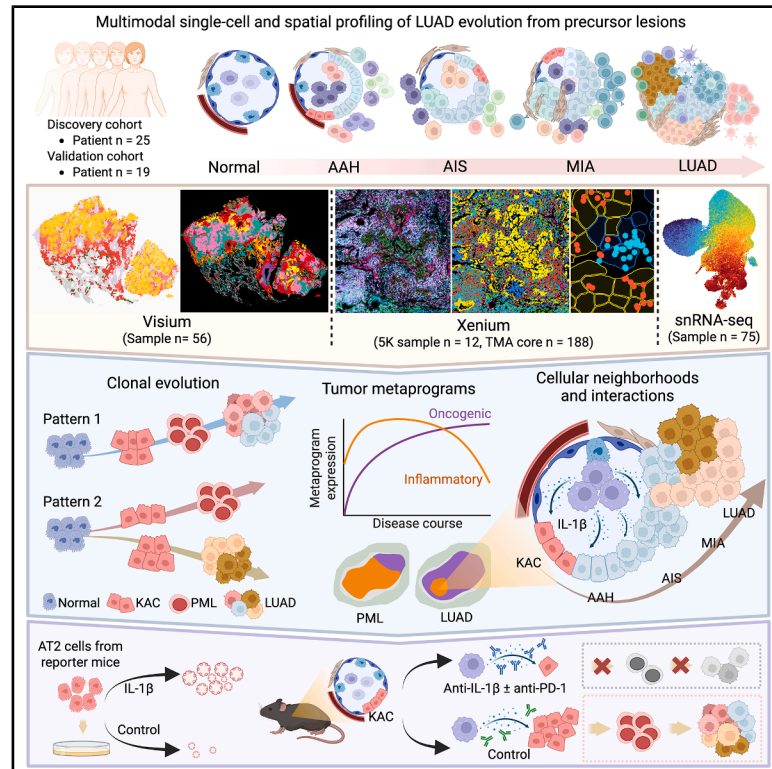


Multimodal spatial-omics reveal co-evolution of alveolar progenitors and proinflammatory niches in progression of lung precursor lesions

Graphical abstract



Authors

Fuduan Peng, Ansam Sinjab, Yibo Dai, ..., Kadir Akdemir, Linghua Wang, Humam Kadara

Correspondence

lwang22@mdanderson.org (L.W.), hkadara@mdanderson.org (H.K.)

In brief

By single-cell and spatial transcriptomics analysis, Peng et al. reveal that alveolar progenitors are the earliest predecessor of LUAD and reside in epithelial-proinflammatory niches that are more common in premalignant stages and enriched with IL-1 β -IL1R1 signaling. Targeting IL-1 β and these niches in mice reduces alveolar progenitors and development of LUAD.

Highlights

- KACs or reactive pneumocytes are early precursors in the evolution of LUAD
- KACs reside in proinflammatory niches with high IL1B-IL1R1
- Epithelial-proinflammatory niches are more common in precursor lesions than in LUAD
- Targeting IL-1 β with PD-1 in precancerous stages reduces KACs and LUAD pathogenesis



Article

Multimodal spatial-omics reveal co-evolution of alveolar progenitors and proinflammatory niches in progression of lung precursor lesions

Fuduan Peng,^{1,22} Ansum Sinjab,^{2,22} Yibo Dai,^{1,3,22} Warapen Treekitkarnmongkol,^{2,22} Sujuan Yang,² Lorena I. Gomez Bolanos,² Tieling Zhou,² Minyue Chen,^{2,3} Alejandra G. Serrano,² Avantika Krishna,^{3,4} Nastaran Karimi,⁴ Manvi Sharma,² Akshay Basi,⁵ Guangsheng Pei,¹ Jianlong Liao,¹ Yunhe Liu,¹ Jiping Feng,² Zahraa Rahal,² Yang Liu,¹ Jiahui Jiang,^{1,3} Kai Yu,¹ Tala Noun,^{2,3} Yuejiang Liu,^{2,3} Khaja Khan,² Kyung Serk Cho,¹ Jichao Chen,⁶ Luisa M. Solis,² Sarah Mazzilli,⁷ Steven Dubinett,⁸ Tina Cascone,⁹ Avrum E. Spira,^{7,10} Stephen Swisher,¹¹ Naoh Jimbo,¹² Takuo Hayashi,¹³ Satsuki Kishikawa,¹³ Kazuya Takamochi,¹⁴ Tomoo Itoh,¹² Takashi Yao,¹³ Kenji Suzuki,¹⁴ Neda Kalhor,¹⁵ Ignacio I. Wistuba,² Mingyao Li,¹⁶ Seyed Javad Moghaddam,^{3,4} Junya Fujimoto,¹⁷ Jared Burks,⁴ Jeffrey Myers,¹⁸ Kadir Akdemir,¹⁹ Linghua Wang,^{1,3,20,21,*} and Humam Kadara^{2,3,18,23,*}

¹Department of Genomic Medicine, UT MD Anderson Cancer Center, Houston, TX, USA

²Department of Translational Molecular Pathology, UT MD Anderson Cancer Center, Houston, TX, USA

³The University of Texas MD Anderson Cancer Center UTHealth Houston Graduate School of Biomedical Sciences, Houston, TX, USA

⁴Department of Pulmonary Medicine, UT MD Anderson Cancer Center, Houston, TX, USA

⁵Department of Leukemia, UT MD Anderson Cancer Center, Houston, TX, USA

⁶Department of Pediatrics, Perinatal Institute Division of Pulmonary Biology, University of Cincinnati and Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

⁷Section of Computational Biomedicine, Chobanian & Avedisian School of Medicine, Boston University, Boston, MA, USA

⁸Department of Medicine, The University of California, Los Angeles, Los Angeles, CA, USA

⁹Department of Thoracic Head & Neck Medical Oncology, UT MD Anderson Cancer Center, Houston, TX, USA

¹⁰Lung Cancer Initiative at Johnson & Johnson, Boston, MA, USA

¹¹Department of Cardiovascular and Thoracic Surgery, UT MD Anderson Cancer Center, Houston, TX, USA

¹²Department of Diagnostic Pathology, Kobe University Hospital, Kobe, Japan

¹³Department of Human Pathology, Juntendo University Graduate School of Medicine, Tokyo, Japan

¹⁴Department of General Thoracic Surgery, Juntendo University Graduate School of Medicine, Tokyo, Japan

¹⁵Department of Anatomic Pathology, UT MD Anderson Cancer Center, Houston, TX, USA

¹⁶Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

¹⁷Clinical Research Center, Hiroshima University Hospital, Hiroshima, Japan

¹⁸Department of Head and Neck Surgery, UT MD Anderson Cancer Center, Houston, TX, USA

¹⁹Department of Neurosurgery, UT MD Anderson Cancer Center, Houston, TX, USA

²⁰The James P. Allison Institute, UT MD Anderson Cancer Center, Houston, TX, USA

²¹Institute for Data Science in Oncology, UT MD Anderson Cancer Center, Houston, TX, USA

²²These authors contributed equally

²³Lead contact

*Correspondence: lwang22@mdanderson.org (L.W.), hkadara@mdanderson.org (H.K.)

<https://doi.org/10.1016/j.ccell.2025.10.004>

SUMMARY

The co-evolution of different cell subsets in the progression of precursor lesions to lung adenocarcinoma (LUAD) is incompletely understood. We generated spatial transcriptomic maps of 56 human precursor lesions and LUADs from 25 patients and of an independent cohort of 36 lesions from 19 patients, analyzing a total of 486,519 spots and 5.4 million cells. We identify region-specific programs that distinguish precursors from LUADs. Spatially resolved clonal architectures reveal patient-specific heterogeneity in evolution of precursors to LUADs. We find epithelial alveolar progenitors expressing tumor-associated meta-programs and residing in niches enriched with proinflammatory subsets including *IL1B* high macrophages. Epithelial-proinflammatory niches are prevalent in precursor lesions but become less frequent in LUADs. These niches are conserved in mice and promote alveolar progenitor growth. Targeting inflammation alone or in combination with immune checkpoint blockade in precancerous phase reduces alveolar progenitors. Epithelial-inflammatory niches are stage-specific, shape early LUAD development and represent promising targets for interception.



INTRODUCTION

Lung adenocarcinoma (LUAD) is increasingly being diagnosed in smokers and never-smokers and has become the dominant histological subtype of lung cancer.¹ Despite clinical advances, mortality of patients with advanced LUAD remains relatively high.² Improving outcomes for patients with LUAD is hindered by the limited understanding of the progression of precursor (pre-malignant or pre-invasive) lesions to invasive carcinoma, including minimally invasive adenocarcinoma (MIA) and LUAD. Lung precursor lesions, such as atypical adenomatous hyperplasia (AAH) and adenocarcinoma *in situ* (AIS), represent a window of opportunity for clinical interception.^{3,4} Yet, their heterogeneous progression trajectories—ranging from indolent regression to aggressive transformation—pose significant challenges for clinical risk stratification and therapeutic decision-making.^{3,4}

Molecular and inflammatory changes are enriched in normal-appearing epithelial niches of LUADs,^{5–8} suggestive of changes that may underlie the development of LUAD. Studies from our group and others using mostly bulk sequencing and imaging approaches have unraveled molecular and immune trajectories along the continuum of normal lung tissue to precursor lesions and invasive LUAD, with many of these alterations occurring early in the precursors.^{9–12} Despite these insights, our understanding of the specific epithelial-immune cell interactions and crosstalk that drive LUAD development remains limited, as does our knowledge of the mechanisms through which they can be targeted for early intervention. This gap in knowledge is largely because the earliest human LUAD precursors, such as AAHs, are rarely surgically resected and are for the most part incidentally found in surgical specimens.^{13,14}

Recent advancements in spatial multi-omics technologies, including spatially resolved transcriptomics, have revolutionized our ability to interrogate tissue architecture with molecular precision.¹⁵ Yet, a comprehensive spatial atlas delineating the complex ecosystems of lung premalignancy and cancer across disease stages remains elusive. Key unanswered questions include how lung precursor cells progress and whether their niches and cellular neighborhoods are distinct from those attributed to more advanced lesions such as MIA and LUAD. Furthermore, the role of geographic heterogeneity—such as regional variations in pre-cancerous and cancer cell states and immune cell properties—in driving evolution remains poorly characterized.

Here, we employed a multimodal omics strategy to generate high-resolution spatial maps that allow us to better understand progression along the pathologic spectrum of normal-appearing lung to precursor lesions up to LUAD. Our study not only advances our fundamental understanding of lung carcinogenesis but also provides a spatial roadmap for context-aware (stage-specific) therapeutic strategies tailored to intercept lung malignancy at its most vulnerable stages.

RESULTS

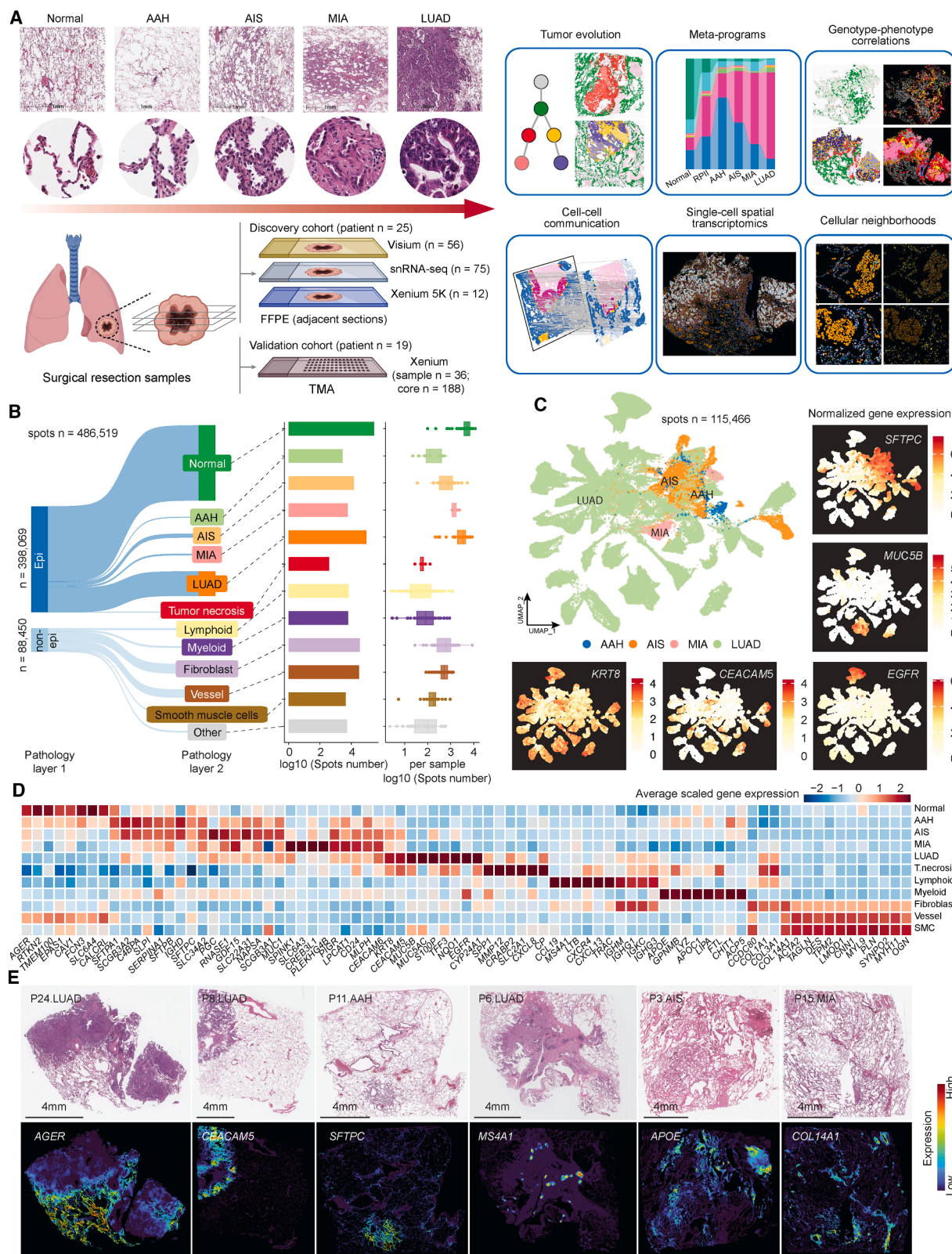
Multimodal spatial-omics reveal diverse region-specific expression patterns mapping to precursor lesions and invasive LUADs

We performed ST analysis using the Visium CytAssist platform (10× Genomics) of 56 tissue samples from 25 patients with

invasive LUAD and that represent distinct stages in the pathologic sequence of LUAD development in order generate high-resolution spatial maps across the pathological continuum of normal tissue to lung adenomatous precursor and invasive lesions (Figure 1A; Table S1). These included 11 AAH, 14 AIS, 4 MIA, and 26 LUAD as well as one lesion-adjacent normal-appearing tissue that was rich with lymphocytic infiltration. Twenty cases had paired lung precursor lesions (AAH and/or AIS) and LUAD and four cases had paired MIA and LUAD. Further pathology review identified, by consensus, that one lung precursor tissue is a LUAD (P7 with 2 LUADs). To enhance spatial resolution, we integrated ST data with snRNA-seq profiles from a matched set of 75 samples, including uninvolved normal lung tissues, from 23 of the 25 patients (Figure 1A, Table S1). We also performed single-cell spatial analysis in a subset of cases ($n = 6$ cases, 12 samples, Table S1). For in-depth validation, we performed single-cell ST of a tissue microarray (TMA) composed of 36 lung precursor and invasive lesions from 19 patients and spotted as 188 tissue cores (Figure 1A; Table S1).

After quality control (QC) and filtering, 486,519 spots from the Visium ST dataset were retained for downstream analyses. We performed detailed pathological annotation of all spots to rigorously integrate gene expression with histopathological information. This analysis identified 398,069 spots that mapped to regions rich with normal epithelia and lesions as well as 88,450 spots that covered diverse non-epithelial subsets and structures (Figure 1B). Spots from precursor lesions (AAH and AIS) largely clustered closely together whereas spots from invasive LUADs showed markedly increased heterogeneity (Figures 1C and S1A). **Relative to spots mapping to invasive lesions, AAH and AIS spots expressed higher levels of genes associated with alveolar type II (AT2) cells such as *SFTPC*, and lower levels of markers associated with AT2 dedifferentiation such as *KRT8* and *CEACAM5* (Figures 1C and 1D). LUAD-associated spots also showed cluster-specific expression of tumor markers such as *EGFR*, *NQO1*, and *TFF3* as well as elevated expression of mucin-related genes such as *MUC5B*, suggestive of increased plasticity and alveolar dedifferentiation in invasive LUAD compared to lung precursor lesions (Figures 1C, 1D, and S1A).**

Genes differentially expressed across all 486,519 spots were tightly associated with specific cell lineages or lineage-restricted structures, including lymphoid/lymphoid aggregates (*MS4A1*, *CXCL13*), myeloid (*APOE*, *GPMNB*), and fibroblast/stromal (*COL1A1*, *ACTA2*, *DES*). Genes also closely denoted disease stage such as normal-appearing lung (*AGER*, *RTKN2*, markers of alveolar type I cells-AT1), precursor lesions (*SFTPC*, *PGC*), MIA (*CREB3L1*) and LUAD (*MUC5AC*, *TFF3*) (Figure 1D). Spots corresponding to AAH, AIS, and MIA shared more expression features among each other than with LUAD spots (Figure 1D). By enhancing spatial resolution with iStar,¹⁶ we found that differentially expressed genes (DEGs) showed robust spatial associations with distinct tissue regions, e.g., normal tissue (*SFTPC* and *AGER*), tumor (*CEACAM5*), as well as immune and stromal components (*MS4A1* for lymphoid; *APOE* for myeloid; *COL1A1* for stromal) (Figure 1E).



(legend on next page)

Spatial landscape of copy number alterations and clonal architectures in precursor lesions and LUADs

To understand spatial clonal heterogeneity and evolutionary trajectories, we applied SpatialInferCNV¹⁷ to infer genome-wide copy number alterations (CNAs) in epithelial spots from Visium ST data. We then reconstructed clonal architectures for each patient by integrating CNAs from paired precursor and invasive lesions, followed by correlative analysis with histological features and tumor properties (Figures 2A, S1B, S1C, and S2A–S2D). For QC, we assessed the influence of tumor spot purity on the performance of CNA inference using Visium ST data, (Figure S2A) and confirmed inferred CNAs in select cases using paired snRNA-seq and WES data (Figure S2B). We stratified cases into two broad patterns based on phylogenetic trees (Figure 2B) and shared clones between precursor and invasive lesions, with representative examples shown in Figure 2C. In pattern 1a ($n = 5$), clones detected in precursor (preinvasive) lesions (e.g., AAH, AIS) were also present in paired invasive lesions which acquired additional subclones (Figures 2B, 2C, and S3). Unlike pattern 1a, pattern 1b ($n = 13$) exhibited shared clones between preinvasive and invasive lesions, while it also harbored stage-specific clones (Figures 2B, 2C, and S3). In pattern 2 ($n = 5$), no clones were shared between preinvasive and invasive lesions (Figures 2B, 2C, and S3). Additionally, one case (P22) was chromosomally stable, and another (P7), which had only two LUAD samples, was uncategorized (Figure S3). As expected, LUAD lesions showed significantly higher levels of aneuploidy, with CNAs detected in all but one case (P22). Notably, the two LUADs from P7 had markedly different CNA profiles (Figure S3).

To further dissect spatial dynamics, we mapped each clone by color-coding its corresponding spots, alongside histology annotations, pseudotemporal trajectories and inferred differentiation states from CytoTRACE¹⁸ (Figures 2C, right, S2D, and S3). Co-localization analysis revealed relationships between histological, spatial, and molecular features. We found that, in most cases, the earliest clones mapped to reactive pneumocytes (RPII) (Figures 2C, right, and S3) that we previously suggested to be the histological resemblant of KRT8 high alveolar intermediate cells that function as LUAD progenitors.¹⁹ In case P20 (pattern 1a), clone A was enriched in RPII and AAH, shared with the LUAD, and was spatially aggregated. In the paired LUAD, clones B and C dominated the LUAD which localized to the tumor core (Figure 2C). In case P23 (pattern 2), preinvasive clones B (B1–4) detected in the two AIS lesions were distinct from invasive clones C and D. None of the preinvasive clones B were detected in the paired LUAD, which was dominated by clone C (high aneuploidy) and a small clone D, spatially adjacent to clone C but with

additional CNAs on chromosome 6 (Figure 2C, right). Of note, precursor-specific and invasive-specific clones were both found in different RPII from P23 (Figure 2C, right), further suggesting the involvement of these pneumocytes in early stages of LUAD development. When incorporating driver mutation information, we found that mutations in *EGFR*, *KRAS*, and *MET* were predominantly observed in pattern 1b cases, but were less frequent in pattern 1a and 2 cases. Notably, *KRAS* mutations were entirely absent from pattern 1a cases (Figure 2D). Our findings demonstrate patient-specific heterogeneity in clonal relationships between paired precursor and invasive lung lesions and in the manner of previous studies.^{11,12,20}

Reactive pneumocytes are KRT8⁺ alveolar cells that represent early precursors of LUAD

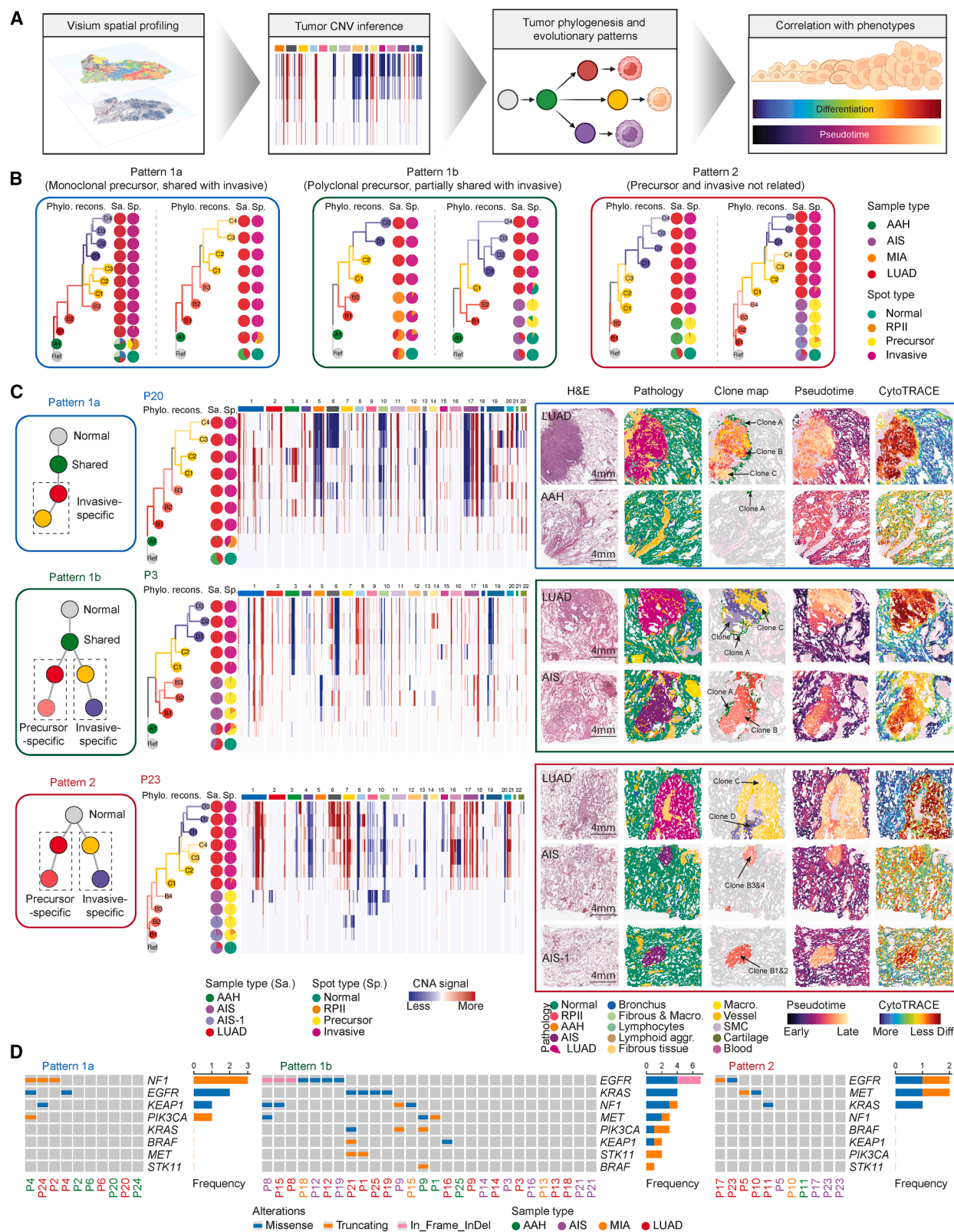
We performed snRNA-seq analysis of cells derived from fixed tissue sections adjacent to those profiled by ST. After rigorous QC and filtering, 401,635 nuclei representing diverse lineages were retained. We identified changes in various immune and stromal cell subsets by disease stage (Figures S4A–S4C). These included increased frequencies of B, plasma, CD4⁺ T, T follicular helper (Tfh), Th17, and Treg cells, as well as reductions in NK cells and monocytes with disease stage (Figure S4D). We then analyzed 139,663 nuclei from epithelial cells identifying distinct lineages including ciliated, club, basal, alveolar (AT1, AT2) and alveolar intermediate cells (AICs). In line with our recent report,¹⁹ we also noted a KRT8 high alveolar intermediate cell (KAC) subset positioned between AT2 and tumor cells (Figure 3A), and that clustered closely with cells from precursor lesions (Figure 3B). KACs represented the intersection between two routes, an AT2 to AT1 trajectory as well as one that led to invasive cells showing progressive pseudotime increases (Figure 3C) and loss of differentiation (Figure 3D).

Non-negative matrix factorization (NMF) analysis of snRNA-seq data identified 9 distinct consensus meta-programs (MPs), each representing unique transcriptional identities (Figure S4E; Table S2) and that were associated with major cell lineages (Figures S4E and S4F). Similar analysis of the epithelial cell lineages also defined 9 MPs (Figure 3E; Table S2). Epithelial MPs showed lineage-specific patterns (e.g., basal, ciliated, club, and AT1 MPs) as well as association with lung lesions (e.g., MP6). Notably, the MP landscape of KACs resembled that of lesions, including reduced MP2 (alveolar) and more prevalent tumor-associated MP6 (Figure 3F). We correlated our derived MPs with those described in two recent studies of pan-cancer cell states^{21,22} (Figure 3G). We found strong correlation between our MPs and previously reported MPs with respect to both

Figure 1. Spatial transcriptomics analysis identifies region-specific expression patterns that distinguish lung precursor and invasive adenocarcinoma lesions

(A) Schematics illustrating the strategy for multimodal single-cell and spatial analysis of multiple sections from tissues representing different stages in the sequence of development of invasive lung adenocarcinoma (LUAD) including atypical adenomatous hyperplasia (AAH), adenocarcinoma *in situ* (AIS), and minimally invasive adenocarcinoma (MIA) along with the different types of analyses conducted to generate spatial maps along this pathologic continuum. (B) Sankey diagram of epithelial and non-epithelial spots that were analyzed by Visium ST and that map to each of the indicated histological and pathological annotations. Box middle lines, median; box limits, upper and lower quartiles; box whiskers, 1.5× the interquartile range. (C) UMAP projection of tumor (precursor and invasive) epithelial spots color coded by lesion stage, with feature plots of select tumor genes. (D) Heatmap showing genes that are differentially expressed between spots with distinct histological and pathological annotations. (E) Images of H&E-stained tissues representative of different lung precursor and invasive lesions, along with their expression of select marker genes at super-pixel resolution.

See also Figure S1 and Table S1.



(legend on next page)

lineage (e.g., ciliated) and phenotype (e.g., inflammatory/stress; Figure 3G). We observed dynamic shifts in MP along the normal-to-malignant transition (Figures 3H and S4G–S4J). MP2-AT2 and MP5-AT1 significantly and progressively decreased, while MP6-tumor/KAC increased by lesion stage (Figures 3H and S4J). We observed similar patterns when we integrated snRNA-seq and ST data and analyzed MPs spatially. Non-invasive regions (i.e., normal lung or precursor lesions) showed higher MP2 and MP7 and lower MP6 compared to invasive regions (Figure 3I). Also, there was significant and progressive decrease in MP2-AT2 and MP5-AT1, and an increase in MP6-tumor/KAC along the spectrum of normal-appearing tissue to precursor lesions (AAH and AIS) and MIA up to LUAD (Figure S4K).

By comprehensive pathological annotation of epithelial niches in the ST data, we found that RPII, which we found previously (Figures 2 and S3) to be involved in evolution of LUAD, were indeed the histological ressemblant of KACs *in situ*.¹⁹ RPII sparsely distributed in lesion-adjacent regions, retained alveolar characteristics (MP2-AT2) but co-expressed MP6-tumor/KAC and MP7-inflammatory/stress, implicating their role in early phenotypic deviation from normal alveolar cells to tumors (Figure 3J). We then mapped MPs and clonal information onto ST data with resolution enhanced from spot to super-pixel level using iStar¹⁶ (Figures 3K and S4L) and found that RPII were confined to regions surrounding lung precursor and invasive lesions and displayed MPs both unique to alveolar space (MP2) as well as shared with lesions (e.g., MP6) (Figures 3J and 3K). RPII had a landscape of ST MPs similar to precursor lesions and KACs in our snRNA-seq data (Figures 3L, S5A, and S5B). We also noted a tendency for different clones to display variable expression of MPs, such as alveolar MP2 and MP6, even within the same lesions (e.g., P12 AIS and LUAD; Figure 3K), suggesting a tight association between clonal and lineage evolution. RPII were often found in regions with shared clones between normal tissue and precursor lesions (e.g., P24 AAH; Figure 3K) as well as between paired AAH and LUAD in cases such as P24 with clonal evolution pattern 1a (Figures 3K and S3). Along this theme, particularly in cases with evolution pattern 1a, most of the clones mapping to RPII were shared across all tissues and lesions (Figures 3M and S5C). In addition, spots located in shared clones, as well as in precursor- and invasive-specific clones, showed significantly increased levels of a KAC score we previously derived¹⁹ compared with reference spots (Figures 3M and S5C; Table S3). We also noted that, in some cases, RPII resided in spots whose clones were specific to either the lung precursor or invasive lesion, such as in cases of evolution patterns 1b and 2 (Figures 3K, 3M, 3N, S5C, and S5D, P12 AIS as a representative). These findings highlight alveolar progenitors (RPII) as an entity that partakes in the pathologic sequence of LUAD development earlier than recognizable lung precursor lesions.

Alveolar progenitors and precursor lung cells reside in proinflammatory niches

To begin to understand mechanisms of KAC oncogenesis, we performed pathways analysis of epithelial cells. We found that proinflammatory gene sets were most evidently enriched in KACs compared to other epithelial cell subsets (Figure 4A). To discern cell phenotypes that influence the pathobiology of RPII, we categorized expression features of all 486,519 spots into 11 MPs (Table S2). These MPs were enriched for genes specific to cell lineages such as normal alveolar (normal parenchyma; MP1 and MP10) and myeloid (MP5) cells, as well as genes pertinent to lung precursor (MP3) and invasive (MP9) lesion spots (Figure 4B). We also identified strong correlations between MPs and histologically annotated lineages or structures (e.g., MP4 lymphoid mapping to lymphocytes and myeloid MP5 mapping to macrophages). When considering lesion stage, RPII and lung precursor lesions (AAH and AIS) showed the strongest correlation with precursor MP3 and, conversely, MIA and LUAD correlated most highly with MP9 (Figure 4B). We then tested correlations among MPs and found significant positive correlations between lymphoid MP4 and plasma cell MP11 ($\rho = 0.606$, adjusted p value <0.0001) and between precursor MP3 and normal alveolar cell MP10 ($\rho = 0.374$, adjusted p value $= 0.013$). Notably, we also found a significant positive correlation between precursor MP3 and myeloid MP5 (macrophages) ($\rho = 0.325$, adjusted p value $= 0.037$). In this theme, RPII and lung precursor lesions were enriched for myeloid MP5 spots. These findings suggest that RPII and lung precursor lesions are in close spatial proximity to myeloid cell subsets.

Given our previous findings, we next enumerated the frequencies of ligand-receptor interactions between myeloid and epithelial cells focusing on those involving inflammatory response genes that were upregulated in KACs. We found that *IL1B-IL1R1* was among the top ligand-receptor interactions enriched in precursors (Figure S5E). Within the top 8 ligand-receptor pairs, *IL1B-IL1R1* axis displayed two unique properties. Unlike other epithelial ligands or receptors (e.g., *CD47*, *AREG*), *IL1R1* was significantly and highly expressed in KACs and precursor cells, and most evidently in precursor lesions, and absent in normal AT1 and AT2 subsets (Figures 4C, 4D, and S5F). We also found that genes (*RELA*, *RELB*, and *NFKB1*) encoding different subunits of NF- κ B, which interplays with pro-inflammatory cytokines including IL-1 β ,²³ were evidently increased in KACs and precursor lesions compared to normal alveolar subsets or LUAD cells (Figure S5F). *IL1B* itself in *IL1B* high macrophages was increased in earlier lesions compared to LUAD (Figure S5G). When we stratified analysis of ligand-receptor interactions by patient-level mutation status of *KRAS*, we found that in *KRAS*-mutant cases, *IL1B-IL1R1* interactions were

Figure 2. Identification of tumor phylogeny and evolutionary patterns in progression of precursor lesions to invasive LUADs

- (A) Overview of analysis showing inference of tumor spatial copy number alterations from ST data.
(B) Identification of two broad evolutionary patterns defined by whether the clones are fully, partially or not shared between precursor (preinvasive) and invasive samples.
(C) Representative cases for each evolutionary pattern. Right panels show corresponding paired preinvasive and invasive tissues with H&E staining, pathology annotations for reactive pneumocytes (RPII), precursor and invasive lesions, and other cell subsets and structures, as well as *in situ* clones, pseudotime, and CytoTRACE distribution.
(D) Frequency of select driver mutations across individual patient samples and grouped by evolutionary pattern.
See also Figures S1–S3.

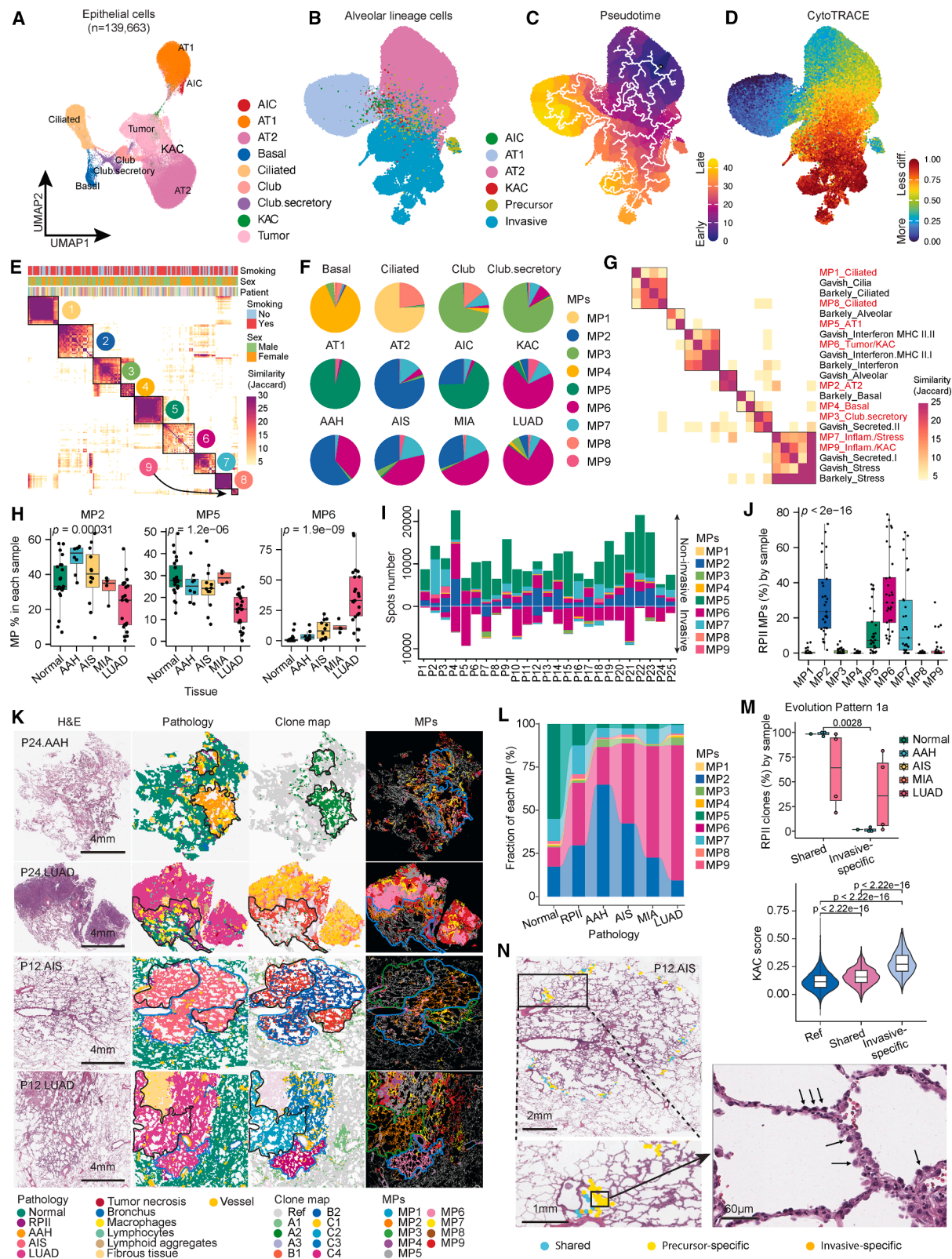


Figure 3. snRNA-seq and ST identify reactive pneumocytes as early precursors of LUAD

(A) UMAP distribution of epithelial cells from all tissues (n = 75) analyzed by snRNA-seq.

(B) UMAP clustering of cells of alveolar, alveolar intermediate, precursor, or invasive lineages.

(legend continued on next page)

markedly more prevalent in precursor lesions (100%) than in invasive lesions (Figure S5H). Among precursor-specific ligand-receptor pairs, *IL1B-IL1R1* was the highest among precursors of *KRAS*-mutant cases. Importantly, this precursor lesion-specific pattern for enrichment of *IL1B-IL1R1* (and other ligand-receptor pairs) was not observed in *KRAS* wild type (WT) cases (Figure S5H). We also found that precursor lesion spots of *KRAS*-mutant cases showed significantly higher levels of signatures of NF- κ B, which promotes *KRAS*-mutant tumorigenesis and inflammation,^{24,25} and interferon signaling compared to spots mapping to precursor lesions of *KRAS* WT cases (Figure S5I; Table S3). This observation extended to RPII and even normal regions (Figure S5I). In contrast, invasive (LUAD) spots showed the opposite trend where they showed significantly reduced NF- κ B and interferon signature levels in *KRAS*-mutant compared to WT cases (Figure S5I).

Given these findings, we thus performed more in-depth analysis of spatial patterns of *IL1B-IL1R1* signaling. When we spatially analyzed *IL1B* and *IL1R1* interaction including expression levels of both genes at super-pixel resolution and using samples from P24 and P3 as representative examples, we found that *IL1R1* mapped to precursor lesions and more evidently so when compared to matched LUAD, whereas *IL1B* predominantly mapped to areas within and nearby lesions (Figures 4E, 4F, and S6A–S6F). Of note, *IL1R1* expression mapped to the edge of the invasive LUAD at the interface with normal parenchyma (Figures 4F and S6D) and this overlapped with areas rich with *IL1B-IL1R1* interactions (Figures 4E, S6A, S6C, and S6E). We then mapped the 9 consensus epithelial MPs at super-pixel resolution using iStar¹⁶ along with the myeloid MP (here denoted as MP10), and which we also visualized with *IL1B* expression (Figure 4G, S6B, S6F, and S6G). This analysis showed that RPII near the paired AAH and LUAD as well as the AAH cells themselves were in niches enriched with *IL1B* high macrophages (Figures 4G, S4B, and S6F–S6G).

Epithelial-proinflammatory niches are stage-specific and enriched in precursor lesions

We then performed single-cell *in situ* analysis using the Xenium platform with a predesigned 5K Prime gene panel and supplemented by a custom-designed 100-gene add-on (Table S4) on 6 pairs of lung precursor lesions (3 AAH and 3 AIS) and their

matched invasive LUADs (Table S1). This analysis yielded 4,598,777 cells that comprised diverse lineages including epithelial, lymphoid, myeloid, stromal, and endothelial subsets (Figures 5A and S6H–S6I). As expected, normal epithelial cells ($n = 498,152$) comprised alveolar and other lineages such as ciliated, basal, club, goblet and neuroendocrine cells (Figure S6H). We identified KACs by analysis of transcriptomic features of alveolar lineage cells, which were positioned between AT1 and AT2 cells (Figure 5B). This analysis also revealed an inflamed AT2 subcluster, marked by elevated *CXCL2* and *NFKBIA/Z* expression (Figure 5B; Table S4). KACs indeed displayed the highest expression of the KAC score we previously derived,¹⁹ while the inflamed AT2 subcluster showed peak expression of an inflammatory signature (Figure 5C; Table S3; see STAR Methods), thus supporting our subclustering of alveolar lineage cells *in situ*. Epithelial cells, especially KACs and precursor and invasive cells, displayed varying structures of cell neighborhoods (Figure 5D).

We then subclustered myeloid cells (Figure S6I, right) and found that two macrophage subclusters (C15 and C16) displayed the highest expression of *IL1B* (Figure S6J) as well as elevated expression of an inflammatory signature (Figure S6K; Table S3). Cell neighborhoods of KACs, AAH, and inflamed AT2 cells contained significantly higher numbers of C15 cells compared to all other epithelial subsets, including normal alveolar cells and more advanced tumor cells (Figure 5E). Similar observations were found for C16 albeit with an attenuated trend (Figure S6L). Of note, C15 showed higher *IL1B* expression than C16 (Figure S6J). KAC clusters mapped to regions harboring RPII, which expressed high levels of *IL1R1* and indeed were in close proximity to macrophages with high *IL1B* expression (Figure 5F). Of note, AAH cells also displayed high *IL1R1* levels and were frequently neighboring *IL1B*-high macrophages (Figures 5F and S6M). It is important to note that we found other proinflammatory cell subsets that were enriched in the neighborhoods of KACs and/or early precursor lesions including *CCL2*⁺, *IL18*⁺, *NFKB1*⁺ and *CSF1*⁺ immune populations (Figure S7A).

We also performed Xenium analysis of a predesigned 298-gene human lung panel along with a custom-made add-on of 100 genes in a uniquely designed TMA composed of 36 lung precursor and invasive lesions (including normal lung parenchyma indicated as normal lung) from 19 patients and that

(C and D) Pseudotime score (C) and CytoTRACE score distribution (D) for cells in (B).

(E) Heatmap showing Jaccard similarity indices among 168 robust non-negative matrix factorization (NMF) programs derived from snRNA-seq of epithelial cells.

(F) Meta-program (MP) distribution by cell type in snRNA-seq data.

(G) Correlation between 9 epithelial MPs identified here and those defined by two previous reports on pan-cancer cell states that are indicated by lead authors.

(H) Percentage of MP 2, 5, and 6 by sample and across normal lung tissues ($n = 24$) and lesion types (n AAH = 9, AIS = 14, MIA = 4, LUAD = 24). Box middle lines, median; box limits, upper and lower quartiles; box whiskers, 1.5 $\times$ the interquartile range. p values were calculated using Kruskal-Wallis tests.

(I) ST spot numbers labeled by snRNA-seq-derived MPs and distributed across non-invasive (normal tissues and precursor lesions) tissues and invasive (MIA and LUAD) lesions.

(J) Percentage of MPs across RPII spots per sample analyzed ($n = 34$ for each group). Box middle lines, median; box limits, upper and lower quartiles; box whiskers, 1.5 $\times$ the interquartile range. p value was calculated using the Kruskal-Wallis test.

(K) Mapping of MPs on ST data in representative tissues at super-pixel resolution and showing correlation with pathology annotations and clone types.

(L) Fractional distribution of MPs across spots covering normal regions, RPII and lesions.

(M) Percentage of clone type in RPII across samples from patients characterized by evolution pattern 1a (top) and KAC score assessment in the clone groups (bottom). Box middle lines, median; box limits, upper and lower quartiles; box whiskers, 1.5 $\times$ the interquartile range. p values were calculated using unpaired Student's t tests.

(N) H&E image of P12 AIS tissue with labeled RPII and their clone types.

See also Figures S4 and S5; Tables S2 and S3.

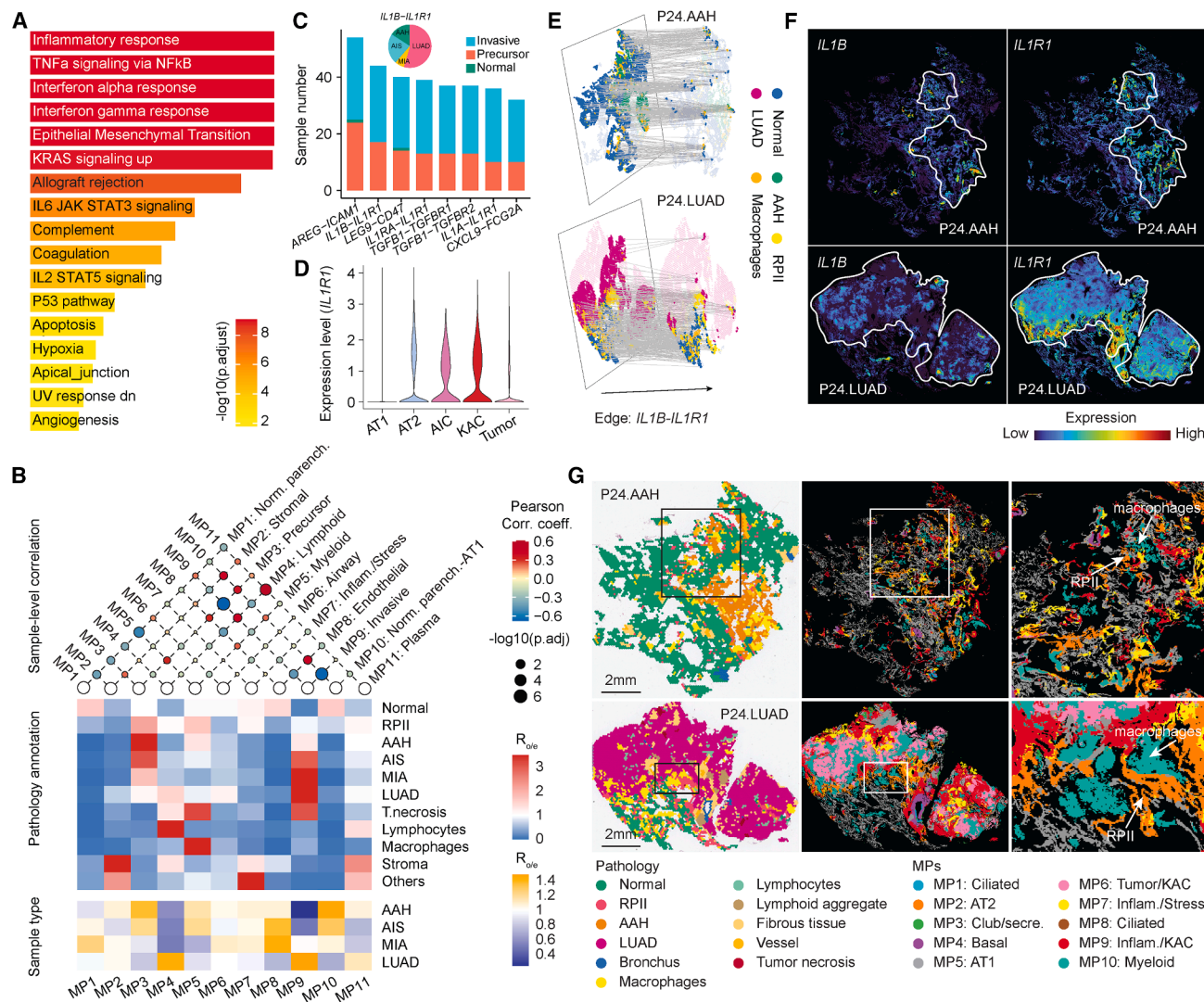
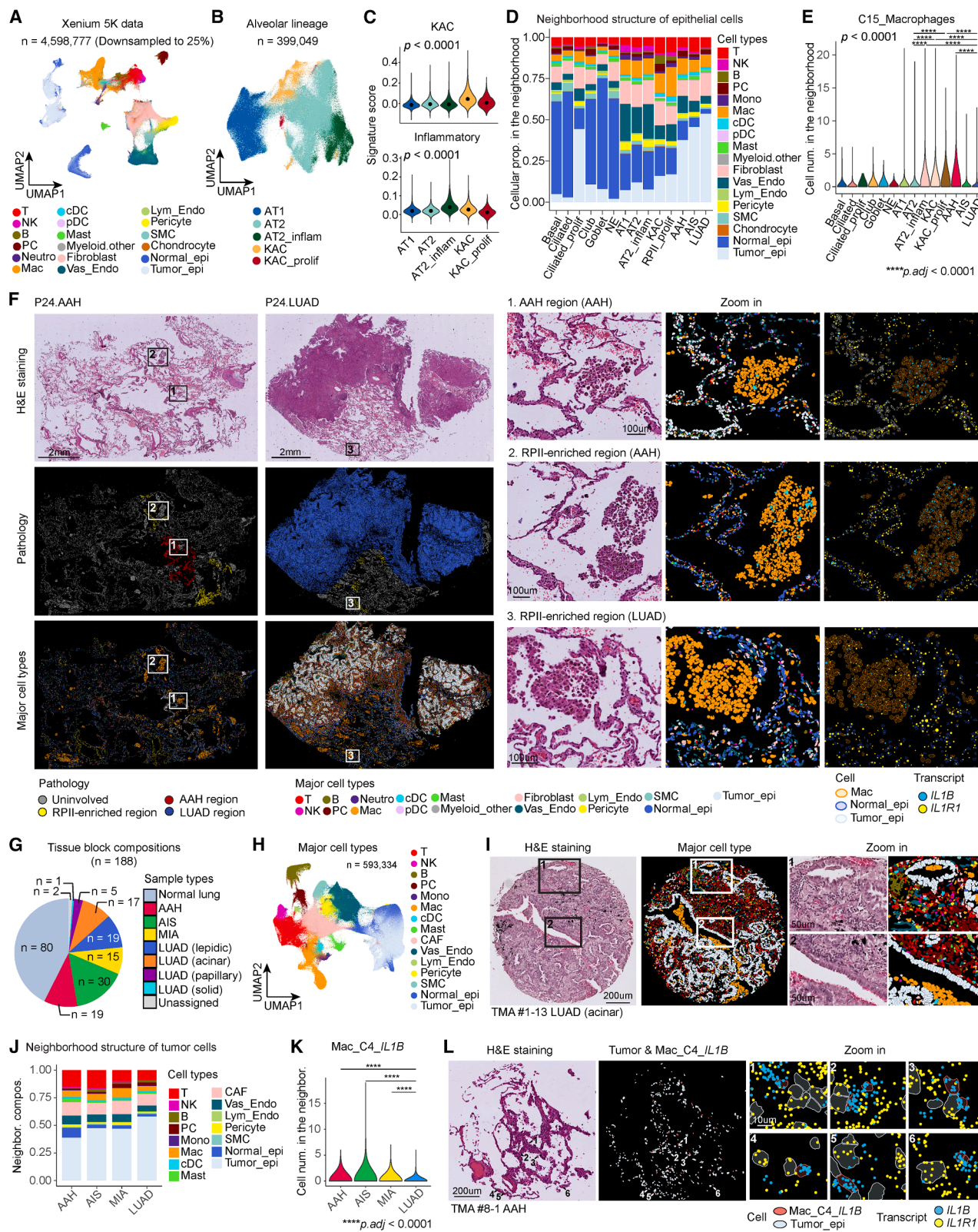


Figure 4. Reactive pneumocytes reside in niches enriched with proinflammatory macrophages

(A) Pathway enrichment results of KACs relative to other epithelial subsets from snRNA-seq data.
 (B) Top: sample-level correlation between 11 MPs derived from Visium ST data. Middle: correlation of MPs with cell lineages, structures and specific histological and pathology annotations. Bottom: correlation of MPs with disease stage (bottom).
 (C) Frequencies of select ligand-receptor pair interactions between myeloid and epithelial cells in ST data. Embedded pie chart represents *IL1B-IL1R1* interactions across tissue types.
 (D) snRNA-seq-derived expression levels of *IL1R1* in AT1, AT2, AIC, KAC and tumor cells.
 (E) Spatial analysis of *IL1B-IL1R1* interactions in P24 AAH and P24 LUAD. RPII: reactive type II pneumocytes.
 (F) Expression levels of *IL1B* and *IL1R1* at super-pixel resolution and in samples shown in (E).
 (G) Super-pixel resolution spatial mapping of the 9 epithelial MPs (from Figure 3E) as well as myeloid MPs generated from snRNA-seq of all cells, here labeled as MP10 in P24. Arrows point to RPII and macrophages in magnified regions.
 See also Figures S5 and S6 and Tables S2 and S3.

were spotted as 188 1 mm cores (Figure 5G; Tables S1 and S4). These samples comprised 593,334 cells that were found across regions of varying tissues and lesions (normal, AAH, AIS, MIA, and LUAD) including distinct histological subtypes of LUAD (Figure S7B). Once again, cells comprised diverse lineages (Figures 5H and 5I) and showed varying proportions across different disease stages (Figure S7C). NK cells were depleted across all lesions as early as AAH whereas T regulatory cells and a C1 subcluster of cancer-associated fibroblasts (CAFs)

increased with disease stage (Figure S7D). Of note, an inflammatory CAF (iCAF) subcluster was most abundant in lung precursor lesions (AAH, AIS) compared to normal lung and invasive cancers (Figure S7D), in line with our findings on increased inflammatory cues in the niche of early lung lesions. As previously shown, lung lesions exhibited distinct neighborhood structures (Figure 5J). Upon subclustering macrophages, we again identified a C4 subcluster with high *IL1B* expression (Figures S7E–S7F) which was significantly more enriched in the



330 Cancer Cell 44, 321–339, February 9, 2026

neighborhoods of AAH, AIS, and even MIA, compared to LUAD (Figure 5K). In line with our findings previously, the C4 macrophage subcluster expressing high levels of *IL1B* was present in the local niche of early lung lesions, such as AAH (Figure 5L) and AIS (Figure S7G) which in turn displayed high expression of the IL-1 β receptor *IL1R1*. Like before (Figure S7A), neighborhoods of precursor lesions, when compared to LUAD, exhibited increased numbers of other proinflammatory subsets (Figure S7H). These findings reveal epithelial-proinflammatory niches that are not only involved in LUAD pathogenesis but are also stage-specific, being prevalent in early lung precursor lesions.

Epithelial-proinflammatory niches are enriched early on in *Kras*-mutant lung oncogenesis

We next interrogated the functional role of proinflammatory niches in the oncogenesis of early LUAD precursor cells. First, we analyzed samples from our previously reported scRNA-seq dataset of enriched epithelial populations derived from a carcinogenesis mouse model of LUAD,¹⁹ in which mice with knockout of the lung lineage gene, *Gprc5a* (*Gprc5a*^{-/-}), show accelerated development of lung tumors with somatic *Kras* mutations upon exposure to nicotine-derived nitrosamine ketone (NNK), a tobacco carcinogen (Figure 6A). We analyzed the lungs of mice exposed to NNK or control saline at the end of exposure (EOE) and at 3 and 7 months post EOE, the time points of onset of early lesions and LUADs, respectively (Figures 6A, S8A, and S8B). Tumor cells clustered close to alveolar subsets, and KACs were tucked at the center of AT1, AT2, and tumor cells (Figures 6B and S8C). Compared to other alveolar and tumor subsets, KACs showed upregulation of various pathways including those associated with inflammation such as TNF α signaling via NF- κ B (Figure 6C). In line with our human data (Figure S5F), we found that mouse KACs, compared to AT1, AT2, and even tumor cells, expressed the highest levels of *Il1r1* and genes encoding different subunits of NF- κ B, particularly *Rela* and *Relb* (Figure 6D). This pattern was not observed with *Stat3* (Figure S8D), another major driver of pro-tumor inflammation in LUAD.²⁶ Of note, *Il1b* expression in macrophages increased

with time post-NNK, whereas *Il1r1* somewhat peaked at 3 months post-NNK at the time of development of early lesions (Figure S8E).

We next analyzed ST data from lungs of NNK-exposed mice collected at EOE and at 7 months post exposure¹⁹ (Figures 6E and S8F). Contrary to tissues analyzed at EOE, which lacked visible lesions, tumors were consistently surrounded by macrophage-enriched niches (Figures 6E and S8F). Colonies of RPII were abundant in regions surrounding these early lesions and were occasionally dispersed in normal-appearing parenchyma (Figure 6E). Super-pixel resolution analysis using iStar¹⁶ and visualization of both our previously reported mouse KAC signature¹⁹ and an inflammatory signature showed that KAC and proinflammatory signals were concurrently high and co-localized to tumor regions and their peritumoral macrophage-rich niches (Figures 6F and S8F; Table S3), very much in line with patterns we observed for human RPII and precursor lung lesions. Further analysis of ligand-receptor interactions in these tissues identified highly condensed *Il1b*-*Il1r1* signals originating from areas surrounding the lesions, which we have previously shown to be rich with macrophages, and consistently shuttling toward the tumor cores (Figure 6G).

Given the similar patterns we found for proinflammatory niches of KACs and early lung tumors in human tissues as well as in lungs of mice that are prone for development of *Kras*-mutant tumors, we focused our attention to functionally interrogate the role of the IL-1 β -IL1R1 axis and its effects on KACs and LUAD development. Knockout of *Il1r1* in mice with lung epithelium-driven expression of *Kras*^{LSL-G12D} (previously referred to as CC-LR²⁷), significantly reduced *Kras*-mutant lung tumor development as well as abundance of LAMP3⁺/KRT8⁺ cells in normal alveolar regions (Figures S8G–S8I). To examine effects of IL-1 β on KACs, we derived KAC-rich organoids (KRT8⁺) organoids from parent AT2 cells of NNK-exposed *Gprc5a*^{-/-} mice that harbor KRT8⁺ lineage reporters (*Gprc5a*^{-/-}; *Krt8*^{CreER}; *Rosa*^{tdT/+})¹⁹ (Figure 6H), and which were labeled following *in vitro* exposure to 4-hydroxytamoxifen (Figure 6H). Treatment with recombinant IL-1 β or coculture with lung macrophages (interstitial macrophages, IM) that were isolated from the same

Figure 5. Single-cell spatial transcriptomics demonstrate that *IL1R1* high epithelial-*IL1B* high proinflammatory niches are enriched in lung precursor lesions

- UMAP embeddings of cells (down sampled to 25% for visualization) analyzed by Xenium using the Human 5K Prime panel.
- UMAP embeddings of alveolar cells, including AT1, AT2, AT2 inflamed (AT2_inflam), KAC and proliferating KAC (KAC_prolif).
- KAC and Inflammatory signature scores across alveolar cell clusters. *p* values were calculated using Kruskal-Wallis tests.
- Proportions of different cell subsets in neighborhoods of each of the indicated epithelial cell types.
- The number of *IL1B*⁺ macrophages (C15) in the neighborhoods of different epithelial cell types. The overall *p* value was calculated using the Kruskal-Wallis test, and the *p* values between groups were calculated using unpaired Student's *t* tests.
- Representative images from paired lesions of P24 and showing visual analysis of the local niche of RPII near AAH and LUAD as well as of AAH cells themselves (left). Zoom out panels show (left, top to bottom): H&E-stained tissues, pathology annotations, and major cell types. Zoomed in images of regions 1, 2, and 3 and visualization of *IL1B* and *IL1R1* transcripts (right).
- Summary of composition of 188 TMA cores indicating presence of different lesions, as well as LUAD histological subtypes, if present.
- UMAP embeddings of 593,334 cells from 8 TMAs analyzed by Xenium.
- A representative core with an acinar LUAD and showing post Xenium H&E staining and corresponding visualization of major lineage cell types. Cell type color legend is the same as in (H).
- Neighborhood structure of tumor cells in different lesions.
- Distribution of *IL1B*⁺ macrophages (C4) in the neighborhoods of tumor cells in different lesions. *p* values between groups were calculated using unpaired Student's *t* tests.
- A representative core with an AAH and showing post Xenium H&E staining of the core and corresponding visualization of *IL1B* transcripts expressed in *IL1B*⁺ macrophages (C4) and *IL1R1* transcripts expressed in atypical hyperplastic cells of an AAH.

See also Figures S6 and S7 and Tables S1, S3, and S4.

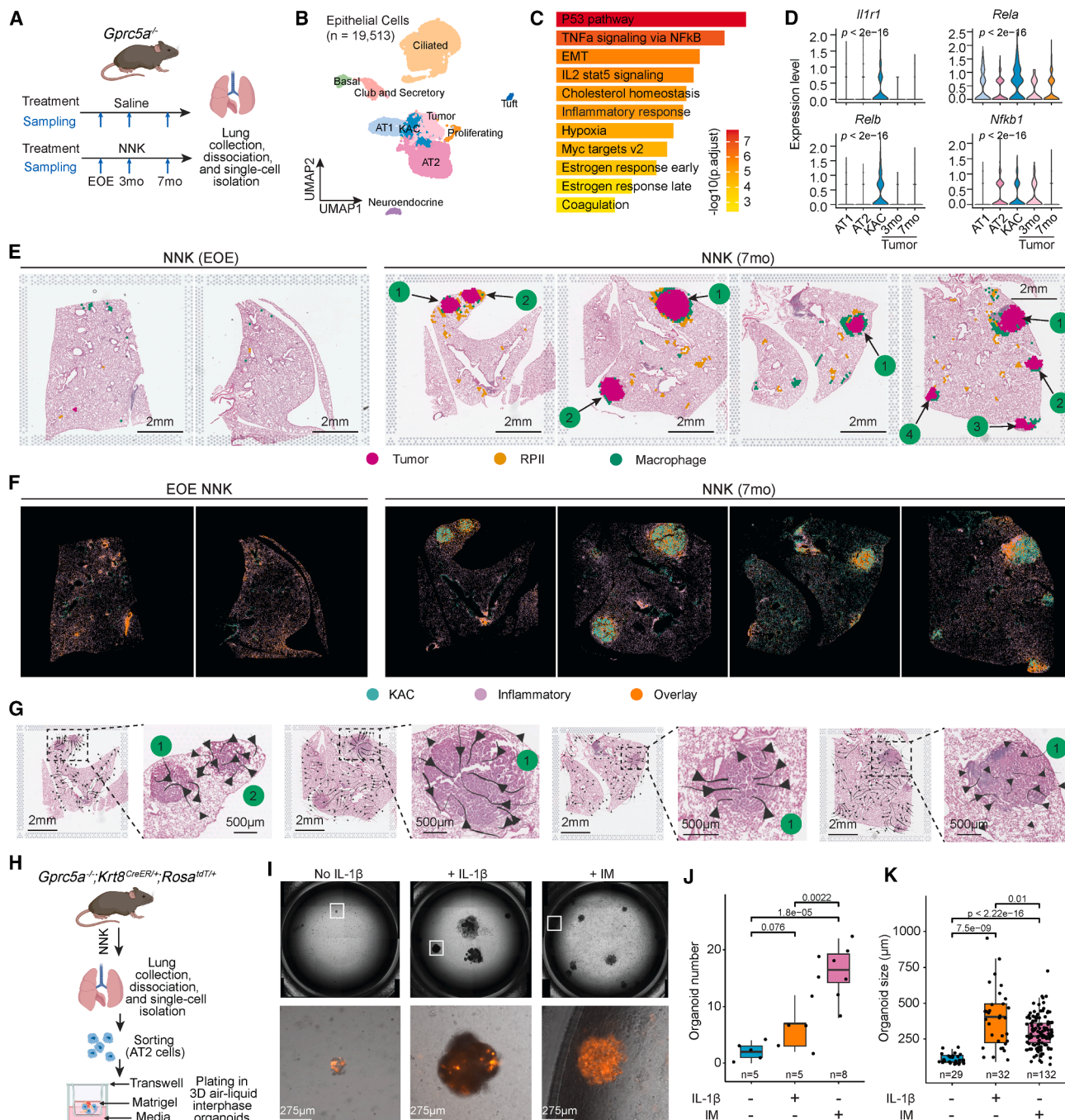


Figure 6. Proinflammatory signaling in macrophage-rich niches drives KAC growth early in tobacco carcinogen-mediated LUAD pathogenesis

(A) Schematic diagram showing tobacco carcinogen (NNK) or control saline exposure in *Gprc5a*^{-/-} mice and analysis of lung single cells at end of exposure (EOE) and at 3 and 7 months post-exposure.

(B) UMAP clustering of epithelial cells.

(C) Pathways enrichment results showing the top 11 pathways increased in KACs relative to other alveolar and tumor subsets.

(D) *Il1r1*, *Rela*, *Relb* and *Nfkb1* expression levels in AT1, AT2, KACs, and tumor cells analyzed by scRNA-seq. *p* values were calculated using Kruskal-Wallis tests.

(E) H&E-stained mouse lung tissues analyzed by Visium ST at EOE (n = 2 samples) and at 7 months post-NNK (n = 4 samples) and showing pathology annotation of tumor, RPII and macrophage spots. Arrows point to lesions in tumor-bearing lung tissues.

(F) Super-pixel resolution spatial mapping of KAC and Inflammatory signatures, as well as their overlay, in the same samples shown in (E).

(G) Analysis (globally and locally) of *Il1b-Il1r1* signaling in tumors and their macrophage-enriched neighborhoods.

(H) Schematic diagram showing strategy for AT2 cell isolation and generation of AT2-derived organoids from NNK-treated KRT8 lineage reporter mice (*Gprc5a*^{-/-}; *Krt8*^{CreER}; *Rosa*^{tdT/+}).

(legend continued on next page)

mice significantly increased the number and size of organoids compared to the untreated control group (Figures 6I–6K). These findings highlight the role of IL-1 β -IL1R1 axis and proinflammatory cues in driving oncogenesis of KACs during early LUAD pathogenesis.

Therapeutic targeting of IL-1 β suppresses pro-tumor inflammation and inhibits LUAD development in conjunction with immune checkpoint blockade

To interrogate whether intercepting IL-1 β signaling could decrease alveolar transformation and, thus, impede LUAD pathogenesis we treated NNK-exposed *Gprc5a*^{-/-} mice with either anti-IL-1 β or control IgG (Figure 7A). To evaluate the effects of IL-1 β blockade alongside conventional immune checkpoint inhibition, we studied additional groups of mice receiving either anti-PD-1 alone or in combination with anti-IL-1 β (Figure 7A). IL-1 β and/or PD-1 blockade were implemented in two experimental settings depending on treatment window, both covering precancerous phases: (1) prevention of early lesion development with antibodies administered immediately after NNK exposure for 3 months, and (2) interception of early lesion progression with treatment beginning at 3 months post-NNK exposure and lasting for 4 months (Figure 7A). Comparative analyses revealed that anti-IL-1 β was more efficacious in reducing the volume of lung adenomas and adenocarcinomas compared to anti-PD-1 monotherapy, in both early and late time points (Figure 7B). Notably, combined IL-1 β and PD-1 blockade led to significantly enhanced tumor suppression relative to monotherapy especially PD-1 blockade and particularly in the interception phase (Figure 7B). We further found in a syngeneic mouse model that antibody-mediated neutralization of IL-1 β failed to reduce the growth of mouse LUAD cells *in vivo* (Figure S8J), further emphasizing that targeting anti-IL-1 β is relevant in early phases in development of LUAD.

Anti-IL-1 β significantly diminished levels of pro-tumor inflammatory cytokines, including CCL3, IL33, and IL17A, in bronchoalveolar lavage fluid (BALF) at 7 months post-NNK (Figure 7C). Combinatorial therapy in lungs analyzed at 7 months showed the most pronounced effects, whereby we also noted significant reductions in IL23, IL21, and CXCL1 (Figures 7C and S8K). ST analysis further demonstrated reduced lesion size and decreased macrophage infiltration in both lesion-adjacent peritumoral regions, in anti-IL-1 β -treated mice compared to controls (Figure 7D). scRNA-seq of epithelial cells enriched from the lungs of treated mice identified distinct clusters, including club, ciliated, basal, neuroendocrine, AT1, AT2, and intermediate alveolar cell populations (e.g., KACs), as well as tumor cells (Figures 7E and S8L–S8M; Table S5). At 3 months post-NNK, KAC fractions were only statistically significantly reduced in the combinatorial treatment group, whereas at 7 months, reduced KAC abundance was more pronounced across all treatment groups (Figures 7F–7G). Reductions in tumor cells were evident across all treatment groups, with the combinatorial approach showing the greatest

impact (Figure S8N). Intriguingly, early preventative exposure to anti-IL-1 β treatment also led to reduced alveolar macrophage fractions, an effect that was extended to all treatment groups in the late treatment window (Figure S8O), whereas cytotoxic CD8⁺ T cell fractions increased with the combination treatment, as did plasma cells albeit only at 7 months (Figures S8P and S8Q).

Immunofluorescence analyses of lung TMAs constructed from a subset of mice treated with the 7 months endpoint highlighted enrichment of KRT8⁺ epithelial cells as well as macrophage- and cytotoxic CD8⁺ T cell-rich niches around tumors (Figure 7H; Table S6). Anti-IL-1 β marginally decreased macrophage fractions relative to controls, while anti-PD-1 alone elevated macrophage abundance (Figure 7H). Strikingly, combinatorial therapy not only counteracted PD-1-associated macrophage enrichment but also produced the highest infiltration of cytotoxic CD8⁺ T cells compared to monotherapy, which were both in line with macrophage and CD8⁺ T cell fractions analyzed by scRNA-seq (Figures 7H, 7I, S8O, and S8P). These findings demonstrate that targeting inflammation by inhibiting IL-1 β signaling during precancerous stages likely disrupts alveolar cell transformation into tumor progenitors and effectively reduces LUAD pathogenesis.

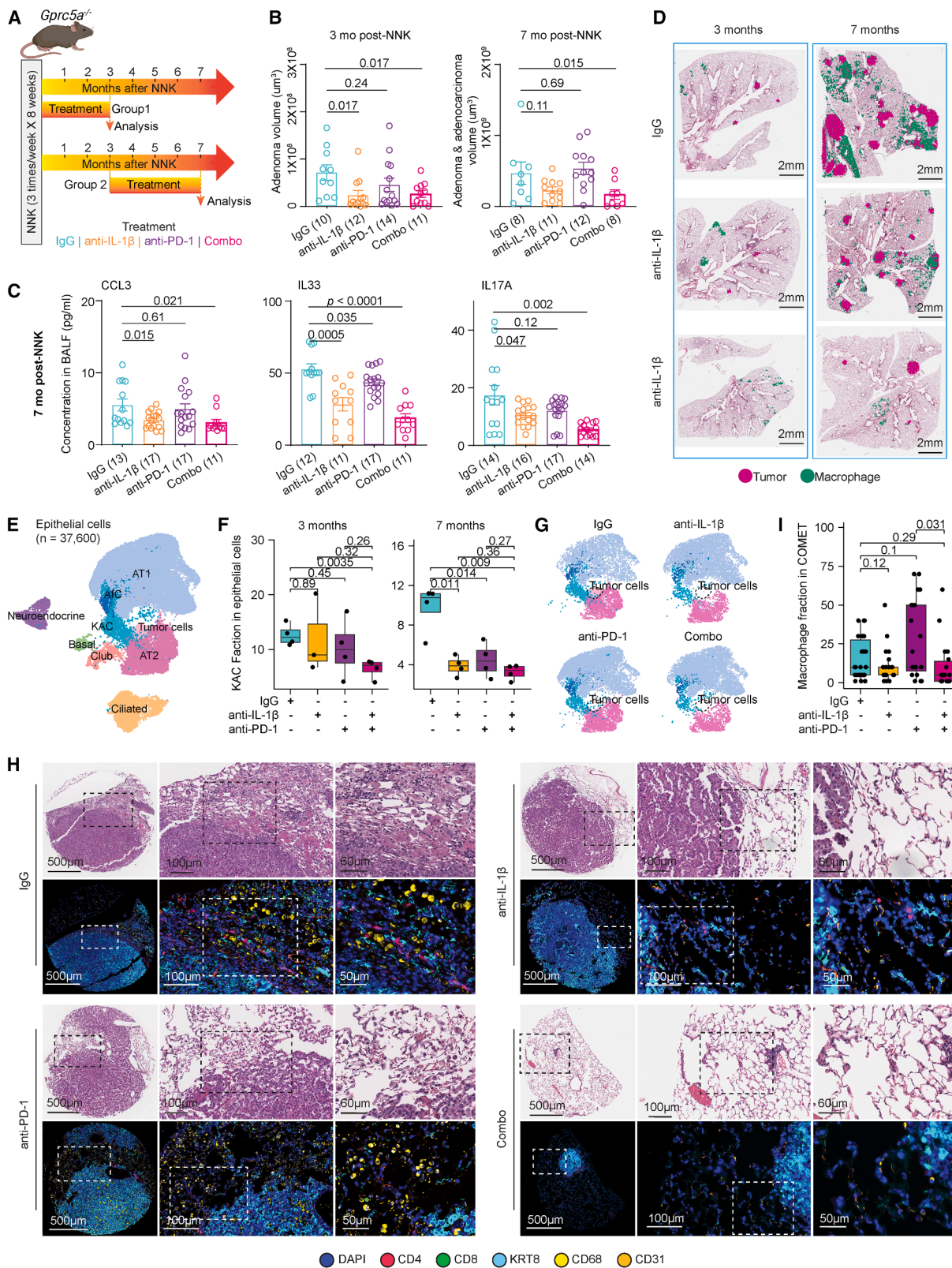
DISCUSSION

In our present study using multimodal single-cell and spatial-omics analysis of lung precursor and invasive lesions coupled with functional validation in experimental and carcinogenesis models of LUAD, we find that epithelial-proinflammatory niches likely drive LUAD initiation and are pertinent for disease interception.

Our spatial genotype-phenotype correlation analyses found that KRT8-high KACs (pathologically denoted as RPII), which we previously reported to act as likely cells-of-origin for LUAD in mice,¹⁹ are more ancestral than invasive and even precursor lung epithelial cells. Alveolar progenitor niches spatially mapped to clones that were mostly shared with precursor lesions, suggesting the role of this niche in tumor initiation and which may manifest in *normal-appearing* regions as *field effects*.²⁸ A precancerous lesion was suggested to be a “morphologically altered tissue in which cancer is more likely to occur than in its apparently normal counterpart”.⁷ Yet, in a recent compelling review, the Precancer Think Tank Team suggested that the term “precancer” is used to broadly describe any type of condition, state or lesion that may or may not be histologically characterized.²⁹ In this theme, our multimodal analysis showed that alveolar progenitors (i.e., RPII), despite exhibiting subtle morphological changes, possess robust characteristics and perhaps denote the earliest lesions in LUAD pathogenesis.

A major finding was that alveolar progenitors commonly resided in niches enriched with proinflammatory macrophages. While earlier work showed that coculture with macrophages increased growth of AT2 cells from mouse models of pollution exposure,³⁰ an interesting observation in our study was that

(I) Representative images of transwell inserts with growing AT2-derived KRT8 high lineage-labeled organoids and their varying numbers and sizes when cultured in control medium, with recombinant IL-1 β , or cocultured with interstitial macrophages (IM). Organoids are imaged at low (top) and high (bottom) magnification. (J and K) The number (J) and size (K) of organoids in the three indicated conditions. Each dot represents one well. Box middle lines, median; box limits, upper and lower quartiles; box whiskers, 1.5 $\times$ the interquartile range. *p* values were calculated using unpaired Students *t* tests. See also Figure S8 and Table S3.



(legend on next page)

alveolar progenitors displayed much higher expression of IL1R1 compared to other alveolar subsets suggesting that they are more receptive of proinflammatory signaling and thus partaking in tumor initiation than AT2 cells. It is important to note that previous studies using models of acute lung injury found that proinflammatory signaling was important for generation and/or maintenance of KRT8 high alveolar intermediates in the lung in the context of pulmonary fibrosis.^{31,32} It is not clear whether epithelial-proinflammatory niches are a fork in the road between resolution of lung injury or tumor initiation.³³ Another main finding of ours was that epithelial-proinflammatory niches are likely to be more operative in earlier (KAC or precancerous) phases in LUAD development. It is important to note that pathological conditions of the lung such as chronic obstructive pulmonary disease (COPD) that are causally linked to heightened inflammation are associated with increased risk of LUAD.³⁴ Notably, immune profiles in the COPD stage were shown to be quite distinct from advanced stage of LUAD, and they were enriched with proinflammatory immune cell subsets.³⁴ While our findings are concordant with prior observations on inflammation as a driver of KRAS-mediated oncogenesis,²⁵ they also provide new insights on proinflammatory signaling that is pertinent to precancerous stages in LUAD development. They also raise the intriguing supposition that epithelial-proinflammatory niches represent tumor-initiating *spatial field effects*, synonymous with genomic and epithelial-intrinsic field effects such as mutant KRAS we and others described previously.³⁵

We also provide clinical value for the association between the epithelial-proinflammatory niche and early phases of LUAD development. It is important to note that the CANTOS trial designed to test lipid lowering-independent strategies for treatment of atherosclerosis previously showed that the IL-1 β neutralizing antibody canakinumab reduced incident lung cancer.³⁶ In sharp contrast, clinical trials showed that addition of canakinumab to first-line pembrolizumab and chemotherapy, to chemotherapy after progression on platinum-based doublet chemotherapy and immunotherapy, or after surgery and adjuvant cisplatin-based chemotherapy did not provide clinical

benefits in patients with NSCLC.^{37–39} It is indeed plausible that targeting inflammation, perhaps even when with immune checkpoint blockade, is more clinically effective in precancerous stages preceding LUAD development.

Limitations of the study

Despite our broad conclusions on patient-level heterogeneity in clonal relationships between lung precursor lesions and LUADs being concordant with previous reports,^{11,12,20} our analysis of CNAs based on Visium ST data have limitations. Confident assessment of the presence of specific clones from Visium-based CNA inference is limited by substantial cell-type heterogeneity in spots (e.g., admixture of tumor and non-tumoral cells). This can lead to noise or clonal misassignments which may impact construction of accurate phylogenetic trees. Future studies employing spatial (spot- or single-cell-level) DNA sequencing-based approaches to profile SNVs and CNAs will help validate our major conclusions as well as better reconstruct phylogenies and elucidate clonal evolution of lung precancer and LUAD in space. Sampling bias of precursors and LUADs cannot be neglected as both are often detected synchronously in resected specimens from patients, which could influence the interpretation of evolutionary dynamics of LUAD. Future studies incorporating lineage tracing and spatial perturbation *in vivo* could dissect the causal role of epithelial-proinflammatory in LUAD initiation.

In conclusion, our multimodal spatial-omics mapping identified epithelial-proinflammatory niches that are likely tumor initiating and are prevalent in precancerous but not invasive phases of LUAD development. We find that these niches not only represent the co-evolution of alveolar progenitors and proinflammatory immune subsets but also provide value for early interception strategies in LUAD.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to the lead contact, Humam Kadara (hkadara@mdanderson.org).

Figure 7. Blocking IL-1 β in combination with immune checkpoint blockade decreases KACs, pro-tumor inflammation, and LUAD pathogenesis

- (A) Schematic diagram showing two windows for treatment of *Gprc5a*^{-/-} mice following exposure to tobacco carcinogen (NNK).
 (B) Histopathological analysis of lung adenoma and adenocarcinoma volumes at 3 and 7 months post-NNK and in mice receiving single agent or combination antibody treatments. The number of replicates for each group is indicated between parentheses. Bars represent mean values, and error bars indicate standard deviation. *p* values were calculated using unpaired Student's *t* tests.
 (C) Cytokine concentration in bronchoalveolar lavage fluid (BALF) of mice analyzed at the 7 months post-NNK time point. The number of replicates for each group is indicated between parentheses. Bars represent mean values, and error bars indicate standard deviation. *p* values were calculated using unpaired Student's *t* tests.
 (D) ST analysis of mouse lungs at 3 and 7 months post-end of NNK, and after receiving control IgG or anti-IL-1 β shown in (A). Tumor spots and macrophage infiltration are annotated.
 (E) UMAP clustering of enriched epithelial cells from lungs of mice shown in (A).
 (F) Fraction of KACs among epithelial cells across the four different antibody treatment groups and at 3 and 7 months post-NNK (*n* = 4 for each group, except for *n* = 3 for anti-IL-1 β in the 3-month group). Box middle lines, median; box limits, upper and lower quartiles; box whiskers, 1.5 $\times$ the interquartile range. *p* values were calculated using unpaired Student's *t* tests.
 (G) UMAP clustering showing changes in fractions of AT1, AT2, AIC, KAC, and tumor cell subsets among all epithelial cells in (E).
 (H) Representative images of 4 TMA cores from each of the four treatment groups analyzed by seq-IF at increasing magnification, with the indicated select markers.
 (I) Quantification of macrophage fractions (among all cells) by treatment group. Results include data from all 96 cores of the 4 TMAs (*n*, control IgG = 30, anti-IL-1 β = 23, anti-PD-1 = 23, Combo = 18) analyzed by seq-IF. Box middle lines, median; box limits, upper and lower quartiles; box whiskers, 1.5 $\times$ the interquartile range. *p* values were calculated using unpaired Student's *t* tests.

See also Figure S8 and Tables S5 and S6.

Materials availability

The study did not generate any new unique reagents.

Data and code availability

The codes and processed data used to produce the results of this paper are available at GitHub (https://github.com/FudianPeng/LungPCA_Code.git) and Zenodo (<https://doi.org/10.5281/zenodo.17172149>). Mouse data are available in the Gene Expression Omnibus (GEO) (datasets GSE222901, GSE300288, GSE300293). Processed human data are available in GEO (datasets GSE307534, GSE308103, GSE307529) and Zenodo (<https://doi.org/10.5281/zenodo.15670280>). Any additional information required to re-analyze the data reported in this paper is available from the [lead contact](#) upon request.

ACKNOWLEDGMENTS

We would like to acknowledge Yanshuo Chu, Guangchun Han, and Ruiping Wang who provided valuable suggestions for analysis and participated in discussions. We also would like to acknowledge Vivian Ha and Rachel Yang for their assistance. This work was supported in part by the National Cancer Institute (NCI) grants R01CA272863 (to H.K. and S.J.M.), U01CA264583 (to H.K. and L.W.), R01CA266280 (to L.W.), U01CA294518 (to L.W.), U24CA274274 (to L.W.), R01CA225977 (to S.J.M.), and the Cancer Center Support Grant P30CA016672, Cancer Prevention Research Institute of Texas (CPRIT) grant RP220101 (to H.K. and L.W.), American Lung Association (ALA)-LUNGevity (to H.K., S.M., S.D. and A.E.S.), ALA Lung Cancer Discovery Award grant LCD-821433 (to S.J.M.), Johnson and Johnson (to H.K.), and the MD Anderson Cancer Center Office of Strategic Research Priorities (to H.K.). H.K., L.W., and T.C. are Andrew Sabin Family Foundation fellows. H.K. was also supported by the Elza A. and Ina Shackelford Freeman Endowed Professorship in Lung Cancer. L.W. was also supported in part by research funding provided by the James P. Allison Institute, the Institute for Data Science in Oncology, Break Through Cancer, and the University of Texas MD Anderson Cancer Center. L.W. is an Associate Member of the James P. Allison Institute. A.S. and M.C. were supported by the NCI training grant T32217789. J.B. was supported in part by the NCI Research Specialist award R50CA243707. M.S. was supported by UTHealth Houston Innovation for Cancer Prevention Research Training Program Postdoctoral Fellowship (CPRIT grant RP210042). MD Anderson's Advanced Technology Genomics Core (ATGC), Institutional Tissue Bank, Veterinary Medicine's Pathology core and the Flow Cytometry and Cellular Imaging Core Facility were supported by the Cancer Center Support Grant P30CA016672 and the ATGC core was also supported by the National Institutes of Health instrument grant 1S10OD024977.

AUTHOR CONTRIBUTIONS

A.S., T.Z., J. Feng., Z.R., K.K., S.S., N.J., T.H., S.K., K.T., T. I., T.Y., K.S., and N.K. were involved in human sample collection, tissue processing and clinical data curation. A.S., W.T., T.Z., and J. Feng performed sample processing for ST, Xenium, scRNA-seq, snRNA-seq, as well as WES. A.S., W.T., M.C., A.K., N. Karimi, M.S., Z.R., Yuejiang L., and S.J.M. performed analyses of animal models. L.I.G.B., A.G.S., L.M.S., N.K., I.I.W. and J. Fujimoto performed detailed histopathological annotation and integration of spatial data. A.S. and A.B. performed seq-IF experiments. A.S., M.C., and T.N. developed workflows for KRT8 reporter tracing in cells derived from mouse lung tissues. F.P., A.S., Y.D., S.Y., L.W., and H.K. performed the formal analysis and data visualization of ST, Xenium, scRNA-seq, snRNA-seq, and WES datasets. F.P., Y.D., S.Y., G.P., J.L., Yunhe Liu., Yang Liu., J. J., K. Y., K.S.C., and L.W. developed analysis workflows and bioinformatics approaches used in this study. S.J.M., J.C., M.L., J.B., K.A., L.W., and H.K. developed and provided support for various workflows used in this study. S.M., S.D., T.C., A.E.S., S.S., N.J., T.H., S.K., K.T., T.I., T.Y., K.S., and J.M. provided data interpretation and resources pertaining to lung cancer cohorts. L.W. and H.K. oversaw the project administration, supervised, and acquired funding resources for the study. H.K. provided strategic oversight and conceived the study. F.P., A.S., Y.D., L.W., and H.K. drafted the manuscript. All authors approved the manuscript.

DECLARATION OF INTERESTS

H.K. reports research funding from Johnson and Johnson. L.W. reports outside the scope of the submitted work honoraria for service on the Scientific Advisory Board of SELLAS Life Sciences. T.C. reports outside the scope of the submitted work speaker fees/honoraria (including travel/meeting expenses) from ASCO Post, AstraZeneca, Bio Ascend, Bristol Myers Squibb, Clinical Care Options, IDEology Health, Medical Educator Consortium, Medscape, OncLive, PEAK Medicals, PeerView, Physicians' Education Resource, Targeted Oncology; advisory role/consulting fees (including travel/meeting expenses) from AstraZeneca, Bristol Myers Squibb, Daiichi Sankyo, Genentech, Merck, oNko-innate, Pfizer, and RAPT Therapeutics; institutional research funding from AstraZeneca and Bristol Myers Squibb. S.M.D. served on an advisory board and holds stock in Early Diagnostics and LungLife AI; and received research support from Johnson & Johnson. A.E.S. is an employee of Johnson and Johnson. I.I.W. reports grants and personal fees from Genentech/Roche, grants and personal fees from Bayer, grants and personal fees from Bristol-Myers Squibb, grants and personal fees from AstraZeneca, grants and personal fees from Pfizer, grants and personal fees from HTG Molecular, personal fees from Asuragen, grants and personal fees from Merck, grants and personal fees from GlaxoSmithKline, grants and personal fees from Guardant Health, personal fees from Flame, grants and personal fees from Novartis, grants and personal fees from Sanofi, personal fees from Daiichi Sankyo, grants and personal fees from Amgen, personal fees from Oncocyte, personal fees from MSD, personal fees from Platform Health, grants from Adaptive, grants from Adaptimmune, grants from EMD Serono, grants from Takeda, grants from Karus, grants from Johnson & Johnson, grants from 4D, from Iovance and from Akoya, outside the submitted work.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **EXPERIMENTAL MODELS AND STUDY PARTICIPANT DETAILS**
 - Patient samples
 - Animal models
 - Organoids
 - Cell lines
- **METHOD DETAILS**
 - Histopathological processing of tissues
 - Visium CytAssist ST profiling
 - Visium CytAssist ST data analysis
 - Analysis of spatial copy number alterations
 - Spatial resolution enhancement using iStar
 - MP identification from visium ST data
 - Analysis of ligand-receptor signaling
 - snRNA-seq data generation
 - snRNA-seq data analysis
 - WES data processing and analysis
 - Xenium data generation
 - Xenium data analysis
 - KAC, NF-κB and interferon signature scores
 - Mouse scRNA-seq data generation
 - Mouse scRNA-seq data analysis
 - Visium ST profiling of mouse lung tissues
 - Visium ST analysis of mouse lung tissues
 - Carcinogen and antibody treatments in mice
 - Histopathology analysis of mouse lung tissues
 - Immunofluorescence analysis of mouse tissues
 - Analysis of AT2-derived KRT8 high organoids
 - IL-1β blockade in syngeneic model of LUAD
 - Cytokine and chemokine profiling
 - Sequential immunofluorescence (seq-IF) analysis
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.ccell.2025.10.004>.

Received: April 20, 2025

Revised: August 27, 2025

Accepted: October 15, 2025

Published: November 6, 2025

REFERENCES

1. Thai, A.A., Solomon, B.J., Sequist, L.V., Gainor, J.F., and Heist, R.S. (2021). Lung cancer. *Lancet* 398, 535–554. [https://doi.org/10.1016/S0140-6736\(21\)00312-3](https://doi.org/10.1016/S0140-6736(21)00312-3).
2. Saab, S., Zalzal, H., Rahal, Z., Khalifeh, Y., Sinjab, A., and Kadara, H. (2020). Insights Into Lung Cancer Immune-Based Biology, Prevention, and Treatment. *Front. Immunol.* 11, 159. <https://doi.org/10.3389/fimmu.2020.00159>.
3. Mazzilli, S.A., Rahal, Z., Rouhani, M.J., Janes, S.M., Kadara, H., Dubinett, S.M., and Spira, A.E. (2025). Translating premalignant biology to accelerate non-small-cell lung cancer interception. *Nat. Rev. Cancer* 25, 379–392. <https://doi.org/10.1038/s41568-025-00791-1>.
4. Travis, W.D., Brambilla, E., Noguchi, M., Nicholson, A.G., Geisinger, K.R., Yatabe, Y., Beer, D.G., Powell, C.A., Riely, G.J., Van Schil, P.E., et al. (2011). International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *J. Thorac. Oncol.* 6, 244–285. <https://doi.org/10.1097/JTO.0b013e318206a22101243894-201102000-00004>.
5. Jakubek, Y., Lang, W., Vattathil, S., Garcia, M., Xu, L., Huang, L., Yoo, S. Y., Shen, L., Lu, W., Chow, C.W., et al. (2016). Genomic Landscape Established by Allelic Imbalance in the Cancerization Field of a Normal Appearing Airway. *Cancer Res.* 76, 3676–3683. <https://doi.org/10.1158/0008-5472.CAN-15-3064>.
6. Kadara, H., Fujimoto, J., Yoo, S.Y., Maki, Y., Gower, A.C., Kabbout, M., Garcia, M.M., Chow, C.W., Chu, Z., Mendoza, G., et al. (2014). Transcriptomic architecture of the adjacent airway field cancerization in non-small cell lung cancer. *J. Natl. Cancer Inst.* 106, dju004. <https://doi.org/10.1093/jnci/dju004>.
7. Kadara, H., Scheet, P., Wistuba, I.I., and Spira, A.E. (2016). Early Events in the Molecular Pathogenesis of Lung Cancer. *Cancer Prev. Res.* 9, 518–527. <https://doi.org/10.1158/1940-6207.CAPR-15-0400>.
8. Kadara, H., Sivakumar, S., Jakubek, Y., San Lucas, F.A., Lang, W., McDowell, T., Weber, Z., Behrens, C., Davies, G.E., Kallhor, N., et al. (2019). Driver Mutations in Normal Airway Epithelium Elucidate Spatiotemporal Resolution of Lung Cancer. *Am. J. Respir. Crit. Care Med.* 200, 742–750. <https://doi.org/10.1164/rccm.201806-1178OC>.
9. Dejima, H., Hu, X., Chen, R., Zhang, J., Fujimoto, J., Parra, E.R., Haymaker, C., Hubert, S.M., Duose, D., Solis, L.M., et al. (2021). Immune evolution from preneoplasia to invasive lung adenocarcinomas and underlying molecular features. *Nat. Commun.* 12, 2722. <https://doi.org/10.1038/s41467-021-22890-x>.
10. Krysan, K., Tran, L.M., Grimes, B.S., Fishbein, G.A., Seki, A., Gardner, B. K., Walser, T.C., Salehi-Rad, R., Yanagawa, J., Lee, J.M., et al. (2019). The Immune Contexture Associates with the Genomic Landscape in Lung Adenomatous Premalignancy. *Cancer Res.* 79, 5022–5033. <https://doi.org/10.1158/0008-5472.CAN-19-0153>.
11. Sivakumar, S., San Lucas, F.A., Jakubek, Y.A., McDowell, T.L., Lang, W., Kallsen, N., Peyton, S., Davies, G.E., Fukuoka, J., Yatabe, Y., et al. (2019). Genomic landscape of allelic imbalance in premalignant atypical adenomatous hyperplasias of the lung. *EBioMedicine* 42, 296–303. <https://doi.org/10.1016/j.ebiom.2019.03.020>.
12. Sivakumar, S., Lucas, F.A.S., McDowell, T.L., Lang, W., Xu, L., Fujimoto, J., Zhang, J., Futreal, P.A., Fukuoka, J., Yatabe, Y., et al. (2017). Genomic Landscape of Atypical Adenomatous Hyperplasia Reveals Divergent Modes to Lung Adenocarcinoma. *Cancer Res.* 77, 6119–6130. <https://doi.org/10.1158/0008-5472.CAN-17-1605>.
13. Kadara, H., Tran, L.M., Liu, B., Vachani, A., Li, S., Sinjab, A., Zhou, X.J., Dubinett, S.M., and Krysan, K. (2021). Early Diagnosis and Screening for Lung Cancer. *Cold Spring Harb. Perspect. Med.* 11, a037994. <https://doi.org/10.1101/cshperspect.a037994>.
14. Sinjab, A., Han, G., Wang, L., and Kadara, H. (2020). Field Carcinogenesis in Cancer Evolution: What the Cell Is Going On? *Cancer Res.* 80, 4888–4891. <https://doi.org/10.1158/0008-5472.CAN-20-1956>.
15. Moffitt, J.R., Lundberg, E., and Heyn, H. (2022). The emerging landscape of spatial profiling technologies. *Nat. Rev. Genet.* 23, 741–759. <https://doi.org/10.1038/s41576-022-00515-3>.
16. Zhang, D., Schroeder, A., Yan, H., Yang, H., Hu, J., Lee, M.Y.Y., Cho, K.S., Susztak, K., Xu, G.X., Feldman, M.D., et al. (2024). Inferring super-resolution tissue architecture by integrating spatial transcriptomics with histology. *Nat. Biotechnol.* 42, 1372–1377. <https://doi.org/10.1038/s41587-023-02019-9>.
17. Erickson, A., He, M., Berglund, E., Marklund, M., Mirzazadeh, R., Schultz, N., Kvastad, L., Andersson, A., Bergenstr hle, L., Bergenstr hle, J., et al. (2022). Spatially resolved clonal copy number alterations in benign and malignant tissue. *Nature* 608, 360–367. <https://doi.org/10.1038/s41586-022-05023-2>.
18. Gulati, G.S., Sikandar, S.S., Wesche, D.J., Manjunath, A., Bharadwaj, A., Berger, M.J., Ilagan, F., Kuo, A.H., Hsieh, R.W., Cai, S., et al. (2020). Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* 367, 405–411. <https://doi.org/10.1126/science.aax0249>.
19. Han, G., Sinjab, A., Rahal, Z., Lynch, A.M., Trekitkammongkol, W., Liu, Y., Serrano, A.G., Feng, J., Liang, K., Khan, K., et al. (2024). An atlas of epithelial cell states and plasticity in lung adenocarcinoma. *Nature* 627, 656–663. <https://doi.org/10.1038/s41586-024-07113-9>.
20. Hu, X., Fujimoto, J., Ying, L., Fukuoka, J., Ashizawa, K., Sun, W., Reuben, A., Chow, C.W., McGranahan, N., Chen, R., et al. (2019). Multi-region exome sequencing reveals genomic evolution from preneoplasia to lung adenocarcinoma. *Nat. Commun.* 10, 2978. <https://doi.org/10.1038/s41467-019-10877-8>.
21. Barkley, D., Moncada, R., Pour, M., Liberman, D.A., Dryg, I., Werba, G., Wang, W., Baron, M., Rao, A., Xia, B., et al. (2022). Cancer cell states recur across tumor types and form specific interactions with the tumor microenvironment. *Nat. Genet.* 54, 1192–1201. <https://doi.org/10.1038/s41588-022-01141-9>.
22. Gavish, A., Tyler, M., Greenwald, A.C., Hoefflin, R., Simkin, D., Tschernichovsky, R., Galili Darnell, N., Somech, E., Barbolin, C., Antman, T., et al. (2023). Hallmarks of transcriptional intratumour heterogeneity across a thousand tumours. *Nature* 618, 598–606. <https://doi.org/10.1038/s41586-023-06130-4>.
23. Liu, T., Zhang, L., Joo, D., and Sun, S.C. (2017). NF-κB signaling in inflammation. *Signal Transduct. Target. Ther.* 2, 17023. <https://doi.org/10.1038/sigtrans.2017.23>.
24. Hamarsheh, S., Gro , O., Brummer, T., and Zeiser, R. (2020). Immune modulatory effects of oncogenic KRAS in cancer. *Nat. Commun.* 11, 5439. <https://doi.org/10.1038/s41467-020-19288-6>.
25. Kitajima, S., Thummalapalli, R., and Barbie, D.A. (2016). Inflammation as a driver and vulnerability of KRAS mediated oncogenesis. *Semin. Cell Dev. Biol.* 58, 127–135. <https://doi.org/10.1016/j.semdb.2016.06.009>.
26. Yu, H., Pardoll, D., and Jove, R. (2009). STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat. Rev. Cancer* 9, 798–809. <https://doi.org/10.1038/nrc2734>.
27. Caetano, M.S., Hassane, M., Van, H.T., Bugarin, E., Cumpian, A.M., McDowell, C.L., Cavazos, C.G., Zhang, H., Deng, S., Diao, L., et al. (2018). Sex specific function of epithelial STAT3 signaling in pathogenesis of K-ras mutant lung cancer. *Nat. Commun.* 9, 4589. <https://doi.org/10.1038/s41467-018-07042-y>.
28. Rahal, Z., Sinjab, A., Wistuba, I.I., and Kadara, H. (2022). Game of clones: Battles in the field of carcinogenesis. *Pharmacol. Ther.* 237, 108251. <https://doi.org/10.1016/j.pharmthera.2022.108251>.

29. Faupel-Badger, J., Kohaar, I., Bahl, M., Chan, A.T., Campbell, J.D., Ding, L., De Marzo, A.M., Maitra, A., Merrick, D.T., Hawk, E.T., et al. (2024). Defining precancer: a grand challenge for the cancer community. *Nat. Rev. Cancer* 24, 792–809. <https://doi.org/10.1038/s41568-024-00744-0>.
30. Hill, W., Lim, E.L., Weeden, C.E., Lee, C., Augustine, M., Chen, K., Kuan, F.-C., Marongiu, F., Evans, E.J., Moore, D.A., et al. (2023). Lung adenocarcinoma promotion by air pollutants. *Nature* 616, 159–167. <https://doi.org/10.1038/s41586-023-05874-3>.
31. Choi, J., Park, J.E., Tsagkogeorga, G., Yanagita, M., Koo, B.K., Han, N., and Lee, J.H. (2020). Inflammatory Signals Induce AT2 Cell-Derived Damage-Associated Transient Progenitors that Mediate Alveolar Regeneration. *Cell Stem Cell* 27, 366–382.e7. <https://doi.org/10.1016/j.stem.2020.06.020>.
32. Strunz, M., Simon, L.M., Ansari, M., Kathiriyai, J.J., Angelidis, I., Mayr, C. H., Tsidiridis, G., Lange, M., Mattner, L.F., Yee, M., et al. (2020). Alveolar regeneration through a Krt8+ transitional stem cell state that persists in human lung fibrosis. *Nat. Commun.* 11, 3559. <https://doi.org/10.1038/s41467-020-17358-3>.
33. Gosens, R., Giangreco, A., Sahai, E., and Chambers, R.C. (2017). Mechanistic overlap between chronic lung injury and cancer: ERS Lung Science Conference 2017 report. *Eur. Respir. Rev.* 26, 170060. <https://doi.org/10.1183/16000617.0060-2017>.
34. King, P.T. (2015). Inflammation in chronic obstructive pulmonary disease and its role in cardiovascular disease and lung cancer. *Clin. Transl. Med.* 4, 68. <https://doi.org/10.1186/s40169-015-0068-z>.
35. Kadara, H., and Wistuba, I.I. (2012). Field cancerization in non-small cell lung cancer: implications in disease pathogenesis. *Proc. Am. Thorac. Soc.* 9, 38–42. <https://doi.org/10.1513/pats.201201-004MS>.
36. Ridker, P.M., MacFadyen, J.G., Thuren, T., Everett, B.M., Libby, P., and Glynn, R.J.; CANTOS Trial Group (2017). Effect of interleukin-1 β inhibition with canakinumab on incident lung cancer in patients with atherosclerosis: exploratory results from a randomised, double-blind, placebo-controlled trial. *Lancet* 390, 1833–1842. [https://doi.org/10.1016/S0140-6736\(17\)32247-X](https://doi.org/10.1016/S0140-6736(17)32247-X).
37. Garon, E.B., Lu, S., Goto, Y., De Marchi, P., Paz-Ares, L., Spigel, D.R., Thomas, M., Yang, J.C.H., Ardizzoni, A., Barlesi, F., et al. (2024). Canakinumab as Adjuvant Therapy in Patients With Completely Resected Non-Small-Cell Lung Cancer: Results From the CANOPY-A Double-Blind, Randomized Clinical Trial. *J. Clin. Oncol.* 42, 180–191. <https://doi.org/10.1200/jco.23.00910>.
38. Paz-Ares, L., Goto, Y., Wan-Teck Lim, D., Halmos, B., Chul Cho, B., Cobo, M., Luis González Larriba, J., Zhou, C., Demedts, I., Atmaca, A., et al. (2024). Canakinumab in combination with docetaxel compared with docetaxel alone for the treatment of advanced non-small cell lung cancer following platinum-based doublet chemotherapy and immunotherapy (CANOPY-2): A multicenter, randomized, double-blind, phase 3 trial. *Lung Cancer* 189, 107451. <https://doi.org/10.1016/j.lungcan.2023.107451>.
39. Tan, D.S.W., Felip, E., de Castro, G., Solomon, B.J., Greystoke, A., Cho, B. C., Cobo, M., Kim, T.M., Ganguly, S., Carcereny, E., et al. (2024). Canakinumab Versus Placebo in Combination With First-Line Pembrolizumab Plus Chemotherapy for Advanced Non-Small-Cell Lung Cancer: Results From the CANOPY-1 Trial. *J. Clin. Oncol.* 42, 192–204. <https://doi.org/10.1200/jco.23.00980>.
40. Gibbons, D.L., Lin, W., Creighton, C.J., Rizvi, Z.H., Gregory, P.A., Goodall, G.J., Thilaganathan, N., Du, L., Zhang, Y., Pertsemilidis, A., and Kurie, J.M. (2009). Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression. *Genes Dev.* 23, 2140–2151. <https://doi.org/10.1101/gad.1820209>.
41. Moghaddam, S.J., Li, H., Cho, S.N., Dishop, M.K., Wistuba, I.I., Ji, L., Kurie, J.M., Dickey, B.F., and Demayo, F.J. (2009). Promotion of lung carcinogenesis by chronic obstructive pulmonary disease-like airway inflammation in a K-ras-induced mouse model. *Am. J. Respir. Cell Mol. Biol.* 40, 443–453. <https://doi.org/10.1165/rcmb.2008-0198OC>.
42. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., and Satija, R. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* 42, 293–304. <https://doi.org/10.1038/s41587-023-01767-y>.
43. Gaujoux, R., and Seoighe, C. (2010). A flexible R package for nonnegative matrix factorization. *BMC Bioinf.* 11, 367. <https://doi.org/10.1186/1471-2105-11-367>.
44. Hanzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf.* 14, 7. <https://doi.org/10.1186/1471-2105-14-7>.
45. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* 16, 1289–1296. <https://doi.org/10.1038/s41592-019-0619-0>.
46. Cable, D.M., Murray, E., Zou, L.S., Goeva, A., Macosko, E.Z., Chen, F., and Irizarry, R.A. (2022). Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* 40, 517–526. <https://doi.org/10.1038/s41587-021-00830-w>.
47. Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* 32, 381–386. <https://doi.org/10.1038/nbt.2859>.
48. Fujimoto, J., Nunomura-Nakamura, S., Liu, Y., Lang, W., McDowell, T., Jakubek, Y., Ezzeddine, D., Kapere Ochieng, J., Petersen, J., Davies, G., et al. (2017). Development of Kras mutant lung adenocarcinoma in mice with knockout of the airway lineage-specific gene Gprc5a. *Int. J. Cancer* 141, 1589–1599. <https://doi.org/10.1002/ijc.30851>.
49. Tao, Q., Fujimoto, J., Men, T., Ye, X., Deng, J., Lacroix, L., Clifford, J.L., Mao, L., Van Pelt, C.S., Lee, J.J., et al. (2007). Identification of the retinoic acid-inducible Gprc5a as a new lung tumor suppressor gene. *J. Natl. Cancer Inst.* 99, 1668–1682. <https://doi.org/10.1093/jnci/djm208>.
50. Butnor, K.J., Beasley, M.B., Cagle, P.T., Grunberg, S.M., Kong, F.M., Marchevsky, A., Okby, N.T., Roggli, V.L., Suster, S., Tazelaar, H.D., and Travis, W.D. (2009). Protocol for the Examination of Resection Specimens From Patients With Primary Non-Small Cell Carcinoma, Small Cell Carcinoma, or Carcinoid Tumor of the Lung. *Arch. Pathol. Lab Med.* 133, 1552–1559.
51. Nicholson, A.G., Tsao, M.S., Beasley, M.B., Borczuk, A.C., Brambilla, E., Cooper, W.A., Dacic, S., Jain, D., Kerr, K.M., Lantuejoul, S., et al. (2022). The 2021 WHO Classification of Lung Tumors: Impact of Advances Since 2015. *J. Thorac. Oncol.* 17, 362–387. <https://doi.org/10.1016/j.jtho.2021.11.003>.
52. Rehkman, N. (2020). “Napoleon Hat” Sign: A Distinctive Cytologic Clue to Reactive Pneumocytes. *Arch. Pathol. Lab Med.* 144, 443–445. <https://doi.org/10.5858/arpa.2019-0615-SA>.
53. Xu, H., Ricciotti, R.W., and Mantilla, J.G. (2022). Practical Lung Pathology (Springer), pp. 13–19. <https://doi.org/10.1007/978-3-031-14402-8>.
54. Shilo, K., Colby, T.V., Travis, W.D., and Franks, T.J. (2007). Exuberant type 2 pneumocyte hyperplasia associated with spontaneous pneumothorax: secondary reactive change mimicking adenocarcinoma. *Mod. Pathol.* 20, 352–356. <https://doi.org/10.1038/modpathol.3800744>.
55. McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. Preprint at arXiv. <https://doi.org/10.48550/arXiv.1802.03426>.
56. Schliep, K.P. (2011). phangorn: phylogenetic analysis in R. *Bioinformatics* 27, 592–593. <https://doi.org/10.1093/bioinformatics/btq706>.
57. Greenwald, A.C., Darnell, N.G., Hoefflin, R., Simkin, D., Mount, C.W., Gonzalez Castro, L.N., Harnik, Y., Dumont, S., Hirsch, D., Nomura, M., et al. (2024). Integrative spatial analysis reveals a multi-layered organization of glioblastoma. *Cell* 187, 2485–2501.e26. <https://doi.org/10.1016/j.cell.2024.03.029>.
58. Liu, J., Manabe, H., Qian, W., Wang, Y., Gu, Y., Yan Chu, A.K., Gadhvi, G., Song, Y., Ono, N., and Welch, J.D. (2024). CytoSignal Detects Locations and Dynamics of Ligand-Receptor Signaling at Cellular Resolution from Spatial Transcriptomic Data. Preprint at bioRxiv. <https://doi.org/10.1101/2024.03.08.584153>.

59. Wolock, S.L., Lopez, R., and Klein, A.M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* 8, 281–291.e9. <https://doi.org/10.1016/j.cels.2018.11.005>.
60. Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1, 417–425. <https://doi.org/10.1016/j.cels.2015.12.004>.
61. Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* 43, 11. <https://doi.org/10.1002/0471250953.bi1110s43>.
62. Chu, Y., Dai, E., Li, Y., Han, G., Pei, G., Ingram, D.R., Thakkar, K., Qin, J.J., Dang, M., Le, X., et al. (2023). Pan-cancer T cell atlas links a cellular stress response state to immunotherapy resistance. *Nat. Med.* 29, 1550–1562. <https://doi.org/10.1038/s41591-023-02371-y>.
63. Zheng, L., Qin, S., Si, W., Wang, A., Xing, B., Gao, R., Ren, X., Wang, L., Wu, X., Zhang, J., et al. (2021). Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* 374, abe6474. <https://doi.org/10.1126/science.abe6474>.
64. Cheng, S., Li, Z., Gao, R., Xing, B., Gao, Y., Yang, Y., Qin, S., Zhang, L., Ouyang, H., Du, P., et al. (2021). A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell* 184, 792–809.e23. <https://doi.org/10.1016/j.cell.2021.01.010>.
65. Wang, R., Song, S., Qin, J., Yoshimura, K., Peng, F., Chu, Y., Li, Y., Fan, Y., Jin, J., Dang, M., et al. (2023). Evolution of immune and stromal cell states and ecotypes during gastric adenocarcinoma progression. *Cancer Cell* 41, 1407–1426.e9. <https://doi.org/10.1016/j.ccell.2023.06.005>.
66. Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740. <https://doi.org/10.1093/bioinformatics/btr260>.
67. Sinjab, A., Han, G., Treekitkarnmongkol, W., Hara, K., Brennan, P.M., Dang, M., Hao, D., Wang, R., Dai, E., Dejima, H., et al. (2021). Resolving the Spatial and Cellular Architecture of Lung Adenocarcinoma by Multiregion Single-Cell Sequencing. *Cancer Discov.* 11, 2506–2523. <https://doi.org/10.1158/2159-8290.CD-20-1285>.
68. Cang, Z., Zhao, Y., Almet, A.A., Stabell, A., Ramos, R., Plikus, M.V., Atwood, S.X., and Nie, Q. (2023). Screening cell-cell communication in spatial transcriptomics via collective optimal transport. *Nat. Methods* 20, 218–228. <https://doi.org/10.1038/s41592-022-01728-4>.
69. Nikitin, A.Y., Alcaraz, A., Anver, M.R., Bronson, R.T., Cardiff, R.D., Dixon, D., Fraire, A.E., Gabrielson, E.W., Gunning, W.T., Haines, D.C., et al. (2004). Classification of proliferative pulmonary lesions of the mouse: recommendations of the mouse models of human cancers consortium. *Cancer Res.* 64, 2307–2316.
70. Liu, Y., Sinjab, A., Min, J., Han, G., Paradiso, F., Zhang, Y., Wang, R., Pei, G., Dai, Y., Liu, Y., et al. (2025). Conserved spatial subtypes and cellular neighborhoods of cancer-associated fibroblasts revealed by single-cell spatial multi-omics. *Cancer Cell* 43, 905–924.e6. <https://doi.org/10.1016/j.ccell.2025.03.004>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-CD21 antibody [CD21]	Cell Signaling	Cat# ab271855; RRID: AB_3676138
CD4 Monoclonal Antibody [CD4]	ThermoFisher	Cat# 14-9766-82; RRID: AB_2573008
CD11c (D1V9Y) Rabbit mAb [CD11c]	Cell Signaling	Cat# 39143; RRID: AB_2924836
FoxP3 (D6O8R) Rabbit mAb [Foxp3]	Cell Signaling	Cat# 72338SF; RRID: AB_3713188
CD163 Monoclonal Antibody (10D6) [CD163]	ThermoFisher	Cat# MA5-11458; RRID: AB_10982556
Granzyme B (D2H2F) Rabbit mAb [GzmB]	Cell Signaling	Cat# 43701; RRID: AB_3713178
Rabbit anti-CD8 alpha Recombinant Monoclonal Antibody [CD8]	Bethyl Fortis	Catalog # A700-173; RRID: AB_3713184
CD3e (E4T1B) XP® Rabbit mAb [CD3e]	Cell Signaling	Cat# 73484; RRID: AB_3698885
Ki-67 (8D5) Mouse mAb [Ki67]	Cell Signaling	Cat# 62548; RRID: AB_2935678
Cd68 antibody [CD68]	ProteinTech	Cat# 28058-1-AP; RRID: AB_2881049
Claudin 4 Polyclonal Antibody (ZMD.306) [Cldn4]	ThermoFisher	Cat# 36-4800; RRID: AB_2533262
Anti-PD-L1 antibody [PDL1]	Abcam	Cat# ab233482; RRID: AB_2811045
PD-1 (Intracellular Domain) (D7D5W) XP® Rabbit mAb [PD1]	Cell Signaling	Cat# 55789sf; RRID: AB_3713186
CD31/PECAM-1 Antibodies [CD31]	Cell Signaling	Cat# 92841SF; RRID: AB_2940919
Keratin, type II; cytokeratin 8/18; EndoA antibody [KRT8]	DSHB	Cat# TROMA-I; RRID: AB_531826
F4/80 (D2S9R) XP® Rabbit mAb [F480]	Cell Signaling	Cat# 25514; RRID: AB_369910
CD45R (B220) Monoclonal Antibody [B220]	ThermoFisher	Cat# 14-0452-82; RRID: AB_467254
CD11B/Integrin Alpha M antibody [CD11b]	ProteinTech	Cat# 21851-1-AP; RRID: AB_2878927
Pan-Keratin (C11) Mouse mAb [PanCK]	Cell Signaling	Cat# 17171; RRID: AB_3697316
PE/Cyanine7 anti-mouse CD45 [CD45-PE/Cy7]	BioLegend	Cat# 103114; RRID: AB_312979
ICAM-2 Monoclonal Antibody (3C4) [ICAM2-A647]	Life Technologies	Cat# A15452; RRID: AB_2534465
Brilliant Violet 421(TM) anti-mouse CD326 (Ep-CAM) [EPCAM-BV421]	BioLegend	Cat# 118225; RRID: AB_2563983
CD324 (E-Cadherin) Monoclonal Antibody (DECMA-1) [ECAD-A488]	eBioscience	Cat# 53-3249-80; RRID: AB_10671270
IgG antibody	Invivogen	Cat# mabg1-ctrlm; RRID: AB_11203233
anti-IL-1b	Invivogen	Cat# mil1b-mab9-1; RRID: AB_3713052
anti-PD-1	BioXCell	Cat# BE0273; RRID: AB_2687796
Rat Anti-Mouse CD31 Monoclonal Antibody [CD31]	BioLegend	Cat# 102423; RRID: AB_2562186
TER-119 Monoclonal Antibody [TER119]	BioLegend	Cat# 116233; RRID: AB_10933426
Rat Anti-Granulocytes (Gr-1 antigen) Monoclonal Antibody [GR-1]	BioLegend	Cat# 108434; RRID: AB_2562219
Rat Anti-Mouse CD45 (Leukocyte Common Antigen, Ly-5) Monoclonal Antibody [CD45]	BioLegend	Cat# 103130; RRID: AB_893339
CD326 (EpCAM) Monoclonal Antibody [EpCAM]	BioLegend	Cat# 118230; RRID: AB_2629758
CD64 Monoclonal Antibody [CD64]	BioLegend	Cat# 139306; RRID: AB_11219391
Rat Anti-Siglec-F Monoclonal Antibody [SiglecF]	BD Biosciences	Cat# 740158; RRID: AB_2739911

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Rat Anti-Rat MHC Class II IA/IE Monoclonal [MHCII]	BioLegend	Cat# 107605; RRID: AB_313320
Mouse Anti-CD271/NGFR Monoclonal Antibody [CD271]	Invitrogen	Cat# 63-9400-42; RRID: AB_2762622
Rat Anti-CD11b Monoclonal Antibody [CD11b]	BioLegend	Cat# 101216; RRID: AB_312799
Anti-LAMP3 Antibody	Synaptic Systems	Cat# 391005; RRID: AB_2713987
Biological samples		
Lung adenocarcinoma samples	Juntendo University, Kobe University Hospital, MD Anderson Cancer Center	N/A
Chemicals, peptides, and recombinant proteins		
NNK	TargetMol	T20533
IL-1 β	Peptrotech	211-11B
Critical commercial assays		
Chromium Fixed RNA Human transcriptome kits	10x Genomics	4 rxns $\times$ 4 BC 1000475
Chromium Fixed RNA Human transcriptome kits	10x Genomics	4 rxns $\times$ 16 BC 100476
Chromium Fixed RNA Profiling Reagent kits	10x Genomics	CG000527
Qiagen AllPrep mini kit	Qiagen	80204
Qubit TM 1X dsDNA HS Assay Kit	Thermo Fisher Scientific	Q33231
Agilent High Sensitivity DNA Kit	Agilent	5067-4626
Chromium Next GEM Single Cell 5' Gel Bead kit v.1.1	10x Genomics	1000169
lung dissociation kit	Miltenyi Biotec	130-095-927
Bio-Plex Pro Mouse Th17 Cytokine 10-Plex Panel	BioRad	12010828
Bio-Plex Pro Mouse Cytokine 31-Plex Panel	BioRad	12009159
Mouse IL-17A (homodimer) Uncoated ELISA	Invitrogen	88-7371
Deposited data		
Human snRNA-seq processed data	This paper	GSE308103
Human Visium spatial transcriptomics processed data	This paper	GSE307534
Human WES processed data	This paper	GSE307529
Human Xenium data	This paper	Zenodo (https://doi.org/10.5281/zenodo.15670280)
Mouse scRNA-seq raw and processed data	Han et al. ¹⁹	GSE222901
Mouse Visium spatial transcriptomics raw and processed data	Han et al. ¹⁹	GSE222901
Mouse scRNA-seq raw and processed data	This paper	GSE300288
Mouse Visium spatial transcriptomics raw and processed data	This paper	GSE300293
Experimental models: Cell lines		
344SQ cells	Gibbons et al. ⁴⁰	N/A
Experimental models: Organoids		
Mouse: AT2 organoids from <i>Gprc5a</i> ^{-/-} ; <i>Krt8</i> ^{CreER/+} ; <i>Rosa</i> ^{tdT/+} mice	Han et al. ¹⁹	N/A
Experimental models: Organisms/strains		
Mouse: <i>Gprc5a</i> ^{-/-}	Han et al. ¹⁹	N/A
Mouse: <i>Krt8</i> ^{CreER/+}	Han et al. ¹⁹	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse: <i>Rosa</i> ^{tdT/+}	Han et al. ¹⁹	N/A
Mouse: <i>Kras</i> ^{LSL-G12D}	Moghaddam et al. ⁴¹	N/A
Mouse: CCSP-Cre	Moghaddam et al. ⁴¹	N/A
Mouse: <i>Il1r1</i> ^{fl/fl}	Jackson Laboratory	028398
Mouse: 129Sv	Jackson laboratory	002448
Software and algorithms		
R/4.3.1	The Comprehensive R Archive Network	https://cran.r-project.org/
cellranger/7.1.0	10× Genomics	https://10xgenomics.com/
Spaceranger/2.0.0	10× Genomics	https://10xgenomics.com/
Seurat/5.1.0	Hao et al. ⁴²	http://www.satijalab.org/seurat
loupe browser/7.1	10× Genomics	https://www.10xgenomics.com/support/software/loupe-browser/latest
SpatialInferCNV	Erickson et al. ¹⁷	https://github.com/aerickso/SpatialInferCNV
NMF/0.28	Gaujoux et al. ⁴³	https://cran.r-project.org/web/packages/NMF/index.html
ggplot2/3.5.1	The Comprehensive R Archive Network	https://cran.r-project.org/web/packages/ggplot2/index.html
GSVA/1.50.5	Hänzelmann et al. ⁴⁴	http://bioconductor.org/packages/release/bioc/html/GSVA.html
pheatmap/1.0.12	The Comprehensive R Archive Network	https://cran.r-project.org/web/packages/pheatmap/
harmony/1.2.0	Korsunsky et al. ⁴⁵	https://github.com/immunogenomics/harmony
msigdb/7.5.1	Dolgalev I (2025). <i>msigdb: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format</i> .	https://github.com/igordot/msigdb
iStar	Zhang et al. ¹⁶	https://github.com/daviddaiweizhang/istar
CytoTRACE/0.3.3	Gulati et al. ¹⁸	https://cytotrace.stanford.edu
spacexr/2.2.1	Cable et al. ⁴⁶	https://github.com/dmccable/spacexr
monocle3/1.3.7	Trapnell et al. ⁴⁷	https://cole-trapnell-lab.github.io/monocle3/docs/introduction/

EXPERIMENTAL MODELS AND STUDY PARTICIPANT DETAILS

Patient samples

Study participants were evaluated at Kobe University Hospital (Kobe, Japan), Juntendo University (Juntendo, Japan), or MD Anderson Cancer Center (Houston, TX, USA). A total of 56 samples comprising normal lung parenchyma tissues (NL), precursor (AAH, AIS) and minimally invasive (MIA) lesions, as well as LUADs were acquired from 25 patients for multimodal omics analysis (Table S1). In addition, we profiled by single-cell ST analysis (Xenium platform) an independent cohort of 8 TMAs of lung precursors and invasive lesions ($n = 36$ samples spotted as 188 cores) from 19 patients with early-stage LUAD. Samples from patients evaluated at MD Anderson Cancer Center were obtained from banked or residual tissues under informed consent and approved by MD Anderson institutional review board protocols. Specimens from all other patients were approved for study by Institutional Review Boards and according to the international ethical guidelines for biomedical research involving human subjects (Council for International Organizations of Medical Sciences). Informed written consents were received from all subjects wherever necessary. Immediately following surgery, resected tissues were processed by an experienced pathologist assistant. Sample collection was initiated with normal lung tissues followed by precursor lesions and then invasive lesions (MIA, LUAD) to minimize cross-contamination during collection. Specimens were formalin-fixed and paraffin-embedded (FFPE) and stained by hematoxylin and eosin (H&E). Clinicopathologic features of patients whose samples were selected for whole tissue multimodal (discovery) analysis or incorporated into TMAs are summarized in Table S1. For the whole tissue multimodal analysis approach ($n = 25$ patients), numbers of tissue samples analyzed by each platform are summarized by patient and lesion type in Table S1, and their numbers are as follows: ST, 56 samples; snRNA-seq, 75 samples; WES, 79 samples; and Xenium, 6 samples. For Xenium analysis of the validation cohort of 8 TMAs ($n = 19$ patients), the type and number of lesions analyzed are summarized by TMA and by patient ID in Table S1.

Animal models

Animal experiments were conducted according to the Institutional Animal Care and Use Committee (IACUC)-approved protocols at the University of Texas MD Anderson Cancer Center. Mice were maintained in a pathogen-free animal facility. In all animal experiments, sex- and age-matched mice were randomized to treatment groups. For all experiments and until endpoints were reached (up to 7 months after exposure to NNK in carcinogenesis experiments, or when subcutaneous tumor volume reached $\sim 200 \text{ mm}^3$ in experiments with syngeneic models), mice were monitored for signs of ill health, and their body weight was measured to ensure weight loss did not exceed 20% of body weight over a period of 72 h following any treatment. None of the mice developed these symptoms; therefore, they were all euthanized upon reaching IACUC-approved endpoints which were never exceeded in any of the experiments. Analysis of data from animal experiments was performed in a blind fashion. Upon necropsy, lung tissues and/or BALF were collected for downstream assays as described in subsequent sections.

For tobacco carcinogenesis experiments, *Gprc5a*^{-/-} and *Gprc5a*^{-/-};*Krt8*^{CreER/+};*Rosa*^{tdT/+} mice were used. *Gprc5a*^{-/-} mice were generated as previously described.^{48,49} *Gprc5a*^{-/-};*Krt8*^{CreER/+};*Rosa*^{tdT/+} mice used for analysis of lineage-labelled KRT8⁺ cells *ex vivo* were generated as previously described.¹⁹ *Kras*^{LSL-G12D} (CC-LR) mice used here were previously developed by crossing mice harboring *Kras*^{LSL-G12D} allele with those containing Cre recombinase inserted into the Club Cell Secretory Protein (CCSP) locus.⁴¹ These mice express mutant *Kras* in lung epithelium. *Kras*^{LSL-G12D};*Il1r1*^{fl/fl} mice were developed by crossing the CC-LR mice with mice harboring conditional allele of *Il1r1* (*Il1r1*^{fl/fl}) which were purchased from the Jackson Laboratory (stock 028398). At 14 weeks, *Kras*^{LSL-G12D} and *Kras*^{LSL-G12D};*Il1r1*^{fl/fl} mice (*n* = 6 per group) were dissected, their lung surface tumors were enumerated and then harvested as have previously described.⁴¹ Lungs were then formalin fixed and paraffin embedded for histopathology and immunofluorescence analysis.

Organoids

Organoids were derived from male and female *Gprc5a*^{-/-};*Krt8*^{CreER/+};*Rosa*^{tdT/+} mice that were treated with NNK and euthanized at end of exposure (*n* = 16) following collection and disassociation of lungs into single-cell suspensions as well as sorting for AT2 fractions. Organoids were then cultured in air-liquid interface and their numbers and size evaluated following treatment with IL-1 β or co-culture with interstitial macrophages.

Cell lines

Early passages of 344SQ mouse LUAD cells which were previously derived from a subcutaneous metastasis in a *Kras*^{LA1/+}; *p53*^{R172H Δ G} mouse model of LUAD⁴⁰ were subcutaneously implanted in six-week-old wild-type male 129sv mice that were obtained from the Jackson laboratory (strain#: 002448).

METHOD DETAILS

Histopathological processing of tissues

Paired archived H&E-stained slides from cases of surgically resected early-stage LUAD were assessed. These comprised whole tissue H&E-stained tissue samples obtained from patients with suspect invasive (LUAD) or non-invasive (precursor) lesions, as well as tissues from uninvolved (lesion-free) normal regions. If a separate normal tissue block was not available, a normal region adjacent to a precursor lesion was used as normal. H&E-stained slides, scanned at 40 $\times$ and digitally acquired using an Aperio ScanScope Turbo slide scanner (Leica Microsystems), were assessed by 3 independent experienced pathologists to determine the presence of and type of lesion (invasive, precursor), as well as identification of normal adjacent lung tissue. HALO software (Indica Labs) was used for pathological annotation. Regions of interest (ROIs) of 11 $\times$ 11mm were selected for multimodal analysis (ST, Xenium, snRNA-seq, WES), as well as an ROI for matched normal region for analysis by snRNA-seq and WES. ROIs were annotated on whole tissue SVS images using the square tool from Aperio AT2 (Leica Biosystems, Vista, CA, USA). Paired samples were selected and classified as invasive, precursor, or normal.

Inclusion criteria for LUAD selection comprised the presence of invasive adenocarcinoma occupying at least 50% of the ST ROI, and a clear interface with adjacent normal tissue. We characterized adenocarcinomas based on their predominant invasive histologic pattern (lepidic, papillary, micropapillary, acinar, or solid).⁵⁰ MIA consisted of lepidic predominant tumors measuring $\leq 3\text{cm}$ with an invasive component of $\leq 0.5\text{cm}$, lacking necrosis and lymphovascular invasion.⁵⁰ AAHs were identified as localized proliferations of thickened alveolar septa lined discontinuously by mild to moderate atypical type II pneumocytes or Clara cells with a heterogeneous round, cuboidal- or hobnail-shaped morphology, lacking stratification and mitosis and measuring $\leq 5\text{mm}$.⁵¹ AIS were classified as monomorphic or pleomorphic neoplastic epithelial cells continuously lining mildly thickened alveolar walls in a lepidic pattern with stratification or overlapping nuclei with coarse chromatin, mitosis and prominent nucleoli, without stromal or alveolar space invasion and measuring $\leq 30\text{mm}$.⁵¹ Inclusion criteria for precursor (preinvasive) lesions (AAH and AIS) also included the presence of a complete lesion within the ST ROI, and a clear interface with adjacent normal tissue. The presence of RPII was also established and identified as monotonous, cuboidal-shaped, enlarged type II pneumocytes with a proportional increase of nucleus and cytoplasm.⁵² For both preinvasive, minimally invasive and invasive ROIs, presence of specific cellular structures of interest such as intra-alveolar macrophages and RPII was also accounted for. For normal adjacent lung tissue, inclusion criteria comprised the presence of bronchiolar structures, vessels, and morphologically normal pneumocytes lining empty alveoli. Exclusion criteria for LUAD selection comprised presence of extensive necrotic areas and absence of normal lung tissue, whereas exclusion criteria for normal adjacent lung tissue,

comprised presence of extensive inflammation, fibrosis or solitary non-epithelial lesions. Once criteria were met for matching normal, LUAD, and at least one preinvasive tissue, a case was considered viable for the discovery cohort and downstream sectioning of respective FFPE blocks for the multimodal interrogation was initiated. Of note, twenty cases had paired lung precursor lesions (AAH and/or AIS) and LUAD and four cases had paired MIA and LUAD (Table S1). Further pathology review identified, by consensus, that one lung precursor tissue is a LUAD (P7 with 2 LUADs) (Table S1).

Detailed histopathologic assessment was conducted in a similar fashion on H&E images from an independent cohort of samples and for construction of human TMAs and subsequent analysis by Xenium single-cell ST (see more details below). After initial H&E analysis, selected blocks were slightly grooved with a sterile blade to mark and select for the ROI. Each FFPE block was then subsequently sectioned in the following order: 25 μ m scrolls for bulk DNA isolation, 5 μ m section for ST, followed by 2 $\times$ 25 μ m scrolls for snRNA-seq. For some samples where whole tissue Xenium analysis was conducted, an additional 5 μ m section was obtained and placed on Xenium slides for single-cell ST analysis. As previously mentioned, an independent cohort of 8 TMAs of lung precursor and invasive lesions ($n = 36$ samples spotted as 188 cores) from 19 patients with early-stage LUAD were profiled by single-cell ST analysis (Xenium platform). TMAs were constructed using 3D TMA Grandmaster (Histech), each with 25 cores. H&E images were obtained for each of the TMAs for careful histopathological examination and confirmation of presence and identity of lesions detected in original tissues.

Visium CytAssist ST profiling

5 μ m FFPE tissue sections were placed on regular plus stained slides and were stained with H&E using Leica ST5010 Autostainer XL using manufacturer-recommended protocols. H&E-stained slides were then cover-slipped and scanned at 40 $\times$ magnification with previously standardized high-resolution settings on an Aperio ScanScope Turbo slide scanner (Leica Microsystems). Before proceeding with CytAssist assay, each high-resolution image was carefully evaluated by a pathologist to determine the presence of previously selected regions, including precursor and invasive lesions, as well as the quality of the tissue: sectioning (e.g., correct tissue area placement, absence of tissue folds or contamination), staining (intensity and balance), and scanning (resolution and sharpness). Tissue samples with small or exhausted precursor lesions were excluded from the study.

ST profiling of human FFPE tissues from P1-P25 was performed using the CytAssist-enabled Visium platform with 11 $\times$ 11 mm fiducial frames (each frame comprises roughly 14,000 spots) from 10 $\times$ Genomics according to the manufacturer's recommendations. Spot-level histopathological annotation and visualization were generated using loupe browser (v.7.1). In brief, cloupe files generated from spaceranger (v2.0.0) were loaded into the loupe browser. Visualization of annotation was then generated in svg formats using the export plot tool. ST RNA-seq libraries were generated according to the manufacturer's instructions with an average of $\sim 11,000$ uniquely barcoded spots/tissue. Libraries were sequenced on NovaSeq 6000 or NovaSeq X platforms (Illumina) to achieve an average depth of $\sim 65,000$ mean reads per spot and $\sim 3,400$ median genes per spot.

Visium CytAssist ST data analysis

Cloupe files were manually annotated on Loupe Browser v7.1 with tissue categories based on histological morphology. Tumor histology was assessed using the WHO Classification of Lung Tumors recommendations.⁵¹ An experienced pathologist (A.G.S.) then checked and annotated each spot by using the Loupe Browser with the corresponding high-resolution images. A spot was annotated as "tumor" (invasive or precursor tumor cells) if it contained at least one tumor cell identified by analysis of H&E image of the same tissue section. A spot categorized as tumor may include some non-malignant cells (e.g., immune, stromal, or vascular cells). However, a spot annotated as normal epithelial should not contain any tumor cells. Spots categorized as having structures such as fibrosis, vessels, lymphoid aggregates among others were devoid of epithelial cells (tumor or normal epithelial cells). These annotations included context-specific assessments of cell type proportions by tissue and disease stage (which are ground truth). Importantly, since a primary focus here is to understand epithelial progenitor and precursor cells, snRNA-seq-derived annotations of all lung epithelial cells, including tumor, precursor and rare subsets (e.g., reactive pneumocytes) were transferred to Visium data. The category "reactive type II pneumocytes" (i.e., RPII) was assigned to spots with hyperplastic pneumocytes. This category is defined as an increase in the number of non-neoplastic pneumocytes, with or without superposition, covering alveoli in a continuous manner, or sometimes with gaps and exhibiting distinguishable regenerative cell changes, such as increased cell size compared to normal pneumocytes, hyperchromatic nuclei, small nucleoli, and mild to moderate irregularities. RPII were typically found near inflammatory or injury-related changes, such as fibrosis and intra-alveolar macrophages or hemorrhage. These reactive areas often showed thickened alveolar septa with lightly prominent stromal cells, and inflammatory infiltrates.^{53,54} In addition, the spot annotations were corroborated when using iStar¹⁶ pixel signature expression level alignment.

Demultiplexed reads were pre-processed using Space Ranger (v2.0.0, 10 $\times$ Genomics) and aligned to the GRCh38 reference genome. The resulting unique molecular identifier (UMI) count matrices were analyzed using Seurat (v5.1.0) in the R platform. Spots annotated as "Space", "Foreign body", "Tissue artifacts" were filtered out during QC processing.¹⁶ Detailed QC metrics were generated and evaluated, and spots were carefully filtered to obtain high-quality data for downstream analyses. Briefly, for basic quality filtering, spots with nUMI <500 and nFeature <200 and relative high percentage of mitochondrial genes >15% were filtered out. As a result, low-quality spots (only $\sim 1.05\%$) were also excluded from subsequent analyses, which reflected the overall high quality of the dataset.

We applied 'SCTransform' function of Seurat to normalize the gene-spot matrix to identify highly variable genes (HVGs). Top 3,000 HVGs were selected for Principal Component Analysis (PCA) and downstream unsupervised spot clustering. To identify spatially

distinct domains, the 'FindNeighbors' function of Seurat was used to construct the Shared Nearest Neighbor (SNN) Graph, based on unsupervised clustering performed with 'FindClusters' function with a resolution set to 0.8. Subsequently, 2-D visualization of the spot clusters was performed using Uniform Manifold Approximation and Projection (UMAP)⁵⁵ with Seurat function 'RunUMAP'.⁴² The top 30 Principal Components (PCs) were used to calculate the embedding. Clustering analysis was performed independently for each patient. Analysis of differentiation trajectories of lung epithelial spots was performed using Monocle 3 by inferring the pseudotemporal ordering of spots according to their transcriptome similarity.⁴⁷

Analysis of spatial copy number alterations

To infer genome-wide copy number alterations in the epithelial spots from Visium data, the 'SpatialInferCNV' R package¹⁷ was applied. To ensure the reliability of our results, an integrative approach was used to select suitable reference spots for CNA calling and clonal architecture reconstruction. Namely, the 'without reference' mode was first used to classify spots into different clones based on clustering patterns (see below inferring phylogenetic structure). Second, the inferred clone information for each spot was then mapped onto the UMAP embedding. Third, spots exhibiting the lowest CNA signal or variation, which mapped precisely to the annotated "normal lung epithelial" population, were selected as reference spots for the final round of SpatialInferCNV analysis. If no such cluster was identified, steps 1–3 were repeated with finer clone clustering until a pure normal lung epithelial cluster was selected. Upon further evaluation of spatial mapping of identified clones, for some patients (P1, P2, P3, P4, P6, P9, P14, P15, P17, P20, P21, P24) all spots annotated as normal lung epithelial were used as reference. Finally, SpatialInferCNV was run in the "with reference" mode using the selected reference spots and by applying the following parameters: cutoff = 0.1, cluster_by_groups = TRUE, HMM = TRUE, and denoise = TRUE.

To infer clonal structure and construct phylogenetic trees for each patient, the guidance provided in the SpatialInferCNV¹⁷ GitHub repository's user guide was followed. Briefly, the denoised dendrogram tree for each patient was visualized, with each node being assigned a numerical identifier. Based on the phylogenetic tree of CNA signals and corresponding heatmap, a set of nodes (as ancestor) using the 'SelectingSubTreeData' function was digitally selected. Subsequently, clone/subclone identities were assigned to these ancestor nodes. Based on their relationship of descendant nodes, each spot was attributed to the same clone identity. Seurat functions were used to spatially visualize the clones assigned to the spots.

To infer phylogenetic relationships, CNA signals were pooled into a unified subclone consensus set. Furthermore, consensus CNA scores falling within the range of 0.98–1.02, which corresponded to chromosomally stable clones, were scaled to a value of 1 to mitigate the impact of signal noise. For each patient, a phylogenetic clone tree was constructed using the Neighbor-Joining method,⁵⁶ alongside the paired reference consensus set that displayed minimal CNA signals as root. Heatmaps were generated to represent consensus CNA signals by R 'pheatmap' function.

To assess the impact of non-tumor cell admixture on CNA inference, robust cell type decomposition (RCTD), a computational approach that leverages cell type profiles derived from single-cell RNA-seq to decompose cell type mixtures,⁴⁶ was applied. RCTD was used to estimate the tumor cell fraction within each tumor (LUAD) spot. Based on these estimates, tumor spots were stratified into groups according to the proportion of intermixing with non-tumor cells, with 0% indicating spots composed exclusively of tumor cells, and 10–20% indicating spots containing 10–20% non-tumor cells, etc. For each patient, CNA calling was then performed using the same set of reference normal spots as in the overall analysis, and the inferred CNA clone signals were compared across the subgroups of spots with varying levels of admixture.

Based on the defined clones, patients were further classified into 2 broad evolutionary patterns according to the presence of shared clones between paired precursor (preinvasive) and invasive samples from the same patient (or in rarer situations paired minimally invasive/MIA and invasive LUAD), precursor-specific clones, or invasive-specific clones. A clone was classified as precursor-specific if fewer than 25 spots were present in the paired LUAD sample and over 95% of the total spots within the clone originated from the precursor sample. Similarly, a clone was defined as invasive-specific if fewer than 25 spots were present in the paired precursor sample and over 95% of the total spots within the clone originated from the LUAD sample. Clones that did not meet these criteria were classified as shared clones between preinvasive and invasive samples.

To calculate the RPII clone fraction within each sample, clones were grouped into 3 major categories: 'precursor & invasive shared', 'precursor-specific', and 'invasive-specific'. For example, if clones A1 and A2 were the only two shared clones, they were merged into the 'precursor & invasive shared' category. Similarly, if clones B1, B2, and B3 were classified as 'precursor-specific', they were combined into a single 'precursor-specific' clone. Finally, the fraction of RPII within each merged clone was calculated relative to the total number of RPII in the same sample.

Spatial resolution enhancement using iStar

To enhance tissue cellular resolution, iStar,¹⁶ an integrative approach that combines spatial transcriptomics with histological features, was used. iStar predicts spatial gene expression at a near-single-cell level, surpassing the conventional spot-level resolution of spatial transcriptomics. By leveraging hierarchical image feature extraction, it refines gene expression predictions, enabling high-resolution mapping of the tissue's cellular and molecular architecture. This advanced resolution supports detailed analyses in tissue segmentation, cell-type identification, and tumor microenvironment (TME) characterization. iStar was used herein to visualize gene expression, signature scores, and meta-program (MP) scores for each sample.

MP identification from visium ST data

The 'sctransform' function was applied on Visium ST data, followed by NMF⁴³ on the normalized gene expression matrix of each sample separately, to capture continuous patterns of gene expression variation. To minimize the influence of a single k parameter selection, NMF was carried out with multiple k values ranging from 2 to 18, generating 170 NMF programs per sample. Each NMF program was summarized by the top 50 genes based on their NMF scores.

Consensus MPs were generated to integrate gene programs across the whole cohort. To identify robust gene programs across samples, per-sample NMF gene programs were clustered based on their Jaccard index (i.e., intersection/union of gene sets). The clustering approach used to define MPs followed the methodology described in Gavish et al.²² and Greenwald et al.⁵⁷ Given the large number of NMF programs, only the most robust and non-redundant programs were included in clustering.

MPs were retained within a sample if they had at least 35/50 gene overlap with another program generated from a different k value. To avoid redundancy, if multiple programs within a sample had $\geq 10\%$ gene overlap, only one was retained. To ensure cross-sample consistency, a program was considered robust across samples if it had $\geq 10\%$ gene overlap with a program from another sample. Using this approach, we identified 11 MPs across covering all major cell lineages (Table S2). Gene set variation analysis was used to calculate scores for the 11 MPs for each spot. Each spot was assigned an MP label based on the highest-scored MP. Then, sample-level pairwise correlations of MPs were analyzed based on the proportion of each MP in each sample.

Analysis of ligand-receptor signaling

CytoSignal⁵⁸ was used to infer the spatial locations and dynamics of cell-cell communication at a cellular resolution using human Visium ST data. This approach enables the spatial resolution of signaling scores, identification of spatial gradients in signaling strength, quantification of contact-dependent and diffusible interactions, and detection of signaling-associated DEGs.

snRNA-seq data generation

snRNA-seq was done on the same samples analyzed by ST (Table S1), using FFPE scrolls consecutive to ST-analyzed sections. Scrolls were processed for single nucleus extraction following instructions demonstrated in 10 $\times$ Genomics' protocol for sample preparation from FFPE tissue sections and for Chromium Fixed RNA profiling (CG000632). Specifically, scroll digestion was achieved using gentleMACS Octo Dissociator with heaters (Miltenyi Biotec, 130-096-427), operating under the program settings recommended by the manufacturer for optimal tissue dissociation. Throughout the process of digestion and washing of nuclei, efficiency of digestion and enumeration of nuclei was evaluated by staining with ViaStain PI Staining Solution (Revvity Health Sciences Inc, CS101095ML) and assessing nuclei samples on a Countess II FL Automated Cell Counter (Thermofisher Scientific, AMAQAF1000). Cell counts following each of the tissue homogenization, probe hybridization, and pooling steps were an average of 3 per sample. For sample multiplexing, sample barcoding (BC) and pooling was done using Chromium Fixed RNA Human transcriptome kits (10 $\times$ Genomics, 4 rxns $\times$ 4 1000475 BC, or 4 rxns $\times$ 16 100476 BC). To capture more cells from precursor and invasive samples with a lower undetected multiplet rate, cells from each of AAH, AIS, MIA and LUAD samples were sub-pooled over two probe barcodes each, to target up to 20,000 cells per sample. Cells from normal lung tissue samples were assigned one barcode, to target up to 10,000 cells. Pooled samples were processed Chromium X instrument (10 $\times$ Genomics, 1000331) and libraries were generated following manufacturer's recommendations outlined in Chromium Fixed RNA Profiling Reagent kits (10 $\times$ Genomics, CG000527). Libraries were sequenced on NovaSeq 6000 or NovaSeq X platforms (Illumina) to achieve a depth of at least 50,000 mean read pairs per cell.

snRNA-seq data analysis

Raw data were pre-processed (demultiplex cellular barcodes, read alignment and generation of gene count matrix) using Cell Ranger Single Cell Software Suite provided by 10 $\times$ Genomics. For alignment of reads, human reference GRCh38 (hg38) was used. Detailed QC metrics were generated and evaluated, and cells were carefully and rigorously filtered to obtain high-quality data for downstream analyses. In brief, for basic quality filtering, cells with low-complexity libraries (in which detected transcripts were aligned to <200 genes such as cell debris, empty drops and low-quality cells) were filtered out and excluded from subsequent analyses, and genes detected in less than 3 cells were also excluded. Nuclei with less than 500 detected genes or with less than 1000 reads count or with a mitochondrial gene fraction that is $\geq 20\%$ were filtered out using Seurat (v5.1.0). Doublets were identified based on library complexity. First, a python-based software Scrublet was applied to identify the doublets by each sample.⁵⁹ Second, based on cluster distribution and marker gene expression, doublets forming distinct clusters with hybrid expression features were also removed. Expression levels and proportions of canonical lineage-related marker genes in each identified cluster were carefully reviewed. Clusters co-expressing discrepant lineage markers were identified and removed.

Data normalization and scale transformation were performed using method 'SCTransform' in Seurat. Top 3,000 HVGs were selected for PCA and downstream unsupervised clustering. The top 50 PCs were used to calculate the embedding. Harmony (version 1.2.0)⁴⁵ was run with default parameters to remove batch effects. To identify distinct cell types, the 'FindNeighbors' function of Seurat was used to construct the SNN Graph, based on unsupervised clustering performed with 'FindClusters' function with a resolution set to 0.5–0.8. Subsequently, 2-D visualization of the spot clusters were performed using UMAP⁵⁵ with Seurat function 'RunUMAP'.⁴² Differentially expressed genes (DEGs) of each cluster were identified by using the 'FindAllMarkers' function in Seurat. Genes expressed in at least 10% of each group were kept for further analysis. Differentially expressed genes in human KACs relative to other alveolar lineage and tumor cells were further used for HALLMARK pathway enrichment by using R package 'msigdb'. HALLMARK

pathways were downloaded from the MSigDB database.⁶⁰ Analysis of differentiation trajectories of lung alveolar and malignant cells was performed using Monocle 3 by inferring the pseudotemporal ordering of cells according to their transcriptome similarity.⁴⁷

NMF was performed on each patient separately to capture continuous patterns of gene expression variation. Negative values in each centered expression matrix were transformed to zero. To minimize the influence of a single k parameter selection, we ran NMF with multiple k values ranging from 2 to 11, generating 65 NMF programs per patient. Each NMF program was summarized by the top 50 genes based on their NMF scores. To integrate gene programs across patients, MPs were generated in a similar manner as described above for Visium ST data, except using a 20% gene overlap cutoff for program retention and robustness. Using this approach, 9 MP clusters across all cell types were identified including 9 MP clusters that were specific for epithelial cells (Table S2). 'AddModuleScore' function in Seurat was later used to calculate scores for the 9 MPs. Each cell was assigned an MP label based on the highest corresponding MP score. When applying the MPs to human Visium ST data, the MP label for each spot was determined using the same approach.

WES data processing and analysis

Total DNA was isolated from 25 μ m FFPE scrolls of the same human lung tissues used for multimodal analysis. Deparaffinization was done using Deparaffinization Solution (Qiagen, 939018) followed by DNA isolation using Qiagen AllPrep mini kit (Qiagen, 80204) and according to the manufacturer's recommendations. DNA yield was assessed using Qubit 1 $\times$ dsDNA HS Assay Kit (Thermo Fisher Scientific, Q33231) on a Qubit 4 Fluorometer (Thermo Fisher Scientific). DNA quality was assessed using Agilent High Sensitivity DNA Kit (Agilent, 5067-4626) on an Agilent 2100 Bioanalyzer System. TWIST-WES was performed on a NovaSeq 6000 platform (Illumina) at a depth of 100 $\times$ for tumor samples and 300 $\times$ for normal, pre-invasive, or mildly invasive samples and using matched normal lung tissues as germline control (total samples = 79).

Reads were mapped to the human reference genome GRCh38 (hg38) using BWA-mem. Duplicates were marked for filtering and base quality scores were recalibrated using GATK.⁶¹ Somatic single nucleotide variants and indels were called using MuTect2 with default settings followed by post-calling filtering, and they were subsequently annotated using GATK Funcotator and ANNOVAR. The following criteria were applied to filter variants: (i) only variants marked as "PASS" by MuTect2 were considered; (ii) a minimum coverage of 10 $\times$ for normal samples and 20 $\times$ for non-normal samples (AAH, AIS, MIA and LUAD) was required; (iii) variants were required to have at least 4 alternate reads or a variant allele frequency (VAF) of $\geq 5\%$ in the non-normal regions, while having fewer than 3 alternate reads and a VAF of $< 5\%$ in the corresponding normal regions, and a VAF ratio (normal/non-normal) of < 0.2 ; (iv) only variants with a median distance of at least 5 bases from the read ends were considered; (v) common variants reported by gnomAD (the Genome Aggregation Database, <https://gnomad.broadinstitute.org/>), Phase-3 1000 Genome Project (http://phase3browser.1000genomes.org/Homo_sapiens/Info/Index) or the NHLBI GO Exome Sequencing Project (ESP6500) (<http://evs.gs.washington.edu/EVS/>) with minor allele frequencies greater than 0.5% were removed. Somatic mutations located within introns, untranslated regions (UTRs), UTR-flanking regions, or those classified as silent mutations were also excluded from subsequent analyses. The somatic mutation burden for each sample was calculated based on 3 major mutation types (missense, truncating, and in-frame indels), as previously described.¹⁹

Xenium data generation

Pathology QC prior to Xenium processing was performed by evaluating scanned H&E-stained sections from human whole tissues (matching 3 AAH, 3 AIS, and 6 LUADs) and from TMAs (8 TMAs comprising 36 cores: 8 AAH, 10 AIS, 5 MIA, and 13 LUADs) for quality of tissue sectioning, staining, scanning and resolution for each tissue and each TMA core. Quality assessment of tissue sectioning evaluated the presence or absence of tissue detachment, folding or tearing. Staining QC evaluated over- or under-staining and differences in stain concentration. Scanning and resolution QC evaluated blurriness and contrast. Histopathology review was completed on each tissue or tissue core to confirm the presence of precursor (AAH, AIS) and minimally invasive (MIA) lesions, as well as LUADs, as previously identified during H&E review of the initial/archived whole tissue section from each patient sample. For samples passing QC and chosen for whole tissue Xenium analysis, 11 $\times$ 11 mm regions of interest were marked, the same regions studied by ST for the respective chosen samples. 11 $\times$ 11 mm regions from each of the selected whole tissue samples were sectioned and placed in pairs (3 AAH/LUAD pairs and 3 AIS/LUAD pairs) on Xenium slides. Samples were processed on Xenium Analyzer (10 $\times$ Genomics, 1000481) following manufacturer instructions outlined in Xenium Prime (5K) *In situ* Gene Expression (10 $\times$ Genomics, CG000760). For processing of TMAs ($n = 8$) on Xenium, two TMAs (containing cores from matched patient tissues) were fitted per Xenium slide and processed on Xenium Analyzer (10 $\times$ Genomics, 1000481) following manufacturer instructions outlined in Xenium *In situ* Gene Expression (10 $\times$ Genomics, CG000749). Each of the two abovementioned Xenium assays was paired with multimodal cell segmentation and one of two in-house custom-designed 100 gene panel add-ons that include additional genes for phenotyping major subsets of various lung epithelial cells, including alveolar and alveolar intermediate cells (see Table S4).

After each run, the Xenium slides underwent H&E staining, scanning and pathology QC. Each post-Xenium H&E-stained slide was converted to an OME.TIFF file using QuPath v0.5.1 and for compatibility with Xenium Analyzer. Converted H&E-stained images were imported to Xenium Explorer v3.2.0 or later for alignment with tissue images stained with morphology markers (DAPI for nucleus, ATP1A1/CD45/E-cadherin for boundary, 18S rRNA/ α SMA/Vimentin for interior), based on which a multi-modal cell segmentation approach was applied to obtain cellular boundaries and centroids. Precise image alignment was manually obtained by placing a minimum of 30 key points on both images, and identifying corresponding landmarks on the tissues such as nuclear edges. Once the alignment was complete, cell segmentations were overlaid onto the H&E images for visualization. For pathology annotations of lesion

regions, freehand selection on Xenium Explorer was performed, with polygons drawn around each premalignant or malignant lesion and around groups of RPII cells. Then, cell barcodes within each selected region were extracted for downstream analysis and visualization.

Xenium data analysis

Xenium 5K data (12 samples from 6 patients) including gene count matrices, metadata, cellular centroid and boundary coordinates, as well as transcript coordinates, were downloaded from Xenium Analyzer (10× Genomics, CG000584). The gene count matrix was used for transcriptomic analysis with the R software Seurat (v5.1.0).⁴² For initial QC, cells with a gene count of less than 30 were filtered out. Furthermore, cells lacking the expression of defined canonical markers were also excluded from subsequent analysis. Finally, a total of 4,598,777 characterizable high-quality cells were used for downstream analysis. Then, data normalization was performed with the function ‘SCTransform’ in Seurat.⁴² A linear dimension reduction by PCA was performed based on HVGs. Elbow plots were drawn with the ‘ElbowPlot’ function to determine the number of dimensions used for downstream analysis. Then, slide-specific batch effects were evaluated, and the R package ‘Harmony’ (version 1.2.0)⁴⁵ was used for data integration in the PCA space for each major cell compartment. Following that, a nearest-neighbor graph was constructed using the function ‘FindNeighbors’ in Seurat,⁴² and unsupervised cell clustering was conducted with the Louvain algorithm with the function FindClusters. Further non-linear dimension reduction was performed with the UMAP method,⁵⁵ implemented with the ‘RunUMAP’ function in Seurat.⁴²

DEGs among cell clusters were identified using the ‘FindAllMarkers’ function in Seurat⁴² and the top DEGs of each cell cluster were reviewed in detail. Additionally, expression of canonical lineage markers, as reported previously,^{19,62–65} was manually checked. Subsequently, major cell compartments were identified, including T and NK cells (*CD3D*, *CD2*, etc. for T cells; *NCAM1*, *FCGR3A*, etc. for NK cells), B and plasma cells (*MS4A1*, *CD79A*, etc. for B cells; *MZB1*, *XBPI1*, etc. for plasma cells), myeloid cells (*CD68*, *CD163*, etc. for macrophages; *CSF3R* for neutrophils; *KIT*, *MS4A2*, etc. for mast cells; *CD1C*, *CLEC9A*, etc. for conventional dendritic cells-cDC, *LILRA4*, *IL3RA*, etc. for plasmacytoid dendritic cells-pDC), stromal cells (*COL5A1*, *PDGFRA*, etc. for fibroblast; *VWF*, *PECAM1*, etc. for vascular endothelial cells; *PROX1*, *RELN*, etc. for lymphatic endothelial cells; *PDGFRB*, *RGS5*, etc. for pericytes; *MYH11* for smooth muscle cells/SMC; *CCN3*, *CHODL*, etc. for chondrocytes), and epithelial cells (highly expressing *EPCAM*, *CDH1*, etc.). Then, each major cell compartment was further subclustered for identification of detailed cell types and states, with similar methods as above. For epithelial cells, normal airway and alveolar cell lineages as well as atypical and tumor epithelial cells were identified by integrating information on gene expression profiles and histological features. Cell type labels were mapped back to each tissue image in Xenium Explorer for visualization of corresponding cellular morphological features and for validation of the accuracy of cell types identified.

The Xenium TMA data for 36 samples from 19 patients, including gene count matrices, metadata, cellular centroid and boundary coordinates, as well as transcript coordinates, were downloaded from Xenium Analyzer. Data QC, processing, as well as cell type and state identification, were performed with a similar approach as described in Xenium 5K data analysis with minor differences. Briefly, cells with UMI counts less than 20 or gene counts less than 10 were removed. After quality filtering, a total of 593,334 characterizable high-quality cells were retained for downstream analysis, which were further subject to dimension reduction and unsupervised clustering. Major cell lineages were subsequently identified, including T, NK, B and plasma, myeloid (monocyte, macrophage, cDC, and mast), stromal (CAF, vascular endothelial, lymphatic endothelial, pericyte, SMC), along with normal, atypical and tumor epithelial cells.

Cellular neighborhood analysis was performed on Xenium 5K and on Xenium TMA data. We defined the neighborhood of each cell as a collection of surrounding cells whose centroid lies within 80 μm of the centroid of the center cell. Each cell in the dataset was iteratively selected as the center cell, and then a cell-by-cell state composition matrix was calculated, with each column representing a cell and each row representing the count of a cell state within each cell’s neighborhood. Subsequently, the neighborhood compositions of different epithelial cell lineages, as well as atypical and tumor epithelial cells, were quantified and compared.

To quantify the enrichment of different cell or spot types across different groups (including cellular compositions, pathology categories, or lesion types), and to examine whether a certain cell type or spots expressing certain MPs are enriched or depleted, the ratio of observed to expected cell or spot numbers of each cell or spot type across different groups were calculated using previously reported methods.⁶⁵ First, the contingency table of cell or spot types by groups was calculated. The chi-square test was then used to calculate expected distributions of cell or spot types across different groups and to evaluate the deviation of actual from expected distribution. Following that, $R_{o/e}$ values for each cell or spot type - group combination were calculated using the following formula:

$$R_{o/e} = \frac{\text{Observed}}{\text{Expected}}$$

Here, $R_{o/e}$ is the ratio of observed to expected cell numbers for a particular cell or spot type in a certain group. For a particular cell type, $R_{o/e} > 1$ indicates that this cell or spot type is more frequently observed than the random expectations in a specific group (i.e., enriched), while $R_{o/e} < 1$ indicates that the number of cells or spots of this type is less than the random expectation and thus depleted. The results were shown as heatmaps using the R package ‘pheatmap’ (version 1.0.12).

KAC, NF- κ B and interferon signature scores

The KAC (RP11) score was derived previously.¹⁹ The NF- κ B score was computed using genes extracted from the NF- κ B pathway in the BioCarta database curated in MSigDB.⁶⁶ The interferon score was derived from the interferon meta-program reported by Barkley et al.²¹ All three scores were computed using GSVA⁴⁴ with the inner “ssgsea” method and default parameters. Genes constituting all scores are summarized in Table S3.

Mouse scRNA-seq data generation

For experiments investigating effect of tobacco carcinogen exposure on LUAD development at the single-cell level and which we have previously reported,¹⁹ 4 *Gprc5a*^{-/-} mice per treatment group (saline or NNK) and analysis timepoint (end of exposure/EOE, 3 months or 7 months post exposure) were used for scRNA-seq (Figure 6A, see below section for more details on NNK exposure). For experiments comparing control IgG, anti-IL-1 β , anti-PD-1 and combination of IL-1 β and PD-1 blockade (Figure 7A), 4 NNK-exposed *Gprc5a*^{-/-} mice per treatment group and timepoint/strategy were used for scRNA-seq analysis and upon QC one sample was excluded ($n = 31$). Lungs were obtained at endpoint, processed into single cells, and sorted to enrich for epithelial cells as previously described.¹⁹ Namely, mouse lung single cells were stained with Sytox Green and the following cocktail of antibodies: CD45-PE/Cy7 (BioLegend, 103114), ICAM2-A647 (Life Technologies, A15452), EPCAM-BV421 (BioLegend, 118225) and ECAD-A488 (eBioscience, 53-3249-80). To enrich epithelial cells, live (Sytox⁻) epithelial singlets (CD45⁻; ICAM2; EpCAM⁺ &/or ECAD⁺) were sorted and mixed at a 1:1 ratio with live cells from each sample. Mixtures were then loaded on a Chromium Controller (10 $\times$ Genomics, 110211) to achieve an average target of $\sim 5,500$ cells per sample. 5' single-cell gene expression libraries were prepared using the Chromium Next GEM Single Cell 5' Gel Bead kit v.1.1 (10 $\times$ Genomics, 1000169) as previously described.⁶⁷ Sequencing was performed on NovaSeq 6000 platform (Illumina) to achieve an average depth of $\sim 50,000$ mean reads per cell.

Mouse scRNA-seq data analysis

Raw scRNA-seq data were pre-processed (demultiplex cellular barcodes, read alignment and generation of gene count matrix) using Cell Ranger Single Cell Software Suite provided by 10 $\times$ Genomics. For read alignment of mouse data, mouse reference GRCm38 (mm10) was used. Data processing steps similar to those described for the human snRNA-seq data were employed for detailed QC, including the primary filtering of cells with < 200 genes and genes expressed in < 3 cells. In addition, cells with fewer than 500 detected genes, with less than 1000 reads count, or with a mitochondrial gene fraction that is $\geq 15\%$ were filtered out using Seurat (v5.1.0). Doublets in each sample were identified by Scrublet⁵⁹ and subsequently removed. In downstream analysis, clusters co-expressing discrepant lineage markers were identified and removed. ‘SCTransform’ was still used to normalize and scale the data and top 3,000 HVGs were selected for PCA and downstream unsupervised clustering. The top 30 PCs were used to calculate the embedding. Due to similar genomic backgrounds, no additional batch effect adjustment method was further applied. The functions ‘FindNeighbors’, ‘FindClusters’, ‘RunUMAP’, and ‘FindAllMarkers’ were implemented as in human data processing.¹⁹ Differentially expressed genes in mouse KACs relative to other alveolar lineage and tumor cells were further used for HALLMARK pathway enrichment by using R package ‘msigdb’. HALLMARK pathways were downloaded from the MSigDB database.⁶⁰

Visium ST profiling of mouse lung tissues

For *Gprc5a*^{-/-} mouse lung tissues analyzed by ST, mice at EOE and at 7 months post exposure to either saline or NNK were studied (2 mice per treatment and timepoint with each mouse lung split across two 6.5×6.5 mm capture areas; $n = 8$ mice; $n = 16$ libraries total; Figure 6A). ST data were generated from 2 regions per mouse (i.e., 4 regions per treatment and timepoint) using 10 $\times$ Genomics’ Visium platform (6.5×6.5 mm ROI and capture areas). Histopathological QC and data generation for all samples were done as previously described for 3 tumor-bearing lung tissues analyzed at 7 months post-NNK.¹⁹

For ST profiling of lungs of NNK-exposed *Gprc5a*^{-/-} mice which received anti-IL-1 β or control IgG and analyzed at 3 and at 7 months post end of NNK (Figure 7A), the Visium CytAssist platform from 10 $\times$ Genomics was used. Six samples representative of the tumor volume analysis were carefully selected by analyzing their respective H&E images for presence of lesions and quality of the tissue, namely, sectioning (e.g., correct tissue area placement, absence of tissue folds or contamination), H&E staining (intensity and balance), and scanning (resolution and sharpness). Chosen samples comprised lungs of mice treated with IgG and analyzed at 3 months and 7 months post end of NNK (1 mouse per timepoint), and mice treated with anti-IL-1 β analyzed at 3 months and 7 months post end of NNK (2 mice per timepoint). One 11×11 mm ROI was selected for each mouse lung for downstream analysis by ST. ST was performed on the selected archived and de-cover-slipped H&E slides by carefully positioning the 10 $\times$ Genomics CytAssist gaskets to select and process these regions of interest, thereby targeting an average of $\sim 8,500$ uniquely barcoded spots per tissue. ST profiling was conducted as described above for human tissues, except for using 10 $\times$ Genomics mouse transcriptome-specific kits. Libraries were sequenced on NovaSeq 6000 platform (Illumina) to achieve an average depth of $\sim 80,000$ mean reads per spot and $\sim 3,800$ median genes per spot.

Visium ST analysis of mouse lung tissues

Demultiplexed reads were pre-processed using spaceranger (v.2.0.0, 10 $\times$ Genomics) and aligned to the GRCm38 (mm10) reference genome. Post-ST spot-level annotation was done by an experienced pathologist (L.I.G.B.) using Loupe Browser with the corresponding high-resolution images, as described earlier for human samples. The resulting unique molecular identifier (UMI) count matrices were analyzed using Seurat (v5.1.0) in the R platform. QC metrics were generated and evaluated, and spots were carefully

filtered to obtain high-quality data for downstream analyses. Briefly, for basic quality filtering, spots with nUMI <500 and nFeature <200 and relatively high percentage of mitochondrial genes >20% were filtered out. Downstream data processing was carried out in the same manner by which human samples were analyzed by ST (see previous section).

KAC¹⁹ and inflammatory⁶⁰ signature scores (Table S3) were visualized using iStar based on enhanced gene expression. The inflammatory score was computed from genes curated from the Hallmark Inflammatory Response Pathway in MSigDB. These scores were then mapped onto the H&E image at the super-pixel level. Scores above the 90th percentile for each signature were considered as ‘overlay.’ Scores between the 70th and 90th percentiles were classified as either ‘KAC’ or ‘Inflammatory,’ depending on which value was higher. Scores below the 70th percentile were not displayed and were represented with a black background.

The COMMOT algorithm⁶⁸ was used to infer *Il1b-Il1r1* ligand-receptor interactions between neighboring cells within the mouse Visium ST dataset. To explore potential *Il1b-Il1r1* signaling in the tumor microenvironment, we restricted our analysis to interactions between macrophages, RPII, and tumor spots. Additionally, we collapsed dimers into individual genes using the CellChat database, defining the signaling type as “Cell-Cell Contact”. All other parameters in COMMOT were set to default values.

Carcinogen and antibody treatments in mice

For NNK exposure, sex- and age-matched 8-week-old mice of either genotype were intraperitoneally injected with 75 mg/kg of body weight NNK (TargetMol, T20533), 3 times per week for 8 weeks. Following 8 weeks of NNK exposure, sex- and age-matched *Gprc5a*^{-/-} mice were split to receive one of two treatment windows. The first group received antibodies for 3 months beginning immediately following end of NNK exposure (prevention setting). The second group received the same antibodies starting at 3 months following end of NNK exposure and for 4 months (interception setting). Within each setting, mice were split into 4 treatment groups receiving one of the following *in vivo*-grade antibodies intraperitoneally: control mouse IgG antibody (Invivogen, mabg1-ctrlm), anti-IL-1 β (Invivogen, clone 7E3, mabg-mil1b), anti-PD-1 (BioXCell, clone 29F.1A12, BP0273), or a combination of both anti-IL-1 β and anti-PD-1. Mice were weighed on a weekly basis. Anti-IL-1 β was administered twice a week at 10 mg/kg whereas anti-PD-1 was administered twice a week at 250 μ g per dose. Endpoints were reached immediately after end of antibody treatments for both strategies, i.e., for the prevention setting, at 3 months post end of NNK after a total of 24 doses of antibody treatments, and in the interception setting, at 7 months post end of NNK after 36 doses of antibody treatments.

Histopathology analysis of mouse lung tissues

Lungs of *Gprc5a*^{-/-} mice (at least $n = 8$ per treatment and time point, total $n = 86$ mice) were inflated with formalin by gravity drip inflation, excised, examined for lung surface lesions by macroscopic observation and processed for FFPE, sectioning and H&E staining. For lesion volume assessment, a 4 μ m surface H&E section was obtained, followed by 5 unstained sections (skipping a total of 20 μ m) and another deeper section for H&E. The process was repeated twice to achieve a total 1 surface H&E and 3 deep H&E's at 28 μ m, 52 μ m, and 76 μ m depths. H&E-stained slides were cover-slipped and digitally scanned using an Aperio ScanScope Turbo slide scanner (Leica Microsystems) at 40 $\times$ magnification and visualized using ImageScope software (Leica Microsystems). Enumeration of lung tumors from each mouse group was conducted through histopathological analysis of the 4 H&E-stained slides.⁶⁹ Tumor volume (measured in mm³) was calculated by adding, from the different sections, the product area of tumor lesions (measured in mm²) by the thickness of tissue sections on each slide. Percent tumor area in *Kras*^{LSL-G12D} and *Kras*^{LSL-G12D};*Il1r1*^{fl/fl} mice was similarly determined based on the sum of areas of lung lesions from total lung area.

Immunofluorescence analysis of mouse tissues

Unstained lung tissue sections from *Kras*^{LSL-G12D} and *Kras*^{LSL-G12D};*Il1r1*^{fl/fl} mice were analyzed by immunofluorescence (IF) analysis. Deparaffinized slides underwent antigen retrieval using 1 $\times$ BioCare Diva Decloaker RTU (BioCare, DV2004G1) buffer in a Decloaking Chamber at 110°C for 15 min. They were then blocked in PBS with 0.3% Triton X-100 and 5% normal goat serum (Jackson ImmunoResearch, 005-000-121) and incubated overnight in a humidified chamber at 4°C with primary antibodies diluted in PBS with 0.3% Triton X-100 and raised against: LAMP3 (Synaptic Systems, 391005) and KRT8 (University of Iowa DSHB, TROMA-I). Slides were then stained with fluorophore-conjugated secondary antibodies (Jackson ImmunoResearch) and 4',6'-diamidino-2-phenylindole (DAPI, D9542-10MG). Sections were mount with Vectashield Antifade Mounting Medium (VectorLabs, H-1000-10), cover slipped, imaged using a Nikon A1plus confocal microscope and analyzed using QuPath software. LAMP3⁺/KRT8⁺ cells were then quantified in each of 3 *Kras*^{LSL-G12D} and 3 *Kras*^{LSL-G12D};*Il1r1*^{fl/fl} mice and the mean number of cells was plotted.

Analysis of AT2-derived KRT8 high organoids

Gprc5a^{-/-};*Krt8*^{CreER/+};*Rosa*^{tdT/+} mice were generated as described previously¹⁹ and treated with NNK as described above and euthanized at end of exposure ($n = 16$). Lungs were immediately collected and dissociated into single-cell suspensions using the lung dissociation kit (Miltenyi Biotec, 130-095-927) following the guidelines provided by the manufacturer. Tissues were processed in a gentleMACS Octo Dissociator with heaters (Miltenyi Biotec, 130-096-427), operating under the program settings recommended by the manufacturer for optimal tissue dissociation while preserving cell viability. The resulting cell suspensions were filtered through a 70 μ m cell strainer (Miltenyi Biotec, 130-098-462) to remove debris and aggregates. Subsequently, red blood cells were lysed using ACK lysis buffer (Thermo Fisher Scientific, A1049201) for 3 min at room temperature. Lysed suspensions were then centrifuged, resuspended in ice-cold PBS, and filtered (using 40 μ m FlowMi tip filters; Millipore, H13680-0040). Cell suspensions were stained with the following antibodies: CD31 (BioLegend, clone:390, 102423), TER119 (BioLegend, clone:TER-119, 116233), GR-1 (BioLegend,

clone:RB6-8C5, 108434), CD45 (BioLegend, clone:30-F11,103130), EpCAM (BioLegend, clone:G8.8, 118230), CD64 (BioLegend, clone: $\times 54$ -5/7.1,139306), SiglecF (BD Biosciences, clone: E50-2440, 740158), MHCII (BioLegend, clone: M5/114.15.2, 107605), CD271 (Invitrogen, clone: ME20.4, 63-9400-42), CD11b (BioLegend, clone: M1/70, 101216). Live (Sytox blue-negative) AT2 cells (EpCAM⁺, MHCII⁺, CD31⁺, TER119⁺, CD271⁺) and interstitial macrophages (CD45⁺, CD64⁺, SiglecF⁺, CD11b⁺, CD31⁺, TER119⁺, GR-1⁺) were sorted by flow cytometry using a BD FACS Aria III cell sorter. Collected AT2 cells were immediately washed and resuspended at a concentration of 5,000 cells per 50 μ L of 3D culture medium (DMEM/F12 medium supplemented with insulin, transferrin and selenium, 10% FBS, penicillin-streptomycin and L-glutamine). AT2 cells were mixed at a 1:1 ratio (by volume) with 25,000 mouse endothelial cells (collected from mouse lungs by CD31 selection and expanded *in vitro* (as previously described¹⁹) and resuspended in 50 μ L of Matrigel growth factor reduced basement membrane matrix (Corning, 356231). Then, 100 μ L of 1:1 AT2:endothelial cell mixture was plated on transwell inserts with 0.4 μ m pores (Corning, 3450) and allowed to solidify for 30 min in a humidified CO₂ incubator. Each well's lower chamber was then supplemented with 3D medium containing ROCK inhibitor (Millipore, Y-27632) and recombinant mouse FGF-10 (R&D Systems, 6224-FG), and plates were incubated at 37°C in a humidified CO₂ incubator. 4-Hydroxytamoxifen (Sigma-Aldrich, T176) was given *in vitro* at 2 μ M, and wells were replenished with 3D medium every other day. Organoids were monitored and analyzed once a week using an EVOS M7000 imaging system (Thermo Fisher Scientific), whereby the numbers and sizes of organoids greater than 100 μ m in diameter were recorded.

For *ex vivo* IL-1 β treatment of lung AT2 cells, 20 ng/ml IL-1 β (Peprotech, 211-11B) was added to the organoid medium in the lower chamber and changed every other day. For coculture of AT2 cells and macrophages, AT2 cells (EpCAM⁺, MHCII⁺, CD31⁺, TER119⁺, CD271⁺) were plated with interstitial macrophages (CD45⁺, CD64⁺, SiglecF⁺, CD11b⁺, CD31⁺, TER119⁺, GR-1⁺) at a 1:6 ratio. Image acquisition as well as documentation of the numbers and sizes of organoids was done as described above.

IL-1 β blockade in syngeneic model of LUAD

We used 344SQ mouse LUAD cells which were previously derived from a subcutaneous metastasis in a *Kras*^{LA1/+}; *p53*^{R172H Δ G} mouse model of LUAD.⁴⁰ Early passage 344SQ cells were cultured in DMEM (Gibco, 11995-065) supplemented with 10% FBS (Gibco, 16000-044), 2 mmol/L L-glutamine (Sigma-Aldrich, G7513), and 1% penicillin-streptomycin (Gibco, 15140-122), and they were maintained at 37°C in a humidified 5% CO₂ atmosphere. Cells were sub-cultured every 2 to 3 days to prevent over confluence. Six-week-old wild-type male 129sv mice were obtained from the Jackson laboratory (strain#: 002448) and underwent a 3-day quarantine, followed by a 1-week acclimation period. Mice were subcutaneously injected with 5×10^5 344SQ cells at 8 weeks of age. 344SQ cells in culture medium were suspended with Matrigel (Corning 356237) at a 1:1 ratio (by volume), with a total injection volume of 100 μ L per mouse. Tumor growth was assessed using caliper measurements, with tumor volume calculated as $\frac{1}{2} \times \text{length} \times \text{width}^2$. Once tumor volume reached ~ 200 mm³, mice (7/group) were randomly divided to receive intraperitoneal injections of either anti-mIL-1 β (InvivoGen, mil1b-mab9-10) or control mlgG1 monoclonal antibody (InvivoGen, bgal-mab9-10). Injections were administered at a dosage of 250 μ g per mouse twice per week. Tumors were measured before each treatment until reaching the experimental endpoint (tumor volume = 2000 mm³).

Cytokine and chemokine profiling

Upon necropsy, BALF was collected from NNK-treated *Gprc5a*^{-/-} mice treated with control IgG, anti-IL-1 β , anti-PD-1, or combination of both antibodies, and from either group of mice, i.e., those receiving the antibodies immediately after end of NNK as well as those receiving them starting at 3 months after end of NNK. Briefly, mice were tracheotomized with a luer stub adapter cannula, and BALF was obtained by instilling and collecting 1 mL of PBS. BALF was centrifuged at $1,250 \times g$ for 10 min, and supernatants were collected and stored at -80°C for downstream cytokine and chemokine assays. Assays were conducted on BALF samples using the Bio-Plex Pro Mouse Th17 Cytokine 10-Plex Panel (BioRad, 12010828, for IL21, IL23, and IL33), Bio-Plex Pro Mouse Cytokine 31-Plex Panel (BioRad, 12009159, for CCL3 and CXCL1), and Mouse IL-17A (homodimer) Uncoated ELISA (Invitrogen, 88-7371), following the manufacturer's instructions.

Sequential immunofluorescence (seq-IF) analysis

H&E-stained slides from NNK treated *Gprc5a*^{-/-} mice receiving antibody treatments at 3 months post end of NNK and analyzed at 7 months were evaluated to construct FFPE TMAs. 21 mouse lung tissues were chosen, comprising at least 5 mice per treatment group. A total of 96 1.5 mm ROIs (with lesions, normal areas adjacent to regions, and immune infiltrates) were carefully annotated to include as cores in the TMAs. In total, 4 TMAs were constructed using 3D TMA Grandmaster (Histech), each with 24 cores representing different lesions and at least one mouse from each of the 4 groups. H&E's were obtained for each of the TMAs for careful histopathological examination and confirmation of lesions.

5 μ m unstained FFPE tissue sections from mouse TMAs were baked, deparaffinized, and rehydrated, followed by autofluorescence quenching and antigen retrieval (pH9 EZ AR 2 Elegance, Biogenex) according to in-house optimized protocols for lung tissues. Slides were then loaded onto a COMET stainer according to manufacturer's recommendations. Samples underwent iterative staining with primary antibodies, secondary antibodies, and subsequent imaging, followed by elution of the primary and secondary antibodies, as previously described.⁷⁰ An antibody panel was optimized for mouse lung tissues and is summarized in Table S6 which lists details of the primary antibodies used (vendor, clone, species, dilutions, and acquisition channels) arranged by staining cycle. After staining completion, imaging outputs from each cycle, namely multi-layer OME.TIFF images, were automatically stitched and

aligned. Images were viewed and analyzed using Lunaphore Viewer after elimination of background fluorescence. Macrophage fractions (out of total cells) were estimated by visualizing CD68⁺ cells in each core and TMA.

QUANTIFICATION AND STATISTICAL ANALYSIS

In addition to the algorithms and statistical analyses described above, all other basic statistical analyses were performed in the R statistical environment (v.4.3.1). The Kruskal–Wallis H -test was used to compare variables of interests across 3 or more groups. Wilcoxon rank-sum test or T-test was used to compare other continuous variables such as gene expression levels and signature scores between groups. Spearman's correlation coefficient was calculated to assess associations between two continuous variables. Benjamin–Hochberg method was used to control for multiple hypothesis testing. All statistical tests performed in this study were two-sided. Results were considered significant at p values or FDR q values < 0.05 . When a p value reported by R was smaller than $2.2\text{e-}16$, it was reported as $p < 2.2\text{e-}16$.