Transcriptome Analyses Identify an RNA Binding Protein Related Prognostic Model for Clear Cell Renal Cell Carcinoma

Yue Wu1,2, Xian Wei1,2, Huan Feng1,2, Bintao Hu1,2, Bo Liu3, Yang Luan1,2, Yajun Ruan1,2, Xiaming Liu1,2, Zhuo Liu1,2, Shaogang Wang1,2, Jihong Liu1,2 and Tao Wang1,2*

1 Department of Urology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, 2 Institute of Urology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, 3 Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

RNA binding proteins (RBPs) play a key role in post-transcriptional gene regulation. They have been shown to be dysfunctional in a variety of cancers and are closely related to the occurrence and progression of cancers. However, the biological function and clinical significance of RBPs in clear cell renal carcinoma (ccRCC) are unclear. In our current study, we downloaded the transcriptome data of ccRCC patients from The Cancer Genome Atlas (TCGA) database and identified differential expression of RBPs between tumor tissue and normal kidney tissue. Then the biological function and clinical value of these RBPs were explored by using a variety of bioinformatics techniques. We identified a total of 40 differentially expressed RBPs, including 10 down-regulated RBPs and 30 up-regulated RBPs. Eight RBPs (APOBEC3G, AUH, DAZL, EIF4A1, IGF2BP3, NR0B1, RPL36A, and TRMT1) and nine RBPs (APOBEC3G, AUH, DDX47, IGF2BP3, MOV10L1, NANOS1, PIH1D3, TDRD9, and TRMT1) were identified as prognostic related to overall survival (OS) and disease-free survival (DFS), respectively, and prognostic models for OS and DFS were constructed based on these RBPs. Further analysis showed that OS and DFS were worse in high-risk group than in the low-risk group. The area under the receiver operating characteristic curve of the model for OS was 0.702 at 3 years and 0.726 at 5 years in TCGA cohort and 0.783 at 3 years and 0.795 at 5 years in E-MTAB-1980 cohort, showing good predictive performance. Both models have been shown to independently predict the prognosis of ccRCC patients. We also established a nomogram based on these prognostic RBPs for OS and performed internal validation in the TCGA cohort, showing an accurate prediction of ccRCC prognosis. Stratified analysis showed a significant correlation between the prognostic model for OS and ccRCC progression.

Keywords: clear cell renal cell carcinoma, RNA binding proteins, prognostic model, survival analysis, bioinformatics

Abbreviations: RCC, renal cell carcinoma; ccRCC, clear cell renal cell carcinoma; TCGA, The Cancer Genome Atlas; RBPs, RNA binding proteins; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; FC, fold change; OS, overall survival; DFS, disease-free survival; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic; AUC, area under the receiver operating characteristic curve; FDR, false discovery rate.
INTRODUCTION

Renal cell carcinoma (RCC) accounts for 2.4% of all malignancies, with an estimated 400,000 new cases and 175,000 deaths worldwide each year (Bray et al., 2018; Siegel et al., 2018). The clear cell renal cell carcinoma (ccRCC) is the most common subtype of RCC, accounting for approximately 70–80% and presents a high risk of heterogeneity and metastasis (Rini et al., 2011; Ljungberg et al., 2019). Although surgical resection can effectively resolve the early stage of ccRCC, 30% of patients still have recurrence or metastasis after surgery, and the late stage of ccRCC has a high mortality rate due to insensitivity to traditional radiotherapy or chemotherapy (Battaglia and Lucarelli, 2015; Moch et al., 2016; Tamma et al., 2019). Therefore, further understanding of the molecular mechanisms of ccRCC and the discovery of more effective molecular biomarkers are essential for early screening, diagnosis, monitoring for metastasis, recurrence, and quality of life in patients.

Post-transcriptional regulation of RNA is an important aspect of gene expression regulation. RNA binding proteins (RBPs) form ribonucleoprotein (RNP) complexes through binding at different sites or random interaction with target RNA, thus strictly regulating RNA metabolism (Iadevaia and Gerber, 2015; Hentze et al., 2018). Currently, there are 1,542 RBP coding genes, accounting for 7.5% of all human protein-coding genes, which have been verified by experiments (Gerstberger et al., 2014). These RBPs regulate a variety of biological processes including RNA processing, splicing, mRNA stability, output, localization, and translation, thus maintaining the physiological balance of the cell (Masuda and Kuwano, 2019). Given this, it comes as no surprise that RBPs dysfunction has been linked to a variety of human diseases. Ribosomal diseases caused by ribosomal protein and tRNA biogenic factor defects, such as Diamond–Blackfan anemia and Shwachman–Diamond syndrome, affect the same tissues and exhibit similar pathology precisely because RBPs bind to the same type of RNA (Narla and Ebert, 2010). Mutations in mRBPs or their targets in neurons lead to abnormal aggregation of proteins or RNA, resulting in a variety of neurodegenerative and neuromuscular diseases (Scheper et al., 2007). However, the role of RBPs in tumor genesis and development is rare. Some studies have shown that RBPs are abnormally expressed in tumor tissues compared with normal tissues and are associated with patient prognosis (Patry et al., 2003; Busà et al., 2007; Ortiz-Zapater et al., 2011). In lung cancer, QKI inhibits tumor cell proliferation by competing with the splicing factor SF1 (Zong et al., 2014). In melanoma, CPEB4 promotes tumor cell proliferation by regulating polyadenylation and promoting the translation of melanoma drivers (Pérez-Guizarro et al., 2016). Knockdown SAM68 in breast cancer cells inhibited tumor cell proliferation by upregulation of cell cycle inhibitors P21 and CDKN1B/P27 (Song et al., 2010). However, in the field of ccRCC, existing studies only described the effect of RBPs on the overall survival (OS) of ccRCC patients (Hua et al., 2020; Zhu et al., 2020), and few RBPs models can be used to predict the prognosis of ccRCC patients. The development of new RBPs models has gradually become an effective method to explore novel therapeutic targets. Therefore, in our current study, we systematically and deeply analyzed the molecular biological function and clinical significance of RBPs in ccRCC to promote our understanding of ccRCC progress, and established risk score models for OS and disease-free survival (DFS), which may provide new biomarkers for disease diagnosis and treatment prognosis.

MATERIALS AND METHODS

Preprocessing Data and Identifying Differential Expression RBPs

Transcriptome data of 72 normal renal tissue specimens and 539 ccRCC specimens were downloaded from The Cancer Genome Atlas database (TCGA1). We then used the edgeR package2 to preprocess the raw data and identify the differentially expressed RBPs based on $|\log_2$ fold change (FC)| $> 1.0$ and false discovery rate (FDR) $< 0.05$. We also downloaded the E-MTAB-1980 dataset from the ArrayExpress database3 and downloaded the transcriptome data of 436 ccRCC patients containing DFS information from the cBioportal database4.

Function and Pathway Enrichment Analysis

We used the WEB-based Gene Set Analysis Toolkit (WebGestalt5) online analysis tool to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of these differentially expressed RBPs (Liao et al., 2019). The GO terms including biological process, cellular component, and molecular function. All analysis results were screened according to the criteria of $P < 0.05$ and gene number $> 5$.

Selection of Prognostic Related RBPs

To identify RBPs with important prognostic significance, we first performed univariate Cox regression analysis of all these differentially expressed RBPs. The least absolute shrinkage and selection operator (LASSO) regression analysis was then used for further screening. Finally, multivariate Cox regression analysis was used to further screen out RBPs with important prognostic value. A $P < 0.05$ was considered significant.

Construction and Evaluation of Prognostic Model for OS

We constructed a multivariate Cox proportional hazards regression model to predict the prognosis of ccRCC patients based on these prognostic related RBPs. The risk score for each patient in the model was calculated using the following formula:

\[
\text{Risk score} = \sum_{i=1}^{n} \text{Expi} \beta_i,
\]

1https://portal.gdc.cancer.gov/
2http://www.bioconductor.org/packages/release/bioc/html/edgeR.html
3https://www.ebi.ac.uk/arrayexpress/
4https://www.cbioportal.org/datasets
5http://www.webgestalt.org/
In this formula, $Exp$ represents the expression value of each gene, and $\beta$ represents the corresponding regression coefficient. We then divided ccRCC patients from the TCGA cohort into low-risk and high-risk subgroups based on the median risk score, and compared OS between the two groups to initially assess the predictive power of the model. In addition, we used the Survival ROC R package to establish the ROC curve to assess the prognostic efficacy of the model and used the rms R package to draw the nomogram to predict OS. Finally, we divided the 539 samples in the TCGA cohort into the training group and the validation group as internal validation and the E-MTAB-1980 cohort with 101 sample information as external validation to evaluate the stability and predictive efficacy of the model.

**Gene Set Enrichment Analysis**
We divided the patients into low-risk and high-risk groups based on the median risk score of the prognostic model, and then performed gene set enrichment analysis (GSEA) by using GSEA_4.0.3 software\footnote{https://www.gsea-msigdb.org/gsea/downloads.jsp}. A $P < 0.05$ and FDR $< 0.25$ were considered to be significant differences.

**Express Level and Prognostic Significance Verification of Prognostic Related RBPs**
We used The Human Protein Atlas (HPA\footnote{https://www.proteinatlas.org/}) online database to verify the protein expression levels of these prognostic related RBPs. And the Kaplan–Meier plotter\footnote{https://kmplot.com/analysis/} online tool was used to assess the prognostic significance of these prognostic related RBPs in ccRCC patients.

**Construction and Evaluation of Prognostic Model for DFS**
Since DFS is also important for the prognosis of tumor patients, we constructed a prognostic model for DFS. We downloaded

---

**FIGURE 1** | The research diagram for analyzing RBPs in ccRCC.
transcriptomic data from the cBioportal database for 436 ccRCC patients with DFS information. Then the prognostic RBPs were screened by Cox regression analysis and LASSO regression analysis and a prognostic model for DFS was constructed.

**Statistical Analysis**

R software (Version 4.0.0) was used for statistical analysis. The differentially expressed genes in tumor tissues and normal tissues were analyzed by “edgeR” package. Cox regression analysis was used to screen for genes associated with prognosis. The OS and DFS of patients were analyzed by Kaplan–Meier method and log-rank test. The “survival ROC” package was used to analyze the ROC curve. The "rms" package was used to draw the nomogram. The Student’s t-test or non-parametric Mann–Whitney rank sum test was used to compare the correlation between risk score, prognostic genes, and clinicopathological variables. P < 0.05 was considered statistically significant.

**RESULTS**

**Screening Differentially Expressed RBPs in ccRCC**

The analysis process of this study was shown in Figure 1. Transcriptome data of ccRCC patients were downloaded from the TCGA database, including 72 normal renal tissue samples and 539 tumor tissue samples (Supplementary Table S1). The edger R package was used to process the data and identify the
differentially expressed RBPs. Of the 1542 RBPs (Gerstberger et al., 2014), 40 met our criteria (log2 FC > 1.0, FDR < 0.05), including 10 down-regulated RBPs and 30 up-regulated RBPs. Figure 2 showed the expression and distribution of these differentially expressed RBPs.

Function and Pathway Enrichment Analysis of These Differentially Expressed RBPs

We performed GO and KEGG enrichment analyses for these differentially expressed RBPs using the WebGestalt online analysis tool to investigate the biological functions and molecular mechanisms of these genes. The analysis results were shown in Table 1. The biological processes analysis showed that these RBPs were significantly enriched in RNA catabolic process, posttranscriptional regulation of gene expression, translational initiation, regulation of cellular amide metabolic process, protein localization to endoplasmic reticulum, meiotic cell cycle, gene silencing, cellular process involved in reproduction in multicellular organism, transposition, and regulation of mRNA metabolic process. The cellular component showed that these RBPs were significantly enriched in polysome, ribosome, ribonucleoprotein granule, cytosolic part, and rough endoplasmic reticulum. In terms of molecular function, these RBPs were significantly enriched in mRNA binding, catalytic activity, acting on RNA, structural constituent of ribosome, helicase activity, nuclease activity, translation regulator activity, snRNA binding, double-stranded RNA binding, ATPase activity, and nucleotidytransferase activity. Moreover, KEGG analysis showed that these RBPs were mainly enriched in ribosome, RNA transport, influenza A, mRNA surveillance pathway, herpes simplex infection, and ribosome biogenesis in eukaryotes.

### Table 1 | KEGG pathway and GO enrichment analysis of differentially expressed RNA binding proteins.

| GO term                                | P-value |
|----------------------------------------|---------|
| Biological processes                   |         |
| RNA catabolic process                   | 2.81e-12|
| Posttranscriptional regulation of gene expression | 2.34e-10|
| Translational initiation                | 3.72e-8 |
| Regulation of cellular amide metabolic process | 5.31e-8 |
| Protein localization to endoplasmic reticulum | 0.000002|
| Meiotic cell cycle                     | 0.000004|
| Gene silencing                         | 0.000016|
| Cellular process involved in reproduction in multicellular organism | 0.000034|
| Transposition                          | 0.000057|
| Regulation of mRNA metabolic process   | 0.000073|
| Polyosome                              | 3.38e-10|
| Ribosome                               | 4.88e-8 |
| Ribonucleoprotein granule              | 6.13e-7 |
| Cytosolic part                         | 0.000001|
| Rough endoplasmic reticulum            | 0.019853|
| mRNA binding                           | 4.25e-10|
| Catalytic activity, acting on RNA      | 2.15e-8 |
| Structural constituent of ribosome     | 1.84e-7 |
| Helicase activity                      | 0.000052|
| Nuclease activity                      | 0.000228|
| Translation regulator activity         | 0.000561|
| snRNA binding                          | 0.005157|
| Double-stranded RNA binding            | 0.018541|
| ATPase activity                        | 0.031468|
| Nucleotidytransferase activity         | 0.049070|
| Ribosome                               | 1.03e-8 |
| RNA transport                          | 0.000494|
| Influenza A                            | 0.000553|
| mRNA surveillance pathway              | 0.001136|
| Herpes simplex infection               | 0.008472|
| Ribosome biogenesis in eukaryotes       | 0.014958|

**GO**, Gene Ontology; **KEGG**, Kyoto Encyclopedia of Genes and Genomes.

**Prognostic Related RBPs Selection**

We performed a univariate Cox regression analysis on all these differentially expressed RBPs and obtained 25 prognostic related RBPs (Figure 3). We further performed LASSO regression analysis on these 25 genes to screen the RBPs with prognostic significance, and obtained 9 RBPs including APOBEC3G, AUH, DAZL, DDX47, EIF4A1, IGF2BP3, NR0B1, RPL36A, and TRMT1 (Supplementary Figure S2). And multivariate Cox regression analysis showed that 8 of the 9 RBPs, namely, APOBEC3G, AUH, DAZL, EIF4A1, IGF2BP3, NR0B1, RPL36A, and TRMT1 independently predicted prognosis of ccRCC patients.

**Prognostic Related Risk Score Model for OS Construction and Evaluation**

We used these eight genes screened from multivariate Cox regression analysis to establish a prognostic model for OS (Table 2). Each ccRCC patient’s risk score was calculated according to the following formula:

\[
Risk \, score = (0.0951 \times \text{Exp} \, APOBEC3G) + (0.1190 \times \text{Exp} \, AUH) + (0.0945 \times \text{Exp} \, DAZL) + (0.1571 \times \text{Exp} \, EIF4A1) + (0.1722 \times \text{Exp} \, RPL36A) + (0.2380 \times \text{Exp} \, TRMT1)
\]

Based on the median risk score, 539 ccRCC patients in the TCGA cohort were divided into low-risk and high-risk subgroups for survival analysis to assess the predictive power of the model. Survival analysis showed that patients in the high-risk group had lower OS than those in the low-risk group (P = 9.556e-13, Figure 4A). We then performed the time-dependent receiver operating characteristic (ROC) analysis to further evaluate the predictive performance of the eight RBPs signature, and the area under the ROC curve (AUC) of the model was 0.729 at 1 year, 0.702 at 3 years, and 0.726 at 5 years (Figure 4B). Figure 4C showed the survival status of each patient in the TCGA cohort assessed by risk score. Subsequently, to evaluate the applicability and stability of the prognostic model for OS, these 539 ccRCC patients in the TCGA cohort were randomly divided into a training data set and a validation data set. We then used the same formula to calculate the risk score of each patient to assess the predictive performance of the model. The results showed that patients in the high-risk group in the training data set had worse OS than those in the low-risk group (P = 1.908e-05, Figure 5A).
We found that the AUC was 0.750 at 1 year, 0.697 at 3 years, and 0.759 at 5 years (Figure 5B). And patients in the validation data set had similar results (Figures 5C,D). In addition, to assess whether the model has similar predictive power in other ccRCC patient cohorts, the same risk score formula was used for the E-MTAB-1980 dataset. Survival analysis also showed that patients in the high-risk group had lower OS than those in the low-risk group ($P = 0.00033$, Figure 6A), and the AUC of the model was 0.788 at 1 year, 0.783 at 3 years, and 0.795 at 5 years (Figure 6B). And Figure 6C showed the survival status of each patient in the E-MTAB-1980 cohort assessed by risk score. These results showed that the signature of these eight RBPs has good predictive performance and stability.

Next, we used the risk score in the prognostic model as the label to explore the functional differences between the two subgroups by conducting GSEA. The results showed that ribosome was mainly enriched in the high-risk group (Figure 7A), indicating that the regulation of RBPs was mainly involved in high-risk ccRCC patients. In addition, we performed a univariate Cox regression analysis for different clinical characteristics of ccRCC patients to evaluate their respective predictive significance. The results showed that age, tumor grade, tumor stage, primary tumor location, regional lymph node invasion, distant metastasis, and risk score were all associated with the OS of ccRCC patients (Figure 7B). However, multiple regression analysis showed that only age ($P < 0.001$), tumor grade ($P = 0.020$), tumor stage ($P < 0.001$), and risk score ($P < 0.001$) were independent prognostic factors related to OS of ccRCC patients (Figure 7C). Moreover, to establish a quantitative prognostic approach for ccRCC patients, we drew a nomogram based on the risk score and other clinical variables (Figure 7D). By drawing a vertical line between each prognosis axis and the total point axis, we can predict the survival probability of ccRCC patients at 1, 3, and 5 years. We also constructed calibration curves to evaluate the predictive performance of the model.

| Gene     | Coef   | Exp(coef) | se(coef) | z     | $Pr (> | z |)$ |
|----------|--------|-----------|----------|-------|--------|
| APOBEC3G | 0.0951 | 1.0998    | 0.0844   | 1.1264| 0.2600 |
| AUH      | -0.1621| 0.8504    | 0.1318   | -1.2299| 0.2187 |
| DAZL     | 0.0945 | 1.0991    | 0.0695   | 1.3597| 0.1739 |
| EIF4A1   | 0.1571 | 1.1701    | 0.1701   | 2.2138| 0.0268 |
| IGF2BP3  | 0.1190 | 1.1264    | 0.0346   | 3.4376| 0.0006 |
| NR0B1    | 0.0998 | 1.0150    | 0.0366   | 2.7241| 0.0064 |
| RPL36A   | 0.1722 | 1.1879    | 0.1493   | 1.1552| 0.2488 |
| TRMT1    | 0.2380 | 1.2687    | 0.1632   | 1.4583| 0.1448 |

**TABLE 2** | Multivariate Cox regression analysis to identify prognosis-related RNA binding proteins.

Coef, coefficient.
the nomogram, and the results showed that there was high consistency between the predicted results and the actual results (Figures 7E–G). And we used the TCGA and E-MTAB-1980 cohorts to verify the accuracy and stability of nomogram to expand its clinical application and availability. Survival analysis showed that nomogram could better distinguish ccRCC patients with low survival rates in TCGA and E-MTAB-1980 cohorts ($P < 0.001$ and $P = 2.32e-05$, Figures 7H,J). Based on the nomogram, the AUC in the TCGA cohort was 0.867 at 1 year, 0.806 at 3 years and 0.778 at 5 years (Figure 7I), and the AUC in the E-MTAB-1980 cohort was 0.910 at 1 year, 0.917 at 3 years, and 0.892 at 5 years (Figure 7K).

**Prognostic Value of the Prognostic Model for OS Stratified by Clinical Parameters**

To explore the clinical significance of the signature based on these eight RBPs in the ccRCC patients stratified by different clinical parameters, we stratified ccRCC patients from TCGA database according to age, gender, grade, stage, T stage, M stage, and N stage. Kaplan–Meier survival curve analysis showed that the OS was significantly shorter for the ccRCC patients in the high-risk group compared to the low-risk group ccRCC patients (Figure 8). These results indicate that the signature of these eight RBPs can predict the prognosis of ccRCC patients without considering clinical parameters.

**Relationship Between Prognostic Model for OS and Clinical Parameters**

We analyzed the correlation between the prognostic model based on these eight RBPs and clinical parameters to explore whether the prognostic model might influence the progression of ccRCC. The results showed no significant correlation between age and prognostic model (Figure 9A). However, the risk score of females was significantly lower than that of male (Figure 9B), the risk score of G1-2 was significantly lower.
FIGURE 5 | Risk score analysis of the eight RBPs prognostic model for OS in the training and validation data set. (A) Kaplan–Meier OS curve analysis in the training data set; (B) time dependent ROC curve analysis in the training data set; (C) Kaplan–Meier OS curve analysis in the validation data set; (D) time dependent ROC curve analysis in the validation data set.

than that of G3-4 (Figure 9C), the risk score of stage I-II was significantly lower than that of stage III-IV (Figure 9D), the risk score of T1-2 was significantly lower than that of T3-4 (Figure 9E), the risk score of M0 was significantly lower than that of M1 (Figure 9F) (The N1 in the N stage is very small and cannot be analyzed). These results showed that prognostic model for OS was significantly correlated with ccRCC tumor progression.

Correlation Between Prognostic RBPs and Clinical Parameters

Based on the above results, we also analyzed the relationship between prognostic RBPs for OS and clinical parameters to further investigate the role of prognostic RBPs in ccRCC. The results showed that AUH, EIF4A1, IGF2BP3, and RPL36A were significantly correlated with gender; APOBEC3G, AUH, IGF2BP3, RPL36A, and TRMT1 were significantly correlated with grade; APOBEC3G, AUH, DAZL, IGF2BP3, RPL36A, and TRMT1 were significantly correlated with stage; APOBEC3G, AUH, DAZL, IGF2BP3, NR0B1, RPL36A, and TRMT1 were significantly correlated with T stage; APOBEC3G, AUH, IGF2BP3, RPL36A, and TRMT1 were significantly correlated with M stage. However, there was no significant correlation between NR0B1 and these clinical parameters (Table 3).

Express Level and Prognostic Significance Verification of Prognostic Related RBPs

To assess the prognostic significance of these prognostic related RBPs in ccRCC patients, we used the Kaplan–Meier plotter online tool to confirm the relationship between these genes and OS. The results showed that all the eight RBPs were related to the OS in ccRCC patients (Figure 10). Subsequently, we used the HPA online database to verify the protein expression levels of these prognostic related RBPs, the results showed that APOBEC3G, EIF4A1, and TRMT1 were significantly increased in ccRCC tissue compared with normal renal tissue (Figures 11A,D,G). And AUH, DAZL, IGF2BP3, and RPL36A were significantly reduced in ccRCC tissue compared with normal renal tissue (Figures 11B,C,E,F). However, the protein expression level of NR0B1 was not available on the HPA online database.

Construction of a Prognostic Model for DFS

In view of the important influence of DFS on the prognosis of ccRCC, we also constructed a prognostic model for DFS. The expression data of 436 ccRCC patients and the corresponding DFS information were download from the cBioportal database. We then identified 9 prognostic RBPs including APOBEC3G,
AUH, DDX47, IGF2BP3, MOV10L1, NANOS1, PIH1D3, TDRD9, and TRMT1 by univariate Cox regression analysis, LASSO regression analysis and multivariate Cox regression analysis. We then constructed a prognostic model for DFS based on these nine prognostic RBPs and calculated each patient’s risk score based on the following formula: Risk score = (0.0852 × Exp APOBEC3G) + (−0.3683 × Exp AUH) + (0.4195 × Exp DDX47) + (0.1445 × Exp IGF2BP3) + (−0.2077 × Exp MOV10L1) + (0.4206 × Exp NANOS1) + (0.7675 × Exp PIH1D3) + (−0.1011 × Exp TDRD9) + (0.2895 × Exp TRMT1). Based on the median risk score, these 436 ccRCC patients were divided into high-risk and low-risk groups for survival analysis to assess the predictive performance of the prognostic model. The results showed that patients in the high-risk group had worse DFS than those in the low-risk group (P = 1.110e-16, Figure 12A). We found that the AUC for DFS was 0.718 at 1 year, 0.763 at 3 years, and 0.813 at 5 years (Figure 12D). These results showed that the RBPs associated prognostic model for DFS has good predictive performance.

In addition, we randomly divided the whole dataset into a training data set (n = 218) and a validation data set (n = 218) to assess the applicability and stability of the prognostic model for DFS. We used the same formula to calculate each patient’s risk score. Survival analysis showed that patients in the high-risk group in the training data set had worse DFS than those in the low-risk group (P = 1.127e-10, Figure 12B). The AUC for DFS was 0.718 at 1 year, 0.763 at 3 years, and 0.813 at 5 years (Figure 12E). Patients in the validation data set had similar results (Figures 12C,F).

Moreover, the prognostic value of the prognostic model for DFS and different clinical parameters were evaluated by Cox regression analysis. The results indicated that the tumor grade, tumor stage, primary tumor location, distant metastasis, and risk score of ccRCC patients were significantly correlated with DFS.

**FIGURE 6 |** Risk score analysis of the eight RBPs prognostic model for OS in the E-MTAB-1980 cohort. (A) Kaplan–Meier OS curve analysis; (B) time dependent ROC curve analysis; (C) survival status of each patient.
However, multiple regression analysis revealed that tumor grade, tumor stage, and risk score were independent prognostic factors associated with DFS \( (P < 0.001) \), Figure 12G). These results suggested that the RBPs associated prognostic model for DFS was also a good predictor of ccRCC patient outcomes.
FIGURE 8 | Kaplan–Meier survival curves analysis stratified by different clinical parameters. (A) Age ≤ 65; (B) Age > 65; (C) Female; (D) Male; (E) Grade1-2; (F) Grade3-4; (G) Stage I-II; (H) Stage III-IV; (I) T stage1-2; (J) T stage3-4; (K) M stage0; (L) M stage1; (M) N stage0; (N) N stage1-X.

FIGURE 9 | Relationship between prognostic model for OS and clinical parameters. (A) Age; (B) Gender; (C) Grade; (D) Stage; (E) T stage; (F) M stage.
TABLE 3 | The relationship between prognostic related RNA binding proteins and clinicopathologic parameters.

| Gene   | Gender | Grade | Stage | T stage | M stage |
|--------|--------|-------|-------|---------|---------|
|        | Female |       |       |         |         |
|        | Male   |       |       |         |         |
| N      |        |       |       |         |         |
| APOBEC3G | t-value |     |       |         |         |
|        | P-value |     |       |         |         |
| AUH   | t-value |     |       |         |         |
|        | P-value |     |       |         |         |
| DAZL  | t-value |     |       |         |         |
|        | P-value |     |       |         |         |
| EIF4A1 | t-value |     |       |         |         |
|        | P-value |     |       |         |         |
| IGF2BP3 | t-value |     |       |         |         |
|        | P-value |     |       |         |         |
| NR0B1 | t-value |     |       |         |         |
|        | P-value |     |       |         |         |
| RPL36A | t-value |     |       |         |         |
|        | P-value |     |       |         |         |
| TRMT1 | t-value |     |       |         |         |
|        | P-value |     |       |         |         |

NA, not available. *Non-parametric Mann–Whitney rank sum test.

FIGURE 10 | Validation the prognostic value of the prognostic RBPs for OS in ccRCC by Kaplan–Meier plotter. (A) APOBEC3G; (B) AUH; (C) DAZL; (D) EIF4A1; (E) IGF2BP3; (F) NR0B1; (G) RPL36A; (H) TRMT1.

DISCUSSION

Malignant tumor is a kind of complex heterogeneous diseases, apart from the classic view that affect cancer or tumor suppressor gene signal channel change decision, It has also been found to be associated with post-transcriptional hijacking by tumor cells, enabling them to rapidly and stably regulate protein expression levels in response to intracellular and extracellular signaling changes to adapt to local microenvironments (Pereira et al., 2017). RBPs are a key player in post-transcriptional events, participating in almost all post-transcriptional regulation, controlling intracellular transcript metabolism and function, and maintaining homeostasis. Multiple studies have reported that RBPs are dysregulated in cancers and regulate cancers...
progression through a variety of mechanisms, including transcriptional and posttranscriptional regulation, genomic change, and posttranslational modification (Patry et al., 2003; Busà et al., 2007; Ortiz-Zapater et al., 2011; Song et al., 2010; Zong et al., 2014; Pérez-Guijarro et al., 2016). However, the expression pattern and role of RBPs in ccRCC are rarely reported. In this study, we systematically analyzed the transcriptome data of ccRCC patients from TCGA database, and identified differential expression RBPs between tumor tissue and normal kidney tissue. We then performed functional enrichment analysis to evaluate their biological function, and performed univariate Cox regression analysis, LASSO regression analysis and multivariate Cox regression analysis to screen prognostic related RBPs and constructed a prognostic risk score model for OS based on these RBPs. In addition, we also built a prognostic model for DFS to predict ccRCC prognosis based on prognostic related RBPs.

The biological function and pathway enrichment analysis of differentially expressed RBPs showed that these genes were significantly enriched in posttranscriptional regulation of gene expression, translational initiation, transposition, protein localization to endoplasmic reticulum, RNA catabolic process, regulation of cellular amide metabolic process, regulation of mRNA metabolic process, gene silencing, ribonucleoprotein granule, mRNA binding, ribosome, polysome, catalytic activity, acting on RNA, translation regulator activity, nuclease activity, double-stranded RNA binding, nucleotidyltransferase activity, RNA transport, and mRNA surveillance pathway, which involved RNA processing, splicing, localization, RNA metabolism and subsequent translation regulation. Previous studies have shown

**FIGURE 11** | The expression status of the prognostic RBPs proteins in ccRCC and normal renal tissues in the HPA database. (A) APOBEC3G; (B) AUH; (C) DAZL; (D) EIF4A1; (E) IGF2BP3; (F) RPL36A; (G) TRMT1.
that multiple RBPs regulatory mechanisms have been identified in cancers, including transcriptional and posttranscriptional regulation, genomic change, and posttranslational modification (Patry et al., 2003; Trabucchi et al., 2009; Ishii et al., 2014; Preca et al., 2015). In lung adenocarcinoma, splice regulator RBM10 inhibits tumor cell proliferation and Notch signaling activity (Bechara et al., 2013). Cancer transcription factor MYC up-regulates the mRNA expression of hnRNPA1 and hnRNPA2 in gliomas, which promotes the synthesis of pyruvate kinase M subtype 2 (PKM2) and participates in glycolytic transformation (Clower et al., 2010). IMP1 has been reported to be elevated in multiple tumors, and reduced IMP1 expression can impair the normal transmission and local translation of adhesive and motif-related target mRNAs (Gu et al., 2009). EIF4E is a key factor in mRNA cycling and translation, and it has been found that EIF4E is overexpressed in a variety of tumors and is associated with poor prognosis (Ruggero et al., 2004). These results suggest that RBPs may influence the occurrence and progression of tumors by regulating multiple biological processes including RNA processing, RNA metabolism, RNA transport, translation regulation and mRNA surveillance pathway.

In addition, we performed univariate Cox regression analysis, LASSO regression analysis and multivariate Cox regression analysis on these differentially expressed RBPs, and 8 prognostic related RBPs including APOBEC3G, AUH, DAZL, EIF4A1, IGF2BP3, NR0B1, RPL36A, and TRMT1 were selected. APOBEC3G, a member of the Apolipoprotein B mRNA editing enzyme-catalyzed polypeptide (APOBEC) family, was found to be overexpressed in renal carcinoma tissues and cell lines (Komohara et al., 2007), consistent with our results. Olson et al. (2018) found that this family is the source of somatic mutations in tumor cells that drive tumor evolution and may be associated with tumor cell recurrence, metastasis, and treatment resistance. AUH was found to be under-expressed in RCC and significantly associated with poorer survival in patients (Zhang et al., 2019), which is similar to our results. The DAZL mutation was found to be associated with testicular cancer (Ruark et al., 2013). The main function of EIF4A1 is to release mRNA structure in combination with other translation factors (Qi et al., 2013). EIF4A1 has been reported to be associated with malignant phenotypes of tumor cells, tumor-specific survival, and susceptibility to therapeutic drugs.
to patient heterogeneity. Second, the model construction and different characteristics of different platforms may lead to the limitation of our study. First, our study is mainly based on a single bioomics information platform, which may affect the reliability of statistics. Finally, the prognostic models for OS and DFS based on prognostic RBPs showed good predictive performance. However, the exact molecular mechanisms of these prognostic RBPs involved in the occurrence, progression, and prognosis of renal cancer are still unclear, and the possible molecular mechanism and biological function need to be further explored.

CONCLUSION

We systematically analyzed the biological function and prognostic value of RBPs in ccRCC by using a variety of bioinformatics techniques. These RBPs may be involved in the pathogenesis, progression, and metastasis of tumors. For the first time, we established prognostic risk score models for OS and DFS based on prognostic RBPs, and revealed they are independent prognostic factors related to OS and DFS in ccRCC patients. Our results are helpful to understand the molecular mechanism of ccRCC from a new perspective and to develop new prognostic markers or therapeutic targets.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://portal.gdc.cancer.gov/.

AUTHOR CONTRIBUTIONS

YW designed the study and performed the data analysis. XW, HF, BH, and BL performed the data analysis. YL, YR, XL, ZL, SW, and JL performed the data analysis and revised the manuscript. TW designed the study and revised the manuscript. All the authors read and approved the final manuscript.

FUNDING

This study was supported by the Medical Youth Top Talent Program of Hubei Provincial.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.617872/full#supplementary-material
REFERENCES

Alishabi, A. M., Vastrad, B., Shaikh, I. A., and Vastrad, C. (2019). Identification of crucial candidate genes and pathways in Glioblastoma multiform by bioinformatics analysis. *Biomolecules* 9:201. doi: 10.3390/biom9050201

Battaglia, M., and Lucarelli, G. (2015). The role of renal surgery in the era of targeted therapy: the urologist's perspective. *Urologia* 82, 137–138. doi: 10.5301/uro.5000105

Bechara, E. G., Sebestyén, E., Bernardis, I., Eyras, E., and Valcárcel, J. (2013). RBM5, a novel prognostic marker that correlates with colon cancer progression and pathogenesis. *Ann. Surg. Oncol.* 16, 3499–3506. doi: 10.1002/ason.20134-009-0648-5

Li, D., Zhou, Y., Chen, Y., Ke, G., Wen, H., and Wu, X. (2014). Decreased expression of EIF4A1 after preoperative brachytherapy predicts better tumor specific survival in cervical cancer. *Int. J. Gynecol. Cancer* 24, 908–915. doi: 10.1097/IGC.0000000000000152

Liao, Y., Wang, J., Jiaehng, E. J., Shi, Z., and Zhang, B. (2019). WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* 47, W199–W205. doi: 10.1093/nar/gkz401

Ljungberg, B., Albigeis, L., Abu-Ghanem, Y., Bensalah, K., Dabestani, S., Fernández-Pello, S., et al. (2019). European association of urology guidelines on renal cell Carcinoma: the 2019 update. *Eur. Urol.* 75, 799–810. doi: 10.1016/j.euro.2019.02.011

Masuda, K., and Kuwano, Y. (2019). Diverse roles of RNA-binding proteins in cancer traits and their implications in gastrointestinal cancers. *Wiley Interdiscip. Rev. RNA* 10, e1520. doi: 10.1002/wrna.1520

Moch, H., Cubilla, A. L., Humphrey, P. A., Reuter, V. E., and Ulbright, T. M. (2016). The 2016 WHO classification of tumours of the urinary system and male genital organs: part a, penile, and testicular tumours. *Eur. Urol.* 70, 93–105. doi: 10.1016/j.euro.2016.02.029

Nagel, S., Venturini, L., Meyer, C., Kaufmann, M., Scherr, M., Drexler, H. G., et al. (2010). Multiple mechanisms induce ectopic expression of LYL1 in subsets of T-ALL cell lines. *Leuk Res.* 34, 521–528. doi: 10.1016/j.leukres.2009.06.020

Nakamura, Y., Suzuki, T., Arai, Y., and Sasano, H. (2009). Nuclear receptor DAX1 in human prostate cancer: a novel independent biological modulator. *Endo. J.* 5, 39–44. doi: 10.1007/endoj.1087-1777

Narla, A., and Ebert, B. L. (2010). Ribosomopathies: human disorders of ribosome dysfunction. *Blood* 115, 3196–3205. doi: 10.1182/blood-2009-10-171829

Oda, T., Tian, T., Inoue, M., Ikeda, J., Qiu, Y., Okumura, M., et al. (2009). Tumorogenic role of orphan nuclear receptor NR0B1 in lung adenocarcinoma. *Ann. J. Pathol.* 175, 1235–1245. doi: 10.2353/ajpath.2009.090010

Olson, E. M., Harris, R. S., and Harki, D. A. (2018). APOBEC Enzymes as Targets for Virus and Cancer Therapy. *Cell Chem. Biol.* 25, 36–49. doi: 10.1016/j.chembiol.2017.10.007

Ortíz-Zapater, E., Pineda, D., Martínez-Bosch, N., Fernández-Miranda, G., Iglesias, M., Alameda, F., et al. (2011). Key contribution of CPEB4-mediated translational control to cancer progression. *Nat. Med.* 17, 83–90. doi: 10.1038/nm.2540

Patry, C., Bouchard, L., Labrecque, P., Gendron, D., Lemieux, B., Toutant, J., et al. (2003). Small interfering RNA-mediated reduction in heterogeneous nuclear ribonucleoparticle A1/A2 proteins induces apoptosis in human cancer cells but not in normal maternal cells. *Cancer Res.* 63, 7679–7688.

Pereira, B., Billaud, M., and Almeida, R. (2017). RNA-binding proteins in cancer: old players and new actors. *Trends Cancer* 3, 506–528. doi: 10.1016/j.trecan.2017.05.003

Pérez-Guijarro, E., Karras, P., Cifalaov, M., Martínez-Herranz, R., Cañón, E., Grata, 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

Preca, B. T., Bajgard, K., Mock, K., Sundararajan, V., Pfannstiel, J., Maurer, J., et al. (2015). A self-enforcing CD44v6/CEB1 feedback loop maintains EMT and stemness properties in cancer cells. *Int. J. Cancer* 137, 2566–2577. doi: 10.1002/ijc.29462

Qi, M., Qi, Y., Ma, Y., He, R., Ji, Y., Sun, Z., et al. (2013). Over-expression of human cytomegalovirus miR-US25-2-3p downregulates eIF4A1 and inhibits HCMV replication. *FERS Lett.* 587, 2266–2271. doi: 10.17160/ferslet.2013.05.057

Rini, B. I., Campbell, S. C., and Escudier, B. (2009). Renal cell carcinoma. *Lancet* 373, 1119–1122. doi: 10.1016/S0140-6736(09)62029-4

Rukar, E., Seal, S., McDonald, H., Zhang, F., Elliott, A., Lau, K., et al. (2013). Identification of nine new susceptibility loci for testicular cancer, including variants near DAZL and PRDM14. *Nat. Genet.* 45, 686–689. doi: 10.1038/ng.1014

Ruggiero, D., Montanaro, L., Ma, L., Xu, W., Londei, P., Cordon-Cardo, C., et al. (2004). The translation factor eIF-4E promotes tumor formation and cooperates with c-Myc in lymphomagenesis. *Nat. Med.* 10, 484–486. doi: 10.1038/ natm1042

Saito, S., Ito, K., Suzuki, T., Utsunomiya, H., Akahira, J., Sugihashi, Y., et al. (2005). Orphan nuclear receptor DAX-1 in human endometrium and its disorders. *Cancer Sci.* 96, 645–652. doi: 10.1111/j.1349-7006.2005.00101.x
Scheper, G. C., van der Knaap, M. S., and Proud, C. G. (2007). Translation matters: protein synthesis defects in inherited disease. Nat. Rev. Genet. 8, 711–723. doi: 10.1038/nrg2142

Seo, D. C., Sung, J. M., Cho, H. J., Yi, H., Seo, K. H., Choi, I. S., et al. (2007). Gene expression profiling of cancer stem cell in human lung adenocarcinoma A549 cells. Mol. Cancer 6:75. doi: 10.1186/1476-4598-6-75

Siegel, R. L., Miller, K. D., and Jemal, A. (2018). Cancer statistics, 2018. CA Cancer J. Clin. 68, 7–30. doi: 10.3322/caac.21442

Song, L., Wang, L., Li, Y., Xiong, H., Wu, J., Li, J., et al. (2010). Sam68 up-regulation correlates with, and its down-regulation inhibits, proliferation and tumourigenicity of breast cancer cells. J. Pathol. 222, 227–237. doi: 10.1002/path.2751

Tamma, R., Rutigliano, M., Lucarelli, G., Annese, T., Ruggieri, S., Cascardi, E., et al. (2019). Microvascular density, macrophages, and mast cells in human clear cell renal carcinoma with and without bevacizumab treatment. Urol. Oncol. 37, e1–e355. doi: 10.1016/j.urolonc.2019.01.025

Trabucchi, M., Briata, P., Garcia-Mayoral, M., Haase, A. D., Filipowicz, W., Ramos, A., et al. (2009). The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. Nature 459, 1010–1014. doi: 10.1038/nature08025

Tschirdewahn, S., Panic, A., Püllen, L., Harke, N. N., Hadaschik, B., Riesz, P., et al. (2019). Circulating and tissue IMP3 levels are correlated with poor survival in renal cell carcinoma. Int. J. Cancer 145, 531–539. doi: 10.1002/ijc.32124

Wang, T., Fan, L., Watanabe, Y., McNeill, P. D., Moulton, G. G., Bangur, C., et al. (2003). L523S, an RNA-binding protein as a potential therapeutic target for lung cancer. Br. J. Cancer 88, 887–894. doi: 10.1038/sj.bjc.6600806

Wei, W., Cao, W., Zhan, Z., Yan, L., Xie, Y., and Xiao, Q. (2019). MiR-1284 suppresses gastric cancer progression by targeting EIF4A1. Oncol. Targets Ther. 12, 3965–3976. doi: 10.2147/OTT.S191015

Zhang, B., Wu, Q., Wang, Z., Xu, R., Hu, X., Sun, Y., et al. (2019). The promising novel biomarkers and candidate small molecule drugs in kidney renal clear cell carcinoma: evidence from bioinformatics analysis of high-throughput data. Mol. Genet. Genomic Med. 7:e607. doi: 10.1002/mg3.607

Zhu, Z., He, A., Lin, L., Xu, C., Cai, T., and Lin, J. (2020). Biological functions and prognostic value of RNA binding proteins in clear cell renal cell Carcinoma. J. Cancer 11, 6591–6600. doi: 10.7150/jca.49175

Zong, F. Y., Fu, X., Wei, W. J., Luo, Y. G., Heiner, M., Cao, L. J., et al. (2014). The RNA-binding protein QKI suppresses cancer-associated aberrant splicing. PLoS Genet. 10:e1004289. doi: 10.1371/journal.pgen.1004289

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 © 2021 Wu, Wei, Feng, Hu, Liu, Luan, Ruan, Liu, Liu, Wang, Liu and Wang. 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.