IVYSPRING INTERNATIONAL PUBLISHER

JOURNAL OF CANCER

2026; 17(9): 1626-1638. doi: 10.7150/jca.138884

Research Paper

Smoking-Associated Enrichment of CXCL13$^+$ Tfh Cells and VCAM1$^+$ Fibroblasts Is Linked to Tertiary Lymphoid Structure Formation in Lung Adenocarcinoma

Yong Jun Choi[1], Chi Young Kim[1], Eun Young Kim[2], Sang Hoon Lee[2], Min Kyung Park[1], and Yoon Soo Chang[1]✉

1. Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea.
2. Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea.

✉ Corresponding author: YSC, Room 413, Research Center for Future Medicine, 63-gil 20, Eonju-ro Gangnam-gu, 06229 Seoul, South Korea. **Tel:** +82-2-2019-3309, **Fax:** +82-2-3463-3882, **E-mail:** yschang@yuhs.ac.

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). See https://ivyspring.com/terms for full terms and conditions.

Received: 2026.06.05; Accepted: 2026.08.24; Published: 2026.09.03

Abstract

Background: Smoking history is associated with improved immune checkpoint inhibitor treatment responses. We investigated smoking-associated differences in the tumor microenvironment (TME) of lung adenocarcinoma (LUAD) and identified biomarkers associated with tertiary lymphoid structure (TLS) formation.

Methods: Single-cell RNA-sequencing datasets from tumors and matched adjacent normal tissues from five current smokers and five never-smokers with LUAD were integrated, and the T- and NK-cell compartments were analyzed in detail. Key findings were further evaluated using public LUAD scRNA-seq datasets and validated by immunofluorescence in formalin-fixed paraffin-embedded LUAD tissues using QuPath-based image analysis.

Results: In current smokers, follicular helper T (Tfh) cells were enriched in the TME, whereas CD4$^+$ tissue-resident memory T cells were reduced ($R_{O/E}$, median [interquartile range (IQR)]: 2.1 [1.5–5.7] vs. 1.0 [0.5–1.5], p=0.040; and 0.7 [0.6–1.1] vs. 1.7 [1.6–2.4], p<0.001, respectively). The Tfh cluster showed high expression of *CXCL13* and increased expression of *PDCD1* and *CTLA4*. Immunofluorescence confirmed increased CXCL13$^+$CD4$^+$ cells in current smokers (83.0 [10.0–218.5] vs. 10.0 [2.0–60.5] cells/0.25 mm², p<0.001). Gene set analysis showed enrichment of a B-cell chemotaxis signature in the Tfh cluster (normalized enrichment score, 1.693; adjusted p=0.023), and CXCL13$^+$CD4$^+$ T-cell density correlated with CD20$^+$ B-cell density (Spearman ρ=0.445, p<0.001). Current smokers also showed greater TLS burden, with higher TLS counts and larger TLS areas, and had increased VCAM1$^+$ fibroblast abundance (283.0 [118.5–580.5] vs. 98.5 [32.5–191.0] cells/0.25 mm², p<0.001), which strongly correlated with CD20$^+$ cell density (Spearman ρ=0.766, p<0.001).

Conclusions: Smoking-associated enrichment of CXCL13$^+$CD4$^+$ Tfh cells was associated with B-cell accumulation and TLS formation in LUAD. These findings identify the CXCL13$^+$ Tfh–B-cell axis, together with VCAM1$^+$ fibroblast-rich stromal niches, as potential biomarkers and therapeutic targets in LUAD.

Keywords: cancer-associated fibroblast, lung adenocarcinoma, tumor microenvironment, Tfh cells, CXCL13, tertiary lymphoid structure

Introduction

Lung cancer is one of the leading causes of cancer-related mortality worldwide, with an estimated 2 million diagnoses and 1.8 million deaths [1,2]. Smoking plays a pivotal role in the initiation and progression of lung adenocarcinoma (LUAD) and squamous cell carcinoma [3,4]. A single cigarette

contains over 7000 toxic compounds and 60 carcinogens, including polycyclic aromatic hydrocarbon carcinogens such as benzo[a]pyrene and nicotine-derived nitrosamine ketone [4].

Recent trends in lung cancer therapy have shifted toward immune checkpoint inhibitors (ICIs) targeting programmed cell death protein 1 (PD-1), programmed death-ligand 1 (PD-L1), and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), demonstrating promising efficacy in treating metastatic non-small cell lung cancer (NSCLC) [5-7]. These agents have demonstrated clinical activity, particularly in patients with high PD-L1 expression, and have been associated with improved overall survival [8]. However, the predictive performance of PD-L1 as a biomarker remains limited, because a substantial proportion of patients with high PD-L1 expression fail to respond to ICIs [9,10]. In addition to PD-L1, other potential biomarkers, including tumor mutational burden and features of the tumor immune microenvironment, have been investigated to predict ICI responses [11]. However, more robust biomarkers are required to accurately predict the response to ICIs and improve patient selection for immunotherapy.

Despite the well-established harmful effects of smoking, several studies have highlighted enhanced efficacy of ICIs in smokers [12,13]. In particular, higher densities of tumor- and stromal-infiltrating $CD4^+$ and $CD8^+$ T cells were observed in smokers, suggesting a more active antitumor immune response [14]. Recent studies have demonstrated that T cells can actively promote the formation of tertiary lymphoid structures (TLS), which are associated with improved antitumor immunity [15,16]. TLS-rich tumor microenvironments exhibit organized lymphoid aggregates with enhanced local antigen presentation and coordinated adaptive immune responses, leading to more effective tumor control. However, the precise mechanisms underlying enhanced responsiveness to ICIs in smokers remain unclear.

This study aimed to investigate the biological mechanisms underlying the favorable responsiveness to ICIs associated with smoking by comparing T-cell and NK-cell profiles in the tumor microenvironment (TME) between current smokers and never-smokers. We analyzed single-cell RNA sequencing (scRNA-seq) data from tumor and matched adjacent normal lung tissues obtained from patients with early-stage LUAD who underwent surgical resection.

Methods

Study population and patient recruitment

Samples were collected from patients who underwent curative surgical resection for LUAD at Gangnam Severance Hospital, Yonsei University College of Medicine, between 2019 and 2022 (Fig. 1A) [17-19]. The discovery cohort consisted of ten patients with LUAD, including five current smokers and five never-smokers. Current smokers were defined as patients with a smoking history of at least 30 pack-years who had continued smoking until LUAD diagnosis and had stopped smoking for less than one month before surgery. Never-smokers were defined as those who had never smoked; former smokers were excluded to minimize biological heterogeneity associated with variable smoking cessation durations and cessation-related changes in the tumor microenvironment. Eligible patients met the following criteria: (1) surgically resectable stage LUAD; (2) no evidence of distant metastasis on staging workup, including positron emission tomography-computed tomography and brain magnetic resonance imaging; (3) no history of malignancy or anticancer treatment; (4) availability of residual tumor (Tu) and adjacent normal lung (NL) tissue for analysis; and (5) pathological confirmation of LUAD. Patients who had received neoadjuvant or adjuvant chemotherapy were excluded. This cohort was used as the discovery cohort for the subsequent analyses.

Tissue processing, scRNA-seq library construction, and quality control

Paired tumor and adjacent NL tissues were obtained from ten patients in the discovery cohort. Intraoperative specimens were transferred to the pathology laboratory, where frozen section examinations were performed to confirm LUAD. Fresh tumor tissue and adjacent NL tissue collected more than 2 cm from the tumor margin were cut into approximately $0.5 \times 0.5 \times 0.5$ cm^3 pieces and processed by a pathologist. Samples were preserved in MACS Tissue Storage Solution at 4 °C and sent to Macrogen Korea (Seoul, Republic of Korea) for downstream processing according to the Chromium Single Cell 3′ Gene Expression protocol (10x Genomics). Single-cell suspensions were prepared using a gentleMACS Octo Dissociator with a Heater and a Multi-Tissue Dissociation Kit 1 (Miltenyi Biotec). Libraries were generated by loading up to 10,000 cells per sample using a Chromium Single Cell 3′ GEM Library and Gel Bead Kit v3. After polyadenylated mRNA capture with poly(dT) primers, barcoded full-length cDNA was synthesized and paired-end sequencing was performed on an Illumina platform. Raw base-call files were demultiplexed into FASTQ files using Cell Ranger mkfastq and counted. Quality control was performed using Seurat (v5.0.1), excluding cells with > 10% mitochondrial gene content or < 200 genes. Expected doublets were identified and removed from

each sample using DoubletFinder (v2.0.4). After quality control and doublet removal, the scRNA-seq data from all samples were integrated using the Harmony algorithm to minimize batch effects.

External validation cohorts

Three independent external validation approaches were used to assess the robustness of the findings. First, for transcriptomic validation, we analyzed publicly available LUAD scRNA-seq datasets comprising four never-smokers and three current smokers from previously published studies [20,21]. These datasets were processed from raw FASTQ files using the same preprocessing and downstream analytical workflow as the discovery cohort, including read alignment, quality control, data integration, clustering, and cell-type annotation. Second, for spatial validation, an independent cohort of formalin-fixed, paraffin-embedded LUAD tissues was used for the immunofluorescence-based evaluation of key cellular markers and tissue architecture. Third, for external transcriptomic validation at the bulk level, publicly available LUAD RNA-sequencing cohorts were analyzed to examine whether the key gene expression patterns identified in the discovery analyses were recapitulated in larger patient cohorts from cBioPortal/GDC, including the CPTAC GDC 2025, OncoSG 2020, and TCGA GDC cohorts [22-24]. Kaplan–Meier survival analysis was performed separately using the TCGA GDC cohort.

Immunofluorescence staining and image analysis

Immunofluorescence (IF) staining was performed using formalin-fixed, paraffin-embedded LUAD tissues from an independent validation cohort approved by IRB No. 3-2024-0456. The following primary antibodies were used: anti-CXCL13/BCA1 (GeneTex, GTX108471), anti-CD4 (Ventana, 790-4423), anti-CD20 (Ventana, 760-2531), anti-VCAM1 (Abcam, ab134047), anti-FAP (Invitrogen, PA5-99458), and anti-PD-L1 (Ventana, 790-4905). Nuclei were counterstained with DAPI using ProLong™ Gold Antifade Mountant (Thermo Fisher Scientific, P36931). Fluorescence images were acquired using a confocal microscope (ZEISS LSM 980; ZEISS, Oberkochen, Germany) and a whole-slide scanner (ZEISS Axioscan 7; ZEISS). Digital image analysis was performed using QuPath software (version 0.6.0). Cells were first detected using the Cell Detection function, after which positive cells were classified using the Create Single Measurement Classifier function according to marker intensity. The co-expression of multiple markers was then determined using the Create Composite Classifier function. For quantitative analysis, ten randomly selected fields per patient (0.25 mm^2 per field) were evaluated. A TLS was defined as a dense nodular aggregate of lymphoid cells resembling a lymphoid structure, with or without a germinal center, and with a minimum diameter of 150 μm [15,25].

Bioinformatic and statistical analysis

All downstream single-cell analyses were performed using Seurat R package (version 5.0.1). Cell clustering was performed using the FindNeighbors and FindClusters functions, and marker genes were identified using the FindAllMarkers and FindMarkers functions. Cell-type annotation was based on the expression of canonical marker genes and previously reported markers. Ligand–receptor interaction analysis was performed using CellChat (version 1.6.1) and gene set enrichment analysis was conducted using clusterProfiler (version 4.8.3). Adjusted p-values for gene set enrichment analysis were calculated using the Benjamini–Hochberg method.

To compare cluster distribution according to smoking status, the ratio of observed to expected counts ($R_{O/E}$) was calculated for each T/NK-cell subcluster in each sample based on the chi-square-derived observed and expected counts. Group-wise differences in $R_{O/E}$ values and other continuous variables between current smokers and never-smokers were evaluated using the Wilcoxon rank-sum test unless otherwise specified. Correlations between cellular density and gene expression levels were assessed using Spearman's rank correlation analysis.

Trajectory and pseudotime analyses of $CD4^+$ T-cell subsets were performed using the R packages Slingshot (v2.18.0) and condiments (v1.18.0). Differences in the pseudotime distributions between groups were assessed using the Wilcoxon rank-sum test. For external bulk RNA-seq validation, gene expression values were z-score-normalized and patients were stratified into low- and high-expression groups (Q1 vs. Q4) for Kaplan–Meier survival analysis. Differences in survival were compared using log-rank tests. All statistical tests were two-sided, and $p<0.05$ was considered statistically significant.

Ethics

This study was approved by the Institutional Review Board of our institution (IRB No. 3-2019-0299). The acquisition of patient specimens for external validation was approved by a separate protocol (IRB No. 3-2024-0456). Written informed consent was obtained from all the participants.

Results

Data integration and annotation

To characterize the T- and NK-cell compartments in the TME of LUAD, tumors and matched adjacent normal tissues were obtained from five current smokers and five never-smokers (Fig. 1A). After quality control, scRNA-seq data from these samples were integrated using the Harmony algorithm [26]. For validation, we integrated publicly available scRNA-seq datasets from patients with non-advanced-stage LUAD with available smoking history data, including four never-smokers and three current smokers. The discovery dataset revealed eight major cell lineages defined by the expression of canonical marker genes. The major cell lineages identified in the discovery dataset include T/NK cells, B cells, myeloid cells, epithelial cells, endothelial cells, fibroblasts, mast cells, and proliferative cells (Fig. 1B). Subsequently, the discovery dataset was used as a reference to classify cell types in the validation dataset (Supplementary Fig. s1A and B). The baseline characteristics of the discovery dataset are presented in Supplementary Tables s1 and s2. No statistically significant differences were observed between current smokers and never-smokers in age, TNM stage, or *EGFR* mutation status. However, sex distribution differed significantly between the groups: all five current smokers were male, whereas four of the five never-smokers were female (p=0.048).

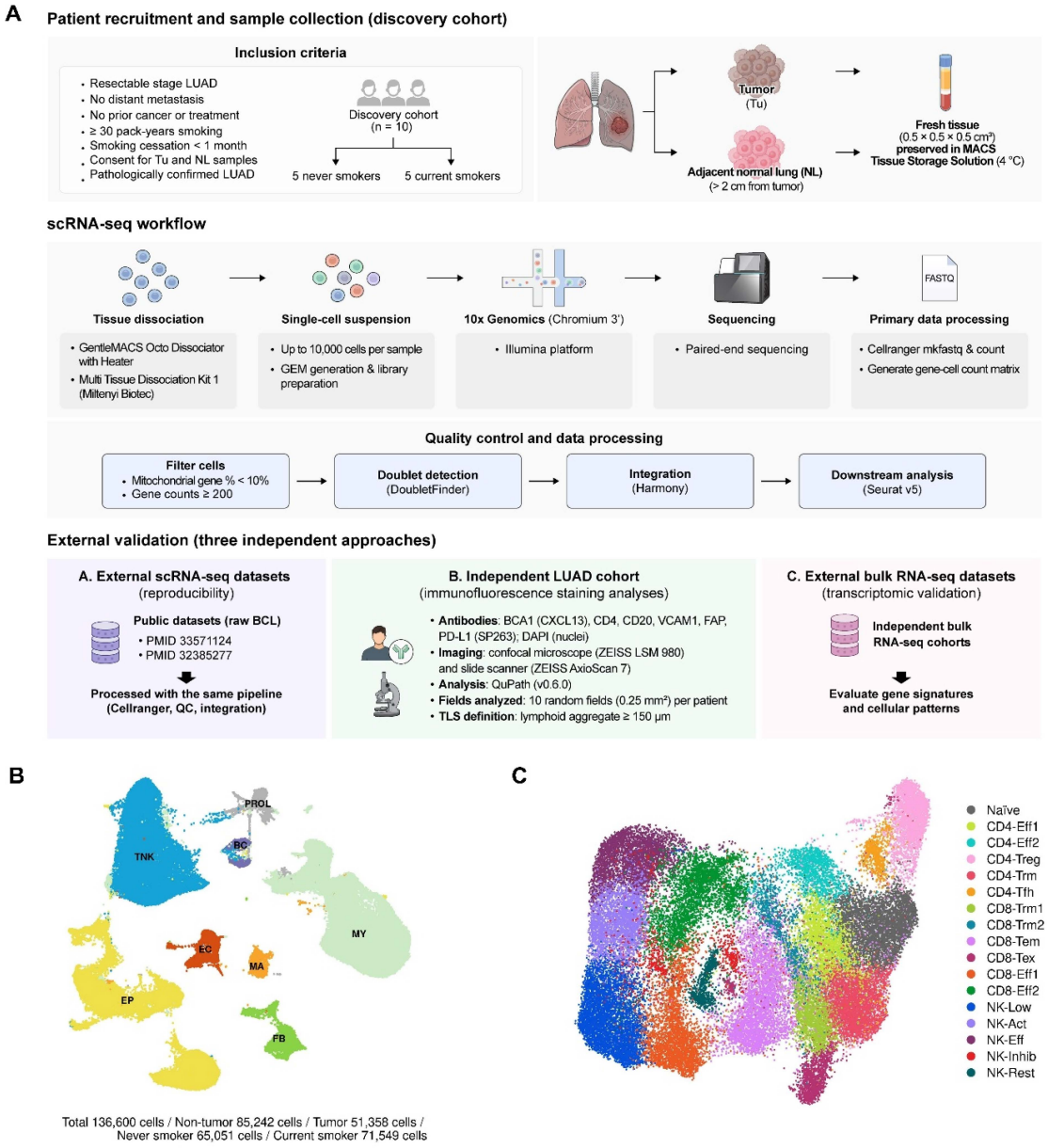

Figure 1. Single-cell landscape of T and NK cells in lung cancer. (A) Overview of sequencing datasets, including sample sources, preprocessing workflow, and distribution across cohorts. (B) UMAP visualization of major cell lineages. (C) UMAP visualization of all T and NK cells.

After quality control, 136,600 cells were retained for the discovery dataset, and 52,443 cells were retained for the validation dataset (Supplementary Table s3). The T/NK-cell compartments were extracted from both datasets for downstream analysis, yielding 45,538 cells in the discovery dataset and 23,306 cells in the validation dataset (Supplementary Table s4). After excluding putative alveolar-cell and B-cell contaminating clusters, 43,328 cells in the discovery dataset and 21,075 cells in the validation dataset were retained for subsequent T/NK-cell subtype analyses (Supplementary Table s5). To confirm annotation consistency, we compared our cell identities with the Azimuth Lung reference atlas v2.0 (Supplementary Fig. s1C) [27]. The proportions of major cell lineages in the discovery and validation datasets are summarized in Supplementary Table s4.

CXCL13$^+$ T follicular helper (Tfh) cells are enriched in the TME of current smokers

The T/NK compartment was resolved into 17 subclusters based on canonical marker gene expression (Fig. 1C and 2A, Supplementary Fig. s2A and B, and Supplementary Table s5). To compare differences in cluster distribution according to smoking status, the ratio of observed to expected counts ($R_{O/E}$) was calculated for each T/NK subcluster in each sample using the chi-square test. In the TME of current smokers, the CD4$^+$ Tfh cluster was significantly enriched, whereas the CD4$^+$ Trm cluster was markedly reduced (current smoker vs. never-smoker, CD4$^+$ Tfh: $R_{O/E,}$ median [interquartile range (IQR)], 2.11 [1.53–5.68] vs. 1.04 [0.53–1.48], p=0.040; CD4$^+$ Trm: 0.66 [0.56–1.10] vs. 1.75 [1.64–2.37], p<0.001; Fig. 2A, B; Supplementary Fig. s3). A similar compositional trend was observed in the validation dataset. No substantial differences were observed in the proportions of the remaining T-cell or NK-cell subclusters between current smokers and never-smokers. The CD4$^+$ Tfh cluster showed increased expression of multiple activation-associated markers, including *CD2, CD6, CD82, TNFRSF4* (OX40), *TNFRSF18* (GITR), *ICOS, CD27,* and *CD28*, and was characterized by prominent *CXCL13* expression (Supplementary Fig. s2B). Consistent with these transcriptomic findings, IF analysis of an independent patient cohort demonstrated a significantly higher abundance of CXCL13$^+$CD4$^+$ T cells in tumors from current smokers than in those from never-smokers (83.00 [10.00–218.50] vs. 10.00 [2.00–60.50] cells per 0.25 mm², p<0.001; Fig. 2B, C). CD4$^+$ Tfh cells also exhibited relatively high expression of immune checkpoint-related molecules, including *PDCD1* (PD-1) and *CTLA4* (CTLA-4) (Supplementary Fig. s2B). Consistent with these findings, the external RNA-seq validation cohort revealed positive correlations between *CXCL13* expression and *PDCD1* and *CTLA4* across cohorts, and with *CD274 (PD-L1*) in some of the external datasets (Supplementary Fig. s4) [22-24]. In addition, IF staining of LUAD samples demonstrated a positive correlation between CXCL13$^+$CD4$^+$ T cells and PD-L1 expression **(**Supplementary Fig. s4B–D).

Trajectory analysis of CD4$^+$ T cells identified three major differentiation paths originating from naïve cells and leading to CD4$^+$ Eff, CD4$^+$ Trm, and CD4$^+$ Tfh states (Fig. 2D). Within the naïve-to-Tfh trajectory, cells from current smokers occupied later pseudotime positions than those from never-smokers (5.9 [4.4–7.8] vs. 5.0 [3.4–7.1], p<0.001). Along this inferred trajectory, the cells progressively upregulated *CXCL13, TIGIT, TSHZ2, ICA1*, and *IGFL2*, consistent with the acquisition of a Tfh-like activated and immune regulatory program, whereas *CCL5, ANXA1, VIM, GZMA*, and *CD52* were downregulated, indicating the attenuation of cytotoxic and migratory features (Fig. 2E).

CD4$^+$ Tfh cell enrichment in the TME is associated with B-cell accumulation and TLS development

Gene set enrichment analysis showed that the CD4$^+$ Tfh cluster was preferentially enriched for B-cell-related programs, including B-cell proliferation, activation, receptor signaling, and chemotaxis (Fig. 3A; Supplementary Fig. s2C–E). Among these, the B-cell chemotaxis signature was significantly enriched (normalized enrichment score (NES) = 1.693, adjusted p=0.023; Fig. 3B).

Consistent with these transcriptomic findings, IF imaging demonstrated more prominent accumulation of CXCL13$^+$CD4$^+$ T cells and CD20$^+$ cells in tumors from current smokers than in those from never-smokers (Fig. 3C). In addition, the number of CXCL13$^+$CD4$^+$ T cells was positively correlated with CD20$^+$ cell density (Spearman ρ = 0.445, p<0.001; Fig. 3D), and CD20$^+$ B-cell density was significantly higher in current smokers than in never-smokers (41.5 [9.0–190.0] vs. 5.0 [0.0–36.0] cells per 0.25 mm², p<0.001; Fig. 3E).

Current smokers also exhibited greater TLS burden than never-smokers, with significantly higher TLS counts and larger TLS areas on field-based quantification (0.0 [0.0–1.0] vs. 0.0 [0.0–0.0] per 0.25 mm², p=0.040; and 0.0 [0.0–35011.3] vs. 0.0 [0.0–0.0] μm² per 0.25 mm², p=0.021; Fig. 3F). Moreover, CXCL13$^+$CD4$^+$ T-cell density was significantly higher in TLS areas than in non-TLS areas (19.0 [3.0–60.0] vs. 3.0 [0.5–10.0] per 0.25 mm², p<0.001; Fig. 3F), further supporting the association of CXCL13$^+$ Tfh cells with TLS development in the TME.

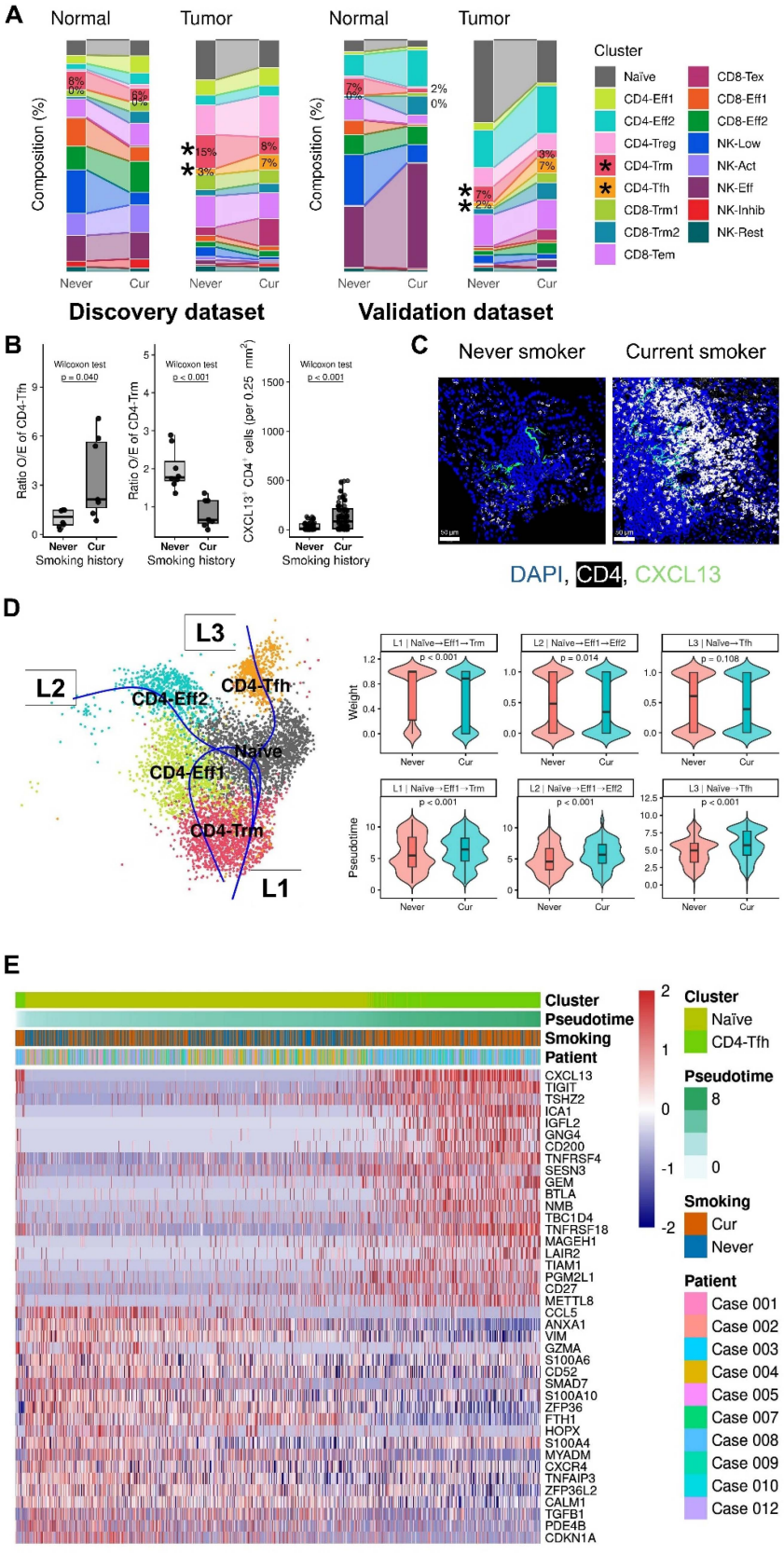

Figure 2. Expansion of CXCL13⁺CD4⁺ Tfh cells in smokers. (A) Proportional composition of T/NK-cell clusters stratified by smoking status in normal and tumor tissues across cohorts. Asterisks (*) indicate clusters showing significant differences in both cell proportion and ratio of observed to expected ($R_{O/E}$) values between groups. In the tumor tissues of both the discovery and validation sets, CD4-Trm cells were less abundant in current smokers, whereas CD4-Tfh cells were more abundant in current smokers. (B) Comparison of CD4-Tfh and CD4-Trm cell enrichment between the smoking groups. Single-cell RNA sequencing–derived $R_{O/E}$ values are shown alongside validation using an independent cohort, where CXCL13⁺CD4⁺ cell abundance was quantified by immunofluorescence. Current smokers showed increased enrichment of CD4-Tfh cells and decreased enrichment of CD4-Trm cells compared to never-smokers. Consistently, CXCL13⁺CD4⁺ cells were more abundant in smokers by immunofluorescence analysis. p-

values were calculated using the Wilcoxon rank-sum test. (C) Representative immunofluorescence images showing CXCL13$^+$CD4$^+$ T cells (CXCL13, green; CD4, white; DAPI, blue) in the tumor tissues. Scale bar: 50 μm. (D) Trajectory inference analysis of CD4$^+$ T cells showing differentiation from the naïve to effector and Tfh states. Distinct lineage paths (L1–L3) are indicated, with comparisons of trajectory weights and pseudotimes between smoking groups. (E) Heatmap of the gene expression dynamics along the naïve-to-Tfh trajectory ordered by pseudotime. Cells were annotated based on cluster identity, smoking status, and patient.

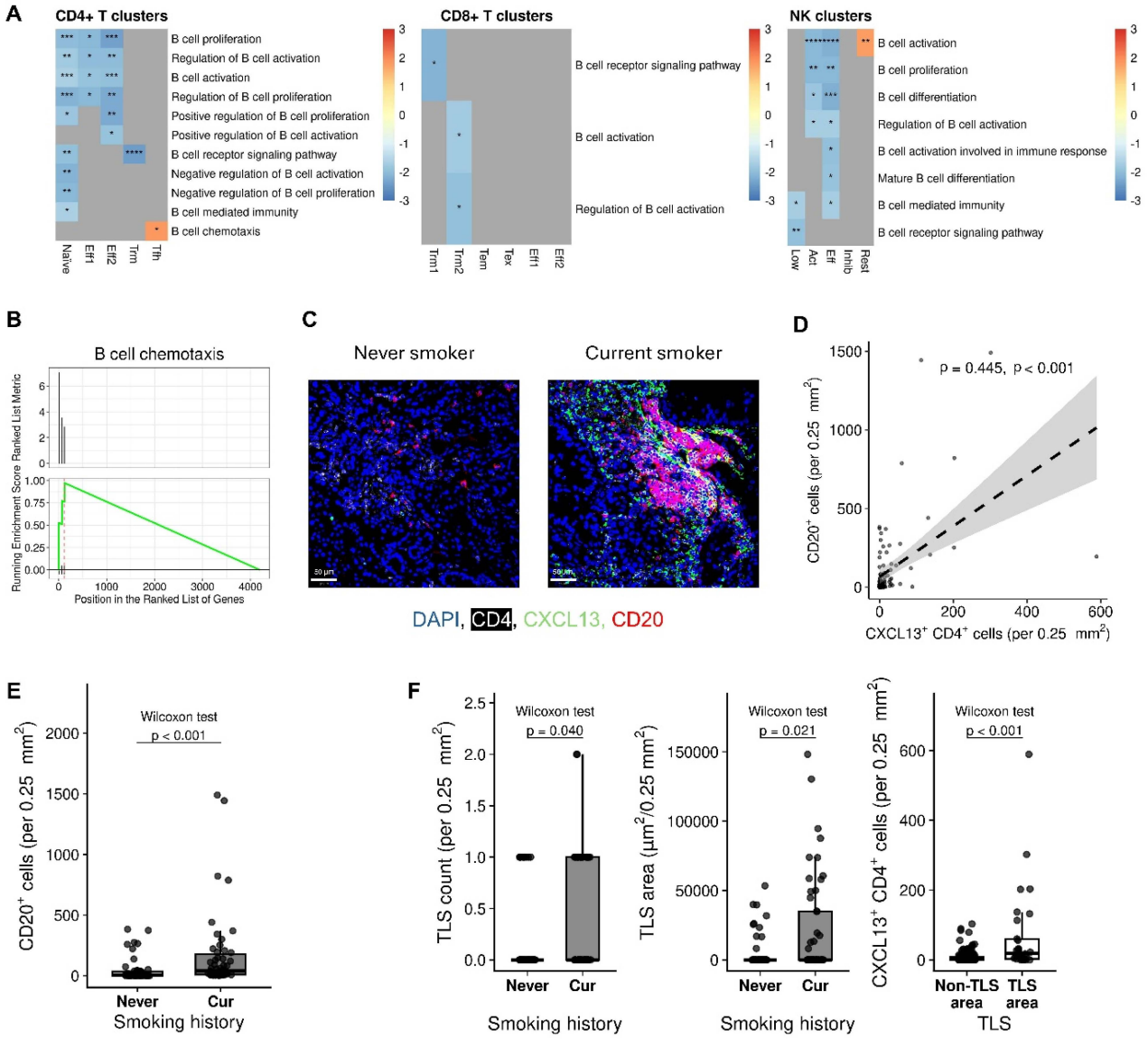

Figure 3. Smoking-associated enrichment of CXCL13$^+$CD4$^+$ T cells and B-cell recruitment with tertiary lymphoid structure (TLS) formation. (A) Gene set enrichment analysis (GSEA) of B-cell-related pathways across CD4$^+$, CD8$^+$, and NK-cell clusters. B-cell activation, proliferation, and receptor signaling pathways were predominantly enriched in specific CD4$^+$ T-cell clusters. (B) GSEA plot demonstrating the enrichment of the B-cell chemotaxis gene set in CD4$^+$ Tfh cells. (C) Representative multiplex immunofluorescence images showing DAPI (blue), CD4 (white), CXCL13 (green), and CD20 (red). Co-localization of CXCL13$^+$CD4$^+$ T cells with CD20$^+$ B cells is observed in tumor regions. Scale bar: 50 μm. (D) Correlation between CXCL13$^+$CD4$^+$ T cells and CD20$^+$ B cells, indicating a positive association. (E) Comparison of CD20$^+$ B-cell density according to smoking history. Current smokers showed significantly higher levels of CD20$^+$ cell infiltration. (F) TLS-related analysis. Current smokers had increased TLS counts and larger TLS areas. CXCL13$^+$CD4$^+$ T cells were enriched in TLS regions compared with non-TLS regions.

VCAM1$^+$ fibroblasts are associated with CXCL13$^+$ Tfh enrichment and TLS development

To explore potential intercellular communication patterns associated with the enrichment of CD4$^+$ Tfh cells, B cells, and TLS formation in tumors from current smokers, we performed CellChat-based ligand–receptor interaction analysis across TLS-related cell populations [28]. Cells were stratified according to tissue type (adjacent normal vs. tumor) and smoking status (never-smoker vs. current smoker) in the discovery dataset. Across these subgroup analyses, 96 putative ligand–receptor interactions were inferred, with the SPP1, VCAM, BTLA, and MHC-I signaling axes showing relatively greater predicted activity in tumors from current smokers (Supplementary Fig. s5 and Fig. 4A). Among the analyzed cell types, fibroblasts showed the most prominent predicted outgoing signaling activity,

indicating a fibroblast-centered communication pattern in the CellChat analysis (Supplementary Fig. s5).

Representative IF images of tumor tissues from ten never-smokers and ten current smokers showed more abundant VCAM1$^+$FAP$^+$ fibroblasts in current smokers, frequently localized within regions enriched in CXCL13$^+$ or CD20$^+$ cells (Fig. 4B, C). Quantitative analysis confirmed that VCAM1$^+$FAP$^+$ fibroblasts were significantly more abundant in current smokers than in never-smokers (283.0 [118.5–580.5] vs. 98.5 [32.5–191.0] cells per 0.25 mm², p<0.001; Fig. 4D). In addition, VCAM1$^+$FAP$^+$ fibroblast density was positively correlated with both CXCL13$^+$ cell density and CD20$^+$ cell density (Spearman ρ = 0.362 and 0.766, respectively; Fig. 4E). TLS regions also showed significantly higher infiltration of VCAM1$^+$FAP$^+$ cells compared with non-TLS regions (369.5 [166.0–577.0] vs. 39.0 [5.0–134.0] cells per 0.25 mm², p<0.001; Fig. 4D). Furthermore, the abundance of VCAM1$^+$FAP$^+$ cells was positively correlated with the total TLS area per 0.25 mm² (Spearman ρ = 0.454, p<0.001; Fig. 4E).

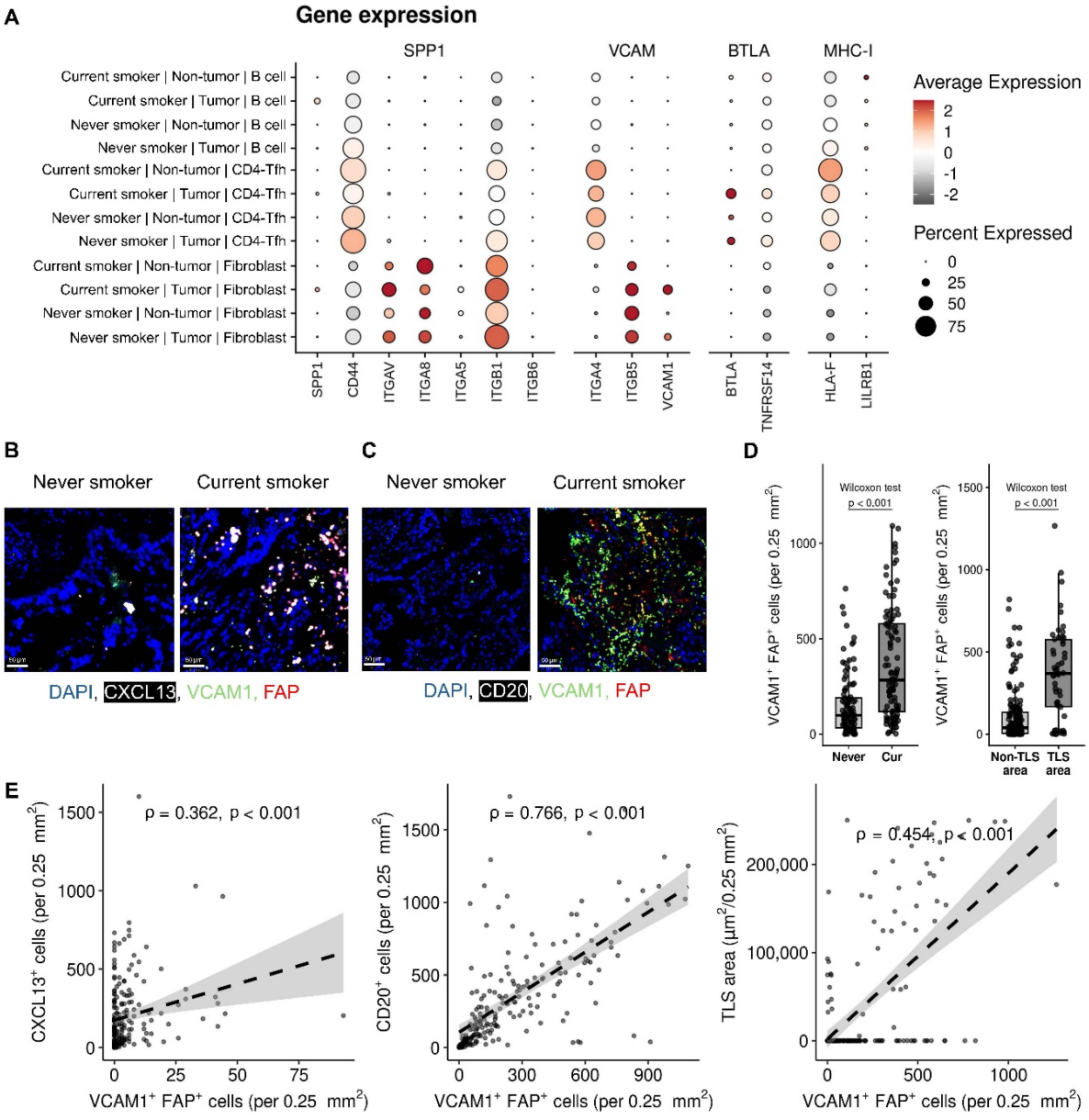

Figure 4. VCAM1$^+$FAP$^+$ stromal cells are associated with CXCL13$^+$CD4$^+$ Tfh cells and B-cell infiltration in smokers. (A) Dot plot showing the expression of ligand–receptor genes related to the SPP1, VCAM, BTLA, and MHC-I signaling pathways across major cell populations stratified by smoking status and tissue type. VCAM-related genes are predominantly expressed in stromal compartments, including fibroblasts. (B) Representative multiplex immunofluorescence images showing co-localization of CXCL13$^+$ cells with VCAM1$^+$FAP$^+$ stromal cells. DAPI (blue), CXCL13 (white), VCAM1 (green), and FAP (red) staining. Scale bar: 50 μm. (C) Representative multiplex immunofluorescence images showing co-localization of CD20$^+$ cells with VCAM1$^+$FAP$^+$ stromal cells. DAPI (blue), CD20 (white), VCAM1 (green), and FAP (red) staining. Scale bar: 50 μm. (D) Quantification of VCAM1$^+$FAP$^+$ stromal cell density based on multiplex immunofluorescence according to smoking history and TLS status. Current smokers exhibited significantly higher VCAM1$^+$FAP$^+$ cell infiltration than never-smokers, and VCAM1$^+$FAP$^+$ cells were enriched in TLS regions compared with non-TLS regions (Wilcoxon test, p<0.001 for both). (E) Correlation analyses showing positive associations of VCAM1$^+$FAP$^+$ stromal cells with CXCL13$^+$ cells, CD20$^+$ cells, and TLS area, as assessed by Spearman's correlation analysis.

To further assess the relevance of these signaling pathways in external cohorts, we examined the correlation between *CXCL13* and genes involved in the four selected ligand–receptor axes. In these RNA-seq cohorts, several genes within the VCAM, BTLA, and MHC-I signaling pathways positively correlated with *CXCL13* expression, whereas *SPP1*-related genes showed less consistent associations (Supplementary Fig. s6). These findings support a potential association between VCAM1$^+$ fibroblast-rich stromal niches and CXCL13$^+$ Tfh/B-cell–enriched immune organization in tumors of current smokers.

Expression of *CXCL13* and selected TLS-associated genes is associated with overall survival in LUAD

To explore the potential clinical relevance of the molecular profiles identified in our analyses, we evaluated their association with 5-year overall survival in an external large-scale LUAD bulk RNA-seq cohort from the TCGA GDC portal (Fig. 5). Candidate genes were selected from key molecules highlighted across the preceding Tfh-, immune checkpoint-, and ligand–receptor–related analyses among TLS-related cell clusters in the TME of smokers and never-smokers. Gene expression levels were z-score-normalized and patients were stratified into low- and high-expression groups (Q1 vs. Q4) for Kaplan–Meier survival analysis. High *CXCL13* expression was significantly associated with improved overall survival, supporting the potential prognostic relevance of the CXCL13$^+$ Tfh-related immune program in LUAD. Among the selected candidate genes, higher expression of immune regulatory genes, including *CTLA4, BTLA*, and *TNFRSF14*, as well as adhesion-related genes, including *ITGA8, ITGA5, ITGB1*, and *ITGB5*, was associated with favorable survival outcomes. In contrast, several other genes, including *PDCD1, CD274, VCAM1*, and *FAP*, were not significantly associated with overall survival (Fig. 5). These findings suggest that the CXCL13-associated immune program, together with selected stromal interaction pathways, may be associated with favorable prognosis in LUAD.

Discussion

In this study, we characterized the TME of early-stage LUAD according to smoking status, using integrated single-cell transcriptomic and spatial analyses. Current smoking was associated with enrichment of CXCL13$^+$CD4$^+$ T cells, increased B-cell accumulation, and enhanced TLS formation. However, these findings should not be interpreted as suggesting that smoking confers a favorable prognosis in NSCLC, nor should they support smoking behavior. It is important to distinguish between the intrinsic biological mechanisms identified in this study and any perceived beneficial effects of smoking.

In addition, cigarette smoking may remodel the LUAD TME in an exposure-duration–dependent manner [29]. Ongoing or recent smoke exposure can rapidly induce oxidative stress, epithelial injury, and the release of pro-inflammatory cytokines, resulting in transient activation and recruitment of innate immune cells [30]. In contrast, prolonged exposure may establish more persistent changes through cumulative genomic damage, epigenetic reprogramming, sustained adaptive immune dysregulation, and remodeling of fibroblasts and the extracellular matrix [31,32]. Therefore, the TME observed in current smokers may reflect both acute inflammatory responses to recent exposure and chronic molecular and stromal alterations caused by long-term smoking. Several key findings were obtained in this study.

First, current smokers showed enrichment of CXCL13$^+$ Tfh cells with relative depletion of CD4$^+$ Trm cells, indicating smoking-associated remodeling of the CD4$^+$ T-cell landscape toward a follicular helper–like phenotype. Previous studies have reported heterogeneous smoking-related immune alterations including T-cell dysfunction and inflammatory remodeling [33,34]. Smoking is also linked to increased *CXCL13* expression and TLS formation in inflammatory lung diseases. Although direct evidence linking smoking to CXCL13$^+$ Tfh cells and TLS formation in lung cancer remains limited, osteopontin, encoded by *SPP1*, has been implicated in Tfh differentiation and chronic inflammatory remodeling [35-37]. In our study, CellChat analysis also predicted relatively greater SPP1 signaling activity in tumors from smokers, consistent with previous reports showing smoking-associated *SPP1* upregulation in lung tissues. Together with previous reports, these computational observations raise the possibility that smoking-associated inflammatory niches may contribute to CXCL13$^+$ Tfh enrichment and TLS organization.

Second, CXCL13$^+$ Tfh cells highly expressed immune checkpoint-related molecules, including *PDCD1* and *CTLA4*. In external validation datasets, *CXCL13* expression was also positively associated with checkpoint-related programs. This checkpoint-enriched phenotype is consistent with previous studies of CXCL13-producing Tfh-like cells [38,39]. In addition, recent studies have associated CXCL13$^+$ cells and TLS-rich TMEs with improved responses to ICIs across several cancers [40-42]. Consistent with these findings, higher *CXCL13* expression was associated with favorable survival outcomes in this study. These observations may partially explain the improved

responsiveness to ICIs observed in smokers, although further prospective validation is required.

Third, CXCL13$^+$ Tfh cells and VCAM1$^+$ fibroblasts were spatially associated with TLS-related regions. In parallel, CellChat analysis predicted relatively greater VCAM-related signaling involving fibroblasts in the tumors from current smokers. Together, these observations suggest an association between VCAM1$^+$ fibroblasts and TLS-related stromal organization, but do not establish direct cellular crosstalk or causality.

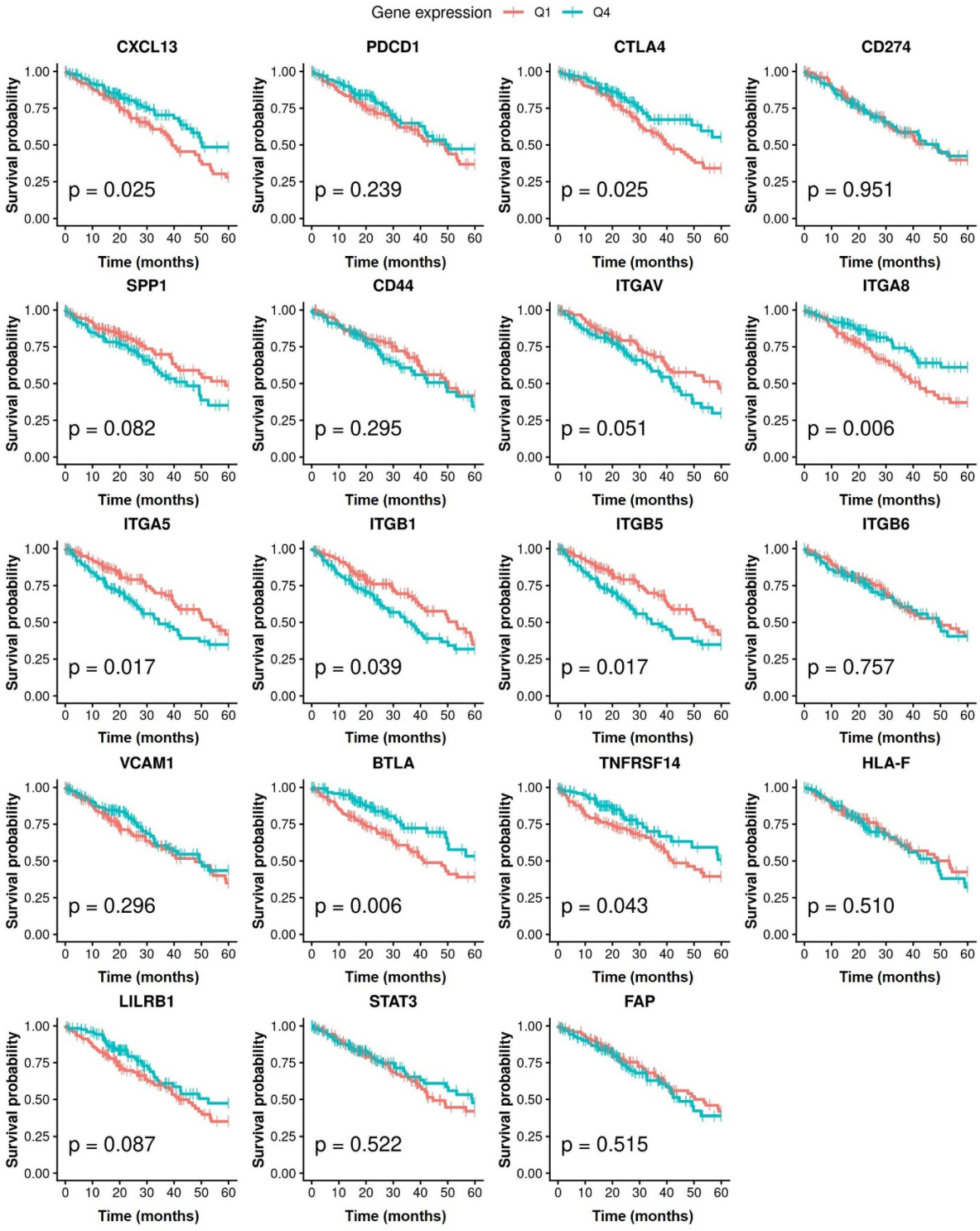

Figure 5. Prognostic significance of CXCL13-associated signaling genes. Kaplan–Meier survival curves comparing overall survival between patients with high (Q4) and low (Q1) expression of *CXCL13* and genes related to the SPP1, VCAM, BTLA, and MHC-I signaling pathways. Patients with higher expression of *CXCL13* and immune regulatory genes such as *CTLA4* and *BTLA* showed significantly improved survival, whereas several adhesion-related genes (e.g., integrins) also demonstrated prognostic associations. p-values were calculated using the log-rank test.

Previous studies have highlighted the importance of CXCL13$^+$ T cells and fibroblasts in TLS development across multiple cancer types [43-48]. Although our findings are associative, they support the potential role of VCAM1$^+$ fibroblasts in smoking-associated TLS formation. However, neither bulk *VCAM1* nor *FAP* expression was significantly associated with overall survival in the TCGA LUAD cohort. This finding indicates that the spatial association of VCAM1$^+$FAP$^+$ fibroblasts with TLS-related immune regions should not be interpreted as evidence of a protective or favorable prognostic function. Cancer-associated fibroblasts (CAFs) are functionally heterogeneous, and *VCAM1* expression may be shared by fibroblast states with distinct or even opposing effects on tumor immunity and progression. Furthermore, bulk RNA-sequencing cannot distinguish *VCAM1* expression in VCAM1$^+$FAP$^+$ fibroblasts from that in other stromal, endothelial, or immune cell populations, nor can it estimate the abundance of this specific fibroblast subset. Therefore, the functional and prognostic significance of VCAM1$^+$FAP$^+$ fibroblasts remains uncertain and requires cell-type-specific and functional validation.

This study has several limitations. First, our conclusions regarding cellular interactions were primarily based on associations observed through immunofluorescence staining, rather than direct functional validation. Although we complemented these findings with single-cell transcriptomic analyses and external validation using publicly available datasets, we were unable to perform *in vitro* or *in vivo* functional experiments such as cell-based assays or animal models. Therefore, additional mechanistic studies are required to establish causal relationships. An additional limitation is that smoking exposure cannot be experimentally controlled in human subjects for ethical reasons. Accordingly, this study was necessarily observational in design, and residual confounding cannot be fully excluded. To mitigate this limitation, we compared clearly defined current smokers and never-smokers, used paired tumor and adjacent normal tissues, and incorporated orthogonal validation across independent public scRNA-seq datasets, an external immunofluorescence cohort, and external bulk RNA-sequencing cohorts. Second, multiplex immunofluorescence was conducted using a four-channel platform, which limited our ability to directly interrogate complex cell–cell interactions at higher dimensional resolution. To address this limitation, we performed high-resolution whole-slide scanning and detailed spatial quantification, which enabled comprehensive assessment of cell distribution and TLS-related features. Nevertheless, future studies using higher-plex imaging technologies may provide a more precise characterization of cellular interactions within the TLS. Third, the discovery dataset was derived from a single center and included only ten patients, which limited statistical power and increased susceptibility to patient-specific or outlier effects. Although no statistically significant differences were observed in age, TNM stage, or *EGFR* mutation status between current smokers and never-smokers, the small sample size precludes firm conclusions regarding baseline comparability. In addition, sex distribution was significantly imbalanced: all current smokers were male, whereas four of the five never-smokers were female. This pattern partly reflects the strong association between smoking status and sex in the Korean population; however, sex may independently influence the tumor immune microenvironment and therefore represents an important potential confounding factor. Accordingly, the observed differences should be regarded as exploratory and require confirmation in larger, sex-balanced, multicenter cohorts.

Our findings should not be interpreted as supporting smoking behavior. Understanding the intrinsic mechanisms underlying smoking-associated immune remodeling may provide insights into tumor immune organization and inform future immunotherapeutic strategies. Emerging evidence suggests that CXCL13 has clinical potential not only as a biomarker for favorable prognosis, but also as an adjunctive immunotherapeutic target [49,50]. By contrast, direct therapeutic evidence for VCAM-1 remains limited; however, given its role in leukocyte recruitment and retention, fibroblast-associated VCAM-1 may represent a stromal program that cooperates with CXCL13-driven immune organization and could be explored in future combinatorial strategies [51-53].

Conclusions

In this study, CXCL13$^+$ Tfh cells were enriched in the TME of current smokers with LUAD and exhibited increased expression of ICI-related genes. CXCL13$^+$ Tfh cells and VCAM1$^+$FAP$^+$ fibroblasts were enriched in tumors from current smokers and spatially associated with TLS-rich regions. These findings suggest that they are components of smoking-associated TLS-related immune organization, although their functional interactions and clinical implications remain to be established. These findings may provide biological relevance, rather than direct evidence, for the differential ICI responsiveness previously reported according to smoking history and support further evaluation of immune and stromal features as candidate biomarkers in ICI-treated

cohorts. Further mechanistic and prospective clinical studies are needed to validate these findings.

Abbreviations

CAF: Cancer-associated fibroblast; **ICI:** immune checkpoint inhibitor; **LUAD**: lung adenocarcinoma; **NSCLC:** non-small cell lung cancer; **TME:** tumor microenvironment; **TLS:** tertiary lymphoid structure.

Supplementary Material

Supplementary figures and tables.
https://www.jcancer.org/v17p1626s1.pdf

Acknowledgments

We thank MID (Medical Illustration & Design), a member of the Medical Research Support Services of Yonsei University College of Medicine, for providing excellent support with medical illustration.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grant (RS-2023-NR076411) awarded to YS Chang.

Ethics approval and consent to participate

This study was approved by the Institutional Review Board (IRB) of Gangnam Severance Hospital (IRB No. 3-2019-0299). The acquisition of patient specimens for external validation was approved by a separate protocol (IRB No. 3-2024-0456). The data were collected in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all the participants.

Availability of data and materials

The datasets are available in the NCBI Sequence Read Archive under BioProject accession numbers PRJNA773987 and PRJNA901260.

Author contributions

Yong Jun Choi: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Visualization, Writing - original draft.

Chi Young Kim: Investigation, Data curation, Writing – review & editing.

Sang Hoon Lee: Data curation, Investigation, Writing – review & editing.

Eun Young Kim: Conceptualization, Data curation, Methodology, Writing – review & editing.

Min Kyung Park: Investigation (experimental work), Writing – review & editing.

Yoon Soo Chang: Conceptualization, Supervision, Funding acquisition, Resources, Writing – review & editing.

All authors read and approved the final manuscript.

Competing Interests

The authors have declared that no competing interest exists.

References

1. Thandra KC, Barsouk A, Saginala K, et al. Epidemiology of lung cancer. Contemp Oncol (Pozn). 2021;25(1):45–52.
2. Kim Y, Hwang JY, Na K, et al. Mutation-Driven Immune Microenvironments in Non-Small Cell Lung Cancer: Unrevealing Patterns through Cluster Analysis. Yonsei Med J. 2024;65(12):683–694.
3. Yoshida K, Gowers KHC, Lee-Six H, et al. Tobacco smoking and somatic mutations in human bronchial epithelium. Nature. 2020;578(7794):266–272.
4. Corke LK, Li JJN, Leighl NB, et al. Tobacco Use and Response to Immune Checkpoint Inhibitor Therapy in Non-Small Cell Lung Cancer. Curr Oncol. 2022;29(9):6260–6276.
5. Tan S, Day D, Nicholls SJ, et al. Immune Checkpoint Inhibitor Therapy in Oncology: Current Uses and Future Directions: JACC: CardioOncology State-of-the-Art Review. JACC CardioOncol. 2022;4(5):579–597.
6. Kim M, Chang C-S, Choi MC, et al. Rechallenge with Anti-PD-1 Inhibitors in Patients with Recurrent Gynecologic Malignancies. Yonsei Med J. 2023;64(10):587–592.
7. Zhuang Y, Tong L, Sun X. Acidic Tumor Microenvironments and Emerging Therapeutic Strategies for Cancer Therapy. Yonsei Med J. 2025;66(10):619–629.
8. Reck M, Rodríguez-Abreu D, Robinson AG, et al. Pembrolizumab versus Chemotherapy for PD-L1–Positive Non–Small-Cell Lung Cancer. N Engl J Med. 2016;375(19):1823–1833.
9. Topalian SL, Taube JM, Anders RA, et al. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. Nat Rev Cancer. 2016;16(5):275–287.
10. Doroshow DB, Bhalla S, Beasley MB, et al. PD-L1 as a biomarker of response to immune-checkpoint inhibitors. Nat Rev Clin Oncol. 2021;18(6):345–362.
11. Havel JJ, Chowell D, Chan TA. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. Nat Rev Cancer. 2019;19(3):133–150.
12. Chen DL, Li QY, Tan QY. Smoking history and the efficacy of immune checkpoint inhibitors in patients with advanced non-small cell lung cancer: a systematic review and meta-analysis. J Thorac Dis. 2021;13(1):220–231.
13. Wang X, Ricciuti B, Alessi JV, et al. Smoking History as a Potential Predictor of Immune Checkpoint Inhibitor Efficacy in Metastatic Non-Small Cell Lung Cancer. J Natl Cancer Inst. 2021;113(12):1761–1769.
14. Reuben A, Zhang J, Chiou SH, et al. Comprehensive T cell repertoire characterization of non-small cell lung cancer. Nat Commun. 2020;11(1):603.
15. Sautès-Fridman C, Petitprez F, Calderaro J, et al. Tertiary lymphoid structures in the era of cancer immunotherapy. Nat Rev Cancer. 2019;19(6):307–325.
16. Helmink BA, Reddy SM, Gao J, et al. B cells and tertiary lymphoid structures promote immunotherapy response. Nature. 2020;577(7791):549–555.
17. Cha YJ, Kim EY, Choi YJ, et al. Accumulation of plasmacytoid dendritic cell is associated with a treatment response to DNA-damaging treatment and favorable prognosis in lung adenocarcinoma. Front Immunol. 2023;14:1154881.
18. Kim EY, Cha YJ, Lee SH, et al. Early lung carcinogenesis and tumor microenvironment observed by single-cell transcriptome analysis. Transl Oncol. 2022;15(1):101277.
19. Cha YJ, Lee EH, Kim CY, et al. Anticancer Treatment Influences TREM2 in Tumor-Associated Macrophages in Lung Cancer. Cancer Res Treat. 2026;58(2):465–480.
20. Xing X, Yang F, Huang Q, et al. Decoding the multicellular ecosystem of lung adenocarcinoma manifested as pulmonary subsolid nodules by single-cell RNA sequencing. Sci Adv. 2021;7(5):eabd9738.
21. Kim N, Kim HK, Lee K, et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. Nat Commun. 2020;11(1):2285.
22. Satpathy S, Clark NM, Chen YJ, et al. Integrative analysis of lung adenocarcinoma across diverse ethnicities and exposures. Cancer Cell. 2025;43(9):1731–1757.e17.
23. Chen J, Yang H, Teo ASM, et al. Genomic landscape of lung adenocarcinoma in East Asians. Nat Genet. 2020;52(2):177–186.
24. Collisson EA, Campbell JD, Brooks AN, et al. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014;511(7511):543–550.
25. Elfving H, Yu H, Fessehatsion KK, et al. Spatial distribution of tertiary lymphoid structures in the molecular and clinical context of non-small cell lung cancer. Cell Oncol (Dordr). 2025;48(3):801–813.
26. Korsunsky I, Millard N, Fan J, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods. 2019;16(12):1289–1296.
27. Sikkema L, Ramírez-Suástegui C, Strobl DC, et al. An integrated cell atlas of the lung in health and disease. Nat Med. 2023;29(6):1563–1577.
28. Sautès-Fridman C, Lawand M, Giraldo NA, et al. Tertiary Lymphoid Structures in Cancers: Prognostic Value, Regulation, and Manipulation for Therapeutic Intervention. Front Immunol. 2016;7:407.

29. Saint-André V, Charbit B, Biton A, et al. Smoking changes adaptive immunity with persistent effects. Nature. 2024;626(8000):827–835.
30. Fenercioglu AK, Uzun H, Unal DO. The Convergent Immunopathogenesis of Cigarette Smoke Exposure: From Oxidative Stress to Epigenetic Reprogramming in Chronic Disease. Int J Mol Sci. 2026;27(1):187.
31. Yang M, Lei X, Ren D, et al. Epigenetic regulation of the tumor microenvironment in lung cancer: mechanism insights and therapeutic prospects. Front Immunol. 2025;16:1697428.
32. Dhillon P, Carroll M, Cheng H. Tumorigenesis and Tumor Microenvironment in Lung Cancer. Curr Issues Mol Biol. 2026;48(3):247.
33. Hernandez CP, Morrow K, Velasco C, et al. Effects of cigarette smoke extract on primary activated T cells. Cell Immunol. 2013;282(1):38–43.
34. Feng Y, Kong Y, Barnes PF, et al. Exposure to cigarette smoke inhibits the pulmonary T-cell response to influenza virus and Mycobacterium tuberculosis. Infect Immun. 2011;79(1):229–237.
35. Shan M, Yuan X, Song LZ, et al. Cigarette smoke induction of osteopontin (SPP1) mediates T(H)17 inflammation in human and experimental emphysema. Sci Transl Med. 2012;4(117):117ra119.
36. Park J, Lee J, Hur Y, et al. ETV5 promotes lupus pathogenesis and follicular helper T cell differentiation by inducing osteopontin expression. Proc Natl Acad Sci U S A. 2024;121(24):e2322009121.
37. Dai Z, Lin L, Xu Y, et al. Extracellular vesicle dynamics in COPD: understanding the role of miR-422a, SPP1 and IL-17 A in smoking-related pathology. BMC Pulm Med. 2024;24(1):173.
38. Garaud S, Dieu-Nosjean M-C, Willard-Gallo K. T follicular helper and B cell crosstalk in tertiary lymphoid structures and cancer immunotherapy. Nat Commun. 2022;13(1):2259.
39. Cho J-W, Park S, Kim G, et al. Dysregulation of TFH-B-TRM lymphocyte cooperation is associated with unfavorable anti-PD-1 responses in EGFR-mutant lung cancer. Nat Commun. 2021;12(1):6068.
40. Vahidian F, Lamaze FC, Bouffard C, et al. CXCL13 Positive Cells Localization Predict Response to Anti-PD-1/PD-L1 in Pulmonary Non-Small Cell Carcinoma. Cancers (Basel). 2024;16(4):708.
41. Xu S, Li D, Ning T, et al. High expression of CXCL13 predicts a favorable response to immunotherapy by upregulating CXCR5+CD8+ T-cell infiltration in gastric cancer. Front Immunol. 2025;16:1551259.
42. Ouyang W, Peng Q, Lai Z, et al. Synergistic role of activated CD4(+) memory T cells and CXCL13 in augmenting cancer immunotherapy efficacy. Heliyon. 2024;10(5):e27151.
43. Schumacher TN, Thommen DS. Tertiary lymphoid structures in cancer. Science. 2022;375(6576):eabf9419.
44. Ukita M, Hamanishi J, Yoshitomi H, et al. CXCL13-producing CD4+ T cells accumulate in the early phase of tertiary lymphoid structures in ovarian cancer. JCI Insight. 2022;7(12):e157215.
45. Groeneveld CS, Fontugne J, Cabel L, et al. Tertiary lymphoid structures marker CXCL13 is associated with better survival for patients with advanced-stage bladder cancer treated with immunotherapy. Eur J Cancer. 2021;148:181–189.
46. Zhang Y, Liu G, Zeng Q, et al. CCL19-producing fibroblasts promote tertiary lymphoid structure formation enhancing anti-tumor IgG response in colorectal cancer liver metastasis. Cancer Cell. 2024;42(8):1370–1385.e9.
47. Taniguchi K, Yoshikawa T, Yanagita M. The roles of tertiary lymphoid structures in orchestrating immune responses in peripheral organs. Inflamm Regen. 2025;45(1):28.
48. Yoshikawa T, Yanagita M. Single-Cell Analysis Provides New Insights into the Roles of Tertiary Lymphoid Structures and Immune Cell Infiltration in Kidney Injury and Chronic Kidney Disease. Am J Pathol. 2025;195(1):40–54.
49. Sorin M, Karimi E, Rezanejad M, et al. Single-cell spatial landscape of immunotherapy response reveals mechanisms of CXCL13 enhanced antitumor immunity. J Immunother Cancer. 2023;11(2):e005545.
50. Thommen DS, Koelzer VH, Herzig P, et al. A transcriptionally and functionally distinct PD-1(+) CD8(+) T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade. Nat Med. 2018;24(7):994–1004.
51. Onder L, Papadopoulou C, Lütge A, et al. Fibroblastic reticular cells generate protective intratumoral T cell environments in lung cancer. Cell. 2025;188(2):430–446.e20.
52. Barone F, Gardner DH, Nayar S, et al. Stromal Fibroblasts in Tertiary Lymphoid Structures: A Novel Target in Chronic Inflammation. Front Immunol. 2016;7:477.
53. Guo S, Yuan J, Meng X, et al. Cancer-associated fibroblasts: Just on the opposite side of antitumour immunity? Int Immunopharmacol. 2023;122:110601.