Molecular Characterization and Clinical Relevance of RNA Binding Proteins in Colorectal Cancer

Zhen Zhang[1,2†], Ling Wang[3†], Quan Wang[1,2], Mengmeng Zhang[1,2], Bo Wang[1,2], Kewei Jiang[1,2], Yingjiang Ye[1,2], Shan Wang[1,2]* and Zhanlong Shen[1,2]*

[1] Department of Gastroenterological Surgery, Peking University People's Hospital, Beijing, China, [2] Laboratory of Surgical Oncology, Beijing Key Laboratory of Colorectal Cancer Diagnosis and Treatment Research, Peking University People's Hospital, Beijing, China, [3] Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Hepatopancreatobiliary Surgery Department I, Peking University Cancer Hospital and Institute, Beijing, China

Abnormal expression of RNA binding proteins (RBPs) has been reported across various cancers. However, the potential role of RBPs in colorectal cancer (CRC) remains unclear. In this study, we performed a systematic bioinformatics analysis of RBPs in CRC. We downloaded CRC data from The Cancer Genome Atlas (TCGA) database. Our analysis identified 242 differentially expressed RBPs between tumor and normal tissues, including 200 upregulated and 42 downregulated RBPs. Next, we found eight RBPs (RRS1, PABPC1L, TERT, SMAD6, UPF3B, RP9, NOL3, and PTRH1) related to the prognoses of CRC patients. Among these eight prognosis-related RBPs, four RBPs (NOL3, PTRH1, UPF3B, and SMAD6) were selected to construct a prognostic risk score model. Furthermore, our results indicated that the prognostic risk score model accurately predicted the prognosis of CRC patients [area under the receiver operating characteristic curve (AUC)for 3- and 5-year overall survival (OS) and was 0.645 and 0.672, respectively]. Furthermore, we developed a nomogram based on a prognostic risk score model. The nomogram was able to demonstrate the wonderful performance in predicting 3- and 5-year OS. Additionally, we validated the clinical value of four risk genes in the prognostic risk score model and identified that these risk genes were associated with tumorigenesis, lymph node metastasis, distant metastasis, clinical stage, and prognosis. Finally, we used the TIMER and Human Protein Atlas (HPA)database to validate the expression of four risk genes at the transcriptional and translational levels, respectively, and used a clinical cohort to validate the roles of NOL3 and UPF3B in predicting the prognosis of CRC patients. In summary, our study demonstrated that RBPs have an effect on CRC tumor progression and might be potential prognostic biomarkers for CRC patients.

Keywords: colorectal cancer, RNA binding protein, prognostic model, transcriptomics, TCGA

INTRODUCTION

Colorectal cancer (CRC) is one of the most common cancers of the gastrointestinal tract. It is the third leading cause of cancer-related death worldwide (Siegel et al., 2020). Although surgical and adjuvant therapies have improved, the 5-year overall survival (OS) rate of CRC patients ranges from 90 to 10% (Van Cutsem et al., 2014). The poor prognosis of CRC is primarily due to tumor distant metastasis and recurrence

(Roth et al., 2010; Kow, 2019). Therefore, understanding the mechanisms that lead to CRC initiation and progression is necessary for diagnosis, therapeutic interventions, and prognostic prediction.

RNA binding proteins (RBPs) are a class of proteins involved in splicing, modifications, transport, localization, stability, degradation, and translation of RNAs (Mitchell and Parker, 2014; Perron et al., 2018). In fact, more than 1,500 human RBPs have been validated by high-throughput screens and experiments (Gerstberger et al., 2014; Neelamraju et al., 2015). RBPs play vital roles in several essential cellular processes by interacting with their target RNAs (Moore, 2005). The target RNAs of RBPs are diverse and include microRNAs, transfer RNAs (tRNAs), small interfering RNA, small nucleolar RNAs, and small nuclear RNAs (Hentze et al., 2018). Abnormally expressed RBPs regulate the expression and function of oncogenes and tumor-suppressor genes via post-transcriptional regulatory mechanisms across various cancers. For example, aberrant hnRNPM expression promoted breast cancer metastasis by controlling CD44 splice isoform switching during epithelial–mesenchymal transition (EMT) (Xu et al., 2014). In melanomas, CPEB4 increased the translation of melanoma drivers MITF and RAB72A, which helps promote tumor proliferation (Pérez-Guijarro et al., 2016). In hepatocellular carcinoma, HuR/methyl-HuR and AUF1 modulate MAT1A and MAT2A expression through post-translational regulation of their messenger RNAs (mRNAs), thus impacting tumor progression (Vázquez-Chantada et al., 2010). In CRC, overexpression of IMP-1 increased proliferation by directly binding to and stabilizing c-Myc (Mongroo et al., 2011).

With the rapid development of high-throughput sequencing, researchers have have been able to perform a systematic functional analysis of RBPs using high-throughput bioinformatics profiling. Recent bioinformatics studies have implied that RBPs can predict the prognosis of different cancers, such as breast cancer, lung adenocarcinoma, glioma, hepatocellular carcinoma, and leukemia (Li et al., 2019a,b, 2020; Saha et al., 2019; Wang K. et al., 2019, Wang Z. et al., 2020). In this study, we downloaded CRC data from the Cancer Genome Atlas (TCGA) database. Next, we selected differentially expressed RBPs to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses. Furthermore, we identified prognostic RBPs that enable the construction of a prognostic risk score model. Furthermore, we built a nomogram based on the prognostic risk score model. Finally, we explored the clinical value of these risk genes. Our study detected that several RBPs are involved in CRC, which might be used to predict the prognosis of CRC patients and inhibit tumor progression in the future.

MATERIALS AND METHODS

Data Acquisition
RNA sequencing data and the corresponding clinical data were downloaded from the TCGA database[1]. Overall, 568 CRC

patients were used to analyze the differentially expressed RBPs in the TCGA database. Then, 489 CRC patients were selected for clinical analyses, as these patients had complete clinical information, including primary tumor, lymph node metastasis, distant metastasis, clinical stage, and follow-up for at least 1 month. Additionally, we downloaded RNA-seq data of the GSE29623 cohort from the Gene Expression Omnibus (GEO) database[2]. In total, 1,542 RBPs were included in our study (Gerstberger et al., 2014; Neelamraju et al., 2015).

Clinical Samples and Immunohistochemistry Staining
In total, tumor and normal tissues of 44 CRC patients were obtained from the Peking University People's Hospital. All tissues were histopathologically confirmed by pathologists. This study was granted approval by the ethics committee of Peking University People's Hospital. Immunohistochemistry (IHC)staining was performed according to prior published protocols (Zhang et al., 2019). Briefly, the staining index scores were assigned as follows: staining intensity (negative: 0; weak: 1; moderate: 2; strong: 3) and positive staining (<5%: 0; 5–25%: 1; 26–50%: 2; 51–75%: 3; > 75%: 4). The staining index scores were calculated by multiplying the staining intensity score by the positive staining score, which ranged from 0 to 12. The antibodies used in this study included anti-NOL3 (Proteintech, United States) and anti-UPF3B (Proteintech, United States).

Functional Enrichment Analysis of Differentially Expressed RNA Binding Proteins
The GO and KEGG pathway analyses were performed to analyze the biological functions of these differentially expressed RBPs. The GO terms include biological process (BP), cellular component (CC), and molecular function (MF).

Building a Prognostic Model
A univariate Cox regression analysis was performed to identify prognostic RBPs. Subsequently, a multivariable Cox regression analysis was carried out to construct a prognostic risk score model. Furthermore, the risk score was calculated as follows:

$$riskscore = \sum_{i=1}^{n} coefficieent(i) * expression(i),$$

where coefficient(i) and expression(i) represent the regression coefficient and expression levels of selected genes in the prognostic risk score model, respectively. The time-dependent receiver operating characteristic (ROC) analysis was used to assess the prognostic ability of the prognostic risk score model. Then, a nomogram was built to predict the OS of CRC patients. The calibration plot and concordance index (C−index) were used to evaluate the performance of the nomogram.

---

[1]https://portal.gdc.cancer.gov/

[2]https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29623

External Validation of the Expression and Genetic Alterations of the Risk Genes

The TIMER database[3] and the Human Protein Atlas (HPA) database[4] were utilized to explore and validate gene expression at the transcriptional and translational level, respectively. The cBioportal for Cancer Genomics[5] was used to identify each genetic alteration, which included four CRC studies.

Statistical Analyses

Statistical analyses were performed using the R software v3.5.1. Kaplan–Meier analysis was used to construct survival curves, and the log-rank test was utilized to assess the significance of differences. *T*-test and Wilcoxon signed-rank test were used to explore quantitative variables. $P < 0.05$ represents statistical significance.

RESULTS

Identifying Differentially Expressed RNA Binding Proteins in Colorectal Cancer

Figure 1A shows detailed designs of the study. We downloaded transcriptomic files of CRC from the TCGA database encompassing 568 tumors and 44 normal samples. The Wilcoxon signed-rank test was applied to identify significantly differentially expressed RBPs. The R package Limma was used to identify differentially expressed RBPs according to the following parameters: $|\log_2 FC| > 1$ and false discovery rate (FDR) < 0.05. Among the 1,542 RBPs, we detected 242 differentially expressed RBPs, including 200 upregulated and 42 downregulated RBPs between tumor and normal tissues (**Figures 1B,C**).

Functional Enrichment Analysis of Differentially Expressed RNA Binding Proteins in Colorectal Cancer

To investigate the biological significance of RBPs in CRC, we performed GO and KEGG pathway analyses of 242 differentially expressed RBPs using the R package clusterProfiler. We displayed the top 10 significantly enriched GO terms (**Figure 2A**). Results revealed that these differentially expressed RBPs were significantly enriched in biological processes such as non-coding RNA processing, ribonucleoprotein complex biogenesis, ribosome biogenesis, and ribosomal RNA metabolism. They were mainly located in the preribosome, nucleolar part, small-subunit processome, 90S preribosome, and cytoplasmic ribonucleoprotein granule. They were found to participate in various molecular functions, including catalytic activity (acting on RNA), catalytic activity (acting on a tRNA), ribonuclease activity, single-stranded RNA binding, and poly(U) RNA binding. For KEGG pathway analysis, we found that these differentially expressed RBPs were mainly associated with ribosome biogenesis in eukaryotes, RNA degradation, RNA

---

[3]https://cistrome.shinyapps.io/timer/
[4]http://www.proteinatlas.org
[5]http://www.cbioportal.org/

transport, mRNA surveillance pathway, and aminoacyl-tRNA biosynthesis (**Figure 2B**).

Identifying Prognostic RNA Binding Proteins in Colorectal Cancer

To assess the prognostic significance of RBPs in CRC, we performed a univariate Cox regression analysis of 242 differentially expressed RBPs. We detected eight RBPs (RRS1, PABPC1L, TERT, SMAD6, UPF3B, RP9, NOL3, and PTRH1) that were related to the prognosis of CRC patients ($p < 0.05$, **Figure 3A**). Also, all these eight prognostic RBPs were protective factors because CRC patients with high expression of these RBPs had poor prognosis. The expression of eight prognostic RBPs in tumor and normal tissues of CRC patients are shown in **Figure 3B**.

Constructing a Prognostic Risk Score Model in Colorectal Cancer

We constructed the optimum prognostic risk score model for prediction of OS of CRC patients by using multivariate Cox regression analysis. The identified eight prognostic RBPs were used to construct the prognostic risk score model. Among the eight prognostic RBPs, we identified UPF3B, SMAD6, NOL3, and PTRH1 as risk genes in the prognostic risk score model. Furthermore, coefficients of four risk genes are shown in **Figure 4A**. We calculated the risk scores using regression coefficient and expression levels of the risk genes according to this equation: risk score = 0.4257 * expression (UPF3B) + 0.5463 * expression (SMAD6) + 0.562 * expression (NOL3) + 2.368 * expression (PTRH1). Then, all CRC patients were divided into either high-risk or low-risk groups according to the risk scores. We demonstrated that patients in the high-risk group had a shorter OS time compared with those in the low-risk group (**Figure 4B**). We measured the prognostic ability of the risk score model through the use of an ROC analysis, which was conducted using the R package survivalROC. Also, our results indicated that the area under the ROC curve for 3- and 5-year OS was 0.645 and 0.672, respectively (**Figure 4C**). Heat map of mRNA expression indicated that all four were upregulated in the high-risk group (**Figure 4D**). Distribution of risk scores and survival status of patients are shown in **Figures 4E,F**. In addition, we found a higher percentage of deaths in the high-risk group.
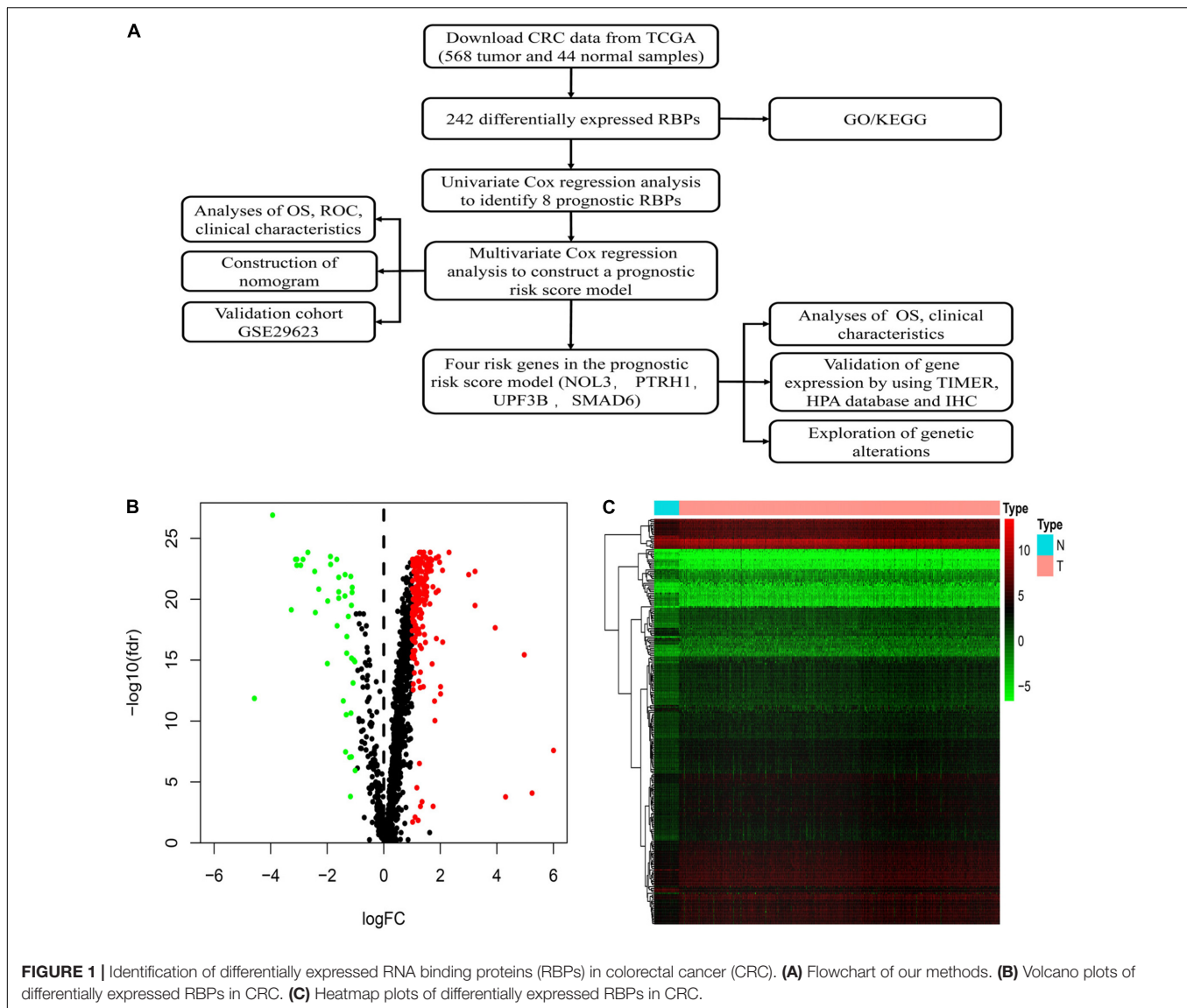
We next assessed whether the four RBP-related gene models can predict the survival prognosis of CRC patients among additional CRC cohorts. We calculated the risk scores using the same formula that was used in the GSE29623 cohort. The results demonstrated that patients in the high-risk group had shorter OS time compared with those in the low-risk group (**Supplementary Figures 1A–D**), which is consistent with the TCGA CRC cohort. The results demonstrated that the four RBP-related gene models can accurately predict the prognosis of CRC patients.

Additionally, we explored the relationship between risk scores and clinical features of CRC patients. These results implied that risk scores were significantly higher in CRC patients that have deeper tumor infiltration (**Figure 4G**), distant metastasis

FIGURE 1 | Identification of differentially expressed RNA binding proteins (RBPs) in colorectal cancer (CRC). **(A)** Flowchart of our methods. **(B)** Volcano plots of differentially expressed RBPs in CRC. **(C)** Heatmap plots of differentially expressed RBPs in CRC.

(**Figure 4H**), and lymph node metastasis (**Figure 4I**) and are at a late clinical stage (**Figure 4J**).

Building a Nomogram Based on the Prognostic Risk Score Model

Univariate and multivariate Cox regression analyses demonstrated that age, clinical stage, and the risk scores obtained from the prognostic risk score model were independent prognostic factors of CRC patients (**Figures 5A,B**). Therefore, based on these three independent prognostic factors, we constructed a nomogram that would be able to predict 3- and 5-year OS of CRC patients using the R package rms (**Figure 6A**). The calibration plots showed that the probability of predicting 3- or 5-year OS through nomogram agreed with actual observation (**Figures 6B,C**). Furthermore, the C-index for predicting OS through nomogram was 0.75 (95% CI: 0.69–0.81).

Validating Clinical Value of the Four Risk Genes in the Prognostic Risk Score Model

To further explore the clinical value of these four risk genes, we analyzed the relationship between expression of these genes and clinical features of the CRC patients. The survival curves showed that high expression of NOL3, PTRH1, SMAD6, and UPF3B is associated with poor prognosis of CRC patients (**Figures 7A–D**). We also found that NOL3 was significantly overexpressed in patients with lymph node metastasis (**Figure 7E**), distant metastasis (**Figure 7F**), and at a late clinical stage (**Figure 7G**). PTRH1 was found to be significantly overexpressed in patients with deeper tumor infiltration (**Figure 7H**). UPF3B was significantly upregulated in patients with deeper tumor infiltration (**Figure 7I**), as well as late clinical stage (**Figure 7J**). However, SMAD6 was downregulated in patients with deeper tumor infiltration (**Figure 7K**).
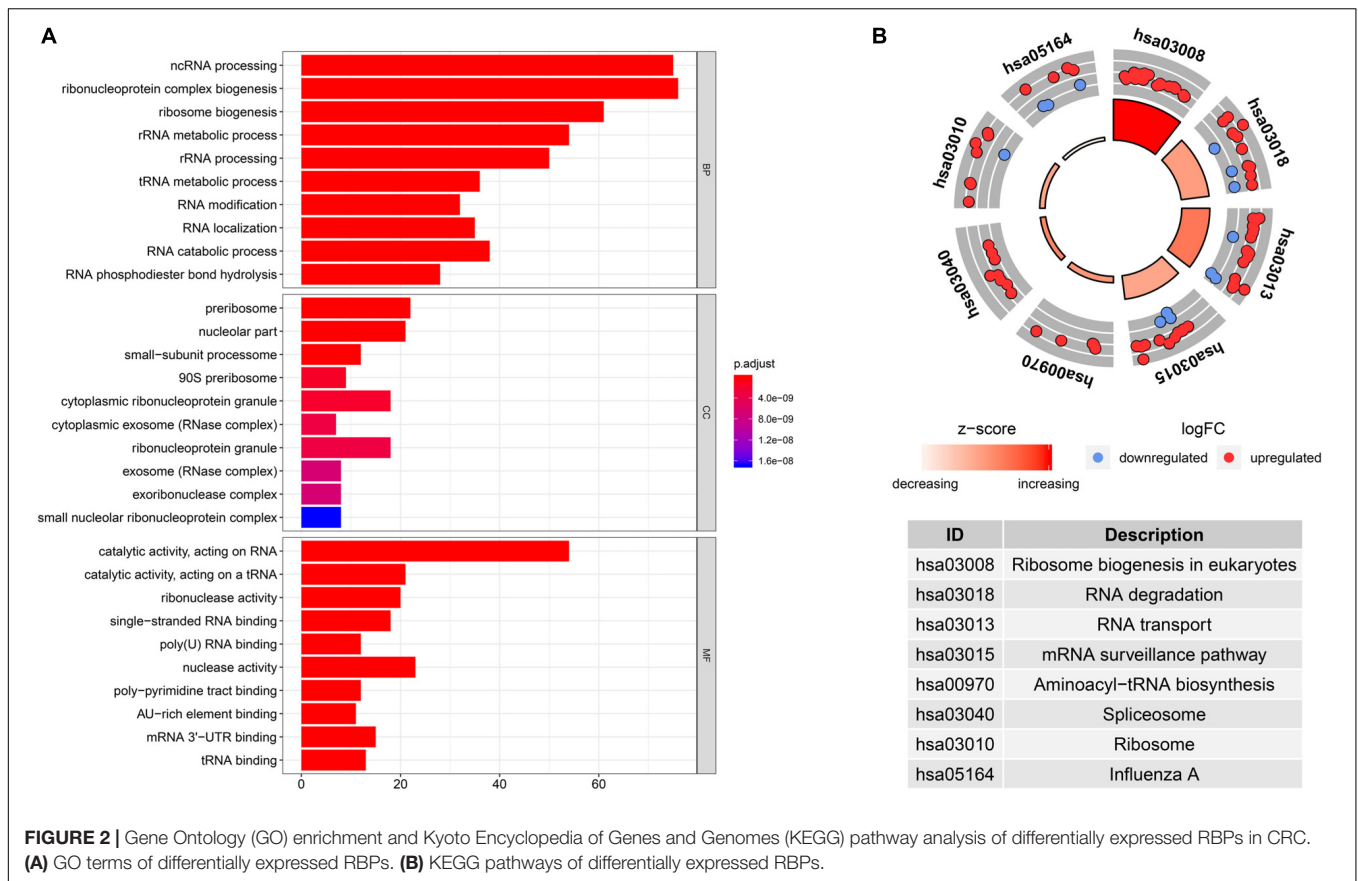
FIGURE 2 | Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of differentially expressed RBPs in CRC. **(A)** GO terms of differentially expressed RBPs. **(B)** KEGG pathways of differentially expressed RBPs.
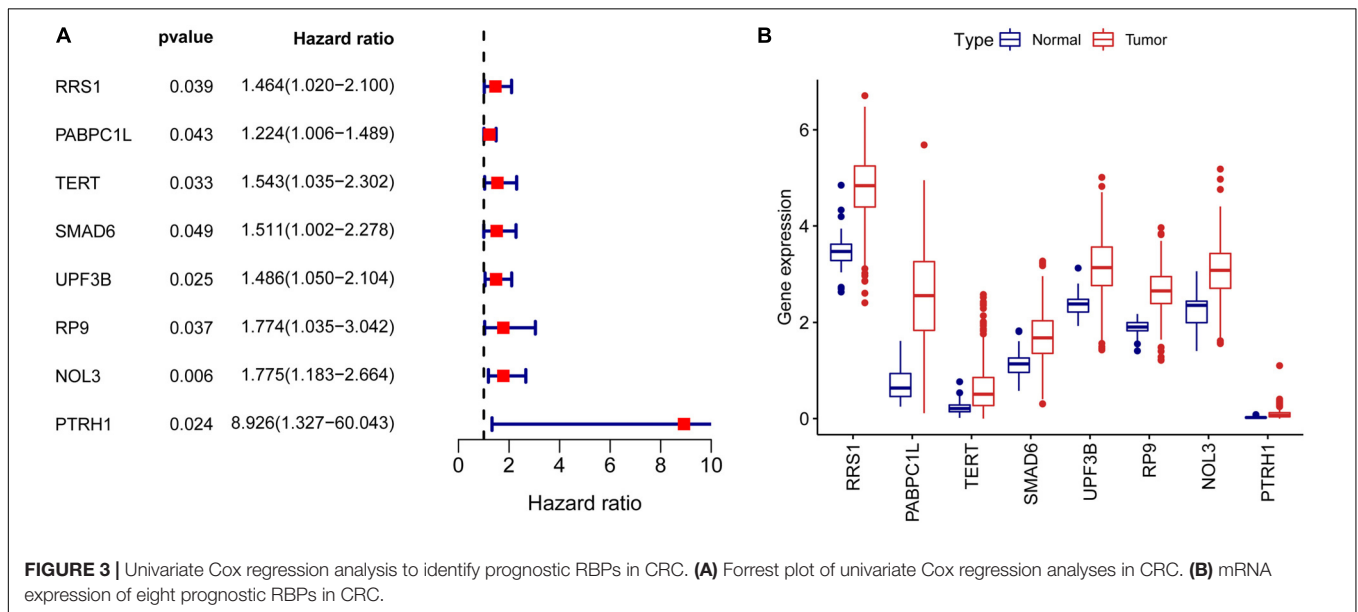

FIGURE 3 | Univariate Cox regression analysis to identify prognostic RBPs in CRC. **(A)** Forrest plot of univariate Cox regression analyses in CRC. **(B)** mRNA expression of eight prognostic RBPs in CRC.

External Validation of the Four Risk Genes

Consistent with results from the TCGA database, each of the four risk genes was found to be significantly upregulated in both colon and rectal cancer, according to the TIMER database (**Figure 8**).

Interestingly, we identified that the expression of these four genes is not the same across different cancers. For instance, UPF3B is upregulated in esophageal carcinoma, stomach adenocarcinoma, hepatocellular carcinoma, and breast invasive carcinoma, whereas UPF3B was downregulated in kidney chromphobe, prostate adenocarcinoma, and thyroid carcinoma. Additionally,
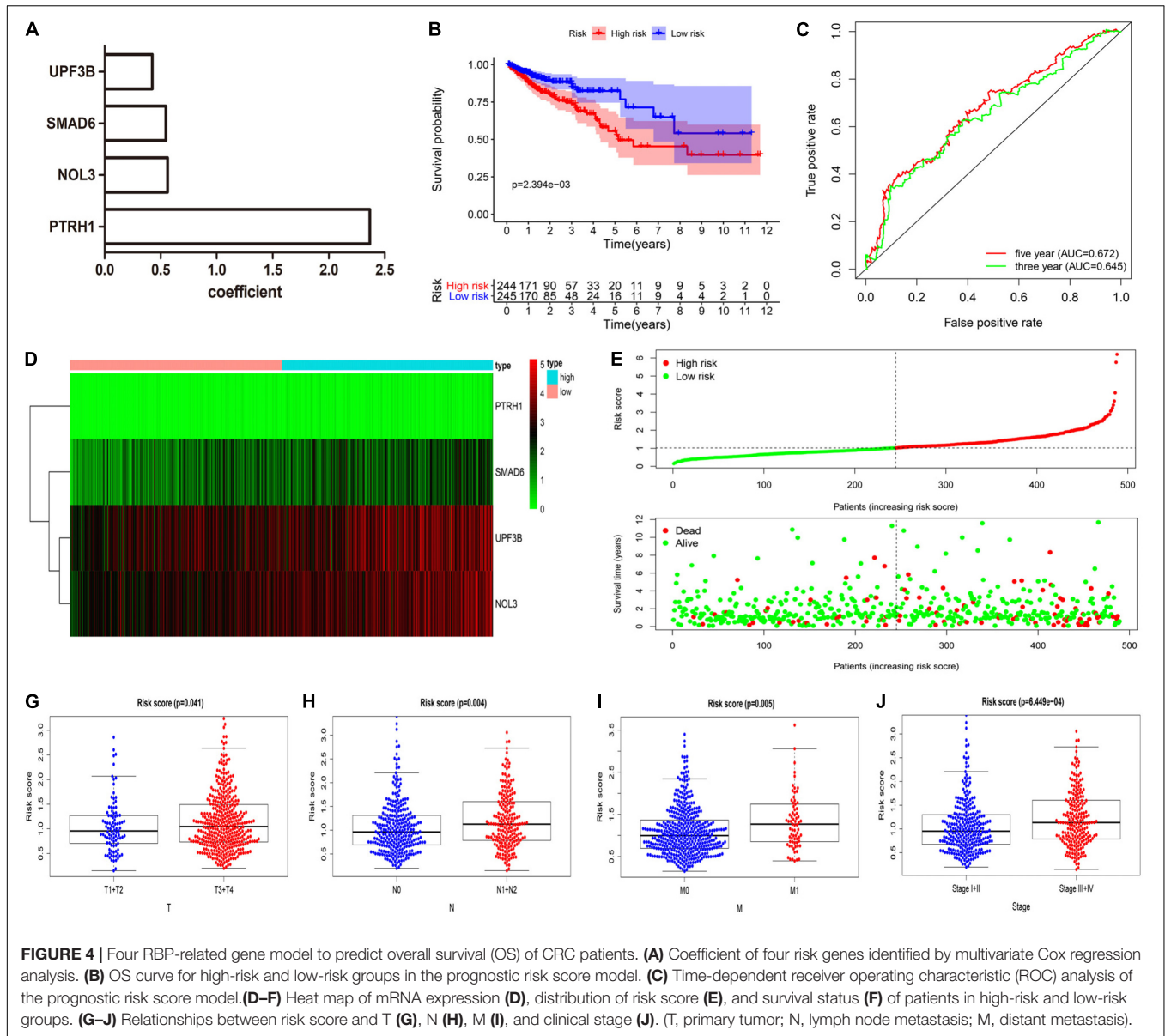
FIGURE 4 | Four RBP-related gene model to predict overall survival (OS) of CRC patients. (A) Coefficient of four risk genes identified by multivariate Cox regression analysis. (B) OS curve for high-risk and low-risk groups in the prognostic risk score model. (C) Time-dependent receiver operating characteristic (ROC) analysis of the prognostic risk score model.(D–F) Heat map of mRNA expression (D), distribution of risk score (E), and survival status (F) of patients in high-risk and low-risk groups. (G–J) Relationships between risk score and T (G), N (H), M (I), and clinical stage (J). (T, primary tumor; N, lymph node metastasis; M, distant metastasis).
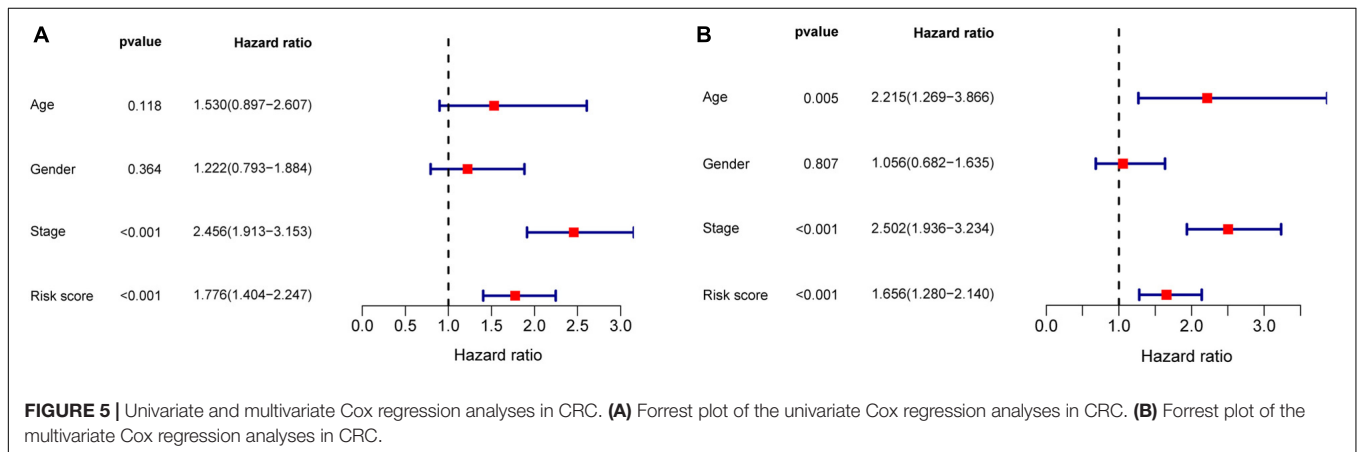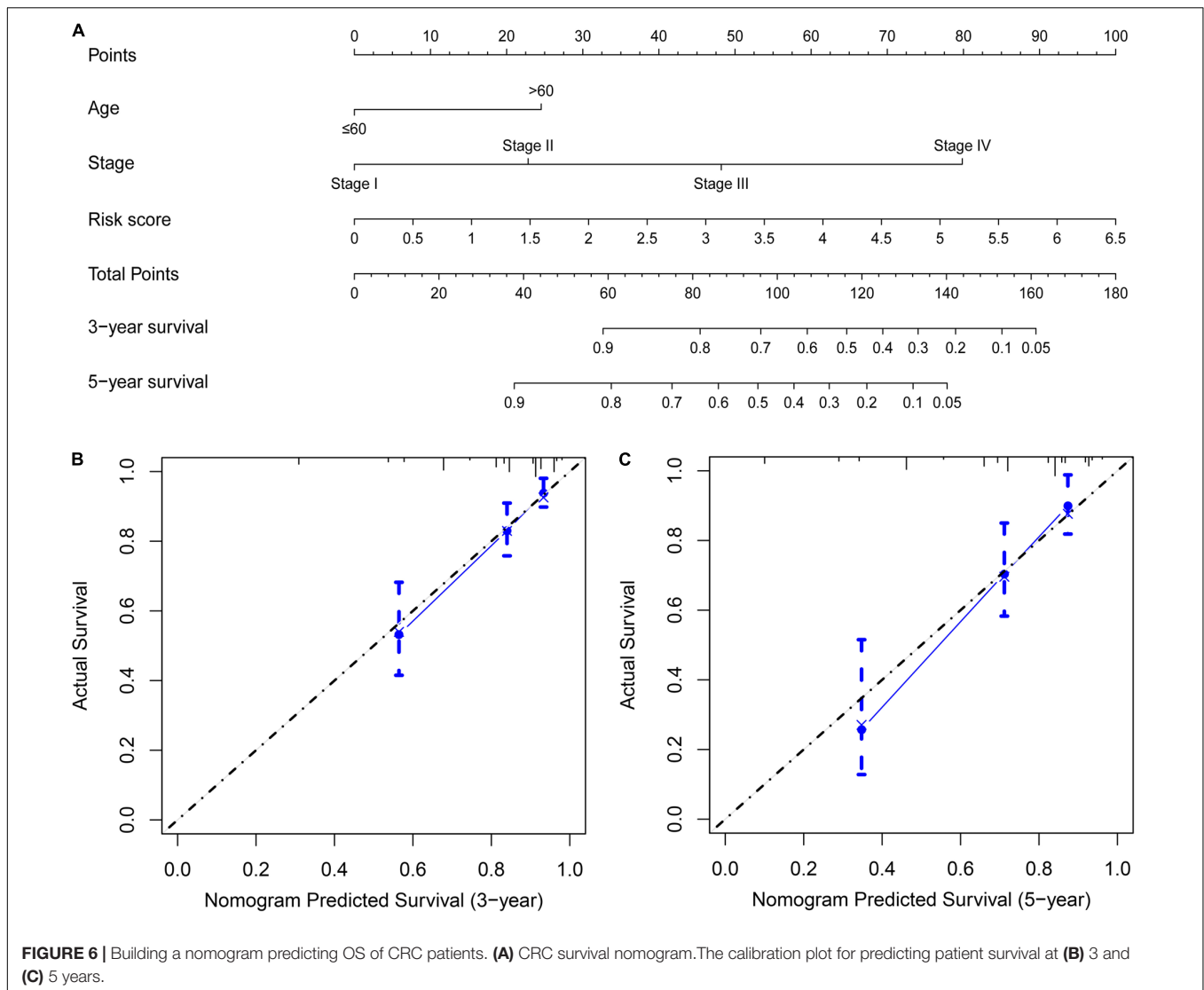

FIGURE 5 | Univariate and multivariate Cox regression analyses in CRC. (A) Forrest plot of the univariate Cox regression analyses in CRC. (B) Forrest plot of the multivariate Cox regression analyses in CRC.

FIGURE 6 | Building a nomogram predicting OS of CRC patients. **(A)** CRC survival nomogram. The calibration plot for predicting patient survival at **(B)** 3 and **(C)** 5 years.

the expression of UPF3B was not found to be significantly different in pancreatic adenocarcinoma, kidney renal clear cell carcinoma, and kidney renal papillary cell carcinoma.

To further validate the expression of the four risk genes at the translational level, we analyzed the protein expression of these genes in the HPA database. The results indicated that NOL3 and UPF3B were overexpressed in CRC tumor tissues compared with normal tissues (**Figure 9A**). PTRH1 expression was not significantly different between the CRC tumor tissues and normal tissues (**Figure 9A**). However, information of SMAD6 levels were not found on the website. In addition, genetic alterations of the four risk genes were found to rarely occur (**Figure 9B**).
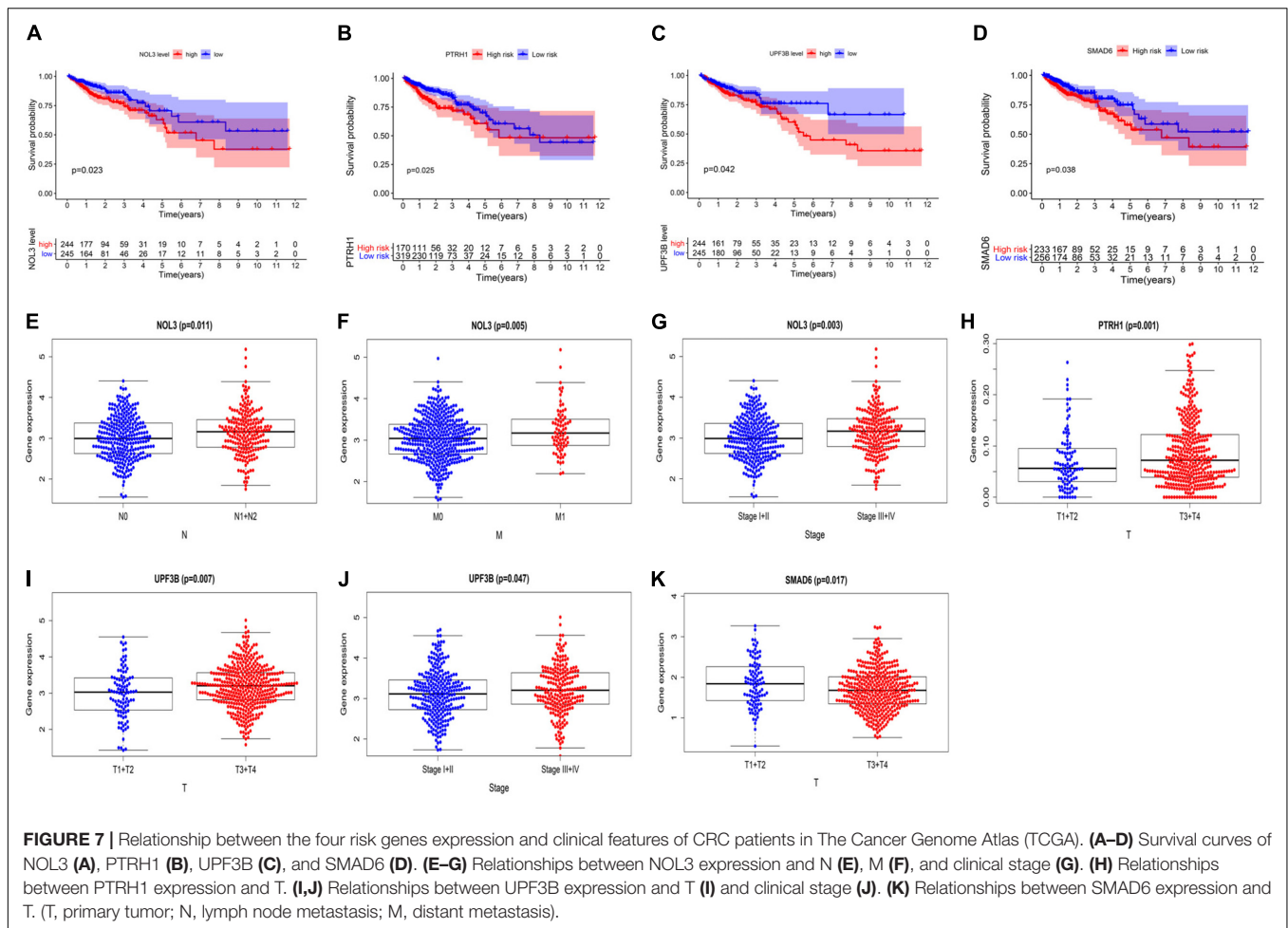
Validation of the Clinical Significance of NOL3 and UPF3B in Colorectal Cancer Patients by Immunohistochemistry

We obtained 44 pairs of CRC samples from Peking University People's Hospital to validate protein expression of the two key

RBPs (NOL3 and UPF3B) in CRC. Immunohistochemical staining results displayed that NOL3 and UPF3B were upregulated in CRC tumor tissues (**Figures 10A–C**). In addition, our results showed that overexpression of NOL3 and UPF3B was associated with poor prognosis of CRC patients (**Figures 10D,E**). These results indicated that NOL3 and UPF3B play vital roles in predicting the prognosis of CRC patients.

DISCUSSION

Recently, numerous studies have focused on certain characteristics, such as autophagy and metabolic reprogramming, to identify gene signatures that are able to predict the mortality risk of cancer (Wan et al., 2019; Lin et al., 2020; Liu et al., 2020). In this study, we identified four RBP-related genes that were able to predict the OS of CRC patients. First, we detected 242 differentially expressed RBPs from a total of

FIGURE 7 | Relationship between the four risk genes expression and clinical features of CRC patients in The Cancer Genome Atlas (TCGA). **(A–D)** Survival curves of NOL3 **(A)**, PTRH1 **(B)**, UPF3B **(C)**, and SMAD6 **(D)**. **(E–G)** Relationships between NOL3 expression and N **(E)**, M **(F)**, and clinical stage **(G)**. **(H)** Relationships between PTRH1 expression and T. **(I,J)** Relationships between UPF3B expression and T **(I)** and clinical stage **(J)**. **(K)** Relationships between SMAD6 expression and T. (T, primary tumor; N, lymph node metastasis; M, distant metastasis).

1,542 RBPs. Then, we performed univariate and multivariable Cox regression analyses and selected four prognosis-related RBPs to construct a prognostic risk score model. Also, results showed that the prognostic risk score model accurately predicted prognosis of CRC patients. Furthermore, we built a nomogram based on independent prognostic factors (including the risk score obtained from the prognostic risk score model, age, and clinical stage). Also, the nomogram performed well in predicting the 3- and 5-year OS. We also validated the clinical value of the four risk genes and found that they were associated with tumorigenesis, lymph node metastasis, distant metastasis, clinical stage, and OS. Finally, we confirmed the vital roles of NOL3 and UPF3B in predicting prognosis of CRC patients using IHC in a clinical cohort.

Based on four prognosis-related RBPs, we constructed a prognostic risk score model to predict OS of CRC patients. Some of these genes were found to be related to tumorigenesis and progression of CRC and other malignancies. NOL3 was strongly upregulated across multiple cancers. In particular, NOL3 was found to be highly expressed in AML and was associated with poor prognosis of AML patients (Carter et al., 2011, 2019; Mak et al., 2014a,b). Overexpression of NOL3 was

also related to poor prognosis of nasopharyngeal carcinoma patients (Wu et al., 2013). Studies also found that NOL3 promoted tumorigenesis, metastasis, and chemoresistance in breast cancer, all of which contributed to worse patient prognosis (Medina-Ramirez et al., 2011). NOL3 was a direct target of miR-185 in gastric cancer (Li et al., 2014). Recent discoveries have also identified that upregulation of NOL3 was associated with worse prognosis among CRC patients (Mercier et al., 2008; Tóth et al., 2016), which is consistent with our results. Finally, NOL3 might be modulated by known cancer signaling proteins including Ras (Wu et al., 2010) and HIF-1 (Ao et al., 2012), and the lncRNA PCAT6 (Huang et al., 2019), in CRC.

SMAD6, a member of the SMAD family, negatively modulates the transforming growth factor-β signaling pathway (Jung et al., 2013). SMAD6 is predictive of patient survival in oral squamous cell carcinoma (Mangone et al., 2010). SMAD6 was found to be overexpressed in glioma, and its overexpression is associated with poor patient survival (Jiao et al., 2018). SMAD6 correlated with poor patient survival among non-small cell lung cancer, and its knockdown inhibited cell proliferation and increased apoptosis in the lung cancer cell line (Jeon et al., 2008). Our study indicated
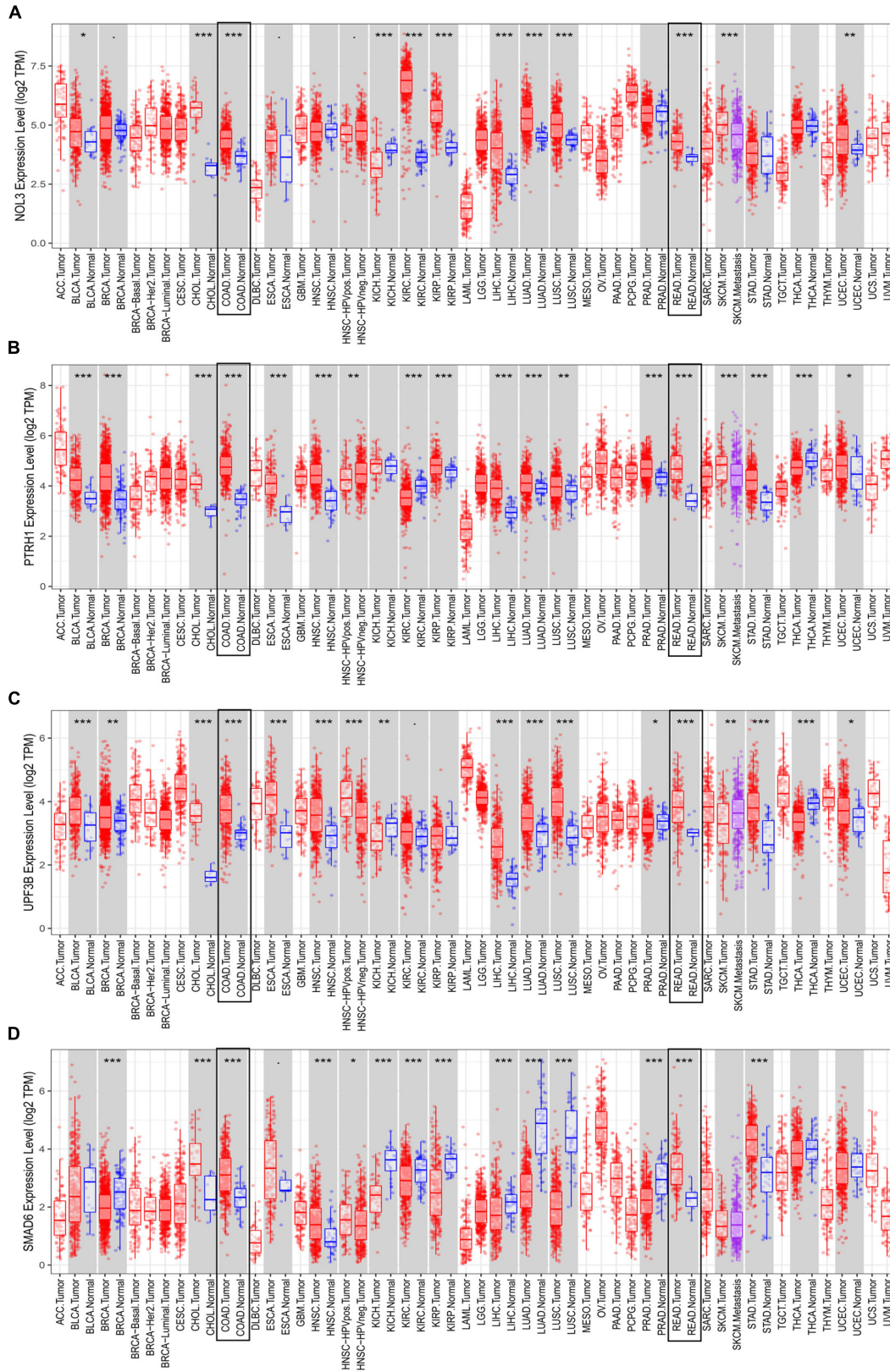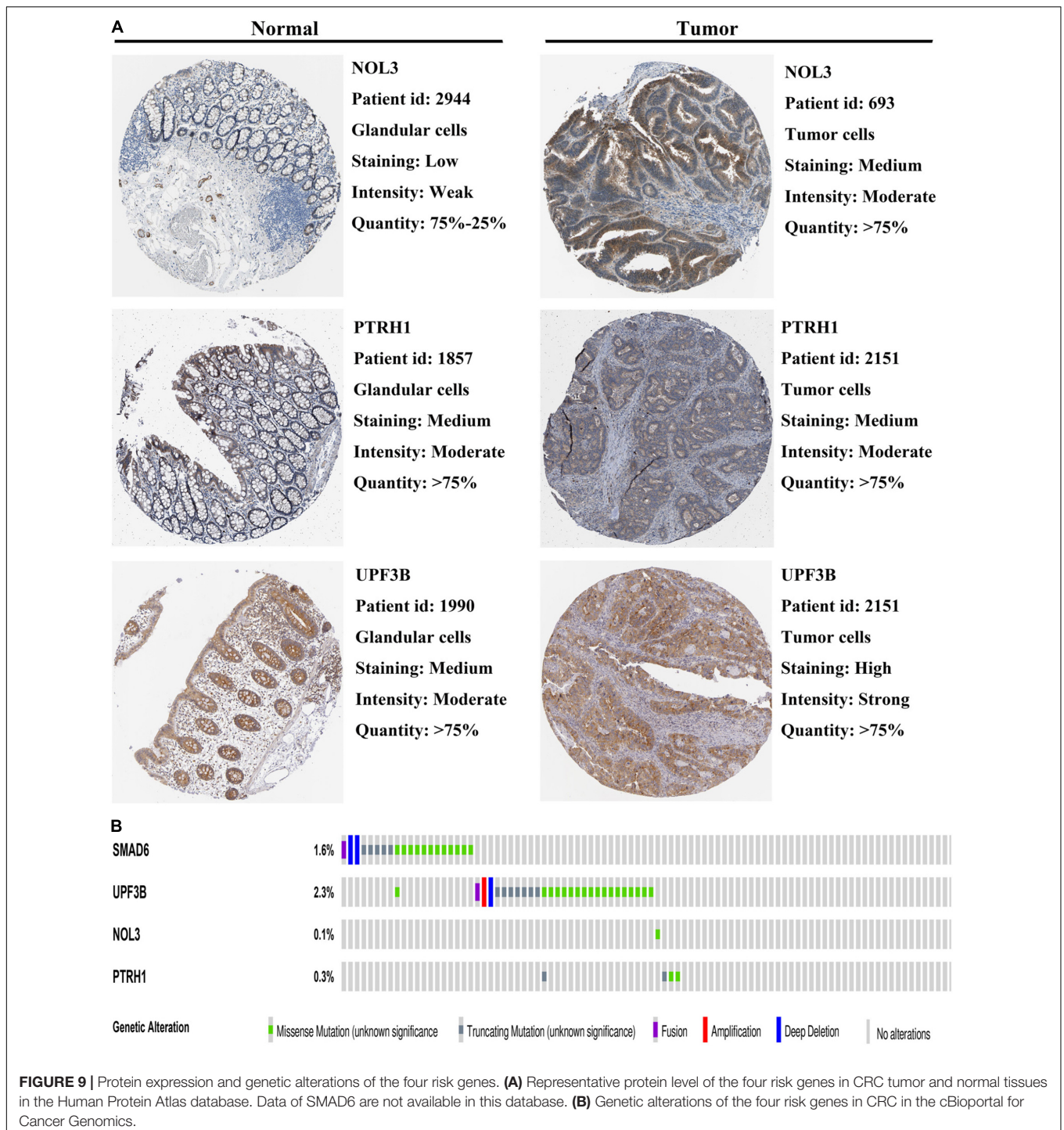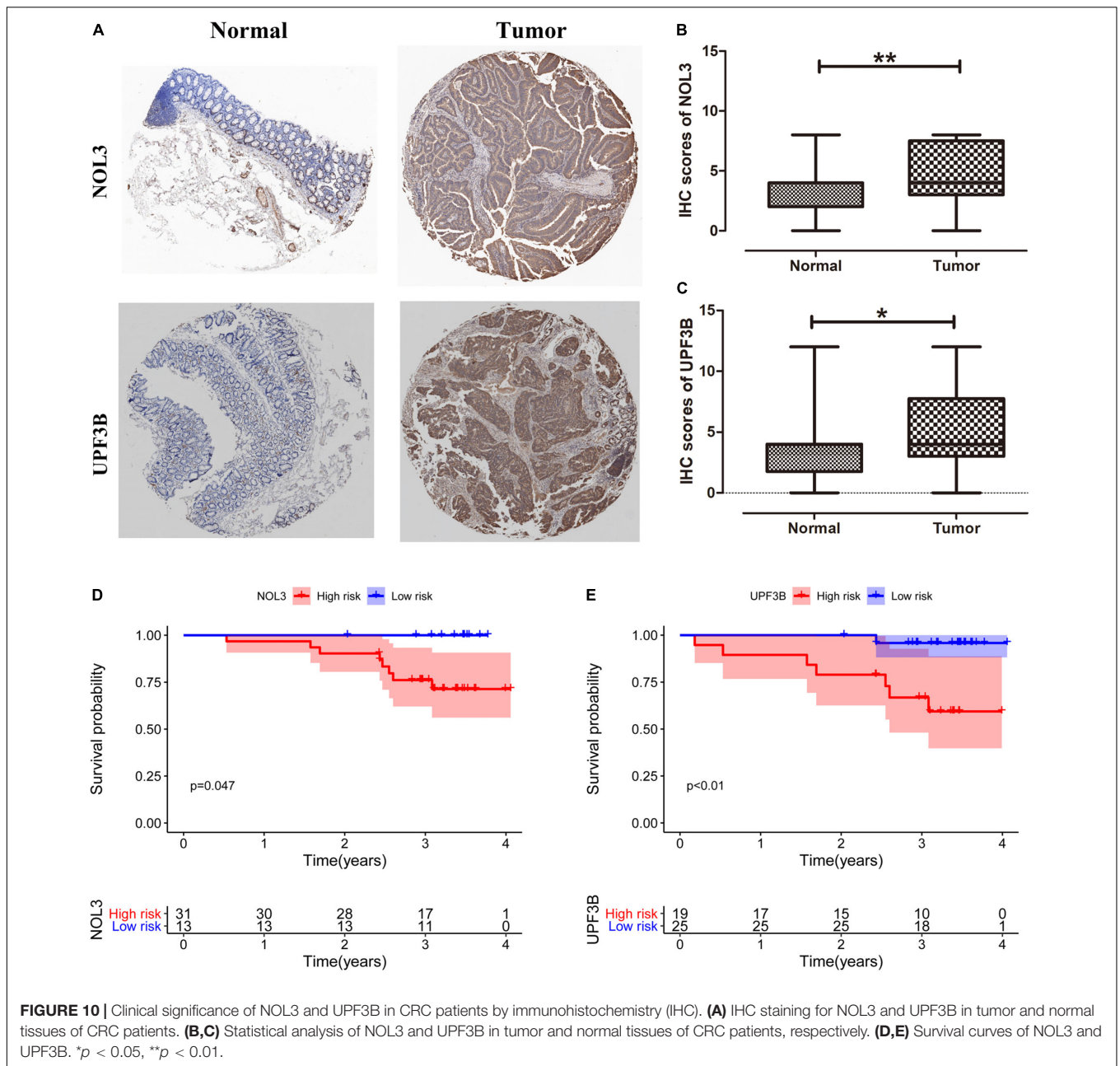
FIGURE 8 | Expression of the four risk genes in multiple cancers. Expression of **(A)** NOL3, **(B)** PTRH1, **(C)** UPF3B, and **(D)** SMAD6 in TIMER. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

FIGURE 9 | Protein expression and genetic alterations of the four risk genes. **(A)** Representative protein level of the four risk genes in CRC tumor and normal tissues in the Human Protein Atlas database. Data of SMAD6 are not available in this database. **(B)** Genetic alterations of the four risk genes in CRC in the cBioportal for Cancer Genomics.

that overexpression of SMAD6 is related to worse patient prognosis in CRC.

Our study identified that upregulated PTRH1 and UPF3B correlated with worse prognosis of CRC patients. However, we only found a few studies about these two genes in CRC and other cancers. UPF3B, a member of the UP-frameshift proteins, mediates nonsense-mediated mRNA decay (Raimondeau et al., 2018). UP-frameshift proteins include

UPF1, UPF2, UPF3A, and UPF3B (Raimondeau et al., 2018). Although UPF3B is less well-studied, we found several studies about additional UP-frameshift proteins in cancers. UPF1 regulates tumor progression via diverse mechanisms across different kinds of cancers, including CRC (Bordonaro and Lazarova, 2019), hepatocellular carcinoma (Chang et al., 2016; Zhang et al., 2017), pancreatic adenosquamous carcinoma (Liu et al., 2014), glioblastoma (Shao et al., 2019), and

FIGURE 10 | Clinical significance of NOL3 and UPF3B in CRC patients by immunohistochemistry (IHC). (A) IHC staining for NOL3 and UPF3B in tumor and normal tissues of CRC patients. (B,C) Statistical analysis of NOL3 and UPF3B in tumor and normal tissues of CRC patients, respectively. (D,E) Survival curves of NOL3 and UPF3B. $*p < 0.05$, $**p < 0.01$.

endometrial carcinoma (Xing et al., 2020). UPF3A partially contributes to the effect of calcium homeostasis endoplasmic reticulum protein (CHERP) in promoting tumorigenesis in CRC (Wang Q. et al., 2019). UPF3B and other UP-frameshift proteins can be interacted (Raimondeau et al., 2018). These data, combined with results from our study, suggest that UPF3B regulates tumor progression of CRC and may represent a potential prognostic biomarker for CRC patients. However, the mechanism of its effect in CRC requires further study.

Although our study indicates that RBPs prominently contribute to the prognosis of CRC patients, several limitations need to be pointed out. First, the clinical cohort contains fewer patients, which may lead to deviation. Additionally, the mechanisms of how these RBPs regulate the progression of CRC require further exploration.

CONCLUSION

In conclusion, we performed a comprehensive bioinformatics analysis of RBPs and identified several potential prognostic RBPs in CRC. The prognostic risk score model, including four RBPs, is an independent prognostic factor for CRC. These four RBPs are involved in tumorigenesis, progression, and prognosis of CRC. RBPs represent an alternative strategy to interfere

with tumor progression and predict the prognosis of CRC patients in the future.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Peking University People's Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ZZ, LW, SW, and ZS designed the study. ZZ, LW, QW, MZ, and BW analyzed the data. ZZ and LW wrote the manuscript. KJ, YY, SW, and ZS reviewed and edited the manuscript. All authors read and approved the final manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.580149/full#supplementary-material

Supplementary Figure 1 | Test of four RBPs-related gene model in the GSE29623 cohort. **(A)** Overall survival curve for high-risk and low-risk groups in the GSE29623 cohort. **(B–D)** Heat map of mRNA expression **(B)**, distribution of risk score **(C)**, and survival status **(D)** of patients in high-risk and low-risk groups in the GSE29623 cohort, respectively.

REFERENCES

Ao, J. E., Kuang, L. H., Zhou, Y., Zhao, R., and Yang, C. M. (2012). Hypoxia-inducible factor 1 regulated ARC expression mediated hypoxia induced inactivation of the intrinsic death pathway in p53 deficient human colon cancer cells. *Biochem. Biophys. Res., Commun.* 420, 913–917. doi: 10.1016/j.bbrc.2012.03.101

Bordonaro, M., and Lazarova, D. (2019). Amlexanox and UPF1 modulate Wnt signaling and apoptosis in HCT-116 colorectal cancer cells. *J. Cancer* 10, 287–292. doi: 10.7150/jca.28331

Carter, B. Z., Mak, P. Y., Wang, X., Tao, W., Ruvolo, V., Mak, D., et al. (2019). An ARC-Regulated IL1β/Cox-2/PGE2/β-Catenin/ARC circuit controls leukemia-microenvironment interactions and confers drug resistance in AML. *Cancer Res.* 79, 1165–1177. doi: 10.1158/0008-5472.Can-18-0921

Carter, B. Z., Qiu, Y. H., Zhang, N., Coombes, K. R., Mak, D. H., Thomas, D. A., et al. (2011). Expression of ARC (apoptosis repressor with caspase recruitment domain), an antiapoptotic protein, is strongly prognostic in AML. *Blood* 117, 780–787. doi: 10.1182/blood-2010-04-280503

Chang, L., Li, C., Guo, T., Wang, H., Ma, W., Yuan, Y., et al. (2016). The human RNA surveillance factor UPF1 regulates tumorigenesis by targeting Smad7 in hepatocellular carcinoma. *J. Exp. Clin. Cancer Res.* 35:8. doi: 10.1186/s13046-016-0286-2

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

Hentze, M. W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* 19, 327–341. doi: 10.1038/nrm.2017.130

Huang, W., Su, G., Huang, X., Zou, A., Wu, J., Yang, Y., et al. (2019). Long noncoding RNA PCAT6 inhibits colon cancer cell apoptosis by regulating anti-apoptotic protein ARC expression via EZH2. *Cell Cycle* 18, 69–83. doi: 10.1080/15384101.2018.1558872

Jeon, H. S., Dracheva, T., Yang, S. H., Meerzaman, D., Fukuoka, J., Shakoori, A., et al. (2008). SMAD6 contributes to patient survival in non-small cell lung cancer and its knockdown reestablishes TGF-beta homeostasis in lung cancer cells. *Cancer Res.* 68, 9686–9692. doi: 10.1158/0008-5472.Can-08-1083

Jiao, J., Zhang, R., Li, Z., Yin, Y., Fang, X., Ding, X., et al. (2018). Nuclear Smad6 promotes gliomagenesis by negatively regulating PIAS3-mediated STAT3 inhibition. *Nat. Commun.* 9:2504. doi: 10.1038/s41467-018-04936-9

Jung, S. M., Lee, J. H., Park, J., Oh, Y. S., Lee, S. K., Park, J. S., et al. (2013). Smad6 inhibits non-canonical TGF-β1 signalling by recruiting the deubiquitinase A20 to TRAF6. *Nat. Commun.* 4:2562. doi: 10.1038/ncomms3562

Kow, A. W. C. (2019). Hepatic metastasis from colorectal cancer. *J. Gastrointest. Oncol.* 10, 1274–1298. doi: 10.21037/jgo.2019.08.06

Li, Q., Wang, J. X., He, Y. Q., Feng, C., Zhang, X. J., Sheng, J. Q., et al. (2014). MicroRNA-185 regulates chemotherapeutic sensitivity in gastric cancer by targeting apoptosis repressor with caspase recruitment domain. *Cell Death Dis.* 5:e1197. doi: 10.1038/cddis.2014.148

Li, W., Gao, L. N., Song, P. P., and You, C. G. (2020). Development and validation of a RNA binding protein-associated prognostic model for lung adenocarcinoma. *Aging* 12, 3558–3573. doi: 10.18632/aging.102828

Li, Y., McGrail, D. J., Xu, J., Li, J., Liu, N. N., Sun, M., et al. (2019a). MERIT: systematic analysis and characterization of mutational effect on RNA interactome topology. *Hepatology* 70, 532–546. doi: 10.1002/hep.30242

Li, Y., Xiao, J., Bai, J., Tian, Y., Qu, Y., Chen, X., et al. (2019b). Molecular characterization and clinical relevance of m(6)A regulators across 33 cancer types. *Mol. Cancer* 18:137. doi: 10.1186/s12943-019-1066-3

Lin, Q. G., Liu, W., Mo, Y. Z., Han, J., Guo, Z. X., Zheng, W., et al. (2020). Development of prognostic index based on autophagy-related genes analysis in breast cancer. *Aging* 12, 1366–1376. doi: 10.18632/aging.102687

Liu, C., Karam, R., Zhou, Y., Su, F., Ji, Y., Li, G., et al. (2014). The UPF1 RNA surveillance gene is commonly mutated in pancreatic adenosquamous carcinoma. *Nat. Med.* 20, 596–598. doi: 10.1038/nm.3548

Liu, G. M., Xie, W. X., Zhang, C. Y., and Xu, J. W. (2020). Identification of a four-gene metabolic signature predicting overall survival for hepatocellular carcinoma. *J. Cell Physiol.* 235, 1624–1636. doi: 10.1002/jcp.29081

Mak, P. Y., Mak, D. H., Mu, H., Shi, Y., Ruvolo, P., Ruvolo, V., et al. (2014a). Apoptosis repressor with caspase recruitment domain is regulated by MAPK/PI3K and confers drug resistance and survival advantage to AML. *Apoptosis* 19, 698–707. doi: 10.1007/s10495-013-0954-z

Mak, P. Y., Mak, D. H., Ruvolo, V., Jacamo, R., Kornblau, S. M., Kantarjian, H., et al. (2014b). Apoptosis repressor with caspase recruitment domain modulates second mitochondrial-derived activator of caspases mimetic-induced cell death through BIRC2/MAP3K14 signalling in acute myeloid leukaemia. *Br. J. Haematol.* 167, 376–384. doi: 10.1111/bjh.13054

Mangone, F. R., Walder, F., Maistro, S., Pasini, F. S., Lehn, C. N., Carvalho, M. B., et al. (2010). Smad2 and Smad6 as predictors of overall survival in oral

squamous cell carcinoma patients. *Mol. Cancer* 9:106. doi: 10.1186/1476-4598-9-106

Medina-Ramirez, C. M., Goswami, S., Smirnova, T., Bamira, D., Benson, B., Ferrick, N., et al. (2011). Apoptosis inhibitor ARC promotes breast tumorigenesis, metastasis, and chemoresistance. *Cancer Res.* 71, 7705–7715. doi: 10.1158/0008-5472.Can-11-2192

Mercier, I., Vuolo, M., Jasmin, J. F., Medina, C. M., Williams, M., Mariadason, J. M., et al. (2008). ARC (apoptosis repressor with caspase recruitment domain) is a novel marker of human colon cancer. *Cell Cycle* 7, 1640–1647. doi: 10.4161/cc.7.11.5979

Mitchell, S. F., and Parker, R. (2014). Principles and properties of eukaryotic mRNPs. *Mol. Cell* 54, 547–558. doi: 10.1016/j.molcel.2014.04.033

Mongroo, P. S., Noubissi, F. K., Cuatrecasas, M., Kalabis, J., King, C. E., Johnstone, C. N., et al. (2011). IMP-1 displays cross-talk with K-Ras and modulates colon cancer cell survival through the novel proapoptotic protein CYFIP2. *Cancer Res.* 71, 2172–2182. doi: 10.1158/0008-5472.Can-10-3295

Moore, M. J. (2005). From birth to death: the complex lives of eukaryotic mRNAs. *Science* 309, 1514–1518. doi: 10.1126/science.1111443

Neelamraju, Y., Hashemikhabir, S., and Janga, S. C. (2015). The human RBPome: from genes and proteins to human disease. *J. Proteomics* 127(Pt A), 61–70. doi: 10.1016/j.jprot.2015.04.031

Pérez-Guijarro, E., Karras, P., Cifdaloz, M., Martínez-Herranz, R., Cañón, E., Graña, O., et al. (2016). Lineage-specific roles of the cytoplasmic polyadenylation factor CPEB4 in the regulation of melanoma drivers. *Nat. Commun.* 7:13418. doi: 10.1038/ncomms13418

Perron, G., Jandaghi, P., Solanki, S., Safisamghabadi, M., Storoz, C., Karimzadeh, M., et al. (2018). A general framework for interrogation of mRNA stability programs identifies RNA-binding proteins that govern cancer transcriptomes. *Cell Rep.* 23, 1639–1650. doi: 10.1016/j.celrep.2018.04.031

Raimondeau, E., Bufton, J. C., and Schaffitzel, C. (2018). New insights into the interplay between the translation machinery and nonsense-mediated mRNA decay factors. *Biochem. Soc. Trans.* 46, 503–512. doi: 10.1042/bst20170427

Roth, A. D., Tejpar, S., Delorenzi, M., Yan, P., Fiocca, R., Klingbiel, D., et al. (2010). Prognostic role of KRAS and BRAF in stage II and III resected colon cancer: results of the translational study on the PETACC-3, EORTC 40993, SAKK 60-00 trial. *J. Clin. Oncol.* 28, 466–474. doi: 10.1200/jco.2009.23.3452

Saha, S., Murmu, K. C., Biswas, M., Chakraborty, S., Basu, J., Madhulika, S., et al. (2019). Transcriptomic analysis identifies RNA binding proteins as putative regulators of myelopoiesis and leukemia. *Front. Oncol.* 9:692. doi: 10.3389/fonc.2019.00692

Shao, L., He, Q., Liu, Y., Liu, X., Zheng, J., Ma, J., et al. (2019). UPF1 regulates the malignant biological behaviors of glioblastoma cells via enhancing the stability of Linc-00313. *Cell Death Dis.* 10:629. doi: 10.1038/s41419-019-1845-1

Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics, 2020. *CA Cancer J. Clin.* 70, 7–30. doi: 10.3322/caac.21590

Tóth, C., Meinrath, J., Herpel, E., Derix, J., Fries, J., Buettner, R., et al. (2016). Expression of the apoptosis repressor with caspase recruitment domain (ARC) in liver metastasis of colorectal cancer and its correlation with DNA mismatch repair proteins and p53. *J. Cancer Res. Clin. Oncol.* 142, 927–935. doi: 10.1007/s00432-015-2102-3

Van Cutsem, E., Cervantes, A., Nordlinger, B., and Arnold, D. (2014). Metastatic colorectal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 25(Suppl. 3), iii1–iii9. doi: 10.1093/annonc/mdu260

Vázquez-Chantada, M., Fernández-Ramos, D., Embade, N., Martínez-Lopez, N., Varela-Rey, M., Woodhoo, A., et al. (2010). HuR/methyl-HuR and AUF1 regulate the MAT expressed during liver proliferation, differentiation, and carcinogenesis. *Gastroenterology* 138, 1943–1953. doi: 10.1053/j.gastro.2010.01.032

Wan, B., Liu, B., Yu, G., Huang, Y., and Lv, C. (2019). Differentially expressed autophagy-related genes are potential prognostic and diagnostic biomarkers in clear-cell renal cell carcinoma. *Aging* 11, 9025–9042. doi: 10.18632/aging.102368

Wang, K., Li, L., Fu, L., Yuan, Y., Dai, H., Zhu, T., et al. (2019). Integrated bioinformatics analysis the function of RNA binding proteins (RBPs) and their prognostic value in breast cancer. *Front, Pharmacol.* 10:140. doi: 10.3389/fphar.2019.00140

Wang, Q., Wang, Y., Liu, Y., Zhang, C., Luo, Y., Guo, R., et al. (2019). U2-related proteins CHERP and SR140 contribute to colorectal tumorigenesis via alternative splicing regulation. *Int. J. Cancer* 145, 2728–2739. doi: 10.1002/ijc.32331

Wang, Z., Tang, W., Yuan, J., Qiang, B., Han, W., and Peng, X. (2020). Integrated analysis of RNA-Binding proteins in glioma. *Cancers* 12:892. doi: 10.3390/cancers12040892

Wu, L., Nam, Y. J., Kung, G., Crow, M. T., and Kitsis, R. N. (2010). Induction of the apoptosis inhibitor ARC by Ras in human cancers. *J. Biol. Chem.* 285, 19235–19245. doi: 10.1074/jbc.M110.114892

Wu, P., Tang, Y., He, J., Qi, L., Jiang, W., and Zhao, S. (2013). ARC is highly expressed in nasopharyngeal carcinoma and confers X-radiation and cisplatin resistance. *Oncol. Rep.* 30, 1807–1813. doi: 10.3892/or.2013.2622

Xing, T. R., Chen, P., Wu, J. M., Gao, L. L., Yang, W., Cheng, Y., et al. (2020). UPF1 participates in the progression of endometrial cancer by inhibiting the expression of lncRNA PVT1. *Onco Targets Ther.* 13, 2103–2114. doi: 10.2147/ott.S233149

Xu, Y., Gao, X. D., Lee, J. H., Huang, H., Tan, H., Ahn, J., et al. (2014). Cell type-restricted activity of hnRNPM promotes breast cancer metastasis via regulating alternative splicing. *Genes Dev.* 28, 1191–1203. doi: 10.1101/gad.241968.114

Zhang, H., You, Y., and Zhu, Z. (2017). The human RNA surveillance factor Up-frameshift 1 inhibits hepatic cancer progression by targeting MRP2/ABCC2. *Biomed. Pharmacother.* 92, 365–372. doi: 10.1016/j.biopha.2017.05.090

Zhang, Z., Zheng, X., Li, J., Duan, J., Cui, L., Yang, L., et al. (2019). Overexpression of UBR5 promotes tumor growth in gallbladder cancer via PTEN/PI3K/Akt signal pathway. *J. Cell Biochem.* 120, 11517–11524. doi: 10.1002/jcb.28431