Title

Cell type-restricted activity of hnRNPM promotes breast cancer metastasis via regulating alternative splicing

Permalink

https://escholarship.org/uc/item/3jn699jb

Journal

Genes & Development, 28(11)

ISSN

0890-9369

Authors

Xu, Yilin
Gao, Xin D
Lee, Jae-Hyung
et al.

Publication Date

2014-06-01

DOI

10.1101/gad.241968.114

Peer reviewed

Cell type-restricted activity of hnRNPM promotes breast cancer metastasis via regulating alternative splicing

Yilin Xu,[1] Xin D. Gao,[1] Jae-Hyung Lee,[2,8,9] Huilin Huang,[1] Haiyan Tan,[3,4] Jaegyoon Ahn,[2] Lauren M. Reinke,[1] Marcus E. Peter,[1] Yue Feng,[5] David Gius,[6] Kalliopi P. Siziopikou,[7] Junmin Peng,[3,4] Xinshu Xiao,[2] and Chonghui Cheng[1,10]

[1]Division of Hematology/Oncology, Robert H. Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60611, USA; [2]Department of Integrative Biology and Physiology, University of California at Los Angeles, Los Angeles, California 90095, USA; [3]Department of Structural Biology, St. Jude Proteomics Facility, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, USA; [4]Department of Developmental Neurobiology, St. Jude Proteomics Facility, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, USA; [5]Department of Pharmacology, Emory University, Atlanta, Georgia 30322, USA; [6]Department of Radiation Oncology, Robert H. Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60611, USA; [7]Department of Pathology, Robert H. Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60611, USA

Tumor metastasis remains the major cause of cancer-related death, but its molecular basis is still not well understood. Here we uncovered a splicing-mediated pathway that is essential for breast cancer metastasis. We show that the RNA-binding protein heterogeneous nuclear ribonucleoprotein M (hnRNPM) promotes breast cancer metastasis by activating the switch of alternative splicing that occurs during epithelial–mesenchymal transition (EMT). Genome-wide deep sequencing analysis suggests that hnRNPM potentiates TGFβ signaling and identifies CD44 as a key downstream target of hnRNPM. hnRNPM ablation prevents TGFβ-induced EMT and inhibits breast cancer metastasis in mice, whereas enforced expression of the specific CD44 standard (CD44s) splice isoform overrides the loss of hnRNPM and permits EMT and metastasis. Mechanistically, we demonstrate that the ubiquitously expressed hnRNPM acts in a mesenchymal-specific manner to precisely control CD44 splice isoform switching during EMT. This restricted cell-type activity of hnRNPM is achieved by competition with ESRP1, an epithelial splicing regulator that binds to the same *cis*-regulatory RNA elements as hnRNPM and is repressed during EMT. Importantly, hnRNPM is associated with aggressive breast cancer and correlates with increased CD44s in patient specimens. These findings demonstrate a novel molecular mechanism through which tumor metastasis is endowed by the hnRNPM-mediated splicing program.

Tumor metastasis involves the spreading of cancer cells from the primary site to distant organs. Despite recent advances in cancer therapy, tumor metastasis remains the leading cause of death in cancer patients. Thus, understanding the mechanisms underlying metastasis is crucial for developing effective strategies to treat this deadly disease.

During metastasis, tumor cells frequently hijack the developmental program of epithelial–mesenchymal transition (EMT) to invade surrounding tissues and migrate to distant organs (Thiery and Sleeman 2006; Yang and Weinberg 2008). EMT is induced by the cytokine TGFβ and a group of transcription factors, including Twist, Snail, and Zeb1/2 (De Craene and Berx 2013). These factors suppress the expression of the adherens junction molecule E-cadherin, resulting in characteristics of EMT that include a morphological change from a cobblestone-like epithelial appearance to a spindle-like fibroblast

Present address: [8]Department of Life and Nanopharmaceutical Sciences, [9]Department of Maxillofacial Biomedical Engineering, School of Dentistry, Kyung Hee University, Seoul 130-701, Korea.
[10]Corresponding author
E-mail chengc@northwestern.edu

phenotype, loss of E-cadherin at cell junctions, increased expression of mesenchymal N-cadherin, and acquisition of a migratory phenotype.

Emerging evidence has suggested that the regulation of alternative RNA splicing plays a causal role in EMT. By studying alternative splicing of the cell adhesion molecule CD44, we demonstrated previously that CD44 isoform switching occurs during EMT, resulting in production of the mesenchymal-specific CD44 standard (CD44s) splice isoform. We further showed that this CD44 isoform switching is required in order for cells to undergo EMT (Brown et al. 2011). Genome-wide analysis also revealed that alternative splicing of a group of genes is associated with epithelial or mesenchymal phenotypes and invasive breast tumors (Warzecha et al. 2010; Shapiro et al. 2011). These data suggest that alternative splicing is important for EMT and EMT-associated invasiveness. However, a significant question that has remained is whether aberrant splicing is functionally critical for cancer metastasis.

In the present study, we set out to identify splicing factors that are crucial for breast cancer metastasis. Using a cis-acting RNA element important for EMT-associated alternative splicing as a bait, we identified the heterogeneous nuclear ribonucleoprotein M (hnRNPM) as an essential splicing factor that potentiates TGFβ signaling, drives cells to undergo EMT, and, more importantly, promotes breast cancer metastasis in animals. We demonstrate that the molecular pathway of hnRNPM-mediated CD44 splice isoform switching is essential for breast cancer metastasis. Mechanistically, hnRNPM exerts a mesenchymal-specific activity to stimulate CD44 exon skipping, and this is achieved at least in part by the antagonizing function of an epithelial splicing regulator, ESRP1, whose expression is repressed during EMT. Significantly, increased hnRNPM expression is associated with high-grade human breast tumors and positively correlates with elevated expression of the CD44s splice isoform. Together, these studies provide compelling evidence demonstrating that a molecular pathway through hnRNPM-mediated alternative splicing causally contributes to breast cancer metastasis, functionally connecting the molecular program of splicing regulation and cancer metastasis in patients.

Results

Identification of the splicing regulator hnRNPM that favors an EMT phenotype

To identify splicing regulators that favor a mesenchymal phenotype, a cis-acting RNA element, I-8, was used as a trap for RNA-binding proteins in mesenchymal cells (Fig. 1A). This I-8 element is located in the intron downstream from the CD44 variable exon 8 (v8) and contains a stretch of GU-rich motifs that is necessary for mediating an epithelial splicing pattern in epithelial cells (Reinke et al. 2012). As such, it was proposed that these trapped mesenchymal-expressed splicing factors are candidates for suppressing epithelial splicing, allowing the production of mesenchymal splicing isoforms necessary
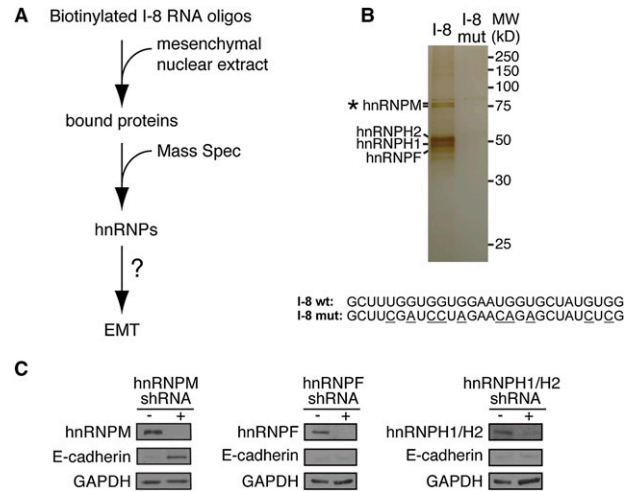
Figure 1. Identification of the splicing regulator hnRNPM that favors an EMT phenotype. (A) Schematic of the RNA pull-down coupled with MS approach to identify RNA-binding proteins that interact with the EMT-associated CD44 I-8 cis-acting element. (B, top) A silver stain image showing unique proteins identified by MS. (Bottom) Sequences of the GU-rich I-8 element and its mutant, in which the GU-rich motifs are disrupted by A or C nucleotide substitution (underlined). (C) Immunoblot analysis showing that silencing hnRNPM, but not the other hnRNPs, in mesenchymal Mes10A cells causes an increase in expression of the epithelial marker E-cadherin. GAPDH served as a loading control.

for EMT. An RNA pull-down assay was performed using a biotin-labeled 28-nucleotide (nt) I-8 RNA together with nuclear extracts of the mesenchymal MDA-MB-231 breast cancer cells. Silver-staining and mass spectrometry (MS) analysis were used to identify proteins bound to the I-8 RNA as well as a mutant I-8 probe as a control for nonspecific protein binding (Fig. 1B).

The four most abundant I-8-interacting proteins, hnRNPM, hnRNPH1/H2, and hnRNPF (Fig. 1B), are members of the hnRNP family, which consists of RNA-binding proteins frequently involved in splicing suppression (Han et al. 2010; Busch and Hertel 2012). Immunoblot analysis confirmed the association between these hnRNPs and the I-8 motif (Supplemental Fig. S1A), consistent with previous reports that hnRNPM recognizes GU-rich sequences, and hnRNPH and hnRNPF recognize G tracts (Swanson and Dreyfuss 1988; Datar et al. 1993; Matunis et al. 1994; Huelga et al. 2012).

shRNA knockdown of each of these hnRNPs in mesenchymal Mes10A cells, derived by treating MCF10A cells with TGFβ (Brown et al. 2011), revealed that depletion of hnRNPM was sufficient to promote a marked increase in the epithelial marker E-cadherin, while no increase in E-cadherin was associated with depletion of the other I-8-trapped hnRNPs (Fig. 1C). Further analysis revealed that hnRNPM ablation also increased expression of epithelial markers γ-catenin and occludin and decreased expression of the mesenchymal marker vimentin (Supplemental Fig. S1B), suggesting that the splicing factor hnRNPM is important for a mesenchymal phenotype.

Depletion of hnRNPM impairs EMT

To determine the role of hnRNPM in regulating EMT, a well-established inducible EMT model system was used in which TGFβ treatment of immortalized human mammary epithelial cells (HMLEs) induces EMT (Mani et al. 2008). Knockdown of *hnRNPM* using two shRNAs, M2 and M4 (Fig. 2A; Supplemental Fig. S2A), impaired TGFβ-induced EMT, indicated by (1) preservation of high levels of epithelial markers E-cadherin, γ-catenin, and occludin (Fig. 2A); (2) failed up-regulation of the mesenchymal markers N-cadherin and vimentin (Fig. 2A); (3) preserved

Figure 2. Depletion of hnRNPM impairs EMT. (*A*) Immunoblot analysis of EMT markers in HMLEs expressing control shRNA or hnRNPM shRNAs (M2 or M4) before (untreated) and after 18 d of TGFβ treatment. Upon TGFβ treatment, hnRNPM knockdown cells show impaired down-regulation of epithelial markers E-cadherin, γ-catenin, and occludin and diminished up-regulation of mesenchymal markers N-cadherin and vimentin compared with control. GAPDH served as a loading control. (*B*) Immunofluorescence images (63×) indicating that, unlike control cells, hnRNPM shRNA-expressing cells maintain E-cadherin expression and localization at cell junctions after 18 d of TGFβ treatment. Green staining indicates E-cadherin. DAPI staining (blue) indicates nuclei. (*C*) Phase contrast images (10×) illustrating impaired morphological changes in cells expressing hnRNPM shRNAs (M2 or M4) after 18 d of TGFβ treatment. (*D*) Immunoblot analysis of EMT markers showing that hnRNPM knockdown inhibits TGFβ-induced EMT in H358 cells. (*E*) Immunoblot analysis of MDA-MB-231 cells that express control and hnRNPM M2 shRNAs indicating that hnRNPM knockdown results in increased expression of E-cadherin. (*F*) Scratch wound assay showing that depletion of hnRNPM inhibits wound closure. The width of the wound was measured at 0 h, 16 h, and 24 h after the scratch was made. The percentage of wound healed was plotted. (*G*) Boyden chamber assay showing that hnRNPM knockdown inhibits tumor cell invasion. The numbers of invaded cells were plotted. Error bars in *F* and *G* indicate SEM. *n* > 4. (***) *P* < 0.001; (*) *P* < 0.05.

E-cadherin localization at cell–cell junctions (Fig. 2B); and (4) maintenance of a cobblestone-like epithelial morphology (Fig. 2C). These observations demonstrate that hnRNPM is necessary for TGFβ-induced EMT.

Two additional inducible EMT systems, the TGFβ-induced EMT in H358 lung epithelial cells (Thomson et al. 2011) and the tamoxifen-induced EMT in HMLEs that express a Twist-ER (estrogen receptor) fusion protein (Mani et al. 2008), were also used to validate that *hnRNPM* silencing prevents epithelial cells from transitioning to a mesenchymal state, as measured by EMT markers (Fig. 2D; Supplemental Fig. S2B). Furthermore, *hnRNPM* knockdown in mesenchymal MDA-MB-231 breast cancer cells converted these cells to express increased levels of the epithelial marker E-cadherin (Fig. 2E). These results echo those shown in Figure 1C (left panel), suggesting that hnRNPM is important for the maintenance of a mesenchymal phenotype. In addition, depletion of hnRNPM in MDA-MB-231 cells impaired wound closure in a scratch wound assay (Fig. 2F; Supplemental Fig. S2C) and showed diminished tumor cell invasion (Fig. 2G). These results support the function of hnRNPM in tumor cell motility and invasion and are consistent with its role in the acquisition of a mesenchymal phenotype.

Genome-wide analysis of hnRNPM-mediated pathways suggests its role in TGFβ signaling and EMT

To identify downstream pathways and targets of hnRNPM that direct EMT, deep RNA sequencing (RNA-seq) was performed using the mesenchymal MDA-MB-231-derived lung metastatic LM2 clonal line (Minn et al. 2005) that expressed either control or hnRNPM shRNA. hnRNPM knockdown increased the expression of 181 genes and decreased 263 by at least 2.5-fold (Supplemental Table S1). Gene ontology (GO) enrichment analysis revealed that hnRNPM-dependent genes were involved in biological processes such as response to wounding, cell migration, and cell motility (Fig. 3A; Supplemental Table S2A), supporting a role for hnRNPM in cell migration and wound healing. Pathway analysis detected significant enrichment of two molecular pathways (Fig. 3A; Supplemental Table S2B), one of which is the TGFβ signaling cascade, a key pathway that drives EMT and metastasis. Validation experiments showed that hnRNPM silencing reduced expression of TGFβ-positive regulators, including *TGFβRII* and *ACVR1*, and increased expression of the TGFβ-negative regulators *BMP4*, *BMP5*, and *Smad6* (Supplemental Fig. S3A). Notably, depletion of hnRNPM inhibited TGFβ-induced phosphorylation of its downstream effector, Smad3 (Fig. 3B), suggesting that hnRNPM potentiates TGFβ signaling. Furthermore, our gene set enrichment analysis (GSEA) indicated that hnRNPM is required for global enrichment of EMT- and TGFβ-related gene signatures and is associated with negative enrichment of the epithelial CDH1 (E-cadherin) gene signature (Fig. 3C; McBryan et al. 2007; Onder et al. 2008; Sarrio et al. 2008). Collectively, these data suggest that hnRNPM functionally promotes TGFβ signaling and EMT.
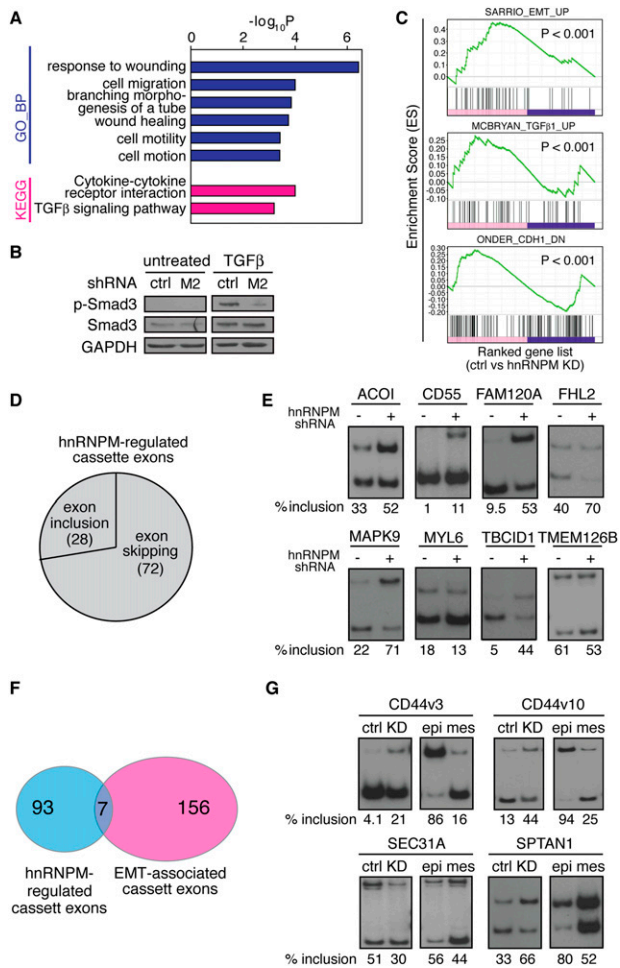
Figure 3. Genome-wide analysis of hnRNPM targets suggests its role in TGFβ signaling and EMT and identifies CD44 as a downstream target. (*A*) RNA isolated from control and hnRNPM knockdown cells was subjected to RNA-seq analysis. Genes showing a ≥2.5-fold change were used for GO analysis, which indicated enrichment of hnRNPM-mediated functions and pathways that are associated with EMT. (*B*) Immunoblot analysis of p-Smad3 levels in control and hnRNPM shRNA-expressing cells stimulated with 5 ng/mL TGFβ for 30 min showing that silencing hnRNPM inhibits TGFβ-induced Smad3 phosphorylation. (*C*) GSEA in control versus hnRNPM knockdown (KD) cells showing enrichment of a published EMT gene signature (*top*; normalized enrichment score [NES] = 1.0) and a TGFβ1 gene signature (*middle* panel; NES = 1.0) as well as negative enrichment of the CDH1 epithelial gene signature (*bottom* panel; NES = 1.08). (*D*) A pie chart showing the distribution of hnRNPM-regulated cassette exon events with a false discovery rate (FDR) cutoff of <5%. (*E*) Images of RT–PCR results depicting validation of hnRNPM splicing targets. The *top* and *bottom* bands in each image indicate the included and skipped splicing product, respectively. (*F*) Venn diagram showing overlapping exon-skipping events that are regulated by hnRNPM and associated with EMT. (*G*) Images of RT–PCR results depicting validation of hnRNPM- and EMT-regulated overlapping cassette exons. KD indicates hnRNPM knockdown. epi and mes represent RT–PCR products from epithelial HMLEs and Twist-induced mesenchymal cells, respectively. Quantification of PCR products is presented as the percentage of inclusion, showing that hnRNPM depletion promoted exon inclusion of CD44v3, CD44v10, and SPTAN1 and exon skipping of SEC31A, whereas these exons were preferentially skipped during EMT.

CD44 is a downstream target of hnRNPM

To specifically identify hnRNPM-dependent alternative splicing events that are critical for EMT, exon skipping events were annotated, as these account for the majority of alternative splicing activity. The change in the percentage of exon inclusion (Ψ) caused by hnRNPM knockdown was calculated, which identified 100 hnRNPM-regulated cassette exons. These results demonstrated that hnRNPM promotes skipping of 72 exons and inclusion of 28 exons (Fig. 3D; Supplemental Table S3A). RT–PCR validation of nine cassette exons was performed. Eight confirmed RNA-seq results (Fig. 3E), and one failed in PCR reaction.

Next, our data set was compared with a previously published data set comprised of exons whose splicing is differentially regulated during EMT (Shapiro et al. 2011). This analysis identified seven overlapping cassette exons (overlap $P < 5 \times 10^{-3}$, hypergeometric test) (Fig. 3F; Supplemental Table S3B), among which hnRNPM directed skipping of six exons and inclusion of one exon. While the SEC31A exon, for which hnRNPM promoted inclusion, was skipped during EMT, all of the other six cassette exons for which hnRNPM promoted exon skipping were also observed to be skipped in the EMT data set (Shapiro et al. 2011), suggesting a function for hnRNPM in stimulating the switch of alternative splicing that occurs during EMT. Furthermore, RT–PCR validation of RNA-seq data confirmed all seven hnRNPM-mediated events and five out of seven EMT-associated events (Fig. 3G; Supplemental Fig. S3B,C).

The *CD44* gene encodes nine variable exons clustered between its constitutive exons. Inclusion of one or more of the variable exons generates CD44 variant isoforms (CD44v), whereas skipping all of the variable exons produces CD44s. CD44 exon skipping that generates CD44s was shown to be required for EMT (Brown et al. 2011). Among the seven hnRNPM- and EMT-regulated overlapping cassette exons, two were specific for CD44. Quantitation of the RT–PCR splice products showed that these two CD44 variable exons, v3 and v10, ranked as the second and third most-skipped hnRNPM-regulated exons. hnRNPM knockdown caused a sixfold and fivefold increase in the ratio of inclusion to skipping of v3 and v10, respectively, which indicated that hnRNPM promotes skipping of these exons (Fig. 3G; Supplemental Fig. S3C). In addition, the v3 and v10 exons showed the highest changes in exon skipping during EMT, exhibiting a roughly 30-fold and 50-fold decrease in exon inclusion to skipping ratios, respectively (Supplemental Fig. S3C). We suspect that the difference in the magnitude of CD44 exon-skipping changes between hnRNPM knockdown and EMT experiments is due to the fact that certain epithelial splicing factors, such as ESRP1 (see below), that also contribute to CD44 splicing during EMT are absent in the LM2 mesenchymal cells in which hnRNPM was ablated. Because CD44 variable exons are frequently included or skipped as a cluster, the effect of hnRNPM knockdown on other CD44 variable exons was examined. Quantitative RT–PCR (qRT–PCR)

results showed that hnRNPM knockdown increased CD44 v5 and v8 exon inclusion by 12-fold and ninefold, respectively (Supplemental Fig. S3D). Thus, these results imply that hnRNPM favors CD44 exon skipping, generating the mesenchymal-specific CD44s splice isoform.

hnRNPM promotes CD44 exon skipping

To directly demonstrate that hnRNPM promotes CD44 exon skipping, a minigene splicing reporter construct that contains the CD44 v8 exon and its flanking introns, including the I-8 GU-rich elements, was used (Fig. 4A, top panel). Splicing of the v8 exon was gauged using standard RT–PCR followed by gel electrophoresis to measure the

percentage of v8 inclusion (Fig. 4A, bottom panel). In parallel, qRT–PCR with specific primer sets that amplify individual v8-included and -skipped products was performed for quantification of inclusion to skipping ratios (Fig. 4B). Cotransfection of 293 cells with the minigene construct and hnRNPM demonstrated that hnRNPM increased v8 skipping in a dose-dependent manner, resulting in a decrease in v8 inclusion from 94% in control cells to 17% in cells transfected with the highest dosage of hnRNPM (Fig. 4A, bottom panel). qRT–PCR analysis of v8-included and -skipped products confirmed these results (Fig. 4B). Mutating the I-8 GU motifs in the splicing minigene construct reduced hnRNPM's splicing repression, while swapping the 200-nt intronic region, including I-8, with a random 200-nt sequence almost
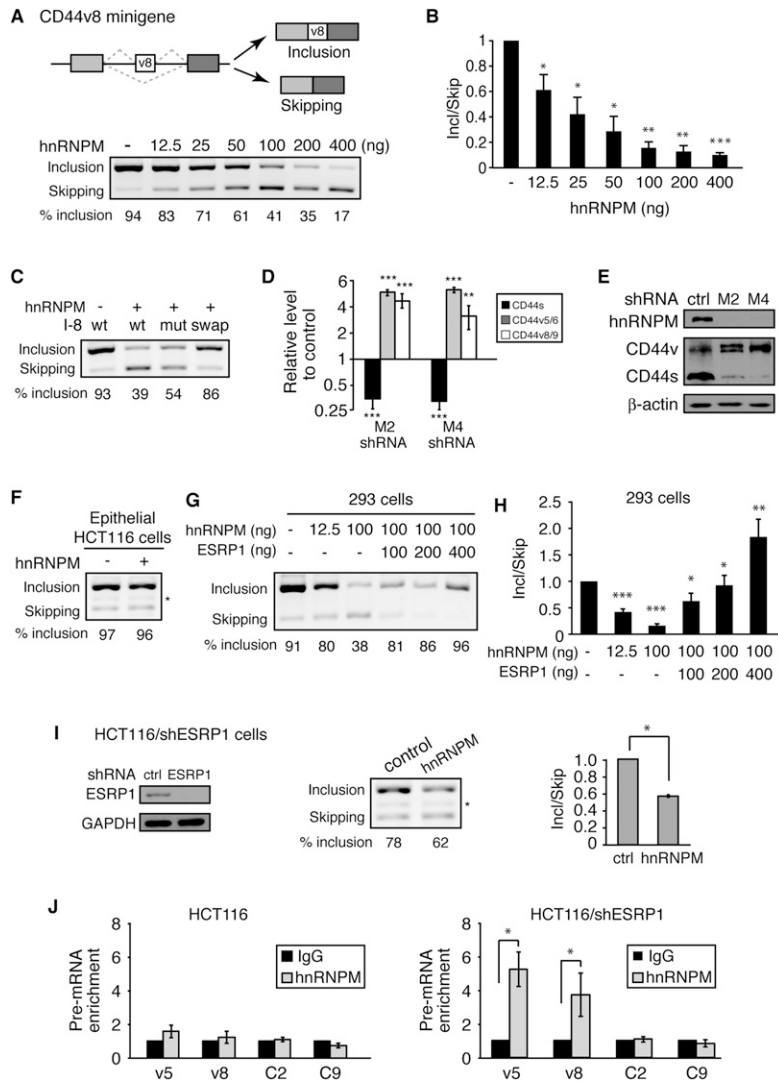
Figure 4. hnRNPM promotes CD44 exon skipping in a mesenchymal-specific manner. (*A, top*) Schematic of the CD44v8 minigene construct. The v8 exon and its flanking introns are inserted between two constitutive exons. (*Bottom*) RT–PCR analysis of RNA harvested from 293 cells cotransfected with the v8 minigene and the indicated amounts of hnRNPM showing that hnRNPM promotes CD44 exon skipping in a dose-dependent manner. (*B*) qRT–PCR analysis of CD44 v8 exon skipping using RNA samples shown in *A*. Ratios of v8 exon inclusion and skipping were plotted. (*C*) RT–PCR analysis of cells transfected with 200 ng of hnRNPM and CD44v8 minigene constructs that contained either mutations within the I-8 *cis*-element (mut) or a 200-nt substitution immediately downstream from the v8 5′ splice site (swap) showing that the I-8 sequence is critical for mediating hnRNPM's exon-skipping activity. (*D*) qRT–PCR analysis showing that hnRNPM knockdown in Mes10A cells promoted CD44 variable exon inclusion. Ratios of CD44 isoforms (CD44s, CD44v5/v6, and CD44v8/v9) between hnRNPM knockdown cells and control cells were plotted. (*E*) Immunoblot analysis indicating an increase in endogenous CD44v expression and a decrease in CD44s level upon silencing hnRNPM in Mes10A cells. (*F*) RT–PCR analysis showing that hnRNPM does not promote CD44 exon skipping in the epithelial HCT116 cells (*G,H*) Cotransfection experiment in 293 cells showing that ESRP1 inhibits hnRNPM's exon-skipping activity. Images of RT–PCR (*G*) and relative ratios of v8 inclusion and skipping from qRT–PCR (*H*) are shown using RNA harvested from cells cotransfected with the CD44v8 reporter construct and the indicated amounts of hnRNPM and ESRP1. (*I*) RT–PCR analysis of CD44v8 minigene transfection in ESRP1-knockdown HCT116 cells showing that removal of ESRP1 in epithelial HCT116 cells restores hnRNPM's ability to promote exon skipping. (*Left* panel) ESRP1 knockdown efficiency. (*Middle* panel) RT–PCR images. (*Right* panel) qRT–PCR results. (*J, left* panel) CLIP analysis showing that hnRNPM is not enriched in CD44 variable exon regions in HCT116 cells. (*Right* panel) ESRP1 knockdown in HCT116 cells restores hnRNPM's association with the v5 and v8 regions of CD44 pre-mRNA. RNA was isolated from immunoprecipitated fractions. Results of RT–PCR analysis were plotted as fold change of PCR signals in different regions of CD44 pre-mRNA. Asterisks in images in *F* and *I* indicate nonspecific PCR products. Error bars in *B*, *D*, *H*, *I*, and *J* indicate SEM. $n \geq 3$. (***) $P < 0.001$; (**) $P < 0.01$; (*) $P < 0.05$.

completely abolished the ability of hnRNPM to promote exon skipping (Fig. 4C, cf. lanes 2 and 4; Supplemental Fig. S4A). Finally, knockdown of endogenous hnRNPM in the mesenchymal Mes10A cells showed significantly decreased CD44s mRNA with a concomitant increase in CD44v5/6 and v8/9 mRNA (Fig. 4D). Immunoblotting confirmed that hnRNPM ablation converted the protein expression from CD44s to CD44v (Fig. 4E). Thus, hnRNPM is necessary and sufficient to stimulate CD44 exon skipping via its interaction with GU-rich motifs located in introns downstream from variable exons.

ESRP1 inhibits the ability of hnRNPM to promote CD44 exon skipping

hnRNPM is a ubiquitously expressed protein in both epithelial and mesenchymal cells. However, epithelial cells predominantly express CD44v, the splice product that is inhibited by hnRNPM. These results suggest that hnRNPM is functionally inactive, at least in terms of CD44 alternative splicing, in epithelial cells. Accordingly, hnRNPM cotransfected with the CD44v8 minigene construct as well as hnRNPM knockdown in HCT116 epithelial cells showed no effects on CD44 exon skipping (Fig. 4F; Supplemental Fig. S4B,C).

ESRP1, a splicing regulator whose expression is restricted to epithelial cells (Warzecha et al. 2009), binds to the same GU-rich cis-acting motifs present in the I-8 sequence that hnRNPM interacts with (Reinke et al. 2012). In direct contrast to hnRNPM's activity, ESRP1 promotes CD44 exon inclusion (Warzecha et al. 2009; Brown et al. 2011; Reinke et al. 2012). Increasing amounts of ESRP1 cotransfected with hnRNPM in 293 cells, which do not express endogenous ESRP1, suppressed hnRNPM's effect on exon skipping of CD44 v8 in a dose-dependent manner (Fig. 4G,H). Removal of ESRP1 in epithelial HCT116 cells enabled hnRNPM to repress

CD44 exon inclusion (Fig. 4I). Finally, hnRNPM's localization and phosphorylation status in epithelial HCT116 and mesenchymal MDA-MB-231 cells showed no detectable differences (Supplemental Fig. S4D,E).

These results suggest that ESRP1 constrains hnRNPM to regulate alternative splicing in a cell type-specific manner. This idea was further supported by crosslinking and immunoprecipitation (CLIP) analysis of hnRNPM occupancy on CD44 pre-mRNA (Supplemental Fig. S4F). While hnRNPM did not display any enrichment in association with the CD44 variable and constitutive exons in the epithelial HCT116 cells (Fig. 4J, left panel; Supplemental Fig. S4G), depletion of ESRP1 in these cells led to an increase in hnRNPM's association with CD44 pre-mRNA, including the variable exon v5 and v8 regions, by fivefold and fourfold, respectively (Fig. 4J, right panel; Supplemental Fig. S4G). Together, these results demonstrate that ESRP1 prevents the activity of hnRNPM on CD44 exon skipping in epithelial cells.

Enforced expression of CD44s rescues the impaired EMT phenotype by hnRNPM depletion

CD44 isoform switching occurs during TGFβ-induced EMT (Brown et al. 2011; Reinke et al. 2012), resulting in decreased expression of CD44v and increased expression of the exon-skipped CD44s isoform (Fig. 5A, cf. left lanes before and after TGFβ treatment). hnRNPM knockdown completely abolished this isoform switching such that the hnRNPM-ablated cells maintained high expression of CD44v even after TGFβ treatment (Fig. 5A; Supplemental Fig. S5A). Importantly, forced expression of the CD44s splice isoform (Fig. 5B) rescued the impaired EMT phenotype that otherwise occurred in the hnRNPM-depleted cells (Fig. 5C–E). The CD44s-re-expressing cells progressed to a mesenchymal phenotype in response to TGFβ treatment, indicated by (1) the acquisition of
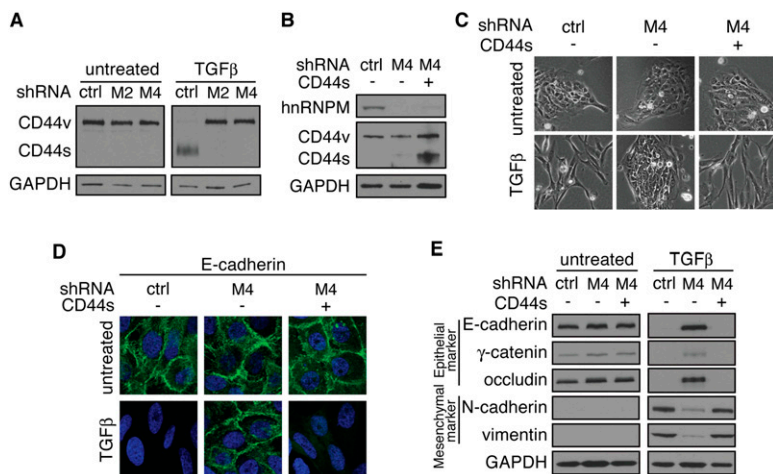
Figure 5. Enforced expression of CD44s rescues the impaired EMT phenotype by hnRNPM depletion. (A) Immunoblot analysis of CD44 splice isoforms showing that silencing hnRNPM inhibits TGFβ-induced CD44 isoform switching from CD44v to CD44s in HMLEs. (B) Immunoblot analysis showing levels of hnRNPM and CD44 isoforms in control, hnRNPM ablated, and CD44s-re-expressed HMLEs. (C) Phase contrast images (10×) illustrating impaired morphological changes in cells expressing hnRNPM M4 shRNA after 18 d of TGFβ treatment and rescued morphological phenotype upon re-expressing CD44s. (D) Immunofluorescence images (63×) indicating that unlike the hnRNPM-depleted cells in which E-cadherin remains localized at cell junctions, CD44s re-expression in hnRNPM knockdown cells results in loss of E-cadherin localization at cell junctions after 18 d of TGFβ treatment, similar to control cells. Green staining indicates E-cadherin. DAPI staining (blue) indicates nuclei. (E) Immunoblot analysis of EMT markers showing that CD44s re-expression rescues the impaired EMT caused by hnRNPM knockdown.

a spindle-like morphology (Fig. 5C); (2) the disappearance of E-cadherin at cell junctions (Fig. 5D); (3) the decrease in expression of the epithelial markers E-cadherin, γ-catenin, and occludin; and (4) the increase in expression of mesenchymal markers N-cadherin and vimentin to levels similar to that observed in the control cells (Fig. 5E). These results demonstrate that CD44s is a critical target of the hnRNPM-dependent splicing program that promotes EMT.
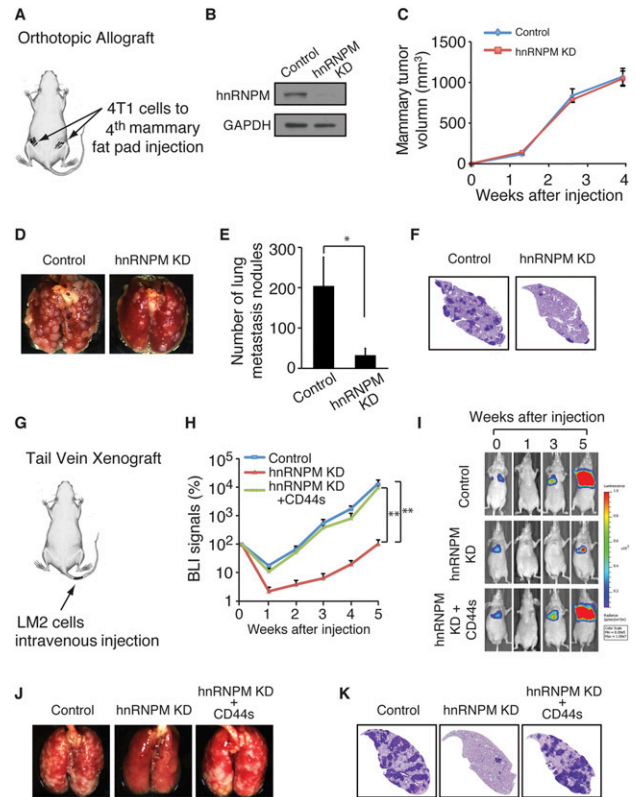
hnRNPM is required for breast cancer lung metastasis in vivo

EMT is a hallmark of tumor metastasis, including breast cancer. To directly address the function of hnRNPM in metastasis, a murine allograft breast cancer lung metastasis model was used. 4T1 breast cancer cells (Aslakson and Miller 1992) were implanted into the fourth mammary fat pad of a syngeneic mouse, resulting in mammary tumor formation and lung metastasis in 4 wk (Fig. 6A). Mice injected with 4T1 cells expressing either control or *hnRNPM* shRNAs (Fig. 6B; Supplemental Fig. S6A) formed tumors at similar growth rates (Fig. 6C; Supplemental Fig. S6B). In contrast, control mice formed large numbers of metastatic nodules in the lungs, whereas a fivefold decrease was observed in mice injected with *hnRNPM* knockdown cells (Fig. 6D, 6E). These results are supported by H&E staining of lung sections of these mice (Fig. 6F). In addition, an experimental lung metastasis assay using intravenous tail injection of 4T1 cells showed that silencing *hnRNPM* caused a sixfold reduction of metastatic lung lesions (Supplemental Fig. S6C–F).

A murine xenograft breast cancer lung metastasis model using the human MDA-MB-231 derivative LM2 cells (Minn et al. 2005) also showed a decrease in lung metastasis in response to hnRNPM ablation, as monitored by bioluminescent imaging (BLI) (Fig. 6G–K). This decrease was observed at the first week after injection, suggesting that hnRNPM is important for tumor cell extravasation and seeding. At week five, a >160-fold reduction of lung metastasis was observed (Fig. 6H–K). In addition, ectopic expression of CD44s in hnRNPM-depleted LM2 cells (Supplemental Fig. S6G) restored the ability to form metastatic lung lesions (Fig. 6H–K). These findings indicate that hnRNPM promotes breast cancer lung metastasis at least partly through its regulation of the CD44s splice isoform.

hnRNPM expression is associated with aggressive breast cancers and correlates with CD44s levels and the mesenchymal status of patient breast tumors

Molecular classification of breast cancers divides them into the luminal A, luminal B, HER2+, normal-like, and basal-like subtypes (Perou et al. 2000; Sorlie et al. 2001). Most of the basal-like tumors are ER-negative. These tumors are highly aggressive and are associated with poor prognosis without effective therapeutics (Rakha et al. 2009; Prat and Perou 2011). GSEA of breast cancer patient gene signatures revealed that gene signatures of basal-like

Figure 6. hnRNPM is required for breast cancer lung metastasis in vivo. (*A*) Schematic of allograft experiment using the mouse breast cancer 4T1 cells. (*B*) Immunoblot analysis showing knockdown efficiency of hnRNPM in 4T1 cells. (*C*) Plot of tumor sizes showing that depletion of hnRNPM does not affect the rate of mammary tumor growth. (*D–F*) hnRNPM knockdown inhibits breast cancer metastasis to the lungs, as indicated by microscopic images of gross lungs (*D*), numbers of metastasis nodules (*E*), and histology images of lungs (*F*) from mice implanted with control or hnRNPM shRNA-expressing 4T1 cells. Error bars in *C* and *E* indicate SEM. *n* = 6. (*) *P* < 0.05. (*G*) Schematic of xenograft experiment using the human breast cancer LM2 cells. (*H–K*) hnRNPM knockdown impairs human breast cancer lung metastasis, and this defect is rescued by enforced expression of CD44s, the isoform produced hnRNPM-mediated splicing. Normalized BLI signals (*H*) and representative BLI images (*I*) at the indicated time points, gross lung lobe images (*J*), and lung H&E (*K*) are shown in mice injected intravenously with LM2 cells expressing control and hnRNPM shRNA or coexpressing hnRNPM shRNA and CD44s cDNA. Five to seven mice were used per group. (**) *P* < 0.01. Mann Whitney *U*-test.

breast cancer, ER-negative breast cancer, and high-grade breast cancer (van 't Veer et al. 2002; Farmer et al. 2005; Sotiriou et al. 2006) were significantly enriched in control versus hnRNPM knockdown cells (Fig. 7A). Furthermore, enrichment of metastatic gene signatures in other cancer types (Bidus et al. 2006; Rickman et al. 2008) was also found in control versus hnRNPM knockdown cells (Supplemental Fig. S7). These data show a strong association between hnRNPM and an aggressive phenotype of breast cancers in patient samples.
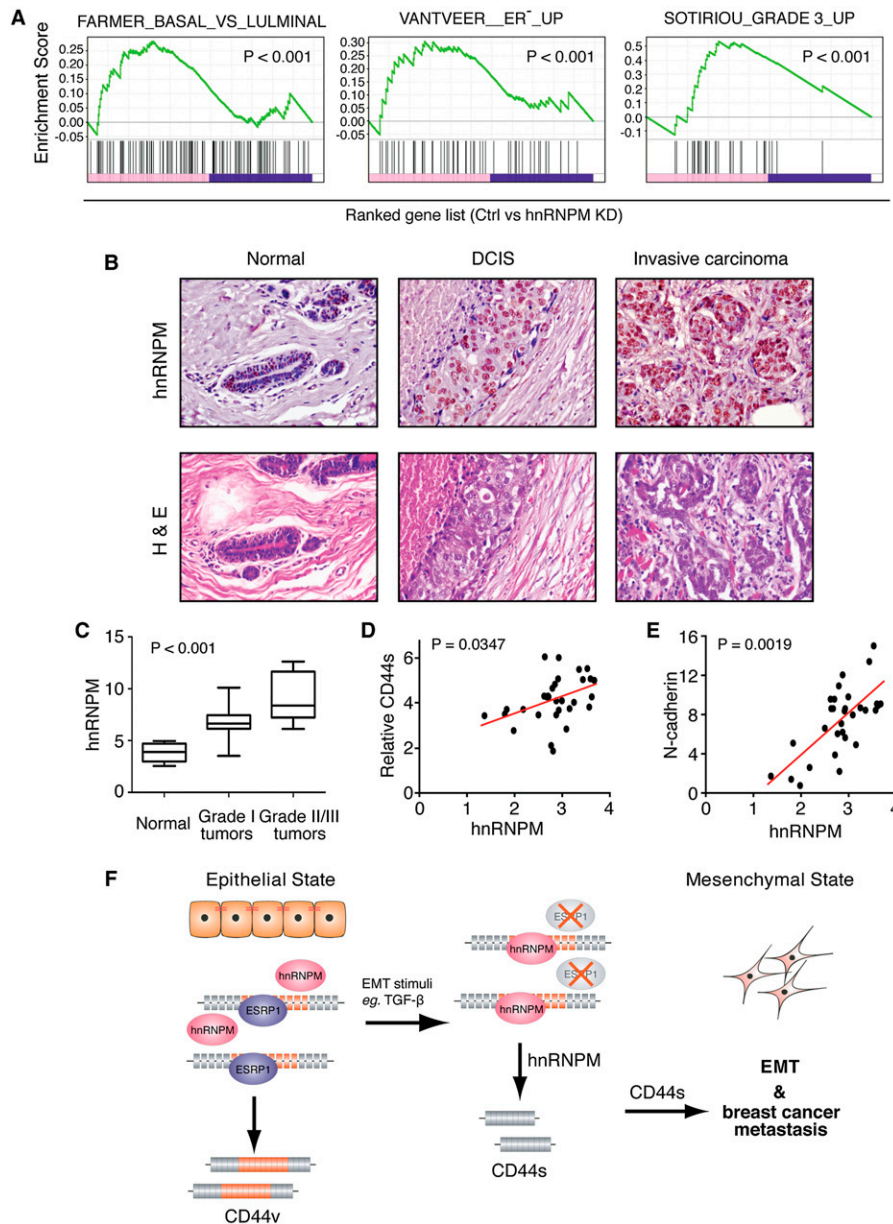
Figure 7. hnRNPM expression is associated with aggressive breast cancers and correlates with levels of CD44s and the mesenchymal status of patient breast tumors. (*A*) GSEA in control compared with hnRNPM knockdown (KD) breast cancer cells showing enrichment of gene signatures from basal-like, ER⁻, and high-grade (grade 3 vs. grade 1) breast tumors. NES = 1.0 in all three gene signatures. (*B*, *top* panel) IHC analysis showing that hnRNPM expression (brown) increases in both DCIS and invasive breast cancer specimens. Blue indicates nuclei. Corresponding H&E staining in the same or nearby regions is shown in the *bottom* panel. (*C*) The relative levels of hnRNPM mRNA in individual human samples from normal breast tissue or breast tumors were compared in different groups according to tumor grade. *P*-values were calculated using a one-way ANOVA test. (*D*,*E*) Relative mRNA levels of hnRNPM and the corresponding levels of CD44s (ratio of CD44s/CD44 total mRNA; *D*), and N-cadherin (*E*) were plotted in log₂ scale in each patient sample. Linear regression-fitted lines are shown in red. *P*-values were obtained from Spearman correlation analysis. (*F*) Model of hnRNPM-mediated regulation of EMT and breast cancer metastasis. See the Discussion for details.

Immunohistochemistry (IHC) analysis of hnRNPM (Fig. 7B) showed that hnRNPM was moderately expressed in benign breast tissues, with a pattern of patchy staining; i.e., positive-stained ductal cells were surrounded by unstained ductal cells, and <30% of the normal luminal/ductal cells were stained positive. In contrast, an increased staining of hnRNPM was observed in ductal carcinoma in situ (DCIS) cells, and invasive breast cancer specimens showed extensive and uniform staining of hnRNPM, ranging from 50% to 100% of all malignant cells.

Assessment of mRNA levels from normal human breast tissues and patient breast tumors revealed increased *hnRNPM* expression, which was significantly

correlated with tumor grade ($P < 0.001$, ANOVA) (Fig. 7C). *hnRNPM* was increased 1.7-fold in low-grade (grade I) and 2.4-fold in advanced-grade (grades II and III) tumors, demonstrating that *hnRNPM* expression is associated with advanced tumor grades. Interestingly, *hnRNPM* expression positively correlated with *CD44s* ($P = 0.0347$, $r = 0.3805$, Spearman correlation) (Fig. 7D), corroborating our observation that hnRNPM promotes the production of CD44s. Furthermore, the positive correlation between *hnRNPM* and the mesenchymal marker *N-cadherin* in these clinical samples ($P = 0.0019$, $r = 0.5353$, Spearman correlation) (Fig. 7E) suggests a role for hnRNPM in promoting EMT in human breast cancer patients. Together with our experimental findings, these results demonstrate that the hnRNPM-mediated splicing program for the production of the CD44s splice isoform is a prevailing molecular pathway that drives EMT and breast cancer metastasis.

Discussion

Abnormal activation of the developmental program EMT is frequently observed in tumor metastasis, a disease complication that remains the major cause of cancer-related death. Therefore, understanding the mechanisms of EMT and metastasis is of critical need to develop effective interventions for the treatment of metastatic cancers. In this study, we used an unbiased RNA-binding biochemical approach and identified the splicing factor hnRNPM as a crucial player in EMT and breast cancer metastasis. Our data demonstrate that hnRNPM is essential for both EMT induction and maintenance. Genome-wide analysis of hnRNPM-regulated downstream pathways suggests an active role for hnRNPM in promoting TGFβ activation, a critical pathway for EMT and metastasis. Using different mouse models of breast cancer lung metastasis, our results clearly show that hnRNPM plays a central role in establishing breast metastatic nodules in the lung. Significantly, hnRNPM expression is associated with gene signatures of aggressive breast cancer, is abnormally elevated in patient breast cancer specimens, and positively correlates with breast tumor mesenchymal status as gauged by N-cadherin expression. Together, these data provide strong evidence indicating that hnRNPM is a crucial regulator driving breast cancer metastasis.

hnRNPM has long been known as a splicing factor and was well characterized for binding GU-rich RNA *cis*-elements (Swanson and Dreyfuss 1988; Datar et al. 1993; Hovhannisyan and Carstens 2007; Huelga et al. 2012). However, the biological function of its RNA-binding activity has not been reported. In this study, we demonstrate that hnRNPM exerts its central function in EMT and metastasis by directing the switch in alternative splicing that occurs during EMT. Notably, our results establish the hnRNPM–CD44s splice isoform axis as a novel molecular pathway that drives EMT and tumor metastasis. Consistent with this notion, our clinical sample analysis revealed a significant positive correlation between hnRNPM and CD44s in breast cancer speci-

mens. It is interesting to note that, in addition to its activity as a splicing regulator, hnRNPM was reported as a cell surface receptor expressed in Kupple macrophages that interacts with carcinoembryonic antigen produced by colorectal cancer cells, which is potentially important for tumor metastasis (Thomas et al. 2011). Thus, hnRNPM may function through multiple mechanisms to promote tumor metastasis.

CD44 is a cell surface molecule and a widely used cancer stem cell marker in various cancer types (Al-Hajj et al. 2003; Jin et al. 2006; Dalerba et al. 2007; Li et al. 2007; Liu et al. 2007; Prince et al. 2007). Isoforms of CD44 produced by alternative splicing have been implicated in different cellular pathways. The CD44s isoform plays an indispensable role in EMT. It potentiates PI3K/Akt signaling and cell death resistance (Brown et al. 2011; Herishanu et al. 2011; Mima et al. 2012; Fedorchenko et al. 2013) and is required for bone metastasis of breast cancer (Hiraga et al. 2013). On the other hand, the CD44v isoform, in particular the v6 isoform, facilitates activation of Ras/MAPK signaling and maintains a proliferative cellular state via a positive feedback loop regulation (Matter et al. 2002; Cheng and Sharp 2006; Cheng et al. 2006). These results highlight the diverse contributions of different CD44 splice isoforms in signaling pathways and cancer processes. Our findings that CD44 isoform switching is controlled by hnRNPM during tumor metastasis functionally prove the concept that the splicing regulatory network serves as a crucial mechanism for cancer phenotypes.

At the molecular level, hnRNPM directly binds to CD44 pre-mRNA at GU-rich sequences and stimulates skipping of CD44 variable exons. Our results further demonstrate that hnRNPM is inactive for CD44 exon skipping in epithelial cells due to ESRP1 antagonization, which is likely through competing binding to the same GU-rich motifs in the CD44 introns. This antagonization in turn converts the ubiquitously expressed hnRNPM to a mesenchymal-specific splicing regulator. We suspect that ESRP1 exhibits a higher binding affinity than that of hnRNPM, precluding hnRNPM binding to CD44 pre-mRNA. While warranting further investigation, this view is supported by our CLIP results showing that hnRNPM does not bind to CD44 pre-mRNA in epithelial cells where ESRP1 exists but does bind when ESRP1 is depleted. These data together with previous findings suggest a working model for hnRNPM's mesenchymal splicing activity during EMT (Fig. 7F). In epithelial cells, ESRP1 outcompetes hnRNPM to bind to CD44 pre-mRNA and stimulates inclusion of CD44 variable exons, resulting in the production of CD44v (Warzecha et al. 2009; Brown et al. 2011; Reinke et al. 2012). When cells are exposed to EMT stimuli such as TGFβ, ESRP1 is transcriptionally repressed (Horiguchi et al. 2012; Reinke et al. 2012). hnRNPM then promotes CD44 isoform switching by activating its exon skipping, thus driving the EMT program. This mode of regulation illustrates a novel model in which the ubiquitously expressed hnRNPM exerts a restricted cell-type activity that allows for precise control of the splice isoform switching that occurs during EMT. This restricted activity is accomplished

at least in part by competition with the cell type-specific splicing factor ESRP1, which binds to the same *cis*-regulatory RNA elements. The mechanism presented here is distinguished from previously reported tissue-specific splicing factors, such as the neuronal-specific NOVA, nPTB, and the brain- and muscle-specific RBFOX1 and RBFOX2 (Buckanovich et al. 1993; Markovtsov et al. 2000; Polydorides et al. 2000; Jin et al. 2003; Underwood et al. 2005), in which tissue-specific activities are largely controlled by their differential expression patterns in various tissue types. Furthermore, our genome-wide analysis of hnRNPM-mediated pathways revealed that TGFβ signaling is significantly stimulated by hnRNPM. Since TGFβ activation suppresses the expression of ESRP1, which in turn promotes the activity of hnRNPM, these results suggest a complex gene regulatory network in which the early phase of TGFβ activation enables hnRNPM's activity. Consequently, hnRNPM continues to stimulate the TGFβ signaling, amplifying the net activity that results in EMT and tumor metastasis.

Investigation of the mechanisms of breast cancer metastasis has been largely focused on pathways mediated by transcription factors, kinases, and cytokines. Increasing attention has recently been paid to the splicing regulatory mechanisms that might influence metastasis-associated pathways. Previous bioinformatic analysis suggested a role for the splicing factor RBFOX2 in promoting mesenchymal-specific splicing (Venables et al. 2013). Experimental approaches showed contradictory results, which demonstrated an active role for RBFOX2 in ENAH epithelial-specific splicing and an inhibitory role in FGFR2 mesenchymal-specific splicing (Mauger et al. 2008; Dittmar et al. 2012; Lovci et al. 2013). Given the importance of the hnRNPM-mediated pathway in breast cancer metastasis, future investigation into whether these and other splicing regulators functionally contribute to EMT and metastasis will lead to better understanding of the splicing regulatory pathways and mechanisms in cancer metastasis.

In conclusion, dysregulation of alternative splicing has been increasingly recognized in cancer-related pathways (Valacca et al. 2010; Anczukow et al. 2012; Das et al. 2012; Bonomi et al. 2013). It is thus critical to investigate the functional significance of splicing regulation in the context of cancer. The present study demonstrates that the splicing regulator-mediated switch in alternative splicing serves as an essential mechanism for TGFβ-induced EMT and breast cancer metastasis. Furthermore, our results describe a novel molecular mechanism underlying the competition between two splicing regulators (hnRNPM and ESRP1) that is critical for restricting the activity of hnRNPM to mesenchymal cells and for the production of the CD44s splice isoform favoring EMT and metastasis. Since TGFβ signaling suppresses ESRP1 expression and hnRNPM potentiates TGFβ activity, in future studies, it will be exciting to investigate whether the TGFβ signaling network is sustained by the hnRNPM-mediated splicing program, conferring breast cancer metastasis through a positive feedback circuit.

Materials and methods

Cell lines and EMT induction

Maintenance of the immortalized HMLEs HMLE and HMLE/Twist-ER; induction of EMT by TGFβ or tamoxifen in HMLE or HMLE/Twist-ER cells, respectively; and maintenance of the TGFβ-induced mesenchymal MCF10A cells (Mes10A) were described previously (Brown et al. 2011). The human bronchioalveolar carcinoma cell line H358 was grown in RPMI 1640 supplemented with 10% FBS, L-glutamine, penicillin, and streptomycin and induced to undergo EMT by 10 ng/mL TGFβ treatment for 16–18 d. The human breast carcinoma cell line MDA-MB-231, embryonic kidney cell line 293, and colorectal carcinoma cell line HCT116 were grown in DMEM supplemented with 10% FBS, L-glutamine, penicillin, and streptomycin. The murine 4T1 breast cancer cells were grown in RPMI supplemented with 5% FBS, 5% CS, nonessential amino acids, L-glutamine, penicillin, and streptomycin.

Antibodies

Antibodies used for Western blotting and immunofluorescence assays were as follows: hnRNPM (OriGene), hnRNPF (Santa Cruz Biotechnology, Inc.), hnRNPH1/H2 (Santa Cruz Biotechnology, Inc.), p-Smad3 (Cell Signaling), Smad3 (Cell Signaling), CD44 (R&D Systems and Santa Cruz Biotechnology, Inc.), E-cadherin (Cell Signaling Technology), γ-catenin (Cell Signaling Technology), occludin (Santa Cruz Biotechnology, Inc.), N-cadherin (BD), vimentin (NEOMarkers), β-actin (Sigma-Aldrich), and GAPDH (Millipore). The ESRP1 antibody was a gift from Dr. R. Carstens (University of Pennsylvania).

RNA pull-down assay

Biotin-labeled RNA oligonucleotides were synthesized by Integrated DNA Technologies. The I-8 sequence is 5′-biotin-GCUUUGGUGGUGGAAUGGUGCUAUGUGG-3′. The I-8 mutant sequence is 5′-biotin-GCUUCGAUCCUAGAACAGAGCUAUCUCG-3′ (mutated residues underlined). Each biotin-labeled RNA oligonucleotide (10 μL at 40 μM) was first immobilized onto 50 μL of streptavidin beads (50% slurry; Pierce) in a final volume of 300 μL of 1× binding buffer (20 mM Tris, 200 mM NaCl, 6 mM EDTA, 5 mM potassium fluoride, 5 mM β-glycerophosphate, 2 μg/mL aprotinin at pH 7.5) for 2 h at 4°C followed by incubation with 200 μg of MDA-MB-231 nuclear extract in 1× binding buffer in a final volume of 400 μL for 2 h at 4°C. The beads were washed three times with binding buffer, resuspended in 60 μL of 2× sodium dodecyl sulfate sample buffer, and boiled for 5 min. Ten microliters of the sample was analyzed under denaturing conditions on a 10% SDS gel and detected by silver staining or Western blot analysis.

Liquid chromatography-tandem MS (LC-MS/MS) analysis

Proteins isolated by RNA pull-down assay using biotin-labeled RNA I-8 or I-8 mutant oligonucleotides in MDA-MB-231 nuclear extract were identified by LC-MS/MS as described previously (Xu et al. 2009).

RNA-seq data analysis

Paired-end RNA-seq reads were mapped to human genome (hg19) and transcriptome (Ensembl release 68) using a method described previously (Bahn et al. 2012; Li et al. 2012). Only uniquely mapped pairs of reads were retained for analysis of alternative splicing.

Differentially spliced exons across samples were identified using the MATS package (Shen et al. 2012). Exons with inclusion levels differing by at least 5% at a false discovery rate (FDR) cutoff of 5% were defined as differentially spliced exons across samples. RNA-seq data have been deposited in Gene Expression Omnibus under the accession number GSE57243.

CD44 splicing reporter minigene assays

CD44 splicing reporter minigene assays were performed in 293 and HCT116 cells. Briefly, 2.25×10^5 cells were seeded in a 24-well plate 24 h before transfection using Lipofectamine 2000 (Invitrogen). Different amounts of hnRNPM were cotransfected with the CD44 v8 minigene construct. Cells were collected 24 h after transfection, and RNA was isolated using an RNA isolation kit (Omega Bio-Tek) for subsequent RT–PCR analysis.

Animal studies

All animal experiments were performed in compliance with guidelines of the Institutional Animal Care and Use Committee at Northwestern University. Injections of 4T1 breast tumor cells into the fourth mammary fat pad were performed on 8-wk-old female Balb/C mice. Tumor cells were injected bilaterally into the fourth mammary gland. Mice were sacrificed 4 wk after tumor cell injection. Mammary tumors were removed and weighed individually. Lungs were inflated with paraformaldehyde and fixed. H&E staining of lung sections was used for counting metastatic lung lesions. Intravenous injections of the LM2 breast cancer cells were performed on 6-wk-old female Balb/c nude mice. Development of metastasis in lungs was monitored by BLI with the IVIS spectrum (Caliper LifeSciences) and analyzed by measuring photon flux of the region of interest. Mice were sacrificed 5 wk after tumor cell injection, lungs were inflated with paraformaldehyde, and lung metastasis nodules were counted under a dissection microscope.

Statistics

Statistical analyses, including two-tailed, unpaired Student's *t*-tests, were done in Excel. One-way ANOVA and Spearman correlation were performed using GraphPad Prism software. Mann-Whitney *U*-tests were performed to analyze nonnormally distributed data sets. *P*-values <0.05 were considered significant. $P \leq 0.05$ (*), $P \leq 0.01$ (**), and $P \leq 0.001$ (***) were indicated.

Acknowledgments

We thank A. Lemond (University of Dundee) for providing the hnRNPM plasmid, R. Carstens (University of Pennsylvania) for the ESRP1 antibody, J. Yang (University of California at San Diego) and Y. Kang (Princeton University) for breast cancer cell lines, and Wensheng Liu (Northwestern University Cell Imaging Facility) for imaging. This research was supported in part by grants from the American Cancer Society (RSG-09-181 to J.P.), the US National Institutes of Health (R01HG006264 and U01HG007013 to X.X.), and the American Cancer Society (RSG-09-252), the US National Institutes of Health (R01GM110146 and R01CA182467), the H Foundation, the A Sister's Hope Foundation, and the Lynn Sage Breast Cancer Research Foundation and Scholar Award (to C.C.).

References

Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. 2003. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci* **100:** 3983–3988.

Anczukow O, Rosenberg AZ, Akerman M, Das S, Zhan L, Karni R, Muthuswamy SK, Krainer AR. 2012. The splicing factor SRSF1 regulates apoptosis and proliferation to promote mammary epithelial cell transformation. *Nat Struct Mol Biol* **19:** 220–228.

Aslakson CJ, Miller FR. 1992. Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor. *Cancer Res* **52:** 1399–1405.

Bahn JH, Lee JH, Li G, Greer C, Peng G, Xiao X. 2012. Accurate identification of A-to-I RNA editing in human by transcriptome sequencing. *Genome Res* **22:** 142–150.

Bidus MA, Risinger JI, Chandramouli GV, Dainty LA, Litzi TJ, Berchuck A, Barrett JC, Maxwell GL. 2006. Prediction of lymph node metastasis in patients with endometrioid endometrial cancer using expression microarray. *Clin Cancer Res* **12:** 83–88.

Bonomi S, di Matteo A, Buratti E, Cabianca DS, Baralle FE, Ghigna C, Biamonti G. 2013. HnRNP A1 controls a splicing regulatory circuit promoting mesenchymal-to-epithelial transition. *Nucleic Acids Res* **41:** 8665–8679.

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.

Buckanovich RJ, Posner JB, Darnell RB. 1993. Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. *Neuron* **11:** 657–672.

Busch A, Hertel KJ. 2012. Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdiscip Rev RNA* **3:** 1–12.

Cheng C, Sharp PA. 2006. Regulation of CD44 alternative splicing by SRm160 and its potential role in tumor cell invasion. *Mol Cell Biol* **26:** 362–370.

Cheng C, Yaffe MB, Sharp PA. 2006. A positive feedback loop couples Ras activation and CD44 alternative splicing. *Genes Dev* **20:** 1715–1720.

Dalerba P, Dylla SJ, Park IK, Liu R, Wang X, Cho RW, Hoey T, Gurney A, Huang EH, Simeone DM, et al. 2007. Phenotypic characterization of human colorectal cancer stem cells. *Proc Natl Acad Sci* **104:** 10158–10163.

Das S, Anczukow O, Akerman M, Krainer AR. 2012. Oncogenic splicing factor SRSF1 is a critical transcriptional target of MYC. *Cell reports* **1:** 110–117.

Datar KV, Dreyfuss G, Swanson MS. 1993. The human hnRNP M proteins: identification of a methionine/arginine-rich repeat motif in ribonucleoproteins. *Nucleic Acids Res* **21:** 439–446.

De Craene B, Berx G. 2013. Regulatory networks defining EMT during cancer initiation and progression. *Nat Rev Cancer* **13:** 97–110.

Dittmar KA, Jiang P, Park JW, 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.

Farmer P, Bonnefoi H, Becette V, Tubiana-Hulin M, Fumoleau P, Larsimont D, Macgrogan G, Bergh J, Cameron D, Goldstein D, et al. 2005. Identification of molecular apocrine breast tumours by microarray analysis. *Oncogene* **24:** 4660–4671.

Fedorchenko O, Stiefelhagen M, Peer-Zada AA, Barthel R, Mayer P, Eckei L, Breuer A, Crispatzu G, Rosen N, Landwehr T, et al. 2013. CD44 regulates the apoptotic response and promotes disease development in chronic lymphocytic leukemia. *Blood* **121:** 4126–4136.

Han SP, Tang YH, Smith R. 2010. Functional diversity of the hnRNPs: past, present and perspectives. *Biochem J* **430:** 379–392.

Herishanu Y, Gibellini F, Njuguna N, Hazan-Halevy I, Farooqui M, Bern S, Keyvanfar K, Lee E, Wilson W, Wiestner A. 2011. Activation of CD44, a receptor for extracellular matrix components, protects chronic lymphocytic leukemia cells from spontaneous and drug induced apoptosis through MCL-1. *Leuk Lymphoma* **52:** 1758–1769.

Hiraga T, Ito S, Nakamura H. 2013. Cancer stem-like cell marker CD44 promotes bone metastases by enhancing tumorigenicity, cell motility, and hyaluronan production. *Cancer Res* **73:** 4112–4122.

Horiguchi K, Sakamoto K, Koinuma D, Semba K, Inoue A, Inoue S, Fujii H, Yamaguchi A, Miyazawa K, Miyazono K, et al. 2012. TGF-β drives epithelial–mesenchymal transition through δEF1-mediated downregulation of ESRP. *Oncogene* **31:** 3190–3201.

Hovhannisyan RH, Carstens RP. 2007. Heterogeneous ribonucleoprotein m is a splicing regulatory protein that can enhance or silence splicing of alternatively spliced exons. *J Biol Chem* **282:** 36265–36274.

Huelga S, Vu A, Arnold J, Liang T, Liu P, Yan B, 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 Reports* **1:** 167–178.

Jin Y, Suzuki H, Maegawa S, Endo H, Sugano S, Hashimoto K, Yasuda K, Inoue K. 2003. A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG. *EMBO J* **22:** 905–912.

Jin L, Hope KJ, Zhai Q, Smadja-Joffe F, Dick JE. 2006. Targeting of CD44 eradicates human acute myeloid leukemic stem cells. *Nat Med* **12:** 1167–1174.

Li C, Heidt DG, Dalerba P, Burant CF, Zhang L, Adsay V, Wicha M, Clarke MF, Simeone DM. 2007. Identification of pancreatic cancer stem cells. *Cancer Res* **67:** 1030–1037.

Li G, Bahn JH, Lee JH, Peng G, Chen Z, Nelson SF, Xiao X. 2012. Identification of allele-specific alternative mRNA processing via transcriptome sequencing. *Nucleic Acids Res* **40:** e104.

Liu R, Wang X, Chen GY, Dalerba P, Gurney A, Hoey T, Sherlock G, Lewicki J, Shedden K, Clarke MF. 2007. The prognostic role of a gene signature from tumorigenic breast-cancer cells. *N Engl J Med* **356:** 217–226.

Lovci MT, Ghanem D, Marr H, Arnold J, Gee S, Parra M, Liang TY, Stark TJ, Gehman LT, Hoon S, et al. 2013. Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. *Nat Struct Mol Biol* **20:** 1434–1442.

Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, Brooks M, Reinhard F, Zhang CC, Shipitsin M, et al. 2008. The epithelial–mesenchymal transition generates cells with properties of stem cells. *Cell* **133:** 704–715.

Markovtsov V, Nikolic JM, Goldman JA, Turck CW, Chou MY, Black DL. 2000. Cooperative assembly of an hnRNP complex induced by a tissue-specific homolog of polypyrimidine tract binding protein. *Mol Cell Biol* **20:** 7463–7479.

Matter N, Herrlich P, Konig H. 2002. Signal-dependent regulation of splicing via phosphorylation of Sam68. *Nature* **420:** 691–695.

Matunis MJ, Xing J, Dreyfuss G. 1994. The hnRNP F protein: unique primary structure, nucleic acid-binding properties, and subcellular localization. *Nucleic Acids Res* **22:** 1059–1067.

Mauger DM, Lin C, Garcia-Blanco MA. 2008. hnRNP H and hnRNP F complex with Fox2 to silence fibroblast growth factor receptor 2 exon IIIc. *Mol Cell Biol* **28:** 5403–5419.

McBryan J, Howlin J, Kenny PA, Shioda T, Martin F. 2007. ERα-CITED1 co-regulated genes expressed during pubertal mammary gland development: implications for breast cancer prognosis. *Oncogene* **26:** 6406–6419.

Mima K, Okabe H, Ishimoto T, Hayashi H, Nakagawa S, Kuroki H, Watanabe M, Beppu T, Tamada M, Nagano O, et al. 2012. CD44s regulates the TGF-β-mediated mesenchymal phenotype and is associated with poor prognosis in patients with hepatocellular carcinoma. *Cancer Res* **72:** 3414–3423.

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.

Onder TT, Gupta PB, Mani SA, Yang J, Lander ES, Weinberg RA. 2008. Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. *Cancer Res* **68:** 3645–3654.

Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, et al. 2000. Molecular portraits of human breast tumours. *Nature* **406:** 747–752.

Polydorides AD, Okano HJ, Yang YY, Stefani G, Darnell RB. 2000. A brain-enriched polypyrimidine tract-binding protein antagonizes the ability of Nova to regulate neuron-specific alternative splicing. *Proc Natl Acad Sci* **97:** 6350–6355.

Prat A, Perou CM. 2011. Deconstructing the molecular portraits of breast cancer. *Mol Oncol* **5:** 5–23.

Prince ME, Sivanandan R, Kaczorowski A, Wolf GT, Kaplan MJ, Dalerba P, Weissman IL, Clarke MF, Ailles LE. 2007. Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proc Natl Acad Sci* **104:** 973–978.

Rakha EA, El-Sayed ME, Reis-Filho J, Ellis IO. 2009. Pathobiological aspects of basal-like breast cancer. *Breast Cancer Res Treat* **113:** 411–422.

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

Rickman DS, Millon R, De Reynies A, Thomas E, Wasylyk C, Muller D, Abecassis J, Wasylyk B. 2008. Prediction of future metastasis and molecular characterization of head and neck squamous-cell carcinoma based on transcriptome and genome analysis by microarrays. *Oncogene* **27:** 6607–6622.

Sarrio D, Rodriguez-Pinilla SM, Hardisson D, Cano A, Moreno-Bueno G, Palacios J. 2008. Epithelial–mesenchymal transition in breast cancer relates to the basal-like phenotype. *Cancer Res* **68:** 989–997.

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 JW, Huang J, Dittmar KA, Lu ZX, Zhou Q, Carstens RP, Xing Y. 2012. MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data. *Nucleic Acids Res* **40:** e61.

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, et al. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci* **98:** 10869–10874.

Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, Nordgren H, Farmer P, Praz V, Haibe-Kains B, et al. 2006. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* **98:** 262–272.

Swanson MS, Dreyfuss G. 1988. RNA binding specificity of hnRNP proteins: a subset bind to the 3′ end of introns. *EMBO J* **7:** 3519–3529.

Thiery JP, Sleeman JP. 2006. Complex networks orchestrate epithelial–mesenchymal transitions. *Nat Rev Mol Cell Biol* **7:** 131–142.

Thomas P, Forse RA, Bajenova O. 2011. Carcinoembryonic antigen (CEA) and its receptor hnRNP M are mediators of metastasis and the inflammatory response in the liver. *Clin Exp Metastasis* **28:** 923–932.

Thomson S, Petti F, Sujka-Kwok I, Mercado P, Bean J, Monaghan M, Seymour SL, Argast GM, Epstein DM, Haley JD. 2011. A systems view of epithelial–mesenchymal transition signaling states. *Clin Exp Metastasis* **28:** 137–155.

Underwood JG, Boutz PL, Dougherty JD, Stoilov P, Black DL. 2005. Homologues of the *Caenorhabditis elegans* Fox-1 protein are neuronal splicing regulators in mammals. *Mol Cell Biol* **25:** 10005–10016.

Valacca C, Bonomi S, Buratti E, Pedrotti S, Baralle FE, Sette C, Ghigna C, Biamonti G. 2010. Sam68 regulates EMT through alternative splicing-activated nonsense-mediated mRNA decay of the SF2/ASF proto-oncogene. *J Cell Biol* **191:** 87–99.

van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, et al. 2002. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415:** 530–536.

Venables JP, Brosseau JP, Gadea G, Klinck R, Prinos P, Beaulieu JF, 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.

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

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.

Xu P, Duong DM, Peng J. 2009. Systematical optimization of reverse-phase chromatography for shotgun proteomics. *J Proteome Res* **8:** 3944–3950.

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