

Multi-Omics Dissection of Gastric Cancer Stroma Identifies MMP1+ Cancer-Associated Fibroblasts as Drivers of Tumor Angiogenesis via GDF15–TGFR2 Axis

Zeyu Li

The First Affiliated Hospital of China Medical University

Na Han

The First Affiliated Hospital of China Medical University

Hanwen Liu

The First Affiliated Hospital of China Medical University

Ruiying Wang

The First Affiliated Hospital of China Medical University

Mohan zhang

The First Affiliated Hospital of China Medical University

Xuancheng Lin

The First Affiliated Hospital of China Medical University

Chi Xue

The First Affiliated Hospital of China Medical University

Fanghao Xiao

The First Affiliated Hospital of China Medical University

HIDASA IKO

The First Affiliated Hospital of China Medical University

Xiaolu Wang

The First Affiliated Hospital of China Medical University

Rui Guo

Cancer Hospital of Dalian University of Technology, Liaoning Cancer Hospital & Institute

Zhi Zhu

The First Affiliated Hospital of China Medical University

Ziming Gao

The First Affiliated Hospital of China Medical University

Kai Li

kli@cmu.edu.cn

Research Article

Keywords: Gastric cancer, Single cell RNA sequencing, Angiogenesis, Cancer-associated fibroblasts

Posted Date: October 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11146303/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Gastric cancer (GC) progression is driven not only by malignant epithelial cells but also by the surrounding stromal microenvironment, in which cancer-associated fibroblasts (CAFs) contribute to angiogenesis and immune evasion. However, the specific CAF states and the intercellular signaling through which they remodel this microenvironment remain incompletely characterized. Here, we integrated single-cell and spatial transcriptomics of an in-house GC cohort with public bulk, immunotherapy, and pan-cancer datasets, together with functional assays, to dissect CAF heterogeneity and its role in GC. Among nine fibroblast subpopulations, we identified an MMP1⁺ CAF subset that was enriched in tumor tissue, accumulated with tumor stage, and was associated with poor prognosis. MMP1⁺ CAFs occupied the initiation point of the fibroblast pseudotime trajectory and exhibited a stem-like phenotype enriched in extracellular matrix remodeling, epithelial–mesenchymal transition, and angiogenesis programs. Mechanistically, MMP1⁺ CAFs signaled to immature vascular endothelial cells through the GDF15–TGFB2 axis, activating angiogenic pathways and the downstream transcription factor TCF12 to promote neovascularization; this effect was abrogated in vitro by the GDF15 inhibitor Pongromab. In parallel, MMP1⁺ CAFs engaged tumor cells to upregulate VEGFA, which in turn acted on endothelial cells via VEGFR1/VEGFR2. Beyond angiogenesis, MMP1⁺ CAFs interacted with FOLR2⁺ M2-like macrophages to establish an immunosuppressive microenvironment and to predict resistance to immunotherapy. Finally, pan-cancer analysis revealed the broad conservation of MMP1⁺ CAFs and their generally adverse prognostic association across colorectal, lung, hepatocellular, and pancreatic cancers. Together, these findings define MMP1⁺ CAFs as a stromal state coordinating angiogenesis in gastric cancer and highlight the GDF15–TGFB2 axis as a candidate therapeutic target.

1. Introduction

Gastric cancer (GC) ranks as the fifth most commonly diagnosed malignancy and the fifth leading cause of cancer-related mortality worldwide, accounting for approximately 0.97 million new cases and 0.66 million deaths in 2022^[1, 2]. Despite recent advances in diagnosis and treatment, the prognosis of gastric cancer remains poor, with most patients diagnosed at an advanced stage and deriving limited benefit from conventional therapies that primarily target tumor cell proliferation^[3]. The development and progression of gastric cancer is a complex, multi-stage process increasingly recognized as being driven not only by malignant epithelial cells but also by the surrounding tumor microenvironment (TME)^[4–6]. The TME comprises a variety of cellular components, including tumor cells, immune cells, endothelial cells, and fibroblasts, which interact through intricate signaling networks to collectively regulate tumor growth, invasion, metastasis and therapeutic resistance.

Tumor angiogenesis plays a pivotal role in the progression of gastric cancer within the tumor microenvironment (TME). This process not only provides the oxygen and nutrients essential for tumor growth but also establishes vascular routes that facilitate tumor cell dissemination and metastasis^[7]. Angiogenesis is a highly coordinated process involving extensive interactions among endothelial cells, pericytes, vascular smooth muscle cells, tumor cells, tumor-associated immune cells, and cancer-

associated fibroblasts^[8, 9]. Although anti-angiogenic agents such as VEGF/VEGFR-targeted therapies have been incorporated into the clinical management of gastric cancer, their long-term efficacy remains limited by tumor adaptive mechanisms, the activation of compensatory pro-angiogenic pathways, and the post-treatment emergence of more aggressive phenotypes^[10]. This underscores the necessity to move beyond blockade of a single angiogenic factor and to elucidate the cellular sources and regulatory mechanisms that sustain neovascularization within the gastric TME. Therefore, a thorough understanding of the regulatory mechanisms underlying tumor angiogenesis is essential for identifying novel therapeutic targets and improving the treatment effect in gastric cancer.

In this context, cancer-associated fibroblasts (CAFs) have attracted increasing attention as key stromal orchestrators of tumor angiogenesis. Comprising the primary stromal component of the TME, CAFs play essential roles in extracellular matrix (ECM) remodeling, metastatic progression, angiogenesis and therapeutic resistance^[11, 12]. CAFs are now recognized as a highly heterogeneous compartment encompassing multiple transcriptionally and functionally distinct subsets—including myofibroblastic (myCAF), inflammatory (iCAF), vascular (vCAF), and antigen-presenting (apCAF) states—that arise from diverse cellular origins and occupy discrete niches within the tumor^[13, 14]. Through remodeling ECM, secreting growth factors, cytokines, and chemokines, and engaging in reciprocal crosstalk with malignant epithelial, endothelial, and immune cells, CAFs actively promote various aspects of tumor progression. The abundance of CAFs correlates with advanced stage, chemoresistance, and unfavorable prognosis in gastric cancer^[15–17]. However, the contributions of specific CAF subpopulations and their effector molecules to tumor angiogenesis remain poorly understood.

In this study, by integrating single-cell and spatial transcriptomics with public cohorts and functional assays in gastric cancer (Fig. 1A), we identify a novel CAFs subpopulation, MMP1⁺ CAFs, which accumulates progressively during tumor progression and is significantly associated with poor prognosis. Trajectory and functional analyses reveal that MMP1⁺ CAFs occupy a low-differentiated, stem-like state and are enriched in ECM remodeling, epithelial-mesenchymal transition (EMT), and angiogenesis programs. Mechanistically, MMP1⁺ CAFs act on immature vascular endothelial cells through the GDF15-TGFR2 axis, activating endothelial angiogenic pathways and modulating key transcription factors, including TCF12, to promote directed endothelial migration and neovascularization; the GDF15 inhibitor Ponesimab abrogates this pro-angiogenic activity in vitro. In parallel, MMP1⁺ CAFs recruit and activate EMT tumor cells to upregulate VEGFA secretion, which further enhances the pro-angiogenic function of immature endothelial cells through VEGFR1/VEGFR2. Beyond angiogenesis, MMP1⁺ CAFs interact with FOLR2⁺ M2-type macrophages to establish an immunosuppressive microenvironment, and their abundance serves as an effective and reliable predictor of immunotherapy resistance. Pan-cancer analysis further reveals that MMP1⁺ CAFs are widely present across multiple cancer types, including colorectal cancer, lung adenocarcinoma, hepatocellular carcinoma, and pancreatic cancer, and are generally correlated with poor clinical outcomes. Our study highlights the heterogeneity of CAFs and

their dual role in driving angiogenesis and immune suppression in gastric cancer, offering a potential therapeutic target with broad applicability across cancers.

2. Results

2.1. Single-Cell RNA Sequencing Reveals Progressive Fibroblast Infiltration During Gastric Cancer Progression

To elucidate the cellular heterogeneity and molecular features of gastric cancer (GC), we performed single-cell RNA sequencing (scRNA-seq) on primary tumor tissues from eight GC patients (Fig. 1A, Table S1). After quality-control filtering, a total of 54,321 cells were retained for downstream analysis (Figure S1A-C). To eliminate batch effects among samples, the Harmony algorithm was applied for data integration. Following dimensionality reduction and clustering, nine major cell types were identified (Fig. 1B, Figure S1D-F), including B lymphocytes (BLC; *CD79A*, *MS4A1*), epithelial cells (*EPCAM*, *KRT18*), endothelial cells (*PECAM1*, *CLDN5*), fibroblasts (*DCN*, *LUM*), mast cells (*CPA3*, *TPSAB1*), myeloid cells (*CD14*, *CD68*), plasma cells (*MZB1*, *IGHA1*), cycling cells (*MKI67*, *TOP2A*), and T lymphocytes (TLC; *CD3D*, *CD3E*, *CD2*) (Fig. 1C). Differentially expressed genes among cell types were computed using FindAllMarkers, extracting the significantly up- and down-regulated genes for each cell type (Fig. 1D). Notably, fibroblasts and endothelial cells exhibited a significant positive correlation within the microenvironment, suggesting an important role of the stromal microenvironment in gastric cancer (Fig. 1E). Further quantification of intergroup cell proportions revealed that fibroblast infiltration increased with tumor stage (Fig. 1F). Immunohistochemistry confirmed the progressive accumulation of fibroblasts along with tumor progression (Fig. 1G,H).

Figure 1. Single-cell transcriptomic landscape of the gastric cancer tumor microenvironment. (A) Schematic of the study design, covering scRNA-seq analysis of an in-house GC cohort, validation in public bulk and spatial transcriptomic datasets, experimental validation (IHC, IF, in vitro assays), and pan-cancer analysis. (B) t-SNE plot showing nine major cell types across 54,321 single cells. (C) Heatmap illustrating the expression of top marker genes in the corresponding cell types, with color indicating the row-scaled Z-score. (D) Circos plot showing significantly up-regulated (red) and down-regulated (blue) differentially expressed genes for each cell type. (E) Correlation matrix depicting pairwise associations among the nine cell types; blue and red indicate negative and positive correlations, respectively, with fibroblasts and endothelial cells showing a significant positive correlation. (F) Stacked bar plot showing the proportion of major cell types across tumor stages, revealing increasing fibroblast infiltration with tumor progression. (G) Representative α -SMA IHC staining in GC tissues of increasing stage (II–IV). Scale bars, 300 μ m. (H) Quantification of the α -SMA-positive area across stages (II–IV). Data are mean $\pm$ SD; * $P < 0.05$, *** $P < 0.001$, **** $P < 0.0001$.

2.2. MMP1^+ CAFs Are Key Subpopulations of Cancer-Associated Fibroblasts Associated with Poor Prognosis

To further investigate the heterogeneity of fibroblasts within the gastric cancer tumor microenvironment, we isolated and reclustered the fibroblast population to identify pro-tumorigenic fibroblast subpopulations (Fig. 2A, Figure S2A-C). The nine fibroblast subclusters were annotated based on their signature genes (Fig. 2B, Table S2). We used the typical markers of myCAFs, iCAFs, vCAFs and apCAFs to score these nine subclusters, but most of them were not included in these four fibroblast subtypes (Figure S2D). GO enrichment of nine fibroblasts subclusters revealed that MMP1⁺ CAFs were mostly enriched in extracellular matrix and collagen fibril organization pathways, indicating its crucial role in stromal microenvironment organization (Figure S3). MMP1⁺ CAFs highly expressed marker genes such as *CST1*, *MMP1*, *MMP3*, *MMP11*, and *CXCL8*. Functional enrichment analysis using PROGENy revealed that MMP1⁺ CAFs significantly activated pro-tumorigenic functional phenotypes, including VEGF, Hypoxia, and TNF- α (Fig. 2C). Furthermore, GSVA and GSEA results showed that MMP1⁺ CAFs were enriched in pro-tumorigenic pathways such as epithelial-mesenchymal transition (EMT) and angiogenesis (Fig. 2D,E). Consistently, spatial transcriptomic data revealed that MMP1⁺ CAFs significantly co-localized with the angiogenesis functional score in space (Fig. 2F). These results suggest that MMP1⁺ CAFs may activate angiogenic functions in gastric cancer.

Deciphering the origin of MMP1⁺ CAFs, we performed pseudotime analysis on all fibroblast populations using Monocle2, Monocle3 combined with CytoTRACE. The results revealed that MMP1⁺ CAFs were located at the initiation point of differentiation, suggesting a stem-like phenotype (Fig. 2G,H, Figure S4A,B). Pseudotime trajectory analysis of MMP1⁺ CAFs showed significant enrichment of gene modules associated with ECM remodeling and cell-matrix adhesion (Figure S4C,D). To assess the clinical significance of MMP1⁺ CAFs, we combined the TCGA-STAD dataset and performed ssGSEA scoring to estimate the abundance of MMP1⁺ CAFs in tissues using top30 highly expressed genes. The results revealed that, compared with adjacent normal tissues, MMP1⁺ CAFs infiltrated more abundantly in tumors, and their infiltration increased within tumor stages (Fig. 2I). Survival analysis demonstrated that MMP1⁺ CAFs were associated with shorter survival, indicating that MMP1⁺ CAFs may promote tumor progression (Fig. 2J). In addition, immunofluorescence (IF) staining of tumor and adjacent normal tissue sections, using α -SMA as a pan-fibroblast marker and MMP1 as a specific marker, confirmed the predominant presence of MMP1⁺ CAFs within the tumor microenvironment, supporting their potential functional relevance (Fig. 2K). Collectively, these findings indicate that MMP1⁺ CAFs exert a tumor-promoting role in the gastric cancer tumor microenvironment, participating in tumor progression through the activation of angiogenic functions.

Figure 2. MMP1⁺ CAF is a key subpopulation of cancer-associated fibroblasts associated with poor prognosis. (A) t-SNE plot showing nine fibroblast subclusters (C1_MMP1_CAF–C9_SFRP2_CAF). (B) Dot plot illustrating the expression of signature genes across the nine fibroblast subclusters, with dot color indicating mean expression and dot size the fraction of cells. (C) PROGENy heatmaps showing pathway activity (left) and functional program scores (right) across fibroblast subclusters, with MMP1⁺ CAFs enriched in VEGF, Hypoxia, TNF α , angiogenesis, and epithelial-mesenchymal transition. (D, E) GSEA plots

showing enrichment of angiogenesis (D) and epithelial-mesenchymal transition (E) gene sets in MMP1⁺ CAFs. (F) Spatial transcriptomic feature plots showing the co-localization of the MMP1⁺ CAF signature with the angiogenesis score (red boxes). (G, H) Monocle2 pseudotime trajectory of fibroblasts colored by pseudotime (G) and by subcluster (H), placing MMP1⁺ CAFs at the initiation point of differentiation. (I) Box plot of the MMP1⁺ CAF signature score across tumor stages (I–IV) in TCGA-STAD (Kruskal-Wallis, $P = 0.043$). (J) Kaplan-Meier curves showing overall survival stratified by MMP1⁺ CAF signature (score-high vs. score-low; $P = 0.0086$). (K) Representative immunofluorescence staining of MMP1 (green) and α -SMA (red) with DAPI (blue) in normal and GC tissues, confirming MMP1⁺ CAF enrichment in the tumor microenvironment.

2.3. MMP1⁺ CAFs Interact with Vascular Endothelial Cells to Promote Tumor Angiogenesis

To elucidate the mechanism by which MMP1⁺ CAFs influence the gastric cancer tumor microenvironment (TME), we conducted a comprehensive cell-cell communication analysis across all cell types, revealing that the GDF signaling pathway was particularly active in MMP1⁺ CAFs (Fig. 3A). Among these, the GDF15-TGFBR2 axis was identified as the sole driver of this pathway activity (Fig. 3B), and we found that MMP1⁺ CAFs exerted the most significant effect on endothelial cells through the GDF15-TGFBR2 axis (Fig. 3C,D).

To further characterize the endothelial cell subpopulations, we isolated endothelial cells and reclustered them based on canonical and cell-specific marker genes, yielding five distinct subpopulations: arterial endothelial cells (*GJA5*, *TSPAN2*), venous endothelial cells (*ACKR1*, *CPE*, *POSTN*), lymphatic endothelial cells (*PROX1*, *LYVE1*, *PDPN*), cycling endothelial cells (*MKI67*, *TOP2A*), and immature endothelial cells (*HSPG2*, *VWA1*, *APLN*) (Fig. 3E). Among these, immature endothelial cells increased significantly with tumor stage progression (Fig. 3F). Notably, *TGFBR2* was significantly enriched in immature endothelial cells (Fig. 3G), suggesting that MMP1⁺ CAFs may preferentially act on this endothelial subpopulation. GO enrichment analysis revealed that immature endothelial cells were enriched in pathways such as tissue migration and adhesion, while GSVA and GSEA analyses revealed that immature endothelial cells activated angiogenesis-related pathways, including the TGF- β pathway, the complement pathway, and WNT signaling (Fig. 3H,I). Multiple bulk transcriptomic datasets confirmed a significant positive correlation between MMP1⁺ CAFs and immature endothelial cells (Figure S5A-F). Moreover, *GDF15* and *TGFBR2* expression co-localized in space, and MMP1⁺ CAFs also co-localized with immature endothelial cells (Figure S5G-J).

To explore the molecular interactions between MMP1⁺ CAFs and immature endothelial cells, we applied NicheNet analysis (Fig. 3H). The results showed that the potential target genes in immature endothelial cells were enriched in biological processes such as the regulation of angiogenesis, smooth muscle cell proliferation, and endothelial cell migration. These results are consistent with the functional characteristics of immature endothelial cells and further support their key role in angiogenesis (Fig. 3I).

In addition, we confirmed the co-localization of GDF15-TGFBR2 in gastric cancer by multiplex immunofluorescence, demonstrating the communication between MMP1⁺ CAFs and immature endothelial cells in gastric cancer (Fig. 3J). Collectively, these findings reveal that MMP1⁺ CAFs act on immature vascular endothelial cells through the GDF15-TGFBR2 axis, activating the angiogenic pathways of endothelial cells and promoting the neovascularization process in gastric cancer.

Figure 3. MMP1⁺ CAFs act on immature vascular endothelial cells through the GDF15-TGFBR2 axis.

(A) Dot plot showing the outgoing communication patterns of secreting cells across all cell types, with MMP1⁺ CAFs displaying broad signaling activity. (B) Bar plot showing the relative contribution of each ligand-receptor pair to the GDF signaling pathway, identifying GDF15-TGFBR2 as the sole driver. (C) Circle plot showing MMP1⁺ CAF-derived GDF15 signaling to other cell types, with the strongest effect on endothelial cells. (D) Heatmap of the GDF signaling pathway network showing the roles (sender, receiver, mediator, influencer) of each cell type. (E) t-SNE plot showing five endothelial cell subpopulations: arterial, venous, lymphatic, cycling, and immature endothelial cells. (F) Violin plot showing *TGFBR2* expression across endothelial subpopulations, with significant enrichment in immature endothelial cells (Kruskal-Wallis, $P = 2.3e - 08$). (G) GSEA bar plot showing up-regulated (orange) and down-regulated (blue) pathways in immature endothelial cells, including angiogenesis-related programs. (H) NicheNet heatmap showing the regulatory potential of prioritized MMP1⁺ CAF ligands on predicted target genes in immature endothelial cells. (I) GO enrichment of the predicted target genes, showing enrichment in regulation of angiogenesis, vasculature development, and endothelial cell migration. (J) Representative immunofluorescence staining of GDF15 (green) and TGFBR2 (red) with DAPI (blue), confirming GDF15-TGFBR2 co-localization in GC tissues.

2.4. MMP1⁺ CAFs Activate the Angiogenic Function of Endothelial Cells Through the GDF15-TGFBR2 Axis

There was evidence indicates that GDF15 is one of the key drivers of angiogenesis^[18, 19]. Given that immature vascular endothelial cells play a guiding role in angiogenesis through directed migration, and considering the established function of the GDF15-TGFBR2 axis in regulating cell motility and convergent extension^[20], we hypothesized that MMP1⁺ CAFs may promote the migration of immature endothelial cells through this signaling pathway. Immunofluorescence (IF) staining of tumor tissue sections showed that MMP1⁺ CAFs were spatially adjacent to endothelial cells, providing anatomical evidence for potential paracrine signaling (Fig. 4A). Further Transwell migration and tube formation assays demonstrated that GDF15 significantly enhanced the migratory and tube-forming capacities of human umbilical vein endothelial cells (HUVECs). Conversely, the GDF15 inhibitor Ponsegromab impaired both processes, underscoring the key role of the GDF15 signaling pathway in endothelial activation (Fig. 4B-E).

We further analyzed the transcription factors active in endothelial cells. SCENIC results showed that the activity of multiple transcription factors was significantly elevated in immature endothelial cells, with the

top five transcription factors being *RFX1*, *SP3*, *ZNF253*, *TCF12*, and *HIVEP3* (Fig. 4F). These transcription factors mainly regulated the specific signature genes of immature endothelial cells. Among them, *TCF12* regulated most of the specific signature genes, and has been reported in the literature to participate in the regulation of angiogenesis^[21, 22]. In addition, we found that *TCF12* regulated the myosin *MYO1D*, which plays an important role in cell migration (Fig. 4G). We therefore speculate that *TCF12* may play an important role in angiogenesis and endothelial cell migration. We confirmed that transcription factors such as *TCF12* exhibited the highest activity in immature endothelial cells (Fig. 4H, Figure S6A). TCGA-STAD data validated that the top five transcription factors were positively correlated with the GDF15-TGFBR2 axis, suggesting an upstream-downstream relationship between this axis and the transcription factors of immature endothelial cells (Fig. 4I, Figure S6B). Furthermore, we experimentally validated that GDF15 could activate the expression of the aforementioned transcription factors in endothelial cells, whereas the inhibitor reversed this process (Fig. 4J, Figure S6C).

Figure 4. Experimental validation that the GDF15-TGFBR2 axis activates the angiogenic function of endothelial cells. (A) Representative multiplex immunofluorescence of MMP1 (green), α -SMA (red), and CD34 (purple) with DAPI (blue) in GC tissues, showing the spatial adjacency of MMP1⁺ CAFs to endothelial cells. (B) Representative images of Transwell migration (top) and tube formation (bottom) assays of HUVECs treated with PBS, GDF15, GDF15 + Pongegromab, or Pongegromab. Scale bars, 300 μ m. (C–E) Quantification of migrated cell number (C), number of junctions (D), and number of meshes (E) across treatment groups. (F) SCENIC heatmap showing transcription factor regulon activity across endothelial subpopulations, with *TCF12* enriched in immature endothelial cells. (G) Chord diagram showing the regulatory relationships between transcription factors and the signature genes of immature endothelial cells. (H) Violin plot showing *TCF12* regulon activity across endothelial subpopulations, highest in immature endothelial cells. (I) Correlation between *TCF12* and *TGFBR2* expression in TCGA-STAD ($R = 0.39$, $P = 1.6e - 16$). (J) qRT-PCR quantification of *TCF12* expression in endothelial cells across treatment groups. Data are mean $\pm$ SD; ns, not significant; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

2.5. MMP1⁺ CAFs Promote VEGFA Release by Tumor Cells, Thereby Promoting Angiogenesis in Gastric Cancer

Abundant literature indicates that cancer-associated fibroblasts can promote the malignant transformation of tumor cells, leading to tumor progression and poor prognosis^[23]. To elucidate the role of MMP1⁺ CAFs in promoting the malignant phenotype of gastric cancer, we comprehensively evaluated their effects on tumor cells. Dimensionality reduction and subclustering of tumor cells yielded 15 cell clusters, named Epi 1 to Epi 15. Cell-cell communication analysis revealed that MMP1⁺ CAFs communicated most significantly with Epi 10 (Fig. 5A,B, Figure S7A,B), with the significant ligand-receptor pairs mainly concentrated in the Collagen pathway, including *COL1A1*, *COL1A2*, *COL6A1*, *COL6A2*, and *COL6A3*, and *SDC4* was the most significant tumor cell receptor (Figure S7C-E).

Subsequent gene-set scoring revealed that Epi 10 had the highest angiogenesis score, suggesting that it may mediate the angiogenic process (Fig. 5C), and GSEA enrichment likewise showed that Epi 10 significantly upregulated angiogenic functional pathways (Fig. 5D). We assessed the malignancy of tumor cells using inferCNV and found that Epi 10 had the highest CNV score, indicating the highest degree of malignancy and validating it as the tumor cell subpopulation involved in gastric cancer progression (Fig. 5E,F). To further explore the molecular mechanism by which MMP1⁺ CAFs act on the tumor cell subpopulation Epi 10, NicheNet results showed that MMP1⁺ CAFs activated the target gene *VEGFA* in Epi 10 (Fig. 5G). GO enrichment analysis showed that the target genes of Epi 10 were enriched in biological processes related to the cell cycle and proliferation, suggesting their role in promoting tumor cell proliferation. *VEGFA* is a classic marker gene of angiogenesis, and abundant literature has reported its promoting effect on the tube formation of vascular endothelial cells and the underlying mechanisms^[24]. We found that *VEGFA* was specifically highly expressed in Epi 10 (Fig. 5H,I). This led us to hypothesize that MMP1⁺ CAFs promote the production of *VEGFA* by Epi 10 tumor cells through these signaling pathways, thereby promoting angiogenesis.

Further communication analysis between tumor cells and endothelial cells showed that *VEGFA* signaling originated primarily from Epi 10 tumor cells and acted on immature endothelial cells and arterial endothelial cells (Fig. 5J). This interaction was mediated by *VEGFR1* and *VEGFR2*, which regulate the function of endothelial cells (Fig. 5K). Correlation analysis indicated that *VEGFA* signaling was positively correlated with the expression of *TCF12* (Fig. 5L,M). In summary, we propose that MMP1⁺ CAFs recruit and activate Epi 10 tumor cells to upregulate the production of *VEGFA*, thereby influencing immature endothelial cells and promoting angiogenesis.

Figure 5. The interaction mechanism by which MMP1⁺ CAFs promote *VEGFA* release by tumor cells.

(A) t-SNE plot showing 15 tumor cell subclusters (Epi 1–Epi 15). (B) Heatmap showing the interaction strength between MMP1⁺ CAFs and each tumor cell subcluster, with the strongest communication observed with Epi 10. (C) Box plot of the angiogenesis score across tumor cell subclusters, highest in Epi 10. (D) GSEA bar plot showing up-regulated (orange) and down-regulated (blue) pathways in Epi 10, including angiogenesis. (E) inferCNV heatmap showing large-scale copy number variations across tumor cell subclusters, with T cells as the reference. (F) Box plot of the CNV score across tumor cell subclusters, with Epi 10 exhibiting the highest CNV burden. (G) NicheNet heatmap showing the regulatory potential of MMP1⁺ CAF ligands on predicted target genes in Epi 10, including *VEGFA*. (H) Feature plot showing *VEGFA* expression across tumor cell subclusters. (I) Dot plot showing *VEGFA* expression across tumor cell subclusters, specifically enriched in Epi 10. (J) Circle plot showing *VEGFA* signaling from Epi 10 tumor cells to endothelial cell subpopulations. (K) Chord diagram of the *VEGFA* signaling network, showing Epi 10 as the primary source acting on immature and arterial endothelial cells. (L, M) Correlation of *TCF12* expression with *VEGFR1* (L; $R = 0.33$, $P = 7.7e - 12$) and *VEGFR2* (M; $R = 0.2$, $P = 3.6e - 05$) in TCGA-STAD.

2.6. MMP1⁺ CAFs Mediate Immunosuppression and Can Serve as a Predictive Indicator of Immunotherapy Resistance

To investigate the association between MMP1⁺ CAFs and the tumor immune microenvironment, we evaluated the differences in immune cell infiltration between the MMP1⁺ CAF high and low-infiltration groups based on the ssGSEA and CIBERSORT algorithms. The results showed that in gastric cancer patients with high MMP1⁺ CAF infiltration, immunosuppressive cells such as M2 macrophages were more abundant, whereas memory B cells and naive T cells were less abundant. This suggests that MMP1⁺ CAFs may shape the immunosuppressive microenvironment of gastric cancer through their interaction with M2 macrophages (Fig. 6A).

Myeloid cells were extracted, and using differentially expressed genes and signature marker genes, myeloid cells were classified into five populations: monocytes (*VCAN*, *FCN1*), macrophages (*C1QA*, *C1QB*), neutrophils (*FCGR3B*, *CXCR2*), cDC2 (*CD1C*, *FCER1A*), cDC3 (*CCR7*, *LAMP3*), and pDC (*GZMB*, *CLIC3*) (Figure S8A,B). The macrophages were further subclustered into IL4I1⁺ Macro, IGLC2⁺ Macro, FOLR2⁺ Macro, JUNB⁺ Macro, and the proliferative MKI67⁺ Macro (Fig. 6B, Figure S8C-E). Cell-cell communication analysis revealed that MMP1⁺ CAFs exhibited the strongest communication with FOLR2⁺ Macro (Fig. 6C), and transcriptomic correlation analysis showed that FOLR2⁺ Macro had the highest correlation coefficient with MMP1⁺ CAFs (Figure S8F). Gene-set scoring revealed that FOLR2⁺ Macro exhibited the characteristics of immunosuppressive M2-type macrophages (Fig. 6D). These results suggest that MMP1⁺ CAFs may act on FOLR2⁺ M2-type macrophages, mediating the generation of immunosuppression in gastric cancer.

To explore the relationship between MMP1⁺ CAFs infiltration and immunotherapy, we performed TIDE analysis based on the TCGA-STAD data and found that the low MMP1⁺ CAFs group had more predicted immunotherapy responders than the high MMP1⁺ CAFs group (Fig. 6E). In tumors with high immune cell infiltration, there exist dysfunctional effector cytotoxic T cells that could otherwise effectively kill tumor cells, but whose function is suppressed for certain reasons (dysfunction). In addition, although the immune system can recognize tumors, T cells cannot effectively infiltrate into the tumor interior and are instead excluded (exclusion). We found that the high MMP1⁺ CAF infiltration group had higher TIDE immunosuppression scores (Fig. 6F), more dysfunctional cytotoxic effector T cells (Fig. 6G), and more T cells excluded from the tumor (Fig. 6H). Microsatellite instability (MSI) is considered one of the most promising biomarkers for immunotherapy; we found that the high MMP1⁺ CAF infiltration group had lower MSI scores, suggesting immunotherapy resistance in this subset of patients (Fig. 6I). We observed a trend for higher MMP1⁺ CAF scores in non-responders from an anti-PD1 therapy cohort of GC patients (PRJEB25780)^[25], suggesting that MMP1⁺ CAFs may serve as an indicator of immunotherapy resistance (Fig. 6J,K). To further confirm our findings, we tested the predictive capacity of MMP1⁺ CAFs in the

immunotherapy cohort, and the results showed that MMP1⁺ CAFs were effective and reliable in predicting immunotherapy resistance (Fig. 6L). Collectively, these results reveal that MMP1⁺ CAFs may serve as an indicator of immunotherapy resistance, which may help us identify populations potentially sensitive to immunotherapy.

Figure 6. MMP1⁺ CAF can serve as a predictive indicator of immunotherapy resistance. (A) Box plot comparing immune cell infiltration between the MMP1⁺ CAF high- and low-infiltration groups based on CIBERSORT, with M2 macrophages enriched in the high group. (B) t-SNE plot showing five macrophage subpopulations (C1_IL4I1_Macro, C2_IGLC2_Macro, C3_FOLR2_Macro, C4_JUNB_Macro, C5_MKI67_Macro). (C) Heatmap showing the interaction strength between MMP1⁺ CAFs and macrophage subpopulations, strongest with FOLR2⁺ Macro. (D) Box plot of the M2 score across macrophage subpopulations, highest in FOLR2⁺ Macro (Kruskal-Wallis, $P < 2.2e - 16$). (E) Stacked bar plot showing the proportion of predicted immunotherapy responders in the MMP1⁺ CAF high versus low groups. (F–I) Violin plots comparing TIDE (F), Dysfunction (G), Exclusion (H), and MSI (I) scores between the MMP1⁺ CAF high- and low-infiltration groups. (J) Waterfall plot showing MMP1⁺ CAF scores in responders and non-responders from the anti-PD-1 GC cohort (PRJEB25780). (K) Violin plot comparing MMP1⁺ CAF scores between non-responders and responders. (L) ROC curve evaluating the predictive performance of MMP1⁺ CAF score for immunotherapy response (AUC = 0.639). ns, not significant; **** $P < 0.0001$.

2.7. MMP1⁺ CAFs Are Widely Present Across Multiple Cancers

To investigate the presence of MMP1⁺ CAFs across various cancer types, we performed a pan-cancer analysis. Single-cell data from colorectal cancer (CRC), lung adenocarcinoma (LUAD), hepatocellular carcinoma (HCC), and pancreatic cancer (PDAC) were subjected to dimensionality reduction and clustering analysis (Fig. 7A-D). The fibroblasts therein were extracted for subpopulation analysis (Fig. 7E,G,I,K). We constructed a gene set using the top 30 signature genes of MMP1⁺ CAFs. The results showed that MMP1⁺ CAFs predominantly overlapped with CST11⁺ CAFs in colorectal cancer, with MMP11⁺ CAFs in lung adenocarcinoma, and with PDGFRA⁺ CAFs in hepatocellular carcinoma. In pancreatic cancer, MMP1⁺ CAFs were primarily associated with COL11A1⁺ CAFs, followed by CST2⁺ CAFs and IFIT3⁺ CAFs (Fig. 7F,H,J,L).

We further applied the ssGSEA algorithm to estimate the proportion of MMP1⁺ CAFs in the corresponding cancers within the TCGA datasets and explored their relationship with tumor progression. Analysis based on patient paired samples showed that MMP1⁺ CAFs infiltrated more abundantly in the tumor tissues of colorectal cancer, lung adenocarcinoma, and hepatocellular carcinoma, whereas no statistically significant difference was observed in pancreatic cancer. Furthermore, in lung adenocarcinoma, hepatocellular carcinoma, and pancreatic cancer, the infiltration score of MMP1⁺ CAFs

increased with tumor progression, although the stage-specific differences in lung adenocarcinoma and hepatocellular carcinoma did not reach statistical significance. Survival analysis indicated that MMP1⁺ CAFs were associated with poor prognosis in lung adenocarcinoma, hepatocellular carcinoma, and pancreatic cancer; interestingly, an opposite trend was observed in colorectal cancer. These findings suggest that MMP1⁺ CAFs are associated with tumor progression in certain cancers and are generally associated with poor clinical outcomes across multiple malignancies, highlighting their potential as a universal therapeutic target in cancer treatment (Fig. 7M-P).

Figure 7. MMP1⁺ CAF is widely present across multiple cancer types. (A–D) t-SNE plots showing major cell types in colorectal cancer (CRC, A), lung adenocarcinoma (LUAD, B), hepatocellular carcinoma (HCC, C), and pancreatic cancer (PDAC, D). (E, G, I, K) t-SNE plots showing fibroblast subclusters in CRC (E), LUAD (G), HCC (I), and PDAC (K). (F, H, J, L) Box plots of the MMP1⁺ CAF signature score (top 30 signature genes) across fibroblast subclusters in CRC (F), LUAD (H), HCC (J), and PDAC (L), showing predominant overlap with CST1⁺ CAFs, MMP11⁺ CAFs, PDGFRA⁺ CAFs, and COL11A1⁺ CAFs, respectively (Kruskal-Wallis). (M–P) Analysis of MMP1⁺ CAF infiltration in the corresponding TCGA cohorts—COAD (M), LUAD (N), LIHC (O), and PAAD (P): box plots (top) showing MMP1⁺ CAF abundance in paired tumor versus normal tissues and across tumor stages, and Kaplan-Meier curves (bottom) showing overall survival stratified by MMP1⁺ CAF signature (score-high vs. score-low), with poor prognosis in LUAD, LIHC, and PAAD.

3. Discussion

Cancer-associated fibroblasts are increasingly recognized as active regulators of tumor progression, yet their functional heterogeneity remains a major barrier to therapeutic translation. In this study, we identified an MMP1-expressing CAF subpopulation that progressively accumulated during gastric cancer progression and was associated with adverse clinical outcomes. Integrating single-cell and spatial transcriptomics with public cohorts and functional assays, we further defined a multicellular program through which MMP1⁺ CAFs may coordinate angiogenesis and immune suppression. These findings position MMP1⁺ CAFs as a clinically relevant stromal state and suggest that targeting upstream fibroblast-derived signals may complement current anti-angiogenic and immunotherapeutic strategies.

A central observation of our work is that MMP1⁺ CAFs occupy a low-differentiated, stem-like position along the fibroblast differentiation trajectory and are strongly enriched in ECM remodeling, epithelial-mesenchymal transition, and angiogenesis programs. Notably, MMP1⁺ CAFs did not conform to the classical myCAF, iCAF, vCAF, or apCAF subtypes, indicating that they represent a distinct fibroblast state that expands the existing taxonomy of gastric CAFs. The progressive accumulation of MMP1⁺ CAFs with tumor stage, confirmed at the protein level and in the TCGA-STAD cohort, and their association with shorter overall survival reinforce their pathological significance. A similar MMP1⁺ CAFs cluster, significantly enriched in tumor tissue, has previously been identified in esophageal squamous cell carcinoma (ESCC)^[26]. This population expresses markers including CXCL8, CXCL10, PTGDS, CCL19, and

RGS5 and activates growth factor and ERK signaling pathways. Comparative analysis of the two populations reveals overlap in marker gene expression, suggesting that our MMP1⁺ CAF subpopulation shares several features with the previously identified CAF cluster. Both subpopulations are enriched in genes involved in ECM remodeling. The observed discrepancies in marker expression may reflect tumor type-specific differences, tissue origin variations, or methodological differences in clustering approaches, which could contribute to the heterogeneity between these populations.

Mechanistically, our data delineate a coordinated, multicellular angiogenic circuit centered on MMP1⁺ CAFs. On one arm, MMP1⁺ CAFs signal directly to immature vascular endothelial cells through the GDF15-TGFBR2 axis. Immature endothelial cells, which expanded with tumor stage and were selectively enriched for TGFBR2, are well suited to act as the responsive vascular compartment, given their guiding role in directed endothelial migration during sprouting angiogenesis. GDF15 plays a multifaceted role in cancer. Previous study demonstrated that GDF15⁺ CAFs promoted immunosuppression via the GFRAL/RET cascade in melanoma, thereby contributing to tumor aggressiveness and immunotherapy resistance^[27]. In this study, we identified a distinct mechanism whereby MMP1⁺ CAFs regulate the migration of immature EC through a GDF15-mediated signaling axis. Consistent with prior reports that GDF15 promotes angiogenesis and that the GDF15-TGFBR2 axis regulates cell motility and convergent extension^[28], we found that GDF15 enhanced HUVEC migration and tube formation, whereas the clinical-stage GDF15-neutralizing antibody Pongegromab abrogated these effects. Downstream, SCENIC and validation experiments implicated a set of transcription factors—most prominently TCF12—as effectors of this axis; TCF12 regulated the majority of immature-endothelial signature genes, including the motility-associated myosin MYO1D, which has independently been linked to angiogenesis in colorectal cancer^[29]. On a second arm, MMP1⁺ CAFs engaged Epi 10 tumor cells—the epithelial subpopulation with the highest CNV burden and angiogenic score—through collagen–SDC4 interactions, driving upregulation of VEGFA, which in turn acted on immature endothelial cells via VEGFR1/VEGFR2. Thus, MMP1⁺ CAFs promote neovascularization both directly, via GDF15-TGFBR2, and indirectly, by co-opting malignant epithelial cells into a VEGFA-secreting, pro-angiogenic state. This two-pronged architecture offers a plausible explanation for why single-agent VEGF/VEGFR blockade achieves only limited durable benefit in gastric cancer^[30], and suggests that targeting the upstream stromal driver—or the GDF15-TGFBR2 axis specifically—may complement conventional anti-angiogenic therapy.

Beyond their vascular functions, MMP1⁺ CAFs contributed to an immunosuppressive microenvironment. Previous research revealed CAFs can induce macrophage M2-like polarization, and in turn, M2 macrophages can also regulate the activation and progression of CAFs^[15, 31, 32]. In our study, high MMP1⁺ CAFs infiltration was associated with enrichment of immunosuppressive M2 macrophages and depletion of memory B and naive T cells, and cell-cell communication analysis identified FOLR2⁺ M2-like macrophages as the preferential myeloid partner of MMP1⁺ CAFs. Functionally, high MMP1⁺ CAF abundance corresponded to higher TIDE scores, greater T-cell dysfunction and exclusion, lower microsatellite-instability scores, and a paucity of predicted immunotherapy responders; in an

independent anti-PD-1 gastric cancer cohort, MMP1⁺ CAF abundance was enriched among non-responders and predicted immunotherapy resistance. These observations position MMP1⁺ CAF not only as a mechanistic driver but also as a candidate biomarker that could help stratify patients for immune checkpoint blockade—an unmet need in gastric cancer, where response rates remain modest and reliable predictive markers are scarce^[33]. The convergence of pro-angiogenic and immunosuppressive functions within a single fibroblast subpopulation further raises the possibility that MMP1⁺ CAFs coordinate vascular and immune remodeling as complementary facets of a unified tumor-promoting program.

Finally, pan-cancer analysis demonstrated that MMP1⁺ CAF-like populations are broadly present across colorectal cancer, lung adenocarcinoma, hepatocellular carcinoma, and pancreatic cancer, overlapping with previously described matrix-remodeling CAF states and generally correlating with poor prognosis, with the notable exception of colorectal cancer. This broad conservation suggests that the MMP1⁺ CAF program reflects a recurrent, potentially target of stromal malignancy rather than a gastric-specific phenomenon, while the divergent behavior in colorectal cancer underscores that stromal function is context-dependent and warrants tissue-specific interrogation before therapeutic translation.

Several limitations should be acknowledged. First, our single-cell cohort was small and derived from a single institution; although key findings were corroborated across multiple public bulk, spatial, and immunotherapy cohorts, validation in larger multi-center cohorts is warranted. Second, the causal contribution of MMP1⁺ CAFs to angiogenesis, immune suppression, and therapeutic resistance was inferred largely from computational analyses and in vitro assays; genetic or pharmacological perturbation of MMP1⁺ CAFs in vivo, ideally in immunocompetent models, will be required to establish causality and to test the therapeutic potential of GDF15-TGFBR2 blockade. Third, the precise ontogeny of MMP1⁺ CAFs and the upstream signals that induce this state remain to be defined.

4. Conclusion

In summary, through integrated single-cell and spatial transcriptomic profiling of gastric cancer complemented by multi-cohort and experimental validation, we identified MMP1⁺ CAFs as a low-differentiated, stem-like cancer-associated fibroblast subpopulation that accumulates with tumor progression and predicts poor prognosis. MMP1⁺ CAFs drive tumor angiogenesis through two convergent mechanisms—directly activating immature vascular endothelial cells via the GDF15-TGFBR2–TCF12 axis, and indirectly inducing VEGFA secretion from Epi 10 tumor cells through collagen–SDC4 signaling to engage VEGFR1/VEGFR2—while simultaneously fostering an immunosuppressive microenvironment via crosstalk with FOLR2⁺ M2 macrophages that predicts resistance to immunotherapy. The conservation of MMP1⁺ CAFs across multiple cancer types and their consistent association with adverse outcomes highlight their broad clinical relevance. Collectively, these findings illuminate the dual pro-angiogenic and immunosuppressive role of a defined CAF subpopulation in

gastric cancer and establish MMP1⁺ CAFs, and the GDF15-TGFBR2 axis in particular, as promising targets for anti-angiogenic and immunomodulatory therapy in gastric cancer.

5. Methods

Patients and Samples Collection

A total of 8 patients with pathologically confirmed gastric cancer were enrolled at the First Affiliated Hospital of China Medical University. Fresh tissues were obtained from the tumor core of all 8 patients by surgical resection for single-cell RNA sequencing (scRNA-seq). Written informed consent was obtained from all participants. This study was approved by the Research Ethics Committee (approval no. : 2026-124-2) and conducted in accordance with the relevant ethical guidelines. None of the patients had received chemotherapy or radiotherapy prior to tumor resection. All patients were staged according to the 8th edition of the AJCC guidelines, and their clinical characteristics are summarized in Table S1.

Publicly available single-cell datasets used in this study can be accessed from the following sources: GEO datasets GSE132257 (CRC), GSE202642 (HCC), and GSE189357 (LUAD). The PDAC single-cell dataset is available at <https://doi.org/10.24433/CO.0121060.v1>. The spatial transcriptomic dataset can be obtained from GSE251950. The Affymetrix microarray datasets used in this study include GSE13861, GSE14210, GSE15459, GSE26942 and GSE51105. Bulk Transcriptomic data of gastric cancer immunotherapy cohort were obtained through accession numbers: Kim_PRJEB25780. Furthermore, TCGA datasets for STAD, COAD, LUAD, LIHC, and PAAD were obtained from easyTCGA R package (v0.0.4.2000, <https://github.com/ayueme/easyTCGA>).

Single-cell RNA library construction and sequencing

DNBelab C Series High-throughput Single-Cell RNA Library Preparation Set (MGI, #940-000519-00) was utilized for scRNA-seq library preparation. In brief, the single-cell suspensions were converted to barcoded scRNA-seq libraries through steps including droplet encapsulation, emulsion breakage, mRNA captured beads collection, reverse transcription, cDNA amplification and purification. cDNA production was sheared to short fragments with 300–500 bp, and indexed sequencing libraries were constructed according to the manufacturer's protocol. Qualification was performed using Qubit ssDNA Assay Kit (Thermo Fisher Scientific) and Agilent Bioanalyzer 2100. All libraries were further sequenced by the MGISEQ-2000 or DNBSEQ-T7 sequencing platform with pair-end sequencing. The sequencing reads contained 30-bp read 1 (including the 10-bp cell barcode 1, 10-bp cell barcode 2 and 10-bp unique molecular identifiers (UMI)), 100-bp read 2 for gene sequences and 10-bp barcodes read for sample index.

Single-Cell RNA Sequencing Data Processing

Raw single-cell RNA sequencing data were processed using Cell Ranger (v7.1.0) to obtain unique molecular identifiers (UMIs) and to generate feature–barcode expression matrices. After quality control,

sequencing reads were aligned to the reference human genome GRCh38 under default parameters. The UMI count matrices were imported into Seurat (v5.0.1) for downstream analysis. Cells expressing fewer than 500 or more than 7,500 genes, as well as those with a mitochondrial gene proportion exceeding 20%, were filtered out. After removal of low-quality cells, a total of 54,321 high-quality single cells were retained for downstream analysis.

The data were first normalized using the `NormalizeData` function, and 3,000 highly variable genes were identified with the `FindVariableFeatures` function. Standardization was performed using the `ScaleData` function. Principal component analysis (PCA) was then conducted on the 3,000 highly variable genes, and the top 30 principal components were selected for t-SNE and UMAP dimensionality reduction and visualization. Cell clustering was performed using the `FindNeighbors` and `FindClusters` functions based on the shared nearest neighbor (SNN) algorithm, with 30 principal component dimensions and a resolution of 0.8. Batch effects across samples were corrected using the Harmony package^[34]. Finally, clusters were visualized by UMAP analysis, and 9 major cell types were manually annotated according to canonical marker genes.

Differential Gene Expression Analysis

Differential gene expression analysis was employed to explore transcriptional differences between distinct cell types or states, thereby facilitating a deeper understanding of the tumor microenvironment. Differentially expressed genes among cell subpopulations were identified using the `FindAllMarkers` function in combination with the Wilcoxon rank-sum test. Genes with $\log_2FC > 1$ and an FDR-adjusted p value < 0.05 were considered significantly differentially expressed.

Inference of Copy Number Variations (CNVs) in Epithelial Cells

To characterize epithelial cells and their chromosomal copy number variations (CNVs), we performed unsupervised clustering of the single-cell RNA sequencing data using `inferCNV` (v1.18.1)^[35] under standard parameters and following the developers' guidelines. This analysis focused on large-scale CNV events (e.g., chromosomal insertions or deletions) at the cell lineage and patient levels, with T cells used as the reference.

Trajectory Analysis

To delineate the dynamic functional transitions and differentiation trajectories of cell subpopulations, three complementary approaches were employed: `Monocle2` (v2.22.0)^[36], `Monocle3` (v1.3.1), and `CytoTRACE` (v0.3.3).

For the `Monocle2` analysis, a single-cell dataset object was initialized with the `newCellDataSet` function and restricted to highly variable genes (mean expression ≥ 0.1 across a minimum of 10 cells). Genes exhibiting significant variation (q value < 0.01) were retained, after which `reduceDimension()` was applied to project the cells into a low-dimensional space. The reconstructed differentiation paths were then

displayed using `plot_cell_trajectory()`. Genes displaying pronounced pseudo temporal dynamics (q value < 10^{-10}) were ordered along the trajectory and rendered with `plot_pseudotime_heatmap()`, and were subsequently grouped according to their expression kinetics.

For the Monocle3 analysis, a `CellDataSet` object was assembled from the expression matrix together with the cell and gene metadata derived from the corresponding Seurat object. Following normalization, identification of highly variable genes, dimensionality reduction, and clustering, cellular trajectories were inferred with the `learn_graph` function, and each cell was assigned a pseudotime value along the resulting graph. Genes varying significantly across pseudotime were detected using the `graph_test` function (q value < 0.05), and co-regulated gene modules associated with differentiation were extracted via `find_gene_modules` and visualized as heatmaps across the distinct cell populations.

For the CytoTRACE analysis, the expression matrix was first retrieved from the scRNA-seq dataset and filtered to retain only genes detected in at least five cells. This processed matrix was supplied to the CytoTRACE function to estimate the differentiation potential of each cell. CytoTRACE scores span a continuum from 0 to 1, in which higher values denote greater stemness and lower values correspond to more differentiated cellular states.

Cell - Cell Interaction Analysis

The analysis of cell - cell interaction was performed using the CellChat (v2.1.2)^[37], NicheNet (v2.0.2)^[38] R packages. CellChat infers communication relationships based on the expression levels of ligand-receptor pairs. The single-cell RNA sequencing data were processed by sequentially applying the `identifyOverExpressedGenes`, `identifyOverExpressedInteractions`, `computeCommunProb`, and `aggregateNet` functions. The resulting cell communication networks were visualized using the `netVisual_circle`, `netVisual_heatmap`, `netVisual_bubble`, and `netVisual_aggregate` functions.

The NicheNet package was employed to analyze ligand activity, receptor interaction, and potential target genes between sender and receiver cell groups. For receiver cells, the `nichenet_seuratobj_aggregate` function was used to compute ligand-receptor activity and identify potential target genes with the following thresholds: `expression_pct` = 0.10, `lfc_cutoff` = 0.25, and `cutoff_visualization` = 0.33.

Pathway Analysis

To assess functional enrichment, gene set scores were calculated for each cell type using the `AddModuleScore` function in the Seurat package. Differentially expressed genes (DEGs) between two cell populations were identified using the `FindMarkers` function. Gene Ontology (GO), KEGG, and gene set enrichment analysis (GSEA) were subsequently performed using the ClusterProfiler package (v3.18.1)^[39]. PROGENy^[40] was performed using `progeny`(v1.24.0). Gene set variation analysis (GSVA) was performed using GSVA package (v1.50.5)^[41]. Gene sets representing specific functions were selected from the MSigDB database. GSVA scores were calculated to explore the functions and mechanisms of each cell types.

SCENIC Analysis

To investigate differences in transcription factor activity among cell subpopulations, SCENIC analysis was performed using pySCENIC (v0.10.3)^[42]. First, co-expression modules between transcription factors and their target genes were inferred using GENIE3. Regulons, each comprising a transcription factor and its directly regulated target genes, were then identified using RcisTarget. Finally, regulon activity in each cell was scored using AUCell.

Immune Infiltration And Immunotherapy Efficacy

Immune infiltration in TCGA-STAD samples was assessed using CIBERSORTx (<https://cibersortx.stanford.edu/>)^[43]. The CIBERSORTx LM22 signature matrix includes 22 validated leukocyte gene signatures^[44]. Quantile normalization was disabled in accordance with the recommendations provided by the CIBERSORTx web interface, and the number of permutations was set to 500. The TIDE online tool (<http://tide.dfci.harvard.edu/>) was then employed to predict the likelihood of immunotherapy response of GC patients. In addition, we used the IMvigor210 cohort to analyze the immunotherapy efficacy based on the infiltration levels of MMP1⁺ CAFs.

Survival Analysis

To explore the prognostic value of cell subpopulations, the infiltration level of each cell type was estimated using the single-sample gene set enrichment analysis (ssGSEA) method, based on the top 30 genes specifically expressed in the single-cell RNA sequencing data. Their prognostic significance was subsequently evaluated by Kaplan–Meier survival analysis.

Spatial Transcriptomics Data Analysis

Spatial Transcriptomics slides can simultaneously capture approximately 3000–5000 markers with two identical capture areas and obtain spatial location and sequencing results. The “Load10X_Spatial” function was used to read the spatial transcriptome data, and exclude markers with the total number of genes less than 500 or the proportion of mitochondrial genes more than 30%. Subsequently, the “SCTransform” function was used for standardization, screening of highly variable genes and normalization. The “RunPCA” function was then used to identify the main sources of change in the spatial transcriptome data, and the first 30 PCs were selected for clustering, with a unified resolution of 1.2. “FindNeighbors” and “FindClusters” were used for unsupervised cluster analysis, and “RunUMAP” was applied to generate a coordinate matrix that visualized the location of individual marker points onto a two-dimensional plane. Spatial map ping between ST and single-cell datasets was performed using the CellTrek package (v0.0.94)^[45] to determine the spatial localization of specific cell types within tissue sections.

Immunohistochemical/Immunofluorescence Staining

For IHC staining, formalin-fixed and paraffin-embedded gastric cancer tissue samples were sectioned at a thickness of 5 μm . Following deparaffinization and rehydration, antigen retrieval was performed by heating the sections in 10 mM sodium citrate buffer (pH 6) at 100°C for 20 minutes. Endogenous peroxidase activity was blocked using 0.3% hydrogen peroxide for 10 minutes. Subsequently, the sections were incubated overnight at 4°C with anti-SMA (MXB, China; MAB-0890). The next day, the sections were treated with horseradish peroxidase-conjugated secondary antibodies at room temperature for 30 minutes, and the signal detection was achieved using 3,3'-diaminobenzidine (DAB) as a chromogen agent. The sections were counterstained with hematoxylin, dehydrated, and imaged. The positively stained area and the number of microvessels at 20 $\times$ magnification was quantified using ImageJ software (v1.53e, U.S. National Institutes of Health).

For mIHC staining, formalin-fixed sections were stained with antibodies against GDF15 (Invitrogen, USA; CL0328), MMP1 (CST, USA; E9S9N), CD34 (Invitrogen, USA; QBEND/10), SMA (MXB, China; MAB-0890), and TGFBR2 (proteintech, USA; 2D5H7) according to the manufacturer's instructions for the multicolor immuno-fluorescence kit (huilانبio, China; RC0086-34RM) and imaged with the Nikon AXR fluorescence microscope (Nikon, Japan).

Transwell Migration Assay

To evaluate cell migration, 500 μL of complete 10% fetal bovine serum (FBS) DMEM medium containing different treatments was added to the lower chamber of an 8 μm pore size Transwell insert (Corning Transwell, 24-well plate). The treatments included 10 ng/mL GDF15 (MedChemExpress, New Jersey, USA; HY-P75170), 10 $\mu\text{g}/\text{mL}$ Pongegromab (MedChemExpress, New Jersey, USA; HY-P99241), a combination of the two agents, and PBS as a control. In the upper chamber, 100 μL of serum-free DMEM containing 30 000 HUVEC cells was seeded. The cells were incubated at 37°C for 24 h. After incubation, the Transwell inserts were removed, and the medium was carefully aspirated. Non-migrated cells on the upper surface of the membrane were gently wiped off using a cotton swab. The inserts were then fixed with 4% paraformaldehyde for 20–30 min. After fixation, the inserts were rinsed with PBS. Migrated cells on the lower side of the membrane were stained with 0.1% crystal violet for 5–10 min, followed by three PBS washes to remove unbound dye. After air-drying, the stained cells were observed and imaged using EVOS™ M5000 Imaging System (Thermo Fisher Scientific, USA) at 20 $\times$ magnification, selecting 3–5 random fields per sample. Cell quantification was subsequently performed using ImageJ software (v1.53e, U.S. National Institutes of Health).

Angiogenesis Assay

A mixture of Matrigel (Corning, NY, USA; 354230) and serum - free DMEM (Gibco; 31885) was prepared at a 1:1 ratio. A total of 50 μL of this mixture was added to each well of a 96-well cell culture plate and allowed to solidify at 37°C for 3 h. HUVEC cells at passage five were seeded into the wells in a volume of 100 μL of medium, containing 30 000 cells per well. The cells were treated with different agents: 10 ng/mL GDF15 (MedChemExpress, New Jersey, USA; HY-P75170), 10 $\mu\text{g}/\text{mL}$ Pongegromab (MedChemExpress, New Jersey, USA; HY-P99241), a combination of the two agents and PBS as a

control. Tube formation was monitored hourly under EVOS™ M5000 Imaging System, and representative images were captured. Tube formation metrics were quantified with ImageJ software (v1.53e, U.S. National Institutes of Health) using the Angiogenesis Analyzer plug-in (<http://image.bio.methods.free.fr/ImageJ/?Angiogenesis-Analyzer-for-ImageJ>).

Quantitative real-time PCR

qPCR was conducted to evaluate the mRNA levels of genes. Total RNA was extracted using TRIzol reagent (Invitrogen, 15596026CN) according to the manufacturer's instructions. The RNA concentration and purity were assessed using a NanoDrop spectrophotometer (Thermo Fisher Scientific, USA). Reverse transcription was performed using the HiScript III All-in-One RT SuperMix Perfect for qPCR (Vazyme, China; R333-01) following the manufacturer's protocol. qRT-PCR was conducted using the ChamQ SYBR qPCR Master Mix (Vazyme, China; Q311) on a QuantStudio Dx Real-Time PCR System (Applied Biosystems, USA). The reaction conditions were as follows: 95°C for 30 s, followed by 40 cycles of 95°C for 10 s and 60°C for 30 s. Gene-specific primers used for qRT-PCR are listed in Supplementary Table 3. The analysis of mRNA expression level was quantified by the $2^{-\Delta\Delta CT}$ method. Each experiment was performed in triplicate, and data were presented as mean $\pm$ standard deviation (SD).

Statistical analysis

All statistical analyses and graph generation were performed using R software (v4.3.3; <http://www.r-project.org>) and GraphPad Prism 9.0. For comparisons between two groups, either an unpaired two-tailed Student's t test (equal variance) or Welch's corrected unpaired t test (unequal variance) with a 95% confidence interval was used. One-way ANOVA followed by Tukey's post hoc test (for all pairwise comparisons) or Dunnett's test (for comparisons against a single control group) was used for multigroup analyses. Long-rank test was used on survival analysis. For all analyses, statistically significant P values were indicated as follows: no significance, $P \geq 0.05$; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ and **** $P < 0.0001$.

Declarations

Data availability

All data included in this study are available upon request by contacting the corresponding author. Additionally, publicly available datasets were analyzed, which can be accessed at GEO: GSE251950, GSE132257, GSE202642, GSE189357, GSE13861, GSE14210, GSE15459, GSE26942 and GSE51105. The gastric cancer immunotherapy cohort is available under accession number PRJEB25780. The PDAC single-cell dataset is available at <https://doi.org/10.24433/CO.0121060.v1>. TCGA datasets for STAD, COAD, LUAD, LIHC, and PAAD were obtained from the easyTCGA R package (<https://github.com/ayueme/easyTCGA>).

Code availability

This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Ethics approval and consent to participate

This study was approved by the Ethics Committee of the First Hospital of China Medical University (protocol no. 2026-124-2).

Competing interests

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

Acknowledgements

Not applicable

Funding

This work was supported by grant No. 82573674 from the National Natural Science Foundation of China; grant No. 82503930 from the National Natural Science Foundation of China; grant No. 2024JH2/102600312 from the Science and Technology Joint Plan of Liaoning Province; grant No. 2025-MS-190 from the Natural Science Foundation of Liaoning Province.

Author information

These authors contributed equally: Zeyu Li, Na Han, Hanwen Liu, Ruiying Wang.

Authors and Affiliations

Department of Surgical Oncology and General Surgery, Key Laboratory of Precision Diagnosis and Treatment of Gastrointestinal Tumors, Ministry of Education, the First Affiliated Hospital of China Medical University, Shenyang, China

Zeyu Li, Na Han, Hanwen Liu, Mohan Zhang, Chi Xue, Fanghao Xiao, HIDASA IKO, Xuancheng Lin, Zhi Zhu, Ziming Gao & Kai Li

Department of Ultrasound, the First Affiliated Hospital of China Medical University, Shenyang, China

Ruiying Wang

Department of Colorectal Surgery, Cancer Hospital of China Medical University, Cancer Hospital of Dalian University of Technology, Liaoning Cancer Hospital & Institute, Shenyang, China

Rui Guo

Contributions

K.L., Z.G., and Z.Z. conceived and supervised the study. Z.L., N.H., H.L., and R.W. drafted the manuscript and performed the bioinformatic analysis. Z.L., M.Z., C.X., and F.X. conducted the data collection and single-cell data processing. H.L., R.W., H.I., and X.L. contributed to the IHC/IF staining and in vitro experiments. N.H., C.X., X.L., X.W., and R.G. contributed to data interpretation and visualization. K.L., Z.G., and Z.Z. reviewed and revised the manuscript. All authors have reviewed and approved the final manuscript.

Corresponding authors

Correspondence to Kai Li, Ziming Gao or Zhi Zhu.

References

1. Bray, F., et al., *Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. CA Cancer J Clin, 2024. **74**(3): p. 229-263.
2. Smyth, E.C., et al., *Gastric cancer*. Lancet, 2020. **396**(10251): p. 635-648.
3. Kumar, V., K. Ramnarayanan, and R. Sundar, *Prognosis and Treatment of Gastric Cancer: A 2024 Update*. Cancers (Basel), 2024. **16**(9): p. 1708.
4. Hanahan, D., *Hallmarks of Cancer: New Dimensions*. Cancer Discov, 2022. **12**(1): p. 31-46.
5. Yasuda, T. and Y.A. Wang, *Gastric cancer immunosuppressive microenvironment heterogeneity: implications for therapy development*. Trends Cancer, 2024. **10**(7): p. 627-642.
6. DeBerardinis, R.J., *Tumor Microenvironment, Metabolism, and Immunotherapy*. N Engl J Med, 2020. **382**(9): p. 869-871.
7. Folkman, J., *Tumor angiogenesis: therapeutic implications*. N Engl J Med, 1971. **285**(21): p. 1182-1186.
8. Newman, A.C., et al., *The requirement for fibroblasts in angiogenesis: fibroblast-derived matrix proteins are essential for endothelial cell lumen formation*. Mol Biol Cell, 2011. **22**(20): p. 3791-3800.
9. Lugano, R., M. Ramachandran, and A. Dimberg, *Tumor angiogenesis: causes, consequences, challenges and opportunities*. Cellular and Molecular Life Sciences, 2020. **77**(9): p. 1745-1770.
10. Viallard, C. and B. Larrivée, *Tumor angiogenesis and vascular normalization: alternative therapeutic targets*. Angiogenesis, 2017. **20**(4): p. 409-426.
11. de Visser, K.E. and J.A. Joyce, *The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth*. Cancer Cell, 2023. **41**(3): p. 374-403.
12. Chhabra, Y. and A.T. Weeraratna, *Fibroblasts in cancer: Unity in heterogeneity*. Cell, 2023. **186**(8): p. 1580-1609.
13. Foster, D.S., et al., *Multiomic analysis reveals conservation of cancer-associated fibroblast phenotypes across species and tissue of origin*. Cancer Cell, 2022. **40**(11): p. 1392-1406.e7.
14. Cheng, P.S.W., M. Zaccaria, and G. Biffi, *Functional heterogeneity of fibroblasts in primary tumors and metastases*. Trends Cancer, 2025. **11**(2): p. 135-153.

15. Li, X., et al., *Single-cell RNA sequencing reveals a pro-invasive cancer-associated fibroblast subgroup associated with poor clinical outcomes in patients with gastric cancer*. *Theranostics*, 2022. **12**(2): p. 620-638.
16. Lu, G., et al., *Cancer associated fibroblast derived SLIT2 drives gastric cancer cell metastasis by activating NEK9*. *Cell Death Dis*, 2023. **14**(7): p. 421.
17. Qu, X., et al., *Loss of cancer-associated fibroblast-derived exosomal DACT3-AS1 promotes malignant transformation and ferroptosis-mediated oxaliplatin resistance in gastric cancer*. *Drug Resist Updat*, 2023. **68**: p. 100936.
18. Zheng, X., et al., *Exosomes derived from 5-fluorouracil-resistant colon cancer cells are enriched in GDF15 and can promote angiogenesis*. *J Cancer*, 2020. **11**(24): p. 7116-7126.
19. Park, H., et al., *Ionizing radiation-induced GDF15 promotes angiogenesis in human glioblastoma models by promoting VEGFA expression through p-MAPK1/SP1 signaling*. *Front Oncol*, 2022. **12**: p. 801230.
20. Wu, X., et al., *Deciphering of intra-tumoural heterogeneity and the interplay between metastasis-associated meta-program and myofibroblasts in gastric cancer*. *Clin Transl Med*, 2025. **15**(5): p. e70319.
21. Chen, Y., et al., *TCF12 enhances angiogenesis and affects sorafenib response in liver cancer via HIF-1 α interaction*. *Biomol Biomed*, 2025. **25**(8): p. 1868-1881.
22. Li, W., et al., *TCF12 Transcriptionally Activates SPHK1 to Induce Osteosarcoma Angiogenesis by Promoting the S1P/S1PR4/STAT3 Axis*. *Mol Cell Biol*, 2024. **44**(5): p. 178-193.
23. Ma, Z., et al., *Interferon-dependent SLC14A1(+) cancer-associated fibroblasts promote cancer stemness via WNT5A in bladder cancer*. *Cancer Cell*, 2022. **40**(12): p. 1550-1565.e7.
24. Hsu, W.H., et al., *Oncogenic KRAS Drives Lipofibrogenesis to Promote Angiogenesis and Colon Cancer Progression*. *Cancer Discov*, 2023. **13**(12): p. 2652-2673.
25. Kim, S.T., et al., *Comprehensive molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric cancer*. *Nat Med*, 2018. **24**(9): p. 1449-1458.
26. Yang, W., et al., *Single-cell and spatial transcriptomics reveal SPRR2A(+) epithelial-MMP1(+) fibroblast interactions as key drivers of Esophageal Squamous Cell Carcinoma progression*. *Int J Biol Macromol*, 2026: p. 152654.
27. Zhao, Z., et al., *Ectopic expression of GDF15 in cancer-associated fibroblasts enhances melanoma immunosuppression via the GFRAL/RET cascade*. *J Immunother Cancer*, 2025. **13**(6).
28. Sriramareddy, S., et al., *GDF15 Reprograms the Microenvironment to Drive Liver Metastasis of Uveal Melanoma*. *Cancer Res*, 2026.
29. Chen, Y.H., et al., *Myo1b promotes tumor progression and angiogenesis by inhibiting autophagic degradation of HIF-1 α in colorectal cancer*. *Cell Death Dis*, 2022. **13**(11): p. 939.
30. Saeed, A., R. Park, and W. Sun, *The integration of immune checkpoint inhibitors with VEGF targeted agents in advanced gastric and gastroesophageal adenocarcinoma: a review on the rationale and*

- results of early phase trials.* J Hematol Oncol, 2021. **14**(1): p. 13.
31. Gu, X., et al., *Lactate-induced activation of tumor-associated fibroblasts and IL-8-mediated macrophage recruitment promote lung cancer progression.* Redox Biol, 2024. **74**: p. 103209.
 32. Fu, Y., et al., *Decoding the CAF-TAM axis: multi-omics dissection and therapeutic targeting of stromal-immune crosstalk in the tumor microenvironment.* Cell Death Dis, 2026. **17**(1).
 33. Joshi, S.S. and B.D. Badgwell, *Current treatment and recent progress in gastric cancer.* CA Cancer J Clin, 2021. **71**(3): p. 264-279.
 34. Korsunsky, I., et al., *Fast, sensitive and accurate integration of single-cell data with Harmony.* Nat Methods, 2019. **16**(12): p. 1289-1296.
 35. Patel, A.P., et al., *Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma.* Science, 2014. **344**(6190): p. 1396-1401.
 36. Trapnell, C., et al., *The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells.* Nat Biotechnol, 2014. **32**(4): p. 381-386.
 37. Jin, S., et al., *Inference and analysis of cell-cell communication using CellChat.* Nat Commun, 2021. **12**(1): p. 1088.
 38. Browaeys, R., W. Saelens, and Y. Saeys, *NicheNet: modeling intercellular communication by linking ligands to target genes.* Nat Methods, 2020. **17**(2): p. 159-162.
 39. Yu, G., et al., *clusterProfiler: an R package for comparing biological themes among gene clusters.* OMICS, 2012. **16**(5): p. 284-287.
 40. Schubert, M., et al., *Perturbation-response genes reveal signaling footprints in cancer gene expression.* Nat Commun, 2018. **9**(1): p. 20.
 41. Hänzelmann, S., R. Castelo, and J. Guinney, *GSVA: gene set variation analysis for microarray and RNA-seq data.* BMC Bioinformatics, 2013. **14**: p. 7.
 42. Van de Sande, B., et al., *A scalable SCENIC workflow for single-cell gene regulatory network analysis.* Nat Protoc, 2020. **15**(7): p. 2247-2276.
 43. Chen, B., et al., *Profiling Tumor Infiltrating Immune Cells with CIBERSORT.* Methods Mol Biol, 2018. **1711**: p. 243-259.
 44. Steen, C.B., et al., *Profiling Cell Type Abundance and Expression in Bulk Tissues with CIBERSORTx.* Methods Mol Biol, 2020. **2117**: p. 135-157.
 45. Wei, R., et al., *Spatial charting of single-cell transcriptomes in tissues.* Nat Biotechnol, 2022. **40**(8): p. 1190-1199.

Figures

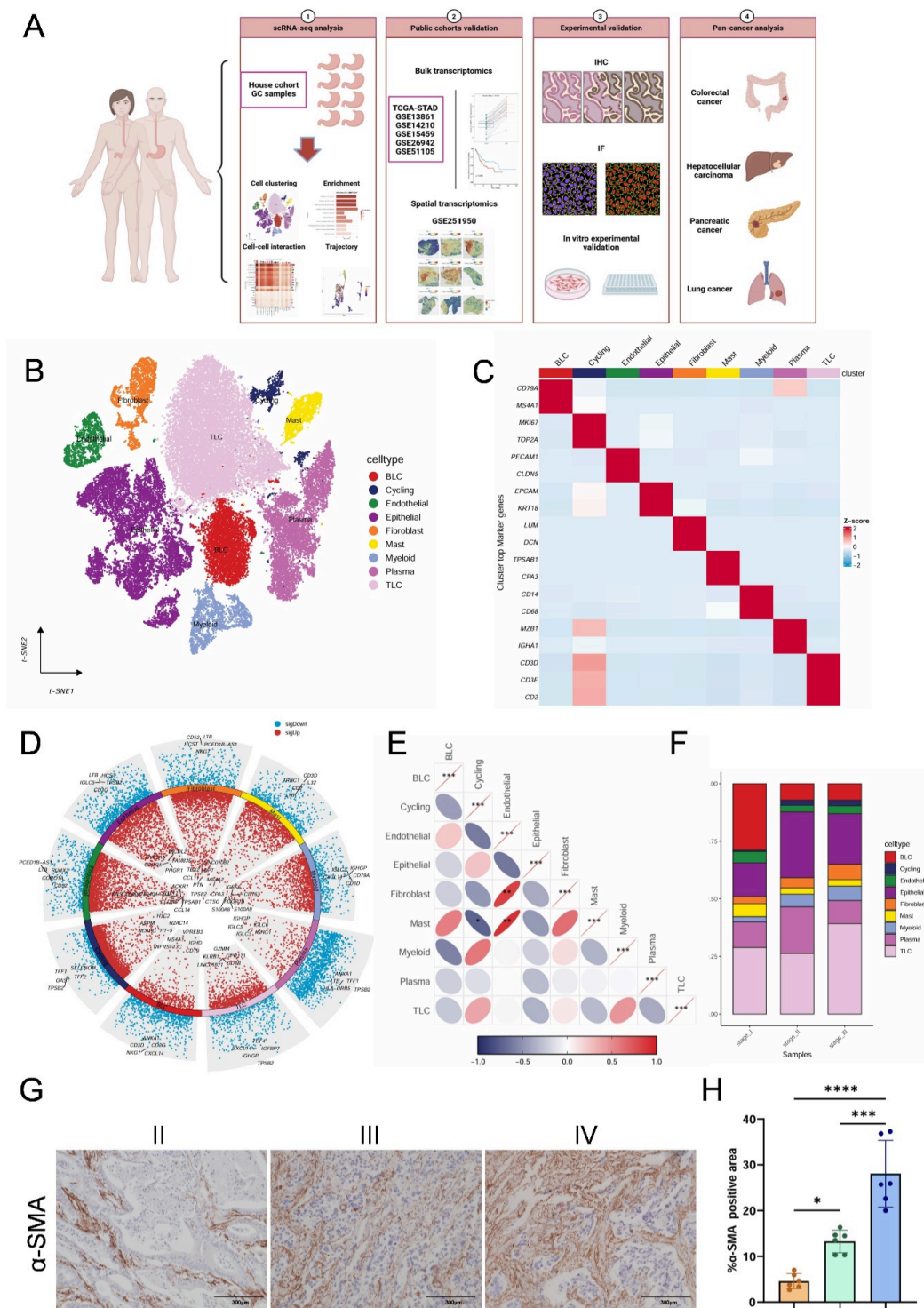


Figure 1

Single-cell transcriptomic landscape of the gastric cancer tumor microenvironment. (A) Schematic of the study design, covering scRNA-seq analysis of an in-house GC cohort, validation in public bulk and spatial transcriptomic datasets, experimental validation (IHC, IF, in vitro assays), and pan-cancer analysis. (B) t-SNE plot showing nine major cell types across 54,321 single cells. (C) Heatmap illustrating the expression of top marker genes in the corresponding cell types, with color indicating the row-scaled

Z-score. (D) Circos plot showing significantly up-regulated (red) and down-regulated (blue) differentially expressed genes for each cell type. (E) Correlation matrix depicting pairwise associations among the nine cell types; blue and red indicate negative and positive correlations, respectively, with fibroblasts and endothelial cells showing a significant positive correlation. (F) Stacked bar plot showing the proportion of major cell types across tumor stages, revealing increasing fibroblast infiltration with tumor progression. (G) Representative α -SMA IHC staining in GC tissues of increasing stage (II–IV). Scale bars, 300 μ m. (H) Quantification of the α -SMA–positive area across stages (II–IV). Data are mean $\pm$ SD; *P < 0.05, ***P < 0.001, ****P < 0.0001.

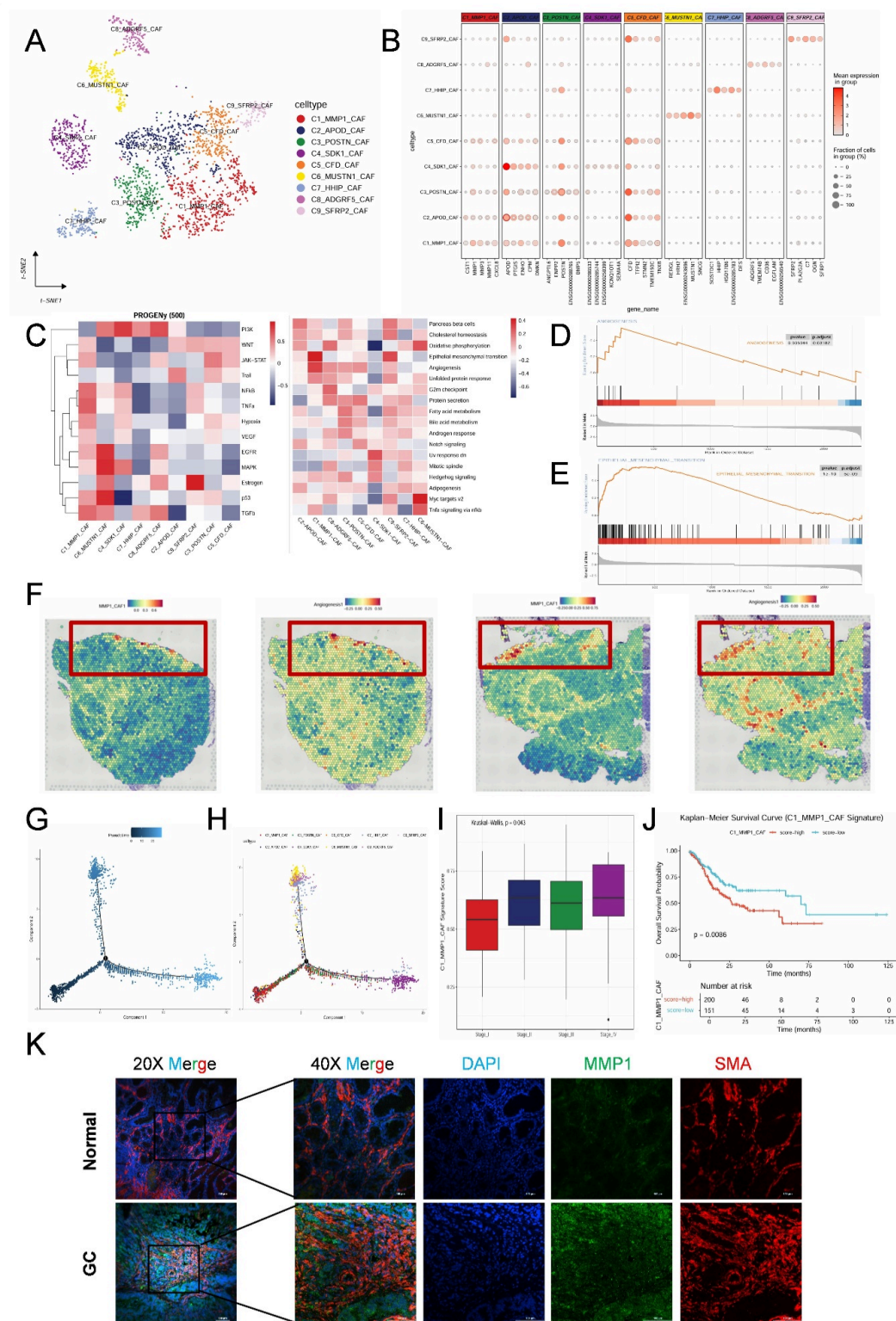


Figure 2

MMP1⁺ CAF is a key subpopulation of cancer-associated fibroblasts associated with poor prognosis.

(A) t-SNE plot showing nine fibroblast subclusters (C1_MMP1_CAF–C9_SFRP2_CAF). (B) Dot plot illustrating the expression of signature genes across the nine fibroblast subclusters, with dot color indicating mean expression and dot size the fraction of cells. (C) PROGENy heatmaps showing pathway activity (left) and functional program scores (right) across fibroblast subclusters, with MMP1⁺ CAFs

enriched in VEGF, Hypoxia, TNF α , angiogenesis, and epithelial-mesenchymal transition. (D, E) GSEA plots showing enrichment of angiogenesis (D) and epithelial-mesenchymal transition (E) gene sets in MMP1⁺ CAFs. (F) Spatial transcriptomic feature plots showing the co-localization of the MMP1⁺ CAF signature with the angiogenesis score (red boxes). (G, H) Monocle2 pseudotime trajectory of fibroblasts colored by pseudotime (G) and by subcluster (H), placing MMP1⁺ CAFs at the initiation point of differentiation. (I) Box plot of the MMP1⁺ CAF signature score across tumor stages (I–IV) in TCGA-STAD (Kruskal-Wallis, P = 0.043). (J) Kaplan-Meier curves showing overall survival stratified by MMP1⁺ CAF signature (score-high vs. score-low; P = 0.0086). (K) Representative immunofluorescence staining of MMP1 (green) and α -SMA (red) with DAPI (blue) in normal and GC tissues, confirming MMP1⁺ CAF enrichment in the tumor microenvironment.

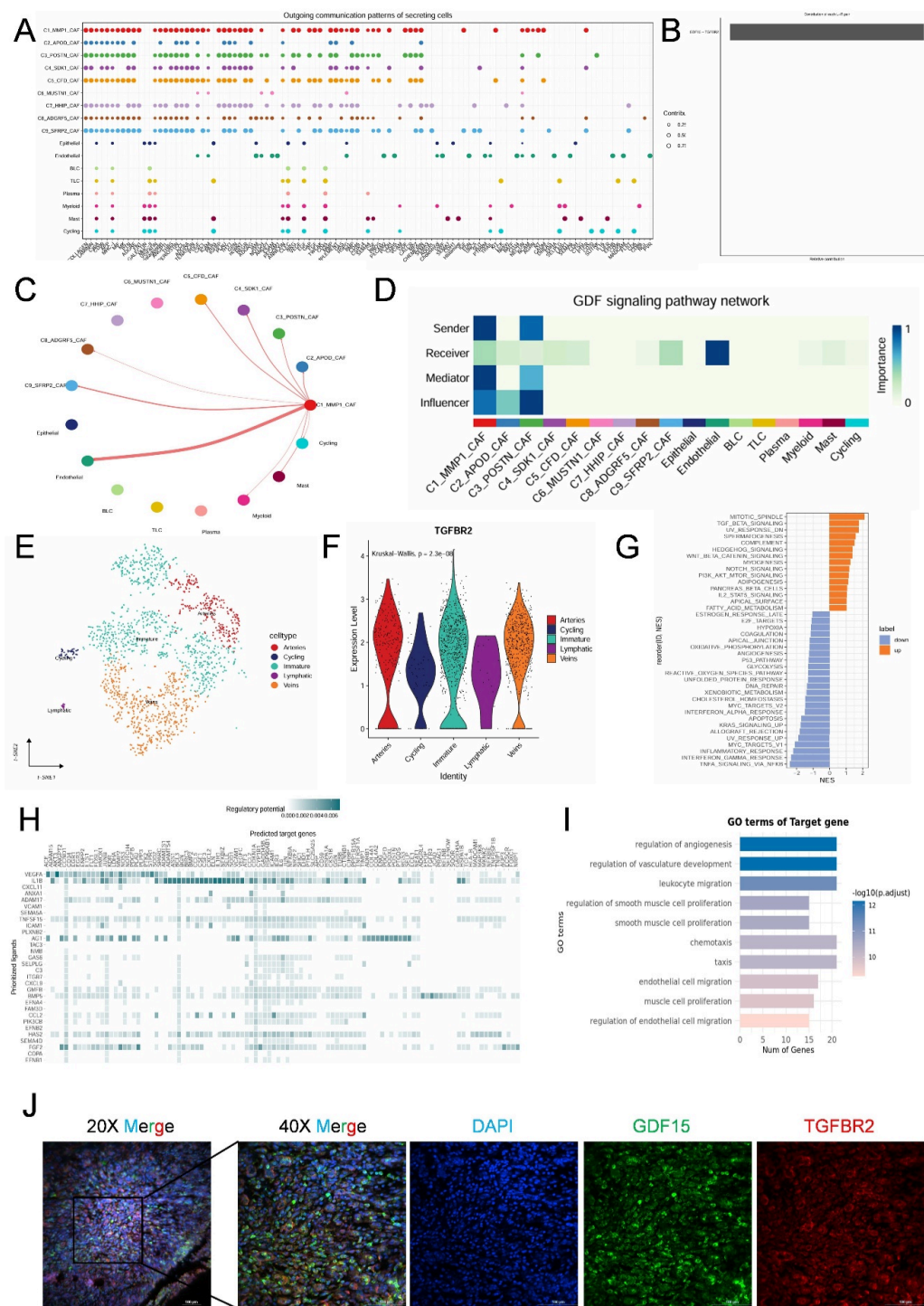


Figure 3

MMP1⁺ CAFs act on immature vascular endothelial cells through the GDF15-TGFR2 axis. (A) Dot plot showing the outgoing communication patterns of secreting cells across all cell types, with MMP1⁺ CAFs displaying broad signaling activity. (B) Bar plot showing the relative contribution of each ligand-receptor pair to the GDF signaling pathway, identifying GDF15-TGFR2 as the sole driver. (C) Circle plot showing MMP1⁺ CAF-derived GDF15 signaling to other cell types, with the strongest effect on endothelial cells.

(D) Heatmap of the GDF signaling pathway network showing the roles (sender, receiver, mediator, influencer) of each cell type. (E) t-SNE plot showing five endothelial cell subpopulations: arterial, venous, lymphatic, cycling, and immature endothelial cells. (F) Violin plot showing *TGFBR2* expression across endothelial subpopulations, with significant enrichment in immature endothelial cells (Kruskal-Wallis, $P = 2.3e-08$). (G) GSEA bar plot showing up-regulated (orange) and down-regulated (blue) pathways in immature endothelial cells, including angiogenesis-related programs. (H) NicheNet heatmap showing the regulatory potential of prioritized MMP1⁺ CAF ligands on predicted target genes in immature endothelial cells. (I) GO enrichment of the predicted target genes, showing enrichment in regulation of angiogenesis, vasculature development, and endothelial cell migration. (J) Representative immunofluorescence staining of GDF15 (green) and TGFBR2 (red) with DAPI (blue) confirming GDF15-TGFBR2 co-localization in GC tissues.

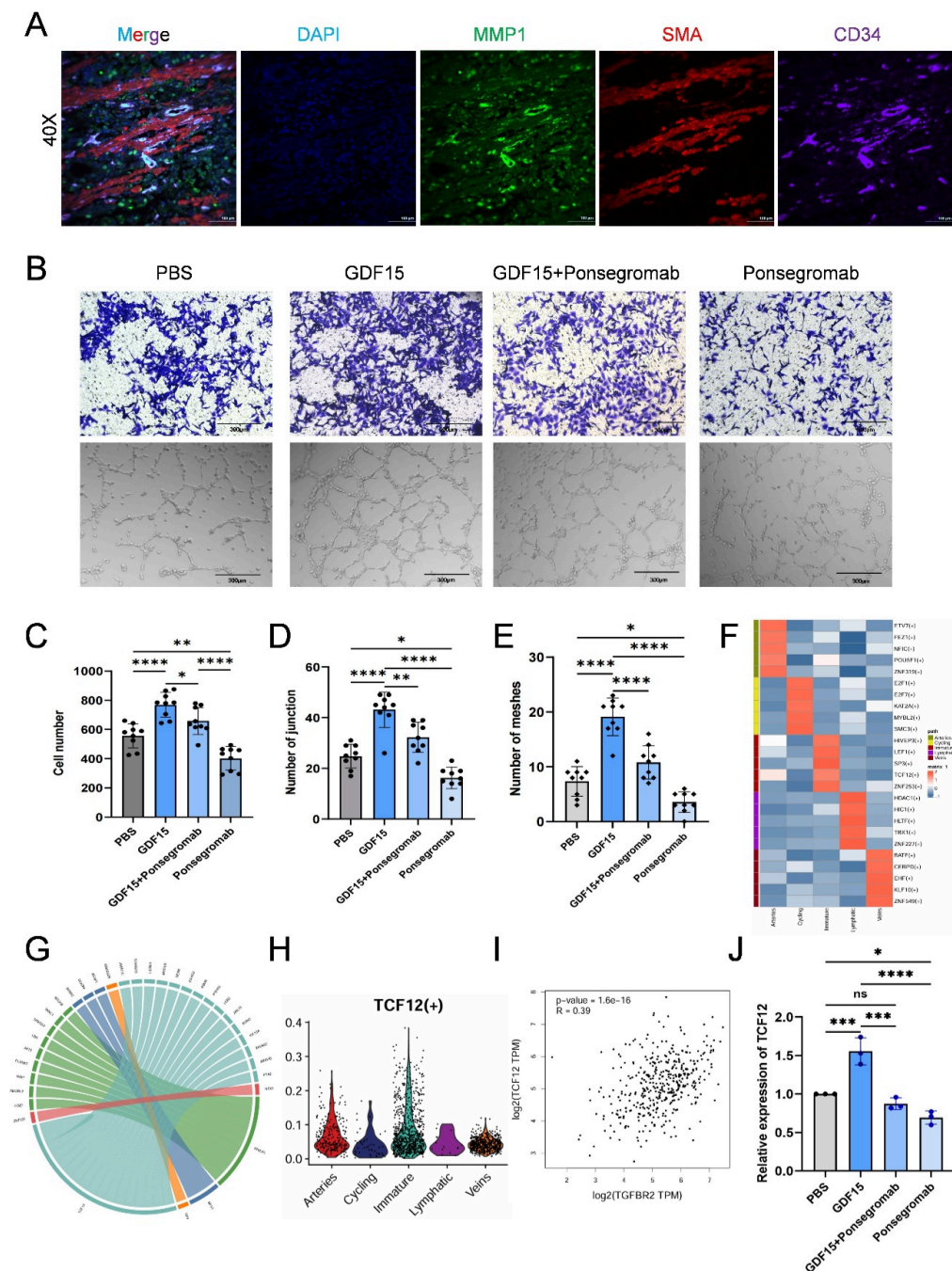


Figure 4

Experimental validation that the GDF15-TGFB2 axis activates the angiogenic function of endothelial cells. (A) Representative multiplex immunofluorescence of MMP1 (green), α -SMA (red), and CD34 (purple) with DAPI (blue) in GC tissues, showing the spatial adjacency of MMP1⁺ CAFs to endothelial cells. (B) Representative images of Transwell migration (top) and tube formation (bottom) assays of HUVECs treated with PBS, GDF15, GDF15 + Ponsegromab, or Ponsegromab. Scale bars, 300 μ m. (C–E)

Quantification of migrated cell number (C), number of junctions (D), and number of meshes (E) across treatment groups. (F) SCENIC heatmap showing transcription factor regulon activity across endothelial subpopulations, with TCF12 enriched in immature endothelial cells. (G) Chord diagram showing the regulatory relationships between transcription factors and the signature genes of immature endothelial cells. (H) Violin plot showing TCF12 regulon activity across endothelial subpopulations, highest in immature endothelial cells. (I) Correlation between *TCF12* and *TGFBR2* expression in TCGA-STAD ($R = 0.39$, $P = 1.6e-16$). (J) qRT-PCR quantification of *TCF12* expression in endothelial cells across treatment groups. Data are mean $\pm$ SD; ns, not significant; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

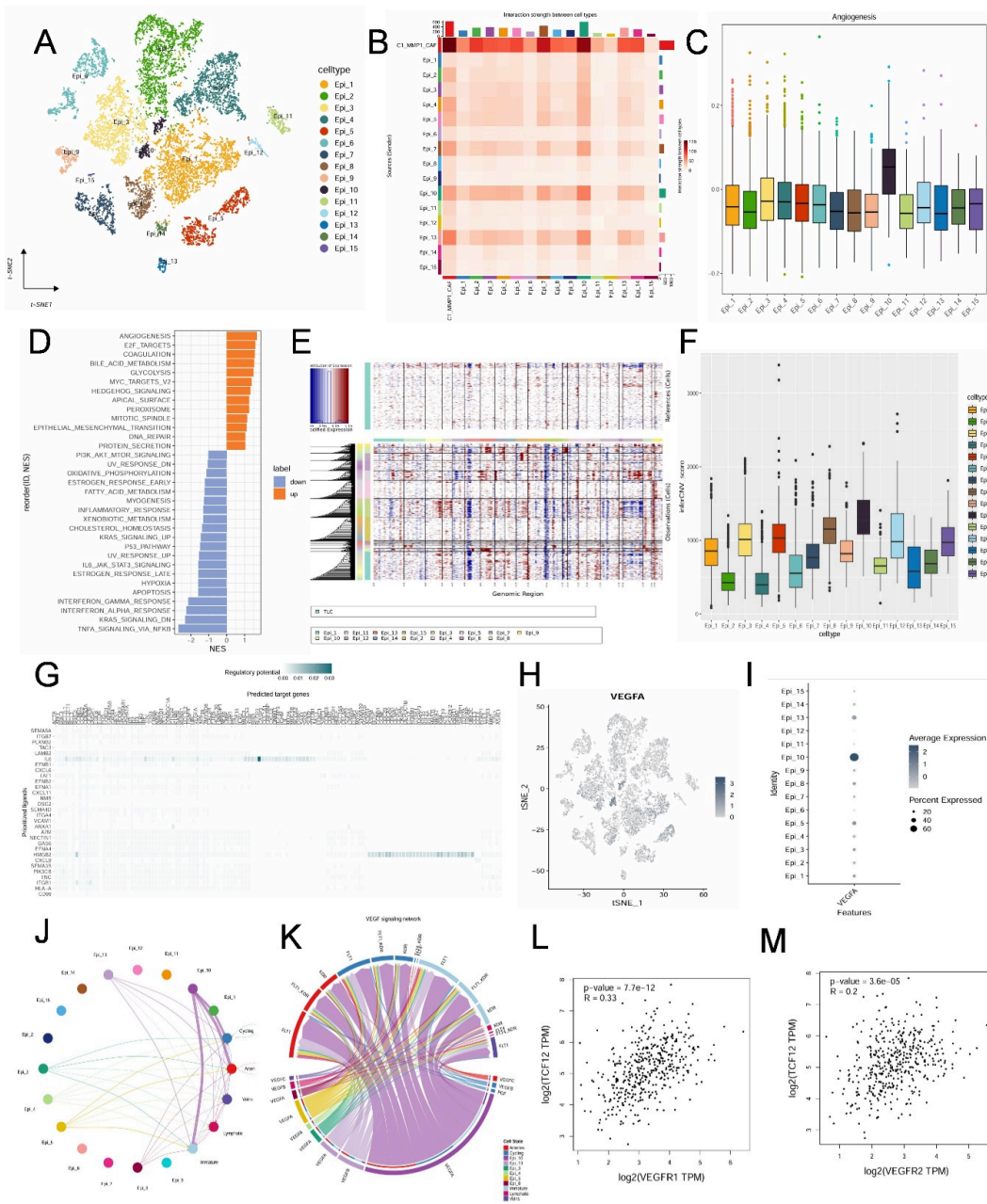


Figure 5

The interaction mechanism by which MMP1⁺ CAFs promote VEGFA release by tumor cells. (A) t-SNE plot showing 15 tumor cell subclusters (Epi 1–Epi 15). (B) Heatmap showing the interaction strength between MMP1⁺ CAFs and each tumor cell subcluster, with the strongest communication observed with Epi 10. (C) Box plot of the angiogenesis score across tumor cell subclusters, highest in Epi 10. (D) GSEA bar plot showing up-regulated (orange) and down-regulated (blue) pathways in Epi 10, including

angiogenesis. (E) inferCNV heatmap showing large-scale copy number variations across tumor cell subclusters, with T cells as the reference. (F) Box plot of the CNV score across tumor cell subclusters, with Epi 10 exhibiting the highest CNV burden. (G) NicheNet heatmap showing the regulatory potential of MMP1⁺ CAF ligands on predicted target genes in Epi 10, including *VEGFA*. (H) Feature plot showing *VEGFA* expression across tumor cell subclusters. (I) Dot plot showing *VEGFA* expression across tumor cell subclusters, specifically enriched in Epi 10. (J) Circle plot showing *VEGFA* signaling from Epi 10 tumor cells to endothelial cell subpopulations. (K) Chord diagram of the *VEGFA* signaling network, showing Epi 10 as the primary source acting on immature and arterial endothelial cells. (L, M) Correlation of *TCF12* expression with *VEGFR1* (L; $R = 0.33$, $P = 7.7e-12$) and *VEGFR2* (M; $R = 0.2$, $P = 3.6e-05$) in TCGA-STAD.

macrophage subpopulations, strongest with FOLR2⁺ Macro. (D) Box plot of the M2 score across macrophage subpopulations, highest in FOLR2⁺ Macro (Kruskal-Wallis, $P < 2.2e-16$). (E) Stacked bar plot showing the proportion of predicted immunotherapy responders in the MMP1⁺ CAF high versus low groups. (F–I) Violin plots comparing TIDE (F), Dysfunction (G), Exclusion (H), and MSI (I) scores between the MMP1⁺ CAF high- and low-infiltration groups. (J) Waterfall plot showing MMP1⁺ CAF scores in responders and non-responders from the anti-PD-1 GC cohort (PRJEB25780). (K) Violin plot comparing MMP1⁺ CAF scores between non-responders and responders. (L) ROC curve evaluating the predictive performance of MMP1⁺ CAF score for immunotherapy response (AUC = 0.639). ns, not significant; **** $P < 0.0001$.

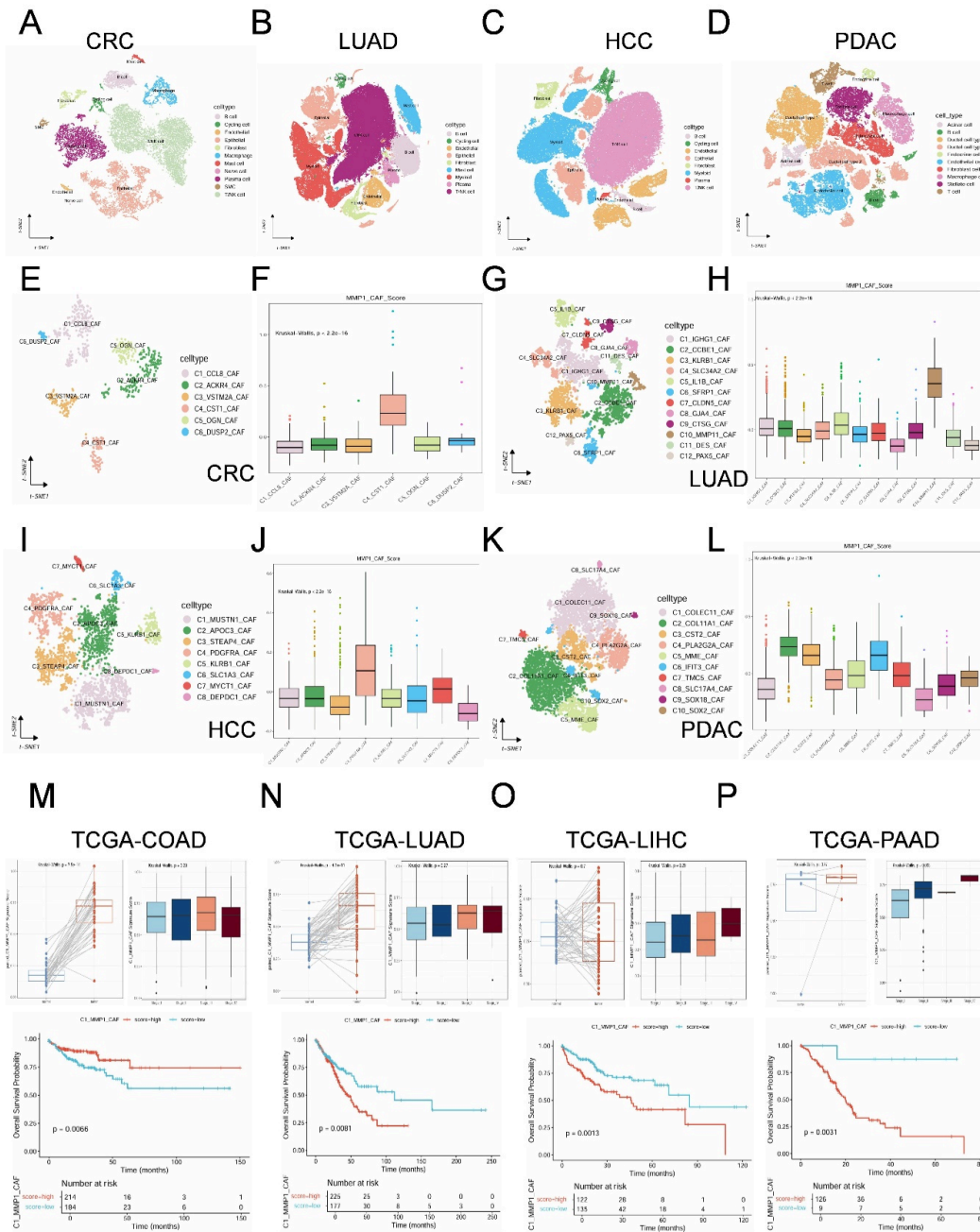


Figure 7

MMP1⁺ CAF is widely present across multiple cancer types. (A–D) t-SNE plots showing major cell types in colorectal cancer (CRC, A), lung adenocarcinoma (LUAD, B), hepatocellular carcinoma (HCC, C), and pancreatic cancer (PDAC, D). (E, G, I, K) t-SNE plots showing fibroblast subclusters in CRC (E), LUAD (G), HCC (I), and PDAC (K). (F, H, J, L) Box plots of the MMP1⁺ CAF signature score (top 30 signature genes) across fibroblast subclusters in CRC (F), LUAD (H), HCC (J), and PDAC (L), showing predominant overlap

with CST1⁺ CAFs, MMP11⁺ CAFs, PDGFRA⁺ CAFs, and COL11A1⁺ CAFs, respectively (Kruskal-Wallis). (M–P) Analysis of MMP1⁺ CAF infiltration in the corresponding TCGA cohorts—COAD (M), LUAD (N), LIHC (O), and PAAD (P): box plots (top) showing MMP1⁺ CAF abundance in paired tumor versus normal tissues and across tumor stages, and Kaplan-Meier curves (bottom) showing overall survival stratified by MMP1⁺ CAF signature (score-high vs. score-low), with poor prognosis in LUAD, LIHC, and PAAD.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation.pdf](#)