Am J Cancer Res 2022;12(12):5532-5551
www.ajcr.us /ISSN:2156-6976/ajcr0146097

Original Article

Risk stratification of patients with right-sided colorectal cancer based on the tumor-infiltrating M1 macrophage

Dong Min Lim[1*], Junho Kang[2*], Soo-Yeon Woo[3], Hee-Sun Choi[3], Munju Kwon[3], Jayoung Kim[3], Hae Ryoun Park[6], Keehoon Jung[7], Ninib Baryawno[8#], Hyung-Sik Kim[9,10#], Dongjun Lee[3#], Yun Hak Kim[4,5,10#]

[1]Interdisciplinary Program of Genomic Data Science, Pusan National University, Yangsan, Republic of Korea; [2]Medical Research Institute, Pusan National University, Busan, Republic of Korea; Departments of [3]Convergence Medicine, [4]Anatomy, [5]Biomedical Informatics, School of Medicine, Pusan National University, Yangsan, Republic of Korea; [6]Department of Oral Pathology, School of Dentistry, Pusan National University, Yangsan, Republic of Korea; [7]Department of Anatomy and Cell Biology and Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Republic of Korea; [8]Childhood Cancer Research Unit, Department of Women's and Children's Health, Karolinska Institutet, Stockholm 17177, Sweden; [9]Department of Oral Biochemistry, [10]Dental and Life Science Institute, School of Dentistry, Pusan National University, Yangsan 50612, Republic of Korea. [*]Equal contributors. [#]Co-senior authors.

Received August 23, 2022; Accepted November 16, 2022; Epub December 15, 2022; Published December 30, 2022

Abstract: The homing of M1 and M2 macrophages may play distinct roles in the tumor microenvironment (TME). However, these roles of macrophages in the TME remain unclear. We downloaded RNA sequencing data from The Cancer Genome Atlas (TCGA) database for patients with CRC. Subsequently, Kaplan-Meier survival curves were generated to assess the differential infiltration of M1 and M2 macrophages based on CRC location. Differentially expressed gene (DEG) and functional analyses were performed to screen the roles of DEGs. Critical prognostic genes were identified using least absolute shrinkage and selection operator regression. The risk scores were calculated for each patient. In patients with right-sided CRC, reduced M1 macrophage infiltration was associated with poor prognosis. M1 macrophage infiltration positively correlated with CD8+ T cell infiltration. A risk model was developed and validated for performance using GSE103479 and GSE72970. Nine genes were identified as independent prognostic genes that could be potential biomarkers for effectively predicting survival in patients with right-sided CRC. Kaplan-Meier curves for overall survival and progression-free survival analyses revealed that the high-risk group of patients with right-sided CRC had a poor prognosis. This novel M1 macrophage-related risk model may provide a gene signature for predicting the survival outcomes of patients with right-sided CRC and facilitate further studies examining the relationship between infiltration of M1 macrophages and the prognosis of such patients.

Keywords: Colorectal cancer, colorectal cancer location, M1 macrophage, tumor microenvironment, LASSO regression

Introduction

Globally, colorectal cancer (CRC) is the second and third most common cancer among women and men, respectively [1], accounting for approximately 10% of cancer-related deaths [2]. By 2035, the worldwide incidence of CRC is estimated to increase to 2.5 million new cases [3]. The risk of CRC is reportedly enhanced by negative lifestyle choices, including smoking, excessive alcohol intake, obesity, and consumption of red and processed meat [3, 4]. Distant metastasis primarily occurs in the liver and is a major cause of death in patients with CRC [5]. Adenomatous polyposis coli (*APC*), *KRAS*, and *p53* are the predominantly mutated genes in CRC. Furthermore, activation of several pathways downstream of the epidermal growth factor receptor, including the RAS/MAPK, PI3K/AKT, PLC, signal transducer and activator of transcription (STAT), and SRC/FAK pathways, can affect tumor proliferation, angiogenesis, and cell survival [6-9].

CRC differs in incidence, pathogenesis, molecular pathways, and prognosis depending on the

primary tumor location as a crucial prognostic factor. Recently, clinical and biological differences between right- and left-sided CRC have been extensively discussed [10-13]. In right-sided CRC, mutations in the DNA mismatch repair (MMR) system, *BRAF* mutations, and an overall worse prognosis have been typically documented; despite the high immune activity, the prognosis of right-sided CRC remains poor. In left-sided CRC, chromosomal instability pathway-related mutations, including *APC*, *PIK3CA*, and *p53* mutations, mainly occur, and the overall prognosis of patients is better than that of patients with right-sided CRC. Differences between right- and left-sided CRC, along with differences in molecular pathways, have been noted in the tumor stage at diagnosis. Patients with right-sided CRC are diagnosed at a more advanced tumor stage than patients with left-sided CRC owing to the flat morphology of the colon as determined by colonoscopic screening [14, 15]; in addition, changes in bowel habits are more evident in left-sided CRC [12]. Therefore, it is essential to predict the survival outcomes of patients with right-sided CRC.

The tumor microenvironment (TME) comprises various immune cells, stromal cells, extracellular matrix molecules, and cytokines [16]. CRC exhibits heterogeneity in its pathogenic mechanisms, corresponding to diverse types and subtypes of tumor cell phenotypes, therapeutic outcomes, and prognosis. Although the role of TME cells in CRC is being explored, it remains unclear [17]. Currently, there exists a broad consensus that anatomical sites are crucial factors in CRC [18]. However, the difference between right- and left-sided TME in CRC remains poorly understood. Therefore, a comparative analysis of the TME according to CRC location is required. The objective of the present study was to determine the association between the abundance of immune cell infiltration and patient survival according to the CRC location, and to identify a prognostic gene signature.

Material and methods

Data acquisition and processing

Gene expression data (HTSeq-FPKM and HTSeq-Counts) and clinical data were downloaded from TCGA-COAD (colorectal adenocarcinoma) and TCGA-READ (rectal adenocarcinoma) databases. Data are available on the Genomic Data Commons data portal (https://portal.gdc.cancer.gov). We filtered the data by "primary tumor". Data regarding overall survival (OS) and progression-free survival (PFS) were downloaded from the UCSC Xena Browser (http://xena.ucsc.edu). We selected GSE103-479 [19] and GSE72970 [20] from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov) as validation sets. Data pre-processing was performed using R software (version 4.1.1) [21].

Quantification of tumor-infiltrating immune cells

We estimated the proportions of infiltrating immune cell types in each sample using bulk RNA sequencing (HTSeq-FPKM) data, performed using a quanTIseq deconvolution algorithm in the R package "quantiseqr". QuanTIseq quantifies fractions of immune cell types from bulk RNA sequencing data. QuanTIseq is designed as an analysis pipeline optimized for pre-processing raw RNA-seq data, quantifying and normalizing gene expression, re-annotating genes, and estimating cell fractions. Unlike previous deconvolution algorithm, quanTIseq is specifically designed for RNA-seq data, which is the current advanced technology for quantification of high-throughput data. The results of extensive validation showed that quanTIseq can faithfully and quantitatively infer immune cell fractions from bulk RNA-seq data [22].

Differentially expressed gene (DEG) analysis

DEG analysis was performed between high levels of M1 macrophages and low levels of M1 macrophages in patients with right-sided CRC, and HTSeq counts of gene expression data were used. We used the edgeR R package (v3.34.1) to identify DEGs with an adjusted p-value < 0.05 and |logFC| > 1 as screening conditions [23]. The adjusted p-value was calculated using the Benjamini-Hochberg method. We filtered out unexpressed and low counts using the edgeR function filterByExpr, with minimum count required for at least some samples =10, minimum total count required =15, and minimum proportion of samples in the smallest group that expressed the gene =0.7. Next, trimmed mean of the M-values (TMM) normalization was performed using the edgeR package. TMM-normalized data were statistically

processed using the edgeR function, glmQLFTest. The quasi-likelihood test, which computes data by fitting a quasi-likelihood negative binomial generalized log-linear model, performed a gene-specific statistical test for a given coefficient or contrast. The EnhancedVolcano (v1.10.0) package in R was used to represent DEGs by plotting volcanoes [24].

Protein-protein interaction (PPI) network construction and module selection

To identify relevant pathways and functions of DEGs, PPI networks were constructed using the Search Tool for Interaction Gene Search (STRING) database (http://www.string-db.org) [25] and Cytoscape software (Version 3.8.2) [26]. In the PPI network of DEGs, the minimum required interaction score of ≥ 0.4 was set as the screening condition. The Molecular Complex Detection (MCODE) plugin was utilized for the most significant module selection [27]. The cutoff values of MCODE were set as follows: degree cutoff =2, cluster finding = haircut, node score cutoff =0.2, K-score =2, and maximum depth =100.

Functional analysis of DEGs

Relevant DEG pathways were further explored by gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses. The ClusterProfiler (v4.0.5) package in R was used for GO and KEGG enrichment analyses [28]. Adjusted *p*-values were calculated using the Benjamini-Hochberg method. Statistical significance was defined as a *p*-value of < 0.05. GO, and KEGG enrichment analyses were performed using the enrichment plot (v1.12.3) R package [29].

Least absolute shrinkage and selection operator (LASSO) regression and development of M1 macrophages-related risk score

LASSO regression analysis was performed using the glmnet (v4.1-3) R package to obtain significant prognostic genes and establish an M1 macrophage-related risk model. The LASSO regression is a one of penalty regression analysis, and has the advantage of creating a simpler regression model by penalizing variables that do not significantly contribute to the performance of the model. We performed tenfold cross-validation (CV) to improve the model accuracy and avoid overfitting the model to the training data, only the most important predictors are selected. Optimal lambda values as regularization tuning parameter were acquired from the mean squared error and selected representative prognostic genes. Thus, model parsimony could optimize model accuracy [30, 31]. Risk scores were calculated based on LASSO regression coefficients, and "High-risk" and "Low-risk" groups were classified according to the median risk score. The formula is as follows:

$$\textit{Risk score} = \sum[\textit{gene expression}] \times [\textit{regression coefficient}]$$

To assess the predictive ability of the M1 macrophage-related risk model, the OS and PFS of the risk group were analyzed using Kaplan-Meier curves, and receiver operating characteristic (ROC) curve analysis was performed to evaluate the specificity and sensitivity of the risk model. The pROC (v 1.18.0) package in R was used to display and analyze ROC curves [32].

Validation of the prognostic risk model

To evaluate the prognostic performance of the M1 macrophage-related risk model, two GEO datasets (GSE103479 [19] and GSE72970 [20]) were selected to further validate the established model. LASSO regression analysis was performed to determine the optimal lambda values with tenfold CV. Risk scores were calculated based on LASSO regression coefficients and stratified into high- or low-risk groups according to the median value. Kaplan-Meier and ROC curves were analyzed using OS and PFS in patients with right-sided CRC to validate the predictive ability of the M1 macrophage-related risk model.

Statistical analysis

Box plots were plotted using the ggpubr (v0.4.0) function, ggboxplot [33]. The Wilcoxon signed-rank test was used to compare differences in the mean values of genes according to M1 macrophage infiltration. Statistical significance was set at $P < 0.05$. The five-year OS and PFS rates were calculated using Kaplan-Meier survival analysis, and log-rank *p*-values were calculated. All statistical analyses and visualizations were performed using the R software (version 4.1.1) [21].

Table 1. Characteristics of patients with right- and left-sided CRC (TCGA dataset)

Right-sided CRC			Left-sided CRC		
Variable	N	Mean (SD*)	Variable	N	Mean (SD*)
Age (years)	134	67.5 (12.9)	Age (years)	131	62.14 (12.7)
	N	Percentage (%)		N	Percentage (%)
Sex	134		Sex	131	
Female	55	41	Female	64	48.9
Male	79	59	Male	67	51.1
Race	125		Race	121	
White	93	74.4	White	101	83.5
Asian	4	3.2	Asian	1	0.8
Black	28	22.4	American indian	1	0.8
Primary Disease Stage	128		Black	18	14.9
I-II	76	54.5	Primary Disease Stage	125	
III-IV	52	45.5	I-II	62	49.6
Sample type	134		III-IV	63	50.4
Primary	134	100	Sample type	131	
Origin	134		Primary	131	100
Ascending colon	59	44	Origin	131	
Hepatic flexure of colon	11	8.2	Descending colon	14	10.7
Cecum	64	47.8	Rectosigmoid junction	47	35.9
Death	134		Sigmoid colon	65	49.6
No	95	70.9	Splenic flexure of colon	5	3.8
Yes	39	29.1	Death	131	
			No	102	77.9
			Yes	29	22.1

CRC, Colorectal Cancer; *SD, Standard Deviation.

Results

Patient characteristics

We included 265 patients from The Cancer Genome Atlas (TCGA) project [34, 35]: 134 patients with right-sided CRC (ascending colon, hepatic flexure of the colon, and cecum) and 131 with left-sided CRC (descending colon, rectosigmoid junction, sigmoid colon, and splenic flexure of colon). The patient information used in this study is listed in **Table 1**. We downloaded two validation datasets (GSE-103479 [19] and GSE72970 [20]). Each cohort included tumor origin and survival information of patients with CRC. Data on patients with right-sided CRC was used to validate the results. **Table 2** lists patient information for the validation sets used in the present study.

Tumor-infiltrating immune cells and their prognostic significance in patients with CRC

We used the quanTIseq algorithm [22] to analyze bulk RNA-seq data and identify infiltrated immune cells in the TME. The results of the quanTIseq algorithm are shown in Supplementary Figure 1. Binarizing by median fractions, we grouped patients with CRC into two groups, “High” and “Low”. The Kaplan-Meier survival method was used to examine the association between immune cell infiltration and OS. Kaplan-Meier survival curves for OS analyses revealed that reduced M1 macrophages from the right side of the colon was associated with poor prognosis in patients with right-sided CRC (**Figure 1**, log-rank *P=0.02*). Conversely, infiltration of M2 macrophages in patients with right-sided CRC and M1 and M2 macrophage infiltration in patients with left-sided CRC showed no significant correlation with OS (**Figure 1**). Kaplan-Meier survival analysis of the other infiltrating immune cells in patients with right- (Supplementary Figure 2) and left-sided CRC (Supplementary Figure 3) revealed no significant correlation with OS. As shown in Supplementary Figure 2, infiltration of dendritic cells in patients with right-sided CRC showed significant correlation with OS.

Table 2. Characteristics of patients in validation cohorts (GSE103479 and GSE72970)

GSE103479			GSE72970		
Variable	N	Mean (SD*)	Variable	N	Mean (SD*)
Age (years)	154	69.7 (11.4)	Age (years)	124	61.7 (11.4)
	N	Percentage (%)		N	Percentage (%)
Sex	156		Sex	124	
Female	68	43.6	Female	50	40.3
Male	88	56.4	Male	74	59.7
Contact country	156		Contact country	124	
United Kingdom	156	100	France	124	100
Primary Disease Stage	156		Pathological N stage	124	
II	84	53.8	pNX-pN0	43	34.7
III	72	46.2	pN1-pN2	81	65.3
Sample type	156		Pathological T stage	124	
Primary	156	100	pTx-pT2	37	29.8
Origin	150		pT3-pT4	87	70.2
Left colon	77	51.4	Sample type	124	
Right colon	59	39.3	Primary	124	100
Transverse colon	14	9.3	Origin	124	
Death	155		Left colon	88	71.0
No	95	61.3	Right colon	34	27.4
Yes	60	38.7	Transverse colon	2	1.6
			Death	124	
			No	32	25.8
			Yes	92	74.2

*SD, Standard Deviation.

However, most of the fraction values of dendritic cells in patients with right-sided CRC were 0, which was excluded from the further evaluation.

Molecular differences between high and low levels of M1 macrophages in patients with right-sided CRC

Based on the aforementioned screening conditions, we identified 619 DEGs (405 upregulated genes and 214 downregulated genes in high levels of M1 macrophages) between high and low levels of M1 macrophages in patients with right-sided CRC (Supplementary Table 1). Volcano plots of DEGs are shown in **Figure 2A**. To identify PPIs of DEGs, the STRING network-based protein interaction assay using Cytoscape software was used to create a PPI network. The interaction between DEGs was confirmed by constructing a PPI network. The upregulated PPI network consisted of 293 nodes and 4,119 edges (**Figure 2B**). Three significant modules of upregulated genes were identified using MCODE. Among the PPI network of upregulated genes, the module with the highest MCODE contained 33 genes (*CCL4*, *CCR5*, *CD209*, *CD274*, *CD38*, *CD4*, *CD80*, *CD86*, *CD8A*, *CLEC7A*, *CSF2*, *CTLA4*, *CXCL11*, *CXCL13*, *CYBB*, *FASLG*, *FCGR2A*, *FCGR3A*, *FCGR3B*, *HAVCR2*, *IFNG*, *IL10*, *IL1B*, *IL2RA*, *IL2RB*, *IL6*, *ITGAM*, *ITGAX*, *PDCD1*, *PTPRC*, *SIGLEC1*, *TBX21*, and *TNF*; **Figure 2C**). The second highest module comprised 16 upregulated genes (*CCR2*, *CD163*, *CSF3*, *CXCL1*, *CXCL10*, *CXCL9*, *FCGR1A*, *FCGR2B*, *ICAM1*, *IL10RA*, *IL1A*, *MRC1*, *TLR1*, *TLR2*, *TLR7*, and *TLR8*; **Figure 2D**). The third highest module comprised upregulated human leukocyte antigen (HLA)-related genes (*HLA-DMB*, *HLA-DOA*, *HLA-DMA*, *HLA-DPB1*, *HLA-DPA1*, *HLA-DQA1*, *HLA-DRB5*, *HLA-DQB2*, *HLA-DQA2* and *HLA-DQB1*; **Figure 2E**). The PPI network of downregulated genes consisted of 126 nodes and 204 edges, and the largest module consisted of *APOH*, *CALCA*, *CHGB*, *CPE*, *ELAVL4*, *F7*, *GAP43*, *KIF1A*, *MAPT*, *NEFL*, *PCSK1*, *SCT*, *SERPINA7*, *TAC1*, and *TTR* (**Figure 2F**, **2G**).

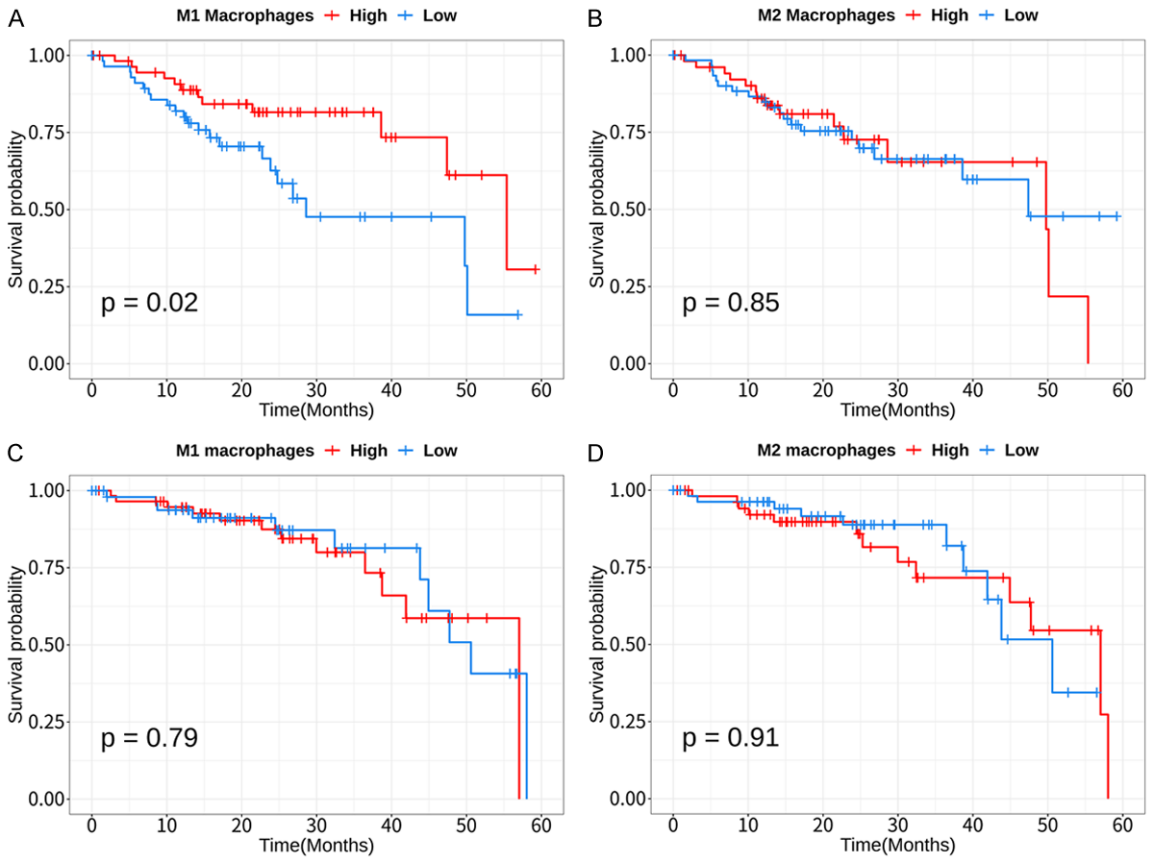


Figure 1. Kaplan-Meier curves for overall survival analysis. Survival curves of patients with right-sided CRC according to expression levels of tumor-infiltrating (A) M1 macrophages and (B) M2 macrophages (n=115). Survival curves according to expression levels of tumor-infiltrating (C) M1 macrophages and (D) M2 macrophages in patients with left-sided CRC (n=114). Data used to generate this figure were downloaded from TCGA (The Cancer Genome Atlas) database. CRC, Colorectal Cancer.

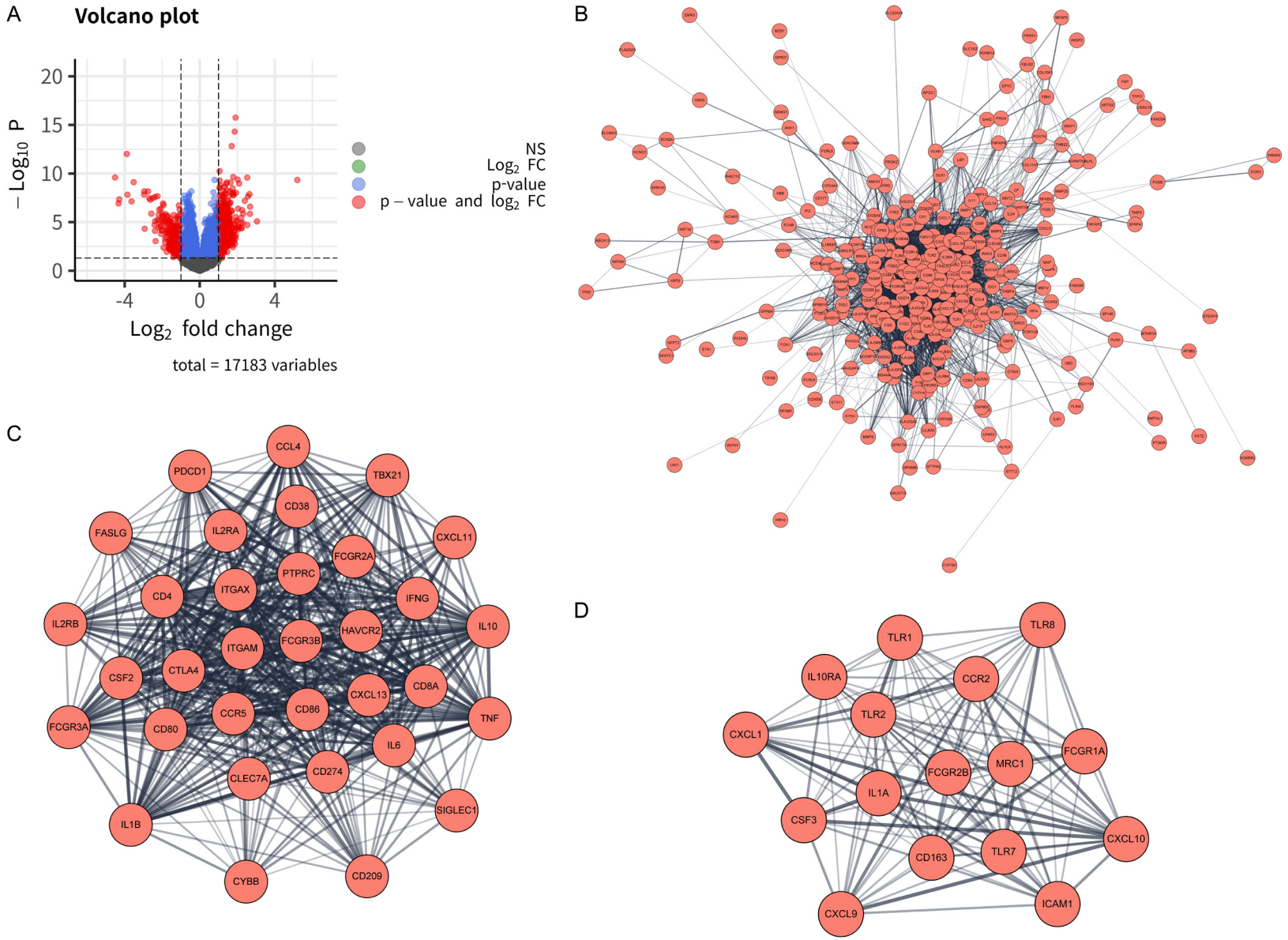
Functional enrichment analysis of DEGs and infiltrated CD8+ T cells

We conducted GO and KEGG enrichment analyses to determine the biological roles of 74 genes (upregulated: 59, downregulated: 15) identified through the PPI network and MCODE. Statistically significant GO and KEGG enrichment analyses are shown in **Figure 3**.

GO analysis was organized into three categories: biological process (BP), cellular component (CC), and molecular function (MF). The top BP of 59 upregulated genes were T-cell activation, cellular response to biotic stimulus, cell adhesion, and cellular response to lipopolysaccharide. The external side of the plasma membrane, endocytic vesicle membrane, MHC class II protein complex, and clathrin-coated endocytic vesicles were enriched in CC, and the MF of these genes included cytokine receptor binding, immune receptor activity, and MHC class II receptor activity (**Figure 3A**). The top 3 GO categories of the top 15 downregulated genes were protein processing, neuronal cell bodies, and hormone activity (**Figure 3B**).

Figure 3C and **3D** show the KEGG pathway analysis of 59 upregulated and 15 downregulated genes, respectively. It was confirmed that the DEGs were mainly associated with tuberculosis, leishmaniasis, rheumatoid arthritis, hematopoietic cell lineage, cell adhesion molecules, phagosomes, graft-versus-host disease, inflammatory bowel disease, type 1 diabetes mellitus, and allograft rejection. Neuroactive ligand-receptor interactions and thyroid hormone synthesis were included in the KEGG analysis of 15 downregulated genes.

Interestingly, GO results revealed that CD8+ T cell activation occurred in patients with high


A
Volcano plot
20
15
10
5
0
− Log10 P
-4
0
4
Log2 fold change
total = 17183 variables
NS
Log2 FC
p-value
p − value and log2 FC
B
C
CCL4
PDCD1
TBX21
CD38
FASLG
IL2RA
CXCL11
FCGR2A
PTPRC
CD4
ITGAX
IFNG
IL2RB
IL10
HAVCR2
ITGAM
FCGR3B
CSF2
CTLA4
CD8A
CXCL13
FCGR3A
CD80
CCR5
CD86
TNF
IL6
CLEC7A
CD274
SIGLEC1
IL1B
CYBB
CD209
D
TLR1
TLR8
IL10RA
CCR2
TLR2
CXCL1
FCGR1A
FCGR2B
MRC1
IL1A
CSF3
CXCL10
CD163
TLR7
ICAM1
CXCL9

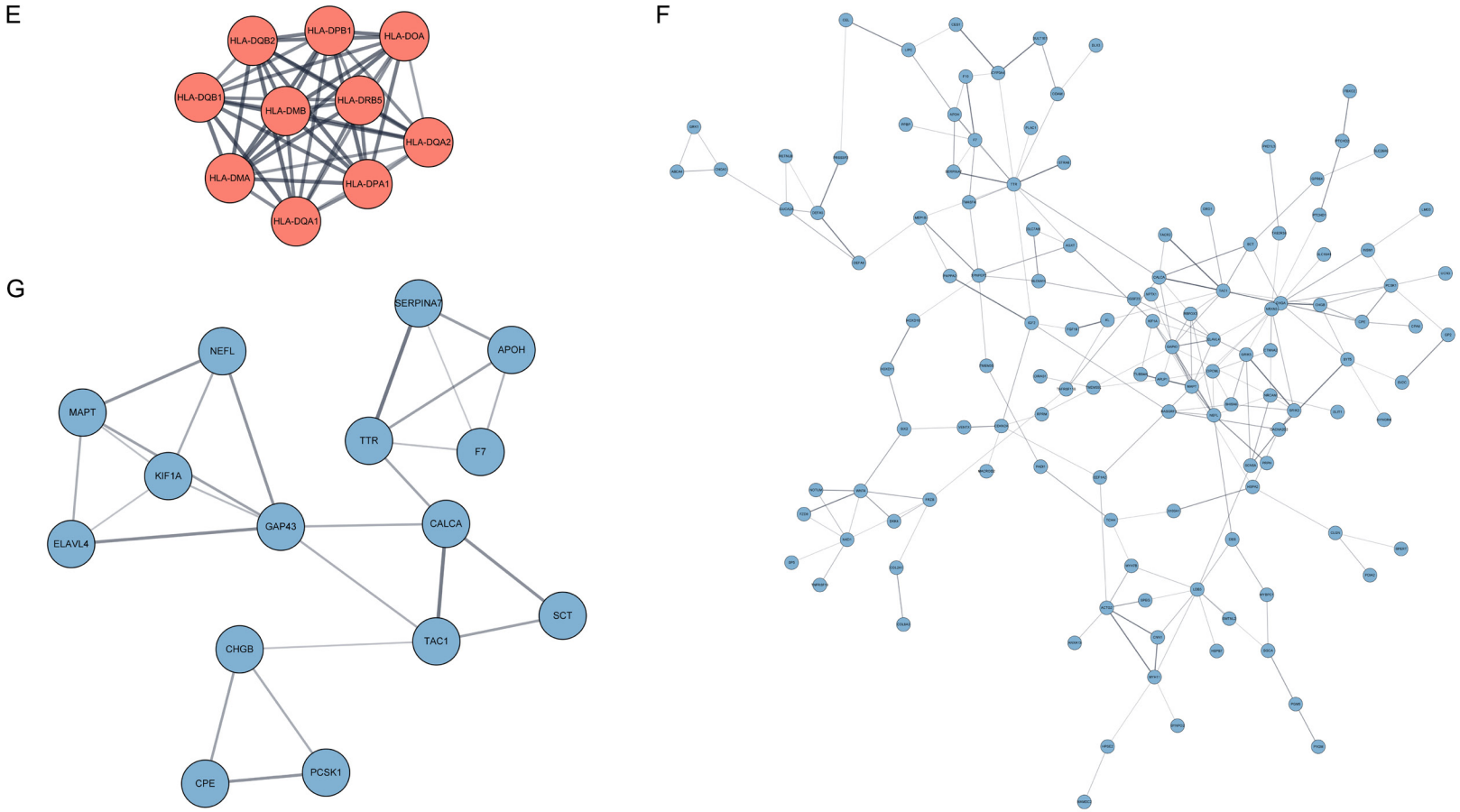


Figure 2. Volcano plot and protein-protein interaction (PPI) network construction and significant module selection involving the common differentially expressed genes (DEGs). A. Volcano plot of DEGs. In the plot, genes that meet the *p*-value and log2 fold-change cutoff value are denoted in red; Genes that pass only the *p*-value are denoted in blue; A gene that passes only the log2 fold change is denoted in green; Genes that pass both the *p*-value and log2 fold change are denoted in black. B. PPI network of upregulated genes. C. The module with the highest Molecular Complex Detection (MCODE) among the PPI network of upregulated genes. D. Module with the second-highest MCODE among the PPI network of upregulated genes. E. Module with the third-highest MCODE among the PPI network of upregulated genes. F. PPI network of downregulated genes. G. The module with the highest MCODE among the PPI network of downregulated genes.

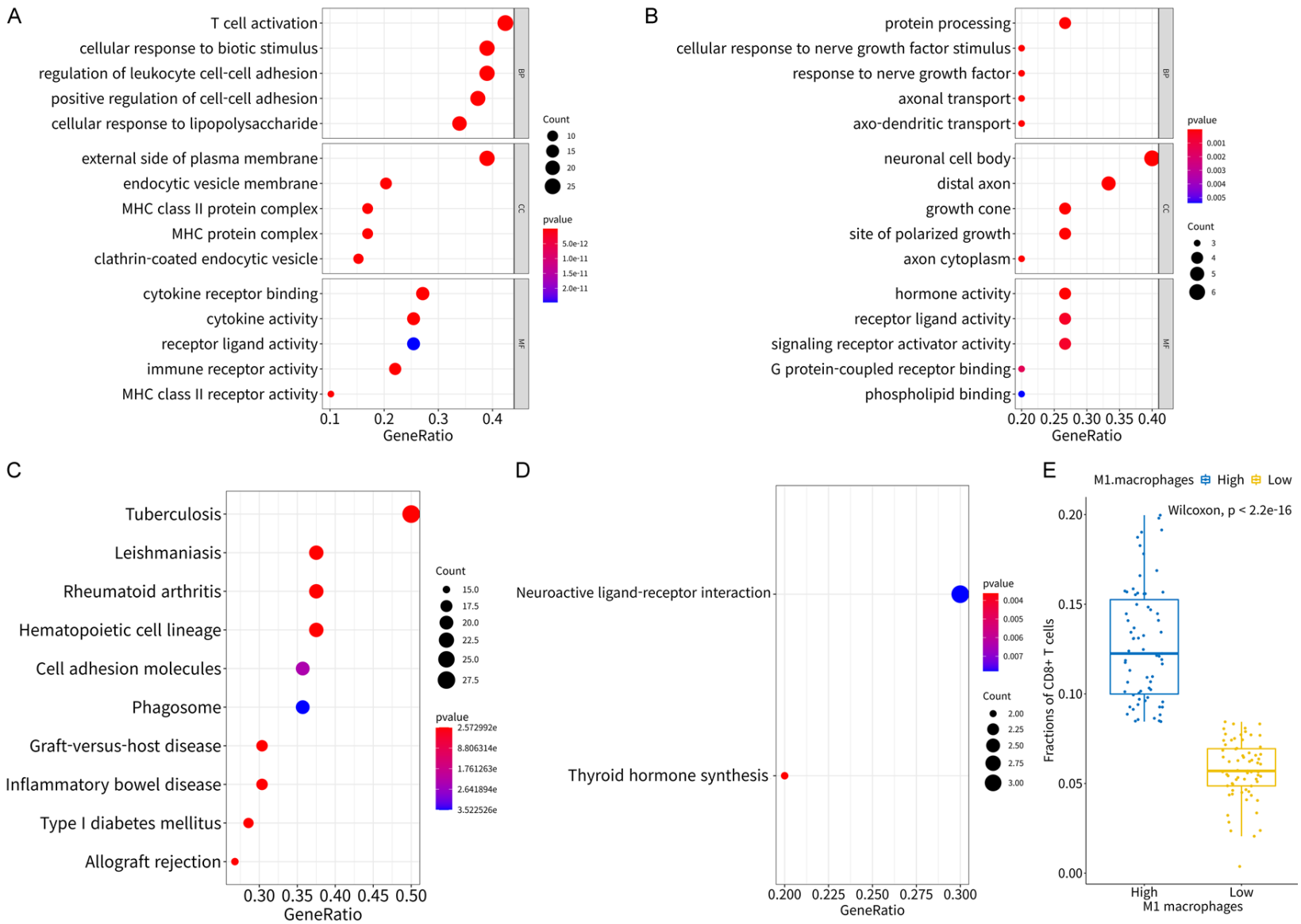


Figure 3. Functional analysis and boxplot. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of 74 genes (upregulated: 59, downregulated: 15) identified through protein-protein interaction network and module with the Molecular Complex Detection (MCODE) section. GO results for (A) upregulated and (B) downregulated genes. The GO analysis result is composed of biological process (BP), cellular component (CC), and molecular

function (MF) terms. Dot plots show *p*-values of < 0.05 for GO results. The size of the circle indicates the number of genes corresponding to each term. The closer to red, the lower the *p*-value. KEGG pathway enrichment analysis of (C) upregulated and (D) downregulated genes. The size of the circle indicates the number of genes corresponding to each term. Dot plots show *p*-values of < 0.05 for KEGG results. The closer to red, the lower the *p*-value. (E) Boxplot of the fractions of CD8+ T cells according to levels of M1 macrophages.

levels of M1 macrophages. **Figure 3E** shows the differences in fractions of CD8+ T cells according to the number of M1 macrophages. In patients with right-sided CRC, the relative mean of fractions was higher for high levels of M1 macrophages than for low levels of M1 macrophages (Wilcoxon signed-rank test, *P* < *2.2e-16*; **Figure 3E**).

LASSO regression analysis of prognosis-related genes and construction of the risk model

Kaplan-Meier survival curves for OS analyses were constructed to identify the prognosis-related DEGs, and 15 prognostic genes were used for LASSO regression analysis (Supplementary Figure 4). LASSO regression analysis was combined with tenfold CV to build a classifier for the M1 macrophage-related risk model. The optimal lambda value was calculated using the penalty regularization parameter. Finally, nine DEGs were identified as independent prognostic genes, including *CTLA4*, *CD274*, *CXCL1*, *CXCL10*, *HLA-DPA1*, *IFNG*, *IL1A*, *IL1B*, and *TLR8* (**Figure 4**). The mean expression levels of nine prognosis-related genes were higher in the high level M1 macrophage group than in the low level M1 macrophage group in patients with right-sided CRC, as shown in **Figure 5** (*CTLA4, P=1.4e-06*; *CD274, P=8e-07*; *CXCL1, P= 1.2e-12*; *CXCL10*, *P=3e-06*; *HLA-DPA1, P= 5.4e-06*; *IFNG, P=0.00021*; *IL1A, P=0.0026*; *IL1B, P=2.2e-05*; *TLR8, P=7e-05*). The coefficients for the nine DEGs were calculated using LASSO analysis, and the M1 macrophage-related risk score was calculated according to the linear combination of the expression of each gene multiplied by LASSO coefficients: (0.0465 × *CTLA4*) + (-0.2261 × *CD274*) + (-0.0038 × *CXCL1*) + (0.0012 × *CXCL10*) + (0.0003 × *HLA-DPA1*) + (0.1002 × *IFNG*) + (-0.0141 × *IL1A*) + (-0.0096 × *IL1B*) + (-0.0439 × *TLR8*). ROC curve analysis was performed to verify the best predictive accuracy of the risk model using nine prognosis-related genes selected by LASSO analysis (Supplementary Figure 5). We divided patients with right-sided CRC into "High-risk" and "Low-risk" groups according to the median risk score.

Construction and validation of M1 macrophages-related risk model

Kaplan-Meier survival curves for OS and PFS analyses showed that the high-risk group had a poorer prognosis than the low-risk group (OS, *P=0.00058*; PFS, *P=0.0029*; **Figure 6A**, **6B**). The ROC curve was used to demonstrate the accuracy of the developed risk model for prognosis in patients with right-sided CRC. Areas under the curve (AUC) values for OS and PFS were 0.743 and 0.702, respectively (**Figure 6C**, **6D**).

GSE103479 [19] and GSE72970 [20] datasets were used to validate the prediction performance of the M1 macrophage-related risk model. In the GSE103479 dataset, Kaplan-Meier survival curves for OS and PFS analyses revealed that the high-risk group had a poorer prognosis than the low-risk group (n=41; OS, P=0.0022; PFS, P=0.035; **Figure 6E**, **6F**), and AUC values from the ROC curves for OS and PFS were 0.619 and 0.726, respectively (**Figure 6H**, **6I**). Kaplan-Meier survival curves for OS analysis of the GSE72970 dataset included 28 patients with right-sided CRC, and OS in the high-risk group showed a poor prognosis (*P=0.02*; AUC=0.917; **Figure 6G** and **6J**).

Functional enrichment analysis of nine genes from the M1 macrophages-related risk model

Table 3 presents GO and KEGG enrichment analyses of the nine genes included in the M1 macrophage-related risk model. The GO enrichment assay showed that DEGs were primarily enriched in positive regulation of T-cell activation (GO: 0050870; *P=1.24e-07*), positive regulation of leukocyte cell-cell adhesion (GO: 1903039; *P=1.24e-07*), external side of the plasma membrane (GO: 0009897; *P=8.44e-04*), and cytokine activity (GO: 0005125; *P=1.20e-06*). The top 10 enriched KEGG pathways included: rheumatoid arthritis (*P=1.15e-08*), graft-versus-host disease (*P=2.06E-06*), type I diabetes mellitus (*P=2.06E-06*), inflammatory bowel disease (*P=6.75E-06*), influenza A (*P=6.75E-06*), leishmaniasis (*P=1.11E-05*),

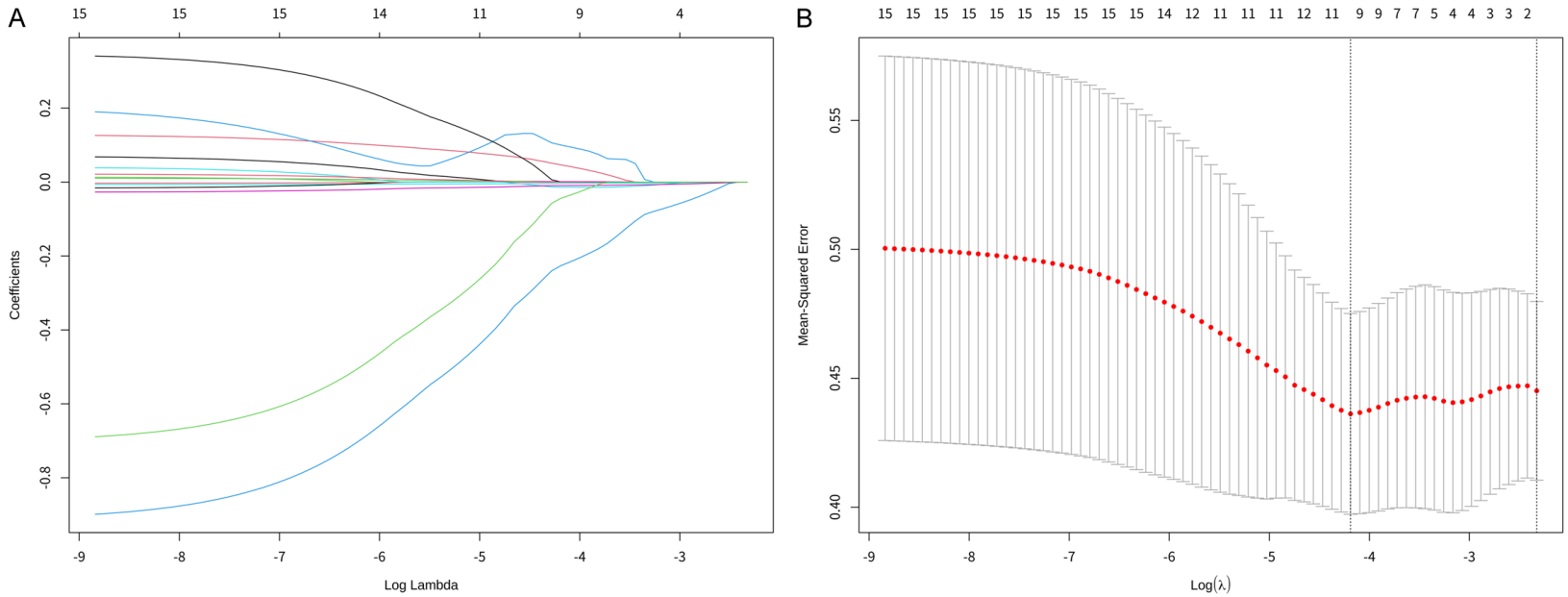


Figure 4. Prognostic gene selection using least absolute shrinkage and selection operator (LASSO) regression model. A. Distribution of LASSO coefficients. B. Tuning parameter selection in the LASSO risk model for tenfold cross-validation. Nine genes significantly related to prognosis were selected.

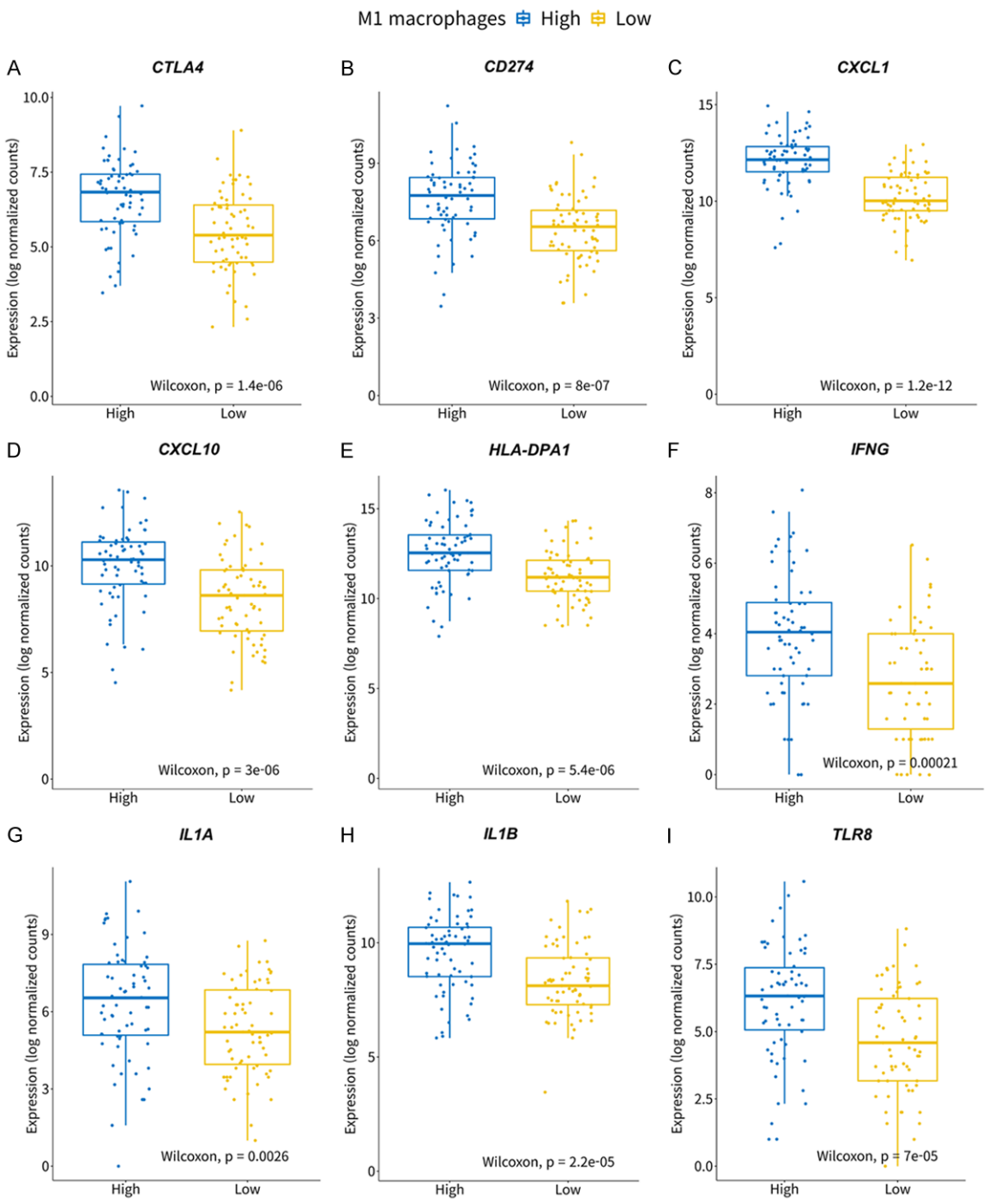
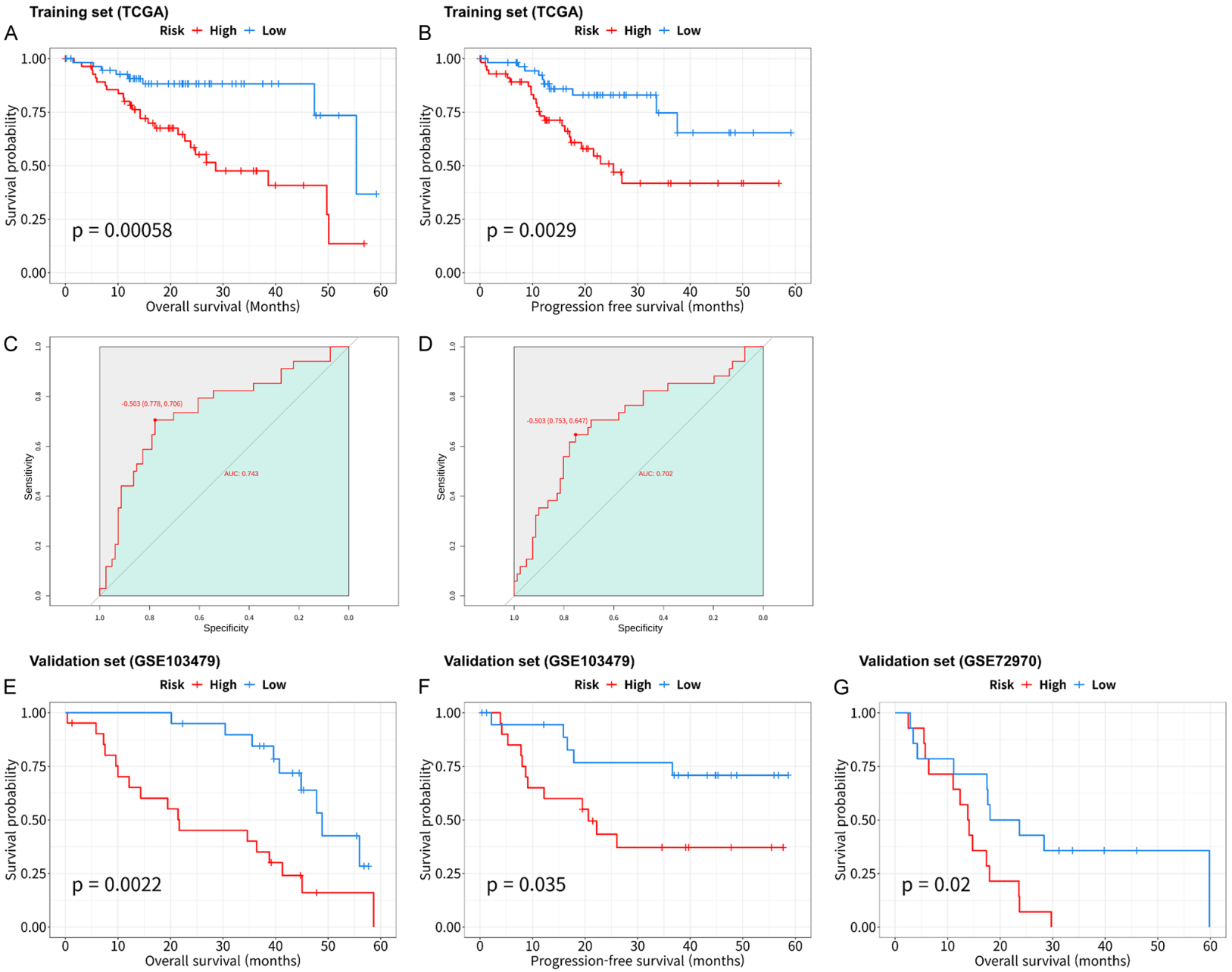


Figure 5. Boxplot of nine prognosis-related genes according to levels of M1 macrophages in patients with right-sided CRC. (A) *CTLA4*, (B) *CD274*, (C) *CXCL1,* (D) *CXCL10*, (E) *HLA-DPA1,* (F) *IFNG*, (G) *IL1A*, (H) *IL1B*, and (I) *TLR8*. Groups were evaluated using the Wilcoxon signed-rank test to generate the *p*-value. CRC, Colorectal Cancer.

interleukin (IL)-17 signaling pathway (*P=2.13E-05*), cytokine-cytokine receptor interaction (*P=6.27E-05*), tuberculosis (*P=2.20E-04*), hematopoietic cell lineage (*P=9.89E-04*).

Discussion

In the present study, we developed an M1 macrophage-related risk model based on M1 mac-


A
Training set (TCGA)
Risk High Low
Survival probability
p = 0.00058
Overall survival (Months)
B
Training set (TCGA)
Risk High Low
Survival probability
p = 0.0029
Progression free survival (months)
C
-0.503 (0.778, 0.706)
AUC: 0.743
Sensitivity
Specificity
D
-0.503 (0.753, 0.647)
AUC: 0.702
Sensitivity
Specificity
E
Validation set (GSE103479)
Risk High Low
Survival probability
p = 0.0022
Overall survival (months)
F
Validation set (GSE103479)
Risk High Low
Survival probability
p = 0.035
Progression-free survival (months)
G
Validation set (GSE72970)
Risk High Low
Survival probability
p = 0.02
Overall survival (months)

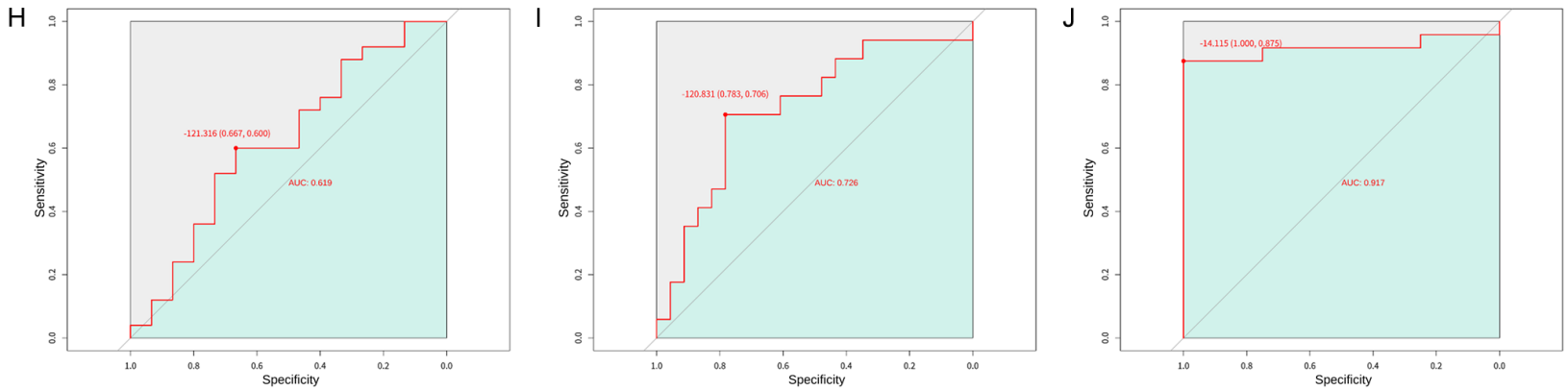


Figure 6. Development and validation of the prognostic risk model. Patients were classified into high-risk and low-risk groups based on their prognostic risk score in The Cancer Genome Atlas (TCGA) training cohort (n=115), GSE103479 validation cohort (n=41), and GSE72970 validation cohort (n=28). A, C. Kaplan-Meier survival curves and receiver operating characteristic (ROC) curves for overall survival (OS) analysis of TCGA training cohorts, respectively. B, D. Kaplan-Meier survival curves and ROC curves for progression-free survival (PFS) analysis of TCGA training cohort, respectively. E, H. Kaplan-Meier survival curves and ROC curves for OS analysis of GSE103479 validation cohort, respectively. F, I. Kaplan-Meier survival curves and ROC curves for PFS analysis of GSE103479 validation cohort, respectively. G, J. Kaplan-Meier survival curves and ROC curves for OS analysis of GSE72970 validation cohort, respectively.

Table 3. GO and KEGG pathway enrichment analysis for nine prognosis-related genes

Category	Description	ID	Count	p.adjust[1]	Gene symbol
GO: Biological process	positive regulation of T cell activation	GO: 0050870	6	1.24E-07	*CD274/CTLA4/HLA-DPA1/IFNG/IL1A/IL1B*
	positive regulation of leukocyte cell-cell adhesion	GO: 1903039	6	1.24E-07	*CD274/CTLA4/HLA-DPA1/IFNG/IL1A/IL1B*
	positive regulation of cell-cell adhesion	GO: 0022409	6	2.24E-07	*CD274/CTLA4/HLA-DPA1/IFNG/IL1A/IL1B*
	regulation of T cell activation	GO: 0050863	6	3.93E-07	*CD274/CTLA4/HLA-DPA1/IFNG/IL1A/IL1B*
	regulation of leukocyte cell-cell adhesion	GO: 1903037	6	3.93E-07	*CD274/CTLA4/HLA-DPA1/IFNG/IL1A/IL1B*
GO: Cellular component	external side of plasma membrane	GO: 0009897	4	8.44E-04	*CD274/CTLA4/CXCL10/TLR8*
	clathrin-coated endocytic vesicle	GO: 0045334	2	6.10E-03	*CTLA4/HLA-DPA1*
	clathrin-coated vesicle	GO: 0030136	2	4.19E-02	*CTLA4/HLA-DPA1*
	endolysosome membrane	GO: 0036020	1	4.19E-02	*TLR8*
	MHC class II protein complex	GO: 0042613	1	4.19E-02	*HLA-DPA1*
GO: Molecular function	cytokine activity	GO: 0005125	5	1.20E-06	*CXCL1/CXCL10/IFNG/IL1A/IL1B*
	cytokine receptor binding	GO: 0005126	5	1.20E-06	*CXCL1/CXCL10/IFNG/IL1A/IL1B*
	receptor ligand activity	GO: 0048018	5	1.18E-05	*CXCL1/CXCL10/IFNG/IL1A/IL1B*
	signaling receptor activator activity	GO: 0030546	5	1.18E-05	*CXCL1/CXCL10/IFNG/IL1A/IL1B*
	interleukin-1 receptor binding	GO: 0005149	2	1.63E-04	*IL1A/IL1B*
KEGG pathway	Rheumatoid arthritis	hsa05323	6	1.15E-08	*CTLA4/CXCL1/HLA-DPA1/IFNG/IL1A/IL1B*
	Graft-versus-host disease	hsa05332	4	2.06E-06	*HLA-DPA1/IFNG/IL1A/IL1B*
	Type I diabetes mellitus	hsa04940	4	2.06E-06	*HLA-DPA1/IFNG/IL1A/IL1B*
	Inflammatory bowel disease	hsa05321	4	6.75E-06	*HLA-DPA1/IFNG/IL1A/IL1B*
	Influenza A	hsa05164	5	6.75E-06	*CXCL10/HLA-DPA1/IFNG/IL1A/IL1B*
	Leishmaniasis	hsa05140	4	1.11E-05	*HLA-DPA1/IFNG/IL1A/IL1B*
	IL-17 signaling pathway	hsa04657	4	2.13E-05	*CXCL1/CXCL10/IFNG/IL1B*
	Cytokine-cytokine receptor interaction	hsa04060	5	6.27E-05	*CXCL1/CXCL10/IFNG/IL1A/IL1B*
	Tuberculosis	hsa05152	4	2.20E-04	*HLA-DPA1/IFNG/IL1A/IL1B*
	Hematopoietic cell lineage	hsa04640	3	9.89E-04	*HLA-DPA1/IL1A/IL1B*

1. *P*-values adjusted by the Benjamini-Hochberg correction. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

rophage infiltration in patients with right-sided CRC using the TCGA RNA-seq dataset and performed external validation of the model performance using GSE103479 and GSE72970 microarray data. Based on our observed findings, the M1 macrophage-related risk model, including nine prognostic genes (*CTLA4*, *CD274*, *CXCL1*, *CXCL10*, *HLA-DPA1*, *IFNG*, *IL1A*, *IL1B*, and *TLR8*), was an independent prognostic factor in predicting the outcomes of patients with CRC. The performance of the risk model was assessed using AUCs of the ROC curve for both training and validation sets. The nine identified prognostic genes could be potential biomarkers for effectively predicting the outcomes of patients with right-sided CRC.

Herein, our findings indicated that reduced infiltration of M1 macrophages from the right side of the colon could be associated with poor prognosis in patients with CRC. The homing of pro-inflammatory (M1) and anti-inflammatory (M2) macrophages reportedly display different roles during inflammation [36]. More importantly, M1 and M2 macrophages exert distinct effects on the TME. In particular, M1 macrophage activation is triggered by interferon-γ (IFN-γ), bacterial lipopolysaccharide (LPS), or tumor necrosis factor α and can be induced by several signal transduction pathways involving STAT and nuclear factor kappa-light-chain-enhancer of activated B cells (NFκB) [37]. Chemokines are major inducers of macrophage chemotaxis; however, their involvement in the differential regulation of M1 and M2 macrophages and CD8+ T cell recruitment in the TME of patients with CRC remains unclear. M1 macrophages are potent effector cells that kill microorganisms, inhibit tumor growth, and produce pro-inflammatory cytokines [38, 39]. Although most tumor-associated macrophages exhibit immunosuppressive M2 macrophages, they retain plasticity for polarization. Moreover, efforts are ongoing to re-polarize tumor-promoting M2 macrophages from tumoricidal M1 macrophages. One potential antitumor strategy involves altering the tumor immune microenvironment (TIME) by engineering macrophages. M1 macrophages alter immune suppression within the TME by facilitating antitumor T cells and directly attacking tumor cells. Specifically, chimeric antigen receptor macrophages (CAR-Ms) can induce antigen-specific phagocytosis and tumor removal *in vitro*. In humanized mouse models, CAR-Ms have been shown to induce a pro-inflammatory TME and promote antitumor T-cell activity [40-42]. In the current study, M1 macrophage is significantly associated with prognosis of right-sided CRC patients and it is correlated with CD8+ T cells. Although the survival analysis result is not statistically significant, patients with high CD8+ T cells tend to live longer than patients with low CD8+ T cells. In the gene set analysis, M1 macrophage in right-sided CRC induces T cell activation and has potent inflammatory function. Therefore, we believe that the M1 macrophage affects the survival of right-sided CRC patients in various ways such as anti-tumor function of M1 macrophage and CD8+ T cells activation. Future studies need to generate additional evidence regarding the association between M1 macrophage infiltration, T-cell activation, and patient survival according to CRC location.

We observed that high expression levels of *CTLA4* and *CD274* were associated with a good prognosis in patients with right-sided CRC. This finding is consistent with the report that both overexpression of cytotoxic T-lymphocyte-associated antigen 4 (CTLA-4) and programmed death ligand-1 (PD-L1, encoded by the *CD274* gene) can be crucial suppressors of antitumor immunity, associated with a better therapeutic response. Moreover, Li et al. have reported that overexpression of PD-1 and PD-L1 correlated with better prognosis in patients with CRC. Tumor-infiltrating lymphocyte (TIL)-PD-1 is an independent prognostic factor for survival in patients with CRC [43]. Interestingly, KEGG pathway analysis revealed that *IFNG*, *IL1A*, *IL1B*, and *HLA-DPA1* are associated with chronic diseases, including rheumatoid arthritis, graft-versus-host disease, type I diabetes mellitus, and inflammatory bowel disease. IFN-γ plays an essential role in the innate immune response to intracellular bacterial pathogens. In addition, IFN-γ coordinates several biological responses, primarily involved in host defense and immune surveillance, as well as in the establishment of adaptive immunity and modulation of pro-inflammatory macrophages. IFN-γ reportedly induces M1 macrophages. Upregulation of MHC class I by IFN-γ plays a crucial role in the host response to intracellular pathogens and tumor cells owing to cytotoxic T-cell activation, promoting cell-mediated immunity. IFN-γ functions as a cytotoxic CD8+ T

cell differentiation signal and is essential for inducing cytotoxic T-cell precursor proliferation [44, 45]. HLA-DPA1 belongs to HLA class II alpha chain paralogs. High expression level of *HLA-DPA1* can be associated with a good host immune response, positive regulation of chronic inflammation, and a better prognosis in CRC [46]. Recently, Singh et al. have reported that five genes, *CXCL3*, *MET*, *IL1A*, *IL1B*, and *TNS1*, identified through an integrated bioinformatics approach, can potentially serve as biomarkers for the regulatory function of the inflammatory response and cellular homeostasis in patients with early-age onset CRC. Consistently, our data suggest that *IL1A* and *IL1B* play important roles in patients with right-sided CRC [47].

Wang et al. have reported that *CXCL1* and *CXCL10* are pivotal for initiating malignancy and cancer progression. The authors suggested that high *CXCL1* expression level is significantly correlated with breast cancer lymph node metastasis, poor OS, and the basal-like subtype [48]. However, in our study, we determined that high expression levels of *CXCL1* and *CXCL10* were associated with better survival rates in patients with CRC, not only in the training set but also in the two validation sets. Accordingly, additional evidence and discussions are needed to comprehensively clarify how chemokines function as major inducers of macrophages. In addition, the role of macrophage phenotypes on the prognosis of patients with CRC needs to be elaborated. Toll-like receptors are the most frequently examined RNA sensors for intestinal diseases. He et al. have reported lower expression levels of TLR7 and TLR8 in CRC tissues than in controls, indicating potential antitumor effects [49]. This result is consistent with the conclusions of the present study.

Various studies have developed risk models using LASSO regression [16, 50, 51], demonstrating the performance of risk models using tenfold CV to avoid overfitting. In the present study, we systematically developed a risk model based on M1 macrophage infiltration in right-sided CRC. Our study had the advantage of evaluating the performance of the risk model, given that it was validated in two GEO datasets, including patients with right-sided CRC. Both OS and PFS were evaluated using ROC curves. AUC values revealed that the M1 macrophage-related risk model demonstrated good accuracy in predicting survival in TCGA training set and two GEO validation sets. Multifaceted strategies are warranted to develop precision medicine for cancer therapy. We suggest that the M1 macrophage-related risk model may provide valuable prognosis of patients with right-sided CRC.

There are several limitations to this study. First, we performed only based on bioinformatics methods for the analysis. Second, we couldn't evaluate association with between immune cell infiltration and drug responsiveness in patients due to the lack of datasets including drug responsiveness. Therefore, further in vivo and in vitro experiments are needed to validate patient prognosis as well as drug responsiveness. Nevertheless, the risk model for predicting OS and PFS in patients with right sided CRC we developed showed superior performance in both TCGA and two GEO cohorts. The prognostic gene signature we generated may provide a direction for further research finding about mechanism of TME in CRC.

In Conclusions, our study constructed an M1 macrophage-related risk model to predict the outcomes of patients with right-sided CRC. The score generated from our risk signature based on the nine prognosis-related genes was an independent prognostic marker for predicting OS and PFS in both TCGA and GEO cohorts. This study provides a novel gene signature for the prognosis of patients with right-sided CRC and offers a vital principle for future studies to examine the relationship between infiltration of M1 macrophages and the prognosis of patients with right-sided CRC.

Acknowledgements

This study was supported by the National Research Foundation of Korea (NRF-2018R1-A5A2023879, 2020R1A2C1005203, 2021R1-A2C4001466, and 2022R1A5A2027161), and this research was supported by the Korean Fund for Regenerative Medicine (KFRM) grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Health & Welfare, code: 22A0205L1-11). This work was also supported by the 2021 Post-Doctoral Development Program of Pusan National University.

Disclosure of conflict of interest

None.

Address correspondence to: Ninib Baryawno, Childhood Cancer Research Unit, Department of Women's and Children's Health, Karolinska Institutet, K6 Kvinnors och barns hälsa, K6 Barnonkologi och Barnkirurgi Johnsen, Stockholm 17177, Sweden. Tel: +46-(0)-76-589-77-40; Fax: +46-(0)-76-589-77-40; E-mail: n.baryawno@ki.se; Hyung-Sik Kim, Department of Oral Biochemistry, School of Dentistry, Pusan National University, Room #411, 49 Busandaehak-ro, Mulgeum-eup, Yangsan-si, Gyeongsangnam-do 50612, Republic of Korea. Tel: +82-51-510-8231; Fax: +82-51-510-8210; E-mail: hskimcell@pusan.ac.kr; Dongjun Lee, Department of Convergence Medicine, School of Medicine, Pusan National University, Room #502, 49 Busandaehak-ro, Mulgeum-eup, Yangsan-si, Gyeongsangnam-do 50612, Republic of Korea. Tel: +82-51-510-8042; Fax: +82-51-510-8526; E-mail: lee.dongjun@pusan.ac.kr; Yun Hak Kim, Dental and Life Science Institute, School of Dentistry, Pusan National University, Room #701, 49 Busandaehak-ro, Mulgeum-eup, Yangsan-si, Gyeongsangnam-do 50612, Republic of Korea. Tel: +82-51-510-8091; Fax: +82-51-510-8049; E-mail: yunhak10510@pusan.ac.kr

References

[1] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D and Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015; 136: E359-386.

[2] Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, van de Velde CJ and Watanabe T. Colorectal cancer. Nat Rev Dis Primers 2015; 1: 15065.

[3] Dekker E, Tanis PJ, Vleugels JLA, Kasi PM and Wallace MB. Colorectal cancer. Lancet 2019; 394: 1467-1480.

[4] Huxley RR, Ansary-Moghaddam A, Clifton P, Czernichow S, Parr CL and Woodward M. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. Int J Cancer 2009; 125: 171-180.

[5] Biasco G, Derenzini E, Grazi G, Ercolani G, Ravaioli M, Pantaleo MA and Brandi G. Treatment of hepatic metastases from colorectal cancer: many doubts, some certainties. Cancer Treat Rev 2006; 32: 214-228.

[6] Lievre A, Blons H and Laurent-Puig P. Oncogenic mutations as predictive factors in colorectal cancer. Oncogene 2010; 29: 3033-3043.

[7] Markowitz SD and Bertagnolli MM. Molecular basis of colorectal cancer. N Engl J Med 2009; 361: 2449-2460.

[8] Sartore-Bianchi A, Martini M, Molinari F, Veronese S, Nichelatti M, Artale S, Di Nicolantonio F, Saletti P, De Dosso S, Mazzucchelli L, Frattini M, Siena S and Bardelli A. PIK3CA mutations in colorectal cancer are associated with clinical resistance to EGFR-targeted monoclonal antibodies. Cancer Res 2009; 69: 1851-1857.

[9] Mehlen P and Fearon ER. Role of the dependence receptor DCC in colorectal cancer pathogenesis. J Clin Oncol 2004; 22: 3420-3428.

[10] Meguid RA, Slidell MB, Wolfgang CL, Chang DC and Ahuja N. Is there a difference in survival between right-versus left-sided colon cancers? Ann Surg Oncol 2008; 15: 2388-2394.

[11] Nawa T, Kato J, Kawamoto H, Okada H, Yamamoto H, Kohno H, Endo H and Shiratori Y. Differences between right- and left-sided colon cancer in patient characteristics, cancer morphology and histology. J Gastroenterol Hepatol 2008; 23: 418-423.

[12] Petrelli F, Tomasello G, Borgonovo K, Ghidini M, Turati L, Dallera P, Passalacqua R, Sgroi G and Barni S. Prognostic survival associated with left-sided vs right-sided colon cancer: a systematic review and meta-analysis. JAMA Oncol 2017; 3: 211-219.

[13] Yaeger R, Chatila WK, Lipsyc MD, Hechtman JF, Cercek A, Sanchez-Vega F, Jayakumaran G, Middha S, Zehir A, Donoghue MTA, You D, Viale A, Kemeny N, Segal NH, Stadler ZK, Varghese AM, Kundra R, Gao J, Syed A, Hyman DM, Vakiani E, Rosen N, Taylor BS, Ladanyi M, Berger MF, Solit DB, Shia J, Saltz L and Schultz N. Clinical sequencing defines the genomic landscape of metastatic colorectal cancer. Cancer Cell 2018; 33: 125-136, e3.

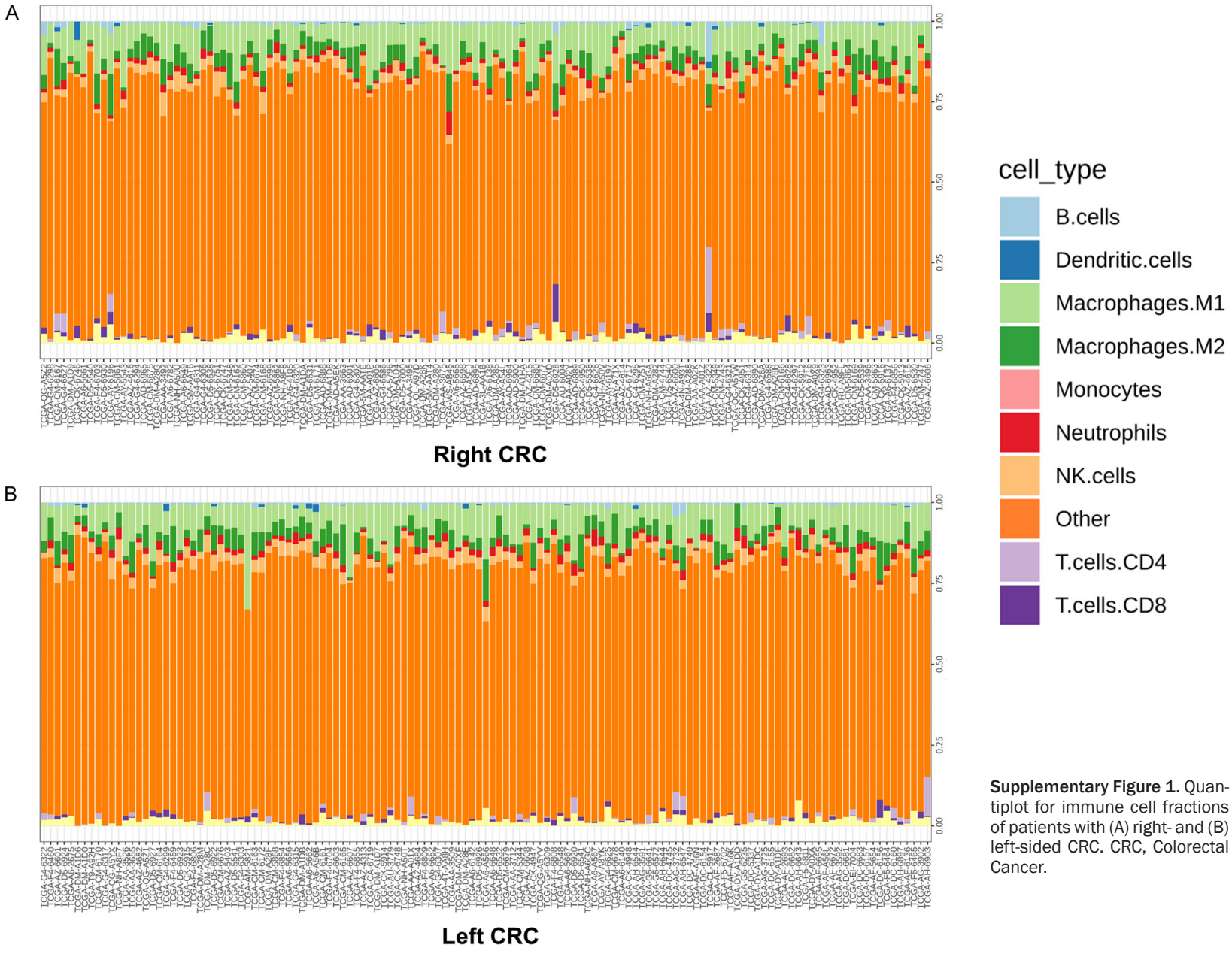
[14] Baran B, Mert Ozupek N, Yerli Tetik N, Acar E, Bekcioglu O and Baskin Y. Difference between left-sided and right-sided colorectal cancer: a focused review of literature. Gastroenterology Res 2018; 11: 264-273.

[15] Weiss JM, Pfau PR, O'Connor ES, King J, LoConte N, Kennedy G and Smith MA. Mortality by stage for right- versus left-sided colon cancer: analysis of surveillance, epidemiology, and end results-Medicare data. J Clin Oncol 2011; 29: 4401-4409.

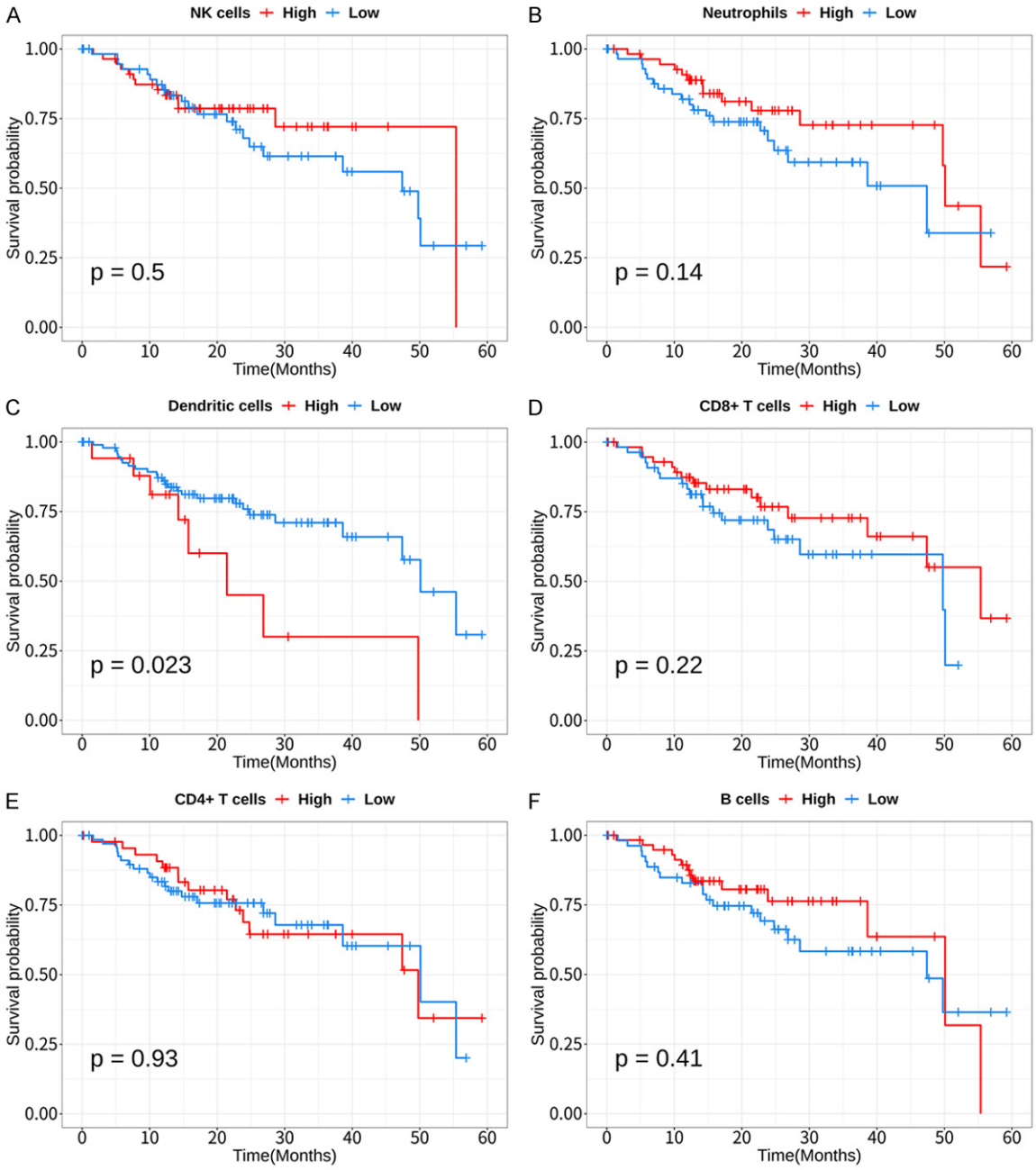
[16] Wu J, Li L, Zhang H, Zhao Y, Zhang H, Wu S and Xu B. A risk model developed based on tumor microenvironment predicts overall survival and associates with tumor immunity of patients

with lung adenocarcinoma. Oncogene 2021; 40: 4413-4424.
[17] Colangelo T, Polcaro G, Muccillo L, D'Agostino G, Rosato V, Ziccardi P, Lupo A, Mazzoccoli G, Sabatino L and Colantuoni V. Friend or foe? The tumour microenvironment dilemma in colorectal cancer. Biochim Biophys Acta Rev Cancer 2017; 1867: 1-18.
[18] Lee GH, Malietzis G, Askari A, Bernardo D, Al-Hassi HO and Clark SK. Is right-sided colon cancer different to left-sided colorectal cancer?-A systematic review. Eur J Surg Oncol 2015; 41: 300-308.
[19] Allen WL, Dunne PD, McDade S, Scanlon E, Loughrey M, Coleman H, McCann C, McLaughlin K, Nemeth Z, Syed N, Jithesh P, Arthur K, Wilson R, Coyle V, McArt D, Murray GI, Samuel L, Nuciforo P, Jimenez J, Argiles G, Dienstmann R, Tabernero J, Messerini L, Nobili S, Mini E, Sheahan K, Ryan E, Johnston PG, Van Schaeybroeck S, Lawler M and Longley DB. Transcriptional subtyping and CD8 immunohistochemistry identifies poor prognosis stage II/III colorectal cancer patients who benefit from adjuvant chemotherapy. JCO Precis Oncol 2018; 2018: PO.17.00241.
[20] Del Rio M, Mollevi C, Bibeau F, Vie N, Selves J, Emile JF, Roger P, Gongora C, Robert J, Tubiana-Mathieu N, Ychou M and Martineau P. Molecular subtypes of metastatic colorectal cancer are associated with patient response to irinotecan-based therapies. Eur J Cancer 2017; 76: 68-75.
[21] Team RC. R: a language and environment for statistical computing. 2013.
[22] Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, Krogsdam A, Loncova Z, Posch W, Wilflingseder D, Sopper S, Ijsselsteijn M, Brouwer TP, Johnson D, Xu Y, Wang Y, Sanders ME, Estrada MV, Ericsson-Gonzalez P, Charoentong P, Balko J, de Miranda NFDCC and Trajanoski Z. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. Genome Med 2019; 11: 34.
[23] Robinson MD, McCarthy DJ and Smyth GK. EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010; 26: 139-140.
[24] Blighe K, Rana S and Lewis M. EnhancedVolcano: publication-ready volcano plots with enhanced colouring and labeling. R Package Version 2019; 1.
[25] Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ and von Mering C. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res 2015; 43: D447-D452.
[26] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003; 13: 2498-2504.
[27] Bader GD and Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 2003; 4: 2.
[28] Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X and Yu G. ClusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innovation (Camb) 2021; 2: 100141.
[29] Yu G. Enrichplot: visualization of functional enrichment result. R Package Version 2019; 1.
[30] Friedman J, Hastie T and Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010; 33: 1-22.
[31] Pak K, Kim YH, Suh S, Goh TS, Jeong DC, Kim SJ, Kim IJ, Han ME and Oh SO. Development of a risk scoring system for patients with papillary thyroid cancer. J Cell Mol Med 2019; 23: 3010-3015.
[32] Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC and Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011; 12: 77.
[33] Wickham H. ggplot2: elegant graphics for data analysis. Springer; 2016.
[34] Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C and Stuart JM. The cancer genome atlas pan-cancer analysis project. Nat Genet 2013; 45: 1113-1120.
[35] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C and Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012; 2: 401-404.
[36] Xuan W, Qu Q, Zheng B, Xiong S and Fan GH. The chemotaxis of M1 and M2 macrophages is regulated by different chemokines. J Leukoc Biol 2015; 97: 61-69.
[37] Tang X, Mo C, Wang Y, Wei D and Xiao H. Anti-tumour strategies aiming to target tumour-associated macrophages. Immunology 2013; 138: 93-104.
[38] Saccani A, Schioppa T, Porta C, Biswas SK, Nebuloni M, Vago L, Bottazzi B, Colombo MP, Mantovani A and Sica A. p50 nuclear factor-κB

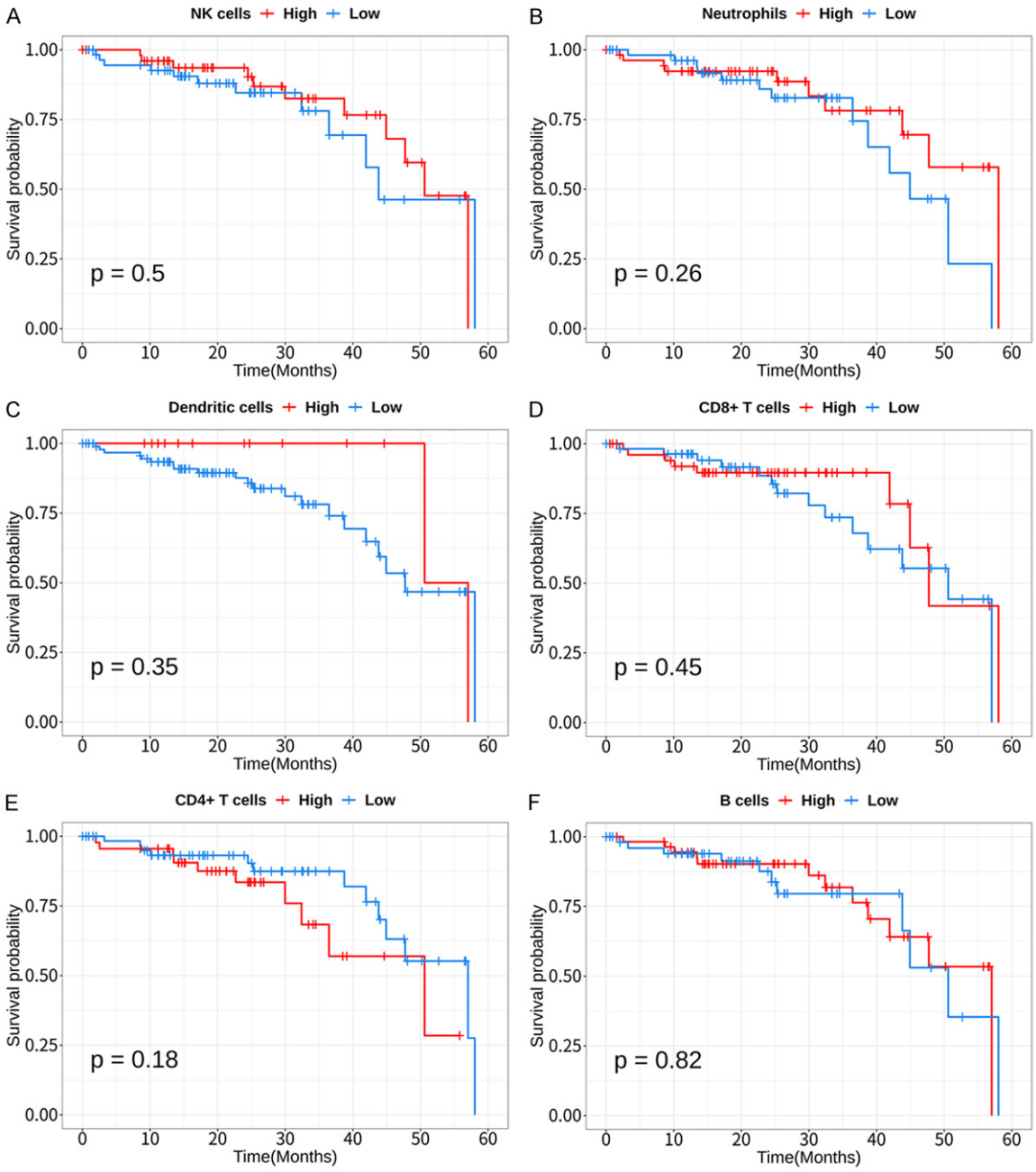
overexpression in tumor-associated macrophages inhibits M1 inflammatory responses and antitumor resistance. Cancer Res 2006; 66: 11432-11440.
[39] Wang YC, He F, Feng F, Liu XW, Dong GY, Qin HY, Hu XB, Zheng MH, Liang L, Feng L, Liang YM and Han H. Notch signaling determines the M1 versus M2 polarization of macrophages in antitumor immune responses. Cancer Res 2010; 70: 4840-4849.
[40] Biswas SK, Allavena P and Mantovani A. Tumor-associated macrophages: functional diversity, clinical significance, and open questions. Semin Immunopathol 2013; 35: 585-600.
[41] Klichinsky M, Ruella M, Shestova O, Lu XM, Best A, Zeeman M, Schmierer M, Gabrusiewicz K, Anderson NR, Petty NE, Cummins KD, Shen F, Shan X, Veliz K, Blouch K, Yashiro-Ohtani Y, Kenderian SS, Kim MY, O'Connor RS, Wallace SR, Kozlowski MS, Marchione DM, Shestov M, Garcia BA, June CH and Gill S. Human chimeric antigen receptor macrophages for cancer immunotherapy. Nat Biotechnol 2020; 38: 947-953.
[42] Xia Y, Rao L, Yao H, Wang Z, Ning P and Chen X. Engineering macrophages for cancer immunotherapy and drug delivery. Adv Mater 2020; 32: e2002054.
[43] Li Y, Liang L, Dai W, Cai G, Xu Y, Li X, Li Q and Cai S. Prognostic impact of programed cell death-1 (PD-1) and PD-ligand 1 (PD-L1) expression in cancer cells and tumor infiltrating lymphocytes in colorectal cancer. Mol Cancer 2016; 15: 55.
[44] Castro F, Cardoso AP, Gonçalves RM, Serre K and Oliveira MJ. Interferon-gamma at the crossroads of tumor immune surveillance or evasion. Front Immunol 2018; 9: 847.
[45] Berg RE, Crossley E, Murray S and Forman J. Memory CD8+ T cells provide innate immune protection against Listeria monocytogenes in the absence of cognate antigen. J Exp Med 2003; 198: 1583-1593.
[46] Ji L, Fu J, Hao J, Ji Y, Wang H, Wang Z, Wang P and Xiao H. Proteomics analysis of tissue small extracellular vesicles reveals protein panels for the reoccurrence prediction of colorectal cancer. J Proteomics 2021; 249: 104347.
[47] Singh MP, Rai S, Singh NK and Srivastava S. Transcriptomic landscape of early age onset of colorectal cancer identifies novel genes and pathways in Indian CRC patients. Sci Rep 2021; 11: 11765.
[48] Wang N, Liu W, Zheng Y, Wang S, Yang B, Li M, Song J, Zhang F, Zhang X, Wang Q and Wang Z. CXCL1 derived from tumor-associated macrophages promotes breast cancer metastasis via activating NF-kappaB/SOX4 signaling. Cell Death Dis 2018; 9: 880.
[49] He L, Wang F, Tian H, Xie Y, Xie L and Liu Z. The expression profile of RNA sensors in colorectal cancer and its correlation with cancer stages. Transl Cancer Res 2019; 8: 1351-1363.
[50] Yang J, Kim H, Shin K, Nam Y, Heo HJ, Kim GH, Hwang BY, Kim J, Woo S, Choi HS, Ko DS, Lee D and Kim YH. Molecular insights into the development of hepatic metastases in colorectal cancer: a metastasis prediction study. Eur Rev Med Pharmacol Sci 2020; 24: 12701-12708.
[51] Ye Y, Dai Q and Qi H. A novel defined pyroptosis-related gene signature for predicting the prognosis of ovarian cancer. Cell Death Discov 2021; 7: 71.



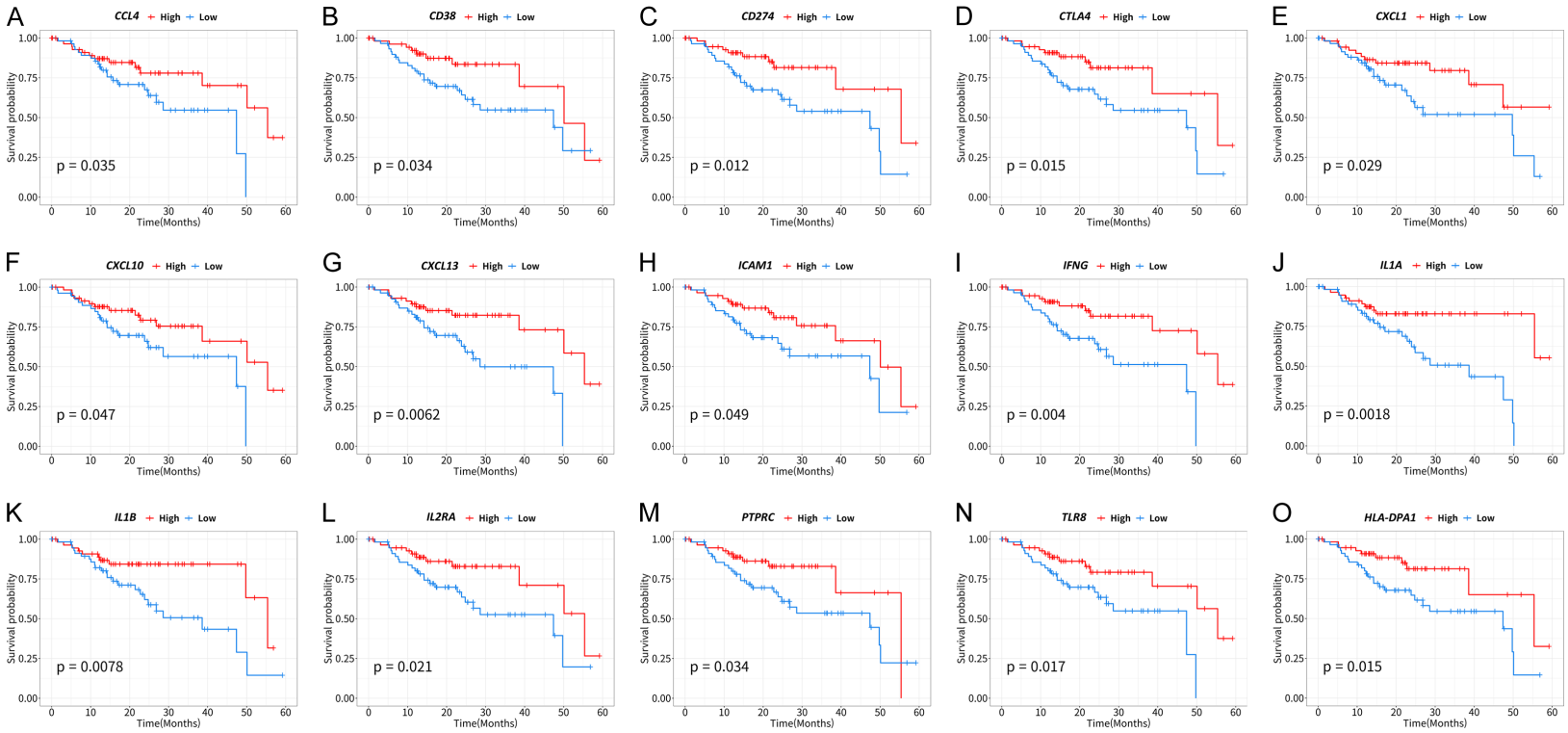
Supplementary Figure 1. Quantiplot for immune cell fractions of patients with (A) right- and (B) left-sided CRC. CRC, Colorectal Cancer.



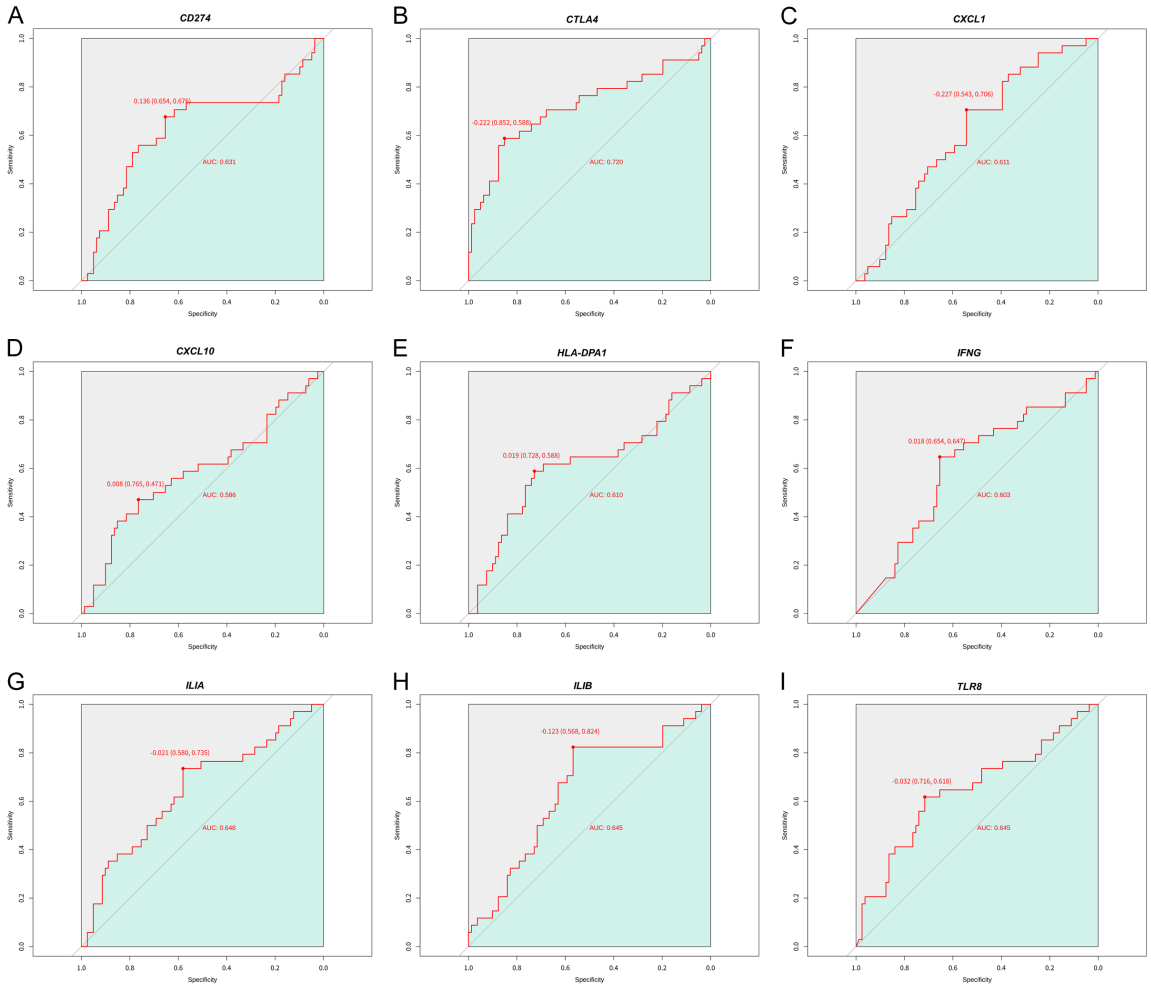
Supplementary Figure 2. Kaplan-Meier curves for overall survival analysis of patients with right-sided CRC according to expression levels of tumor-infiltrating (A) NK cells, (B) Neutrophils, (C) Dendritic cells, (D) CD8+ T cells, (E) CD4+ T cells, and (F) B cells. CRC, Colorectal Cancer; NK, Natural Killer.



Supplementary Figure 3. Kaplan-Meier curves for overall survival analysis of patients with left-sided CRC according to expression levels of tumor-infiltrating (A) NK cells, (B) Neutrophils, (C) Dendritic cells, (D) CD8+ T cells, (E) CD4+ T cells, and (F) B cells. CRC, Colorectal Cancer; NK, Natural Killer.



Supplementary Figure 4. Kaplan-Meier curves for overall survival analysis of 15 prognosis-related differentially expressed genes (DEGs). (A) *CCL4*, (B) *CD38*, (C) *CD274*, (D) *CTLA4*, (E) *CXCL1*, (F) *CXCL10*, (G) *CXCL13*, (H) *ICAM1*, (I) *IFNG*, (J) *IL1A*, (K) *IL1B*, (L) *IL2RA*, (M) *PTPRC*, (N) *TLR8* and (O) *HLA-DPA1*.



Supplementary Figure 5. ROC curves of nine prognosis-related genes used for risk score calculation. (A) *CD274*, (B) *CTLA4*, (C) *CXCL1*, (D) *CXCL10*, (E) *HLA-DPA1*, (F) *IFNG*, (G) *IL1A*, (H) *IL1B*, and (I) *TLR8*. ROC, Receiver Operating Characteristic.