8370–8383 *Nucleic Acids Research, 2021, Vol. 49, No. 14* *Published online 9 July 2021*
https://doi.org/10.1093/nar/gkab595

RBFOX2 alters splicing outcome in distinct binding modes with multiple protein partners

Delong Zhou [1], Sonia Couture[1], Michelle S. Scott [2,*] and Sherif Abou Elela [1,*]

[1]Département de microbiologie et d'infectiologie, Faculté de médecine et des sciences de la santé, Université de Sherbrooke, Sherbrooke, QC J1E 4K8, Canada and [2]Département de biochimie et génomique fonctionnelle, Faculté de médecine et des sciences de la santé, Université de Sherbrooke, Sherbrooke, QC J1E 4K8, Canada

Received March 01, 2021; Revised June 22, 2021; Editorial Decision June 23, 2021; Accepted June 28, 2021

ABSTRACT

RBFOX2 controls the splicing of a large number of transcripts implicated in cell differentiation and development. Parsing RNA-binding protein datasets, we uncover that RBFOX2 can interact with hnRNPC, hnRNPM and SRSF1 to regulate splicing of a broad range of splicing events using different sequence motifs and binding modes. Using immunoprecipitation, specific RBP knockdown, RNA-seq and splice-sensitive PCR, we show that RBFOX2 can target splice sites using three binding configurations: single, multiple or secondary modes. In the single binding mode RBFOX2 is recruited to its target splice sites through a single canonical binding motif, while in the multiple binding mode RBFOX2 binding sites include the adjacent binding of at least one other RNA binding protein partner. Finally, in the secondary binding mode RBFOX2 likely does not bind the RNA directly but is recruited to splice sites lacking its canonical binding motif through the binding of one of its protein partners. These dynamic modes bind distinct sets of transcripts at different positions and distances relative to alternative splice sites explaining the heterogeneity of RBFOX2 targets and splicing outcomes.

INTRODUCTION

Alternative splicing decisions are regulated by up to one thousand RNA-binding proteins (RBPs) that bind to pre-mRNAs in the nucleus (1–4). Many of these splicing factors are expressed in a tissue, condition or disease-specific manner leading to programmable modification of splicing patterns (5–13). One of the most studied alternative splicing factors is the master regulator of tissue-specific alternative splicing RBFOX2 (7,14–17). The RBFOX family also includes RBFOX1 and RBFOX3 (18). All RBFOX proteins contain a central RNA recognition motif that recognizes a consensus sequence, (U)GCAUG, usually found in the introns that flank target exons (19,20). RBFOX proteins promote exon skipping when they bind upstream of the alternative exon but inclusion when they bind downstream of this exon (15,16,21,22). Despite their common consensus RNA-binding motif, these proteins are only partially redundant in vitro, and in vivo they are mostly expressed in different tissues (7,14,18). RBFOX1 is expressed in the brain, heart and muscles but not in ovaries, and RBFOX3 is expressed exclusively in the brain (7,14,18). RBFOX2 is more widely expressed, especially in muscle and mesenchyme/mesoderm, and it is the only member of the RBFOX family that is found in ovary and breast tissues (16,21).

RBFOX2 was implicated in the development of ovarian and breast cancer as well as epithelial to mesenchymal transition (EMT), which is essential for cancer metastasis (21,23,24). EMT-associated splicing is only partially regulated by RBFOX2 and requires the expression of a large number of RNA-binding proteins including ESRPs, CELF, MBNL, hnRNPs, SAM68 and SRSF1 (17,24,25). The expression of RBFOX2 increases after induction of EMT (26) and RBFOX2-depleted cells undergo an EMT with no obvious morphological differences from the control cells (23). Nevertheless, cell invasion is significantly reduced in RBFOX2-depleted tumor cells, suggesting that RBFOX2 enables epithelial cells to gain invasive properties during EMT (23). The role of RBFOX2 in regulating EMT appears to be linked to its role in cancer development. RBFOX2 expression is significantly lower in epithelial ovarian cancer than in normal tissues (17,21). Decreasing the level of RBFOX2 in cancer cell lines shifts many alternative splicing events in the same direction as in breast and ovarian cancer tissues. Similarly, RBFOX2 also regulates subtype-specific splicing in a panel of breast cancer cell lines (21). The relative importance of RBFOX2 in cancer might be partially due to differences in cell lines. However, several reports have shown that differences in RBFOX2 expression could also be observed between normal and cancer cells of the same type. For example, laser-dissected tissue of the normal ovary and

*To whom correspondence should be addressed. Tel: +1 819 821 8000 (Ext 75275); Email: sherif.abou.elela@usherbrooke.ca
Correspondence may also be addressed to Michelle S. Scott. Email: michelle.scott@usherbrooke.ca

ovarian tumor microenvironment have clear differences in RBFOX2 expression (16).

In the last few years, it has become clear that RBFOX2 recognizes different sets of alternatively spliced RNAs in different tissues and not all of the identified splicing targets of RBFOX2 harbor its RNA-binding motif (7,15,19,27). Indeed, several studies have identified RBFOX2-dependent alternative exons that are not located near a canonical (U)GCAUG motif as well as the presence of the motif not resulting in the expected splicing outcome (16,27–29). For example, in several documented cases the presence of the binding site upstream of the exon results in exon inclusion rather than exclusion (17,21,29,30). The origin of this behavior is not clear, but it was suggested that it might be explained in part by a large number of RBFOX2-interacting proteins that can act as recruiters, chaperones or co-factors capable of modifying the alternative splicing activity of RBFOX2. Indeed, it was shown that RBFOX2 forms a part of a 'large assembly of splicing regulators' (LASR), a multimeric complex containing the proteins hnRNPM, hnRNPH, hnRNPC, Matrin3, NF110/NFAR-2, NF45, and DDX5 (28). Strikingly, only the binding motifs of hnRNPC and hnRNPM, among these proteins, were found enriched near RBFOX2-binding sites and their contribution to RBFOX2 function remains limited to a few examples (28). In other cases, RBFOX2 was found to interact with, and its splicing target was affected by, the 3′ end formation factor CPSF2 (31). These studies suggest that the effect of RBFOX2 on splicing is mediated by other proteins. However, it is not clear how and when these proteins influence RBFOX2 function.

Here, we systematically analyzed the contribution of the proteins that physically or functionally interact with RBFOX2 on mRNA binding and splicing to identify the main proteins associated with RBFOX2 binding and function. Our results indicate that RBFOX2 is recruited to its targeted RNA using different binding modes and protein partners permitting the regulation of a broad range of splice sites exhibiting different sequence motifs and functional outcomes.

MATERIALS AND METHODS

Identification of RBFOX2 associated proteins

To examine the systems biology of the archetypal splicing factor RBFOX2 we collected proteins that physically interact with RBFOX2 by mass spectrometry (32) or co-sedimentation assays (28) or have similar effect on splicing (31) and examined the compiled dataset using DAVID v6.8 web server (38) and those accumulating in the nucleus were retained for further analysis (Supplementary Table S1).

Motif detection in RBP binding sites

This study used a panoply of published crosslinking and immunoprecipitation sequencing (CLIP-seq) experiments. The binding sites of RBPs were obtained from enhanced CLIP-seq (eCLIP-seq) experiments in HepG2 cell lines from the ENCODE project (33,34), except for CPSF2 for which no eCLIP-seq dataset is yet available. The peaks were defined by the narrowbed files for two replicates from the ENCODE project. We then merged the peaks using Bedtools v2.25.0 (35), and the centers of the merged peaks were considered to be the binding positions of the RBP. This process defined between ∼70 000 (RBFOX2) and ∼800 000 (hnRNPC) binding sites for each RBP eCLIP-seq experiment. For CPSF2, the binding sites were obtained from individual nucleotide resolution CLIP-seq (iCLIP-seq) experiments performed by Misra *et al.* (31). The centers of their CPSF2 binding windows were considered to be the binding positions, which gave ∼1,700,000 binding sites. Based on the frequency of the known motifs near the protein binding sites we selected an 80 nt window, located 60 nt upstream to 20 nt downstream of the binding center (illustrated in Supplementary Figure S1A), and scanned it for the presence of the RBP motif. The motifs considered for each RBP were those defined in the web supplement of Ray *et al.* (36), and listed in Supplementary Table S2. Three datasets, each containing 1 000 000 randomly chosen genomic windows of 80 nt, were used as background sequences for motif enrichment analysis. Statistical significance was determined by chi-squared tests with R v3.3.0.

Identifying overlapping RBP binding sites

The binding sites of two RBPs were considered as overlapping if the distance between the two centers was within 20nt. They were considered as adjacent if the distance was between 20nt and 60nt. For enrichment analysis, the extent of overlap, between RBFOX2 and other RBPs considered, was compared to the extent of overlap between two RBPs that are unrelated to splicing or to RBFOX2: FXR2 or NCBP2. Statistical significance was determined using chi-squared tests with R v3.3.0.

Formaldehyde crosslinking and immunoprecipitation (IP)

For each IP reaction 2×10^7 SKOV3.ip1 cells were suspended in crosslinking buffer (HEPES 20 mM, KCl 10 mM, $MgCl_2$ 1.5 mM) at 1×10^6 cells/ml and incubated with 0.2% (v/v) fresh formaldehyde for 10min on ice, then quenched with glycine (final concentration 0.25 M) for another 10min on ice. Cells were then broken with a Branson sonicator. Whole cell extract was treated with 15 ug of RNase A to eliminate protein-protein interactions via RNA bridges. IP was performed as previously described (37). Briefly, 5μg of antibody (RBFOX2: BETHYL A300-864A; hnRNPC: MBL RN052PW; hnRNPM: SANTA CRUZ sc-20001; SRSF1: invitrogen Catalog #32-4500) was incubated with 50μl Dynabeads (Invitrogen catalog #11041 or #11203D) for 30min at room temperature. The beads-antibody complex was then incubated, with a cell lysate from 2×10^7 SKOV3.ip1 cells, overnight at 4°C. The mix was washed to prevent unspecific binding. The protein was then eluted with 0.1 M glycine, pH 2.5 for 3 min at room temperature. Tris buffer pH 8.0 and Laemmli buffer were added to the elution to neutralize the pH and preserve the sample. Samples were heated for 1 h at 95°C to reverse cross-linking before Western Blot.

RNA immunoprecipitation

RNA-IP was performed as described in the previous section using whole cell extracts but with no formaldehyde crosslinking and no RNase treatment. During the last wash, 10% of the mix was taken to confirm the IP efficiency by Western Blot. The remaining mix was treated with proteinase K and the RNA was extracted using acidic phenol/chloroform/isoamyl alcohol mix (125:24:1 pH 4.3 Thermofisher, Canada) followed by chloroform extraction. The RNA was precipitated with 95% ethanol containing 2% potassium acetate. The RT-qPCR was performed in three independent biological replicates each with three technical replicates. Reverse transcription (RT) was performed using immunoprecipitated RNA, random hexamers (3 μM), dNTP (1 mM), 20 units of RNaseOUT (Thermofisher, Canada) and 1 unit of locally produced Moloney Murine Leukemia Virus-RT (MMULV-RT) in a total volume of 20 μl. The RT reaction was carried out for a single cycle of 25°C/5′, 42°C/60′ and 65°C/20′. cDNA was diluted using 580 μl of nuclease free H_2O and stored at –20°C. Titer tests were performed to determine optimal cDNA concentration and reduce sample-specific RT and PCR biases. The PCR reactions was performed in an Eppendorf Mastercycler RealPlex. The 10 μl reaction mix included 3 μl of diluted cDNA, 200 nM primers and 5μl of BioRad iTaq Universal SYBER Green Supermix. No-RT and no-template reaction mixes were used as negative controls. The targeted sequences were amplified for 50 cycles (60°C/30′, 72°C/30′ and 72°C/30′). The primers used for RT-qPCR are listed in Supplementary Table S9. Primer specificity and PCR efficiency were validated as previously described (38,39). The RNA concentration detected in the samples precipitated using antibodies against the protein of interest (IP) was compared to that obtained after mock precipitation using IgG. The fold enrichment was calculated using the following equation: Enrichment $= 2\wedge(Cq_{IgG} - Cq_{IP})$ where Cq is the quantification cycle. The statistical significance of difference in RNA concentrations between conditions were calculated using two-sided Student's t-test assuming unequal variances performed on R v3.3.0 software.

RBP knockdowns

RBP knockdowns were performed using siRNA obtained from IDT and their sequences are listed in Supplementary Table S3. SKOV3.ip1 cells were transfected with 10nM of each siRNA using Invitrogen Lipofectamine 2000. RNA was extracted from mock and siRNA-transfected cells, 72 h post-transfection as previously described (38,39). RNA quantity, purity and integrity were assessed on a NanoDrop™ 8000 Spectrophotometer and an Agilent 2200 TapeStation. Samples with RNA integrity number (RIN) higher than 9 were reverse transcribed and 1 ng of cDNA was used for the RT-qPCR as described above. The PCR primers used are listed in Supplementary Table S4. The relative mRNA expression was calculated using the following equation: $Rel\_Exp_{knockdown} = 2\wedge(Cq_{mock} - Cq_{knockdown})_{Gene}/2\wedge(Cq_{mock} - Cq_{knockdown})_{House\ keeping}$. Only the knockdowns resulting in >60% reduction in RNA amount as detected by RT-qPCR were considered for further analyses.

RNA-seq analysis

Strand-specific cDNA libraries were prepared using the previously described TGIRT-seq protocol (40) and sequenced using Illumina Nextseq 500. We obtained an average of 98M paired-end reads per sample, with a minimum of 75M reads. The quality of paired-end RNA-seq reads was examined using Fastqc v0.10.1, then trimmed using Trimmomatic v0.38 (41) with the following parameters LEADING:30 TRAILING:30 SLIDINGWINDOW:4:15 MINLEN:50. Reads retained by the trimming were further trimmed to a fixed length of 50 nt. The reads were aligned against human reference genome hg19 (obtained from UCSC genome browser golden path ftp://hgdownload.soe.ucsc.edu/goldenPath/hg19/chromosomes/) using Tophat (v2.0.9, Bowtie 2.1.0.0, Samtools 0.1.19.0) with default parameters (42–44). The sequencing depths and alignment rates for each sample are listed in Supplementary Table S5. The output files were then sorted and indexed using Samtools. Insert sizes were estimated using PICARD (v2.10.9-SNAPSHOT, http://broadinstitute.github.io/picard/; Broad Institute). Gene expression levels were calculated using CoCo v0.2.1p2 and Ensembl genome GRCh37 release 87 (45,46).

Alternative splicing analysis and validations

Differential alternative splicing analysis of cassette exons was performed using MISO (0.4.9) with the annotation for hg19 provided by MISO (47). In each biological replicate, an event was considered affected by an siRNA against an RBP if (a) at least 150 reads were mapped to the event in the mock sample, (b) the absolute splicing shift was greater than 15%, and (c) the Bayes Factor was >10. An event was considered affected by the knockdown of an RBP if it was affected by both siRNAs used in the gene knockdown, as described in the RBP knockdown section above, and when the shift was in the same direction in each case. Splicing events detected, by RNA-seq, to be regulated by RBPs are listed in Supplementary Table S6. Splice sensitive end-point RT-PCR was used to validate the splicing events detected by RNA-seq analysis as previously described (16). The primer sequences and tested events are listed in Supplementary Table S7.

Identifying splicing events bound by RBPs

The genomic positions of splicing events obtained from the MISO analysis were compared with eCLIP binding sites of RBPs described above using the bedmap program of Bedtools v2.25.0. A splicing event was considered as bound by an RBP if the cassette exon or its flanking introns intersected with at least one binding site of the RBP. For each event, the binding sites closest to the upstream exon, cassette exon or downstream exon are listed in Supplementary Table S8. The Table contains information on the exons near which the closest binding sites were found, including the distance between the centre of the binding sites and the exons (negative values for upstream binding, 0 for binding within exons, and positive values for downstream binding). Supplementary Table S8 also contains information on the RBPs that bind to the binding sites, presence of the motifs, and the

binding category. The density of binding sites, defined as the number of binding sites found in each window divided by the number of splicing events, was then plotted for the cassette exon and 2500 nt intronic regions flanking the splice sites. Intronic regions were divided into windows of 100 nt, while the cassette exons were divided into: 5′ (up to 100 nt), 3′ (up to 100 nt) or middle (rest of the exon when the exon is longer than 200 nt).

Classification of RBP binding sites

Binding sites were classified as (a) single-binding sites when they bind and contain the motif of a single RBP, (b) multiple-binding sites if they bind and contain the motif of more than one RBP, (c) secondary-binding sites if they bind more than one RBP but harbour the motif of only one of these interacting proteins or (d) Unknown sites for those that do not meet any of the previous criteria.

Gene ontology

Gene ontology analysis was performed on the Metascape v6.8 web server on a list of genes associated with the splicing events of interest (48) using splicing events not affected by RBP knockdown as negative control (with at least 150 reads mapped to the events in the mock samples and $< 5\%$ change in PSI in any replicate).

Domain prediction

Human genome annotation files and protein sequences were obtained from Ensembl (GRCh38, release 103). InterProScan v5.40–77.0 (49) was used with default options to predict domains and important sites in each protein. Predicted domain location in the protein was compared to the genomic annotation to obtain its corresponding genomic location. The genomic locations of events identified by RNA-seq (in human genome version hg19) were first converted to GRCh38 using the UCSC genome browser LiftOver function (https://genome.ucsc.edu/cgi-bin/hgLiftOver) then compared to the locations of predicted domains to identify domains that overlap with alternatively spliced exons regulated by each binding mode, as well as exons not affected by the depletion of any of the four investigated RBPs. Chi-squared tests were performed to investigate whether the proportion of alternatively spliced events overlapping with each domain is significantly different from the non-affected events, and *P*-values were adjusted using the Benjamini-Hochberg Procedure.

RESULTS

RBFOX2 identifies its target RNAs as part of a modular complex of RNA binding proteins

Most of the splicing events regulated by RBFOX2 do not contain a canonical RBFOX2-binding motif (15,50,51). It was previously proposed that at least some of these events are regulated through the binding of RBFOX2 chaperones or cofactors like CPSF2 or hnRNPM (28,31). Indeed, RBFOX2 interacts with many different proteins but their contribution to the selection of RBFOX2 splicing targets, and a full characterization of the functional complex responsible for RBFOX2 RNA selectivity, is still lacking (28,31). To understand the mechanism by which RBFOX2 selects its target RNAs, we scanned the literature and public databases to compile RBFOX2-interacting proteins, with the goal of examining their contribution to the binding, and splicing, of RBFOX2 targets. Nuclear proteins predicted to physically interact with RBFOX2 by mass spectrometry (32) or co-sedimentation assays (28), or to have similar effect on splicing (31) were identified. Those with established RNA-binding patterns, as defined by previously published RNA crosslinking (e.g. eCLIP or iCLIP) and RNA binding motifs (37,38) were retained for further analysis, as described in Figure 1A. In total this literature search identified 630 RBFOX2-interacting proteins (Supplementary Table S1).

Most of the interacting partners are likely not RNA-binding proteins and just eleven of the 630 RBFOX2-interacting proteins had established RNA-binding motifs and binding sites, all of which were identified using eCLIP and iCLIP (Figure 1A and Supplementary Table S2). Of these eleven RBFOX2 binding proteins with established binding patterns, only three (hnRNPM, CPCF2 and hnRNPC) were previously suggested to function as RBFOX2 splicing cofactors (28,31) that may influence its splicing and or binding activity, while the remaining 8 (hnRNPK, RBM5, TIA1, Sam68, U2AF1, QKI, IGF2BP1 and SRSF1) were not previously shown to affect RBFOX2 binding. Intriguingly, we found that 60% of RBFOX2-binding sites (∼40 nt long) that contain the canonical RBFOX2 binding motif (GCAUG) overlapped with the binding sites of one or more of these RBFOX2-interacting proteins (centers of binding sites within 20 nt). Furthermore, an even larger proportion (i.e. 75%) of RBFOX2-binding sites lacking the canonical motif overlap with binding sites for the other proteins (Figure 1B and Supplementary Figure S1B). These observations suggest that RBFOX2-interacting proteins have the potential to influence the selection of RBFOX2 targets and may function, at least in some cases, as chaperones that recruit RBFOX2 to RNA in the absence of its canonical RNA-binding motif.

Comparing the overlap between RBFOX2-binding sites and each of its eleven interacting proteins to the overlap observed with non-RBFOX2-interacting and non-splicing-regulating RNA-binding proteins (e.g. FXR2 and NCBP2) indicated that the putative binding sites for all but one RBFOX2-interacting proteins (CPSF2) are significantly enriched $p<2.2e\text{-}16$ near RBFOX2-binding sites (Figure 1C). Nine of the eleven interacting proteins each overlapped with $>10\%$ of the RBFOX2-binding sites, and the highest level of overlap with RBFOX2 binding sites is observed for RBM5, TIA1, IGF2BP1 and SRSF1 (Supplementary Figure S1B). In the majority of cases the RNA-binding proteins tested bound to the RNA within 20 nucleotides from the center of RBFOX2 binding sites, except for CPSF2, which is found to bind adjacent to RBFOX2 binding sites (60 nucleotides away from the center of the canonical RBFOX2-binding site, Supplementary Figure S1C). The canonical motifs of seven RBFOX2-interacting proteins (hnRNPC, hnRNPM, Sam68, RBM5, U2AF1, QKI and SRSF1) are enriched within their RBFOX2-overlapping binding sites (Figure 1D and Supplementary Figure S1D). This suggests

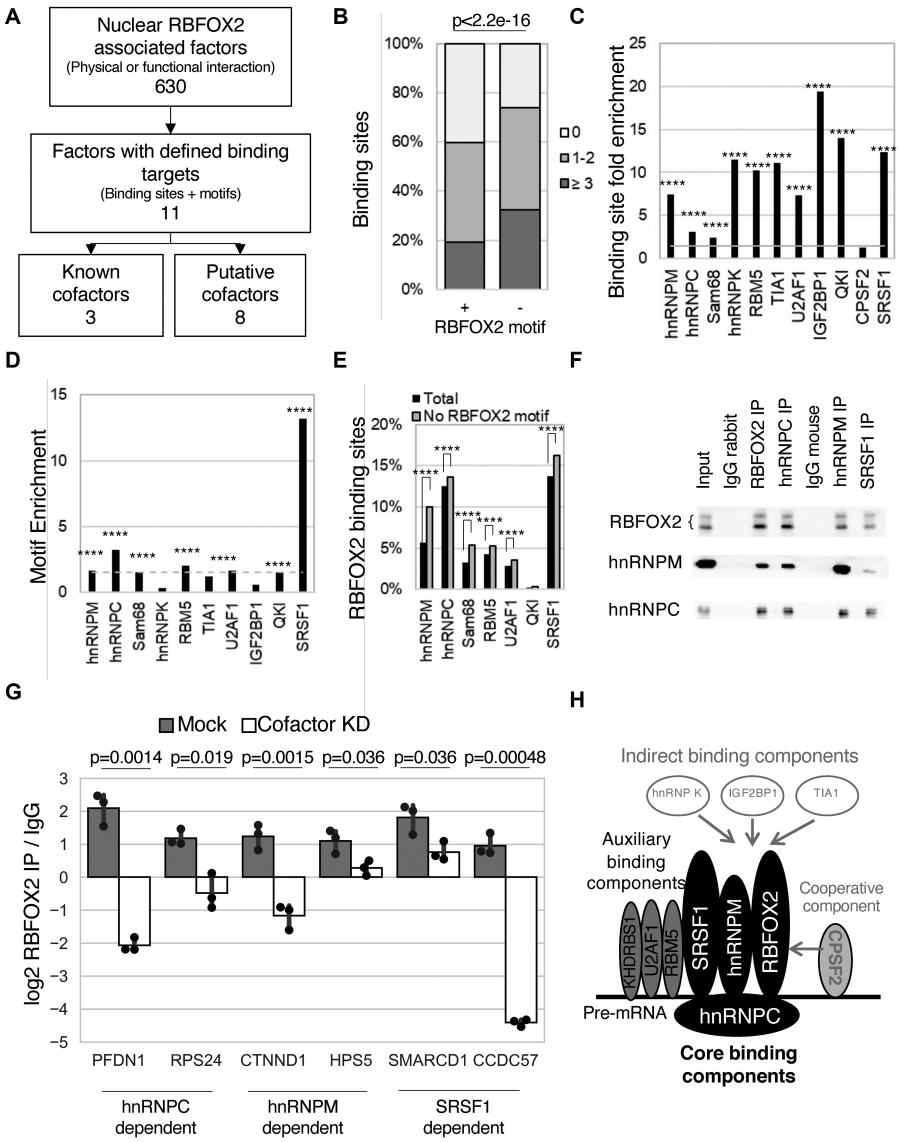

Figure 1. Identification of RBFOX2 binding cofactors. (**A**) Strategy for the identification of potential RBFOX2 cofactors. Proteins with predicted physical or functional association with RBFOX2 were identified from the literature (28,32,54) and their overlap with RBFOX2 binding sites was examined. The number of known and newly identified (putative) cofactors are indicated at the bottom. (**B**) RBFOX2-interacting proteins differentially associate with RBFOX2-binding sites lacking canonical RBFOX2 motifs. The proportion of RBFOX2-binding sites containing (+, n = 31,344) or lacking (−, *n* = 42,360) RBFOX2-binding motifs that overlaps with the binding sites of zero, 1–2 or more than 2 interacting proteins is shown as bar graphs. The statistical significance of the difference between the binding sites containing or lacking an RBFOX2 binding motif was calculated using chi-squared test and the *P*-value indicated on top. (**C**) Binding of RBFOX2-interacting proteins is significantly enriched at RBFOX2-binding sites. The fold-enrichment of RBFOX2-interacting proteins in RBFOX2-binding sites was calculated in comparison with their distribution to the overlap with binding sites of non-splicing related RBPs and normalizing by the number of binding sites established for each RBP. The four asterisks indicate p< 2.2e-16 determined by chi-squared tests. (**D**) Identification of the binding motifs of RBFOX2-interacting proteins that are enriched in RBFOX2-binding sites. The fold-enrichment, of the binding motif of each of the RBFOX2-interacting proteins in their RBFOX2 overlapping binding sites, was calculated relative to the background of random genomic sequence. The *p*-value of the difference in the motif distribution between the overlapping binding sites and the genome was calculated and indicated by the asterisks as in (C). (**E**) The motifs of RBFOX2-interacting proteins are differentially enriched in binding sites lacking RBFOX2 motifs. The motif enrichment of RBFOX2-interacting proteins in RBFOX2-binding sites containing and lacking RBFOX2 motifs were compared and significance of the difference calculated and indicated as in (C). (**F**) Validation of RBFOX2 core protein interactions in the ovarian cancer cell line SKOV3.ip1. Cell extracts were prepared from SKOV3.ip1 cells and RBFOX2, hnRNPC, hnRNPM and SRSF1 were immunoprecipitated and the associated proteins were immunoblotted using protein-specific antibodies. The two RBFOX2 protein isoforms are indicated by the open bracket. (**G**) The core RBFOX2 interacting proteins recruit RBFOX2 to RNA lacking an RBFOX2 motif. The SRSF1, hnRNPC and hnRNPM genes were knocked down in SKOV3.ip1 cell lines and the co-immunoprecipitation of RBFOX2 with RNA lacking canonical RBFOX2 binding motifs was examined by RT-qPCR. Enrichments were calculated relative to values obtained from mock precipitation with rabbit IgG. The error bar indicates the standard deviation obtained from three independent biological replicates. The significance of difference before and after the siRNA knockdown was evaluated by two-sided Student's t-tests assuming unequal variances and the resulting *P*-value is indicated on top. (**H**) Schematic illustration of RBFOX2 partner proteins and their contribution to RNA binding. The RBFOX2-interacting proteins identified in (C) are presented by the ellipses in relation to the degree of their association with the RBFOX2 protein's RNA-binding site. Black, gray, light grey and white ellipses indicate major recruiters of RBFOX2, minor recruiters of RBFOX2, cooperative factors that stabilize RBFOX2-RNA interaction by binding to adjacent sequence and factors that do not directly bind to RBFOX2-targeted RNA sequences, respectively.

that in most cases these seven proteins bind in proximity of RBFOX2 binding sites using their own canonical binding motifs and as such they may influence RBFOX2 binding. Surprisingly, the highest level of motif enrichment is found in RBFOX2-binding sites overlapping with the binding sites of the canonical SR protein splicing factor SRSF1, which has not previously been linked to RBFOX2 (Figure 1D). Out of the 7 RBPs only 3 (hnRNPC, hnRNPM and SRSF1) overlapped with more than 10% of RBFOX2 binding sites lacking its motif (Figure 1E). This indicates that out of the 11 proteins tested hnRNPC, hnRNPM and SRSF1 are the most likely to recruit RBFOX2 to splice sites lacking its binding motif. Indeed, RBFOX2-binding sites lacking its motif featured more motif-containing binding sites of hnRNPC, hnRNPM and SRSF1 than RBFOX2 motif-containing binding sites. This further supports the idea that these proteins may act as recruiters of RBFOX2 (Figure 1E). To directly confirm the interaction of RBFOX2 with hnRNPC, hnRNPM and SRSF1, we examined the co-precipitation of these proteins by pulling them down using antibodies against each of the four proteins in the ovarian cancer cell line SKOV3.ip1, which was previously used to link RBFOX2 to splicing in ovarian cancer (38,52). As indicated in Figure 1F, antibodies against any one of these four 'core' proteins co-immunoprecipitated all three other components, consistent with the view that they are either a part of the same complex or at the very least form a dynamic network of flexible and interchangeable binary complexes.

To determine the contribution of the three core proteins (hnRNPC, hnRNPM and SRSF1) to RBFOX2 target selectivity, we next examined the impact of their knockdowns on the binding of RBFOX2 when its binding motif is absent. We used antibodies against RBFOX2 to immunoprecipitate it before and after the knockdown of each of the three core proteins in SKOV3.ip1cells and monitored the co-immunoprecipitated RNA by RT-qPCR. As indicated in Figure 1G, RBFOX2 binding to RNAs lacking canonical RBFOX2 motifs required the presence of the partner protein associated with the binding motif found near the targeted splice site. We conclude that SRSF1, hnRNPC, and hnRNPM can indeed function as bona fide recruiters of RBFOX2. As illustrated in Figure 1H, overall, the data indicate that the RBFOX2 complex selects its RNA targets mostly through the combinatorial binding of hnRNPC, hnRNPM and SRSF1 to their nearby binding sites. However, based on the protein interaction and RNA binding data described in Figure 1C–E, it is also possible that, in a smaller number of cases, RBFOX2 may be recruited to RNA by RBM5 or U2AF1 and its binding stabilized by the 3′ end formation factor CPSF2. Since hnRNPK, IGF2BP1 and TIA1 interact with RBFOX2 but do not directly bind to most RBFOX2-binding sites it is more likely that these proteins either indirectly or allosterically influence RBFOX2 function rather than acting as direct recruiters of RBFOX2.

RBFOX2 may alter splicing through the binding of its protein partners

Analysis of the RBFOX2-interacting protein binding patterns suggests that each of these proteins can bind either directly through its canonical binding motif or by protein-protein interactions as part of the complex but it is not clear whether both of these binding modes contribute to the splicing decision. Therefore, we examined the impact of depleting each of the RBFOX2 core interacting proteins (RBFOX2, hnRNPC, hnRNPM and SRSF1) on the splicing of simple cassette exons and investigated the effect of each knockdown on splicing with respect to their binding and motif distributions. Each RBP was knocked down in SKOV3.ip1cells using two independent siRNAs and the knockdown efficiency was confirmed by RT-qPCR and Western Blot as indicated in Supplementary Figure S2A and S2B. In general, depleting one RBP had little effect on the expression of the other RBPs, indicating that the effects on splicing are specific to the depleted RBP and not an indirect effect of cross-regulation between the interacting proteins (Supplementary Figure S2A–C). RNA samples from the successful knockdown of two independent biological replicates were sequenced to identify splicing events regulated by each RBP. 45 to 65 such altered splicing events for each RBP as determined by RNA-seq were then verified using splice sensitive endpoint RT-PCR (Supplementary Figure S2D and Supplementary Table S7). As indicated in Figure 2A, the knockdowns resulted in varying effects on the splicing of cassette exons. As measured by RNA-seq, the largest number of affected splicing events arose from the depletion of hnRNPC followed by RBFOX2, SRSF1 and hnRNPM. Depletion of hnRNPC or SRSF1 caused more exon exclusion than inclusion. In contrast, hnRNPM depletion had the opposite effect leading to more exon inclusion, while RBFOX2 depletion resulted in equal numbers of exon inclusion and exclusion events. This indicates that RBFOX2 and its interacting partners affect splicing in a modular and RBP-specific manner and not as a fixed complex with a single splicing outcome.

To determine the contribution of their binding to the effect of the different RBPs on splicing we examined the binding motif distribution of the different RBPs near the affected splicing events. First, we verified the accuracy of the binding patterns predicted by eCLIP by re-examining the association of the predicted binding sites found near the affected splicing events with the different RBPs, using the same SKOV3.ip1cell lines used for sequencing. As shown in Figure 2B and Supplementary Figure S3A, all the mRNA carrying the predicted binding sites by eCLIP that we randomly chose for testing were enriched in the protein fractions precipitated by their respective RBPs. Based on these results we then compared the splicing events affected by the knockdown of each RBP with their respective binding sites predicted by eCLIP. As expected, most splicing events affected by the depletion of one RBP had at least one binding site of this RBP and binding sites were enriched in alternatively spliced regions (Figure 2C). Most canonical binding sites (binding sites containing canonical binding motif) are found in events affected by hnRNPC knockdown and most sites lacking canonical motifs are found in RBFOX2-affected splicing events. Indeed, 40% of the splicing events affected by RBFOX2 knockdown were not associated with clear RBFOX2 binding site. This trend was also maintained when examining the motif distribution in the binding sites of the different RBPs. Once again for all RBPs except RBFOX2, most binding sites contained the cognate protein's

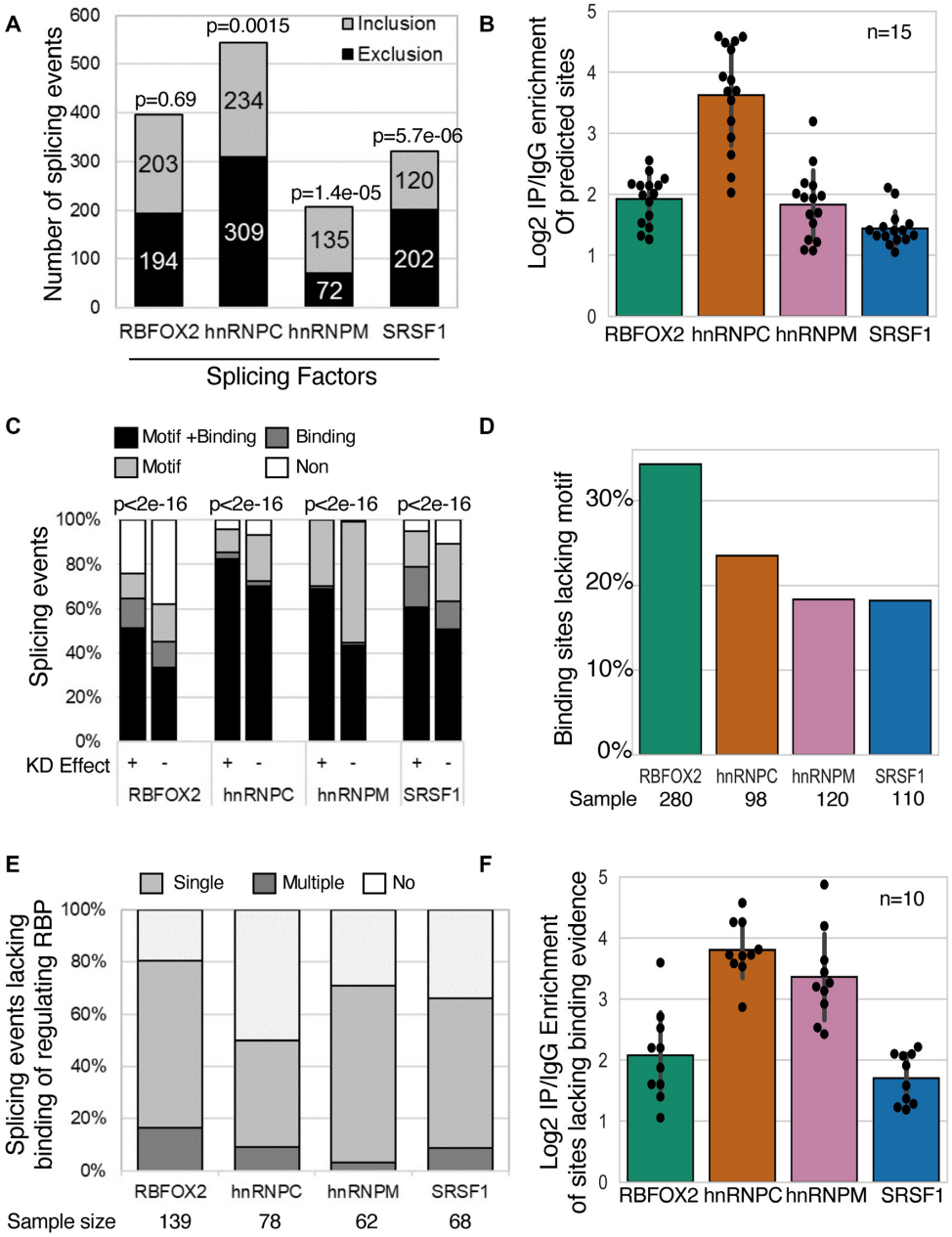

Figure 2. RBFOX2-interacting proteins alter splicing using different combinations of binding sites and motifs. (**A**) Impact of the knockdown of RBFOX2 and its interacting proteins on splicing. RBFOX2, hnRNPC, hnRNPM and SRSF1 were knocked down using two distinct siRNAs in SKOV3.ip1 cells and the RNA was sequenced using TGIRT-seq. Cassette exon events with at least 150 aligned reads and that were significantly altered by the RBP knockdown ($|\Delta\Psi| \geq 15\%$, Bayes factor ≥ 10) were identified and the number of affected inclusion and exclusion events presented in the form of a bar graph. Binominal tests were performed to evaluate the significance of changes in splicing ratio. (**B**) Confirmation that RBFOX2 core proteins interact with their splicing targets in SKOV3.ip1 cells. RBFOX2 and its protein partners were immunoprecipitated as described in Figure 1G and the enrichment of 15 randomly selected splicing events affected by each protein knockdown was evaluated using RT-qPCR. Error bars indicate standard deviation of the enrichment distribution. (**C**) Identification of the binding sites and motifs located near the splice targets of RBFOX2 core proteins. The distribution of the motifs and binding sites near splice sites affected by the knockdown of RBFOX2 (n = 391), hnRNPC (n = 540), hnRNPM ($n = 207$) and SRSF1 (n = 322) were identified and compared to those obtained with unaffected splice sites ($n = 14\,457$). Black, light grey, dark grey, and white columns indicate splice sites featuring a binding site with a canonical motif, binding sites without motif, motifs without binding or neither motif nor binding site. The significance of difference between the affected and unaffected splice sites was evaluated using Chi-squared tests and the obtained p-values indicated on top. (**D**) RBFOX2-interacting proteins bind to RBFOX2-targeted splice sites lacking the RBFOX2 binding motif. The percent overlap between motifs containing and motifs lacking binding sites near each RBP targeted splice sites was determined and is shown in the form of a bar graph. (**E**) Binding of RBFOX2-interacting proteins to splice sites affected by the knockdown of non-binding RBPs. The splice sites affected by, but not bound to, each RBP were identified and their association with the other interacting proteins examined and presented in the form of a bar graph. Single, Multiple and 'No', indicate splice sites associated with one, more than one or no proteins, respectively. (**F**) Validation of RBFOX2 core proteins binding to targeted splicing events lacking eCLIP sites. The association of RBFOX2 and its partners with targeted splice sites that lack eCLIP binding was monitored using immunoprecipitation as described in Figure 1G and the data shown in the form of a bar graph. 10 splicing events are randomly selected for each RBP and the error bars indicate the standard deviation of the enrichment distribution.

expected canonical motifs (Figure 2C). Comparison of the motif's distribution in the binding site with their random distribution in the genome indicates that the reduced motif enrichment of RBFOX2 relative to its protein partner is not due to variation in the redundancy/simplicity of motifs. In 35% of the cases, RBFOX2 binding in the absence of a canonical motif could be explained by the presence of a canonical binding site of one of its partner proteins (Figure 2D). This motif-independent binding was less evident in the case of hnRNPC, hnRNPM and SRSF1. Similarly, most (80%) RBFOX2-dependent splicing events that do not feature its binding sites include a canonical binding site of one of its partners, which could be co-immunoprecipitated with RBFOX2 (Figure 2E and F and Supplementary Figure S3B). These data indicate that in most cases hnRNPC, hnRNPM and SRSF1 act as recruiters of RBFOX2 to splice sites lacking its binding site. This mode of indirect recruitment (binding through the binding motif of a protein partner) is also observed in the cases of hnRNPC, hnRNPM and SRSF1 (Figure 2E and F and Supplementary Figure S3B). Collectively, >90% of all the splicing events affected by RBFOX2, hnRNPC, hnRNPM and SRSF1 included a canonical binding site and motif of at least one of the four interacting proteins. Together these data indicate that RBFOX2′s effect on splicing is not limited to splicing events featuring its canonical binding site but it could also modulate splicing after indirect recruitment by one of its protein partners.

RBFOX2 identifies splicing targets using three different binding modes

To better define the different binding configurations of RBFOX2, we classified the binding events based on the presence of the binding motif and the overlap between the different RBFOX2-interacting protein binding sites. As indicated in Figure 3A, the splicing events affected by RBFOX2 could be divided into three groups. The first include a single canonical RBP binding site (single), the second include overlapping canonical binding sites of two or more proteins (multiple), and the third (secondary binding) includes two overlapping binding sites, one of which contains a canonical binding motif and the second lacking known motifs. To determine the contribution of the different proteins to the different modes of RBFOX2 binding, we monitored the effect of knocking down RBFOX2-interacting proteins on the association of RBFOX2 with its targeted RNA. Overall, the knockdowns of RBFOX2-interacting proteins were more likely to inhibit RBFOX2 binding in the multiple mode than in the single and secondary modes of binding (Figure 3B and Supplementary Figure S4). Together these data suggest that the contribution of RBFOX2-interacting proteins to RNA binding varies between binding modes and the partner protein is most often required for splicing control in the multiple binding mode.

The splicing outcome of RBFOX2 core proteins is determined by the identity and position of their primary binding site

To evaluate the functional relevance of the binding modes we examined their distribution near the different splice sites and their association with exon inclusion. In the majority of cases, at least one protein binding site was identified in the 500 nucleotides adjacent to the splice sites affected by the proteins' knockdown (Figure 4A, listed in Supplementary Table S9). The majority of splice sites (635 sites, 70.4%) featured a single motif-containing binding site (single binding mode), while 133 sites or 14.7% of the splice sites had two or more motif containing binding sites (multiple binding mode). Finally, 144 sites or 14.9% included one motif containing binding site and one motif lacking binding site (secondary binding mode). The location of binding sites and percent inclusion of the adjacent cassette exons varied based on the identity of primary binding protein. For example, while most canonical RBFOX2, hnRNPC and hnRNPM-binding sites were found in the intronic sequence, many more SRSF1 primary binding sites were found within the cassette exon. The splicing outcome of RBFOX2 core proteins also varied based on the binding location, and the identity, of the primary binding protein (Figure 4B). Binding of RBFOX2 downstream of the cassette exon increased the chance of exon inclusion (knockdown increases exclusion), while binding of hnRNPC and hnRNPM downstream of the exon or the binding of SRSF1 to the exon increased exclusion (knockdown increased inclusion). The distance between the binding site and the splice site also had an effect on splicing in a protein-dependent manner. For example, binding far from the splice site increased exon inclusion when hnRNPC and hnRNPM are the primary binders (decreased inclusion upon knockdown) and inhibited inclusion when RBFOX2 is the primary binder (increased inclusion upon knockdown, Figure 4B). Notably, the binding location and splicing outcome also varied between binding modes (Figure 4C and D and Supplementary Figure S5). As summarized in Figure 4D, RBFOX2 single binding is more likely to cause exon inclusion when bound downstream than upstream of the exon. In contrast, in secondary mode RBFOX2 binds predominantly within the cassette exon or downstream of the exon via the canonical binding of SRSF1. Similarly, SRSF1 usually causes exon inclusion by binding to the cassette exon in the single binding mode, while it binds more frequently to the flanking introns in the multiple and secondary modes. We conclude that exon inclusion is determined at least in part by the location and identity of the RBFOX2-interacting protein's primary binding site.

RBFOX2 binding modes regulate the splicing of genes with different functions

To explore the possible functional implications of the variation in RBFOX2 binding modes, we identified the gene ontology terms and reactome pathways of the genes affected by each binding mode. As indicated in Figure 5A and Supplementary Figure S10, genes targeted by the different binding modes featured distinct gene ontology terms. The single binding mode associated with genes implicated in cell proliferation and survival (e.g. cell division, apoptosis and cell adhesion) while the multiple and secondary modes associated with genes implicated in different levels of RNA synthesis and metabolism (e.g. transcription, translation and splicing). The function of the targeted genes also varied be-

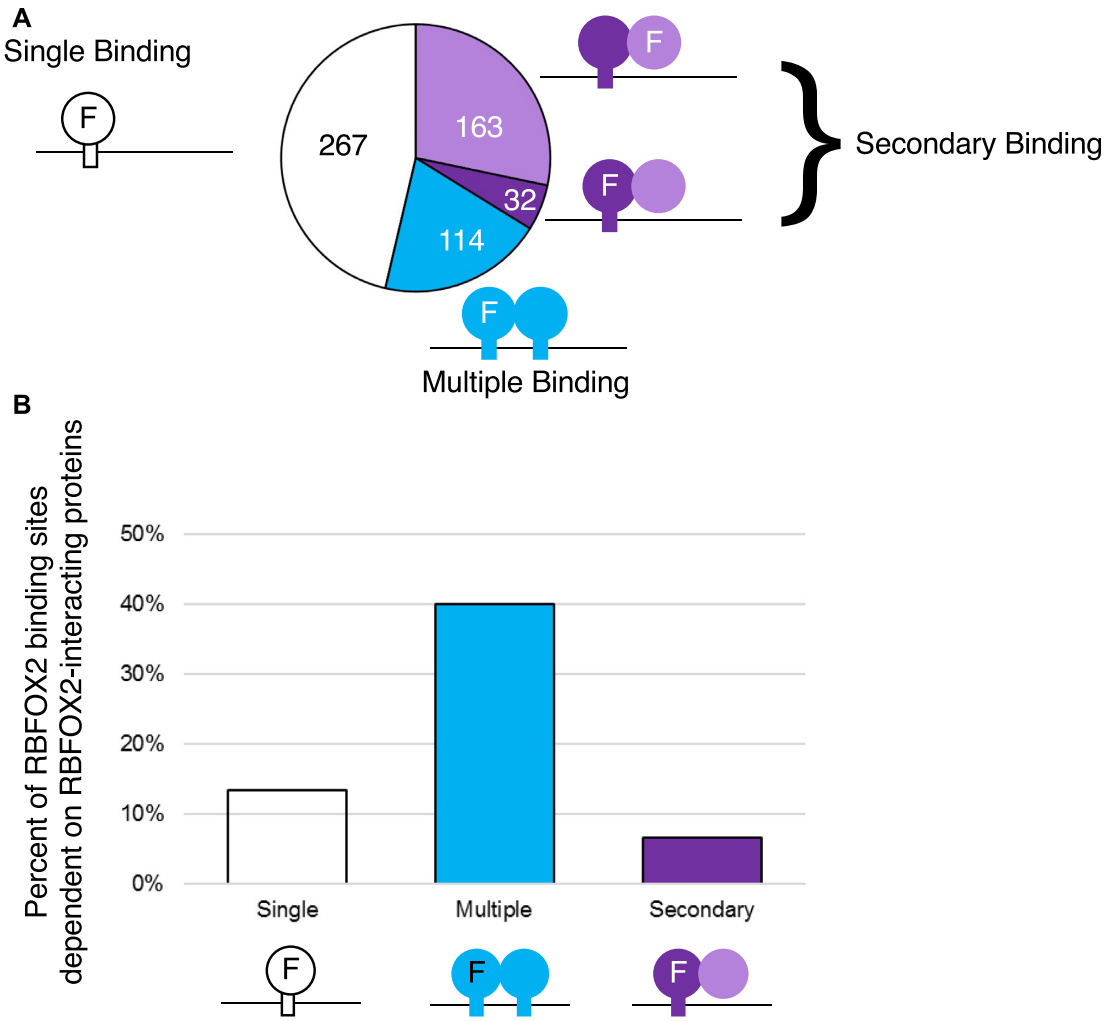

Figure 3. RBFOX2 regulates splicing using different binding configurations. (**A**) Types and distribution of RBFOX2 binding modes. The splice sites affected by the RBFOX2 knockdowns were separated into groups of single, multiple and secondary binding modes based on the number of binding site and the presence of canonical motifs. RNA binding proteins, binding motifs and mRNA are indicated by circles, rectangles and solid lines, respectively. (**B**) Effect of knocking down RBFOX2 protein partners on the different binding modes of RBFOX2. RBFOX2 was immunoprecipitated in SKOV3.ip1 cells before and after the knockdown of hnRNPC, hnRNPM and SRSF1 and the association of splice sites with RBFOX2 was examined using RT-qPCR. Shown are the percent of RBFOX2 binding sites in each binding mode that bind less efficiently to RBFOX2 after the knockdown of RBFOX2-interacting proteins. Binding was considered affected when the association with RBFOX2 was reduced by 40% and with p-value of <0.05. The different modes of binding are illustrated at the bottom. F indicates RBFOX2.

tween the primary binding proteins, albeit to a lesser extent. Most primary binding protein-dependent differences were detected in the single mode, while most functional overlap was observed in the secondary binding mode (Figure 5B). The functional specialization of the binding modes is also supported by differences in the function of the protein domains targeted by the different binding modes (Supplementary Figure S6). The single-binding-targeted domains were implicated in cytoskeleton remodeling consistent with the gene ontology's predicted effect of single binding on cell proliferation. Similarly, the multiple mode affected the splicing of domains implicated in transcription, corroborating the corresponding proposed gene ontology effect. Together these data suggest that the cellular effects of RBFOX2 and its partner proteins may vary based on the binding mode and identity of the primary binding protein.

DISCUSSION

In this study, we demonstrate that RBFOX2 splice site selection and outcome is modified at least in part by a set of core RNA binding proteins that include hnRNPC, hnRNPM and SRSF1 (Figures 1 and 2). The results confirm previously proposed interactions between RBFOX2 and hnRNPM and hnRNPC (28) and identify new physical and functional interactions between RBFOX2 and SRSF1, which is implicated in development and cell differentiation (53–56). SR proteins are considered to be 'essential' alternative splicing factors whereas RBFOX2 is the archetypal 'facultative' alternative splicing factor and as such connecting SRSF1 with RBFOX2 places these two proteins' interactions at the very hub of alternative splicing. RBFOX2-interacting proteins do not only recruit RBFOX2 to RNAs lacking an RBFOX2 binding motif but they may also mod-

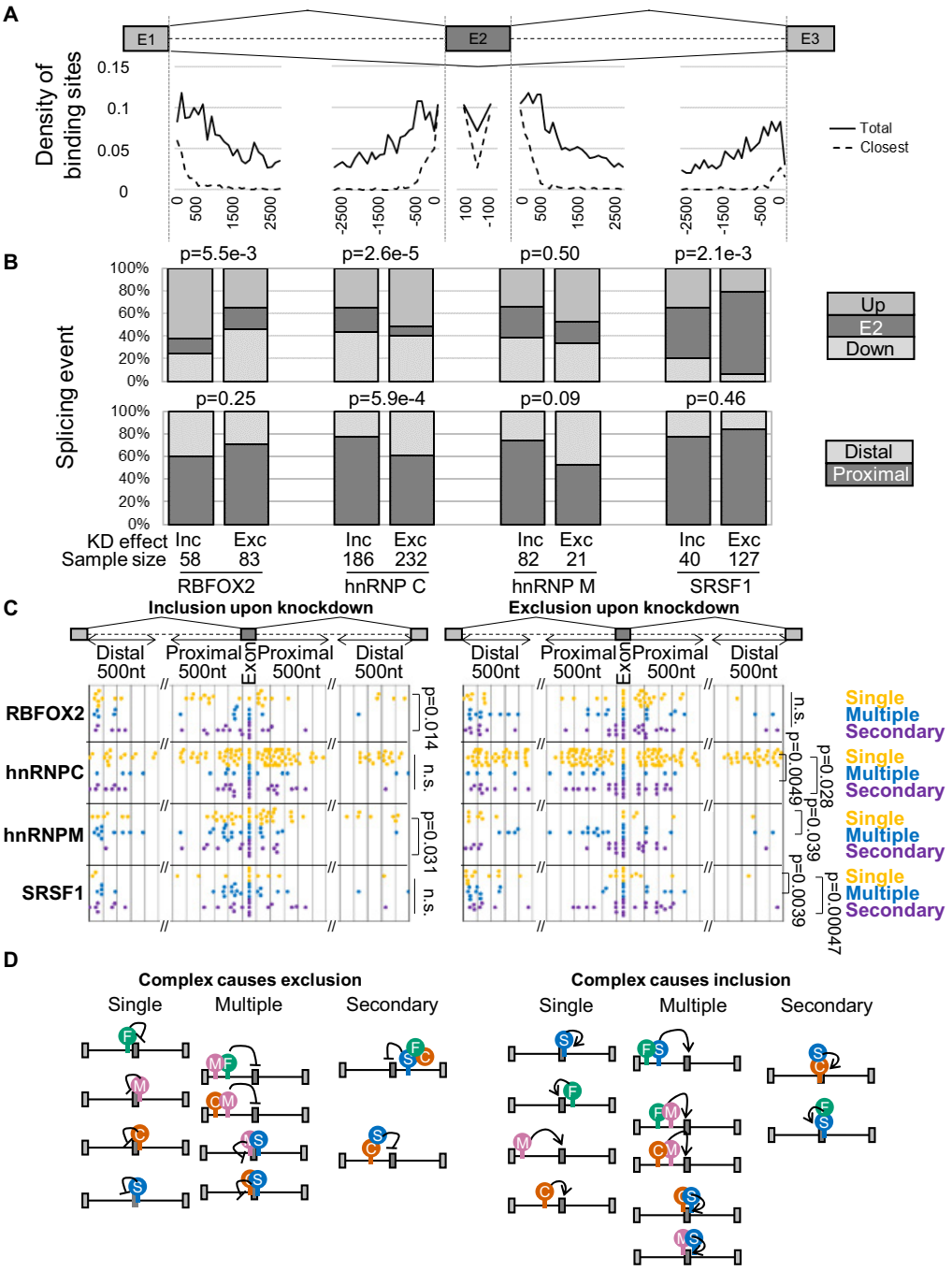

Figure 4. The location and splicing outcome of RBFOX2 core proteins is determined by the identity of the primary binding protein. (**A**) Distribution of RBFOX2 core proteins' binding sites relative to alternatively spliced exons. The distribution of the different protein binding sites in the 2500 nucleotides adjacent to the splice sites affected by the knockdown of RBFOX2 core proteins is plotted as a line chart. The solid line indicates all detected binding sites while the dashed line indicates the closest binding sites to the exon/intron boundary. The positions of the exons (boxes) and introns (dashed line) are indicated on top. (**B**) Correlation between the identity of the primary binding protein and the splicing outcome. The percent exon inclusion (Inc) and exclusion (Exc) associated with the closest motif containing binding site was determined and plotted as a function of the cassette exon (upper panel) or the exon/intron boundaries (lower panel). Up, E2 and Down indicate binding sites found within 500 nucleotides upstream of the cassette exon, the cassette exon, and 500 downstream of the cassette exon, respectively. Proximal and 'distal' indicate binding sites found within 500 nucleotides of the cassette exon and more than 500 nucleotides from the cassette exon, respectively. The number of events in each category is indicated at the bottom of the panel. The significance of the difference between the proportion of inclusion and exclusion events associated with each primary binding protein was calculated using Chi-squared tests and the *p*-value indicated on top. (**C**) Comparison between the location of the primary binding site and splicing outcome of the different RBFOX2 binding modes. The location of the primary binding sites (colored dots) of RBFOX2 and its partner proteins was determined in the single, multiple and secondary binding modes and plotted relative to their targeted splice sites. The significance of difference between the binding-site-distribution in each binding mode was calculated using Chi-squared tests and indicated on the right. (**D**) Schematic representation of the effect of the different RBFOX2 binding configurations on splicing. Green (F), orange (C), magenta (M) and blue (S) circles represent RBFOX2, hnRNPC, hnRNPM and SRSF1, respectively.

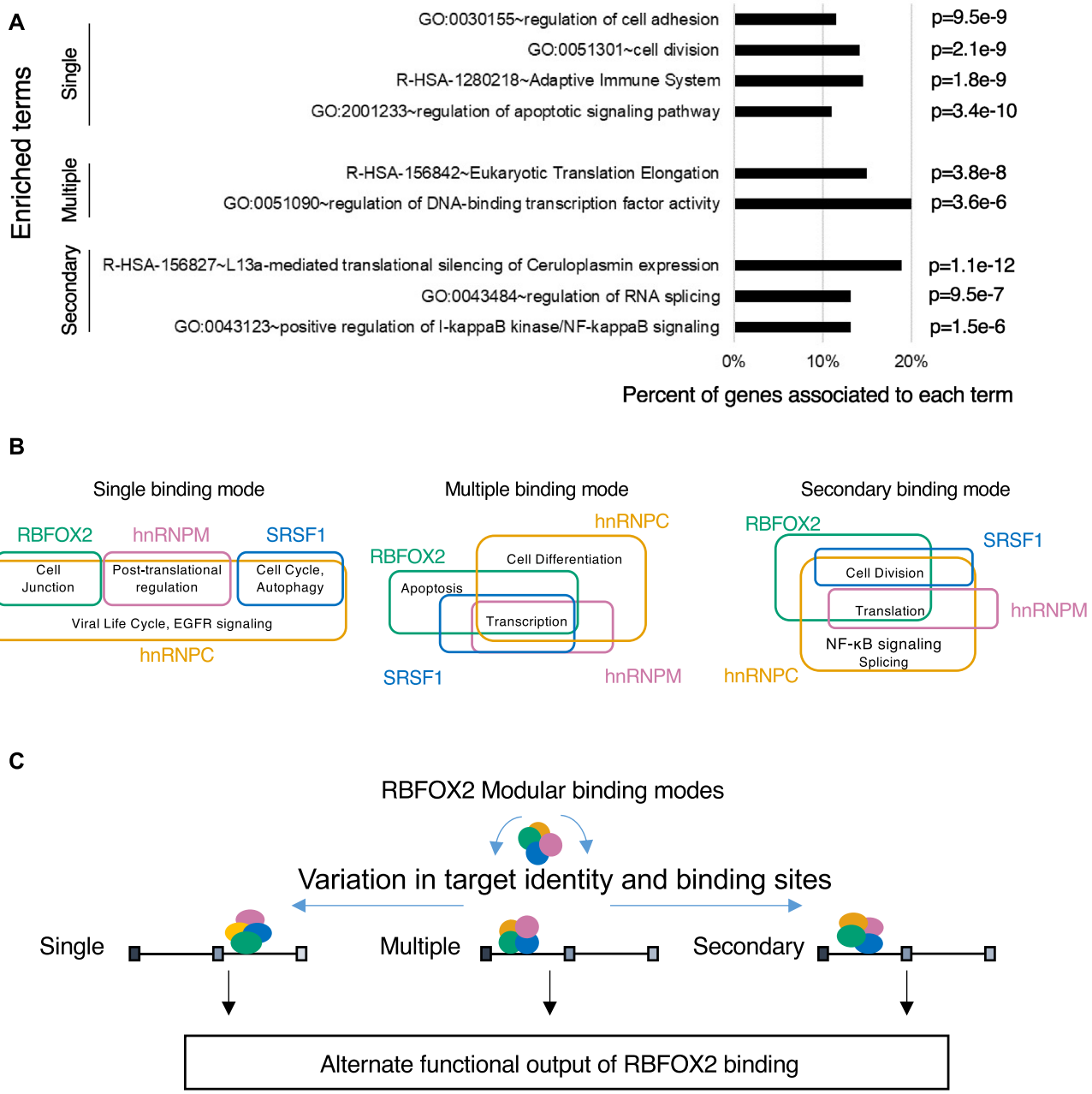

Figure 5. RBFOX2 binding modes target genes with different cellular functions. (**A**) Gene ontology analysis of genes targeted by the different RBFOX2 binding modes. Genes undergoing splicing events regulated by the different binding modes were determined and the enriched functions evaluated using gene ontology. Terms with Benjamini-adjusted *p*-value <0.05 were retained. The significance of difference in gene ontology between the different binding modes is calculated using Chi-squared tests and indicated on the right. (**B**) Descriptive representation of the gene ontology-predicted cellular functions affected by the different RBFOX2 binding modes as a function of the primary binder. The identity of the primary binding protein is indicated near its targeted cellular functions. (**C**) Proposed model of RBFOX2 binding and function. RBFOX2, hnRNPC, hnRNPM and SRSF1 are depicted as green, orange, magenta and blue ovals, respectively. The boxes and lines indicate exons and introns respectively.

ify the position of RBFOX2 binding relative to the splice site and its effect on the adjacent exon (Figure 1 and 4). RBFOX2-interacting proteins do not affect splicing as one static entity with a single splice outcome but rather through modular binding configurations with different targets and functional outcomes. This modular binding mode explains why many RBFOX2-targeted splicing events do not contain the RBFOX2 canonical (U)GCAUG binding motif and why RBFOX2 does not always bind at a fixed distance from the targeted alternative exons. We have shown the RBFOX2-interacting proteins may recruit it to splice sites lacking its binding motif and modify RBFOX2 binding position and splice outcome (Figures 1 and 4). The different binding configurations of RBFOX2 target genes with different cellular functions explain how a single RNA binding protein may contribute to different and even opposing

cellular functions (Figure 5). Together our results suggest a model of cooperativity between RNA-binding proteins that broadens the range of splice targets and functional outcomes beyond the limits of any single protein or splicing factor.

We find that only 64% of all splicing events affected by the depletion of RBFOX2 are actually bound directly by RBFOX2 and only 51% are bound through the canonical RBFOX2 motif (Figure 2). Traditionally, splicing events affected, but not bound, by a given RBP were considered artifacts caused by strain or technique variations or by secondary cascading effects of RBP depletion. However, the results presented here indicate that these events are likely direct targets of other components of the complex. This is supported by the fact that more than 80% of the RBFOX2-dependent splice sites were bound by at least one of its interacting proteins. Furthermore, we have shown that the alternatively spliced exons affected by RBFOX2 directly interact with RBFOX2 core proteins in the same cell line exhibiting the splicing event (Figure 2B and F). Therefore, we argue that most of the splicing events affected by the depletion of RBFOX2 in cell lines are genuine splicing targets of this protein. Indeed, given that we have only tested two out of the five proteins forming the stable (LASR) complex, it is quite possible that all RBFOX2-dependent splicing sites are bound to one, or more, RBFOX2-interacting proteins. However, it is also possible that RBFOX2 influences splicing through flexible and exchangeable configurations that may include a single protein, distinct binary complexes or multi-protein complexes. As the binding patterns of other RBFOX2-binding proteins are established, we will then be able to directly assess this hypothesis. That said, it remains formally possible that, in certain cases, weak or non-canonical RBFOX2 motifs contribute to the binding of these orphan splice sites lacking the canonical motif (19,57). However, these possible weak motifs are unlikely to be sufficient for binding since only one suboptimal motif (GCACG) and one weak motif (GCCUG) (57) are slightly enriched in orphan sites lacking the motif of all four RBFOX2 core complex proteins (Supplementary Figure S7) and none are enriched in the binding sites related to splicing.

Most alternative splicing factors tested experimentally were found to bind many sites across the transcriptome, even in unspliced RNA, making the interpretation of the functional or biological significance of these binding sites difficult (15). This also extends to cases where many binding sites of an RNA-binding protein are found in the same intron, making it difficult to identify the effective binding site regulating splicing. Here we have demonstrate that in most cases at least one RBP-binding site is found within 500 nucleotides of the splice site, suggesting that binding near the splice site is critical, if not essential, for the function of the RBFOX2 complex in splicing (Figure 4). This, of course, does not necessarily mean that distal binding sites are not effective or functional. Indeed, it was previously shown, using mini-genes, that in certain cases, distal binding sites may affect splicing, and several mechanisms were suggested for this effect, like RNA looping-out (58). However, the distribution of RBFOX2-binding sites near their targeted splicing events suggests that, at least for the RBFOX2 complex, the distal binding sites are rarely sufficient for splicing in the absence of any other proximal sites (Figure 4).

The systems analysis of RBFOX2′s splicing events, RNA-binding sites and interacting proteins presented here allows us to formulate a model where the cellular impact of RBFOX2 is defined by its mode of binding (Figure 5C). According to this model, RBFOX2 uses a variety of protein partners in various configurations to identify RNA featuring different motifs and binding sites. Depending on the composition of the binding complex and the identity of the primary binding sites, RBFOX2 and protein partners may target genes with different functions. According to this model, altering the expression or activity of RBFOX2 and partner proteins increases the regulatory spectrum of RBFOX2 and permits independent or differential regulation of different subsets of RNA targets. Indeed, it is now possible to explain how RBFOX2 programs the splicing of a specific group of mRNAs in different cells with the same level of RBFOX2 expression. Clearly there are considerable challenges ahead to decipher the repertoire of complex functional interactions of splicing factors, which are also exacerbated by the existence of multiple alternative isoforms of the splicing factors themselves.

DATA AVAILABILITY

RNA-seq data used in this study are available through the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo) under the accession number GSE140679.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Benoit Chabot and Julian Venables for critical reading of the manuscript.

FUNDING

Canadian Institutes of Health Research (CIHR) [PJT 153171 to S.A. and M.S.S.]; S.A. holds Canada Research Chair in RNA Biology and Cancer Genomics; M.S.S. holds a Fonds de Recherche du Québec – Santé (FRQS) Research Scholar Junior 2 Career Award. Funding for open access charge: CIHR.
Conflict of interest statement. None declared.

REFERENCES

1. Lee,Y. and Rio,D.C. (2015) Mechanisms and regulation of alternative pre-mRNA splicing. *Annu. Rev. Biochem.*, **84**, 291–323.
2. Roy,B., Haupt,L.M. and Griffiths,L.R. (2013) Review: Alternative splicing (AS) of genes as an approach for generating protein complexity. *Curr. Genomics*, **14**, 182–194.
3. Shukla,S. and Oberdoerffer,S. (2012) Co-transcriptional regulation of alternative pre-mRNA splicing. *Biochim. Biophys. Acta*, **1819**, 673–683.
4. Witten,J.T. and Ule,J. (2011) Understanding splicing regulation through RNA splicing maps. *Trends Genet.: TIG*, **27**, 89–97.
5. Gracida,X., Norris,A.D. and Calarco,J.A. (2016) Regulation of tissue-specific alternative splicing: *C. elegans* as a model system. *Adv. Exp. Med. Biol.*, **907**, 229–261.

6. Bezzi,M., Teo,S.X., Muller,J., Mok,W.C., Sahu,S.K., Vardy,L.A., Bonday,Z.Q. and Guccione,E. (2013) Regulation of constitutive and alternative splicing by PRMT5 reveals a role for Mdm4 pre-mRNA in sensing defects in the spliceosomal machinery. *Genes Dev.*, **27**, 1903–1916.
7. Gehman,L.T., Meera,P., Stoilov,P., Shiue,L., O'Brien,J.E., Meisler,M.H., Ares,M. Jr, Otis,T.S. and Black,D.L. (2012) The splicing regulator Rbfox2 is required for both cerebellar development and mature motor function. *Genes Dev.*, **26**, 445–460.
8. Calarco,J.A., Superina,S., O'Hanlon,D., Gabut,M., Raj,B., Pan,Q., Skalska,U., Clarke,L., Gelinas,D., van der Kooy,D. *et al.* (2009) Regulation of vertebrate nervous system alternative splicing and development by an SR-related protein. *Cell*, **138**, 898–910.
9. Fagnani,M., Barash,Y., Ip,J.Y., Misquitta,C., Pan,Q., Saltzman,A.L., Shai,O., Lee,L., Rozenhek,A., Mohammad,N. *et al.* (2007) Functional coordination of alternative splicing in the mammalian central nervous system. *Genome Biol.*, **8**, R108.
10. Clark,T.A., Schweitzer,A.C., Chen,T.X., Staples,M.K., Lu,G., Wang,H., Williams,A. and Blume,J.E. (2007) Discovery of tissue-specific exons using comprehensive human exon microarrays. *Genome Biol.*, **8**, R64.
11. Nagao,K., Togawa,N., Fujii,K., Uchikawa,H., Kohno,Y., Yamada,M. and Miyashita,T. (2005) Detecting tissue-specific alternative splicing and disease-associated aberrant splicing of the PTCH gene with exon junction microarrays. *Hum. Mol. Genet.*, **14**, 3379–3388.
12. Anko,M.L. (2014) Regulation of gene expression programmes by serine-arginine rich splicing factors. *Semin. Cell Dev. Biol.*, **32C**, 11–21.
13. Han,N., Li,W. and Zhang,M. (2013) The function of the RNA-binding protein hnRNP in cancer metastasis. *J. Cancer Res. Ther.*, **9**, S129–134.
14. Gallagher,T.L., Arribere,J.A., Geurts,P.A., Exner,C.R., McDonald,K.L., Dill,K.K., Marr,H.L., Adkar,S.S., Garnett,A.T., Amacher,S.L. *et al.* (2011) Rbfox-regulated alternative splicing is critical for zebrafish cardiac and skeletal muscle functions. *Dev. Biol.*, **359**, 251–261.
15. 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.
16. Brosseau,J.P., Lucier,J.F., Nwilati,H., Thibault,P., Garneau,D., Gendron,D., Durand,M., Couture,S., Lapointe,E., Prinos,P. *et al.* (2013) Tumor microenvironment-associated modifications of alternative splicing. *RNA*, **20**, 189–201.
17. Venables,J.P., Brosseau,J.P., Gadea,G., Klinck,R., Prinos,P., Beaulieu,J.F., 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.
18. Conboy,J.G. (2017) Developmental regulation of RNA processing by Rbfox proteins. *Wiley Interdiscip. Rev. RNA*, **8**, https://doi.org/10.1002/wrna.1398.
19. Lambert,N., Robertson,A., Jangi,M., McGeary,S., Sharp,P.A. and Burge,C.B. (2014) RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol. Cell*, **54**, 887–900.
20. Ponthier,J.L., Schluepen,C., Chen,W., Lersch,R.A., Gee,S.L., Hou,V.C., Lo,A.J., Short,S.A., Chasis,J.A., Winkelmann,J.C. *et al.* (2006) Fox-2 splicing factor binds to a conserved intron motif to promote inclusion of protein 4.1R alternative exon 16. *J. Biol. Chem.*, **281**, 12468–12474.
21. 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.
22. Huang,S.C., Ou,A.C., Park,J., Yu,F., Yu,B., Lee,A., Yang,G., Zhou,A. and Benz,E.J. Jr (2012) RBFOX2 promotes protein 4.1R exon 16 selection via U1 snRNP recruitment. *Mol. Cell. Biol.*, **32**, 513–526.
23. Braeutigam,C., Rago,L., Rolke,A., Waldmeier,L., Christofori,G. and Winter,J. (2014) The RNA-binding protein Rbfox2: an essential regulator of EMT-driven alternative splicing and a mediator of cellular invasion. *Oncogene*, **33**, 1082–1092.
24. Shapiro,I.M., Cheng,A.W., Flytzanis,N.C., Balsamo,M., Condeelis,J.S., Oktay,M.H., Burge,C.B. and Gertler,F.B. (2011) An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype. *PLoS Genet.*, **7**, e1002218.
25. Aparicio,L.A., Abella,V., Valladares,M. and Figueroa,A. (2013) Posttranscriptional regulation by RNA-binding proteins during epithelial-to-mesenchymal transition. *Cell. Mol. Life Sci.*, **70**, 4463–4477.
26. Venables,J.P., Vignal,E., Baghdiguian,S., Fort,P. and Tazi,J. (2012) Tissue-specific alternative splicing of Tak1 is conserved in deuterostomes. *Mol. Biol. Evol.*, **29**, 261–269.
27. 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.
28. Damianov,A., Ying,Y., Lin,C.H., Lee,J.A., Tran,D., Vashisht,A.A., Bahrami-Samani,E., Xing,Y., Martin,K.C., Wohlschlegel,J.A. *et al.* (2016) Rbfox proteins regulate splicing as part of a large multiprotein complex LASR. *Cell*, **165**, 606–619.
29. Singh,R.K., Xia,Z., Bland,C.S., Kalsotra,A., Scavuzzo,M.A., Curk,T., Ule,J., Li,W. and Cooper,T.A. (2014) Rbfox2-coordinated alternative splicing of Mef2d and Rock2 controls myoblast fusion during myogenesis. *Mol. Cell*, **55**, 592–603.
30. Lovci,M.T., Ghanem,D., Marr,H., Arnold,J., Gee,S., Parra,M., Liang,T.Y., Stark,T.J., Gehman,L.T., Hoon,S. *et al.* (2013) Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. *Nat. Struct. Mol. Biol.*, **20**, 1434–1442.
31. Misra,A., Ou,J., Zhu,L.J. and Green,M.R. (2015) Global analysis of CPSF2-mediated alternative splicing: Integration of global iCLIP and transcriptome profiling data. *Genom Data*, **6**, 217–221.
32. Brannan,K.W., Jin,W., Huelga,S.C., Banks,C.A., Gilmore,J.M., Florens,L., Washburn,M.P., Van Nostrand,E.L., Pratt,G.A., Schwinn,M.K. *et al.* (2016) SONAR Discovers RNA-binding proteins from analysis of large-scale protein-protein interactomes. *Mol. Cell*, **64**, 282–293.
33. Van Nostrand,W.E., Farrow,J.S., Wagner,S.L., Bhasin,R., Goldgaber,D., Cotman,C.W. and Cunningham,D.D. (1991) The predominant form of the amyloid beta-protein precursor in human brain is protease nexin 2. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 10302–10306.
34. Sundararaman,B., Zhan,L., Blue,S.M., Stanton,R., Elkins,K., Olson,S., Wei,X., Van Nostrand,E.L., Pratt,G.A., Huelga,S.C. *et al.* (2016) Resources for the comprehensive discovery of functional RNA elements. *Mol. Cell*, **61**, 903–913.
35. Quinlan,A.R. (2014) BEDTools: the Swiss-Army Tool for genome feature analysis. *Curr. Protoc. Bioinformatics*, **47**, 11.12.1–11.12.34.
36. Ray,D., Kazan,H., Cook,K.B., Weirauch,M.T., Najafabadi,H.S., 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.
37. Keene,J.D., Komisarow,J.M. and Friedersdorf,M.B. (2006) RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nat. Protoc.*, **1**, 302–307.
38. Brosseau,J.P., Lucier,J.F., Nwilati,H., Thibault,P., Garneau,D., Gendron,D., Durand,M., Couture,S., Lapointe,E., Prinos,P. *et al.* (2014) Tumor microenvironment-associated modifications of alternative splicing. *RNA*, **20**, 189–201.
39. Prinos,P., Garneau,D., Lucier,J.F., Gendron,D., Couture,S., Boivin,M., Brosseau,J.P., Lapointe,E., Thibault,P., Durand,M. *et al.* (2011) Alternative splicing of SYK regulates mitosis and cell survival. *Nat. Struct. Mol. Biol.*, **18**, 673–679.
40. Boivin,V., Deschamps-Francoeur,G., Couture,S., Nottingham,R.M., Bouchard-Bourelle,P., Lambowitz,A.M., Scott,M.S. and Abou-Elela,S. (2018) Simultaneous sequencing of coding and noncoding RNA reveals a human transcriptome dominated by a small number of highly expressed noncoding genes. *RNA*, **24**, 950–965.
41. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.
42. Kim,D., Pertea,G., Trapnell,C., Pimentel,H., Kelley,R. and Salzberg,S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.*, **14**, R36.

43. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
44. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R. and Genome Project Data Processing, S. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
45. Deschamps-Francoeur,G., Boivin,V., Elela,S.A. and Scott,M.S. (2019) CoCo: RNA-seq read assignment correction for nested genes and multimapped reads. *Bioinformatics*, **35**, 5039–5047.
46. Zerbino,D.R., Achuthan,P., Akanni,W., Amode,M.R., Barrell,D., Bhai,J., Billis,K., Cummins,C., Gall,A., Giron,C.G. *et al.* (2018) Ensembl 2018. *Nucleic Acids Res.*, **46**, D754–D761.
47. Katz,Y., Wang,E.T., Airoldi,E.M. and Burge,C.B. (2010) Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat. Methods*, **7**, 1009–1015.
48. Zhou,Y., Zhou,B., Pache,L., Chang,M., Khodabakhshi,A.H., Tanaseichuk,O., Benner,C. and Chanda,S.K. (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.*, **10**, 1523.
49. Jones,P., Binns,D., Chang,H.Y., Fraser,M., Li,W., McAnulla,C., McWilliam,H., Maslen,J., Mitchell,A., Nuka,G. *et al.* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics*, **30**, 1236–1240.
50. Klinck,R., Bramard,A., Inkel,L., Dufresne-Martin,G., Gervais-Bird,J., Madden,R., Paquet,E.R., Koh,C., Venables,J.P., Prinos,P. *et al.* (2008) Multiple alternative splicing markers for ovarian cancer. *Cancer Res.*, **68**, 657–663.
51. Yeo,G.W., Van Nostrand,E.L. and Liang,T.Y. (2007) Discovery and analysis of evolutionarily conserved intronic splicing regulatory elements. *PLos Genet.*, **3**, e85.
52. Klinck,R., Bramard,A., Inkel,L., Dufresne-Martin,G., Gervais-Bird,J., Madden,R., Paquet,E.R., Koh,C., Venables,J.P., Prinos,P. *et al.* (2008) Multiple alternative splicing markers for ovarian cancer. *Cancer Res.*, **68**, 657–663.
53. Blanco,F.J. and Bernabeu,C. (2012) The splicing factor SRSF1 as a marker for endothelial senescence. *Front Physiol*, **3**, 54.
54. Das,S. and Krainer,A.R. (2014) Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. *Mol. Cancer Res.*, **12**, 1195–1204.
55. Goncalves,V. and Jordan,P. (2015) Posttranscriptional regulation of splicing factor SRSF1 and its role in cancer cell biology. *Biomed. Res. Int.*, **2015**, 287048.
56. Sokol,E., Boguslawska,J. and Piekielko-Witkowska,A. (2017) The role of SRSF1 in cancer. *Postepy Hig. Med. Dosw (Online)*, **71**, 422–430.
57. Begg,B.E., Jens,M., Wang,P.Y., Minor,C.M. and Burge,C.B. (2020) Concentration-dependent splicing is enabled by Rbfox motifs of intermediate affinity. *Nat. Struct. Mol. Biol.*, **27**, 901–912.
58. Nasim,F.U., Hutchison,S., Cordeau,M. and Chabot,B. (2002) High-affinity hnRNP A1 binding sites and duplex-forming inverted repeats have similar effects on 5' splice site selection in support of a common looping out and repression mechanism. *RNA*, **8**, 1078–1089.