Development and Validation of a Unique Gignature of Cancer Stemness-related Genes for Predicting the Overall survival of Colorectal Cancer Patients

Ran Wei
Chinese Academy of Medical Sciences & Peking Union Medical College
Jichuan Quan
Chinese Academy of Medical Sciences & Peking Union Medical College
Shuofeng Li
Chinese Academy of Medical Sciences & Peking Union Medical College
Zhao Lu
Chinese Academy of Medical Sciences & Peking Union Medical College
Xu Guan
Chinese Academy of Medical Sciences Institute of Hematology and Blood Diseases Hospital
Zheng Jiang
Chinese Academy of Medical Sciences & Peking Union Medical College
Xishan Wang (wxshan_1208@126.com)
Chinese Academy of Medical Sciences and Peking Union Medical College

Research

Keywords: colorectal cancer, cancer stem cell, tumor immune microenvironment, machine-learning

Posted Date: December 15th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-125323/v1

Abstract

Background: Cancer stem cells (CSCs), which are characterized by self-renewal and plasticity, are highly correlated with tumor metastasis and drug resistance. To fully understand the role of CSCs in colorectal cancer (CRC), we evaluated the stemness traits and prognostic value of stemness-related genes in CRC.

Methods: In this study, the data from 616 CRC patients from The Cancer Genome Atlas (TCGA) were assessed and subtyped based on the mRNA expression-based stemness index (mRNAsi). The correlations of cancer stemness with the immune microenvironment, tumor mutational burden (TMB) and N6-methyladenosine (m6A) RNA methylation regulators were analyzed. Weighted gene co-expression network analysis (WGCNA) was performed to identify the crucial stemness-related genes and modules. Furthermore, a prognostic expression signature was constructed using Lasso-penalized Cox regression analysis. The signature was validated via multiplex immunofluorescence staining of tissue samples in an independent cohort of 48 CRC patients.

Results: This study suggests that high mRNAsi scores are associated with poor overall survival in stage CRC patients. Moreover, the levels of TMB and m6A RNA methylation regulators were positively correlated with mRNAsi scores, and low mRNAsi scores were characterized by increased immune activity in CRC. The analysis identified 2 key modules and 34 key genes as prognosis-related candidate biomarkers. Finally, a 3-gene prognostic signature (PARPBP, KNSTRN and KIF2C) was explored together with specific clinical features to construct a nomogram, which was successfully validated in an external cohort.

Conclusions: There is a unique correlation between CSCs and the prognosis of CRC patients, and the novel biomarkers related to cell stemness could accurately predict the clinical outcomes of these patients.

Background

Colorectal cancer is among the most common and lethal cancers of the digestive system [1]. Although neoadjuvant chemo-radiotherapy and immunotherapy offer good prospects for operable colorectal cancer cases, the 5-year survival rates remain low in cases of advanced disease [2]. With advances in individualized tumor treatment, the remarkable tumor heterogeneity was shown to be closely associated with chemoresistance, radiosensitivity and patient survival [3]. Cancer stem cell (CSC) traits, a crucial part of cancer heterogeneity, are considered to be crucial drivers of the prognosis and response to therapy [3, 4].

Mounting evidence suggests the existence of CSC in colorectal cancer, with studies revealing their roles in metastasis, drug resistance, and continually helping the cancer adapt to the changing microenvironment [5, 6]. It has been demonstrated that the accumulated epigenetic and genetic variability allows CSC to evolve and thereby continue tumor growth and maintenance, which is closely associated with the alteration of the tumor microenvironment (TME) [7, 8]. Recent studies demonstrated that colon cancer cells with stem-cell-like properties can promote drug resistance, migration and invasion, which is

regulated by the TME and not a fully autonomous behavior of individual cells [9, 10]. Cancer stemness encompasses both the stemness phenotype with bona fide CSCs and the intrinsic potential for differentiation into colon cancer cells, which is defined as a fundamental underlying property of malignancy [11, 12]. Studies investigated the levels of cancer cell stemness markers CD133 and CD44 d in cell culture models in vitro revealed that a large number of cytokines, cellular and molecular mechanisms play significant roles in mediating cancer cell stemness and CSC properties. These include cancer-associated fibroblasts, IL13, exosomes and c-Jun signaling in colorectal cancer [13, 14]. However, most of theories haven't been confirmed in vivo or translated into clinical research, because of the integrated and complex cancer ecosystem.

The interaction between the immune system and CSCs is still controversial, as the increased tumorigenicity of CSCs reveals that they can promote oncogenic immunomodulation in colon cancer [15, 16]. Moreover, cancer cells and CSCs with high stemness can decrease the expression of major histocompatibility complex (MHC) molecules, thereby promoting immune evasion and decreasing the activity of anti-tumorigenic immune cells [17]. However, an integrated understanding of colorectal cancer stemness, including its interactions with the tumor immune microenvironment, still requires further research. To evaluate the role of stemness in tumor pathogenesis and the vital factors leading to dedifferentiation and acquisition of stem-cell-like properties in colorectal cancer, artificial intelligence and bioinformatic methods could be employed to further understand cancer stemness [18, 19]. One-Class Logistic Regression (OCLR) can be sued to extract epigenetic and transcriptomic features from normal stem cells and their differentiated progeny, including induced pluripotent stem cells and embryonic stem cells with different level of stemness. Notably, this approach could also be used to identify the stem cell signatures and quantify cancer stemness via a multi-omics analysis.

In this study, we hypothesized that cancer stemness may confer immunosuppressive properties on tumors and mediate the prognosis. To verify it, we developed the immune score construct and identified the proportions of immune cells using vector regression. We then assessed of correlations the stemness indices with molecular features, and identified a stemness-related gene signature that might be helpful in evaluating the prognosis of colon and rectal cancer patients.

Materials And Methods

Study design

To explore the mRNAsi in colon and rectal cancer, we obtained the transcriptome profiling for the gene expression and clinical information of colon and rectal cases from the TCGA database, and the analysis workflow was shown (Figure 1). We employed the "Cell type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT)" algorithm and “Estimation of Stromal and Immune cells in Malignant Tumors using Expression data (ESTIMATE)” to develop the immune score model construction and identify the proportions of immune cells with support vector regression. And we have analyzed the alteration spectrum reversible N6-methyladenosine (m6A) RNA methylation regulators in different level of

mRNAsi colon and rectal cancer group. Besides we assessed stemness-related genes sets with WGCNA. We identified signatures from a key gene module, and their prognostic role in colon and rectal cancer. Our studies applied some multi-omics approaches to identify the features of cancer stemness and define the roles of stemness-related genes in colorectal cancer.

Data sources

Within the The Cancer Genome Atlas (TCGA) database (https://tcga-data.nci.nih.gov/tcga/), we identified the transcriptome profiling by RNA-sequencing (RNA-seq), single nucleotide variants (SNV) and the corresponding prognostic and clinicopathological information of the colon and rectal cancer set. And the Ensembl IDs were converted to gene symbols based on the Ensemble database (http://asia.ensembl.org/index.html). The RNA-seq results of 433 tissues and 408 colon cancer samples were obtained from COAD, and 221 tissues and 208 rectal cancer samples were obtained from READ with the FPKM method [20]. The SNV results were obtained from COAD and READ with the VarScan2 Variant Aggregation and Masking.

Colorectal cancer patients

The colorectal cancer specimens who underwent the surgical resection from January 2006 through December 2012 were approved by the Pathology Department of Cancer Institute and Hospital, Chinese Academy of Medical Sciences. All of the patients have gotten informed consent. Patients who met the following criteria were included: (A) no getting neoadjuvant radiotherapy or chemotherapy before the surgery; (B) with complete clinicopathological information, including the sex, age, tumor location, TNM classification, follow-up time and so on. And the specimens in this study were followed up every three months until 1 January 2018. Forty-eight tumor tissues and adjacent normal tissues were collected from colorectal cancer patients with untreated stage □ to stage IV.

mRNAsi in subtypes

Unsupervised learning, as a machine learning technique, is employed in the field of medical data mining, and Wiznerowicz et al. [19] draw two stemness indexes based on reflected epigenetic regulation features and transcriptome, respectively, named mRNAsi and mDNAsi with OCLR. The mRNAsi is used to identify cancer stemness features and assess the degree associated with oncogenic dedifferentiation, which could be a quantitative form of cancer CSCs. Higher mRNAsi scores are related with cancer biological processes in CSCs and more tumor dedifferentiation based on the histopathological grades. Besides these studies further evaluated annotated cancer stemness in nearly 12,000 samples, covering 33 tumor types from TCGA. And the stemness feature was performed employing colon and rectal cancer specimens from TCGA-COAD and -READ to get the comprehensive molecular subtyping of each specimen

[19]. And we divided colon and rectal cancer patients into high-mRNAsi group and low-mRNAsi group based on the cut-off value determined by the median mRNAsi index.

Differentially expressed genes (DEGs)

The “limma” packages in R language was employed to identified differentially expressed genes in the expression data from the TCGA, and Wilcox Test was used in the analysis processing [21]. The |log2-fold change|>1 and false discovery rate (FDR)<0.05 were considered to be the cut-offs criteria for DEGs selection between colon and rectal cancer and normal sets. The volcano plot and heat-map were drawn with the “pheatmap” package in R.

Estimation of immune microenvironment and infiltrating immune cells

We employed the ESTIMATE to evaluate the immune score, immune cell infiltration level, for each colon and rectal cancer samples from TCGA [22]. And then, we evaluated the enrichment levels of the 29 immune signatures in each colon and rectal cancer specimen by the single-sample gene-set enrichment analysis (ssGSEA) score [23-25]. The colon and rectal cases could be classified into three subgroups based on the ssGSEA score and hierarchical clustering. Besides, we evaluated the 22 human immune cell subsets of every colon and rectal cancer samples with CIBERSORT web portal (http://cibersort.stanford.edu/) and 1000 permutations [26]. The CIBERSORT deconvolution algorithm output had a p-value<0.05 was considered accurate and successful deconvolution. And the output estimates would be normalized for each sample to add up to one, enabling their direct interpretation as cell fractions for comparison across different groups. The package “Genefilter” R was employed to identify each sample.

m6A RNA methylation regulators

Twenty-one m6A RNA methylation regulators, including *METTL16, METTL14, METTL3, RBM15, KIAA1429, WTAP, ZC3H13, YTHDC2, YTHDC1, YTHDF2, YTHDF1, HNRNPC, YTHDF3, FTO, HNRNPC, IGF2BP1, HNRNPA2B1, RBMX, ALKBH5, IGF2BP3, IGF2BP2* [27,28], were used in this study. And mRNA expression data of these m6A RNA methylation regulators were compared between high mRNAsi scores and low mRNAsi scores groups. The package “corrplot” R was employed to identify the Pearson correlations between genes. The correlation between the expression of these genes could be evaluated in different groups.

Weighted Gene Co-expression Network Analysis (WGCNA)

We used the “WGCNA” R package to establish the co-expression network of differentially expressed genes [29]. We employed Pearson’s correlation matrices, co-expression similarity matrix and average linkage method to evaluate the correlations among the included genes. We used the function Amn=|Cmn|β (Cmn=Pearson’s correlation between gene-m and gene-n; Amn=adjacency between gene m and gene n; β representing soft thresholding parameter) to distinguish the strength of correlations and build a weighted adjacency matrix with a scale-free co-expression network. We used a topological overlap matrix (TOM) to evaluate the connectivity and dissimilarity of the co-expression network established with an appropriate β value. Based on the TOM dissimilarity measurements, the average hierarchical linkage clustering could be established, and we set the minimum genome to 30 to build module dendrograms.

In order to confirm the key modules and genes, we set the module membership (MM) and gene significance (GS) to be the measure used to identify the correlation between genes and mRNAsi and epigenetically regulated mRNAsi (EREG-mRNAsi). The module eigengenes (MEs) were defined as the significant components of the principal component analysis (PCA) for each gene module, where the expression level of every gene could be grouped with a distinct feature. We used a log10 transformation of the p-value (GS=lgp) for the linear regression of correlations between clinical phenotypes and gene expression. We used module significance (MS) to represent the correlation between clinical traits and gene expression calculated using the average GS in the module. We used a cut-off of < 0.25 to merge highly similar modules, which could help cluster the key genes. The thresholds for the screening of key module genes were set as cor. gene GS > 0.5 and cor. gene MM > 0.6.

Gene set, ontology and pathway enrichment analysis

The package “clusterProfiler” in R was employed to evaluate the enrichment analysis of Kyoto Encyclopedia of Genes and Genomes (KEGG) and gene ontology (GO) to reveal the key biological functions of the module genes [30]. We set p-value < 0.05 and an FDR < 0.05 as the statistically criteria to output. The whole transcriptome of all tumor samples was employed for GSEA, and only gene sets with NOM p < 0.05 and FDR q < 0.05 were set as statistically criteria.

Protein–Protein Interaction (PPI) Network

We used Search Tool for the Retrieval of Interacting Genes (STRING) Version 11.0 (https://string-db.org/) to establish the PPI network, in order to determine the co-expression relationship between key module genes [31].

Oncomine and Cbioportal database

Oncomine (http://www.oncomine.org/) was employed to identify differences in transcriptomic RNA-sequencing data of module genes between tumors and normal tissues in colorectal cancer. The threshold

limits were as follows: p-value, 0.05; fold change, all; gene level, 10%; data type, mRNA. We compared the transcriptome level of key module genes in clinical cancer specimens were determined by Student's t test. And Cbioportal (http://www.cbioportal.org/) was used to identified the somatic mutations of module genes in colorectal cancer.

Multiplex immunofluorescence staining and image analysis

To evaluate the expression and distribution of three key genes related to tumor stemness in colorectal cancer and normal tissues, we performed multiplex immunofluorescence staining using the PANO 7-plex IHC kit (cat. No. 0004100100; Panovue, Beijing, China) and Tyramide Signal Amplification Fluorescence Kit (Panovue, Beijing, China) [32]. We established the colorectal cancer tissue microarrays (TMAs) consisting of primary tumor, metastatic tumor and matched normal tissue from cancer patients who had been confirmed by pathological examination with hematoxylin and eosin (H&E) staining [6]. Each of these tissues was cut into pieces of 1.0 mm and attached to the slides (5 mm thick) from the TMAs. The TMAs were incubated with anti-PARPBP (ab211634; 1:100; Abcam; Cambridge, UK), anti-KNSTRN (ab122769; 1:100; Abcam; Cambridge, UK) and anti-KIF2C (12139-1-AP; 1:200; Proteintech, Rosemont, IL) antibodies at 4°C overnight, and then with horseradish peroxidase-conjugated secondary antibody and tyramide. A microwave was uesd to heat-regenerate the TMAs after each Tyramide Signal Amplification step. We used 4',6-diamidino-2-phenylindole (DAPI) to counter-stain the cell nuclei.

The Mantra System (PerkinElmer, Waltham, Massachusetts, US) was used to capture the multispectral immunofluorescence images with the fluorescence spectra at 20-nm wavelength intervals from 420 to 720 nm with the same exposure time, which were then composited to establish a single stack image. To capture images of sections without autofluorescence, we extracted the spectrum of autofluorescence of TMAs and each fluorescein from the images of unstained and single-stained sections, which were used to establish the spectral library for multispectral unmixing using inForm image analysis software (PerkinElmer, Waltham, Massachusetts, US). Two independent pathologists analyzed and counted single-positive cells and the expression of the three genes in each tissue of TMAs at 200× magnification in a blinded manner. The nucleated stained cells were quantified and expressed as the number of cells in TMAs. The positive rate of single index and single index intensity score were employed to evaluate the expression and distribution of the identified key genes in colorectal cancer, which were calculated by multiplication of the multiplex immunofluorescence staining intensity (percentage of single index %= Number of positive cells/ Total number of cells; The 25% staining was taken as the threshold of the strength score; 25–49%, strength I; 50–74%, strength II; 75–100%, strength III; single index strength score= [(strength I*positive rate of single index) *1+(strength II*positive rate of single index) *2+(strength III*positive rate of single index) *3] *100) [33].

Statistical analysis

In this study, the colon and rectal cancer specimens were allocated into training and testing group randomly through the 2:1 ratio with the package “caret”, in order to promote the generalization ability of model. And we established the LASSO-penalized Cox regression model to identify the most significantly survival-related gene signatures with the patients’ overall survival (OS). We set 10-fold cross validation as the criteria to prevent overfitting with the penalty parameter lambda.1se [34]. Then the time dependent receiver operating characteristic (ROC) curve and the area under the curve (AUC) were employed to identify the prognostic accuracy of the three gene signature model in training and testing group with the package “survival ROC” [35]. The median of risk scores was set as the cut-off value to the separate patients into high risk and low risk score group. We employed the Kaplan–Meier survival analysis and the log-rank test to evaluate the difference in OS between different groups. The final (forward and backward elimination methods) Multivariate Cox regression analysis was employed to evaluate the independence of predictors and established three signatures and nomogram. The we validate the performance of the Cox model internally and externally with bootstrap method. Bootstrap-corrected OS rates were calculated by averaging the Kaplan-Meier estimates based on 2000 bootstrap samples.

Results

1. Stemness properties of colon and rectal cancer cells

Colon and rectal cancer patients could be subtyped using the mRNAsi

The mRNAsi of colon and rectal cancer tissues was significantly higher than that of normal tissues (Wilcoxon rank sum test, p<0.05) (Figures 2A and H), which suggested that the level of stemness is associated with tumor development. The mRNAsi among different stages of colon and rectal cancer did not show significant differences (Mann-Whitney U test, p<0.05) (Figures 2B, C, I and J), which suggested the mRNAsi may not be closely related to the clinical stage in many tumor types [36,37]. In our study, the colon and rectal cancer patients were divided into high-mRNAsi and low-mRNAsi groups based on the cut-off value determined by the median mRNAsi index. Stage □ colon cancer patients with higher mRNAsi had a significantly shorter OS than those with lower mRNAsi (Log-rank test. P < 0.05, Figures 2D, G, K, and N). These results suggested that the mRNAsi of colon and rectal cancer is closely associated with the prognosis of stage □ colon cancer patients.

Differences in cell stemness and mutations among the mRNAsi subtypes

We compared the mutational landscape between colon and rectal cancer patients in the high- and low-mRNAsi groups. A higher proportion of APC (80.7%) and TP53 (61.9%) mutations were found in the high-mRNAsi group than the proportion of APC (70.7%) and TP53 (51.2%) mutations in the low-mRNAsi group

of colon cancer patients, and the difference was statistically significant (Figure 3, Supplementary Figure 1). Furthermore, oncogenic KRAS mutations, signature factors of the receptor tyrosine kinase (RTK) pathway, were found in similar proportions in the two groups (Figure 3, Supplementary Figure 1).

Differences of cell stemness and immune microenvironment between the mRNAsi subtypes

We used 29 immune-associated gene sets to quantify the enrichment levels of different immune cell functions, pathways and types in colon and rectal cancer samples based on ssGSEA scores [23]. The ssGSEA scores of the 29 gene sets were employed for hierarchical clustering of immune subtypes into high, moderate and low groups (Figure 4A, 4B). We also evaluated the immune score using ESTIMATE, and the immune scores were higher in the high-subtype than those in the low-subtype [23]. The mRNAsi in the high-subtype was higher than in the low-subtype for colon cancer patients. However, there is no significant difference for rectal cancer patients (Kruskal–Wallis test, Wilcoxon rank sum test, $P < 0.001$) (Figures 4C and F). Moreover, we found that the Immune Score was lower in the high-mRNAsi group than in the low-mRNAsi group for both colon and rectal cancer (Wilcoxon rank sum test, $P < 0.001$) (Figures 4D and G). In addition, when comparing the Tumor Mutational Burden (TMB) in groups with different levels of stemness index, we observed the opposite trend, with the level of TMB increasing from the low-mRNAsi group to the high-mRNAsi group (low-mRNAsi < high-mRNAsi) (Wilcoxon rank sum test, $P < 0.001$) (Figures 4E and H).

We investigated the proportions and differences of tumor infiltrating immune cell subsets between the high- and low-mRNAsi subgroups of colon and rectal cancer patients using the CIBERSORT algorithm and the LM22 gene signature (Figures 5A and E). To further confirm the relationships among 22 tumor-infiltrating immune cell types, we performed correlation analysis. The results revealed a positive correlation between CD8+ T cells and M1 macrophages in the high-mRNAsi subgroup. The CD8+ T cells were more strongly negatively correlated with mast cells and resting memory CD4+ T cells in the high-mRNAsi group than the low-mRNAsi group of colon cancer patients (Figures 5B, C, F, and G). Eight types of tumor-infiltrating immune cells were correlated with the mRNAsi in colon cancer and five types in rectal cancer (Wilcoxon rank sum, $P<0.05$) (Figures 5D and H). Among them, six types of tumor-infiltrating immune cells were positively correlated with high mRNAsi in colon cancer, including CD8+ T cells, resting NK cells, activated memory activated CD4+ T cells, follicular helper T cells, as well as resting and activated dendritic cells. Two types of immune cells were positively correlated with low mRNAsi, including regulatory T cells and M1 macrophages.

Therefore, the high-mRNAsi group exhibited higher levels of TMB and a lower immune score, which indicated that tumor stemness may be negatively correlated with tumor immunity and positively correlated with TMB, which is related closely with the production of neoantigens in tumors.

Cell stemness and the response of immunotherapy in subtypes

The immune checkpoint molecules are significant for immunotherapy, including the major histocompatibility complex (MHC) class I and II. And we investigated the potential correlation between the mRNAsi and immune checkpoint molecules. Human leukocyte antigen (HLA) gene expression was enriched in low-mRNAsi subgroup in colon and rectal cancer (Figure 6 A-F). Most HLA genes were significantly abnormally expressed between subgroups. Compared to the high-mRNAsi subgroup, KIAA1429, HLA.DOB, HLA.DQA1, HLA.DOA, HLA.DPA1, HLA.DPB1, HLA.DRB1, HLA.DMA, HLA.DRA, HLA.DMB, HLA.DQB1, HLA.DRB5, HLA.DRB6, HLA.DPB2, HLA.DQA2 and HLA.DQB2 were upregulated in low-mRNAsi subgroup (Wilcoxon rank sum test, *p < 0.05; **p < 0.01; ***p < 0.001) (Figure 6 A-B). To further understand the relationship among this HLA genes, we performed correlation analysis in high and low-mRNAsi groups for colon and rectal cancer patients (Figure 6 C-F). All these results revealed that the level of HLA could be associated negatively with the cancer stemness.

Cell stemness and the expression of m6A RNA methylation regulative factors in subtypes

Accumulated evidence suggested that m6A mRNA methylation is significant for tumor development, self-renewal and regulating the development of cancer stem cells [38,39]. High-mRNAsi was significantly associated with increasing gene expression, such as *METTL3*, *WTAP*, *YTHDF1*, *YTHDF2*, *RBM15*, *ALKBH5*, *HNRNPC* and *FTO*, which are involved in m6A mRNA methylation in colon and rectal (Wilcoxon rank sum test, *p < 0.05; **p < 0.01; ***p < 0.001) (Figure 6 G-H). Correlation analysis was performed in high and low-mRNAsi subgroups of colon and rectal cancer patients to further investigate the relationship among this m6A RNA methylation regulative factors (Figure 6 I-L). The result suggested that the high cancer stemness could be significantly with the high expression of m6A RNA methylation regulative factors and regulated their correlation.

2. The mRNAsi-related gene signature

The screening of DEGs and identification of mRNAsi-related modules

We screened DEGs in datasets of colon and rectal cancer tissues and normal tissues. We identified 6498 DEGs in colon cancer, 4528 of which were up- and 1970 downregulated (Figure 7A). In rectal cancer, 2072 DEGs were up- and 1776 were downregulated (Figure 7F).

A gene co-expression network was established to classify the genes into biologically significant modules based on the average linkage hierarchical clustering strongly associated with the mRNAsi in colon and rectal cancer. In this model, we set β=6 (scale-free R2=0.95) as the soft threshold to construct the scale-free network (Supplementary Figure 2-3), which yielded 9 gene modules in colon cancer and 14 in rectal cancer (Figures 7B, C, G, and H). To further identify the relationship between key gene models and mRNAsi, we set MS as the overall gene expression level of a certain module to identify correlations with stemness phenotypes. We selected the yellow module with a correlation of more than 0.7 in colon cancer and brown module with a correlation of almost 0.6 in rectal cancer, for subsequent analyses (Figures 7B, D, G, and I). We set cor. MM > 0.6 and cor. GS > 0.5 as the selection threshold, and identified 61 key genes significantly related to mRNAsi in colon cancer, as well as 41 key genes in rectal cancer. GO and KEGG pathway enrichment analyses were performed to evaluate the principal biological functions of the key genes in colon and rectal cancer (Supplementary Figure 4-5). The results showed that the keys gene are related to DNA replication, ATPase activity, and condensed chromosome kinetochores in colon cancer (Supplementary Figure 4-5). The key genes were found to be involved in regulation of chromosome segregation, condensed chromosomes, catalytic activity and DNA helicase activity in rectal cancer (Supplementary Figure 4-5). Additionally, we established a PPI network to validate protein interactions using the STRING database, based on the key gene selection in both colon and rectal cancer (Supplementary Figure 4-5). We also analyzed the interactions among these key genes in colon and rectal cancer (Figures 7E and J).

Analysis of key gene expression and mutational landscape

By overlapping the mRNAsi-related genes in colon cancer (Figure 7E) and those in rectal cancer (Figure 7J), we obtained 34 overlapped genes and established the PPI network (Figure 8A, supple Figure). We employed two databases, Cbioportal and Oncomine, to systematically understand the key gene mutational landscape and expression. The results showed that most of genes selected were significantly upregulated in more than one cancer type in addition to colorectal cancer with the analysis of Oncomine (Figure 8 B). We also investigated the mutational landscape of key genes in colorectal cancer with Cbioportal (Supplementary Figure6).

3. Establishment of a three-gene prognostic signature

Prognostic value of genes in the mRNAsi-related modules in colon and rectal cancer

The colon and rectal cancer patients were divided into training and validation cohorts, and 34 key genes were selected to established a prognostic stemness risk score model using the LASSO algorithm and final (forward and backward elimination methods) multivariate Cox analysis in the training cohort (Figure 9A,

Supplementary Figure7). As shown in Figure 8A, the poly (ADP-ribose) polymerase 1 binding protein (PARPBP), kinetochore-localized astrin/SPAG5 binding protein (KNSTRN) and kinesin family member 2C (KIF2C) were correlated with a poor prognosis in patients from the training cohort according to the multivariate Cox regression analysis. The risk scores were calculated based on the sum of loci β values and the risk coefficient in the risk prediction model with discrete clinical outcomes with regards to OS (Figure 9B). The prognostic index formula for colon and rectal cancer patients in the training cohorts was as follows: Risk score= [Status of PARPBP * (-0.6921)] + [Status of KNSTRN * (1.4204)] + [Status of KIF2C * -0.9696]. This prediction model based on cancer stemness could be a valuable tool for distinguishing among colon and rectal cancer patients. We divided the patients in the training cohorts into a high-risk group and a low-risk group on the basis of the median risk score, which was set as the cut-off value to determine whether the risk of the patient is high or low in the validation cohorts. The association of the risk prediction model and the prognosis is shown in Figures 9. The results of survival analysis proved that the OS of the high-risk group was significantly lower than that of the low-risk according to the Kaplan–Meier curves of the training cohorts (log-rank test, $p<0.01$) (Figures 9C), as well as the Kaplan–Meier curves of the validation cohorts (log-rank test, $p<0.01$) (Figures 9D). The heat maps of three significant genes and the risk scores for each sample in the training and validation cohorts are shown in Figure 9E-H, and PCA was used to reveal a clear boundary between the low- and high-risk groups of colon and rectal cancer patients (Figures 9B). Then, ROC curves were employed to verify the validity of the stemness gene-based prediction model in the training and validation cohorts (Figures 9I and J, Supplementary Figure8). The AUCs were equal to 0.769 at 1 years, 0.683 at 3 years, and 0.728 at 5 years in training group (Figure 9I, Supplementary Figure8). Similarly, the AUCs were equal to 0.685 at 1 years, 0.656 at 3 years, 0.708 at 5 years in validation group (Figure 9J, Supplementary Figure8), which showed that the model could achieve satisfactory predictive accuracy in both the training and validation cohorts.

Validation of the three-gene signature in stage IV colorectal cancer

GSEA was conducted to analyze potential biological characteristics of the three signature genes in colon and rectal cancer patients. As shown in Figure 10A, according to HALLMARK collection defined by MSigDB, the genes in the high-risk group were mainly enriched in cancer stemness-related pathways, such as DNA repair, glycolysis, MYC targets and PI3K-Akt-mTOR signaling. According to GO collection defined by MSigDB, the genes were enriched in functions related to tumor development such as methylation, mitotic nuclear division, signal transduction by P53 class mediators and DNA damage (Figure 10B). According to KEGG collection defined by MSigDB, the genes were enriched in apoptosis, cell cycle, cysteine and methionine metabolism and P53 signaling pathway (Figure 10C). According to the immunological gene set collection defined by MSigDB, multiple immune-function gene sets were enriched in the high-risk group (Figure 10D). To further explore the relationship between the expression of the three

signature genes and the tumor immune microenvironment, we analyzed the corrections between the three genes and 22 types of immune cell infiltration profiles (Supplementary Figure9).

Establishment and validation of the nomogram

A Cox regression model was applied to the training cohort to identify the predictors of OS. Univariate analyses indicated that age, stage-T, stage-N, stage-M, and cancer stemness risk score group were associated with OS in colorectal cancer patients ($p < 0.1$ in all cases) (Table 1). Next, the final (forward and backward elimination methods) multivariate Cox analyses found that age, stage-T, stage-M, and cancer stemness risk score group were independent risk factors for OS (Table 1).

Table 1
Univariable and Multivariable Cox Regression Analysis of OS in CRC Patients

Characteristic (OS)	Univariable analysis		Multivariable analysis	
	HR (95%CI)	P value	HR (95%CI)	P value
Age (years)				
< 60	1		1	
≥ 60	1.44 (0.97–2.11)	0.067	1.68 (1.14–2.50)	0.008
Stage-T				
T1	1		1	
T2	1.19 (0.39–3.63)	0.754	1.53 (0.50–4.67)	0.455
T3	1.44 (0.82–2.53)	0.204	1.28 (0.72–2.30)	0.399
T4	4.35 (2.30–8.22)	< 0.001	2.46 (1.24–4.91)	0.01
Stage-N				
N0	1		1	
N1	1.54 (1.04–2.29)	0.033	1.12 (0.72–1.75)	0.594
N2	2.69 (1.84–3.93)	< 0.001	1.57 (1.01–2.48)	0.042
Stage-M				
M0	1		1	
M1	3.27 (2.26–4.72)	< 0.001	2.41 (1.54–3.75)	< 0.001
Mx	1.69 (1.03–2.73)	0.038	1.69 (0.96–2.73)	0.067
Gene-risk				
Low	1		1	
High	2.43 (1.71–3.45)	< 0.001	2.29 (1.61–3.27)	< 0.001

A nomogram for predicting the 1-, 3- and 5-year OS was established using these independent variables (Fig. 11A). Because age, stage-T, stage-M, and cancer stemness risk score group were predictive for OS in multivariate analysis, these variables were further included in the nomogram. A weighted total score was calculated from these factors, which was applied to predict the 1-, 3- and 5-year OS of colorectal cancer patients.

The nomogram for predicting the 1-, 3- and 5-year OS of colorectal cancer patients was developed based on the multivariate model. The model showed good accuracy for predicting the OS, and internal validation was performed using the training cohort with a C-index of 0.738. Calibration curves for the probability of OS at 1, 3 and 5 years indicated satisfactory consistency between actual observation and

nomogram-predicted OS probabilities in both the training cohort and validation cohort (Fig. 11B, Supplementary Fig. 10). Furthermore, the Decision Curve Analysis (DCA) results of the nomograms also confirmed their clinical applicability for predicting the OS, with superior performance compared to AJCC TNM stage. Thus, the results showed that the nomogram expressed a higher net prognostic benefit than the TNM staging system (Fig. 11C, Supplementary Fig. 10).

Evaluation of the expression and distribution of the three key tumor stemness-related genes and performance of a stemness-related genetic model

Next, we detected the expression and distribution of PARPBP, KNSTRN and KIF2C in colorectal cancer, and performed a multiplex immunofluorescence staining in TMAs (Fig. 12A). We assessed the single index and single index strength score of the three genes in cancer tissue and matched normal tissue samples, which showed that the expression of these genes was higher in cancer tissues based on the single index strength score (Student's two-tailed t-tests, ***$p < 0.001$) (Fig. 12B). To further evaluate the performance of the stemness-related genetic model, we calculated the risk scores based on the single index strength score of the three genes and the prognostic index formula (risk scores = [Status of PARPBP * (-0.6921)] + [Status of KNSTRN * (1.4204)] + [Status of KIF2C * -0.9696]). We classified the patients into high and lowrisk groups according to the risk score (Table 2), and the Kaplan–Meier OS curves of the two groups were significantly different (log-rank test, $p < 0.05$) (Figs. 12C). Additionally, we also established calibration curves for the probability of OS at 1, 3 and 5 years, which indicated satisfactory consistency between actual observation and nomogram-predicted OS probabilities in this cohort (Fig. 12D, Supplementary Fig. 11).

Table 2
Comparisons of Baseline Characteristics of CRC Patients in Low-or High-Risk Group by Risk Scores

Characteristic	Low-Risk (n = 24)	High-Risk (n = 24)	P-value
Age at RC diagnosis, No. (%), years			0.057
< 60	14 (58.3)	20 (83.3)	
≥ 60	10 (41.7)	4 (16.7)	
Sex, No. (%)			0.562
Female	10 (41.7)	12 (50.0)	
Male	14 (58.3)	12 (50.0)	
Tumor location, No. (%)			0.768
Colon	9 (37.5)	10 (41.7)	
Rectal	15 (62.5)	14 (58.3)	
Stage-T, No. (%)			0.033
T1	1 (4.1)	0	
T3	16 (66.7)	10 (41.7)	
T4	7 (29.2)	14 (58.3)	
Stage-N, No. (%)			0.162
N0	4 (16.7)	4 (16.7)	
N1	12 (50.0)	6 (25.0)	
N2	8 (33.3)	14 (58.3)	
Stage-M, No. (%)			0.683
M0	4 (16.7)	3 (12.5)	
M1	20 (83.3)	21 (87.5)	
Tumor grade, No. (%)			0.017
Grade III	19 (79.2)	11 (45.8)	
Grade IV	5 (20.8)	13 (54.2)	

NOTE. P-value was calculated using χ^2 test for categorical variables.

Discussion

Colorectal cancer is a highly aggressive disease, with an urgent need for novel therapeutic strategies and biomarkers for its prognosis and early detection in clinical practice. According to the recently developed cancer stem-cell (CSC) hypothesis, tumor cells are suggested to originate from a stem cell population with self-renewal capacity [8]. These CSCs have been reported to be involved in resistance to cytotoxic conditions, promoting the propagation and recurrence of cancer [40]. Identifying key genes driving the transformation of tumor CSC and underlying biological mechanisms in colorectal cancer may uncover unprecedented therapeutic targets. In this study, we firstly revealed the association between the mRNAsi and prognosis of colorectal cancer patients. The results showed that stemness was significantly higher in colon and rectal cancer tumor tissues than normal tissues containing intestinal stem cells.

Recent studies indicated that cancer CSCs may be a dynamic population continuously influenced by cooperating forces such as microenvironmental, epigenetic and genetic factors [41–43]. The stem cells in normal colonic crypts are continuously and random replaced by other homologous cells, which can provide an advantage for the accumulation of oncogenic mutations through complex stem-cell dynamics [44–45]. Poulsom et al. [46] demonstrated that oncogenic mutations accumulated in stem cells may trigger the rapid development of aggressive sub-clones in colon adenomas. Accumulating evidence suggests that the cancer stem cells can compete with normal stem cells during the early phases of tumor growth, and this competition then continues among neoplastic stem cells during cancer progression, which would be promoted genetic mutations as well as environmental pressures (including radiotherapy and chemotherapy) [47]. We analyzed the mutational landscape in different mRNAsi groups of colon and rectal cancer. The results revealed a higher proportion of APC and TP53 mutant cells in the high-mRNAsi group than in the low-mRNAsi group of colon cancer patients. Previous research showed that mutations of TP53 can shut down its tumor suppressor function, promoting the self-renewal, reprogramming and differentiation of cancer stem cells [48, 49]. Mutant TP53 can inhibit cancer development and promote the acquisition of stem-cell like phenotypes and features to augment the invasion, metastasis and cell death resistance of cancer cells [50].

Analysis of tumor infiltration by different immune cells populations in the high- and low-mRNAsi groups indicated increased infiltration by CD8 + and CD4 + T cells in the high-mRNAsi group of colon cancer. Cancer infiltration by CD8 + T cells may predict higher sensitivity to immunotherapy and better prognosis. However, CD8 + T cells depend on several transcriptional networks to enhance their maintenance and terminal differentiation, as well as the production of stem cell-like central memory cells [51, 52]. A recent study revealed stem-like CD8 + T cell populations that are able to proliferate and produce high levels of checkpoint molecules under stimulation, with the ability to clonally expand to functional effector T cells and self-renew [53–55]. Tumors with high levels of stemness may have higher levels of infiltration by immune cells such as CD8 + and CD4 + T cells, as well as having the potential to produce neoantigens that sensitize them to treatment with immune checkpoint inhibitors. Although TME and stemness were both identified as significant features of cancer in recent years, their covariation across cancers has not been systematically investigated. We found that the TMB is higher in the high-mRNAsi group than in the

low-mRNAsi group of both colon and rectal cancer. The TMB can be defined as the amount of non-synonymous mutations in protein coding regions, which may promote the production of neoantigens by tumor cells [56]. Recent studies established the significance of TMB as a predictive biomarker for the success of treatment with immune checkpoint inhibitors (ICIs) such as anti–programmed cell death (PD)-1 and anti–programmed death-ligand 1 (PD-L1) therapy [57–59]. However, there is accumulating evidence that cancer patients with a high TMB may have worse survival than patients with low TMB in the absence of treatment with ICIs [60, 61]. In this study, we found that tumors with a high mRNAsi may potentially be more easily recognized by the immune system, and therefore more sensitive to treatment with immune checkpoint inhibitors.

M6A, methylation at the N6 position of adenosine, is one of the most abundant and common internal modifications found in long noncoding RNAs (lncRNAs) and mRNAs [62]. It can regulate RNA stability, alternative splicing, decay, abundance, translation and nuclear export [63]. The m6A binding proteins ("readers"), demethylases ("erasers") and methyltransferases ("writers") can remove or add m6A sites, and can therefore participate in regulating every aspect of RNA metabolism [64]. Our results show that a number of m6A RNA methylation regulators were expressed at higher levels in the high-mRNAsi groups of both colon and rectal cancer patients. YTHDF1, a "reader" of m6A methylation, plays a significant role in regulating the colon cancer stem cell-like features via the Wnt/β-catenin pathway and regulating the translation of selectively recognized mRNAs [65]. Furthermore, ALKBH5 was found to promote tumor formation and increase cancer cell stemness by regulating mRNA stability of pluripotency factors via m6A demethylation in hypoxic environments [66]. Additionally, FTO plays a significant role in immune evasion and self-renewal of cancer stem cells, and FTO upregulation can protect cancer cells from T cell cytotoxicity [67].

To further explore the prognostic value and biological mechanisms of potential therapeutic targets, we built a cancer stemness-related prognostic model to provide novel insights into treatment options for colon and rectal cancer. The prognostic index is based on the fractions of three genes identified among the cancer stemness-related key module genes, which need to be further investigated to uncover their potential molecular biological mechanisms. PARPBP, a significant inhibitor of homologous recombination (HR), has been demonstrated to be related to increased AFP levels, proliferation, differentiation, as well as poor prognosis of lung and hepatocellular carcinoma patients [68, 69]. PARPBP was shown to inhibit the formation of DNA double-strand break repair protein (RAD51)-DNA structures, and thereby decrease the recombination of replicating chromosomes, promoting genomic instability and DNA damage hypersensitivity in pancreatic cancer [70]. PARPBP promotes tumor cell migration and invasion by enhancing mutagenic replication, extravasation, anoikis resistance and self-renewal in lung cancer [71]. KNSTRN, an important component of the mitotic spindle, was found to regulate chromosome segregation and anaphase onset during mitosis in cancer cells [72]. Furthermore, recent studies demonstrated that accumulation of KNSTRN mutations may be an early event in cancer development that accelerates tumor growth in cutaneous squamous cell carcinoma and melanoma [73, 74]. KIF2C is an important regulator of chromosome segregation, bipolar spindle formation and microtubule depolymerization during mitosis, and it may be related with poor prognosis in non-small cell lung cancer [75]. Additionally, KIF2C was

found to inhibit apoptosis and cellular senescence, promoting the proliferation and cell cycle progression of lung and hepatocellular cancer cells via a P53-depended pathway [76–78]. Moreover, KIF2C was shown to act as a tumor antigen that can elicit spontaneous and frequent CD41 T cell responses of the Th1-type in colorectal cancer, in a process that is influenced by peripheral T regulatory cells [79].

Furthermore, we established a nomogram to better predict the survival of colon and rectal cancer patients and visualize the prediction results, which can further improve the compliance and therapeutic efficacy for patients. Additionally, we compared the prognostic accuracy of the TNM stage and nomogram model with DCA, which showed that the nomogram model consisting of stemness-related genes and clinical phenotype could have a higher prognostic ability than the TNM stage. The present results suggest that the model based on three stemness-related genes may have reliable prognostic accuracy in conjunction with the clinical phenotype. However, some of data in this retrospective analysis released in publicly available datasets may be limited, and the incomplete clinicopathological information may cause potential bias that would influence the evaluation of the prognostic ability. Data from TCGA are from western countries, and all of the datasets consisted of a mutational landscape and transcriptome, which may hinder the clinical translation and generalization of the prognostic model. Consequently, we performed multiplex immunofluorescence staining in TMAs of colorectal cancer patients from China, and verified the prognostic ability of the nomogram model in an advanced colorectal cancer cohort. As shown in Fig. 2, stage colon cancer patients with higher mRNAsi values had a lower apparent survival probability than those with lower mRNAsi values, and the difference was statistically significant, which suggested a correlation between CSCs and advanced tumor. The evaluation of CSCs-related genes may be a prognostic marker for selecting patients at high-risk of metastasis from CRC, that are likely to benefit from treatment. Immunofluorescence results could be employed to evaluate the expression and distribution of the three key stemness-related genes identified by proteomics, with the potential to make this model more convenient and reliable in clinical practice.

Several limitations in this study should be kept in mind. The stemness-related genes and corresponding downstream signaling pathways need further functional validation, since lacking of in vitro or in vivo experiments greatly reduces the reliability of present research. Although the risk score and nomogram based on the multivariate Cox proportional hazards regression analysis were used widely to evaluate significant factors related with clinical prognosis, some machine learning algorithms might have a better prognostic ability. Prominent examples of such algorithms include Naïve Bayes, Random Forest and Decision Tree, which we will test in our next stage of research. Many molecular mechanisms contributing to cancer stemness remain to be explored in the future, and will help elucidate the development of colorectal cancer.

Conclusions

Taken together, our study highlights a robust correlation between the level of cancer stemness and traits related to tumor heterogeneity, including the immune microenvironment, TMB, and the expression of m6A RNA methylation regulatory factors in colorectal cancer cells. The prognostic signature based on mRNAsi

may contribute to personalized prognosis of clinical outcomes in colorectal cancer and act as a potential prognostic biomarker for responses to differentiation therapies in clinical practice. The proposed stemness-related genetic model could provide great assistance in formulating efficient therapeutic strategies for the personalized management of colorectal cancer.

Abbreviations

CSC: Cancer stem cell; TME: Tumor microenvironment; MHC: Major histocompatibility complex; OCLR: One-Class Logistic Regression; CIBERSORT: Cell type Identification by Estimating Relative Subsets of RNA Transcripts; ESTIMATE: Estimation of Stromal and Immune cells in Malignant Tumors using Expression data; m6A: N6-methyladenosine; TCGA: The Cancer Genome Atlas; SNV: Single nucleotide variants; RNA-seq: RNA-sequencing; DEGs: Differentially expressed genes; WGCNA: Weighted Gene Co-expression Network Analysis; TOM: Topological overlap matrix; MM: Module membership; GS: Gene significance; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene ontology; TMAs: Tissue microarrays; OS: Overall survival; ROC: Receiver operating characteristic; AUC: Area under the curve; lncRNA: Long noncoding RNA

Declarations

Acknowledgements

The authors would appreciate all the members of the Department of Colorectal Surgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College.

Authors’ contributions

Conception and design: WR, QJC, JZ and WXS

Collection and assembly of data: WR, QJC, JZ and WXS

Data analysis and interpretation: WR, QJC and LSF

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

Funding

None

Availability of data and materials

Not applicable

Ethics approval and consent to participate

This study was approved by Institutional Review Board of National Cancer Center (NCC), Chinese Academy of Medical Sciences (approval no. 2017–20). All the patients' samples used in this study have obtained necessary consent.

Competing interests

The authors declare no conflict of interest.

Consent for publication

Written informed consent for publication was obtained from all the participants.

References

1. Sanoff HK, Bleiberg H, Goldberg RM. Managing older patients with colorectal cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25(14):1891-7. Epub 2007/05/10. doi: 10.1200/jco.2006.10.1220. PubMed PMID: 17488988.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2018;68(6):394-424. Epub 2018/09/13. doi: 10.3322/caac.21492. PubMed PMID: 30207593.
3. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. CA: a cancer journal for clinicians. 2016;66(2):115-32. Epub 2016/01/26. doi: 10.3322/caac.21338. PubMed PMID: 26808342.
4. Park IJ, You YN, Agarwal A, Skibber JM, Rodriguez-Bigas MA, Eng C, et al. Neoadjuvant treatment response as an early response indicator for patients with rectal cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2012;30(15):1770-6. Epub 2012/04/12. doi: 10.1200/jco.2011.39.7901. PubMed PMID: 22493423; PubMed Central PMCID: PMCPMC3383178 found at the end of this article.
5. Zeuner A, Todaro M, Stassi G, De Maria R. Colorectal cancer stem cells: from the crypt to the clinic. Cell stem cell. 2014;15(6):692-705. Epub 2014/12/07. doi: 10.1016/j.stem.2014.11.012. PubMed PMID: 25479747.

6. Lenos KJ, Miedema DM, Lodestijn SC, Nijman LE, van den Bosch T, Romero Ros X, et al. Stem cell functionality is microenvironmentally defined during tumour expansion and therapy response in colon cancer. Nature cell biology. 2018;20(10):1193-202. Epub 2018/09/05. doi: 10.1038/s41556-018-0179-z. PubMed PMID: 30177776; PubMed Central PMCID: PMCPMC6163039.
7. Kreso A, Dick JE. Evolution of the cancer stem cell model. Cell stem cell. 2014;14(3):275-91. Epub 2014/03/13. doi: 10.1016/j.stem.2014.02.006. PubMed PMID: 24607403.
8. Wicha MS, Liu S, Dontu G. Cancer stem cells: an old idea--a paradigm shift. Cancer research. 2006;66(4):1883-90; discussion 95-6. Epub 2006/02/21. doi: 10.1158/0008-5472.can-05-3153. PubMed PMID: 16488983.
9. Ricci-Vitiani L, Lombardi DG, Pilozzi E, Biffoni M, Todaro M, Peschle C, et al. Identification and expansion of human colon-cancer-initiating cells. Nature. 2007;445(7123):111-5. Epub 2006/11/24. doi: 10.1038/nature05384. PubMed PMID: 17122771.
10. Choi D, Lee HW, Hur KY, Kim JJ, Park GS, Jang SH, et al. Cancer stem cell markers CD133 and CD24 correlate with invasiveness and differentiation in colorectal adenocarcinoma. World journal of gastroenterology. 2009;15(18):2258-64. Epub 2009/05/14. doi: 10.3748/wjg.15.2258. PubMed PMID: 19437567; PubMed Central PMCID: PMCPMC2682242.
11. Michieli P, Mazzone M, Basilico C, Cavassa S, Sottile A, Naldini L, et al. Targeting the tumor and its microenvironment by a dual-function decoy Met receptor. Cancer cell. 2004;6(1):61-73. Epub 2004/07/21. doi: 10.1016/j.ccr.2004.05.032. PubMed PMID: 15261142.
12. Cheng C, Yaffe MB, Sharp PA. A positive feedback loop couples Ras activation and CD44 alternative splicing. Genes & development. 2006;20(13):1715-20. Epub 2006/07/05. doi: 10.1101/gad.1430906. PubMed PMID: 16818603; PubMed Central PMCID: PMCPMC1522067.
13. Li Z. CD133: a stem cell biomarker and beyond. Experimental hematology & oncology. 2013;2(1):17. Epub 2013/07/03. doi: 10.1186/2162-3619-2-17. PubMed PMID: 23815814; PubMed Central PMCID: PMCPMC3701589.
14. MacArthur BD. Collective dynamics of stem cell populations. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(10):3653-4. Epub 2014/03/05. doi: 10.1073/pnas.1401030111. PubMed PMID: 24591612; PubMed Central PMCID: PMCPMC3956180.
15. Baharom F, Ramirez-Valdez RA, Tobin KKS, Yamane H, Dutertre CA, Khalilnezhad A, et al. Intravenous nanoparticle vaccination generates stem-like TCF1(+) neoantigen-specific CD8(+) T cells. Nature immunology. 2020. Epub 2020/11/04. doi: 10.1038/s41590-020-00810-3. PubMed PMID: 33139915.
16. Condeelis J, Pollard JW. Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. Cell. 2006;124(2):263-6. Epub 2006/01/28. doi: 10.1016/j.cell.2006.01.007. PubMed PMID: 16439202.
17. Sultan M, Coyle KM, Vidovic D, Thomas ML, Gujar S, Marcato P. Hide-and-seek: the interplay between cancer stem cells and the immune system. Carcinogenesis. 2017;38(2):107-18. Epub 2016/11/21. doi: 10.1093/carcin/bgw115. PubMed PMID: 27866156.

18. Sokolov A, Paull EO, Stuart JM. ONE-CLASS DETECTION OF CELL STATES IN TUMOR SUBTYPES. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing. 2016;21:405-16. Epub 2016/01/19. PubMed PMID: 26776204; PubMed Central PMCID: PMCPMC4856035.
19. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. Cell. 2018;173(2):338-54.e15. Epub 2018/04/07. doi: 10.1016/j.cell.2018.03.034. PubMed PMID: 29625051; PubMed Central PMCID: PMCPMC5902191.
20. Zeng D, Zhou R, Yu Y, Luo Y, Zhang J, Sun H, et al. Gene expression profiles for a prognostic immunoscore in gastric cancer. The British journal of surgery. 2018;105(10):1338-48. Epub 2018/04/25. doi: 10.1002/bjs.10871. PubMed PMID: 29691839; PubMed Central PMCID: PMCPMC6099214.
21. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nature communications. 2013;4:2612. Epub 2013/10/12. doi: 10.1038/ncomms3612. PubMed PMID: 24113773; PubMed Central PMCID: PMCPMC3826632.
22. Song K, Li L, Zhang G. Bias and Correction in RNA-seq Data for Marine Species. Marine biotechnology (New York, NY). 2017;19(5):541-50. Epub 2017/09/09. doi: 10.1007/s10126-017-9773-5. PubMed PMID: 28884399.
23. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research. 2015;43(7):e47. Epub 2015/01/22. doi: 10.1093/nar/gkv007. PubMed PMID: 25605792; PubMed Central PMCID: PMCPMC4402510.
24. Amarasinghe KC, Li J, Hunter SM, Ryland GL, Cowin PA, Campbell IG, et al. Inferring copy number and genotype in tumour exome data. BMC genomics. 2014;15(1):732. Epub 2014/08/30. doi: 10.1186/1471-2164-15-732. PubMed PMID: 25167919; PubMed Central PMCID: PMCPMC4162913.
25. He Y, Jiang Z, Chen C, Wang X. Classification of triple-negative breast cancers based on Immunogenomic profiling. Journal of experimental & clinical cancer research : CR. 2018;37(1):327. Epub 2018/12/31. doi: 10.1186/s13046-018-1002-1. PubMed PMID: 30594216; PubMed Central PMCID: PMCPMC6310928.
26. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature. 2009;462(7269):108-12. Epub 2009/10/23. doi: 10.1038/nature08460. PubMed PMID: 19847166; PubMed Central PMCID: PMCPMC2783335.
27. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC bioinformatics. 2013;14:7. Epub 2013/01/18. doi: 10.1186/1471-2105-14-7. PubMed PMID: 23323831; PubMed Central PMCID: PMCPMC3618321.
28. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nature methods. 2015;12(5):453-7. Epub 2015/03/31. doi:

10.1038/nmeth.3337. PubMed PMID: 25822800; PubMed Central PMCID: PMCPMC4739640.

29. Chai RC, Wu F, Wang QX, Zhang S, Zhang KN, Liu YQ, et al. m(6)A RNA methylation regulators contribute to malignant progression and have clinical prognostic impact in gliomas. Aging. 2019;11(4):1204-25. Epub 2019/02/28. doi: 10.18632/aging.101829. PubMed PMID: 30810537; PubMed Central PMCID: PMCPMC6402513.
30. Yang Y, Hsu PJ, Chen YS, Yang YG. Dynamic transcriptomic m(6)A decoration: writers, erasers, readers and functions in RNA metabolism. Cell research. 2018;28(6):616-24. Epub 2018/05/24. doi: 10.1038/s41422-018-0040-8. PubMed PMID: 29789545; PubMed Central PMCID: PMCPMC5993786.
31. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics. 2008;9:559. Epub 2008/12/31. doi: 10.1186/1471-2105-9-559. PubMed PMID: 19114008; PubMed Central PMCID: PMCPMC2631488.
32. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics : a journal of integrative biology. 2012;16(5):284-7. Epub 2012/03/30. doi: 10.1089/omi.2011.0118. PubMed PMID: 22455463; PubMed Central PMCID: PMCPMC3339379.
33. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic acids research. 2019;47(D1):D607-d13. Epub 2018/11/27. doi: 10.1093/nar/gky1131. PubMed PMID: 30476243; PubMed Central PMCID: PMCPMC6323986.
34. Goeman JJ. L1 penalized estimation in the Cox proportional hazards model. Biometrical journal Biometrische Zeitschrift. 2010;52(1):70-84. Epub 2009/11/26. doi: 10.1002/bimj.200900028. PubMed PMID: 19937997.
35. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics. 2000;56(2):337-44. Epub 2000/07/06. doi: 10.1111/j.0006-341x.2000.00337.x. PubMed PMID: 10877287.
36. Pei J, Wang Y, Li Y. Identification of key genes controlling breast cancer stem cell characteristics via stemness indices analysis. Journal of translational medicine. 2020;18(1):74. Epub 2020/02/14. doi: 10.1186/s12967-020-02260-9. PubMed PMID: 32050983; PubMed Central PMCID: PMCPMC7014665.
37. Bai KH, He SY, Shu LL, Wang WD, Lin SY, Zhang QY, et al. Identification of cancer stem cell characteristics in liver hepatocellular carcinoma by WGCNA analysis of transcriptome stemness index. Cancer medicine. 2020;9(12):4290-8. Epub 2020/04/21. doi: 10.1002/cam4.3047. PubMed PMID: 32311840; PubMed Central PMCID: PMCPMC7300398.
38. Cui Q, Shi H, Ye P, Li L, Qu Q, Sun G, et al. m(6)A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. Cell reports. 2017;18(11):2622-34. Epub 2017/03/16. doi: 10.1016/j.celrep.2017.02.059. PubMed PMID: 28297667; PubMed Central PMCID: PMCPMC5479356.
39. Dai D, Wang H, Zhu L, Jin H, Wang X. N6-methyladenosine links RNA metabolism to cancer progression. Cell death & disease. 2018;9(2):124. Epub 2018/01/28. doi: 10.1038/s41419-017-0129-

x. PubMed PMID: 29374143; PubMed Central PMCID: PMCPMC5833385.

40. Aghaalikhani N, Rashtchizadeh N, Shadpour P, Allameh A, Mahmoodi M. Cancer stem cells as a therapeutic target in bladder cancer. Journal of cellular physiology. 2019;234(4):3197-206. Epub 2018/11/25. doi: 10.1002/jcp.26916. PubMed PMID: 30471107.
41. Lei J, Levin SA, Nie Q. Mathematical model of adult stem cell regeneration with cross-talk between genetic and epigenetic regulation. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(10):E880-7. Epub 2014/02/07. doi: 10.1073/pnas.1324267111. PubMed PMID: 24501127; PubMed Central PMCID: PMCPMC3956188.
42. MacArthur BD, Lemischka IR. Statistical mechanics of pluripotency. Cell. 2013;154(3):484-9. Epub 2013/08/06. doi: 10.1016/j.cell.2013.07.024. PubMed PMID: 23911316.
43. Meacham CE, Morrison SJ. Tumour heterogeneity and cancer cell plasticity. Nature. 2013;501(7467):328-37. Epub 2013/09/21. doi: 10.1038/nature12624. PubMed PMID: 24048065; PubMed Central PMCID: PMCPMC4521623.
44. Snippert HJ, Schepers AG, van Es JH, Simons BD, Clevers H. Biased competition between Lgr5 intestinal stem cells driven by oncogenic mutation induces clonal expansion. EMBO reports. 2014;15(1):62-9. Epub 2013/12/21. doi: 10.1002/embr.201337799. PubMed PMID: 24355609; PubMed Central PMCID: PMCPMC3983678.
45. Song X, Gao H, Lin Y, Yao Y, Zhu S, Wang J, et al. Alterations in the microbiota drive interleukin-17C production from intestinal epithelial cells to promote tumorigenesis. Immunity. 2014;40(1):140-52. Epub 2014/01/15. doi: 10.1016/j.immuni.2013.11.018. PubMed PMID: 24412611.
46. Myant KB, Cammareri P, McGhee EJ, Ridgway RA, Huels DJ, Cordero JB, et al. ROS production and NF-κB activation triggered by RAC1 facilitate WNT-driven intestinal stem cell proliferation and colorectal cancer initiation. Cell stem cell. 2013;12(6):761-73. Epub 2013/05/15. doi: 10.1016/j.stem.2013.04.006. PubMed PMID: 23665120; PubMed Central PMCID: PMCPMC3690525.
47. Humphries A, Cereser B, Gay LJ, Miller DS, Das B, Gutteridge A, et al. Lineage tracing reveals multipotent stem cells maintain human adenomas and the pattern of clonal expansion in tumor evolution. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(27):E2490-9. Epub 2013/06/15. doi: 10.1073/pnas.1220353110. PubMed PMID: 23766371; PubMed Central PMCID: PMCPMC3704042.
48. Emmink BL, Verheem A, Van Houdt WJ, Steller EJ, Govaert KM, Pham TV, et al. The secretome of colon cancer stem cells contains drug-metabolizing enzymes. Journal of proteomics. 2013;91:84-96. Epub 2013/07/10. doi: 10.1016/j.jprot.2013.06.027. PubMed PMID: 23835434.
49. Brosh R, Rotter V. When mutants gain new powers: news from the mutant p53 field. Nature reviews Cancer. 2009;9(10):701-13. Epub 2009/08/21. doi: 10.1038/nrc2693. PubMed PMID: 19693097.
50. Muller PA, Vousden KH. Mutant p53 in cancer: new functions and therapeutic opportunities. Cancer cell. 2014;25(3):304-17. Epub 2014/03/22. doi: 10.1016/j.ccr.2014.01.021. PubMed PMID: 24651012; PubMed Central PMCID: PMCPMC3970583.

51. Solomon H, Madar S, Rotter V. Mutant p53 gain of function is interwoven into the hallmarks of cancer. The Journal of pathology. 2011;225(4):475-8. Epub 2011/10/26. doi: 10.1002/path.2988. PubMed PMID: 22025211.
52. Aloni-Grinstein R, Shetzer Y, Kaufman T, Rotter V. p53: the barrier to cancer stem cell formation. FEBS letters. 2014;588(16):2580-9. Epub 2014/02/25. doi: 10.1016/j.febslet.2014.02.011. PubMed PMID: 24560790.
53. Skoulidis F, Goldberg ME, Greenawalt DM, Hellmann MD, Awad MM, Gainor JF, et al. STK11/LKB1 Mutations and PD-1 Inhibitor Resistance in KRAS-Mutant Lung Adenocarcinoma. Cancer discovery. 2018;8(7):822-35. Epub 2018/05/19. doi: 10.1158/2159-8290.cd-18-0099. PubMed PMID: 29773717; PubMed Central PMCID: PMCPMC6030433.
54. Rizvi H, Sanchez-Vega F, La K, Chatila W, Jonsson P, Halpenny D, et al. Molecular Determinants of Response to Anti-Programmed Cell Death (PD)-1 and Anti-Programmed Death-Ligand 1 (PD-L1) Blockade in Patients With Non-Small-Cell Lung Cancer Profiled With Targeted Next-Generation Sequencing. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2018;36(7):633-41. Epub 2018/01/18. doi: 10.1200/jco.2017.75.3384. PubMed PMID: 29337640; PubMed Central PMCID: PMCPMC6075848.
55. Wu X, Gu Z, Chen Y, Chen B, Chen W, Weng L, et al. Application of PD-1 Blockade in Cancer Immunotherapy. Computational and structural biotechnology journal. 2019;17:661-74. Epub 2019/06/18. doi: 10.1016/j.csbj.2019.03.006. PubMed PMID: 31205619; PubMed Central PMCID: PMCPMC6558092.
56. Pai SG, Carneiro BA, Chae YK, Costa RL, Kalyan A, Shah HA, et al. Correlation of tumor mutational burden and treatment outcomes in patients with colorectal cancer. Journal of gastrointestinal oncology. 2017;8(5):858-66. Epub 2017/12/01. doi: 10.21037/jgo.2017.06.20. PubMed PMID: 29184690; PubMed Central PMCID: PMCPMC5674261.
57. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science (New York, NY). 2015;348(6230):124-8. Epub 2015/03/15. doi: 10.1126/science.aaa1348. PubMed PMID: 25765070; PubMed Central PMCID: PMCPMC4993154.
58. Rimm DL, Han G, Taube JM, Yi ES, Bridge JA, Flieder DB, et al. A Prospective, Multi-institutional, Pathologist-Based Assessment of 4 Immunohistochemistry Assays for PD-L1 Expression in Non-Small Cell Lung Cancer. JAMA oncology. 2017;3(8):1051-8. Epub 2017/03/10. doi: 10.1001/jamaoncol.2017.0013. PubMed PMID: 28278348; PubMed Central PMCID: PMCPMC5650234.
59. Jansen CS, Prokhnevska N, Master VA, Sanda MG, Carlisle JW, Bilen MA, et al. An intra-tumoral niche maintains and differentiates stem-like CD8 T cells. Nature. 2019;576(7787):465-70. Epub 2019/12/13. doi: 10.1038/s41586-019-1836-5. PubMed PMID: 31827286; PubMed Central PMCID: PMCPMC7108171.

60. Jadhav RR, Im SJ, Hu B, Hashimoto M, Li P, Lin JX, et al. Epigenetic signature of PD-1+ TCF1+ CD8 T cells that act as resource cells during chronic viral infection and respond to PD-1 blockade. Proceedings of the National Academy of Sciences of the United States of America. 2019;116(28):14113-8. Epub 2019/06/23. doi: 10.1073/pnas.1903520116. PubMed PMID: 31227606; PubMed Central PMCID: PMCPMC6628832.
61. Zhang JY, Zhao YL, Lv YP, Cheng P, Chen W, Duan M, et al. Modulation of CD8(+) memory stem T cell activity and glycogen synthase kinase 3β inhibition enhances anti-tumoral immunity in gastric cancer. Oncoimmunology. 2018;7(4):e1412900. Epub 2018/04/11. doi: 10.1080/2162402x.2017.1412900. PubMed PMID: 29632726; PubMed Central PMCID: PMCPMC5889281.
62. Lugli E, Galletti G, Boi SK, Youngblood BA. Stem, Effector, and Hybrid States of Memory CD8(+) T Cells. Trends in immunology. 2020;41(1):17-28. Epub 2019/12/08. doi: 10.1016/j.it.2019.11.004. PubMed PMID: 31810790; PubMed Central PMCID: PMCPMC6934921.
63. Radpour R, Riether C, Simillion C, Höpner S, Bruggmann R, Ochsenbein AF. CD8(+) T cells expand stem and progenitor cells in favorable but not adverse risk acute myeloid leukemia. Leukemia. 2019;33(10):2379-92. Epub 2019/03/17. doi: 10.1038/s41375-019-0441-9. PubMed PMID: 30877275.
64. Zhao X, Cui L. Development and validation of a m(6)A RNA methylation regulators-based signature for predicting the prognosis of head and neck squamous cell carcinoma. American journal of cancer research. 2019;9(10):2156-69. Epub 2019/11/14. PubMed PMID: 31720080; PubMed Central PMCID: PMCPMC6834477.
65. Geng Y, Guan R, Hong W, Huang B, Liu P, Guo X, et al. Identification of m6A-related genes and m6A RNA methylation regulators in pancreatic cancer and their association with survival. Annals of translational medicine. 2020;8(6):387. Epub 2020/05/02. doi: 10.21037/atm.2020.03.98. PubMed PMID: 32355831; PubMed Central PMCID: PMCPMC7186697.
66. Zhang C, Samanta D, Lu H, Bullen JW, Zhang H, Chen I, et al. Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m□A-demethylation of NANOG mRNA. Proceedings of the National Academy of Sciences of the United States of America. 2016;113(14):E2047-56. Epub 2016/03/24. doi: 10.1073/pnas.1602883113. PubMed PMID: 27001847; PubMed Central PMCID: PMCPMC4833258.
67. Bai Y, Yang C, Wu R, Huang L, Song S, Li W, et al. YTHDF1 Regulates Tumorigenicity and Cancer Stem Cell-Like Activity in Human Colorectal Carcinoma. Frontiers in oncology. 2019;9:332. Epub 2019/05/28. doi: 10.3389/fonc.2019.00332. PubMed PMID: 31131257; PubMed Central PMCID: PMCPMC6509179.
68. Yan Y, Liu F, Han L, Zhao L, Chen J, Olopade OI, et al. HIF-2α promotes conversion to a stem cell phenotype and induces chemoresistance in breast cancer cells by activating Wnt and Notch pathways. Journal of experimental & clinical cancer research : CR. 2018;37(1):256. Epub 2018/10/21. doi: 10.1186/s13046-018-0925-x. PubMed PMID: 30340507; PubMed Central PMCID: PMCPMC6194720.

69. Su R, Dong L, Li Y, Gao M, Han L, Wunderlich M, et al. Targeting FTO Suppresses Cancer Stem Cell Maintenance and Immune Evasion. Cancer cell. 2020;38(1):79-96.e11. Epub 2020/06/13. doi: 10.1016/j.ccell.2020.04.017. PubMed PMID: 32531268; PubMed Central PMCID: PMCPMC7363590.
70. Yu B, Ding Y, Liao X, Wang C, Wang B, Chen X. Overexpression of PARPBP Correlates with Tumor Progression and Poor Prognosis in Hepatocellular Carcinoma. Digestive diseases and sciences. 2019;64(10):2878-92. Epub 2019/04/06. doi: 10.1007/s10620-019-05608-4. PubMed PMID: 30949905.
71. Xu D, Tao Z, Tang X, He JK. Poly (ADP-ribose) polymerase-1 Binding Protein Facilitates Lung Adenocarcinoma Cell Proliferation and Correlates with Poor Prognosis. Annals of clinical and laboratory science. 2019;49(5):574-80. Epub 2019/10/16. PubMed PMID: 31611199.
72. O'Connor KW, Dejsuphong D, Park E, Nicolae CM, Kimmelman AC, D'Andrea AD, et al. PARI overexpression promotes genomic instability and pancreatic tumorigenesis. Cancer research. 2013;73(8):2529-39. Epub 2013/02/26. doi: 10.1158/0008-5472.can-12-3313. PubMed PMID: 23436799; PubMed Central PMCID: PMCPMC3630264.
73. Choi EB, Yang AY, Kim SC, Lee J, Choi JK, Choi C, et al. PARP1 enhances lung adenocarcinoma metastasis by novel mechanisms independent of DNA repair. Oncogene. 2016;35(35):4569-79. Epub 2016/02/24. doi: 10.1038/onc.2016.3. PubMed PMID: 26898760.
74. Schmitz L, Grinblat B, Novak B, Hoeh AK, Händschke K, von Dobbeler C, et al. Somatic mutations in kinetochore gene KNSTRN are associated with basal proliferating actinic keratoses and cutaneous squamous cell carcinoma. Journal of the European Academy of Dermatology and Venereology : JEADV. 2019;33(8):1535-40. Epub 2019/04/12. doi: 10.1111/jdv.15615. PubMed PMID: 30972880.
75. Lee CS, Bhaduri A, Mah A, Johnson WL, Ungewickell A, Aros CJ, et al. Recurrent point mutations in the kinetochore gene KNSTRN in cutaneous squamous cell carcinoma. Nature genetics. 2014;46(10):1060-2. Epub 2014/09/10. doi: 10.1038/ng.3091. PubMed PMID: 25194279; PubMed Central PMCID: PMCPMC4324615.
76. Jaju PD, Nguyen CB, Mah AM, Atwood SX, Li J, Zia A, et al. Mutations in the Kinetochore Gene KNSTRN in Basal Cell Carcinoma. The Journal of investigative dermatology. 2015;135(12):3197-200. Epub 2015/09/09. doi: 10.1038/jid.2015.339. PubMed PMID: 26348826; PubMed Central PMCID: PMCPMC4747638.
77. Gan H, Lin L, Hu N, Yang Y, Gao Y, Pei Y, et al. KIF2C exerts an oncogenic role in nonsmall cell lung cancer and is negatively regulated by miR-325-3p. Cell biochemistry and function. 2019;37(6):424-31. Epub 2019/07/23. doi: 10.1002/cbf.3420. PubMed PMID: 31328811.
78. Chen J, Li S, Zhou S, Cao S, Lou Y, Shen H, et al. Kinesin superfamily protein expression and its association with progression and prognosis in hepatocellular carcinoma. Journal of cancer research and therapeutics. 2017;13(4):651-9. Epub 2017/09/14. doi: 10.4103/jcrt.JCRT_491_17. PubMed PMID: 28901309.
79. Gwon MR, Cho JH, Kim JR. Mitotic centromere-associated kinase (MCAK/Kif2C) regulates cellular senescence in human primary cells through a p53-dependent pathway. FEBS letters.

2012;586(23):4148-56. Epub 2012/10/27. doi: 10.1016/j.febslet.2012.10.012. PubMed PMID: 23098759.

Figures

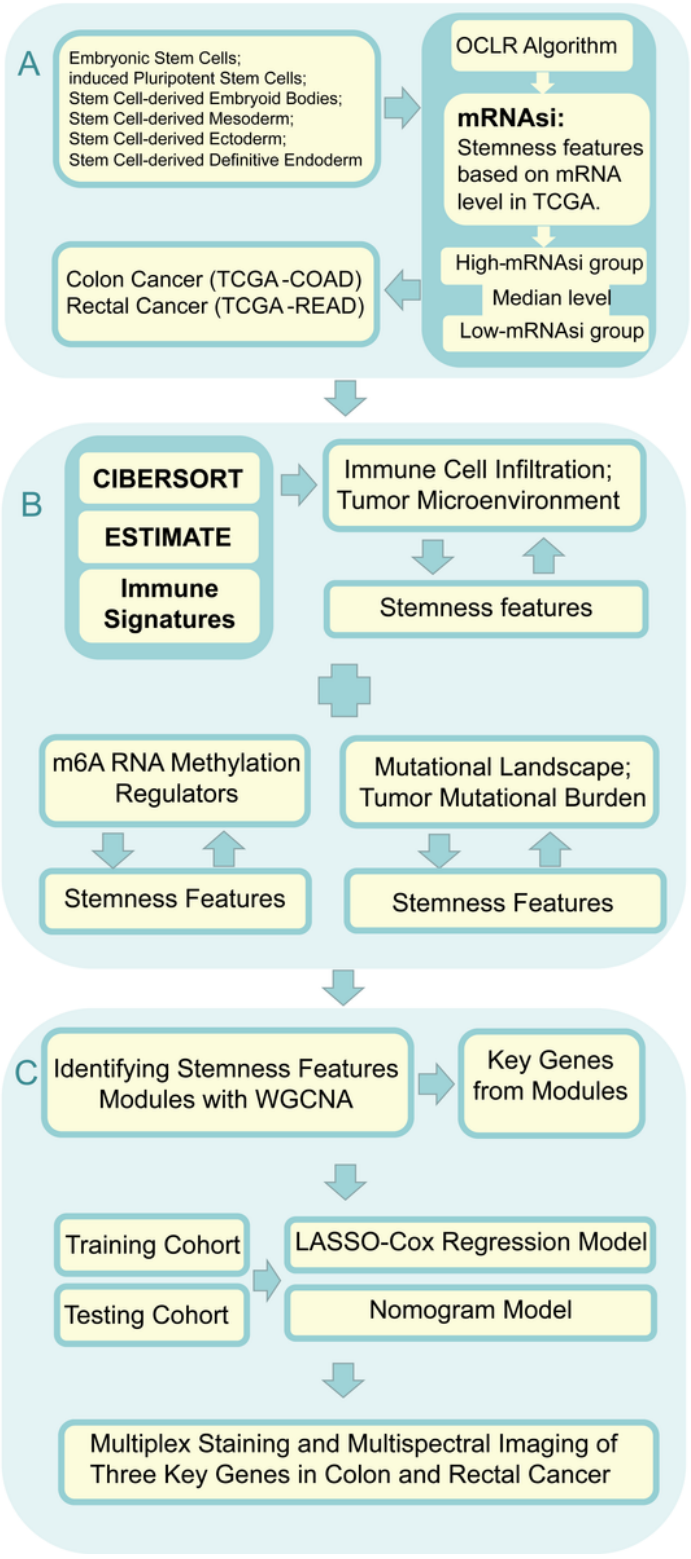

Figure 1

The analysis process of this study.

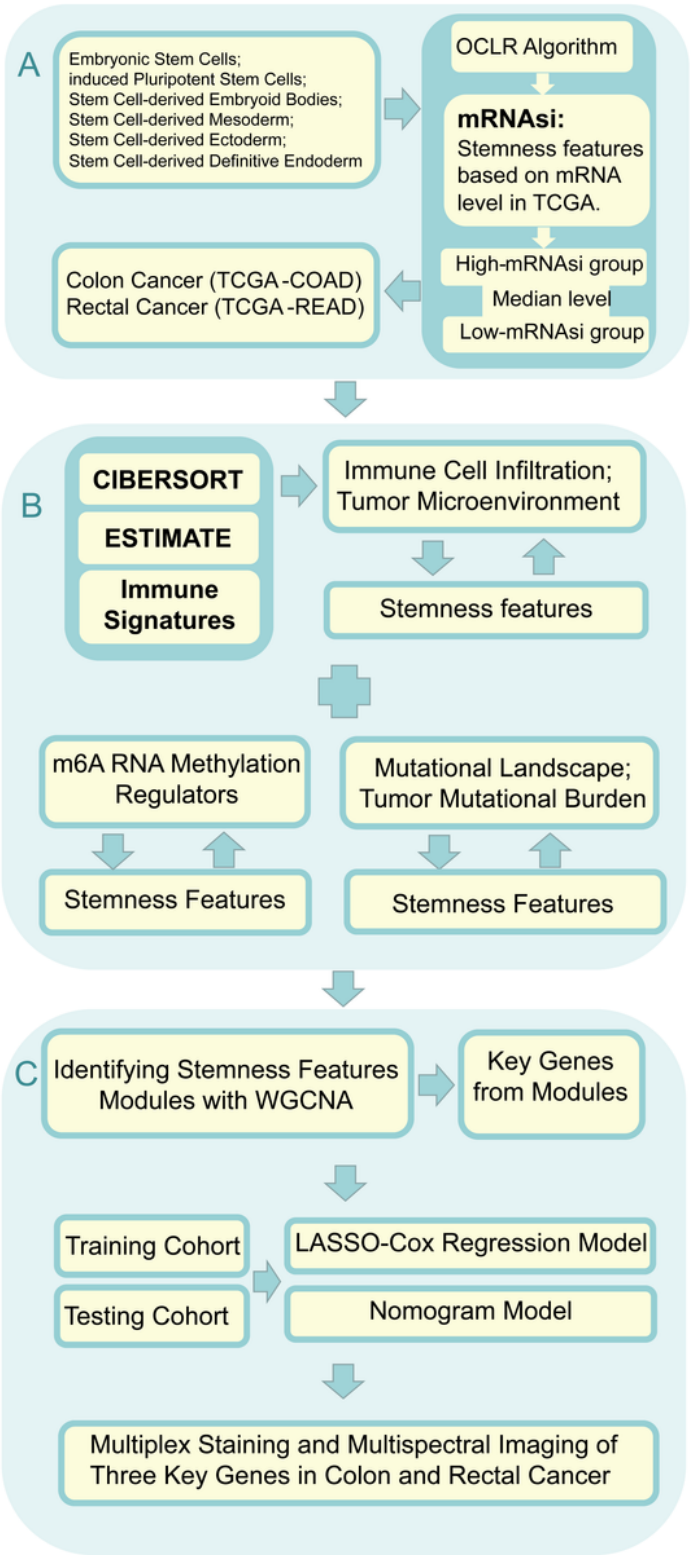

Figure 1

The analysis process of this study.

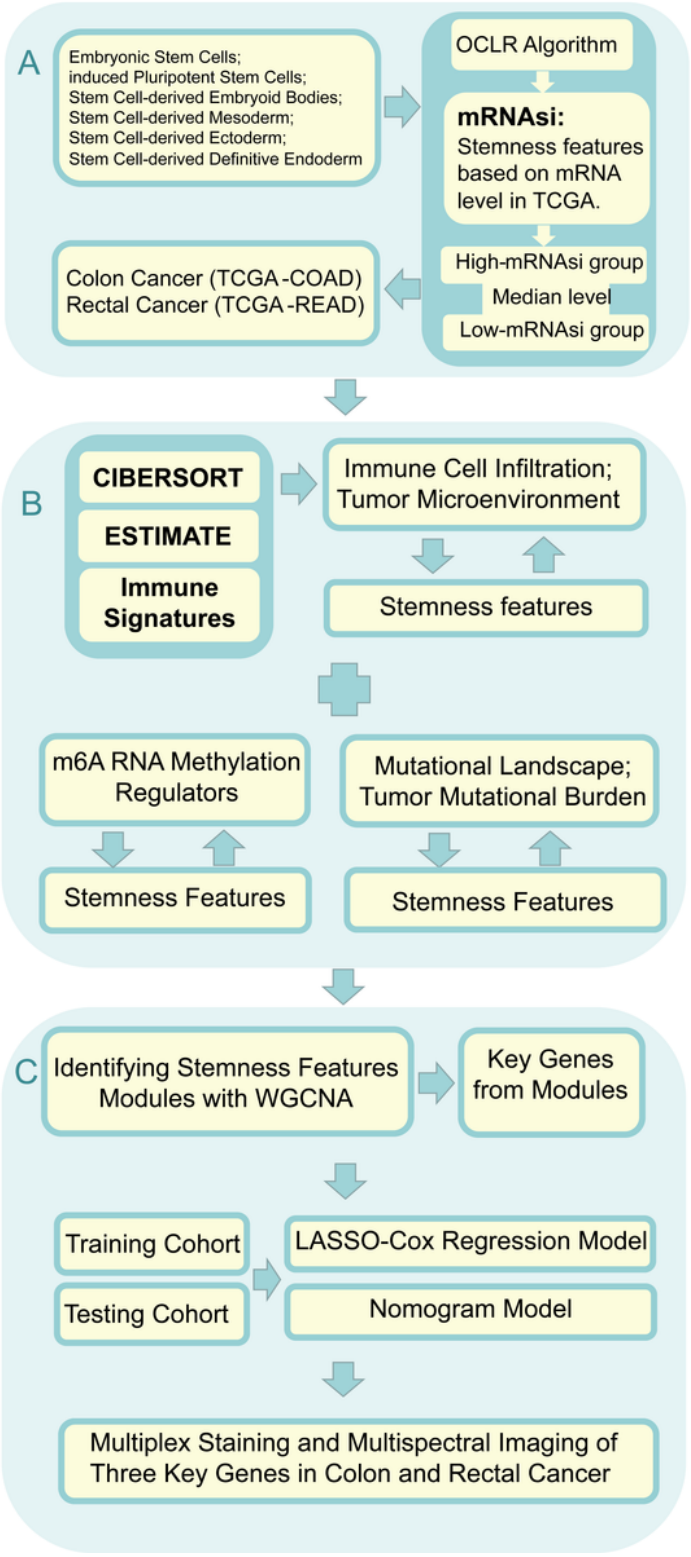

Figure 1

The analysis process of this study.

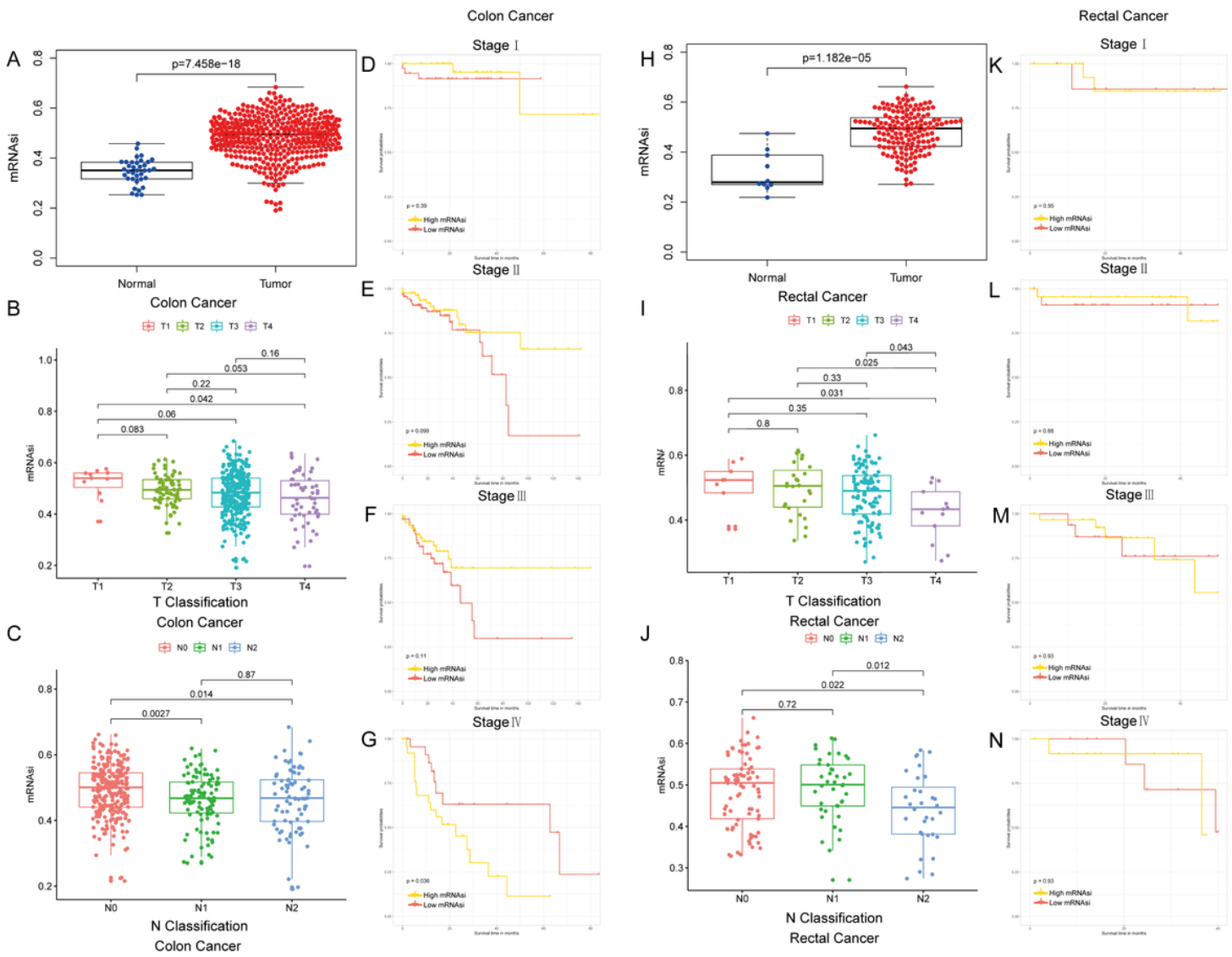

Figure 2

Colon and rectal cancer patients could be subtyped using the mRNAsi. (A) Differences in mRNAsi between normal and colon cancer tissues. (B, C) Comparison of mRNAsi in different T stages (B) or N stages (C) in colon cancer. (D-G) Kaplan–Meier survival curves of mRNAsi in colon cacner. $p < 0.05$ indicates statistical significance. (H) Differences in mRNAsi between normal and colon cancer tissues. (I, J) Comparison of mRNAsi in different T stages (I) or N stages (J) in rectal cancer. (K-N) Kaplan–Meier survival curves of mRNAsi in rectal cancer. $p < 0.05$ indicates statistical significance.

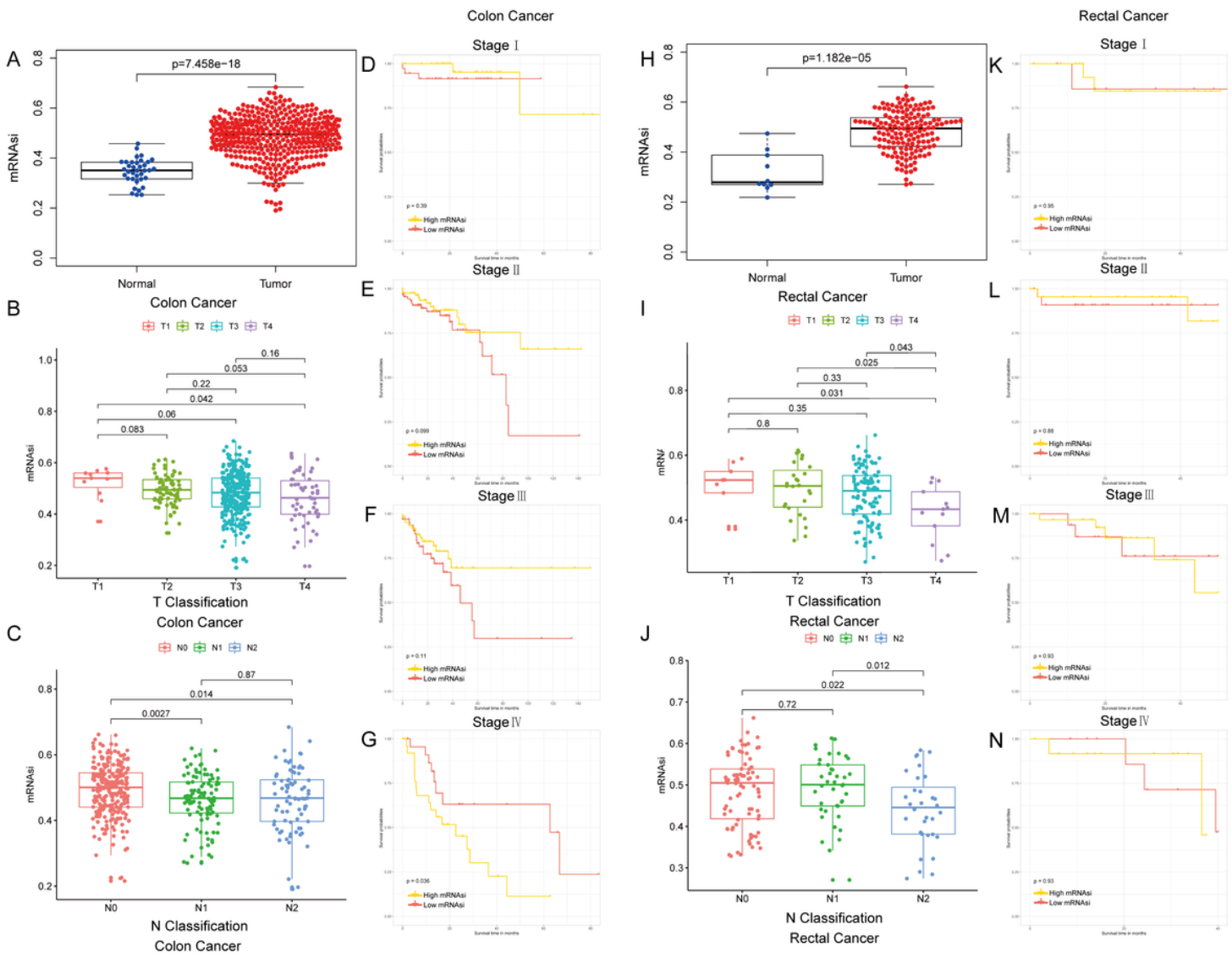

Figure 2

Colon and rectal cancer patients could be subtyped using the mRNAsi. (A) Differences in mRNAsi between normal and colon cancer tissues. (B, C) Comparison of mRNAsi in different T stages (B) or N stages (C) in colon cancer. (D-G) Kaplan–Meier survival curves of mRNAsi in colon cacner. $p < 0.05$ indicates statistical significance. (H) Differences in mRNAsi between normal and colon cancer tissues. (I, J) Comparison of mRNAsi in different T stages (I) or N stages (J) in rectal cancer. (K-N) Kaplan–Meier survival curves of mRNAsi in rectal cancer. $p < 0.05$ indicates statistical significance.

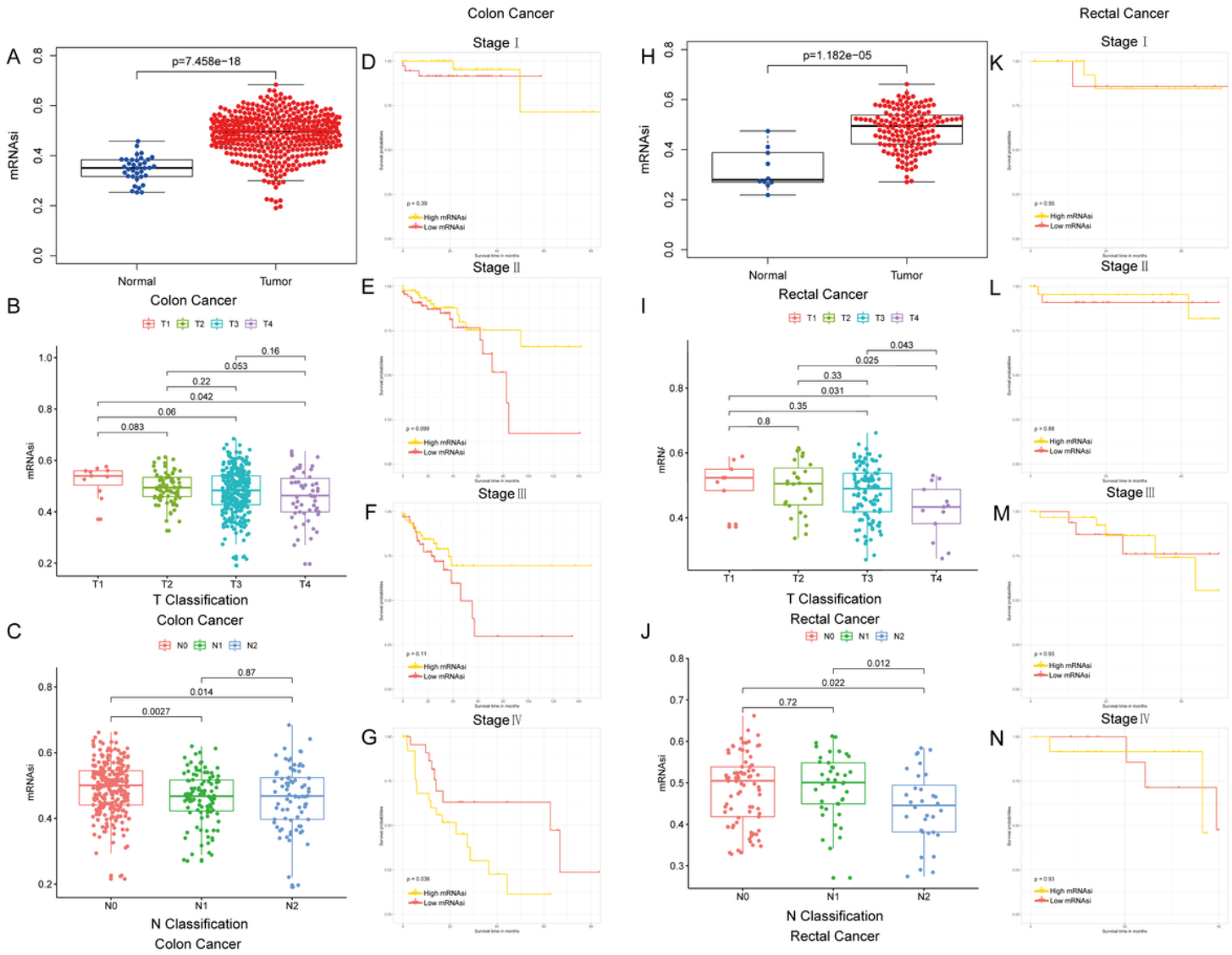

Figure 2

Colon and rectal cancer patients could be subtyped using the mRNAsi. (A) Differences in mRNAsi between normal and colon cancer tissues. (B, C) Comparison of mRNAsi in different T stages (B) or N stages (C) in colon cancer. (D-G) Kaplan–Meier survival curves of mRNAsi in colon cacner. $p < 0.05$ indicates statistical significance. (H) Differences in mRNAsi between normal and colon cancer tissues. (I, J) Comparison of mRNAsi in different T stages (I) or N stages (J) in rectal cancer. (K-N) Kaplan–Meier survival curves of mRNAsi in rectal cancer. $p < 0.05$ indicates statistical significance.

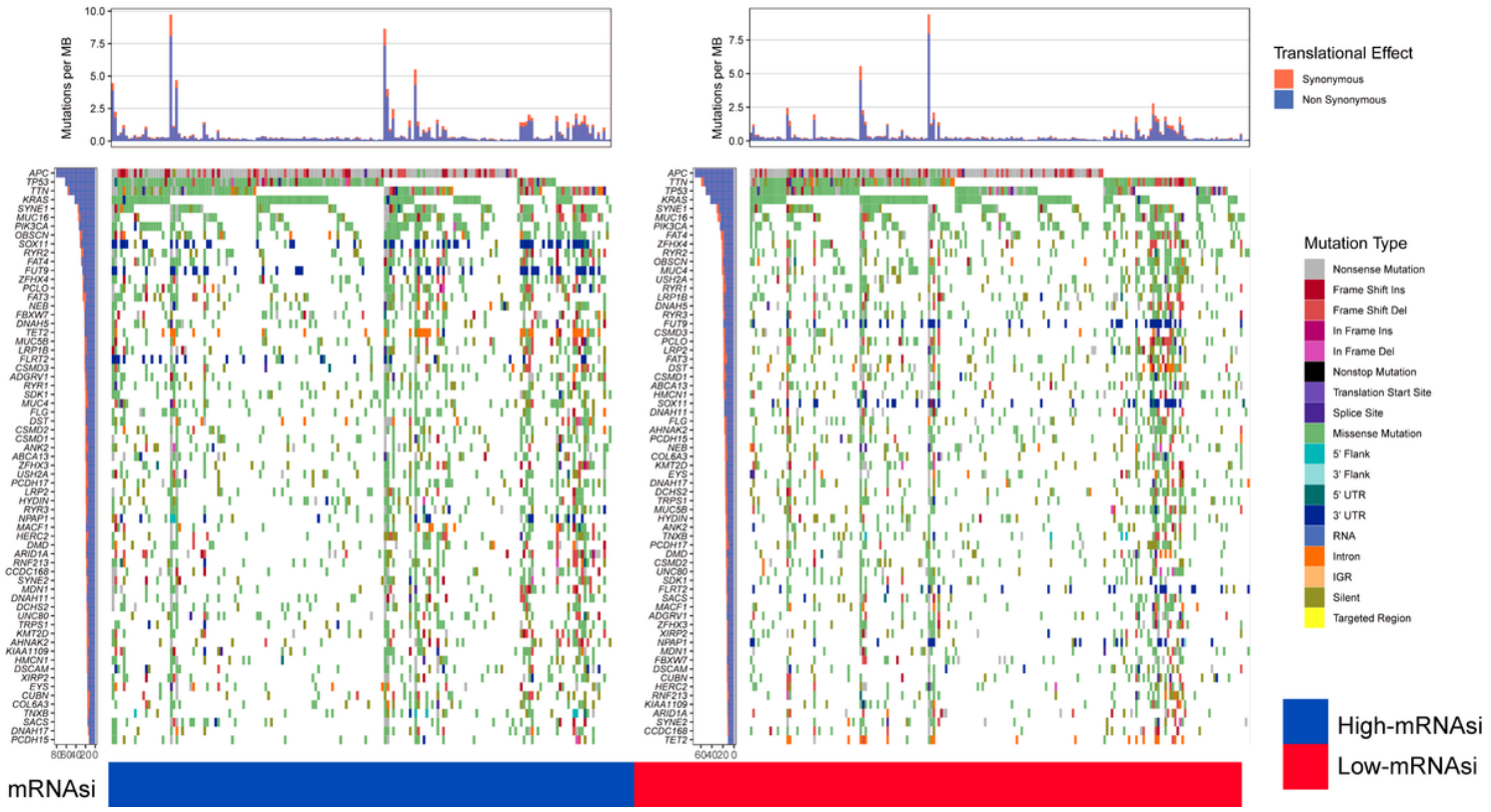

Figure 3

Association of the mRNAsi index with the colon cancer mutational landscape.

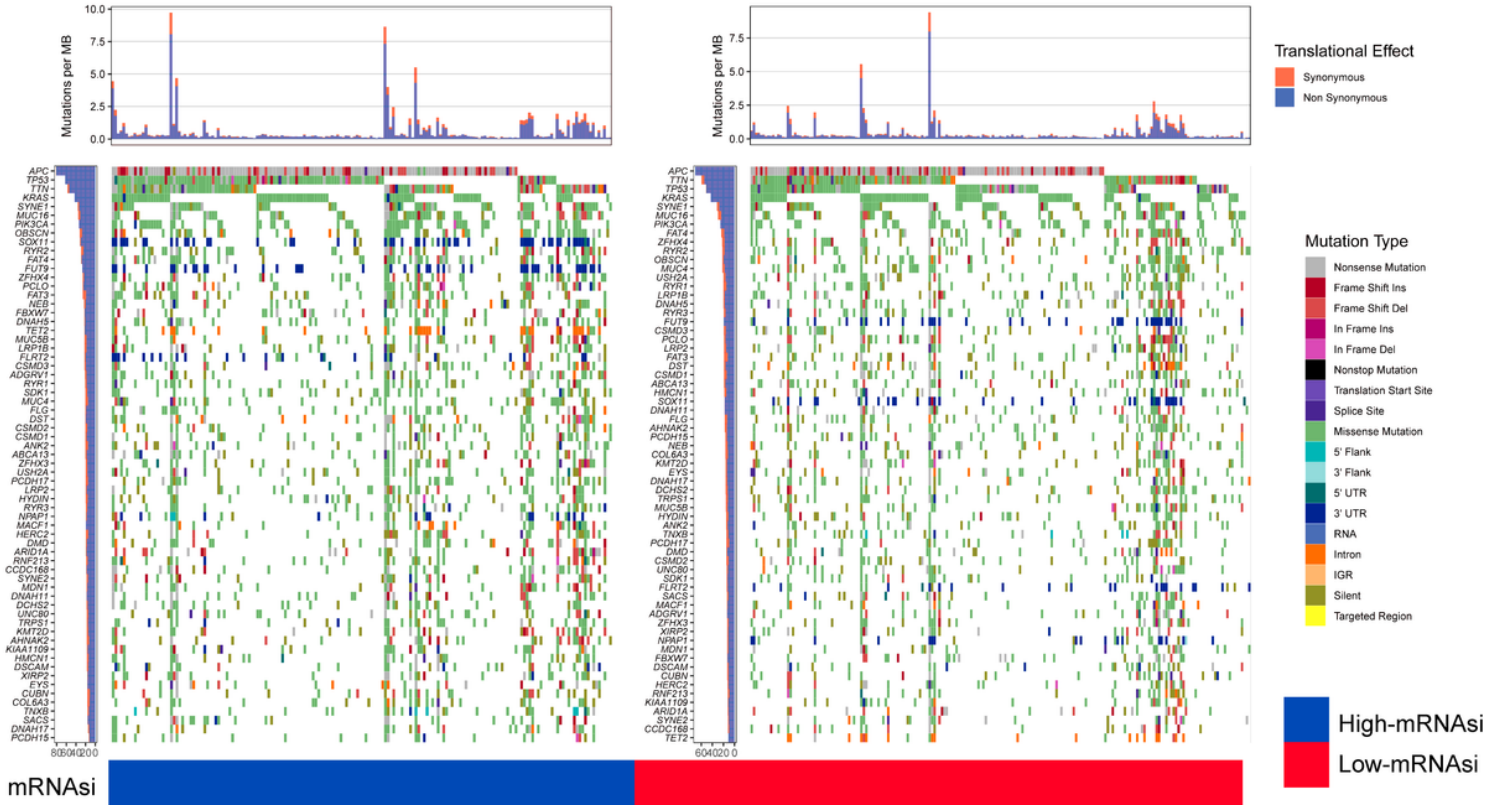

Figure 3

Association of the mRNAsi index with the colon cancer mutational landscape.

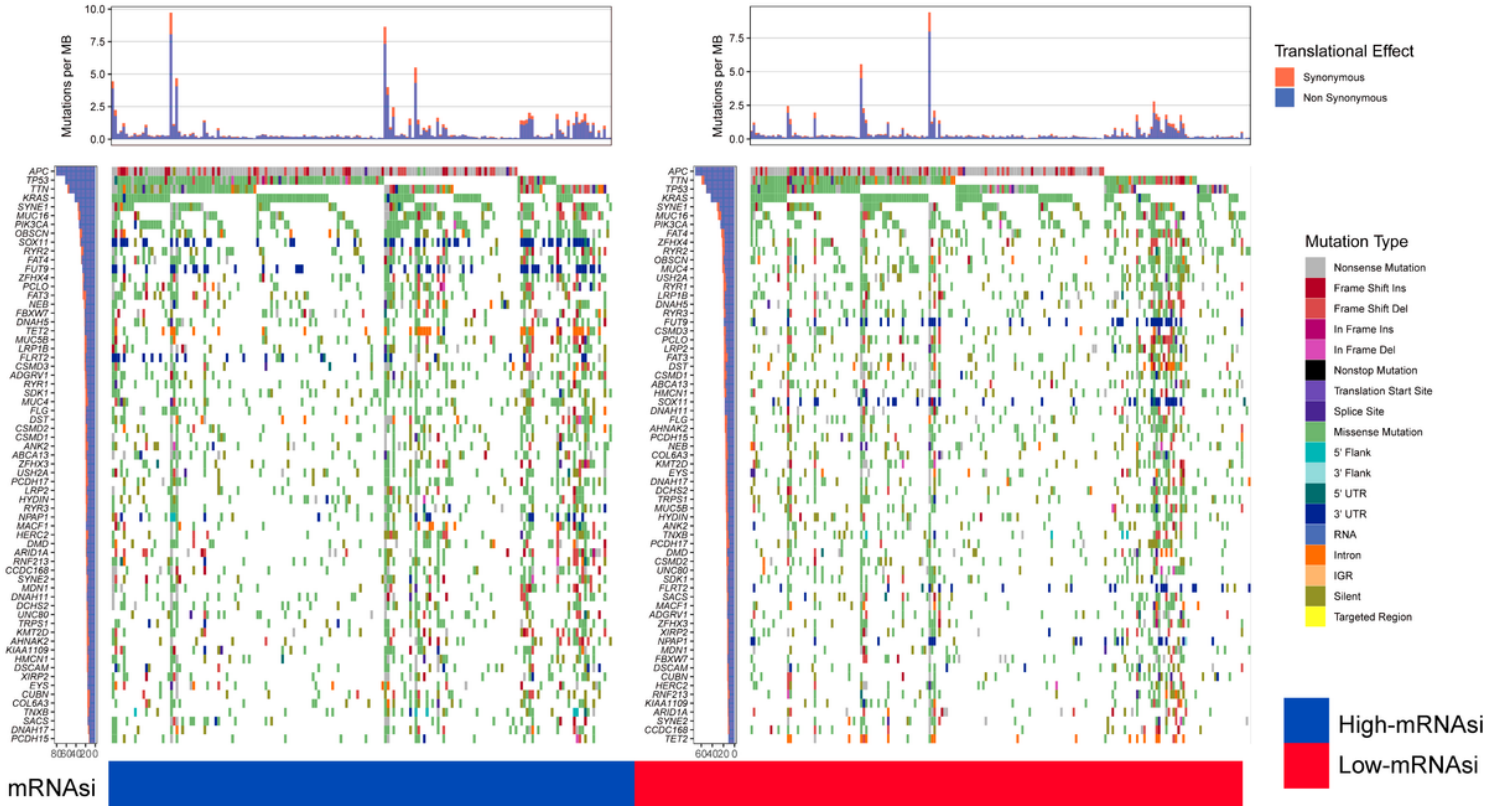

Figure 3

Association of the mRNAsi index with the colon cancer mutational landscape.

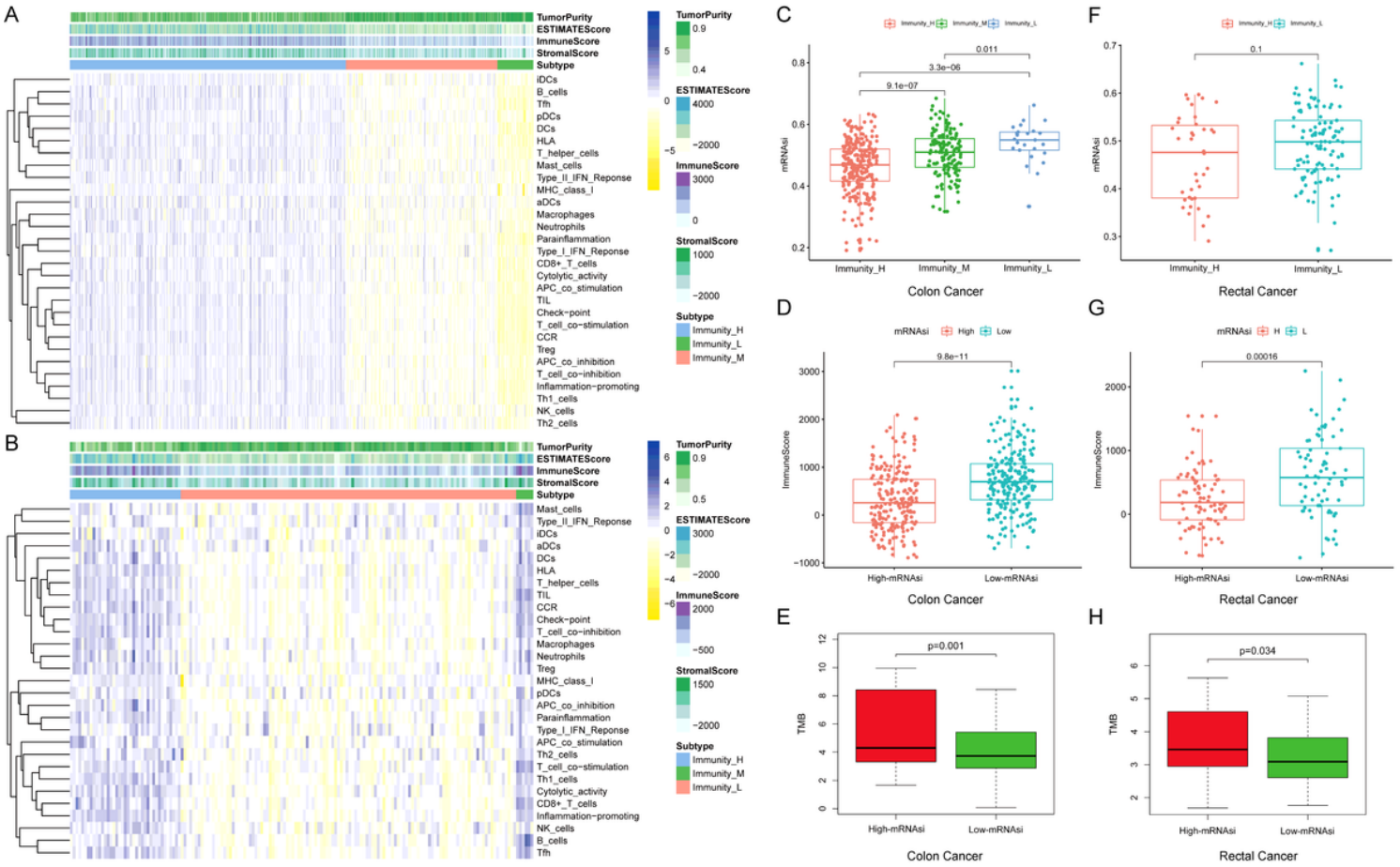

Figure 4

Hierarchical clustering of colon and rectal cancer yields three stable subtypes in four different datasets named Immunity_H, Immunity_M and Immunity_L. Tumor purity, Stromal score, and Immune score were evaluated by ESTIMATE. (A) The colon cancer patients in TCGA-COAD database. (B) The rectal cancer patients in TCGA-READ database. (C, F) Comparison of the mRNAsi levels between three colorectal cancer subtypes. Mann–Whitney U test. (D, G) Comparison of the immune cell infiltration levels between different mRNAsi group of colorectal cancer. Mann–Whitney U test. (E, H) Comparison of the TMB levels between different mRNAsi group of colorectal cancer. Kruskal–Wallis rank sum test. **p < 0.01; ***p < 0.001; p ≥ 0.05, not significant.

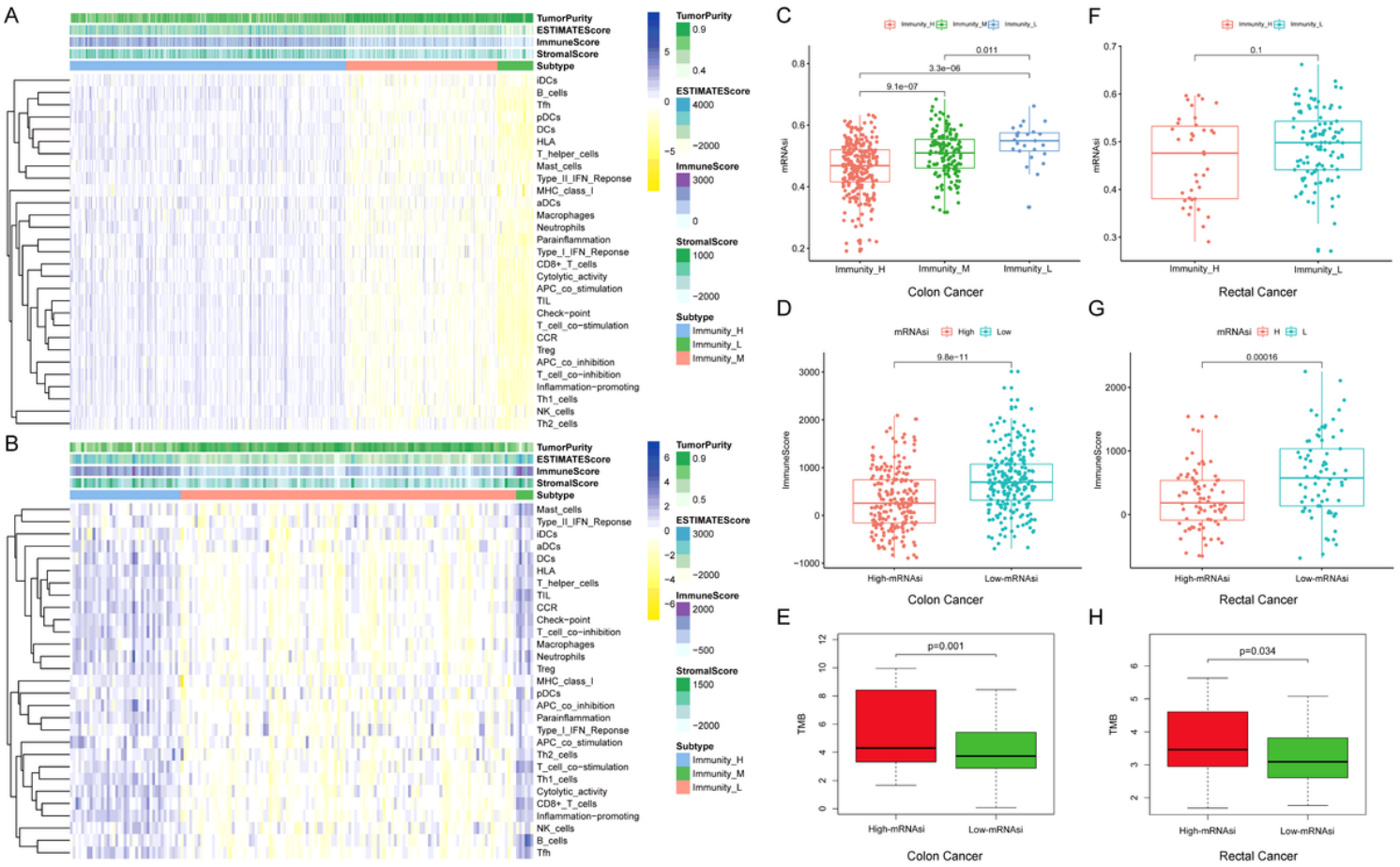

Figure 4

Hierarchical clustering of colon and rectal cancer yields three stable subtypes in four different datasets named Immunity_H, Immunity_M and Immunity_L. Tumor purity, Stromal score, and Immune score were evaluated by ESTIMATE. (A) The colon cancer patients in TCGA-COAD database. (B) The rectal cancer patients in TCGA-READ database. (C, F) Comparison of the mRNAsi levels between three colorectal cancer subtypes. Mann–Whitney U test. (D, G) Comparison of the immune cell infiltration levels between different mRNAsi group of colorectal cancer. Mann–Whitney U test. (E, H) Comparison of the TMB levels between different mRNAsi group of colorectal cancer. Kruskal–Wallis rank sum test. **p < 0.01; ***p < 0.001; p ≥ 0.05, not significant.

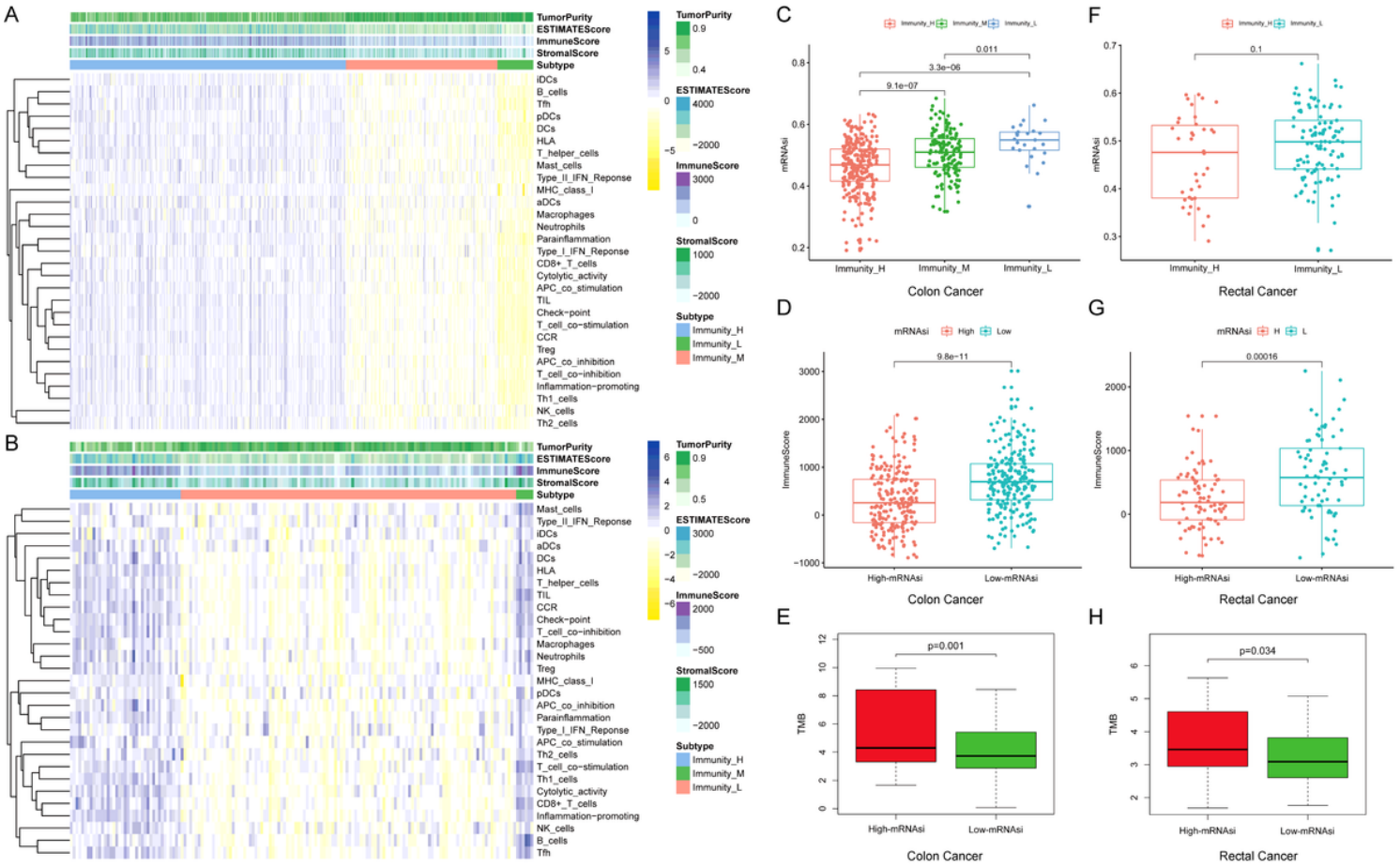

Figure 4

Hierarchical clustering of colon and rectal cancer yields three stable subtypes in four different datasets named Immunity_H, Immunity_M and Immunity_L. Tumor purity, Stromal score, and Immune score were evaluated by ESTIMATE. (A) The colon cancer patients in TCGA-COAD database. (B) The rectal cancer patients in TCGA-READ database. (C, F) Comparison of the mRNAsi levels between three colorectal cancer subtypes. Mann–Whitney U test. (D, G) Comparison of the immune cell infiltration levels between different mRNAsi group of colorectal cancer. Mann–Whitney U test. (E, H) Comparison of the TMB levels between different mRNAsi group of colorectal cancer. Kruskal–Wallis rank sum test. **p < 0.01; ***p < 0.001; p ≥ 0.05, not significant.

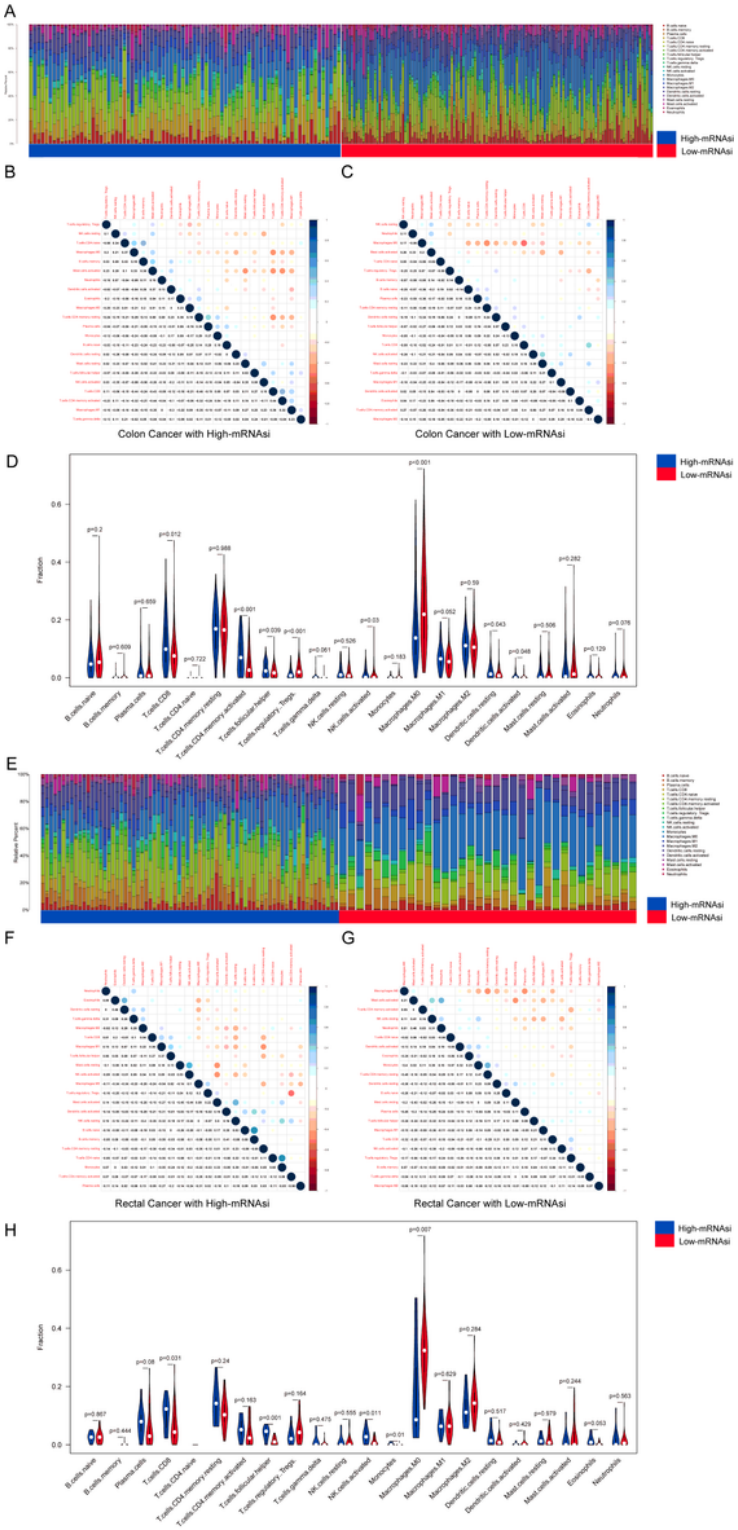

Figure 5

Composition of immune cells in different mRNAsi group and correlation analysis. (A) Barplot showing the fractions of 22 immune cells of colon cancer patients in TCGA-COAD database. Column names of plot were sample ID. (B) Heatmap showing the correlation between immune cells of colon cancer cases in high mRNAsi group. (C) Heatmap showing the correlation between immune cells of colon cancer cases in low mRNAsi group. The shade of each tiny color box represented corresponding correlation value

between two cells. (D) Comparison of the proportions of immune cell subsets between different mRNAsi group of colon cancer. ANOVA test, P values are shown. (E) Barplot showing the fractions of 22 immune cells of rectal cancer patients in TCGA-READ database. Column names of plot were sample ID. (F) Heatmap showing the correlation between immune cells of rectal cancer cases in high mRNAsi group. (G) Heatmap showing the correlation between immune cells of rectal cancer cases in low mRNAsi group. The shade of each tiny color box represented corresponding correlation value between two cells. (H) Comparison of the proportions of immune cell subsets between different mRNAsi group of rectal cancer. ANOVA test, P values are shown.

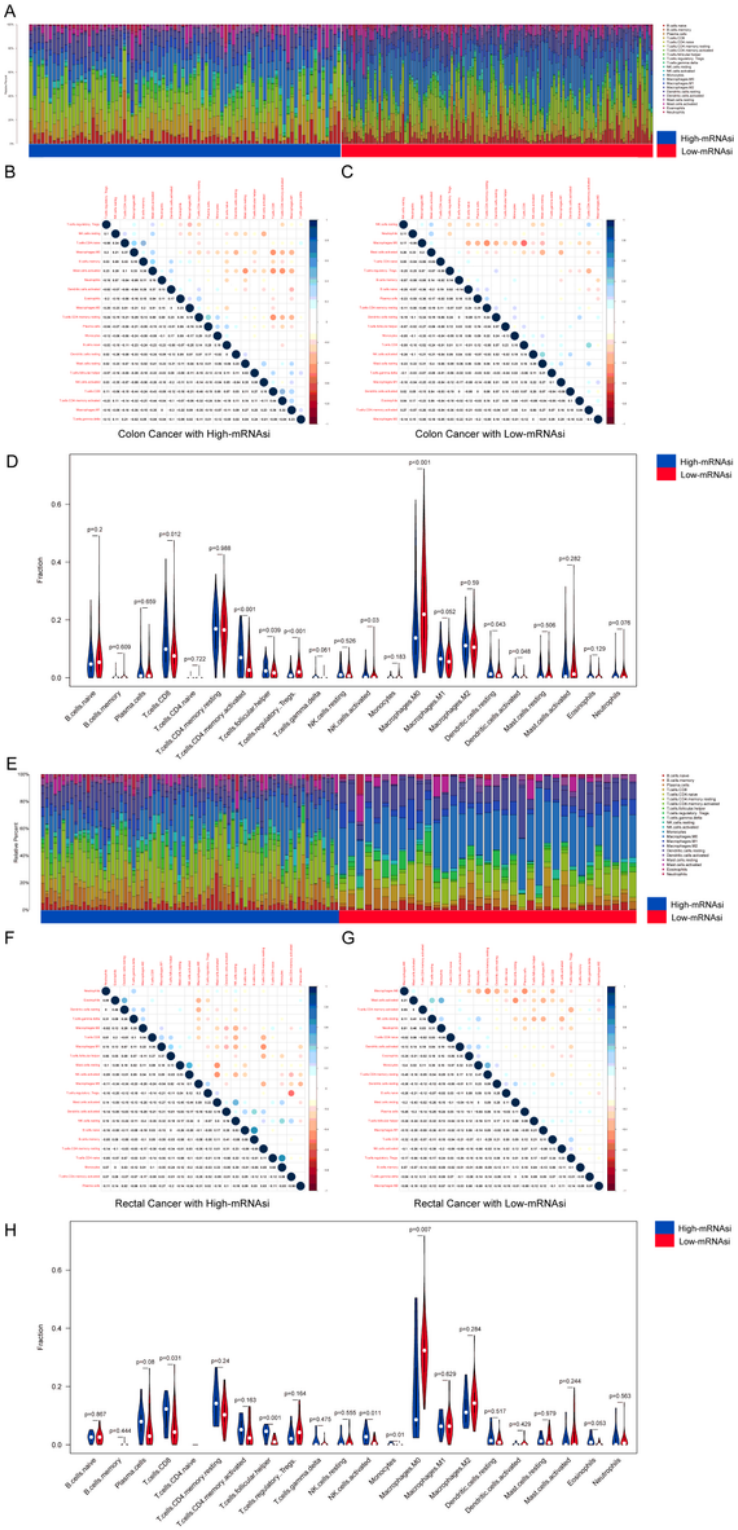

Figure 5

Composition of immune cells in different mRNAsi group and correlation analysis. (A) Barplot showing the fractions of 22 immune cells of colon cancer patients in TCGA-COAD database. Column names of plot were sample ID. (B) Heatmap showing the correlation between immune cells of colon cancer cases in high mRNAsi group. (C) Heatmap showing the correlation between immune cells of colon cancer cases in low mRNAsi group. The shade of each tiny color box represented corresponding correlation value

between two cells. (D) Comparison of the proportions of immune cell subsets between different mRNAsi group of colon cancer. ANOVA test, P values are shown. (E) Barplot showing the fractions of 22 immune cells of rectal cancer patients in TCGA-READ database. Column names of plot were sample ID. (F) Heatmap showing the correlation between immune cells of rectal cancer cases in high mRNAsi group. (G) Heatmap showing the correlation between immune cells of rectal cancer cases in low mRNAsi group. The shade of each tiny color box represented corresponding correlation value between two cells. (H) Comparison of the proportions of immune cell subsets between different mRNAsi group of rectal cancer. ANOVA test, P values are shown.

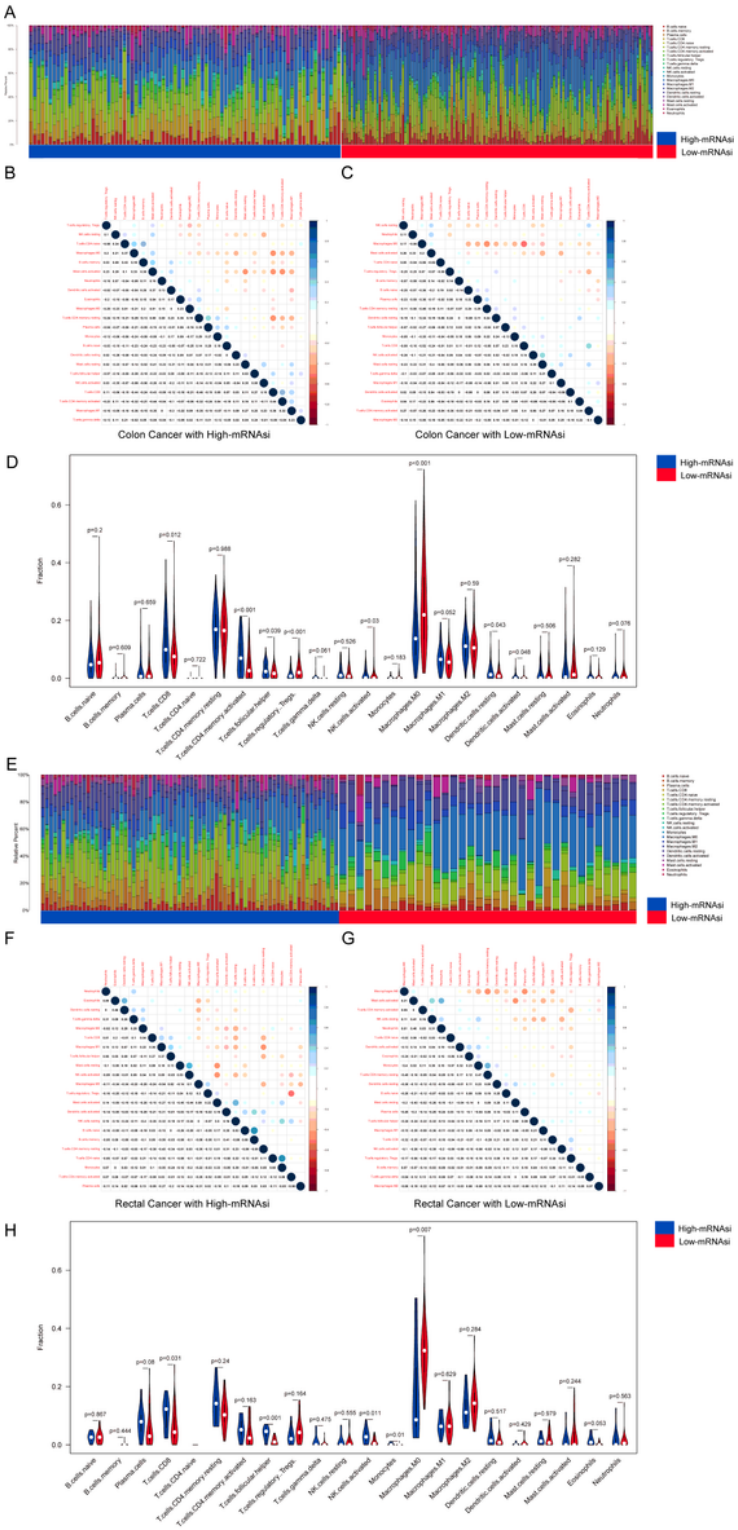

Figure 5

Composition of immune cells in different mRNAsi group and correlation analysis. (A) Barplot showing the fractions of 22 immune cells of colon cancer patients in TCGA-COAD database. Column names of plot were sample ID. (B) Heatmap showing the correlation between immune cells of colon cancer cases in high mRNAsi group. (C) Heatmap showing the correlation between immune cells of colon cancer cases in low mRNAsi group. The shade of each tiny color box represented corresponding correlation value

between two cells. (D) Comparison of the proportions of immune cell subsets between different mRNAsi group of colon cancer. ANOVA test, P values are shown. (E) Barplot showing the fractions of 22 immune cells of rectal cancer patients in TCGA-READ database. Column names of plot were sample ID. (F) Heatmap showing the correlation between immune cells of rectal cancer cases in high mRNAsi group. (G) Heatmap showing the correlation between immune cells of rectal cancer cases in low mRNAsi group. The shade of each tiny color box represented corresponding correlation value between two cells. (H) Comparison of the proportions of immune cell subsets between different mRNAsi group of rectal cancer. ANOVA test, P values are shown.

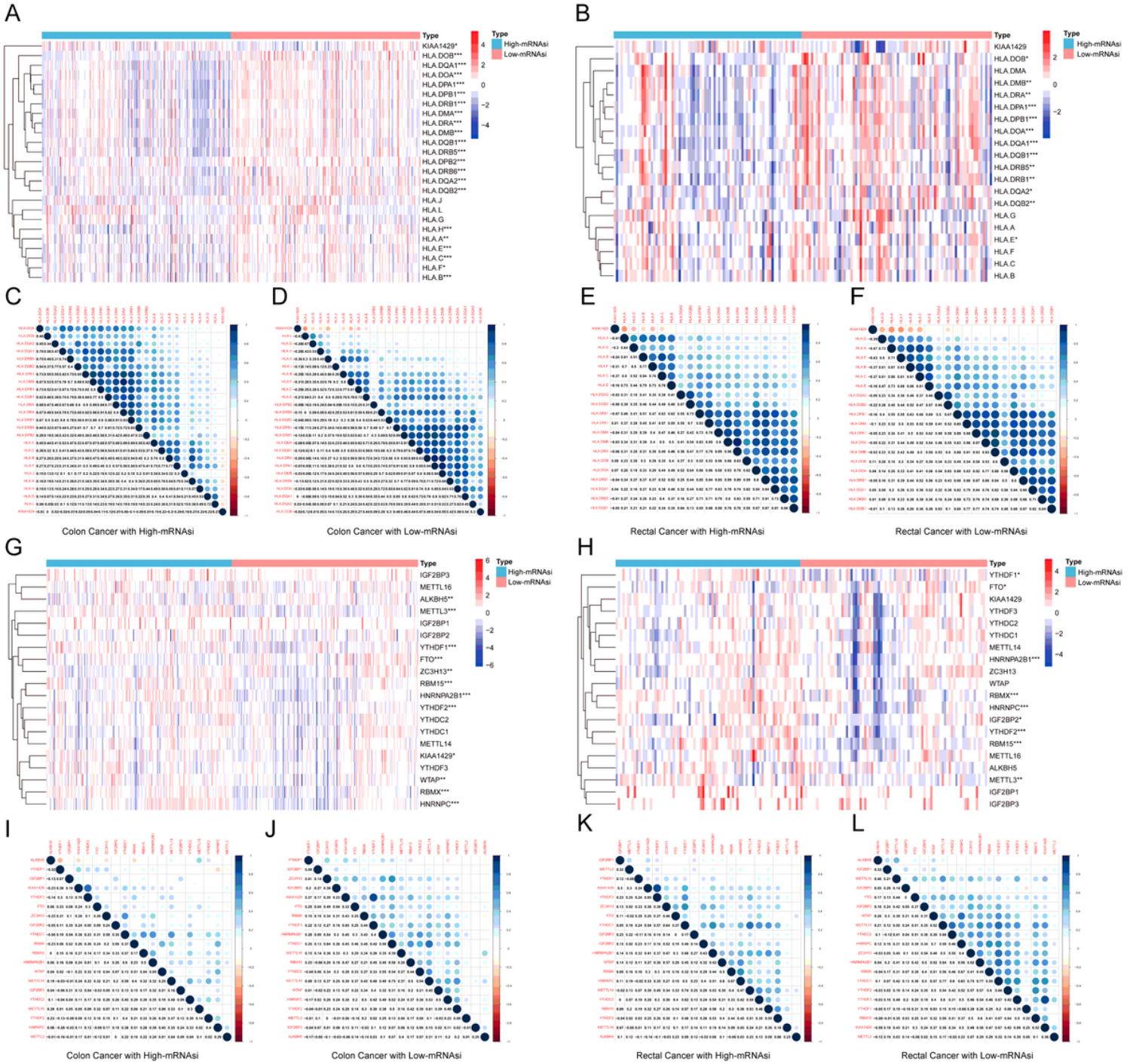

Figure 6

The expression of m6A RNA methylation regulative factors and HLA between different level of mRNAsi index. (A, B) The heatmap was used to visualize the expression levels of HLA RNA methylation regulators in different level of mRNAsi index group of colon and rectal cancer. (C-F) The Pearson correlation analysis was used to determine the correlation among HLA RNA methylation regulators. (G, H) The heatmap was used to visualize the expression levels of m6A RNA methylation regulators in different level of mRNAsi index group of colon and rectal cancer. (I-L) The Pearson correlation analysis was used to determine the correlation among m6A RNA methylation regulators.

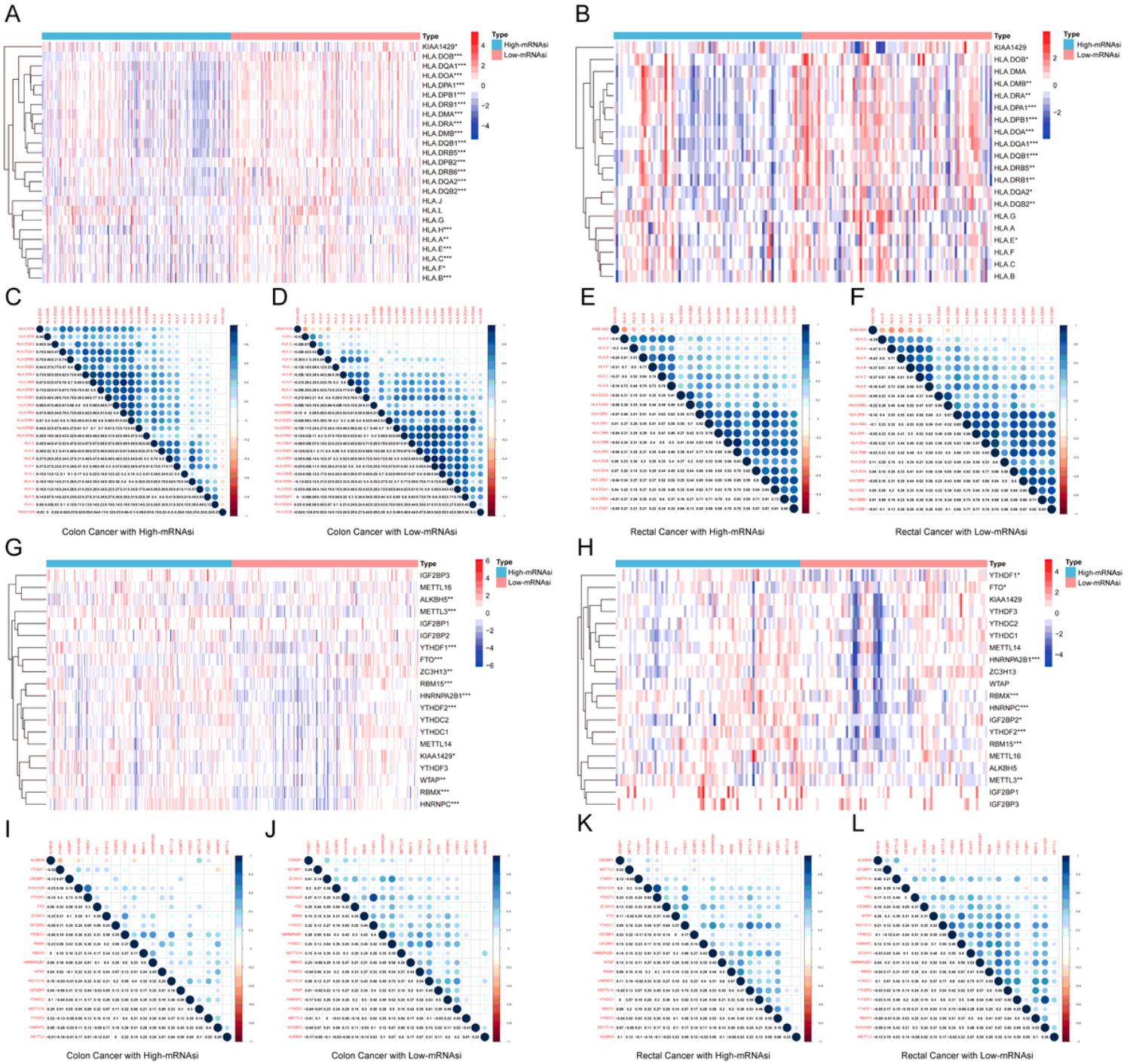

Figure 6

The expression of m6A RNA methylation regulative factors and HLA between different level of mRNAsi index. (A, B) The heatmap was used to visualize the expression levels of HLA RNA methylation regulators in different level of mRNAsi index group of colon and rectal cancer. (C-F) The Pearson correlation analysis was used to determine the correlation among HLA RNA methylation regulators. (G, H) The heatmap was used to visualize the expression levels of m6A RNA methylation regulators in different level of mRNAsi index group of colon and rectal cancer. (I-L) The Pearson correlation analysis was used to determine the correlation among m6A RNA methylation regulators.

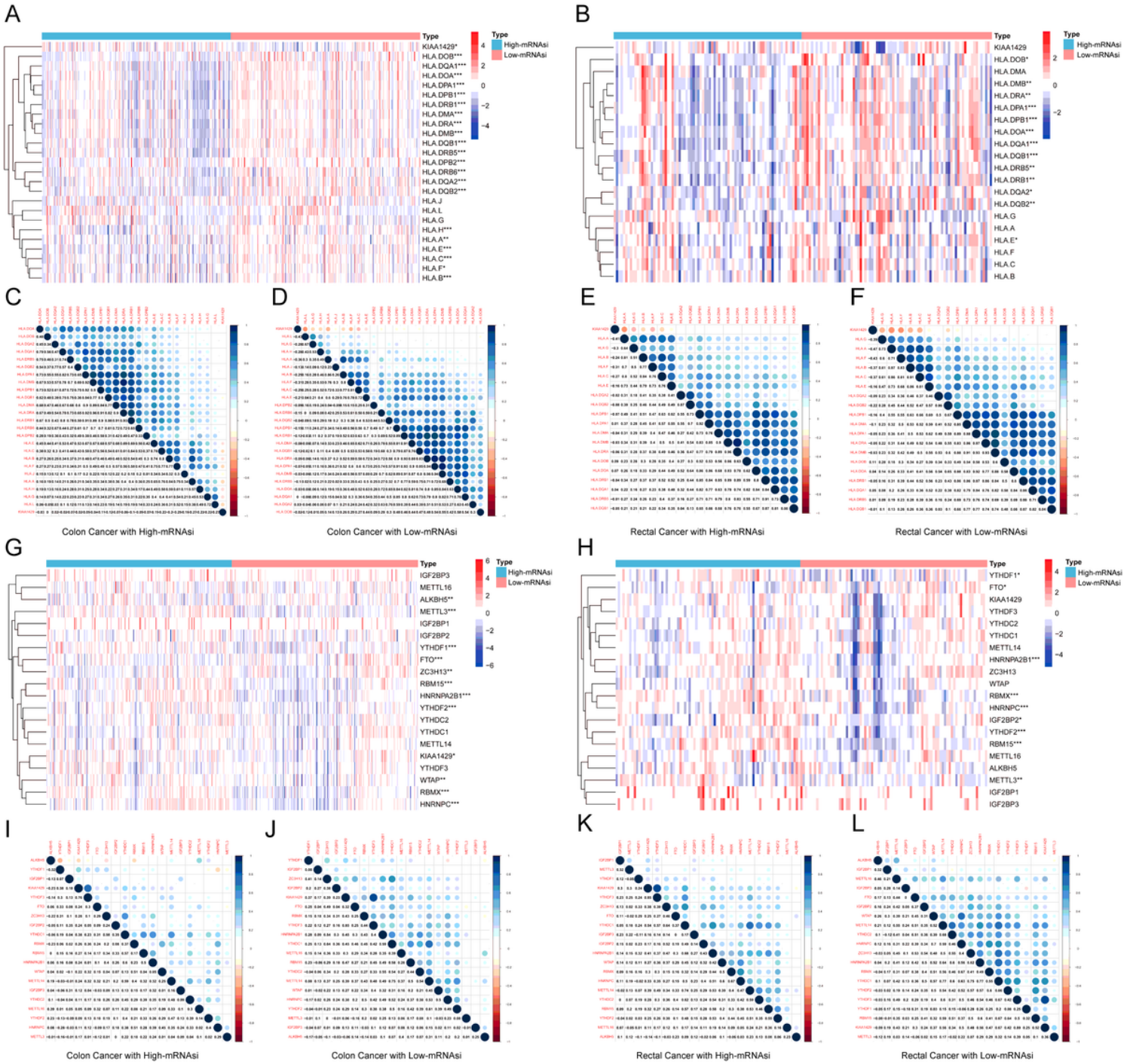

Figure 6

The expression of m6A RNA methylation regulative factors and HLA between different level of mRNAsi index. (A, B) The heatmap was used to visualize the expression levels of HLA RNA methylation regulators in different level of mRNAsi index group of colon and rectal cancer. (C-F) The Pearson correlation analysis was used to determine the correlation among HLA RNA methylation regulators. (G, H) The heatmap was used to visualize the expression levels of m6A RNA methylation regulators in different level of mRNAsi index group of colon and rectal cancer. (I-L) The Pearson correlation analysis was used to determine the correlation among m6A RNA methylation regulators.

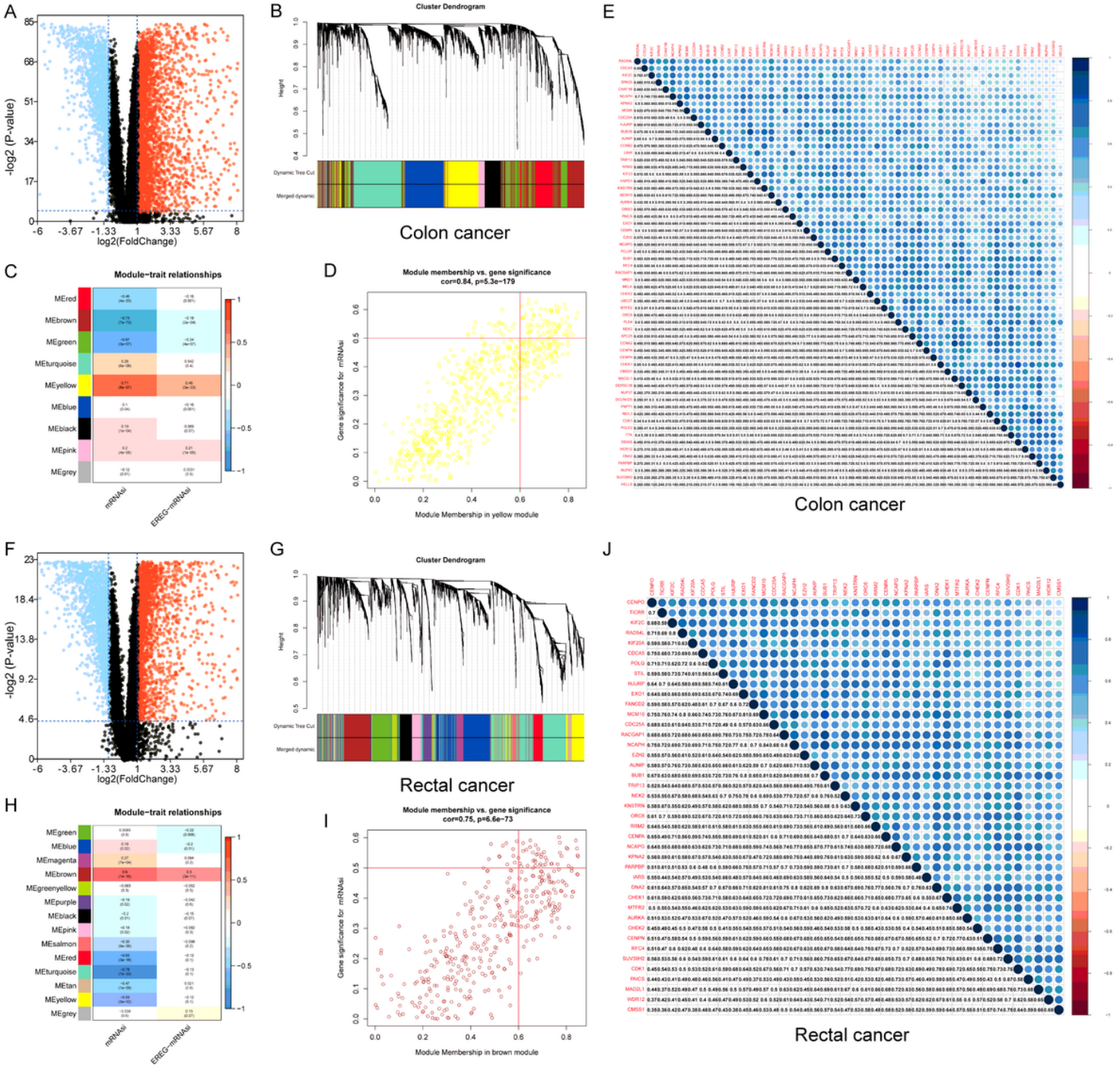

Figure 7

Identification of DEGs and stemness-related key modules in colon and rectal cancer. (A, F) Differentially expressed genes: red indicates upregulated genes; blue indicates downregulated genes and black indicates genes excluded by DEG screening criteria. (B, C, G, H) Identification of a co-expression module in colon and rectal cancer. Each piece of the leaves on the cluster dendrogram corresponded to a gene, and those genes with similar expression patterns compose a branch. Correlation between gene modules and mRNAsi scores or EREG-mRNAsi. The upper row in each cell indicates the correlation coefficient ranging from -1 to 1 of the correlation between a certain gene module and mRNAsi or EREG-mRNAsi. The lower row in each cell indicates the p-value. (D, I) The scatter plot of the top three important gene modules, and those circles located in the upper right indicate the key genes in these modules.

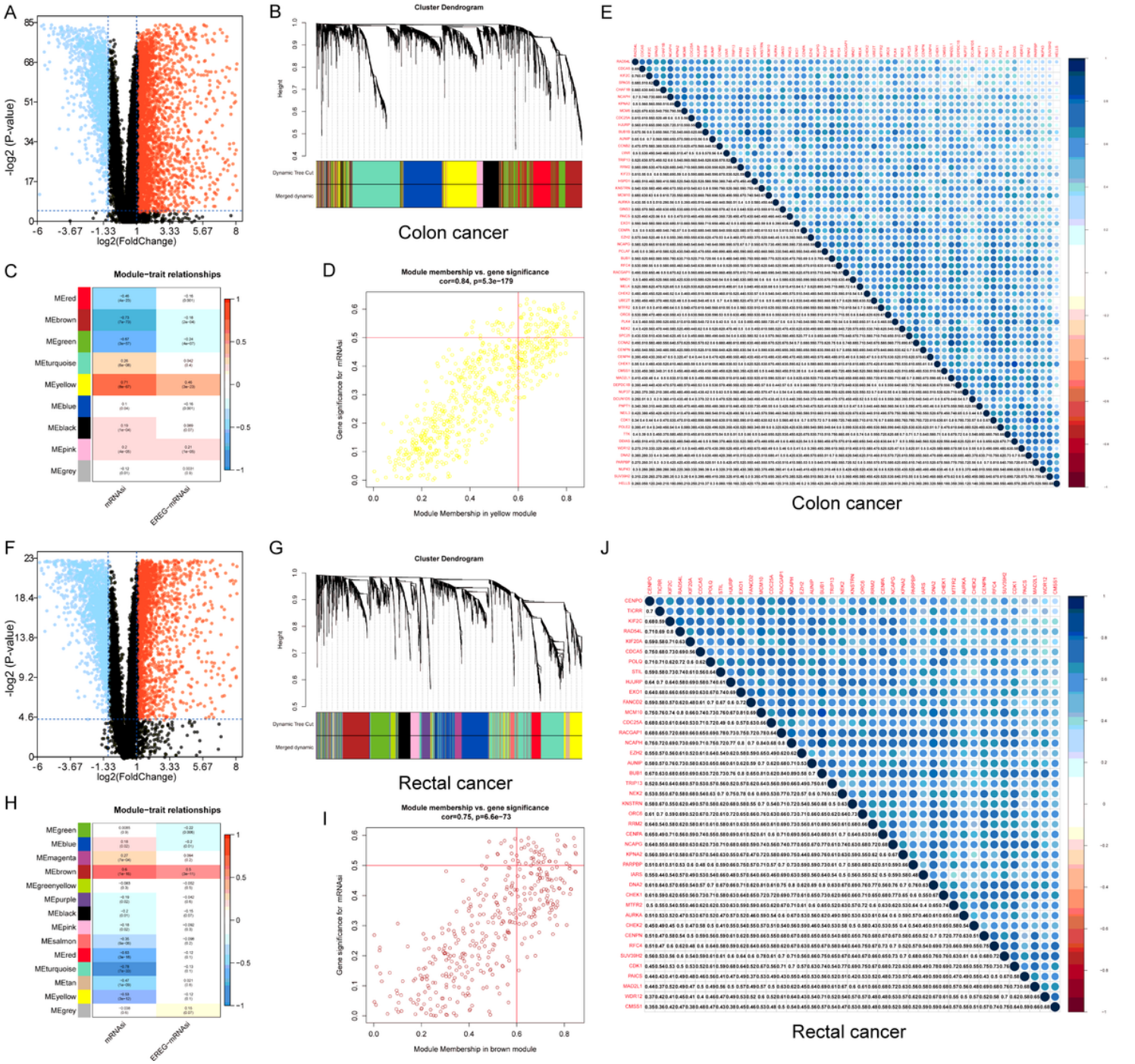

Figure 7

Identification of DEGs and stemness-related key modules in colon and rectal cancer. (A, F) Differentially expressed genes: red indicates upregulated genes; blue indicates downregulated genes and black indicates genes excluded by DEG screening criteria. (B, C, G, H) Identification of a co-expression module in colon and rectal cancer. Each piece of the leaves on the cluster dendrogram corresponded to a gene, and those genes with similar expression patterns compose a branch. Correlation between gene modules and mRNAsi scores or EREG-mRNAsi. The upper row in each cell indicates the correlation coefficient ranging from -1 to 1 of the correlation between a certain gene module and mRNAsi or EREG-mRNAsi. The lower row in each cell indicates the p-value. (D, I) The scatter plot of the top three important gene modules, and those circles located in the upper right indicate the key genes in these modules.

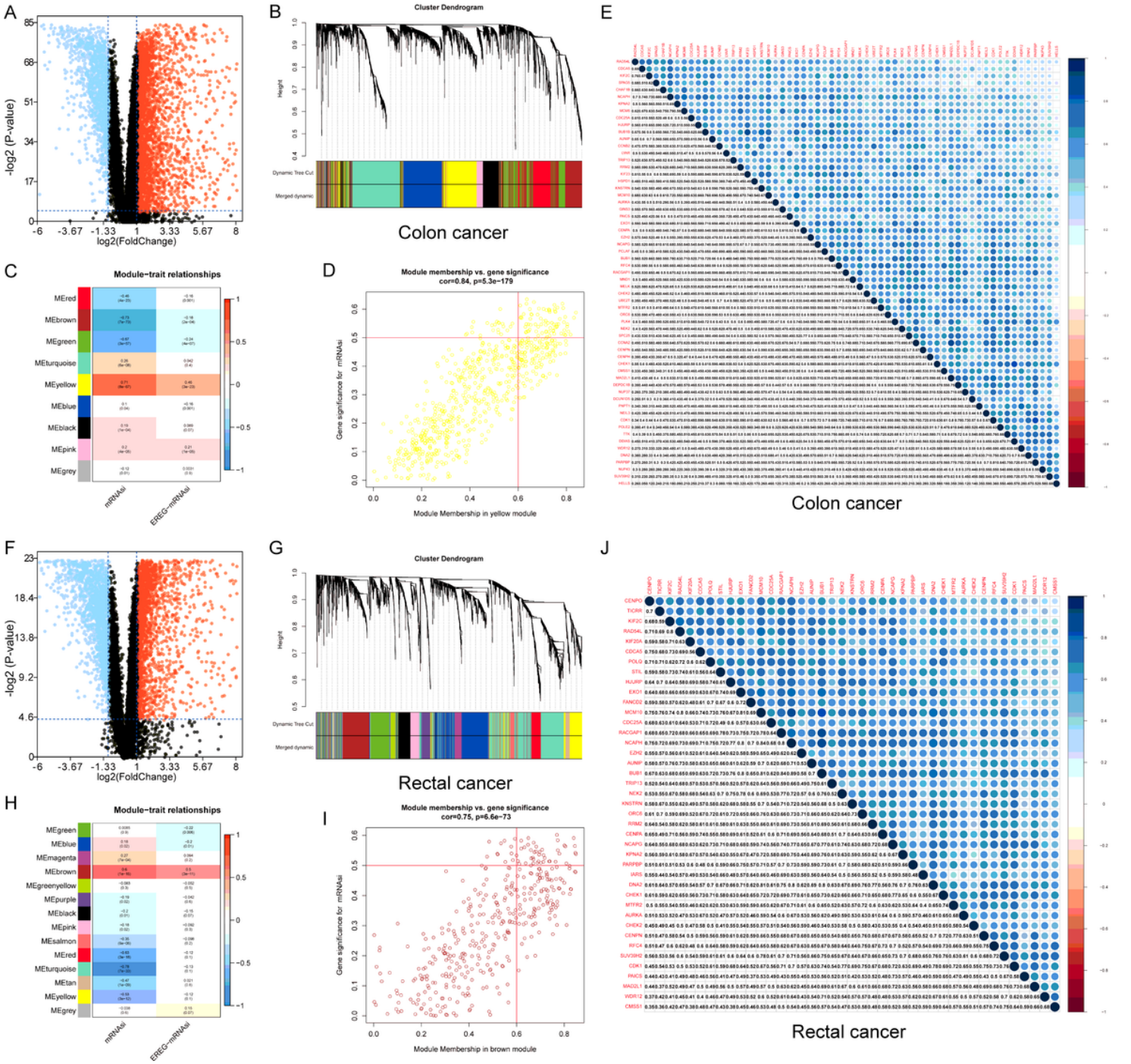

Figure 7

Identification of DEGs and stemness-related key modules in colon and rectal cancer. (A, F) Differentially expressed genes: red indicates upregulated genes; blue indicates downregulated genes and black indicates genes excluded by DEG screening criteria. (B, C, G, H) Identification of a co-expression module in colon and rectal cancer. Each piece of the leaves on the cluster dendrogram corresponded to a gene, and those genes with similar expression patterns compose a branch. Correlation between gene modules and mRNAsi scores or EREG-mRNAsi. The upper row in each cell indicates the correlation coefficient ranging from -1 to 1 of the correlation between a certain gene module and mRNAsi or EREG-mRNAsi. The lower

row in each cell indicates the p-value. (D, I) The scatter plot of the top three important gene modules, and those circles located in the upper right indicate the key genes in these modules.

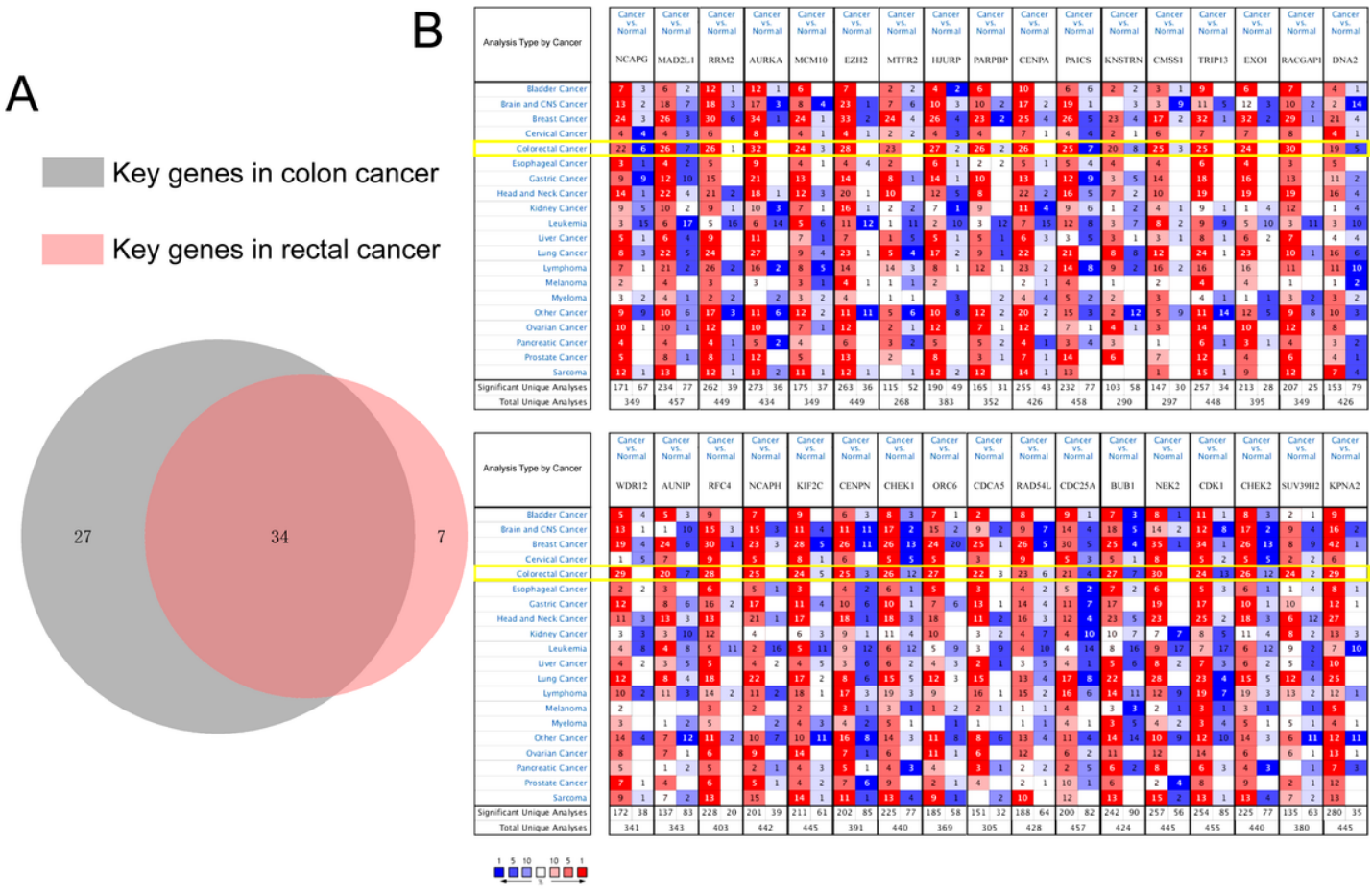

Figure 8

Expression validation of key genes. (A) The venn diagram. (B) The mRNA expression of key genes in multiple cancer types from the Oncomine database. The number in the cells represents the number of analyses meeting the thresholds. Red indicates a higher expression level of target genes in tumor tissues than in normal tissues, and blue indicates an opposite expression pattern. The color depth of each cell indicates the gene rank, and the deeper the color depth is, the higher the gene rank.

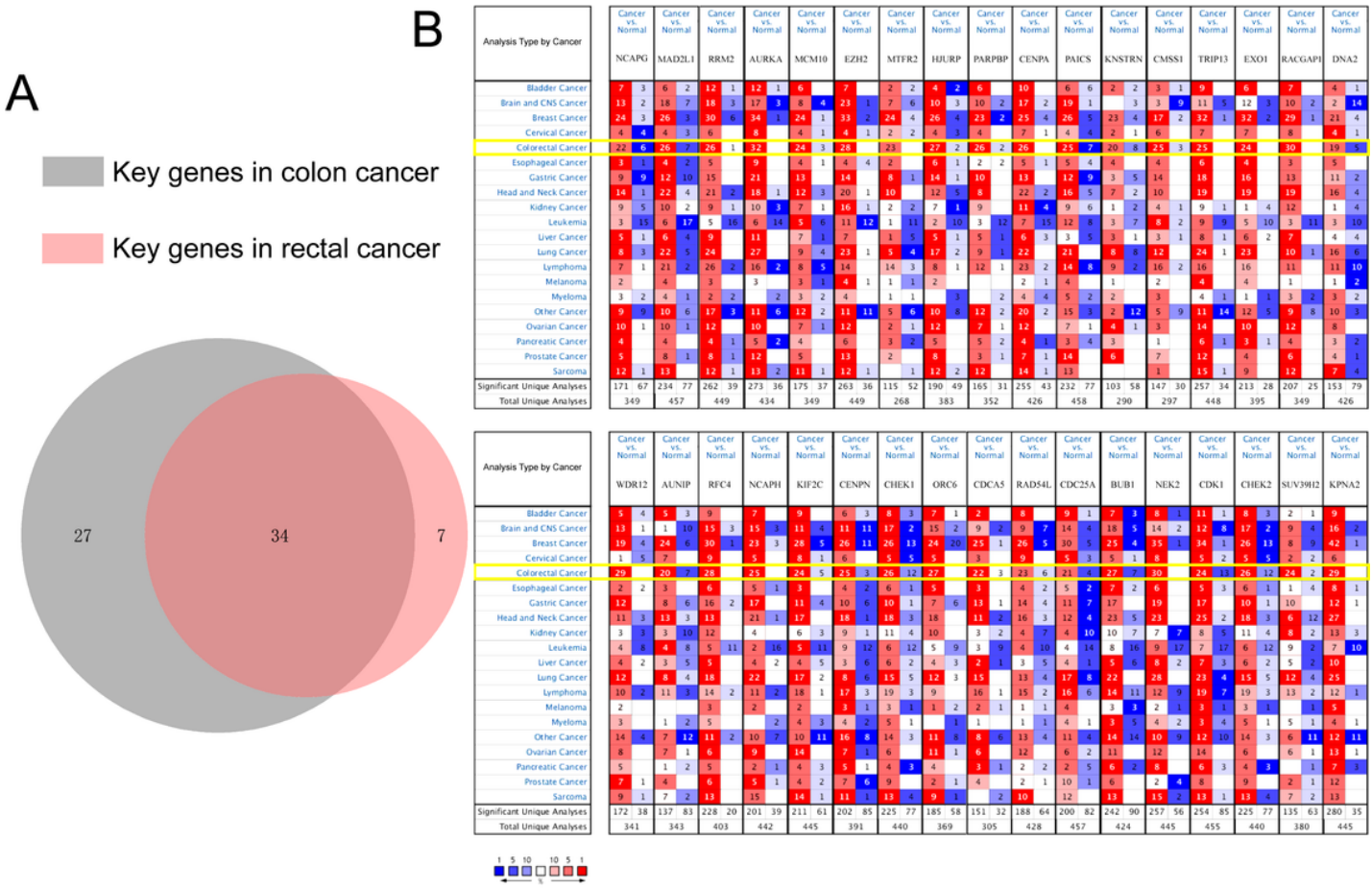

Figure 8

Expression validation of key genes. (A) The venn diagram. (B) The mRNA expression of key genes in multiple cancer types from the Oncomine database. The number in the cells represents the number of analyses meeting the thresholds. Red indicates a higher expression level of target genes in tumor tissues than in normal tissues, and blue indicates an opposite expression pattern. The color depth of each cell indicates the gene rank, and the deeper the color depth is, the higher the gene rank.

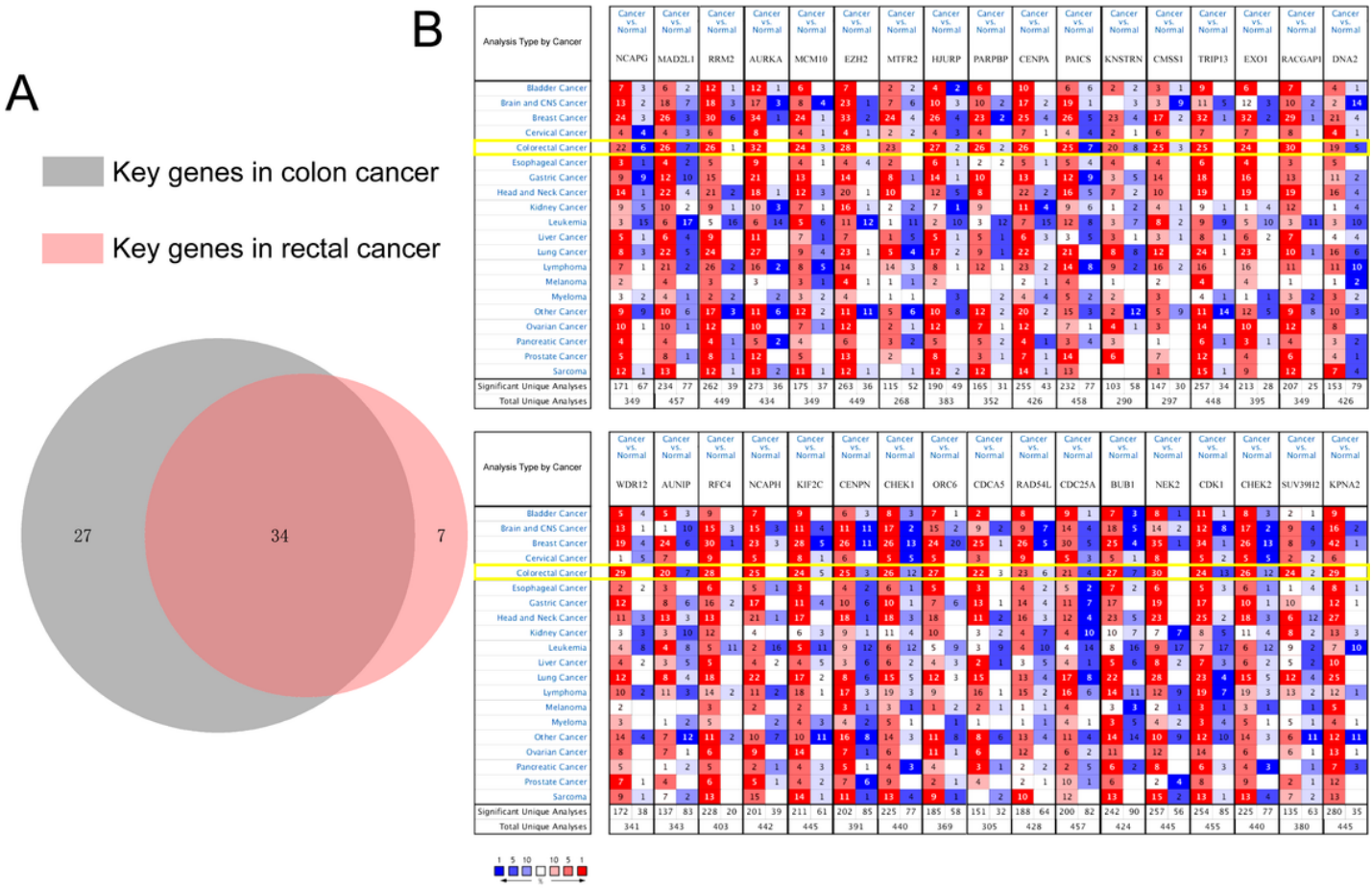

Figure 8

Expression validation of key genes. (A) The venn diagram. (B) The mRNA expression of key genes in multiple cancer types from the Oncomine database. The number in the cells represents the number of analyses meeting the thresholds. Red indicates a higher expression level of target genes in tumor tissues than in normal tissues, and blue indicates an opposite expression pattern. The color depth of each cell indicates the gene rank, and the deeper the color depth is, the higher the gene rank.

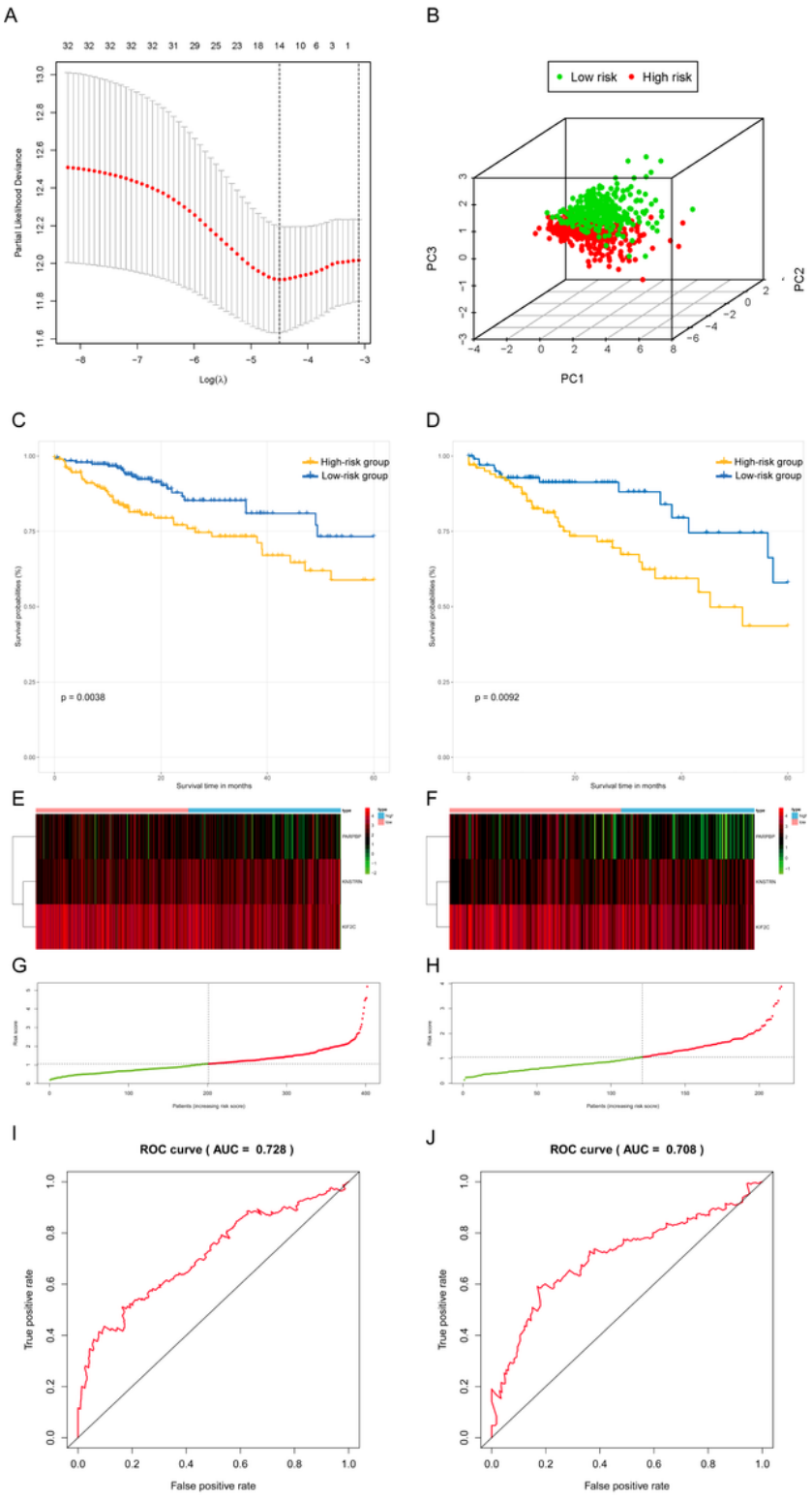

Figure 9

The Survival analysis and prognostic performance of the three-gene signature of colon and rectal cancer. (A) LASSO of the risk score for colon and rectal cancer between high-risk and low-risk patients in training group. (B) PCA of the risk score for colon and rectal cancer between high-risk and low-risk patients in training group. (C) The Kaplan–Meier test of the risk score for the overall survival of colon and rectal cancer between high-risk and low-risk patients in training group (log-rank test, $p < 0.001$); (D) The

prognostic value of the risk score showed by the time-dependent receiver operating characteristic (ROC) curve for predicting the 5 years overall survival training group; (E) Risk score curve of the three-gene signature of colon and rectal cancer in training group; (F) Heatmap showed the expression of three genes by risk score of colon and rectal cancer in training group; (G) The Kaplan–Meier test of the risk score for the overall survival of colon and rectal cancer between high-risk and low-risk patients in testing group (log-rank test, $p < 0.001$); (H) The prognostic value of the risk score showed by the time-dependent ROC curve for predicting the 5 years overall survival in testing group; (I) Risk score curve of the three-gene signature of colon and rectal cancer in testing group; (J) Heatmap showed the expression of three genes by risk score of rectal cancer in testing group.

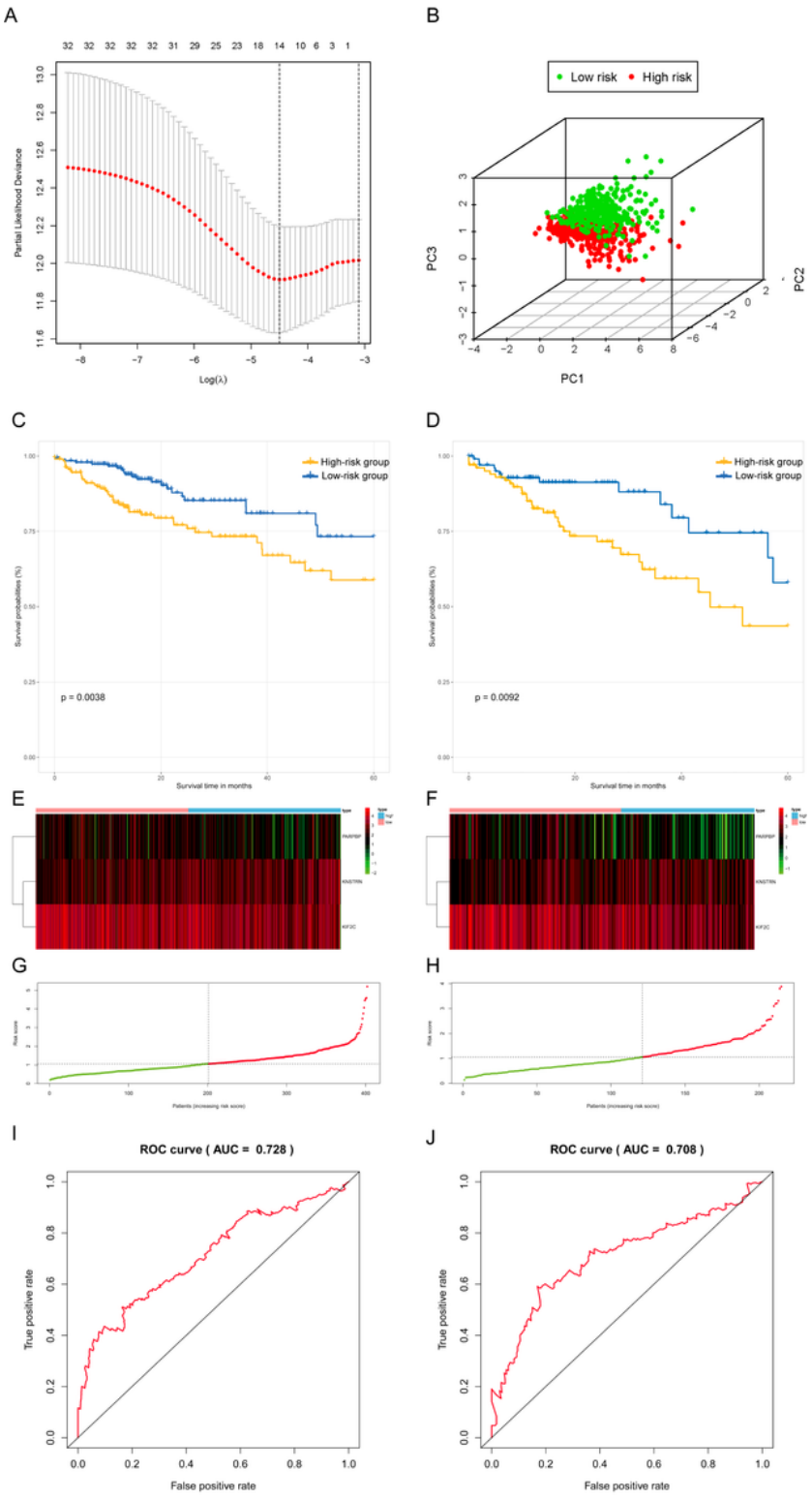

Figure 9

The Survival analysis and prognostic performance of the three-gene signature of colon and rectal cancer. (A) LASSO of the risk score for colon and rectal cancer between high-risk and low-risk patients in training group. (B) PCA of the risk score for colon and rectal cancer between high-risk and low-risk patients in training group. (C) The Kaplan–Meier test of the risk score for the overall survival of colon and rectal cancer between high-risk and low-risk patients in training group (log-rank test, $p < 0.001$); (D) The

prognostic value of the risk score showed by the time-dependent receiver operating characteristic (ROC) curve for predicting the 5 years overall survival training group; (E) Risk score curve of the three-gene signature of colon and rectal cancer in training group; (F) Heatmap showed the expression of three genes by risk score of colon and rectal cancer in training group; (G) The Kaplan–Meier test of the risk score for the overall survival of colon and rectal cancer between high-risk and low-risk patients in testing group (log-rank test, $p < 0.001$); (H) The prognostic value of the risk score showed by the time-dependent ROC curve for predicting the 5 years overall survival in testing group; (I) Risk score curve of the three-gene signature of colon and rectal cancer in testing group; (J) Heatmap showed the expression of three genes by risk score of rectal cancer in testing group.

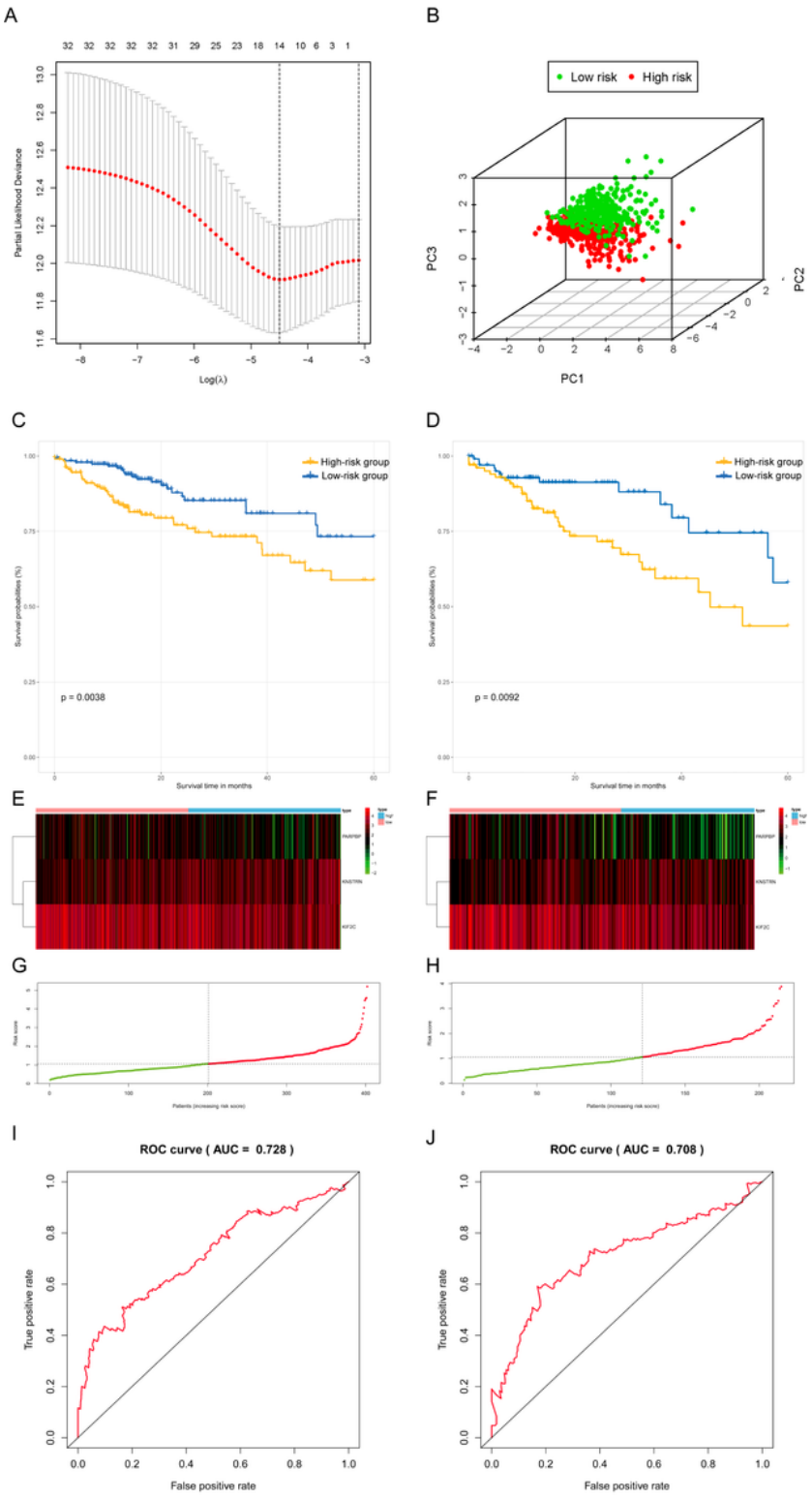

Figure 9

The Survival analysis and prognostic performance of the three-gene signature of colon and rectal cancer. (A) LASSO of the risk score for colon and rectal cancer between high-risk and low-risk patients in training group. (B) PCA of the risk score for colon and rectal cancer between high-risk and low-risk patients in training group. (C) The Kaplan–Meier test of the risk score for the overall survival of colon and rectal cancer between high-risk and low-risk patients in training group (log-rank test, $p < 0.001$); (D) The

prognostic value of the risk score showed by the time-dependent receiver operating characteristic (ROC) curve for predicting the 5 years overall survival training group; (E) Risk score curve of the three-gene signature of colon and rectal cancer in training group; (F) Heatmap showed the expression of three genes by risk score of colon and rectal cancer in training group; (G) The Kaplan–Meier test of the risk score for the overall survival of colon and rectal cancer between high-risk and low-risk patients in testing group (log-rank test, $p < 0.001$); (H) The prognostic value of the risk score showed by the time-dependent ROC curve for predicting the 5 years overall survival in testing group; (I) Risk score curve of the three-gene signature of colon and rectal cancer in testing group; (J) Heatmap showed the expression of three genes by risk score of rectal cancer in testing group.

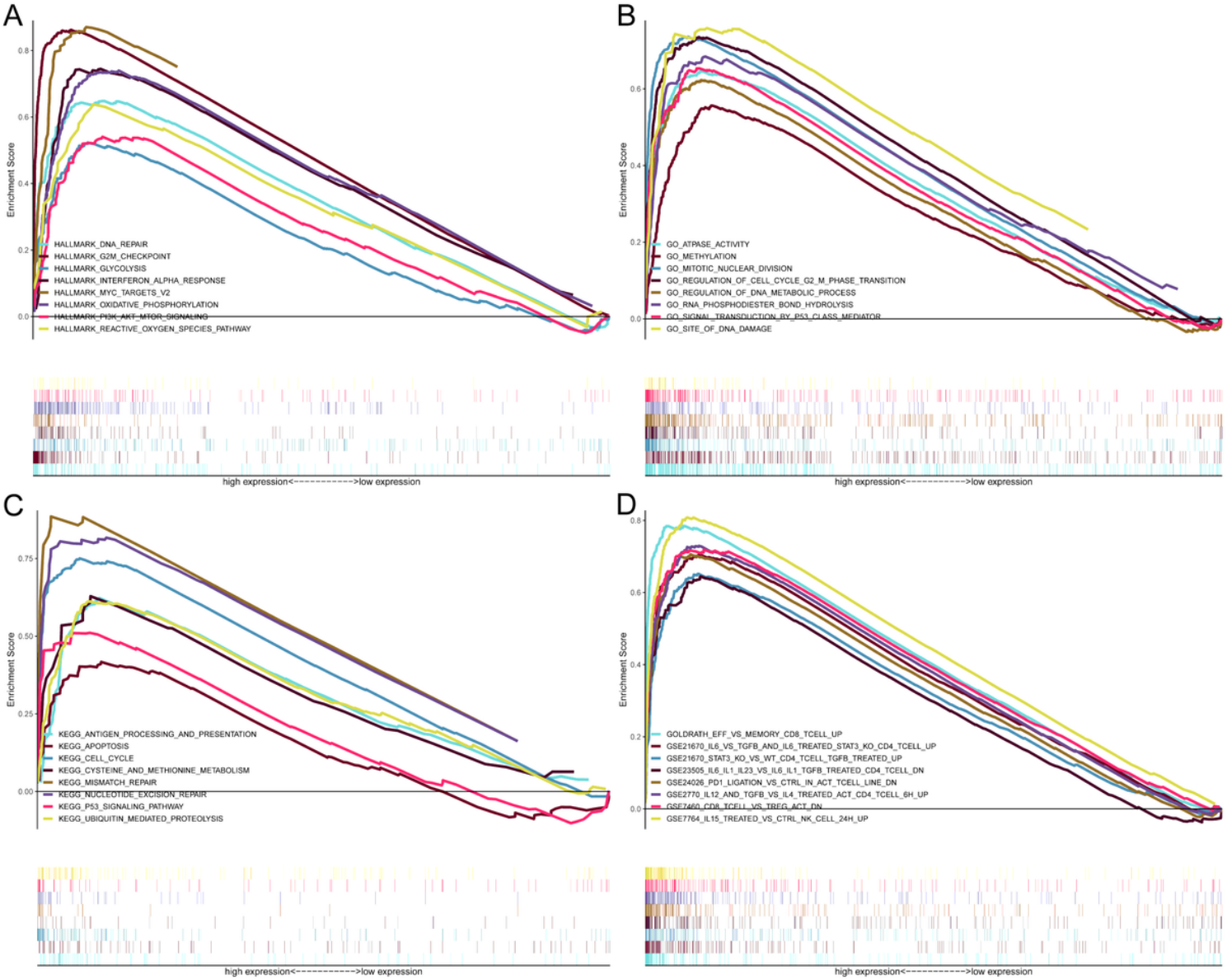

Figure 10

GSEA for samples with high-risk and low-risk based on the prognostic index of three-gene signature. (A) The enriched gene sets in GO collection by the high-risk sample. Each line representing one particular gene set with unique color, and up-regulated genes located in the left approaching the origin of the

coordinates, by contrast the down-regulated lay on the right of x-axis. Only gene sets with NOM p < 0.05 and FDR q < 0.05 were considered significant. And only several leading gene sets were displayed in the plot. (B) The enriched gene sets in KEGG by samples with high-risk sample. And only several leading gene sets were displayed in the plot. (C) Enriched gene sets in HALLMARK collection by samples of high-risk sample. Only several leading gene sets are shown in plot. (D) Enriched gene sets in C7 collection, the immunologic gene sets, by samples of high-risk sample. Only several leading gene sets are shown in plot.

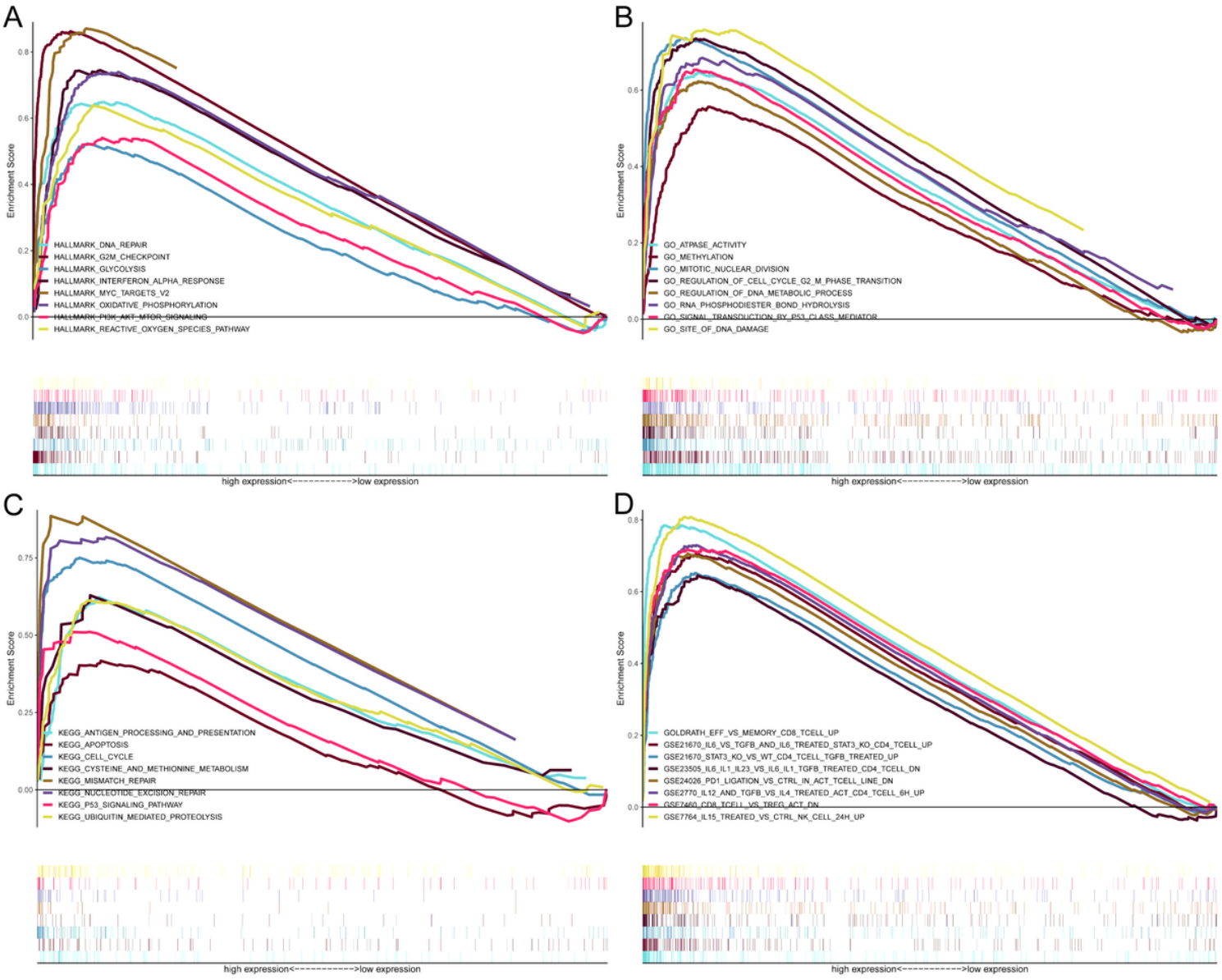

Figure 10

GSEA for samples with high-risk and low-risk based on the prognostic index of three-gene signature. (A) The enriched gene sets in GO collection by the high-risk sample. Each line representing one particular gene set with unique color, and up-regulated genes located in the left approaching the origin of the coordinates, by contrast the down-regulated lay on the right of x-axis. Only gene sets with NOM p < 0.05 and FDR q < 0.05 were considered significant. And only several leading gene sets were displayed in the plot. (B) The enriched gene sets in KEGG by samples with high-risk sample. And only several leading gene

sets were displayed in the plot. (C) Enriched gene sets in HALLMARK collection by samples of high-risk sample. Only several leading gene sets are shown in plot. (D) Enriched gene sets in C7 collection, the immunologic gene sets, by samples of high-risk sample. Only several leading gene sets are shown in plot.

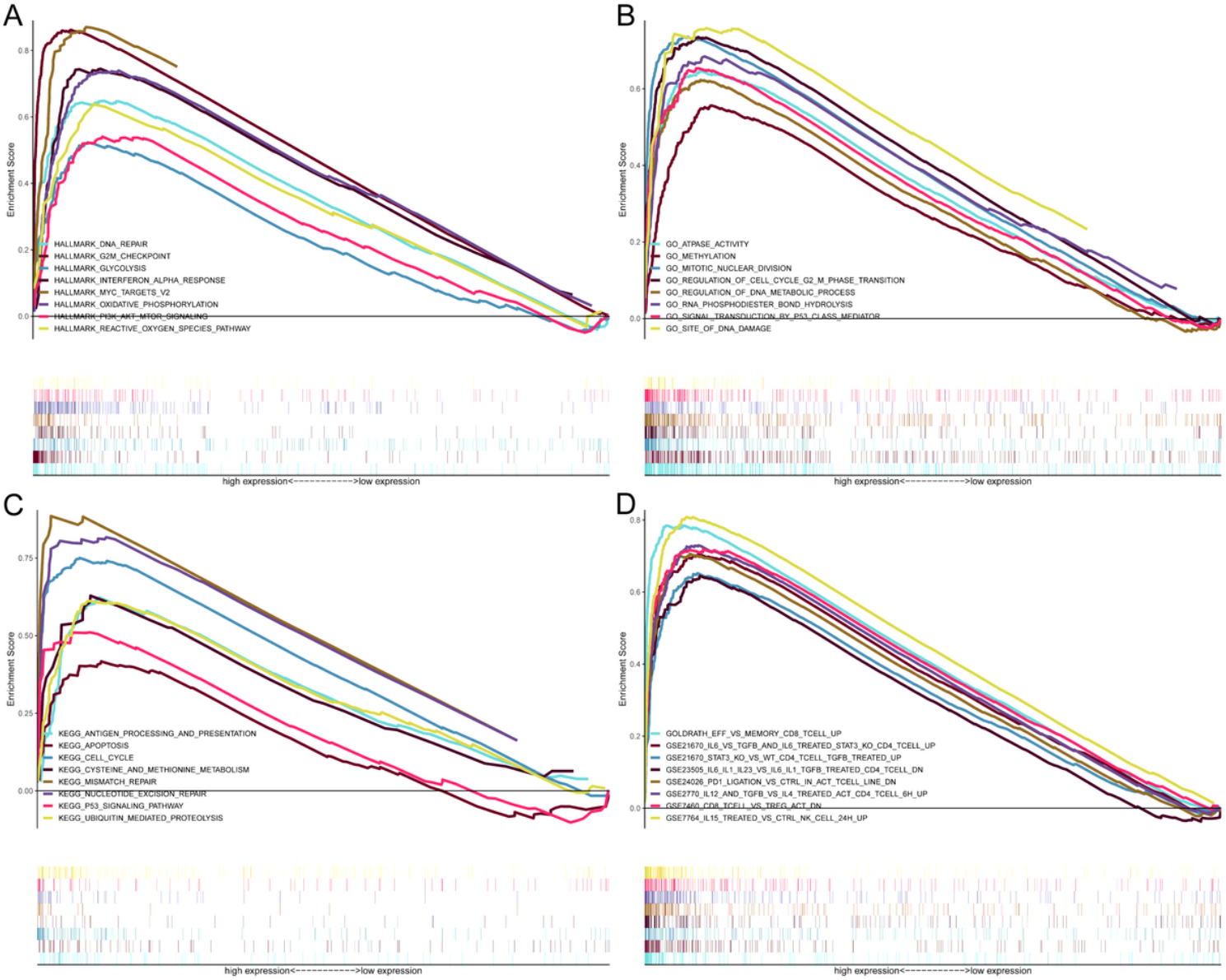

Figure 10

GSEA for samples with high-risk and low-risk based on the prognostic index of three-gene signature. (A) The enriched gene sets in GO collection by the high-risk sample. Each line representing one particular gene set with unique color, and up-regulated genes located in the left approaching the origin of the coordinates, by contrast the down-regulated lay on the right of x-axis. Only gene sets with NOM p < 0.05 and FDR q < 0.05 were considered significant. And only several leading gene sets were displayed in the plot. (B) The enriched gene sets in KEGG by samples with high-risk sample. And only several leading gene sets were displayed in the plot. (C) Enriched gene sets in HALLMARK collection by samples of high-risk sample. Only several leading gene sets are shown in plot. (D) Enriched gene sets in C7 collection, the immunologic gene sets, by samples of high-risk sample. Only several leading gene sets are shown in plot.

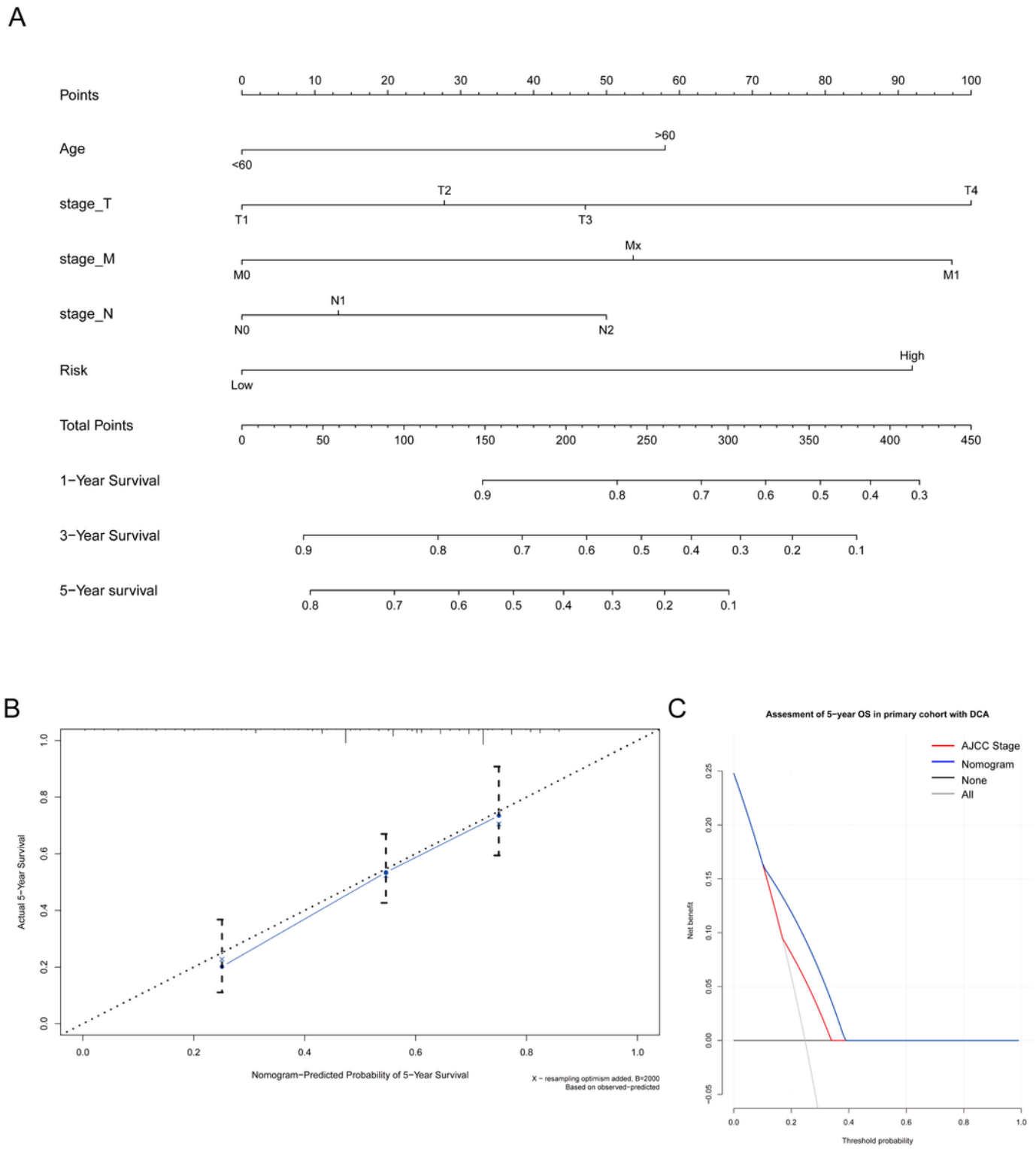

Figure 11

(A) A Nomogram for predicting 1-, 3- and 5-year OS in colon cancer. To calculate probability of OS, first determine the value for each factor by drawing a vertical line from that factor to the points scale. “Points” is a scoring scale for each factor, and “total points” is a scale for total score. Then sum all of the individual values and draw a vertical line from the total points scale to the 1-, 3- and 5-year OS probability lines to obtain OS estimates. (B) The nomogram cohort was divided into three equal groups for

validation. The gray line represents the perfect match between the actual (y-axis) and nomogram-predicted (x-axis) survival probabilities. Black circles represent nomogram-predicted probabilities for each group, and X's represent the bootstrap-corrected estimates. Error bars represent the 95% CIs of these estimates. A closer distance between two curves suggests higher accuracy. (C) The DCA of nomogram in training set for 5 years OS.

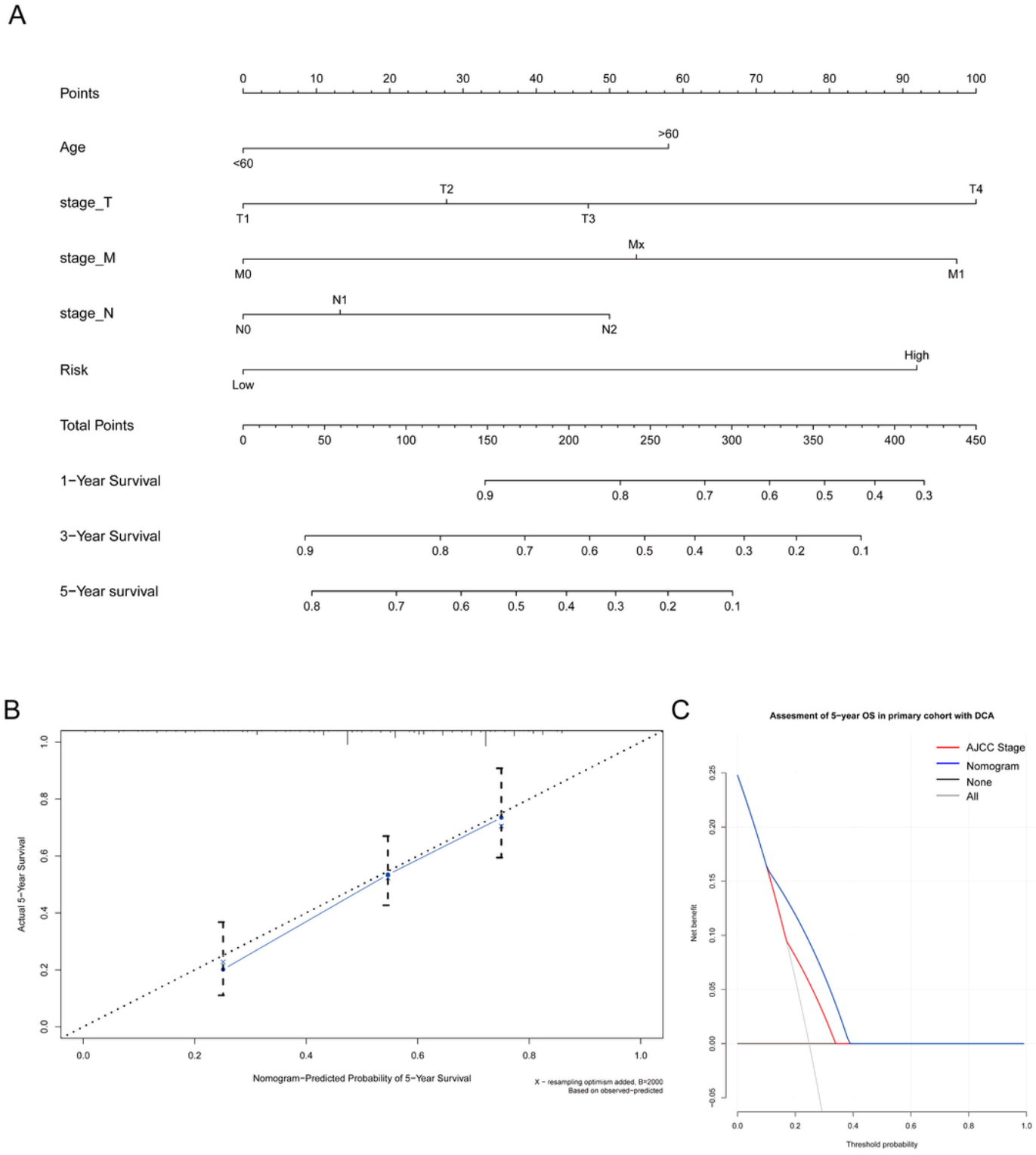

Figure 11

(A) A Nomogram for predicting 1-, 3- and 5-year OS in colon cancer. To calculate probability of OS, first determine the value for each factor by drawing a vertical line from that factor to the points scale. “Points” is a scoring scale for each factor, and “total points” is a scale for total score. Then sum all of the individual values and draw a vertical line from the total points scale to the 1-, 3- and 5-year OS probability lines to obtain OS estimates. (B) The nomogram cohort was divided into three equal groups for validation. The gray line represents the perfect match between the actual (y-axis) and nomogram-predicted (x-axis) survival probabilities. Black circles represent nomogram-predicted probabilities for each group, and X’s represent the bootstrap-corrected estimates. Error bars represent the 95% CIs of these estimates. A closer distance between two curves suggests higher accuracy. (C) The DCA of nomogram in training set for 5 years OS.

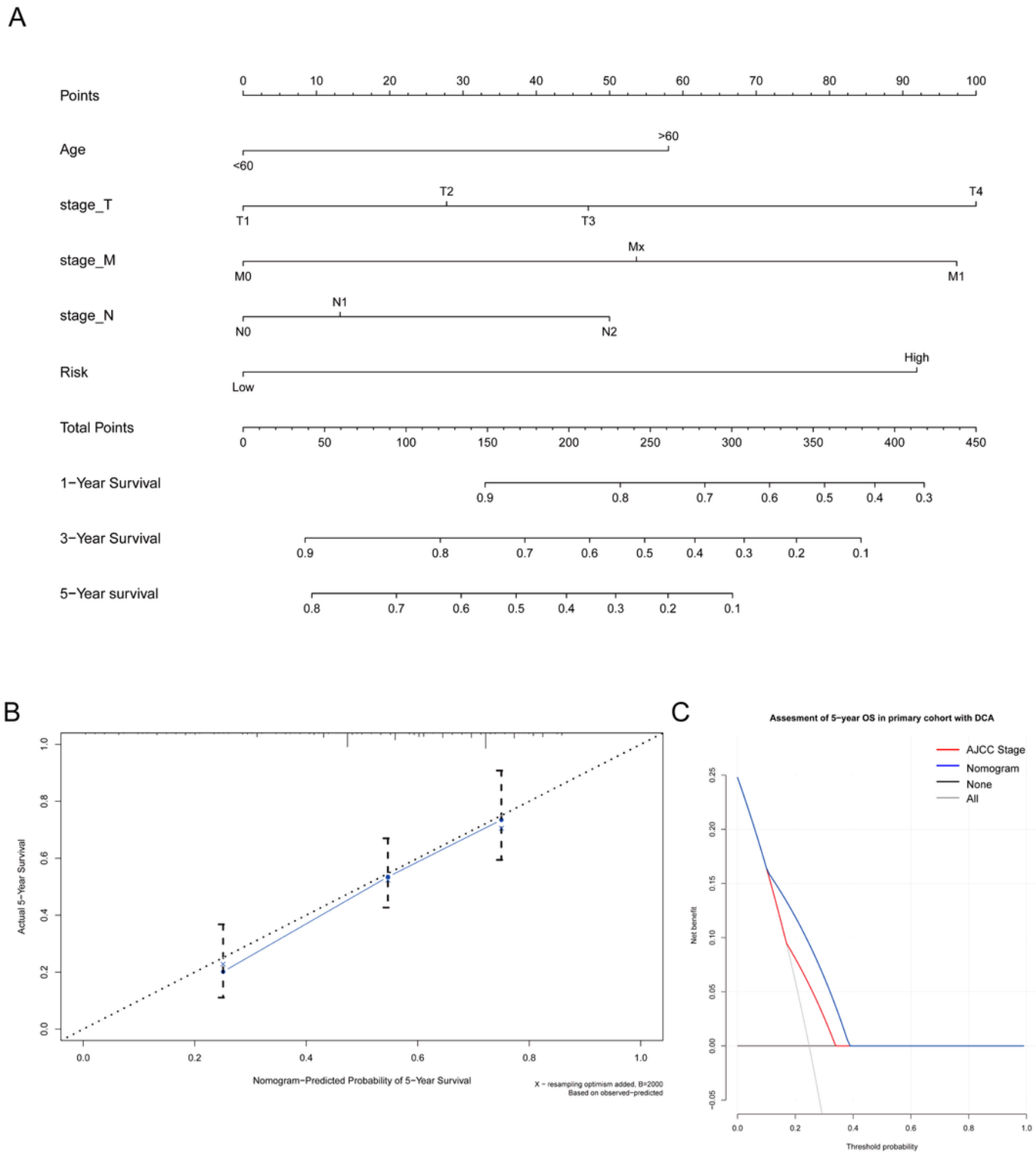

Figure 11

(A) A Nomogram for predicting 1-, 3- and 5-year OS in colon cancer. To calculate probability of OS, first determine the value for each factor by drawing a vertical line from that factor to the points scale. “Points” is a scoring scale for each factor, and “total points” is a scale for total score. Then sum all of the individual values and draw a vertical line from the total points scale to the 1-, 3- and 5-year OS probability lines to obtain OS estimates. (B) The nomogram cohort was divided into three equal groups for

validation. The gray line represents the perfect match between the actual (y-axis) and nomogram-predicted (x-axis) survival probabilities. Black circles represent nomogram-predicted probabilities for each group, and X's represent the bootstrap-corrected estimates. Error bars represent the 95% CIs of these estimates. A closer distance between two curves suggests higher accuracy. (C) The DCA of nomogram in training set for 5 years OS.

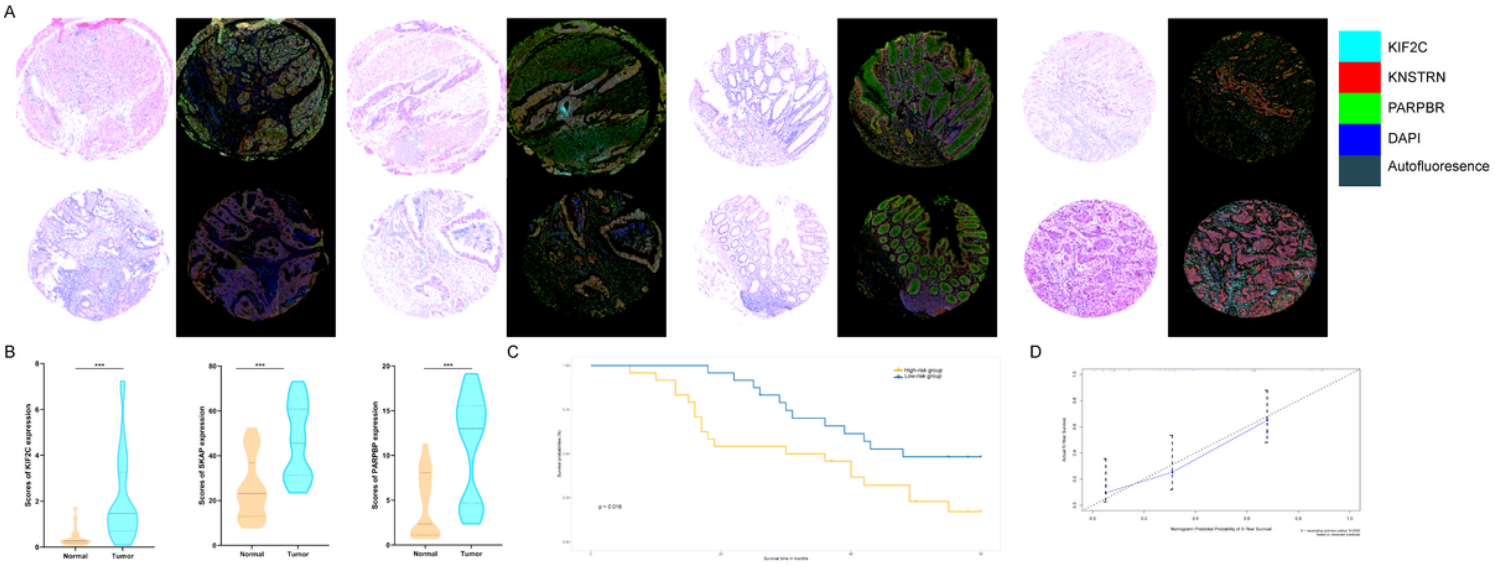

Figure 12

(A) Tumor tissue and corresponding non-tumorous adjacent tissue were collected from colorectal cancer patients in TMAs. The expression of PARPBP (green), KNSTRN (red), and KIF2C (cyan) indicates the lipid droplets in tumor-infiltrating myeloid cells. Nuclei (blue) were stained using DAPI. The absolute number of positive cells was quantified in whole fields (hpf; scale bar = 20 lm). (B) Compared the expression of PARPBP, KNSTRN, and KIF2C in tumor tissue and non-tumorous adjacent tissue based on the single index strength score (Student's two-tailed t-tests, ***p < 0.001). (C) The Kaplan–Meier test of the risk score for the overall survival between high-risk and low-risk group (log-rank test, p < 0.05). (D) The calibration curves for identifying the consistency between actual observation and nomogram-predicted OS probabilities.

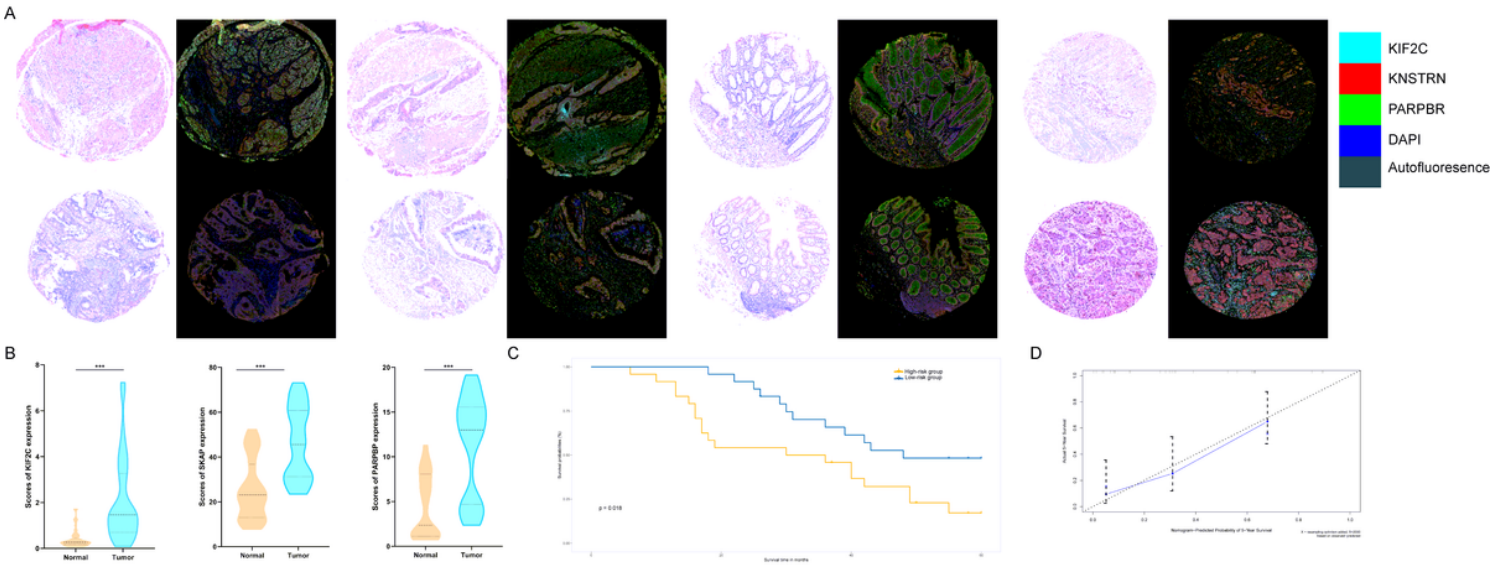

Figure 12

(A) Tumor tissue and corresponding non-tumorous adjacent tissue were collected from colorectal cancer patients in TMAs. The expression of PARPBP (green), KNSTRN (red), and KIF2C (cyan) indicates the lipid droplets in tumor-infiltrating myeloid cells. Nuclei (blue) were stained using DAPI. The absolute number of positive cells was quantified in whole fields (hpf; scale bar = 20 lm). (B) Compared the expression of PARPBP, KNSTRN, and KIF2C in tumor tissue and non-tumorous adjacent tissue based on the single index strength score (Student's two-tailed t-tests, ***p < 0.001). (C) The Kaplan–Meier test of the risk score for the overall survival between high-risk and low-risk group (log-rank test, p < 0.05). (D) The calibration curves for identifying the consistency between actual observation and nomogram-predicted OS probabilities.

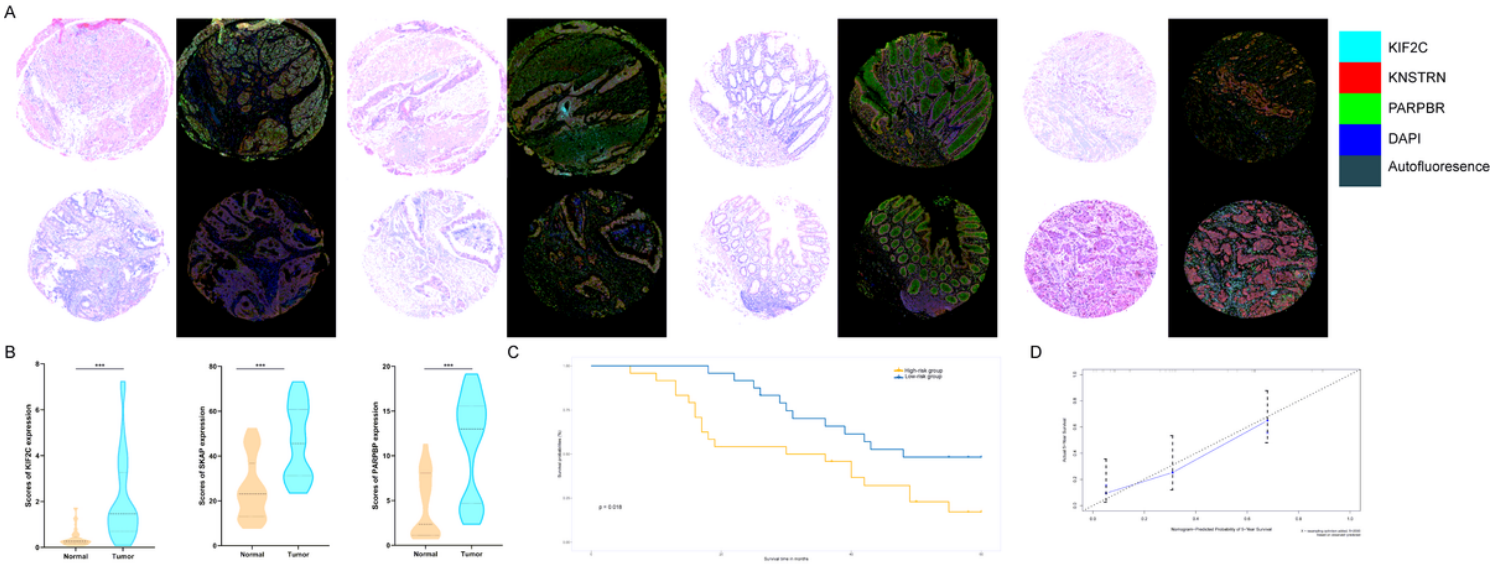

Figure 12

(A) Tumor tissue and corresponding non-tumorous adjacent tissue were collected from colorectal cancer patients in TMAs. The expression of PARPBP (green), KNSTRN (red), and KIF2C (cyan) indicates the lipid

droplets in tumor-infiltrating myeloid cells. Nuclei (blue) were stained using DAPI. The absolute number of positive cells was quantified in whole fields (hpf; scale bar = 20 lm). (B) Compared the expression of PARPBP, KNSTRN, and KIF2C in tumor tissue and non-tumorous adjacent tissue based on the single index strength score (Student's two-tailed t-tests, ***p < 0.001). (C) The Kaplan–Meier test of the risk score for the overall survival between high-risk and low-risk group (log-rank test, p < 0.05). (D) The calibration curves for identifying the consistency between actual observation and nomogram-predicted OS probabilities.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarydata202011.docx
- Supplementarydata202011.docx
- Supplementarydata202011.docx