UCLA
UCLA Previously Published Works

Title

Coregulation of alternative splicing by hnRNPM and ESRP1 during EMT

Permalink

https://escholarship.org/uc/item/15t262v9

Journal

RNA, 24(10)

ISSN

1355-8382

Authors

Harvey, Samuel E
Xu, Yilin
Lin, Xiaodan
et al.

Publication Date

2018-10-01

DOI

10.1261/rna.066712.118

Peer reviewed

Coregulation of alternative splicing by hnRNPM and ESRP1 during EMT

SAMUEL E. HARVEY,[1,2] YILIN XU,[2] XIAODAN LIN,[1,2] XIN D. GAO,[2] YUSHAN QIU,[1,2] JAEGYOON AHN,[3,4] XINSHU XIAO,[3] and CHONGHUI CHENG[1,2]

[1]Lester and Sue Smith Breast Center, Department of Molecular and Human Genetics, Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas 77030, USA

[2]Division of Hematology/Oncology, Robert H. Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60611, USA

[3]Department of Integrative Biology and Physiology and the Molecular Biology Institute, University of California Los Angeles, Los Angeles, California 90095, USA

ABSTRACT

The epithelial–mesenchymal transition (EMT) is a fundamental developmental process that is abnormally activated in cancer metastasis. Dynamic changes in alternative splicing occur during EMT. ESRP1 and hnRNPM are splicing regulators that promote an epithelial splicing program and a mesenchymal splicing program, respectively. The functional relationships between these splicing factors in the genome scale remain elusive. Comparing alternative splicing targets of hnRNPM and ESRP1 revealed that they coregulate a set of cassette exon events, with the majority showing discordant splicing regulation. Discordant splicing events regulated by hnRNPM show a positive correlation with splicing during EMT; however, concordant events do not, indicating the role of hnRNPM in regulating alternative splicing during EMT is more complex than previously understood. Motif enrichment analysis near hnRNPM–ESRP1 coregulated exons identifies guanine–uridine rich motifs downstream from hnRNPM-repressed and ESRP1-enhanced exons, supporting a general model of competitive binding to these cis-elements to antagonize alternative splicing. The set of coregulated exons are enriched in genes associated with cell migration and cytoskeletal reorganization, which are pathways associated with EMT. Splicing levels of coregulated exons are associated with breast cancer patient survival and correlate with gene sets involved in EMT and breast cancer subtyping. This study identifies complex modes of interaction between hnRNPM and ESRP1 in regulation of splicing in disease-relevant contexts.

Keywords: RNA binding proteins; hnRNPM; ESRP1; alternative splicing; EMT

INTRODUCTION

Alternative RNA splicing is a fundamental mechanism of functional genome diversity that enables the nearly 21,000 protein-coding genes in the human genome to give rise to over 100,000 transcripts (The Encode Consortium 2012; Harrow et al. 2012). Deep transcriptome sequencing has revealed that ~95% of all human multiexon transcripts can undergo alternative splicing, positioning alternative splicing as a critical form of post-transcriptional gene regulation in a variety of cellular and biological processes (Pan et al. 2008; Wang et al. 2008; Barash et al. 2010). Dysregulation of alternative splicing is increasingly implicated in a variety of human diseases, including cancer progression and survival (Liu and Cheng 2013; Cieply and Carstens 2015).

Alternative splicing has emerged as a central regulatory process during the epithelial–mesenchymal transition (EMT) (Warzecha et al. 2010; Brown et al. 2011; Shapiro et al. 2011; Reinke et al. 2012; Yang et al. 2016). EMT is a developmental program whereby epithelial cells transit to a mesenchymal phenotype, which occurs in natural processes such as organogenesis and wound healing (Thiery 2003; Nieto et al. 2016). A mounting body of evidence suggests that EMT is aberrantly activated in cancer cells to mediate tumor recurrence and metastasis (Yang and Weinberg 2008; Thiery et al. 2009). Study of the molecular mechanism of EMT has been largely restricted to cellular signaling and transcriptional regulation. Recently, work from our group has demonstrated that alternative splicing

of the gene *CD44* causally contributes to EMT and breast cancer metastasis (Brown et al. 2011; Reinke et al. 2012; Xu et al. 2014; Zhao et al. 2016). Further evidence has also emerged to show the essential role of alternative splicing of other genes in controlling EMT (Lu et al. 2013; Hernandez et al. 2015). A variety of splicing regulatory proteins have also been implicated in EMT alternative splicing; however, few have been shown to have essential functional roles during EMT (Warzecha et al. 2010; Braeutigam et al. 2013; Xu et al. 2014; Yang et al. 2016).

Investigating the mechanisms underlying the regulation of *CD44* alternative splicing led us to identify antagonistic roles between two splicing factors, heterogeneous nuclear ribonucleoprotein M (hnRNPM) and epithelial splicing regulatory protein 1 (ESRP1). hnRNPM promotes *CD44* variable exon skipping and favors a mesenchymal phenotype, whereas ESRP1 stimulates *CD44* variable exon inclusion and promotes an epithelial cellular state (Warzecha et al. 2009b; Brown et al. 2011; Reinke et al. 2012; Xu et al. 2014). In addition, a recent study showed that hnRNPM promotes a set of alternative splicing events associated with cell survival and resistance to inhibition of the PI3K-Akt pathway, two traits that are associated with EMT (Grille et al. 2003; Passacantilli et al. 2017). Interestingly, hnRNPM is ubiquitously expressed but functions in a mesenchymal cell-state–specific manner to regulate *CD44* alternative splicing. This cell-state restricted activity of hnRNPM is guided in part by competition with ESRP1, which is expressed in epithelial tissues but not in mesenchymal tissues (Warzecha et al. 2009a; Xu et al. 2014; Bebee et al. 2015). hnRNPM and ESRP1 share common guanine–uridine-rich (GU-rich) binding sites (Dittmar et al. 2012; Huelga et al. 2012), and the presence of ESRP1 suppresses the activity of hnRNPM by binding to the same GU-rich *cis*-elements near *CD44* variable exons (Xu et al. 2014). Given the critical roles hnRNPM and ESRP1 play in modulating EMT, we hypothesized that these two splicing factors compete to regulate not only *CD44* alternative splicing, but also many other splicing events which may be associated with EMT. The balance between hnRNPM and ESRP1 splicing regulation may therefore control the phenotypic switch between an epithelial state and a mesenchymal state.

In this study, we analyzed splicing events coregulated by both hnRNPM and ESRP1. Our results show that hnRNPM and ESRP1 exhibit inverse activities in regulating most coregulated splicing events. Unexpectedly, they also display concordant activities when regulating a subset of coregulated splicing events. Importantly, our results reveal that hnRNPM and ESRP1 regulate a set of cassette exons to promote and inhibit EMT, respectively. Cassette exons regulated discordantly by hnRNPM and ESRP1 are enriched in GU-rich motifs specifically in the downstream intron, corresponding with known hnRNPM and ESRP1

binding motifs and likely sites of competitive splicing regulation. This competitive mode of regulation is more widespread than previously appreciated. Coregulated cassette exons also stratify breast cancer patients by overall survival and correlate with cancer-relevant gene sets, highlighting the importance of hnRNPM and ESRP1 splicing regulation in cancer biology.

RESULTS

Splicing factors ESRP1 and hnRNPM coregulate a set of cassette exons

We previously showed that ESRP1 antagonizes the splicing activity of hnRNPM on *CD44* alternative splicing and that hnRNPM stimulates EMT-associated splicing in a mesenchymal cell-type–specific manner (Xu et al. 2014). In an effort to better understand how ESRP1 and hnRNPM functionally interact with each other globally, we compared alternative splicing events from transcriptome-profiling data sets in response to hnRNPM or ESRP1 perturbation. We performed differential alternative splicing analysis when hnRNPM was perturbed by shRNA in two well-established mesenchymal cell lines: MDA-MB-231-derived lung-metastatic LM2 and bone-metastatic BM1 (Kang et al. 2003; Minn et al. 2005). We chose mesenchymal cell lines to investigate hnRNPM-mediated alternative splicing to avoid masking of hnRNPM function by ESRP1, which is not expressed in mesenchymal tissues (Warzecha et al. 2009a; Bebee et al. 2015). We focused on cassette exons because they are the most common form of alternative splicing (Wang et al. 2008). Cassette exon splicing levels are reported using the percent spliced in (PSI) metric, which is a measure of the relative abundance of the exon inclusion isoform. We obtained a set of 1635 hnRNPM-regulated alternative cassette exons by taking the union of all significantly regulated exons (FDR < 0.05, $|\Delta PSI| \geq 0.1$, average junction reads per cassette event ≥ 10) after hnRNPM knockdown in LM2 and BM1 cells. Using the union of cassette exons identified from previously published alternative splicing profiling data sets for ESRP1, including ESRP1 overexpression in MDA-MB-231 cells and ESRP1/2 knockdown in prostate epithelial PNT2 and lung non-small cell carcinoma H358 cells (Warzecha et al. 2010; Dittmar et al. 2012; Yang et al. 2016), we derived a corresponding set of 1300 ESRP1-regulated cassettes. A full accounting of all data sets with corresponding replicates, splicing factor perturbations, cell types, and the number of cassette exons detected in each data set is listed in Supplemental Table S1. Although a variety of cell lines were considered in our analysis, out of the 1635 cassette exons significantly regulated by hnRNPM, 1548 (95%) were detected in a background set of all exons expressed in the ESRP1 data sets. Similarly, out of the 1300 cassette exons regulated by ESRP1, 1124 (86%) were detected

in a background set of exons from the hnRNPM data sets. These results show that the combined data sets, although encompassing different cell types, contained the vast majority of assayed cassettes and were suitable for comparative analysis.

To determine cassette exons regulated by both hnRNPM and ESRP1, we intersected the hnRNPM-regulated cassettes with ESRP1-regulated cassettes and obtained a statistically significant overlap of 213 coregulated cassette exons (Fig. 1A, $P = 7.9 \times 10^{-115}$, hypergeometric test; Supplemental Table S2). Cassette exons where hnRNPM and ESRP1 both promote exon inclusion or exon skipping were defined as concordant exons. Cassette exons where hnRNPM and ESRP1 promote opposite splice isoforms were defined as discordant exons. Nearly two-thirds of the coregulated exons (134/213, 63%) were regulated discordantly, while the remaining exons (79/213, 37%) were regulated concordantly (Fig. 1A,B). The fact that the majority of coregulated events show discordant regulation mirrors the antagonistic role that ESRP1 and hnRNPM play in favoring cell-state–specific splicing programs (Xu et al. 2014). In contrast, the concordant splicing regulation by ESRP1 and hnRNPM suggests that they cooperate to control splicing within a subset of genes. Thus, the coregulation of splicing between ESRP1 and hnRNPM is not purely antagonistic and may be more complex than previously understood.

As hnRNPM and ESRP1 play important roles in regulating alternative splicing during EMT, we overlapped the coregulated exons with a set of EMT regulated alternative splicing events derived from previous studies (Shapiro et al. 2011; Yang et al. 2016). Over 30% (67/213) of the coregulated exons overlap with EMT, representing a statistically significant overlap (Fig. 1C, $P = 2.63 \times 10^{-80}$, hypergeometric test; Supplemental Table S2). We then analyzed the regulatory roles of hnRNPM and ESRP1 on
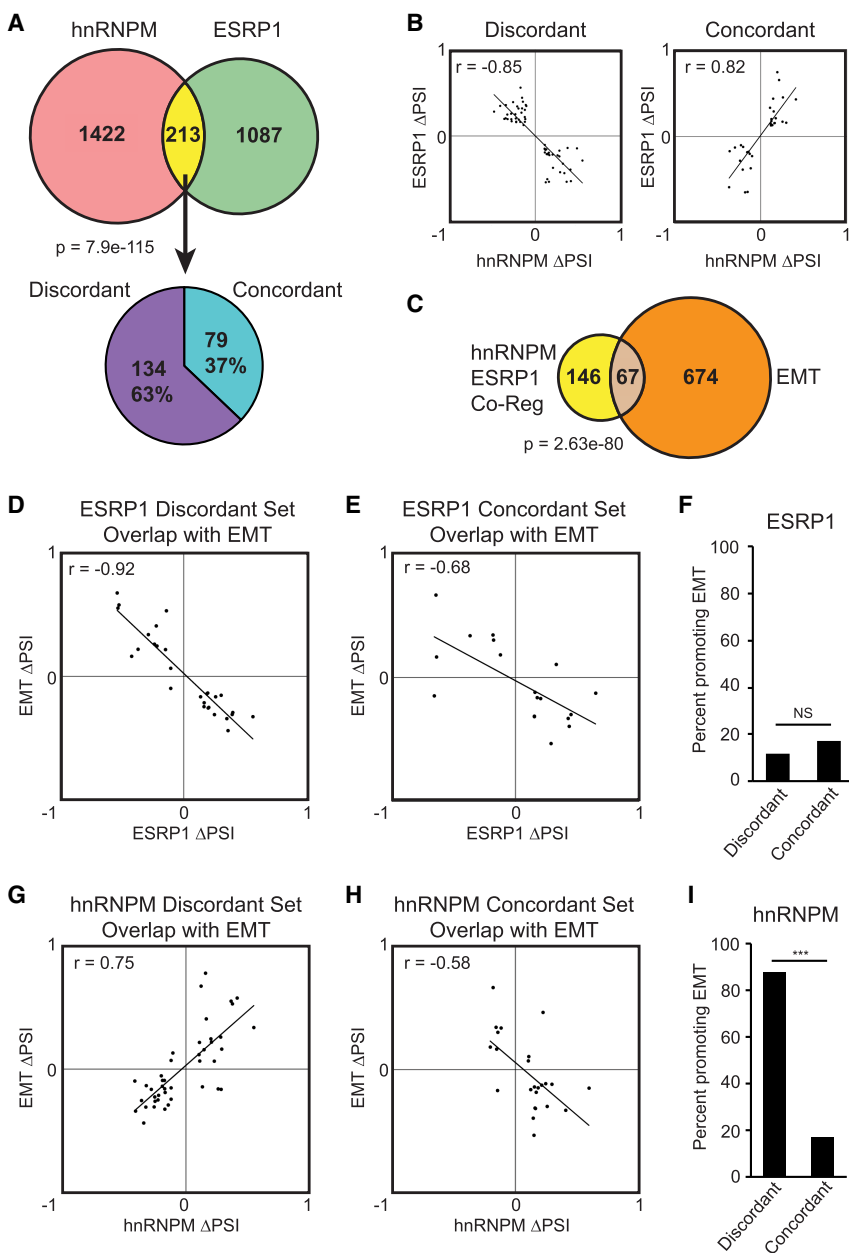

FIGURE 1. hnRNPM and ESRP1 coregulate a set of cassette exons in discordant and concordant manners. (A) RNA-sequencing analysis of hnRNPM knockdown and ESRP1 perturbation data sets identified splicing factor–dependent cassette exon events. In total, 213 cassette exon events showed overlapping regulation by hnRNPM and ESRP1 (P-value by hypergeometric test). The majority, 63% (134/213) cassette exons, were regulated discordantly by hnRNPM and ESRP1. The minority, 37% (79/213) cassette exons, were regulated concordantly. (B) Discordant exons show a negative correlation of hnRNPM versus ESRP1 ΔPSI splicing changes, while concordant exons show a positive correlation. Positive ΔPSI indicates promotion of exon skipping, while negative ΔPSI indicates promotion of exon inclusion. (C) 67 of the hnRNPM-ESRP1 coregulated cassette exons overlap with cassette exons regulated during EMT (P-value by hypergeometric test). (D,E) Both ESRP1 discordantly regulated exons (D) and concordantly regulated exons (E) show a negative correlation with EMT splicing. Positive and negative ΔPSI indicate promotion of exon skipping and inclusion, respectively. (F) 12% of ESRP1-regulated discordant exons promote EMT compared to 17% of concordant exons. (G) hnRNPM discordantly regulated exons show a positive correlation with EMT. (H) hnRNPM concordant exons show a negative correlation with EMT. (I) 88% of hnRNPM discordant exons promote EMT compared to 17% of hnRNPM concordant exons (P < 0.001 by Fisher's exact test).

these EMT-associated splicing events. For both discordant and concordant exons coregulated by ESRP1 and hnRNPM, ESRP1-mediated splicing inversely correlated with the EMT-associated splicing, in line with the role of ESRP1 as an epithelial specific splicing regulator (Fig. 1D–F; Warzecha et al. 2009a, 2010; Bebee et al. 2015). Interestingly however, hnRNPM showed bidirectional correlation. For the hnRNPM–ESRP1 discordant exons, we observed a positive correlation between hnRNPM-mediated splicing and EMT-associated splicing, indicating that hnRNPM promotes splicing that occurs during EMT (Fig. 1G). For the hnRNPM–ESRP1 concordant exons, however, hnRNPM's activity inversely correlated with EMT-associated splicing (Fig. 1H). This bidirectional difference in EMT-splicing regulation was statistically significant (Fig. 1I). These results suggest that although the majority of hnRNPM-regulated events are consistent with its role in driving a mesenchymal splicing program in opposition to ESRP1, hnRNPM may also be involved in a small subset of splicing events regulated in favor of an epithelial splicing pattern when functioning in concert with ESRP1.

Validation of ESRP1 and hnRNPM coregulated cassette exons

We chose to validate hnRNPM splicing activity in the mesenchymal cell line LM2 that does not express ESRP1 (Fig. 2A), and we chose human mammary epithelial cells (HMLE) to validate ESRP1 activity where ESRP1 is robustly expressed (Fig. 2A). hnRNPM is expressed similarly in LM2 and HMLE cells (Fig. 2A). We experimentally validated four of the concordant and four of the discordant coregulated splicing events using RT-PCR upon shRNA-mediated hnRNPM or ESRP1 knockdown to confirm splicing regulation observed in the RNA sequencing studies (Fig. 2B,C; Supplemental Table S2). The validation set confirms that hnRNPM and ESRP1 coregulate splicing events in both concordant and discordant manners (Fig. 2C). Moreover, we examined the specificity of the hnRNPM and ESRP1 splicing regulatory relationships by using two splicing minigenes,

with one containing a discordantly regulated exon at *CD44* variable exon 5 and the other harboring a concordant exon at *MARK3* exon 17. Cotransfection experiments of the *CD44v5* minigene with hnRNPM or ESRP1 in 293FT cells showed that hnRNPM promotes v5 exon skipping, whereas ESRP1 inhibits it (Fig. 2D), supporting discordant regulation by hnRNPM and ESRP1. In contrast, cotransfection of the concordant *MARK3* exon 17 minigene with hnRNPM or ESRP1 both resulted in dose-dependent increases in exon skipping, mirroring the concordant regulation of
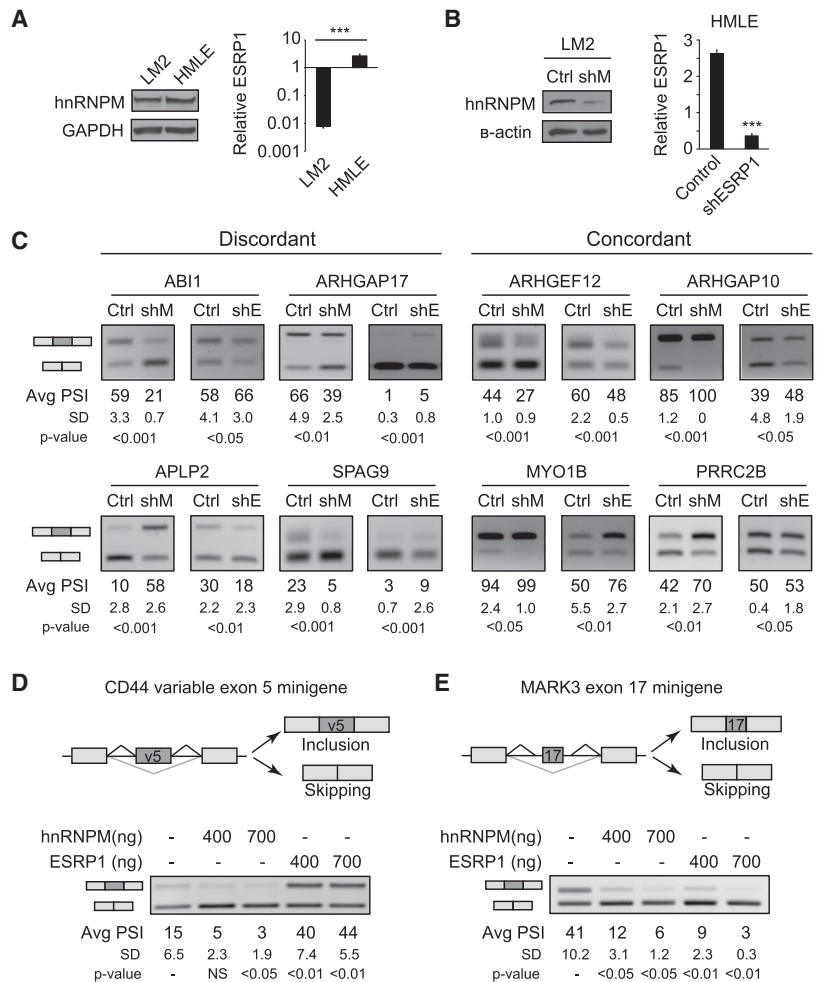

FIGURE 2. Validation of hnRNPM and ESRP1 coregulation of cassette exons. (*A*) Western blot showing expression of hnRNPM in LM2 and HMLE cells. qPCR showing expression of ESRP1 relative to TBP in LM2 and HMLE cells (Error bars = S.E.M, $n = 3$, $P < 0.001$). (*B*) Western blot showing hnRNPM knockdown in LM2 cells and qPCR showing ESRP1 knockdown relative to TBP in HMLE cells (error bars = S.E.M, $n = 3$, $P < 0.001$). (*C*) Cassette exon splicing events regulated by hnRNPM and ESRP1. hnRNPM validation was conducted in LM2 cells, while ESRP1 validation was conducted in HMLE cells. Avg PSI represents an average of three experiments. SD represents standard deviation of PSI values (*P*-value calculated by Student's *t*-test). (*D*) Splicing minigene analysis showing cotransfection of *CD44* variable exon 5 minigene with hnRNPM in 293FT cells promotes exon skipping, while ESRP1 promotes exon inclusion. (*E*) Cotransfection of *MARK3* exon 17 minigene and hnRNPM or ESRP1 in 293FT cells shows that both promote exon skipping. Avg PSI represents an average of three experiments. SD represents standard deviation of PSI values (*P*-value calculated by Student's *t*-test).

splicing observed in the RNA-sequencing data (Fig. 2E). These results show that hnRNPM and ESRP1 function in both discordant and concordant fashions that are dependent on splicing substrates.

ESRP1 and hnRNPM discordantly regulated exons are enriched in shared GU-rich binding sites

In order to better understand the functional relationship between hnRNPM and ESRP1 in coregulating alternative splicing, we performed motif enrichment analysis on the introns near all hnRNPM and ESRP1 coregulated splicing events (Fig. 3A). Both hnRNPM and ESRP1 are known to bind GU-rich cis-elements primarily in introns (Dittmar et al. 2012; Huelga et al. 2012; Bebee et al. 2015; Yang et al. 2016). We observed selective enrichment of multiple GU-rich hexamers downstream from hnRNPM-repressed and ESRP1-enhanced events, with both sets of events showing enrichment of two identical motifs, GUGUGU and GUGGUG, among the top three enriched motifs (Fig. 3A–C). The observation that GU-rich motif enrichment was observed downstream from exons regulated oppositely by hnRNPM and ESRP1 suggests that GU-rich motifs are enriched near discordantly regulated exons. These data support a model where hnRNPM and ESRP1 compete for shared binding sites directly downstream from cassette exons to regulate alternative splicing antagonistically.

We also noted significant enrichment of a UGCAUG motif downstream from hnRNPM-enhanced and ESRP1-repressed events (Fig. 3A). This sequence corresponds to the well-known binding motif of the RBFOX family of RNA binding proteins, of which only RBFOX2 is expressed in the cell lines used in this study (Jangi et al. 2014; Weyn-Vanhentenryck et al. 2014). In total, 67/213 (31%) of hnRNPM–ESRP1 coregulated splicing events contain a splice-site proximal UGCAUG motif within 250 nt of a splice site. The motif was more enriched downstream from the cassette exons investigated, with 45/213 (21%) containing a motif downstream, while 30/213 (14%) contained a motif upstream of the cassette exon.

To experimentally examine the binding relationships of hnRNPM and ESRP1, we analyzed their ability to bind to the GU-rich motifs downstream from the discordantly regulated APLP2 cassette exon 7, which was validated in Figure 2C. APLP2 contains multiple occurrences of GU-rich motifs identified from the motif analysis within 250 nt downstream from APLP2 cassette exon 7 (Fig. 3D). RNA pull-down assays were conducted in MDA-MB-231 cells, which do not express ESRP1, and MDA-MB-231 cells ectopically expressing HA-tagged ESRP1 (Fig. 3E). Both cell lines express hnRNPM equally (Fig. 3E). As predicted from the RNA-seq data and validation experiments, an increase in APLP2 exon 7 PSI was observed upon ESRP1 overexpression (Fig. 3F). In order to examine binding of hnRNPM and ESRP1 to the GU-rich motifs

downstream from APLP2 exon 7, we designed two 5′ biotinylated RNA probes: GU1 and GU2 (Fig. 3G, top panel). hnRNPM binds GU2 much more strongly than GU1, while ESRP1 bound GU1 and GU2 relatively equally (Fig. 3G, bottom panel). GU2 contains more GU-nucleotide stretches than GU1, and this may have contributed to the binding differences of hnRNPM on these probes. As both ESRP1 and hnRNPM do not have clearly defined binding sites, instead binding to degenerate GU-rich sequences, there could be variations in their affinity to different GU repeats. These binding activities are specific because binding of both proteins was abolished upon disruption of GU-rich motifs (Fig. 3G). To determine whether hnRNPM binding is decreased in the presence of ESRP1, we compared relative hnRNPM binding on the GU2 probe in parental MDA-MB-231 cells to that in MDA-MB-231 cells overexpressing ESRP1-HA. By comparing the amount of hnRNPM that was associated with the GU2 probe relative to input, we found a 70% reduction in hnRNPM binding in the ESRP1-HA-expressing MDA-MB-231 cells (Fig. 3H), suggesting that ESRP1 is capable of competing with hnRNPM for the same binding sites on the APLP2 pre-mRNA.

ESRP1 and hnRNPM coregulated exons are enriched in EMT processes and correlate with breast cancer signatures and patient survival

In order to better understand the relevance of the complex regulation of splicing by hnRNPM and ESRP1 to disease phenotypes, we performed gene ontology analysis using DAVID on the set of 213 coregulated exons (Huang da et al. 2009a,b). We observed significant GOTERMs associated with cell polarity, cell adhesion, and cytoskeletal dynamics (Fig. 4A), all processes that are critical for EMT (Supplemental Table S3). To assess the contribution of hnRNPM and ESRP1 coregulation of splicing in breast cancer patient samples, we mined the publicly available TCGA (The Cancer Genome Atlas) RNA-sequencing data for invasive breast carcinoma (BRCA) and calculated the PSI values for hnRNPM and ESRP1 coregulated splicing events. After stratifying the patients based on PSI levels for each exon via 2-means clustering, we observed alternative splicing events that are positively or negatively correlated with patient overall survival (FDR < 0.05, log rank test). We found that increased SPAG9 exon 24 inclusion levels predict poorer patient survival, while increased ZMYND8 exon 22 inclusion is associated with a better prognosis in breast cancer (Fig. 4B). SPAG9 exon 24 PSI increases during EMT, with exon inclusion stimulated by hnRNPM and inhibited by ESRP1 (Supplemental Table S2). Conversely, ZMYND8 exon 22 exon skipping is promoted by hnRNPM and antagonized by ESRP1 (Supplemental Table S2).

Intrigued by the role these exons played in predicting breast cancer patient survival, we performed gene set
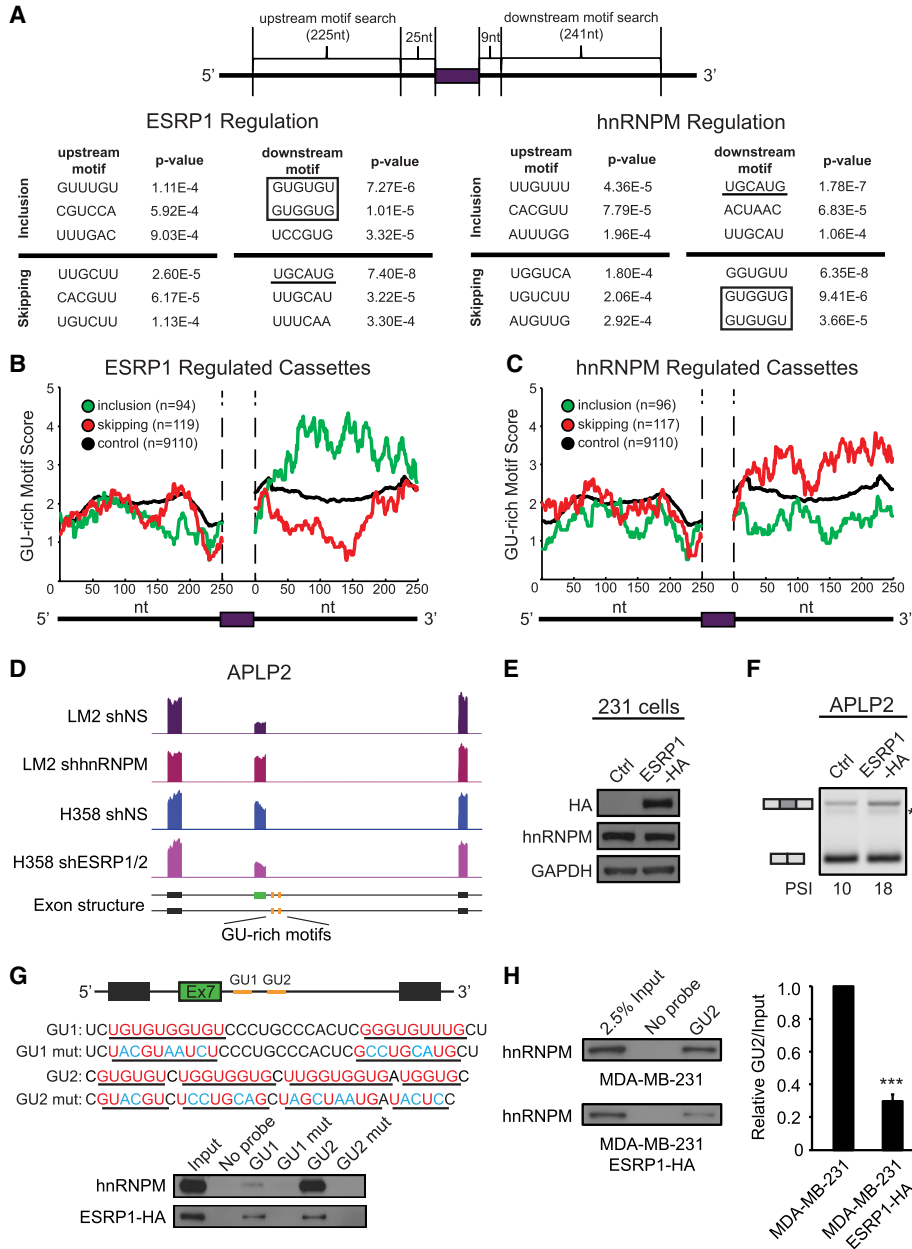
FIGURE 3. hnRNPM and ESRP1 show common motif enrichment downstream from discordantly regulated exons and compete for shared binding sites. (*A*) K-mer enrichment analysis showing the top three enriched 6-mers in introns flanking hnRNPM and ESRP1 coregulated cassette exons. Two identical GU-rich motifs (GUGUGU and GUGGUG, black box) were enriched downstream from cassette exons in ESRP1-enhanced and hnRNPM-repressed events. The RBFOX motif (UGCAUG, underlined) was enriched downstream from ESRP1-repressed and hnRNPM-enhanced events. (*B,C*) RNA motif map analysis of GU-rich motifs in introns flanking hnRNPM-ESRP1 coregulated exons with respect to ESRP1 regulation (*B*) and hnRNPM regulation (*C*) reveals enrichment of GU-rich motifs in downstream introns of ESRP1-enhanced and hnRNPM-repressed cassette exon splicing events. Inclusion events (green). Skipping events (red). Control events (black). (*D*) Genome browser plot of RNA sequencing data sets showing hnRNPM knockdown promotes *APLP2* exon 7 inclusion and ESRP1 depletion promotes skipping. Black bars indicate constitutive exons. Green bar indicates variable exon 7. Yellow bars indicate location of two clusters of GU-rich motifs within 250 nucleotides (nt) downstream from *APLP2* exon 7. (*E*) Immunoblot of HA-tagged ESRP1 overexpression in MDA-MB-231 cells and endogenous hnRNPM. (*F*) ESRP1-HA overexpression in MDA-MB-231 cells results in increased *APLP2* exon 7 inclusion. (*) Indicates nonspecific band. (*G*) (*Upper* panel) The two GU-rich regions identified within 250 nt downstream from *APLP2* exon 7 were used to design RNA probes GU1 and GU2 containing stretches of GU nucleotides underlined and in red with mutant probes GU1-mut and GU2-mut with mutated sequences colored in blue. (*Lower* panel) RNA pull-down analysis using RNA probes blotting for endogenous hnRNPM and overexpressed ESRP1-HA in MDA-MB-231 cells shows that hnRNPM and ESRP1 bind common GU-rich sequences. (*H*) RNA pull-down experiments using a static amount of the biotinylated GU2 RNA probe and cell lysate assaying for hnRNPM in the MDA-MB-231 cell line, which does not express ESRP1, and the same line with ESRP1-HA overexpression. A total of 2.5% input was provided as a loading control for both samples. Overexpression of ESRP1 leads to less hnRNPM binding, suggesting that ESRP1 competes for the same GU2 binding site. (Error bars = S.E.M, *n* = 3, [***] *P*-value < 0.001 by Student's *t*-test).
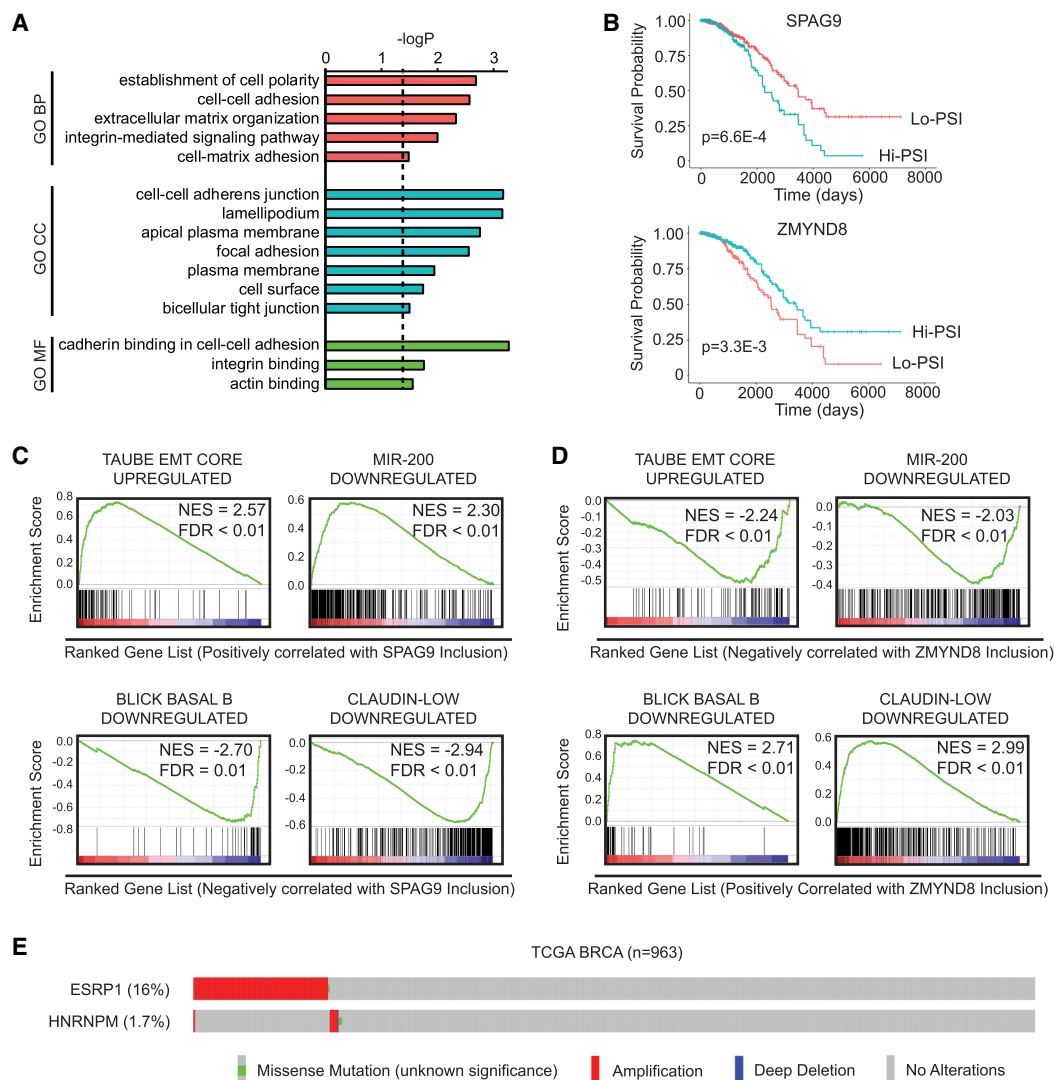
FIGURE 4. hnRNPM-ESRP1 coregulated exons are associated with EMT and breast cancer survival. (*A*) Gene ontology analysis of genes that contain hnRNPM-ESRP1 coregulated exons identified significant terms associated with cell polarity, adhesion, migration, and the cytoskeleton. Direct GOTERMS BP, MF, and CC were queried using DAVID. (*B*) PSI levels of *SPAG9* exon 24 and *ZMYND8* exon 22 stratify breast cancer patients by overall survival. (*C*) Genes positively correlated with *SPAG9* exon 24 inclusion are up-regulated during EMT and down-regulated by mir-200, while genes negatively correlated with *SPAG9* exon 24 inclusion are down-regulated in basal and claudin-low breast cancer subtypes. (*D*) Genes up-regulated during EMT and down-regulated by mir-200 are negatively correlated with *ZMYND8* exon 22 inclusion, while genes down-regulated in basal and claudin-low breast cancers are positively correlated with *ZMYND8* exon 22 inclusion. (*E*) Visualization of TCGA BRCA tumors with different genomic alterations in *ESRP1* or *hnRNPM*. Alterations in *ESRP1* and *hnRNPM* are mostly mutually exclusive.

enrichment analysis (GSEA) on gene signatures that are correlated with *SPAG9* exon 24 and *ZMYND8* exon 22 inclusion. Within the TCGA data set, a published EMT gene signature was positively correlated with *SPAG9* exon 24 inclusion, while this signature was negatively correlated with *ZMYND8* exon 22 inclusion (Fig. 4C,D; Taube et al. 2010). Genes down-regulated by mir-200, a potent repressor of EMT (Korpal and Kang 2008; Park et al. 2008), also showed a positive correlation with *SPAG9* exon 24 inclusion but a negative correlation with *ZMYND8* exon 22 inclusion. Moreover, genes down-

regulated in basal subtype and claudin-low subtype breast cancers, which are generally non-hormone dependent and resistant to conventional therapies (Blick et al. 2010; Prat et al. 2010), were negatively correlated with *SPAG9* exon 24 inclusion. Conversely, these genes were positively correlated with *ZMYND8* exon 22 inclusion.

As hnRNPM and ESRP1 showed primarily opposing functions in splicing regulation, we investigated the occurrence of genomic alterations in the genes encoding these two proteins in 963 TCGA breast tumors where genome sequencing and copy number analyses were available (Fig.

4E). Alterations in *ESRP1* occurred in 16% of patients compared to 1.7% with *hnRNPM* alterations, with the majority of alterations representing gene amplification through copy number variation. Amplification or mutations in these genes were mostly mutually exclusive, with only two patients showing amplification of both genes.

DISCUSSION

In summary, this study reveals widespread coregulation of alternative splicing by hnRNPM and ESRP1, both identified as key regulators of EMT splicing programs (Warzecha et al. 2010; Xu et al. 2014). hnRNPM-regulated cassette exons significantly overlap with ESRP1-regulated cassette exons, with the majority of coregulated events showing discordant splicing regulation, suggesting that hnRNPM and ESRP1 largely serve to functionally antagonize one another. We also observed a subset of splicing events regulated concordantly by hnRNPM and ESRP1 that inversely correlate with EMT splicing. Although these events represent a minority of coregulated events, these results suggest that hnRNPM is partly correlated with antagonistic regulation of splicing during EMT. These results are surprising as we observed that hnRNPM is required for cells to undergo EMT (Xu et al. 2014). Interestingly, this scenario is reminiscent of that observed for RBM47, an RNA binding protein that inhibits EMT (Vanharanta et al. 2014). RBM47 showed primarily discordant regulation of splicing events compared to EMT, but also concordant regulation of a subset of splicing events that promote EMT (Yang et al. 2016). These findings highlight the importance of understanding the combinatorial regulation of splicing by different factors with respect to a complex biological process such as EMT. Whether the hnRNPM-regulated splicing events that oppose EMT play a functional role during EMT or are important for regulating cellular processes that are highly active in epithelial cell states will be an interesting area for future study.

Some of the alternative splicing events coregulated by hnRNPM and ESRP1 have been investigated in detail to understand their functional contributions to EMT. *CD44* contains multiple variable exons that undergo extensive alternative splicing during EMT. Exon skipping of all variable exons of *CD44* to generate the CD44s isoform is required for EMT and has been shown to promote Akt-signaling, mediate invadopodia activity, and attenuate degradation of EGFR to promote sustained RTK signaling, all of which have implications during EMT (Brown et al. 2011; Zhao et al. 2016; Liu and Cheng 2017; Wang et al. 2017). ESRP1 and hnRNPM directly regulate alternative splicing of *CD44* in a discordant manner, with ESRP1 driving exon inclusion and hnRNPM promoting exon skipping. Another hnRNPM–ESRP1 coregulated splicing event that plays a role during EMT is alternative splicing of *EXOC7* exon 7 and exon 8. Skipping of *EXOC7* exons 7 and 8

produces a mesenchymal isoform capable of promoting actin polymerization and increased cell invasion compared to the epithelial isoform where part of exon 8 is included (Lu et al. 2013). ESRP1 was shown to regulate *EXOC7* splicing to promote production of the epithelial isoform, although the role of hnRNPM is not known. In addition, hnRNPM and ESRP1 coregulate exon skipping of *TCF7L2* exon 4. The exon-4 skipped *TCF7L2* isoform is capable of greater activation of Wnt/β-catenin target gene promoters compared to inclusion isoforms (Weise et al. 2010), and this isoform is up-regulated in the mesenchymal state. It is worth noting that in our study the *CD44* and *TCF7L2* splicing events were discordantly regulated by hnRNPM and ESRP1, with ESRP1 promoting production of the epithelial isoform and hnRNPM promoting the mesenchymal splicing pattern. While the precise molecular consequences of most of the hnRNPM–ESRP1 regulated splicing events during EMT are unknown, our observation that hnRNPM–ESRP1 coregulated splicing events are enriched in genes associated with cytoskeleton remodeling and cell adhesion present interesting avenues for further study.

Analysis of motifs near hnRNPM and ESRP1 coregulated cassettes identified enrichment of GU-rich motifs downstream from hnRNPM-repressed and ESRP1-enhanced events. Previous studies of hnRNPM-mediated alternative splicing identified enrichment of GU-rich sequences and hnRNPM binding sites by cross-linking and immunoprecipitation (CLIP) analyses specifically near hnRNPM-repressed but not hnRNPM-enhanced events, which is consistent with our study (Huelga et al. 2012). Similarly, GU-rich motifs have been previously identified downstream from ESRP1-enhanced and upstream of ESRP1-repressed exons, indicating a position-specific function of ESRP1 in regulating alternative splicing (Warzecha et al. 2010; Dittmar et al. 2012; Bebee et al. 2015; Yang et al. 2016). The enrichment of GU-rich motifs in hnRNPM–ESRP1 discordant exons supports a model of competitive binding between the two proteins, which is supported by our binding analysis downstream from variable exon 7 of *APLP2* as well as our previous study of *CD44* alternative splicing (Xu et al. 2014).

We also identified striking enrichment of UGCAUG, the motif recognized by the RBFOX family of RNA binding proteins, near hnRNPM–ESRP1 coregulated cassettes. RBFOX motifs have been previously shown to enrich near exons regulated by hnRNPM (Damianov et al. 2016) and ESRP1 (Warzecha et al. 2010; Dittmar et al. 2012; Yang et al. 2016). RBFOX proteins were recently shown to function in a complex with hnRNPM, along with other splicing factors, where the presence of both factors was necessary for their full splicing activity (Damianov et al. 2016). We found a small but significant overlap (~7%, $P = 4.52 \times 10^{-7}$, Fisher's exact test) between the hnRNPM–ESRP1 coregulated exons and a recently

published set of cassette exons regulated by RBFOX2 in HeLa cells (Weyn-Vanhentenryck et al. 2014). These preliminary results suggest that RBFOX2 is involved in regulating the set of hnRNPM–ESRP1 coregulated exons. Since RBFOX2 has been shown to promote alternative splicing during EMT (Shapiro et al. 2011; Braeutigam et al. 2013; Venables et al. 2013), we speculate that RBFOX2 may function in concert with hnRNPM and against ESRP1 to regulate EMT-associated alternative splicing. A detailed analysis of RBFOX2's regulatory relationship with hnRNPM and ESRP1 is an interesting area for future study.

Through analysis of hnRNPM and ESRP1 coregulated cassette exons using breast cancer TCGA data, we identified splicing events that stratify breast cancer patients by overall survival and observed correlations with EMT and breast cancer subtype gene sets. These data indicate that hnRNPM and ESRP1 splicing function is associated with breast cancer progression and prognosis. We also show that in breast tumors few genomic alterations exist in *hnRNPM*, while copy number variations in *ESRP1* are present in a subset of tumors, and these alterations appear to be mutually exclusive. We previously showed that hnRNPM expression positively correlates with breast tumor grade (Xu et al. 2014), and others have shown that hnRNPM expression promotes tumor aggressiveness and predicts poor prognosis (Sun et al. 2017). The relationship between ESRP1 expression and cancer progression is more complicated, with previous studies reporting a positive correlation between ESRP1 expression and longer patient survival in colorectal, pancreatic, and breast cancers (Ueda et al. 2014; Lu et al. 2015). Conversely, other studies have shown directly opposite correlations with ESRP1 and survival in breast cancers, ovarian cancers, and melanomas (Yae et al. 2012; Yao et al. 2016; Jeong et al. 2017). These conflicting results implicate both increased and decreased ESRP1 expression in tumor progression and suggest that ESRP1 function in cancer is highly tissue and context dependent.

In-depth analysis of hnRNPM and ESRP1 splicing regulation informs our understanding of splicing factor binding and functional dynamics in the context of disease-relevant splicing programs and indicates the importance of understanding the competitive and cooperative mechanisms of splicing regulation that allow precise modulation of alternative splicing. Taken together, we show that coregulation of alternative splicing by hnRNPM and ESRP1 is widespread and primarily antagonistic, although a subset of events is regulated concordantly. Furthermore, we demonstrate that in controlling hnRNPM–ESRP1 discordantly regulated events, hnRNPM promotes alternative splicing in the same direction as EMT. hnRNPM and ESRP1 splicing antagonism is explained, at least in part, by competition for GU-rich elements downstream from coregulated exons. Lastly, hnRNPM–ESRP1 coregulated splicing events correlate with EMT and breast cancer–associated gene sets and

predict breast cancer patient survival. Taken together, this study highlights the complex regulation of alternative splicing by ESRP1 and hnRNPM as well as the relevance of this regulatory interaction in EMT and cancer.

MATERIALS AND METHODS

Cell lines

Maintenance of immortalized human mammary epithelial cells (HMLE) cells was conducted as previously described (Brown et al. 2011). Human embryonic kidney 293FT, human breast carcinoma MDA-MB-231, and MDA-MB-231 metastatic derivative lines 4175 (LM2) and 1833 (BM1) were grown in DMEM supplemented with 10% FBS, L-glutamine, penicillin, and streptomycin. The ESRP1-HA overexpressing MDA-MB-231 cell line was described previously (Reinke et al. 2012).

Plasmids, shRNAs, and ESRP1 overexpression

Two expression plasmids, pcDNA3-FLAG-hnRNPM and pcDNA3-FLAG-ESRP1, were subcloned from pECFP-hnRNPM (Lleres et al. 2010) and pBRIT-ESRP1 (Brown et al. 2011), respectively. The *CD44* variable exon 5 minigene was described previously (Brown et al. 2011). The *MARK3* exon 17 minigene was constructed through PCR amplification of *MARK3* exon 17 and ~400 nt of flanking intron followed by cloning into the BamH1 site of the *CD44*v5 minigene. Primers for *MARK3* minigene construction are listed in Supplemental Table S4. The control, hnRNPM, and ESRP1 shRNAs were described previously (Brown et al. 2011; Xu et al. 2014). Control shRNA: 5′-CCCGAATTAGCTGGACACT CAA-3′. hnRNPM shRNA: 5′-GGAATGGAAGGCATAGGATTT-3′. The ESRP1 shRNA was obtained from Open Biosystems (clone V2LHS_155255) with sequence 5′-CGCATAAGATCTTGGAAT AATA-3′.

Transfection, semi-quantitative RT-PCR, and qRT-PCR (qPCR)

Briefly, 2.25×10^5 HEK293FT cells were plated in 24-well plates 24 h prior to transfection. Cotransfection of hnRNPM and/or ESRP1 with splicing minigenes was performed using Lipofectamine 2000 (Invitrogen) per the manufacturer's instructions. For HEK293FT transfections, 100 ng of minigene was used. RNA was extracted from cells using the E.Z.N.A. Total RNA Kit (Omega Bio-Tek). RNA concentration was measured using a Nanodrop 2000 (Thermo Fisher Scientific). cDNA was generated via reverse transcription using the GoScript Reverse Transcription System (Promega) with 1 µL GoScript RT and 250 ng of RNA in a total volume of 20 µL followed by incubation at 25°C for 5 min, 42°C for 30 min, and 70°C for 15 min. Semi-quantitative RT-PCR assaying for splicing products was performed using Hot StarTaq DNA polymerase (Qiagen), and PCR cycles were run for 30 or fewer cycles.

Primers for semi-quantitative analysis were designed in constitutive exons flanking each variable exon (Harvey and Cheng 2016). Semi-quantitative PCR generates both exon inclusion and skipping products which were separated through agarose

gel electrophoresis. PCR product intensity was measured using ImageJ image analysis software. qPCR was performed using GoTaq qPCR Master Mix (Promega) per the manufacturer's instructions on a CFX Connect Real-Time PCR system (BioRad) using a two-step protocol and supplied software. For every qPCR sample per biological replicate, two technical replicates were performed and Ct counts were averaged. Normalization and quantification of qPCR data was done using the $2^{-\Delta Ct}$ method relative to TATA-binding protein (TBP) expression (Livak and Schmittgen 2001). Primer specificity was verified with melt-curve analysis during qPCR. Primers for semi-quantitative and qPCR analysis were designed to generate products spanning introns (Supplemental Table S4).

Quantitative immunoblotting

Whole cell lysates or RNA pull-down samples were separated by 10% SDS-PAGE, transferred to a PVDF membrane (BioRad), and probed with the appropriate antibody. Primary antibodies used in western blots included HA-HRP (Roche Applied Science), hnRNPM (OriGene). GAPDH (GE) and β-actin (Sigma-Aldrich) were used as loading controls. After incubation with HRP-tagged secondary antibodies, if appropriate, blots were visualized via chemiluminescence (Thermo Fisher).

RNA pull-down assays

5′-biotinylated nucleotides were used for RNA pull-down experiments. The *APLP2* exon 7 associated probes include GU1: 5′-biotin-UCUGUGUGGUGUCCCUGCCCACUCGGGUGUUUGCU, which was mutated to GU1 mut: 5′-biotin-UCUACGUAAUCUCCCUGCCCACUCGCCUGCAUGCU and GU2: 5′-biotin-CGUGUGUCUGGUGGUGCUUGGUGGUGAUGGUGC, which was mutated to GU2 mut: 5′-biotin-CGUACGUCUCCUGCAGCUAGCUAAUGAUACUCC. Biotinylated RNA oligos (10 μL at 40 μM) were immobilized on 50 μL of streptavidin beads (50% slurry; Thermo Fisher) in a total volume of 400 μL 1× binding buffer (20 mM Tris, 200 mM NaCl, 6 mM EDTA, 5 mM sodium fluoride, 5 mM β-glycerophosphate, 2 mg/mL aprotinin, pH 7.5) for 2 h at 4°C in a rotating shaker. After immobilization, beads were washed three times in 1× binding buffer, then 200 μg MDA-MB-231 or MDA-MB-231 ESRP1-HA cell lysates were suspended with the beads in 400 μL of 1× binding buffer for incubation at 4°C overnight. Beads were then washed three times in 1× binding buffer, resuspended in 60 μL of 2× Laemmli sample buffer (Bio-Rad), and boiled for 5 min. Ten microliters of sample was analyzed under denaturing conditions on 10% SDS-PAGE and detected via immunoblotting.

RNA sequencing analysis

Total RNA was extracted from LM2 or BM1 cells stably expressing control or hnRNPM shRNAs using TRIzol, and poly(A)-selected RNA-seq libraries were generated using TruSeq Stranded mRNA Library Preparation Kits (Illumina) and subjected to 100-bp PE stranded RNA-seq on an Illumina HiSeq 4000. RNA-seq reads were aligned to the human genome (GRCh37, primary assembly) and transcriptome (Gencode version 24 backmap 37

comprehensive gene annotation) using STAR version 2.5.3a (Dobin et al. 2013) with the alignEndsType parameter set as EndToEnd. Differential alternative splicing of cassette exons was quantified using rMATS version 3.2.5 (Shen et al. 2014), and differential splicing events were collected using the following cutoffs: FDR < 0.05, |ΔPSI| ≥ 0.1, and average junction reads per cassette event ≥10. Control cassette exons not differentially spliced were defined by the following filters: FDR > 0.5, minimum PSI for control or hnRNPM knockdown samples <0.85, maximum PSI > 0.15, and average junction reads per cassette event ≥10. RNA sequencing data for H358 shESRP1/2 was processed in the same way after retrieving raw reads from GEO record GSE75492. After rMATS analysis, cassette exons events were collapsed if they shared identical cassette exon coordinates. Other ESRP1-regulated cassette exons were obtained from the supplemental material of published studies (Warzecha et al. 2010; Dittmar et al. 2012). EMT-regulated cassette exons were obtained from the supplemental material of two previous studies (Shapiro et al. 2011; Yang et al. 2016). Sequencing data sets for hnRNPM are deposited in the Gene Expression Omnibus at GSE112516.

Motif enrichment and RNA motif maps

K-mer enrichments were calculated using 250 bp of the sequence flanking the cassette exons. To avoid enrichment of canonical splice site motifs, 9 nt downstream from 5′ splice sites and 25 nt upstream of 3′ splice sites were removed. To assess enrichment of hexamers, a previously published motif analysis method was used (Coelho et al. 2015). For the GU-rich motif RNA map analysis, the top 12 6-mer motifs from an ESRP1-SELEX-Seq analysis were obtained (Dittmar et al. 2012). The motifs were UGGUGG, GGUGGG, GUGGUG, GUGGGG, GUGUGG, GGUGUG, UGUGGG, GGUGGU, GUGGGU, UGGGGU, GGGGGU, UGGGGG. The 250 bp of sequence flanking upstream and downstream hnRNPM and ESRP1 regulated cassette exons as well as control exons obtained from the rMATS alternative splicing analysis were also obtained. The GU-rich motif score was computed in a custom python script by counting the number of nucleotides covered by any of the GU-rich motifs in a sliding window of 50 bp shifted 1 nt at a time across the 250 bp interval in all of the regulated and control cassette exons. The GU-rich motif score was set equal to the percent of nucleotides covered by the motifs in each of the sliding windows and plotted for regulated exons, stratified by inclusion or skipping, and control exons.

TCGA BRCA survival analysis, genomic alterations, GSEA, and gene ontology

Processed TCGA BRCA level 3 RNA Seq V2 data for exon junctions and gene expression were downloaded from the Genomic Data Commons Legacy Archive (Cancer Genome Atlas Network 2012). Cassette exon PSI values in each patient sample were calculated using the following equation from the exon junction files: PSI = (Inclusion junction reads/2)/([Inclusion junction reads/2] + [Skipping junction reads]). Patients were clustered into two groups using *K*-means clustering by PSI value. Kaplan–Meier survival analysis was conducted between these two groups using overall survival. *P*-values were computed using log-rank tests.

For GSEA analysis, correlation values between the given cassette exon PSI values and all genes in the TCGA BRCA RNA-seq V2 data sets were computed. Genes were then ranked by correlation and GSEA was performed using the Broad Institute javaGSEA desktop application (Mootha et al. 2003; Subramanian et al. 2005).

For analysis of genomic alterations in *ESRP1* and *hnRNPM*, the TCGA BRCA Provisional data set containing tumor samples with genome sequencing and copy number variation analysis (n = 963) was queried using cBioPortal (www.cbioportal.org). The genomic alterations figure was derived from the cBioPortal Oncoprint analysis page.

Gene ontology analysis was conducted using DAVID v6.8 with the gene list composed of all genes containing an exon in the 213 hnRNPM-ESRP1 coregulated exons (Huang da et al. 2009a,b). GOTERM enrichment was restricted to GOTERM BP-DIRECT, GOTERM MF-DIRECT, and GOTERM CC-DIRECT. The background gene set was composed of all genes in hnRNPM and ESRP1 RNA sequencing data sets with FPKM > 3 in at least one sample. All significant GOTERMs are listed in Supplemental Table S3.

Statistics

Statistical analyses included two-tailed Student's *t*-tests and hypergeometric testing unless otherwise noted. *P*-values <0.05 were considered significant: $P < 0.05$ (*), $P < 0.01$ (**), and $P < 0.001$ (***).

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

Author contributions: S.E.H. and C.C. designed experiments. Y.X. generated cell lines. X.L. and X.D.G. cloned splicing minigenes. Y.Q. performed splicing quantification from TCGA data. S.E.H. performed experiments, bioinformatics analysis, and analyzed data. J.A. assisted with analysis of RNA-sequencing data. X.X. supervised bioinformatics analysis. S.E.H. and C.C. wrote the manuscript. C.C. supervised the study.

REFERENCES

Barash Y, Calarco JA, Gao W, Pan Q, Wang X, Shai O, Blencowe BJ, Frey BJ. 2010. Deciphering the splicing code. *Nature* **465:** 53–59.

Bebee TW, Park J, Sheridan KI, Warzecha CC, Cieply BW, Rohacek AM, Xing Y, Carstens RP. 2015. The splicing regulators Esrp1 and Esrp2 direct an epithelial splicing program essential for mammalian development. *Elife* **4:** e08954.

Blick T, Hugo H, Widodo E, Waltham M, Pinto C, Mani SA, Weinberg RA, Neve RM, Lenburg ME, Thompson EW. 2010. Epithelial mesenchymal transition traits in human breast cancer cell lines parallel the CD44^{hi}/CD24$^{lo/-}$ stem cell phenotype in human breast cancer. *J Mammary Gland Biol Neoplasia* **15:** 235–252.

Braeutigam C, Rago L, Rolke A, Waldmeier L, Christofori G, Winter J. 2013. The RNA-binding protein Rbfox2: an essential regulator of EMT-driven alternative splicing and a mediator of cellular invasion. *Oncogene* **33:** 1082–1092.

Brown RL, Reinke LM, Damerow MS, Perez D, Chodosh LA, Yang J, Cheng C. 2011. CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression. *J Clin Invest* **121:** 1064–1074.

Cancer Genome Atlas Network. 2012. Comprehensive molecular portraits of human breast tumours. *Nature* **490:** 61–70.

Cieply B, Carstens RP. 2015. Functional roles of alternative splicing factors in human disease. *Wiley Interdiscip Rev RNA* **6:** 311–326.

Coelho MB, Attig J, Bellora N, König J, Hallegger M, Kayikci M, Eyras E, Ule J, Smith CWJ. 2015. Nuclear matrix protein Matrin3 regulates alternative splicing and forms overlapping regulatory networks with PTB. *EMBO J* **34:** 653–668.

Damianov A, Ying Y, Lin C-H, Lee J-A, Tran D, Vashisht AA, Bahrami-Samani E, Xing Y, Martin KC, Wohlschlegel JA, et al. 2016. Rbfox proteins regulate splicing as part of a large multiprotein complex LASR. *Cell* **165:** 606–619.

Dittmar KA, Jiang P, Park J, Amirikian K, Wan J, Shen S, Xing Y, Carstens RP. 2012. Genome-wide determination of a broad ESRP-regulated posttranscriptional network by high-throughput sequencing. *Mol Cell Biol* **32:** 1468–1482.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* **29:** 15–21.

The Encode Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489:** 57–74.

Grille SJ, Bellacosa A, Upson J, Klein-Szanto AJ, van Roy F, Lee-Kwon W, Donowitz M, Tsichlis PN, Larue L. 2003. The protein kinase Akt induces epithelial mesenchymal transition and promotes enhanced motility and invasiveness of squamous cell carcinoma lines. *Cancer Res* **63:** 2172–2178.

Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, et al. 2012. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* **22:** 1760–1774.

Harvey SE, Cheng C. 2016. Methods for characterization of alternative RNA splicing. *Methods Mol Biol* **1402:** 229–241.

Hernandez JR, Kim JJ, Verdone JE, Liu X, Torga G, Pienta KJ, Mooney SM. 2015. Alternative CD44 splicing identifies epithelial prostate cancer cells from the mesenchymal counterparts. *Med Oncol* **32:** 159.

Huang da W, Sherman BT, Lempicki RA. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37:** 1–13.

Huang da W, Sherman BT, Lempicki RA. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4:** 44–57.

Huelga SC, Vu AQ, Arnold JD, Liang TY, Liu PP, Yan BY, Donohue J, Shiue L, Hoon S, Brenner S, et al. 2012. Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Rep* **1:** 167–178.

Jangi M, Boutz PL, Paul P, Sharp PA. 2014. Rbfox2 controls autoregulation in RNA-binding protein networks. *Genes Dev* **28:** 637–651.

Jeong HM, Han J, Lee SH, Park HJ, Lee HJ, Choi JS, Lee YM, Choi YL, Shin YK, Kwon MJ. 2017. ESRP1 is overexpressed in ovarian cancer

and promotes switching from mesenchymal to epithelial pheno-type in ovarian cancer cells. *Oncogenesis* **6:** e391.

Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, Cordon-Cardo C, Guise TA, Massague J. 2003. A multigenic program me-diating breast cancer metastasis to bone. *Cancer Cell* **3:** 537–549.

Korpal M, Kang Y. 2008. The emerging role of miR-200 family of microRNAs in epithelial-mesenchymal transition and cancer me-tastasis. *RNA Biol* **5:** 115–119.

Liu S, Cheng C. 2013. Alternative RNA splicing and cancer. *Wiley Interdiscip Rev RNA* **4:** 547–566.

Liu S, Cheng C. 2017. Akt signaling is sustained by a CD44 splice isoform–mediated positive feedback loop. *Cancer Res* **77:** 3791–3801.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* **25:** 402–408.

Lleres D, Denegri M, Biggiogera M, Ajuh P, Lamond AI. 2010. Direct interaction between hnRNP-M and CDC5L/PLRG1 proteins affects alternative splice site choice. *EMBO Rep* **11:** 445–451.

Lu H, Liu J, Liu S, Zeng J, Ding D, Carstens RP, Cong Y, Xu X, Guo W. 2013. Exo70 isoform switching upon epithelial-mesenchymal transition mediates cancer cell invasion. *Dev Cell* **27:** 560–573.

Lu ZX, Huang Q, Park JW, Shen S, Lin L, Tokheim CJ, Henry MD, Xing Y. 2015. Transcriptome-wide landscape of pre-mRNA alternative splicing associated with metastatic colonization. *Mol Cancer Res* **13:** 305–318.

Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, Viale A, Olshen AB, Gerald WL, Massague J. 2005. Genes that mediate breast cancer metastasis to lung. *Nature* **436:** 518–524.

Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, et al. 2003. PGC-1α-responsive genes involved in oxidative phosphory-lation are coordinately downregulated in human diabetes. *Nat Genet* **34:** 267–273.

Nieto AM, Huang R, Jackson RA, Thiery J. 2016. EMT: 2016. *Cell* **166:** 21–45.

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* **40:** 1413–1415.

Park SM, Gaur AB, Lengyel E, Peter ME. 2008. The miR-200 family de-termines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev* **22:** 894–907.

Passacantilli I, Frisone P, De Paola E, Fidaleo M, Paronetto MP. 2017. hnRNPM guides an alternative splicing program in response to inhibition of the PI3K/AKT/mTOR pathway in Ewing sarcoma cells. *Nucleic Acids Res* **45:** 12270–12284.

Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, He X, Perou CM. 2010. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* **12:** R68.

Reinke LM, Xu Y, Cheng C. 2012. Snail represses the splicing regulator epithelial splicing regulatory protein 1 to promote epithelial-mes-enchymal transition. *J Biol Chem* **287:** 36435–36442.

Shapiro IM, Cheng AW, Flytzanis NC, Balsamo M, Condeelis JS, Oktay MH, Burge CB, Gertler FB. 2011. An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype. *PLoS Genet* **7:** e1002218.

Shen S, Park J, Lu Z-x, Lin L, Henry MD, Wu Y, Zhou Q, Xing Y. 2014. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-seq data. *Proc Natl Acad Sci* **111:** E5593–E5601.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci* **102:** 15545–15550.

Sun H, Liu T, Zhu D, Dong X, Liu F, Liang X, Chen C, Shao B, Wang M, Wang Y. 2017. HnRNPM and CD44s expression affects tumor aggressiveness and predicts poor prognosis in breast cancer with axillary lymph node metastases. *Genes Chromosomes Cancer* **56:** 598–607.

Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, Yang J, Hartwell K, Onder TT, Gupta PB, Evans KW, et al. 2010. Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc Natl Acad Sci* **107:** 15449–15454.

Thiery J. 2003. Epithelial-mesenchymal transitions in development and pathologies. *Curr Opin Cell Biol* **15:** 740–746.

Thiery JP, Acloque H, Huang RY, Nieto MA. 2009. Epithelial-mesen-chymal transitions in development and disease. *Cell* **139:** 871–890.

Ueda J, Matsuda Y, Yamahatsu K, Uchida E, Naito Z, Korc M, Ishiwata T. 2014. Epithelial splicing regulatory protein 1 is a favor-able prognostic factor in pancreatic cancer that attenuates pancre-atic metastases. *Oncogene* **33:** 4485–4495.

Vanharanta S, Marney CB, 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, Brosseau J-P, Gadea G, Klinck R, Prinos P, Beaulieu J-F, Lapointe E, Durand M, Thibault P, Tremblay K, et al. 2013. RBFOX2 is an important regulator of mesenchymal tissue-specific splicing in both normal and cancer tissues. *Mol Cell Biol* **33:** 396–405.

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456:** 470–476.

Wang W, Zhang H, Liu S, Kim C, Xu Y, Hurley LA, Nishikawa R, Nagane M, Hu B, Stegh AH, et al. 2017. Internalized CD44s splice isoform attenuates EGFR degradation by targeting Rab7A. *Proc Natl Acad Sci* **114:** 8366–8371.

Warzecha CC, Sato TK, Nabet B, Hogenesch JB, Carstens RP. 2009a. ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing. *Mol Cell* **33:** 591–601.

Warzecha CC, Shen S, Xing Y, Carstens RP. 2009b. The epithelial splicing factors ESRP1 and ESRP2 positively and negatively regu-late diverse types of alternative splicing events. *RNA Biol* **6:** 546–562.

Warzecha CC, Jiang P, Amirikian K, Dittmar KA, Lu H, Shen S, Guo W, Xing Y, Carstens RP. 2010. An ESRP-regulated splicing programme is abrogated during the epithelial-mesenchymal transition. *EMBO J* **29:** 3286–3300.

Weise A, Bruser K, Elfert S, Wallmen B, Wittel Y, Wohrle S, Hecht A. 2010. Alternative splicing of Tcf7l2 transcripts generates protein variants with differential promoter-binding and transcriptional ac-tivation properties at Wnt/β-catenin targets. *Nucleic Acids Res* **38:** 1964–1981.

Weyn-Vanhentenryck SM, Mele A, Yan Q, Sun S, Farny N, Zhang Z, Xue C, Herre M, Silver PA, Zhang MQ, et al. 2014. HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory net-work linked to brain development and autism. *Cell Rep* **6:** 1139–1152.

Xu Y, Gao XD, Lee JH, Huang H, Tan H, Ahn J, Reinke LM, Peter ME, Feng Y, Gius D, et al. 2014. Cell type-restricted activity of hnRNPM promotes breast cancer metastasis via regulating alternative splic-ing. *Genes Dev* **28:** 1191–1203.

Yae T, Tsuchihashi K, Ishimoto T, Motohara T, Yoshikawa M, Yoshida GJ, Wada T, Masuko T, Mogushi K, Tanaka H, et al. 2012. Alternative splicing of CD44 mRNA by ESRP1 enhances lung colonization of metastatic cancer cell. *Nat Commun* **3:** 883.

Yang J, Weinberg RA. 2008. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell* **14:** 818–829.

Yang Y, Park J, Bebee TW, Warzecha CC, Guo Y, Shang X, Xing Y, Carstens RP. 2016. Determination of a comprehensive alternative splicing regulatory network and combinatorial regulation by key factors during the epithelial-to-mesenchymal transition. *Mol Cell Biol* **36:** 1704–1719.

Yao J, Caballero OL, Huang Y, Lin C, Rimoldi D, Behren A, Cebon JS, Hung MC, Weinstein JN, Strausberg RL, et al. 2016. Altered expression and splicing of ESRP1 in malignant melanoma correlates with epithelial-mesenchymal status and tumor-associated immune cytolytic activity. *Cancer Immunol Res* **4:** 552–561.

Zhao P, Xu Y, Wei Y, Qiu Q, Chew T-L, Kang Y, Cheng C. 2016. The CD44s splice isoform is a central mediator for invadopodia activity. *J Cell Sci* **129:** 1355–1365.