

# Ano1 is a Prognostic Biomarker That is Correlated with Immune Infiltration in Colorectal Cancer

Jun Chen<sup>1,\*</sup>, Hongli Wang<sup>2,\*</sup>, Fang Peng<sup>3,\*</sup>, Haiyan Qiao<sup>1</sup>, Linfeng Liu<sup>1</sup>, Liang Wang<sup>1</sup>, Bingbing Shang<sup>4</sup>

<sup>1</sup>Laboratory Animal Center, Dalian Medical University, Dalian, 116044, Liaoning Province, People's Republic of China; <sup>2</sup>Cardiology Department, The Second Hospital of Dalian Medical University, Dalian, 116023, Liaoning Province, People's Republic of China; <sup>3</sup>Pathology Department, The Second Hospital of Dalian Medical University, Dalian, 116023, Liaoning Province, People's Republic of China; <sup>4</sup>Emergency Department, The Second Hospital of Dalian Medical University, Dalian, 116023, Liaoning Province, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Bingbing Shang; Liang Wang, Tel +86-17709875175; +86-13332225676, Email shangbingbing@dmu.edu.cn; wangliang-dy@dmu.edu.cn

**Background:** Anoctamin 1 (ANO1) has been observed to be overexpressed in gastrointestinal and pulmonary epithelial cells, as well as in a number of cancers. Although Ano1 is involved in the prognosis of colorectal cancer (CRC), its mechanism of action in metastatic CRC has not been fully elucidated.

**Methods:** The expression of Ano1 was assessed in samples obtained from The Cancer Genome Atlas (TCGA) database. Then, we used Gene Ontology (GO) enrichment, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, Gene set enrichment analysis (GSEA), Gene set variation analysis (GSVA), and Weighted Correlation Network Analysis (WGCNA) to determine the functions of Ano1. Additionally, random survival forest, Cox multivariate analysis, Kaplan Meier analysis, and ROC were used to determine the predictive value of Ano1 on clinical outcomes in CRC patients. Finally, HE staining, immunohistochemical (IHC) analysis and qRT-PCR were used to explore the expression of the Ano1 gene in CRC tissue.

**Results:** The expression level of Ano1 in CRC was significantly elevated, and the prognosis was poor. The modules with a higher proportion of upregulated genes tended to be positively correlated with Ano1-high. KNG1, GNG4, F2, POSTN, THBS2, SPP1 and FGA were identified as hub proteins of the PPI network. The heatmap showed that the expression level of the Ano1-high group was significantly negatively correlated with immune infiltrate. The overexpression of the Ano1 gene in CRC tissue samples was also confirmed by HE staining, immunohistochemical (IHC) analysis and qRT-PCR.

**Conclusion:** High expression of Ano1 is closely related to a poor prognosis in patients with colorectal cancer. Ano1 may participate in the metastasis and progression, as well as the immune regulation of CRC. In summary, Ano1 can act as a potential prognostic biomarker and a novel target for CRC therapy.

**Keywords:** ANO1, biomarker, colorectal cancer, prognosis, Immune infiltration

## Introduction

At present, the incidence rate of colorectal cancer is the third highest among all cancers worldwide. CRC is followed by lung cancer, which is the second most common cause of cancer-related deaths.<sup>1</sup> In recent years, the incidence rate of colorectal cancer in China has shown an increase, with 387,600 CRC new patients being diagnosed during the year 2015.<sup>2</sup> The effective methods of treatment for CRC include radiotherapy, chemotherapy, immunotherapy, and surgery. The high cost of therapy for CRC is a heavy social and economic burden. The tumor stage affects the prognosis of CRC. Although multiple methods of treatment have been used, the prognosis of patients diagnosed with metastatic colorectal cancer (mCRC) remains very poor. When CRC patients are at the metastatic phase, the 5-year overall survival decreases to about 13%.<sup>3,4</sup> Methods of therapy for patients with unresectable or mCRC are increasingly chosen using specific

biomarkers.<sup>5</sup> Therefore, there is a necessity to identify reliable biomarkers to predict prognosis and identify novel immunotherapy targets in patients with mCRC.

Ano1, also known as TMEM16A, TAOS2, ORAOV2, and FLJ10261 (DOG1, discovered on GIST-1), encodes a hypothetical protein and belongs to the calcium-activated chloride channels family.<sup>6–9</sup> Studies have found Ano1 is expressed in exocrine glands, nervous system, smooth muscles, and other tissues.<sup>10,11</sup> As a transmembrane protein, Ano1 is involved in a variety of biological functions, such as cell proliferation, attachment, and movement. Since Ano1 forms the molecular base of CaCCs in multiple cell types, it is regarded as a novel strategy to treat many diseases, including tumors.<sup>8</sup> Recent studies have shown that a high expression of Ano1 is associated with breast cancer, prostate carcinoma, squamous cell carcinomas, and gastrointestinal stromal tumors.<sup>12–17</sup> However, in human malignancies, especially in mCRC, the mechanism of action followed by ANO1 remains to be elucidated.

We aimed to outline the correlation between Ano1 and CRC using TCGA. We performed a correlation analysis between the expression of the Ano1 gene and immune cell infiltration using CIBERSORT. Furthermore, we used GSEA and GSVA to perform the enrichment analyses. WGCNA was also used to perform the co-expression analyses. A PPI network was established to screen for hub proteins. Moreover, we developed prognostic models to predict patient prognosis.

## Materials and Methods

### Data Source and Preprocessing

Gene expression data (644 tumor samples and 51 para-cancerous control samples) and relevant clinical data were downloaded from TCGA database using R software via the TCGA BioLinks package<sup>18</sup> from GDC. Overall, the HTSeq-Counts data of 617 CRC patients with complete overall survival information were retained and further analyzed. Ethical approval and informed consent were not required because our data were provided by TCGA.

### Immunohistochemistry (IHC) of Ano1

We used HE staining and immunohistochemical (IHC) analysis to explore the expression of the Ano1 gene in CRC tissue samples collected from The Second Hospital of Dalian Medical University (Dalian, Liaoning province, China). Our study was approved by the Ethics Committee of the hospital (License number: 2021-079).

Human CRC tissues were fixed in 10% formalin overnight and then embedded in paraffin. Next, the tissue section (3  $\mu$ m in thickness) was dewaxed, rehydrated, and subjected to antigen retrieval in EDTA citrate buffer. The sections were covered with 3% hydrogen peroxide and then incubated overnight at 4°C with anti-Ano1 (1:200, TA805184, Zhongshan Jinqiao Biological, Beijing, China). The bound antibodies were detected using Secondary antibody binding to horseradish peroxidase and developed with 3,3'-diaminobenzidine using a two-step test kit (PV-9000, Zhongshan Jinqiao Biological, Beijing, China).

### Differential Expression Analysis and Prognostic Analysis of Ano1 in TCGA-CRC

The loading, post refinement, filtering, and normalization of RNA-Seq Count data were performed using the edgeR package. Count per million read (CPM) values were calculated and were used to extract the expression values of the Ano1 gene. Differential expression analysis of the tumor/normal tissue was conducted using Student's *t*-test in R software.  $P < 0.05$  was defined as a significant result.

Based on the results of all clinical prognostic analyses including overall survival (OS) and OS status, survival analyses were performed on the Ano1 gene. Ano1 expression was stratified into two groups, Ano1-high and Ano1-low groups, at the median value of tumor samples. The Log rank test was used for the statistical analysis of survival using the survival R package. A log-rank  $p$  value  $< 0.05$  was defined as the threshold of significance.

### Pan Cancer Differential Gene Expression Analyses of Ano1

To further study the expression of Ano1 in multiple tumors, Pan-cancer RNA-seq data and clinical data were downloaded from the UCSC Xena database. The processed data were used to extract the expression values of the

Ano1 gene, and the last three tumor samples were preserved as the control group. In detail, differential gene expression analyses of Ano1 in each cancer type was also performed using the Student's *t*-test. A P value < 0.05 was regarded as the threshold of significance. ROC curves were plotted to evaluate the efficacy of Ano1 to differentiate tumors from controls.

## Differential Gene Expression Analysis Based on the Expression Profiles of Ano1

To further explore the underlying mechanism of Ano1 in tumors, we classified TCGA-CRC tumor samples into two groups (high- and low-Ano1), based on the median value of the Ano1 gene, and performed a differential analysis. We loaded, collected, filtered, and normalize the raw count data using the edgeR statistical analysis package. Then, voom, linear models and empirical Bayes of the limma package indicated that we should perform a differential analysis between the two groups.<sup>19</sup>

The adj. p value was obtained via BH corrections. A P value < 0.05 and an absolute value of log2 (fold change) of more than 1 were the thresholds of significant differences in differential gene expression. We have demonstrated this differential expression and clustering by drawing a volcano plot, PCA and heat map of the differential gene, which were produced using the R packages: ggplot2, ggfortify, and pheatmap, separately.

## Functional Enrichment and Analysis

The cluster Profiler R package<sup>3</sup> was used to test the statistical enrichment of upregulated and downregulated differentially expressed genes in the GO enrichment and KEGG pathways. The cancer hallmark gene set<sup>20</sup> “h.all.v7.4.entrez.gmt” within the Molecular Signatures Database (MSigDB) was selected as references for GSEA. The Cancer hallmark was also used as reference when performing the GSVA using the R package.<sup>21</sup> The threshold of significance for enrichment was set at adj.p value < 0.05.

## WGCNA Analysis

Then, we extracted the Ano1 grouped correlated gene modules from the differential gene expression values matrix via WGCNA performed using the “WGCNA” package in R.

We first estimated the best Soft Thresholding Power  $\beta$  to obtain the adjacency matrix and topological matrix. The topological matrix clustered the genes based on the dissimilarity between genes, from which module assignments were made using a dynamic tree cutting method. The feature vector gene of each module was calculated and was regarded as the first principal component (eigengene) of a module that producing the gene expression levels in the module. Finally, a statistical analysis was performed to analyze the gene and the module.

## Key Nodes in the Protein-Protein Interaction (PPI) Network

In this study, we choose protein-protein interactions downloaded from the human STRING database as the background and assigned the differential genes of the WGCNA module to the background network. Then, we constructed the PPI network of the differential genes with a threshold of (combined score > 0.4 (Medium)). We used Cytoscape software to construct the network plots after the PPI relationship was obtained. The Cytoscape<sup>22</sup> plugin was used for topological analysis of the nodes in the network. The results included Betweenness centrality (BC), Degree Centrality (DC), and Closeness centrality (CC). The ranking of nodes based on their topology properties was used to identify the hub proteins<sup>23</sup> involved in the PPI network.

## Correlation Between Ano1 and Immune Cell Infiltrates

CIBERSORT is a tool used for the deconvolution of the transcriptome expression matrix based on the principle of linear support vector regression. We identified the composition and abundance<sup>24</sup> of immune cells in a mixed cell sample. Gene expression data were uploaded onto the CIBERSORT tool and the output with a P value less than 0.05 was filtered. Then, immune cells with the degree of infiltration at 0 were deleted. Thereafter, we derived the matrix of infiltrating immune cells. The R package, pheatmap, was used to draw the heatmaps. The images illustrate the distribution of immune cell infiltrates in each sample.

In addition, we performed a correlation analysis between the expression of the Ano1 gene and immune cell infiltration, which was visualized using scatter plots.

## Construction and Evaluation of the Prognostic Models

Based on the gene expression values in tumors, the DEGs obtained in method 2.3 were used to construct prognosis-related DEG models using the random survival forest (rsf) method. Random survival forest is obtained by adding survival analysis to random forest analysis, using the bootstrap resampling method to randomly extract N samples from the original dataset. Then, we built a survival trees model to obtain the variable importance measure (VIMP) of each variable. The larger the VIMP value, the stronger the prediction ability. The cumulative hazard function (CHF) was obtained from the mean of each survival tree, which reflects the cumulative probability.<sup>25</sup> The data of the DEGs and clinical variables were randomly divided into test and validation groups using the R package, random Forest SRC, to perform the RSF analysis.

We calculated the risk score of CHF based on test set and used the median risk score as the threshold for the high- and low-risk groups. The prognosis-related differential gene set was used as the validation set. The risk score of each sample was obtained by fitting it onto the same model parameters, and the samples were stratified into high- and low-risk groups using the Risk score threshold. We plotted the Kaplan-Meier curve using the “survminer” R package and performed a Log rank test. Time-dependent ROC analyses were conducted using the timeROC package in R. Time-dependent ROC analysis was used to predict the classification efficiency of the RiskScore. Subsequently, we analyzed the classification efficiency of the 1, 3, and 5-year overall survival. A multivariate Cox regression analysis using clinical data was conducted. We used the R package, caret, to conduct the probabilistic calibration analysis and plotted a calibration curve.

## qRT-PCR for CRC and Normal Tissues

The fresh tissue was grinded, then total RNA from fresh tissues was extracted using RNAiso plus (Takara, Dalian, China). Reverse transcription was performed using PrimeScript™ RT reagent Kit with gDNA Eraser (Takara, Dalian, China) according to the manufacturer's instructions. Quantitative PCR (qPCR) assay was performed using SYBR® Premix Ex Taq™ II (Takara, Dalian, China) on a qPCR system (Mx3000P, Agilent Technologies). The gene-specific qPCR primers are listed below: Ano1, forward 5'-GATCCCATCCAGCCCAAAGTG-3', reverse 5'-CGGGTTTGTCTGTC GAAAAAGGA-3'; 18S, forward 5'-CAGCCACCCGAGATTGAGCA-3', reverse 5'-TAGTAGCGACGGGCGGTGT-3'.

Statistical analysis was performed by Two-way-ANOVA and Student's *t* test using GraphPad Prism 7.0 or Microsoft Excel. Relative quantification was also determined using the  $2^{-(\Delta\Delta Ct)}$  method. Ribosomal RNA 18S was used as the reference gene. Data were presented as mean±standard deviation (SD). Differences were considered statistically significant if  $P < 0.05$ .

## Results

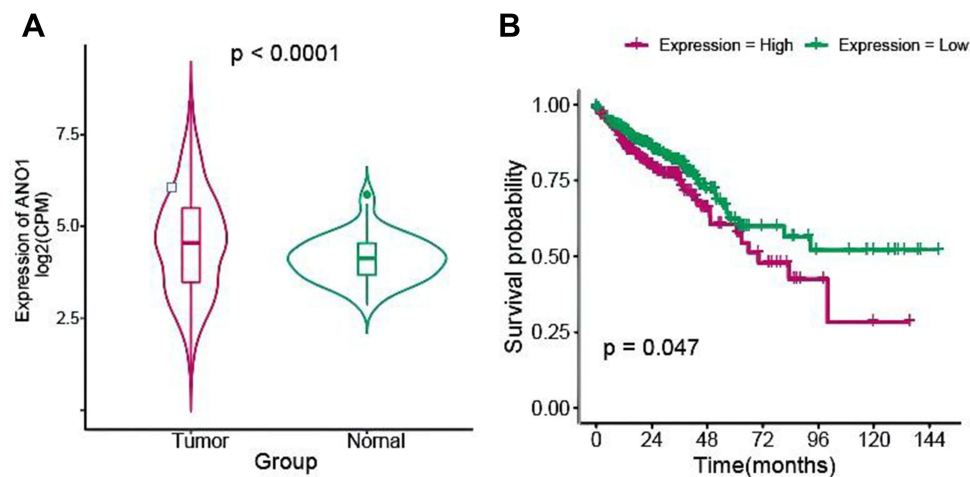
### The Differential Expression and Prognosis Analysis of Ano1 in CRC

The results of the differential expression analysis demonstrated that the expression of the Ano1 gene was significantly elevated in the CRC samples ( $p$  value =  $2.98 \times 10^{-5}$ ). The average value for the tumor group ( $n=644$ ) was 4.558 and that for the paracancerous control group ( $n=51$ ) was 4.110 (Figure 1A).

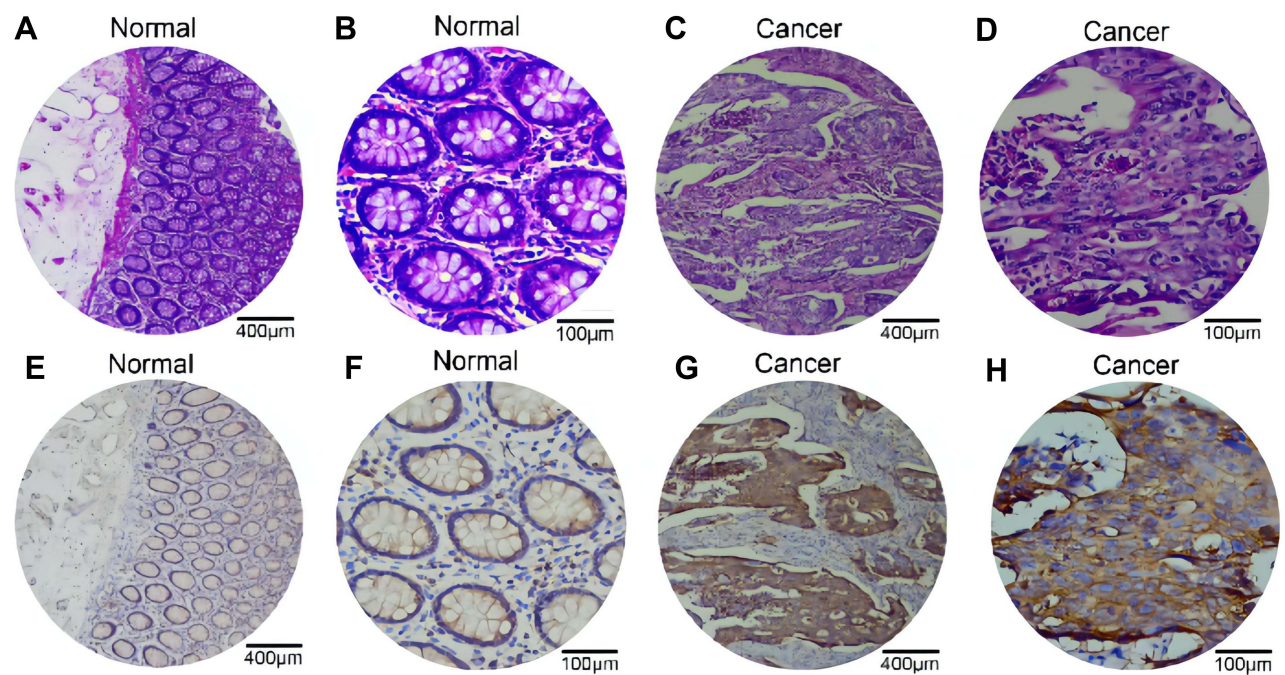
Survival analysis showed that patients with high Ano1 gene expression had poorer survival rates than those in the low-expression group. That is that higher the value, the worse the prognosis (Figure 1B).

### IHC Results of the Ano1 Gene

The Ano1 gene was validated using HE staining and IHC (Figure 2). Using HE staining, we found that the glands in the normal colonic mucosa were polarly arranged (Figure 2A and B). The tumor tissue in the colorectal adenocarcinoma was irregularly nested (Figure 2C). High-level magnification showed local necrosis of the colon adenocarcinoma tissue, increase in tumor cell volume, large nucleus that was highly stained, increase in nucleoplasm ratio, and mitotic images



**Figure 1** Differentially expressed AnoI gene and Kaplan-Meier curve. **(A)** Violin plot shows the differential expression value of AnoI between tumor and paracancerous control group. The difference is statistically significant with the P-value in Student's t-test. **(B)** The Kaplan-Meier survival curve shows a significant difference in survival between these two groups. Statistical significance P value was calculated using the log-rank test ( $P=0.047$ ).

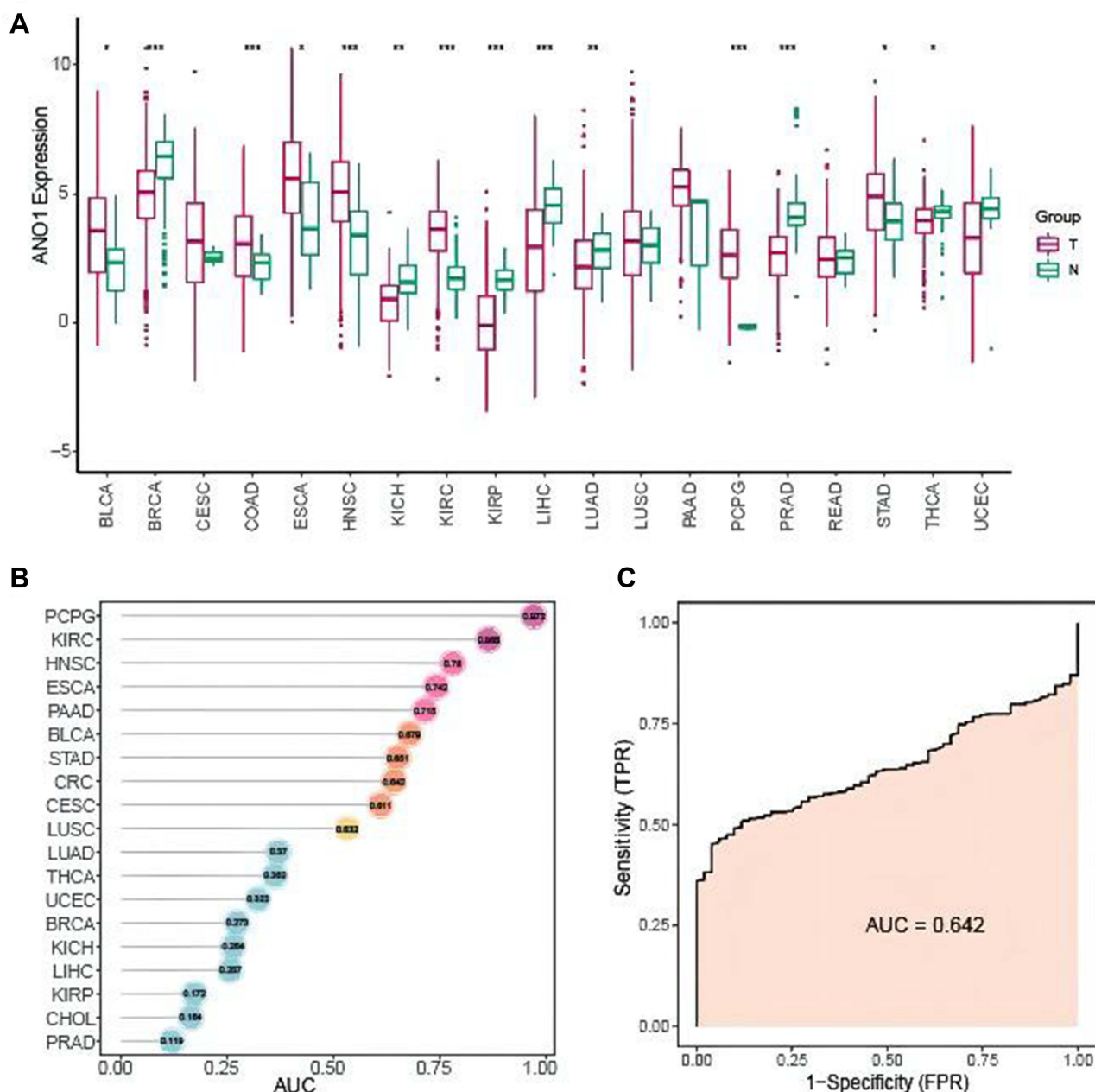


**Figure 2** HE staining and IHC validated AnoI was expressed higher in CRC tissues than normal tissues. HE staining: **(A and B)** Normal colonic mucosa with polarly arranged glands. **(C)** Colorectal adenocarcinoma with irregular nesting of tumor tissue. **(D)** At high magnification, the colon adenocarcinoma showed local necrosis, increased tumor cell volume, hyperchromatic nuclei, increased nucleo-plasma ratio and mitotic images. Immunohistochemical staining of ANO1 protein: **(E and F)** The expression of ANO1 protein was negative in normal colon mucosal glandular epithelial cells. **(G and H)** The expression of ANO1 protein was strongly positive in the cytoplasm of colorectal adenocarcinoma cells, but negative in the surrounding stroma and lymphocytes.

were visible (Figure 2D). We observed the expression of ANO1 protein was negative in normal colon mucosal glandular epithelial cells (Figure 2E and F). The expression of ANO1 protein was strongly positive in the cytoplasm of colorectal adenocarcinoma cells, but negative in the surrounding stroma and lymphocytes (Figure 2G and H).

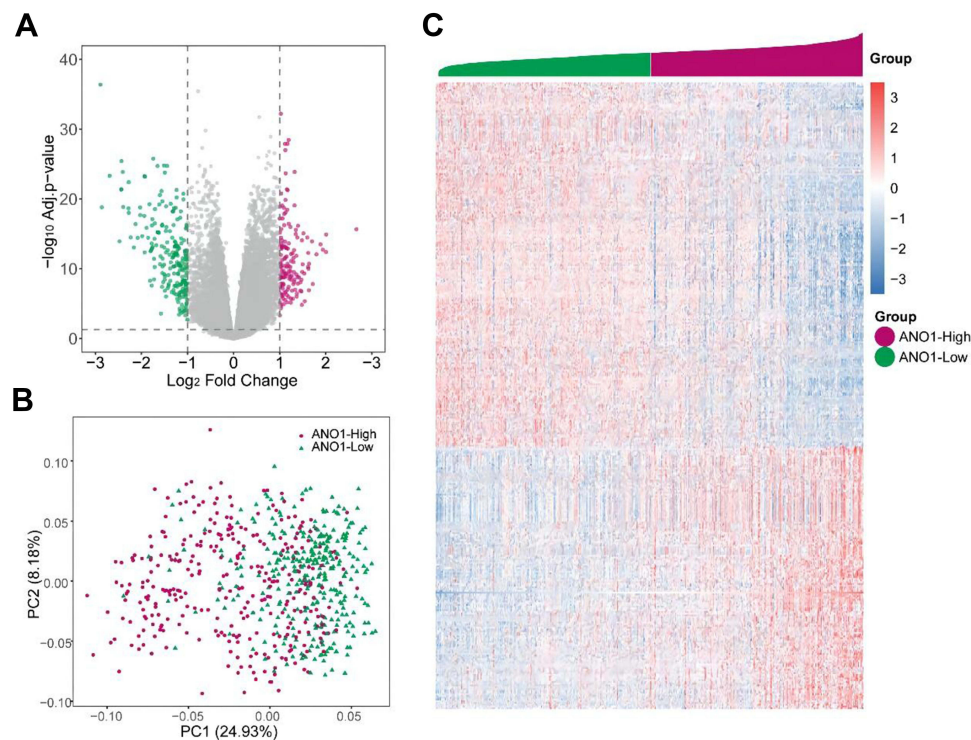
## Pan Cancer Differential Expression Analysis of AnoI in TCGA

After the sample was screened as described in 2.2, 19 tumor types were finally enrolled. A total of 7833 samples were obtained, including 7126 tumor and 707 adjacent tissue samples. The least number of samples were CHOL ( $n = 45$ ) and



**Figure 3** Differential expression and classification of ANO1 gene in multiple tumors. **(A)** The box plot shows gene expression levels of Ano1. The abscissa represents various tumor types, where the ordinate axis is the expression value of Ano1 (log2 FPKM). The tumor group (T) is represented by red and the control group is represented by green (N). **(B)** AUC is used to demonstrate the classification of Ano1 for tumor and normal samples. Using AUC as a horizontal coordinate, and tumor type as a longitudinal coordinate. The ability of classification decreased gradually from the top to the bottom. **(C)** The ROC curve shows the diagnostic value of cAno1 for CRC. **Note:** \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

the most were BRCA ( $n = 1204$ ). The differential expression analysis showed that the Ano1 gene had significant gene expression differences across the 14 tumor types (Figure 3A). The results of the ROC-AUC further revealed that its classification effect on the nine tumors was quite good ( $AUC > 0.6$ ). The difference was more pronounced in tumor types (such as PCPG, KIRC, and HNSC) with better classification efficacy ( $AUC > 0.75$ ) (Figure 3B). In addition, Ano1 showed a better classification efficacy for CRC ( $AUC = 0.642$ ) (Figure 3C).



**Figure 4** Analysis of discrepancies based on AnoI gene expression profiles. **(A)** Genes differentially expressed of AnoI-high group is displayed by volcano plot. Here, red represents high expression and green represents low expression. **(B)** highlights the results of PCA. The red dots represent AnoI-high and the green triangles represent AnoI-low. **(C)** A heatmap of differential gene expression. The red represents AnoI-high group and the green represents AnoI-low group. In this heatmap, AnoI gene expression values increased gradually from left to right. The redder the color is, the higher the expression level.

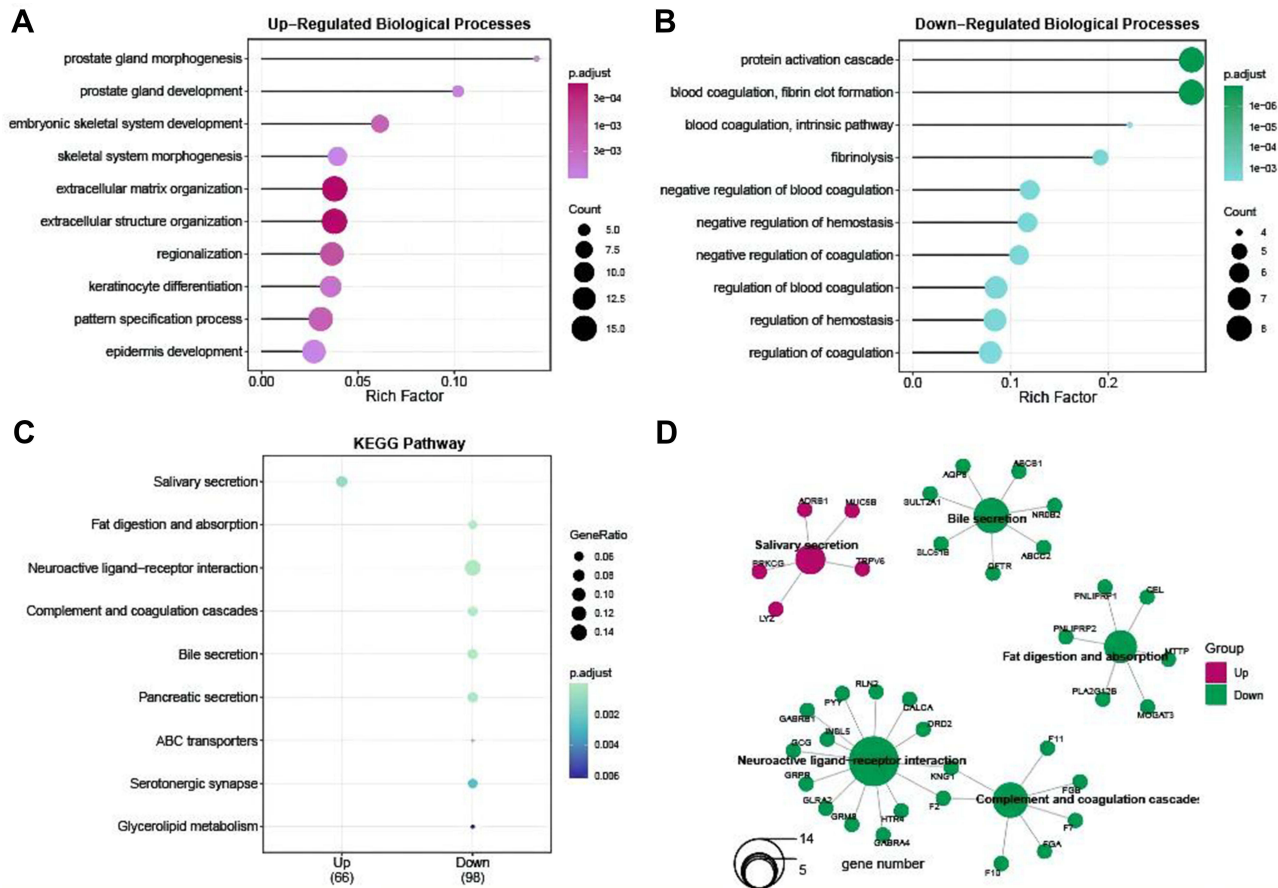
## Differential Analysis of the AnoI Expression Profiles

The TCGA-CRC dataset was pre-processed by filtering the dataset as described in 2.3, and a total of 15,057 genes remained. Thereafter, we obtained 358 differential genes through differential gene screening (Ano1-high VS Ano1-low), which included 150 upregulated genes and 208 downregulated genes (Figure 4A). We analyzed the differential gene expression values using Principal component analysis (PCA). The results showed that samples from the same category tended to aggregate. Therefore, we can use the differential gene expression value to discriminate between the Ano1-high and Ano1-low groups (Figure 4B). Based on the different expression levels of AnoI gene, the differential genes showed a convergent or opposite trend in the heatmaps (Figure 4C).

## Functional Enrichment and Analyses of AnoI

The GO enrichment analysis showed that the upregulated genes were enriched in a variety of biological processes, such as extracellular matrix tissue, regionalization, embryonic skeletal, and system development. The downregulated genes were enriched in the protein activation cascade, coagulation, fibrin clot formation, fibrinolysis, and hemostatic negative regulation (Figure 5A and B, Schedule 1). The KEGG analysis indicated that the upregulated genes were mainly involved in the salivary secretion pathway, while the downregulated genes are mainly concentrated in fat digestion and absorption, neuroactive ligand receptor interaction, complement and the coagulation cascade pathway (Figure 5C and D, Schedule 2).

The MSigDB Hallmark GSEA analysis showed that the differentially expressed genes were positively correlated with Cancer Hallmarks. The pathways involved included allograft rejection, epithelial mesenchymal transition, and inflammatory response. The DEGs were negatively correlated with pathways including myc targets v1 and oxidative phosphorylation (Figure 6A–I, Schedule 3). The MSigDB Hallmark GSVA analysis also showed the degree of variation of each tumor sample at the level of partial Cancer Hallmark gene sets (Figure 7A). We found that the 5 gene sets labeled with the cancer markers included interferon gamma response, allograft rejection, TNF- $\alpha$  signaling via NF- $\kappa$ B, inflammatory



**Figure 5** GO and KEGG-pathway analysis of differential genes. **(A)** GO-BP enrichment analysis for up-regulated genes. The color represents the adjusted p value; the size of dots represents the proportion of enriched genes. **(B)** GO-BP enrichment analysis for down-regulated genes. **(C)** KEGG pathway enrichment analysis. **(D)** KEGG pathway and gene network are presented. The color to guide grouping the differential genes as shown. The size of central node presents the number of genes in these pathways.

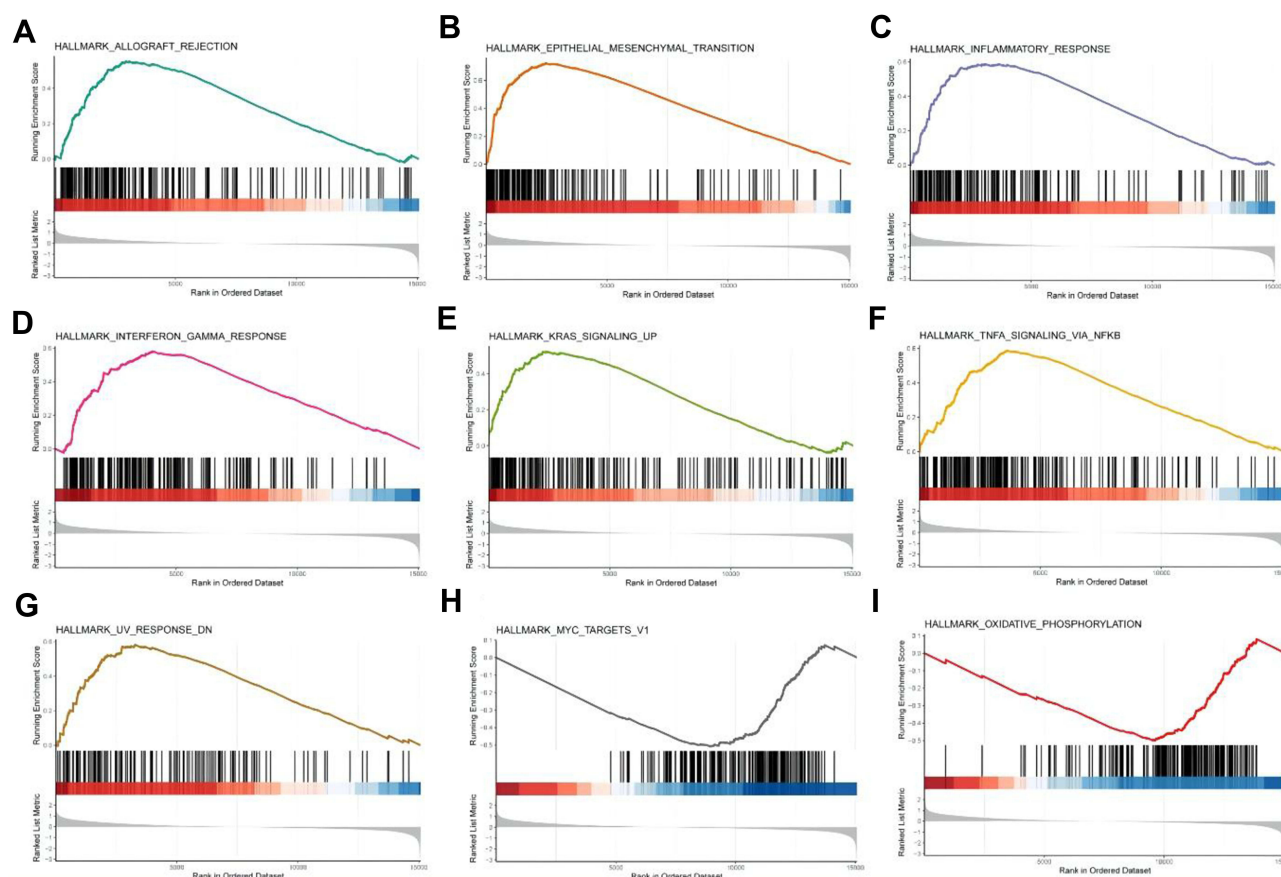
response, and KRAS signaling, which were clustered. The other 5 gene sets included ultraviolet response downregulation, apical junction, epithelial mesenchymal transition, MYC targets v1, and oxidative phosphorylation, which were also clustered. Details of the Cancer Hallmark gene sets are provided on the website. In addition, we performed a correlation analysis between the ANO1 expression level and the variation of all Hallmark gene sets. We observed a weak correlation between the level of Ano1 expression and variation in the Estrogen Response Early/Late pathway (Figure 7B and C).

# WGCNA Analysis

The  $\lambda$  value (Soft Thresholding Power) of the set of 5 was chosen as previously described. Then, we obtained 7 modules (besides the grey module), which contained a total of 341 genes (gold: 145; steel blue: 58; turquoise: 56; slate blue: 23; salmon: 21; tan: 20; light green: 18) (Figure 8A). The module-sample trait correlation analysis showed that there was a high-level of correlation between the Ano1 group and each module ( $|r| > 0.3$ ). Five of these modules were negatively correlated with Ano1-high (gold, turquoise, slate blue, salmon, and light green), and the other two were positively correlated with Ano1-high (steel blue, tan) (Figure 8B). Then, we counted the categories of genes (including upregulated and downregulated genes) in each module. The results corresponded with the results of the correlation analysis. The module with a higher proportion of upregulated genes tended to be positively correlated with Ano1-high (Figure 8C).

# PPI Network Construction and Key Node Analysis

Based on the description provided in the methods section, we constructed the PPI network of the top three modules (gold, steel blue, and turquoise) in terms of the number of genes (Figure 9A–C). PPI-gold contains 94 nodes (42 upregulated



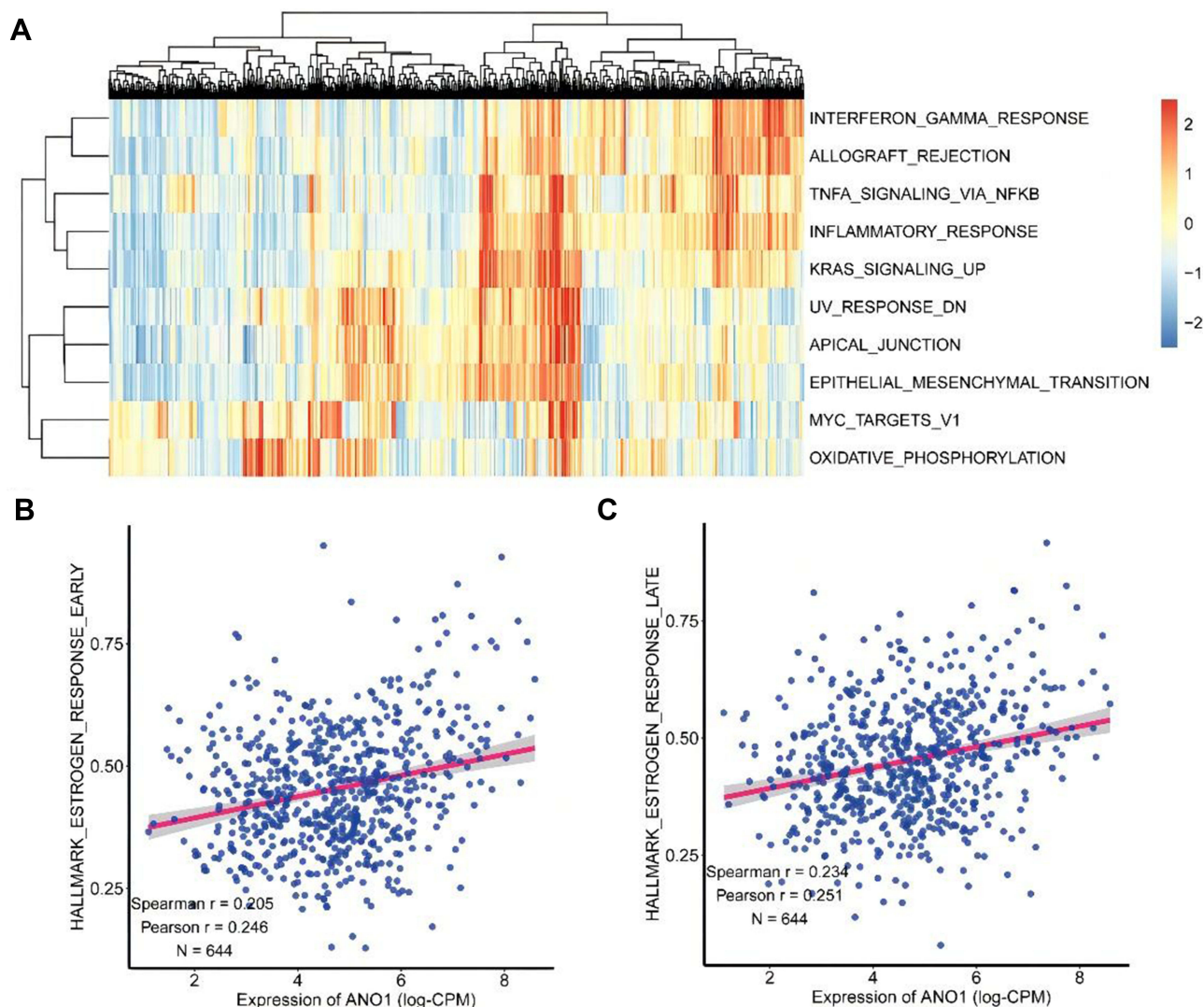
**Figure 6** The Top 9 results of the GSEA. (A–I) correspond to the top 9 pathways in Schedule 3. (A) Allograft rejection, pathway. (B) Epithelial mesenchymal transition. (C) Inflammatory response. (D) Interferon gamma (IFN- $\gamma$ ) response. (E) KRAS signaling pathway. (F) TNF- $\alpha$  signaling via NF- $\kappa$ B pathway. (G) UV response DN. (H) MYC targets V1. (I) Oxidative phosphorylation.

and 52 downregulated) with 205 relationships; PPI-steel blue contains 33 nodes (31 upregulated and 2 downregulated) with 42 relationships; while PPI-turquoise contains 26 nodes (1 upregulated and 25 downregulated) with 43 relationships. The Hub nodes with higher level rankings were PPI-gold: KNG1 (21), GNG4 (19), F2 (15); PPI-steel blue: POSTN (8), THBS2 (7), SPP1 (6); and PPI-turquoise: FGA (7). These genes frequently interact with other proteins in each module and are important nodes in the network.

## Analysis of Immune Cell Infiltration and Its Association with a Related Diagnostic Markers

The heatmap demonstrates significant differences in partial immune cell infiltration between the Ano1-high and Ano1-low samples. Plasma cells, activated memory CD4(+) T cells, and activated dendritic cells were found in Ano1-high. The degree of infiltration in the group was significantly lower. The infiltration of cells such as activated mast cells and resting memory CD4(+) T cell was significantly higher in the Ano1-high group (Figure 10A). This result indicates that in Ano1-low tumors, cells such as plasma cells, activated CD4(+) T cells, and activated dendritic cells show a high level of infiltration, while in Ano1-high tumors, the degree of infiltration of cells such as activated mast cells and resting CD4(+) T cells is higher in Ano1-low tumors.

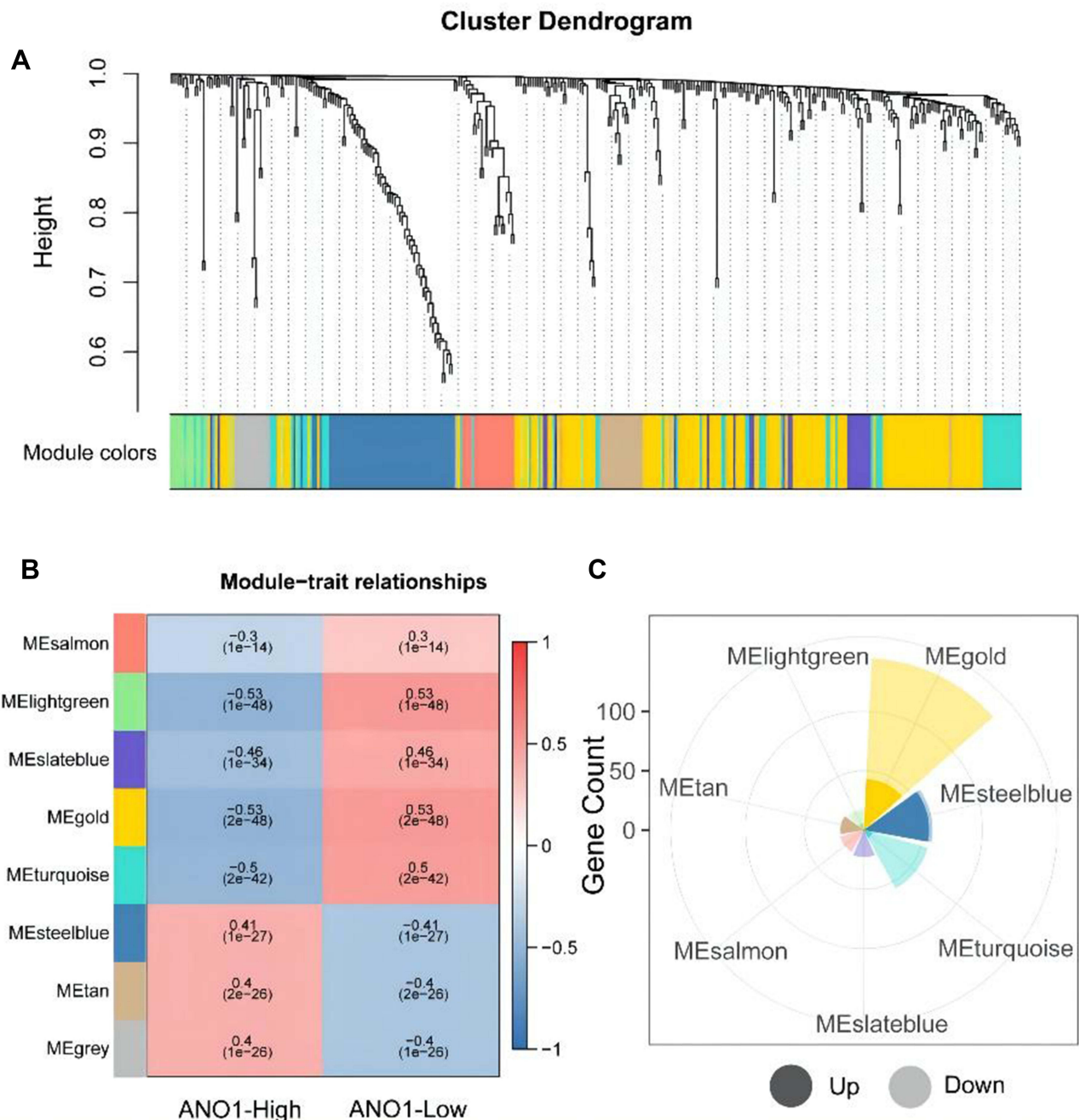
Further analysis of the correlation between immune cell types in the heatmap and ANO1 gene expression showed that plasma cells, activated memory CD4(+) T cells and ANO1 gene expression were significantly negatively correlated (Figure 10B and C), which is consistent with the heatmap results, implying that ANO1 may be involved in the function inhibition of helper T cells.



**Figure 7** Gene set variance analysis. (A) The heat map exhibits partial Cancer Hallmark Set (Hallmark is consistent with Figure 5). (B) Correlation analysis between Ano1 and Estrogen Response Early. The red straight line represents the regression line, with shading to convey 95% confidence intervals. (C) Correlation analysis between Ano1 and Estrogen Response Late.

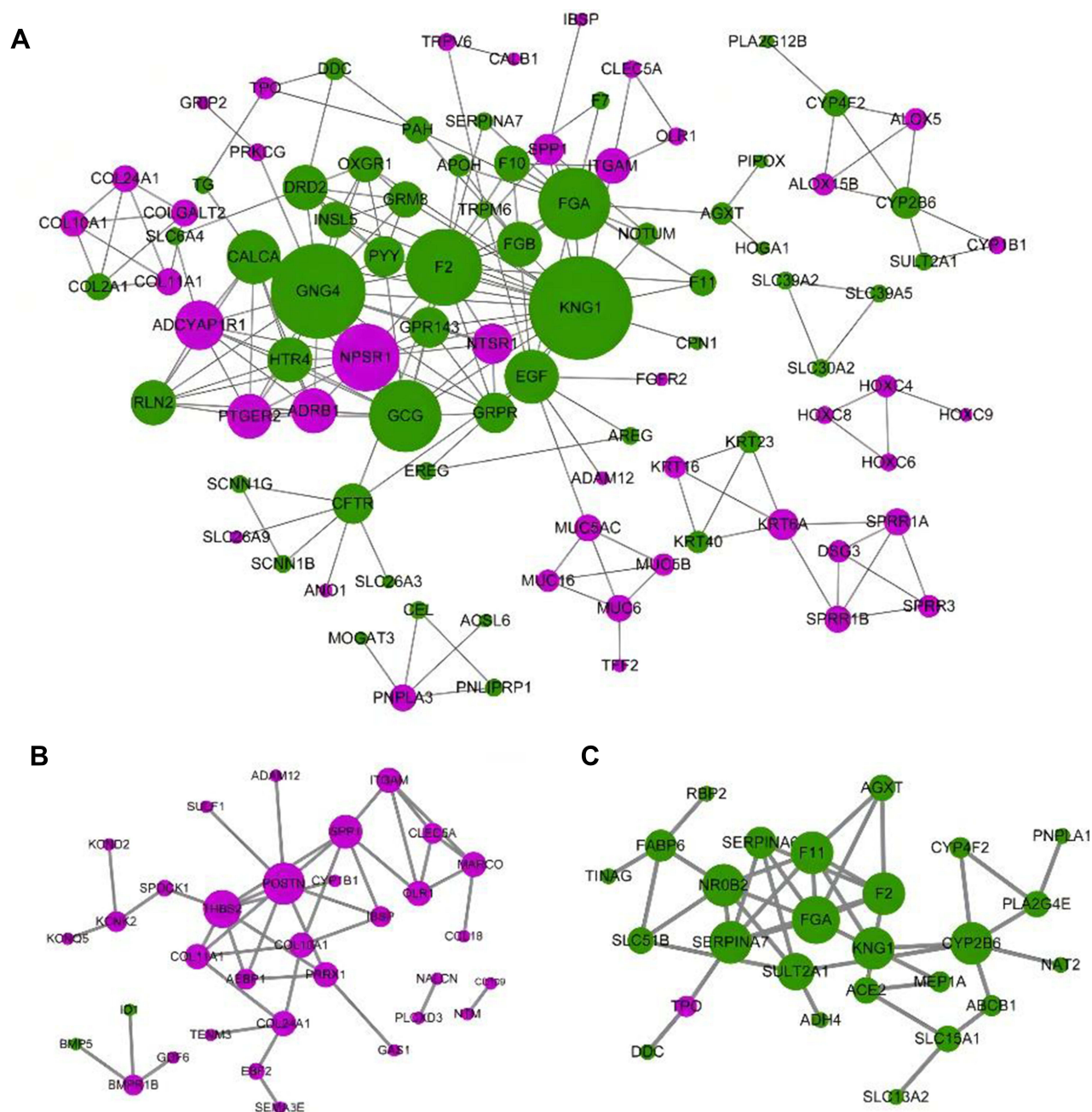
## Construction and Evaluation of a Prognostic Model

The patients were randomly divided into an 80% training set ( $n=493$ ) and a 20% test set ( $n=124$ ). To determine the contribution of each gene in the model (VIMP), the RSF module was constructed based on differential gene expression values. In general, the larger the VIMP, the better the predictive ability; whereas a value close to zero or a negative value indicated a lack of predictive ability. The top 20 ranked genes were used to reconstruct the RSF module (Figure 11A and B). The risk score (Range: 0.012–0.691) for each patient was computed by screening the CHF of important differential genes. The samples were divided into two groups based on their Ano1 expression level, using the median value. The high risk group ( $n=246$ ) was defined as those with a value  $>0.163$  and the low risk group ( $n=267$ ) as those with a value  $\leq 0.163$ . There was a statistical significance in the survival rate between the two groups (log-rank  $P < 0.0001$ ). Compared with the low-risk group, a lower survival trend was observed in the high-risk group (Figure 11C). In the validation set, we used the same module and parameters and used the risk score (0.163) as the threshold. Based on a similar method, the samples were divided into a high-risk group ( $n=48$ ) and a low-risk group ( $n=76$ ). Statistically significant differences in the survival rate were also observed between the two groups (log-rank  $P < 0.0001$ ). The survival rate of the high-risk group was significantly lower compared with the low-risk group (Figure 11D).



**Figure 8** WGCNA analysis. **(A)** WGCNA modules with different colors representing different modules. **(B)** The correlation heat map demonstrates a correlation between module and Anol-group. The rows represent modules and different colors represent different modules. The columns represent Anol-group. In the square boxes, the values present correlation and p-value. **(C)** The statistics for the different categories of genes within modules. The sector area represents the number of genes. The different transparency represents different categories. The dark colors correspond to up-regulated genes and the light colors correspond to down-regulated genes.

We integrated other known clinical variants (such as Age, Tumor Stage, Pathology TNM Stage, Gender, Race and BMI) with risk score described above and constructed a multivariate COX model. After adjusting for other factors, the risk score remained as an independent prognostic risk factor for CRC patients (HR= 59.3,  $p$  value= 7.87E-05). Moreover, the effect of Age, Gender and Pathological M loaded on prognosis was significant (Figure 12A). The calibration curve of the Cox model showed that the predicted 1-year, 3-year, and 5-year values deviated very little from real values. This indicated a reasonably good predictive accuracy (Figure 12B). A ROC curve is constructed using the risk score and the

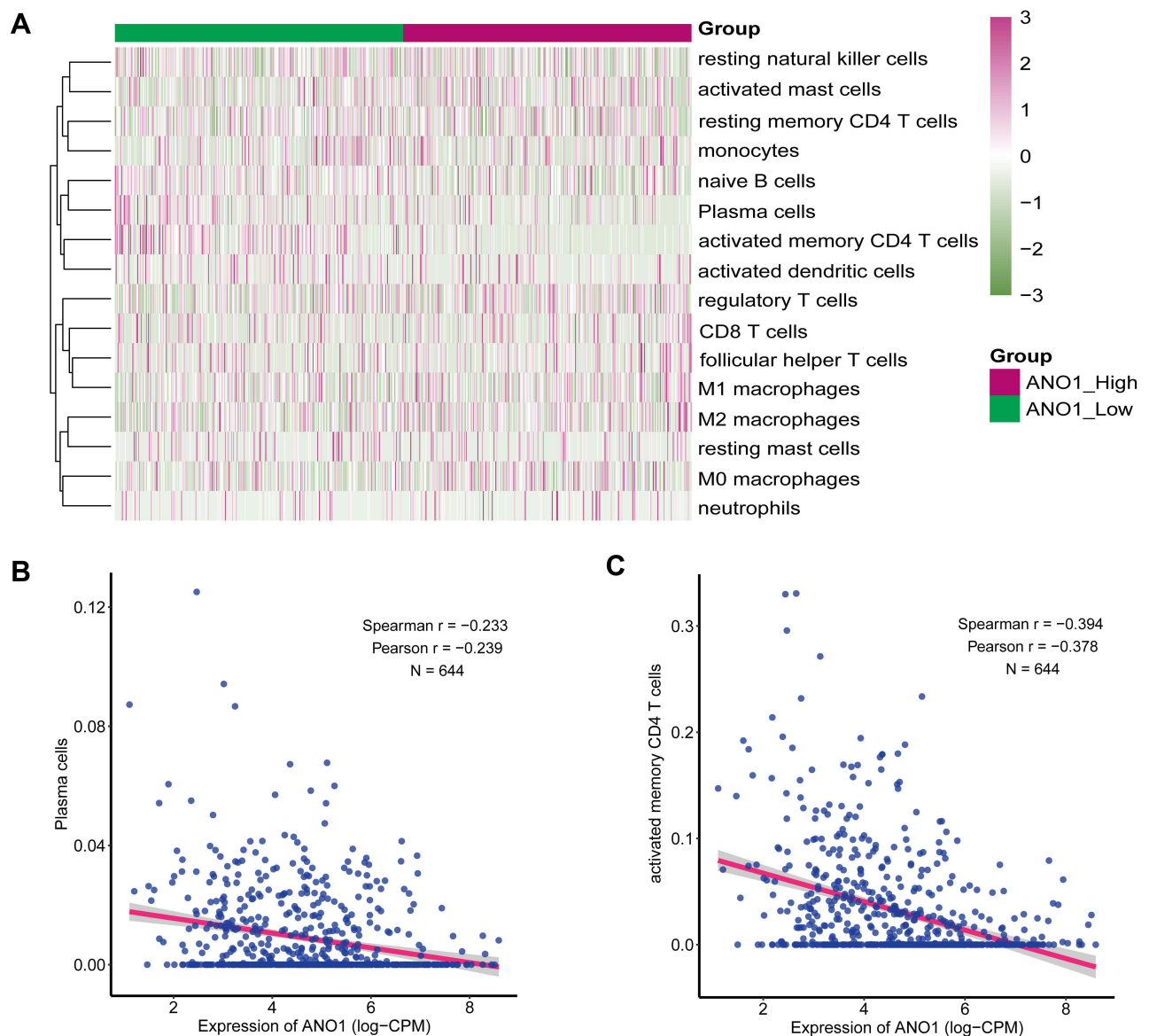


**Figure 9** PPI network. (A) PPI-gold; (B) PPI-steel blue; (C) PPI-turquoise. Here, nodes represent genes and edges represent PPI relationships. Up-regulated genes are shown in red, whereas those down-regulated are in green.

AUC was determined. The AUC value of 1-year, 3-year and 5-year survival were all above 0.95, which indicates that the model possessed a good discrimination ability (Figure 12C).

## External Validation in Expression

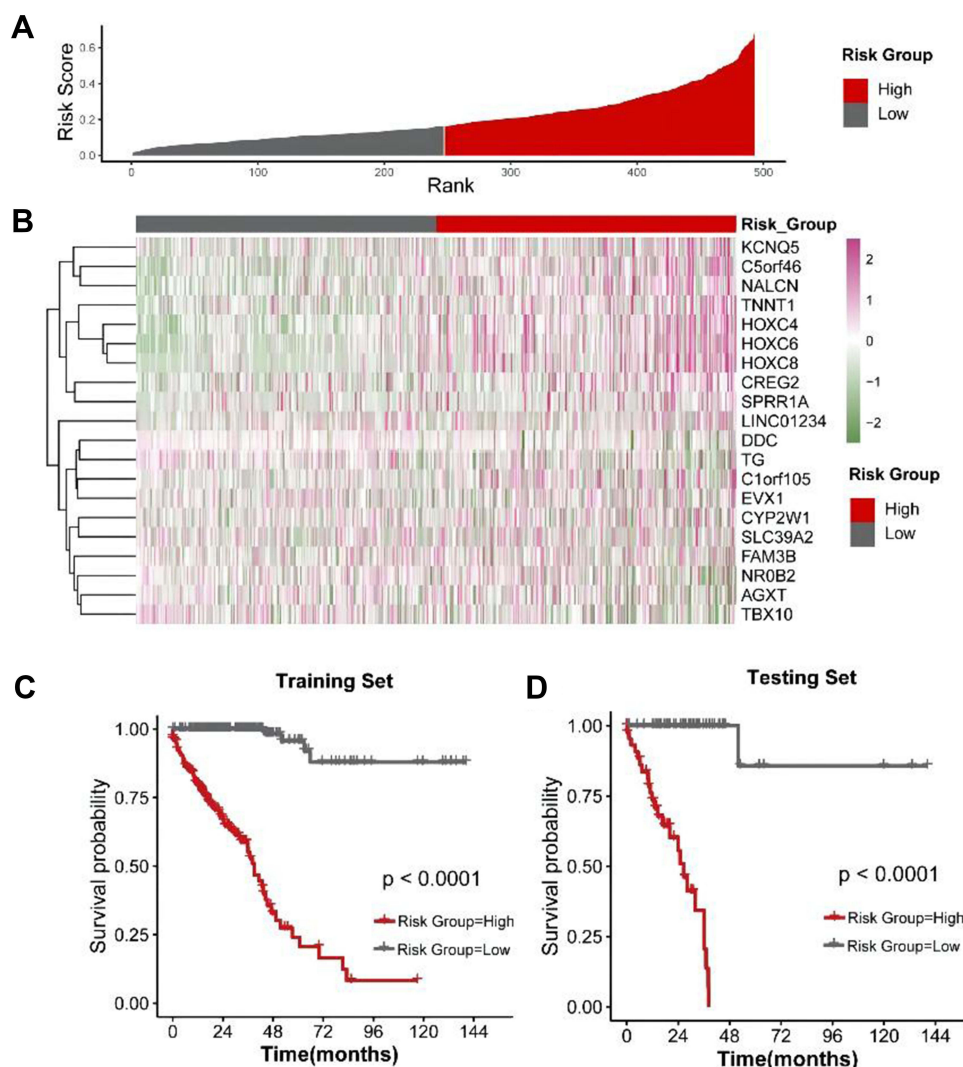
qRT-PCR further revealed significant differential mRNA expression for ANO1 between 6 pairs of CRC cancer and paired paracancer tissues. (Supporting Information Figure S1A and B). In general, the relative expression of Anol mRNA was relatively significantly overexpressed in CRC cancer compared to the paired paracancer tissues.



**Figure 10** Immune cell infiltration can be visualized in heatmap and scatter plot. **(A)** The cluster heat map for the ratio of immune cell infiltration. AnoI-high group and AnoI-low group are indicated by red and green, respectively. The depth of the colors in the heatmap is proportional to the degree of immune cell infiltration. **(B)** The correlations between Plasma cells infiltration and AnoI expression is displayed as a heat map. **(C)** The correlations between activated memory CD4(+) T cell infiltration and AnoI expression is displayed as a heat map.

## Discussion

As a pathogenetically heterogeneous disease, CRC mainly occurs in elderly patients.<sup>26</sup> The prognosis of patients is extremely poor, especially in patients with metastatic colorectal cancer. CRC ranks the second highest for deaths caused by cancer.<sup>1</sup> Distant metastasis typically triggers death and the liver is the most common site of metastasis. For metastatic patients, the 5-year overall survival is approximately 13%.<sup>4,27</sup> Comprehensive treatment strategies for CRC include surgical procedure, chemotherapy, radiation therapy, targeted therapy, and immunotherapy.<sup>28,29</sup> Immunotherapy is a promising field for the further development of treatment methods for CRC.<sup>30</sup> However, there are few targeted or immunotherapy drugs that can serve as first-line drugs for CRC.<sup>31</sup> Recent studies have shown that anti-EGFR antibodies and anti-VEGF antibodies can be used for the treatment of metastatic CRC.<sup>32,33</sup> In this study, we aimed to identify approaches that could be used to target VEGF and EGFR for the treatment of metastatic colorectal cancer.<sup>34,35</sup>



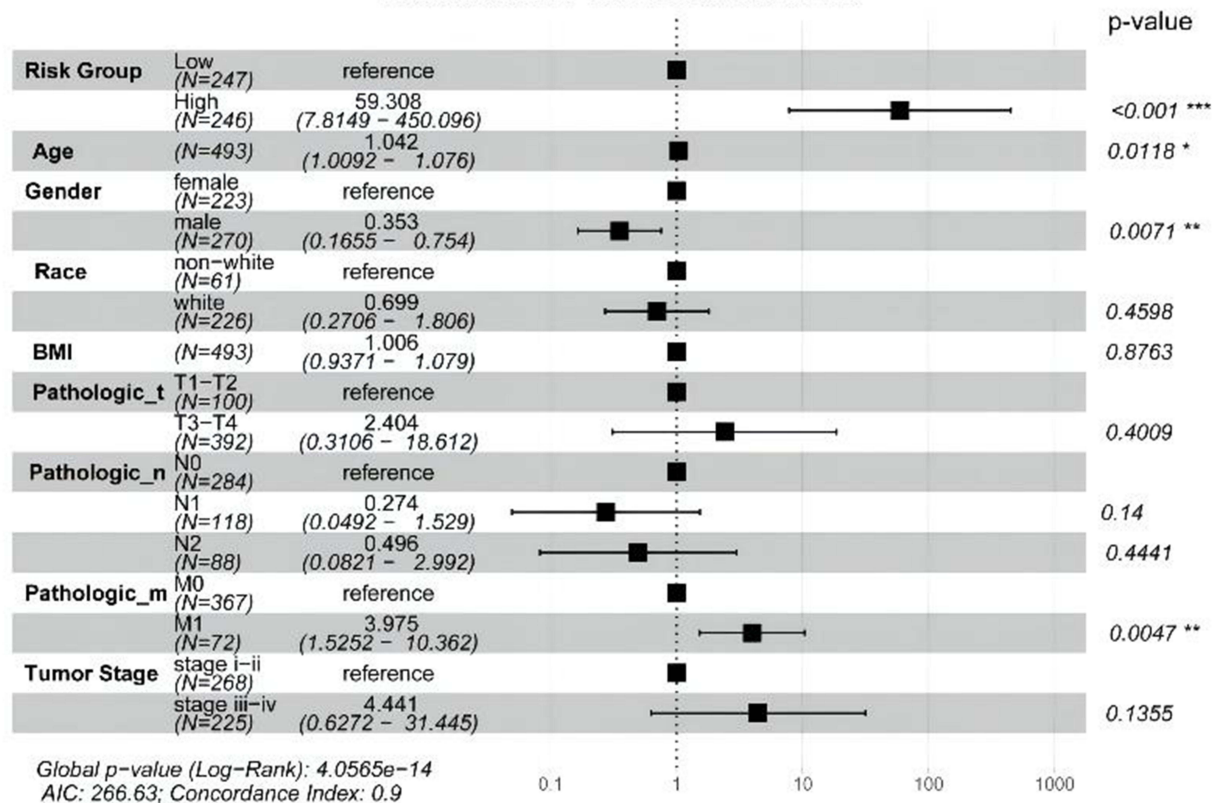
**Figure 11** Results of RSF. **(A)** Risk Score distribution. **(B)** The clustering heat map of expression level for 20 key genes based on RSF model. **(C)** The survival curve of Risk Score for test set. **(D)** The survival curve of validation set Risk Score.

ANO1, which encodes a 960-amino acid protein with 8 transmembrane domains, is highly deregulated in different tumors.<sup>7,8,13,36</sup> As an oncogenic factor, Ano1 can promote cell movement, adhesion, and invasion. Previous studies have found that ANO1 can inhibit the growth and invasion of ovarian cancer by blocking the PI3K/Akt signaling pathway, which is useful in prognosis analysis.<sup>14</sup> The upregulation of ANO1 expression can be used as a risk index for head and neck squamous cell carcinoma (HNSCC).<sup>15</sup> One study showed that ANO1 plays a crucial role in CRC progression through miR-132.<sup>37</sup> As a calcium-activated chloride channel, Ano1 may be a potential novel target for cancer therapy. However, the mechanism of action of Ano1 in distant metastatic CRC needs to be explored further.

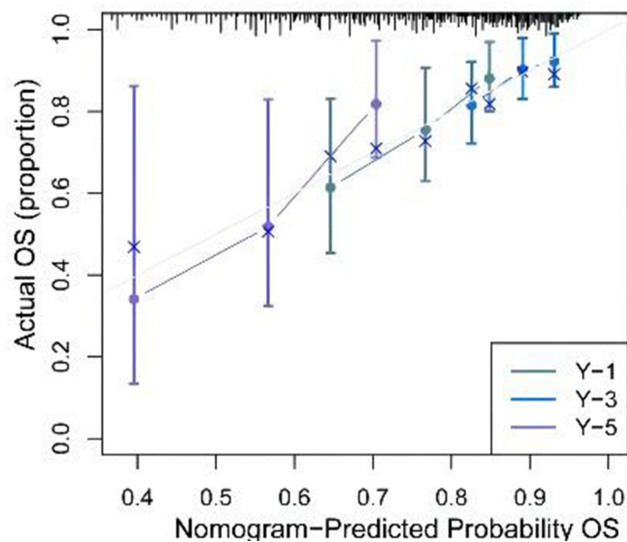
In this paper, Ano1 was found to be associated with several biological processes through a gene co-expression analysis. We know that Ano1 can control a variety of cell functions, including cancer cell proliferation and migration, smooth muscle contraction, epithelial cell secretion and neuronal excitation.<sup>18</sup> To study the potential function of Ano1 in CRC, we performed a GO enrichment analysis, while KEGG, GSEA, and GSVA analyses were performed on the co-expressed genes. The GSEA analysis showed that the differentially expressed genes were positively correlated with epithelial mesenchymal transition. Epithelial mesenchymal transition (EMT) plays an important role in promoting tumor invasion and metastasis.<sup>38,39</sup> Studies have shown that the upregulation of EMT can promote the invasion and metastasis of colon cancer.<sup>40</sup>

A

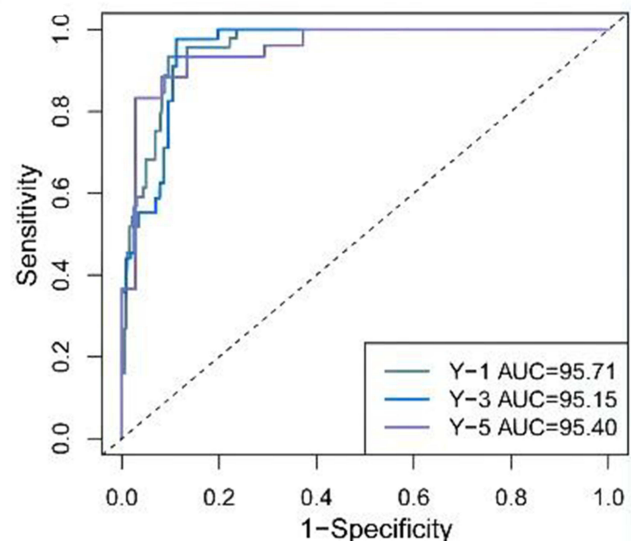
## Hazard Ratio of TCGA-CRC Cox Model



B



C



**Figure 12** The results and evaluation of multivariable Cox model. (A) Forest plots of multivariable Cox model. The point represents HR, and the line represents 95% CI. (B) The calibration curve of the model. The abscissa is the predicted survival rate, and the ordinate is the actual survival rate. The diagonal line represents that the predicted probability is equal to the actual probability. The farther deviated from the diagonal line, the larger the prediction error. 1-year, 3-year, and 5-year survival rates are colored in green, blue, and purple, respectively. (C) The ROC curve for classification efficacy of survival rate based on Risk Score. 1-year, 3-year, and 5-year AUC curves are colored in green, blue, and purple, respectively.

**Note:** \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

Thereafter, we used WGCNA to establish seven modules to study the relationship between Ano1 and its co-expressed genes. The results showed that Ano1 was highly expressed and the proportion of upregulated genes was high. Through screening, KNG1, GNG4, F2, POSTN, THBS2, and SPPIFGA, were mined as the main nodes in the PPI network. This provides a basis for studying the mechanism of action of Ano1. Tumor infiltrating lymphocytes (TILs), which include macrophages, T cells and NK cells, are closely related to the progression and prognosis of CRC. These cells can secrete various factors that affect the microenvironment outside the tumor to regulate tumor behavior.<sup>41</sup> We found that there were significant differences in the infiltration levels of certain types of immune cells between the two groups, which demonstrated that the expression of Ano1 can affect immune cell infiltration. Further studies found that the expression level of Ano1 was negatively correlated with the infiltration degree of plasma cells and memory activated CD4 + T cells. This indicates that Ano1 can affect the prognosis of CRC by affecting the level of immune infiltration. In addition, we used the random survival forest (RSF) method to construct a model of the DEGs related to prognosis. The ROC curves of the model for 1-year, 3-year and 5-year OS show that the model has a good prediction ability.

However, this study has certain limitations. First, the samples were obtained from TCGA database and not from the Second Hospital of Dalian Medical University. Therefore, we need to validate our hypothesis using animal models. We will perform in vivo and in vitro tests to further confirm our conclusions.

In conclusion, our study verified the value of Ano1 for the diagnosis and for predicting the prognosis of CRC. Ano1 may not only participate in metastasis and progression but may also perform an immune regulation function in CRC. Higher expression levels of Ano1 were correlated with a poorer prognosis. Thus, Ano1 can act as a potential prognostic biomarker and a novel target for CRC therapy. To explore the mechanism of action of Ano1 for the pathogenesis of CRC and immunotherapy for CRC seems to be promising. Therefore, we need to carry out further in-depth research before our results can be applied in a clinical setting.

## Conclusion

In summary, an integrated bioinformatics analysis was performed in the current study to explore the relationship between Ano1 and metastatic colorectal cancer. We confirmed that the high expression of Ano1 is closely related to the poor prognosis of patients with colorectal cancer. Ano1 can act as a potential prognostic biomarker and a novel target for CRC therapy. However, additional molecular experiments are required to further validate the findings of the present study.

## Research Ethics and Consent

All participants have been informed why the research is being conducted, whether or not anonymity is assured, and how the data they are collecting is being stored. Subjects have accurately informed the subjects of all the details related to this study, especially of the possible risks and benefits arising from participation in this study. Subjects read the informed consent form carefully and comprehensively, had the opportunity to raise questions and get answers, and gave voluntary consent to participate in this study. Ethics approval from ethics committee of Second Hospital of Dalian Medical University (approval number: 2021-079) has been obtained prior to conducting study.

## Author Contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

No funding was received for this article.

## Disclosure

The authors declare no conflict of interest.

## References

- Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249. doi:10.3322/caac.21660
- Yang Y, Han Z, Li X, Huang A, Shi J, Gu J. Epidemiology and risk factors of colorectal cancer in China. *Chin J Cancer Res*. 2020;32(6):729–741. doi:10.21147/j.issn.1000-9604.2020.06.06
- Tricoli JV, Boardman LA, Patidar R, et al. A mutational comparison of adult and adolescent and young adult (AYA) colon cancer. *Cancer*. 2018;124(5):1070–1082. doi:10.1002/cncr.31136
- Wasan HS, Gibbs P, Sharma NK, et al. First-line selective internal radiotherapy plus chemotherapy versus chemotherapy alone in patients with liver metastases from colorectal cancer (FOXfire, SIRFLOX, and FOXfire-global): a combined analysis of three multicentre, randomised, Phase 3 trials. *Lancet Oncol*. 2017;18(9):1159–1171. doi:10.1016/S1470-2045(17)30457-6
- Wookey V, Grothey A. Update on the role of pembrolizumab in patients with unresectable or metastatic colorectal cancer. *Therap Adv Gastroenterol*. 2021;14:17562848211024460. doi:10.1177/17562848211024460
- Nguyen M, Tipping Smith SS, Lam M, et al. An update on the use of immunotherapy in patients with colorectal cancer. *Expert Rev Gastroenterol Hepatol*. 2021;15(3):291–304. doi:10.1080/17474124.2021.1845141
- Jansen K, Büschek F, Moeller K, et al. DOG1 is commonly expressed in pancreatic adenocarcinoma but unrelated to cancer aggressiveness. *PeerJ*. 2021;9:e11905. doi:10.7717/peerj.11905
- Duvvuri U, Shiwarshi DJ, Xiao D, et al. TMEM16A induces MAPK and contributes directly to tumorigenesis and cancer progression. *Cancer Res*. 2012;72(13):3270–3281. doi:10.1158/0008-5472.Can-12-0475-t
- Kashyap MK, Marimuthu A, Kishore CJ, et al. Genomewide mRNA profiling of esophageal squamous cell carcinoma for identification of cancer biomarkers. *Cancer Biol Ther*. 2009;8(1):36–46. doi:10.4161/cbt.8.1.7090
- Benedetto R, Cabrita I, Schreiber R, Kunzelmann K. TMEM16A is indispensable for basal mucus secretion in airways and intestine. *FASEB J*. 2019;33(3):4502–4512. doi:10.1096/fj.201801333RRR
- Peters CJ, Gilchrist JM, Tien J, et al. The sixth transmembrane segment is a major gating component of the TMEM16A calcium-activated chloride channel. *Neuron*. 2018;97(5):1063–1077.e1064. doi:10.1016/j.neuron.2018.01.048
- Filippou A, Pehkonen H, Karhemo PR, et al. ANO1 expression orchestrates p27Kip1/MCL1-mediated signaling in head and neck squamous cell carcinoma. *Cancers*. 2021;13:5. doi:10.3390/cancers13051170
- Song Y, Gao J, Guan L, Chen X, Gao J, Wang K. Inhibition of ANO1/TMEM16A induces apoptosis in human prostate carcinoma cells by activating TNF- $\alpha$  signaling. *Cell Death Dis*. 2018;9(6):703. doi:10.1038/s41419-018-0735-2
- Nishida T, Sakai Y, Takagi M, et al. Adherence to the guidelines and the pathological diagnosis of high-risk gastrointestinal stromal tumors in the real world. *Gastric Cancer*. 2020;23(1):118–125. doi:10.1007/s10120-019-00966-4
- Wang H, Yao F, Luo S, et al. A mutual activation loop between the Ca(2+)-activated chloride channel TMEM16A and EGFR/STAT3 signaling promotes breast cancer tumorigenesis. *Cancer Lett*. 2019;455:48–59. doi:10.1016/j.canlet.2019.04.027
- Zhang C, Li H, Gao J, Cui X, Yang S, Liu Z. Prognostic significance of ANO1 expression in cancers. *Medicine*. 2021;100(4):e24525. doi:10.1097/md.00000000000024525
- Jiang Y, Cai Y, Shao W, et al. MicroRNA-144 suppresses aggressive phenotypes of tumor cells by targeting ANO1 in colorectal cancer. *Oncol Rep*. 2019;41(4):2361–2370. doi:10.3892/or.2019.7025
- Colaprico A, Silva TC, Olsen C, et al. TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res*. 2016;44(8):e71. doi:10.1093/nar/gkv1507
- Law CW, Alhamdoosh M, Su S, et al. RNA-seq analysis is easy as 1-2-3 with limma, glimma and edgeR. *F1000Research*. 2016;5:1408. doi:10.12688/f1000research.9005.3
- Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417–425. doi:10.1016/j.cels.2015.12.004
- Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform*. 2013;14:7. doi:10.1186/1471-2105-14-7
- Tang Y, Li M, Wang J, Pan Y, Wu FX. CytoNCA: a Cytoscape plugin for centrality analysis and evaluation of protein interaction networks. *Bio Systems*. 2015;127:67–72. doi:10.1016/j.biosystems.2014.11.005
- He X, Zhang J. Why do hubs tend to be essential in protein networks? *PLoS Genet*. 2006;2(6):e88. doi:10.1371/journal.pgen.0020088
- Newman AM, Steen CB, Liu CL, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol*. 2019;37(7):773–782. doi:10.1038/s41587-019-0114-2
- Villanueva A, Portela A, Sayols S, et al. DNA methylation-based prognosis and epidrivers in hepatocellular carcinoma. *Hepatology*. 2015;61(6):1945–1956. doi:10.1002/hep.27732
- Aparicio T, Bouche O, Taieb J, et al. Bevacizumab+chemotherapy versus chemotherapy alone in elderly patients with untreated metastatic colorectal cancer: a randomized Phase II trial-PRODIGE 20 study results. *Ann Oncol*. 2018;29(1):133–138. doi:10.1093/annonc/mdx529
- de Andrea CE, Schalper KA, Sanmamed MF, Melero I. Immunodivergence in metastatic colorectal cancer. *Cancer Cell*. 2018;34(6):876–878. doi:10.1016/j.ccell.2018.11.012
- Sun W, Nie W, Wang Z, Zhang H, Li Y, Fang X. Lnc HAGLR promotes colon cancer progression through sponging miR-185-5p and activating CDK4 and CDK6 in vitro and in vivo. *Onco Targets Ther*. 2020;13:5913–5925. doi:10.2147/OTT.S246092
- Lu H, Shi T, Wang M, et al. B7-H3 inhibits the IFN- $\gamma$ -dependent cytotoxicity of V $\gamma$ 9V $\delta$ 2 T cells against colon cancer cells. *Oncoimmunology*. 2020;9(1):1748991. doi:10.1080/2162402X.2020.1748991
- Malonis RJ, Lai JR, Vergnolle O. Peptide-based vaccines: current progress and future challenges. *Chem Rev*. 2020;120(6):3210–3229. doi:10.1021/acs.chemrev.9b00472
- Lin H, Wang B, Yu J, Wang J, Li Q, Cao B. Protein arginine methyltransferase 8 gene enhances the colon cancer stem cell (CSC) function by upregulating the pluripotency transcription factor. *J Cancer*. 2018;9(8):1394–1402. doi:10.7150/jca.23835
- Angelbello AJ, Chen JL, Childs-Disney JL, Zhang P, Wang Z-F, Disney MD. Using genome sequence to enable the design of medicines and chemical probes. *Chem Rev*. 2018;118(4):1599–1663. doi:10.1021/acs.chemrev.7b00504

33. Woolston A, Khan K, Spain G, et al. Genomic and transcriptomic determinants of therapy resistance and immune landscape evolution during anti-EGFR treatment in colorectal cancer. *Cancer Cell*. 2019;36(1):35–50.e39. doi:10.1016/j.ccell.2019.05.013
34. Azad T, Janse van Rensburg HJ, Lightbody ED, et al. A LATS biosensor screen identifies VEGFR as a regulator of the Hippo pathway in angiogenesis. *Nat Commun*. 2018;9(1):1061. doi:10.1038/s41467-018-03278-w
35. Zhang Y-F, Li C-S, Zhou Y, Lu X-H. Effects of propofol on colon cancer metastasis through STAT3/HOTAIR axis by activating WIF-1 and suppressing Wnt pathway. *Cancer Med*. 2020;9(5):1842–1854. doi:10.1002/cam4.2840
36. Liu JJ, He F, Guo SB. TMEM16A overexpression indicates poor prognosis in colorectal cancer patients. *Rev Esp Enferm Dig*. 2021. doi:10.17235/reed.2021.8292/2021
37. Mokutani Y, Uemura M, Munakata K, et al. Down-regulation of microRNA-132 is associated with poor prognosis of colorectal cancer. *Ann Surg Oncol*. 2016;23(Suppl5):599–608. doi:10.1245/s10434-016-5133-3
38. Fukusumi T, Guo TW, Sakai A, et al. The NOTCH4-HEY1 pathway induces epithelial-mesenchymal transition in head and neck squamous cell carcinoma. *Clin Cancer Res*. 2018;24(3):619–633. doi:10.1158/1078-0432.CCR-17-1366
39. Xue J, Li R, Gao D, Chen F, Xie H. CXCL12/CXCR4 axis-targeted dual-functional nano-drug delivery system against ovarian cancer. *Int J Nanomedicine*. 2020;15:5701–5718. doi:10.2147/IJN.S257527
40. Zhang L-M, Wang P, Liu X-M, Zhang Y-J. LncRNA SUMO1P3 drives colon cancer growth, metastasis and angiogenesis. *Am J Transl Res*. 2017;9(12):5461–5472.
41. Li J, Byrne KT, Yan F, et al. Tumor cell-intrinsic factors underlie heterogeneity of immune cell infiltration and response to immunotherapy. *Immunity*. 2018;49(1):178–193.e177. doi:10.1016/j.immuni.2018.06.006

## International Journal of General Medicine

Dovepress

### Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>