Biomolecules and Biomedicine

ISSN: 2831-0896 (Print) | ISSN: 2831-090X (Online)

Journal Impact Factor® (2024): 2.2

CiteScore® (2024): 5.2

www.biomolbiomed.com | blog.biomolbiomed.com

RESEARCH ARTICLE

Wang et al: TMGs in CRC prognosis

Role of telomere maintenance genes as a predictive biomarker for colorectal cancer immunotherapy response and prognosis

Zhikai Wang[1#], Chunyan Zhao[2#], Yifen Huang[3*], Chong Li[2*]

[1]Department of Gastrointestinal Surgery, Henan Provincial People's Hospital, Zhengzhou, China

[2]Department of Oncology, The Affiliated Dazu's Hospital of Chongqing Medical University, Chongqing, China

[3]Outpatient Department, The Affiliated Dazu's Hospital of Chongqing Medical University, Chongqing, China

*Correspondence to **Yifen Huan**: 3516882473@qq.com and **Chong Li**: 152389@hospital.cqmu.edu.cn

[#]Equally contributed to this work: **Zhikai Wang** and **Chunyan Zhao**

DOI: https://doi.org/10.17305/bb.2025.12053

ABSTRACT

Colorectal cancer (CRC) represents a significant global health challenge. Although telomere maintenance plays a crucial role in tumorigenesis, the prognostic value and immunotherapeutic relevance of telomere maintenance genes (TMGs) in CRC remain poorly understood. In this study, relevant data were retrieved from The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO) databases. TMG scores were calculated using the single-sample gene set enrichment analysis (ssGSEA) method, and TMGs associated with prognosis were subsequently identified. TCGA-CRC samples were classified into subtypes via consensus clustering (ConsensusClusterPlus). A risk prediction model was then constructed using univariate and Lasso Cox regression analyses. Survival analysis was performed using Kaplan–Meier curves generated with the survival package. Key genes were validated in vitro using cellular models. Immune cell infiltration was evaluated through ssGSEA, TIMER, and MCP-Counter tools, and chemotherapy responses were predicted using the pRRophetic package. From 28 prognosis-related TMGs, two distinct CRC subtypes were established, with subtype C1 demonstrating more favorable clinical outcomes. Additionally, a risk model incorporating seven TMG-related genes (*CDC25C, CXCL1, RTL8C, FABP4, ITLN1, MUC12*, and *ERI1*) was developed for CRC prognosis. Differential mRNA expression levels of these genes were confirmed between CRC cell lines and normal control cells. Furthermore, silencing *MUC12* suppressed CRC cell migration and invasion in vitro. Importantly, CRC patients classified as low-risk exhibited superior responses to immunotherapy, whereas high-risk patients showed increased sensitivity to conventional anti-cancer treatments. This study represents the first systematic evaluation of TMGs in CRC prognosis and immunotherapy, providing novel insights that could inform personalized therapeutic strategies.

Keywords: Immunotherapy; Risk model; Prognostic signature; Telomere maintenance gene; Colorectal cancer.

INTRODUCTION

Colorectal cancer (CRC) ranks among the most prevalent malignancies of the digestive system [1, 2], with statistics suggesting it will affect ~2.5 million individuals by 2035 [3]. Globally, CRC represents 9.6% of all cancer cases, making it the third most frequently detected cancer [4]. It has been reported that CRC has the second highest death rate (9.3%), resulting in an annual death number of 900,000 all over the world [4, 5]. As a result of its asymptomatic nature in early stages, timely diagnosis and treatment remain difficult, which requires more advanced screening methods [6, 7]. Though studies identified transcription factors (e.g., Nrf2) and ferroptosis-related genes (GSH, GPX4, P53) as potential therapeutic targets for CRC [8], a lack of reliable biomarkers continues to lower the accuracy of prognostic prediction [9]. Thus, discovering biomarkers specific to CRC and developing prognostic models are essential for the early detection and prognosis of CRC [10].

Telomeres, composed of protective proteins and TTAGGG repeats [11, 12], are regarded as nucleoprotein complexes locating at the termini of human chromosomes. Study has revealed crucial roles of telomeres in protecting chromosomal and genome stabilities[13]. However, under certain cell division or disease states, the length of telomeres shortens gradually [14, 15]. In order to maintain a normal length, telomerase adds TTAGGG DNA sequence to the location of the end of chromosomes[16, 17]. Studies have increasingly shown the relevance between the shortening of telomere length and the incidence of various diseases such as tumors [13, 18, 19]. It has been observed that TMGs influence cancer occurrence through regulating the mutations in telomerase reverse transcriptase promoter (TERT), highlighting the potential of TMGs to act as cancer biomarkers[20, 21]. In cancers, two telomere elongation mechanisms [22], namely, telomerase activation[23] and alternative lengthening of telomeres (ALT)[24], have been found. Telomerase overexpression is frequently observed in many tumor cells [25, 26], for instance, thyroid cancer abnormal activation of ALT detected in telomerase-negative tumors [27]. Despite these results, in order to improve the current therapeutic strategies, the prognostic value of TMGs in CRC awaits to be clarified.

Hence, the current study was designed to systematically evaluate the expression patterns, molecular subtypes, and prognostic significance of TMGs in CRC, along with their

immune microenvironment interactions and treatment response correlations. Utilizing the transcriptomic data of CRC in public databases, the ssGSEA method was employed to compute the TMG expression scores, and key survival-related genes were identified. Next, molecular subtypes were classified by the prognostically correlated TMGs applying consensus clustering method, followed by performing differential expression analysis and functional enrichment analysis to analyze the potential biological mechanisms. Next, a RiskScore model was created to evaluate the performance of the TMGs in evaluating the immune cell infiltration, immunotherapy response and chemotherapy sensitivity, and survival for CRC patients. Finally, *in vitro* experiments were carried out for the purpose of verifying the expressions and biological functions of the key genes. Collectively, our discoveries revealed the essential role of TMGs in immune regulation and CRC development, offering novel insights for personalized treatment and risk assessment.

MATERIAL AND METHODS

Acquisition and preprocessing of data

CRC data including clinical information (both colon adenocarcinoma and rectum adenocarcinoma), copy number variations (CNVs), and somatic mutations were retrieved from TCGA database (https://portal.gdc.cancer.gov/). 589 tumor samples with survival time longer than 0 day were retained after filtering out samples lacking survival. The RNA-seq expression profiles were converted to TPM format and log2-transformed. Additionally, we downloaded GSE17537 microarray data from GEO (https://www.ncbi.nlm.nih.gov/geo) and transformed the probes to symbols based on the annotation file. 55 tumor samples in GSE17537 were retained after excluding those without clinical follow-up data or survival time. Finally, 2093 TMGs were extracted from a past study [28].

Identification of CRC-related TMGs and analysis of their mutations and CNVs

Using ssGSEA method, the TMG scores for the TCGA-CRC dataset were computed. DEGs between tumor and adjacent tissues were identified and intersected with TMG signatures. Prognosis-correlated TMGs were then selected through univariate Cox regression analysis. CNV and mutation data from the TCGA database were collected. Briefly, MuTect2

[29] was employed to analyze the mutational landscape of TMG in primary CRC samples from the TCGA-CRC cohort, followed by visualizing the mutational status into a waterfall plot with the maftools R package[30]. Then, the CNV status of the TMGs in primary CRC samples were detected by ADTEx[31].

Molecular clustering

Using the ConsensusClusterPlus package [32], we performed consensus clustering on the tumor samples with the parameters of hierarchical clustering (clusterAlg = "hc") and Pearson correlation distance (distance = "pearson"). The analysis involved 500 iterations with a resampling rate of 80%. The optimal number of clusters (k) was identified based on stabilized cumulative distribution function (CDF) curve, minimal incremental gains in the delta area plots, and high intra-cluster consensus and clear inter-cluster separation in consensus matrices. Finally, the clinical features (M.stage, N.stage, T.stage, stage, status, age and gender) and prognosis in different molecular subtypes were systematically assessed.

Enrichment analysis

DEGs between the C1 and C2 were identified by the limma package (FDR < 0.05 & |log2FC| >log2(1.5)) [33] to select common genes for enrichment analysis. The TCGA-CRC cohort was subdivided using the DEGs into two molecular subtypes. Next, the clusterProfiler R package[34] was used for conducting GO and KEGG enrichment analysis [35].

Establishment of a risk model

Prognosis-correlated DEGs were screened through univariate Cox regression analysis, followed by refining the risk model with Lasso Cox regression in the glmnet package[36] and 10-fold cross-validation. A RiskScore model was formulated applying stepwise multivariate Cox regression analysis as follow:

$$RiskScore = \sum \beta i \times Expi$$

Where, βi represents the coefficient of a gene in the Cox regression model, and Expi represents the gene expression. The samples were assigned by the median RiskScore as the threshold into low- and high-risk groups, followed by analyzing the survival difference

between the two groups using KM curves with the survival package[37]. The prognostic classification by the RiskScore model was validated using receiver operating characteristic (ROC) analysis in timeROC R package[38] and principal component analysis (PCA). Furthermore, the prognostic differences in the two risk groups across gender, age, and stage were compared by calculating the RiskScores for all the patients.

Independent clinical and pathological factors (stage, M.stage, N.stage, T.stage, status, age and gender) for CRC prognosis were selected by univariate and multivariate Cox regression analyses along with the RiskScore model. To predict 1-, 3-, and 5-year survival for CRC patients, a prognostic nomogram was developed using the rms package[39]. Finally, the clinical reliability of the nomogram was tested by calibration curve analysis and decision curve analysis (DCA).

Tumor microenvironment (TME) differences across the risk groups

Immune cell infiltration in each risk groups was comprehensively examined using the ssGSEA algorithm, the MCPcounter package[40], and the TIMER online tool (http://cistrome.org/TIMER). Specifically, the TIMER offer six main analysis modules to enable the correlation analysis between a range of factors and immune infiltration.

Culture and transfection of cells

Human colon adenocarcinoma cells SW1116 (BNCC100262) and human normal colon epithelial cells NCM460 (BNCC339288) ordered from BeiNa Culture Collection (Beijing, China) were cultured in RPMI Medium 1640 (31800, Solarbio Lifesciences, Beijing, China) or Leibovitz's L-15Cell Culture Medium (LA9510, Solarbio Lifesciences) supplemented with 10% FBS (S9020, Solarbio Lifesciences). The temperature was constantly maintained at 37°C to ensure the optimal cell growth condition. Additionally, NCM460 culture was performed in 5%CO_2.

The siRNA was synthesized by GenePharma (Suzhou, China) and transfected into MUC12 cells applying Lipofectamine 2000 (11668027, Invitrogen, Carlsbad, CA), and the negative control with scramble target sequence (target sequence: 5'-ACCAGTATTGGAGGTAATACAAC-3') was purchased [41].

Wound healing assay

A controlled artificial wound was made on the cell monolayers with a 10 μL pipette tip after culturing the transfected CRC cells SW1116 (1×10^5 cells/well) to complete confluence in a 6-well plate. Next, the cells were cultured in serum-free medium at 37°C. The cells were photographed at 0 hour (h) and 48 h under an Eclipse Ts2R microscope (Nikon, Tokyo, Japan).

Transwell assay

Matrigel (50 μL, Corning, Inc.) was used to pre-coat the transwell chambers (pore size: 8 μm, 3422, Corning, Inc., Corning, NY), which were placed on 24-well plates. Next, the transfected SW1116 cells in serum-free medium (200μL) were planted into the upper chamber at 1×10^5 cells/well, whereas the lower transwell chamber contained 700 μL complete medium. After cell culture for 48 h, the cells were fixed by 4% paraformaldehyde (P1110, Solarbio Lifesciences) for 30 minutes (min), followed by using 0.1% crystal violet solution (G1063, Solarbio Lifesciences) for cell staining another 15 min. Finally, the cells were observed and quantified under the microscope used in the previous assay.

QRT-PCR analysis

Total RNAs from NCM460 and SW1116 cells were extracted using the TriZol kit (15596026CN, Invitrogen), followed by reverse transcription into cDNA with a first-strand cDNA synthesis kit (1708890, Bio-Rad Laboratories, Hercules, CA). Using a CFX384 qPCR System (1855484, Bio-Rad Laboratories), qCR amplification was performed with the use of SYBR Green Supermix (1708880, Bio-Rad Laboratories) under the cycling conditions starting with initial denaturation at 95°C for 2 min, followed by 40 cycles at 95°C for 15 seconds (s), at 60°C for 30 s, and at 72°C for 30 s. With GAPDH as a normalization control, relative mRNA expression was calculated by the method of $2^{-\Delta\Delta C}$ [42]. See **Table. S1** for the primers.

Immunotherapy and drug sensitivity

Immunotherapy responses were predicted by standardizing the transcriptome data applying TIDE (http://tide.dfci.harvard.edu/) to calculate the TIDE scores, with higher TIDE

scores showing greater possibility of immune escape and less immunotherapy benefit. Next, chemotherapy sensitivity in the TCGA-CRC dataset and the differences of patients' responses were analyzed and compared by the pRRophetic software package[43]. Patients' sensitivity in different risk groups to chemotherapy agents was evaluated with IC50.

Statistical analysis

All the analyses were conducted in R language (version 3.6.0). Data distribution normality was verified by Shapiro-Wilk test prior to the test of variance. For comparisons between two independent groups of continuous variables, the Wilcoxon rank-sum test was employed. Kruskal-Wallis test was used to examine the differences in continuous variables among three groups. Chi-square test was used to assess the disparities in categorical variables across different groups. Additionally, the log-rank test was adopted to compare the differences of survival time among patients in different groups. $P < 0.05$ denoted a statistical significance. For *in vitro* cellular experiments, the difference between normal and experimental group cells and the experimental data were analyzed using Student's t-test and GraphPad Prism 8.0 software. Data were expressed by the mean ± standard deviation (SD). SangerBox (http://sangerbox.com/) offered this study analytical assistance [44].

RESULTS

Genomic landscape of the TMGs in CRC

Using ssGSEA analysis, we first computed the TMG scores in the TCGA-CRC dataset. It was found that the tumor tissues demonstrated significantly elevated TMG scores than the adjacent non-tumor tissues (**Figure. 1A**). Further analysis filtered 317 DEGs between the two types of tissues (**Figure. 1B**), among which 28 TMGs showing significant prognostic relevance in CRC were determined through univariate Cox regression analysis (**Figure. 1C**). Furthermore, analysis on the mutational status and CNVs on these 28 TMGs in the tumor samples revealed that only 23.69% of the samples carried gene mutations linked to telomere maintenance (**Figure. 1D**), with *SNAI1* and *RBL1* having a comparatively higher frequency of copy number amplification (**Figure. 1E**).

Identification of CRC molecular subtypes based on TMGs

Here, we use consistency clustering to cluster the TCGA-CRC samples with a combination of CDF curves, Delta Area plots and consistency matrix heatmaps. In this study, we use consistency clustering to analyze the data in clusters. The stability and reasonableness of different clustering numbers (k) are evaluated by combining CDF curves, Delta Area plots, and heatmaps of the consistency matrix. First, the CDF curves showed that as the number of clusters increased from k=2 to k=3, the CDF curves shifted to the upper right corner, and the clustering consistency improved, but the enhancement was significantly less than after k=3 (**Figure. 2A**). **Figure. 2B** further shows that the increment in CDF area is the largest between k=2 and k=3, but the increment flattens out at k≥4, suggesting that there is limited gain from continuing to increase the number of clusters. In addition, as shown in **Figure. 2C**, with k=2, the heatmap shows high cluster consistency of samples within clusters, clear separation between clusters, and a more robust clustering structure. Therefore, we finally determined k = 2 as the optimal number of clusters and divided the samples into two clusters (C1 and C2). Overall survival was significantly better in C1 than in C2 (**Figure. 2D,** $P = 0.0015$). Analysis of the clinical features revealed that the two subtypes differed in M.stage and status. Comparison on TMG expression profiles between C1 and C2 showed higher expressions of multiple genes in C1, including *RBL1*, *CHEK1*, *BRCA1*, *HMMR*, *KPNA2*, *CCNA2*, *NCAPG*, *TKT*, *TRAP1*, *ORC1*, *CDC45*, *CCT2*, and *CCNB1*(**Figure. 2E**). These data verified a robust classification of the CRC samples into two molecular subtypes with significant survival differences and heterogeneity of clinical features, providing a solid basis for subsequent study of molecular subtyping.

Enrichment analysis results of the DEGs

Differential expression analysis using limma package[45] identified 538 DEGs between C1 and C2 (282 upregulated in C1, 256 upregulated in C2). GO and KEGG enrichment analysis showed that C1-associated genes were mainly enriched in proliferation-related pathways including DNA replication and cell cycle (**Figure. 3A-D**), while C2-associated genes were mainly enriched in pathways related to cancer metastasis and invasion such as focal adhesion, extracellular matrix organization, extracellular matrix (ECM)-receptor

interaction (**Figure. 3E-H**). These results demonstrated significant differences between the two subtypes in terms of potential therapeutic response and biological behaviors in CRC.

Development of a risk model based using the seven TMGs

Through univariate Cox analysis, we identified 101 prognostic DEGs ($P < 0.05$), which were finally refined to seven key genes (*CDC25C, CXCL1, RTL8C, FABP4, ITLN1, MUC12,* and *ERI1*) by Lasso Cox regression analysis with 10-fold cross-validation (Figure. 4A) and stepwise multivariate Cox regression analysis (**Figure. 4B**). The formula of the RiskScore model was as follow:

$$\text{RiskScore} = (0.155 * \text{RTL8C}) + (0.080 * \text{FABP4}) + (-0.062 * \text{ITLN1}) + (-0.100 * \text{CXCL1}) + (-0.122 * \text{MUC12}) + (-0.282 * \text{ERI1}) + (-0.285 * \text{CDC25C})$$

Patients were allocated by the median value of RiskScore into low-risk and high-risk groups. KM survival curve demonstrated that the two risk groups in the TCGA-CRC training cohort differed significantly in patients' survival, with those having a higher RiskScore showing shorter overall survival (OS) (**Figure. 4C**). The timeROC package[38] was employed in ROC analysis for further validating the prognosis classification of the RiskScore. The AUC values in the training dataset for 1-, 3-, and 5-year survival were 0.63, 0.68, and 0.72, respectively, which suggested a highly accurate survival evaluation by the model (**Figure. 4C**). Furthermore, PCA results also showed a distinct separation between the two risk groups in the TCGA-CRC cohort (**Figure. 4D**), further supporting the performance of the RiskScore in identifying CRC patients with different risks.

The robustness of the RiskScore was confirmed using the validation dataset of GSE17537 (**Figure. 4E-F**). Consistently, the value of the RiskScore was negatively correlated with survival outcomes of the samples in this dataset. Comparison on the performance of the low- and high- risk groups across different clinical factor subgroups revealed distinct differences ($P < 0.05$) between the two risk groups, independent of stage classification (I + II or III + IV), gender (male or female), or age (> 67 or ≤ 67). This reflected an independent risk classification, which was less likely to be influenced by other clinical factors (**Figure. 4G-I**).

Validation of the TMG-related risk model

Analysis of the TCGA cohort revealed close associations between the risk groups and pathological staging, with high-risk group having more patients with C2 subtype and metastasis cases. This finding was consistent with our previous research, which showed that the C2 cluster was enriched in metastasis-related pathways. Further analysis revealed that the RiskScore and more advanced clinical staging (T.stage, N.stage, and stage) were positively correlated. **Figure. 5A** and **Figure. 5B** show the correlation between clinical features and RiskScore and the violin plot, respectively. Both univariate and multivariate Cox regression analysis identified M.stage, stage, the RiskScore, and age as significant factors for CRC prognosis (**Figure. 5C-D**). The nomogram was developed combining other clinical pathological characteristics and the RiskScore for quantifying the survival and risk for CRC patients (**Figure. 5E**). The findings showed the greatest influence of the RiskScore on the survival evaluation. Calibration curves exhibited that the 1-, 3-, and 5-year prediction curves closely aligned with the standard curve (**Figure. 5F**), manifesting a strong prediction performance of the nomogram. Additionally, DCA also demonstrated a better clinical effectiveness and reliability of the nomogram (**Figure. 5G**).

In vitro verification of the key genes for CRC prognosis

The relative expressions of the seven genes (*CDC25C*, *CXCL1*, *RTL8C*, *FABP4*, *ITLN1*, *MUC12*, and *ERI1*) in SW1116 and NCM460 cells were measured. It was found that the expressions of *ITLN1* and *ERI1* were notably downregulated in SW1116 cells, while the mRNA expressions of *CDC25C*, *CXCL1*, *RTL8C*, *FABP4*, and *MUC12* were significantly higher in SW1116 cells than in control NCM460 cells (**Figure. 6A-G**, $P < 0.05$).

Previous study found the potential of *MUC12* as a molecular marker for the prognosis of CRC [46, 47]. Therefore, this study performed wound healing and transwell assays to evaluate the potential effects of *MUC12* knockdown on CRC cells. As shown in **Figure. 6H-I**, *MUC12* knockdown notably suppressed the migration and invasion abilities of SW1116 cells ($P < 0.01$). This result was consistent with the cancer-promoting role of *MUC12*, which further supported the clinical significance of the RiskScore model developed based on TMGs.

Differences in the TME between CRC patients with different risks

The ssGSEA showed that the infiltration of Type 17 T helper cell, Neutrophil, Activated B cell, Type 2 T helper cell, Activated CD4 T cell, Activated CD8 T cell, which typically exhibit killing effects on tumor cells [48], was lower in the high-risk group than in the low-risk group and this may contribute to the protection of tumor cells. Compared to the low-risk group, high-expressed Regulatory T cell in the high-risk group could promote tumor development (**Figure. 7A**). The TIMER analysis demonstrated that CD8_Tcell and B_cell was low-expressed in the high-risk group (**Figure. 7B**). The MCPcounter analysis (**Figure. 7C**) unveiled significant infiltration differences of some cell types between the two groups. Specifically, high-risk group had significantly lower infiltration of NK cells, cytotoxic lymphocytes, T cells, Neutrophil, B lineage. This implied that the absence of immune effector cells may shape a "cold-immune" TME of the high-risk patients, which may explain their poorer prognosis.

Immunotherapy and drug sensitivity analysis for CRC patients in different risk groups

TIDE analysis revealed significantly lower TIDE scores in the low-risk group (**Figure. 8A**), suggesting more active immune response and less immune evasion possibility in those patients. Further analysis showed that the low-risk group had a significantly higher expression level of the immune checkpoint inhibitor CD274 (PD-L1) than the high-risk group (**Figure. 8B**), indicating a better response of low-risk patients to immune checkpoint blockade therapy. Based on the ssGSEA algorithm and an established gene signature [49], the responsiveness to treatments including anticancer immunotherapy and chemotherapy was analyzed. It was observed that low-risk CRC patients had a higher reactivity across multiple treatment-correlated gene sets (**Figure. 8C-E**), while high-risk CRC patients had higher sensitivities to common anti-cancer drugs, for instance, Phenformin, MG.132, Cyclopamine, and Sorafenib (**Figure. 8F**). These findings highlighted that patients in different risk groups responded differently to the therapeutic strategies, with low-risk patients benefiting more from immunotherapy and high-risk patients benefiting more from conventional anti-tumor drug treatments.

DISCUSSION

CRC diagnosis remains challenging due to its asymptomatic symptoms in early stages, which also results in a poor prognosis. This highlights the need for effective prognostic biomarkers to reduce the mortality of CRC patients [7]. Telomere length in lymphocytes is closely involved in tumor development, and telomere shortening is considered as a prognostic marker for CRC [50]. While genes like *RCN3* have emerged as prognostic markers, their specific mechanism still requires in-depth investigation [51]. Hence, this study developed a novel TMG-based risk model for the prognostic assessment in CRC to contribute to this filed.

Clustering analysis identified two distinct CRC subtypes (C1, C2) based on TCGA-CRC samples, with C2 subtype having higher copy number and expression of *SNAI1* and poorer prognosis. *SNAI1* has been found to play a pivotal part in maintaining telomere integrity[52] and its absence promotes telomerase activity in mesenchymal stem cells, highlighting the potential of *SNAI1* as a crucial TMG in the process[52]. *SNAI1* could also regulate epithelial-mesenchymal transition (EMT) [53] during which epithelial cells lose connections and polarity but acquire mesenchymal properties and invasive ability [54]. Such phenotypic changes during EMT contribute to tumorigenesis. The expression level and function of *SNAI1* have been widely studied in many types of cancers including CRC. *SNAI1* drives stem cell properties, metabolic alterations, cancer invasion and chemoresistance in epithelial ovarian cancer,[55], promotes metastasis in breast tumors [56]. Moreover, high-expressed *SNAI1* is considered as a clinical biomarker in gastric cancer,[57]. In CRC, intestinal epithelial *SNAI1* facilitates CRC development through EMT and Wnt/β-catenin signaling pathway[58]. Furthermore, study on both CRC patients and *in vitro* experiment showed that the expression of *SNAI1* is predictive of the patients' response to drug treatment [59]. This study found that *SNAI1* had a higher CNV, which was consistent with previous finding that CNVs occur frequently in the regions containing genes with crucial functions in CRC and therefore could be regarded as biomarkers for cancer detection[60]. This study proposed SNAI1 as a promising candidate for early CRC detection.

Recent studies highlight the multifaceted roles of the 7 identified TMGs (*CDC25C*, *CXCL1*, *RTL8C*, *FABP4*, *ITLN1*, *MUC12*, and *ERI1*) in carcinogenesis. For instance, *CDC25C* plays a critical part in regulating the G2/M phase of the cell cycle, and its

expression changes are implicated in the growth of cancer [61]. *CXCL1*, a CXC chemokine subfamily member, demonstrates clinical significance in many cancer types[62]. *FABP4*, abundant in adipocytes, shows an upregulated expression in multiple solid tumors and this indicates a poor prognosis[63]. *RTL8C* has the potential to be considered as a promising pan-cancer biomarker [64]. *ERI1* expressed at a high level is linked to an improved OS of CRC [65]. *ITLN*, primarily generated by stromal vascular fraction cells, is crucially involved in cancer growth [66]. *MUC12* is a type of transmembrane mucin typically expressed in the normal colon and less expressed in the pancreas. Study reported a lower mRNA level of *MUC12* in certain CRC tissues in comparison to that in normal colonic tissues [67, 68]. Notably, *MUC12* presented functional complexity in this study. On one hand, the results of multivariate Cox regression analysis revealed that *MUC12* behaved as an independent protective prognostic factor, with higher expression associated with longer overall survival. On the other hand, *in vitro* experiments showed that downregulated *MUC12* significantly suppressed the migration and invasion of CRC cells, suggesting that it may have a metastasis-promoting effect. The difference between "statistically protective" and "functionally pro-carcinogenic" potentially indicated a dual role of *MUC12* in different pathological stages or microenvironmental contexts. In some early stages, *MUC12* may play a protective role by maintaining the epithelial barrier function, while its overexpression may be involved in EMT and microenvironmental remodeling to promote tumor cell metastasis and invasion during tumor progression [69, 70]. This suggested that the specific role of *MUC12* still needs to be elucidated with more *in vivo* mechanistic studies.

For cancer pathogenesis, the TME is one of the essential factors whose compositional changes can indicate patients' responses to immunotherapy [71, 72]. In the current study, high-risk CRC group had a worse prognostic outcome, which was consistent with its enrichment in the C2 cluster. As we found that the enriched pathways in C2 were largely linked to cancer metastasis, this indicated that a worse prognosis of high-risk patients may be caused by metastasis. High-risk patients also showed suppressed abilities of immune system defense, resulting in an upregulated expression of immune cells, however, low-risk patients demonstrated a stronger immune response. The high-risk group also showed higher infiltration of endothelial cells and fibroblasts. Endothelial cells is a primary cell type that

plays a crucial role in angiogenesis in cancer tumor [59], which is a vital process that provides essential oxygen and nutrients for tumors [73]. Fibroblasts are also a type of mesenchymal cell that participates in tissue homeostasis and disease processes[74]. Varying infiltration of different immune cell types within the TME may contribute to the distinction between the two risk groups. The synergistic effects between fibroblasts and endothelial cells could cause tumor metastasis and diffusion, resulting a worse prognosis to high-risk CRC patients. Notably, the high-risk patients had a "suppressed" immune system and exhibited greater sensitivity to conventional anti-tumor agents. This also suggested that in patients with significant immunosuppression, small molecule targeted agents therapy could be prioritized over the use of immunotherapy alone.

This study had several limitations to be acknowledged. Firstly, while we utilized clinical information and large-scale RNA-seq data from public databases, potential biases may arise from inter-sample heterogeneity, differences in sequencing platforms, and incomplete clinical annotations, which could all affect the generalizability of the current model. Future multicenter prospective studies with larger independent cohorts are needed to validate and improve the clinical applicability and robustness of our risk model. Secondly, although we revealed significant differences in immune cell infiltration and immune checkpoint expressions between the two risk groups, the precise molecular mechanisms through which TMGs modulated the TME or immune evasion remained unclear. Subsequent investigations should integrate single-cell transcriptomics, multi-omics approaches, and functional experiments to systematically elucidate the immunoregulatory roles of the key TMGs and their potential as targets for combination immunotherapy. Finally, our *in vitro* validation primarily focused on expression profiling and limited functional characterization, therefore, comprehensive studies involving knockout or overexpression of other key genes and animal models are required to strengthen the mechanistic evidence to support our findings.

CONCLUSION

This study was the first to systematically identify TMGs closely linked to the prognosis of CRC at the whole genome level and to construct a CRC prognostic risk model consisting

of seven key TMGs (*CDC25C*, *CXCL1*, *RTL8C*, *FABP4*, *ITLN1*, *MUC12*, and *ERI1*). The model not only had a strong prediction ability in multiple independent cohorts, but also could effective identify immune microenvironment differences and drug sensitivity of CRC patients. Our findings supported that TMGs influenced the clinical outcomes of patients with CRC by modulating tumor immune escape mechanisms. Combined with *in vitro* experiments, the expressions of the key genes were found to be closely related to the invasive ability of CRC cells, further enhancing the biological explanatory power of the model. This study innovatively combined the telomere maintenance mechanism with the immunotherapy potential, providing candidate targets and a novel theoretical basis for the management and development of targeted therapy for CRC.

Conflicts of interest: Authors declare no conflicts of interest.

Funding: The study was supported by Chongqing Natural Science Foundation General Project (CSTB2023NSCQ-MSX0502) and Chongqing Dazu District Science and Technology Development Project (DZKJ2022JSYJ-KWXM1003).

Data availability: The datasets generated and/or analyzed during the current study are available in the [GSE17537] repository, [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE17537].

Submitted: 14 January, 2025

Accepted: 22 May, 2025

Published online: 02 July, 2025

REFERENCES

1. Li JX, Han T, Wang X, Wang YC, Chen X, Chen WS, et al. Identification of prognostic immune-related lncRNA signature predicting the overall survival for colorectal cancer. Scientific Reports. 2023;13(1):1333.
2. Ruan Y, Lu G, Yu Y, Luo Y, Wu H, Shen Y, et al. PF-04449913 Inhibits Proliferation and Metastasis of Colorectal Cancer Cells by
Down-regulating MMP9 Expression through the ERK/p65 Pathway. Current Molecular Pharmacology. 2023;17.

3. Sharma P, Bora K, Kasugai K, Balabantaray B-K. Two Stage Classification with CNN for Colorectal Cancer Detection. Oncologie. 2020;22(3):129--45.
4. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2024;74(3):229-63.
5. Shi X, Wang Y, Li C, Fu W, Zhang X, Gong A. Knockdown of RCN1 contributes to the apoptosis of colorectal cancer via regulating IP3R1. Biocell. 2024;48(5):835-45.
6. Malki A, ElRuz RA, Gupta I, Allouch A, Vranic S, Al Moustafa A-E. Molecular Mechanisms of Colon Cancer Progression and Metastasis: Recent Insights and Advancements. International Journal of Molecular Sciences. 2021;22(1):130.
7. Dekker E, Tanis PJ, Vleugels JL, Kasi PM, Wallace MB. Colorectal cancer. The Lancet Diabetes Endocrinology. 2019;394(10207):1467-80.
8. Wang Y, Zhang Z, Sun W, Zhang J, Xu Q, Zhou X, et al. Ferroptosis in colorectal cancer: Potential mechanisms and effective therapeutic targets. Biomedicine & Pharmacotherapy. 2022;153:113524.
9. Duan L, Yang W, Wang X, Zhou W, Zhang Y, Liu J, et al. Advances in prognostic markers for colorectal cancer*. Expert Review of Molecular Diagnostics. 2019;19(4):313-24.
10. Cherri S, Oneda E, Zanotti L, Zaniboni A. Optimizing the first-line treatment for metastatic colorectal cancer. Frontiers in Oncology. 2023;13:1246716.
11. Makarov VL, Hirose Y, Langmore JP. Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. Cellular and Molecular Life Sciences. 1997;88(5):657-66.
12. Brabletz T, Kalluri R, Nieto MA, Weinberg RA. EMT in cancer. Nature Reviews Cancer. 2018;18(2):128-34.
13. Nassour J, Schmidt TT, Karlseder J. Telomeres and cancer: resolving the paradox. Annual Review of Cancer Biology. 2021;5:59-77.
14. Alder JK, Armanios M. Telomere-mediated lung disease. Physiological Reviews. 2022.
15. Brandt M, Dörschmann H, Khraisat Sa, Knopp T, Ringen J, Kalinovic S, et al. Telomere shortening in hypertensive heart disease depends on oxidative DNA damage and predicts impaired recovery of cardiac function in heart failure. Hypertension. 2022;79(10):2173-84.
16. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, et al. Specific association of human telomerase activity with immortal cells and cancer. Science. 1994;266(5193):2011-5.
17. Wang F, Tang ML, Zeng ZX, Wu RY, Xue Y, Hao YH, et al. Telomere-and telomerase-interacting protein that unfolds telomere G-quadruplex and promotes telomere extension in mammalian cells. Proceedings of the National Academy of Sciences. 2012;109(50):20413-8.
18. in der Stroth L, Tharehalli U, Guenes C, Lechel A. Telomeres and telomerase in the development of liver cancer. Cancers. 2020;12(8):2048.
19. Savage SA. Telomere length and cancer risk: finding Goldilocks. Biogerontology. 2024;25(2):265-78.
20. Xiao Y, Xu D, Jiang C, Huili Y, Nie S, Zhu H, et al. Telomere maintenance-related genes are important for survival prediction and subtype identification in bladder cancer. Frontiers in Genetics. 2023;13.
21. Günes C, Wezel F, Southgate J, Bolenz C. Implications of TERT promoter mutations and

telomerase activity in urothelial carcinogenesis. Nature Reviews Urology. 2018;15(6):386-93.
22. Liu X, Wang J, Su D, Wang Q, Li M, Zuo Z, et al. Development and validation of a glioma prognostic model based on telomere-related genes and immune infiltration analysis. Translational cancer research. 2024;13(7):3182-99.
23. Shay JW, Wright WE. Role of telomeres and telomerase in cancer. Seminars in Cancer Biology. 2011;21(6):349-53.
24. Bryan TM, Englezou A, Dalla-Pozza L, Dunham MA, Reddel RR. Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. Nature Medicine. 1997;3(11):1271-4.
25. Heidenreich B, Kumar R. TERT promoter mutations in telomere biology. Mutation Research/Reviews in Mutation Research. 2017;771:15-31.
26. Yan M, Zhang Z, Wang L, Huang H, Wang J, Zhu C, et al. Cross-talk of Three Molecular Subtypes of Telomere Maintenance Defines Clinical Characteristics and Tumor Microenvironment in Gastric Cancer. Journal of Cancer. 2024;15(10):3227-41.
27. Barthel FP, Wei W, Tang M, Martinez-Ledesma E, Hu X, Amin SB, et al. Systematic analysis of telomere length and somatic alterations in 31 cancer types. Nature genetics. 2017;49(3):349-57.
28. Zou JJ, Chu SD, Bao QF, Zhang YY. Telomere maintenance genes-derived prognosis signature characterizes immune landscape and predicts prognosis of head and neck squamous cell carcinoma. Medicine. 2023;102(31):e34586.
29. Patiyal S, Dhall A, Raghava GPS. Prediction of risk-associated genes and high-risk liver cancer patients from their mutation profile: benchmarking of mutation calling techniques. Biology methods & protocols. 2022;7(1):bpac012.
30. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome research. 2018;28(11):1747-56.
31. Zheng S, Wang X, Fu Y, Li B, Xu J, Wang H, et al. Targeted next-generation sequencing for cancer-associated gene mutation and copy number detection in 206 patients with non-small-cell lung cancer. Bioengineered. 2021;12(1):791-802.
32. Li J, Xie L, Xie Y, Wang F. Bregmannian consensus clustering for cancer subtypes analysis. Computer Methods and Programs in Biomedicine. 2020;189:105337.
33. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences. 2003;100(16):9440-5.
34. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics: A Journal of Integrative Biology. 2012;16(5):284-7.
35. Xia S, Tao H, Su S, Chen X, Ma L, Li J, et al. DNA Methylation Variation Is Identified in Monozygotic Twins Discordant for Congenital Heart Diseases. Congenital Heart Disease. 2024;19(2):247-56.
36. Li Q, Chu Y, Yao Y, Song Q. A Treg-related riskscore model may improve the prognosis evaluation of colorectal cancer. The Journal of Gene Medicine. 2024;26(2):e3668.
37. Therneau TM, Lumley T. Package 'survival'. R: Get the Top Words and Documents in Each Topic. 2015;128(10):28-33.
38. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing

risks. Statistics in Medicine. 2013;32(30):5381-97.
39. Harrell Jr FE, Harrell Jr MFE, Hmisc D. Package 'rms'. Vanderbilt University. 2017;229(Q8).
40. 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 Biology. 2016;17(1):218.
41. Zhang X, Jin M, Yao X, Liu J, Yang Y, Huang J, et al. Upregulation of LncRNA WT1-AS Inhibits Tumor Growth and Promotes Autophagy in Gastric Cancer via Suppression of PI3K/Akt/mTOR Pathway. Current Molecular Pharmacology. 2024;17.
42. Amuthalakshmi S, Sindhuja S, Nalini CN. A Review on PCR and POC-PCR - A Boon in the Diagnosis of COVID-19. Current Pharmaceutical Analysis. 2022;18(8):745-64.
43. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. Plos One. 2014;9(9):e107468.
44. Shen W, Song Z, Zhong X, Huang M, Shen D, Gao P, et al. Sangerbox: a comprehensive, interaction-friendly clinical bioinformatics analysis platform. Imeta. 2022;1(3):e36.
45. Song ZG, Yu JB, Wang MM, Shen WT, Wang CC, Lu TY, et al. CHDTEPDB: Transcriptome Expression Profile Database and Interactive Analysis Platform for Congenital Heart Disease. Congenital Heart Disease. 2023;18(6):693--701.
46. Zhang L, Li D, Du F, Huang H, Yuan C, Fu J, et al. A panel of differentially methylated regions enable prognosis prediction for colorectal cancer. Genomics. 2021;113(5):3285-93.
47. Iranmanesh H, Majd A, Nazemalhosseini Mojarad E, Zali MR, Hashemi M. Investigating the Relationship between the Expression Level of Membrane-Bound Mucin (MUC12) and Clinicopathological Characterization of Colorectal Cancer. Archives of Advances in Biosciences. 2021;12(1):31-6.
48. Hedrick CC, Malanchi I. Neutrophils in cancer: heterogeneous and multifaceted. Nature Reviews Immunology. 2022;22(3):173-87.
49. Hu J, Yu A, Othmane B, Qiu D, Li H, Li C, et al. Siglec15 shapes a non-inflamed tumor microenvironment and predicts the molecular subtype in bladder cancer. Theranostics. 2021;11(7):3089-108.
50. Tomasova K, Kroupa M, Forsti A, Vodicka P, Vodickova L. Telomere maintenance in interplay with DNA repair in pathogenesis and treatment of colorectal cancer. Mutagenesis. 2020;35(3):261-71.
51. Ma SZ, Zhu XM, Xin C, Cao F, Xu MN, Han XL, et al. RCN3 Expression Indicates Prognosis in Colorectal Cancers. Oncologie. 2022;24(4):823--33.
52. Louzon M, Zahn S, Capelli N, Massemin S, Coeurdassier M, Pauget B, et al. Impact of ageing and soil contaminants on telomere length in the land snail. Ecotoxicology and Environmental Safety. 2020;201:110766.
53. Hambleton P, Barbieri M-A. A hypothesis for a novel role of RIN1-the modulation of telomerase function by the MAPK signaling pathway. BioCell. 2020;44(4):525--34.
54. Huang Z, Zhang Z, Zhou C, Liu L, Huang C. Epithelial–mesenchymal transition: The history, regulatory mechanism, and cancer therapeutic opportunities. MedComm. 2022;3(2):e144.
55. Suzuki T, Conant A, Curow C, Alexander A, Ioffe Y, Unternaehrer JJ. Role of epithelial-mesenchymal transition factor SNAI1 and its targets in ovarian cancer aggressiveness.

Journal of Cancer Metastasis and Treatment. 2023;9.
56. Tran HD, Luitel K, Kim M, Zhang K, Longmore GD, Tran DD. Transient SNAIL1 Expression Is Necessary for Metastatic Competence in Breast Cancer. Cancer Research. 2014;74(21):6330-40.
57. Fang J, Ding Z. SNAI1 is a prognostic biomarker and correlated with immune infiltrates in gastrointestinal cancers. Aging. 2020;12(17):17167-208.
58. Qing F, Xue J, Sui L, Xiao Q, Xie T, Chen Y, et al. Intestinal epithelial SNAI1 promotes the occurrence of colorectal cancer by enhancing EMT and Wnt/β-catenin signaling. Medical oncology (Northwood, London, England). 2023;41(1):34.
59. Galindo-Pumariño C, Collado M, Castillo ME, Barquín J, Romio E, Larriba MJ, et al. SNAI1-expressing fibroblasts and derived-extracellular matrix as mediators of drug resistance in colorectal cancer patients. Toxicology and Applied Pharmacology. 2022;450:116171.
60. Molparia B, Oliveira G, Wagner JL, Spencer EG, Torkamani A. A feasibility study of colorectal cancer diagnosis via circulating tumor DNA derived CNV detection. Plos One. 2018;13(5):e0196826.
61. Liu K, Zheng M, Lu R, Du J, Zhao Q, Li Z, et al. The role of CDC25C in cell cycle regulation and clinical cancer therapy: a systematic review. Cancer Cell International. 2020;20(1):213.
62. Korbecki J, Bosiacki M, Barczak K, Łagocka R, Brodowska A, Chlubek D, et al. Involvement in Tumorigenesis and Clinical Significance of CXCL1 in Reproductive Cancers: Breast Cancer, Cervical Cancer, Endometrial Cancer, Ovarian Cancer and Prostate Cancer. International Journal of Molecular Sciences. 2023;24(8):7262.
63. Sun NH, Zhao X. Therapeutic implications of FABP4 in cancer: an emerging target to tackle cancer. Frontiers in Pharmacology. 2022;13:948610.
64. Chen HP, Chen SY, Chen C, Li AF, Wei ZX. Leucine zipper downregulated in cancer 1 may serve as a favorable prognostic biomarker by influencing proliferation, colony formation, cell cycle, apoptosis, and migration ability in hepatocellular carcinoma. Frontiers in Genetics. 2022;13:900951.
65. Li H, Liu J, Liu WH, Zheng L, Chen JH. Investigation of potential prognostic biomarkers for colorectal cancer. Archives of Medical Science. 2023.
66. Paval DR, Di Virgilio TG, Skipworth RJ, Gallagher IJ. The emerging role of intelectin-1 in cancer. Frontiers in Oncology. 2022;12:767859.
67. Iranmanesh H, Majd A, Mojarad EN, Zali MR, Hashemi M. Investigating the Relationship Between the Expression Level of Mucin Gene Cluster (MUC2, MUC5A, and MUC5B) and Clinicopathological Characterization of Colorectal Cancer. Galen medical journal. 2021;10:e2030.
68. Williams SJ, McGuckin MA, Gotley DC, Eyre HJ, Sutherland GR, Antalis TM. Two novel mucin genes down-regulated in colorectal cancer identified by differential display. Cancer Res. 1999;59(16):4083-9.
69. Cox KE, Liu S, Lwin TM, Hoffman RM, Batra SK, Bouvet M. The Mucin Family of Proteins: Candidates as Potential Biomarkers for Colon Cancer. Cancers (Basel). 2023;15(5).
70. Pothuraju R, Chaudhary S, Rachagani S, Kaur S, Roy HK, Bouvet M, et al. Mucins, gut microbiota, and postbiotics role in colorectal cancer. Gut microbes. 2021;13(1):1974795.

71. Tiwari A, Trivedi R, Lin S-Y. Tumor microenvironment: barrier or opportunity towards effective cancer therapy. Journal of Biomedical Science. 2022;29(1):83.
72. De Visser KE, Joyce JA. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. Cancer Cell International. 2023;41(3):374-403.
73. Folkman J. Role of angiogenesis in tumor growth and metastasis: Elsevier; 2002. 15-8 p.
74. Plikus MV, Wang X, Sinha S, Forte E, Thompson SM, Herzog EL, et al. Fibroblasts: Origins, definitions, and functions in health and disease. Cell Division. 2021;184(15):3852-72.

TABLES AND FIGURES WITH LEGENDS

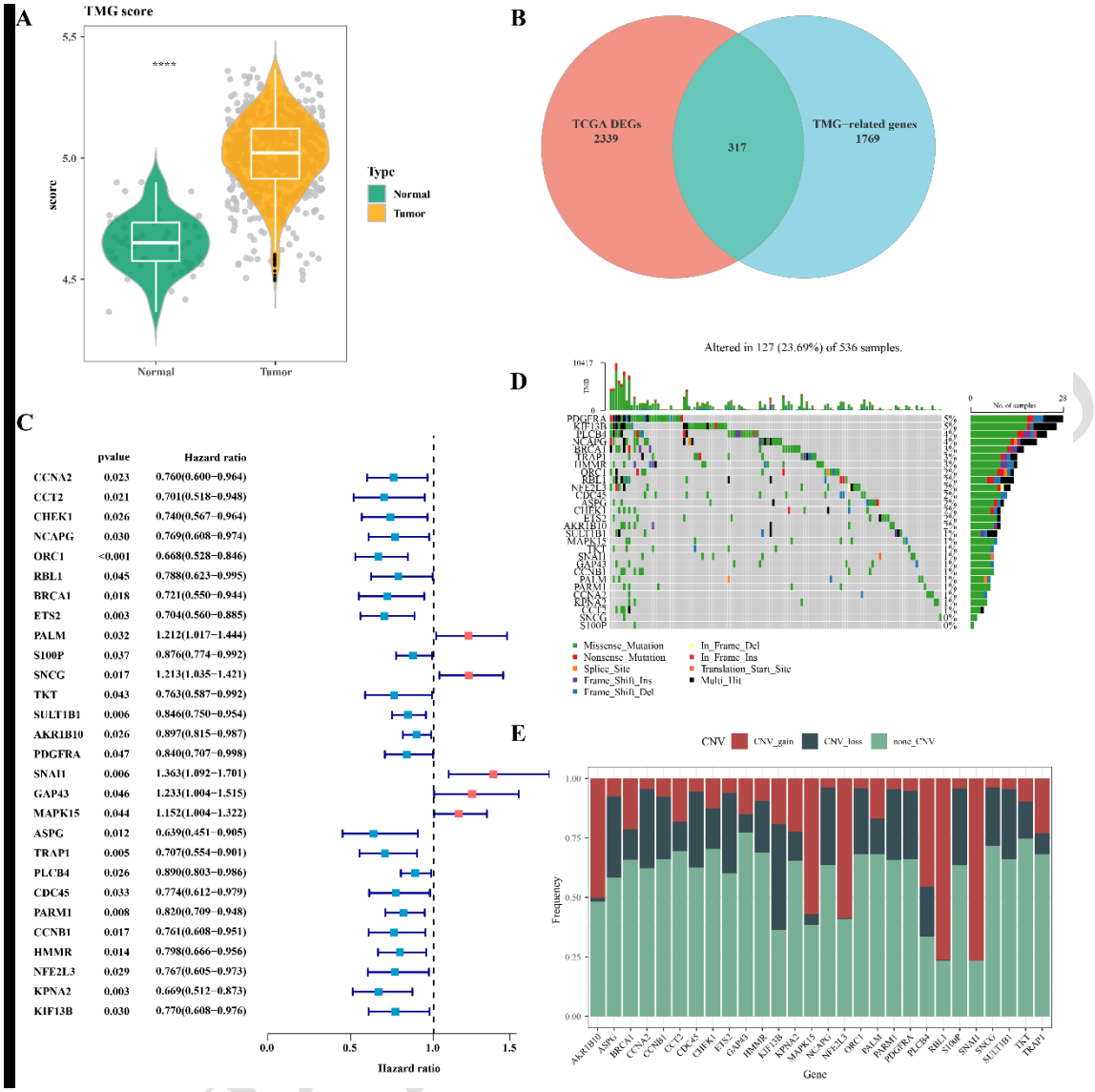


Figure 1. Analysis of the genomic landscape of TMGs in CRC.

(A) The expression of MG scores in CRC and non-cancerous adjacent control samples. (B) The differential genes of tumor tissue and adjacent tissue were intersected with TMG. (C) 28 TMGs were closely linked to CRC prognosis. (D) Mutation status of TMGs in CRC. (E) CNVs of TMGs in CRC.

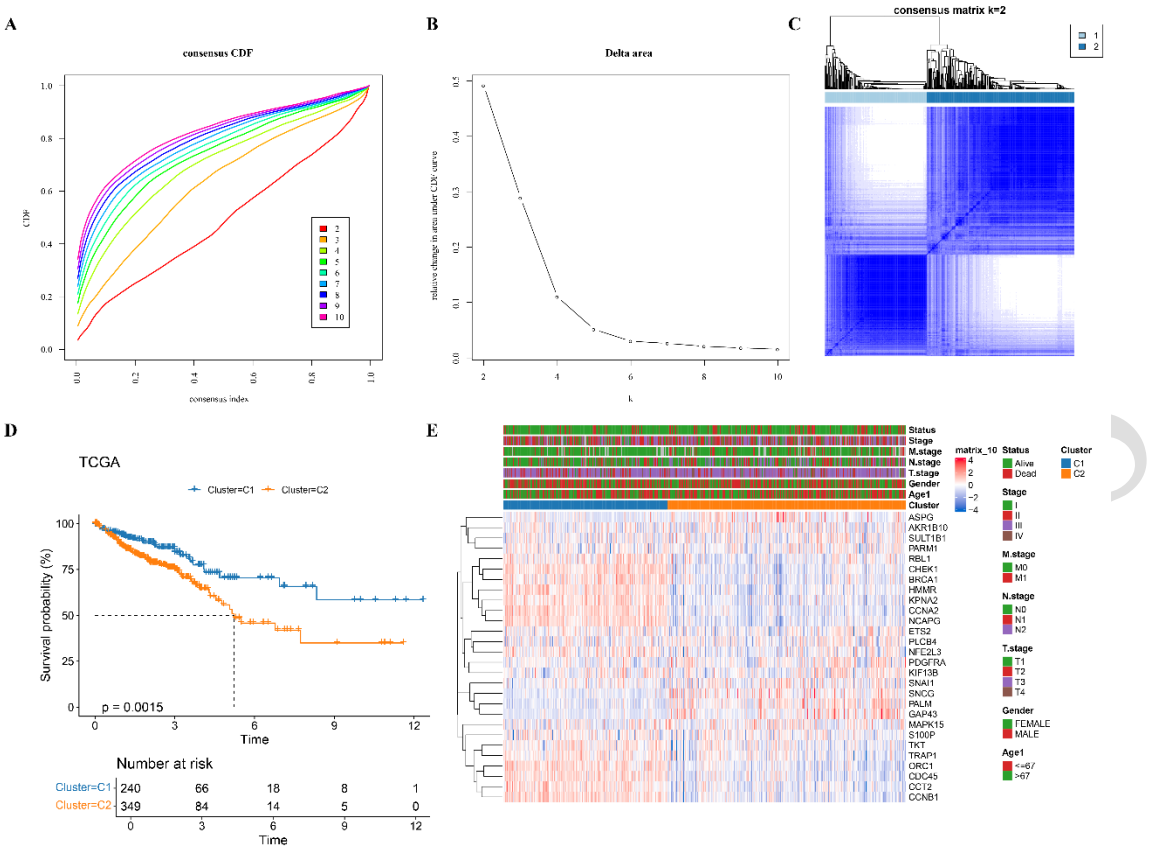


Figure 2. Classification and prognostic differences of TCGA-CRC samples.

(A) CDF curve was plotted for the TCGA-CRC cohort samples. (B) CDF delta area curve was plotted for the TCGA-CRC cohort samples. (C) At consensus k=2, heatmap of sample clustering in the TCGA-CRC cohort was generated. (D) KM curve displaying the relationship between OS and two subtypes in the TCGA-CRC cohort. (E) Heat maps of clinical features (status, stage, M.stage, N.stage, T.stage, age, and gender) and expression between subtypes in the TCGA-CRC cohort.

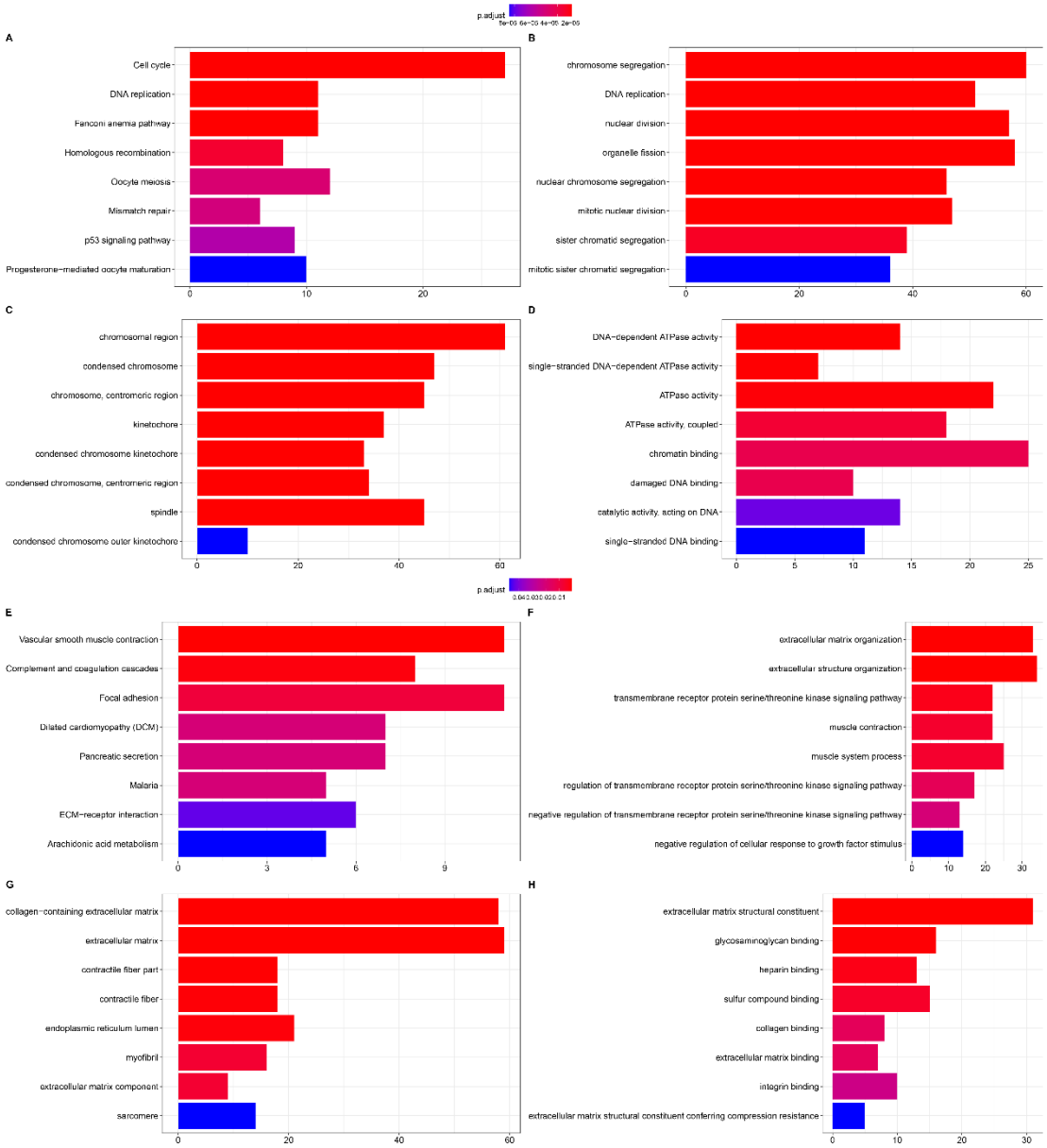


Figure 3. Differential gene identification and enrichment analysis.

(A-D) In TCGA-CRC cohort, DEGs enrichment analysis of C1 subtype. (E-H) In TCGA-CRC cohort, DEGs enrichment analysis of C2 subtype.

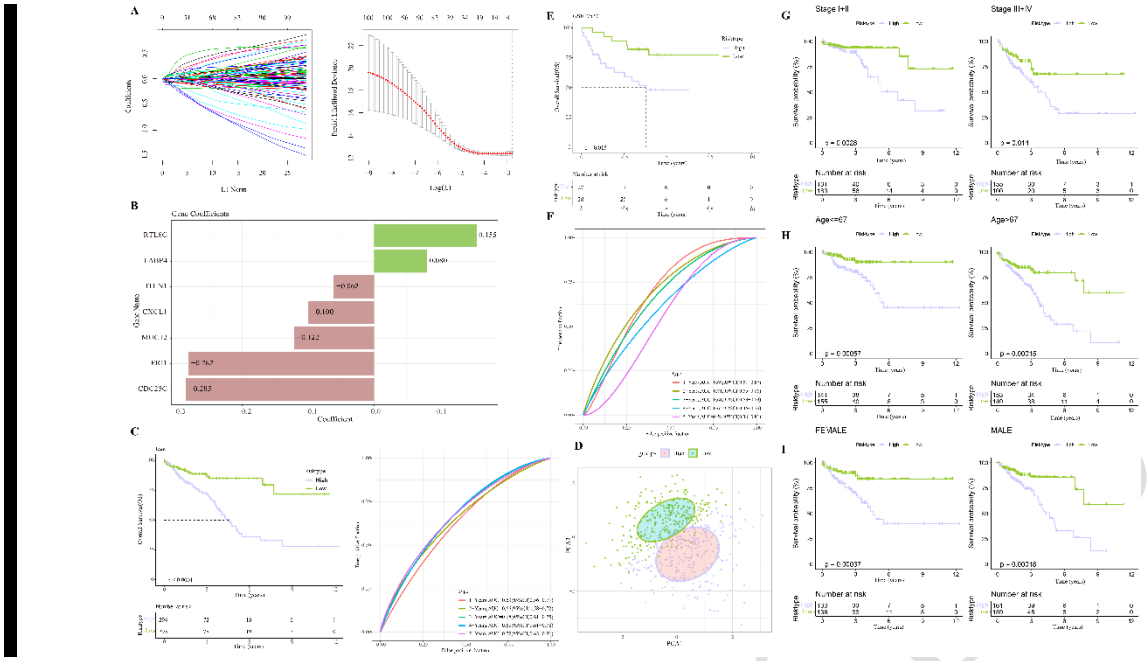
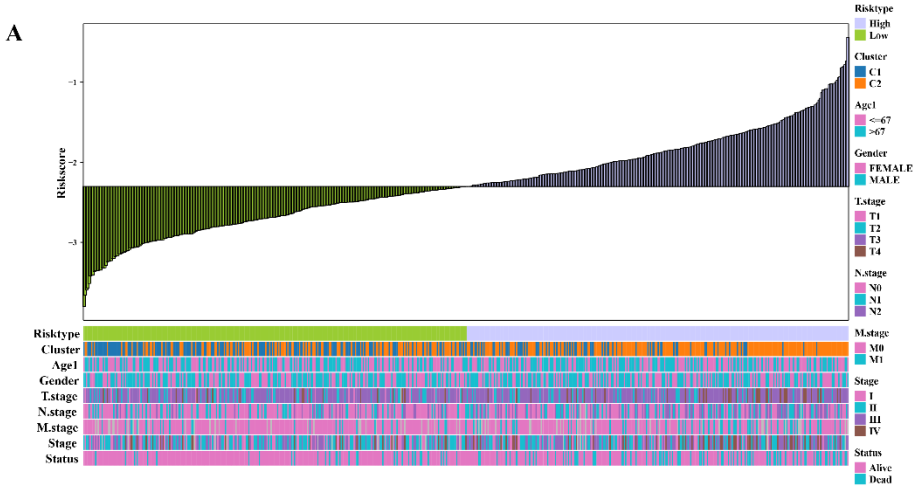
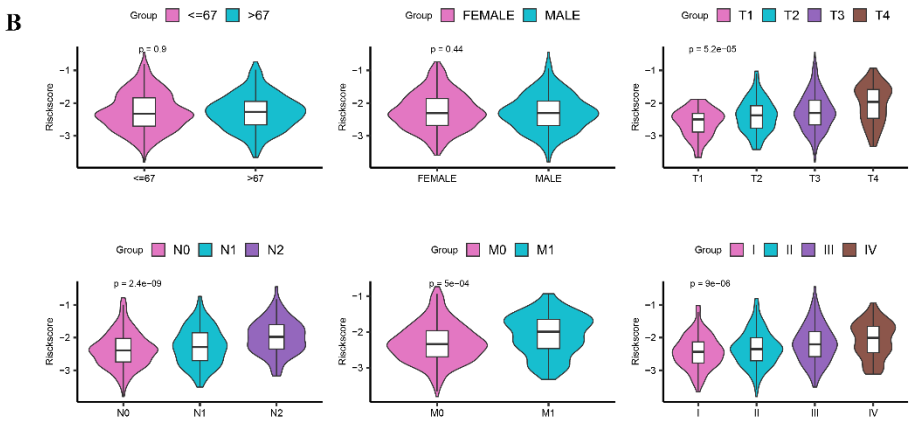


Figure 4. Development of TMGs-based risk model and verification.

(A) LASSO Cox regression analysis was performed to analyze the DEGs linked to CRC prognosis in the TCGA-CRC training cohort. (B) Risk coefficients of key genes in the TCGA-CRC training cohort. (C) KM survival curve and ROC curve for 1-, 3- and 5-year prognostic prediction for TCGA-CRC training cohort. (D) PCA comparing low- and high-risk groups within the TCGA-CRC training cohort. (E-F) KM survival curves and ROC curves for the model based on the GEO testing dataset. (G-I) Prognosis differences between the two risk groups across different tumor stages (G), ages (H), and genders (I).



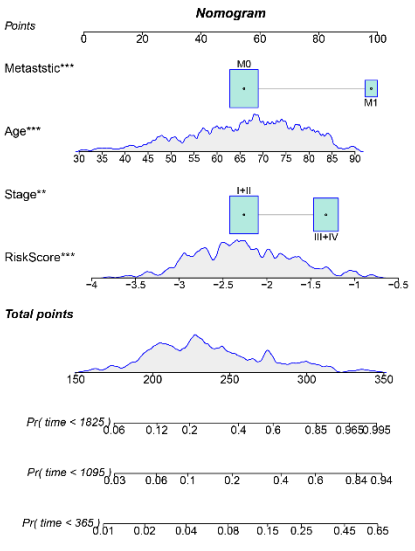


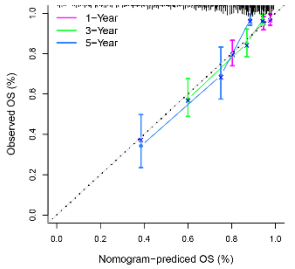
C

Names	p.value	Hazard Ratio(95% CI)
Age	<0.001	1.03(1.01,1.05)
Gender	0.738	1.06(0.74,1.53)
T.stage	0.006	2.5(1.3,4.79)
N.stage	<0.001	2.76(1.89,4.04)
M.stage	<0.001	4.55(3.04,6.81)
Stage	<0.001	3.21(2.15,4.82)
RiskScore	<0.001	2.72(1.99,3.71)

D

Names	p.value	Hazard Ratio(95% CI)
Age	<0.001	1.04(1.02,1.06)
T.stage	0.117	1.98(0.84,4.64)
N.stage	0.077	0.42(0.16,1.1)
M.stage	<0.001	2.83(1.73,4.61)
Stage	0.007	4.58(1.53,13.72)
RiskScore	<0.001	2.13(1.5,3.03)





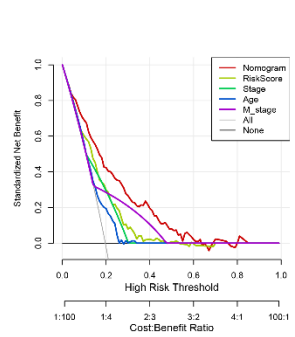


Figure 5. Development of a nomogram to assess CRC prognosis.

(A) Relationship between RiskScore and clinical characteristics (M.stage, N.stage, T.stage, stage, status, age, and gender). (B) Violin plots illustrating the distribution of clinical characteristics (status, M.stage, N.stage, T.stage, stage, age and gender) between low-risk and high-risk groups. (C-D) Univariate and multivariate COX regression analyses were performed to determine the effect of RiskScore and clinical characteristics (status, stage, M.stage, N.stage, T.stage, age and gender). (E) Nomogram to predict the 1-, 3-, and 5-year OS of CRC patients. (F) Calibration curve used to verify established nomogram. (G) DCA of nomogram. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

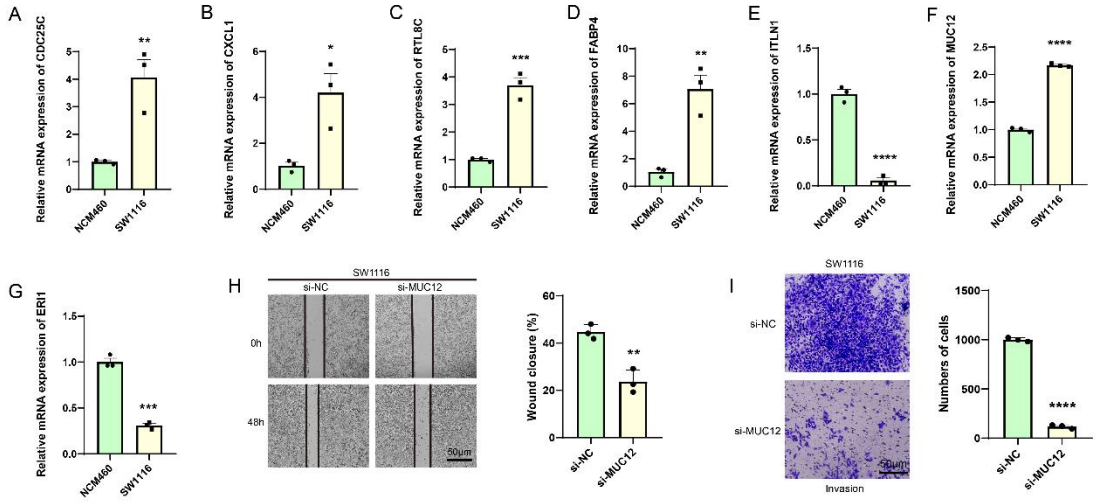


Figure 6. *In-vitro* validation using CRC cells

(A-G) Quantified expression levels of 7 biomarkers CDC25C (A), CXCL1 (B), RTL8C (C) FABP4 (D), ITLN1 (E), MUC12 (F) and ERI1 (G) in CRC cells SW1116 and human normal colonic epithelial cells NCM460 via qRT-PCR. (H) Effects of MUC12 silencing on the migration of CRC cells SW1116 tested via wound healing assay. (I) Effects of MUC12 silencing on the invasion of CRC cells SW1116 tested via transwell assay. All the data from three independent experimental sets were shown as mean ± standard deviation. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

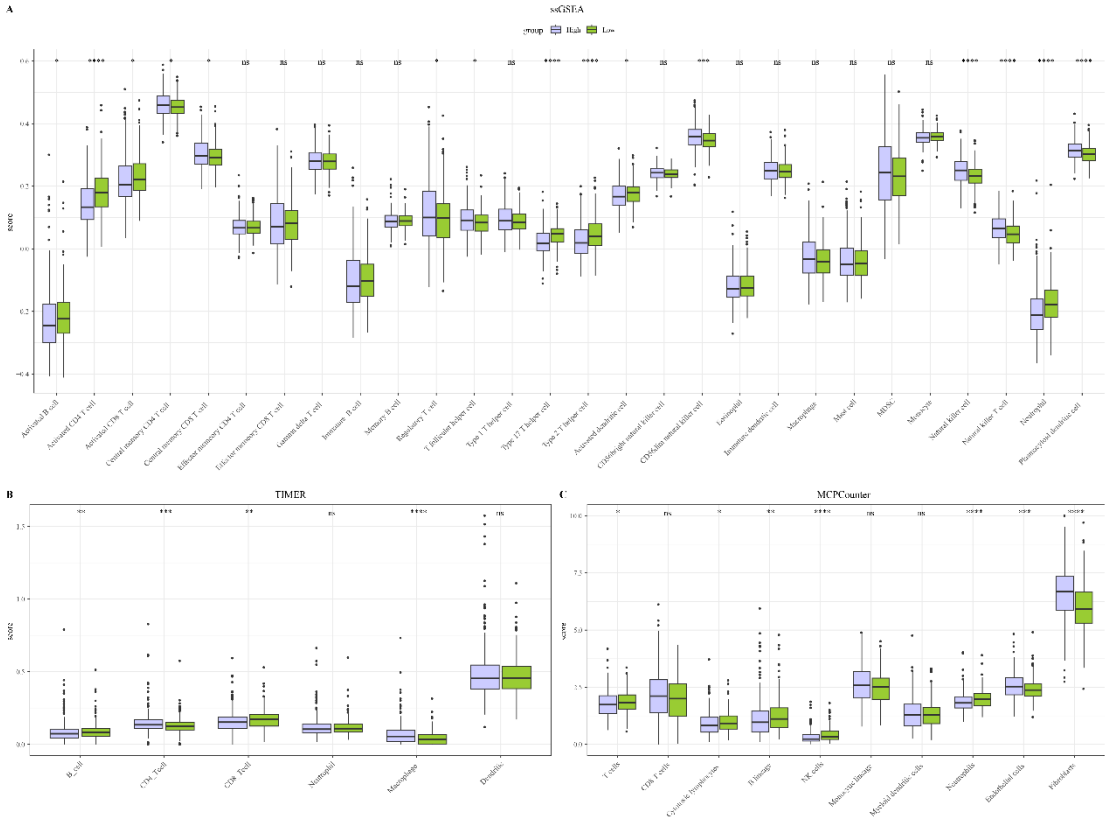


Figure 7. Analysis of differences in immune infiltration levels between high- and low-risk groups

(A-C) The method of (A) ssGSEA, (B)TIMER and (C)MCPcounter was used to calculate the immune infiltration levels of the two risk groups, respectively.

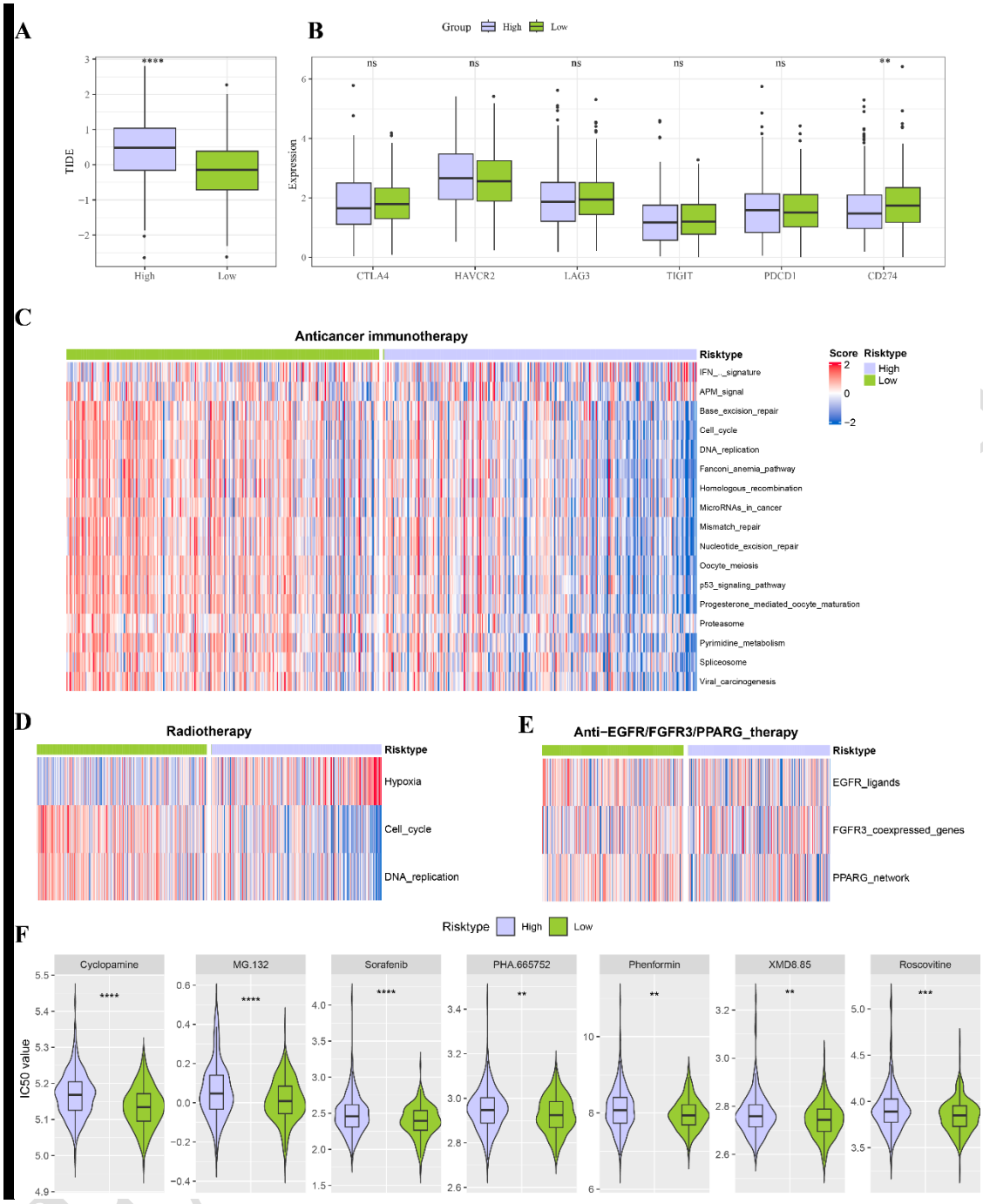


Figure 8. Immunotherapy and drug sensitivity studies between high and low risk groups.

(A) Differences in TIDE scores between high and low risk groups. (B) Variations in common immune checkpoint expressions between high and low risk groups. (C-E) Results of differences in response to (C) Anticancer immunotherapy，(D) Radiotherapy，(E) Anti−EGFR/FGFR3/PPARG_therapy between high and low risk groups. (F) Differences in drug sensitivity of Cyclopamine, MG-132, Sorafenib, PHA.665752, Phenformin, XMD8.85 and Roscovitine were observed between high-risk and low-risk groups.

SUPPLEMENTAL DATA

Table S1. Primer sequences used in this study.

Gene	Accession No.	Primers (5'-3') Forward	Reverse
CDC25C	NM_001790	AAGGCGGCTACAGAGACTTCTT	AGAGTTGGCTGGCTTGTGAGA
CXCL1	NM_001511	TGCTGCTCCTGCTCCTGGTA	GCTTTCCGCCCATTCTTGAGTG
RTL8C	NM_001078171	AAGCGAGGAGCAGCGATGGA	TGTGAGGCGGGTGATGAGGAA
FABP4	NM_001442	TGCAGCTTCCTTCTCACCTTGA	TGACGCATTCCACCACCAGTT
ITLN1	NM_017625	AACGCCTTGTGTGCTGGAATGA	ATCTCACGGCTGCTGCTGTAAC
MUC12	NM_001164462	CCTCAACTCACACGACGCCTTC	TGCTGCTGTAGACGGTGGTAGA
ERI1	NM_153332	ATCCTCTTGCCTCAGCCTCCT	TTCAAGACCAGCCTGACCAACA
GAPDH	NM_002046	GTCTCCTCTGACTTCAACAGCG	ACCACCCTGTTGCTGTAGCCAA