



OPEN ACCESS

EDITED BY

Qi Zhao,
Sun Yat-sen University Cancer Center
(SYSUCC), China

REVIEWED BY

Zeyu Chen,
Dana-Farber Cancer Institute,
United States
Ye Qiao Zhou,
University of Pennsylvania, United States

*CORRESPONDENCE

Hongzhu Qu,
quhongzhu@big.ac.cn
Xiangdong Fang,
fangxd@big.ac.cn
Nan Wu,
nanwu@bjmu.edu.cn

[†]These authors have contributed equally
to this work

SPECIALTY SECTION

This article was submitted to Cancer
Cell Biology,
a section of the journal
Frontiers in Cell and Developmental
Biology

RECEIVED 24 April 2022

ACCEPTED 04 July 2022

PUBLISHED 15 August 2022

CITATION

Liu B, Wang C, Fang Z, Bai J, Qian Y,
Ma Y, Ruan X, Yan S, Li S, Wang Y,
Dong B, Yang X, Li M, Xia X, Qu H, Fang X
and Wu N (2022), Single-cell RNA
sequencing reveals the cellular and
molecular changes that contribute to
the progression of
lung adenocarcinoma.
Front. Cell Dev. Biol. 10:927300.
doi: 10.3389/fcell.2022.927300

COPYRIGHT

© 2022 Liu, Wang, Fang, Bai, Qian, Ma,
Ruan, Yan, Li, Wang, Dong, Yang, Li, Xia,
Qu, Fang and Wu. This is an open-
access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

Single-cell RNA sequencing reveals the cellular and molecular changes that contribute to the progression of lung adenocarcinoma

Bing Liu^{1†}, Chen Wang^{2,3†}, Zhanjie Fang^{2,3†}, Jing Bai⁴,
Ying Qian^{2,3}, Yuanyuan Ma¹, Xiuyan Ruan², Shi Yan¹, Shaolei Li¹,
Yaqi Wang¹, Bin Dong⁵, Xin Yang⁶, Meng Li², Xuefeng Xia⁴,
Hongzhu Qu^{2,3*}, Xiangdong Fang^{2,3*} and Nan Wu^{1*}

¹Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Department of Thoracic Surgery II, Peking University Cancer Hospital & Institute, Beijing, China, ²CAS Key Laboratory of Genome Sciences and Information, Beijing Key Laboratory of Genome and Precision Medicine Technologies, Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bioinformatics, Beijing, China, ³University of Chinese Academy of Sciences, Beijing, China, ⁴Geneplus-Beijing Institution, Peking University Medical Industrial Park, Zhongguancun Life Science Park, Beijing, China, ⁵Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Central Laboratory, Peking University Cancer Hospital and Institute, Beijing, China, ⁶Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Department of Pathology, Peking University Cancer Hospital and Institute, Beijing, China

Pure ground glass nodules (GGNs) and solid nodules (SNs) represent early and relatively late stages of lung adenocarcinoma (LUAD) in radiology, respectively. The cellular and molecular characteristics of pure GGNs and SNs have not been comprehensively elucidated. Additionally, the mechanism driving the progression of lung adenocarcinoma from pure GGN to SN in radiology is also elusive. In this study, by analyzing the single-cell transcriptomic profiles of 76,762 cells from four pure GGNs, four SNs, and four normal tissues, we found that anti-tumor immunity mediated by NK and CD8+T cells gradually weakened with the progression of LUAD and humoral immunity mediated by plasma B cells was more active in SNs. Additionally, the proliferation ability of some special epithelial cell increased during the progression process from pure GGN to SN. Furthermore, stromal cells and M2 macrophages could assist the progression of LUAD. Through comprehensive analyses, we revealed dynamic changes in cellular components and intercellular interactions during the progression of LUAD. These findings could facilitate our understanding of LUAD and discovery of novel therapeutic targets.

KEYWORDS

single-cell RNA sequencing, molecular, cellular, progression, lung adenocarcinoma

Introduction

Lung cancer is the leading cause of cancer-related mortality, among which lung adenocarcinoma (LUAD) represents the major histological subtype, both globally and in China (Chen et al., 2016; Siegel et al., 2021). The disease spectrum has gradually shifted with population-based lung cancer screening. More and more early-stage LUADs manifesting as ground glass nodule (GGN) in radiology have been detected (Gao et al., 2017), and they are further classified into pure GGNs and mixed GGNs, which represent a special entity characterized by slow growth and indolent non-invasive features (Chang et al., 2013). Previous studies have demonstrated that pure GGNs could develop into mixed GGNs and then solid nodules (SNs) over time (Silva et al., 2012; Cho et al., 2016). Patients with GGN can harbor excellent survival by surgical resection, and some GGNs do not affect the lifespan of patients under regular surveillance. However, LUADs manifesting as SN are often more aggressive and exhibit a higher recurrence rate after resection (Zhang et al., 2020). Additionally, although multiple treatment modalities, including surgical resection, chemotherapy, targeted therapy, and immunotherapy, have significantly improved the clinical outcome of patients with LUAD, treatment effect varies considerably ranging from permanent cure to rapid progression (Hirsch et al., 2017; Billan et al., 2020). That is predominantly attributed to the limited understanding of the underlying mechanisms of tumor progression and inter-tumoral and intra-tumoral cellular and molecular heterogeneity of LUAD (Hua et al., 2020; Yuan et al., 2021).

Conventional bulk sequencing has uncovered abundant molecular aberrations that drive carcinogenesis and the progression of LUAD. For instance, Chen et al. revealed that the frequency of TP53, arm-level copy number alterations, and HLA loss of heterozygosity increased during the progression from pre-invasive lesion to invasive LUAD (Chen et al., 2019). However, LUAD is composed of various types of cells, including epithelial cells, stromal cells, and immune cells, and they constitute a complex ecosystem (Lambrechts et al., 2018). Due to methodological limitations of bulk sequencing, important molecular alterations in cancer cells may be undervalued. Additionally, rare but functionally important cell subpopulations (e.g., cancer stem cells and tumor-infiltrated immune cells) may not be identified. In order to comprehensively decipher the underlying molecular mechanism and reveal the heterogeneity of LUAD, particularly at different development stages, more advanced techniques, such as single-cell RNA sequencing (scRNA-seq), are required.

Recent advances in scRNA-seq enable exploring the microenvironment and intra-tumoral heterogeneity of human solid tumors and allow for assessment of heterogeneous cell populations at single-cell resolution (Lei et al., 2021). Notably, scRNA-seq has also been employed to dissect the molecular

characteristics of lung cancer, including LUAD (Lavin et al., 2017; Guo et al., 2018). A series of pioneering studies comprehensively mapped the immune landscape of LUAD, and revealed unprecedented intra-tumoral heterogeneity. This provided deeper insights into the biology of LUAD and guided immunotherapy design (Lavin et al., 2017). One recent study of advanced metastatic LUAD revealed the molecular and cellular reprogramming during metastasis, which provided diagnostic and therapeutic targets (Kim et al., 2020). Another study provided single-cell transcriptomic profiling of subsolid nodules and facilitated the understanding of their indolent nature (Xing et al., 2021). These studies have decoded comprehensive cellular and molecular characteristics of special stages of LUAD or individual cell types, such as T cells. Nevertheless, it is still ambiguous about the dynamic evolutionary procedure of LUAD from pure GGNs to SNs, especially at single-cell resolution. A comprehensive understanding of this procedure is of significance to develop reasonable and effective treatment strategies for patients at different stages of LUAD.

In the present study, we performed scRNA-seq on fresh surgical specimens from eight resectable LUADs, among which four were presented as pure GGNs and the other four were SNs on computed tomography (CT) image. They represent different development phases of LUAD in radiology. By comprehensively analyzing their cellular and molecular profiles, we revealed dynamic changes in cellular components and intercellular interactions during the evolution of LUAD. We found that anti-tumor immunity mediated by NK and CD8⁺ T cells gradually weakened and humoral immunity mediated by plasma B cells was more active in late stages. These findings may facilitate our understanding of LUAD and the discovery of novel therapeutic targets.

Materials and methods

Patients and specimens

Eight treatment-naïve patients with early-stage LUAD in radiology who underwent radical resection between December 2018 and November 2019 at the Department of Thoracic Surgery II, Peking University Cancer Hospital & Institute were enrolled in this study. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and approved by the Ethics Committee of Peking University Cancer Hospital and Institute (Institutional Review Board No. 2019KT59). All patients provided written informed consent before surgery. Demographic and clinicopathological characteristics of these patients are listed in [Supplementary Table S1](#).

Tumor tissues were collected from four patients with pure GGN (WDF, WST, JYS, and YL) and four patients with SN (HYP, LHY, QDL, and CGC). Additionally, non-cancerous

adjacent normal tissues were collected from one patient with SN (LHY), one with pure GGN (YL) and two other patients (YLF and LY). Fresh tumor or non-cancerous adjacent normal tissues were collected immediately after a qualified pathologist confirmed the pathological diagnosis of these nodules *via* fast frozen pathology during surgery. Non-cancerous adjacent normal tissues were usually collected at the edge of the resected lobe and were macroscopically normal. In the cases of sublobar resection, they were collected > 2 cm apart from the tumor edge.

To validate the results of scRNA-seq, formalin-fixed paraffin-embedded (FFPE) tumor tissues from another cohort comprising fifteen patients with pure GGN and fifteen with SN who also underwent surgery at the Department of Thoracic Surgery II, Peking University Cancer Hospital & Institute were evaluated for the expression of selected markers at protein level using immunohistochemistry (IHC), and their brief clinicopathological characteristics of them are listed in [Supplementary Table S2](#) (validation cohort). All patients provided written informed consent before surgery.

Tissue dissociation and sequencing

Fresh tissues were preserved in tissue storage solution (Miltenyi Biotech, Cat. No. 130-100-008) and were transported on ice to preserve viability, and all tissues were dissociated within 2 h after surgery. After being washed 3 times with phosphate buffered saline (PBS; Gibco, Cat. No. 8120002), tissues were cut into small pieces of 1–2 mm³ in size and then transferred to a gentleMACS C tube containing 5 ml enzyme mix (Miltenyi Biotech, Cat. No. 130-095-929). The gentleMACS programs h_tumor_01, h_tumor_02, and h_tumor_02 were run with two 30-min incubations on MACSmix™ Tube Rotator at 37 °C. Then, the samples were centrifuged, and the resuspended samples were filtered to a 50 ml tube through MACS SmartStrainer (70 μm). The filtered suspension was centrifuged at 300 g for 7 min, and the supernatant was completely aspirated. After that, red blood cells were removed by RBC lysis buffer (BD, Cat. No. 349202), and cell suspension was centrifuged at 300 g for 7 min again. The cell pellet was used for the next experiments. Cell suspension was processed to prepare scRNA-seq libraries by Chromium Single cell 3' Reagent V2 kits with a cell recovery target of 8,000, following the manufacturer's instructions (10× Genomics, Pleasanton, CA, United States). Libraries were sequenced on an Illumina Xten.

Raw data processing and quality control

Raw sequencing data were mapped to the GRCh38 human reference genome and quantified at gene level to produce gene

expression matrix cross-cells with Cell Ranger toolkit (v3.1.0). The expression matrixes were first converted to Seurat object with Seurat (v4.0.3). Next, cells with less than 200 genes or more than 10% mitochondrial-derived unique molecular identifiers (UMI) were considered as low-quality cells and filtered out. Additionally, those with more than 5,000 gene expressions were also excluded, which were thought to be doublets.

Unsupervised dimensional reduction and clustering

The gene expression matrix was normalized by the `NormalizeData` function. The cell cycle score was obtained by `CellCycleScoring` function and then the difference between G2M and S-phase score was eliminated. Next, `ElbowPlot` function was employed to identify the true dimensionality, and the top 30 principal components (PCs) were involved in `RunPCA` function. Then, these cells were clustered by `FindNeighbors` and `FindClusters` function with resolution set as default. Finally, cell clusters were displayed as Uniform Manifold Approximation and Projection (UMAP) plot with `DimPlot` function. `FindAllMarkers` function was used to find the highly expressed or differentially expressed genes. According to the canonical marker genes of each cluster, these cell clusters were annotated.

Estimation of CNV

The R package “inferCNV” (v1.8.1) was used to infer the CNV of each cell type, and the cells derived from stromal cells and immune cells were used as normal controls. The R code was provided in <https://github.com/broadinstitute/inferCNV> with the default parameters except for cut-off = 0.1.

Inference of cell state through trajectory analysis

The Monocle2 (v2.20.0) package was used to analyze single-cell trajectories to discover the cell-state transitions. The Seurat object was first converted to the monocle object, and then the significantly changed genes calculated by `differentialGeneTest` function were used to determine the differential cell states. The state where most normal cells clustered was considered as the root state and used to specify the start of the pseudotime trajectory. The position of cells on the trajectory was determined by `orderCells` function. The visualization of cell differentiation and gene expression changes with pseudotime was performed by function `plot_cell_trajectory`, `plot_genes_in_pseudotime` and `plot_pseudotime_heatmap`, respectively.

Cell–cell interaction network

The interaction network between cells was analyzed by using CellPoneDB2 (<https://www.cellponedb.org/>). The gene expression matrix of single cells and the corresponding cell type information were used as input files to infer the potential interaction strength between every cell type, with interaction set as default (interactions = 1,000). The interactions between cells were evaluated by means and *p*-value from the output file ($p \leq 0.05$).

Regulons network analysis

SCENIC (v1.2.4) was used to analyze the regulons networks in our data based on single-cell RNA datasets and produce the expressed regulons of each cell. Next, upregulated or downregulated regulons of different clusters were calculated by limma package (v3.48.3) with adj. *p*-value < 0.05 and then were used for subsequent analysis.

Public data analysis

RNA-seq data and clinical information of LUAD from the Cancer Genome Atlas (TCGA) database were downloaded from UCSC Xena (<https://xena.ucsc.edu/>). The average expression differences between tumor and normal tissues in TCGA of special marker genes identified from our samples were compared by *t*-test. Additionally, the non-negative matrix factorization (NMF) model was used to reduce dimensionality of expression matrix of marker genes among TCGA LUAD cohort, and survival curve was constructed by the R package “Survminer” (0.4.9) and “Survival” (3.2–13) to assess the relationship between special marker genes expression and prognosis. The average expression differences between two gene expression matrixes generated by NMF was compared by *t*-test.

Pathway analysis

The marker genes of each cell subsets and the differentially expressed genes of tissues of different origin were, respectively, calculated through FindAllmarkers function ($p \leq 0.05$, $\log_{2}FC \geq 0.25$) and limma package. The functional enrichment was performed on GO enrichment. Gene set variation analysis (GSVA) was used to assess differential pathways, and the activity scores of special cells were also calculated by GSVA (v1.40.1).

Inferring malignant cells

The CNV expression matrix of epithelial cells of each sample was obtained by R package inferCNV. The CNV score of each cell

was calculated by the mean square of CNV obtained by “inferCNV.” The cells were sorted according to the CNV score, and the top 5% cells were regarded as the reference cells. The average CNV value of top 5% cells was calculated firstly, and then the correlation coefficient of CNV scores between each cell and average CNV was calculated as described above. The cells with the correlation coefficient over than 0.3 were regarded as malignant cells.

Immunohistochemistry

Serial 4 μm sections of FFPE samples were cut onto glass slides and subjected to IHC staining following standard protocol. Firstly, the slides were deparaffinized by using xylenes and ethanol gradient. Antigen retrieval was performed in citric acid buffer in a 95°C water bath for 20 min, and then the slides were incubated in 3% H_2O_2 for 25 min, then added with 3% BSA for 30 min at room temperature to block the non-specific binding sites. After that, the slides were incubated with primary antibodies followed by HRP-linked secondary antibodies and diaminobenzidine (DAB, Servicebio, Cat. No. G1211) staining. After counterstaining with hematoxylin, slides were dehydrated with sequential ethanol washes.

The following antibodies were used to detect the respective proteins: anti-RGS5 (rabbit, 1:500, ab196799, ABCAM), anti-Ki67 (mouse, 1:500, #9449, CST), anti-HLA-DR (rabbit, 1:500, ab92511, ABCAM), anti-TOP2A (mouse, 1:500, #12286, CST), anti-PTGDS (rabbit, 1:200, PA018969HA01HU, Cusabio Technology), and anti-Syndecan-1 (CD138, rabbit, 1:500, ab128936, ABCAM).

The intensity of RGS5, TOP2A, PTGDS, and Sydecan-1 was scored as 0 (no staining), 1 (weak), 2 (moderate), and 3 (strong) by experienced pathologists. The Ki-67 expression level was estimated by counting the ratio of positive tumor cells in representative regions. As for HLA-DR, the slides were scanned by whole slide digital scanning using a digital pathology scanner (Aperio VERSA, Leica Biosystems, Buffalo Grove, IL, United States), and scoring was assessed on an Aperio Scanscope (Aperio Technologies, United States). The average positive cell ratio was measured in 8 intra-tumoral non-overlapping fields using fixed areas of 0.078 square millimeters.

Results

Single-cell transcriptomic profiling of pure ground glass nodules and solid nodules

To elucidate the cellular and molecular characteristics of different developmental phases of LUAD, tumor tissues from four pure GGNs and four SNs were prospectively collected for

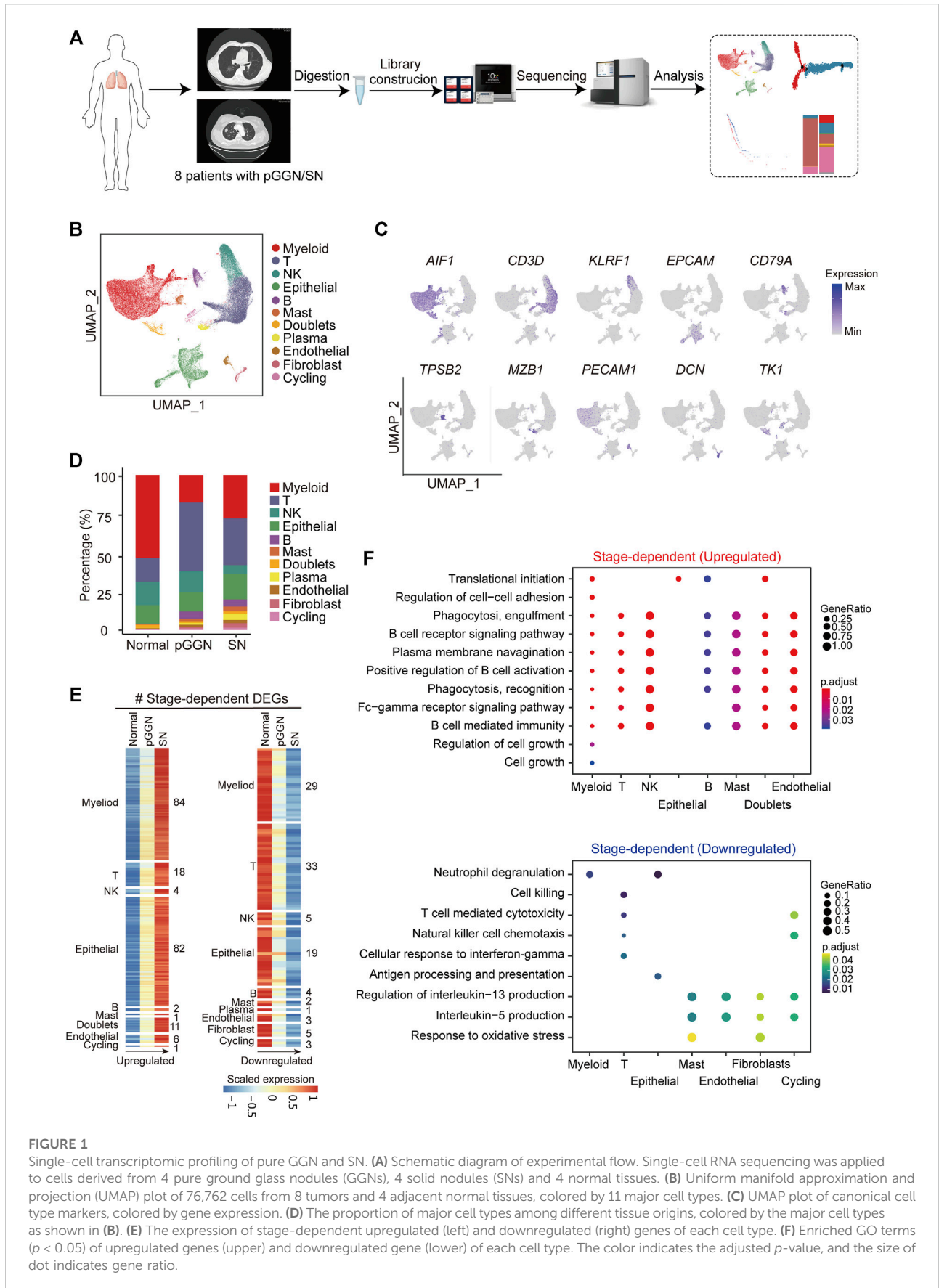


FIGURE 1

Single-cell transcriptomic profiling of pure GGN and SN. **(A)** Schematic diagram of experimental flow. Single-cell RNA sequencing was applied to cells derived from 4 pure ground glass nodules (GGNs), 4 solid nodules (SNs) and 4 normal tissues. **(B)** Uniform manifold approximation and projection (UMAP) plot of 76,762 cells from 8 tumors and 4 adjacent normal tissues, colored by 11 major cell types. **(C)** UMAP plot of canonical cell type markers, colored by gene expression. **(D)** The proportion of major cell types among different tissue origins, colored by the major cell types as shown in **(B)**. **(E)** The expression of stage-dependent upregulated (left) and downregulated (right) genes of each cell type. **(F)** Enriched GO terms ($p < 0.05$) of upregulated genes (upper) and downregulated gene (lower) of each cell type. The color indicates the adjusted p -value, and the size of dot indicates gene ratio.

scRNA-seq in this study. Additionally, non-cancerous adjacent normal tissues from one patient with pure GGN, one with SN, and two other patients were also obtained as controls (Supplementary Table S1). The CT images of each tumor are shown in Supplementary Figure S1A. Single-cell solutions and DNA libraries of each tissue were carefully constructed using 10X Genomics and were then sequenced. The single-cell transcriptomic profiles of normal tissues, pure GGNs, and SNs were characterized and compared using comprehensive bioinformatic analyses (Figure 1A).

A total of 76,762 high-quality cells were obtained, including 22,053 cells from pure GGNs, 21,306 cells from SNs, and 33,403 cells from normal tissues. The correlation coefficient between samples calculated in the same group was high (≥ 0.89), suggesting that the batch effect among samples was relatively small (Supplementary Figure S1). According to the expression of canonical marker genes, all of these cells were assigned to 10 well-known cell types (Figure 1B): myeloid cells (*LYZ+*, *CD163+*, and *CD68+*), T cells (*CD3D+* and *TRAC+*), NK cells (*NGK7+*, *GNLY+*, and *KLRF1+*), epithelial cells (*EPCAM+*, *KRT19+*, and *KRT18+*), B cells (*CD79A+*), mast cells (*MS4A2+*, *TPSB2+*, and *GATA2+*), plasma B cells (*MZB1* and *IGHA1+*), endothelial cells (*PECAM1+* and *FLT1+*), fibroblasts (*COL1A2+*, *COL6A2+*, and *DCN+*), and cycling cells (*TK1+*, *MKI67+*, and *TOP2A+*) (Figure 1C; Supplementary Figure S1C). It was observed that the proportion of plasma B cells, B cells, fibroblasts, and endothelial cells increased sequentially from normal tissues to pure GGNs and then to SNs, while that of NK cells decreased (Figure 1D). The proportion of T cells was relatively higher and that of myeloid cells was lower in pure GGNs than in SNs. When we classified immune cells into innate immunity and adaptive immunity, the adaptive immune response was significantly activated, whereas the innate immunity was significantly inactivated in both pure GGNs and SNs compared with that of normal tissues (Supplementary Figure S1D), indicating that innate immune response mainly mediated by NK cells might gradually weaken during the progression of LUAD. However, adaptive immune response was more highly activated in pure GGNs than in SNs, suggesting that adaptive immune response to tumors was relatively stronger in the earlier stage of LUAD development (Supplementary Figure S1D).

Progression of LUAD was accompanied by transcriptome changes. In this study, a total of 1,113 upregulated genes and 1,240 downregulated genes ($|\text{avg_logFC}| \geq 0.25$ and $p_{\text{val_adj}} \leq 0.05$) were identified during the progression from pure GGNs to SNs (Supplementary Figure S1E). We compared differentially expressed genes (DEGs) among the total cells from three types of tissues and obtained 34 upregulated and 29 downregulated genes related to the progression of LUAD (Supplementary Figure S1F). Furthermore, we comprehensively characterized DEGs according to distinct cell types with the progression of LUAD, namely, stage-dependent DEGs, and

identified 209 upregulated DEGs in 9 cell types and 104 downregulated in 10 cell types (Figure 1E). Gene ontology (GO) enrichment analysis revealed that the upregulated stage-dependent DEGs were mainly related to cell growth, regulation of cell–cell adhesion, B cell receptor signaling pathway, phagocytosis and other pathways, while the downregulated stage-dependent DEGs were mainly related to neutrophil degranulation, cell killing response, T cell-mediated cytotoxicity, and response to oxidative stress (Figure 1F).

Proliferation ability of special epithelial cells increases during the progression of lung adenocarcinoma

A total of 3,969 epithelial cells from normal tissues were obtained, and based on canonical marker genes, these epithelial cells were subclustered into 5 subsets, including ciliated cells (*TPPP3+*), club cells (*SGCB1A1+*), alveolar cells (*SFTPA1+* and *AGER+*), alveolar type II cells (*AT2*; *SFTPA1+*), and alveolar type I cells (*AT1*; *AGER+*) (Figures 2A,B). Among them, alveolar cells were the mixture of AT1 and AT2, expressing *SFTPA1* and *AGER* simultaneously. The distribution of these cells in each sample is shown in Figure 2C. The 6,215 epithelial cells from tumor tissues were divided into malignant and nonmalignant cells based on the CNV expression matrix by inferring CNV (Supplementary Figure S2A). Obviously, the proportion of malignant cells in SNs was higher than that in pure GGNs (Supplementary Figure S2B,C). Totally, 756 genes were upregulated and 487 genes were downregulated in malignant cells compared with nonmalignant cells ($|\text{avg_logFC}| \geq 0.25$ and $p_{\text{val_adj}} \leq 0.05$) (Supplementary Figure S2D). GO enrichment analysis showed that the response to hypoxia, epithelial tube morphogenesis, and regulation of protein serine/threonine kinase activity pathways were upregulated in malignant cells. Notably, the regulation of humoral immune response, neutrophil-mediated immunity, and regulation of T-cell activation pathways were activated in nonmalignant cells (Supplementary Figure S2E; Supplementary Table S3). Furthermore, gene set variation analysis (GSVA) inferred that the function of the malignant cells from tumors, was different from that of epithelial cells from normal tissues and nonmalignant cells from tumors (Figure 2D). The G2M checkpoint, TGF beta signaling, and glycolysis were upregulated in malignant cells from SNs. Spermatogenesis and KRAS signaling pathways were upregulated in epithelial cells from normal and nonmalignant cells from tumors, suggesting that the proliferation ability of malignant cells from SNs was enhanced.

To further assess malignant cells, they were re-clustered into six cell subsets (Figure 2E). Among them, the proportion of C1 and C6 increased from pure GGNs to SNs, while those of C2, C3, C4, and C5 decreased (Figure 2F). The marker genes of

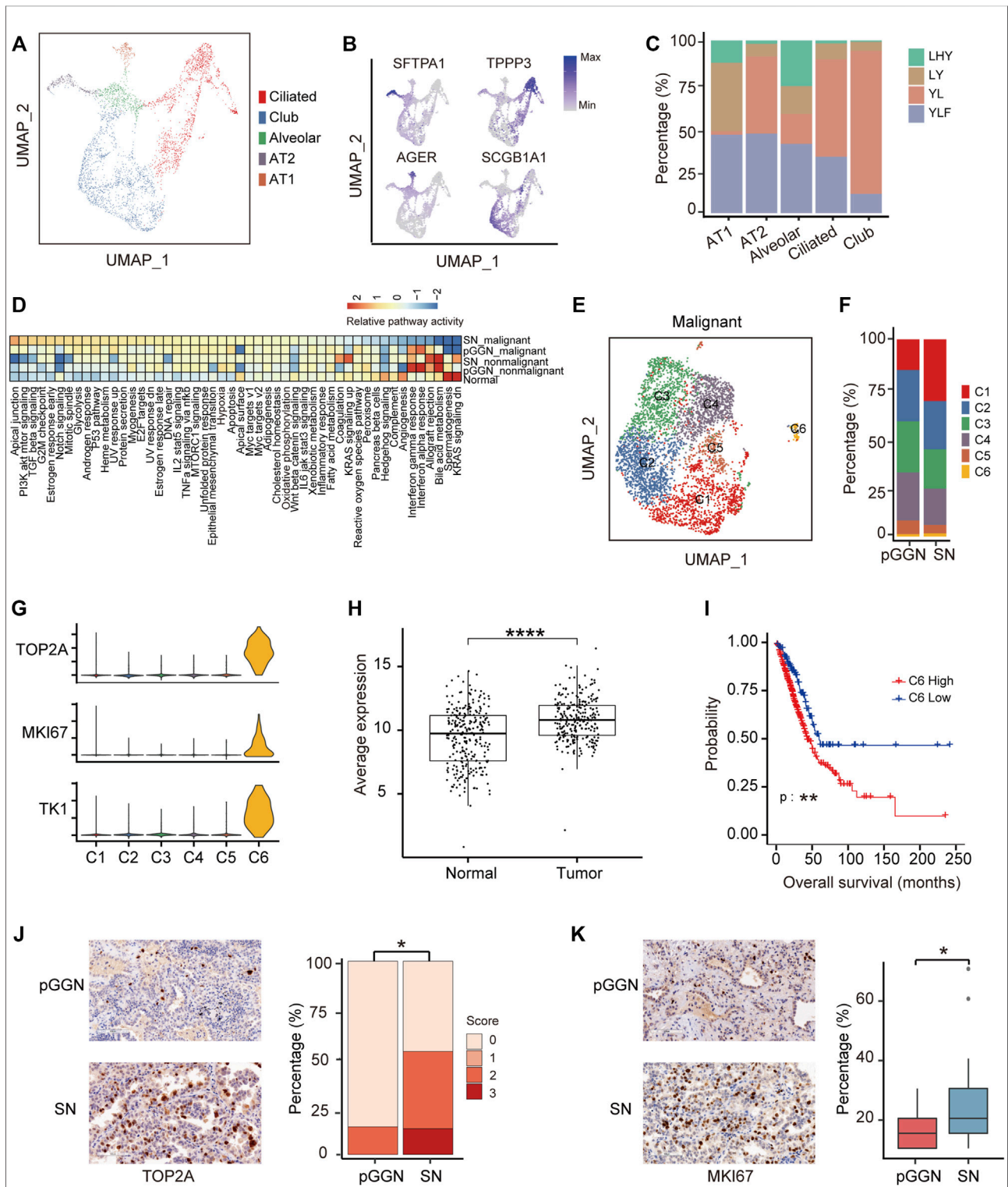


FIGURE 2

Proliferation ability of special epithelial cells increases during the progression of LUAD. (A) UMAP plot of epithelial cells from 4 normal adjacent tissues, colored by 5 cell subtypes. (B) UMAP plot of canonical cell type markers, colored by gene expression. Four highly expressed genes were used to define 5 cell types, of which one cell type is a mixture. (C) The proportion of normal samples among each cell type, colored by patient origins. (D) The enrichment of hallmark pathways by GSVA analysis among normal epithelial cells, nonmalignant cells, and malignant cells. The scores were sorted by relative pathway activity. (E) UMAP plot of epithelial malignant cells, colored by 6 cell subsets. Epithelial cells from tumor tissues were divided into malignant cells and nonmalignant cells by inferCNV. (F) The proportion of cell types among each tissue origin, colored by cell subtypes. (G) The

(Continued)

FIGURE 2

expression of proliferative markers (*TOP2A*, *MKI67*, and *TKI*) in epithelial malignant cells subsets. **(H)** The average expression of marker genes of C6 in TCGA LUAD cohort. It was significantly higher in tumor tissue than in normal tissue. **(I)** The survival curves of two groups of TCGA LUAD patients with high and low expression of C6 cluster marker genes. The *p*-value was calculated by two-sided log-rank test. **(J)** The expression of *TOP2A* in pure GGNs and SNs by immunohistochemical examination, and the left was the representative images. Score 0, 1, 2, and 3 represents no, weak, moderate, and strong staining, indicating the relative expression level of *TOP2A* from low to high. *TOP2A* is highly expressed in SNs than in pure GGNs. **(K)** The expression of *MKI67* in pure GGNs and SNs by immunohistochemical examination, and the left was the representative images. The ratio of positive expressed tumor cells in pure GGNs and SNs was compared. The comparison in **(H,J)** and **(K)** was performed using t-test. *, *p* ≤ 0.05; **, *p* ≤ 0.001; ***, *p* ≤ 0.0001. A *p*-value < 0.05 was considered statistically significant.

C1–C5 were involved in similar functions, such as positive regulation of cell–cell adhesion and T-cell activation, while marker genes of C6 were mainly related to the cell cycle G2/M phase transition and positive regulation of the cell cycle (Supplementary Figure S2F). The expression of marker genes of C6 cluster cells, including a total of 276 genes such as *TOP2A*, *MKI67*, and *TKI* (Figure 2G), was higher in tumor tissues than in normal tissues from TCGA database (Figure 2H). In addition, patients with LUAD from TCGA with higher expression of these genes in tumor tissues had a shorter overall survival compared with those patients with lower expression (Figure 2I; Supplementary Figure S2G), suggesting that these genes could serve as promising prognostic biomarkers. Furthermore, the expression of *TOP2A* and *MKI67* was also higher in the SN group than in the pure GGN group at protein level as estimated by immunohistochemistry (IHC) in the validation cohort (Figures 2J,K).

Additionally, we extracted epithelial cells from tumor tissues to infer the evolutionary trajectory and found that nonmalignant cells were located in the early pseudo time (Supplementary Figure S2H). The heatmap of gene changes accompanied by pseudo time showed that *TPPP3*, *SCGB3A1*, and MHC family genes such as *HLA-DRB5* and *HLA-DRB1* were upregulated in the early pseudo time, and that *SOX4* and *SFTPA1* were upregulated in the end of the pseudo time (Supplementary Figure S2I). These indicated that these genes might be involved in the malignant transformation of epithelial cells.

Stromal cells assist the progression of lung adenocarcinoma

A total of 923 fibroblasts were divided into five subsets (Figure 3A; Supplementary Figure S3A), including inflammatory cancer-associated fibroblasts (Fibro-C1, also known as iCAFs; *PTGDS*+) secreting cytokines such as *CXCL12*, *CCL2*, and *CXCL14* (Supplementary Figure S3B), pericytes (Fibro-C2; *RGS5*+ and *PDGFRB*+), myofibroblasts (Fibro-C3; *ACTA2*+), antigen-presenting CAFs (Fibro-C4; *CXCR4*+ and *HLA-DRA*+), and other fibroblasts (Fibro-C5; *COL1A1*+ and *MMP2*+). The abundance of pericytes (Fibro-C2) increased sequentially from normal tissues to pure GGNs, and then to SNs, while that of iCAFs (Fibro-C1) decreased

(Figure 3B). The proportion of Fibro-C4 was the highest in pure GGNs, and that of Fibro-C3 was the lowest in pure GGNs. Notably, Fibro-C5 existed only in tumor tissues (Figure 3B). GSVA analysis revealed that metabolic pathways, such as glycosaminoglycan biosynthesis and glycolysis were upregulated in fibroblasts derived from SNs (Figure 3C).

Previous studies have demonstrated that iCAFs promote anti-tumor immunity (Biffi and Tuveson, 2021; 2019). Hence, the decrease of iCAFs (Fibro-C1) implied that the anti-tumor immune response might gradually decrease with tumor progression. (Figure 3B). Additionally, higher expression of marker genes of iCAFs was correlated with longer overall survival in the TCGA LUAD cohort (Supplementary Figure S3C; Figure 3D). Previous studies have reported that pericytes (Fibro-C2) significantly contribute to cancer invasion and metastasis by transitioning to fibroblasts (Hosaka et al., 2016; Zeltz et al., 2020). The proportion of Fibro-C2 increased sequentially from normal tissues to pure GGNs, and then to SNs, suggesting that the ability of cancer invasion and metastasis mediated by Fibro-C2 gradually increased with tumor progression, and targeting Fibro-C2 might offer a new treatment option to inhibit cancer metastasis. We detected the expression of *PTGDS* and *RGS5*, the marker genes of Fibro-C1 and Fibro-C2, respectively, by IHC in the validation cohort, and found that *PTGDS* was highly expressed in pure GGNs and the expression of *RGS5* was lower in pure GGNs than that in SNs (Figure 3E), which was consistent with the results of scRNA-seq. Fibro-C4 expressed high levels of MHC family genes, such as *HLA-DRA*, indicating the features of antigen presentation (Supplementary Figure S3A). The high proportion of Fibro-C4 in pure GGNs indicated that the antigen presentation process was more active in the initial stage of LUAD.

Except for fibroblasts, a total of 937 endothelial cells were also obtained, which were subclustered into five subsets (Figure 3F; Supplementary Figure S3D), including extra-alveolar capillary ECs (Endo-C1; *EDN1*+), tumor ECs (Endo-C3; *INSR*+), and other ECs (Endo-C2, *ICAM1*+; Endo-C4, *MAP1B*+; Endo-C5, *CXCR4*+). Specifically, the proportion of Endo-C1 and Endo-C4 decreased gradually, whereas that of Endo-C2 increased gradually with tumor progression. Endo-C5 was the least abundant cells in pure GGNs (Figure 3G). GO analysis showed that Endo-C1 exhibited the characteristics of endothelium development, and pathways such as extracellular

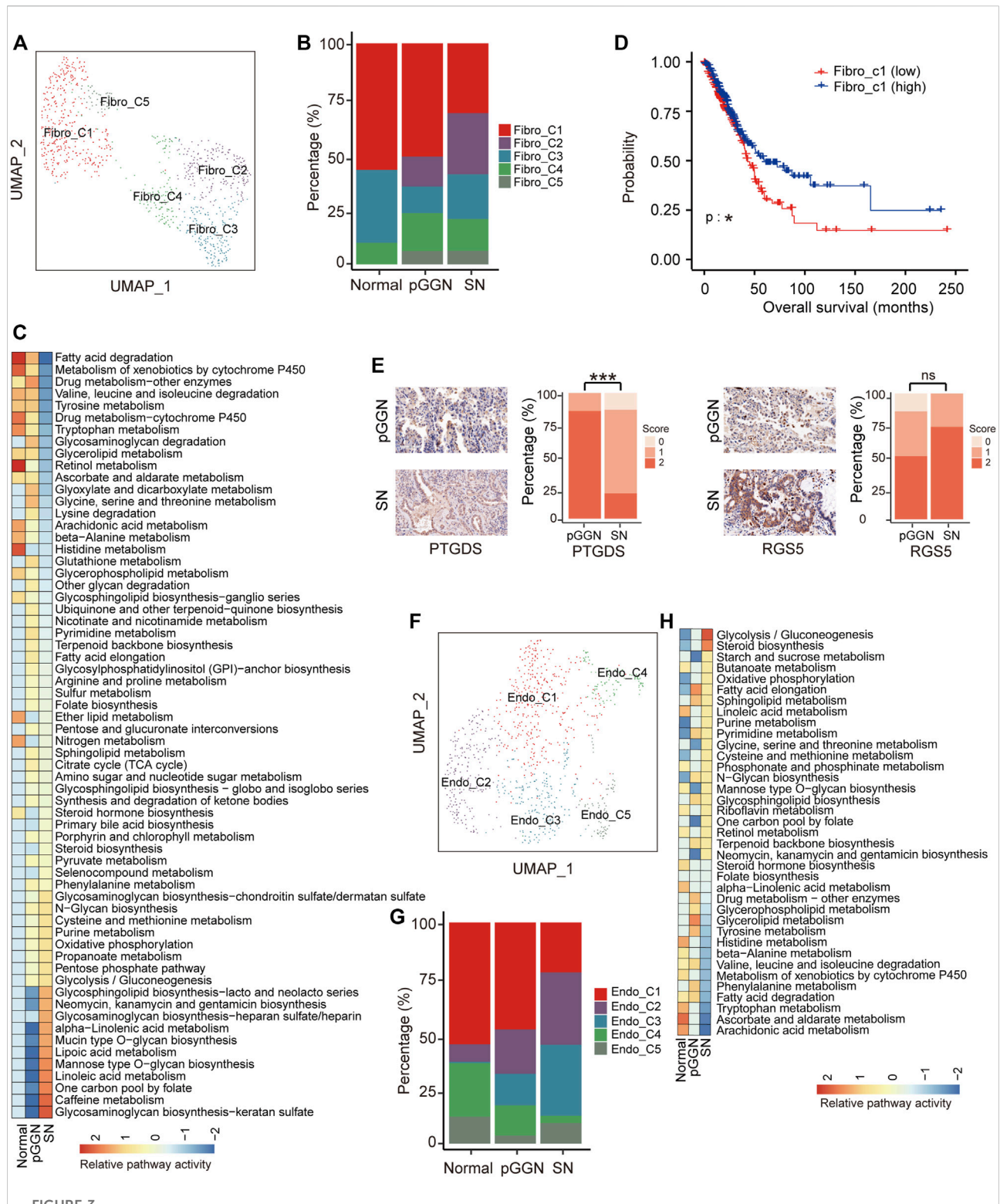
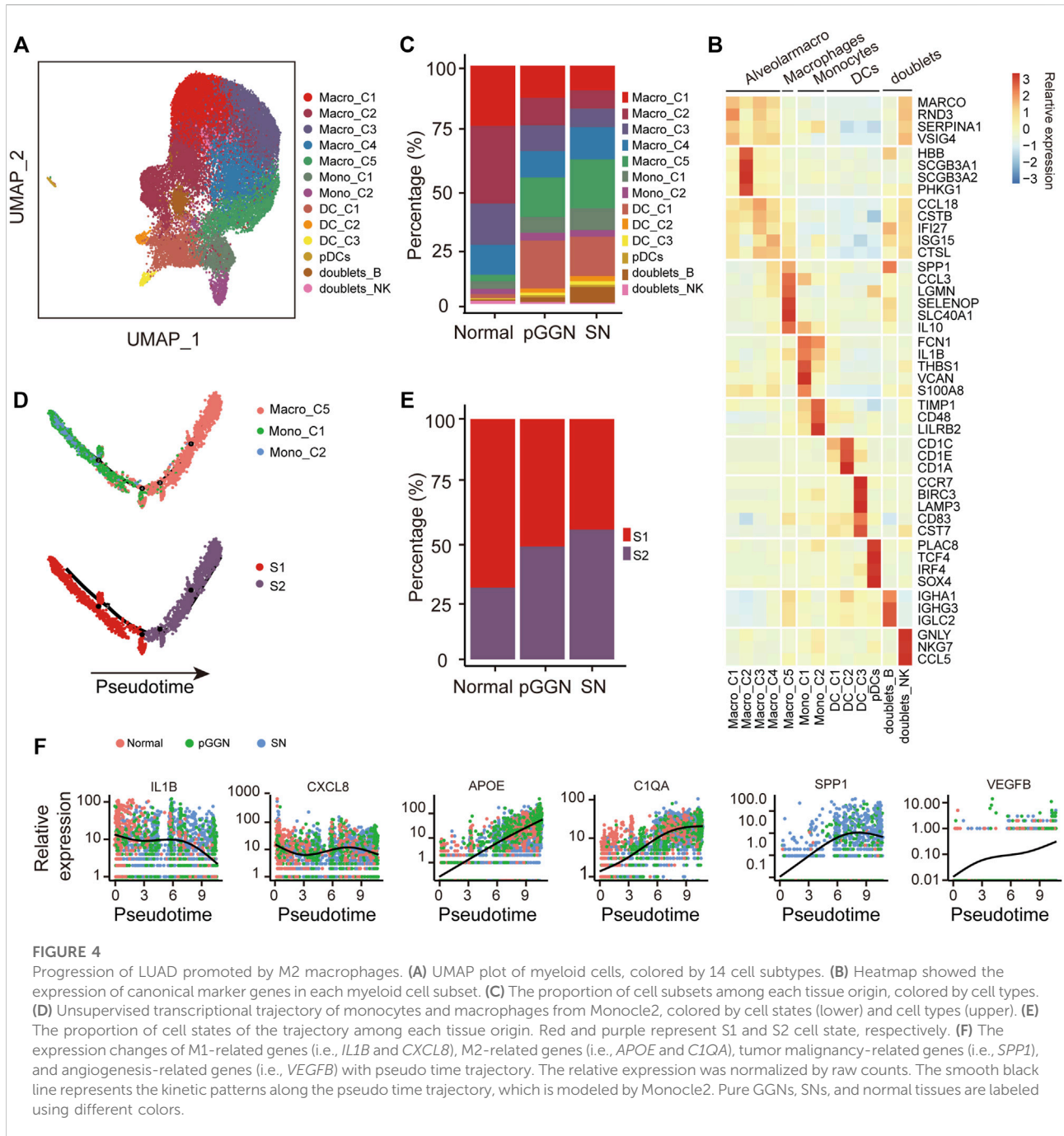


FIGURE 3

Stromal cells assist progression of tumor during the progression of LUAD. (A) UMAP plot of fibroblasts, colored by 5 cell subtypes. (B) The proportion of major cell subsets among each tissue origin. (C) Heatmap showing the activity of metabolic pathways for fibroblasts, among pure GGNs, SNs, and normal tissues. Many metabolic pathways were upregulated in fibroblasts derived from SNs. (D) The survival curves of two groups of TCGA LUAD patients with high and low expression of Fibro-C1 cluster marker genes. The p -value was calculated by two-sided log-rank test. (E) The expression of PTGDS and RGS5 in the pure GGNs and SN tissues by immunohistochemical examination. The comparison was performed using t -test. (F) UMAP plot of endothelial cells, colored by 5 major cell types. (G) The proportion of major cell subsets among each tissue origin. (H) The enriched metabolic pathways by GSVA analysis in endothelial cells from different tissue origin. The scores were sorted by pathway activity in malignant cells from SN. ns, $p > 0.05$; *, $p \leq 0.05$; **, $p \leq 0.001$; ***, $p \leq 0.0001$. A p -value < 0.05 was considered statistically significant.



structure organization were upregulated in Endo-C4. Notably, negative regulation of the immune system process and regulation of vasculature development were upregulated in Endo-C2 and Endo-C3, respectively (Supplementary Figure S3E). In addition, a comparison of the metabolic pathways by GSVA analysis between endothelium from normal tissues, pure GGNs and SNs revealed that signaling pathways that could assist tumor growth, such as glycolysis, were enriched in the endothelium in SNs (Figure 3H).

Progression of lung adenocarcinoma is promoted by M2 macrophages

Myeloid cells play a critical role in maintaining tissue homeostasis and regulating inflammation in the lung (Bissonnette et al., 2020; Hou et al., 2021). Additionally, previous studies also confirmed that they had impacts on the progression of LUAD (Lavin et al., 2017; Zheng et al., 2020). A total of 27,758 myeloid cells were detected in this study (Figures

4A,B), and they were subclustered into 13 subsets based on canonical marker genes, including alveolar macrophages (AM; *MARCO*+; including 4 subsets), macrophages (Macro-C5; *LGMN*+ and *IL10*+), monocytes (*FCN1*+ and *IL1B*+; including 2 subsets), dendritic cells (DC; *CD1C*+; including 3 subsets), pDCs (*TCF4*+), and doublets (including 2 subsets). Among them, the doublets-B subset expressed both B cell markers and myeloid markers and the doublets-NK subset expressed both myeloid markers and NK cell markers. Notably, we found that AMs were mainly derived from normal tissues, whereas the abundance of Macro-C5 increased from normal tissues to pure GGNs, and then to SNs (Figure 4C).

Numerous studies have demonstrated that macrophages are involved in tumorigenesis and drug resistance (Schmall et al., 2015; Wang et al., 2017). Monocytes and macrophages are both the members of the mononuclear phagocyte system, and monocytes are the precursors of macrophages (Ginhoux and Jung, 2014). In order to understand the relationship between monocytes and macrophages in the progression of LUAD, an unsupervised trajectory analysis was performed through the Monocle2 package to assess transcriptional changes from monocytes to macrophages. We found that this trajectory evolved from monocytes to macrophages, and it consisted of two transcriptional states (S1 and S2, Figure 4D). S1 was observed to be distinctly positioned at the first half of the trajectory, while S2 was located at the second half of the trajectory. Intriguingly, the cells of S1 gradually decreased from normal tissues to pure GGNs, and then to SNs, whereas those of S2 gradually increased and were significantly enriched in SNs (Figure 4E). These results suggest that monocytes differentiated into macrophages along with the progression of LUAD. Additionally, along with the trajectory, the expression of M1 macrophage marker genes, including *IL1B* and *CXCL8*, decreased, while the expression of M2 macrophage marker genes, including *APOE* and *CIQA*, increased (Figure 4F), indicating that macrophages might gradually polarize from M1 to M2 as the trajectory evolved. Previous studies have demonstrated that M2 macrophages are involved in wound repair, tissue fibrosis, and angiogenesis. In this study, the expression of *VEGFB* also gradually increased as the trajectory evolved (Figure 4F), and the expression of *SPPI1*, which could promote tumor progression (Chiou et al., 2019; He et al., 2021), also increased with the evolution of pseudo time, suggesting that M2 macrophages might promote angiogenesis during the progression of LUAD.

Furthermore, we investigated the regulatory activities of transcription factors (TFs) in different trajectory states to understand the regulation of macrophage evolution. Notably, the expression of *KLF3* and *BATF3* was downregulated, whereas the expression of *STAT1* and *TCF4* was upregulated with pseudo time of the trajectory, which might be employed to promote the M2 polarization process (Supplementary Figure S4).

Proportion of Tregs and plasma B cells increases with the progression of lung adenocarcinoma

With 30,662 cells detected, NK and T cells were the most prevalent cell types. They were subclustered into 9 subsets (Figure 5A) and designated as NK cells (NK-C1 and NK-C2; *TYROBP*+), regulatory T cells (Tregs, CD4-C1; *FOXP3*+), CD4+ naïve T cells (CD4-C2, CD4-C3, CD4-C4, CD4-C5; *LEF1*+ and *CCR7*+), and CD8+ T cells (CD8-C1, CD8-C2; *CD8B*+ and *CD8A*+ (Figure 5B). Notably, NK cells were depleted from normal tissues to SNs, as previously mentioned (Lavin et al., 2017), while Tregs increased from normal tissues to SNs (Figures 5C,D). Besides, the expression of cytotoxic effector of CD8+ T cells decreased in SNs and the expression of naïve markers of CD8+ T cells was highly expressed in pure GGNs, suggesting that the cytotoxic function of T cells became weak in SNs (Figure 5E). It is established that Tregs function primarily by repressing the functionality of CD8+ T cells in cancer, including LUAD (Ganesan et al., 19502013; Adamczyk et al., 2021). Except for reactivation of CD8+ T cells, depleting Tregs may be also therapeutically beneficial.

B cells are also pivotal components of adaptive immunity, except for T cells. Previous studies have also demonstrated that tumor-infiltrating B cells play a critical role in anti-tumor immunity. A total of 3,223 B cells were identified in this study, and they were subclustered into 3 separate subsets based on canonical marker genes: follicular B cells (FLB; *HLA-DRA*+), plasma B cells (*MZB1*+ and granzyme B cells (GrB; *GZMB*+ (Figure 5F; Supplementary Figure S5A). Among these cells, follicular B cells and plasma B cells accounted for most of the total cells. Follicular B cells exhibited high expression of HLA-DRs, and its proportion was relatively higher in pure GGNs than in normal tissues and SNs, indicating that the abundance of follicular B cells increased at an early stage of LUAD (Figure 5G). IHC analysis also supported this result (Figure 5H). Previous studies have certified that follicular B cells are the precursor of plasma B cells, and when they encounter antigen and T-cell help, they can become plasma B cells or memory cells (Pieper et al., 2013). These results imply that follicular B cells may differentiate into plasma B cells with the continuous stimulation of tumor antigens during the progression of LUAD, which leads to the decrease of its abundance.

Plasma B cells are the major resource of antibodies, which are key components of humoral immunity. Of note, the abundance of plasma B cells was relatively higher in SNs (Figure 5G). IHC analysis also revealed that plasma B cells markers like CD138 (Syndecan-1) were upregulated in SNs (Supplementary Figure S5B). These indicated that plasma B cells involved in the progression of LUAD.

In order to further confirm the evolution trajectory of plasma B cells, we extracted follicular B cells and plasma B cells to infer the trajectory and also found that the proportion of plasma B cells from SNs increased in the end of pseudo time (Figure 5I). SCENIC analysis showed that TFs, such as *MYC* and *ELF1*, were depressed, while *XBPI* was activated as pseudo time progressed (Supplementary Figure S5C). Previous studies have implied that

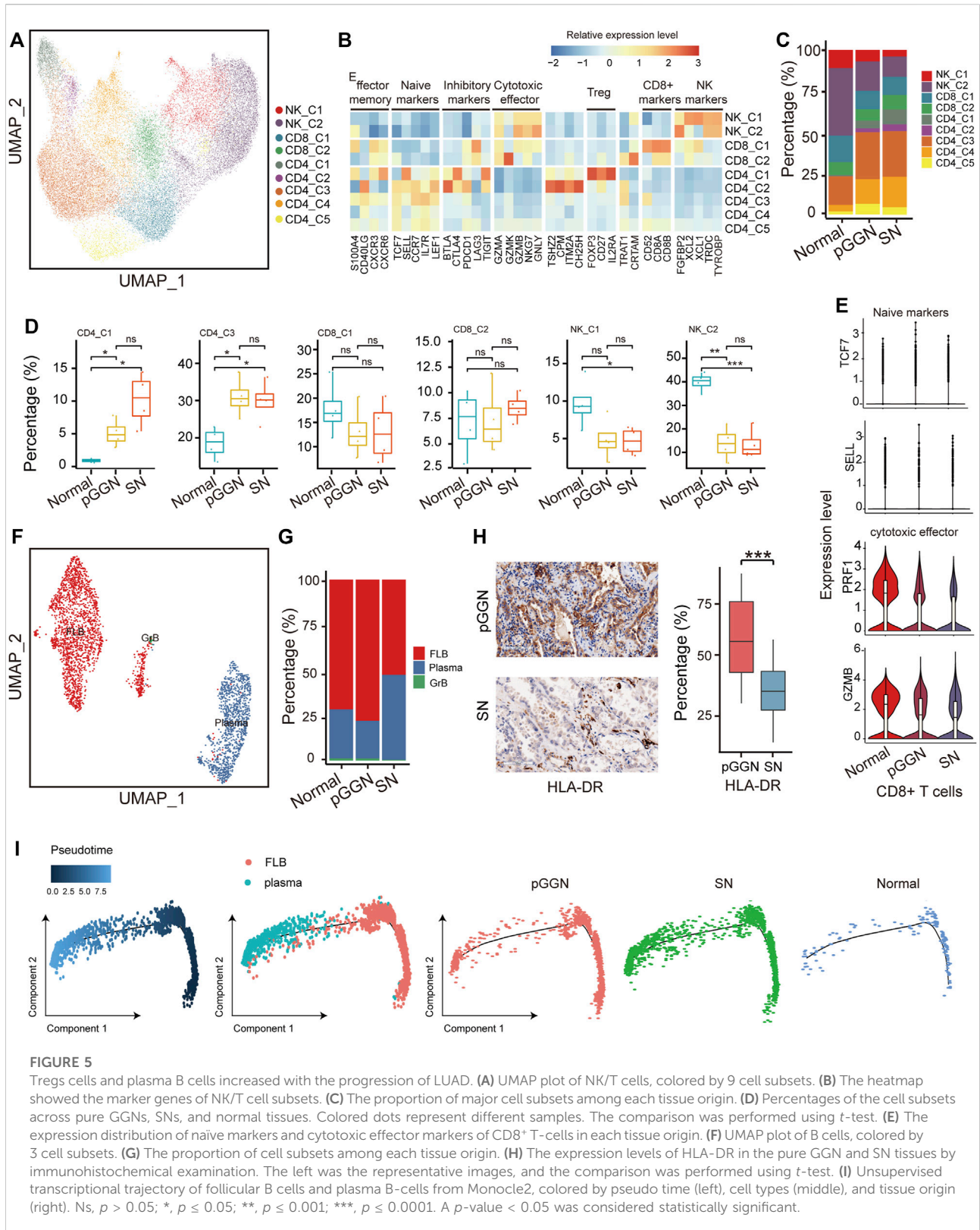
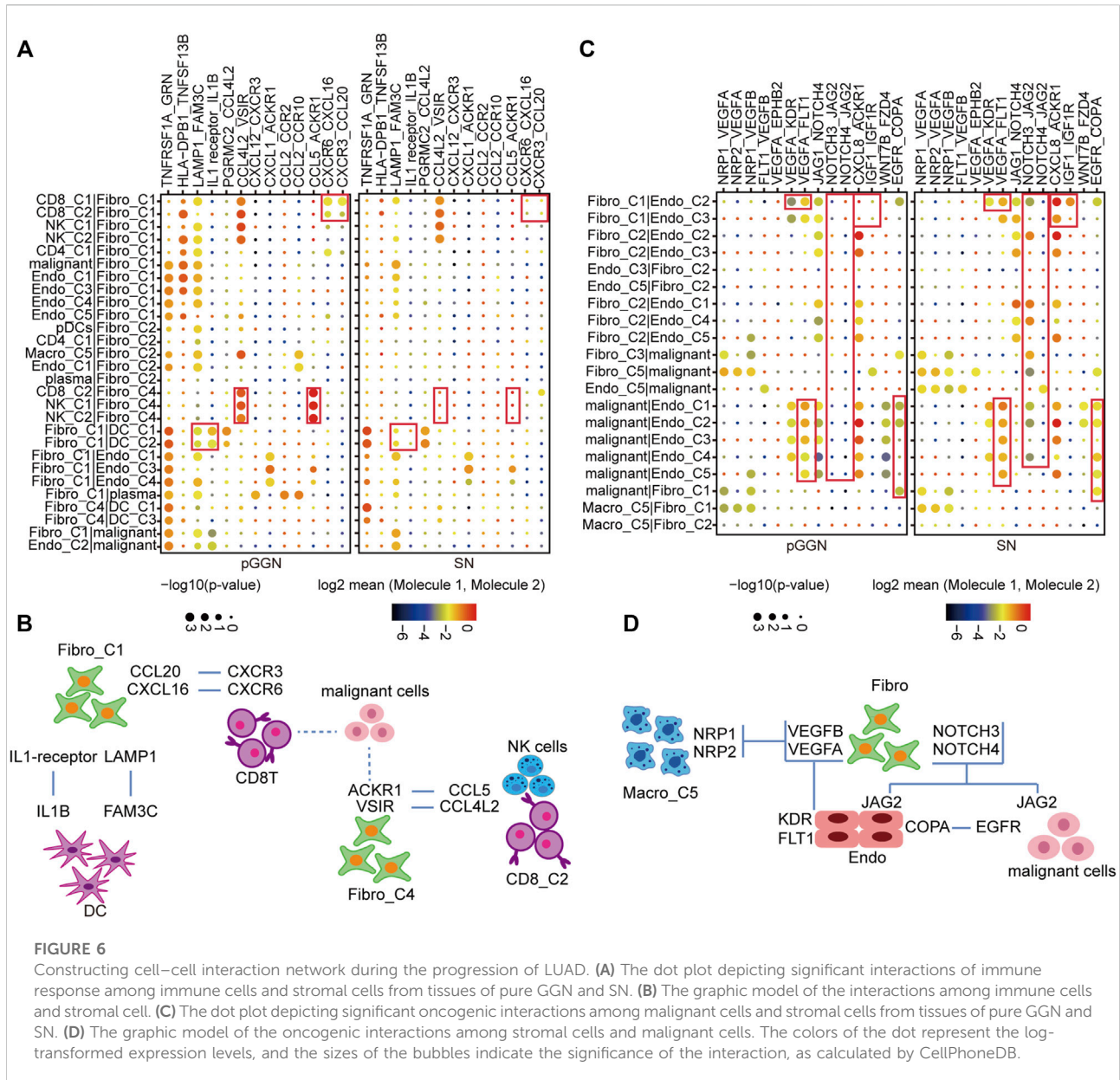


FIGURE 5

Tregs cells and plasma B cells increased with the progression of LUAD. (A) UMAP plot of NK/T cells, colored by 9 cell subsets. (B) The heatmap showed the marker genes of NK/T cell subsets. (C) The proportion of major cell subsets among each tissue origin. (D) Percentages of the cell subsets across pure GGNs, SNs, and normal tissues. Colored dots represent different samples. The comparison was performed using *t*-test. (E) The expression distribution of naïve markers and cytotoxic effector markers of CD8⁺ T-cells in each tissue origin. (F) UMAP plot of B cells, colored by 3 cell subsets. (G) The proportion of cell subsets among each tissue origin. (H) The expression levels of HLA-DR in the pure GGN and SN tissues by immunohistochemical examination. The left was the representative images, and the comparison was performed using *t*-test. (I) Unsupervised transcriptional trajectory of follicular B cells and plasma B-cells from Monocle2, colored by pseudo time (left), cell types (middle), and tissue origin (right). *Ns*, *p* > 0.05; *, *p* ≤ 0.05; **, *p* ≤ 0.001; ***, *p* ≤ 0.0001. A *p*-value < 0.05 was considered statistically significant.



XBPI may be involved in the development of lymphocytes, including plasma B cells (Manso et al., 2019). We hypothesized that activated *XBPI* and inhibited *MYC* and *ELF1* might promote the differentiation of follicular B cells into plasma B cells.

Construction of cell–cell interaction networks during the progression of lung adenocarcinoma

To characterize the transformation of intercellular interactions with the progression of LUAD, we, respectively,

inferred putative cell-to-cell interactions in pure GGNs and SNs based on ligand-receptor signaling.

Fibro-C1 could act as an immunosuppressive medium to deliver inflammatory factors and elicit anti-tumor immune response. In the network, CD8 + T cells with high expression of CXCR6 and CXCR3 formed strong cellular communication with Fibro-C1 with high expression of CXCL16 and CCL20. DC interacted with Fibro-C1 in the form of LAMP1_FAM3C and IL1 receiver_IL1B. These interactions were shown to be enhanced in patients with pure GGNs. Moreover, the interactions between CD8⁺ T cells, NK cells, and Fibro-C4 by ligand and receptor binding of CCL4L2_VSIR and

CCL5_ACKR1 were more significant in pure GGNs than in SNs (Figures 6A,B).

Furthermore, the interactions that promoted the development of the tumor itself are observed (Figures 6C,D). For instance, VEGF signaling pathway-related interactions, such as VEGFA_KDR and VEGFA_FLT1, were very strong between endothelial cells and fibroblasts. In addition, the NOTCH signaling pathway-related interactions, such as NOTCH3_JAG2 and NOTCH4_JAG2, were detected in malignant cells, endothelial cells, and fibroblasts, which were stronger in SNs than in pure GGNs. Additionally, the EGFR-related interactions were obvious between malignant cells and endothelial cells in SNs. We thought that these intercellular interactions were essential for the development of LUAD.

Meanwhile, we noted that a subset of genes was highly expressed in specific cell subtypes at specific stages. In pure GGNs, LAMP1 and FAM3C were highly expressed in Fibro-C1 and DC-C1, respectively; CCL4L2 and CCL5 were highly expressed in CD8-C2, and VSIR and ACKR1 were highly expressed in Fibro-C4. In SNs, some genes, such as VEGFA, NOTCH3/4, and EGFR, were highly expressed in multiple cell subtypes, which were consistent with our previous findings (Supplementary Figures S6A,B).

We generated a complex interaction network from different origins, including angiogenesis and immune response. The diverse interactions at different development stages of LUAD highlight potential therapeutic targets.

Discussion

A robust increase of GGNs detected during lung cancer screening program raises the blaring sirens in battle with the most notorious malignancy in the world. Interestingly, these small indolent lesions demonstrate a long history of development into aggressive lesions with systematic spreading potency. The indications for surgical resection are obscure in small pure GGNs, yet routine annual monitoring of these nodules with CT years by years causes considerable psychological stress in patients and significant consumption of medical resources (Zhang et al., 2020). Moreover, synchronous multiple nodules pose a substantial challenge in decision-making during surgery. Hence, there is an urgent need to better understand molecular mechanisms underlying the progression of these small lesions to more invasive entities. In the present study, we depicted the cellular and molecular landscape and dynamics of LUAD from pure GGN to SN using scRNA-seq technology. This study also highlighted different cellular and molecular networks in pure GGNs and SNs, which could facilitate further discovery of therapeutic targets.

Along with progression of LUAD in radiology, we discovered that epithelial cells harbor more malignant characteristics, including the accumulation of CNV and the ability to proliferate. Additionally,

the antigen-presenting ability mediated by the MHC molecules is weakened. Usually, the intrinsic characteristics of malignant cells determine tumor biology (Li et al., 2018; Wang et al., 2018). However, efforts to understand the tumor microenvironment have revealed its pivotal role in tumorigenesis (Casey et al., 2015). We found that during the progression of LUAD, innate immunity mediated by NK cells was gradually weakened, and adaptive immunity exhibited significant changes, including the development of T-cell exhaustion and the differentiation from follicular B cells to plasma B cells. Notably, immunosuppression that was mainly depending on myofibroblasts, Tregs, and macrophages was enhanced during this process, and similarly, pro-angiogenesis mediated by M2 macrophages and tECs was also enhanced. Previous studies have also demonstrated that stage I LUADs already harbor significantly altered T-cell and NK-cell components, which support our conclusions. Altogether, the progression of LUAD involves the dynamic reprogramming of multiple cellular elements, including cancer cells, immune cells, and stromal cells. The proliferation of malignant epithelial cells is weak, and immunosurveillance functions well in the early stages. These may together account for the indolent nature of pure GGNs.

T cells are a subpopulation of the most abundant lymphocytes in LUAD and are the main components of adaptive immunity (Guo et al., 2018). They have been widely studied and therapeutically targeted for immunotherapy in various solid tumors. Similarly, B cells are also key components in adaptive immunity with diverse functions. Unfortunately, there are few studies on B cells in LUADs, and inconsistent effects of B cells on tumor development have been reported (Budczies et al., 2021; Cui et al., 2021). B cells mainly produce antibodies and induce T-cell activation and proliferation through antigen presentation to regulate immune and inflammatory response. One previous study revealed that the depletion of B cells *via* anti-IgM antibodies could reduce the tumor burden in mice with colorectal cancer, suggesting a role of B cells in promoting the progression of colorectal cancer (House and Maley, 1986). Another study demonstrated that B cells could enhance the metastatic ability of breast cancer cells by activating the CXCR4/SDF1 α axis in tumor cells *via* secreting HSPA4-targeting immunoglobins (Gu et al., 2019). Conversely, several studies have revealed that B cells potentially exhibited anti-tumor functions (Helmink et al., 2020). For instance, depletion of B cells through anti-CD20 antibodies could enhance the progression of melanoma in mice. One recent study based on scRNA-seq provided multiple novel insights of B cells in non-small cell lung cancer, and identified several subtypes of tumor-infiltrating B cells with versatile functions (Chen et al., 2020). In our study, we identified five separate subsets of B cells and found that the abundance of follicular B cells was relatively higher in pure GGNs, while that of plasma B cells was relatively higher in SNs. We thought that this may be due to the differentiation from follicular B cells to plasma B cells with the continuous stimulation of tumor antigens during tumor development. Together, our

study identifies that except for T cells, B cells, especially plasma B cells, also play pivotal roles in the progression of LUAD.

In recent years, immunotherapy mainly based on immune checkpoint inhibitors (ICIs), such as anti-PD1/PD-L1 and anti-CTLA-4 antibodies, has significantly altered the landscape of cancer treatment, including lung cancer (Borghaei et al., 2015; Billan et al., 2020). Unfortunately, only a minority of unselected patients benefit from immunotherapy. This is mainly due to the significant complexity of the tumor microenvironment of solid tumor. First, at different developmental stages, immune cells that act as determinants vary. In this study, we found that NK cells were more abundant in pure GGNs, but gradually decreased. This indicates that innate immunity mediated by NK cells is still relatively active at an early stage of LUAD. In the clinic, there are more and more synchronous and metachronous multiple early-stage primary LUAD presenting as pure GGNs or mixed GGNs. Although surgery is the optimal treatment strategy for early-stage LUAD, it is very difficult to resect all GGNs within individual patients because of the limitation of surgical technology or the physical condition of the patient. This poses a substantial challenge for clinicians. One recent study reported the effect of postoperative EGFR-TKI treatment on residual GGO lesions after lung cancer resection (Cheng et al., 2021). Although some patients may benefit from this modality, there are several inherent limitations to this strategy, due to drug resistance and inter-tumor heterogeneity in at genetic level. Based on our study, we propose that NK cell-mediated immunotherapy may be an ideal and effective treatment approach for patients with pure GGN who could tolerate surgery and for patients with multiple GGNs. Fortunately, a wide variety of sources of therapeutic NK cells are currently being clinically tested (Myers and Miller, 2021). Additionally, several other stromal cells, such as CAFs and endothelial cells, also present dynamic changes during the progression of LUAD. These immune and stromal cells constitute a complex ecosystem which facilitates the tumor development. Therefore, combination therapy targeting two or more immune cells and stromal cells may be a reasonable strategy for the refractory advanced stage LUADs that have received ICI monotherapy.

There are some limitations in this study. First, the clonality of T cells and B cells is important in the process of tumor development, however, we did not investigate this. Therefore, further studies based on single-cell TCR and BCR sequencing are imperative in the future. Second, mixed GGN is composed of GGN and solid component and is highly heterogeneous. Therefore, we did not analyze mixed GGNs. Third, due to the technical limitations of scRNA-seq, we could not identify the spatial distribution of immune cells and stromal cells in individual tumor, and new technologies of spatial transcriptomics may be helpful to solve this problem. Fourth, due to technological limitations, we could not further confirm the detailed functions of

specific cell types, such as B cells. Finally, there are not sufficient functional experiments to verify our results. Therefore, further investigations on cellular and molecular biology are still required in the future.

In summary, we comprehensively revealed the cellular and molecular landscape and dynamics of different stages of resectable LUAD. Anti-tumor immunity mediated by NK and CD8+T cells gradually weakened during the progression of LUAD, and humoral immunity mediated by plasma B cells was more active in late stages. Additionally, stromal cells, including fibroblasts and macrophages, also played pivotal roles in this process. Moreover, there were complex cell-cell interactions among these cells, and they constituted a complex ecosystem to determine the development of LUAD. Our study may facilitate our understanding of LUAD and the discovery of novel therapeutic targets.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: CNCB-NGDC (National Genomics Data Center, China National Center for Bioinformation, <https://ngdc.cncb.ac.cn/gsa/>) and the accession number is PRJCA006719.

Ethics statement

This study was approved by the Ethics Committee of Peking University Cancer Hospital & Institute (Institutional Review Board No. 2019KT59). The patients/participants provided their written informed consent to participate in this study.

Author contributions

BL, CW, and ZF contributed equally to this work. NW, XF, and HQ designed the study and supervised the implementation of this study. BL, SY, SL, YM, and YW collected the samples and clinical information. XR and ML constructed cell solutions. BD and XY performed pathological analysis and IHC analysis. CW, ZF, JB, and BL contributed to the analysis and interpretation of the data. BL and CW wrote the draft of the manuscript. All authors read and approved the final version of the manuscript.

Funding

This study was supported by the National Key Research and Development Program of China (Grant Nos. 2018YFC0910700, 2018YFC0910702, and 2020YFC2003405), Beijing Natural

Science Foundation (No. 7192036), Special Fund of Beijing Municipal Administration of Hospitals Clinical Medicine Development (No. XMLX201841), and Beijing Municipal Administration of Hospital's Ascent Plan (No. DFL20191101).

Acknowledgments

The authors thank all the participants in this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Adamczyk, A., Pastille, E., Kehrmann, J., Vu, V. P., Geffers, R., Wasmer, M. H., et al. (2021). GPR15 facilitates recruitment of regulatory T cells to promote colorectal cancer. *Cancer Res.* 81 (11), 2970–2982. doi:10.1158/0008-5472.can-20-2133
- Biffi, G., and Tuveson, D. A. (2021). Diversity and biology of cancer-associated fibroblasts. *Physiol. Rev.* 101 (1), 147–176. doi:10.1152/physrev.00048.2019
- Billan, S., Kaidar-Person, O., and Gil, Z. (2020). Treatment after progression in the era of immunotherapy. *Lancet. Oncol.* 21 (10), e463–e476. doi:10.1016/s1470-2045(20)30328-4
- Bissonnette, E. Y., Lauzon-Joset, J. F., Debley, J. S., and Ziegler, S. F. (2020). Cross-Talk between alveolar macrophages and lung epithelial cells is essential to maintain lung homeostasis. *Front. Immunol.* 11, 583042. doi:10.3389/fimmu.2020.583042
- Borghaei, H., Paz-Ares, L., Horn, L., Spigel, D. R., Steins, M., Ready, N. E., et al. (2015). Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N. Engl. J. Med.* 373 (17), 1627–1639. doi:10.1056/NEJMoa1507643
- Budczies, J., Kirchner, M., Kluck, K., Kazdal, D., Glade, J., Allgäuer, M., et al. (2021). A gene expression signature associated with B cells predicts benefit from immune checkpoint blockade in lung adenocarcinoma. *Oncoimmunology* 10 (1), 1860586. doi:10.1080/2162402x.2020.1860586
- Casey, S. C., Amedei, A., Aquilano, K., Azmi, A. S., Benencia, F., Bhakta, D., et al. (2015). Cancer prevention and therapy through the modulation of the tumor microenvironment. *Semin. Cancer Biol.* 35 (Suppl. uppl), S199–S223. doi:10.1016/j.semcancer.2015.02.007
- Chang, B., Hwang, J. H., Choi, Y. H., Chung, M. P., Kim, H., Kwon, O. J., et al. (2013). Natural history of pure ground-glass opacity lung nodules detected by low-dose CT scan. *Chest* 143 (1), 172–178. doi:10.1378/chest.11-2501
- Chen, H., Carrot-Zhang, J., Zhao, Y., Hu, H., Freeman, S. S., Yu, S., et al. (2019). Genomic and immune profiling of pre-invasive lung adenocarcinoma. *Nat. Commun.* 10 (1), 5472. doi:10.1038/s41467-019-13460-3
- Chen, J., Tan, Y., Sun, F., Hou, L., Zhang, C., Ge, T., et al. (2020). Single-cell transcriptome and antigen-immunoglobulin analysis reveals the diversity of B cells in non-small cell lung cancer. *Genome Biol.* 21 (1), 152. doi:10.1186/s13059-020-02064-6
- Chen, W., Zheng, R., Baade, P. D., Zhang, S., Zeng, H., Bray, F., et al. (2016). Cancer statistics in China, 2015. *Ca. Cancer J. Clin.* 66 (2), 115–132. doi:10.3322/caac.21338
- Cheng, B., Li, C., Zhao, Y., Li, J., Xiong, S., Liang, H., et al. (2021). The impact of postoperative EGFR-TKIs treatment on residual GGO lesions after resection for lung cancer. *Signal Transduct. Target. Ther.* 6 (1), 73. doi:10.1038/s41392-020-00452-9
- Chiou, J., Chang, Y. C., Tsai, H. F., Lin, Y. F., Huang, M. S., Yang, C. J., et al. (2019). Follistatin-like protein 1 inhibits lung cancer metastasis by preventing proteolytic activation of osteopontin. *Cancer Res.* 79 (24), 6113–6125. doi:10.1158/0008-5472.can-19-0842
- Cho, J., Kim, E. S., Kim, S. J., Lee, Y. J., Park, J. S., Cho, Y. J., et al. (2016). Long-Term follow-up of small pulmonary ground-glass nodules stable for 3 Years: Implications of the proper follow-up period and risk factors for subsequent growth. *J. Thorac. Oncol.* 11 (9), 1453–1459. doi:10.1016/j.jtho.2016.05.026
- Cui, C., Wang, J., Fagerberg, E., Chen, P. M., Connolly, K. A., Damo, M., et al. (2021). Neoantigen-driven B cell and CD4 T follicular helper cell collaboration promotes anti-tumor CD8 T cell responses. *Cell.* 184 (25), 6101–6118.e13. doi:10.1016/j.cell.2021.11.007
- Ganesan, A. P., Johansson, M., Ruffell, B., Yagui-Beltrán, A., Lau, J., Jablons, D. M., et al. (19502013), 191. Baltimore, 2009–2017. doi:10.4049/jimmunol.1301317 Tumor-infiltrating regulatory T cells inhibit endogenous cytotoxic T cell responses to lung adenocarcinoma. *J. Immunol.*
- Gao, J. W., Rizzo, S., Ma, L. H., Qiu, X. Y., Warth, A., Seki, N., et al. (2017). Pulmonary ground-glass opacity: Computed tomography features, histopathology and molecular pathology. *Transl. Lung Cancer Res.* 6 (1), 68–75. doi:10.21037/tlcr.2017.01.02
- Ginhoux, F., and Jung, S. (2014). Monocytes and macrophages: Developmental pathways and tissue homeostasis. *Nat. Rev. Immunol.* 14 (6), 392–404. doi:10.1038/nri3671
- Gu, Y., Liu, Y., Fu, L., Zhai, L., Zhu, J., Han, Y., et al. (2019). Tumor-educated B cells selectively promote breast cancer lymph node metastasis by HSPA4-targeting IgG. *Nat. Med.* 25 (2), 312–322. doi:10.1038/s41591-018-0309-y
- Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., et al. (2018). Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat. Med.* 24 (7), 978–985. doi:10.1038/s41591-018-0045-3
- He, C., Sheng, L., Pan, D., Jiang, S., Ding, L., Ma, X., et al. (2021). Single-cell transcriptomic analysis revealed a critical role of SPP1/CD44-mediated crosstalk between macrophages and cancer cells in glioma. *Front. Cell. Dev. Biol.* 9, 779319. doi:10.3389/fcell.2021.779319
- Helmink, B. A., Reddy, S. M., Gao, J., Zhang, S., Basar, R., Thakur, R., et al. (2020). B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* 577 (7791), 549–555. doi:10.1038/s41586-019-1922-8
- Hirsch, F. R., Scagliotti, G. V., Mulshine, J. L., Kwon, R., Curran, W. J., Jr., Wu, Y. L., et al. (2017). Lung cancer: Current therapies and new targeted treatments. *Lancet (London, Engl.)* 389 (10066), 299–311. doi:10.1016/s0140-6736(16)30958-8
- Hosaka, K., Yang, Y., Seki, T., Fischer, C., Dubey, O., Fredlund, E., et al. (2016). Pericyte-fibroblast transition promotes tumor growth and metastasis. *Proc. Natl. Acad. Sci. U. S. A.* 113 (38), E5618–E5627. doi:10.1073/pnas.1608384113
- Hou, F., Xiao, K., Tang, L., and Xie, L. (2021). Diversity of macrophages in lung homeostasis and diseases. *Front. Immunol.* 12, 753940. doi:10.3389/fimmu.2021.753940
- House, A. K., and Maley, M. A. (1986). Colorectal carcinoma in a rat model: Suppression of tumour development and altered host immune status following

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2022.927300/full#supplementary-material>

- treatment with anti B-lymphocyte serum. *J. Surg. Oncol.* 32 (4), 256–262. doi:10.1002/jso.2930320416
- Hua, X., Zhao, W., Pesatori, A. C., Consonni, D., Caporaso, N. E., Zhang, T., et al. (2020). Genetic and epigenetic intratumor heterogeneity impacts prognosis of lung adenocarcinoma. *Nat. Commun.* 11 (1), 2459. doi:10.1038/s41467-020-16295-5
- Kim, N., Kim, H. K., Lee, K., Hong, Y., Cho, J. H., Choi, J. W., et al. (2020). Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat. Commun.* 11 (1), 2285. doi:10.1038/s41467-020-16164-1
- Lambrechts, D., Wauters, E., Boeckx, B., Aibar, S., Nittner, D., Burton, O., et al. (2018). Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat. Med.* 24 (8), 1277–1289. doi:10.1038/s41591-018-0096-5
- Lavin, Y., Kobayashi, S., Leader, A., Amir, E. D., Elefant, N., Bigenwald, C., et al. (2017). Innate immune landscape in early lung adenocarcinoma by paired single-cell analyses. *Cell.* 169 (4), 750–765. doi:10.1016/j.cell.2017.04.014
- Lei, Y., Tang, R., Xu, J., Wang, W., Zhang, B., Liu, J., et al. (2021). Applications of single-cell sequencing in cancer research: Progress and perspectives. *J. Hematol. Oncol.* 14 (1), 91. doi:10.1186/s13045-021-01105-2
- Li, J., Byrne, K. T., Yan, F., Yamazoe, T., Chen, Z., Baslan, T., et al. (2018). Tumor cell-intrinsic factors underlie heterogeneity of immune cell infiltration and response to immunotherapy. *Immunity* 49 (1), 178–193. doi:10.1016/j.immuni.2018.06.006
- Manso, B. A., Zhang, H., Mikkelsen, M. G., Gwin, K. A., Secreto, C. R., Ding, W., et al. (2019). Bone marrow hematopoietic dysfunction in untreated chronic lymphocytic leukemia patients. *Leukemia* 33 (3), 638–652. doi:10.1038/s41375-018-0280-0
- Myers, J. A., and Miller, J. S. (2021). Exploring the NK cell platform for cancer immunotherapy. *Nat. Rev. Clin. Oncol.* 18 (2), 85–100. doi:10.1038/s41571-020-0426-7
- Pieper, K., Grimbacher, B., and Eibel, H. (2013). B-cell biology and development. *J. Allergy Clin. Immunol.* 131 (4), 959–971. doi:10.1016/j.jaci.2013.01.046
- Schmall, A., Al-Tamari, H. M., Herold, S., Kampschulte, M., Weigert, A., Wietelmann, A., et al. (2015). Macrophage and cancer cell cross-talk via CCR2 and CX3CR1 is a fundamental mechanism driving lung cancer. *Am. J. Respir. Crit. Care Med.* 191 (4), 437–447. doi:10.1164/rccm.201406-1137OC
- Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer statistics, 2021. *Ca. Cancer J. Clin.* 71 (1), 7–33. doi:10.3322/caac.21654
- Silva, M., Sverzellati, N., Manna, C., Negrini, G., Marchianò, A., Zompatori, M., et al. (2012). Long-term surveillance of ground-glass nodules: Evidence from the MILD trial. *J. Thorac. Oncol.* 7 (10), 1541–1546. doi:10.1097/JTO.0b013e3182641bba
- Wang, D. H., Lee, H. S., Yoon, D., Berry, G., Wheeler, T. M., Sugarbaker, D. J., et al. (2017). Progression of EGFR-mutant lung adenocarcinoma is driven by alveolar macrophages. *Clin. Cancer Res.* 23 (3), 778–788. doi:10.1158/1078-0432.ccr-15-2597
- Wang, G., Zhao, D., Spring, D. J., and DePinho, R. A. (2018). Genetics and biology of prostate cancer. *Genes. Dev.* 32 (17-18), 1105–1140. doi:10.1101/gad.315739.118
- Xing, X., Yang, F., Huang, Q., Guo, H., Li, J., Qiu, M., et al. (2021). Decoding the multicellular ecosystem of lung adenocarcinoma manifested as pulmonary subsolid nodules by single-cell RNA sequencing. *Sci. Adv.* 7 (5), eabd9738. doi:10.1126/sciadv.abd9738
- Yuan, C., Chen, H., Tu, S., Huang, H. Y., Pan, Y., Gui, X., et al. (2021). A systematic dissection of the epigenomic heterogeneity of lung adenocarcinoma reveals two different subclasses with distinct prognosis and core regulatory networks. *Genome Biol.* 22 (1), 156. doi:10.1186/s13059-021-02376-1
- Zeltz, C., Primac, I., Erusappan, P., Alam, J., Noel, A., Gullberg, D., et al. (2020). Cancer-associated fibroblasts in desmoplastic tumors: Emerging role of integrins. *Semin. Cancer Biol.* 62, 166–181. doi:10.1016/j.semcancer.2019.08.004
- Zhang, Y., Fu, F., and Chen, H. (2020). Management of ground-glass opacities in the lung cancer spectrum. *Ann. Thorac. Surg.* 110 (6), 1796–1804. doi:10.1016/j.athoracsur.2020.04.094
- Zheng, X., Weigert, A., Reu, S., Guenther, S., Mansouri, S., Bassaly, B., et al. (2020). Spatial density and distribution of tumor-associated macrophages predict survival in non-small cell lung carcinoma. *Cancer Res.* 80 (20), 4414–4425. doi:10.1158/0008-5472.can-20-0069