Resource

Functional Splicing Network Reveals Extensive Regulatory Potential of the Core Spliceosomal Machinery

Panagiotis Papasaikas,[1,2,4] J. Ramón Tejedor,[1,2,4] Luisa Vigevani,[1,2] and Juan Valcárcel[1,2,3,*]
[1]Centre de Regulació Genòmica, Dr. Aiguader 88, 08003 Barcelona, Spain
[2]Universitat Pompeu Fabra, Dr. Aiguader 88, 08003 Barcelona, Spain
[3]Institució Catalana de Recerca i Estudis Avançats, Passeig Lluis Companys 23, 08010 Barcelona, Spain
[4]Co-first authors
*Correspondence: juan.valcarcel@crg.eu
http://dx.doi.org/10.1016/j.molcel.2014.10.030

SUMMARY

Pre-mRNA splicing relies on the poorly understood dynamic interplay between >150 protein components of the spliceosome. The steps at which splicing can be regulated remain largely unknown. We systematically analyzed the effect of knocking down the components of the splicing machinery on alternative splicing events relevant for cell proliferation and apoptosis and used this information to reconstruct a network of functional interactions. The network accurately captures known physical and functional associations and identifies new ones, revealing remarkable regulatory potential of core spliceosomal components, related to the order and duration of their recruitment during spliceosome assembly. In contrast with standard models of regulation at early steps of splice site recognition, factors involved in catalytic activation of the spliceosome display regulatory properties. The network also sheds light on the antagonism between hnRNP C and U2AF, and on targets of antitumor drugs, and can be widely used to identify mechanisms of splicing regulation.
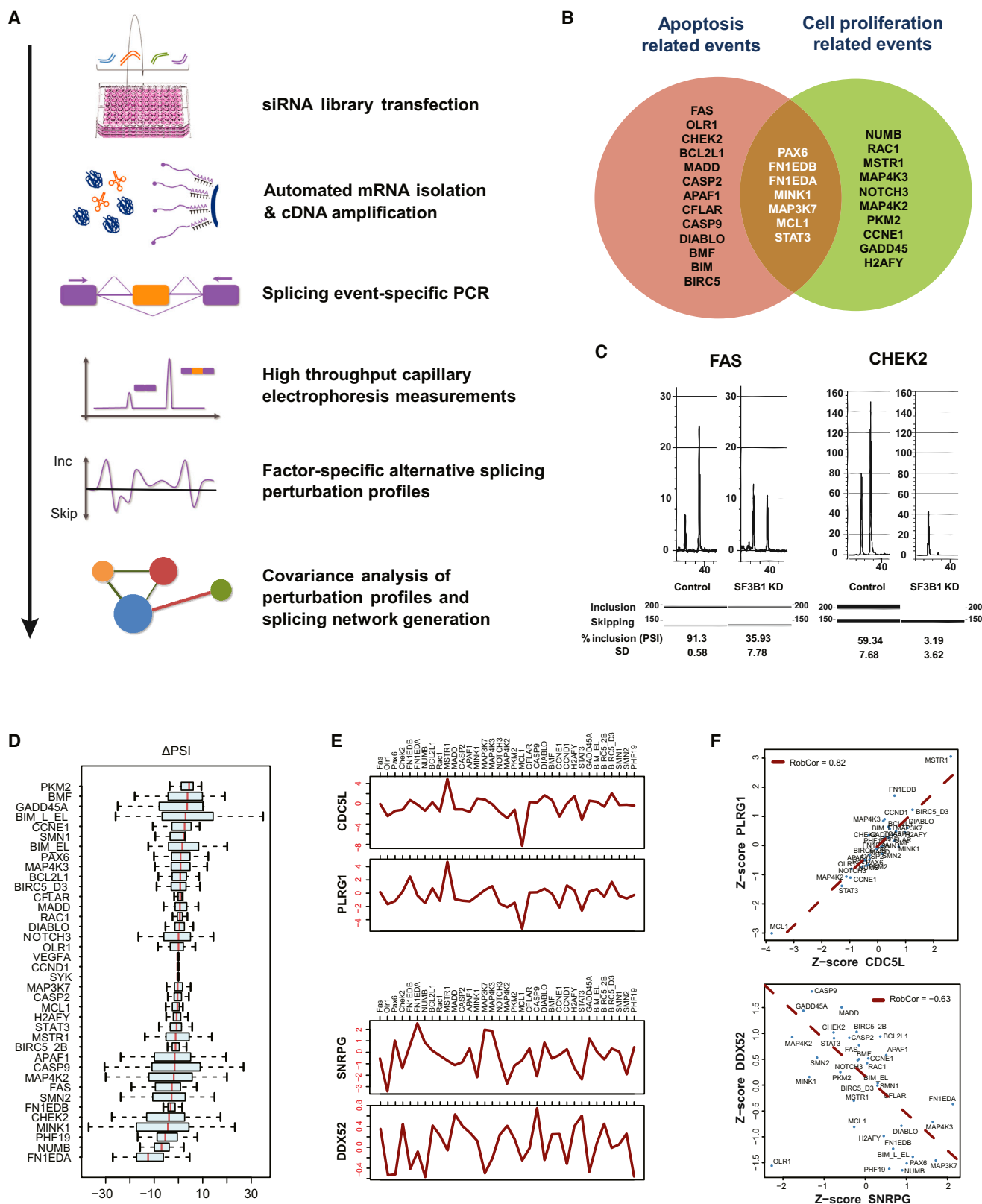
INTRODUCTION

Pre-mRNA splicing is carried out by the spliceosome, one of the most complex molecular machineries of the cell, composed of five small nuclear ribonucleoprotein particles (U1, U2, U4/5/6 snRNP) and about 150 additional polypeptides (reviewed by Wahl et al., 2009). Detailed biochemical studies using a small number of model introns have delineated a sequential pathway for the assembly of spliceosomal subcomplexes. For example, U1 snRNP recognizes the 5′ splice site, and U2AF—a heterodimer of 35 and 65 kDa subunits—recognizes sequences at the 3′ end of introns. U2AF binding helps to recruit U2 snRNP to the upstream branchpoint sequence, forming complex A. U2 snRNP binding involves interactions of pre-mRNA sequences with U2 snRNA as well as with U2 proteins (e.g., SF3B1). Subse-

quent binding of preassembled U4/5/6 tri-snRNP forms complex B, which after a series of conformational changes forms complexes Bact and C, concomitant with the activation of the two catalytic steps that generate splicing intermediates and products. Transition between spliceosomal subcomplexes involves profound dynamic changes in protein composition as well as extensive rearrangements of base-pairing interactions between snRNAs and between snRNAs and splice site sequences (Wahl et al., 2009). RNA structures contributed by base-pairing interactions between U2 and U6 snRNAs serve to coordinate metal ions critical for splicing catalysis (Fica et al., 2013), implying that the spliceosome is an RNA enzyme whose catalytic center is only established upon assembly of its individual components.

Differential selection of alternative splice sites in a pre-mRNA (alternative splicing, AS) is a prevalent mode of gene regulation in multicellular organisms, often subject to developmental regulation (Nilsen and Graveley, 2010) and frequently altered in disease (Cooper et al., 2009; Bonnal et al., 2012). Substantial efforts made to dissect mechanisms of AS regulation on a relatively small number of pre-mRNAs have provided a consensus picture in which protein factors recognizing cognate auxiliary sequences in the pre-mRNA promote or inhibit early events in spliceosome assembly (Fu and Ares, 2014). These regulatory factors include members of the hnRNP and SR protein families, which often display cooperative or antagonistic functions depending on the position of their binding sites relative to the regulated splice sites. Despite important progress (Barash et al., 2010; Zhang et al., 2010), the combinatorial nature of these contributions complicates the formulation of integrative models for AS regulation.

To systematically capture the contribution of splicing regulatory factors, including components of the core spliceosome (Clark et al., 2002; Park et al., 2004; Pleiss et al., 2007; Saltzman et al., 2011) to AS regulation, we set up a high-throughput screen to evaluate the effects of knocking down each individual splicing component or regulator on 36 functionally important AS events (ASEs) and developed a framework for data analysis and network modeling. Network-based approaches can provide rich representations for real-world systems that capture their organization more adequately than more traditional approaches or mere cataloguing of pairwise relationships. Numerous studies have utilized them as a platform for deriving comprehensive transcriptional, metabolic, and physical interaction maps in several

A

siRNA library transfection

Automated mRNA isolation & cDNA amplification

Splicing event-specific PCR

High throughput capillary electrophoresis measurements

Factor-specific alternative splicing perturbation profiles

Covariance analysis of perturbation profiles and splicing network generation

B

Apoptosis related events

FAS
OLR1
CHEK2
BCL2L1
MADD
CASP2
APAF1
CFLAR
CASP9
DIABLO
BMF
BIM
BIRC5

PAX6
FN1EDB
FN1EDA
MINK1
MAP3K7
MCL1
STAT3

Cell proliferation related events

NUMB
RAC1
MSTR1
MAP4K3
NOTCH3
MAP4K2
PKM2
CCNE1
GADD45
H2AFY

C

	FAS		CHEK2	
	Control	SF3B1 KD	Control	SF3B1 KD
Inclusion	200	200		
Skipping	150	150		
% inclusion (PSI)	91.3	35.93	59.34	3.19
SD	0.58	7.78	7.68	3.62

D

ΔPSI

E

F

(legend on next page)

organisms and contexts (e.g., Chatr-Aryamontri et al., 2013; Franceschini et al., 2013; Wong et al., 2012). Their application has also proven invaluable for deciphering regulatory relationships in transcriptional circuits, for studying their properties upon perturbation, and for connecting physiological responses and diseases to their molecular underpinnings (e.g., Kim et al., 2012; Watson et al., 2013; Yang et al., 2014).

Key to our approach is the premise that the distinct profiles of splicing changes caused by perturbation of regulatory factors depend on the functional bearings of those factors and can, therefore, be used as proxies for inferring functional relationships. We derive a network that recapitulates the topology of known splicing complexes and provides an extensive map of >500 functional associations among ~200 splicing factors (SFs). We use this network to identify general and particular mechanisms of AS regulation and to identify key SFs that mediate the effects on AS of perturbations induced by iron (see accompanying manuscript by Tejedor et al., 2014, in this issue of *Molecular Cell*) or by antitumor drugs targeting components of the splicing machinery. Our study offers a unique compendium of functional interactions among SFs, a discovery tool for coupling cell-perturbing stimuli to splicing regulation, and an expansive view of the splicing regulatory landscape and its organization.

RESULTS

Results from a genome-wide siRNA screen to identify regulators of Fas/CD95 AS revealed that knockdown of a significant fraction of core SFs (defined as proteins identified in highly purified spliceosomal complexes assembled on model, single-intron pre-mRNAs; Wahl et al., 2009) caused changes in Fas/CD95 exon 6 inclusion (Tejedor et al., 2014). The number of core factors involved and the extent of their regulatory effects exceeded those of classical splicing regulatory factors like SR proteins or hnRNPs. To systematically evaluate the contribution of different classes of splicing regulators to AS regulation, we set up a screen in which we assessed the effects of knockdown of each individual factor on 38 ASEs relevant for cell proliferation and/or apoptosis (Figures 1A and 1B). The ASEs were selected due to their clear impact on protein function and their documented biological relevance (see Table S1 available online). A library of 270 siRNA pools (Table S2) was used, corresponding to genes encoding core spliceosomal components as well as auxiliary regulatory SFs and factors involved in other RNA-processing steps, including RNA stability, export, or polyadenylation. In addition, 40 genes involved in chromatin structure mod-

ulation were included, to probe for possible functional links between chromatin and splicing regulation (Luco et al., 2011).

The siRNA pools were individually transfected in biological triplicates in HeLa cells using a robotized procedure (Supplemental Experimental Procedures). Seventy-two hours posttransfection, RNA was isolated, cDNA libraries generated using oligo-dT and random primers, and the patterns of splicing probed by PCR using primers flanking each of the alternatively spliced regions (Figure 1A). PCR products were resolved by high-throughput capillary electrophoresis (HTCE) and the ratio between isoforms calculated as the percent spliced in (PSI) index (Katz et al., 2010), which represents the percentage of exon inclusion/alternative splice site usage. Knockdowns of factors affecting one ASE (Fas/CD95) were also validated with siRNA pools from a second library and by a second independent technique (Illumina HiSeq sequencing; Tejedor et al., 2014). Conversely, the effects of knocking down core SFs (P14, SNRPG, and SF3B1) on the ASEs were robust when different siRNAs and RNA purification methods were used (Figure S1A). Several pieces of evidence indicate that the changes in AS are not due to the induction of cell death upon depletion of essential SFs. First, cell viability was not generally compromised after 72 hr of knockdown of 13 individual core SFs (Figure S1C). Second, cells detached from the plate were washed away before RNA isolation. Third, splicing changes upon induction of apoptosis with staurosporine were distinct from those observed upon knockdown of SFs (Figure S1D).

Overall the screen generated a total of 33,288 PCR data points. These measurements provide highly robust and sensitive estimates of the relative use of competing splice sites. Figure 1D shows the spread of the effects of knockdown for all the factors in the screen for the ASEs analyzed. Three events (VEGFA, SYK, and CCND1) were excluded from further analysis, because they were not affected in the majority of the knockdowns (median absolute ΔPSI <1, Figure 1D). From these data we generated, for every knockdown condition, a perturbation profile that reflects its impact across the 35 ASEs in terms of the magnitude and direction of the change. Figures 1E and 1F show examples of such profiles and the relationships between them: in one case, knockdown of CDC5L or PLRG1 generates very similar perturbation profiles (upper diagrams in Figures 1E and 1F), consistent with their well-known physical interactions within the PRP19 complex (Makarova et al., 2004). In contrast, the profiles associated with the knockdown of SNRPG and DDX52 are to a large extent opposite to each other (lower diagrams in Figures 1E and 1F), suggesting antagonistic functions. Perturbation profiles and relationships between factors were not significantly changed

Figure 1. Comprehensive Mapping of Functional Interactions between SFs in ASEs Implicated in Cell Proliferation and Apoptosis
(A) Flowchart of the pipeline for network generation. See main text for details.
(B) Genes with ASEs relevant for the regulation of cell proliferation and/or apoptosis used in this study.
(C) Examples of HTCE profiles for FAS/CD95 and CHEK2 ASEs under conditions of knockdown of the core SF SF3B1. Upper panels show the HTCE profiles, lower panels represent the relative intensities of the inclusion and skipping isoforms. PSI values indicate the median and standard deviation of three independent experiments.
(D) Spread of PSI changes observed for each of the events analyzed in this study upon the different knockdowns.
(E) Examples of splicing perturbation profiles for different knockdowns. The profiles represent change toward inclusion (>0) or skipping (<0) for each ASE upon knockdown of the indicated factors, quantified as a robust Z score (see Experimental Procedures).
(F) Robust correlation estimates and regression of the perturbation profiles of strongly correlated (PLRG1 versus CDC5L, upper panel) or anticorrelated (DDX52 versus SNRPG, lower panel) factors. See also Figures S1 and S2, Table S1, and Table S2.
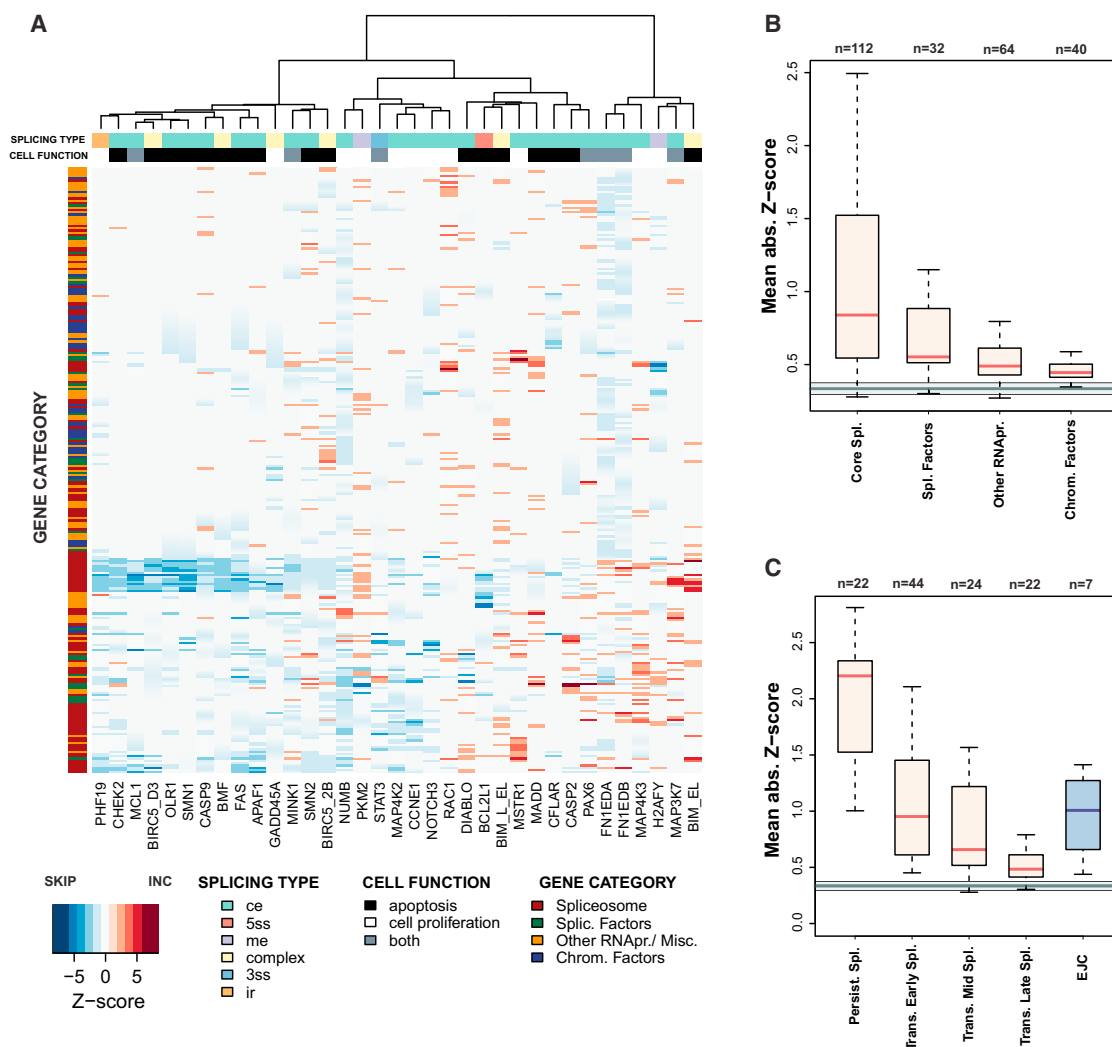
Figure 2. Coordinated Regulation of Splicing Events by Coherent Subsets of SFs

(A) Heatmap representation of the results of the screen. ASEs used in this study are on the x axis, while knockdown conditions on the y axis. Data are clustered on both dimensions (ward linkage, similarity measure based on Pearson correlation). Information about ASEs, (a) type (ce, cassette exon; 5ss, alternative 5′ splice site; me, mutually exclusive exon; complex, multiexonic rearrangement; 3ss, alternative 3′ splice site; ir, intron retention) and (b) involvement in apoptosis and/or cell proliferation regulation, as well as information on the category of genes knocked down, is color coded.

(B) Boxplot representation of the mean absolute Z score of AS changes induced by the knockdown of particular classes of factors, including core and noncore SFs, other RNA processing factors, and chromatin remodeling factors. Median and spread (interquartile range) of AS changes of mock siRNA conditions are represented as a green line and box, respectively.

(C) As in (B) for spliceosomal factors that assemble early and stay as detectable components through the spliceosome cycle; factors that are present only transiently at early, mid, or late stages of assembly; and components of the exon junction complex (EJC).

See also Figure S3, Table S2, Table S3, and Table S4.

when knockdown of a subset of factors was carried out for 48 hr (Figure S2A), arguing against AS changes being the consequence of secondary effects of SF knockdown in these cases (see also below the discussion on IK/SMU1).

Pervasive and Distinct Effects of Core Spliceosome Components on Alternative Splicing Regulation

The output of the screening process is summarized in the heatmap of Figure 2A. Three main conclusions can be derived from these results. First, a large fraction of the SF knockdowns have

noticeable but distinct effects on AS, indicating that knockdown of many components of the spliceosome causes switches in splice site selection, rather than a generalized inhibition of splicing of every intron. Second, a substantial fraction of AS changes correspond to higher levels of alternative exon inclusion, again arguing against simple effects of decreasing splicing activity upon knockdown of general SFs, which would typically favor skipping of alternative exons that often harbor weaker splice sites. These observations suggest extensive versatility of the effects of modulating the levels of SFs on splice site

selection. Third, despite the diversity of the effects of SF knockdowns on multiple ASEs, similarities can also be drawn (Figure 2A). For example, a number of ASEs relevant for the control of programmed cell death (e.g., Fas, APAF1, CASP9, or MCL1) cluster together and are similarly regulated by a common set of core spliceosome components. This observation suggests coordinated regulation of apoptosis by core factors, although the functional effects of these AS changes appear to be complex (Figure S3A). Another example is two distinct ASEs in the fibronectin gene (cassette exons EDA and EDB, of paramount importance to control aspects of development and cancer progression; Muro et al., 2003) which cluster close together, suggesting common mechanisms of regulation of the ASEs in this locus.

To explore the effects of different classes of SFs on AS regulation, we first grouped them in four categories: genes coding for core spliceosome components, noncore SFs/regulators, RNA-processing factors not directly related to splicing, and chromatin-related factors. Core SFs display the greater spread and average magnitude of effects, followed by noncore SFs (Figure 2B). Factors related with chromatin structure and remodeling showed a narrower range of milder effects (Figure 2B), although their values were clearly above the range of changes observed in mock knockdown samples. Of interest, knockdown of core factors leads to stronger exon skipping effects, while noncore SFs and other RNA processing and chromatin-related factors show more balanced effects (Figures S3B and S3C).

To further dissect the effects of core components, we subdivided them into four categories depending on the timing and duration of their recruitment to splicing complexes during spliceosome assembly (Wahl et al., 2009): (1) persistent components (e.g., 17S U2 snRNPs, Sm proteins) that enter the assembly process before or during complex A formation and remain until completion of the reaction, (2) transient early components that join the reaction prior to B complex activation but are scarce at later stages of the process, (3) transient middle components joining during B-act complex formation, and (4) transient late components that are only present during or after C complex formation. Our results indicate that knockdowns of factors that persist during spliceosome assembly cause AS changes of higher magnitude than those caused by knockdowns of transient factors involved in early spliceosomal complexes, which in turn are stronger and more dispersed than those caused by factors involved in middle complexes and in complex C formation or catalytic activation (Figure 2C). As observed for core factors in general, knockdown of persistent factors tends to favor skipping, while the effects of knockdown of more transient factors are more split between inclusion and skipping (Figures S3B and S3C).

Interestingly, knockdowns of components of the exon junction complex (EJC) display a range of effects that resembles those of transient early or mid splicing complexes (Figure 2C). These results are in line with recent reports documenting effects of EJC components in AS, in addition to their standard function in post-splicing processes (Ashton-Beaucage et al., 2010).

To test whether the knockdown of core factors causes a general inhibition of splicing, we measured the levels of introns (both in constitutive and alternatively spliced regions) relative to exons in four genes, upon the knockdown of four SFs (BRR2, SF3B1, SNRPG, and SLU7). The results shown in Figure 3 reveal a complex picture in which retention of certain introns is clearly increased and retention of other introns is very mild, and there is even one case (the two introns flanking the fibronectin EDI alternative exon) in which the low levels of intron detected under control conditions are reduced upon knockdown of core SFs (notably, this ASE typically displays a distinct pattern of AS upon knockdown of core SFs, Figure 2A). Therefore, rather than a general uniform inhibition of splicing, reduction in the levels of core SFs induces differential effects in different introns, a concept reminiscent of the differential effects observed for antitumor drugs targeting core components like SF3B1 (Bonnal et al., 2012).

Reconstruction of a Functional Network of Splicing Factors

To systematically and accurately map the functional relationships of SFs on AS regulation, we quantified the similarity between AS perturbation profiles for every pair of factors in our screen using a robust correlation estimate for the effects of the factors' knockdowns across the ASEs. This measure captures the congruence between the shapes of perturbation profiles, while it does not take into account proportional differences in the magnitude of the fluctuations. Crucially, it discriminates between biological and technical outliers and is resistant to distorting effects of the latter (Supplemental Experimental Procedures).

We next employed glasso, a regularization-based algorithm for graphical model selection (Friedman et al., 2008; Supplemental Experimental Procedures), to reconstruct a network from these correlation estimates (Figures 4A and S4). Glasso seeks a parsimonious network model for the observed correlations. This is achieved by specifying a regularization parameter, which serves as a penalty that can be tuned in order to control the number of inferred connections (network sparsity) and therefore the false discovery rate (FDR) of the final model. For a given regularization parameter, both the number of inferred connections and the FDR are a decreasing function of the number of ASEs assayed (Figure 4B). Random sampling with different subsets of the real or reshuffled versions of the data indicates that network reconstruction converges to ∼500 connections with a FDR <5% near 35 events (Figure 4B), implying that analysis of additional ASEs in the screening would have only small effects on network sparsity and accuracy.

The complete network of functional interactions among the screened factors is shown in Figure 4A. It is comprised of 196 nodes representing screened factors and 541 connections (average degree 5.5), of which 518 correspond to positive and 23 to negative functional associations. The topology of the resulting network has several key features. First, two classes of factors are easily distinguishable. The first class encompasses factors that form densely connected clusters comprised of core spliceosomal components (Figure 4A). Factors physically linked within U2 snRNP and factors functionally related with U2 snRNP activity, and with the transition from complex A to B, form a tight cluster (orange nodes in the inset of Figure 4A). An adjacent but distinct highly linked cluster corresponds mainly to factors physically associated within U5 snRNP or the U4/5/6
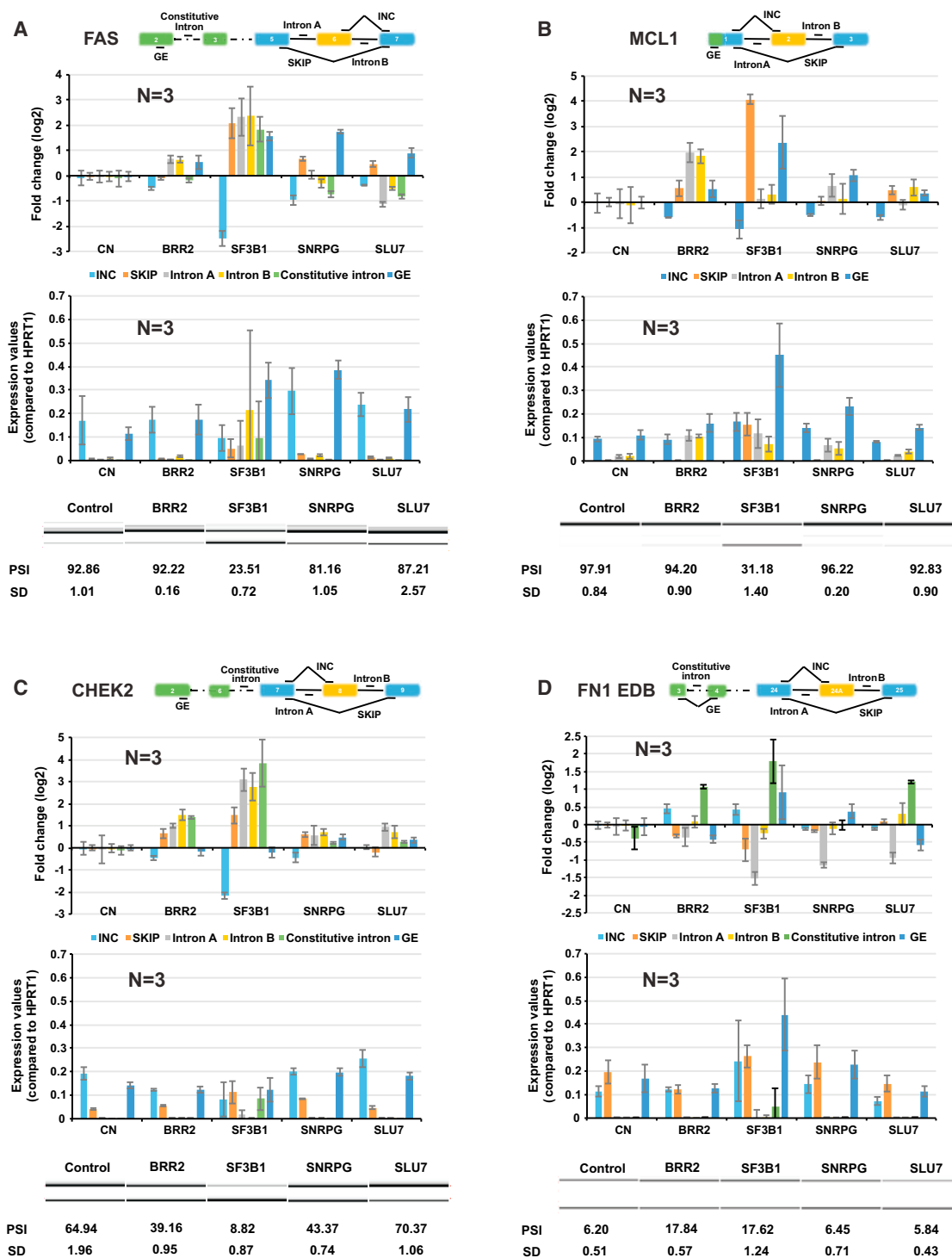
Figure 3. Variety of Effects of Knockdown of Core SFs on Intron Retention

(A–D) Real-time quantification of AS, intron retention, and gene expression for four genes included in the splicing network (FAS, A; MCL1, B; CHEK2, C; and FN1, D) upon knockdown of four core spliceosomal components (BRR2, SF3B1, SNRPG, and SLU7). In each of the panels, the graphs represent the following: top, scheme of genomic locations, AS patterns, and amplicons used; middle top, quantification of fold changes in expression of the different isoforms or introns (indicated by the color code) compared to mock siRNA conditions; middle bottom, relative levels of the different isoforms or introns (indicated by the color code) compared to HPRT1 housekeeping gene mRNAs; and bottom, AS changes measured by HTCE. PSI values are indicated. For all assays, values represent the mean, and error bars the standard deviation of three independent biological replicas.

tri-snRNP (blue nodes in Figure 4A, inset). Factors in these clusters and their mutual links encompass 23% of the network nodes but 63% of the total inferred functional associations (average degree 14.5). A second category includes factors that form a periphery of lower connectivity that occasionally projects to the core. This category includes many of the classical splicing regulators (e.g., SR proteins, hnRNPs) as well as several chromatin factors. An intermediate category—more densely interconnected but not reaching the density of interactions of the central modules—includes multiple members of the RNA-dependent DEAD/X box helicase family. While the number of connections (degree) is significantly higher for core spliceosomal factors in general (Figure 4C), this is especially clear for factors that represent persistent components of splicing complexes along the assembly pathway (Figures 4D and 4E). Persistent factors are particularly well connected to each other (50% of the possible functional connections are actually detected, Figure 4E). Factors that belong to consecutive complexes are significantly more linked to each other than factors that belong to early and late complexes, and factors that transiently assemble toward the final stages of spliceosome assembly display the least number of connections (Figures 4D and 4E).

Collectively, these results indicate that core SFs, particularly those that persist during assembly, display highly related functions in splice site recognition and fluctuations in their levels or activities have related consequences on the modulation of splice site choice. Of the observed connections, 20% can be attributed to known physical interactions and 82% of these are among core components (Figure 4A). Examples include tight links between the two subunits of U2AF, between the interacting partners PLRG1 and CDC5L or IK and SMU1. We estimate that about 50% of the functional links detected in our network could be predicted by previous knowledge of the composition or function of splicing complexes.

Of relevance, a substantial number of the functional associations closely recapitulate the composition and even—to some extent—the topology of known spliceosomal complexes (Figure S4; see also below and Discussion). These observations suggest that differences in the sensitivity of ASEs to SF perturbations reflect the specific role of these factors in the splicing process. Conversely, these results warrant the use of splicing perturbation profiles as proxies for inferring physical or functional links between factors in the splicing process.

Generality of the Functional Interactions in the AS Regulation Network

The above network was derived from variations in 35 ASEs selected because of their relevance for cell proliferation and apoptosis, but they represent a heterogeneous collection of AS types and, possibly, regulatory mechanisms. To identify functional connections persistent across AS types, we undertook a subsampling approach. We sampled subsets of 17 from the original 35 ASEs and iteratively reconstructed the network of functional associations from each subset (see Supplementary methods). Following 10,000 iterations, we asked how many functional connections were recovered in at least 90% of the sampled networks. The retrieved set of connections are shown in Figure 5A and they can be considered to represent a "basal"

or indispensable splicing circuitry that captures close similarities in function between SFs regardless of the subset of splicing substrates considered. The majority of these connections correspond to links between core components, with modules corresponding to major spliceosomal complexes. For example, the majority of components of U2 snRNP constitute the most prominent of the core modules (Figure 5A). Strikingly, the module is topologically subdivided in two submodules, the first corresponding to components of the SF3a and SF3b complexes and the second corresponding to proteins in the Sm complex. The connectivity between U2 snRNP components remarkably recapitulates known topological features of U2 snRNP organization, including the assembly of SF3a/b complexes in stem loops I and II and the assembly of the Sm ring in a different region of the U2 snRNA (Behrens et al., 1993) (Figures 5A and S4). That the connectivity between components derived from assessing the effects on AS of depletion of these factors correlates with topological features of the organization of the snRNP further argues that the functions in splice site selection are tightly linked to the structure and function of this particle, highlighting the potential resolution of our approach for identifying meaningful structural and mechanistic links.

A second prominent "core" module corresponds to factors known to play roles in conformational changes previous to/concomitant with catalysis in spliceosomes from yeast to human, including PRP8, PRP31 or U5-200. Once again some of the topological features of the module are compatible with physical associations between these factors (e.g., PRP8 with PRP31 or U5-116K or PRP31 with C20ORF14). That knockdown of these late-acting factors, which coordinate the final steps of the splicing process (Bottner et al., 2005; Häcker et al., 2008; Wahl et al., 2009), causes coherent changes in splice site selection strongly argues that the complex process of spliceosome assembly can be modulated at late steps, or even at the time of catalysis, to affect splice site choice. Other links in the "core" network recapitulate physical interactions, including the aforementioned modules involving the two subunits of the 3′ splice site-recognizing factor U2AF, the interacting partners PLRG1 and CDC5L and IK and SMU1.

To further evaluate the generality of our approach, we tested whether the functional concordance derived from our network analysis and from the iterative selection of a "core" network could be recapitulated in a different cell context and genome-wide. We focused our attention in the strong functional association between IK and SMU1, two factors that have been described as associated with complex B and its transition to Bact (Bessonov et al., 2008) (Figure 5A). Analysis of the effects of knockdown of these factors in the same sets of ASEs in HEK293 cells revealed a similar set of associations, despite the fact that individual events display different splicing ratios upon IK/SMU1 knockdown in this cell line compared to HeLa (Figures 5B and 5C). To explore the validity of this functional association in a larger set of splicing events, RNA was isolated in biological triplicates from HeLa cells transfected with siRNAs against these factors, and changes in AS were assessed using genome-wide splicing-sensitive microarrays. The results revealed a striking overlap between the effects of IK and SMU1 knockdown, both on gene expression changes and on

A

B

C

D
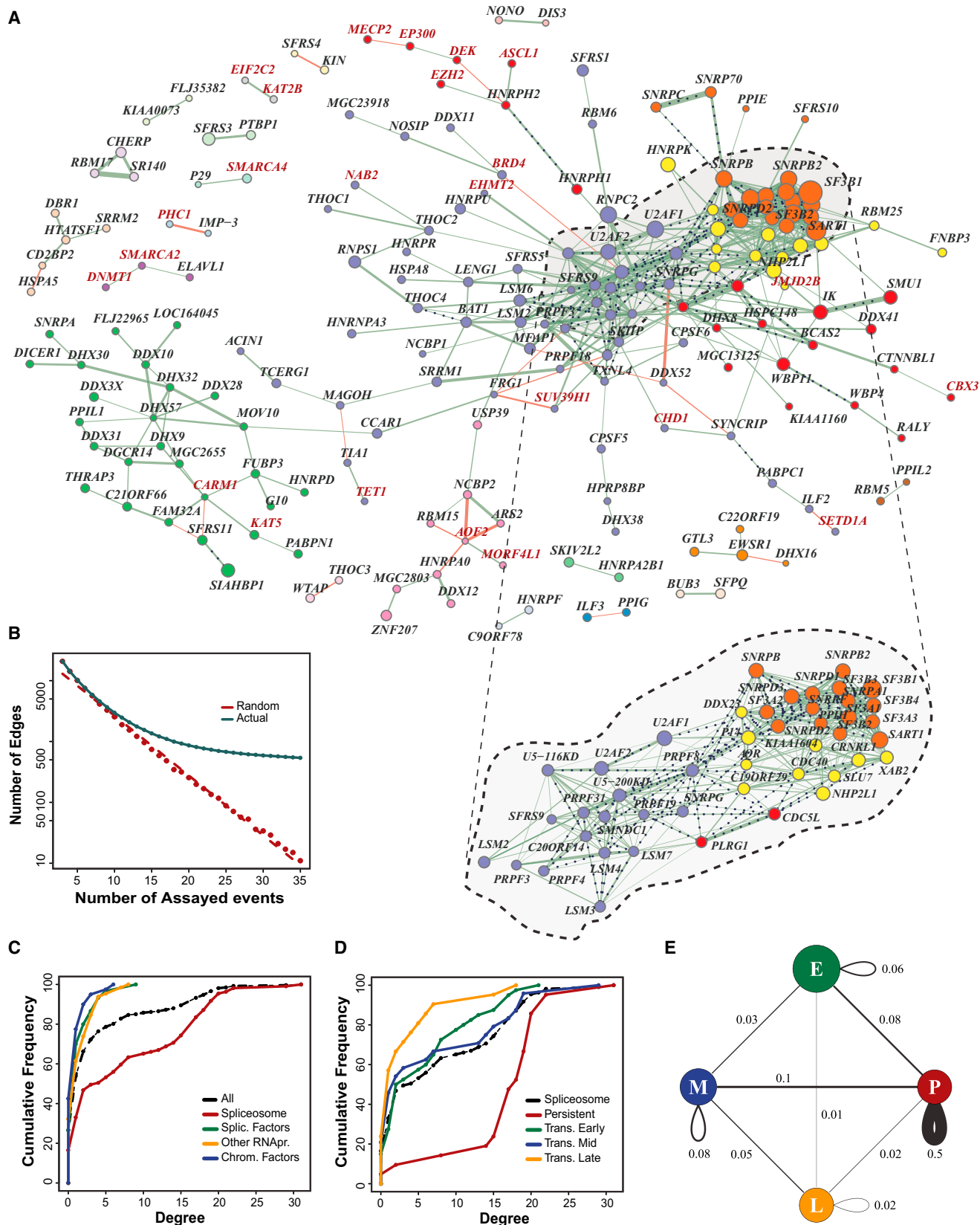
E

AS changes (61% overlap, p < 1E-20) covering a wide range of gene expression levels, fold differences in isoform ratios and AS types (Figures 5D–5F, S5B, and S5C). The overlap was much more limited with RBM6, a splicing regulator (Bechara et al., 2013) located elsewhere in the network. These results validate the strong functional similarities between IK and SMU1 detected in our analysis. The molecular basis for this link could be explained by the depletion of each protein (but not their mRNA) upon knockdown of the other factor, consistent with formation of a heterodimer of the two proteins and destabilization of one partner after depletion of the other (Figure S5A). Notably, the best correlation between the effects of these factors was between IK knockdown for 48 hr and SMU1 knockdown for 72 hr, perhaps reflecting a stronger functional effect of IK on AS regulation (Figures S2B and S2D).

Gene ontology analysis revealed a clear common enrichment of AS changes in genes involved in cell death and survival (Figure S5D), suggesting that IK/SMU1 could play a role in the control of programmed cell death through AS. To evaluate this possibility, we tested the effects of knockdown of these factors, individually or combined, on cell proliferation and apoptotic assays. The results indicated that their depletion activated cleavage of PARP, substantially reduced cell growth, and increased the fraction of cells undergoing apoptosis (Figures S5E–S5G). We conclude that functional links revealed by the network can be used to infer common mechanisms of splicing regulation and provide insights into the biological framework of the regulatory circuits involved.

Ancillary Functional Network Interactions Reveal Alternative Mechanisms of AS Regulation

The iterative generation of networks from subsets of ASEs explained above could, in addition to identifying general functional interactions, be used to capture functional links that are specific to particular classes of events characterized by distinct regulatory mechanisms. To explore this possibility, we sought the set of functional connections that are robustly recovered in a significant fraction of ASEs subsets but are absent in others (see Supplemental Experimental Procedures). Figure 6A shows the set of such identified interactions, which therefore represents specific functional links that only emerge when particular subsets of ASEs are considered. The inset represents graphically the results of principal component analysis (PCA) showing the variability in the strength of these interactions depending on

the presence or absence of different ASEs across the subsamples. The most discriminatory set of ASEs for separating these interactions is also shown (see also Figure S6). While this analysis may be constrained by the limited number of events analyzed and higher FDR due to multiple testing, it can provide a valuable tool for discovering regulatory mechanisms underlying specific classes of ASEs as illustrated by the example below.

A synergistic link was identified between hnRNP C and U2AF1 (the 35 kDa subunit of U2AF, which recognizes the AG dinucleotide at 3′ splice sites) for a subset of ASEs (Figures 6A and S6A). We explored a possible relationship between the effects of knocking down these factors and the presence of consensus hnRNP C binding sites (characterized by stretches of uridine residues) or 3′ splice site-like motifs in the vicinity of the regulated exons. The results revealed a significant correlation between the two factors (0.795), associated with the presence of composite elements comprised of putative hnRNP C binding sites within sequences conforming to a 3′ splice site motif (see Experimental Procedures) upstream and/or downstream of the regulated exons (Figures 6B and S6B). No such correlation was observed for ASEs lacking these sequence features (Figures 6C and S6C). These observations are in line with results from a recent report revealing that hnRNP C prevents exonization of Alu elements by competing the binding of U2AF to Alu sequences (Zarnack et al., 2013). Our results extend these findings by capturing functional relationships between hnRNP C and U2AF in ASEs that harbor binding sites for these factors in the vicinity of the regulated splice sites, both inside of Alu elements or independent of them (Figures 6B and 6D). Our data are compatible with a model in which uridine-rich and 3′ splice site-like sequences sequester U2AF away from the regulated splice sites, causing alternative exon skipping, while hnRNP C displaces U2AF away from these decoy sites, facilitating exon inclusion. In this model, depletion of U2AF and hnRNP C are expected to display the same effects in these ASEs, as observed in our data (Figures 6B and 6D). In contrast, in ASEs in which hnRNP C binding sites are located within the polypyrimidine tract of the 3′ splice site of the regulated exon, the two factors display antagonistic effects (Figures 6C and 6D), as expected from direct competition between them at a functional splice site. This example illustrates the potential of the network approach to identify molecular mechanisms of regulation on the basis of specific sequence features and the interplay between cognate factors.

Figure 4. Functional Splicing Regulatory Network

(A) Graphical representation of the reconstructed splicing network. Nodes (circles) correspond to individual factors and edges (lines) to inferred functional associations. Positive or negative functional correlations are represented by green or red edges, respectively. Edge thickness signifies the strength of the functional interaction, while node size is proportional to the overall impact (median *Z* score) of a given knockdown in the regulation of AS. Node coloring depicts the network's natural separation in coherent modules (see Supplemental Experimental Procedures). Known physical interactions as reported in the STRING database (Franceschini et al., 2013) are represented in black dotted lines between factors. The inset is an expanded view of factors in the U2 and U4/5/6 snRNP complexes.

(B) Number of recovered network edges using actual or randomized data sets (red and green points respectively) as a function of the number of ASEs used. Lines represent the best fit curves for the points.

(C and D) Network degree distribution. The plots represent the cumulative frequency of the number of connections (degree) for different classes of factors (as in Figures 2B and 2C).

(E) Functional crosstalk between different categories of spliceosome components. Values represent the fraction of observed functional connections out of the total possible connections among factors that belong to the different categories: persistent, P; transient early, E; transient mid, M; or transient late, L factors. See also Figure S4E and Table S5.
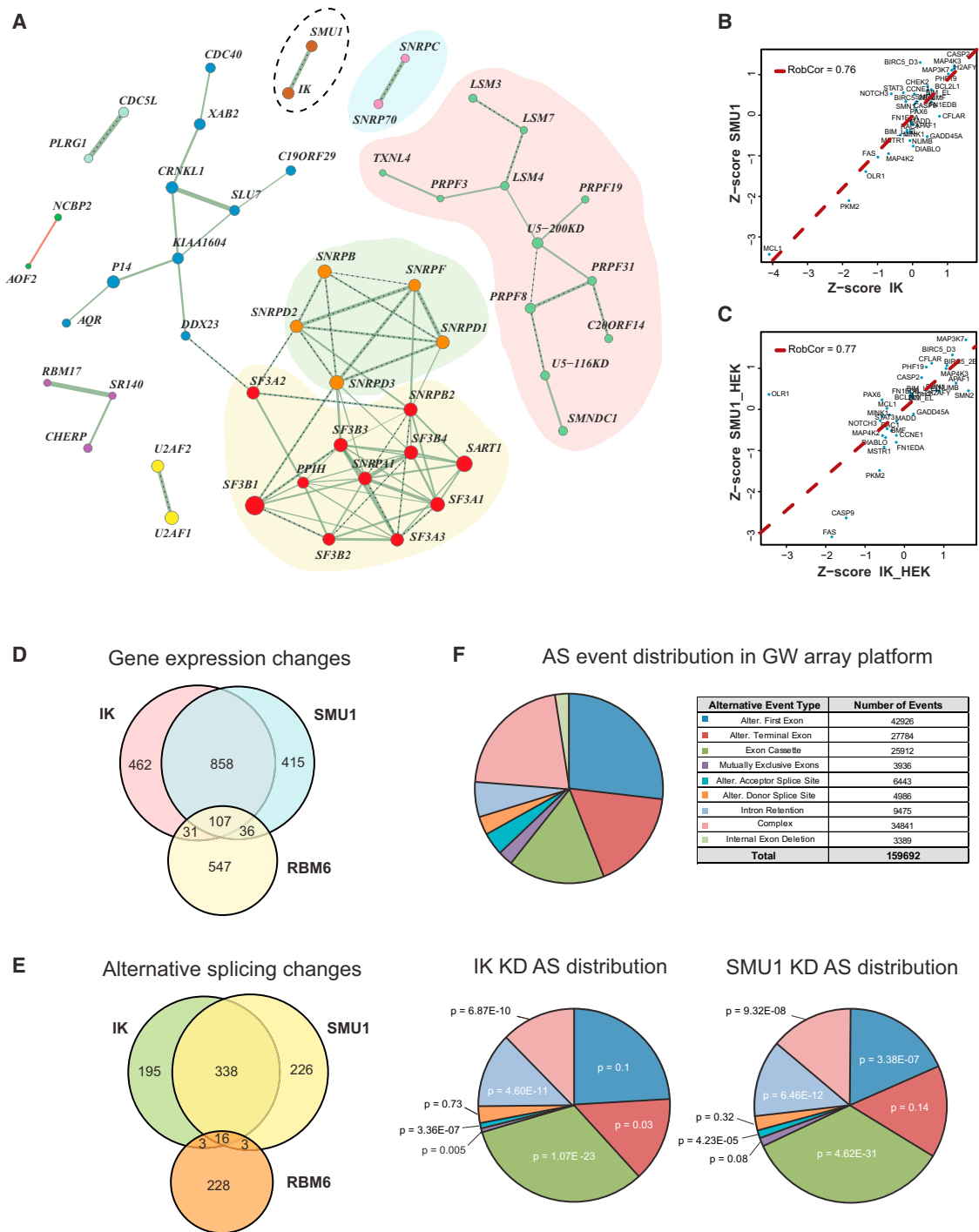
Figure 5. Core Network Involved in AS Regulation

(A) Graphical representation of the functional connections present in at least 90% of 10,000 networks generated by iterative selection of subsets of 17 out of the 35 ASEs used to generate the complete network. Network attributes represented as in Figure 4A. Known spliceosome complexes are highlighted by shadowed areas (U1 snRNP, blue; U2 snRNP, yellow; SM proteins, green; and tri-snRNP proteins, red).

(B and C) Consistency of inferred functional interactions in different cell lines. Knockdown of IK or SMU1 was carried out in parallel in HeLa (B) or HEK293 cells (C), and changes for the 35 ASEs were analyzed by RT-PCR and HTCE. Robust correlation estimates and regression for the AS changes observed in IK versus SMU1 knockdowns are shown for each of the cell lines.

(D) Venn diagrams of the overlap between the number of gene expression changes upon IK, SMU1, or RBM6 knockdowns.

(legend continued on next page)

Mapping Drug Targets within the Spliceosome

Another application of our network analysis is to identify possible targets within the splicing machinery of physiological or external (e.g., pharmacological) perturbations that produce changes in AS that can be measured using the same experimental setting. The similarity between profiles of AS changes induced by such perturbations and the knockdown of particular factors can help to uncover SFs that mediate their effects. To evaluate the potential of this approach, cells were treated with drugs known to affect AS decisions and the patterns of changes in the 35 ASEs induced by these treatments were assessed by the same robotized procedure for RNA isolation and RT-PCR analysis used to locate their position within the network.

The results indicate that structurally similar drugs like Spliceostatin A (SSA) and Meayamycin cause AS changes that closely resemble the effects of knocking down components of U2 snRNP (Figure 7A), including close links between these drugs and SF3B1, a known physical target of SSA (Kaida et al., 2007; Hasegawa et al., 2011) previously implicated in mediating the effects of the drug through alterations in the AS of cell cycle genes (Corrionero et al., 2011). Of interest, changes in AS of MCL1 (leading to the production of the proapoptotic mRNA isoform) appear as prominent effects of both drugs (Figure 7B), confirming and extending recent observations obtained using Meayamycin (Gao et al., 2013) and suggesting that this ASE can play a key general role in mediating the antiproliferative effects of drugs that target core splicing components like SF3B1 (Bonnal et al., 2012). A strong link is also captured between each of the drugs and PPIH (Figure 7A), a peptidyl-prolyl *cis-trans* isomerase which has been found to be associated with U4/5/6 tri-snRNP (Horowitz et al., 1997; Teigelkamp et al., 1998) but had not been implicated directly in 3′ splice site regulation. Indeed, the network indicates a close relationship between PPIH and multiple U2 snRNP components, particularly in the SF3a and SF3b complexes, further arguing that PPIH plays a role in 3′ splice site definition closely linked to branchpoint recognition by U2 snRNP. Treatment of the cells with the Clk (Cdc2-like) kinase family inhibitor TG003, which is also known to modulate AS (Muraki et al., 2004) and causes changes in ASEs analyzed in our network (Figure 7B), led to links with a completely different set of SFs (Figure 7A), attesting to the specificity of the results. These data confirm the potential of the network analysis to identify bona fide SF targets of physiological or pharmacological perturbations and further our understanding of the the underlying molecular mechanisms of regulation.

DISCUSSION

The combined experimental and computational approach described in this manuscript provides a rich source of information and a powerful toolset for studying the splicing process and its regulation. Similar approaches could be useful to systematically capture information about other aspects of transcrip-

tional or posttranscriptional gene regulation. The data provide comprehensive information about SFs (and some additional chromatin factors) relevant to understand the regulation of 35 ASEs important for cell proliferation and apoptosis. Second, the network analysis reveals functional links between SFs, which in some cases are based upon physical interactions. The reconstruction of the known topology of some complexes (e.g., U2 snRNP) by the network suggests that the approach captures important aspects of the operation of these particles and therefore has the potential to provide novel insights into the composition and intricate workings of multiple Spliceosomal subcomplexes. Third, the network can serve as a resource for exploring mechanisms of AS regulation induced by perturbations of the system, including genetic manipulations, internal or external stimuli, or exposure to drugs. As an example, the network was used in the accompanying manuscript by Tejedor et al. (2014) to identify a link between AS changes induced by variations in intracellular iron and the function of SRSF7, which can be explained in terms of iron-induced changes in the RNA binding and splicing regulatory activity of this zinc-knuckle-containing SR protein. Other examples include (1) the identification of the extensive regulatory overlap between the interacting proteins IK/RED and SMU1, relevant for the control of programmed cell death; (2) evidence for a general role of uridine-rich hnRNP C binding sequences (some associated with transposable elements; Zarnack et al., 2013) in the regulation of nearby alternative exons via antagonism with U2AF; and (3) delineation of U2 snRNP components and other factors, including PPIH, that are functionally linked to the effects of antitumor drugs on AS.

Our data reveal changes of AS regulation mediated either by core splicing components or by classical regulatory factors like proteins of the SR and hnRNP families. While the central core of functional links is likely to operate globally, alternative configurations (particularly in the periphery of the network) are likely to emerge if similar approaches are applied to other cell types, genes, or biological contexts. We report considerable versatility in the effects of core spliceosome factors on alternative spicing. Precedents exist for a role of core SFs on the regulation of splicing efficiency or splice site selection in yeast, *Drosophila*, and mammalian cells (Clark et al., 2002; Park et al., 2004; Pleiss et al., 2007; Saltzman et al., 2011). Despite their assumed general roles in the splicing process, depletion or mutation of particular core factors led to differential splicing effects in these studies, and at least in some cases the differential effects could be attributed to particular features of the regulated pre-mRNAs. For example, Clark et al. (2002) found that the core components Prp17p and Prp18p are dispensable for splicing of yeast introns with short branchpoint to 3′ splice site distances. Pleiss et al. (2007) found that introns of yeast ribosome protein genes are particularly sensitive to mutation of various core factors, including two DEAD/X box family of RNA helicases (PRP2 and PRP5) involved in spliceosome conformational transitions and PRP8, a highly conserved protein involved in catalytic activation

---

(E) Venn diagrams of the overlap between the number of AS changes upon knockdown of IK, SMU1, or RBM6.

(F) Distribution of ASEs in the Affymetrix array platform (upper panel) and distribution of the AS changes observed upon IK (lower left) or SMU1 (lower right) knockdown. p values correspond to the difference between the distribution of AS changes upon knockdown and the distribution of ASEs in the array. See also Figure S5.

Figure 6. Variable Functional Interactions Extracted from Networks Generated from Subsets of ASEs

(A) Network of ancillary functional connections characteristic of subsets of ASEs. Lines link SFs that display functional similarities in their effects on subsets of ASEs. Line colors indicate the positions of the interactions in the PCA shown in the inset, which captures the different relative contributions of the ASEs in defining these connections (see also Figure S6 and Experimental Procedures). For clarity, only the top ten discriminatory ASEs are shown.

(B and C) Correlation values and regression for the effects of U2AF1 versus HNRPC knockdowns for ASEs in the presence (B) or absence (C) of upstream composite HNRPC binding/3′ splice site-like elements.

(legend continued on next page)

(Collins and Guthrie, 1999; Mozaffari-Jovin et al., 2012). Distinct effects of the mutations on different ribosomal genes were also observed, arguing again for specific effects on splicing of certain genes. DEAD/H box proteins as well as components of U1, U2, and U4/6 snRNP were found to modulate particular ASEs in an RNAi screen of RNA binding proteins in *Drosophila* cells (Park et al., 2004). In contrast with the standard function of classical splicing regulators acting through cognate binding sites specific to certain target RNAs, core factors could both carry out general, essential functions for intron removal and in addition display regulatory potential if their levels become limiting for the function of complexes, leading to differences in the efficiency or kinetics of assembly on alternative splice sites. This could be the case for the knockdown of SmB/B′, a component of the Sm complex present in most spliceosomal snRNPs, which leads to autoregulation of its own pre-mRNA as well as to effects on hundreds of other ASEs, particularly in genes encoding RNA processing factors (Saltzman et al., 2011).

The results of our genome-wide siRNA screen for regulators of Fas AS (Tejedor et al., 2014) provide unbiased evidence that an extensive number of core SFs have the potential to contribute to the regulation of splice site selection. This is, in fact, the most populated category of screen hits. Similarly, the results of our network highlight coherent effects on alternative splice site choice of multiple core components, revealing also that the extent of their effects on alternative splice site selection can be largely attributed to the duration and order of their recruitment in the splicing reaction.

Remarkably, effects on splice site selection are associated with depletion not only of complexes involved in early splice site recognition but also of factors involved in complex B formation or even in catalytic activation of fully assembled spliceosomes. While particular examples of AS regulation at the time of the transition from complex A to B and even at later steps had been reported (House and Lynch, 2006; Bonnal et al., 2008; Lallena et al., 2002), our results indicate that the implication of late factors is quite general. Such links, e.g., those involving PRP8 and other interacting factors closely positioned near the reacting chemical groups at the time of catalysis, like PRP31 or U5 200KD, are actually part of the persistent functional interactions that emerge from the analysis of any subsample of ASEs analyzed, highlighting their general regulatory potential in splice site selection. How can factors involved in late steps of the splicing process influence splice site choices? It is conceivable that limiting amounts of late spliceosome components would favor splicing of alternative splice site pairs harboring early complexes that can more efficiently recruit late factors, or influence the kinetics of conformational changes required for catalysis. In this context, the regulatory plasticity revealed by our network analysis may be contributed, at least in part, by the emerging realization that a substantial number of SFs contain disordered regions when analyzed in isolation (Korneta and Bujnicki, 2012; Chen and Moore, 2014). Such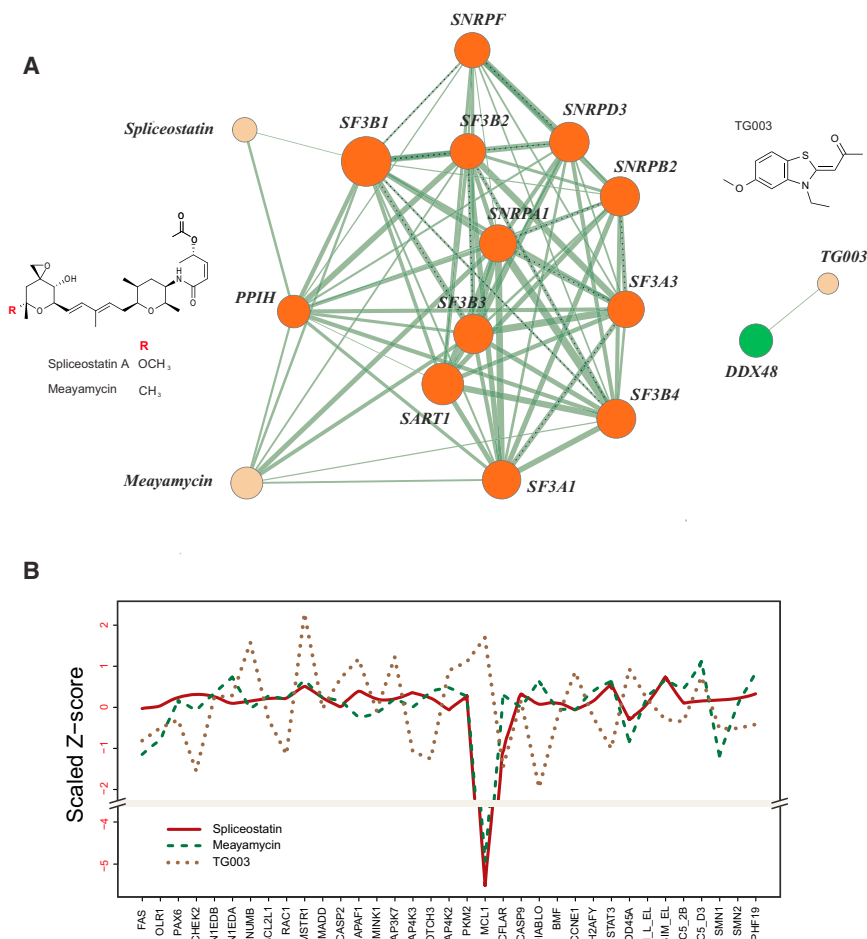 regions may be flexible to adopt different conformations in the presence of other spliceosomal components, allowing alternative routes for spliceosome assembly on different introns, with some pathways being more sensitive than others to the depletion of a general factor.

Kinetic effects on the assembly and/or engagement of splice sites to undergo catalysis would be particularly effective if spliceosome assembly is not an irreversible process. Results of single molecule analysis are indeed compatible with this concept (Tseng and Cheng, 2008; Abelson et al., 2010; Hoskins et al., 2011; Hoskins and Moore, 2012; Shcherbakova et al., 2013), thus opening the extraordinary complexity of conformational transitions and dynamic compositional changes of the spliceosome as possible targets for regulation. Interestingly, while depletion of early factors tends to favor exon skipping, depletion of late factors causes a similar number of exon inclusion and skipping effects, suggesting high plasticity in splice site choice at this stage of the process.

Given this potential, it is conceivable that modulation of the relative concentration of core components can function as a physiological mechanism for splicing regulation, for example during development and cell differentiation. Indeed, variations in the relative levels of core spliceosomal components have been reported (Wong et al., 2013). Furthermore, recent reports revealed the high incidence of mutations in core SFs in cancer. For example, the gene encoding SF3B1, a component of U2 snRNP involved in branchpoint recognition, has been described as among the most highly mutated in myelodysplastic syndromes (Yoshida et al., 2011), chronic lymphocytic leukemia (Quesada et al., 2012), and other cancers (reviewed in Bonnal et al., 2012), correlating with different disease outcomes. Remarkably, SF3B1 is the physical target of antitumor drugs like Spliceostatin A and—likely—Meayamycin, as captured by our study. These observations are again consistent with the idea that depletion, mutation, or drug-mediated inactivation of core SFs may not simply cause a collapse of the splicing process—at least under conditions of limited physical or functional depletion—but rather lead to changes in splice site selection that can contribute to physiological, pathological, or therapeutic outcomes. SF3B1 participates in the stabilization and proofreading of U2 snRNP binding to the branchpoint region (Corrionero et al., 2011), and it is therefore possible that variations in the activity of this protein result in switches between alternative 3′ splice sites harboring branch sites with different strengths and/or flanking decoy binding sites. Similar concepts may apply to explain the effects of mutations in other SFs linked to disease, including, for example, mutations in PRP8 leading to retinitis pigmentosa (Pena et al., 2007).

In summary, the data and methodological approaches presented in this study provide a rich resource for understanding the function of the spliceosome and the mechanisms of AS regulation, including the identification of targets of physiological, pathological, or pharmacological perturbations within the complex splicing machinery.

(D) Schematic representation of the distribution of composite HNRPC binding sites/3′ splice site-like sequences in representative examples of the splicing events analyzed. The relative position of Alu elements in these regions is also shown. Genomic distances are drawn to scale. The direction of the arrows indicates up or downregulation of cassette exons upon knockdown of U2AF1 (black arrows) or HNRPC (white arrows). See also Figure S6.

A

SNRPF

Spliceostatin

SF3B1   SF3B2   SNRPD3

SNRPB2

SNRPA1

SF3A3

PPIH   SF3B3

SART1   SF3B4

Meayamycin   SF3A1

Spliceostatin A   OCH₃
Meayamycin   CH₃

TG003

TG003

DDX48

B

Scaled Z-score

— Spliceostatin
-- Meayamycin
···· TG003

FAS OLR1 PAX6 CHEK2 FN1EDB FN1EDA NUMB BCL2L1 RAC1 MSTR1 MADD CASP2 APAF1 MINK1 MAP3K7 MAP4K3 NOTCH3 MAP4K2 PKM2 MCL1 CFLAR CASP9 DIABLO BMF CCNE1 H2AFY STAT3 GADD45A BIM_L_EL BIM_EL BIRC5_2B BIRC5_D3 SMN1 SMN2 PHF19

Figure 7. Mapping the Effects of Pharmacological Treatments to the Splicing Network
(A) Functional connections of splicing inhibitory drugs with the splicing machinery. See text for details.
(B) Splicing perturbation profiles for Spliceostatin A, Meayamycin, or TG003 treatments across the 35 ASEs analyzed.

each knockdown-AS pair. From this value x_i the effect of each treatment was summarized as a robust Z score (Birmingham et al., 2009):

$$Z_i' = \frac{x_i - \tilde{x}}{\hat{\sigma}}$$

where $\tilde{x}$ and $\hat{\sigma}$ are the particular ASE's sample median and median absolute deviation (*MAD*)-based consistent estimator of the standard deviation ($\hat{\sigma} \approx 1.483\ MAD$), respectively.

Robust Sample Correlation Estimation

Robust correlation estimation is based on an iterative weighting algorithm that discriminates between technical outliers and reliable measurements with high leverage. The weighting for measurements relies on calculation of a reliability index that takes into account their cumulative influence on correlation estimates of the complete data set. Algorithmic details are provided in Supplemental Experimental Procedures.

Network Reconstruction

The process of network reconstruction relies on the regularization-based graphical lasso (glasso) algorithm for graphical model selection (Friedman et al., 2008). The glasso process for network reconstruction was implemented using the glasso R package by Jerome Friedman, Trevor Hastie, and Rob Tibshirani (http://statweb.stanford.edu/~tibs/glasso/). A detailed exposition of the algorithmic steps for network reconstruction and analysis, glasso, and its use for graphical model selection is provided in Supplemental Experimental Procedures.

Principal Component Analysis of Ancillary Connections

PCA for ancillary edges was performed using R-mode PCA (R function princomp). The input data set is the scaled 95 × 35 matrix containing the absolute average correlation for each of the 95 edges over the subsets of the 10,000 samples that included each of the 35 events. The coordinates for projecting the 95 edges are directly derived from their scores on the first two principal components. The arrows representing the top ten discriminatory events (based on the vector norm of the first two PC loadings) were drawn from the origin to the coordinate defined by the first two PC loadings scaled by a constant factor for display purposes.

Splice Site Scoring, Identification of Probable HNRPC Binding Sites and Alu Elements

For the identification and scoring of 3′ ss we used custom-built position weight matrices (PWMs) of length 21−8 intronic plus three exonic positions. The matrices were built using a set of human splice sites from constitutive, internal exons compiled from the hg19 UCSC annotation database (Karolchik et al., 2014). Background nucleotide frequencies were estimated from a set of strictly intronic regions. Threshold was set based on a FDR of 0.5/kbp of random intronic sequences generated using a second-order Markov chain of actual intronic regions.

We considered as probable HNRPC binding sites consecutive stretches of U{4} as reported in Zarnack et al. (2013). Annotation of ALU elements was

EXPERIMENTAL PROCEDURES

siRNA Library Transfection and mRNA Extraction

HeLa cells were transfected in biological triplicates with siRNAs pools (ON TARGET plus smartpool, Dharmacon, Thermo Scientific) against 270 splicing and chromatin remodeling factors (see Table S2). Endogenous mRNAs were purified 72 hr posttransfection by using oligo dT-coated 96-well plates (mRNA catcher PLUS, Life Technologies) following the manufacturer's instructions. RNA samples corresponding to treatments with splicing arresting drugs were isolated with the Maxwell 16 LEV simplyRNA kit (Promega).

RT-PCR and High-Throughput Capillary Electrophoresis

Cellular mRNAs were reverse transcribed using Superscript III Retro Transcriptase (Invitrogen, Life Technologies) following the manufacturer's recommendations. PCR reactions for every individual splicing event analyzed were carried out using forward and reverse primers in exonic sequences flanking the alternatively spliced region of interest and further reagents provided in the GOTaq DNA polymerase kit (GoTaq, Promega). Primers used in this study are listed in Table S6.

HTCE measurements for the different splicing isoforms were performed in 96-well format in a Labchip GX Caliper workstation (Caliper, Perkin Elmer) using a HT DNA High Sensitivity LabChip chip (Perkin Elmer). Data values were obtained using the Labchip GX software analysis tool (version 3.0).

Quantification of AS Changes from HTCE Measurements

Robust estimates of isoform ratios upon siRNA or pharmacological treatments were obtained using the median PSI indexes of the biological triplicates for

based on the RepeatMasker track developed by Arian Smit (http://www. repeatmasker.org) and the Repbase library (Jurka et al., 2005).

More detailed experimental procedures are provided in Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures, six tables, and Supplemental Experimental Procedures and can be found with this article at http://dx.doi. org/10.1016/j.molcel.2014.10.030.

AUTHOR CONTRIBUTIONS

P.P. set up the computational pipeline to generate the splicing network. J.R.T. set up and carried out the experimental analyses. L.V. contributed to the network analyses involving drugs. P.P., J.R.T., and J.V. designed the project and wrote the manuscript.

ACKNOWLEDGMENTS

REFERENCES

Abelson, J., Blanco, M., Ditzler, M.A., Fuller, F., Aravamudhan, P., Wood, M., Villa, T., Ryan, D.E., Pleiss, J.A., Maeder, C., et al. (2010). Conformational dynamics of single pre-mRNA molecules during in vitro splicing. Nat. Struct. Mol. Biol. *17*, 504–512.

Ashton-Beaucage, D., Udell, C.M., Lavoie, H., Baril, C., Lefrançois, M., Chagnon, P., Gendron, P., Caron-Lizotte, O., Bonneil, E., Thibault, P., and Therrien, M. (2010). The exon junction complex controls the splicing of MAPK and other long intron-containing transcripts in Drosophila. Cell *143*, 251–262.

Barash, Y., Calarco, J.A., Gao, W., Pan, Q., Wang, X., Shai, O., Blencowe, B.J., and Frey, B.J. (2010). Deciphering the splicing code. Nature *465*, 53–59.

Bechara, E.G., Sebestyén, E., Bernardis, I., Eyras, E., and Valcárcel, J. (2013). RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation. Mol. Cell *52*, 720–733.

Behrens, S.E., Tyc, K., Kastner, B., Reichelt, J., and Lührmann, R. (1993). Small nuclear ribonucleoprotein (RNP) U2 contains numerous additional proteins and has a bipartite RNP structure under splicing conditions. Mol. Cell. Biol. *13*, 307–319.

Bessonov, S., Anokhina, M., Will, C.L., Urlaub, H., and Lührmann, R. (2008). Isolation of an active step I spliceosome and composition of its RNP core. Nature *452*, 846–850.

Birmingham, A., Selfors, L.M., Forster, T., Wrobel, D., Kennedy, C.J., Shanks, E., Santoyo-Lopez, J., Dunican, D.J., Long, A., Kelleher, D., et al. (2009). Statistical methods for analysis of high-throughput RNA interference screens. Nat. Methods *6*, 569–575.

Bonnal, S., Martínez, C., Förch, P., Bachi, A., Wilm, M., and Valcárcel, J. (2008). RBM5/Luca-15/H37 regulates Fas alternative splice site pairing after exon definition. Mol. Cell *32*, 81–95.

Bonnal, S., Vigevani, L., and Valcárcel, J. (2012). The spliceosome as a target of novel antitumour drugs. Nat. Rev. Drug Discov. *11*, 847–859.

Bottner, C.A., Schmidt, H., Vogel, S., Michele, M., and Käufer, N.F. (2005). Multiple genetic and biochemical interactions of Brr2, Prp8, Prp31, Prp1 and Prp4 kinase suggest a function in the control of the activation of spliceosomes in Schizosaccharomyces pombe. Curr. Genet. *48*, 151–161.

Chatr-Aryamontri, A., Breitkreutz, B.J., Heinicke, S., Boucher, L., Winter, A., Stark, C., Nixon, J., Ramage, L., Kolas, N., O'Donnell, L., et al. (2013). The BioGRID interaction database: 2013 update. Nucleic Acids Res. *41* (Database issue), D816–D823.

Chen, W., and Moore, M.J. (2014). The spliceosome: disorder and dynamics defined. Curr. Opin. Struct. Biol. *24*, 141–149.

Clark, T.A., Sugnet, C.W., and Ares, M., Jr. (2002). Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. Science *296*, 907–910.

Collins, C.A., and Guthrie, C. (1999). Allele-specific genetic interactions between Prp8 and RNA active site residues suggest a function for Prp8 at the catalytic core of the spliceosome. Genes Dev. *13*, 1970–1982.

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

Corrionero, A., Miñana, B., and Valcárcel, J. (2011). Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A. Genes Dev. *25*, 445–459.

Fica, S.M., Tuttle, N., Novak, T., Li, N.S., Lu, J., Koodathingal, P., Dai, Q., Staley, J.P., and Piccirilli, J.A. (2013). RNA catalyses nuclear pre-mRNA splicing. Nature *503*, 229–234.

Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C., and Jensen, L.J. (2013). STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. *41* (Database issue), D808–D815.

Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. Biostatistics *9*, 432–441.

Fu, X.D., and Ares, M., Jr. (2014). Context-dependent control of alternative splicing by RNA-binding proteins. Nat. Rev. Genet. *15*, 689–701.

Gao, Y., Vogt, A., Forsyth, C.J., and Koide, K. (2013). Comparison of splicing factor 3b inhibitors in human cells. ChemBioChem *14*, 49–52.

Häcker, I., Sander, B., Golas, M.M., Wolf, E., Karagöz, E., Kastner, B., Stark, H., Fabrizio, P., and Lührmann, R. (2008). Localization of Prp8, Brr2, Snu114 and U4/U6 proteins in the yeast tri-snRNP by electron microscopy. Nat. Struct. Mol. Biol. *15*, 1206–1212.

Hasegawa, M., Miura, T., Kuzuya, K., Inoue, A., Won Ki, S., Horinouchi, S., Yoshida, T., Kunoh, T., Koseki, K., Mino, K., et al. (2011). Identification of SAP155 as the target of GEX1A (Herboxidiene), an antitumor natural product. ACS Chem. Biol. *6*, 229–233.

Horowitz, D.S., Kobayashi, R., and Krainer, A.R. (1997). A new cyclophilin and the human homologues of yeast Prp3 and Prp4 form a complex associated with U4/U6 snRNPs. RNA *3*, 1374–1387.

Hoskins, A.A., and Moore, M.J. (2012). The spliceosome: a flexible, reversible macromolecular machine. Trends Biochem. Sci. *37*, 179–188.

Hoskins, A.A., Friedman, L.J., Gallagher, S.S., Crawford, D.J., Anderson, E.G., Wombacher, R., Ramirez, N., Cornish, V.W., Gelles, J., and Moore, M.J. (2011). Ordered and dynamic assembly of single spliceosomes. Science *331*, 1289–1295.

House, A.E., and Lynch, K.W. (2006). An exonic splicing silencer represses spliceosome assembly after ATP-dependent exon recognition. Nat. Struct. Mol. Biol. *13*, 937–944.

Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., and Walichiewicz, J. (2005). Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. *110*, 462–467.

Kaida, D., Motoyoshi, H., Tashiro, E., Nojima, T., Hagiwara, M., Ishigami, K., Watanabe, H., Kitahara, T., Yoshida, T., Nakajima, H., et al. (2007). Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. Nat. Chem. Biol. *3*, 576–583.

Karolchik, D., Barber, G.P., Casper, J., Clawson, H., Cline, M.S., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M., et al. (2014). The UCSC Genome Browser database: 2014 update. Nucleic Acids Res. *42* (Database issue), D764–D770.

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.

Kim, D., Kim, M.S., and Cho, K.H. (2012). The core regulation module of stress-responsive regulatory networks in yeast. Nucleic Acids Res. *40*, 8793–8802.

Korneta, I., and Bujnicki, J.M. (2012). Intrinsic disorder in the human spliceosomal proteome. PLoS Comput. Biol. *8*, e1002641.

Lallena, M.J., Chalmers, K.J., Llamazares, S., Lamond, A.I., and Valcárcel, J. (2002). Splicing regulation at the second catalytic step by Sex-lethal involves 3′ splice site recognition by SPF45. Cell *109*, 285–296.

Luco, R.F., Allo, M., Schor, I.E., Kornblihtt, A.R., and Misteli, T. (2011). Epigenetics in alternative pre-mRNA splicing. Cell *144*, 16–26.

Makarova, O.V., Makarov, E.M., Urlaub, H., Will, C.L., Gentzel, M., Wilm, M., and Lührmann, R. (2004). A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing. EMBO J. *23*, 2381–2391.

Mozaffari-Jovin, S., Santos, K.F., Hsiao, H.H., Will, C.L., Urlaub, H., Wahl, M.C., and Lührmann, R. (2012). The Prp8 RNase H-like domain inhibits Brr2-mediated U4/U6 snRNA unwinding by blocking Brr2 loading onto the U4 snRNA. Genes Dev. *26*, 2422–2434.

Muraki, M., Ohkawara, B., Hosoya, T., Onogi, H., Koizumi, J., Koizumi, T., Sumi, K., Yomoda, J., Murray, M.V., Kimura, H., et al. (2004). Manipulation of alternative splicing by a newly developed inhibitor of Clks. J. Biol. Chem. *279*, 24246–24254.

Muro, A.F., Chauhan, A.K., Gajovic, S., Iaconcig, A., Porro, F., Stanta, G., and Baralle, F.E. (2003). Regulated splicing of the fibronectin EDA exon is essential for proper skin wound healing and normal lifespan. J. Cell Biol. *162*, 149–160.

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

Park, J.W., Parisky, K., Celotto, A.M., Reenan, R.A., and Graveley, B.R. (2004). Identification of alternative splicing regulators by RNA interference in Drosophila. Proc. Natl. Acad. Sci. USA *101*, 15974–15979.

Pena, V., Liu, S., Bujnicki, J.M., Lührmann, R., and Wahl, M.C. (2007). Structure of a multipartite protein-protein interaction domain in splicing factor prp8 and its link to retinitis pigmentosa. Mol. Cell *25*, 615–624.

Pleiss, J.A., Whitworth, G.B., Bergkessel, M., and Guthrie, C. (2007). Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. PLoS Biol. *5*, e90.

Quesada, V., Conde, L., Villamor, N., Ordóñez, G.R., Jares, P., Bassaganyas, L., Ramsay, A.J., Beà, S., Pinyol, M., Martínez-Trillos, A., et al. (2012). Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat. Genet. *44*, 47–52.

Saltzman, A.L., Pan, Q., and Blencowe, B.J. (2011). Regulation of alternative splicing by the core spliceosomal machinery. Genes Dev. *25*, 373–384.

Shcherbakova, I., Hoskins, A.A., Friedman, L.J., Serebrov, V., Corrêa, I.R., Jr., Xu, M.Q., Gelles, J., and Moore, M.J. (2013). Alternative spliceosome assembly pathways revealed by single-molecule fluorescence microscopy. Cell Rep. *5*, 151–165.

Teigelkamp, S., Achsel, T., Mundt, C., Göthel, S.F., Cronshagen, U., Lane, W.S., Marahiel, M., and Lührmann, R. (1998). The 20kD protein of human [U4/U6.U5] tri-snRNPs is a novel cyclophilin that forms a complex with the U4/U6-specific 60kD and 90kD proteins. RNA *4*, 127–141.

Tejedor, J.R., Papasaikas, P., and Valcárcel, J. (2014). Genome-wide identification of Fas/CD95 alternative splicing regulators reveals links with iron homeostasis. Mol. Cell *57*, this issue, 23–38.

Tseng, C.K., and Cheng, S.C. (2008). Both catalytic steps of nuclear pre-mRNA splicing are reversible. Science *320*, 1782–1784.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell *136*, 701–718.

Watson, E., MacNeil, L.T., Arda, H.E., Zhu, L.J., and Walhout, A.J. (2013). Integration of metabolic and gene regulatory networks modulates the C. elegans dietary response. Cell *153*, 253–266.

Wong, A.K., Park, C.Y., Greene, C.S., Bongo, L.A., Guan, Y., and Troyanskaya, O.G. (2012). IMP: a multi-species functional genomics portal for integration, visualization and prediction of protein functions and networks. Nucleic Acids Res. *40* (Web Server issue), W484–W490.

Wong, J.J., Ritchie, W., Ebner, O.A., Selbach, M., Wong, J.W., Huang, Y., Gao, D., Pinello, N., Gonzalez, M., Baidya, K., et al. (2013). Orchestrated intron retention regulates normal granulocyte differentiation. Cell *154*, 583–595.

Yang, Y., Han, L., Yuan, Y., Li, J., Hei, N., and Liang, H. (2014). Gene co-expression network analysis reveals common system-level properties of prognostic genes across cancer types. Nat. Commun. *5*, 3231.

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

Zarnack, K., König, J., Tajnik, M., Martincorena, I., Eustermann, S., Stévant, I., Reyes, A., Anders, S., Luscombe, N.M., and Ule, J. (2013). Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. Cell *152*, 453–466.

Zhang, C., Frias, M.A., Mele, A., Ruggiu, M., Eom, T., Marney, C.B., Wang, H., Licatalosi, D.D., Fak, J.J., and Darnell, R.B. (2010). Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. Science *329*, 439–443.