

Research Paper

Ferroptosis-associated molecular classification characterized by distinct tumor microenvironment profiles in colorectal cancer

Wenqin Luo^{1,2†}, Weixing Dai^{1,2†}, Qingguo Li^{1,2†}, Shaobo Mo^{1,2†}, Lingyu Han^{1,2}, Xiuying Xiao³, Ruiqi Gu^{1,2}, Wenqiang Xiang^{1,2}, Li Ye^{1,2}, Renjie Wang^{1,2}, Ye Xu^{1,2}, Sanjun Cai^{1,2}, Guoxiang Cai^{1,2}✉

1. Department of Colorectal Surgery, Fudan University Shanghai Cancer Center, Shanghai, China.
2. Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China.
3. State Key Laboratory of Oncogenes and Related Genes, Shanghai Cancer Institute, Department of Oncology, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China.

† Wenqin Luo, Weixing Dai, Qingguo Li and Shaobo Mo contributed equally to this work.

✉ Corresponding author: Guoxiang Cai, MD, PhD. Department of Colorectal Surgery, Fudan University Shanghai Cancer Center Department of Oncology, Shanghai Medical College, Fudan University, 270 Dong'an Road, Shanghai, 200032, China. E-mail: gxcaifusc@163.com

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). See <http://ivyspring.com/terms> for full terms and conditions.

Received: 2021.12.07; Accepted: 2022.01.20; Published: 2022.02.07

Abstract

Ferroptosis is a non-apoptotic form of cell death recognized in recent years. Nonetheless, the potential role of ferroptosis-associated genes in immune regulation and tumor microenvironment formation remains unknown. In this study, we characterized the ferroptosis-associated patterns of colorectal cancer through integrative analyses of multiple datasets with transcriptomics, genomics, and single-cell transcriptome profiling. Three distinct ferroptosis-associated clusters (FAC1, FAC2 and FAC3) were identified from 1251 CRC bulk samples, which were associated with different clinical outcomes and biological pathways. The TME characterization revealed that the three patterns were highly consistent with known immune profiles: immune-desert (FAC1), immune-inflamed (FAC2) and immune-excluded (FAC3), respectively. Ferroptosis-associated immune and stromal-activated genes were obtained and characterized by corresponding function in CRC tumorigenesis. Further single-cell analyses identified the ferroptosis-associated immune responding tumor cells and ferroptosis-associated stromal cells infiltration pattern. Based on the Fersig score, which was extracted from the ferroptosis phenotype-related signature, patients with lower Fersig score were characterized by prolonged survival time and effective immune responses. Collectively, we uncovered the ferroptosis-associated patterns associated with TME diversity and immune response phenotype. The Fersig we constructed could be the potential therapeutic target genes to improve the efficacy of patients' immunotherapy. The Fersig scoring scheme could enhance the understanding of TME infiltration associated with ferroptosis and prediction of immunotherapy efficacy.

Key words: Colorectal cancer; Ferroptosis; Tumor microenvironment; Immune activation; Stromal cells infiltration; Myofibroblast; Immunotherapy.

Introduction

Colorectal cancer (CRC) is one of the most common malignancies worldwide, which is the main cause of cancer death [1]. Once diagnosed with CRC at an advanced stage, patients would have extremely poor prognosis [2]. Thus, more valuable prognostic markers and potential therapeutic targets are required to help patients benefit from more aggressive therapy.

Ferroptosis is an iron- and lipid reactive oxygen species (ROS)-dependent form of programmed cell death [3, 4], which is involved in multiple diseases, including neurological disorders and diverse cancers. Despite recent advances in cancer therapeutics, some patients remain resistant to therapy because of anti-apoptosis. Thus, the induction of ferroptosis may

be an effective intervention strategy to the treatment of cancers [5-7]. Recently, various molecules have been implicated in ferroptosis during pathogenesis of CRC. RSL3 inhibited GPX4 to induce ferroptosis in CRC [8]. Knockout of SLC7A11 induced ferroptotic cell death to reduce chemotherapy resistance [9]. Chen et al. reported that combination of β -elemene, a ferroptosis inducer, with anti-EGFR treatment could trigger ferroptosis to produce anti-tumor effects in CRC patients with RAS mutation [10, 11]. These studies highlighted the promising strategy of inducing ferroptosis in cancer therapy. However, there is a lack of comprehensive understanding of ferroptosis-associated genes in pathogenesis of CRC. In this study, we aimed to integrate transcriptomics and genomic data for systematical analysis of ferroptosis-associated genes to enhance our understanding of ferroptosis in CRC tumorigenesis.

Increasing evidence indicated that the tumor microenvironment (TME) contexture (including infiltrating immune cells, stromal cells, and secreted cytokines et al) played an important role in the tumorigenesis [12, 13]. In CRC, disease relapse and mortality could be predicted by the enrichment level of lymphocyte populations (cytotoxic CD8⁺ T cells) at the center or margin of tumor [14]. Recently, immunotherapy has been recognized as a revolution in cancer treatments, which can modulate immune responses against tumors. Immune-checkpoint blockade (ICB) therapy, represented by CTLA-4 and PD-1, has achieved a marked durable response in CRC patients [15]. In CRC, deficient mismatch repair or high microsatellite instability (dMMR/MSI-H) is a well-established biomarker for the potential response to ICB immunotherapy [16]. However, only a small proportion of dMMR/MSI-H patients showed positive responses to immunotherapy, thus there is an urgent clinical need to find more precise and reliable biomarkers to identify patients who may benefit from immunotherapy, and monitor their therapeutic responses [17]. Given that there are classes of TME contexture associated with ICB responsiveness, comprehensively understanding the microenvironment infiltration patterns can facilitate the management of ICB therapy.

Recent studies revealed the correlation between TME cells infiltration and ferroptosis. For example, M2-polarized tumor-associated macrophages (TAMs) can act as "iron-donators" to promote cancer progression [18]. Ferroportin (FPN-1), as an iron exporter, can reduce the cellular iron of M2-polarized TAMs, thereby providing iron for tumor cells [19]. Wang's group demonstrated that the treatment of PD-L1 inhibitors could increase lipid ROS in CD45-IDB cells and suppress tumor progression [20].

Further understanding the interactions between TME cells infiltration and ferroptosis can provide new angles for cancer therapy, such as ferroptosis-based immunotherapy [21].

In this study, we integrated TCGA (COREAD), GEO (GSE14333 [22], GSE37892 [23] and GSE39582 [24]), and single-cell RNA (scRNA) sequencing (GSE144735 [25]) databases to evaluate the association between ferroptosis patterns and TME cell-infiltrating characteristics. We identified three distinct patterns using unsupervised clustering, which were closely linked to three previously reported immune phenotypes: immune-inflamed, immune-excluded, and immune-desert [26]. We also identified the concrete genes associated with different TME infiltration phenotypes, quantified the ferroptosis patterns of individual tumors, and predicted their correlation with the response to immunotherapy. These findings will help understanding ferroptosis in shaping diverse TME profiles and provide new angles for the therapeutic intervention in CRC.

Results

Ferroptosis-associated molecular classification in colorectal cancer

The analytical process in this study is illustrated in Figure S1. We investigated the roles of 259 ferroptosis-associated genes derived from FerrDb database [27] in CRC, including 108 drivers, 69 suppressors and 111 markers. We retained totally 228 human genes for analyses (Figure 1A). As expected, gene ontology (GO) enrichment analyses showed that these genes were characterized by the biological processes of ferroptosis and cell death (Figure 1B). Principal component analysis (PCA) (Figure 1C) and hierarchical clustering of gene expression (Figure S2C) based on paired tumor-normal samples showed that ferroptosis-associated genes could distinguish CRC samples from normal samples, suggesting that ferroptosis might play a significant role in regulating CRC tumorigenesis. We next utilized consensus clustering analysis of the non-negative matrix factorization (NMF) algorithm [28] to stratify tumor samples based on the expression of the 228 genes (Figure S2A). Accordingly, we identified three distinct clusters and referred to these clusters as ferroptosis-associated clusters (FAC), including 148 cases in FAC1, 132 cases in FAC2 and 58 cases in FAC3 (Figure 1D-F, Figure S2B, Table S1), among which FAC3 had the worst prognosis in TCGA-COREAD cohort (Figure 1F, overall survival (OS) within 2000 days, $P = 0.00034$; Figure S2B, overall survival (OS) of all samples, $P = 0.0092$, log-rank test). In addition, we combined three GEO datasets with

available clinical data (GSE39582, GSE14333 and GSE37892) into a meta-GEO cohort and obtained similar results of classification and prognosis (Figure

1F, relapse-free survival (RFS) within 2000 days, $P < 0.0001$; Figure S2B, relapse-free survival (RFS) of all samples, $P < 0.0001$, log-rank test).

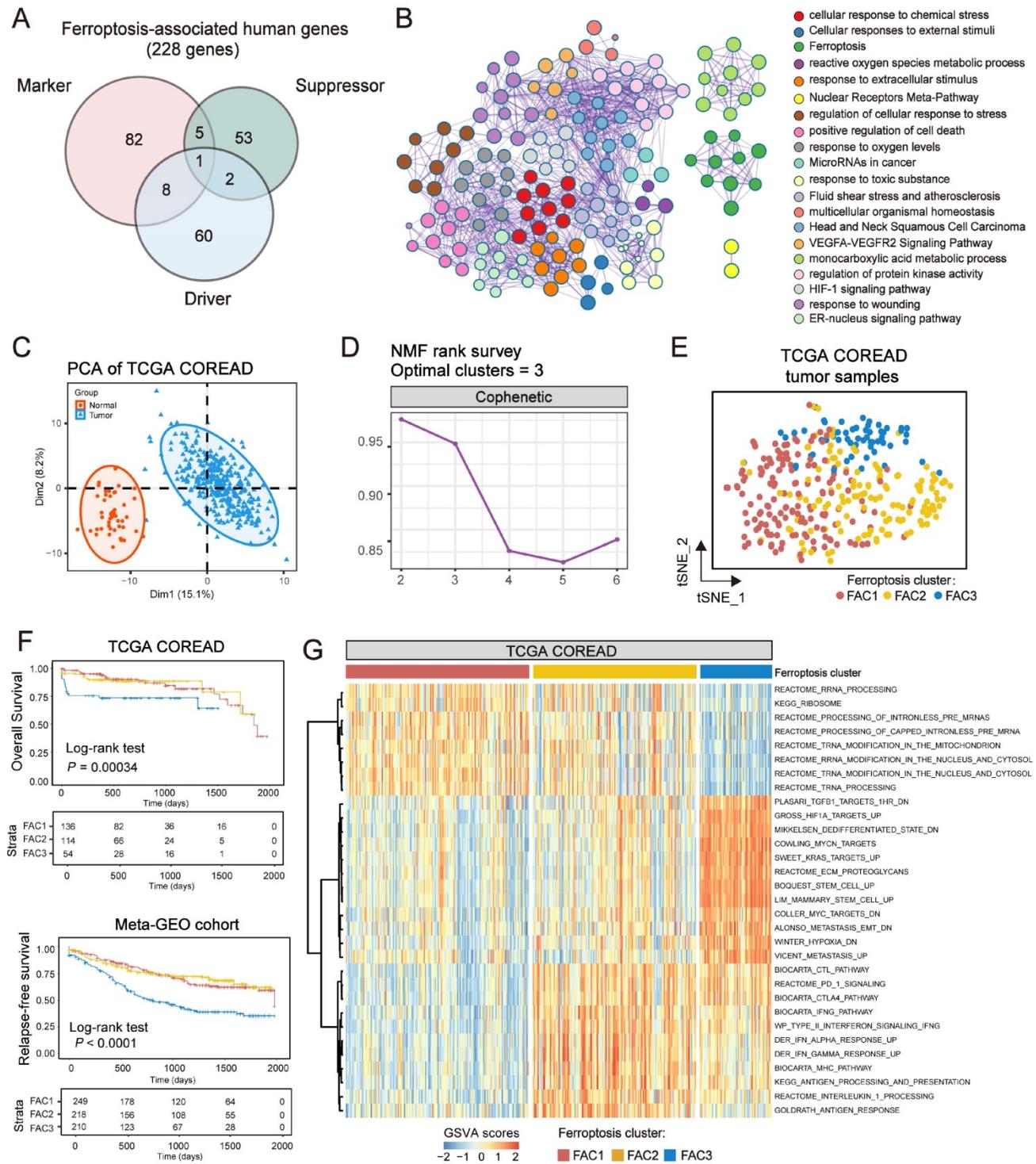


Figure 1. Ferroptosis-associated molecular classification in colorectal cancer. (A) Venn diagram shows three subtypes of ferroptosis-associated human genes, including 71 drivers, 61 suppressors and 96 markers. **(B)** Metascape enrichment network visualization shows the clusters of enriched GO terms of ferroptosis-associated genes. Cluster annotations are shown in the color code. **(C)** Principal component analysis of ferroptosis-associated genes to distinguish tumors from normal patients. All samples: n=382; tumor: n=338; normal: n=44. **(D)** NMF rank survey was shown. The optimal number of clusters: rank=3. **(E)** t-SNE visualization of three NMF clusters in tumors of TCGA-COREAD cohort, n=338. Dots are colored according to clusters. **(F)** Kaplan-Meier curves for overall survival of three NMF clusters in TCGA and for relapse-free survival of three clusters in meta-GEO cohort. P value was determined by the log-rank test. Patients' survival days within 2000 days are shown. **(G)** Pathway activities of three clusters using GSVA analyses.

To understand the biological discrepancy among three distinct clusters, we performed gene set variation analysis (GSVA) [29] on tumor samples of TCGA (Figure 1G). The results showed that FAC1 was enriched in pathways related to RNA modification and processing. Interestingly, FAC2 was remarkably enriched in immune activation and immunotherapy-related pathways, such as antigen response, interferon gamma/alpha response, PD-1 and CTLA4 signaling. However, FAC3 presented enrichment pathways mainly correlated with tumor-specific and stromal pathways such as TGF- β , stem cell and epithelial-mesenchymal transition (EMT), supporting its worse prognosis.

In terms of clinical characteristics (Figure 2A), most patients in FAC3 had advanced stage tumor (M1 and stage IV) (15.52% in FAC-3 patients, $P = 0.0063$, Pearson's chi-square test), only few patients were with early-stage tumor (stage I) (5.17% in FAC3, $P = 0.0087$, Pearson's chi-square test). The lowest proportion of advanced stage tumors (M1 and stage IV) (3.79% in FAC2 patients, $P = 0.0063$, Pearson's chi-square test), but highest proportion of early-stage tumors (stage I) (22.73% in FAC2 patients, $P = 0.0087$, Pearson's chi-square test) were observed in FAC2 samples. Similar results in meta-GEO cohort indicated that the less patients with early-stage CRC (Stage 0-I) were in FAC1 and FAC2, the more patients with advanced stage (Stage IV) would be in FAC3 (Figure 2C).

Biological molecular changes underlying different clusters in CRC

To further investigate molecular changes underlying these clusters, we utilized a consensus molecular subtypes (CMS) classifier, referred to as CMScaller[30], to stratify TCGA tumor samples. We found that nearly all patients in FAC3 were enriched in CMS4 subtype, while FAC2 contained most of the CMS1 patients (Figure 2B). For microsatellite instability (MSI) status, patients in FAC2 encompassed the majority of MSI-H tumors, reminiscent of CMS1 characteristics [31]. We next applied single-sample gene set enrichment (ssGSEA) analysis [29] using signatures extracted from CMScaller package (Figure 2B and Figure S3A-B). The results showed that FAC3 presented high enrichment of gene signatures associated with CMS4-mesenchymal, such as LGR5⁺ stem cells, TGF- β signaling, and EMT process. Moreover, we found that CMS3-metabolic characteristics such as fatty acids and glycolysis signature were enriched in FAC2, conforming to the biological features of ferroptosis [32]. Patients in FAC1 were mainly enriched in CMS2-canonical characteristics, marked by WNT

signaling activation. We next turned our attention to GSE39582 dataset with six molecular subtypes (CIN_{Immune-Down}, CIN_{WntUp}, CIN_{normal}, CSC, dMMR, and KRASm) [24]. In this dataset, dMMR tumor was associated with immune activation. CIN_{WntUp} tumor was associated with upregulation of WNT signaling. CIN_{normal} and CSC were associated with EMT, stem cells and stromal pathways. Consistent with previous findings, patients of CSC and CIN_{normal} subtype were mainly clustered into FAC3. CIN_{WntUp} and CIN_{Immune-Down} subtypes were predominantly clustered into FAC1. dMMR subtype was concentrated within FAC2 patients (Figure 2D). All these findings indicated that marked differences in the intrinsic biological underpinnings of ferroptosis-associated classification in CRC.

Given that our classification was based on the expression of ferroptosis-associated genes, we then calculated the ssGSEA score using the gene set with 228 genes, which we named as ferroptosis score. This analysis revealed that ferroptosis score was significantly upregulated in FAC2 and FAC3 (Figure 2B and 2E). Moreover, we found ferroptosis had positive correlation with both CMS1 (MSI signature included) and CMS4-characteristics (TGF- β signaling and EMT signature included) (Figure 2F and Figure S3C). Therefore, we postulated that ferroptosis-associated genes may regulate CRC tumorigenesis mainly through immune activation and stromal pathways activation.

Single-cell transcriptome profiling of CRC cells based on ferroptosis-associated molecular classification

Since bulk-tissue contains signatures of distinct cell populations, we then analyzed single-cell RNA (scRNA) dataset (GSE144735) [25] to investigate whether our classification could be precisely reflected in CRC tumor single cells. The dataset, Lee *et al.* (GSE144735), contained the SMC and KUL3 cohorts with scRNA expression data. We firstly isolated 17,469 tumor epithelial cells from the SMC cohort, and then scored them using differentially expressed genes (DEGs) of FAC1-3 bulks (Figure 3A). We found that CRC cells could only be assigned into FAC1 and FAC2-like populations, of which FAC2-like tumor cells showed upregulated inflammatory signature (Figure 3A), marked by high expression of chemokines and interferon-related genes as previously described [33]. For CMS classification (Figure 3B), FAC1-like tumor cells were mainly enriched in CMS2 subtype, while FAC2-like tumor cells were predominantly enriched in CMS1 subtype through both bulk and single-cell level classification. For mutation type, FAC2-like tumor cells were mainly

clustered into MSI-H and BRAF mutation groups, which has also been reported as CMS1 characteristics [34]. In contrast, APC mutation, referred to as characteristics of CIN subtype (CIN_{WntUp}) in GSE39582 dataset, was mainly enriched in FAC1-like tumor cells.

Considering the immune activation in FAC2 bulk samples, we checked the expression of MHC-I (HLA-A and HLA-B) and MHC-II (HLA-DRA)

molecules, and found their upregulation in FAC2-like tumor cells (Figure 3C). These results indicated that FAC2-like cells displayed immune response [35], as confirmed by gene ontology (GO) analysis (Figure 3D). In contrast, FAC1-like displayed enrichment of RNA processing (Figure 3D), in line with the description in Figure 1G. Finally, examination of ferroptosis-associated gene score (Figure 3C) showed that ferroptosis had positive correlation with immune

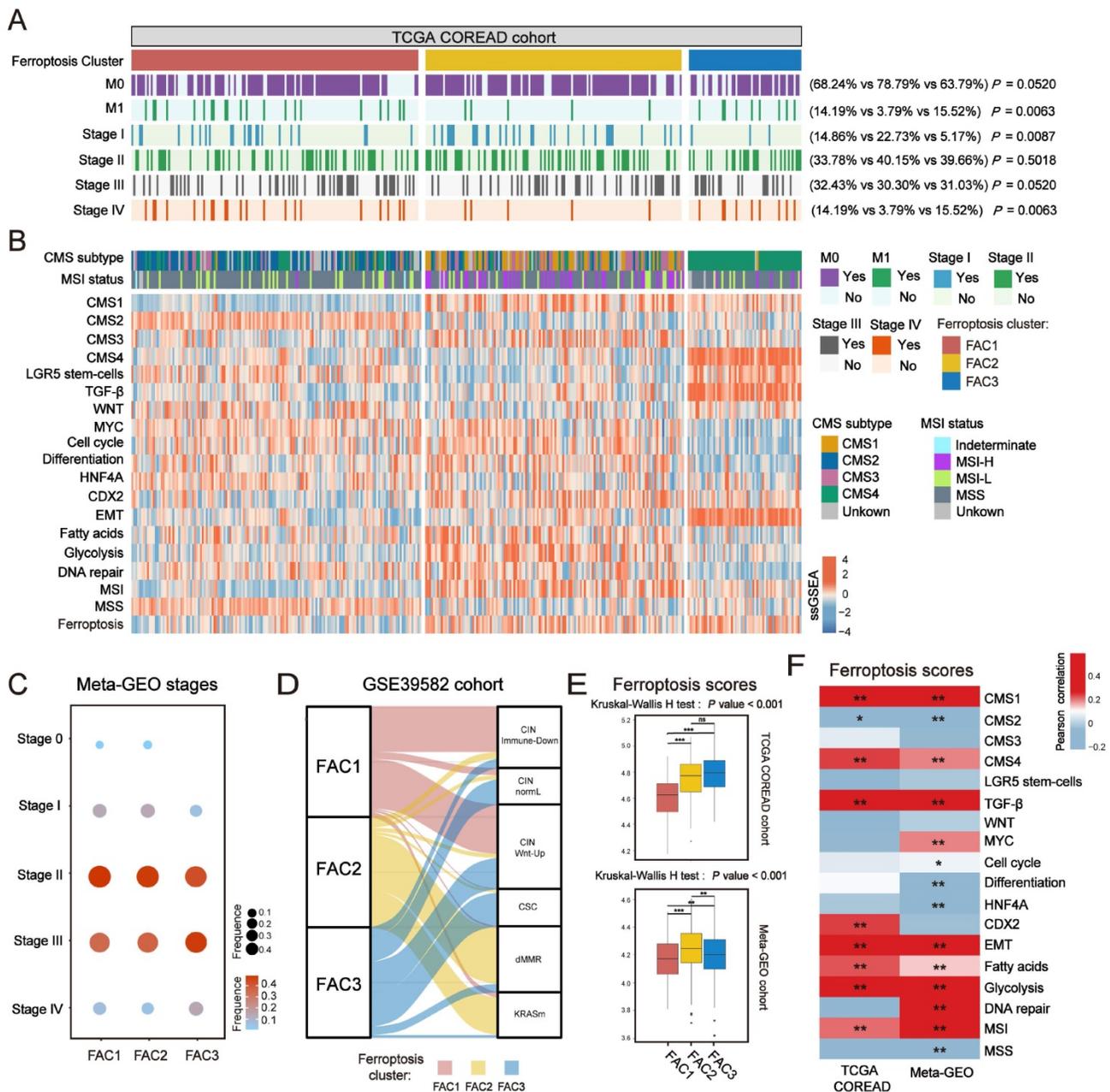


Figure 2. Clinical characteristics and biological molecular changes underlying three clusters in CRC. (A) Heatmap shows clinical characteristics of NMF clusters in TCGA cohort. P value was determined by Pearson's chi-square test. (B) Heatmap shows molecular characteristics of NMF clusters in TCGA cohort. (C) Dot plot shows the proportions of different stages among NMF clusters of meta-GEO cohort. Both color and size represent the fraction of stages in different clusters. (D) Alluvial diagram of clusters in groups with different molecular subtypes (CIN_{Immune-Down}, CIN_{WntUp}, CIN_{normL}, CSC, dMMR, and KRAS_m). (E) Barplot shows ssGSEA score of ferroptosis-associated genes in the NMF clusters. The statistical difference of three clusters was compared by the Kruskal-Wallis H test. *P < 0.05; **P < 0.01; ***P < 0.001. The difference of two clusters was compared by the wilcox test. *P < 0.05; **P < 0.01; ***P < 0.001. (F) Correlations between ferroptosis-associated ssGSEA score and biological gene signatures correlated with CMS subtypes using Pearson analysis.

responding cells. Similar results were obtained in the KUL3 cohort (Figure S4A-D). Overall, FAC1 and FAC2-like tumor cells presented similar molecular

patterns with previous analyses of TCGA and GEO cohort. Notably, FAC3-like tumor single cells were not detected in SMC and KUL3 cohorts.

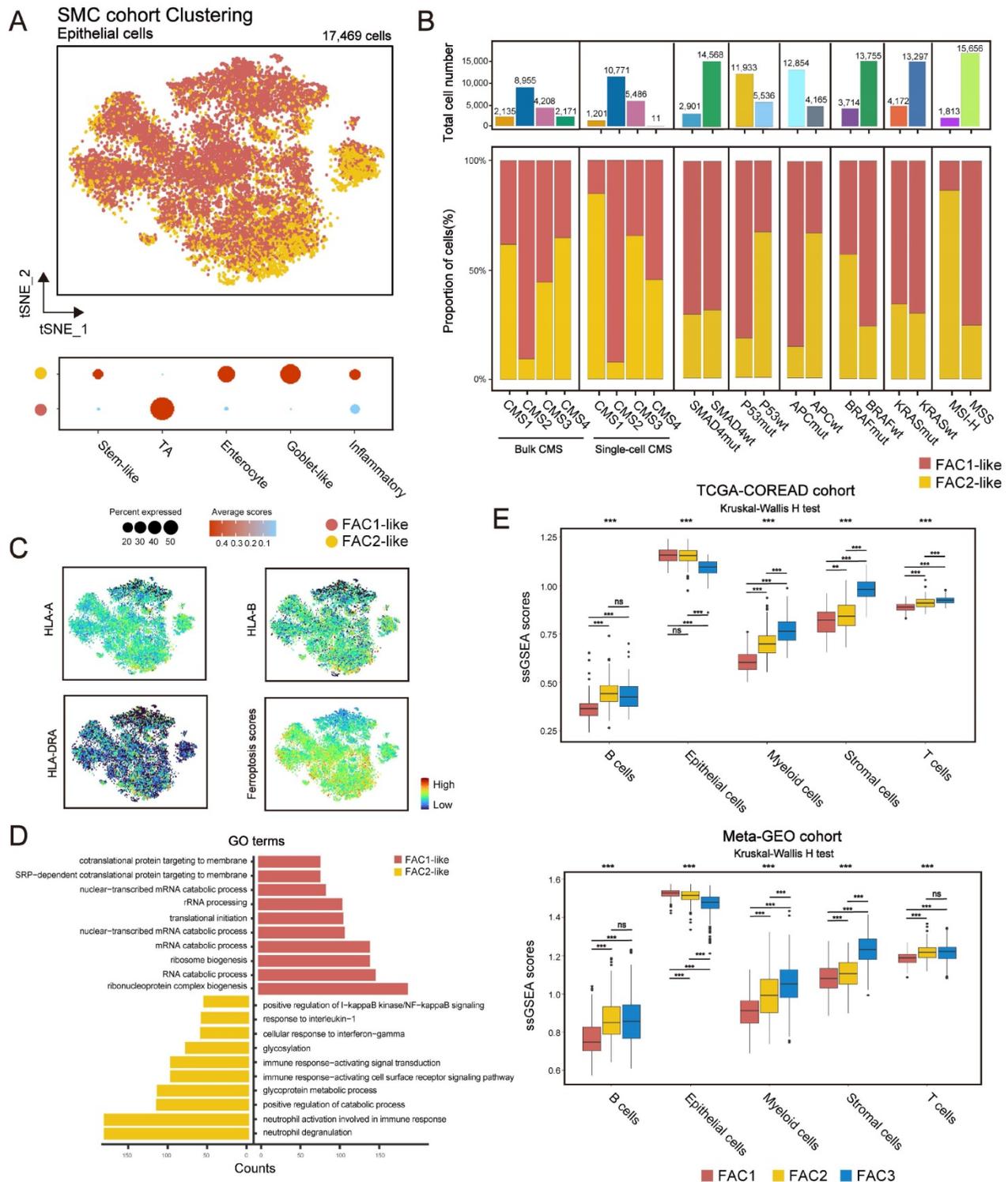


Figure 3. Single-cell transcriptome profiling of CRC cells based on ferroptosis-associated molecular classification. (A) t-SNE visualization of 17,469 tumor epithelial cells from SMC single-cell cohort (top panel). Cells are colored according to clusters. Dot plot for the score of gene signatures associated with CRC cellular phenotype and responses to therapy in each cell type. Color represents the mean score in each cell cluster, and size indicates the fraction of cells expressing gene score (bottom panel). **(B)** Barplot shows the proportion of different molecular characteristics in CRC tumor cells. Bars are colored according to clusters. **(C)** t-SNE visualization of MHC-I (HLA-A, HLA-B), MHC-II (HLA-DRA) molecular expression and the score of ferroptosis-associated genes. **(D)** GO analysis of differential expressed genes between FAC1-like and FAC2-like CRC tumor cells. **(E)** ssGSEA score of 5 cell types (epithelial cells, T cells, B cells, Myeloid cells and stromal cells) among three clusters in TCGA and meta-GEO cohorts. The statistical difference of three clusters was compared by the Kruskal-Wallis H test. *P < 0.05; **P < 0.01; ***P < 0.001. The difference of two clusters was compared by the wilcox test. *P < 0.05; **P < 0.01; ***P < 0.001.

Lee *et al.* has indicated that single-cell representation of CMS classification included few CMS4-like tumor cells [25]. Just as shown in Figure 3B, although CMS4 bulk subtype was detected, few CMS4-like tumor cells were presented at single-cell level, even for those from CMS4 bulks. Lee *et al.* verified that remarkable stromal cells infiltration led to scarce CMS4-like tumor cells in CMS4 tissues. Based on this finding, we estimated the influences of single-cell types on our bulk classification (Figure 3E). Here we demonstrated that stromal cell types (fibroblasts and endothelial cells included) were highly enriched in FAC3 tumors, while epithelial cells were enriched in FAC1-2 tumors. Subsequent ESTIMATE algorithm [36] confirmed low tumor purity and remarkable stromal infiltration in FAC3 tumors (Figure S4E). Therefore, FAC3, marked by remarkable stromal infiltration, couldn't be reflected in tumor cells at single cell level.

Distinct tumor microenvironment infiltration in three tumor clusters

In Figure 3E, tumor microenvironment (TME) infiltration displayed significant differences among three ferroptosis-associated clusters. To further characterize their microenvironment heterogeneity, we performed ssGSEA and CIBERSORT analyses using 31 cell types detected in single-cell research [25] (Figure 4A-B; Figure S5A). For immune infiltration, ssGSEA score of 14 cell subtypes was upregulated in both FAC2 and FAC3 (Figure 4A and Figure S5A). We next examined the functional status of CD8⁺ T cells (Figure S5B) and found that FAC2 had the highest CD8⁺ T cytotoxicity score, indicating greater functionality [37]. We also evaluated immune cell infiltration using cell types from the study of Charoentong [38] (Figure S5C) and found that FAC2 was rich in adaptive immune cells such as activated CD4⁺ and CD8⁺ T cells, while FAC3 was enriched in innate immune cells, such as natural killer cells, MDSC and regulatory T cells, recognized as the major components of immune suppressive TME [39]. Nearly all stromal cell types were mainly enriched in FAC3 (Figure 4A and S5A), among which myofibroblasts were the most abundant (Figure 4B). Subsequent survival analysis showed that myofibroblasts could predict a poor prognosis (Figure 4C; overall survival (OS) in TCGA, $P = 0.028$; relapse-free survival (RFS) in meta-GEO, $P < 0.0001$; log-rank test). Previous findings in Figure 2B showed that FAC3 tumors displayed stromal pathways activation such as TGF- β and EMT signaling. Given that stromal myofibroblasts can suppress immune microenvironment by TGF- β production [25], and TGF- β can promote angiogenesis, metastasis, tumor cell EMT and

fibroblast activation [40], we further evaluated the stromal-related signature enrichment using gene sets (angiogenesis, EMT2, EMT3 and pan-fibroblast TGF- β included) derived from Mariathasan *et al.* [40], and then confirmed their activation in FAC3 tumors (Figure S5D). Moreover, previous studies indicated that immune cells could be retained in the stroma surrounding tumor cell nests rather than penetrating their parenchyma, so that the antitumor immunity could be rendered ineffective [26]. Therefore, although FAC3 tumors had an abundance of immune cells, FAC3 tumors were recognized as immune-excluded phenotype [41, 42].

Previous studies have indicated that MSI/dMMR status and tumor mutation burden (TMB) level could predict the response to immunotherapy [43]. We next checked them in our classification and identified their high level (MSI-H/dMMR and TMB-H) in FAC2 (Figure 4D). Given that PD-L1 expression could predict the response to anti-PD-1/L1 treatment [44], we compared the PD-L1 expression among the three clusters, and then observed the higher expression of PDL-1 in FAC2 and FAC3 (Figure 4E). Based on previous findings, FAC1 subtype was recognized as immune-deserted characterized by a paucity of immune and stromal cells. FAC2 subtype was recognized as immune-inflamed characterized by adaptive immune cell infiltration and immune activation, which predicted a positive response to immunotherapy, such as anti-PD-L1/PD-1 therapy. However, FAC3 subtype was recognized as immune-excluded characterized by innate immune cell infiltration and remarkable stromal cell infiltration. Therefore, FAC3 subtype, though with upregulated PD-L1 expression, might show ineffective response to anti-PD-L1/PD-1 therapy (Figure 4E). Correlation between detailed cell types and ferroptosis-associated gene score demonstrated the strong influences of ferroptosis on CRC stromal and immune infiltration (Figure 4F).

Identifying specific ferroptosis-associated genes correlated with immune activation

We assessed the transcriptional signature of ferroptosis correlated with immune activation. By overlapping ferroptosis-associated genes with DEGs upregulated in FAC2 (Figure 5A), we obtained 35 genes related to immune activation, such as T cells activation and interferon gamma response (Figure 5B), and then recognized them as immune-activated Fersig. All these genes displayed positive correlation (Figure 5C), implying their similar functional category.

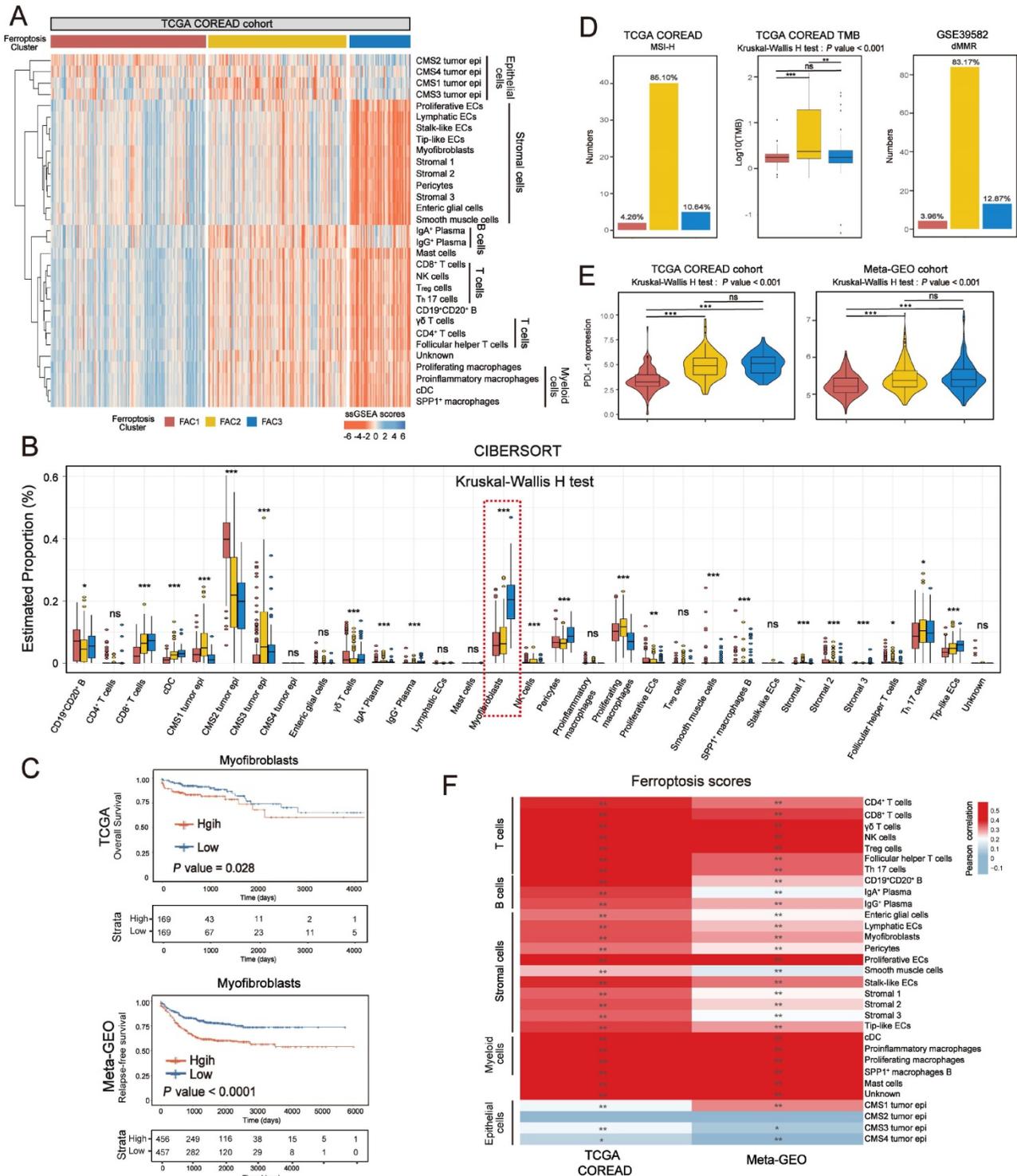


Figure 4. Distinct tumor microenvironment infiltration in three tumor clusters. (A) Heatmap shows the ssGSEA score of 31 cell subtypes in three ferroptosis-associated clusters. **(B)** Relative abundance of 31 cell subtypes of three clusters in TCGA cohort. The statistical difference of three clusters was compared by the Kruskal-Wallis H test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. **(C)** Kaplan-Meier curves for overall survival of TCGA and for relapse-free survival of meta-GEO cohort. The high and low groups were divided by the median value of the ssGSEA score of DEGs of myfibroblasts. P value was determined by the log-rank test. **(D)** Proportion of patients with MSI-H in the NMF clusters of TCGA cohort. TMB level of the NMF clusters in TCGA cohort. Proportion of patients with dMMR in the NMF clusters of meta-GEO cohort. The statistical difference of three clusters was compared by the Kruskal-Wallis H test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. **(E)** Expression of PD-L1 among three clusters of TCGA and meta-GEO cohorts. The statistical difference of three clusters was compared by the Kruskal-Wallis H test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. The difference of two clusters was compared by the wilcox test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. **(F)** Correlations between ssGSEA score of ferroptosis-associated and DEGs of 31 cell subtypes of CRC using Pearson analysis.

Analysis of copy number variation (CNV) showed a widespread CNV alteration (Figure S6A)

among them. Therefore, we investigated their influence on clinical and molecular patterns using

ssGSEA score. We found that patients with high immune-activated Fersig expression displayed a paucity of advanced stage (stage IV) (30.56% vs 69.44% in TCGA; 40% vs 60% in meta-GEO) (Figure 5D). Moreover, patients with high immune-activated Fersig expression presented immune-activated characteristics, such as remarkable enrichment of CMS1 (96.08% in TCGA), MSI-H (87.23% in TCGA), dMMR subtypes (88.12% in GSE39582) (Figure 5E), and significant immune cells infiltration (Figure 5F). We observed that stromal cell infiltration and stromal pathways were downregulated in high immune-activated Fersig score group (Figure 5F). Somatic mutation analysis showed that high immune-activated Fersig score group was associated with higher mutation rates of common mutant genes in CRC, such as TTN (54% vs 38%), MUC16 (37% vs 18%) and PIK3CA (32% vs 19%) (Figure 5H). Given that the high immune-activated Fersig score group displayed a higher level of PD-L1 expression and TMB quantification (Figure 5G), we postulated that immune-activated Fersig presented a predictive advantage in response to immunotherapy.

To find the potential genes in regulating immune activation for future study, we performed univariate Cox analyses of immune-activated Fersig on both TCGA (based on overall survival (OS)) and meta-GEO (based on relapse-free survival (RFS)) cohort (Figure S6B). The results showed that 11 genes were associated with better prognosis, among which DUOX2, XBP1, EIF2S1, SOCS1, DRD5, NOS2 and PSAT1 have been reported to be correlated with the promotion of ferroptosis, whereas RRM2, TFRC, CISD2 and GCH1 have been reported to be associated with the inhibition of ferroptosis. These 11 genes were highly expressed in FAC2 group, confirming their main role in immune activation (Figure S6C). Subsequent correlation analysis showed that XBP1, CISD2, EIF2S1, RRM2, GCH1, SOCS1, NOS2 and PSAT1 displayed positive correlation with CD8⁺ T cytotoxicity score and negative correlation with myofibroblasts score (Figure S6D), suggesting these genes could trigger immune activation. The correlation between their regulation of ferroptosis and immune activation in CRC need to be explored in future studies.

Identifying specific ferroptosis-associated genes correlated with stromal activation

Identical analyses were performed to obtain a stromal-activated Fersig, including 49 genes, characterized by stromal processes enrichment such as endothelial cell proliferation (Figure 6A-C). Prevalent CNV alteration also indicated their influence on CRC progression (Figure S7A). Contrary

to immune-activated Fersig, patients with high stromal-activated Fersig expression were enriched in more advanced stage (stage IV) (61.11% vs 38.89% in TCGA; 58.33% vs 38.89% in meta-GEO) (Figure 6D). For molecular characteristics, high stromal-activated Fersig score group was prominently enriched in stromal subtypes, including CMS4 (90.91% in TCGA) and CSC (94.92% in GSE39582). We also observed the immune cell infiltration and higher expression of PD-L1 in high stromal-activated Fersig score group (Figure 6E-G). However, high stromal-activated Fersig score group displayed remarkable stromal cell infiltration, EMT and TGF- β pathways activation, and showed no significant differences in somatic mutation and TMB (Figure 6F-H). Therefore, stromal-activated Fersig may promote CRC progression and predict an ineffective response to immunotherapy.

Next, univariate Cox analyses of stromal-activated Fersig on both TCGA (based on overall survival (OS)) and meta-GEO cohort (based on relapse-free survival (RFS)) obtained 16 genes associated with significant worse prognosis. Among them, NOX4, ALOX15, CDO1, ATF3, TXNIP, PLIN4, HIC1 and ATP6V1G2 have been reported to be correlated with the promotion of ferroptosis, whereas DUSP1, RGS4, SLC2A3, SLC2A6, SLC2A14, AKR1C2, HSPB1 and ZFP36 have been reported to be associated with the inhibition of ferroptosis (Figure S7B). These 16 genes were highly expressed in FAC3 group, confirming their main role in stromal infiltration (Figure S7C). Subsequent correlation analysis showed that PLIN4, RGS4, SLC2A3, SLC2A14, HIC1, DUSP1, CDO1, NOX4 and SLC2A6 displayed negative correlation with CD8⁺ T cytotoxicity score and positive correlation with myofibroblasts as well as EMT gene score (Figure S7D), suggesting these genes could contribute to the suppressive tumor microenvironment. The correlation between their regulation of ferroptosis and the suppressive tumor microenvironment in CRC need to be explored in future studies.

Examining expression of immune-activated and stromal-activated Fersig at single-cell level

We explored the expression level of immune-activated and stromal-activated Fersig at single-cell level (Figure 7A and S8A). The results showed that immune-activated ferroptosis-associated genes were mainly expressed in tumor cells and T cells (Figure 7B-7C and Figure S8B-S8C). However, stromal-activated ferroptosis-associated genes were predominantly expressed in stromal cells (Figure 7B-7C and Figure S8B-S8C). Consistent with the previous mapping FAC1-3 subtypes on CRC tumor cells, the expression of stromal-activated genes in

CRC tumor epithelial cells were lower than the immune-activated genes. These analyses further confirmed the definition of immune-activated and stromal-activated genes.

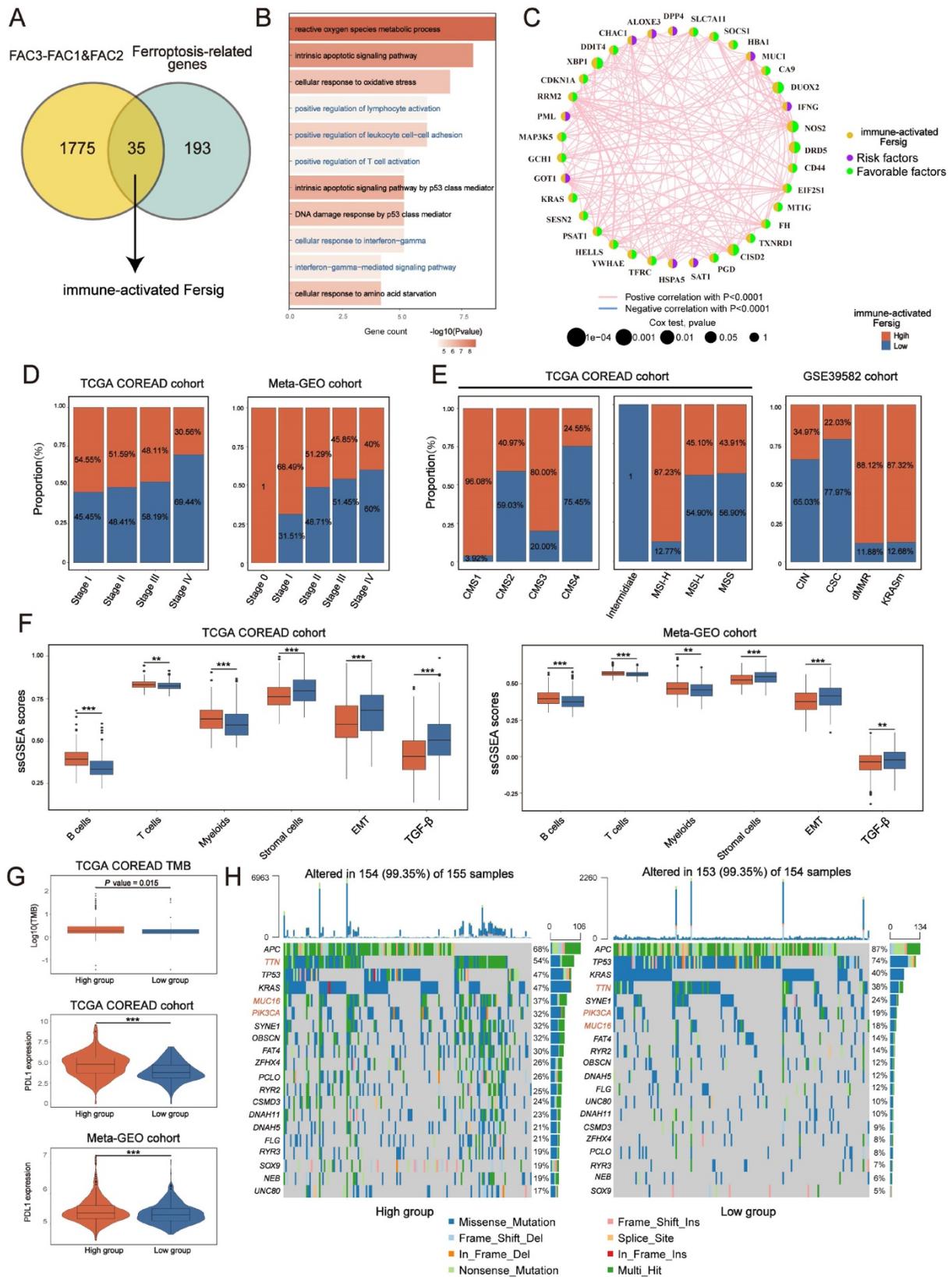


Figure 5. Identifying specific ferroptosis-associated genes correlated with immune activation. (A) Venn diagram shows shared genes between DEGs upregulated in FAC2 and ferroptosis-associated genes, which obtained 35 genes termed immune-activated Fersig. (B) GO analysis of immune-activated Fersig. (C) Correlation between 35 genes. The size of each gene represents survival impact (log-rank test P values indicated). Favorable factors for overall survival are indicated in green, and risk factors are indicated

in purple. The thickness of the line represents the strength of correlation estimated by Pearson correlation analysis. Positive correlation is indicated in pink and negative correlation in blue. **(D)** The high and low groups were divided by the median value of the ssGSEA score of immune-activated Fersig. Barplots show the proportion of clinical characteristics between high and low groups of immune-activated Fersig. **(E)** Barplots show the proportion of molecular characteristics between high and low groups of immune-activated Fersig. **(F)** ssGSEA score of signatures of TME cell types, EMT and TGF- β between high and low groups in TCGA and meta-GEO cohorts. The difference of two clusters was compared through the wilcox test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. **(G)** Comparison of TMB and PD-L1 expression between high and low groups. The difference of two clusters was compared through the wilcox test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. **(H)** Oncoplots show genetic alterations of common mutant genes in CRC between high and low groups. The number on the right indicates the mutation frequency in each regulator. Each column represents individual patient. High group: $n=155$; low group: $n=154$.

Ferroptosis phenotype-related DEGs in colorectal cancer

To further confirm the underlying molecular and clinical patterns determined by ferroptosis-associated genes, we overlapped 854 DEGs among the three ferroptosis-associated clusters and recognized them as ferroptosis phenotype-related signature (Figure S9A). We next performed unsupervised consensus clustering analyses based on a univariate Cox regression analysis of the 854 DEGs (Method section). Both in TCGA and meta-GEO cohorts, this analysis divided patients into three ferroptosis phenotype-related signature groups that had different clinicopathologic subgroups, which were defined as gene-cluster A, B and C (Figure 8A; Figure S9B and Figure S10A; Table S5). By hierarchical clustering and GO analysis (Figure 8A and Figure S9C), the 854 DEGs were further grouped into signature genes A, B and C (Table S5). Genes A were associated with stromal biological processes such as muscle system process and TGF- β signaling. Genes B were associated with immune cells activation. Genes C were enriched in catabolic process and RNA processing. We observed that gene-cluster A presented worst prognosis (Figure 8B; overall survival (OS) in TCGA, $P = 0.0086$; Figure S10A, relapse-free survival (RFS) in meta-GEO, $P = 0.03$, log-rank test) with significant enrichment of advanced stage patients (M1 and stage IV) (16.88% in gene-cluster A, $P = 0.0122$, Pearson's chi-square test) (Figure 8A). We also found that gene-cluster A contained most of the previous FAC3 tumors. Moreover, it was also characterized by CMS4 subtype enrichment (Figure 8A), remarkable stromal cells infiltration (especially myofibroblasts infiltration) (Figure S9E) and stromal-related pathways activation (TGF- β , EMT and angiogenesis included) (Figure 8C, S9D and S10B). Gene-cluster B had most of the FAC2 tumors, characterized by CMS1 subtype enrichment, strong immune cells infiltration (Figure 8A, 8C, S9E and S10B) and especially highest enrichment of cytotoxic CD8⁺ T cells (Figure 8D and S10C). Gene-cluster C possessed most of the FAC1 tumors, and displayed the paucity of TME cells (Figure 8A, 8C, S9E and S10B). Subsequent analysis showed that our immune-activated Fersig tended to be upregulated in gene-cluster B, while stromal-activated Fersig tended to be upregulated in gene-cluster A (Figure 8E and S10D). This result was

consistent with previous findings that ferroptosis played a main role in both immune and stromal infiltration. Taken together, based on ferroptosis-associated genes regulation, CRC could be divided into three stable distinct phenotypes, including immune-deserted, immune-inflamed, and immune-excluded phenotypes.

Further construction of Fersig score and exploration of its role in predicting immunotherapeutic benefits

We have characterized the role of ferroptosis-associated genes in prognosis, immune and stromal infiltration modulation based on patient population. After that, we developed a scoring scheme termed Fersig score (Methods section), which is based on ferroptosis phenotype-related signature, to quantify ferroptosis regulation in individual CRC patient. Since the regulation of ferroptosis focused on immune and stromal infiltration, we calculated the Fersig score based on immune-related genes B and stromal-related genes A in Figure 8A. In TCGA and GEO cohort, we observed that the high Fersig score group had the worse prognosis (Figure 9A and S11A; overall survival (OS) in TCGA, $P = 0.0064$; relapse-free survival (RFS) in meta-GEO, $P = 0.0021$; log-rank test). For molecular characteristics, we found that nearly all CMS4 (98.19% in TCGA) and CSC subtype (94.92% in GSE39582) were clustered into the high Fersig score group (Figure 9B and S11B). Subsequent TME infiltration analysis showed that although the high Fersig score group displayed immune cell infiltration (Figure 9C and S11C, S11E), it presented stromal-related signaling activation and remarkable stromal cell infiltration (Figure 9C, S11C and S11E), especially myofibroblasts infiltration (Figure S11E). We then demonstrated that the high Fersig score group upregulated the expression of stromal-activated Fersig and downregulated the expression of immune-activated Fersig (Figure 9D and S11D), indicating the high Fersig score group was like an immune-excluded phenotype. Finally, we turned our attention to the somatic mutation patterns and found that the high Fersig score group had no advantage in gene mutation and downregulated the TMB (Figure 9E). Moreover, high Fersig score group was enriched in more death (Figure 9F), and advanced stage (M1 and stage IV) of CRC patients (Figure 9F and S11B)), consistent with the worse prognosis in the high Fersig

score group. Therefore, although PD-L1 was upregulated (Figure 9E), the high Fersig score group

predicted an ineffective response to immunotherapy due to its remarkable stromal cell infiltration.

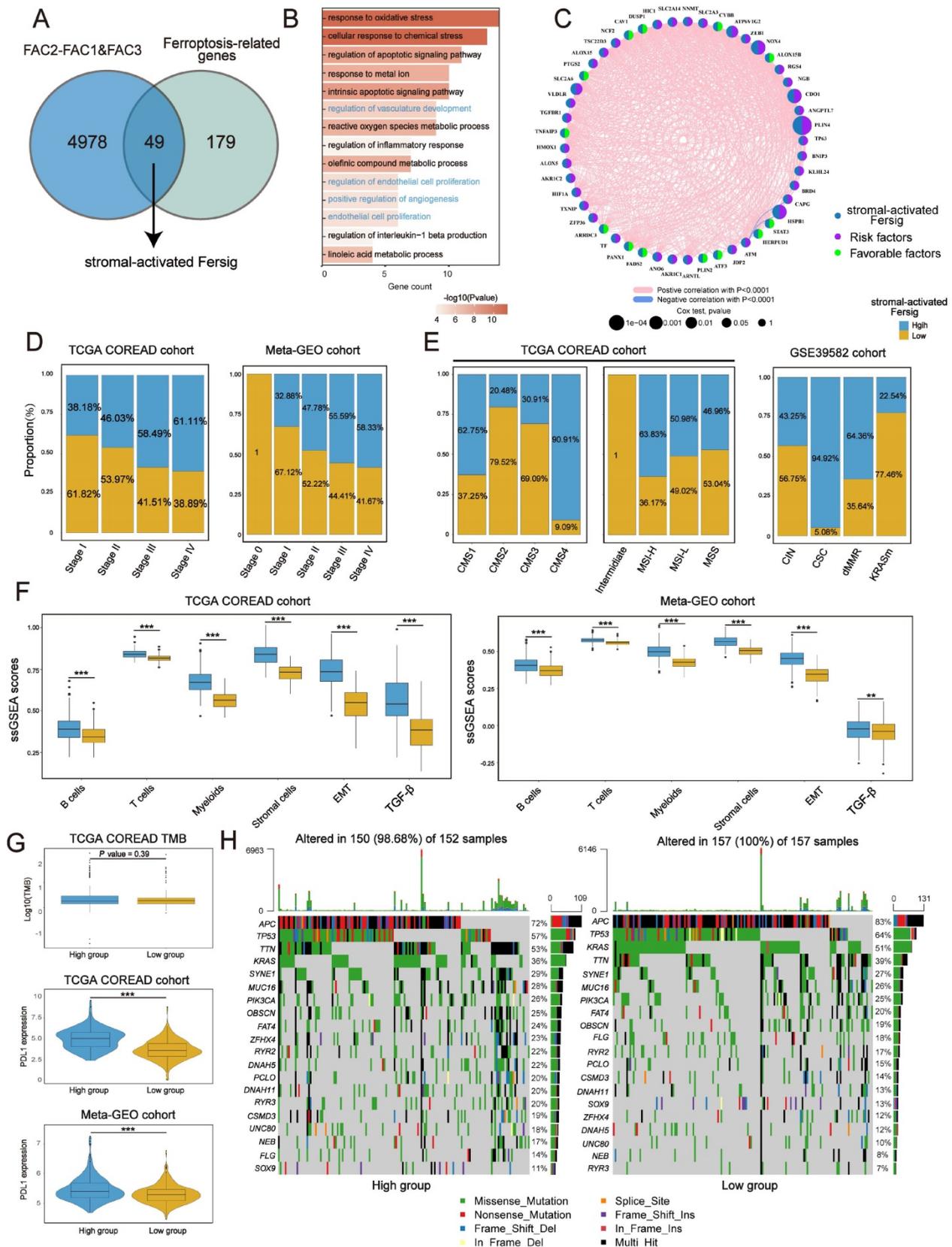


Figure 6. Identifying specific ferroptosis-associated genes correlated with stromal activation. (A) Venn diagram shows shared genes between DEGs upregulated in FAC3 and ferroptosis-associated genes, which obtained 49 genes termed stromal-activated Fersig. (B) GO analysis of stromal-activated Fersig. (C) Correlation between 49

genes. The size of each gene represents survival impact (log-rank test P values indicated). Favorable factors for overall survival are indicated in green, and risk factors are indicated in purple. The thickness of the line represents the strength of correlation estimated by Pearson correlation analysis. Positive correlation is indicated in pink and negative correlation in blue. **(D)** The high and low groups were divided by the median value of the ssGSEA score of stromal-activated Fersig. Barplots show the proportion of clinical characteristics between high and low groups of stromal-activated Fersig. **(E)** Barplots show the proportion of molecular characteristics between high and low groups of stromal-activated Fersig. **(F)** ssGSEA score of signatures of TME cell types, EMT and TGF- β between high and low groups in TCGA and meta-GEO cohort. The difference of two clusters was compared through the wilcox test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. **(G)** Comparison of TMB and PD-L1 expression between high and low groups. The difference of two clusters was compared by the wilcox test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. **(H)** Oncoplots show genetic alterations of common mutant genes in CRC between high and low groups. The number on the right indicates the mutation frequency in each regulator. Each column represents individual patient. High group: $n=155$; low group: $n=154$.

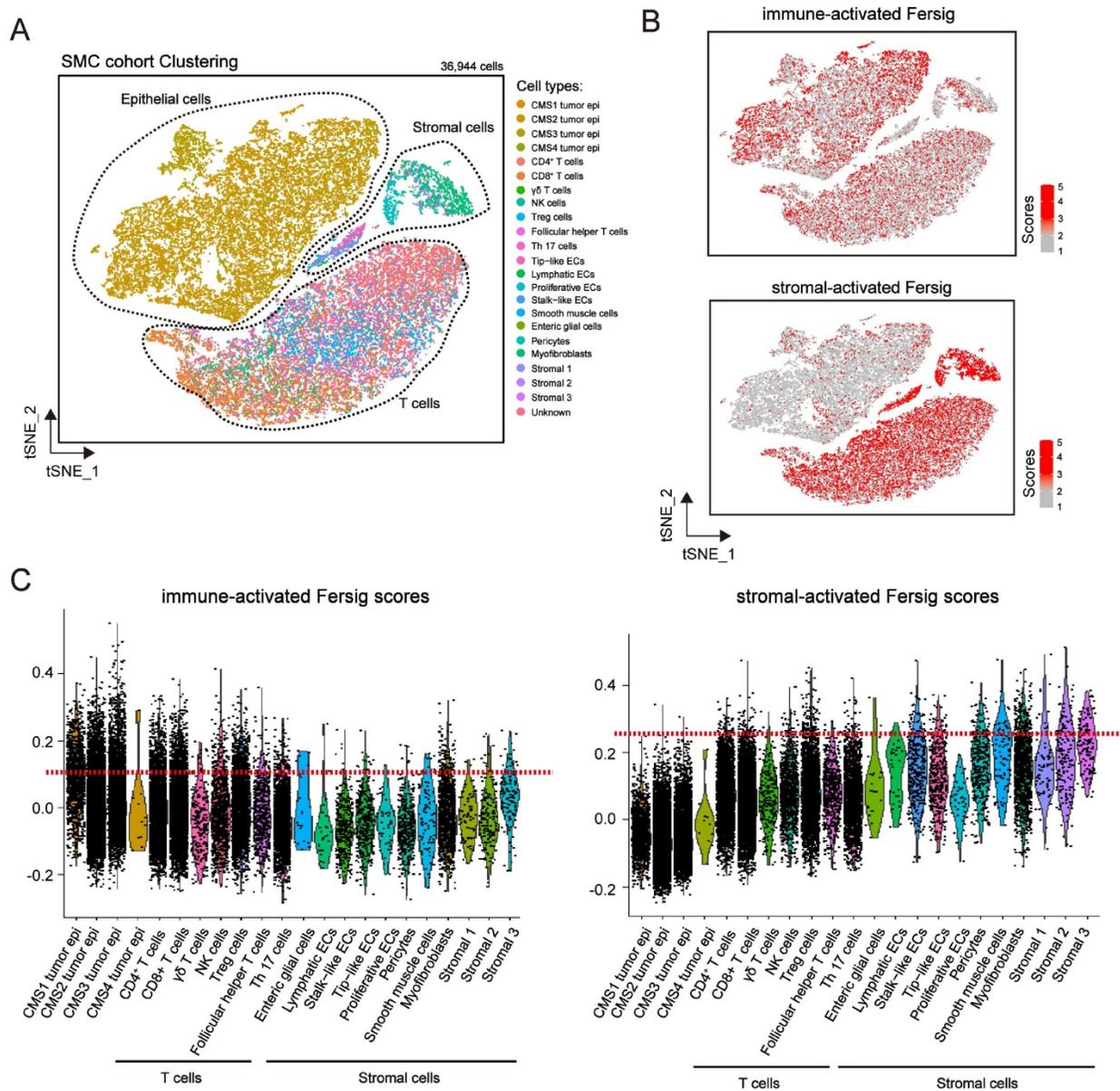


Figure 7. Examining expression of immune-activated and stromal-activated Fersig at single-cell level. (A) t-SNE visualization of 36,944 single cells of CRC. Cells are colored according to cell types. **(B)** t-SNE visualization of score of immune-activated Fersig and stromal-activated Fersig. **(C)** Violin plots shows the score of immune-activated Fersig and stromal-activated Fersig among different cell subtypes.

Immunotherapies represented by PD-1/CTLA-4 inhibitors have been widely used in cancer therapy [45]. We next investigated whether our Fersig scoring algorithm could predict effective response to these two immune checkpoint blockade therapies. In both anti-PD-1 cohort (Liu *et al.* study) [46] and anti-CTLA-4 cohort (Vanallen *et al.* study) [47], patients with low Fersig score exhibited significantly better prognosis (Figure 10A, $P = 0.0044$; Figure 10B, P

$= 0.016$; log-rank test). Furthermore, patients with low Fersig score displayed more effective clinical response to anti-PD-1/CTLA-4 immunotherapies, compared to those with high Fersig score (response rate of anti-PD-1 cohort: 54.10% vs. 26.67%, Figure 10A; response rate of anti-CTLA-4 cohort: 66.67% vs. 38.46%, Figure 10B). Taken together, the Fersig score we constructed could be used to predict the prognosis of patients and their response to immunotherapies.

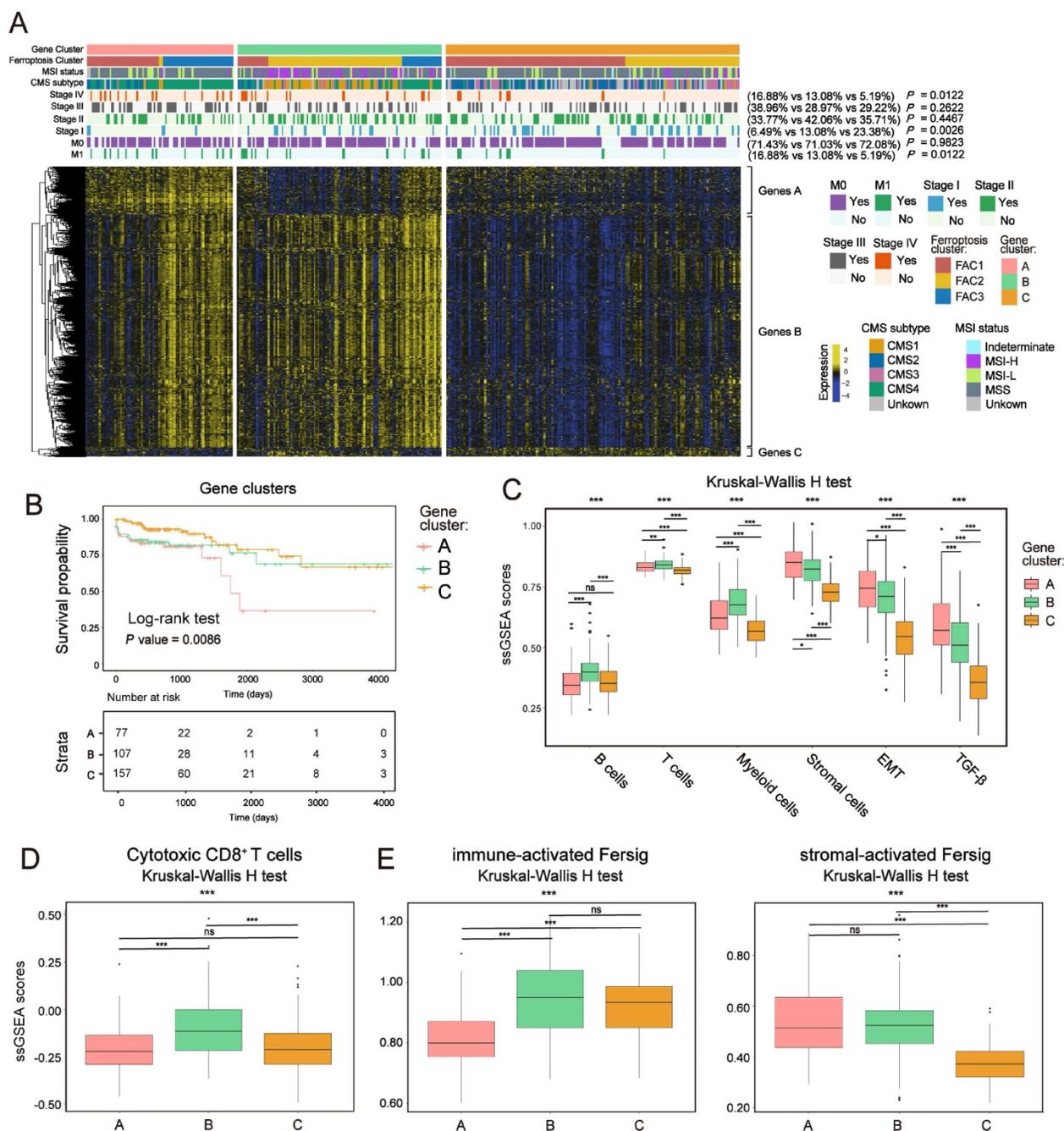


Figure 8. Ferroptosis phenotype-related DEGs in colorectal cancer. (A) NMF clustering of TCGA tumor samples using ferroptosis phenotype-related signature. Clinical and molecular characteristics are shown on the top. The difference of three gene clusters was compared through the Pearson's chi-square test. (B) Kaplan-Meier curves for overall survival of three gene clusters in TCGA. P value was determined by the log-rank test. (C) ssGSEA score of signatures of TME cell types, EMT and TGF- β among three gene clusters in TCGA. The statistical difference of three clusters was compared by the Kruskal-Wallis H test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. The difference of two clusters was compared by the wilcox test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. (D-F) ssGSEA score of signatures of cytotoxic CD8⁺ T cells, immune-activated and stromal-activated ferroptosis-associated genes among three gene clusters. The statistical difference of three clusters was compared by the Kruskal-Wallis H test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. The difference of two clusters was compared by the wilcox test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

We previously identified immune-activated and stromal-activated ferroptosis-associated genes. To further understand their correlation with immunotherapies, we compared the expression level of genes with significant prognosis between low and high Fersig score groups of the anti-PD-1 cohort (Liu et al. study) (Figure S12A-S12B). The results showed that immune-activated genes with better prognosis,

such as SOCS1, were upregulated in the low Fersig score group (Figure 10C), while stromal-activated genes with worse prognosis, such as CDO1, were upregulated in the high Fersig score group (Figure 10E). Subsequent correlation analyses showed that the expression of SOCS1 was positively correlated with the infiltration of cytotoxic CD8⁺ T cells, while the expression of CDO1 didn't show significant

correlation (Figure 10D and 10F). We also utilized tissue microarray (TMA), containing 308 patients, as a patient cohort to explore the correlation between SOCS1/CDO1 and the prognosis of patients. The results showed that SOCS1 could predict the better prognosis, while CDO1 may predict the worse prognosis (Figure 10G and 10I). As our team previously reported[48], the cancer tissues on our TMA could be divided into high or low intraepithelial or stromal CD8⁺ TILs groups. We found that high intraepithelial or stromal CD8⁺ TILs groups (CD8e or

CD8s) were positively correlated with the expression of SOCS1, while CDO1 didn't show significant association with high CD8e or CD8s groups (Figure 10H and 10J; Figure S12C-S12D). These data indicated the considerable immunotherapeutic benefit of CRC with high SOCS1 expression. Overall, immune-activated, and stromal-activated ferroptosis-associated genes we constructed could be the target genes to enhance the outcomes of patients' immunotherapy.

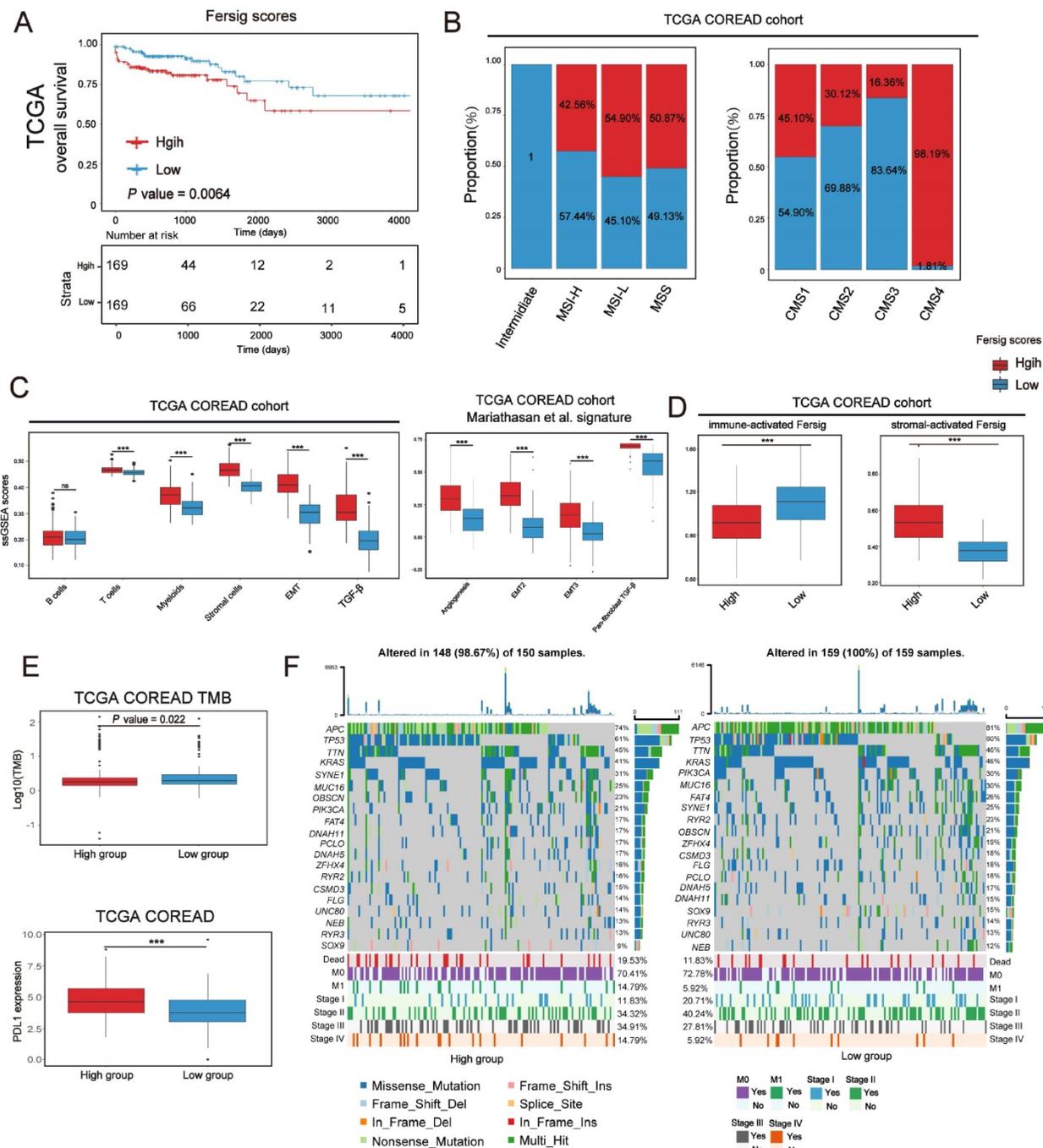


Figure 9. Further construction of Fersig score. (A) Kaplan-Meier curves for overall survival of TCGA. The high and low groups were divided by the median value of the PCA score of Fersig. P value was determined by the log-rank test. **(B)** Barplots show the proportion of molecular characteristics between high and low groups of Fersig. **(C)**

ssGSEA score of signatures of TME cell types, EMT, TGF- β and other stromal-related signatures between high and low groups in TCGA cohort. The difference of two clusters was compared by the wilcox test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. **(D)** ssGSEA score of immune-activated and stromal-activated Fergis between high and low groups in TCGA cohort. The difference of two clusters was compared by the wilcox test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. **(E)** Comparison of TMB and PD-L1 expression between high and low groups. The difference of two clusters was compared by the wilcox test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$. **(F)** Oncoplots show genetic alterations of common mutant genes in CRC between high and low groups. The number on the right indicates the mutation frequency in each regulator. Each column represents individual patient. High group: $n=150$; low group: $n=159$.

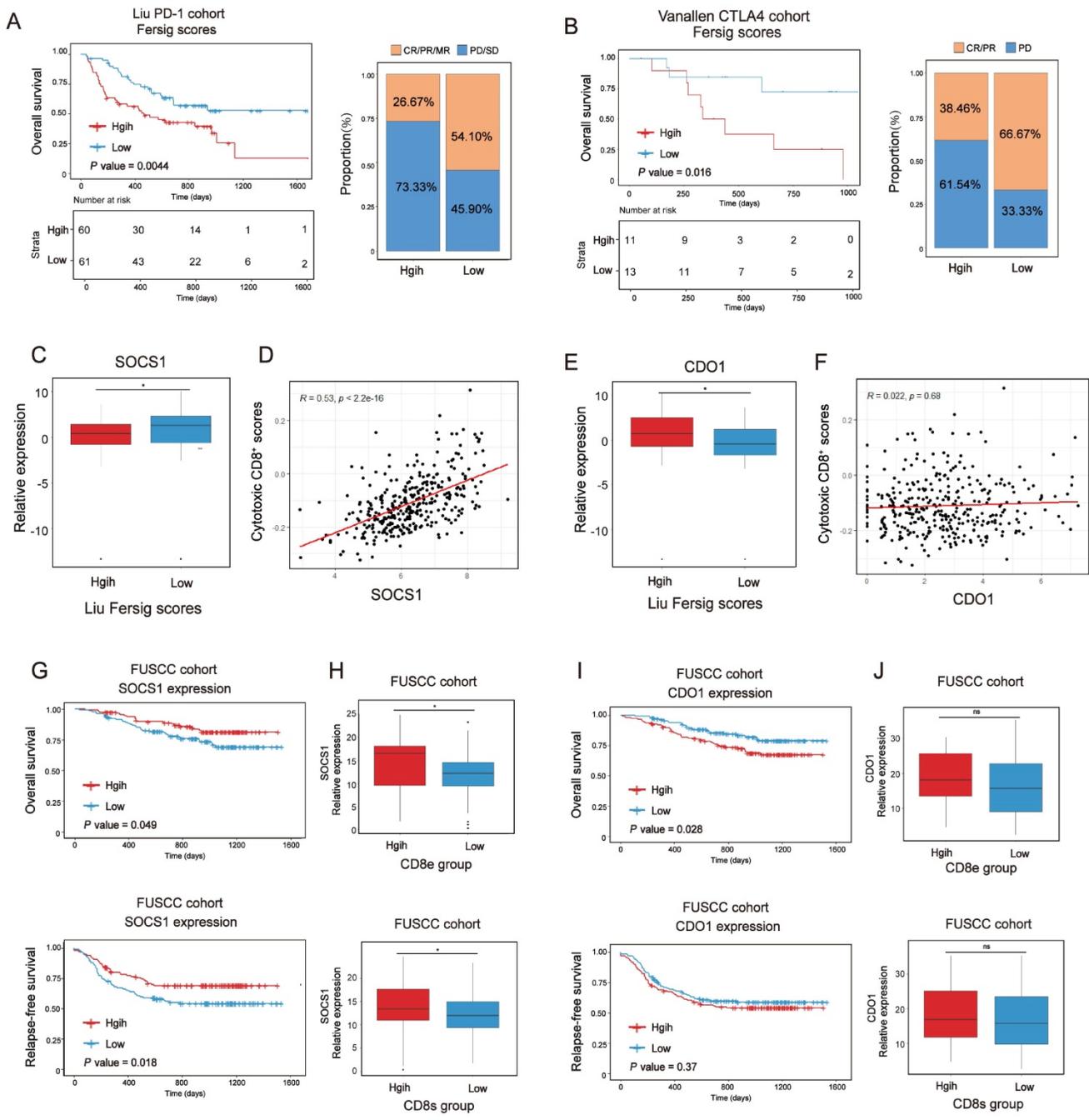


Figure 10. Exploration of Fergis score role in predicting immunotherapeutic benefits. **(A)** Kaplan-Meier curves for high and low Fergis score patient groups in the Liu et al.'s cohort. Log-rank test, $P = 0.0044$. The fraction of patients with clinical response to anti-PD-1 immunotherapy in low or high Fergis score groups of Liu et al.'s cohort. **(B)** Kaplan-Meier curves for high and low Fergis score patient groups in the Vanallen et al.'s cohort. Log-rank test, $P = 0.016$. The fraction of patients with clinical response to anti-CTLA-4 immunotherapy in low or high Fergis score groups of Vanallen et al.'s cohort. **(C, E)** Barplots show the expression of SOCS1/CDO1 in high and low Fergis score patient groups of the Liu et al.'s cohort. **(D, F)** Dotplots show the correlation between the expression of SOCS1/CDO1 and cytotoxic CD8+ T score. **(G, I)** The high and low groups were divided by the median value of gene expression. Kaplan-Meier curves for high and low patient groups in FUSCC TMA cohort. P value was determined by the log-rank test. **(H, J)** The gene expression of SOCS1 and CDO1 in CD8e or CD8s high and low groups. The difference of two groups was compared by the wilcox test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Discussion

Evidences have showed that ferroptosis could regulate the antitumor immunity by their interaction with different immune cell types, which is potentially recognized as a concept of immunogenic death (ICD) [49, 50]. Although various studies have revealed the regulation of ferroptosis-associated genes in TME, a landscape of TME characteristics mediated by ferroptosis-associated genes have not been comprehensively recognized. Therefore, characterizing ferroptosis-associated phenotypes in TME would promote our understanding of how ferroptosis impacts tumor immunity, which may be applied to the existing immunotherapies to enhance their efficacy.

In this study, we identified three distinct ferroptosis-associated molecular patterns characterized by different TME phenotypes, which were related to diverse antitumor immunity in CRC patients. The ferroptosis-associated cluster FAC1 was characterized by paucity of TME cells infiltration, corresponding to the immune-desert phenotype. The FAC2 was characterized by immune activation, consistent with the immune-inflamed and CMS1-like phenotype. The FAC3 was characterized by remarkable stromal cells infiltration, together with EMT, TGF- β signaling pathway activation, identical to the immune-excluded and CMS4-like phenotype. We also performed analyses of scRNA databases to confirmed the biological characteristics of FAC1 and FAC2 on CRC tumor single cells, and highlighted the remarkable infiltration of stromal cells in FAC3 patients. Given that baseline levels of tumor-infiltrating immune, stromal cells and inflammatory cytokines secretion have been showed to predict the immune response [51]. We also identified that FAC2 was significantly associated with elevated cytotoxic CD8⁺ T cells infiltration and PD-L1 expression, supporting its predictive advantage in response to immunotherapy. However, FAC3 patients presented enrichment of myofibroblasts, which was confirmed to be associated with worse prognosis in CRC patients. Previous studies reported that stromal myofibroblasts contributed to an immune suppressive microenvironment by TGF- β production [25, 52]. Although FAC3 patients upregulated the expression of PD-L1, we postulated that CRC patients with FAC3 pattern couldn't effectively respond to immunotherapy because the TGF- β pathways and myofibroblasts impeded the penetration of cytotoxic CD8⁺ T cells into the tumor parenchyma. Therefore, therapies targeting TGF- β might restore the effective response to immunotherapy in CRC patients with FAC3 pattern.

Next, we identified two subsets of ferroptosis-associated genes, recognized as immune-activated and stromal-activated genes. By comprehensive analyses of clinical characteristics, TME infiltration, stromal-related pathways enrichment and TMB level, we found that patients with upregulated expression of immune-activated genes exhibited clinical advantages, immune activation, and the high level of TMB, implying their effective response to immunotherapy. Conversely, patients with upregulated expression of stromal-activated genes displayed clinical disadvantages, stromal-related pathways activation and remarkable stromal cells infiltration, implying their ineffective response to immunotherapy. Notably, consistent with the expression patterns of PD-L1 in FAC2 and FAC3 groups, both patients with upregulation of immune-activated or stromal-activated genes highly expressed PD-L1. Previous studies indicated that the prognostic of PD-L1 expression in CRC is contradictory [53-55]. Blocking PD-L1/PD-1 interaction can prolong tumor suppression or stabilize the progression of cancers [56]. Based on our findings of two phenotypes with distinct TME infiltration in patients with upregulation of PD-L1, we postulated that we should determine the response of CRC patients to immunotherapy mainly based on TME cells infiltration, especially in patients with high expression of PD-L1. In our study, we found that the two subsets of immune-activated and stromal-activated ferroptosis-associated genes could distinguish better or worse response to immunotherapy in CRC patients with high expression of PD-L1.

We also identified the potential ferroptosis-associated genes of immune-activated and stromal-activated subsets based on their significant influence on the prognosis of CRC. In immune-activated subset, XBP1, CISD2, EIF2S1, RRM2, GCH1, SOCS1, NOS2 and PSAT1 displayed positive correlation with prognosis and cytotoxic CD8⁺ T cells infiltration, implying their advantages in immunotherapy. However, these genes have been reported as drivers or suppressors in ferroptosis. Similarly, in stromal-activated subset, PLIN4, RGS4, SLC2A3, SLC2A14, HIC1, DUSP1, CDO1, NOX4 and SLC2A6 exhibited negative correlation with CD8⁺ T cytotoxicity score and positive correlation with myofibroblasts as well as EMT signature score, implying their disadvantages in immunotherapy. They also contained drivers or suppressors of ferroptosis. Therefore, the role of these genes in ferroptosis for CRC should be further explored. Overall, these findings suggested that we could apply the intervention on these target genes to enhance the outcomes of patients' immunotherapy.

Further unsupervised consensus clustering analyses based on ferroptosis phenotype-related signature confirmed the stable TME infiltration patterns in CRC patients, and highlighted the main role of ferroptosis-related process in immune activation and stromal cells infiltration. We also constructed a scoring scheme termed the Fersig score, and verified its positive correlation with stromal-related pathways activation and stromal cells infiltration. Given the absence of CRC immunotherapy dataset, we analyzed different immunotherapy regimens (anti-PD-1/CTLA-4) across different datasets, and verified that low level of Fersig score could predict better response to immunotherapies. As a PCA algorithm, this scoring scheme showed an advantage of focusing the score on the well-correlated genes, and it could downregulate contributions of anti-correlated genes, which could improve the accuracy and efficiency of its prognostic and immunotherapy prediction. In clinical practice, the Fersig scoring scheme could be used to comprehensively investigate the ferroptosis-regulated patterns and corresponding TME infiltration characterization in individual patient. It could not only evaluate patients' clinicopathological features such as stages, molecular subtypes and TMB, but also identify the immune phenotypes of tumors. The Fersig scoring scheme may guide the more effective clinical practice to improve patients' response to immunotherapies.

In conclusion, we comprehensively explored the ferroptosis regulation patterns among 1,251 CRC tumor bulk samples and 29 single-cell samples, and identified their correlation with TME cell-infiltrating characteristics. These integrated analyses highlighted the main role of ferroptosis in immune activation and stromal cells infiltration during CRC development, which will contribute to understanding the TME infiltration based on ferroptosis and provide an interesting insight into immunotherapeutic efficacy.

Methods

Immunohistochemistry (IHC)

Cancer tissue microarray including 308 CRC cases were obtained from the Fudan University Shanghai Cancer Center (FUSCC). This study was approved by the Ethical Committee of the FUSCC and all participants signed written informed consent in accordance with the regulations of the Institutional Review Boards of the FUSCC. Immunohistochemically staining was performed according to standard protocol. Briefly, paraffin-embedded samples were cut into 4 μm sections and placed on polysine coated slides. Paraffin sections were baked overnight at 58°C, dewaxed in xylene, rehydrated

through a graded series of ethanol, quenched for endogenous peroxidase activity in 0.3% hydrogen peroxide for 15 mins. Antigen retrieval was performed by high-pressure cooking in citrate buffer (pH=6.0) for about 20 mins, then allowed to cool to room temperature, blocking the nonspecific antibody binding sites in 5% normal goat serum for 2 hrs. Sections were incubated at 37°C for 1.5 hrs with rabbit polyclonal antibody against SOCS1 (1:100, Affinity, AF5378), and CDO1 (1:100, Proteintech, 12589-1-AP).

Preparation of bulk RNA expression datasets

Colorectal cancer databases (gene expression data and clinical data included) were obtained from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>) and TCGA (<https://xenabrowser.net/datapages/TCGA-COREAD>). For GEO database, a total of 913 patients from GSE39582, GSE14333, and GSE37892 were identified and fully studied in this study. Combat method was used to remove the batch effects among three different GEO datasets. We finally combined three GEO datasets as a meta-GEO cohort.

For TCGA database, RNA expression data of a total 382 patients (338 tumor and 44 normal) were downloaded from the UCSC Xena (<https://xenabrowser.net/datapages/TCGA-COREAD>). The somatic mutation data of TCGA-COREAD were curated from <https://portal.gdc.cancer.gov/repository>. The copy number variation data was curated from the UCSC Xena database. The description of patients with bulk RNA expression data was shown in Table S1.

Preparation of single-cell RNA expression datasets

scRNA data was downloaded from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>). Two cohorts of GSE144735 [25] were identified and fully studied in this study.

We performed Seurat-based analysis to preprocess the single-cell data. The method of quality control (including the threshold of number of detected genes per cell and the percentage of mitochondrial genes expressed) followed Lee *et al.*'s study [25]. For batch correcting, we applied Harmony integration as previously described [57]. The basic information for single-cell datasets of all samples were shown in Table S2.

Ferroptosis-associated genes used for analysis

259 ferroptosis-associated genes were obtained from FerrDb database (<http://www.zhounan.org/ferrdb/>), including 108 drivers, 69 suppressors and 111 markers. The biological information of these genes was shown in Table S3. Only human genes were analyzed in our study. There are totally 71

drivers, 61 suppressors and 96 markers.

Consensus molecular clustering by NMF

To obtain consensus clustering of colorectal cancer samples, we performed consensus clustering with NMF to identify distinct classification based on the expression of ferroptosis-associated human genes. The optimal number of clusters was selected according to cophenetic coefficients (Fig S1A). When rank=3, the consensus matrix heatmap still kept sharp boundaries, indicating its stable clustering. Therefore, rank=3 was finally selected as optimal number of clusters. The NMF clusters of TCGA and meta-GEO were shown in Table S1.

Gene set variation analysis (GSVA) and single-sample gene set enrichment (ssGSEA) analysis

In Figure 1C, we quantified the pathway activities of tumor samples by applying GSVA analysis. The pathway signatures derived from the Hallmark gene sets and C2 curated gene sets (download from MSigDB database v7.4). The significantly expressed pathways in three distinct clusters were shown by heatmap.

The gene signatures in Figure 2B were obtained from CMScaller package, and the gene sets in Figure 8C were constructed by Mariathan *et al.*. The signatures named EMT and TGF- β through our study were obtained from CMScaller package.

Gene signatures of single cell types used for calculating ssGSEA score were obtained from DEGs of cell types detected in Lee *et al.*'s study [25]. The expression difference on a natural log scale was set at least 0.25.

Cytotoxic marker genes were CST7, GZMA, GZMB, IFNG, NKG7 and PRF1. MHC-I marker genes were HLA-A, HLA-B and HLA-C. MHC-II marker genes were HLA-DMA, HLA-DOA, HLA-DRA, HLA-DOB, HLA-DMB, HLA-DPA1, HLA-DPB1, HLA-DQB1, HLA-DQB2, HLA-DPA1, HLA-DPB2, HLA-DQB1 and HLA-DRB.

CMS classification for bulk RNA-seq

We utilized CMScaller to stratify TCGA-COREAD tumor samples. The CMS subtypes annotation were shown in Table S1.

Mapping ferroptosis-associated molecular classification of bulk samples onto CRC tumor cells at single-cell level

We first identified the specific genes among three NMF clusters of TCGA-COREAD cohort using FindAllMarkers function. All genes were probed and the expression difference on a natural log scale was set at least 0.5. Then we subset tumor cells from SMC

and KUL3 cohorts in Lee *et al.*'s (GSE144735) study and scored them with the specific genes of NMF clusters using AddModuleScore function. We finally annotated each cell type based on their maximum gene score. The annotation of single cell was shown in Table S4.

Molecular characteristics of single cells identification

The molecular characteristics of each sample in SMC and KUL3 cohorts have been described in Lee *et al.*'s (GSE144735) study, including CMS subtypes at bulk and single-cell levels, SMAD4 mutation, P53 mutation, Braf mutation, KRAS mutation and MSI status. The molecular types of each CRC tumor cell were annotated using their sample characteristics.

TME infiltration evaluation using ssGSEA and CIBERSORT

In order to precisely quantify the abundance of cell types in CRC tumor, we utilized 5 cell types (epithelial cells, T cells, B cells, Stromal cells and Myeloid cells included) and 31 cell subtypes detected in a single-cell CRC study (GSE144735), and calculated these cell type scores in TCGA and GEO bulk samples using ssGSEA analysis. We further calculated relative abundance of each cell type using the deconvolution approach CIBERSORT. To confirm the stable TME infiltration patterns of NMF clusters, we also evaluated immune cell infiltration with cell types from the study of Charoentong [38] using ssGSEA analysis.

Somatic mutation analyses

The somatic mutation data of TCGA-COREAD were curated from <https://portal.gdc.cancer.gov/repository>. Varscan file format was selected for analyses. The copy number variation data was curated from UCSC Xena database. We identified significantly mutated genes and calculated TMB using 'maftool'.

Constructing immune-activated and stromal-activated Fersig

We utilized 'limma' package to calculate differentially expressed genes (DEGs) between FAC2 and the other two clusters, recognized as DEGs upregulated in FAC2. Then we overlapped DEGs of FAC2 and ferroptosis-associated genes, and obtained 35 genes termed immune-activated subset. We obtained stromal-activated subset by identical analyses.

In Figure 5 and 6, we divided bulk samples into high and low groups by setting the median value of ssGSEA score of immune-activated and stromal-activated genes as the threshold.

Defining cell score

We used the AddModuleScore function in the Seurat R package to evaluate the score of immune-activated and stromal-activated genes on single cells in Figure 7 and S7.

Collection of genomic and clinical information of the ICI-based cohort

Two immunotherapeutic cohorts were included in our study: metastatic melanoma treated with anti-PD-1[46] or anti-CTLA-4 mcAb[47]. The gene expression profiles of pre-therapy biopsy samples were curated and transformed into the TPM format for further analysis.

Construction of the Fersig score

We developed a Fersig scoring scheme to quantify the ferroptosis-associated genes expression level of individual patients by using principal component analysis (PCA) as previously described [58]. Firstly, each DEG among NMF clusters FAC1-3 was calculated across all samples in the TCGA cohort. The overlapping DEGs identified from different clusters were selected and used to perform prognostic analysis using a univariate Cox regression model. NMF unsupervised clustering method using prognostic DEGs was employed to classify patients into three gene clusters in the TCGA and meta-GEO cohort. Then we performed PCA analysis in bulk samples using the DEGs correlated to immune and stromal activation respectively, and principal component 1 (PC1) was extracted to serve as the signature score. After obtaining the prognostic value of each PC1 of immune-related and stromal-related DEGs, we applied a method similar to GGI [59] to define the Fersig score of each patient:

$$\text{Fersig score} = \sum PC1_i - \sum PC1_j$$

where i is the signature score of clusters whose Cox coefficient is positive, and j is the expression of genes whose Cox coefficient is negative. The Fersig score of each database was shown in Table S6.

Statistical analyses

Statistical analysis was performed using R (version 4.0.0) and GraphPad Prism (version 7.04). The Kruskal-Wallis H test, Pearson's chi-square test, wilcox test and logrank test were used in this study. Detailed descriptions of statistical tests are specified in the figure legends and in the results.

Abbreviations

CRC: Colorectal cancer; scRNA: single-cell RNA; TMB: Tumor mutation burden; CNV: Copy number variation; EMT: Epithelial-mesenchymal transition;

GEO: Gene-expression omnibus; TCGA: The cancer genome atlas; GSVA: Gene set variation analysis; HR: Hazard ratios; ICB: Immunological checkpoint blockade; MSI: Microsatellite instability; dMMR: deficient mismatch repair; DEGs: Differentially expressed genes; ssGSEA: single-sample gene-set enrichment analysis; TGF- β : Transforming growth factor beta; TME: Tumor microenvironment.

Supplementary Material

Supplementary figures and tables.

<https://www.ijbs.com/v18p1773s1.pdf>

Acknowledgements

Author Contributions

CGX designed this work and revised this manuscript. MSB, XWQ, HLY, YL, ZWS, LCZ, YR, WRJ, XY and CSJ prepared the data. LWQ, DWX, LQG, GRQ and XXY integrated and analyzed the data. LWQ, DWX, LQG and MSB wrote this manuscript. All authors approved this manuscript.

Funding

This work was supported by the grants of Science and Technology Commission of Shanghai Municipality (No. 19140902100), National Natural Science Foundation of China (No. 8217112384, No. 81871958 AND NO. 82103554), Shanghai Municipal Health Commission (No. SHWJRS 2021-99) and Fudan University (NO. IDF159052 and No. DGF8 28007).

Availability of data and materials

All data in this study can be obtained from the Gene-Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>), the GDC portal (<https://portal.gdc.cancer.gov/>) and the UCSC Xena (<https://xenabrowser.net/datapages/>).

Competing Interests

The authors have declared that no competing interest exists.

References

1. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, et al. Colorectal cancer statistics, 2017. *CA Cancer J Clin.* 2017; 67: 177-93.
2. Siegel RL, Miller KD, Fuchs HE, Jemal A. *Cancer Statistics, 2021.* *CA Cancer J Clin.* 2021; 71: 7-33.
3. Wang H, Cheng Y, Mao C, Liu S, Xiao D, Huang J, et al. Emerging mechanisms and targeted therapy of ferroptosis in cancer. *Mol Ther.* 2021; 29: 2185-208.
4. Dixon SJ, Lemberg KM, Lamprecht MR, Skouta R, Zaitsev EM, Gleason CE, et al. Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell.* 2012; 149: 1060-72.
5. Komatsu M, Kurokawa H, Waguri S, Taguchi K, Kobayashi A, Ichimura Y, et al. The selective autophagy substrate p62 activates the stress responsive transcription factor Nrf2 through inactivation of Keap1. *Nat Cell Biol.* 2010; 12: 213-23.

6. Brown CW, Amante JJ, Chhoy P, Elaimy AL, Liu H, Zhu LJ, et al. Prolimin2 Drives Ferroptosis Resistance by Stimulating Iron Export. *Dev Cell*. 2019; 51: 575-86.e4.
7. Zhu S, Zhang Q, Sun X, Zeh HJ, 3rd, Lotze MT, Kang R, et al. HSPA5 Regulates Ferroptotic Cell Death in Cancer Cells. *Cancer Res*. 2017; 77: 2064-77.
8. Sui X, Zhang R, Liu S, Duan T, Zhai L, Zhang M, et al. RSL3 Drives Ferroptosis Through GPX4 Inactivation and ROS Production in Colorectal Cancer. *Front Pharmacol*. 2018; 9: 1371.
9. Zhang L, Liu W, Liu F, Wang Q, Song M, Yu Q, et al. IMCA Induces Ferroptosis Mediated by SLC7A11 through the AMPK/mTOR Pathway in Colorectal Cancer. *Oxid Med Cell Longev*. 2020; 2020: 1675613.
10. Chen P, Li X, Zhang R, Liu S, Xiang Y, Zhang M, et al. Combinative treatment of β -elemene and cetuximab is sensitive to KRAS mutant colorectal cancer cells by inducing ferroptosis and inhibiting epithelial-mesenchymal transformation. *Theranostics*. 2020; 10: 5107-19.
11. Serebrniiskii IG, Connelly C, Frampton G, Newberg J, Cooke M, Miller V, et al. Comprehensive characterization of RAS mutations in colon and rectal cancers in old and young patients. *Nat Commun*. 2019; 10: 3722.
12. Fridman WH, Zitvogel L, Sautès-Fridman C, Kroemer G. The immune contexture in cancer prognosis and treatment. *Nat Rev Clin Oncol*. 2017; 14: 717-34.
13. Turley SJ, Cremasco V, Astarita JL. Immunological hallmarks of stromal cells in the tumour microenvironment. *Nat Rev Immunol*. 2015; 15: 669-82.
14. Pagès F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, et al. International validation of the consensus Immunoscope for the classification of colon cancer: a prognostic and accuracy study. *Lancet*. 2018; 391: 2128-39.
15. Xiao Y, Freeman GJ. The microsatellite instable subset of colorectal cancer is a particularly good candidate for checkpoint blockade immunotherapy. *Cancer Discov*. 2015; 5: 16-8.
16. Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, et al. Immunotherapy in colorectal cancer: rationale, challenges and potential. *Nat Rev Gastroenterol Hepatol*. 2019; 16: 361-75.
17. Kreidieh M, Mukherji D, Temraz S, Shamseddine A. Expanding the Scope of Immunotherapy in Colorectal Cancer: Current Clinical Approaches and Future Directions. *Biomed Res Int*. 2020; 2020: 9037217.
18. Pfeifhofer-Obermair C, Tymoszuk P, Petzer V, Weiss G, Nairz M. Iron in the Tumor Microenvironment-Connecting the Dots. *Front Oncol*. 2018; 8: 549.
19. Cairo G, Recalcati S, Mantovani A, Locati M. Iron trafficking and metabolism in macrophages: contribution to the polarized phenotype. *Trends Immunol*. 2011; 32: 241-7.
20. Wang W, Green M, Choi JE, Gijón M, Kennedy PD, Johnson JK, et al. CD8(+) T cells regulate tumour ferroptosis during cancer immunotherapy. *Nature*. 2019; 569: 270-4.
21. Yu B, Choi B, Li W, Kim DH. Magnetic field boosted ferroptosis-like cell death and responsive MRI using hybrid vesicles for cancer immunotherapy. *Nat Commun*. 2020; 11: 3637.
22. Jorissen RN, Gibbs P, Christie M, Prakash S, Lipton L, Desai J, et al. Metastasis-Associated Gene Expression Changes Predict Poor Outcomes in Patients with Dukes Stage B and C Colorectal Cancer. *Clin Cancer Res*. 2009; 15: 7642-51.
23. Laibe S, Lagarde A, Ferrari A, Monges G, Birnbaum D, Olschwang S. A seven-gene signature aggregates a subgroup of stage II colon cancers with stage III. *Oncotarget*. 2012; 3: 560-5.
24. Marisa L, de Reyniès A, Duval A, Selves J, Gaub MP, Vescovo L, et al. Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med*. 2013; 10: e1001453.
25. Lee HO, Hong Y, Etioglu HE, Cho YB, Pomella V, Van den Bosch B, et al. Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer. *Nat Genet*. 2020; 52: 594-603.
26. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature*. 2017; 541: 321-30.
27. Zhou N, Bao J. FerrDb: a manually curated resource for regulators and markers of ferroptosis and ferroptosis-disease associations. *Database (Oxford)*. 2020; 2020.
28. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature*. 1999; 401: 788-91.
29. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics*. 2013; 14: 7.
30. Eide PW, Bruun J, Lothe RA, Svein A. CMScaller: an R package for consensus molecular subtyping of colorectal cancer pre-clinical models. *Scientific Reports*. 2017; 7: 16618.
31. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. *Nature Medicine*. 2015; 21: 1350-6.
32. Battaglia AM, Chirillo R, Aversa I, Sacco A, Costanzo F, Biamonte F. Ferroptosis and Cancer: Mitochondria Meet the "Iron Maiden" Cell Death. *Cells*. 2020; 9.
33. Sadanandam A, Lyssiotis CA, Homicsko K, Collisson EA, Gibb WJ, Wullschlegel S, et al. A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nat Med*. 2013; 19: 619-25.
34. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012; 487: 330-7.
35. Neeffes J, Jongsma MLM, Paul P, Bakke O. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature Reviews Immunology*. 2011; 11: 823-36.
36. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-García W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature Communications*. 2013; 4: 2612.
37. Guo X, Zhang Y, Zheng L, Zheng C, Song J, Zhang Q, et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat Med*. 2018; 24: 978-85.
38. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep*. 2017; 18: 248-62.
39. Lindau D, Gielen P, Kroesen M, Wesseling P, Adema GJ. The immunosuppressive tumour network: myeloid-derived suppressor cells, regulatory T cells and natural killer T cells. *Immunology*. 2013; 138: 105-15.
40. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGF β attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature*. 2018; 554: 544-8.
41. Hegde PS, Karanikas V, Evers S. The Where, the When, and the How of Immune Monitoring for Cancer Immunotherapies in the Era of Checkpoint Inhibition. *Clin Cancer Res*. 2016; 22: 1865-74.
42. Herbst RS, Soria JC, Kowanetz M, Fine GD, Hamid O, Gordon MS, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature*. 2014; 515: 563-7.
43. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol*. 2019; 30: 44-56.
44. Sun C, Mezzadra R, Schumacher TN. Regulation and Function of the PD-L1 Checkpoint. *Immunity*. 2018; 48: 434-52.
45. Seidel JA, Otsuka A, Kabashima K. Anti-PD-1 and Anti-CTLA-4 Therapies in Cancer: Mechanisms of Action, Efficacy, and Limitations. *Frontiers in oncology*. 2018; 8: 86.
46. Liu D, Schilling B, Liu D, Sucker A, Livingstone E, Jerby-Arnon L, et al. Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nature medicine*. 2019; 25: 1916-27.
47. Miao D, Margolis CA, Gao W, Voss MH, Li W, Martini DJ, et al. Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science (New York, NY)*. 2018; 359: 801-6.
48. Feng Y, Li Y, Cai S, Peng J. Immunological nomograms predicting prognosis and guiding adjuvant chemotherapy in stage II colorectal cancer. *Cancer management and research*. 2019; 11: 7279-94.
49. Tang R, Xu J, Zhang B, Liu J, Liang C, Hua J, et al. Ferroptosis, necroptosis, and pyroptosis in anticancer immunity. *Journal of Hematology & Oncology*. 2020; 13: 110.
50. Xu H, Ye D, Ren M, Zhang H, Bi F. Ferroptosis in the tumor microenvironment: perspectives for immunotherapy. *Trends in molecular medicine*. 2021; 27: 856-67.
51. Topalian SL, Taube JM, Anders RA, Pardoll DM. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nature reviews Cancer*. 2016; 16: 275-87.
52. Koliariaki V, Pallangyo CK, Gretten FR, Kollias G. Mesenchymal Cells in Colon Cancer. *Gastroenterology*. 2017; 152: 964-79.
53. Huang CY, Chiang SF, Ke TW, Chen TW, You YS, Chen WT, et al. Clinical significance of programmed death 1 ligand-1 (CD274/PD-L1) and intra-tumoral CD8+ T-cell infiltration in stage II-III colorectal cancer. *Scientific reports*. 2018; 8: 15658.
54. Lee KS, Kim BH, Oh HK, Kim DW, Kang SB, Kim H, et al. Programmed cell death ligand-1 protein expression and CD274/PD-L1 gene amplification in colorectal cancer: Implications for prognosis. *Cancer science*. 2018; 109: 2957-69.
55. Enkhbat T, Nishi M, Takasu C, Yoshikawa K, Jun H, Tokunaga T, et al. Programmed Cell Death Ligand 1 Expression Is an Independent Prognostic Factor in Colorectal Cancer. *Anticancer research*. 2018; 38: 3367-73.

56. Brahmer JR, Tykodi SS, Chow LQ, Hwu WJ, Topalian SL, Hwu P, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *The New England journal of medicine*. 2012; 366: 2455-65.
57. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature Methods*. 2019; 16: 1289-96.
58. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, et al. Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures. *Cancer Immunol Res*. 2019; 7: 737-50.
59. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst*. 2006; 98: 262-72.