



OPEN ACCESS

EDITED BY

Sujit Kumar Bhutia,
National Institute of Technology
Rourkela, India

REVIEWED BY

Jindong Xie,
Sun Yat-sen University Cancer Center
(SYSUCC), China
Fanen Yuan,
University of Pittsburgh, United States

*CORRESPONDENCE

Min Yang
✉ minyang@imm.ac.cn

RECEIVED 18 August 2023

ACCEPTED 27 September 2023

PUBLISHED 19 October 2023

CITATION

Zhang F, Zhang R, Zong J, Hou Y, Zhou M,
Yan Z, Li T, Gan W, Lv S, Yang L, Zeng Z,
Zhao W and Yang M (2023) Computational
identification and clinical validation of a
novel risk signature based on coagulation-
related lncRNAs for predicting prognosis,
immunotherapy response, and
chemosensitivity in colorectal
cancer patients.
Front. Immunol. 14:1279789.
doi: 10.3389/fimmu.2023.1279789

COPYRIGHT

© 2023 Zhang, Zhang, Zong, Hou, Zhou,
Yan, Li, Gan, Lv, Yang, Zeng, Zhao and Yang.
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.

Computational identification and clinical validation of a novel risk signature based on coagulation-related lncRNAs for predicting prognosis, immunotherapy response, and chemosensitivity in colorectal cancer patients

Fang Zhang¹, Rixin Zhang¹, Jinbao Zong^{2,3}, Yufang Hou¹,
Mingxuan Zhou¹, Zheng Yan¹, Tiegang Li¹, Wenqiang Gan¹,
Silin Lv¹, Liu Yang¹, Zifan Zeng¹, Wenyi Zhao¹ and Min Yang^{1*}

¹State Key Laboratory of Bioactive Substances and Function of Natural Medicine, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, ²Clinical Laboratory, The Affiliated Hospital of Qingdao University, Qingdao, China, ³Qingdao Hospital of Traditional Chinese Medicine, The Affiliated Qingdao Hiser Hospital of Qingdao University, Qingdao, China

Background: Coagulation is critically involved in the tumor microenvironment, cancer progression, and prognosis assessment. Nevertheless, the roles of coagulation-related long noncoding RNAs (CRLs) in colorectal cancer (CRC) remain unclear. In this study, an integrated computational framework was constructed to develop a novel coagulation-related lncRNA signature (CRLncSig) to stratify the prognosis of CRC patients, predict response to immunotherapy and chemotherapy in CRC, and explore the potential molecular mechanism.

Methods: CRC samples from The Cancer Genome Atlas (TCGA) were used as the training set, while the substantial bulk or single-cell RNA transcriptomics from Gene Expression Omnibus (GEO) datasets and real-time quantitative PCR (RT-qPCR) data from CRC cell lines and paired frozen tissues were used for validation. We performed unsupervised consensus clustering of CRLs to classify patients into distinct molecular subtypes. We then used stepwise regression to establish the CRLncSig risk model, which stratified patients into high- and low-risk groups. Subsequently, diversified bioinformatics algorithms were used to explore prognosis, biological pathway alteration, immune microenvironment, immunotherapy response, and drug sensitivity across patient subgroups. In addition, weighted gene coexpression network analysis was used to construct an lncRNA–miRNA–mRNA competitive endogenous network. Expression levels of CRLncSig, immune checkpoints, and immunosuppressors were determined using RT-qPCR.

Results: We identified two coagulation subclusters and constructed a risk score model using CRLncSig in CRC, where the patients in cluster 2 and the low-risk group had a better prognosis. The cluster and CRLncSig were confirmed as the independent risk factors, and a CRLncSig-based nomogram exhibited a robust prognostic performance. Notably, the cluster and CRLncSig were identified as the indicators of immune cell infiltration, immunoreactivity phenotype, and immunotherapy efficiency. In addition, we identified a new endogenous network of competing CRLs with microRNA/mRNA, which will provide a foundation for future mechanistic studies of CRLs in the malignant progression of CRC. Moreover, CRLncSig strongly correlated with drug susceptibility.

Conclusion: We developed a reliable CRLncSig to predict the prognosis, immune landscape, immunotherapy response, and drug sensitivity in patients with CRC, which might facilitate optimizing risk stratification, guiding the applications of immunotherapy, and individualized treatments for CRC.

KEYWORDS

colorectal cancer, coagulation, long noncoding RNA, prognostic signature, tumor microenvironment, immunotherapy, chemosensitivity

Introduction

According to the global cancer report released by the International Agency for Research on Cancer of the World Health Organization, colorectal cancer (CRC) has become the third most common and second most lethal cancer in the world, with more than 1.9 million new cases and approximately 900,000 deaths every year (1). With the development of gene profiles, targeted therapy, and immunotherapy, the treatment of advanced CRC has achieved great progress, but the 5-year survival rate of those patients is still less than 15%. Research of molecular mechanisms has confirmed that CRC is not a single type of disease but a group of diseases with highly heterogeneous and complex biological characteristics (2). Differentiating the different subtypes of CRC helps to understand the biological behavior of CRC at the molecular level and provides individualized treatment to patients. Currently, the common staging strategies for CRC are mainly the tumor size, node, and metastasis (TNM) staging system following the American Joint Committee on Cancer (3) standard and the consensus molecular subtype (CMS) system based on gene expression (4). However, there is still a wide variance in the clinical outcome of patients belonging to the same subtype. Immune-checkpoint therapy has revolutionized cancer therapy and become one of the most important anticancer immunotherapy methods. Immune-checkpoint inhibitors (ICIs) can reactivate disabled T cells and block the process of tumor immune escape to inhibit the malignant progression. However, the current benefit of ICIs therapy in CRC is limited to patients with high microsatellite instability, which prevents 85% of CRC patients from benefiting from immunotherapy (5–7). In recent years, with the deepening of the concept of precision medicine, the detection of related biomarkers is an effective supplement for CRC screening and plays an important

role in the judgment of individualized prognosis, precision medicine, and efficacy prediction.

There is a certain connection between the occurrence of tumors and the formation of thrombosis. The disorder of the coagulation system often leads to the occurrence of malignant tumors, so patients with malignant tumors are prone to venous thromboembolism (VTE) (8). However, not only can cancer cause VTE, but it may also lead to systemic coagulopathy, thereby leading to disseminated intravascular coagulation or thrombotic microangiopathy (9, 10). Specific expression promoted by tumor cell-associated clots is a unique feature of malignant tumors. The prognostic role of coagulation-related molecules in various cancers has been reported. The models constructed from coagulation-related molecules have good prognostic value in invasive ductal carcinoma and colon cancer (11, 12). A novel platelet-related gene signature has also been developed to predict the prognosis of patients with triple-negative breast cancer (13). In studies related to skin cutaneous melanoma, coagulation molecules could predict patients' prognosis and tumor microenvironment (TME) (14). It has been shown that the coagulation cascade plays an important role in the TME of hepatocellular carcinoma (HCC). There is a clear correlation between coagulation and TME in HCC, and the coagulation-related risk score can be used as a reliable prognostic biomarker to provide therapeutic benefits for chemotherapy and immunotherapy and help in clinical decision-making for HCC patients (15). In addition, in gastric cancer, the expression level of procoagulant genes is increased, resulting in increased angiogenesis, epithelial–mesenchymal transition (EMT), and TGF- β signaling, thus resulting in a poor prognosis (16). Collectively, the disorder of the coagulation system plays a very important role in the occurrence and development of tumors, and coagulation-related molecules are highly likely to become biomarkers for the prediction of prognosis and treatment in patients with CRC.

Long noncoding RNAs (lncRNAs) are noncoding products with a length of more than 200 nucleotides (17, 18). There is increasing evidence that lncRNAs play roles of oncogenes or antioncogenes in the regulation of tumorigenesis and development (19). However, there are few studies on coagulation-related lncRNAs (CRLs) in cancer. Coagulation Factor XI Antisense RNA 1 (*F11-AS1*) is an lncRNA that has been detected in ovarian cancer and pancreatic cancer (20, 21). It has been found that *F11-AS1* is downregulated in liver hepatocellular carcinoma (LIHC) tissues and cells and can inhibit the proliferation and migration of LIHC cells. *F11-AS1* regulates the expression of *PTEN* through competitive binding with *miR-3146*, thereby inhibiting the progression of LIHC (22). Another CRL, *ncRuPAR* (non-protein coding RNA, upstream of coagulation factor II thrombin receptor [F2R]/protease-activated receptor-1 [PAR-1]), can inhibit tumor cell proliferation and promote apoptosis of human gastric cancer cells by inhibiting PAR-1, PI3K/Akt signaling, and cyclin D1, and its downregulation can promote angiogenesis, invasion, metastasis, and progression (23, 24). However, it is not clear whether CRLs have a regulatory effect on the occurrence and development of CRC. Considering the important role of the coagulation system and lncRNAs in the development of cancer, we assume that CRLs may become a clinically valuable biomarker for CRC.

In this study, we screened the lncRNAs related to the coagulation pathway and successfully established the cluster of CRLs as well as the coagulation-related lncRNA signature (CRLncSig) in CRC. We comprehensively studied the prognostic value, potential mechanism, immune microenvironment, immunotherapy response rate, and drug sensitivity of the cluster and signature constructed using CRLs in CRC and constructed a competing endogenous RNA (ceRNA) network based on CRLs. Our findings suggest that CRLs can be used as novel biomarkers to predict the mortality risk and can elucidate signaling pathways and mechanisms involved in CRC progression. In addition, CRLs can predict the response to immunotherapy and drug sensitivity of CRC patients, which is conducive to the personalized treatment of CRC patients.

Materials and methods

Data sources for research

The gene expression profiles and clinical data of 619 CRC tissues and 50 adjacent normal tissues were downloaded from The Cancer Genome Atlas (TCGA, <https://www.cancer.gov/tcga>). CMSs for the patients in the cohort were downloaded from Colorectal Cancer Subtyping Consortium Synapse (4). We used the annotation file downloaded from GENCODE (<https://www.genencodegenes.org>) to acquire the gene list of lncRNA. Furthermore, we used the software R to merge the gene list of the lncRNA and the gene expression profiles to obtain the expression of lncRNAs in CRC patients. RNA-sequencing and clinical data for the external validation cohorts GSE147602 (25) and GSE198103 (26), an immunotherapy cohort (GSE181815) (27), and a single-cell RNA-sequencing (scRNA-seq) cohort (GSE136394) (28) were downloaded from the Gene Expression Omnibus (GEO) database

(<https://www.ncbi.nlm.nih.gov/geo/>). The mutation data for TCGA-CRC patients were acquired from TCGA database and analyzed using the package “maftools” (29).

Identification of CRLs

A total of 454 coagulation-related genes (CRGs) were extracted from the Molecular Signatures Database (MSigDB, <https://www.gsea-msigdb.org/gsea/msigdb>). We used the Spearman correlation analysis to screen the CRLs. With a threshold of $P\text{-value} < 0.05$ and an absolute value of correlation coefficient > 0.3 , we confirmed 2680 CRLs.

Consensus clustering based on CRLs

First, we performed univariate Cox regression analysis in TCGA cohort to obtain CRLs that were related to the prognosis of CRC patients ($P < 0.05$). Then consensus clustering analysis was performed to identify the distinct coagulation patterns based on the expression of 97 prognosis-related CRLs. The univariate Cox regression and consensus clustering analyses were performed using the R packages “survival” and “ConsensusClusterPlus,” respectively.

Development of a coagulation-related prognostic model

Among the prognosis-related CRLs ($P < 0.01$), a stepwise regression analysis was used for the development of a coagulation-related prognostic model. Afterward, we obtained 10 CRLs (dynein light chain roadblock-type 2 antisense RNA 1 (*DYNLRB2-AS1*), EF-hand calcium binding domain 13 divergent transcript (*EFCAB13-DT*), Ewing sarcoma-associated transcript 1 (*EWSAT1*), long intergenic non-protein coding RNA 645 (*LINC00645*), long intergenic non-protein coding RNA 901 (*LINC00901*), long intergenic non-protein coding RNA 1496 (*LINC01496*), long intergenic non-protein coding RNA 1738 (*LINC01738*), long intergenic non-protein coding RNA 2962 (*LINC02962*), LDL receptor-related protein 1 antisense RNA (*LRP1-AS*), and *PATJ* divergent transcript (*PATJ-DT*)) to construct a prognostic model. The risk score of each patient was computed through multivariate Cox regression analysis. Based on the median risk scores, we divided the CRC patients into two groups, namely, the group with a high-risk score and the group with a low-risk score. The stepwise and multivariate Cox regression analyses were performed using the R packages “MASS” and “survival,” respectively. The risk score for each patient was calculated using the following formula:

$$\text{Risk score} = \sum_{i=1}^{10} \text{Expression}(\text{lncRNA}_i) \times \text{Coefficient}(\text{lncRNA}_i)$$

Validation of the prognostic value and construction of a nomogram

To evaluate the prognostic value of the clusters and risk scores in TCGA cohort, the Kaplan-Meier (K-M) survival curves were

plotted to compare the overall survival (OS), progression-free survival (PFS), or disease-specific survival (DSS) between the clusters and risk groups using the R package “survminer.” To better demonstrate the reliability of our results, we randomly divided the samples in TCGA into training and validation sets at a ratio of 4:6. In addition, to determine the stability of the clusters and risk scores in predicting OS, we analyzed the correlation between the clusters and clinical characteristics and plotted K–M curves of risk scores for different clinical groups.

We constructed a nomogram model to predict the 1-, 3-, 5-, and 7-year OS of TCGA-CRC patients by selecting the risk scores and certain clinical features from the multivariate Cox regression analysis ($P < 0.05$) and tested the predictive ability of this model through calibration and receiver operating characteristic (ROC) curves. The nomogram was constructed using the R packages “survival” and “regplot,” and the calibration and ROC curves were drawn using the R packages “rms” and “timeROC,” respectively.

Functional enrichment analysis of differentially expressed genes

We used the R package “limma” to mine differentially expressed genes (DEGs) in different clusters and risk groups, with filtering criteria of an adjusted P -value < 0.05 and an absolute value of \log_2 (fold change (FC)) > 0.5 . We performed pathway enrichment analysis of DEGs using the R package “clusterProfiler” (30), including the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, Gene Ontology (GO) analysis, and Gene Set Variation Analysis (GSVA). GSVA was performed using the R package “GSVA” (31).

Immune profile analysis

First, we used the Estimation of STromal and Immune cells in Malignant Tumours using Expression data (ESTIMATE) algorithm to evaluate the immunity scores of different clusters and risk groups, including estimate, immune, and stromal scores (32). Subsequently, we downloaded the immune cell infiltration data of TCGA-CRC patients from the TIMER 2.0 database (Tumor Immune Estimation Resource, <https://cistrome.shinyapps.io/timer/>), including multiple deconvolution algorithms to assess the extent of immune cell infiltration of patients, namely, the Estimating the Proportion of Immune and Cancer cells (EPIC) (33), the Microenvironment Cell Populations-counter (MCP-counter) (34), the quantification of the Tumor Immune contexture from human RNA-seq data (quanTIseq) (35), TIMER (36), and Xcell (37) algorithms. We used the GSVA algorithm to evaluate the score of six immune-related gene sets [hematopoietic cell kinase (HCK), lymphocyte-specific protein tyrosine kinase, immunoglobulin G (IgG), major histocompatibility complex (MHC) I and II, and signal transducer and activator of transcription 1 (STAT1)] in different clusters (38) and compared the expression of different immunomodulatory genes among the clusters and risk groups. The gene list of immunomodulatory genes was

downloaded from the TISIDB database (<http://cis.hku.hk/TISIDB/download.php>) (39).

Analysis of immunotherapy efficacy

The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm can be used to predict a patient’s response to ICI therapy. The TIDE score is superior to the recognized immunotherapy biomarkers (tumor mutation burden (TMB), PDL1 level, and interferon γ) for measuring the effect of anti-PD1 and anti-CTLA4 treatment (40). The TIDE score, T-cell exclusion score, and the infiltration of cancer-associated fibroblast (CAF) were retrieved from the TIDE portal (<http://tide.dfci.harvard.edu>) by entering the normalized transcriptome data of TCGA-CRC patients. The Cytolytic activity (CYT) score is a novel immunotherapy biomarker that reflects the antitumor immune activity of CD8+ cytotoxic T cells and macrophages, which can be assessed by calculating the geometric mean of the expression of *GZMA* and *PRF1* (41). The T-cell-inflamed gene expression profile (GEP) score is a predictive score of the efficacy of ICIs widely used in preclinical studies. The GEP score was calculated using the 18 characteristic gene sets of IFN-gamma inflammation. Log2 log conversion was performed on the TPM values of each gene, and the average of the 18 genes was calculated (42).

Construction of a ceRNA network

To construct an lncRNA–miRNA–mRNA competitive endogenous network (ceRNA), we used weighted gene coexpression network analysis (WGCNA) to screen out the CRLs most correlated with the clusters and risk scores (hub-CRLs) and the immune-related CRGs (immune-CRGs) (43). Next, the miRcode database was used to obtain the miRNA family that bound to hub-CRLs (44). To improve the reliability of the results, mRNAs bound to miRNAs were obtained from the TargetScan (45), miRDB (46), and miRTarBase databases (47), and the results of the three databases were intersected with the immune-CRGs obtained previously. Finally, the ceRNA network was visualized using the software Cytoscape (version 3.9.1, U.S. National Institute of General Medical Sciences).

Drug-sensitivity prediction

To predict the drug sensitivity of TCGA-CRC patients according to the risk scores, the drug information from three databases, namely the Genomics of Drug Sensitivity in Cancer (GDSC) (48), the Cancer Therapeutics Response Portal (CTRP) (49–51), and the Profiling Relative Inhibition Simultaneously in Mixtures (PRISM), was included in our research (52, 53). We also acquired the transcription data of cell lines from DepMap Public 22Q2 (<https://depmap.org/portal/download/>), and the prediction was made using the R package “oncoPredict” (54).

scRNA-seq analysis

We used the scRNA-seq cohort GSE136394, which contained single-cell sequencing data of tumor-infiltrating lymphocytes from five CRC patients to perform scRNA-seq analysis. The R packages “Seurat” (55) and “harmony” (56) were used to read sample data and remove batch effects between the samples, and then we used the TSNE method for dimension reduction processing to obtain the clusters. Finally, we performed cell type annotation through Cellmarker 2.0 website (<http://bio-bigdata.hrbmu.edu.cn/CellMarker/index.html>).

Cell lines and cell culture

For the validation of our previous results, we used two human CRC cell lines (LOVO cells and HCT116 cells) and a normal human epithelial cell line (HcoEpic cells), which were purchased from BNBIO Company (Beijing, China). All cells were cultured in a constant temperature and humidity cell incubator at 37°C and 5% CO₂. The corresponding medium for the cells was as follows: LOVO cells [F-12K medium (Gibco, Invitrogen, Paisley, UK)], HCT116 cells [F12/Dulbecco’s modified Eagle medium (Gibco, Invitrogen, Paisley, UK)], and HcoEpic cells [F12/Dulbecco’s modified Eagle medium (Gibco, Invitrogen, Paisley, UK)]. The medium was supplemented with 10% fetal bovine serum (Corning, NY, USA) and 2% penicillin–streptomycin (10,000 units/mL penicillin, 10,000 µg/mL streptomycin; Gibco, Invitrogen, Paisley, UK).

RNA extraction and real-time quantitative PCR (RT-qPCR) validation

In addition to the three cell types mentioned above, we used samples from 51 patients with CRC. A total of 51 CRC and matched adjacent normal (distance to cancer greater than 5 cm) tissue samples used for RT-qPCR assay were obtained from patients who had been diagnosed with CRC by pathological examination of tissue biopsy and undergone operations at the Affiliated Hospital of Qingdao University. No radiotherapy or chemotherapy was applied before tissue collection. Informed consent was obtained from all of the participating patients. This work was approved by the Research Ethics Committee of The Affiliated Hospital of Qingdao University and was performed following the 1964 Helsinki Declaration and its later revisions. We used an RNeasy kit (Beyotime, Shanghai, China, R0027) following the manufacturer’s instructions to extract RNA from the cells and tissues. Then, we reverse-transcribed 1 µg of total RNA using SuperScript II reverse transcriptase (TaKaRa, Japan, RR047). Quantitative PCR analysis was performed using SYBR Green Master Mix (TaKaRa, Japan, RR820) with an ABI 7900 HT real-time PCR system. The primer sequences for RT-qPCR are listed in [Supplementary Table S1](#).

Statistical analysis

We used the software IBM SPSS Statistics 25, RStudio 4.2, and GraphPad Prism 8 for the analysis and visualization of the data. Differences were analyzed using the two-sided Wilcoxon rank-sum test for non-normally distributed continuous data and the chi-square test for categorical data. The correlation between non-normally distributed variables was analyzed using Spearman correlation analyses. The difference in survival between different groups was calculated using the log-rank test, and *P*-values were corrected using the Benjamini–Hochberg method in functional enrichment analysis. A *P*-value lower than 0.05 was considered statistically significant.

Results

Consensus clustering identifies two clusters based on CRLs

The flowchart of the entire research is shown in [Figure 1](#). First, we used Spearman correlation analysis of 454 CRGs and obtained a total of 2680 CRLs ($P < 0.05$, $|r| > 0.3$). Then, we used univariate Cox regression analysis and obtained 97 of the 2680 CRLs associated with prognosis ($P < 0.05$, [Figure 2A](#)). Subsequently, the 97 CRLs were used for unsupervised cluster analysis. According to the consensus CDF curve, the optimal cluster number was determined to be 2 ([Figure 2B](#)). After unsupervised clustering, we identified two clusters within TCGA cohort that showed different expression patterns of CRLs ([Figure 2C](#)). The expression levels of CRLs varied between the different clusters ([Supplementary Figure 1](#)). Specifically, CRLs with a positive correlation with CRGs were highly expressed in cluster 2, while CRLs with a negative correlation were poorly expressed in cluster 2.

Prognostic significance of the clusters based on CRLs

According to the results of the K–M survival analysis of OS, PFS, and DSS in TCGA cohort, the patients in cluster 2 had a higher survival probability than those in cluster 1 ([Figures 2D, E](#); [Supplementary Figure 2](#)). Then, we performed a correlation analysis between the clusters and clinical characteristics. As shown in [Figure 2F](#), the clusters significantly correlated with OS, MSI, venous invasion, clinical stage, TNM stage, and CMS classification. To explore the prognostic value of the clusters in TCGA patients, Cox regression analysis was performed. First, the clusters and clinical characteristics were incorporated into the univariate Cox regression model, and the *P*-values of the cluster for OS and PFS were lower than 0.05 ([Figures 2G–H](#)). Subsequently, the factors with $P < 0.05$ in the univariate Cox regression analysis were included in the multivariate Cox regression analysis. As the clinical stage, lymphatic metastasis, and vascular invasion are not

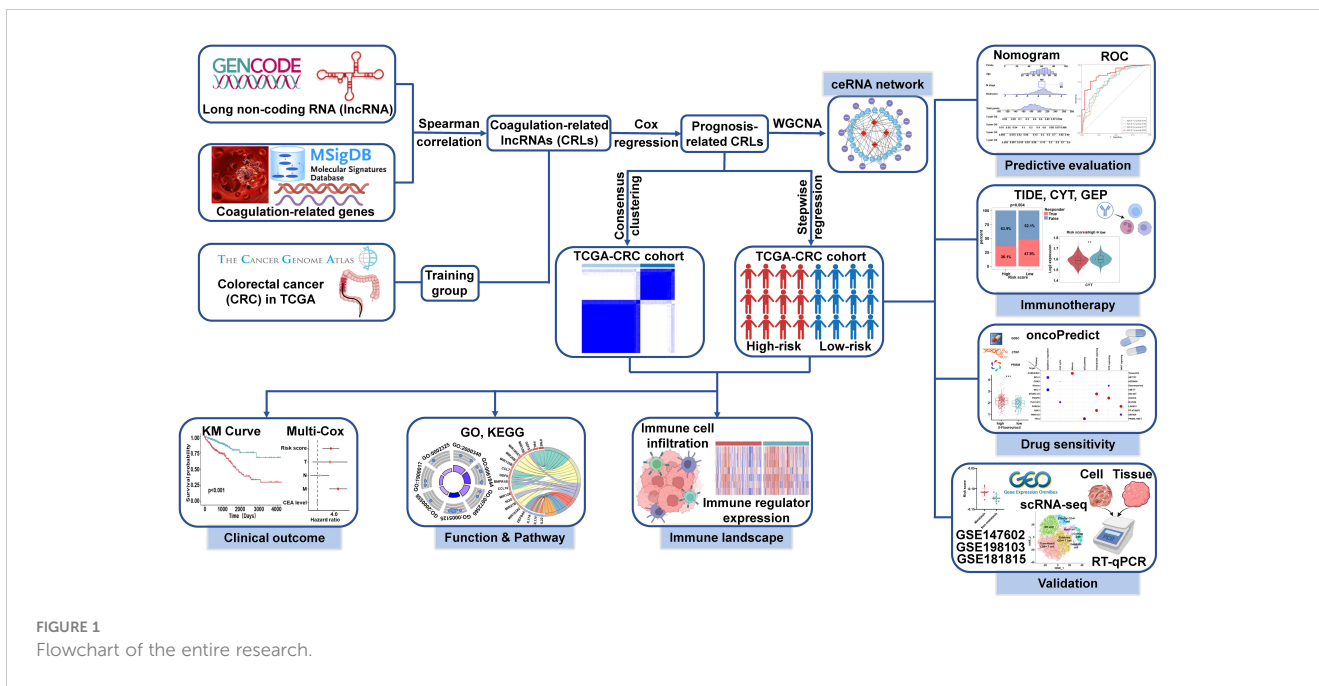


FIGURE 1
Flowchart of the entire research.

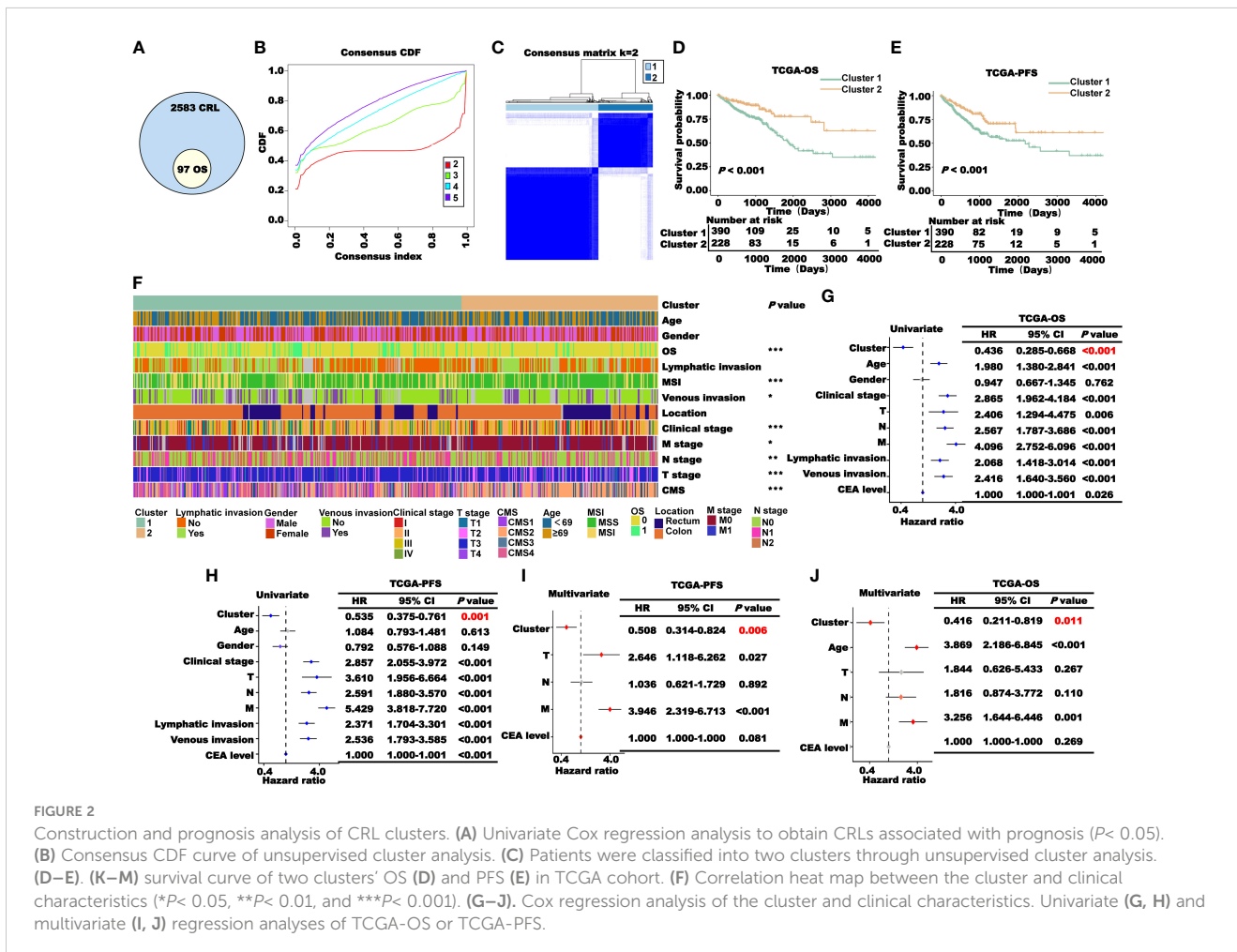


FIGURE 2
Construction and prognosis analysis of CRL clusters. (A) Univariate Cox regression analysis to obtain CRLs associated with prognosis ($P < 0.05$). (B) Consensus CDF curve of unsupervised cluster analysis. (C) Patients were classified into two clusters through unsupervised cluster analysis. (D–E). (K–M) survival curve of two clusters' OS (D) and PFS (E) in TCGA cohort. (F) Correlation heatmap between the cluster and clinical characteristics ($*P < 0.05$, $**P < 0.01$, and $***P < 0.001$). (G–J). Cox regression analysis of the cluster and clinical characteristics. Univariate (G, H) and multivariate (I, J) regression analyses of TCGA-OS or TCGA-PFS.

independent of the TNM stage, the TNM stage was included in this study. As shown in Figures 2I, J, the cluster was a significant independent prognostic factor for CRC.

Biological pathways and functional enrichment analysis of the clusters based on CRLs

To explore the potential mechanism of the coagulation-related cluster as an independent prognostic factor in CRC, we performed mutation analysis, GO, and KEGG pathway enrichment analysis. Because the occurrence of cancer is closely related to gene mutation, we first performed mutation analysis of clusters 1 and 2 and showed the top 20 mutant genes with a waterfall plot (Supplementary Figure 3A). Then, we compared the mutation percentages of 10 common genes closely related to the occurrence of cancer in the two clusters and found that there was no significant difference in the mutation percentages of those genes, except *BRAF* and *RNF43* (Supplementary Figure 3B).

To further explore the potential mechanism, we first used the R package “limma” to mine the DEGs in clusters 1 and 2, and a total of 1624 genes upregulated in cluster 1 and 377 genes upregulated in cluster 2 were obtained (Figure 3A). Subsequently, gene name and log₂FC of the DEGs in the clusters were used for enrichment analysis of GO and KEGG pathways. In the KEGG pathway analysis, the pathways mainly enriched in cluster 1 included cytokine–cytokine receptor interaction, microRNAs in cancer, and the TGF-beta signaling pathway. The pathways mainly enriched in cluster 2 included inflammatory bowel disease, IL-17 signaling pathway, and Th17 cell differentiation (Figure 3B).

According to the results of GO analysis, immune-related pathways, such as positive regulation of antimicrobial humoral response, response to interleukin-17, and positive regulation of humoral immune response, were mainly enriched in cluster 2 (Figure 3C). The above results suggest that the better prognosis of CRC patients in cluster 2 may be due to the anticancer effects of the immune-related pathways.

Immune profile analysis of the clusters based on CRLs

Previous pathway enrichment results suggest that the prognostic value of the clusters may be related to immunity. In addition, tumor immune microenvironments play a crucial role in the development and progression of tumors. Therefore, we performed the immune profile analysis of the clusters. First, we calculated the estimate, immune, and stromal scores for clusters 1 and 2 using the R package “Estimate.” As shown in Figure 4A, the scores of cluster 1 were all higher than those of cluster 2. Subsequently, we downloaded multiple deconvolution algorithms from the TIMER2.0 database to evaluate immune cell infiltration in patients with TCGA-CRC. In cluster 1, CAFs, M2 macrophages, myeloid dendritic cells, and other immunosuppressive cells were infiltrated to a high degree (Figure 4B). Furthermore, to investigate the immune system metagene, we examined crucial genes involved in inflammatory activities and immune molecules. We used the GSVA algorithm to calculate the scores of several key gene sets and then compared the scores between clusters 1 and 2. The radar plot presented in Figure 4C shows that the HCK score of cluster 1 was significantly higher than that of cluster 2, while the IgG, MHC-I,

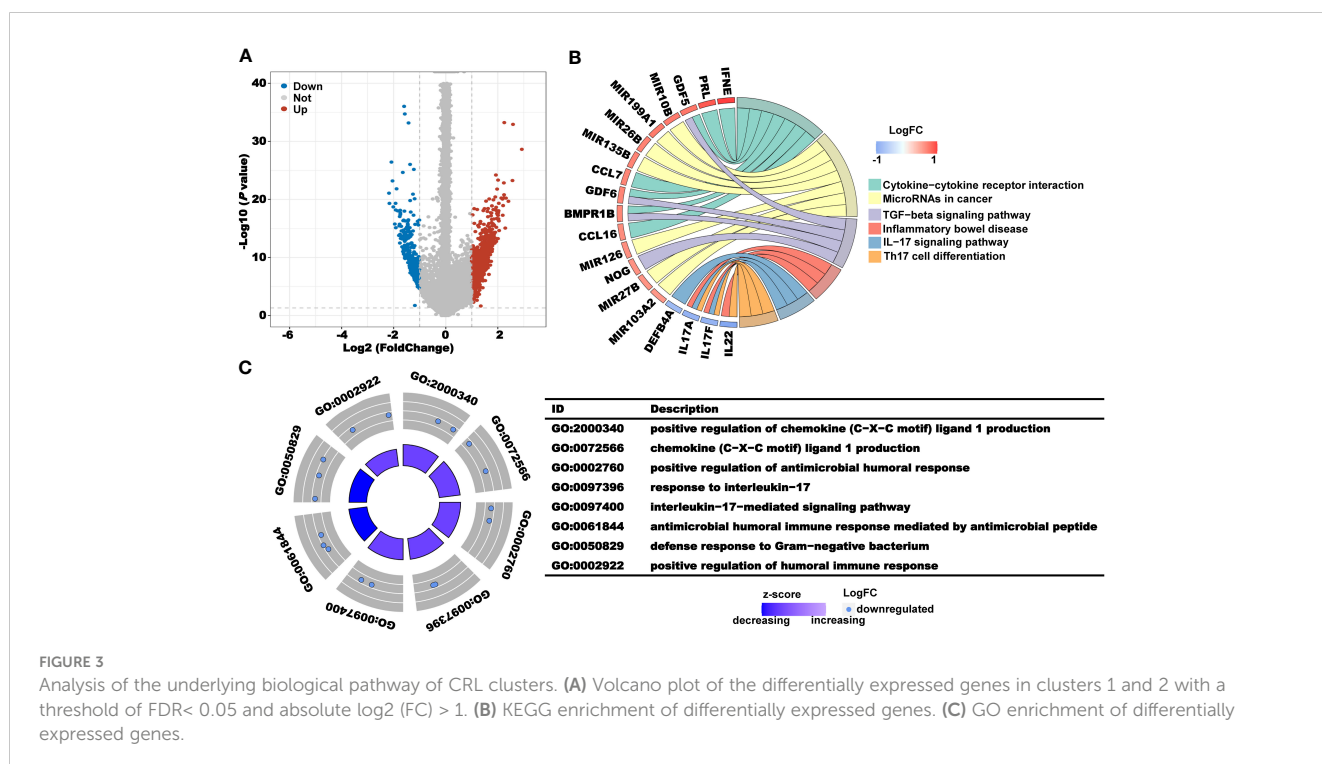
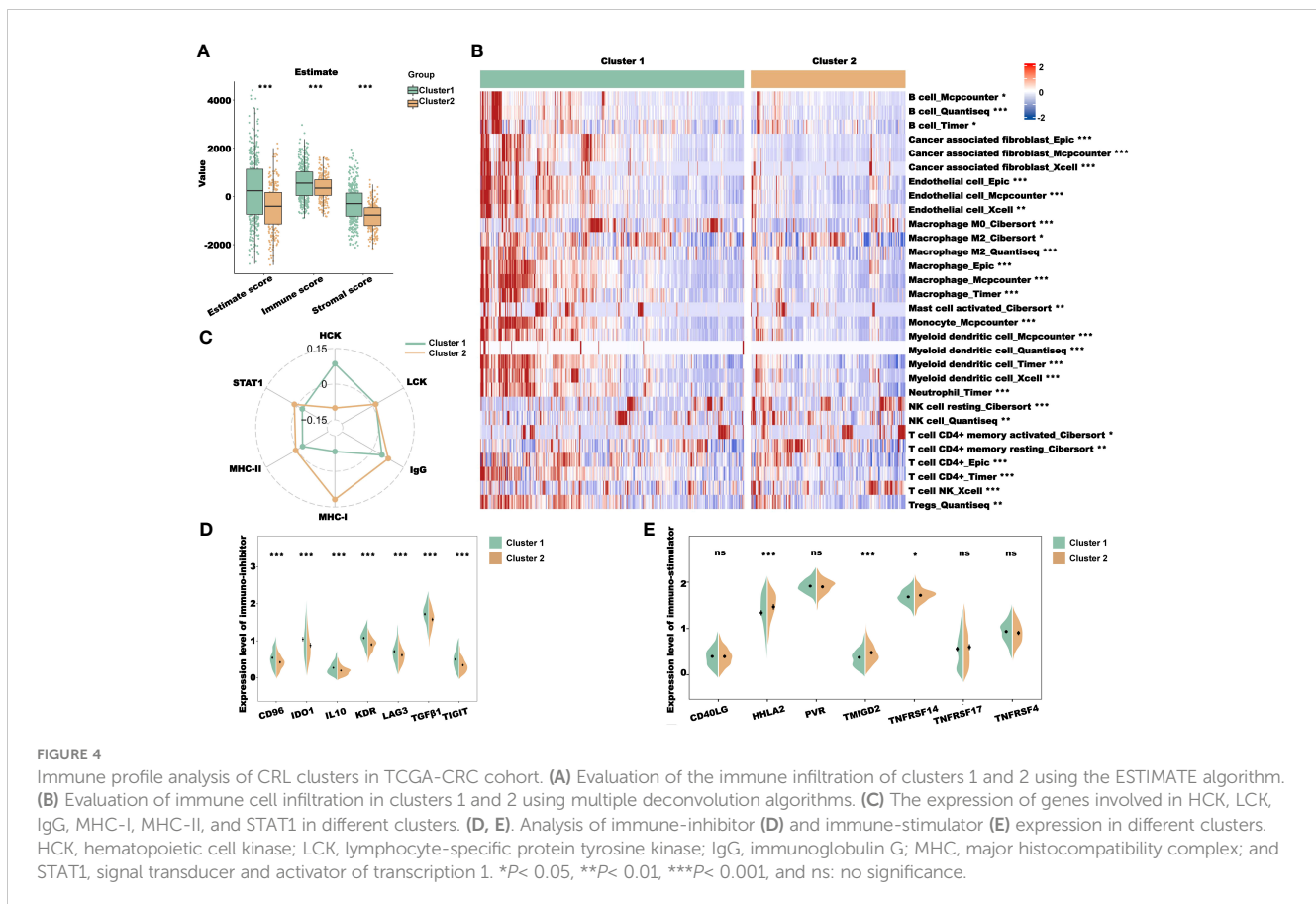


FIGURE 3 Analysis of the underlying biological pathway of CRL clusters. (A) Volcano plot of the differentially expressed genes in clusters 1 and 2 with a threshold of FDR < 0.05 and absolute log₂ (FC) > 1. (B) KEGG enrichment of differentially expressed genes. (C) GO enrichment of differentially expressed genes.



MHC-II, and STAT1 scores of cluster 2 were significantly higher than those of cluster 1. To further explore the association between the prognostic value of cluster and immunity, we analyzed the differences in the expression levels of immunostimulatory and immunosuppressive genes in clusters 1 and 2. We found that the expression levels of most immunosuppressive genes, such as *CD96*, *IDO1*, *IL10*, and *KDR*, increased in cluster 1 (Figure 4D), while the expression levels of some immune stimulatory genes, such as *HHLA2*, *TMIGD2*, and *TNFRSF14*, increased in cluster 2 (Figure 4E). These results further indicated that the prognostic value of clusters was related to immune regulation. The higher immune score of cluster 1 may be due to the higher level of immunosuppressive cell infiltration and immunosuppressive gene expression, which presented an immunosuppressive TME and led to a worse prognosis for patients.

Construction of a prognostic CRLncSig

To construct a prognostic model of CRLs, the aforementioned prognostic CRLs ($P < 0.01$) were used for stepped-regression dimension reduction analysis, and 10 CRLs (*DYNLRB2-AS1*, *EFCAB13-DT*, *EWSAT1*, *LINC00645*, *LINC00901*, *LINC01496*, *LINC01738*, *LINC02962*, *LRP1-AS*, and *PATJ-DT*) were finally obtained for the prognostic model construction. The results of the univariate Cox regression analysis of the 10 CRLs are shown in Figure 5A. Next, a multivariate Cox regression analysis was

employed to derive regression coefficients for individual genes and compute patient risk scores (Figure 5B). Figure 5C displays the distribution of risk scores among TCGA-CRC patients. The patients were categorized into high- and low-risk groups based on the median risk score. Subsequently, a correlation analysis was performed to determine the relationship between the risk score and 10 CRLs (Figure 5D) and the expression differences of the 10 CRLs between the high- and low-risk groups (Figure 5E). The results indicated a high correlation among most of the CRLs. Specifically, the expression levels of *DYNLRB2-AS1*, *EFCAB13-DT*, *EWSAT1*, *LINC01496*, *LINC01738*, *LRP1-AS*, and *PATJ-DT* were significantly higher in the high-risk group than in the low-risk group, while the expression levels of *LINC00645*, *LINC00901*, and *LINC02962* were higher in the low-risk group than in the high-risk group.

Prognostic significance of the CRLncSig

Next, we examined the prognostic role of the risk scores in TCGA-CRC cohort. The results of the K–M survival analysis showed that the patients in the high-risk group had a lower survival probability than those in the low-risk group in terms of OS, DSS, and PFS (Figure 6A). In addition, the results of the K–M survival analysis of the training and validation sets showed that patients with high risk scores had worse survival than other patients (Supplementary Figure 4). To verify the stability of the model, different clinical characteristics were stratified, and K–M survival

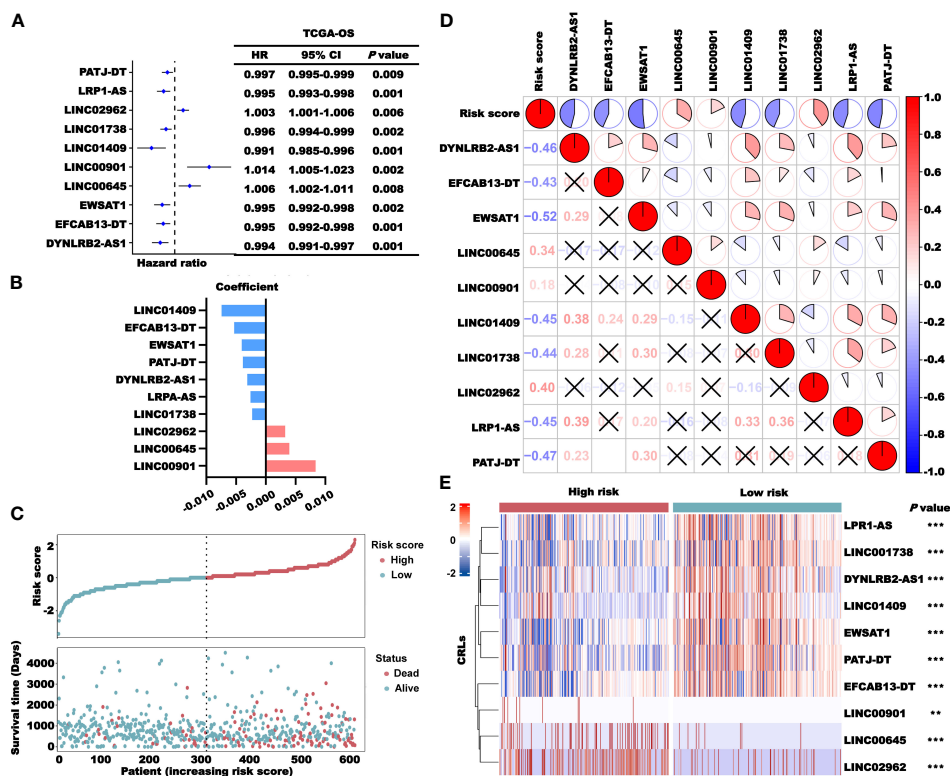


FIGURE 5 Construction of the CRLncSig. (A) Univariate Cox regression analysis of TCGA-OS of the 10 lncRNAs used to construct the signature. (B) Coefficients of the 10 lncRNAs in the prognosis signature. (C) Distribution of the risk score (high and low) and status (dead and alive) in TCGA-CRC cohort. (D) The correlation between the risk score and 10 lncRNAs. (E) Expression profiles of the lncRNAs in the high- and low-risk groups. ** $P < 0.01$ and *** $P < 0.001$.

analysis was performed independently (Supplementary Figures 5, 6). With different clinical characteristics, the patients in the high-risk group still had a worse survival probability. Subsequently, univariate and multivariate Cox regression analyses were performed to verify whether the risk score was an independent prognostic factor for CRC. The results showed that the p -values of the risk score were all lower than 0.05 (Figure 6B), indicating that the risk score constructed based on CRLncSig is an independent prognostic factor for CRC.

Construction and validation of a nomogram combining clinical characteristics

We combined the factors with $P < 0.05$ in TCGA-OS multivariate Cox regression analysis to construct a nomogram model based on the risk score to predict the 1-, 3-, 5-, and 7-year OS of CRC patients (Figure 7A). We then used the calibration curve and ROC curve to verify the accuracy of the nomogram prediction (Figures 7B, C). The calibration curve coincided well with the diagonal, and the area under the curve (AUC) values of the 1-, 3-, 5-, and 7-year nomograms were all larger than 0.7, indicating that the nomogram had a good predictive ability.

Biological pathways and functional enrichment analysis of the CRLncSig

We used the R package “maftool” for the mutation analysis of the high- and low-risk groups (Supplementary Figure 7). There was no significant difference in the frequency of the key gene mutations in the process of cancer occurrence between the two groups, indicating that the difference in prognosis between the two groups may not be related to gene mutations. Subsequently, we used the R package “limma” to excavate the DEGs between the high- and low-risk groups. We found that a total of 2840 genes were highly expressed in the high-risk group and 790 genes were highly expressed in the low-risk group (Figure 8A). We used these DEGs for GO and KEGG pathway enrichment analyses in the high- and low-risk groups. The results of GO enrichment analysis showed that the main enrichment pathways in the high-risk group were cell migration involved in sprouting angiogenesis, ERBB2-ERBB3 signaling pathway, positive regulation of epithelial cell migration, natural killer cell inhibitory signaling pathway, regulation of epithelial cell proliferation, and regulation of platelet-derived growth factor receptor-alpha signaling pathway. In the low-risk group, the mainly enriched pathways were positive regulation of chemokine (C-X-C motif) ligand 1 production, antimicrobial humoral immune response mediated by antimicrobial peptide,

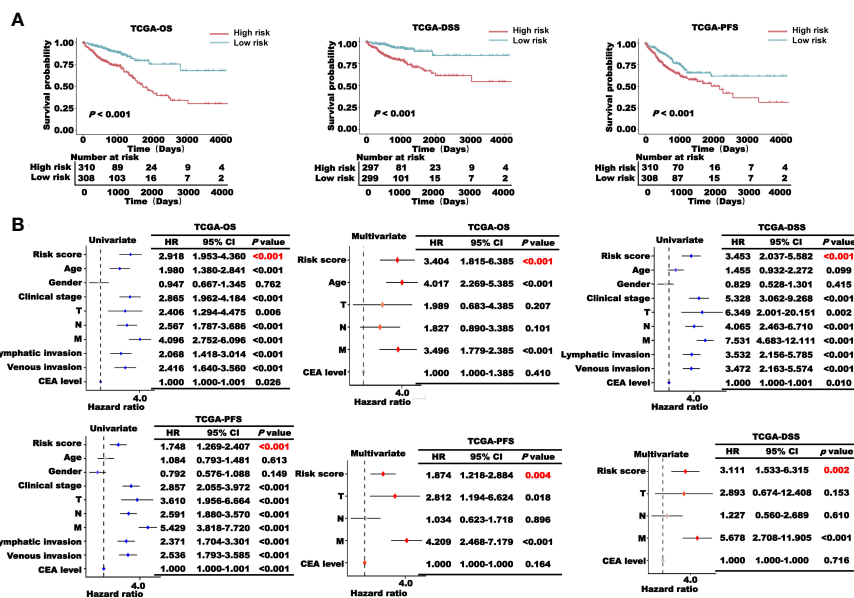


FIGURE 6 Prognosis analysis of the CRLncSig in TCGA-CRC cohort. (A) K–M survival curve of OS, DSS, and PFS of the high- and low-risk groups. (B) Univariate and multivariate regression analyses of TCGA-OS, TCGA-PFS, and TCGA-DSS of the risk score and clinical characteristics.

positive regulation of immunoglobulin production in mucosal tissue, positive regulation of cytokine production involved in inflammatory response, and natural killer cell differentiation involved in immune response (Figure 8B). Similarly, the KEGG enrichment analysis showed the enrichment of cancer immune-

related pathways. microRNAs in cancer, chemical carcinogenesis-receptor activation, chemical carcinogenesis-receptor activation, and JAK–STAT signaling pathway were enriched in the high-risk group, and inflammatory bowel disease was enriched in the low-risk group (Figure 8C). From those enrichment results, we

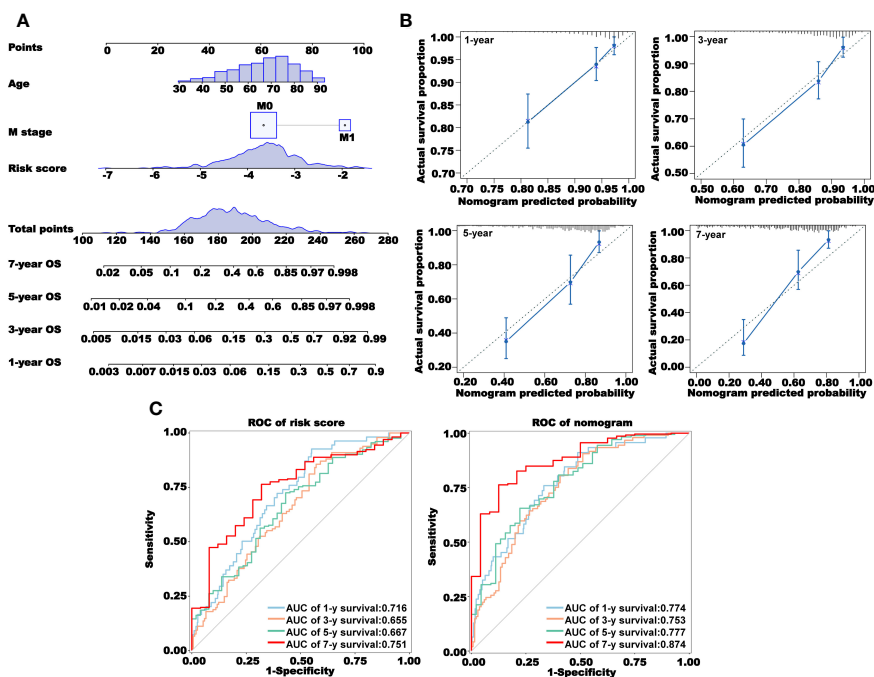
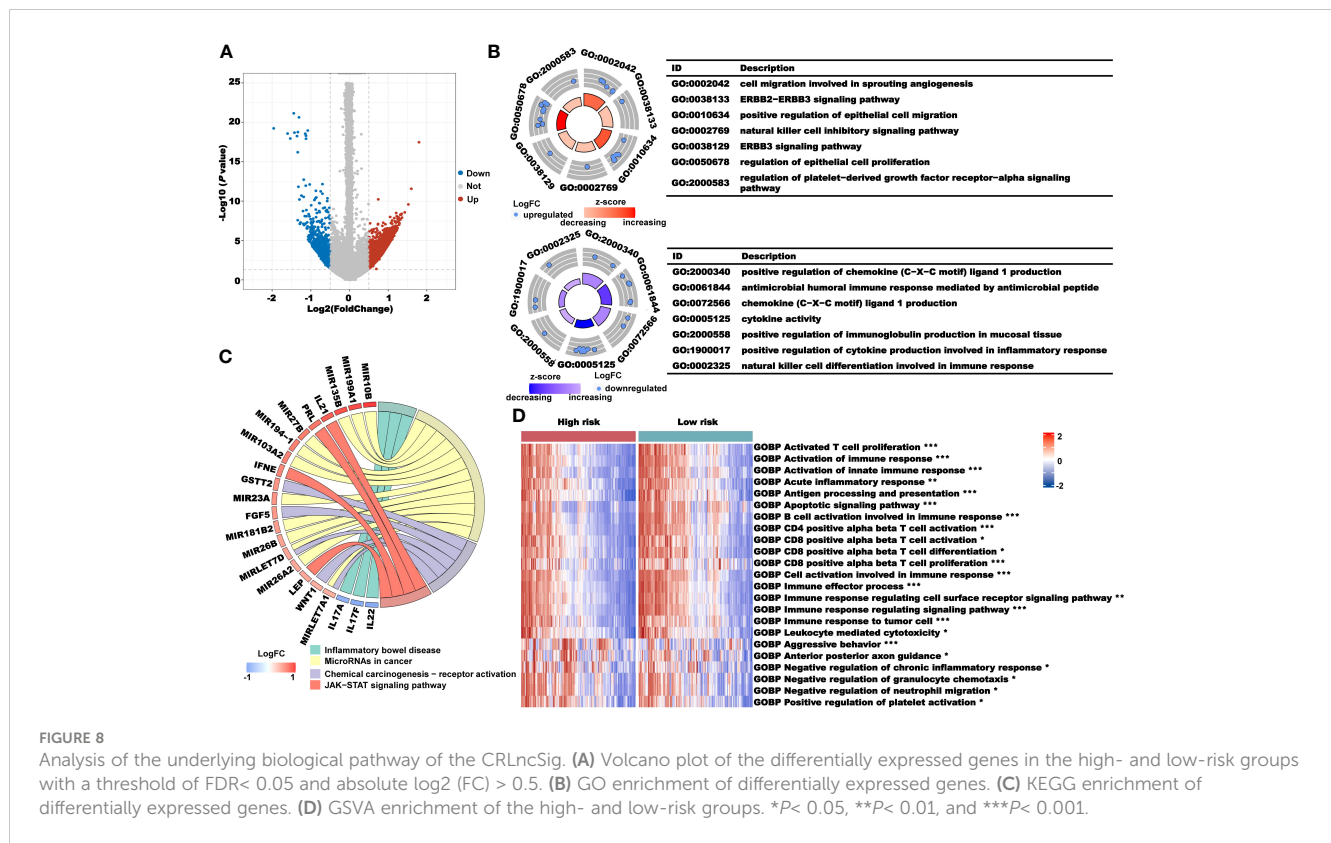


FIGURE 7 Construction and validation of a nomogram. (A) Prediction of 1-, 3-, 5-, and 7-year overall survival using a nomogram constructed using three independent prognostic factors. (B) Calibration curves for 1-, 3-, 5-, and 7-year survival. (C) Receiver operating characteristic (ROC) curves showing the prediction accuracy of the risk score and nomogram. AUC, area under the curve.



speculated that the differential prognosis of the high- and low-risk groups may be mediated through cancer immune pathways. To further verify our conjecture, we performed a GSVA pathway enrichment analysis. As shown in Figure 8D, many immune-related pathways were upregulated in the low-risk group, such as activated T-cell proliferation, activation of the immune response, antigen processing and presentation, CD4-positive alpha beta T-cell activation, CD8-positive alpha beta T-cell activation, differentiation and proliferation, and immune response to tumor cells. The pathways that were significantly upregulated in the high-risk group were aggressive behavior, negative regulation of chronic inflammatory response, and positive regulation of platelet activation. These results further illustrated that the prognostic value of our CRL signature in the occurrence and development of CRC was closely related to cancer immunity.

Immune profile analysis of the CRLncSig

Next, we performed an immune profile analysis of the CRLncSig. The results of the Estimate showed that the estimate, immune, and stromal scores of the high-risk group were higher than those of the low-risk group (Figure 9A). We used several algorithms, such as EPIC, MCP-counter, quantIseq, TIMER, and Xcell, to evaluate the degree of immune cell infiltration in the high- and low-risk groups. We found that the immune infiltration of CAFs, M2 macrophages, myeloid dendritic cells, and regulatory T cells in the high-risk group was higher than that in the low-risk group (Figures 9B–F). In addition, the expression levels of IgG,

MHC family, and STAT1 molecules were higher in the low-risk group than in the high-risk group (Figure 9G). However, most of the immunosuppressive genes, such as *CD96*, *IDO1*, *IL10*, *KDR*, *LAG3*, *TGFB1*, and *TIGIT*, were expressed at higher levels in the high-risk group than in the low-risk group, whereas most of the immune stimulatory genes showed no significant differences between the two groups (Figure 9H). Therefore, we hypothesized that the patients in the high-risk group had higher immune infiltration and presented an immunosuppressed TME, leading to a worse prognosis.

Analysis of immunotherapy response based on the CRLncSig

From the previous results, we saw a significant correlation between the risk score and the patient’s immune microenvironment. Next, we analyzed the expression levels of several key immune checkpoints in the high- and low-risk groups. As shown in Figure 10A, the expression levels of *CTLA4*, *PDI*, *PDL1*, and *PDL2* in the high-risk group were all increased, suggesting that the risk score may be related to the patient’s response to immunotherapy. We used the TIDE algorithm to predict the sensitivity to immunotherapy for TCGA-CRC patients. The TIDE, Exclusion, and CAF scores of the low-risk group were all lower than those of the high-risk group, indicating that the sensitivity of the low-risk group to immunotherapy was higher than that of the high-risk group (Figure 10B). We also used the CYT and GEP scores to predict immunotherapy sensitivity in TCGA-CRC patients. Both the CYT and GEP scores of the low-risk group were higher than those of the high-

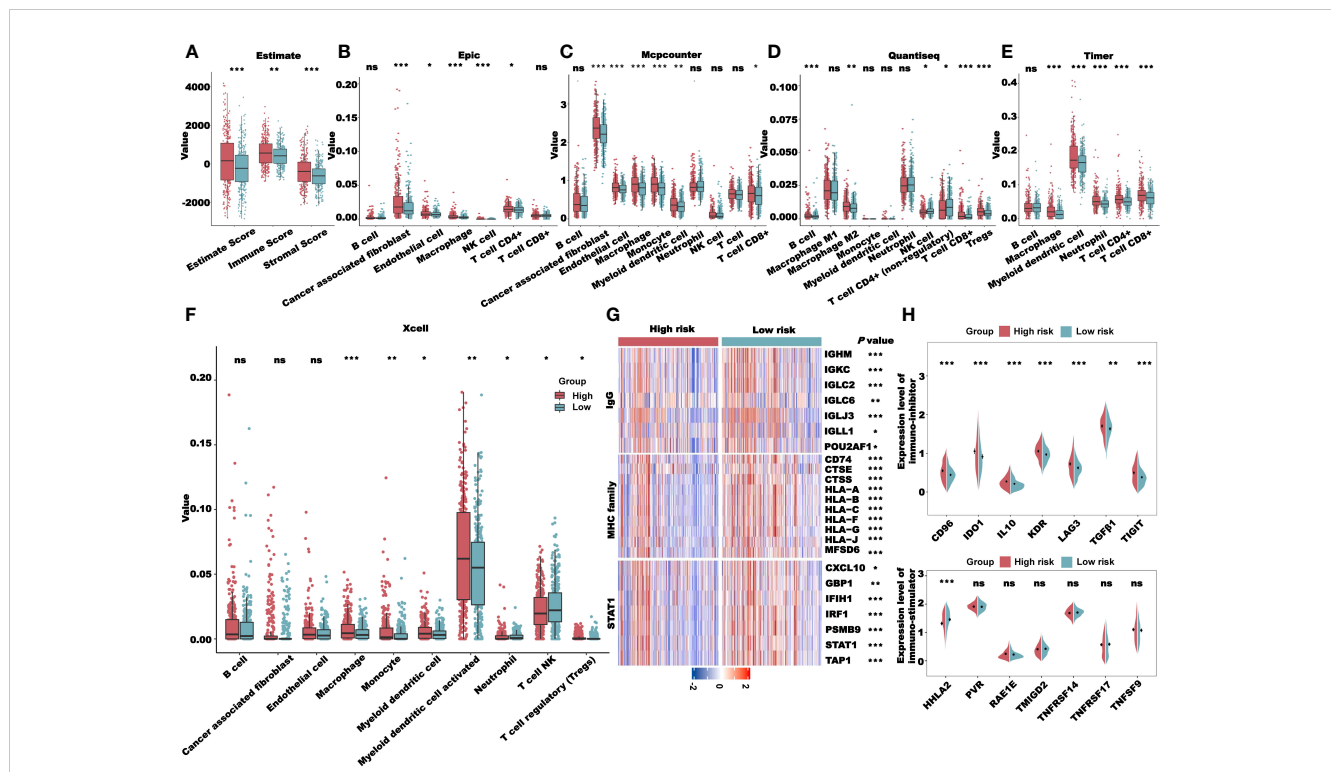


FIGURE 9 Immune profile analysis of the CRLncSig in TCGA-CRC cohort. (A) Evaluation of the immune infiltration of the high- and low-risk groups using the ESTIMATE algorithm. (B–F). Evaluation of immune cell infiltration in the high- and low-risk groups using multiple deconvolution algorithms: EPIC (B), MCP-counter (C), quanTiseq (D), TIMER (E), and Xcell (F). (G) The expression of genes involved in IgG, MHC family, and STAT1 in different groups. (H) Analysis of immune-inhibitor and immune-stimulator expression in different groups. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and ns: no significance.

risk group, which further indicated that the sensitivity of the low-risk group to immunotherapy was higher than that of the high-risk group (Figures 10C–D). The TIDE algorithm can also predict whether a patient will respond to immunotherapy (False/True). As expected, patients with a False responder had a significantly higher risk score than those with a True responder (Figure 10E). Likewise, the proportion of patients with a True responder was significantly higher in the low-risk group than in the high-risk group (Figure 10F). Subsequently, we validated the TIDE results using a thymic cancer cohort receiving immunotherapy. The results showed that immunotherapy response rates were lower in patients with high-risk scores and higher in patients with low-risk scores. Then, we used the risk score to predict the immunotherapy response of patients in GSE181815. The results showed that the risk score had an excellent prediction effect (Supplementary Figure 8).

To explore the relationship between the coagulation-related clusters and immunotherapy sensitivity, we first compared the expression levels of several key immune checkpoints in clusters 1 and 2 and found that the expression levels of immune checkpoints in cluster 1 were higher than those in cluster 2 (Figure 10G). In addition, the TIDE, Exclusion, and CAF scores of cluster 1 were all higher than those of cluster 2, suggesting that cluster 2 patients were more sensitive to immunotherapy (Figure 10H). Similarly, in cluster 2, the proportion of patients with a True responder was significantly higher than that of the high-risk group (Figure 10J), further indicating that the patients in cluster 2 were more sensitive to

immunotherapy. The risk score of cluster 1 was significantly higher than that of cluster 2 (Figure 10I), and the proportion of patients belonging to cluster 1 in the high-risk group was significantly higher than the proportion of patients belonging to cluster 2 (Figure 10K). These data indicated that our cluster and risk score had a certain correlation. To prove this correlation, we combined the OS, cluster, and risk score of TCGA-CRC patients to draw a Sankey diagram (Figure 10L). It was clear that most of the dead patients belonged to cluster 1 and the high-risk group.

Construction of the ceRNA network based on CRLs

We used WGCNA to obtain the CRLs that most correlated with the risk score and cluster, and selected the CRLs in the METurquoise module as the candidate lncRNAs (Figure 11A). To further reduce the number of candidate lncRNAs, we screened the CRLs with the highest correlation in the module (Figure 11B). The red part is the final candidate CRLs, abbreviated as hub-CRLs. WGCNA was also used to obtain the most CRGs associated with the stromal, immune, and estimate scores. As shown in Figure 11C, we chose CRGs in the MEblue as candidate genes, abbreviated as immune-CRGs. Next, we used the miCode database to obtain the miRNA family binding to hub-CRLs, and then obtained the mRNA binding to miRNA through the TargetScan, miCode, and miRTarBase databases,

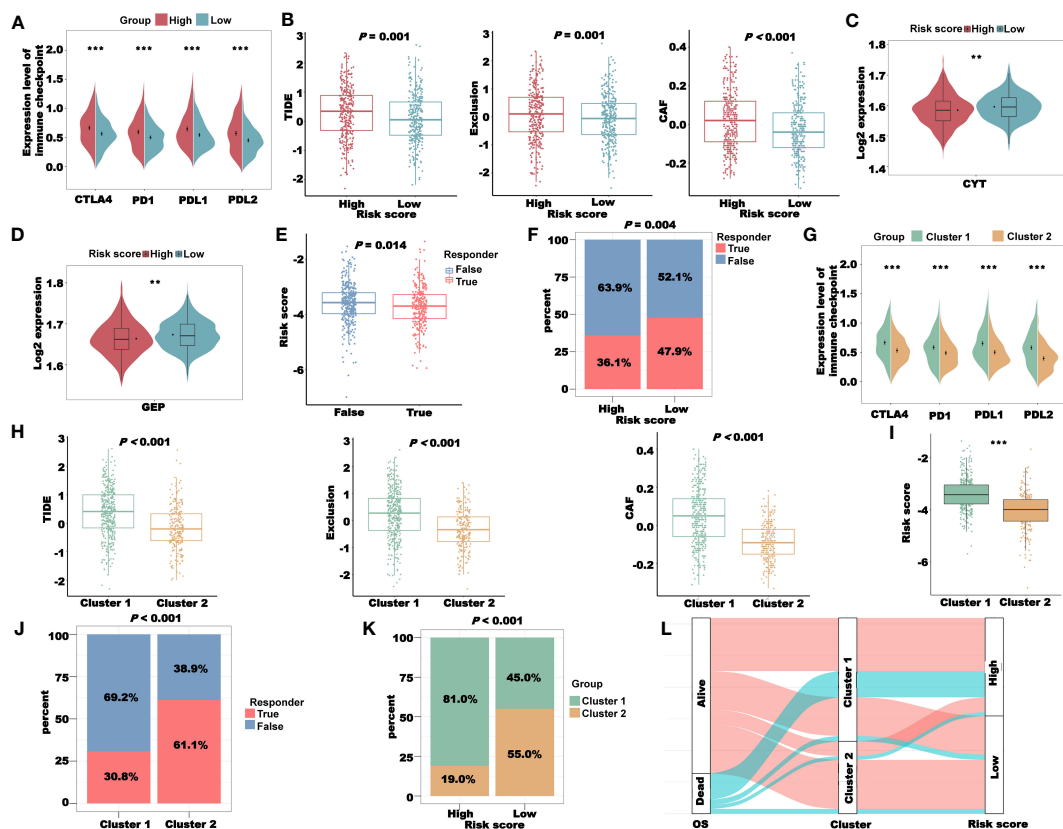


FIGURE 10

Immunotherapy response analysis of the CRLncSig in TCGA-CRC cohort. (A) The expression of key immune checkpoints in the high- and low-risk groups. (B) Prediction of patient's response to immunotherapy in different groups, including TIDE, Exclusion, and CAF scores, using the TIDE algorithm. (C) The CYT score of patients in different groups. (D) The GEP score of patients in different groups. (E) Prediction of the risk score of patients in different responders using the TIDE algorithm. (F) The distribution of different responders in high- and low-risk groups. (G) The expression of key immune checkpoints in clusters 1 and 2. (H) The TIDE, Exclusion, and CAF scores in clusters 1 and 2. (I) The difference in risk score between clusters 1 and 2. (J) The distribution of different responders in clusters 1 and 2. (K) The distribution of clusters in the high- and low-risk groups. (L) Sankey diagram combining the OS, Cluster, and risk scores. ** $P < 0.01$ and *** $P < 0.001$.

and took the intersection with immune-CRGs. Finally, four hub-CRLs, 25 miRNAs, and 14 immune-CRGs were obtained. Subsequently, we drew the ceRNA network using Cytoscape software (Figure 11D). We also analyzed the expression of 14 immune-CRGs in different risk groups and clusters, and the expression levels of most genes, including *THBS1*, *PLAU*, *SPARC*, *GATA3*, *ITP1*, *PDGFRA*, and *PRKG1*, were different among different groups (Figure 11E).

Drug-sensitivity prediction in the CRC patients in the high- and low-risk groups

To better apply our CRLncSig to the clinic, we used the R package "oncoPredict" to predict drug sensitivity. The half-maximal inhibitory concentration (IC_{50}) in GDSC and the area under the dose-response curve (AUC) in CTRP and PRISM negatively correlated with drug sensitivity. Among 198 compounds in the GDSC database, 135 compounds showed significant differences in the IC_{50} values between the high- and low-risk groups ($P < 0.05$). As

shown in Figure 12, the patients in the low-risk group were more sensitive to common chemotherapy drugs, including 5-fluorouracil, docetaxel, oxaliplatin, paclitaxel, and vinorelbine. Compounds targeting the EGFR signaling, ERK/MAPK signaling, and RTK signaling, such as afatinib, AZD3759, erlotinib, ERK_2440, selumetinib, and SB505124, were predicted to be more effective in the high-risk group. However, osimertinib, sapitinib, trametinib, ulixertinib, VX-11e, AZD4547, cediranib, crizotinib, and savolitinib were predicted to be more effective in the low-risk group.

In Figure 13A, we selected compounds with a strong correlation with the risk score from the GDSC database to draw bubble plots to show both pathway and target information of the compounds. To screen more drugs suitable for high-risk patients, the intersection of drugs with more sensitivity in the high-risk group in the CTRP and PRISM databases was conducted, and a total of 38 compounds were obtained (Figure 13B). Figures 13C, D showed some of these compounds, such as afatinib, bosutinib, erlotinib, and sorafenib. The above results proved that our CRLncSig can well predict drug sensitivity in patients of TCGA-CRC cohort and is expected to become a drug guidance strategy for patients with CRC in the future.

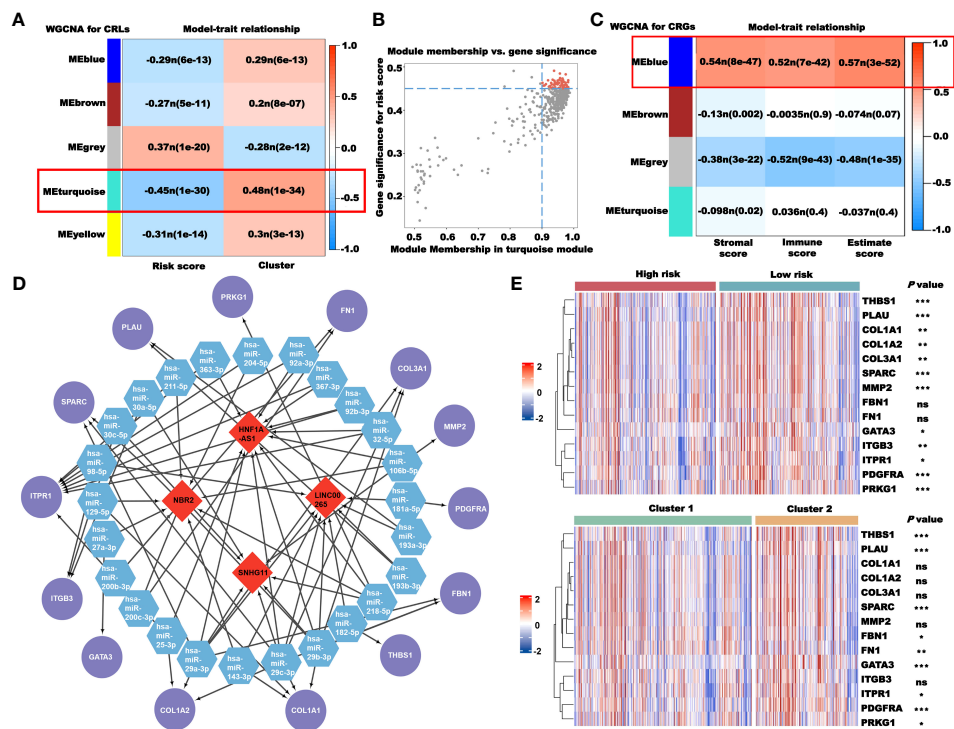


FIGURE 11 Construction of the ceRNA network based on CRLs. (A) Selection of the CRLs that correlated with the risk score and clusters using WGCNA. (B) Screening of key CRLs in METurquoise using correlation scatter map. (C) Selection of the CRGs that correlated with immunity using WGCNA. (D) The ceRNA network constructed using Cytoscape (red: CRLs; blue: miRNAs; and purple: CRGs). (E) The expression of 14 CRGs in different risk groups and clusters. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and ns: no significance.

Validation of the CRLncSig

We next validated the clinical significance of our CRLncSig. First, we validated the prognostic value of the risk score in CRC using the GEO dataset GSE147602. The results showed that in GSE147602, CRC patients without metastasis had higher risk scores than those with metastasis, while there were no significant differences in the risk score between patients of different genders and ages (Figure 14A). We divided the patients in GSE147602 into the high- and low-risk groups based on the median and found that the proportion of patients without metastasis in the low-risk group was much higher than that with metastasis (Figure 14B). Then, we used another GEO dataset, GSE198103, to further verify the ability of this model to predict the prognosis of patients. In GSE198103, the risk scores of patients with a distant metastasis stage were significantly higher than those of patients with a localized stage (Supplementary Figure 9). The ability of the coagulation-related lncRNA signature to predict the prognosis of CRC patients was validated with these two cohorts. In addition, we examined the expression of the 10 lncRNAs used to construct the signature in two types of human CRC cells (HCT116 and LOVO) and one type of human normal epithelial cells (HcoEpic) (Figures 14C, D). The expression levels of *DYNLRB2-AS1*, *EFCAB13-DT*, *EWSAT1*, *LINC00645*, *LINC00901*, *LINC01738*, *LINC02962*, and *LRP1-AS* in

HCT116 and LOVO cells were significantly downregulated compared with those in HcoEpic cells. However, the expression level of *LINC01409* in HCT116 cells was not significantly different from that in HcoEpic cells but was significantly upregulated in LOVO cells. In contrast, there was no significant difference in the expression level of *PATJ* between LOVO and HcoEpic cells, but it was significantly downregulated in HCT116 cells. We also confirmed our previous results in 51 CRC patients. As shown in Figure 14E, *DYNLRB2-AS1*, *EFCAB13-DT*, *EWSAT1*, *LINC01409*, *LINC01738*, *LINC02962*, *LRP1-AS*, and *PATJ-DT* were downregulated in the tumor tissues, while there were no significant differences in the expression level of *LINC00645* and *LINC00901* between normal tissue and tumor tissue. We also plotted the correlation graph between the risk score and 10 lncRNAs (Figure 14F). In addition, we examined the expression levels of *PD1*, *PDL1*, *CTLA4*, *TGFB1*, *TIGIT*, *IDO1*, and *LAG3* in 51 tumor samples. According to the median, 51 patients were divided into the high- and low-risk groups. We found that the expression levels of *PD1*, *PDL1*, *TGFB1*, *TIGIT*, *IDO1*, and *LAG3* in the high-risk group were significantly higher than those in the low-risk group (Figure 14G). The risk score showed a significant negative correlation with the expression levels of *PD1*, *CTLA4*, *TGFB1*, and *TIGIT* (Figure 14H), which further proved our previous findings.

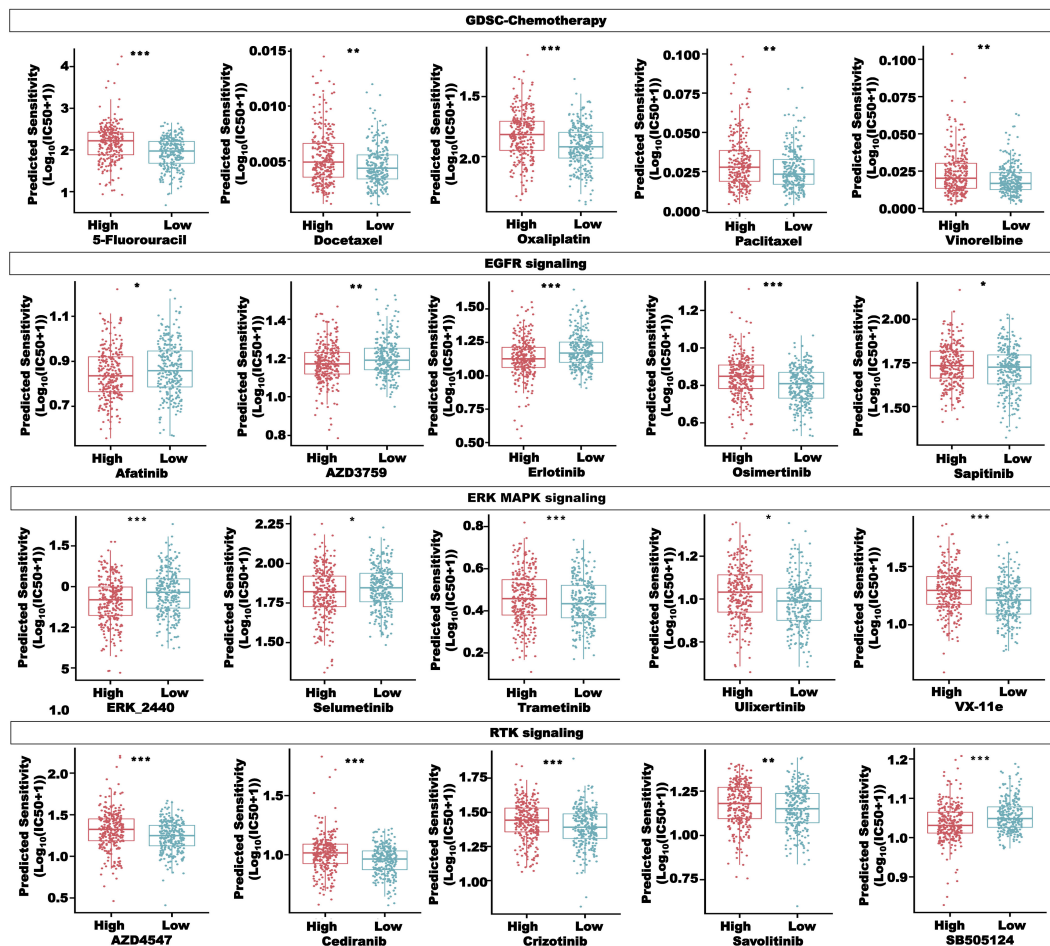


FIGURE 12 Drug sensitivity analysis in the high- and low-risk groups, including compounds in the GSDC database, targeting chemotherapy, EGFR signaling, ERK MAPK signaling, and RTK signaling. * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$.

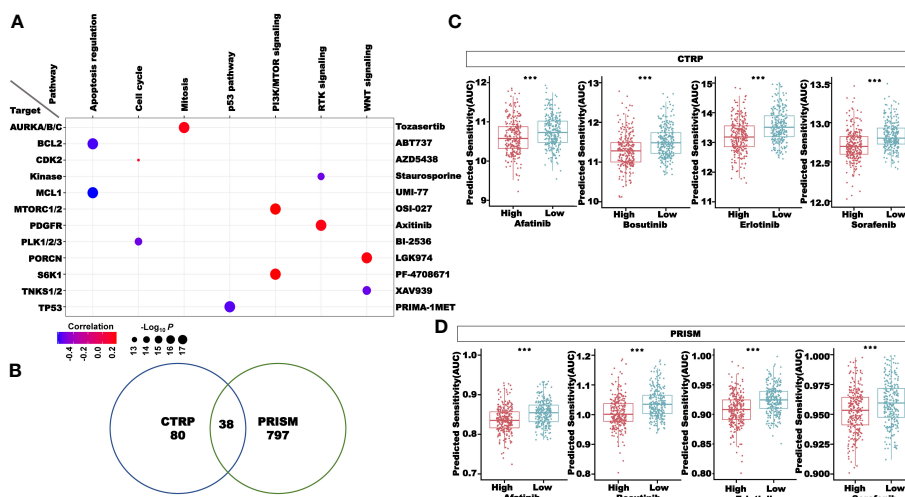


FIGURE 13 Drug sensitivity analysis in the high- and low-risk groups. (A) The correlation between the risk score and compounds in the GDSC database. (B) Sensitive compounds for the high-risk group shared by the CTRP and PRISM databases. (C) Four compounds sensitive for the high-risk group in the CTRP database. (D) Four compounds sensitive for the high-risk group in the PRISM database. *** $P < 0.001$.

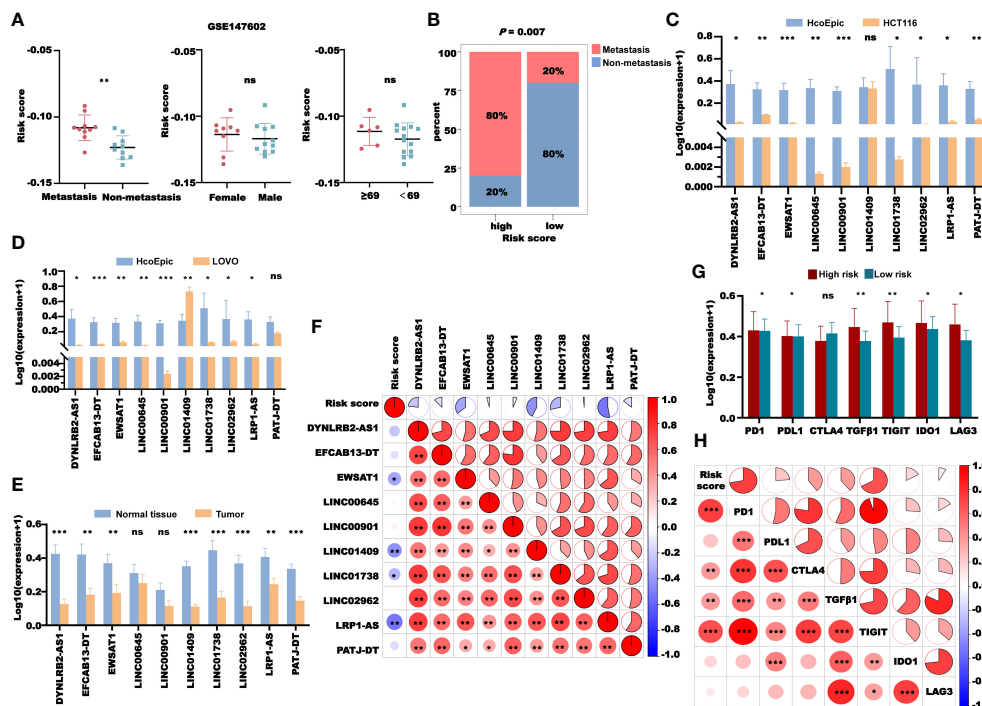


FIGURE 14 Validation of the CRL signature. (A) Risk score distribution in GSE147602 with different clinical features. (B) The distribution of patients with or without metastasis in the high- and low-risk groups in GSE147602. (C) Comparison of the expression of 10 lncRNAs between HcoEpic cells and HCT116 cells. (D) Comparison of the expression of 10 lncRNAs between HcoEpic cells and LOVO cells. (E) Comparison of the expression of 10 lncRNAs in normal and tumor clinical samples. (F) The correlation between the risk score and 10 lncRNAs in tumor samples. (G) The expression of PD1, PDL1, CTLA4, TGFβ1, TIGIT, IDO1, and LAG3 in the tumor samples of the high- and low-risk groups. (H) The correlation between the risk score and the expression of PD1, PDL1, CTLA4, TGFβ1, TIGIT, IDO1, and LAG3 in tumor samples. * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$, and ns: no significance.

scRNA-seq analysis of the CRLncSig

We used the scRNA-seq cohort GSE136394 to conduct a single-cell analysis of the CRL signature. After TSNE reduction and cell annotation, we obtained seven cell clusters, including transitional CD8 + T cells, cytotoxic CD4+ T cells, NK cells, cycling cells, effector CD4+ T cells, dendritic cells, and mast cells (Figure 15A). A total of five samples were included in the GSE136394 dataset, among which four channels were used for single-cell sequencing for the sample GSM4047944. Figure 15B shows the proportion of different cells in each sample. Next, we explored the cellular distribution of the 10 lncRNAs used to build the signature. We found that *DYNLRB2-AS1* was mainly distributed in transitional CD8+ T cells, cytotoxic CD4+ T cells, and NK cells. *EFCAB13-DT* was mainly distributed in cytotoxic CD4+ T cells and NK cells, and *EWSAT1* was mainly distributed in transitional CD8+ T cells and effector CD4+ T cells. *LINC01409* was distributed in all of the clustered cells, while *PATJ-DT* was mainly distributed in transitional CD8+ T cells, NK cells, and effector CD4+ T cells. However, the content of *LINC00645*, *LINC00901*, *LINC01738*, *LINC02962*, and *LRP1-AS* in these cells was zero (Figure 15C).

Discussion

Tumor cells can activate the coagulation system and induce a hypercoagulable state in patients with malignant tumors. Previous

studies have shown that the incidence of VTE is significantly higher in patients with cancer than in those without cancer (57). It has been reported that the gastrointestinal tract is one of the sites of high incidence of blood coagulation disorders (58, 59). Thromboembolic events are a common cause of death in patients with CRC, even in those with a good cancer prognosis (60). The significant relationship between CRC and thrombosis raises the possibility that the coagulation pathway could be a therapeutic target. lncRNAs are a kind of highly stable noncoding RNAs, and they are also a kind of reliable biomarker of cancer. Currently, the construction of models to predict the prognosis of patients has been proven to be effective and meaningful in many types of cancer (61, 62). In this study, we identified two CRC subtypes based on CRLs and constructed a CRC prognostic signature consisting of 10 CRLs.

The CRLncSig was composed of *DYNLRB2-AS1*, *EFCAB13*, *EWSAT1*, *LINC00645*, *LINC00901*, *LINC01409*, *LINC01738*, *LINC02962*, *LRP1-AS*, and *PATJ-DT*, and six of them (*DYNLRB2-AS1*, *EFCAB13*, *LINC01409*, *LINC01738*, *LRP1-AS*, and *PATJ-DT*) had not been previously reported for their role in tumors. *EWSAT1* has been reported to play a pro-tumor role in a variety of cancers. *EWSAT1* can contribute to the proliferation and invasion of glioma (63), promote HCC metastasis (64), and promote the progression of ovarian cancer (65). Studies on *EWSAT1* in CRC have also been reported. In CRC, the increased expression of *EWSAT1* promotes the proliferation, invasion, and epithelial-mesenchymal transition of CRC cells (66), and also promotes the progression of CRC by

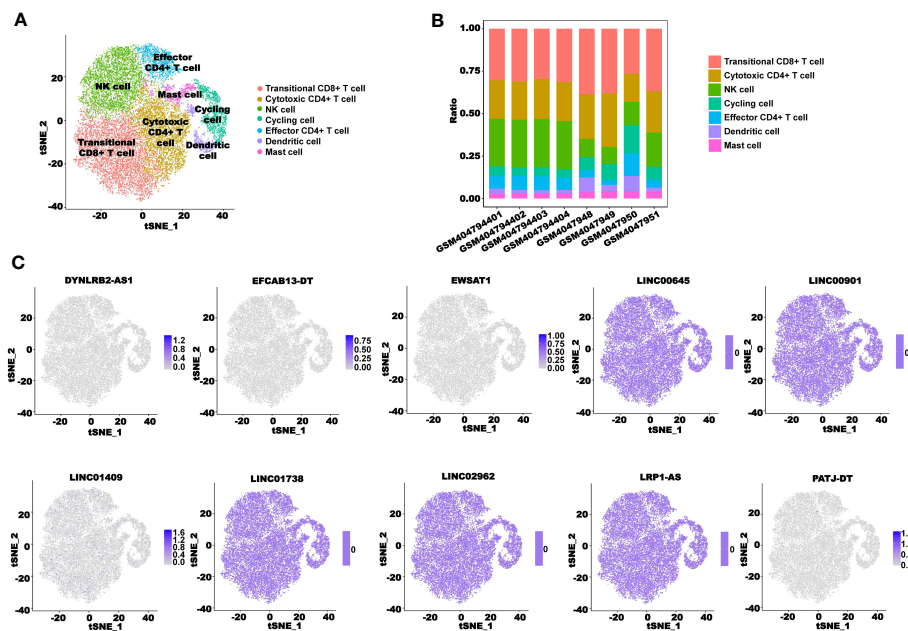


FIGURE 15
scRNA-seq analysis of the CRLncSig. (A) TSNE plot visualization of all cell subtypes from five CRC patients in GSE136394. (B) Bar plots of the cell proportions in each CRC patient. (C) TSNE plot visualization of the distribution of 10 CRLs.

regulating *FBXL20* expression through sponging *miR-326* (67). Li et al. (68) found that *LINC00645* promoted TGF- β -induced EMT in gliomas by regulating the *miR-205-3p-ZEB1* axis, thereby promoting tumor malignant progression. The study also found that the abnormal expression of *LINC00901* could promote the growth and invasion of pancreatic cancer cells (69). *LINC02962*, also known as *lnc-CCDST*, was found to be significantly downregulated in cervical cancer tissues and bind to pre-carcinogenic DHX9 and E3 ubiquitin ligase MDM2, thereby affecting cervical cancer cell invasion and angiogenesis (70). There are few studies on CRLs in CRC, which has a very good prospect. The signature constructed based on the CRLs also showed excellent clinical value.

Our results indicated that the subtypes and signatures based on CRLs can well distinguish the prognosis of CRC patients. CRC is a highly heterogeneous malignant tumor. The current pathological staging system completely depends on the anatomical degree of the tumor and cannot fully reflect the biological heterogeneity of CRC patients, which affects the accuracy of traditional methods to predict the prognosis of CRC. Herein, we constructed a nomogram model with age, M stage, and risk score to help determine disease progression and personalize diagnosis. The constructed nomogram showed excellent predictive power.

The crosstalk between coagulation and innate immunity is complicated and has been reported in many studies (71, 72). Coagulation is a key component of innate immunity because it prevents the spread of bacteria and can cause inflammation (73). By performing the pathway enrichment analysis, we showed that inflammatory bowel disease and immune-related pathways were enriched in the established signature and clusters. In addition, our immune profile results showed that cluster 1 and the high-risk group

were characterized by immunosuppressive microenvironments, in which multiple algorithms indicated a higher proportion of CAFs. CAFs are one of the key components of tumor mesenchyme, which not only provide physical support for epithelial cells but also serve as key functional regulators in cancer. It has been reported that CAFs are associated with poor prognosis and chemotherapy resistance in multiple solid tumors (74, 75). Furthermore, we found that the expression levels of immunosuppressive factors were significantly increased in cluster 1 and the high-risk group, which means that our CRLs in CRC can inhibit the anticancer immune response of patients, thereby leading to the malignant progression.

Interestingly, our study highlighted the potential role of the CRLncSig in predicting the response to immunotherapy in CRC. TIDE, CYT, and GEP scores can predict the response rate of cancer patients to immunotherapy. A greater TIDE score and lower CYT and GEP scores are associated with a greater possibility for immune evasion and exclusion and a lower chance of immunotherapy benefits (40–42). The patients in the high-risk group had higher TIDE scores and lower CYT and GEP scores than the patients in the low-risk group, indicating that the patients in the high-risk group had a lower response rate to immunotherapy. Meanwhile, the TIDE algorithm showed a substantial difference between the nonresponders and responders and predicted a much greater probability of ICI responders among the low-risk patients, explaining why the high-risk patients had a poor prognosis. Additionally, the investigation of drug sensitivity revealed significant differences in the responses to various chemotherapeutic treatments and molecularly targeted medications between the high- and low-risk patients. These results demonstrate that using the risk signature to direct the application of chemotherapy and targeted therapy may be beneficial. As a result, an approach optimizing regimens of a combination of immunotherapy,

chemotherapy, and targeted therapy based on the novel CRL signature may be effective for the individualized treatment of patients with CRC.

The ceRNA network is an essential lncRNA-mediated regulatory molecular mechanism that may be used for the investigation of biomarkers, therapeutic targets, and molecular mechanisms (76). In cancer-related studies, most lncRNAs can participate in tumor progression by targeting microRNAs. It has been reported that *lncRNA-CDC6* can promote breast cancer progression by sponging *microRNA-215* (77). *lncRNA-RMRP* can promote the proliferation, migration, and invasion of bladder cancer through *miR-206* (78). Considering the interactions between lncRNAs and microRNAs, we constructed an lncRNA-miRNA-mRNA ceRNA network to explore the potential regulatory mechanisms of CRLs. Through screening, we obtained four key CRLs, namely, *HNF1A-AS1*, *LINC00265*, *NBR2*, and *SNHG11*. Interestingly, *HNF1A-AS1* can induce resistance to 5-FU in gastric cancer cells via the *miR-30b-5p/EIF5A2* pathway (79). In CRC, *LINC00265* promotes glycolysis and lactate production by regulating the *miR-216b-5p/TRIM44* axis (80), and *NBR2* inhibits cell migration and invasion by downregulating *miRNA-21* (81). In addition, *SNHG11* could enhance bevacizumab resistance in CRC by mediating the *miR-1207-5p/ABCC1* axis (82). In this study, we identified a new endogenous network of competing CRLs with microRNA/mRNA, which will provide new ideas for future mechanistic studies of CRLs in the malignant progression of CRC.

In the validation of the CRLncSig at the cellular level, the expression levels of most lncRNAs were decreased in CRC cells (HCT116 and LOVO cells) compared with normal human epithelial cells (HcoEpic cells). These results were verified in CRC and adjacent normal tissue samples. However, we noticed that the expression level of *LINC01409* was significantly higher in LOVO cells than in HcoEpic cells. CRC tissues contained cancer cells and stromal cells such as immune cells, fibroblasts, and blood vessels, which may explain the inconsistencies in RT-qPCR results from our cell and tissue samples. To detect the cell distribution of lncRNAs used to construct the signature in patients with CRC, we performed a single-cell sequencing analysis. From the results, we could see that *LINC01409* was distributed in most immune cells, while other lncRNAs were relatively less distributed in immune cells. Moreover, RT-qPCR was used to detect the expression levels of immune checkpoints and immunosuppressors in CRC tissues, which was consistent with the results in TCGA cohort, further confirming the reliability of our CRL model.

The results showed that our CRLncSig has a good prospect in the precision diagnosis and treatment of CRC. Nonetheless, there are still some limitations to our study. First, most of the current public datasets of cancer-related sequencing fail to contain complete lncRNA expression information; as a result, the number of independent verification sets available is extremely limited. Therefore, only three independent validation sets, GSE147602, GSE198103, and GSE181815, were used to validate our signature in this study, and more independent CRC cohorts, as well as multicenter, large-scale prospective studies, should be integrated to confirm our findings in the future. Second, the mechanisms by which these CRLs affect the immune microenvironment and drug

sensitivity of CRC patients are still unclear. Despite our functional enrichment and ceRNA network construction, the biological functions and interactions of these CRLs still need further experimental studies *in vivo* and *in vitro*.

Collectively, this is the most systematic exploration of the clinical significance of CRLs in CRC patients to date. We successfully developed and validated coagulation subtypes and a novel CRL model in CRC. The model performed accurately in predicting the prognosis, immune status, immunotherapy response, and drug sensitivity of CRC patients. This study may provide an innovative perspective for clinical prognosis prediction of patients with CRC and could help deepen the theoretical basis for immunotherapeutic improvement and individualized antitumor therapy.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Research Ethics Committee of The Affiliated Hospital of Qingdao University. The patients/participants provided their written informed consent to participate in this study. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

FZ: Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft. RZ: Data curation, Methodology, Writing – original draft. JZ: Data curation, Resources, Writing – original draft. YH: Methodology, Software, Writing – review & editing. MZ: Formal analysis, Validation, Writing – review & editing. ZY: Investigation, Methodology, Writing – review & editing. TL: Software, Visualization, Writing – review & editing. WG: Data curation, Resources, Writing – review & editing. SL: Investigation, Methodology, Writing – review & editing. LY: Validation, Writing – review & editing. ZZ: Data curation, Writing – review & editing. WZ: Formal analysis, Writing – review & editing. MY: Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Natural Science Foundation of China (NSFC)

Grant (No. 81973325, No. 81773750) and CAMS Innovation Fund for Medical Sciences (CIFMS) Grant (No. 2021-I2M-1-028).

Acknowledgments

The authors are thankful for the publicly available databases.

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.

References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
- Sottoriva A, Kang H, Ma Z, Graham TA, Salomon MP, Zhao J, et al. A big bang model of human colorectal tumor growth. *Nat Genet* (2015) 47(3):209–16. doi: 10.1038/ng.3214
- Olawaiye AB, Baker TP, Washington MK, Mutch DG. The new (Version 9) american joint committee on cancer tumor, node, metastasis staging for cervical cancer. *CA: Cancer J Clin* (2021) 71(4):287–98. doi: 10.3322/caac.21663
- Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Song S, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med* (2015) 21(11):1350–6. doi: 10.1038/nm.3967
- Waldman AD, Fritz JM, Lenardo MJ. A guide to cancer immunotherapy: from T cell basic science to clinical practice. *Nat Rev Immunol* (2020) 20(11):651–68. doi: 10.1038/s41577-020-0306-5
- Mahoney KM, Rennert PD, Freeman GJ. Combination cancer immunotherapy and new immunomodulatory targets. *Nat Rev Drug Discovery* (2015) 14(8):561–84. doi: 10.1038/nrd4591
- Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, et al. Immunotherapy in colorectal cancer: rationale, challenges and potential. *Nat Rev Gastroenterol Hepatol* (2019) 16(6):361–75. doi: 10.1038/s41575-019-0126-x
- Serrano PE, Parpia S, Linkins LA, Elit L, Simunovic M, Ruo L, et al. Venous thromboembolic events following major pelvic and abdominal surgeries for cancer: A prospective cohort study. *Ann Surg Oncol* (2018) 25(11):3214–21. doi: 10.1245/s10434-018-6671-7
- Levi M. Cancer-related coagulopathies. *Thromb Res* (2014) 133 Suppl 2:S70–5. doi: 10.1016/s0049-3848(14)50012-6
- Levi M. Disseminated intravascular coagulation in cancer patients. *Best Pract Res Clin Haematol* (2009) 22(1):129–36. doi: 10.1016/j.beha.2008.12.005
- Li J, Du J, Wang Y, Jia H. A coagulation-related gene-based prognostic model for invasive ductal carcinoma. *Front Genet* (2021) 12:722992. doi: 10.3389/fgene.2021.722992
- Ma Y, Wang B, He P, Qi W, Xiang L, Maswikiti EP, et al. Coagulation- and fibrinolysis-related genes for predicting survival and immunotherapy efficacy in colorectal cancer. *Front Immunol* (2022) 13:1023908. doi: 10.3389/fimmu.2022.1023908
- Xie J, Zou Y, Ye F, Zhao W, Xie X, Ou X, et al. A novel platelet-related gene signature for predicting the prognosis of triple-negative breast cancer. *Front Cell Dev Biol* (2021) 9:795600. doi: 10.3389/fcell.2021.795600
- Song B, Chi H, Peng G, Song Y, Cui Z, Zhu Y, et al. Characterization of coagulation-related gene signature to predict prognosis and tumor immune microenvironment in skin cutaneous melanoma. *Front Oncol* (2022) 12:975255. doi: 10.3389/fonc.2022.975255
- He Q, Yang J, Jin Y. Immune infiltration and clinical significance analyses of the coagulation-related genes in hepatocellular carcinoma. *Briefings Funct* (2022) 23(4):1–21. doi: 10.1093/bib/bbac291
- Oshi M, Sarkar J, Tokumaru Y, Yan L, Kosaka T, Akiyama H, et al. Higher intratumoral expression of pro-coagulation genes is a predictor of angiogenesis, epithelial mesenchymal transition and worse patient survival in gastric cancer. *Am J Cancer Res* (2022) 12(8):4001–14.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1279789/full#supplementary-material>

- Guttman M, Rinn JL. Modular regulatory principles of large non-coding rnas. *Nature* (2012) 482(7385):339–46. doi: 10.1038/nature10887
- Batista PJ, Chang HY. Long noncoding rnas: cellular address codes in development and disease. *Cell* (2013) 152(6):1298–307. doi: 10.1016/j.cell.2013.02.012
- Yan H, Bu P. Non-coding rna in cancer. *Essays Biochem* (2021) 65(4):625–39. doi: 10.1042/ebc20200032
- Song J, Zhang W, Wang S, Liu K, Song F, Ran L. A panel of 7 prognosis-related long non-coding Rnas to improve platinum-based chemoresistance prediction in ovarian cancer. *Int J Oncol* (2018) 53(2):866–76. doi: 10.3892/ijo.2018.4403
- Yu X, Lin Y, Sui W, Zou Y, Lv Z. Analysis of distinct long noncoding rna transcriptional fingerprints in pancreatic ductal adenocarcinoma. *Cancer Med* (2017) 6(3):673–80. doi: 10.1002/cam4.1027
- Du J, Chen M, Liu J, Hu P, Guan H, Jiao X. Lncrna F11-as1 suppresses liver hepatocellular carcinoma progression by competitively binding with mir-3146 to regulate pten expression. *J Cell Biochem* (2019) 120(10):18457–64. doi: 10.1002/jcb.29163
- Liu L, Yan B, Yang Z, Zhang X, Gu Q, Yue X. Ncrupar inhibits gastric cancer progression by down-regulating protease-activated receptor-1. *Tumour Biol J Int Soc Oncodevelopmental Biol Med* (2014) 35(8):7821–9. doi: 10.1007/s13277-014-2042-6
- Xu R, Yu J, Song S, Sun D, Xiu L, Xu J, et al. Long non-coding rna Ncrupar regulates gastric cancer cell proliferation and apoptosis via phosphoinositide 3-kinase/protein kinase B signaling. *Int J Med Sci* (2022) 19(13):1835–46. doi: 10.7150/ijms.76664
- Zhi Q, Wan D, Ren R, Xu Z, Guo X, Han Y, et al. Circular rna profiling identifies Circ102049 as a key regulator of colorectal liver metastasis. *Mol Oncol* (2021) 15(2):623–41. doi: 10.1002/1878-0261.12840
- Glibetic N, Shvetsov YB, Aan FJ, Peplowska K, Hernandez BY, Matter ML. Transcriptome profiling of colorectal tumors from patients with sepsis reveals an ethnic basis for viral infection risk and sepsis progression. *Sci Rep* (2022) 12(1):20646. doi: 10.1038/s41598-022-24489-8
- He Y, Ramesh A, Gusev Y, Bhuvaneshwar K, Giaccone G. Molecular predictors of response to pembrolizumab in thymic carcinoma. *Cell Rep Med* (2021) 2(9):100392. doi: 10.1016/j.xcrm.2021.100392
- Lu YC, Jia L, Zheng Z, Tran E, Robbins PF, Rosenberg SA. Single-cell transcriptome analysis reveals gene signatures associated with T-cell persistence following adoptive cell therapy. *Cancer Immunol Res* (2019) 7(11):1824–36. doi: 10.1158/2326-6066.Cir-19-0299
- Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* (2018) 28(11):1747–56. doi: 10.1101/gr.239244.118
- Yu G, Wang LG, Han Y, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. *Omics J Integr Biol* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118
- Hänzelmann S, Castelo R, Guinney J. Gsva: gene set variation analysis for microarray and Rna-seq data. *BMC Bioinf* (2013) 14:7. doi: 10.1186/1471-2105-14-7
- Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-García W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* (2013) 4:2612. doi: 10.1038/ncomms3612
- Xiao H, Zhang J, Wang K, Song K, Zheng H, Yang J, et al. A cancer-specific qualitative method for estimating the proportion of tumor-infiltrating immune cells. *Front Immunol* (2021) 12:672031. doi: 10.3389/fimmu.2021.672031

34. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* (2016) 17(1):218. doi: 10.1186/s13059-016-1070-5
35. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Med* (2019) 11(1):34. doi: 10.1186/s13073-019-0638-6
36. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, et al. Timer2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res* (2020) 48(W1):W509–w14. doi: 10.1093/nar/gkaa407
37. Aran D, Hu Z, Butte AJ. Xcell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol* (2017) 18(1):220. doi: 10.1186/s13059-017-1349-1
38. Zhu M, Zeng Q, Fan T, Lei Y, Wang F, Zheng S, et al. Clinical significance and immunometabolism landscapes of a novel recurrence-associated lipid metabolism signature in early-stage lung adenocarcinoma: A comprehensive analysis. *Front Immunol* (2022) 13:783495. doi: 10.3389/fimmu.2022.783495
39. Ru B, Wong CN, Tong Y, Zhong JY, Zhong SSW, Wu WC, et al. Tisidb: an integrated repository portal for tumor-immune system interactions. *Bioinf (Oxford England)* (2019) 35(20):4200–2. doi: 10.1093/bioinformatics/btz210
40. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med* (2018) 24(10):1550–8. doi: 10.1038/s41591-018-0136-1
41. Narayanan S, Kawaguchi T, Yan L, Peng X, Qi Q, Takabe K. Cytolytic activity score to assess anticancer immunity in colorectal cancer. *Ann Surg Oncol* (2018) 25(8):2323–31. doi: 10.1245/s10434-018-6506-6
42. Cristescu R, Mogg R, Ayers M, Albright A, Murphy E, Yearley J, et al. Pan-tumor genomic biomarkers for Pd-1 checkpoint blockade-based immunotherapy. *Sci (New York NY)* (2018) 362(6411). doi: 10.1126/science.aar3593
43. Langfelder P, Horvath S. Wgcna: an R package for weighted correlation network analysis. *BMC Bioinf* (2008) 9:559. doi: 10.1186/1471-2105-9-559
44. Jeggari A, Marks DS, Larsson E. Mircode: A map of putative microRNA target sites in the long non-coding transcriptome. *Bioinf (Oxford England)* (2012) 28(15):2062–3. doi: 10.1093/bioinformatics/bts344
45. McGeary SE, Lin KS, Shi CY, Pham TM, Bisaria N, Kelley GM, et al. The biochemical basis of microRNA targeting efficacy. *Sci (New York NY)* (2019) 366(6472). doi: 10.1126/science.aav1741
46. Chen Y, Wang X. Mirdb: an online database for prediction of functional microRNA targets. *Nucleic Acids Res* (2020) 48(D1):D127–d31. doi: 10.1093/nar/gkz757
47. Huang HY, Lin YC, Li J, Huang KY, Shrestha S, Hong HC, et al. Mirtarbase 2020: updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res* (2020) 48(D1):D148–d54. doi: 10.1093/nar/gkz896
48. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of drug sensitivity in cancer (Gdsc): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* (2012) 41(D1):D955–D61. doi: 10.1093/nar/gks1111%JNucleicAcidsResearch
49. Rees MG, Seashore-Ludlow B, Cheah JH, Adams DJ, Price EV, Gill S, et al. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat Chem Biol* (2016) 12(2):109–16. doi: 10.1038/nchembio.1986
50. Seashore-Ludlow B, Rees MG, Cheah JH, Cokol M, Price EV, Coletti ME, et al. Harnessing connectivity in a large-scale small-molecule sensitivity dataset. *Cancer Discovery* (2015) 5(11):1210–23. doi: 10.1158/2159-8290.Cd-15-0235
51. Basu A, Bodycombe NE, Cheah JH, Price EV, Liu K, Schaefer GI, et al. An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. *Cell* (2013) 154(5):1151–61. doi: 10.1016/j.cell.2013.08.003
52. Yu C, Mannan AM, Yvone GM, Ross KN, Zhang Y-L, Marton MA, et al. High-throughput identification of genotype-specific cancer vulnerabilities in mixtures of barcoded tumor cell lines. *Nat Biotechnol* (2016) 34(4):419–23. doi: 10.1038/nbt.3460
53. Corsello SM, Nagari RT, Spangler RD, Rossen J, Kocak M, Bryan JG, et al. Discovering the anticancer potential of non-oncology drugs by systematic viability profiling. *Nat Cancer* (2020) 1(2):235–48. doi: 10.1038/s43018-019-0018-6
54. Maeser D, Gruener RF, Huang RS. Oncopredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. *Briefings Bioinf* (2021) 22(6). doi: 10.1093/bib/bbab260
55. Butler A, Hoffman P, Smibert P, Papalexis E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* (2018) 36(5):411–20. doi: 10.1038/nbt.4096
56. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with harmony. *Nat Methods* (2019) 16(12):1289–96. doi: 10.1038/s41592-019-0619-0
57. Alikhan R, Cohen AT, Combe S, Samama MM, Desjardins L, Eldor A, et al. Risk factors for venous thromboembolism in hospitalized patients with acute medical illness: analysis of the medenox study. *Arch Intern Med* (2004) 164(9):963–8. doi: 10.1001/archinte.164.9.963
58. Green KB, Silverstein RL. Hypercoagulability in cancer. *Hematology/oncol Clinics North America* (1996) 10(2):499–530. doi: 10.1016/s0889-8588(05)70349-x
59. Caine GJ, Stonelake PS, Lip GY, Kehoe ST. The hypercoagulable state of Malignancy: pathogenesis and current debate. *Neoplasia (New York NY)* (2002) 4(6):465–73. doi: 10.1038/sj.neo.7900263
60. Rees PA, Clouston HW, Duff S, Kirwan CC. Colorectal cancer and thrombosis. *Int J Colorectal Dis* (2018) 33(1):105–8. doi: 10.1007/s00384-017-2909-2
61. Wang Y, Hou K, Jin Y, Bao B, Tang S, Qi J, et al. Lung adenocarcinoma-specific three-integrin signature contributes to poor outcomes by metastasis and immune escape pathways. *J Trans Internal Med* (2021) 9(4):249–63. doi: 10.2478/jtim-2021-0046
62. Zou Y, Xie J, Zheng S, Liu W, Tang Y, Tian W, et al. Leveraging diverse cell-death patterns to predict the prognosis and drug sensitivity of triple-negative breast cancer patients after surgery. *Int J Surg (London England)* (2022) 107:106936. doi: 10.1016/j.ijso.2022.106936
63. Yang H, Chen W, Jiang G, Yang J, Wang W, Li H. Long non-coding rna ewsat1 contributes to the proliferation and invasion of glioma by sponging Mir-152-3p. *Oncol Lett* (2020) 20(2):1846–54. doi: 10.3892/ol.2020.11716
64. He X, Chen J, Zhou J, Mao A, Xu W, Zhu H, et al. Lncrna-ewsat1 promotes hepatocellular carcinoma metastasis via activation of the src-yap signaling axis. *FASEB J Off Publ Fed Am Soc Exp Biol* (2022) 36(12):e22663. doi: 10.1096/fj.202200825R
65. Fu X, Zhang L, Dan L, Wang K, Xu Y. Lncrna ewsat1 promotes ovarian cancer progression through targeting mir-330-5p expression. *Am J Trans Res* (2017) 9(9):4094–103.
66. Zhang R, Li JB, Yan XF, Jin K, Li WY, Xu J, et al. Increased ewsat1 expression promotes cell proliferation, invasion and epithelial-mesenchymal transition in colorectal cancer. *Eur Rev Med Pharmacol Sci* (2018) 22(20):6801–8. doi: 10.26355/eurrev_201810_16146
67. Liu J, Huang S, Liao X, Chen Z, Li L, Yu L, et al. Lncrna ewsat1 promotes colorectal cancer progression through sponging mir-326 to modulate fbxl20 expression. *Oncotargets Ther* (2021) 14:367–78. doi: 10.2147/ott.S272895
68. Li C, Zheng H, Hou W, Bao H, Xiong J, Che W, et al. Long non-coding rna linc00645 promotes Tgf-B-induced epithelial-mesenchymal transition by regulating Mir-205-3p-Zeb1 axis in glioma. *Cell Death Dis* (2019) 10(10):717. doi: 10.1038/s41419-019-1948-8
69. Peng WX, Liu F, Jiang JH, Yuan H, Zhang Z, Yang L, et al. N6-methyladenosine modified linc00901 promotes pancreatic cancer progression through Igf2bp2/Myc axis. *Genes Dis* (2023) 10(2):554–67. doi: 10.1016/j.gendis.2022.02.014
70. Ding X, Jia X, Wang C, Xu J, Gao SJ, Lu C. A Dhx9-Lncrna-Mdm2 interaction regulates cell invasion and angiogenesis of cervical cancer. *Cell Death Differ* (2019) 26(9):1750–65. doi: 10.1038/s41418-018-0242-0
71. Delvaeye M, Conway EM. Coagulation and innate immune responses: can we view them separately? *Blood* (2009) 114(12):2367–74. doi: 10.1182/blood-2009-05-199208
72. Esmon CT. Interactions between the innate immune and blood coagulation systems. *Trends Immunol* (2004) 25(10):536–42. doi: 10.1016/j.it.2004.08.003
73. Ryan TAJ, O'Neill LAJ. Innate immune signaling and immunothrombosis: new insights and therapeutic opportunities. *Eur J Immunol* (2022) 52(7):1024–34. doi: 10.1002/eji.202149410
74. Kobayashi H, Enomoto A, Woods SL, Burt AD, Takahashi M, Worthley DL. Cancer-associated fibroblasts in gastrointestinal cancer. *Nat Rev Gastroenterol Hepatol* (2019) 16(5):282–95. doi: 10.1038/s41575-019-0115-0
75. Chen YF, Yu ZL, Lv MY, Cai ZR, Zou YF, Lan P, et al. Cancer-associated fibroblasts impact the clinical outcome and treatment response in colorectal cancer via immune system modulation: A comprehensive genome-wide analysis. *Mol Med (Cambridge Mass)* (2021) 27(1):139. doi: 10.1186/s10020-021-00402-3
76. Chan JJ, Tay Y. Noncoding rna:Rna regulatory networks in cancer. *Int J Mol Sci* (2018) 19(5):1310. doi: 10.3390/ijms19051310
77. Kong X, Duan Y, Sang Y, Li Y, Zhang H, Liang Y, et al. Lncrna-cdc6 promotes breast cancer progression and function as Cerna to target Cdc6 by sponging microRNA-215. *J Cell Physiol* (2019) 234(6):9105–17. doi: 10.1002/jcp.27587
78. Cao HL, Liu ZJ, Huang PL, Yue YL, Xi JN. Lncrna-Rmrp promotes proliferation, migration and invasion of bladder cancer via Mir-206. *Eur Rev Med Pharmacol Sci* (2019) 23(3):1012–21. doi: 10.26355/eurrev_201902_16988
79. Jiang L, Zhang Y, Su P, Ma Z, Ye X, Kang W, et al. Long non-coding Rna Hnf1a-As1 induces 5-Fu resistance of gastric cancer through Mir-30b-5p/Eif5a2 pathway. *Trans Oncol* (2022) 18:101351. doi: 10.1016/j.tranon.2022.101351
80. Sun S, Li W, Ma X, Luan H. Long noncoding rna linc00265 promotes glycolysis and lactate production of colorectal cancer through regulating of Mir-216b-5p/Trim44 axis. *Digestion* (2020) 101(4):391–400. doi: 10.1159/000500195
81. Bai J, Xu J, Zhao J, Zhang R. Lncrna Nbr2 suppresses migration and invasion of colorectal cancer cells by downregulating Mirna-21. *Hum Cell* (2020) 33(1):98–103. doi: 10.1007/s13577-019-00265-1
82. Huang W, Zhang H, Tian Y, Li Y, Li J, Zhong X, et al. Lncrna Snhg11 enhances bevacizumab resistance in colorectal cancer by mediating Mir-1207-5p/Abcc1 axis. *Anti-cancer Drugs* (2022) 33(6):575–86. doi: 10.1097/cad.0000000000001289