Biomedical Research 2024; 35 (2): 1-13

ISSN 0970-938X
www.biomedres.info

Gastric cancer: Profiling CAFs and prognostic signature development through single-cell RNA sequencing and TCGA analysis.

Xiaofei Sun, Song Gao, Lili Hu, Yanli Liu*

Dongying Traditional Chinese Medicine Hospital, Dongying, Shandong 257099, China

Abstract

Background: Cancer-associated Fibroblasts (CAFs) play critical roles in tumor growth, angiogenesis, metastasis, and therapy resistance. This study aimed to investigate the characteristics of CAFs in Gastric Cancer (GC) and develop a CAF-based risk signature for predicting the prognosis of GC patients.

Methods: Utilizing scRNA-seq data from GEO and survival prognosis data from The Cancer Genome Atlas (TCGA), CAF clusters were identified with Seurat R based on unique markers. DEGs between normal and tumor samples in TCGA were pinpointed. Pearson correlation analysis unveiled DEGs associated with CAF clusters, followed by univariate Cox regression to identify prognostic CAF-related genes. A risk signature was then crafted using Lasso regression on these genes. Finally, an integrated scoring model was developed, merging the risk signature with clinicopathological factors.

Results: Analyzing scRNA-seq data in GC, we identified six distinct clusters of CAFs. Five of these clusters significantly correlated with GC prognosis. We pinpointed 557 DEGs strongly linked to these CAF clusters and derived a refined risk signature of six key genes. These genes are primarily involved in 39 crucial pathways such as angiogenesis, apoptosis, and hypoxia. Our risk signature shows notable associations with stromal and immune scores, as well as specific immune cell types. Multivariate analysis confirms its independent prognostic value in GC, suggesting potential for predicting immunotherapy outcomes. Integrating stage with the CAF-based risk signature, we created a novel scoring model with robust predictive performance for GC prognosis.

Conclusion: The CAF-derived risk signature emerges as a potent prognostic tool for GC, offering valuable insights into the intricate landscape of CAFs within the tumor microenvironment. Such comprehensive profiling may hold promise in guiding personalized immunotherapeutic strategies and refining treatment modalities for GC.

Keywords: Gastric cancer, CAFs, scRNA-seq, TCGA, Prognostic signature.

Abbreviations: CAFs: Cancer-associated Fibroblasts; GC: Gastric Cancer; scRNA-seq: Single-cell RNA Sequencing; DEGs: Differentially Expressed Genes; TME: Tumor Microenvironment; HCC: Hepatocellular Carcinoma; TAMs: Tumor-associated Macrophages; TILs: Tumor-infiltrating Lymphocytes; GEO: Gene Expression Omnibus; SNV: Single-nucleotide Variant; CNV: Copy Number Variants; TCGA: The Cancer Genome Atlas; UMAP: Uniform Manifold Approximation and Projection; TSNE: T-distributed Stochastic Neighbour Embedding; KEGG: Kyoto Encyclopedia of Genes and Genomes; FDR: False Discovery Rate; ROC: Receiver Operating Characteristic; DCA: Decision Curve Analysis; ICB: Immune Checkpoint Blockade; SERPINE1: Serpin Family E Member 1; MATN3: Matrilin 3; OLFM3: Olfactomedin 3; LINC02408: Long Intergenic Non-Protein Coding RNA 2408; GPX3: Glutathione Peroxidase 3; AUC: Area Under the Curve; CR: Complete Response; PR: Partial Response; SD: Stable Disease; PD: Progressive Disease; HR: Hazard Ratio; CI: Confidence Interval; TimeROC: Time-dependent Receiver Operating Characteristic; TIME: Tumor Immune Microenvironment; CAR-T: Chimeric Antigen Receptor T.

Accepted on April 11, 2024

Introduction

According to the latest data released by GLOBOCAN, the global annual incidence of gastric cancer reached 1.089 million in 2020, ranking fifth among all malignant tumors. In the same year, 769,000 people died from gastric cancer (age-standardized death rate of 7.7 per 100,000 population), ranking fourth among all cancer types, following lung cancer, colorectal cancer, and liver cancer. A study published in 2022 estimated that by 2040, the number of new cases worldwide will increase by 62% to reach

1.77 million [1]. Despite the availability of surgical and pharmacological treatments, gastric cancer still leads to high mortality and morbidity rates with poor prognosis [2]. Research indicates that infection with Helicobacter pylori is the primary cause of gastric cardia cancer [3]. Current research evidence suggests that secondary prevention strategies for gastric cancer, primarily using gastroscopy as the main screening method, are only cost-effective in regions or high-risk populations with a high incidence of gastric cancer [4]. With the advancements in genomics, proteomics, and microbiomics, novel and updated multi-gene features have shown significant value in predicting the prognosis and recurrence of gastric cancer.

The Tumor Microenvironment (TME) consists of tumor cells and stromal cells, playing pivotal roles in cancer progression [5]. Among the stromal elements, CAFs are central players, originating from diverse cell sources including bone marrow-derived mesenchymal stem cells, adipocytes, and cancer cells themselves [6]. CAFs interact intricately with cancer cells, impacting tumor growth, spread, and response to treatment by secreting various factors [7]. In GC, CAFs have been linked to lymph node metastasis, tumor staging, and progression [8,9]. Despite extensive study, a comprehensive understanding of CAFs' systemic characteristics and their influence on GC prognosis and response to immunotherapy remains elusive. In this investigation, we utilized GC scRNA-seq data and TCGA transcriptome data to identify distinct subclusters of CAFs [10-12]. We then developed a CAF-centered risk signature specific to GC, assessing its clinical relevance and association with the immune landscape and immunotherapy response. To aid clinical application, we crafted a novel nomogram integrating the CAF-based risk signature with clinicopathological features [13]. Our study provides fresh insights into GC pathophysiology, potentially guiding tailored therapeutic approaches for improved patient outcomes [14].

Materials and Methods

Data acquisition and processing

For this study, scRNA-seq data were sourced from the Gene Expression Omnibus (GEO) database, specifically from dataset GSE167297. This dataset included 10 primary tumor samples, 2 portal vein tumor thrombi samples, 1 metastatic lymph node sample, and 8 non-tumor liver samples. Preprocessing of the scRNA-seq data followed specific criteria: gene expression was required in at least 3 cells, with each cell expressing a minimum of 250 genes. Mitochondrial and rRNA proportions were evaluated using the percentage feature set function within the Seurat R package. Further refinement involved filtering single cells based on the expression of a minimum of 6000 genes with UMI>100, resulting in a final count of 22088 cells for subsequent analyses.

Transcriptome data, Single-nucleotide Variant (SNV) data, Copy Number Variants (CNV) data for Masked Copy Number Segment, and clinical information pertaining to GC were sourced from TCGA database. Samples lacking survival data and outcome status were excluded, leaving a total of 333 tumor samples and 30 para-cancerous samples for our analyses.

To delineate cancer-relevant pathways, we curated a list of ten pathways based on literature review. This comprehensive approach aimed to elucidate the molecular landscape of GC, providing valuable insights into potential therapeutic targets and prognostic indicators [15].

Definition of CAF

To define the Cancer-associated Fibroblast (CAF) signature, we re-analyzed GC scRNA-seq data using the Seurat package [16]. Cells with over 6000 or below 250 expressed genes were filtered out, followed by log normalization of the remaining genes. Batch effects among the 14 samples were mitigated using the FindIntegrationAnchors function. We then employed the Uniform Manifold Approximation and Projection (UMAP) for dimensional reduction with 15 principal components and a resolution of 0.2. Subsequently, clustering of single cells into subgroups was performed using the FindNeighbors and FindClusters functions (dim=40, resolution=0.2). Identification of CAF subgroups through ACTA2, FAP, PDGFRB, and NOTCH3 marker genes. Fibroblasts underwent further clustering with FindNeighbors and FindClusters functions, followed by TSNE dimensionality reduction. Functional pathway analysis of CAF clusters was conducted using Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis, facilitated by the cluster Profiler package [17]. Furthermore, we examined CNV profiles within CAF clusters using the CopyKAT R package. This analysis provided insights into distinct features between tumor cells and adjacent normal cells within each sample, revealing the dynamic interplay within the tumor microenvironment [18].

Identification of CAF hub genes

Differential gene expression analysis between tumor and normal tissue was conducted using the limma package, with an FDR threshold <0.05 and log2(Fold Change)>1. Next, correlations between DEGs and CAF clusters were assessed, identifying key CAF-related genes with a p-value<0.001 and correlation coefficient >0.4. Prognosis-related genes were identified through univariate Cox regression analysis ($p<0.05$). To reduce gene numbers, lasso Cox regression analysis was employed, followed by multivariate Cox regression with stepwise regression. Based on the multivariate Cox model, a risk signature was constructed: risk score=$\Sigma bi*Expi$, where 'i' represents genes in the signature, 'Expi' represents their expression, and 'bi' represents the coefficients from the Cox model. Patients were stratified into high- and low-risk groups after zero-mean normalization. The predictive performance of the risk signature was assessed using ROC curve analysis with the timeROC package in both training and validation cohorts.

Immune landscape assessment

The CIBERSORT algorithm was used to determine the distribution of 22 immune cell subtypes in the TCGA cohort. Additionally, immune and stromal scores were derived using the ESTIMATE algorithm to provide insights into the TME [19].

Risk signature construction and nomogram development

A prognostic nomogram was constructed using univariate and multivariate Cox regression analyses on clinicopathological factors and the risk signature. Variables with $p<0.05$ in the multivariate Cox model were included. The nomogram was developed using the rms package, and its accuracy was assessed with calibration curves and Decision Curve Analysis (DCA) [20].

Assessment of immunotherapy responsiveness

Transcriptomic data from GC patients treated with atezolizumab (IMvigor210 cohort) [21] and pre-treatment melanomas receiving anti-PD-1 therapy (GSE78220 cohort) were used to evaluate the risk signature's potential in predicting response to Immune Checkpoint Blockade (ICB) [22].

Statistical analysis

All analyses were performed in R (v4.2.1). Pearson or Spearman correlation analysis was used for correlation matrices. The Wilcoxon test compared two groups, while survival disparities were assessed with Kaplan-Meier curves and Log-rank tests ($p<0.05$ considered significant).

Results

Identifying CAFs in scRNA-seq samples

Data from GSE167297, downloaded from GEO, underwent quality control and merging, resulting in 22,088 cells for analysis. After logarithmic normalization and dimensionality reduction, 21 subclusters were identified. Among them, 2 CAF subclusters were distinguished. Subsequent clustering and dimensionality reduction within these 2 CAF subclusters revealed six distinct CAF clusters. None of these clusters expressed epithelial cell-specific genes, confirming accurate CAF identification. The t-SNE plot in Figure 1A illustrates the distribution of the 14 samples, showing the presence of 6 distinct CAF clusters for further analysis (Figure 1B).

A total of 557 Differentially Expressed Genes (DEGs) were identified within these CAF clusters, with expression profiles of top five DEGs (as cluster markers) depicted in Figure 1C. Proportions of the six clusters in each cohort are shown in Figure 1D. KEGG analysis revealed enrichment of DEGs in pathways such as cytokine receptor interaction, chemokine signaling, and MAPK signaling (Figure 1E). Additionally, based on CNV features [18], the six CAF clusters included 435 tumor cells and 226 normal cells (Figure 1F).

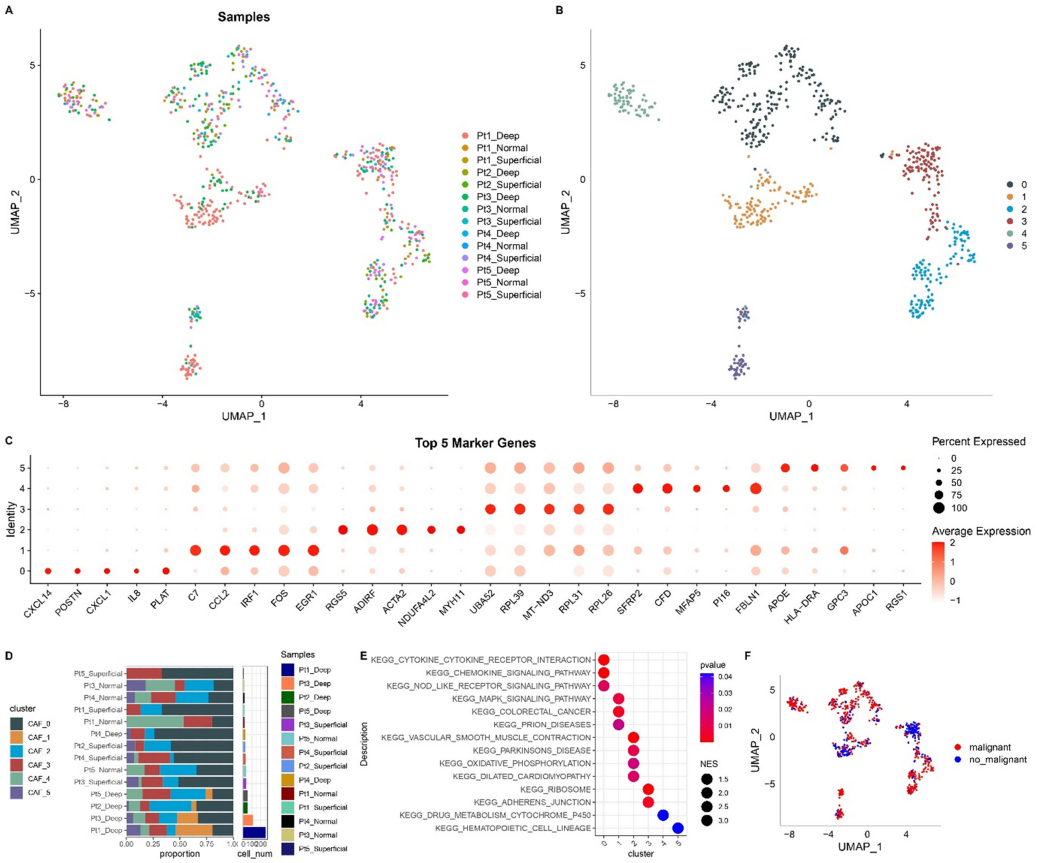

Figure 1. *Analysis of CAF clusters in GC patients using scRNA-seq data: (A) t-SNE plot of sample distribution (14 samples); (B) t-SNE plot of 6 fibroblast clusters post-clustering; (C) dot plot showing top 5 marker gene expression in subgroups; (D) subgroups within cancer tissue, with proportions and cell counts in adjacent tissue; (E) KEGG enrichment analysis of 6 fibroblast subsets; (F) t-SNE distribution map predicting malignant and non-malignant cells using CopyKAT.*

The expression of cancer-related pathways in CAF

To explore the link between CAF clusters and tumor progression, we analyzed ten tumor-related pathways within the six CAF clusters. Figure 2A shows GSVA scores for these pathways across the clusters. Notably, CAF_0, CAF_2, CAF_4, and CAF_5 had significantly higher proportions of malignant cells than the other clusters (Figure 2B). We also examined GSVA scores of the ten pathways across malignant and non-malignant cells in each CAF cluster, revealing nuanced variations (Figure 2C).

To assess prognostic significance, we calculated ssGSEA scores for marker genes (top 5 DEGs from Figure 1C) of each CAF cluster in the TCGA cohort. Results showed notable differences, with CAF_0 displaying elevated scores in tumor samples but the remaining clusters showing higher scores in normal samples (Figure 3A). Using the survminer R package to find optimal cut-off values, we stratified TCGA Gastric Cancer (GC) samples into high- and low-CAF score groups. Except for CAF_3, other clusters showed better prognosis in the high-CAF score group (Figure 3B). These findings suggest that while CAF_3 differs in GC and normal samples, its role in GC progression may be limited.

Identification of hub genes associated with CAF

To create a risk signature, we initially identified DEGs between tumor and normal tissues. Figure 4A shows a total of 36,782 DEGs, with 9,122 up-regulated and 17,889 down-regulated. Among these, 2,808 genes correlated significantly with prognosis-related CAF clusters. We then conducted univariate Cox regression, identifying 189 genes of prognostic significance (Figures 4A and 4B).

Refining gene selection, Lasso Cox regression narrowed to 5 genes with a lambda of 0.0851 (Figures 4C and 4D). Finally, using multivariate Cox regression with stepwise selection, we included 5 genes in the risk signature: SERPINE1, MATN3, OLFM3, LINC02408, and GPX3 (Figure 4E). Risk scores were calculated for each sample, separating into high- and low-risk groups after z-mean normalization. Model AUCs for 1- to 5-year survival ranged from 0.65 to 0.8 in TCGA (Figure 4F). Kaplan-Meier analysis revealed significantly poorer survival for high-risk vs. low-risk patients in TCGA (Figure 4G).

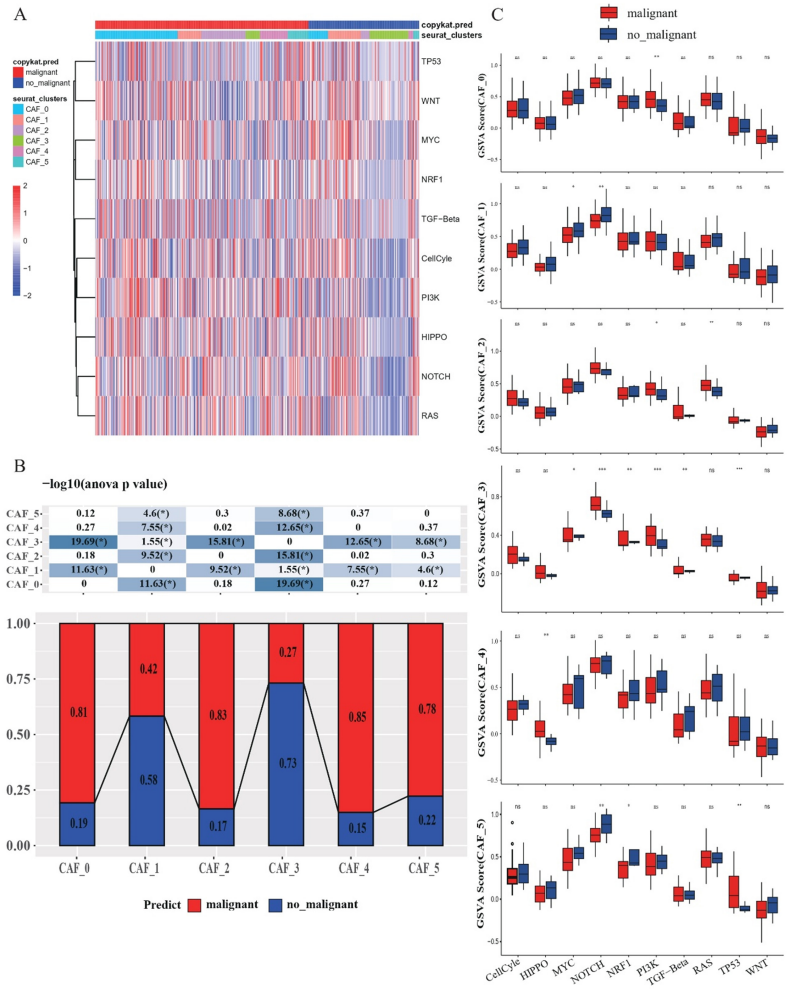

Figure 2. *Tumor-related pathway characteristics in CAF clusters: (A) Heatmap displaying enrichment scores of 10 tumor-related pathways in CAF cells; (B) Comparison of CAF clusters in malignant and non-malignant cells; GSVA score comparisons for each pathway within CAF_0, CAF_1, CAF_2, CAF_3, CAF_4, and CAF_5 clusters (C) Statistical significance was determined using Wilcoxon test.*
Note: **P<0.05; **P<0.01; ***P<0.001; and ****P<0.0001, ns: not significant.*

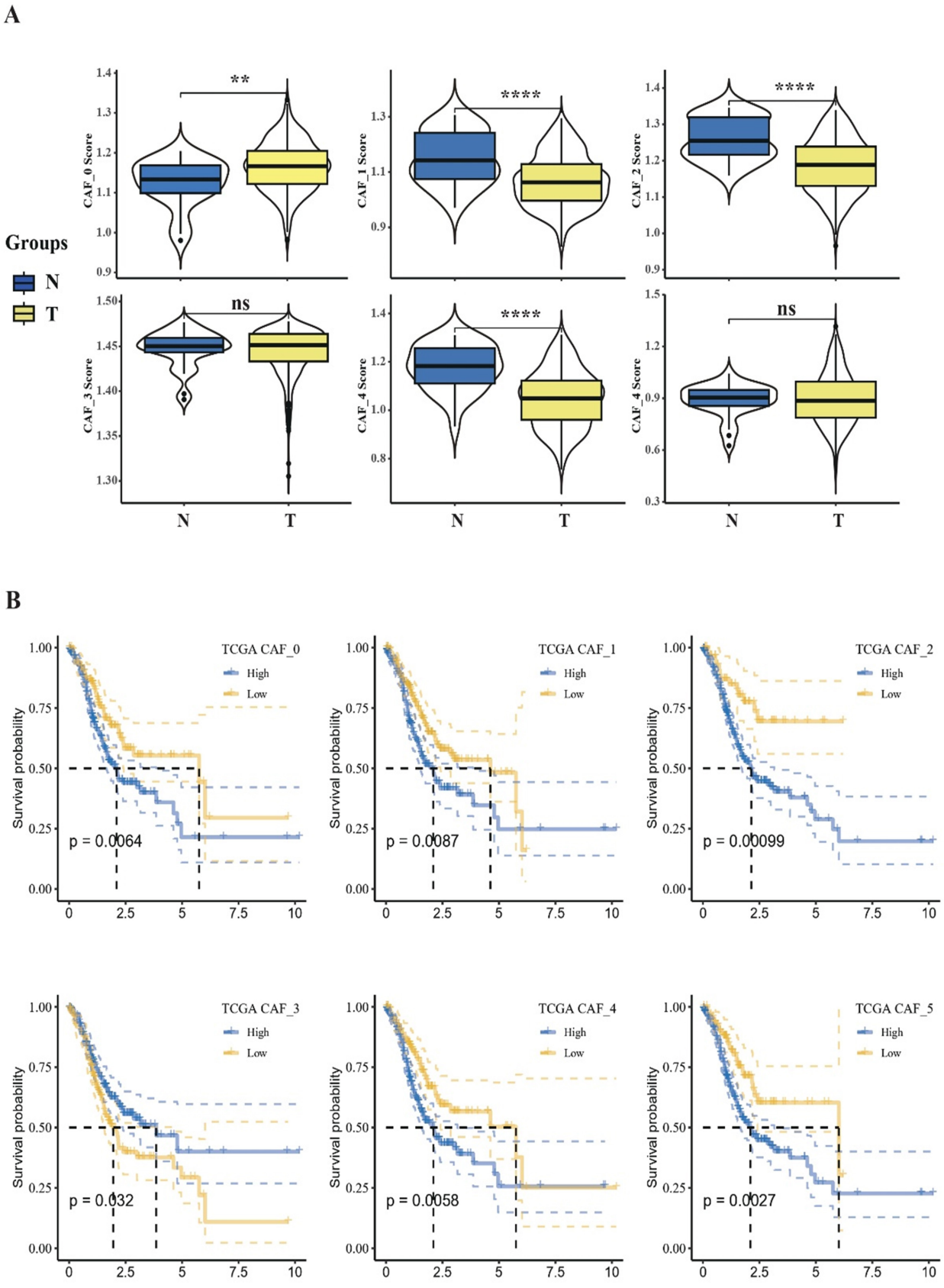

Figure 3. *Prognostic implications of the six distinct CAF clusters in GC patients: (A) Comparative evaluation of CAF scores in cancer and normal tissues; (B) Kaplan-Meier survival curves depicting outcomes for high and low CAF score cohorts within the 6 CAF clusters.*
Note: *Statistical significance is denoted by **P<0.01 and ****P<0.0001.*

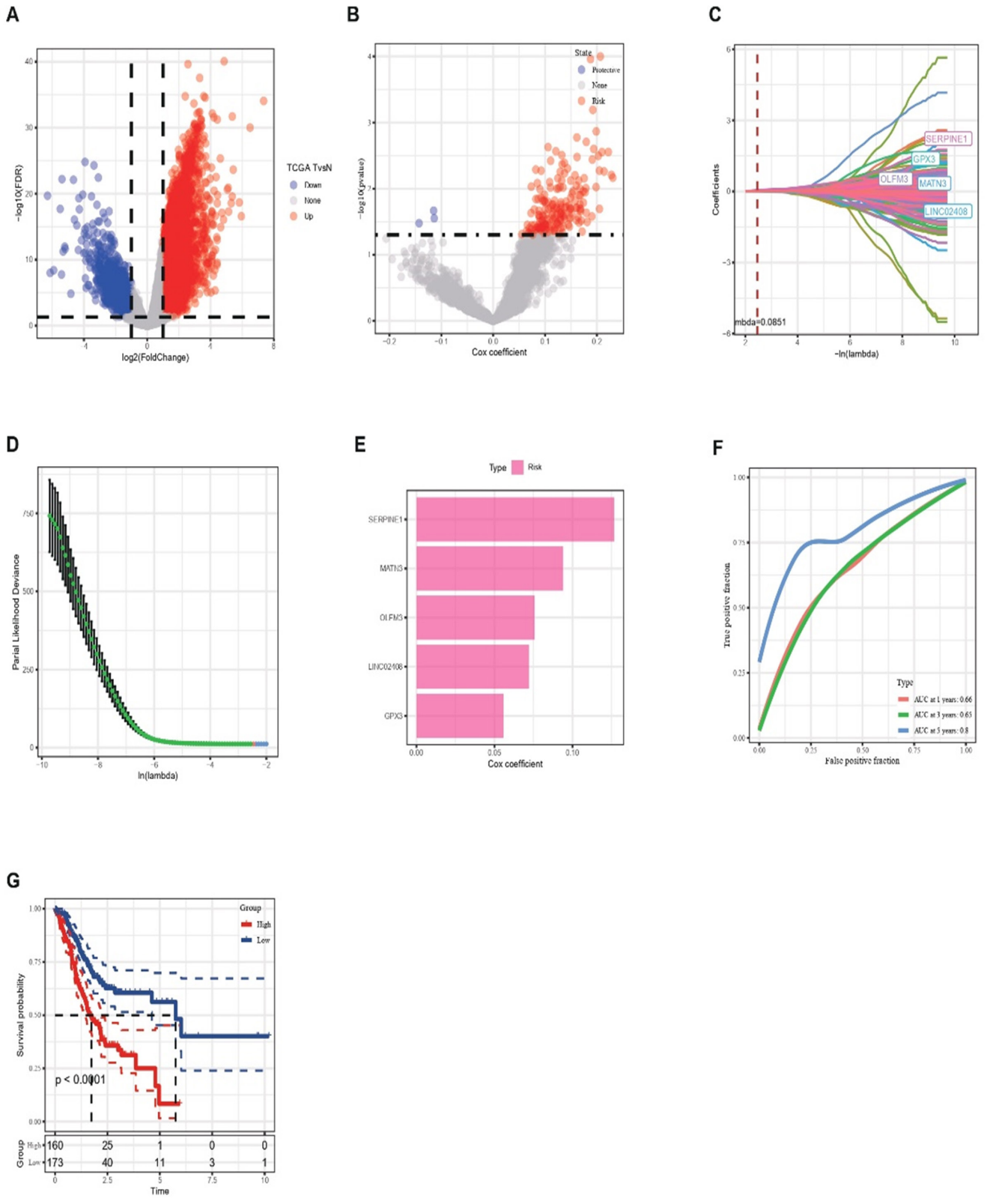

Figure 4. *Identification of pivotal predictive genes for the development of a risk signature involved the following steps: (A) Visualization of differentially expressed genes in cancer and normal tissues using a TCGA cohort volcano plot; (B) Construction of a volcano plot displaying prognosis-related genes identified from univariate Cox regression analysis; (C) Trajectory visualization of each independent variable with lambda values; (D) Plot presentation of coefficient distributions for the logarithmic (lambda) series used in parameter selection; (E) Illustration of multivariate Cox coefficients for each gene in the risk signature; (F) Assessment of the risk model's performance, comprising five genes, through ROC curves in TCGA; (G) Analysis of the risk model's prognostic capability via Kaplan-Meier curves in TCGA cohorts.*

Mutation and pathway analysis of hub genes

Our analysis of SNV mutations within the 5 genes of our risk signature revealed varying frequencies. Specifically, SERPINE1, MATN3, OLFM3, and GPX3 showed higher SNV mutation rates, while LINC02408 exhibited none (Figure 5A). We also investigated potential co-occurrence among these pivotal genes and the 10 most frequently mutated genes. Figure 5B indicates no significant co-occurrence among these five genes, except for MATN3, which displayed a noteworthy probability of co-occurrence with LRP1B mutation. Regarding CNV, only a limited number of samples showed gain/loss among the five genes (Figure 5C). To further understand the relationship between the risk genes and Gastric Cancer (GC), we explored their correlations with various GC molecular signatures. Notably, MATN3 exhibited significant positive correlations with Aneuploidy Score, Homologous Recombination Defects, Fraction Altered, and Number of Segments. Conversely, GPX3 and OLFM3 showed negative correlations with Aneuploidy Score, Fraction Altered, and Nonsilent Mutation Rate (Figure 5D). Furthermore, we delved into the potential pathways associated with each risk gene. Figures 6A and 6B present 39 pathways significantly correlated with these 5 genes, including critical pathways such as angiogenesis, apical junction, and apoptosis, among others.

The correlation between hub genes and immune response

Our study unveiled notable associations among the risk genes and immune-related scores. Specifically, SERPINE1, LINC02408, and GPX3 showed significant positive correlations with stromal score, immune score, and estimate score, while MAT3 and OLFM3 exhibited no significant associations with the immune score (Figure 7A). In addition, we divided the samples into high and low expression groups based on the median expression value of each gene. Comparing the immune score between these groups, we found that the high expression group of SERPINE1, OLFM3, LINC02408, and GPX3 displayed significantly higher scores compared to the low expression group (Figure 7B). Further correlation analysis revealed significant negative associations of the 5 genes with T cells follicular helper and NK cells resting. Moreover, except for OLFM3, the other four genes displayed positive correlations with Macrophages M2 (Figure 7C). Additionally, significant differences in various immune cells were observed between the high and low expression groups of the risk genes (Figure 7D). These findings highlight the intricate interplay between the risk genes and the immune microenvironment in GC.

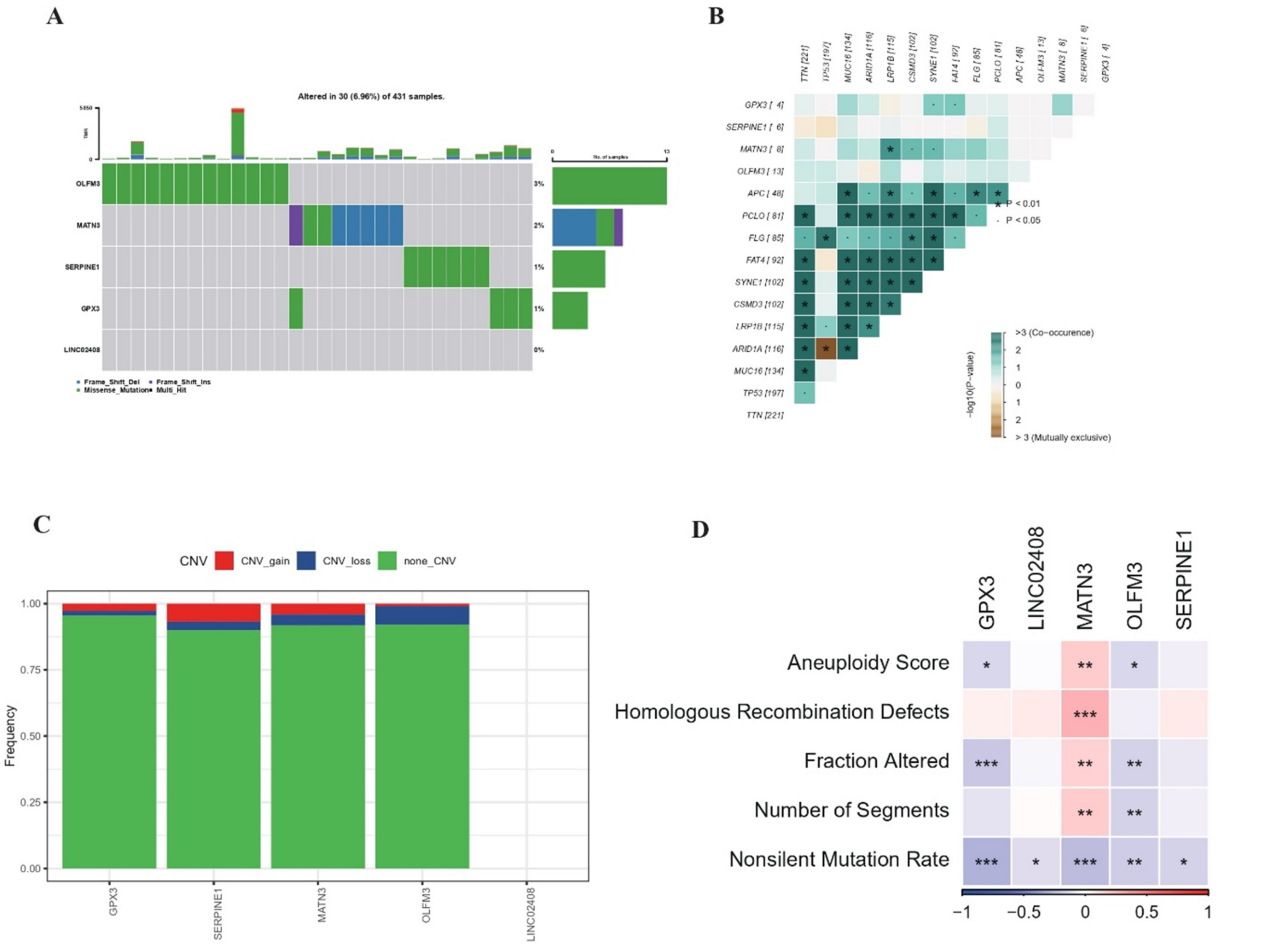

Figure 5. *Mutational analysis of risk signature genes: (A) Waterfall plot showing SNV mutations in the 5 key genes; (B) Analysis of colinearity and mutual exclusion with the 10 most mutated genes; (C) Evaluation of CNV mutations (gain, loss, none) in the 5 key genes; (D) Correlation heatmap of mutations in the 5 key genes with genomic characteristics.*

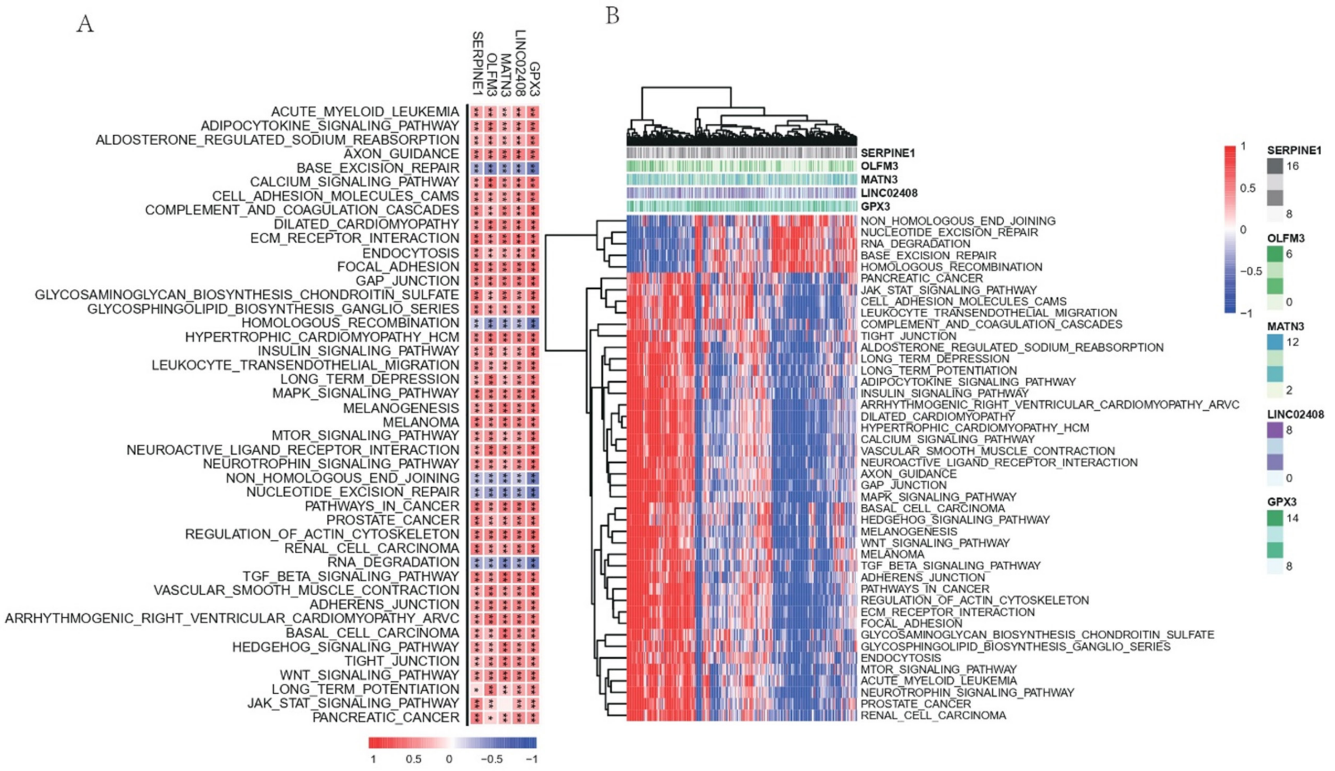

Figure 6. *Pathway analysis of risk genes: (A) Gene-pathway correlation heatmap; (B) Enrichment score heatmap for key pathways.* ***Note:*** *Significance levels: *P<0.05, **P<0.01, ***P<0.001.*

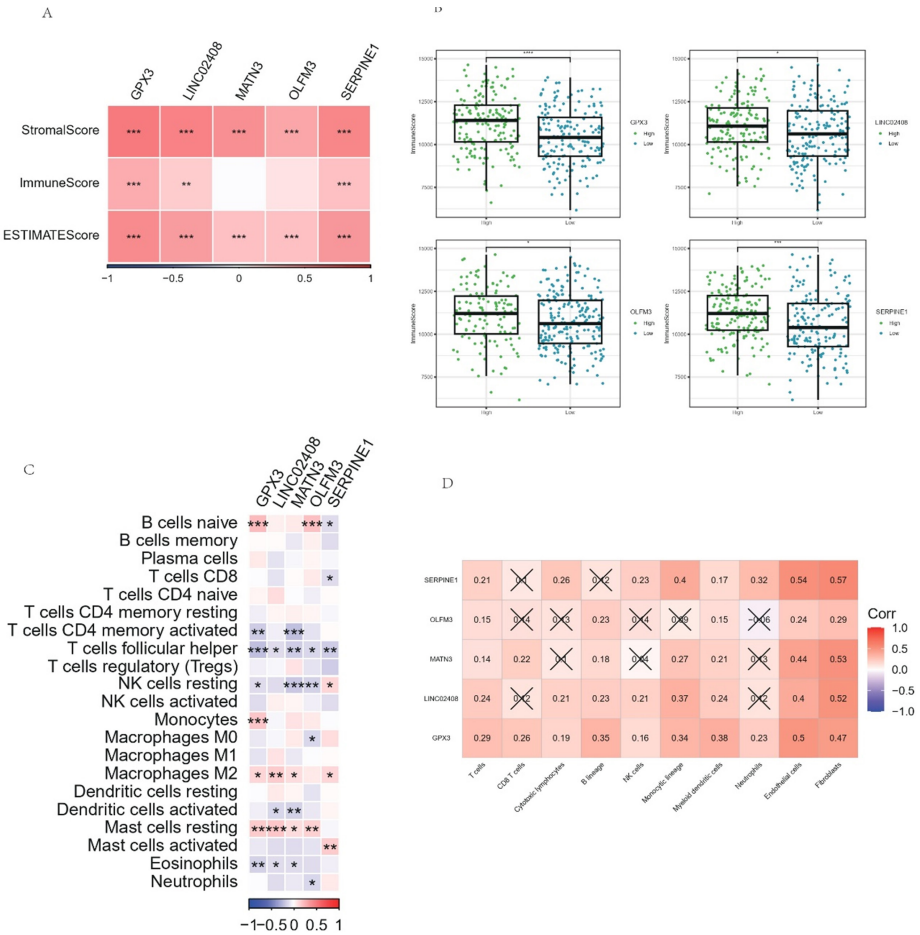

Figure 7. *Analysis of risk genes and the immune landscape: (A) Correlation matrix displaying relationships with stromal score, immune score, and estimate score; (B) Wilcoxon test comparison of immune scores between high and low expression groups of key genes; (C) Correlation analysis with immune cell scores predicted by CIBERSORT; (D) Comparison of 22 immune cell scores between high and low expression groups of key genes. Note: Statistical significance: *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.*

Prognostic value of risk signature in PD-L1 blockade immunotherapy response

The responsiveness of our risk signature to PD-L1 blockade immunotherapy was assessed using the IMvigor210 and GSE78220 cohorts, focusing on T-cell immunotherapy combined with immune checkpoint inhibitors [23]. Within the IMvigor210 cohort of 348 patients receiving anti-PD-L1 receptor blockers, varying responses were observed, including Complete Response (CR), Partial Response (PR), Stable Disease (SD), and Progressive Disease (PD). Notably, patients with SD/PD demonstrated higher risk scores compared to those with CR/PR (Figure 8A). Furthermore, the high-risk group exhibited a higher proportion of SD/PD patients compared to the low-risk group (Figures 8B and 8C). Analysis of the IMvigor210 cohort unveiled significant clinical outcomes. Patients within the low-risk group experienced substantial benefits with extended overall survival compared to their high-risk counterparts (Figure 8D, p=0.29). Particularly in Stage I+II patients, distinct survival disparities were evident across risk groups (Figure 8D, p=0.29), highlighting the risk score's sensitivity in this early-stage cohort. However, in Stage III+IV patients, no significant differences were observed (Figure 8E, p=0.047), indicating a potential limitation of the risk score's utility in advanced-stage settings. These findings underscore the differential impact of the risk score based on disease stage, emphasizing its potential as a valuable prognostic indicator, especially in early-stage gastric cancer patients.

Prognostic integration and nomogram development

To enhance the predictive accuracy of our risk signature, we conducted univariate and multivariate Cox regression analyses to integrate clinicopathological characteristics with the risk score. The multivariate analysis identified the risk signature as the most significant independent prognostic factor for osteosarcoma, with a Hazard Ratio (HR) of 2.819 (95% Confidence Interval [CI]: 1.893-4.198, p<0.001). Additionally, metastatic status emerged as the second most significant independent prognostic factor, with an HR of 2.718 (95% CI: 1.874-3.942, P=0.002) (Figures 9A and 9B). Utilizing these insights, we developed a nomogram that integrates the stage and risk score, as illustrated in Figure 9C. The calibration plot demonstrated the nomogram's effectiveness in accurately predicting survival outcomes (Figure 9D). Furthermore, Decision Curve Analysis (DCA) indicated the superior discriminative ability of the nomogram in identifying high-risk patients compared to using the risk score or stage alone (Figure 9E). Time-dependent Receiver Operating Characteristic (TimeROC) analysis revealed that the Area Under the Curve (AUC) of the risk score and nomogram surpassed that of other indicators in the TCGA cohort (Figure 9F).

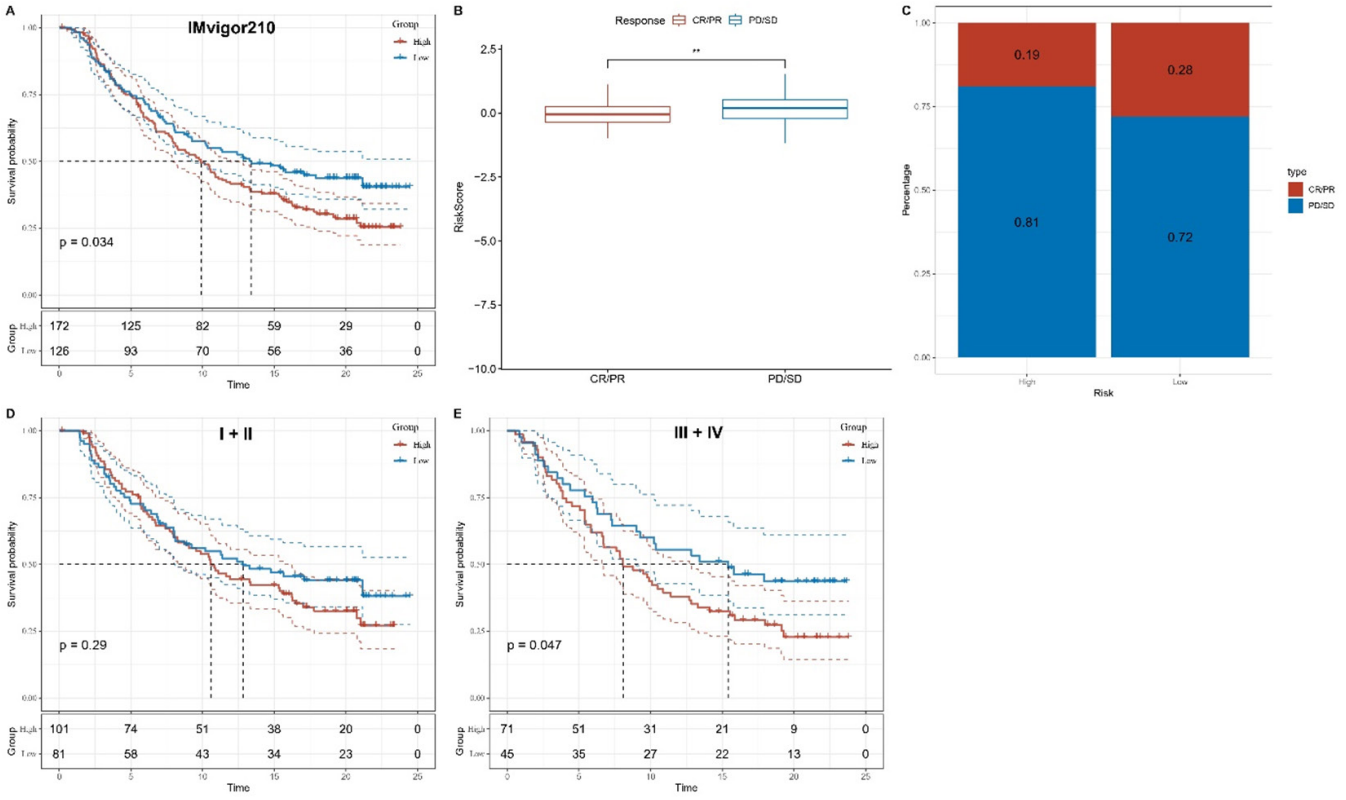

Figure 8. *Responsiveness of risk score to PD-L1 blockade immunotherapy in the IMvigor210 cohort: (A) Differences in risk scores across immunotherapy responses; (B) Distribution of immunotherapy responses among risk score groups; (C) Prognostic differences among risk score groups; (D) Prognostic disparities in early-stage patients; (E) Prognostic variations in advanced-stage patients. Note: Significance: ****P<0.0001.*

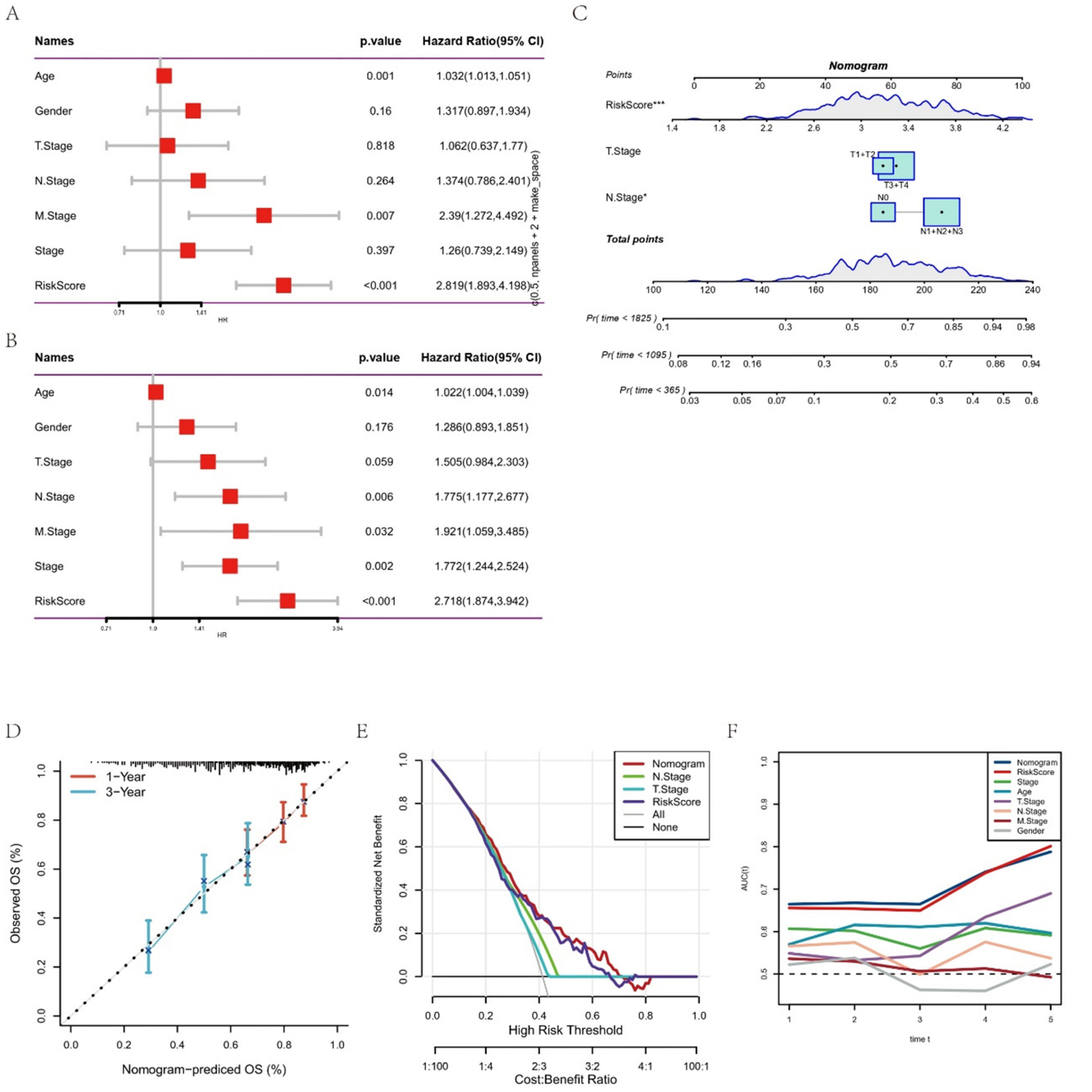

Figure 9. *Development of a prognostic nomogram for GC prognosis: (A,B) Univariate and multivariate Cox analyses with risk score and clinicopathological characteristics; (C) Nomogram model integrating risk score and stage; (D) Calibration curves for 1, 3, and 5-year predictions; (E) Decision curve analysis for the nomogram; (F) Time-ROC analysis comparing predictive capacity with clinicopathological features.*
Note: *Significance: ***P<0.001.*

Discussion

The intricate interplay between tumor cells and stromal components, especially CAFs, has garnered substantial attention for its pivotal role in driving tumor progression [24]. CAFs contribute significantly to tumor proliferation, angiogenesis, metastasis, and chemotherapy resistance by releasing diverse factors into the TME [25]. Our study focused on unravelling the heterogeneity of CAFs in GC through a comprehensive analysis of scRNA-seq data.

Our analysis identified six unique CAF clusters, each with distinct properties likely governing various aspects of TME biology. Notably, five of these CAF clusters demonstrated significant associations with GC prognosis, as evidenced by a score derived from differentially expressed genes across the clusters. Our exploration of pathways revealed notable differences, particularly highlighting the aberrant activation of crucial pathways such as the Hippo signaling pathway, known for its profound influence on gastric cancer development [26]. Additionally, the MYC-mediated axis, pivotal in GC proliferation, migration, invasion, and drug resistance, was prominently featured [27]. Expanding

on the prognostic potential of these CAF clusters, we formulated a CAF-based risk signature incorporating five genes: SERPINE1, MATN3, OLFM3, LINC02408, and GPX3. Our analysis uncovered Single Nucleotide Variant (SNV) mutations in some of these genes, potentially influencing GC progression [28-30]. Although significant co-occurrence among these mutations was not observed, our findings suggest their plausible roles in advancing GC. Notably, recent studies have developed gene signatures based on genomic instability, with SERPINE1 among those predicting GC prognosis [31]. Further correlation analyses with 39 pathways unveiled distinct signatures for protective and risk genes. Protective genes were positively associated with pathways such as allograft rejection and myogenesis, while risk genes correlated with fatty acid metabolism, xenobiotic metabolism, and adipogenesis [32,33]. Alterations in fatty acid metabolism have been implicated in GC pathogenesis, with demonstrated prognostic value of fatty acid metabolism-related genes [34-36]. Similarly, polymorphisms in xenobiotic metabolism-related genes have been linked to increased GC risk [37,38]. The role of adipogenesis, often associated with obesity, has also been highlighted in various cancers [39,40]. These insights pave the way for deeper investigations into the regulation of these risk genes in GC.

Recent evidence has emphasized the intricate interplay between Cancer-associated Fibroblasts (CAFs) and the Tumor Immune Microenvironment (TIME) in tumor progression [41-43]. In our study, three predictive genes displayed significant positive correlations with the immune score, suggesting potential crosstalk with the TIME in GC and positioning these genes as promising therapeutic targets. The TIME, housing a diverse array of immune cells within the TME, collectively shapes the anti-tumor immune response. CAFs can modulate these immune cells, creating an immunosuppressive TME that aids tumor cells in evading immune surveillance [44]. Notably, our risk signature showcased a negative association between the predictive genes and various T cell subsets, pivotal players in tumor progression. The therapeutic potential of T cell-based therapies, including checkpoint blockade and Chimeric Antigen Receptor T (CAR-T) cell therapy, has been well-established [45].

However, the widespread challenge of immunotherapy resistance necessitates precise patient stratification [46]. Our study suggests that the risk signature identified herein effectively stratifies patients likely to benefit from immunotherapies. Prior research has reported the role of endosialin, expressed by CAFs, in regulating macrophage recruitment and polarization in GC [47]. Within our signature, risk genes exhibited positive correlations with M2 macrophages and negative correlations with M0 macrophages, indicating potential involvement in macrophage polarization and, consequently, the pro-tumorigenic milieu. Notably, these genes also displayed negative correlations with helper T cells, further underscoring their potential pro-tumorigenic roles. Additionally, CAFs in GC have been reported to upregulate PD-L1 expression through the IL-8-mediated mechanism, thus posing as a potential therapeutic target to overcome immune resistance [48]. Intriguingly, our data also suggests that the CAF-based signature could predict responsiveness to anti-PD-L1 immunotherapy, offering novel insights into the CAF-mediated reshaping of the cancer microenvironment and immune status within the TME of GC [49,50].

In summary, our study illuminates the potential prognostic significance of the CAF-based risk signature in gastric cancer. However, a nuanced interpretation of our findings necessitates acknowledging certain limitations. The retrospective nature of our data drawn from public databases underscores the need for rigorous validation through large-scale prospective studies across diverse patient cohorts and multiple research centers. Furthermore, our focus on prognostic assessment underscores the need for deeper investigations into the mechanisms by which the CAF-based signature contributes to gastric cancer development. Addressing these gaps will refine the reliability of our findings, fostering a more comprehensive understanding of the intricate interplay between cancer-associated fibroblasts and the pathogenesis of gastric cancer.

Conclusion

In conclusion, we conducted a comprehensive characterization of Cancer-associated Fibroblasts (CAFs) in Gastric Cancer (GC) and identified six distinct CAF clusters with unique properties. Our analysis revealed differential gene expression patterns, highlighting enrichment in signaling pathways such as vascular smooth muscle contraction, focal adhesion, oxytocin, and PPARG signaling. Notably, five of these CAF clusters exhibited significant associations with GC prognosis, leading to the establishment of a CAF-based prognostic risk signature comprising five genes. Importantly, this signature showed strong correlations with the immune landscape, suggesting its potential for predicting responses to PD-L1 blockade immunotherapy. Additionally, we developed a novel nomogram integrating the risk signature with clinicopathological features, offering a reliable tool for predicting clinical outcomes in GC patients. This study sheds light on the intricate role of CAFs in GC pathogenesis and highlights the potential clinical utility of CAF-based signatures in personalized treatment strategies for GC.

Acknowledgment

Not applicable.

Conflict of Interest

The authors declare that they have no conflict of interest.

Consent to Publication

Not applicable.

Ethical Approval and Consent to Participate

Not applicable.

Funding

Not applicable.

Author Contribution

Xiaofei Sun: Conceptualization, Methodology; Song Gao: Data curation, Writing and original draft; Lili Hu: Investigation, Visualization; Yanli Liu: Supervision, Writing-review and editing.

Data Availability

Our data involved in this study were obtained from the following sources: TCGA STAD cohort (Cancer Genome Atlas Research) and Jeong *et al.* (GEO: GSE167297).

References

1. Thrift AP, Wenker TN, El-Serag HB. Global burden of gastric cancer: Epidemiological trends, risk factors, screening and prevention. Nat Rev Clin Oncol 2023; 20: 338-349.
2. Alsina M, Arrazubi V, Diez M, Tabernero J. Current developments in gastric cancer: From molecular profiling to treatment strategy. Nat Rev Gastroenterol Hepatol 2023; 20: 155-170.
3. Bessède E, Mégraud F. Microbiota and gastric cancer. Semin Cancer Biol 2022; 86: 11-17.
4. Yang L, Ying X, Liu S, Lyu G, Xu Z, Zhang X, Li H, Li Q, Wang N, Ji J. Gastric cancer: Epidemiology, risk factors and prevention strategies. Chin J Cancer Res 2020; 32: 695-704.
5. Wu T, Dai Y. Tumor microenvironment and therapeutic response. Cancer Lett 2017; 387: 61-68.
6. Nombela-Arrieta C, Isringhausen S. The role of the bone marrow stromal compartment in the hematopoietic response to microbial infections. Front Immunol 2016; 7: 689.
7. Chen X, Song E. Turning foes to friends: Targeting cancer-associated fibroblasts. Nat Rev Drug Discov 2019; 18: 99-115.
8. Bartoschek M, Oskolkov N, Bocci M, Lövrot J, Larsson C, Sommarin M, Madsen CD, Lindgren D, Pekar G, Karlsson G, Ringnér M. Spatially and functionally distinct subclasses of breast cancer-associated fibroblasts revealed by single cell RNA sequencing. Nat Commun 2018; 9: 5150.
9. Liao Z, Tan ZW, Zhu P, Tan NS. Cancer-associated fibroblasts in tumor microenvironment-Accomplices in tumor malignancy. Cell Immunol 2019; 343: 103729.
10. Qi R, Bai Y, Li K, Liu N, Xu Y, Dal E, Wang Y, Lin R, Wang H, Liu Z, Li X. Cancer-associated fibroblasts suppress ferroptosis and induce gemcitabine resistance in pancreatic cancer cells by secreting exosome-derived ACSL4-targeting miRNAs. Drug Resist Updat 2023; 68: 100960.
11. Mak TK, Li X, Huang H, Wu K, Huang Z, He Y, Zhang C. The cancer-associated fibroblast-related signature predicts prognosis and indicates immune microenvironment infiltration in gastric cancer. Front Immunol 2022; 13: 951214.
12. Fiori ME, Di Franco S, Villanova L, Bianca P, Stassi G, De Maria R. Cancer-associated fibroblasts as abettors of tumor progression at the crossroads of EMT and therapy resistance. Mol Cancer 2019; 18: 70.
13. Li C, Teixeira AF, Zhu HJ, Ten Dijke P. Cancer associated-fibroblast-derived exosomes in cancer progression. Mol Cancer 2021; 20: 154.
14. Ma M, Sun J, Liu Z, Ouyang S, Zhang Z, Zeng Z, Li J, Kang W. The immune microenvironment in gastric cancer: Prognostic prediction. Front Oncol 2022; 12: 836389.
15. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, Dimitriadoy S, Liu DL, Kantheti HS, Saghafinia S, Chakravarty D. Oncogenic signaling pathways in the cancer genome atlas. Cell 2018; 173: 321-337.
16. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 2018; 36: 411-420.
17. Yu G, Wang LG, Han Y, He QY. clusterProfiler: An R package for comparing biological themes among gene clusters. OMICS 2012; 16: 284-287.
18. Gao R, Bai S, Henderson YC, Lin Y, Schalck A, Yan Y, Kumar T, Hu M, Sei E, Davis A, Wang F. Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. Nat Biotechnol 2021; 39: 599-608.
19. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. Methods Mol Biol 2018; 1711: 243-259.
20. Zhang S, Tong YX, Zhang XH, Zhang YJ, Xu XS, Xiao AT, Chao TF, Gong JP. A novel and validated nomogram to predict overall survival for gastric neuroendocrine neoplasms. J Cancer 2019; 10: 5944-5954.
21. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, Kadel Iii EE, Koeppen H, Astarita JL, Cubas R, Jhunjhunwala S. TGFbeta attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature 2018; 554: 544-548.
22. Hugo W, Zaretsky JM, Sun LU, Song C, Moreno BH, Hu-Lieskovan S, Berent-Maoz B, Pang J, Chmielowski B, Cherry G, Seja E. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. Cell 2016; 165: 35-44.
23. Waldman AD, Fritz JM, Lenardo MJ. A guide to cancer immunotherapy: from T cell basic science to clinical practice. Nat Rev Immunol 2020; 20: 651-668.
24. Cha YJ, Koo JS. Roles of omental and bone marrow adipocytes in tumor biology. Adipocyte 2019; 8: 304-317.
25. Altorki NK, Markowitz GJ, Gao D, Port JL, Saxena A, Stiles B, McGraw T, Mittal V. The lung microenvironment: an important regulator of tumour growth and metastasis. Nat Rev Cancer 2019; 19: 9-31.
26. Liu X, Wang Y, Chen B, Chan WN, Mui CW, Cheung AH, Zhang J, Wong KY, Yu J, Kang W, To KF. Targeting the hippo pathway in gastric cancer and other malignancies in the digestive system: From bench to bedside. Biomedicines 2022; 10: 2512.
27. Dong Y, Li X, Lin Z, Zou W, Liu Y, Qian H, Jia J. HOXC-AS1-MYC regulatory loop contributes to the growth and metastasis in gastric cancer. J Exp Clin Cancer Res 2019; 38: 502.
28. Zhu Z, Qin J, Dong C, Yang J, Yang M, Tian J, Zhong X. Identification of four gastric cancer subtypes based on genetic analysis of cholesterogenic and glycolytic pathways. Bioengineered 2021; 12: 4780-4793.

29. Li L, Xing R, Cui J, Li W, Lu Y. Investigation of frequent somatic mutations of MTND5 gene in gastric cancer cell lines and tissues. Mitochondrial DNA B Resour 2018; 3: 1002-1008.
30. Chen H, Wang J, Zhuang Y, Wu H. Identification of the potential molecular mechanism and driving mutations in the pathogenesis of familial intestinal gastric cancer by whole exome sequencing. Oncol Rep 2018; 40: 2316-2324.
31. Xu X, Zhang L, Qian Y, Fang Q, Xiao Y, Chen G, Cai G, Abula A, Wang Z, Zhai E, Chen J. SERPINE1-based immune gene signature predicts prognosis and immunotherapy response in gastric cancer. Pharmaceuticals (Basel) 2022; 15: 1401.
32. Yao Q, Wang C, Wang Y, Xiang W, Chen Y, Zhou Q, Chen J, Jiang H, Chen D. STXBP3 and GOT2 predict immunological activity in acute allograft rejection. Front Immunol 2022; 13: 1025681.
33. Lin-Wang HT, Cipullo R, Dinkhuysen JJ, Finger MA, Rossi JM, Correia EB, Hirata MH. Down regulation of protective genes is associated with cellular and antibody-mediated rejection. Clin Transplant 2017; 31: 1-10.
34. Xu W, Ding H, Zhang M, Liu L, Yin M, Weng Z, Xu C. The prognostic role of fatty acid metabolism-related genes in patients with gastric cancer. Transl Cancer Res 2022; 11: 3593-3609.
35. Zhang Y, Liu W, Feng W, Wang X, Lei T, Chen Z, Song W. Identification of 14 differentially-expressed metabolism-related genes as potential targets of gastric cancer by integrated proteomics and transcriptomics. Front Cell Dev Biol 2022; 10: 816249.
36. Oh HR, An CH, Yoo NJ, Lee SH. Somatic mutations of amino acid metabolism-related genes in gastric and colorectal cancers and their regional heterogeneity--a short report. Cell Oncol (Dordr) 2014; 37: 455-461.
37. Yu G, Liang B, Yin K, Zhan M, Gu X, Wang J, Song S, Liu Y, Yang Q, Ji T, Xu B. Identification of metabolism-related gene-based subgroup in prostate cancer. Front Oncol 2022; 12: 909066.
38. Ferreira GD, Fernandes GM, Penteado C, Coria VR, Galbiatti-Dias AL, Russo A, Castanhole-Nunes MM, Silva RF, Silva RD, Pavarino EC, Torreglosa Ruiz Cintra M. Polymorphisms in xenobiotic metabolism-related genes in patients with hepatocellular carcinoma: A case-control study. Xenobiotica 2021; 51: 737-744.
39. Oshi M, Tokumaru Y, Angarita FA, Lee L, Yan L, Matsuyama R, Endo I, Takabe K. Adipogenesis in triple-negative breast cancer is associated with unfavorable tumor immune microenvironment and with worse survival. Sci Rep 2021; 11: 12541.
40. Nguyen TD, Miyatake Y, Yoshida T, Kawahara H, Hanayama R. Tumor-secreted proliferin-1 regulates adipogenesis and lipolysis in cachexia. Int J Cancer 2021; 148: 1982-1992.
41. Kwon JT, Bryant RJ, Parkes EE. The tumor microenvironment and immune responses in prostate cancer patients. Endocr Relat Cancer 2021; 28: T95-T107.
42. Melaiu O, Lucarini V, Cifaldi L, Fruci D. Influence of the tumor microenvironment on NK cell function in solid tumors. Front Immunol 2019; 10: 3038.
43. Sokratous G, Polyzoidis S, Ashkan K. Immune infiltration of tumor microenvironment following immunotherapy for glioblastoma multiforme. Hum Vaccin Immunother 2017; 13: 2575-2582.
44. Mao X, Xu J, Wang W, Liang C, Hua J, Liu J, Zhang B, Meng Q, Yu X, Shi S. Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: New findings and future perspectives. Mol Cancer 2021; 20: 131.
45. Kudo M. Immuno-oncology therapy for hepatocellular carcinoma: Current status and ongoing trials. Liver Cancer 2019; 8: 221-238.
46. Desbois M, Wang Y. Cancer-associated fibroblasts: Key players in shaping the tumor immune microenvironment. Immunol Rev 2021; 302: 241-258.
47. Yang F, Wei Y, Han D, Li Y, Shi S, Jiao D, Wu J, Zhang Q, Shi C, Yang L, Song W. Interaction with CD68 and regulation of GAS6 expression by endosialin in fibroblasts drives recruitment and polarization of macrophages in hepatocellular carcinoma. Cancer Res 2020; 80: 3892-3905.
48. Lou M, Iwatsuki M, Wu X, Zhang W, Matsumoto C, Baba H. Cancer-associated fibroblast-derived IL-8 upregulates PD-L1 expression in gastric cancer through the NF-kappaB pathway. Ann Surg Oncol 2023; 1-3.
49. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature 2013; 499: 43-49.
50. Jeong HY, Ham IH, Lee SH, Ryu D, Son SY, Han SU, Kim TM, Hur H. Spatially distinct reprogramming of the tumor microenvironment based on tumor invasion in diffuse-type gastric cancers. Clin Cancer Res 2021; 27; 6529-6542.

***Correspondence to:**

Yanli Liu

Dongying Traditional Chinese Medicine Hospital

Dongying

Shandong 257099

China