in Supplemental Experimental Procedures) overnight at 4°C. Protein G-coated Dynabeads (50 μl; Life Technologies) were added to each lysate-antibody complex, incubated for 2 hr, spun, and washed three times with IP lysis buffer. Protein complexes were eluted by boiling with Laemmli buffer. Blots were probed with the same antibodies listed in Supplemental Experimental Procedures. In preliminary experiments, the amount of antibody used for each immunoprecipitation was adjusted so that near-equivalent signals were obtained in the final immunoblot.

RNA-Seq

Samples were prepared using a TruSeq RNA Sample Preparation Kit (Illumina) according to manufacturer's instructions. The raw sequencing data have been deposited in NCBI's GEO and are accessible through GEO Series accession number GSE60392.

Genomic iCLIP Assays

iCLIP experiments were carried out essentially as previously described (König et al., 2010), with some modifications (see Supplemental Experimental Procedures). The raw data have been deposited in NCBI's GEO and are accessible through GEO Series accession number GSE60392.

iCLIP Motif Analysis

The human genome was partitioned by a window size of 40 nt. Genomic regions with one or more hits by the 5′ end of the iCLIP data were selected for motif search. Parameters supplied for motif finding with HOMER's findMotifsGenome program (Heinz et al., 2010) were -size 40 -len 5,6,7,8 -rna -local 4 -S 10 -gc.

ACCESSION NUMBERS

The GEO accession number for the RNA-seq and genomic iCLIP data reported in this paper is GSE60392 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE60392).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2015.03.016.

AUTHOR CONTRIBUTIONS

A.M. and M.R.G. designed the experiments. A.M. performed all the experiments. J.O. and L.J.Z. performed the bioinformatics analysis. A.M. and M.R.G. interpreted the data and wrote the manuscript.

ACKNOWLEDGMENTS

We thank Christopher Burge for reagents, the UMass Deep Sequencing Core for sequencing the RNA-seq and iCLIP samples, the UMass RNAi Core Facility for providing shRNA clones and libraries, and Sara Deibler for editorial assistance. This work was supported by NIH grant R01GM035490 to M.R.G. M.R.G. is an investigator of the Howard Hughes Medical Institute.

Received: October 30, 2014
Revised: February 12, 2015
Accepted: March 11, 2015
Published: April 23, 2015

REFERENCES

Anders, S., Reyes, A., and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. Genome Res. *22*, 2008–2017.

Barabino, S.M., Hübner, W., Jenny, A., Minvielle-Sebastia, L., and Keller, W. (1997). The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. Genes Dev. *11*, 1703–1716.

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

Black, D.L., and Graveley, B.R. (2006). Splicing bioinformatics to biology. Genome Biol. *7*, 317.

Blanchette, M., Labourier, E., Green, R.E., Brenner, S.E., and Rio, D.C. (2004). Genome-wide analysis reveals an unexpected function for the Drosophila splicing factor U2AF50 in the nuclear export of intronless mRNAs. Mol. Cell *14*, 775–786.

Chan, S., Choi, E.A., and Shi, Y. (2011). Pre-mRNA 3′-end processing complex assembly and function. Wiley Interdiscip. Rev. RNA. *2*, 321–335.

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

Colgan, D.F., and Manley, J.L. (1997). Mechanism and regulation of mRNA polyadenylation. Genes Dev. *11*, 2755–2766.

Cooper, T.A., Wan, L., and Dreyfuss, G. (2009). RNA and disease. Cell *136*, 777–793.

Del Gatto-Konczak, F., Olive, M., Gesnel, M.C., and Breathnach, R. (1999). hnRNP A1 recruited to an exon in vivo can function as an exon splicing silencer. Mol. Cell. Biol. *19*, 251–260.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Huelga, S.C., Vu, A.Q., Arnold, J.D., Liang, T.Y., Liu, P.P., Yan, B.Y., Donohue, J.P., Shiue, L., Hoon, S., Brenner, S., et al. (2012). Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. Cell Rep. *1*, 167–178.

Huh, G.S., and Hynes, R.O. (1994). Regulation of alternative pre-mRNA splicing by a novel repeated hexanucleotide element. Genes Dev. *8*, 1561–1574.

Iioka, H., Loiselle, D., Haystead, T.A., and Macara, I.G. (2011). Efficient detection of RNA-protein interactions using tethered RNAs. Nucleic Acids Res. *39*, e53.

Irimia, M., and Blencowe, B.J. (2012). Alternative splicing: decoding an expansive regulatory layer. Curr. Opin. Cell Biol. *24*, 323–332.

Izquierdo, J.M., Majós, N., Bonnal, S., Martínez, C., Castelo, R., Guigó, R., Bilbao, D., and Valcárcel, J. (2005). Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. Mol. Cell *19*, 475–484.

Jensen, K.B., and Darnell, R.B. (2008). CLIP: crosslinking and immunoprecipitation of in vivo RNA targets of RNA-binding proteins. Methods Mol. Biol. *488*, 85–98.

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

Keller, W., Bienroth, S., Lang, K.M., and Christofori, G. (1991). Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3′ processing signal AAUAAA. EMBO J. *10*, 4241–4249.

Keren, H., Lev-Maor, G., and Ast, G. (2010). Alternative splicing and evolution: diversification, exon definition and function. Nat. Rev. Genet. *11*, 345–355.

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

Kornblihtt, A.R., Schor, I.E., Alló, M., Dujardin, G., Petrillo, E., and Muñoz, M.J. (2013). Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat. Rev. Mol. Cell Biol. *14*, 153–165.

Kyburz, A., Friedlein, A., Langen, H., and Keller, W. (2006). Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3′ end processing and splicing. Mol. Cell *23*, 195–205.

Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., et al. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature *456*, 464–469.

Lim, L.P., and Sharp, P.A. (1998). Alternative splicing of the fibronectin EIIIB exon depends on specific TGCATG repeats. Mol. Cell. Biol. *18*, 3900–3906.

Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., and Fu, X.D. (2008). The splicing factor SC35 has an active role in transcriptional elongation. Nat. Struct. Mol. Biol. *15*, 819–826.

Mandel, C.R., Kaneko, S., Zhang, H., Gebauer, D., Vethantham, V., Manley, J.L., and Tong, L. (2006). Polyadenylation factor CPSF-73 is the pre-mRNA 3′-end-processing endonuclease. Nature *444*, 953–956.

Mandel, C.R., Bai, Y., and Tong, L. (2008). Protein factors in pre-mRNA 3′-end processing. Cell. Mol. Life Sci. *65*, 1099–1122.

Martin, G., Gruber, A.R., Keller, W., and Zavolan, M. (2012). Genome-wide analysis of pre-mRNA 3′ end processing reveals a decisive role of human cleavage factor I in the regulation of 3′ UTR length. Cell Rep. *1*, 753–763.

Martinson, H.G. (2011). An active role for splicing in 3′-end formation. Wiley Interdiscip. Rev. RNA. *2*, 459–470.

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

Murthy, K.G., and Manley, J.L. (1995). The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3′-end formation. Genes Dev. *9*, 2672–2683.

Nilsen, T.W., and Graveley, B.R. (2010). Expansion of the eukaryotic proteome by alternative splicing. Nature *463*, 457–463.

Paz, I., Akerman, M., Dror, I., Kosti, I., and Mandel-Gutfreund, Y. (2010). SFmap: a web server for motif analysis and prediction of splicing factor binding sites. Nucleic Acids Res. *38*, W281–W285.

Séraphin, B., Kretzner, L., and Rosbash, M. (1988). A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5′ cleavage site. EMBO J. *7*, 2533–2538.

Siliciano, P.G., and Guthrie, C. (1988). 5′ splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. Genes Dev. *2*, 1258–1267.

Sullivan, K.D., Steiniger, M., and Marzluff, W.F. (2009). A core complex of CPSF73, CPSF100, and Symplekin may form two different cleavage factors for processing of poly(A) and histone mRNAs. Mol. Cell *34*, 322–332.

Tollervey, J.R., Wang, Z., Hortobágyi, T., Witten, J.T., Zarnack, K., Kayikci, M., Clark, T.A., Schweitzer, A.C., Rot, G., Curk, T., et al. (2011). Analysis of alternative splicing associated with aging and neurodegeneration in the human brain. Genome Res. *21*, 1572–1582.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat. Biotechnol. *31*, 46–53.

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

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

Venables, J.P., 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.

Wang, Z., Rolish, M.E., Yeo, G., Tung, V., Mawson, M., and Burge, C.B. (2004). Systematic identification and analysis of exonic splicing silencers. Cell *119*, 831–845.

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

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

Zhang, C., Zhang, Z., Castle, J., Sun, S., Johnson, J., Krainer, A.R., and Zhang, M.Q. (2008). Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2. Genes Dev. *22*, 2550–2563.

Zhou, H.L., and Lou, H. (2008). Repression of prespliceosome complex formation at two distinct steps by Fox-1/Fox-2 proteins. Mol. Cell. Biol. *28*, 5507–5516.

Zhu, J., Mayeda, A., and Krainer, A.R. (2001). Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. Mol. Cell *8*, 1351–1361.