scientific reports

Check for updates

OPEN

Identifying colorectal cancer subtypes and establishing a prognostic model using metabolic plasticity and ferroptosis genes

Jingwen Guan[1]✉, Simin Min[2], Yan Xia[1], Zhiguo Guo[3] & Xiaolan Zhou[3]

Metabolic plasticity and ferroptosis are essential for colorectal cancer (CRC) progression. The effects and prognostic value of metabolic plasticity- and ferroptosis-related genes (MPFRGs) in CRC remain unclear. We established a prognostic model for CRC patients by identifying important genes in metabolic plasticity and ferroptosis. Data of CRC patients were retrieved from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus; MPFRG data were obtained from GeneCards and FerrDb. We performed functional (to explore differences between the two metabolic subtypes) and single-sample gene set (to assess the immune environment) enrichment analyses. Immunophenotype, tumor immunological dysfunction, and exclusion scores were assessed to determine patient immune responses. A least absolute shrinkage and selection operator-Cox regression model comprising 10 significant differentially expressed genes of metabolic plasticity and ferroptosis (MPFDEGs) was constructed using TCGA training cohort and validated using the GSE17536 and GSE39582 datasets. We established a nomogram comprising metabolic plasticity- and ferroptosis-based signatures, revealing the clinical application and potential molecular mechanisms underlying the role of MPFRGs in CRC. Our model (developed based on 10 MPFDEGs) is efficient for calculating the overall survival of CRC patients. Our findings provide new strategies for the clinical management and individualized treatment of these patients.

Keywords Bioinformatics, Colorectal cancer, Ferroptosis, Immune infiltration, Metabolic plasticity, Prognostic factor

Colorectal cancer (CRC) is the most common malignancy of the digestive system[1]. The prognosis of this tumor is poor, and its pathogenesis remains unclear. Treatment can be tailored to individual patients to reliably identify the risk of progression among those with CRC. The effectiveness of CRC-specific prognostic models, constructed based on multiple gene expression, has been demonstrated in previous studies[2–4]. Each step in cancer progression requires tumor cells to reprogram their metabolic state. When survival microenvironments are threatened, tumor cells genetically regulate and alter their metabolic pathways to adapt to harsh environments. This process is called metabolic plasticity[5]. Metabolic plasticity is a hallmark of cancer[6]. The metabolic plasticity of tumor cells is mediated by oncogenes and tumor suppressors, which are involved in signaling pathways and are crucial influencing molecules that activate aerobic glycolysis or lipid metabolism[7]. Oncogene activation can impact several metabolic plasticity-related genes (MRGs), such as *c-MYC*, which activates many aerobic glycolytic enzyme genes, promoting oxidative phosphorylation. Another oncogene, *AKT*, directly promotes oxidative phosphorylation and glycolytic uncoupling reactions by upregulating the genes encoding aerobic glycolytic enzymes[8,9]. Metabolic plasticity includes lipid metabolism, which is vital in tumor invasion and metastasis[10,11]. As the tumor immune microenvironment markedly influences tumor development and treatment outcomes, metabolic plasticity is closely associated with it[12]. Metabolic plasticity influences T-cell differentiation and function and macrophage polarization and function in the tumor microenvironment[13]. The emergence of new drug resistance mechanisms in tumors may be related to abnormal lipid metabolism remodeling[14]. Recent studies have explored the relationship between metabolic plasticity and patient survival[4,8,9,15]. Furthermore, multi-model risk assessments based on MRGs have been used for gastric cancer[16], CRC[17], and osteosarcoma[18].

[1]Department of Pathology, Suzhou Hospital of Anhui Medical University, Anhui, China. [2]Department of Science and Education Section, Suzhou Hospital of Anhui Medical University, Anhui, China. [3]Department of Gastroenterology, Suzhou Hospital of Anhui Medical University, Anhui, China. ✉email: gjw595525671@163.com

Therefore, MRGs can be used as diagnostic markers and therapeutic targets. However, whether metabolic plasticity functions as a diagnostic, therapeutic, or prognostic tool for CRC remains unknown.

Ferroptosis, a newly discovered mode of cellular autophagy, is caused by lipid peroxide overload on cellular membranes[19]. It is closely associated with tumor initiation and development and various cellular processes, such as redox homeostasis maintenance, iron treatment, sugar and lipid metabolisms, and mitochondrial activity[20]. In addition, multiple tumor suppressors, such as *p53* and BRCA1-associated protein 1 (BAP1), are affected by ferroptosis, confirming that ferroptosis inherently hinders cancer progression[21]. In treatment-resistant tumor cells, metabolic plasticity and ferroptosis occur frequently. Thus, the induction and inhibition of ferroptosis and the regulatory pathways of cell metabolism are significant in treating patients with drug-resistant cancers[22]. Metabolic plasticity and ferroptosis are associated with therapeutic immunity[13,23]. Blocking immune checkpoints has recently been reported as an effective anticancer treatment approach[24]. Studies on CRC based on metabolic plasticity and ferroptosis have significant clinical applications and therapeutic prospects. However, the effects and prognostic value of metabolic plasticity- and ferroptosis-related genes (MPFRGs) in CRC remain unexplored. In this study, we aimed to investigate the relationship among CRC, metabolic plasticity, and ferroptosis to better understand the underlying mechanisms of and improve the effectiveness of individual treatments for CRC. Although studies have discussed the prognosis of patients with colon and rectal adenocarcinoma (COADREAD) using gene expression and clinical features, a comprehensive risk-scoring model is lacking for MPFRGs. In this study, least absolute shrinkage and selection operator (LASSO) regression model and multivariate Cox regression analysis were used to screen related genes, and the performance of the model was evaluated using a calibration curve and receiver operating characteristic curve, aiming to build a risk-scoring model based on MPFRGs to predict the prognosis of COADREAD patients and evaluate their response to immunotherapy through risk scoring.

Results

Impact of MPFRGs on CRC in the cancer genome atlas (TCGA)-CRC that includes a COADREAD dataset

First, CRC-related genes were obtained from TCGA database, MPFDEGs were screened through differential expression analysis, and Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed. The prognostic model of CRC was constructed and divided into two subgroups. Subsequently, the correlation between this model and immune infiltration and clinical features was determined; the specific flowchart of this process is shown in Fig. S1. Further, mRNA expression analyses in normal and cancer samples from TCGA-COADREAD indicated that over 70% of the MPFRGs exhibited significant differences ($p < 0.05$) between the two categories (Fig. 1a). Analyses involving volcano plots and heatmaps revealed a considerable disparity in the MPFRGs between the COADREAD and normal cohorts, indicating low glutathione S-transferase M1 (*GSTM1)* expression and high interleukin 1B (*IL1B)* expression in the COADREAD group (Fig. 1b, c). Single-nucleotide mutation analysis showed that MPFRGs were mutated in 317 samples, with a mutation frequency of 70.76%, with 62% being the maximum mutation rate in the *TP53* gene (Fig. 1d). We downloaded the data of "Masked Copy Number Segment" from TCGA database using TCGAbiolinks package and collected the copy number variation data of 611 patients with COADREAD. Then, we used the ggplot2 package in R to visually analyze these data. During the analysis, we counted the gain and loss of gene copy number in each sample and drew the corresponding frequency distribution map. The results show that there are changes in copy number in most samples, mainly focusing on the loss of copy number (Fig. 1e). Finally, we analyzed the positions of the 11 MPFRGs on human chromosomes and drew chromosomal localization maps (Fig. 1f).

Construction of molecular subtypes of MPFRGs in TCGA-COADREAD dataset

A correlation heatmap showed a positive correlation among most MPFRGs (Fig. 2a). *IL6* exhibited a significantly positive correlation with *IL1* (Fig. 2b), whereas *PTEN* showed a significantly negative correlation with *G6PD* (Fig. 2c). We used a consensus clustering technique to group the specimens of all patients with TCGA-COADREAD according to the expression patterns of the 11 MPFRGs (Fig. 2d, e). Accordingly, we clustered all samples into two categories and performed a prognostic analysis. Survival was significantly different in both subtypes (log-rank $p = 0.022$; Fig. 2f), demonstrating the accuracy of the clustering results. The differential expression box diagram of MPFRGs showed that, excluding *PPARA* and *G6PD*, other MPFRGs showed significant differences in expression among different groups (Fig. 2g, h). *HIF1A*, *IL6*, and *IL1B* were highly expressed in the C2 class with a good prognosis; however, they showed low expression in the C1 class with a poor prognosis.

Gene set variation analysis (GSVA) of the MPFRG model in TCGA-COADREAD dataset

GSVA was used to analyze the expression profiles of the two subtypes for hallmark gene set enrichment (Fig. 3a, Supplementary Table S1). The ggplot function in R was used to plot the proportion of disease locations in the different MPFRG molecular subtypes. C2, with a better prognosis, accounted for more tumors in the colon than in the rectum (C2: 29.4% in the colon and 25.0% in the rectum; Fig. 3b). Subsequently, using the ggplot function in R, a boxplot of differences in the body mass index (BMI) among MPFRG molecular subtypes was plotted. Patients in C1 with a negative prognosis showed significantly elevated BMI compared with those in C2 with a positive prognosis ($p = 0.015$; Fig. 3c). A Sankey diagram was created using the R ggalluvial package to illustrate the correlation among staging, MPFRG molecular subtypes, and N staging (Fig. 3d). The results showed that most patients at the C2 stage with good prognosis were in stages I, II, and III and the N0 stage with good prognosis; most of the N1 and N2 stage patients with poor prognosis were included in the C1 stage.

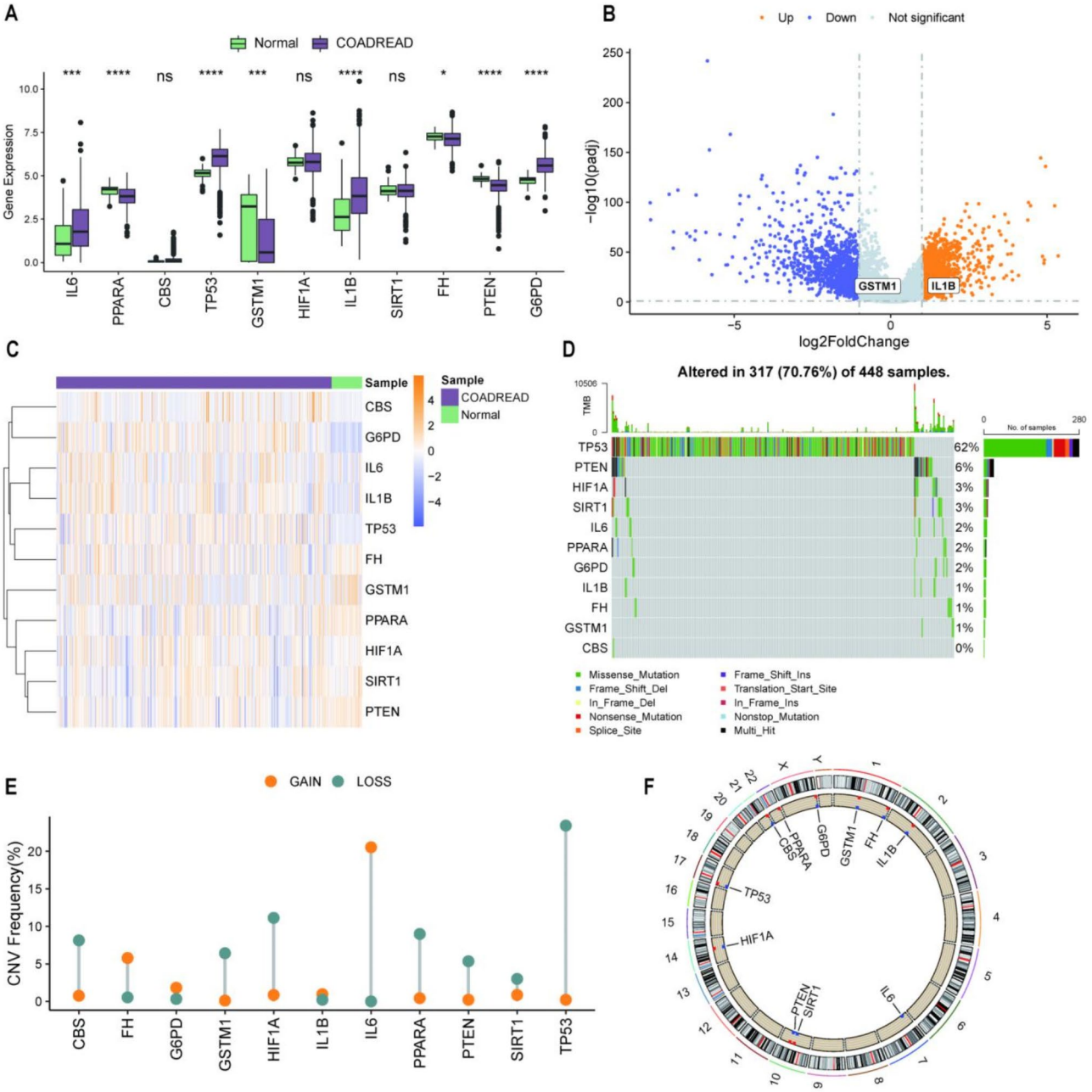

Fig. 1. Impact of metabolic plasticity- and ferroptosis-related genes (MPFRGs) on colorectal cancer (CRC) in The Cancer Genome Atlas (TCGA)-colon and rectal adenocarcinoma (COADREAD) dataset (**a**) mRNA levels in normal and cancer samples are compared using TCGA-COADREAD data. (**b**,**c**) The heat map and volcano plots of MPFRGs; (**c**) Single-nucleotide mutations of MPFRGs. (**d**–**f**) Genetic mutation landscape in the normal and CRC groups. $^{*}p < 0.05$, $^{***}p < 0.001$, $^{****}p < 0.0001$.

Functional enrichment analysis of MPFDEG subtypes in TCGA-COADREAD

In patients with CRC, we identified 850 significant MPFDEGs (Supplementary Table S2). Unsupervised clustering was performed on all patient samples using a consistent clustering algorithm. These two clusters produced the best clustering effect (Fig. 4a). Subsequently, we performed a prognostic evaluation of the two sample types. We observed a significant variation in survival rates between the two categories (log-rank $p = 0.019$; Fig. 4b), validating the accuracy of the clustering results. Based on GO and KEGG functional enrichment analyses, 850 MPFDEGs were identified. The results showed that the primary biological processes were "leukocyte migration", "leukocyte cell-cell adhesion", "cell chemotaxis", "myeloid leukocyte migration", "leukocyte chemotaxis", and "regulation of cell-cell adhesion" (Fig. 4c); the major cellular component included "membrane microdomain", "endocytic vessel", "collagen trimer", "endoplasmic reticulum lumen", tertiary granule", and "membrane raft" (Fig. 4d); enrichment of molecular function terms was observed in "cytokine binding", "glycosaminoglycan

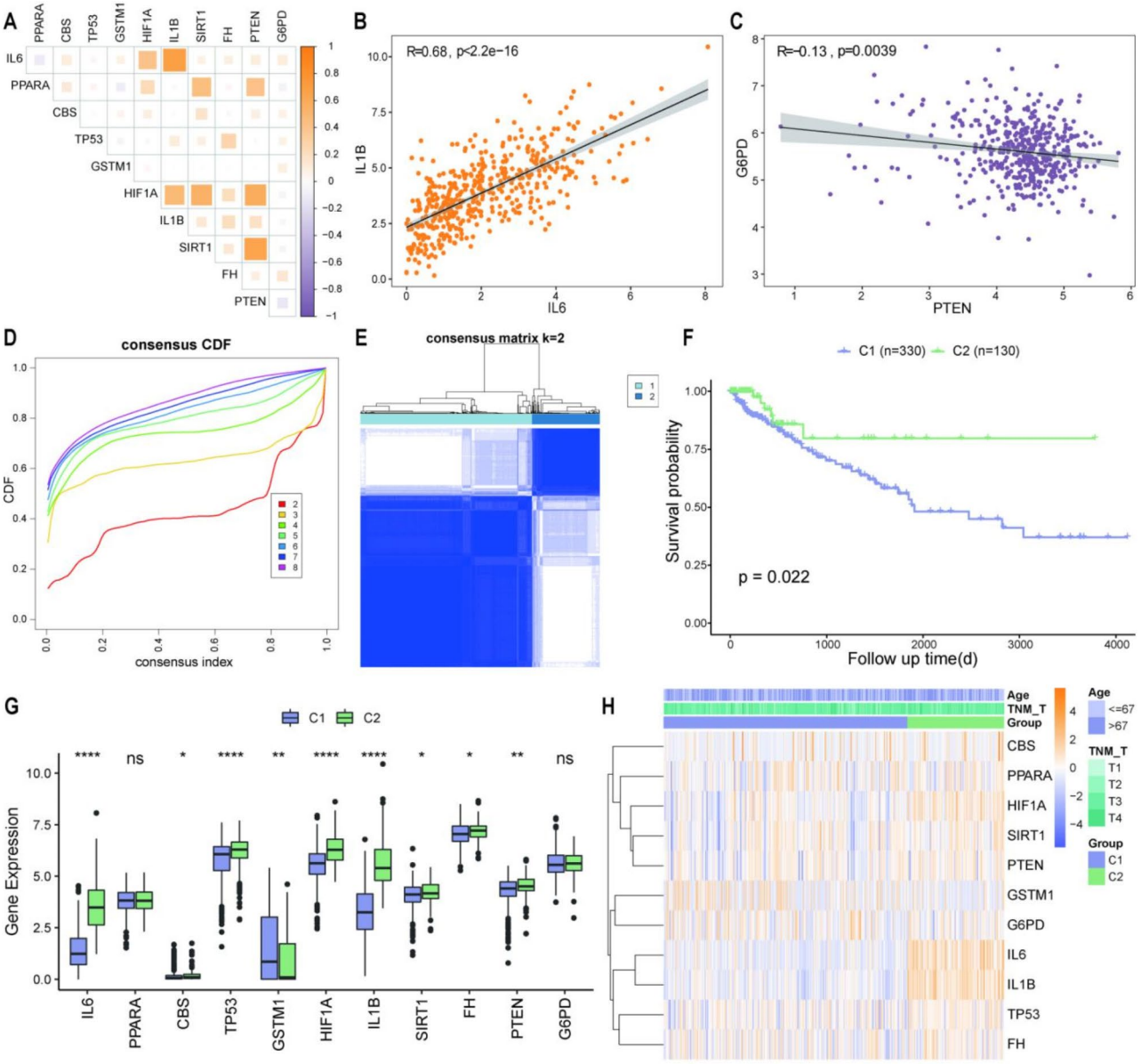

Fig. 2. Construction of molecular subtypes of metabolic plasticity- and ferroptosis-related genes (MPFRGs) in the cancer genome atlas (TCGA)-colon and rectal adenocarcinoma (COADREAD) dataset. (**a**) The correlation between MPFRGs: Positively correlated genes (**b**) and negatively associated genes (**c**). **(d)** A consensus cumulative distribution function (CDF) plot and consensus index for k = 2 to 8. **(e)** Consensus matrix heatmap of two subtypes (k = 2). **(f,g)** Survival analysis of two subtypes of samples (**f**); box plots of differential expression of MPFRGs among different groups (**g**). (**h**) Consensus matrix heatmap includes age, group, and TNM. $*p < 0.05$, $**p < 0.01$, $****p < 0.0001$, ns: non-statistically significant.

binding", "growth factor binding", "cytokine receptor binding", "cytokine activity", and "collagen binding" (Fig. 4e; Table 1). The pathway enrichment analysis of MPFDEGs showed enrichment in KEGG pathways, such as viral protein interaction with cytosine and cytosine receptors, phagosome, *Staphylococcus aureus* infection, leishmaniasis, osteoclast differentiation, and chemokine (Fig. 4f; Table 2), PI3K-Akt (Fig. 4g), and tumor necrosis factor (TNF) signaling pathways (Fig. 4h). Correlation analysis showed the age and sex proportion of individuals with different MPFDEG subtypes (Fig. 4i, j). Among individuals aged ≤ 67 years, the proportion of G1 with better survival was higher than that of G2 with worse survival. Among individuals aged ≤ 67 years, a greater percentage of G1 indicated a more favorable prognosis than that in those aged > 67 years; the proportion of G1 with a better outlook was higher in women than in men. A Sankey diagram was drawn using the R ggalluvial function to illustrate the correlation among N staging, MPFDEG molecular subtypes, and stages (Fig. 4k). Most G1 patients with good prognosis were at the N0 stage, whereas most stages III and IV patients with poor prognosis were included in the G2 group with poor prognosis.

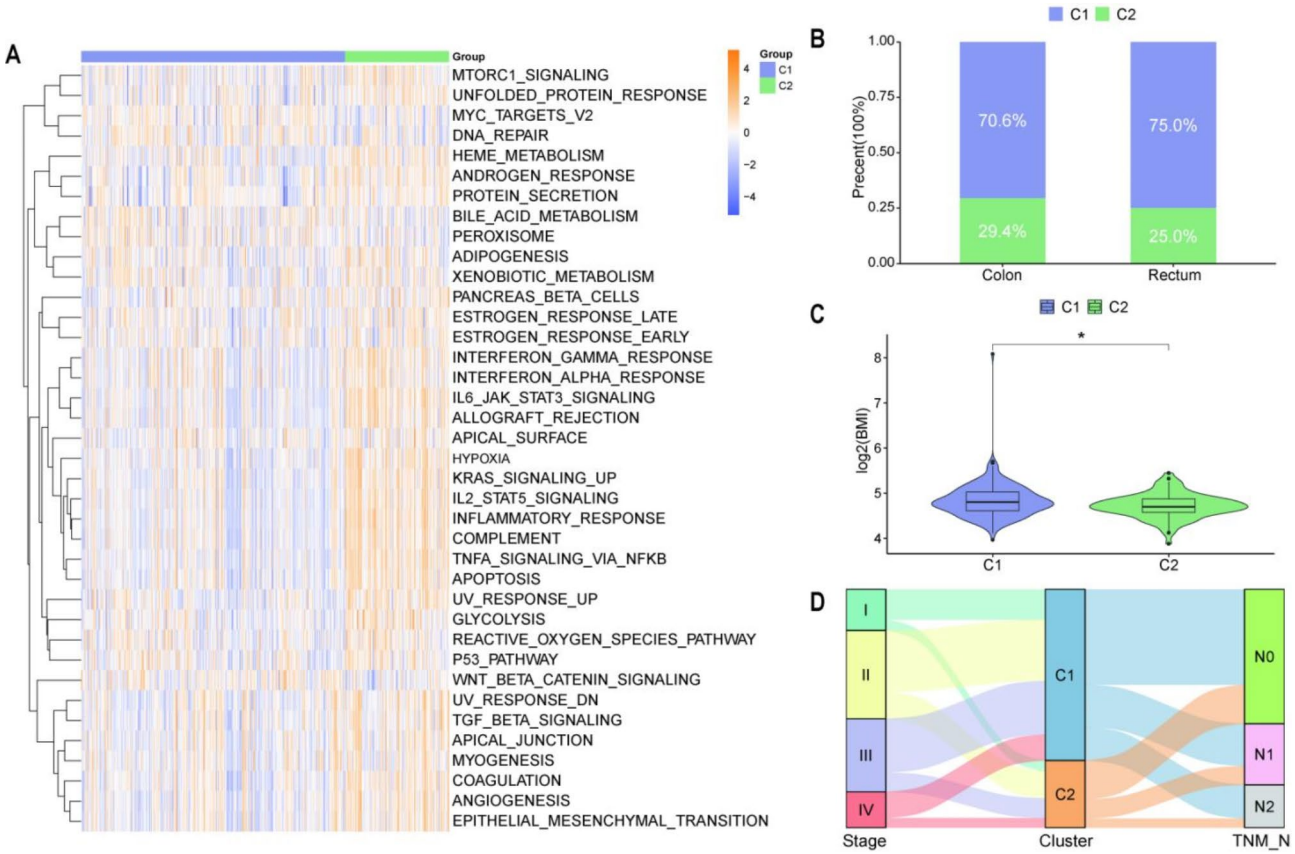

Fig. 3. Gene set variation analysis (GSVA) of the metabolic plasticity- and ferroptosis-related genes (MPFRG) model in The Cancer Genome Atlas (TCGA)-colon and rectal adenocarcinoma (COADREAD) dataset (**a**) Hallmark gene set enrichment analysis. (**b**) The proportion of disease location in different molecular subtypes. (**c**) The boxplot shows the body mass index (BMI) of the MPFRG molecular subtypes. (**d**) The Sankey diagram includes the cluster, stage, and TNM_N. $^{*}p < 0.05$.

Construction of the MPFDEG model and gene set enrichment analysis (GSEA) using TCGA-COADREAD dataset

A risk assessment model was developed using the expression levels of 850 MPFDEGs. First, singlE−factor Cox regression was used to screen 850 MPFDEGs from TCGA-COADREAD dataset, retaining 34 genes at < 0.1 for subsequent analyses (Fig. 5a, Supplementary Table S3). The optimal lambda value was determined and validated through a ten-fold cross-validation, prompting further analyses of 19 genes (Fig. 5b). In the stepwise regression analysis, the best combination of 19 genes was identified through multivariate Cox regression analysis, with a model constructed using 10 genes (Fig. 5c).

To confirm the efficacy of the developed model, TCGA-COADREAD training set patients were separated into high- and low-risk groups using a risk score cutoff value of 1.974474. Survival analysis in the TCGA dataset showed that the individuals classified as low-risk had notably longer survival periods than those classified as high-risk ($p < 0.0001$; Fig. 5d). Subsequently, the cutoffs of the risk score (GSE17536: -0.7467838; GSE39582: -0.3060653) were used to divide the validation sets of GSE17536 and GSE39582 patients into high- and low-risk clusters. The low- and high-risk clusters exhibited significant differences in GSE17536 according to survival analysis (Fig. 5e). In the GSE39582 dataset, the low-risk group exhibited notably higher survival rates than did the high-risk group (Fig. 5f). GSEA was used to analyze the TCGA data to examine the biological processes associated with gene expression in the high- and low-risk cohorts. The findings indicated that allograft rejection (Fig. 5g), protein secretion (Fig. 5h), inflammatory response (Fig. 5i), interferon-gamma response (Fig. 5j), and pathways, such as G2M checkpoint (Fig. 5k) and androgen response (Fig. 5l), were markedly enriched in the low-risk cohort (Table 3).

Immune cell infiltration in patients from TCGA-COADREAD dataset

We evaluated MPFDEG model risk scores for the immune cell infiltration levels in the COADREAD dataset. In TCGA-COADREAD dataset, the infiltration score of immune cells was visualized using the heatmap function in R using the single-sample gene set enrichment analysis (ssGSEA) algorithm to quantify and label each infiltrating immune cell type (Fig. 6a, Supplementary Table S4). Neutrophils and type 17 T helper cells were more abundant in the high-risk group than in the low-risk group. By contrast, central memory CD4 T cells, macrophages, and activated dendritic cells were more abundant in the low-risk group than in the high-risk group (Fig. 6a).

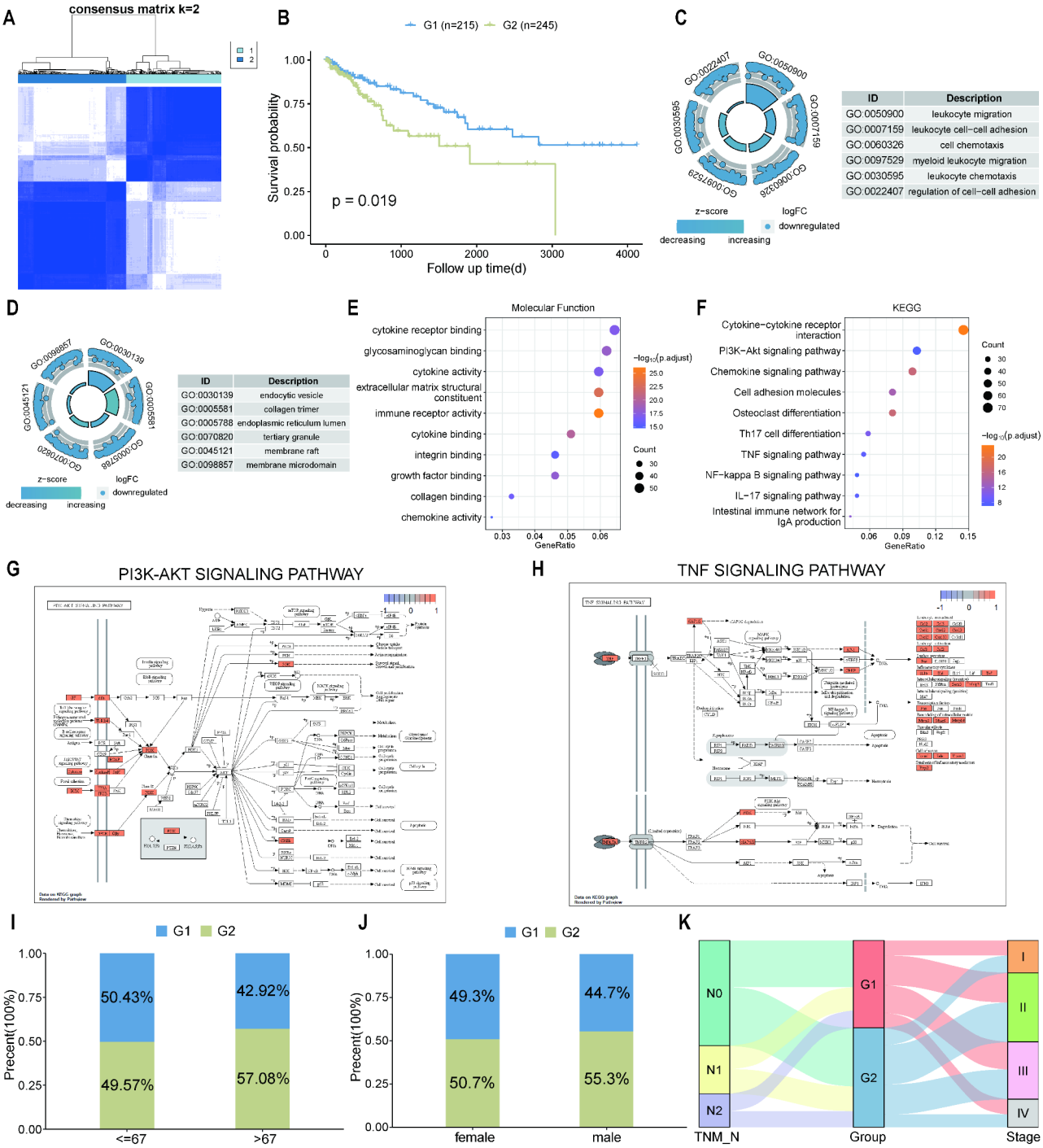

Fig. 4. Molecular subtype construction and functional enrichment analysis of differentially expressed genes of metabolic plasticity and ferroptosis (MPFDEGs) in The Cancer Genome Atlas (TCGA)-colon and rectal adenocarcinoma (COADREAD) dataset (**a**) Determination of the optimal cluster number; k = 2 is regarded as the best clustering number. (**b**) Using consensus clustering to generate classifications for survival analysis. (**c**–**f**) Gene Ontology (GO)/Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. (**g**,**h**) PI3K-Akt signaling pathway (**g**) and tumor necrosis factor (TNF) signaling pathway (**h**) are reported to be enriched. (**i**,**j**) The proportion of the two subtypes concerning age group (**i**) and sex (**j**). (**k**) The Sankey diagram includes group, stage, and TNM_N.

Subsequently, the correlation of the immune cells was plotted using the R function corrplot, revealing that most immune cells exhibited a strong correlation with each other. A negative relationship was observed between $CD56^{dim}$ natural killer cells and effector memory CD4 T cells (cor=-0.276, $p < 0.0001$; Fig. 6b). Immune B cells were negatively correlated with the risk score (cor=-0.299, $p < 0.0001$). Ten correlation heat maps between

ONTOLOGY	ID	Description	p_{adj}
BP	GO:0050900	Leukocyte migration	7.40E−61
BP	GO:0007159	Leukocyte cell–cell adhesion	1.88E−48
BP	GO:0060326	Cell chemotaxis	2.09E−47
BP	GO:0097529	Myeloid leukocyte migration	5.38E−46
BP	GO:0045785	Positive regulation of cell adhesion	2.11E−45
BP	GO:0030595	Leukocyte chemotaxis	2.11E−45
BP	GO:1903037	Regulation of leukocyte cell-cell adhesion	3.22E−39
BP	GO:0022407	Regulation of cell-cell adhesion	4.19E−39
BP	GO:0022409	Positive regulation of cell-cell adhesion	8.36E−39
BP	GO:1903039	Positive regulation of leukocyte cell-cell adhesion	1.11E−38
CC	GO:0062023	Collagen-containing extracellular matrix	9.47E−40
CC	GO:0009897	External side of plasma membrane	3.76E−39
CC	GO:0030667	Secretory granule membrane	5.56E−25
CC	GO:0030139	Endocytic vesicle	1.61E−20
CC	GO:0005581	Collagen trimer	3.29E−19
CC	GO:0005788	Endoplasmic reticulum lumen	1.18E−15
CC	GO:0070820	Tertiary granule	2.01E−15
CC	GO:0045121	Membrane raft	2.74E−14
CC	GO:0098857	Membrane microdomain	2.77E−14
CC	GO:0070821	Tertiary granule membrane	5.48E−14
MF	GO:0140375	Immune receptor activity	8.92E−27
MF	GO:0005201	Extracellular matrix structural constituent	1.39E−23
MF	GO:0019955	Cytokine binding	9.63E−21
MF	GO:0005539	Glycosaminoglycan binding	5.59E−19
MF	GO:0019838	Growth factor binding	1.17E−18
MF	GO:0005126	Cytokine receptor binding	8.65E−18
MF	GO:0005125	Cytokine activity	1.19E−17
MF	GO:0005518	Collagen binding	4.08E−17
MF	GO:0005178	Integrin binding	3.63E−16
MF	GO:0008009	Chemokine activity	1.79E−15

Table 1. GO enrichment analysis results of differential molecular subtypes of MPFRGs using TCGA-COADREAD dataset. *TCGA* The Cancer Genome Atlas, *COADREAD* colorectal adenocarcinoma, *GO* Gene Ontology, *BP* biological process, *MF* molecular function, *CC* cellular component, *MPFRGs* metabolic plasticity and ferroptosis-related genes.

ID	Description	p_{adj}
hsa05323	Rheumatoid arthritis	6.04E−32
hsa04060	Cytokine–cytokine receptor interaction	7.43E−24
hsa04640	Hematopoietic cell lineage	7.43E−24
hsa04061	Viral protein interaction with cytokine and cytokine receptor	9.09E−24
hsa04145	Phagosome	3.52E−19
hsa05150	Staphylococcus aureus infection	1.95E−18
hsa05140	Leishmaniasis	5.44E−18
hsa04380	Osteoclast differentiation	2.42E−17
hsa04062	Chemokine signaling pathway	3.09E−17
hsa05144	Malaria	3.71E−14

Table 2. KEGG enrichment analysis results of differential molecular subtypes of MPFRGs using TCGA-COADREAD dataset. *TCGA* The Cancer Genome Atlas, *COADREAD* colorectal adenocarcinoma, *KEGG* Kyoto Encyclopedia of Genes and Genomes, *MPFRGs* metabolic plasticity and ferroptosis-related genes.

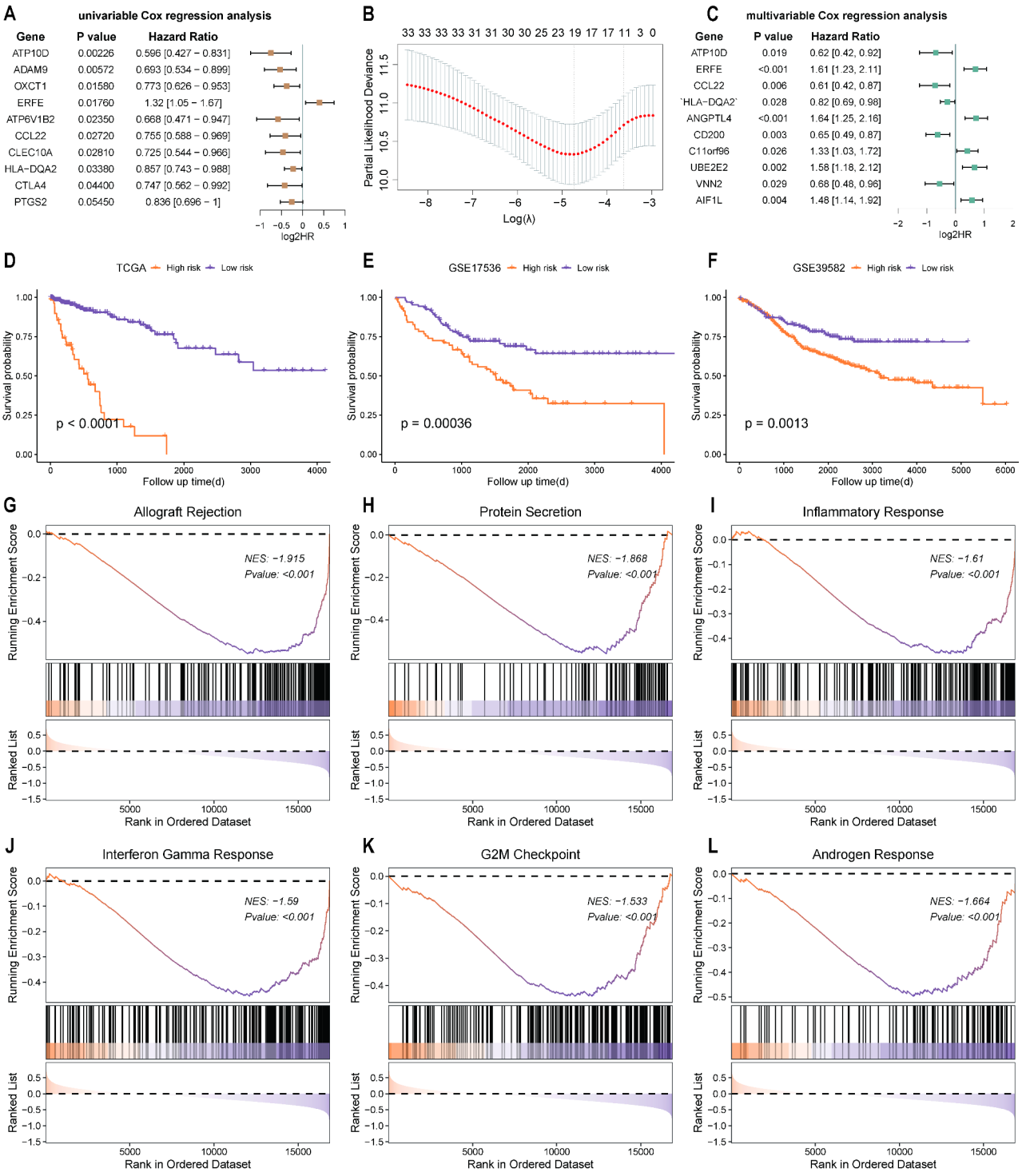

Fig. 5. Construction of differentially expressed genes of metabolic plasticity and ferroptosis (MPFDEG)-based model and gene set enrichment analysis (GSEA) in The Cancer Genome Atlas (TCGA)-colon and rectal adenocarcinoma (COADREAD) dataset. (**a**) Univariate Cox regression analysis of MPFDEGs. (**b**) Partial likelihood deviance for least absolute shrinkage and selection operator (LASSO) coefficient profiling. (**c**) Multivariate Cox regression analysis of MPFDEGs. (**d**–**f**) Survival probability curves of low- and high-risk groups in TCGA database and the GSE17536 and GSE39582 datasets. (**g**–**l**) Various biological processes (Allograft Rejection, Protein Secretion, Inflammatory Response, Interferon-Gamma Response, G2M Checkpoint, and Androgen Response) associated with gene expression in TCGA data.

ID	NES	p_{adj}
HALLMARK_ALLOGRAFT_REJECTION	− 1.914666497	4.60E−09
HALLMARK_PROTEIN_SECRETION	− 1.868347081	8.28E−07
HALLMARK_INFLAMMATORY_RESPONSE	− 1.609816713	1.66E−05
HALLMARK_INTERFERON_GAMMA_RESPONSE	− 1.590186238	1.66E−05
HALLMARK_G2M_CHECKPOINT	− 1.533218802	0.000209592
HALLMARK_ANDROGEN_RESPONSE	− 1.66351912	0.000322062
HALLMARK_KRAS_SIGNALING_UP	− 1.499031338	0.000508299
HALLMARK_MITOTIC_SPINDLE	− 1.494384742	0.000620597
HALLMARK_E2F_TARGETS	− 1.4673566	0.000620597
HALLMARK_IL6_JAK_STAT3_SIGNALING	− 1.626938781	0.001152558

Table 3. Results of GSEA enrichment analysis for high and low risk of risk scores based on TCGA-COADREAD dataset. *TCGA* The Cancer Genome Atlas, *COADREAD* colorectal adenocarcinoma, *GSEA* gene set enrichment analysis.

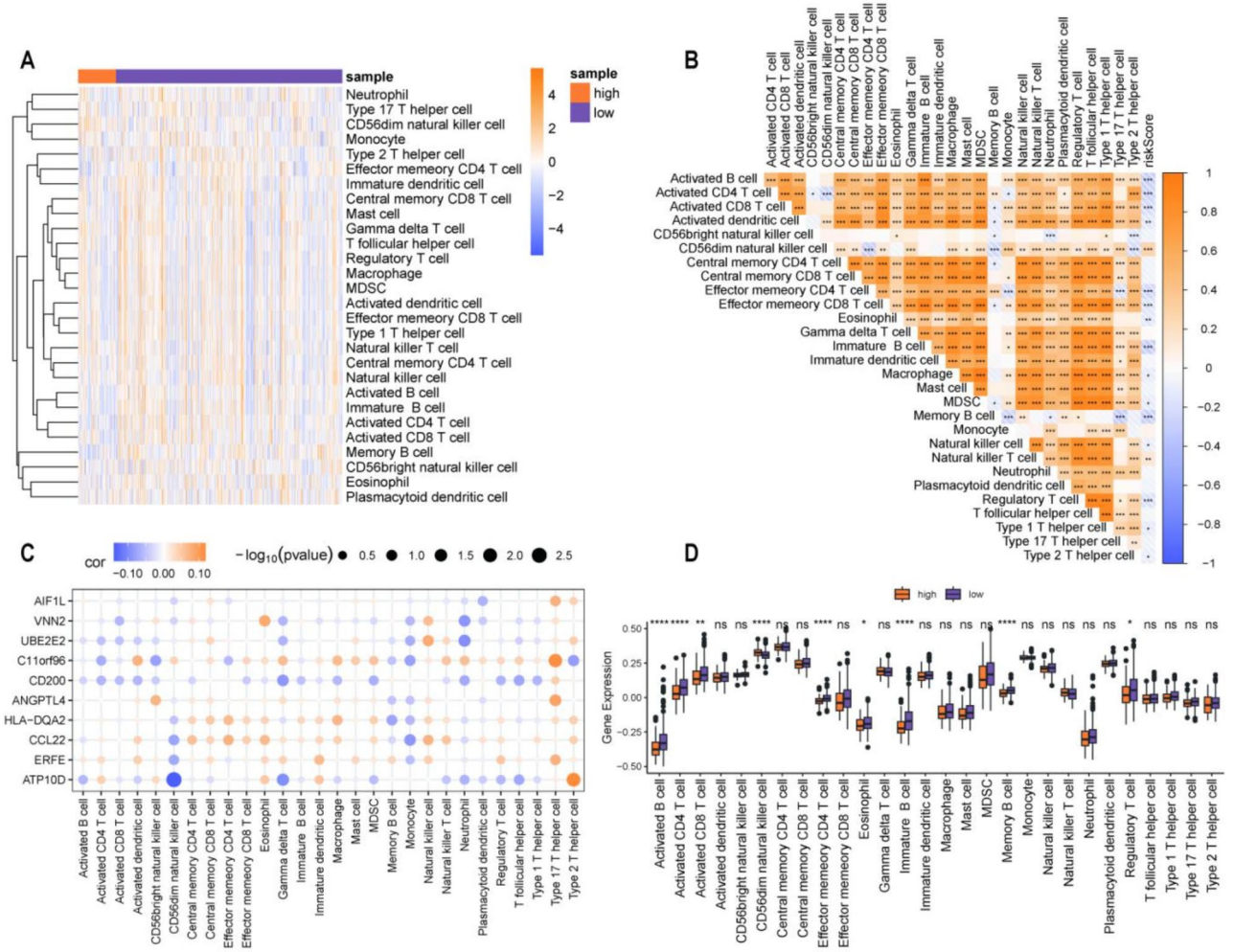

Fig. 6. Infiltration of immune cells in patients in The Cancer Genome Atlas (TCGA)-colon and rectal adenocarcinoma (COADREAD) dataset. (**a**) Immune cell infiltration scores in the high-risk and low-risk groups. (**b**) Overview and comparison of the infiltration scores of immune cells. (**c**) Heat map of the correlation between 10 differentially expressed genes of metabolic plasticity and ferroptosis (MPFDEGs) and immune cells. (**d**) Box plot depicting the differences in immune cell infiltration between the high- and low-risk groups. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$, *n.s* non-statistically significant.

MPFDEGs and immune cells were plotted using the ggplot function. The results revealed that genes such as *C11ORF96* were highly expressed in most immune cells, such as type 17 T helper cells (cor = 0.113, $p = 0.015$); genes such as *ATP10D* were marginally expressed in most immune cells, such as CD56^{dim} natural killer cells (cor=-0.139, $p = 0.003$; Fig. 6c). Next, we examined whether immune cell infiltration differed between the high- and low-risk cohorts by drawing a difference box plot using the ggplot function. Over 25% of the immune cells showed notable differences between the low- and high-risk individuals ($p < 0.05$). Additionally, the low-risk group had significantly higher infiltration levels of immune cells than did the high-risk group ($p < 0.05$; Fig. 6d).

Analysis of TCGA-COADREAD dataset to predict immunotherapy efficacy

We divided the IMvigor210 patients based on their risk score (-0.6124537) into high- and low-risk clusters. Survival was higher among the low-risk group than among the high-risk group ($p = 0.036$; Fig. 7a). The tumor immune dysfunction and exclusion (TIDE) algorithm results showed that the TIDE scores of the high-risk group were notably higher than those of the low-risk group (Fig. 7b, Supplementary Table S5).

For a more in-depth examination of the predictive impact of the risk assessment model on immunotherapy, we downloaded the immunophenotype scores (IPSs; Supplementary Table S6) related to CRC from The Cancer Imaging Archive (TCIA) database. We plotted a box plot displaying the variances in scores between the high- and low-risk clusters of TCGA-COADREAD and the IPSs using the ggplot function of R. The results showed that the high-risk individuals had significantly higher IPSs than did the low-risk individuals ($p = 0.046$; Fig. 7c). Most immune checkpoint genes exhibited notably different expression levels in differential boxplots (Fig. 7d–i).

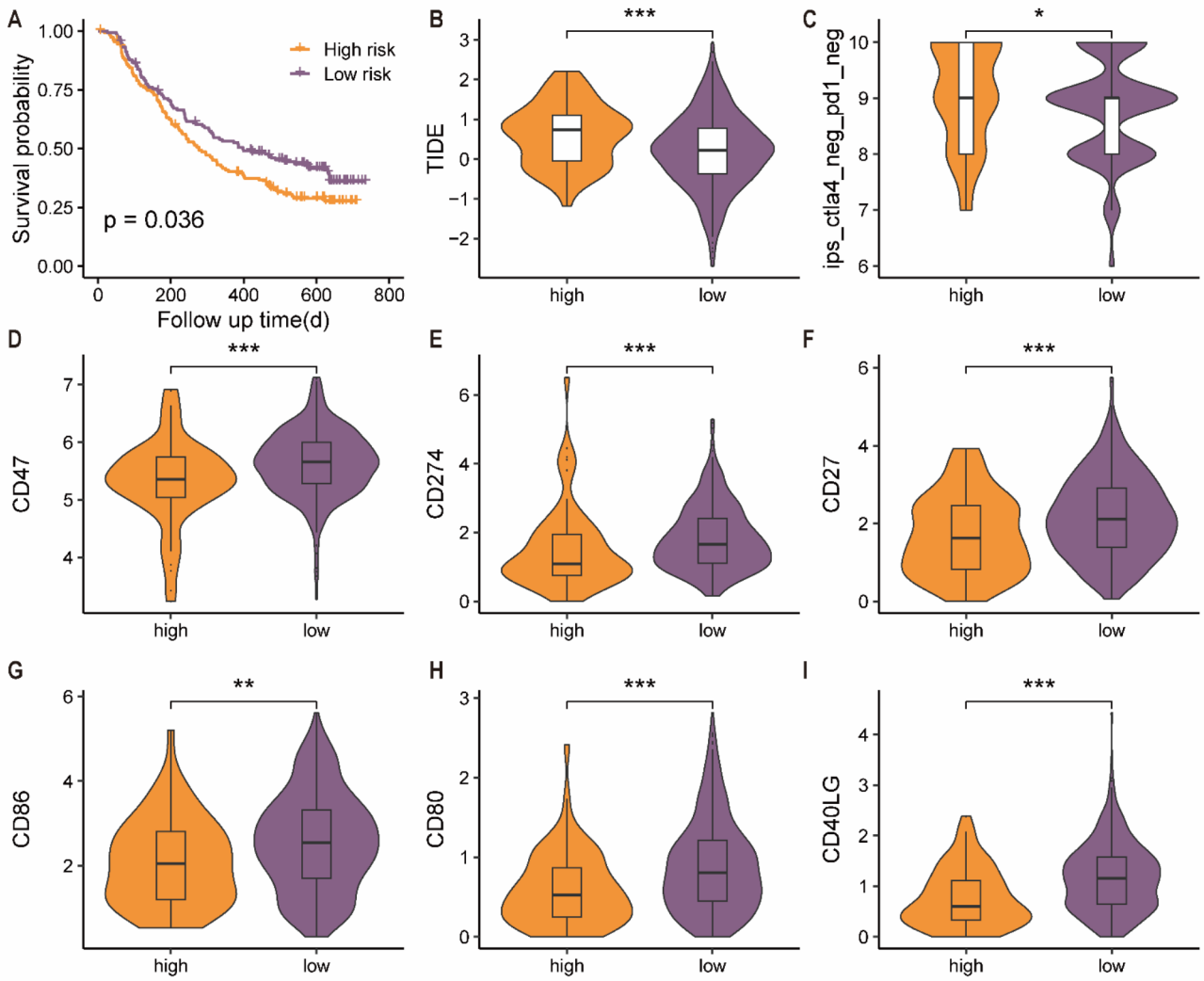

Fig. 7. Prediction of immunotherapy efficacy based on risk scores in The Cancer Genome Atlas (TCGA)-colon and rectal adenocarcinoma (COADREAD) dataset. Differences in survival probability (**a**), tumor immune dysfunction and exclusion (**b**), and immunophenotype score (IPS)-CTLA4-PD1 (**c**), between the low-risk and high-risk patients. (**d**–**i**) Differences in immune checkpoint genes between the high-risk and low-risk patients. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

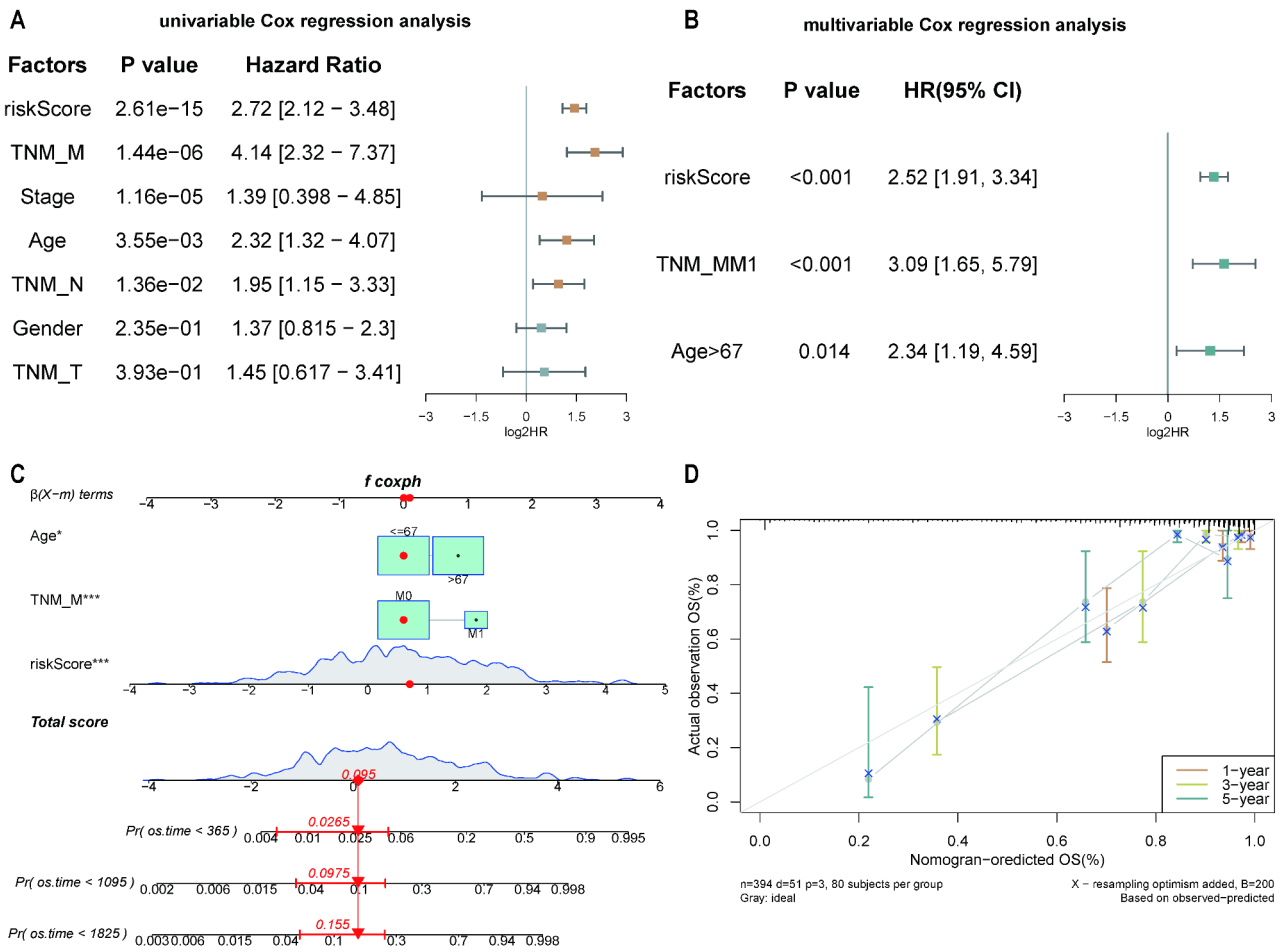

Fig. 8. Development of a clinical prediction model using The Cancer Genome Atlas (TCGA)-colon and rectal adenocarcinoma (COADREAD) risk scores. (**a**,**b**) Univariate Cox regression analysis consists of risk score, TNM_M, stage, age, TNM_N, gender, TNM_T; and multivariate Cox regression analysis consists of risk score, TNM_MM1, and age. (**c**) The nomogram shows the overall survival status of patients with COADREAD. (**d**) 1-, 3-, and 5-year survival rates of the patients predicted using established nomograms.

	Hazard ratio (HR)	Lower 95% CI	Upper 95% CI	*p* value
RiskScore	2.72	2.12	3.48	2.61E−15
TNM_M	4.14	2.32	7.37	1.44E−06
Stage	1.39	0.398	4.85	1.16E−05
Age	2.32	1.32	4.07	0.00355
TNM_N	1.95	1.15	3.33	0.0136
Gender	1.37	0.815	2.3	0.235
TNM_T	1.45	0.617	3.41	0.393

Table 4. Univariate Cox regression results of metabolic plasticity and ferroptosis gene-related score on clinical prognosis. *TNM_M* tumor node metastasis_metastasis, *TNM_N* tumor node metastasis_ node, *TNM_T* tumor node metastasis_ tumor, *CI* confidence interval.

Analysis of TCGA-COADREAD dataset for clinical prediction

Univariate Cox regression outcomes revealed that the risk score ($p < 0.001$), M-stage ($p < 0.001$), stage ($p < 0.001$), age group ($p = 0.004$), and N-stage ($p = 0.014$) were associated with the overall survival (OS) of patients with COADREAD (Fig. 8a; Table 4). The multivariate Cox regression outcomes revealed that the risk score ($p < 0.001$), M-stage ($p < 0.001$), and age group ($p = 0.014$) were markedly associated with the OS of patients with COADREAD (Fig. 8b; Table 5). Our column chart was used to predict the lifespan of patients with COADREAD based on the M-stage, risk score, and age group (Fig. 8c). Calibration curves were constructed for the column

	Hazard ratio (HR)	Lower 95% CI	Upper 95% CI	*p* value
RiskScore	2.52	1.91	3.34	< 0.001
TNM_MM1	3.09	1.65	5.79	< 0.001
Age > 67	2.34	1.19	4.59	0.014

Table 5. Multivariate Cox regression results of metabolic plasticity and ferroptosis gene-related score on clinical prognosis. *TNM_MM1* tumor node metastasis_metastasis M1, *CI* confidence interval.

charts. A comparison of the estimated OS values of patients at 1, 3, and 5 years with the actual observed values revealed good consistency between the two groups (Fig. 8d).

Discussion

Clinical evidence has confirmed the unfavorable CRC outcomes. Establishing a clinical prediction model to accurately predict the survival of patients with CRC can help better stratify and predict the risk and develop personalized treatment strategies[25]. Iron-mediated cell death links metabolism, diseases, immune cells, and targeted therapies[26]. Cancer cells exhibit metabolic plasticity, as established earlier[27]. Metabolic plasticity confers a survival advantage to cancer cells for using nutrients, rapidly producing ATP, synthesizing biomass, and balancing reactive oxygen species[28]. Metabolic plasticity and ferroptosis are common metabolic pathways in tumor cells and are extensively studied in other diseases[29–31].

Through these correlation analysis results, we preliminarily understand the interaction between MPFRGs in CRC. Most MPFRGs showed a positive correlation, indicating that these genes might act synergistically in similar biological pathways to regulate the development of CRC. *IL-6* and *IL-1*, as inflammatory factors, may jointly promote cancer progression in the microenvironment of CRC, suggesting that we need to further study the specific mechanism of these inflammatory factors in cancer. The significant negative correlation between PTEN and G6PD suggests that they may inhibit each other in the processes involving metabolism and growth regulation. PTEN is an important tumor suppressor gene, and G6PD plays a key role in cell metabolism, which provides us with new potential therapeutic targets. Through these observations, we can better understand the role of MPFRGs in CRC and provide valuable directions for future research. GO analysis showed that the abnormal expression of key genes in MPFDEGs primarily involves immune receptor activity and cytokine receptor binding, considering molecular functions. KEGG analysis showed that MPFDEGs primarily involve the PI3K/Akt and cytokine–cytokine receptor interaction signaling pathways, with key genes such as *ATP10D*, *ERFE*, and *ATP10D* confirmed to be closely related to lipomics and metabolomics[32]. Alexander et al.[33] found that *ATP10D* can reduce obesity and enhance the responsiveness to insulin in mice fed a high-fat diet and that *ATP10D* transgenic mice exhibit lipid- and metabolism-related changes in the liver and plasma. Some researchers have found that new genetic variations discovered in *ATP10D* are present in individuals with extreme phenotypes, such as heightened or reduced susceptibility to tobacco-induced non-small cell lung cancer; however, studies on *ATP10D* related to metabolic plasticity and ferroptosis in CRC are lacking. Additionally, *ERFE* is involved in tumor progression and invasion through its association with *TP53* mutations[34], co-participation in the P13K/AKT and mTOR signaling pathways[35] with IL-11, and by possibly activating the NOTCH-related signaling pathway[36].

The genes related to tumor differentiation, metastatic risk, and patient survival in MPFDEGs may serve as potential candidates for developing new therapeutic targets for CRC, such as angiopoietin-like 4 (*ANGPTL4*), *UBE2E2*, *ERFE*, *VNN2*, *CCL22*, and *C11orf96*. *ANGPTL4* is related to lipid metabolism and regulates the distribution of lipoproteins and fatty acids in peripheral tissues by inhibiting the activity of lipoprotein lipase in various environments[37]. Similarly, *ANGPTL4* is related to cardiovascular and lipid metabolic diseases, such as atherosclerosis and type 2 diabetes, and tumorigenesis, tumor angiogenesis, and tumor cell migration[38,39]. The levels of ROS, MMP-1, and MMP-9 are substantially reduced upon ANGPTL4 deletion, inhibiting the transactivation of NADPH oxidase 4 (NOX4) by c-Jun, causing an enhancement in the migration ability of CRC cells mediated by oleic acid and further affecting the survival rates of patients with CRC[40]. This observation supports the results of our study. *UBE2E2*, a member of the ubiquitin-coupled enzyme (UBE2) family, is upregulated in ovarian malignant tumors and negatively correlated with patient prognosis. Ovarian cancer cells are stimulated to proliferate, migrate, and undergo epithelial–mesenchymal transition by activating UBE2, possibly through the UBE2E2-Nrf2-p62-Snail signaling axis[41]. *ERFE*, also known as *CTRP15*, is a member of the C1q tumor necrosis factor-related protein family. In a pan-cancer study, Xu et al.[42] found that *ERFE* overexpression indicated poor prognosis of breast cancer, colon cancer, pancreatic cancer, and renal clear cell carcinoma, consistent with the results of this study. This aspect requires further functional analysis and may have potential clinical value, as *ERFE* could be developed as a therapeutic target for CRC. However, reports on the role of *ERFE* in ferroptosis and metabolic plasticity in tumors are unavailable. The glycosylphosphatidylinositol-anchored cell surface protein vanin-2 (*VNN2*) may be an independent predictor of recurrence in most patients with acute lymphoblastic leukemia, and its expression may increase cell chemotherapy resistance[43]. Chen et al. found that the *CCL22* chemokines produced by tumor-associated macrophages are positively correlated with tumor invasion and low survival rates in patients with esophageal squamous cell carcinoma[44]. The possible mechanism is that *CCL22* is expressed through the chemokine receptor diacylglycerol kinase α. By connecting the FAK signaling pathway and activating the FAK/AKT pathway, CCL22-stimulated FAK overactivation accelerates the progression and invasion of malignant tumor cells. Moreover, some researchers have found that *CCL22* may act as an independent prognostic factor for patients with CRC[45]. Other studies report that patients with lung and breast cancers may be treated by inhibiting *C11orf96* expression[46,47].

The immune system of patients with CRC is generally suppressed; among the MPFDEGs, *ERFE*, *CCL22*, *CD200*, and *HLA-DQA2* may be the potential immune regulation-related genes. The aberrant expression of these genes may affect tumor immune escape and the efficacy of immunotherapy. Srole et al.[48] have found that *ERFE* overexpression can cause enrichment reactions in Th2 immune cells. Chapoval et al.[49] showed that tumor regulatory T cells (Tregs) overexpressing *ERFE* were not significantly enriched. In this case, *EFRE* overexpression caused no increased infiltration of type 2 T helper cells. The chemokine *CCL22*, primarily produced by immune dendritic cells, binds to the CCR4 receptor to recruit factors in tumor cells to regulate Treg migration, affecting T-cell immune processes[49]. *HLA-DQA2*, a class II human leukocyte antigen (HLA), is crucial in antigen presentation and immune response initiation. As the invasive ability of early lung adenocarcinoma increases, its antigen presentation ability is also enhanced, and the *HLA-DQA2* genes involved in antigen presentation are upregulated[50]. *HLA-DQA2* expression is significantly increased in breast cancer[51]; however, its correlation with tumor prognosis remains unknown. In this case, *HLA-DQA2* was reported for the first time as an MPFDEG that is an independent predictor of prognosis. The cell surface glycoprotein CD200 (OX-2; OX-91) is an emerging checkpoint ligand; however, it is understudied[52]. CD200R promotes myeloid cell expansion and limits antitumor immunity (e.g., in brain and ovary tumors)[52,53].

According to the immune escape mechanism of tumors, clinically significant cancers choose fewer immunogenic cancer cells (immune selection) to avoid antitumor immune responses, hindering cancer cell recognition by $CD8^+$ T cells. In this study, the expression of immune checkpoint CTLA-4 and PD-1 inhibitors was notably higher in the high-risk cohort than in the low-risk cohort ($p < 0.05$). As CD4 memory resting T cells differentiate, they can help $CD8^+$ T cells reject tumors, inhibit harmful immune responses to antigens, inhibit $CD8^+$ T-cell activation, and kill natural killer cells. Thus, the risk score is consistent with immune checkpoint expression assessed in tumor-infiltrating immune cells, indicating that the model has good clinical efficacy.

This study has some limitations. The results are unvalidated in wet experiments, and corresponding clinical correlation studies are lacking, complicating the analysis of the findings in conjunction with clinical information. Moreover, this study included numerous datasets, which may have caused unavoidable inter-batch differences or insufficient sample sizes during the analysis. Given the above shortcomings, we will collect many clinical specimens to test the accuracy and robustness of this model. Cell and animal experiments were performed to study the effect of the overexpression or knockout of related genes on the occurrence and development of CRC.

However, this study also has some strengths. Our findings revealed the clinical application and potential molecular mechanisms of MPFDEGs in CRC. A model based on 10 MPFDEGs (*ATP10D*, *ERFE*, *CCL22*, *HLA-DQA2*, *ANGPTL4*, *CD200*, *C11orf96*, *UBE2E2*, *VNN2*, and *AIF1L*) was established that can be used to predict the OS of patients with COADREAD; these genes could serve as potential therapeutic targets. The performance of the developed risk model was validated using multiple validation methods, which confirmed its robustness and accuracy. Additionally, the effects of MPFDEG expression on the tumor immune microenvironment were established.

Methods

Data

Expression profiles and corresponding clinical information, including data on copy number variations, were obtained from TCGA (https://portal.gdc.cancer.gov/; Table 6). Data were obtained from the Gene Expression Omnibus database to verify the prognostic potential of genetic risk scores. Datasets GSE17536 and GSE39582, related to CRC (COADREAD), were retrieved for *Homo sapiens*, and the data platform was GPL570. No datasets are available for CRC immunotherapy; therefore, we referred to a study on colon cancer[54] and verified with the immunotherapy cohort data for bladder cancer. MRGs (Supplementary Table S7) and ferroptosis-related genes (Supplementary Table S8) were determined using the GeneCards (https://www.genecards.org/) database and FerrDb (http://www.zhounan.org/ferrdb/current/), respectively. Eleven MPFRGs were obtained after merging and duplicating (Supplementary Table S9). Moreover, chromosome maps help to identify key genes with frequent copy number changes in cancer and visualize their specific locations on chromosomes. Simultaneously, these positional data can provide reference for us to further study the function and mechanism of these genes in CRC.

Construction of molecular subtypes based on MPFRGs

Microarray gene expression datasets were clustered using consistency clustering. TCGA expression profile data were clustered using MPFRGs in the Consensus Cluster Plus package (Version 4.2.2, https://www.r-project.org/) of R to better distinguish CRC subtypes.

GSVA

GSVA enrichment analysis was performed on the gene expression profile dataset of patients in TCGA-COADREAD dataset to explore the differences in biological processes among the different clusters. In TCGA-COADREAD dataset, the gene set "h. all. v7.5.1. symbols" is derived from MSigDB for GSVA of the MPFRG molecular subtypes.

Differential gene analysis related to metabolic plasticity and ferroptosis

Using the limma R package (v3.34.9), differential analysis of the grouping was conducted to determine the differences between molecular subtypes. We set -log10 (p_{adj} value) < 0.0001 and |logFC| > 0.5 as the thresholds to screen for MPFDEGs. Among them, genes with |logFC| > 0.5 and -log10 (p_{adj} value) < 0.0001 were considered as upregulated MPFDEGs, whereas those with |logFC|<-0.5 and -log10 (p value) < 0.0001 were considered as downregulated MPFDEGs.

	Alive (401)	Dead (59)	Total (460)
Age			
Mean	64.7	68.4	65.2
Median	66	71	67
≤67	217 (54%)	17 (29%)	234 (51%)
>67	184 (46%)	42 (71%)	226 (49%)
Gender			
Female	182 (45%)	25 (42%)	207 (45%)
Male	219 (55%)	34 (58%)	253 (55%)
T			
T1	14 (3.5%)	2 (3.4%)	16 (3.5%)
T2	72 (18%)	4 (6.8%)	76 (16.5%)
T3	275 (68.6%)	40 (67.8%)	315 (68.5%)
T4	39 (9.7%)	12 (20.3%)	51 (11.1%)
TX	1 (0.2%)	1 (1.7%)	2 (0.4%)
N			
N0	234 (58.4%)	23 (39%)	257 (55.9%)
N1	99 (24.7%)	18 (30.5%)	117 (25.4%)
N2	66 (16.4%)	17 (28.8%)	83 (18%)
NX	2 (0.5%)	1 (1.7%)	3 (0.7%)
M			
M0	298 (74.3%)	31 (52.5%)	329 (72%)
M1	45 (11.2%)	20 (33.9%)	65 (14%)
MX	58 (14.5%)	8 (13.6%)	66 (14%)
Stage			
I	73 (18.2%)	3 (5%)	76 (16.5%)
II	146 (36.4%)	17 (28.8%)	163 (35.4%)
III	120 (30%)	15 (25.4%)	135 (29.3%)
IV	46 (11.4%)	20 (34%)	66 (14.4%)
Site			
Colon	290 (72%)	50 (85%)	340 (74%)
Rectum	111 (28%)	9 (15%)	120 (26%)

Table 6. TCGA dataset information. *TCGA* The Cancer Genome Atlas.

Functional enrichment analysis

ClusterProfiler was used to perform GO annotation and KEGG pathway enrichment analyses of MPFDEGs in TCGA-COADREAD dataset; the critical value of FDR < 0.05 was considered statistically significant. Data with p_{adj} and q values < 0.05 were considered entry screening criteria, and the Benjamini–Hochberg method was used to correct the *p*-values.

MPFDEG-based risk model construction

We screened for genes associated with prognosis through univariate or multivariate Cox and LASSO regression analyses and constructed a prognostic model. The cutoff function in the R survey was used to calculate the optimal cutoff to group patients by risk score; the cutoffs were entered into TCGA-COADREAD portal to separate the high- and low-risk patients. Using the R survival package, Kaplan–Meier (KM) analysis and a logarithmic rank test were conducted on the TCGA OS rate. Based on the cutoffs, GSE17536 and GSE39582 risk scores were calculated, and the data were grouped into risk clusters.

$$\begin{aligned}\text{Risk Score} = {} & (0.48 \times\ mRNA\ Expression\,(ERFE)) + (0.5 \times\ mRNA\ Expression\,(ANGPTL4)) \\ & + (0.29 \times\ mRNA\ Expression\,(C11or96)) + (0.46 \times\ mRNA\ Expression\,(UBE2E2)) \\ & + (0.39 \times\ mRNA\ Expression\,(AIF1L)) + (-0.5 \times\ mRNA\ Expression\,(CCL22)) \\ & + (-0.48 \times\ mRNA\ Expression\,(ATP10D)) + (-0.2 \times\ mRNA\ Expression\,(HLA-DQA2)) \\ & + (-0.43 \times\ mRNA\ Expression\,(CD200)) + (-0.39 \times\ mRNA\ Expression\,(VNN2))\end{aligned}$$

GSEA

TCGA-COADREAD gene expression profiles were used to conduct GSEA to compare the biological processes between low- and high-risk groups. Our analysis of TCGA-COADREAD dataset relied on the MSigDB gene set "h. all. v7.5.1. symbols".

Analysis of immune infiltration

We downloaded immune-related genes, including 28 cell types and 782 genes, such as macrophages, Tregs, and other immune cell subtypes, from a study by Charoentong et al.[55]. The GSVA package and ssGSEA method were used to analyze TCGA-COADREAD expression profile. Immune cell correlation maps were drawn using the corrplot R package.

Immunotherapy analysis

The IPS data for CRC were obtained from TCIA database and analyzed using the R ggplot2 package. Using the TIDE algorithm, the high- and low-risk treatment responses of patients were compared. The TIDE scores and immunotherapy responses of each cancer sample were calculated using standardized TCGA-COADREAD expression profile data as the input.

Modeling clinical prediction based on MPFDEGs

Clinicopathological features can be combined with risk scores to provide a prognostic assessment tailored to each patient. Using TCGA-COADREAD expression profile data for univariate and multivariate Cox regression analyses, we examined the ability of risk scores combined with clinicopathological features of patients with CRC to predict OS. Following the selection of a risk-scoring model based on clinical pathology features that were notably correlated with OS rates ($p < 0.05$), a clinical prediction column chart was constructed using the rms package in R.

Statistical analysis

Data were analyzed using R (version 4.2.2). Two sets of continuous variables conforming to the normal distribution were compared using the Student's t-test. The Mann – Whitney U test (Wilcoxon rank-sum test) was used to process the data of non-normally distributed variables. The KM survival curve represented the survival difference. The survival time difference was evaluated using a logarithmic rank test. A p value < 0.05 was considered statistically significant, with all p-values being two-tailed.

Data availability

All data are publicly accessible, open access, and can be accessed upon reasonable request. Data relevant to this study are included in the article or uploaded as supplemental information. The CRC series used in this study is available in TCGA (https://portal.gdc.cancer.gov/), which are public repositories for functional genomics data. Gene expression data from the GEO repository (https://www.ncbi.nlm.nih.gov/geo/) include the following dat asets: GSE17536, GSE39582.

Received: 17 April 2024; Accepted: 31 October 2024
Published online: 08 November 2024

References

1. Siegel, R. L., Miller, K. D., Wagle, N. S. & Jemal, A. Cancer statistics, 2023. *CA Cancer J. Clin.* **73**, 17–48 (2023).
2. Ahluwalia, P., Kolhe, R. & Gahlay, G. K. The clinical relevance of gene expression based prognostic signatures in colorectal cancer. *Biochim. Biophys. Acta Rev. Cancer* **1875**, 188513 (2021).
3. Kamel, H. F. M. & Al-Amodi, H. S. A. B. Exploitation of gene expression and cancer biomarkers in paving the path to era of personalized medicine. *Genom. Proteom. Bioinform.* **15**, 220–235 (2017).
4. Ye, Z., Zou, S., Niu, Z., Xu, Z. & Hu, Y. A novel risk model based on lipid metabolism-associated genes predicts prognosis and indicates immune microenvironment in breast cancer. *Front. Cell. Dev. Biol.* **9**, 691676 (2021).
5. Jia, D. et al. Elucidating cancer metabolic plasticity by coupling gene regulation with metabolic pathways. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 3909–3918 (2019).
6. Ward, P. S. & Thompson, C. B. Metabolic reprogramming: a cancer hallmark even Warburg did not anticipate. *Cancer Cell* **21**, 297–308 (2012).
7. Kreuzaler, P., Panina, Y., Segal, J. & Yuneva, M. Adapt and conquer: metabolic flexibility in cancer growth, invasion and evasion. *Mol. Metab.* **33**, 83–101 (2020).
8. Faubert, B., Solmonson, A. & DeBerardinis, R. J. Metabolic reprogramming and cancer progression. *Science* **368** (2020).
9. Tong, Y., Gao, W. Q. & Liu, Y. Metabolic heterogeneity in cancer: an overview and therapeutic implications. *Biochim. Biophys. Acta Rev. Cancer* **1874**, 188421 (2020).
10. Liu, R. Z. & Godbout, R. An amplified fatty acid-binding protein gene cluster in prostate cancer: emerging roles in lipid metabolism and metastasis. *Cancers (Basel)* **12**, 3823 (2020).
11. Luo, X. et al. Emerging roles of lipid metabolism in cancer metastasis. *Mol. Cancer* **16**, 76 (2017).
12. Li, X. et al. Navigating metabolic pathways to enhance antitumour immunity and immunotherapy. *Nat. Rev. Clin. Oncol.* **16**, 425–441 (2019).
13. Xia, L. et al. The cancer metabolic reprogramming and immune response. *Mol. Cancer* **20**, 28 (2021).
14. Zhu, K., Xiaoqiang, L., Deng, W., Wang, G. & Fu, B. Development and validation of a novel lipid metabolism-related gene prognostic signature and candidate drugs for patients with bladder cancer. *Lipids Health Dis.* **20**, 146 (2021).
15. Ohshima, K. & Morii, E. Metabolic reprogramming of cancer cells during tumor progression and metastasis. *Metabolites* **11**, 28 (2021).
16. Yu, S. et al. Seven-gene signature based on glycolysis is closely related to the prognosis and tumor immune infiltration of patients with gastric cancer. *Front. Oncol.* **10**, 1778 (2020).
17. Lin, D. et al. Molecular subtype identification and prognosis stratification by a metabolism-related gene expression signature in colorectal cancer. *J. Transl. Med.* **19**, 279 (2021).
18. Qian, H., Lei, T., Hu, Y. & Lei, P. Expression of lipid-metabolism genes is correlated with immune microenvironment and predicts prognosis in osteosarcoma. *Front. Cell. Dev. Biol.* **9**, 673827 (2021).
19. Yan, H. F. et al. Ferroptosis: mechanisms and links with diseases. *Signal. Transduct. Target. Ther.* **6**, 49 (2021).
20. Jiang, X., Stockwell, B. R. & Conrad, M. Ferroptosis: mechanisms, biology and role in disease. *Nat. Rev. Mol. Cell. Biol.* **22**, 266–282 (2021).
21. Zhang, Y. et al. BAP1 links metabolic regulation of ferroptosis to tumour suppression. *Nat. Cell. Biol.* **20**, 1181–1192 (2018).

22. Lei, G., Zhuang, L. & Gan, B. Targeting ferroptosis as a vulnerability in cancer. *Nat. Rev. Cancer* **22**, 381–396 (2022).
23. Dang, Q. et al. Ferroptosis: a double-edged sword mediating immune tolerance of cancer. *Cell. Death Dis.* **13**, 925 (2022).
24. Morad, G., Helmink, B. A., Sharma, P. & Wargo, J. A. Hallmarks of response, resistance, and toxicity to immune checkpoint blockade. *Cell* **184**, 5309–5337 (2021).
25. Liu, J., Liu, Z., Li, J., Tian, S. & Dong, W. Personalizing prognostic prediction in early-onset colorectal cancer. *J. Cancer* **11**, 6727–6736 (2020).
26. Li, S. & Huang, Y. Ferroptosis: an iron-dependent cell death form linking metabolism, diseases, immune cell and targeted therapy. *Clin. Transl. Oncol.* **24**, 1–12 (2022).
27. Liberti, M. V. & Locasale, J. W. The Warburg effect: how does it benefit cancer cells? *Trends Biochem. Sci.* **41**, 211–218 (2016).
28. Wu, Z., Lee, Y. F., Yeo, X. H., Loo, S. Y. & Tam, W. L. Shifting the gears of metabolic plasticity to drive cell state transitions in cancer. *Cancers (Basel)* **13**, 1316 (2021).
29. Chen, Y., Li, X., Wang, S., Miao, R. & Zhong, J. Targeting iron metabolism and ferroptosis as novel therapeutic approaches in cardiovascular diseases. *Nutrients* **15**, 591 (2023).
30. Huang, R. et al. Adaptive changes allow targeting of ferroptosis for glioma treatment. *Cell. Mol. Neurobiol.* **42**, 2055–2074 (2022).
31. Tong, J. et al. Targeting a novel inducible GPX4 alternative isoform to alleviate ferroptosis and treat metabolic-associated fatty liver disease. *Acta Pharm. Sin. B* **12**, 3650–3666 (2022).
32. Sigruener, A. et al. Lipidomic and metabolic changes in the P4-type ATPase ATP10D deficient C57BL/6J wild type mice upon rescue of ATP10D function. *PLoS ONE*. **12**, e0178368 (2017).
33. Fusco, J. P. et al. Genomic characterization of individuals presenting extreme phenotypes of high and low risk to develop tobacco-induced lung cancer. *Cancer Med.* **7**, 3474–3483 (2018).
34. Donehower, L. A. et al. Integrated analysis of TP53 gene and pathway alterations in the cancer genome atlas. *Cell. Rep.* **28**, 3010 (2019).
35. Widjaja, A. A. et al. IL11 stimulates ERK/P90RSK to inhibit LKB1/AMPK and activate mTOR initiating a mesenchymal program in stromal, epithelial, and cancer cells. *iScience* **25**, 104806 (2022).
36. Asnaghi, L. et al. Notch signaling promotes growth and invasion in uveal melanoma. *Clin. Cancer Res.* **18**, 654–665 (2012).
37. Aryal, B., Price, N. L. & Suarez, Y. Fernández-Hernando, C. ANGPTL4 in metabolic and cardiovascular disease. *Trends Mol. Med.* **25**, 723–734 (2019).
38. Fernández-Hernando, C. & Suárez, Y. ANGPTL4: a multifunctional protein involved in metabolism and vascular homeostasis. *Curr. Opin. Hematol.* **27**, 206–213 (2020).
39. Hui, B. et al. RREB1-induced upregulation of the lncRNA AGAP2-AS1 regulates the proliferation and migration of pancreatic cancer partly through suppressing ANKRD1 and ANGPTL4. *Cell. Death Dis.* **10**, 207 (2019).
40. Shen, C. J. et al. Oleic acid-induced NOX4 is dependent on ANGPTL4 expression to promote human colorectal cancer metastasis. *Theranostics* **10**, 7083–7099 (2020).
41. Hong, X. et al. UBE2E2 enhances snail-mediated epithelial-mesenchymal transition and Nrf2-mediated antioxidant activity in ovarian cancer. *Cell. Death Dis.* **14**, 100 (2023).
42. Xu, Q. et al. In silico pan-cancer analysis reveals prognostic role of the erythroferrone (ERFE) gene in human malignancies. *Int. J. Mol. Sci.* **24**, 1725 (2023).
43. Bornhauser, B. et al. The hematopoietic stem cell marker VNN2 is associated with chemoresistance in pediatric B-cell precursor ALL. *Blood Adv.* **4**, 4052–4064 (2020).
44. Chen, J. et al. Tumor-associated macrophage (TAM)-derived CCL22 induces FAK addiction in esophageal squamous cell carcinoma (ESCC). *Cell. Mol. Immunol.* **19**, 1054–1066 (2022).
45. Chen, W. et al. Identification of a tumor microenvironment-related gene signature indicative of disease prognosis and treatment response in colon cancer. *Oxid. Med. Cell. Longev.* 6290261 (2021).
46. Dan, W. C. et al. Integrative analyses of radiation-related genes and biomarkers associated with breast cancer. *Eur. Rev. Med. Pharmacol. Sci.* **27**, 256–274 (2023).
47. Maimaiti, A., Xingliang, L. & Shi, L. In vitro and in vivo anti-lung cancer activity of emodin: a RNA-seq transcriptome analysis. *Curr. Med. Chem.* (2022).
48. Srole, D. N., Jung, G., Waring, A. J., Nemeth, E. & Ganz, T. Characterization of erythroferrone structural domains relevant to its iron-regulatory function. *J. Biol. Chem.* **299**, 105374 (2023).
49. Chapoval, S., Dasgupta, P., Dorsey, N. J. & Keegan, A. D. Regulation of the T helper cell type 2 (Th2)/T regulatory cell (Treg) balance by IL-4 and STAT6. *J. Leukoc. Biol.* **87**, 1011–1018 (2010).
50. Liu, J. et al. Immune microenvironment analysis and novel biomarkers of early-stage lung adenocarcinoma evolution. *Front. Oncol.* **13**, 1150098 (2023).
51. Wu, G. et al. Bioinformatics analysis of the clinical significance of HLA class II in breast cancer. *Medicine (Baltim.)* **101**, e31071 (2022).
52. Wright, G. J. et al. Characterization of the CD200 receptor family in mice and humans and their interactions with CD200. *J. Immunol.* **171**, 3034–3046 (2003).
53. Gordon, M. Y. et al. Clinical heterogeneity in chronic myeloid leukaemia reflecting biological diversity in normal persons. *Br. J. Haematol.* **122**, 424–429 (2003).
54. Zhang, H. C. et al. Identification and validation in a novel quantification system of ferroptosis patterns for the prediction of prognosis and immunotherapy response in left- and right-sided colon cancer. *Front. Immunol.* **13**, 855849 (2022).
55. Charoentong, P. et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell. Rep.* **18**, 248–262 (2017).

Acknowledgements

We would like to thank Editage (https://www.editage.cn/) for English language editing.We would like to thank Dr. Salunke for your review of our paper and your valuable comments.

Author contributions

Conceptualization, Jingwen.Guan and Simin.Min; Methodology, Jingwen.Guan; Software, Jingwen.Guan, Zhiguo.Guo; Writing—Original Draft Preparation, Jingwen.Guan; Writing—Review & Editing, Jingwen.Guan, Simin.Min, and Xiaolan.Zhou; Project Administration, Jingwen Guan and Yan Xia. All authors have read and agreed to the published version of the manuscript.

Funding

This study was supported by the Science Foundation of Anhui Medical University for Young Scholars under Grant No. 2021xkj080 (Anhui Medical University 2021 Campus-level Research Project) and the Research Project of Suzhou Science and Technology under Grant No. SZKJXM202319.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-78505-0.

Correspondence and requests for materials should be addressed to J.G.

Reprints and permissions information is available at www.nature.com/reprints.

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