ESRP1 and ESRP2 Are Epithelial Cell-Type-Specific Regulators of FGFR2 Splicing

Claude C. Warzecha,1,2 Trey K. Sato,3 Behnam Nabet,1 John B. Hogenesch,3,5 and Russ P. Carstens1,2,4,*
1Renal Division, Department of Medicine
2Cell and Molecular Biology Graduate Group
3Department of Pharmacology, Institute for Translational Medicine and Therapeutics
4Department of Genetics
University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA
5Penn Genome Frontiers Institute, University of Pennsylvania, Philadelphia, PA 19104, USA
*Correspondence: russcars@upenn.edu
DOI 10.1016/j.molcel.2009.01.025

SUMMARY
Cell-type-specific expression of epithelial and mesenchymal isoforms of Fibroblast Growth Factor Receptor 2 (FGFR2) is achieved through tight regulation of mutually exclusive exons IIIb and IIIc, respectively. Using an application of cell-based cDNA expression screening, we identified two paralogous epithelial cell-type-specific RNA-binding proteins that are essential regulators of FGFR2 splicing. Ectopic expression of either protein in cells that express FGFR2-IIIc caused a switch in endogenous FGFR2 splicing to the epithelial isoform. Conversely, knockdown of both factors in cells that express FGFR2-IIIb by RNA interference caused a switch from the epithelial to mesenchymal isoform. These factors also regulate splicing of CD44, p120-Catenin (CTNND1), and hMena (ENAH), three transcripts that undergo changes in splicing during the epithelial-to-mesenchymal transition (EMT). These studies suggest that Epithelial Splicing Regulatory Proteins 1 and 2 (ESRP1 and ESRP2) are coordinators of an epithelial cell-type-specific splicing program.

INTRODUCTION
Alternative splicing generates multiple mRNAs from a single gene transcript, greatly expanding proteomic diversity in complex organisms. Regulation of splicing is achieved by auxiliary cis-elements that bind regulatory proteins that enhance or silence splicing of adjacent exons (Black, 2003; Matlin et al., 2005). Most known splicing regulators are RNA-binding proteins (RBPs), such as the SR and hnRNP family of proteins, that are expressed fairly ubiquitously, albeit with some differences in expression between tissues (Hanamura et al., 1998). Modulation of these proteins’ activity and subcellular localization by post-translational modifications contributes to cell-type-specific splicing decisions (Allemand et al., 2005; Stamm, 2008). Posttranscriptional regulation of the mRNAs encoding RBPs by microRNAs can also influence cell-type-specific splicing decisions (Boutz et al., 2007; Makeyev et al., 2007). Furthermore, alternatively spliced exons are subject to combinatorial control by numerous splicing regulators with both negative and positive effects on splicing (Black, 2003; Smith and Valcarcel, 2000). A very limited number of mammalian splicing factors have been identified that are cell-type specific. Nonetheless, the identification of additional splicing regulatory proteins with distinct cell-type-specific differences in expression remains an elusive goal.

An alternative splicing event with well-defined cell-type specificity and functional consequences is the choice between mutually exclusive exons IIIb and IIIc of fibroblast growth factor receptor 2 (FGFR2) (Figure 1A). The FGFR2-IIIb splice variant is exclusive to epithelial cells, while FGFR2-IIIc is mesenchymal, and the resulting receptors have distinct differences in ligand-binding specificity (Orr-Urtreger et al., 1993; Zhang et al., 2006). The compartment-specific expression of these FGFR2 splice variants and their ligands is essential for regulation of cell proliferation and differentiation during development (Eswaran et al., 2005). Previous studies identified a number of auxiliary cis-elements and RBPs that regulate FGFR2 exon IIIb and exon IIIc splicing (Hovhannisyan and Carstens, 2007; Mauger et al. [2008] and references therein). To identify additional splicing factors that promote FGFR2-IIIb expression, we carried out a genome-wide, high-throughput cDNA overexpression screen. Among the previously uncharacterized splicing factors identified using this approach were epithelial cell-type-specific splicing regulators.

RESULTS
Identification of Splicing Regulatory Proteins Using a High-Throughput cDNA Expression Screen
We used luciferase-based reporter minigenes to carry out a high-throughput cDNA expression screen for factors that promote the epithelial pattern of FGFR2 splicing. A heterologous minigene, PKC-neg-40B-IF3-Luc, consists of a 40 nt exon (40B) whose inclusion in spliced transcripts is required for translation of the luciferase coding sequence (Figure 1B). This exon is included in ~3% of spliced transcripts in stably transfected cells. Insertion of an FGFR2 intron 8 fragment (Intron Fragment 3, IF3) confers epithelial cell-type-specific splicing enhancement to the heterologous exon and a corresponding increase in luciferase activity.
Epithelial-Specific FGFR2 Splicing Regulators

(Hovhannisyan et al., 2006 and data not shown). As 293T cells express the mesenchymal FGFR2-IIIc isoform, we generated a 293T cell clone that stably expresses the PKC-neg-40B-IF3-Luc minigene and employed a reverse transfection approach to screen ~15,000 cDNAs within the Mammalian Gene Collection (MGC) library for factors that promote exon 40B inclusion (Figure 1C). Using a median 6-fold increase in luciferase activity greater than 14,000 human or mouse cDNAs in the screen, corresponding to 22 unique genes (Table 1). Four of these hits were false positives, but the remaining 18 cDNAs enhanced luciferase activity and increased exon 40b splicing (see Figures S1A and S1B available online).

Figure 1. Identification of FGFR2 Splicing Regulators Rbm35a, Esrp1, and Rbm35b, Esrp2, in a High-Throughput cDNA Expression Screen

(A) Schematic of the FGFR2 protein (top) and the pre-mRNA in the region encoding exons IIlb and IIIc. IG, immunoglobulin-like domains; TM, transmembrane domain; TK, tyrosine kinase domains. (B) Schematic of the reporter minigene and screening strategy. An FGFR2 intron 8 fragment (IF3) required for IIlb inclusion in epithelial cells is positioned downstream of a heterologous exon. Auxilary intronic cis-elements are indicated by hatched boxes. (C) Flowchart summarizing the cell-based screen. (D) Schematic of the minigenes stably expressed in 293T-clone 3 (top). Cells were transiently transfected with empty vector (EV) or cDNAs for Fox-1, Rbm35a (Esrp1), Rbm35b (Esrp2), Rbm38, or Fusilli, and exon inclusion was determined by RT-PCR. Average percentages of exon 40B inclusion with standard deviations compiled from three experiments are indicated below a representative gel. The Rbm35a, Rbm35b, and Rbm38 cDNAs represented here are the MGC clones from the screening collection. (E) 293T cells were transduced with pMXs-based retroviruses containing cDNAs for Rbm35a (Esrp1), Rbm35b (Esrp2), Fusilli, or EGFP. The Rbm35a cDNA used here is a full-length clone containing alternative exons 14 and 15 (2A). FGFR2 splice variant analysis was determined by an RT-PCR protocol in which products are digested with Aval (A) or HincII (H), which specifically cut exon IIlb- and exon IIlc-containing products, respectively. U indicates undigested PCR products. Percent exon IIlb inclusion is calculated as Exon IIlb Product (lane H)/Exon IIlb + Exon IIlc (lane A) product. The average percentage of IIlb inclusion with standard deviations was compiled from three experiments. (F) Expression of Esrp1, Esrp2, and Fusilli protein in cells from (E) as determined by immunoblotting with anti-FLAG and anti-ESRP1 (RBM35A) and anti-ESRP2 (RBM35B) antibodies. β-actin is used as a loading control. Arrows indicate the respective protein bands. The asterisk represents a variant of Esrp2 suspected to result from use of a alternative translational start site.

Satisfyingly, 15 of the 18 validated hits encoded RBPs, including eight known mammalian splicing regulators.

Two Paralogous Splicing Factors Identified in the Screen, Rbm35a and Rbm35b, Require the FGFR2 Auxiliary cis-Element ISE/ISS-3 to Modulate Splicing of the Reporter Minigene

Subsequent validations showed that most of the factors identified in the screen did not require FGFR2 intron 8 sequence elements to enhance exon 40B splicing (Figure S1B). Among the seven cDNAs that demonstrated a dependence on FGFR2 intron 8 sequences, the two with the most robust enhancement...
of splicing were two RNA Recognition Motif (RRM)-containing proteins, Rbm35b and Rbm38. A paralog of Rbm35b, Rbm35a, was also a hit in the screen, and we therefore directed further investigation into these three gene products. Rbm38 (RNPC1) has not been implicated in mammalian splicing, although a C. elegans ortholog, SUP-12, regulates splicing of the worm FGFR Egl-15 (Kuroyanagi et al., 2007). Rbm35a and Rbm35b are also previously uncharacterized mammalian splicing factors, but are also orthologous with a C. elegans splicing regulator, sym-2 (Barberan-Soler and Zahler, 2008). In previous work, we characterized an auxiliary cis-element, ISE/ISS-3 (Intronic Splicing Enhancer/Intronic Splicing Silencer-3), that functions specifically in epithelial cell types to enhance splicing of the upstream exon IIIb and silence the downstream exon IIIc (Hovhannisyan and Carstens, 2005). ISE/ISS-3 is located in intron 8 downstream of a UGCAUG motif that is a binding site for the Fox family of splicing regulators. Binding of Fox-2 to this element has been shown to play an important role in FGFR2 splicing regulation (Baraniak et al., 2006). In order to determine whether Rbm35a, Rbm35b, or Rbm38 require ISE/ISS-3 and/or the UGCAUG motif, we transfected these cDNAs in a 293T cell clone containing a previously described mRFP-based heterologous minigene that also contains the full FGFR2 Intron Fragment 3 (IF3) (Newman et al., 2006). The same cell clone also stably expresses a control EGFP-based minigene in which ISE/ISS-3 and the UGCAUG motif have been deleted (Intron Fragment 5: IF5).

As predicted, transient transfection of a cDNA encoding Fox-1 enhanced exon inclusion only in the minigene that contained its binding site (Figure 1D). Rbm38 promoted a similar degree of exon 40B splicing from both minigenes, suggesting that it did not bind to or require ISE/ISS-3 or the UGCAUG motif.

### Table 1. cDNAs that Promoted a $\geq$6-Fold Change in Luciferase Activity in the Screen

<table>
<thead>
<tr>
<th>Gene Symbola</th>
<th>Fold Change in Screen</th>
<th>MGC Clone Number</th>
<th>RNA Binding Domain</th>
<th>Validated Change in Splicing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zgpat (1)b</td>
<td>30.4</td>
<td>BC021513.1</td>
<td>CCCH zinc finger, G patch</td>
<td>yes</td>
</tr>
<tr>
<td>Rbm38</td>
<td>26.6</td>
<td>BC006687.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>Esrp1 (Rbm35a)*</td>
<td>25.5</td>
<td>BC031468.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>Zgpat (2)b</td>
<td>14.1</td>
<td>BC027218.1</td>
<td>CCCH zinc finger, G patch</td>
<td>yes</td>
</tr>
<tr>
<td>SFRS3b,c</td>
<td>13.5</td>
<td>BC009914.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>Cugbp2c</td>
<td>12.5</td>
<td>BC026856.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>SEPT9</td>
<td>12.0</td>
<td>BC054004.1</td>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>Vgl4</td>
<td>11.4</td>
<td>BC048841.1</td>
<td>RRM</td>
<td>no</td>
</tr>
<tr>
<td>BOLL</td>
<td>11.0</td>
<td>BC033674.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>Sfrs3b,c</td>
<td>10.6</td>
<td>BC071196.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>Esrp2 (Rbm35b)*</td>
<td>10.6</td>
<td>BC031444.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>Khdrbs1d</td>
<td>10.5</td>
<td>BC005051.1</td>
<td>KH domain</td>
<td>yes</td>
</tr>
<tr>
<td>RBM4b,c</td>
<td>9.5</td>
<td>BC032735.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>Rcc1</td>
<td>9.2</td>
<td>BC019807.1</td>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>Prkaca</td>
<td>8.6</td>
<td>BC054834.1</td>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>Brunol4</td>
<td>8.4</td>
<td>BC048405.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>Sdb</td>
<td>8.4</td>
<td>BC030820.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>Rbm4b,c</td>
<td>8.1</td>
<td>BC019488.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>BRUNOL5c</td>
<td>7.8</td>
<td>BC028101.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>RNPS1</td>
<td>7.7</td>
<td>BC018316.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>Zcchc10</td>
<td>7.4</td>
<td>BC025078.1</td>
<td>CCCH zinc finger</td>
<td>yes</td>
</tr>
<tr>
<td>SNRP70b</td>
<td>6.8</td>
<td>BC001315.1</td>
<td>RRM</td>
<td>yes</td>
</tr>
<tr>
<td>PRKACA</td>
<td>6.8</td>
<td>BC039846.1</td>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>ZGPATb</td>
<td>6.4</td>
<td>BC032612.1</td>
<td>CCCH zinc finger, G patch</td>
<td>yes</td>
</tr>
<tr>
<td>WNT1</td>
<td>6.3</td>
<td>BC005449.1</td>
<td></td>
<td>no</td>
</tr>
<tr>
<td>Pspc1</td>
<td>6.1</td>
<td>BC026772.1</td>
<td>RRM</td>
<td>no</td>
</tr>
<tr>
<td>Zbtb7c</td>
<td>6.1</td>
<td>BC018187.1</td>
<td>C2H2 zinc finger</td>
<td>no</td>
</tr>
<tr>
<td>Zgpat (2)b</td>
<td>6.0</td>
<td>BC027218.1</td>
<td>CCCH zinc finger, G patch</td>
<td>yes</td>
</tr>
</tbody>
</table>

*Esrp1 (Rbm35a) and Esrp2 (Rbm35b) are marked with asterisks.

All caps denotes a human cDNA, and lowercase represents mouse cDNAs.

Genes represented by multiple, independent cDNAs or transfections.

Known regulators of mammalian splicing.

Not tested directly, but inferred from the results obtained with human SFRS3.
and most likely interacts with upstream FGFR2 sequence elements (Figure 1D). However, the ability of Rbm35a and Rbm35b to promote splicing was dependent upon the presence of ISE/ISS-3 and/or the UGCAUG motif to enhance exon 40B splicing. A cDNA encoding the D. melanogaster ortholog for Rbm35a and Rbm35b, Fusilli, achieved a similar splicing outcome.

Ectopic Expression of Rbm35a or Rbm35b Causes a Switch in Endogenous FGFR2 Splicing from the Mesenchymal-to-Epithelial Isoform

Transduction of 293T cells with retroviral vectors encoding FLAG-tagged cDNAs for Rbm35a or Rbm35b induced a substantial switch in splicing of endogenous FGFR2 transcripts from exon IIIc to exon IIIb splicing (Figure 1E). Immunoblotting with anti-FLAG, anti-RBM35A, and anti-RBM35B antibodies confirmed that this switch corresponded with expression of proteins of the size predicted for full-length Rbm35a and Rbm35b (Figure 1F). Transduction with virus expressing a cDNA for D. melanogaster Fusilli also caused a switch in endogenous FGFR2 splicing. The MGC Rbm35a cDNA clone used in the previous transient transfection (Figure 1D) is truncated near the 5' end. In Figure 1E and subsequent studies, we used full-length Rbm35a cDNAs, including two different splice variants, that displayed greater splicing activity than either the truncated clone of Rbm35a or Rbm35b (Figure S2).

Rbm35a and Rbm35b Are Epithelial Cell-Type-Specific Splicing Regulatory Proteins

Analysis of RBM35A and RBM35B mRNA expression in a panel of cell lines available in the lab demonstrated that expression of both genes correlated with the epithelial FGFR2-IIIb isoform with very low to undetectable expression in FGFR2-IIIc expressing cell lines (Figure 2A and Figure S3). Furthermore, microarray data from the NCI60 panel of cell lines revealed that cell lines classified as ‘‘epithelial’’ based on the E-cadherin/vimentin protein ratio expressed substantially higher levels of RBM35A and/or RBM35B than cells classified as ‘‘mesenchymal’’ (Figure S4) (Park et al., 2008). A number of these ‘‘epithelial’’ cell lines have been confirmed to express FGFR2-IIIb in the data presented here as well as in a previous study (Figures S3B and S3G).
Epithelial-Specific FGFR2 Splicing Regulators

(Cha et al., 2008). Epithelial-specific expression of Rbm35a within brain sections was previously shown by in situ hybridization analysis in mice (McKee et al., 2005). To more comprehensively examine Rbm35a mRNA expression, we carried out in situ hybridization analysis of whole P1 and adult mouse tissue sections as well as panels of tissues with defined epithelial cell layers. These studies revealed distinct epithelial-specific expression in diverse tissues and organs with particularly notable levels of expression in skin and gastrointestinal epithelia (Figures 2B and 2C and Figures S5–S8). Evidence that Rbm35a expression is also epithelial specific in mouse embryos is available in an online in situ hybridization database (Visel et al., 2004).

The mammalian ESRPs and their orthologs in chicken, D. melanogaster, and C. elegans contain three RNA Recognition Motifs (RRMs) and display significant phylogenetic sequence conservation within these domains, particularly RRM1 (Figure 2D and Figure S9). Our demonstration that D. melanogaster Fusilli can regulate endogenous FGFR2 splicing provides evidence that this fly protein is a splicing factor. Furthermore, its ability to substitute for the mammalian protein in human cells suggests that these orthologs bind similar RNA target sequences and interact with conserved components of multiprotein splicing regulatory complexes. Epithelial-specific expression of Fusilli in the stromodeum and proctodeum was previously noted by in situ hybridization analysis of fly embryos (Wakabayashi-Ito et al., 2001). Collectively, these data provide strong evidence that these factors are evolutionarily conserved epithelial cell-type–specific splicing proteins. We therefore propose the names Epithelial Splicing Regulatory Proteins 1 and 2 (ESRP1 and ESRP2) to replace the generic gene symbols RBM35A and RBM35B, respectively.

The ESRPs Are Required for Expression of FGFR2-IIIb

To determine whether expression of ESRP1 or ESRP2 is required for expression of epithelial FGFR2-IIIb we performed RNA interference (RNAi) to deplete these gene products in the human prostatic epithelial cell line PNT2, in which we could achieve effective knockdown (Figure 3A). Two separate combinations of effective siRNAs against ESRP1 and ESRP2 caused a significant switch from exon IIIb to exon IIIc splicing. Average percentages of exon IIIb inclusion with standard deviations from three transfections are indicated below a representative RT-PCR.

(A) Quantitative RT-PCR showing efficient knockdown of ESRP1 and ESRP2 in PNT2 cells using two different combinations of siRNAs targeting each transcript compared with control siRNAs against EGFP. Mean expression values from triplicate assays ± SD are shown relative to the EGFP control.

(B) ESRP1 and ESRP2 knockdown by RNAi causes a switch in endogenous FGFR2 from exon IIIb to exon IIIc splicing. Average percentages of exon IIIb inclusion with standard deviations from three transfections are indicated below a representative RT-PCR.

(C) PNT2 cells were transduced with cDNAs for EGFP, mouse Esr1, or Esr2, and subjected to siRNAs against human ESRP1 and ESRP2. FGFR2 splice variant analysis from three experiments is shown along with a representative RT-PCR gel.

(D) Expression of Esr1, Esr2, and Fusilli protein in cells from (C) as determined by immunoblotting with anti-FLAG and anti-ESRP1 (RBM35A) and anti-ESRP2 (RBM35B) antibodies. β-actin is used as a loading control. Arrows indicate the respective protein bands. The asterisk is as described in the legend to Figure 1F.

(E) Depletion of ESRP1 and ESRP2 also causes a switch in splicing of CD44, p120-catenin (CTNND1), and ENAH splicing. The splice variants are indicated by boxes at the right of each gel. The solid triangle in exon 3 or exon 5 of CTNND1 indicates the different translational start sites used in the respective isoforms.

Figure 3. ESRP1 and/or ESRP2 Is Necessary for Expression of FGFR2-IIIb and the Regulation of CD44, CTNND1, and ENAH Splicing

To further establish that this switch was due to the loss of ESRP1 and ESRP2 expression, we carried out a “rescue experiment” with mouse cDNAs for Esr1 and Esr2 that were immune to knockdown by these human siRNAs. PNT2 cells were transduced with an EGFP control or vectors containing the cDNAs encoding FLAG-tagged mouse Esr1 or Esr2. The cells were subsequently transfected with siRNAs
against human ESRP1 and ESRP2 and endogenous FGFR2 isoform analysis was performed. In contrast to the controls, cells transduced with RNAi resistant cDNAs for Esrp1 maintained predominant FGFR2-IIIb expression (Figure 3C). However, the cDNA for Esrp2 only partially preserved FGFR2-IIIb expression when both endogenous factors were depleted. Independent knockdown of ESRP1 in PNT2 cells caused a partial switch toward FGFR2-Illic splicing whereas knockdown of ESRP2 alone caused no change in FGFR2 splicing (data not shown). Thus, at least in these cells, it appears that the preservation of the FGFR2-Illb splicing pathway is more dependent on ESRP1 than ESRP2. Expression of the FLAG-tagged Esrp1 and Esrp2 protein in these experiments was verified by immunoblotting with anti-FLAG antibodies as well as antibodies against each protein (Figure 3D). These data thus provide fairly conclusive evidence that the expression of at least one of these epithelial cell-type-specific protein paralogs is required for FGFR2-Illb expression.

The ESRPs Regulate Splicing of CD44, CTNND1, and ENAH

The profound switch in endogenous FGFR2 splicing from the epithelial to mesenchymal isoform upon depletion of ESRP1 and ESRP2 suggested that these cell-type-specific factors might regulate additional epithelial-specific transcript variants. Inclusion of several “variable” exons of CD44 transcripts, including exons 8–10 (V8-V10), has been shown to be epithelial specific (Ponta et al., 2003). Depletion of ESRP1 and ESRP2 in PNT2 cells resulted in a significant decrease in the inclusion of CD44 exons 8–10 and increase in the standard isoform (CD44s) in which all of the variable exons are skipped (Figure 3E). Delta catenin (CTNND1), also known as p120-Catenin, expresses mesenchymal specific splice variants that contain alternative exons 2 and 3 (Keirsebilck et al., 1998). Skipping of these exons in epithelial cells results in a shorter protein isoform that initiates translation in exon 5. Expression of the mesenchymal p120-Catenin isoform is induced during the epithelial-to-mesenchymal transition (EMT) (Ohkubo and Ozawa, 2004). Knockdown of ESRP1 and ESRP2 in PNT2 cells also induced expression of the mesenchymal isoform of p120-Catenin (Figure 3E). ENAH contains an alternative exon, 11a, that is predominantly included in epithelial cell lines and skipped in mesenchymal cells (Pino et al., 2008). Knockdown of ESRP1 and ESRP2 led to a significant decrease in ENAH exon 11a inclusion (Figure 3E). These three examples of additional targets of the ESRPs suggest that they regulate a larger number of epithelial versus mesenchymal splice variants. Furthermore, they illustrate examples of regulated targets in which they can promote epithelial-specific exons (CD44 and ENAH), silence mesenchymal exons (CTNND1), or both (FGFR2).

The ISE/ISS-3 Auxiliary cis-Element Is a Target Binding Site for ESRP1 and ESRP2

To investigate whether these proteins bind specifically to ISE/ISS-3, we performed UV crosslinking experiments with radiolabeled ISE/ISS-3 RNAs and previously defined functional mutants and control RNAs (Hovhannisyan et al., 2006). The RNAs were incubated with nuclear extracts from 293T cells transiently transfected with cDNAs encoding FLAG-tagged Esrp1 or empty vector. In Esrp1 transfected cells, a band of the size predicted for the FLAG-tagged Esrp1 was crosslinked to the wild-type ISE/ISS-3 (WT), but not a mutant in which GU or UG dinucleotides in the 5’ half of the element were replaced with AC (AC), or an unrelated RNA (BS) (Figures 4A and 4B). Immunoprecipitation with anti-FLAG antibodies confirmed the identity of the crosslinked Esrp1 protein. An RNA sequence containing three tandem copies of the 5’ end of ISE/ISS-3 (3X WT) is nearly equivalent in function to the full-length ISE/ISS-3, and three more discrete GU to AC mutations (3X MT) abrogate splicing activity (Figure 4C) (Hovhannisyan and Carstens, 2007). Crosslinking of overexpressed Esrp1 and Esrp2 was observed with the wild-type, but not the functional, mutant of this sequence, suggesting that they bind specific GU-rich sequence motifs (Figure 4D). To verify that this difference in crosslinking was due to differential binding, we also carried out competition experiments using 3X WT and 3X MT unlabeled “cold” competitor RNAs. These experiments show that crosslinking of Esrp1 to the wild-type sequence is effectively competed by the wild-type, but not the mutant, competitor (Figure 4E). We validated that the overexpressed cDNAs produce Esrp1 and Esrp2 protein by immunoblotting of these extracts with anti-ESRP1 and anti-ESRP2 antibodies (Figure 4F). To determine whether endogenous Esrp1 and Esrp2 protein binds to this sequence, we carried out RNA affinity pull-down experiments using FGFR2-Illb expressing KATO III cells, for which preparation of bulk nuclear extracts is feasible. Proteins that bound to beads containing the 3X WT or 3X MT RNAs, or beads alone, were analyzed by SDS-PAGE gels and immunoblotting using anti-ESRP1 and anti-ESRP2 antibodies. Immunoreactive bands of the expected size for ESRP1 and ESRP2 were identified in pull downs from the wild-type, but not the mutant, 3X ISE/ISS-3 sequence or to beads alone (Figure 4G). Coomassie staining showed that similar amounts of total protein bound to both RNA columns. In the case of ESRP1, the same protein was recognized by two different antibodies, providing additional evidence that this protein represents endogenous ESRP1. We had difficulty unambiguously identifying endogenous ESRP1 or ESRP2 by direct western analysis with these antibodies. These results, therefore, also allowed us to verify ESRP1 and ESRP2 protein expression in an epithelial cell line. It is noteworthy that both of these proteins eluded detection in mass spectrometry analysis of pulldowns from KATO III nuclear extracts using this same assay. We suspect that promiscuous binding by more abundant RNA-binding proteins that bind this sequence masked the identification of peptides corresponding to the ESRPs.

Downregulation of the ESRPs Coincides with Loss of Epithelial Splicing during the EMT, and Ectopic Expression of ESRP1 in Mesenchymal Cells Restores an Epithelial Splicing Program

A switch from FGFR2-Illb to FGFR2-Illc has been shown during the EMT, a developmental process associated in pathophysiological conditions with fibrosis and cancer metastasis (Thiery and Sleeman, 2006). Changes in CD44 and CTNND1 splicing also occur in models of EMT, and the same changes in splicing occur upon knockdown of the ESRPs. We therefore investigated whether ESRP expression is lost in a human mammary epithelial
cell line, HMLE, during the induction of an EMT by transcription factor Twist (Yang et al., 2004). Analysis of FGFR2 splicing following Twist-induced EMT showed a partial switch from FGFR2-IIIb to FGFR2-IIIc that coincided with a morphological EMT, decrease in epithelial markers, and increase in mesenchymal markers (Figures 5A–5D). These changes were associated with downregulation of the mRNAs for ESRP1 and ESRP2 (Figure 5E). We also noted changes in CD44, CTNND1, and ENAH splicing that were similar to those seen after knockdown of ESRP1 and ESRP2 (Onder et al., 2008). Using published microarray data from this experiment, we examined ESRP expression during the EMT along with that of previously published epithelial and mesenchymal markers using gene cluster analysis. We also performed a similar analysis using microarray data from a breast cancer model in which MCF10F cells selected for invasive properties in vitro underwent an EMT. A switch from epithelial-to-mesenchymal CD44 splicing was previously observed in this model of the EMT (Huang et al., 2007). In both analyses, ESRP1 and ESRP2 clustered with epithelial markers and correlated inversely with expression of mesenchymal markers during the EMT (Figure S10). Furthermore, ESRP1 and ESRP2 mRNA expression correlated highly with that of E-Cadherin in the microarray analyses of the NCI60 panel of cell lines (Table S1). These results suggest that a decrease in ESRP expression may be a general feature of the EMT and that the ESRPs are a component of an epithelial gene signature.

The loss of ESRP expression during the EMT suggested that ectopic expression of the ESRPs in mesenchymal cells would

Figure 4. ESRP1 and ESRP2 Bind to GU-Rich Sequences in ISE/ISS-3

(A) At top, the sequence of ISE/ISS-3 (WT) is shown, and the sequences that were mutated to AC or CA are underlined. The resulting mutant sequence (Mut [AC]) is shown at bottom.

(B) Rbm39a binds specifically to ISE/ISS-3 (WT), but not to the mutated sequence (AC) or unrelated RNA (BS). Results of direct crosslinking (Direct XL) and crosslinking followed by IP of the FLAG-tagged proteins are shown.

(C) At top, the sequence of the wild-type 5′ half (43 nt) of ISE/ISS-3 present as 3 tandem copies in the 3× WT sequence. A more discreet set of 3 GU to AC mutations is underlined. The resulting sequence present as three tandem copies in the 3× MT mutant sequence is shown at bottom.

(D) Esrp1 and Esrp2 bind three tandem copies of the 5′ half of ISE/ISS-3 in a sequence-specific manner. UV crosslinking was performed using nuclear extracts from 293T cells transiently transfected with FLAG tagged cDNAs for the Esrp1-2A, Esrp2, or empty vector (EV). Arrows indicate the crosslinked bands. W, 3× WT sequence; M, 3× MT sequence; kD, sizes of molecular weight markers.

(E) UV crosslinking competition using nuclear extracts from 293T cells transiently transfected with Esrp1-2A. Extracts were incubated with radiolabeled 3× WT RNAs and unlabelled 3× WT or 3× MT competitor RNAs with the molar fold-excess of competitor (Fold comp) indicated.

(F) Nuclear extracts from 293T cells transiently transfected with cDNAs for two splice variants of Esrp1, Esrp2, and empty vectors immunoblotted with antibodies against endogenous Esrp1 and Esrp2. Two different antibodies against Esrp1 specifically detected the overexpressed protein.

(G) RNA pull-down assays from KATO III cell nuclear extracts show sequence-specific binding of endogenous ESRP1 and ESRP2 to the 3× WT (W), but not 3× MT (M). Proteins eluted from beads containing the indicated RNAs or beads alone (B) were immunoblotted with antibodies against ESRP1 or ESRP2. A Coomassie stain showing total protein content in the RNA pull-down assay is also shown.
Epithelial-Specific FGFR2 Splicing Regulators

DISCUSSION

We identified two paralogous epithelial cell-type-specific splicing proteins, ESRP1 and ESRP2, that are required for the expression of epithelial FGFR2-IIlb. We also determined that ISE/ISS-3 is an ESRP binding site and mutations that abolish the function of this element abrogate binding. This cis-element restore an epithelial splicing program. To test this hypothesis, we transduced the mesenchymal MDA-MB-231 cell line with virus encoding FLAG-tagged ESRP1 or control EGFP virus. Examination of FGFR2, CD44, CTNND1, and ENAH splicing demonstrated that the expression of ESRP1 caused a switch toward the epithelial splicing pathway for each of these transcripts (Figure 5G). Similar changes in splicing were also observed with ectopic expression of ESRP1 in the mesenchymal MDA-MB-435 cell line (Figure S11).

Figure 5. Expression of ESRP1 and ESRP2 Is Abrogated during the EMT

(A) The morphologies of the HMLE cells expressing either the empty pBabe-Puro or pBabe-Puro-Twist by phase contrast microscopy. (B) Expression of epithelial genes E-cadherin and Keratin 15 and mesenchymal genes N-cadherin and Foxc2 by RT-PCR of RNAs from HMLE cells expressing either the control vector (EV) or Twist. (C) Immunoblot confirming Twist-induced down-regulation of E-Cadherin protein and upregulation of the mesenchymal Vimentin protein. β-actin is included as a loading control. (D) Downregulation of the epithelial miR-200c during Twist-induced EMT as determined by qRT-PCR. Mean expression values from triplicate assays ± SD are shown relative to control vector. (E) Downregulation of ESRP1 and ESRP2 mRNA during Twist-induced EMT determined by qRT-PCR. Mean expression values from triplicate assays ± SD are shown relative to control vector. (F) Twist-induced EMT causes a change in splicing of FGFR2, CD44, CTNND1, and ENAH. Asterisks indicate additional known minor splice variants. (G) Ectopic expression of Esrp1 in mesenchymal MDA-MB-231 cells causes opposite changes in splicing of FGFR2, CD44, CTNND1, and ENAH to those that occur in the EMT. (H) Expression of FLAG-tagged Esrp1 in MDA-MB-231 cells as determined by immunoblotting with anti-FLAG antibodies.

was previously shown to enhance splicing of the upstream exon IIIb and silence the downstream exon IIIC, and our studies indicate that the ESRPs participate in both of these functions. Fox-2 has also been shown to both enhance exon IIlb and silence exon IIIC splicing, and the proximity of a Fox-binding site to ISE/ISS-3 suggests that cooperative interactions between these proteins may be required for both functions. Several proteins, including hnRNP M, hnRNP F, and hnRNP H silence exon IIIC but do not directly affect exon IIlb splicing (Hovhannisyan and Carstens, 2007; Mauger et al., 2008). It is noteworthy that hnRNP F/H are close homologs of the ESRPs and both function in FGFR2 splicing regulation (Figure S9). However, while hnRNP F and -H bind to G triplets in exon IIIC, mutations in ISE/ISS-3 that abrogate ESRP binding do not directly affect G triplets, suggesting that they have somewhat different binding specificity. Although its binding site in FGFR2 is not defined, cotransfection studies suggest that RBM38 enhances exon IIlb splicing in the absence of effects on exon IIIC splicing (data not shown). Because knockdown of both ESRP1 and ESRP2 in an epithelial cell line caused a switch in endogenous FGFR2 splicing toward the mesenchymal pathway, we propose that in the absence of these proteins, the combinatorial balance of other splicing regulators favors the mesenchymal FGFR2 splicing.
pathway. The ESRP proteins, however, most likely collaborate with other, more ubiquitously expressed splicing regulators to remodel the RNP complexes assembled on FGFR2 transcripts in epithelial cells and tip the balance toward exon IIIb inclusion and exon IIIC repression (Figure 6).

Although a very limited number of cell-type-specific splicing regulators have been identified in mammals, there are almost certain to be numerous additional splicing factors with cell-type-specific expression that will be found among the large number of uncharacterized RBPs (David and Manley, 2008). To date, the best characterized mammalian cell-type- or tissue-specific splicing factors such as Nova, nPTB, Fox-1/2, Muscleblind (MBNL), and CELF family members regulate neuronal- or muscle-specific splicing events (Li et al., 2007; Matlin et al., 2005). Studies of the neural-specific Nova proteins revealed that they coregulate splicing of numerous targets encoding proteins that function at neuronal synapses (Ule et al., 2005). Studies of the neural and muscle-enriched Fox proteins have similarly shown that their targets are enriched for gene products that function at neuromuscular junctions and that are implicated in neurological and muscular diseases (Zhang et al., 2008). Such examples show that identification of coregulated splicing events that comprise these splicing regulatory networks can potentially identify proteins that are functionally linked with roles in common cellular and developmental processes (Keene, 2007). The limited examples of coregulated targets of the ESRPs we have identified here suggest that they similarly coregulate the splicing of numerous transcripts whose different splice variants have distinct roles in epithelial and mesenchymal cells that contribute to the unique functions and characteristics of each cell type.

Epithelial- and mesenchymal-specific isoforms that are regulated by the ESRPs are likely to participate in epithelial-mesenchymal crosstalk during early vertebrate development and to have important roles in EMTs during development as well as in disease processes such as cancer metastasis and tissue fibrosis. Interesting examples of coordination among the ESRP targets described here are consistent with these hypotheses. For example, expression of epithelial FGFR2-IIIb in the apical ectodermal ridge (AER) is required for limb induction in response to mesenchyme-derived FGF-10 (Xu et al., 1998). Expression of the epithelial variant of CD44 in the AER was shown to facilitate a reciprocal interaction of AER-derived FGF-8 with receptors in the underlying mesenchyme that most likely include FGFR2-IIIC as well as FGFR1 (Sherman et al., 1998). p120-Catenin (CTNND1) associates with E-cadherin at the plasma membrane and promotes cell-cell adhesion, but paradoxically can also promote cell motility and invasion in cells that have lost E-cadherin expression (Yanagisawa et al., 2008). These seemingly contradictory functions of p120-catenin appear to be due to the different activities of the splice variants that predominate in epithelial versus mesenchymal cells as the mesenchymal p120-catenin isoforms, but not the epithelial isoforms, promote cell motility and invasion in mesenchymal cells (such as MDA-MB-231) that have lost E-cadherin expression (Yanagisawa et al., 2008). Although there is some controversy regarding the relationship between the EMT and tumor metastasis, there are numerous demonstrations that the EMT is one mechanism that can contribute to the metastatic process (Yang and Weinberg, 2008). Our studies showing coordinated changes in splicing that accompany the EMT suggest that the differential functions of these epithelial and mesenchymal splice variants may contribute to metastasis. Enhanced cell motility and invasive capacity mediated by the mesenchymal isoform of p120-catenin is just one example of a splicing change during the EMT that might contribute to metastasis. However, whether these changes in splicing are required for the EMT to occur will require further investigation.

Given the limitations of biochemical approaches, a number of groups have developed fluorescence-based splicing reporter minigenes to establish genetic screens for splicing regulators (Newman et al. [2006] and references therein). An application of this approach in C. elegans led to the identification of the previously noted SUP-12, whose mammalian ortholog, RNPC1 (RBM38), was not known to be a splicing regulator prior to our study (Kuroyanagi et al., 2007). Mammalian cell-based screens were successful in identification of two factors that regulate Tau exon 10 inclusion through expression cloning with a pooled cDNA library (Kar et al., 2006; Wu et al., 2006). Cell-based
screening approaches also recently uncovered hnRNP L-like (hnRNP LL) as a splicing regulator of CD45 in activated T cells (Oberdoerffer et al., 2008; Topp et al., 2008). Further application of genome-wide, high throughput screening approaches has the potential to further expand the number of known splicing regulators. Parallel overexpression of multiple cDNAs at the same time using reverse transfection in a gain-of-function approach has now been scaled up to permit genome-wide screens in high-throughput format, an approach that we applied successfully in this study (Luesch, 2006). While the relative merits and liabilities of an RNAi screen versus a cDNA expression screen can be debated, functionally redundant proteins such as ESRP1 and ESRP2 may elude detection in an arrayed RNAi screen. In addition to the ESRPs, we also discovered several additional, previously uncharacterized mammalian splicing regulators, including RBM38, that also merit further investigation. These studies thus illustrate the potential of mammalian genetic screens in high-throughput format to complete a more comprehensive catalog of mammalian splicing regulators.

**EXPERIMENTAL PROCEDURES**

**Plasmids**

Minigene PKC-neg-40B-IF3-Luc-Puro was derived from PKC-neg-40B-IF3-EGFP-Puro (Newman et al., 2006) by substitution of a firefly luciferase CDS from pGL4 for EGFP. Additional details of plasmid construction are provided in the Supplemental Data.

**Cell Culture and Transfection**

293T cells used in the screen and validations were maintained in DMEM with 5%-10% FBS and transfected with Transit 293 (Mirus). Details on additional cell lines used are described in the Supplemental Data.

**Viral Production and Transduction**

Retroviral transductions were performed using the 293T-N16 packaging cell line. Cells were transfected in 6 cm dishes with 5.7 μg of the retroviral expression vector and 0.3 μg of pCMV-VSV-G using Transit 293. After 16-20 hr, the media was replaced with fresh DMEM with 5% FBS, and virus was harvested after an additional 24 hr. Target cells were transduced with a 50/50 mix of viral supernatant and growth media. Selection was carried out using g/ml puromycin or 40 μg/ml blasticidin.

**RT-PCR**

RNA was isolated using Trizol (Invitrogen). Reverse transcription, PCR, and minigene splicing analysis were performed essentially as described (Newman et al., 2006). The identity of all splice variants was confirmed by DNA sequencing.

**Real-Time qRT-PCR**

One microgram of total RNA was used to generate cDNA with the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Analysis was performed in triplicate using TaqMan Gene Expression Assays for human RBM35A and RBM35B, normalized to 18 s rRNA (Applied Biosystems). The ddCt method of relative quantification was used on a 7500 Fast Real-Time PCR System (Applied Biosystems). MicroRNA expression levels were measured in quadruplicate using the microRNA TaqMan Assays for human miR-200c (Applied Biosystems), normalized to RNU48 (snoRNA, C/D box 48).

**RNA Interference**

PNT2 cells were seeded in 12-well plates and transfected with siRNAs twice over a period of 48 hr using Lipofectamine 2000. RNA was extracted 48 hr after the second siRNA transfection. QIAGEN siRNAs and Ambion siRNAs were used at final concentrations of 40 nM and 10 nM, respectively. For the RNAi rescue experiments, PNT2 cells were transduced with pMXs-ires-blast2 vectors and selected in Blasticidin for 5 to 7 days prior to transfection with siRNAs.

**Antibodies and Immunoblotting**

Total cell extracts were harvested in RIPA buffer. Immunoblotting was performed as described (Hovhannisyan and Carstens, 2007). Full details on the antibodies are in the Supplemental Data.

**UV Crosslinking and Immunoprecipitation**

The constructs and protocol for UV crosslinking and immunoprecipitation were previously described, but with further modifications as described in the Supplemental Data (Hovhannisyan and Carstens, 2007).

**SUPPLEMENTAL DATA**

The Supplemental Data include 11 figures, one table, Supplemental Experimental Procedures, and Supplemental References and can be found with this article online at http://www.cell.com/molecular-cell/supplemental/S1097-2765(09)00069-0.

**ACKNOWLEDGMENTS**

We thank Kunio Inoue, S. Ananth Karumanchi, Tosho Kitamura, Reini F.Luco, Tom Misteli, Marcus Peter, and Robert Weinberg for reagents; Tom Jongens, Steve Liebhaber, and Rebecca Wells for critical review of the manuscript; Jeanne Geskes for assistance with the high-throughput screen; John Tobias for assistance with microarray analysis; Jose Russo for use of microarray data; and Sendurai Mani, Scott Dessain, and Michael Beers for helpful discussions. C.C.W. is supported by Genetics Training Grant NIGMS T32-GM08216. This research was supported by the Pennsylvania Department of Health and NINDS R01 NS054794 (J.B.H.), DOD grant PC 991539 (R.P.C.), NIH grant R01 CA093769 (R.P.C.), and a Penn Genome Frontiers Institute Seed Grant (R.P.C.).

Received: September 5, 2008
Revised: December 11, 2008
Accepted: January 12, 2009
Published: March 12, 2009

**REFERENCES**


