

RESEARCH

Open Access



Identification and validation of a CD4⁺ T cell-related prognostic model to predict immune responses in stage III-IV colorectal cancer

Mengting Li^{1,2†}, Weining Zhu^{1,2†}, Yuanyuan Lu^{2,3†}, Yu Shao^{1,2}, Fei Xu^{1,2*}, Lan Liu^{1,2*} and Qiu Zhao^{1,2*}

Abstract

Background CD4⁺ T cells play an indispensable role in anti-tumor immunity and shaping tumor development. We sought to explore the characteristics of CD4⁺ T cell marker genes and construct a CD4⁺ T cell-related prognostic signature for stage III-IV colorectal cancer (CRC) patients.

Method We combined scRNA and bulk-RNA sequencing to analyze stage III-IV CRC patients and identified the CD4⁺ T cell marker genes. Unsupervised cluster analysis was performed to divide patients into two clusters. The LASSO and multivariate Cox regression were performed to establish a prognostic-related signature. RT-qPCR and immunofluorescence staining were performed to examine the expression of ANXA2 in CRC tissue.

Result We found a higher infiltration abundance of activated memory CD4⁺ T cells was associated with improved prognosis in stage III-IV CRC patients. Patients were divided into two subgroups with distinct clinical and immunological behaviors based on CD4⁺ T cell marker genes. And then a prognostic signature consisting of six CD4⁺ T cell marker genes was established, which stratified patients into high- and low-risk groups. Immune spectrum showed that the low-risk group had higher immune cell infiltration than the high-risk group. Furthermore, the risk score of this signature could predict the susceptibility of stage III-IV CRC patients to immune checkpoint inhibitors and chemotherapy drugs. Finally, we validated that ANXA2 was enriched in Tregs and was associated with infiltration of Tregs in CRC tumor microenvironment.

Conclusion The CD4⁺ T cell-related prognostic signature established in the study can predict the prognosis and the response to immunotherapy in stage III-IV CRC patients. Our findings provide new insights for tumor immunotherapy of advanced CRC patients.

Keywords CD4⁺ T cell, scRNA-seq, Bulk data, Colorectal cancer, Prognostic signature

[†]Mengting Li, Weining Zhu and Yuanyuan Lu contributed equally to this work.

*Correspondence:
Fei Xu
328221059@qq.com
Lan Liu
lliugi@whu.edu.cn

Qiu Zhao

qiuzhao@whu.edu.cn

¹Department of Gastroenterology, Zhongnan Hospital of Wuhan University, Wuhan, China

²Hubei Clinical Center and Key Lab of Intestinal and Colorectal Diseases, Wuhan, China

³Department of Gastroenterology, Wuhan Third Hospital, Tongren Hospital of Wuhan University, Wuhan, China



Introduction

The incidence and mortality rate of colorectal cancer (CRC) continues to rise globally. By 2040, there will be 3.2 million new cases and 1.6 million deaths. This trend highlights the urgency of research and treatment for colorectal cancer [1]. In recent years, great progress has been made in surgical therapy, systemic therapy, and other therapeutic strategies, which have greatly improved the overall survival time of CRC patients in an early stage [2]. However, many CRC patients were initially diagnosed at advanced stages with their 5-year survival rate ranging from 14–71% [3]. There are few effective therapies and prognostic biomarkers for CRC patients at advanced stages, thus resulting in inappropriate decision-making and ultimately a poorer prognosis [4]. Therefore, it is urgent to find effective biomarkers related to prognosis and treatment for stage III-IV CRC patients. Currently, existing studies based on microarray or RNA-seq detection data have provided a broad data basis for searching for effective biomarkers. Zheng et al. demonstrated that CALD1 could be used as a prognostic biomarker and a prospective therapeutic target for stage III/IV pMMR CRCs through multiple bioinformatic analyses and cell-level assays [5]. Moreover, Liu et al. established a nomogram model of the SMAD4, ZFH3, and PREX2 mutation status, pathological location, and preoperative CEA value to accurately predict the risk of Stage III/IV CRC patients [6]. Although these biomarkers had made some achievements in monitoring the prognosis of advanced CRC patients, due to the heterogeneity and complex composition of tumors, conventional bulk data detection usually failed to distinguish tumor tissue from non-tumor tissue, such as tumor microenvironment or paracancer tissue, which lead to a certain bias in the predictive performance of these genetic markers for prognostic outcomes in patients with stage III-IV CRC.

Single-cell RNA sequencing (scRNA-seq) allows the definition of molecularly distinct cell subpopulations and reveals the molecular characteristics of diverse cells in the TME at the single-cell level, greatly improving the limitations of traditional bulk data. Previous studies have reported that exploring gene expression signatures based on molecular characteristics of immune cells derived from scRNA-seq data provides new insights into cancer immunity [7, 8]. Since current data obtained by single-cell technology cannot directly link cell types to the clinical phenotype of cancer patients. Several studies have combined single-cell data with bulk data to identify reliable biomarkers that have shown benefits in predicting prognosis and sensitivity to chemotherapeutic drugs in patients with multiple types of cancer, such as hepatocellular carcinoma [9], lung adenocarcinoma [10], breast carcinoma [11]. Therefore, based on bulk RNA-seq and scRNA-seq analysis, comprehensive identification of

predictive biomarkers and novel molecular targets for diagnosis and treatment of stage III-IV CRC patients will contribute to the precise stratification of patients.

During the initiation and progression of tumors, tumor cells are surrounded by a complex array of TME, including kinds of immune cells, stromal cells, extracellular matrix, and various cytokines [12]. Abundant evidence confirms that the heterogeneity of immune cell infiltration in TME is an important factor affecting the efficacy and prognosis of CRC and other tumor types [13, 14]. In the TME, T cells play an important role in immune monitoring and tumor eradication [15]. As one of the most common T cell subtypes, CD4⁺ T cells are closely associated with anti-tumor responses, as they can enhance the killing activity of other anti-tumor effector cells, such as CD8⁺ T cells and macrophages [16, 17]. In several studies, adoptive transfer of tumor-specific CD4⁺ T cells has shown impressive efficacy in tumor suppression [18, 19]. It has been found that circulating CD4⁺CTLA-4⁺ T cells increased in patients with advanced stage, which might be used as a potential biomarker to predict the progression and treatment response of CRC [20]. There are few studies on the antitumor immunity of CD4⁺T cells in advanced CRC. Therefore, it is necessary to investigate the relationship between the gene expression profile of CD4⁺T cell-related markers and the prognosis and therapeutic response of stage III-IV CRC patients. In our study, we performed an integrative analysis of scRNA-seq and bulk RNA-seq of stage III-IV CRC patients to identify CD4⁺ T cell marker genes. We verified the predictive power of the CD4⁺ T cell-related signature and further compared the differences in immune cell infiltration, immune checkpoint response, and sensitivity to chemotherapeutic drugs between the two risk groups. Our study will help to provide recommendations for the clinical treatment of stage III-IV CRC patients.

Method

Data collection

In this study, two scRNA-seq datasets containing clinical information of CRC patients were downloaded from the Gene expression omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>) and used to screen CD4⁺ T cell marker genes of stage III-IV CRC patients. GSE166555 includes 6 stage III-IV CRC patients, and GSE144735 includes 2 stage III-IV CRC patients. To analyze the relationship between the proportion of various immune cells and the prognosis of stage III-IV CRC patients, we downloaded the expression matrix and clinical information of 194 stage III-IV CRC patients from The Cancer Genome Atlas (TCGA) database and the expression matrix and clinical information of 422 stage III-IV CRC patients from the GEO database, including GSE17538 ($n=132$), GSE12945 ($n=26$) and GSE39582 ($n=264$). In addition,

to explore the value of CD4⁺ T cell-related signature in predicting immunotherapy response, 298 samples treated with immunotherapy were obtained from the IMvigor210 cohort. The data of stage III-IV CRC patients in TCGA, GSE17538, GSE12945, and GSE39582 were combined, and the ComBat method was conducted for batch correction. Subsequent analyses involved a total of 616 stage III-IV CRC patients.

Acquisition of CD4⁺ T cell marker genes by scRNA-seq analysis

We performed scRNA-seq data analysis by R packages, including “Seurat” and “SingleR” [21]. To retain high-quality scRNA-seq data, unqualified cells in the integrated dataset were excluded according to the following quality control standards: (1) Cells with mitochondrial gene content of more than 5% were removed, (2) Cells that expressed less than 50 genes and clusters with cell count less than 3 were also removed. After normalizing the data using the “NormalizeData” package, we performed principal component analysis (PCA) to screen out the top 15 principal components (PCs) based on the top 2000 highly variable genes. T-distributed stochastic neighbor embedding (t-SNE) was then used for unsupervised clustering, and cell subpopulations were visualized using T-SN-1 and T-SN-2. The “FindAllMarkers” function in the “Seurat” package was used to calculate the differentially expressed genes (DEGs) of each cluster compared to all other clusters. $|\log_2(\text{fold change})| > 1$ and adjusted P -value < 0.05 were used to identify the marker genes for each cluster. Finally, the MonacoImmuneData dataset was applied to annotate the cell subpopulations of the different clusters. The common CD4⁺ T cell marker genes in GSE166555 and GSE144735 were obtained by venn diagram and included in subsequent analysis.

Consensus clustering and pathway enrichment analysis of CD4⁺ T cell marker genes

We used the “ConsensusClusterPlus” package in the R studio program to classify stage III-IV CRC patients into different molecular subgroups based on CD4⁺ T cell marker gene expression by unsupervised cluster analysis. The consensus clustering algorithm was used to establish the quantity and consistency of clusters. The optimal number of clusters was determined by the k-means algorithm. The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were conducted using the R package “clusterProfiler”.

Construction and validation of a CD4⁺ T cell-related prognostic signature

A univariate Cox regression analysis was used to identify CD4⁺ T cell marker genes associated with prognosis

in stage III-IV CRC patients. Subsequently, to minimize overfitting, the least absolute shrinkage and selection operator (LASSO) Cox regression was performed by the “glmnet” package to determine significant prognostic genes in the training set. The optimum value of penalty parameters was determined by 10-fold cross validation method. Finally, based on the genes selected by LASSO Cox regression analysis, we used stepwise multivariate Cox regression analysis to identify key prognostic genes and establish the CD4⁺ T cell-related signature. According to the expression levels and corresponding coefficients of the signature genes (Table S1), the risk score of each stage III-IV CRC patient was calculated. The formula is: risk score = expression level of a gene (a) × correspondence coefficient (a) + expression level of a gene (b) × correspondence coefficient (b) + expression level of a gene (n) × correspondence coefficient (n). The patients were divided into high-risk group and low-risk groups based on the median risk score. The Kaplan-Meier survival curve was generated using the R package “survminer” to compare the survival differences between the high-low risk groups. The time-dependent receiver operating characteristic (ROC) curves and the area under the curve (AUC) were measured by the “survival-ROC” package to assess the accuracy of the prognostic signature. The nomogram was established by combining clinicopathological characteristics and risk scores using the “rms” package. The calibration curve analysis was performed to confirm the accuracy of the established nomogram.

Evaluation of tumor immune microenvironment

To assess the specific cellular components of the immune microenvironment, the CIBERSORT algorithm was applied to calculate the relative proportions of 22 immune cells in the CRC samples obtained in each sample, and samples with $p < 0.05$ were screened for subsequent analysis. The components of the immune and stromal cells in the tumor microenvironment (TME) of stage III-IV CRC patients were calculated to compare the differences in microenvironment characteristics between different CD4⁺ T cell-related subgroups or risk groups by the ESTIMATE algorithm. Single sample Gene Set Enrichment Analysis (ssGSEA) was used to further compare the activity of immune cells and the immune function of each sample in different risk groups. Moreover, we calculated the tumor mutation burden (TMB) score for each CRC patient in the high-risk and low-risk groups.

Drug sensitivity analysis

The Drug sensitivity data were downloaded from the Genomics of Drug Sensitivity in Cancer database (GDSC, <https://www.cancerrxgene.org/>). The “oncoDirect” R package was used to evaluate differences in the

half-maximal inhibitory concentration (IC50) values of common CRC chemotherapeutic drugs between the high-risk and low-risk groups.

Patients

CRC patients undergoing radical resection of CRC in Zhongnan Hospital of Wuhan University were collected as study objects. The acquisition of clinical information and tissue samples of patients has been approved by the Ethics Committee of Zhongnan Hospital of Wuhan University and informed consent of patients (ID:2020110). We collected tumor and paracancer tissue samples from 5 CRC patients undergoing endoscopic or surgical resection from Zhongnan Hospital of Wuhan University. The inclusion criteria were as follows: 1. Age ≥ 18 years old, ≤ 75 years old, male and female; 2. Patients with unresectable metastatic colon or rectum adenocarcinoma confirmed by histopathology; 3. Unresectable metastases have not received any systemic antitumor therapy; 4. For subjects who have previously received neoadjuvant or adjuvant therapy, the date of first discovery of progression must be at least 6 months away from the date of last administration of neoadjuvant or adjuvant therapy.

Immunofluorescence staining

Colorectal tumor and adjacent normal paraffin sections of CRC patients were incubated sequentially with two primary antibodies, ANXA2 (rabbit, 1:200, ProteinTech, Wuhan, China), Foxp3 (mouse, 1:500, Abcam, Shanghai, China). The nucleus was counterstained with DAPI (Biaqiandu technology, Wuhan, China). The stained sections were observed using a fluorescence microscope (Olympus, Japan).

Quantitative real-time PCR

Total RNA from CRC tumor and adjacent colon or rectum tissues was extracted by TRIzol reagent (Invitrogen, USA). The first-strand cDNA was synthesized using total RNA with a reverse transcription kit (TOYOBOO, Japan). Real-time PCR was performed using SYBR Green RT-PCR kit (CWbio, China). The primers were synthesized by Tsingke Biological Technology (Beijing, China). The expression levels of the examined ANXA2 (primer forward: 5'-AGAGTTTCCCGCTTGGTTGA-3', primer reverse: 5'-CCATATGCACTTGGGGGTGT-3') transcripts were compared to that of GAPDH (primer forward: 5'-AATGGGCAGCCGTTAGGAAA-3', primer reverse: 5'-GCGCCAATACGACCAAATC-3') and normalized to the mean value of the controls.

Statistical analysis

All statistical analyses and graphics in the study were conducted in R software (version 4.2.2, <https://www.r-project.org/>). Variables between two groups and three

groups were compared by Wilcoxon or Kruskal-Wallis test. Univariate and multivariate Cox regression analyses were used to investigate the prognostic value of the CD4⁺ T cell-related signature. Statistical differences in survival curves were determined by a log-rank test. $P < 0.05$ was considered statistically significant.

Result

Immunogenomic characterization of public CRC datasets reveals critical roles of activated memory CD4⁺ T cells in CRC prognosis

To assess the role of immune cells in the prognosis of CRC patients, we first quantitatively estimated the composition of 22 immune infiltrating cells in 4 publicly available CRC cohorts and evaluated the prognostic effect of each immune cell type on CRC at different clinical stages. Forest maps showed that infiltration of memory B cells (HR = 26.463, 95% CI, 1.388–504.577, $P = 0.029$) and M2 macrophages (HR = 10.013, 95% CI, 1.223–81.977, $P = 0.032$) was significantly associated with poor prognosis of stage III-IV CRC patients ($n = 617$), while infiltration of activated memory CD4⁺ T cells (HR = 0.010, 95% CI, 0.001–0.256, $P = 0.005$) was associated with better prognosis of stage III-IV CRC patients (Fig. 1A). However, the infiltration of these immune cells was not significantly associated with the prognosis of stage I-II CRC patients ($n = 697$, Figure S1). Survival analysis was used to further compare the relationship between activated memory CD4⁺ T cells, M2 macrophages, and memory B cells and the prognosis of stage III-IV CRC patients. As expected, patients with high infiltration of activated memory CD4⁺ T cells had a significantly longer OS than those with low infiltration ($P = 0.012$, Fig. 1B), whereas patients with high infiltration of memory B cells had a shorter OS than those with low infiltration ($P = 0.043$, Fig. 1C). There was no significant difference in OS between patients with high and low infiltrating M2 macrophages ($P = 0.079$, Fig. 1D). These results suggest that activated memory CD4⁺ T cells and memory B cells can effectively predict the prognosis of stage III-IV CRC patients.

Acquisition of CD4⁺ T cell marker genes using scRNA-seq data

After data processing and screening, a total of 6 patients with stage III-IV CRC in GSE166555 and a total of 2 patients with stage III-IV CRC in GSE144735 were included in subsequent analysis. Vlnplots showed the range of detected gene numbers (nFeature), the sequence count per cell (nCount) and the percentage of mitochondrial content (percent.mt) in each sample of GSE166555 and GSE144735 (Figure S2A and S2B). Correlation analysis results showed that nCount was positively correlated with nFeature in GSE166555 and GSE144735

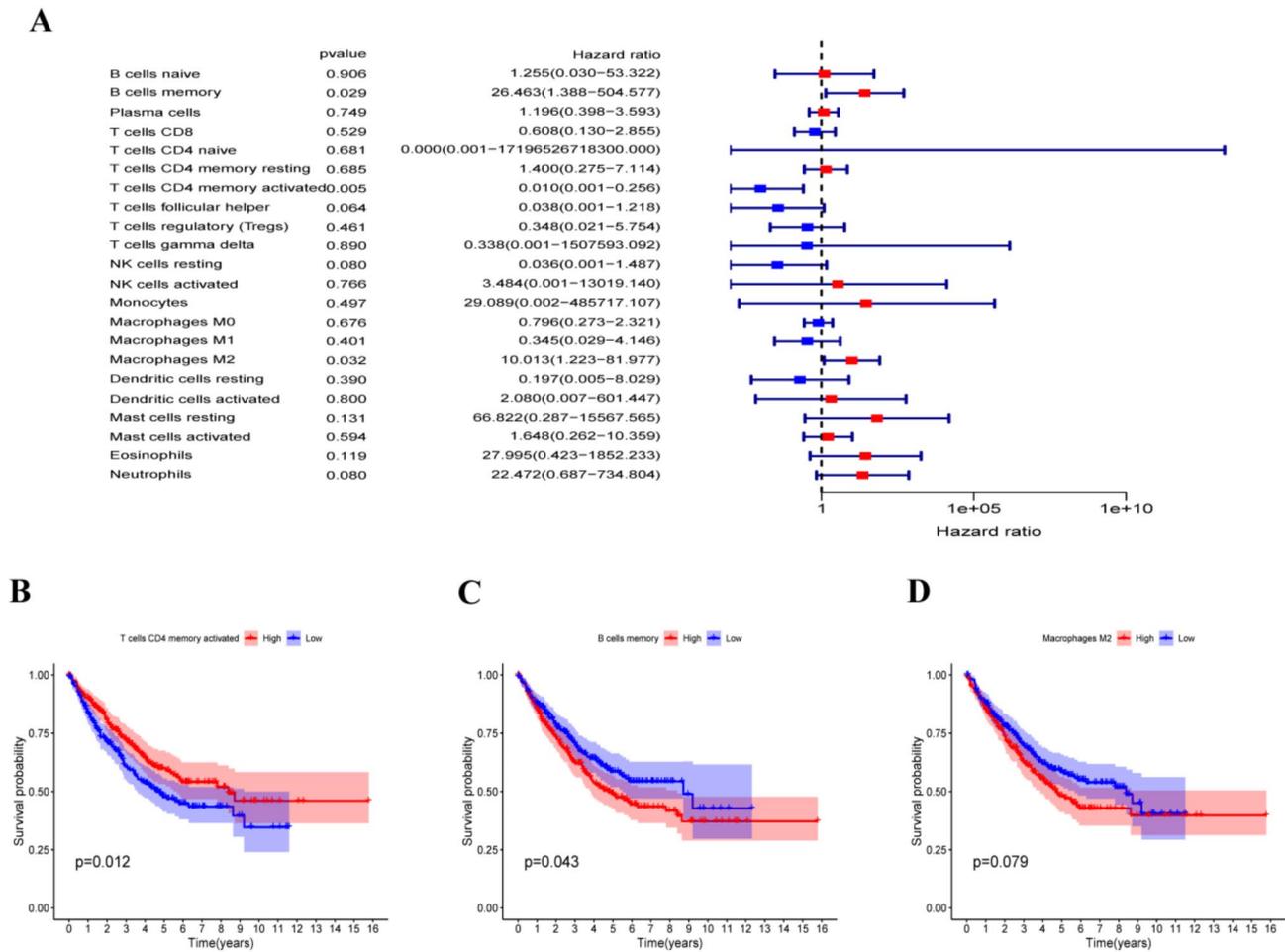


Fig. 1 The role of immunogenomic characterization in the prognosis of stage III-IV CRC. **(A)** The forest plot showed the effect of each immune cell type on the overall survival (OS) in stage III-IV CRC patients. **(B-D)** The Kaplan-Meier (K-M) curves of stage III-IV CRC patients and activated memory CD4⁺ T cells, M2 macrophages, and memory B cells

(Figure S2C and S2D). Then, we selected the top 1500 highly variable genes and plotted them in a scatter plot (Figure S2E and S2F). The PCA method was performed to reduce the dimensionality (Fig. 2A and B), and 15 PCs with P -value < 0.05 were selected for further analysis in both GSE166555 and GSE144735 (Figure S2G and S2H). We identified 16 clusters in GSE166555 and 13 clusters in GSE144735, and visualized these clusters using the t-SNE algorithm (Fig. 2C and D). The reference dataset from the MonacoImmuneData was used to annotate each cluster, and we found that cells from cluster 0 in GSE166555 were classified as CD4⁺ T cells (Fig. 2E), while cells from clusters 0 and 4 in GSE144735 were classified as CD4⁺ T cells (Fig. 2F). The clusters identified as CD4⁺ T cells were also found to have different gene expression profiles compared to the other clusters, with 246 and 184 genes differentially expressed in GSE166555 and GSE144735, respectively (Figure S3A and Figure S3B), which were identified as stage III-IV CRC-related CD4⁺ T cell marker genes (Table S2 and Table S3).

Identification of CD4⁺ T cell-related subtypes of stage III-IV CRC

A total of 54 CD4⁺ T cell marker genes of stage III-IV CRC were identified by Venn diagram (Fig. 3A, Table S4). To understand the biological function and prognostic value of these CD4⁺ T cell marker genes, we integrated the expression matrix and clinical data of 4 bulk datasets (TCGA-CRC, GSE17538, GSE12945, GSE39582), and finally included a total of 48 CD4⁺ T cell marker genes for follow-up analysis. As expected, the results of GO analysis showed that these CD4⁺ T cell marker genes were closely related to immune and inflammatory response, such as “positive regulation of leukocyte activation”, “T cell receptor binding” and “MHC protein complex” (Fig. 3B). Similarly, KEGG pathway analysis showed that these CD4⁺ T cell marker genes were mainly enriched in immune-related pathways like “Hematopoietic cell lineage”, “Th17 cell differentiation” and “Th1 and Th2 cell differentiation” (Fig. 3C). The network plot presented a comprehensive landscape of the relationship

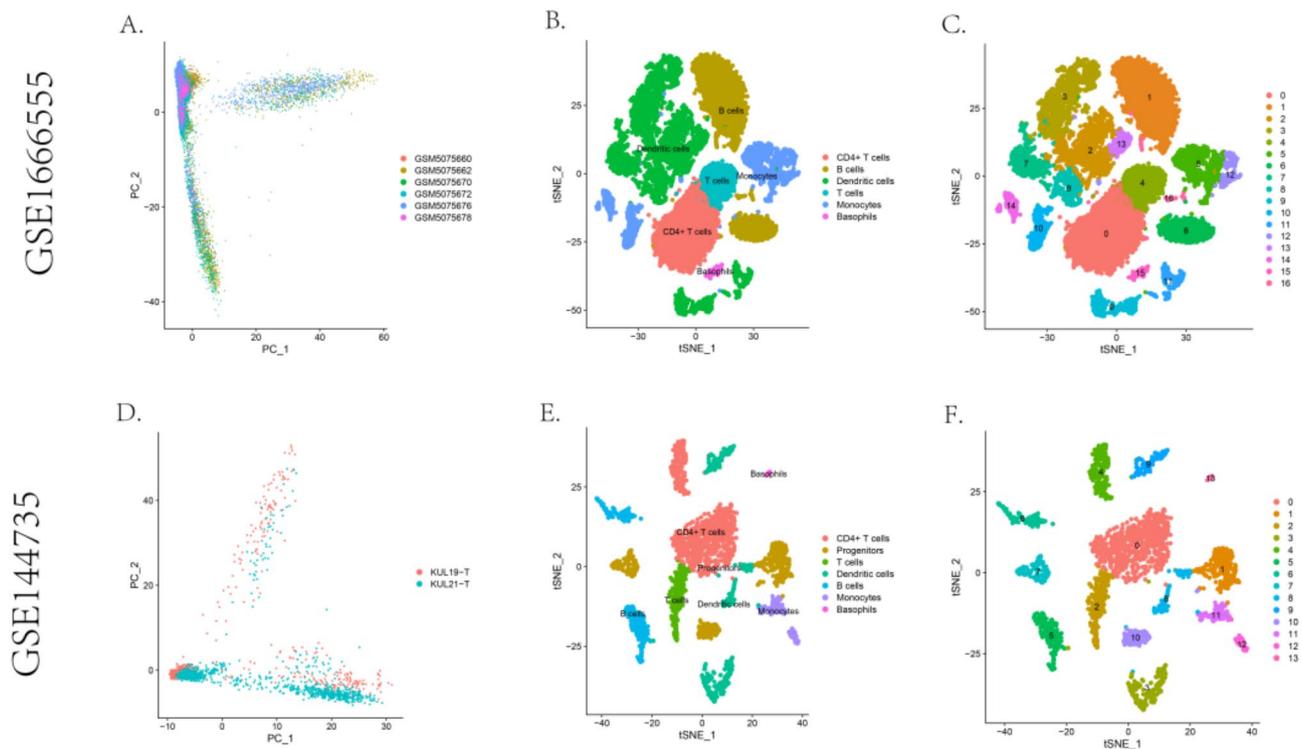


Fig. 2 Acquisition of CD4⁺ T cell marker genes using scRNA-seq data. (A–B) PCA was used for dimensionality reduction in GSE166655 and GSE144735. (C–D) 16 clusters in GSE166655 and 13 clusters in GSE144735 were visualized based on t-SNE algorithm. (E–F) Cell subpopulations identified by marker genes

between CD4⁺ T cell marker genes and the prognostic value of stage III–IV CRC (Figure S4A). Univariate Cox regression analysis revealed 20 of the 48 CD4⁺ T cell marker genes were significantly associated with survival ($P < 0.05$, Figure S4B). To further evaluate the role of CD4⁺ T cell marker genes in predicting the prognosis of stage III–IV CRC, we used 20 prognostic CD4⁺ T cell marker genes for unsupervised clustering on stage III–IV CRC patients. As shown in Figure S4, when $k = 2$ or $k = 3$, the cohort could be well divided into different subtypes. Overall survival analysis showed significant survival differences between the two subtypes ($P = 0.043$, Figure S4E) and no significant survival differences among 3 subtypes. We observed significant differences in TME between the two subtypes, and cluster A was remarkably abundant in most immune cell infiltrates, such as activated B cell, activated CD4 T cell, activated CD8 T cell, activated dendritic cell, CD56bright natural killer cell, and others (Fig. 3F). In addition, we calculated the TME score for each stage III–IV CRC patient using the ESTIMATE algorithm, and compared immune scores, stroma scores, ESTIMATE scores, and tumor purity between the two subtypes. According to the findings, cluster B patients had higher immune scores (Fig. 3G), stroma scores (Fig. 3H), and ESTIMATE scores (Fig. 3I) than those in cluster A patients, while tumor purity was lower than those in cluster A patients (Fig. 3J). In conclusion, cluster

B patients had higher levels of immune cell infiltration and TME scores, which explained the better prognosis to some extent.

Construction and evaluation of the prognostic signature based on CD4⁺ T cell marker genes

To further evaluate the reliability of CD4⁺ T cell markers for predicting prognosis in patients with III–IV CRC, we established a CD4⁺ T cell-related prognostic signature. The LASSO Cox regression was performed to build a CD4⁺ T cell-related prognostic signature in the training set based on 20 prognostic CD4⁺ T cell marker genes (Figure S5A and S5B). Finally, we used multivariate Cox regression analysis to optimize the prognostic signature, including only the six most predictive genes. A CD4⁺ T cell marker gene risk score was built: Risk score = $(-0.216 \times \text{CD2 expression}) + (-0.429 \times \text{CD9 expression}) + (0.447 \times \text{ANXA2 expression}) + (0.388 \times \text{RORA expression}) + (0.472 \times \text{RGS1 expression}) + (-0.349 \times \text{HLA-DRA expression})$. We used the above formula to calculate the risk score for each patient and divided patients into high-risk group ($n = 154$) and low-risk groups ($n = 154$) based on the median cutoff point. Kaplan–Meier (K–M) curves showed that the overall survival of high-risk patients was significantly worse than that of low-risk patients in the training set (Fig. 4A), testing set (Fig. 4B), and total set (Fig. 4C). From the distribution of risk scores and

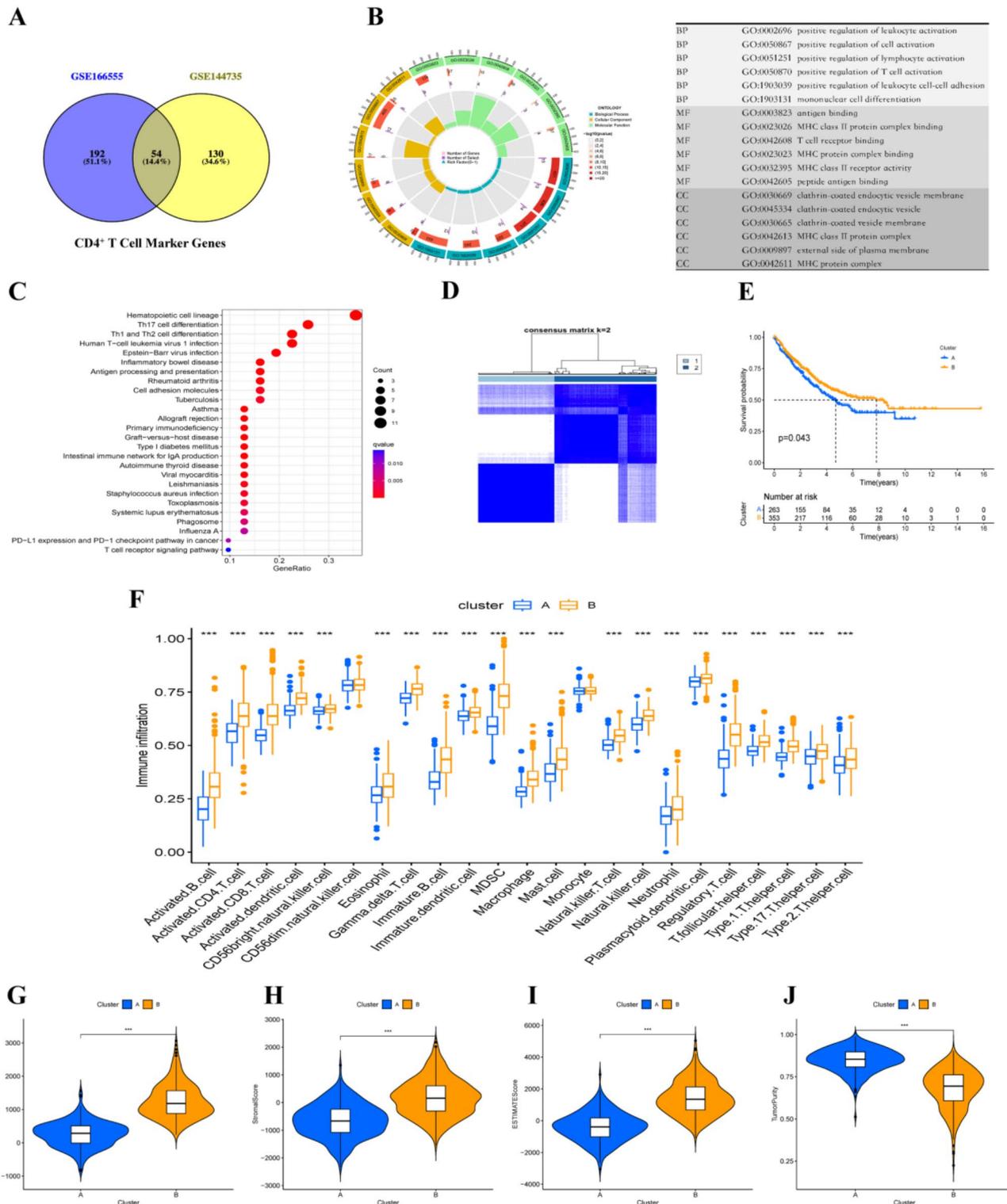


Fig. 3 Identification of CD4⁺ T cell-related subtypes of stage III-IV CRC. **(A)** Venn diagram shows the intersection of 54 CD4⁺ T cell marker genes in the GSE166555 and GSE144735. **(B)** Gene Ontology (GO) biopathway of the CD4⁺ T cell marker genes. **(C)** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of the CD4⁺ T cell marker genes. **(D)** The stage III-IV CRC patients were divided into two clusters by applying a consensus clustering algorithm. **(E)** Overall survival analyses for the two clusters ($p=0.043$). **(F)** The abundance of immune cells in the two clusters. **(G-J)** Differences in immune scores, stroma scores, ESTIMATE scores, and tumor purity of the two clusters

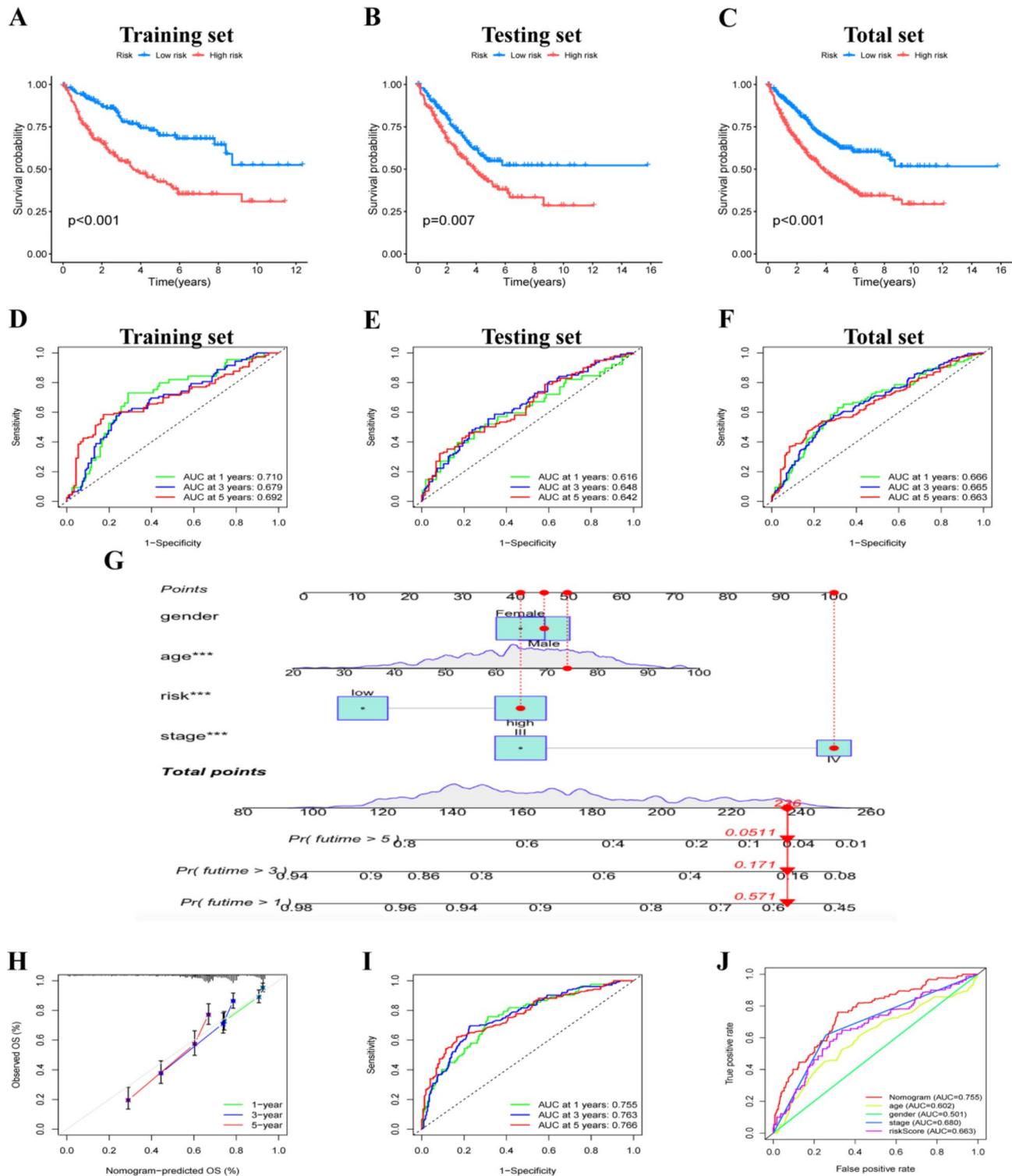


Fig. 4 Construction and evaluation of the prognostic signature based on CD4⁺ T cell marker genes. **(A-C)** The Kaplan-Meier (K-M) curves showed the different survival outcomes in the train, test, and total set. **(D-F)** Time-dependent receiver operating characteristics (ROC) curves for OS at 1-, 3-, and 5-years of the three sets. **(G)** The nomogram was used to predict the prognosis of patients based on their clinical information and risk score. **(H-I)** Calibration plot for the validation of the nomogram, with an AUC of 0.755, 0.763, and 0.766 at 1-, 3-, and 5-years, respectively. **(J)** Nomogram was a better predictor of 1-year OS than other clinical parameters

survival status of each patient, we found that higher risk scores were associated with more deaths in the training set (Figure S5C and S5F), testing set (Figure S4D and S4G), and total set (Figure S5E and S5H). The heat map showed detailed expression levels of the six genes (Figure S5I-K). Time-dependent ROC curves were utilized to evaluate the predictive accuracy of the prognostic signature. As shown in Fig. 4D, the 1-, 3-, and 5-year AUC values of stage III-IV CRC patients in the training set were 0.710, 0.679, and 0.692, respectively. Moreover, the 1, 3, and 5-year AUC values in the testing set and the total set also confirmed that this CD4⁺ T cell-related prognostic signature could well predict the prognosis of stage III-IV CRC patients (Fig. 4E-F). For better clinical application, we constructed a nomogram based on the risk score of the prognostic signature and other clinicopathological factors to more comprehensively predict patient survival

(Fig. 4G). The calibration curve and ROC curves showed that the nomogram had strong predictive performance, with an AUC of 0.755, 0.763, and 0.766 at 1, 3, and 5 years, respectively (Fig. 4H-I). Furthermore, the multiple-ROC curve illustrated that the nomogram predicted OS with better accuracy than other clinicopathological factors at 1 year (Fig. 4).

Relationship between the prognostic signature and clinicopathological factors

To understand the relationship between the CD4⁺ T cell-related signature and clinicopathological factors, we showed the relationship between CD4⁺ T cell-related subtypes, the prognostic signature, and patient survival status. We found that cluster B patients were more concentrated in the low-risk group with fewer deaths (Fig. 5A). Moreover, we found that risk scores of stage

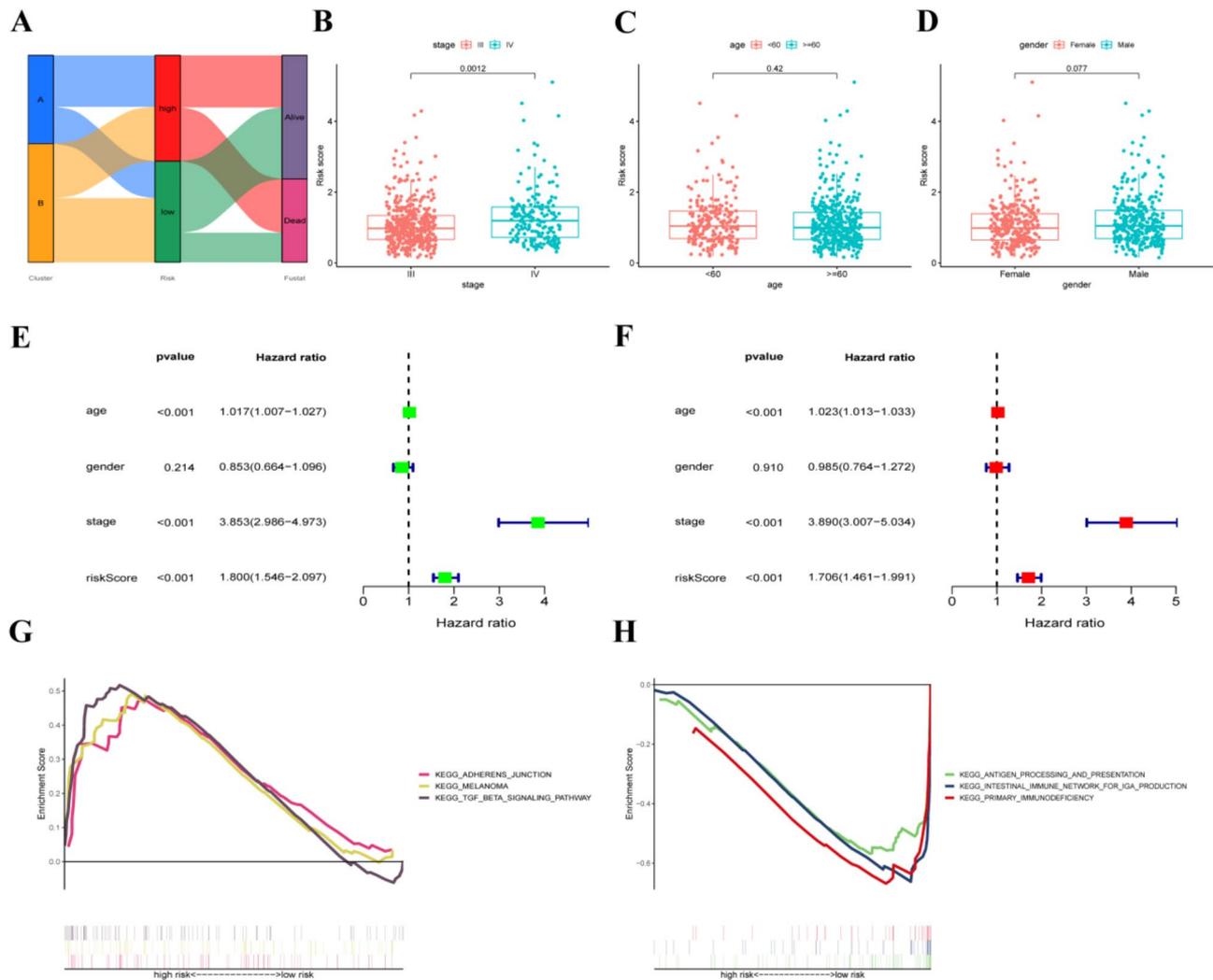


Fig. 5 Relationship between the prognostic signature and clinicopathological factors. **(A)** Sankey diagram of subtype and survival status. **(B-D)** The relationship between the risk scores and clinical stage, age, and gender. **(E-F)** Univariate and multivariate Cox regression analysis. **(G-H)** Gene Set Enrichment Analysis (GSEA) between high- and low-risk groups

IV patients were significantly higher than those of stage III patients ($P=0.0012$, Fig. 5B), while the risk score had no significant relationship with age ($P=0.42$, Fig. 5C) and gender ($P=0.077$, Fig. 5D). The results of univariate and multivariate Cox regression analysis suggested that the risk score was an independent prognostic factor for OS of stage III-IV CRC patients (Fig. 5E and F).

We further explored differences in biological processes and pathways between high- and low-risk groups. The results of the GSEA analysis showed that the enriched pathways in the high-risk group were mainly concentrated in adherens junction, melanoma and TGF-beta signaling pathway (Fig. 5G), and the enriched pathways in the low-risk group were focused on antigen processing and presentation, intestinal immune network for IGA production and primary immunodeficiency (Fig. 5H).

Association between the prognostic signature and immune cell infiltration

Since CD4⁺ T cells play a crucial role in anti-tumor immune response, we investigated the relationship between the CD4⁺ T cell-related prognostic signature and immune cell infiltration in stage III-IV CRC patients.

As shown in Fig. 6A, the six CD4⁺ T cell marker genes constructed the prognostic signature were significantly correlated with the infiltration levels of most immune cells. The EPIC analysis also showed significantly correlation between immune infiltration and 6 maker genes (Fig. 6S). Moreover, we found that memory B cells, plasma cells, resting memory CD4 T cells, and neutrophils were significantly upregulated in the high-risk group, and naive B cells, CD8 T cells, activated memory CD4 T cells, follicular helper T cells, resting NK cells, and resting dendritic cells were significantly upregulated in the low-risk group (Fig. 6B). The results of ssGSEA analysis showed that the enrichment of immune-related functions was significantly higher in the low-risk group than in the high-risk group, such as APC co-stimulation, CCR, Checkpoint, Cytolytic activity, HLA, inflammation promoting, MHC class I, T cell co-inhibition, T cell co-stimulation, aDCs, B cells, CD8⁺ T cells, DCs, iDCs, pDCs, T helper cells, Tfh, Th1 cells, TIL (Fig. 6C and D). In addition, by using the ESTIMATE algorithm, we found that the immune scores of high-risk patients were lower than those of low-risk patients, while stromal scores were higher than those of low-risk patients (Fig. 6E). These

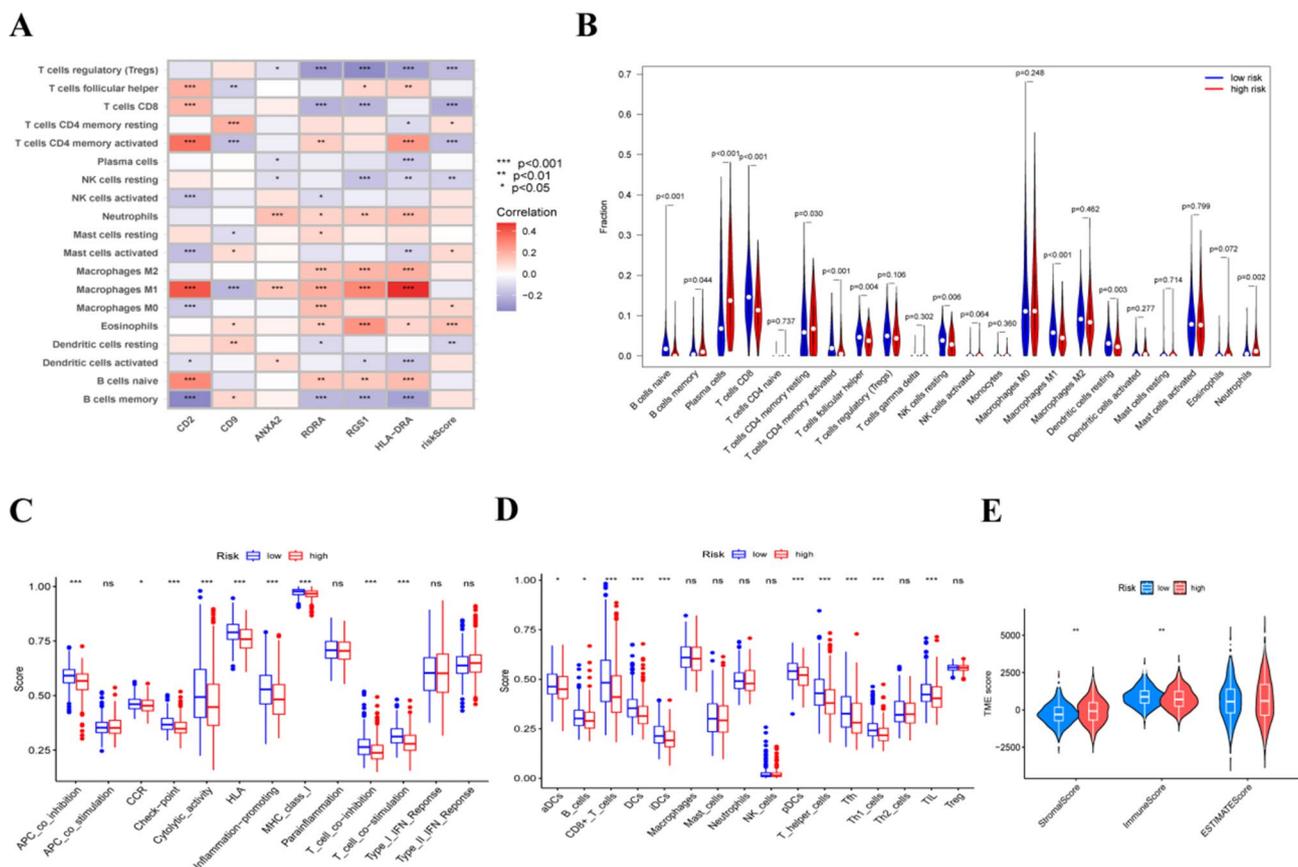


Fig. 6 Association between the prognostic signature and immune cell infiltration. (A) The correlation of 6 key genes expression and immune cell infiltration. (B) CIBERSORT analysis showed the correlation between 6 key genes and immune infiltration. (C-D) The abundance of immune-related functions in the high- and low- risk groups. (E) Comparison of TME score in the two risk groups

results suggested that the risk score was negatively correlated with the level of immune cell infiltration.

Relationship between the CD4⁺ T cell-related signature with immunotherapy response and drug sensitivity

Given the important role of CD4⁺ T cell-based immunotherapeutic strategies in anti-tumor treatment, we explored whether the CD4⁺ T cell-related signature could predict responses of stage III-IV CRC patients to chemotherapy and immunotherapy. We first used the GDSC [22] database to predict the chemotherapy drug sensitivity of CRC patients in high-risk and low-risk groups. The results showed that patients in high-risk group were more sensitive to Oxaliplatin and 5-Fluorouracil than those in the low-risk group, but there were no significant differences in Irinotecan sensitivity between the two groups (Fig. 7A-C). Next, we assessed the prognostic role of risk score for chemotherapy and chemotherapy-combined bevacizumab-immunotherapy in GSE72970. As shown in Fig. 7D and E, the low-risk group had higher proportion of Response patients (R) and lower proportion of Non-Response patients (NR) than the high-risk group, indicating that the low-risk group was more sensitive to chemotherapy and chemotherapy + bevacizumab than the high-risk group. We also found significant differences in OS and progression-free survival (PFS) between these two groups. (Fig. 7F-I) These results suggest that the CD4⁺ T-cell-related signature is a good predictor of chemotherapy and immunotherapy response in CRC patients.

ANXA2 is associated with the infiltration of Tregs in stage III-IV CRC

Among all CD4⁺T cell genes in the risk score cohort, we found that ANXA2 is significantly positive correlated with poor prognosis in stage III-IV CRC patients. Subsequently, we decided to explore the relationship between ANXA2 and the infiltration of CD4⁺ T cell subsets. After primarily classification of GEO dataset GSE166555 (Fig. 8A), we further conducted subclustering analyses on CD4⁺ T cell clusters. We found that ANXA2 was enriched in Treg subsets (Fig. 8B) and the proportion of Tregs significantly increased in ANXA2^{high} group, compared to ANXA2^{low} group (Fig. 8C). Immunofluorescence staining of FOXP3 and ANXA2 on 5 pairs of CRC tumor and adjacent normal tissues showed that the colocalization of FOXP3 and ANXA2 in CRC tumor tissues was increased compared with that in adjacent normal tissues (Fig. 8D and E and Figure S7). Additionally, RT-PCR results showed that ANXA2 mRNA expression in colorectal cancer tissues was higher than that in normal tissues (Fig. 8F). These results indicated that the expression of ANXA2 increased in CRC tissue, and it's associated with the infiltration of CRC Tregs.

Discussion

Immunotherapy has achieved remarkable results in anti-tumor treatment. More and more studies have focused on biomarkers that can predict immunotherapy responses to improve the prognosis of patients. TME has been widely recognized as an important factor affecting the efficacy of cancer immunotherapy, and TME-related biomarkers have also received increasing attention. However, TME-related biomarkers based on tumorigenesis to predict immunotherapy response and prognosis in patients with advanced CRC are still lacking. The scRNA-seq technique is a powerful tool for exploring tumor heterogeneity and tumor immune microenvironment, which is important for identifying potential therapeutic targets [23]. The abundance of T cells especially CD4⁺ T cell infiltration in tumors is closely related to the prognosis of patients with various solid tumors [24–26]. Recently, the signature of CD4⁺ T cell-related lncRNA constructed by Ning et al. has been proven to have potential predictive value for the prognosis and immunotherapy response of breast cancer [27]. Based on the important role of CD4⁺ T cells in the prognosis and immunotherapy of cancer patients, in our study, we tried to explore the relationship between CD4⁺ T cell marker genes and the prognosis of stage III-IV CRC patients.

We systematically analyzed the relationship between CD4⁺ T cells and the prognosis of patients with stage III-IV CRC patients by combining bulk RNA-seq and clinical information data. As expected, CD4⁺ T cells are associated with a better prognosis of patients with stage III-IV CRC patients, but not with the prognosis of patients with stage I-II CRC patients. This suggests that CD4⁺ T cells may be a key biomarker for the prognosis of advanced CRC. Subsequently, we used scRNA-seq analysis and found out CD4⁺ T cell marker genes in stage III-IV CRC patients. Based on the CD4⁺ T cell marker genes, we used unsupervised clustering to divide patients into two subsets. It was found that there were significant differences in overall survival and immune cell infiltration between the two subsets. Patients in cluster B have a better prognosis and more immune cell infiltration compared with patients in cluster A. Consistently, immune scores and stromal scores of cluster B patients were also higher than those of cluster A patients. A large number of studies have shown that in patients treated with checkpoint inhibitors, high infiltration of immune cells is related to better tumor response and long-term survival [28, 29]. In addition, it has been reported that a TME of CRC with a high immune score is associated with better survival [30]. Therefore, the better prognosis of cluster B patients might be related to the high infiltration of immune cells.

We further established a CD4⁺ T cell-related prognostic signature for stage III-IV CRC patients using the training set. The prognostic signature consisted of six

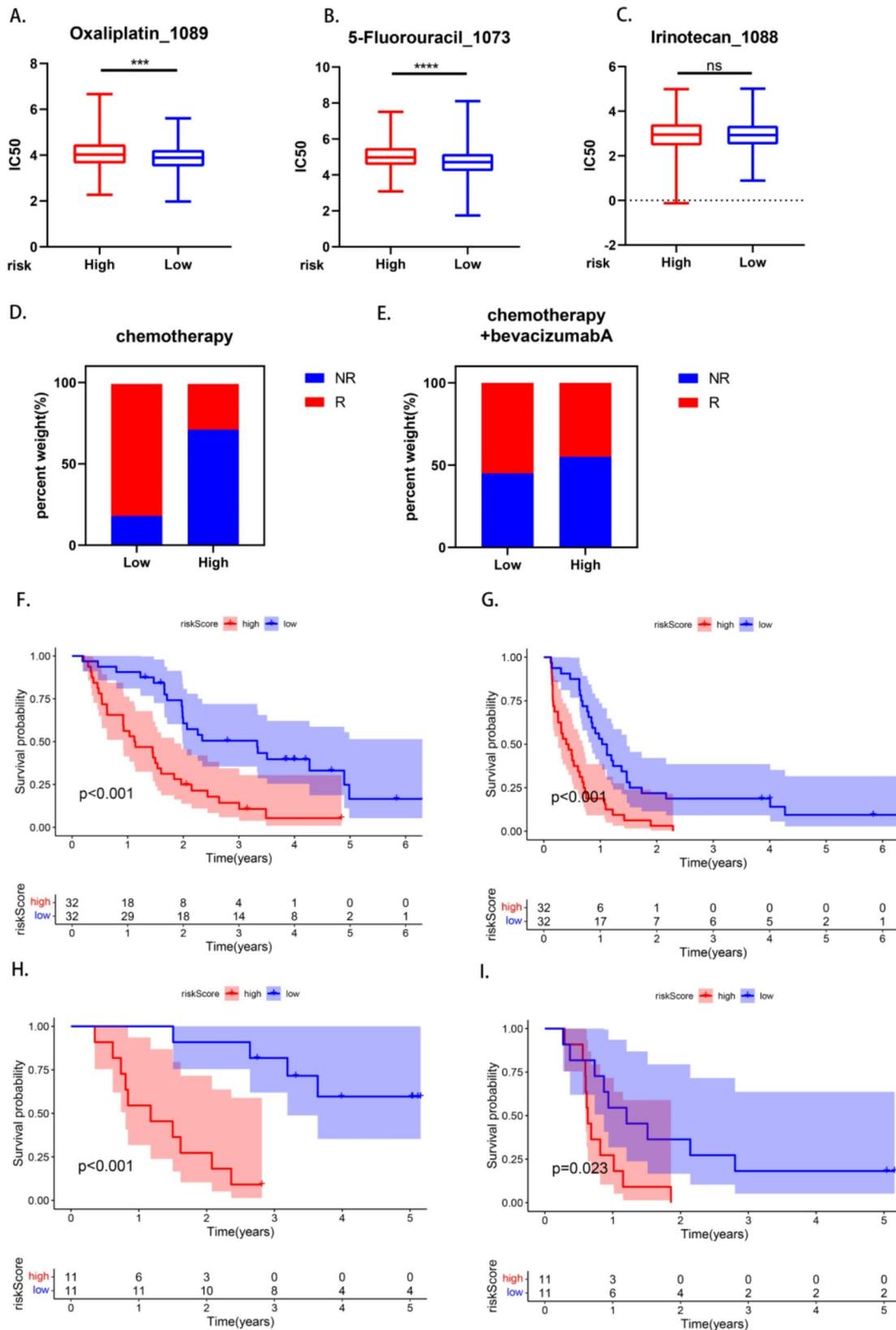


Fig. 7 Evaluation of predicting immunotherapy response and drug sensitivity. Estimated IC50 of Oxaliplatin_1089 (A), 5-fluorouracil_1073 (B) and Irinotecan_1088 (C) based on GDSC database in high-risk and low-risk groups. (n=616) Response of chemotherapy (D) (n=64) and chemotherapy + bevacizumab (E) (n=22) in high-risk and low-risk groups. Overall survival analysis (F) and progression free survival analysis (G) of chemotherapy between two risk groups. (n=64) Overall survival analysis (H) and progression free survival analysis (I) of chemotherapy + bevacizumab between two risk groups. (n=22)

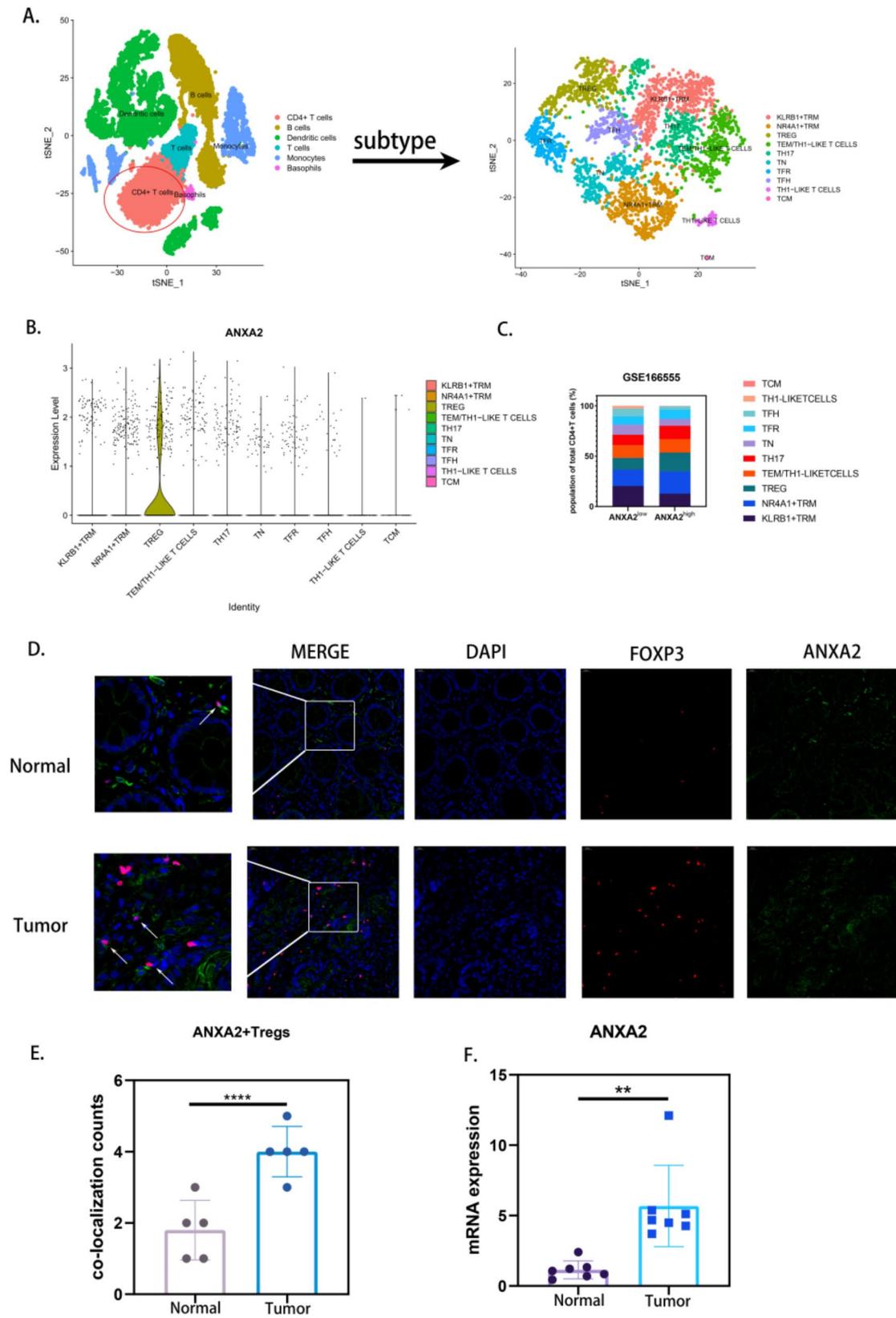


Fig. 8 (See legend on next page.)

(See figure on previous page.)

Fig. 8 ANXA2 is associated with infiltration of Tregs in Stage III-IV CRC. **(A)** GSE166555 CD4⁺ T cells subset was subclustered into 10 clusters which were visualized based on t-SNE algorithm. **(B)** Violin plot of ANXA2 distribution in 10 subcluster of GSE166555 CD4⁺ T cells subset. **(C)** Proportion of CD4⁺ T cell subclusters in ANXA2^{high} group and ANXA2^{low} group. **(D)** Representative image of immunofluorescence staining of ANXA2 and FOXP3 on paraffin sections of normal and CRC tissue. (Scale bar: 50 μm) **(E)** Data analysis of the number of cells colocalized with ANXA2 and FOXP3. (n = 5) **(F)** Expression of ANXA2 mRNA in CRC and normal tissue

CD4⁺ T cell marker genes, including CD2, CD9, ANXA2, RORA, RGS1, and HLA-DRA. The predictive power of the CD4⁺ T cell-related signature was verified in both the testing set and the total set. The prognostic signature showed good robustness and repeatability in both cohorts. Moreover, we established a nomogram by combining clinicopathological factors and risk scores, and verified its predictive performance through a variety of validation methods. This nomogram could accurately predict the prognosis of stage III-IV CRC patients and guide the establishment of individualized examination procedures. Based on the excellent predictive power of this signature, we performed a GSEA analysis to explore the underlying mechanism. The low-risk group was more concentrated on pathways related to immune response, while the high-risk group mainly focused on pathways closely related to tumor progression, such as adhesion junction and TGF-beta signaling pathway. The alteration of adhesion junction is usually associated with metastasis of colorectal cancer cells [31], Disruption of the TGF-β signaling pathway in advanced CRC has been reported to lead to an aggressive phenotype of CRC and a poor prognosis [32]. The inferior prognosis of the high-risk patients may be partly attributed to the significant enrichment of these pathways.

Due to the key role of TME in anti-tumor response, the differences in TME between the high-risk and low-risk groups were compared. We observed that naïve B cells, CD8 T cells, activated memory CD4 T cells, follicular helper T cells, resting NK cells, and resting dendritic cells were significantly upregulated in the low-risk group. Numerous studies have shown that these cells play an antitumor role in various solid tumors [33–36]. Moreover, the enrichment of immune-related functions and immune scores in the low-risk group were significantly higher than those in the high-risk group. A high level of immune cell infiltration and strong immune response usually indicates a strong anti-tumor response [37, 38]. The low level of immune cell infiltration can promote tumor cells to evade immune monitoring and thus promote tumor progression, which may partly explain the significantly decreased survival rate of high-risk patients. In addition, to evaluate the impact of risk score on chemotherapy and immunotherapy, we compared the differences between the high-risk and low-risk groups in the We have analyzed the sensitivity of high-risk and low-risk patients to common chemotherapy drugs for CRC patients and have shown that the low-risk group

patients is more sensitive to Oxaliplatin, 5-Fluorouracil, and Irinotecan than the high-risk group patients. What's more, we obtained the expression matrix of 84 stage III-IVCRC patients receiving chemotherapy or chemotherapy + bevacizumab-immunotherapy from the GEO data set GSE72970 and analyzed responses to these therapies between the high-risk and low-risk groups. As expected, patients in the low-risk group were more sensitive to chemotherapy or chemotherapy + bevacizumab-immunotherapy. We also found significant differences in overall survival (OS) and progression-free survival (PFS) between these two groups. These results suggest that the CD4⁺ T-cell-related signature can potentially be applied to predict the sensitivity of CRC patients to chemotherapy and immunotherapy drugs. Nevertheless, there are certain limitations to this study. Due to the lack of relevant data, we were unable to conduct the colorectal cancer cohort analysis of multiple immunotherapy combinations or immune check-point PD-L1 as was done in the melanoma [39] and lung adenocarcinoma [40] studies. The six CD4⁺ T cell marker genes used to construct the prognostic signature are closely related to the development and progression of CRC. It has been reported that the highest CD9-CD63 was observed in CRC patients, which is an independent prognostic factor for progression-free survival and overall survival of CRC patients [41]. Increased ANXA2 levels have been found in stage IV tumors and metastasis compared to stage I-III tumors, and silicon analysis confirmed that ANXA2 overexpression was associated with consensus molecular subtypes (CMS) with the worst prognosis [42]. Previous studies have shown that the methylation status of the RORA promoter region was significantly correlated with unfavorable CRC stages (stage III and stage IV), which may be a useful biomarker for highly advanced CRC patients [43]. Lange et al. demonstrated that RGS1 is a potential marker of CRC tissue quality through microarray and qPCR analysis [44]. HLA-DRA expression was increased in the right side CRC and was associated with patient prognosis [45]. These reports indicated that these CD4⁺ T cell marker genes might be potential therapeutic targets for stage III-IV CRC patients and that the novel prognostic signature constructed by them might be useful for clinical applications. Additionally, by subpopulation analysis of CD4⁺ T cells, we found that ANXA2 expression was mainly distributed on Treg cells. Finally, we confirmed that ANXA2 expression was significantly up-regulated in CRC tissues, and ANXA2 co-localization

with Foxp3 (a Treg cell specific marker) was significantly increased in CRC compared with healthy controls. As we all know, Treg cells are a subpopulation of CD4⁺ T cells, which play a catalytic role in tumor development. Numerous studies have shown that an increased number of Treg cells in tumors is associated with tumor development, immunotherapy failure, and poor prognosis in CRC [46, 47]. Reducing or depleting Tregs in CRC has been shown to inhibit tumor growth, suggesting that targeting Tregs would be an effective anti-tumor approach for CRC [48]. At present, the relationship between ANXA2 and Treg cells has not been reported. ANXA2 has been demonstrated to be a key regulator of CRC progression [42]. While, the relationship between ANXA2 and Treg cells has not been reported. We have shown for the first time that ANXA2 can be used as a Treg-associated marker gene to predict the prognosis of advanced CRC patients and may be a potential target for immunotherapy in tumors.

Although this study provides new insights for predicting the prognosis and treatment of CRC patients, it still has some limitations. First, this is a retrospective study that needs further validation in a prospective cohort study. Second, due to the small number of samples in the current scRNA dataset and the lack of clinical survival information, we didn't validate TME differences with single-cell data further. Third, future research must explore the potential mechanism between the expression of CD4⁺ T cell marker genes and the prognosis of stage III-IV CRC patients. Fourthly, when discussing the relationship between risk score and chemotherapy drug sensitivity, there was a lack of relevant experiments to further verify the conclusion. Finally, lacking of single cell RNA sequencing data from stage III-IV CRC patients leads to analyze TME differences among different risk groups.

In conclusion, based on the comprehensive analysis of scRNA and bulk RNA-sequencing, we found that CD4⁺ T cells had a significant relationship with the prognosis of stage III-IV CRC patients, and then developed a novel prognostic signature composed of six CD4⁺ T cell marker genes, which could effectively predict the prognosis and the response to immunotherapy in stage III-IV CRC patients. The CD4⁺ T cell-related signature could serve as a prognostic biomarker for stage III-IV CRC patients and provide new insights for clinical decision-making of individualized immunotherapy.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12876-025-03716-2>.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We acknowledge the free use of TCGA and GEO databases.

Author contributions

Mengting Li and Qiu Zhao conceived and designed the study. Yuanyuan Lu and Yu Shao collected the data, Mengting Li and Weining Zhu analyzed the data. Weining Zhu supervised the data and provided statistical advice. Mengting Li, Weining Zhu and Yuanyuan Lu wrote the paper, Fei Xu and Lan Liu reviewed the paper. All authors have read and approved the final manuscript.

Funding

This study was supported by the Hubei Provincial Natural Science Foundation (No. 2023AFB203), Fundamental Research Funds for the Central Universities (No. 2042023kf0086) and Scientific Research Project of Wuhan Science and Technology Bureau (No.2024040801020371).

Data availability

The raw data of this study are derived from GEO (<https://www.ncbi.nlm.nih.gov/>) and the TCGA database (<https://portal.gdc.cancer.gov/>). All the databases above are open access and can be login in directly without relevant accession numbers.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 15 September 2024 / Accepted: 19 February 2025

Published online: 11 March 2025

References

- Eng C, Yoshino T, Ruiz-García E et al. Colorectal cancer. *Lancet*. 2024;404(10449):294–310. [https://doi.org/10.1016/S0140-6736\(24\)00360-X\(2024\)](https://doi.org/10.1016/S0140-6736(24)00360-X(2024)).
- Shinji S, et al. Recent advances in the treatment of colorectal cancer: A review. *J Nippon Med Sch*. 2022;89:246–54. https://doi.org/10.1272/jnms.JNMS.2022_89-310.
- Ahmed M, Colon Cancer. A clinician's perspective in 2019. *Gastroenterol Res*. 2020;13:1–10. <https://doi.org/10.14740/gr1239>.
- De Greef K, et al. Multidisciplinary management of patients with liver metastasis from colorectal cancer. *World J Gastroenterol*. 2016;22:7215–25. <https://doi.org/10.3748/wjg.v22.i32.7215>.
- Zheng H, et al. Weighted gene Co-expression network analysis identifies CALD1 as a biomarker related to M2 macrophages infiltration in stage III and IV mismatch Repair-Proficient colorectal carcinoma. *Front Mol Biosci*. 2021;8:649363. <https://doi.org/10.3389/fmolb.2021.649363>.
- Liu K, et al. Nomogram model characterized by mutant genes and clinical indexes to identify high-risk patients with stage III/IV colorectal cancer. *J Gastrointest Oncol*. 2020;11:1214–23. <https://doi.org/10.21037/jgo-20-548>.
- Zhong R, et al. Immune cell infiltration features and related marker genes in lung cancer based on single-cell RNA-seq. *Clin Transl Oncol*. 2021;23:405–17. <https://doi.org/10.1007/s12094-020-02435-2>.
- Liang L, et al. Integration of scRNA-Seq and bulk RNA-Seq to analyse the heterogeneity of ovarian Cancer immune cells and Establish a molecular risk model. *Front Oncol*. 2021;11:711020. <https://doi.org/10.3389/fonc.2021.711020>.
- Yu L, et al. Characterization of cancer-related fibroblasts (CAF) in hepatocellular carcinoma and construction of CAF-based risk signature based on single-cell RNA-seq and bulk RNA-seq data. *Front Immunol*. 2022;13:1009789. <https://doi.org/10.3389/fimmu.2022.1009789>.
- Song P, et al. Identification and validation of a novel signature based on NK cell marker genes to predict prognosis and immunotherapy response in

- lung adenocarcinoma by integrated analysis of Single-Cell and bulk RNA-Sequencing. *Front Immunol.* 2022;13:850745. <https://doi.org/10.3389/fimmu.2022.850745>.
11. Ruan Z, et al. Development and validation of a prognostic model and gene co-expression networks for breast carcinoma based on scRNA-seq and bulk-seq data. *Ann Transl Med.* 2022;10:1333. <https://doi.org/10.21037/atm-22-5684>.
 12. Hanahan D, Coussens LM. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell.* 2012;21:309–22. <https://doi.org/10.1016/j.ccr.2012.02.022>.
 13. Binnewies M, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med.* 2018;24:541–50. <https://doi.org/10.1038/s41591-018-0014-x>.
 14. AlMusawi S, Ahmed M, Nateri AS. Understanding cell-cell communication and signaling in the colorectal cancer microenvironment. *Clin Transl Med.* 2021;11:e308. <https://doi.org/10.1002/ctm2.308>.
 15. Liu YT, Sun ZJ. Turning cold tumors into hot tumors by improving T-cell infiltration. *Theranostics.* 2021;11:5365–86. <https://doi.org/10.7150/thno.58390>.
 16. Ahrends T, et al. CD4(+) T cell help confers a cytotoxic T cell effector program including coinhibitory receptor downregulation and increased tissue invasiveness. *Immunity.* 2017;47:848–e861845. <https://doi.org/10.1016/j.immuni.2017.10.009>.
 17. Miggelbrink AM, et al. CD4 T-Cell exhaustion: does it exist and what are its roles in cancer? *Clin Cancer Res.* 2021;27:5742–52. <https://doi.org/10.1158/1078-0432.Ccr-21-0206>.
 18. Perez-Diez A, et al. CD4 cells can be more efficient at tumor rejection than CD8 cells. *Blood.* 2007;109:5346–54. <https://doi.org/10.1182/blood-2006-10-51318>.
 19. Melenhorst JJ, et al. Decade-long leukaemia remissions with persistence of CD4(+) CAR T cells. *Nature.* 2022;602:503–9. <https://doi.org/10.1038/s41586-021-04390-6>.
 20. Toor SM, et al. Immune checkpoints in Circulating and Tumor-Infiltrating CD4(+) T cell subsets in colorectal Cancer patients. *Front Immunol.* 2019;10:2936. <https://doi.org/10.3389/fimmu.2019.02936>.
 21. Hao Y, et al. Integrated analysis of multimodal single-cell data. *Cell.* 2021;184:3573–e35873529. <https://doi.org/10.1016/j.cell.2021.04.048>.
 22. Yang W, Soares J, Greninger P, et al. Genomics of drug sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* 2013;41(Database issue):D955–61. <https://doi.org/10.1093/nar/gks1111>. (2013).
 23. Papalexli E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol.* 2018;18:35–45. <https://doi.org/10.1038/nri.2017.76>.
 24. Eberst G, et al. Prognostic value of CD4+ T lymphopenia in non-small cell lung Cancer. *BMC Cancer.* 2022;22. <https://doi.org/10.1186/s12885-022-09628-8>.
 25. Wakabayashi O, et al. CD4+ T cells in cancer stroma, not CD8+ T cells in cancer cell nests, are associated with favorable prognosis in human non-small cell lung cancers. *Cancer Sci.* 2003;94:1003–9. <https://doi.org/10.1111/j.1349-7006.2003.tb01392.x>.
 26. Ahlén Bergman E, et al. Increased CD4(+) T cell lineage commitment determined by CpG methylation correlates with better prognosis in urinary bladder cancer patients. *Clin Epigenetics.* 2018;10:102. <https://doi.org/10.1186/s13148-018-0536-6>.
 27. Ning S, et al. Identification of CD4(+) conventional T Cells-Related lncRNA signature to improve the prediction of prognosis and immunotherapy response in breast Cancer. *Front Immunol.* 2022;13:880769. <https://doi.org/10.3389/fimmu.2022.880769>.
 28. Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nat Rev Cancer.* 2019;19:133–50. <https://doi.org/10.1038/s41568-019-0116-x>.
 29. Mezheyeuski A, et al. The immune landscape of colorectal Cancer. *Cancers (Basel).* 2021;13. <https://doi.org/10.3390/cancers13215545>.
 30. Yuan W, Cai W, Huang X, Peng S. Prognostic value of immune scores in the microenvironment of colorectal cancer. *Oncol Lett.* 2020;20:256. <https://doi.org/10.3892/ol.2020.12119>.
 31. Zhu W, et al. Study on the action mechanism of the Yifei Jianpi Tongfu formula in treatment of colorectal Cancer lung metastasis based on network analysis, molecular docking, and experimental validation. *Evid Based Complement Alternat Med.* 2022;2022(6229444). <https://doi.org/10.1155/2022/6229444>.
 32. Itatani Y, Kawada K, Sakai Y. Transforming growth Factor- β signaling pathway in colorectal Cancer and its tumor microenvironment. *Int J Mol Sci.* 2019;20. <https://doi.org/10.3390/ijms20235822>.
 33. Makita Y, et al. Activated iNKT cells enhance the anti-tumor effect of antigen specific CD8 T cells on mesothelin-expressing salivary gland cancer. *BMC Cancer.* 2018;18:1254. <https://doi.org/10.1186/s12885-018-5179-7>.
 34. Fu C, Jiang A. Dendritic Cells. CD8 T cell immunity in tumor microenvironment. *Front Immunol.* 2018;9:3059. <https://doi.org/10.3389/fimmu.2018.03059>.
 35. Wang Y, et al. Dendritic cell biology and its role in tumor immunotherapy. *J Hematol Oncol.* 2020;13. <https://doi.org/10.1186/s13045-020-00939-6>.
 36. Li X, et al. Low-Dose decitabine augments the activation and Anti-Tumor immune response of IFN- γ (+) CD4(+) T cells through enhancing I κ Ba degradation and NF- κ B activation. *Front Cell Dev Biol.* 2021;9:647713. <https://doi.org/10.3389/fcell.2021.647713>.
 37. Lu C, Liu Y, Ali NM, Zhang B, Cui X. The role of innate immune cells in the tumor microenvironment and research progress in anti-tumor therapy. *Front Immunol.* 2022;13:1039260. <https://doi.org/10.3389/fimmu.2022.1039260>.
 38. Pansy K, et al. Immune regulatory processes of the tumor microenvironment under malignant conditions. *Int J Mol Sci.* 2021;22. <https://doi.org/10.3390/ijms222413311>.
 39. Larkin J, et al. Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *N Engl J Med.* 2015;373:23–34. <https://doi.org/10.1056/NEJMoa1504030>.
 40. Schoenfeld AJ, et al. Clinical and molecular correlates of PD-L1 expression in patients with lung adenocarcinomas. *Ann Oncol.* 2020;31:599–608. <https://doi.org/10.1016/j.annonc.2020.01.065>.
 41. Wei P, et al. Plasma extracellular vesicles detected by single molecule array technology as a liquid biopsy for colorectal cancer. *J Extracell Vesicles.* 2020;9:1809765. <https://doi.org/10.1080/20013078.2020.1809765>.
 42. Rocha MR, et al. Annexin A2 overexpression associates with colorectal cancer invasiveness and TGF- β induced epithelial mesenchymal transition via Src/ANXA2/STAT3. *Sci Rep.* 2018;8:11285. <https://doi.org/10.1038/s41598-018-29703-0>.
 43. Kano H, Takayama T, Midorikawa Y, Nagase H. Promoter hypomethylation of RAR-related orphan receptor α 1 is correlated with unfavorable clinicopathological features in patients with colorectal cancer. *Biosci Trends.* 2016;10:202–9. <https://doi.org/10.5582/bst.2016.01097>.
 44. Lange N, et al. Identification and validation of a potential marker of tissue quality using gene expression analysis of human colorectal tissue. *PLoS ONE.* 2015;10:e0133987. <https://doi.org/10.1371/journal.pone.0133987>.
 45. Hu Y, et al. Differential expression and prognostic correlation of immune related factors between right and left side colorectal Cancer. *Front Oncol.* 2022;12:845765. <https://doi.org/10.3389/fonc.2022.845765>.
 46. Aristin Revilla S, Kranenburg O, Coffey PJ. Colorectal Cancer-Infiltrating regulatory T cells: functional heterogeneity, metabolic adaptation, and therapeutic targeting. *Front Immunol.* 2022;13:903564. <https://doi.org/10.3389/fimmu.2022.903564>.
 47. van Herk EH, Te Velde AA. Treg subsets in inflammatory bowel disease and colorectal carcinoma: characteristics, role, and therapeutic targets. *J Gastroenterol Hepatol.* 2016;31:1393–404. <https://doi.org/10.1111/jgh.13342>.
 48. Wang R, et al. Inhibition of PCSK9 enhances the antitumor effect of PD-1 inhibitor in colorectal cancer by promoting the infiltration of CD8(+) T cells and the exclusion of Treg cells. *Front Immunol.* 2022;13:947756. <https://doi.org/10.3389/fimmu.2022.947756>.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.