Molecular Cell
Article

Genome-wide Analysis Reveals SR Protein Cooperation and Competition in Regulated Splicing

Shatakshi Pandit,[1,4] Yu Zhou,[1,4] Lily Shiue,[2,4] Gabriela Coutinho-Mansfield,[1] Hairi Li,[1] Jinsong Qiu,[1] Jie Huang,[3] Gene W. Yeo,[1] Manuel Ares, Jr.,[2,*] and Xiang-Dong Fu[1,3,*]

[1]Department of Cellular and Molecular Medicine and Institute for Genomic Medicine, University of California, San Diego, La Jolla, CA 92093-0651, USA

[2]Center for Molecular Biology of RNA and Department of Molecular, Cell, and Developmental Biology, University of California at Santa Cruz, Santa Cruz, CA 95064, USA

[3]State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan, Hubei 430072, China

[4]These authors contributed equally to this work

*Correspondence: ares@ucsc.edu (M.A.), xdfu@ucsd.edu (X.-D.F.)

http://dx.doi.org/10.1016/j.molcel.2013.03.001

SUMMARY

SR proteins are well-characterized RNA binding proteins that promote exon inclusion by binding to exonic splicing enhancers (ESEs). However, it has been unclear whether regulatory rules deduced on model genes apply generally to activities of SR proteins in the cell. Here, we report global analyses of two prototypical SR proteins, SRSF1 (SF2/ASF) and SRSF2 (SC35), using splicing-sensitive arrays and CLIP-seq on mouse embryo fibroblasts (MEFs). Unexpectedly, we find that these SR proteins promote both inclusion and skipping of exons in vivo, but their binding patterns do not explain such opposite responses. Further analyses reveal that loss of one SR protein is accompanied by coordinated loss or compensatory gain in the interaction of other SR proteins at the affected exons. Therefore, specific effects on regulated splicing by one SR protein actually depend on a complex set of relationships with multiple other SR proteins in mammalian genomes.

INTRODUCTION

SR proteins are among the best-characterized splicing regulators in higher eukaryotic cells. They are involved in both constitutive and alternative precursor messenger RNA (pre-mRNA) splicing, making this family of RNA binding proteins (RBPs) unique compared to other RBPs that function only in regulated splicing (Lin and Fu, 2007; Long and Caceres, 2009; Zhong et al., 2009). SR proteins are RNA recognition motif (RRM)-containing RBPs, and each harbors a signature RS-rich domain(s). Both RRMs and RS domains have been implicated in the early steps of spliceosome assembly in which the RRM is responsible for RNA binding and the RS domain for protein-protein interactions. However, additional biochemical evidence suggests that RRMs and RS domains may also participate in protein-protein and protein-RNA interactions, respectively (Hertel and Graveley, 2005; Shen et al., 2004).

It has been generally accepted that SR proteins regulate alternative splicing by binding to exonic splicing enhancers (ESEs) to promote exon inclusion, which has been unequivocally demonstrated by tethering SR proteins to an alternative exon (Graveley and Maniatis, 1998). SR protein binding to ESEs is thought to enhance the recognition of weak splice sites by the splicing machinery, an activity equally applicable to both constitutive and regulated exons (Shen and Green, 2006). Heterogeneous nuclear ribonucleoprotein (hnRNP) A/B proteins can antagonize the positive effect of SR proteins on splice site selection (Mayeda and Krainer, 1992; Mayeda et al., 1994).

Despite the fact that SR proteins have been extensively characterized on model genes, it has been unclear whether and how various regulatory rules deduced from biochemical studies apply to endogenous transcripts. For example, in vitro binding or in vivo functional systematic evolution of ligands by exponential enrichment (SELEX) experiments have deduced consensus binding sequences for several SR proteins (Cavaloc et al., 1999; Liu et al., 2000, 1998; Schaal and Maniatis, 1999; Tacke and Manley, 1995). However, different assays appear to reveal distinct consensus motifs, suggesting that SR-RNA interactions may be quite degenerate, context sensitive, or assay dependent.

Such uncertainty in SR protein binding consensus makes it difficult to accurately predict SR protein binding sites in mammalian transcriptomes. For example, the ESEfinder program based on functional SELEX predicts SR ESEs on the majority of expressed pre-mRNAs (Cartegni et al., 2003; Lim et al., 2011). However, mapping of in vivo binding sites for SRSF3 (SRp20) and SRSF4 (SRp75) indicates that the two SR proteins only bind small distinct subsets of endogenous transcripts (Ankö et al., 2010; Änkö et al., 2012). A similar observation has also been made with *Drosophila* dASF and dSRp55 (Gabut et al., 2007), implying that the interactions of SR proteins with RNAs in vivo may be more selective than previously thought. Alternatively, these initial analyses may not have reached saturation nor had the sensitivity to capture most SR protein binding events on expressed transcripts.

The effect of SR proteins on alternative splicing has also remained a subject of debate. Through binding to ESEs, SR proteins function as positive splicing regulators to promote exon inclusion. However, increasing evidence suggests that they are also involved in exon-skipping events (Gallego et al., 1997; Ghigna et al., 2005; Lemaire et al., 1999; Solis et al., 2008), possibly through competition between the alternative exon and flanking constitutive exons (Sanford et al., 2009). For example, tethering an SR protein to the alternative exon induces its inclusion, but anchoring the SR protein in a flanking exon promotes alternative exon skipping (Han et al., 2011a). These results suggest that SR proteins control alternative exon inclusion in opposite directions depending on where they bind. The question remains as to how widely SR proteins use this strategy to regulate splicing in the cell and whether specific splicing outcomes could be directly linked to the positional effect of SR protein binding in vivo.

Here, we utilized crosslinking immunoprecipitation sequencing (CLIP-seq) in combination with splicing-sensitive arrays to address the in vivo RNA binding properties and functions of two classic SR proteins in regulated splicing. We found extensive overlap between SRSF1 and SRSF2 in binding to exons, and we detected extensive induction of both exon-inclusion and -skipping events in response to depletion of either SR protein. Surprisingly, we found little correlation between SR protein binding and induced splicing changes. Further analysis revealed that the loss of RNA binding by one SR protein induces changes in RNA binding by another SR protein, suggesting compensatory or synergistic actions of the remaining SR proteins and other splicing regulators in specific SR protein-depleted cells. Mutational analysis revealed that SR protein depletion-induced exon inclusion could be switched to skipping by preventing such compensatory responses, thus suggesting a general regulatory principle that emphasizes the collective contribution of multiple SR proteins to regulated splicing in mammalian transcriptomes.

RESULTS

Preferential and Overlapping Binding of SRSF1 and SRSF2 on Exons

We previously generated two MEF cell lines from conditional SRSF1 and SRSF2 knockout mice (Lin et al., 2005). Each of these MEF lines is functionally complemented with the respective HA-tagged exogenous gene expressed from a Tet-Off promoter at the level equivalent to that of the endogenous counterpart, as shown earlier (Lin et al., 2005) and reconfirmed before beginning this study (data not shown). This system permits controlled depletion of SR protein with doxycycline (Dox) and CLIP-seq analysis for each protein using the same anti-HA antibody, which is efficient in immunoprecipitation (Figure 1A). Both immunoprecipitated SR proteins could be labeled with γ-^{32}P-labeled ATP, likely due to associated SR protein kinases, as noted earlier (Sanford et al., 2008). This provides the location reference for bulk immunoprecipitated SR proteins. Using T4 kinase, we detected protein-RNA adducts that were sensitive to nuclease trimming (Figure 1B). We isolated such adducts linked with 30–60 nt RNA and performed deep sequencing using the established CLIP-seq protocol (Ule et al., 2005; Xiao et al., 2012; Xue et al., 2009; Yeo et al., 2009). We initially obtained 3,694,535 and 4,874,935 nonredundant tags for SRSF1 and SRSF2, respectively, that could be uniquely mapped to the mouse genome (version mm9).

The two SR proteins under investigation bind both exons and introns. However, since exons are much shorter than introns, we detected an enrichment of ~15-fold in exons after length normalization, suggesting that SR proteins prefer exonic sequences, as expected based on their known RNA binding properties (Figure 1C). This profile largely agrees with data for SRSF1 in HEK293T cells (Sanford et al., 2009), but our data provide more coverage and allow comparison between the two SR proteins. We detected 0.3–0.4 million tags for SRSF1 and SRSF2 on intron-exon and exon-exon boundaries, indicating that both SR proteins bind pre-mRNA and spliced mRNA (Figure 1D). Considering SRSF2 is a nonshuttling SR protein (Cáceres et al., 1998), its binding to spliced mRNA suggests it is removed or replaced by other shuttling SR proteins prior to mRNA export.

The most striking observation is an extensive overlap between the two SR proteins in binding to both constitutive and alternative exons, as seen on a representative gene (Figure 1E). Some binding events are unique to one or the other SR protein and detectable in both intronic and exonic regions (red arrows in Figure 1E). We observed only background binding on MEFs not expressing any HA-tagged protein (Figure 1E). The observed binding pattern is also completely different from that of multiple other HA-tagged RNA binding proteins we analyzed by CLIP-seq (data not shown). Thus, the exon-central binding profiles of the two SR proteins are unlikely to result from nonspecific binding to the HA tag or immunoglobulin G (IgG) beads.

We deduced 50,983 and 56,336 binding clusters (peaks; Benjamini-Hochberg corrected p value $< 1 \times 10^{-5}$) for SRSF1 and SRSF2, respectively, in various classes of RNA including intron-containing large intergenic noncoding RNAs (lincRNAs); however, the predominant class, by far, is intron-containing, protein-coding pre-mRNA (Figure S1A). The interaction of SR proteins with microRNA is consistent with the reported role of SRSF1 in the regulation of microRNA biogenesis (Wu et al., 2010). SRSF1 and SRSF2 each bind, on average, ~30% of the exons among 7,385 (SFRS1) and 7,226 (SFRS2) expressed transcripts that contain introns at the current tag density (Figure S1B). Gene expression (detected by Affymetrix exon junction arrays, see Experimental Procedures) positively correlates with the degree of collective SR binding events on exons, but much less in intronic regions (Figures 1F, S1C, and S1D). SR protein binding is significantly enriched on the internal exons compared to binding on either the first or the last exon (Figures S1E and S1F). Reduced binding on the first exon likely reflects a functional interplay of SR proteins with DNA and RNA at the promoter-proximal region (X. Ji and X.-D.F., unpublished data). Less frequent SR binding on the last exon may reflect the size of the last exon, which tends to be bigger than internal exons; thus, while the total number of binding events may be similar, the averaged binding density at individual locations may be lower.

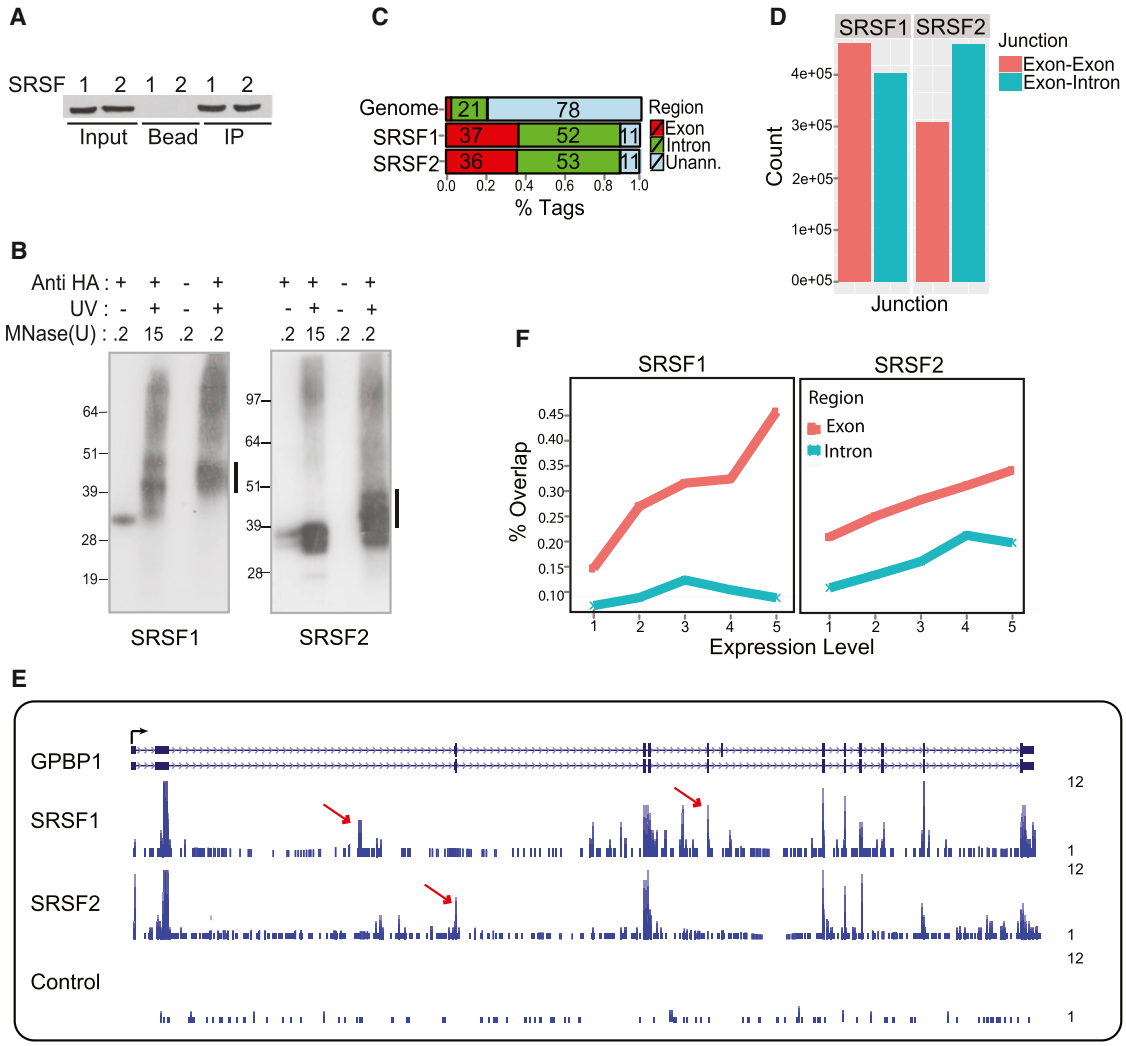


Figure 1. Genomic Landscape of SRSF1 and SRSF2 Binding

(A) Western blot of SRSF1-HA and SRSF2-HA using HA antibody, showing efficient immunoprecipitation of both SR proteins in MEFs.

(B) Autoradiograph of ^{32}P-labeled RNA-SR protein adducts with or without crosslinking. The line on the right side of each gel indicates the excised region for CLIP-seq.

(C) The genomic distribution of SRSF1 and SRSF2 tags relative to different proportions of specific genomic regions.

(D) The number of SRSF1 and SRSF2 tags mapped on exon-exon and exon-intron junctions.

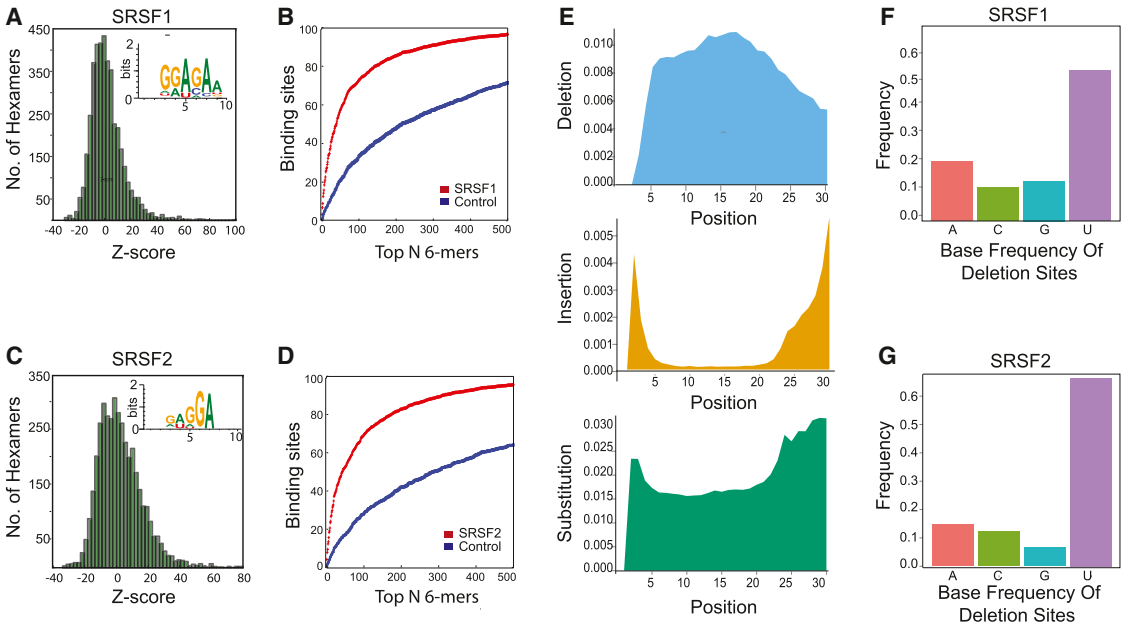
(E) The University of California, Santa Cruz (UCSC) Genome Browser view of SRSF1 and SRSF2 binding events on the *GPBP1* gene. Control indicates background tags from anti-HA CLIP on a wild-type MEF line that does not carry any HA-tagged protein. Arrows indicate unique binding peaks for a specific SR protein.

(F) Overlap in SR protein binding within exons and introns as a function of gene expression, showing that the overlap is more extensive in exons than introns. See also Figure S1.

In Vivo RNA Binding Specificity of SRSF1 and SRSF2

The mapped in vivo binding sites for SRSF1 and SRSF2 afford us the opportunity to address their preferred binding motifs in mammalian cells. While it is generally accepted that SRSF1 binds to GA-rich sequences, the RNA binding consensus for SRSF2 has eluded clear description. The in vitro binding SELEX experiments suggest that the RRM of SRSF2 binds to diverse sequences, some of which are also purine-rich (Cavaloc et al., 1999; Schaal and Maniatis, 1999; Tacke and Manley, 1995), but functional selection for SRSF2-responsive elements indicates that this SR protein prefers GC-rich sequences (Liu et al., 2000) that closely resemble the newly proposed consensus SSNG (where S is C or G and N corresponds to any nucleotide) from structural analysis of SRSF2 (Daubner et al., 2012). It has therefore remained unclear which of these motifs reflect the action of SRSF2 in vivo.

To identify the binding site consensus sequences for SRSF1 and SRSF2 proteins, we first took the standard approach of searching for overrepresented hexamer sequences in detected peaks against randomly sampled genomic sequence as



H

SR Protein	Motif	CIMS			Peak			Functional SELEX winners
		Target (%)	Bg (%)	p-val	Target (%)	Bg (%)	p-val	
SRSF1 Motif 1		46.83	18.88	1e-36	41.18	25.14	7.33e-191	
SRSF1 Motif 2		55.61	33.00	1e-20	59.76	44.38	1.62e-181	SRSASGA
SRSF1 Motif 3		35.12	17.57	1e-16	39.88	25.02	1.11e-180	(S= G/C; R= G/A; Y=C/U)
SRSF2 Motif 1		74.10	5.18	1e-478	13.55	8.63	9.01e-160	GRYYCSYR
SRSF2 Motif 2		39.36	13.95	1e-59	45.89	18.32	1.17e-216	UGCNGYY
SRSF2 Motif 3		31.98	14.76	1e-29	47.69	21.70	3.96e-210	

Figure 2. Motif Analysis for SR Protein Binding Sites in the Mouse Transcriptome

(A) Histogram of *Z* scores for hexamers in CLIP-seq peaks for SRSF1. Insert indicates the deduced consensus based on top hexamers.

(B) Percentage of SRSF1 peaks that contain top hexamers.

(C) Histogram of *Z* scores for hexamers in CLIP-seq peaks for SRSF2. Insert indicates the deduced consensus based on top hexamers.

(D) Percentage of SRSF2 peaks that contain top hexamers.

(E) Positional profiles of deletion (blue), insertion (yellow), and substitution (green) in sequenced tags from SRSF2 CLIP-seq tags. A similar pattern was also observed for SRSF1 CLIP-seq.

(F and G) Base frequency at deletion sites for SRSF1 (F) and SRSF2 (G). Deletion of uracil is predominant in both SR CLIP-seq data sets.

(H) Top three enriched motifs for SRSF1 and SRSF2 deduced by word counting and CIMS analysis and their representations in deletion-containing sequences or peaks relative to general background sequences.

background (Figures 2A and 2C). Based on *Z* scores, we identified the top 143 and 173 hexamers that cover ~80% of all detected peaks for SRSF1 and SRSF2, respectively (Figures 2B and 2D). We next used these top-ranking hexamers to deduce binding consensus for each SR protein (inserts in Figure 2A and 2B). We identified the GGAGA motif for SRSF1, which is consistent with the functional SELEX results (Liu et al., 1998) and with the consensus deduced from the initial CLIP-seq study of SRSF1 (Sanford et al., 2009). We could only detect the GA-rich motif for SRSF2, indicating that SRSF2 may bind to RNA with more degenerate motifs (Figure 2C).

SR Protein Binding Consensus Based on Induced Deletions and Genomic Distribution

Because the two proteins show extensive overlap in RNA binding, weaker motifs deduced solely from the CLIP sequences may be incidentally derived from binding sites closely occupied by different SR proteins. To address this, we applied a recently published approach that relies on crosslinking-induced mutation sites (CIMS) introduced during the CLIP-seq procedure (Zhang and Darnell, 2011) to more precisely locate the binding site. From all mapped binding events in the mouse transcriptome, we found that 3.5% (SRSF1) and 8% (SRSF2) of tags

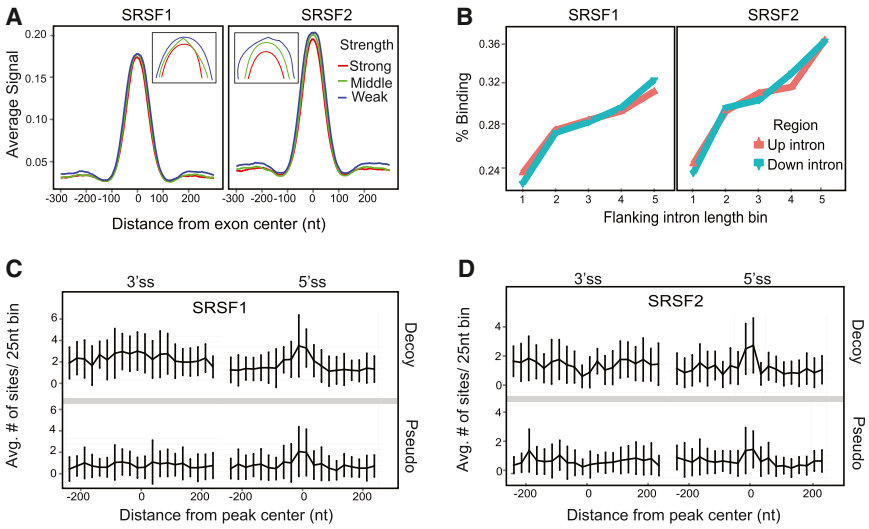


Figure 3. Correlation between SR Protein Binding and Splicing Features

(A) SR protein binding versus splice site strength. Inserts highlight a reverse correlation between SR protein bindings and splice site strength.

(B) Positive correlation between SR protein binding and the length of both upstream introns (Up intron) and downstream introns (Down intron). Intron bin 1: <350 nt; bin 2: 350 nt–1 kb; bin 3: 1 kb–2 kb; bin 4: 2 kb–4 kb; bin 5: >4 kb.

(C and D) Association of SR protein binding events with decoy or pseudo exons for SRSF1 (C) and SRSF2 (D) in the mouse genome. See also Figure S2.

have one or more internal deletions, the former being somewhat lower than the fractions (8%–20%) reported for Nova and Ago2 CLIP tags by Zhang and Darnell (2011). However, these internal deletions are typically distributed within the sequenced tags, distinct from the distribution of insertions and point mutations that were frequently found on the 3′ side of the tags due to sequencing errors (Figure 2E). Furthermore, U residues were more frequently depleted compared to the other three nucleotides (Figures 2F and 2G). Because U is most susceptible to UV crosslinking in general, the fact that SRSF1 binding motifs often do not contain Us possibly accounts for the relatively low deletion frequency we observe for this SR protein. This highlights the need to use multiple approaches for deriving motifs from CLIP data, as the crosslinking efficiency of individual nucleotides at variable positions in the motif may influence the numbers and kinds of reads detected.

To complement the analysis of crosslinking-induced deletions, we selected deletion-containing tags that were also mapped within SR binding peaks to deduce consensus motifs by using the HOMER (hypergeometric optimization of motif enrichment) software (Heinz et al., 2010), which is related to the MEME motif identification program (Bailey et al., 2009). This approach again yielded GA-rich consensus motifs for SRSF1, which are well represented in tags with internal deletions and those within mapped SRSF1 binding peaks (Figure 2H). In contrast, the most enriched motif for SRSF2 among tags containing internal deletions is a CU-rich sequence, followed by two less well-enriched motifs that contain a GC core. However, the top CU-rich motif accounts for only a small fraction (~14%) of the sequences within mapped SRSF2 peaks, while the next two enriched motifs are more representative (close to 50%) (Figure 2H). We note that the CU- and GC-rich sequences agree with the consensus sequences deduced from both binding and functional (in vitro splicing) SELEX experiments (Cavaloc et al., 1999; Liu et al., 2000). Our approach suggests a general approach for motif analysis based on CLIP-seq data by coupling the CIMS strategy with binding frequencies in mammalian transcriptomes.

Compensation for Weak Splice Sites and Long Introns by the Action of SR Proteins

The high-resolution mapping results permitted us to test a long-held theory concerning the ability of SR proteins to compensate for weak 5′ and 3′ splice sites in mammalian cells. We divided the 5′ and 3′ splice sites in the mouse genome into three categories, each based on their average maximum entropy score (Yeo and Burge, 2004), and queried CLIP-seq signals for SRSF1 and SRSF2 within each category. We observed an inverse correlation between SR protein binding and the strength of the splice sites (i.e., weaker splice sites showed stronger SR protein binding), which is true for both SRSF1 and SRSF2 (Figure 3A). We also performed a similar set of analyses on individual exons, with or without dividing them into constitutive or alternative exons, finding a similar trend (Figure S2). We noted in these analyses that the differences are rather small, implying that the contribution of SR proteins to the selection of weak splice sites might depend not on a single SR protein at each exon, but rather on the collective action of multiple SR proteins on each exon. As a result, the contribution of a given SR protein is less evident from such a metagene analysis.

As the size of flanking introns has been shown to influence the ability of an internal exon to be recognized in *Drosophila* (Fox-Walsh et al., 2005), we next tested whether SR protein binding on exons might compensate for the size of introns. For this purpose, we divided exons into five bins according to the length of the upstream or downstream intron and compared SRSF1 or SRSF2 binding on exons against the length of flanking introns. Strikingly, we found that binding of both SRSF1 and SRSF2 to exons correlates positively with the length of upstream and downstream introns (Figure 3B). These observations suggest that exons with longer flanking introns require stronger SR binding to be functionally defined in the mouse transcriptome.

Because half of binding sites for SRSF1 and SRSF2 were mapped to intronic regions, we asked whether some intronic binding events might reflect regulatory activities of SR proteins in splicing. A number of studies have implicated decoy exons in alternative splicing (Buratti et al., 2007; Havlioglu et al., 2007). Decoy exons only contain one potential 5′ or 3′ splice site, whereas pseudo exons carry both potential splice sites

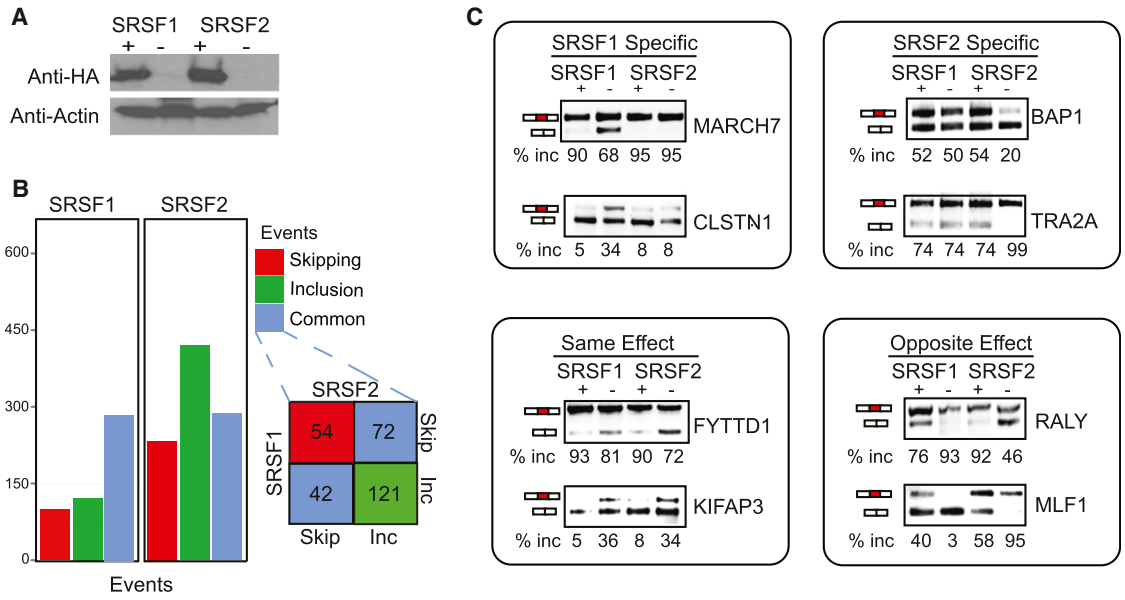


Figure 4. Induction of Exon Inclusion and Skipping in SR Protein-Depleted Cells

(A) Western analysis of SR protein depletion using Dox. Actin served as a loading control. "+" and "−" indicate the presence or absence, respectively, of the SR protein in mock-treated and Dox-treated cells.

(B) Bar graph presentation of altered alternative events detected by exon junction array and RASL-seq. Depletion of either SR proteins causes both exon inclusion and skipping. Splicing events affected by both are highlighted with similar or opposite responses to depletion of either SR protein on the right.

(C) Validation of SRSF1 and SRSF2 depletion-induced splicing of cassette exons in four different classes. The SR proteins exhibited direct binding in various locations on all of these illustrated genes. See also Figure S3 and Table S1.

separated by a sequence of up to 250 nt (Yeo and Burge, 2004). After centering the mapped intronic SR protein binding sites, we examined the frequency of potential 3′ or 5′ splice sites upstream or downstream of the peak. We found a small fraction of potential 5′ splice sites, but not 3′ splice sites, near intronic binding peaks for both SRSF1 and SRSF2 (Figures 3C and 3D). These observations indicate that only a small fraction of SR protein binding events in introns might represent decoy or pseudo exons. In this context, it is interesting to note that intronic SR protein binding events have been found to inhibit exon inclusion even though they are not part of pseudo exons or decoy exons (Erkelenz et al., 2013).

Extensive Involvement of SR Proteins in Exon-Inclusion and -Skipping Events In Vivo

We next investigated how specific SR protein binding events might be linked to the regulation of alternative splicing in the mouse transcriptome. As SRSF1 and SRSF2 were each expressed from a Tet-Off promoter, we were able to efficiently deplete them by adding Dox to the culture media (Figure 4A). After Dox-induced depletion, total RNA was isolated for profiling on splicing-sensitive exon junction microarrays (Du et al., 2010). To aid in validation of these candidate events, we divided the data in four bins and selected representative events (colored in Table S1) in each bin for RT-PCR validation. The result indicates that the overall validation rate for events in bins 1–3 is 88% (out of 84 events examined, 10 failed). We also complemented the array-based method with the RASL-seq technology (RNA-mediated oligonucleotide annealing, selection, and ligation coupled with deep sequencing) that we recently developed (Li et al., 2012; Zhou et al., 2012) and validated a large set of candidate events at the rate of 93.5% (46 examined, 3 failed) (Table S1). A representative set of validated events is shown for both SRSF1 (Figure S3A) and SRSF2 (Figure S3B).

In total, we detected 498 SRSF1-dependent and 912 SRSF2-dependent alternative splicing events, consistent with the function of these SR proteins as splicing regulators. Interestingly, among 498 altered splicing events detected in SRSF1-depleted MEFs, 225 events showed increased exon skipping and 276 exhibited increased exon inclusion. Similarly, in response to Dox-induced depletion of SRSF2, 312 exons increased in skipping while 601 exons increased in inclusion (Figure 4B). Therefore, both SR proteins appear to be more frequently involved in repressing, rather than promoting, exon inclusion. On the surface, this contradicts the widely perceived roles for SR proteins in promoting exon inclusion. We compared events affected by loss of each protein and found that SRSF1 and SRSF2 jointly regulate 288 alternative splicing events. Of these shared events, 39% are regulated in opposite directions (Figure 4B), suggesting that the two SR proteins function in a combinatorial fashion to regulate a subset of splicing events in the mouse transcriptome.

Our data show all possible combinations of alternative splicing control by the two SR proteins (Figure 4C). For example, SRSF1 depletion induced exon skipping (*MARCH7* exon 7) or inclusion (*CLSTN1* exon 3) while SRSF2 depletion showed no effect on these exons (Figure 4C, top left panel). The converse

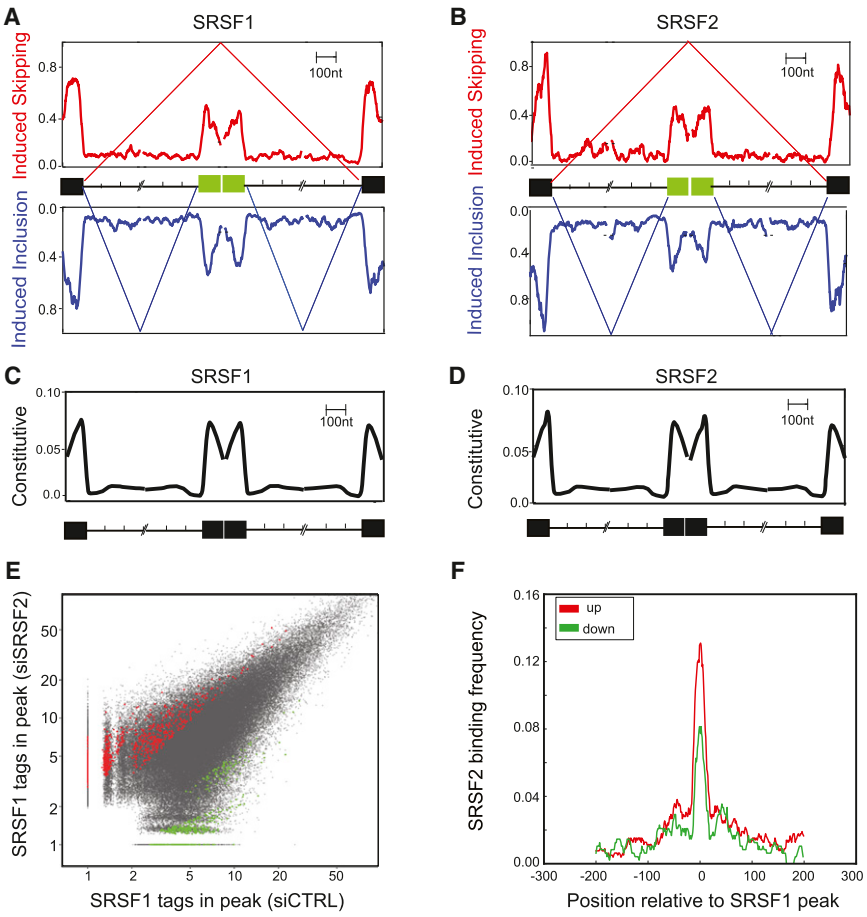


Figure 5. Composite RNA Maps for SRSF1 and SRSF2 and Interplay between the Two SR Proteins

(A and B) Composite maps for SRSF1 (A) and SRSF2 (B) on induced exon-inclusion or -skipping events based on the most reliable RASL-seq data streamlined by RT-PCR in addition to array-derived and RT-PCR-validated events.

(C and D) Composite maps of SRSF1 (C) and SRSF2 (D) constructed on constitutive exons.

(E) Scatter plot of SRSF1 binding peaks before (siCtrl) and after (siSRSF2) knockdown of SRSF2. x and y axes indicate the number of tags associated with each peak. Red and green indicate peaks that showed a significant increase or decrease, respectively, in SRSF1 binding in response to SRSF2 depletion.

(F) SRSF2 binding frequency on increased (red) and decreased (green) SRSF1 binding peaks.

was true on *BAP1* exon 5 and *TRA2α* exon 2 (Figure 4C, top right panel), which responded to SRSF2, but not SRSF1, depletion. These exons exhibited restricted responses to only one of the two SR proteins. On *FYTTD1* exon 7 and *KIFAP3* exon 20, depletion of either SR protein had the same effect (Figure 4C, bottom left panel). On *RALY* exon 4, SRSF1 depletion resulted in exon inclusion, but SRSF2 depletion induced exon skipping, and the converse was true on the *MLF1* exon 3 (Figure 4C, bottom right panel). These cases exemplify complex functional requirements for different SR proteins at different exons. Although each of these examples deserves detailed molecular dissection to understand the basis for the specific effects detected, the data clearly illustrate more complicated biological functions for SR proteins in regulated splicing than previously appreciated based on studies with model genes.

Functional Consequences Not Linked to Simple Positional Effects of SR Proteins

Recent genome-wide studies suggest that many splicing regulators influence splice site selection in a position-sensitive manner (Han et al., 2011b; Witten and Ule, 2011). For SR proteins, the tethering approach has demonstrated that SR protein binding on an internal cassette exon enhances inclusion of that internal exon, whereas tethering to a flanking constitutive exon stimulates skipping of the internal alternative exon (Han et al., 2011a). Our in vivo mapping results coupled with a large number of SR protein-dependent exon-inclusion or -skipping events presented an opportunity to explore potential positional effects of SR proteins from the global perspective.

Based on the large number of validated splicing events, in addition to those reliably detected by RASL-seq, we constructed cassette exon models separately for induced exon-inclusion and -skipping events and mapped SR binding information to the internal alternative and flanking constitutive exons as well as surrounding intronic regions (for the exon-proximal 300 nt, see Figures 5A and 5B). For comparison, we constructed a similar cassette exon model from a set of randomly selected constitutively spliced exons (Figures 5C and 5D). Unexpectedly, we could not see any relationship between SR protein binding and functional outcomes. The only observable feature common for both SR proteins is lower binding on the internal alternative exon compared to the flanking constitutive exons in the regulated cassette exon model, but this trend is not evident on the constitutive exon model (compare Figures 5A and 5B to 5C and 5D). This is consistent with those internal exons being categorized as alternative due to insufficient SR protein activity to ensure their constitutive inclusion. The puzzle, then, is why depletion of a specific SR protein would induce exon inclusion in one set of these genes but cause exon skipping in another set.

We hypothesized that specific functional outcomes might depend on the collective contribution of multiple SR proteins acting on individual alternatively spliced exons. As such, depletion of one specific SR protein might alter the competition between the alternative and flanking constitutive splice sites. For example, as illustrated earlier (Han et al., 2011a), the skipping of an alternative exon might increase in response to the removal of an SR protein if the SR protein contributes more to the

selection of the alternative exon than the flanking exons. Conversely, inclusion of the exon may increase if the depleted SR protein mainly acts on the flanking constitutive exon(s) while other SR proteins bind to the internal alternative exon. This would cause weakening of the flanking exon(s), thereby increasing the competitiveness of the alternative exon, leading to increased inclusion. In other words, the specific outcome after depletion of an SR protein for each exon under consideration depends on the activities of other functionally related SR proteins and their binding sites on the exon itself and on its flanking exons.

Compensatory Changes in SR Protein Binding to RNA

To test the above hypothesis, we mapped SRSF1 binding sites by CLIP-seq in MEFs depleted of endogenous SRSF2 by RNAi (RNA interference). We performed two biological replicates, yielding a total of 9,206,798 (no depletion) and 8,912,433 (SRSF2-depleted) nonredundant and uniquely mapped tags for SRSF1 (see Experimental Procedures). Because SR protein binding is positively correlated with levels of gene expression, we focused on transcripts with similar overall levels of gene expression and SR binding (<1.5-fold changes in total CLIP-seq tag density per transcript). We normalized each data set to the same number of total counts (5 million) and identified statistically significant changes in SRSF1 binding that were induced by SRSF2 depletion. This identified 786 increased (red) and 479 decreased (green) SRSF1 binding events (Figure 5E), indicating that binding of SRSF1 is influenced by the action of SRSF2 at many locations. These altered SRSF1 binding events do not seem enriched for alternative exons versus constitutive exons, although such changes on regulated exons might more readily lead to measurable functional outcomes. It is important to emphasize that CLIP-seq tags measure SR-RNA interactions before and after splicing and, therefore, do not strictly reflect final levels of spliced mRNA.

To further understand how SRSF2 binding influences the interaction of SRSF1 with RNA, we analyzed SRSF1 binding peaks that responded to SRSF2 depletion by determining the binding frequency of SRSF2 around SRSF1 binding sites. This analysis revealed that strong SRSF2 binding sites occurred more frequently on and near the sites that showed increased SRSF1 binding after SRSF2 depletion. In contrast, the frequency of strong SRSF2 binding sites was lower on and near sites where SRSF1 binding was reduced in SRSF2-depleted cells (Figure 5F). We interpret these observations to indicate that SRSF1 and SRSF2 compete for binding at many sites. On the sites that show less binding by both SR proteins, however, reduction in SRSF2 binding might indirectly enhance other RNA binding proteins, such as hnRNP A/B, to compete with adjacent SRSF1 binding events, thus causing a coordinated loss in SR binding.

Both coordinated loss and compensatory gain of SR proteins can be better appreciated on some representative genes (Figures 6A–6C). In the case of *HNRNPA2B1* transcripts, for example, SRSF1 and SRSF2 bound similarly to both alternative and flanking exons. Upon SRSF2 knockdown, SRSF1 binding was reduced in the middle alternative exon (arrow in Figure 6A) while little change was detected on the two flanking exons. This coordinated loss of both SR proteins likely contributes to weakening of the alternative exon, thus causing its efficient skipping in SRSF2-depleted cells (Figure 6D, left panel with a model below). On *CDC45I* transcripts, the reduction of SRSF2 binding allowed SRSF1 to bind better on two locations: one on the alternative exon and the other on the downstream constitutive exon (arrows in Figure 6B). The preferential gain in SRSF1 binding on the internal alternative exon may thus be responsible for the increased inclusion of the exon in response to SRSF2 depletion (Figure 6D, right panel with a model below). Similarly, on *CCNL1* transcripts, the reduction of SRSF2 binding allowed SRSF1 to bind better on the alternative exon (arrows in Figure 6C), possibly explaining the increased inclusion of the alternatively spliced exon in SRSF2-depleted cells (Figure 6D, right panel). These observations expose the complex contributions of individual SR proteins to regulation of alternative splicing that arise from competition or collaboration with one another and with other splicing regulators at specific sites in pre-mRNAs within the cell.

Conversion from SR Protein-Dependent Exon Inclusion to Skipping

Loss of SR protein binding on the internal alternative exon explains induced exon skipping after depletion according to the well-known function of SR proteins in promoting exon definition. Here, we wished to further explore the mechanism by which exon inclusion increases upon SR protein depletion. Given the gain in SRSF1 at some sites upon SRSF2 depletion, we wished to test whether an SRSF2 depletion-dependent increase in SRSF1 binding could be responsible for increased exon inclusion in SRSF2-depleted cells. We thus constructed two minigenes containing the alternative and flanking exons from the *CDC45I* and *CCNL1* genes, which were spliced similarly to the corresponding regions of their endogenous genes and showed altered SRSF1 binding after SRSF2 depletion (Figures 6E and 6F, bottom panel, first two lanes). We next introduced a deletion mutation (indicated by X) to the site in the alternative or downstream flanking constitutive exon that showed the alteration in SRSF1 binding after SRSF2 depletion. On *CDC45I* exon 5, SRSF2 knockdown slightly enhanced SRSF1 binding at the upstream site but reduced its binding at the downstream site. By mutating site 5.1 where SRSF2 depletion induced a minor gain of SRSF1 binding, we found that the mutant had a similar splicing ratio to the wild-type construct before SRSF2 knockdown and caused a small change in the inclusion of the internal alternative exon after SRSF2 knockdown (Figure 6E, bottom panel). The minor impact by the mutation in this flanking exon might be due to other sequence features that make this exon constitutive; in fact, this construct served as a control for other mutations analyzed.

We next tested mutations in the alternative exon 4 of *CDC45I*. The prediction is that specific deletion mutations in site 4.1 (the predominant binding site for SRSF2) and 4.2 (the predominant binding site for SRSF1) would each weaken the alternative exon in wild-type cells. A much bigger effect would be predicted in SRSF2-depleted cells because the Ex4.1 mutation would prevent the gain of SRSF1 binding in SRSF2-depleted cells, and the Ex4.2 mutation would weaken SRSF1 binding, which may synergize with the loss of SRSF2 binding. As predicted,

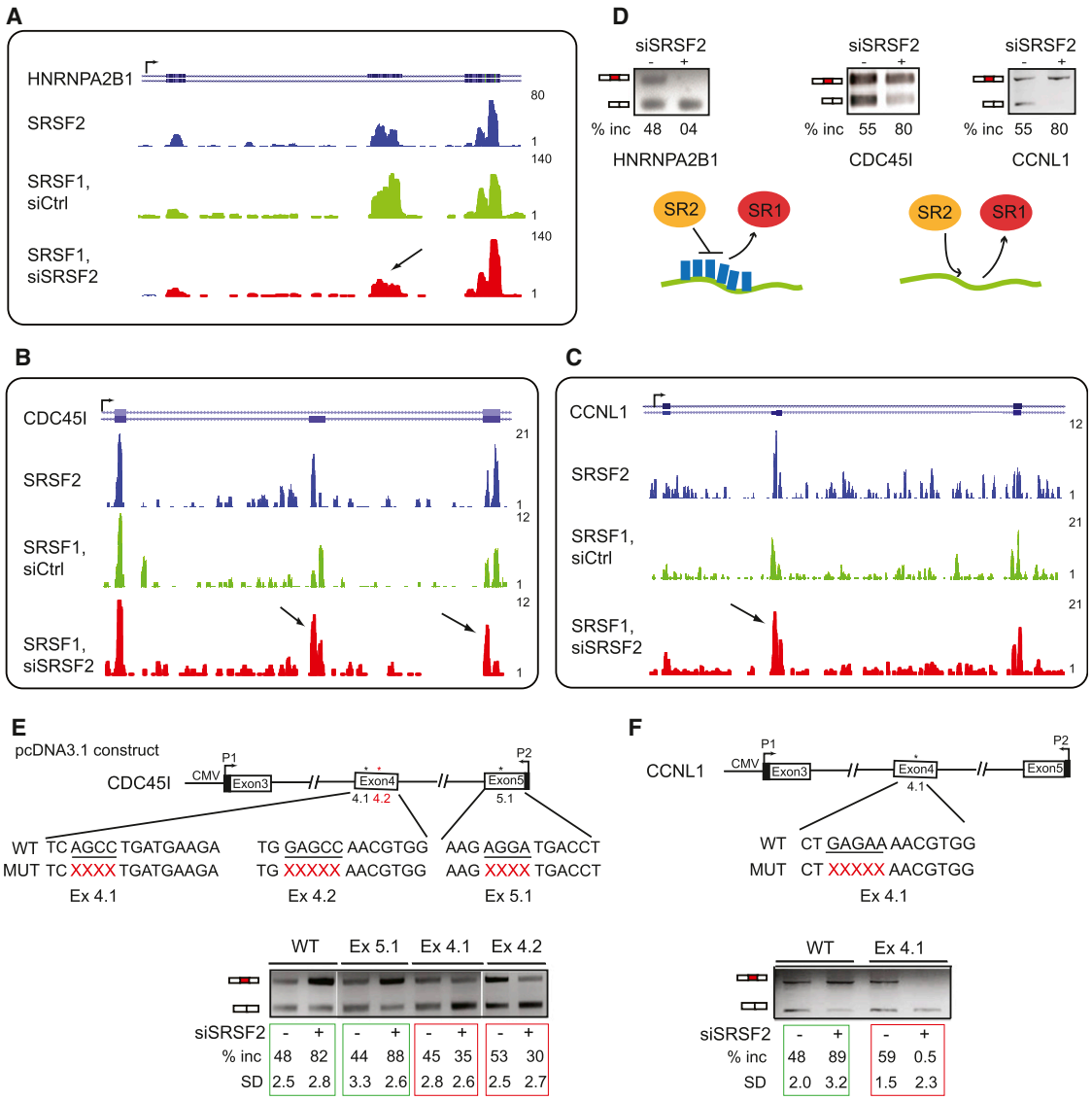


Figure 6. Synergistic and Compensatory Interactions of SR Proteins with RNA and Associated Functional Consequences

(A–C) UCSC Genome Browser views of the *HNRNPA2B1* (A), *CDC45I* (B), and *CCNL1* (C) genes with binding by SRSF2 (blue), SRSF1 before SRSF2 knockdown (green), and SRSF1 after SRSF2 knockdown (red).

(D) The splicing response of *HNRNPA2B1* (left), *CDC45I*, and *CCNL1* (right) to SRSF2 depletion. The proposed model on the left indicates potential competition of SR protein binding by other RNA binding proteins (blue bricks), which may quickly occupy the binding site vacated by an SR protein (SR1), thus preventing the compensatory binding by another SR protein (SR2). The proposed model on the right suggests that the site vacated by the first SR protein (SR1) is quickly occupied by a second SR protein (SR2).

(E) Graphical representation of the *CDC45I* minigene and specific deletion mutations (labeled X) introduced in SRSF1 binding sites in exon 5 (Ex5.1) and two locations in exon 4 (Ex4.1 and Ex4.2). P1 and P2 indicate PCR primers for splicing analysis. RT-PCR analyses were performed on wild-type (WT) and mutant minigenes before and after SRSF2 knockdown (bottom). The mutations in exon 4 caused the switch of SRSF2 depletion-induced exon inclusion (green boxes) to exon skipping (red boxes).

(F) Graphical representation of the *CCNL1* minigene and SRSF1 binding site deletion mutations (labeled X) introduced in exon 4 (Ex4.1). P1 and P2 indicate PCR primers for splicing analysis. Splicing of WT and mutant minigene before and after SRSF2 knockdown was analyzed by RT-PCR. The deletion mutation in exon 4 caused the switch of SRSF2 depletion-induced exon inclusion (green box) to exon skipping (red box).

we found that both Ex4.1 and Ex4.2 mutants showed a modest effect in wild-type cells. In contrast, both mutants flipped from SRSF2 depletion-induced exon inclusion to skipping (Figure 6E, bottom panel). We performed a similar analysis on the *CCNL1* minigene containing the alternative exon 4 and found that a deletion mutation introduced in the exon also switched SRSF2 depletion-induced exon inclusion to skipping (Figure 6F, bottom panel). Together, these data demonstrated that the splicing

response to changes in function of one SR protein could be profoundly influenced by compensatory actions of another SR protein on regulated exons.

DISCUSSION

The Ser/Arg-rich protein family of splicing factors is believed to play a key role in all regulated splicing events studied to date (Lin and Fu, 2007; Long and Caceres, 2009). The well-established concepts for the function of SR protein in regulated splicing include the following: (1) SR proteins preferentially bind to exonic sequences to promote splice site selection; (2) each SR protein has preferred binding site sequences in RNA, although different SR proteins may have shared binding sites; (3) individual SR proteins appear to have both unique and redundant functions in splicing; (4) SR proteins are largely responsible for promoting exon inclusion; and (5) the function of SR proteins in splice site selection can be antagonized by hnRNP proteins (Lin and Fu, 2007). By determining in vivo targets for the two classic SR family members and linking their RNA binding activities to functions in regulated splicing, this study has now elucidated several principles for the function and mechanism of action of SR proteins in mammalian cells.

Most Exons Are Recognized by More than One SR Protein

Consistent with their roles in constitutive and regulated splicing by binding exonic splicing enhancers, we found that both SRSF1 and SRSF2 crosslink extensively and preferentially to exons. Interestingly, both of these abundant and ubiquitous SR proteins show extensive overlap in their interactions with RNA, which is quite distinct from the reported binding profiles for another two SR proteins SRSF3 and SRSF4 (Ankö et al., 2010; Änkö et al., 2012). This raises an intriguing possibility that the SR family members might be composed of two classes: one whose members bind broadly to most exonic sequences and another whose members bind to more restricted sets of exons. It is also interesting that SRSF2 binds to both pre-mRNA and spliced mRNA. Because SRSF2 is a nonshuttling SR protein, this observation implies that it has to be removed or exchanged with other shuttling SR proteins prior to mRNA export, as noted earlier (Lin et al., 2005), which may constitute a step in the nucleus for the regulation of mRNA export.

SR proteins are believed to bind RNA with distinct sequence specificity. However, it has been difficult to obtain well-defined consensus motifs for each SR protein—information important for understanding splicing regulatory networks. For example, early SELEX studies based on in vitro binding or in vitro splicing function provide binding specificities for SR proteins that appear quite degenerate, although some general trends can be deduced (Cavaloc et al., 1999; Liu et al., 2000, 1998; Tacke and Manley, 1995). A degenerate binding mode for SRSF2 has recently been reinforced by the proposed consensus of the core SSNG motif from structural analysis (Daubner et al., 2012). Our in vivo mapping data largely confirmed the proposed binding specificity for the two SR proteins, suggesting that SRSF1 generally prefers GA-rich sequences while SRSF2 recognizes a GC-rich core. A possible selective advantage for having multiple SR proteins function through degenerate recognition sequences may be to ensure efficient inclusion (1) while still accommodating evolutionary variations in exon sequence due to protein coding and (2) as levels of individual SR proteins vary in different cell types or during development.

RNA Binding of SR Proteins Compensates for Weak Splice Sites and Intron Length

Another proposed function of SR proteins is to compensate for weak splice sites in mammalian genomes. Interestingly, the majority of exons in the human genome are short and introns are long (Sakharkar et al., 2004). While we found that SR protein binding (as measured by CLIP peak strength) is inversely correlated with the strength of nearby splice sites, a more striking feature of our data is the correlation between SR protein binding and intron length. The effect of intron length has not been well studied in vitro because substrates with short introns are easier to prepare, yet authentic introns vary considerably with many up to 100 kb in length. As intron size increases, the ability of the cell to identify authentic exons from an increasing mass of intron sequence may demand more efficient SR protein binding.

Genome-wide comparison of constitutive and alternative exons has revealed that alternative exons contain fewer ESEs, which, coupled with weaker splice sites, render their inclusion inefficient (Fairbrother et al., 2002; Itoh et al., 2004; Wang et al., 2005; Zhang and Chasin, 2004). This is consistent with our finding that the major SR proteins bind more weakly on alternative exons than flanking constitutive exons. This also agrees with the idea that alternative exons are less efficiently recognized by general splicing factors, including the ubiquitous SR proteins, so that they can become included only upon the appearance of appropriate cell-type-specific or developmentally regulated splicing factors in the cell. Our results reinforce the idea that exon recognition involves a complex interplay of splice site strength, intron length, and binding of splicing enhancing factors such as SR proteins.

SR Proteins Create Appropriate Exon Usage by Influencing Both Exon Inclusion and Skipping In Vivo

Perhaps one of the most striking findings here is the involvement of SR proteins in both exon-inclusion and -skipping events, which contradicts the common assumption that SR proteins mainly function in promoting exon inclusion. Because SRSF1 and SRSF2 often bind to the same exons, we were further surprised by the apparently specialized functions of SR proteins in including discrete sets of exons in the mouse transcriptome. The two major SR proteins exhibit nonredundant functions on some exons and jointly regulate a smaller fraction of events.

A similar observation was also made recently with the hnRNP family of splicing regulators (Huelga et al., 2012). Again, contrary to the expectation that hnRNP proteins antagonize SR proteins to cause exon skipping, many hnRNP proteins appear to function in promoting both exon inclusion and skipping in vivo at different exons. These observations indicate the need to revise the idea that SR proteins are predominantly positive splicing regulators while hnRNP proteins are predominantly negative splicing regulators, given their in vivo effects on both exon

inclusion and skipping as well as our evidence that loss of one SR protein remodels the binding profile of another.

Coordinated Action of SR Proteins Is Responsible for Complex Splicing Outcomes

We initially considered the complex effects of SR proteins on regulated splicing to reflect their position-dependent activities, which have been demonstrated on model genes (Han et al., 2011a; Sanford et al., 2009). However, the lack of the anticipated binding site distributions on competing exons from the composite RNA map generated for both SRSF1 and SRSF2 challenges this interpretation. We then reasoned that this simple idea might be confounded by compensatory responses of other SR proteins and splicing regulators upon depletion of an individual member of the family. To test this idea in principle, we profiled SRSF1 binding before and after knocking down SRSF2 and found that depletion of SRSF2 caused both gain and loss of SRSF1 binding at many exons. This means that in order to understand SR protein function, we must consider that specific binding events by one SR protein are in competition with binding of other SR proteins and splicing regulators nearby. We further showed by mutational analysis that prevention of such compensatory binding is sufficient to alter the splicing response from SR protein depletion-induced exon inclusion to exon skipping. Therefore, the complex response observed in cells depleted of a specific SR protein likely results from the loss of function of that SR protein in combination with coordinated alteration of binding (and activity in splicing) of other splicing regulators at many exons near the binding sites of the missing protein. These findings provide a conceptual framework to understand complex effects of SR proteins, and perhaps other splicing regulators, in vivo.

EXPERIMENTAL PROCEDURES

Cell Culture, RNA Extraction, and Splicing Profiling

The SRSF1 and SRSF2 Tet-repressible MEFs were maintained in Dulbecco's modified Eagle's medium (DMEM) plus 10% fetal bovine serum (FBS) without tetracycline. SRSF1 or SRSF2 protein was depleted by growing the cells for the indicated period of time in the presence of 10 μl/ml of Dox. RNA isolation, RT-PCR, and splicing profiling were carried out according to established standard protocols, which are detailed in the Supplemental Experimental Procedures.

Plasmid Construction

CDC45I exons 3–5 and *CCNL* exons 3–5 were PCR amplified and cloned into pcDNA 3.1(+) at the BamH1 and HindIII sites. The mutant plasmids were generated using deletion primers adjacent to the binding site. SRSF2 MEF cells were transfected with the indicated plasmids using Lipofectamine 2000 (Life Technologies). At 6 hr after incubating Lipofectamine-plasmid complexes with the cells, the media was changed to serum either containing Dox or without it (to induce SRSF2 depletion). After 48 hr, the cells were harvested for RNA and protein analyses.

Analysis of Genome-wide Data

CLIP-seq was carried out as previously described (Xue et al., 2009; Yeo et al., 2009) using anti-HA antibody (Abcam ab9110). For surveying SR protein binding in the absence of another SR protein, we treated SRSF1-HA MEFs with siRNA against SRSF2 (Dharmacon ON-TARGETplus J-044306-05). All analysis of CLIP-seq data was done using custom Python and R scripts, BEDTools (Quinlan and Hall, 2010), and the Kent source package (Kent et al., 2002) as detailed in the Supplemental Experimental Procedures, which also contains the procedures for correlative analysis of SR protein binding and splicing signals, motif analysis, construction of RNA maps, and in vivo RNA interactions of one SR protein in the absence of another SR protein.

ACCESSION NUMBERS

The Gene Expression Omnibus number for the microarray data for SRSF1 and SRSF2 reported in this paper is GSE44583.

The Gene Expression Omnibus number for the CLIP-seq data for SRSF1 and SRSF2 reported in this paper is GSE44591.

SUPPLEMENTAL INFORMATION

Supplemental Information includes three figures, one table, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2013.03.001.

ACKNOWLEDGMENTS

We thank Dr. Chris Brenner for advice on motif analysis and members of the Fu lab for ideas and extensive discussion over the course of this investigation. This work was supported by NIH grants GM049369 and HG004659, the China 973 program grant 2011CB811300, and the China 111 program grant B06018 to X.-D.F. as well as NIH grant GM084317 to G.W.Y. and M.A.

Received: October 10, 2012
Revised: January 2, 2013
Accepted: February 25, 2013
Published: April 4, 2013

REFERENCES

Ankö, M.L., Morales, L., Henry, I., Beyer, A., and Neugebauer, K.M. (2010). Global analysis reveals SRp20- and SRp75-specific mRNPs in cycling and neural cells. Nat. Struct. Mol. Biol. *17*, 962–970.

Änkö, M.L., Müller-McNicoll, M., Brandl, H., Curk, T., Gorup, C., Henry, I., Ule, J., and Neugebauer, K.M. (2012). The RNA-binding landscapes of two SR proteins reveal unique functions and binding to diverse RNA classes. Genome Biol. *13*, R17.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. *37*(Web Server issue), W202–W208.

Buratti, E., Stuani, C., De Prato, G., and Baralle, F.E. (2007). SR protein-mediated inhibition of CFTR exon 9 inclusion: molecular characterization of the intronic splicing silencer. Nucleic Acids Res. *35*, 4359–4368.

Cáceres, J.F., Screaton, G.R., and Krainer, A.R. (1998). A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. Genes Dev. *12*, 55–66.

Cartegni, L., Wang, J., Zhu, Z., Zhang, M.Q., and Krainer, A.R. (2003). ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res. *31*, 3568–3571.

Cavaloc, Y., Bourgeois, C.F., Kister, L., and Stévenin, J. (1999). The splicing factors 9G8 and SRp20 transactivate splicing through different and specific enhancers. RNA *5*, 468–483.

Daubner, G.M., Cléry, A., Jayne, S., Stevenin, J., and Allain, F.H. (2012). A syn-anti conformational difference allows SRSF2 to recognize guanines and cytosines equally well. EMBO J. *31*, 162–174.

Du, H., Cline, M.S., Osborne, R.J., Tuttle, D.L., Clark, T.A., Donohue, J.P., Hall, M.P., Shiue, L., Swanson, M.S., Thornton, C.A., and Ares, M., Jr. (2010). Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. Nat. Struct. Mol. Biol. *17*, 187–193.

Erkelenz, S., Mueller, W.F., Evans, M.S., Busch, A., Schöneweis, K., Hertel, K.J., and Schaal, H. (2013). Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms. RNA

19, 96–102. Published online November 21, 2012. http://dx.doi.org/10.1261/rna.037044.112.

Fairbrother, W.G., Yeh, R.F., Sharp, P.A., and Burge, C.B. (2002). Predictive identification of exonic splicing enhancers in human genes. Science *297*, 1007–1013.

Fox-Walsh, K.L., Dou, Y., Lam, B.J., Hung, S.P., Baldi, P.F., and Hertel, K.J. (2005). The architecture of pre-mRNAs affects mechanisms of splice-site pairing. Proc. Natl. Acad. Sci. USA *102*, 16176–16181.

Gabut, M., Dejardin, J., Tazi, J., and Soret, J. (2007). The SR family proteins B52 and dASF/SF2 modulate development of the Drosophila visual system by regulating specific RNA targets. Mol. Cell. Biol. *27*, 3087–3097.

Gallego, M.E., Gattoni, R., Stévenin, J., Marie, J., and Expert-Bezançon, A. (1997). The SR splicing factors ASF/SF2 and SC35 have antagonistic effects on intronic enhancer-dependent splicing of the beta-tropomyosin alternative exon 6A. EMBO J. *16*, 1772–1784.

Ghigna, C., Giordano, S., Shen, H., Benvenuto, F., Castiglioni, F., Comoglio, P.M., Green, M.R., Riva, S., and Biamonti, G. (2005). Cell motility is controlled by SF2/ASF through alternative splicing of the Ron protooncogene. Mol. Cell *20*, 881–890.

Graveley, B.R., and Maniatis, T. (1998). Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. Mol. Cell *1*, 765–771.

Han, J., Ding, J.H., Byeon, C.W., Kim, J.H., Hertel, K.J., Jeong, S., and Fu, X.D. (2011a). SR proteins induce alternative exon skipping through their activities on the flanking constitutive exons. Mol. Cell. Biol. *31*, 793–802.

Han, J., Xiong, J., Wang, D., and Fu, X.D. (2011b). Pre-mRNA splicing: where and when in the nucleus. Trends Cell Biol. *21*, 336–343.

Havlioglu, N., Wang, J., Fushimi, K., Vibranovski, M.D., Kan, Z., Gish, W., Fedorov, A., Long, M., and Wu, J.Y. (2007). An intronic signal for alternative splicing in the human genome. PLoS ONE *2*, e1246.

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.

Hertel, K.J., and Graveley, B.R. (2005). RS domains contact the pre-mRNA throughout spliceosome assembly. Trends Biochem. Sci. *30*, 115–118.

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.

Itoh, H., Washio, T., and Tomita, M. (2004). Computational comparative analyses of alternative splicing regulation using full-length cDNA of various eukaryotes. RNA *10*, 1005–1018.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res. *12*, 996–1006.

Lemaire, R., Winne, A., Sarkissian, M., and Lafyatis, R. (1999). SF2 and SRp55 regulation of CD45 exon 4 skipping during T cell activation. Eur. J. Immunol. *29*, 823–837.

Li, H., Qiu, J., and Fu, X.D. (2012). RASL-seq for massively parallel and quantitative analysis of gene expression. Curr. Protoc. Mol. Biol. *98*, 4.13.1–4.13.9.

Lim, K.H., Ferraris, L., Filloux, M.E., Raphael, B.J., and Fairbrother, W.G. (2011). Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes. Proc. Natl. Acad. Sci. USA *108*, 11093–11098.

Lin, S., and Fu, X.D. (2007). SR proteins and related factors in alternative splicing. Adv. Exp. Med. Biol. *623*, 107–122.

Lin, S., Xiao, R., Sun, P., Xu, X., and Fu, X.D. (2005). Dephosphorylation-dependent sorting of SR splicing factors during mRNP maturation. Mol. Cell *20*, 413–425.

Liu, H.X., Zhang, M., and Krainer, A.R. (1998). Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. Genes Dev. *12*, 1998–2012.

Liu, H.X., Chew, S.L., Cartegni, L., Zhang, M.Q., and Krainer, A.R. (2000). Exonic splicing enhancer motif recognized by human SC35 under splicing conditions. Mol. Cell. Biol. *20*, 1063–1071.

Long, J.C., and Caceres, J.F. (2009). The SR protein family of splicing factors: master regulators of gene expression. Biochem. J. *417*, 15–27.

Mayeda, A., and Krainer, A.R. (1992). Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell *68*, 365–375.

Mayeda, A., Munroe, S.H., Cáceres, J.F., and Krainer, A.R. (1994). Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. EMBO J. *13*, 5483–5495.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Sakharkar, M.K., Chow, V.T., and Kangueane, P. (2004). Distributions of exons and introns in the human genome. In Silico Biol. (Gedrukt) *4*, 387–393.

Sanford, J.R., Coutinho, P., Hackett, J.A., Wang, X., Ranahan, W., and Caceres, J.F. (2008). Identification of nuclear and cytoplasmic mRNA targets for the shuttling protein SF2/ASF. PLoS ONE *3*, e3369.

Sanford, J.R., Wang, X., Mort, M., Vanduyn, N., Cooper, D.N., Mooney, S.D., Edenberg, H.J., and Liu, Y. (2009). Splicing factor SFRS1 recognizes a functionally diverse landscape of RNA transcripts. Genome Res. *19*, 381–394.

Schaal, T.D., and Maniatis, T. (1999). Selection and characterization of pre-mRNA splicing enhancers: identification of novel SR protein-specific enhancer sequences. Mol. Cell. Biol. *19*, 1705–1719.

Shen, H., and Green, M.R. (2006). RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans. Genes Dev. *20*, 1755–1765.

Shen, H., Kan, J.L., and Green, M.R. (2004). Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. Mol. Cell *13*, 367–376.

Solis, A.S., Peng, R., Crawford, J.B., Phillips, J.A., 3rd, and Patton, J.G. (2008). Growth hormone deficiency and splicing fidelity: two serine/arginine-rich proteins, ASF/SF2 and SC35, act antagonistically. J. Biol. Chem. *283*, 23619–23626.

Tacke, R., and Manley, J.L. (1995). The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. EMBO J. *14*, 3540–3551.

Ule, J., Jensen, K., Mele, A., and Darnell, R.B. (2005). CLIP: a method for identifying protein-RNA interaction sites in living cells. Methods *37*, 376–386.

Wang, J., Smith, P.J., Krainer, A.R., and Zhang, M.Q. (2005). Distribution of SR protein exonic splicing enhancer motifs in human protein-coding genes. Nucleic Acids Res. *33*, 5053–5062.

Witten, J.T., and Ule, J. (2011). Understanding splicing regulation through RNA splicing maps. Trends Genet. *27*, 89–97.

Wu, H., Sun, S., Tu, K., Gao, Y., Xie, B., Krainer, A.R., and Zhu, J. (2010). A splicing-independent function of SF2/ASF in microRNA processing. Mol. Cell *38*, 67–77.

Xiao, R., Tang, P., Yang, B., Huang, J., Zhou, Y., Shao, C., Li, H., Sun, H., Zhang, Y., and Fu, X.D. (2012). Nuclear matrix factor hnRNP U/SAF-A exerts a global control of alternative splicing by regulating U2 snRNP maturation. Mol. Cell *45*, 656–668.

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

Yeo, G., and Burge, C.B. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J. Comput. Biol. *11*, 377–394.

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, X.H., and Chasin, L.A. (2004). Computational definition of sequence motifs governing constitutive exon splicing. Genes Dev. *18*, 1241–1250.

Zhang, C., and Darnell, R.B. (2011). Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. Nat. Biotechnol. *29*, 607–614.

Zhong, X.Y., Wang, P., Han, J., Rosenfeld, M.G., and Fu, X.D. (2009). SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. Mol. Cell *35*, 1–10.

Zhou, Z., Qiu, J., Liu, W., Zhou, Y., Plocinik, R.M., Li, H., Hu, Q., Ghosh, G., Adams, J.A., Rosenfeld, M.G., and Fu, X.D. (2012). The Akt-SRPK-SR axis constitutes a major pathway in transducing EGF signaling to regulate alternative splicing in the nucleus. Mol. Cell *47*, 422–433.