Original Article

Molecular characterization of colorectal cancer: A five-gene prognostic signature based on RNA-binding proteins

Qiankun Liang[1], Xiaojuan Du[1], Lanfang Mao[1,2], Guopan Wang[3]

[1]Gansu University of Chinese Medicine, [2]Affiliated Hospital of Gansu University of Chinese Medicine, Lanzhou, [3]Zunyi Medical University, Zunyi, China

Abstract

Background: Colorectal cancer (CRC) is one of the most common cancers worldwide. RNA-binding proteins (RBPs) regulate essential biological processes and play essential roles in a variety of cancers. The present study screened differentially expressed RBPs, analyzed their function and constructed a prognostic model to predict the overall survival of patients with CRC.

Methods: We downloaded CRC RNA-sequencing data from the Cancer Genome Atlas (TCGA) portal and screened differentially expressed RBPs. Then, functional analyses of these genes were performed, and a risk model was established by multivariate Cox regression.

Results: We obtained 132 differentially expressed RBPs, including 66 upregulated and 66 downregulated RBPs. Functional analysis revealed that these genes were significantly enriched in RNA processing, modification and binding, ribosome biogenesis, post-transcriptional regulation, ribonuclease and nuclease activity. Additionally, some RBPs were significantly related to interferon (IFN)-alpha and IFN-beta biosynthetic processes and the Toll-like receptor signaling pathway. A prognostic model was constructed and included insulin like growth factor 2 messenger ribonucleic acid binding protein 3 (IGF2BP3), poly (A) binding protein cytoplasmic 1 like (PABPC1L), peroxisome proliferator activated receptor gamma coactivator 1 alpha (PPARGC1A), peptidyl- transfer ribonucleic acid hydrolase 1 homolog (PTRH1) and tudor domain containing 7 (TDRD7). The model is an independent risk factor for clinicopathological characteristics.

Conclusion: Our study provided novel insights into the pathogenesis of CRC and constructed a prognostic gene model, which may be helpful for determining the prognosis of CRC.

Keywords: Colorectal cancer, prognostic signature, RNA-binding proteins, survival rate

Address for correspondence: Dr. Qiankun Liang, Gansu University of Chinese Medicine, Lanzhou 730020, China.
E-mail: 1531271264@qq.com

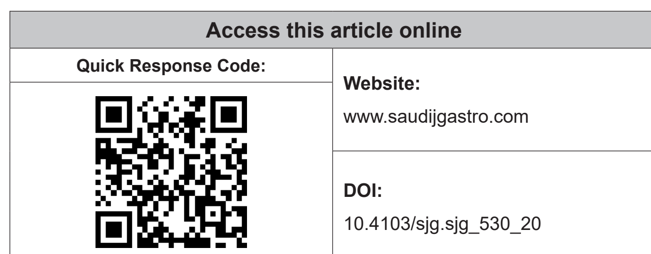
Submitted: 24-Sep-2020 **Revised:** 27-Mar-2021 **Accepted:** 03-Apr-2021 **Published:** 23-Jun-2021

INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer with more than 2.2 million new cases and 1.1 million deaths estimated to occur by 2030.[1] It is predicted that the incidence of CRC in young individuals will increase by as much as 90% by 2030.[2] There are no obvious symptoms in the early stage of CRC, but changes in stool habits, emaciation, hematochezia, anemia, colonic fistula or intestinal perforation can be found during the

Access this article online

Quick Response Code:



Website:
www.saudijgastro.com

DOI:
10.4103/sjg.sjg_530_20

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

How to cite this article: Liang Q, Du X, Mao L, Wang G. Molecular characterization of colorectal cancer: A five-gene prognostic signature based on RNA-binding proteins. Saudi J Gastroenterol 2021;27;223-33.

progression and distant metastasis of CRC.[3] Nearly 5% of CRC cases are genetic syndromes, including familial adenomatous polyposis (FAP), hereditary nonpolyposis colorectal cancer (HNPCC), or mutY deoxyribonucleic acid glycosylase-associated polyposis (MAP), and 15-20% of the cases have a family history of hereditary diseases. However, the majority of cases are sporadic diseases without a family history of CRC, which is closely associated with microsatellite instability and lifestyle factors.[4-6] Tumor recurrence is the primary barrier to improving overall survival.[7] To reduce the recurrence rate and mortality of patients with CRC, it is crucial to improve the surveillance ability following surgery. The traditional prognostic criterion for CRC is the American Joint Committee on Cancer/Union Internationale Contre le Cancer (AJCC/UICC) tumor-node-metastasis (TNM) staging system, but the system provides limited prognostic information and cannot predict the response to therapy due to significantly different clinical outcomes possibly emerging in the same TNM stages.[1] Therefore, some new methods to classify cancer, including molecular pathways, mutation status, cell origin, gene expression-based stratification and the TNM immune staging system, have been used.[1,8,9] In recent years, biomarkers have also been applied for early diagnosis and prognosis, and therefore, it is vital to find new biomarkers that reflect the occurrence, progression and prognosis of CRC.

RNA-binding proteins (RBPs) play essential roles in post-transcriptional events and regulate essential biological processes, including RNA localization, stability, transport, editing, degradation and translation, by forming ribonucleoprotein (RNP) complexes with target RNA.[10] There are 1,542 RBPs that have been experimentally validated, accounting for approximately 7.5% of all protein-coding genes.[11] RBPs are ubiquitous in cells and are evolutionarily conserved so that they can play a central role in essential biological functions.[11] Because the regulation of RBPs to RNAs occurs in rapid and efficient ways, small changes in RBPs can alter gene expression and cell fate.[12,13]

Abnormalities in RBPs can induce diseases, including cancer, and correlate with patient prognosis.[10,14-16] Previous studies have found that several RBPs are significantly related to CRC, such as lin-28 homolog B (LIN28B), human antigen R (HuR) and ribonucleic acid binding motif protein 3 (RBM3). LIN28B is overexpressed in nearly 30% of CRC, which indicates worse survival and recurrence.[17,18] The expression of HuR is significantly enhanced in colon tumors by regulating target transcribed mRNAs and is vital for neoplastic transformation and cancer development.[19] Over-expressed RBM3 in colorectal cancer cells increases proliferation and engenders chemotherapy resistance.[20] RBPs are closely related to CRC, and therefore, it is useful to perform a systematic analysis of RBPs in colon tumors.

With the development of high-throughput sequencing technology, a vast number of data sets have been generated and conserved in public portals, such as TCGA and Gene Expression Omnibus (GEO). However, the mining of these data is limited as it requires professional software and advanced expertise for researchers to process and analyze.[21] In the present study, we downloaded CRC data from the TCGA database and used bioinformatics to screen differentially expressed RBPs in tumor and normal tissue and further analyze their function in CRC. We also constructed a prognostic model to predict the overall survival of patients with CRC, which may have prognostic value in the clinic in future.

MATERIALS AND METHODS

Data acquisition and analysis

The corresponding RNA-sequencing data containing 473 colon tumor samples and 41 normal colon samples were downloaded from the Genomic Data Commons data portal (*https://portal.gdc.cancer.gov/*). The R package "edgeR" was used to normalize the above data and analyze the expression of each RBP between tumor and normal samples. A false discovery rate (FDR) <0.05 and $|\log_2$ fold change (FC)$| \geq 1$ were set as filter conditions for differentially expressed genes (DEGs). The Gene Expression Omnibus series GSE17538 dataset was obtained from the Gene Expression Omnibus (GEO) database (*https://www.ncbi.nlm.nih.gov/geo/*) and used as a validation cohort.

Gene Ontology, Kyoto Encyclopedia of Genes and Genomes analysis, and protein-protein interaction analysis

The differentially expressed RBPs were uploaded to the Database for Annotation Visualization and Integrated Discovery (DAVID) version 6.7 for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. Both $P < 0.05$ and FDR < 0.05 were considered statistically significant. GO analysis included biological processes, cellular components and molecular functions. Subsequently, the STRING online tool (*http://www.string-db.org/*) was used to investigate the protein-protein interactions (PPIs) of differentially expressed RBPs. Visualization of the interaction network was performed by Cytoscape 3.8.0 software. Module analyses were performed by Molecular

Complex Detection (MCODE). Modules with scores and node numbers greater than 5 were selected as key modules. The DAVID online tool and R package "ggplot2" were used to analyze and visualize GO terms and KEGG pathways of hub genes and key module genes.

Establishment of the prognostic gene signature for predicting the survival rate

Multivariate Cox regression analysis was performed to construct a prognostic gene signature. The result was shown as risk score = (Coefficient $_{mRNA1}$ × expression of $mRNA_1$) + (Coefficient $_{mRNA2}$ × expression of $mRNA_2$) + ... + (Coefficient $_{mRNAn}$ × expression of $mRNA_n$). The R packages "survival" and "survminer" were used to draw the Kaplan–Meier survival curve. To test the independence of the prognostic gene signature of other clinical parameters in TCGA, Cox regression analyses of clinical parameters were performed.

Validation of the gene change

The expression of prognostic genes was validated at the mRNA level by the Tumor Immune Estimation Resource database (TIMER) database (*https://cistrome.shinyapps.io/timer/*) and at the protein level by the Human Protein Atlas database (HPA) (*http://www.proteinatlas.org*). Additionally, genetic alterations of prognostic genes were determined by the cBioPortal online tool (*http://www.cbioportal.org/*).

Statistical analysis

In this study, we used SPSS 19.0 (IBM, Armonk, New York) and R software (R Foundation for Statistical Computing, Vienna, Austria) to process and analyze the data. For the Kaplan–Meier estimates, a significant difference in two-group survival curves was assessed by a log-rank test. $P < 0.05$ was considered statistically significant.

RESULTS

Identification of differentially expressed RBPs

To compare the differentially expressed RBPs (1,542 selected) between 41 normal colon samples and 473 tumor colon samples from the TCGA database, we used an R package to analyze the data. We obtained 132 differentially expressed RBPs ($P < 0.05$, $|\log_2 FC| \geq 1.0$), including 66 upregulated RBPs and 66 downregulated RBPs [Supplementary Table S1]. We also constructed an expression heat map for all differentially expressed RBPs [Figure 1].

Gene Ontology and KEGG pathway analysis of differentially expressed RBPs

To explore the functions and mechanisms of the above RBPs, the upregulated and downregulated RBPs were uploaded to DAVID for GO and KEGG pathway analysis, respectively. As shown in Table 1, for biological process (BP) analysis, upregulated RBPs were significantly enriched in RNA processing (ncRNA metabolic process, ncRNA processing, rRNA processing, rRNA metabolic process, tRNA processing and tRNA metabolic process), ribosome biogenesis, RNP complex biogenesis, RNA modification, post-transcriptional regulation of gene expression, RNA localization and regulation of translation. The downregulated RBPs were significantly enriched in the post-transcriptional regulation of gene expression, mRNA metabolic process and mRNA processing. For GO cellular component (CC) analysis [Table 1], upregulated RBPs were significantly enriched in the nuclear lumen, nucleolus, intracellular organelle lumen, membrane-enclosed lumen, RNP complex, intracellular nonmembrane-bound organelle, nuclear pore and pore complex; but the downregulated RBPs were not significantly enriched in any cellular component. The

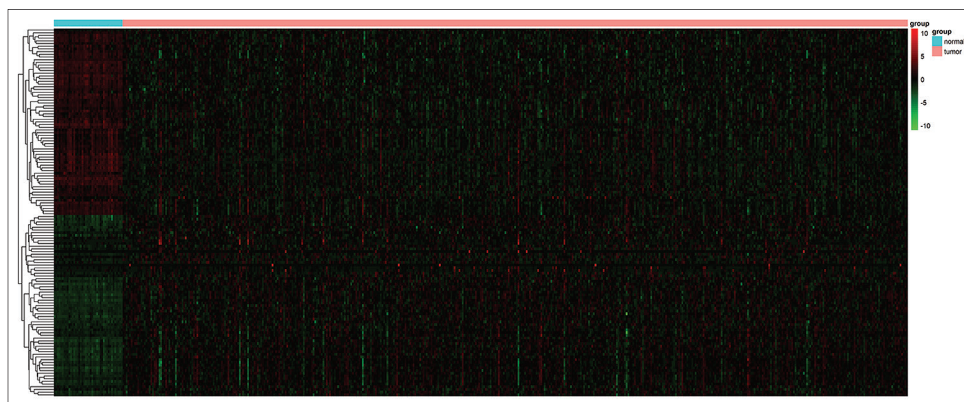


Figure 1: The differentially expressed RBPs in colorectal cancer. The columns are samples and the rows are RBPs. Green represents downregulation, while red represents upregulation

Table 1: GO enrichment and KEGG pathway analysis results of differentially upregulated genes and downregulated genes

	Category	GO term/KEGG pathway	*P*	FDR
Upregulated RBPs				
	BP	ncRNA metabolic process	5.36E-25	7.55E-22
	BP	ncRNA processing	1.15E-23	1.62E-20
	BP	RNA processing	1.38E-15	1.88E-12
	BP	ribosome biogenesis	2.15E-12	3.02E-09
	BP	ribonucleoprotein complex biogenesis	7.15E-12	1.01E-08
	BP	rRNA processing	1.22E-10	1.71E-07
	BP	rRNA metabolic process	1.79E-10	2.52E-07
	BP	tRNA processing	7.78E-10	1.10E-06
	BP	tRNA metabolic process	2.64E-08	3.72E-05
	BP	RNA modification	1.51E-06	0.002132
	BP	Post-transcriptional regulation of gene expression	2.29E-06	0.003226
	BP	RNA localization	3.38E-06	0.004756
	BP	regulation of translation	2.07E-05	0.029152
	CC	nuclear lumen	2.73E-17	3.12E-14
	CC	nucleolus	3.90E-16	5.11E-13
	CC	intracellular organelle lumen	6.27E-15	7.09E-12
	CC	organelle lumen	1.14E-14	1.30E-11
	CC	membrane-enclosed lumen	1.90E-14	2.17E-11
	CC	ribonucleoprotein complex	1.50E-13	1.71E-10
	CC	intracellular nonmembrane-bounded organelle	1.23E-12	1.40E-09
	CC	non-membrane-bounded organelle	1.23E-12	1.40E-09
	CC	nucleolar part	1.67E-06	0.001909
	CC	nuclear pore	5.64E-06	0.006438
	CC	pore complex	1.40E-05	0.015924
	MF	RNA binding	1.45E-27	1.72E-24
	MF	ribonuclease activity	1.21E-08	1.43E-05
	MF	nuclease activity	3.49E-07	4.13E-04
	KEGG	RNA degradation	0.096550	42.642106
Downregulated RBPs				
	BP	Post-transcriptional regulation of gene expression	2.05E-06	0.00294
	BP	mRNA metabolic process	8.50E-06	0.012212
	BP	mRNA processing	3.18E-05	0.045609
	MF	RNA binding	6.29E-28	7.40E-25
	MF	double-stranded RNA binding	2.46E-07	2.89E-04
	MF	single-stranded RNA binding	4.71E-06	0.005538
	KEGG	Toll-like receptor signaling pathway	0.010117	6.510642

GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; CC: cellular component; BP: biological processes; MF: molecular function; RBP: RNA-binding protein

GO molecular function (MF) analysis showed that the upregulated RBPs were significantly related to RNA binding, ribonuclease activity and nuclease activity, and downregulated RBPs were significantly related to RNA binding, including double-stranded RNA binding and single-stranded RNA binding [Table 1]. KEGG pathway results showed that the upregulated RBPs were related to RNA degradation, and downregulated RBPs were related to the Toll-like receptor signaling pathway, but neither was significant [Table 1].

Protein-protein interaction network construction

We uploaded these differentially expressed RBPs to the STRING database and constructed a coexpression network with 131 nodes and 411 edges. Then, the degree of each protein was calculated, and 10 genes were screened as hub genes according to the top eight scores of the degree, including nucleolar protein 56 (NOP56), ribosomal ribonucleic acid processing 9 (RRP9), bystin like (BYSL), dyskeratosis congenita 1 (DKC1), ribosome biogenesis regulator 1 homolog (RRS1), U3 small nucleolar ribonucleic acid-associated protein 14 homolog A (UTP14A), block of proliferation 1 (BOP1) and damage specific deoxyribonucleic acid binding protein 1 and cullin 4 associated factor 13 (DCAF13) [Figure 2a]. We further selected three possible key modules with the MCODE plug-in with scores and node numbers greater than 5 [Figure 2b-2d]. Then, GO and KEGG pathway analyses of 10 hub genes and key module genes were performed. As Figure 3 shows, for biological processes, these genes were significantly enriched in the ncRNA and rRNA processes, ribosome complex biogenesis, and regulation of IFN-alpha and IFN-beta biosynthetic processes. For cellular components, hub genes and selected module genes were significantly enriched in the nuclear lumen, nucleolus, intracellular organelle lumen, membrane-enclosed lumen and intracellular nonmembrane-bound organelle. For MF, these genes were significantly related to RNA binding and

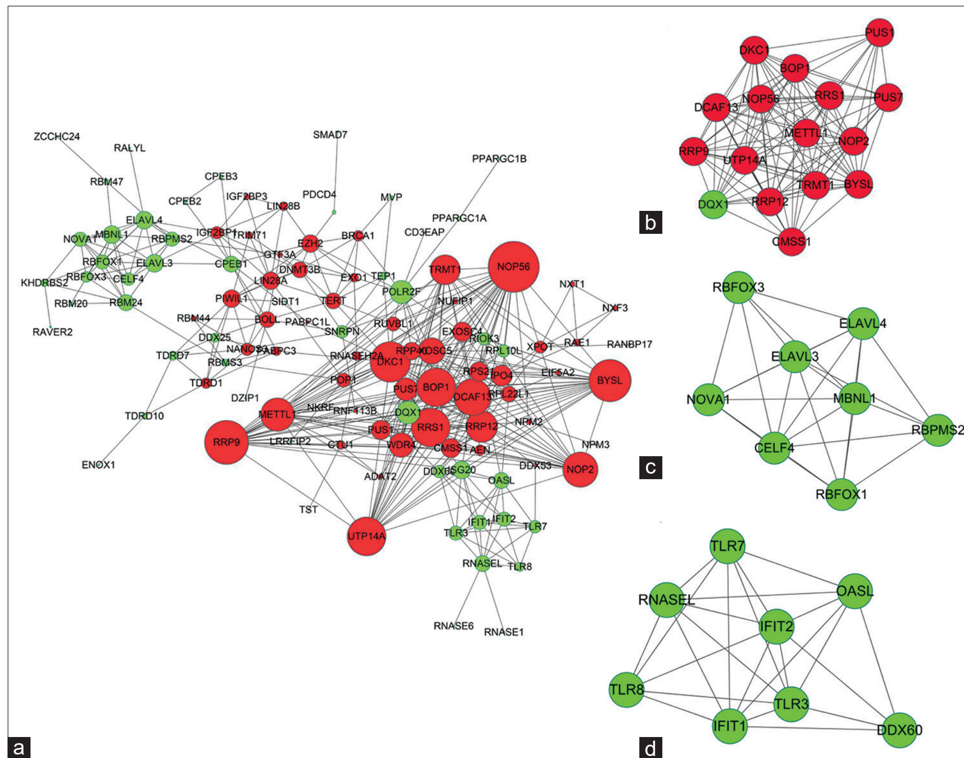


Figure 2: Establishment of a protein-protein interaction (PPI) network and module analyses. (a) PPI network of RBPs; (b-d) Elected necessary modules with MCODE score >5 and nodes >5

double-stranded RNA binding. For the KEGG pathway, these genes were significantly related to the Toll-like receptor signaling pathway.

Establishment and validation of the five-gene signature predicting the survival rate

To predict the survival rate of patients with colon cancer, we constructed a risk score model based on multiple RBP expression signatures. The genes included in the model were insulin like growth factor 2 messenger ribonucleic acid binding protein 3 (IGF2BP3), poly (A) binding protein cytoplasmic 1 like (PABPC1L), peroxisome proliferator activated receptor gamma coactivator 1 alpha (PPARGC1A), peptidyl- transfer ribonucleic acid hydrolase 1 homolog (PTRH1) and tudor domain containing 7 (TDRD7). The calculation formula of the risk score is as follows: risk score = 0.05323 × expression of IGF2BP3 + 0.15438 × expression of PABPC1L − 0.16234 × expression of PPARGC1A + 0.30132 × expression of PTRH1 − 0.34485 × expression of TDRD7. Patients were divided into high-risk and low-risk groups with a cutoff of 0.99980 (the median of the risk score). The GSE17538 dataset was obtained to validate the risk score model. Each patient's risk score was calculated, and the cutoff of the high-risk and low-risk groups was − 0.63054 (the median of the risk score). The overall survival and expression heat maps of the high-risk and low-risk groups were plotted with R software [Figure 4]. Compared with the low-risk group, the overall survival was significantly poorer in the high-risk group in TCGA ($P < 0.001$) and the GEO data set ($P < 0.05$).

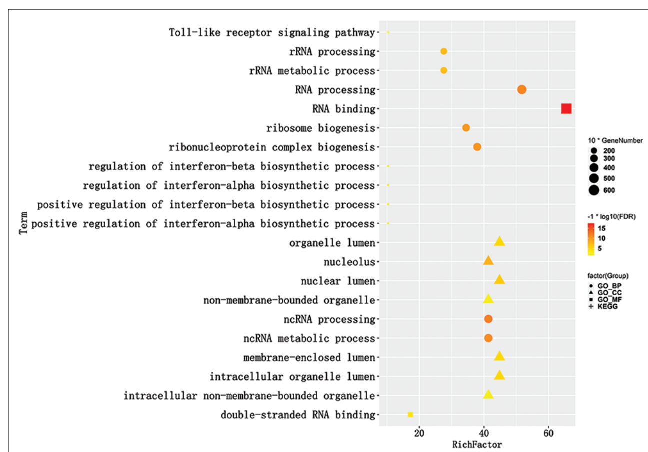


Figure 3: GO and KEGG pathway enrichment analysis of hub genes and key module genes. GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; CC: cellular component; BP: biological processes; MF: molecular function; RBP: RNA-binding protein

Independent risk factors for clinicopathological characteristics to predict overall survival

To identify independent clinical risk factors, including age, sex, TNM pathological stage, primary site (left side and right side) and our risk score model, univariate and multivariate

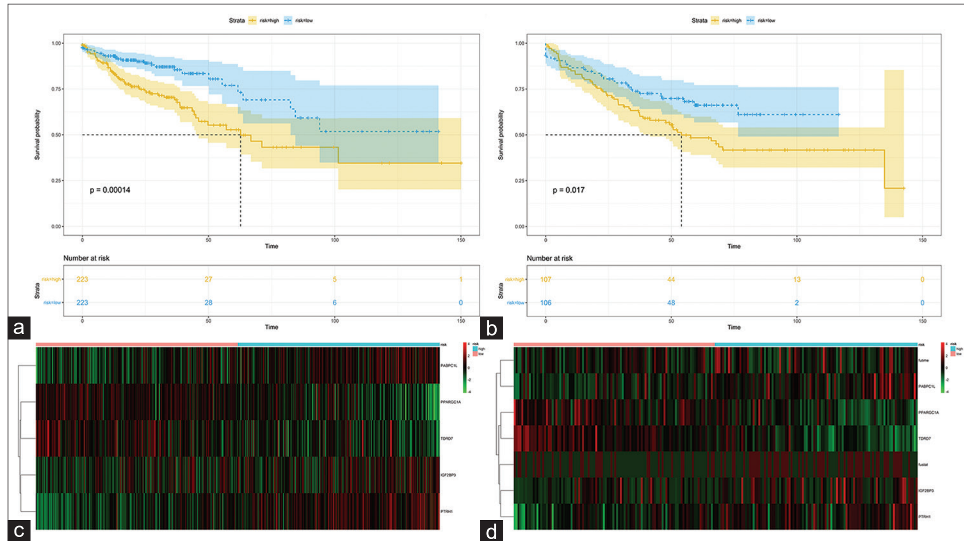


Figure 4: Prognostic risk assessment of five RBPs for patients with colorectal cancer. The assessment of the risk model for overall survival in the TCGA cohort (a) and GSE17538 cohort (b). The expression of the prognostic model genes in the TCGA cohort (c) and GSE17538 cohort (d)

Cox regression analyses were performed with the TCGA and GEO data sets. The right side of the colon includes the cecum, hepatic flexure of colon, ascending colon and transverse colon, and the left side colon extends from the splenic flexure, descending colon and sigmoid.[22] As shown in Figure 5, our risk score model and TNM pathological stage were both independent prognostic factors for the overall survival.

External validation using online databases

To validate the expression levels and investigate genetic alterations of five prognostic model genes, HPA (*http://www.proteinatlas.org/*), TIMER (*https://cistrome.shinyapps.io/timer/*) and cBioPortal database (*https://www.cbioportal.org/*) were used in this study. We selected 392 patients with colon cancer to investigate the genetic alterations of the five genes in the cBioPortal database. As shown in Figure 6a, PABPC1L possessed the most frequent genetic alterations (10%), and amplification mutation was the most common. In accordance with our results, the expression of IGF2BP3 and PTRH1 was significantly enhanced, and TDRD7 was decreased in the tumor samples compared with the normal samples in the HPA and TIMER databases [Figures 6b and 7]. There was no immunohistochemistry information about the PABPC1L and PPARGC1A proteins in HPA. The TIMER database showed that the PABPC1L mRNA level was significantly increased and that PPARGC1A was decreased in the tumor group compared with the normal group [Figure 7].

DISCUSSION

CRC is one of the most commonly diagnosed types of cancers but also the most preventable and treatable cancer.[23] Therefore, it is vital to screen for early CRC. Previous studies have shown that incidence and mortality steadily declined in countries with programmatic screening.[24,25] Colonoscopy may be the most effective strategy to screen for CRC, with 0.022 life-years gained and 1,068 CRCs prevented, but it is associated with rare and severe complications.[26,27] However, colonoscopy may be suitable as the second procedure following a positive first test.[28] Therefore, noninvasive tests are needed as supplements. Fecal deoxyribonucleic acid (DNA)

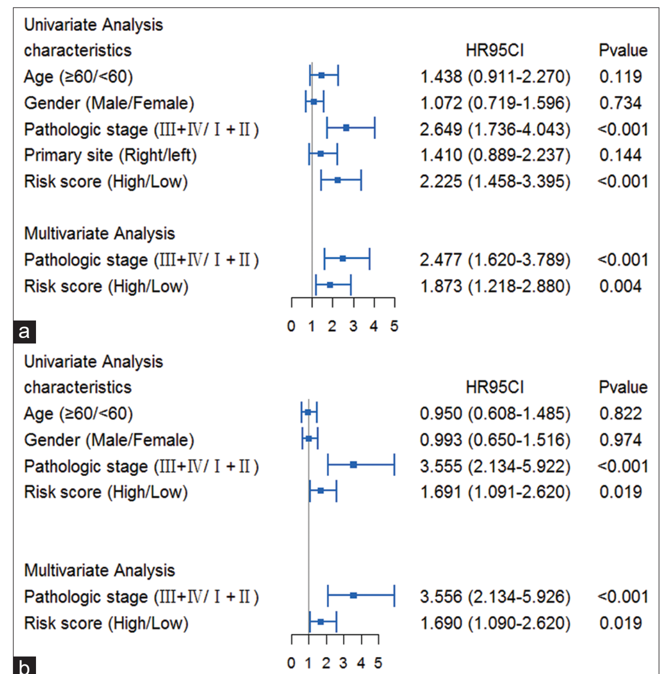


Figure 5: Forrest plot of the univariate and multivariate Cox regression analysis in colorectal cancer. Forrest plot of Cox regression analysis in TCGA database (a) and GEO database (b)

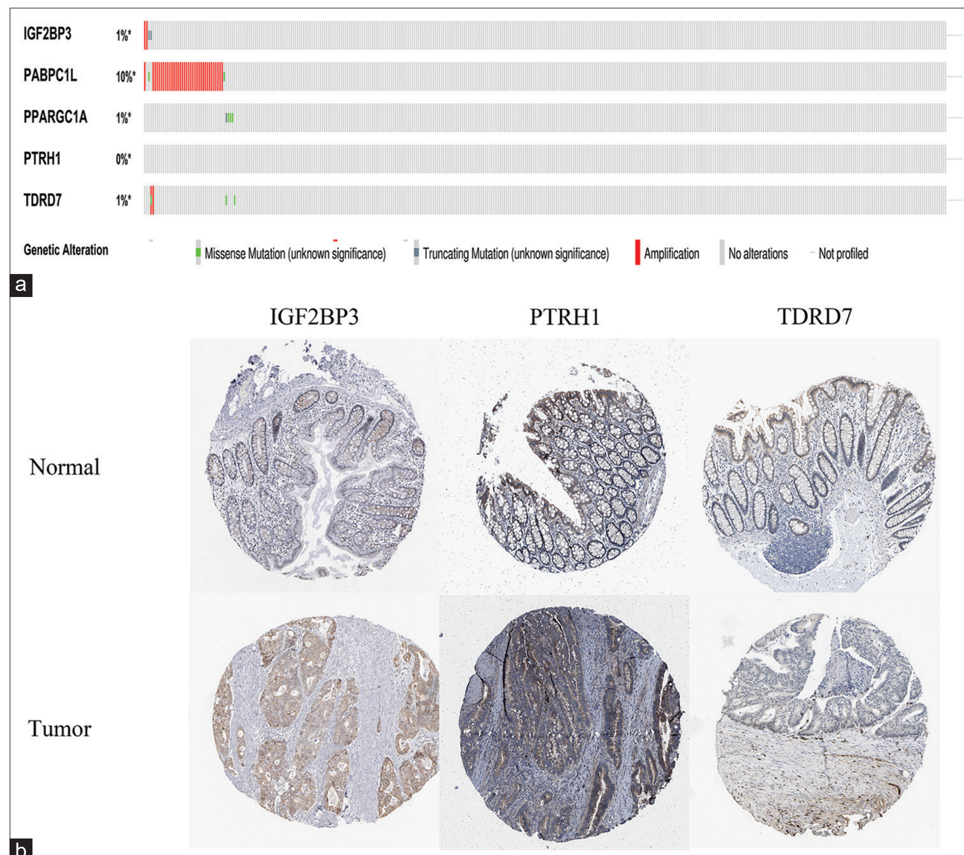


Figure 6: Expression and genetic alterations of the five model genes. (a) Genetic alterations of the genes in colorectal cancer. Data were from the cBioportal for Cancer Genomics. (b) Protein expression of genes in the colorectal cancer and normal colon tissues. Data were from the Human Protein Atlas

testing is a new non-invasive strategy for screening, and Cologuard®-targeted multiple DNA sequences, including mutant kirsten rat sarcoma viral oncogene homolog (KRAS), actin, fecal immunochemical test (FIT), aberrantly methylated bone morphogenetic protein 3 (BMP3) and the N-myc downstream regulated gene family member 4 (NDRG4) promoter regions, have been approved for CRC screening.[29]

Compared with the FIT, a noninvasive screening method, Cologuard® had a higher sensitivity (74% vs 92%) but lower specificity (95-96% vs 87-90%) for detection of CRC.[29] Therefore, finding more sensitive and specific biomarkers is essential for a fecal DNA test to diagnose CRC. Similar to molecular mutation and methylation, gene expression-based methods have also been used as tools for diagnosis, classification, treatment and prognosis.[8,30-33] In this study, we used the gene expression data of RBPs from TCGA to identify the differentially expressed RBPs and analyze their function as well as construct a multiple gene signature predicting overall survival in patients with CRC.

We obtained 132 differentially expressed RBPs, including 66 upregulated and 66 downregulated RBPs. The biological processes of these DEGs were enriched in RNA processing and RNA metabolic process, ribosome biogenesis, RNP complex biogenesis, RNA modification, post-transcriptional regulation of gene expression and RNA localization. The gene mutations of RBPs changed their function, potentially inducing cancer. For example, dicer 1, ribonuclease III (DICER1), an endoribonuclease, processes precursor RNA molecules into mature forms, which is vital for miRNA and RNA interference biogenesis pathways.[34] The mutation of DICER1 impairs miRNA-mediated gene suppression, which can induce tumor susceptibility.[35] Post-transcriptional modifications frequently occur in RNA-binding elements within RBPs, which can change RBP binding properties, function and subcellular localization, and therefore, may be among the most critical mechanisms of RBP dysfunction in cancer.[36]

For GO CC analysis, significantly differentially expressed RBP enrichment occurred in the nuclear lumen, nucleolus, intracellular organelle lumen, membrane-enclosed lumen,

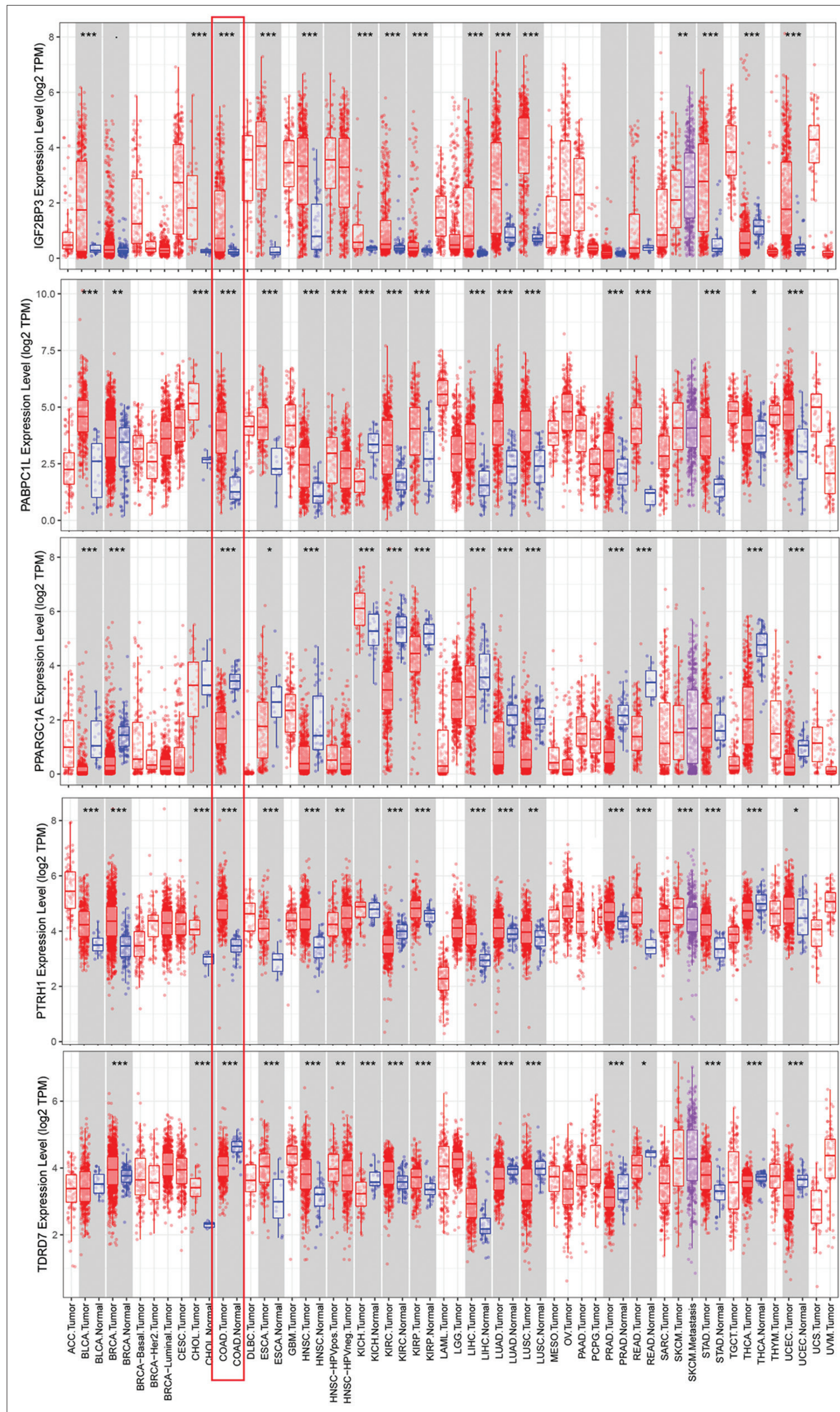


Figure 7: The expression of the model genes in cancers. Data extracted from the TIMER database

RNP complex, intracellular non-membrane-bound organelle, nuclear pore and pore complex. RBPs are ubiquitous in cells, and therefore, regulate a variety of biological functions. Nuclear maturation of transcripts, nuclear export of RNA, and RNA processing require nuclear RBPs.[37] In addition, several small nuclear

ribonucleoproteins (U1, U2, U4, U5, and U6 snRNP) form a spliceosome to participate in pre-mRNA splicing, and abnormalities in RNA splicing frequently occur at the origin of many disorders, including cancer.[38,39] The RNP mutation regulates p53 expression, which induces many diseases, including cancer.[40] Furthermore, heterogeneous nuclear RNP D-like is aberrantly expressed in CRC, which promotes the growth of these cells by activating cell cycle progression.[41] RBPs can mediate organelle-coupled translation as post-transcriptional regulators.[42]

The GO MF analysis showed that the differentially expressed RBPs were significantly related to RNA binding, ribonuclease activity and nuclease activity. RBPs can bind targeted RNAs by RNA-binding domains composed of 60-100 amino acids. As stated above, ribonuclease processes precursor RNA molecules into mature forms, and mutation of special ribonucleases, such as DICER1, induces cancer.[35]

We screened eight hub genes, including NOP56, RRP9, BYSL, DKC1, RRS1, UTP14A, BOP1 and DCAF13, by constructing a protein-protein interaction (PPI) network and calculating the degree score. NOP56 is required for rRNA methylation, and abnormalities in NOP56 induce altered ribosome biogenesis and cell cycle progression.[43] Additionally, NOP56 is hyperactivated and essential for Myc-induced cell transformation in Burkitt's lymphoma.[44] Deacetylation of RRP9 is a prerequisite for pre-rRNA processing,[45] but there is no definite report about the relevance between RRP9 and cancer. BYSL participates in ribosomal processing of 18S rRNA and is significantly enhanced in hepatocellular carcinoma specimens and is essential for nucleogenesis in cancer cell proliferation.[46] DKC1 is remarkably increased in CRC samples, and DKC1 enhances angiogenesis by promoting hypoxia inducible factor-1 (HIF-1α) transcription and facilitates metastasis in CRC.[47] RRS1 is necessary for ribosome biogenesis and is significantly increased in CRC samples to promote cancer development.[48] UTP14A is essential for 18S rRNA processing and promotes p53 degradation by binding p53.[49] Nucleolar UTP14A is significantly higher in CRC samples than in normal samples.[50] BOP1 is a direct Wnt/β-catenin target gene and can induce epithelial-mesenchymal transition (EMT), cell migration and experimental metastasis of CRC cells.[51] DCAF13 is upregulated in breast cancer, lung cancer and hepatocellular carcinoma and indicates poor overall survival.[52-54] As mentioned above, DKC1, RRS1, UTP14A and BOP1 were related to CRC, which was reported in previous studies, but there has been no detailed study about the relationship between the NOP56, RRP9, BYSL and DCAF13 genes and CRC. Further research on the mechanism of the NOP56, RRP9, BYSL or DCAF13 genes in CRC may be helpful.

To investigate the functions of these hub genes and three key module genes, GO and KEGG analyses were performed. The results indicated that these genes possibly function by regulating RNA processes, RNA binding, ribosome complex biogenesis, IFN-alpha and IFN-beta biosynthetic processes and the Toll-like receptor signaling pathway. TLR3, TLR7 and TLR8 genes in the third module [Figure 2d] were enriched in IFN-alpha and IFN-beta biosynthetic processes and the Toll-like receptor signaling pathway, and these three mRNAs were significantly decreased in CRC tissues compared with normal tissues. Toll-like receptors (TLRs) are widely expressed in tumor cells and may play a significant role in cancer biology.[55] Clinically, TLR agonists have been used as anticancer agents that activate immune cells in the tumor microenvironment.[56] IFN-α, -β, and -γ induce growth inhibition or cancer cell death.[57] Genetic variation of IFN is associated with the risk of colon cancer development.[58]

Additionally, we constructed a multiple gene risk model predicting overall survival, including IGF2BP3, PABPC1L, PPARGC1A, PTRH1 and TDRD7. The model is an independent risk factor for clinicopathological characteristics to predict overall survival. IGF2BP3 regulates IGF2 translation and function and is expressed in testis tissues.[59] IGF2BP3 is correlated with several cancers and is significantly upregulated in CRC tissues.[59] PABPC1L is an important paralog of PABPC1 that regulates and stabilizes mRNA translation.[60] A study reported that PABPC1L is overexpressed in CRC and promotes CRC cell proliferation and migration by regulating the protein kinase B (AKT)-signaling pathway.[60] A recent study validated that PABPC1L might promote colon tumorigenesis by regulating mRNA splicing.[61] Mutation of PPARGC1A is related to the risk of CRC, and PPARGC1A protects against tumorigenesis by regulating the fate of the enterocyte cells.[62,63] PTRH1 is peptidyl-tRNA hydrolase 1 and salvages tRNA from peptidyl-tRNA, which is toxic to cells when accumulating.[64] TDRD7 is a scaffold protein and can induce congenital cataracts.[64] However, there is no definitive study about the relationship between PTRH1 or TDRD7 and cancer.

Finally, the expression of risk model genes was validated using online databases. IGF2BP3, PABPC1L and PTRH1 were significantly upregulated in CRC tissues, but PPARGC1A and TDRD7 were downregulated compared with the normal group. Our results were consistent with those in the online databases.

One limitation of the study is the lack of experimental data validating the variations in the risk model genes. Although, all of the model gene mRNA levels in CRC were validated by the TIMER database, only IGF2BP3, PTRH1 and TDRD7 protein levels were validated by the HPA database. Therefore, it would be more convincing if a related experiment was performed to investigate the protein expression levels of the model genes, especially PABPC1L and PPARGC1A. Another limitation is the lack of more clinical data validating the prognostic signature established in the study. GSE17538 was used as the validation cohort, but it is necessary to collect more data and samples in the clinic to validate the signature before clinical application.

In conclusion, we screened differentially expressed RBPs, identified hub genes and analyzed the functions and related pathways in CRC. Subsequently, we constructed a risk score model including the IGF2BP3, PABPC1L, PPARGC1A, PTRH1 and TDRD7 genes. The prognostic model may help predict the overall survival of patients with CRC and be applied in the clinic in the future.

Acknowledgment

We thank Zhu Peng for the valuable discussion.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. Gut 2017;66:683-91.
2. Cavestro GM, Mannucci A, Zuppardo RA, Di Leo M, Stoffel E, Tonon G. Early onset sporadic colorectal cancer: Worrisome trends and oncogenic features. Dig Liver Dis 2018;50:521-32.
3. Yang Z, Chen Y, Wu D, Min Z, Quan Y. Analysis of risk factors for colon cancer progression. Onco Targets Ther 2019;12:3991-4000.
4. Burt RW. Colon cancer screening. Gastroenterology 2000;119:837-53.
5. Cannon-Albright LA, Skolnick MH, Bishop DT, Lee RG, Burt RW. Common inheritance of susceptibility to colonic adenomatous polyps and associated colorectal cancers. N Engl J Med 1988;319:533-7.
6. Atef N, Alieldin N, Sherif G, Loay I, Mostafa Mahmoud A, Mohamed G. Microsatellite instability and life style factors in sporadic colorectal cancer. Asian Pac J Cancer Prev 2020;21:1471-80.
7. Gerger A, Zhang W, Yang D, Bohanes P, Ning Y, Winder T, *et al*. Common cancer stem cell gene variants predict colon cancer recurrence. Clin Cancer Res 2011;17:6934-43.
8. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, *et al*. The consensus molecular subtypes of colorectal cancer. Nat Med 2015;21:1350-6.
9. Sidaway P. Immunoscore provides a more accurate prognosis. Nat Rev Clin Oncol 2018;15:471.
10. Wang ZL, Li B, Luo YX, Lin Q, Liu SR, Zhang XQ, *et al*. Comprehensive genomic characterization of RNA-binding proteins across human cancers. Cell Rep 2018;22:286-98.
11. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. Nat Rev Genet 2014;15:829-45.
12. Smith CW, Valcárcel J. Alternative pre-mRNA splicing: The logic of combinatorial control. Trends Biochem Sci 2000;25:381-8.
13. Turner M, Díaz-Muñoz MD. RNA-binding proteins control gene expression and cell fate in the immune system. Nat Immunol 2018;19:120-9.
14. Yang G, Fu H, Zhang J, Lu X, Yu F, Jin L, *et al*. RNA-binding protein quaking, a critical regulator of colon epithelial differentiation and a suppressor of colon cancer. Gastroenterology 2010;138:231-40.e1-5.
15. Hopkins TG, Mura M, Al-Ashtal HA, Lahr RM, Abd-Latip N, Sweeney K, *et al*. The RNA-binding protein LARP1 is a post-transcriptional regulator of survival and tumorigenesis in ovarian cancer. Nucleic Acids Res 2016;44:1227-46.
16. Kang D, Lee Y, Lee JS. RNA-binding proteins in cancer: Functional and therapeutic perspectives. Cancers (Basel) 2020;12:2699.
17. King CE, Cuatrecasas M, Castells A, Sepulveda AR, Lee JS, Rustgi AK. LIN28B promotes colon cancer progression and metastasis. Cancer Res 2011;71:4260-8.
18. Madison BB, Jeganathan AN, Mizuno R, Winslow MM, Castells A, Cuatrecasas M, *et al*. Let-7 represses carcinogenesis and a stem cell phenotype in the intestine via regulation of Hmga2. PLoS Genet 2015;11:e1005408.
19. Brody JR, Dixon DA. Complex HuR function in pancreatic cancer cells. Wiley Interdiscip Rev RNA 2018;9:e1469.
20. Chatterji P, Rustgi AK. RNA binding proteins in intestinal epithelial biology and colorectal cancer. Trends Mol Med 2018;24:490-506.
21. Deng M, Brägelmann J, Schultze JL, Perner S. Web-TCGA: An online platform for integrated analysis of molecular cancer data sets. BMC Bioinformatics 2016;17:72.
22. Meguid RA, Slidell MB, Wolfgang CL, Chang DC, Ahuja N. Is there a difference in survival between right- versus left-sided colon cancers? Ann Surg Oncol 2008;15:2388-94.
23. Tucker A, Tucker SP. Increasing colorectal cancer screening compliance through community education. Gastroenterol Nurs 2012;35:416-9.
24. Center MM, Jemal A, Smith RA, Ward E. Worldwide variations in colorectal cancer. CA Cancer J Clin 2009;59:366-78.
25. Makar GS, Makar M, Obinero C, Davis W, Gaughan JP, Kwiatt M. Refusal of cancer-directed surgery in patients with colon cancer: Risk factors of refusal and survival data. Ann Surg Oncol 2021;28:606-16.
26. Buskermolen M, Cenin DR, Helsingen LM, Guyatt G, Vandvik PO, Haug U, *et al*. Colorectal cancer screening with faecal immunochemical testing, sigmoidoscopy or colonoscopy: A microsimulation modelling study. BMJ 2019;367:l5383.
27. Carethers JM. Fecal DNA testing for colorectal cancer screening. Annu Rev Med 2020;71:59-69.
28. Issa IA, Noureddine M. Colorectal cancer screening: An updated review of the available options. World J Gastroenterol 2017;23:5086-96.
29. Ahlquist DA. Multi-target stool DNA test: A new high bar for noninvasive screening. Dig Dis Sci 2015;60:623-33.
30. Kulasingam V, Diamandis EP. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. Nat Clin Pract Oncol 2008;5:588-99.
31. Galon J, Mlecnik B, Bindea G, Angell HK, Berger A, Lagorce C, *et al*. Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. J Pathol 2014;232:199-209.
32. Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. N Engl J Med 2009;360:790-800.
33. Yokota M, Kojima M, Higuchi Y, Nishizawa Y, Kobayashi A, Ito M, *et al*. Gene expression profile in the activation of subperitoneal fibroblasts reflects prognosis of patients with colon cancer. Int J Cancer 2016;138:1422-31.
34. Lin S, Gregory RI. MicroRNA biogenesis pathways in cancer. Nat Rev Cancer 2015;15:321-33.
35. Pereira B, Billaud M, Almeida R. RNA-binding proteins in cancer: Old

players and new actors. Trends Cancer 2017;3:506-28.
36. Castello A, Fischer B, Frese CK, Horos R, Alleaume AM, Foehr S, *et al*. Comprehensive identification of RNA-binding domains in human cells. Mol Cell 2016;63:696-710.
37. Coppin L, Leclerc J, Vincent A, Porchet N, Pigny P. Messenger RNA life-cycle in cancer cells: Emerging role of conventional and non-conventional RNA-binding proteins? Int J Mol Sci 2018;19:650.
38. Matera AG, Wang Z. A day in the life of the spliceosome. Nat Rev Mol Cell Biol 2014;15:108-21.
39. Ohno K, Takeda JI, Masuda A. Rules and tools to predict the splicing effects of exonic and intronic mutations. Wiley Interdiscip Rev RNA 2018;9. doi: 10.1002/wrna. 1451.
40. Goudarzi KM, Lindström MS. Role of ribosomal protein mutations in tumor development (Review). Int J Oncol 2016;48:1313-24.
41. Zhang P, Ji D, Hu X, Ni H, Ma W, Zhang X, *et al*. Oncogenic heterogeneous nuclear ribonucleoprotein D-like promotes the growth of human colon cancer SW620 cells via its regulation of cell cycle. Acta Biochim Biophys Sin (Shanghai) 2018;50:880-7.
42. Béthune J, Jansen RP, Feldbrügge M, Zarnack K. Membrane-associated RNA-binding proteins orchestrate organelle-coupled translation. Trends Cell Biol 2019;29:178-88.
43. Lykke-Andersen S, Ardal BK, Hollensen AK, Damgaard CK, Jensen TH. Box C/D snoRNP autoregulation by a cis-Acting snoRNA in the NOP56 pre-mRNA. Mol Cell 2018;72:99-111.e5.
44. Cowling VH, Turner SA, Cole MD. Burkitt's lymphoma-associated c-Myc mutations converge on a dramatically altered target gene response and implicate Nol5a/Nop56 in oncogenesis. Oncogene 2014;33:3519-27.
45. Chen S, Blank MF, Iyer A, Huang B, Wang L, Grummt I, *et al*. SIRT7-dependent deacetylation of the U3-55k protein controls pre-rRNA processing. Nat Commun 2016;7:10734.
46. Wang H, Xiao W, Zhou Q, Chen Y, Yang S, Sheng J, *et al*. Bystin-like protein is upregulated in hepatocellular carcinoma and required for nucleologenesis in cancer cell proliferation. Cell Res 2009;19:1150-64.
47. Hou P, Shi P, Jiang T, Yin H, Chu S, Shi M, *et al*. DKC1 enhances angiogenesis by promoting HIF-1α transcription and facilitates metastasis in colorectal cancer. Br J Cancer 2020;122:668-79.
48. Wu XL, Yang ZW, He L, Dong PD, Hou MX, Meng XK, *et al*. RRS1 silencing suppresses colorectal cancer cell proliferation and tumorigenesis by inhibiting G2/M progression and angiogenesis. Oncotarget 2017;8:82968-80.
49. Hu L, Wang J, Liu Y, Zhang Y, Zhang L, Kong R, *et al*. A small ribosomal subunit (SSU) processome component, the human U3 protein 14A (hUTP14A) binds p53 and promotes p53 degradation. J Biol Chem 2011;286:3119-28.
50. Zhang J, Ren P, Xu D, Liu X, Liu Z, Zhang C, *et al*. Human UTP14a promotes colorectal cancer progression by forming a positive regulation loop with c-Myc. Cancer Lett 2019;440-441:106-15.
51. Qi J, Yu Y, Akilli Öztürk Ö, Holland JD, Besser D, Fritzmann J, *et al*. New Wnt/β-catenin target genes promote experimental metastasis and migration of colorectal cancer cells through different signals. Gut 2016;65:1690-701.
52. Li W, Li N, Gao L, You C. Integrated analysis of the roles and prognostic value of RNA binding proteins in lung adenocarcinoma. PeerJ 2020;8:e8509.
53. Wang K, Li L, Fu L, Yuan Y, Dai H, Zhu T, *et al*. Integrated bioinformatics analysis the function of RNA binding proteins (RBPs) and their prognostic value in breast cancer. Front Pharmacol 2019;10:140.
54. Cao J, Hou P, Chen J, Wang P, Wang W, Liu W, *et al*. The overexpression and prognostic role of DCAF13 in hepatocellular carcinoma. Tumour Biol 2017;39:1010428317705753.
55. Shi M, Chen X, Ye K, Yao Y, Li Y. Application potential of toll-like receptors in cancer immunotherapy: Systematic review. Medicine (Baltimore) 2016;95:e3951.
56. Braunstein MJ, Kucharczyk J, Adams S. Targeting toll-like receptors for cancer therapy. Target Oncol 2018;13:583-98.
57. Umemoto S, Haruta M, Sakisaka M, Ikeda T, Tsukamoto H, Komohara Y, *et al*. Cancer therapy with major histocompatibility complex-deficient and interferon β-producing myeloid cells derived from allogeneic embryonic stem cells. Cancer Sci 2019;110:3027-37.
58. Slattery ML, Lundgreen A, Bondurant KL, Wolff RK. Interferon-signaling pathway: Associations with colon and rectal cancer risk and subsequent survival. Carcinogenesis 2011;32:1660-7.
59. Shantha Kumara H, Kirchoff D, Caballero OL, Su T, Ahmed A, Herath SA, *et al*. Expression of the cancer testis antigen IGF2BP3 in colorectal cancers; IGF2BP3 holds promise as a specific immunotherapy target. Oncoscience 2015;2:607-14.
60. Wu YQ, Ju CL, Wang BJ, Wang RG. PABPC1L depletion inhibits proliferation and migration via blockage of AKT pathway in human colorectal cancer cells. Oncol Lett 2019;17:3439-45.
61. Zhang B, Babu KR, Lim CY, Kwok ZH, Li J, Zhou S, *et al*. A comprehensive expression landscape of RNA-binding proteins (RBPs) across 16 human cancer types. RNA Biol 2020;17:211-26.
62. D'Errico I, Salvatore L, Murzilli S, Lo Sasso G, Latorre D, Martelli N, *et al*. Peroxisome proliferator-activated receptor-gamma coactivator 1-alpha (PGC1alpha) is a metabolic regulator of intestinal epithelial cell fate. Proc Natl Acad Sci U S A 2011;108:6603-8.
63. Cho YA, Lee J, Oh JH, Chang HJ, Sohn DK, Shin A, *et al*. Genetic variation in PPARGC1A may affect the role of diet-associated inflammation in colorectal carcinogenesis. Oncotarget 2017;8:8550-8.
64. Selvaraj M, Singh NS, Roy S, Sangeetha R, Varshney U, Vijayan M. Cloning, expression, purification, crystallization and preliminary X-ray analysis of peptidyl-tRNA hydrolase from Mycobacterium tuberculosis. Acta Crystallogr Sect F Struct Biol Cryst Commun 2006;62(Pt 9):913-5.

Table S1: Differentially expressed RBPs in CRC tumors

	logFC	logCPM	*P*	FDR
Upregulated genes				
TRIM71	5.252092033	2.124678675	9.79E-13	5.52E-12
DDX53	5.128346312	−0.054815076	1.53E-07	5.07E-07
PIWIL1	4.83886599	5.578536266	2.89E-17	2.21E-16
RNF113B	4.433580112	−0.370795097	1.44E-08	5.55E-08
LIN28B	4.413644757	0.82379011	1.13E-05	3.01E-05
IGF2BP1	4.36561002	4.61413343	2.57E-11	1.29E-10
IGF2BP3	3.945034577	4.719395618	5.65E-13	3.28E-12
LIN28A	3.697972027	0.032742618	7.16E-07	2.24E-06
NANOS3	3.569228566	3.456647586	2.85E-19	2.74E-18
DAZ1	3.350030833	−1.184596163	0.000111317	0.000254009
BOLL	2.884140365	−0.434401514	5.47E-06	1.53E-05
AZGP1	2.812186463	8.63612364	2.05E-25	3.27E-24
NXF3	2.752221573	4.047676951	1.58E-06	4.78E-06
PABPC1L	2.710261612	7.200354027	5.42E-29	1.19E-27
MEX3A	2.542479708	6.387621406	1.53E-28	3.18E-27
ZC3HAV1L	1.874408723	5.42511512	7.56E-39	3.09E-37
RPL22L1	1.774744326	8.44434824	2.79E-16	2.01E-15
BOP1	1.598183922	9.423269293	7.49E-22	9.18E-21
PABPC3	1.557671526	3.687150556	5.18E-10	2.34E-09
GTF3A	1.55732465	10.02434717	2.36E-23	3.22E-22
ZNF239	1.547431548	5.248492378	6.48E-23	8.60E-22
PTRH1	1.520944693	0.480847081	1.29E-10	6.13E-10
CD3EAP	1.471948231	5.578793212	4.20E-42	1.94E-40
NPM2	1.454557228	3.678553176	6.95E-11	3.42E-10
PUS7	1.405900191	7.463647345	8.91E-47	5.55E-45
RANBP17	1.367449355	5.050817554	1.42E-10	6.65E-10
CTU1	1.342375408	5.41828813	1.12E-10	5.34E-10
DNMT3B	1.302205155	5.107081478	7.72E-18	6.18E-17
EXO1	1.289001074	6.269997084	2.37E-27	4.36E-26
TDRD1	1.282895129	2.316847696	0.010904959	0.017838819
NPM3	1.262189959	7.809369459	8.12E-18	6.46E-17
PUS1	1.261090894	8.003907669	3.08E-26	5.14E-25
ADAT2	1.25863411	6.533245296	2.13E-22	2.71E-21
TERT	1.253008914	3.809619762	1.42E-07	4.74E-07
WDR4	1.224610552	6.9845416	5.65E-25	8.53E-24
BYSL	1.213992186	7.469396478	1.12E-36	4.11E-35
DKC1	1.186280907	9.625946252	1.11E-49	7.21E-48
RPS21	1.176678106	12.23290432	1.48E-09	6.39E-09
RNASEH2A	1.175551544	7.909875893	1.67E-20	1.80E-19
EIF5A2	1.16515838	5.888026788	6.45E-17	4.90E-16
CMSS1	1.160142376	7.175366188	3.22E-40	1.40E-38
RNASE7	1.158878615	1.155198822	0.000614424	0.001254231
POP1	1.136162109	6.40586659	2.93E-26	4.94E-25
RRP9	1.128266644	7.984953818	1.41E-21	1.64E-20
NOP56	1.123190457	9.755925619	2.19E-28	4.48E-27
NXT1	1.112391285	7.277514032	2.95E-16	2.12E-15
AEN	1.110808127	8.007551354	2.30E-25	3.59E-24
NKRF	1.11002803	6.579241127	1.27E-32	3.50E-31
RRP12	1.101723077	8.372373387	2.25E-31	5.85E-30
NOP2	1.09919167	8.495440418	3.54E-36	1.21E-34
XPOT	1.08459342	9.255088614	2.43E-31	6.23E-30
METTL1	1.082529018	6.935630618	5.64E-31	1.39E-29
EXOSC5	1.074013455	7.696109804	1.27E-15	8.83E-15
EXOSC4	1.062443431	7.65030706	3.37E-11	1.68E-10
RUVBL1	1.059453166	8.558068427	1.92E-43	9.48E-42
TRMT1	1.057786733	8.066260156	5.52E-20	5.62E-19
NUFIP1	1.056410149	6.302977769	6.33E-21	6.98E-20
UTP14A	1.049426319	7.889421464	2.24E-25	3.53E-24
RRS1	1.042385669	7.936299522	7.39E-22	9.13E-21
RPP40	1.027915589	5.570807288	5.83E-24	8.53E-23
IPO4	1.014828039	5.240761805	6.46E-17	4.90E-16
RAE1	1.01441098	8.398347334	3.13E-24	4.62E-23
EZH2	1.013098708	7.618504345	1.96E-33	5.86E-32
BRCA1	1.007746875	7.255928368	2.04E-19	1.98E-18

Contd...

Table S1: Contd...

	logFC	logCPM	*P*	FDR
DCAF13	1.004870648	8.498221702	3.32E-23	4.45E-22
RBM44	1.002443309	1.674897423	0.00036559	0.000771561
Downregulated genes				
SPATS2L	−1.018711652	9.381277297	1.46E-44	7.76E-43
MRPL35	−1.021146683	8.194693085	9.67E-71	1.98E-68
RNASEK	−1.02135772	4.991526262	5.91E-28	1.16E-26
TLR8	−1.024775724	4.042673828	1.96E-05	5.06E-05
AFF1	−1.03109313	8.68311084	4.17E-20	4.27E-19
TDRD7	−1.031278158	7.543353529	1.01E-44	5.59E-43
IFIT2	−1.032077937	5.936624667	2.75E-09	1.15E-08
AUH	−1.032900902	6.536551042	1.21E-50	8.68E-49
ISG20	−1.034250731	7.135576114	1.12E-11	5.83E-11
NUDT16	−1.046325651	8.181935779	1.77E-42	8.44E-41
ENDOU	−1.062598483	−0.004575165	3.06E-06	8.89E-06
RNASE6	−1.064261911	5.752743539	3.32E-18	2.85E-17
ZC3H12D	−1.073679084	5.353361389	5.64E-18	4.64E-17
MBNL1	−1.075785462	9.4055544	5.34E-28	1.06E-26
POLR2F	−1.090844126	−0.11558887	1.00E-06	3.12E-06
SIDT1	−1.135163316	6.467405631	7.92E-13	4.52E-12
LRRFIP2	−1.138718087	8.168415741	2.33E-57	2.39E-55
CPEB2	−1.146052023	6.680102289	1.65E-21	1.90E-20
IFIT1	−1.150260371	6.156479839	1.37E-08	5.31E-08
SAMD4A	−1.175497088	6.545515214	1.37E-14	8.82E-14
TEP1	−1.188112107	8.170597653	2.06E-36	7.38E-35
SECISBP2L	−1.215358029	7.495299122	1.34E-27	2.49E-26
MVP	−1.222810402	11.32647701	2.65E-35	8.64E-34
SNRPN	−1.251220236	6.263888265	2.52E-12	1.37E-11
SMAD7	−1.256486047	7.156267483	9.96E-33	2.80E-31
RBFOX1	−1.26740694	1.783445713	0.000185201	0.000408298
SIDT2	−1.285276523	8.013266473	4.77E-44	2.44E-42
ZCCHC24	−1.322832986	7.19865387	9.69E-18	7.67E-17
DQX1	−1.33521449	6.046592594	2.89E-14	1.83E-13
RNASE1	−1.357335127	9.494884607	2.12E-21	2.42E-20
DZIP1	−1.413625591	4.776508966	3.16E-19	3.01E-18
RAVER 2	−1.435878882	8.024897588	8.10E-69	1.45E-66
ENOX1	−1.483998611	3.159585265	1.05E-18	9.70E-18
RBMS3	−1.533057774	4.768315059	1.84E-15	1.26E-14
RBM47	−1.569350087	9.903733603	1.87E-98	8.95E-96
CPEB3	−1.569508522	5.060620275	1.74E-45	9.99E-44
A1CF	−1.569649279	7.285646369	7.51E-23	9.78E-22
RNASEL	−1.571605874	6.334090368	2.58E-57	2.46E-55
PPARGC1B	−1.575478851	6.898803885	6.04E-54	4.81E-52
AFF2	−1.590884551	1.07330892	7.04E-10	3.12E-09
ZC3H12C	−1.597876992	6.219690847	2.11E-23	2.91E-22
PABPC5	−1.643140524	1.470718935	5.35E-18	4.43E-17
DDX25	−1.689538826	−0.543449745	5.68E-10	2.53E-09
OASL	−1.766689825	6.688054671	7.82E-27	1.40E-25
DDX60	−1.78664835	7.895634882	5.64E-34	1.72E-32
RIOK3	−1.79956147	8.886852168	2.36E-80	6.76E-78
TST	−1.844266359	9.964738094	1.35E-59	1.49E-57
TLR7	−1.94116463	3.868143012	2.34E-26	4.04E-25
TDRD10	−1.994318385	1.078344946	9.54E-36	3.18E-34
EIF4E3	−1.998820321	7.264031251	4.52E-55	4.05E-53
PPARGC1A	−2.021604371	6.482288359	2.17E-33	6.34E-32
RBM24	−2.150961737	3.018653031	4.00E-28	8.08E-27
TLR3	−2.324869449	6.156211797	1.94E-72	4.63E-70
RPL10L	−2.345603898	1.234731438	1.20E-27	2.25E-26
PDCD4	−2.352722098	9.300681351	5.69E-189	8.16E-186
CELF4	−2.654819968	1.001503732	1.47E-30	3.45E-29
RBM20	−2.667853092	2.192803562	8.08E-39	3.21E-37
RALYL	−2.946160789	−0.586432324	5.34E-19	4.97E-18
AFF3	−3.135612519	3.658371055	2.16E-67	3.09E-65
NOVA1	−3.215254166	2.889620781	1.02E-52	7.68E-51
CPEB1	−3.321445537	1.968416617	2.09E-63	2.49E-61
ELAVL4	−3.364897978	1.948050664	4.59E-68	7.31E-66

Contd...

Table S1: Contd...

	logFC	logCPM	*P*	FDR
ELAVL3	−3.548646359	0.03139513	1.07E-54	8.98E-53
RBPMS2	−3.612398239	4.71748561	1.85E-87	6.62E-85
KHDRBS2	−4.204963195	−0.03288386	4.17E-65	5.43E-63
RBFOX3	−5.008485622	3.927742185	5.15E-122	3.69E-119