

# NMB as a Novel Prognostic Biomarker in Colorectal Cancer Correlating with Immune Infiltrates

**Jiaxin Fan**

Shanxi Medical University

**Chaowei Liang**

Shanxi Medical University

**Jiajia Wang**

Shanxi Medical University

**Chaojie Liang** (✉ [liangchaojie8@126.com](mailto:liangchaojie8@126.com))

First Hospital of Shanxi Medical University

**Jiansheng Guo**

First Hospital of Shanxi Medical University

---

## Research Article

**Keywords:** colorectal cancer, NMB, biomarker, immune infiltration, nomogram

**Posted Date:** January 10th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1233841/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

**Background:**Neuromedin B(NMB) is associated with the occurrence and development of a variety of cancers, However, the role of NMB in colorectal cancer is lacking in further studies.

**Methods:**Transcriptome data and clinical data of CRC were downloaded and analyzed from the TCGA database and GEO database to study the differential expression of NMB. We analyzed the relationship between NMB expression and survival in patients with colorectal cancer using 8 public datasets from the Gene Expression Integration (GEO) database and the TCGA database. Meta-analysis was performed on the analysis results of TCGA and GEO data to determine the role of NMB in CRC. The receiver operating characteristic (ROC) curve was used to evaluate the accuracy of NMB in predicting survival rate in CRC patients. Wilcox. Test and Kruskal. Tests were used to study the relationship between clinicopathological features and the expression of NMB. Cox regression analysis was used to analyze the effect of NMB expression on survival. Gene collection enrichment analysis (GSEA) was performed using the TCGA database to screen the signaling pathway regulated by NMB. The Linkedomics platform was used to identify NMB co-expressed genes and explore the potential mechanisms of NMB mediation. Tumor Immune Estimation Resource (TIMER) site database was used to analyze the relationship between NMB expression level and immune infiltration. Related genes were identified by co-expression analysis, and four genes (NDUFB10, SERF2, DPP7, and NAPRT) were screened out as a prognostic signature. The relationship between risk score and OS were studied to explore the predictive value of risk score for CRC. Nomogram was constructed to predict 1 - and 3-year survival in colorectal cancer patients.

**Results:**NMB was highly expressed in colorectal cancer, suggested a poor prognosis. The ROC curve proved that NMB had a high accuracy in predicting the survival rate of CRC patients. Multivariate regression analysis demonstrated that NMB was an independent predictor of survival in patients with CRC.GSEA identified the pathways involved in NMB regulation, including the P53 Signaling pathway, VEGF Signaling pathway, JAK-STAT Signaling pathway, MAPK Signaling pathway,mTOR Signaling pathway, TGF-BETA Signaling pathway, and WNT Signaling pathway, etc. Then,6512 co-expressed genes were identified through the Linkedomics Platform to investigate the potential mechanisms of NMB regulation, including Hepatocellular carcinoma cell cycle, EGF/EGFR Signaling Pathway, VEGFA-VEGFR2 Signaling Pathway, etc. We also conclude that NMB is correlated with T cells CD8, T cells CD4 memory resting, Macrophages M0. Different mutational forms of NMB were associated with the immune infiltration of 6 leukocytes. We determined the relationship between NMB and immune marker sets in colorectal cancer, such as CCR7, CD3E, CTLA4, HAVCR2, HLA-DPB1. The predictive ability of the risk score was significantly better than that of T, N, and M stages. A new nomogram for predicting the 1-year and 3-year OS of CRC patients was constructed, showing good reliability and accuracy for improved treatment decisions. In addition, NMB may contribute to drug resistance in CRC.

**Conclusion:**NMB is highly expressed in CRC and provides a potential biomarker for the diagnosis and prognosis of CRC.

## Introduction

Cancer plays an important role in influencing morbidity and mortality in both developed and developing countries[1]. Due to the lack of understanding of the pathological process and regulatory mechanism of cancer, the mortality rate caused by cancer will be on the rise in the future.[2]Although the prognosis of cancer has been improved by various treatments, in many cancers, the prognosis has always been less than satisfactory. Among various types of cancer, colorectal cancer(CRC)is one of its major roles.CRC is the third most common cancer and the second most common cause of cancer-related death in the world.1.2 million patients are diagnosed with colorectal cancer and more than 600,000 die from the disease[3]. It has been reported that the 5-year survival rate for CRC is still very low, although there are many treatments available.[4, 5] Therefore, it is of great necessity to identify key new biomarkers and potential mechanisms associated with CRC.

NMB (Neuromedin B), a protein-coding gene, encodes members of a bomb-like family of neuropeptides that negatively regulate eating behavior. The protein encoded by this gene regulates smooth muscle contractions in the colon by binding to its cognitive receptor, the neurotransmitter B receptor (NMBR).NMBR is widespread in the gastrointestinal tract.[6]Currently, many studies have shown that NMB is closely related to the occurrence, development, recurrence, and metastasis of a variety of cancers. Including breast cancer cells[7], and lung cancer[8]. Previous studies have shown that in normal and malignant colonic epithelial cells, NMB and its receptors are co-expressed in proliferating cells in an autocrine manner[9].

In our study, we used the Cancer Genome Atlas (TCGA) database to investigate the diagnostic and prognostic value of NMB expression in CRC. In addition, we also analyzed the relationship between NMB expression and clinicopathological characteristics of patients with CRC and its prognostic value in CRC. We also further evaluated biological pathways to study the NMB regulatory pathway related to the pathogenesis of CRC by gene collection enrichment analysis (GSEA).To further investigate the underlying mechanisms of NMB regulation, we used the Linkedomics platform to identify the co-expressed genes of NMB and then investigate the regulatory pathways of these genes.TIMER database was used to study the relationship between NMB and immune infiltration. A prognostic signature was established through genes related to

NMB, and then the risk score of each patient was calculated, and the relationship between the risk score and OS was explored. This study reveals the potential role of NMB in CRC, which will help us further understand the possible mechanism of CRC.

## Materials And Methods

### Study design and data processing

Our study design was shown in a flowchart (Figure.1), including TCGA-based data collection and multiple bioinformatics analyses.

All microarray data sets (612, including 44 normal samples and 568 tumor samples) and clinical information related to survival time of 711 CRC patients were downloaded from The Cancer Genome Atlas(TCGA)(<https://portal.gdc.cancer.gov/repository>). Get the data set using GEO.Expression data and clinical information of NMB were extracted from 8 datasets(GSE17536,GSD17538,GSE29623,GSE38832,GSE40967,GSE71187,GSE87211,GSE103479).Perl programming language was used to match gene expression information with clinical information and delete unknown or incomplete clinical information.

Because sequencing data and clinical information are obtained through public databases, there are no ethical issues.

### Relative NMB expression levels between CRC and normal tissues.

survival package, beeswarm package , and limma package had been utilized for analyzing whether there is a significant difference in the expression of NMB between CRC and normal tissues.Wilcox.Test was used to determine the P-value. Boxplot was used to plot the scatter difference plot and the paired difference plot to visualize the results. R Software (v. 4.0.4) was used to analyze the above. In addition, we verified the differential expression of NMB in colorectal cancer and normal tissues through Gene expression profiling interactive analysis(GEPIA).

### Survival Analysis.

Patients were divided into two groups according to their median NMB expression level (low NMB expression group and high NMB expression group). Survival analysis was performed using survival and survival software packages and the result was visualized using the Kaplan-Meier survival curve. The survival package,survminer package, and timeROC package were used to draw a receiver operating characteristic (ROC) curve to evaluate the accuracy of NMB in predicting the survival of CRC patients. At the same time, we evaluated the relationship between NMB expression level and overall survival (OS) using the Survival module in GEPIA(Quartile NMB expression, cutoff-high, 75%; Cutoff - Low, 25%). The Kaplan-Meier (K-M) method was used to compare different survival curves(cutoff by quartile expression level of NMB, Cutoff-High:75%, Cutoff-Low:25%). The results of the TCGA and GEO databases were meta-analyzed to determine the relationship between the expression of NMB and the survival prognosis of patients with colorectal cancer.

### Clinical correlation analysis

Wilcox. Test was used to analyze the relationship between age, sex, distant metastasis , and the expression of NMB.Kruskal. test was used to analyze the relationship between T, N, stage, and the expression level of NMB.At last. We use a boxplot to visualize the results. In addition, logistic regression analysis was performed to investigate the association between NMB and clinicopathological features.

### Univariate and Multivariate Cox Regression Analyses.

Univariate and multivariate analyses were performed using Cox proportional hazard regression model. We used univariate analysis to assess the independent predictors of survival of clinicopathological parameters and NMB expression. Multivariate Cox analysis was used to evaluate whether NMB could be used as an independent prognostic factor for CRC. The survival package and Survminer package were used for data analysis, and the results were visualized using a forestplot.

### Gene Set Enrichment Analysis (GSEA)

GSEA(version4.1.0) was used to analyze the signaling pathway associated with NMB in CRC. Gene expression enrichment analysis was performed among different phenotypes determined by NMB expression level.Select the annotated gene set (c2.cp.kegg.v7.2.symbols.gmt) as reference gene sets.The sequence of gene sets was analyzed 1000 times. Normalized enrichment score (NES), nominal P-value, and false discovery rate (FDR) q-value were used to rank the enrichment paths for each group. Eventually, multiple cancer-related pathways were identified.

**To determine the co-expression genes of NMB and the potential mechanisms of NMB mediation.**

Using the Linkedomics Platform to identify the co-expressed genes of NMB. The GO\_BP/CC/MF, KEGG pathway, Wiki pathway, and Reactome pathway were analyzed by overrepresentation enrichment analysis (ORA), The GO\_BP/CC/MF, KEGG pathway were analyzed by gene set enrichment (GSEA). Forecast its potential function.

### **Analysis of association between NMB expression level and Immune Infiltrates**

To further explore the mechanism by which NMB is involved in the pathological progression of CRC, we investigated the relationship between NMB expression and immune infiltration. Firstly, the E1071 package, preprocessCore package, and Limma package were used to determine the content of immune cells in each sample, and then a histogram was plotted using the Corrplot package to show the results of the immune infiltration. We divided the patients into the high expression group and the low expression group according to the expression level of NMB, and then used the Limma package and Vioplot package to conduct the difference test of immune cells (pFilter=0.05). Then, the limma package, ggplot2 package, ggpubr package, and ggExtra package were used to test the correlation between the expression of NMB and the content of immune cells. (pFilter=0.05). Finally, we intersected the results of the two tests to obtain immune cells that were correlated with the expression of NMB. In addition, the SCNA module was used to explore the correlation between somatic copy number alteration (SCNA) of NMB and the immune abundance of 6 leukocytes.

### **Correlation analysis of NMB expression level and tumor-infiltrating immune cells gene markers**

Tumor Immune Estimation Resource (TIMER), A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. Including 10,897 samples, covering 32 cancer types from the Cancer Genome Atlas (TCGA). Correlation module was used to analyze the correlation between NMB expression level and Tumor-Infiltrating immune cells gene markers.

### **Construction and Evaluation of the Prognostic Signatures of NMB**

First, we obtained the 162 differential genes by constructing a co-expression network of NMB, (fdrFilter=0.0000001, logFCfilter=10), and then performed Univariate COX analysis on these DEGs, and determined DEGs that are significantly different from OS, (pFilter=0.05). Finally, we performed multivariate Cox regression analysis on DEGs screened by univariate COX analysis. 4 genes (NDUFB10, SERF2, DPP7, and NAPRT) were selected as a predictive signature. We used the risk score calculation formula to calculate the risk score of each patient: risk score = coef gene 1 × gene 1 expression + coef gene 2 × gene 2 expression + ... + coef gene N × gene N expression. The risk score is obtained by weighting the expression level of the gene and the regression coefficient (coef). The risk ratio (HR) of the multivariate Cox regression analysis was log-transformed to calculate the coef value, and the expression of each gene involved in the prognostic characteristics was defined as the expression of N gene. Patients were divided into high-risk and low-risk groups based on the median risk score.

The difference of OS between the high-risk group and the low-risk group was analyzed by the K-M method, and the survival curve was obtained. We then plotted Receiver Operating Characteristic (ROC) curves using the survivalROC package to assess the ability of risk score and other clinical features to predict CRC and to assess sensitivity and specificity by the area below the curve (AUC values). In order to determine whether the risk score can be used as an independent predictor of the prognosis of CRC patients, we incorporated age, gender, stage, T, M, N, and risk score into univariate and multivariate Cox regression analysis.

### **Construction and Validation of Nomogram Based on Risk Score**

Nomogram can intuitively calculate the survival rate of colorectal cancer patients, which has important value in clinical application. We screened the prognostic factors of colorectal cancer patients, constructed a nomogram using the survival package and RMS package to predict 1-year and 3-year survival rates of colorectal cancer patients, and drew calibration curves to evaluate the accuracy of the nomogram. Finally, the ROC curve of the nomogram was drawn and the area under the curve (AUC) was calculated.

### **Drug sensitivity analysis**

We investigated drug sensitivity of NMB-related genes by using Gene Set Cancer Analysis (GSCA) (<http://bioinfo.life.hust.edu.cn/web/GSCALite/>), a web-based platform for Gene Set Cancer Analysis GSCALite.

### **Statistical Analysis.**

R software (version 4.0.4) was used for statistical analysis. Wilcoxon Test and Kruskal-Wallis Test were used to analyze clinicopathological parameters and the expression level of NMB. Kaplan-Meier analysis was used to investigate the relationship between survival rate and NMB expression level. Univariate and multivariate survival analyses were performed using Cox proportional hazard regression model.

## **Results**

## The differential expression of NMB between CRC tumor tissues and normal tissues

Scatter difference plot( $p < 0.001$ , Figure.2A) and paired difference plot( $p < 0.001$ , Figure.2B) showed that the expression level of NMB in CRC tumor tissues was significantly higher than that in normal tissues. The same results were confirmed in GEPIA( $p < 0.001$ , Figure.2C).

## High expression of NMB in CRC tumor tissue suggests poor overall survival

We evaluated the association between NMB expression and prognosis in CRC patients using Kaplan-Meier risk estimates. Compared with low NMB expression, high NMB expression was associated with significantly poorer overall survival(OS)( $p < 0.001$ , Figure.2D). The same results were obtained in GEPIA( $p = 0.012$ , Figure.2E). We used expression data from 44 normal samples and 568 tumor samples to draw ROC curves(Figure.2F) to evaluate the diagnostic value of NMB in CRC. The area under the ROC curve(AUC) AUC at 1 years: 0.622; AUC at 3 years: 0.576; AUC at 5 years: 0.669. Therefore, NMB has considerable diagnostic value for CRC.

We performed survival analysis on 8 datasets in the GEO database using Kaplan-Meier risk estimates. (Figure.3A-H). Then, Meta-analysis was performed on the above analysis results( HR =1.0459,95%-CI:1.0067-1.0865, $z = 2.31$ ,  $p$ -value=0.0211). The results of the meta-analysis were visualized using the forest map(Figure.3I). Because  $I^2 < 50\%$  and  $P > 0.05$ ( $I^2 = 1\%$ ,  $P = 0.43$ ), we chose the fixed effect model. The results of this meta-analysis suggest that NMB is indeed a high-risk gene in colorectal cancer.(HR=1.05,95%-CI:1.01-1.09)

## Relationship between expression of NMB and clinicopathological features.

Clinical data from 711 patients with CRC from TCGA were analyzed and unknown and incomplete clinical information was deleted. The expression of NMB was only correlated with age( $p = 0.01$ , Figure.4A), but not with gender( $p = 0.293$ , Figure.4B), T classification( $p = 0.209$ , Figure.4C) and N classification( $p = 0.088$ , Figure.4D), distant metastasis(M)( $p = 0.838$ , Figure.4E) and clinical stage( $p = 0.919$ , Figure.4F). Similarly, Logistic regression analysis showed that the expression of NMB in colorectal cancer was only correlated with age(OR = 1.504 for  $\geq 65$  versus  $< 65$ ,  $P = 0.019$ ), and there was no significant correlation with gender(OR = 1.178 for Male versus female,  $P = 0.343$ ), T classification(OR = 0.327 for T2 versus T1,  $P = 0.071$ ; OR = 0.368 for T3 versus T1,  $P = 0.092$ ; OR = 0.310 for T4 versus T1,  $P = 0.066$ ), N classification(OR = 0.807 for N1 versus N0,  $P = 0.306$ ; OR = 1.475 for N2 versus N0,  $P = 0.104$ ), distant metastasis(M)(OR = 1.139 for M1 versus M0,  $P = 0.603$ ), and stage (OR = 0.969 for stage II versus stage I,  $P = 0.901$ ; OR = 0.927 for stage III versus stage I,  $P = 0.0.775$ ; OR = 1.059 for stage IV versus stage I,  $P = 0.855$ ) (Table 1).

## NMB Is an Independent Predictor of Poor Survival in CRC

The univariate and multivariate Cox proportional hazard regression analyses were used to evaluate whether the expression of NMB could be an independent predictor of poor survival in CRC patients. After the deletion of unknown or incomplete clinical information. We performed univariate and multivariate Cox proportional hazard regression analyses in 472 patients with CRC. The univariate analysis showed that age (HR, 1.039; 95% CI, 1.017-1.062;  $p < 0.001$ ), stage (HR, 2.293, 95% CI, 1.794-2.932,  $p < 0.001$ ), T classification (HR, 2.890; 95% CI, 1.887-4.427;  $p < 0.001$ ), lymph node (HR, 2.073; 95% CI, 1.606-2.675;  $p < 0.001$ ), distant metastasis (HR, 4.512; 95% CI, 2.888-7.049;  $p < 0.001$ ), and high-NMB expression (HR, 1.046; 95% CI, 1.005-1.088;  $p = 0.027$ ) were important predictors of survival(Figure 5A, Table 2). The multivariate analysis showed that age (HR, 1.47; 95% CI, 1.025-1.069;  $p < 0.001$ ), T classification (HR, 1.733; 95% CI, 1.068-2.813;  $p = 0.026$ ) and high-NMB expression (HR, 1.054; 95% CI, 1.008-1.102;  $p = 0.021$ ) were the important independent predictors of poor overall survival of CRC(Figure 5B and Table 2).

## GSEA identifies the NMB-related signaling pathway.

To identify differentially activated signaling pathways in CRC, we performed the Gene Set Enrichment Analysis (GSEA) between low NMB and high NMB expression datasets. Various cancer-related KEGG pathways are enriched in high NMB phenotypes(Figure 6), for example, Cell cycle, DNA replication, P53 Signaling pathway, and VEGF signaling pathway. While KEGG pathways enriched in low NMB phenotypes were Colorectal cancer, ERBB Signaling pathway, JAK-STAT Signaling pathway, MAPK Signaling pathway, MTOR Signaling pathway, NOTCH Signaling pathway, pancreatic cancer, pathways in cancer, TGF-BETA Signaling pathway, and WNT Signaling pathway.

## Enrichment analysis of NMB co-expressed genes

The co-expressed genes have similar functions and mechanisms. To further investigate the underlying mechanisms of NMB regulation, we used the Linkedomics Platform to identify 6512 significant genes. The volcano map (Figure 7A) shows that there is a correlation between global genes and NMB by Pearson test. Heat maps (Figure 7B) show the top 50 genes in CRC that are negatively and positively correlated with NMB. ORA showed that co-expressed genes were involved in Hepatocellular carcinoma cell cycle, RNA transport, RNA processing, Metabolism of RNA, negative regulation of gene expression, chromosome, EGF/EGFR Signaling Pathway, VEGFA-VEGFR2 Signaling Pathway, TGF-beta Signaling Pathway, Negative regulation of NOTCH4 signaling, Gene expression (Transcription), and Metabolism of proteins(Figure 8). Then, GSEA was performed to investigate the potential functions and pathways of NMB induction. We explored three main types of GO enrichment:

biological process (BP), cellular component (CC), and molecular function (MF). In the BP category, we explored ribonucleoprotein complex biogenesis, tRNA metabolic process, generation of precursor metabolites and energy, establishment of protein localization to membrane, protein-containing complex disassembly, regulation of GTPase activity, regulation of vasculature development, cell-cell adhesion via plasma-membrane adhesion molecules, peptidyl-serine modification, and Ras protein signal transduction(Figure 9A).In the CC category, we explored mitochondrial inner membrane, mitochondrial matrix, cytosolic part, condensed chromosome, cell-cell junction, early endosome, cell leading edge, and apical part of cell(Figure 9B).In the MF category, we explored structural constituent of ribosome, catalytic activity, acting on RNA, lyase activity, guanyl-nucleotide exchange factor activity, cofactor binding, catalytic activity, acting on DNA, modification-dependent protein binding, protein serine/threonine kinase activity, nucleoside-triphosphatase regulator activity, and phospholipid binding(Figure 9C).For KEGG pathway analysis, we explored Ribosome, Proteasome, Spliceosome, Metabolic pathways, RNA transport, Parathyroid hormone synthesis, secretion and action,Proteoglycans in cancer, ECM-receptor interaction, Ras signaling pathway, and JAK-STAT signaling pathway. (Figure 9D)

### Relationship Between NMB Expression and Immune Infiltration in CRC

Tumor infiltrating lymphocytes is an independent prognostic factor for survival in cancer patients.Like breast cancer[10], ovarian cancer, colorectal cancer, and gastric cancer. Therefore, we investigated the relationship between NMB expression and Immune Infiltration in colorectal cancer. The histogram shows the number of immune cells in each sample(Figure 10A). The difference test between the expression of NMB and immune cells showed that T cells CD8( $p < 0.001$ ), T cells CD4 memory resting( $p = 0.023$ ), NK cells activated( $p = 0.029$ ), and Macrophages M0 ( $p = 0.016$ )were different in the high expression group and the low expression group(Figure 10B).Correlation test showed that the expression of NMB was correlated with T cells CD8( $R = 0.24, p < 0.001$ ),T cells CD4 naive( $R = -0.14, p = 0.014$ ),T cells CD4 memory resting( $R = -0.14, p = 0.011$ ),T cells CD4 memory activated( $R = 0.16, p = 0.0029$ ),T cells follicular helper( $R = 0.19, p < 0.001$ ),Macrophages M0( $R = -0.13, p = 0.017$ ) and Macrophages M1( $R = 0.12, p = 0.026$ ).After the intersection of the two test results, it can be concluded that the expression level of NMB is significantly correlated with T cells CD8, T cells CD4 memory resting, Macrophages M0(Figure 10C).Besides,different mutational forms of NMB in CRC were associated with immune infiltration of 6 leukocytes(B cell,CD4+ T cells,CD8+ T cells,macrophage,neutrophil,dendritic).So it follows that NMB plays an important role in immune infiltration in CRC.

### Correlation Between Expression Level of NMB and Immune Marker Sets

To further investigate the relationship between NMB expression and immunoinfiltrating cells in colorectal cancer,we used TIMER database to detect the immune markers of T cells,CD8 + T cells, B cells,monocytes,neutrophils, NK cells, TAMs,M1 and M2 macrophages, and dendritic cells.Then,we also analyzed T cells with different functions, such as Th1 cells, Th2 cells, Tregs, Tfh cells, Th17 cells, and depleted T cells.Our results show that the expression level of NMB in CRC is closely related to immune marker sets in most immune cells(Table 3).After adjustment for correlation of tumor purity,It turns out that CD3E,HLA-DPB1,CD3D,CD79A,TGFB1,CD2,HLA-DPA1,ITGAX,HLA-DRA,CD86,CCR7,CD19,NRP1,HAVCR2,CTLA4,TBX21 showed a significant correlation with NMB expression in colon cancer( $P < 0.001$ ; Cor value  $\geq 0.40$ ).And in rectal cancer,CD3E,HAVCR2,IL10,CD1C,CEACAM8,CCR7,STAT3,CTLA4,ITGAM,HLA-DPB1,FOXP3,CCL2,KIR3DL1,STAT1 showed a significant correlation with NMB expression( $P < 0.001$ ; Cor value  $\geq 0.40$ ).

### Construction of the Prognostic Signature for CRC Patients

Using the survival information of CRC patients to perform univariate Cox regression analysis on the 162 DEGs, it was found that there are 6 DEGs with significant prognostic differences. Perform multivariate Cox analysis on 6 genes with prognostic significance, and construct a prognostic signature composed of 4 genes, including NDUFB10, SERF2, DPP7, and NAPRT.Based on the prognosis signature, the risk score calculation formula was obtained: risk score =  $(0.547443833 \times \text{DPP7 expression}) + (-1.222239025 \times \text{NDUFB10 expression}) + (0.400602782 \times \text{NAPRT expression}) + (0.792523684 \times \text{SERF2 expression})$ .The risk score was calculated for each CRC patient, and patients were divided into high a risk-group ( $n = 270$ ) and a low-risk group ( $n = 270$ ) based on the median.We constructed a heat map to show the expression of 4 genes in the high-risk group and the low-risk group, and the expression of 4 genes in the high-risk patients was higher than that in the low-risk patients(Figure.11A).Figure.11B shows the distribution of risk scores for CRC patients. Patients are divided into two groups, with risk scores increasing from left to right.Figure.11C shows the distribution of risk scores for CRC patients. Patients are divided into two groups, with risk scores increasing from left to right. K-M curve was used to compare the difference in OS time between the high-risk group and the low-risk group (Figure.11D). Results showed that CRC patients with a high-risk score had significantly lower OS than GC patients with a low-risk score( $P < 0.001$ ). We plotted a time-dependent ROC curve to predict survival in CRC patients, showing that the risk score had high sensitivity and specificity.AUC of risk score(AUC=0.711) was higher than that of age(AUC=0.646),gender(AUC=0.433),stage(AUC=0.709),T stage(AUC=0.673),N stage(AUC=0.656) and M stage(AUC=0.647) (Figure.11E).Figure.11F reflects the univariate Cox analysis of the relationship between the clinical features, risk score, and OS of CRC patients.Age( $P < 0.001$ ),stage( $P < 0.001$ ),T( $P < 0.001$ ),N( $P < 0.001$ ),M( $P < 0.001$ ) and risk score( $P < 0.001$ ) significantly affect the prognosis of GC patients.Figure.11G reflects a multivariate Cox analyzed the relationship between the clinical features, risk score, and OS of GC patients. Age ( $P < 0.001$ ), T( $P = 0.019$ ) and risk score ( $P < 0.001$ ) are independent prognostic risk factors for CRC.

### Construction and Validation of the Nomogram

We used factors such as age, stage, T, M, N, and risk score to construct a nomogram to predict the survival rate of CRC patients more conveniently (Figure.11H). According to the nomogram, the scores of CRC patients are calculated and then added to obtain the total score, thereby predicting the survival probability of 1 year and 3 years, which is beneficial to guide clinical decision-making. Because the closer the calibration curve is to the diagonal, the more accurate the prediction result will be. The calibration curves of the nomogram show that the nomogram has good accuracy in predicting survival rates at 1 and 3 years (Figure.11I, J). The 1-year (AUC = 0.711) and 3-year (AUC = 0.712) ROC curves also show that the forecasting ability of the nomogram is very accurate (Figure.11K).

### Analysis of drug sensitivity

We selected 162 NMB-related genes through the co-expression network ( $\text{fdrFilter}=0.0000001, \text{logFCfilter}=10$ ), and performed Spearman correlation analysis with small molecule/drug sensitivity (IC<sub>50</sub>) to explore the correlation between DEGs and drug sensitivity. The results showed that NMB was significantly related to the drug resistance of many chemotherapeutic drugs and tumor-targeted drugs, including 5-Fluorouracil, Methotrexate, Belinostat, CUDC-101, vorinostat, and so on (Figure.12).

## Discussion

Although cancer deaths have continued to decline since 1991 [4], Cancer continues to play an important role in influencing human morbidity and mortality [1]. Colorectal cancer is the most common tumor of the digestive system, and the diagnosis of colorectal cancer currently relies mainly on Colonoscopy [11]. However, this diagnostic method has certain limitations [12]. Early diagnosis of cancer is key to improving overall survival, reducing disease-free progression, and reducing the risk of recurrence. Christopher E Barbieri's research shows that precision medicine has the greatest potential to affect the health of patients. Define reliable predictive biomarkers and new therapeutic targets to truly improve the prognosis of patients. [13] In addition, Fang-Ze Wei's research shows that CLCA1 can be used as a prognostic marker of CRC and is related to immune infiltration. It may be a potential therapeutic target for CRC to improve the prognosis of patients [14]. Wei Xu's research also shows that Circulating lncRNA SNHG11 is related to the early diagnosis and prognosis of colorectal cancer [15]. These studies have proven that biomarkers play a key role in the early identification of colorectal cancer, helping to predict disease progression and response to treatment [16]. Therefore, it is necessary to research colorectal cancer biomarkers. Matusiak et al. used RT-qPCR to investigate the expression of the NMB gene in colon cancer cell lines (aco-2 and HT-29 cells), as well as normal colon epithelial cells (NCM-460). The results showed that the expression level of NMB in colon cancer cell lines was significantly higher than that in colon epithelial cells. In addition, the results of immunohistochemistry also showed that the expression of NMB in colon cancer cell lines was higher than that in normal colon epithelium. [9] However, no studies have been conducted on the diagnostic and prognostic value of NMB in colorectal cancer. Therefore, the potential role of NMB in CRC is the focus of this study.

In our study, we downloaded CRC transcriptome data (612, including 44 normal samples and 568 tumor samples) and clinical information related to survival time of 711 CRC patients from the TCGA database. Scatter plots and paired plots were drawn by using the Survival package, Beeswarm package and Limma package, and the results showed that the expression level of NMB in colorectal cancer was significantly higher than that in normal tissues. In addition, we confirmed the differential expression of NMB in colorectal cancer and normal tissues through GEPIA. Then, the K-M curve was drawn to analyze the relationship between NMB expression and OS, and the ROC curve was drawn to analyze the predictive value of NMB expression in colorectal cancer prognosis. The results showed that high expression of NMB was significantly associated with poor overall survival. We performed a meta-analysis on the survival analysis results of the TCGA database and eight datasets from the GEO database. due to  $I^2 < 50\%$  and  $P > 0.05$  ( $I^2 = 1\%$ ,  $P = 0.43$ ), we chose a fixed-effect model. The results of the meta-analysis showed that NMB was indeed a high-risk gene for colorectal cancer (HR=1.05, 95%-CI: 1.01-1.09). To investigate the relationship between NMB expression and clinicopathological features, histograms were drawn and logistic regression analysis was performed. The results showed that the expression level of NMB in colorectal cancer was only related to age and had no significant association with other clinicopathologic features. We then used Cox proportional hazard regression analyses to conclude that NMB was the important independent predictor of poor overall survival of CRC. The differential expression of NMB can provide a new perspective for the process of studying CRC and serve as a meaningful diagnostic biomarker for CRC.

To investigate the potential regulatory mechanism of NMB in colorectal cancer. KEGG analysis was performed using GSEA4.1.0. The results showed many well-known KEGG pathways related to cancer. such as Cell cycle, DNA replication, P53 Signaling pathway [17], and VEGF Signaling pathway [18], were enriched in high NMB phenotypes. KEGG signaling pathways were enriched in low NMB phenotypes, including Colorectal cancer, ERBB Signaling pathway [19], JAK-STAT Signaling pathway [20], MAPK Signaling pathway [21], MTOR Signaling pathway [22], NOTCH Signaling pathway [23], pancreatic cancer, pathways in cancer, TGF-BETA Signaling pathway [24], and WNT Signaling pathway [25].

Through the Linkedomics Platform, we identified the co-expressed genes of NMB for further study. By ORA and GSEA analysis of these co-expressed genes, we have gained a further understanding of the underlying regulatory mechanisms of NMB. ORA showed that co-expressed genes were involved in a variety of cancer-related pathways, such as Hepatocellular carcinoma cell cycle, EGF/EGFR Signaling Pathway [26],

VEGFA-VEGFR2 Signaling Pathway[27], TGF-beta Signaling Pathway[24], Negative regulation of NOTCH4 signaling. There is already evidence that morphine promotes tumorigenesis and cetuximab resistance through EGFR signal activation in human colorectal cancer[28]. DONG CHUL KIM's research shows that Notch-4 has a significant correlation with P2Y2R, and P2Y2R plays an important role in tumor progression and metastasis[29].

By using GSEA to perform three types of GO enrichment analysis, our study identified the regulatory mechanisms related to cancer. Such as tRNA metabolic process, regulation of GTPase activity, regulation of vasculature development, cell-cell adhesion via plasma-membrane adhesion molecules, Ras protein signal transduction, guanyl-nucleotide exchange factor activity, protein serine/threonine kinase activity, and phospholipid binding. Related studies have reported that the tRNA-yW Synthesizing Protein 2 (TYW2) undergoes promoter hypermethylation-associated transcriptional silencing in human cancer, particularly in colorectal tumors[30]. The research data of Tezcan G, Garanina EE et al. showed the role of Rab5 in the activation of inflammasomes, suggesting that this GTPase may be a potential therapeutic target for inhibiting inflammation in CRC [31]. The study of Shun Li et al. found that the vascular system in advanced tumors is more abundant and further enhances the proliferation of cancer cells[32]. An existing study confirmed that cell-cell adhesion leads to increased cell migration and invasion[33]. The study by Urosevic J et al. reported that genes regulated by RAS-ERK1/2 signaling mediate the recurrence of CRC [34]. As a guanine nucleotide exchange factor (GEF), GEF16 was confirmed by Bei Yu and others to have the ability to stimulate in vitro proliferation and migration in colorectal cancer cells [35]. The study by Wing-Hung Chan et al. reported that receptor tyrosine kinase fusion is an important alternative driving factor for serological pathways in the development of colorectal cancer[36]. Daojun Hu reported that Annexin A2, as a calcium-dependent phospholipid binding protein, can be used as a non-invasive and promising biomarker for the diagnosis of CRC, and the combined detection with CEA has good clinical application value in the diagnosis of CRC[37]. KEGG pathway analysis using GSEA showed that Proteoglycans in cancer, ECM-receptor interaction[38], Ras signaling pathway[39], and JAK-STAT signaling pathway[20] were enriched. Taken together, these results indicate that NMB is involved in the tumor-related KEGG pathway and GO pathway.

To further study the relationship between NMB expression and prognosis of colorectal cancer. We investigated the relationship between NMB and immune infiltration and immune markers. We tested the difference and correlation between the expression of NMB and immune cells, and then intersected the results of the two tests. The results showed that the expression of NMB was significantly correlated with T cells CD8, T cells CD4 memory resting, Macrophages M0. Finally, we analyzed the relationship between NMB expression and immune marker sets with TIMER. It turns out that CD3E,HLA-DPB1,CD3D,CD79A,TGFB1,CD2,HLA-DPA1,ITGAX,HLA-DRA,CD86,CCR7,CD19,NRP1,HAVCR2,CTLA4,TBX21 showed a significant correlation with NMB expression in colon cancer( $P < 0.001$ ; Cor value  $\geq 0.40$ ). And in rectal cancer,CD3E,HAVCR2,IL10,CD1C,CEACAM8,CCR7,STAT3,CTLA4,ITGAM,HLA-DPB1,FOXP3,CCL2,KIR3DL1,STAT1 showed a significant correlation with NMB expression( $P < 0.001$ ; Cor value  $\geq 0.40$ ). The research of Yang Xiong et al. showed that CCR7 promotes lymph node metastasis of urinary bladder cancer (UBC) through lymphangiogenic effects and the interaction of CCL21/CCR7 promotes the migration and invasion of UBC cells through the MEK/ERK1/2 signaling pathway[40]. These correlations may indicate the potential mechanism of NMB in regulating immune cell function in colorectal cancer.

We can calculate the risk score of each patient based on the prognostic characteristics. The study found that risk score is an independent risk factor affecting prognosis and can be used to predict OS in CRC. High-risk patients have a poor prognosis for lower-risk patients. Therefore, we speculate that the 4 DEGs that constitute prognostic signals are involved in the progression of CRC. Through the ROC curve analysis of OS of CRC patients, we found that this prognostic signature has a good predictive value for CRC (AUC = 0.711) and can be used to predict the prognosis of CRC patients. To facilitate clinical application and better predict the prognosis of patients with CRC, we constructed a nomogram to predict the survival rate of patients at 1 and 3 years. Judging from the calibration curve and ROC curve, the nomogram has a better predictive effect.

At present, studies have confirmed that abnormal gene expression is related to chemotherapy resistance[41, 42]. We used 162 DEGs to study in the GSCA database, and the results showed that the high expression of NMB may be related to the resistance of CRC chemotherapy and targeted therapy drugs, such as 5-Fluorouracil[43], Methotrexate[44], Belinostat[45], CUDC-101[46], vorinostat[45] and so on. Therefore, NMB can provide a new perspective on the research process of CRC and serve as a meaningful diagnostic biomarker for CRC.

In conclusion, we identified NMB as a prognostic marker for CRC and as associated with immune infiltration. It may be a potential therapeutic target for CRC. To better predict the survival rate of patients for clinical application, we have constructed a nomogram with a higher predictive value. Our bioinformatics studies based on public databases provide directions for in vitro trials, improve the diagnosis rate of colorectal cancer, and provide new strategies for gene-targeted therapy to better prevent and treat colorectal cancer.

## Declarations

### Ethics approval and consent to participate

Not applicable

## Consent for publication

Not applicable

## Availability of data and materials

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article.

## Competing interests

The authors declare that they have no competing interests.

## Funding

Youth Science and Technology Research Fund of Shanxi Province 201901D211480

## Authors' contributions

Jiaxin Fan and Chaojie Liang: conception and design. Jiaxin Fan and Jiajia Wang: acquisition, analysis, and interpretation of data. Jiaxin Fan and Chaowei Liang: figures drawing. Jiaxin Fan and Jiansheng Guo: writing and revision of manuscript. Jiansheng Guo: study supervision. All authors read and approved the final manuscript.

## Acknowledgements

My deepest gratitude is my teacher, Professor Liang Chaojie, for his continuous encouragement. He guided me through all stages of this thesis. Without his consistent and enlightening guidance, this paper would not have reached its current form. I am also very grateful to Teacher Huang He and Liang Chaowei for their guidance and assistance in the past two years. Of course, thanks to the support of the Youth Science and Technology Research Fund. Finally, I would like to thank my sweetheart, Ms. Wang Jiajia, for her care and great trust in me over the years, for giving me help and time during the difficult process of my thesis, listening to my opinions, and helping me to solve the problem.

## References

1. Wu S, Powers S, Zhu W, et al. Substantial contribution of extrinsic risk factors to cancer development[J]. *Nature*, 2016,529(7584):43-47. DOI: 10.1038/nature16166.
2. Russnes HG, Lønning PE, Børresen-Dale AL, et al. The multitude of molecular analyses in cancer: the opening of Pandora's box[J]. *Genome Biol*, 2014,15(9):447. DOI: 10.1186/s13059-014-0447-6.
3. Brenner H, Kloor M, Pox CP. Colorectal cancer[J]. *Lancet*, 2014,383(9927):1490-1502. DOI: 10.1016/S0140-6736(13)61649-9.
4. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019[J]. *CA Cancer J Clin*, 2019,69(1):7-34. DOI: 10.3322/caac.21551.
5. Wang W, Kandimalla R, Huang H, et al. Molecular subtyping of colorectal cancer: Recent progress, new challenges and emerging opportunities[J]. *Semin Cancer Biol*, 2019,55:37-52. DOI: 10.1016/j.semcancer.2018.05.002.
6. Ma Z, Su J, Guo T, et al. Neuromedin B and Its Receptor: Gene Cloning, Tissue Distribution and Expression Levels of the Reproductive Axis in Pigs[J]. *PLoS One*, 2016,11(3):e0151871. DOI: 10.1371/journal.pone.0151871.
7. Park HJ, Kim MK, Choi KS, et al. Neuromedin B receptor antagonism inhibits migration, invasion, and epithelial-mesenchymal transition of breast cancer cells[J]. *Int J Oncol*, 2016,49(3):934-942. DOI: 10.3892/ijo.2016.3590.
8. Moody TW, Berna MJ, Mantey S, et al. Neuromedin B receptors regulate EGF receptor tyrosine phosphorylation in lung cancer cells[J]. *Eur J Pharmacol*, 2010,637(1-3):38-45. DOI: 10.1016/j.ejphar.2010.03.057.
9. Matusiak D, Glover S, Nathaniel R, et al. Neuromedin B and its receptor are mitogens in both normal and malignant epithelial cells lining the colon[J]. *Am J Physiol Gastrointest Liver Physiol*, 2005,288(4):G718-728. DOI: 10.1152/ajpgi.00156.2004.
10. Stanton SE, Disis ML. Clinical significance of tumor-infiltrating lymphocytes in breast cancer[J]. *J Immunother Cancer*, 2016,4:59. DOI: 10.1186/s40425-016-0165-6.
11. Zhou D, Tian F, Tian X, et al. Diagnostic evaluation of a deep learning model for optical diagnosis of colorectal cancer[J]. *Nat Commun*, 2020,11(1):2961. DOI: 10.1038/s41467-020-16777-6.
12. Lech Pedersen N, Mertz Petersen M, Ladd JJ, et al. Development of blood-based biomarker tests for early detection of colorectal neoplasia: Influence of blood collection timing and handling procedures[J]. *Clin Chim Acta*, 2020,507:39-53. DOI: 10.1016/j.cca.2020.03.035.

13. Barbieri CE, Chinnaiyan AM, Lerner SP, et al. The Emergence of Precision Urologic Oncology: A Collaborative Review on Biomarker-driven Therapeutics[J]. *Eur Urol*, 2017,71(2):237-246. DOI: 10.1016/j.eururo.2016.08.024.
14. Wei FZ, Mei SW, Wang ZJ, et al. Differential Expression Analysis Revealing CLCA1 to Be a Prognostic and Diagnostic Biomarker for Colorectal Cancer[J]. *Front Oncol*, 2020,10:573295. DOI: 10.3389/fonc.2020.573295.
15. Xu W, Zhou G, Wang H, et al. Circulating lncRNA SNHG11 as a novel biomarker for early diagnosis and prognosis of colorectal cancer[J]. *Int J Cancer*, 2020,146(10):2901-2912. DOI: 10.1002/ijc.32747.
16. Ogunwobi OO, Mahmood F, Akingboye A. Biomarkers in Colorectal Cancer: Current Research and Future Prospects[J]. *Int J Mol Sci*, 2020,21(15)DOI: 10.3390/ijms21155311.
17. Yin H, Xie J, Jiang P, et al. Chiauranib selectively inhibits colorectal cancer with KRAS wild-type by modulation of ROS through activating the p53 signaling pathway[J]. *Am J Cancer Res*, 2020,10(11):3666-3685.
18. Yue YC, Yang BY, Lu J, et al. Metabolite secretions of *Lactobacillus plantarum* YYC-3 may inhibit colon cancer cell metastasis by suppressing the VEGF-MMP2/9 signaling pathway[J]. *Microb Cell Fact*, 2020,19(1):213. DOI: 10.1186/s12934-020-01466-2.
19. Imoto H, Zhang S, Okada M. A Computational Framework for Prediction and Analysis of Cancer Signaling Dynamics from RNA Sequencing Data-Application to the ErbB Receptor Signaling Pathway[J]. *Cancers (Basel)*, 2020,12(10)DOI: 10.3390/cancers12102878.
20. Yue Y, Zhang Q, Wu S, et al. Identification of key genes involved in JAK/STAT pathway in colorectal cancer[J]. *Mol Immunol*, 2020,128:287-297. DOI: 10.1016/j.molimm.2020.10.007.
21. Lei L, An G, Zhu Z, et al. C8orf48 inhibits the tumorigenesis of colorectal cancer by regulating the MAPK signaling pathway[J]. *Life Sci*, 2021,266:118872. DOI: 10.1016/j.lfs.2020.118872.
22. Cheng H, Jiang X, Zhang Q, et al. Naringin inhibits colorectal cancer cell growth by repressing the PI3K/AKT/mTOR signaling pathway[J]. *Exp Ther Med*, 2020,19(6):3798-3804. DOI: 10.3892/etm.2020.8649.
23. Tyagi A, Sharma AK, Damodaran C. A Review on Notch Signaling and Colorectal Cancer[J]. *Cells*, 2020,9(6)DOI: 10.3390/cells9061549.
24. Lv X, Zhang J, Zhang J, et al. A Negative Feedback Loop Between NAMPT and TGF- $\beta$  Signaling Pathway in Colorectal Cancer Cells[J]. *Onco Targets Ther*, 2021,14:187-198. DOI: 10.2147/OTT.S282367.
25. Bian J, Dannappel M, Wan C, et al. Transcriptional Regulation of Wnt/ $\beta$ -Catenin Pathway in Colorectal Cancer[J]. *Cells*, 2020,9(9)DOI: 10.3390/cells9092125.
26. Yang YC, Pan KF, Lee WJ, et al. Circulating Proteoglycan Endocan Mediates EGFR-Driven Progression of Non-Small Cell Lung Cancer[J]. *Cancer Res*, 2020,80(16):3292-3304. DOI: 10.1158/0008-5472.CAN-20-0005.
27. Refolo MG, Lotesoriere C, Lolli IR, et al. Molecular mechanisms of synergistic action of Ramucirumab and Paclitaxel in Gastric Cancers cell lines[J]. *Sci Rep*, 2020,10(1):7162. DOI: 10.1038/s41598-020-64195-x.
28. Lu H, Zhang H, Weng ML, et al. Morphine promotes tumorigenesis and cetuximab resistance via EGFR signaling activation in human colorectal cancer[J]. *J Cell Physiol*, 2021,236(6):4445-4454. DOI: 10.1002/jcp.30161.
29. Kim DC, Jin H, Lee JS, et al. P2Y(2)R has a significant correlation with Notch-4 in patients with breast cancer[J]. *Oncol Lett*, 2020,20(1):647-654. DOI: 10.3892/ol.2020.11630.
30. Rosselló-Tortella M, Llinàs-Arias P, Sakaguchi Y, et al. Epigenetic loss of the transfer RNA-modifying enzyme TYW2 induces ribosome frameshifts in colon cancer[J]. *Proc Natl Acad Sci U S A*, 2020,117(34):20785-20793. DOI: 10.1073/pnas.2003358117.
31. Tezcan G, Garanina EE, Zhuravleva MN, et al. Rab GTPase Mediating Regulation of NALP3 in Colorectal Cancer[J]. *Molecules*, 2020,25(20)DOI: 10.3390/molecules25204834.
32. Li S, Liu M, Do MH, et al. Cancer immunotherapy via targeted TGF- $\beta$  signalling blockade in T(H) cells[J]. *Nature*, 2020,587(7832):121-125. DOI: 10.1038/s41586-020-2850-3.
33. Amable G, Martínez-León E, Picco ME, et al. Metformin inhibition of colorectal cancer cell migration is associated with rebuilt adherens junctions and FAK downregulation[J]. *J Cell Physiol*, 2020,235(11):8334-8344. DOI: 10.1002/jcp.29677.
34. Urosevic J, Blasco MT, Llorente A, et al. ERK1/2 Signaling Induces Upregulation of ANGPT2 and CXCR4 to Mediate Liver Metastasis in Colon Cancer[J]. *Cancer Res*, 2020,80(21):4668-4680. DOI: 10.1158/0008-5472.CAN-19-4028.
35. Yu B, Xu L, Chen L, et al. FYN is required for ARHGEF16 to promote proliferation and migration in colon cancer cells[J]. *Cell Death Dis*, 2020,11(8):652. DOI: 10.1038/s41419-020-02830-1.
36. Chan AW, Pan Y, Tong JH, et al. Receptor tyrosine kinase fusions act as a significant alternative driver of the serrated pathway in colorectal cancer development[J]. *J Pathol*, 2020,251(1):74-86. DOI: 10.1002/path.5418.
37. Hu D, Shen B, Yu M, et al. Altered Serum Annexin A2 Might Be a New Potential Diagnostic Biomarker in Human Colorectal Cancer[J]. *Ann Clin Lab Sci*, 2020,50(6):726-733.

38. Wang J, Yu S, Chen G, et al. A novel prognostic signature of immune-related genes for patients with colorectal cancer[J]. *J Cell Mol Med*, 2020,24(15):8491-8504. DOI: 10.1111/jcmm.15443.
39. Song K, Lu H, Jin L, et al. Qualitative Ras pathway signature for cetuximab therapy reveals resistant mechanism in colorectal cancer[J]. *FEBS J*, 2020,287(23):5236-5248. DOI: 10.1111/febs.15306.
40. Xiong Y, Huang F, Li X, et al. CCL21/CCR7 interaction promotes cellular migration and invasion via modulation of the MEK/ERK1/2 signaling pathway and correlates with lymphatic metastatic spread and poor prognosis in urinary bladder cancer[J]. *Int J Oncol*, 2017,51(1):75-90. DOI: 10.3892/ijo.2017.4003.
41. Litviakov NV, Cherdynitseva NV, Tsyganov MM, et al. Changing the expression vector of multidrug resistance genes is related to neoadjuvant chemotherapy response[J]. *Cancer Chemother Pharmacol*, 2013,71(1):153-163. DOI: 10.1007/s00280-012-1992-x.
42. Klintman M, Buus R, Cheang MC, et al. Changes in Expression of Genes Representing Key Biologic Processes after Neoadjuvant Chemotherapy in Breast Cancer, and Prognostic Implications in Residual Disease[J]. *Clin Cancer Res*, 2016,22(10):2405-2416. DOI: 10.1158/1078-0432.CCR-15-1488.
43. Hashimoto Y, Yoshida Y, Yamada T, et al. Current Status of Therapeutic Drug Monitoring of 5-Fluorouracil Prodrugs[J]. *Anticancer Res*, 2020,40(8):4655-4661. DOI: 10.21873/anticancer.14464.
44. Widemann BC, Adamson PC. Understanding and managing methotrexate nephrotoxicity[J]. *Oncologist*, 2006,11(6):694-703. DOI: 10.1634/theoncologist.11-6-694.
45. Kenny RG, Ude Z, Docherty JR, et al. Vorinostat and Belinostat, hydroxamate-based anti-cancer agents, are nitric oxide donors[J]. *J Inorg Biochem*, 2020,206:110981. DOI: 10.1016/j.jinorgbio.2019.110981.]
46. Ji M, Li Z, Lin Z, et al. Antitumor activity of the novel HDAC inhibitor CUDC-101 combined with gemcitabine in pancreatic cancer[J]. *Am J Cancer Res*, 2018,8(12):2402-2418.

## Tables

Table1 | Relationship between clinical features and NMB expression in CRC patients (logistic regression).

Clinicopathological parameters	Total (N)	Odds ratio in NMB expression	P-value
Age ≥65 vs. <65	540	1.504(1.070-2.121)	0.019
Gender Male vs. female	540	1.1779(0.840-1.653)	0.343
T classification			
T2 vs. T1	105	0.327(0.086-1.032)	0.071
T3 vs. T1	383	0.368(0.100-1.100)	0.092
T4 vs. T1	78	0.310(0.080-1.015)	0.066
N classification			
N1vs.N0	446	0.807(0.534-1.216)	0.306
N2vs.N0	410	1.475(0.926-2.369)	0.104
distant metastasis			
M1vs.M0	477	1.139(0.697-1.867)	0.603
Stage			
Stage II vs. Stage I	300	0.969(0.594-1.582)	0.901
Stage III vs. Stage I	241	0.927(0.551-1.558)	0.775
Stage IV vs. Stage I	170	1.058(0.578-1.940)	0.855

Table 2 | Associations between overall survival and clinicopathological characteristics in CRC using univariate analysis and multivariate analysis.

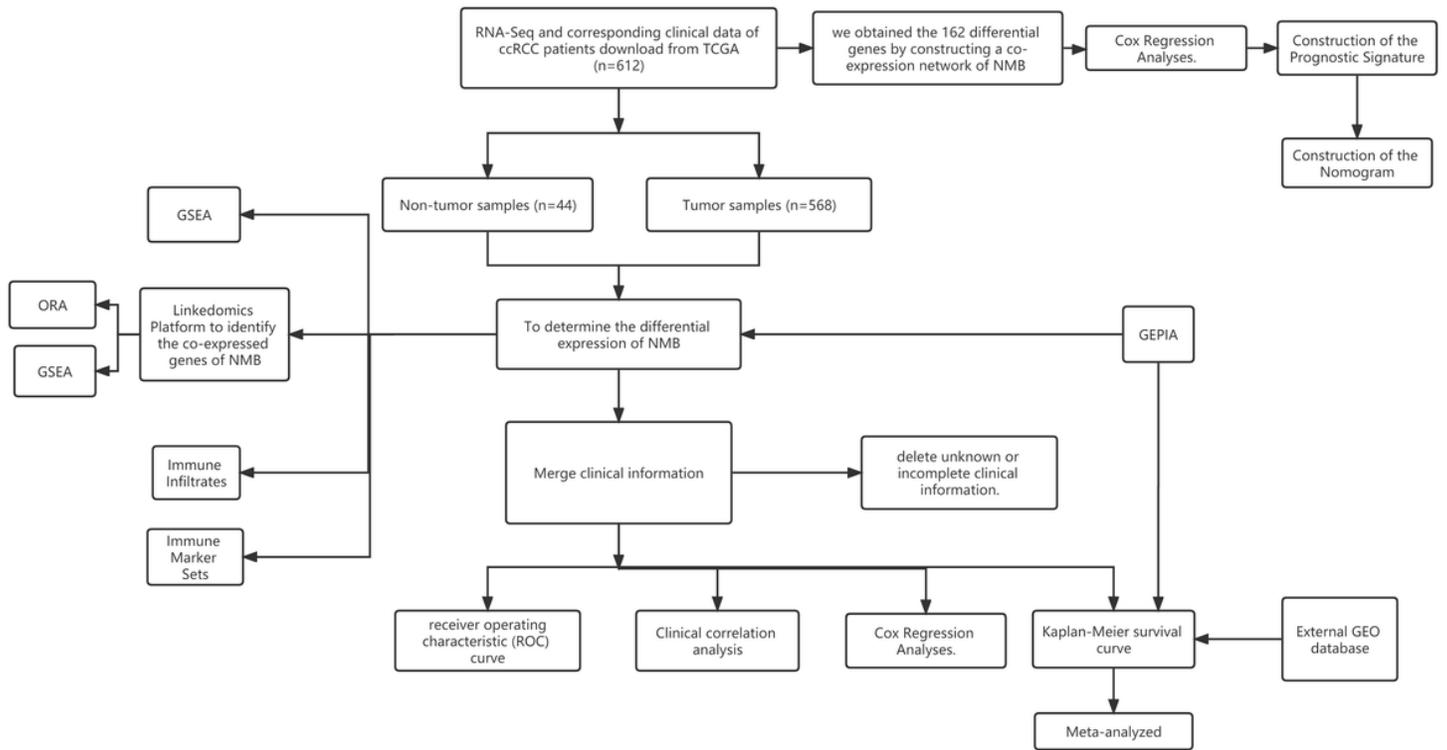
Parameters	univariate analysis				multivariate analysis			
	HR	95%CI		P-value	HR	95%CI		P-value
		Lower	Upper			Lower	Upper	
Age (continuous)	1.039	1.017	1.062	0.000	1.047	1.025	1.069	0.000
Gender	1.036	0.671	1.599	0.873				
stage	2.293	1.794	2.932	0.000	1.580	0.772	3.235	0.211
Tumor size	2.890	1.887	4.427	0.000	1.733	1.068	2.813	0.026
Lymph node	2.073	1.606	2.675	0.000	1.178	0.756	1.835	0.469
Distant metastasis	4.512	2.888	7.049	0.000	1.709	0.632	4.622	0.291
NMB expression level	1.046	1.005	1.088	0.027	1.054	1.008	1.102	0.021

Table 3 | Correlation analysis between NMB and relate genes and markers of immune cells in TIMER.

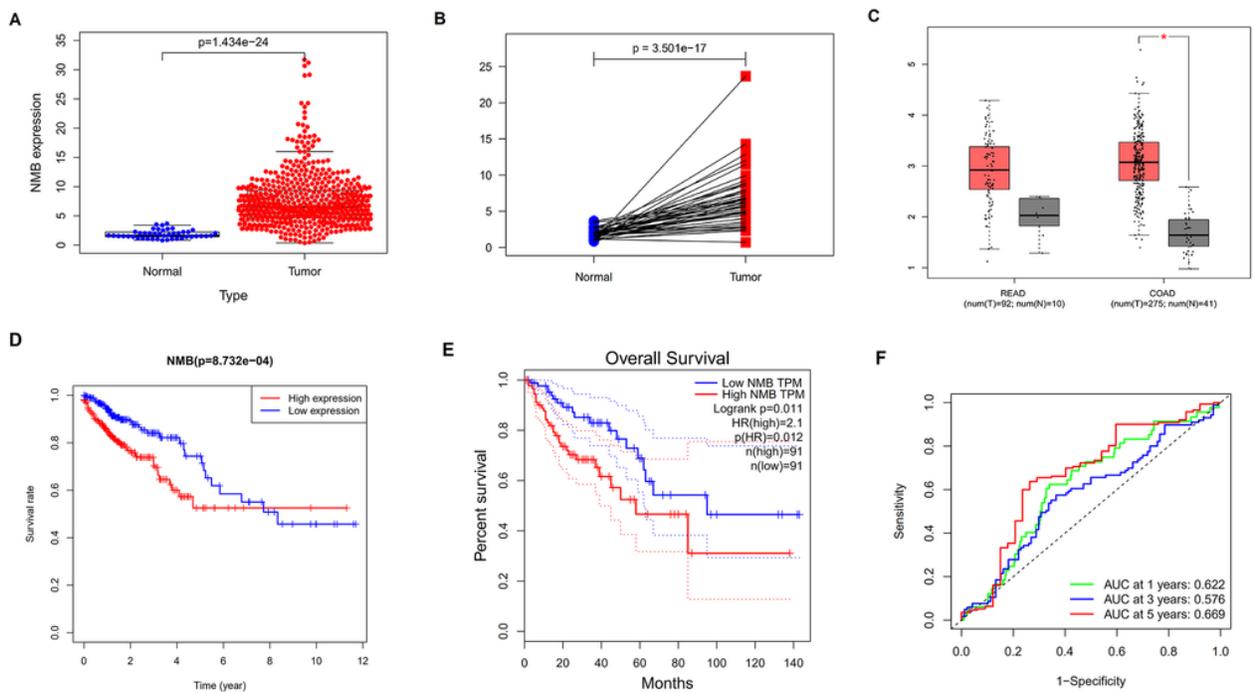
Description	Gene markers	COAD						READ					
		None		Purity		Age		None		Purity		Age	
		Cor	P										
CD8 + Tcell	CD8A	0.119	0.011	-0.397	0.000	-0.093	0.047	-0.093	0.233	-0.402	0.000	0.024	0.761
	CD8B	0.15	0.001	-0.221	0.000	-0.015	0.757	-0.035	0.657	-0.127	0.134	-0.041	0.598
T cell (general)	CD3D	0.062	0.187	-0.447	0.000	0.057	0.224	-0.082	0.295	-0.412	0.000	0.045	0.560
	CD3E	0.043	0.354	-0.477	0.000	0.051	0.277	-0.126	0.105	-0.444	0.000	0.104	0.183
	CD2	0.083	0.077	-0.435	0.000	0.061	0.191	-0.093	0.234	-0.404	0.000	0.05	0.521
B cell	CD19	0.025	0.594	-0.414	0.000	-0.052	0.271	-0.063	0.422	-0.35	0.000	-0.056	0.471
	CD79A	0.005	0.912	-0.443	0.000	-0.025	0.596	-0.198	0.011	-0.443	0.000	0.108	0.165
Monocyte	CD86	0.077	0.099	-0.42	0.000	0.056	0.220	-0.046	0.556	-0.407	0.000	-0.058	0.456
	CSF1R	0.017	0.713	-0.366	0.000	0.078	0.099	-0.124	0.113	-0.414	0.000	0.011	0.888
TAM	CCL2	0.062	0.182	-0.354	0.000	0.058	0.214	-0.051	0.510	-0.474	0.000	-0.01	0.903
	CD68	0.072	0.123	-0.317	0.000	0.063	0.181	-0.055	0.478	-0.355	0.000	0.026	0.735
	IL10	0.093	0.046	-0.303	0.000	0.106	0.024	0.024	0.762	-0.381	0.000	-0.026	0.743
M1 Macrophage	NOS2	0.176	0.000	-0.131	0.008	0.076	0.105	-0.003	0.967	-0.064	0.451	0.112	0.151
	IRF5	0.091	0.051	-0.093	0.061	-0.009	0.852	0.114	0.145	-0.117	0.167	-0.002	0.979
	PTGS2	0.125	0.007	-0.221	0.000	0.062	0.190	0.109	0.162	-0.258	0.002	-0.077	0.322
M2 Macrophage	CD163	0.051	0.279	-0.389	0.000	0.106	0.024	0.081	0.296	-0.379	0.000	-0.015	0.845
	VSIG4	0.085	0.070	-0.355	0.000	0.109	0.020	0.010	0.894	-0.327	0.000	0.052	0.502
	MS4A4A	0.074	0.111	-0.379	0.000	0.109	0.020	-0.078	0.320	-0.407	0.000	0.013	0.866
Neutrophils	CEACAM8	0.099	0.034	0.103	0.038	0.076	0.106	0.152	0.051	0.109	0.202	0.116	0.137
	ITGAM	-0.021	0.655	-0.355	0.000	0.054	0.253	-0.096	0.218	-0.361	0.000	-0.017	0.828
	CCR7	0.028	0.552	-0.415	0.000	-0.052	0.267	-0.090	0.246	-0.375	0.000	-0.102	0.192
Natural killer cell	KIR2DL1	0.121	0.009	-0.189	0.000	0.053	0.262	-0.038	0.627	-0.174	0.039	0.077	0.321
	KIR2DL3	0.091	0.053	-0.196	0.000	0.127	0.007	-0.083	0.288	-0.186	0.028	-0.087	0.267
	KIR2DL4	0.118	0.012	-0.322	0.000	0.112	0.017	-0.04	0.609	-0.348	0.000	0.039	0.617
	KIR3DL1	0.113	0.015	-0.249	0.000	0.083	0.079	0.007	0.930	0.161	0.057	-0.067	0.391
	KIR3DL2	-0.003	0.954	-0.264	0.000	0.056	0.236	-0.133	0.088	-0.295	0.000	0.091	0.245
	KIR3DL3	0.039	0.405	-0.063	0.203	0.016	0.732	-0.077	0.327	-0.084	0.323	-0.063	0.422
Dendritic cell	KIR2DS4	0.108	0.021	-0.116	0.019	0.048	0.307	0.093	0.243	-0.098	0.250	0.116	0.138
	HLA-DPB1	0.062	0.188	-0.450	0.000	0.115	0.014	-0.120	0.123	-0.414	0.000	0.010	0.896
	HLA-DQB1	0.154	0.000	-0.334	0.000	0.056	0.041	0.057	0.468	-0.267	0.001	0.033	0.676
	HLA-DRA	0.102	0.030	-0.422	0.000	0.103	0.028	-0.049	0.530	-0.445	0.000	-0.003	0.972
	HLA-DPA1	0.108	0.021	-0.431	0.000	0.118	0.012	-0.121	0.120	-0.451	0.000	-0.006	0.937
	CD1C	0.012	0.805	-0.331	0.000	-0.068	0.151	-0.218	0.005	-0.394	0.000	0.000	0.997

Th1	NRP1	0.010	0.827	-0.414	0.000	-0.016	0.728	-0.069	0.377	-0.382	0.000	-0.141	0.070
	ITGAX	0.023	0.628	-0.427	0.000	0.016	0.732	-0.017	0.830	-0.410	0.000	-0.022	0.779
	TBX21	0.084	0.074	-0.400	0.000	0.102	0.030	-0.098	0.209	-0.358	0.000	0.067	0.389
	STAT1	0.088	0.059	-0.274	0.000	0.076	0.108	0.009	0.911	-0.318	0.000	-0.068	0.385
	IFNG	0.144	0.002	-0.241	0.000	0.151	0.001	-0.096	0.219	-0.307	0.000	0.075	0.339
	TNF	0.093	0.046	-0.255	0.000	0.063	0.180	0.171	0.028	-0.313	0.000	-0.082	0.295
Th2	GATA3	0.027	0.567	-0.364	0.000	-0.024	0.614	-0.061	0.432	-0.331	0.000	-0.017	0.828
	STAT6	-0.139	0.003	0.012	0.815	0.011	0.818	-0.130	0.096	0.060	0.482	0.174	0.025
	STAT5A	0.044	0.345	-0.161	0.001	0.037	0.434	0.052	0.503	-0.222	0.008	0.042	0.595
	IL13	0.140	0.003	-0.197	0.000	0.007	0.886	0.024	0.763	-0.219	0.009	-0.035	0.655
Tfh	BCL6	0.005	0.916	-0.357	0.000	0.011	0.807	-0.131	0.093	-0.222	0.008	0.062	0.427
	IL21	0.054	0.246	-0.134	0.007	0.083	0.078	0.016	0.841	-0.195	0.021	-0.066	0.399
Th17	STAT3	-0.054	0.250	-0.225	0.000	-0.062	0.186	-0.131	0.093	-0.245	0.004	-0.090	0.250
	IL17A	0.017	0.711	-0.012	0.814	0.023	0.630	-0.036	0.650	0.061	0.471	0.158	0.042
Treg	FOXP3	0.052	0.265	-0.382	0.000	0.018	0.707	-0.117	0.133	-0.383	0.000	0.038	0.626
	CCR8	0.017	0.721	-0.336	0.000	-0.021	0.657	-0.088	0.258	-0.390	0.000	-0.100	0.201
	STAT5B	-0.077	0.100	-0.054	0.279	-0.121	0.010	-0.073	0.648	-0.203	0.016	-0.117	0.132
	TGFB1	0.138	0.003	-0.435	0.000	0.043	0.357	0.076	0.331	-0.393	0.000	-0.001	0.989
T cell exhaustion	PDCD1	0.100	0.032	-0.384	0.000	0.110	0.019	-0.086	0.273	-0.387	0.000	0.005	0.954
	CTLA4	0.050	0.282	-0.402	0.000	0.055	0.240	-0.134	0.085	-0.337	0.000	-0.088	0.257
	LAG3	0.144	0.002	0.391	0.000	0.123	0.009	0.006	0.935	-0.333	0.000	0.079	0.309
	HAVCR2	0.096	0.040	-0.408	0.000	0.068	0.148	-0.004	0.958	-0.402	0.000	-0.032	0.684
	GZMB	0.172	0.000	-0.057	0.254	-0.057	0.254	0.053	0.497	-0.137	0.106	0.027	0.731

## Figures

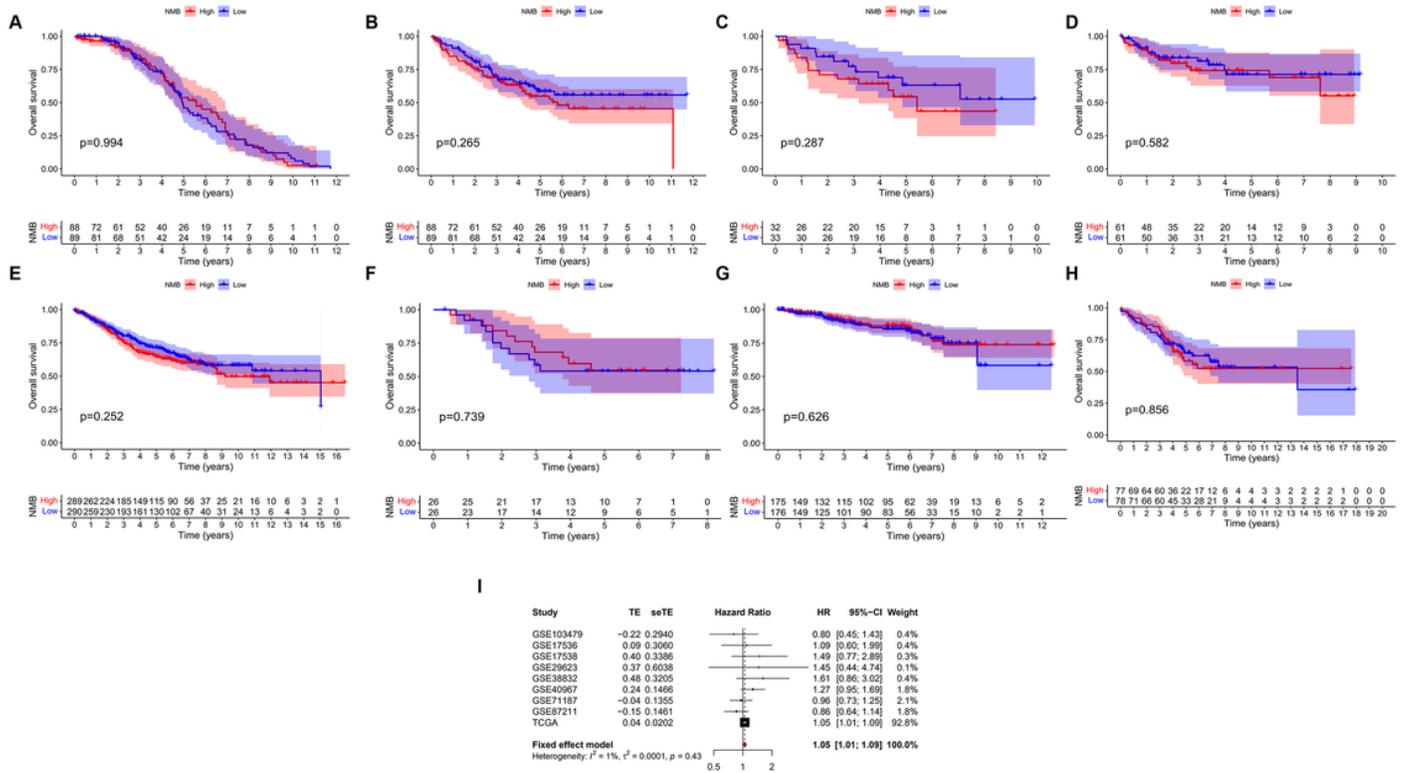


**Figure 1**  
Overall design of this study



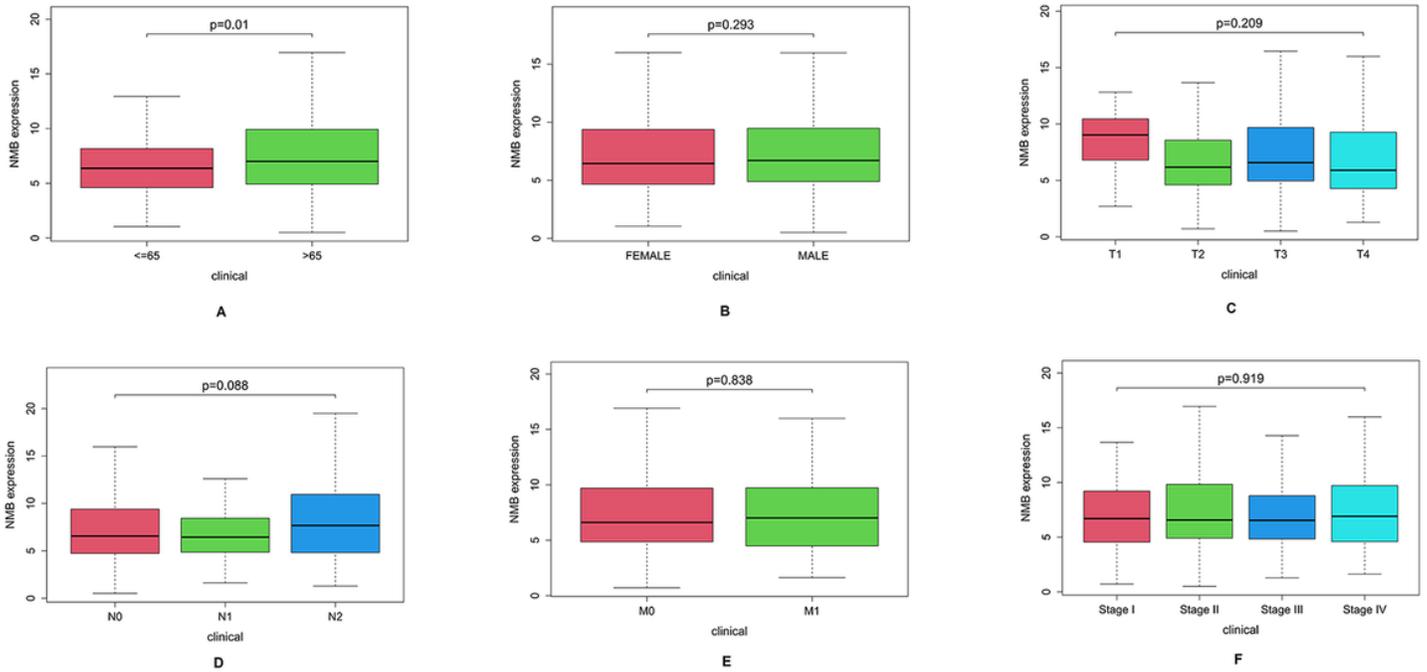
**Figure 2**

(A) Scatter plot shows the differential expression of NMB gene in CRC tissues (N = 568) and tumor-adjacent normal tissues (N = 44)( $p < 0.001$ ). (B) Paired plot shows the differential expression of the NMB gene in CRC tissues (N = 568) and tumor-adjacent normal tissues (N = 44) ( $p < 0.001$ ). (C) GEPIA confirmed differential expression of NMB. (D) The Kaplan-Meier survival curve of the CRC patients with high and low NMB expression levels. (Median NMB expression, cutoff-high, 50%; Cutoff - Low, 50%). (E) The Kaplan-Meier survival curve of the CRC patients with high and low NMB expression levels (Quartile NMB expression, cutoff-high, 75%; Cutoff - Low, 25%). (F) Receiver operating characteristic (ROC) curve for NMB expression.



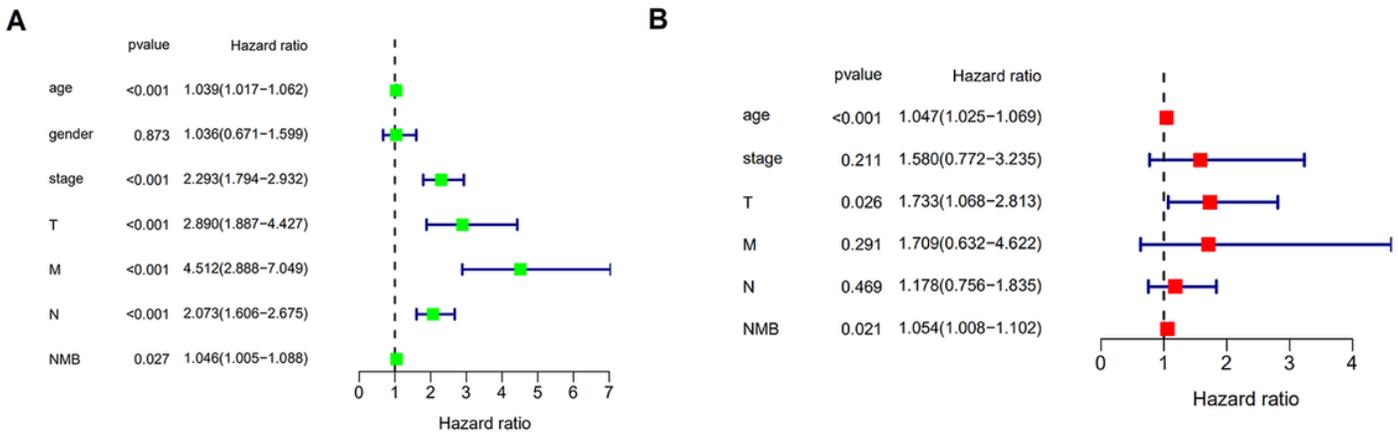
**Figure 3**

Survival analysis based on eight datasets from the GEO database. (A) GSE17536, (B) GSE17538, (C) GSE29623, (D) GSE38832, (E) GSE40967, (F) GSE71187, (G) GSE87211, (H) GSE103479. (I) Meta-analysis of prognosis in patients with CRC.



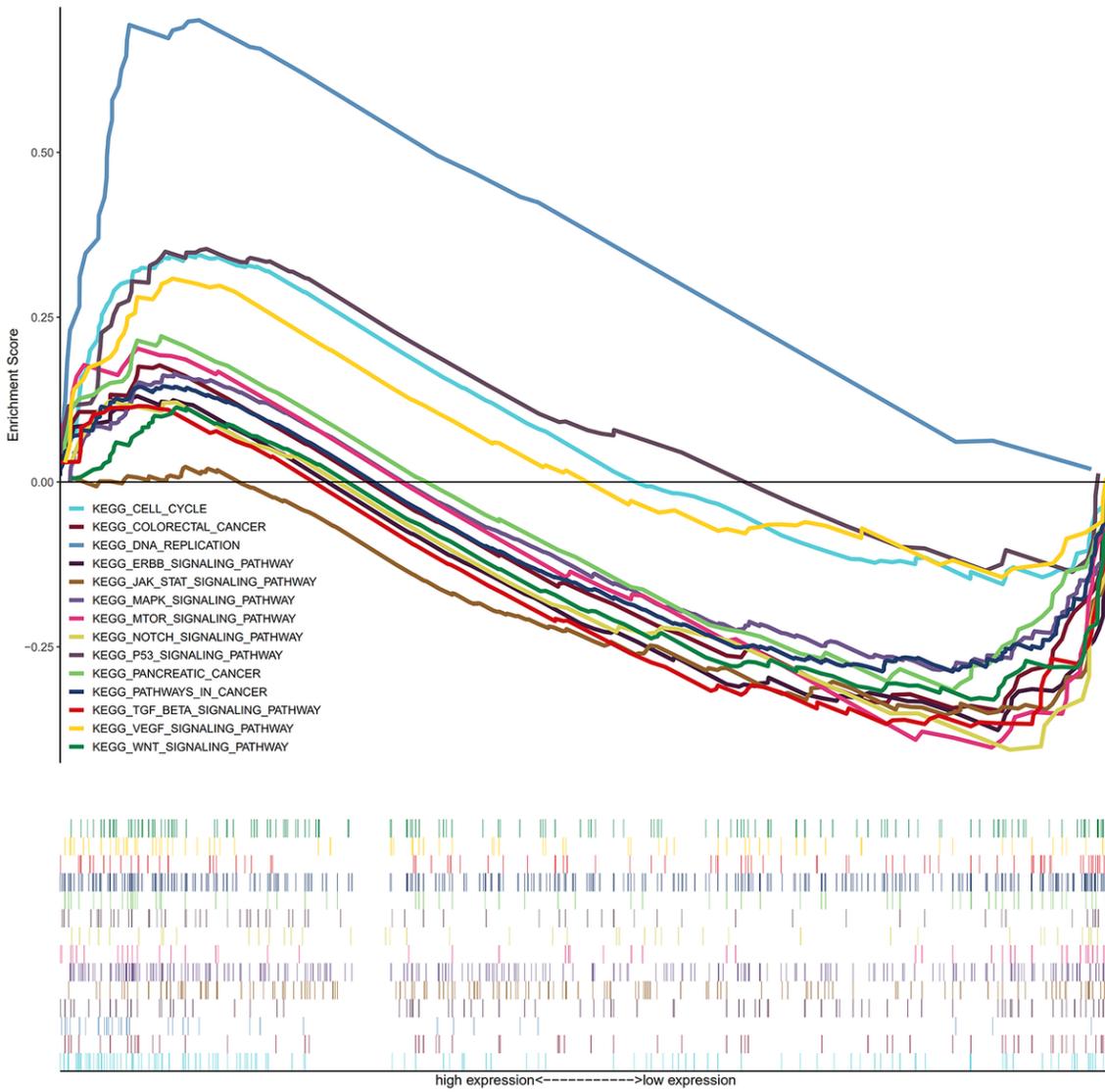
**Figure 4**

Clinical correlation analysis of CRC patients.(A) the expression of NMB in colorectal cancer was correlated with age(OR = 1.504 for  $\geq 65$  versus  $< 65$ ,  $P = 0.019$ ).(B)-(F) There was no statistically significant difference between NMB and gender,T classification,N classification,distant metastasis(M)and stage.



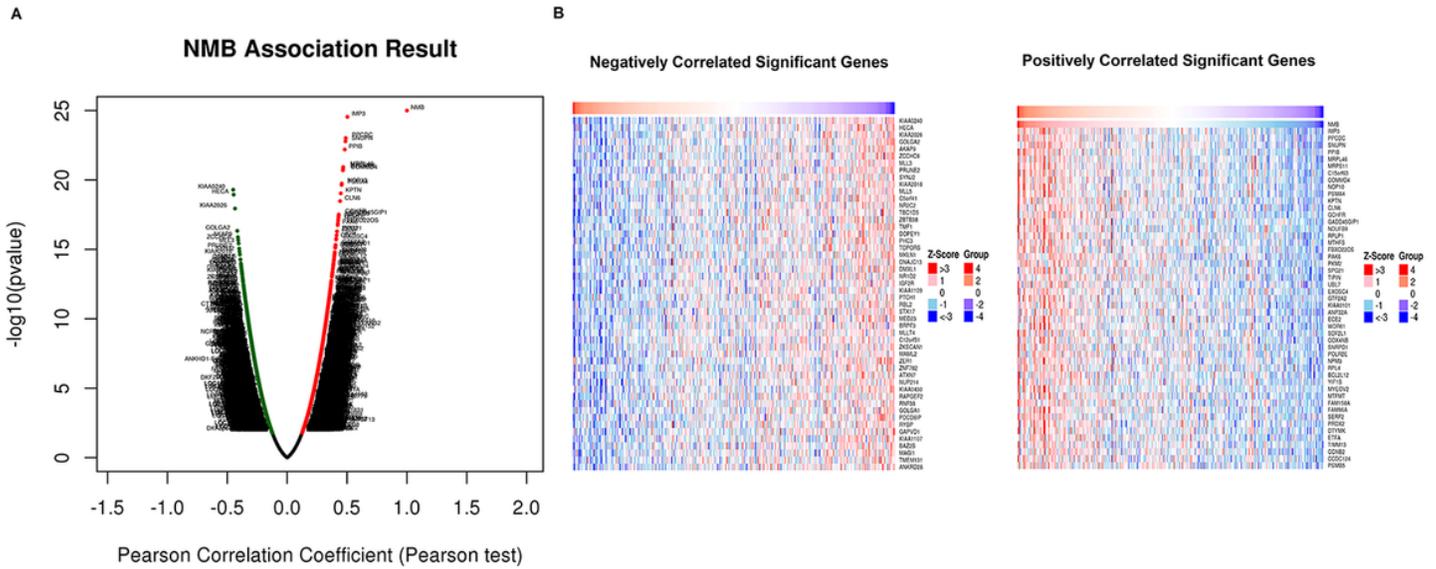
**Figure 5**

Cox regression analysis.(A)The forest map shows the results of Univariate Cox regression analysis.(B)The forest map shows the results of multivariate Cox regression analysis.



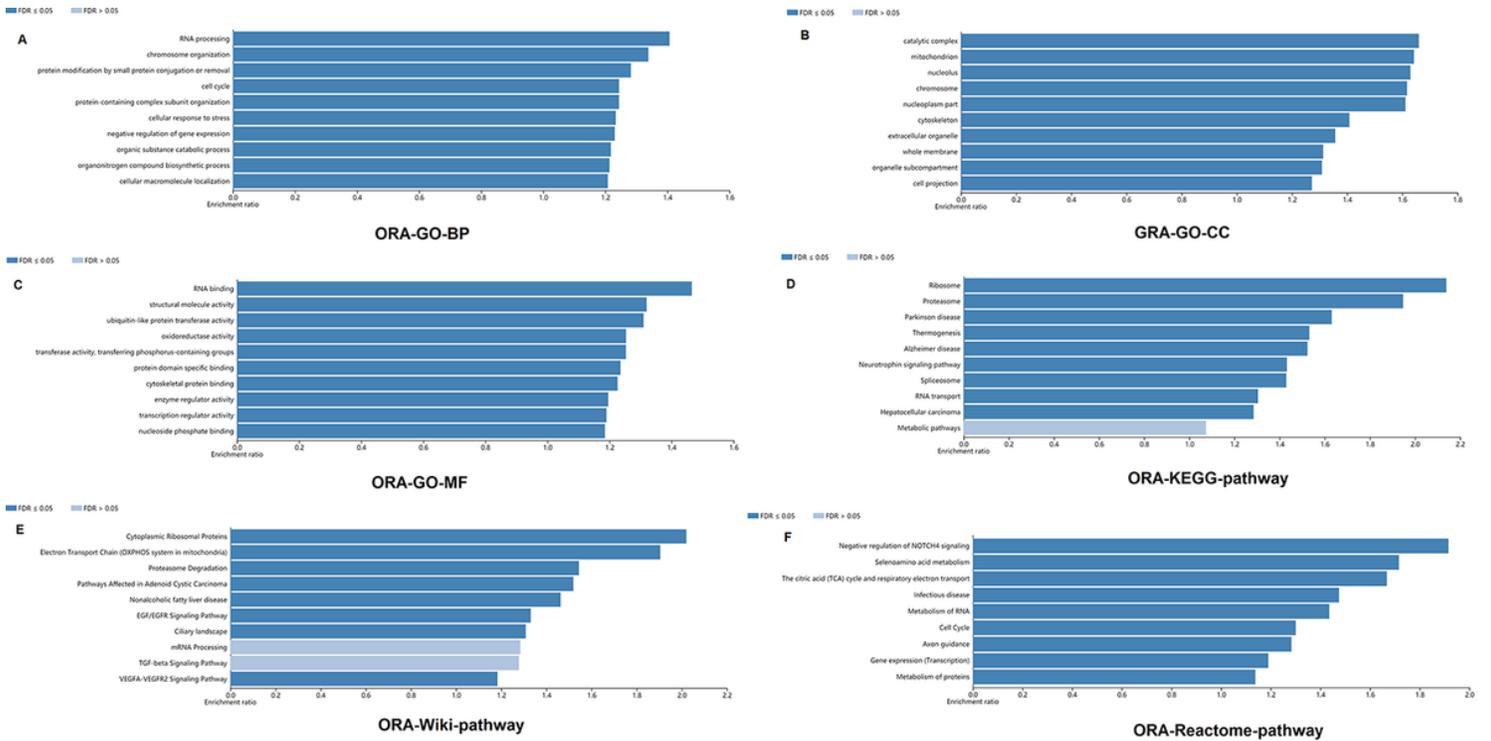
**Figure 6**

A merged enrichment plot from gene set enrichment analysis (GSEA) including enrichment score and gene sets. 14 cancer-related pathways are shown here.



**Figure 7**

NMB co-expression genes in CRC.(A) The global NMB significantly correlated genes in the CRC cohort were identified by LinkedOmics.(B) Heatmaps showing top 50 genes positively and negatively correlated with NMB in CRC. Red dot, positively correlated gene; blue dot, negatively correlated genes.



**Figure 8**

Enrichment analysis of significantly related genes in CRC cohorts by ORA.(A) ORA-GO-BP.(B) ORA-GO-CC.(C) ORA-GO-MF.(D) ORA-KEGG-pathway. (E) ORA-Wiki-pathway.(F) ORA-Reactome-pathway.

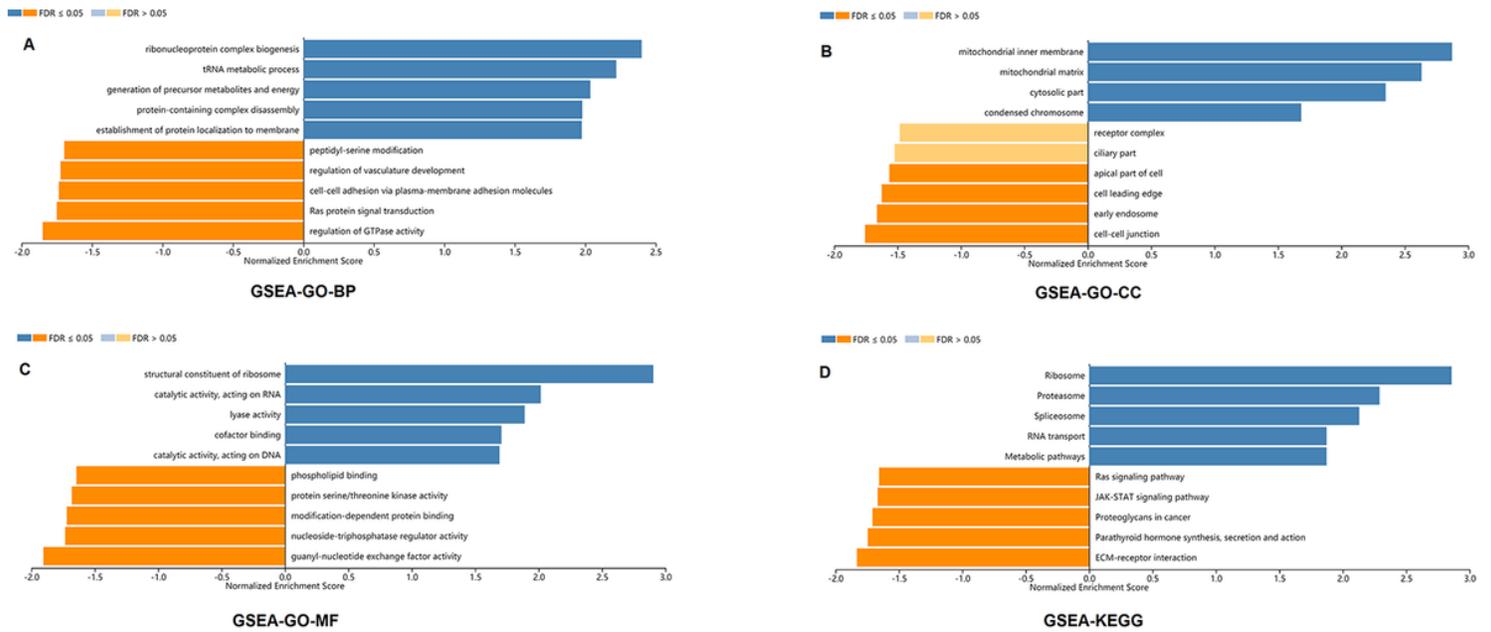


Figure 9

Enrichment analysis of significantly related genes in CRC cohorts by GSEA. (A) GSEA-GO-BP. (B) GSEA-GO-CC. (C) GSEA-GO-MF. (D) GSEA-KEGG.

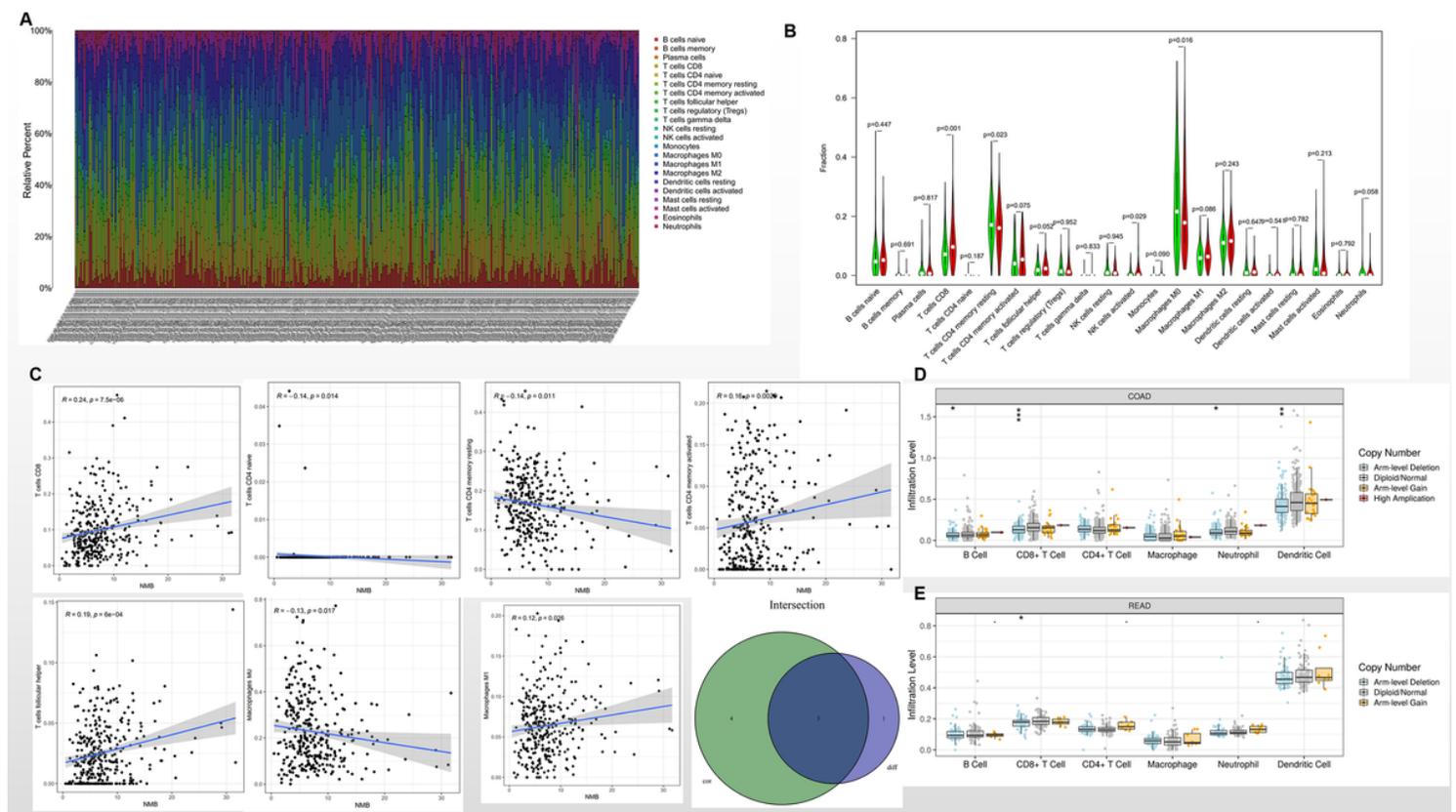


Figure 10

(A) The number of immune cells in each sample. (B) Results of differential analysis of immune cells. Green represents the low expression group and red represents the high expression group. (C) NMB was significantly correlated with T cells CD8 ( $R=0.24, p<0.001$ ), T cells CD4 naive ( $R=-0.14, p=0.014$ ), T cells CD4 memory resting ( $R=-0.14, p=0.011$ ), T cells CD4 memory activated ( $R=0.16, p=0.0029$ ), T cells follicular

helper(R=0.19,p<0.001),Macrophages M0(R=-0.13,p=0.017) and Macrophages M1(R=0.12,p=0.026);The intersection of the two test results.(D-E)The relationships between infiltration levels of 6 immune cells and copy number of NMB in CRC.

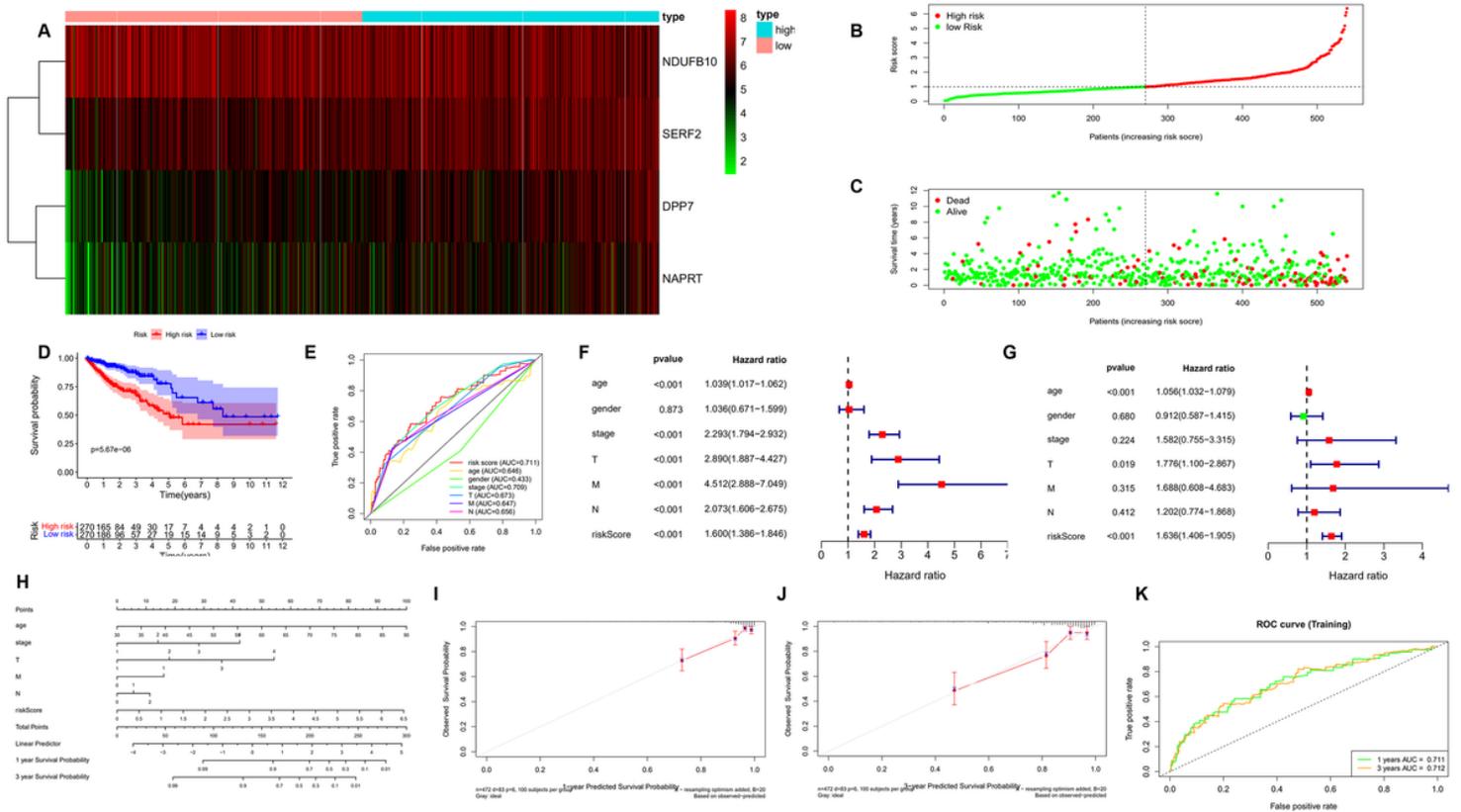


Figure 11

(A) The heatmap of the 4 related gene expression profiles in high- and low-risk CRC patients. (B) Distribution of risk scores of high- and low-risk CRC patients. (C) Scatter plot shows the correlation between survival time and risk score. (D) The K-M curve reflects that the OS of high-risk CRC patients is significantly lower than that of low-risk patients ( $P < 0.001$ ). (E) ROC curve reflects the high predictive value of the risk score (AUC = 0.711). (F) The forest plot reflects the univariate Cox analysis of the relationship between the clinical features, risk score, and OS of CRC. (G) Forest plot reflects multivariate Cox analyzed the relationship between the clinical features, risk score and OS of CRC patients. risk score ( $P < 0.001$ ) is an independent prognostic risk factor for CRC. (H) Calculate the scores of each item of CRC patients according to the nomogram, and the total scores obtained after addition can predict the 1- and 3-year survival probability. (I)(J) The 1- and 3-year calibration curves of the nomogram. (K) The ROC curves of 1- and 3-year nomogram (AUC = 0.711 of 1 year, AUC = 0.712 of 3 years).

