

Supplementary Information

Additional file 2: Supplementary Materials for: Tumor Microenvironment Remodeling after Neoadjuvant Immunotherapy in Non-small Cell Lung Cancer Revealed by Single-Cell RNA Sequencing. This file includes Supplementary Figures: Fig. S1-S8

Fig. S1: Sample collection, single-cell clustering and marker gene visualization for major lineages.

Fig. S2: Immunohistochemistry (IHC) staining of immune cells.

Fig. S3: Cluster marker genes, tumor cell identification, cellular fractions and hematoxylin eosin (H&E) staining for epithelial cells.

Fig. S4: Upregulated genes and signatures for malignant cells among TN, MPR and NMPR patients and serum estradiol dynamics.

Fig. S5: Marker genes, differentially expressed genes and RNA velocity for T/NK subsets

Fig. S6: Marker genes, immunofluorescence, and NicheNet analysis for B-cell subsets.

Fig. S7: Marker genes and survival analysis for myeloid subsets.

Fig. S8: Monocle, SCENIC and cellular fraction analysis for neutrophil subsets.

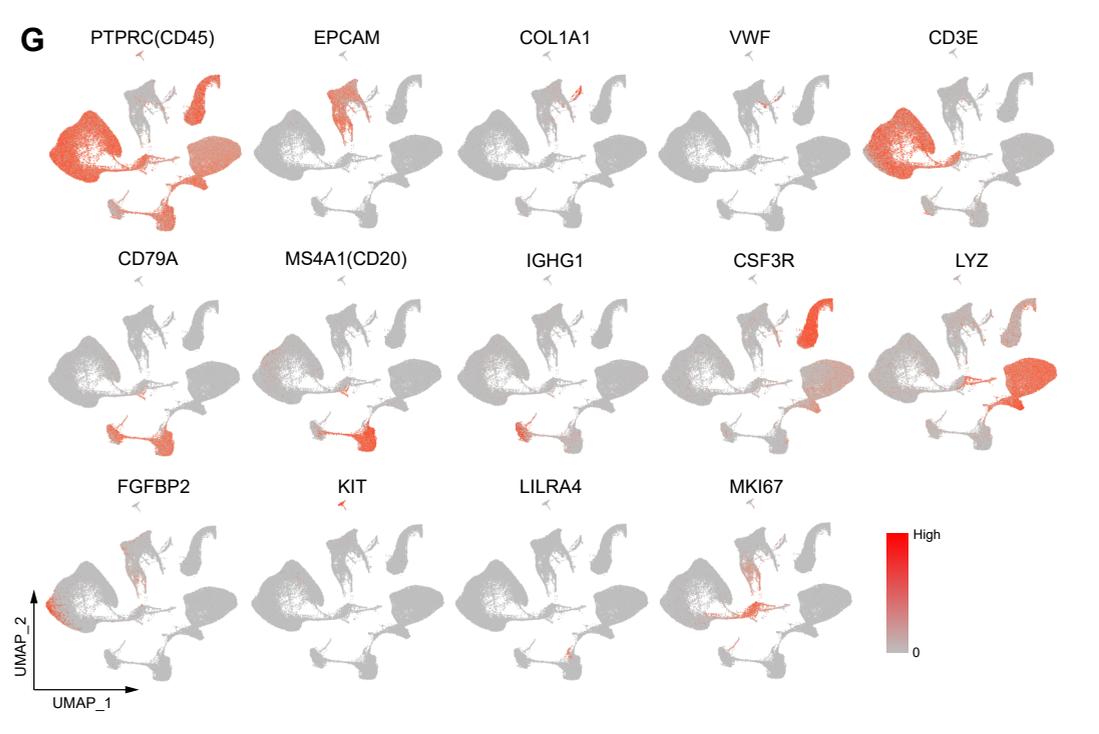
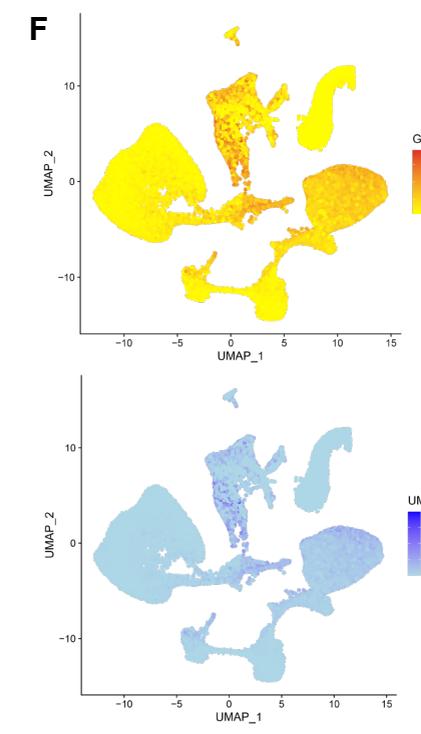
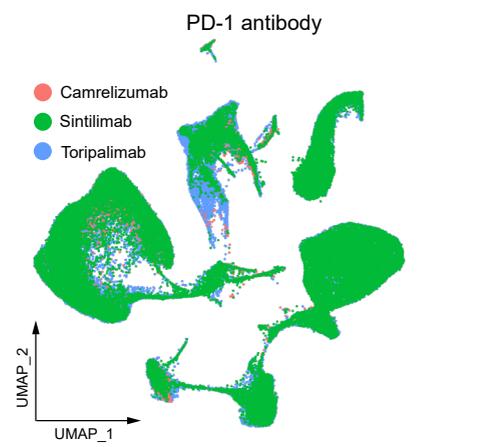
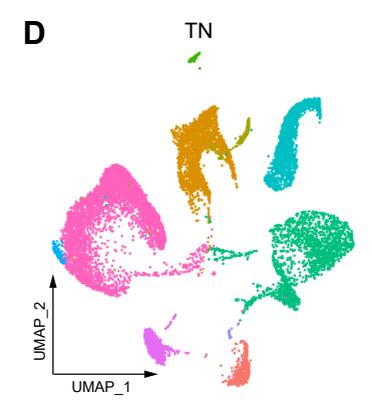
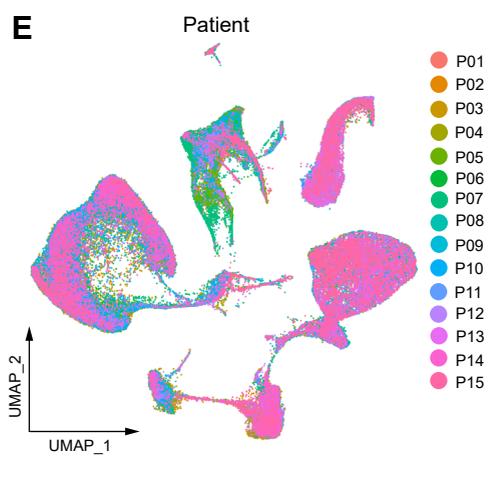
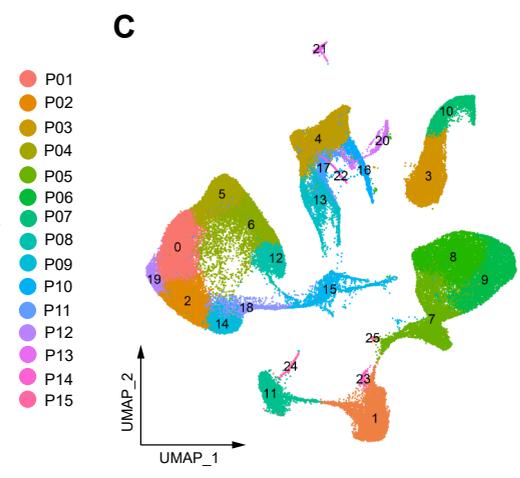
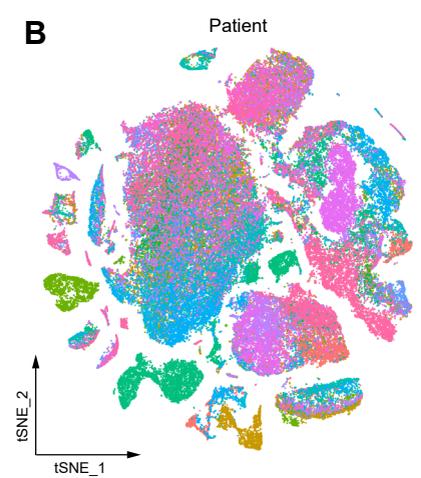
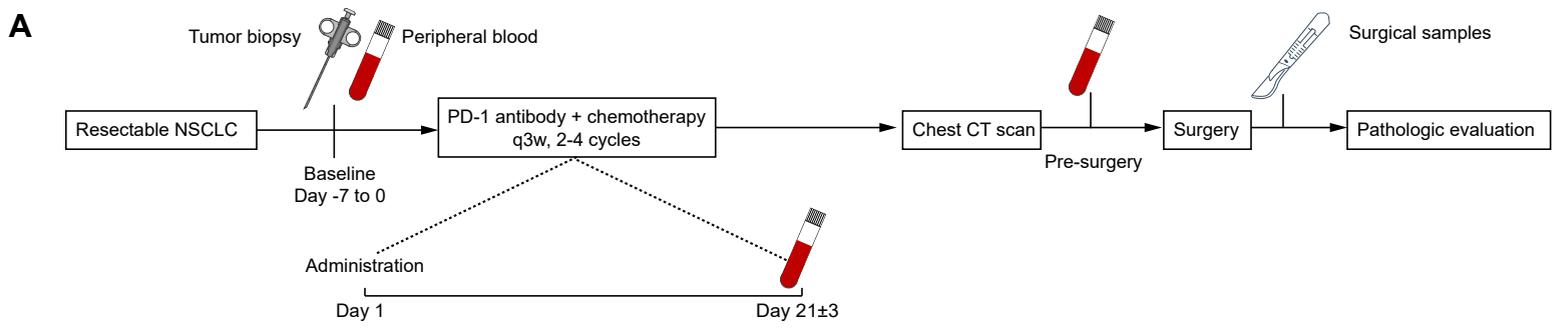


Fig. S1: Sample collection, single-cell clustering and marker gene visualization for major lineages.

A, Schematic representation of trial treatment.

B, tSNE plot showing batch effects from patients.

C, UMAP plot of all cells categorized into 26 clusters.

D, UMAP plots of the cells from treatment naïve (TN), major pathologic response (MPR) and non-major pathologic response (NMPR) patients.

E, UMAP plots of all cells colored by patient origin and PD-1 antibody.

F, UMAP plots showing the number of detected genes and UMIs for all cells.

G, UMAP plots showing the expression of selected marker genes for the defined cell types.

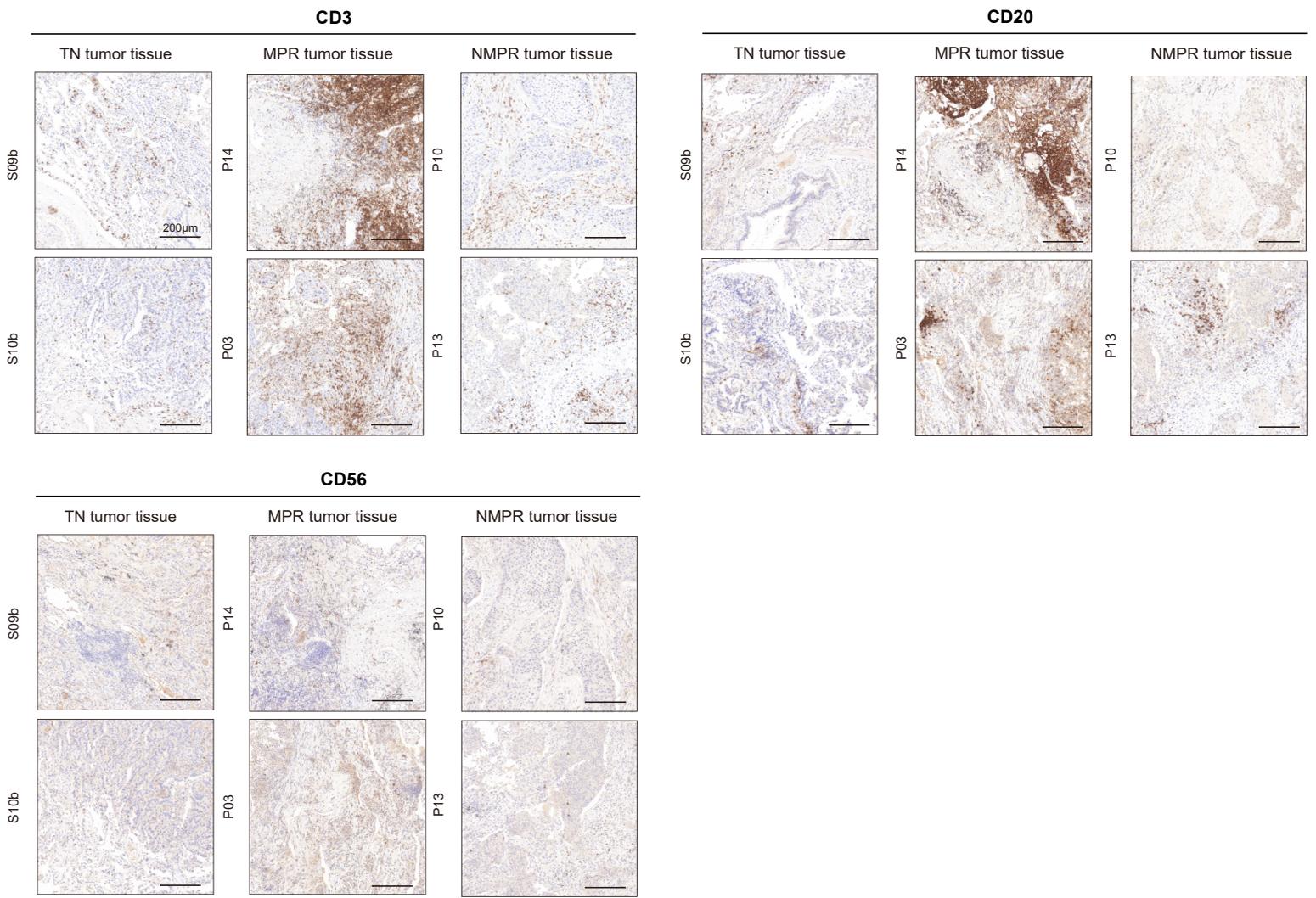


Fig. S2: Immunohistochemistry (IHC) staining of immune cells. Representative IHC staining of T (CD3), NK (CD56) and B (CD20) cells in another two TN, MPR and NMPR patients, respectively.

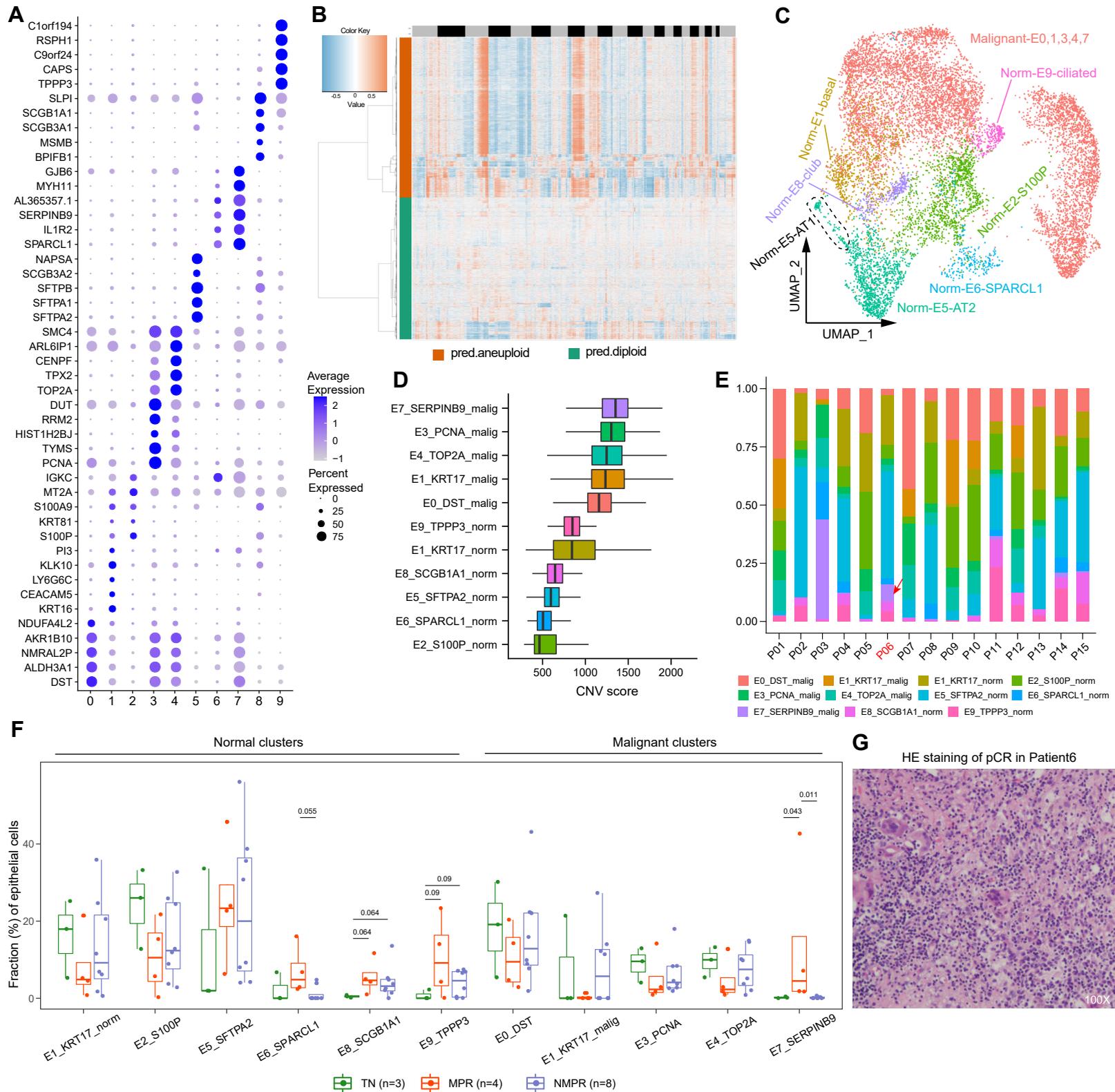


Fig. S3: Cluster marker genes, tumor cell identification, cellular fractions and hematoxylin eosin (H&E) staining of epithelial cells. A, Dot plots of top 5 marker genes for each epithelial cluster. B, Heatmap showing large-scale CNVs for individual epithelia. Each row represented a cell and the columns represented chromosomal regions. The cells in orange were predicted to be malignant cells and green were normal cells. C, UMAP plots of epithelia colored by malignant (in red) and normal clusters. The cells in the black dash line were alveolar type I (AT1) cells. D, Box plots of CNV scores for each epithelial cluster. Center line indicates the median, lower and upper hinges represent the 25th and 75th percentiles, respectively, and whiskers denote $1.5 \times$ interquartile range. Two-sided unpaired Wilcoxon test was used. E, The proportions of the cells in different clusters for each patient. The red arrow marked the malignant E7_SERPINB9 cells in Patient 6 whose tumor was assessed to be no viable tumor cells after therapy (pCR). F, Boxplot showing cellular fractions of each epithelial cluster in TN (n=3), MPR (n=4) and NMPR (n=8) patients. All differences with adjusted $P < 0.05$ are indicated; One-sided unpaired Wilcoxon test was used, and the P values were adjusted by the false discovery rate (FDR) method. G, H&E staining image of P06 whose tumor was assessed to be pCR.

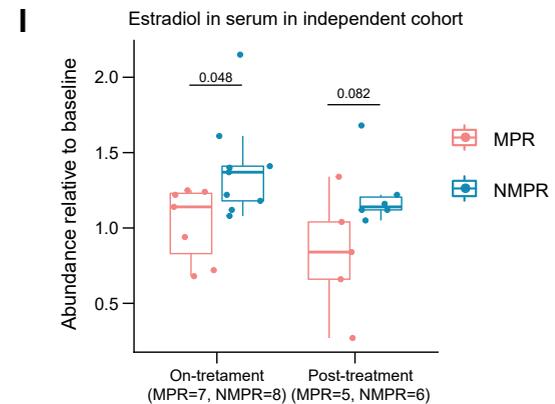
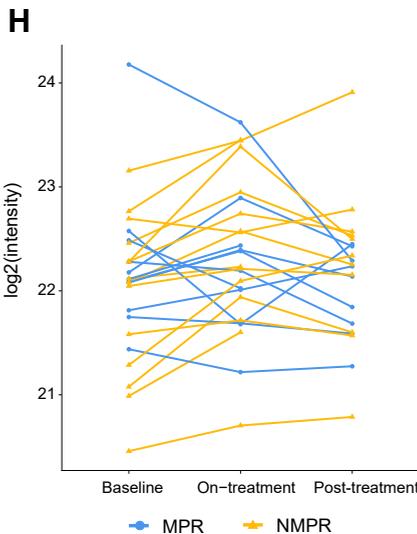
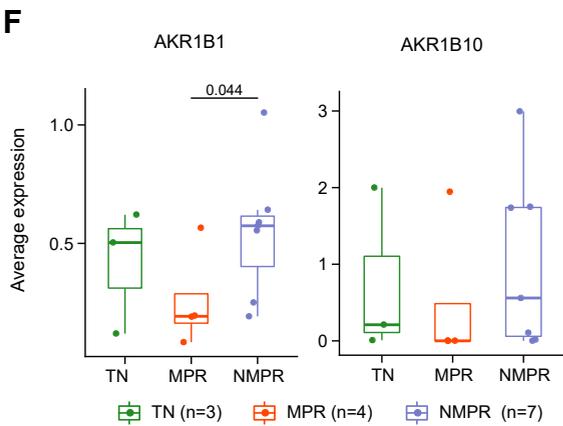
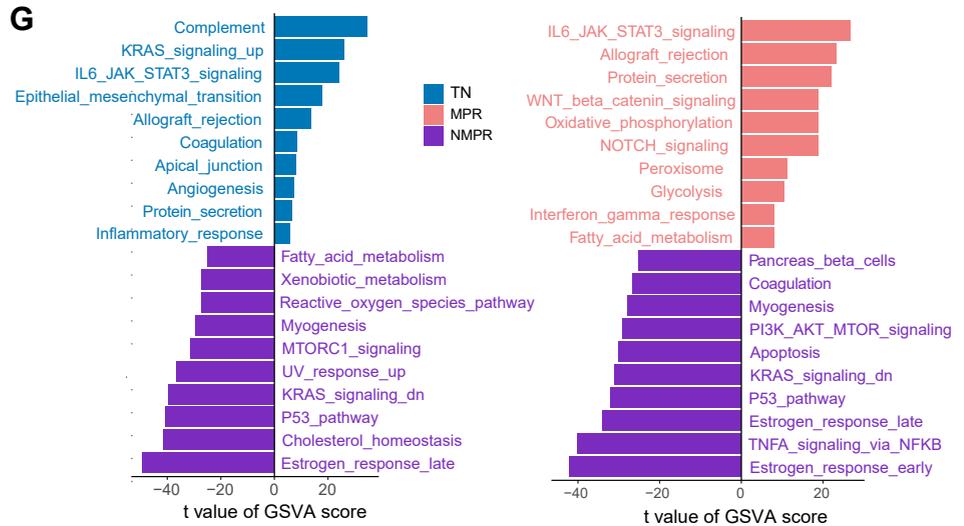
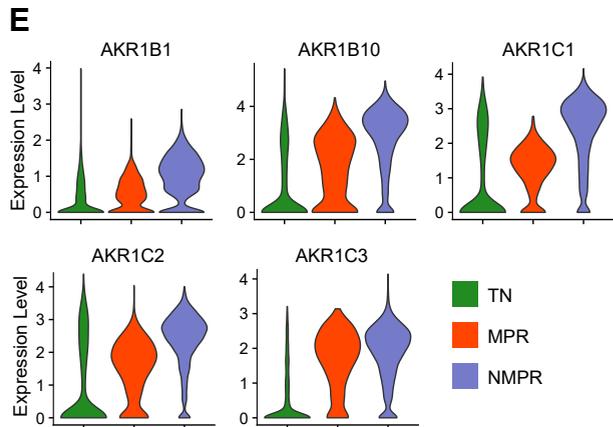
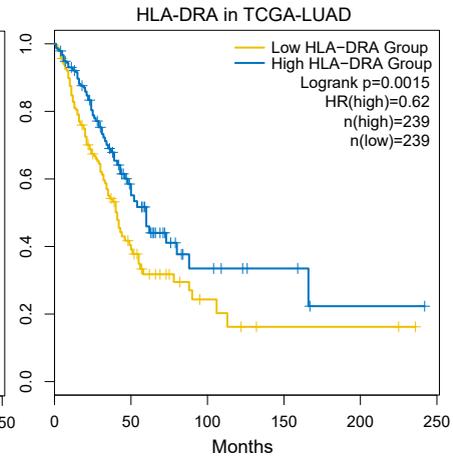
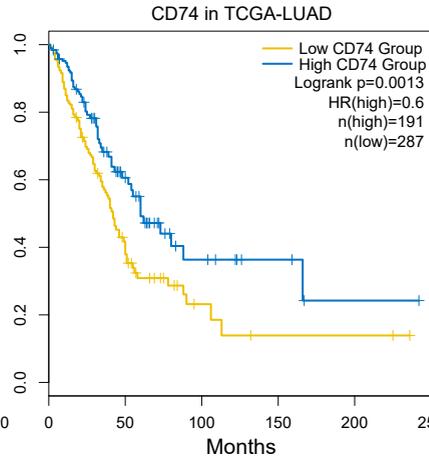
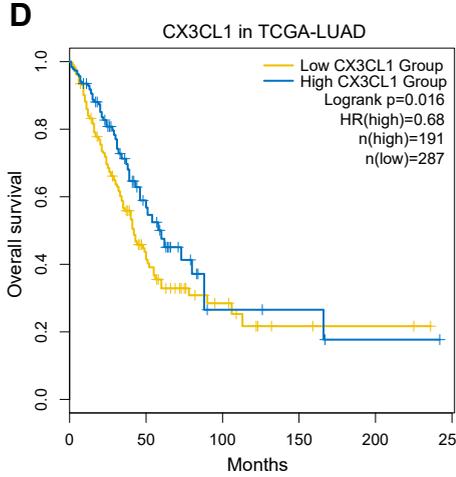
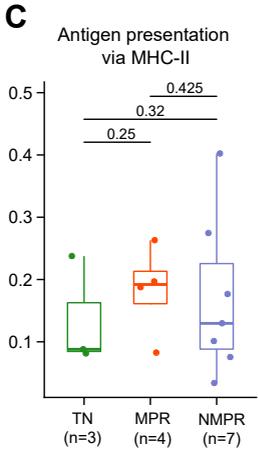
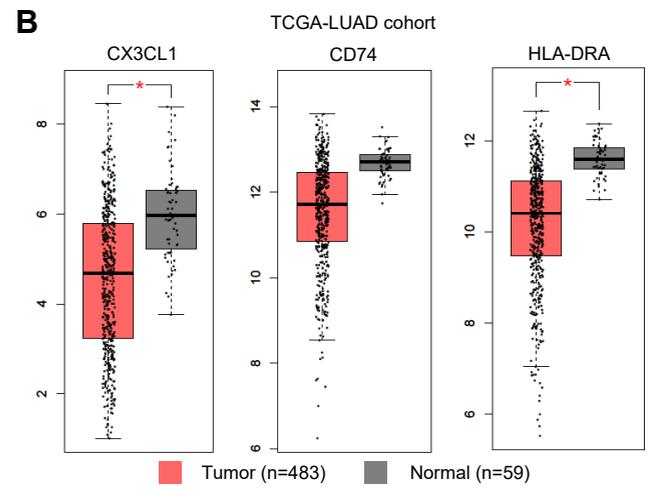
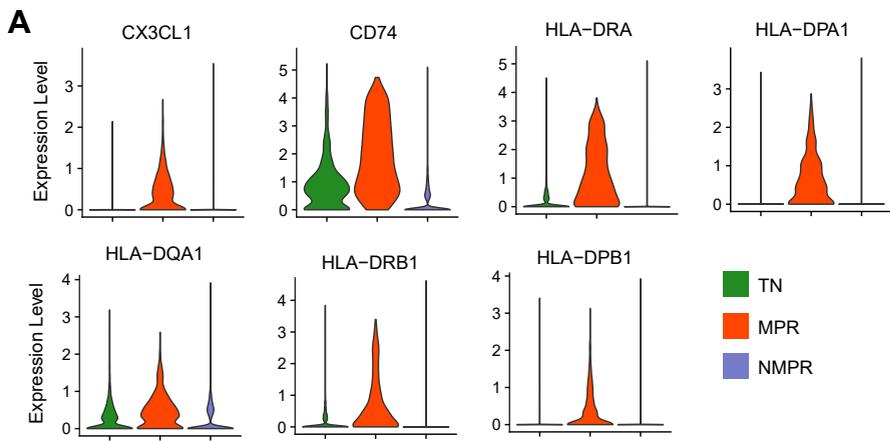


Fig. S4: Upregulated genes and signatures for malignant cells among TN, MPR and NMPR patients and serum estradiol dynamics.

- A, Violin plots of upregulated genes in malignant cells in MPR patients.
- B, Boxplots of the expression of CX3CL1, CD74 and HLA-DRA in TCGA-LUAD and normal lung tissues. Center line indicates the median, lower and upper hinges represent the 25th and 75th percentiles, respectively, and whiskers denote $1.5 \times$ interquartile range. Two-sided unpaired Wilcoxon test was used. * $P < 0.05$, ANOVA test.
- C, Boxplot of the average signature scores of antigen presentation via MHC-II for cancer cells in TN (n=3), MPR (n=4) and NMPR (n=7) patients. One-sided t-test was used.
- D, Kaplan–Meier survival curve of the expression of CX3CL1, CD74 and HLA-DRA in TCGA lung adenocarcinoma (LUAD) cohort. Two-sided log-rank test was used.
- E, Violin plots of upregulated genes in malignant cells in NMPR patients.
- F, Boxplots of the average expression of ARK1B1 and ARK1B10 in malignant cells in TN (n=3), MPR (n=4) and NMPR (n=7) patients. One-sided t-test was used.
- G, Top 10 differences in pathway activities scored per cell by GSVA comparing malignant cells between TN (n=1077) and NMPR (n=5368) patients (left panel) and between MPR (n=997) and NMPR patients (right panel).
- H, Line plot of estradiol abundance (\log_2 intensity) in baseline, on-treatment and post-treatment for each patient.
- I, Boxplots of the β -estradiol abundance relative to baseline in 15 patients independent against scRNA-seq cohort (7 patients were assessed as MPR and 8 as NMPR after surgery) at on-treatment and post-treatment timepoint.

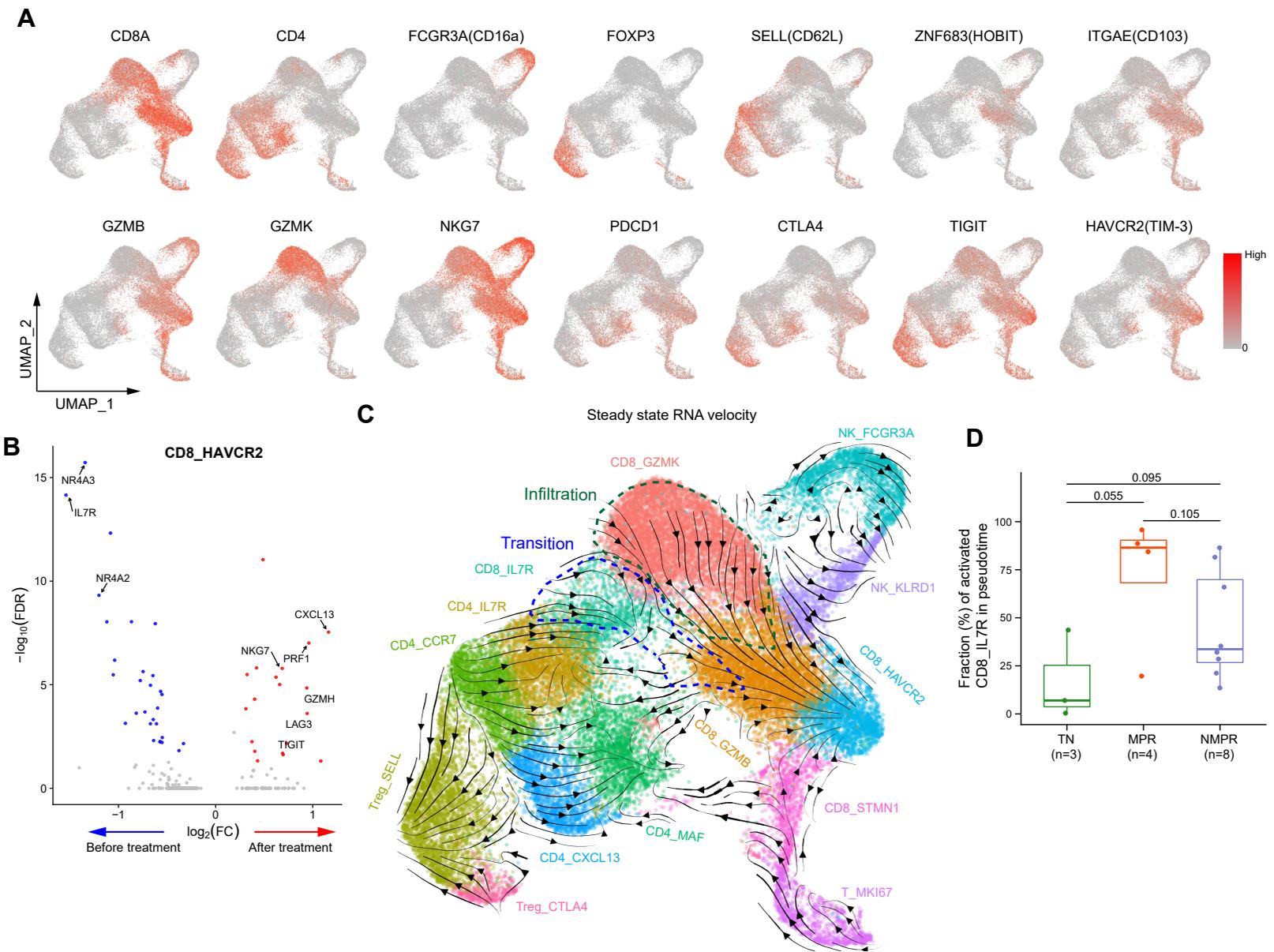


Fig. S5: Marker genes, differentially expressed genes and RNA velocity for T/NK subsets

- A, UMAP plots showing the expression of selected canonical marker genes for T/NK cells.
- B, Volcano plot of differentially expressed genes of CD8_HAVCR2 cells before and after therapy.
- C, Steady-state RNA velocity of T/NK cell clusters. The blue dash line marked the memory T cells transition into cytotoxic cells, and the blackish-green dash line marked the differentiation of effector memory cells into residual memory cells and exhausted cells.
- D, Box plots of fractions of activated memory T (CD8_IL7R) cells in pseudo-time in TN (n=3), MPR (n=4) and NMPR (n=8) patients. Center line indicates the median, lower and upper hinges represent the 25th and 75th percentiles, respectively, and whiskers denote 1.5× interquartile range. One-sided unpaired Wilcoxon test was used.

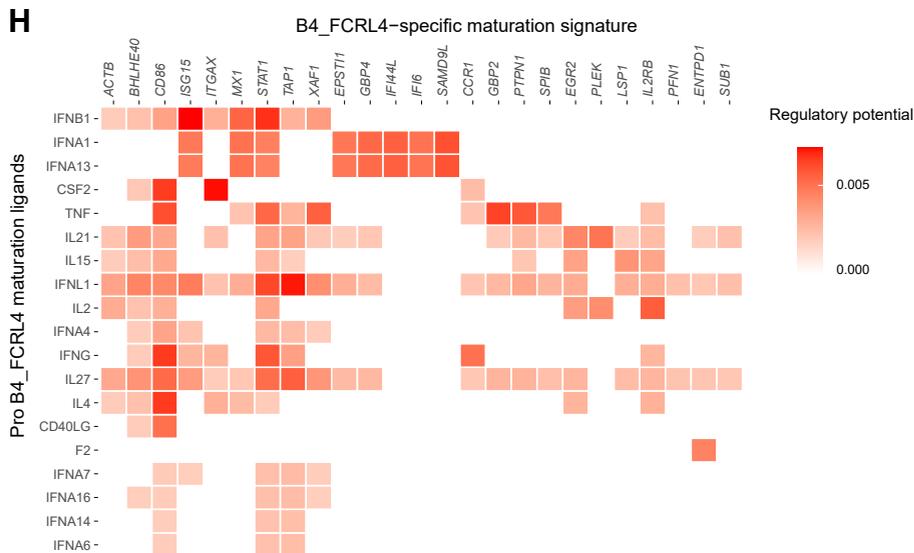
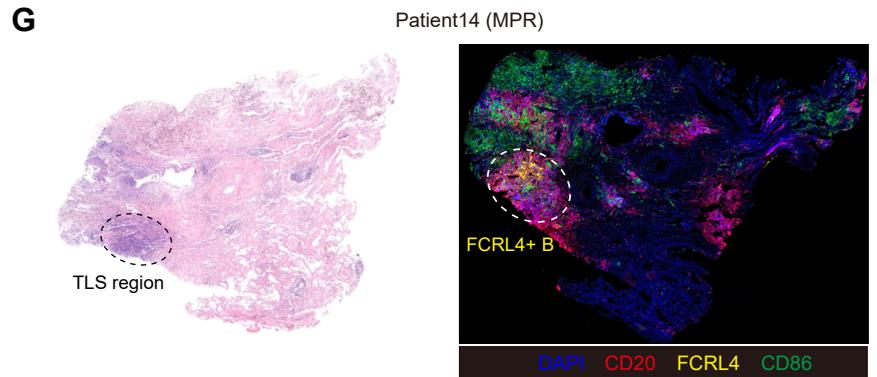
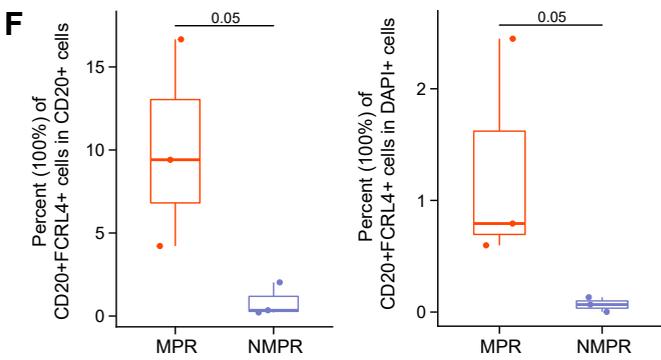
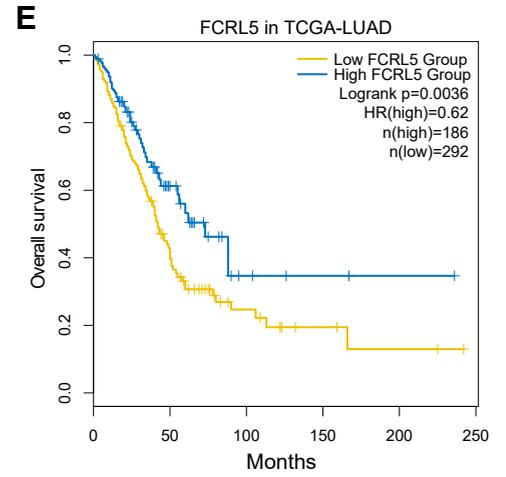
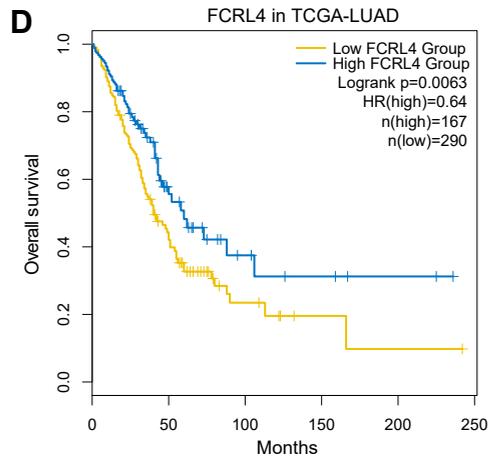
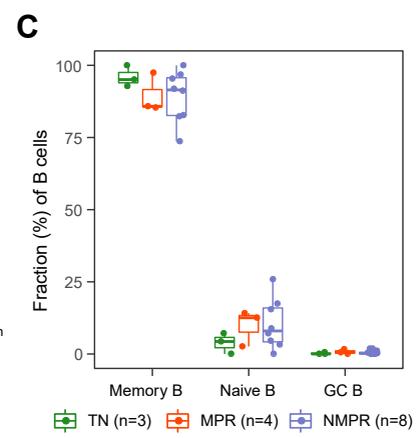
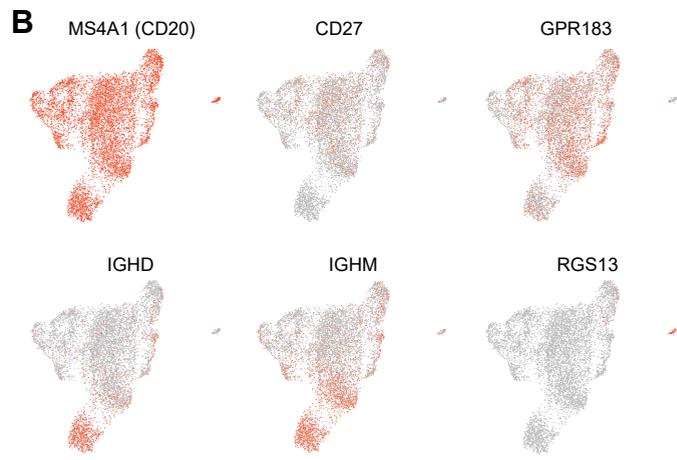
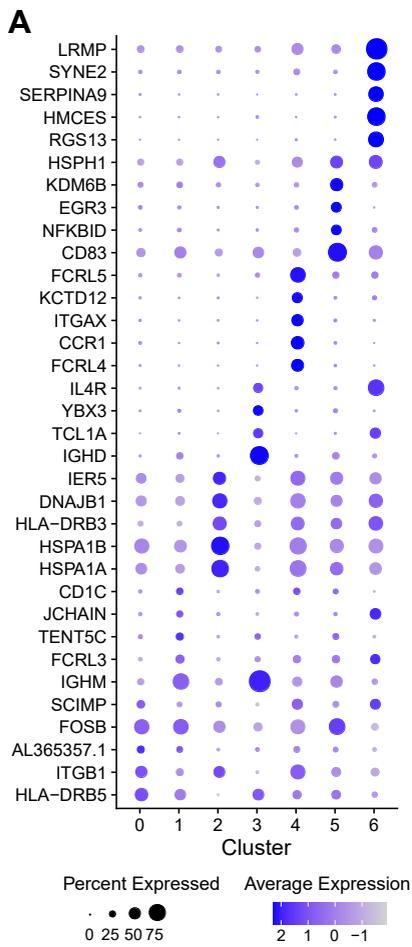


Fig. S6: Marker genes, immunofluorescence, and NicheNet analysis for B-cell subsets.

A, Dot plot of top 5 marker genes for each B cluster.

B, UMAP plots showing the expression of selected marker genes for the defined cell types.

C, Boxplot showing cellular fractions of each B cluster in TN (n=3), MPR (n=4) and NMPR (n=8) patients.

Center line indicates the median value, lower and upper hinges represent the 25th and 75th percentiles, respectively, and whiskers denote 1.5× interquartile range. All P values > 0.05, one-sided unpaired Wilcoxon test.

D, Kaplan–Meier survival curve of the expression of FCRL4 in TCGA-LUAD cohort.

E, Kaplan–Meier survival curve of the expression of FCRL5 in TCGA-LUAD cohort.

F, Quantification of FCRL4+ B cells from the multiplex immunofluorescence images.

G, HE and multiplex IF showing the location of tertiary lymphoid structures (TLS) and FCRL4+ B cells.

H, Heatmap showing potential ligands driving the phenotype of B4_FCRL4 cells.

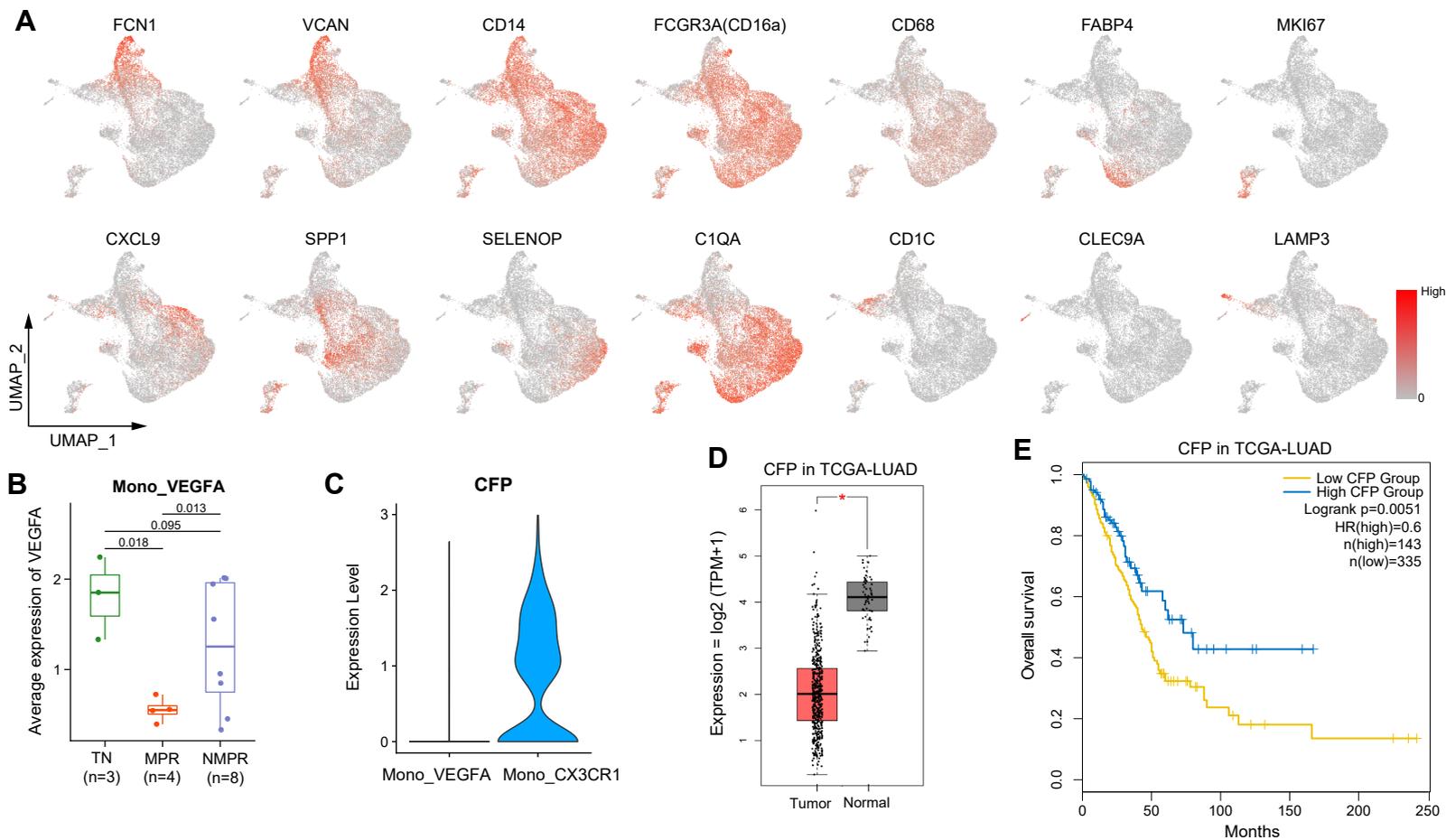


Fig. S7: Marker genes and survival analysis for myeloid subsets.

A, UMAP plots showing the expression of selected canonical marker genes for myeloid cells.

B, Boxplots of the average expression of VEGFA in Mono_VEGFA cells in TN (n=3), MPR (n=4) and NMPR (n=8) patients. Center line indicates the median, lower and upper hinges represent the 25th and 75th percentiles, respectively, and whiskers denote $1.5 \times$ interquartile range. Two-sided unpaired Wilcoxon test was used. One-sided t-test was used.

C, Violin plot of the expression of CFP in Mono_VEGFA and Mono_CX3CR1 cells.

D, Boxplots of the expression of CFP in TCGA-LUAD and normal lung tissues. *P < 0.05, ANOVA test.

E, Kaplan–Meier survival curve of the expression of CFP in TCGA-LUAD cohort.

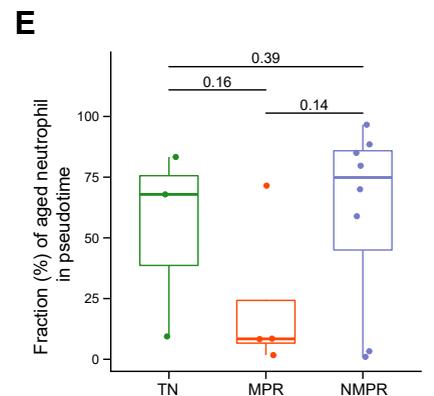
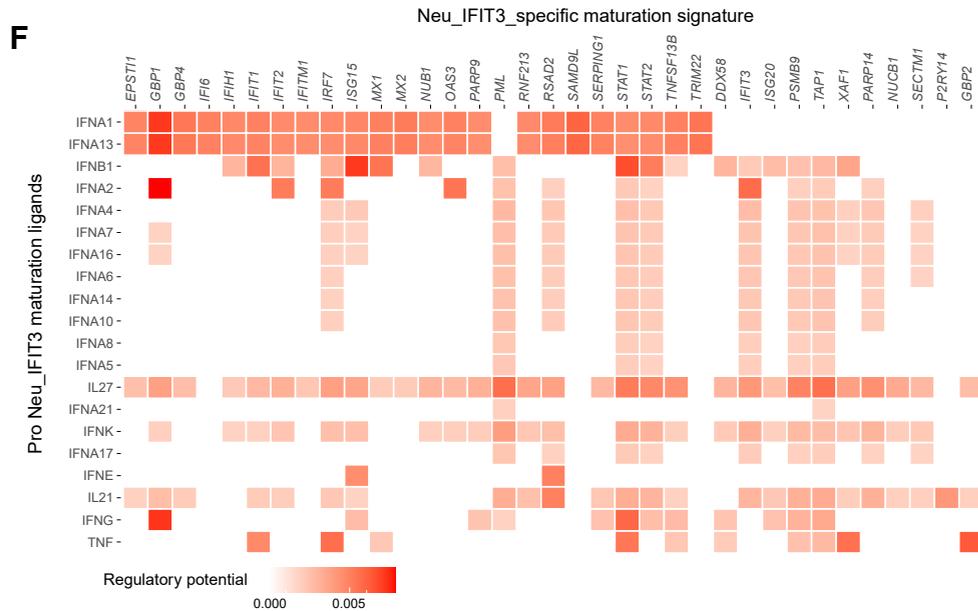
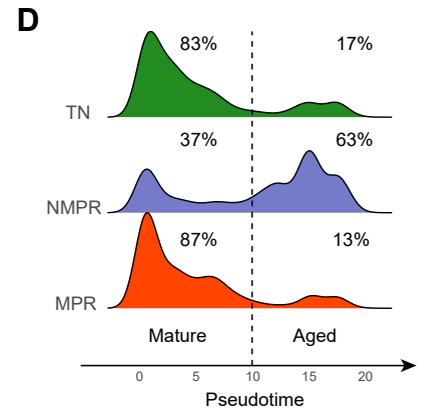
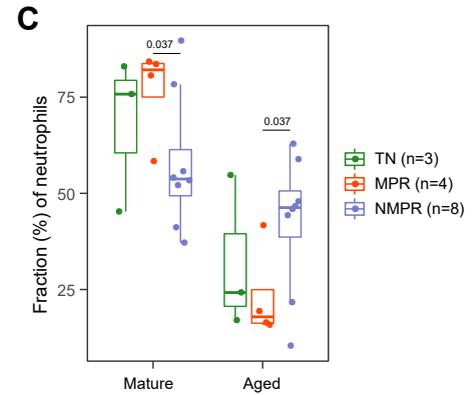
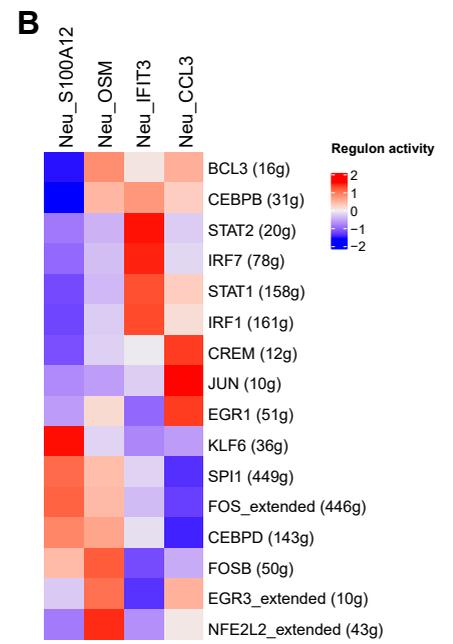
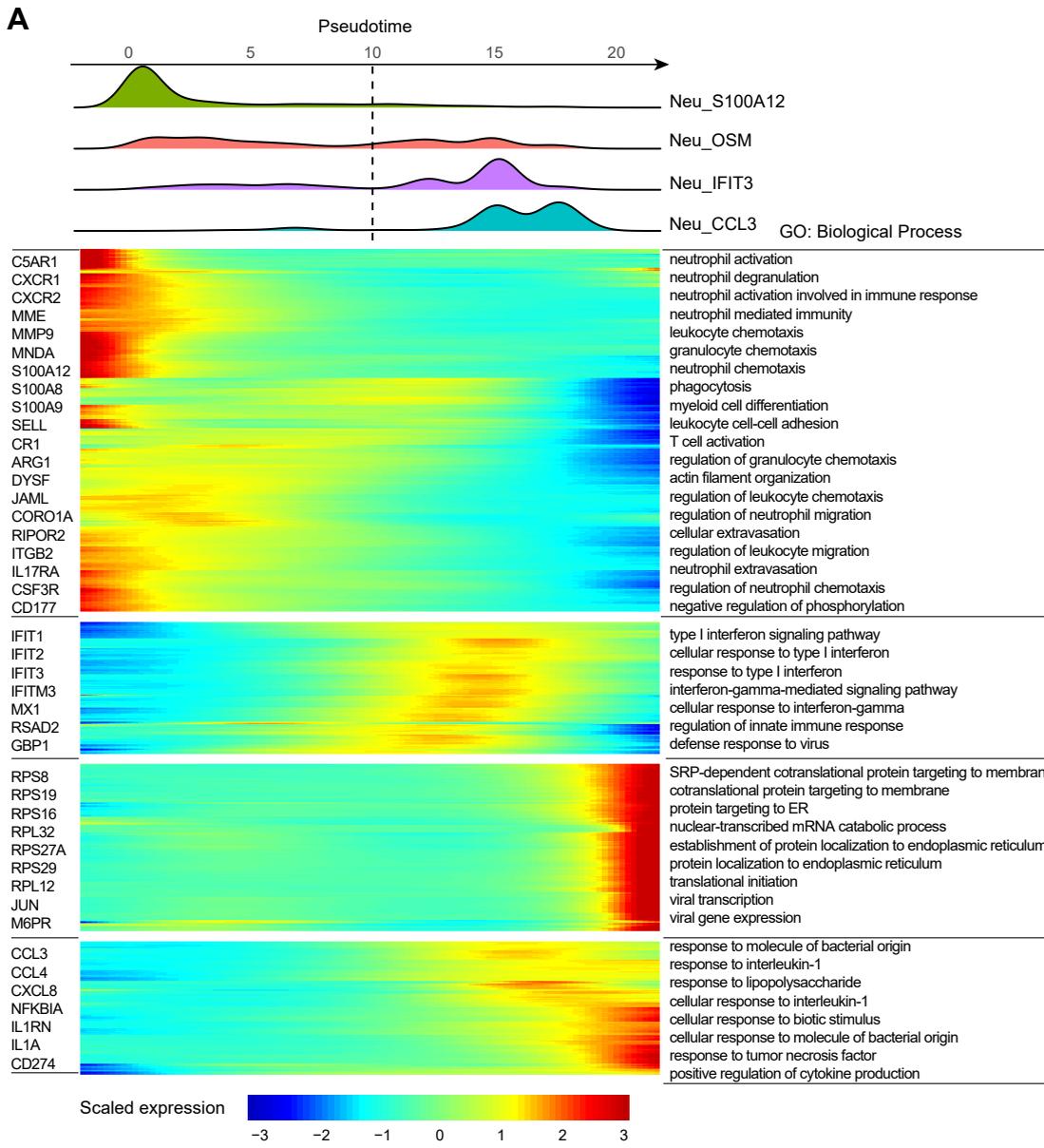


Fig. S8: Monocle, SCENIC and cellular fraction analysis for neutrophil subsets.

- A, Heatmap of the top differential genes in neutrophils along the pseudo-time (lower panel). The distribution of neutrophils during the transition along with the pseudo-time (upper panel). Subsets were labeled by colors (upper panel).
- B, Heatmap showing transcription factor activity for each neutrophil subset.
- C, Boxplot showing cellular fractions of mature (Neu_S100A12 and Neu_OSM) and aged (Neu_CCL3 and Neu_IFIT3) cells in TN (n=3), MPR (n=4) and NMPR (n=8) patients. Center line indicates the median value, lower and upper hinges represent the 25th and 75th percentiles, respectively, and whiskers denote 1.5× interquartile range. All differences with $P < 0.05$ are indicated; One-sided unpaired Wilcoxon test was used.
- D, The distribution of neutrophils during the transition (divided into 2 phases: mature and aged) along with the pseudo-time in TN, MPR and NMPR patients.
- E, Box plots of fractions of aged neutrophils in pseudo-time in TN (n=3), MPR (n=4) and NMPR (n=8) patients. One-sided unpaired Wilcoxon test was used.
- F, Heatmap showing potential ligands driving the phenotype of Neu_IFIT3 cells.