

Identification KCNE4 and DNASE1L3 for Prognostic Utility in Colorectal Cancer: A Bioinformatics Analysis

Deming Liu

Ningxia Medical University

Xue Xiang

Ningxia Medical University

Yaqiong Chen

Gansu Provincial Hospital

Yajun Jiao

Ningxia Medical University

Liuli Wang

Lanzhou University

Yongfeng Wang

Gansu Provincial Hospital

Haizhong Ma

Gansu Provincial Hospital

Shixun Ma

Gansu Provincial Hospital

Hui Cai

Gansu Provincial Hospital

Tiankang Guo (✉ tiankanguo2020@163.com)

Ningxia Medical University <https://orcid.org/0000-0003-2750-6629>

Primary research

Keywords: colorectal cancer, prognostic biomarkers, immune infiltration, immune regulation, predictive indicators.

Posted Date: November 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1039855/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Objective: Immunotherapy has improved the prognosis of cancer patients who did not benefit from conventional treatment and decreased mortality, which has become an effective treatment for multiple carcinomas in their advanced stages. Patients with colorectal cancer (CRC) were mostly in the state of mismatch repair-proficient or microsatellite stability, with limited benefit in immunotherapy. Hence, immunotherapy strategies for CRC still need to be explored continuously.

Methods: Based on comprehensive analysis of the immune infiltration associated genes in the Cancer Genome Atlas (TCGA) database and differential genes for the progression and metastasis of CRC in the Gene Expression Omnibus (GEO) database, as well as validation analysis on several online databases, we obtained two genes (KCNE4 and DNASE1L3) as novel immune-related molecules for CRC.

Results: High expression of KCNE4 and DNASE1L3 were significantly correlated with CRC progression and prognosis, and were strongly associated with the infiltration of different types of macrophages in the CRC tumor microenvironment. In addition, we analysed the guidance value of KCNE4 and DNASE1L3 in anti-PD-1 therapy using data from the IMvigor210 group, and predicted the significant value of KCNE4 and DNASE1L3 in CRC immunotherapy by analyzing the correlation of KCNE4 and DNASE1L3 with the expression of CRC immune checkpoint markers.

Conclusions: We hypothesized that DNASE1L3 and KCNE4 would be potential prognostic biomarkers and predictors of immunotherapy in CRC. Our results may serve as an indication of survival prognosis and provide a new assessment indicator for the choice of immunotherapy for CRC patients.

Introduction

Colorectal cancer (CRC) remains the third most prevalent cancer and the second leading cause of cancer-related death worldwide [1]. Although the precision of traditional treatment such as surgery, chemotherapy, and radiotherapy has improved significantly, the clinical prognosis of individual cases remains unsatisfactory. Chemotherapy and radiotherapy have played an essential role in the treatment of advanced CRC by significantly improving the survival prognosis and reducing mortality over the past decades[2, 3]. However, chemotherapy and radiotherapy treatments target both healthy and cancer cells, and therefore result in significant side effects in patients. The development of resistance to chemotherapy or radiotherapy during treatment has resulted in another crisis in oncology treatment [4]. Therefore, we need to urgently develop other effective strategies to address these problems. Immunotherapy is a novel achievement in the field of oncology research, and the best candidate to overcome current problems. Immunotherapy eliminates carcinoma cells via mobilizing the patient's own immune cells to recognize specific antigens on the surface of cancer cells, decreasing the incidence of side effects associated with traditional adjuvant therapy[5]. Immunotherapy has shown encouraging results in patients with melanomas and breast cancer, and is an impactful new treatment that could revolutionize cancer treatment [6, 7]. Identification of immune related genes is particularly essential for

the effective application of immunotherapy. The benefits of providing immunotherapy to CRC patients remain debatable. Previously, immunotherapy was thought to be less effective against malignancies such as CRC, which was not an immunogenic cancer, such as melanoma and breast cancer [8]. Nevertheless, multiple clinical studies confirmed that patients with high levels of microsatellite instability (MSI-H) were sensitive to immunotherapy compared to the patients with low levels of microsatellite instability (MSI-L) and microsatellite stability (MSS) [9]. Currently, numerous immunotherapy drugs, such as anti-CTLA-4 monoclonal antibodies (ipilimumab and tremelimumab), anti-PD-1 mAbs (nivolumab and pembrolizumab), and anti-PD-L1 mAbs (atezolizumab, durvalumab, and avelumab), have been approved for the treatment of patients with progressive CRC, and have been used effectively for the treatment of metastatic CRC (mCRC) [10]. However, there were less than 20% of CRC patients who benefited from these immunotherapeutic drugs [11]. Therefore, it is necessary to explore new immune target molecules. Among the first wave of immune cytokines, Granulocyte macrophage colony-stimulating factor (GM-CSF) serves a promising target, which may be the key factors inducing cytokine storm. In our previous study, we found that GM-CSF promoted metastasis in colorectal cancer by inducing epithelial mesenchymal transition (EMT)[12].

In this study, we used the ESTIMATE algorithm to calculate the stromal and immune scores for CRC patients from the Cancer Genome Atlas (TCGA) database, and individually evaluated the correlation between stromal and immune scores and the prognosis of CRC patients. The results showed that patients with a high immune score had a favorable prognosis. Subsequently, immune genes significantly associated with CRC progression were filtered out by comparing the differential expression of genes in different immune score groups and performing soft clustering analysis. In addition, the GSE39582 was used as additional sample data, to filter out genes associated with CRC progression. The candidate genes in both cohorts were analysed by univariate cox regression analysis and intersected. Finally, we identified two novel immune-related molecules (KCNE4, DNASE1L3) and investigated their potential value for guidance and targeted therapy in CRC immunotherapy using the IMvigor210 cohort and TIMER2.0 database.

Deoxyribonuclease 1-like 3 (DNASE1L3) is a $\text{Ca}^{2+}/\text{Mg}^{2+}$ -dependent endonuclease that is lowly expressed in multiple cancers and has a significant correlation with the infiltration of immune cells in the tumor microenvironment. Analysis of previous studies showed that downregulation of DNASE1L3 may promote the immune escape process of tumor cells by affecting the CCR7/CCL19 axis, suggesting that DNASE1L3 may play an important role in the immune regulation of tumors[13, 14]. Potassium Voltage-Gated Channel Subfamily E Regulatory Subunit 4 (KCNE4), an ancillary peptide of Kv1.3, plays a critical function in the physiological processes of the immune system by controlling the abundance and localization of K^{+} channel proteins on the cell membrane surface. Some scholarly work showed that KCNE4 may be involved in the immune regulation of the tumor microenvironment in CRC[15, 16]. Our study evaluated for the first time the immunomodulatory role of KCNE4 in the tumor microenvironment of CRC and analysed the predictive value of immunotherapy efficacy. The results of this study would help clinicians to make more effective decisions regarding immunotherapy treatment for CRC patients.

Materials And Methods

Data acquisition and preprocessing

RNA-seq data for CRC patients and clinical information from TCGA were downloaded from the UCSC Xena (<http://xena.ucsc.edu/>) [17]. Gene expression data were converted from the FPKM format to transcripts per million (TPM) using the formula $TPM_i = (FPKM_i / \sum_j FPKM_j) \times 10^6$. GSE39582, a microarray expression dataset of CRC[18], Affymetrix Human Genome U133 Plus 2.0 microarray expression series matrix files were downloaded from the Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>) database. Probes were transformed to gene symbols using the annotation file from the GPL570 platform. The highest average was retained when multiple probes were matched to a gene. In both the datasets, we included only the data for CRC patients for whom complete information regarding clinicopathological characteristics and survival status were available, and normalized the expression data.

Calculation of stromal/immune score and Kaplan-Meier survival analysis

The immune, stromal, and ESTIMATE scores of CRC patients were determined from the TCGA cohort using the ESTIMATE algorithm, values were classified into the high and low immune/stromal score groups based on optimal cut-off values [19]. The optimal cut-off value was calculated using the “res.cut” function in the package “survminer”. Moreover, the Kaplan-Meier plot depicting the overall survival in the two groups was analysed, and the results were evaluated using the Wilcoxon log-rank test.

Analysis of the correlation between immune score and clinicopathological characteristics

The correlation between immune scores and clinicopathological characteristics such as age, gender, depth of tumor infiltration, lymphatic metastasis, distal metastasis, and pathologic stage were evaluated separately in CRC patients from the TCGA cohort using the Wilcoxon test or Kruskal Wallis test. The correlation graphs were produced using the ‘ggplot2’ package in R 4.0.3 software.

Evaluation of tumor infiltrating immune cells (TIICs)

TIICs in the TCGA CRC patients were assessed using CIBERSORT, a bioanalytical tool that uses a deconvolution algorithm to analyze gene expression data from 22 immune cell types. The gene expression signature matrices of 22 immune cells were downloaded from the CIBERSORT platform (<https://cibersortx.stanford.edu/>)[20]. The matrix data showing the CRC gene expression levels were compared with the feature matrix data of 22 TIICs from the CIBERSORT platform, using support vector regression, to generate a matrix showing the proportions of 22 TIICs in the high and low immune score groups, the algorithm used a default signature matrix after 1000 permutations. In addition, the algorithm was used to obtain a p-value for the deconvolution of each sample and a confidence measure for the data. Only data for which $p < 0.05$ were selected for the subsequent analysis.

Identifying differential genes

For CRC RNA-seq data from the TCGA cohort, differential expression analysis of the high and low immune score groups was performed using the 'limma' package of R-4.0.3 software. The expression data were imported into the 'lmFit' and 'eBayes' functions, to calculate the statistical differences, and genes with significant expression related differences were selected for subsequent analysis, based on the filter criteria: $|\log \text{ fold change}(\log \text{FC})| > 3$ and $p\text{-value} < 0.05$. Then, the "ggplot2" package of R 4.0.3 was utilized to create volcano plots for the differential genes (DEGs).

For the GSE39582, we classified the pathologic stage into early (I-II) and advanced (III-IV) stage groups, and the pathologic M stages into non-metastatic (M0) and metastatic (M1) groups. First, expression values were normalized using the 'normalizeBetweenArrays' function of the 'limma' package. Differential analysis was performed using the 'limma' package. The DEGs in the advanced and early-stage groups were distinguished by whether the $p\text{-value} < 0.05$ and $|\log \text{ FC}| > 0.15$. Screening for non-metastatic (M0) and metastatic (M1) groups was performed using the following screening criteria: $p\text{ value} < 0.05$ and $|\log \text{ FC}| > 0.2$. Subsequently, genes that promote or suppress CRC progression were identified separately via cross-tabulation with Venn diagrams.

Functional Enrichment Analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed by the "enrichGo" and "enrichKEGG" functions in the "clusterProfiler" package. GO enrichment analysis consisted of three components: biological process (BP), cellular component (CC), and molecular function (MF)[21]. If $p\text{-values} < 0.05$, values were considered statistically significant. We had only performed the analysis for DEGs that were positively regulated during tumor progression.

Soft clustering analysis of gene expression data in a time-series way

After determining the correlation between immune score and clinicopathological characteristics in the TCGA cohort, the pathologic stage was classified as I, II, III, and IV, and pathologic M stages were classified as M0, M1, M1a, M1b. The 'Mfuzz' package in R was used as a tool to classify DEGs with high/low immune score groups into clusters representing different stages. Unlike GO and KEGG, Mfuzz used a fuzzy c-mean algorithm based on the minimization of a weighted squared error function in the 'mfuzz' function to achieve the clustering of genes at different periods or states, based on gene expression profiles[22]. To prevent the clustering of too much random data, we set the following parameters: minimum standard deviation=0.5, score=0.5. Finally, genes in clusters that showed a consistent expression trend were consolidated, and genes that promoted/suppressed CRC progression were merged separately using a Venn diagram.

Univariate Cox Regression Analyses

Univariate Cox regression was performed for each of the DEGs obtained from the GSE39582 dataset and for genes available via soft clustering from the TCGA cohort, genes with p-values < 0.05 were screened for subsequent analysis. Univariate analysis was carried out using the 'survival' package in R. Finally, meaningful genes obtained from the GSE39582 dataset and the TCGA cohort dataset were intersected using intersection graphics.

Differential expression analysis and survival analysis

The UALCAN (<http://ualcan.path.uab.edu/index.html>) database was used to analyze the differential expression of candidate genes in healthy and cancerous patients, as well as the correlation between candidate genes and tumor progression. Their association with overall survival (OS) and disease-free survival (DFS) was analysed simultaneously in Colon adenocarcinoma (COAD) patients. The UALCAN database facilitates the comprehensive analysis of cancer OMICS data using the TCGA database, MT500, and CPTAC[23].

In addition, the association between potential candidate genes and OS and DFS in COAD was analysed utilizing the GEPIA (Gene Expression Profiling Interactive Analysis) database (<http://gepia.cancer-pku.cn/index.html>), using the median value as the cutoff value. GEPIA, developed by Peking University, contains expression data from 9,736 tumor patients derived from TCGA, and allows for functional analyses, including differential genes expression analysis, correlation analysis, patient survival analysis, and similar gene detection and dimensionality reduction analysis[24].

Analysis of protein expression

The Human Protein Analysis (HPA) database (<https://www.proteinatlas.org>) was applied to evaluate candidate genes expression levels with regard to the OS of CRC patients. Candidate genes expression levels were determined in healthy and cancerous tissues using immunohistochemical images from 'The Tissue Atlas' and 'The Pathology Atlas' [25]. The HPA is a large proteomics database based on RNA sequencing and immunohistochemical analysis, which is designed to map information regarding all human proteins in cells, tissues, and organs via the integration of various omics technologies. It can be used for the analysis of differential protein expression in healthy and cancerous tissues, and the subcellular localization of proteins.

Immune Infiltration Analysis

The correlation between the infiltration level of macrophages and CD8+ T cells with candidate genes expression was assessed using the Tumor Immune Estimation Resource 2.0 (TIMER 2.0) database. Unlike TIMER, which uses one algorithm, TIMER 2.0 (<http://timer.cistrome.org/>) uses six deconvolution algorithms to provide better estimates of the level of immune infiltration for the TCGA tumor profile. It also provides four modules for investigating the relationship between immune infiltration and clinical features[26]. We assessed the correlation of the infiltration of M1 macrophages, M2 macrophages, and CD8+ T cells with candidate genes expression using the QUANTISEQ deconvolution algorithm, which

uses the Spearman method. Because this deconvolution algorithm did not provide information regarding M0 macrophage infiltration, we assessed M0 macrophage infiltration using the CIBERSORT algorithm.

Analysis of the potential effect in immunotherapeutic responses

The "Gene_corr" module in the TIMER2.0 database was used to analyze the correlation between DNASE1L3 and KCNE4 and eight immune checkpoint markers (currently common PDCD1, PD-L1, CTLA-4, TIGIT, and TIM-3, as well as potential markers CSF1, CSF1R, and SIRPA) in COAD and Rectum adenocarcinoma (READ). The IMvigor210 cohort from the 'IMvigor210CoreBiologies' package (<http://research-pub.gene.com/IMvigor210CoreBiologies>) (including 298 urothelial cancer cases who received immunotherapy with complete clinical information and expression data) was used to evaluate the effect of candidate genes in immunotherapy.

Statistical Analysis

All statistical analyses were performed in R 4.0.3 software and online databases. The Wilcoxon rank sum test was used to analyze the expression signatures of paired samples from the GEO and TCGA cohorts, and the Kruskal Wallis test was used to compare differences in expression among multiple tumor stages. All statistical tests were two-sided and p-values < 0.05 were considered to be statistically significant.

Results

Processing of expression and clinical data

The flowchart of the analysis process for this study was shown in Fig. 1. Expression and clinical information for 458 CRC patients from the TCGA cohort were downloaded from the UCSC Xena (<http://xena.ucsc.edu/>). After excluding patients with missing clinicopathological and follow-up information, 375 CRC patients were included in the follow-up analysis. Probe conversion files for the UCSC Xena database were downloaded, along with expression data. The microarray dataset GSE39582, which contained 585 CRC patients, was downloaded from the GEO database. Thus, a total of 546 samples were included. The clinical characteristics of the patients included in the above two cohorts were described thoroughly in Table 1.

Table 1

Detailed information on the clinical characteristics of eligible CRC patients in the TCGA cohort and GSE39582 cohort.

Clinical characteristics of CRC patients in TCGA datasets (Total=375)			Clinical characteristics of CRC patients in TCGA datasets (Total=546)		
Characteristics	Group	N (%)	Characteristics	Group	N (%)
age	>=60	273 (72.8%)	age	>=60	398 (72.9%)
	<60	102 (27.2%)		<60	148 (27.1%)
gender	female	173 (46.1%)	gender	female	250 (45.8%)
	male	202 (53.9%)		male	296 (54.2%)
Pathologic T	T1	9 (2.40%)	Pathologic T	T1	12 (2.20%)
	T2	65 (17.3%)		T2	47 (8.61%)
	T3	260 (69.3%)		T3	372 (68.1%)
	T4	23 (6.13%)		T4	115 (21.1%)
	T4a	13 (3.47%)			
	T4b	5 (1.33%)			
Pathologic N	N0	224 (59.7%)	Pathologic N	N0	305 (55.9%)
	N1	60 (16.0%)			
	N1a	12 (3.20%)		N1	136 (24.9%)
	N1b	12 (3.20%)			
	N1c	2 (0.53%)		N2	99 (18.1%)
	N2	53 (14.1%)			
	N2a	5 (1.33%)		N3	6 (1.10%)
	N2b	7 (1.87%)			
Pathologic M	M0	315 (84.0%)	Pathologic M	M0	486 (89.0%)
	M1	48 (12.8%)			
	M1a	9 (2.40%)		M1	60 (11.0%)
	M1b	3 (0.80%)			
Pathologic stage	I	36 (6.59%)	Pathologic stage	I	36 (6.59%)
	II	257 (47.1%)		II	257 (47.1%)

Clinical characteristics of CRC patients in TCGA datasets (Total=375)		Clinical characteristics of CRC patients in TCGA datasets (Total=546)	
III	194 (35.5%)	III	194 (35.5%)
IV	59 (10.8%)	IV	59 (10.8%)
Supplementary Material			
Supplemental Table S1.xls: Cellular Component;			
Supplemental Table S2.xls: Molecular Function;			
Supplemental Table S3.xls: Biological Process;			
Supplemental Table S4.xls: Kyoto Encyclopedia of Genes and Genomes;			
Supplemental Table S5.xls: Results of univariate regression analysis for the TCGA cohort			
Supplemental Table S6.xls: Results of univariate regression analysis for the GSE39582 dataset			

Immune scores were significantly correlated with the overall survival of CRC patients

The Kaplan-Meier curve showed that the survival rate of CRC patients in the high immune score group was significantly higher than that in the low immune score group (p-value = 0.039, Fig. 2a). Conversely, the OS of CRC patients with high stromal scores was lower than that in those with low stromal scores, but values were not statistically significant (p-value = 0.071, Fig. 2b). Furthermore, the ESTIMATE score was not significantly correlated with survival in CRC patients (p-value = 0.12, Fig. 2c).

Significant correlation between immune scores and pathological characteristics of CRC patients

While analyzing the results, we concluded that the immune scores were significantly correlated with pathologic stage progression. Lower immune scores were associated with stage IV disease, compared to scores associated with stages I, II, and III (p-value = 0.013, Fig. 3f). In addition, there was a significant association between immune scores and tumor metastasis (pathologic M stage) (p-value = 0.0022, Fig. 3e). Patients with low immune scores were more likely to develop metastases. However, there was no statistically significant correlation between the immune scores and patient age (p-value = 0.19, Fig. 3a), gender (p-value = 0.26, Fig. 3b), pathologic T stage (p-value = 0.63, Fig. 3c), and pathologic N stage (p-value = 0.2, Fig. 3d). Data were estimated by the Wilcoxon rank-sum test or Kruskal Wallis test.

Composition of tumor infiltrating cells in the high/low immune score groups

We applied the CIBERSORT deconvolution algorithm to evaluate the compositional framework of 22 TIICs in the TCGA cohort. Based on a calculated P-value < 0.05, a total of 171 patients satisfied the criteria, including 160 and 11 patients with high and low immune scores, respectively. Results for immune cells fractions showed that the immune cells infiltrating around the CRC tissues included mainly macrophages and T cells (Fig. 4a). The differential infiltration of 22 types of immune cells in the high/low immune score groups was demonstrated in Fig. 4b. The proportions of infiltrating CD8+ T cells (p-value = 0.0048), M1 macrophages (p-value = 0.0002), M2 macrophages (p-value = 0.004), and neutrophils (p-value = 0.0022) in the high immune score group were higher than those in the low immune score group. In contrast, the infiltration percentages of memory B cells (p-value = 0.0071), naive CD4+ T cells (p-value = 0.0022), and resting NK cells (p-value = 0.046) were decreased. Notably, the proportion of infiltrating M0 macrophages was down-regulated in the high immune score group, while the proportions of activated CD4+ memory T cells, follicular helper T cells, and activated NK cells were up-regulated, however, no statistically significant differences were observed. The correlation for TIICs (Fig. 4c) showed that there was a weak to moderate correlation between different immune cells infiltrating into the tumor tissue. Based on the proportion of immune cells which were up-regulated in the high immune fraction group, we focused on the immune cells associated with infiltration of CD8+ T cells, M1 macrophages and M2 macrophages. Positive correlations with M1 macrophages, including activated CD4+ memory T cells (R=0.35), activated NK cells (R=0.33), and resting mast cells (R=0.35) were observed, while negative correlations with M0 macrophages (R=-0.38) and activated mast cells (R=-0.37) were observed. Positive correlations with M2 macrophages, including resting dendritic cells (R=0.24), eosinophils (R=0.21), and M1 macrophages (R=0.21) were observed, and negative correlations with regulatory T cells (Tregs) (R=-0.23) and resting NK cells (R=-0.23) were observed. Positive correlations with CD8+ T cells, including follicular helper T cells (R=0.44), activated CD4+ memory T cells (R=0.49), and M1 macrophages (R=0.29) were observed; negatively correlations with M0 macrophages (R=-0.48), activated mast cells (R=-0.34), and resting CD4+ memory T cells (R=-0.32) were observed.

Identification of differential genes in high/low immune score groups and functional enrichment analysis

To determine immune correlated DEGs in CRC patients, 3311 up-regulated and 5125 down-regulated genes were screened based on the filter conditions: $|\log_2FC| > 3$ and $p\text{-value} < 0.05$, and the results were shown in the volcano plot (Fig. 5a). Subsequently, we analysed DEGs via GO functional enrichment analysis at three biological levels: MF, BP, and CC. The results revealed that the DEGs were significantly associated with several immune regulatory processes, including the lymphocyte mediated immune response, regulation of lymphocyte activation, immunoglobulin receptor binding, and T cell receptor complex. The enrichment analysis of the regulatory pathways involved in DEGs showed that DEGs affected the process of immune regulation during tumor progression through multiple major biological pathways, including the chemokine, NF-kappa B, and T cell receptor signaling pathways. In addition, we also discovered that DEGs were involved in PD-L1 expression and the PD-1 check-point pathway in cancer

patients. Some of the results of enrichment analysis were shown in Fig. 5b. The results of enrichment analysis were displayed in detail in **Supplementary Table S1-4**.

For the microarray dataset GSE39582, there were 1182 DEGs, of which 732 were up-regulated and 450 were down-regulated in advanced CRC samples, compared to early CRC patients. In addition, a total of 1030 DEGs were selected in the progression of pathologic M stage. Of these, 632 were up-regulated and 398 were down-regulated. Subsequently, intersections were generated for genes that promoted (up-regulated) or inhibited (down-regulated) CRC progression. The results showed that 136 genes suppressed tumor progression, while 275 genes promoted tumor progression. These results were displayed using volcano plots and figures illustrating the intersections (Fig. 5c-e).

Soft Composition of tumor infiltrating cells in the high/low immune score groups

During soft clustering analysis of the pathologic stage (Fig. 6a), gene expression in clusters 1 and 8 showed a consistent increasing trend with pathological stage progression, while gene expression in clusters 2 and 9 showed a consistent decreasing tendency. The genes in clusters 1 and 8, a total of 1411 genes associated with the promotion of the pathologic stage were combined. The genes in clusters 2 and 9, a total of 1485 genes associated with the inhibition of pathologic stage progression were combined. Besides, during the soft clustering analysis of the pathologic M stage (Fig. 6b), gene expression in clusters 5 and 10 showed a persistent increasing trend with progression of the pathologic M stage, while gene expression in clusters 4 and 6 showed a persistent decreasing trend. Similarly, the genes in clusters 5 and 10, a total of 1759 genes, which were pathologic M stage promoting genes, were combined. The genes in clusters 4 and 6, a total of 1321 genes, which suppressed pathologic M stage progression, were combined. Finally, we intersected genes that promoted the progression of a certain pathologic stage and pathologic M stage, a total of 327 genes, identified as immune related genes that promoted CRC progression. Besides, 640 genes suppressed the progression of the pathologic stage and pathologic M stage, and inhibited CRC progression. These results were presented in a Venn diagram (Fig. 6c).

Univariate Cox regression analysis

Univariate Cox regression analysis of the genes obtained from the above two cohorts was carried out using the survival package of R software. The results showed that a total of 97 genes were associated with patient survival in the TCGA cohort, along with 62 risk factors and 35 protective factors (**Supplementary Table S5**). A total of 181 genes were associated with patient survival in the GSE39582, along with 100 risk factors and 81 protective factors (**Supplementary Table S6**). We intersected the above genes with the same role. Three genes were eventually identified, of which two were risk factors (DNASE1L3, LGALS4) and one was a protective factor (KCNE4) (Fig. 7).

Validation of survival analysis and differences in expression

Differential analysis of CRC and healthy patients showed that high levels of DNASE1L3 and LGALS4 were expressed in healthy patients, while high levels of KCNE4 were expressed in CRC patients. During pathologic stage progression in CRC patients, DNASE1L3 and LGALS4 were negatively correlated with pathologic stage progression, while KCNE4 was significantly positively correlated with pathologic stage progression (Fig. 8a-b). KCNE4, DNASE1L3, and LGALS4 overexpression in tumors were significantly associated with OS and disease-free survival (DFS) in COAD patients. Patients with high LGALS4 expression levels had favorable DFS (HR=0.34, p=0.0051) and OS (HR=0.58, p=0.0028), and patients with high expression levels of KCNE4 had poor OS (HR=1.8, p=0.021) and DFS (KCNE4: HR=1.7, p=0.025), according to the GEPIA database. Moreover, patients with high LGALS4 (p=0.013) and DNASE1L3 (p=0.013) expression levels exhibited a good OS, while patients with high expression levels of KCNE4 (p=0.002) had a poor OS, according to the ULCAN database (Fig. 8c-e).

Analysis of expression of candidate gene associated proteins

We performed immunohistochemical analysis of 3 candidate genes using the HPA database. The results revealed that the expression levels of DNASE1L3 and LGALS4 in the healthy colon and rectal tissues were higher than those in CRC tissues, and the survival curves showed that genes expressed at high levels had a higher survival rate. DNASE1L3 was expressed at moderate levels in one healthy colon tissue sample and one healthy rectal tissue sample, and at low levels in 6 of 11 CRC samples (Fig. 9a). When the same antibody (antibody CAB025653) was used to detect LGALS4 expression, high LGALS4 expression levels were observed in 3 healthy colon and 3 healthy rectal samples, LGALS4 was not detected in 11 of 12 CRC samples (Fig. 9b). KCNE4 was moderately expressed in three colon tissues and highly expressed in two rectal tissues. KCNE4 was expressed at moderate levels in 10 of 12 CRC tissues, and at low levels in 2 CRC tissues. However, the survival curve showed that the prognosis of patients with high expression levels of KCNE4 was poorer than that of patients with low expression levels of KCNE4 (Fig. 9c).

Correlation of macrophage and CD8+ T cell infiltration with candidate genes expression

We used the QUANTISEQ deconvolution algorithm in the TIMER2.0 database to validate the correlation between candidate genes expression and the infiltration of CD8+ T cells and three types of macrophages in COAD and rectum adenocarcinoma (READ) patients. The results showed that the expression levels of DNASE1L3 in COAD (r=0.239, p<0.01) and READ (r=0.324, p<0.01) patients were significantly positively correlated with M1 macrophage infiltration (Fig. 10a). The expression levels of KCNE4 in both COAD (r=0.466, p<0.01) and READ (r=0.532, p<0.01) patients were significantly positively correlated with M2 macrophage infiltration (Fig. 10c). There was no significant correlation between the expression of LGALS4 and the infiltration of both the immune cells (Fig. 10b).

The role of candidate genes in the prediction of immunotherapeutic benefits

Based on the above results, we analysed the correlation of the expression between two candidate genes (KCNE4, DNASE1L3) and immune checkpoint markers using the TIMER2.0 database. The graph (Fig. 11a) showed that in both COAD and READ, there was a significant correlation between the expression of KCNE4 and various immune checkpoint markers, especially TIM-3(HAVCR2), CSF1, CSF1R, SIRPA, compared to DNASE1L3, while the correlation with the expression of PD-1(PDCD1) was lower. In addition, in order to evaluate the contribution of candidate genes in the immunotherapy benefit population, we incorporated patients in the IMVigor210 cohort receiving anti-PD-L1 immunotherapy into further study analyses. The correlation map of gene expression in the IMVigor210 cohort revealed a stronger correlation between DNASE1L3 and PDCD1 expression ($r=0.5$) compared to KCNE4 ($r=0.2$) (Fig. 11b). Similarly in survival prognosis and treatment effects, patients with high DNASE1L3 had a superior prognosis and were better able to benefit from anti-PD-1 therapy (Fig. 11c), while patients with high KCNE4 had the opposite result (Fig. 11d).

Discussions

CRC is one of the common lethal cancers in the world, and the treatment is mainly based on surgery, with chemotherapy and radiotherapy as adjuvant modalities[27]. As most patients developed recurrence or metastasis during the follow-up treatment, it directly decreased the survival rate of patients[28, 29]. Therefore, it is a continuous goal to explore new therapeutic means for the treatment of CRC. With the progress of oncology research, immunotherapy has become a novel member of antitumor therapy, and meanwhile the importance of the immune microenvironment in tumor progression was gradually unveiled. Notably, the tumor immune microenvironment (TIM) played a double-edged role in determining the 'life' of malignancy and was significantly associated with the survival of patients with malignant tumors[30]. Previous studies had revealed that the composition of the immune microenvironment in CRC directly affects the prognosis of CRC patients, for instance, the enrichment of lymphocytes and CD8+ T lymphocytes was beneficial to improve the survival of CRC patients. However, high levels of mast cells, tumor-infiltrating Th17 cells, cause poor prognosis in CRC patients[31, 32]. Currently, the stromal/immune scores were the most commonly used measure for the bioinformatic evaluation of immune cell infiltration in the tumor microenvironment[19]. Here, the ESTIMATE algorithm was used to calculate the stromal/immune score of CRC patients from the TCGA cohort, and we found that patients in the high immune score group had a higher survival rate compared to the low immune score group. The correlation between the immune scores and clinical characteristics of CRC patients was showed that there was a significant negative correlation between the immune score and some of the clinical characteristics (pathologic stage and pathologic M stage) of patients, and that the immune score progressively decreased with tumor progression. High immune scores, i.e., immune cell infiltration played a positive role in prolonging CRC patient survival, and this finding was consistent with previous researches on immune infiltration in CRC[31].

Macrophages reportedly played a vital role in the progression of multiple malignancies, and different cell subtypes played distinct roles in the progression of malignancy. M1 macrophages can activate the immune response by secreting pro-inflammatory factors such as IL-4 and INF- γ to inhibit the progression of malignant tumors, whereas M2 macrophages can inhibit the immune response by secreting cytokines such as IL-10 and TGF- β , to enhance the immune escape of malignant tumor cells and promote tumor progression[33–35]. In contrast to the findings of previous studies regarding the promotion of tumor progression by M2 macrophages, certain studies had shown that M2 macrophages, with CD163 as a marker, might have synergistic effects and participate in tumor suppression, and were positively correlated with good prognosis[36, 37]. Our study showed similar results, with an increased proportion of M2 macrophages infiltrating in the high immune score group. This may be attributable to the heterogeneity of malignant tumors and the low accuracy of RNA-seq based algorithms. CD8+ T cells, also known as cytotoxic T lymphocytes (CTLs), have been recognized as the main anti-tumor immune effector cells, and their high levels of infiltration were associated with a favorable prognosis in various cancers [38]. Minority studies demonstrated that CD8+ T cells had a significant positive correlation with the chemotherapeutic effectiveness and immunotherapy response rates[39, 40]. Results of our study were consistent with the extant consensus, which showed that the proportion of CD8+ T cell infiltration was second only to that of macrophages. The proportion of infiltrated CD8+ T cells was significantly higher in the high immune score group, compared to that in the low immune score group, for which the prognosis was worse. Correlation plots for the infiltration of 22 TIICs showed that CD8+ T cells were positively correlated with the infiltration of M1 macrophages ($r=0.29$), activated memory CD4+ T cells ($r=0.49$), follicular helper T cells ($r=0.44$), and were negatively correlated with M0 macrophages($r=-0.48$). This also illustrates the complexity of the interaction of immune infiltrating cells during tumor progression.

Based on these results, we identified the DEGs between high and low immune score groups and performed functional enrichment analysis for DEGs with a positive regulatory role in the high immune score group. The results showed that DEGs were significantly associated with immune regulation in terms of the molecular functions, biological processes, and cellular components, and participated in immune-related pathways such as the PD-1 checkpoint pathway and chemokine signaling pathway. Programmed cell death 1 (PD1) and its ligand (PDL1) are critical regulatory physiological immune checkpoints. The PD-L1 present on tumor cells may inhibit T cell mediated immune responses by binding to the PD-1 receptor on the surface of T cells[41]. Antibody drugs that target PD-L1 or PD-1 to block the PD-1/PD-L1 pathway are highly effective for the enhancement of T cell anti-tumor immunotherapy[42]. The chemokine signaling pathway is an essential pro-inflammatory pathway, in which chemokines recruit leukocytes to infiltrate into cells and produce local inflammatory effects in response to an exogenous pathogen attack. Chemokine signals are conveyed by chemokine receptors (G-protein coupled receptors) expressed on immune cells. Upon the activation of these receptors, the α - and β - γ -subunits of G-proteins dissociate to activate different downstream pathways, leading to cell polarization and actin reorganization[43]. Thus, CRC patients can respond effectively to immunotherapies such as anti-PD-1 owing to the immune response generated by immune cells. Our results had demonstrated that immune scores were significantly negatively correlated with the progression of certain pathologic stages and the

pathologic M stage in patients. Hence, we re-clustered DEGs by their pathologic stage and pathologic M stage in chronological order using a soft clustering analysis algorithm. Then, we analysed the clustered genes using univariate Cox regression to identify genes that were significantly associated with CRC prognosis. To avoid biased results resulting from the small sample size and homogeneity of the data platform, we downloaded the GSE39582 dataset from the GEO database as an additional dataset for the screening of variables. Upon the intersection of the candidate genes generated from the above two datasets, three immune-related genes were identified. Of these, DNASE1L3 and LGALS4 were protective factors and KCNE4 was a risk factor.

The results of candidate genes expression analysis showed that the expression levels of DNASE1L3 and LGALS4 in healthy patients were significantly higher than those in CRC patients, and were negatively correlated with tumor progression, patients with high expression levels of DNASE1L3 and LGALS4 received a favorable prognosis. However, expression levels of KCNE4 were significantly high in CRC patients, and were positively correlated with tumor progression, patients with a high expression level of KCNE4 received a worse prognosis. The expression levels of proteins associated with the three candidate genes were evaluated using the immunohistochemical specimens available in the HPA database. Moderate to high expression levels of DNASE1L3 and LGALS4 related proteins were observed in healthy colon/rectal tissues, while a low or negative expression level was observed in CRC tissues. KCNE4 related proteins were mainly moderately expressed in both healthy and tumor tissues. Since the number of samples identified via immunohistochemistry analysis of the database was limited, the differential expression of the relevant proteins could not be identified in healthy and cancer tissues. Therefore, the evaluation of varied expression levels in different tissues based on the only available sample size could result in somewhat inaccurate results, and this method was only used for a preliminary assessment of cellular localization and the expression level of the protein of interest. Subsequently, we primarily evaluated the correlation between the expression of three candidate genes and the infiltration of macrophages and CD8+ T cells. The results demonstrated that the expression of DNASE1L3 was positively correlated with the infiltration of M1 macrophages and negatively correlated with the infiltration of M0 macrophages. KCNE4 expression was positively correlated with the infiltration of M2 macrophages. It was suggested that the overexpression of DNASE1L3 might be associated with the differentiation or recruitment of M1 macrophages, while the overexpression of KCNE4 might be associated with the differentiation or recruitment of M2 macrophages. Based on the above comprehensive screening and analysis, we ultimately identified two immune related genes associated with survival in CRC patients, namely DNASE1L3 and KCNE4.

In the tumor microenvironment, macrophages constitute the highest percentage of cells, and they were involved in the progression of malignancy through direct or indirect processes, including angiogenesis, tumor cell invasion, metastasis, and regulation of the tumor microenvironment[44]. The surface molecules of macrophages performed an essential role in tumor progression. CSF1R (colony-stimulating factor 1 receptor) could assist tumor-associated macrophages to infiltrate into the tumor surroundings, of which associated ligand colony-stimulating factor 1 (CSF1) can also promote macrophage differentiation[45]. CD47 on the surface of tumor cells interacts with (signal regulatory protein alpha)

SIRPA on the surface of macrophages to promote the immune escape of tumor cells[46]. Targeting molecular markers on the surface of macrophages could effectively reduce macrophage recruitment in the tumor microenvironment, differentiation, and immune escape of tumor cells[47]. As a primary member of the body's immune regulation, T cells protect the body from pathogens, tumors and other damage by recognizing specific antigens. However, the combination of tumor cells and tumor microenvironment during tumor progression can inhibit the anti-tumor effect of T cells, and even cause exhaustion of T cells[48, 49]. TIM-3 (T cell immunoglobulin and mucin-domain containing-3), also known as HAVCR2, which was produced on the surface of exhausted T cells undergoing chronic tumor antigen stimulation, had been used as a marker of T cell exhaustion, and targeted inhibition of TIM-3 may be expected to restore the anti-tumor capability of T cells[50, 51]. The information from the IMVigor210 cohort was utilized as an efficacy evaluation dataset for immunotherapy, and the analysis showed that there was a higher positive correlation between DNASE1L3 and PDCD1, compared to KCNE4, and that patients with high DNASE1L3 were able to benefit more from treatment with anti-PD-1. The survival prognosis of patients in the IMVigor210 cohort also confirmed the benefit, as the results showed that patients with high DNASE1L3 had prolonged survival. The correlation between candidate genes and immune checkpoint markers predicts, in part, the effect of immunotherapy. In the analysis of the TCGA CRC queue, there was a significant positive correlation between the expression of KCNE4 and multiple immune checkpoint markers, except for PDCD1. In the TCGA CRC cohort, DNASE1L3 was unsatisfactory in immunotherapy prediction, which may be related to tumor heterogeneity and limited samples. Nevertheless, DNASE1L3 correlated higher with PDCD1, CSF1R, and TIGIT compared to other immune checkpoint markers. Furthermore, the correlation of DNASE1L3 with PDCD1 was higher than that of KCNE4, which was generally similar to the results of the IMVigor210 cohort, and these findings suggested the potential predictive value of DNASE1L3 in predicting anti-PD-1 therapy in CRC patients.

DNASE1L3 is a member of the deoxyribonuclease 1 family, which has the capacity to degrade chromosomal DNA strands [14, 52]. DNASE1L3 is mainly secreted by bone marrow cells, and functional nuclear localization signals facilitate its transportation into the nucleus, where it participates in the fragmentation of nuclear endosomal DNA during apoptosis and necrosis[53, 54]. Previous studies have shown that DNASE1L3 was closely associated with the development of autoimmune diseases, and the absence or mutation of DNASE1L3 resulted in an anti-DNA response in the body. This in turn resulted in autoimmune diseases such as systemic lupus erythematosus and systemic sclerosis[55, 56]. In recent years, the prevalence of bioinformatics technology has resulted in the identification of the role of DNASE1L3 in cancer progression. For instance, DNASE1L3 was involved in signal transduction in breast cancer patients, and could be an independent prognostic factor for patient survival after the radical resection of liver cancer [57, 58]. The role of DNASE1L3 in colorectal cancer has been examined to a lesser extent. In our study, DNASE1L3 was found to be significantly associated with M1 macrophage infiltration in the CRC tumor microenvironment and significantly correlated with CRC progression and survival, hence, it might represent a new therapeutic target for CRC treatment.

KCNE4 is the fourth and largest member of the KCNE Kv channel regulatory subunit family [59, 60]. A previous study has shown that it was mainly expressed in abundant amounts in the human brain [61].

KCNE4 acts as an ancillary peptide of Kv1.3 and negatively regulates K⁺ channels on the surface of immune cells, mainly by controlling the abundance and localization of K⁺ channel proteins on the cell membrane surface, inhibiting outward K⁺ currents, and accelerating channel inactivation[15]. It plays an essential role in the proliferation, activation, and apoptosis of leukocytes and other physiological processes of the immune system, as well as in angiogenesis [62, 63]. The role of KCNE4 in malignant tumors has rarely been studied. Using bioinformatics techniques, we determined that KCNE4 was significantly expressed in CRC patients and was positively correlated with tumor progression. In addition, KCNE4 expression was also significantly correlated with patient prognosis and positively correlated with M2 macrophage infiltration in the tumor immune microenvironment. KCNE4 might play an important role in the tumor microenvironments of CRC patients and could represent a novel target molecule.

Although we have assessed several noteworthy immune-related factors that might be useful for CRC immunotherapy treatment, certain limitations were associated with our study. First, the heterogeneity of the immune microenvironment in malignant tumors inevitably leads to biased results. Second, despite combining the data from the TCGA and GEO databases for the screening of genes, the databases for validating the results were established based on data from the TCGA database with relatively small sample size. In addition, the immune scores of most malignant tumor tissues and the proportions of infiltrating immune cells were calculated using the ESTIMATE and CIBERSOTR algorithms, but because these algorithms are not completely stable, the results need to be confirmed via additional experiments.

Conclusion

we identified two immune-related genes that were significantly associated with the progression and prognosis of CRC, and assessed the predictive value of DNASE1L3 and KCNE4 in response to immunotherapy. Our findings may provide new assessment indicators for the choice of immunotherapy for CRC patients.

Abbreviations

CRC
colorectal cancer
TCGA
the Cancer Genome Atlas
GEO
Gene Expression Omnibus
MSI-H
microsatellite instability
MSI-L
microsatellite instability
MSS
microsatellite stability

mCRC
metastatic CRC
GM-CSF
Granulocyte macrophage colony-stimulating factor Figures
EMT
epithelial mesenchymal transition
DNASE1L3
Deoxyribonuclease 1-like 3
KCNE4
Voltage-Gated Channel Subfamily E Regulatory Subunit 4
GO
Gene Ontology
KEGG
Kyoto Encyclopedia of Genes and Genomes
TIICs
tumor infiltrating immune cells
OS
overall survival
DFS
disease-free survival
COAD
Colon adenocarcinoma
READ
Rectum adenocarcinoma
GEPIA
Gene Expression Profiling Interactive Analysis
HPA
Human Protein Analysis
TIMER2.0
Tumor Immune Estimation Resource 2.0
TIM
immune microenvironment
CTLs
cytotoxic T lymphocytes
PD1
Programmed cell death 1
CSF1
colony-stimulating factor 1.

Declarations

Acknowledgements

The authors thank the Key Laboratory of Molecular Diagnosis and Precision Medicine in Surgical Oncology of Gansu Province and the DaVinci Surgery System Database (DSSD, www.davincisurgerydatabase.com) for their guidance and support in the methodology.

Author contributions

Deming Liu and Xue Xiang contribute equally to the article. All authors contributed to the concept and design of the study Deming Liu and Xue Xiang drafted the manuscript; Yaqiong Chen and Yajun Jiao downloaded and analysed the data; Deming Liu and Liuli Wang performed the visual mapping of the results; Yongfeng Wang, Haizhong Ma and Shixun Ma interpreted the data for the work; Hui Cai proofread and revised the manuscript; Tiankang Guo is responsible for the integrity of the entire work. All authors read and approved the final version of the manuscript.

Funding

This study was supported by Natural Science Foundation of Gansu Province, China (No.18JR3RA052), Lanzhou Talent Innovation and Entrepreneurship Project Task Contract (No. 2016-RC-56), Gansu Da Vinci robot high-end diagnosis and treatment team construction project, National Key Research and Development Program (No. 2018YFC1311500), National Natural Science Foundation of China (No. ZYYDDFFZZJ-1) and 2021 Graduate Innovation Fund Project of Gansu University of Chinese Medicine (No. 2021CX54).

Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no conflicts of interest.

References

1. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2020**. *CA: a cancer journal for clinicians* 2020, **70**(1):7–30.
2. Aschele C, Bergamo F, Lonardi S: **Chemotherapy for operable and advanced colorectal cancer**. *Cancer treatment reviews* 2009, **35**(6):509–516.
3. Cantero-Muñoz P, Urién MA, Ruano-Ravina A: **Efficacy and safety of intraoperative radiotherapy in colorectal cancer: a systematic review**. *Cancer letters* 2011, **306**(2):121–133.
4. Cui C, Yang J, Li X, Liu D, Fu L, Wang X: **Functions and mechanisms of circular RNAs in cancer radiotherapy and chemotherapy resistance**. *Molecular cancer* 2020, **19**(1):58.
5. Johdi NA, Sukor NF: **Colorectal Cancer Immunotherapy: Options and Strategies**. *Frontiers in immunology* 2020, **11**:1624.
6. Sullivan RJ, Flaherty KT: **Immunotherapy: Anti-PD-1 therapies-a new first-line option in advanced melanoma**. *Nature reviews Clinical oncology* 2015, **12**(11):625–626.
7. Esteva FJ, Hubbard-Lucey VM, Tang J, Pusztai L: **Immunotherapy and targeted therapy combinations in metastatic breast cancer**. *The Lancet Oncology* 2019, **20**(3):e175-e186.
8. Banerjea A, Bustin SA, Dorudi S: **The immunogenicity of colorectal cancers with high-degree microsatellite instability**. *World journal of surgical oncology* 2005, **3**:26.
9. Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, Diaz LA, Jr.: **Immunotherapy in colorectal cancer: rationale, challenges and potential**. *Nature reviews Gastroenterology & hepatology* 2019, **16**(6):361–375.
10. Ciardiello D, Vitiello PP, Cardone C, Martini G, Troiani T, Martinelli E, Ciardiello F: **Immunotherapy of colorectal cancer: Challenges for therapeutic efficacy**. *Cancer treatment reviews* 2019, **76**:22–32.
11. Feng M, Zhao Z, Yang M, Ji J, Zhu D: **T-cell-based immunotherapy in colorectal cancer**. *Cancer letters* 2021, **498**:201–209.
12. Chen Y, Zhao Z, Chen Y, Lv Z, Ding X, Wang R, Xiao H, Hou C, Shen B, Feng J *et al*: **An epithelial-to-mesenchymal transition-inducing potential of granulocyte macrophage colony-stimulating factor in colon cancer**. *Scientific reports* 2017, **7**(1):8265.
13. Deng Z, Xiao M, Du D, Luo N, Liu D, Liu T, Lian D, Peng J: **DNASE1L3 as a Prognostic Biomarker Associated with Immune Cell Infiltration in Cancer**. *OncoTargets and therapy* 2021, **14**:2003–2017.
14. Serpas L, Chan RWY, Jiang P, Ni M, Sun K, Rashidfarrokhi A, Soni C, Sisirak V, Lee WS, Cheng SH *et al*: **Dnase1l3 deletion causes aberrations in length and end-motif frequencies in plasma DNA**. *Proceedings of the National Academy of Sciences of the United States of America* 2019, **116**(2):641–649.
15. Vallejo-Gracia A, Sastre D, Colomer-Molera M, Solé L, Navarro-Pérez M, Capera J, Roig SR, Pedrós-Gámez O, Estadella I, Szilágyi O *et al*: **KCNE4-dependent functional consequences of Kv1.3-related**

- leukocyte physiology.** *Scientific reports* 2021, **11**(1):14632.
16. Wu B, Tao L, Yang D, Li W, Xu H, He Q: **Development of an Immune Infiltration-Related Eight-Gene Prognostic Signature in Colorectal Cancer Microenvironment.** *BioMed research international* 2020, **2020**:2719739.
 17. Lee CM, Barber GP, Casper J, Clawson H, Diekhans M, Gonzalez JN, Hinrichs AS, Lee BT, Nassar LR, Powell CC *et al*: **UCSC Genome Browser enters 20th year.** *Nucleic acids research* 2020, **48**(D1):D756-d761.
 18. Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, Etienne-Grimaldi MC, Schiappa R, Guenot D, Ayadi M *et al*: **Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value.** *PLoS medicine* 2013, **10**(5):e1001453.
 19. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA *et al*: **Inferring tumour purity and stromal and immune cell admixture from expression data.** *Nature communications* 2013, **4**:2612.
 20. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA: **Robust enumeration of cell subsets from tissue expression profiles.** *Nature methods* 2015, **12**(5):453–457.
 21. Yu G, Wang LG, Han Y, He QY: **clusterProfiler: an R package for comparing biological themes among gene clusters.** *Omics: a journal of integrative biology* 2012, **16**(5):284–287.
 22. Kumar L, M EF: **Mfuzz: a software package for soft clustering of microarray data.** *Bioinformatics* 2007, **2**(1):5–7.
 23. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, Varambally S: **UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses.** *Neoplasia (New York, NY)* 2017, **19**(8):649-658.
 24. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z: **GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses.** *Nucleic acids research* 2017, **45**(W1):W98-w102.
 25. Asplund A, Edqvist PH, Schwenk JM, Pontén F: **Antibodies for profiling the human proteome-The Human Protein Atlas as a resource for cancer research.** *Proteomics* 2012, **12**(13):2067–2077.
 26. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, Li B, Liu XS: **TIMER2.0 for analysis of tumor-infiltrating immune cells.** *Nucleic acids research* 2020, **48**(W1):W509-w514.
 27. Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, van de Velde CJ, Watanabe T: **Colorectal cancer.** *Nature reviews Disease primers* 2015, **1**:15065.
 28. Tauriello DV, Calon A, Lonardo E, Batlle E: **Determinants of metastatic competency in colorectal cancer.** *Molecular oncology* 2017, **11**(1):97–119.
 29. Sørensen CG, Karlsson WK, Pommergaard HC, Burcharth J, Rosenberg J: **The diagnostic accuracy of carcinoembryonic antigen to detect colorectal cancer recurrence - A systematic review.** *International journal of surgery (London, England)* 2016, **25**:134–144.
 30. Quail DF, Joyce JA: **Microenvironmental regulation of tumor progression and metastasis.** *Nature medicine* 2013, **19**(11):1423–1437.

31. Malka D, Lièvre A, André T, Taïeb J, Ducreux M, Bibeau F: **Immune scores in colorectal cancer. Where are we?** *European journal of cancer (Oxford, England: 1990)* 2020, **140**:105-118.
32. Wang D, Yu W, Lian J, Wu Q, Liu S, Yang L, Li F, Huang L, Chen X, Zhang Z *et al*: **Th17 cells inhibit CD8(+) T cell migration by systematically downregulating CXCR3 expression via IL-17A/STAT3 in advanced-stage colorectal cancer patients.** *Journal of hematology & oncology* 2020, **13**(1):68.
33. Hume DA: **The Many Alternative Faces of Macrophage Activation.** *Frontiers in immunology* 2015, **6**:370.
34. Pollard JW: **Tumour-educated macrophages promote tumour progression and metastasis.** *Nature reviews Cancer* 2004, **4**(1):71–78.
35. Ge P, Wang W, Li L, Zhang G, Gao Z, Tang Z, Dang X, Wu Y: **Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of colorectal cancer.** *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie* 2019, **118**:109228.
36. Edin S, Wikberg ML, Dahlin AM, Rutegård J, Öberg Å, Oldenborg PA, Palmqvist R: **The distribution of macrophages with a M1 or M2 phenotype in relation to prognosis and the molecular characteristics of colorectal cancer.** *PloS one* 2012, **7**(10):e47045.
37. Gomez-Brouchet A, Illac C, Gilhodes J, Bouvier C, Aubert S, Guinebretiere JM, Marie B, Larousserie F, Entz-Werlé N, de Pinieux G *et al*: **CD163-positive tumor-associated macrophages and CD8-positive cytotoxic lymphocytes are powerful diagnostic markers for the therapeutic stratification of osteosarcoma patients: An immunohistochemical analysis of the biopsies from the French OS2006 phase 3 trial.** *Oncoimmunology* 2017, **6**(9):e1331193.
38. Farhood B, Najafi M, Mortezaee K: **CD8(+) cytotoxic T lymphocytes in cancer immunotherapy: A review.** *Journal of cellular physiology* 2019, **234**(6):8509–8521.
39. Asare-Werehene M, Communal L, Carmona E, Han Y, Song YS, Burger D, Mes-Masson AM, Tsang BK: **Plasma Gelsolin Inhibits CD8(+) T-cell Function and Regulates Glutathione Production to Confer Chemoresistance in Ovarian Cancer.** *Cancer research* 2020, **80**(18):3959–3971.
40. Krishna S, Lowery FJ, Copeland AR, Bahadiroglu E, Mukherjee R, Jia L, Anibal JT, Sachs A, Adebola SO, Gurusamy D *et al*: **Stem-like CD8 T cells mediate response of adoptive cell immunotherapy against human cancer.** *Science (New York, NY)* 2020, **370**(6522):1328–1334.
41. Payandeh Z, Khalili S, Somi MH, Mard-Soltani M, Baghbanzadeh A, Hajiasgharzadeh K, Samadi N, Baradaran B: **PD-1/PD-L1-dependent immune response in colorectal cancer.** *Journal of cellular physiology* 2020, **235**(7-8):5461–5475.
42. Brahmer JR, Tykodi SS, Chow LQ, Hwu WJ, Topalian SL, Hwu P, Drake CG, Camacho LH, Kauh J, Odunsi K *et al*: **Safety and activity of anti-PD-L1 antibody in patients with advanced cancer.** *The New England journal of medicine* 2012, **366**(26):2455–2465.
43. Kanehisa M, Sato Y: **KEGG Mapper for inferring cellular functions from protein sequences.** *Protein science: a publication of the Protein Society* 2020, **29**(1):28–35.
44. Duan Z, Luo Y: **Targeting macrophages in cancer immunotherapy.** *Signal transduction and targeted therapy* 2021, **6**(1):127.

45. Pyonteck SM, Akkari L, Schuhmacher AJ, Bowman RL, Sevenich L, Quail DF, Olson OC, Quick ML, Huse JT, Teijeiro V *et al*: **CSF-1R inhibition alters macrophage polarization and blocks glioma progression.** *Nature medicine* 2013, **19**(10):1264–1272.
46. Pathria P, Louis TL, Varner JA: **Targeting Tumor-Associated Macrophages in Cancer.** *Trends in immunology* 2019, **40**(4):310–327.
47. Beltraminelli T, De Palma M: **Biology and therapeutic targeting of tumour-associated macrophages.** *The Journal of pathology* 2020, **250**(5):573–592.
48. McLane LM, Abdel-Hakeem MS, Wherry EJ: **CD8 T Cell Exhaustion During Chronic Viral Infection and Cancer.** *Annual review of immunology* 2019, **37**:457–495.
49. Garber K: **Driving T-cell immunotherapy to solid tumors.** *Nature biotechnology* 2018, **36**(3):215–219.
50. Kataoka S, Manandhar P, Lee J, Workman CJ, Banerjee H, Szymczak-Workman AL, Kvorjak M, Lohmueller J, Kane LP: **The costimulatory activity of Tim-3 requires Akt and MAPK signaling and its recruitment to the immune synapse.** *Science signaling* 2021, **14**(687).
51. Das M, Zhu C, Kuchroo VK: **Tim-3 and its role in regulating anti-tumor immunity.** *Immunological reviews* 2017, **276**(1):97–111.
52. Sisirak V, Sally B, D'Agati V, Martinez-Ortiz W, Özçakar ZB, David J, Rashidfarrokhi A, Yeste A, Panea C, Chida AS *et al*: **Digestion of Chromatin in Apoptotic Cell Microparticles Prevents Autoimmunity.** *Cell* 2016, **166**(1):88–101.
53. Shiokawa D, Shika Y, Tanuma S: **Identification of two functional nuclear localization signals in DNase gamma and their roles in its apoptotic DNase activity.** *The Biochemical journal* 2003, **376**(Pt 2):377–381.
54. Guo D, Ma D, Liu P, Lan J, Liu Z, Liu Q: **DNASE1L3 arrests tumor angiogenesis by impairing the senescence-associated secretory phenotype in response to stress.** *Aging* 2021, **13**(7):9874–9899.
55. Zhao Q, Yang C, Wang J, Li Y, Yang P: **Serum level of DNase1l3 in patients with dermatomyositis/polymyositis, systemic lupus erythematosus and rheumatoid arthritis, and its association with disease activity.** *Clinical and experimental medicine* 2017, **17**(4):459–465.
56. Orvain C, Assassi S, Avouac J, Allanore Y: **Systemic sclerosis pathogenesis: contribution of recent advances in genetics.** *Current opinion in rheumatology* 2020, **32**(6):505–514.
57. Xu B, Lv W, Li X, Zhang L, Lin J: **Prognostic genes of hepatocellular carcinoma based on gene coexpression network analysis.** *Journal of cellular biochemistry* 2019.
58. Wang S, Ma H, Li X, Mo X, Zhang H, Yang L, Deng Y, Yan Y, Yang G, Liu X *et al*: **DNASE1L3 as an indicator of favorable survival in hepatocellular carcinoma patients following resection.** *Aging* 2020, **12**(2):1171–1185.
59. Roig SR, Solé L, Cassinelli S, Colomer-Molera M, Sastre D, Serrano-Novillo C, Serrano-Albarrás A, Lillo MP, Tamkun MM, Felipe A: **Calmodulin-dependent KCNE4 dimerization controls membrane targeting.** *Scientific reports* 2021, **11**(1):14046.

60. Solé L, Roig SR, Sastre D, Vallejo-Gracia A, Serrano-Albarrás A, Ferrer-Montiel A, Fernández-Ballester G, Tamkun MM, Felipe A: **The calmodulin-binding tetraleucine motif of KCNE4 is responsible for association with Kv1.3.** *FASEB journal: official publication of the Federation of American Societies for Experimental Biology* 2019, **33**(7):8263-8279.
61. Sun L, Hui AM, Su Q, Vortmeyer A, Kotliarov Y, Pastorino S, Passaniti A, Menon J, Walling J, Bailey R *et al*: **Neuronal and glioma-derived stem cell factor induces angiogenesis within the brain.** *Cancer cell* 2006, **9**(4):287–300.
62. Cheong A, Bingham AJ, Li J, Kumar B, Sukumar P, Munsch C, Buckley NJ, Neylon CB, Porter KE, Beech DJ *et al*: **Downregulated REST transcription factor is a switch enabling critical potassium channel expression and cell proliferation.** *Molecular cell* 2005, **20**(1):45–52.
63. Solé L, Vallejo-Gracia A, Roig SR, Serrano-Albarrás A, Marruecos L, Manils J, Gómez D, Soler C, Felipe A: **KCNE gene expression is dependent on the proliferation and mode of activation of leukocytes.** *Channels (Austin, Tex)* 2013, **7**(2):85–96.

Figures

Figure 2

Kaplan-Meier survival curves. (2a-c) There was a significant correlation between immune score and OS of CRC patients ($p=0.039$), with a favorable prognosis in the high immune score group compared to the low immune score group.

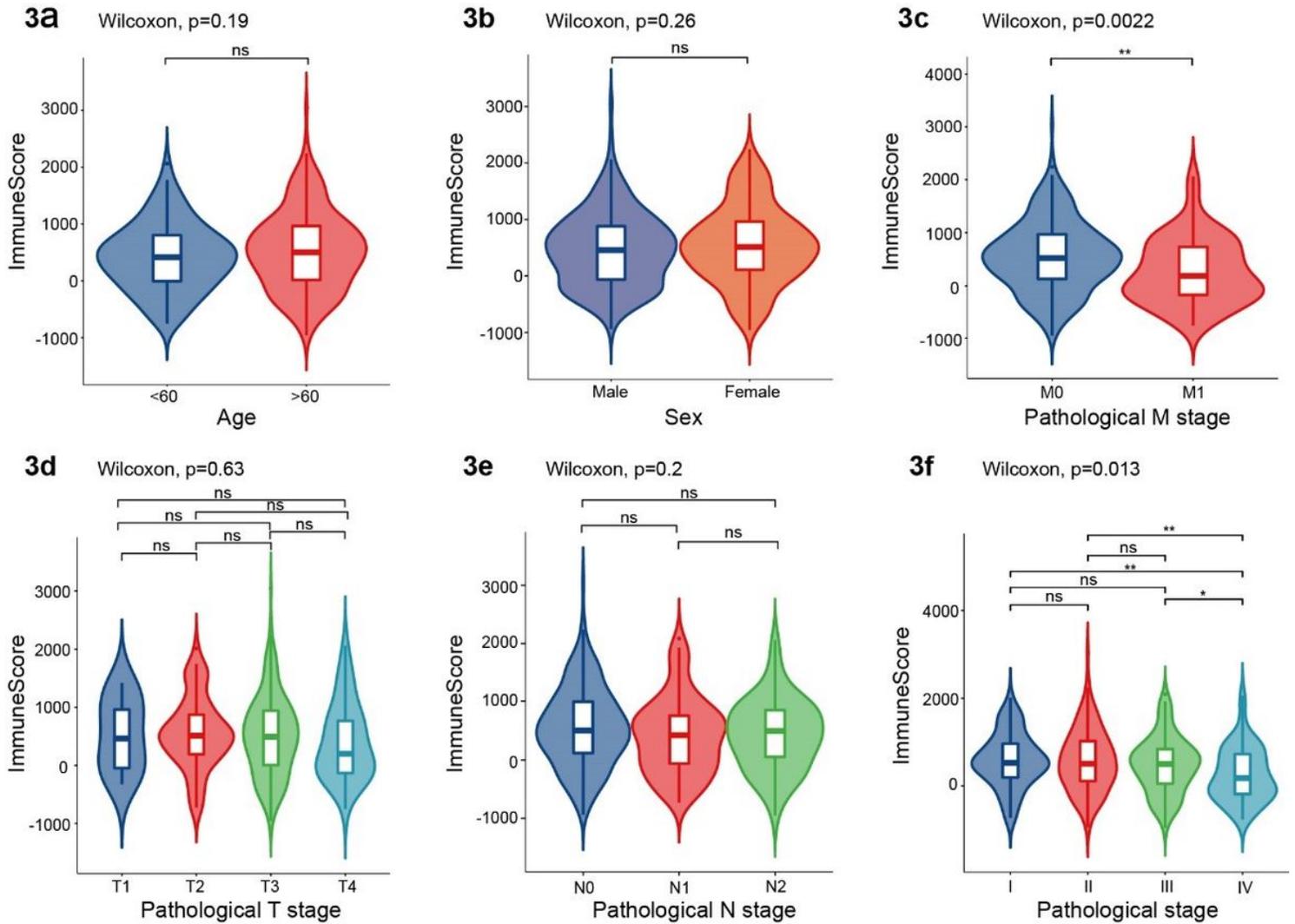


Figure 3

Violin plot of the correlation between immune scores and clinical characteristics of CRC patients. Horizontal and vertical axes represent clinical characteristics and immune scores. The immune scores were significantly correlated with the progression of pathological stages and pathologic M stage.

* $p<0.05$, ** $p<0.01$, ns, no significance.

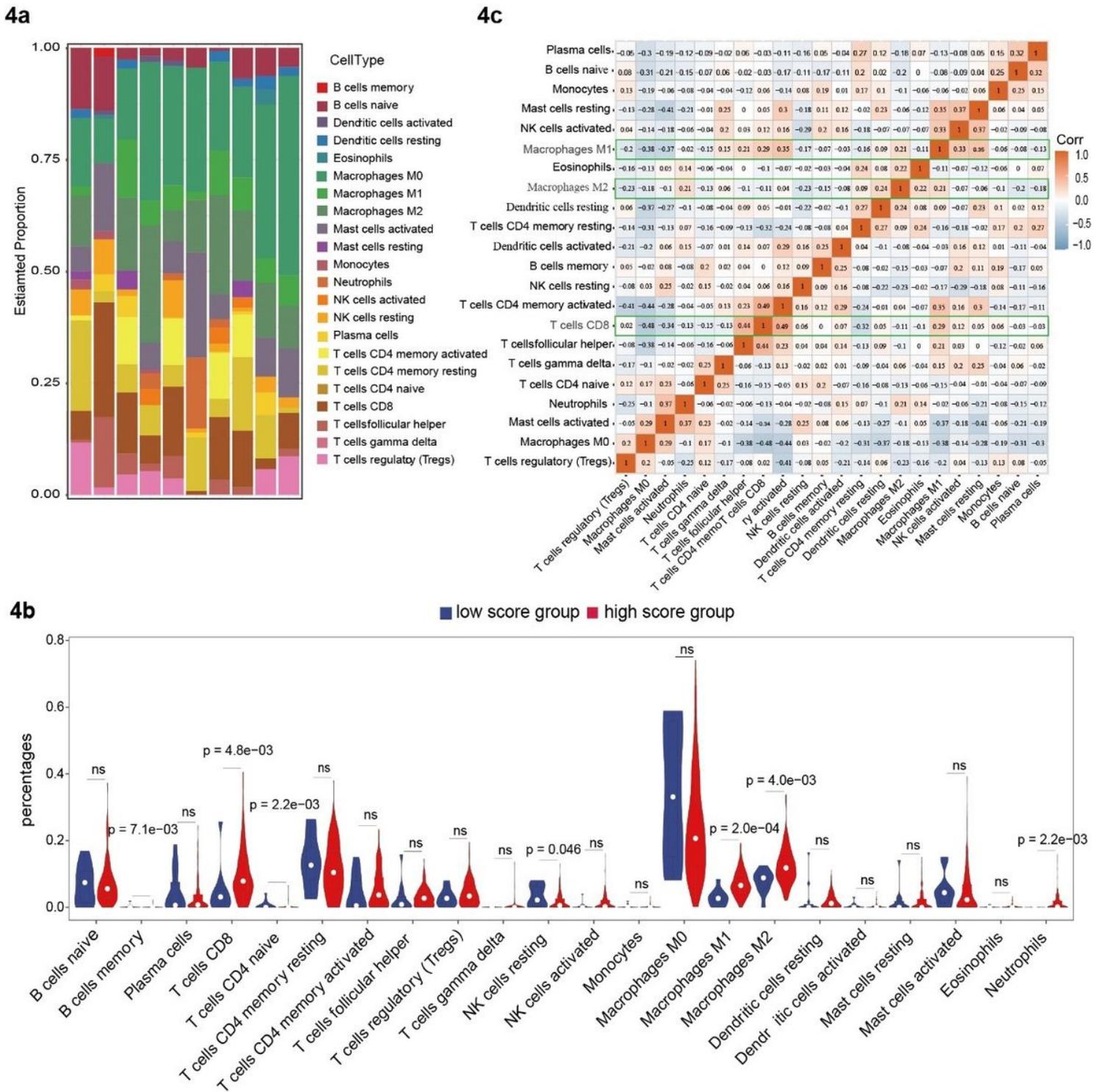


Figure 4

Composition of TIICs between the high and low immune score groups in the TCGA cohort. (4a) Fraction of immune cells in the former 10 CRC patients. (4b) Comparisons of immune cells between high immune score group and low immune score group. The high immune score group was shown in red, while blue represents the low immune score group, horizontal and vertical axes represent 22 immune cells and the percentage of cells infiltration, ns, for not statistically significant. (4c) Correlation matrix of all 22 TIICs proportions, the progressively deeper red and blue colors represent positive and negative correlations

respectively. The green boxes represent the correlation between the three immune cells (CD8+ T, M1 macrophages, M2 macrophages) and 22 immune cells.

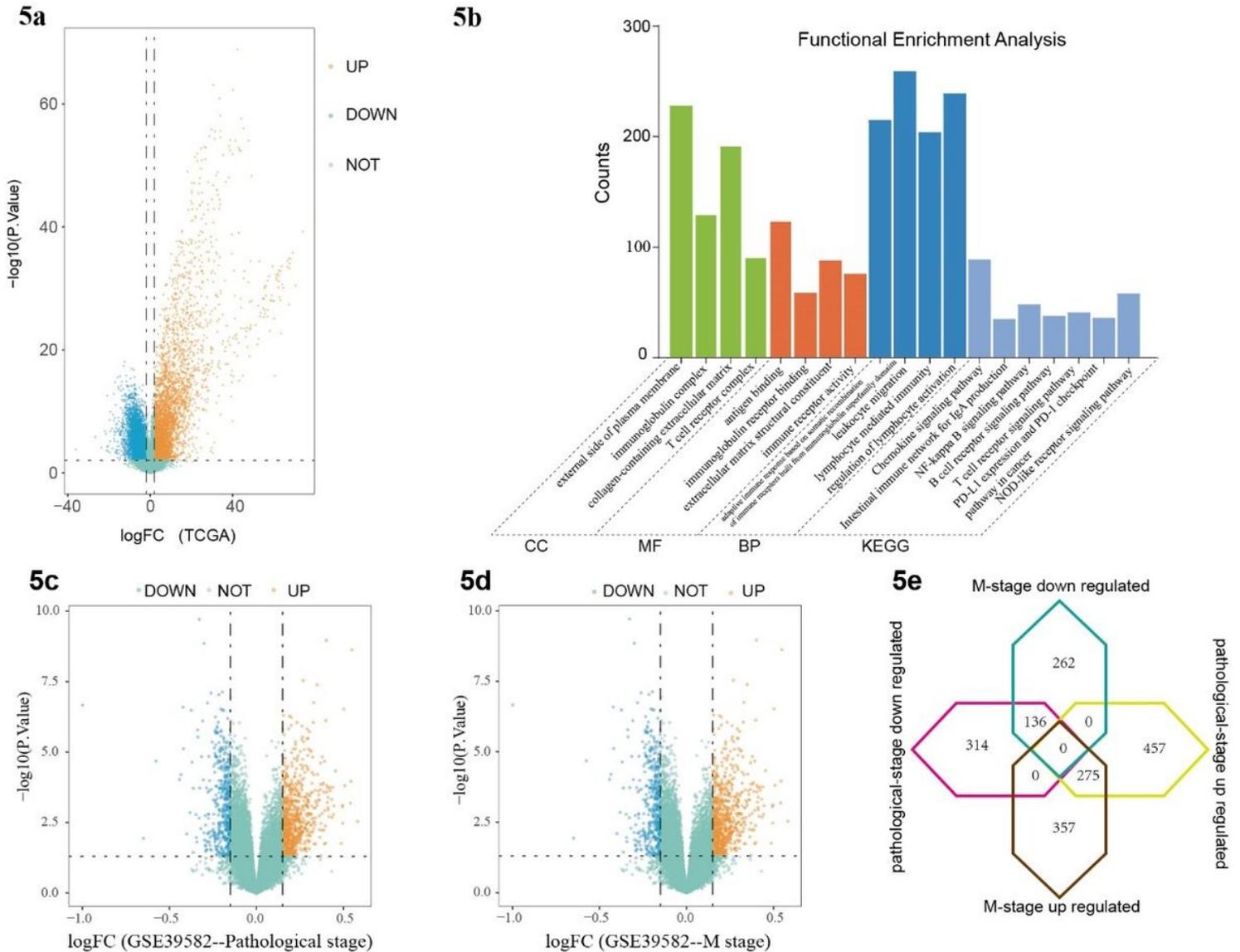


Figure 5

Identification of DEGs and results of functional enrichment analysis. (5a) Volcano plot showing up- (orange) and down-regulated (blue) immune genes in the high immune score group and low immune score group. (5b) GO and KEGG analysis of positively regulated DEGs in the high immune score group. Horizontal axis represents the results of the partial enrichment analysis, vertical axis represents the counts of genes enriched, the color of the bar chart represents the different enrichment categories. (5c-5d) Volcano map displaying differential up- (orange) or down-regulated (blue) genes across pathologic stage (pathological stage and pathologic M stage) in patients from the GSE39582 dataset. (5e) Intersection chart of up- or down-regulated DEGs in the progression of different pathologic stage (pathological stage and pathologic M stage).

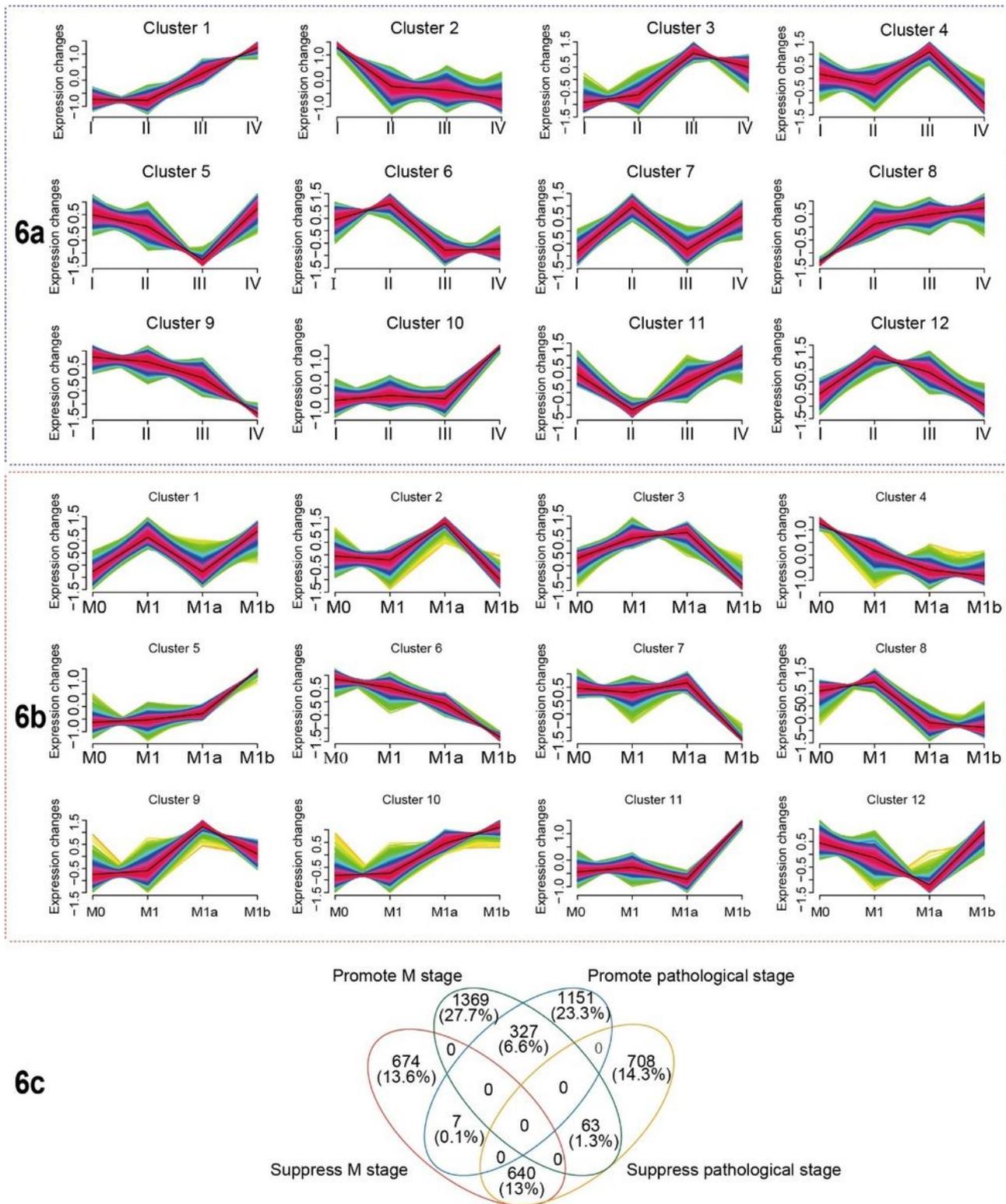


Figure 6

Results of soft clustering analysis. The horizontal axis shows the pathological staging and the vertical axis presents the normalized intensity ratio of the gene expression values. (6a) The genes in clusters 1 and 8 showed a persistent increasing trend as the pathological stage progressed, and clusters 2 and 9 showed a persistent decreasing trend. Genes with the same trend were combined as genes that promote or suppress pathological stage. (6b) The genes in clusters 5 and 10 showed a persistent increasing trend

as the pathological stage progressed, and clusters 4 and 6 showed a persistent decreasing trend. Genes with the same trend were combined as genes that promote or suppress pathological M stage. (6c) Venn diagram displayed results of intersection of genes that promote or suppress CRC progression.

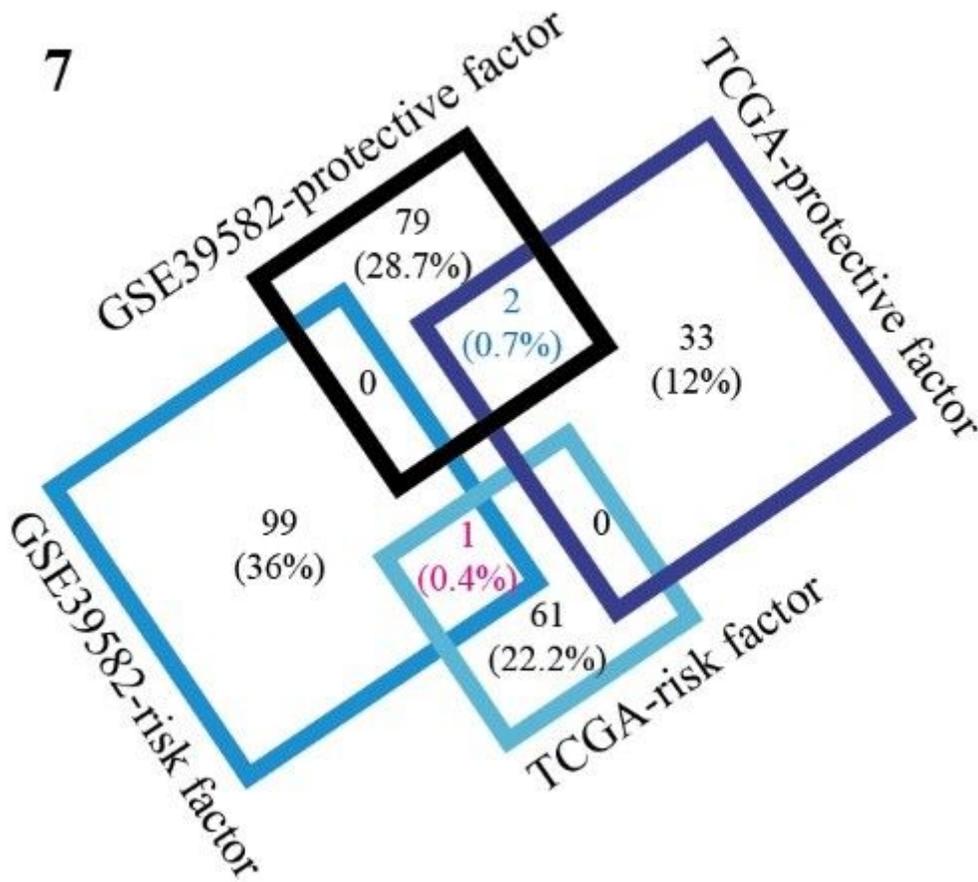


Figure 7

Intersection gene chart. The intersection figure showed the protective and risk factors that were observed in both the TCGA cohort and the GSE39582 dataset. containing one risk factors (red) and two protective factors (blue).

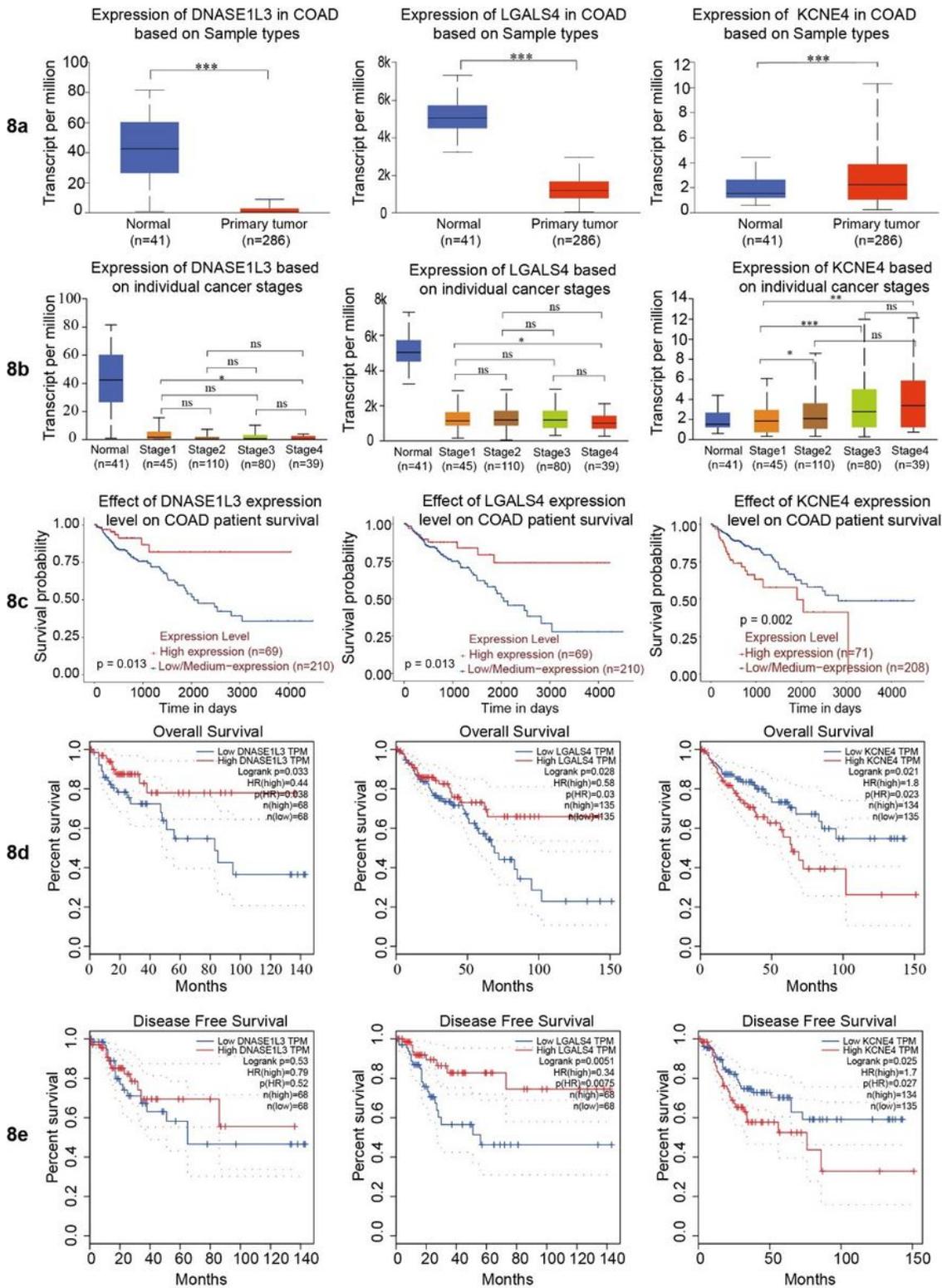


Figure 8

Expression analysis graphs and survival curves of COAD. (8a) Graph of differential expression analysis of three candidate genes in the UALCAN database in normal and CRC tissues, blue for normal tissues and red for tumor tissues. (8b) Expression analysis of three candidate genes in the UALCAN database in the progression of CRC pathological stage. (8c) In the UALCAN database, the highly expressed LGALS4 ($p=0.013$) and DNASE1L3 ($p=0.013$) had good OS and the highly expressed KCNE4 ($p=0.002$) had poor

OS. (8d-e) In the GEPIA database, the highly expressed LGALS4 had a favorable DFS (HR=0.34, p=0.0051) and OS (HR=0.58, p=0.0028), and the highly expressed KCNE4 had a poor OS [KCNE4: HR=1.8, p=0.021] and DFS (KCNE4: HR=1.7, p=0.025).

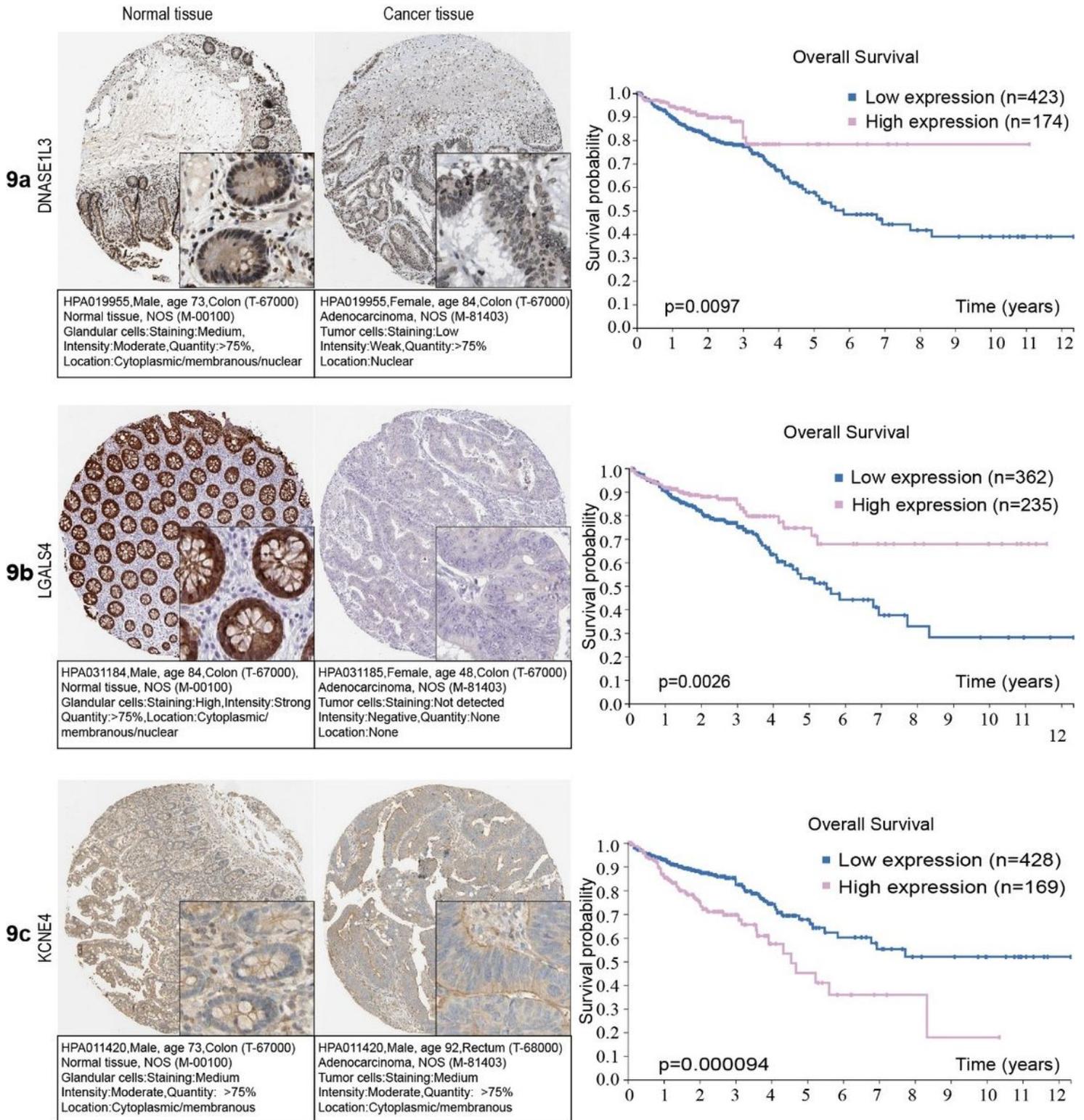


Figure 9

Immunohistochemistry of candidate genes and OS analysis. (9a-9b) The associated proteins of DNASE1L3, LGALS4 were moderate to highly expressed in normal tissues and were not detected in CRC

tissues. Patients with high expression of these genes had a higher survival rate. (9d) KCNE4 showed moderate expression in both normal tissue and CRC tissues. Patients with high expression of KCNF4 had a lower survival rate.

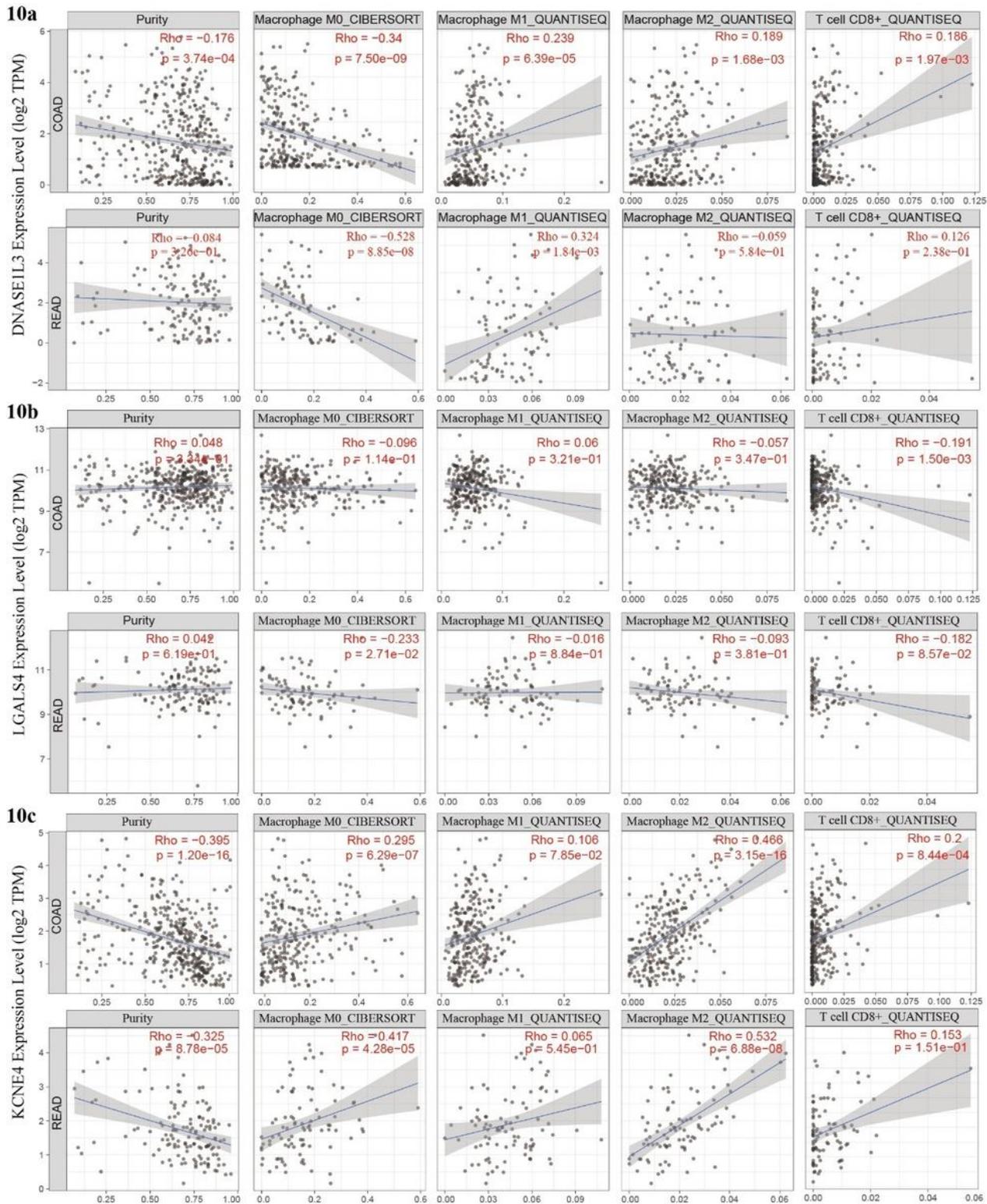


Figure 10

Correlation between the candidate genes with the infiltration of CD8+ T cells and macrophages. (10a) The expression of DNASE1L3 in COAD and READ patients showed a significant positive correlation with

the infiltration of M1 macrophages and a negative correlation with the infiltration of M0 macrophages. (10b) The expression of LGALS4 in COAD and READ patients have no significant correlation with the infiltration of macrophages and CD8+ T cells. (10c) The expression of DNASE1L3 in COAD and READ patients showed a significant positive correlation with the infiltration of M2 macrophages and M0 macrophages.

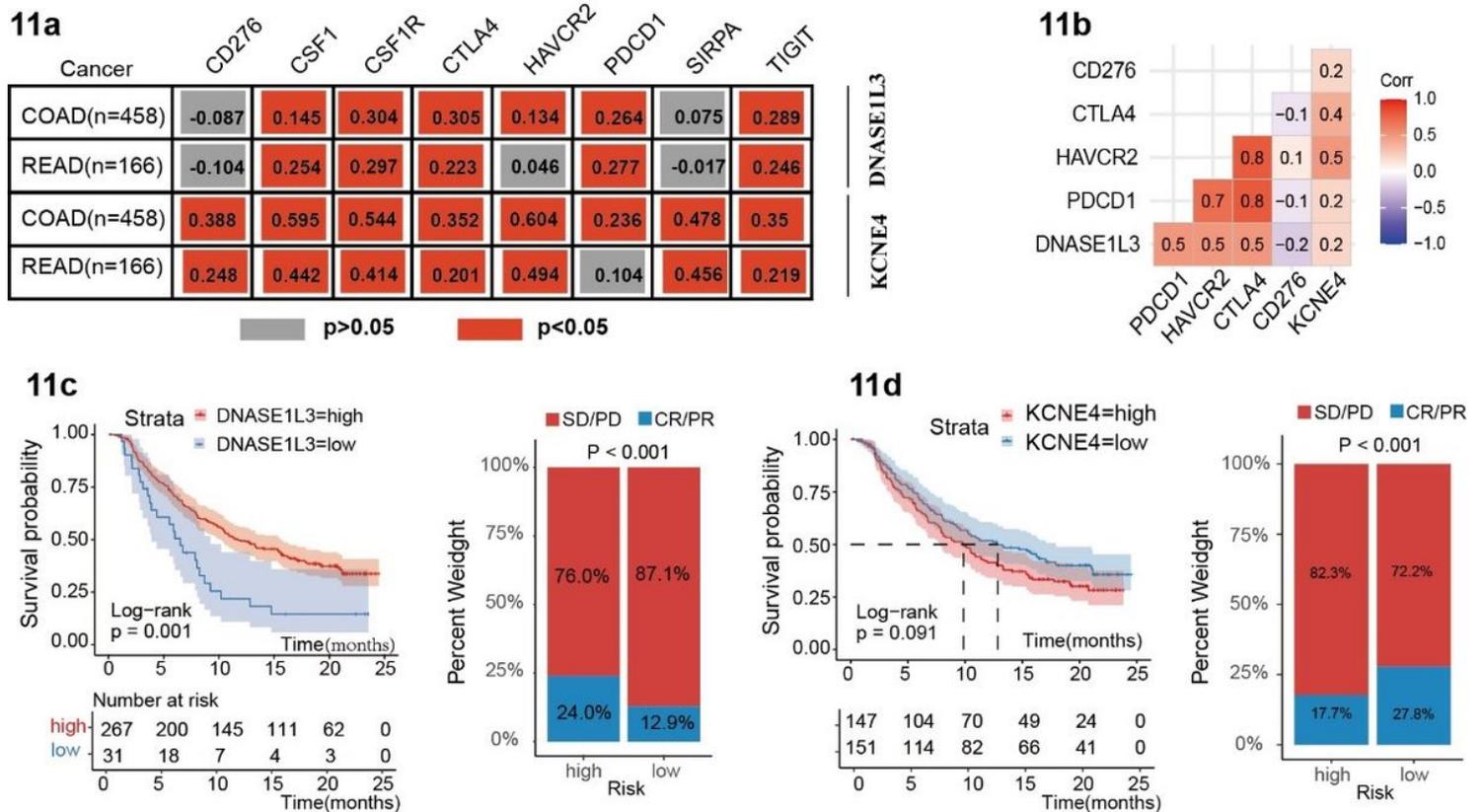


Figure 11

Correlation between candidate genes and immune checkpoint markers and the role of candidate genes in the immunotherapeutic response. (11a) The correlation between the expression of candidate genes and CSF1, CSF1R, SIRPA, PD-1 (PDCD1), PD-L1(CD274), CTLA4 and TIM-3 (HAVCR2) in COAD and READ cohorts. (11b) The correlation between the expression of candidate genes and PD-1 (PDCD1), PD-L1(CD274), CTLA4 and TIM-3 (HAVCR2) in the IMVigor210 cohort. (11c) Kaplan-Meier curves of patients with high or low DNASE1L3 and rate of clinical benefits (CR/PR and SD/PD) of anti-PD-1 therapy in the IMVigor210 cohort. (11d) Kaplan-Meier curves of patients with high or low KCNE4 and rate of clinical benefits (CR/PR and SD/PD) of anti-PD-1 therapy in the IMVigor210 cohort.

Supplementary Files

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

- [SupplementalTableS16.xls](#)