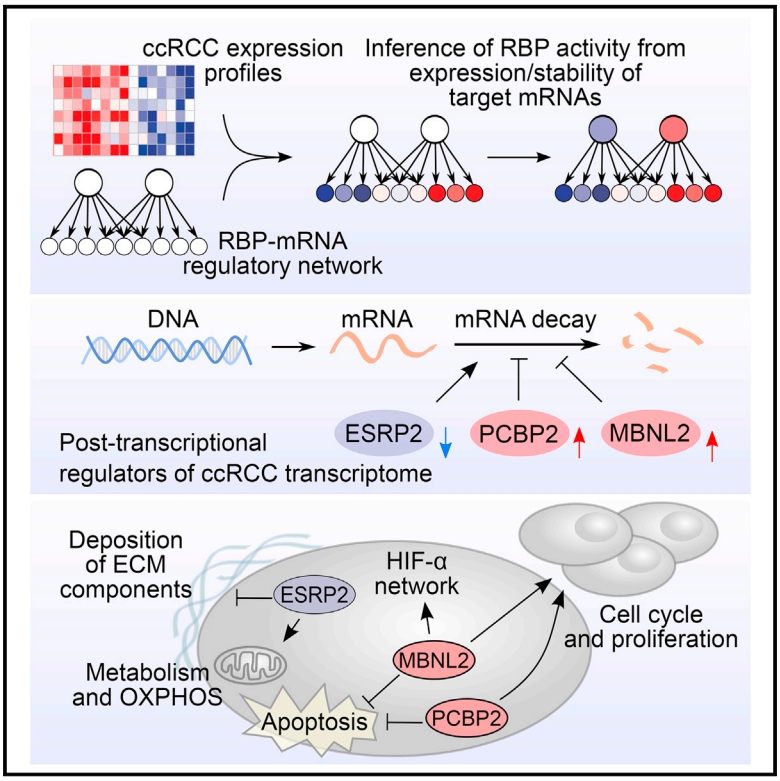
Article

Cell Reports

A General Framework for Interrogation of mRNA Stability Programs Identifies RNA-Binding Proteins that Govern Cancer Transcriptomes

Graphical Abstract



Authors

Gabrielle Perron, Pouria Jandaghi, Shraddha Solanki, ..., Fadi Brimo, Hamed S. Najafabadi, Yasser Riazalhosseini

Correspondence

hamed.najafabadi@mcgill.ca (H.S.N.), yasser.riazalhosseini@mcgill.ca (Y.R.)

In Brief

Perron et al. develop a computational approach that models the functional activity of RBPs in individual cancer samples by monitoring their associated RNA stability programs. Applying this method to renal cell carcinoma transcriptomes, the authors identify RBPs that enhance cancer-associated pathways including hypoxia and cell cycle.

Highlights

- Computational modeling of mRNA stability programs suggests a role of RBPs in cancer
- Dysregulation of these programs leads to perturbation of cancer-associated pathways
- RBP modulation in cell lines partially mirrors the transcriptome remodelling in cancer

Data and Software Availability

GSE83999

Perron et al., 2018, Cell Reports *23*, 1639–1650
May 8, 2018
https://doi.org/10.1016/j.celrep.2018.04.031

CellPress

Cell Reports
Article

A General Framework for Interrogation of mRNA Stability Programs Identifies RNA-Binding Proteins that Govern Cancer Transcriptomes

Gabrielle Perron,[1,2,11] Pouria Jandaghi,[1,2,11] Shraddha Solanki,[3] Maryam Safisamghabadi,[1,2] Cristina Storoz,[3] Mehran Karimzadeh,[1,2] Andreas I. Papadakis,[4,5] Madeleine Arseneault,[1,2] Ghislaine Scelo,[6] Rosamonde E. Banks,[7] Jorg Tost,[8] Mark Lathrop,[1,2] Simon Tanguay,[9] Alvis Brazma,[10] Sidong Huang,[4,5] Fadi Brimo,[3] Hamed S. Najafabadi,[1,2,12,*] and Yasser Riazalhosseini[1,2,*]

[1]Department of Human Genetics, McGill University, Montreal, QC H3A 1B1, Canada
[2]McGill University and Genome Quebec Innovation Centre, Montreal, QC H3A 0G1, Canada
[3]Department of Pathology, McGill University, Montreal, QC H3A 2B4, Canada
[4]Department of Biochemistry, McGill University, Montreal, QC H3G 1Y6, Canada
[5]Rosalind and Morris Goodman Cancer Research Centre, McGill University, Montreal, QC H3A 1A3, Canada
[6]International Agency for Research on Cancer (IARC), 150 cours Albert Thomas, Lyon 69008, France
[7]Leeds Institute of Cancer and Pathology, University of Leeds, Cancer Research Building, St. James's University Hospital, Leeds LS9 7TF, UK
[8]Laboratory for Epigenetics & Environment, Centre National de Recherche en Génomique Humaine, CEA-Institut de Biologie Francois Jacob, 2 rue Gaston Crémieux, 91000 Evry, France
[9]Department of Urology, McGill University, Montreal, QC H3G 1A4, Canada
[10]European Molecular Biology Laboratory, European Bioinformatics Institute, EMBL-EBI, Wellcome Trust Genome Campus, Hinxton CB10 1SD, UK
[11]These authors contributed equally
[12]Lead Contact
*Correspondence: hamed.najafabadi@mcgill.ca (H.S.N.), yasser.riazalhosseini@mcgill.ca (Y.R.)
https://doi.org/10.1016/j.celrep.2018.04.031

SUMMARY

Widespread remodeling of the transcriptome is a signature of cancer; however, little is known about the post-transcriptional regulatory factors, including RNA-binding proteins (RBPs) that regulate mRNA stability, and the extent to which RBPs contribute to cancer-associated pathways. Here, by modeling the global change in gene expression based on the effect of sequence-specific RBPs on mRNA stability, we show that RBP-mediated stability programs are recurrently deregulated in cancerous tissues. Particularly, we uncovered several RBPs that contribute to the abnormal transcriptome of renal cell carcinoma (RCC), including PCBP2, ESRP2, and MBNL2. Modulation of these proteins in cancer cell lines alters the expression of pathways that are central to the disease and highlights RBPs as driving master regulators of RCC transcriptome. This study presents a framework for the screening of RBP activities based on computational modeling of mRNA stability programs in cancer and highlights the role of post-transcriptional gene dysregulation in RCC.

INTRODUCTION

Cancer is a complex disorder, with a large array of molecular factors involved. Global deregulation of genes is a hallmark of cancer cells, and a large number of the diagnostic, prognostic, and therapeutic approaches target "master regulators" of gene expression or their downstream effectors (Goolam et al., 2016; Kim and Roberts, 2016; Perera et al., 2015). Nevertheless, unlike extensive studies on diverse molecular factors involved in transcriptional gene regulation, such as aberrations involving transcription factors (TFs) or epigenetic patterns, our knowledge of the mechanisms underlying post-transcriptional gene deregulation in cancer is scarce. RNA-binding proteins (RBPs), as key factors that modulate the stability and splicing of thousands of mRNAs (Gerstberger et al., 2014), play central roles in post-transcriptional gene regulation. Previous studies have linked a few RBPs with transcriptome remodeling and disease progression in cancer. For example, TARBP2 (Goodarzi et al., 2014) and NELFE (Dang et al., 2017) have recently been identified as oncogenic RBPs whose abnormal activation results in the progression of breast and hepatocellular carcinomas, respectively, by dysregulating the stability of their target mRNA. Likewise overexpression of SNRPB leads to abnormal splicing of many genes involved in DNA repair and chromatin remodeling pathways, thereby contributing to enhanced cell proliferation in glioblastoma (Correa et al., 2016).

However, a systematic interrogation of RBP-mediated regulatory programs in cancer and the extent to which they shape the cancer gene expression signature has not been performed, despite evidence for widespread dysregulation of these proteins in cancer (Kechavarzi and Janga, 2014). A recent analysis of eukaryotic RBPs has provided a regulatory model that links the binding of 35 human RBPs to the decay rate of their target mRNAs (Ray et al., 2013). This model identifies high-confidence

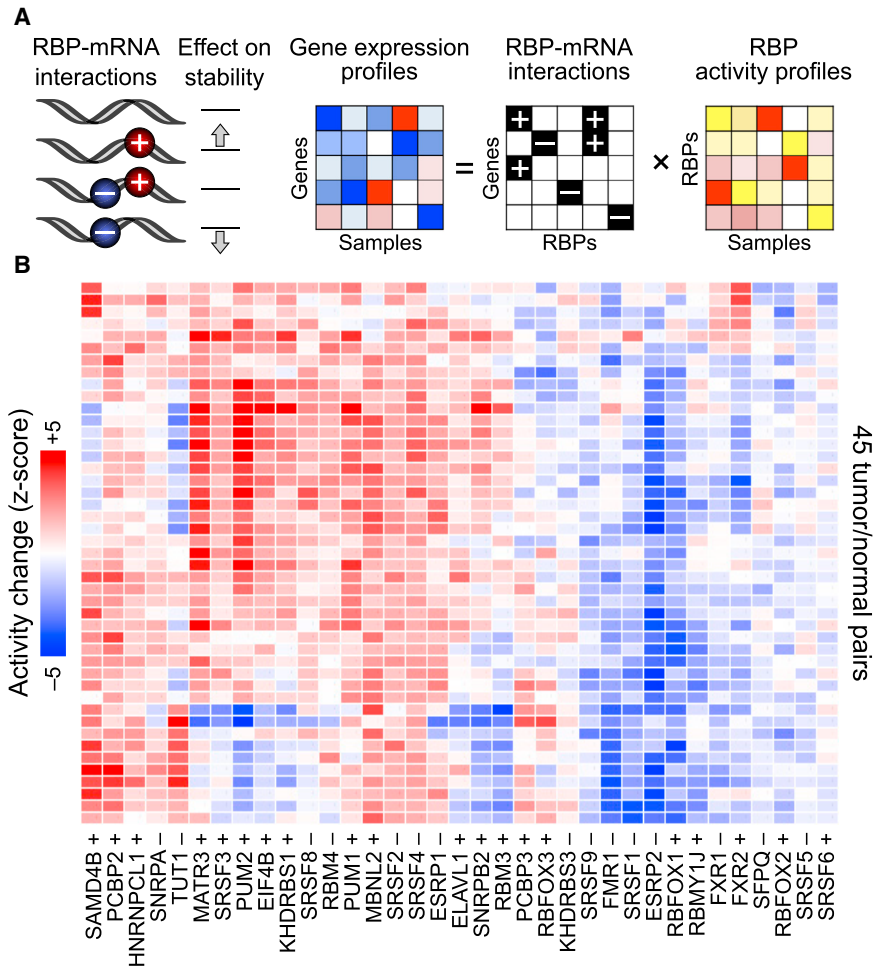


Figure 1. Inferring Dysregulated Post-transcriptional Programs from ccRCC Expression Data

(A) Schematic representation of inferring RBP activities by combining the stability code (Ray et al., 2013) with gene expression data. The plus (+) signs represent stabilizing RBPs, and the minus (−) signs stand for destabilizing RBPs.

(B) The inferred activity profiles of RBPs in ccRCC tumors. Each row represents one patient from the CAGEKID cohort (Scelo et al., 2014), and each column represents one RBP. Plus (+) and minus (−) signs denote RBPs that stabilize or destabilize their targets, respectively. The color gradient corresponds to the inferred ccRCC-associated activity change (*Z* score obtained by bootstrapping; see Experimental Procedures). Red denotes increased activity of RBP in ccRCC tumor relative to matching normal tissue, and blue denotes decreased activity.

See also Figure S1.

RESULTS

Widespread Dysregulation of Post-Transcriptional Regulatory Programs in ccRCC

To identify RBP stability programs that are dysregulated in ccRCC, we used the RNA sequencing (RNA-seq) profiles of tumor and patient-matched normal tissue sample pairs from 45 ccRCC patients in the Cancer Genomics of the Kidney (CAGEKID) cohort (Scelo et al., 2014). These expression profiles quantitatively measure the extent of increase or decrease in the abundance of each mRNA in tumor cells compared to matching normal kidney tissue. We modeled the abundance of each mRNA as a function of the combinatorial effect of RBPs that bind to that mRNA (see Experimental Procedures), with the mRNA-RBP interactions derived from a previously reported "high-confidence" stability network (Ray et al., 2013). Following previously described methods for modeling gene regulatory networks (Lee and Bussemaker, 2010), we assumed that, in the logarithmic scale, the differential stability of each mRNA reflects the additive effect of the differential activity of RBPs that bind to that mRNA (Figure 1A). In this model, an increase in the "activity" of a stabilizing RBP or a decrease in the activity of a destabilizing RBP can result in an increase in the abundance of their target mRNAs. Conversely, a decrease in the activity of a stabilizing RBP or an increase in the activity of a destabilizing RBP can lead to a decrease in the abundance of their target mRNAs. By combining this model with ccRCC gene expression data, we predicted the change in the activity of RBPs based on the change in the abundance of their target mRNAs (Figure 1A). This model only considers post-transcriptional regulation by RBPs and, therefore, can only partially explain the changes in mRNA abundance. Despite this limitation, as we will show here, this model regulatory targets of each RBP, including mRNAs that have a functional RBP binding site at their 3′ UTRs, and indicates whether the binding of each RBP "stabilizes" or "destabilizes" its regulatory targets. Here, using this model, we present a general computational framework for analyzing post-transcriptional gene regulatory programs in cancer to identify key RBPs whose aberrant function plays a central role in the establishment of cancer-associated expression signatures.

We applied our method to gene expression profiles of clear cell renal cell carcinoma (ccRCC), which is the most common type of kidney cancer. The incidence rate of ccRCC is increasing worldwide, and its metastatic form is resistant to chemotherapy and radiotherapies (Capitanio and Montorsi, 2016). A hallmark of ccRCC is dramatic dysregulation of the transcriptome, which is partly mediated by transcriptional programs governed by the hypoxia-inducible transcription factor (HIF) and the malfunction of the epigenome machinery (Riazalhosseini and Lathrop, 2016). However, our understanding of the mechanisms that underlie the post-transcriptional gene regulation in ccRCC is very limited. We show that RBPs contribute significantly to the regulation of ccRCC transcriptome through modulation of mRNA stability, and we uncover key RBPs that regulate central ccRCC pathways.

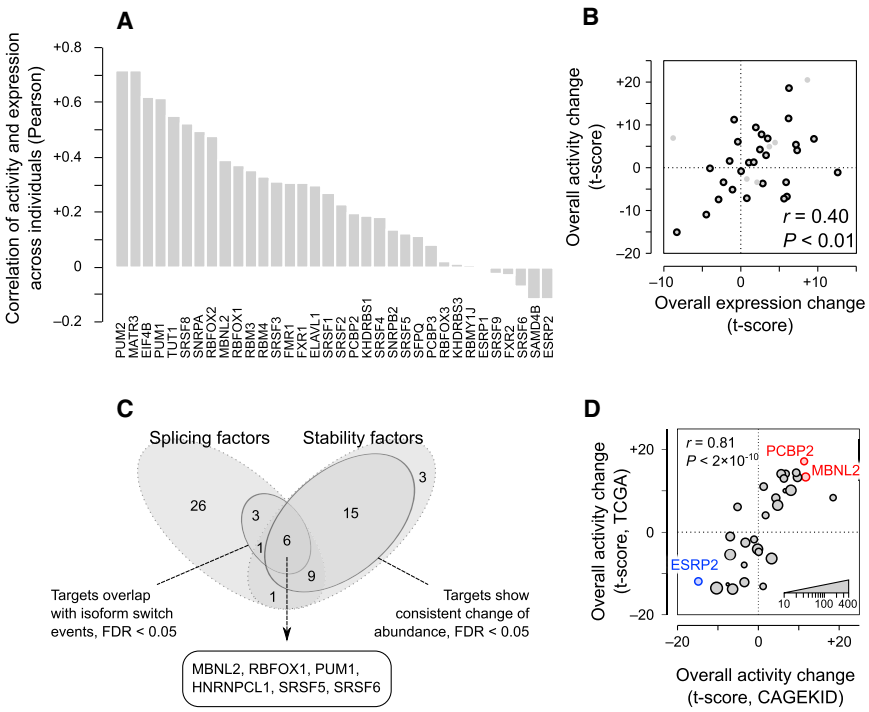


Figure 2. Orthogonal Information Supports the Inferred ccRCC-Associated RBP Activity Levels

(A) Correlation of the inferred activity profile of each RBP with its mRNA abundance across the CAGEKID cohort. See Figure S2 for individual scatterplots for each RBP.

(B) Overall change in inferred activity versus mRNA expression of RBPs. Each dot represents one RBP. The x axis corresponds to the t score of difference of mRNA abundance in tumor versus matching normal tissues (paired t test of log-transformed expression). The y axis corresponds to the t score of inferred ccRCC-associated activity changes: a positive value indicates that the average of the corresponding column in Figure 1B is above zero, and a negative value indicates that the average is below zero. Circles with black borders represent RBPs that have a motif that is based on direct experimental evidence or based on an analysis of a human paralog with >80% sequence similarity at the RNA-binding domain (Ray et al., 2013).

(C) Overlap of the RBPs whose stability targets are significantly dysregulated (FDR < 0.01, t test of average inferred RBP activity change) and RBPs whose splicing targets are significantly enriched for isoform switch events (FDR < 0.01, Fisher's exact test).

(D) Comparison of the inferred activity of RBPs, based on the stability code analysis, between CAGEKID and TCGA cohorts. The axes show the t statistic for average inferred RBP activity change between tumor and normal tissue. RBPs with a direct motif or with a motif that is derived from a human paralog with >80% sequence similarity (Ray et al., 2013) are shown. The circle size denotes the number of stability targets for each RBP.

See also Figure S2 and Table S1.

successfully identifies drivers of aberrant post-transcriptional regulatory programs.

We observed that regulatory programs that are associated with a large number of RBPs appear to be dysregulated in ccRCC, based on the predicted activity levels of RBPs in different patients (Figure 1B). To examine whether these estimates are confounded by transcription, we also used a recent approach for deconvolution of transcription and decay rates based on the intronic and exonic reads in the RNA-seq data (Alkallas et al., 2017). The RBP activities that were inferred from estimated mRNA stability profiles were highly consistent with the activities that we inferred from mRNA abundances (Figure S1), suggesting that our inferences are not confounded by the impact of transcription. These activity patterns were different across patients for some of the RBPs, which may reflect the inter-individual heterogeneity in ccRCC (Figure 1B). As expected, the predicted activity levels for a majority of RBPs correlated with the RBP mRNA abundances across individuals ($p < 4 \times 10^{-7}$, t test for average Pearson correlation; Figures 2A and S2). Furthermore, the overall increase or decrease of activity of RBPs in ccRCC was consistent with their overall up- or downregulation at the mRNA level ($r = 0.4$, $p < 0.01$; Figure 2B). These results suggest that the predicted RBP activity levels are, in part, functional consequences of the change in the expression of RBPs and can serve to predict RBPs whose dysregulation govern abnormal post-transcriptional gene regulation in ccRCC.

Identification and Validation of Recurrent Dysregulated Post-Transcriptional Programs

We used multiple orthogonal sources of information to further validate our prediction of the RBPs that contribute to the aberrant post-transcriptional gene regulation in ccRCC. First, we reasoned that, for RBPs that have a dual role in regulating mRNA stability and splicing (Ray et al., 2013), a change in the RBP activity should, at the same time, affect the stability of the mRNAs that have a binding site for that RBP in their 3′ UTRs, as well as the splicing of the mRNAs with a binding site near their exon-intron junctions. We identified RBPs whose splicing targets, as defined previously based on the presence of binding sites near exon-intron junctions of alternatively spliced exons (Ray et al., 2013), overlap significantly with ccRCC-associated "isoform switch events" (Scelo et al., 2014) (Table S1). These switch events represent changes in the identity of the most abundant spliced variant of a given gene in tumors compared to patient-matched normal tissue. Of the 10 RBPs whose splicing target set had a significant (false discovery rate [FDR] < 0.05) overlap with the ccRCC isoform switch events, six RBPs were also among the ones whose stability programs were dysregulated (Figure 2C), in agreement with the previously observed dual role of these RBPs in regulating both mRNA splicing and stability (Ray et al., 2013).

Next, we examined the reproducibility of RBP activity inference across different ccRCC datasets. For this, we used the expression profiles of 72 matching pairs of ccRCC tumor-normal

tissues from The Cancer Genome Atlas (TCGA) dataset (Cancer Genome Atlas Research Network, 2013) and inferred the change in the RBP activity levels in tumor versus normal samples of each patient. We observed a strong correlation in the ccRCC-associated change in RBP activity levels between the CAGEKID and TCGA cohorts ($r = 0.81$ for RBPs with a direct motif or a motif derived from a highly similar human paralog, $p < 2 \times 10^{-10}$; Figure 2D), confirming that our measurements are not dataset specific.

This comparative analysis revealed multiple RBPs with consistently large up- or downregulation of activity levels in both the CAGEKID and TCGA cohorts (Figure 2D; Table S1). In this work, we focus on three RBPs with the largest and most reproducible dysregulation of cancer-associated activity: ESRP2, which showed the greatest decrease in activity; and MBNL2 and PCBP2, which showed the largest increase in activity.

Modulation of RBP Levels Supports Their Role in Remodeling ccRCC Transcriptome

We further investigated the roles of ESRP2, PCBP2, and MBNL2 in remodeling the ccRCC transcriptome and the potential contribution to malignancy by analyzing pathways affected by their modulation, along with associated cellular functions. For this purpose, we performed an integrated analysis of ccRCC gene expression profiles and the transcriptomes of the cells that were specifically deficient (through RNAi) for candidate RBPs. Overall, for all the three RBPs, we observed a significant agreement between the transcriptome response that RBP knockdown triggered in the cell and what we expected based on our analysis of ccRCC expression profiles, as described in the following sections.

Modulation of ESRP2 Partially Recreates the Transcriptome Changes in ccRCC

In line with the reduced activity of ESRP2 that we had identified in both CAGEKID and TCGA datasets (Figure 2D), we observed that the mRNA of *ESRP2* is significantly downregulated in tumor tissues in both cohorts (Figure 3A). Likewise, we detected much higher expression levels of *ESRP2* in normal primary renal proximal tubule epithelial cells (PRPTECs) compared to several ccRCC cell lines (Figure 3B). Accordingly, we knocked down the *ESRP2* gene in PRPTECs using two independent short hairpin RNAs (shRNAs) (Figure 3C) and investigated the changes of the transcriptome using RNA-seq. We observed widespread and reproducible remodeling of the PRPTEC transcriptome as a result of ESRP2 knockdown (Figures 3D and S3A); specifically, we observed overall upregulation of the stability of the mRNAs that have potential ESRP2 binding sites in their 3′ UTRs, particularly for genes that are also upregulated in ccRCC tumors at the stability level (Figure S3B). This observation is consistent with the previously suggested role of ESRP2 in destabilizing its target mRNAs (Ray et al., 2013), suggesting that downregulation of ESRP2 in ccRCC tumors leads to upregulation of its targets. We note that the effect of this deregulation may spread beyond the direct targets of ESRP2 and lead to the upregulation and downregulation of downstream genes. Consistent with this notion, we observed dysregulation of the mRNA abundances for a considerable number of biological processes and genes after ESRP2 knockdown, which largely mirrored those that are dysregulated in ccRCC tumors (Figure 3E). These included upregulation of components of extracellular matrix (ECM), in line with previous reports on an increased deposition of ECM in cancer (reviewed by Gilkes et al., 2014), as well as downregulation of processes related to mitochondrial metabolism and oxidative phosphorylation, which are most often downregulated in ccRCC tumors (Cancer Genome Atlas Research Network, 2013; Scelo et al., 2014) (Figure 3F). We further validated upregulation of COL1A1, a major component of ECM, at the protein level upon ESRP2 depletion (Figure 3G). These findings further support the role of ESRP2 deficiency in the abnormal activity of ccRCC-associated pathways.

PCBP2 Positively Regulates Cell-Cycle Progression

Our analysis of activity patterns of RBPs also highlighted PCBP2 as one of the RBPs with the most elevated activity in ccRCC (Figure 2D). Interestingly, however, analysis of the *PCBP2* mRNA levels did not show a significant change in the tumors as compared to normal samples in any of the two cohorts (Figure 4A). To rectify this apparent discrepancy, we set out to directly measure the protein abundance of PCBP2 in tumor and normal tissues of an independent cohort of ccRCC patients, using tissue microarrays (TMAs) (Xu et al., 2013). Through immunohistochemistry (IHC) analysis, we were able to measure the abundance of PCBP2 in a panel of 58 morphologically normal and 77 tumor tissue samples from ccRCC patients. Our results from these experiments revealed that PCBP2 shows a strong increase in abundance in tumor cells (5.8-fold increase, $p < 7 \times 10^{-8}$; Figures 4B and 4C), in agreement with the predicted increase in its activity levels. These findings suggested that upregulation of PCBP2 occurs at post-transcriptional level. Indeed, our analysis of a previously published dataset (Loayza-Puch et al., 2016) revealed a strong increase (4.9-fold) in ribosome occupancy of PCBP2 mRNA in a ccRCC tumor compared to matching normal kidney tissue, in line with translational upregulation of PCBP2 in tumors (Figure S4A).

To gain insight about the functional significance of PCPB2 abnormal activation in ccRCC, we explored cellular pathways whose activity is influenced by PCPB2. We observed that, in HepG2 cells, whose gene expression profile is correlated with that of common ccRCC cell lines (Figure S4B), PCBP2 knockdown induces a change in the transcriptome that is in the opposite direction of the transcriptome remodeling in ccRCC ($r = -0.07$, $p < 6 \times 10^{-24}$; Figure 4D; PCBP2-knockdown expression data are from the ENCODE Project Consortium [2012]). Furthermore, genes that are downregulated after PCBP2 knockdown are enriched for pathways that are relevant to cell proliferation in cancer, including cell cycle and related pathways (Figure 4E). The genes in these pathways are also correlated with inferred PCBP2 activity levels across the 45 CAGEKID samples ($p < 4 \times 10^{-5}$, t test for average Pearson correlation), but not with *PCBP2* mRNA levels (Figure 4F), further highlighting the biological relevance of the inferred activity levels in contrast to mRNA abundance for this RBP. In line with these results, we observed G1 cell-cycle arrest upon the inhibition of PCBP2 through RNAi in 786-O and A-498 renal cancer cell lines (Figures 4G and 4H), which resulted in reduced cell viability and induction

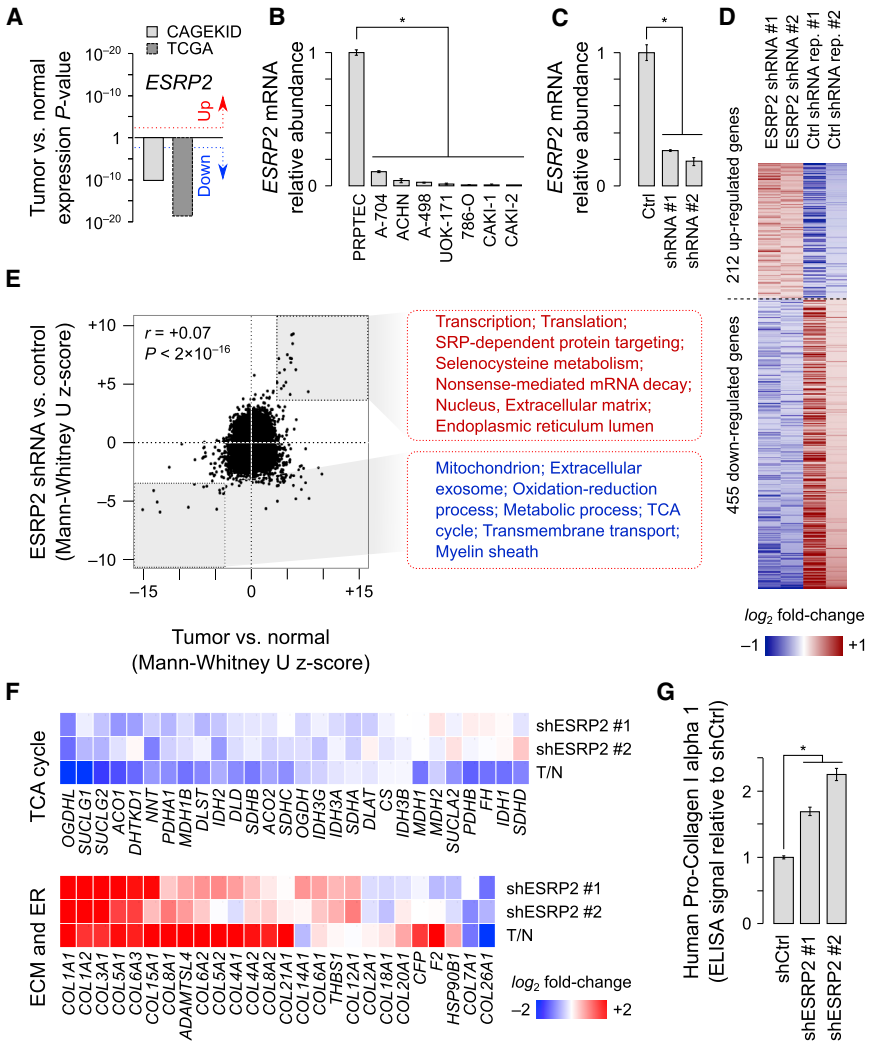


Figure 3. ESRP2 Deficiency Results in Dysregulation of ECM and Metabolic Pathways

(A) The ccRCC-associated mRNA abundance change for *ESRP2*. The y axis corresponds to a p value obtained from paired t test of mRNA abundance between ccRCC tumors and matching normal tissue, either from the CAGEKID cohort (light gray) or the TCGA cohort (dark gray). The p value of upregulation is shown above baseline, and downregulation is shown below baseline, in logarithmic scale.

(B) Comparison of the abundance of *ESRP2* mRNA between normal primary renal proximal tubule epithelial cells (PRPTECs) and seven common cell line models of RCC, measured by qRT-PCR.

(C) *ESRP2* mRNA abundance, measured by qRT-PCR, in PRPTECs infected with a control shRNA (Ctrl) or with two different shRNAs that specifically target *ESRP2*.

(D) Widespread and reproducible change in mRNA abundance as a result of ESRP2 knockdown in PRPTECs. Each row represents one gene, and each column represents either transfection of PRPTECs with an *ESRP2*-specific shRNA or with a control shRNA. Only genes that had at least 10 reads in each RNA-seq dataset and showed a minimum of 20% increase or decrease of expression in both of the *ESRP2* shRNA transfections compared to both of the control shRNA transfections are included. A total of 667 genes match these criteria for reproducible up-/downregulation of expression; the probability of observing this number of reproducible changes by chance is 4×10^{-6} based on permutation of expression fold-change values within each condition.

(E) Comparison of dysregulated processes and cell compartments in ccRCC tumor and in ESRP2-knockdown PRPTECs. Each dot represents either a Gene Ontology (GO) (The Gene Ontology Consortium, 2015) biological process or a GO cell compartment. The axes show the *Z* score from Mann-Whitney U test for change in the abundance of mRNAs of each GO term, with positive values representing overall upregulation of the GO term and negative values representing downregulation. The x axis corresponds to dysregulation in ccRCC tumor relative to normal tissue (CAGEKID dataset), and the y axis corresponds to dysregulation in ESRP2 knockdown (average of two shRNAs) relative to the Ctrl condition. GO terms that are significantly upregulated (*Z* score $\geq$ 3.5, red) or downregulated (*Z* score $\leq$ −3.5, blue) in both ccRCC tumor and ESRP2-knockdown cells are highlighted.

(F) The mRNA abundance of genes that belong to both extracellular matrix (ECM) and endoplasmic reticulum lumen (ER) compartments (bottom) or genes that belong to the tricarboxylic acid (TCA) cycle (top). The color gradient represents the average logarithm of fold change between tumor and normal tissue (based on CAGEKID) or logarithm of fold change for each of the *ESRP2* shRNA infections relative to control shRNA. The color map is shown on the right.

(G) Abundance of COLA1 after ESRP2 knockdown in PRPTECs, as measured by ELISA.

All error bars represent SD. *p < 0.05 (one-sided Wilcoxon test).

See also Figures S3 and S5.

of apoptosis (Figure 4I). These observations confirmed that PCBP2 enhances cell-cycle progression and proliferation in ccRCC.

MBNL2 Is Essential for ccRCC Cells

Lastly, we investigated the dysregulation and the role of MBNL2 in ccRCC, which had exhibited a strong enhanced activity in both CAGEKID and TCGA datasets (Figure 2D). We noted that *MBNL2* mRNA is significantly upregulated only in the CAGEKID cohort, with the TCGA cohort showing an overall consistent, albeit not statistically significant, trend (Figure 5A). Therefore, we interrogated MBNL2 protein levels using IHC as described earlier. Measuring MBNL2 expression in a panel of 35 morphologically normal and 76 tumor tissue samples from ccRCC patients showed a strong overexpression in tumor cells (16.8-fold increase for MBNL2, $p < 7 \times 10^{-8}$; Figures 5B and 5C), corroborating the predicted increase in its activity levels.

To study MBNL2 function, we inhibited the expression of MBNL2 in 786-O and A-498 cells (Figure 5D), which resulted in a global effect on the transcriptome, with significant overlap between genes dysregulated in the two cell lines (Figure S3C). Given the well-characterized role of MBNL2 in regulating splicing

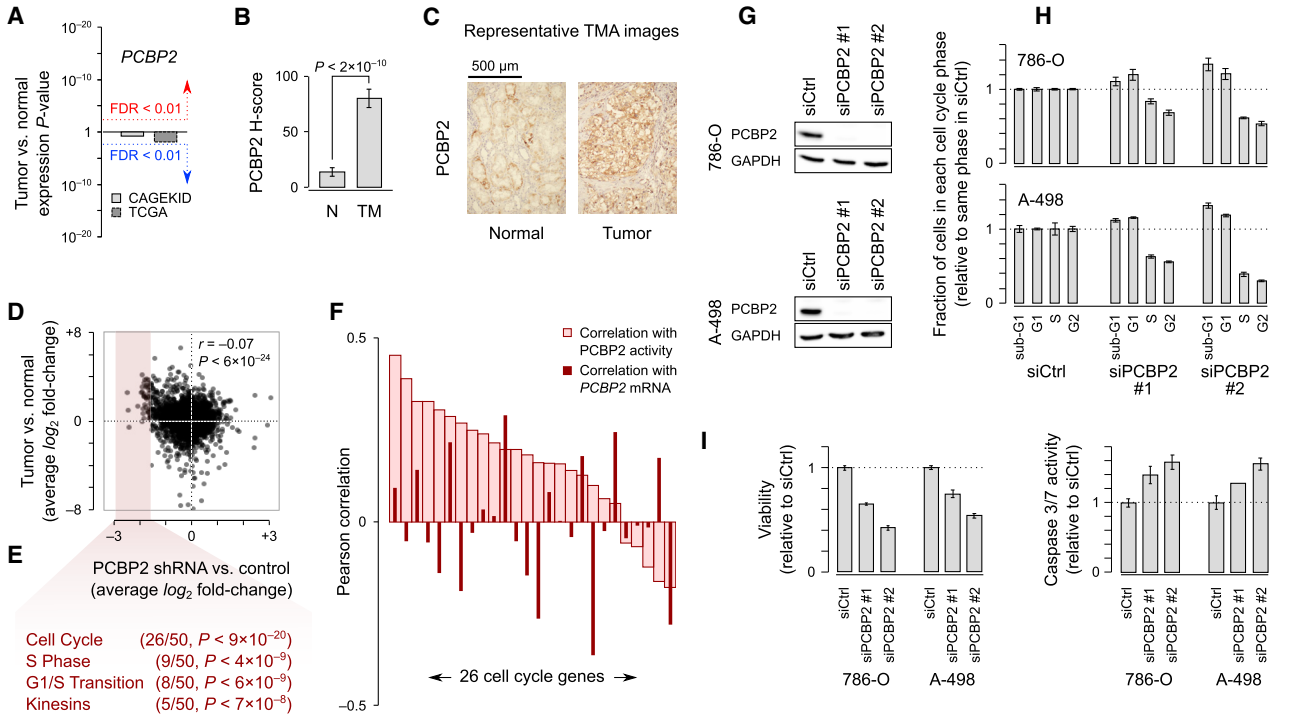


Figure 4. PCPB2 Is Upregulated at Translational Levels in ccRCC and Contributes to the Cell-Cycle Progression

(A) *PCBP2* mRNA abundance does not show a significant change in ccRCC tumor compared to matched normal kidney tissue samples. See Figure 3A for the plot details.

(B) Comparison of protein abundance between ccRCC tumor (TM) and normal kidney (N) tissue for PCBP2, measured using tissue microarrays (TMAs), shows upregulation in tumors. See the Supplemental Experimental Procedures for the definition of H-score. p values are based on t test.

(C) Representative TMA images for measuring the abundance of PCBP2 in normal and tumor tissues.

(D) Comparison of mRNA abundance change between tumor and normal tissues (CAGEKID dataset, y axis) and mRNA abundance change as a result of PCBP2 knockdown in HepG2 cells (x axis, average of two shRNAs).

(E) Representative pathways that are significantly enriched among the top 50 most downregulated genes after PCBP2 knockdown. The fraction of the 50 genes that belong to each pathway is shown in parentheses, followed by enrichment p value (Fisher's exact test). Pathway annotations are based on Reactome (Croft et al., 2014).

(F) Correlation of the ccRCC-associated mRNA abundance of cell-cycle genes with the mRNA abundance of *PCBP2* (solid bars) or inferred PCBP2 activity (open bars). Correlations are calculated across 45 matching pairs of tumor-normal tissues from the CAGEKID dataset. Only cell-cycle genes that are among the 50 most downregulated genes after PCBP2 knockdown are shown.

(G) Western blot analysis of PCBP2 protein in 786-O and A-498 cells after knockdown using two different siRNAs, as well as using a control (non-targeting) siRNA.

(H) Change in the proportion of cells in different cell-cycle phases after PCBP2 knockdown in 786-O and A-498 cells.

(I) Viability (left) and caspase-3/7 activity (right) after PCBP2 knockdown.

Values are normalized to control (non-targeting) siRNA in (H) and (I). Error bars in (B) represent SEM. Other error bars represent SD.

See also the Supplemental Experimental Procedures and Figure S4.

(Charizanis et al., 2012), we first examined the splicing targets of this protein and observed significant dysregulation of their transcript isoform distributions after MBNL2 knockdown (Figure S3D). Next, we examined the effect of MBNL2 knockdown on its stability targets. Inference of mRNA stability from the RNA-seq data revealed an overall destabilization of predicted MBNL2-binding mRNAs after MBNL2 knockdown (Figure S3E), consistent with a role of MBNL2 in stabilizing its targets. Interestingly, MBNL2 knockdown partially reversed the gene expression signature of ccRCC, leading to an overall decrease in the expression of genes that are upregulated in ccRCC ($r = -0.11$, $p < 3 \times 10^{-62}$; Figure S3F). These genes are highly enriched for pathways with well-established roles in ccRCC, including the HIF-α TF network (Harris, 2002), the FOXM1 TF network (Wu et al., 2013; Xue et al., 2012), and the PLK1 signaling pathway (Ding et al., 2011; Zhang et al., 2013) (Figure 5E). The HIF-α network genes that showed the greatest downregulation after MBNL2 knockdown included vascular endothelial growth factor A (VEGFA), a well-known angiogenic factor, along with four other genes (Figure 5F). In agreement with a potential role of MBNL2 in regulating these genes, they showed strong correlation with MBNL2 expression across the 45 tumor samples from the CAGEKID dataset (Figure 5G), with VEGFA showing the strongest correlation ($r = 0.49$, $p < 3 \times 10^{-4}$). Through a pull-down assay using an antibody against MBNL2, we confirmed that MBNL2 directly binds to *VEGFA* mRNA (Figure 5H), consistent with a previously identified interaction between mouse Mbnl2 and the 3′ UTR of *Vegfa* (Charizanis et al., 2012). We also

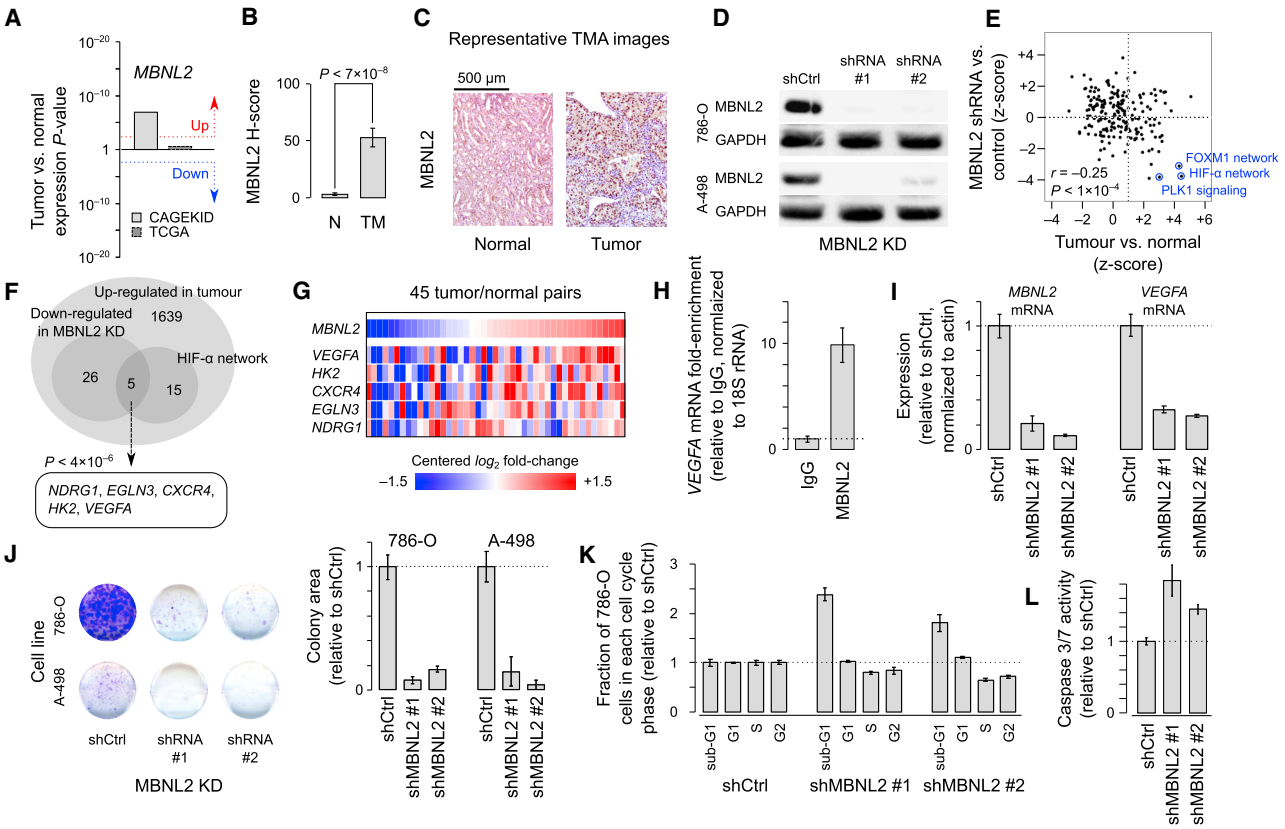


Figure 5. MBNL2 Upregulation Enhances HIF-Responsive Gene Expression and Cell Proliferation Pathways in ccRCC

(A) The ccRCC-associated mRNA abundance change for *MBNL2*. See Figure 3A for the plot details.

(B) Comparison of protein abundance between ccRCC tumor (TM) and normal kidney (N) tissue for MBNL2, measured using TMAs. See the Supplemental Experimental Procedures for the definition of H-score. p values are based on Student's t test.

(C) Representative TMA images for measuring the abundance of MBNL2 in normal and tumor tissues.

(D) Western blot analysis of MBNL2 abundance in 786-O and A-498 cells infected with a control shRNA (Ctrl) or with two different shRNAs specifically targeting *MBNL2*.

(E) Comparison of dysregulated pathways in ccRCC tumor and MBNL2-knockdown ccRCC cell lines. Each dot represents a pathway from the PID database (Schaefer et al., 2009). The x axis shows the Mann-Whitney U test *Z* score for up- or downregulation of pathways in ccRCC tumor versus normal tissue (CAGEKID dataset), and the y axis shows the *Z* score for dysregulation of pathways after MBNL2 knockdown relative to control shRNA.

(F) Among genes that are upregulated in tumor (average $\log_2$ fold change $\geq$ 1.5), the overlap of the HIF-α network genes (based on PID) with those that are downregulated after MBNL2 knockdown (average $\log_2$ fold-change ≤ -1.5) is shown.

(G) Comparison of the expression of HIF-α network genes with MBNL2 across 45 ccRCC tumors in the CAGEKID dataset. Each row represents a gene, and each column represents a matching pair of tumor-normal tissues. Columns are sorted based on MBNL2 expression. The color gradient denotes $\log_2$ of fold change between ccRCC tumor and normal tissue, with values transformed so that the average of each column becomes zero. Only the HIF-α network genes that are upregulated in ccRCC tumor and downregulated in MBNL2-knockdown cells are shown.

(H) Interaction of *VEGFA* mRNA with MBNL2, as measured by qRT-PCR in an MBNL2 pull-down assay.

(I) Downregulation of *MBNL2* after knockdown (left) leads to downregulation of *VEGFA* (right), as measured by qRT-PCR.

(J) Colony formation assay for 786-O and A-498 cells infected with a control shRNA (shCtrl) or with two different shRNAs that specifically inhibit *MBNL2*. Representative images are shown on the left, and colony area measurements from independent experiments are shown using the bar plots (n = 3).

(K) Change in the cell-cycle distribution after MBNL2 knockdown in 786-O cells.

(L) Caspase-3/7 activity after MBNL2 knockdown.

Values are normalized to control (non-targeting) shRNA in (I-L). Error bars in (B) represent SEM. Other error bars represent SD.

See also the Supplemental Experimental Procedures and Figures S3 and S5.

observed significant reduction of *VEGFA* mRNA levels after MBNL2 knockdown (Figure 5I), corroborating a role of MBNL2 in stabilizing the *VEGFA* mRNA.

In addition, we observed that MBNL2 inhibition significantly reduced the colony-forming ability of both 786-O and A-498 cells (Figure 5J), in agreement with the downregulation of FOXM1 and PLK1 pathways, which regulate cell-cycle progression and cell proliferation in ccRCC (Ding et al., 2011; Xue et al., 2012). Further analysis of cell-cycle status revealed that knockdown of MBNL2 in 786-O cells impairs cell cycle and induces an increase in the sub-G1 population of cells, indicative of increased apoptotic cells (Figure 5K). Indeed, we observed an induction of activated

caspase-3/7 in MBNL2-deficient cells, confirming the activation of the apoptosis pathway (Figure 5L).

DISCUSSION

By combining ccRCC expression profiles with a model of the RBP-regulated stability programs (Ray et al., 2013), we identified several RBPs that have significant roles in modulating the ccRCC gene expression signature, affecting the abundance of mRNAs in pathways that are critical to the development and progression of cancer. We investigated three RBPs with highly significant cancer-associated activities—ESRP2, PCBP2 and MBNL2—for their function and potential mechanisms of action in ccRCC. PCBP2 is of particular interest, given that the mRNA level of this protein is not different between ccRCC tumor and normal kidney tissues. We observed a very significant increase in the abundance of PCBP2 protein in ccRCC tumors, despite a lack of change at the mRNA level, suggesting that this RBP is potentially upregulated at the translation level, which is supported by data showing increased PBCP2 ribosome occupancy in ccRCC. This candidate would be missed had we relied on differential expression at the mRNA level for discovering RBPs with a gene regulatory function in cancer. This example highlights the value of inferring RBP activity levels based on modeling the post-transcriptional regulatory programs in order to identify RBPs driving transcriptome aberrations. Interestingly, PCBP2 has recently been shown to be upregulated at both the mRNA and protein levels in glioma where its inhibition slowed down cancer cell growth by impairing cell-cycle progression (Han et al., 2013b). Results from our transcriptomic analysis of ccRCC tumors and cell-cycle analysis of knockdown cells corroborate the oncogenic function of PCBP2 shown in this report and suggest that the expression of RBPs can be dysregulated by different mechanisms in different tissues; e.g., at the transcriptional level in glioma and at the post-transcriptional level in ccRCC.

Our study has also highlighted ESRP2 as a putative tumor suppressor in ccRCC with reduced activity levels in tumors. Supporting our results, a recent report has shown that the ESRP2 function suppresses ccRCC progression (Mizutani et al., 2016). ESRP2 expression is also associated with favorable prognosis in ccRCC (Figure S5A). Moreover, suggesting a general tumor-suppressive function for ESRP2, previous literature has reported that the deficiency of this protein promotes epithelial-mesenchymal transition (EMT), a process with a central role in cancer progression and metastasis (Shapiro et al., 2011; Warzecha et al., 2009, 2010). In line with the fact that EMT involves the dysregulation and remodeling of ECM (Gonzalez and Medici, 2014), we observed elevated expression of collagens, which are among the major constituents of ECM (Frantz et al., 2010), following ESRP2 suppression in kidney epithelial cells. Furthermore, whereas previous studies have focused only on the splicing regulatory function of ESRP2, our data reveal that the functional consequences of ESRP2 deficiency are broader than previously recognized; involve dysregulation of its stability activity; and lead to additional aberrations, including a metabolic shift toward the Warburg-effect, a characteristic feature of ccRCC (Zaravinos and Deltas, 2014).

Finally, MBNL2 displayed significant elevated activities in ccRCC in both the CAGEKID and TCGA datasets. MBNL2 plays an important role in the differentiation of embryonic stem cells (Han et al., 2013a), as well as in neuronal differentiation (Taliaferro et al., 2016). Although not previously linked to ccRCC, MBNL2 expression increases the migratory capacity of lung cancer cells (Adereth et al., 2005), suggesting a role in metastasis. Our IHC analyses also show that MBNL2 is upregulated in metastatic ccRCC tissues compared to primary tumors (Figure S5B). Furthermore, MBNL2 positively regulates levels of colony-stimulating factor receptor 1 (CSF1R) (Adereth et al., 2005), whose inhibition decreases the proliferation of renal cancer cells (Menke et al., 2012). Corroborating these reports, our analysis of the ccRCC transcriptome suggests an oncogenic function for MBNL2, which is at least partly due to the effect of MBNL2 on the abundance of cancer-associated transcripts and is supported by further functional validation in ccRCC cell lines. Specifically, we established that suppression of MBNL2 impairs cell proliferation and activates the apoptosis pathway in renal cancer cells.

PCBP2, MBNL2, and ESRP2 have widespread effects on the transcriptome based on our knockdown experiments followed by RNA-seq. Our analyses confirmed that the mRNAs that have target sites in their 3′ UTRs for these RBPs are enriched in the altered transcriptome following knockdown experiments. Of significance, we established that MBNL2 directly binds to the mRNA that encodes VEGFA and is essential for maintaining its expression in ccRCC cells. VEGFA is a prominent angiogenic factor that plays important roles in cancer metastasis. In addition to direct binding and dysregulation of *VEGFA* in MBNL2-knockdown cells, we observed a significant correlation between MBNL2 activity and the expression of *VEGFA* in tumor samples, supporting the notion that the upregulation of MBNL2 leads to higher levels of VEGFA. We note the potential clinical applications of this finding, given that VEGFA and downstream pathways are therapeutic targets in ccRCC (Riazalhosseini and Lathrop, 2016).

In addition to the direct effects of RBPs, a large fraction of the global remodeling of the transcriptome in our knockdown experiments appears to reflect indirect downstream effects, since many of the affected mRNAs do not have a binding site for the RBPs. This widespread dysregulation may potentially stem from the effect of RBPs on regulatory hubs, which can amplify and propagate across thousands of genes. For example, in addition to direct targeting by MBNL2 (e.g., in the case of *VEGFA*), downregulation of HIF-α network genes after inhibition of MBNL2 can be a result of downregulation of the HIF-α TF itself. Indeed, we observed that in A-498 cells, the mRNAs encoding the HIF-1α and HIF-2α proteins, which share many transcriptional targets in the HIF-α network (Schödel et al., 2011; Shinojima et al., 2007), are reproducibly downregulated after MBNL2 knockdown (Figure S5C). Similarly, in 786-O cells, which only express the full-length HIF-2α (Maxwell et al., 1999; Shinojima et al., 2007), MBNL2 knockdown results in downregulation of this gene (Figure S5C). This is further highlighted by a strong correlation between the MBNL2 abundance/activity and the abundance of HIF-1/2-α transcripts in the ccRCC tumors (Figure S5D). In addition, by analyzing genes that have a binding site for HIF-1α

but not for HIF-2α (Schödel et al., 2011), we found that these genes are responsive to MBNL2 only in A-498 cells, which express full-length HIF-1α transcript (Shinojima et al., 2007). In contrast, in 786-O cells, which do not express a full-length HIF-1α transcript (Maxwell et al., 1999; Shinojima et al., 2007), MBNL2 knockdown does not alter the expression of genes that are specifically bound by HIF-1α (Figure S5E), further supporting an indirect effect of MBNL2 on these genes that requires modulation of HIF-α TFs. Genes that are specifically bound by HIF-2α but not HIF-1α, on the other hand, were downregulated after MBNL2 knockdown in both A-498 and 786-O cells, consistent with the notion that HIF-2α is expressed and functional in both of these cell lines (Shinojima et al., 2007). The importance of hypoxia in progression of ccRCC and its driving role in cancer metastasis, as extensively described previously (reviewed by Keith et al., 2011), warrants the need for further studies to delineate the mechanisms through which MBNL2 can directly or indirectly modulate the expression of HIF-α TFs and their downstream targets.

It is notable that the model of RBP-mediated stability network that we used covers only 9% of all sequence-specific human RBPs with known RNA-binding domains, and even a smaller fraction of all human RBPs (Gerstberger et al., 2014). Therefore, the actual landscape of RBP dysregulation in ccRCC is probably more complex than the analysis presented here. Nonetheless, our study presents a general framework for analysis of RBP stability programs in ccRCC and highlights the regulatory mechanisms of the central pathways that underlie ccRCC progression, including hypoxia signaling and the switch to aerobic glycolysis. Given that these pathways are among the most attractive therapeutic targets for ccRCC (Riazalhosseini and Lathrop, 2016; Srinivasan et al., 2015), our results may open new avenues for targeted therapy in ccRCC. Similarly, our approach can be used for systematic analysis of post-transcriptional gene regulation in other cancers, and in other diseases in general.

EXPERIMENTAL PROCEDURES

Additional details on the experimental procedures and reagents used in this study are provided in the Supplemental Experimental Procedures and Table S2.

Expression Data and the Stability Code

Expression data were obtained for 45 pairs of ccRCC tumors and patient-matched normal tissues from CAGEKID (Scelo et al., 2014) and for 72 pairs of tumors and patient-matched normal tissues from TCGA (Cancer Genome Atlas Research Network, 2013). Gene-level measurements (fragments per kilobase per million [FPKM] for CAGEKID and RSEM-based transcript per million [TPM] for TCGA) were used to calculate $\log_{10}$ of fold change between each tumor and its patient-matched normal tissue (after adding a pseudocount equal to 1% of the median of each sample). Gene symbols were mapped to Ensembl v84 gene IDs, and genes with ambiguous ID mapping were discarded.

We obtained the set of high-confidence stability targets of 35 RBPs from a previous study (Ray et al., 2013), consisting of 3,483 mRNAs that contain at least one functional binding site in their 3′ UTR for at least one RBP. Of these, 3,376 genes had gene-level measurements in the CAGEKID dataset (1,151 for the TCGA dataset). This stability code was converted to a 3,376 × 35 "binding" matrix B (Figure 1A), in which +1 and −1 indicate functional binding of RBPs resulting in the stabilization and destabilization of the mRNA, respectively, and in which zero indicates no functional binding.

Modeling mRNA Stability and Inferring RBP Activity

For simplicity of modeling mRNA stability, we assume that all mRNA isoforms encoded by each gene have the same decay rate. This assumption is reasonable, considering that the stability code that we used consists of genes that have no alternative 3′ UTRs (Ray et al., 2013), and therefore, 3′ UTR-mediated regulation of mRNA stability should apply uniformly to all isoforms of each gene. With this assumption, at steady state, the abundance of the mRNA for gene g in tissue t can be expressed as:

$$[g]_t = r_{p,t}(g)/r_{d,t}(g),$$

where $r_{p,t}$ and $r_{d,t}$ denote rates of production (transcription) and decay in tissue t, respectively. Therefore, the logarithm of abundance can be expressed as:

$$log[g]_t = logr_{p,t}(g) - logr_{d,t}(g) = logr_{p,t}(g) + logs_{g,t},$$

where $s_{g,t}$ represents the stability of mRNA g in tissue t, defined as the multiplicative inverse of decay rate. We model the mRNA stability as the combinatorial effect of binding of RBPs to the 3′ UTR, similar to previous studies on modeling gene expression based on TF binding (Lee and Bussemaker, 2010):

$$s_{g,t} = \prod_i a_{i,t}^{b(g,i)}$$

Here, $a_{i,t}$ represents the activity of RBP i in tissue t, and $b(g,i)$ denotes the functional binding of RBP i to the 3′ UTR of mRNA g (+1 for stabilizing binding, 0 for no binding, and −1 for destabilizing binding, as indicated earlier). Therefore, the abundance of g can be modeled as:

$$log[g]_t = logr_{p,t}(g) + \sum_i log(a_{i,t}) \times b(g,i),$$

and the differential abundance of g between patient-matched tumor and normal tissue can be modeled as:

$$\Delta log[g] = \Delta logr_p(g) + \sum_i \Delta log(a_i) \times b(g,i),$$

where $\Delta log(a_i)$ is the differential activity of RBP i in tumor versus normal tissue. We assume that the differential rate of transcription is independent of the differential rate of decay (Alkallas et al., 2017), allowing us to estimate the values of $\Delta log(a_i)$ for the RBPs in each tumor-normal tissue pair by regression, using $\Delta log\ [g]$ (logarithm of fold change) as the response variable and the matrix B (RBP binding matrix) as the covariate. In other words, we assume that $\Delta log\ r_p(g)$ is not correlated with $\Delta log\ r_{d,t}(g)$, and, therefore, the regression coefficients in $\Delta log\ [g] \sim \Sigma_i\ \Delta log(a_i)$ are equal to the regression coefficients in the full model. For each tumor-normal tissue pair, least-squares regression was performed with bootstrapping to estimate the RBP differential activities as well as the confidence intervals. Specifically, we repeated the regression 1,000 times, each time sampling 3,376 mRNAs for CAGEKID (or 1,151 mRNAs for TCGA) with replacement. The average (μ) and SD (σ) of regression coefficients were estimated from the 1,000 bootstrapped regressions and were used to calculate a Z score in each tissue for each RBP as μ/σ. This Z score forms the core measurement of inferred differential RBP activities in this article.

To estimate the overall increase or decrease of RBP activities across all individuals, we calculated the Student's t score of the set of Z scores obtained for each RBP across the individuals. A large positive t score indicates that the inferred Z scores are overall above zero across the individuals, and a large negative t score indicates that the inferred Z scores are overall below zero. All data files and scripts for measuring RBP activities are available at http://csg.lab.mcgill.ca/sup/ccRCC_RBP/.

TMAs

TMA specimens used in this study have been described previously (Xu et al., 2013). Immunohistochemistry (IHC) was performed using a Ventana automated system and the iVIEW DAB Detection Kit (Ventana Medical Systems, Tucson, AZ, USA) at the IHC laboratory of McGill University Health Centre (MUHC). See the Supplemental Experimental Procedures for additional details.

Cell Culture

The established renal cancer cell lines 786-O and A-498 and the non-tumorigenic PRPTEC cells were purchased from the American Type Culture Collection (ATCC; Rockville, MD, USA). These cell lines were cultured according to the recommendations of ATCC in the appropriate cell culture media and were incubated at 37°C in a humidified incubator with 5% CO_2.

Real-Time qPCR

For each sample, 1 μg total RNA was used for reverse transcription followed by real-time qPCR reactions in triplicates. Expression of each mRNA was normalized to the expression of the housekeeping gene actin and was reported as $2^{-\Delta\Delta Ct}$. See the Supplemental Experimental Procedures for additional details.

RBP Knockdown

The shRNA-mediated knockdown (KD) of *ESRP2* and *MBNL2* was carried out based on previously described protocols (Papadakis et al., 2015). Briefly, lentiviral supernatants produced by HEK293T cells were used to infect cells. In parallel, supernatant harboring pLKO.1 was generated and used as control. Infected cells were selected with puromycin (2 μg/mL, Sigma) for 4–5 days. For PCBP2 knockdown, reverse transfection with 35 nM siRNA using Lipofectamine RNAiMAX (Invitrogen) was performed according to the manufacturer's instructions.

Western Blot Analysis

Protein extraction and western blotting from cell lines were performed as previously described (Jandaghi et al., 2016). See the Supplemental Experimental Procedures for additional details.

ESRP2- and MBNL2-Knockdown RNA-Seq

Total RNA was extracted from cells infected with *MBNL2* or *ESRP2* or control shRNAs, using the miRNeasy Mini Kit (QIAGEN) according to the supplier protocols. The RNA was used to generate rRNA-depleted first-strand cDNA libraries using TruSeq Stranded Total RNA-LT (Ribo-Zero Gold, Illumina). The libraries were sequenced on a HiSeq 2000 platform, producing a minimum of 50 million paired-end 100-bp reads per sample. Reads were mapped to the GRCh37 human genome assembly using STAR (Dobin et al., 2013). Number of reads mapping to each gene was calculated using HTSeq (Anders et al., 2015) based on gene annotations from GENCODE v19 (Harrow et al., 2012), and genes with fewer than an average of 10 reads per sample were discarded. The variance-stabilized logarithm of read counts was calculated using DESeq (Anders and Huber, 2010), and the values were normalized to have an average of zero for each sample. The effect of knockdown on the abundance of each mRNA was calculated as the difference of normalized values between KD and control shRNA samples, representing logarithm of fold change between the two conditions.

For functional enrichment analysis, Gene Ontology (GO) (The Gene Ontology Consortium, 2015) and pathway interaction database (PID) (Schaefer et al., 2009) annotations were obtained from ConsensusPathDB (Kamburov et al., 2011). The annotations were intersected with the expression measurements, and genes with no annotations were removed from the analysis. To measure the overall upregulation or downregulation of a particular gene category (i.e., a GO biological process, GO cell compartment, or PID pathway), the distribution of logarithm of fold change of genes that belonged to that category was compared to the distribution of logarithm of fold change of genes that did not belong to that category, using the Mann-Whitney U test. The p values were corrected for multiple testing using the Benjamini-Hochberg procedure.

RNA-seq data are deposited under accession number GEO: GSE83999.

RBP Immunoprecipitation

MBNL2 immunoprecipitation was performed using the Magna RIP Kit (Millipore) according to the manufacturer's protocol with slight modifications, as explained in the Supplemental Experimental Procedures. RNA abundances were normalized to the 18S rRNA as the internal control.

Pro-Collagen I Alpha 1 Detection by ELISA

An ELISA Kit from Abcam was used to quantify human COL1A1 protein in cell lysate. Total protein of the non-tumorigenic PRPTEC cells was isolated using M-PER lysis buffer and quantified using the BCA Protein Assay. 0.1 μg total protein and the antibody cocktail were loaded into each well of the 96-well plate and shaken gently for 1 hr at room temperature. Each well was washed four times with wash buffer and incubated 10 min with TMB ELISA Substrate. To detect horseradish peroxidase (HRP) activity, stop solution was added into each well, and the optical density (OD) was recorded at 450 nm using Tecan Infinite Microplate Reader.

Colony-Formation Assay

Cells infected with shRNA constructs were trypsinized and plated in a 6-well plate as single cells (2,000 cells per well), and colony formation was assessed after 10 days. See the Supplemental Experimental Procedures for additional details.

Fluorescence-Activated Cell Sorting Cell-Cycle Analysis

After 48 hr, knockdown and control cells were harvested with Accutase cell detachment solution (Sigma) and washed with cold PBS. Cold Nicoletti buffer (Nicoletti et al., 1991) was used for staining, and the DNA content of single nuclei was analyzed by BD FACSCanto II Flow Cytometry Analyzer Systems with collection of at least 10,000 events for each sample. The experiments were performed in triplicates.

Cell Viability and Caspase-3/7 Activity

Cell viability and apoptosis induction were examined using CellTiter-Glo and Caspase-Glo 3/7 Assay Kits (Promega), respectively, 48 hr (apoptosis) or 96 hr (cell viability) after transfection (PCBP2) or puromycin cell selection (MBNL2), as previously described (Jandaghi et al., 2016).

DATA AND SOFTWARE AVAILABILITY

The accession number for the RNA-seq data generated in this study and reported in this paper is GEO: GSE83999. Additional data files are available at http://csg.lab.mcgill.ca/sup/ccRCC_RBP/.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and two tables and can be found with this article online at https://doi.org/10.1016/j.celrep.2018.04.031.

ACKNOWLEDGMENTS

This work was supported by funds from the Canadian Institutes of Health Research (PJT-155966 to H.S.N. and Y.R.); an Alfred. P. Sloan Research fellowship (to H.S.N.); Compute Canada resource allocations (to H.S.N); the European 7th Framework Programme (241669, the CAGEKID project, http://www.cng.fr/cagekid); McGill University; Génome Québec; and the Ministère de l'Enseignement supérieur, de la Recherche, de la Science et de la Technologie (MESRST). G.P. is supported by a training scholarship from Fonds de recherche du Québec – Santé (FRQS).

AUTHOR CONTRIBUTIONS

G.P. and H.S.N. analyzed the data and developed the computational methods. P.J. designed and performed the functional experiments with contributions from M.S., A.I.P., and M.A. S.S., C.S., and F.B. performed the IHC experiments and TMA analysis. P.J. and M.K. contributed to data analysis. G.S., S.H., R.E.B., J.T., A.B., and S.T. contributed reagents and materials and edited the manuscript. M.L. provided critical advice on data analysis and statistical approaches. H.S.N. and Y.R. conceived and directed the study, designed the experiments, and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 28, 2017
Revised: March 3, 2018
Accepted: April 6, 2018
Published: May 8, 2018

REFERENCES

Adereth, Y., Dammai, V., Kose, N., Li, R., and Hsu, T. (2005). RNA-dependent integrin alpha3 protein localization regulated by the Muscleblind-like protein MLP1. Nat. Cell Biol. *7*, 1240–1247.

Alkallas, R., Fish, L., Goodarzi, H., and Najafabadi, H.S. (2017). Inference of RNA decay rate from transcriptional profiling highlights the regulatory programs of Alzheimer's disease. Nat. Commun. *8*, 909.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Cancer Genome Atlas Research Network (2013). Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature *499*, 43–49.

Capitanio, U., and Montorsi, F. (2016). Renal cancer. Lancet *387*, 894–906.

Charizanis, K., Lee, K.Y., Batra, R., Goodwin, M., Zhang, C., Yuan, Y., Shiue, L., Cline, M., Scotti, M.M., Xia, G., et al. (2012). Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. Neuron *75*, 437–450.

Correa, B.R., de Araujo, P.R., Qiao, M., Burns, S.C., Chen, C., Schlegel, R., Agarwal, S., Galante, P.A., and Penalva, L.O. (2016). Functional genomics analyses of RNA-binding proteins reveal the splicing regulator SNRPB as an oncogenic candidate in glioblastoma. Genome Biol. *17*, 125.

Croft, D., Mundo, A.F., Haw, R., Milacic, M., Weiser, J., Wu, G., Caudy, M., Garapati, P., Gillespie, M., Kamdar, M.R., et al. (2014). The Reactome pathway knowledgebase. Nucleic Acids Res. *42*, D472–D477.

Dang, H., Takai, A., Forgues, M., Pomyen, Y., Mou, H., Xue, W., Ray, D., Ha, K.C.H., Morris, Q.D., Hughes, T.R., et al. (2017). Oncogenic activation of the RNA binding protein NELFE and MYC signaling in hepatocellular carcinoma. Cancer Cell *32*, 101–114.e8.

Ding, Y., Huang, D., Zhang, Z., Smith, J., Petillo, D., Looyenga, B.D., Feenstra, K., Mackeigan, J.P., Furge, K.A., and Teh, B.T. (2011). Combined gene expression profiling and RNAi screening in clear cell renal cell carcinoma identify PLK1 and other therapeutic kinase targets. Cancer Res. *71*, 5225–5234.

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

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Frantz, C., Stewart, K.M., and Weaver, V.M. (2010). The extracellular matrix at a glance. J. Cell Sci. *123*, 4195–4200.

The Gene Ontology Consortium (2015). Gene Ontology Consortium: going forward. Nucleic Acids Res. *43*, D1049–D1056.

Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. Nat. Rev. Genet. *15*, 829–845.

Gilkes, D.M., Semenza, G.L., and Wirtz, D. (2014). Hypoxia and the extracellular matrix: drivers of tumour metastasis. Nat. Rev. Cancer *14*, 430–439.

Gonzalez, D.M., and Medici, D. (2014). Signaling mechanisms of the epithelial-mesenchymal transition. Sci. Signal. *7*, re8.

Goodarzi, H., Zhang, S., Buss, C.G., Fish, L., Tavazoie, S., and Tavazoie, S.F. (2014). Metastasis-suppressor transcript destabilization through TARBP2 binding of mRNA hairpins. Nature *513*, 256–260.

Goolam, M., Scialdone, A., Graham, S.J.L., Macaulay, I.C., Jedrusik, A., Hupalowska, A., Voet, T., Marioni, J.C., and Zernicka-Goetz, M. (2016). Heterogeneity in Oct4 and Sox2 targets biases cell fate in 4-cell mouse embryos. Cell *165*, 61–74.

Han, H., Irimia, M., Ross, P.J., Sung, H.K., Alipanahi, B., David, L., Golipour, A., Gabut, M., Michael, I.P., Nachman, E.N., et al. (2013a). MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. Nature *498*, 241–245.

Han, W., Xin, Z., Zhao, Z., Bao, W., Lin, X., Yin, B., Zhao, J., Yuan, J., Qiang, B., and Peng, X. (2013b). RNA-binding protein PCBP2 modulates glioma growth by regulating FHL3. J. Clin. Invest. *123*, 2103–2118.

Harris, A.L. (2002). Hypoxia–a key regulatory factor in tumour growth. Nat. Rev. Cancer *2*, 38–47.

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

Jandaghi, P., Najafabadi, H.S., Bauer, A.S., Papadakis, A.I., Fassan, M., Hall, A., Monast, A., von Knebel Doeberitz, M., Neoptolemos, J.P., Costello, E., et al. (2016). Expression of DRD2 is increased in human pancreatic ductal adenocarcinoma and inhibitors slow tumor growth in mice. Gastroenterology *151*, 1218–1231.

Kamburov, A., Pentchev, K., Galicka, H., Wierling, C., Lehrach, H., and Herwig, R. (2011). ConsensusPathDB: toward a more complete picture of cell biology. Nucleic Acids Res. *39*, D712–D717.

Kechavarzi, B., and Janga, S.C. (2014). Dissecting the expression landscape of RNA-binding proteins in human cancers. Genome Biol. *15*, R14.

Keith, B., Johnson, R.S., and Simon, M.C. (2011). HIF1α and HIF2α: sibling rivalry in hypoxic tumour growth and progression. Nat. Rev. Cancer *12*, 9–22.

Kim, K.H., and Roberts, C.W. (2016). Targeting EZH2 in cancer. Nat. Med. *22*, 128–134.

Lee, E., and Bussemaker, H.J. (2010). Identifying the genetic determinants of transcription factor activity. Mol. Syst. Biol. *6*, 412.

Loayza-Puch, F., Rooijers, K., Buil, L.C., Zijlstra, J., Oude Vrielink, J.F., Lopes, R., Ugalde, A.P., van Breugel, P., Hofland, I., Wesseling, J., et al. (2016). Tumour-specific proline vulnerability uncovered by differential ribosome codon reading. Nature *530*, 490–494.

Maxwell, P.H., Wiesener, M.S., Chang, G.W., Clifford, S.C., Vaux, E.C., Cockman, M.E., Wykoff, C.C., Pugh, C.W., Maher, E.R., and Ratcliffe, P.J. (1999). The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. Nature *399*, 271–275.

Menke, J., Kriegsmann, J., Schimanski, C.C., Schwartz, M.M., Schwarting, A., and Kelley, V.R. (2012). Autocrine CSF-1 and CSF-1 receptor coexpression promotes renal cell carcinoma growth. Cancer Res. *72*, 187–200.

Mizutani, A., Koinuma, D., Seimiya, H., and Miyazono, K. (2016). The Arkadia-ESRP2 axis suppresses tumor progression: analyses in clear-cell renal cell carcinoma. Oncogene *35*, 3514–3523.

Nicoletti, I., Migliorati, G., Pagliacci, M.C., Grignani, F., and Riccardi, C. (1991). A rapid and simple method for measuring thymocyte apoptosis by propidium iodide staining and flow cytometry. J. Immunol. Methods *139*, 271–279.

Papadakis, A.I., Sun, C., Knijnenburg, T.A., Xue, Y., Grernrum, W., Hölzel, M., Nijkamp, W., Wessels, L.F., Beijersbergen, R.L., Bernards, R., and Huang, S. (2015). SMARCE1 suppresses EGFR expression and controls responses to MET and ALK inhibitors in lung cancer. Cell Res. *25*, 445–458.

Perera, R.M., Stoykova, S., Nicolay, B.N., Ross, K.N., Fitamant, J., Boukhali, M., Lengrand, J., Deshpande, V., Selig, M.K., Ferrone, C.R., et al. (2015). Transcriptional control of autophagy-lysosome function drives pancreatic cancer metabolism. Nature *524*, 361–365.

Ray, D., Kazan, H., Cook, K.B., Weirauch, M.T., Najafabadi, H.S., Li, X., Guerroussov, S., Albu, M., Zheng, H., Yang, A., et al. (2013). A compendium of RNA-binding motifs for decoding gene regulation. Nature *499*, 172–177.

Riazalhosseini, Y., and Lathrop, M. (2016). Precision medicine from the renal cancer genome. Nat. Rev. Nephrol. *12*, 655–666.

Scelo, G., Riazalhosseini, Y., Greger, L., Letourneau, L., Gonzàlez-Porta, M., Wozniak, M.B., Bourgey, M., Harnden, P., Egevad, L., Jackson, S.M., et al.

(2014). Variation in genomic landscape of clear cell renal cell carcinoma across Europe. Nat. Commun. *5*, 5135.

Schaefer, C.F., Anthony, K., Krupa, S., Buchoff, J., Day, M., Hannay, T., and Buetow, K.H. (2009). PID: the Pathway Interaction Database. Nucleic Acids Res. *37*, D674–D679.

Schödel, J., Oikonomopoulos, S., Ragoussis, J., Pugh, C.W., Ratcliffe, P.J., and Mole, D.R. (2011). High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. Blood *117*, e207–e217.

Shapiro, I.M., Cheng, A.W., Flytzanis, N.C., Balsamo, M., Condeelis, J.S., Oktay, M.H., Burge, C.B., and Gertler, F.B. (2011). An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype. PLoS Genet. *7*, e1002218.

Shinojima, T., Oya, M., Takayanagi, A., Mizuno, R., Shimizu, N., and Murai, M. (2007). Renal cancer cells lacking hypoxia inducible factor (HIF)-1alpha expression maintain vascular endothelial growth factor expression through HIF-2alpha. Carcinogenesis *28*, 529–536.

Srinivasan, R., Ricketts, C.J., Sourbier, C., and Linehan, W.M. (2015). New strategies in renal cell carcinoma: targeting the genetic and metabolic basis of disease. Clin. Cancer Res. *21*, 10–17.

Taliaferro, J.M., Vidaki, M., Oliveira, R., Olson, S., Zhan, L., Saxena, T., Wang, E.T., Graveley, B.R., Gertler, F.B., Swanson, M.S., and Burge, C.B. (2016). Distal alternative last exons localize mRNAs to neural projections. Mol. Cell *61*, 821–833.

Warzecha, C.C., Sato, T.K., Nabet, B., Hogenesch, J.B., and Carstens, R.P. (2009). ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing. Mol. Cell *33*, 591–601.

Warzecha, C.C., Jiang, P., Amirikian, K., Dittmar, K.A., Lu, H., Shen, S., Guo, W., Xing, Y., and Carstens, R.P. (2010). An ESRP-regulated splicing programme is abrogated during the epithelial-mesenchymal transition. EMBO J. *29*, 3286–3300.

Wu, X.R., Chen, Y.H., Liu, D.M., Sha, J.J., Xuan, H.Q., Bo, J.J., and Huang, Y.R. (2013). Increased expression of forkhead box M1 protein is associated with poor prognosis in clear cell renal cell carcinoma. Med. Oncol. *30*, 346.

Xu, B., Abourbih, S., Sircar, K., Kassouf, W., Mansure, J.J., Aprikian, A., Tanguay, S., and Brimo, F. (2013). Enhancer of zeste homolog 2 expression is associated with metastasis and adverse clinical outcome in clear cell renal cell carcinoma: a comparative study and review of the literature. Arch. Pathol. Lab. Med. *137*, 1326–1336.

Xue, Y.J., Xiao, R.H., Long, D.Z., Zou, X.F., Wang, X.N., Zhang, G.X., Yuan, Y.H., Wu, G.Q., Yang, J., Wu, Y.T., et al. (2012). Overexpression of FoxM1 is associated with tumor progression in patients with clear cell renal cell carcinoma. J. Transl. Med. *10*, 200.

Zaravinos, A., and Deltas, C. (2014). ccRCC is fundamentally a metabolic disorder. Cell Cycle *13*, 2481–2482.

Zhang, G., Zhang, Z., and Liu, Z. (2013). Polo-like kinase 1 is overexpressed in renal cancer and participates in the proliferation and invasion of renal cancer cells. Tumour Biol. *34*, 1887–1894.