Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks

Endre Sebestyén,1,5 Babita Singh,1,5 Belén Miñana,1,2 Amadís Pagès,1 Francesca Mateo,3 Miguel Angel Pujana,3 Juan Valcárcel,1,2,4 and Eduardo Eyras1,4

1Universitat Pompeu Fabra, E08003 Barcelona, Spain; 2Centre for Genomic Regulation, E08003 Barcelona, Spain; 3Program Against Cancer Therapeutic Resistance (ProCURE), Catalan Institute of Oncology (ICO), Bellvitge Institute for Biomedical Research (IDIBELL), E08908 L’Hospitalet del Llobregat, Spain; 4Catalan Institution for Research and Advanced Studies, E08010 Barcelona, Spain

Alternative splicing is regulated by multiple RNA-binding proteins and influences the expression of most eukaryotic genes. However, the role of this process in human disease, and particularly in cancer, is only starting to be unveiled. We systematically analyzed mutation, copy number, and gene expression patterns of I348 RNA-binding protein (RBP) genes in II solid tumor types, together with alternative splicing changes in these tumors and the enrichment of binding motifs in the alternatively spliced sequences. Our comprehensive study reveals widespread alterations in the expression of RBPs, as well as novel mutations and copy number variations in association with multiple alternative splicing changes in cancer drivers and oncogenic pathways. Remarkably, the altered splicing patterns in several tumor types recapitulate those of undifferentiated cells. These patterns are predicted to be mainly controlled by MBNL1 and involve multiple cancer drivers, including the mitotic gene NUMAI. We show that NUMAI alternative splicing induces enhanced cell proliferation and centrosome amplification in nontumorigenic mammary epithelial cells. Our study uncovers novel splicing networks that potentially contribute to cancer development and progression.

[Supplemental material is available for this article.]

Alternative splicing alterations are emerging as important signatures to further understand tumor formation and to develop new therapeutic strategies (Grosso and Carmo-Fonseca 2014). Specific alternative splicing changes that confer tumor cells with a selective advantage may be caused by mutations in splicing regulatory sequences (Dorman et al. 2014) and/or regulatory factors (Brooks et al. 2014). Various splicing factors have been described to be mutated in cancer, including SF3B1, SRSF2, ZRSR2, and U2AF1 in myelodysplastic syndromes and lymphoid leukemias (Yoshida et al. 2011), RBM10 and U2AF1 in lung tumors (Imlenlski et al. 2012; Brooks et al. 2014), and SF3B1 in breast tumors (Ellis et al. 2012; Maguire et al. 2015). These mutations generally impair the recognition of regulatory sites, thereby affecting the splicing of multiple genes, including oncogenes and tumor suppressors (Kim et al. 2015). On the other hand, increasing evidence shows that changes in the relative concentration of splicing factors can also trigger oncogenic processes. For instance, splicing factors from the SR and hnRNP families are overexpressed in multiple tumor types and induce splicing changes that contribute to cell proliferation (Karni et al. 2007; Golan-Gerstl et al. 2011). Similarly, down-regulation of splicing factors that act as tumor suppressors has also been observed (Wang et al. 2014; Zong et al. 2014).

Importantly, specific alternative splicing events can substantially recapitulate cancer-associated phenotypes linked to mutations or expression alterations of splicing factors. This is the case of NUMAI, for which the reversal of the splicing change induced by RBM10 mutations in lung cancer cells can revert the proliferative phenotype (Bechara et al. 2013). Events that contribute to cancer are often controlled by multiple factors, like the exon skipping event of MST1R involved in cell invasion, which is controlled by SRSF1 (Ghigna et al. 2005), HNRNP2B1 (Golan-Gerstl et al. 2011), HNRNP1, and SRSF2 (Moon et al. 2014). Furthermore, some events may be affected by both mutations and expression changes in splicing factors. For instance, mutations in RBM10 or down-regulation of QKI lead to the same splicing change in NUMAI that promotes cell proliferation (Bechara et al. 2013; Zong et al. 2014).

Alternative splicing changes that characterize and contribute to the pathophysiology of cancer (Sebestyén et al. 2015) are thus potentially triggered by alterations in a complex network of RNA binding proteins, which remain to be comprehensively described. To elucidate the complete set of alterations in these factors and how they globally affect alternative splicing that may contribute to cancer, we analyzed RNA and DNA sequencing data from The Cancer Genome Atlas (TCGA) project for 11 solid tumor types.

Results

RBPs are frequently de-regulated and characterize tumor types

Using TCGA data for 11 solid tumor types (Supplemental Table S1), we analyzed the differential gene expression between normal and tumor sample pairs of 1348 genes encoding known and unknown RBPs.
predicted RNA binding proteins (RBPs) (Methods; Supplemental Table S2). The majority of these genes (1143, 84.8%) show significant differential expression in at least one tumor type (Supplemental Fig. S1; Supplemental Table S3). Examining in detail 162 RBP genes annotated as known or putative splicing factors (SFs), we find they can be separated into three groups. One group is frequently up-regulated, another one down-regulated, and a third one shows opposite patterns in the three kidney tumor types (KICH, KIRC, KIRP) compared to other tumor types (Fig. 1A). Moreover, 132 (80%) of them are differentially expressed in at least one tumor type and 45 were previously associated with oncogenic or tumor suppressor activities (Fig. 1A, labeled in red; Supplemental Table S4). We also found apparent discrepancies with previous literature. For instance, although SRSF5 was described as oncogenic (Huang et al. 2007), it is down-regulated in six tumor types; and TRA2B, reported as oncogenic in breast cancer (Watermann et al. 2006), is up-regulated in LUSC but down-regulated in KICH and thyroid carcinoma (THCA). Also, the oncogenic SRSF2, SRSF3, and SRSF6 (Xiao et al. 2007; Jia et al. 2010; Jensen et al. 2014) are down-regulated in KICH, while the oncogenic SRSF1 and the tumor suppressor RBM4 do not show any significant expression changes. New patterns also emerge, including up-regulation of genes from the RBM family and down-regulation of the genes from the MBNL family (Fig. 1A). Unsupervised clustering of all the 4442 tumor samples using normalized expression values per sample for SFs (Supplemental Fig. S2) or for all RBPs (Supplemental Fig. S3) largely separates samples by tumor type. This pattern is reproduced when using a different gene set of similar size (Supplemental Fig. S3). Closer inspection of the BRCA and COAD samples reveals that the expression patterns of RBPs also reproduce tumor subtypes (Supplemental Figs. 4, 5). Collectively, expression analyses indicate that RBPs are frequently and specifically altered in human tumors, and they represent one of the multiple expression profiles that characterize the intrinsic properties of the tumor types and their tissue of origin.

RBPs de-regulation is partially driven by genomic alterations

To further define the extent to which RBP genes are altered in human cancer, the TCGA data were analyzed for protein-affecting mutations and copy number variations (CNVs). Although most of the RBP genes are mutated in tumors (Supplemental Table S4), they are generally mutated in fewer samples compared to candidate cancer drivers (listed in Supplemental Tables S5, S6) and other genes (Fig. 1B). We confirmed 5.4% (25/458) of LUAD tumor samples to have protein-affecting mutations for RBM10, in agreement with previous studies (Imielinski et al. 2012). Using this case as a reference, we observed 205 (15.2%) of all RBPs (13 or 8% of the 162 SFs) mutated in >5% of samples in a given tumor type (Table 1; Supplemental Table S3). In general, there is a weak association between mutations and expression changes, whereas a number of genes show mutual exclusion of mutations and expression changes.
changes (Table 1; Supplemental Table S3). On the other hand, CNVs are highly recurrent across samples, with gains more frequent than losses (Fig. 1C; complete list of CNVs available in Supplemental Table S3). Up-regulated RBPs had generally more CNV gains than nonregulated RBPs (Mann–Whitney U test P-value < 2.2 × 10^{-15}), and 90% of the up-regulated RBPs had CNV gains in >50% of the samples from the same tumor type. The splicing factors ESRP1, PARP1, and KHDRBS3, which are near the reported 8q24 amplification (The Cancer Genome Atlas Network 2012b), show CNVs in BRCA, COAD, HNSC, LIHC, and LUAD (Fig. 1C), with ESRP1 and PARP1 showing frequent association with up-regulation (Table 1). We also observed the frequent amplification of TRA2B, IGF2BP2, and FXR1 in HNSC and LUSC, as reported before (The Cancer Genome Atlas Network 2012a, 2015). However, the association with up-regulation is only observed in LUSC (Table 1). CNV losses are less frequent than gains and show weaker association with down-regulation (Table 1; Supplemental. Table S3). Finally, only a low fraction of the CNVs were detected as focal (Table 1; Supplemental Table S3). These results indicate that many of the detected expression alterations of RBPs may be explained by CNVs, although most often in association with large chromosomal events, thereby highlighting the potential relevance of these alterations in shaping the tumor phenotypes.

RBP mutations associate with abnormal splicing patterns in cancer

Next, we investigated the patterns of differential splicing that may occur as a consequence of the alterations described above. To determine those possibly related to expression alterations in RBPs, we first evaluated the significant splicing changes in tumors compared to normal tissue, considering five major event types: skipping exon (SE), alternative 3′-splice-site (AS), alternative 5′-splice-site (A3), mutually exclusive exon (MX), and retained intron (RI) events (Fig. 2A; Supplemental Tables S7, S8). We examined the splicing patterns of cancer drivers in more detail. From the 937 drivers collected (Supplemental Tables S5, S6), 653 (69.7%) have annotated events and 292 (31.2%) have at least one differentially spliced event, with a number of them occurring in multiple tumor types (Supplemental Fig. S6). Moreover, seven of the 11 tumor types show enrichment of differentially spliced events in drivers (Fig. 2B, in red; Supplemental Table S9). Additionally, various cancer hallmarks (Liberzon et al. 2015) are enriched in differentially spliced events but not in differentially expressed genes (Fig. 2C; Supplemental Table S10). These results suggest that alternative splicing contributes to cancer development independently of mutations or expression alterations.

To determine the splicing changes related to mutations in RBPs, we compared the inclusion levels (percent spliced in, PSpI) of the events between samples with or without protein-affecting mutations for each RBP (Fig. 2D,E; Supplemental Fig. S8; Supplemental Table S11). We found less differentially spliced events compared with the tumor-normal comparison, and some of the detected events affect cancer drivers (Table 2). HNRNPL had 16 mutations in COAD, 13 of them indels in an RNA recognition motif, causing frameshifts (Fig. 2F). These mutations had 42 events associated with them (Z-score = 41.35), including the cancer driver CASP8 (Fig. 2G). In COAD, we also found the putative RBP MACF1 (Baltz et al. 2012) with mutations along the entire protein (42 events, Z-score = 45.35) (Supplemental Fig. S9). MACF1 is a component of the WNT-pathway known to affect splicing (Bordonaro 2013). In LUAD, we found 53 events associated with RBM10 (Z-score = 52.35), 83 events associated with U2AF1 (Z-score = 81.35), and 49 events associated with the transcription factor and putative RBP TCF20 (Castello et al. 2012) (Z-score = 50.35). Most of the TCF20 alterations consist of an insertion in a glycin-rich region at the N terminus (Supplemental Fig. S9). As a comparison, we analyzed SF3BI and found 64 differentially spliced events (Z-score = 63.35) in BRCA, despite being mutated in only 1.7% of the tested samples (16 out of 956 for which we had mutation and RNA-seq data) (Supplemental Fig. S9). Notably, compared with the proportions of the different event types, there is a significant enrichment of A3 events associated with SF3BI (Fisher’s exact test P-value = 1.45 × 10^{-11}), of A5 events for TCF20 (P-value = 2.09 × 10^{-8}), and of SE events for RBM10 (P-value = 0.003) (Supplemental Fig. S9; Supplemental Table S12). There is also a significant depletion of SE events for TCF20 (P-value = 0.0014), of SE events for SF3BI (P-value = 2.24 × 10^{-8}), and of A5 events for U2AF1 (P-value = 1.67 × 10^{-3}) and HNRNPL (P-value = 0.005) (Fig. 2H). Furthermore, 17 (20%) of the genes detected to change splicing with SF3BI, 32 (38%) associated with U2AF1, and 21 (30%) related with RBM10 mutations coincide with those obtained in recent analyses (Supplemental Table S13; Brooks et al. 2014; Darman et al. 2015; Maguire et al. 2015). The limited overlap could be due to the use of different methods for differential splicing analysis and the use of different sample sets. In summary, we validated known RBP mutations affecting alternative splicing and described new ones, whose mutations potentially affect splicing patterns. Our analyses reveal a rich source of new information about alternative splicing events associated with RBP alterations with a potential relevance in cancer.

Common and specific cancer patterns of differential splicing mediated by RBPs

The above results suggest that mutations in RBPs are not the single cause of splicing changes in tumors. Therefore, we further characterized the splicing changes between tumor and normal samples to determine their association with the expression of RBPs. We first identified common patterns of splicing changes between pairs of tumor types by selecting events with a strong correlation with a differentially expressed SF in the two tumor types (Methods). PSI changes (APSI) for these events show high correlation between tumor pairs (Fig. 3A) and indicate potential common regulators (Fig. 3B; Supplemental Fig. S10; Supplemental Table S14). BRCA and LUAD share 229 events associated with various factors, including QKI and SRSF5 (Fig. 3B, upper left panel), whereas KIRC and PRAD share 78 anti-correlating events associated with ESRP2 and MBNL1 (Fig. 3B, upper right panel). RBM47, described as a tumor suppressor (Vanharanta et al. 2014), appears as the main common SF between KIRC and HNSC, with 61 associated events (Fig. 3B, lower left panel). HNSC and LUSC share 141 events, 63 of them associated with RBM28 (Fig. 3B, lower right panel). These results suggest functional connections between RBPs and the splicing changes detected in cancer. Next, we also studied whether there are tumor-specific events (Methods) and found 380 events that largely separate the 4442 tumor samples by type (Fig. 3C; Supplemental Fig. S11; Supplemental Table S15). These splicing changes may be indicative of tumor-type-specific oncogenic mechanisms.

Enriched RBP motifs in differentially spliced events of cancer

To further understand the link between the observed splicing patterns and the RBPs, we tested the enrichment of binding motifs in differentially spliced events in each tumor type. We assigned
binding motifs from RNAcompete (Ray et al. 2013) to 104 of the analyzed RBPs and tested their enrichment (see Methods; Supplemental Figs. S12, S13). Considering enriched motifs of differentially expressed RBPs in the same tumor type, we observed the CELF, RBFOX, and MBNL motifs among the most frequently enriched across tumor types, as well as in luminal breast tumors (Fig. 4A; Supplemental Figs. S14–S18). Down-regulated RBP genes in inclusion events and up-regulated ones in skipping events show more frequently enriched motifs in upstream and exonic regions, consistent with a positional effect (Fig. 4B). On the other hand, down-regulated RBP genes in skipping events and up-regulated ones in inclusion events show enriched motifs most frequently on exons, suggesting that RBPs more often enhance the inclusion of bound exons.

To define candidate target events for differentially expressed RBPs, we selected differentially spliced events whose PSI correlates with the RBP expression (|R| > 0.5, Spearman) and that contain the corresponding RNA binding motif. We could assign between 20% and 80% of the differentially spliced events to at least one RBP (Fig. 4C; Supplemental Tables S16, S17). We ranked the RBPs in each expression group according to the total number of cancer driving events with differentially spliced events across all tumor types (Fig. 4D; Supplemental Fig. S19). The results of this analysis emphasize the relevance of CELF2, ESRP1, ELAVL1, RBFOX2, PTBP1, QKI, and 80% of the differentially spliced events to at least one RBP.
TRA2B, and RBM47 in relation to alterations of splicing in cancer and highlight novel prominent factors, including HNRNPC, SYNCRIP (also known as HNRNPQ), and genes from the CPEB and MBNL families. We confirmed the role of MBNL1, QKI, RBFOX2, PTBP1, RBM47, and ESRP1 in some of these tumors by comparing the ΔPSI values of the events with those obtained by single knockdown or overexpression experiments in cell lines (Fig. 4E, upper panels; Supplemental Figs. S20–S22; Supplemental Table S18). Additionally, comparing the tumor events with those differentially spliced between human embryonic stem cells (hESC) and differentiated cells or tissues (Han et al. 2013) yielded a high positive correlation of ΔPSI values in BRCA, PRAD, LIHC, and breast luminal tumors, to a lesser extent in COAD and LUAD, and an anti-correlation in KIRC. This is in agreement with the MBNL1 and MBNL2 expression patterns in these tumors and with the majority of the correlating events containing the MBNL motif (Fig. 4E, lower panels; Supplemental Fig. S23; Supplemental Table S19). These results highlight the prominent role of RBP expression changes in the alterations of splicing in cancer and suggest new mechanisms of regulation.

Network analysis uncovers overlapping RBP-mediated regulatory modules in cancer
To identify splicing regulatory networks relevant in cancer, we built clusters with the 162 SFs using the correlation between
Table 2. Events in cancer drivers associated with protein-affecting mutations in RBPs

| Tumor type | Gene   | Cancer driver | Event type | ΔPSI | Pval |
|------------|--------|---------------|------------|------|------|
| BRCA       | CDK12  | TEBB          | SE         | −1.00| 0.0001|
| BRCA       | MDN1   | EBF1          | SE         | −0.23| 0.0343|
| BRCA       | SF3B1  | BCLL1         | A3         | 0.23| 4.14 × 10⁻⁶|
| BRCA       | SF3B1  | MEFA2         | SE         | −0.21| 0.0026|
| COAD       | BAT211 | MLH1          | SE         | 0.70| 0.0109|
| COAD       | CHD2   | CAST          | SE         | −0.15| 0.0396|
| COAD       | HNRPNL | CAPS8         | A5         | 0.17| 0.0271|
| COAD       | KIAA0802| MKN2         | A3         | −0.13| 0.0340|
| COAD       | KIAA0802| MYH11        | MX         | −0.14| 0.0004|
| COAD       | MACF1  | TIP2          | SE         | 0.16| 0.0251|
| COAD       | MHH9   | AURKA         | A5         | 0.16| 0.0462|
| COAD       | MIOY18A| MDM4          | A3         | 0.54| 0.0462|
| COAD       | MIOY18A| PDCD1LG2      | A5         | 0.21| 0.0403|
| COAD       | SPEN   | TITL9         | SE         | −0.17| 0.0118|
| COAD       | YLPM1  | CD44          | SE         | 0.12| 0.0268|
| COAD       | ZC3H18 | ACSL6         | MX         | −0.84| 0.0114|
| COAD       | ZC3H18 | MB1D          | SE         | −0.11| 0.1737|
| HNSC       | DSP    | PAX5          | MX         | −0.12| 0.0034|
| HNSC       | EPPK1  | TAF1          | SE         | 0.24| 0.0007|
| HNSC       | EPPK1  | PTCH1         | SE         | −1.00| 0.0058|
| LIHC       | CHD2   | BU2B1         | SE         | 1.00| 0.0020|
| LIHC       | EPPK1  | NEDD4L        | SE         | −0.26| 0.0339|
| LIHC       | HUE1   | HLAA          | A5         | 0.35| 0.0281|
| LUAD       | CHD2   | SELP          | A3         | −1.00| 0.0008|
| LUAD       | NOM1   | TFG           | AS         | −0.23| 0.0035|
| LUAD       | RBM10  | BLM           | SE         | 0.13| 0.3400|
| LUAD       | RBM10  | CTNND1        | A3         | −0.34| 0.0064|
| LUAD       | RBM10  | MUC1          | SE         | 0.13| 0.3337|
| LUAD       | RBM10  | WNK1          | A3         | 0.22| 0.0033|
| LUAD       | TWISTNB| CIPI1         | MX         | −0.20| 0.0435|
| LUAD       | U2AF1  | BCOR          | SE         | −0.12| 0.3516|
| LUAD       | U2AF1  | CHCHD7        | A3         | −0.11| 0.0005|
| LUAD       | U2AF1  | CTNBN1        | A3         | 0.26| 0.0002|
| LUAD       | U2AF1  | CTNBN1        | RE         | 0.19| 0.0056|
| LUAD       | U2AF1  | MUC1          | MX         | −0.26| 0.0472|
| LUAD       | U2AF1  | PATZ2         | A3         | −0.15| 0.0347|
| LUAD       | U2AF1  | PCMI          | SE         | 0.12| 0.0016|
| LUAD       | U2AF1  | RIPA2         | SE         | −0.26| 0.0008|
| LUAD       | U2AF1  | RIT1          | SE         | −0.13| 0.0018|

For each tumor type and for each RBP with more than 10 associated differentially spliced events, we indicate the events in cancer drivers predicted to have a significant splicing change. The table includes the PSI change between mutated and nonmutated samples (ΔPSI) and the P-value of the comparison after correcting for multiple testing (Pval). SF3B1 is included for comparison. Coordinates for the events are given in Supplemental Table S7.

gene expression and event PSI values. We linked these clusters to differentially spliced genes in enriched cancer hallmarks (Methods). This analysis revealed modules of splicing regulation, with one or two genes as the main regulator of each hallmark across different tumor types. Myogenesis and EMT networks are enriched in almost all tumor types and controlled by similar factors (Fig. S5A−B; Supplemental Figs. S24, S25). Other relevant regulatory modules are the G2 checkpoint (G2M), which includes NUMA1, a gene involved in spindle formation (Zheng et al. 2010), and the Wnt/beta-catenin pathway, which includes NUMB, whose alternative splicing is linked to cell proliferation (Bechara et al. 2013). We predicted the splicing of these genes to be controlled by MBNL1, among other factors (Fig. S5C,D). Interestingly, angiogenesis, which is an enriched hallmark in COAD for splicing but not for gene expression, includes an event in SERPINAS, an inhibitor of serine proteases involved in homeostasis and thrombosis (Suzuki 2008), which we predict to be controlled by RBM47, PTPB1, and RBM28 (Fig. 5E). This analysis reveals new roles of RBPs and splicing in cancer-relevant processes.

MBNL1 contributes to cell proliferation through alternative splicing regulation of NUMAI

MBNL1 emerges as a potentially key regulator of splicing for multiple cancer drivers, especially in luminal breast tumors (Supplemental Fig. S26). In particular, MBNL1 is predicted to control an exon skipping event in NUMAI, whose PSI value correlates with MBNL1 expression (Spearman R = 0.65/0.66 in luminal A/B) and contains the MBNL1 motif (Fig. 6A; Supplemental Fig. S27). The same event has increased PSI value in KIRC (ΔPSI = 0.11, corrected P-value = 3.11 × 10⁻⁶), where MBNL1 is significantly up-regulated compared to normal tissues, providing further support for the dependence of NUMAI splicing on MBNL1. We detected MBNL1 protein in the breast epithelial cell line MCF10A, but not in the luminal-like MCF7 (Supplemental Fig. S28). MBNL1 expression depletion in MCF10A cells with two different siRNAs induced skipping of exon 16 in NUMAI (Supplemental Fig. S28), recapitulating the pattern observed in the tumor samples (Fig. 6B, upper panel) and in MCF7 (Supplemental Fig. S28). This result was validated in the same samples by real-time quantitative RT-PCR (Supplemental Fig. S29). As a control for NUMAI gene expression level, we measured the expression of constitutive exon 5 and 7 using semiquantitative (Supplemental Fig. S28) and real-time quantitative (Supplemental Fig. S29) RT-PCR and found no change upon MBNL1 depletion. This is in agreement with the lack of differential expression observed for NUMAI in all breast tumor samples (corrected P-value = 0.68) and in luminal tumors (corrected P-value = 0.50), compared to the matched normal samples. We also tested the alternative splicing of NUMB exon 9, which we predicted to be dependent on MBNL1 in BRCA luminal tumors. The depletion of MBNL1 recapitulates the NUMB splicing pattern in luminal samples (Fig. 6B, middle panel).

For a comparison with MBNL1, we evaluated the role of QKI, which we also observed down-regulated in BRCA luminal tumors and whose product was detected in MCF10A but not in MCF7 cells (Supplemental Fig. S28). Upon QKI expression depletion, NUMAI exon 16 inclusion changed by a small but reproducible amount in the direction opposite to that with MBNL1 depletion (Fig. 6B, upper panel). Although we did not find a QKI motif on the NUMAI event, this is consistent with the negative correlation found with QKI expression (R = −0.11) in BRCA. Next, we also tested NUMB exon 9, which we predicted to be controlled by QKI in BRCA luminal tumors. QKI depletion induces exon 9 inclusion, recapitulating the pattern in luminal samples (Fig. 6B, middle panels).

NUMAI alternative splicing is likely controlled by other factors (Fig. 5C). For instance, NUMAI exon 16 contains an RBM42 motif and its PSI anti-correlates with RBM42 expression in breast luminal tumors; hence, RBM42 potentially acts as a repressor of exon inclusion. Consistent with this, depletion of RBM42 in MCF7 cells leads to increased inclusion of NUMAI exon 16 (Supplemental Fig. S30).

To measure whether the splicing change in NUMAI had any effect on cancer cell hallmarks, we used antisense oligonucleotides (AONs) targeting specifically the 5′ and 3′ splice-sites of NUMAI exon 16. These AONs promote exon skipping, recapitulating in the MCF10A cells the splicing pattern observed in BRCA luminal tumors, with the AON against the 3′ splice-site being more efficient (Supplemental Fig. S31), as validated using real-time quantitative RT-PCR (Supplemental Fig. S32). Importantly, the AONs did not affect NUMAI expression as measured by semiquantitative.
Figure 3. Common and specific events in tumors. (A) Common events and splicing factors between pairs of tumor types. For each pair of tumor types and for each splicing factor differentially expressed in both tumor types, we plot the correlation of ΔPSI values for events that have a correlation of |R| > 0.5 (Spearman) with these splicing factors in both tumor types. Only factors with more than 50 associated events in both tumor types are shown. Each event is only plotted once, and the color of the plot corresponds to the most common correlating splicing factor. Correlations between alternative splicing events colored by tumor type.

(Supplemental Fig. S31) and real-time quantitative (Supplemental Fig. S32) RT-PCR in the same samples. We then measured the proliferation/viability of MCF10A cells transfected with the AONs targeting NUMA1 exon 16 or with siRNAs against MBNL1 and QKI. We observed a significant increase in cell proliferation/viability at 72, 96, and 120 h upon depletion of MBNL1 or QKI, compared with controls (Student’s t-test P-value < 0.05) (Fig. 6C; Supplemental Table S20), and when transfecting cells with the AON against the 5’ splice site (Student’s t-test P-values < 0.05) (Fig. 6C). Using only the 3’ splice site or both AONs, there was also an increase, albeit not statistically significant, possibly due to a smaller effect of the 3’ AON on NUMA1 splicing and the shared concentration of both AONs. Furthermore, although QKI has a mild effect on NUMA1 exon 16 splicing, QKI down-regulation
Figure 4. Enriched RNA binding motifs in differentially spliced events. (A) Enriched RNA binding motifs in differentially spliced skipping exon events in each tumor type, separated by inclusion (upper panels) or skipping (lower panels) events, and by upstream (left), exonic (middle), or downstream (right) regions. Only enriched motifs for splicing factors that are differentially expressed are indicated in each tumor type. RBPs (y-axis) are grouped according to Figure 1A. Gene up- and down-regulation is indicated in red and blue, respectively. The color intensity indicates the type (according to Figure 1A.

Similar plots for the other event types are given in Supplemental Figures S14–S18. (B) Proportion of enriched motifs in inclusion ($\Delta$PSI > 0.1) (right panel) and skipping ($\Delta$PSI < 0.1) (left panel) events, in each of the event regions (x-axis); upstream (U), exon (E), and downstream (D). Proportions are separated according to whether the RBP gene is up- (red) or down- (blue) regulated. (C) Total proportion (y-axis) of differentially spliced events in each tumor type (x-axis) that are predicted as targets of one or more differentially expressed RBPs with significance Z-score > 1.96. (D) Proportion of differentially spliced events (marked in green) that are predicted as targets of one or more differentially expressed RBPs with significance Z-score > 1.96. RBPs are grouped according to Figure 1A. Upper panels indicate the number of unique cancer drivers with differentially spliced events predicted as targets of each RBP across all tumors. RBPs are ordered according to this value within each group. (E) Correlation (Pearson R) of $\Delta$PSI values (x-axis) in breast tumors (BRCA) and prostate tumors (PRAD) with the $\Delta$PSI of events (y-axes) from the knockdown of MBNL1 and MBNL2 in HeLa cells (lower panels) and from the comparison of stem cells (ESC)- with differentiated cells (CL) (upper panels). Events with a predicted MBNL binding motif are indicated in blue.

has a much stronger effect on NUMB alternative splicing (Fig. 6B), which strongly promotes cell proliferation in a variety of cell types (Misquilla-Ali et al. 2011; Bechara et al. 2013). We further studied the possible effects of the alternative splicing of NUMA1 exon 16 on centrosome amplification, a hallmark of breast carcinogenesis. Using the AON against the S’ splice-site, there is a significant increase in the number of cells with centrosome amplification compared with controls in MCF10A cells (Fig. 6D; Supplemental Table S21). In contrast, the siRNA against MBNL1 did not yield a significant difference, possibly due to the superposition of indirect effects. We also observed an inverse corre-

lation between a signature for chromosome instability (Carter et al. 2006) and the NUMA1 event PSI in luminal tumors (Fig. 6E), which was absent in other tumor types (Supplemental Fig. S33), providing further support for a relationship between NUMA1 alternative splicing and the fidelity of centrosome formation. The exon skipping described in NUMA1 is the only coding difference between the tumor and the normal isoforms. While it is unclear whether this 14-amino acid change alone could explain the observed effects, e.g., we could not detect any protein domain or disordered region (Dosztányi et al. 2005), using GPS (Xue et al. 2011) we predicted loss of a high scoring threonine phosphorylation site
nostic and clinical studies is that the definition of the functional development. An important implication of our study for prog-
splicing as a complementary molecular mechanism to explain tu-
previous reports on the impact of splicing on cancer hallmarks
from those linked to gene expression changes. This agrees with
mal samples. The splicing changes detected are predicted to have
mutations is lower than those changing between tumor and nor-
observed in tumors, since the number of events affected by RBP
appear as major contributors of the alternative splicing changes
CASP9
2008). As HNRNPL also controls
indels in an RNA binding domain would affect the splicing of
a gene involved in programmed cell death (Yang et al. 2008). As HNRNPL also controls CASP9 alternative splicing
(Goehe et al. 2010), this suggests a relevant role of HNRNPL in ap-
optosis. On the other hand, the expression changes in RBP genes
appear as major contributors of the alternative splicing changes
observed in tumors, since the number of events affected by RBP
mutations is lower than those changing between tumor and nor-
mal samples. The splicing changes detected are predicted to have an
impact on many cancer hallmarks, some of them different from
those linked to gene expression changes. This agrees with
previous reports on the impact of splicing on cancer hallmarks
(Oltean and Bates 2013) and highlights the relevance of alternative
splicing as a complementary molecular mechanism to explain tu-
mor development. An important implication of our study for prog-
nostic and clinical studies is that the definition of the functional
impact of somatic genetic and epigenetic alterations should be ex-
panded to include changes in the alternative splicing of the genes in
cancer-relevant pathways.

We identified many RBPs whose expression alteration has
some potential relevance in cancer. For instance, TRA2B is fre-
quently amplified and up-regulated in LUSC and its binding motif
enriched in differentially spliced events, including an SE event in
the DNA damage-response gene CHEK1, reported recently to be
controlled by TRA2 proteins (Best et al. 2014). Despite the similar
genetic alterations between squamous tumors (Hoadley et al.
2014), only LUSC shows overexpression of TRA2B, which could
explain some of the differences in splicing patterns between these
two tumor types. We also found that genes from the MBNL family
are frequently down-regulated in tumors and their binding motif is
enriched in differentially spliced events, especially in breast and
prostate tumors. Additionally, splicing changes related to MBNL1
recapitulate the splicing patterns of undifferentiated cells, in agree-
ment with recent studies describing MBNL1 and MBNL2 as regula-
 tors of a stem cell-related splicing program (Han et al. 2013;
Venables et al. 2013). MBNL1 potentially controls multiple genes
that participate in cancer-related pathways, including the mitotic
gene NUMA1. We described that NUMA1 alternative splicing leads
to higher proliferation and increased centrosome amplification in
normal cells. Notably, the alternative splicing change is not asso-
ciated with an expression change in NUMA1, as measured from
the tumor and normal samples and by semi-quantitative and quan-
titative real-time RT-PCR in cells after MBNL1 depletion or modifi-
cation of the NUMA1 splicing by anti-sense oligonucleotides.
Further, NUMA1 does not change expression between MCF7 and
MCF10 cells (Allo et al. 2014). Although the NUMA1 locus was re-
lated before to breast cancer risk (Kammerer et al. 2005), a clear
mechanism explaining its relevance in cancer is still lacking.
NUMA1 produces a protein component of the nuclear matrix,
which is dependent on threonine phosphorylation to regulate
the orientation of mitotic spindles and ensure symmetric cell

**Discussion**

Our study reveals that alterations in genes encoding RNA binding
proteins are pervasive in cancer, characterize the different tumor
types, and can explain many of the alternative splicing changes
observed in tumors. Many of the RBP expression changes appear
related to large copy number alterations. In contrast, mutations in
RBPs are particularly low in frequency compared to mutation
rates in other genes. We measured the potential relevance of RBP
alterations by looking at the associated splicing changes. Only a
few RBPs beyond the previously described cases show mutations
associated with genome-wide effects on alternative splicing. One
of the novel cases is HNRNPL, for which we predict that frequent
indels in an RNA binding domain would affect the splicing of
CASP9, a gene involved in programmed cell death (Yang et al.
2008). As HNRNPL also controls CASP9 alternative splicing
(Goehe et al. 2010), this suggests a relevant role of HNRNPL in ap-
optosis. On the other hand, the expression changes in RBP genes
appear as major contributors of the alternative splicing changes
observed in tumors, since the number of events affected by RBP
mutations is lower than those changing between tumor and nor-
mal samples. The splicing changes detected are predicted to have an
impact on many cancer hallmarks, some of them different from
those linked to gene expression changes. This agrees with
previous reports on the impact of splicing on cancer hallmarks
(Oltean and Bates 2013) and highlights the relevance of alternative
splicing as a complementary molecular mechanism to explain tu-
mor development. An important implication of our study for prog-
nostic and clinical studies is that the definition of the functional
impact of somatic genetic and epigenetic alterations should be ex-
panded to include changes in the alternative splicing of the genes in
cancer-relevant pathways.

We identified many RBPs whose expression alteration has
some potential relevance in cancer. For instance, TRA2B is fre-
quently amplified and up-regulated in LUSC and its binding motif
enriched in differentially spliced events, including an SE event in
the DNA damage-response gene CHEK1, reported recently to be
controlled by TRA2 proteins (Best et al. 2014). Despite the similar
genetic alterations between squamous tumors (Hoadley et al.
2014), only LUSC shows overexpression of TRA2B, which could
explain some of the differences in splicing patterns between these
two tumor types. We also found that genes from the MBNL family
are frequently down-regulated in tumors and their binding motif is
enriched in differentially spliced events, especially in breast and
prostate tumors. Additionally, splicing changes related to MBNL1
recapitulate the splicing patterns of undifferentiated cells, in agree-
ment with recent studies describing MBNL1 and MBNL2 as regula-
 tors of a stem cell-related splicing program (Han et al. 2013;
Venables et al. 2013). MBNL1 potentially controls multiple genes
that participate in cancer-related pathways, including the mitotic
gene NUMA1. We described that NUMA1 alternative splicing leads
to higher proliferation and increased centrosome amplification in
normal cells. Notably, the alternative splicing change is not asso-
ciated with an expression change in NUMA1, as measured from
the tumor and normal samples and by semi-quantitative and quan-
titative real-time RT-PCR in cells after MBNL1 depletion or modifi-
cation of the NUMA1 splicing by anti-sense oligonucleotides.
Further, NUMA1 does not change expression between MCF7 and
MCF10 cells (Allo et al. 2014). Although the NUMA1 locus was re-
lated before to breast cancer risk (Kammerer et al. 2005), a clear
mechanism explaining its relevance in cancer is still lacking.
NUMA1 produces a protein component of the nuclear matrix,
which is dependent on threonine phosphorylation to regulate
the orientation of mitotic spindles and ensure symmetric cell
Cancer splicing networks

**Figure 6.** Regulation of NUMA1 alternative splicing by MBNL1 in breast luminal tumors. (A) PSI value distributions in tumor and paired normal sample for luminal A (LA) and luminal B (LB) breast tumors for the events in NUMA1 (LA: ΔPSI = −0.22, P-value = 7.81 × 10^{-7}, LB: ΔPSI = −0.23, P-value = 0.037) and NUMB (LA: ΔPSI = 0.28, P-value = 0.0001, LB: ΔPSI = 0.28, P-value = 0.016). All P-values given are corrected for multiple testing. (B) Semiquantitative RT-PCR isoform analysis upon knockdowns of MBNL1 (lanes 2–10) and QKI (lanes 12–14) at different siRNA concentrations and the controls with scrambled siRNAs (lanes 1, 11). The diagrams to the right indicate the position of the alternatively spliced exons. The bottom lanes correspond to the RT-PCR amplification of RNA from the PUM gene in the same samples, which are used as a control for RNA expression. (C) Resazurin-based assays of cell viability/proliferation. Measurements were performed in triplicate at 72, 96, and 120 h. The plot shows measurements upon knockdowns of MBNL1 (siMBNL1) and QKI (siQKI), upon transfection of AONs targeting the 3′ and 5′ splice-sites independently and both together, and the corresponding controls (scrambled siRNA and random AON). (D, left panel) Graph showing the results of the evaluation of centrosome amplification upon knockdown of MBNL1 (siMBNL1) or upon transfection of AONs targeting 5′ splice-sites (5′ ss AON), compared to the corresponding controls siScrambled (P = 0.4271) and random AON (P = 0.04356), respectively (one-sided Mann-Whitney U test). (Right panels) Representative merged (TUBG1 and DAPI) images of immunofluorescence assays. (E) Correlation of NUMA1 event PSI (x-axis) with the CIN25 signature of aneuploidy (y-axis) across the tumor (red) and normal (blue) samples for luminal A (upper panel) (R = −0.4, Spearman) and B (lower panel) (R = −0.33, Spearman).

division (Kotak et al. 2013). NUMA1 alternative splicing removes a potential threonine phosphorylation site; hence, one attractive hypothesis is that this splicing change affects NUMA1 phosphorylation, thereby impairing correct spindle positioning, leading to increased genome instability. Besides MBNL1, our analyses provided evidence for the control of NUMA1 splicing by RBM42, in agreement with its proposed role in cell cycle (Suvorova et al. 2013) and pointing to an involvement of this factor in alternative splicing and cancer worth investigating further.

The origin of the expression changes in RBPs remains to be described. In particular, those cases that cannot be explained by DNA alterations in the gene locus may originate from alterations in the pathways that control the regulation of RBPs. It was recently postulated that RBPs involved in development and differentiation controlled by common enhancers could act as master regulators (Jangi and Sharp 2014). These included MBNL1, RBFOX2, RBM24, RBM38, RBM20, RBFOX1, ZNF638, and RBMS3, which we found frequently down-regulated in tumors, with enriched motifs in differentially spliced events and potential targets in multiple cancer drivers. This suggests a general mechanism in tumors by which the de-activation of one or more RBPs through the alteration of a common enhancer would lead to the reversal of multiple RNA splicing patterns to trigger or sustain an undifferentiated phenotype. Taken together, our results provide a rich resource of information about novel networks of RBPs that trigger common and specific alternative splicing changes in several solid tumors that may be relevant to understand the molecular basis of, and potentially reverse, the oncogenic properties of tumor cells.

**Methods**

**Data sets**

Data sets were downloaded from the TCGA data portal (https://tcga-data.nci.nih.gov/tcga/) (Supplemental Table S1). Details are provided as Supplemental Material. The 1348 genes coding for RNA binding proteins analyzed include those with high confidence for RNA binding (Balz et al. 2012; Castello et al. 2012; Kwon et al. 2013) and those annotated as RNA-binding in Ensembl (Cunningham et al. 2014; Supplemental Table S2). From this set, a subset of 162 known and potential auxiliary splicing factors was selected.

**Differential expression**

Quantile normalization and voom (Law et al. 2014) transformation were performed on gene read counts. Differential expression was analyzed using the limma package (Smyth 2005), and P-values were corrected for multiple testing using the Benjamini-Hochberg method. Genes were considered differentially expressed if the |log2-fold change| > 0.5 in absolute value and the corrected P-value
<0.05. An expression Z-score per gene and per tumor sample was calculated using the quantile-normalized and voom-transformed read-counts.

Mutation and copy number variation analysis
The frequency of somatic mutations across all samples with available data was calculated per gene and tumor type. The association between expression regulation and mutations was measured using a Jaccard index. Mutual exclusion was measured using the number of samples having an RBP mutation and no expression change and the number of samples having expression change but no RBP mutation. For samples with CNV data available, the overlap of each RBP with the annotated CNVs was calculated, requiring a CNV score > \log_2(3) or < \log_2(1) for gain or loss, respectively. We required that the full locus of the RBP fall within a copy number region and defined focal CNVs to be those smaller than 5 Mb. Using a Jaccard index, an association was calculated between the expression up- or down-regulation and CNV gain or loss, respectively. More details are provided as Supplemental Material.

Alternative splicing events
A total of 30,820 alternative splicing events were calculated from the gene annotation using SUPPA (Alamancos et al. 2015): 16,232 exon skipping events, 4978 alternative 5’ splice-site events, 6336 alternative 3’ splice-site events, 1478 mutually exclusive exon events, and 1787 retained intron events. Differentially spliced events were obtained by comparing the event percent spliced-in value distributions between normal and tumor samples using a Wilcoxon signed rank test, removing samples with missing PSI values, using at least 10 paired samples, and correcting for multiple testing with the Benjamini-Hochberg method. Events were considered differentially spliced if the difference between the tumor and normal median PSIs > 0.1 in absolute value and a corrected P-value < 0.05. Events associated with mutations in RBPs were calculated in a similar way separating samples according to whether they had or did not have mutations in an RBP. Tests were performed using protein-affecting mutations. Only RBPs with mutations in at least 10 samples were tested. An enrichment Z-score was calculated per RBP and tumor type by comparing the number of events changing significantly, with the median value obtained using all RBPs tested. Tumor-type–specific alternative splicing events were calculated by comparing their PSI values for pairs of tumor types using an information theoretic measure. More details of the analyses are given as Supplemental Material.

Gene sets
Annotations for 50 cancer hallmarks were obtained from the Molecular Signatures Database v4.0 (Libерzon et al. 2015), and a Fisher’s exact test was performed per hallmark using genes with annotated events and genes with differentially spliced events in each tumor type. The lists of 82 genes whose alternative splicing was linked before to cancer (Supplemental Table S5) and of 889 cancer drivers based on mutations and CNVs (34 in common with the previous set) (Supplemental Table S6) were collected from the literature (more details in the Supplemental Material).

RNA binding motif enrichment
Details on how motifs from RNAcompete (Ray et al. 2013) were assigned to the RBP genes are given in the Supplemental Material. The tool FIMO (Bailey and Elkan 1994) was used to scan the motifs in the event regions using a P-value < 0.001. Motif enrichment analysis was performed by comparing the frequency of regions in differentially spliced events with a motif with 100 random subsamples of the same size from equivalent regions in nondifferentially spliced events controlling for similar G+C content. An enrichment Z-score per motif, region, and direction of change (ΔPSI > 0.1, ΔPSI < –0.1) was calculated from the observed frequency and the 100 random control sets. A differentially spliced event was considered to be a potential target of a differentially expressed RBP if the correlation between the event PSI value and the gene expression robust Z-score was |Z| > 0.5 (Spearman) and the event contained the RBP binding motif. To assess significance, the same number of differentially spliced events in a tumor type was randomly selected from all events 100 times and events associated with the RBPs calculated each time as described previously. A Z-score was calculated from the mean and standard deviation. Cases with Z-score > 1.96 were considered significant.

Networks of RBPs and events
Networks of RBPs and events were built using the correlations between RBPs through events. For each pair of RBPs, a correlation was calculated using the Spearman correlation values with all differentially spliced events in the same tumor type. RBP clusters were built by calculating an inverse covariance matrix of these correlations using the glasso algorithm (Friedman et al. 2008) and then searching for dense, highly connected subgraphs with a greedy algorithm (Clauset et al. 2004). Events were associated with a network if they had |Z| > 0.8 (Spearman) or |Z| > 0.5 plus motif for any of the RBPs in an RBP cluster.

Experimental procedures
Details on the cell cultures, siRNA transfections, Western blot analyses, and semiquantitative RT-PCR experiments are described in the Supplemental Material.

Antisense oligonucleotides treatment
2′-O-Methyl RNA oligonucleotides were designed with full phosphorothioate linkage, antisense to the 5′ or 3′ splice sites of NUMA1 alternative exon 16 (hg19 coordinates Chr11: 71723447–71723488) optimizing GC content to 45%–60%. Custom modified and HPLC-purified RNA oligos were ordered in a 0.2 µM scale from Sigma-Aldrich:

- NUMA1_ex16_5′ss: 5′-ggcagauacucugcuauuguugc-3′
- NUMA1_ex16_3′ss: 5′-ccucuaagcugcuaccaccguu-3′
- RANDOM 2′-O-Methyl RNA oligo: 5′-GCAauggttccagaaggugUCAugugugucG-3′

Antisense RNA oligos were transfected in triplicate at 20-nM final concentration using 2 µL Lipofectamine RNAiMax (Life Technologies, 13778150) per 1 mL of total volume of transfection in OPTIMEM (Life Technologies, 13778150). After 5 h of treatment, media was replaced by DMEM-F12 containing 10% FBS and Pen-Strep.

Cell proliferation/viability assay
2500 MCF10A cells/well were seeded the night before treatment in 96-well plates (NUNC, 167008) in 100 µL complete DMEM-F12 medium. Wells with none, half, or double the amount of cells were also seeded for fluorescence calibration. Cells were transfected with siRNA or AON oligos as described. Resazurin (Sigma, R7017) treatment was performed 72, 96, and 120 h after transfection, in seven replicates and incubated for 4 h in a 37°C incubator. Fluorescence was measured after 4 h of incubation, using a TECAN infinite m200 device with 530-nm excitation wavelength, 590-nm
emission wavelength, 30-nm emission bandwidth, and set to optimal gain. The medium was replaced by complete DMEM-F12 after measurements.

Centrosome count and aneuploidy signature
The number of centrosomes was determined by immunofluorescence assays using an anti-gamma-tubulin (TUBG1) antibody (clone GTU-88, Sigma-Aldrich; dilution 1:1000). The expected immunostaining pattern of this centrosomal marker in normal cells is one or two foci proximal to the nucleus. The cells were fixed in cold methanol for 10 min and washed in phosphate-buffered saline. The secondary antibody was Alexa Fluor 488 (Molecular Probes, Life Technologies), and the cells were mounted using VECTASHIELD with DAPI. The results correspond to at least five independent fields and >200 cells analyzed. The significance of the results was assessed using the one-sided Mann-Whitney U test (Supplemental Table S21). The chromosome instability signature CIN25 (Carter et al. 2006) was used by calculating the mean value of the normalized expression robust Z-score values for the 25 genes from the signature in each sample.

Acknowledgments
We thank P. Papasakas, B. Blencowe, M. Irinina, and Q. Morris for comments and discussions. E.S., B.S., A.P., and E.E. were supported by the Ministerio de Economía y Competitividad (MINECO) and European Commission (FEDER) (BIO2014-52566-R), Consolider RNAREG (CSD2009-00080), by Agencia de Gestión d’Ajusts Universitaris i de Recerca (AGAUR) (SGR2014-1121), and by the Sandra Ibarra Foundation for Cancer (FIS2013). J.V. and B.M. were supported by Fundación Botín, by Banco de Santander through its Santander Universities Global Division, and by Consolider RNAREG (CSD2009-00080), MINECO, and AGAUR. F.M. and M.A.P. were supported by AECC (Hereditary Cancer), AGAUR (SGR2014-364), the Instituto de Salud Carlos III (ISCIII), the MINECO, and FEDER (PIE13/00022-ONCOPROFILE, P11/ 00854, and RTICC RD12/0036/0008).

References
Alamancos GP, Pagès A, Trincado JL, Bellora N, Eysa E. 2015. Leveraging transcript quantification for fast computation of alternative splicing profiles. RNA 21: 1521–1531.
Allo M, Agirre E, Bessonov L, Bertucci P, Gomez Acuna L, Buggiano V, Bessonov S, Bertucci P, Gomez Acuna L, Buggiano V, Allo M, Agirre E, Bessonov S, Bertucci P, Gomez Acuna L, Buggiano V, Allo M, Agirre E, Bessonov S, Bertucci P, Gomez Acuna L, Buggiano V, Allo M, Agirre E, Bessonov S, Bertucci P, Gomez Acuna L, Buggiano V, Allo M, Agirre E, Bessonov S, Bertucci P, Gomez Acuna L, Buggiano V.

The Cancer Genome Atlas Network. 2012b. Comprehensive molecular characterization of squamous cell lung cancers. Nature 489: 519–525.

The Cancer Genome Atlas Network. 2012b. Comprehensive molecular portrait of human breast tumours. Nature 490: 61–70.

The Cancer Genome Atlas Network. 2015. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 517: 576–582.

Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z. 2006. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. Nat Genet 38: 1043–1048.

Cancer Genome Atlas Network. 2012b. Comprehensive molecular characterization of squamous cell lung cancers. Nature 489: 519–525.

The Cancer Genome Atlas Network. 2012. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 517: 576–582.

Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z. 2006. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. Nat Genet 38: 1043–1048.

Cancer Genome Atlas Network. 2012b. Comprehensive molecular characterization of squamous cell lung cancers. Nature 489: 519–525.

The Cancer Genome Atlas Network. 2012. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 517: 576–582.
Kotak S, Busso C, Gönçzy P. 2013. NuMA phosphorylation by CDK1 couples mitotic progression with cortical dynein function. *EMBO J* **32**: 2517–2529.

Kwon SC, Yi H, Eichelaum K, Führ S, Fischer B, You KT, Castello A, Kriegweld J, Hentze MW, Kim VN. 2013. The RNA-binding protein repertoire of embryonic stem cells. *Nat Struct Mol Biol* **20**: 1122–1130.

Law CW, Chen Y, Shi W, Smyth GK. 2014. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**: R29.

Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. 2015. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**: 417–425.

Maguire SL, Leonidou A, Wai P, Marchio C, Ng CK, Sapino A, Salomon A-V, Reis-Filho JS, Weigelt B, Natrajan RC. 2015. *SF3B1* mutations constitute a novel therapeutic target in breast cancer. *J Pathol* **235**: 571–580.

Misquitta-Ali CM, Cheng E, O’Hanlon D, Liu N, McGlade CJ, Tsao MS, Blencowe BJ. 2011. Global profiling and molecular characterization of alternative splicing events misregulated in lung cancer. *Mol Cell Biol* **31**: 138–150.

Moon H, Cho S, Loh TJ, Oh HK, Jang HN, Zhou J, Kwon Y-S, Liao DJ, Jun Y, Eom S, et al. 2014. SRSF2 promotes splicing and transcription of exon 11 included isoform in Ron proto-oncogene. *Biochim Biophys Acta* **1839**: 1132–1140.

Oltean S, Bates DO. 2013. Hallmarks of alternative splicing in cancer. *Oncogene* **1–8.

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

Sebestyén E, Zawisza M, Eyras E. 2015. Detection of recurrent alternative splicing switches in tumor samples reveals novel signatures of cancer. *Nucleic Acids Res* **43**: 1345–1356.

Smyth GK. 2005. limma: Linear models for microarray data. In *Bioinformatics and computational biology solutions using R and Bioconductor* (ed. Gentleman R, et al.), pp. 397–420. Springer, New York.

Suvorova ES, Croken M, Kratzer S, Ting LM, de Felipe MC, Balu B, Markillie ML, Weiss LM, Kim K, White MW. 2013. Discovery of a splicing regulator required for cell cycle progression. *PLoS Genet* **9**: e1003505.

Suzuki K. 2008. The multi-functional serpin, protein C inhibitor: beyond thrombosis and hemostasis. *J Thromb Haemost* **6**: 2017–2026.

Vanharanta S, Marney CR, Shu W, Valiente M, Zou Y, Mele A, Darnell RB, Massagué J. 2014. Loss of the multifunctional RNA-binding protein RBM47 as a source of selectable metastatic traits in breast cancer. *eLife* **3**: e02734.

Venables JP, Lapasset L, Gadea G, Fort P, Klinc R, Irimia M, Vignal E, Thibault P, Prinos P, Chabot B, et al. 2013. MBNL1 and RBFOX2 cooperate to establish a splicing programme involved in pluripotent stem cell differentiation. *Nat Commun* **4**: 2480.

Wang Y, Chen D, Qian H, Tsai YS, Shao S, Liu Q, Dominguez D, Wang Z. 2014. The splicing factor RBM4 controls apoptosis, proliferation, and migration to suppress tumor progression. *Cancer Cell* **26**: 374–389.

Watermann DO, Tang Y, Zur Hausen A, Jager M, Stamm S, Stickeler E. 2006. Splicing factor Tra2-b1 is specifically induced in breast cancer and regulates alternative splicing of the CD44 gene. *Cancer Res* **66**: 4774–4780.

Xiao R, Sun Y, Ding J-H, Lin S, Rose DW, Rosenfield MG, Fu X-D, Li X. 2007. Splicing regulator SC35 is essential for genomic stability and cell proliferation during mammalian organogenesis. *Mol Cell Biol* **27**: 5393–5402.

Xue Y, Liu Z, Cao J, Ma Q, Gao X, Wang Q, Jin C, Zhou Y, Wen L, Ren J. 2011. GPS 2.1: enhanced prediction of kinase-specific phosphorylation sites with an algorithm of motif length selection. *Protein Eng Des Sel* **24**: 255–260.

Yang M, Sun T, Wang L, Yu D, Zhang X, Miao X, Liu J, Zhao D, Li H, Tan W, et al. 2008. Functional variants in cell death pathway genes and risk of pancreatic cancer. *Clin Cancer Res* **14**: 3230–3236.

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.

Zheng Z, Zhu H, Wan Q, Liu J, Xiao Z, Siderovski DP, Du Q. 2010. LGN regulates mitotic spindle orientation during epithelial morphogenesis. *J Cell Biol* **189**: 275–288.

Zong F-Y, Fu X, Xu X, Wei W-J, Luo Y-G, Heiner M, Cao L-J, Fang Z, Fang R, Lu D, Ji H, et al. 2014. The RNA-binding protein QKI suppresses cancer-associated aberrant splicing. *PLoS Genet* **10**: e1004289.

Received September 20, 2015; accepted in revised form April 11, 2016.