Cell Reports Medicine, Volume *4*

Supplemental information

Distinct immune microenvironment of lung adenocarcinoma in never-smokers from smokers

Wenxin Luo, Zhen Zeng, Yang Jin, Lan Yang, Ting Fan, Zhoufeng Wang, Yitong Pan, Ying Yang, Menglin Yao, Yangqian Li, Xue Xiao, Gang Wang, Chengdi Wang, Shuai Chang, Guowei Che, Li Zhang, Yalun Li, Yong Peng, and Weimin Li

Supplemental Figures

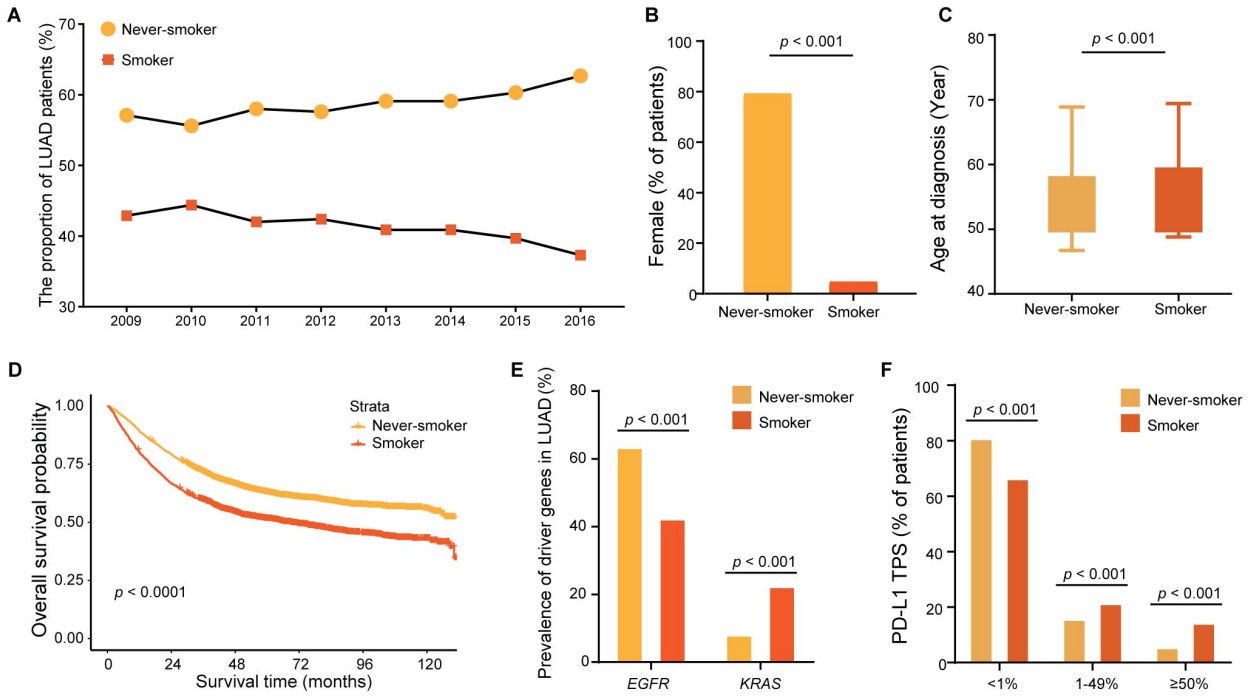

Figure S1. Clinicopathological features and prognostic differences between never-smoker and smoker LUADs from West China Hospital, Sichuan University, related to Figure 1.

A, The epidemiological changes of never-smoker and smoker LUADs from 2009 to 2016. **B,** The proportion of female in never-smoker and smoker groups, Chi-square test. **C,** The mean age (with SD) at diagnosis of never-smoker and smoker lung cancer patients, two-sided Wilcoxon Signed-Ranked test. **D,** The Kaplan-Meier survival curves of never-smoker and smoker patients. **E,** The proportion of LUAD patients harboring *EGFR* or *KRAS* driver mutations in never-smoker and smoker groups, Chi-square test. **F,** The proportion of patients with different PD-L1 expression, Chi-square test (tumor proportion score [TPS]: < 1%, 1-49%, and ≥ 50%).

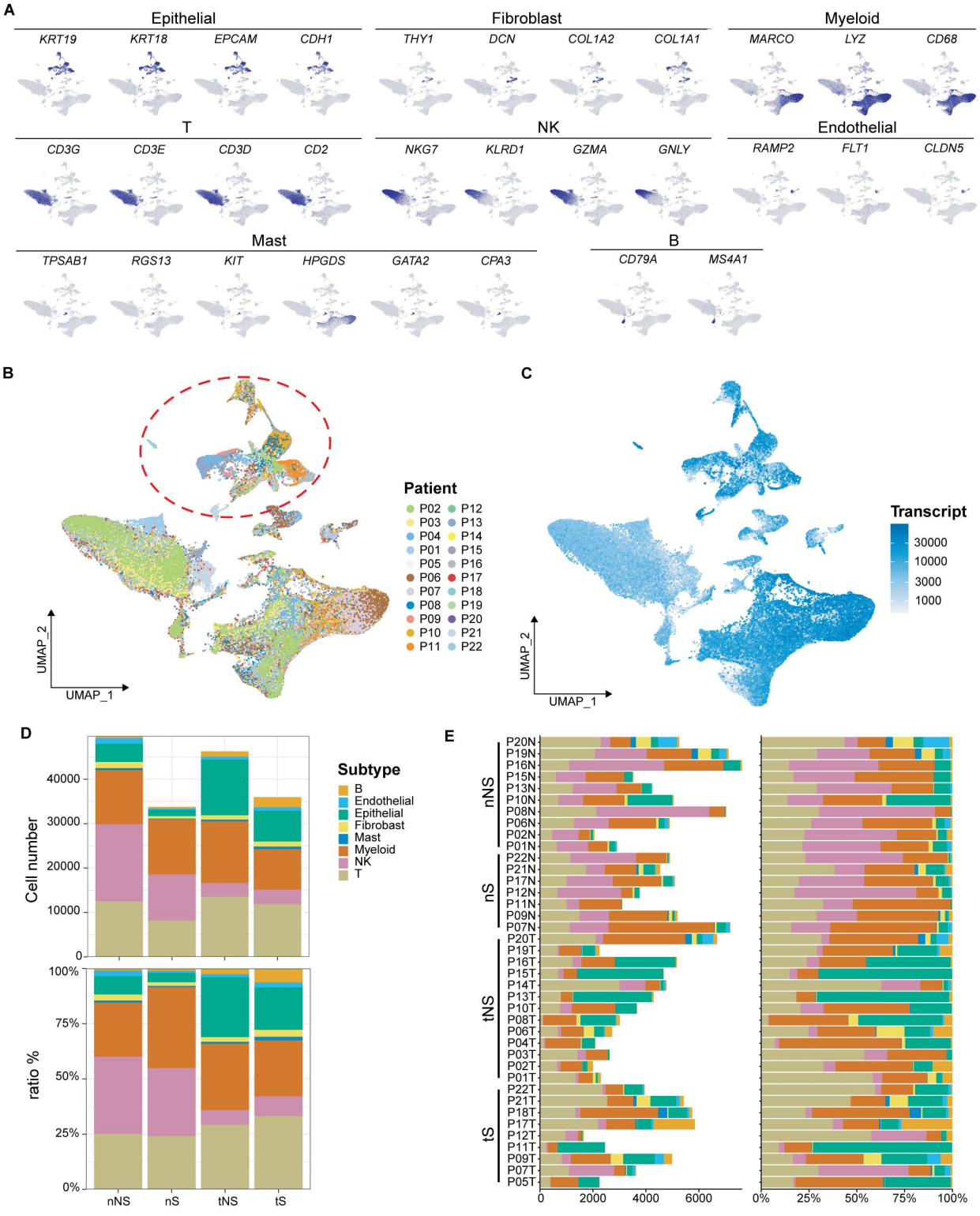
A
Epithelial
KRT19
KRT18
EPCAM
CDH1
Fibroblast
THY1
DCN
COL1A2
COL1A1
Myeloid
MARCO
LYZ
CD68
T
CD3G
CD3E
CD3D
CD2
NK
NKG7
KLRD1
GZMA
GNLY
Endothelial
RAMP2
FLT1
CLDN5
Mast
TPSAB1
RGS13
KIT
HPGDS
GATA2
CPA3
B
CD79A
MS4A1
B
Patient
P02 P12
P03 P13
P04 P14
P01 P15
P05 P16
P06 P17
P07 P18
P08 P19
P09 P20
P10 P21
P11 P22
UMAP_2
UMAP_1
C
Transcript
30000
10000
3000
1000
UMAP_2
UMAP_1
D
Subtype
B
Endothelial
Epithelial
Fibrobast
Mast
Myeloid
NK
T
Cell number
40000
30000
20000
10000
0
ratio %
100%
75%
50%
25%
0%
nNS
nS
tNS
tS
E
nNS
nS
tNS
tS
P20N
P19N
P16N
P15N
P13N
P10N
P08N
P06N
P02N
P01N
P22N
P21N
P17N
P12N
P11N
P09N
P07N
P20T
P19T
P16T
P15T
P14T
P13T
P10T
P08T
P06T
P04T
P03T
P02T
P01T
P22T
P21T
P18T
P17T
P12T
P11T
P09T
P07T
P05T
0
2000
4000
6000
0%
25%
50%
75%
100%

Figure S2. Eight cell types identified in this study, related to Figure 1.

A, Feature plots depicting single cell gene expression of canonical marker genes in eight types of cells. Blue color: high expression. **B** and **C,** The UMAP view of 165,753 cells, colored by 22 patients (**B**) and the number of transcripts (**C**). The epithelial cells were circled with red dotted lines. **D** and **E,** The number and its proportion of different type of cells in the four tissue groups (**D**) and in different tissues from each patient (**E**).

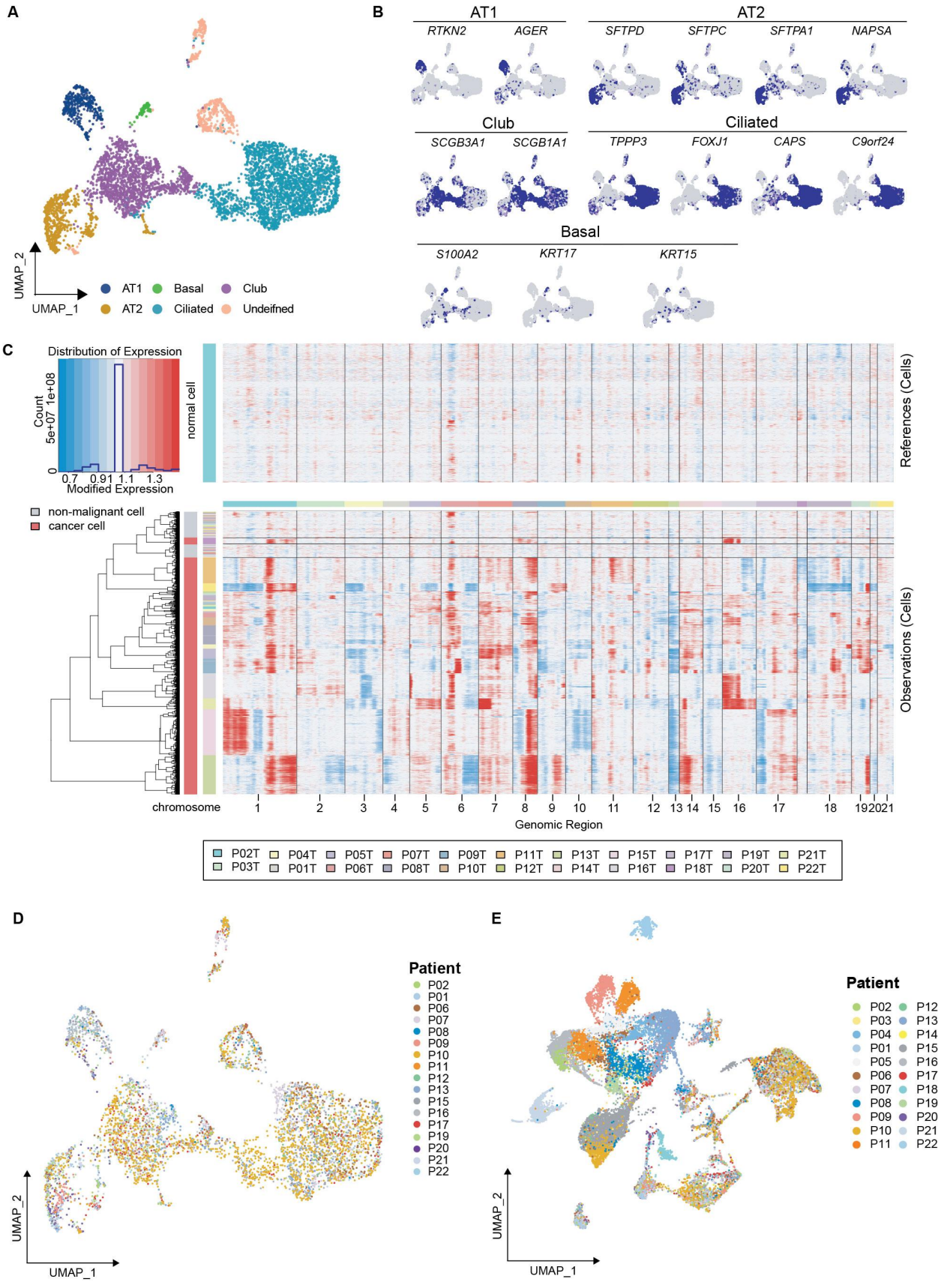
A
UMAP_2
UMAP_1
AT1
AT2
Basal
Ciliated
Club
Undeifned
B
AT1
RTKN2
AGER
AT2
SFTPD
SFTPC
SFTPA1
NAPSA
Club
SCGB3A1
SCGB1A1
Ciliated
TPPP3
FOXJ1
CAPS
C9orf24
Basal
S100A2
KRT17
KRT15
C
Distribution of Expression
Count
5e+07 1e+08
0
0.7 0.91 1.1 1.3
Modified Expression
normal cell
References (Cells)
non-malignant cell
cancer cell
Observations (Cells)
chromosome
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Genomic Region
P02T P03T P04T P01T P05T P06T P07T P08T P09T P10T P11T P12T P13T P14T P15T P16T P17T P18T P19T P20T P21T P22T
D
Patient
P02
P01
P06
P07
P08
P09
P10
P11
P12
P13
P15
P16
P17
P19
P20
P21
P22
UMAP_2
UMAP_1
E
Patient
P02 P12
P03 P13
P04 P14
P01 P15
P05 P16
P06 P17
P07 P18
P08 P19
P09 P20
P10 P21
P11 P22
UMAP_2
UMAP_1

Figure S3. Subtypes of epithelial cells identified in this study, related to Figure 2.

A, UMAP view of epithelial cells from normal lung tissues, colored by six subtypes. **B,** Feature plots colored by expression (blue color: high expression) of canonical marker genes in epithelial cell subsets. **C,** Heatmap showing inferred copy-number variations (CNVs) profiles based on the scRNA-seq data of tumor cells from each patient, distinguishing cancer cells from non-malignant cells. Red, gene amplifications; blue, gene deletions. **D,** UMAP view of epithelial cells from normal lung tissues, colored by 22 LUAD patients. **E,** UMAP view of all epithelial cells from normal lung tissues and tumor tissues, colored by each patient.

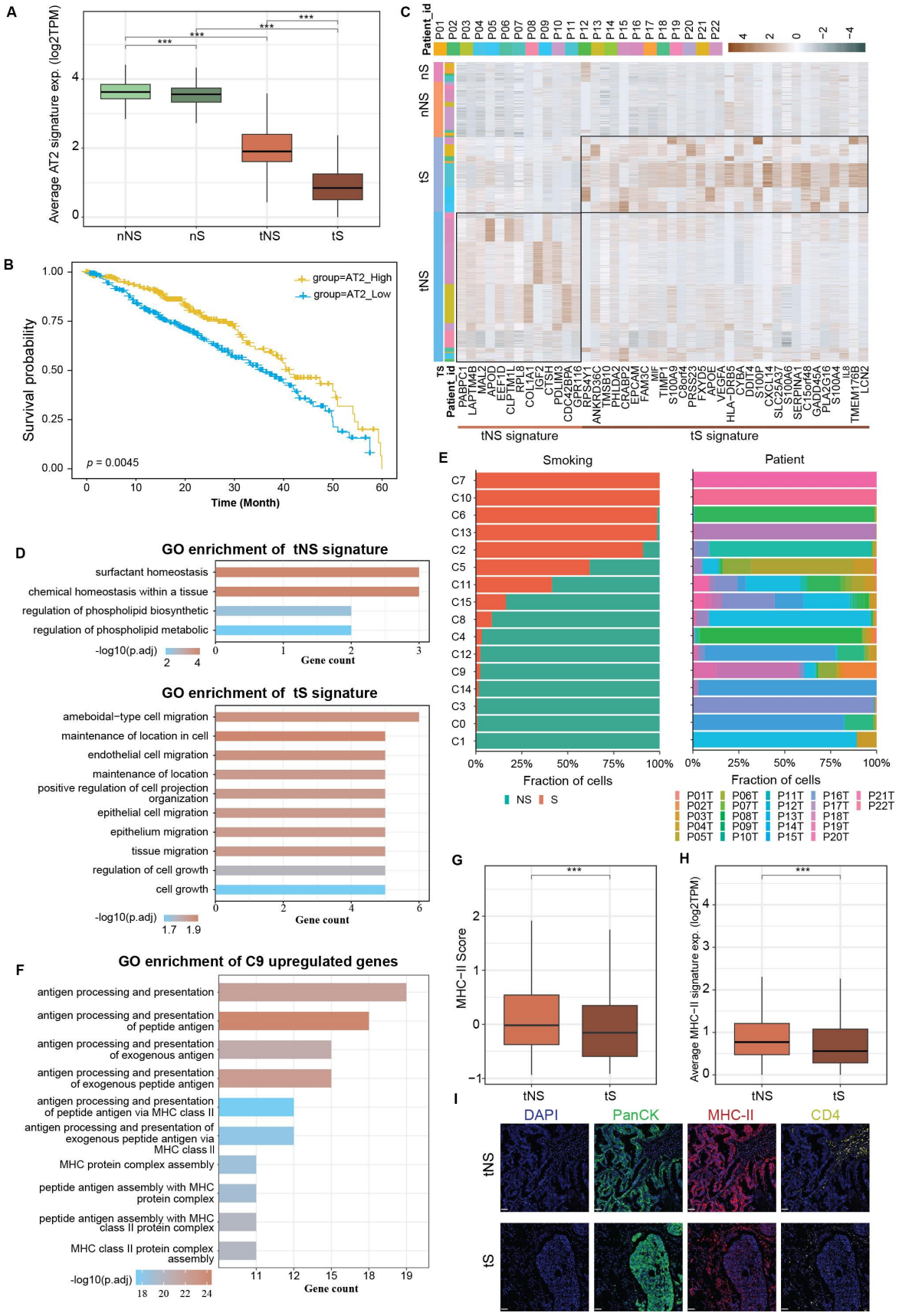
A
Average AT2 signature exp. (log2TPM)
nNS
nS
tNS
tS
B
Survival probability
group=AT2_High
group=AT2_Low
p = 0.0045
Time (Month)
C
Patient_id
nS
nNS
tS
tNS
tNS signature
tS signature
D
GO enrichment of tNS signature
surfactant homeostasis
chemical homeostasis within a tissue
regulation of phospholipid biosynthetic
regulation of phospholipid metabolic
-log10(p.adj)
Gene count
GO enrichment of tS signature
ameboidal−type cell migration
maintenance of location in cell
endothelial cell migration
maintenance of location
positive regulation of cell projection organization
epithelial cell migration
epithelium migration
tissue migration
regulation of cell growth
cell growth
F
GO enrichment of C9 upregulated genes
antigen processing and presentation
antigen processing and presentation of peptide antigen
antigen processing and presentation of exogenous antigen
antigen processing and presentation of exogenous peptide antigen
antigen processing and presentation of peptide antigen via MHC class II
antigen processing and presentation of exogenous peptide antigen via MHC class II
MHC protein complex assembly
peptide antigen assembly with MHC protein complex
peptide antigen assembly with MHC class II protein complex
MHC class II protein complex assembly
E
Smoking
Patient
Fraction of cells
NS
S
G
MHC−II Score
H
Average MHC−II signature exp. (log2TPM)
I
DAPI
PanCK
MHC-II
CD4
tNS
tS

Figure S4. Significance of tNS/tS signatures in cancer cells, related to Figure 2.

A, Boxplots showing the expression levels of AT2 signature in the four tissue types of GSE131907, ***$p < 0.001$, two-sided Wilcoxon Signed-Ranked test. **B**, Kaplan-Meier curves for the overall survival of TCGA-LUAD patients (n = 453), which were divided into two groups according to the mean expression of AT2 signature. +: censored observations, *p*-value was calculated using the two-sided log-rank test. **C,** Heatmap showing the expression levels of tS signature and tNS signature in each patient. **D**, Gene Ontology [GO] enrichment of tNS and tS signatures. **E,** Proportion of cancer cells from smoking status (left) and different patients (right) in each cluster sorted by number of cells. **F,** GO enrichment of C9 up-gregulated genes identified by Seurat FindMarker function. **G and H,** Boxplots showing the expression levels of MHC-II signature in cancer cells from never-smokers and smokers of our dataset (G) and GSE131907 (H). **I**, Multiplex immunohistochemistry (mIHC) staining of PanCK, MHC-II and CD4 in LUAD tumor tissues from never-smokers and smokers. Nuclei (blue), PanCK (green), MHC-II (red), and CD4 (yellow). Scale bars, 50 μm. At least three independent samples per tissue types.

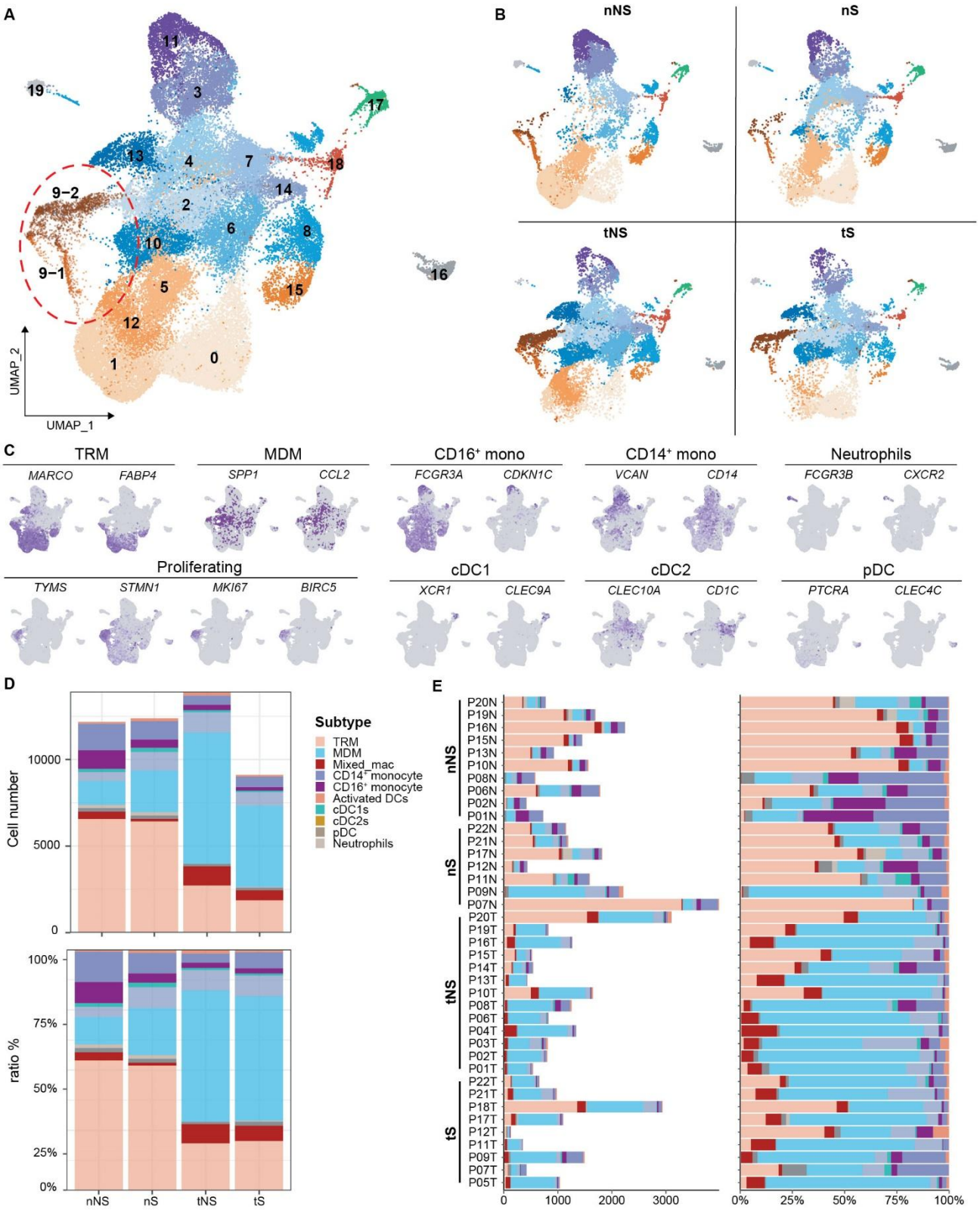
A
11
3
19
17
13
4
7
18
9−2
14
2
6
8
10
16
9−1
5
15
12
1
0
UMAP_2
UMAP_1
B
nNS
nS
tNS
tS
C
TRM
MARCO
FABP4
MDM
SPP1
CCL2
CD16⁺ mono
FCGR3A
CDKN1C
CD14⁺ mono
VCAN
CD14
Neutrophils
FCGR3B
CXCR2
Proliferating
TYMS
STMN1
MKI67
BIRC5
cDC1
XCR1
CLEC9A
cDC2
CLEC10A
CD1C
pDC
PTCRA
CLEC4C
D
Cell number
10000
5000
0
ratio %
100%
75%
50%
25%
0%
nNS
nS
tNS
tS
Subtype
TRM
MDM
Mixed_mac
CD14⁺monocyte
CD16⁺ monocyte
Activated DCs
cDC1s
cDC2s
pDC
Neutrophils
E
nNS
nS
tNS
tS
P20N
P19N
P16N
P15N
P13N
P10N
P08N
P06N
P02N
P01N
P22N
P21N
P17N
P12N
P11N
P09N
P07N
P20T
P19T
P16T
P15T
P14T
P13T
P10T
P08T
P06T
P04T
P03T
P02T
P01T
P22T
P21T
P18T
P17T
P12T
P11T
P09T
P07T
P05T
0
1000
2000
3000
0%
25%
50%
75%
100%

Figure S5. Subsets of myeloid cells identified in this study, related to Figure 3.

A, UMAP view of 47,580 myeloid cells, colored by clusters, cells in cluster 9 (Mixed_mac) were circled. **B,** UMAP view of different myeloid cells by tissue types. **C,** Feature plots colored by expression (blue: high expression) of canonical marker genes in myeloid cell subsets. **D** and **E,** The number and its proportion of different myeloid cell subsets in four tissues (**D**) and in different tissue from each patient (**E**).

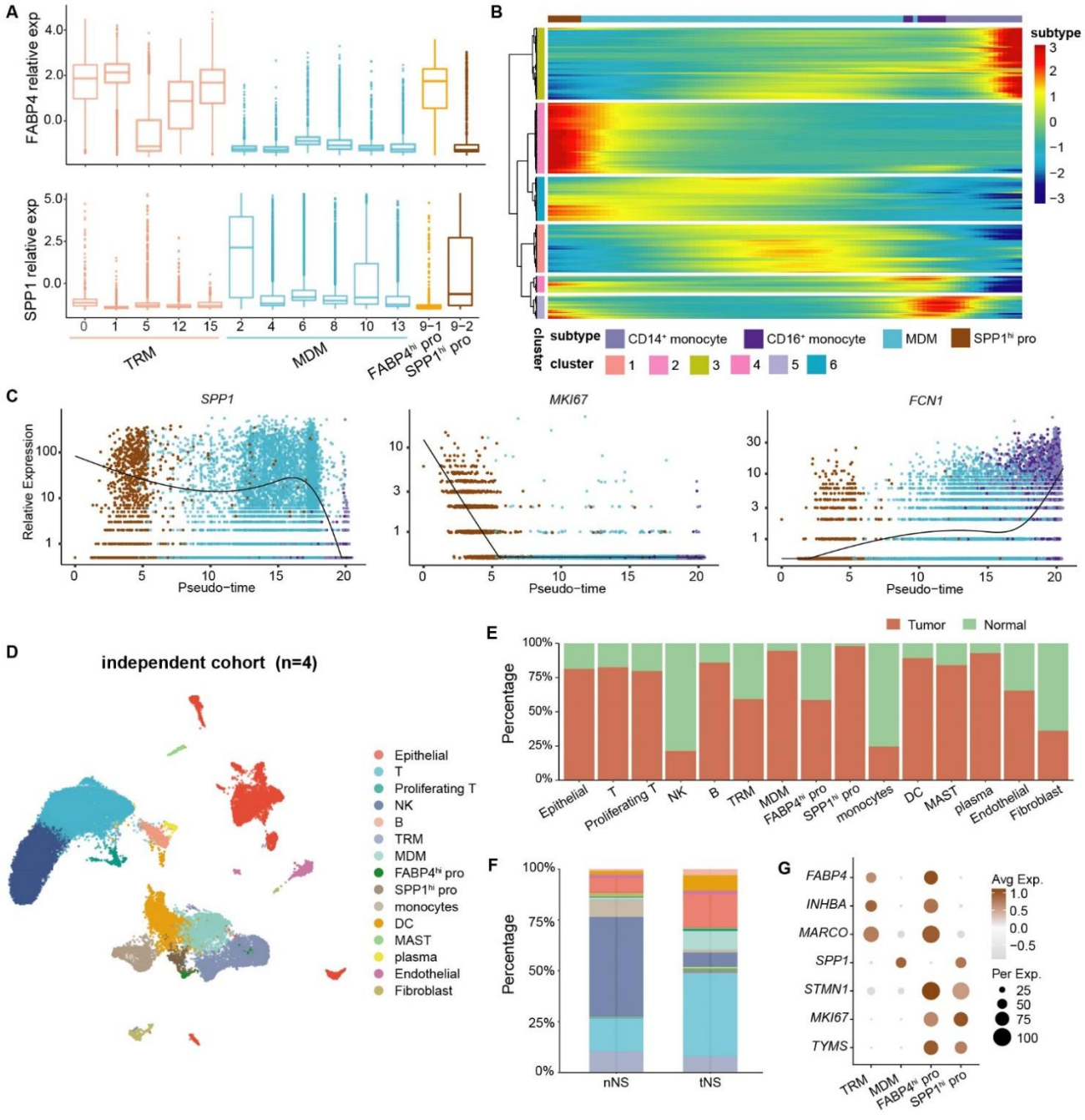

Figure S6. Characteristics of SPP1^{hi} pro cells identified in this study, related to Figure 3.

A, Expression level of *FABP4* (top) and *SPP1* (bottom) in TRM, MDM and FABP4^{hi} pro and SPP1^{hi} pro. **B,** Heatmap of differentially expressed genes arranged in pseudo-temporal patterns. **C,** Relative expression of different marker genes. **D,** UMAP view of cells from other independent LUADs, colored by different cell subtypes. **E,** Proportion of different cell subsets in each tissue type from independent validation datasets. **F,** Relative proportions of cell types in different tissue types. **G,** Dot plot of mean expression of selected marker genes in four populations of macrophages from independent validation dataset.

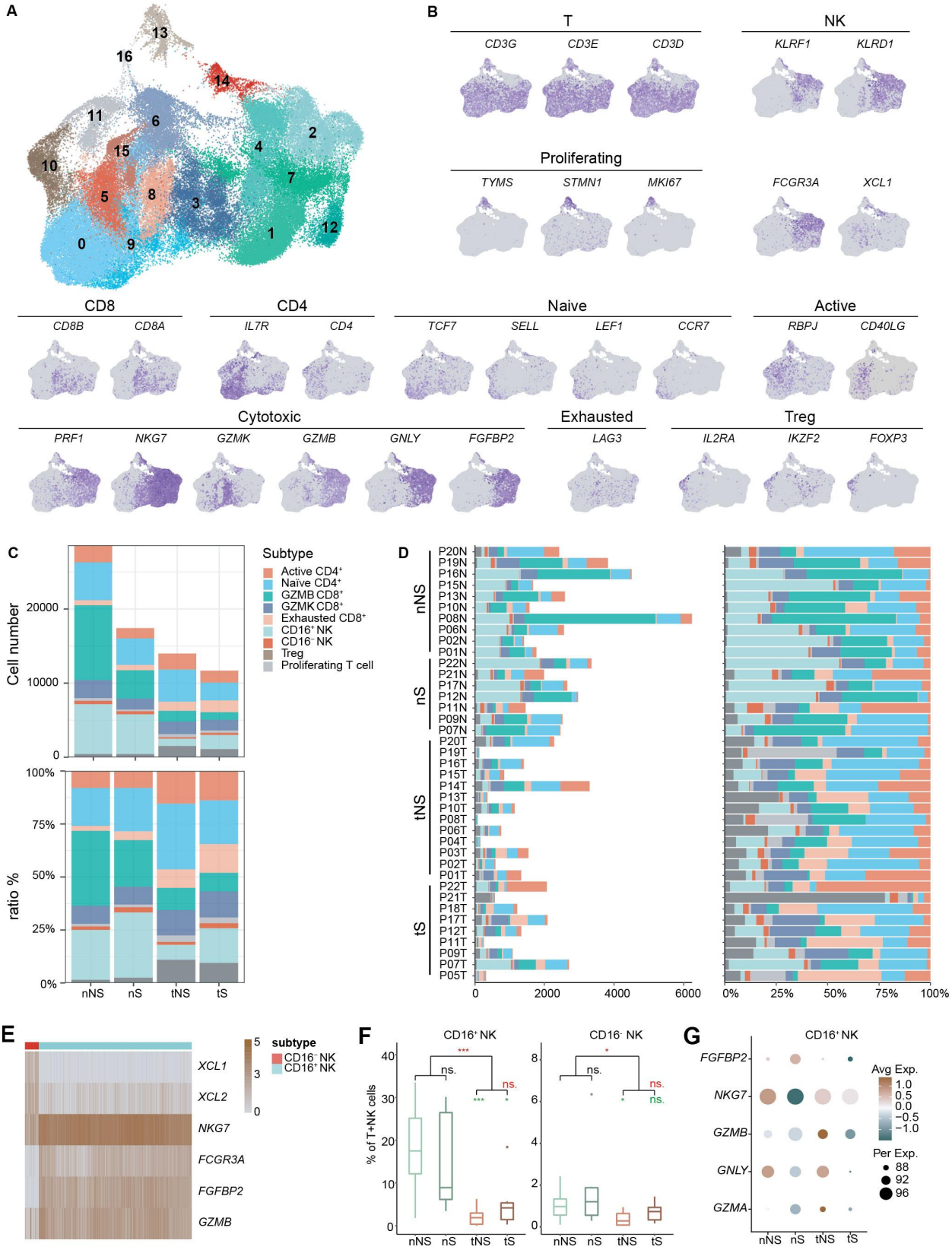
A
B
T
CD3G
CD3E
CD3D
NK
KLRF1
KLRD1
Proliferating
TYMS
STMN1
MKI67
FCGR3A
XCL1
CD8
CD8B
CD8A
CD4
IL7R
CD4
Naive
TCF7
SELL
LEF1
CCR7
Active
RBPJ
CD40LG
Cytotoxic
PRF1
NKG7
GZMK
GZMB
GNLY
FGFBP2
Exhausted
LAG3
Treg
IL2RA
IKZF2
FOXP3
C
Cell number
ratio %
Subtype
Active CD4+
Naïve CD4+
GZMB CD8+
GZMK CD8+
Exhausted CD8+
CD16+ NK
CD16- NK
Treg
Proliferating T cell
nNS
nS
tNS
tS
D
E
subtype
CD16- NK
CD16+ NK
XCL1
XCL2
NKG7
FCGR3A
FGFBP2
GZMB
F
CD16+ NK
CD16- NK
% of T+NK cells
ns.
G
CD16+ NK
FGFBP2
NKG7
GZMB
GNLY
GZMA
Avg Exp.
Per Exp.

Figure S7. Subsets of T/NK cells identified in this study, related to Figure 6.

A, UMAP view of 80,194 T/NK cells, colored by different clusters. **B,** Feature plots colored by expression (blue: high expression) of canonical marker genes in T/NK cell subsets. **C,** The number and its proportion of major T/NK cell subsets across each tissue type. **D,** Cell number and proportion of major T/NK cell subsets in different tissue from each patient. **E,** Heatmap of highly expressed genes in $CD16^{+}$ NK and $CD16^{-}$ NK, respectively. **F,** Proportion of $CD16^{+}$ NK and $CD16^{-}$ NK in T/NK cells in four types of tissues. Asterisks represent the significance of differences in the following comparisons: all tumor tissues vs. all normal lung tissues (brown), nS vs. nNS (black), tS vs. tNS (red), tNS vs. nNS and tS vs. nS (green). **G,** Dot plot of mean expression of cytotoxicity related genes across four tissue types.