Research Article

Identification of 6 Hub Proteins and Protein Risk Signature of Colorectal Cancer

Taohua Yue [iD], Cheng Liu, Jing Zhu, Zhihao Huang, Shihao Guo, Yuyang Zhang, Hao Xu, Yucun Liu, Pengyuan Wang [iD], and Shanwen Chen [iD]

Division of General Surgery, Peking University First Hospital, Peking University, 8 Xi Shiku Street, Beijing 100034, China

Correspondence should be addressed to Pengyuan Wang; pengyuan_wang_bjmu@163.com
and Shanwen Chen; shanwen_chen_bjmu@163.com

Taohua Yue and Cheng Liu contributed equally to this work.

Background. Colorectal cancer (CRC) is the second most common cause of cancer death in the United States and the third most common cancer globally. The incidence of CRC tends to be younger, and we urgently need a reliable prognostic assessment strategy. *Methods.* Protein expression profile and clinical information of 390 CRC patients/samples were downloaded from the TCPA and TCGA database, respectively. The Kaplan-Meier, Cox regression, and Pearson correlation analysis were applied in this study. *Results.* Based on the TCPA and TCGA database, we screened 6 hub proteins and first constructed protein risk signature, all of which were significantly associated with CRC patients' overall survival (OS). The risk score was an independent prognostic factor and significantly related with the size of the tumor in situ (T). 6 hub proteins were differentially expressed in cancer and normal tissues and in different CRC stages, which were validated at the ONCOMINE database. Next, 40 coexpressed proteins of 6 hub proteins were extracted from the TCPA database. In the protein-protein interaction (PPI) network, HER1, HER2, and CTNNB1 were at the center. Function enrichment analysis illustrated that 46 proteins were mainly involved in the EGFR (HER1) tyrosine kinase inhibitor resistance pathway. *Conclusion.* Studies indicated that 6 hub proteins might be considered as new targets for CRC therapies, and the protein risk signature can be used to predict the OS of CRC patients.

1. Introduction

According to the 2020 Colorectal Cancer Statistics, colorectal cancer (CRC) is the second most common cancer death in the United States [1]. The latest report points out that the age of onset of CRC is getting younger, with the median age dropping from 72 years in 2001-2002 to 66 years in 2015-2016 [2]. Especially for young CRC patients, we need tailored clinical management strategies, recognize the patient's risk in early diagnosis, reduce the side effects of treatment in low-risk patients, and perform adjuvant chemoradiotherapy other than surgery in high-risk patients.

Since the completion of the Human Genome Project and the rise of microarray profiling and genome-wide sequencing, increasing studies have predicted the survival of cancer patients at the genetic level [3–5]. Based on the Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO), hundreds of differentially expressed, metastasis-related genes and survival-related genes have been identified in CRC tissues and cell lines [6–8]. Compared with utilizing a single gene to predict survival, the risk signatures or models constructed by multiple genes can predict the overall survival (OS) and disease-free survival more accurately [9]. Unfortunately, protein risk signature has never been constructed to predict the prognosis of CRC patients.

The Cancer Proteome Atlas (TCPA) database provided the protein expression profile by integrating RPPA chip data from TCGA and several independent tumor research projects [10]. In our research, based on the Kaplan-Meier method and Cox regression analysis, 6 hub proteins were identified, and a protein risk signature was firstly constructed. The overall survival (OS) of the high-risk group

was significantly shorter than that of the low-risk group. The receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC) further confirmed the accuracy of the risk signature. Pearson correlation analysis between the signature and clinical parameters indicated that the protein risk score of T_3_4 was significantly higher than that of T_1_2. With the help of the UALCAN website, we further found that 6 hub proteins were differentially expressed in cancer and normal tissues and in different CRC stages, which were validated at the ONCOMINE database. In order to mine the molecular characteristics of 6 hub proteins, we performed protein-protein interaction (PPI) and enrichment analysis based on the 6 hub proteins and their coexpressed proteins. We found that these proteins were mainly enriched in EGFR (HER1) tyrosine kinase inhibitor resistance, regulation of DNA metabolic process, and pathways in cancers, all of which were directly associated with tumorigenesis.

In conclusion, protein risk signature, a novel prognostic assessment tool, had the potential to predict the outcomes of CRC patients in clinical practice, and 6 hub proteins were expected to become novel therapeutic targets in the future.

2. Materials and Methods

2.1. Ethics Statement. All data in this study were obtained from online public databases and did not involve any in vitro or in vivo experiments.

2.2. Data Mining. The protein expression profiles and the matching clinical information of colorectal cancer (CRC) patients were downloaded from the Cancer Proteome Atlas Portal (TCPA) (https://www.tcpaportal.org/tcpa/) [11] and the Cancer Genome Atlas website (TCGA) (https://portal.gdc.cancer.gov/) [12], respectively. The TCPA is a comprehensive resource for accessing, visualizing, and analyzing functional proteomics of patient tumor samples and cancer cell lines. The TCGA is currently the largest cancer genetic information database and molecularly characterized over 20,000 primary cancer and matched normal samples covering 33 cancer types [13]. All the data in this study is the latest data from the above official website sources.

2.3. Identification of Candidate Proteins. The TCPA database had converted the raw data into a recognizable format [14]. Candidate proteins associated with the overall survival (OS) were extracted based on both univariate Cox proportional hazard regression analysis and Kaplan-Meier [15]. Proteins with a p value below 0.05 were defined as significant. Proteins with HR < 1 were defined as candidate protective proteins, while proteins with HR > 1 were considered candidate risky proteins [16].

2.4. Construction of the Prognostic Risk Signature. Based on the step-wise multivariate Cox proportional hazard regression analysis, we obtained 6 hub proteins from 24 candidate risk proteins [17]. The prognostic risk signature was built by combining the expression values of 6 hub proteins weighted by their regression coefficients. The median risk score was set as a cutoff value and divided CRC patients into high- and low-risk groups [18]. R "survival" package (https://CRAN.R-project.org/package=survival) was used to assess the significance of the OS difference between high- and low-risk groups.

2.5. Performance Assessment. To measure the performance of our hub protein risk signature, the receiver operating characteristic (ROC) curve and the corresponding areas under the ROC curve (AUC) were produced using the R "survivalROC" packages [19]. The univariate and multivariate Cox proportional hazard regression analyses were performed to evaluate the independent prognostic potential of protein risk signature. Multivariate Cox analysis adjusted the influences of age, gender, pathological American Joint Committee on Cancer (AJCC) stage, tumor size in situ (T), lymph node metastasis (N), and distant metastasis status (M) on risk signature [20].

2.6. Clinical Parameter Correlation. Based on the TCGA clinical information, we performed the Pearson correlation analysis between risk signature and age, gender, stage, T, N, and M [21].

2.7. Differential Expression Analysis. UALCAN is a user-friendly and interactive web resource for analyzing cancer OMICS data based on level 3 RNA-seq data and clinical data of 31 cancer types from the TCGA database [22]. The differential expression of 6 hub proteins and their encoding genes were analyzed at the UALCAN website (http://ualcan.path.uab.edu/) [23].

The ONCOMINE database (http://www.oncomine.org/) is a cancer microarray database and integrated data-mining platform designed to facilitate discovery from genome-wide expression analysis [24]. In our research, transcriptional data of 6 encoding genes between CRC tissues and normal tissue were obtained from this database. Thresholds were designed as follows: p value: 0.001, fold change: 1.5, gene rank: top 10%, data type: mRNA.

2.8. Protein Coexpression Analysis. Based on 6 hub proteins, we performed Pearson correlation analysis and found their coexpressed proteins. The correlation filtering criteria were p less than 0.001, and Pearson correlation coefficient (PCC) greater than 0.4. PCC > 0 meant a positive correlation with hub protein, and PCC < 0 meant a negative correlation. Next, we generated a Sankey diagram using ggplot2 and ggalluvial package among 6 hub proteins and their coexpressed proteins.

2.9. Molecular Network and Functional Enrichment Analysis. Protein-protein interaction (PPI) and functional enrichment analysis were constructed using the STRING database (https://string-db.org/), Cytoscape software version 3.7.1 [25], and the Metascape (https://metascape.org/gp/index.html) [26]. The STRING database helps to explore the interaction network between proteins and discover core regulatory proteins. If the network contained between 3 and 500 proteins, the Molecular Complex Detection (MCODE) algorithm would be applied to identify densely connected network components [26]. To clearly illustrate the functions and molecular pathways of 6 hub proteins and their
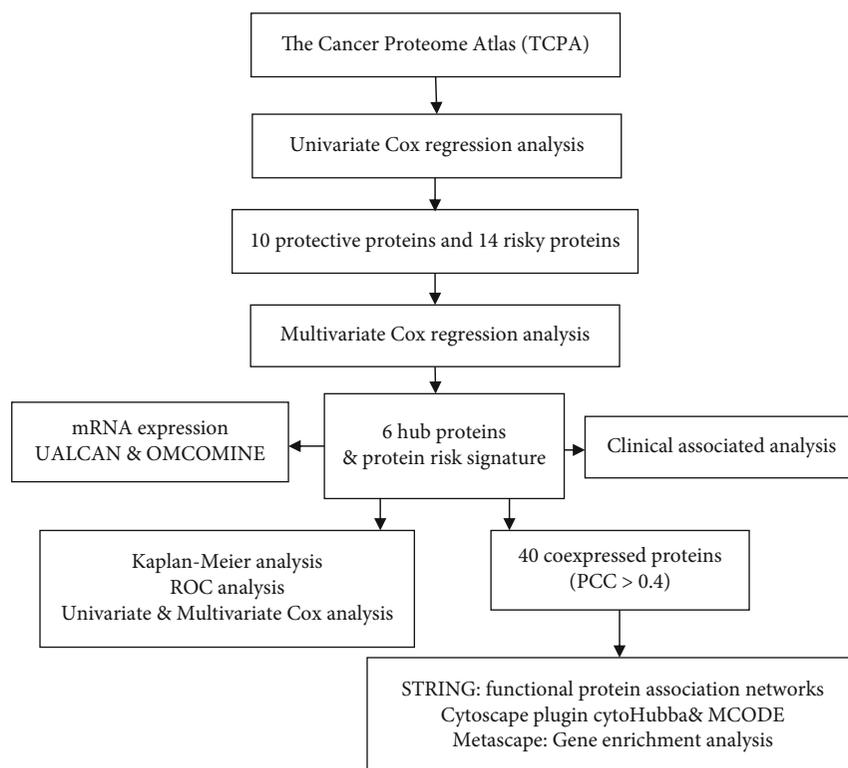
FIGURE 1: Flow chart of our study.

coexpressed proteins, we conducted the Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis at the Metascape website [27].

2.10. Statistical Analysis. In this study, all statistical analyses were performed using *R* software (version 3.6.1; https://www.r-project.org/). Significance was defined as $p < 0.05$.

3. Results

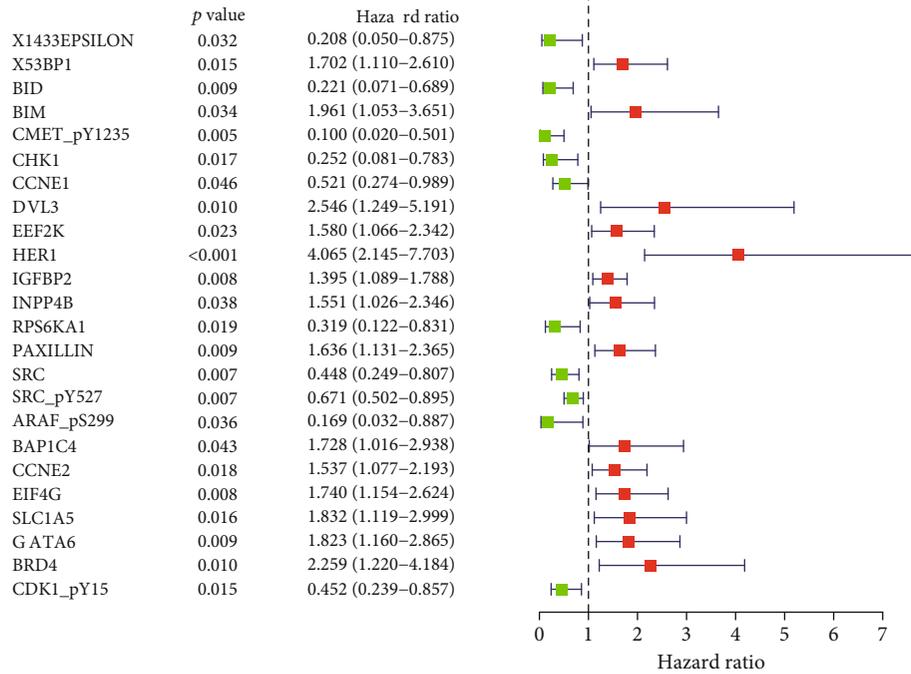*3.1. Identification of Candidate Survival-Related Proteins.* The workflow of our study was illustrated in Figure 1. At the TCPA database, we downloaded protein expression profile of 390 CRC patients/tissues. The corresponding clinical information was downloaded from the TCGA database. There was no requirement for ethical approvals. As predicting the prognosis of CRC that was critical for cancer patients, we performed univariate Cox regression analysis and explored 24 candidate survival-associated proteins, including 10 protective proteins and 14 risky proteins, all of which were displayed in the forest plot (Figure 2(a)) and volcano plot (Figure 2(b)). Risky proteins meant that the higher its expression, the higher the patient's risk of death and the shorter the overall survival (OS).

3.2. Identification of Hub Proteins. To extract hub proteins that were actively involved in the onset and progression of CRC, we further performed multivariate Cox regression analysis and finally identified 6 survival-related proteins, CCNE1, HER1, INPP4B, RPS6KA1, SRC, and SLC1A5, and named them hub proteins ($p < 0.05$). All of the identified
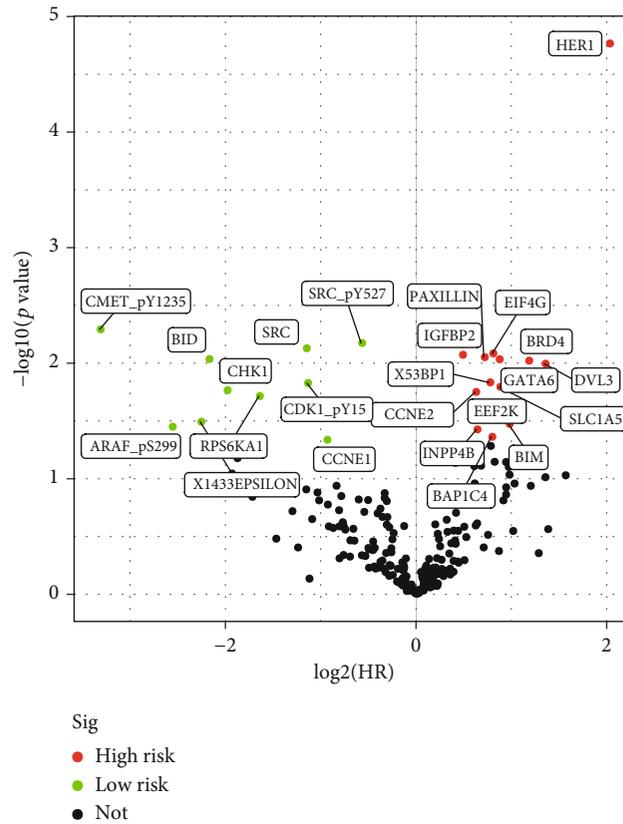
hub proteins were significantly related to the OS of CRC patients and were potential prognostic markers for monitoring patients' outcomes (Figure 3).

3.3. Construction and Validation of a Prognostic Signature. Based on multivariate Cox regression analysis, we further constructed a protein risk signature, and the formula was as follows: [expressions of CCNE1 × (−0.7475)] + [expressions of HER1 × 0.8318] + [expressions of INPP4B × 0.3337] + [expressions of RPS6KA1 × (−0.8763)] + [expressions of SRC × (−0.7350)] + [expressions of SLC1A5 × 0.6579]. Based on the median of the risk score, we separated CRC patients into 2 groups, the high- and low-risk group (Figure 4(a)). CRC patients belonging to the high-risk group doomed to a poor prognosis, and the number of patients who died of CRC was significantly more than that of the low-risk group (Figure 4(b)). The differential expression of 6 hub proteins between the high- and low-risk groups was illustrated in the heat map. HER1, INPP4B, and SLC1A5 were highly expressed in the high-risk group, while CCNE1, RPS6KA1, and SRC were highly expressed in the low-risk group (Figure 4(c)). The protein risk signature was significantly associated with the OS of CRC patients. The OS in the high-risk group was significantly shorter than that of the low-risk group (Figure 4(d)).

The area under ROC curve (AUC) of the receiver operating characteristic (ROC) curve was 0.694, suggesting that its predictive effectiveness of the OS was moderate (Figure 4(e)). In particular, it was worth mentioning that the AUC of our risk signature was larger than the existing clinic parameters stage, *T*, *N*, and *M*, which vigorously filled

	p value	Hazard ratio	
X1433EPSILON	0.032	0.208 (0.050–0.875)	
X53BP1	0.015	1.702 (1.110–2.610)	
BID	0.009	0.221 (0.071–0.689)	
BIM	0.034	1.961 (1.053–3.651)	
CMET_pY1235	0.005	0.100 (0.020–0.501)	
CHK1	0.017	0.252 (0.081–0.783)	
CCNE1	0.046	0.521 (0.274–0.989)	
DVL3	0.010	2.546 (1.249–5.191)	
EEF2K	0.023	1.580 (1.066–2.342)	
HER1	<0.001	4.065 (2.145–7.703)	
IGFBP2	0.008	1.395 (1.089–1.788)	
INPP4B	0.038	1.551 (1.026–2.346)	
RPS6KA1	0.019	0.319 (0.122–0.831)	
PAXILLIN	0.009	1.636 (1.131–2.365)	
SRC	0.007	0.448 (0.249–0.807)	
SRC_pY527	0.007	0.671 (0.502–0.895)	
ARAF_pS299	0.036	0.169 (0.032–0.887)	
BAP1C4	0.043	1.728 (1.016–2.938)	
CCNE2	0.018	1.537 (1.077–2.193)	
EIF4G	0.008	1.740 (1.154–2.624)	
SLC1A5	0.016	1.832 (1.119–2.999)	
G ATA6	0.009	1.823 (1.160–2.865)	
BRD4	0.010	2.259 (1.220–4.184)	
CDK1_pY15	0.015	0.452 (0.239–0.857)	

(a)

(b)

FIGURE 2: 24 candidate proteins were significantly correlated to prognosis in CRC. (a) Forest plot of 24 candidate proteins. (b) Volcano plot of 24 candidate proteins. The green dots indicated low-risk proteins, while the red dots represented high-risk proteins. $p < 0.05$.
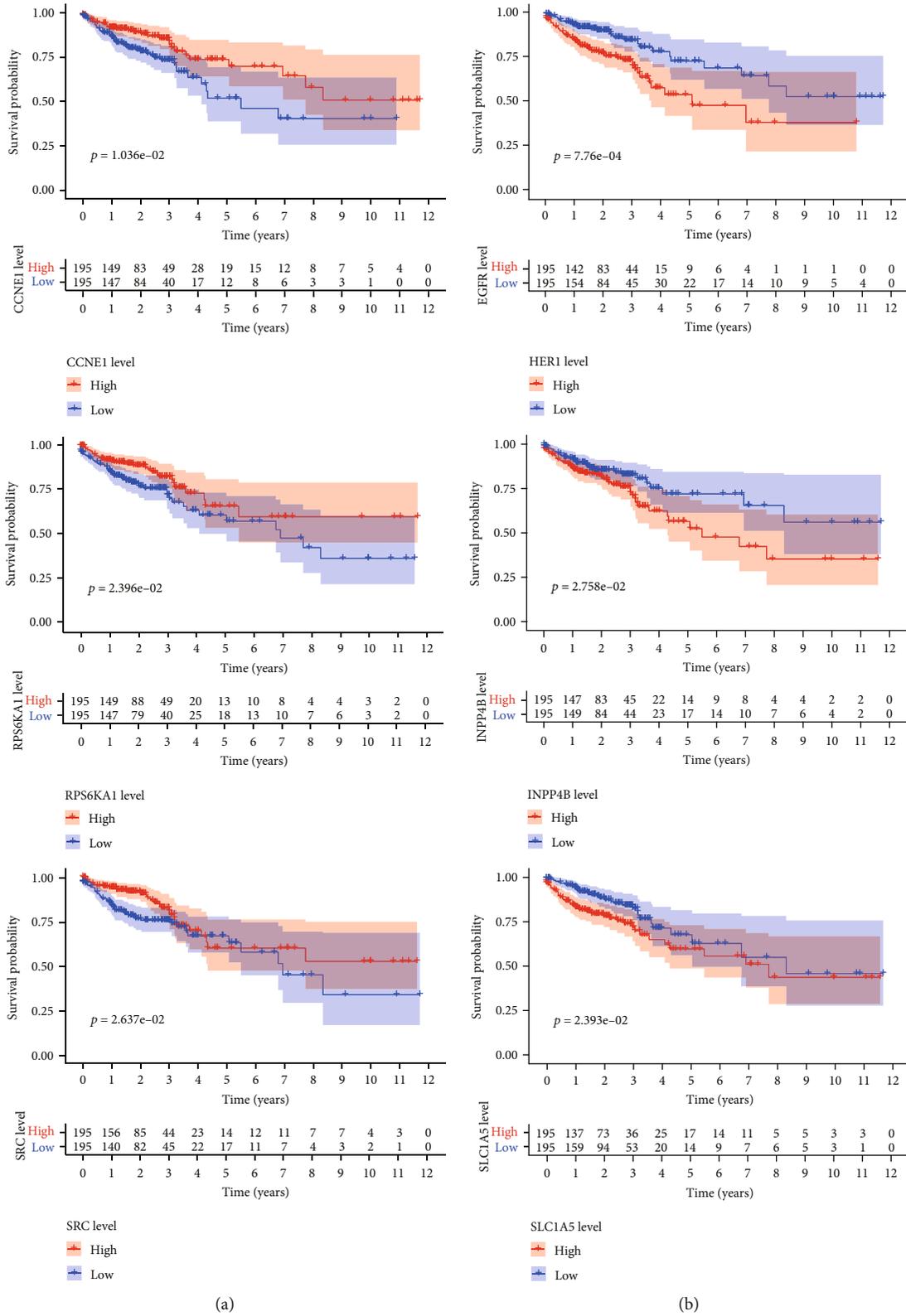
Figure 3: Kaplan-Meier survival curve among high- and low-risk groups based on 6 hub proteins. (a) The higher the expression of CCNE1, RPS6KA1, and SRC, the longer the OS of CRC patients. (b) The higher the expression of HER1, INPP4B, and SLC1A5, the shorter the OS of CRC patients.
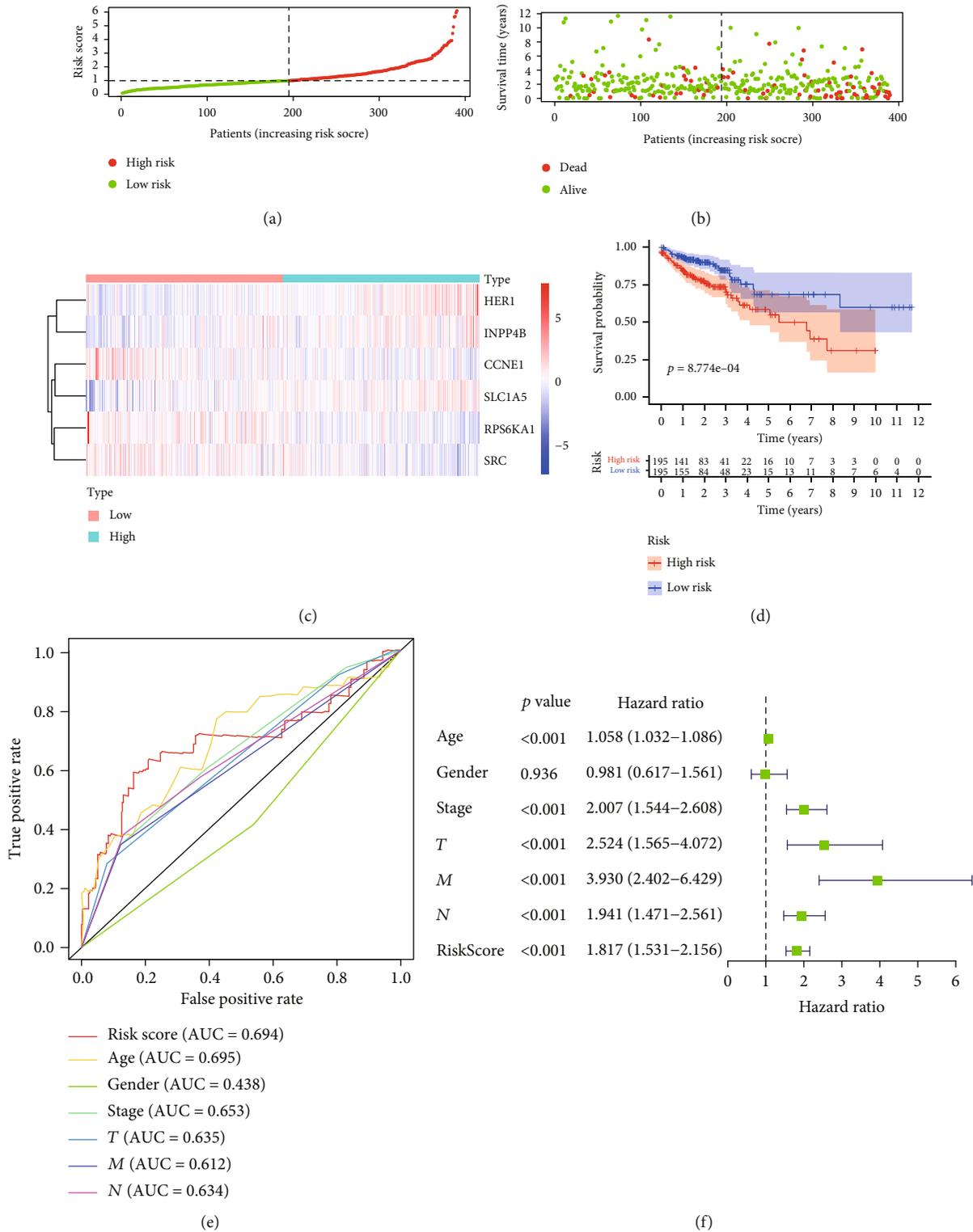
(a)

(b)

(c)

(d)

(e)

(f)

Figure 4: Continued.

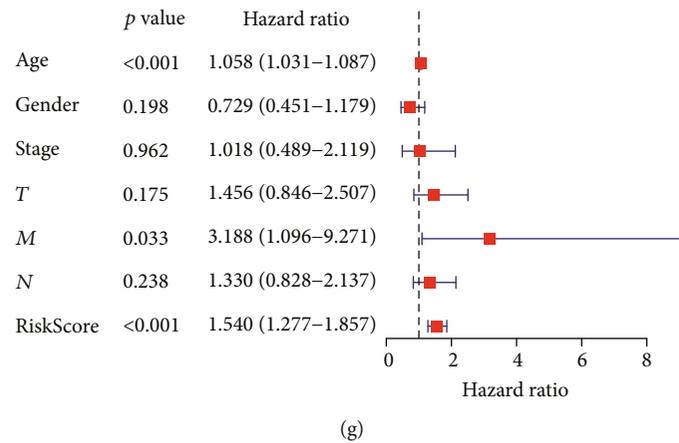|	p value	Hazard ratio
Age	<0.001	1.058 (1.031–1.087)
Gender	0.198	0.729 (0.451–1.179)
Stage	0.962	1.018 (0.489–2.119)
T	0.175	1.456 (0.846–2.507)
M	0.033	3.188 (1.096–9.271)
N	0.238	1.330 (0.828–2.137)
RiskScore	<0.001	1.540 (1.277–1.857)

(g)

FIGURE 4: Construction and validation of protein risk signature. (a) Risk score distribution of CRC patients based on the median of risk score (low, green; high, red). (b) Scatterplots of CRC patients with different survival status in both groups. (c) Heatmap of expression profiles of included 6 hub proteins. (d) Kaplan-Meier curve of the high-risk (red) and low-risk (blue) CRC patients and patients in the high-risk group had a shorter overall survival. (e) The predictive accuracy of risk signature. (f) Univariate and (g) multivariate Cox regression analysis to verify that the risk score was an independent prognostic factor.

the blanks of existing clinical prognosis. To investigate our risk signature's independence, we performed univariate (Figure 4(f)) and multivariate Cox regression analysis (Figure 4(g)) and found that protein risk signature was an independent and reliable prognostic factor. The clinical characteristics of TCGA patients were shown in Table 1. 36 cases with incomplete clinical information were eliminated, and the remaining 354 CRC patients were finally evaluated. The median age of 354 CRC patients was 68 years. The OS was compared between high- and low-risk groups of different clinical characteristics (Figure 5).

3.4. Clinical Relevance Assessment. To understand our protein signature's clinical relevance, we investigated the relationships between the risk signature and age, gender, stage, T, N, and M. The risk score was significantly higher in seniors and advanced T stage cases (Figure 6). As shown in Figure 7(a), expressions of 6 encoding genes were significantly related to CRC patients' pathological stage. In the late stage, expressions of CCNE1, INPP4B, SRC, and SLC1A5 were significantly increased to varying degrees while expressions of HER1 and RPS6KA1 were significantly reduced to varying degrees. The same conclusion was reached at the protein level (Figure 7(b)).

3.5. Differential Expression Analysis of 6 Hub Proteins. Compared with normal samples, mRNA expressions of CCNE1, INPP4B, SRC, and SLC1A5 increased significantly in primary tumors. At the protein expression levels, we found that expressions of CCNE1 and SLC1A5 increased significantly in CRC tissues. Expressions of HER1 and RPS6KA1 in CRC were significantly reduced at both mRNA and protein levels (Figure 8). Next, we validated mRNA expressions of 6 encoding genes at the OMCOMINE website and finally drew the same conclusion (Figure 9).

3.6. Pearson Correlation Analysis of the TCPA Database. In organisms, proteins that have the same function or exist in

TABLE 1: Clinical characteristics of 354 CRC patients included in our study.

Clinical characteristics	No. of patients (%)
Age	
≤68	181 (51.13)
>68	173 (48.87)
Gender	
Male	186 (52.54)
Female	168 (47.46)
Stage-AJCC	
Stage_I_II	208 (58.76)
Stage_III_IV	146 (41.24)
T	
T_1_2	65 (18.36)
T_3_4	289 (81.64)
N	
N0	214 (60.45)
N_1_2_3	140 (39.55)
M	
M0	302 (85.31)
M1	52 (14.69)

the same pathway tend to be coexpressed. Based on the TCPA protein expression profile and 6 hub proteins, we extracted 11 significant coexpressed proteins (pearson correlation coefficient (PCC) was approximately equal to 0.5) (Figure 10(a)) and 29 coexpression proteins (PCC > 0.4). All of them were displayed in the Sankey diagram (Figure 10(b)).

3.7. Molecular Network and Functional Enrichment Analysis. Based on these 46 proteins, we performed protein-protein interaction (PPI) and functional enrichment analysis. PPI network demonstrated that HER1, HER2, and CTNNB1
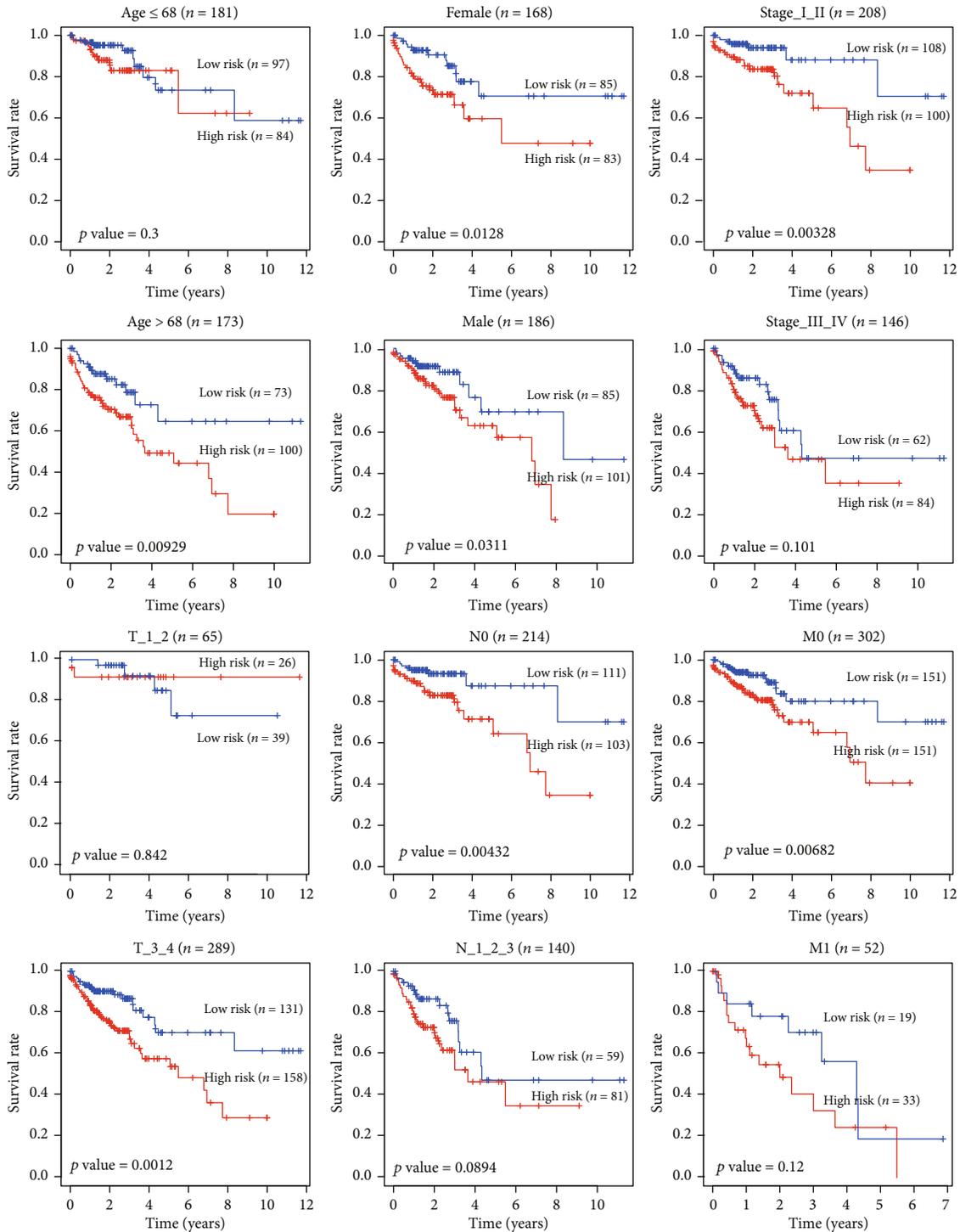
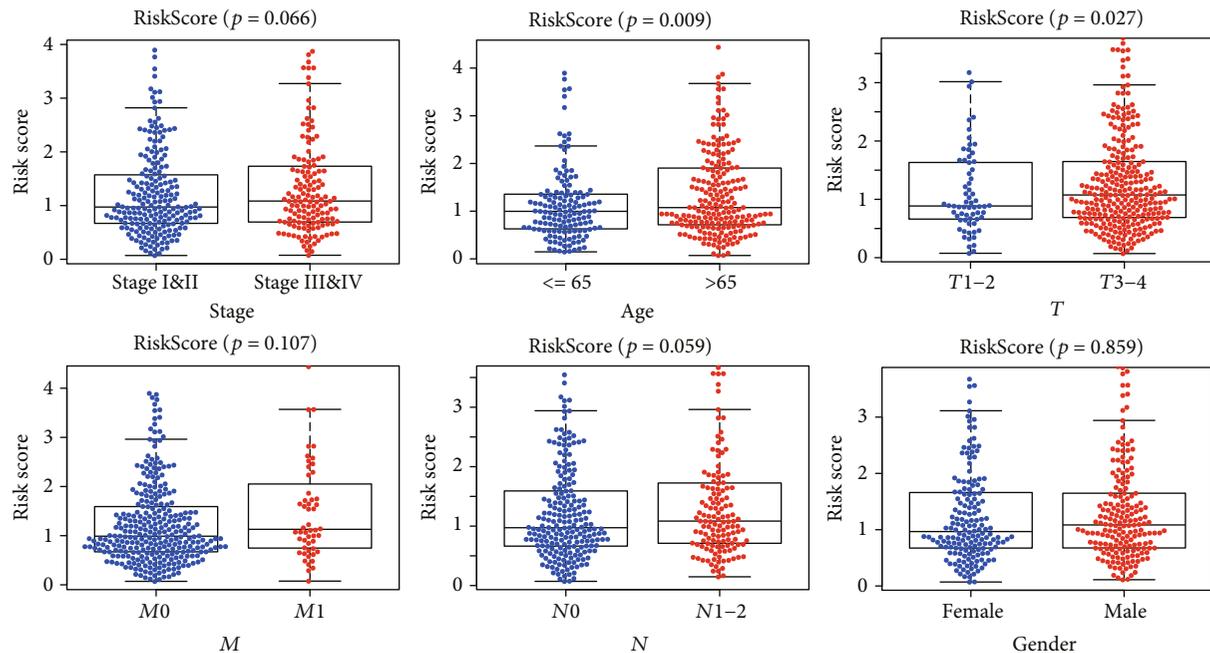FIGURE 5: The OS of high- and low-risk groups in different clinical characteristics.

FIGURE 6: Relationships between protein risk signature and clinical characteristics. The risk score was significantly higher in T_3_4 than T_1_2, and the greater the age, the higher the risk score ($p < 0.05$). In terms of stage and N, the risk score had only a little short of significance ($p > 0.05$).

were the top 3 nodes ranked by degrees calculated by Cytoscape plugin cytoHubba (Figure 11(a)). The Molecular Complex Detection (MCODE) algorithm had been applied to identify the densely connected network components, and the MCODE_1 was displayed in Figure 11(b). At the Metascape, we performed PPI enrichment analysis, and the top 3 clusters were displayed in Table 2. Nodes of MCODE_1 were mainly enriched in pathways in cancer (Table 2). Gene enrichment analysis mainly focused on EGFR (HER1) tyrosine kinase inhibitor resistance and regulation of the DNA metabolic process (Figure 11(c)). Enrichment analysis results at the protein, and gene levels all indicated that the prognostic-related proteins of CRC were related to EGFR tyrosine kinase inhibitor resistance.

4. Discussion

Incidence and mortality of colorectal cancer (CRC) in the United States rank third and second, respectively. The latest report states that CRC is getting younger and younger, and mortality increases with the delay of discovery. Advances in sequencing and microarray technology and the opening of various online databases have promoted the diagnosis and treatment of diseases at the genomic level.

Reviewing the previous literature on the prognosis of CRC, 4 gene signatures [28], 9 gene signatures [7], hypoxia-related signature [29], autophagy score signature [30, 31], somatic mutation signatures [32], metabolism-related signature [33], chemokine/chemokine receptor signature [34], and immune-related signature [35] had been constructed to predict the OS of CRC patients. However, these bioinformatic analyses only predict the prognosis of CRC patients at the

RNA level, and there are few studies on the protein prognostic signature. Besides, these previous RNA bioinformatic studies did not identify the same protein-coding genes of our signature.

In this study, based on the TCPA and TCGA database, we performed K-M and Cox regression analysis. 6 hub proteins were screened, all of which were significantly associated with the overall survival (OS) of CRC patients. Based on these 6 hub proteins, we first developed a protein-related prognostic signature, which was an independent prognostic factor and significantly associated with CRC patients' OS. The receiver operating characteristic (ROC) curve and the areas under the ROC curve (AUC) further validated its accuracy. And our protein risk signature's performance was better than the existing clinical-pathological parameters, including T, N, M, and stage. Based on the median of the risk score, we divided CRC patients into high- and low-risk groups and found that CRC patients in the high-risk group had significantly more deaths than the low-risk group. Differential expression analysis found that 6 hub proteins were significantly differentially expressed in CRC patients with different stages and cancer tissues than normal tissues. Given the potential clinical significance of 6 hub proteins, we performed protein-protein interaction (PPI) and enrichment analysis on 6 hub proteins and their coexpressed proteins. The relationship between EGFR tyrosine kinase inhibitor resistance and CRC's development and prognosis urgently needs more research.

The CCNE1 protein (Cyclin E1) belongs to the highly conserved cyclin family, which forms a complex with CDK2 and functions as its regulatory subunit, whose activity is necessary for the G1/S transition in the cell cycle [36]. Previous studies have shown that CCNE1 amplification and
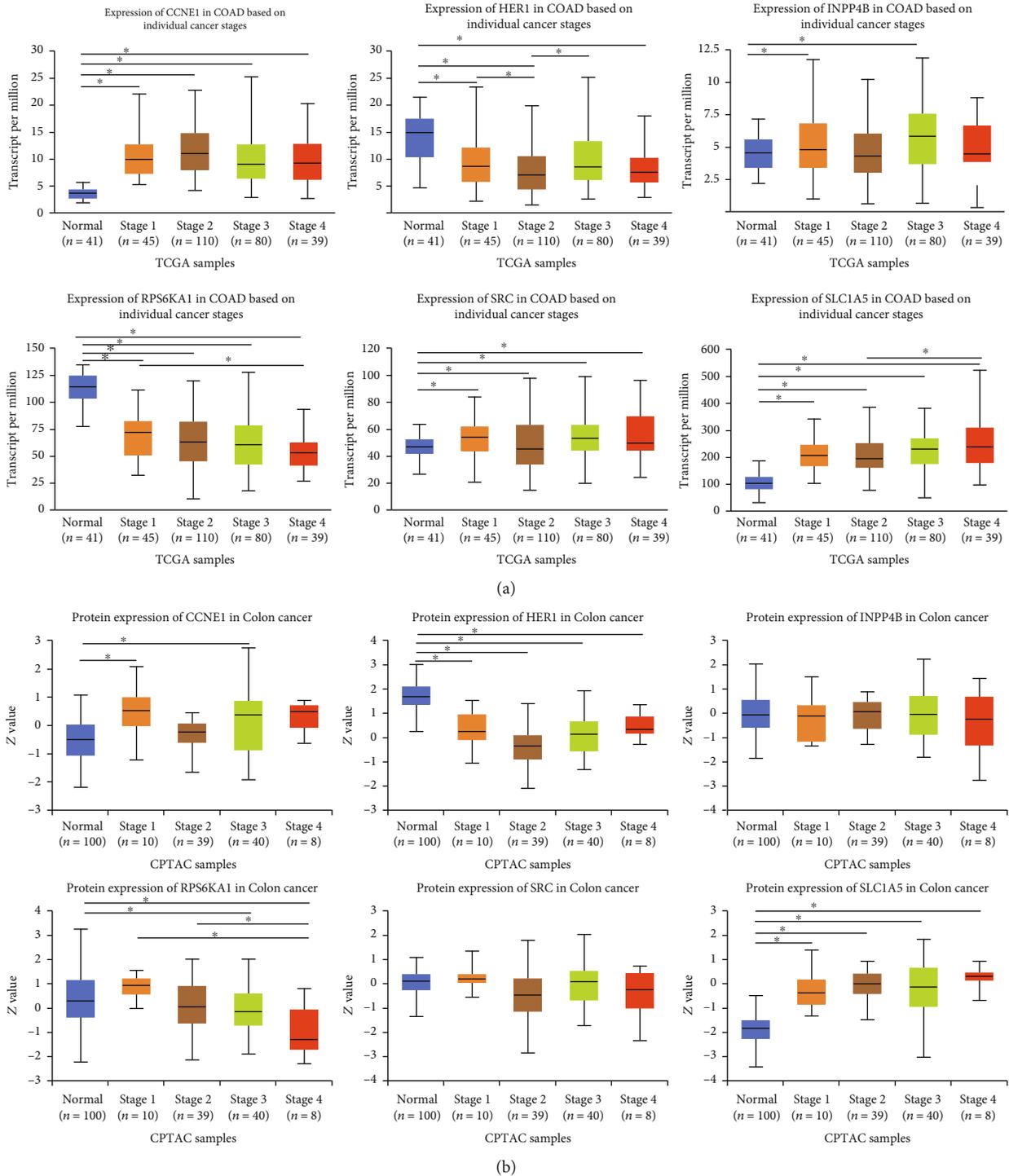
(a)

(b)

Figure 7: Relationships between expressions of 6 encoding genes and hub proteins and individual cancer stages of CRC patients (UALCAN). (a) mRNA expressions of HER1, RPS6KA1, and SLC1A5 were significantly associated with CRC patients' individual cancer stages. Patients who were in advanced stages tended to express higher HER1 and SLC1A5 and lower RPS6KA1. (b) At the protein expression level, the expression of the RPS6KA1 protein was lower in the advanced stage. $^{*}p < 0.05$.
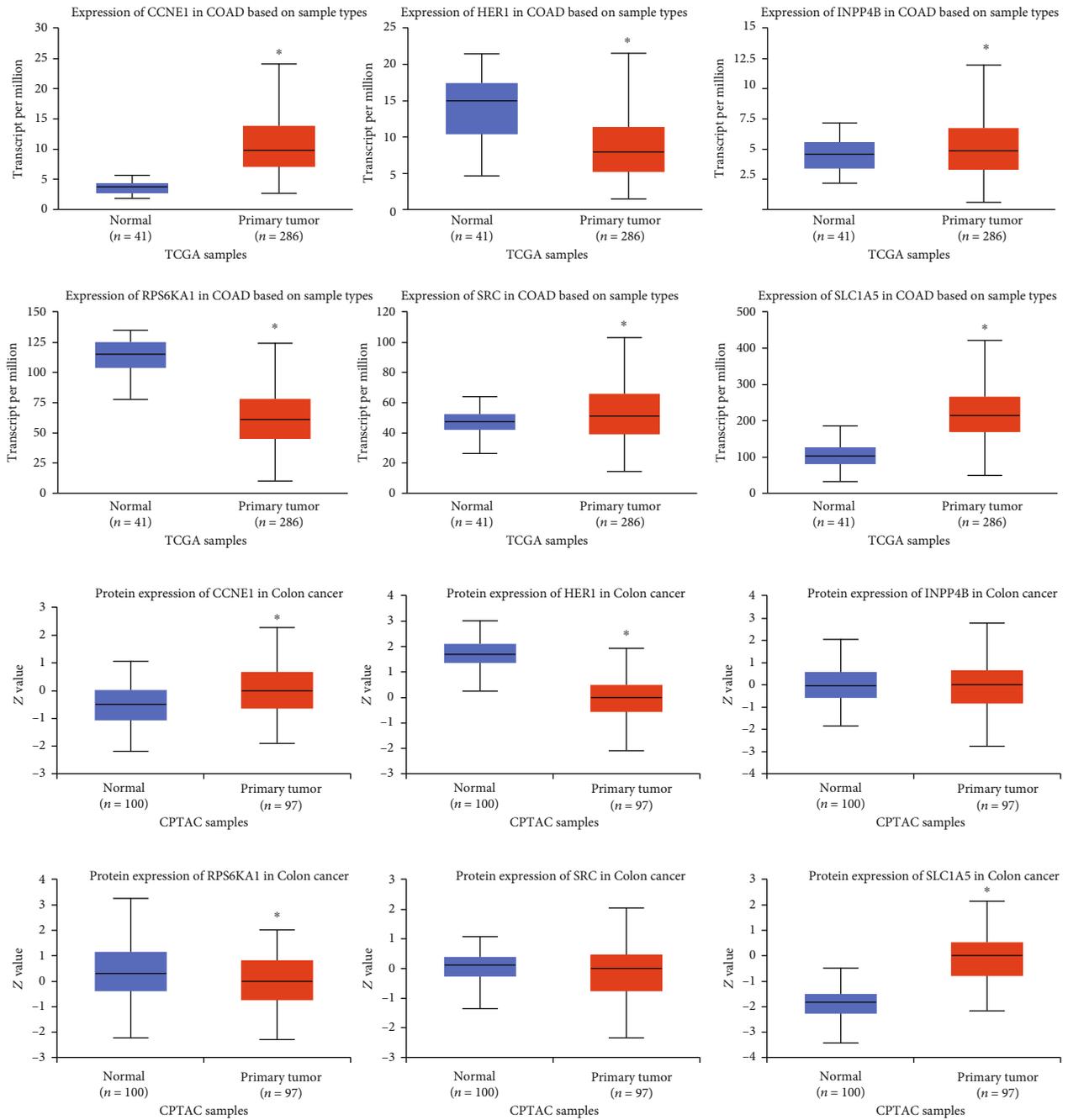
FIGURE 8: Expressions of 6 encoding genes and hub proteins between CRC tissues and normal tissues (UALCAN). CCNE1, INPP4B, SRC, and SLC1A5 were significantly overexpressed in CRC tissues while HER1 and RPS6KA1 were significantly underexpressed in CRC tissues. The mRNA expressions of INPP4B and SRC did not match their protein expression. $^{*}p < 0.05$.

Analysis Type by Cancer	Cancer vs. Normal CCNE1		Cancer vs. Normal HER1		Cancer vs. Normal INPP4B		Cancer vs. Normal RPS6KA1		Cancer vs. Normal SRC		Cancer vs. Normal SLC1A5	
Bladder Cancer	2		1		1		4		1			
Brain and CNS Cancer	1	2	11		1	2	1		1		1	
Breast Cancer	8		1	23	5		13		4		2	1
Cervical Cancer	4						3					
Colorectal Cancer	11			5	1			15	4		3	
Esophageal Cancer	1						5		1			
Gastric Cancer	4			1							1	
Head and Neck Cancer	4		2		3		2		5		1	
Kidney Cancer			7	4	4	1			2	1		
Leukemia		8				1	1	5			1	
Liver Cancer	1					1			1			
Lung Cancer	13		4	1	1			2	1		3	1
Lymphoma	13		5	1	3	2	4	1	5	1	4	
Melanoma			1	5	1				2		1	
Myeloma		1			1			1	1	2	2	
Other Cancer	3	2					4	1				
Ovarian Cancer	7			2	1							
Pancreatic Cancer	1	2			4				1	1		
Prostate Cancer			1	1			2					
Sarcoma	6		1	3	1			1	1	2		3
Significant Unique Analyses	79	15	33	47	25	12	36	25	30	9	22	5
Total Unique Analyses	350		359		353		365		315		338	

FIGURE 9: Transcriptional expressions of 6 encoding genes in 20 types of cancers at the ONCOMINE database. The difference in the mRNA expression was compared by Student's t-test. Red represented the overexpression while blue represented the low expression. The darker the color, the greater the significance of the over- or underexpressed genes.

overexpression result in chromosome instability and further lead to tumorigenesis, which is associated with poor prognosis in CRC [37].

Epidermal growth factor receptor (EGFR, HER1), belonging to the protein kinase superfamily, is a transmembrane glycoprotein that binds to epidermal growth factor, which induces receptor dimerization and tyrosine autophosphorylation and leads to cell proliferation [38, 39]. Another mechanism for the HER1 activation is the overproduction of its ligands. Previous studies have revealed that targeting HER1 can inhibit the proliferation and induce apoptosis of CRC cells [40]. Adding HER1 mAb to chemotherapy or best supportive care improves progression-free survival (moderate to high-quality evidence), the OS (moderate evidence), and tumor response rate (moderate to high-quality evidence) in CRC patients [41].

The inositol polyphosphate-4-phosphatase type IIB (INPP4B) protein belongs to the enzyme of the phosphatidylinositol signaling pathway. The mechanism of INPP4B is to remove the phosphate group of the inositol ring from inositol 3,4-bisphosphate. Previous studies have shown that INPP4B restrains CRC's proliferation and metastasis [42, 43]. However, some studies also demonstrated that NPP4B promotes survival and proliferation of tumor cells, including triple-negative breast cancer and leukemia [44–46].

Ribosomal protein S6 kinase A1 (RPS6KA1) is a member of the RSK (ribosomal S6 kinase) family of serine/threonine kinases and can phosphorylate members of the mitogen-activated kinase (MAPK) signaling pathway. RPS6KA1 participates in the carcinogenic process by regulating cell growth, insulin, and inflammation. Research also has shown that genetic variation of RPS6KA1 is significantly associated with the risk of developing colon cancer [47].

SRC proto-oncogene, nonreceptor tyrosine kinase (SRC), is the protein encoded by the proto-oncogene. Studies have shown that abnormal activation of intracellular tyrosine kinase SRC has been considered as a mechanism for acquired chemotherapy resistance in metastatic CRC [48]. SRC phosphorylation and activation can promote CRC invasion and metastasis [49, 50]. Stabilization of SRC promotes epithelial-mesenchymal transition in CRC [51]. And mutations of SRC can be involved in the malignant progression of colon cancer [52].

The solute carrier family 1 member 5 (SLC1A5) protein is a sodium-dependent neutral amino acid transporter, functioning as a receptor for RD114/type D retrovirus [53, 54]. Studies have found that SLC1A5 is an important transporter of glutamine, and upregulated SLC1A5 promotes the growth and survival of CRC cells [55]. And inhibition of SLC1A5 sensitizes CRC to cetuximab both in vitro and in vivo. The sensitization mechanism is that inhibition of SLC1A5
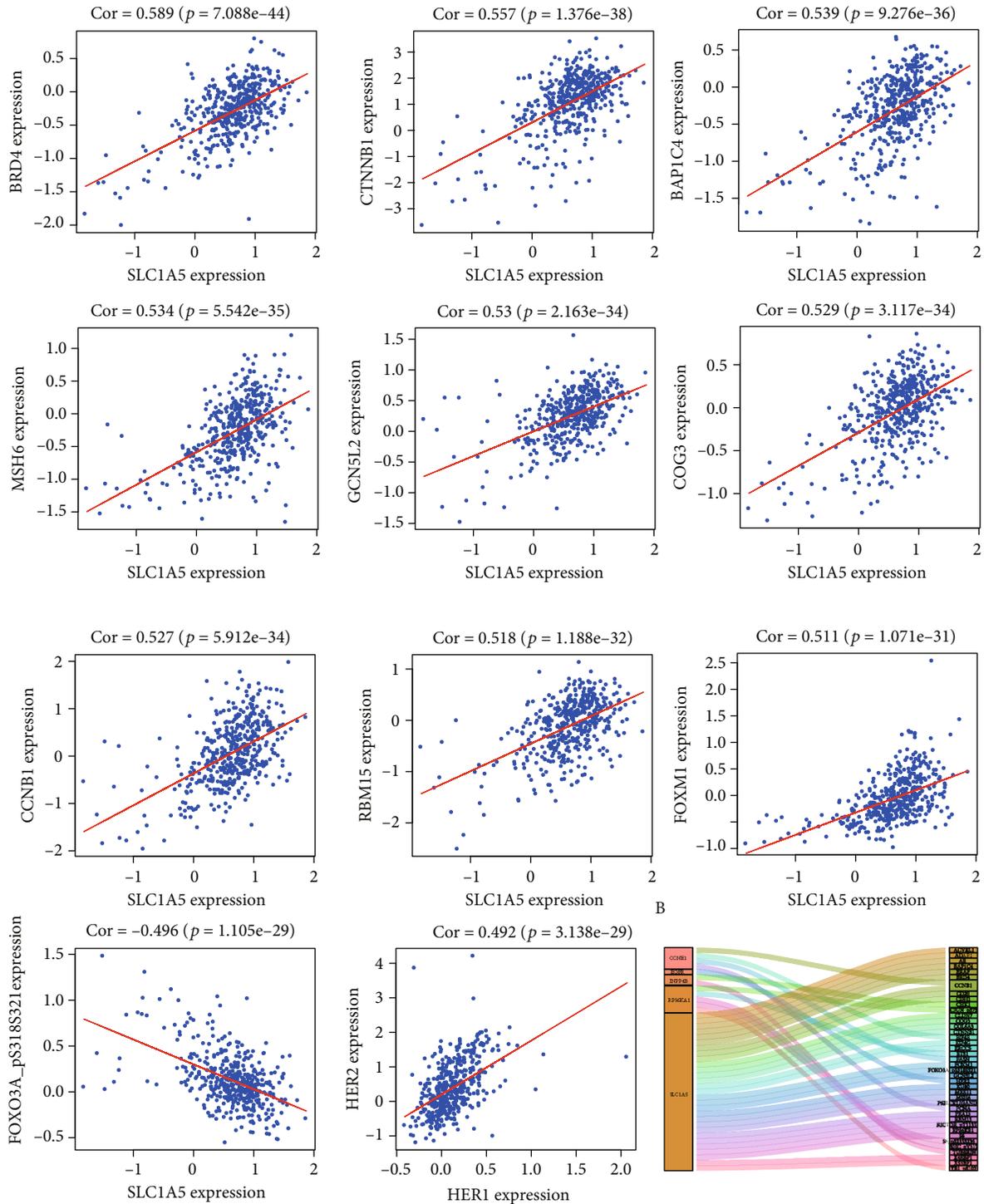
Figure 10: Coexpressed proteins of 6 hub proteins at the TCPA database. (a) 11 coexpressed proteins are moderately related to 6 hub proteins. Pearson correlation coefficient (PCC) was close to 0.5. (b) Sankey diagram of all proteins related to 6 hub proteins at the TCPA database (PCC > 0.4) ($p < 0.05$).
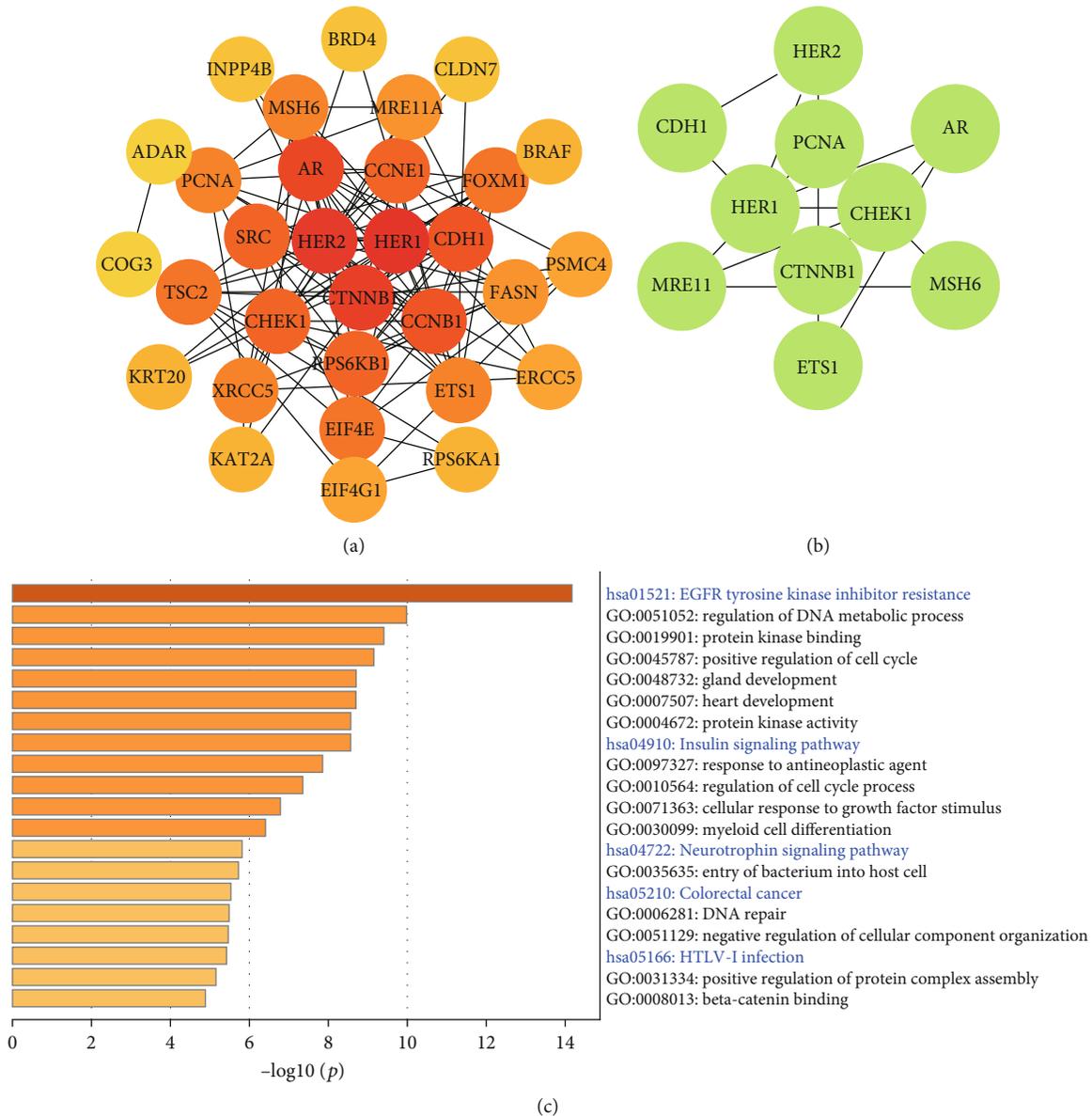
(a)

(b)

(c)

FIGURE 11: Protein-protein interaction (PPI) and enrichment analysis among coexpressed 46 proteins. (a) The PPI was ranked by degrees calculated by Cytoscape plugin cytoHubba. The darker the color, the higher the degree. (b) The significant module MOCDE_1 was built using the MCODE plugin at the Metascape. (c) The pathway and process enrichment analysis among 46 proteins. Blue words represented KEGG analysis, while black words represented GO analysis ($p < 0.05$).

TABLE 2: PPI enrichment analysis at the Metascape.

GO	Description	Log10 (p)
hsa01521	EGFR tyrosine kinase inhibitor resistance	-15.2
hsa04012	ErbB signaling pathway	-10.8
hsa01522	Endocrine resistance	-10.4
MCODE_1_GO	Description	Log10 (p)
hsa05200	Pathways in cancer	-10.5
hsa05213	Endometrial cancer	-8.5
GO : 0048732	Gland development	-8.1

promotes ubiquitin-proteasome degradation of HER1 and reduces the nuclear HER1 expression [56].

Based on previous studies, these 6 proteins are involved in the development and resistance of CRC. Our study further reveals the prognostic roles of these 6 proteins in CRC. More molecular research is urgently needed.

In this study, at the Metascape, enrichment analysis revealed that 6 hub proteins and their coexpressed proteins mainly enriched in EGFR (HER1) tyrosine kinase inhibitor resistance. HER1 serves as a stimulus for cancer growth, and some tyrosine kinase inhibitor (TKI) targeting HER1 has been currently administered. However, TKI resistance is common and leads to the recurrence of tumors. The efficacy of combined targeting 6 hub proteins in patients with CRC needs to be further studied.

In summary, our studies revealed that differentially expressed 6 hub proteins and the protein risk signature were significantly associated with the OS of CRC patients. Our risk signature has the potential for clinical application to predict the outcomes of CRC patients, and 6 hub proteins are expected to become novel therapeutic targets in the future.

Abbreviations

CRC:	Colorectal cancer
TCPA:	The Cancer Proteome Atlas
TCGA:	The Cancer Genome Atlas
PPI:	Protein-protein interaction
ROC curve:	Receiver operating characteristic curve
AUC:	Areas under the ROC curve
KEGG:	Kyoto Encyclopedia of Genes and Genomes
GO:	Gene ontology
OS:	Overall survival
CCNE1:	Cyclin E1
EGFR (HER1):	Epidermal growth factor receptor
INPP4B:	Inositol polyphosphate-4-phosphatase type IIB
RPS6KA1:	Ribosomal protein S6 kinase A1
SRC:	SRC proto-oncogene, nonreceptor tyrosine kinase
SLC1A5:	Solute carrier family 1 member 5.

Data Availability

The data used to support the findings of our study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

Shanwen Chen, Pengyuan Wang, and Yucun Liu designed this research. Taohua Yue and Cheng Liu collected and analyzed the data. Taohua Yue and Jing Zhu wrote this article. Zhihao Huang and Shihao Guo revised the manuscript. All authors have read and approved the final manuscript. Tao-hua Yue and Cheng Liu contributed equally to this work and should be considered as co-first authors.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[2] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA: a Cancer Journal for Clinicians*, vol. 70, no. 1, pp. 7–30, 2020.

[3] M. Zhang, K. Zhu, H. Pu et al., "An immune-related signature predicts survival in patients with lung adenocarcinoma," *Frontiers in Oncology*, vol. 9, p. 1314, 2019.

[4] H. Chen, X. Sun, W. Ge, Y. Qian, R. Bai, and S. Zheng, "A seven-gene signature predicts overall survival of patients with colorectal cancer," *Oncotarget*, vol. 8, no. 56, pp. 95054–95065, 2017.

[5] X. Wang, J. Zhou, M. Xu et al., "A 15-lncRNA signature predicts survival and functions as a ceRNA in patients with colorectal cancer," *Cancer Management and Research*, vol. Volume 10, pp. 5799–5806, 2018.

[6] Z. Huang, Q. Yang, and Z. Huang, "Identification of critical genes and five prognostic biomarkers associated with colorectal cancer," *Medical Science Monitor*, vol. 24, pp. 4625–4633, 2018.

[7] L. Chen, D. Lu, K. Sun et al., "Identification of biomarkers associated with diagnosis and prognosis of colorectal cancer patients based on integrated bioinformatics analysis," *Gene*, vol. 692, pp. 119–125, 2019.

[8] H. Zhu, Y. Ji, W. Li, and M. Wu, "Identification of key pathways and genes in colorectal cancer to predict the prognosis based on mRNA interaction network," *Oncology Letters*, vol. 18, no. 4, pp. 3778–3786, 2019.

[9] J. Bai, X. Zhang, Z. X. Xiang, P. Y. Zhong, and B. Xiong, "Identification of prognostic immune-related signature predicting the overall survival for colorectal cancer," *European Review for Medical and Pharmacological Sciences*, vol. 24, no. 3, pp. 1134–1141, 2020.

[10] P. Wu, Z. J. Heins, J. T. Muller et al., "Integration and analysis of CPTAC proteomics data in the context of cancer genomics in the cBioPortal," *Molecular & Cellular Proteomics*, vol. 18, no. 9, pp. 1893–1898, 2019.

[11] J. Li, R. Akbani, W. Zhao et al., "Explore, visualize, and analyze functional cancer proteomic data using the cancer proteome atlas," *Cancer Research*, vol. 77, no. 21, pp. e51–e54, 2017.

[12] The Cancer Genome Atlas Network, "Comprehensive molecular characterization of human colon and rectal cancer," *Nature*, vol. 487, no. 7407, pp. 330–337, 2012.

[13] K. Tomczak, P. Czerwinska, and M. Wiznerowicz, "Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge," *Contemporary Oncology*, vol. 1A, pp. 68–77, 2015.

[14] J. Li, Y. Lu, R. Akbani et al., "TCPA: a resource for cancer functional proteomics data," *Nature Methods*, vol. 10, no. 11, pp. 1046-1047, 2013.

[15] R. He and S. Zuo, "A robust 8-gene prognostic signature for early-stage non-small cell lung cancer," *Frontiers in Oncology*, vol. 9, p. 693, 2019.

[16] Z. Wang, Q. Song, Z. Yang, J. Chen, J. Shang, and W. Ju, "Construction of immune-related risk signature for renal papillary cell carcinoma," *Cancer Medicine*, vol. 8, no. 1, pp. 289–304, 2019.

[17] J.-A. W. Chapman, M. E. Trudeau, K. I. Pritchard et al., "A comparison of all-subset cox and accelerated failure time models with cox step-wise regression for node-positive breast cancer," *Breast Cancer Research and Treatment*, vol. 22, no. 3, pp. 263–272, 1992.

[18] P. Lin, Y. N. Guo, L. Shi et al., "Development of a prognostic index based on an immunogenomic landscape analysis of papillary thyroid cancer," *Aging*, vol. 11, no. 2, pp. 480–500, 2019.

[19] J. H. Liu, M. Feng, S. Y. Li et al., "Identification of molecular markers associated with the progression and prognosis of endometrial cancer: a bioinformatic study," *Cancer Cell International*, vol. 20, no. 1, p. 59, 2020.

[20] T. Fiolet, B. Srour, L. Sellem et al., "Consumption of ultra-processed foods and cancer risk: results from NutriNet-Santé prospective cohort," *BMJ*, vol. 360, 2018.

[21] K. Lin, J. Huang, H. Luo et al., "Development of a prognostic index and screening of potential biomarkers based on immunogenomic landscape analysis of colorectal cancer," *Aging*, vol. 12, no. 7, pp. 5832–5857, 2020.

[22] D. S. Chandrashekar, B. Bashel, S. A. H. Balasubramanya et al., "UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses," *Neoplasia*, vol. 19, no. 8, pp. 649–658, 2017.

[23] F. Chen, D. S. Chandrashekar, S. Varambally, and C. J. Creighton, "Pan-cancer molecular subtypes revealed by mass-spectrometry-based proteomic characterization of more than 500 human cancers," *Nature Communications*, vol. 10, no. 1, p. 5679, 2019.

[24] D. R. Rhodes, J. Yu, K. Shanker et al., "ONCOMINE: a cancer microarray database and integrated data-mining platform," *Neoplasia*, vol. 6, no. 1, pp. 1–6, 2004.

[25] R. Q. He, X. G. Zhou, Q. Y. Yi et al., "Prognostic signature of alternative splicing events in bladder urothelial carcinoma based on Spliceseq data from 317 cases," *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology.*, vol. 48, no. 3, pp. 1355–1368, 2018.

[26] G. D. Bader and C. W. V. Hogue, "An automated method for finding molecular complexes in large protein interaction networks," *BMC Bioinformatics*, vol. 4, no. 1, p. 2, 2003.

[27] Y. Zhou, B. Zhou, L. Pache et al., "Metascape provides a biologist-oriented resource for the analysis of systems-level datasets," *Nature Communications*, vol. 10, no. 1, p. 1523, 2019.

[28] P. Ahluwalia, A. K. Mondal, C. Bloomer et al., "Identification and clinical validation of a novel 4 gene-signature with prognostic utility in colorectal cancer," *International Journal of Molecular Sciences*, vol. 20, no. 15, p. 3818, 2019.

[29] Y. Yang, A. Qu, Q. Wu et al., "Prognostic value of a hypoxia-related microRNA signature in patients with colorectal cancer," *Aging*, vol. 12, no. 1, pp. 35–52, 2020.

[30] Z. Zhou, S. Mo, W. Dai et al., "Development and validation of an autophagy score signature for the prediction of postoperative survival in colorectal cancer," *Frontiers in Oncology*, vol. 9, p. 878, 2019.

[31] J. Wei, X. Ge, Y. Tang et al., "An autophagy-related long non-coding RNA signature contributes to poor prognosis in colorectal cancer," *Journal of Oncology*, vol. 2020, Article ID 4728947, 13 pages, 2020.

[32] M. Menor, Y. Zhu, Y. Wang, J. Zhang, B. Jiang, and Y. Deng, "Development of somatic mutation signatures for risk stratification and prognosis in lung and colorectal adenocarcinomas," *BMC Medical Genomics*, vol. 12, no. S1, p. 24, 2019.

[33] Y.-L. Sun, Y. Zhang, Y.-C. Guo, Z.-H. Yang, and Y.-C. Xu, "A prognostic model based on six metabolism-related genes in colorectal cancer," *BioMed Research International*, vol. 2020, Article ID 5974350, 16 pages, 2020.

[34] A. Mitchell, S. L. Hasanali, D. S. Morera et al., "A chemokine/-chemokine receptor signature potentially predicts clinical outcome in colorectal cancer patients," *Cancer Biomarkers*, vol. 26, no. 3, pp. 291–301, 2019.

[35] X. Hu, Y. Q. Li, X. J. Ma, L. Zhang, S. J. Cai, and J. J. Peng, "A risk signature with inflammatory and T immune cells infiltration in colorectal cancer predicting distant metastases and efficiency of chemotherapy," *Frontiers in Oncology*, vol. 9, p. 704, 2019.

[36] A. Sewing, V. Ronicke, C. Burger, M. Funk, and R. Muller, "Alternative splicing of human cyclin E," *Journal of Cell Science*, vol. 107, Part 2, pp. 581–588, 1994.

[37] W. Zhu, Y. Si, J. Xu et al., "Methyltransferase like 3 promotes colorectal cancer proliferation by stabilizing CCNE1 mRNA in an m6A-dependent manner," *Journal of Cellular and Molecular Medicine*, vol. 24, no. 6, pp. 3521–3533, 2020.

[38] Z. Wang, "ErbB Receptors and Cancer," *Methods in Molecular Biology*, vol. 1652, pp. 3–35, 2017.

[39] Y. F. Pei, J. Liu, J. Cheng, W. D. Wu, and X. Q. Liu, "Silencing of LAMC2 reverses epithelial-Mesenchymal transition and inhibits angiogenesis in cholangiocarcinoma via inactivation of the epidermal growth factor receptor Signaling pathway," *The American Journal of Pathology*, vol. 189, no. 8, pp. 1637–1653, 2019.

[40] X. Qu, L. Yang, Q. Shi, X. Wang, D. Wang, and G. Wu, "Lidocaine inhibits proliferation and induces apoptosis in colorectal cancer cells by upregulating mir-520a-3p and targeting EGFR," *Pathology, Research and Practice*, vol. 214, no. 12, pp. 1974–1979, 2018.

[41] D. L. H. Chan, E. Segelov, R. S. H. Wong et al., "Epidermal growth factor receptor (EGFR) inhibitors for metastatic colorectal cancer," *The Cochrane Database of Systematic Reviews*, vol. 6, article Cd007047, 2017.

[42] Y. Chen, Z. Sun, M. Qi et al., "INPP4B restrains cell proliferation and metastasis via regulation of the PI3K/AKT/SGK pathway," *Journal of Cellular and Molecular Medicine*, vol. 22, no. 5, pp. 2935–2943, 2018.

[43] E. J. Choi, M. S. Kim, N. J. Yoo, and S. H. Lee, "Inactivating frameshift mutation of INPP4B encoding a PI3K pathway

phosphatase in gastric and colorectal cancers," *Pathology Oncology Research*, vol. 22, no. 3, pp. 653-654, 2016.

[44] H. Jin, L. Yang, L. Wang et al., "INPP4B promotes cell survival via SGK3 activation in NPM1-mutated leukemia," *Journal of Experimental & Clinical Cancer Research: CR*, vol. 37, no. 1, p. 8, 2018.

[45] D. E. Reed and K. M. Shokat, "INPP4B and PTEN loss leads to PI-3,4-P2 accumulation and inhibition of PI3K in TNBC," *Molecular Cancer Research*, vol. 15, no. 6, pp. 765–775, 2017.

[46] L. Chen, Y. Cao, D. Rong, Y. Wang, and Y. Cao, "MicroRNA-605 functions as a tumor suppressor by targeting INPP4B in melanoma," *Oncology Reports*, vol. 38, no. 2, pp. 1276–1286, 2017.

[47] M. L. Slattery, A. Lundgreen, J. S. Herrick, and R. K. Wolff, "Genetic variation in RPS6KA1, RPS6KA2, RPS6KB1, RPS6KB2, and PDK1 and risk of colon or rectal cancer," *Mutation Research*, vol. 706, no. 1-2, pp. 13–20, 2011.

[48] C. M. Parseghian, N. U. Parikh, J. Y. Wu et al., "Dual inhibition of EGFR and c-Src by cetuximab and dasatinib combined with FOLFOX chemotherapy in patients with metastatic colorectal cancer," *Clinical Cancer Research*, vol. 23, no. 15, pp. 4146–4154, 2017.

[49] D. Tan, W. Zhang, Y. Tao, Y. Galiya, and M. Wang, "PZR promotes metastasis of colorectal cancer through increasing FAK and Src phosphorylation," *Acta Biochimica et Biophysica Sinica*, vol. 51, no. 4, pp. 356–364, 2019.

[50] Y. Wang, N. Wu, B. Pang et al., "TRIB1 promotes colorectal cancer cell migration and invasion through activation MMP-2 via FAK/Src and ERK pathways," *Oncotarget*, vol. 8, no. 29, pp. 47931–47942, 2017.

[51] A. Asiri, M. S. Toss, T. P. Raposo et al., "Cten promotes epithelial-mesenchymal transition (EMT) in colorectal cancer through stabilisation of Src," *Pathology International*, vol. 69, no. 7, pp. 381–391, 2019.

[52] Y. Shimada, Y. Muneoka, M. Nagahashi et al., "BRAF V600E and SRC mutations as molecular markers for predicting prognosis and conversion surgery in stage IV colorectal cancer," *Scientific Reports*, vol. 9, no. 1, p. 2466, 2019.

[53] S. Larriba, L. Sumoy, M. D. Ramos et al., "ATB(0)/SLC1A5 gene. Fine localisation and exclusion of association with the intestinal phenotype of cystic fibrosis," *European Journal of Human Genetics*, vol. 9, no. 11, pp. 860–866, 2001.

[54] M. Scalise, L. Pochini, L. Console, M. A. Losso, and C. Indiveri, "The human SLC1A5 (ASCT2) amino acid transporter: from function to structure and role in cell biology," *Frontiers in Cell and Developmental Biology*, vol. 6, p. 96, 2018.

[55] F. Huang, Y. Zhao, J. Zhao et al., "Upregulated SLC1A5 promotes cell growth and survival in colorectal cancer," *International Journal of Clinical and Experimental Pathology*, vol. 7, no. 9, pp. 6006–6014, 2014.

[56] H. Ma, Z. Wu, J. Peng et al., "Inhibition of SLC1A5 sensitizes colorectal cancer to cetuximab," *International Journal of Cancer*, vol. 142, no. 12, pp. 2578–2588, 2018.