

# IRF1-mediated immune cell infiltration is associated with metastasis in colon adenocarcinoma

Yao-jian Shao, MD<sup>a</sup>, Jun-jie Ni, MD<sup>a</sup>, Shen-yu Wei, MD<sup>a</sup>, Xiong-peng Weng, MD<sup>b</sup>, Meng-die Shen, MD<sup>a</sup>, Yi-xin Jia, MD<sup>a</sup>, Li-na Meng, MD<sup>c,d,\*</sup>

## Abstract

**Background:** Evidence suggests that metastasis is chiefly responsible for the poor prognosis of colon adenocarcinoma (COAD). The tumor microenvironment plays a vital role in regulating this biological process. However, the mechanisms involved remain unclear. The aim of this study was to identify crucial metastasis-related biomarkers in the tumor microenvironment and investigate its association with tumor-infiltrating immune cells.

**Methods:** We obtained gene expression profiles and clinical information from The Cancer Genome Atlas database. According to the “Estimation of STromal and Immune cells in Malignant Tumor tissue using Expression data” algorithm, each sample generated the immune and stromal scores. Following correlation analysis, the metastasis-related gene was identified in The Cancer Genome Atlas database and validated in the GSE40967 dataset from Gene Expression Omnibus. The correlation between metastasis-related gene and infiltrating immune cells was assessed using the Tumor Immune Estimation Resource database.

**Results:** The analysis included 332 patients; the metastatic COAD samples showed a low immune score. Correlation analysis results showed that interferon regulatory factor 1 (IRF1) was associated with tumor stage, lymph node metastasis, and distant metastasis. Furthermore, significant associations between IRF1 and CD8+ T cells, T cell (general), dendritic cells, T-helper 1 cells, and T cell exhaustion were demonstrated by Spearman's correlation coefficients and *P* values.

**Conclusions:** The present findings suggest that IRF1 is associated with metastasis and the degree of immune infiltration of CD8+ T cells (general), dendritic cells, T-helper 1 cells, and T cell exhaustion in COAD. These results may provide information for immunotherapy in colon cancer.

**Abbreviations:** COAD = colon adenocarcinoma, ESTIMATE = Estimation of STromal and Immune cells in Malignant Tumor tissues using Expression data, MRG = metastasis-related gene, TME = tumor microenvironment.

**Keywords:** colon adenocarcinoma, immune cell infiltration, IRF1, metastasis

Editor: Bülent Kantarçeken.

This study was supported by the National Natural Science Foundation of China (No. 81373877).

The authors have no conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are publicly available.

<sup>a</sup> First College of Clinical Medical, Zhejiang Chinese Medical University, Hangzhou, <sup>b</sup> Second College of Clinical Medical, Wenzhou Medical University, Wenzhou, <sup>c</sup> Department of Gastroenterology, the First Affiliated Hospital of Zhejiang Chinese Medical University, <sup>d</sup> Key Laboratory of Digestive Pathophysiology of Zhejiang Province, Hangzhou, Zhejiang, PR China.

\* Correspondence: Li-na Meng, Department of Gastroenterology, the First Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou 310006, PR China (e-mail: mln6713@163.com).

Copyright © 2020 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Shao Yj, Ni Jj, Wei Sy, Weng Xp, Shen Md, Jia Yx, Meng Ln. IRF1-mediated immune cell infiltration is associated with metastasis in colon adenocarcinoma. *Medicine* 2020;99:37(e22170).

Received: 8 April 2020 / Received in final form: 9 August 2020 / Accepted: 13 August 2020

<http://dx.doi.org/10.1097/MD.00000000000022170>

## 1. Introduction

Colorectal cancer is a common malignancy of the digestive system, with approximately 147,950 new cases and 53,200 deaths reported in United States of America in 2020.<sup>[1]</sup> Colon adenocarcinoma (COAD) is the predominant pathological classification of colorectal cancer, accounting for >90% of cases.<sup>[2]</sup> The stage of patients at diagnosis is an important predictor of prognosis. The 5-year survival rate ranges from 91% (for patients diagnosed with stage I disease) to 12% (for those diagnosed with stage IV disease).<sup>[3]</sup> The limited effectiveness of treatments in patients with metastatic is responsible for the high mortality rates observed. Accordingly, it is important to investigate the etiopathogenesis of metastasis and develop effective therapies for COAD with metastasis.

Accumulating evidence indicates that the tumor microenvironment (TME) regulates numerous facets of tumor progression, including cell proliferation and metastasis.<sup>[4,5]</sup> The TME is a dynamic system composed of immune cells, stromal cells, cytokines, and extracellular matrix. Among them, immune and stromal cells play indispensable roles as primary producers of cytokines and chemokines.<sup>[6]</sup> Different degrees of infiltration of immune cells can alter tumor behavior. The antigen-presenting

cells with PD-L1 positivity are able to promote the metastasis of tumor cells by impairing the cytotoxicity PD-L1 against CD8+ T cells.<sup>[7,8]</sup> T cell chemokines (such as C-C motif chemokine ligand 3 [CCL3], C-X-C motif chemokine ligand 9 [CXCL9], and CXCL10) and proinflammatory cytokines (such as interleukin 12 [IL12], tumor necrosis factor alpha [TNF $\alpha$ ], and granulocyte-macrophage colony stimulating factor [GM-CSF]) secreted by CD8+ T cells are capable of killing tumor cells.<sup>[9,10]</sup> CD8+ T cells can be activated by neutrophils in the “N1” polarization state to enhance the anti-tumor effect. Nevertheless, the anti-tumor immunity is impaired by neutrophils in the “N2” polarization state, such as the action of pathological Notch signal secretion and regulatory T cell recruitment.<sup>[10–12]</sup> Fibroblasts are the main representative components of stromal cells. Activin A secreted by fibroblasts is a powerful pro-metastatic cytokine, which induces the migration of epithelial cells.<sup>[13]</sup> Fibroblast-derived CXCL12 alters PI3K/AKT signaling transduction to influence the aggressive spread of COAD.<sup>[14]</sup> Recently, immunotherapeutic strategies, such as programmed cell death-1/programmed cell death ligand-1 (PD-1/PD-L1) have been developed and were shown to be effective against non-small cell lung cancer,<sup>[15]</sup> melanoma,<sup>[16]</sup> kidney cancer,<sup>[17]</sup> and microsatellite instability colorectal cancer.<sup>[18]</sup> However, weak responses to treatment with PD-1/PD-L1 checkpoint inhibitors persist in other subtypes of colorectal cancer.

The Estimation of STromal and Immune cells in Malignant Tumor tissues using Expression data (ESTIMATE) algorithm can help to reveal the infiltration degree of immune and stromal cells in the TME.<sup>[19]</sup> The immune and stromal score of each sample represents the infiltration degree of immune and stromal cells. The ESTIMATE algorithm has been employed in glioblastoma,<sup>[20]</sup> bladder cancer,<sup>[21]</sup> and renal clear cell carcinoma.<sup>[22]</sup>

In this study, we obtained 332 complete gene expression profiles of COAD from The Cancer Genome Atlas (TCGA) database and calculated the immune and stromal scores based on the ESTIMATE algorithm. The tumor metastasis-related genes (MRGs) were extracted following verification in the Gene Expression Omnibus (GEO) database. Based on 2 dataset, interferon regulatory factor 1 (IRF1) was differentially expressed in different stages of COAD, and was considered the MRG in COAD. IRF1 is a member of the interferon regulatory factor family; the protein encoded by this gene is considered a tumor suppressor by stimulating the immune response against tumors. Previous research studies have reported that IRF1 can affect the maturation and activity of natural killer cells, development of T helper 1 (Th1) cells, and maturation of CD8+ T-cells. Therefore, the role of IRF1 in the COAD immune microenvironment was further explored in the Tumor IMMune Estimation Resource (TIMER) site.

## 2. Materials and methods

### 2.1. Database

The dataset including the gene expression profile (RNA, Seq V3) and corresponding clinical information of 385 patients with COAD was downloaded from TCGA (<https://gdc.nci.nih.gov/>). After excluding patients with indistinct clinical information, we included 332 samples in the analysis. The ESTIMATE algorithm was applied to calculate the immune/stromal scores of each sample. We analyzed the distribution of immune/stromal scores from 3 different groups: tumor stage, lymph node metastasis (N),

and distant metastasis (M) using the R package. Tumors were classified into stages I, II, III, and IV; lymph node metastases were divided into N0 and N1–2 groups; and distant metastases were divided into M0 and M1 groups.

### 2.2. Identification of differentially expressed genes (DEGs)

According to the immune/stromal scores of the 332 samples, we classified them into high- and low-score groups. The R package Limma was utilized to screen DEGs. Absolute values of log fold-change ( $\log_{2}FC$ )  $\geq 1$  and false discovery rate  $< .05$  were set as thresholds to filter DEGs.

### 2.3. Construction of a protein-protein interaction (PPI) network

We collected information on proteins encoded by DEGs and constructed a PPI network with the highest confidence (.95) interaction score in the STRING database (<http://string-db.org>). Subsequently, the PPI network was imported into the Cytoscape software for reestablishment. The MCODE plugin of Cytoscape was applied to identify the significant modules, which were the densely connected regions of the PPI network.

### 2.4. Enrichment analyses of function and pathway

Functional and pathway analyses of DEGs contained in significant modules were performed using the R package clusterProfiler. The Gene Ontology (GO) includes 3 sub ontologies, namely biological process (BP), cellular component (CC), and molecular function (MF). The adjusted  $P$  value  $< .05$  was set as the cutoff value for GO term and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. The top 10 terms of each sub ontology are shown in the bubble diagram.

### 2.5. Metastasis correlation analysis and verification

Metastasis correlation analysis was performed using the R package beeswarm and a boxplot was used to exhibit the expression of DEGs in the tumor stage group, lymph node metastasis group and distant metastasis group.  $P$  values  $< .05$  denoted statistical significance. We adopted the GSE40967 dataset from the GEO database to verify the characteristics of DEG expression using the same approach. The GSE40967 dataset contains the gene expression profiles and clinical information of 1048 patients with COAD based on the GPL570 platforms (Affymetrix Human Genome U133 Plus 2.0 Array). Finally, only 560 samples with complete TNM stage clinical information were selected for the correlation analysis. The DEGs with significant differences in TCGA database and GEO database were considered MRGs. The prognostic value of MRGs was explored using the Gene Expression Profiling Interactive Analysis (GEPIA; <http://gepia.cancer-pku.cn/>) online tool, and was shown via a Kaplan–Meier survival curve.

### 2.6. Immune infiltration

The TIMER database (<https://cistrome.shinyapps.io/timer/>) was utilized to assess the correlation of the expression of MRGs with different types of immune infiltrates. In the gene module, we analyzed the correlation of MRG expression with tumor-infiltrating immune cells of COAD. In addition, we investigated the correlation between MRG expression and gene markers of

diverse immune cells, which were referenced in previous studies via a correlation module.<sup>[2,3]</sup> We classified infiltrating immune cells in detail according to their different functions. The macrophages were subdivided into tumor-associated, M1, and M2 macrophages. T cells were subdivided into CD8+ T, T (general), Th1, Th2, follicular T helper, Th17, regulatory T cells, and T cell exhaustion. The scatter plots together with Spearman's correlation coefficients and *P* values were generated in both modules to show the associations. The expression of MRGs and levels of gene markers in scatter plots were detected using log2 RSEM. The absolute value of Spearman's correlation coefficient determined the strength of the correlation based on the following guideline (weak: <.40; moderate: .40-.59; strong: >.60). *P* values <.05 denoted statistical significance.

**2.7. Ethics statement**

The datasets in this study are achieved from publicly available database (TCGA and GEO). Ethical approval and informed consent were thus not required.

**3. Results**

**3.1. Immune cell infiltrations are correlated with the clinical characteristics of COAD**

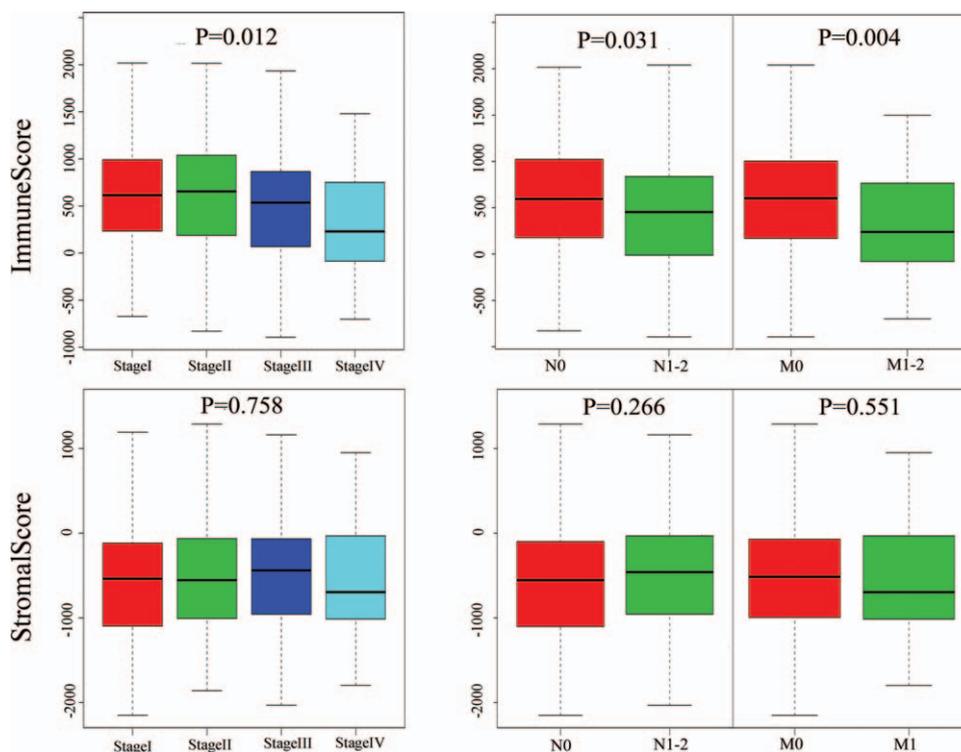
A total of 332 COAD samples with specific gene expression and clinical information were included in this study. The stages of disease for these patients were: stage I (n=58, 17.5%), stage II (n=135, 40.6%), stage III (n=86, 25.9%), and stage IV (n=53, 16.0%). The details of TNM staging and other characteristics are

**Table 1**

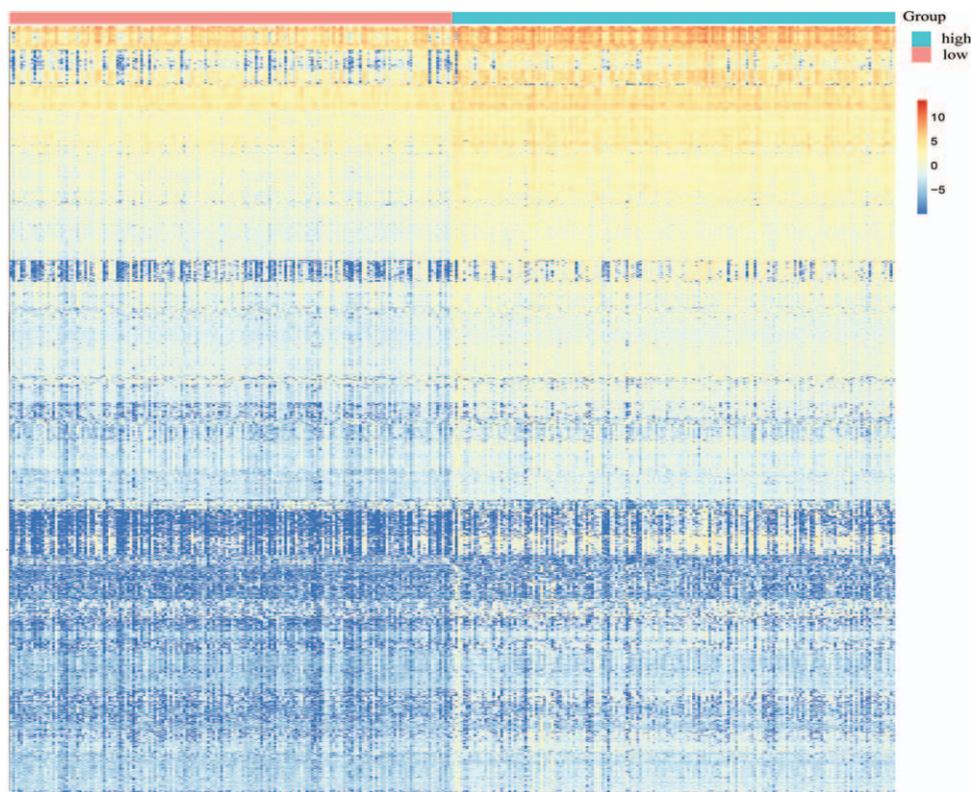
**Characteristics of COAD patients in TCGA database.**

Characteristics	TCGA Data	Characteristics	TCGA data
Age		Tumor	
<=65	133	T1	7
>65	199	T2	58
Gender		T3	231
Male	178	T4	36
Female	154		
Vital status		Nodes	
Alive	274	N0	200
Death	58	N1	75
Stage		Metastasis	
Stage I	58	N2	57
Stage II	135	M0	279
Stage III	86	M1	53
Stage IV	53		

listed in Table 1. We calculated the immune and stromal scores of 332 samples by employing the ESTIMATE algorithm. The distribution of the immune and stromal scores was between -893.55 to 2417.00 and -2151.29 to 1696.53, respectively. To explore the value of infiltrating immune and stromal cells in COAD, we assessed the association between immune/stromal scores and tumor stage, lymph node metastasis, and distant metastasis using the R package. As shown in Figure 1, there was no statistically significant difference found in the stromal scores, whereas the immune score was markedly correlated with stage (*P*=.012), lymph node metastasis (*P*=.031) and distant metastasis (*P*=.004).



**Figure 1.** Clinical correlation analysis of the immune score and stromal score in COAD. The immune score showed significant correlations with stage (*P*=.012), lymphatic metastasis (*P*=.031), and distant metastasis (*P*=.004); however, there was no statistically significant difference found in the stromal scores. COAD = colon adenocarcinoma.



**Figure 2.** Heatmap of the immune scores of 1,655 DEGs (1,608 upregulated DEGs and 47 downregulated DEGs). DEG = differentially-expressed gene.

metastasis ( $P=.004$ ). The distribution of the immune score was as follows: stages I and II > stage III > stage IV. The preliminary investigation revealed that the patients with lymph node metastasis and distant metastasis generated lower immune scores. Subsequently, we selected the immune score for the following studies.

### 3.2. DEG screening based on the immune score

To determine the potential association of immune cell infiltration with gene expression profiles, all COAD samples were divided into 2 different groups by setting the median of the immune score as the threshold value. The differential expression analysis between high and low immune scores was performed using the R package Limma. According to the cutoff criteria (false discovery rate  $<.05$  and  $|\log_{2}FC| \geq 1.0$ ), we obtained 1,655 DEGs (1,608 upregulated DEGs and 47 downregulated DEGs). Figure 2 illustrates the heatmap of DEGs.

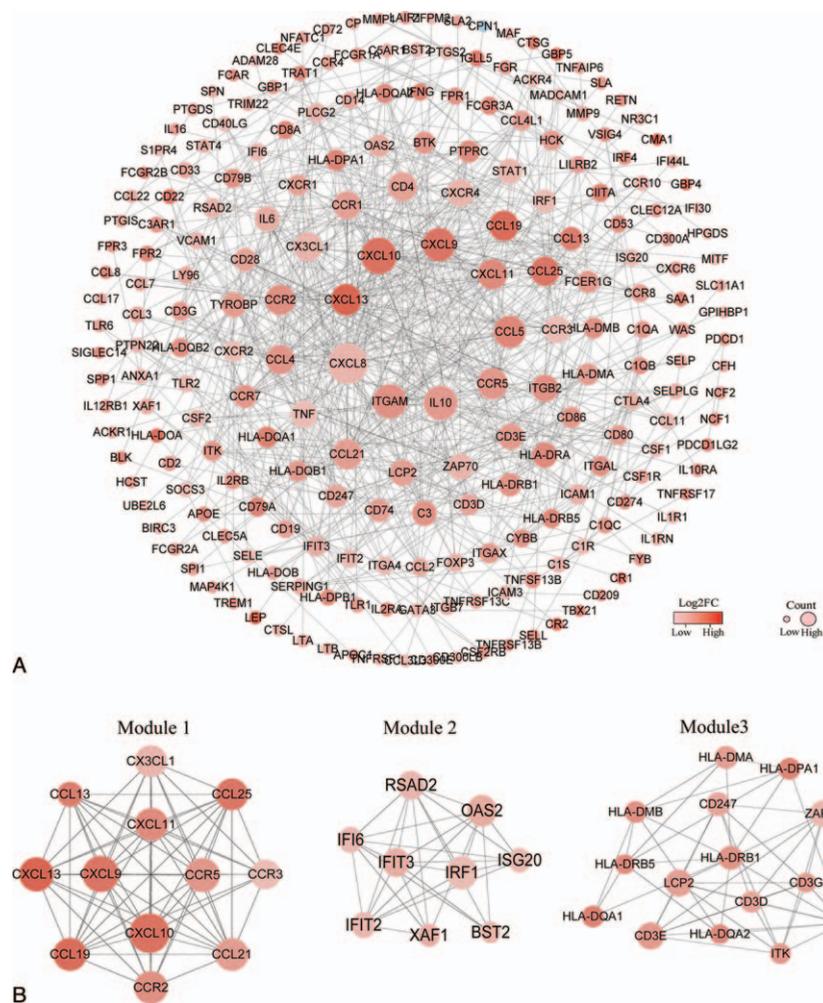
### 3.3. Three significant modules from the PPI network

The PPI network of DEGs was constructed using the STRING database and Cytoscape software (Version 3.7.1). After removing the outlying genes without interaction, the network contained 222 nodes and 674 edges. Node color indicated the  $\log_{2}FC$  of gene expression and the node size increased based on the count of interacting proteins (Fig. 3A). Subsequently, the MCODE plugin of Cytoscape was utilized to identify significant modules of the PPI network; the top 3 modules were identified and termed Modules 1, 2, and 3 for convenience. Module 1 consisted of 12 nodes and 63 edges, including CXCL9, CXCL13, CXCL11,

CXCL10, CX3CL1, CCR5, CCR3, CCR2, CCL25, CCL21, CCL19, and CCL13. Module 2 contained 29 edges involving 9 nodes: bone marrow stromal cell antigen 2 (BST2), interferon alpha inducible protein 6 (IFI6), interferon induced protein with tetratricopeptide repeats 2 (IFIT2), IFIT3, IRF1, interferon stimulated exonuclease gene 20 (ISG20), 2'-5'-oligoadenylate synthetase 2 (OAS2), radical S-adenosyl methionine domain containing 2 (RSAD2), and XIAP associated factor 1 (XAF1). In Module 3, 46 edges were formed in the network involving 15 nodes, namely CD247, CD3D, CD3E, CD3G, CD4, HLA-DMA, HLA-DMB, HLA-DPA1, HLA-DQA1, HLA-DQA2, HLA-DRB1, HLA-DRB5, IL2 inducible T cell kinase (ITK), lymphocyte cytosolic protein 2 (LCP2), and zeta chain of T cell receptor associated protein kinase 70 (ZAP70). The DEGs contained in the 3 modules were all upregulated in the high immune score group (Fig. 3B).

### 3.4. GO term and KEGG pathway enrichment analyses

Potential biological functions of 36 upregulated DEGs in the high immune score group were investigated using the R package clusterProfiler. The GO enrichment analyses consisted of 3 subontologies: BP, CC, and MF. For the BP category, 36 DEGs were significantly enriched in the T cell receptor signaling pathway, response to interferon-gamma, and antigen receptor-mediated signaling pathway (Fig. 4A). The 36 DEGs in the CC category primarily clustered in the external side of the plasma membrane (Fig. 4B). Regarding the MF category, DEGs were mainly involved in chemokine receptor binding, G protein-coupled receptor binding, and cytokine receptor binding (Fig. 4C). Moreover, the result of the KEGG pathway analysis



**Figure 3.** (A) After removing outlying genes without interaction, the PPI network consisted of 222 DEGs. Node color indicates the log<sub>2</sub>FC of gene expression and node size increased based on the count of interacting proteins. (B) Top 3 significant modules of the PPI network. Module 1 consisted of 12 nodes and 63 edges. Module 2 consisted of 9 nodes and 29 edges. Module 3 consisted of 15 nodes and 46 edges. DEG = differentially-expressed gene, FC = fold change, PPI = protein-protein interaction.

suggested that 36 DEGs were significantly related to immune response, such as Th1 and Th2 cell differentiation, Th17 cell differentiation, chemokine signaling pathway, etc. (Fig. 4D).

**3.5. IRF1 expression associated with metastasis**

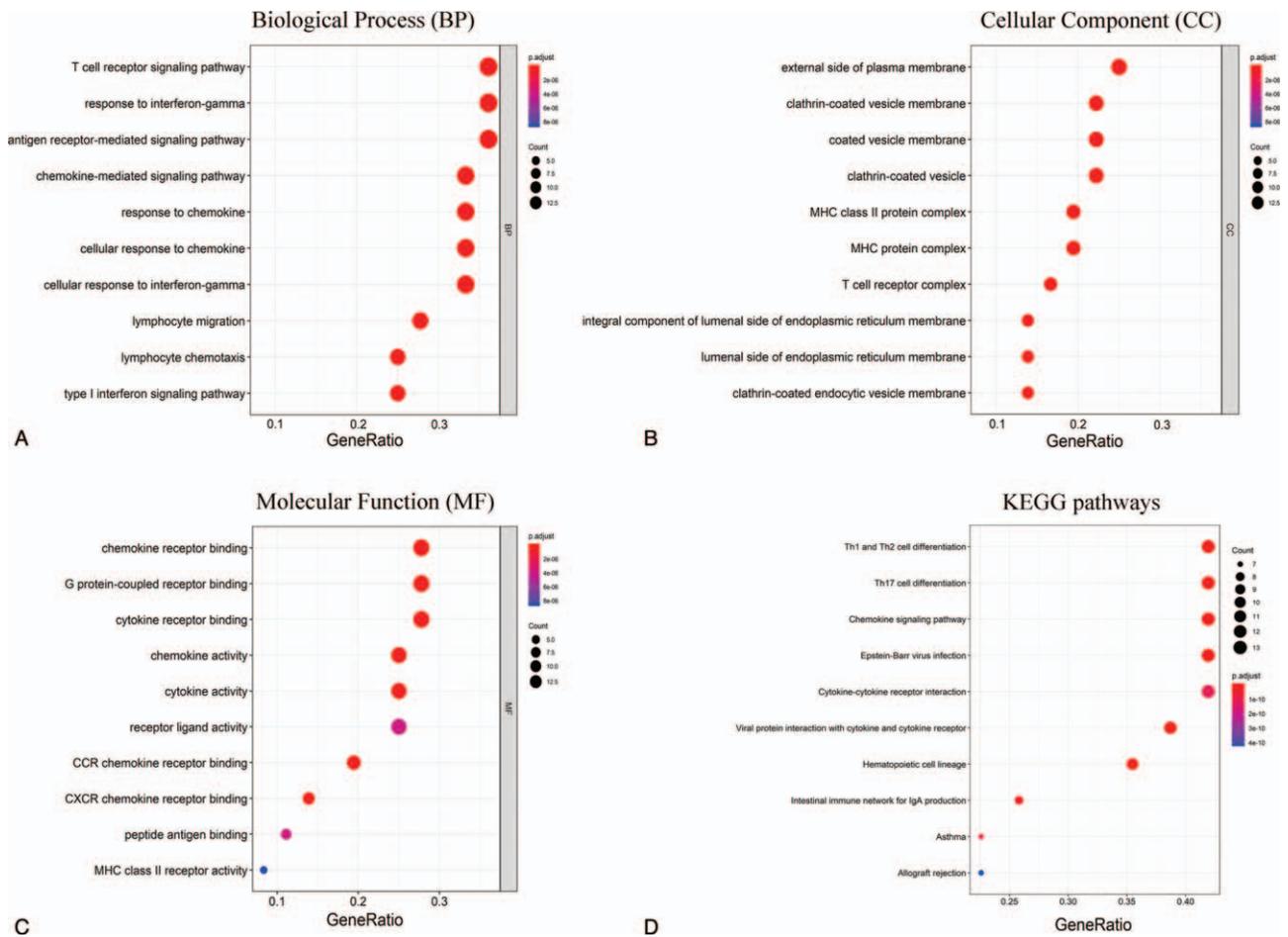
Correlation analysis was performed between 36 upregulated DEGs and clinical characteristics, including tumor stage, lymph node metastasis, and distant metastasis using the R package beeswarm. A total of 12 DEGs showed significant correlation in TCGA database (Table 2). Subsequently, we extracted the gene expression of 12 DEGs from the GSE40967 dataset to verify the reproducibility of its correlation with clinical characteristics. As shown in Table 3, the expression of CD3D, CD3G, CXCL9, and IRF1 were significantly associated with tumor stage and distant metastasis. The expression of IRF1 may be associated with lymph node metastasis ( $P = .052$ ); there were no significant correlations found between the expression of CD3D, CD3G, and CXCL9 and lymph node metastasis in the GSE40967 dataset (Fig. 5A, B). Further Kaplan–Meier survival curves were constructed to investigate the prognostic value of these 4 genes based on

GEPIA. The survival analysis indicated that patients with higher levels of IRF1 expression had a longer overall survival than those with lower levels. There were no statistically significant differences found in the CXCL9, CD3D, and CD3G groups (Fig. 5C).

**3.6. Immune infiltration analysis of IRF1**

The association between IRF1 expression and immune infiltrates was investigated in the TIMER database. The results demonstrated that IRF1 expression was significantly positively correlated with the infiltrating degree of B cells ( $r = .111, P = 2.56e-02$ ), CD8+ T cells ( $r = .229, P = 3.08e-06$ ), neutrophils ( $r = .501, P = 6.52e-27$ ), and dendritic cells (DC) ( $r = .406, P = 2.25e-17$ ). There was a negative correlation between IRF1 expression and tumor purity ( $r = .258, P = 1.24e-07$ ). In addition, there were no significant correlations between IRF1 expression and infiltrating degree of CD4+ T cells and macrophages.

In the correlation module, we focused on exploring the correlations between IRF1 and immune marker gene sets of diverse immune infiltrating cells. Following adjustment by purity, we discovered that IRF1 expression was strongly correlated with



**Figure 4.** Top 10 terms of the GO and KEGG pathway. (A) Bubble diagram of the biological process sub ontologies. (B) Bubble diagram of the cell component sub ontologies. (C) Bubble diagram of the molecular function sub ontologies. (D) Bubble diagram of the top 10 terms of the KEGG pathway. GO = gene ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes.

CD8A of CD8+ T cells, CD3D, CD3E, CD2 of T cell (general), T-box transcription factor 21 (TBX21), signal transducer and activator of transcription 1 (STAT1), and interferon gamma (IFNG) of Th1. The high infiltration degree of CD8+ T cells, T cell (general), and Th1 generally indicated that these are anti-tumor

factors in COAD (Table 4). In addition, the correlations between IRF1 expression and HLA-DRA, HLA-DPA1 of DCs, and PD-1, lymphocyte activating gene 3 (LAG3) of T cell exhaustion were strong. PD-1 and LAGs play important roles in the progression of tumors. Therefore, the results of the correlation analysis further

**Table 2**  
Based on TCGA, 12 DEGs significantly associated with metastasis.

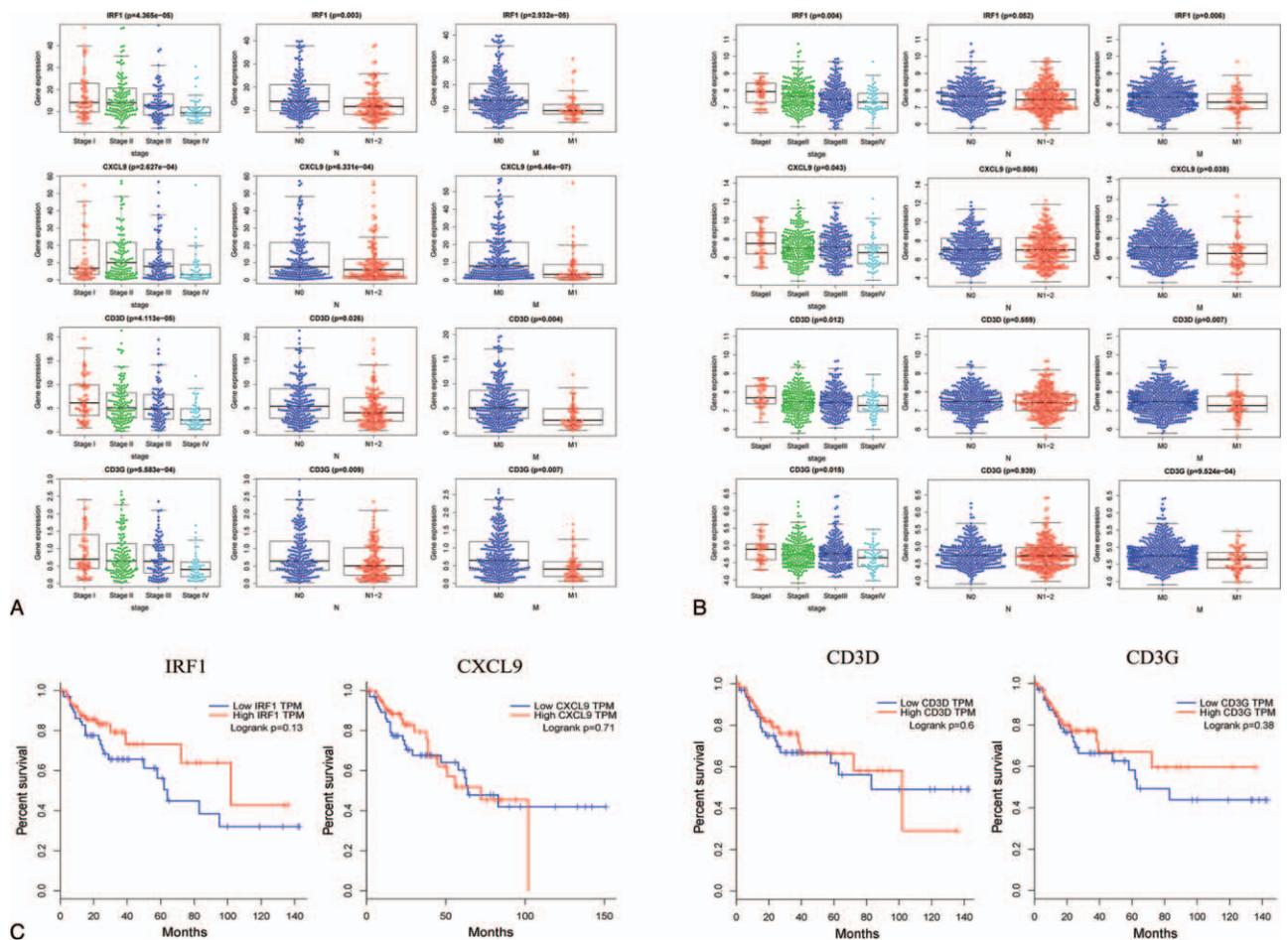
Gene	Stage		N		M	
	Cor	P value	Cor	P value	Cor	P value
CCL25	10.483	.015	2.873	.004	3.269	.001
CCR5	12.998	.005	2.104	.036	3.736	3.11E-04
CD3D	22.962	4.11E-05	2.233	.026	2.989	0.004
CD3G	17.497	5.58E-04	2.624	.009	2.755	0.007
CXCL9	19.084	2.63E-04	3.452	6.33E-04	5.16	6.46E-07
CXCL10	17.59	5.34E-04	2.774	.006	4.328	2.13E-05
CXCL11	16.185	.001	2.617	.009	3.164	0.002
IRF1	22.838	4.37E-05	2.998	.003	4.359	2.93E-05
HLA-DMA	11.092	.011	2.082	.038	3.368	9.59E-04
HLA-DMB	10.023	.018	2.351	.019	3.901	1.49E-04
HLA-DRB1	9.963	.019	2.194	.029	3.715	2.95E-04
HLA-DRB5	10.863	.012	2.814	.005	3.02	.003

Cor = correlation coefficient, M = distant metastasis, N = lymph node metastasis.

**Table 3**  
**Correlation between 12 DEGs and clinical characteristics in GSE40967 dataset.**

Gene	stage		N		M	
	Cor	P value	Cor	P value	Cor	P value
CCL25	7.865	.049	1.688	.092	0.632	.528
CCR5	4.862	.182	-0.456	.648	0.81	.421
CD3D	10.94	.012	0.585	.559	2.767	.007
CD3G	10.497	.015	0.077	.939	3.432	9.52E-04
CXCL10	4.132	.248	-0.253	.8	2.059	.043
CXCL11	2.955	.399	0.573	.567	2.372	.02
CXCL9	8.151	.043	0.246	.806	2.114	.038
IRF1	13.221	.004	1.945	.052	2.832	.006
HLA-DMA	4.607	.203	0.839	.402	1.656	.102
HLA-DMB	3.573	.311	-0.131	.896	1.548	.126
HLA-DRB1	5.47	.14	-1.164	.245	-2.225	.029
HLA-DRB5	3.722	.293	0.632	.527	1.405	.164

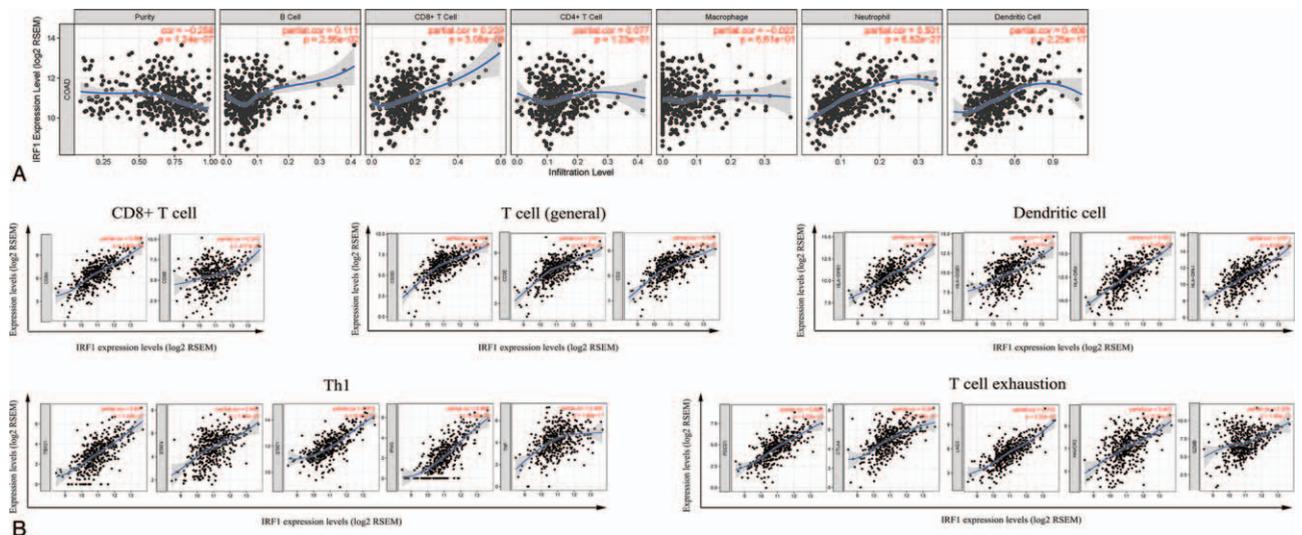
Cor = correlation coefficient, M = distant metastasis, N = lymph node metastasis.



**Figure 5.** Verification of the association between 4 DEGs and metastasis. (A) Boxplots of IRF1, CXCL9, CD3D, and CD3G expression in different groups according to tumor stage, lymphatic metastasis, and distant metastasis based on TCGA database. (B) Boxplots of IRF1, CXCL9, CD3D, and CD3G expression based on the GEO database. CD3D, CD3G, CXCL9, and IRF1 showed significant difference in the tumor stage and distant metastasis groups. IRF1 expression may be associated with lymph node metastasis ( $P = .052$ ) (C) Kaplan-Meier survival curve of IRF1, CXCL9, CD3D, and CD3G based on the GEPIA. The group with high levels of IRF1 expression showed a longer overall survival than that with low IRF1 expression levels; however, there was no statistically significant difference found in the CXCL9, CD3D, and CD3G groups. CXCL9 = C-X-C motif chemokine ligand 9, DEG = differentially-expressed gene, GEO = Gene Expression Omnibus, IRF1 = interferon regulatory factor 1.

**Table 4****Correlation analysis between IRF1 expression and marker gene sets of immune infiltrating cells in TIMER database.**

Description	Gene mraket	IRF1			
		None		Purity	
		Cor	P	Cor	P
CD8+ T cell	CD8A	0.684	0	0.666	1.83E-53
	CD8B	0.29	2.86E-10	0.247	4.71E-07
T cell (general)	CD3D	0.677	0	0.655	3.76E-51
	CD3E	0.65	0	0.611	5.97E-43
	CD2	0.641	0	0.606	3.97E-42
B cell	CD19	0.338	1.07E-13	0.273	2.33E-08
	CD79A	0.34	7.52E-14	0.262	8.3E-08
Monocyte	CD86	0.491	3.79E-29	0.443	5.94E-21
	CD115 (CSF1R)	0.405	0	0.352	2.83E-13
TAM	CCL2	0.277	1.85E-09	0.224	5.38E-06
	CD68	0.417	0	0.367	2.24E-14
	IL10	0.381	2.62E-17	0.358	1.08E-13
M1 Macrophage	INOS (NOS2)	0.476	0	0.461	1E-22
	IRF5	0.175	.000174	0.167	.000747
	COX2(PTGS2)	0.238	2.66E-07	0.229	3.02E-06
M2 Macrophage	CD163	0.414	2.05E-20	0.359	8.58E-14
	VSIG4	0.364	9.17E-16	0.315	8.98E-11
	MS4A4A	0.405	1.49E-19	0.347	6.41E-13
Neutrophils	CD66b (CEACAM8)	-0.014	.769851	0.036	.474654
	CD11b (ITGAM)	0.371	1.64E-16	0.299	8.14E-10
Natural killer cell	CCR7	0.451	0	0.391	2.8E-16
	KIR2DL1	0.311	1E-11	0.27	3.18E-08
	KIR2DL3	0.38	3.62E-17	0.331	7.27E-12
	KIR2DL4	0.582	7.69E-43	0.558	1.33E-34
	KIR3DL1	0.407	1.1E-19	0.376	4.53E-15
	KIR3DL2	0.397	8.86E-19	0.37	1.39E-14
	KIR3DL3	0.214	.000004	0.193	8.89E-05
	KIR2DS4	0.303	3.6E-11	0.269	3.57E-08
	HLA-DPB1	0.614	0	0.578	1.3E-37
	HLA-DQB1	0.502	0	0.457	2.41E-22
Dendritic cell	HLA-DRA	0.678	0	0.653	8.43E-51
	HLA-DPA1	0.637	0	0.611	6.95E-43
	BDCA-1(CD1C)	0.162	.000515	0.078	.117854
	BDCA-4(NRP1)	0.25	6.55E-08	0.168	.000692
	CD11c (ITGAX)	0.445	0	0.395	1.25E-16
	T-bet (TBX21)	0.669	8.59E-61	0.637	1.44E-47
	STAT4	0.502	0	0.458	1.93E-22
Th1	STAT1	0.685	0	0.673	6.26E-55
	IFN- $\gamma$ (IFNG)	0.693	8.24E-67	0.696	5.25E-60
	TNF- $\alpha$ (TNF)	0.433	2.15E-22	0.405	1.89E-17
	GATA3	0.353	6.44E-15	0.303	4.57E-10
	STAT6	0.1	.032677	0.134	.006803
	STAT5A	0.334	2.82E-13	0.283	6.78E-09
Th2	IL13	0.276	2.02E-09	0.26	1.13E-07
	BCL6	0.268	6.67E-09	0.217	1.01E-05
	IL21	0.298	7.2E-11	0.283	6.33E-09
Th17	STAT3	0.307	2.53E-11	0.269	3.59E-08
	IL17A	0.22	2.05E-06	0.238	1.23E-06
Treg	FOXP3	0.491	0	0.44	1.12E-20
	CCR8	0.388	7.04E-18	0.334	4.78E-12
	STAT5B	-0.151	.001159	-0.187	.000153
	TGF $\beta$ (TGFB1)	0.371	1.44E-16	0.303	4.72E-10
	PD-1 (PDCD1)	0.68	1.51E-63	0.665	4.25E-53
T cell exhaustion	CTLA4	0.564	8.12E-40	0.541	3.32E-32
	LAG3	0.713	0	0.705	3.2E-62
	TIM-3 (HAVCR2)	0.496	0	0.447	2.23E-21
	GZMB	0.34	1.02E-13	0.329	1E-11



**Figure 6.** Correlation of IRF1 expression with immune cells in COAD. (A) Scatter plots of the gene module. IRF1 expression levels showing moderate positive correlations with the infiltration levels of neutrophils and dendritic cells. (B) Scatter plots of the correlation module. IRF1 expression showing strong positive correlations with some gene markers of CD8+ T, T (general), Th1 cells, dendritic cells, and T cell exhaustion. COAD = colon adenocarcinoma, IRF1 = interferon regulatory factor 1, Th1 = T helper 1.

confirmed the crucial role of IRF1 in immune infiltration (Table 4). The scatter plots show each pair of IRF1-marker genes with a strong correlation (Fig. 6).

#### 4. Discussion and conclusions

Metastasis of COAD to external organs is responsible for the death of most patients with this disease; hence, metastasis is linked to poor prognosis for patients.<sup>[24]</sup> As new technologies are applied to investigate metastasis, the underlying mechanisms are becoming clear. Tumor metastasis is not always caused by primary tumor cells alone; the microenvironment surrounding primary tumor cells plays a critical role.<sup>[25–27]</sup> Experimental evidence suggested that the TME of progressive tumors exhibits different characteristics from that of early tumors in COAD. The activated transforming growth factor- $\beta$  (TGF- $\beta$ ) signaling in TME drives an important mechanism of immune evasion, which accelerates T cell exhaustion and inhibits acquisition of the Th1 effector phenotype.<sup>[28]</sup> Blockade of TGF- $\beta$  signaling unleashes a potent response of cytotoxic T cells to prevent tumor cell metastasis.<sup>[29]</sup> Increased growth differentiation factor 15 (GDF15) in the TME represents a factor of tumor cell invasion and metastasis.<sup>[30]</sup> In this study, we aimed to identify MRGs by analyzing the TME characteristics of different stages in COAD.

According to the distribution of the immune score in 3 groups, metastatic COAD had a lower immune score than primary COAD. The distribution characteristic of the immune score indicated that there was a powerful immune evasion in the TME of metastatic COAD; however, the factors involved in this process remain unclear. Based on the MCODE module analysis in the PPI network, 36 upregulated DEGs in the high immune score group were identified as the hub genes which drive immune evasion in metastatic COAD. The GO term and KEGG pathway enrichment analyses revealed that these 36 DEGs were mainly concerned with the process of immunity and inflammatory responses, especially for T cell-mediated immunity, such as the T cell receptor signaling pathway (BP) and Th1 and Th2 cell

differentiation (KEGG). Correlation analysis using TCGA and GEO databases suggested that 4 DEGs (IRF1, CD3D, CD3G, and CXCL9) correlated with tumor stage and distant metastasis. The Kaplan–Meier survival curves of these 4 DEGs in the GEPIA showed that upregulation of IRF1 was a protective factor for the prognosis of patients with COAD. The aforementioned results indicated that the expression of IRF1 may serve as an important function in regulating immune evasion.

Indeed, as a member of the interferon regulatory factor family, IRF1 is associated with immune cell infiltration in COAD. Studies have reported that the deletion of IRF1 altered the type and function of immune cells in chronic inflammation of the colon, which increase the susceptibility of colitis-associated colon cancer.<sup>[31]</sup> The activated IRF1 recruited monocytes and macrophages to enhance the anti-tumor immune response by inducing interferon- $\beta$  autocrine signaling.<sup>[32]</sup> However, the mechanisms of IRF1 in immune cell infiltration are not well understood. Further correlation analysis using the TIMER database showed that IRF1 was positively correlated with CD8+ T cells, T cells (general), and Th1 cells. A previous study confirmed that patients with melanoma who were refractory to adoptive T cell therapy had reduced IRF1 expression levels.<sup>[33]</sup> IRF1 influences the antitumor efficacy of cyclophosphamide by regulating the amplification of Th1.<sup>[34]</sup> In contrast to Th1, Th17 cells, which were weakly correlated with IRF1, exerted opposite effects on the survival of patients with colorectal cancer.<sup>[35,36]</sup> The aforementioned findings were consistent with the results of the correlation analysis, which suggested that IRF1 served as an antitumor factor in COAD. Although there was no direct evidence that IRF1 was correlated with HLA-DRA or HLA-DPA1 of DCs, the correlation between IRF1 and major histocompatibility complex class I (MHC-I) molecules has been reported. In aggressive neuroblastoma, IRF1 and nuclear factor- $\kappa$ B rescued the immune escape phenotype by restoring the pathway of MHC class I-restricted tumor antigen processing and presentation to cytotoxic T lymphocytes.<sup>[37]</sup> Certainly, with the improved understanding of tumor immunology, increasing

attention has been focused on the gene markers of the MHC-II restricted neoantigen.<sup>[38]</sup> These correlation between IRF1 and HLA-DRA and HLA-DPA1 indicated a differential perspective of tumor immunology.

The correlation analysis also yielded another unexpected result. IRF1 showed a strong correlation with PD-1 and LAG3 of T cell exhaustion, which are both tumor-promoting factors in the TME.<sup>[39–41]</sup> These results indicated a differential function of IRF1 in COAD. Studies have revealed that PD-1 led to T cell apoptosis, impairing protective inflammatory response and promoting tumor immune evasion.<sup>[42]</sup> Nevertheless, preliminary clinical studies of PD-1/PD-L1 checkpoint inhibitors in colon cancer showed that the effectiveness of these treatments was extremely limited except for the microsatellite instability subtype.<sup>[18,27]</sup> IRF1 is a transcription factor of the interferon receptor signaling pathways, which primarily regulate the expression of PD-L1 (a ligand of PD-1).<sup>[43]</sup> A recent experiment reported that IRF1-deficient tumors exhibited enhanced cytotoxicity of CD8+ T cells through down-regulation of PD-L1 expression.<sup>[44]</sup> The aforementioned results imply the presence of a mechanism to explain the limited efficacy of PD-1/PD-L1 checkpoint inhibitors. Further investigation is warranted to confirm this conjecture.

In conclusion, inhibition of IRF1 expression is correlated with metastasis, which is caused by immune evasion in advanced COAD. A low degree of T cell (including CD8+ T cells and Th1 cells) infiltration caused by IRF1 deficiency is mainly responsible for the immune evasion. This result increases our understanding of the TME and may provide a potential therapeutic target for immunotherapy in colon cancer.

## Acknowledgments

The authors are thankful to all participants in this study.

## Author contributions

**Conceptualization:** Lina Meng.

**Data curation:** Yaojian Shao.

**Formal analysis:** Yaojian Shao.

**Funding acquisition:** Yaojian Shao.

**Methodology:** Lina Meng.

**Investigation:** Yaojian Shao.

**Software:** Yaojian Shao

**Supervision:** Xiong-peng Wen, Meng-die Shen.

**Validation:** Lina Meng.

**Writing – original draft:** Liting Ye, Jun-jie Ni, Shen-yu Wei.

**Writing – review & editing:** Lina Meng.

## References

- [1] Siegel RL, Miller KD. Cancer statistics. *Cancer J Clin* 2020;70:7–30.
- [2] Siegel RL, Miller KD. Colorectal cancer statistics. *Cancer J Clin* 2020;70:145–64.
- [3] Miller KD, Nogueira L, Mariotto AB, et al. Cancer treatment and survivorship statistics. *Cancer J Clin* 2019;69:363–85.
- [4] Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 2013;39:782–95.
- [5] Clark AG, Vignjevic DM. Modes of cancer cell invasion and the role of the microenvironment. *Curr Opin Cell Biol* 2015;36:13–22.
- [6] Pages F, Mlecnik B, Marliot F, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet (London, England)* 2018;391: 2128–39.
- [7] Lazarus J, Maj T, Smith JJ, et al. Spatial and phenotypic immune profiling of metastatic colon cancer. *JCI Insight* 2018;3:e121932.
- [8] Lazarus J, Oneka MD, Barua S, et al. Mathematical modeling of the metastatic colorectal cancer microenvironment defines the importance of cytotoxic lymphocyte infiltration and presence of PD-L1 on antigen presenting cells. *Ann Surg Oncol* 2019;26:2821–30.
- [9] Scapini P, Lapinet-Vera JA, Gasperini S, et al. The neutrophil as a cellular source of chemokines. *Immunol Rev* 2000;177:195–203.
- [10] Fridlender ZG, Sun J, Kim S, et al. Polarization of tumor-associated neutrophil phenotype by TGF-beta: “N1” versus “N2” TAN. *Cancer Cell* 2009;16:183–94.
- [11] Mishalian I, Bayuh R, Eruslanov E, et al. Neutrophils recruit regulatory T-cells into tumors via secretion of CCL17—a new mechanism of impaired antitumor immunity. *Int J Cancer* 2014;135:1178–86.
- [12] Ruland J. Colon cancer: epithelial notch signaling recruits neutrophils to drive metastasis. *Cancer Cell* 2019;36:213–4.
- [13] Bauer J, Emon MAB. Increased stiffness of the tumor microenvironment in colon cancer stimulates cancer associated fibroblast-mediated prometastatic activin A signaling. *Sci Rep* 2020;10:50.
- [14] Ma J, Sun X, Wang Y, et al. Fibroblast-derived CXCL12 regulates PTEN expression and is associated with the proliferation and invasion of colon cancer cells via PI3k/Akt signaling. *Cell Commun Signal* 2019;17:119.
- [15] Ottonello S, Genova C, Cossu I, et al. Association between response to nivolumab treatment and peripheral blood lymphocyte subsets in patients with non-small cell lung cancer. *Front Immunol* 2020;11:125.
- [16] Liu D, Schilling B. Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nat Med* 2019;25:1916–27.
- [17] Raimondi A, Randon G, Sepe P, et al. The evaluation of response to immunotherapy in metastatic renal cell carcinoma: open challenges in the clinical practice. *Int J Mol Sci* 2019;20:4263.
- [18] Yaghoubi N, Soltani A, Ghazvini K, et al. PD-1/PD-L1 blockade as a novel treatment for colorectal cancer. *Biomed Pharmacother* 2019;110:312–8.
- [19] Yoshihara K, Shahmoradgoli M, Martinez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* 2013;4:2612.
- [20] Jia D, Li S, Li D, et al. Mining TCGA database for genes of prognostic value in glioblastoma microenvironment. *Aging* 2018;10:592–605.
- [21] Luo Y, Zeng G, Wu S. Identification of microenvironment-related prognostic genes in bladder cancer based on gene expression profile. *Front Genet* 2019;10:1187.
- [22] Liu S, Li S, Wang Y, et al. Prognostic value of infiltrating immune cells in clear cell renal cell carcinoma (ccRCC). *J Cell Biochem* 2020;121: 2571–81.
- [23] Pan JH, Zhou H, Cooper L, et al. LAYN is a prognostic biomarker and correlated with immune infiltrates in gastric and colon cancers. *Front Immunol* 2019;10:6.
- [24] Kiberstis PA. Metastasis: an evolving story. *Science (New York, N Y)* 2016;352:162–3.
- [25] Lenos KJ, Miedema DM, Lodestijn SC, et al. Stem cell functionality is microenvironmentally defined during tumour expansion and therapy response in colon cancer. *Nat Cell Biol* 2018;20:1193–202.
- [26] Triner D, Devenport SN, Ramakrishnan SK, et al. Neutrophils restrict tumor-associated microbiota to reduce growth and invasion of colon tumors in mice. *Gastroenterology* 2019;156:1467–82.
- [27] Llosa NJ, Cruise M, Tam A, et al. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. *Cancer Discov* 2015;5:43–51.
- [28] Tauriello DVF, Palomo-Ponce S, Stork D, et al. TGFbeta drives immune evasion in genetically reconstituted colon cancer metastasis. *Nature* 2018;554:538–43.
- [29] Jung B, Staudacher JJ, Beauchamp D. Transforming growth factor beta superfamily signaling in development of colorectal cancer. *Gastroenterology* 2017;152:36–52.
- [30] Ding Y, Hao K, Li Z, et al. c-Fos separation from Lamin A/C by GDF15 promotes colon cancer invasion and metastasis in inflammatory microenvironment. *J Cell Physiol* 2020;235:4407–21.
- [31] Jeyakumar T, Fodil N, Van Der Kraak L, et al. Inactivation of interferon regulatory factor 1 causes susceptibility to colitis-associated colorectal cancer. *Scientific Reports* 2019;9:18897.
- [32] Venkatesh D, Hernandez T, Rosetti F, et al. Endothelial TNF receptor 2 induces IRF1 transcription factor-dependent interferon-beta autocrine signaling to promote monocyte recruitment. *Immunity* 2013;38:1025–37.
- [33] Cascone T, McKenzie JA, Mbofung RM, et al. Increased tumor glycolysis characterizes immune resistance to adoptive T cell therapy. *Cell Metab* 2018;27:977–987.e974.

- [34] Buccione C, Fragale A, Polverino F, et al. Role of interferon regulatory factor 1 in governing Treg depletion, Th1 polarization, inflammasome activation and antitumor efficacy of cyclophosphamide. *International Journal of Cancer* 2018;142:976–87.
- [35] Doulabi H, Rastin M, Shabahang H, et al. Analysis of Th22, Th17 and CD4(+) cells co-producing IL-17/IL-22 at different stages of human colon cancer. *Biomed Pharmacother* 2018;103:1101–6.
- [36] Yoshida N, Kinugasa T, Miyoshi H, et al. A High RORgammaT/CD3 ratio is a strong prognostic factor for postoperative survival in advanced colorectal cancer: analysis of helper T cell lymphocytes (Th1, Th2, Th17 and Regulatory T Cells). *Ann Surg Oncol* 2016;23:919–27.
- [37] Lorenzi S, Forloni M, Cifaldi L, et al. IRF1 and NF-kB restore MHC class I-restricted tumor antigen processing and presentation to cytotoxic T cells in aggressive neuroblastoma. *PloS One* 2012;7:e46928.
- [38] Sun Z, Chen F, Meng F, et al. MHC class II restricted neoantigen: a promising target in tumor immunotherapy. *Cancer Lett* 2017;392: 17–25.
- [39] Kollmann D, Schweiger T, Schwarz S, et al. PD1-positive tumor-infiltrating lymphocytes are associated with poor clinical outcome after pulmonary metastasectomy for colorectal cancer. *Oncoimmunology* 2017;6:e1331194.
- [40] Xiao Y, Freeman GJ. The microsatellite instable subset of colorectal cancer is a particularly good candidate for checkpoint blockade immunotherapy. *Cancer Discov* 2015;5:16–8.
- [41] Andrews LP, Marciscano AE, Drake CG, et al. LAG3 (CD223) as a cancer immunotherapy target. *Immunol Rev* 2017;276:80–96.
- [42] Masugi Y, Nishihara R, Yang J, et al. Tumour CD274 (PD-L1) expression and T cells in colorectal cancer. *Gut* 2017;66:1463–73.
- [43] Garcia-Diaz A, Shin DS, Moreno BH, et al. Interferon receptor signaling pathways regulating PD-L1 and PD-L2 expression. *Cell Rep* 2017;19: 1189–201.
- [44] Shao L, Hou W, Scharping NE. IRF1 Inhibits Antitumor Immunity through the Upregulation of PD-L1 in the tumor cell. *Cancer Immunol Res* 2019;7:1258–66.