Spatial Positioning and Matrix Programs of Cancer-Associated Fibroblasts Promote T-cell Exclusion in Human Lung Tumors

John A. Grout, Philémon Sirven, Andrew Leader, Shrisha Maskey, Eglantine Hector, Isabelle Puisieux, Fiona Steffan, Evan Cheng, Navpreet Tung, Mathieu Maurin, et al.

RESEARCH ARTICLE

Spatial Positioning and Matrix Programs of Cancer-Associated Fibroblasts Promote T-cell Exclusion in Human Lung Tumors

John A. Grout[1,2,3], Philemon Sirven[4,5], Andrew M. Leader[1,2,3], Shrisha Maskey[1,2,3], Eglantine Hector[4,5], Isabelle Puisieux[4,5], Fiona Steffan[4,5], Evan Cheng[1,2,3], Navpreet Tung[1,2,3], Mathieu Maurin[4,5], Romain Vaineau[4,5], Lea Karpf[1,2,3], Martin Plaud[1,2,3], Anne-Laure Begue[4,5], Koushik Ganesh[4,5], Jérémy Mesple[4,5], Maria Casanova-Acebes[1,2,3], Alexandra Tabachnikova[1,2,3], Shilpa Keerthivasan[6], Alona Lansky[1,2,3], Jessica Le Berichel[1,2,3], Laura Walker[1,7], Adeeb H. Rahman[1,7,8], Sacha Gnjatic[1,2,7,9], Nicolas Girard[10,11], Marine Lefevre[12], Diane Damotte[13], Julien Adam[14], Jerome C. Martin[1,2,3], Andrea Wolf[15], Raja M. Flores[15], Mary Beth Beasley[16], Rachana Pradhan[17], Soren Muller[17], Thomas U. Marron[1,2,9], Shannon J. Turley[6], Miriam Merad[1,2,3,7], Ephraim Kenigsberg[1,8], and Hélène Salmon[1,2,3,4,5]

ABSTRACT It is currently accepted that cancer-associated fibroblasts (CAF) participate in T-cell exclusion from tumor nests. To unbiasedly test this, we used single-cell RNA sequencing coupled with multiplex imaging on a large cohort of lung tumors. We identified four main CAF populations, two of which are associated with T-cell exclusion: (i) MYH11$^+$αSMA$^+$ CAF, which are present in early-stage tumors and form a single cell layer lining cancer aggregates, and (ii) FAP$^+$αSMA$^+$ CAF, which appear in more advanced tumors and organize in patches within the stroma or in multiple layers around tumor nests. Both populations orchestrate a particular structural tissue organization through dense and aligned fiber deposition compared with T cell–permissive CAF. Yet they produce distinct matrix molecules, including collagen IV (MYH11$^+$αSMA$^+$ CAF) and collagen XI/XII (FAP$^+$αSMA$^+$ CAF). Hereby, we uncovered unique molecular programs of CAF driving T-cell marginalization, whose targeting should increase immunotherapy efficacy in patients bearing T cell–excluded tumors.

SIGNIFICANCE: The cellular and molecular programs driving T-cell marginalization in solid tumors remain unclear. Here, we describe two CAF populations associated with T-cell exclusion in human lung tumors. We demonstrate the importance of pairing molecular and spatial analysis of the tumor microenvironment, a prerequisite to developing new strategies targeting T cell–excluding CAF.

INTRODUCTION

Lung cancer is the leading cause of cancer-related deaths worldwide, accounting for roughly 1.6 million deaths per year, with non–small cell lung carcinoma (NSCLC) being the most prevalent form (1). The partial success of immune-checkpoint blockade in only a subset of patients with NSCLC underscores the need for a better understanding of the determinants controlling antitumor immunity (2). In addition to high tumor mutational burden and PD-L1 expression levels in the tumor, CD8$^+$ T-cell density has been shown as a predictor of immunotherapy response (3, 4). By analyzing T-cell localization within the tumor, recent studies have revealed the importance of T-cell infiltration into tumor nests relative to the surrounding stroma (3, 5, 6). Understanding the mechanisms regulating T-cell exclusion is therefore crucial to improve T cell–based therapies and patient outcomes.

Using real-time imaging of T-cell dynamics in human NSCLC, we previously found that dense fibers oriented parallel to the tumor–stroma interface form a barrier around the tumor mass and limit T-cell contact with tumor cells (7). However, the cellular sources and their extracellular matrix (ECM) programs remain unknown. Fibroblasts are known to shape lymphocyte compartmentalization in secondary lymphoid organs, in which they produce distinct sets of chemokines and a complex ECM conduit system that serves as a scaffold along which dendritic cells and lymphocytes migrate and engage (8–10). Although the role of fibroblasts in restricting immune cell localization is well established in spleen and lymph nodes, only recently has the tumor stroma emerged as a player in regulating local immune responses (11–14).

Given the growing evidence indicating that cancer-associated fibroblasts (CAF) can regulate tumor immunity and progression (11–14), CAF are becoming an important target for cancer treatment. TGFβ blockade and NOX4 inhibition were shown to act on CAF and facilitate T-cell infiltration, leading to better responses to anti–PD-1/PD-L1 treatment in murine cancer models (6, 15, 16). Yet modulating and depleting CAF have led to opposite results in other tumor systems (16, 17) and have not yet managed to achieve clinical benefit in human cancer (18, 19). How to manipulate fibroblast properties for therapeutic purposes remains challenging, largely due to our limited understanding of the tumor CAF compartment and the mechanisms by which distinct CAF populations modulate antitumor immunity, including immune cell spatial organization.

[1]The Precision Immunology Institute, Icahn School of Medicine at Mount Sinai, New York, New York. [2]The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York. [3]Department of Oncological Sciences, Icahn School of Medicine at Mount Sinai, New York, New York. [4]Institut Curie, INSERM, U932, Equipe Leader Fondation ARC, Paris, France. [5]PSL Research University, Paris, France. [6]Department of Cancer Immunology, Genentech, Inc., South San Francisco, California. [7]Human Immune Monitoring Center, Icahn School of Medicine at Mount Sinai, New York, New York. [8]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York. [9]Division of Hematology/Oncology, Icahn School of Medicine at Mount Sinai, New York, New York. [10]Thorax Institute Curie Montsouris, Institut Curie, Paris, France. [11]UVSQ, Paris Saclay University, Versailles, France. [12]Department of Pathology, Institut Mutualiste Montsouris, Paris, France. [13]Department of Pathology, Assistance Publique–Hôpitaux de Paris, Paris Cité University, Paris, France. [14]Department of Pathology, Paris Saint-Joseph Hospital, Paris, France. [15]Department of Thoracic Surgery, Icahn School of Medicine at Mount Sinai, New York, New York. [16]Department of Pathology, Icahn School of Medicine at Mount Sinai, New York, New York. [17]Department of Bioinformatics and Computational Biology, Genentech, Inc., South San Francisco, California.

Note: E. Kenigsberg and H. Salmon are the co-senior authors of this article.

Current address for H. Salmon: Centre de Recherche en Transplantation et Immunologie, UMR 1064, INSERM, Nantes University, Nantes Laboratoire d'Immunologie, Centre Hospitalier Universitaire de Nantes, Nantes, France.

Corresponding Authors: Hélène Salmon, Immunity and Cancer, Institute Curie, 26 rue d'Ulm, Paris 75005, France. Phone: 33 1 56 24 63 29; E-mail: helene.salmon@curie.fr; and Ephraim Kenigsberg, Icahn School of Medicine at Mount Sinai, 1 Gustave L. Levy Place, Box 1498, New York, NY 10029-6504. Phone: 212-659-8540; E-mail: ephraim.kenigsberg@mssm.edu

Cancer Discov 2022;12:1–20

doi: 10.1158/2159-8290.CD-21-1714

The initial characterization of functional heterogeneity of CAF included a description of inflammatory CAF (iCAF) and myofibroblastic CAF (myCAF) in mouse models of pancreatic cancer (20). Transcriptional signatures of these distinct CAF phenotypes have subsequently been found in human pancreatic and breast cancer (21, 22), as well as an additional subset, antigen-presenting CAF (apCAF; ref. 22). iCAF are described as being found distal from the tumor site with a secretory phenotype, whereas myCAF are characterized by activation and contractility genes and their close proximity to tumor cells (20). Prior studies have used single-cell RNA sequencing (scRNA-seq) to profile CAF in various human cancers, including NSCLC (23–26), bladder (27), pancreas (22, 28), breast (21), head and neck (29), and liver (30). Although the diversity of CAF is increasingly appreciated, the molecular programs of human fibroblast subsets and their discrete functional contributions to the tumor organization and T-cell compartmentalization have not been resolved.

We reasoned that pairing scRNA-seq profiling with high-resolution spatial mapping would enable unbiased identification of CAF transcriptional subsets and uncover their spatial organization in the context of the tumor microenvironment (TME). Our scRNA-seq analysis on 15 surgically resected NSCLC samples along with 12 paired adjacent tissue samples identified novel CAF subpopulations that we validated by profiling 35 tumors by multiplexed IHC (31). We analyzed the spatial organization of the stromal and immune cell populations and revealed two CAF subsets with distinct ECM programs that were associated with $CD3^+$ and $CD8^+$ T-cell exclusion from the tumor nests. Importantly, by applying high-resolution histologic profiling on a large NSCLC cohort, our study characterizes both intratumor and intertumor CAF and T-cell heterogeneity.

RESULTS

Paired scRNA-seq and IHC Analysis Identifies Four CAF Populations with Distinct Transcriptional Profiles and Structural Organization in Human NSCLC

To characterize the stromal cell compartment in NSCLC in an unbiased way, we profiled nonimmune, nontumor/epithelial cells isolated from 15 NSCLC samples and 12 paired adjacent tissue samples using the 10X Genomics scRNA-seq platform (Fig. 1A; Supplementary Table S1). Using flow cytometry and mass cytometry by time of flight (CyTOF), we optimized the digestion and sorting protocols to maximize stromal cell recovery while preserving cell integrity (Methods; Supplementary Fig. S1A and S1B). A total of 33,742 cells were sequenced that contained 31,402 stromal cells after excluding contaminating immune cells, epithelial cells, and cells not passing quality control (Supplementary Table S2). Using an unsupervised clustering method, which integrates samples over different conditions and patients while modeling background noise (32, 33), we identified 28 clusters, including 24 stromal cell clusters of variable abundance shared among samples (Supplementary Fig. S1C; Supplementary Table S2) as well as four clusters, containing either contaminating immune cells or cells with high mitochondrial content, that were excluded from future analysis. mRNA counts [unique molecular identifiers (UMI)] per cluster and mitochondrial content per sample were similar (Supplementary Fig. S1C and S1D).

To unbiasedly dissect cell identities, we analyzed the mRNA counts of variably expressed genes across the 24 stromal cell clusters (Supplementary Fig. S1E). The cell clusters represented three major stromal cell compartments and expressed well-reported lineage markers: fibroblasts ($PDGFRA^+$, $MMP2^+$), endothelial cells [EC; including both blood ($CLDN5^+$, $PECAM1^+$) and lymphatic ($TFF3^+$, $PROX1^+$) EC], and perivascular cells [PvC; including pericytes ($MCAM^+$, $COX4I2^+$) and smooth muscle (SM) cells ($MCAM^+$, DES^+); Fig. 1B; Supplementary Fig. S1F; Supplementary Table S3]. Blood EC clusters included arteries, venules, tip cells, as well as two lung capillary subsets recently described as aerocytes and general capillaries (ref. 34; Supplementary Fig. S2A). The PvC clusters enriched in tumor lesions included tumor pericytes, which expressed high amounts of RGS5 and multiple collagens (COL1A1, COL3A1, COL6A3), and a cluster expressing multiple immunomodulatory genes including *CCL19* and *CCL21* (Supplementary Fig. S2B). Notably, IHC showed that the $MCAM^+$ cells were restricted to vascular areas and were not found in the rest of the stroma (Supplementary Fig. S2C).

Further dissection of fibroblast populations identified multiple subsets with distinct transcriptional profiles and uneven abundances in the tumor lesion or the adjacent tissue (Fig. 1C and D; Supplementary Table S4). Based on this scRNA-seq analysis, we identified genes associated with each cluster and defined antibody panels (Supplementary Table S5) that enabled further characterization by multiplexed IHC (Fig. 1E and F; Supplementary Fig. S3A and S3B). Two clusters enriched in the adjacent lung tissue were characterized by coexpression of *MME* (CD10), *FIGF* (VEGFD), and *FGFR4* (Fig. 1C) and were annotated as alveolar fibroblasts (alv. fib.) based on their specific localization to the lung alveoli by IHC (Fig. 1E; Supplementary Fig. S3A). Interestingly, one of these clusters expressed high levels of inflammatory genes, including *IL6* and *ICAM1*, and was thus referred to as inflamed alv. fib. (inf. alv. fib.; Fig. 1C). Another cluster enriched in the adjacent lung was annotated as $PI16^+$ fibroblasts ($PI16^+$ fib.) based on its coexpression of *PI16*, *CD34*, and leptin receptor (*LEPR*), localization to the blood vessel adventitia, and similarity to the universal $PI16^+$ fib. described in Buechler et al. (Fig. 1C and E; Supplementary Fig. S3A; refs. 35–37). The last adjacent tissue cluster, CLU^+ fibroblasts, was characterized by high expression of clusterin (*CLU*; Fig. 1C).

Fibroblast clusters enriched in tumor samples were annotated as CAF. One CAF cluster displayed an expression profile similar to that of alv. fib., including expression of the broad adjacent tissue fibroblast marker alcohol dehydrogenase 1B (*ADH1B*) and lower expression of the canonical CAF marker *FAP*, and was referred to as $ADH1B^+$ CAF (Fig. 1C and E; Supplementary Fig. S3A). $ADH1B^+$ CAF could be distinguished from alv. fib. in IHC by their lack of CD10 expression and localization in the tumor lesion (Fig. 1C and E; Supplementary Fig. S3A–S3C). Three clusters showed strong expression of canonical activated CAF markers *FAP*, *POSTN, LRRC15*, and *GREM1* (23, 28) and were denoted as FAP^+ CAF (Fig. 1C and E; Supplementary Fig. S3B). Another common CAF

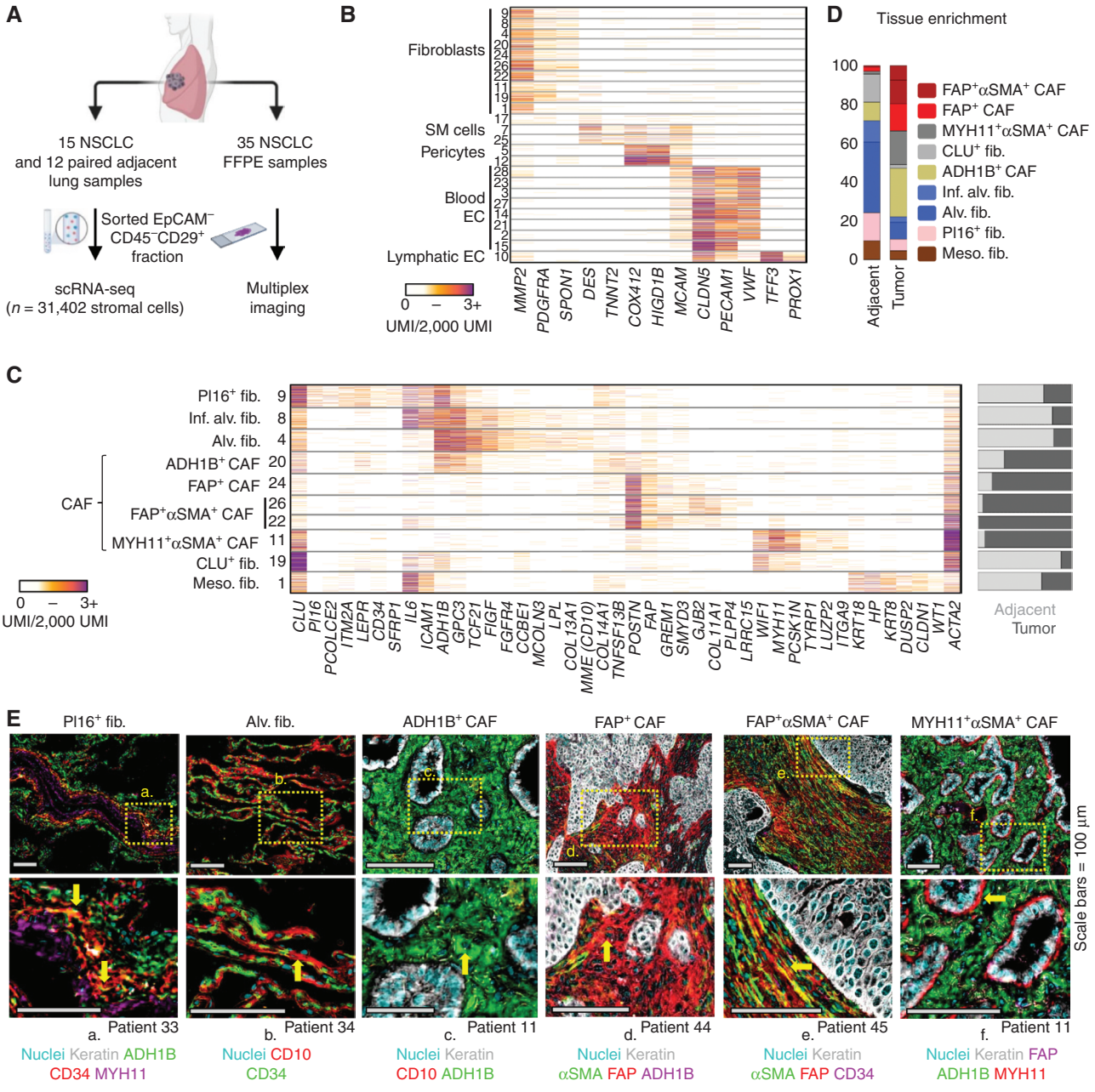


Figure 1. Paired scRNA-seq and IHC analysis identifies four CAF populations with distinct transcriptional profiles and structural organization in human NSCLC. **A,** Tissue processing workflow for scRNA-seq and IHC imaging of formalin-fixed, paraffin-embedded (FFPE) samples. **B,** scRNA-seq mRNA counts (UMIs) per cell (rows) of select stromal lineage marker genes (columns). Fibroblast, SM, pericyte, blood, and lymphatic EC clusters are identified based on the expression of marker genes such as *PDGFRA*, *DES*, *COX4I2*, *PECAM1*, and *TFF3*, respectively. All cells displayed in this figure, and all subsequent similar scRNA-seq figures, were downsampled to 2,000 UMI. **C,** Extended gene lists highlighting gene expression profiles between the fibroblast subsets along with differing propensities for enrichment (right bar plot) in tumor (dark gray) or adjacent tissue (light gray). **D,** Averaged fibroblast composition in adjacent and tumor samples across all patients. The bar graph depicts the percentage of cells from each fibroblast subset among all fibroblasts. CLU⁺ fib., CLU⁺ fibroblast. **E,** FFPE NSCLC sections were stained for fibroblast markers identified in scRNA-seq results. All the scRNA-seq–based fibroblast clusters (**D**) were detected utilizing IHC except meso. fib. and CLU⁺ fib., which were not in the scope of this study. Arrows highlight cells of interest (PI16⁺ fib.: CD34⁺ADH1B⁺MYH11^{neg}; Alv. fib.: CD10⁺CD34^{neg}; ADH1B⁺ CAF: ADH1B⁺CD10^{neg}; FAP⁺ CAF: FAP⁺ADH1B^{neg}αSMA^{neg}; FAP⁺αSMA⁺ CAF: FAP⁺αSMA⁺CD34^{neg}; and MYH11⁺αSMA⁺ CAF: MYH11⁺FAP^{neg}ADH1B^{neg}). See Supplementary Fig. S3 for other stainings. All scale bars, 100 µm. **F,** IHC staining presentation for the main identified fibroblast and CAF clusters.

marker, *ACTA2* (αSMA; ref. 38), was differentially expressed among the FAP$^+$ CAF (Fig. 1C and E; Supplementary Fig. S3B), and clusters with high *ACTA2* expression were designated as FAP$^+$αSMA$^+$ CAF. Notably, a cluster that was highly enriched in a single patient (Supplementary Table S2) shared both fibroblast genes (*PDGFRA*, *MMP2*, *COL1A1*, and *BGN*) and mesothelial cell genes, such as keratins and *WT1*, and was therefore designated as mesothelial-like fibroblasts (meso. fib.; ref. 39; Fig. 1C). An additional CAF cluster, MYH11$^+$αSMA$^+$ CAF, clearly distinct from the other CAF subsets, was characterized by the expression of myosin heavy chain 11 (*MYH11*), *ACTA2*, and intermediate levels of *CD34* while lacking *ADH1B* and *FAP* expression (Fig. 1C). Histologic analysis of matched formalin-fixed, paraffin-embedded (FFPE) tumor samples revealed an MYH11$^+$αSMA$^+$CD34$^+$ADH1B^{neg}FAP^{neg} cell population observed as a single layer of elongated CAF-encapsulating tumor nests, which is in contrast to ADH1B$^+$ CAF and FAP$^+$ CAF that are spread throughout the stroma (Fig. 1E and F; Supplementary Fig. S3A). CyTOF confirmed the presence of the main fibroblast subsets identified through scRNA-seq, including alv. fib., PI16$^+$ fib., MYH11$^+$αSMA$^+$ CAF, and FAP$^+$αSMA$^+$ CAF (Fig. 2A).

Further analysis of ADH1B$^+$ CAF revealed a subset of cells that expressed high levels of the T cell–attracting and T-cell retention genes *CCL19*, *CCL21*, and *VCAM1*, reminiscent of fibroblastic reticular cells present in secondary lymphoid organs (ref. 40; Fig. 2B). IHC staining of CCL19 confirmed the specific localization of these fibroblasts to tertiary lymphoid structures (TLS), with clear preferential enrichment for the T-cell zone (Fig. 2B). In some cases, the B-cell zone was delineated by podoplanin (PDPN) expression, which marks follicular dendritic cells that were not captured by scRNA-seq, likely due to their low abundance (Fig. 2B). Given the report of MHCII-expressing CAF in human NSCLC tumors (41), we considered whether these CCL19$^+$ TLS CAF may be involved in antigen presentation to T cells. The scRNA-seq data showed MHCII gene expression among CAF, with CCL19$^+$ADH1B$^+$ CAF expressing the highest levels, suggesting that CCL19$^+$ TLS CAF are involved in antigen presentation to T cells (Fig. 2C). Notably, we compared CCL19$^+$ CAF with other cells of the TME, including EC and meso. fib. captured in our study and immune cells from our NSCLC immune cell dataset (32). This analysis showed that MHCII expression in CCL19$^+$ CAF is orders of magnitude lower than in EC, as well as in the classic MHCII-expressing cells, type 1 dendritic cells (DC1). Nevertheless, the high density of CCL19$^+$ CAF in TLS may contribute to effective antigen presentation in these areas (Fig. 2D).

Taken together, our combined IHC and single-cell analysis has defined diverse fibroblast populations with distinct molecular and spatial patterns in human NSCLC. By enriching for stromal cells from a large NSCLC cohort, we achieved highly granular scRNA-seq characterization and uncovered CAF populations undescribed to date, including a single layer of MYH11$^+$αSMA$^+$ CAF bordering tumor cells in a fraction of NSCLC lesions. The four CAF subsets described here expand upon the iCAF, myCAF, and apCAF profiles described in pancreatic tumors (20, 22). Our analysis suggests that in human lung tumors, myCAF include FAP$^+$ CAF, FAP$^+$αSMA$^+$ CAF, and MYH11$^+$αSMA$^+$ CAF, highlighting the transcriptomic and spatial complexity of this population (Fig. 2E). The full expression profiles of the different fibroblast populations and histology data are available at https://scdissector.org/grout, allowing for multidimensional exploration.

ADH1B$^+$ CAF and FAP$^+$ CAF Stratify NSCLC into Two Main Stromal Patterns Associated with Tumor Stage and Histology

Analysis of the fibroblast composition as determined by scRNA-seq indicated that low-stage tumors were dominated by ADH1B$^+$ CAF with or without MYH11$^+$αSMA$^+$ CAF, whereas higher-stage tumors were enriched for FAP$^+$ CAF and FAP$^+$αSMA$^+$ CAF (Fig. 3A and B; Supplementary Fig. S4A). To test the dichotomy between ADH1B$^+$ and FAP$^+$ CAF enrichment, we leveraged a larger cohort of 35 patient FFPE samples and quantified the tumor area covered by ADH1B and FAP using IHC. This unbiased analysis showed that the stroma of NSCLC is significantly dominated by either ADH1B$^+$ or FAP$^+$ CAF (hypergeometric test, $P = 0.008$; Fig. 3C; Supplementary Table S6). MYH11$^+$αSMA$^+$ CAF were observed in half of ADH1B$^+$ CAF–rich samples (9/18; Fig. 3C), but they were not observed in FAP$^+$ CAF–rich samples, corroborating the scRNA-seq analysis that showed a correlation between ADH1B$^+$ CAF and MYH11$^+$αSMA$^+$ CAF (Fig. 3A). FAP$^+$ CAF–rich samples showed highly variable stroma coverage by FAP$^+$αSMA$^+$ CAF (Fig. 3B; Supplementary Fig. S4B), which is in line with the variable *ACTA2* expression seen across the scRNA-seq FAP$^+$ CAF clusters (Fig. 1C). Although FAP$^+$αSMA$^+$ CAF could be found throughout the stroma, they were organized as cell layers lining tumor nests in a fraction of tumors (Supplementary Fig. S4C).

To study the transcriptional programs behind ADH1B$^+$ CAF and FAP$^+$ CAF and to better understand their relationship to adjacent tissue fibroblasts, we analyzed gene expression covariance patterns across ADH1B$^+$ CAF, FAP$^+$ CAF, alv. fib., and PI16$^+$ fib. We identified groups of coexpressed genes (gene modules) with distinct expression patterns across these fibroblast populations (Supplementary Fig. S4D). FAP$^+$ CAF upregulated activation genes (modules 14 and 15) including multiple collagen genes (*COL1A1* and *COL3A1*) that contribute to tissue stiffness (42) and other ECM genes such as biglycan (*BGN*) that can promote tissue mineralization (ref. 43; Supplementary Fig. S4D). FAP$^+$ CAF expressed low levels of the alv. fib. genes, including the fibroblast transcription factor *TCF21* (44), the marker *MME* (CD10), as well as the ECM gene elastin (*ELN*), which is critical for normal lung physiology (ref. 45; Fig. 3D). ADH1B$^+$ CAF expressed intermediate levels of FAP$^+$ CAF activation genes (Fig. 3D), and a subset of samples showed a spatial gradient of ADH1B$^+$ CAF to FAP$^+$ CAF from the invasive margin to the tumor center (Fig. 3E) with some cells coexpressing both markers, suggesting that ADH1B$^+$ CAF represent a range of lowly activated fibroblasts. ADH1B$^+$ CAF also shared genes with both PI16$^+$ fib. and alv. fib., which may point toward both lung fibroblast types as their potential cellular sources (Fig. 3D; Supplementary Fig. S5A). Interestingly, the scRNA-seq data showed that ADH1B$^+$ CAF cells express a gradient of FAP$^+$ CAF and PI16$^+$ fib. genes from cells with high expression of PI16 genes and low FAP genes to cells with low PI16 genes and high FAP genes (Fig. 3F). This further supports the hypothesis that

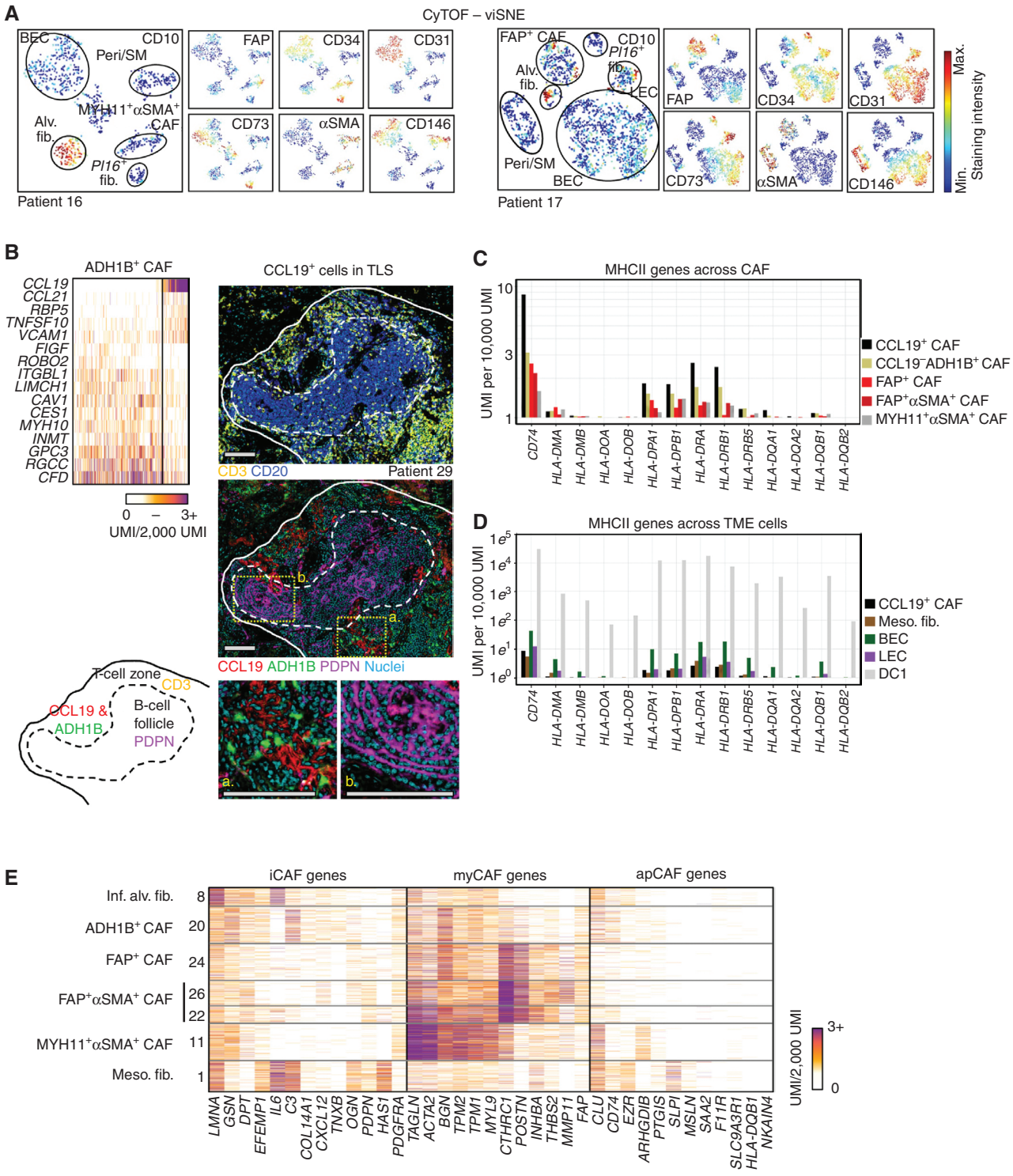
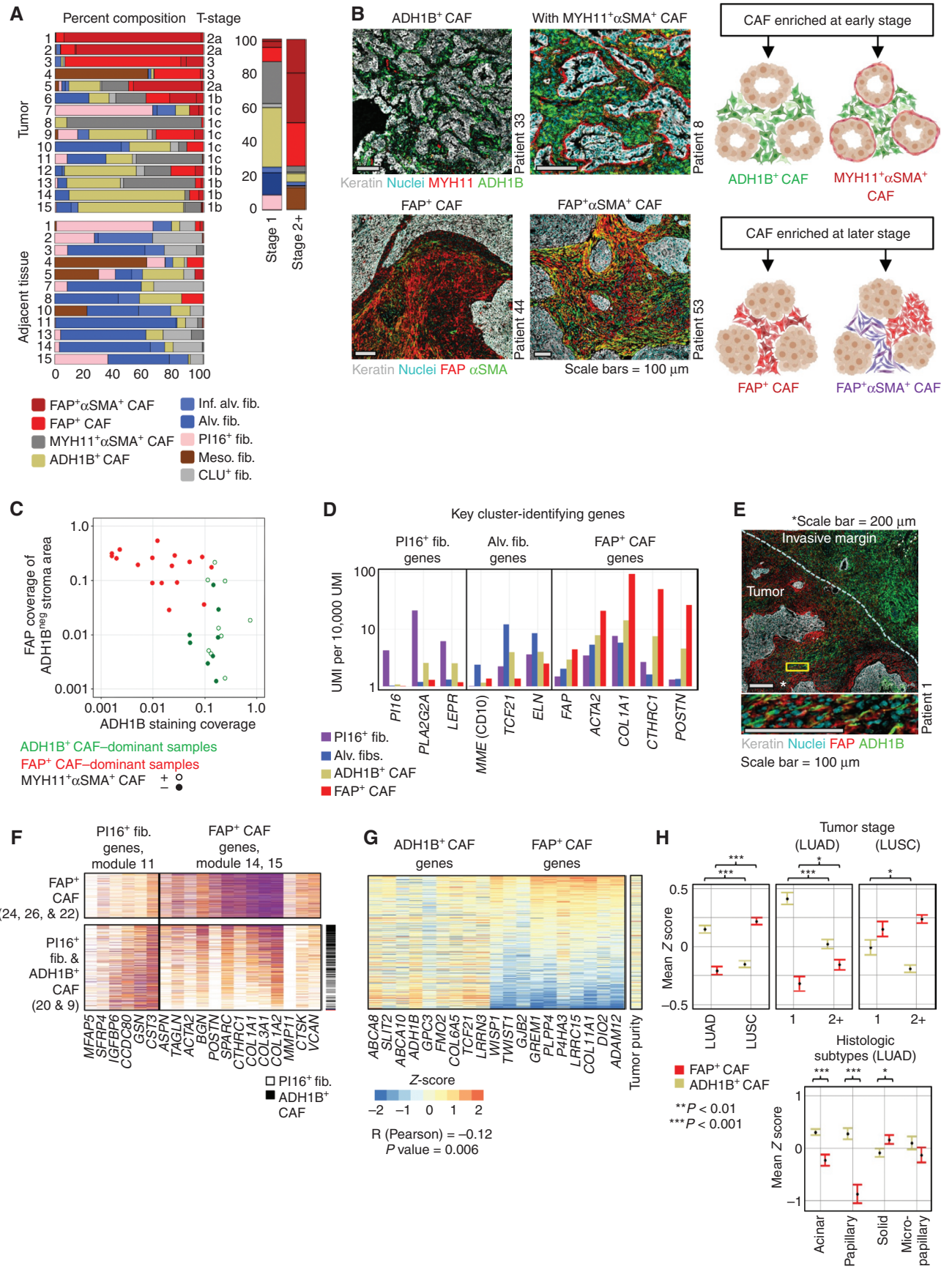


Figure 2. Further characterization of CAF subsets in human NSCLC. **A,** Stromal cell populations visualized with visualization of t-distributed stochastic neighbor embedding (viSNE) in CyTOF. EC, PvC, and multiple fibroblast subsets can be distinguished with relatively few markers (CD10, CD31, CD34, CD73, FAP, CD146, and αSMA). BEC, blood EC; LEC, lymphatic EC; Peri, PvC. **B,** Top left, highlighting CCL19-expressing cells within ADH1B+ CAF. These cells expressed high amounts of *CCL19*, *CCL21*, and *VCAM1* and low levels of certain ADH1B+ CAF genes such as MYH10 and GPC3 (bottom and right). Multiplex IHC of a representative TLS. PDPN and CD20 mark follicular dendritic cells and B cells, respectively, in the B-cell follicle, whereas the T-cell zone is identified with CD3 staining. CCL19 and ADH1B staining show ADH1B+ fibroblasts surrounded by the secreted chemokine CCL19, specifically in the T-cell zone. All scale bars, 100 μm. **C,** Average expression of MHCII genes in each CAF subset. **D,** Average MHCII gene expression in classic antigen-presenting cells, DC1, EC, and meso. fib. **E,** myCAF, iCAF, and apCAF gene signatures (20, 22) projected onto NSCLC CAF clusters.



ADH1B$^+$ CAF are a lowly activated form of fibroblast and may derive from PI16$^+$ fib.

Leveraging our scRNA-seq datasets, we created gene signatures for ADH1B$^+$ CAF and FAP$^+$ CAF by selecting for genes with highly specific expression in their corresponding CAF populations in contrast with all other cell types. With these signatures, we scored The Cancer Genome Atlas (TCGA) lung adenocarcinoma (LUAD) samples by their expression of ADH1B$^+$ CAF and FAP$^+$ CAF genes and revealed an anticorrelation between the two scores ($P = 0.006$; Fig. 3G; Supplementary Fig. S5B; Supplementary Table S7), supporting the two distinct CAF profiles observed across patients with NSCLC in scRNA-seq and histology (Fig. 3A–C; ref. 46). Analysis of tumor purity, estimated by tumor nuclei abundance, did not reveal clear association with ADH1B$^+$ CAF or FAP$^+$ CAF genes (Fig. 3G), confirming that contaminating adjacent tissue was not a major contributor to the ADH1B$^+$ CAF signal. Further analysis of TCGA data showed that ADH1B$^+$ CAF genes were significantly increased in stage 1 tumors, LUAD, and the papillary LUAD subtype, whereas FAP$^+$ CAF were enriched in later-stage tumors, lung squamous cell carcinoma (LUSC), and the LUAD solid subtype (Fig. 3H). LUAD across tumor stages confirmed our observation that ADH1B$^+$ CAF and FAP$^+$ CAF were correlated with lower- and higher-stage tumors, respectively. Similar associations were observed in our in-house FFPE cohort (Supplementary Table S1). Altogether, we showed that ADH1B$^+$ and FAP$^+$ CAF phenotypes were correlated with tumor stage and clinically relevant histologic subtypes (46, 47), suggesting that molecular characterization of fibroblasts could refine the clinical categorization of NSCLC tumors.

To validate the diverse transcriptional programs of CAF observed in our dataset, we studied their expression in four additional public scRNA-seq datasets in NSCLC (23–26). This analysis identified two fibroblast subsets expressing ADH1B$^+$ and FAP$^+$ CAF genes, confirming their enrichment in early- and late-stage NSCLC tumors, respectively (Supplementary Fig. S6A–S6C; Supplementary Table S8). FAP$^+$ cells display variable expression levels of the FAP$^+$αSMA$^+$ CAF program (Supplementary Fig. S6D), reflecting the observed heterogeneity within FAP$^+$ cells (Supplementary Fig. S4D). In Kim et al. (24), we identified a population resembling MYH11$^+$αSMA$^+$ CAF expressing multiple MYH11$^+$αSMA$^+$ CAF genes including *MYH11*, *COL4A1*, and *COL4A2* (Supplementary Fig. S6E). As expected, these cells were not found in the other studies, which were predominantly composed of stage 2+ tumors and squamous cell carcinomas. Altogether, the external datasets examined validated the dominant CAF populations in NSCLC.

Next, we sought to determine if ADH1B$^+$ and FAP$^+$ CAF were present across different tumor types. To search for ADH1B$^+$ and FAP$^+$ CAF in other cancers, we returned to TCGA to analyze the data available for breast, colon, pancreatic, prostate, and ovarian cancers. Unsupervised hierarchical clustering grouped ADH1B$^+$ and FAP$^+$ CAF genes by their coexpression in each dataset separately (Supplementary Fig. S7). In pancreatic, breast, and colon cancers, we observed a significant separation of the two gene groups, suggesting that other cancer types also harbor lowly and highly activated CAF with similar transcriptional profiles as ADH1B$^+$ and FAP$^+$ CAF found in NSCLC.

ADH1B$^+$ CAF and FAP$^+$ CAF Correlate with Immune Cell Composition and Not with T-cell Localization

Given the data showing that CAF contribute to regulating tumor immunity (11, 12, 48), we investigated the different ligands expressed by CAF populations found in human NSCLC (Fig. 4A). Increased expression of the cytokines *IL34* and *CSF1* suggested macrophage regulation (49) by ADH1B$^+$ CAF, whereas FAP$^+$ CAF might attract eosinophils/basophils via *CCL11* (50), as well as CCR5$^+$ T cells and monocytes through *CCL3* and *CCL5* chemokines (51–53). Notably, the high levels of *CCL21* and *TNFSF13B* (BAFF) in ADH1B$^+$ CAF mainly come from the CCL19-expressing ADH1B$^+$ cells specifically found in TLS (Figs. 2B and 4A), likely contributing to naïve T-cell attraction and B-cell survival in these structures (40). MYH11$^+$αSMA$^+$ CAF expressed increased levels of thymic stromal lymphopoietin (*TSLP*), which can stimulate the maturation of immune cells that express both *IL7R* and *CRLF2* genes forming the heterodimeric TSLP receptor, such as certain dendritic cells (54). MYH11$^+$αSMA$^+$ CAF also showed strong expression of *TGFB1*, which has been implicated in reducing cytotoxic T-cell function (55). IHC confirmed the expression of TGFβ1 by MYH11$^+$αSMA$^+$ CAF at the protein level and its lack of expression on ADH1B$^+$ CAF or FAP$^+$ CAF (Supplementary Fig. S8A and S8B).

To further investigate the contribution of different CAF populations to shaping the immune microenvironment, we used multiplex imaging on FFPE tissue to histologically

Figure 3. ADH1B$^+$ CAF and FAP$^+$ CAF stratify NSCLC into two main stromal patterns associated with tumor stage and histology. **A,** Left, fibroblast subset composition, displayed by percentages, in individual tumor and adjacent tissue samples from the 15 scRNA-seq patients. Right, fibroblast distribution in stage 1 and stage 2$^+$ tumors. The bar graph depicts the percentage of cells from each fibroblast subset among all fibroblasts. CLU$^+$ fib., CLU$^+$ fibroblast. **B,** Top left, ADH1B$^+$ CAF–rich patients showing ADH1B presence throughout the stroma. ADH1B$^+$ CAF–rich patients may present with (top right) or without (top left) a distinct single cell layer of MYH11$^+$αSMA$^+$ CAF at the tumor border. Bottom left, FAP$^+$ CAF–rich patients with FAP staining throughout the stroma. The patients shown demonstrate the variable αSMA presentation in FAP$^+$ cells. All scale bars, 100 μm. Right, cartoon illustrating the observed presentation of multiple CAF subsets in NSCLC. **C,** ADH1B and FAP staining in the IHC cohort. ADH1B staining coverage in the stroma is shown on the *x*-axis. FAP staining coverage in the stroma on regions that did not stain for ADH1B is shown on the *y*-axis. Tumors show a significant preference for either ADH1B or FAP, with less than 5% coverage of the opposing stain. The 5% cutoff was selected after performing hypergeometric tests for 10 thresholds, at 5% increments, between 5% and 50%. The Bonferroni correction adjusted *P* value is 0.008. **D,** Mean expression of selected genes highlighting ADH1B$^+$ CAF intermediate expression of PI16$^+$ fib., alv. fib., and FAP$^+$ CAF–associated genes. **E,** Tumor sample with an extensive invasive margin that displays a spectrum of ADH1B to FAP staining. Zoom, bottom panel, cells appearing to transition from ADH1B to FAP expression. Top panel scale bar, 200 μm; bottom panel scale bar, 100 μm. **F,** Expression of PI16$^+$ fib. and FAP$^+$ CAF module genes in PI16$^+$ fib., ADH1B$^+$ CAF, and FAP$^+$ CAF. Based on gene expression patterns, ADH1B$^+$ CAF appear to occupy an intermediate state of activation between PI16$^+$ fib. and FAP$^+$ CAF. **G,** Relative expression, displayed by *Z*-score, of ADH1B$^+$ CAF– and FAP$^+$ CAF–associated genes in TCGA-LUAD bulk RNA sequencing samples. ADH1B$^+$ CAF and FAP$^+$ CAF genes are significantly anticorrelated, (Pearson) $R = -0.12$ and $P = 0.006$. The sample tumor nuclei count is used as a proxy of tumor purity and shows a relatively even distribution. **H,** TCGA-LUAD mean *Z*-score and standard error of mean (SEM) of ADH1B$^+$ CAF and FAP$^+$ CAF gene signatures stratified by tumor subtype (top left and bottom) or stage (top middle and top right). *Z*-score calculation is listed in Methods, and significance is calculated by independent *t* test (right).

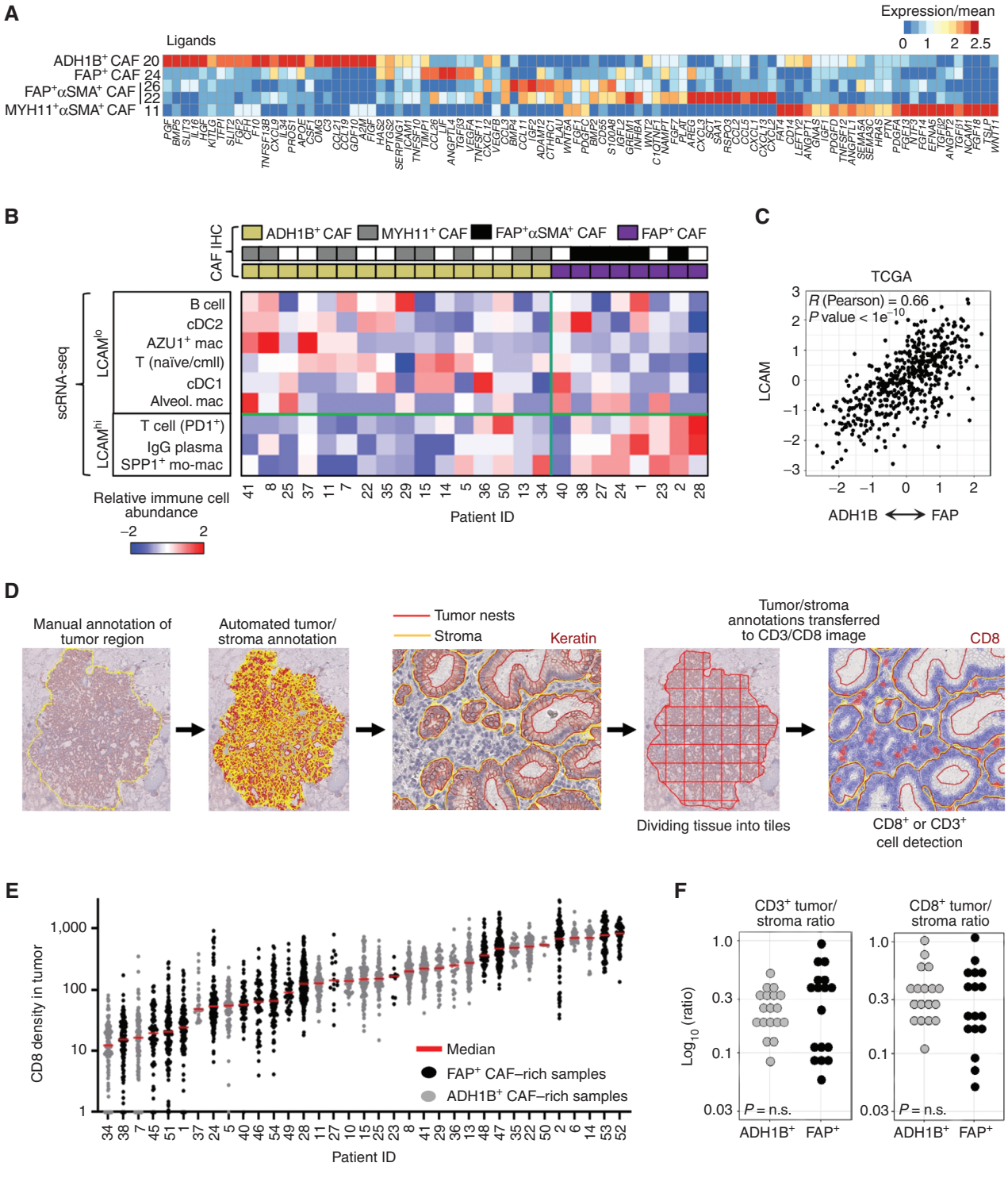


Figure 4. ADH1B+ CAF and FAP+ CAF correlate with immune cell composition and not with T-cell localization. **A,** Gene expression over the mean of highly variable immunomodulatory ligands in CAF clusters. **B,** Immune composition of scRNA-seq tumor samples from ref. 32. CAF phenotype is identified by IHC on the matched FFPE samples and then used to stratify samples. The relative abundance of each cell population within its respective compartment, that is, PD1+ T cells among all T cells, is calculated and then scaled across all tumors for the respective *Z*-score value. The LCAM score is significantly correlated with CAF phenotype, (Pearson) $R = 0.62$ and $P = 0.01$. Alveol. mac, alveolar macrophage; AZU1+ mac, AZU1+ macrophage; SPP1+ mo-mac, SPP1+ monocyte–derived macrophage; T cmll, T central memory II. **C,** Estimating the correlation between CAF phenotype and LCAM in TCGA-LUAD samples. Each patient's mean ADH1B+ CAF gene signature is subtracted from their mean FAP+ CAF gene signature, and the resulting values are correlated with an estimated LCAM score. The corresponding Pearson correlation values are shown. **D,** Schematic of QuPath methodology for tiling and T-cell quantification. **E,** CD8+ cell infiltration into tumor nests in each patient (columns). Each point represents an individual 1,000 μm × 1,000 μm tile (all other tiling is 500 μm × 500 μm). **F,** IHC quantification of the tumor/stroma CD3+ or CD8+ cells per mm² ratio. Tumor samples are stratified by their stroma profile (ADH1B+ CAF–rich or FAP+ CAF–rich), and no significant difference (n.s.; *t* test) was observed.

profile the CAF subset composition of a large cohort of NSCLC samples that we had previously studied using scRNA-seq of purified immune cells (32). We demonstrated a significant association (Pearson, $R = 0.62$, $P = 0.01$) between the presence of FAP$^+$ CAF and the enrichment of inflammatory *SPP1*$^+$ monocyte–derived macrophages, IgG$^+$ plasma cells, and PD1$^+$ T cells (Fig. 4B). These immune cell types were recently described as part of a cellular module termed Lung Cancer Activation Module (LCAM; ref. 32). We then validated this CAF–immune association in the TCGA-LUAD cohort. There was a significant correlation between CAF phenotype and the LCAM score ($R = 0.66$, $P < 1e^{-10}$), supporting that FAP$^+$ CAF–rich samples are linked to more inflammatory and activated immune cells, LCAM^{hi}, in LUAD (Fig. 4C).

Given that the spatial distribution of T cells is a predictor of clinical response to immune-checkpoint blockade (56), we used an unbiased cell quantification method to measure T-cell infiltration in the tumor nests and identified a wide range of infiltration levels across the cohort (Fig. 4D–F). Importantly, there was no observed association between any of the ADH1B$^+$ and FAP$^+$ CAF–rich profiles and CD3$^+$, CD8$^+$, or FOXP3$^+$ T-cell localization (Fig. 4E and F; Supplementary Fig. S8C; Supplementary Table S9). Taken together, ADH1B$^+$ CAF and FAP$^+$ CAF stratify tumor lesions by two levels of fibroblast activation and correlate with the immune phenotype, but not with T-cell spatial distribution.

MYH11$^+$αSMA$^+$ CAF Are Correlated with Decreased T-cell Infiltration in Tumor Nests

The lack of correlation between ADH1B$^+$ CAF or FAP$^+$ CAF with T-cell infiltration contrasts with the general idea that activated fibroblasts orchestrate T-cell exclusion, raising the hypothesis that fibroblast subsets other than ADH1B$^+$ or FAP$^+$ CAF could be involved. To investigate whether MYH11$^+$αSMA$^+$ CAF, which form a single cell layer around tumor nests in a fraction of early-stage tumors, could also affect T-cell tumor infiltration, we subdivided stage 1 patients based on the presence of MYH11$^+$αSMA$^+$ CAF at the tumor border (Fig. 5A). In tumor lesions containing MYH11$^+$αSMA$^+$ CAF, the tumor-to-stroma ratio of infiltrating CD3$^+$ or CD8$^+$ cells was significantly lower (Fig. 5B, right graphs), consistent with a decreased infiltrating CD3$^+$ or CD8$^+$ T-cell density in the tumor (Fig. 5B, left graphs; Supplementary Table S9). FOXP3$^+$ T cells showed a similar trend, although not significant (Supplementary Fig. S9A). In addition, high expression of *TGFB1/TGFB2, WNT5A*, and *WNT11* by MYH11$^+$αSMA$^+$ CAF (Fig. 4A) is in line with previous findings linking TGFβ and WNT/β-catenin pathways with immune cell exclusion in tumors (6, 57–59). These results suggested that MYH11$^+$αSMA$^+$ CAF, with their peritumoral location, may decrease T-cell infiltration into tumor nests.

Within our cohort, MYH11$^+$αSMA$^+$ CAF were found enriched in LUAD samples, especially in the acinar/papillary subtypes, whereas neither the solid subtype of LUAD nor the LUSC samples contained MYH11$^+$αSMA$^+$ CAF lining tumor nests (Fig. 5C). This differential enrichment was also observed within tumor LUAD lesions displaying acinar and solid tumor regions, as annotated by a pathologist (Supplementary Fig. S9B and S9C). Interestingly, the IHC image bank of the Human Protein Atlas showed that a similar peritumoral MYH11 staining pattern as one layer was observed in a fraction of samples of pancreatic and breast cancers (Fig. 5D), suggesting that these CAF may be present in additional cancer types.

Although most tumor lesions were characterized by either high or low MYH11$^+$αSMA$^+$ CAF presence, a fraction of tumors showed local heterogeneity. We assessed the intensity of the CAF barrier at the tumor boundary in 500 μm × 500 μm tiles using the abundance of MYH11$^+$αSMA$^+$ cells in the stroma close to (<10 μm) versus distant from (20 μm–30 μm) tumor cells, which is referred to as the MYH11$^+$αSMA$^+$ CAF score (Fig. 5E and F; Supplementary Fig. S10A–S10C; Supplementary Tables S10 and S11). This automated analysis found that locations where MYH11$^+$αSMA$^+$ CAF were present had significantly lower tumor T-cell density in two independent samples, highlighting that local spatial organization may be driving intertumor differences. Additionally, histologic analysis of the tumor lesion by a pathologist found that regions with high MYH11$^+$αSMA$^+$ CAF scores were predominantly acinar/papillary and lepidic regions at the tumor edge had a lower score (Fig. 5F, bottom). Altogether, these results show that MYH11$^+$αSMA$^+$ CAF are a single layer of elongated cells associated with T-cell marginalization both across NSCLC tumor samples and within tumor lesions.

FAP$^+$αSMA$^+$ CAF Define Regions of Poor T-cell Infiltration within Tumor Lesions and Are Coupled with Dense ECM Deposition

Spatial analysis of FAP$^+$ CAF–rich samples revealed that FAP$^+$αSMA$^+$ CAF, a subset of FAP$^+$ CAF in scRNA-seq, could also explain CD3$^+$ and CD8$^+$ T-cell infiltration within tumors. We measured αSMA coverage and T-cell density in the stroma in 500 μm × 500 μm sections across the tumor lesion (Fig. 6A; Supplementary Table S12) and revealed that regions dense in αSMA are poorly infiltrated by T cells ($R = -0.48$, $P = 1e^{-10}$; Fig. 6B). This anticorrelation was replicated across different tumors (Supplementary Fig. S11A–S11C; Supplementary Tables S12 and S13) and suggested that FAP$^+$αSMA$^+$ CAF directly restrict T-cell motility. Notably, a high fraction of FAP$^+$ CAF–rich samples presented with several layers of FAP$^+$αSMA$^+$ CAF lining tumor nests that delineated regions devoid of T cells (Fig. 6C). In addition, intertumor αSMA heterogeneity showed a trend toward an anticorrelation with T-cell infiltration in tumor nests (Supplementary Fig. S11D and S11E).

Based on prior studies showing that the ECM plays a role in T-cell exclusion and immunosuppression (7, 60, 61), we postulated that FAP$^+$αSMA$^+$ CAF may also express a specific ECM profile involved in regulating T-cell localization. Masson's trichrome staining revealed a high density of fiber deposition at the tumor border in both MYH11$^+$αSMA$^+$ and FAP$^+$αSMA$^+$ CAF–containing tumor lesions, suggesting that these CAF are depositing a fibrillar barrier limiting T-cell access to tumor cells (Fig. 6D). Analysis of the scRNA-seq data showed that MYH11$^+$αSMA$^+$ CAF expressed *COL9A1*, *COL27A1*, and a distinct type of sheet-forming, basement membrane collagens, *COL4A1* and *COL4A2*, which are found in lining vessels and various epithelial layers (refs. 62, 63; Fig. 6E). A thick layer of collagen IV fibers lining tumor nests was frequently found colocalized with MYH11$^+$ CAF in tumor

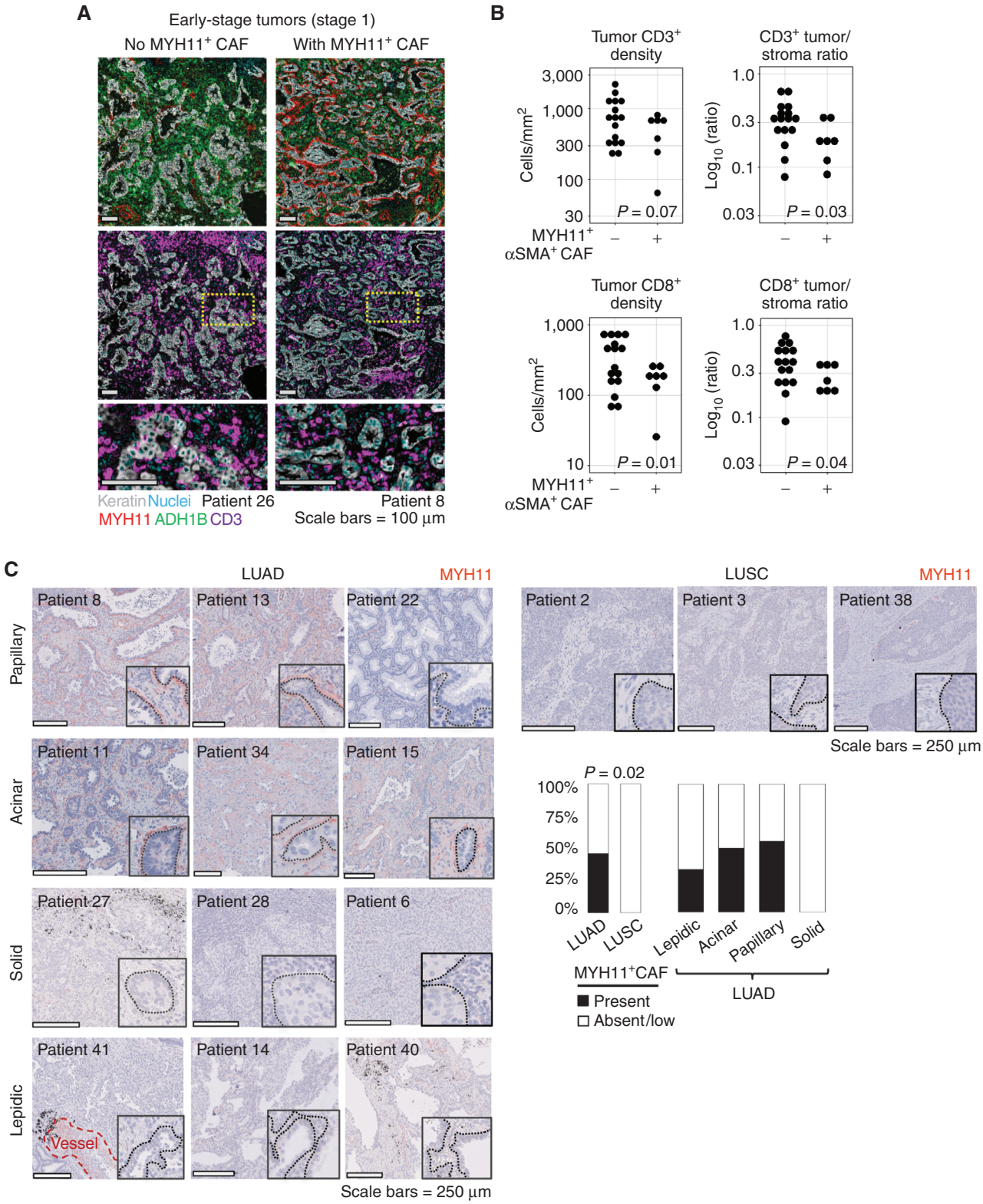


Figure 5. MYH11+αSMA+ CAF are correlated with decreased T-cell infiltration in tumor nests. **A,** Representative examples of IHC stains from NSCLC tumors with and without MYH11+αSMA+ CAF present showing CD3+ cell exclusion from tumor nests when MYH11+αSMA+ CAF are present. **B,** The presence or absence of MYH11+αSMA+ CAF demonstrates significant differences in tumor-infiltrating CD3+ or CD8+ cells per mm^2 (left) and the ratio of CD3+ or CD8+ cells per mm^2 in the tumor versus stroma (right). Only early-stage (tumor stage 1) patients were included to eliminate bias due to MYH11+αSMA+ CAF rarely being found at a later stage. **C,** Representative images of MYH11 staining in multiple pathologies and histologic subtypes. All scale bars, 250 μm. Bar plot, MYH11+αSMA+ CAF distribution in different pathologies and histologic subtypes in NSCLC. Significance is determined by the *t* test. (*continued on following page*)

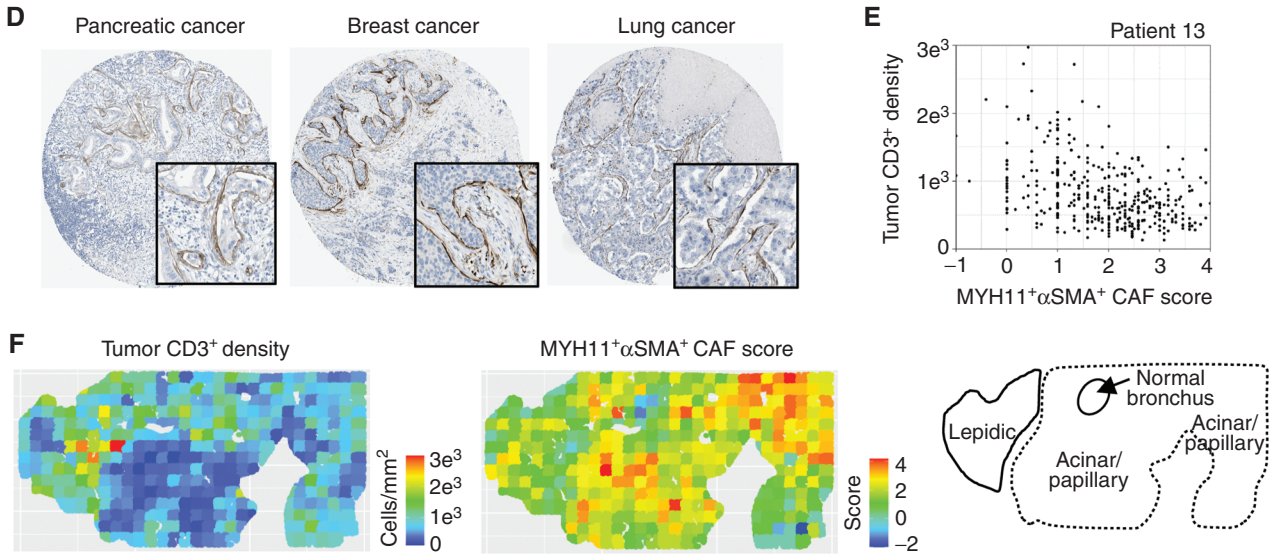


Figure 5. ***(Continued)*** **D,** MYH11 staining from the Human Protein Atlas. A line of MYH11$^+$ cells was observed lining tumor nests in tumor samples from the Human Protein Atlas, including pancreatic cancer (patient ID: 956), breast cancer (patient ID: 2805), and lung cancer (patient ID: 1847). **E,** Quantification of 500 × 500 μm tiles of both MYH11$^+$αSMA$^+$ CAF score, estimating tumor proximity of MYH11$^+$αSMA$^+$ cells by quantifying their enrichment within 10 μm from tumor cells versus regions 20 μm to 30 μm from tumor cells, and tumor-infiltrating CD3$^+$ cells per mm^2. A high MYH11$^+$αSMA$^+$ CAF score is significantly anticorrelated (Pearson) with the number of tumor-infiltrating CD3$^+$ cells relative to the stroma. **F,** Visualization of the tiling described in **E**. Bottom, histologic scoring of a tumor lesion highlighting that a high MYH11$^+$αSMA$^+$ CAF score is associated more with the acinar/papillary phenotype rather than lepidic.

lesions (Supplementary Fig. S12A), whereas samples or tumor regions lacking MYH11$^+$ CAF showed no to low collagen IV deposition (Supplementary Fig. S12B and S12C). There were rare exceptions to this observation (Supplementary Fig. S12D), indicating that additional ECM factors may contribute to the fiber density observed in MYH11$^+$αSMA$^+$ CAF–rich samples by Masson's trichrome (Fig. 6D).

The ECM program of FAP$^+$αSMA$^+$ CAF was distinct from MYH11$^+$αSMA$^+$ CAF. FAP$^+$αSMA$^+$ CAF expressed high levels of the fibrillar collagen *COL11A1* (64) and *COL12A1* (Fig. 6E) compared with other CAF, including FAP$^+$ CAF. We performed collagen fiber staining on FAP$^+$ and FAP$^+$αSMA$^+$ CAF–rich samples and quantified the coverage of collagen XI, XII, and IV, as well as αSMA in 500 μm × 500 μm tiles. This analysis revealed a strong correlation between the stroma coverage of αSMA and collagen XI/XII (Fig. 6F; Supplementary Fig. S13A–S13D; Supplementary Tables S14–S16). Importantly, no positive correlation was observed between collagen XI/XII and FAP (αSMA^{neg}) stroma coverage or between collagen IV and αSMA (Fig. 6F; Supplementary Fig. S13A–S13D). We observed a strong alignment of collagen fibers in areas rich in FAP$^+$αSMA$^+$ CAF (Supplementary Fig. S13A, S13D, and S13E), indicating that these CAF shape not only the local matrix composition but also its structural organization.

In our prior study using live imaging of T cells in viable human NSCLC tumor slices (7), we showed that a dense matrix around tumor nests was anticorrelated with T-cell motility. Collagenase treatment of the tumor tissue from three patients in that study led to increased T-cell contact with tumor cells, demonstrating the functional role of dense matrix fibers in restricting T cell/tumor cell interactions (Fig. 6G, top). We hypothesized that the resolution of our current analysis could enable the identification of the CAF subsets that generated this barrier. We, therefore, retrieved FFPE slides from the same three tumors and stained for CAF markers identified in the current study. Our staining shows multilayer FAP$^+$αSMA$^+$ CAF localized around tumor nests, revealing the cellular source of the causal factor behind T-cell exclusion in those samples (Fig. 6G, bottom). Thus, the specific spatial distribution of FAP$^+$αSMA$^+$ CAF and their unique ECM profiles may drive T-cell exclusion in NSCLC and represent potential therapeutic targets. Combined with MYH11$^+$αSMA$^+$ CAF, this suggests a refined model for CAF phenotypes in NSCLC and a potential mechanism for T-cell exclusion (Fig. 7A and B).

DISCUSSION

A majority of patients fail to achieve clinical benefit using standard immune-checkpoint blockade, and as such, novel combination approaches are required to improve the response (65). Patients with T cell–excluded tumors have a poor response to immune-checkpoint blockade compared with those with T cell–infiltrated tumors (3, 5, 6), which raises the possibility that targeting the mechanism of T-cell exclusion would improve clinical responses. To this end, our study provides a comprehensive map of the fibroblast compartment in human lung tumors at the single-cell level and with spatial resolution. We define the molecular and functional diversity of the fibroblast compartment of lung tumors and determine how distinct CAF subsets may influence immune cell composition as well as T-cell spatial organization.

Our analysis shows that the stroma in NSCLC lesions is dominated by either lowly activated ADH1B$^+$ CAF, with or

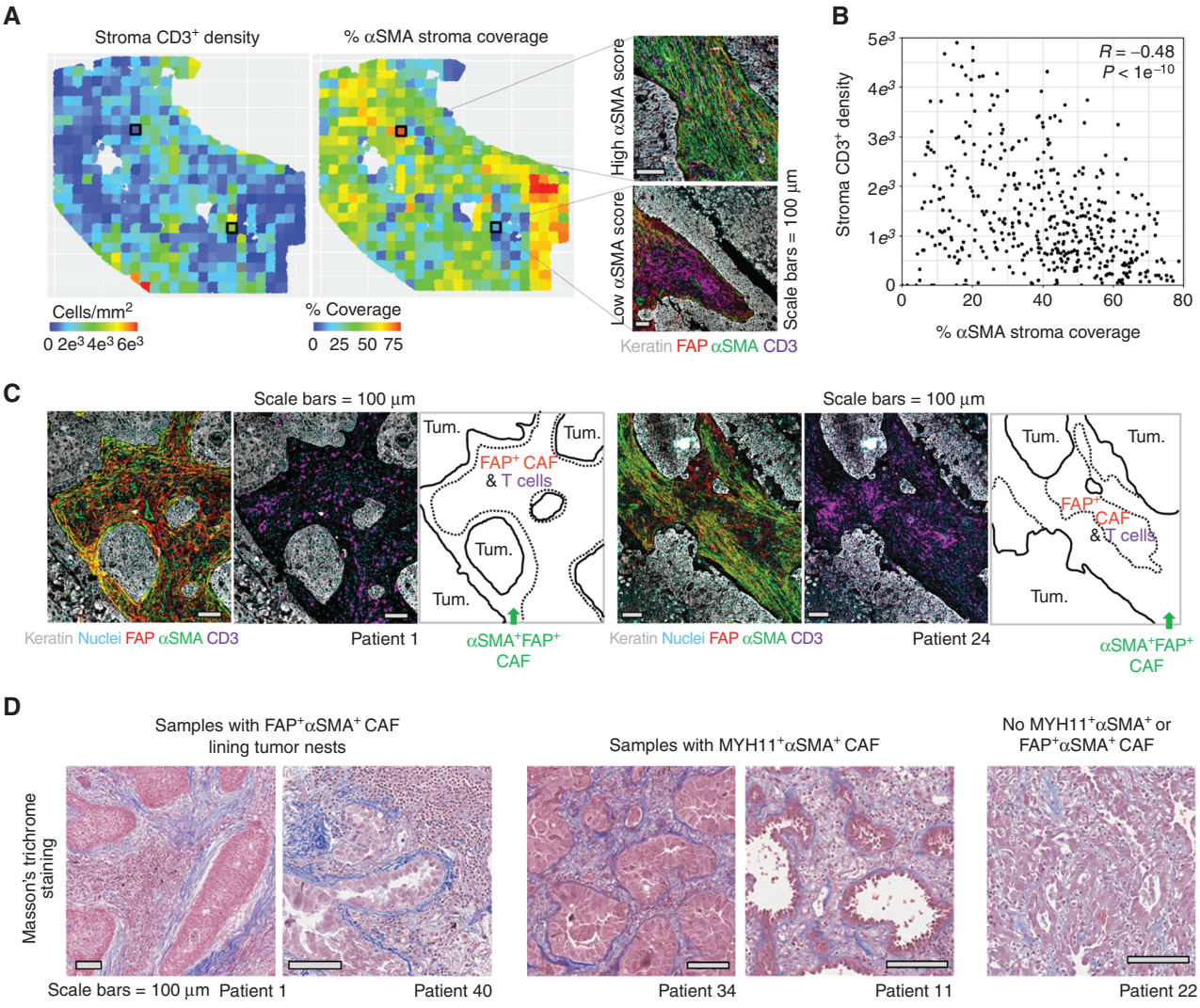


Figure 6. FAP+αSMA+ CAF define patterns of poor T-cell infiltration within tumor lesions. **A,** Left, intratumoral heterogeneity of αSMA coverage (middle) and CD3+ cell density in the stroma in 500 × 500 µm tiles. Right, representative examples of tiles showing regions with high or low levels of αSMA. **B,** Quantification of αSMA coverage and CD3+ density in each tile (points) as defined in **A**, showing a significant anticorrelation (Pearson) of αSMA coverage and CD3+ cell density. **C,** Dense αSMA staining at the tumor border associated with decreased CD3+ cell abundance. The green arrows highlight border regions with high αSMA and low CD3+ cells. Tum., tumor. **D,** Masson's trichrome staining highlights increased ECM at the tumor boundary in samples containing MYH11+αSMA+ or FAP+αSMA+ CAF. (*continued on following page*)

without MYH11+αSMA+ CAF, or highly activated FAP+ CAF, with variable αSMA levels (Fig. 7). ADH1B+ CAF have higher activation levels than fibroblasts found in normal lung tissue, as seen by their expression profile and enrichment in the tumor lesion, as well as by the clear spatial distinction between CD10+ alv. fib. in adjacent tissue and ADH1B+CD10neg CAF in the tumor, as seen in multiplex IHC. FAP+ CAF, by contrast, show dramatic transcriptional differences from adjacent tissue fibroblasts, including high expression of many previously established CAF markers such as *FAP*, *POSTN*, and *COL1A1* (23, 28). FAP+ CAF and FAP+αSMA+ CAF represent higher activation states compared with ADH1B+ CAF and occasionally form spatial gradients of ADH1B+-to-FAP+ CAF, suggesting ADH1B+ CAF may contribute to the FAP+ CAF pool. FAP+αSMA+ CAF are a subpopulation of FAP+ CAF with higher expression of contractility and ECM genes. Our spatial data frequently show increased αSMA staining at the tumor nest boundary, suggesting that the αSMA program is upregulated in FAP+ CAF upon physical or molecular signals from tumor cells. In conjunction with our observation that ADH1B+ CAF transition to FAP+ CAF, this suggests that the tumor cells and TME play a critical role in CAF differentiation. ADH1B+ CAF express transcriptional programs of both alv. fib. and PI16+ fib., suggesting that these two lung tissue cell types could give rise to ADH1B+ CAF. *In vivo* fate-mapping experiments will be needed to further investigate this possibility.

We have also shown that ADH1B+ CAF and FAP+ CAF are associated with pathologic and histologic subtypes as well as tumor stage, with ADH1B+ CAF being associated with the adenocarcinoma papillary subtype and stage 1 and FAP+ CAF being enriched in the adenocarcinoma solid subtype and

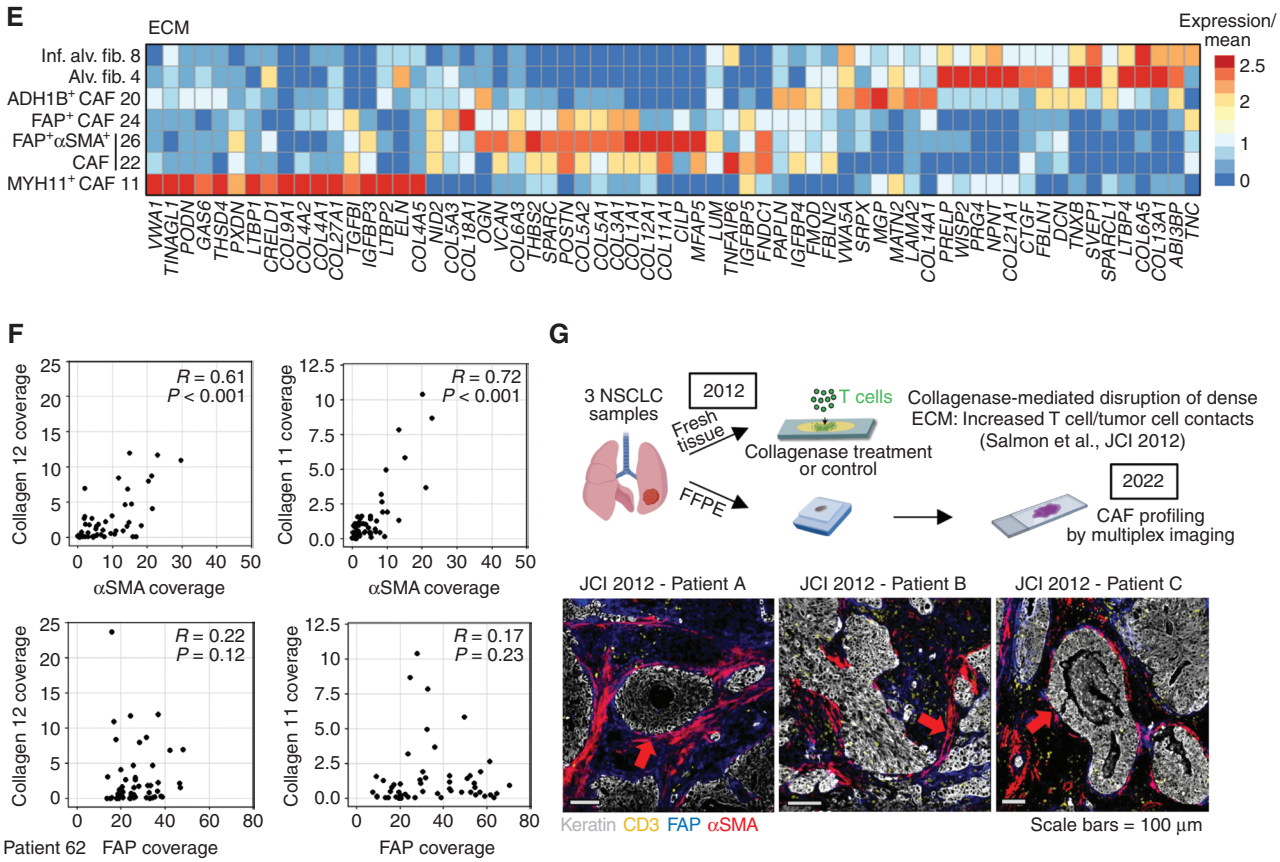


Figure 6. ***(Continued)*** **E,** Averaged gene expression of highly variable ECM genes in CAF clusters. **F,** αSMA coverage of the stroma is significantly correlated (Spearman) with collagen XI and XII deposition, whereas FAP+ CAF show no correlation. **G,** Top, our prior work (7) showed that collagenase treatment of viable slices of NSCLC tumor tissue increased T-cell access to tumor cells. FFPE sections from tumor samples of the same three patients were stained by multiplex IHC for markers of CAF identified in the present study. Bottom, FAP is found throughout the stroma, and αSMA shows increased expression at the tumor border (red arrows).

squamous cell carcinomas and later stage. This association between CAF populations and histologic subtypes, which correlates with prognosis (47), may shed light on the molecular programs behind the different NSCLC subtypes and inform clinical trial inclusion criteria when therapeutically targeting CAF subsets. Furthermore, our staining protocol for these CAF subsets may help refine the categorization of histologic subtypes. Beyond subtype and stage, we found a significant association between FAP+ CAF and the LCAM inflammatory/activated immune phenotype (including *SPP1*+ monocyte–derived macrophages, IgG plasma cells, and PD1+ T cells), which we previously described (32). This observation, in conjunction with the distinct immunomodulatory profiles of ADH1B+ CAF and FAP+ CAF, suggests that CAF participate in shaping the immune response to the tumor.

Prior studies have suggested that activated CAF may play a role in T-cell exclusion (6, 15, 16). Notably, FAP+ CAF do not correlate with the T-cell distribution pattern in NSCLC, an important factor to keep in mind when developing therapeutics and that may explain why strategies targeting FAP+ CAF have failed in human clinical trials so far (18, 66). In contrast, we have found two distinct CAF populations with specific molecular programs and spatial organizations that contribute to T-cell exclusion. First, FAP+αSMA+ CAF are significantly correlated with regions of T-cell exclusion in the tumor stroma and can form multiple layers at the tumor boundary and restrict T-cell contact with tumor cells. On the other hand, MYH11+αSMA+ CAF form a single cell layer lining tumor nests in a fraction of adenocarcinomas and are significantly correlated with immune cell exclusion from tumor regions, both within cancer lesions and across tumor samples. The enrichment of MYH11+αSMA+ CAF in early-stage tumors may suggest that they respond to early tumor cell signals that may be lost upon tumor progression.

FAP+αSMA+ CAF and MYH11+αSMA+ CAF correspond to two clearly distinct fibroblast subsets, as observed through the scRNA-seq data and in line with their presence in distinct tumors. Notably, they display similarities, including high expression levels of ECM and contractility genes, which implies that they influence T-cell spatial distribution through similar mechanisms, with the production of fibers limiting T-cell access to cancer cells (Salmon et al.; ref. 7). The differences in their matrix deposition (including type IV collagen for MYH11+αSMA+ CAF and type XI and XII collagens for FAP+αSMA+ CAF) also indicate that they drive T-cell marginalization by forming different types of barriers to lymphocytes. Our analysis of NSCLC tumor

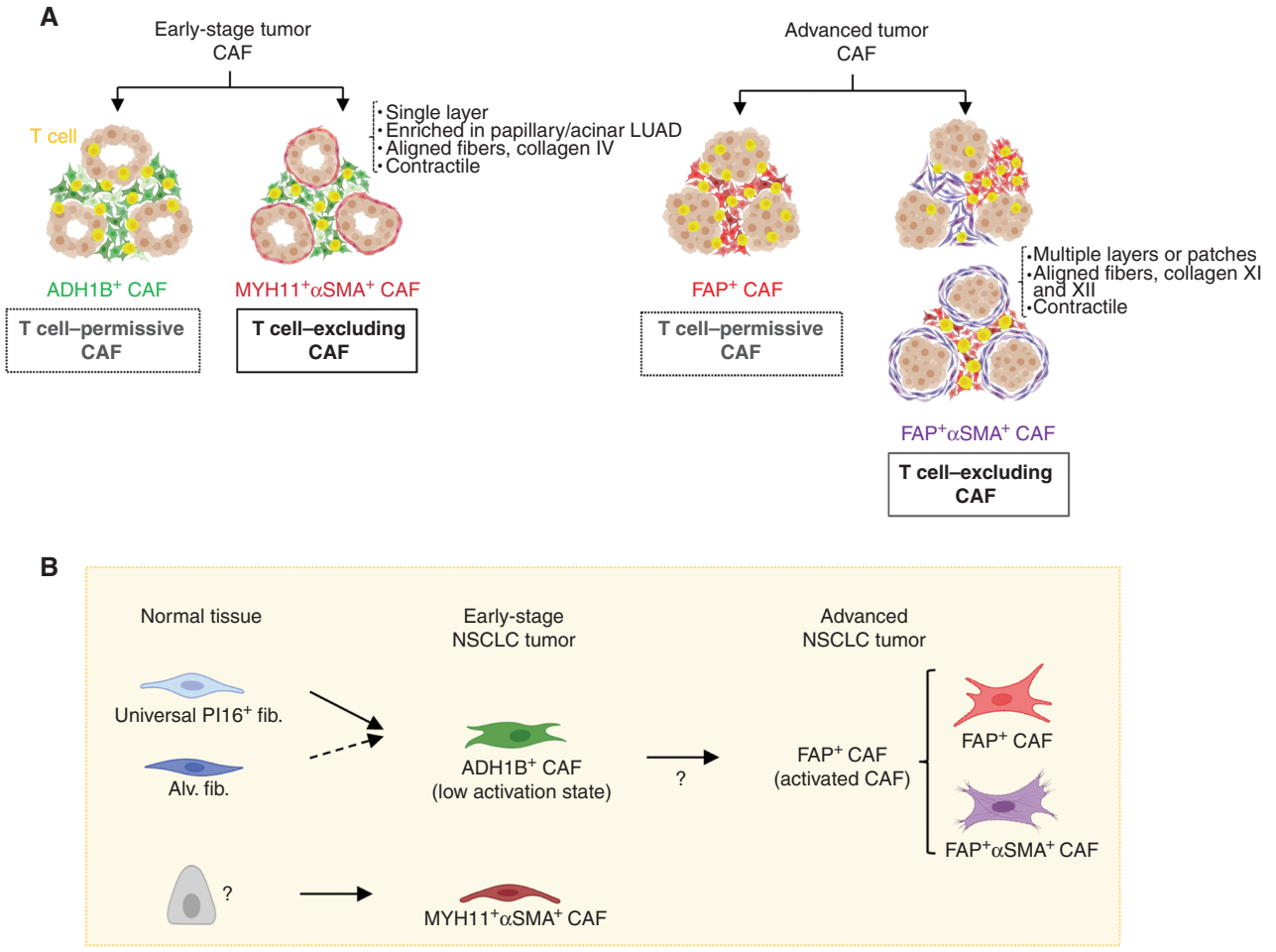


Figure 7. Working model. **A,** Graphical illustration of all stroma presentations found in this study. NSCLC samples enriched in ADH1B+ CAF throughout the stroma can be found with or without a single cell layer of MYH11+αSMA+ CAF lining tumor cell aggregates. Those with MYH11+αSMA+ CAF show increased T-cell exclusion from the tumor nests. NSCLC samples enriched in FAP+ CAF are found with a variable abundance of FAP+αSMA+ CAF. Stromal regions with high αSMA have reduced T-cell accumulation, and tumor nests surrounded by several layers of FAP+αSMA+ CAF show a lower T-cell infiltration. **B,** Cartoon depicting the general distribution of fibroblast and CAF populations in adjacent lung tissue, early-stage NSCLC, and advanced NSCLC, as well as the potential differentiation trajectories.

samples from Salmon et al. (7) demonstrates the causal role of FAP+αSMA+ CAF in excluding T cells, identifying them in areas of dense fiber deposition that were implicated in restricting T-cell interactions with tumor cells (7). Although ECM degradation can improve T-cell infiltration (61), targeting ECM molecules is challenging in patients given the low specificity and the risk of on-target, off-tumor toxicity. Here, our study paves the way to develop novel strategies to differently target the two distinct cellular sources of these ECM molecules.

In summary, our study has identified several CAF populations that show greater heterogeneity than the previously established CAF classification and provides novel therapeutic targets to pursue in order to augment response to cancer immunotherapies. We demonstrate that pairing molecular and spatial analysis is crucial to understanding the true organization of the human TME and to developing novel CAF targeting strategies for efficient antitumor combinations.

METHODS

Human Subjects

In collaboration with the Biorepository and Department of Pathology, tumor and adjacent noninvolved lung samples were obtained from surgical specimens of patients undergoing resection at the Mount Sinai Medical Center (New York, NY). Written informed consent was obtained in accordance with U.S. Common Rule and the following protocols were reviewed and approved by the Institutional Review Board at the Icahn School of Medicine at Mount Sinai: IRB Human Subjects Electronic Research Applications 10-00472 and 10-00135. Additional FFPE NSCLC samples were obtained from Institut Mutualiste Montsouris, Paris, in collaboration with Institut Curie, Paris. Collection of clinical NSCLC specimens at Institut Mutualiste Montsouris was conducted under the umbrella protocol of the pathologic department and biospecimen core facility, established under the reference EUdract 2017-A03081-52 and approved by the Ethics Committee CPP SUD-EST I. FFPE blocks from NSCLC tumors used in ref. 7 were obtained from Assistance Publique–Hôpitaux de Paris (Hôtel Dieu Hospital), with the approval of the IRB CPP Ile de France II, 2008-133 and 2012 06-12, No. 2018 MS1.

Tissue Processing

The noninvolved lung and tumor tissue were weighed and cut into sections of 0.1 to 0.2 g and then placed into 5-mL microtubes (Argos Technologies). Sections were minced with scissors and enzymatically digested in CO_2-independent media (Fisher Scientific, 18045088) with 0.25 mg/mL Collagenase IV (Sigma-Aldrich, C5138-1G), 200 U/mL Collagenase D (Sigma-Aldrich, 11088882001), and 0.1 mg/mL DNase (Sigma-Aldrich, DN25-1G) for 40 minutes at 37°C under 80 rpm agitation. Cell suspensions were passed through a syringe with an 18-gauge needle 8 to 10 times, filtered through a 70-μm cell strainer, and then lysed in red blood cell buffer (Fisher Scientific, NC9067514). The cells were resuspended in a buffer comprised of DPBS (Corning, D8537-6 × 500 mL) with 5% BSA (Equitech-Bio, BAH62-0500) and 1 mmol/L EDTA (Sigma-Aldrich, 46-034-CI) and then counted using a hemocytometer and Trypan blue (Fisher Scientific, MT25900CI).

Flow Cytometry Sorting

Cells were stained for EpCAM (BioLegend, clone 9C4), CD45 (BioLegend, clone HI30), CD29 (BioLegend, clone TS2/16), PDPN (BioLegend, clone NC-08), and Live/Dead blue fluorescent dye (Thermo Scientific, L34963) for 30 minutes at 4°C. Among live cells, EpCAM and CD45 were used to remove epithelial and immune cells, respectively, whereas CD29, present on all stromal cells, was used to enrich for cells with intact surface markers (see Supplementary Fig. S1B). The 1.5-mL collection tubes (Fisher Scientific, 05-408-129) were coated with 10% BSA to improve cell survival after sorting.

scRNA-seq

For each sample, up to an estimated 5,000 cells were loaded directly from the flow cytometry sort onto 10X Chromium chemistry kits. Kit versions for each sample are indicated in Supplementary Table S1. Processing downstream of cell loading was performed by the Human Immune Monitoring Core at the Icahn School of Medicine at Mount Sinai. Libraries were prepared according to the manufacturer's instructions, and quality control of the cDNA and final libraries was performed using the CyberGreen qPCR library quantification assay. Sequencing was performed on Illumina sequencers to a depth of at least 80 million reads per library.

Sequencing Data Analysis and Unsupervised Batch-Aware Clustering

Transcriptomic library reads were aligned to the GRCh38 reference genome and quantified using Cell Ranger (v3.1.0).

Stromal cells isolated from tumor and adjacent lung samples were analyzed using an unsupervised batch-aware clustering method we recently described (33). First, stromal cells were filtered for cell barcodes recording >800 UMI, with <25% mitochondrial gene expression, and with less than defined thresholds of expression for genes associated with red blood cells, epithelial cells, macrophages, T cells, and plasma cells (Supplementary Table S17). Thirteen tumor and 11 adjacent samples were clustered jointly. This expectation-maximization–like algorithm iteratively updates both cluster assignments and sample-wise noise estimates until it converges, using a multinomial mixture model capturing the transcriptional profiles of the different cell states and sample-specific fractions of background noise. We ran the algorithm described in Martin et al. (33) with minor modifications: Training and test set sizes per sample were 7,500 and 2,500, respectively. The best clustering initiation was selected from 1,000 instead of 10,000 k-means+ runs. For this clustering, we included barcodes with more than 800 UMIs and used $K_{reg\_ds} = 0.2$; $(P_1, P_2) =$ (0th, 50th) percentiles; $K_{reg} = 5 \times 10^{-6}$; $k = 28$. Genes with high variability between patients were not used in the clustering. Those genes consisted of mitochondrial, stress, metallothionein genes, immunoglobulin variable chain genes, HLA class I and II genes, and three specific genes with variable/noisy expression: *MALAT1, JCHAIN*, and *XIST* (Supplementary Table S17). Ribosomal genes were excluded only from the k-means clustering (step 2.D as described in Martin et al.; ref. 33; Supplementary Table S17).

Cell Annotation

Using the gene module analysis described earlier, we identified highly variable genes and explored their expression across different clusters. Clusters were annotated by comparing gene expression patterns with profiles reported in prior literature.

For stromal cell clusters, EC expressed multiple identifying markers such as *PECAM1*, *VWF*, *CLDN5*, and *EMCN*, and lymphatics could be identified with *TFF3*, *LYVE1*, and *PROX1*. PvC were identified by a combination of subset-specific markers for and shared expression of contractile genes such as *ACTA2*, *TAGLN*, *MYL9*, and *TPM2*. PvC subset-specific genes included *RGS5*, *COX4I2*, and *HIGD1B* for pericytes and *DES* and *ACTG2* for SM. Identifying fibroblast markers included those listed in the main text, *PDGFRA*, *SPON1*, and *MMP2*, but also *DCN*, *FBLN1*, *LUM*, *COL1A2*, *RARRES2*, and *CTGF*.

A cluster (#13) with contaminating epithelial cells was identified by the high expression of multiple keratin genes including *KRT17* and *KRT19*. Contaminating macrophages were identified in cluster 6 by expression of CD45 (*PTPRC*), *C1QB*, *C1QA*, *C1QC*, and *MARCO*. Clusters 18 and 16 were excluded due to high mitochondrial gene content and hemoglobin genes, respectively. The annotation process for fibroblast subsets is described in the text.

Histologic Staining

Multiplexed IHC was performed according to the protocol developed by Remark et al. (31) with some modifications. Slides were baked at 37°C overnight, deparaffinized in xylene, and then rehydrated. Antigen retrieval was done in citrate buffer (pH 6 or 9; Dako, S2367 or 2369) at 95°C for 30 minutes, followed by incubation in 3% hydrogen peroxide for 15 minutes, and then nonspecific primary antibody binding was blocked using serum-free protein block solution (Dako, X0909) before adding primary antibody for 1 hour at room temperature or overnight at 4°C. The primary antibody was detected using a secondary antibody conjugated to horseradish peroxidase followed by chromogenic revelation using 3-amino-9-ethylcarbazole (AEC; Vector laboratories, SK4200). Slides were counterstained with hematoxylin (Sigma-Aldrich, HHS32-1L) and mounted with a glycerol-based mounting medium (Dako, C0563). Then the same slides were bleached and restained as previously described. Antibody sources can be found in Supplementary Table S5. Masson's trichrome staining was performed by the Biorepository and Pathology Core at the Icahn School of Medicine at Mount Sinai.

Mass CyTOF

Samples were processed to a single-cell suspension according to the tissue processing protocol listed earlier. Cell viability staining was achieved with Rh103 staining for 20 minutes at 37°C, followed by staining with the CyTOF antibodies listed in Supplementary Table S5. Acquisition of the samples was performed by the Human Immune Monitoring Center at Mount Sinai. All analysis of the CyTOF samples, including the creation of the visualization of t-distributed stochastic neighbor embedding (viSNE) plots, was done using the Cytobank platform (https://www.cytobank.org/).

ECM and Immunomodulatory Gene Expression Profiles

Gene lists were sourced from refs. 67 and 68 for ECM and immunomodulatory genes, respectively. Selected genes had to meet a mean expression threshold of one UMI per 2,000 UMIs in 2% of cells in at least one cluster and meet a minimum 3-fold expression change between at least two clusters. Selected genes for display were further refined by qualitative analysis.

Gene Module Analysis

Gene correlation modules were generated using a similar method as previously described in (33). Briefly, cells are downsampled to 2,000 total UMIs and highly variable genes are isolated. A gene–gene correlation matrix for the isolated gene set is computed for each sample over the cell population(s) of interest, and correlation matrices are averaged following a Fisher Z-transformation. Applying the inverse transformation then results in the best-estimate correlation coefficients of gene–gene interactions across the dataset. Genes are clustered into modules using complete linkage hierarchical clustering over correlation distance. Ribosomal, mitochondrial, HLA, and immunoglobulin genes were removed from the analysis prior to the creation of gene modules, as these genes were not of interest in this study, reflected patient genomic variability, or were heavily influenced by contaminating plasma cells.

Acquisition of TCGA Dataset and Histologic Subtypes

The TCGA-LUAD RNA-seq data were downloaded using the GDCquery and GDCdownload functions from the TCGAbiolinks R package. GDCquery options included: project = "TCGA-LUAD," data.category = "Transcriptome Profiling," data.type = "Gene Expression Quantification," workflow.type = "HTSeq–FPKM," experimental. strategy = "RNA-Seq, and legacy = F. Whole-exome sequencing data were downloaded using the GDCquery_Maf function with the arguments tumor = "TCGA-LUAD" and pipelines = "mutect2." Clinical data were downloaded using the GDCquery_clinic function with the arguments project = "TCGA-LUAD" and type = "clinical."

The dominant histologic subtype for each TCGA tumor was sourced from ref. 46.

CAF Gene Signatures and LCAM in Bulk Analysis

Our initial gene list was acquired from gene module analysis, as described above. To define cell type–specific gene signatures, we first excluded genes well expressed in nonfibroblast lineages, such as EC and PvC within our dataset. Next, we utilized the datasets from refs. 32 and 23 to exclude genes found in epithelial and immune cells. For each dataset, we performed in-house clustering and identified stromal clusters, and then excluded any genes from our signatures if they showed higher expression in the nonstromal clusters. We then compared the expression of the genes between fibroblast subsets and kept genes only if they showed high expression in the cluster of interest relative to other clusters. Due to their similarity, FAP$^+$ CAF and FAP$^+$αSMA$^+$ CAF were treated as one group, and due to ADH1B$^+$ CAF similarity to adjacent fibroblast clusters, they were not contrasted with alv. fib. and PI16$^+$ fib. Signatures were further refined by manually checking that expression was consistently enriched in the cell type of interest across at least three patients.

Bulk RNA samples were scored in the following method: Mitochondrial genes, hemoglobins, and Ig variants genes were removed from the data tables. Next, genes were converted to a percent expression of the total reads and regularized with a constant value, 1e–8, added to each gene and log-transformed. *Z*-scores were then calculated for each gene across all samples. Finally, the overall cell type score was calculated by taking the average of all genes within the signature. A complete list containing the signature and removed genes can be found in Supplementary Table S17.

The derivation of LCAM^{hi} or LCAM^{lo} scores is described in (32). In short, the cell types associated with each state were averaged to create an LCAM^{hi} or LCAM^{lo} score. The difference between LCAM^{hi} and LCAM^{lo} was the final LCAM score. All signatures are calculated using only tumor samples, with sample ID ending in "-01A," and signatures were *Z*-scored before graphing or other analysis. TCGA patients with their corresponding signature scores can be found in Supplementary Table S7.

External scRNA-seq Dataset Analysis

Based on our scRNA-seq data we, defined separate gene signatures for pan-fibroblast cells, PvC, alv. fib., meso. fib., MYH11$^+$αSMA$^+$ CAF, ADH1B$^+$ CAF, FAP$^+$ CAF, and FAP$^+$αSMA$^+$ CAF (Supplementary Table S17). Fibroblasts were isolated by selecting those above our pan-fibroblast score threshold and below the PvC score threshold (Supplementary Fig. S6A). Then, alv. fib., meso. fib., and MYH11$^+$αSMA$^+$ CAF were identified and removed from subsequent plots (Supplementary Fig. S6A). The remaining cells were then compared against the ADH1B$^+$, FAP$^+$, and FAP$^+$αSMA$^+$ CAF gene signatures (Supplementary Fig. S6B–S6D). The datasets can be found at the following locations: Lambrechts et al. (23)—ArrayExpress under accessions E-MTAB-6149 and E-MTAB-6653; Wu et al. (26)—Gene Expression Omnibus (GEO) database accession codes GSE13190712 and GSE99254; Kim et al. (24)—GEO database accession code GSE131907; and Laughney et al. (25)—GEO database accession code GSE123904.

Histology Analysis and Overlays

All histology analyses were performed using the open-source image analysis QuPath software (QuPath-0.2.3, https://qupath.github.io/; ref. 69) and ImageJ/Fiji (70, 71).

To create overlayed images, scans were exported from QuPath as OME.TIFF and then imported into ImageJ using the BioFormats plugin (72). Alignment was done using the "Linear Stack Alignment with SIFT" plug-in (73). The AEC and hematoxylin stains were extracted from individual scans using "colour deconvolution" and colored as desired.

Quantifying T-cell Infiltration and Interpatient CAF Heterogeneity. Cropped scans were imported into a newly created project in QuPath and were aligned using the "interactive image alignment" plug-in. Alignment information was saved using QuPath_script_1. Tumor and stroma regions were defined by applying the "create cytokine annotation" function on the keratin scans. The stroma and tumor annotations were transferred onto the aligned CD3, CD8, αSMA, ADH1B, and FAP scans with QuPath_script_2. On the CD3 and CD8 scans, positive cell detection was used to count the CD3$^+$ and CD8$^+$ cells. We manually removed debris spots (which appear positive for any marker) to avoid false positives. Distance to the tumor and stroma annotations was calculated using the "distance to annotation 2D" option, and the measurements were exported as raw data to be analyzed in R. For Fig. 4E, the stroma and tumor annotations were tiled using the function "Create Tiles." For ADH1B and FAP scans, the positively stained area was calculated using the QuPath training classifier.

Quantifying T-cell Infiltration and Intrapatient CAF Heterogeneity. For the intrapatient analysis, we created a separate QuPath project with all the desired scans to analyze. The images were cropped and exported, and then overlays were generated using an ImageJ script with the following steps:

a. Deconvolution (hematoxylin, AEC, residual)
b. Alignment on hematoxylin images
c. Creation of transformation matrix and then application on AEC images
d. Threshold to remove background and then "Stack of Images"

The composite image was transferred back to QuPath for further analysis. Adjacent tissue regions on slides were excluded from the analysis, and in the region to analyze we used the "train pixel classifier" to annotate tumor nests versus stroma, whereas αSMA, FAP, and ADH1B areas were annotated using "QuPath train classifier." A dedicated script automated the tiling and quantification and resulted in three data files: cell information including staining intensity and distances to annotations, tile annotation parameters, and annotation measurements such as ADH1B-stained area within an annotation. The resulting measurements were exported and analyzed in R.

Quantifying MYH11$^+$αSMA$^+$ CAF Boundary Enrichment. The MYH11$^+$αSMA$^+$ CAF score (Fig. 5C and D) approximates the enrichment of these cells at the tumor nest boundary (<10 μm) relative to their distal background density (20–30 μm). Iterating over tiles, we counted the number of MYH11$^+$αSMA$^+$ double-positive cells in each distance bin of 1 μm. The proximal value was defined as the quantile, 0.75, of the number of cells in the bins within a distance of <10 μm. The distal value was similarly defined as the quantile, 0.75, of the number of cells in the bins with distances between 20 and 30 μm. The MYH11$^+$αSMA$^+$ CAF score was defined as $\log_2$(proximal/distal). The 0.75 percentile was selected to maximize the sensitivity of detecting robust high-density regions.

Quantifying FAP$^+$αSMA$^+$ CAF Correlation with Collagens IV, XI, and XII. After the images were cropped and exported, overlays were generated using an ImageJ script as described above. The composite image was transferred to QuPath software for further analysis. The train pixel classifier was used to annotate tumor nests versus stroma. In the stromal annotation, the train pixel classifier was used to annotate FAP$^+$ αSMA$^-$, αSMA$^+$, collagen IV$^+$, collagen XI$^+$, and collagen XII$^+$ regions. A dedicated script automated the tiling and quantification and resulted in a data file containing per tile annotation measurements for each marker. The resulting measurements were exported and analyzed in R.

Quantifying FAP$^+$αSMA$^+$ CAF Tumor Islet Coverage. Each tumor IHC image was divided into 1,500-μm tiles using QuPath; in each tile, the fraction of tumor nest surface covered by FAP$^+$αSMA$^+$ CAF was estimated. Then, using all tiles, the average FAP$^+$αSMA$^+$ tumor islet coverage was calculated for each patient.

Data Availability

Sequencing data are available in the GEO at GSE183219. The code is available on our Github website: https://github.com/effiken/Grout_et_al. Raw histology data are available at https://scdissector.org/grout/.

Authors' Disclosures

A.H. Rahman reports grants from the NIH during the conduct of the study. S. Gnjatic reports grants from Regeneron, Boehringer Ingelheim, Bristol Myers Squibb, Genentech, Inc., Jannsen R&D, Takeda, and EMD Serono outside the submitted work, as well as a patent for multiplex IHC (MICSSS) issued. J. Adam reports grants and personal fees from MSD, personal fees from Amgen and AstraZeneca, and grants from Bayer outside the submitted work. M.B. Beasley reports personal fees and other support from AstraZeneca, and personal fees from AbbVie outside the submitted work. S. Muller reports personal fees from Genentech, Inc. during the conduct of the study, as well as personal fees from Genentech, Inc. outside the submitted work. S.J. Turley reports other support from Genentech, Inc. outside the submitted work. M. Merad reports grants from Regeneron and Boehringer, personal fees from Compugen, Myeloid Therapeutics, Morphic Therapeutics, AsherBio, DrenBio, Nirogy, OncoResponse, Owkin, Pionyr, OSE Immunotherapeutics, and Larkspur, and personal fees and other support from Innate Pharma and Genenta outside the submitted work. H. Salmon reports grants from Genentech, Inc. during the conduct of the study. No disclosures were reported by the other authors.

Disclaimer

This content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Authors' Contributions

J.A. Grout: Data curation, formal analysis, investigation, writing–original draft. **P. Sirven:** Data curation, formal analysis. **A.M. Leader:** Formal analysis. **S. Maskey:** Investigation. **E. Hector:** Formal analysis. **I. Puisieux:** Formal analysis. **F. Steffan:** Formal analysis. **E. Cheng:** Investigation. **N. Tung:** Investigation. **M. Maurin:** Methodology. **R. Vaineau:** Formal analysis. **L. Karpf:** Methodology. **M. Plaud:** Methodology. **A.-L. Begue:** Investigation. **K. Ganesh:** Investigation. **J. Mesple:** Investigation. **M. Casanova-Acebes:** Investigation. **A. Tabachnikova:** Investigation. **S. Keerthivasan:** Methodology. **A. Lansky:** Project administration. **J. Le Berichel:** Project administration. **L. Walker:** Methodology. **A.H. Rahman:** Methodology. **S. Gnjatic:** Methodology. **N. Girard:** Resources. **M. Lefevre:** Resources. **D. Damotte:** Resources. **J. Adam:** Investigation. **J.C. Martin:** Investigation. **A. Wolf:** Resources. **R.M. Flores:** Resources. **M.B. Beasley:** Resources. **R. Pradhan:** Formal analysis. **S. Muller:** Formal analysis. **T.U. Marron:** Writing–review and editing. **S.J. Turley:** Supervision, funding acquisition. **M. Merad:** Conceptualization, funding acquisition, writing–review and editing. **E. Kenigsberg:** Conceptualization, formal analysis, supervision, writing–review and editing. **H. Salmon:** Conceptualization, resources, supervision, funding acquisition, writing–review and editing.

Acknowledgments

We thank the important contributions of patients who participated in this study. This project was supported by Genentech, Inc. and carried out in collaboration with the Fondation ARC pour la recherche sur le cancer. The computational work was supported by the Scientific Computing at the Icahn School of Medicine at Mount Sinai and the Office of Research Infrastructure of the NIH under award number S10OD026880. This manuscript was edited at Life Science Editors, and the cartoon illustrations were created using BioRender. We thank the Mount Sinai Flow Cytometry Core, the Human Immune Monitoring Center, and the Biorepository and Pathology Core for their support. We thank Sarah Lagha, Anne-Sophie Tedesco, Agathe Seguin-Givelet, and Isabelle Sauret for their contribution to obtaining and studying additional FFPE NSCLC samples from Institut Mutualiste Montsouris. We thank Eliane Piaggio, Ana-Maria Lennon-Duménil, and Olivier Lantz for carefully reading and commenting on the manuscript.

The publication costs of this article were defrayed in part by the payment of publication fees. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

Note

Supplementary data for this article are available at Cancer Discovery Online (http://cancerdiscovery.aacrjournals.org/).

Received January 1, 2022; revised June 10, 2022; accepted August 24, 2022; published first August 26, 2022.

REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359–86.
2. Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, et al. Nivolumab versus docetaxel in advanced nonsquamous non–small-cell lung cancer. N Engl J Med 2015;373:1627–39.
3. Galon J, Bruni D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. Nat Rev Drug Discov 2019;18:197–218.
4. Chen DS, Mellman I. Elements of cancer immunity and the cancer–immune set point. Nature 2017;541:321–30.
5. Herbst RS, Soria JC, Kowanetz M, Fine GD, Hamid O, Gordon MS, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. Nature 2014;515:563–7.
6. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature 2018;554:544–8.

7. Salmon H, Franciszkiewicz K, Damotte D, Dieu-Nosjean MC, Validire P, Trautmann A, et al. Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. J Clin Invest 2012;122:899–910.
8. Mueller SN, Germain RN. Stromal cell contributions to the homeostasis and functionality of the immune system. Nat Rev Immunol 2009;9:618–29.
9. Bajénoff M, Egen JG, Koo LY, Laugier JP, Brau F, Glaichenhaus N, et al. Stromal cell networks regulate lymphocyte entry, migration, and territoriality in lymph nodes. Immunity 2006;25:989–1001.
10. Sixt M, Kanazawa N, Selg M, Samson T, Roos G, Reinhardt DP, et al. The conduit system transports soluble antigens from the afferent lymph to resident dendritic cells in the T cell area of the lymph node. Immunity 2005;22:19–29.
11. Kraman M, Bambrough PJ, Arnold JN, Roberts EW, Magiera L, Jones JO, et al. Suppression of antitumor immunity by stromal cells expressing fibroblast activation protein-α. Science 2010;330:827–30.
12. Ziani L, Chouaib S, Thiery J. Alteration of the antitumor immune response by cancer-associated fibroblasts. Front Immunol 2018;9:414.
13. Wu SZ, Roden DL, Wang C, Holliday H, Harvey K, Cazet AS, et al. Stromal cell diversity associated with immune evasion in human triple-negative breast cancer. EMBO J 2020;39:e104063.
14. Becker LM, O'Connell JT, Vo AP, Cain MP, Tampe D, Bizarro L, et al. Epigenetic reprogramming of cancer-associated fibroblasts deregulates glucose metabolism and facilitates progression of breast cancer. Cell Rep 2020;31:107701.
15. Ford K, Hanley CJ, Mellone M, Szyndralewiez C, Heitz F, Wiesel P, et al. NOX4 inhibition potentiates immunotherapy by overcoming cancer-associated fibroblast-mediated CD8 T-cell exclusion from tumors. Cancer Res 2020;80:1846–60.
16. Feig C, Jones JO, Kraman M, Wells RJB, Deonarine A, Chan DS, et al. Targeting CXCL12 from FAP-expressing carcinoma-associated fibroblasts synergizes with anti–PD-L1 immunotherapy in pancreatic cancer. Proc Natl Acad Sci U S A 2013;110:20212–7.
17. Özdemir BC, Pentcheva-Hoang T, Carstens JL, Zheng X, Wu C-C, Simpson TR, et al. Depletion of carcinoma-associated fibroblasts and fibrosis induces immunosuppression and accelerates pancreas cancer with reduced survival. Cancer Cell 2014;25:719–34.
18. Hofheinz RD, al-Batran SE, Hartmann F, Hartung G, Jäger D, Renner C, et al. Stromal antigen targeting by a humanised monoclonal antibody: an early phase II trial of sibrotuzumab in patients with metastatic colorectal cancer. Oncol Res Treat 2003;26:44–8.
19. Chen Y, McAndrews KM, Kalluri R. Clinical and therapeutic relevance of cancer-associated fibroblasts. Nat Rev Clin Oncol 2021;18:792-804.
20. Öhlund D, Handly-Santana A, Biffi G, Elyada E, Almeida AS, Ponz-Sarvise M, et al. Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. J Exp Medicine 2017;214:579–96.
21. Kieffer Y, Hocine HR, Gentric G, Pelon F, Bernard C, Bourachot B, et al. Single-cell analysis reveals fibroblast clusters linked to immunotherapy resistance in cancer. Cancer Discov 2020;10:1330–51.
22. Elyada E, Bolisetty M, Laise P, Flynn WF, Courtois ET, Burkhart RA, et al. Cross-species single-cell analysis of pancreatic ductal adenocarcinoma reveals antigen-presenting cancer-associated fibroblasts. Cancer Discov 2019;9:1102–23.
23. Lambrechts D, Wauters E, Boeckx B, Aibar S, Nittner D, Burton O, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. Nat Med 2018;24:1277–89.
24. Kim N, Kim HK, Lee K, Hong Y, Cho JH, Choi JW, et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. Nat Commun 2020;11:2285.
25. Laughney AM, Hu J, Campbell NR, Bakhoum SF, Setty M, Lavallée V-P, et al. Regenerative lineages and immune-mediated pruning in lung cancer metastasis. Nat Med 2020;26:259–69.
26. Wu F, Fan J, He Y, Xiong A, Yu J, Li Y, et al. Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. Nat Commun 2021;12:2540.
27. Chen Z, Zhou L, Liu L, Hou Y, Xiong M, Yang Y, et al. Single-cell RNA sequencing highlights the role of inflammatory cancer-associated fibroblasts in bladder urothelial carcinoma. Nat Commun 2020;11:5077.
28. Dominguez CX, Müller S, Keerthivasan S, Koeppen H, Hung J, Gierke S, et al. Single-cell RNA sequencing reveals stromal evolution into LRRC15+ myofibroblasts as a determinant of patient response to cancer immunotherapy. Cancer Discov 2020;10:232–53.
29. Puram SV, Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S, et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. Cell 2017;171:1611–24.
30. Affo S, Nair A, Brundu F, Ravichandra A, Bhattacharjee S, Matsuda M, et al. Promotion of cholangiocarcinoma growth by diverse cancer-associated fibroblast subpopulations. Cancer Cell 2021;39:866–82.
31. Remark R, Merghoub T, Grabe N, Litjens G, Damotte D, Wolchok JD, et al. In-depth tissue profiling using multiplexed immunohistochemical consecutive staining on single slide. Sci Immunol 2016;1:aaf6925.
32. Leader AM, Grout JA, Maier BB, Nabet BY, Park MD, Tabachnikova A, et al. Single-cell analysis of human non-small cell lung cancer lesions refines tumor classification and patient stratification. Cancer Cell 2021;39:1594–609.
33. Martin JC, Chang C, Boschetti G, Ungaro R, Giri M, Grout JA, et al. Single-cell analysis of Crohn's disease lesions identifies a pathogenic cellular module associated with resistance to anti-TNF therapy. Cell 2019;178:1493–508.
34. Gillich A, Zhang F, Farmer CG, Travaglini KJ, Tan SY, Gu M, et al. Capillary cell-type specialization in the alveolus. Nature 2020;586:785–9.
35. Corselli M, Chen CW, Sun B, Yap S, Rubin JP, Péault B. The tunica adventitia of human arteries and veins as a source of mesenchymal stem cells. Stem Cells Dev 2012;21:1299–308.
36. Travaglini KJ, Nabhan AN, Penland L, Sinha R, Gillich A, Sit RV, et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. Nature 2020;587:619–25.
37. Buechler MB, Pradhan RN, Krishnamurty AT, Cox C, Calviello AK, Wang AW, et al. Cross-tissue organization of the fibroblast lineage. Nature 2021;593:575–9.
38. Nurmik M, Ullmann P, Rodriguez F, Haan S, Letellier E. In search of definitions: cancer-associated fibroblasts and their markers. Int J Cancer 2020;146:895–905.
39. Walker C, Rutten F, Yuan X, Pass H, Mew DM, Everitt J. Wilms' tumor suppressor gene expression in rat and human mesothelioma. Cancer Res 1994;54:3101–6.
40. Fletcher AL, Acton SE, Knoblich K. Lymph node fibroblastic reticular cells in health and disease. Nat Rev Immunol 2015;15:350–61.
41. Kerdidani D, Aerakis E, Verrou K-M, Angelidis I, Douka K, Maniou M-A, et al. Lung tumor MHCII immunity depends on in situ antigen presentation by fibroblasts. J Exp Med 2022;219:e20210815.
42. Tang VW. Collagen, stiffness, and adhesion: the evolutionary basis of vertebrate mechanobiology. Mol Biol Cell 2020;31:1823–34.
43. Haruyama N, Sreenath TL, Suzuki S, Yao X, Wang Z, Wang Y, et al. Genetic evidence for key roles of decorin and biglycan in dentin mineralization. Matrix Biol 2009;28:129–36.
44. Park J, Ivey MJ, Deana Y, Riggsbee KL, Sörensen E, Schwabl V, et al. The Tcf21 lineage constitutes the lung lipofibroblast population. Am J Physiol Lung Cell Mol Physiol 2019;316:L872–85.
45. Mecham RP. Elastin in lung development and disease pathogenesis. Matrix Biol 2018;73:6–20.
46. Collisson EA, Campbell JD, Brooks AN, Berger AH, Lee W, Chmielecki J, et al. Comprehensive molecular profiling of lung adenocarcinoma. Nature 2014;511:543–50.
47. Warth A, Muley T, Meister M, Stenzinger A, Thomas M, Schirmacher P, et al. The novel histologic International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society Classification System of lung adenocarcinoma is a stage-independent predictor of survival. J Clin Oncol 2012;30:1438–46.
48. Fletcher AL, Turley SJ. Who am I? (re-)Defining fibroblast identity and immunological function in the age of bioinformatics. Immunol Rev 2021;302:5–9.
49. Chihara T, Suzu S, Hassan R, Chutiwitoonchai N, Hiyoshi M, Motoyoshi K, et al. IL-34 and M-CSF share the receptor Fms but are not identical in biological activity and signal activation. Cell Death Differ 2010;17:1917–27.

50. Menzies-Gow A, Ying S, Sabroe I, Stubbs VL, Soler D, Williams TJ, et al. Eotaxin (CCL11) and eotaxin-2 (CCL24) induce recruitment of eosinophils, basophils, neutrophils, and macrophages as well as features of early- and late-phase allergic reactions following cutaneous injection in human atopic and nonatopic volunteers. J Immunol 2002;169:2712–8.
51. Schall TJ, Bacon K, Toy KJ, Goeddel DV. Selective attraction of monocytes and T lymphocytes of the memory phenotype by cytokine RANTES. Nature 1990;347:669–71.
52. Taub D, Conlon K, Lloyd A, Oppenheim J, Kelvin D. Preferential migration of activated CD4+ and CD8+ T cells in response to MIP-1 alpha and MIP-1 beta. Science 1993;260:355–8.
53. Castellino F, Huang AY, Altan-Bonnet G, Stoll S, Scheinecker C, Germain RN. Chemokines enhance immunity by guiding naive CD8+ T cells to sites of CD4+ T cell–dendritic cell interaction. Nature 2006;440:890–5.
54. Liu YJ, Soumelis V, Watanabe N, Ito T, Wang YH, Malefyt R de W, et al. TSLP: an epithelial cell cytokine that regulates T cell differentiation by conditioning dendritic cell maturation. Annu Rev Immunol 2007;25:193–219.
55. Thomas DA, Massagué J. TGF-beta directly targets cytotoxic T cell functions during tumor evasion of immune surveillance. Cancer Cell 2005;8:369–80.
56. Hegde PS, Karanikas V, Evers S. The where, the when, and the how of immune monitoring for cancer immunotherapies in the era of checkpoint inhibition. Clin Cancer Res 2016;22:1865–74.
57. Tauriello DVF, Palomo-Ponce S, Stork D, Berenguer-Llergo A, Badia-Ramentol J, Iglesias M, et al. TGFβ drives immune evasion in genetically reconstituted colon cancer metastasis. Nature 2018;554:538–43.
58. Holtzhausen A, Zhao F, Evans KS, Tsutsui M, Orabona C, Tyler DS, et al. Melanoma-derived Wnt5a promotes local dendritic-cell expression of IDO and immunotolerance: opportunities for pharmacologic enhancement of immunotherapy. Cancer Immunol Res 2015;3:1082–95.
59. Spranger S, Bao R, Gajewski TF. Melanoma-intrinsic β-catenin signalling prevents anti-tumour immunity. Nature 2015;523:231–5.
60. Chakravarthy A, Khan L, Bensler NP, Bose P, De Carvalho DD. TGF-β-associated extracellular matrix genes link cancer-associated fibroblasts to immune evasion and immunotherapy failure. Nat Commun 2018;9:4692.
61. Caruana I, Savoldo B, Hoyos V, Weber G, Liu H, Kim ES, et al. Heparanase promotes tumor infiltration and antitumor activity of CAR-redirected T lymphocytes. Nat Med 2015;21:524–9.
62. Sekiguchi R, Yamada KM. Chapter four basement membranes in development and disease. Curr Top Dev Biol 2018;130:143–91.
63. Kalluri R. Basement membranes: structure, assembly and role in tumour angiogenesis. Nat Rev Cancer 2003;3:422–33.
64. Vázquez-Villa F, García-Ocaña M, Galván JA, García-Martínez J, García-Pravia C, Menéndez-Rodríguez P, et al. COL11A1/(pro)collagen 11A1 expression is a remarkable biomarker of human invasive carcinoma-associated stromal cells and carcinoma progression. Tumor Biol 2015;36:2213–22.
65. Herbst RS, Baas P, Kim DW, Felip E, Pérez-Gracia JL, Han JY, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet 2016;387:1540–50.
66. Sahai E, Astsaturov I, Cukierman E, DeNardo DG, Egeblad M, Evans RM, et al. A framework for advancing our understanding of cancer-associated fibroblasts. Nat Rev Cancer 2020;20:174–86.
67. Naba A, Clauser KR, Lamar JM, Carr SA, Hynes RO. Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters. Elife 2014;3:e01308.
68. Ramilowski JA, Goldberg T, Harshbarger J, Kloppmann E, Kloppman E, Lizio M, et al. A draft network of ligand–receptor-mediated multicellular signalling in human. Nat Commun 2015;6:7866.
69. Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: open source software for digital pathology image analysis. Sci Rep 2017;7:16878.
70. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods 2012;9:676–82.
71. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 2012;9:671–5.
72. Linkert M, Rueden CT, Allan C, Burel JM, Moore W, Patterson A, et al. Metadata matters: access to image data in the real world. J Cell Biol 2010;189:777–82.
73. Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Comput Vision 2004;60:91–110.