Translational Oncology 69 (2026) 102816

Contents lists available at ScienceDirect

Translational Oncology

journal homepage: www.elsevier.com/locate/tranon

Multi-omics interrogation identifies IRF4 as a causal transcription factor in monocytes that suppresses colorectal cancer progression by remodeling the immune microenvironment and inhibiting EMT

Huijuan Wen [a,b,1], Fazhan Li [a,b,1], Qingmin He [a,b], Ke Liu [a,b], Xi Zhang [a,b], Yuanchang Zhu [a,b], Tongtong Ge [a,b], Mengke Rong [a,b], Ruoyu Hu [a,b], Huayuan Xi [a,b], Pengyuan Zheng [a,b,*], Simeng Liu [a,b,*]

[a] Henan Key Laboratory for Helicobacter pylori and Digestive Tract Microecology, The Fifth Affiliated Hospital of Zhengzhou University; Institute of Rehabilitation Medicine, Henan Academy of Innovations in Medical Science; Tianjian Laboratory of Advanced Biomedical Sciences, Zhengzhou University, Zhengzhou, 450000, Henan, China

[b] Department of Gastroenterology, The Fifth Affiliated Hospital of Zhengzhou University, Zhengzhou, 453000, Henan, China

ARTICLE INFO

Keywords:
Colorectal cancer
IRF4
Mendelian randomization
Single-cell sequencing
Transcription factors
Tumor microenvironment

ABSTRACT

Background: Although surgery and chemotherapy remain standard treatments for colorectal cancer (CRC), immunotherapy offers a promising avenue for patients with advanced disease. However, the cellular and molecular mechanisms underlying CRC progression and immune evasion are not fully elucidated.

Methods: We integrated single-cell RNA sequencing (GSE161277), bulk transcriptomic data (GEO and TCGA), and Mendelian randomization (MR) analysis to identify key cell subpopulations and transcription factors (TFs) causally associated with CRC. Spatial transcriptomics, multiple immune deconvolution algorithms, and *in vitro* experiments (including lentiviral-mediated IRF4 modulation, flow cytometry, and ELISA) were performed to validate the expression, immune regulatory function, and prognostic significance of the candidate TF.

Results: Single-cell analysis identified monocytes as the most critical cell subpopulation in CRC pathogenesis. MR analysis revealed that IRF4, monocyte-associated TF, was causally linked to increased CRC risk. IRF4 expression was significantly downregulated in CRC tissues at both mRNA and protein levels, with elevated methylation and association with poor prognosis. Multi-omics immune profiling demonstrated that *IRF4* correlated with immune cell infiltration, cancer immunity cycle scores, and immunotherapy response. Spatially, *IRF4* was enriched in normal or low-malignant regions. Experimentally, *IRF4* overexpression in HCT116 CRC cell enhanced CXCL8 and CCL5 secretion, increased the number of $CD14^{+}IRF4^{+}$ monocytes and $CD3^{+}CD8^{+}IRF4^{+}$ T cells in co-culture systems, and suppressed epithelial-mesenchymal transition (EMT), whereas *IRF4* knockdown reversed these effects.

Conclusions: IRF4, monocyte-associated TF, is a causal factor in CRC development. Its downregulation correlates with adverse prognosis and impaired antitumor immunity. IRF4 modulates the tumor microenvironment by regulating chemokine secretion, immune cell recruitment, and EMT, positioning it as a potential therapeutic target for CRC immunotherapy.

Introduction

Colorectal cancer (CRC), one of the most prevalent cancer worldwide, represented 10% of the total global incidence of cancer and was the cause of 9.4% of all cancer deaths in 2020 [1]. Despite the increasing sophistication of CRC screening available, the insidious nature of CRC implied that a significant proportion of patients being diagnosed at a late stage where cancers are malignant and metastatic [2,3]. Unfortunately, the five-year survival rate for Stage III CRC is approximately 50% [4]. Although surgery and chemotherapy have been demonstrated to be efficacious in a considerable proportion of CRC, immunotherapy offers a promising avenue for advanced stage CRC patients [5].

* Corresponding authors at: Department of Gastroenterology, The Fifth Affiliated Hospital of Zhengzhou University, Zhengzhou, 453000, Henan, China.
E-mail addresses: pyzheng@zzu.edu.cn (P. Zheng), simengliu@zzu.edu.cn (S. Liu).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.tranon.2026.102816
Received 6 March 2026; Received in revised form 4 May 2026; Accepted 11 May 2026

Tumor Microenvironment (TME) refers to immediate surrounding area of a tumor, comprising not only the tumor cells, but also their surrounding supporting cells, extracellular matrix, blood vessels, immune cells, and a complex array of signaling molecules [6]. In addition to influencing tumor growth and metastasis, TME exerts an effect upon tumor response to therapy [7]. The advancement of single-cell sequencing has facilitated a more comprehensive classification of cell populations involved in the TME of CRC. CRC cells have been demonstrated to induce an adverse prognosis in CRC patients through their recruitment of tumor-associated neutrophils (TAN), which resulted in peripheral blood neutrophils becoming $AGR2^{+}$TAN secreting AGR2 [8]. Two discrete tumor-associated macrophages populations (TAMs) were discerned within the CRC, designated as $C1QC^{+}$ and $SPP1^{+}$TAMs. While only $SPP1^{+}$ TAMs demonstrated a marked preference for the tumor, which suggested their pivotal involvement in CRC tumorigenesis [9]. Consequently, a comprehensive understanding of the cellular and molecular mechanisms underlying the remodeling of the CRC tumor microenvironments is essential for identifying potential intervention targets to enhance the efficacy of immunotherapy.

Transcription factors (TFs) as master regulators with multifaceted role, regulated essential molecular and cellular processes, including the precise control of immunological functions, which determined the deregulation of TFs was of great significance in tumorigenesis and identification of the cancer hallmarks [10]. BHLHE40 was involved in the promotion of CRC tumorigenesis and metastasis through its role in the TGFB1-TGFBR1/2-BHLHE40 axis, which resulted in the transformation of malignant epithelial cells into $CXCL1^{+}$ Cancer-Associated Fibroblasts (CAFs) and $SFRP2^{+}$ CAFs via epithelial-mesenchymal transition (EMT) [11]. CNOT3 served as a prognostic factor in CRC induced malignant CRC behaviors through regulating the transcriptional programme that promoted self-renewal [12].

In this study, we integrated single-cell transcriptomics, Mendelian randomization, and comprehensive multi-omics analyses to identify monocytes as a key CRC-associated cell subpopulation and to establish IRF4 as a monocyte-associated TF with a causal role in CRC. We further characterized *IRF4* expression patterns, prognostic value, and immunomodulatory functions using spatial transcriptomics, immune deconvolution, and *in vitro* experiments. Our findings reveal that *IRF4* is downregulated in CRC, correlates with favorable prognosis, and modulates the TME by promoting chemokine secretion, enhancing immune cell recruitment, and suppressing EMT. These results position IRF4 as a potential therapeutic target for CRC immunotherapy.

Methods

Data sources

Single-cell RNA sequencing (scRNA-seq) data were obtained from the Gene Expression Omnibus (GEO) database (Accession: GSE161277), comprising three paired samples derived from colon cancer tissues and their corresponding adjacent normal tissues. To ensure data integrity, raw count matrices were re-processed through our standardized pipeline: low-quality cells were excluded based on predefined thresholds, including a gene count per cell between 100 and 7500, a unique molecular identifier (UMI) count range suitable for 10x Genomics data, and a mitochondrial gene percentage threshold of <10%.The data underwent normalization, and the top 2000 most variable genes were identified for principal component analysis (PCA). Cell clustering was executed using the functions "FindNeighbors" and "FindClusters," while visualization of results was achieved through uniform manifold approximation and projection (UMAP). The annotation of cell clusters was based on SingleR and established immune cell marker genes.

To validate and enhance the analysis of gene expression patterns, we curated a comprehensive bulk RNA sequencing cohort. This included datasets from the GEO repository (GSE18105, GSE21510, GSE39582, GSE41258, GSE71187, GSE77952, and GSE87211) and The Cancer Genome Atlas (TCGA) Colon Adenocarcinoma (COAD) project. The baseline characteristics of these cohorts, including tumor stage, histological subtype, and patient survival status, are summarized in Table S1. Raw data were subjected to inter-array quantile normalization, and stringent batch effect correction was performed using the ComBat-seq algorithm (from the 'sva' package in R) to harmonize the heterogeneity arising from different sequencing platforms (e.g., Affymetrix vs. Illumina) before downstream differential expression analysis.

Identification of key cell subpopulations and genes

To identify cell subpopulations critical for CRC development, generalized linear models (glm) and random forest (RF) models were constructed with cell subpopulations as independent variables and CRC status as the dependent variable. The dalex package was used to interpret model outputs and assess the importance of each cell subpopulation.

To investigate key TF regulatory networks modulating the function of different cell subpopulations, Single-Cell reEgulatory Network Inference and Clustering in Python (pySCENIC) was applied. Co-expression networks were computed using GRNboost, and regulons were identified via RcisTarget. The regulon activity of each cell was quantified using AUCell scores.

Mendelian randomization (MR) analysis

Mendelian randomization was employed to explore genes with causal relationships to CRC. Expression quantitative trait loci (eQTLs) and protein quantitative trait loci (pQTLs) for monocyte TFs were used as exposures. eQTL data were obtained from the eQTLGen consortium (https://eqtlgen.org/), including 16,989 genes from 31,684 blood samples of healthy European ancestry. Only cis-eQTLs, Single Nucleotide Polymorphisms (SNPs) within 1 Mb of the gene encoding the protein, were retained. pQTL data originated from 4907 plasma proteins in an Icelandic cohort. Instrumental variables (IVs) with F-statistics ≥ 10 were selected, and linkage disequilibrium (LD) among SNPs was clumped at $r^2 < 0.01$ and a 10,000 kb distance. Steiger filtering was applied to remove IVs explaining more variance in CRC than in the corresponding gene.

Outcome data for CRC were sourced from the IEU OpenGWAS project (Dataset: ukb-b-20145), comprising 1494 CRC cases and 461,439 controls. SNPs matching the IVs were extracted from the outcome data.

For genes with a single eQTL, the Wald ratio method was used for MR analysis. For those with multiple genetic instruments, inverse variance weighted (MR-IVW) meta-analysis was applied. MR analyses were performed using the "TwoSampleMR" package (version 0.5.10). Horizontal pleiotropy was assessed using MR-Egger intercept. Heterogeneity was evaluated using Cochran's Q statistic, with $P > 0.05$ indicating no significant heterogeneity. Leave-one-SNP-out sensitivity analysis was conducted to assess the stability of the MR effect.

Differential expression and immune landscape of key genes

The mRNA expression of IRF4 in CRC patients and healthy controls was extracted from GEO datasets (GSE18105, GSE21510, GSE39582, GSE41258, GSE71187, GSE77952, and GSE87211) and TCGA. The protein expression of IRF4 was obtained from The Cancer Imaging Archive, and its methylation levels were extracted from TCGA. Differential expression between groups was statistically compared.

CIBERSORTx was used to quantify cell-type abundances in bulk RNA-seq data. Immune infiltration data for TCGA samples were obtained from TIMER2.0. Multiple algorithms (XCELL, QUANTISEQ, MCPCOUNTER, CIBERSORT-ABS, TIMER, and EPIC) were employed to comprehensively assess cell-type abundances and their correlations with IRF4 expression. The Tracking Tumor Immunophenotype (TIP) platform

was used to visualize and analyze anticancer immune status and the proportion of tumor-infiltrating immune cells across the seven-step cancer immunity cycle: cancer cell antigen release (step 1), cancer antigen presentation (step 2), initiation and activation (step 3), immune cell trafficking to tumors (step 4), immune cell infiltration into tumors (step 5), T cell recognition of cancer cells (step 6), and killing of cancer cells (step 7). TIP scores were quantified for each step.

Tumor Immune Single-cell Hub (TISCH2), https://tisch.comp-genomics.org/, was used to explore gene expression and distribution across various cell types in the tumor microenvironment and to analyze cell-cell interactions.

Antitumor immune responses were predicted from RNA-seq data using easier, a tool that predicts immunotherapy responses based on cancer-specific immune response models. Five scores were assessed: CYT (cytolytic activity), TLS (tertiary lymphoid structure signature), IFN-γ (interferon-γ signature), T cell_inflamed (T cell-inflamed microenvironment signature), and Chemokines (chemokine signature).

Spatial transcriptomics

Spatial transcriptomics (ST) data were obtained from The National Omics Data Encyclopedia (OEP001756). To accurately estimate cell composition on 10x Visium slides, deconvolution analysis was performed using ST and scRNA-seq data. scRNA-seq data from multiple samples of the same cancer type were collected to construct a comprehensive scRNA reference library. The quality control of scRNA-seq data was performed based on the number of expressed genes, UMI counts, and mitochondrial RNA percentage, following criteria from the source studies. A signature score matrix was constructed by calculating the average expression of the top 25 specific genes for each cell type in the scRNA-seq reference. Enrichment score matrices were generated using the get_enrichment_matrix and enrichment_analysis functions in the Cottrazm package. Enrichment scores for each cell type were visualized using the SpatialFeaturePlot function in Seurat. Spots were classified as Malignant (enrichment score for malignant cells = 1), Normal (score = 0), or Mixed (otherwise). Wilcoxon Rank Sum Tests were used to assess differences in gene expression among the three groups.

Cell culture, stable cell line construction, and western blot

The human normal colon epithelial cell line NCM460, CRC cell lines HCT116 and SW480, and the monocytic cell line THP-1 were obtained from the Cell Bank of the Chinese Academy of Sciences. Human peripheral blood $CD3^+$ T cells were isolated from healthy volunteers using magnetic bead sorting (Miltenyi Biotec), which abtained the approval of the fifth affiliated hospital of zhengzhou university (No. KY2024033). NCM460 cells were cultured in DMEM/F12 medium, while HCT116, SW480, and THP-1 cells were maintained in RPMI-1640 medium, both supplemented with 10% FBS and 1% penicillin streptomycin. To establish IRF4 stable overexpression (OE) and knockdown (KD) cell lines, HCT116 and SW480 cells were transduced with lentiviral vectors carrying IRF4 or shRNA constructs, respectively, and selected with puromycin. IRF4 protein expression was validated by western blot using anti-IRF4 (11247-2-AP, Proteintech, China) and anti-GAPDH (10494-1-AP, Proteintech, China) antibodies, followed by horseradish peroxidase (HRP) conjugated secondary antibodies and ECL detection. Band intensities were quantified using ImageJ.

Co-culture systems and flow cytometry

For monocyte-CRC interaction studies, THP-1 cells were seeded in the upper chamber of 0.4 μm transwell inserts and co-cultured with HCT116 or SW480 NC/OE/KD cells in the lower chamber for 48 h. For T cell-CRC interaction, $CD3^+$ T cells were directly co-cultured with CRC cells at an effector-to-target ratio of 2:1 for 48 h. After co-culture, cells were harvested and stained with fluorochrome-conjugated antibodies for flow cytometric analysis. THP-1 cells were stained with CD14-FITC and IRF4-PE to assess $CD14^+IRF4^+$ monocytes. T cells were stained with CD3-APC, CD8-FITC, and IRF4-PE to quantify $CD3^+CD8^+IRF4^+$ cells. CRC cells were stained with E-cadherin-PE and Vimentin-APC to evaluate epithelial mesenchymal transition (EMT) phenotypes. Additionally, baseline IRF4 expression in NCM460, HCT116, and SW480 cells was determined by intracellular staining with IRF4-PE. All samples were acquired on a BD FACSCanto II and analyzed with FlowJo software.

ELISA Assay

Culture supernatants from CRC NC/OE/KD cells were collected after 24 h of serum-free culture, and concentrations of CXCL8 (KE00453, Proteintech, China) and CCL5 (KE00093, Proteintech, China) were measured according to the manufacturer's instructions. Briefly, standards and samples were added to the wells, followed by HRP conjugated detection antibody. After incubation and washing, chromogenic substrates were added, and the reaction was terminated with stop solution. Optical density was measured at 450 nm, and cytokine concentrations were calculated from standard curves.

Statistical analysis

Statistical analyses were performed using R and Python. Student's *t*-test was used for continuous normally distributed variables between two groups, and Mann-Whitney U test for non-normally distributed variables. Survival differences were visualized using Kaplan-Meier curves and assessed by Log-rank test. Correlations between variables were analyzed using Spearman's correlation. All reported P values were two-tailed, and $P < 0.05$ was considered statistically significant. For cell experiments, data from three biological and three technical replicates were analyzed using SPSS 26.0 and presented as mean ± standard deviation. Comparisons between two groups were performed using independent samples *t*-test, and among multiple groups using one-way ANOVA with LSD post-hoc test. Graphs were generated using GraphPad Prism 9.0.

Results

Single-cell clustering and identification of cell subtypes

After quality control, normalization, and dimensionality reduction, clustering analysis of six samples identified eleven distinct cell subtypes: $CD8^+$ T cells, Naïve B cells, Monocytes, Epithelial cells, NK cells, Plasma cells, Smooth muscle cells, Germinal center cells, Endothelial cells, Dendritic cells (DC), and Common myeloid progenitor cells (CMP) (Fig. 1A-B). The analysis of cell type proportions between CRC and normal samples revealed that $CD8^+$ T cells were more abundant in CRC samples, while Epithelial cells constituted a larger proportion in normal samples (Fig. 1C-D). Marker gene expression patterns across different cell clusters showed significant heterogeneity (Fig. 2).

Identification of monocyte subpopulation and associated transcription factors

Using the DALEX package, cell subpopulation importance in predicting CRC risk was assessed. In the glm model, epithelial cells ranked as the most important feature, followed by monocytes (Fig. 3A). In the RF model, monocytes were identified as the most critical variable (Fig. 3B). The immune infiltration analysis of CRC transcriptomic data from TCGA revealed monocyte infiltration scores were increased in CRC tissues compared to normal tissues (Fig. 4A). Higher monocyte enrichment was associated with poorer patient survival (Fig. 4B). Consequently, monocytes were selected as the key cell subpopulation for further investigation.

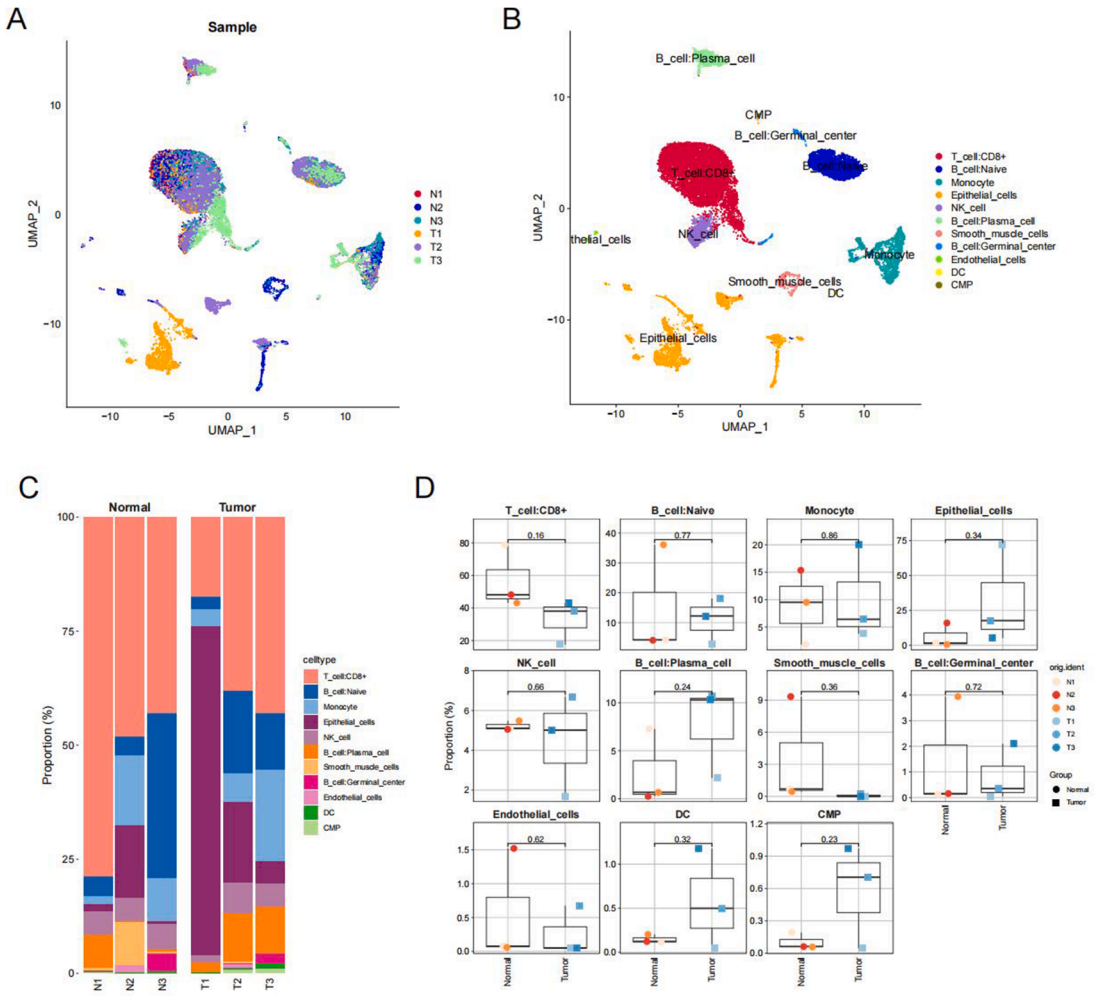

Fig. 1. Single-cell clustering and the identification of cell subtypes in CRC.
(A) The UMAP plot of single-cell transcriptomes from 6 CRC and paired normal samples, with cells colored by cluster identity. (B) The annotation of cell clusters based on SingleR and classical marker genes. 11 distinct cell subtypes were identified: $CD8^+$ T cells, Naïve B cells, Monocytes, Epithelial cells, NK cells, Plasma cells, Smooth muscle cells, Germinal center cells, Endothelial cells, Dendritic cells (DCs), and Common myeloid progenitor cells (CMP). (C) The proportion of each cell subtype in CRC and normal samples. (D) The quantitative comparison of cell subtype abundances between CRC and normal tissues.

Differential gene expression analysis revealed 759 differentially expressed genes in monocytes between normal and CRC samples. *MT1X, RNASE1, LYVE1, MT2A*, and *VSIG4* were significantly upregulated in normal tissues, while *IGHA1, IGKC, IGHA2, IGLC2*, and *C15orf48* were significantly upregulated in CRC samples (Fig. 5A,Table S2).

To explore the regulatory mechanisms underlying these expression differences, pySCENIC analysis was performed to identify TF regulatory networks within monocytes (Fig. 5D). 153 potential TFs were identified in the monocyte subpopulation (Table S3). Regulons controlled by OLIG1, ZNF467, and NFE2 exhibited high regulon activity scores (RAS) and strong regulon specificity scores (RSS) (Fig. 5B). The IRF4-controlled regulon showed high RAS in smooth muscle cells and the highest RSS in plasma B cells (Fig. 5C).

IRF4 Was identified as a key transcription factor in monocytes for CRC

Using monocyte TF eQTLs and pQTLs as instrumental variables, causal relationships with CRC risk were investigated. *ZNF407, ZBTB40, SP3, RFX2*, and *IRF4* were found to be causally associated with CRC. *IRF4* showed the most pronounced expression change, which was significantly downregulated, in CRC compared to the other four TFs, indicating *IRF4* might be a critical protective regulator within monocytes (Fig. 6A,B). Scatter plots illustrated SNP effect sizes for *IRF4* and CRC (Fig. 6D). Leave-one-out analysis indicated no single SNP disproportionately influencing the overall effect (Fig. 6C), and no horizontal pleiotropy was detected (Fig. 6E).

IRF4 Expression and prognosis in CRC

IRF4 mRNA expression was significantly lower in CRC tissues compared to healthy controls across multiple GEO datasets and TCGA (Fig. 7A). Consistently, *IRF4* protein expression was also reduced in CRC, while methylation levels were increased (Figs. 7B-C). Survival analysis based on TCGA-COAD data demonstrated that high *IRF4* expression was associated with longer overall survival (OS), progression-free interval (PFI), and disease-free interval (DFI) (Fig. 7D). These findings were validated in additional independent datasets (Fig. 7E). Restricted cubic spline analysis revealed a nonlinear relationship between *IRF4* expression and survival outcomes, OS, Disease-Specific Survival (DSS), PFI, and DFI (Fig. 7F).

Impact of IRF4 on the tumor microenvironment in CRC

Multi-algorithm immune infiltration analysis showed correlations between *IRF4* expression and the abundances of B cells, DCs, T cells, and monocytes (Figs. 8A, 8C). *IRF4* expression was positively correlated with

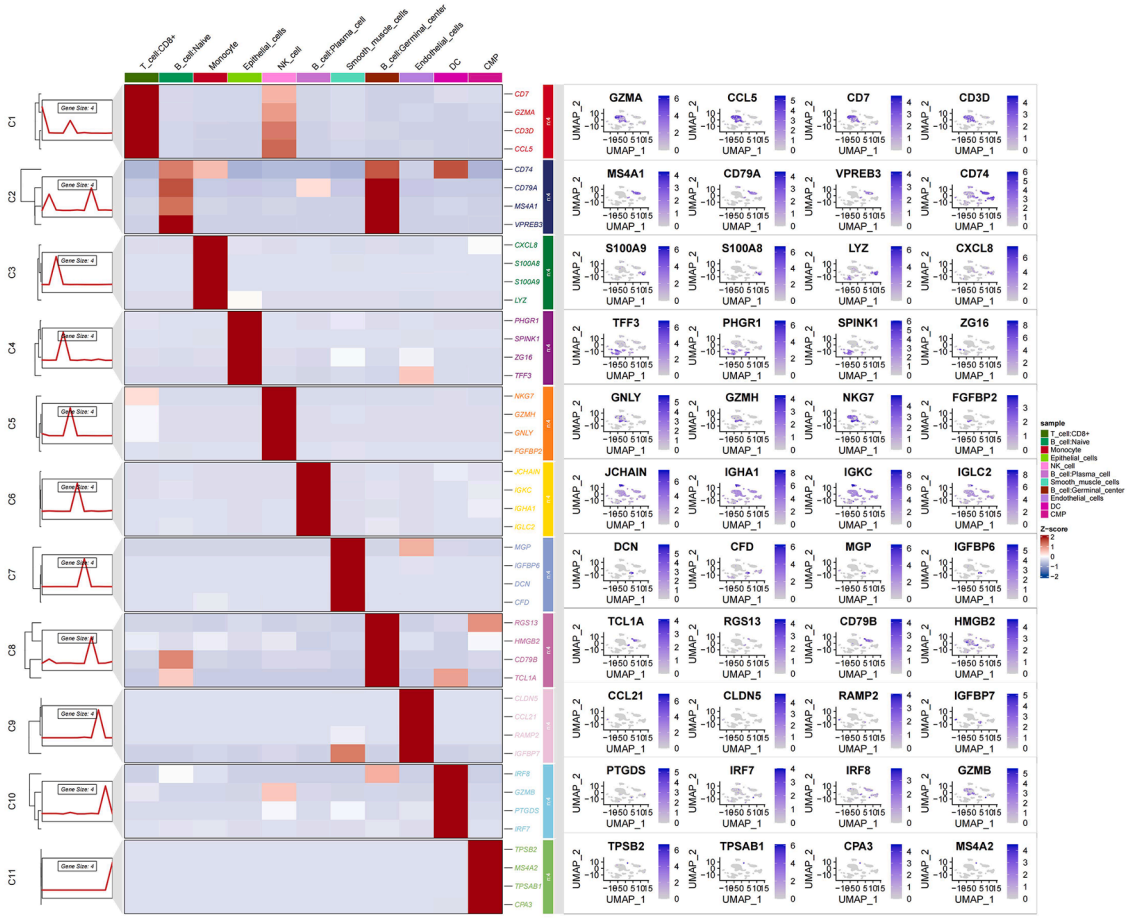

Fig. 2. The marker gene expression patterns across cell clusters.
The expression patterns of canonical marker genes across the 11 identified cell clusters were displayed through heatmap. Each row represented a marker gene, and each column represented a single cell. Color intensity indicated relative expression levels of genes, demonstrating cluster-specific gene signatures.

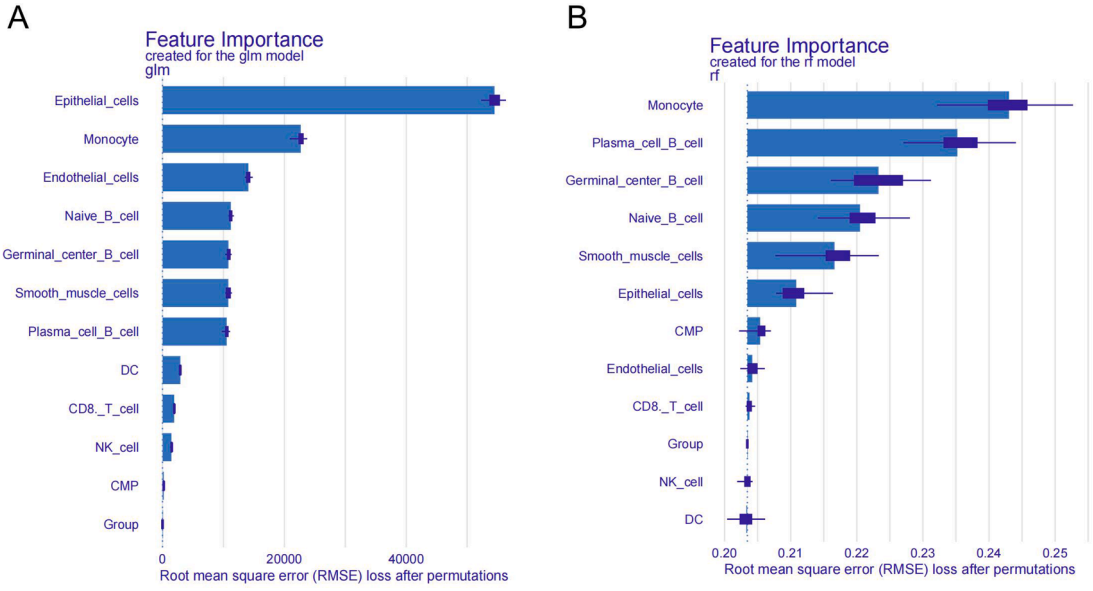

Fig. 3. The identification of key cell subpopulations in CRC pathogenesis.
(A) The contribution of different cell subpopulation to CRC risk prediction was identified through generalized linear model (glm). (B) The contribution of various cell subpopulation to CRC risk prediction was represented by random forest (RF) model.

cancer immunity cycle scores (Fig. 8B). At the single-cell level, *IRF4* expression varied across T cell subtypes (Figs. 8D-F) and was enriched in

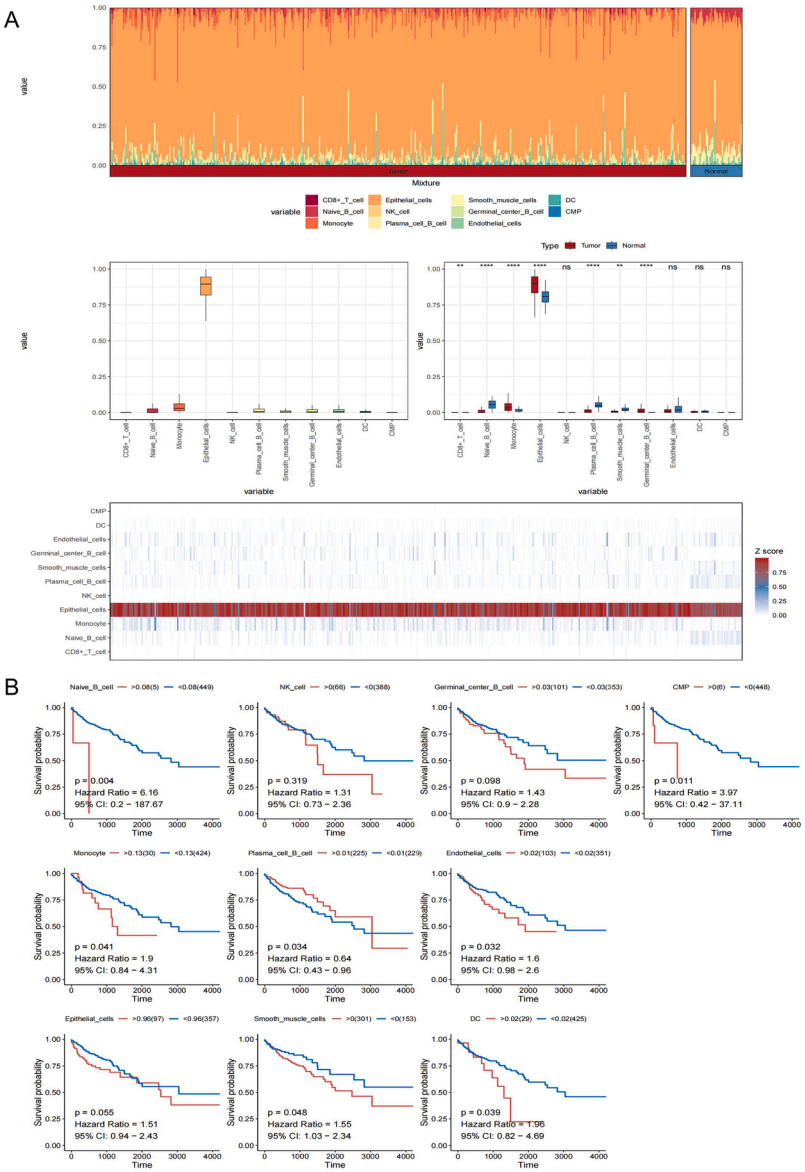

Fig. 4. The immune landscape of CRC and prognostic impact of immune cell subsets.
(A) The relative abundance of immune cell subsets in CRC and normal tissues was quantified by CIBERSORTx. (B) The association between each immune cell subset abundance and patient prognosis was shown by survival analysis.

low-malignant or normal cells within CRC tissues (Figs. 8G-I). Previous literature shows that IRF4 is a key driver of PDCD1 expression [13], we further evaluated the impact of the immune checkpoint molecule PDCD1 on the CRC immune environment. In CRC, the *PDCD1* high-expression group exhibited higher scores of CYT, TLS, IFN-γ, and T cell_inflamed, but lower chemokines produced from T cells scores (Figure S1A-E).

Experimental validation of IRF4 regulatory function in CRC

To validate the bioinformatics findings, *in vitro* experiments were conducted using normal colon epithelial cells (NCM460) and CRC cell lines (HCT116, SW480). Western Blot analysis confirmed significantly lower IRF4 protein expression in CRC cell lines compared to NCM460, with more aggressive HCT116 cell showing lower expression than SW480 cell (Fig. 9A). Single-color flow cytometry corroborated these findings, showing reduced IRF4 MFI in CRC cells (Fig. 9B). Thus, HCT116 cell was furtherly uesd for follwing verification.

In transwell co-culture systems, *IRF4* overexpression in HCT116 cell significantly increased the proportion of $CD14^{+}IRF4^{+}$ monocytes, while *IRF4* knockdown reduced this population (Fig. 9C). Similarly, in direct co-culture with T cells, *IRF4* overexpression increased the percentage of $CD3^{+}CD8^{+}IRF4^{+}$ T cells, whereas that of knockdown decreased it (Fig. 9D).

ELISA assays demonstrated that *IRF4* overexpression upregulated CXCL8 and CCL5 secretion in HCT116 cell, while knocking down *IRF4* suppressed their release (Fig. 9E). Furthermore, flow cytometry analysis of EMT markers showed that *IRF4* overexpression increased the proportion of $E\text{-cadherin}^{+}$ epithelial phenotype cells and decreased $Vimentin^{+}$ mesenchymal phenotype cells, with opposite effects were observed after knocking down *IRF4* in HCT116 cell (Fig. 9F).

Collectively, these experimental results validate that IRF4 is significantly downregulated in CRC and plays a crucial role in modulating monocyte and $CD8^{+}$ T cell phenotypes, promoting chemokine secretion, and inhibiting EMT, which consistent with the bioinformatics analyses.

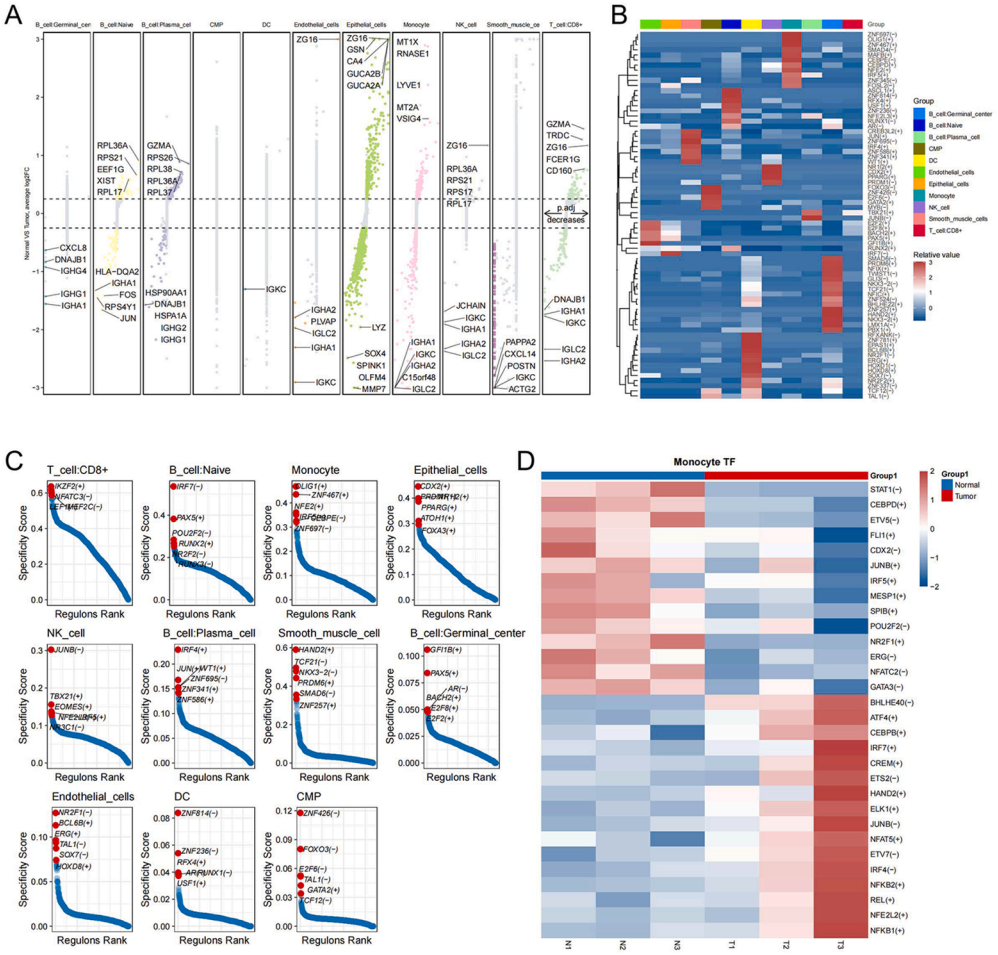

Fig. 5. Transcriptional regulatory networks in monocyte subpopulation.
(A) The differentially expressed genes in above 11 differernt immune cells between normal and CRC tissues were presented through heatmap. (B) The regulon activity scores (RAS) of TF regulatory networks across cell subpopulations, with colors representing RAS levels. (C) The top 6 TFs with highest specificity in each cell subpopulation were highlighted by regulon specificity scores (RSS). (D) The heatmap of the top 30 TFs in monocytes based on regulon activity.

Discussion

In this study, we employed an integrated approach combining single-cell RNA sequencing, Mendelian randomization, multi-omics immune profiling, and *in vitro* experiments to systematically investigate the role of monocyte-associated TFs in CRC. Our findings identify monocytes as the most critical cell subpopulation in CRC pathogenesis and establish IRF4 as a monocyte-associated TF with a causal relationship to CRC risk. We demonstrate that IRF4 is significantly downregulated in CRC tissues, correlates with favorable prognosis, and exerts pleiotropic effects on the tumor microenvironment by modulating chemokine secretion, immune cell recruitment, and EMT.

In line with the findings of ours, Wang et al. found monocytes and macrophages exerted a predominant influence on the TME in CRC [14]. Circulating monocytes exhibited considerable heterogeneity, monocytes can be classified into three major subsets based on their phenotypic and functional characteristics: classical monocytes ($CD14^{++}$ $CD16^{-}$), intermediate monocytes ($CD14^{+}CD16^{+}$) and non-classical monocytes ($CD14^{+}CD16^{++}$) [15]. Distinct monocyte subsets engaged in a range of activities that could either promote or inhibit tumor growth, including, secretion of tumor-inhibitory substances, promotion of new blood vessel growth, remodelling of the extracellular matrix, recruitment of lymphocytes and differentiation into tumor-associated macrophages and DCs [16]. The circulation of monocytes represented a crucial source of tumor-associated macrophages (TAMs). Wang et al. suggested the expression of CXCR2 on peripheral monocytes has been observed to influence the composition of TAMs within the TME of CRC [17]. The presence of high $CD14^{+}$ monocytic cell densities in the CRC stroma was found to be associated with a reduced risk of cancer-specific mortality, even when potential confounding factors were taken into account, including disease stage, mismatch repair status, and the presence of specific mutations [18]. Besides, the distribution of $CCR2^{+}$ and $CD163^{+}$ monocytes in colorectal cancer patients was subject to modulation by neoadjuvant chemotherapy and surgical intervention, which indicated subsets of monocytes were indicative of the effects of therapy in CRC [19].

IRF4 has been considered as a remarkable regulators of monocyte differentiation into monocyte-derived DCs (mo-DCs) [20,21].In addition, IRF4, as a TF, played a pivotal role in regulating normal T and B lymphocyte growth and differentiation, which determined it has been proposed as a prospective therapeutic target for hematological diseases, such as peripheral T-cell lymphoma and T-cell leukemia/lymphoma [22–25]. IRF4, in conjunction with BAFT, exerted control over a molecular program that governed effector Treg differentiation and suppression in tumor [26] Previous studies have demonstrated IRF 4 had a crucial function in the formation of various immunosuppressive cells within the TME [21,27]. In CRC, IRF4, recognized as a TME microenvironment-related genes, displayed a low level of expression and positively correlated with prognosis [28,29], which was consistent with our findings. And Wang, J, et al. suggested upregulated IRF4 could inhibit the growth of CRC cells by promoting the differentiation of Tregs into macrophages [30]. Moreover, our research illustrated IRF4 was

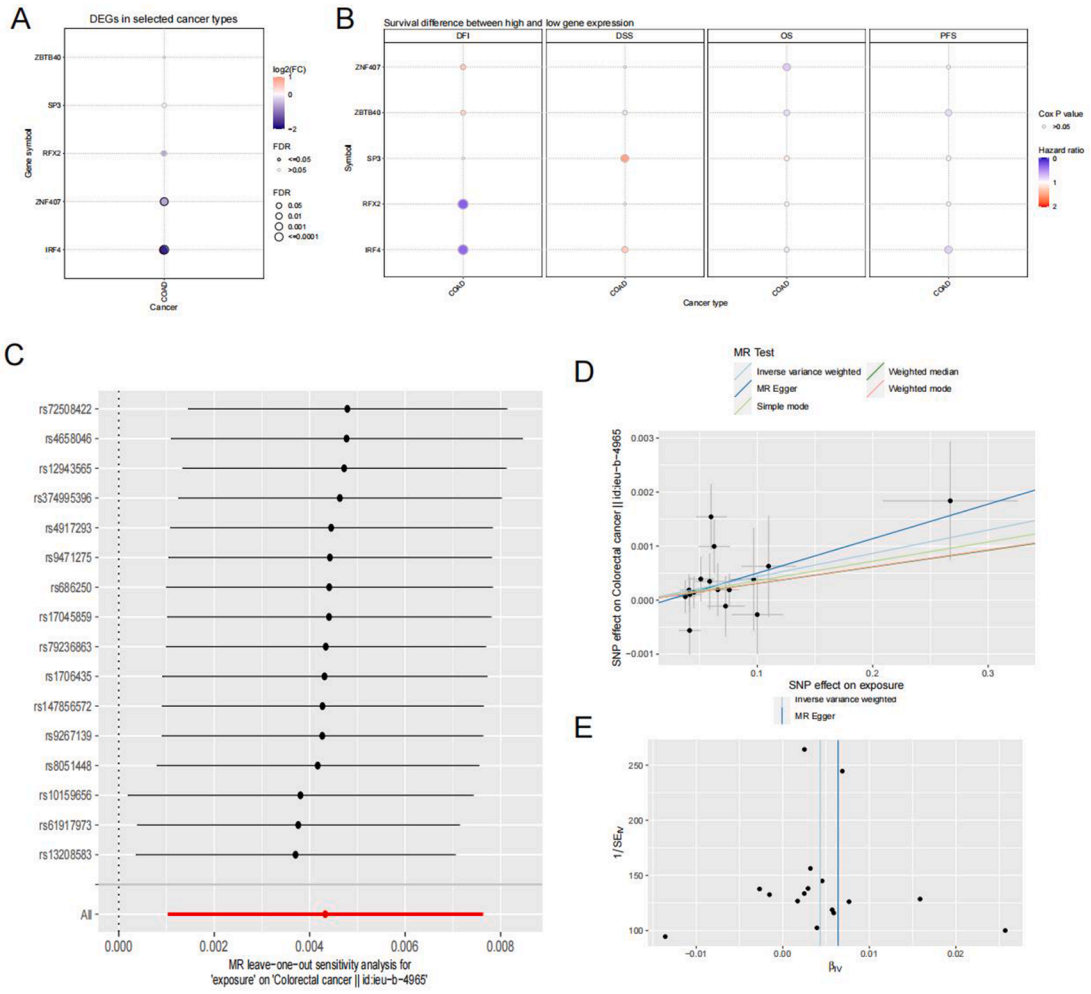

Fig. 6. Mendelian randomization (MR) identified IRF4 as a causal TF for CRC.
(A) The causal effects of monocyte-associated TFs on CRC risk were shown through forest plot. Five mRNAs, ZNF407, ZBTB40, SP3, RFX2, and IRF4, were causally associated with CRC. (B) The relationship of the expression of the above 5 differentially expressed genes in CRC and 4 survival differences, DFI, DSS, OS, and PFS. (C) Leave-one-out sensitivity analysis was performed to analyze whether *IRF4* exerted a single SNP that disproportionately influenced the overall causal estimate. (D) The scatter plot of SNP effects on *IRF4* expression versus CRC risk, with slopes representing causal estimates. (E) The funnel plot was used to assess horizontal pleiotropy for *IRF4*.

linked to the scoring of the cancer immune cycle. Cancer immunotherapies, encompassing checkpoint inhibitors (PDCD1 and CTLA-4) and adoptive cell therapy, utilize the immune system to identify and target cancerous cells [31]. IRF4 has been implicated in numerous facets of T cell exhaustion, including the upregulation of inhibitory receptors (including PDCD1), the reduction of cytokine secretion, and the attenuation of anabolic metabolism [32]. Similarly, Hirsch, T, et al. concluded IRF4 enhanced the expression of PDCD1 and promoted T cell proliferation, while impaired cytokine production of T cells [13]. The *in vitro* experiments provided functional validation of IRF4′s immunomodulatory role. *IRF4* overexpression in CRC cells enhanced the secretion of CXCL8 and CCL5, two chemokines critical for immune cell recruitment. CXCL8 (IL-8) is a chemokine for neutrophils and monocytes [33], while CCL5 (RANTES) recruits T cells, monocytes, and DCs [34]. The increased secretion of these chemokines likely underlies the observed increase in $CD14^{+}IRF4^{+}$monocytes and $CD3^{+A}CD8^{+}IRF4^{+}$ T cells in co-culture systems. These findings demonstrate that *IRF4*-expressing CRC cells could actively remodel the TME by promoting the recruitment and differentiation of immune cells, thereby fostering an antitumor immune microenvironment. This may constitute a potential mechanism through which IRF4 could play a role in the immunotherapeutic treatment of CRC. The immunotherapy response prediction is based on in silico algorithms (TIP and easier) and requires prospective validation in future immunotherapy trials. Although IRF4 overexpression in CRC cells enhanced CXCL8 and CCL5 secretion and concomitantly increased $CD14^{+}IRF4^{+}$ monocytes in co-culture, we did not directly establish a causal link. It remains possible that other IRF4-regulated factors contribute to monocyte recruitment or differentiation. Future studies using chemokine-neutralizing antibodies, recombinant proteins, or chemokine-specific knockout models are required to determine whether CXCL8 and/or CCL5 are necessary and sufficient for the observed monocyte effects.

The reciprocal regulation between IRF4 expression in CRC cells and immune cell phenotypes suggests a complex crosstalk within the TME. CRC cells with high IRF4 expression may create a feed-forward loop that promotes antitumor immunity, while the loss of IRF4 expression may contribute to immune evasion. This concept is supported by our observation that *IRF4* knockdown in HCT116 CRC cell reduced chemokine secretion and decreased the proportion of IRF4-expressing immune cells in co-culture. These findings highlight the importance of IRF4 as a node of communication between tumor cells and the immune microenvironment.

Several limitations of this study should be acknowledged. First, while our MR analysis established a causal relationship between IRF4 and CRC, the specific downstream target genes through which IRF4 exerts its effects in monocytes and other immune cells remain to be identified.

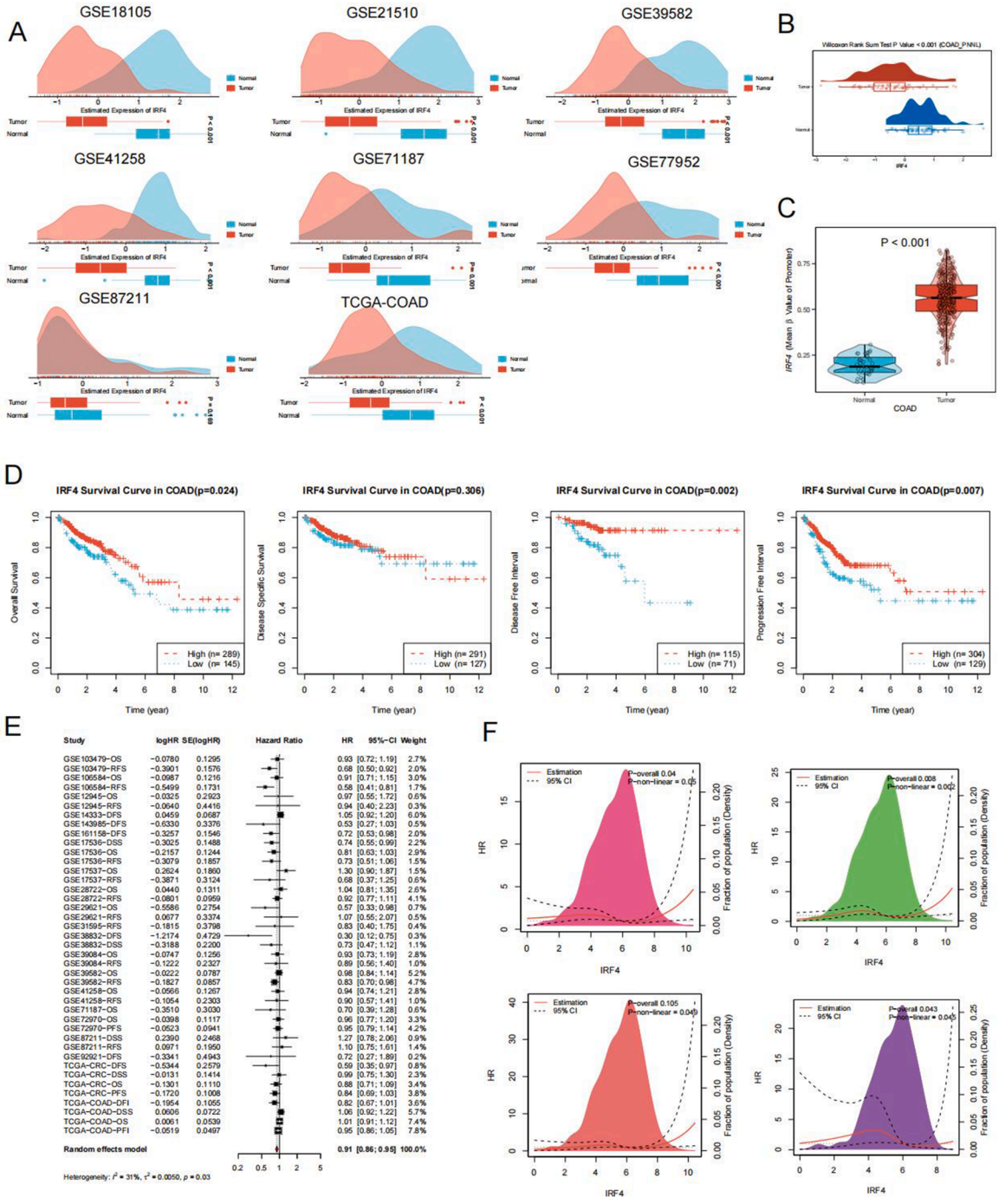

Fig. 7. IRF4 Expression and prognostic value in CRC.
(A) *IRF4* mRNA expression in CRC tissues compared to normal controls across multiple GEO datasets and TCGA. (B) IRF4 protein expression in CRC versus normal tissues from The Cancer Imaging Archive. (C) The methylation levels of *IRF4* in CRC and normal tissues from TCGA. (D) The correlations of *IRF4* expression and overall survival (OS), disease-specific survival (DSS), progression-free interval (PFI), and disease-free interval (DFI) in COAD patients were displayed through Kaplan-Meier survival curves from the TCGA-COAD cohort. (E) The validation of *IRF4* prognostic value in independent GEO datasets. (F) The relationships between *IRF4* expression and survival outcomes (OS, DSS, PFI, and DFI) were depicted by restricted cubic spline analysis.

Second, although we utilized multiple public datasets to validate our findings, further validation in independent clinical cohorts and *in vivo* models is warranted. Third, the precise molecular mechanisms by which IRF4 regulates chemokine expression and EMT in CRC cells require deeper investigation, including chromatin immunoprecipitation sequencing (ChIP-seq) and transcriptomic profiling of IRF4-manipulated cells. In addition, our observation that higher bulk monocyte infiltration correlates with worse prognosis, whereas higher IRF4 expression (detected in monocytes by scRNA-seq) correlates with better prognosis, represents an apparent paradox. This is likely explained by the functional heterogeneity within monocytes. IRF4 promotes monocyte differentiation into immunostimulatory monocyte-derived dendritic cells [20], whereas total monocyte abundance may be dominated by $CD16^{+}$ subsets associated with immunosuppressive TAM polarization. Future studies using single-cell survival analysis or multiplex imaging to link specific monocyte subsets ($IRF4^{+}$ vs. $IRF4^{-}$) to patient outcomes are

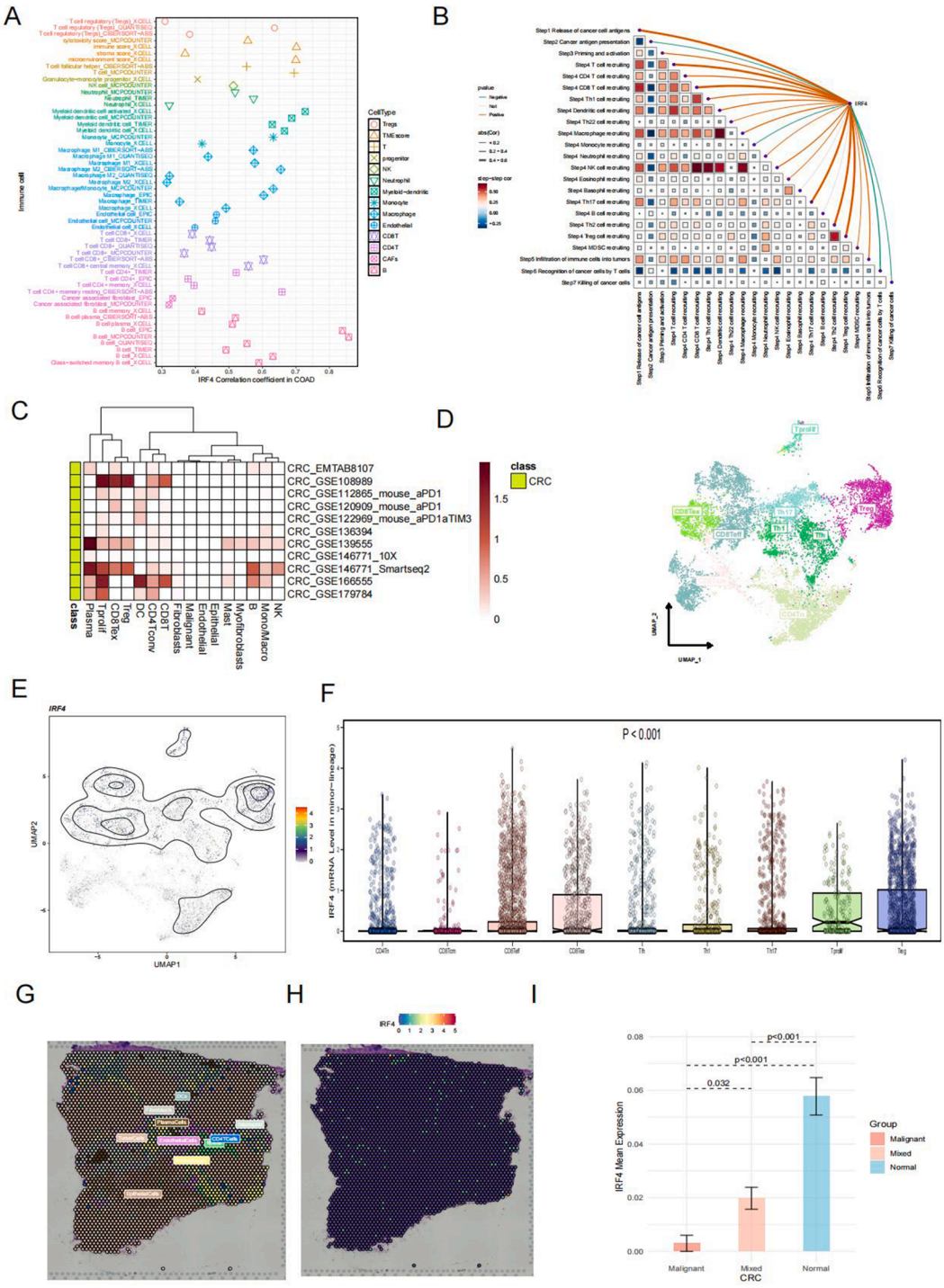

Fig. 8. *IRF4* Correlated with immune infiltration and its spatial distribution in CRC.
(A) The correlations between *IRF4* expression and immune cell abundances assessed by multiple algorithms (XCELL, QUANTISEQ, MCPCOUNTER, CIBERSORT-ABS, TIMER, and EPIC) were shown by heatmap. (B) The correlation network showed the relationships between *IRF4* expression and cancer immunity cycle scores from the TIP database. Red lines indicated positive correlations; green lines indicated negative correlations; line thickness represented correlation strength. (C) *IRF4* expression associations with various cell types in the TME were shown through correlation matrix from TISCH database. (D-F) The single-cell analysis of *IRF4* expression across T cell subtypes from TISCH database. (G-I) *IRF4* mean expression in malignant, mixed and normal regions within CRC tissues was presented through spatial transcriptomics analysis. Spots were classified as Malignant (enrichment score = 1), Normal (score = 0), or Mixed.

needed to resolve this paradox. We caution that bulk immune deconvolution does not capture subset-specific functions."

In conclusion, this study identifies monocytes as a key cellular component in CRC and establishes IRF4 as a monocyte-associated TF with causal relevance to CRC development. The downregulation of IRF4 in CRC, its association with favorable prognosis, and its multifaceted role in modulating the TME through chemokine secretion, immune cell recruitment, and EMT suppression position IRF4 as a promising therapeutic target for CRC immunotherapy [35]. Future studies should focus on elucidating the molecular networks downstream of IRF4 and exploring the potential of IRF4-based therapeutic strategies in preclinical models.

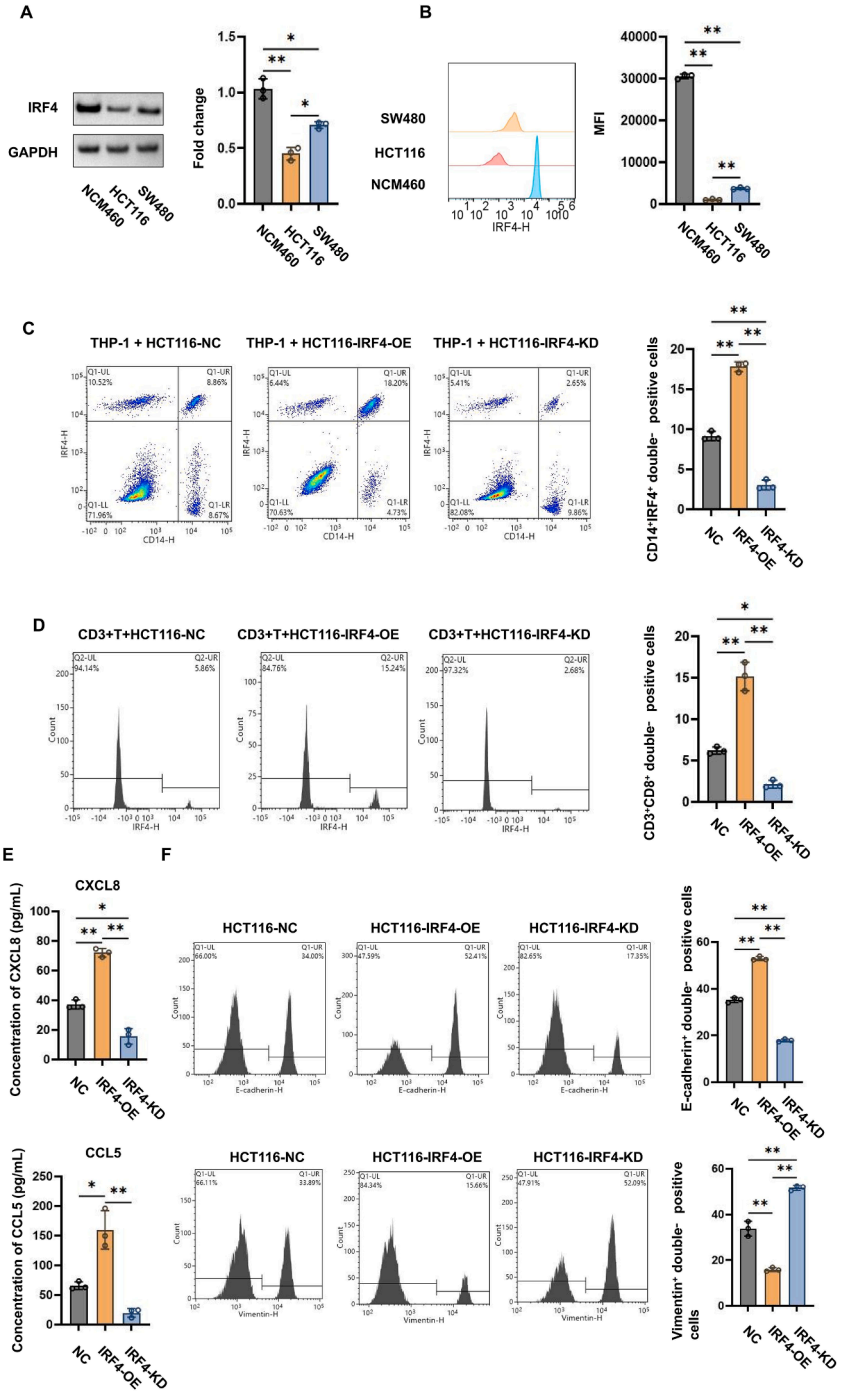

Fig. 9. Experimental validation of IRF4 regulatory function in CRC.
(A) The western blot analysis of IRF4 protein expression in normal colon epithelial cells (NCM460) and CRC cell lines (HCT116, SW480). (B) The single-color flow cytometry analysis of intracellular IRF4 expression (mean fluorescence intensity, MFI) in NCM460, HCT116, and SW480 cells. (C) The two-color flow cytometry analysis of $CD14^{+}IRF4^{+}$ monocytes in transwell co-culture systems with HCT116 CRC cells, empty vector group cells(NC), *IRF4* overexpression group cells (OE), and *IRF4* knockdown group cells (KD). (D) The three-color flow cytometry analysis of $CD3^{+}CD8^{+}IRF4^{+}$ T cells in direct co-culture systems with the above 3 groups cells. (E) The ELISA quantification analysis of CXCL8 and CCL5 secretion from the above 3 cells. (F) The two-color flow cytometry analysis of EMT markers, E-cadherin and Vimentin, from the above 3 groups cells, Data were presented as mean ± SD from three independent experiments. $^{*}P < 0.05$, $^{**}P < 0.01$.

Ethics approval and consent to participate

Our study has been approved by the Fifth Affiliated Hospital of Zhengzhou University, with the ethics number KY2024033.

Consent for publication

All authors agreed to publish.

Data availability

Data used in this study could be obtained upon reasonable request to

the corresponding author.

Funding

This material is based on work supported by Tianjian advanced biomedical laboratory key research and development project, Henan Provincial Medical Science and Technology Research Project (No. SBGJ202503041), Henan Provincial International Science and Technology Cooperation Program (No. 252102520089), Henan Provincial Overseas Training Program for Medical Science and Technology Talents (No. HNMOT2024036) and Henan Provincial Science and Technology Research Project (No. 262102310393).

CRediT authorship contribution statement

Huijuan Wen: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Fazhan Li:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Qingmin He:** Writing – review & editing, Writing – original draft, Formal analysis, Data curation, Conceptualization. **Ke Liu:** Writing – review & editing, Writing – original draft, Data curation, Conceptualization. **Xi Zhang:** Writing – review & editing, Writing – original draft, Data curation, Conceptualization. **Yuanchang Zhu:** Writing – review & editing, Writing – original draft, Conceptualization. **Tongtong Ge:** Writing – review & editing, Writing – original draft, Conceptualization. **Mengke Rong:** Writing – review & editing, Writing – original draft, Conceptualization. **Ruoyu Hu:** Writing – review & editing, Writing – original draft, Conceptualization. **Huayuan Xi:** Writing – review & editing, Writing – original draft, Conceptualization. **Pengyuan Zheng:** Writing – original draft, Supervision, Project administration, Data curation, Conceptualization. **Simeng Liu:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Project administration, Investigation, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

The authors declare no competing financial or non-financial interests.

Acknowledgements

Thanks to the researchers who gave their data for this analysis.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tranon.2026.102816.

References

[1] Y. Xi, P. Xu, Global colorectal cancer burden in 2020 and projections to 2040, Transl. Oncol. 14 (10) (2021) 101174.
[2] X.J. Luo, et al., Novel genetic and epigenetic biomarkers of prognostic and predictive significance in stage II/III colorectal cancer, Mol. Ther. 29 (2) (2021) 587–596.
[3] H. Zhou, et al., Liquid biopsy at the frontier of detection, prognosis and progression monitoring in colorectal cancer, Mol. Cancer 21 (1) (2022) 86.
[4] A. Audisio, et al., Neoadjuvant chemotherapy for early-stage colon cancer, Cancer Treat. Rev. 123 (2024) 102676.
[5] F. Bi, et al., Iparomlimab (QL1604) in patients with microsatellite instability-high (MSI-H) or mismatch repair-deficient (dMMR) unresectable or metastatic solid tumors: a pivotal, single-arm, multicenter, phase II trial, J. Hematol. Oncol. 17 (1) (2024) 109.
[6] K.E. de Visser, J.A. Joyce, The evolving tumor microenvironment: from cancer initiation to metastatic outgrowth, Cancer Cell 41 (3) (2023) 374–403.
[7] M.Z. Jin, W.L. Jin, The updated landscape of tumor microenvironment and drug repurposing, Signal. Transduct. Target. Ther. 5 (1) (2020) 166.
[8] S. Tian, et al., Tumour-associated neutrophils secrete AGR2 to promote colorectal cancer metastasis via its receptor CD98hc-xCT, Gut 71 (12) (2022) 2489–2501.
[9] L. Zhang, et al., Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer, Cell 181 (2) (2020) 442-459.e29.e29.
[10] A. Hasan, et al., Deregulated transcription factors in the emerging cancer hallmarks, Semin. Cancer Biol. 98 (2024) 31–50.
[11] S. Yang, et al., Single-cell and spatial transcriptome profiling identifies the transcription factor BHLHE40 as a driver of EMT in metastatic colorectal cancer, Cancer Res. 84 (13) (2024) 2202–2217.
[12] P. Cejas, et al., Transcriptional regulator CNOT3 defines an aggressive colorectal cancer subtype, Cancer Res. 77 (3) (2017) 766–779.
[13] T. Hirsch, et al., IRF4 impedes human CD8 T cell function and promotes cell proliferation and PD-1 expression, Cell Rep. 43 (7) (2024) 114401.
[14] R.Q. Wang, et al., Single-cell RNA sequencing analysis of the heterogeneity in gene regulatory networks in colorectal cancer, Front. Cell Dev. Biol. 9 (2021) 765578.
[15] L. Ziegler-Heitbrock, et al., Nomenclature of monocytes and dendritic cells in blood, Blood 116 (16) (2010) e74–e80.
[16] C.E. Olingy, H.Q. Dinh, C.C. Hedrick, Monocyte heterogeneity and functions in cancer, J. Leukoc. Biol. 106 (2) (2019) 309–322.
[17] H. Wang, et al., Decreased CXCR2 expression on circulating monocytes of colorectal cancer impairs recruitment and induces re-education of tumor-associated macrophages, Cancer Lett. 529 (2022) 112–125.
[18] J.P. Vayrynen, et al., Prognostic significance of myeloid immune cells and their spatial distribution in the colorectal cancer microenvironment, J. ImmunOther Cancer 9 (4) (2021) e002297.
[19] I. Larionova, et al., PFKFB3 overexpression in monocytes of patients with colon but not rectal cancer programs pro-tumor macrophages and is indicative for higher risk of tumor relapse, Front. Immunol. 13 (2022) 1080501.
[20] C. Goudot, A. Coillard, A.C. Villani, et al., Aryl hydrocarbon receptor controls monocyte differentiation into dendritic cells versus macrophages, Immunity 47 (3) (2017) 582–596.e6.
[21] J. Pan, et al., The bioinformatics-based analysis identifies 7 immune-related genes as prognostic biomarkers for colon cancer, Front. Oncol. 11 (2021) 726701.
[22] K. Ochiai, et al., Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4, Immunity 38 (5) (2013) 918–929.
[23] A.L. Shaffer, et al., IRF4 addiction in multiple myeloma, Nature 454 (7201) (2008) 226–231.
[24] R.L. Boddicker, et al., The oncogenic transcription factor IRF4 is regulated by a novel CD30/NF-kappaB positive feedback loop in peripheral T-cell lymphoma, Blood 125 (20) (2015) 3118–3127.
[25] R.W.J. Wong, et al., Feed-forward regulatory loop driven by IRF4 and NF-kappaB in adult T-cell leukemia/lymphoma, Blood 135 (12) (2020) 934–947.
[26] G. Alvisi, et al., IRF4 instructs effector treg differentiation and immune suppression in human cancer, J. Clin. Invest. 130 (6) (2020) 3137–3150.
[27] J. Lu, et al., Regulatory effects of IRF4 on immune cells in the tumor microenvironment, Front. Immunol. 14 (2023) 1086803.
[28] C. Li, et al., Prognostic value of tumour microenvironment-related genes by TCGA database in rectal cancer, J. Cell Mol. Med. 25 (12) (2021) 5811–5822.
[29] J. Wang, et al., RNF2 promotes the progression of colon cancer by regulating ubiquitination and degradation of IRF4, Biochim. Biophys. Acta 1869 (1) (2022) 119162.
[30] J. Wang, et al., IRF4 overexpression promotes the transdifferentiation of tregs into macrophage-like cells to inhibit the development of colon cancer, Cancer Cell Int. 21 (1) (2021) 58.
[31] L.B. Kennedy, A.K.S. Salama, A review of cancer immunotherapy toxicity, CA Cancer J. Clin. 70 (2) (2020) 86–104.
[32] K. Man, et al., Transcription factor IRF4 promotes CD8(+) T cell exhaustion and limits the development of memory-like T cells during chronic infection, Immunity 47 (6) (2017) 1129-1141.e5.
[33] K. Matsushima, D. Yang, J.J. Oppenheim, Interleukin-8: an evolving chemokine, Cytokine 153 (2022) 155828.
[34] K. Ghasemi, C-C motif glycoprotein ligand 5 (CCL5) and its GPCR CCR5: macromolecular game-changers in cancer biology, Int. J. Biol. Macromol. 329 (Pt 1) (2025) 147737.
[35] N. Schleussner, et al., Transcriptional reprogramming by mutated IRF4 in lymphoma, Nat. Commun. 14 (1) (2023) 6947.