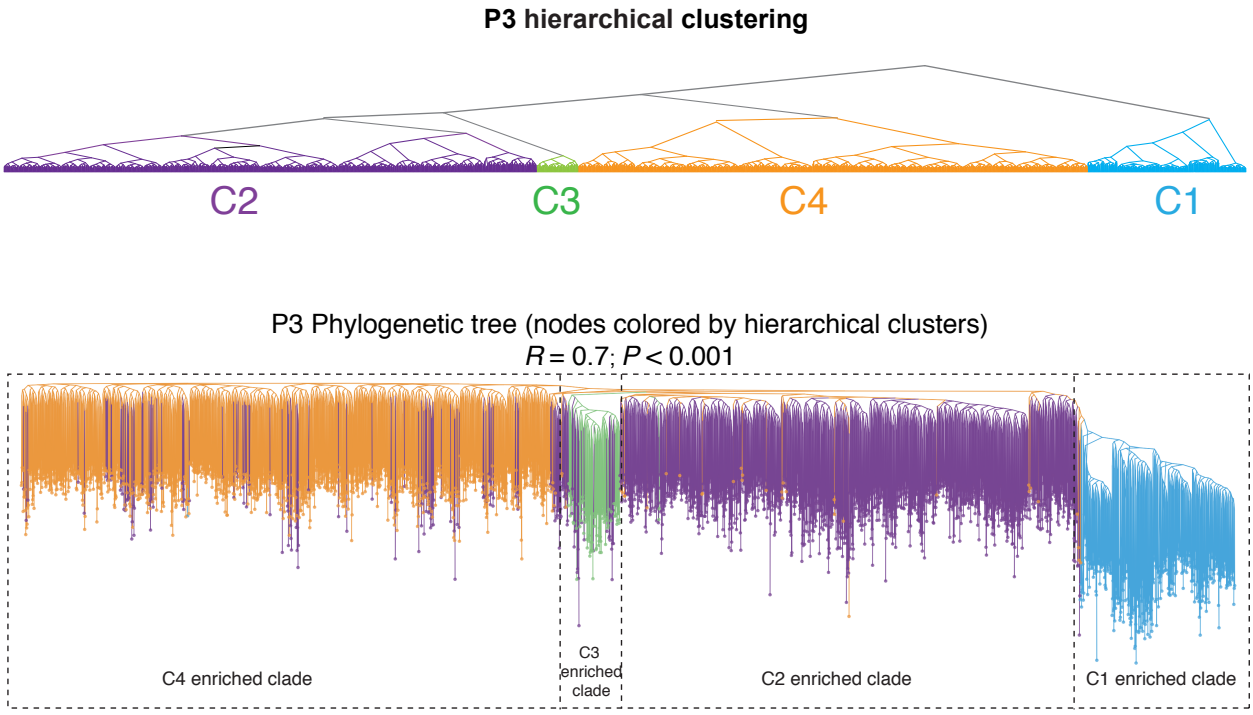
Resolving the spatial and cellular architecture of lung adenocarcinoma by multi-region single-cell sequencing

SUPPLEMENTARY DATA FILE 5

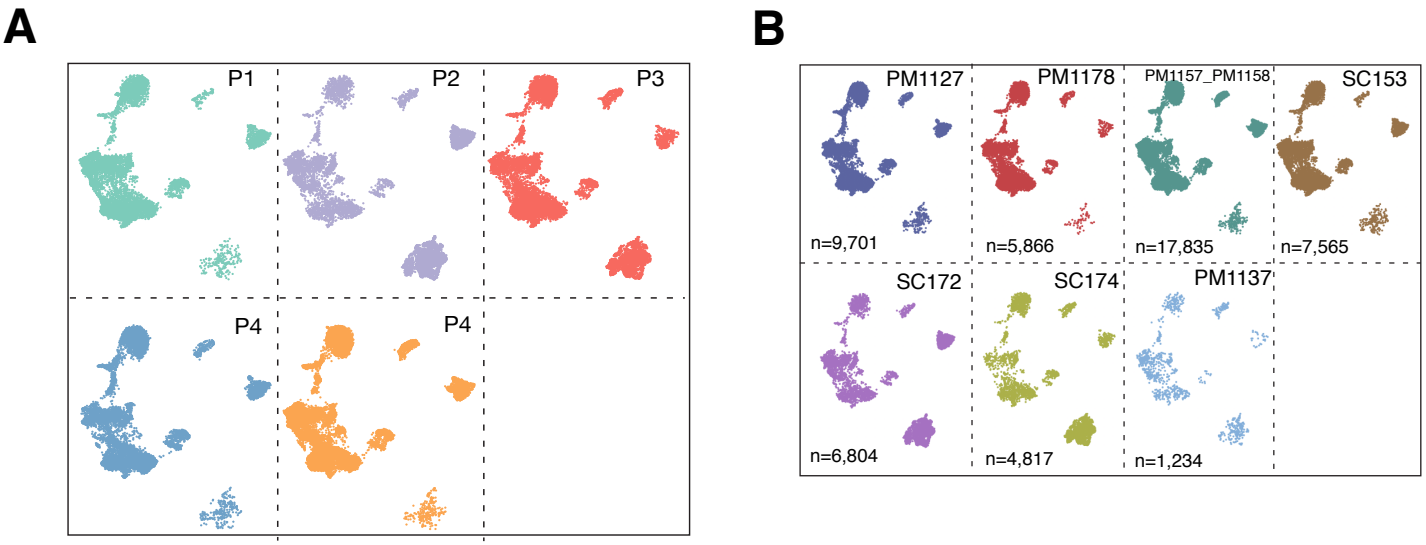
Supplementary Figures S10-S20

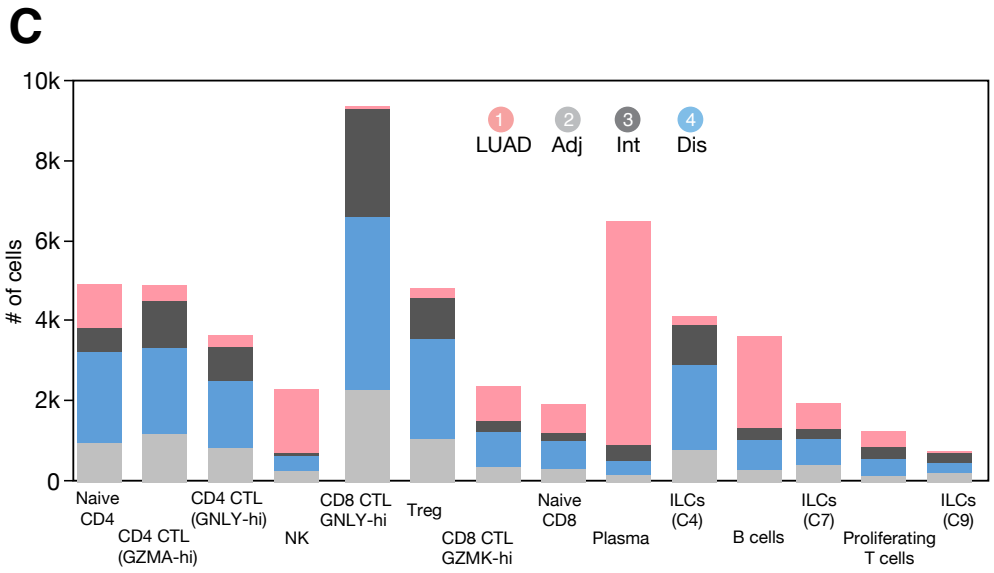
Supplementary Fig. S10



Supplementary Fig. S10. Phylogenetic reconstruction analysis using inferred large-scale CNVs in P3. Hierarchical clustering tree of P3 copy number profile (top) and the phylogenetic tree of the same profile (bottom). *P* – value was calculated using the Pearson's correlation test.

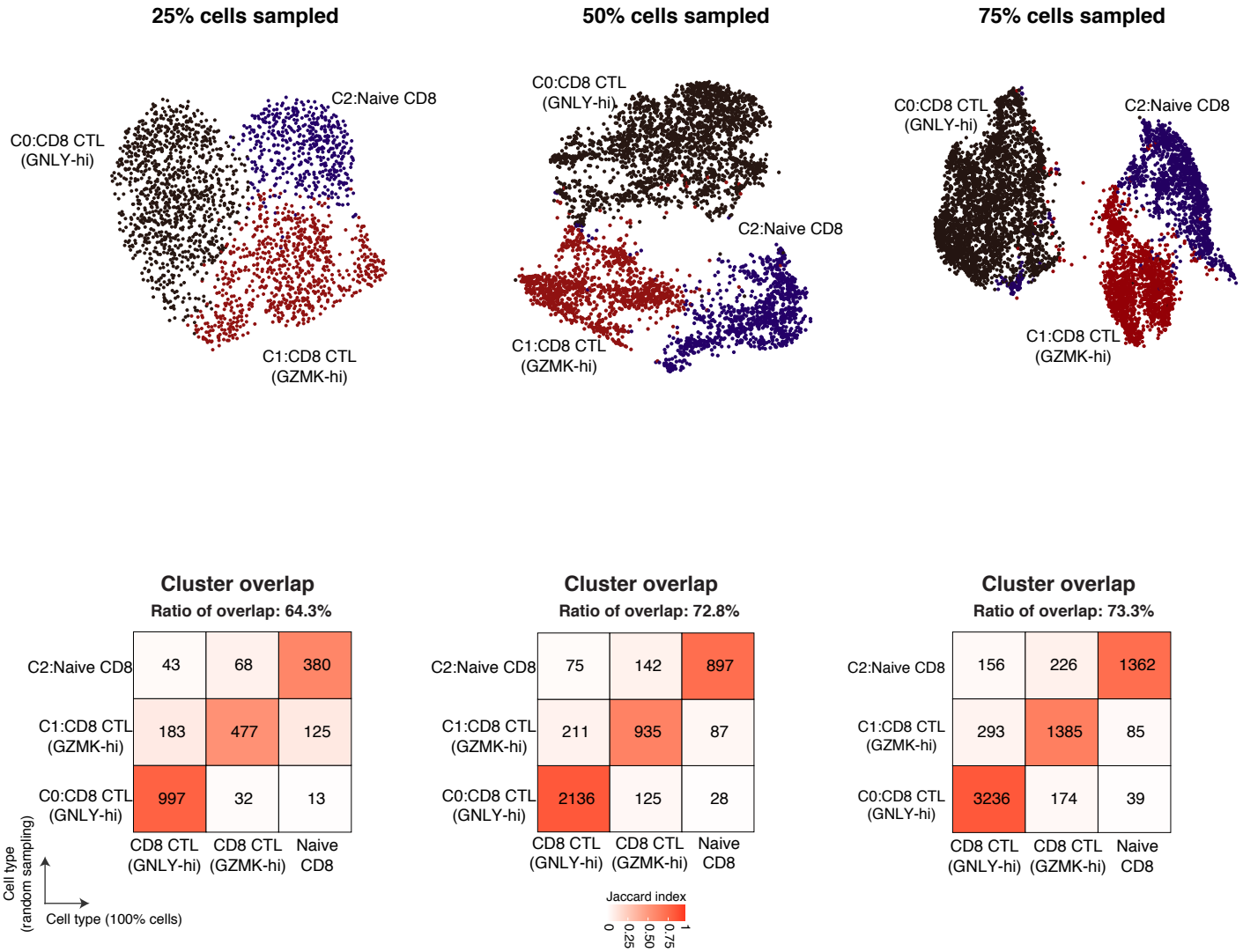
Supplementary Fig. S11





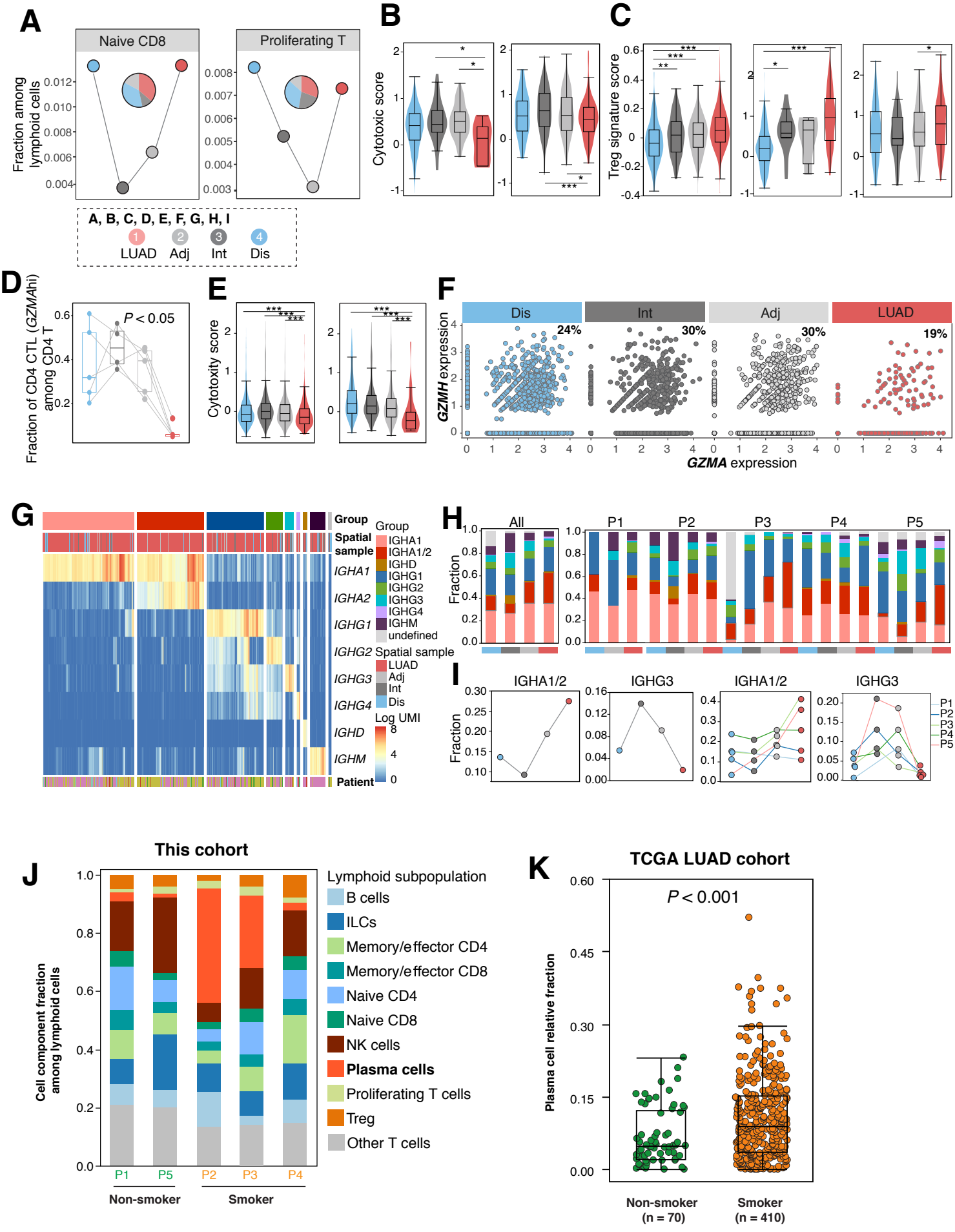
Supplementary Fig. S11. Quality control metrics of cells from lymphoid cell subsets. A-B, UMAP of lymphoid cells colored by patient ID (A) and sequencing/library batch (B). **C,** Bar plots showing the absolute numbers of each lymphoid cell subset in each spatial sample. Dis, distant normal; Int, intermediate normal; Adj, adjacent normal; LUAD, tumor tissue; CTL, cytotoxic T lymphocyte; Treg, T regulatory cell; ILC, innate lymphoid cell; NK, natural killer cell.

Supplementary Fig. S12



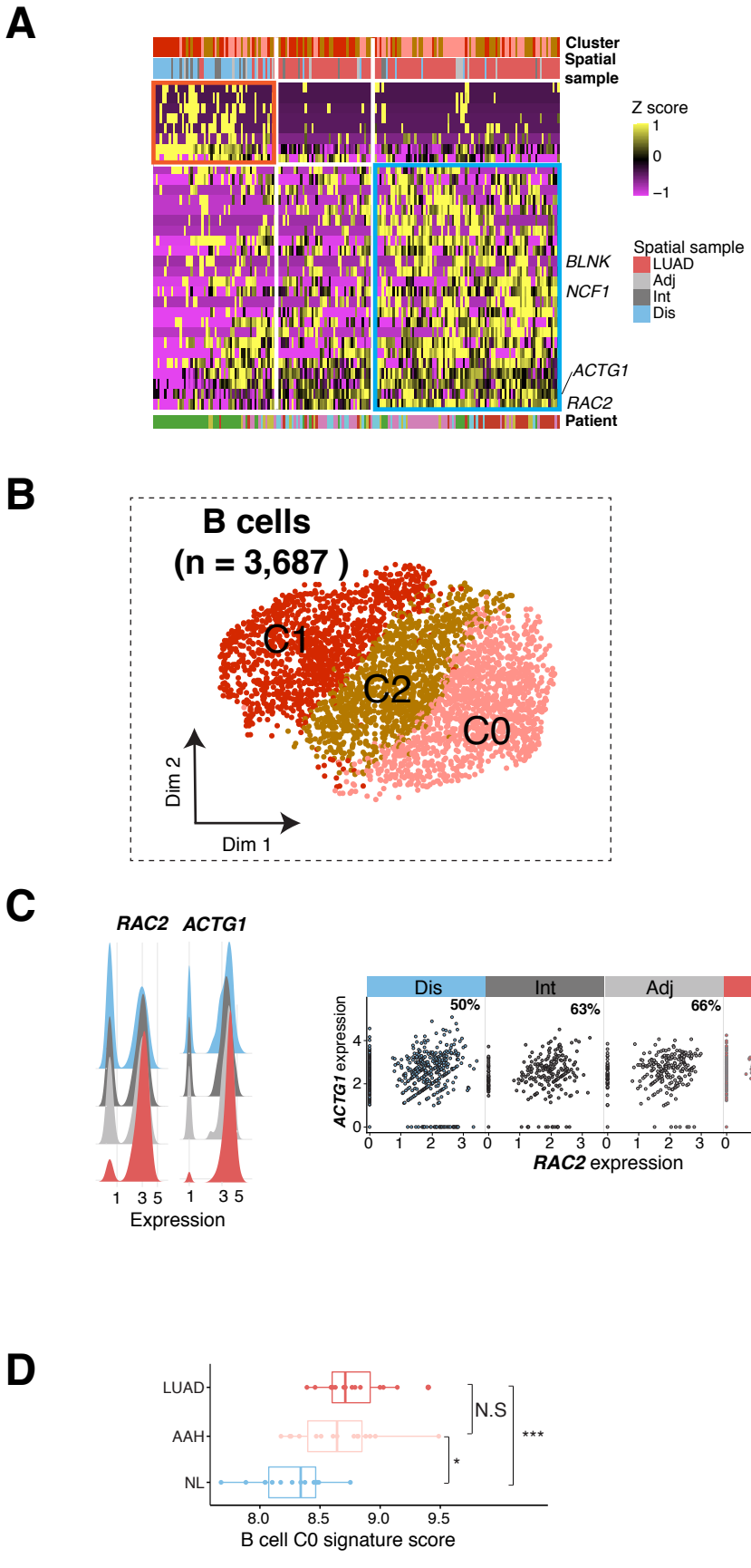
Supplementary Fig. S12. Analysis of robustness of CD8+ T cell clustering. UMAP plots (top) showing the clustering results when using 25% (top left), 50% (top middle), and 75% (top right) of randomly sampled CD8+ T cells. The corresponding heatmaps (bottom) show cluster overlap indices in randomly sampled results (rows) versus the original result using all 9,274 CD8+ T cells (columns).

Supplementary Fig. S13



Supplementary Fig. S13. Reprogramming of lymphoid cell subsets towards an immune suppressive tumor microenvironment in early-stage LUAD. A, Changes in the abundance of specific lymphoid cellular lineages and states across the LUADs and spatial normal samples. Embedded pie charts show the contribution of each spatial sample to the indicated cell subtype/state. **B,** Cytotoxicity signature score in CD8+ *GNLY*_hi CTLs in P3 (left) and in P4 (right) across spatial samples (*, $P < 0.05$; ***, $P < 0.001$ of the Wilcoxon rank sum test). **C,** Treg signature score in T regulatory cells across spatial samples in all patients (left), P3 (middle) and in P5 (right) (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$ of the Wilcoxon rank sum test). **D,** Depletion of CD4+ CTLs in the tumor microenvironment of LUAD. Boxplot showing percentage of CD4+ CTLs *GZMA*-hi among total CD4+ cells from all patients across the spatial samples. Each circle represents a patient sample. *P* – value was calculated using Kruskal-Wallis test. **E,** Violin plots showing cytotoxic signature score in CD4+ CTL *GZMA*-hi cells from all patients (left) and P5 (right; ***, $P < 0.001$ of the Wilcoxon rank sum test). **F,** Frequency of CD4+ CTL *GZMA*-hi cells co-expressing *GZMA* and *GZMH* across the spatial samples. The fractions of *GZMA*+*GZMH*+ *CD4*+ CTLs are labeled on each plot. **G-I,** Expression profiling of different isotypes of plasma cells across the spatial samples. **G,** Heatmap showing isotype genes expression in plasma cells. **H,** Bar plots showing plasma cell isotype composition across spatial samples. **I,** Dot plots showing the fractional change of IGHA1/2 and IGHG3 across all patients and in each patient (left to right). **J,** Stacked bar plot showing changes in the relative fractions of lymphoid cell subsets between non-smokers and smokers. Treg, T regulatory cell; ILC, innate lymphoid cell; NK, natural killer cell; Other T cells, other CD4 and CD8 T cells states. **K,** Boxplot showing change in relative fraction of plasma cells between non-smoker and smokers in TCGA LUAD cohort. *P* – value was calculated using Wilcoxon rank sum test.

Supplementary Fig. S14

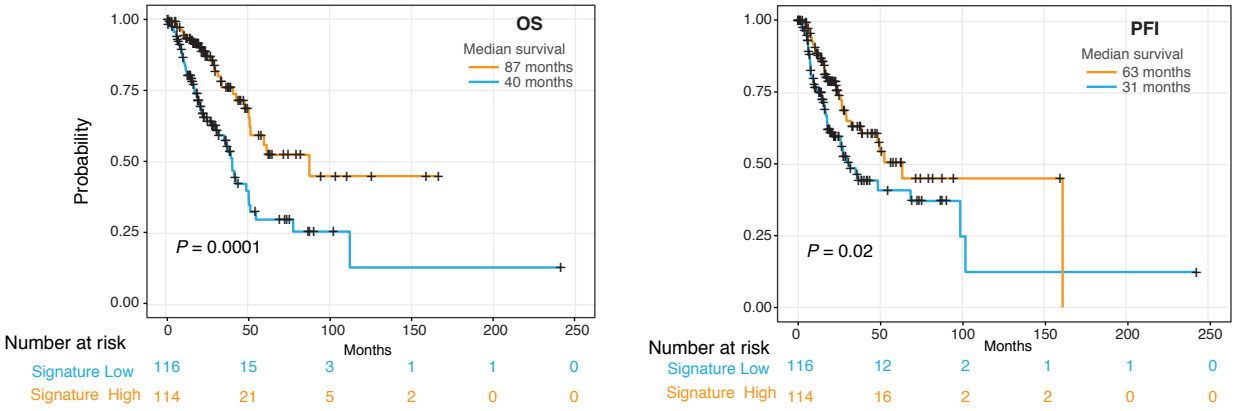


Supplementary Fig. S14. Reprogramming of B cells in the tumor microenvironment of early-stage LUAD. A, Heatmap showing unsupervised clustering of B cells sub-populations. **B,** UMAP showing re-clustering of B cells into sub-populations. **C,** Ridge plots (left) showing *RAC2* and *ACTG* expression levels among B cell sub-clusters, and scatter plots (right) showing the frequency of B cells co-expressing *RAC2*+ and *ACTG*+ across spatial samples. The fractions of *RAC2*+*ACTG*+ B cells are labeled on each plot. **D,** Boxplot showing the B cell C0 signature score in normal lung tissues (NL), premalignant atypical adenomatous hyperplasias (AAH) and LUADs in an independent validation cohort. (*, $P < 0.05$; ***, $P < 0.001$; N.S, $P > 0.05$ of the Wilcoxon rank sum test).

Supplementary Fig. S15

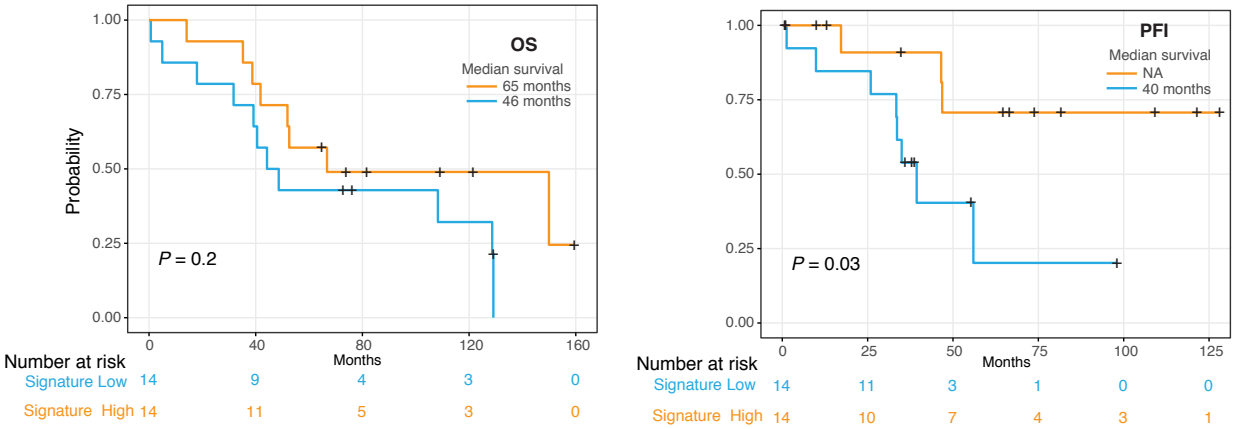
A

B cell signature, TCGA LUAD cohort



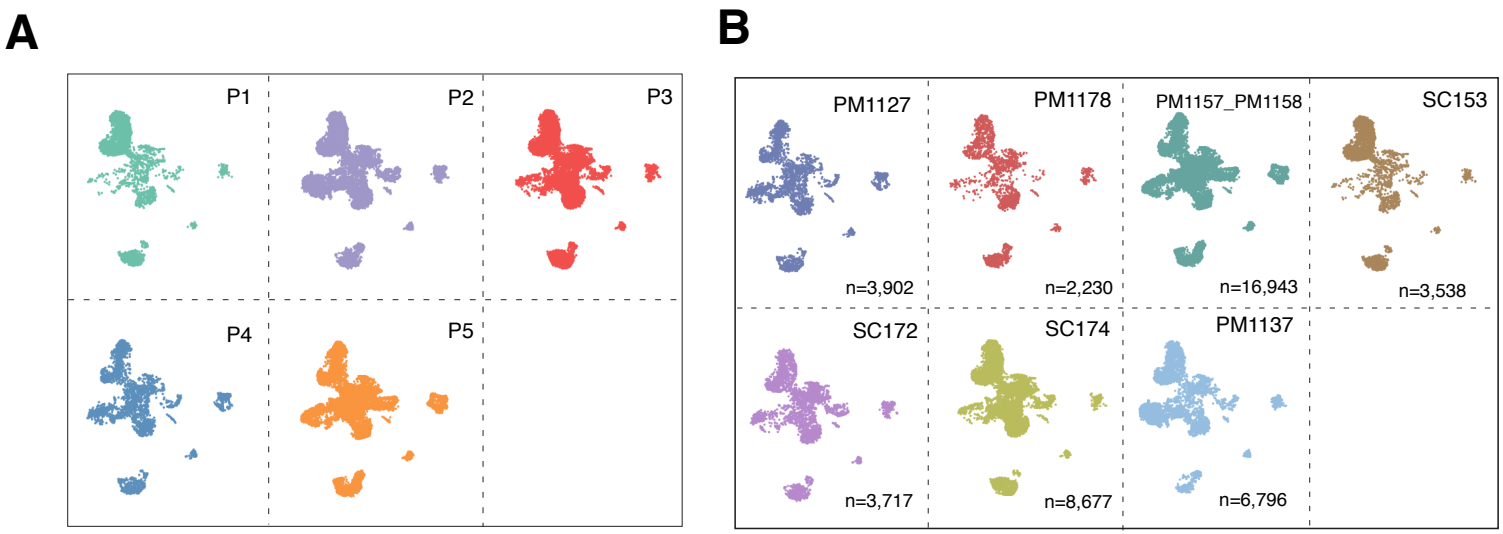
B

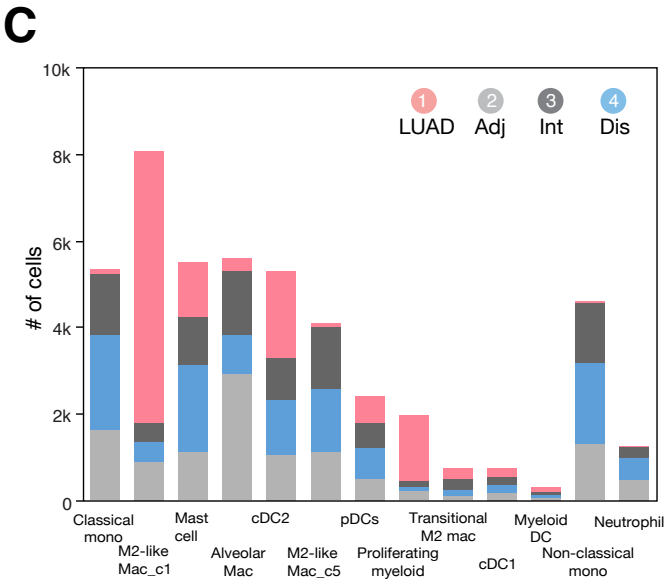
B cell signature, MDACC cohort

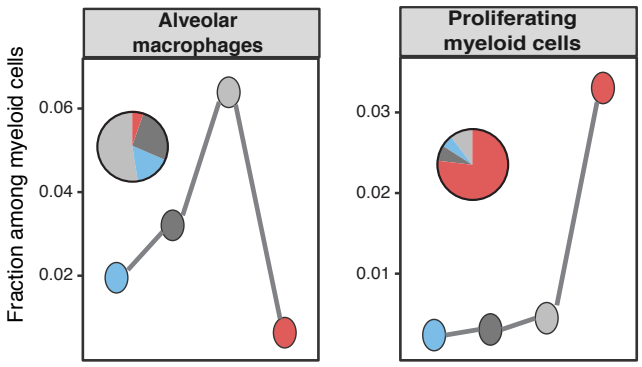


Supplementary Fig. S15. Association of a B cell signature with clinical outcome. Overall survival (OS, left) and progression-free interval (PFI, right) of treatment-naïve LUAD patients grouped by scores of the B cell signature in TCGA (**A**) and MDACC cohorts (**B**). Patients were stratified based on the first and fourth quartiles of B cell signature scores (TCGA B cell low n = 116 and high n = 114; MDACC B cell low n = 14 and high n = 14). Survival analysis was performed using Kaplan–Meier estimates and two-sided log-rank tests.

Supplementary Fig. S16

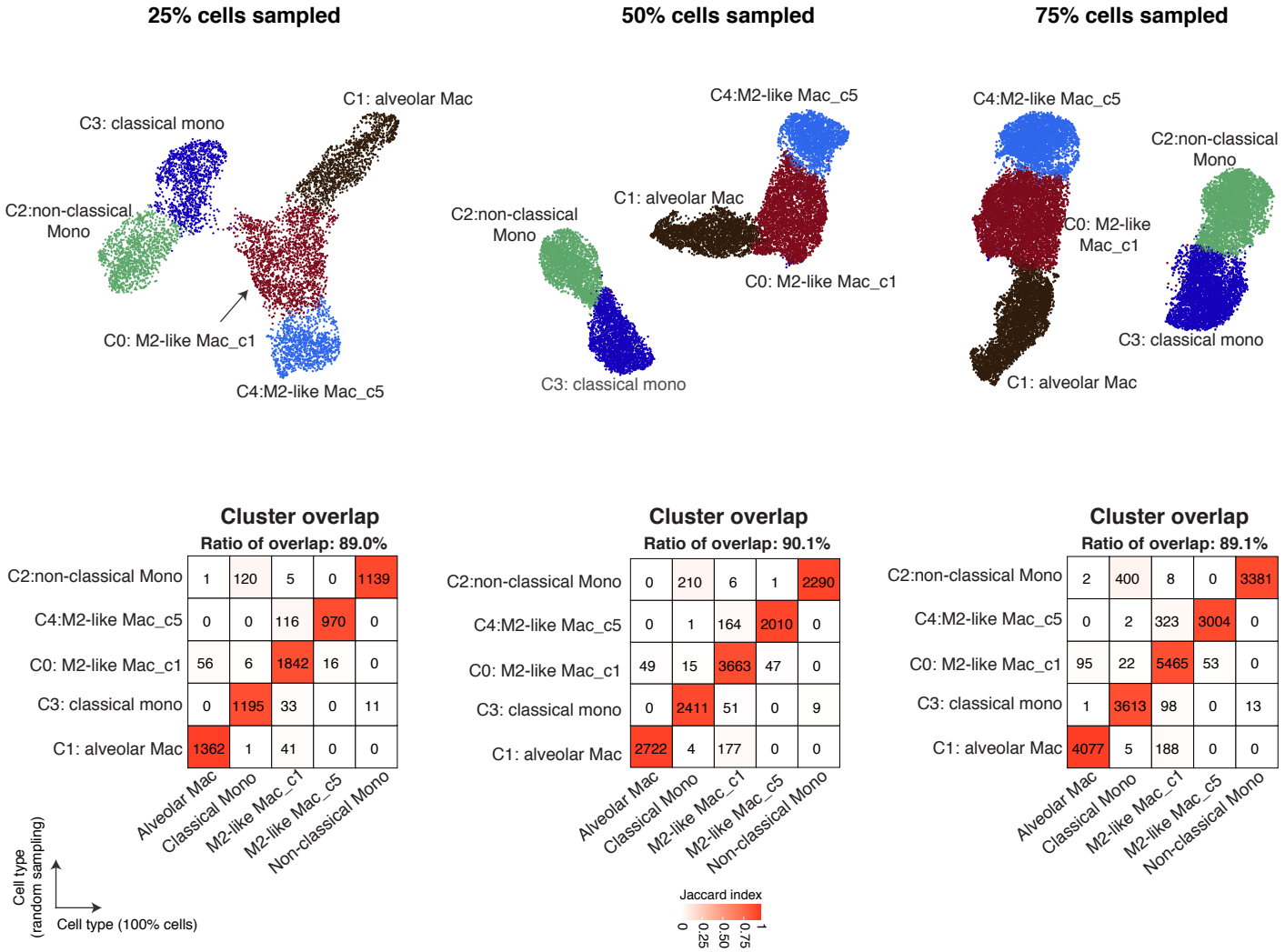






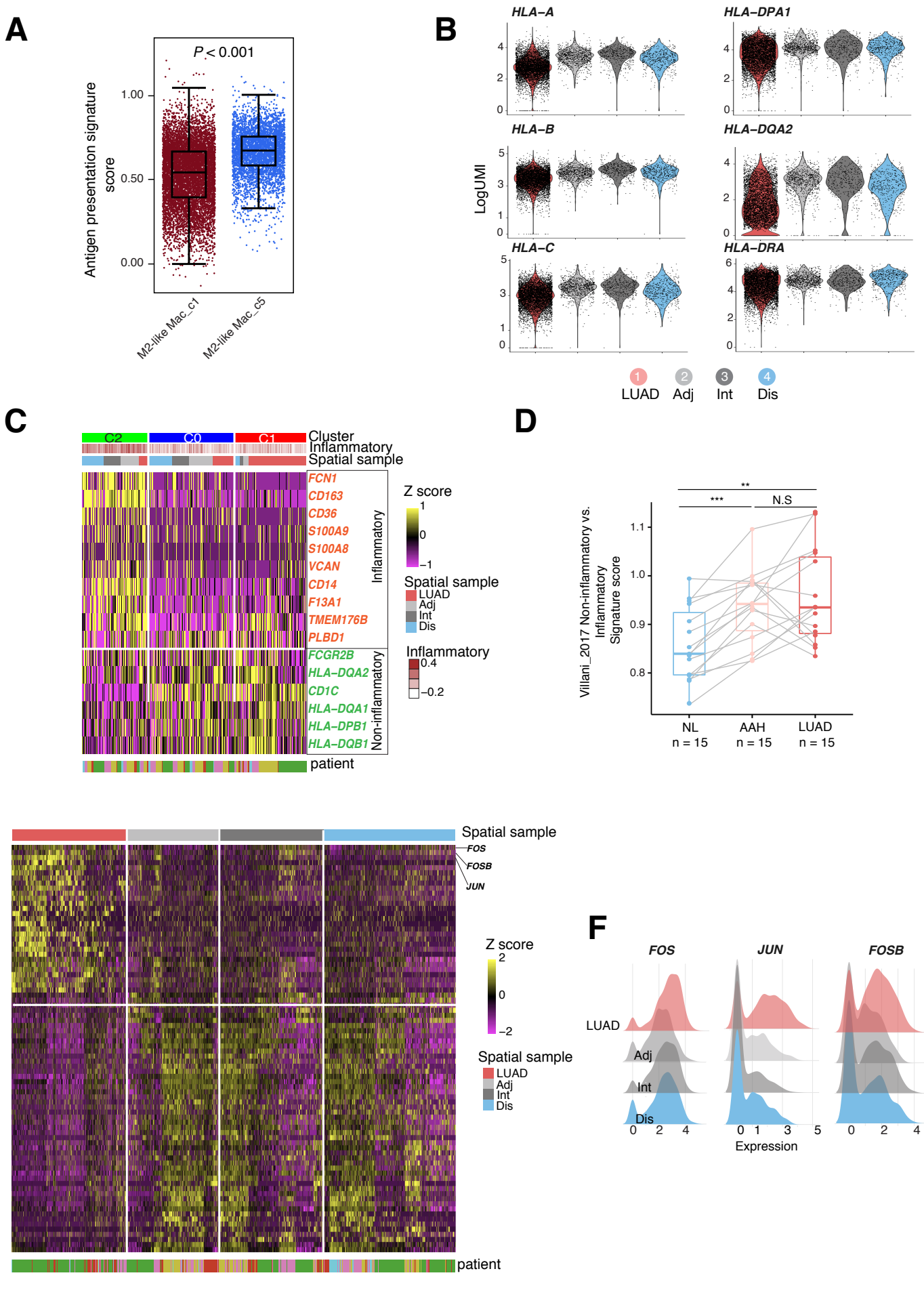
Supplementary Fig. S16. Quality control metrics of cells from myeloid cell subsets. A-B, UMAP view showing myeloid cells colored by patient ID (A) and library/sequencing batch (B). **C,** Stacked bar plots showing absolute numbers of each myeloid cell subset across the spatial samples. Dis, distant normal; Int, intermediate normal; Adj, adjacent normal; LUAD, tumor tissue; mono; monocytes, mac; macrophages, cDC; classical dendritic cells, pDC; plasmacytoid dendritic cell. **D,** Changes in the abundance of specific myeloid cellular lineages and states across the LUADs and spatial normal samples. Embedded pie charts show the contribution of each spatial sample to the indicated cell subtype/state.

Supplementary Fig. S17



Supplementary Fig. S17. Analysis of robustness of monocyte and macrophage clustering. UMAP plots (top) showing the Harmony-based clustering results when using 25% (top, left), 50% (top, middle), and 75% (top, right) of randomly sampled cells. The corresponding heatmaps (bottom) show cluster overlap indices in randomly sampled results (rows) versus when using all 27,664 monocytes and macrophages (columns).

Supplementary Fig. S18

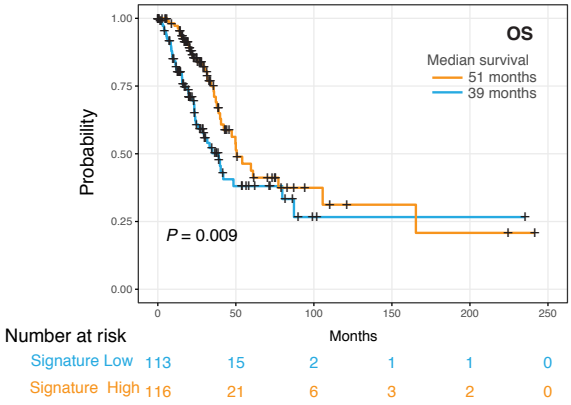


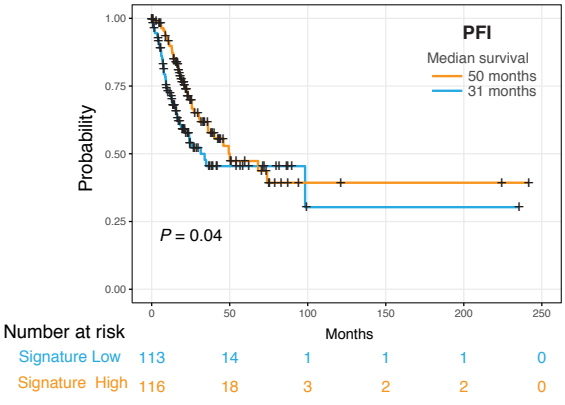
Supplementary Fig. S18. Reprogramming of myeloid cells in the tumor microenvironment of early-stage LUAD. A, Boxplot showing the antigen presentation signature score between M2-like Mac_c1 and M2-like Mac_c5. Mac; macrophages. *P* – value was calculated using Wilcoxon rank sum test. **B,** Violin plots showing the expression level changes among spatial samples for antigen presentation involving genes. **C,** Heatmap showing the expression of inflammatory and non-inflammatory signature genes in cDC2 cells. **D,** Boxplot showing the non-inflammatory versus inflammatory signature score in normal lung tissues (NL), premalignant atypical adenomatous hyperplasias (AAH) and LUADs in an independent validation cohort (**, $P < 0.01$; ***, $P < 0.001$; N.S, $P > 0.05$ of the Wilcoxon rank sum test). **E,** Heatmap showing DEGs between LUAD and normal samples in pDCs. **F,** Ridge plots showing the expression level changes of *FOS*, *JUN* and *FOSB* in pDCs and across spatial samples.

Supplementary Fig. S19

A

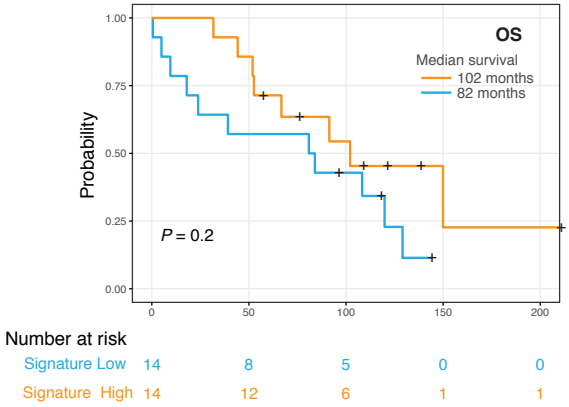
cDC2 signature, TCGA LUAD cohort

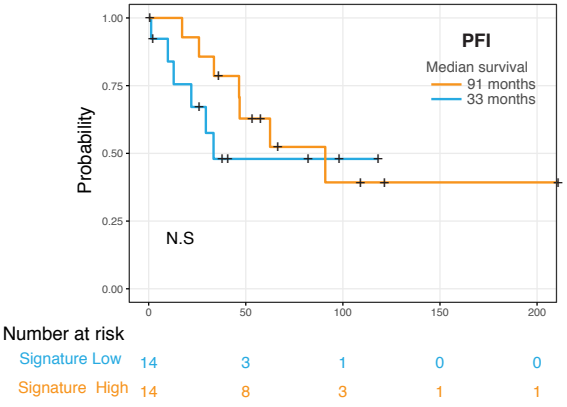




B

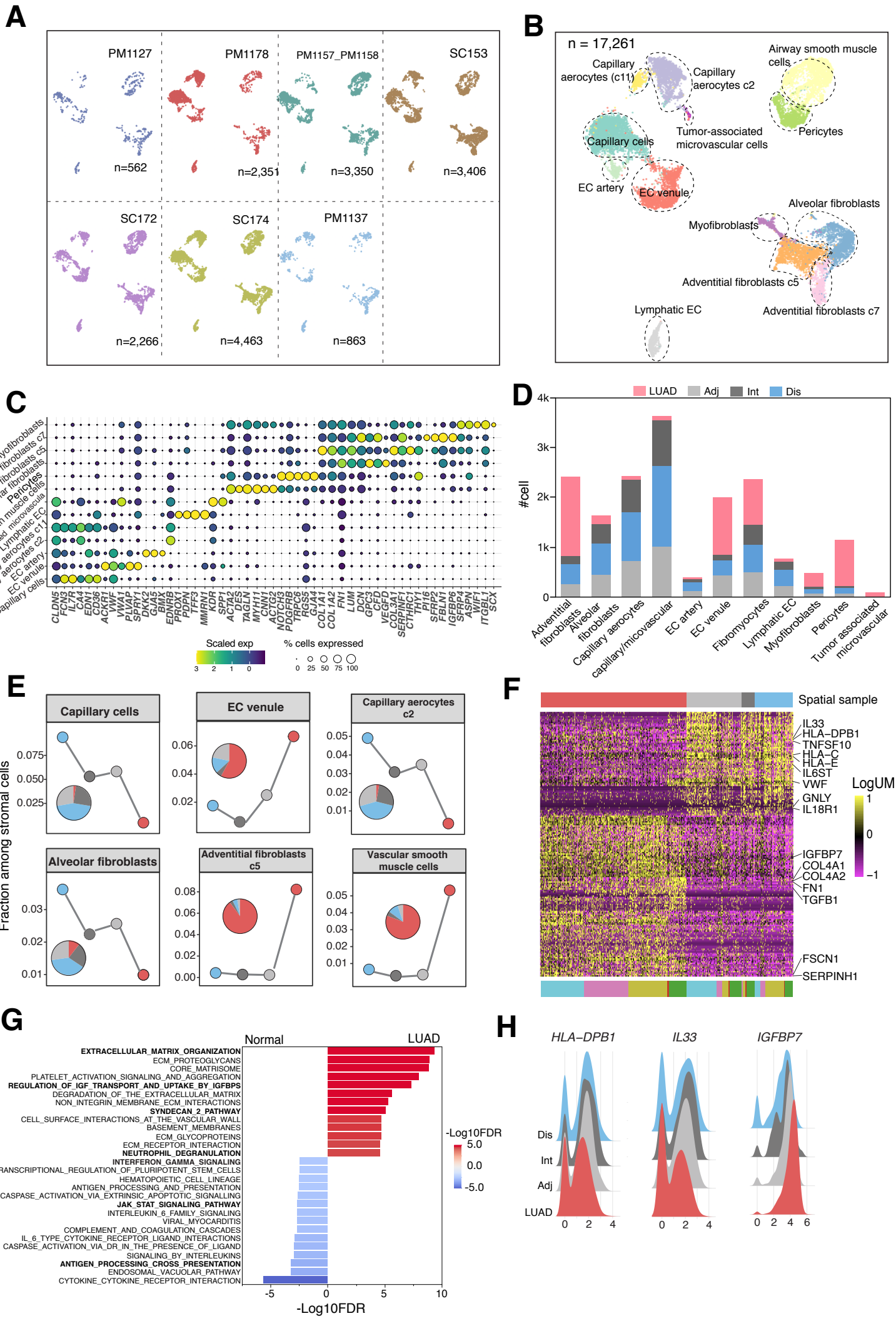
cDC2 signature, MDACC cohort





Supplementary Fig. S19. Survival analysis based on dendritic cell signatures. OS (left) and PFI (right) of treatment-naïve LUAD patients grouped by the cDC2 signature scores (ratio of non-inflammatory to inflammatory expression) in TCGA (**A**) and the MDACC (**B**) cohorts. Patients were grouped into first and fourth quartiles based on the cDC2 signature scores (TCGA cDC2 low n = 113 and high n = 116, MDACC cDC2 low n = 14 and high n = 14). Survival analysis was performed using Kaplan–Meier estimates and two-sided log-rank tests.

Supplementary Fig. S20



Supplementary Fig. S20. Composition and gene expression changes in stromal and endothelial sub-populations in early-stage LUAD. A, UMAP view of stromal and endothelial cells colored by sample batch. **B**, UMAP view of stromal and endothelial sub-populations. **C**, Bubble plot showing the percentage of stromal and endothelial cells expressing lineage markers (indicated by the size of the circle) as well as their scaled expression levels (indicated by the color of the circle). **D,** Bar plot showing the absolute number of cells from each stromal and endothelial subset. Dis, distant normal; Int, intermediate normal; Adj, adjacent normal; LUAD, tumor tissue; EC, endothelial cells. **E**, Changes in the abundance of stromal and endothelial cell lineages across the LUADs and spatial normal samples. Embedded pie charts show the contribution of each spatial sample to the indicated stromal and endothelial subtype. **F**, Heatmap showing DEGs between LUAD and normal samples in EC venule sub-populations. **G**, Bar plot showing significantly enriched pathways of up/down regulated DEGs in panel F. **H,** Ridge plots showing the expression level changes of *HLA-DPB1*, *IL33*, and *IGFBP7* in EC venule sub-populations and across spatial samples.