



Modulator-Dependent RBPs Changes Alternative Splicing Outcomes in Kidney Cancer

Yang Wang^{1,2}, Steven X. Chen¹, Xi Rao¹ and Yunlong Liu^{1*}

¹ Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, United States,

² State Key Laboratory of Biocatalysts and Enzyme Engineering, School of Life Sciences, Hubei University, Wuhan, China

OPEN ACCESS

Edited by:

Yongsheng Kevin Li,
Harbin Medical University, China

Reviewed by:

Xiaowen Chen,
Harbin Medical University, China
Dario Balestra,
University of Ferrara, Italy
Weixin Xie,
The Ohio State University,
United States

*Correspondence:

Yunlong Liu
yunliu@iu.edu

Specialty section:

This article was submitted to
RNA,
a section of the journal
Frontiers in Genetics

Received: 20 January 2020

Accepted: 05 March 2020

Published: 26 March 2020

Citation:

Wang Y, Chen SX, Rao X and
Liu Y (2020) Modulator-Dependent
RBPs Changes Alternative Splicing
Outcomes in Kidney Cancer.
Front. Genet. 11:265.
doi: 10.3389/fgene.2020.00265

Alternative splicing alterations can contribute to human disease. The ability of an RNA-binding protein to regulate alternative splicing outcomes can be modulated by a variety of genetic and epigenetic mechanisms. In this study, we use a computational framework to investigate the roles of certain genes, termed modulators, on changing RBPs' effect on splicing regulation. A total of 1,040,254 modulator-mediated RBP-splicing interactions were identified, including 137 RBPs, 4,309 splicing events and 2,905 modulator candidates from TCGA-KIRC RNA sequencing data. Modulators function categories were defined according to the correlation changes between RBPs expression and their targets splicing outcomes. QKI, as one of the RBPs influencing the most splicing events, attracted our attention in this study: 2,014 changing triplets were identified, including 1,101 modulators and 187 splicing events. Pathway enrichment analysis showed that QKI splicing targets were enriched in tight junction pathway, endocytosis and MAPK signaling pathways, all of which are highly associated with cancer development and progression. This is the first instance of a comprehensive study on how alternative splicing outcomes changes are associated with different expression level of certain proteins, even though they were regulated by the same RBP. Our work may provide a novel view on understanding alternative splicing mechanisms in kidney cancer.

Keywords: alternative splicing, RNA-binding protein, modulation, cancer, dysregulation

INTRODUCTION

Renal cell carcinoma (RCC) is a common malignancy, representing 4.2% of all new cancer cases, with about 73,820 new cases and 14,770 deaths estimated for 2019 in the United States (Siegel et al., 2019). RCC is radiotherapy- and chemotherapy-resistant, and surgery remains first-line therapy (Hsieh et al., 2017; Yin et al., 2019). Despite early surgical treatment, 30% of patients with a localized tumor eventually develop metastases, and 2 years survival rate of patients with metastatic kidney renal clear cell carcinoma (KIRC) is less than 20% (Mickisch, 2002; Janzen et al., 2003). Therefore, identification of underlying molecular mechanisms and metastatic potential of KIRC are essential for improvements in early diagnosis and treatment.

Dysregulation of alternative splicing (AS) is widely considered a new hallmark of cancer and its products are acknowledged as potentially useful biomarkers (Ladomery, 2013). Recent estimates indicate that nine out of every 10 human genes undergo AS in a cell type- or condition-specific manner to create distinct RNA transcripts from the same pre-mRNA molecule (Wang et al., 2008). The key role of AS is further confirmed by the linkage of splicing regulation to numerous human diseases, including neurological disorders and many types of cancer (Scotti and Swanson, 2016).

Regulation of AS is a complicated process in which numerous interacting components are at work, including *cis*-acting elements and *trans*-acting factors, complicated by the functional coupling between transcription and splicing (Wang et al., 2015). Corruption of the process may lead to disruption of normal cellular function and eventually disease. Thus, understanding the regulatory patterns that control AS events has the potential not only to give valuable molecular insights but also to provide solutions for various diseases.

AS events are largely controlled by RNA-binding proteins (RBPs) that recognize specific regulatory sequences embedded in the pre-mRNA transcripts (Gerstberger et al., 2014). However, splicing complexes are intricate molecular machines that process tens to hundreds of RNA target genes. At any given time, depending on the context and cellular stimuli, an RBP will affect only a subset of its RNA target genes. This specificity is often provided by a certain factors we named as “modulators,” such as signaling proteins, microRNAs, lncRNAs that control RBPs activity through several different mechanisms, including: expression level (Payne et al., 2018), protein stability and turnover (Garcia-Maurino et al., 2017), nuclear/cytoplasmic localization (Di Liegro et al., 2014), altered protein interactions (Jankowsky and Harris, 2015), and co-transcriptional regulation (Shukla and Oberdoerffer, 2012). Modulators help a cell combine different external signals and make complex downstream decisions. Elucidating their function is necessary for understanding and controlling cell’s response to external stimuli at transcriptional level.

With the increased availability of large data sets derived from high-throughput experiments and computer algorithms, investigating complex transcriptional dysregulation between RBPs and AS events in complex diseases is now possible. Recently, the ENCODE project published eCLIP data sets for 150 RBPs across K562 and HepG2 cell types (Van Nostrand et al., 2016; Yee et al., 2019). Technological advances have made it possible to define the comprehensive target networks of individual RBPs with high accuracy by integrating global splicing profiles upon depletion of each RBP and genome-wide maps of *in vivo*, direct protein-RNA interactions (Zhang et al., 2010; Weyn-Vanhentenryck et al., 2014).

In this study, we established a computational method for dissecting the relationship among RBPs, alternatives splicing events and a kind of proteins that may influenced the splicing regulation effect of RBPs. Our method is unique in its ability to discover how alternative splicing outcomes is changing when modulator expression is different, even though they were regulated by the same RBP. It is the first time in which a triplet describing the relationships among modulators, RBPs and the outcomes of their alternative splicing targets is reported. The triplet contains three objects: a specific RBP, a splicing target regulated by RBP, and a modulator candidate that may change splicing regulation of the RBP. This method was applied to RCC using TCGA-KIRC dataset to identify modulator-dependent RBPs and their targets splicing outcomes in kidney cancer. QKI, as one of the key RBPs in this study, has the greatest number of influenced splicing events. Functional enrichment analysis showed that the inferred QKI modulators were highly

associated with regulation progress of some hallmark cancer genes, including ARMH4, LINC01268, PDP2, LAPTM4B, and CD7. The results showed that different expression of modulators was associated with the changing roles of RBPs on regulating their targets alternative splicing outcomes. We expect that such integrated analysis could reveal the roles of RBPs and provide novel insights into understanding alternative splicing mechanisms in kidney cancer.

MATERIALS AND METHODS

Identify Alternative Splicing Events and Gene Expression

Paired-end RNA sequencing data from 480 RCC patients was downloaded from The Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma (TCGA-KIRC). The percentage of inclusion (PSI) of spliced events were derived using Mixture of Isoforms (MISO) (Lee et al., 2013). A PSI value was computed for every identified event in each sample, and the original AS events were further processed to generate high-confidence events by retaining events with a PSI value greater than 0.1 in at least 100 samples from 480 (~21% samples in total). Then, events that occurred in both the curated and TCGA datasets were retained to form the final set of AS events. In this study, we only focused on skipped exon (SE) alternative splicing events. We defined an altered skipped exon as any exon of a transcript excluding the first and the last exons. Finally, we only kept the events that at least 100 patients’ have PSI value and coefficient of variation (CV) of PSI was larger than 0.1. Gene differential expression analysis was performed using edgeR (Robinson et al., 2010), and CPM was used to estimated gene expression.

Identify RBPs Targets Using eCLIP Data

We used crosslinking immunoprecipitation (CLIP) data for 150 RBPs profiled in eCLIP peaks (Van Nostrand et al., 2016) downloaded from ENCODE in bed format (Consortium Encode Project, 2012). The peaks in two immortalized human cell types, K562 and HepG2, were filtered by peak enrichment larger than 8 ($\log_2FC \geq 3$) and $p < 10^{-5}$ as recommended (Van Nostrand et al., 2016). Since the agreement between peaks in two replicates was moderate (the median Jaccard distance 25 and 28% in K562 and HepG2, respectively), we took the union of peaks between the two replicates in both cell lines and then pooled the resulting peaks. We defined an RBP-binding splicing event as the region upstream 300 base pairs of the exon to downstream 300 base pairs of the exon.

Genomes and Transcript Annotations

February 2009 assembly of the human genome (hg19, GRCh37) was download from Genome Reference Consortium (Church et al., 2011). The respective transcript annotation v19 was download from GENCODE website (Harrow et al., 2012). Transcript annotations were parsed to extract positions of introns and exons.

Gene Function and Categories Analysis

Functional enrichment analysis was carried out via the hypergeometric test using the clusterProfiler R package. We used Human Gene Ontology annotation provided by Gene Ontology (GO) Consortium (Ashburner et al., 2000; The Gene Ontology Consortium, 2017). GO terms enrichment with adjusted $p < 0.01$ and KEGG pathway enrichment with adjusted $p < 0.05$. The gene types we discussed in this study including immune related genes, which were download from IMMPORT database¹ and TRRUST v2 database.² The lncRNA gene type annotation was based on biomaRt software suite in R.

Construction Modulator-RBP-Splicing Triplets in KIRC

The probabilistic model is similar to Li et al. (2017) as follows:

$$Y_{target} = \beta_0 + \beta_1 X_{rbp} + \beta_2 X_m + \beta_3 X_{rbp} X_m + \varepsilon \quad (1)$$

where, the X_{rbp} , X_m , and Y_{target} are the gene expression of RBP and its modulator, and the splicing outcomes of the affected target gene, respectively. X_{rbp} and X_m represent the effect of RBP and modulator, respectively, on target by themselves alone (main effects), while β_3 represents the effect of their interaction. If an RBP and modulator interaction influences target splicing outcomes, we expect β_3 to be non-zero.

We divide rank-ordered expression values of a gene by tertiles and further discretize the triplets using:

$$x' = \begin{cases} 1 & \text{if } x \text{ is in upper tertile} \\ \text{NULL} & \text{if } x \text{ is in middle tertile} \\ 0 & \text{if } x \text{ is in lower tertile} \end{cases} \quad (2)$$

Values are ranked and transformed by tertials as follows, and coefficients are estimated from the differences in observed proportions of frequencies:

- (1) Splice events are ranked according to PSI.
- (2) RBP is ranked by its expression.
- (3) Each modulator is ranked by their expression.

After discretization, we only consider the eight bins, where none of the genes has “NULL” value, covering ~33% of the samples. This simple strategy has been shown to maximize entropy among groups and the selection of significant triplets' method can be found in Babur et al. (2010).

Statistical Analysis and Software

The data were analyzed and visualized using R statistics software version 3.4.1 and ggplot2 package. Correlations were assessed using Pearson correlation test. Survival curves were generated by the Kaplan Meier method using the median H-score as the cutoff, and differences were analyzed with the log-rank test.

¹<https://www.immport.org>

²<https://www.grnpedia.org>

RESULTS

Category of Modulator Action

We developed a framework to infer the modulators of RBPs whose expression strongly correlates with changes of a RBP's effect on regulating targets splicing outcomes. Here, the transcriptional activity of a RBP was evaluated by the Pearson correlation between the expression level of RBPs and its target splicing outcomes. A schematic diagram of work-flow is provided in **Figure 1**.

The proposed method takes five inputs: gene expression profiles, an RBP of interest, a list of modulator candidates, splicing profiles, and RBPs' binding information. Candidate modulators may include all genes satisfying the criteria. In addition, the expression of the modulator candidates and RBPs were required to be statistically independent. Each possible triplet was then independently tested using the PCCs (Pearson correlation coefficients) estimator, and by comparing $\Delta PCCs$ we defined the subtype of modulation categories. False positives were controlled using appropriate statistical thresholds. Three possible modes of modulator action were identified, depending on whether RBP-splicing correlation increased or decreased as a function of the modulator abundance.

Category of Modulator Action

For each triplet, three possible modes of modulator action were identified depending on whether RBP-splicing correlation increased or decreased as a function of the modulator abundance. The three models are “attenuates splicing,” “enhances splicing,” and “inverts splicing.” Among them, attenuates/enhances splicing modes including two sub-types: attenuates/enhances exon exclusion and attenuates/enhances exon inclusion; inverts splicing mode means that the mode of modulator switches from exon inclusion to exon exclusion or from exon exclusion to exon inclusion. These cases and details interpretations are listed in **Table 1**.

Identify Modulators of QKI in Kidney Cancer

We applied the proposed method for identifying modulators for 150 RBPs. 14,707 exons were selected from 42,485 annotated skipped exons that are derived using the gene structures of ENSEMBL database. We identified 1,040,254 significant modulator-mediated triplets from 40,623,520 potential modulator-RBP-splicing interactions at $FDR \leq 0.01$ using TCGA-KIRC data. The potential interactions consisted of 137 RBPs, 4,309 splicing events, and 2,905 modulator candidates. Among these RBPs, 13 RBPs were filtered out as the PSI distribution among 480 patients did not meet our criteria (the PSI coefficient of variation should be larger than 0.1). RNA-binding protein Quaking (QKI) had the greatest number (68.9%, 199 out of 289) of modulated spliced exons.

We identified 2,014 Modulator-QKI-Splicing triplets. The triplets include 1,101 modulators and 187 splicing events corresponding to 130 genes. According to the correlation between QKI expression and its target splicing

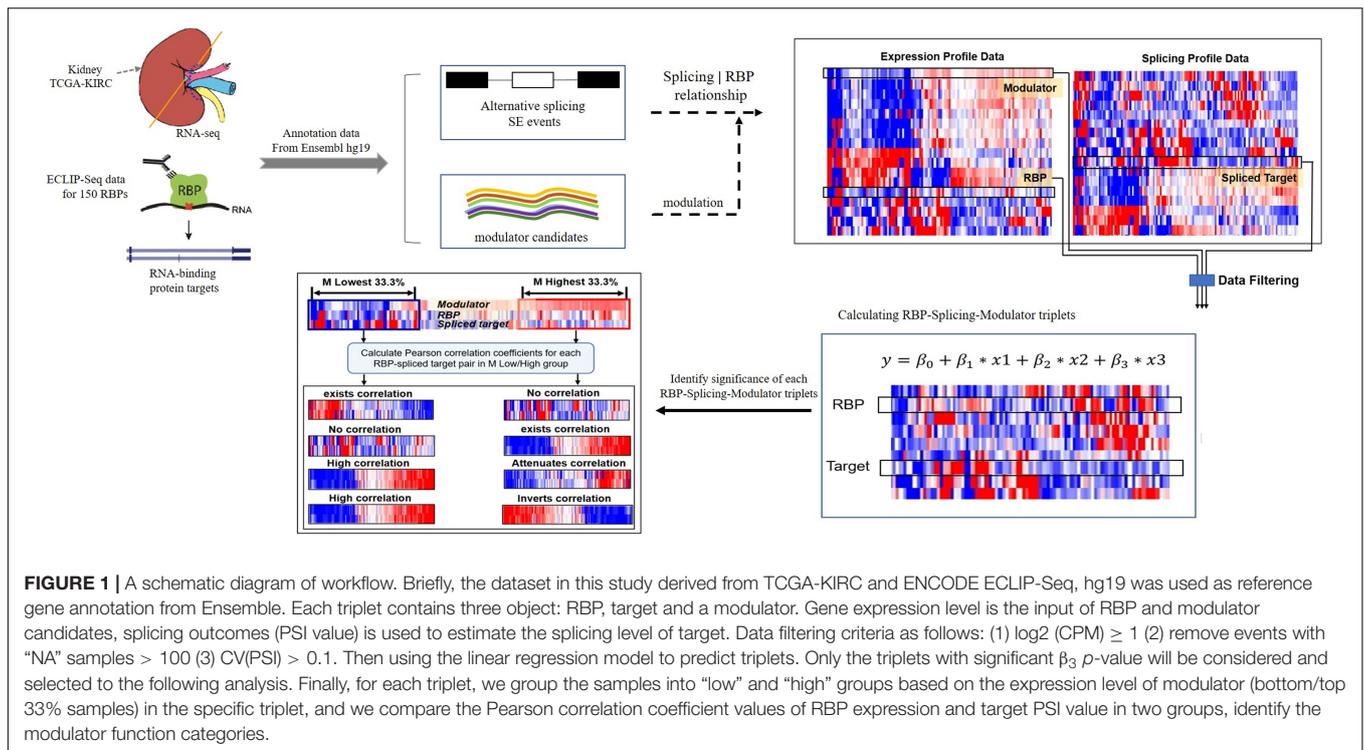


FIGURE 1 | A schematic diagram of workflow. Briefly, the dataset in this study derived from TCGA-KIRC and ENCODE ECLIP-Seq, hg19 was used as reference gene annotation from Ensemble. Each triplet contains three object: RBP, target and a modulator. Gene expression level is the input of RBP and modulator candidates, splicing outcomes (PSI value) is used to estimate the splicing level of target. Data filtering criteria as follows: (1) $\log_2(\text{CPM}) \geq 1$ (2) remove events with “NA” samples > 100 (3) $\text{CV}(\text{PSI}) > 0.1$. Then using the linear regression model to predict triplets. Only the triplets with significant β_3 *p*-value will be considered and selected to the following analysis. Finally, for each triplet, we group the samples into “low” and “high” groups based on the expression level of modulator (bottom/top 33% samples) in the specific triplet, and we compare the Pearson correlation coefficient values of RBP expression and target PSI value in two groups, identify the modulator function categories.

TABLE 1 | Categories of modulator mediated RBP regulations on splicing targets.

Modulation category	PCC _{low}	PCC _{high}	ΔPCCs	Subtype mode
Attenuates splicing	– (EE)	– (EE)	$ PCC_{low} > PCC_{high} $ or $p\text{-value}_{high} > 0.05$	Attenuates exon exclusion (AEE)
Enhances splicing	– (EE)	– (EE)	$ PCC_{low} < PCC_{high} $ or $p\text{-value}_{low} > 0.05$	Enhances Exon exclusion (EEE)
Inverts splicing	+ (EI)	– (EE)		Exon inclusion to exclusion (ExonIE)
Inverts splicing	– (EE)	+ (EI)		Exon exclusion to inclusion (ExonEI)
Enhances splicing	+ (EI)	++(EI)	$ PCC_{low} < PCC_{high} $ or $p\text{-value}_{Low} > 0.05$	Enhances EI (EEI)
Attenuates splicing	+ +(EI)	+(EI)	$ PCC_{low} > PCC_{high} $ or $p\text{-value}_{high} > 0.05$	Attenuates EI (AEI)

“+” and “–” signs in the columns indicate positive and negative values of Pearson correlation coefficient. “EE” represents exon exclusion, and “EI” means exon inclusion. The modulation categories of “attenuates splicing,” “enhances splicing” or “inverts splicing” only refer to the roles of RBP on specific alternative spliced exon.

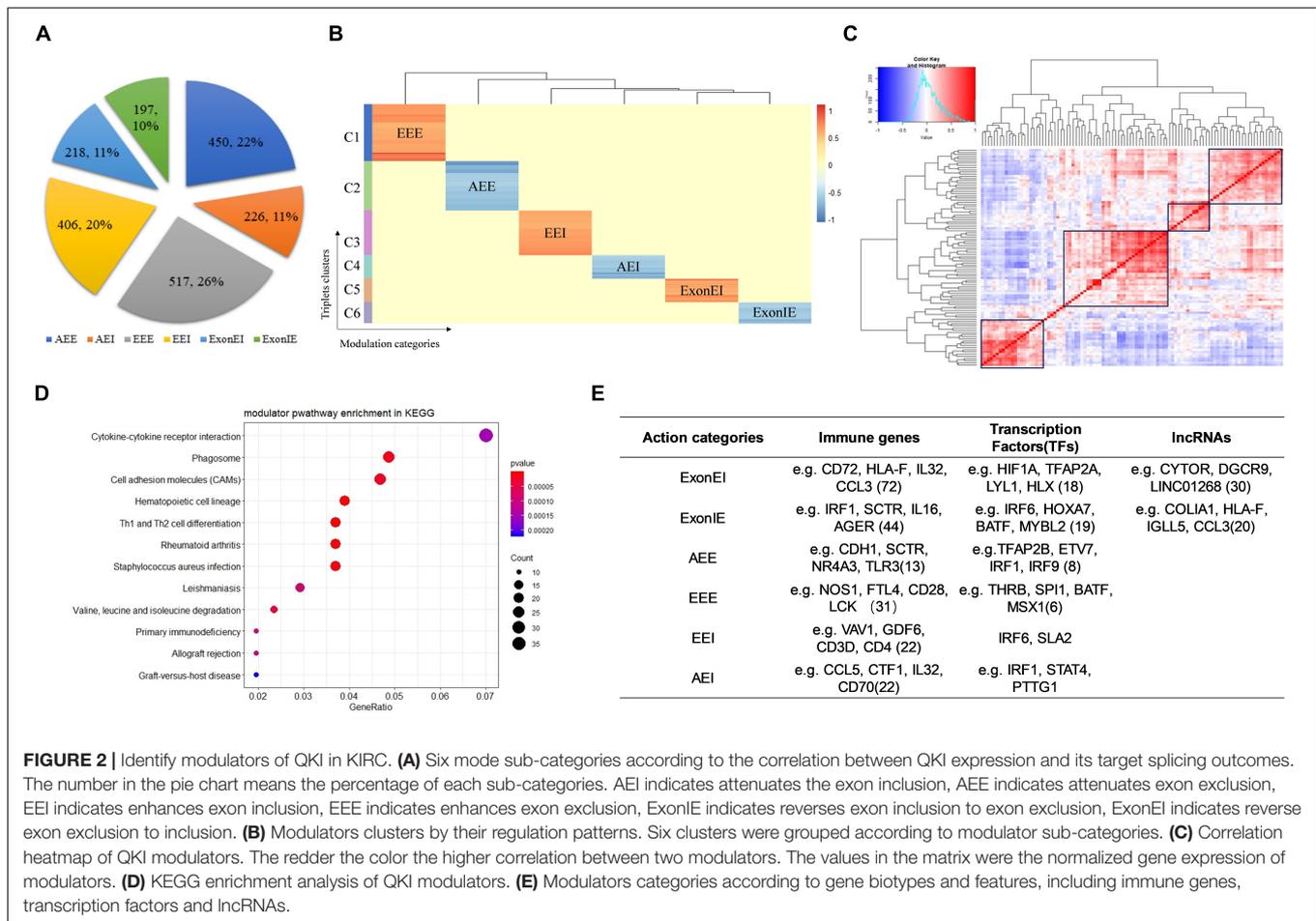
outcomes, six modulator sub-categories were identified, including 450 triplets in “Attenuates_Exon_Exclusion,” 226 triplets in “Attenuates_Exon_Inclusion,” 517 triplets in “Enhances_Exon_Exclusion,” 406 triplets in “Enhances_Exon_Inclusion,” 218 triplets in “Exon_Exclusion_to_Inclusion,” and 197 triplets in “Exon_Inclusion_to_Exclusion” (Figure 2A).

Furthermore, we observed that most modulators affected multiple splicing targets were multimodal, and the same modulator may play opposite roles on different QKI targets. For example, ARMH4 inverts the splicing activity of QKI on its target CLTC: the inclusion of CLTC’s spliced exon was correlated with increasing expression of QKI when ARMH4 is lowly expressed, while such association was inverted when ARMH4 is highly expressed. However, ARMH4 played enhanced exon inclusion role on QKI-STIM1 pair when its expression level changed from low to high. In another case, while the modulator KRT17 influenced 11 splicing targets of QKI, the role of KRT17 on these splicing outcomes changed among

EEE, ExonIE, and ExonEI. The observation indicated that the distinct modulation patterns were triplets dependent rather than depending on specific RBPs or target genes. Our findings support this complexity in modulators typically had many target-specific effects. These findings suggested that more complex models are needed to better elucidate that how splicing is regulated.

Hence, we clustered modulators by their regulation patterns, yielding distinct groups of modulators that mediated splicing dysregulation in specific patterns (Figure 2B). For instance, the modulators in cluster 1–2 tend to reverse the QKI activity on regulating splicing targets’ outcomes, whereas those modulators in cluster 4 and cluster 6 tend to enhance QKI splicing activity. In conclusion, these modulators may work as antagonistic or coactivators to mediate QKI splicing activity.

Among these modulator-mediated triplets, we noticed that many modulators regulate the same QKI splicing targets, this may be because some of the modulators co-express or play similar functions in related pathways. As the result showed in Figure 2C, modulators were grouped into several clusters



according to their expression's correlation. This may be a potential reason why the spliced outcome of same target could be influenced by many modulators. In addition, the pathway enrichment results shown that these modulators were highly enriched in categories which were known to be associated with cancer development and progression, including cytokine-cytokine receptor interaction, Th1 and Th2 cell differentiation, and cell adhesion molecules (Figure 2D).

Furthermore, we classified the modulators of QKI according to gene biotypes and features, including immune related genes, transcription factors and lncRNAs (Figure 2E). The types of categories provide a framework for understanding many types of dysregulation on splicing.

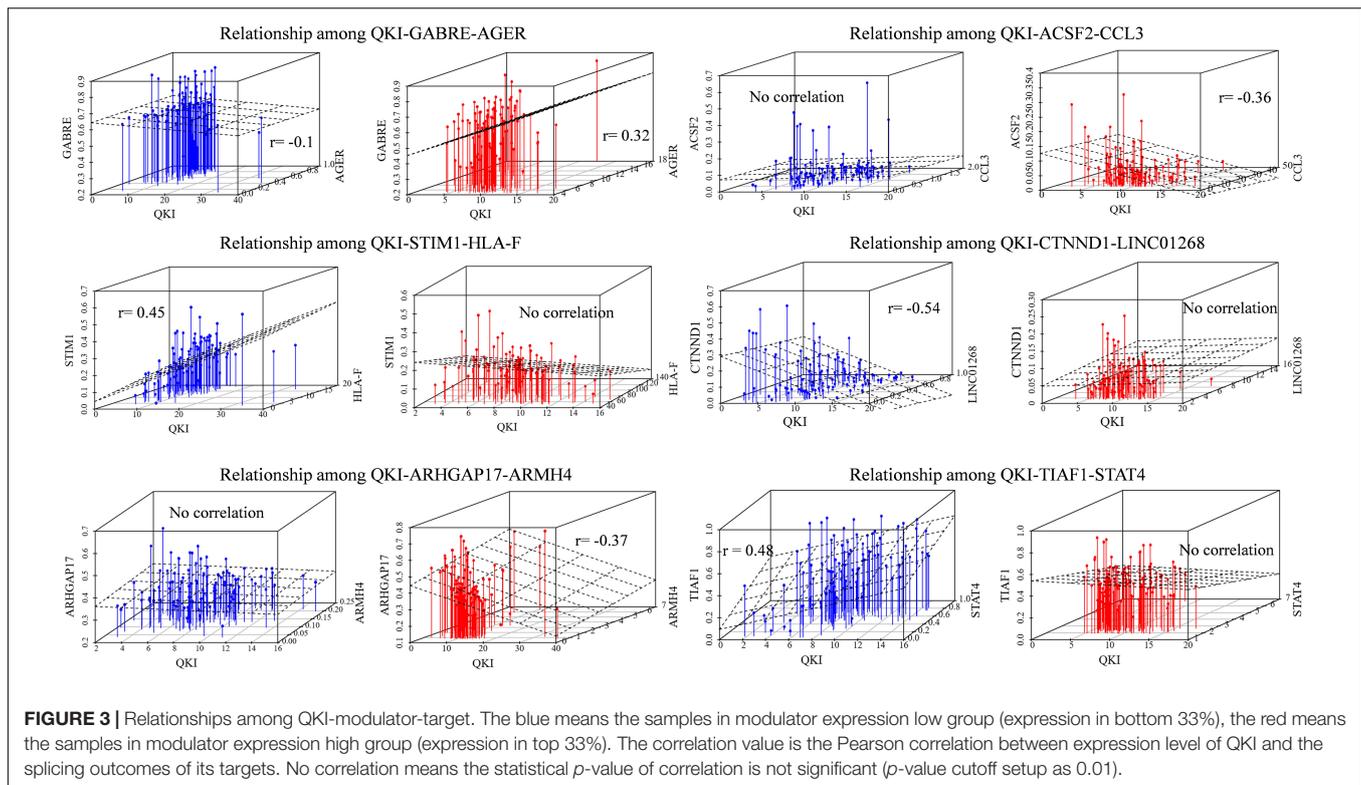
Functional Analysis of QKI Modulators

To confirm the QKI-splicing-modulator triplet signatures as independent predictors, we selected six inferred modulator-influenced triplets to compare the association among QKI-splicing-modulators. The modulators we focused on were obtained from the analysis result 3.3, including immune genes (CCL3, HLA-F, AGER), transcription factors (ARMH4, STAT4), and lncRNAs (LINC01268).

As an example, immune gene AGER as a modulator of QKI, who shown differentially expressed level between cohort and

normal samples in KIRC, played inverts exon exclusion role on regulating the splicing outcomes of GABRE. Comparing the two patterns in different groups, when the expression of modulator AGER is low, the PCCs between QKI expression and GABRE splicing level (PSI) is -0.1 , while such correlation inverts to 0.32 in another group whose AGER expression high. Similar pattern we found that the correlation between QKI and its splicing target STIM1 was lost from 0.45 to no significant correlation when the modulator immune gene HLA-F expression differentially in two groups. Meanwhile, LINC01268 as modulator played attenuated effect on regulating QKI splicing activity. The correlation between QKI and target CTNND1 was -0.54 in LINC01268 expressed low group, while such correlation gone when LINC01268 expression becomes high (Figure 3).

To investigate the association between dysregulated target splicing outcomes and kidney cancer, we performed survival analysis and compared the expression level of these modulators in kidney tumor samples and normal samples based on TCGA-KIRC dataset. Results shown that most of the modulators were differentially expressed between tumor and normal samples and overall survival associated. Although KRT17 as one of the modulators we inferred did not show too much differentially expressed, the clinical information obtained from



TCGA indicated that gene expression in kidney cancer was significantly associated with overall survival outcomes. The gene expression levels and survival analysis of top 10 modulators who has the most influenced targets of QKI were compared (Figures 4A,B).

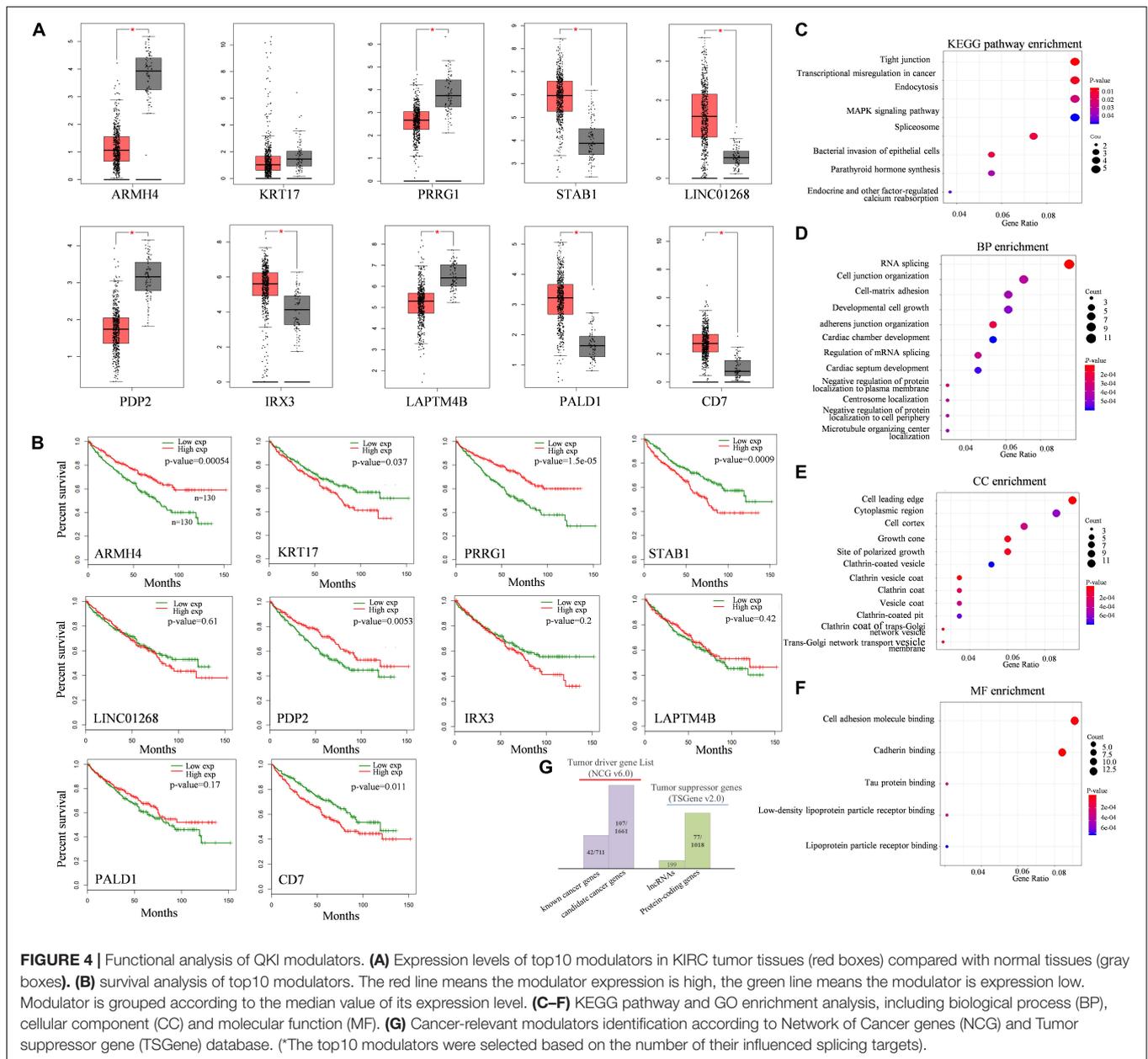
KEGG enrichment analysis revealed that these target genes were enriched in categories known to be related to cancer development and progression (Figure 4C), such as tight junction pathway, transcriptional mis-regulation in cancer, endocytosis, and MAPK signaling pathway. The top enriched GO terms of these influenced target genes were associated with transcriptional regulation progress, including RNA splicing, cell growth and protein binding (Figures 4D-F). The results were reasonable as QKI regulated its target mainly on splicing level, once the expression of QKI was perturbed by the modulators, the roles of QKI on its targets, including binding, splicing, cellular development and transcriptional regulation would be influenced consequently.

In addition, cancer-relevant modulators were identified though tumor associated gene list from the Network of Cancer Genes (NCG, v6.0) database (Repana et al., 2019) and Tumor suppressor gene database (TSGene v2.0) (Zhao et al., 2016), separately (Figure 4G). The 2,372 tumor diver genes obtained from NCG including 711 known cancer genes and 1,661 candidate cancer genes. Among them, 149 genes overlapped with tumor diver genes, almost reaching 13% (149/1,179) of total numbers of modulators we inferred in this study. Meanwhile, approximately 7% (77/1,179) modulators were tumor suppressor genes, and the gene type was protein coding gene.

Analysis the Splicing Outcomes of CTNND1 Influenced by Modulators in Kidney Cancer

In this study, we found that the spliced outcome of the 20th exon of CTNND1 has the most inferred modulators, including 30 lncRNAs and 80 protein coding RNAs. The corresponding AS event is “chr11:57582866: 57582972: + @ chr11: 57583387: 57583473: + @ chr11:57583769: 57586652:+.” Previously study reported that CTNND1 encodes a member of the Armadillo protein family, which function in adhesion between cells and signal transduction (Zhu et al., 2012), multiple CTNND1 isoforms are expressed in cells via alternative splicing, only full-length CTNND1 promotes invasiveness (Yanagisawa et al., 2008). Two modulation categories were identified including 107 attenuates exon exclusion (AEE), 7 enhances exon exclusion (EEE). The alternative spliced exon of CTNND1 and some of it's in each category are shown in Figure 5A.

For example, TNFRSF14 as one of modulators of QKI attenuates the splicing regulation on the 20th exon exclusion of CTNND1. We found that, in TNFRSF14 expression low group, the correlation between QKI expression and CTNND1 PSI is -0.38 ($p = 1.7e-05$), while such correlation is lost when in TNFRSF14 expression high group (correlation = 0.001, $p = 0.98$). This indicated that high expression of modulator TNFRSF14 may play negative effect on changing the splicing activity of QKI. In addition, we found that LAPTM4B, as another modulator of QKI, played enhanced exon exclusion role on regulating the 20th exon splicing outcome when it expression is high (Figure 5B).



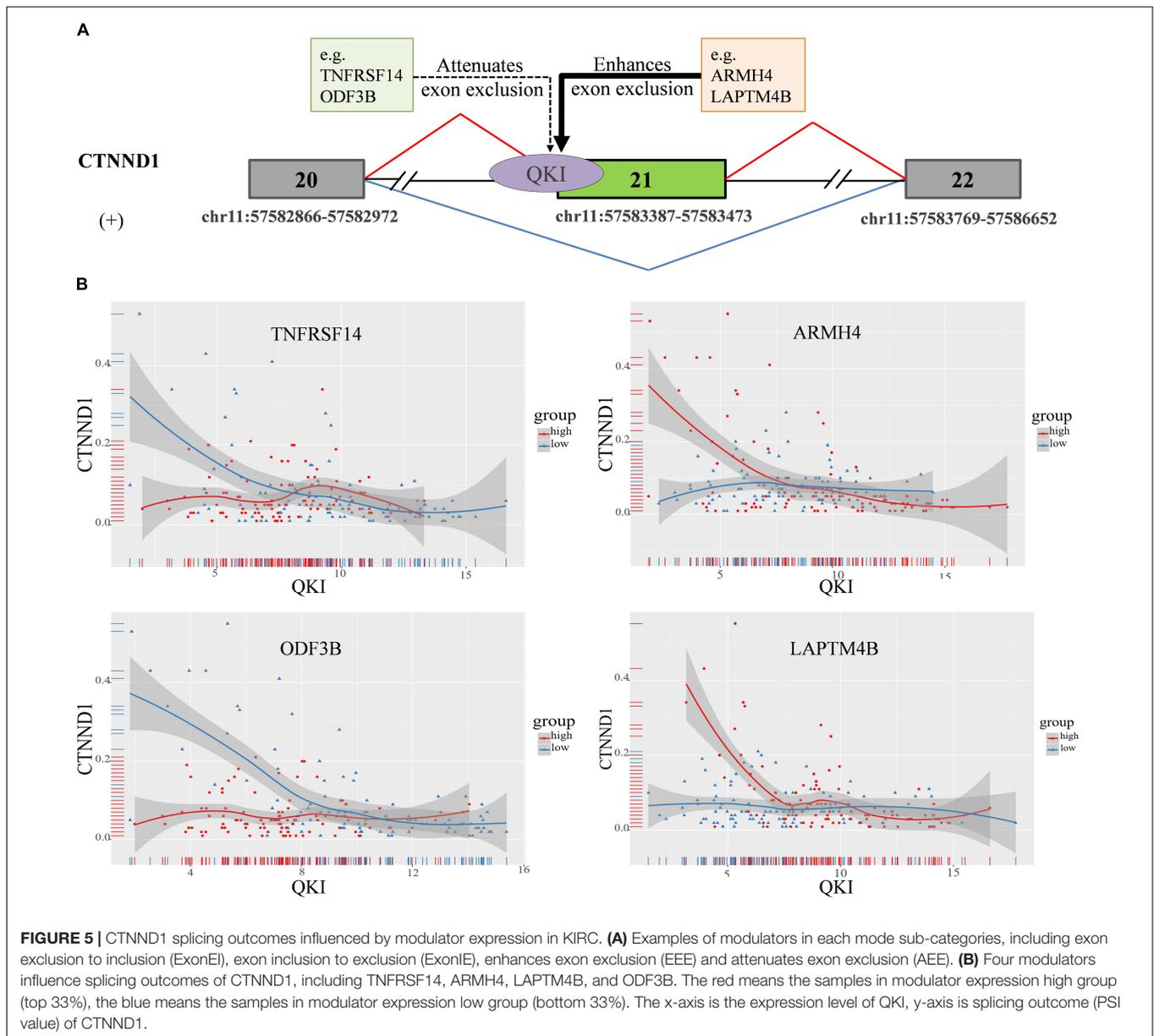
Thus, the results showed that differentially expressed modulators indeed changed the role of QKI on regulating CTNND1’s splicing outcomes, and we believed that this kind of regulation may provide important insights for study dysregulation of splicing outcomes associated with various diseases.

DISCUSSION

Alternative splicing alterations may confer a selective advantage to the tumor, such as angiogenesis (Amin et al., 2011), proliferation (Bechara et al., 2013), cell invasion (Venables et al., 2013), and avoidance of apoptosis (Izquierdo et al., 2005). Some splicing mRNA isoforms could change the reading frame,

resulting in the generation of different protein isoforms with diverse functions and/or localizations (Sutandy et al., 2018). One of the traditional methods to estimate the functions of mRNAs or protein is comparing the difference of gene expression level (Kim et al., 2014; Lorthongpanich et al., 2019; Xu et al., 2019).

However, not all detected alternative splicing events might necessarily result in mRNAs or proteins expression level changing. In addition, global description of alternative splicing networks and demonstration of their functional consequences have now emerged as one of the biggest challenges of the field (Baralle and Giudice, 2017). By integrating gene expression profile with splicing outcomes of alternative splicing events may be one of the possible ways to study the functional consequences for most of the identified splicing events.



In this study, we established a computational method for identification the modulators whose expression is associated with changing the targets splicing outcomes of RBPs in KIRC. Previously, several computational methods have been developed to identify modulators associated with transcription factors (TFs) regulation activity on expression level (Wang et al., 2009; Babur et al., 2010; Li et al., 2016, 2017), these studies discussed the transcriptional activities of TFs can be influenced by the expression level of modulators. Our method is unique in its ability to discover how alternative splicing outcomes is changing when modulator expression is different, even though they were regulated by the same RBP. And the method aimed at dissecting the effects of disruption in RBPs and hopefully it could provide insight into studying alternative splicing networks during development, cell differentiation, and in disease.

During tissue development and cell differentiation specific RBPs are finely regulated at their expression levels, localization, their own splicing, mRNA stability, and translation efficiency (Baralle and Giudice, 2017). RBPs bind to *cis*-elements promoting or inhibiting splice site recognition, hence RBP expression coordinates alternative splicing networks during development. We focused on 150 RBPs in this study, and only 137 RBPs remained in the final analysis due to there was no splicing targets of them within our criteria. Three possible modes of modulator actions were defined in this study, depending on the correlation changes between RBP and its target splicing outcomes when modulator expression is different. Among these RBPs, we found that QKI had the greatest number of influenced spliced exons (68.9%, 199 out of 289), and 2,014 Modulator-QKI-Splicing triplets were finally identified focused on QKI. Results showed

that most modulators affected multiple splicing targets were multimodal, and the same modulator may play opposite roles on different QKI targets.

For example, high expression level of modulator ADAM8 enhanced QKI role on regulating ACSF2 exon exclusion, while it enhanced target FMNL2 exon inclusion in regulating splicing outcomes. Another example, low expression level of AJM1 modulated QKI attenuated exon exclusion on regulating CTNND1 splicing outcomes, while such modulation role changed into enhanced exon inclusion when the target became ATF2. The detailed information about the modulators roles in the triples could be found in **Supplementary Table S1**.

Pathway enrichment results showed that all these influenced splicing target events of QKI were enrolled in cancer development and progression related pathways, including tight junction pathway, transcriptional mis-regulation in cancer, endocytosis, and MAPK signaling pathways. This evidence indicated that these alternative spliced events played crucial roles in kidney cancer, and changes the splicing outcomes of them may result in dysregulation in alternative splicing networks.

Among all these influenced splicing events, CTNND1 attracted more attention as the splicing outcomes of the 20th exon has the most inferred modulators. Previously study reported that CTNND1 was a tumor-driver gene, whose alternative splicing was related to cell invasion and metastasis (Yanagisawa et al., 2008). In addition, CTNND1 (p120) consists of central ARM domain flanked by the N-terminal regulatory (NTR) and C-terminal tail region (CTR) (Ishiyama et al., 2010), and the 20th exon of CTNND1 is in CTR region. Thus, different splicing outcomes of CTNND1 may influence the domain function, resulting in the generation of different protein isoforms with diverse functions.

We identified 114 inferred modulators of QKI-CTNND1 pair, including 30 lncRNAs and 80 protein-coding RNAs. Han et al. (2016) reported that MALAT1 may play as a tumor-suppressor gene in gliomas, and high MALAT1 expression linked to cell proliferation and metastasis. In our results, we noticed that high expression of modulator MALAT1 tended to attenuate QKI regulation role on splicing the 20th exon exclusion in CTNND1. The correlation between QKI expression and CTNND1 PSI was -0.51 ($p = 4.3e-09$) in MALAT1 expression low group, and such correlation changed into -0.20 ($p = 0.02$) in MALAT1 expression high group. LINC00174 as another inferred modulator had been reported that it exerted a tumorigenesis role in glioma. LINC00174 knockdown inhibited cell proliferation, migration, invasion and glycolysis (Shi et al., 2019). We found that, when LINC00174 expression is low, the correlation between QKI expression and splicing outcomes of CTNND1 is -0.37 ($p = 2.6e-05$),

while such correlation was lost ($r = -0.07$, $p = 0.44$) when LINC00174 expression became high. These evidences showed that regulation of alternative splicing outcomes is a complex progress, it different splicing consequence not only associated with RBP but also associated with other proteins expression.

Although the established model in this study and the corresponding results appear helpful for understanding the alternative splicing regulation, there are some limitations. First, many proteins tended to show similar expression pattern, certain RBP-target pairs may have more than two inferred modulators, these results may contain certain false positive modulators. In addition, the function and mechanism of how modulator changes the RBPs regulation on their target splicing outcomes need to be further studied by experiments, for example, modulator co-expressed or physically interacted with RBPs, and this is a long way for us to go. Finally, we expect that our study could provide novel insights for understanding the dysregulation of alternative splicing in cancer.

DATA AVAILABILITY STATEMENT

Paired-end RNA sequencing data from 480 RCC patients was downloaded from the Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma (TCGA-KIRC).

AUTHOR CONTRIBUTIONS

YL and YW designed the study. YW and SC wrote the code, analyzed data, and wrote the manuscript. XR calculated the alternative splicing events in KIRC. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We would appreciate detailed suggestions from reviewers who significantly helped us improved the early version of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00265/full#supplementary-material>

TABLE S1 | Inferred QKI modulators and triplets' information.

- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556
- Babur, O., Demir, E., Gonen, M., Sander, C., and Dogrusoz, U. (2010). Discovering modulators of gene expression. *Nucleic Acids Res.* 38, 5648–5656. doi: 10.1093/nar/gkq287

REFERENCES

- Amin, E. M., Oltean, S., Hua, J., Gammons, M. V., Hamdollah-Zadeh, M., Welsh, G. I., et al. (2011). WT1 mutants reveal SRPK1 to be a downstream angiogenesis target by altering VEGF splicing. *Cancer Cell* 20, 768–780. doi: 10.1016/j.ccr.2011.10.016

- Baralle, F. E., and Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* 18, 437–451. doi: 10.1038/nrm.2017.27
- Bechara, E. G., Sebestyen, E., Bernardis, I., Eyras, E., and Valcarcel, J. (2013). RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation. *Mol. Cell.* 52, 720–733. doi: 10.1016/j.molcel.2013.11.010
- Church, D. M., Schneider, V. A., Graves, T., Auger, K., Cunningham, F., Bouk, N., et al. (2011). Modernizing reference genome assemblies. *PLoS Biol.* 9:e1001091. doi: 10.1371/journal.pbio.1001091
- Consortium Encode Project, (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74. doi: 10.1038/nature11247
- Di Liegro, C. M., Schiera, G., and Di Liegro, I. (2014). Regulation of mRNA transport, localization and translation in the nervous system of mammals (review). *Int. J. Mol. Med.* 33, 747–762. doi: 10.3892/ijmm.2014.1629
- Garcia-Maurino, S. M., Rivero-Rodriguez, F., Velazquez-Cruz, A., Hernandez-Vellisca, M., Diaz-Quintana, A., De la Rosa, M. A., et al. (2017). RNA binding protein regulation and cross-talk in the control of AU-rich mRNA fate. *Front. Mol. Biosci.* 4:71. doi: 10.3389/fmolb.2017.00071
- Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. *Nat. Rev. Genet.* 15, 829–845. doi: 10.1038/nrg3813
- Han, Y., Wu, Z., Wu, T., Huang, Y., Cheng, Z., Li, X., et al. (2016). Tumor-suppressive function of long noncoding RNA MALAT1 in glioma cells by downregulation of MMP2 and inactivation of ERK/MAPK signaling. *Cell Death Dis.* 7:e2123. doi: 10.1038/cddis.2015.407
- Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., et al. (2012). GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* 22, 1760–1774. doi: 10.1101/gr.135350.111
- Hsieh, J. J., Purdue, M. P., Signoretti, S., Swanton, C., Albiges, L., Schmidinger, M., et al. (2017). Renal cell carcinoma. *Nat. Rev. Dis. Primers* 3:17009. doi: 10.1038/nrdp.2017.9
- Ishiyama, N., Lee, S. H., Liu, S., Li, G. Y., Smith, M. J., Reichardt, L. F., et al. (2010). Dynamic and static interactions between p120 catenin and E-cadherin regulate the stability of cell-cell adhesion. *Cell* 141, 117–128. doi: 10.1016/j.cell.2010.01.017
- Izquierdo, J. M., Majos, N., Bonnal, S., Martinez, C., Castelo, R., Guigo, R., et al. (2005). Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. *Mol. Cell.* 19, 475–484. doi: 10.1016/j.molcel.2005.06.015
- Jankowsky, E., and Harris, M. E. (2015). Specificity and nonspecificity in RNA-protein interactions. *Nat. Rev. Mol. Cell Biol.* 16, 533–544. doi: 10.1038/nrm4032
- Janzen, N. K., Kim, H. L., Figlin, R. A., and Beldegrun, A. S. (2003). Surveillance after radical or partial nephrectomy for localized renal cell carcinoma and management of recurrent disease. *Urol. Clin. North Am.* 30, 843–852. doi: 10.1016/s0094-0143(03)00056-9
- Kim, Y., Choi, J. W., Lee, J. H., and Kim, Y. S. (2014). Loss of CDC14B expression in clear cell renal cell carcinoma: meta-analysis of microarray data sets. *Am. J. Clin. Pathol.* 141, 551–558. doi: 10.1309/AJCP4PE4JPSRQBQS
- Ladomery, M. (2013). Aberrant alternative splicing is another hallmark of cancer. *Int. J. Cell Biol.* 2013:463786. doi: 10.1155/2013/463786
- Lee, J. H., Ang, J. K., and Xiao, X. (2013). Analysis and design of RNA sequencing experiments for identifying RNA editing and other single-nucleotide variants. *RNA* 19, 725–732. doi: 10.1261/rna.037903.112
- Li, J., Wang, Y., Rao, X., Wang, Y., Feng, W., Liang, H., et al. (2017). Roles of alternative splicing in modulating transcriptional regulation. *BMC Syst. Biol.* 11(Suppl. 5):89. doi: 10.1186/s12918-017-0465-6
- Li, Y., Wang, Z., Wang, Y., Zhao, Z., Zhang, J., Lu, J., et al. (2016). Identification and characterization of lncRNA mediated transcriptional dysregulation dictates lncRNA roles in glioblastoma. *Oncotarget* 7, 45027–45041. doi: 10.18632/oncotarget.7801
- Lorthongpanich, C., Thumanu, K., Tangkiettrakul, K., Jiamvoraphong, N., Laowtamathron, C., Damkham, N., et al. (2019). YAP as a key regulator of adipo-osteogenic differentiation in human MSCs. *Stem Cell Res. Ther.* 10:402. doi: 10.1186/s13287-019-1494-4
- Mickisch, G. H. (2002). Principles of nephrectomy for malignant disease. *BJU Int.* 89, 488–495. doi: 10.1046/j.1464-410x.2002.02654.x
- Payne, J. L., Khalid, F., and Wagner, A. (2018). RNA-mediated gene regulation is less evolvable than transcriptional regulation. *Proc. Natl. Acad. Sci. U.S.A.* 115, E3481–E3490. doi: 10.1073/pnas.1719138115
- Repana, D., Nulsen, J., Dressler, L., Bortolomeazzi, M., Venkata, S. K., Tournara, A., et al. (2019). The Network of Cancer Genes (NCG): a comprehensive catalogue of known and candidate cancer genes from cancer sequencing screens. *Genome Biol.* 20:1. doi: 10.1186/s13059-018-1612-0
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/bt p616
- Scotti, M. M., and Swanson, M. S. (2016). RNA mis-splicing in disease. *Nat. Rev. Genet.* 17, 19–32. doi: 10.1038/nrg.2015.3
- Shi, J., Zhang, Y., Qin, B., Wang, Y., and Zhu, X. (2019). Long non-coding RNA LINC00174 promotes glycolysis and tumor progression by regulating miR-152-3p/SLC2A1 axis in glioma. *J. Exp. Clin. Cancer Res.* 38:395. doi: 10.1186/s13046-019-1390-x
- Shukla, S., and Oberdoerffer, S. (2012). Co-transcriptional regulation of alternative pre-mRNA splicing. *Biochim. Biophys. Acta* 1819, 673–683. doi: 10.1016/j.bbagr.2012.01.014
- Siegel, R. L., Miller, K. D., and Jemal, A. (2019). Cancer statistics, 2019. *CA Cancer J. Clin.* 69, 7–34. doi: 10.3322/caac.21551
- Sutandy, F. X. R., Ebersberger, S., Huang, L., Busch, A., Bach, M., Kang, H. S., et al. (2018). In vitro iCLIP-based modeling uncovers how the splicing factor U2AF2 relies on regulation by cofactors. *Genome Res.* 28, 699–713. doi: 10.1101/gr.229757.117
- The Gene Ontology Consortium, (2017). Expansion of the gene ontology knowledgebase and resources. *Nucleic Acids Res.* 45, D331–D338. doi: 10.1093/nar/gkw1108
- Van Nostrand, E. L., Pratt, G. A., Shishkin, A. A., Gelboin-Burkhart, C., Fang, M. Y., Sundararaman, B., et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* 13, 508–514. doi: 10.1038/nmeth.3810
- Venables, J. P., Brosseau, J. P., Gadea, G., Klinck, R., Prinos, P., Beaulieu, J. F., et al. (2013). RBFOX2 is an important regulator of mesenchymal tissue-specific splicing in both normal and cancer tissues. *Mol. Cell Biol.* 33, 396–405. doi: 10.1128/MCB.01174-12
- Wang, E. T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., et al. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476. doi: 10.1038/nature07509
- Wang, K., Saito, M., Bisikirska, B. C., Alvarez, M. J., Lim, W. K., Rajbhandari, P., et al. (2009). Genome-wide identification of post-translational modulators of transcription factor activity in human B cells. *Nat. Biotechnol.* 27, 829–839. doi: 10.1038/nbt.1563
- Wang, Y., Liu, J., Huang, B. O., Xu, Y. M., Li, J., Huang, L. F., et al. (2015). Mechanism of alternative splicing and its regulation. *Biomed. Rep.* 3, 152–158. doi: 10.3892/br.2014.407
- Weyn-Vanhentenryck, S. M., Mele, A., Yan, Q., Sun, S., Farny, N., Zhang, Z., et al. (2014). HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. *Cell Rep.* 6, 1139–1152. doi: 10.1016/j.celrep.2014.02.005
- Xu, X., Yu, Y., Zong, K., Lv, P., and Gu, Y. (2019). Up-regulation of IGF2BP2 by multiple mechanisms in pancreatic cancer promotes cancer proliferation by activating the PI3K/Akt signaling pathway. *J. Exp. Clin. Cancer Res.* 38:497. doi: 10.1186/s13046-019-1470-y
- Yanagisawa, M., Huvelde, D., Kreinest, P., Lohse, C. M., Chevillat, J. C., Parker, A. S., et al. (2008). A p120 catenin isoform switch affects Rho activity, induces tumor cell invasion, and predicts metastatic disease. *J. Biol. Chem.* 283, 18344–18354. doi: 10.1074/jbc.M801192200
- Yee, B. A., Pratt, G. A., Graveley, B. R., Van Nostrand, E. L., and Yeo, G. W. (2019). RBP-Maps enables robust generation of splicing regulatory maps. *RNA* 25, 193–204. doi: 10.1261/rna.069237.118
- Yin, L., Li, W., Wang, G., Shi, H., Wang, K., Yang, H., et al. (2019). NR1B2 suppress kidney renal clear cell carcinoma (KIRC) progression by regulation of LATS 1/2-YAP signaling. *J. Exp. Clin. Cancer Res.* 38:343. doi: 10.1186/s13046-019-1344-3

- Zhang, C., Frias, M. A., Mele, A., Ruggiu, M., Eom, T., Marney, C. B., et al. (2010). Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. *Science* 329, 439–443. doi: 10.1126/science.1191150
- Zhao, M., Kim, P., Mitra, R., Zhao, J., and Zhao, Z. (2016). TSGene 2.0: an updated literature-based knowledgebase for tumor suppressor genes. *Nucleic Acids Res.* 44, D1023–D1031. doi: 10.1093/nar/gkv1268
- Zhu, Y. T., Chen, H. C., Chen, S. Y., and Tseng, S. C. (2012). Nuclear p120 catenin unlocks mitotic block of contact-inhibited human corneal endothelial monolayers without disrupting adherent junctions. *J. Cell Sci.* 125(Pt 15), 3636–3648. doi: 10.1242/jcs.103267

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Wang, Chen, Rao and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.