

## Research Article

# Identification of Key Genes and Pathways Involved in Circulating Tumor Cells in Colorectal Cancer

Ruijun Pan,<sup>1,2</sup> Chaoran Yu,<sup>1</sup> Yanfei Shao ,<sup>1</sup> Hiju Hong,<sup>1</sup> Jing Sun ,<sup>1</sup> Zhou Zhang,<sup>3</sup> Peiyong Li ,<sup>2</sup> and Minhua Zheng <sup>1</sup>

<sup>1</sup>Department of General Surgery, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai Minimally Invasive Surgery Center, Shanghai 200025, China

<sup>2</sup>Department of Nuclear Medicine, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China

<sup>3</sup>Department of General Surgery, Shidong Hospital Yangpu District, Shanghai 200025, China

Correspondence should be addressed to Peiyong Li; peiyli@vip.sina.com and Minhua Zheng; zmhtiger@yeah.net

Received 17 March 2021; Revised 28 November 2021; Accepted 22 December 2021; Published 27 January 2022

Academic Editor: Elena Andreucci

Copyright © 2022 Ruijun Pan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Background.** Characterization of the features associated with circulating tumor cells (CTCs) is one of major interests for predicting clinical outcome of colorectal cancer (CRC) patients. However, the molecular features of CTCs remain largely unclear. **Methods.** For identification of key genes and pathways, GSE31023, contained CTCs from six metastatic CRC patients and three controls, was retrieved for differentially expressed gene (DEG) analysis. Protein-protein interaction networks of DEGs were constructed. Hub genes from the network were prognostic analyzed, as well as the association with tumor-infiltrating immune cells. **Results.** 1353 DEGs were identified between the CTC and control groups, with 403 genes upregulated and 950 downregulated. 32 pathways were significantly enriched in KEGG, with ribosome pathway as top. The top 10 hub genes were included, including eukaryotic translation elongation factor 2 (EEF2), ribosomal protein S2 (RPS2), ribosomal protein S5 (RPS5), ribosomal protein L3 (RPL3), ribosomal protein S3 (RPS3), ribosomal protein S14 (RPS14), ribosomal protein SA (RPSA), eukaryotic translation elongation factor 1 alpha 1 (EEF1A1), ribosomal protein S15a (RPS15A), and ribosomal protein L4 (RPL4). The correlation between CD4<sup>+</sup> T cells and RPS14 (correlation = -0.5) was the highest in colon cancer while CD8<sup>+</sup> T and RPS2 (correlation = -0.53) was the highest in rectal cancer. **Conclusion.** This study identified potential role of ribosome pathway in CTC, providing further insightful therapeutic targets and biomarkers for CRC.

## 1. Introduction

Colorectal cancer (CRC) is one of the major digestive malignancies in the world. During the tumor progression, hematogenous tumor cell disseminates and initiates the metastatic cascade of CRC. Circulating tumor cells (CTCs) exist in the peripheral blood of patients with various solid tumors including colorectal cancer and may lead to tumor metastasis [1]. With the development of liquid biopsy, CTCs have been proven to play an important role in detecting early development of metastasis and monitoring the curative effect of adjuvant therapy [2]. Therefore, molecular characterization of

CTCs has been one of major interests for predicting clinical outcome of patients [3].

Due to the low concentration of CTCs in blood, their detection needs highly sensitive and specific methods, including separation (concentration) and identification (detection). At present, CTCs and peripheral hematopoietic cells are generally distinguished according to their biological characteristics (expression and activity of cell surface proteins) and physical characteristics (size, density, charge, and deformability). Compared to the diameter of the blood cells (8  $\mu\text{m}$ ), tumor cells are larger and less likely to deform. Based on these characteristics, many membrane filtration devices appeared for CTC



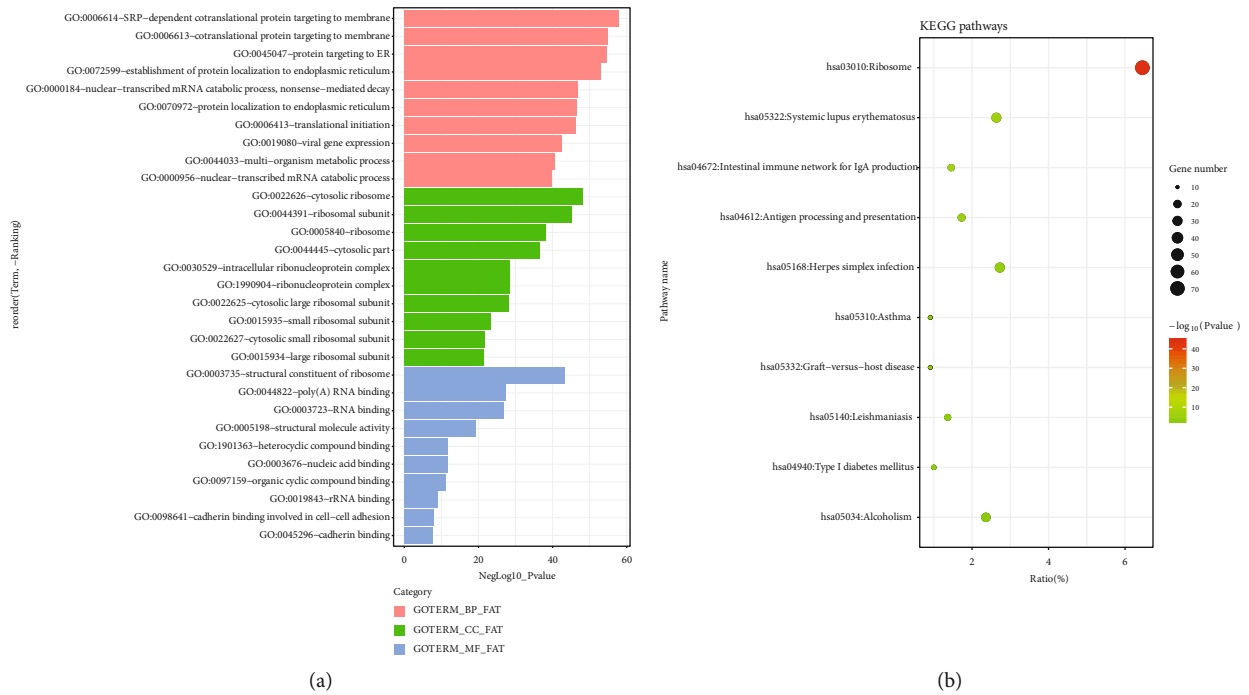


FIGURE 2: Enrichment analysis of DEGs. (a) Gene ontologies and DEGs; (b) Kyoto Encyclopedia of Genes and Genomes enrichment of DEGs.

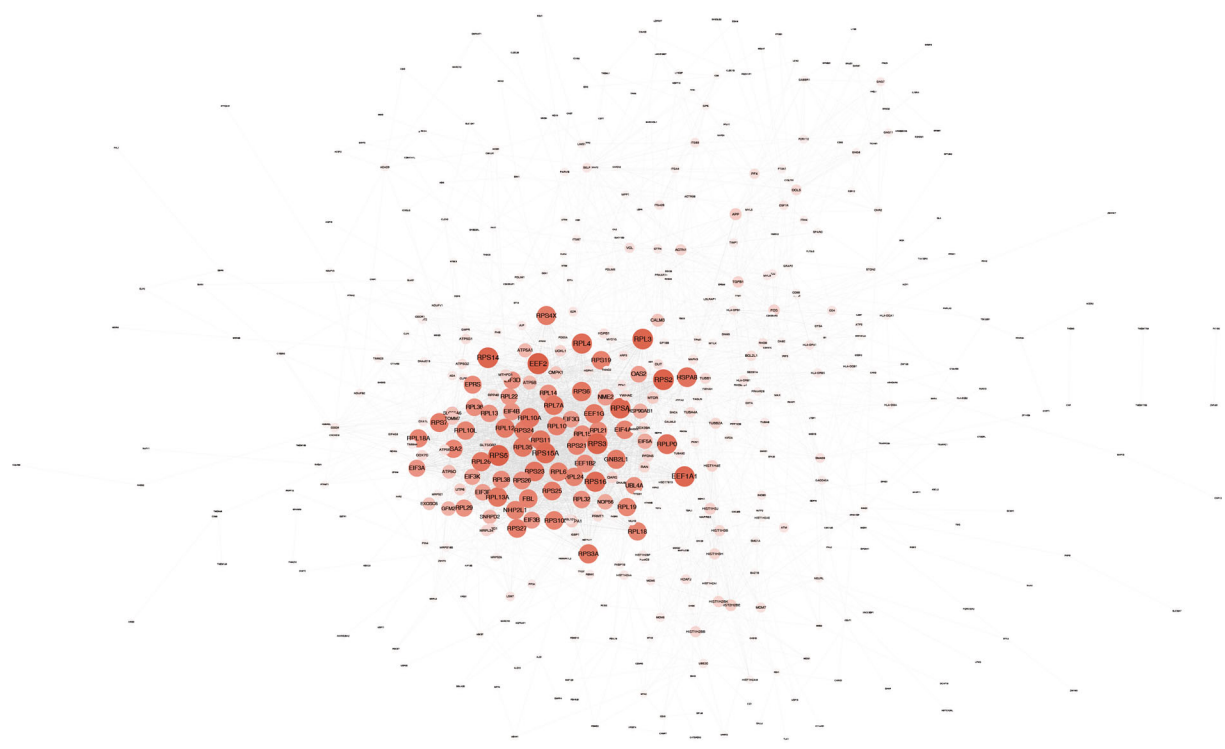


FIGURE 3: The protein-protein interaction (PPI) networks and the network analysis of enrichment results of DEGs. Nodes indicated each DEG. Lines represented in-between interactions. The size of each node and density of color were in proportion to the degrees.

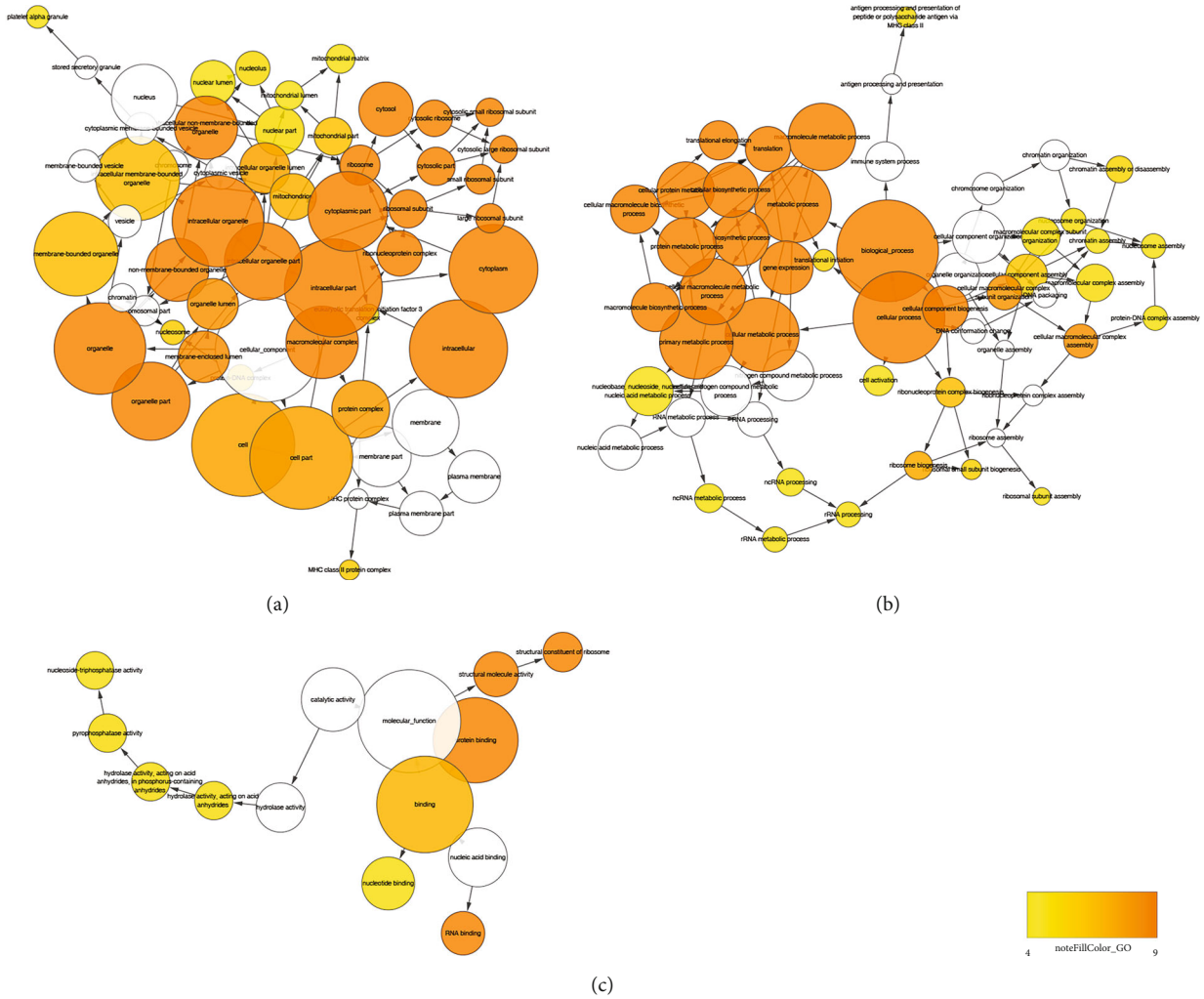


FIGURE 4: The network analysis of enrichment results of DEGs. (a) The enrichment of cellular components. (b) The enrichment of biological processes and the density of color were in proportion to the significance. (c) The network analysis of enrichment results of molecular functions by DEGs. The density of color was in proportion to the significance.

TABLE 1: Hub genes identified by the protein-protein interaction (PPI) networks.

Gene symbol	Log FC	Expression*	Gene name
EEF2	-1.86	Down	Eukaryotic translation elongation factor 2
RPS2	-1.76	Down	Ribosomal protein S2
RPS5	-2.00	Down	Ribosomal protein S5
RPL3	-1.92	Down	Ribosomal protein L3
RPS3	-1.58	Down	Ribosomal protein S3
RPS14	-1.53	Down	Ribosomal protein S14
RPSA	-2.35	Down	Ribosomal protein SA
EEF1A1	-1.88	Down	Eukaryotic translation elongation factor 1 alpha 1
RPS15A	-1.46	Down	Ribosomal protein S15a
RPL4	-1.55	Down	Ribosomal protein L4

\*Gene expression in CTC compared to normal control.

TABLE 2: Top scored three Molecular Complex Detection (MCODE) clusters and significantly enriched pathways. FDR: false discovery rate.

MCODE	Genes	Interactions	Gene symbols	Enriched pathways (FDR < 0.05)
1	61	1791	RPS27, RPS23, RPS14, RPL10L, RPL10, EIF3D, EIF3B, RPL24, RPS21, RPSA, NHP2L1, FBL, RPL7A, RPS16, RPS10, RPL19, RPL13, RPL4, RPS7, RPS3, RPL38, RPS26, RPL6, RPS4X, RPS3A, RPL18, RPL29, RPL14, RPS5, RPL35, RPS25, RPL3, RPL10A, RPS24, RPS15A, RPL12, RPL26, RPS19, RPL32, RPL18A, RPS11, RPL15, RPL36, RPL22, RPLP0, RPS2, RPL21, RPL13A, EIF4B, NSA2, EIF3K, EIF3F, EIF4A1, GNB2L1, EIF3A, EEF1B2, EIF3G, EEF1A1, EEF1G, EEF2, RPS6	hsa03010:ribosome
2	14	45	SPARC, GABBR1, TIMP1, CNR2, GNG7, CCL5, TGFB1, GNG11, GNG8, F13A1, P2RY12, CLU, ITH4, ACTN1	NA
3	28	86	HLA-DPA1, HLA-DPB1, LSM7, HIST1H2AM, DNM3, CTTN, HIST1H2AI, HIST1H2BF, HLA-DRB1, HLA-DRB5, CIITA, H2AFJ, HIST1H2BK, HIST1H2AA, HIST1H2BB, HIST1H2AE, HIST2H2BE, RAN, HLA-DMA, DAB2, LDLRAP1, HLA-DQA1, SNRPD1, REPS1, HLA-DQB1, STON2, PPIH, RBMX	hsa05322:systemic lupus erythematosus hsa05310:asthma hsa05332:graft-versus-host disease hsa05330:allograft rejection hsa04940:type I diabetes mellitus hsa04612:antigen processing and presentation hsa04672:intestinal immune network for IgA production hsa05320:autoimmune thyroid disease hsa05150:Staphylococcus aureus infection hsa05416:viral myocarditis hsa05321:inflammatory bowel disease (IBD) hsa05145:toxoplasmosis hsa05034:alcoholism hsa05140:leishmaniasis hsa05323:rheumatoid arthritis hsa05164:influenza A hsa05152:tuberculosis hsa04514:cell adhesion molecules (CAMs) hsa04145:phagosome hsa05166:HTLV-I infection hsa05168:herpes simplex infection

[13], containing six CTC samples from metastatic CRC patients with three normal controls, to identify potential key genes and pathways associated with CTCs of CRC.

## 2. Materials and Methods

### 2.1. Methods

**2.1.1. Gene Expression Profile GSE31023 for Analysis.** GSE31023 was the gene expression profiling by array, and all corresponding data was downloaded from the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>) [13]. This profile contained CTCs from six metastatic CRC patients and three healthy donors as control. And the related CTCs were isolated from 7.5 mL of peripheral blood by immunomagnetic separation using anti-EpCAM-coated mag-

netic beads (). Briefly, RNA in each sample was extracted and amplified using a whole transcriptome amplification system [13]. GPL13497 (Agilent-026652 Whole Human Genome Microarray 4 \* 44K v2) was the platform for GSE31023.

**2.1.2. Functional Annotation of Differentially Expressed Genes (DEGs).** The DEGs between the CTCs and normal cells were identified using the web tool, GEO2R, with predefined cutoff value  $p$  value < 0.05 and  $|\log \text{fold change} (\log \text{FC})| > 1$  [14]. The gene ontologies (GOs), as well as the Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, were employed for selected DEGs using the Database for Annotation, Visualization, and Integrated Discovery platform (DAVID, <http://david.abcc.ncifcrf.gov/>) [15–18]. Top 10 terms in each category, including biological process (BP), cellular component

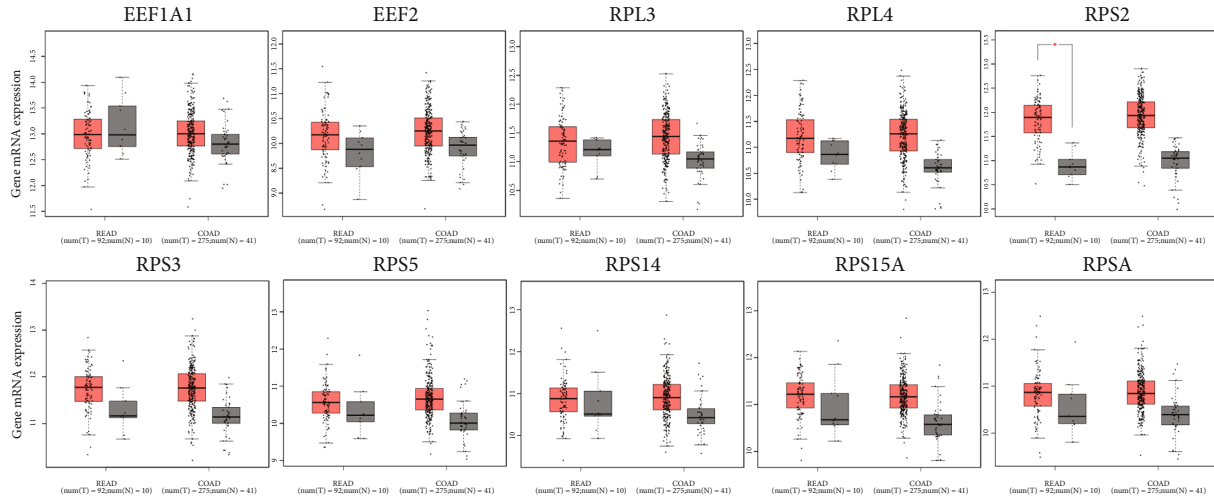


FIGURE 5: The mRNA expression of hub genes between tumor and normal using the GEPIA platform. Boxplot was used for mRNA expression between tumor (red) and normal (grey). READ: rectal adenocarcinoma; COAD: colon adenocarcinoma.

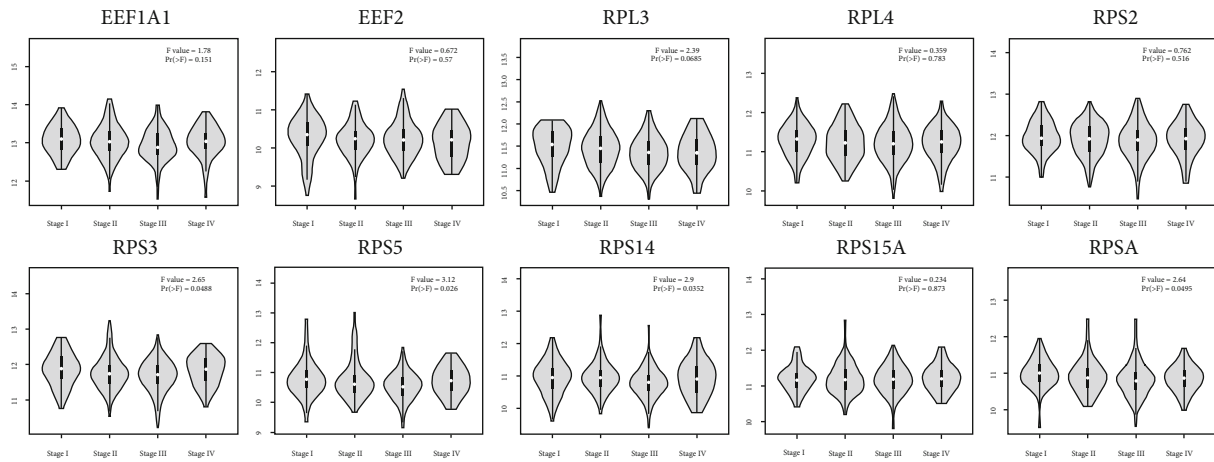


FIGURE 6: The stage-specific mRNA expression of hub genes by violin plot using the GEPIA platform. READ: rectal adenocarcinoma; COAD: colon adenocarcinoma; tumor: red; normal: grey.

(CC), and molecular function (MF), were displayed if more than 10 terms were defined as significant ( $p$  value  $< 0.05$ ).

**2.1.3. Construction of Protein-Protein Interaction (PPI) Networks.** PPI networks of DEGs were performed using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, <http://www.string-db.org/>) and visualized by the Cytoscape software (version 3.6.0) [19, 20]. Moreover, the Molecular Complex Detection (MCODE) program was used for subcluster identification of the PPI [21]. BINGO program was used for the GO presentation in the network analysis [22]. Hub genes were defined as the ten genes with highest degree determined by the PPI network.

**2.1.4. Expression of Hub Genes in The Cancer Genome Atlas (TCGA).** The mRNA expression boxplot of hub genes of TCGA (colon cancer, COAD and rectal cancer, READ) was retrieved from the gene expression profiling interactive analysis platform (GEPIA, <http://gepia.cancer-pku.cn>) [23].

**2.2. Correlation of Tumor-Infiltrating Immune Cells (TIICs) and Hub Genes.** Tumor Immune Estimation Response (TIMER, <https://cistrome.shinyapps.io/timer/>) is a novel platform for analyzing the expression abundance of the immune infiltration cells (CD8<sup>+</sup> T cells, CD4<sup>+</sup> T cells, dendritic cells, macrophages, neutrophils, and B cells) in malignant tumors, which was set up for online comparison based on references in TCGA [24]. Thus, the correlation of hub genes and all immune cells related in tumor was explored via TIMER. The correlation value was corrected by tumor purity [24].

**2.3. Prognostic Values of Hub Gene Signature Defined Risk Groups.** The prognostic values of hub gene signature defined risk groups in both overall survival (OS) and disease-free survival (DFS) were explored via the SurvExpress platform (<http://bioinformatica.mty.itesm.mx:8080/Biomatec/SurvivaX.jsp>) [25]. High- and low-risk groups were determined based on the risk score algorithm [25].

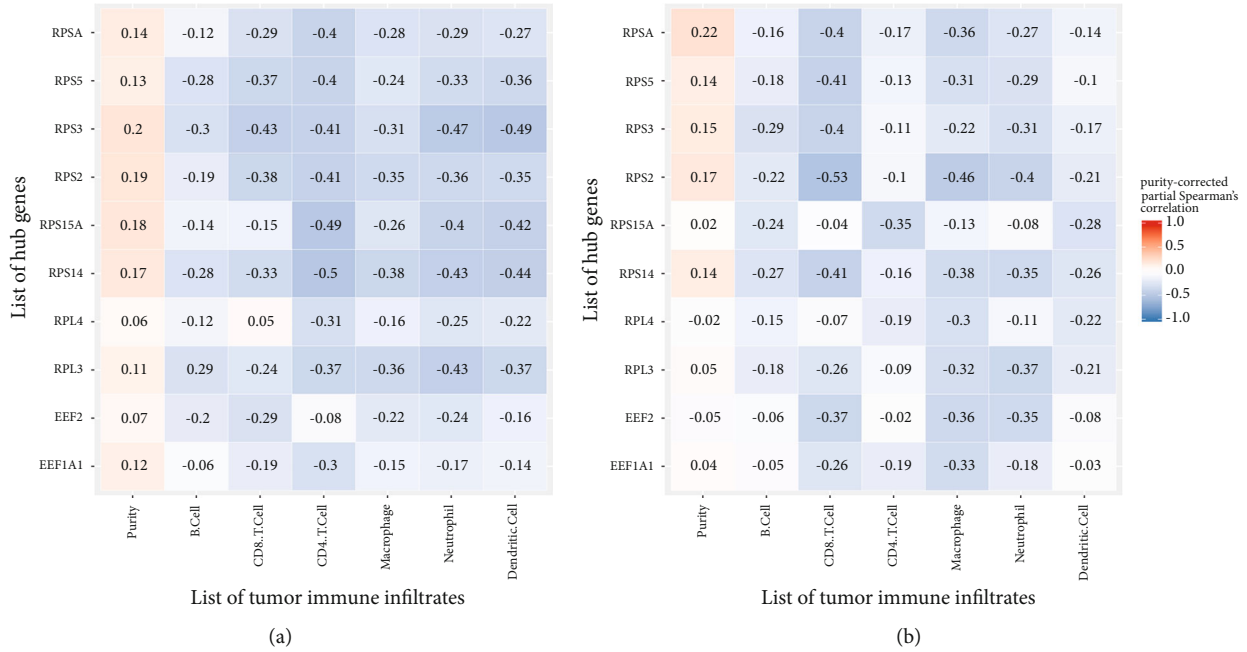


FIGURE 7: The correlation between hub genes and tumor-infiltrating immune cells (TIICs). (a) The correlation displayed in colon cancer of The Cancer Genome Atlas (TCGA); (b) the correlation displayed in rectal cancer of TCGA. Blue indicated negative purity-corrected partial Spearman’s correlation while red indicated positive correlation.

**3. Results**

**3.1. Identification and Functional Enrichment Analysis of DEGs.** A total of 1353 DEGs were identified between the CTCs and control groups, with 403 genes upregulated and 950 downregulated (Figures 1(a) and 1(b)). A total of 547 BP terms were significantly enriched. The most enriched three terms in BP were SRP-dependent cotranslational protein targeting to membrane, cotranslational protein targeting to membrane, and protein targeting to ER. A total of 142 terms were significantly enriched in CC. The most enriched three terms in CC were cytosolic ribosome, ribosomal subunit, and ribosome. A total of 100 terms were significantly enriched in MF. The most enriched three terms in MF were structural constituent of ribosome, poly (A) RNA binding, and RNA binding (Figure 2(a)). Noteworthy, a total of 32 pathways were significantly enriched in KEGG. The top three were ribosome (false discovery rate,  $FDR = 6.41E - 42$ ), systemic lupus erythematosus ( $FDR = 2.39E - 04$ ), and intestinal immune network for IgA production ( $FDR = 6.28E - 04$ ) (Figure 2(b)).

**3.2. PPI Network Establishment of DEGs.** Next, we explored the PPI network of all DEGs. In fact, a total of 496 nodes and 4283 edges were identified within the PPI network (Figure 3). Meanwhile, the functional enrichment network was also displayed (Figures 4(a)–4(c)). The top 10 hub genes include deukaryotic translation elongation factor 2 (EEF2), ribosomal protein S2 (RPS2), ribosomal protein S5 (RPS5), ribosomal protein L3 (RPL3), ribosomal protein S3 (RPS3), ribosomal protein S14 (RPS14), ribosomal protein SA (RPSA), eukaryotic translation elongation factor 1 alpha 1 (EEF1A1), ribosomal protein S15a (RPS15A), and ribosomal protein L4 (RPL4) (Table 1). The top three scored modules

were determined by MCODE and further functionally enriched, which also highlighted the role of ribosome (Table 2). Noteworthy, all the hub genes were found down-regulated in CTCs.

**3.3. Expression of Hub Genes.** Of all the expression comparison between tumor and normal, only RPS2 was upregulated in tumor compared to normal in READ (Figure 5). RPS3, RPS5, RPS14, and RPSA were found significantly stage-specific expressed (Figure 6).

**3.4. The Correlation between Hub Genes and TIICs.** Furthermore, the correlation between hub genes and TIICs was analyzed via the TIMER platform. In colon cancer, the highest correlation was found between  $CD4^+$  T cells and RPS14 (correlation =  $-0.5$ ) and  $CD4^+$  T and RPS15A (correlation =  $-0.49$ ), as well as dendritic cells and RPS3 (correlation =  $-0.49$ ). In rectal cancer, the highest correlation was found between  $CD8^+$  T and RPS2 (correlation =  $-0.53$ ) and macrophage and RPS2 (correlation =  $-0.46$ ) (Figure 7).

**3.5. Prognostic Values of Hub Gene Signature.** Given increasing focus has been found in the prognostic roles of gene signature, this study further explored the prognostic values of hub gene signature via the SurvExpress platform. In OS analysis, significant prognostic roles were found between high-risk and low-risk groups (hazard ratio = 1.99, 95% confidence interval: 1.38-2.87, and  $p = 0.0002$ ) (Figure 8(a)). Meanwhile, the expression comparison was also illustrated between two groups (Figures 8(b) and 8(c)). In DFS analysis, significant prognostic roles were also found between high-risk and low-risk groups (hazard ratio = 1.71, 95% confidence interval: 1.2-2.46, and  $p = 0.003$ ) with expression comparison (Figure 9).

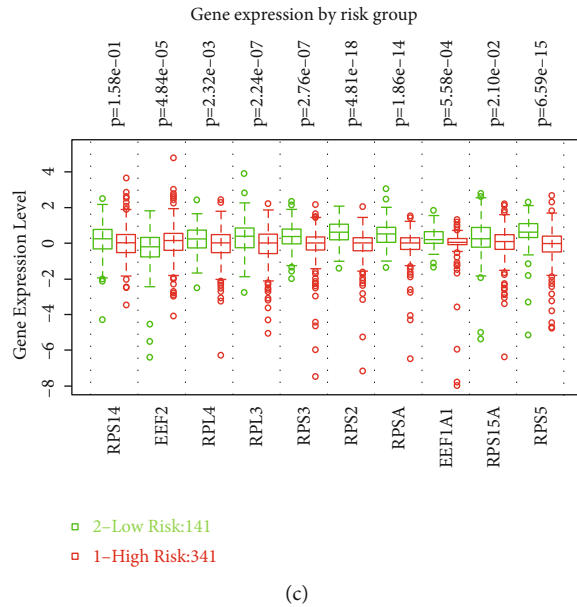
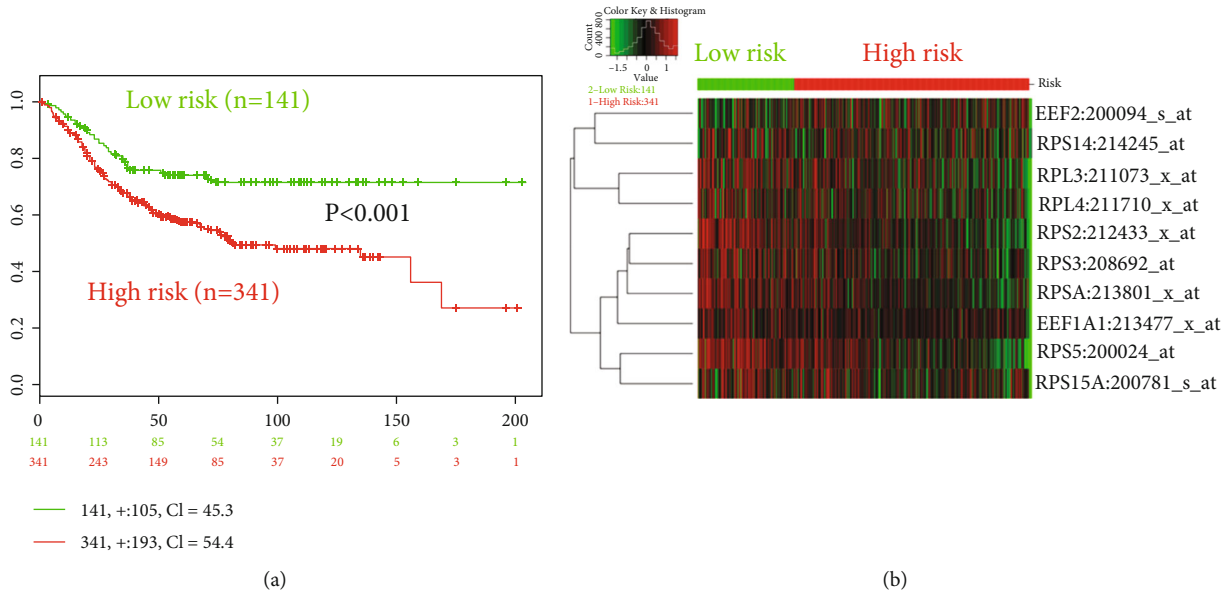


FIGURE 8: The prognostic values (overall survival, OS) of hub gene signature in colorectal cancer via the SurvExpress platform. (a) The KM plot of high-/low-risk groups divided by prognostic index; (b) the mRNA expression of high-/low-risk OS groups in heat map; (c) the mRNA expression comparison between high- and low-risk OS groups. High risk: red; low risk: green.

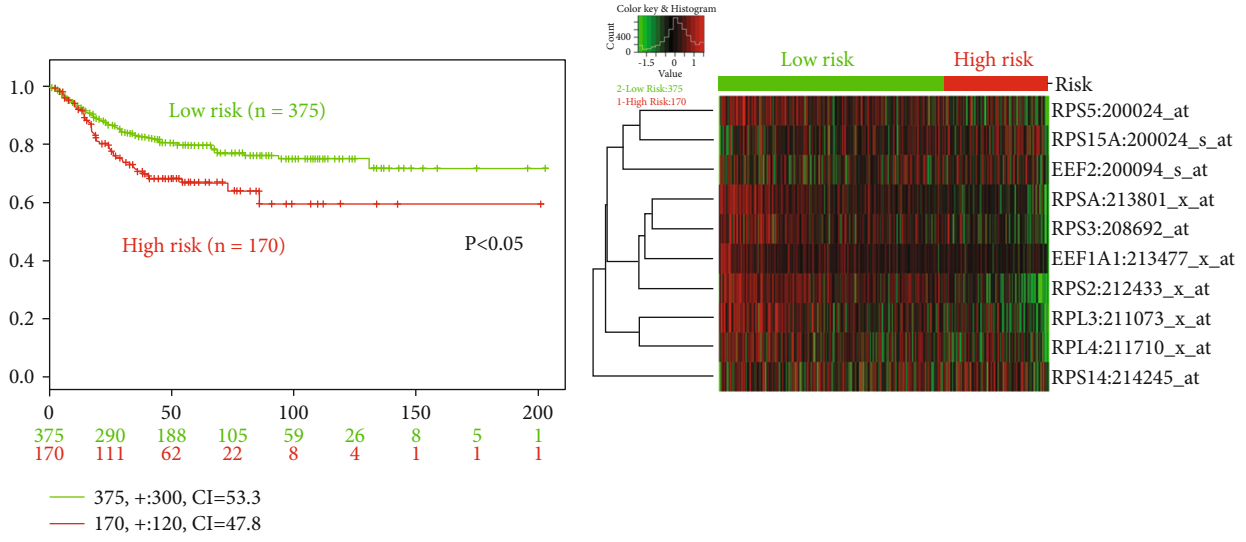
#### 4. Discussion

Commonly, standard patterns for the detection of CTCs in CRC are closely associated with genomic features. In fact, the intrinsic genomic features of metastatic lesions may not be identical to those of primary lesions [26]. During the metastatic progression, tumor cells show reduced adhesion markers and gradually detach from the primary lesion and flow into the circulation system. However, not all of the CTCs could be successfully habited at distant organs. Only a small proportion of tumor cells survives the intrinsic immunological eradication and undergoes profile-change at the secondary lesion. Meanwhile, normal epithelial cells also could join the circulated traveling, guided by inflammation-triggered cytokines [27]. Thus, molecular characterization of CTCs is

needed. However, the reculture of isolated CTCs remains technically difficult. Zhang et al. reported that a population of CTCs from 3 patients with breast cancer could be successfully used to form adherent cell line, with limited survival period and proliferation status [28]. Guan et al. have analyzed 7 GEO datasets (GSE99394, GSE31023, GSE82198, GSE65505, GSE67982, GSE76250, and GSE50746) and found that CTCs mainly change epithelial-mesenchymal transition (EMT), cell adhesion, and apoptosis [29]. Based on the study, we further indicated the key genes and pathways mainly involved in CTCs in CRC and revealed more promising biomarkers in CRC prognosis and immunotherapy.

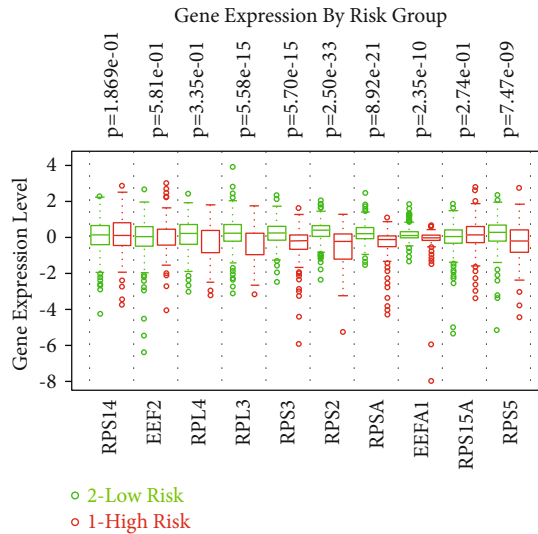
Noteworthy, ribosome pathway was highlighted in this study given the enrichment analysis of DEGs between CTCs and control. Interestingly, most of the hub genes were





(a)

(b)



(c)

FIGURE 9: The prognostic values (disease-free survival, DFS) of hub gene signature in colorectal cancer via the SurvExpress platform. (a) The KM plot of high-/low-risk groups divided by prognostic index; (b) the mRNA expression of high-/low-risk DFS groups in heat map; (c) the mRNA expression comparison between high- and low-risk DFS groups. High risk: red; low risk: green.

closely associated with ribosome pathway and all downregulated in CTCs compared to control. Consistently, expression profiling of breast cancer also highlighted the ribosome-related pathways and terms in genes downregulated in CTCs compared to control [30]. In fact, reduced levels of immune signals and apoptotic pathways were also enriched in CTCs of breast cancer [30]. Moreover, mammalian target of rapamycin pathway, constitutively activated by upstream AKT and PI3K pathways, was one of the key targets for persistent/recurrent epithelial ovarian cancer and closely associated with ribosome protein and eukaryotic translation initiation factor [31]. This study highlighted potential role of ribosome in CTCs of CRC, and the analysis of hub genes has opened up a new question as the therapeutic value of ribosome in CTCs.

For 10 hub genes, remarkable correlations with TIICs and prognostic values had been recognized in this study. However, solid validation remained in another independent CTC cohort, instead of conventional tissue-based genome results. Furthermore, only RPS2 was upregulated in tumor compared to normal in rectal cancer of TCGA, which may due to the different molecular expression characteristics between CTCs and solid tumor cells. Therefore, it is reasonable to further validate the results in an independent CTC cohort study.

Our study had the following strengths. First, we further identified the differentially expressed genes and pathways involved in CTCs in CRC. Second, several external datasets were used to verify that these hub genes can be related to the prognosis and immunotherapy of CRC patients. Besides, the study also has some limitations. First, the databases retrieving

data from studies were conducted in different ways. Second, the direct relationship between these hub genes in CTCs and clinical characteristics has not been further verified.

## 5. Conclusion

This study identified potential role of ribosome pathway in CTC, providing further insightful therapeutic targets for CRC. Moreover, the association between hub genes and CTCs may provide new perspectives for the exploit of new markers.

## Abbreviations

CRC:	Colorectal cancer
CTCs:	Circulating tumor cells
EMT:	Epithelial-to-mesenchymal transition process
MET:	Mesenchymal-to-epithelial transition
ISET:	Isolation by size of epithelial tumor cells
EpCAM:	Epithelial cell adhesion molecule
EGFR:	Epidermal growth factor receptor
GEO:	Gene Expression Omnibus
GOs:	Gene ontologies
KEGG:	Kyoto encyclopedia of genes and genomes
DAVID:	Database for Annotation, Visualization, and Integrated Discovery platform
BP:	Biological process
CC:	Cellular component
MF:	Molecular function
PPI:	Protein-protein interaction
STRING:	Search Tool for the Retrieval of Interacting Genes/Proteins
MCODE:	Molecular Complex Detection
TCGA:	The Cancer Genome Atlas
COAD:	Colon cancer
READ:	Rectal cancer
GEPIA:	Gene expression profiling interactive analysis platform
TIICs:	Tumor-infiltrating immune cells
TIMER:	Tumor immune estimation response
DFS:	Disease-free survival
OS:	Overall survival.

## Data Availability

The datasets supporting the conclusion of this article were included within the article.

## Consent

No consent was necessary.

## Conflicts of Interest

All authors declare no conflict of interest in this study.

## Authors' Contributions

Ruijun Pan, Chaoran Yu, and Yanfei Shao contributed equally to this work.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Nos. 81371598, 81572973, 81402423, 81572818, and 81871984) and Shanghai Municipal Commission of Health and Family Planning (2017YQ062), as well as Shanghai Science and Technology Committee (18695841400).

## References

- [1] L. Cabel, C. Proudhon, H. Gortais et al., "Circulating tumor cells: clinical validity and utility," *International Journal of Clinical Oncology*, vol. 22, pp. 421–430, 2017.
- [2] M. Y. Huang, H. L. Tsai, J. J. Huang, and J. Y. Wang, "Clinical implications and future perspectives of circulating tumor cells and biomarkers in clinical outcomes of colorectal cancer," *Translational Oncology*, vol. 9, pp. 340–347, 2016.
- [3] D. Lin, L. Shen, M. Luo et al., "Circulating tumor cells: biology and clinical significance," *Signal Transduction and Targeted Therapy*, vol. 6, no. 1, p. 404, 2021.
- [4] A. Toss, Z. Mu, S. Fernandez, and M. Cristofanilli, "CTC enumeration and characterization: moving toward personalized medicine," *Annals of translational medicine*, vol. 2, no. 11, p. 108, 2014.
- [5] A. Seeber, G. Untergasser, G. Spizzo et al., "Predominant expression of truncated EpCAM is associated with a more aggressive phenotype and predicts poor overall survival in colorectal cancer," *International Journal of Cancer*, vol. 139, pp. 657–663, 2016.
- [6] Y. Meng, B. Q. Xu, Z. G. Fu et al., "Cytoplasmic EpCAM overexpression is associated with favorable clinical outcomes in pancreatic cancer patients with hepatitis B virus negative infection," *International Journal of Clinical and Experimental Medicine*, vol. 8, pp. 22204–22216, 2015.
- [7] D. Fong, P. Moser, A. Kasal et al., "Loss of membranous expression of the intracellular domain of EpCAM is a frequent event and predicts poor survival in patients with pancreatic cancer," *Histopathology*, vol. 64, pp. 683–692, 2014.
- [8] M. Bulfoni, M. Turetta, F. Del Ben, C. Di Loreto, A. P. Beltrami, and D. Cesselli, "Dissecting the heterogeneity of circulating tumor cells in metastatic breast cancer: going far beyond the needle in the haystack," *International Journal of Molecular Sciences*, vol. 17, p. E1775, 2016.
- [9] S. H. Lim, T. M. Becker, W. Chua, W. L. Ng, P. de Souza, and K. J. Spring, "Circulating tumour cells and the epithelial mesenchymal transition in colorectal cancer," *Journal of Clinical Pathology*, vol. 67, pp. 848–853, 2014.
- [10] M. T. Gabriel, L. R. Calleja, A. Chalopin, B. Ory, and D. Heymann, "Circulating tumor cells: a review of non-EpCAM-based approaches for cell enrichment and isolation," *Clinical Chemistry*, vol. 62, pp. 571–581, 2016.
- [11] P. K. Grover, A. G. Cummins, T. J. Price, I. C. Roberts-Thomson, and J. E. Hardingham, "Circulating tumour cells: the evolving concept and the inadequacy of their enrichment by EpCAM-based methodology for basic and clinical cancer research," *Annals of Oncology*, vol. 25, pp. 1506–1516, 2014.
- [12] Q. Zhang, J. Nong, J. Wang et al., "Isolation of circulating tumor cells and detection of EGFR mutations in patients with non-small-cell lung cancer," *Oncology Letters*, vol. 17, no. 4, pp. 3799–3807, 2019.

- [13] A. Toss, Z. Mu, S. Fernandez, and M. Cristofanilli, "Molecular characterization of circulating tumor cells in human metastatic colorectal cancer," *PLoS one*, vol. 7, no. 7, article e40476, 2012.
- [14] R. Edgar, M. Domrachev, and A. E. Lash, "Gene Expression Omnibus: NCBI gene expression and hybridization array data repository," *Nucleic Acids Research*, vol. 30, no. 1, pp. 207–210, 2002.
- [15] S. Davis and P. S. Meltzer, "GEOquery: a bridge between the Gene Expression Omnibus (GEO) and Bio Conductor," *Bioinformatics*, vol. 23, no. 14, pp. 1846–1847, 2007.
- [16] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources," *Nature Protocols*, vol. 4, no. 1, pp. 44–57, 2009.
- [17] M. Ashburner, C. A. Ball, J. A. Blake et al., "Gene Ontology: tool for the unification of biology," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.
- [18] M. Kanehisa and S. Goto, "KEGG: Kyoto Encyclopedia of Genes and Genomes," *Nucleic Acids Research*, vol. 28, no. 1, pp. 27–30, 2000.
- [19] D. Szklarczyk, A. Franceschini, S. Wyder et al., "STRING v10: protein–protein interaction networks, integrated over the tree of life," *Nucleic Acids Research*, vol. 43, no. D1, pp. D447–D452, 2014.
- [20] P. Shannon, A. Markiel, O. Ozier et al., "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Research*, vol. 13, no. 11, pp. 2498–2504, 2003.
- [21] G. D. Bader and C. W. Hogue, "An automated method for finding molecular complexes in large protein interaction networks," *BMC Bioinformatics*, vol. 4, p. 2, 2003.
- [22] S. Maere, K. Heymans, and M. Kuiper, "BiNGO a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks," *Bioinformatics*, vol. 21, pp. 3448–3449, 2005.
- [23] Z. Tang, C. Li, B. Kang, G. Gao, C. Li, and Z. Zhang, "GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses," *Nucleic Acids Research*, vol. 45, no. W1, pp. W98–W102, 2017.
- [24] T. Li, J. Fan, B. Wang et al., "TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells," *Cancer Research*, vol. 77, no. 21, pp. e108–e110, 2017.
- [25] R. Aguirre-Gamboa, H. Gomez-Rueda, E. Martínez-Ledesma et al., "SurvExpress: an online biomarker validation tool and database for cancer gene expression data using survival analysis," *PLoS One*, vol. 8, no. 9, article e74250, 2013.
- [26] N. C. Bird, D. Mangnall, and A. W. Majeed, "Biology of colorectal liver metastases: a review," *Journal of Surgical Oncology*, vol. 94, pp. 68–80, 2006.
- [27] K. Pantel, E. Denève, D. Nocca et al., "Circulating epithelial cells in patients with benign colon diseases," *Clinical Chemistry*, vol. 58, pp. 936–940, 2012.
- [28] L. Zhang, L. D. Ridgway, M. D. Wetzel et al., "The identification and characterization of breast cancer CTCs competent for brain metastasis," *Science Translational Medicine*, vol. 5, p. 180, 2013.
- [29] Y. Guan, F. Xu, Y. Wang et al., "Identification of key genes and functions of circulating tumor cells in multiple cancers through bioinformatic analysis," *BMC Medical Genomics*, vol. 13, no. 1, p. 140, 2020.
- [30] J. E. Lang, J. H. Scott, D. M. Wolf et al., "Expression profiling of circulating tumor cells in metastatic breast cancer," *Breast Cancer Research and Treatment*, vol. 149, no. 1, pp. 121–131, 2015.
- [31] K. Behbakht, M. W. Sill, K. M. Darcy et al., "Phase II trial of the mTOR inhibitor, temsirolimus and evaluation of circulating tumor cells and tumor biomarkers in persistent and recurrent epithelial ovarian and primary peritoneal malignancies: a Gynecologic Oncology Group study," *Gynecologic Oncology*, vol. 123, no. 1, pp. 19–26, 2011.