

SCIENTIFIC REPORTS



OPEN

The impact of the RBM4-initiated splicing cascade on modulating the carcinogenic signature of colorectal cancer cells

Received: 11 November 2016

Accepted: 06 February 2017

Published: 09 March 2017

Jung-Chun Lin¹, Yuan-Chii Lee^{2,*}, Yu-Chih Liang^{1,*}, Yang C. Fann³, Kory R. Johnson³ & Ying-Ju Lin⁴

A growing body of studies has demonstrated that dysregulated splicing profiles constitute pivotal mechanisms for carcinogenesis. In this study, we identified discriminative splicing profiles of colorectal cancer (CRC) cells compared to adjacent normal tissues using deep RNA-sequencing (RNA-seq). The RNA-seq results and cohort studies indicated a relatively high ratio of exon 4-excluded *neuro-oncological ventral antigen 1* (*Nova1*⁻⁴) and intron 2-retained *SRSF6* (*SRSF6*^{+intron 2}) transcripts in CRC tissues and cell lines. *Nova1* variants exhibited differential effects on eliminating *SRSF6* expression in CRC cells by inducing *SRSF6*^{+intron 2} transcripts which were considered to be the putative target of alternative splicing-coupled nonsense-mediated decay mechanism. Moreover, the splicing profile of *vascular endothelial growth factor (VEGF)165/VEGF165b* transcripts was relevant to *SRSF6* expression, which manipulates the progression of CRC cells. These results highlight the novel and hierarchical role of an alternative splicing cascade that is involved in the development of CRC.

Alternative splicing constitutes a major mechanism in expanding the proteome diversity of eukaryotic cells¹. Recent studies demonstrated that more than 90% of protein-coding genes generate more than one transcript through this meticulously controlled process². In normal tissues, spatiotemporal expression profiles of alternative splicing events determine cell differentiation and specification³, whereas imbalanced or aberrant alternative splicing profiles are highly relevant to hereditary diseases and cancers^{4,5}. The interplay between splicing factors, such as serine-arginine (SR) or the heteronuclear ribonucleoprotein (hnRNP) protein family, and *cis*-elements within cassette exons of alternative splicing events, constitutes the molecular mechanism of programming splicing profiles⁶. Altered expressions or distributions of splicing factors were subsequently noted to be an efficient means for reprogramming splicing profiles of cancer cells^{7,8}.

Statistical studies indicated that while colorectal cancer (CRC) is the third most common and lethal cancer in the human race, the etiology and molecular mechanisms leading to CRC are largely unclear⁹. However, several alternative splicing events are widely observed and relevant to the development of CRC¹⁰. For instance, upregulation of *SRSF10* expression results in an increase in *BCLAF1* exon 5a-included transcripts which facilitates the progression of CRC¹¹. In addition, an increase in the expression of polypyrimidine tract-binding protein 1 (*PTBP1*) leads to a relatively high level of the *PKM2* variant which mediates the Warburg effect in CRC cells¹². Our previous work reported that overexpression of RNA-binding motif protein 4 (*RBM4*) constitutes a regulatory network in reprogramming splicing profiles of the *FGFR2* and *PKM* genes, which modulate the progression and metabolic signature of CRC cells¹³.

With the development of high-throughput approaches, including deep RNA sequencing (RNA-seq) and RNA crosslink immunoprecipitation-coupled proteomic analyses, alternative splicing networks can be investigated on a genome-wide level^{14,15}. In this study, we performed RNA-seq analyses to thoroughly annotate transcriptomes in

¹School of Medical Laboratory Science and Biotechnology, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan. ²Graduate Institute of Biomedical Informatics, Taipei Medical University, Taipei, Taiwan. ³Information Technology and Bioinformatics Program, Division of Intramural Research, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA. ⁴School of Chinese Medicine, China Medical University, Taichung, Taiwan. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.-C.L. (email: lin2511@tmu.edu.tw)

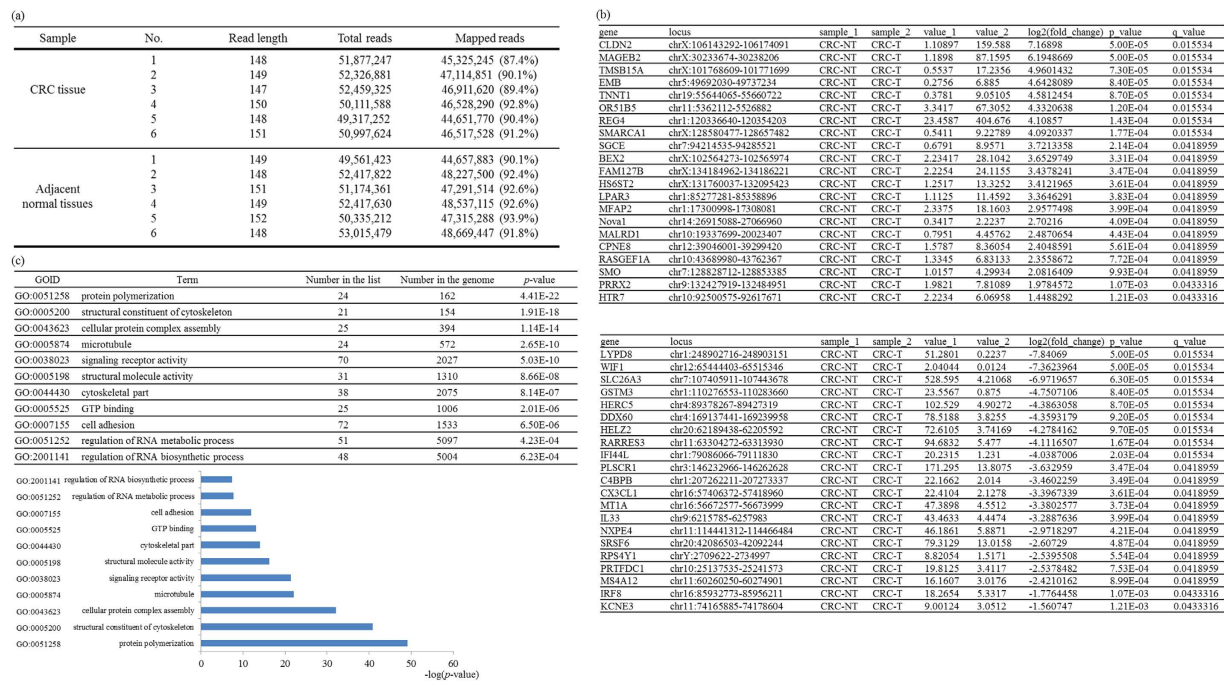


Figure 1. Transcriptome and gene ontology analyses of paired colorectal cancer (CRC) tissues. (a) Statistical summaries of transcriptome analytical results. **(b)** Transcriptome analysis-identified splicing events significantly changed in CRC tissues compared to adjacent normal tissues. **(c)** The top 11 enriched functions for differentially spliced genes in CRC tissues.

CRC tissues and adjacent normal tissues. Among these alternative splicing events, RBM4, Nov1, SRSF6, and vascular endothelial growth factor 165 (VEGF165) comprised a splicing cascade in a consensus sequence-dependent manner which manipulated the migration and angiogenic signature of CRC cells. These results suggested the potential value of splicing events as clinical applications for CRC treatment.

Results

Genome-wide analyses of CRC-associated alternative splicing events. Disclosing an imbalanced or aberrant splicing event may bring about a comprehensive understanding of the genetic causes of distinct malignancies. We therefore performed transcriptome analyses using deep RNA-seq with total RNAs extracted from cancerous and adjacent normal tissues of anonymous CRC patients. The statistical summary of the RNA-seq output showed close numbers of read lengths, amplified reads, and mapping rates within the independent analyses (Fig. 1a, $n = 6$). After filtering differentially spliced transcripts with a false detection rate (FDR)-adjusted p value of < 0.01 , a q value of < 0.05 , and a multiple of change of > 2 , 153 genes from a total of 22,518 splicing events that significantly differed in cancerous tissues compared to adjacent normal counterparts were identified. Figure 1b presents 42 genes which exhibited differential splicing profiles between CRC cancerous tissues and adjacent normal tissues. For instance, the relative level of the *Nov1*⁻⁴ transcript was significantly upregulated in cancerous tissues (Fig. 1b, 2.2237) compared to adjacent normal tissues (Fig. 1b, 0.3417). In contrast, relative expression of the non-coding *SRSF6* transcripts in cancerous tissues (Fig. 1b, 79.3127) was more abundant than that of adjacent normal tissues (Fig. 1b, 13.0158). Gene ontology (GO) analyses were next conducted by examining putative CRC-related splicing events. Figure 1c shows the analytical summary (upper) and the bar chart (lower) presents the p values of the GO analyses. Results showed the enrichment of CRC-related splicing events involved in cytoskeletal, signaling receptor, and cell adhesion terms, which are highly related to the progression of cancer cells. Moreover, enriched terms corresponding to RNA metabolism and biosynthesis also implied a potential impact of RNA-binding proteins, including *Nov1* and *SRSF6* genes (Fig. 1b), on the carcinogenic signature of CRC cells.

Differential expressions and splicing profiles of *Nov1* and *SRSF6* in CRC tissues. To validate the identified results regarding the splicing profiles of *Nov1* and *SRSF6* genes, RT-PCR analyses were conducted with RNAs prepared from cancerous tissues and adjacent normal tissues of anonymous CRC patients (Fig. 2a, $n = 10$). Figure 2a shows the predominant expression of exon 4-excluded *Nov1* transcripts (*Nov1*⁻⁴, 60.1%, white bar) in cancerous tissues, whereas the reduced expression of *Nov1*⁻⁴ transcript (12.5%, white bar) was noted in the adjacent normal tissues of CRC patients. Furthermore, results of the RT-qPCR showed elevated total *Nov1* transcripts in cancerous tissues compared to adjacent normal tissues (by about 4.4-fold; Fig. 2a, qRT-PCR). In addition to *Nov1*, RNA-seq results also indicated differential splicing profiles of the *SRSF6* gene between normal and cancerous tissues (Fig. 1b, lower table). The RT-PCR results showed that adjacent normal tissues exhibited predominant expression of authentic *SRSF6* transcripts (Fig. 2a, NM_006275; 87%, black bar) encoding functional *SRSF6* proteins, whereas splicing profiles of the *SRSF6* gene in cancerous tissues were shifted to non-coding *SRSF6*

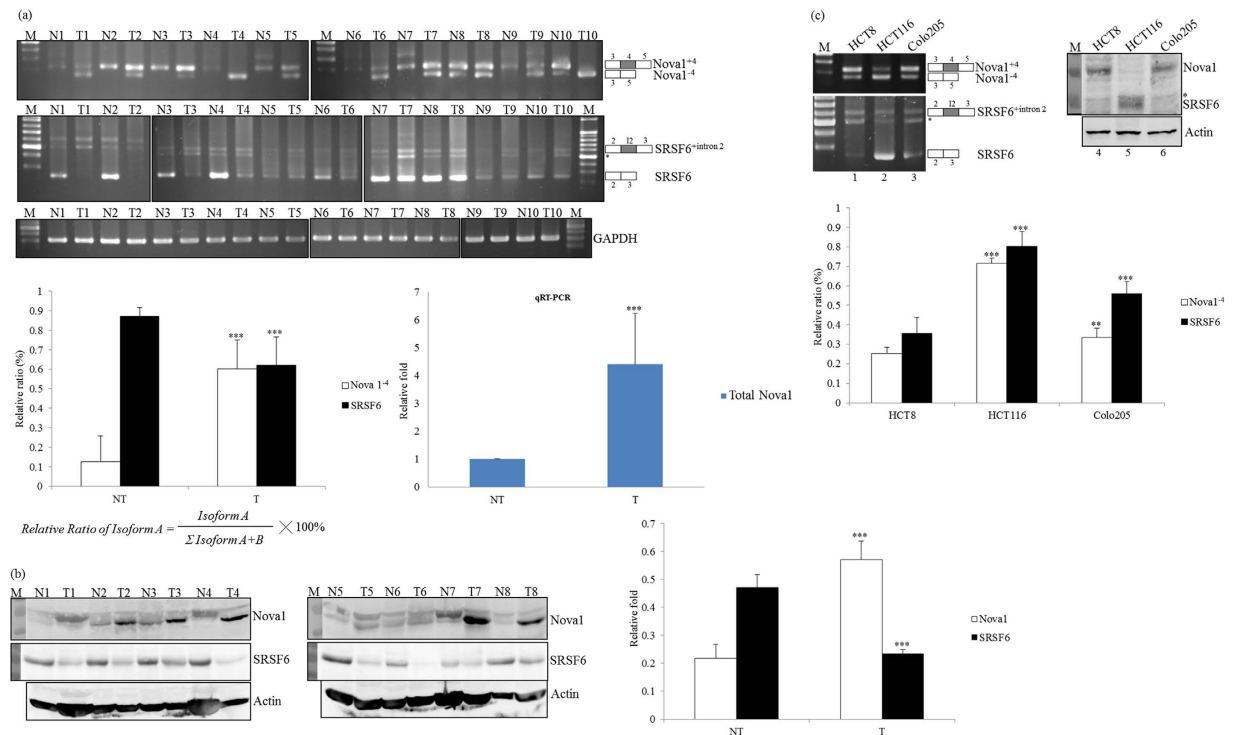


Figure 2. Differential splicing and expression profiles of *Nova1* and *SRSF6* in colorectal cancer (CRC) tissues and derived cell lines. (a) Total RNAs prepared from human CRC tissues (T) and adjacent non-tumorous tissues (NT) were subjected to RT-PCR and RT-qPCR analyses with specific primer sets complementary to the *Nova1* and *SRSF6* transcripts. The bar graphs present relative levels of the *Nova1* and *SRSF6* transcripts (left) or total *Nova1* transcripts (right) in paired CRC tissues ($n = 10$). (b) Total lysates prepared from human CRC tissues (T) and adjacent non-tumorous tissues (NT) were analyzed using an immunoblot assay with the indicated antibodies. The bar graph presents relative levels of *Nova1* and *SRSF6* proteins in paired CRC tissues ($n = 8$). (c) Total RNAs and lysates prepared from CRC-derived cell lines were analyzed by RT-PCR assays with a specific primer set against the *Nova1* and *SRSF6* genes. Western blotting was performed with the indicated antibodies. The bar graph presents relative levels of the *Nova1*⁺⁴ and *SRSF6* transcripts in three independent experiments. The gels shown in this figure were run under the same conditions and were not artificially manipulated. The signal densities were analyzed using TotalLab Quant Software, and the quantitative results are shown as the mean \pm STD ($*p < 0.05$; $**p < 0.01$; $***p < 0.005$).

transcripts (NR_034009, 38%, black bar), which are the potential target of the nonsense-mediated decay (NMD) mechanism¹⁶. As expected, the analytical results of immunoblotting revealed upregulated expression of the *Nova1* protein (by about 2.64-fold) in cancerous tissues (Fig. 2b, white bar). The abundance of the *SRSF6* protein was closely related to the relative level of coding *SRSF6* transcripts in paired CRC tissues (Fig. 2b, black bar). In addition, differential splicing profiles of *Nova1* and *SRSF6* were noted in distinct CRC cell lines. The RT-PCR results indicated close expressions of *Nova1*⁻⁴ (Fig. 2c, 25.2% and 33.5%, white bar) and *SRSF6* transcripts (Fig. 2c, 35.5% and 56.1%, black bar) in the HCT8 and Colo205 colorectal adenocarcinoma cell lines, respectively. In contrast, predominant expressions of the *Nova1*⁻⁴ (Fig. 2c, 71.5%, white bar) and authentic *SRSF6* transcripts (Fig. 2c, 80.2%, black bar) were observed in HCT116 cells derived from colorectal carcinoma. Expression profiles of the *Nova1* and *SRSF6* transcripts were highly relevant to the encoded proteins in distinct CRC cells (Fig. 2c, right panel). The molecular mechanisms contributing to the differential splicing profiles of *Nova1* or *SRSF6* in different CRC cell lines are worthy of further investigation.

RBM4 and *Nova1* constitute an antagonistic circuit toward the selection of *Nova1* exon 4.

Post-transcriptional mechanisms, including alternative splicing and mRNA turnover, constitute multilayer control for regulating gene expressions¹⁷. In our previous study, overexpression of the PTBP2 protein exhibited multiple functions related to the stability and splicing profile of the *Nova1* gene, which abolishes brown adipogenesis¹⁸. The relatively high expression of PTBP2 with a concomitant reduction in the RBM4 protein was documented in HCT8 and Colo205 cells compared to that of HCT116 cells¹³, which may also manipulate splicing profiles of the *Nova1* and *SRSF6* genes (Fig. 2c). The alignment results showed high identity between human and mouse *Nova1* exon 4 and the downstream intron containing multiple YCA motifs (Fig. 3a), although the intronic CU element next to human *Nova1* exon 4 was shortened by interruption with a single guanine residue (Fig. 3a, gray characters). The RT-PCR results revealed the effect of overexpressing PTBP2 on enhancing relative levels of *Nova1*⁻⁴ transcripts (Fig. 3a, lane 2), whereas relatively high expression levels of *Nova1*⁺⁴ transcripts were observed with overexpression of RBM4 and the derived mutant containing the serine-to-alanine (SA) substitution in HCT8 cells

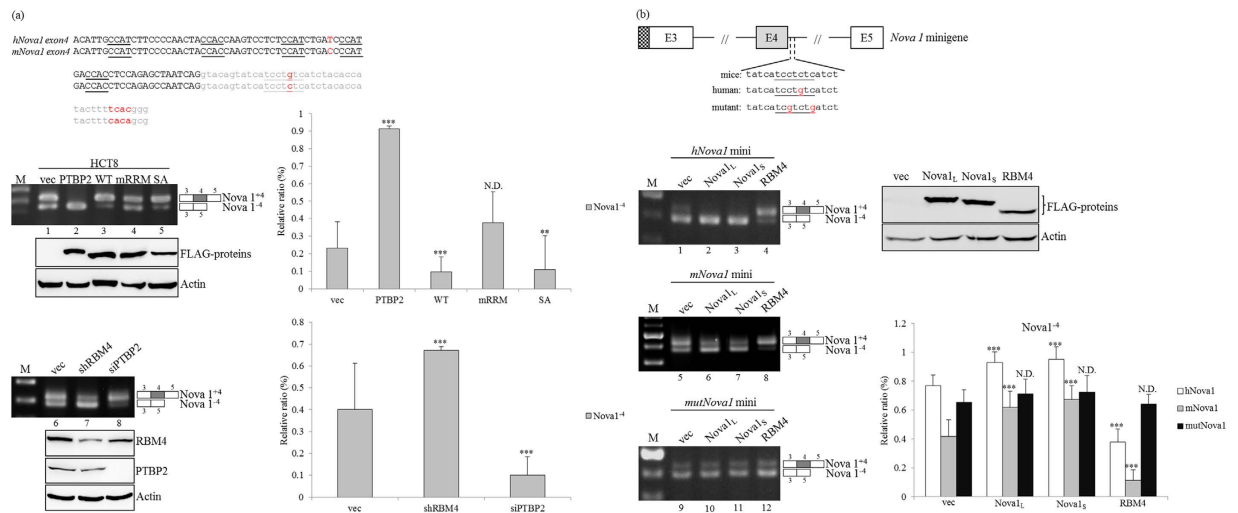


Figure 3. RBM4 and Nova1 exert differential effects on splicing profiles of the *Nova1* gene. (a) The diagram presents the sequence of human and mouse *Nova1* exon 4 and flanking introns. The responsive elements of *Nova1* (underlined black characters) and RBM4 (underlined gray characters) are underlined. HCT-8 cells were transfected with overexpressing or targeting vectors, followed by total RNA and lysate extraction. Splicing profiles of *Nova1* transcripts were analyzed with specific primer sets. Immunoblot assays of indicated proteins were analyzed with specific antibodies. The bar graph shows relative levels of *Nova1*⁻⁴ in three independent experiments. (b) *Nova1* minigenes were cotransfected with expressing vectors into HCT-8 cells. Spliced transcripts of *Nova1* minigenes were analyzed using SV40 oligo and a specific *Nova1* primer as described in the previous panel. The bar graph shows the relative level of *Nova1*⁻⁴ in three independent experiments using TotalLab Quant Software. The quantitative results are shown as the mean ± STD ($p < 0.05$; $**p < 0.01$; $***p < 0.005$). The gels shown in this figure were run under the same conditions and were not artificially manipulated.

(Fig. 3a, lanes 3 and 5). Nevertheless, the splicing profile of *Nova1* showed no response to the presence of an RNA recognition motif (RRM) mutant (Fig. 3a, lane 4), consistently implying the influence of RBM4 on programming splicing profiles through a post-transcriptional mechanism¹⁶. Inversely, depletion of endogenous RBM4 upregulated relatively high levels of *Nova1*⁻⁴ transcripts (Fig. 3a, lane 7), whereas a reduction in PTBP2 mediated an increase in *Nova1*⁺⁴ transcripts (Fig. 3a, lane 8).

The interplay between RBM4 and the intronic CU element next to *Nova1* exon 4 enhances its inclusion in murine brown adipocytes¹⁶. Moreover, *Nova1* isoforms constitute an autoregulatory mechanism on its splicing profiles¹⁹. To evaluate the effects of RBM4 and *Nova1* variants on the human *Nova1* gene, the human *Nova1* minigene containing the cytosine-to-guanine nucleotide substitution within the intronic CU element was derived from the mouse *Nova1* minigene (Fig. 3b, human). The derived mutant containing two cytosine-to-guanine nucleotide substitutions within the intronic CU element was applied in our previous study¹⁶. *In vivo* splicing results showed that more *Nova1*⁺⁴ transcripts were generated from the mouse *Nova1* minigene compared to that of the human *Nova1* reporter (Fig. 3b, lanes 1 and 5). Nevertheless, overexpression of RBM4 enhanced relatively high levels of *Nova1*⁺⁴ transcripts produced from both human and mouse *Nova1* minigenes (Fig. 3b, lanes 4 and 8). Overexpressing *Nova1*_S variants exerted a more-prominent effect (Fig. 3b, lanes 3 and 7; 95.2% and 67.2%) than that of *Nova1*_L variants (Fig. 3b, lanes 2 and 6; 92.9% and 61.8%) on enhancing the relative level of *Nova1*⁻⁴ transcripts generated from mouse or human *Nova1* minigenes. The mutant minigene containing two cytosine-to-guanine nucleotide substitutions within the intronic CU element showed an insubstantial response to the overexpression of RBM4 or *Nova1* variants (Fig. 3b, lanes 10 ~ 12). Collectively, RBM4 constitutes a regulatory mechanism on the utilization of *Nova1* exon 4 through the intronic CU element.

RBM4 and *Nova1* exert differential effects on the migration of CRC cell lines. RBM4 was documented to reduce immortalization, proliferation, and migration of distinct cancer cells by programming the related splicing profiles²⁰. In contrast, upregulation of *Nova1* expression is closely associated with the progression of astrocytomas, hepatocellular carcinoma, and gastric cancer^{21,22}. *In vitro* migration assays were conducted to demonstrate the phenomenon in CRC cell lines. Compared to the empty vector-transfected counterparts, RBM4-overexpressing cells exhibited less migratory activity (Fig. 4a, RBM4). Contrarily, overexpression of *Nova1* variants enhanced the migratory activity of HCT8 cells (Fig. 4a, *Nova1*_S and *Nova1*_L). Furthermore, imaging and statistical results indicated a more-potent effect of the *Nova1*_S isoform than that of *Nova1*_L variants on the migration of CRC cells (Fig. 4a, bar chart). In addition, results of the RT-PCR, RT-qPCR, and immunoblot analyses showed that overexpression of RBM4 enhanced expression of E-cadherin and led to a concomitant decrease in N-cadherin (Fig. 4b, lanes 2 and 6; bar chart) which are critical factors involved in the epithelial-to-mesenchymal transition (EMT)²³. Overexpression of *Nova1* isoforms exerted opposite effects on diminishing E-cadherin expression and the concomitant upregulation of N-cadherin levels in CRC cells (Fig. 4b, lanes 3, 4, 7, and 8).

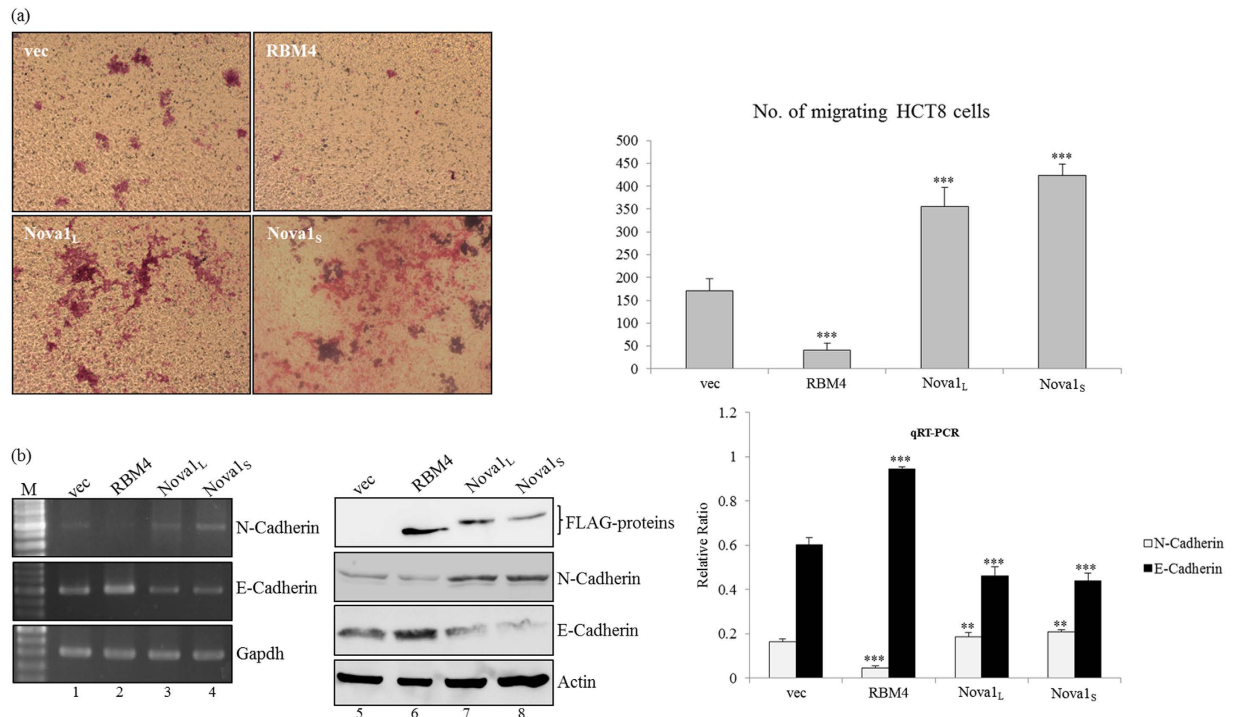


Figure 4. Overexpressing RBM4 and Nova1 variants differentially modulate the migratory activity of colorectal cancer (CRC) cells. (a) The motility of HCT-8 cells transfected with RBM4 or Nova1 variant-expressing vectors was analyzed using an *in vitro* migration assay with a transwell system, followed by Giemsa staining. The counting results of three independent experiments are presented in a bar chart as the mean \pm STD. (b) Total RNAs and lysates were prepared from HCT-8 cells transfected with RBM4- or Nova1-expressing vectors. Expression profiles of cadherin genes and the loading control were analyzed by RT-PCR, RT-qPCR, and immunoblot assays with specific primer sets and antibodies against indicated targets. Western blotting was performed with the indicated antibodies. The gels shown in this figure were run under the same conditions and were not artificially manipulated. The bar graph presents relative levels of each target in three independent RT-qPCR analyses. Signal densities were analyzed using TotalLab Quant Software and the quantitative results are shown as the mean \pm STD (* p < 0.05; ** p < 0.01; *** p < 0.005).

These results suggested that RBM4 and Nova1 proteins potentially constitute an antagonistic mechanism that manipulates the EMT-related pathway in CRC cells.

RBM4 and Nova1 exhibit opposite effects on splicing profiles of SRSF6. In addition to *Nova1*, differential splicing profiles of the *SRSF6* gene in CRC tissues compared to adjacent normal tissues were revealed using RNA-seq (Fig. 1b, lower panel) and RT-PCR analyses (Fig. 2a). *Nova1* was documented to facilitate the inclusion of mouse *SRSF6* intron 2, which shares complete identity with the human *SRSF6* gene (Fig. 5a), in a YCAI element-dependent manner¹⁶. Splicing profiles of human *SRSF6* were therefore examined with the overexpression or depletion of *Nova1* or RBM4 in CRC cells. As shown in Fig. 5b, the presence of RBM4-WT and the derived S309A mutant induced a relatively high level of authentic *SRSF6* transcripts (Fig. 5b, lanes 3 and 5), whereas overexpression of the *Nova1_L* variant reduced the relative expression of coding *SRSF6* transcripts in HCT8 cells (Fig. 5b, lane 2). The splicing pattern of *SRSF6* was sustained with overexpression of the RBM4 RNA recognition motif-mutant containing four mutations (Y37A, F39A, Y113A, and F115A) (Fig. 5b, lane 4). Inversely, ablation of endogenous RBM4 mediated upregulation of *SRSF6*^{+intron 2} transcripts (Fig. 5b, lane 7), whereas upregulated generation of *SRSF6* transcripts was observed in *Nova1*-knockdown HCT8 cells (Fig. 5b, lane 8). To validate the inference of the *Nova1*-regulated mechanism on *SRSF6* splicing, *in vivo* splicing assays were conducted with the *SRSF6* minigene and the previously established derived mutant¹⁶. Increases in intron 2-skipped transcripts generated from the *SRSF6* minigene were observed with RBM4a overexpression compared to empty-vector transfectants (Fig. 5c, lanes 1 and 2). In contrast, relatively high expression levels of *SRSF6*^{+intron 2} transcripts were noted in *Nova1* variants-overexpressing HCT8 cells (Fig. 5c, lanes 3 and 4). The mutant minigene harboring the CA substitution within the CCAC motif (Fig. 5c, diagram) showed an insubstantial response to overexpression of the RBM4 or *Nova1* isoforms (Fig. 5c, lanes 5~8).

SRSF6 was reported to program cancer-related splicing profiles, including *CD44* and *corticotropin-releasing hormone receptor* genes, which subsequently inhibit the progression of cancer cells^{24,25}. Results of the *in vitro* migration assays and statistical analyses showed that *SRSF6* overexpression (Fig. 5d, *SRSF6*) diminished the migratory activity of HCT8 cells, whereas *SRSF6*-knockdown (Fig. 5d, sh*SRSF6*) cells exhibited more-substantial migration compared to empty-vector transfectants. In contrast, ablation of *Nova1* abrogated the progression

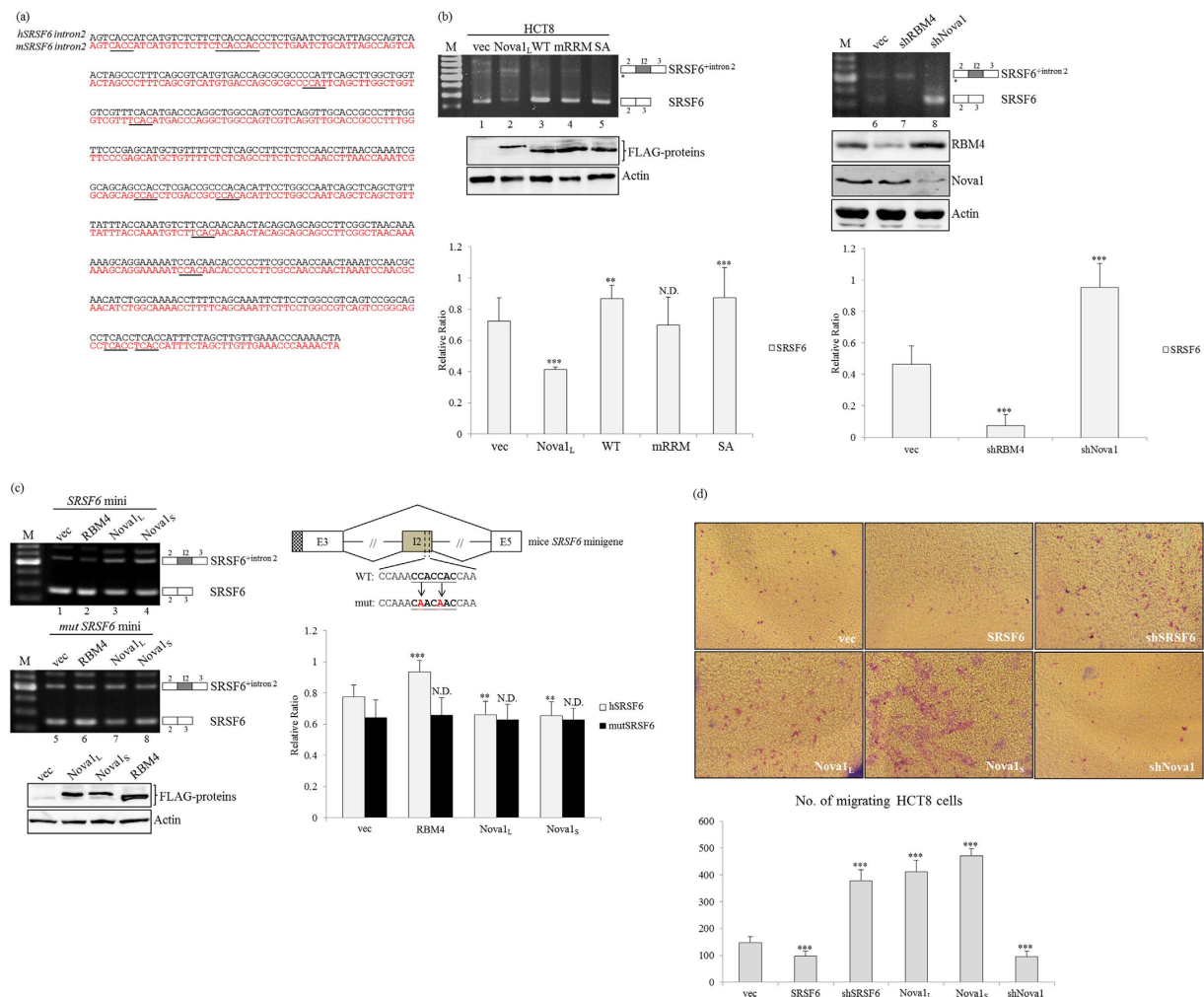


Figure 5. RBM4 and Nova1 variants exhibit opposite effects on splicing profiles of the SRSF6 gene.

(a) The diagram presents the complete identity between sequences of the 602-bp human and mouse *SRSF6* intron 2 fragments. The responsive elements of Nova1 are underlined. (b) HCT-8 cells were transfected with overexpressing or targeting vectors, followed by total RNA and lysate extraction. Splicing profiles of *SRSF6* transcripts were analyzed with specific primer sets. Immunoblot assays of indicated proteins were analyzed with specific antibodies. The bar graph shows relative levels of *SRSF6* coding transcripts in three independent experiments. (c) The *SRSF6* minigene and derived mutant were cotransfected with expressing vectors into HCT-8 cells. Spliced transcripts of the *SRSF6* minigenes were analyzed as described in Fig. 3b. The bar graph shows relative levels of *SRSF6* coding transcripts in three independent experiments. The signal densities were analyzed using TotalLab Quant Software, and the quantitative results are shown as the mean \pm STD (* p < 0.05; ** p < 0.01; *** p < 0.005). The gels shown in this figure were run under the same conditions and were not artificially manipulated. (d) The motility of HCT-8 cells transfected with *SRSF6* or Nova1 variant-expressing vectors or targeting vectors was analyzed using an *in vitro* migration assay and the counting results of three independent experiments are presented in a bar chart as the mean \pm STD.

of CRC cell lines (Fig. 5d, shNova1), which was consistent with previous reports¹³. These results suggested the hierarchical role of the RBM4a-Nova1 interplay in manipulating splicing profiles of the *SRSF6* gene and its physiological significance in CRC cells.

The RBM4-regulated splicing cascade programs splicing profiles and effects of the VEGF gene.

In addition to progressive signatures, active angiogenesis is another hallmark of various cancers, including CRC cells²⁶. Both the expression and splicing profiles of VEGF isoforms regulate the angiogenic capability of colorectal carcinomas²⁷. Proangiogenic VEGF165 and antiangiogenic VEGF165b isoforms are generated from the same gene through alternative splicing mechanisms²⁸. Previous studies documented that the presence of *SRSF6* favors the selection of VEGF exon 8b by directly interacting with the distal 3' splice site, which encodes VEGFb variants²⁹. Even though predominant expression of *VEGF165* was observed in CRC tissues³⁰, differential splicing profiles of *VEGF* were validated in distinct CRC cell lines. Except for adenocarcinoma-derived cell lines, *VEGF165b* transcripts were only generated in HCT116 carcinoma cells (Fig. 6a, lane 2), which may be related

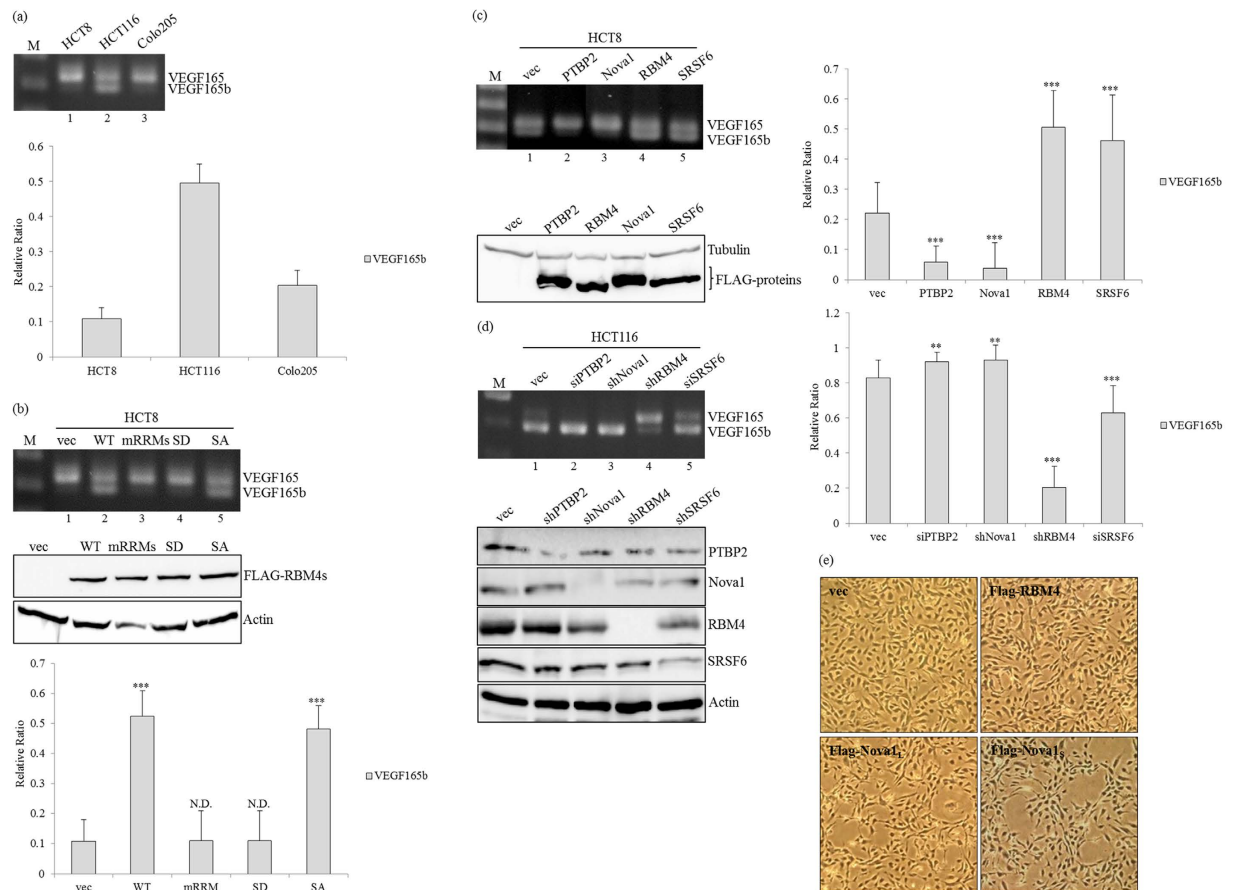


Figure 6. The RBM4-initiated splicing cascade modulates splicing profiles of vascular endothelial growth factor (VEGF) transcripts, which is relevant to the angiogenic activity of colorectal cancer (CRC) cells.

(a) Total RNAs prepared from CRC-derived cell lines were analyzed by an RT-PCR with a specific primer set against *VEGF165* transcripts. The bar graph shows relative levels of *VEGF165b* in three independent experiments. (b) HCT-8 cells were transfected with overexpressing-RBM4 and derived mutants, followed by total RNA and lysate extraction. Splicing profiles of *VEGF165* transcripts were analyzed with specific primer sets. Immunoblot assays of indicated proteins were analyzed with specific antibodies. The bar graph shows relative levels of *VEGF165b* in three independent experiments. (c) Splicing profiles of *VEGF165* transcripts in HCT-8 cells transfected with expressing vectors of splicing regulators were analyzed as described in the last panel. The bar graph shows relative levels of *VEGF165b* in three independent experiments. The overexpressing proteins and loading control were monitored with specific antibodies. (d) HCT-116 cells were transfected with targeting vectors of various splicing regulators, followed by total RNA and lysate extraction. Splicing profiles of *VEGF* transcripts and indicated proteins were examined with specific primer set and antibodies. The bar graph shows relative levels of *VEGF165b* in three independent experiments. The signal densities were analyzed using TotalLab Quant Software, and quantitative results are shown as the mean \pm STD ($*p < 0.05$; $**p < 0.01$; $***p < 0.005$). The gels shown in this figure were run under the same conditions and were not artificially manipulated. (e) Angiogenic activities of HCT-8 cells transfected with RBM4 or Nova1 variant-expressing vectors were analyzed using an *in vitro* tubular formation assay. Cell images were photographed with a microscope system, and the angiogenesis of each transfectant was quantified from four random fields of each well.

to expression profiles of splicing regulators. The overexpression of RBM4-WT and RBM4-SA proteins (Fig. 6b, lanes 2 and 5), but not the RRM-mutant (Fig. 6b, lane 3), induced the relative levels of *VEGF165b* transcripts. The cytoplasmic-enriched RBM4 mutant harboring serine-to-aspartate (S309D) substitution exerted an insubstantial effect compared to the RRM mutant on the splicing profile of the *VEGF* gene (Fig. 6b, lane 4). Moreover, overexpressing the PTBP2 and Nova1 proteins (Fig. 6c, lanes 2 and 3) exerted opposite effects as that of RBM4 and SRSF6 (Fig. 6c, lanes 4 and 5) of further enhancing relative levels of *VEGF165* transcripts. Inversely, the majority of *VEGF165* transcripts were generated with depletion of the endogenous PTBP2 and Nova1 proteins (Fig. 6d, lanes 2 and 3), whereas relative levels of *VEGF165b* transcripts were induced in RBM4- and SRSF6-knockdown HCT116 cells (Fig. 6d, lanes 4 and 5). Collectively, splicing profiles of the *VEGF* gene were meticulously manipulated by the relative expressions of RBM4, Nova1, and SRSF6. The physiological significance of the RBM4 and Nova1 proteins on angiogenic activity was evaluated by conducting *in vitro* tubule formation assays with HUVECs. As expected, overexpression of Nova1 variants induced the fusion of neighboring HUVECs to form

several capillary-like lumens by 24 h post-transfection (Fig. 6e, Flag-Nova1), whereas the cellular morphology remained unchanged in RBM4-transfected cells compared to empty-vector transfectants. Taking these results together, the RBM4-regulated splicing cascade constitutes a novel mechanism involved in the angiogenic activity of CRC cells.

Discussion

Imbalanced splicing events are reportedly relevant to carcinogenic signatures, including immortality, progression, and angiogenesis^{31–33}. In this study, we performed deep RNA-seq to reveal the CRC-related splicing cascade composed of the *Nova1*, *SRSF6*, and *VEGF* genes. The molecular mechanisms involved in programming the splicing profiles of these genes were further investigated. We next observed a marked influence of this splicing cascade on the progression and angiogenesis of CRC-derived cell lines.

We previously reported the impact of the RBM4-regulated splicing cascade on the migration, invasion, and metabolic signatures of CRC cells¹³. Reduced expression of RBM4 was documented to be related to a poor prognosis of gastric cancer patients³⁴. To determine the possible influence of RBM4 on cellular processes, a transcriptome-wide approach was conducted to identify the putative target of RBM4³⁵. To the best of our knowledge, this is the first report of the differential splicing profiles of *Nova1* transcripts revealed in CRC tissues compared to adjacent normal counterparts. The *Nova1* protein was first identified as a splicing factor predominantly expressed in neurons of the central nervous system³⁶. Development of high-throughput sequencing coupled with RNA crosslinking and immunoprecipitation demonstrated the previous prediction that *Nova1* regulated alternative splicing events in a YCAY-dependent manner^{36,37}. More than 800 neuron-specific splicing candidates of *Nova1* were identified with this methodology, which implies the potential function of *Nova1* in neurogenesis³⁸. In our previous study, speckle-enriched localization of the *Nova1_s* variant was highly relevant to its potent influence on modulating brown adipocyte-associated splicing networks, which further enhanced brown adipogenesis¹⁶. Recent studies demonstrated that an increase in *Nova1* expression functioned as a pro-oncogene involved in the proliferation, progression, and immortality of various cancer cells^{21,22,39}. The overexpression of *Nova1* variants was noted to mediate the cadherin switching from E-cadherin to N-cadherin expression, which enhanced the migration of HCT8 cells. The differential splicing profile of *Nova1* might constitute a molecular mechanism in modulating the cadherin switching in other CRC cells, such as HCT116, although more convincing result is required to demonstrate the speculation. The prominent effect of the *Nova1_s* isoform on enhancing the migration of CRC cells could be also achieved by modulating other alternative splicing event. The differential expression network of splicing regulator in distinct cancer cells might modulate the impact of *Nova1* variants on carcinogenic signature by reprogramming the splicing profiles of *Nova1*. Nevertheless, the characterization of *Nova1*-specific candidate is critical for determining its impact in the migration or other carcinogenic signature of CRC.

Autoregulation constitutes a common mechanism of controlling splicing profiles of splicing factors including *Nova1*^{19,36}. *Nova1* variants were reported to exhibit differential effects on the utilization of its own exon 4 in neurons, pancreatic β -cells, and brown adipocytes^{16,19,36}. Nevertheless, RBM4 was demonstrated to play a hierarchical role in controlling splicing profiles of the *Nova1* gene during brown adipogenesis¹⁶. However, differential influences of *Nova1* variants on regulating the human *Nova1* minigene and the derived mutant were insubstantial compared to those of the mouse *Nova1* reporter. The shortened CU element interrupted by one guanine residue may lessen the influence of RBM4, which subsequently strengthened the impact of *Nova1* variants. Nevertheless, the interplay between RBM4 and *Nova1* constituted a molecular mechanism involved in *Nova1* splicing, which was widely identified in different tissues and species.

The alternative splicing-coupled NMD pathway is reportedly a common mechanism for modulating expressions of SR proteins and hnRNP family members^{40,41}. Results of mRNA-seq identified the dominant presence of non-coding *SRSF6* transcripts containing an in-frame premature stop codon within the 268-bp intronic fragment in total RNAs extracted from CRC tissues. The complete identity between the included intron 2 within human and mouse *SRSF6* genes implies that the conserved mechanism contributes to the generation of this non-coding *SRSF6* transcript, which mediated the reduced protein level of *SRSF6* in CRC tissues compared to adjacent normal tissues. Although expression levels of *SRSF6* were documented to be correlated with the progressive activity of various cancers, relative expression levels of *SRSF6* in normal and malignant tissues and its physiological significance were previously debated^{7,42}. An elevated *SRSF6* gene copy number was previously validated using public databases⁷. In our present study, relative expressions of both *SRSF6* transcripts and coding protein were examined with paired tissue samples of anonymous CRC patients. Moreover, the differential splicing profiles of the *SRSF6* gene in distinct CRC cell lines derived from adenocarcinomas or carcinomas also suggested its potential application as a classification marker of CRC patients. Consistently, the upregulated expression of *SRSF6* diminished the progression or angiogenesis of cancer cells^{24,42}.

Our studies show that RBM4 regulates the expression level and splicing profiles of *Nova1* through multi-layer posttranscriptional controls, including mRNA stability and alternative splicing¹³. In CRC tissues and cells, reduced RBM4 expression led to relatively high levels of *Nova1*⁻⁴ transcripts which were greater than upregulated expressions of total *Nova1* mRNAs, subsequently enhancing the migratory activity of CRC cells. *Nova1*⁻⁴ variants demonstrably exerted a prominent effect on mediating the AS-coupled NMD of the *SRSF6* gene, which in turn led to a reduction in antiangiogenic *VEGF165b* variants (Fig. 7). In conclusion, our study first identified CRC-associated splicing profiles of the *Nova1* and *SRSF6* transcripts in paired CRC tissues using mRNA-seq and *in vivo* splicing assays. The molecular mechanisms involved in this RBM4-initiated splicing cascade were further investigated with CRC-derived cells. Correlations between migratory or angiogenic activities and this splicing cascade, composed of RBM4, *Nova1*, *SRSF6*, and *VEGF*, were also demonstrated. To bring the comprehensive scope into the influence of *SRSF6* or *Nova1* on CRC, the potential candidate of these splicing factor is worthy of further identification. To date, the relatively high level of *Nova1*⁻⁴ or reduced *SRSF6* expression could be

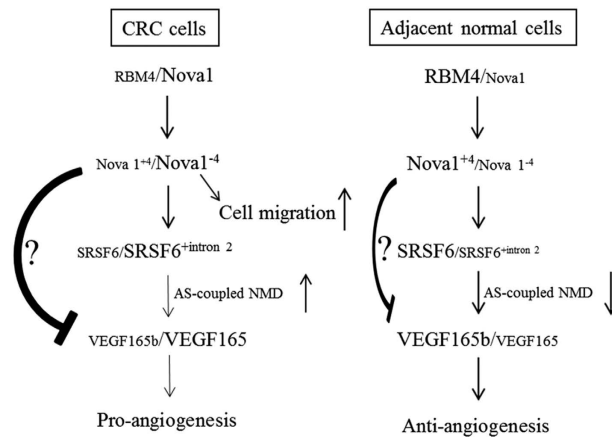


Figure 7. The RBM4-regulated splicing cascade was correlated with the progression of colorectal cancer (CRC) cells. In CRC cells, the downregulated level of RBM4 leads to a concomitant increase in the expression of total *Nova1* transcripts and relative levels of *Nova1*⁻⁴, which subsequently and coordinately strengthened the influence of *Nova1* on its splicing target, including SRSF6. Upregulated *Nova1* variants enhanced the angiogenic activity of CRC cells by indirectly inducing relative expressions of *VEGF165* transcripts, which more greatly enhanced the angiogenesis of CRC cells.

considered as the classificatory or prognostic markers of CRC. Moreover, reprogramming of the CRC-associated splicing cascade can be applied to therapeutic strategies targeting carcinogenic signatures.

Methods

Ethics statement of patient samples and cell culture. The human colorectal tumor samples ($n = 20$) with informed consent were requested as anonymous specimens from the Joint Biobank at Taipei Medical University. All experiments with clinical specimens were performed in accordance with relevant guidelines and regulations. This study was approved by the Joint Institutional Review Board of Taipei Medical University (approval no. 201409044). All of the recommendations of the *Declaration of Helsinki* for biomedical research involving human subjects were followed. HCT-8, Colo205, and HCT-116 cells were cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum (FBS), 600 mg/ml glutamine, 100 U/ml penicillin, and 100 mg/ml streptomycin (Invitrogen, Camarillo, CA, USA). Human umbilical vein epithelial cells (HUVECs) were a kind gift from Dr. Wen-Sen Lee (Taipei Medical University, Taipei, Taiwan). HUVECs were cultured in medium (M)-199 supplemented with 10% FBS, 10 U/ml heparin sodium, 15 mg/L of endothelial cell growth factor, and antibiotics.

RNA extraction and complementary (c)DNA library construction and sequencing. Total RNAs were prepared from clinical tissues using the PureLink RNA mini kit (Invitrogen) according to the manufacturer's protocol. To minimize the contribution of amplified reads that are aligned with intergenic DNA fragments, extracted RNAs were treated with DNase I (Promega, Madison, WI, USA) prior to library construction. The integrity and quantity of the RNAs for deep RNA-seq were assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Redwood, CA, USA). Poly(A) messenger (m)RNAs were enriched from 8 μ g total RNAs with a high integrity number (RIN > 8.0) using oligo (dT)-based affinity matrices, followed by library construction using the NEB Next Ultra RNA Library Prep Kit from Illumina (NEB, Ipswich, MA, USA) according to the manufacturer's instructions. The prepared libraries were sequenced on an Illumina NextSeq 500 platform, and ~150-bp paired-end reads were generated.

Read mapping and transcript annotation and quantification. Preliminary reads were cleaned by trimming the adapter sequences and removing poly-N or low-quality sequences ($Q < 20$). Filtered reads were aligned to the mouse reference genome (GRCm37) using the Tophat v2.0.9 program (available where?). Tolerance parameters were the default setting to allow mismatches of fewer than two bases. Transcriptome assemblies of aligned reads were generated using the Cufflink program. Mutant loci within the assembled transcripts were identified using SAM tools. These transcriptome assemblies generated from individual samples were merged together using the Cuffmerge utility to provide a standard for estimating transcript levels in each condition. Expression levels and the statistical significance of the merged assemblies were calculated using the Cuffdiff analysis.

Plasmid construction. The pUNova1 and pUSRSF6 minigene reporters were constructed as previously described¹⁶. The human *Nova1* and a derived mutant of *Nova1* or SRSF6 minigenes harboring substituted nucleotides within a cassette exon or intron were constructed using the QuikChange site-directed mutagenesis system (Stratagene, Amsterdam, the Netherlands). Expression vectors for the human SRSF6 gene were constructed by placing the coding sequence in-frame into the p3XFLAG-CMV14 vector (Sigma, St. Louis, MO, USA). The human SRSF6 coding region was polymerase chain reaction (PCR)-amplified with the reverse-transcription (RT)

product prepared from the total RNA of HCT8 cells as the template and then inserted into *Hind* III/*Not* I sites of the vector.

Transient transfection, and RT-PCR and quantitative RT-(q)PCR assays. CRC-derived cell lines were grown to 60–70% confluence, and the plasmid was transfected using Lipofectamine 3000 according to the manufacturer's protocol (Invitrogen). Total RNAs were extracted using a PureLink RNA mini kit at 24 h post-transfection (Invitrogen). RNAs prepared from clinical tissues or cultured cells (1 µg) were reverse-transcribed using SuperScriptase III (Invitrogen) in a 10-µl reaction and subjected to a PCR analysis with gene-specific primer sets (Supplementary Table 1). An RT-qPCR was performed with SYBR green fluorescent dye and gene-specific primer sets (Supplementary Table 2) using an ABI One Step™ PCR machine (Applied Biosystems, Foster City, CA, USA). The relative mRNA level was quantitated by the $\Delta\Delta$ -Ct method, and normalized to the level of GAPDH mRNA.

Immunoblot assay. The immunoblot analysis was performed using an enhanced chemiluminescence (ECL) system (Millipore, Billerica, MA, USA), and results were monitored using the LAS-4000 imaging system (Fujifilm, Tokyo, Japan). Primary antibodies used in this study included polyclonal anti-RBM4¹⁶, monoclonal anti-PTBP2 (Abnova, Taipei, Taiwan), monoclonal anti-SRSF6 (EMD, Millipore), monoclonal anti-Nova1 (Abnova), monoclonal anti-actin (Millipore), monoclonal anti- α -tubulin (Abcam, Cambridge, UK), monoclonal anti-E-cadherin (Abcam), monoclonal anti-N-cadherin (Abcam), and monoclonal anti-FLAG M2 (Sigma-Aldrich, St. Louis, MO, USA). Signal intensities were evaluated using TotalLab Quant Software (where available?).

In vitro migration and tubule formation assays. HCT-8 cells (5×10^4 cells/well) transfected with distinct vectors or targeting vectors were seeded in the top chamber containing a polycarbonate membrane (8-µm pore, Corning, Cambridge, MA, USA). The lower chamber contained complete culture medium which served as a chemoattractant. After 24 h, the inner membrane was scraped with a swab, and cells which had migrated to the lower side of the membrane were fixed with 4% paraformaldehyde and stained with a Giemsa solution for counting. For the *in vitro* tubular formation assay, 5×10^4 HUVECs transfected with expressing or targeting vectors were suspended in 500 µl of complete M-199 medium and then seeded on each well of 6-well plates coated with Matrigel. After incubation at 37 °C for 24 h, each well was photographed with an Olympus IX81 microscope (Olympus, Tokyo, Japan). Tubule formation was quantified from four random fields per well.

Statistical analyses. Student's *t*-tests were performed to determine the significance of the above-described experiments. $p < 0.05$ was considered statistically significant.

References

- Nilsen, T. W. & Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**, 457–463 (2010).
- Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**, 1413–1415 (2008).
- Quesnel-Vallièrès, M., Irimia, M., Cordes, S. P. & Blencowe, B. J. Essential roles for the splicing regulator nSR100/SRRM4 during nervous system development. *Genes Dev.* **29**, 746–759 (2015).
- Mavrou, A. *et al.* Serine-arginine protein kinase 1 (SRPK1) inhibition as a potential novel targeted therapeutic strategy in prostate cancer. *Oncogene* **34**, 4311–4319 (2015).
- Hamdollah Zadeh, M. A. *et al.* Alternative splicing of TIA-1 in human colon cancer regulates VEGF isoform expression, angiogenesis, tumour growth and bevacizumab resistance. *Mol. Oncol.* **9**, 167–178 (2015).
- Erkelenz, S. *et al.* Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms. *RNA* **19**, 96–102 (2013).
- Cohen-Eliav, M. *et al.* The splicing factor SRSF6 is amplified and is an oncoprotein in lung and colon cancers. *J. Pathol.* **229**, 630–639 (2014).
- Gonçalves, V. *et al.* Phosphorylation of SRSF1 by SRPK1 regulates alternative splicing of tumor-related Rac1b in colorectal cells. *RNA* **20**, 474–482 (2014).
- Pita-Fernández, S. *et al.* Effect of diagnostic delay on survival in patients with colorectal cancer: a retrospective cohort study. *BMC Cancer* **16**, 664–674 (2016).
- Sveen, A., Kilpinen, S., Ruusulehto, A., Lothe, R. A. & Skotheim, R. I. Aberrant RNA splicing in cancer; expression changes and driver mutations of splicing factor genes. *Oncogene* **35**, 2413–2427 (2016).
- Zhou, X. *et al.* BCLAF1 and its splicing regulator SRSF10 regulate the tumorigenic potential of colon cancer cells. *Nat. Commun.* **5**, 4581–4591 (2014).
- Taniguchi, K. *et al.* PTBP1-associated microRNA-1 and -133b suppress the Warburg effect in colorectal tumors. *Oncotarget* **7**, 18940–18952 (2016).
- Liang, Y. C., Lin, W. C., Lin, Y. J. & Lin, J. C. The impact of RNA binding motif protein 4-regulated splicing cascade on the progression and metabolism of colorectal cancer cells. *Oncotarget* **6**, 38046–38060 (2015).
- Bai, Y. *et al.* Read-Split-Run: an improved bioinformatics pipeline for identification of genome-wide non-canonical spliced regions using RNA-Seq data. *BMC Genomics Suppl* **7**, 503–513 (2016).
- Panwar, B. *et al.* Genome-Wide Functional Annotation of Human Protein-Coding Splice Variants Using Multiple Instance Learning. *J. Proteome Res.* **15**, 1747–1753 (2016).
- Lin, J. C., Chi, Y. L., Peng, H. Y. & Lu, Y. H. RBM4-Nova1-SRSF6 splicing cascade modulates the development of brown adipocytes. *Biochim. Biophys. Acta.* **1859**, 1368–1379 (2016).
- Traunmüller, L., Bornmann, C. & Scheiffele, P. Alternative splicing coupled nonsense-mediated decay generates neuronal cell type-specific expression of SLM proteins. *J. Neurosci.* **34**, 16755–16761 (2014).
- Lin, J. C., Lu, Y. H., Liu, Y. R. & Lin, Y. J. RBM4a-regulated splicing cascade modulates the differentiation and metabolic activities of brown adipocytes. *Sci. Rep.* **6**, 20665 (2016).
- Villate, O. *et al.* Nova1 is a master regulator of alternative splicing in pancreatic beta cells. *Nucleic Acids Res.* **42**, 11818–11830 (2014).
- Wang, Y. *et al.* The splicing factor RBM4 controls apoptosis, proliferation, and migration to suppress tumor progression. *Cancer Cell* **26**, 374–389 (2014).
- Yoon, S. O. *et al.* NOVA1 inhibition by miR-146b-5p in the remnant tissue microenvironment defines occult residual disease after gastric cancer removal. *Oncotarget* **7**, 2475–2495 (2016).

22. Zhi, F. *et al.* MiR-181b-5p downregulates NOVA1 to suppress proliferation, migration and invasion and promote apoptosis in astrocytoma. *PLoS One* **9**, e109124 (2014).
23. Zhao, J. *et al.* Growth differentiation factor 8 induces SKOV3 ovarian cancer cell migration and E-cadherin down-regulation. *Cell Signal* **28**, 1615–1622 (2016).
24. Loh, T. J. *et al.* SR proteins regulate V6 exon splicing of CD44 pre-mRNA. *BMB Rep.* **pii**, 3616 (2016).
25. Lal, S. *et al.* Estrogen alters the splicing of type 1 corticotropin-releasing hormone receptor in breast cancer cells. *Sci. Signal* **6**, ra53 (2013).
26. Ding, C., Li, L., Yang, T., Fan, X. & Wu, G. Combined application of anti-VEGF and anti-EGFR attenuates the growth and angiogenesis of colorectal cancer mainly through suppressing AKT and ERK signaling in mice model. *BMC Cancer* **16**, 791 (2016).
27. Díaz, R. *et al.* p73 Isoforms affect VEGF, VEGF165b and PEDF expression in human colorectal tumors: VEGF165b downregulation as a marker of poor prognosis. *Int. J. Cancer* **123**, 1060–1067 (2008).
28. Renkel, E. S. *et al.* Recombinant human VEGF165b protein is an effective anti-cancer agent in mice. *Eur. J. Cancer* **44**, 1883–1894 (2008).
29. Nowak, D. G. *et al.* Expression of pro- and anti-angiogenic isoforms of VEGF is differentially regulated by splicing and growth factors. *J. Cell Sci.* **121**, 3487–3495 (2008).
30. Kamiya, T. *et al.* The preserved expression of neuropilin (NRP) 1 contributes to a better prognosis in colon cancer. *Oncol. Rep.* **15**, 369–373 (2006).
31. Anczuków, O. *et al.* SRSF1-Regulated Alternative Splicing in Breast Cancer. *Mol. Cell* **60**, 105–117 (2015).
32. Zhang, P. *et al.* CD82 suppresses CD44 alternative splicing-dependent melanoma metastasis by mediating U2AF2 ubiquitination and degradation. *Oncogene* **35**, 5056–5069 (2016).
33. Bullock, N. *et al.* Serine-arginine protein kinase 1 (SRPK1), a determinant of angiogenesis, is upregulated in prostate cancer and correlates with disease stage and invasion. *J. Clin. Pathol.* **69**, 171–175 (2016).
34. Yong, H. *et al.* Prognostic value of decreased expression of RBM4 in human gastric cancer. *Sci. Rep.* **6**, 28222 (2016).
35. Markus, M. A., Yang, Y. H. & Morris, B. J. Transcriptome-wide targets of alternative splicing by RBM4 and possible role in cancer. *Genomics* **107**, 138–144 (2016).
36. Jelen, N., Ule, J., Zivin, M. & Darnell, R. B. Evolution of Nova-dependent splicing regulation in the brain. *PLoS Genet.* **3**, 1838–1847 (2007).
37. Brooks, A. N. *et al.* Conservation of an RNA regulatory map between Drosophila and mammals. *Genome Res.* **21**, 193–202 (2011).
38. Eom, T. *et al.* NOVA-dependent regulation of cryptic NMD exons controls synaptic protein levels after seizure. *Elife* **2**, e00178 (2013).
39. Zhang, Y. A. *et al.* High expression of neuro-oncological ventral antigen 1 correlates with poor prognosis in hepatocellular carcinoma. *PLoS One* **9**, e90955 (2014).
40. Valacca, C. *et al.* Sam68 regulates EMT through alternative splicing-activated nonsense-mediated mRNA decay of the SF2/ASF proto-oncogene. *J. Cell Biol.* **191**, 87–99 (2010).
41. Lin, J. C. & Tarn, W. Y. RBM4 down-regulates PTB and antagonizes its activity in muscle cell-specific alternative splicing. *J. Cell Biol.* **193**, 509–520 (2011).
42. Filippov, V., Filippova, M. & Duerksen-Hughes, P. J. The early response to DNA damage can lead to activation of alternative splicing activity resulting in CD44 splice pattern changes. *Cancer Res.* **67**, 7621–7630 (2007).

Acknowledgements

We greatly appreciate Prof. Wen-Sen Lee (Taipei Medical University, Taiwan) for providing HUVECs. This work was supported by a grant (MOST 105-2311-B-038-003) from the Ministry of Science and Technology of Taiwan.

Author Contributions

Jung-Chun, Lin designed and performed the experiments, analyzed the results and wrote the manuscript. Yuan-Chii Lee, Yang C. Fann and Kory R. Johnson conducted the RNA-seq analyses. Yu-Chih Liang and Ying-Ju Lin analyzed the experimental results.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing Interests: The authors declare no competing financial interests.

How to cite this article: Lin, J.-C. *et al.* The impact of the RBM4-initiated splicing cascade on modulating the carcinogenic signature of colorectal cancer cells. *Sci. Rep.* **7**, 44204; doi: 10.1038/srep44204 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017