

CANCER

Fibroblast atlas: Shared and specific cell types across tissues

Kaidong Liu^{1†}, Yanrui Cui^{1†}, Huiming Han^{1†}, Erliang Guo², Xingyang Shi¹, Kai Xiong¹, Nan Zhang¹, Songmei Zhai¹, Shaocong Sang¹, Mingyue Liu¹, Bo Chen¹, Yunyan Gu^{1*‡}

Understanding the heterogeneity of fibroblasts depends on decoding the complexity of cell subtypes, their origin, distribution, and interactions with other cells. Here, we integrated 249,156 fibroblasts from 73 studies across 10 tissues to present a single-cell atlas of fibroblasts. We provided a high-resolution classification of 18 fibroblast subtypes. In particular, we revealed a previously undescribed cell population, TSPAN8⁺ chromatin remodeling fibroblasts, characterized by high expression of genes with functions related to histone modification and chromatin remodeling. Moreover, TSPAN8⁺ chromatin remodeling fibroblasts were detectable in spatial transcriptome data and multiplexed immunofluorescence assays. Compared with other fibroblast subtypes, TSPAN8⁺ chromatin remodeling fibroblasts exhibited higher scores in cell differentiation and resident fibroblast, mainly interacting with endothelial cells and T cells through ligand *VEGFA* and receptor *F2R*, and their presence was associated with poor prognosis. Our analyses comprehensively defined the shared and specific characteristics of fibroblast subtypes across tissues and provided a user-friendly data portal, Fibroblast Atlas.

INTRODUCTION

The complexity of molecular and physical interactions in biological systems suggests that various cell subtypes perform distinct functions in a context-dependent manner. Tumorigenesis is characterized by high heterogeneity, and efforts to understand this heterogeneity have been largely limited to cancer cells, ignoring those cells in the tumor microenvironment (1). Fibroblasts, a central component of the tumor microenvironment, are canonically referred to as cells that participate in tissue homeostasis and tumorigenesis by producing and maintaining extracellular matrix (2). In recent years, the application of single-cell RNA sequencing (scRNA-seq) has revealed that fibroblasts are also highly heterogeneous in terms of their phenotype, origin, and function (3). The underlying mechanisms responsible for fibroblast heterogeneity remain an active field in cancer research.

In particular, scRNA-seq has advanced our understanding of fibroblast heterogeneity, allowing us to identify previously unidentified fibroblast subtypes in an unbiased manner. For example, Öhlund *et al.* (4) proposed two mutually exclusive and reversible cancer-associated fibroblast (CAF) subtypes in pancreatic cancer, including inflammatory CAF (iCAF) and myofibroblastic CAF (myCAF). Their scRNA-seq study further corroborated the existence of iCAF and myCAF and identified a previously unknown subtype of CAFs, named antigen-presenting CAF (5). In breast cancer, Costa *et al.* (6) identified four different CAF subpopulations (CAF-S1 to CAF-S4) using multicolor flow cytometry. Following this, they further divided CAF-S1 into eight molecularly and phenotypically distinct CAF subtypes by scRNA-seq, and these eight CAF subtypes could reflect some characteristics of

iCAF and myCAF (7). Multiple studies on the same or different tissues have identified similar fibroblast subtypes but assigned them different names, such as the CAF-S4 subtype in breast cancer, the myCAF subtype in pancreatic cancer, and the HNCAF-2 subtype in head and neck squamous cell carcinoma (8). However, it is unclear whether this phenomenon can be extended to other cancer types, and these shared and specific features of CAF subtypes remain unknown.

Buechler *et al.* (9) identified two universal fibroblast subtypes and constructed fibroblast atlases across 17 tissues in healthy and diseased mouse models by integrating 50 scRNA-seq datasets. However, the conservation of fibroblast subtypes in human was only explored in five scRNA-seq datasets from pancreas, colon, and lung tissues. In addition, Luo *et al.* (10) integrated 12 scRNA-seq datasets across 10 solid cancer types in human and identified two normal fibroblast subtypes and six CAF subtypes. The relatively small number of fibroblasts in some studies was insufficient to detect the landscape of fibroblast subtypes, leading to the neglect of some key subtypes (3). Furthermore, the association between the previous fibroblast subtypes with well-defined fibroblast subtypes was unclear. Therefore, it is necessary to integrate as many scRNA-seq datasets as possible and to systematically dissect and compare the relationships between fibroblast subtypes in different tissues.

In this study, we systematically integrated 73 scRNA-seq datasets across 10 tissues to construct a comprehensive atlas of the human fibroblasts, with the ultimate goal of understanding the fibroblast heterogeneity and elucidating the associations between fibroblast subtypes. We identified and defined 18 fibroblast subtypes, and all previously established major fibroblast subtypes were delineated. By contrast, we identified a previously unidentified ubiquitous fibroblast subtype associated with epigenetic modifications, which tended to colocalize with endothelial cells in the spatial transcriptome (ST) data and was more susceptible to T cells through *F2R* receptor. In addition, we provided an online platform to facilitate cancer researchers to further explore the heterogeneity of fibroblasts (<http://medsysbio.org:3838/fibroblasts/>).

Copyright © 2025 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution NonCommercial License 4.0 (CC BY-NC).

¹Department of Systems Biology, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China. ²Department of Thoracic Surgery, Harbin Medical University Cancer Hospital, Harbin 150081, China.

*Corresponding author. Email: guyunyan@ems.hrbmu.edu.cn

†These authors contributed equally to this work.

‡Present address: State Key Laboratory of Frigid Zone Cardiovascular Diseases (SKLFZCD), Harbin Medical University, Harbin, 150081, China.

RESULTS

Construction of a cross-tissue single-cell atlas of fibroblasts

To systematically delineate fibroblast heterogeneity across tissues, we collected and prioritized reported scRNA-seq studies for human tumor and healthy cohorts (Fig. 1A). We obtained 1116 samples from 73 studies, involving 10 tissues and 4 different sample types, including normal, adjacent, tumor, and metastatic samples (Fig. 1B, figs. S1 to S10, and table S1). Fibroblasts extracted from these samples were compiled into a cross-tissue single-cell transcriptome atlas. The integrated fibroblast atlas consisted of 249,156 fibroblasts, of which 65.94% of cells were from tumor samples, 18.67% were from adjacent samples, 11.28% were from normal samples, and 4.11% were from metastatic samples (Fig. 1B).

Functional heterogeneity of fibroblast subtypes

To characterize the heterogeneity of fibroblasts, we performed unsupervised clustering on fibroblasts and identified 33 clusters (Fig. 1C and fig. S11). On the basis of the similarity between cluster-specific genes, we classified 33 fibroblast clusters into 18 common fibroblast subtypes (Fig. 1D). Gene Ontology (GO) enrichment analysis revealed that different fibroblast subtypes were significantly enriched for diverse biological processes ($P < 0.05$, hypergeometric test; fig. S12). Following the hierarchical nomenclature proposed by Lavie *et al.* (3), we delineated 18 fibroblast subtypes: fumarylacetoacetate hydrolase (FAH)⁺ adenosine triphosphate (ATP) synthesis fibroblast, insulin (INS)⁺ inflammatory fibroblast, matrix metalloproteinase 11 (MMP11)⁺ myofibroblast, complement factor D (CFD)⁺ inflammatory fibroblast, interleukin 6 (IL6)⁺ inflammatory fibroblast, pepsinogen A3 (PGA3)⁺ matrix fibroblast, JunB proto-oncogene (JUNB)⁺ ribosome biogenesis fibroblast, secreted phosphoprotein 1 (SPP1)⁺ electron transport fibroblast, ribosomal protein L31 (RPL31)⁺ electron transport fibroblast, tetraspanin 8 (TSPAN8)⁺ chromatin remodeling fibroblast, keratin 14 (KRT14)⁺ electron transport fibroblast, regulator of G protein signaling 5 (RGS5)⁺ myofibroblast, actin alpha 1 (ACTA1)⁺ myofibroblast, phospholipase C gamma 2 (PLCG2)⁺ myofibroblast, aggrecan (ACAN)⁺ ribosome biogenesis fibroblast, CD74 molecule (CD74)⁺ antigen-presenting fibroblast, STEAP family member 1B (STEAP1B)⁺ matrix fibroblast, and tetratricopeptide repeat domain 19 (TTC19)⁺ myofibroblast (Fig. 2A). In particular, we identified and named a previously undescribed subtype of fibroblasts, TSPAN8⁺ chromatin remodeling fibroblasts, characterized by high expression of *TSPAN8*, *GATAD1*, *MDM2*, and *VEGFA*, as well as notable enrichment with histone modification and chromatin remodeling.

Next, we examined transcription factor (TF) expression for fibroblast subtypes. These subtype specific TFs could be used to further understand the function of fibroblast subtypes, as exemplified by the up-regulation of *SP9* in the TSPAN8⁺ chromatin remodeling fibroblasts (Fig. 2B). In addition, we also observed some TFs shared by different fibroblast subtypes, such as *MYOP*, *ZBTB49*, *TFAP4*, and *SOX21*, illustrating common characteristics among fibroblast subtypes (Fig. 2B).

The distribution of fibroblast subtypes in tissue and sample types

We examined the composition of different fibroblast subtypes across tissues (Fig. 2C and fig. S13A). Notably, MMP11⁺ myofibroblasts and CFD⁺ inflammatory fibroblasts accounted for the largest proportion of fibroblasts, with an average of above 20% in most tissues. However, the proportion of fibroblasts showed marked variation

across different tissues. For instance, CD74⁺ antigen-presenting fibroblasts from the stomach exhibited the highest odds ratio (OR) value in tumor samples ($P < 0.001$, Fisher's exact test; fig. S14). Previous work from Pinchuk *et al.* has shown that gastric fibroblasts expressing abundant class II major histocompatibility complex molecules play a crucial role in the activation and proliferation of T helper 17 cells (11). In addition, we further examined the sample distribution of fibroblast subtypes. Among fibroblast subtypes with cell proportions across tissues, we found that CFD⁺ inflammatory fibroblasts were mainly distributed in tumor and adjacent samples, while MMP11⁺ myofibroblasts, RGS5⁺ myofibroblasts, KRT14⁺ electron transport fibroblasts, SPP1⁺ electron transport fibroblasts, and TSPAN8⁺ chromatin remodeling fibroblasts exhibited significantly higher cell proportions in tumor samples (Fig. 2D; $P < 0.05$, Wilcoxon rank-sum test, Fig. 3A). The distribution of different fibroblast subtypes in adjacent and tumor samples may be influenced by the specific roles these cells play in the tumor microenvironment and their involvement in processes during cancer progression (12). Because of the relatively small number of metastatic samples, we did not observe obvious preference of fibroblast subtypes in the metastatic samples (Fig. 2D).

Next, we evaluated fibroblast reprogramming by comparing all fibroblasts in a given pathological tissue to normal fibroblasts. We identified differentially expressed genes (DEGs) for different pathological tissues and found that some chemokines, such as *CXCL1* and *CXCL3*, were highly expressed by fibroblasts from adjacent samples, consistent with the sample distribution of inflammatory fibroblast subtypes. Conversely, fibroblasts from tumor samples showed low expression of these inflammatory factors but high expression of collagen-related genes. These collagen-related genes were also highly expressed by fibroblasts from metastatic samples (Fig. 2E). Pathway analysis revealed that fibroblasts from tumor and metastatic samples shared the highest number of pathways (fig. S13B). These pathways were primarily associated with collagen, such as collagen formation and cross-linking of collagen fibrils (fig. S13C). Fibroblasts played a crucial role in remodeling the extracellular matrix by regulating the balance of collagen fibrils, which were essential for the survival, proliferation, migration, and invasion of cancer cells (13). In contrast, fibroblasts from tumor samples were specifically involved in focal adhesion and smooth muscle contraction pathways (Fig. 2F). After further comparing the differences between fibroblasts in tumor samples and metastatic samples, we observed distinct gene expression patterns. Fibroblasts in tumor samples exhibited high expression of heat shock protein family genes, such as *HSPA1A* and *HSPA1B*, whereas fibroblasts in metastatic samples showed high expression of keratin-associated genes such as *KRT18* and *KRT19* (fig. S15A). Enrichment analysis revealed that fibroblasts in tumor samples were enriched in pathways associated with extracellular matrix, while fibroblasts in metastatic samples were enriched in cytoplasmic translation and ribosome biogenesis pathways (fig. S15, B and C).

Shared and specific characteristics of fibroblast subtypes across tissues

Among fibroblast subtypes with cell proportions across tissues, we found that CFD⁺ inflammatory fibroblast, MMP11⁺ myofibroblast, RGS5⁺ myofibroblast, KRT14⁺ electron transport fibroblast, TSPAN8⁺ chromatin remodeling fibroblast, and SPP1⁺ electron transport fibroblast exhibited significantly higher cell proportions in

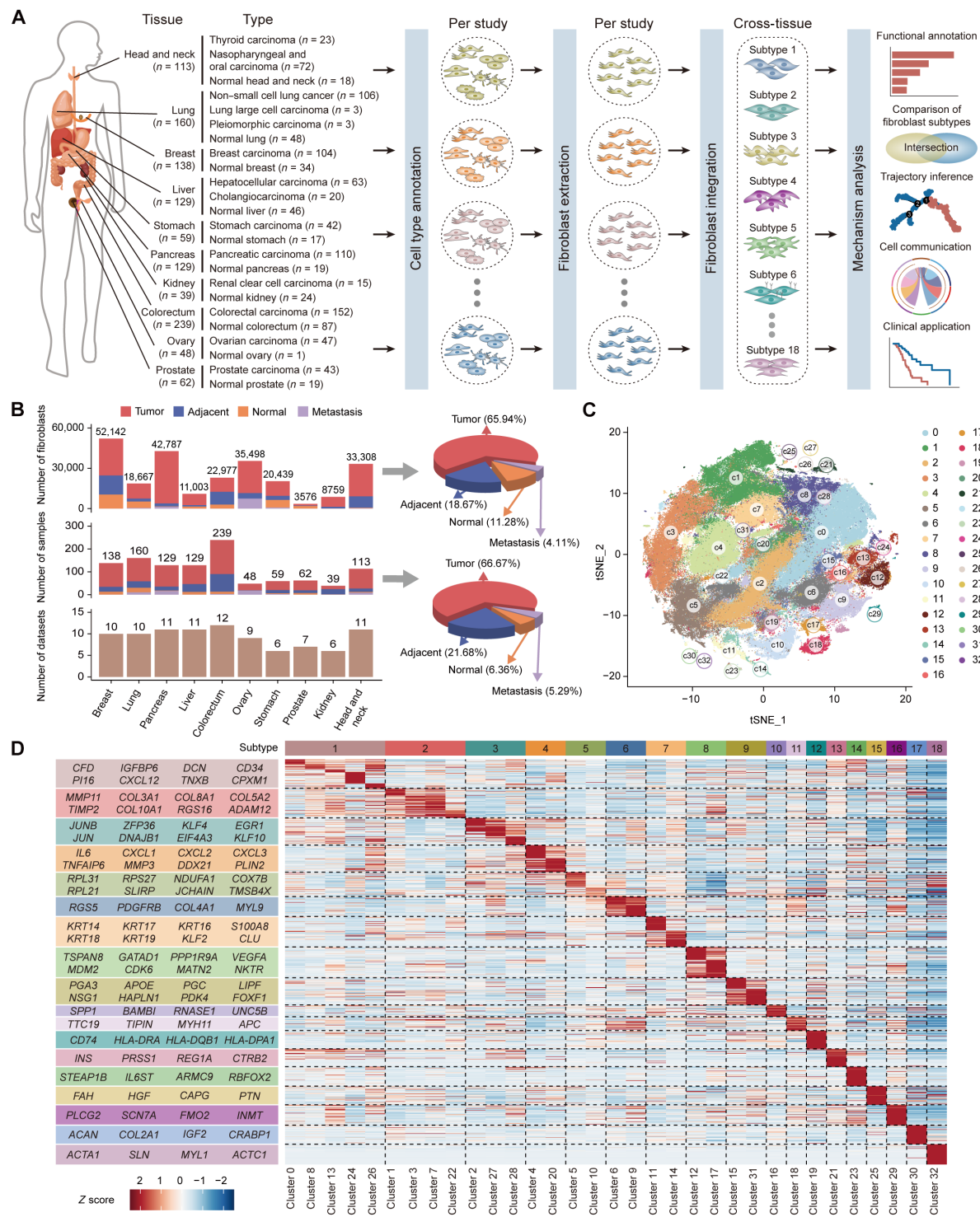


Fig. 1. Defining fibroblast subtypes from a curated collection of scRNA-seq datasets. (A) Workflow of data collection and fibroblast analysis in this study. (B) Barplots showing the number of cells, samples, and datasets across tissues (left), with the exact numbers indicated above each bar. Pie charts depicting summary statistics for the fibroblasts and samples (right). (C) t-Distributed stochastic neighbor embedding (t-SNE) visualization of fibroblast clusters. (D) Heatmap showing the relative expression patterns of marker genes in distinct fibroblast clusters. The top bar representing corresponding fibroblast subtypes. See also table S2.

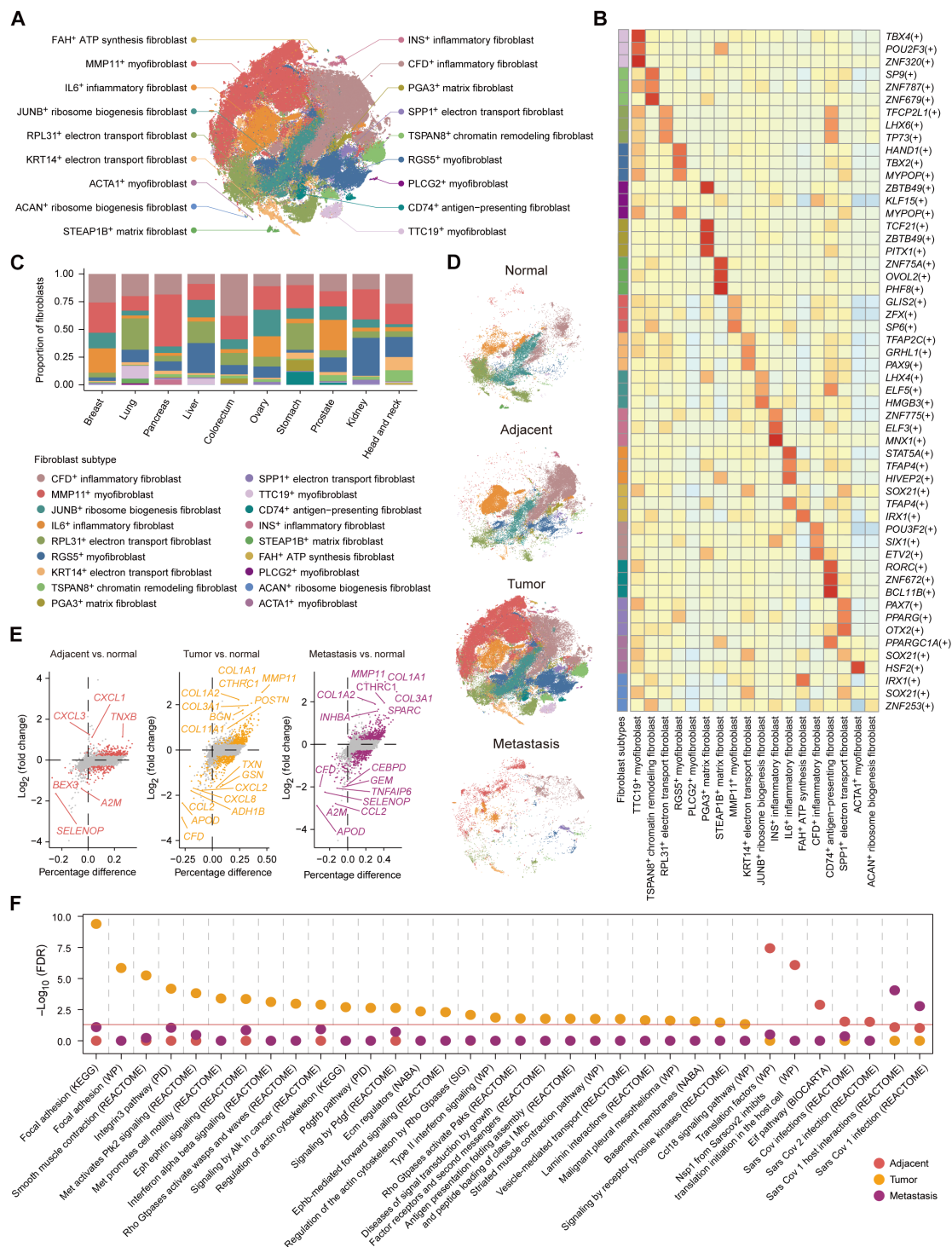


Fig. 2. Transcriptome heterogeneity of fibroblast subtypes. (A) t-SNE visualization of fibroblast subtypes. (B) Heatmap showing the top expression of TFs in each fibroblast subtype. (C) Barplot showing the relative proportion of fibroblast subtypes across tissues. (D) t-SNE plot showing the distribution of fibroblast subtypes across sample groups. (E) Volcano plots showing DEGs between adjacent and normal, tumor and normal, and metastasis and normal fibroblasts. The percentage difference represents the difference in the percentage of fibroblasts expressing the gene. Color denotes genes with the \log_2 fold change > 0.25 and the percentage of fibroblasts expressing the gene > 0.25 . (F) Bubble plot showing specific pathways in the adjacent, tumor, and metastatic fibroblasts. The red line denotes statistical significance. FDR, false discovery rate; KEGG, Kyoto Encyclopedia of Genes and Genomes; WP, WikiPathways; Reactome, Reactome Pathway Knowledgebase; PID, Pathway Interaction Database; SIG, Signaling Gateway; BIOCARTA, BioCarta Pathway Database.

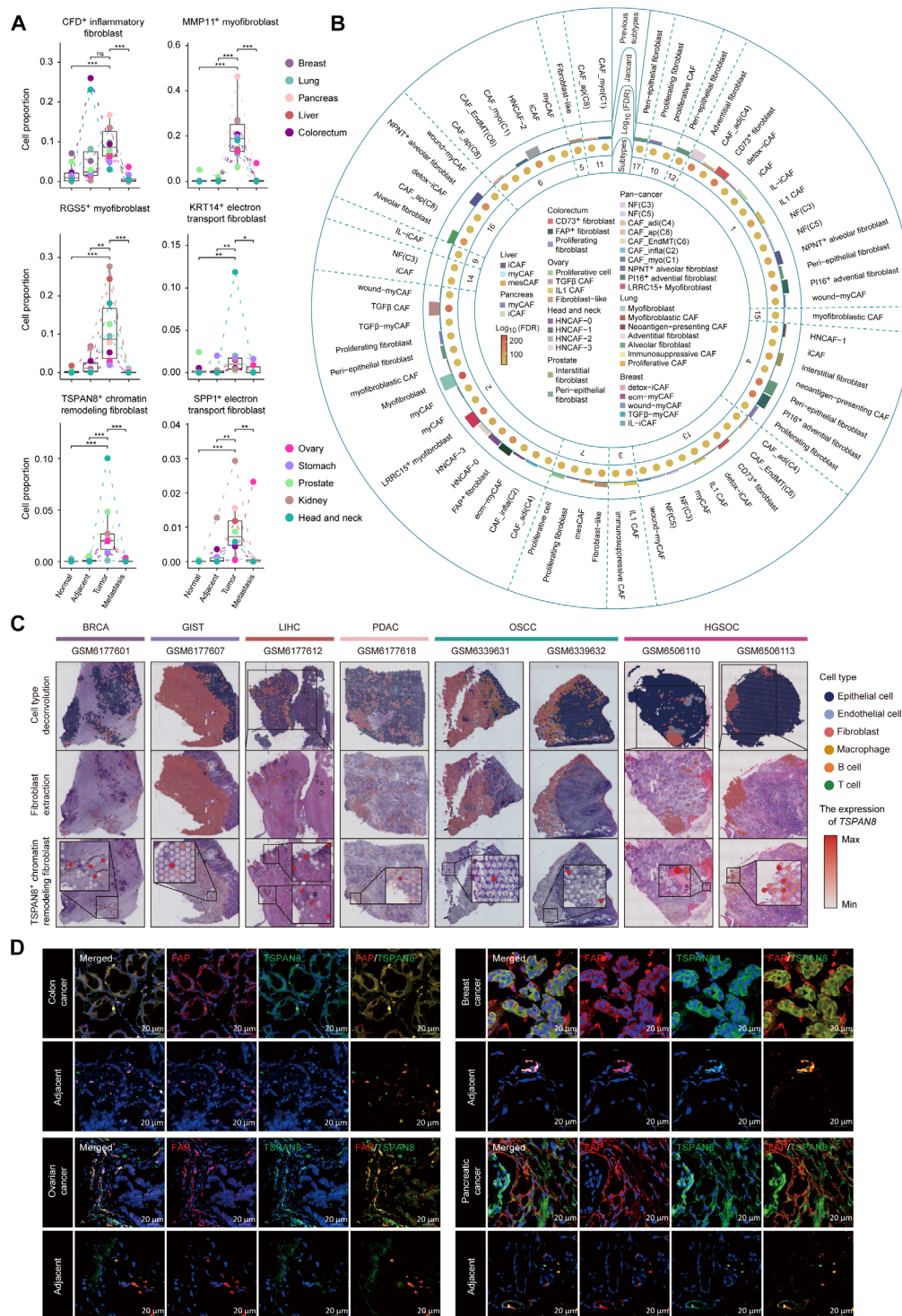


Fig. 3. Identification of TSPAN8⁺ chromatin remodeling fibroblast. (A) Box plots showing cell proportion of six fibroblast subtypes across sample groups. Each point represents a tissue. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; ns, no significance. (B) Comparison between our fibroblast subtypes and previously defined fibroblast subtypes. The color intensity of the points represents the statistical significance of gene set enrichment analysis. The color of bars represents the previously defined fibroblast subtypes. The height of the bars denotes the Jaccard index, calculated by the ratio of gene set intersection to the union. TGF β , transforming growth factor- β . (C) Detection of TSPAN8⁺ chromatin remodeling fibroblast across six cancer types in spatial transcriptomics. Cell type deconvolution of each spot on the histology image (top). Mapping of fibroblasts (middle) and TSPAN8⁺ chromatin remodeling fibroblast (bottom) on the same histology image. BRCA, breast carcinoma; GIST, gastrointestinal stromal tumor; LIHC, liver hepatocellular carcinoma; PDAC, pancreatic ductal adenocarcinoma; OSCC, oral squamous cell carcinoma. (D) Immunofluorescence imaging of TSPAN8⁺ chromatin remodeling fibroblasts in tumor tissues and adjacent tissues of patients diagnosed with colon, breast, ovarian, and pancreatic cancers. TSPAN8 was visualized in green, CAFs were visualized in red, and nuclei were stained in blue. $n = 3$. Scale bars, 20 μ m.

tumor samples ($P < 0.05$, Wilcoxon rank-sum test; Fig. 3A). Thus, we specifically explored the dynamics of different fibroblast subtypes with tumor progression in each tissue. On the basis of the OR results, the aforementioned six fibroblast subtypes showed higher OR values in tumor samples across tissues compared to other sample types (fig. S14). In contrast, the remaining 12 fibroblast subtypes either had no cell proportions in tumor samples or showed no statistical differences when compared to other sample types (fig. S16). Therefore, our results highlighted 6 fibroblast subtypes, including CFD^+ inflammatory fibroblast, MMP11^+ myofibroblast, RGS5^+ myofibroblast, KRT14^+ electron transport fibroblast, TSPAN8^+ chromatin remodeling fibroblast, and SPPI^+ electron transport fibroblast, that were ubiquitous in tumor samples across tissues, while the remaining 12 fibroblast subtypes tended to be tissue specific. This pattern where some fibroblast subtypes are found across various tumor samples, while others exhibit tissue specificity, reflects the complex and diverse roles that fibroblasts play within the tumor microenvironment (14).

We then investigated the concordance of our fibroblast classifications with the 51 fibroblast subtypes defined in 12 previous studies (table S3). Pairwise gene set enrichment analysis indicated that most of our fibroblast classifications were significantly enriched in functionally similar fibroblast categorizations ($P < 0.05$, Hypergeometric test; Fig. 3B). Specifically, CFD^+ inflammatory fibroblasts (subtype 1) matched 13 previously defined fibroblast subtypes, including iCAF in the pancreas. In alignment with the inflammatory function and spatial distribution of CFD^+ inflammatory fibroblasts, iCAF in the pancreas likewise secreted inflammatory mediators and located distantly from tumor cells (4). Furthermore, we found that IL6^+ inflammatory fibroblasts (subtype 4) also matched iCAF in the pancreas. Considering the functional heterogeneity of CFD^+ inflammatory fibroblasts and IL6^+ inflammatory fibroblasts, it is necessary to refine the classification of iCAF. In addition, we also obtained similar results by comparing MMP11^+ myofibroblasts (subtype 2) and RGS5^+ myofibroblasts (subtype 6) with their matched fibroblast subtypes. Regarding KRT14^+ electron transport fibroblasts (subtype 7), the majority of its marker genes were tumor-specific keratins, indicating that fibroblasts might undergo a proliferative process in a specific context. Therefore, it is not unexpected that KRT14^+ electron transport fibroblasts (subtype 7) matched proliferating cell and fibroblast-like subtypes in the ovary, mesCAF subtype in the liver, and proliferating fibroblast subtype in the lung (Fig. 3B). We found that TSPAN8^+ chromatin remodeling fibroblasts did not align with any previously defined fibroblast categorizations (Fig. 3B).

The spatial distribution of TSPAN8^+ chromatin remodeling fibroblasts

TSPAN8^+ chromatin remodeling fibroblasts, as a previously unidentified ubiquitous fibroblast subtype across tissues, exhibit significantly higher cell proportions in tumor samples compared to other sample types ($P < 0.001$, Wilcoxon rank-sum test; Fig. 3A). To validate the existence of TSPAN8^+ chromatin remodeling fibroblasts, we analyzed publicly available ST data across six cancer types: breast carcinoma, gastrointestinal stromal tumor, liver hepatocellular carcinoma, pancreatic ductal adenocarcinoma, oral squamous cell carcinoma, and high-grade serous ovarian carcinoma (HGSOC). Each spot was assigned to a specific cell type based on the highest probabilistic proportion calculated by the robust cell type decomposition (RCTD) method (Fig. 3C). Next, TSPAN8^+ chromatin

remodeling fibroblasts were mapped on the basis of the expression of the marker gene *TSPAN8* within fibroblast-enriched spots. We successfully mapped TSPAN8^+ chromatin remodeling fibroblasts in 21 of the 28 available tissue sections across all six cancer types (Fig. 3C and fig. S17A). In addition, we found that TSPAN8^+ chromatin remodeling fibroblasts were primarily located in the tumor margin area (Fig. 3C and fig. S17A). In the pan-cancer dataset of GSE203612, TSPAN8^+ chromatin remodeling fibroblasts tended to colocalize with endothelial cells rather than other cell types in the tumor microenvironment ($P = 6.58 \times 10^{-08}$, Fisher's exact test; fig. S17B).

Considering the relatively high proportions of TSPAN8^+ chromatin remodeling fibroblasts in breast cancer, ovarian cancer, pancreatic cancer, and colon cancer tissues (Fig. 3A), we conducted multiplexed immunofluorescence assays in these tumor tissues and their adjacent tissues. The up-regulation of *TSPAN8* expression in tumor tissues is observed in fibroblasts tagged with *FAP*, demonstrating the presence of TSPAN8^+ chromatin remodeling fibroblast subtype (Fig. 3D). In addition, tumor tissues exhibited an elevated quantity of TSPAN8^+ chromatin remodeling fibroblasts compared to adjacent tissues (Fig. 3D).

Potential origins of fibroblast subtypes

Since fibroblasts could transition between different functional states, we opted for Monocle2 to infer potential cell differentiation trajectories. The resulting trajectories showed a variety of state transition paths between fibroblast subtypes (Fig. 4A). We defined 11 distinct states within fibroblast subtypes (Fig. 4B). Most fibroblast subtypes, such as CFD^+ inflammatory fibroblasts and MMP11^+ myofibroblasts, were evenly distributed in different states. In contrast, CD74^+ antigen-presenting fibroblasts were largely present in state8, particularly those in the stomach and lung. TSPAN8^+ chromatin remodeling fibroblasts and STEAP1B^+ matrix fibroblasts were dominant in state1, primarily from the lung, pancreas, and head and neck (Fig. 4, C and D, and fig. S18).

To understand the dynamics of fibroblasts, we computed epithelial-mesenchymal transition (EMT) scores and CytoTRACE scores. Fibroblast branches with low EMT scores and high CytoTRACE scores were considered potential roots of the trajectories. We found that fibroblasts in state1, state8, state9, and state10 exhibited lower EMT scores and higher CytoTRACE scores, suggesting that fibroblast subtypes might have distinct origins (Fig. 4, E and F). Using multiple computational methods, we continued to explore the heterogeneous origins of fibroblast subtypes across tissues. On the basis of the gene signatures of resident fibroblasts obtained from the work of Heidegger *et al.* (15), we found that fibroblasts in state8, state9, and state10 exhibited lower resident fibroblast scores than fibroblasts in other states, especially those in state8 ($P = 0.012$, Wilcoxon rank-sum test; Fig. 4G). In contrast, fibroblasts in state1 had similar resident fibroblast scores as fibroblasts at the end of the differentiation trajectories (Fig. 4G). Moreover, in conjunction with the single-cell regulatory network inference and clustering (SCENIC) results, we found that fibroblasts in state8 showed relatively low correlations with fibroblasts in other states (Fig. 4H). Therefore, the trajectory results revealed that fibroblasts in state8, state9, and state10, primarily CD74^+ antigen-presenting fibroblasts, might originate from transitions of other cell types, whereas TSPAN8^+ chromatin remodeling fibroblasts in state1 largely dominated the root of the tissue-resident fibroblasts.

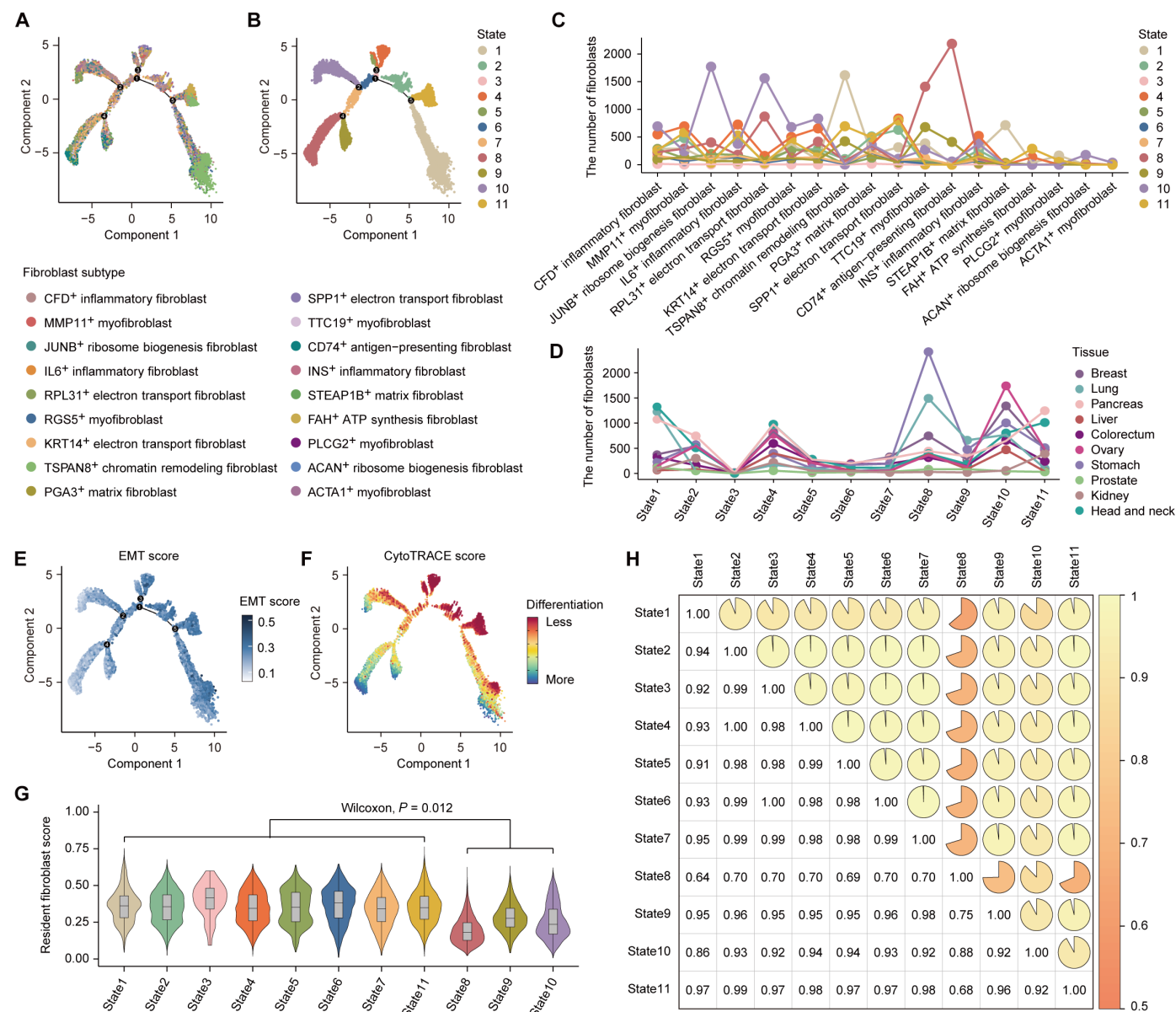


Fig. 4. Differentiation status of fibroblast subtypes. (A and B) Monocle plot showing the differentiation trajectory colored by fibroblast subtypes (A) and states (B). (C) The number of fibroblasts in each fibroblast subtype. Each point represents a state. (D) The number of fibroblasts in each state. Each point represents a tissue. (E and F) Monocle plot showing the differentiation trajectory colored by EMT scores (E) and CytoTRACE scores (F). (G) Violin plots showing the resident fibroblast scores across states. The Wilcoxon rank-sum test was applied to calculate the *P* value between different state groups. (H) Heatmap showing the correlation between TFs in different states. The correlation coefficient was calculated by the Pearson correlation analysis.

The correlation between age and cell senescence of fibroblast subtypes

The distribution of fibroblast subtypes along differentiation trajectories inspired us to study the association between fibroblast subtypes and age. We classified fibroblasts into nine groups based on ages from single-cell sample information. Regardless of gender, inflammatory fibroblasts including CFD⁺ inflammatory fibroblasts, IL6⁺ inflammatory fibroblasts, and INS⁺ inflammatory fibroblasts showed high OR values in the relatively low age distribution range of 30 to 59 years (Fig. 5A and figs. S19 and S20). Conversely, myofibroblasts such as MMP11⁺ myofibroblasts, RGS5⁺ myofibroblasts,

TTC19⁺ myofibroblasts, PLCG2⁺ myofibroblasts, and ACTA1⁺ myofibroblasts displayed high OR values associated with the age distribution ranging from 60 to 79 years (Fig. 5A and figs. S19 and S20). In addition, some fibroblast subtypes exhibited different age distributions between genders. For example, JUNB⁺ ribosome biogenesis fibroblasts exhibited high OR values in females aged 30 to 40 years, while no such trend was observed in males (figs. S19 and S20). The associations between fibroblast subtypes and age could be indicative of the dynamic interplay between these cells and the changing tissue environment over time (2). We also found that CD74⁺ antigen-presenting fibroblasts were distributed

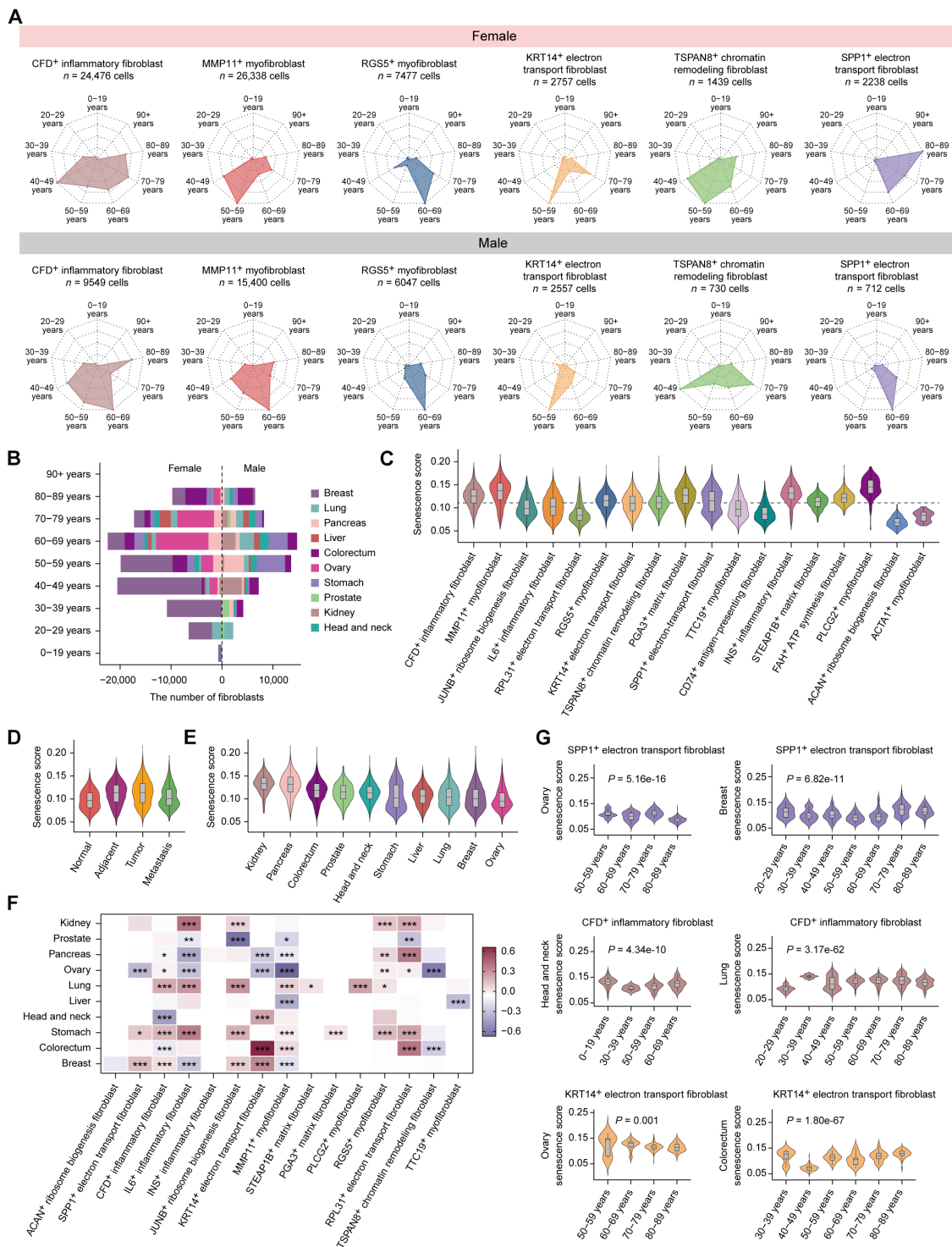


Fig. 5. Aging and cell senescence characteristics of fibroblast subtypes. (A) Radar plots showing the preference of six fibroblast subtypes at different age and gender groups. (B) Barplot showing the number of fibroblasts at different age groups. (C) Violin plots showing the cell senescence scores across fibroblast subtypes. The dashed line denotes the average score of fibroblast senescence. (D and E) Violin plots showing the cell senescence scores across sample groups (D) and tissue groups (E). (F) Heatmap showing the correlation between aging and cellular senescence in each fibroblast subtype across tissues. The Pearson correlation analysis was applied to calculate the correlation coefficients and *P* values. *P* value: * < 0.05; ** < 0.01; *** < 0.001. (G) Violin plots showing the distribution of senescence scores among different age groups. The statistical significance is calculated using a one-way analysis of variance (ANOVA) test.

in 70 to 79 years in females and 80 to 89 years in males (figs. S19 and S20). In contrast, TSPAN8⁺ chromatin remodeling fibroblasts appeared more frequently at 40 to 59 years in contrast to 60+ years (Fig. 5A and fig. S20). Besides, there was no obvious difference in the distribution of fibroblast subtypes with age across tissues (Fig. 5B).

Cellular senescence, a state of cell cycle arrest, showed additional heterogeneity among cell types. Because of the lack of specific markers, we used 526 signature genes overexpressed during cellular senescence to analyze cell-specific senescence levels (16). We implemented 10 scRNA-seq datasets across all tissues and found that fibroblasts had the highest senescence scores compared to other cell types (fig. S21). Next, we further explored the heterogeneity of cellular senescence among the fibroblast subtypes. We found that cellular senescence scores of fibroblast subtypes could reflect different states of the cell differentiation trajectories, that was, fibroblasts at the root of the trajectories often exhibited low cellular senescence scores (Fig. 5C and fig. S18). Specifically, fibroblasts predominantly in state8, state9, and state10, such as RPL31⁺ electron transport fibroblasts, CD74⁺ antigen-presenting fibroblasts, ACAN⁺ ribosome biogenesis fibroblasts, and ACTA1⁺ myofibroblasts, showed the lowest levels of cellular senescence (Fig. 5C). In contrast, tissue-resident–derived TSPAN8⁺ chromatin remodeling fibroblasts did not show a notable difference in the score of cellular senescence (Fig. 5C). Compared with fibroblasts from normal samples, cellular senescence scores of fibroblasts were up-regulated in adjacent samples, tumor samples, and metastatic samples, indicating that tumorigenesis is intertwined with cellular senescence (Fig. 5D). Furthermore, we found that fibroblasts from kidney and pancreas showed relatively high cellular senescence scores, while fibroblasts from ovary had relatively low cellular senescence scores (Fig. 5E).

Unraveling the relationship between aging and cellular senescence will deepen our understanding of the heterogeneity of fibroblast subtypes. In most tissues, the pattern of aging and fibroblast senescence changed in the same direction (Fig. 5F). The prostate, pancreas, ovary, and liver showed opposite directions between cellular senescence and aging changes (Fig. 5F). Regarding fibroblast subtypes, our results highlighted the tissue-specific nature of changes in cellular senescence with aging (Fig. 5F). We also observed significant differences in senescence scores among different age groups in SPP1⁺ electron transport fibroblasts, CFD⁺ inflammatory fibroblasts, and KRT14⁺ electron transport fibroblasts [$P < 0.05$, one-way analysis of variance (ANOVA) test; Fig. 5G]. The exceptions were the RGS5⁺ myofibroblasts and TSPAN8⁺ chromatin remodeling fibroblasts. RGS5⁺ myofibroblasts exhibited consistent positive correlations between cellular senescence and aging across five tissues, while TSPAN8⁺ chromatin remodeling fibroblasts showed significant negative correlations between cellular senescence and aging in the ovary and colon ($P < 0.05$, Pearson correlation analysis; Fig. 5F).

Cell interactions with fibroblast subtypes

We observed specific receptor or ligand genes, such as CXCL12, VEGFA, and PDGFRB, as the marker genes of fibroblast subtypes (Fig. 1D and table S2). Complex intercellular communications among fibroblasts and other cell types are known to play a critical role in tumorigenesis and progression (17). Considering the tissue specificity of certain fibroblast subtypes, we implemented cell interactions with above six ubiquitous fibroblast subtypes in the tumor

samples from 25 scRNA-seq datasets across all tissues, with an average of three datasets per tissue. By CellphoneDB analysis, the interactions between fibroblast subtypes reached the highest number compared to those between fibroblasts and other cell types (Fig. 6A). In addition, regardless of fibroblast subtypes and the tissue types, we detected twice as many fibroblast interactions with endothelial and epithelial cells as with immune cells in the tumor samples (Fig. 6A). Notably, although fibroblasts interacted sparingly with immune cells, fibroblasts presented more frequent interactions with macrophages than T cells and B cells (Fig. 6A).

Subsequently, we further explored receptor-ligand interactions between fibroblast subtypes and other cell types using CellChat. We observed shared receptor-ligand interactions among fibroblast subtypes and other cell types, regardless of the tissue types (Fig. 6B and figs. S22 to S29). However, fibroblast subtypes differed from each other in the expression of these receptor or ligand genes, which could explain the functional heterogeneity observed among fibroblast subtypes (Fig. 6C). Specifically, the ligand CXCL12 was notably enriched in CFD⁺ inflammatory fibroblasts, and they mainly interacted with T cells, B cells, and macrophages through the CXCR4 receptor, highlighting the important role of CFD⁺ inflammatory fibroblasts in immunomodulation (Fig. 6, B and C, and figs. S22 and S23). In addition, we observed specific activation of ligand C3 in CFD⁺ inflammatory fibroblasts, communicating with multiple receptors of C3 in macrophages (e.g., C3AR1, ITGB2, ITGAM, and ITGAX) (Fig. 6C and fig. S23). Our results also indicated that SPP1⁺ electron transport fibroblasts were primarily responsible for the MDK-NCL pathway in fibroblast-epithelial cell interaction (Fig. 6C and fig. S24), which was critical for tumor development and invasion (18). We found that VEGFA (vascular endothelial growth factor A) was specifically enriched in TSPAN8⁺ chromatin remodeling fibroblasts. Consistent with ST analysis, TSPAN8⁺ chromatin remodeling fibroblasts and endothelial cells colocalized in the tumor margin area, suggesting that TSPAN8⁺ chromatin remodeling fibroblasts mainly interacted with endothelial cells through the VEGF signaling pathway (Fig. 6C and fig. S25).

In turn, fibroblast subtypes were also regulated by different cell types. MMP11⁺ myofibroblasts differed from CFD⁺ inflammatory fibroblasts in their communication patterns with macrophages, as they predominantly expressed ITGB5, the receptor for SPP1 on macrophages (fig. S26). This result suggested that different fibroblast subtypes might regulate each other indirectly through other cell types. In contrast, RGS5⁺ myofibroblasts specifically expressed PDGFRB, the receptor for PDGFB in epithelial and endothelial cells, and rarely interacted with immune cells (figs. S27 and S28). We found that B cells rarely communicated to fibroblasts, while the receptor for GZMA in T cells, F2R, was specifically activated in TSPAN8⁺ chromatin remodeling fibroblasts (fig. S29).

The regulation of fibroblast subtypes

Since genomic characteristics were thought to be upstream factors regulating gene expression, we examined the genetic regulation of fibroblast subtypes across tumors. Here, we used bulk data from The Cancer Genome Atlas (TCGA) to investigate the association between fibroblast subtypes and mutations. We found that TP53 and BRAF mutations had the most widespread effects on fibroblast subtypes ($P < 0.05$, Wilcoxon rank-sum test; fig. S30, A and B). In multiple cancer types, TP53 mutation exhibited positive effects on IL6⁺ inflammatory fibroblasts and FAH⁺ ATP synthesis fibroblasts but

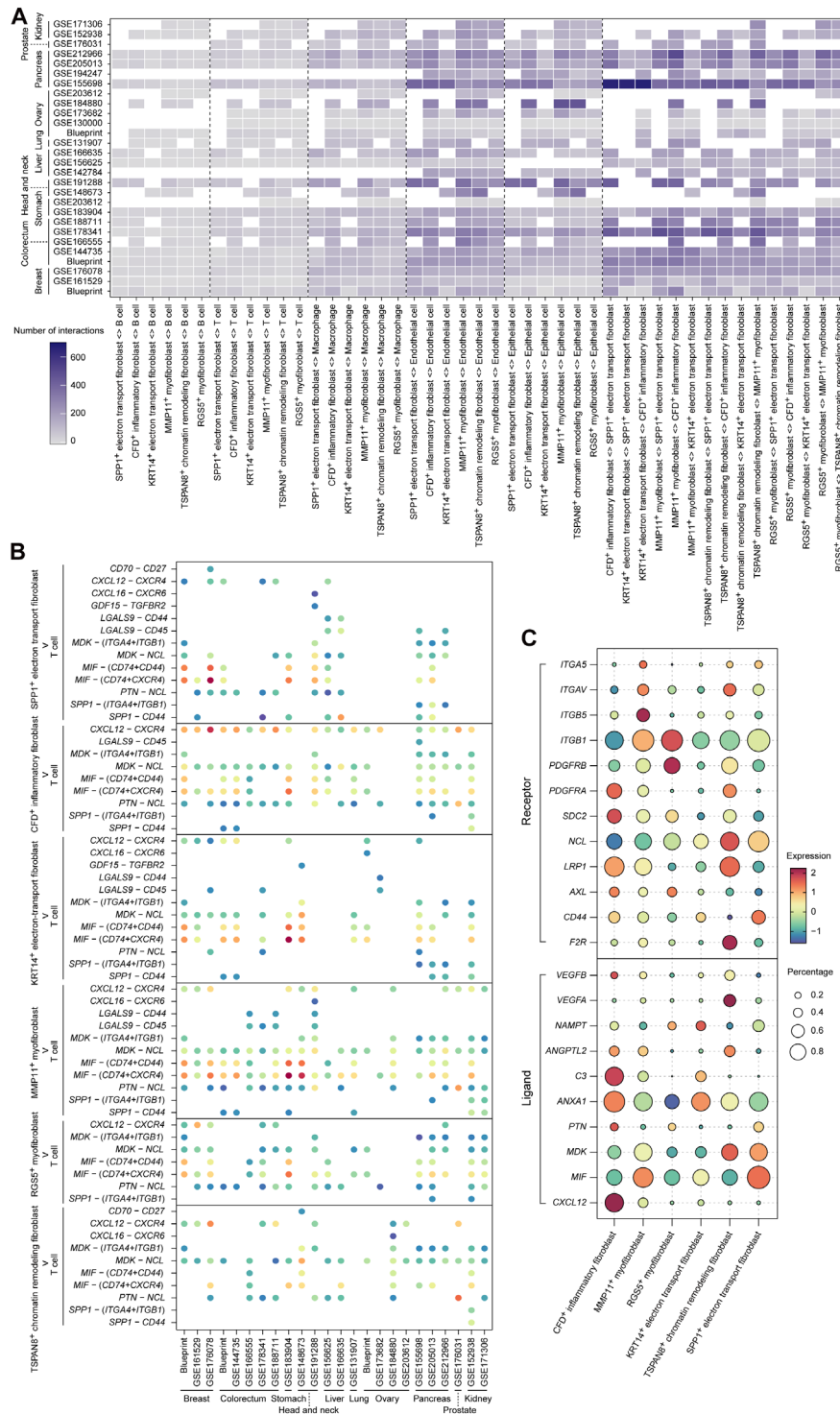


Fig. 6. The cell-cell interactions of fibroblast subtypes. (A) Heatmap showing the interaction strength between two cell types across tissues. The dashed lines were used to separate different cell types. **(B)** Bubble plot showing the specific ligand-receptor interactions across tissues from six fibroblast subtypes to T cells. **(C)** Bubble plot showing the expression pattern of selected ligand and receptor genes in six fibroblast subtypes. The size of each bubble represents the fraction of fibroblasts expressing the gene. The color intensity of bubbles represents the scaled expression level of genes. The black line was used to separate the receptor or ligand genes.

exerted negative effects on RGS5⁺ myofibroblasts and INS⁺ inflammatory fibroblasts. The *BRAF* mutation showed positive rather than negative effects in specific pairings between certain cancer types and fibroblast subtypes. In contrast, we found that TSPAN8⁺ chromatin remodeling fibroblasts were positively associated with *TP53* mutation only in lower-grade glioma (LGG) (fig. S30, A and B).

We next investigated whether fibroblast subtypes were associated with metastasis in TCGA. Compared with mutations, we found associations between metastasis and fibroblast subtypes in only five cancer types (fig. S30C). The most significant correlations with metastasis included KRT14⁺ electron transport fibroblasts, FAH⁺ ATP synthesis fibroblasts, and SPP1⁺ electron transport fibroblasts in kidney clear cell carcinoma, IL6⁺ inflammatory fibroblasts in kidney papillary cell carcinoma (KIRP), and MMP11⁺ myofibroblasts and IL6⁺ inflammatory fibroblasts in thyroid carcinoma ($P < 0.01$, Wilcoxon rank-sum test; fig. S30C). In lung adenocarcinoma, there were significant correlations between IL6⁺ inflammatory fibroblasts and FAH⁺ ATP synthesis fibroblasts with metastasis ($P < 0.05$, Wilcoxon rank-sum test; fig. S30C). By contrast, bladder carcinoma showed moderate correlations between TTC19⁺ myofibroblasts, MMP11⁺ myofibroblasts, and CFD⁺ inflammatory fibroblasts with metastasis (fig. S30C).

The prognostic performance of fibroblast subtypes

We examined prognostic performance of the previously unidentified fibroblast subtype, TSPAN8⁺ chromatin remodeling fibroblasts. The univariate Cox analysis, describing the effect of fibroblast subtype scores on overall survival (OS), demonstrated that TSPAN8⁺ chromatin remodeling fibroblasts were associated with impaired OS in most cancer types (Fig. 7A). In particular, the group with high TSPAN8⁺ chromatin remodeling fibroblast scores was significantly associated with reduced OS in LGG, KIRP, mesothelioma, sarcoma (SARC), ovarian serous cystadenocarcinoma, and stomach adenocarcinoma (STAD) ($P < 0.05$, log-rank test; Fig. 7B). Similarly, we analyzed the associations between patient survival and the remaining 17 fibroblast subtypes. Some associations were context specific, such as the associations of RGS5⁺ myofibroblasts with improved OS in SARC but with reduced OS in STAD (fig. S30D). Other associations were more consistent across cancer types, such as SPP1⁺ electron transport fibroblasts, FAH⁺ ATP synthesis fibroblasts, and MMP11⁺ myofibroblasts correlating with a poor prognosis (fig. S30D).

Association between fibroblast subtypes and immunotherapy response

Since T cells could specifically modulate TSPAN8⁺ chromatin remodeling fibroblasts through the *GZMA-F2R* interaction, we turned to investigate the ability of TSPAN8⁺ chromatin remodeling fibroblasts to predict clinical response to immune checkpoint inhibitors (ICIs). We compiled a set of 20 ICIs therapy datasets that included both bulk transcriptomic data and therapy response information, overall comprising 1079 patients across seven cancer types. We found that higher TSPAN8⁺ chromatin remodeling fibroblast scores were significantly associated with a poorer response to ICIs only in STAD ($P < 0.05$, Fisher's exact test; Fig. 7C). Considering that other fibroblast subtypes were also involved in immunomodulation, we also explored their ability to predict clinical response to ICIs. Notably, we observed some context-specific prediction performance for fibroblast subtypes, such as PGA3⁺ matrix fibroblasts and TTC19⁺

myofibroblasts in non-small cell lung cancer (NSCLC) and melanoma. Unfortunately, these fibroblast subtypes showed unpredictable responses to ICIs when multiple datasets of the same cancer types were combined for NSCLC and melanoma (fig. S31A).

Then, we attempted to further investigate the clinical relevance of TSPAN8⁺ chromatin remodeling fibroblasts in two scRNA-seq datasets of immunotherapy. In the colorectal cancer (CRC) cohort (19), we obtained 52,276 cells consisting of eight cell types after cell annotation, including 2175 fibroblasts (Fig. 7D). Subsequently, a total of 12 subtypes were uniformly assigned to these fibroblasts through the fibroblast atlas built in our study (Fig. 7E). As expected, CFD⁺ inflammatory fibroblasts were found to be enriched in the adjacent samples, whereas other fibroblast subtypes were mainly located in the tumor samples (Fig. 7F). Notably, we observed a high enrichment of RGS5⁺ myofibroblasts and IL6⁺ inflammatory fibroblasts in nonpathological complete response (non-pCR) tumors, while TSPAN8⁺ chromatin remodeling fibroblasts were enriched in pCR tumors (Fig. 7, G and H). In the NSCLC cohort (20), a total of 84,729 cells of nine cell types were analyzed (fig. S31B). We identified 10 fibroblast subtypes using our fibroblast atlas but only found one TSPAN8⁺ chromatin remodeling fibroblast (fig. S31C). In addition, we found that most fibroblasts were derived from patients with stable disease, although patients with partial response accounted for the majority (9 of 15) (fig. S31D). We observed that JUNB⁺ ribosome biogenesis fibroblasts, IL6⁺ inflammatory fibroblasts, and FAH⁺ ATP synthesis fibroblasts were highly enriched in partial response tumors, whereas TSPAN8⁺ chromatin remodeling fibroblasts were enriched in stable disease tumors (fig. S31E).

DISCUSSION

In this study, we constructed a high-resolution fibroblast atlas with well-defined gene signatures and functional states. The large-scale datasets allowed us to precisely elucidate and define 18 fibroblast subtypes, including previously undescribed and overlooked subtypes. Therefore, we were able to further dissect the previous myCAF and iCAF subtypes and discovered the previously unidentified TSPAN8⁺ chromatin remodeling fibroblasts, leading to an improved understanding of the transcriptional heterogeneity of fibroblasts. To facilitate the related research on fibroblasts, we provided a user-friendly, interactive fibroblast portal for visualizing and querying fibroblasts (<http://medsysbio.org:3838/fibroblasts/>).

The comparison with the previous fibroblast subtypes demonstrated that our fibroblast classification provided a higher resolution for distinct fibroblast populations. It is necessary to provide further classification for existing fibroblast subtypes because of the heterogeneity in the origins of myCAF and iCAF subtypes. Accumulating evidence has demonstrated the diverse origins of fibroblasts in different pathological tissues (2, 14, 21). In the trajectory analysis, matched myCAF and iCAF subtypes were distributed in multiple branches that exhibited high differentiation potentials, suggesting that myCAF and iCAF subtypes might have multiple origins. Luo *et al.* (10) revealed a possible evolutionary TAM-CAFAp-CAFmyo path, suggesting that antigen-presenting fibroblasts may originate from macrophages. In addition, they highlighted the possibility of endothelial cells undergoing endothelial-mesenchymal transition as a potential origin for myCAF. Affo *et al.* (22) demonstrated that hepatic stellate cell-derived CAFs exhibited distinct mechanisms of promoting intrahepatic cholangiocarcinoma progression compared

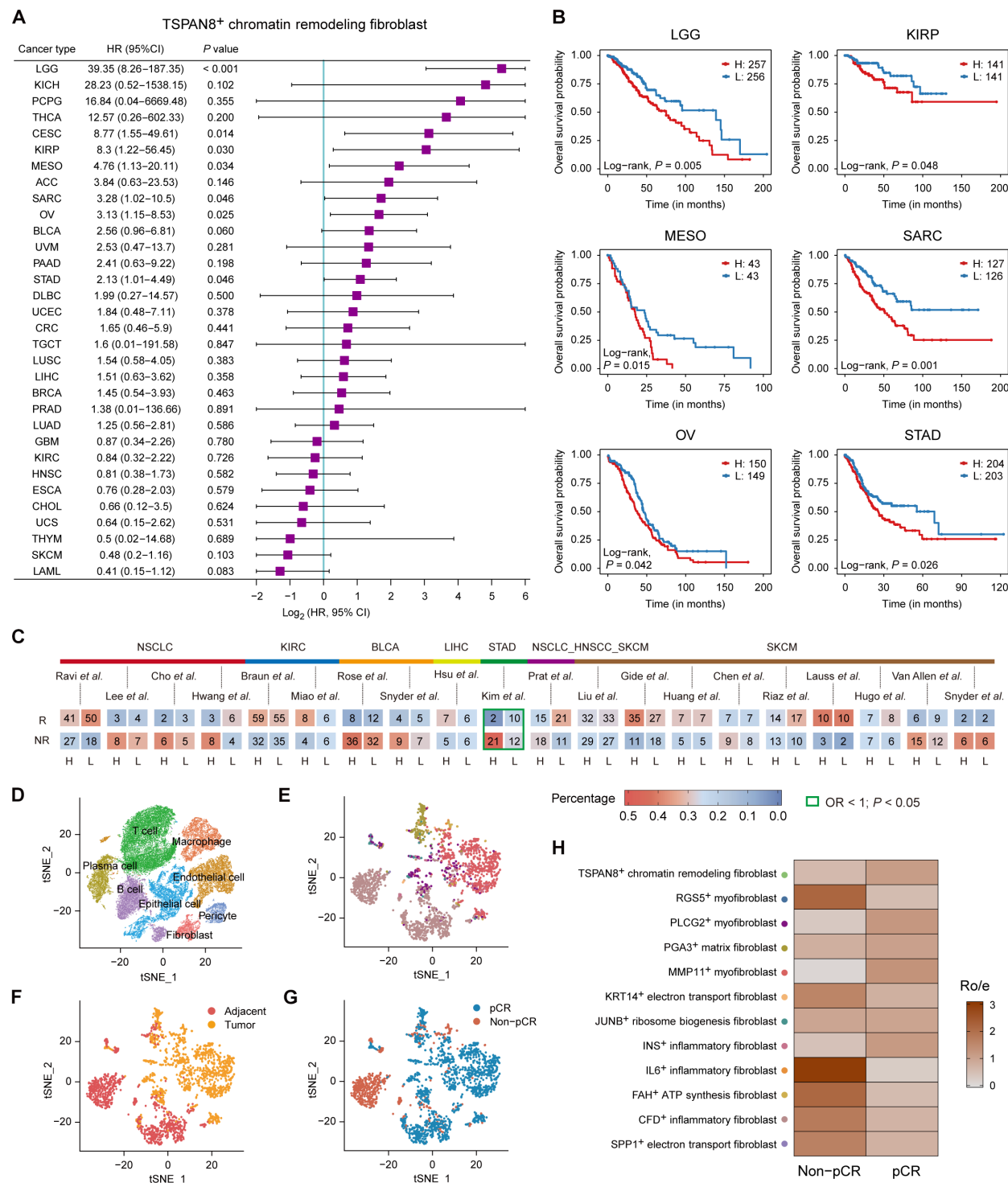


Fig. 7. Clinical significance of TSPAN8⁺ chromatin remodeling fibroblast. (A) Forest plot showing the univariate Cox analysis of TSPAN8⁺ chromatin remodeling fibroblast in TCGA pan-cancer data. HR, hazard ratio; CI, confidence interval. (B) Kaplan-Meier survival curves showing the OS difference between the high and low groups with TSPAN8⁺ chromatin remodeling fibroblast scores in six TCGA cancer types. H, high scoring group; L, low scoring group. (C) Heatmap showing the number of samples between the high and low TSPAN8⁺ chromatin remodeling fibroblast score groups in 20 ICI data. The color intensity is proportional to the number of samples. Green box represents the statistical significance ($P < 0.05$). R, responder; NR, nonresponder; H, high scoring group; L, low scoring group; NSCLC, non-small cell lung cancer; KIRC, kidney renal clear cell carcinoma; BLCA, bladder urothelial carcinoma; HNSCC, head and neck squamous cell carcinoma; SKCM, skin cutaneous melanoma. (D) t-SNE visualization of cell types in the single-cell data of colorectal cancer (CRC). (E to G) t-SNE visualization of fibroblasts only, colored by fibroblast subtypes (E), sample types (F), and immunotherapy response types (G), respectively. pCR, pathological complete response; non-pCR, nonpathological complete response. (H) Heatmap showing the enrichment of fibroblast subtypes in immunotherapy response types.

to portal fibroblast-derived CAFs. In cell interaction analysis of this study, as exemplified by MMP11⁺ myofibroblasts and RGS5⁺ myofibroblasts, they also interacted with different cell types and performed different functions by expressing specific receptor or ligand genes.

In line with the findings of Wu *et al.* (23), our single-cell analysis also revealed that fibroblasts exhibited the highest senescence scores among the various cell types (fig. S21). As reported by Wei *et al.* (24), fibroblasts derived from chronic wounds exhibited a susceptibility to senescence. Accumulating evidence has demonstrated that fibroblast senescence promotes tumor proliferation and invasion (25–27). Furthermore, Chatsirisupachai *et al.* (16) highlighted a tissue-specific association between aging and cellular senescence. Given that fibroblasts exhibit the highest senescence scores compared to other cell types, this tissue-level senescence is likely manifested in fibroblasts. Therefore, when comparing the relationship between age and senescence across different fibroblast subtypes, we identified tissue-specific associations. For example, SPPI1⁺ electron transport fibroblasts exhibited a negative correlation between cellular senescence and aging in the ovary but a positive correlation in the breast (Fig. 5F). One possible reason is that different tissues exhibit distinct mechanisms in response to cellular senescence. For example, senescent cells in the ovary may serve to maintain tissue homeostasis and prevent the excessive production of reactive oxygen species (28). Conversely, in the liver, the accumulation of senescent cells can contribute to fibrotic alterations, ultimately leading to carcinogenesis (29).

In previous scRNA-seq studies, fibroblasts expressing epigenetic-related genes are overlooked due to their relatively small amount (3). In a recent study, Loret *et al.* (30) identified a CAF subtype that highly expressed *STAR*, *TSPAN8*, and *ALDH1A1* after chemotherapy in HGSOc. By integrating 249,156 fibroblasts from multiple independent single-cell platforms, we presented the most comprehensive cross-tissue characterization of TSPAN8⁺ chromatin remodeling fibroblasts. We demonstrated that TSPAN8⁺ chromatin remodeling fibroblasts were detectable in 10x Visium data across six cancer types examined. Notably, TSPAN8⁺ chromatin remodeling fibroblasts tended to colocalize with endothelial cells and specifically expressed ligand *VEGFA*, suggesting that TSPAN8⁺ chromatin remodeling fibroblasts might play an important role in the interaction with endothelial cells. Compared to other fibroblast subtypes, TSPAN8⁺ chromatin remodeling fibroblasts were specially regulated by T cells through the *GZMA-F2R* interaction, indicating their potential as biomarkers for immunotherapy.

Our analysis of TCGA pan-cancer data and bulk ICIs data collectively suggested that TSPAN8⁺ chromatin remodeling fibroblasts showed unfavorable clinical implications in the prognosis of ovarian cancer, gastric cancer, and renal cell carcinoma and resistance to immunotherapy in gastric cancer. In addition, although not statistically significant, TSPAN8⁺ chromatin remodeling fibroblasts exhibited a high ratio of resistance to immunotherapy in other cancer types, including CRC and NSCLC. However, we found that TSPAN8⁺ chromatin remodeling fibroblasts were enriched in pCR tumors following immunotherapy in the scRNA-seq data of CRC. This finding suggests that the potential clinical implications of TSPAN8⁺ chromatin remodeling fibroblasts may be offset by other fibroblast subtypes or other cell types due to their relatively low cell number. Unfortunately, we only found one TSPAN8⁺ chromatin remodeling fibroblast in the NSCLC cohort, which warranted further validation in larger cohorts.

This study has some limitations. First, a canonical marker-based cell annotation may not be sufficient to distinguish fibroblasts from stellate cells, pericytes, and smooth muscle cells, thus additional cell annotation methods need to be combined to improve our cell annotation results. Second, although TCGA pan-cancer analysis indicates that several fibroblast subtypes are associated with metastasis in specific tissues, we have not yet discovered the metastasis mechanism for fibroblast subtypes in single-cell analysis, which requires additional research in larger metastasis cohorts. Third, we only demonstrate the presence of TSPAN8⁺ chromatin remodeling fibroblast in single-cell data, ST data, and multiplexed immunofluorescence assays. Further deep experimental validation may provide a comprehensive understanding of the mechanism of TSPAN8⁺ chromatin remodeling fibroblast involved in cancer progression.

In summary, we systematically delineated fibroblast subtypes across tissues, not only providing a more refined classification for myCAF and iCAF subtypes but also identifying the previously unidentified TSPAN8⁺ chromatin remodeling fibroblasts. Given the differentiation status, cell interactions, age distribution, and senescence level characteristics of the TSPAN8⁺ chromatin remodeling fibroblasts, it will be of great significance to facilitate its biomarker discovery and clinical applications. Notably, our study sheds light on the different roles of fibroblast subtypes in their interactions with other cells, which may guide the future development of targeted therapies based on fibroblast subtypes. Moreover, the integrated analysis of large-scale datasets will be a feasible way for refining subtypes in other cell types.

MATERIALS AND METHODS

Curation of scRNA-seq data

We curated a total of 73 publicly available scRNA-seq datasets generated from different platforms, including 10x, Smart-seq2, Sort-seq, Drop-seq, SCOPE-chip, inDrop, MARS-seq, Seq-well, scFTD-seq, and Microwell-seq. Single-cell transcriptome count matrices for each dataset were selected and downloaded. Surgical resection or biopsy samples from treatment-naïve patients were included in this study.

Preprocessing of scRNA-seq data

scRNA-seq data for each study was separately processed and analyzed using the “Seurat” package (v4.2.1) (31). Raw count matrices underwent filtration by removing cell barcodes with fewer than 200 expressed genes, more than 6 median absolute deviations of expressed genes, and more than 20% mitochondrial transcripts. The remaining cells of these samples were subsequently merged into an aggregate object, and their count expression value was log-normalized. The default 2000 highly variably genes were scaled before performing principal components analysis. In addition, we applied the batch effect correction algorithm, Harmony (32), to integrate gene-cell matrices from different samples. On the basis of the Harmony results, clusters were identified by the FindClusters function at a resolution of 0.8 and visualized using the nonlinear dimensional reduction method t-distributed stochastic neighbor embedding (t-SNE). Cluster-specific genes were identified using the FindAllMarkers function, and the resulting clusters were annotated to cell types by comparing these cluster-specific genes with canonical cell markers. To integrate fibroblasts across tissues, we extracted fibroblasts from each dataset based on cell annotation results. Through Harmony, fibroblasts from different datasets were integrated in a

tissue-specific manner, followed by the integration of all fibroblasts in a cross-tissue manner.

Fibroblast clustering and annotation

Fibroblast clustering was performed using the Seurat package, as described above. We used the clustree R package to visualize clusters at multiple resolutions (range from 0.1 to 1.2). After evaluating the number of clusters at each resolution, we determined 0.4 as the optimal resolution since there was minimal change in the number of clusters beyond that threshold (fig. S11). The DEGs for each sub-cluster were identified using the FindAllMarkers function with the following parameters: only.pos = T and min.pct = 0.25. Cluster-specific genes were defined as a group of genes that were significantly up-regulated in this cluster compared to other clusters. As Cords *et al.* (33) did, by examining the expression levels of cluster-specific genes using the DoHeatmap function in Seurat (v4.2.1), clusters showing similar gene expression patterns were merged into a unified fibroblast subtype. A clustered heatmap was generated from the z-scores of the mean expression values of the top 50 significant DEGs in each cluster. GO enrichment analysis of significant DEGs was performed using the “ClusterProfiler” package (v4.6.2) (34). Last, we took into account both significant DEGs and the GO enrichment results to annotate fibroblast subtypes, aligning with the hierarchical nomenclature proposed by Lavie *et al.* (3). The significant up-regulated DEGs of different fibroblast subtypes were defined as their specific marker genes.

The tissue and sample type preferences of fibroblast subtypes

We calculated the cell proportion of fibroblasts by dividing the number of fibroblasts in each sample type within each tissue by the total number of fibroblasts in that tissue. On the basis of the method from the work of Zheng *et al.* (35), we performed Fisher’s exact test to assess the tissue and sample type preferences of fibroblast subtypes. For each fibroblast subtype, we constructed a 2 by 2 contingency table for each combination of tissue *i* and sample type *j*. This table included the number of fibroblasts from tissue *i* in sample type *j*, the number from fibroblasts of tissue *i* in other sample types, the number of fibroblasts from non-*i* tissue in sample type *j*, and the number of fibroblasts from non-*i* tissue in other sample types. Fisher’s exact test was applied on this contingency table to obtain the OR and the corresponding *P* value. Furthermore, for each fibroblast subtype, we compared the differences in cell proportions between tumor samples and other sample types using the Wilcoxon rank-sum test.

Pathway analysis

We downloaded 3090 canonical pathways from the Molecular Signatures Database (MSigDB, v2023.1.Hs, <https://gsea-msigdb.org/gsea/msigdb>). Each pathway represents a specific biological state or process, curated from online pathway databases or the biomedical literature. DEGs between normal and adjacent fibroblasts, normal and tumor fibroblasts, and normal and metastatic fibroblasts were calculated using the FindMarkers function. Pathway enrichment analysis was assessed by a hypergeometric test, which was used to determine whether the gene sets of these canonical pathways showed statistically significant differences between different fibroblast groups. We further corrected the *P* values by the Benjamini-Hochberg method, and an adjusted *P* less than 0.05 was considered statistically significant.

Comparison with previous fibroblast subtypes

We gathered 12 previous fibroblast categorizations and corresponding marker genes generated from their single-cell data (table S3). To assess the concordance among fibroblast subtypes, we conducted pairwise gene set enrichment analysis on various fibroblast subtypes. Statistical significance among distinct fibroblast subtypes was determined through a hypergeometric test with Benjamini-Hochberg correction for *P* values. In addition, we evaluated the similarity of marker genes between different fibroblast subtypes using the Jaccard index, known as the ratio of gene set intersection to the union.

SCENIC analysis

We performed SCENIC analysis to reveal the TF expression within fibroblast subtypes using the pySCENIC pipeline (36). First, SCENIC inferred gene regulatory networks by identifying a set of genes that were coexpressed with TFs. Then, SCENIC filtered out indirect targets through cis-regulatory motif analysis. The processed gene regulatory networks with significant motif enrichment were identified as regulons. Last, the AUCell algorithm was used to calculate the activity score of each regulon in each cell. The regulon specificity score (RSS) comparing fibroblast subtypes was calculated by calcRSS function in R. The heatmap was used to visualize the RSS of the top three by fold difference between fibroblast subtypes.

Trajectory analysis

The Monocle2 (v2.20.0) (37) algorithm was applied to construct potential cell trajectories for fibroblast subtypes. Subsequently, the DEGs between fibroblast subtypes were used to sort cells in the trajectories. The DDRTree method and orderCells function were applied to reduce dimensionality and identify cell states, respectively. The branch with a high CytoTRACE score and a low EMT score was selected as the root (10, 38). The EMT scores were calculated using the AUCell algorithm based on a set of EMT-related gene sets downloaded from the MSigDB. In addition, we collected a set of gene signatures of resident fibroblasts obtained from the work of Heidegger *et al.* (15). The resident fibroblast scores were then calculated using the AUCell algorithm.

Correlation analysis

In trajectory analysis, we explored the relationship between different functional states of cell differentiation trajectories at the TF level. Specifically, we calculated the average RSS for each regulon in each state and performed Pearson correlation analysis to calculate correlation coefficients of RSS between different states. The heatmap was used to label and visualize the correlation coefficients. In addition, we collected a set of 526 signature genes overexpressed during cellular senescence from a meta-analysis (16). The cellular senescence scores were calculated using the AUCell algorithm. Among fibroblasts with available age information, we filtered out fibroblast subtypes with less than 100 cells in each tissue. Subsequently, we performed Pearson correlation analysis in each tissue to explore the changes in cellular senescence with age. Only fibroblast subtypes with *P* values less than 0.05 for the correlation analysis were visualized in the heatmap.

Cell interaction analysis

In each dataset, fibroblasts were classified into corresponding subtypes based on our defined fibroblast atlas. To investigate cell interactions between fibroblast subtypes and between fibroblast

subtypes and other cell types, we performed CellPhoneDB (v4.1.0) (39) analysis using the Python pipeline. CellPhoneDB inferred the potential interaction strength between two cell types based on the expression levels of ligand or receptor genes. We only selected ligand or receptor genes expressed in more than 10% of cell types. In addition, we randomly permuted the labels of all cells 1000 times to calculate the significance of each receptor-ligand pair. The *P* values less than 0.05 were considered indicative of specific interactions between cell types. Subsequently, we identified and visualized specific receptor-ligand interactions using CellChat (v1.1.3) (40), which systematically categorized receptor-ligand pairs into functionally related signaling pathways. After creating the CellChat objects, we set up a reference database based on the secretory signaling pathways and identified potential receptor-ligand interactions with default parameters.

Association of fibroblast subtypes with genomic alterations and metastasis

To investigate the associations between fibroblast subtypes with genomic alterations and metastasis, we collected TCGA pan-cancer data. Mutation annotations, gene expression profiles, and clinical information of TCGA pan-cancer samples were obtained from the cBioportal (<https://cbioportal.org/>). Functional gene mutations were defined as nonsilent in the variant classification. In each cancer type, we specifically considered gene mutations detected in at least 5% of the samples for subsequent analysis. Metastasis status was determined on the basis of the clinical stage of the patients. For a given fibroblast subtype, each tumor sample was scored according to its marker genes using the single-sample gene set enrichment analysis (ssGSEA). Subsequently, we explored the associations by calculating the difference in fibroblast subtype scores between samples with and without gene mutations or metastasis, and *P* values were computed by the Wilcoxon rank-sum test. The Benjamini-Hochberg method was applied to correct these *P* values, and adjusted *P* values less than 0.05 were considered statistically significant.

Survival analysis

We used the gene expression profiles from TCGA pan-cancer data, and only patients with available survival information were selected for further analysis. Fibroblast subtype scores for each sample were also calculated on the basis of their marker genes using the ssGSEA algorithm. We used univariate Cox regression analysis to evaluate the impact of fibroblast subtypes on OS, and *P* values were obtained through the log-rank test. For a given cancer type, fibroblast subtypes with a hazard ratio (HR) greater than 1 and *P* value < 0.5 were identified as risk factors, while fibroblast subtypes with HR less than 1 and *P* value < 0.5 were identified as protective factors. Subsequently, tumor samples from different cancer types were divided into high and low groups along the median of fibroblast subtype scores, and the Kaplan-Meier method was used to fit the survival curves between the high and low groups.

ST data analysis

We obtained three ST datasets from the Gene Expression Omnibus (<https://ncbi.nlm.nih.gov/geo/>), including GSE211956, GSE208253, and GSE203612, where GSE203612 was obtained from a pan-cancer study (41). The gene expression count matrices and corresponding histology images for the three ST datasets were downloaded and processed using the R package Seurat (v4.2.1). For each ST slice,

spots with fewer than 200 features were excluded from downstream analysis. The raw count matrices were normalized using the function SCTransform in Seurat.

Cell type deconvolution and TSPAN8⁺ chromatin remodeling fibroblast mapping

We used the RCTD (42) method to predict the underlying cell composition at each spot. This method required annotated scRNA-seq data as a reference. For the GSE203612 and GSE211956 datasets, we used their matched scRNA-seq datasets as references. In the case of the GSE208253 dataset, we chose the single-cell oral squamous cell carcinoma dataset GSE172577 as a reference. The preprocessing, clustering, and annotation for the three scRNA-seq datasets were consistent with the above description. Subsequently, RCTD created a spatial map of cell types by fitting the ST data at each spot to a linear combination of individual cell types. We ran the RCTD function with doublet mode and used the first cell type as the identity of the spot. The SpatialDimPlot function in Seurat was applied to project the spatial locations of fibroblasts directly on histology images. After that, we examined the expression of the *TSPAN8* gene at fibroblast-enriched spots using the SpatialFeaturePlot function in Seurat and further mapped the spatial locations of TSPAN8⁺ chromatin remodeling fibroblasts. Within the spots where the first cell type was identified as fibroblasts, we constructed a 2 by 2 contingency table based on whether it belonged to the TSPAN8⁺ chromatin remodeling fibroblast and whether the second cell type was endothelial cell. The Fisher's exact test was then applied on this contingency table to calculate the statistical significance.

Immunofluorescence staining

Human paired tumor samples and adjacent samples were obtained from the Harbin Medical University Cancer Hospital. All of the patients or their guardians provided written consent. The study was approved by the Ethics and Scientific Committees of the Harbin Medical University Cancer Hospital. Tumor tissues and adjacent tissues were fixed in 4% paraformaldehyde, embedded with Tissue-Tek optimal cutting temperature (OCT) (Sakura), and then stored at −80°C. Five-micrometer-thickness OCT sections were acetone-fixed and then blocked in phosphate-buffered saline with 50% goat serum and 0.4% Triton penetration. TSPAN8 (1:200, ABclonal) and FAP (1:100, Santa Cruz) were costaining in OCT sections. Primary antibodies were incubated overnight at 4°C. Alexa Fluor secondary antibodies (1:1000, Cell Signaling Technology) were incubated for 1 hour at room temperature. 4',6-Diamidino-2-phenylindole (1:500, Beyotime) was used to visualize nuclei. Confocal microscopy images were obtained by a Zeiss LSM 800 microscope.

Preprocessing of ICI data

We collected a total of 20 bulk transcriptome datasets on ICIs from public databases and general literature searches (table S4). The collected datasets, each with at least five responders and five nonresponders, were retained for further analysis. In therapy response information, patients with complete response, partial response, and stable disease were considered responders, while patients with progressive disease belonged to nonresponders. In addition, we also collected two scRNA-seq datasets to further evaluate the clinical relevance of fibroblast subtypes in the context of ICIs therapy. After initial filtering, we retained a total of 30 CRC samples and 15 NSCLC samples that received immunotherapy from the GSE205506

and GSE207422 datasets, respectively. For the scRNA-seq dataset GSE205506, we used pCR as responders ($n = 23$) and non-pCR as nonresponders ($n = 7$). For the scRNA-seq dataset GSE207422, we used partial response as responders ($n = 9$) and stable disease as nonresponders ($n = 6$). Single-cell data processing, integration, clustering, and annotation were conducted in the same manner as above. Our defined fibroblast atlas was used as a reference to annotate fibroblast subtypes via the “SingleR” package (v1.6.1) (43), which enabled the matching of marker genes to specific cell types.

Association of fibroblast subtypes with immunotherapy

In the bulk transcriptome data with ICI treatment information, we calculated fibroblast subtype scores for each sample using the ssGSEA algorithm. Next, each sample was classified into a high or low group based on the median of fibroblast subtype scores. For each fibroblast subtype, we used the Fisher’s exact test to evaluate whether the frequency of therapy response in the high group was significantly higher than the frequency in the low group ($P < 0.05$). In the scRNA-seq data, we used the Ro/e (44) method to assess the enrichment or depletion of each fibroblast subtype in the context of therapy response.

Additional statistical analysis

In addition to the statistical analyses described above, all other statistical analyses in this study were performed using R v4.2.2 (<https://r-project.org/>). The significance of P values is shown in the following way: $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; ns: no significance.

Supplementary Materials

The PDF file includes:

Figs. S1 to S31

Legends for tables S1 to S4

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S4

REFERENCES AND NOTES

- D. Lambrechts, E. Wauters, B. Boeckx, S. Aibar, D. Nittner, O. Burton, A. Bassez, H. Decaluwe, A. Pircher, K. Van den Eynde, B. Weynand, E. Verbeken, P. De Leyn, A. Liston, J. Vansteenkiste, P. Carmeliet, S. Aerts, B. Thienpont, Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat. Med.* **24**, 1277–1289 (2018).
- M. V. Plikus, X. Wang, S. Sinha, E. Forte, S. M. Thompson, E. L. Herzog, R. R. Driskell, N. Rosenthal, J. Biernaskie, V. Horsley, Fibroblasts: Origins, definitions, and functions in health and disease. *Cell* **184**, 3852–3872 (2021).
- D. Lavie, A. Ben-Shmuel, N. Erez, R. Scherz-Shouval, Cancer-associated fibroblasts in the single-cell era. *Nat. Cancer* **3**, 793–807 (2022).
- D. Ohlund, A. Handly-Santana, G. Biffi, E. Elyada, A. S. Almeida, M. Ponz-Sarvisé, V. Corbo, T. E. Oni, S. A. Hearn, E. J. Lee, I. I. Chio, C. I. Hwang, H. Tiriach, L. A. Baker, D. D. Engle, C. Feig, A. Kultti, M. Egeblad, D. T. Fearon, J. M. Crawford, H. Clevers, Y. Park, D. A. Tuveson, Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J. Exp. Med.* **214**, 579–596 (2017).
- E. Elyada, M. Bolisetty, P. Laise, W. F. Flynn, E. T. Courtois, R. A. Burkhart, J. A. Teinor, P. Belleau, G. Biffi, M. S. Lucito, S. Sivajothi, T. D. Armstrong, D. D. Engle, K. H. Yu, Y. Hao, C. L. Wolfgang, Y. Park, J. Preall, E. M. Jaffee, A. Califano, P. Robson, D. A. Tuveson, Cross-species single-cell analysis of pancreatic ductal adenocarcinoma reveals antigen-presenting cancer-associated fibroblasts. *Cancer Discov.* **9**, 1102–1123 (2019).
- A. Costa, Y. Kieffer, A. Scholer-Dahirel, F. Pelon, B. Bourachot, M. Cardon, P. Sirven, I. Magagna, L. Fuhrmann, C. Bernard, C. Bonneau, M. Kondratova, I. Kuperstein, A. Zinoviyev, A. M. Givel, M. C. Parrini, V. Soumelis, A. Vincent-Salomon, F. Mechta-Grigoriou, Fibroblast heterogeneity and immunosuppressive environment in human breast cancer. *Cancer Cell* **33**, 463–479.e10 (2018).
- Y. Kieffer, H. R. Hocine, G. Gentric, F. Pelon, C. Bernard, B. Bourachot, S. Lameiras, L. Albergante, C. Bonneau, A. Guyard, K. Tarte, A. Zinoviyev, S. Baulande, G. Zalcman, A. Vincent-Salomon, F. Mechta-Grigoriou, Single-cell analysis reveals fibroblast clusters linked to immunotherapy resistance in cancer. *Cancer Discov.* **10**, 1330–1351 (2020).
- A. Obradovic, D. Graves, M. Korre, Y. Wang, S. Roy, A. Naveed, Y. Xu, A. Luginbuhl, J. Curry, M. Gibson, K. Idrees, P. Hurley, P. Jiang, X. S. Liu, R. Uppaluri, C. G. Drake, A. Califano, Y. J. Kim, Immunostimulatory cancer-associated fibroblast subpopulations can predict immunotherapy response in head and neck cancer. *Clin. Cancer Res.* **28**, 2094–2109 (2022).
- M. B. Buechler, R. N. Pradhan, A. T. Krishnamurthy, C. Cox, A. K. Calviello, A. W. Wang, Y. A. Yang, L. Tam, R. Caothien, M. Roose-Girma, Z. Modrusan, J. R. Arron, R. Bourgon, S. Muller, S. J. Turley, Cross-tissue organization of the fibroblast lineage. *Nature* **593**, 575–579 (2021).
- H. Luo, X. Xia, L. B. Huang, H. An, M. Cao, G. D. Kim, H. N. Chen, W. H. Zhang, Y. Shu, X. Kong, Z. Ren, P. H. Li, Y. Liu, H. Tang, R. Sun, C. Li, B. Bai, W. Jia, Y. Liu, W. Zhang, L. Yang, Y. Peng, L. Dai, H. Hu, Y. Jiang, Y. Hu, J. Zhu, H. Jiang, Z. Li, C. Caulin, J. Park, H. Xu, Pan-cancer single-cell analysis reveals the heterogeneity and plasticity of cancer-associated fibroblasts in the tumor microenvironment. *Nat. Commun.* **13**, 6619 (2022).
- I. V. Pinchuk, K. T. Morris, R. A. Nofchissey, R. B. Earley, J. Y. Wu, T. Y. Ma, E. J. Beswick, Stromal cells induce Th17 during *Helicobacter pylori* infection and in the gastric tumor microenvironment. *PLOS ONE* **8**, e53798 (2013).
- D. Yang, J. Liu, H. Qian, Q. Zhuang, Cancer-associated fibroblasts: From basic science to anticancer therapy. *Exp. Mol. Med.* **55**, 1322–1332 (2023).
- V. Sharma, J. Letson, S. Furuta, Fibrous stroma: Driver and passenger in cancer development. *Sci. Signal.* **15**, eabg3449 (2022).
- G. Biffi, D. A. Tuveson, Diversity and biology of cancer-associated fibroblasts. *Physiol. Rev.* **101**, 147–176 (2021).
- I. Heidegger, G. Fotakis, A. Offermann, J. Goveia, S. Daum, S. Salcher, A. Noureen, H. Timmer-Bosscha, G. Schafer, A. Walenkamp, S. Perner, A. Beatovic, M. Moisse, C. Plattner, A. Krogsdam, J. Haybaeck, S. Soppe, S. Thaler, M. A. Keller, H. Klocker, Z. Trajanoski, D. Wolf, A. Pircher, Comprehensive characterization of the prostate tumor microenvironment identifies CXCR4/CXCL12 crosstalk as a novel antiangiogenic therapeutic target in prostate cancer. *Mol. Cancer* **21**, 132 (2022).
- K. Chatsirisupachai, D. Palmer, S. Ferreira, J. P. de Magalhães, A human tissue-specific transcriptomic analysis reveals a complex relationship between aging, cancer, and cellular senescence. *Aging Cell* **18**, e13041 (2019).
- M. B. Buechler, W. Fu, S. J. Turley, Fibroblast-macrophage reciprocal interactions in health, fibrosis, and cancer. *Immunity* **54**, 903–915 (2021).
- X. Y. Liu, Y. B. Liu, J. C. Xu, Y. F. Zhang, Y. Y. Ruan, Y. Zhao, L. F. Wu, J. W. Hu, Z. Zhang, M. J. He, T. Y. Chen, X. Y. Xu, J. W. Zhang, Y. Q. Zhang, P. H. Zhou, Single-cell transcriptomic analysis deciphers key transitional signatures associated with oncogenic evolution in human intramucosal oesophageal squamous cell carcinoma. *Clin. Transl. Med.* **13**, e1203 (2023).
- J. Li, C. Wu, H. Hu, G. Qin, X. Wu, F. Bai, J. Zhang, Y. Cai, Y. Huang, C. Wang, J. Yang, Y. Luan, Z. Jiang, J. Ling, Z. Wu, Y. Chen, Z. Xie, Y. Deng, Remodeling of the immune and stromal cell compartment by PD-1 blockade in mismatch repair-deficient colorectal cancer. *Cancer Cell* **41**, 1152–1169.e7 (2023).
- J. Hu, L. Zhang, H. Xia, Y. Yan, X. Zhu, F. Sun, L. Sun, S. Li, D. Li, J. Wang, Y. Han, J. Zhang, D. Bian, H. Yu, Y. Chen, P. Fan, Q. Ma, G. Jiang, C. Wang, P. Zhang, Tumor microenvironment remodeling after neoadjuvant immunotherapy in non-small cell lung cancer revealed by single-cell RNA sequencing. *Genome Med.* **15**, 14 (2023).
- L. Han, C. Seward, G. Leone, M. C. Ostrowski, Origin, activation and heterogeneity of fibroblasts associated with pancreas and breast cancers. *Adv. Cancer Res.* **154**, 169–201 (2022).
- S. Affo, A. Nair, F. Brundu, A. Ravichandra, S. Bhattacharjee, M. Matsuda, L. Chin, A. Filliol, W. Wen, X. Song, A. Decker, J. Worley, J. M. Caviglia, L. Yu, D. Yin, Y. Saito, T. Savage, R. G. Wells, M. Mack, L. Zender, N. Arpaia, H. E. Remotti, R. Rabadan, P. Sims, A. L. Leblond, A. Weber, M. O. Riener, B. R. Stockwell, J. Gaublot, J. M. Llovet, R. Kalluri, G. K. Michalopoulos, E. Seki, D. Sia, X. Chen, A. Califano, R. F. Schwabe, Promotion of cholangiocarcinoma growth by diverse cancer-associated fibroblast subpopulations. *Cancer Cell* **39**, 866–882.e11 (2021).
- Z. Wu, B. Uhl, O. Gires, C. A. Reichel, A transcriptomic pan-cancer signature for survival prognostication and prediction of immunotherapy response based on endothelial senescence. *J. Biomed. Sci.* **30**, 21 (2023).
- X. Wei, M. Liu, Z. Zheng, S. Yu, L. Huang, J. Ma, Y. Gao, Y. Peng, L. Chen, R. Tan, Z. She, L. Yang, Defective NCOA4-dependent ferroptosis in senescent fibroblasts retards diabetic wound healing. *Cell Death Discov.* **9**, 138 (2023).
- M. Higashiguchi, H. Murakami, H. Akita, S. Kobayashi, S. Takahama, Y. Iwagami, D. Yamada, Y. Tomimaru, T. Noda, K. Gotoh, Y. Doki, T. Yamamoto, H. Eguchi, The impact of cellular senescence and senescence-associated secretory phenotype in cancer-associated fibroblasts on the malignancy of pancreatic cancer. *Oncol. Rep.* **49**, 98 (2023).
- M. L. Taddei, L. Cavallini, G. Comito, E. Giannoni, M. Folini, A. Marini, P. Gandellini, A. Morandi, G. Pintus, M. R. Raspollini, N. Zaffaroni, P. Chiarugi, Senescent stroma promotes prostate cancer progression: The role of miR-210. *Mol. Oncol.* **8**, 1729–1746 (2014).

27. P. A. Toste, A. H. Nguyen, B. E. Kadera, M. Duong, N. Wu, I. Gawlas, L. M. Tran, M. Bikhchandani, L. Li, S. G. Patel, D. W. Dawson, T. R. Donahue, Chemotherapy-induced inflammatory gene signature and protumorigenic phenotype in pancreatic CAFs via stress-associated MAPK. *Mol. Cancer Res.* **14**, 437–447 (2016).
28. F. Xu, C. Hua, H. M. Tautenhahn, O. Dirsch, U. Dahmen, The role of autophagy for the regeneration of the aging liver. *Int. J. Mol. Sci.* **21**, 3606 (2020).
29. X. Cai, A. Guillot, H. Liu, Cellular senescence in hepatocellular carcinoma: The passenger or the driver? *Cells* **12**, 132 (2022).
30. N. Loret, N. Vandamme, J. De Coninck, J. Taminiau, K. De Clercq, G. Blancke, S. Jonckheere, S. Goossens, K. Lemeire, S. De Prijck, K. Verstaen, R. Seurinck, J. Van Dorpe, S. Weyers, H. Denys, K. Van de Vijver, B. N. Lambrecht, P. Tummers, Y. Saey, G. Berx, Distinct transcriptional programs in ascitic and solid cancer cells induce different responses to chemotherapy in high-grade serous ovarian cancer. *Mol. Cancer Res.* **20**, 1532–1547 (2022).
31. Y. Hao, S. Hao, E. Andersen-Nissen, W. M. Mauck III, S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L. M. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
32. I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P. R. Loh, S. Raychaudhuri, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
33. L. Cords, S. Tietscher, T. Anzeneder, C. Langwieder, M. Rees, N. de Souza, B. Bodenmiller, Cancer-associated fibroblast classification in single-cell and spatial proteomics data. *Nat. Commun.* **14**, 4294 (2023).
34. T. Wu, E. Hu, S. Xu, M. Chen, P. Guo, Z. Dai, T. Feng, L. Zhou, W. Tang, L. Zhan, X. Fu, S. Liu, X. Bo, G. Yu, clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2**, 100141 (2021).
35. L. Zheng, S. Qin, W. Si, A. Wang, B. Xing, R. Gao, X. Ren, L. Wang, X. Wu, J. Zhang, N. Wu, N. Zhang, H. Zheng, H. Ouyang, K. Chen, Z. Bu, X. Hu, J. Ji, Z. Zhang, Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* **374**, abe6474 (2021).
36. B. Van de Sande, C. Flerin, K. Davie, M. De Waegeneer, G. Hulselmans, S. Aibar, R. Seurinck, W. Saelens, R. Cannoodt, Q. Rouchon, T. Verbeiren, D. De Maeyer, J. Reumers, Y. Saey, S. Aerts, A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* **15**, 2247–2276 (2020).
37. X. Qiu, Q. Mao, Y. Tang, L. Wang, R. Chawla, H. A. Pliner, C. Trapnell, Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
38. G. S. Gulati, S. S. Sikandar, D. J. Wesche, A. Manjunath, A. Bharadwaj, M. J. Berger, F. Ilagan, A. H. Kuo, R. W. Hsieh, S. Cai, M. Zabala, F. A. Scheeren, N. A. Lobo, D. Qian, F. B. Yu, F. M. Dirbas, M. F. Clarke, A. M. Newman, Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* **367**, 405–411 (2020).
39. L. Garcia-Alonso, V. Lorenzi, C. I. Mazzeo, J. P. Alves-Lopes, K. Roberts, C. Sancho-Serra, J. Engelbert, M. Mareckova, W. H. Gruhn, R. A. Botting, T. Li, B. Crespo, S. van Dongen, V. Y. Kiselev, E. Prigmore, M. Herbert, A. Moffett, A. Chedotal, O. A. Bayraktar, A. Surani, M. Haniffa, R. Vento-Tormo, Single-cell roadmap of human gonadal development. *Nature* **607**, 540–547 (2022).
40. S. Jin, C. F. Guerrero-Juarez, L. Zhang, I. Chang, R. Ramos, C. H. Kuan, P. Myung, M. V. Plinius, Q. Nie, Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).
41. D. Barkley, R. Moncada, M. Pour, D. A. Liberman, I. Dryg, G. Werba, W. Wang, M. Baron, A. Rao, B. Xia, G. S. Franca, A. Weil, D. F. Delair, C. Hajdu, A. W. Lund, I. Osman, I. Yanai, Cancer cell states recur across tumor types and form specific interactions with the tumor microenvironment. *Nat. Genet.* **54**, 1192–1201 (2022).
42. D. M. Cable, E. Murray, L. S. Zou, A. Goeva, E. Z. Macosko, F. Chen, R. A. Irizarry, Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* **40**, 517–526 (2022).
43. D. Aran, A. P. Looney, L. Liu, E. Wu, V. Fong, A. Hsu, S. Chak, R. P. Naikawadi, P. J. Wolters, A. R. Abate, A. J. Butte, M. Bhattacharya, Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).
44. L. Zhang, X. Yu, L. Zheng, Y. Zhang, Y. Li, Q. Fang, R. Gao, B. Kang, Q. Zhang, J. Y. Huang, H. Konno, X. Guo, Y. Ye, S. Gao, S. Wang, X. Hu, X. Ren, Z. Shen, W. Ouyang, Z. Zhang, Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature* **564**, 268–272 (2018).

Acknowledgments: We thank all members of the Gu laboratory for critical comments on this manuscript. **Funding:** This work was supported by grants from the National Key Research and Development Program of China (grant 2023YFF1204600 to Y.G.), the National Natural Science Foundation of China (grant 32270710 to Y.G.), the Scientific Research Project of Provincial Scientific Research Institutes of Heilongjiang Province (grant CZKYF2024-1-A010 to Y.G.), and the Outstanding Youth Foundation of Heilongjiang Province of China (grant YQ2021H005 to Y.G.). **Author contributions:** K.L. and Y.G. conceived the research; K.L., X.S., and M.L. carried out the data collection; K.L. performed the data processing and analysis; K.L. wrote the manuscript; K.L. and Y.G. revised the manuscript; K.L. and Y.C. developed the web; H.H., N.Z., S.Z., S.S., K.X., and B.C. provided valuable suggestions; All authors read and approved the final manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper, the Supplementary Materials, and/or Fibroblast Atlas (<http://medsysbio.org:3838/fibroblasts/>).

Submitted 12 January 2024
 Accepted 26 February 2025
 Published 2 April 2025
 10.1126/sciadv.ado0173