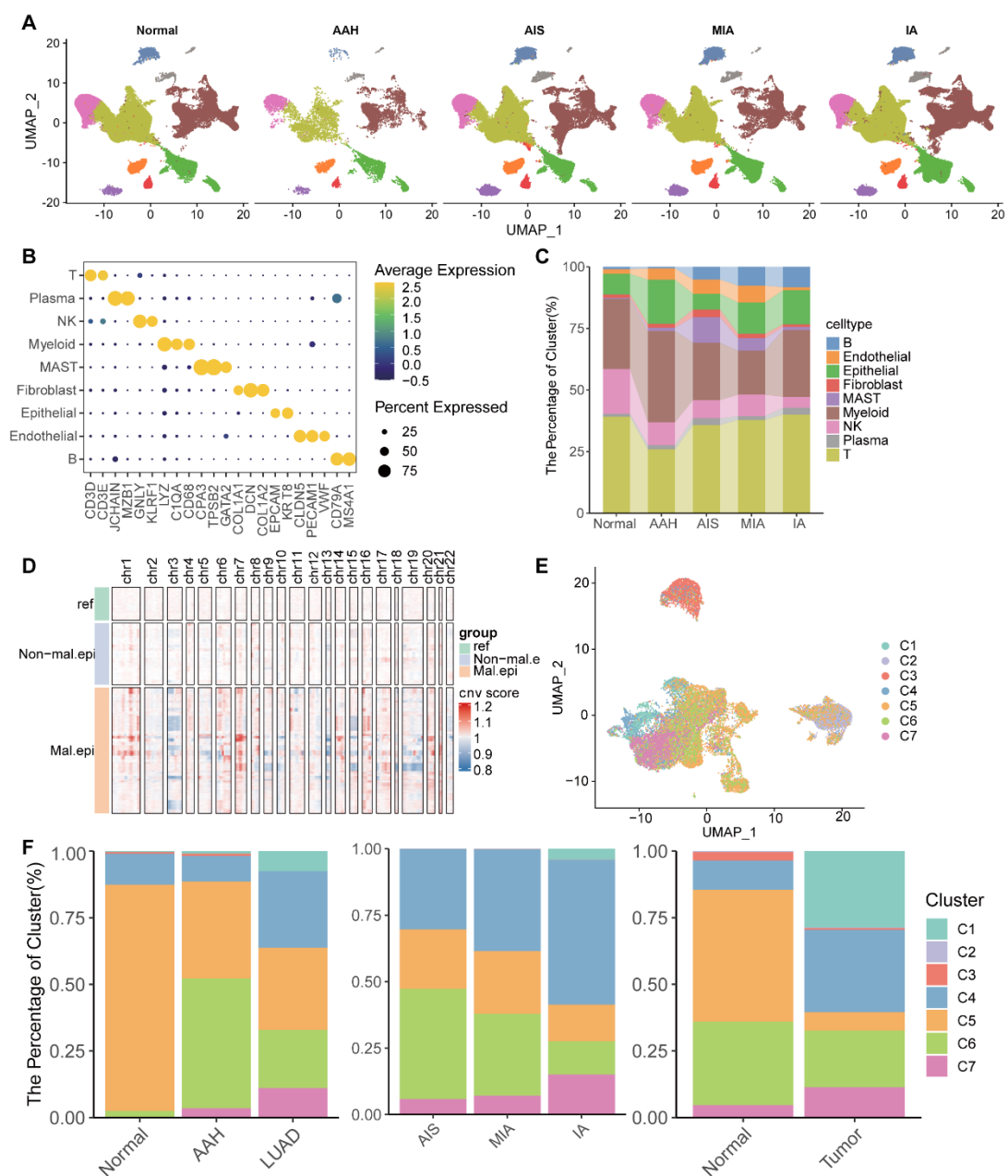


1 **Supplementary Figures**

2 **Figure S1. Global tumor microenvironment composition across GGN stages. A,**
3 The UMAP plot divide by group. **B,** Dot plot showing the markers used for cell-type
4 annotation. **C,** Bar plot showing the proportions of each cell type across different groups.
5 **D,** inferCNV heatmap showing inferred CNV patterns in the reference group (ref), non-
6 malignant epithelial cells (Non-mal.epi), and malignant epithelial cells (Mal.epi). **E,**
7 UMAP plot of malignant epithelial subclusters. **F,** Bar plot showing the proportions of
8 each subtype across different groups. Data for the left and middle panels were derived
9 from the cohort of Zhang et al, while data for the right panel were obtained from the
10 TCGA-LUAD database.

11



12

13

Figure S2. Characterization and functional enrichment analysis of malignant epithelial cells. **A**, Forest plot of the independent validation cohort from Zhang, Y. et al. **Statistical significance: *P < 0.05; **P < 0.01; ***P < 0.001. **B**, UMAP plot showing the distribution of cell cycle phases (G1, S, and G2/M) in malignant epithelial cells. **C**, Box plots showing the percentage of mitochondrial gene expression in different malignant epithelial cell subclusters. **D**, Box plots showing the proportions of C1 malignant epithelial cells across different groups. Each dot represents one sample. **E**, Bubble plot showing GO enrichment analysis of genes upregulated in malignant epithelial cells. **F**, Bubble plot showing KEGG enrichment analysis of genes upregulated in malignant epithelial cells.

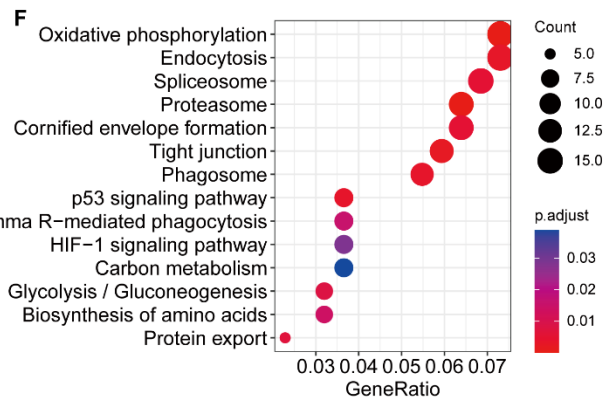
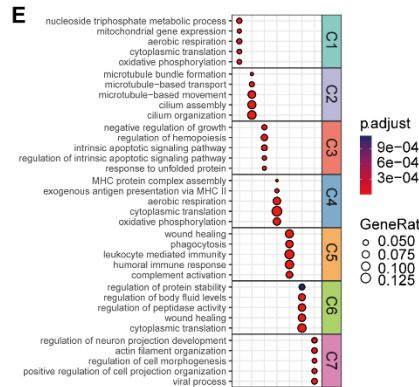
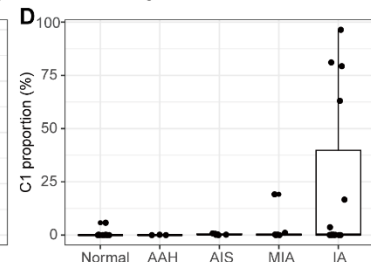
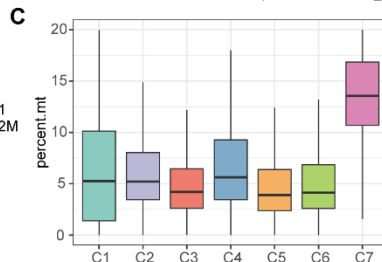
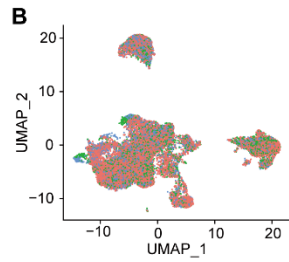
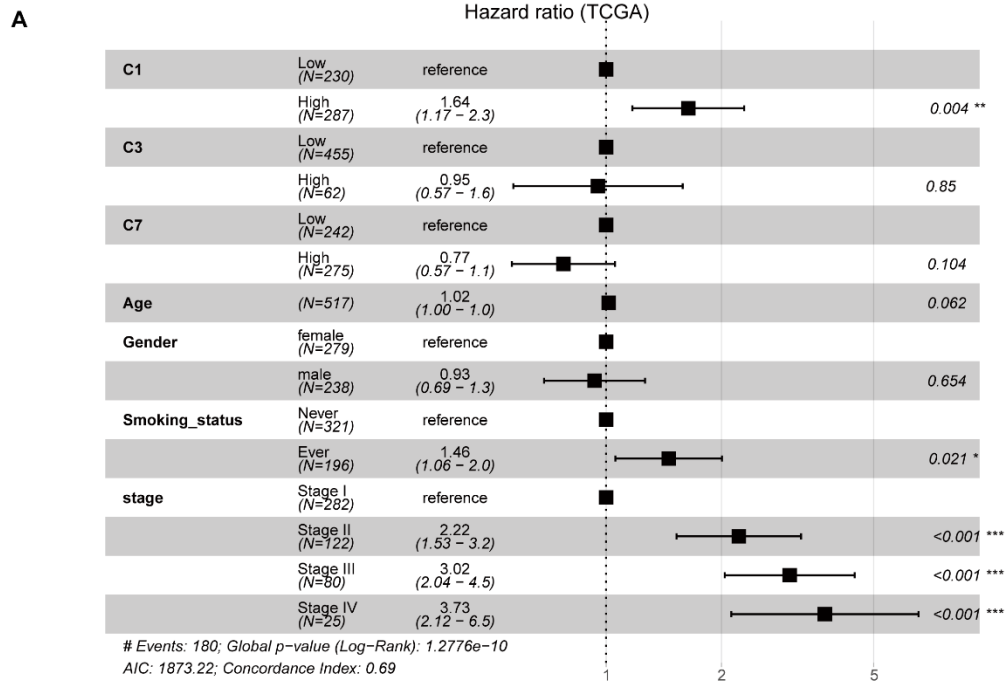
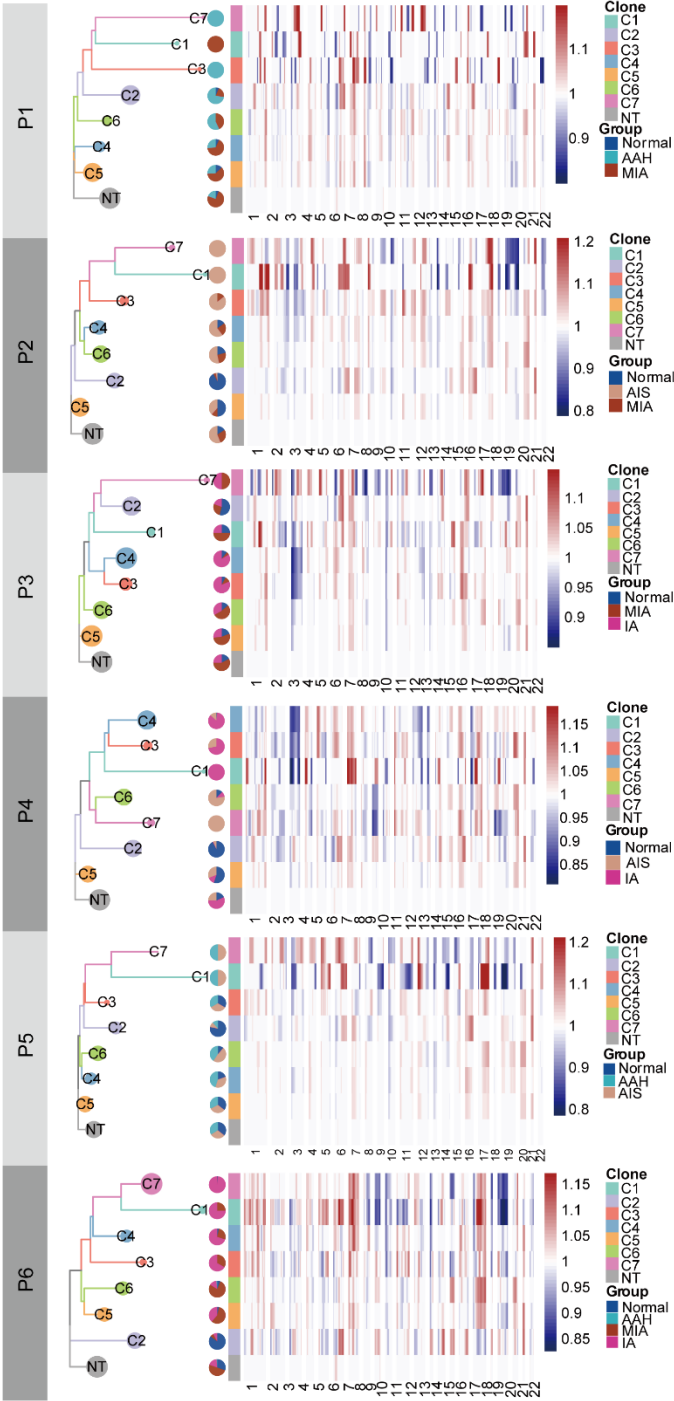


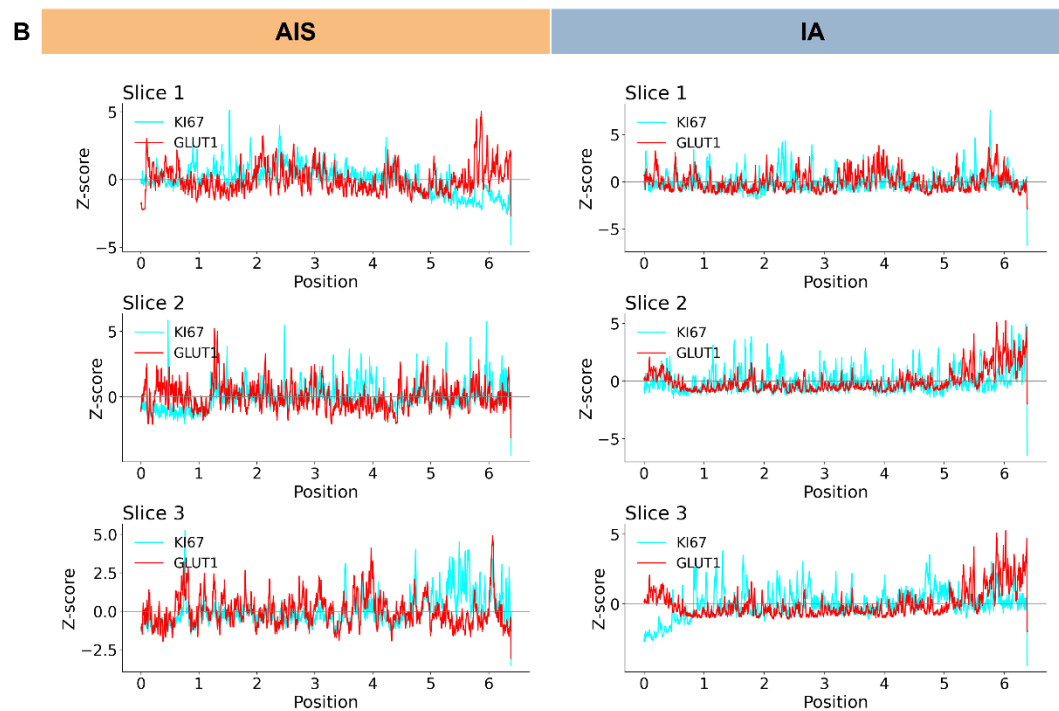
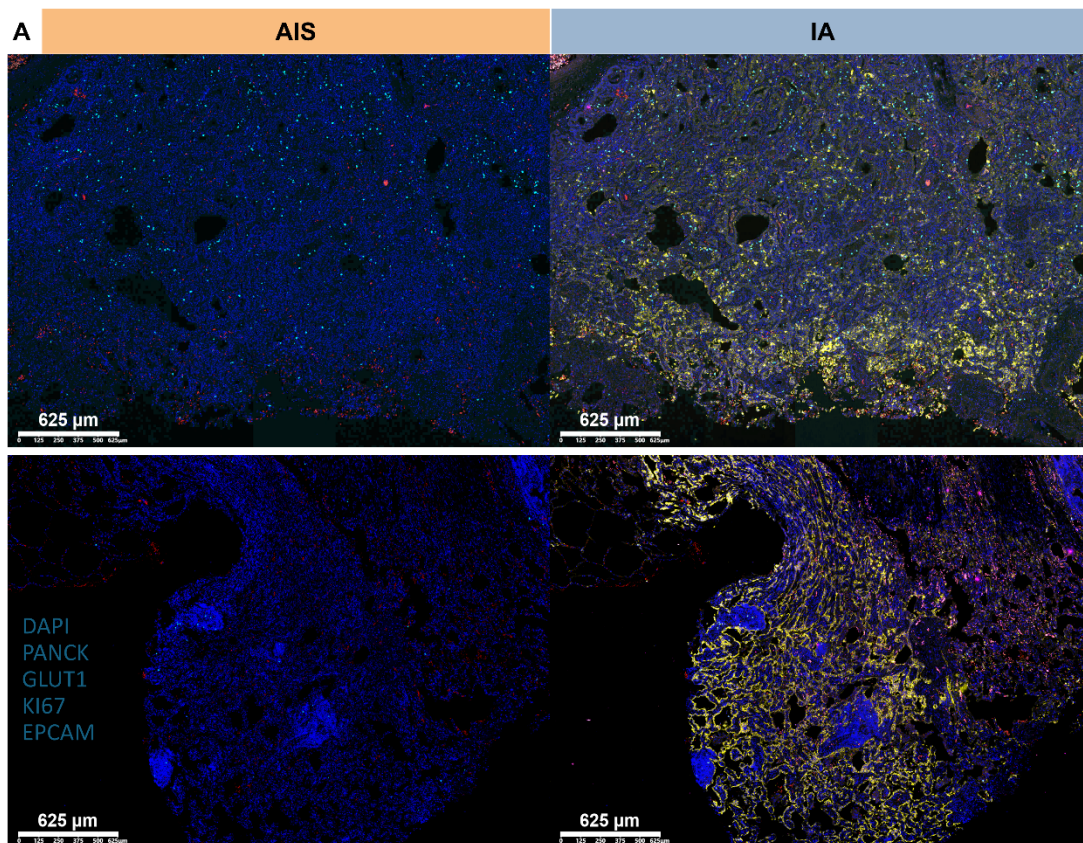
Figure S3. CNV-based phylogenetic analysis of tumor cells from multi-site lung cancer patients. **A**, CNV-based clustering of tumor cell from multi-site patients, with each panel representing one patient. The left panel shows a neighbor-joining (NJ) tree constructed from inferCNV profiles, and the right panel shows the corresponding inferred CNV patterns.



33

34

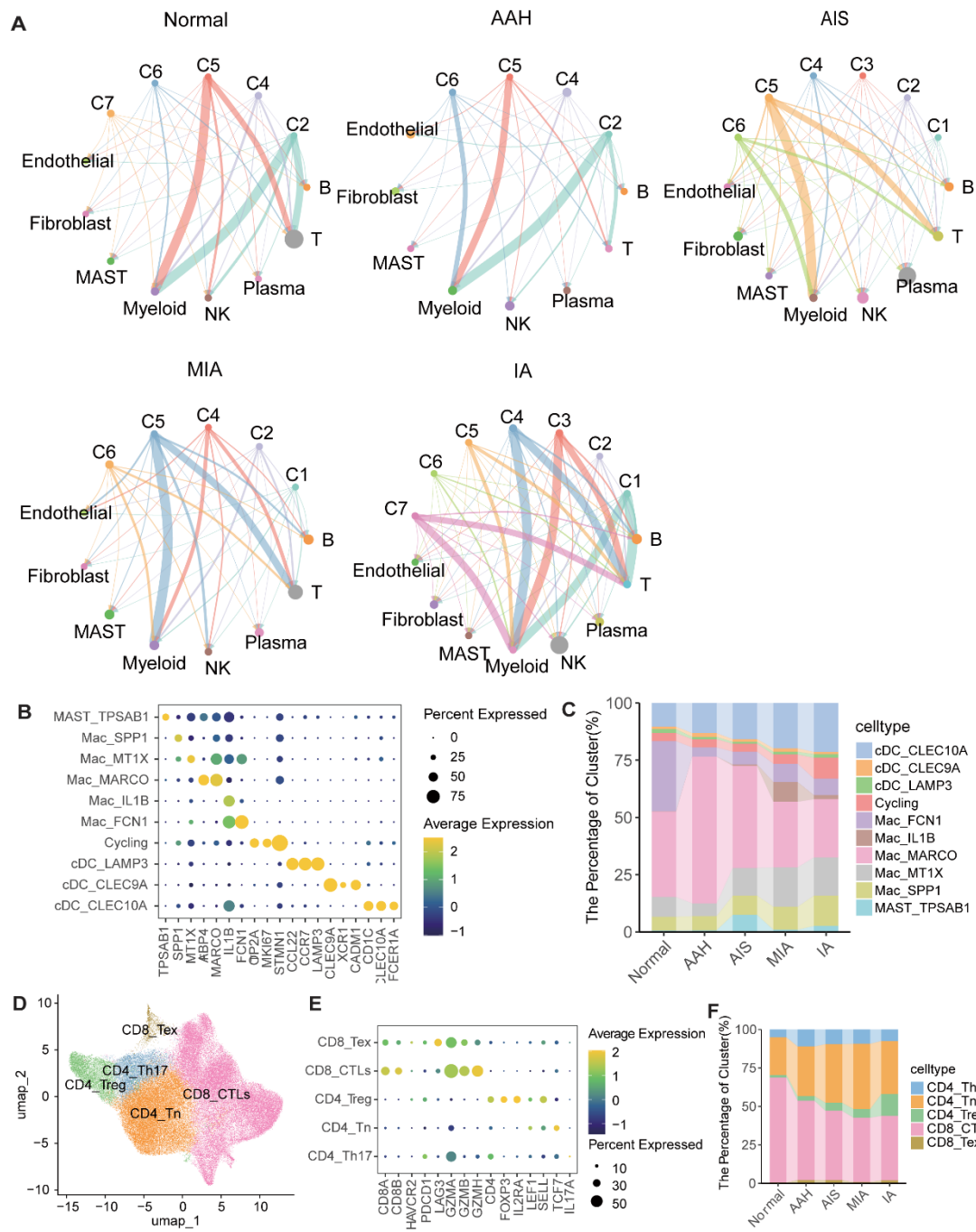
Figure S4. Multiplex immunofluorescence analysis of KI67 and GLUT1 expression in lung cancer tissues. **A**, Representative multiplex immunofluorescence images from two additional samples, including one stage IA lung adenocarcinoma and one AIS sample. The staining panel is the same as in the main figure, including DAPI, epithelial markers, the proliferation marker KI67, and the glycolysis marker GLUT1, showing cells co-expressing KI67 and GLUT1. **B**, Quantification of fluorescence signals from the additional samples. Line plots showing relative fluorescence levels of the indicated markers in IA and AIS samples.



44

45

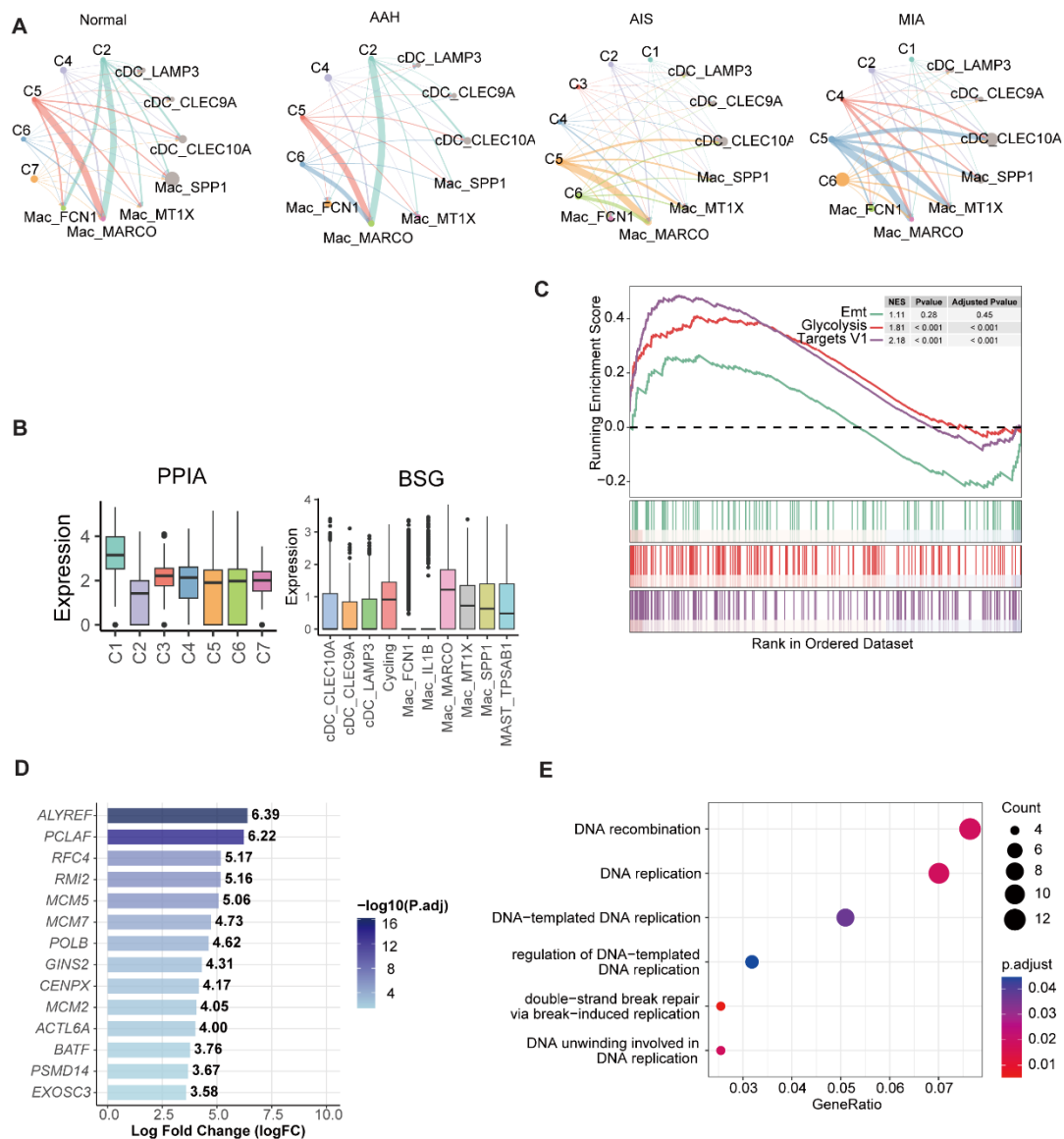
Figure S5. Characterization of immune cell subtypes and cell-cell communication in the lung tumor microenvironment. **A**, Circos plot showing cell-cell communication between malignant epithelial subtypes and other cell populations, with line thickness representing interaction strength. **B**, Bubble plot showing marker gene expression for myeloid cell subtype annotation. **C**, Bar plot showing the proportions of myeloid cell subtypes. **D**, UMAP plot of T-cell subtypes. **E**, Bubble plot showing marker gene expression for T-cell subtype annotation. **F**, Bar plot showing the proportions of T-cell subtypes.



55

56

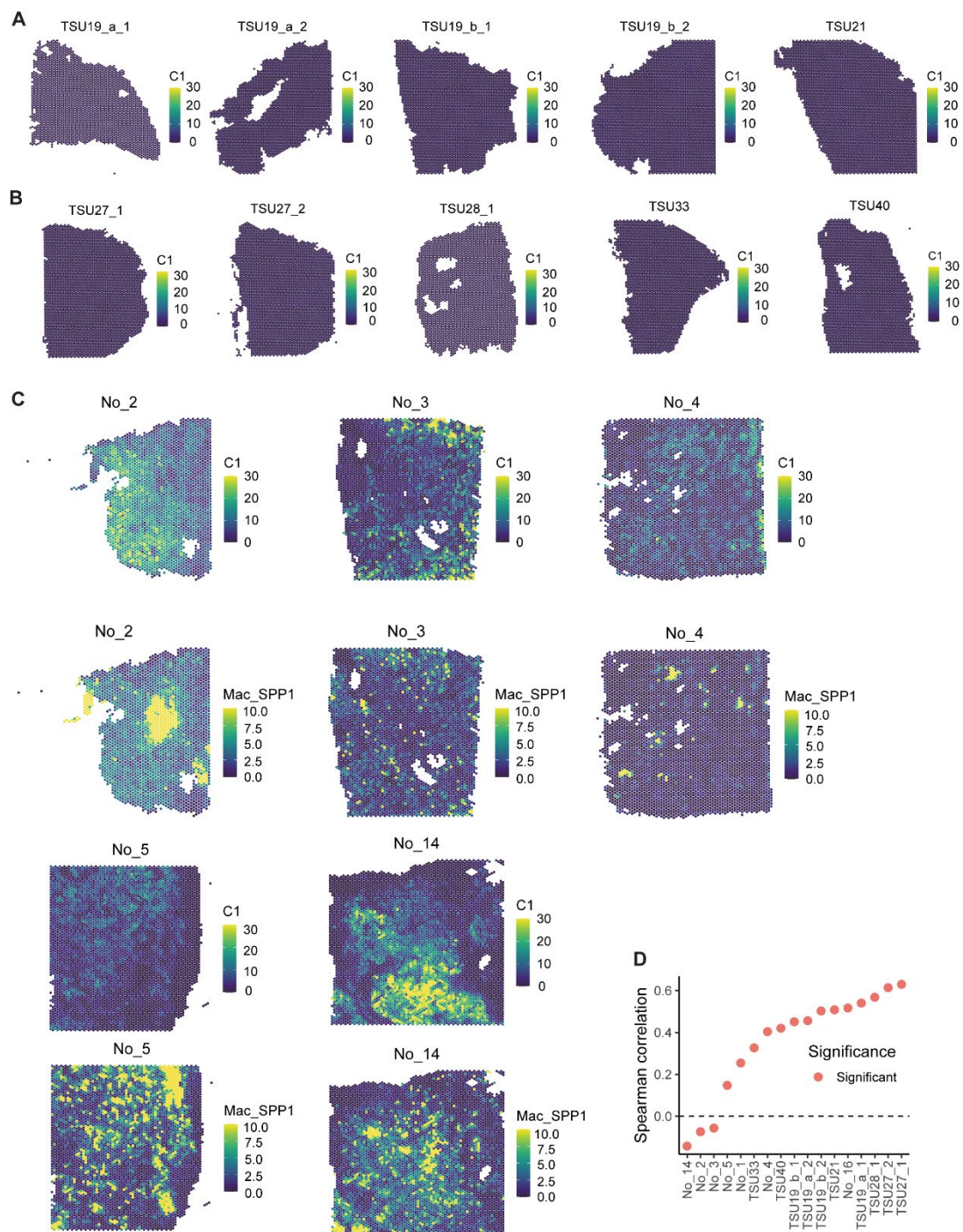
Figure S6. Analysis of PPIA and BSG expression and related functional pathways in malignant epithelial cells and myeloid cells. **A**, Circos plot showing inferred cell-cell communication between malignant epithelial subtypes and myeloid cell subtypes, with line thickness indicating interaction strength. **B**, Box plots showing PPIA expression across malignant epithelial subtypes (left) and BSG expression across myeloid cell subtypes (right). **C**, Curves showing GSEA enrichment results for differentially expressed genes between PPIA⁺BSG⁺ and PPIA⁻BSG⁻ cell group. **D**, Bar plot showing the top 15 genes predicted to be affected by simulated PPIA knockout in C1. Color intensity represents adjusted p-value. **E**, Bubble plot showing GO enrichment analysis of genes predicted to be affected by simulated PPIA knockout.



68

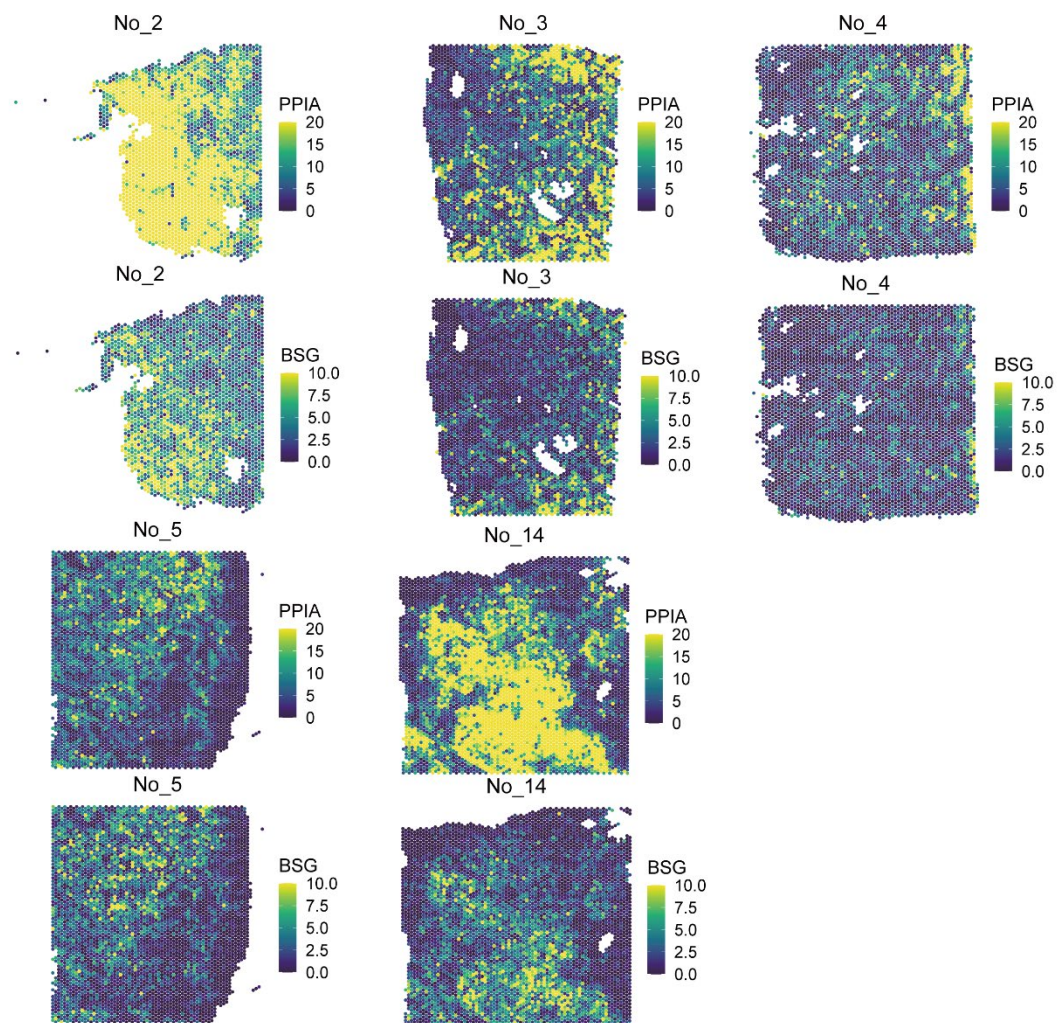
69

Figure S7. Spatial transcriptomic analysis of C1 tumor cells and Mac_SPP1 across stages of lung adenocarcinoma. **A**, Spatial transcriptomics map of stage AIS samples showing the spatial locations of C1 and Mac_SPP1. **B**, Spatial transcriptomics map of stage MIA samples showing the spatial locations of C1 and Mac_SPP1. **C**, Spatial transcriptomics map of stage IA samples showing the spatial locations of C1 and Mac_SPP1. **D**, Scatter plots showing the relationship between C1-scores and Mac_SPP1-scores across spatial transcriptomic samples. Each dot represents one sample, and dot color indicates the significance of the Spearman correlation.



79
80

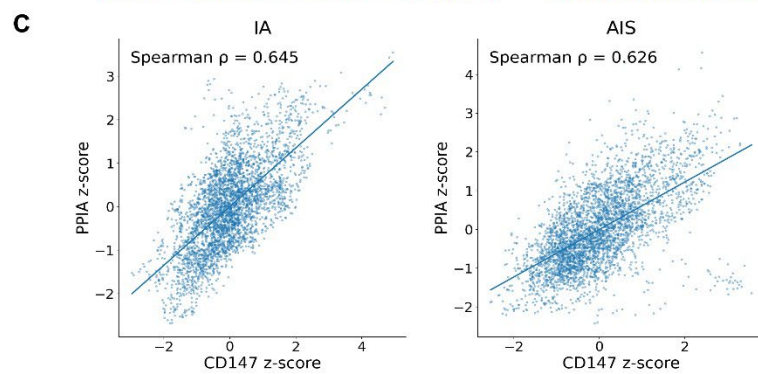
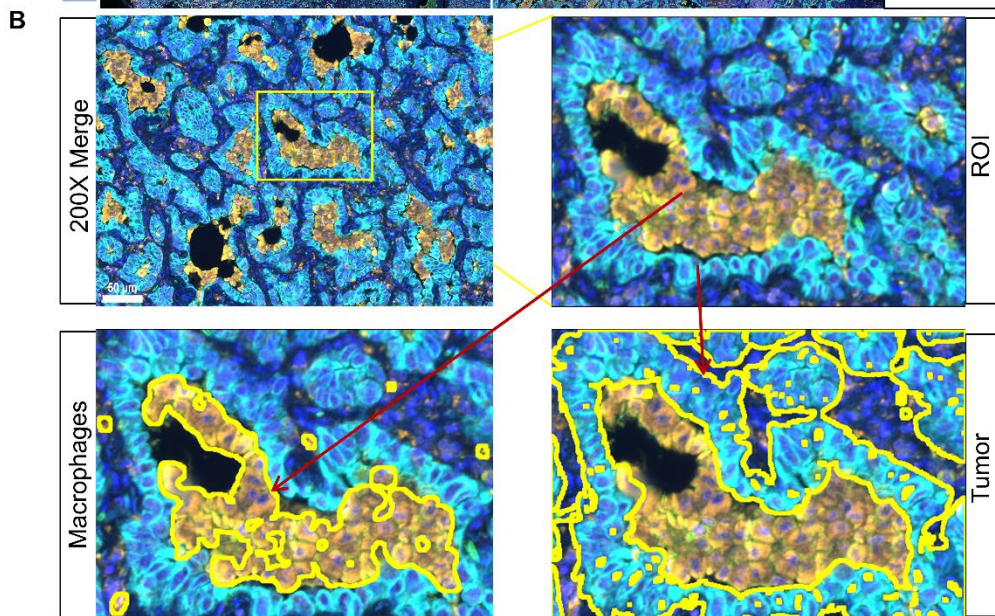
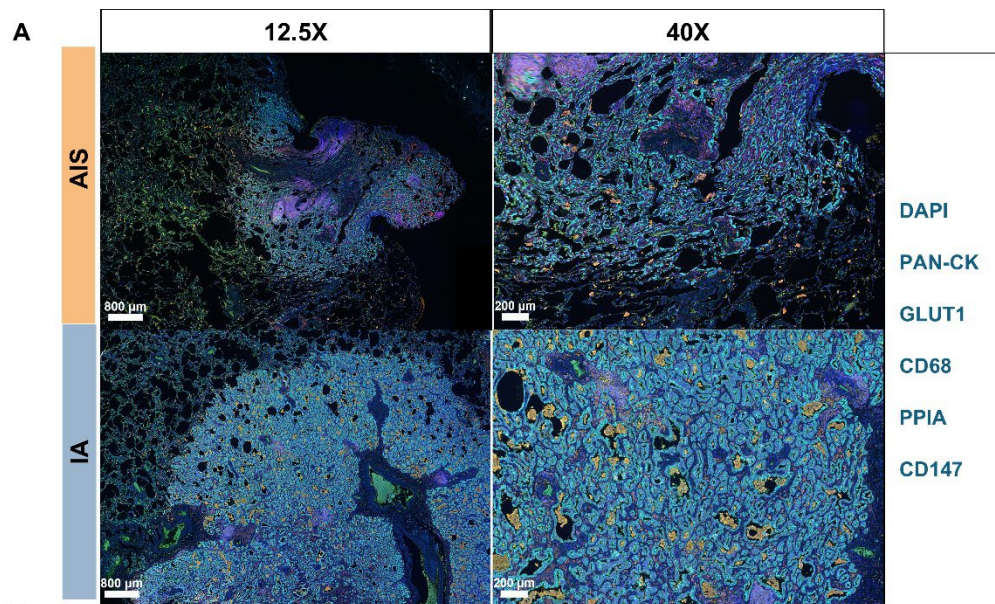
81 **Figure S8. Spatial transcriptomic maps of PPIA and BSG expression in stage IA**
82 **samples.** A, Spatial transcriptomic maps showing PPIA and BSG expression in stage
83 IA samples.
84



85

86

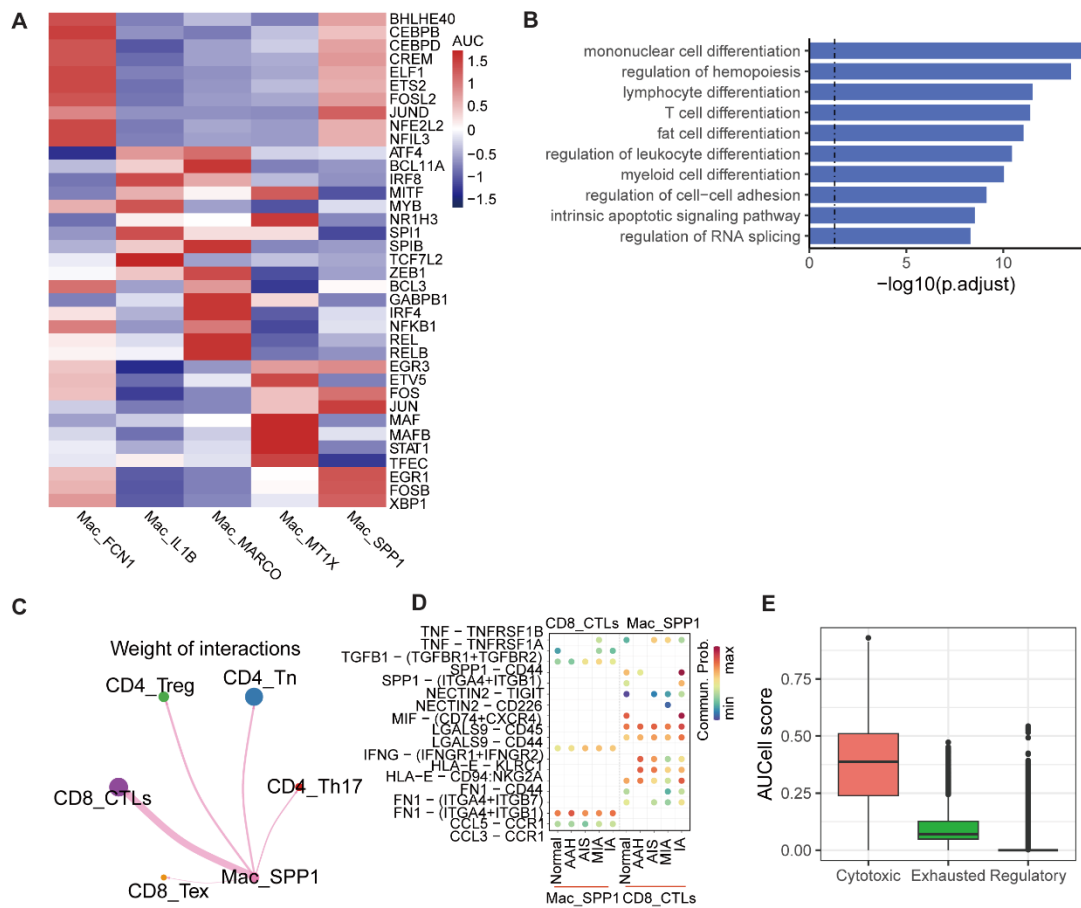
Figure S9. Multiplex immunofluorescence characterization of PPIA and CD147 expression in additional lung tumor samples. **A**, Representative mIF images from two additional lung tumor samples, including one stage IA and one AIS sample. Images are shown at low and high magnifications in the left and right panels, respectively. **B**, Immunofluorescence images showing PPIA and CD147 staining patterns in the additional samples. **C**, Scatter plots showing the correlation between PPIA and CD147 expression. The left panel shows IA samples, and the right panel shows AIS samples. Spearman correlation coefficients are indicated.



96

97

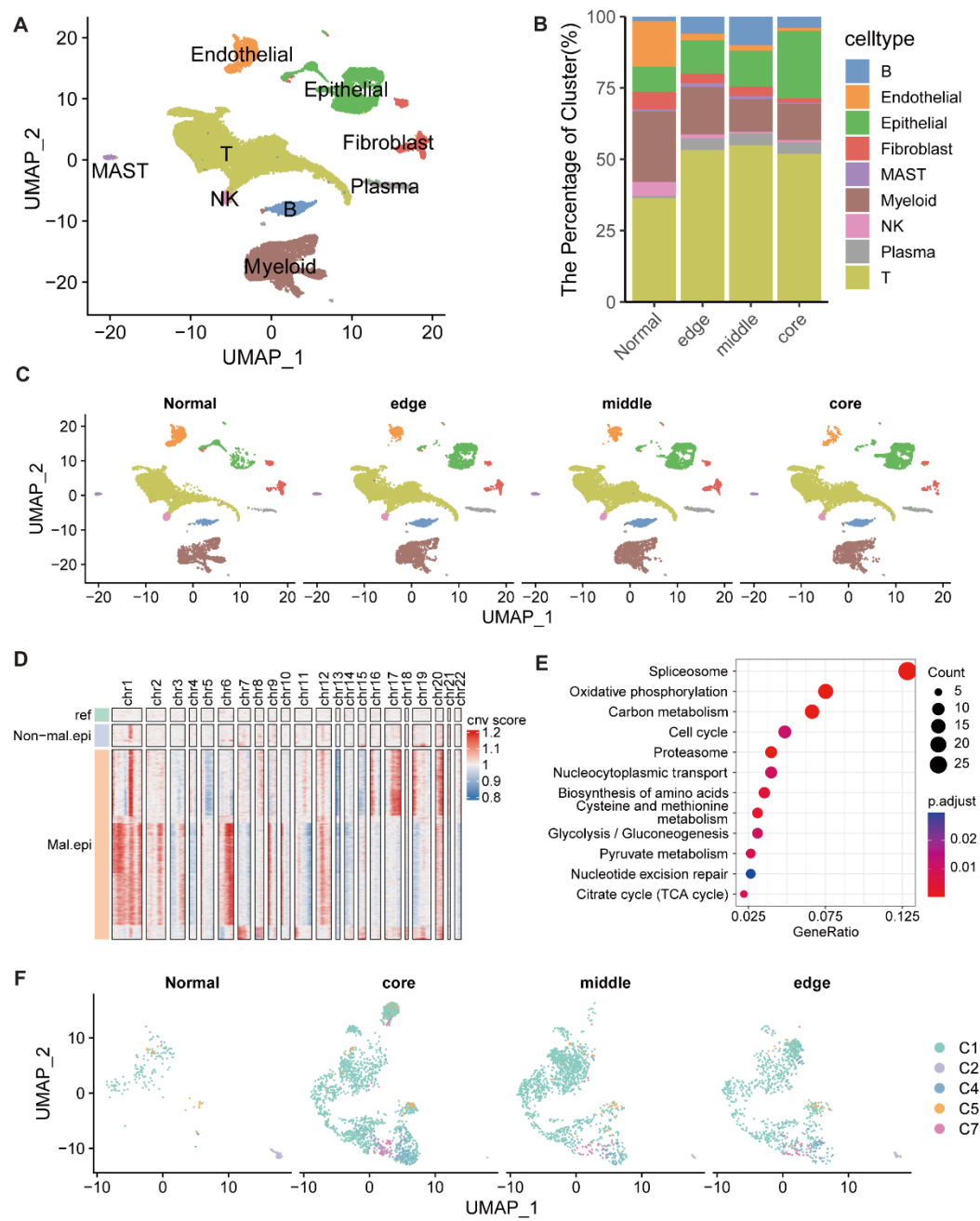
Figure S10. Transcriptional characteristics of Mac_SPP1 macrophages and analysis of communication with CD8⁺ cytotoxic T cells. **A**, Heatmap showing the transcription factor activity across myeloid cell subtypes. **B**, Bar plot showing the GO enrichment analysis of predicted target genes of transcription factors identified in Mac_SPP1. **C**, Circos plot showing inferred cell-cell communication between Mac_SPP1 and T-cell subtypes. Line thickness indicates interaction strength. **D**, Bubble plot showing inferred ligand-receptor interactions between Mac_SPP1 and CD8_CTLs. **E**, Box plots showing cytotoxicity, exhaustion, and regulatory status scores of CD8_CTLs.



108

109

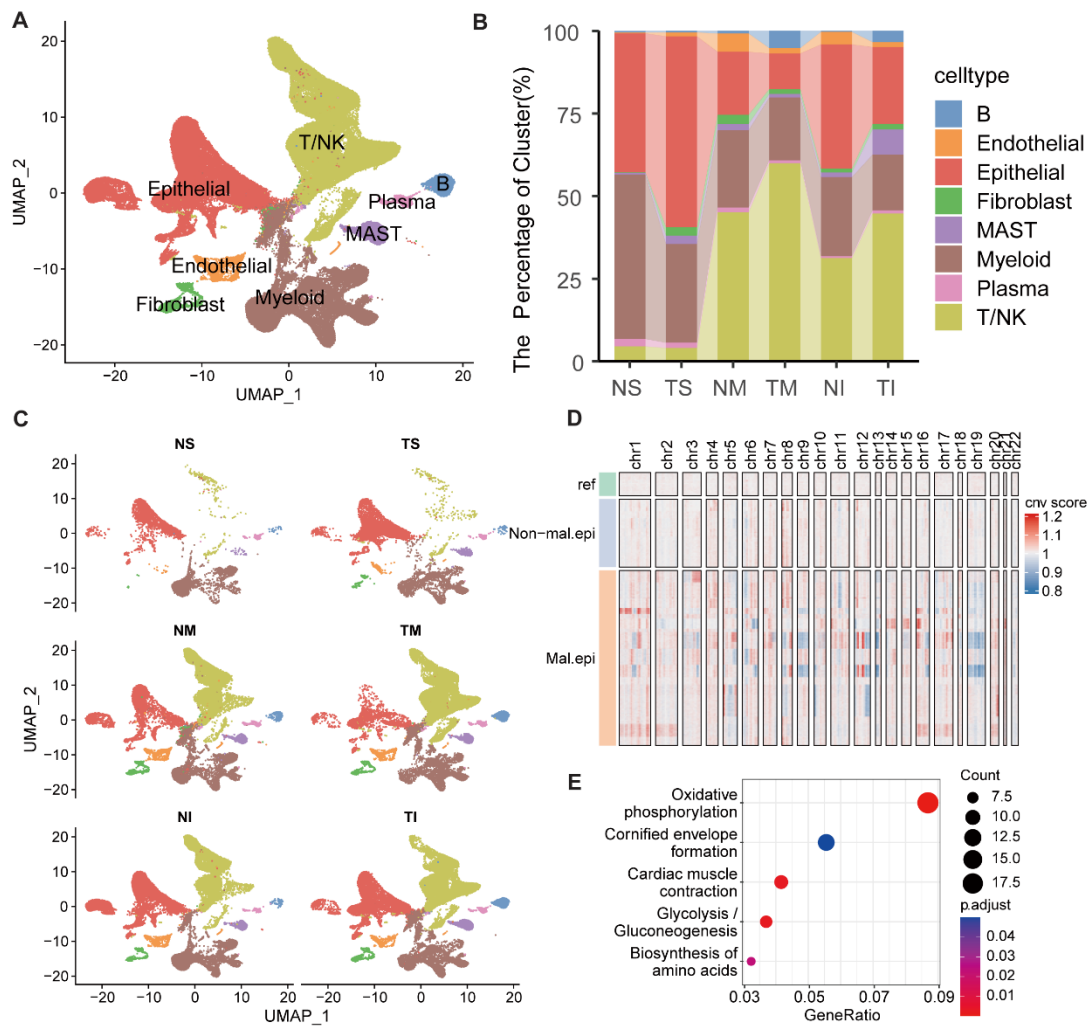
Figure S11. Spatial profiling and functional characterization of C1 malignant epithelial cells in lung tumors. **A**, UMAP plot of single-cell data from different spatial regions of the tumor. **B**, Bar plot showing the proportions of each cell subtype. **C**, UMAP plot of cell subtypes across different groups. **D**, inferCNV heatmap showing inferred CNV patterns in the reference group (ref), non-malignant epithelial cells (Non-mal.epi), and malignant epithelial cells (Mal.epi). **E**, Bubble plot showing GO enrichment analysis of genes upregulated in C1. **F**, UMAP plot of malignant epithelial cells across different groups.



119

120

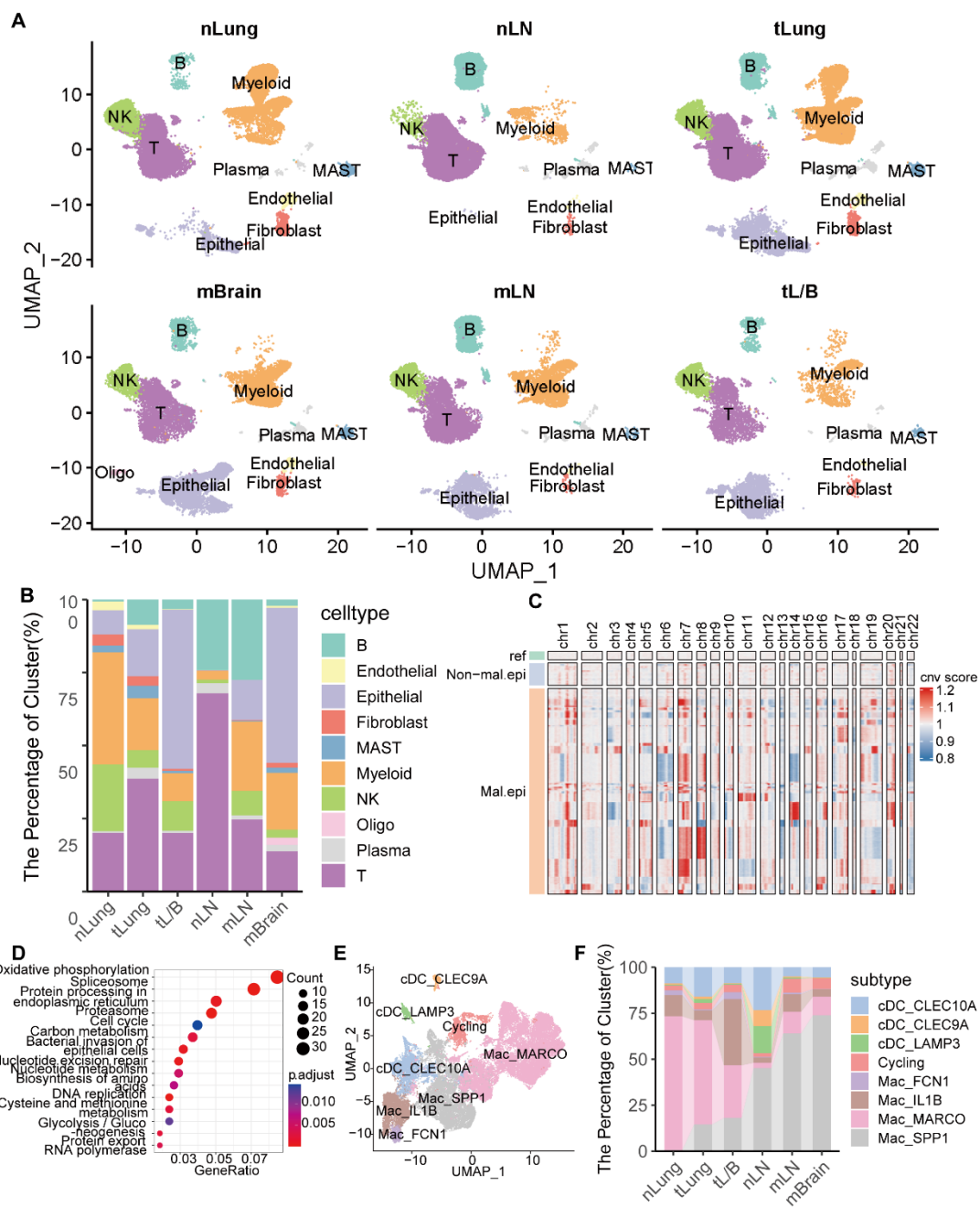
Figure S12. Single-cell characterization of malignant epithelial cells and cellular composition in MPLC. **A**, UMAP plot showing cell-type annotations of the MPLC single-cell dataset. **B**, Bar plot showing the proportions of each cell type. **C**, UMAP plots of cell subtypes across different groups. **D**, inferCNV heatmap showing inferred CNV patterns in the reference group (ref), non-malignant epithelial cells (Non-mal.epi), and malignant epithelial cells (Mal.epi). **E**, Bubble plot showing GO enrichment analysis of genes upregulated in C1.



129

130

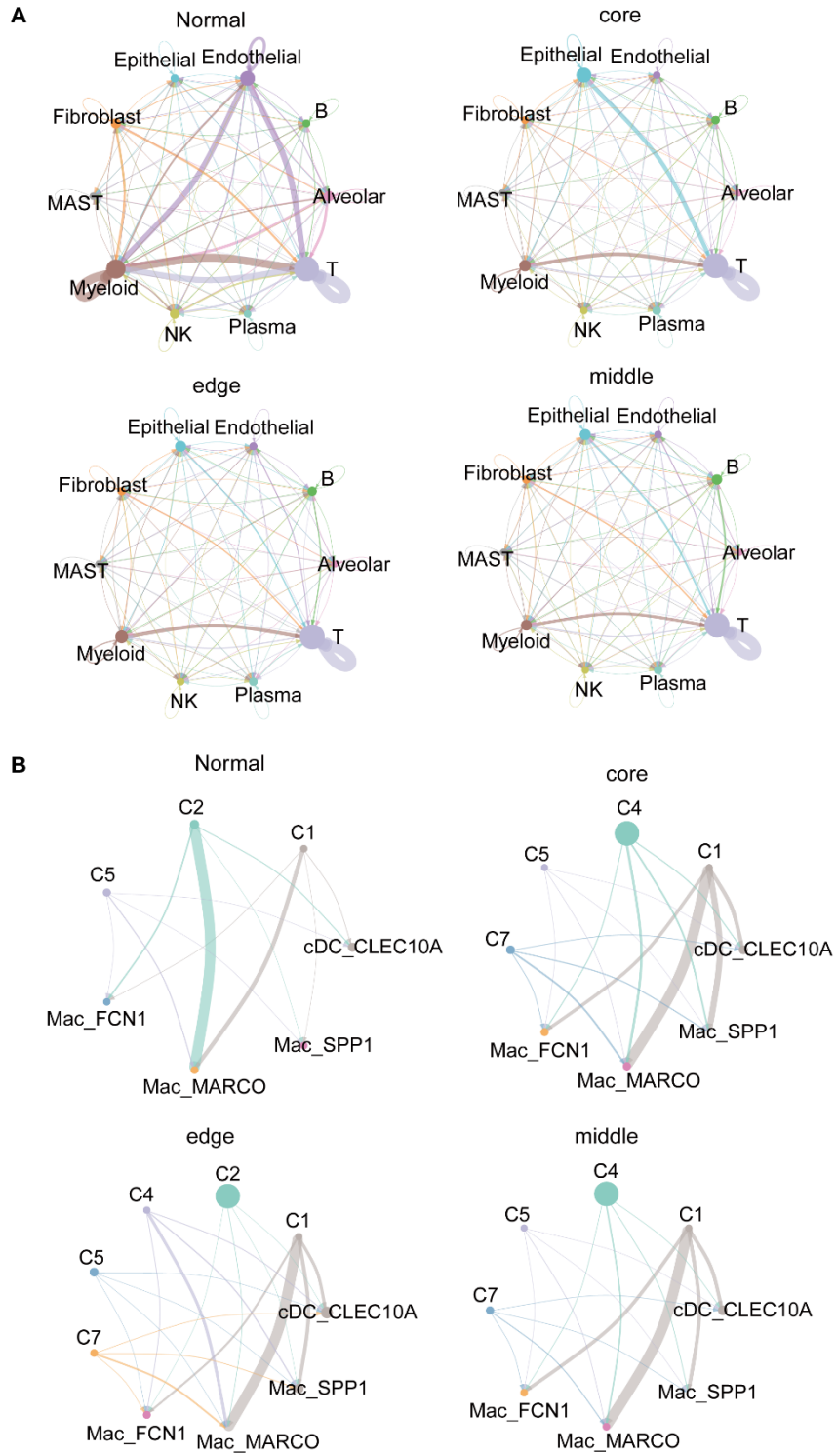
Figure S13. Single-cell characterization of malignant epithelial cells and myeloid subtypes in primary and metastatic lung cancer lesions. **A**, UMAP plot of the single-cell data from primary and metastatic lung cancer lesions. **B**, Bar plot showing the proportions of major cell types. **C**, inferCNV heatmap showing inferred CNV patterns in the reference group (ref), non-malignant epithelial cells (Non-mal.epi), and malignant epithelial cells (Mal.epi). **D**, Bubble plot showing GO enrichment analysis of genes upregulated in C1. **E**, UMAP plot of myeloid cell subtypes. **F**, Bar plot showing the proportions of myeloid cell subtypes.



140

141

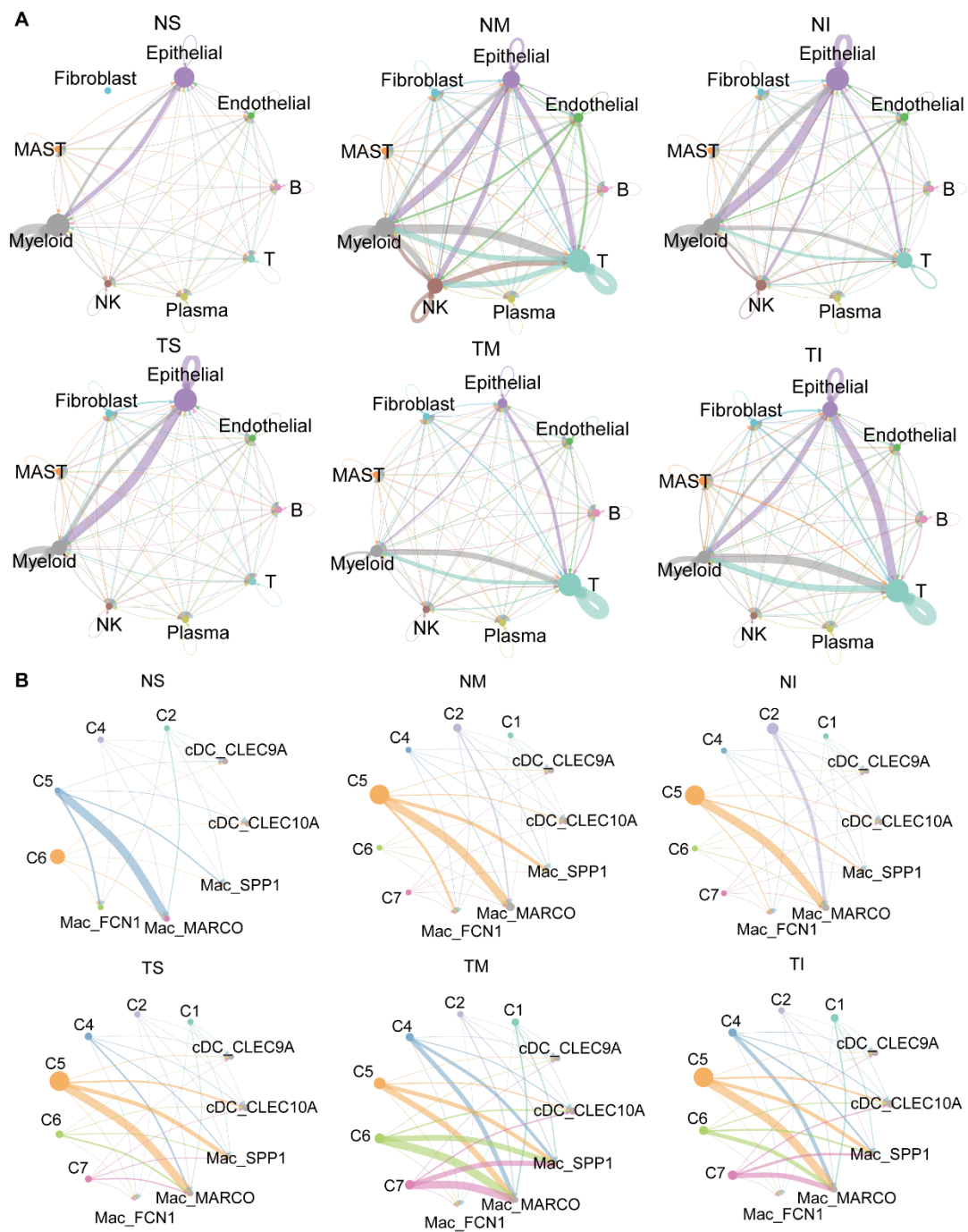
Figure S14. Cell-cell communication between myeloid cell and malignant epithelial cell subtypes. **A**, Circos plot showing inferred cell-cell communication among different cell types across groups. Line thickness indicates interaction strength. **B**, Circos plot showing inferred cell-cell communication between malignant epithelial cell and myeloid cell subtypes. Line thickness indicates interaction strength.



148

149

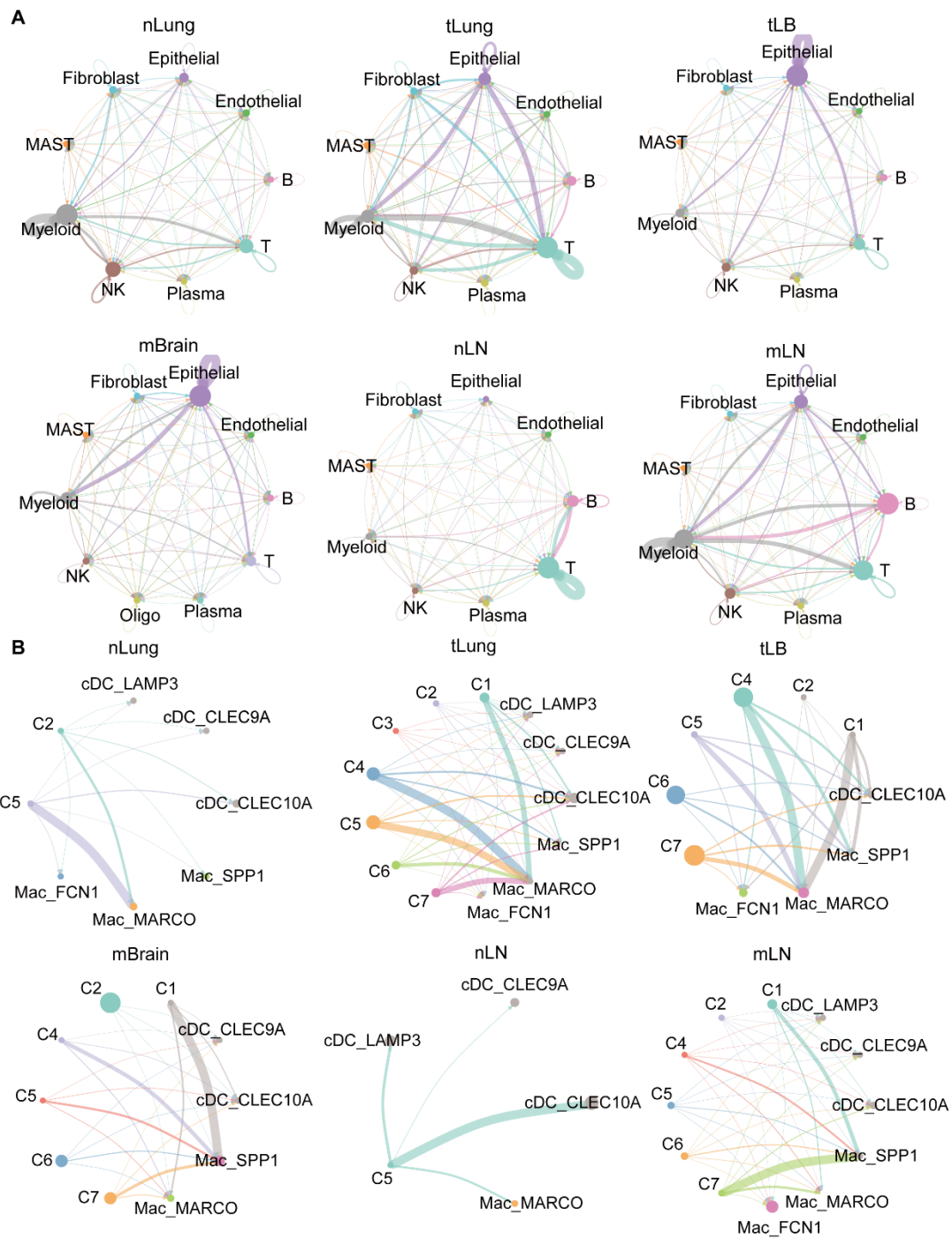
Figure S15. Cell-cell communication between myeloid cell and malignant epithelial cell subtypes. **A**, Circos plot showing inferred cell-cell communication among different cell types across groups. Line thickness indicates interaction strength. **B**, Circos plot showing inferred cell-cell communication between malignant epithelial cell and myeloid cell subtypes. Line thickness indicates interaction strength.



156

157

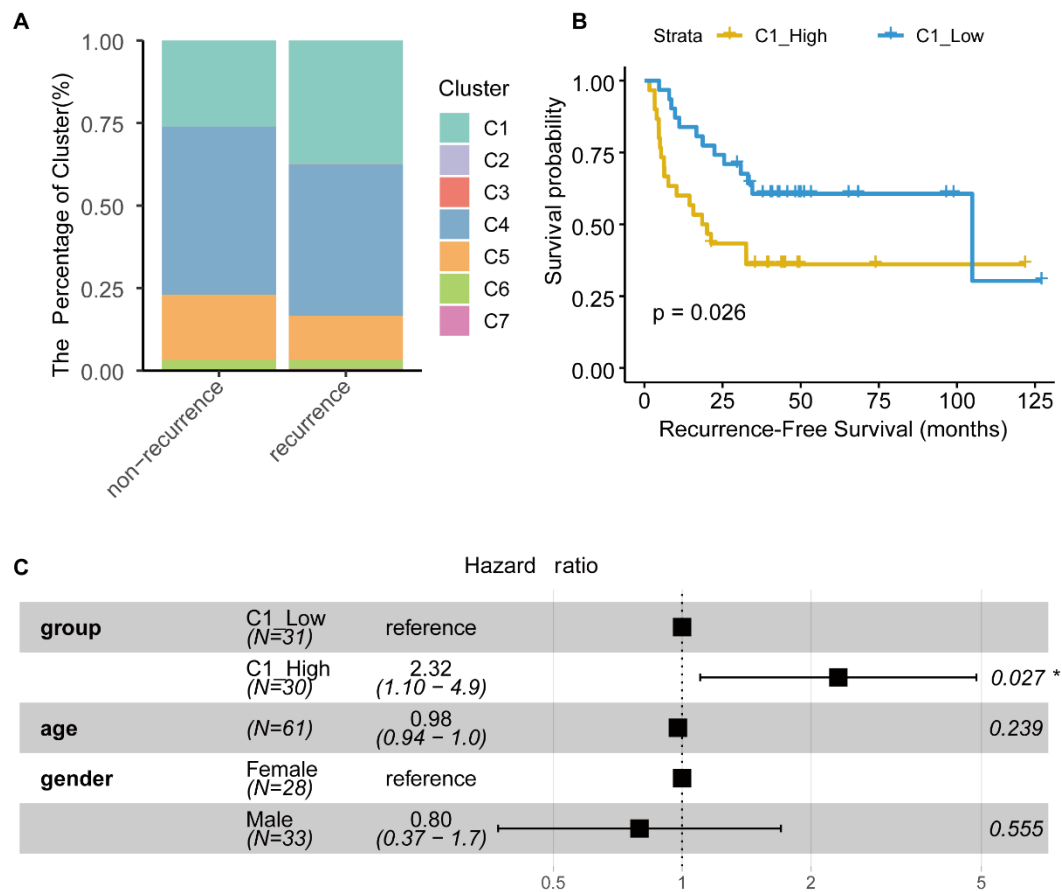
Figure S16. Cell-cell communication between myeloid cell and malignant epithelial cell subtypes. **A**, Circos plot showing inferred cell-cell communication among different cell types across groups. Line thickness indicates interaction strength. **B**, Circos plot showing inferred cell-cell communication between malignant epithelial cell and myeloid cell subtypes. Line thickness indicates interaction strength.



164

165

Figure S17. Deconvolution and survival analyses of malignant epithelial cells in the GSE8894 cohort. **A**, Deconvolution analysis of the bulk RNA-seq dataset showing the relative proportion of malignant epithelial cells across samples. **B**, Kaplan-Meier curves for recurrence-free survival stratified by malignant epithelial cell abundance. **C**, Multivariate Cox regression analysis including malignant epithelial cell abundance and clinical variables.



173

174

175

176