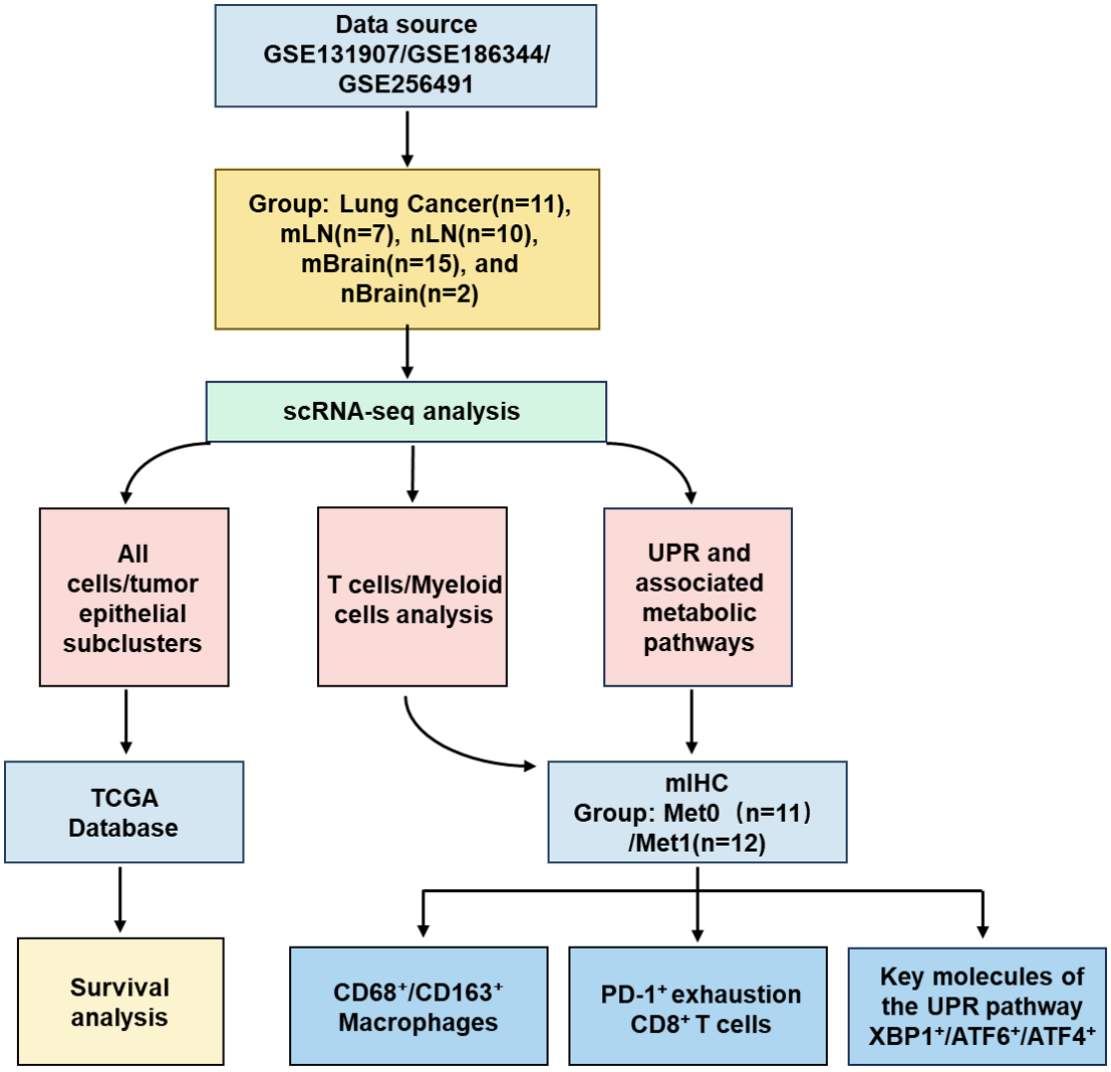
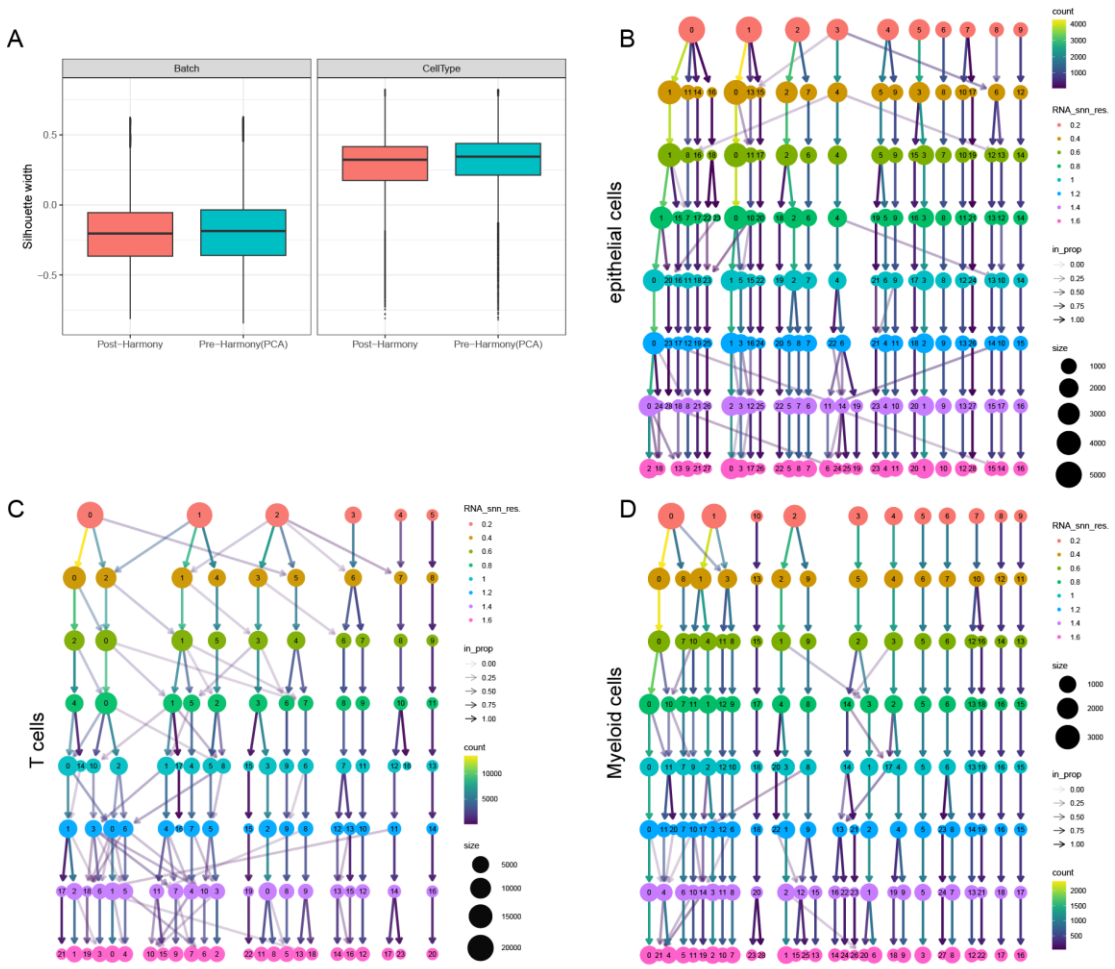
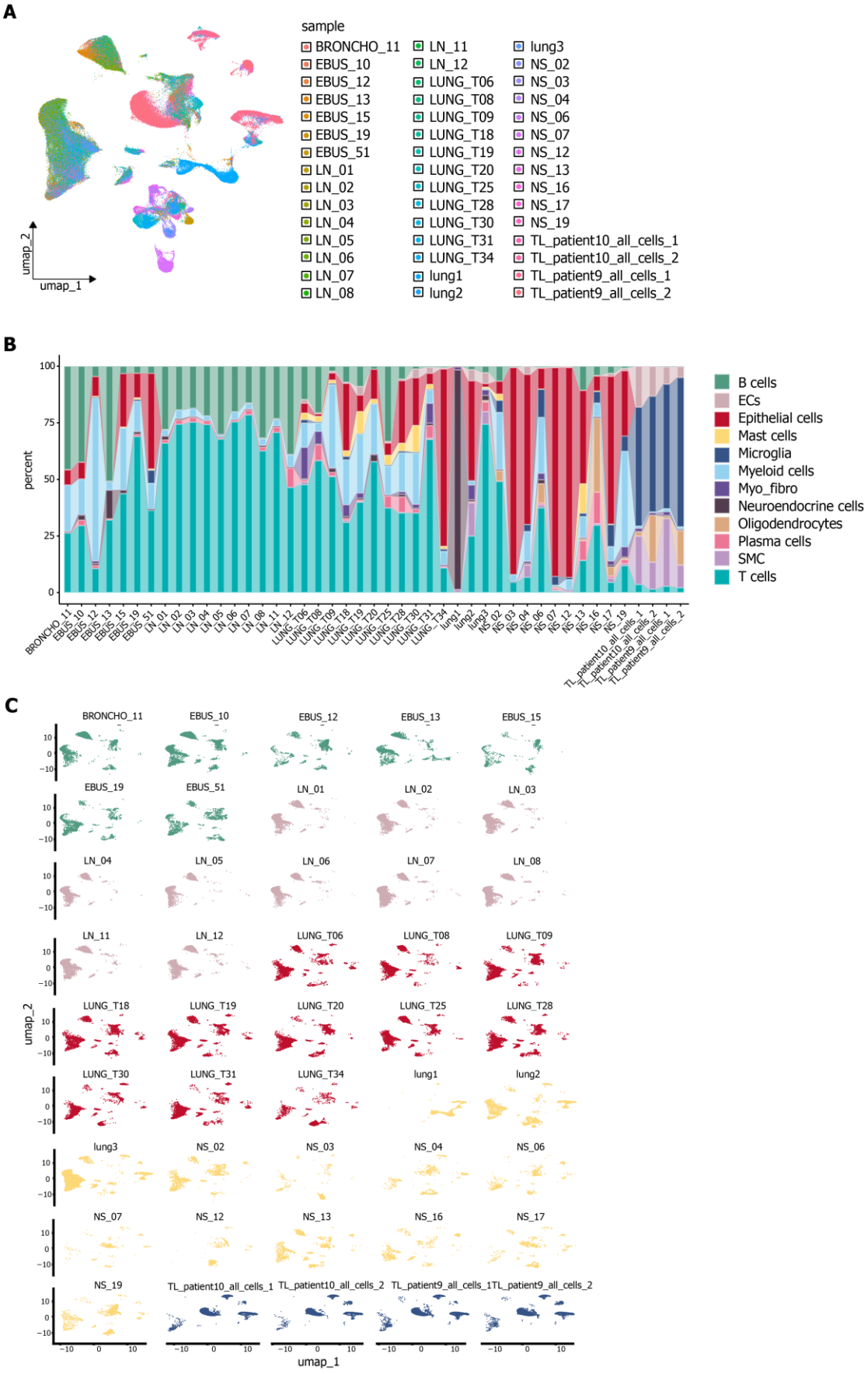
Supplementary Materials:



Supplementary Figure 1. The research workflow.

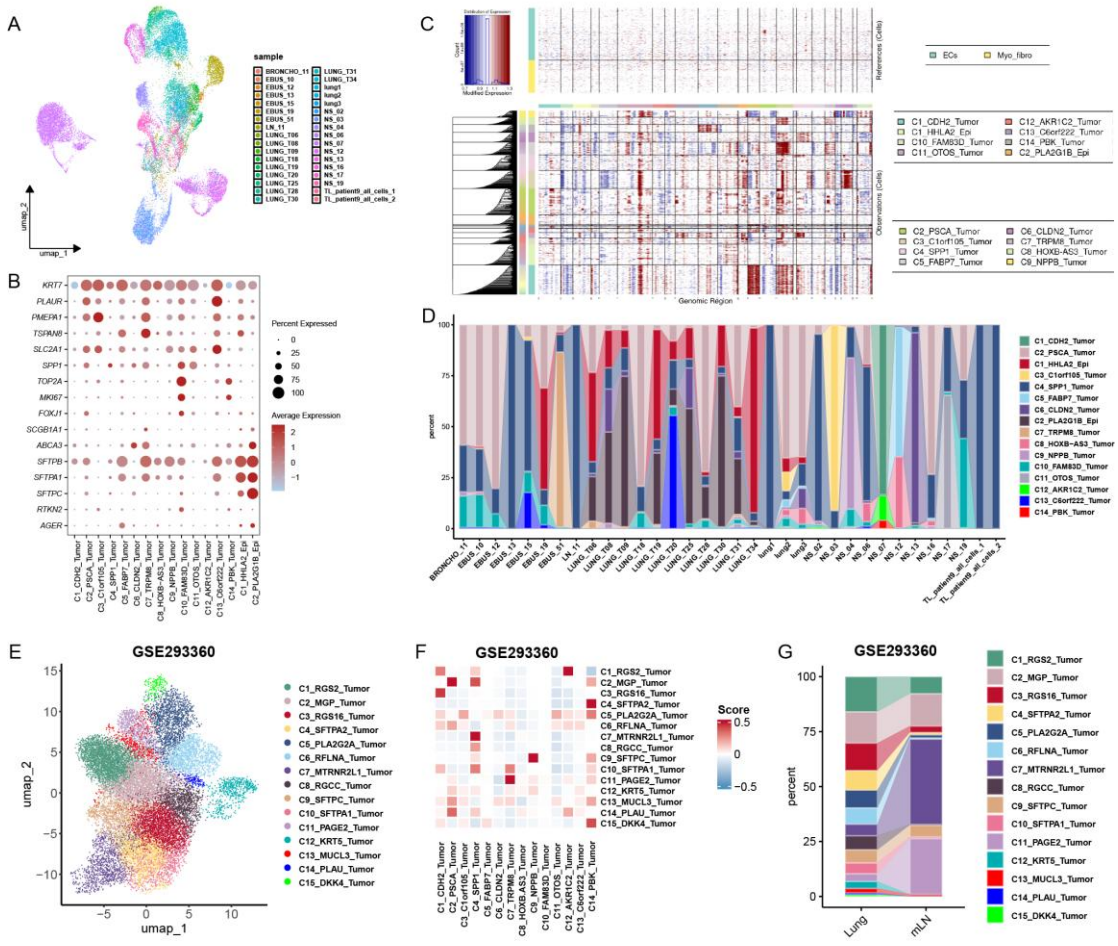


Supplementary Figure 2. Assessment of Harmony batch integration and hierarchical clustering transitions across multiple resolutions. (A) Evaluation of over-correction by Harmony integration based on silhouette width. Left: Batch dimension; Right: CellType dimension. **(B)** Hierarchical tree plot of single-cell clustering across increasing resolutions (epithelial cells). **(C)** Hierarchical tree plot of single-cell clustering across increasing resolutions (T cells). **(D)** Hierarchical tree plot of single-cell clustering across increasing resolutions (Myeloid cells).



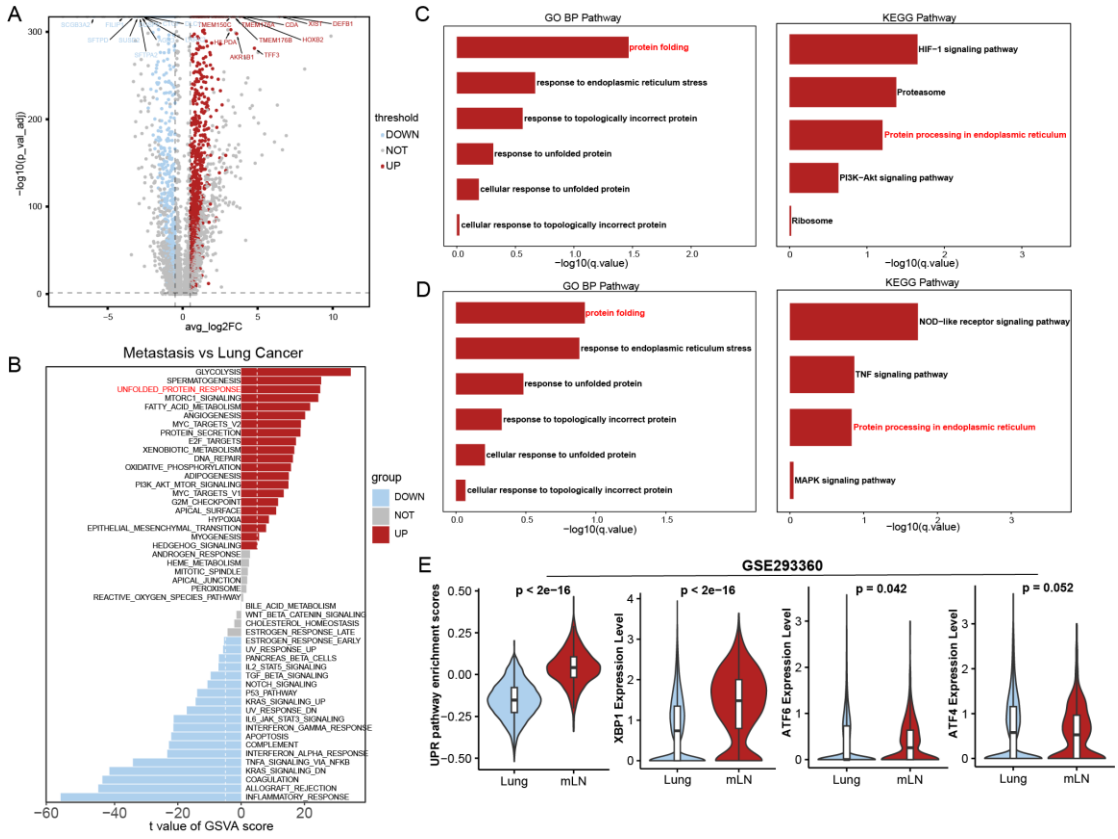
Supplementary Figure 3. Global single-cell transcriptomic landscape and cell type

composition analysis of primary and metastatic NSCLC tissues. (A) UMAP visualization of single-cell RNA sequencing data from all samples, annotated by sample origin, including bronchus (BRONCHO), endobronchial ultrasound biopsy (EBUS), lymph node (LN), lung tissue (LUNG), normal control tissue (NS), and brain tissue from tumor patients (TL). **(B)** Proportions of 12 core cell types in samples of different origins, including B cells, endothelial cells (ECs), epithelial cells, mast cells, microglia, myeloid cells, neurons, neuroendocrine cells, oligodendrocytes, plasma cells, smooth muscle cells (SMC), and T cells. **(C)** UMAP plots of individual samples, showing the distribution of cell subsets in each independent sample

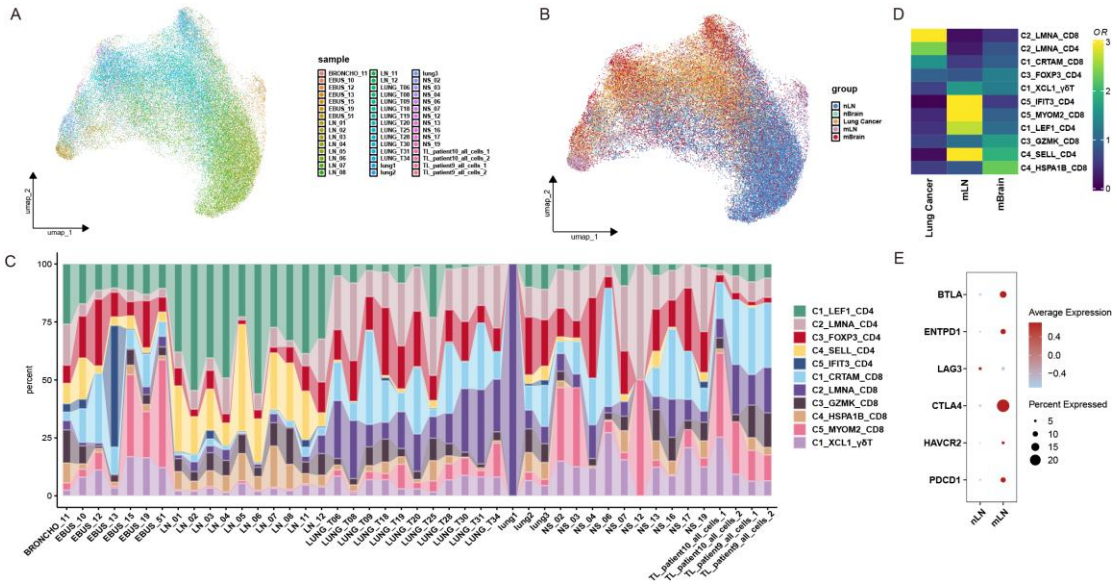


Supplementary Figure 4. Subclustering, sample distribution, and feature correlation analysis of tumor epithelial subpopulations in lymph node and brain metastases of NSCLC. (A) UMAP visualization of subclusters in NSCLC mLN and mBrain, annotated by sample origin, including bronchus, endobronchial ultrasound biopsy, lymph node, lung tissue, normal control tissue, and brain tissue from tumor patients. **(B)** Dot plot showing expression levels of marker genes across identified cell clusters. **(C)** Heatmap of cell CNV inferred based on inferCNV. **(D)** Stacked bar plot showing the proportional distribution of different cell subclusters across individual samples. **(E)** UMAP visualization of 15 tumor subclusters identified by subclustering analysis of 1,540 tumor

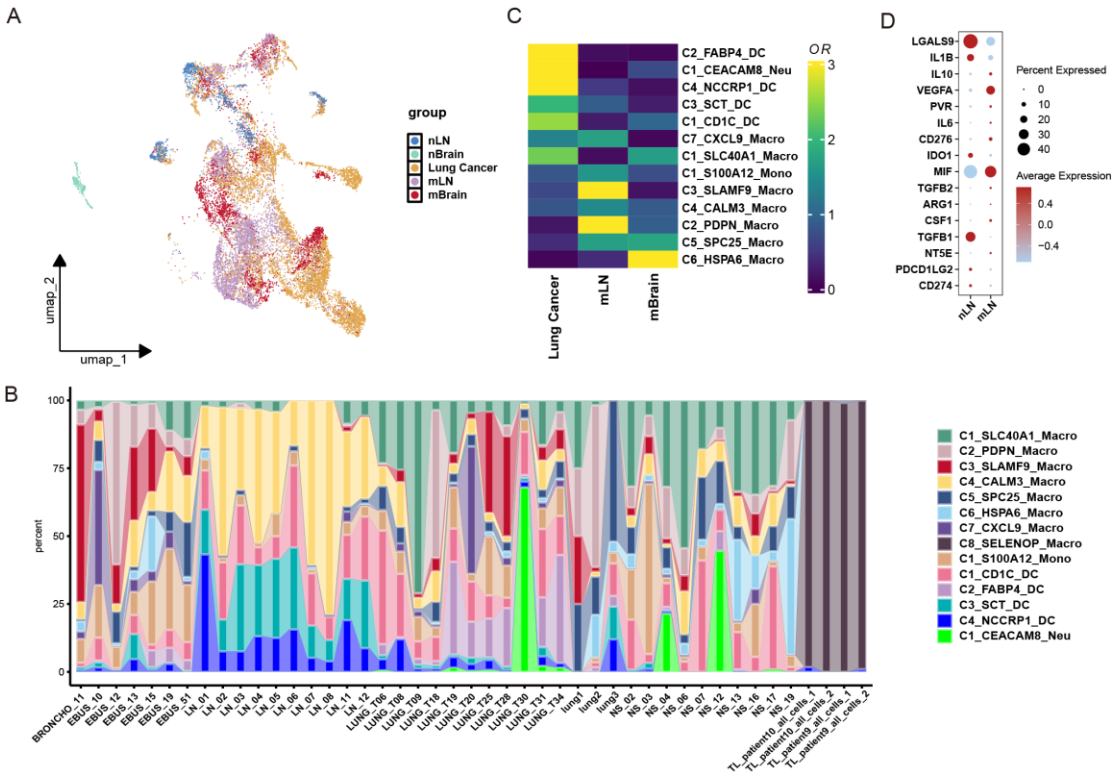
cells derived from NSCLC primary tumors (Lung Cancer), mLN. Each subcluster is annotated by its top marker genes. **(F)** The heatmap shows the consistency of the tumor epithelial subpopulations in the new dataset and the original dataset. **(G)** The stacked bar chart shows the proportions of each tumor subgroup in the Lung Cancer and mLN.



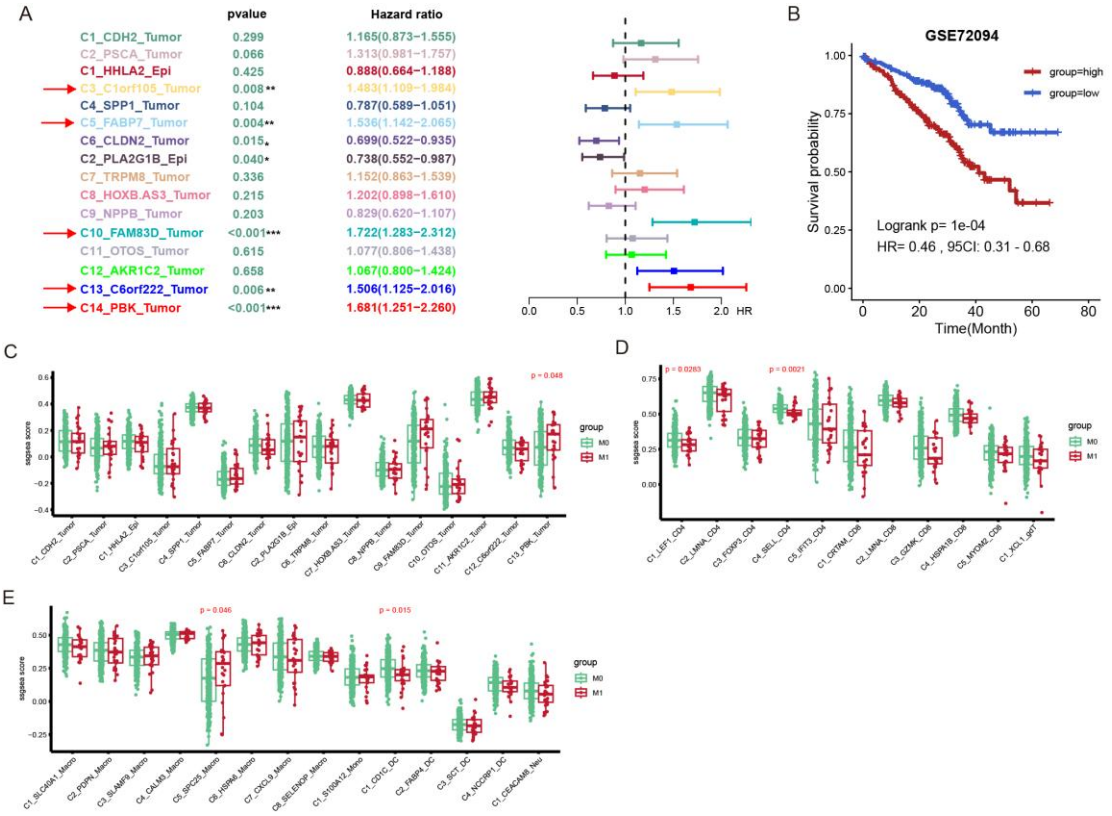
Supplementary Figure 5. Differential expression and pathway analysis between metastatic lesions and primary lung cancer. (A) Volcano plot of differentially expressed genes (DEGs) between metastatic lesions and primary lung cancer. Blue dots represent significantly downregulated genes, red dots represent significantly upregulated genes, and gray dots represent non-significant changes. The horizontal dashed line indicates the significance threshold ($p < 0.05$). **(B)** Bar plot of pathway enrichment analysis using Gene Set Variation Analysis (GSVA) comparing metastatic lesions and primary lung cancer. The x-axis represents the t-value of GSVA scores; red bars indicate upregulated pathways, blue bars downregulated pathways, and gray bars pathways with no significant difference. **(C)** Gene Ontology (GO) and KEGG pathway enrichment analysis of DEGs in mBrain compared to primary tumors. **(D)** GO and KEGG pathway enrichment analysis of DEGs in mLN compared to primary tumors. **(E)** Comparison of UPR pathway enrichment scores and XBP1, ATF6, and ATF4 expression between Lung and mLN samples.



Supplementary Figure 6. T cell subtyping, sample distribution, and tissue enrichment analysis in NSCLC brain and lymph node metastases. (A)UMAP visualization of single-cell RNA sequencing data from all samples annotated for T cells, with sample origin labeled including bronchus (BRONCHO), endobronchial ultrasound biopsy (EBUS), lymph node (LN), lung tissue (lung), normal control tissue (NS), and brain tissue from tumor patients (TL). **(B)** UMAP plot colored by tissue origin (nLN, nBrain, Lung Cancer, mBrain, mLN). **(C)** Proportions of 11 T cell types in samples of different origins. **(D)**Heatmap showing the enrichment of each T cell subset distribution in primary tumors, mBrain, and mLN tissues. Yellow indicates significant enrichment of the subset in the corresponding tissue, while blue indicates lower enrichment. **(E)** Dot plot showing the elevated expression of exhaustion markers, including PD-1 (PDCD1), TIM-3 (HAVCR2), and CTLA4, T cells of metastatic sites (mLN) compared to normal tissues.

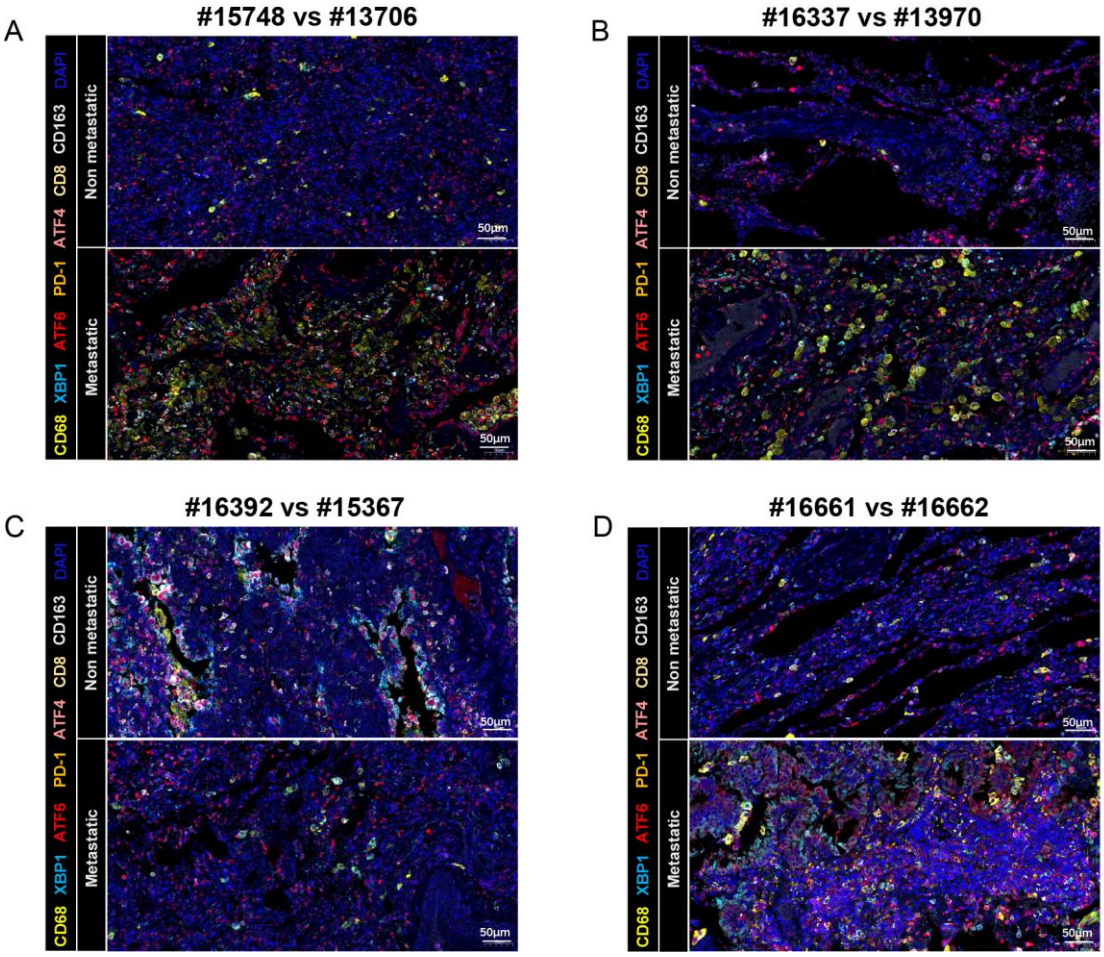


Supplementary Figure 7. Myeloid cell subtyping, sample distribution, and tissue enrichment analysis in NSCLC brain and lymph node metastases. (A) UMAP visualization of myeloid cells in NSCLC mBrain and mLN, grouped by tissue origin. The plot shows the overall distribution of myeloid cells across different tissue microenvironments. **(B)** Stacked bar plot showing the proportional distribution of myeloid cell subsets in each sample. **(C)** Heatmap of myeloid cell subset enrichment across different tissues. Yellow indicates significant enrichment of the subset in the corresponding tissue, while blue indicates lower enrichment. **(D)** Dot plot showing the expression levels of key immunosuppressive in myeloid cells of metastatic sites (mLN) compared to normal tissues.



76

77 **Supplementary Figure 8. Prognostic value of metastasis-associated tumor subclusters in**
78 **TCGA LUAD cohort and box plots of ssGSEA signature scores. (A)** Univariate Cox regression
79 analysis of prognostic factors. Forest plot showing the hazard ratios (HRs) and p values for each
80 tumor (C1 – C14) and epithelial(C1,C2) subcluster in relation to overall survival of LUAD
81 patients from The Cancer Genome Atlas (TCGA). **(B)** The relationship between high/low
82 C14_PBK_Tumor expression and overall survival. **(C)** Box plot for each tumor and epithelial
83 subcuster ssGSEA signatures score. **(D)** Box plot for $CD4^+$ /$CD8^+$ /γδT cells related ssGSEA
84 signatures score. **(E)** Box plot for macrophages, monocytes, DC and neutrophils related ssGSEA
85 signatures score.



Supplementary Figure 9. mIHC images showing the expression and spatial distribution of CD68$^+$/CD163$^+$, PD 1$^+$/ CD8$^+$, XBP1$^+$/ATF6$^+$/ATF4$^+$ makers in Mets0 and Mets1 NSCLC tissues. (A-D) The sample number is indicated at the top of the picture, and the sample information can be found in Supplementary Table 5. Scale bars: 50 μm.