

Figure S1

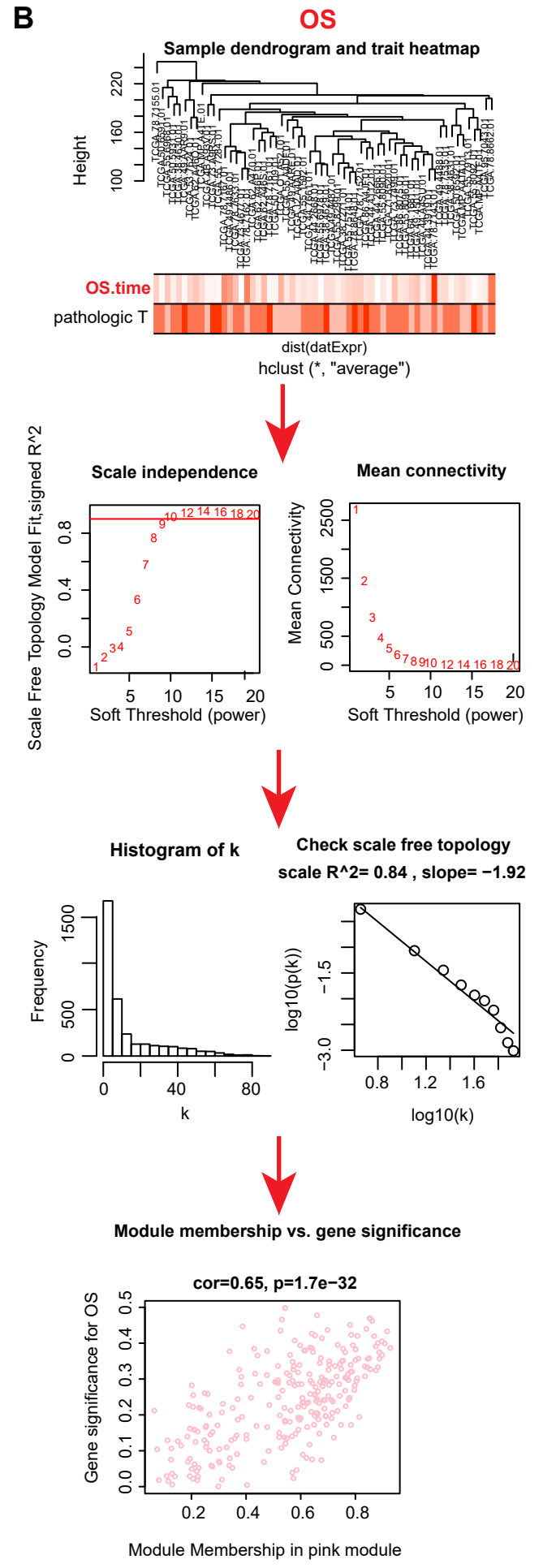
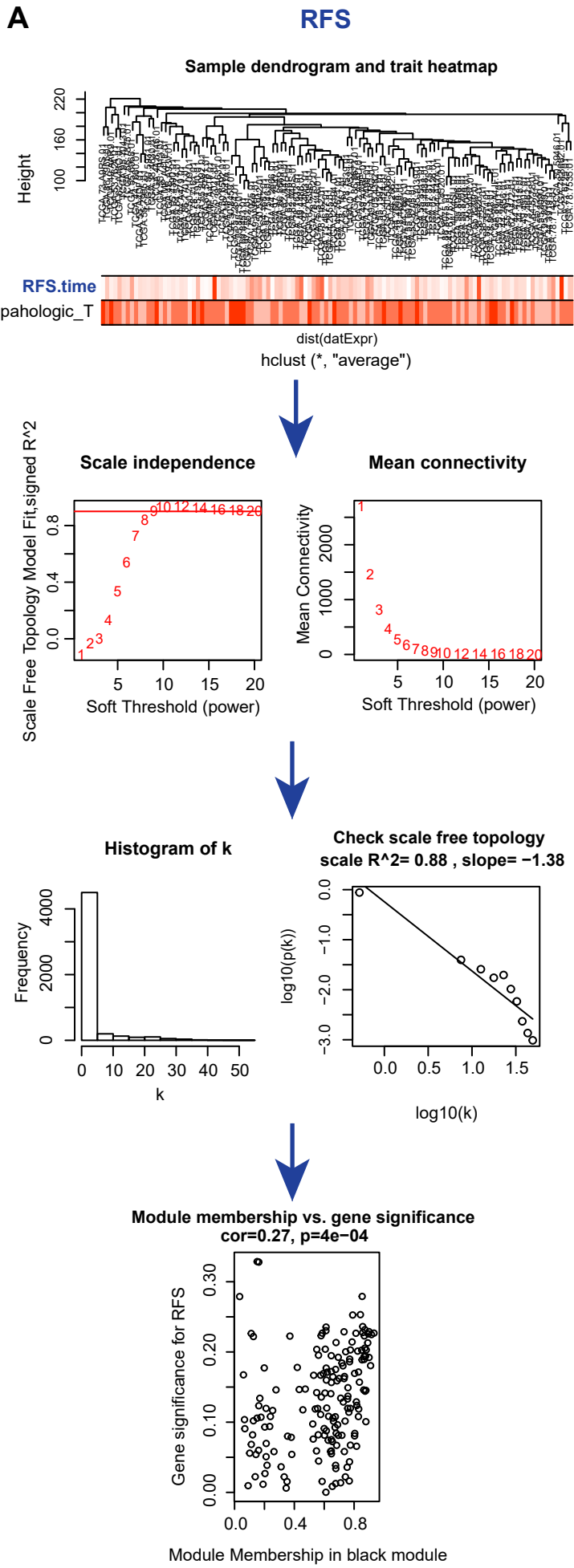


Figure S1. WGCNA analysis showing the gene modules predicting RFS and OS in pN0-stage LUAD

A, B, WGCNA analysis of gene expression generated by unsupervised hierarchical clustering. The upper panel shows the sample dendrogram and trait heatmap. The color row underneath the dendrogram shows the module assignment determined by the dynamic tree cut. The middle panel shows a histogram of network connectivity and log-log plot of the same histogram. The bottom panel shows the correlation between GS (gene signature, correlation between gene significance) and MM (molecular membership, module membership degree) in black and pink modules.

Figure S2

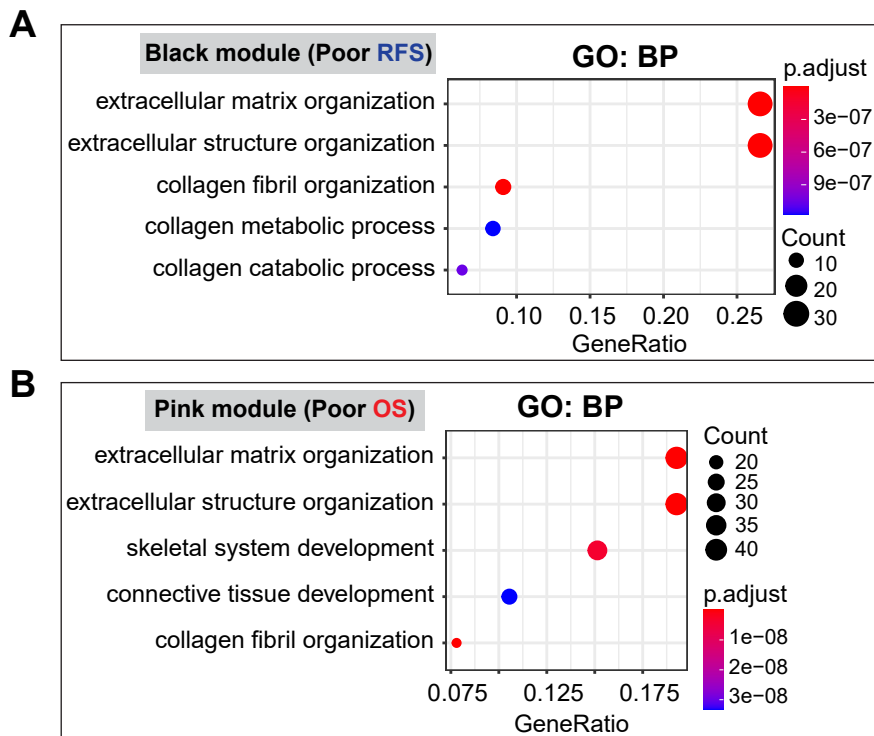


Figure S2. Pathway enrichment analysis of genes in yellow and brown modules

A, B, GO-BP (Gene Ontology Biological Processes) enrichment analyses of genes in black (related to RFS; A) and pink (related to OS; B) modules.

Figure S3

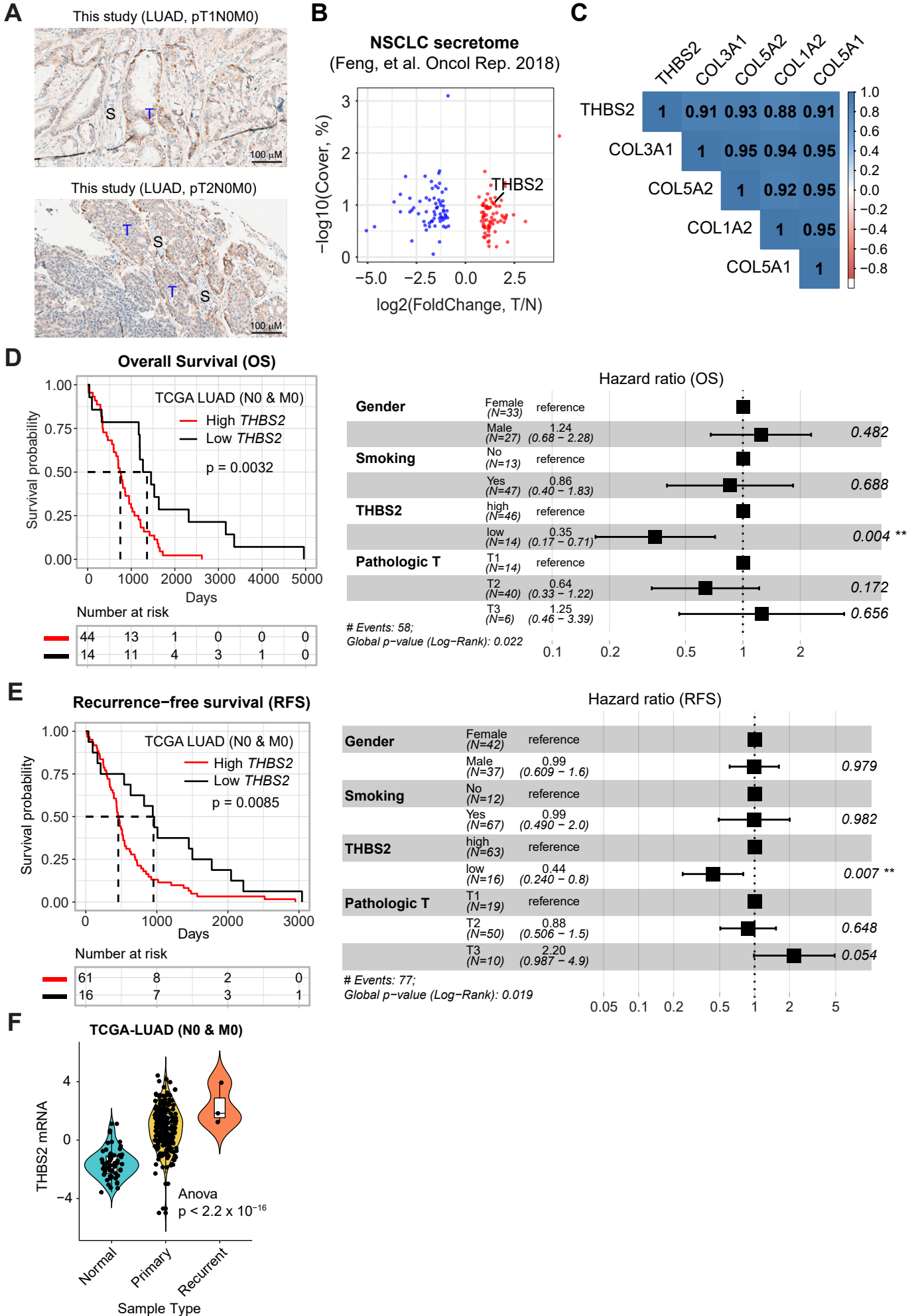


Figure S3. THBS2 predicts the prognosis of pN0-stage LUAD

A, Representative images showing the location and expression of THBS2 in early-stage LUAD.

B, Mining public secretome profiling of non-small cell lung cancer (NSCLC) samples. Red dots indicate the significantly ($p < 0.05$) upregulated secreted proteins from the tissue culture medium of lung cancer patients compared with matched normal lung tissue. Of note, THBS2 is highly secreted by NSCLC tumors.

C, Correlation matrix showing the correlations among the top 5 genes (THBS2, COL3A1, COL5A2, COL1A2, COL5A1) that were predictive RFS and OS (related to Figure 2G).

D, E, Univariate (left) and multivariate (right, forest plot) analyses showing the association of THBS2 expression with OS (D) and RFS (E) in patients with pN0-stage LUAD.

F, THBS2 gene expression across TCGA pN0-stage primary, recurrent LUAD and normal lung tissue.

Figure S4

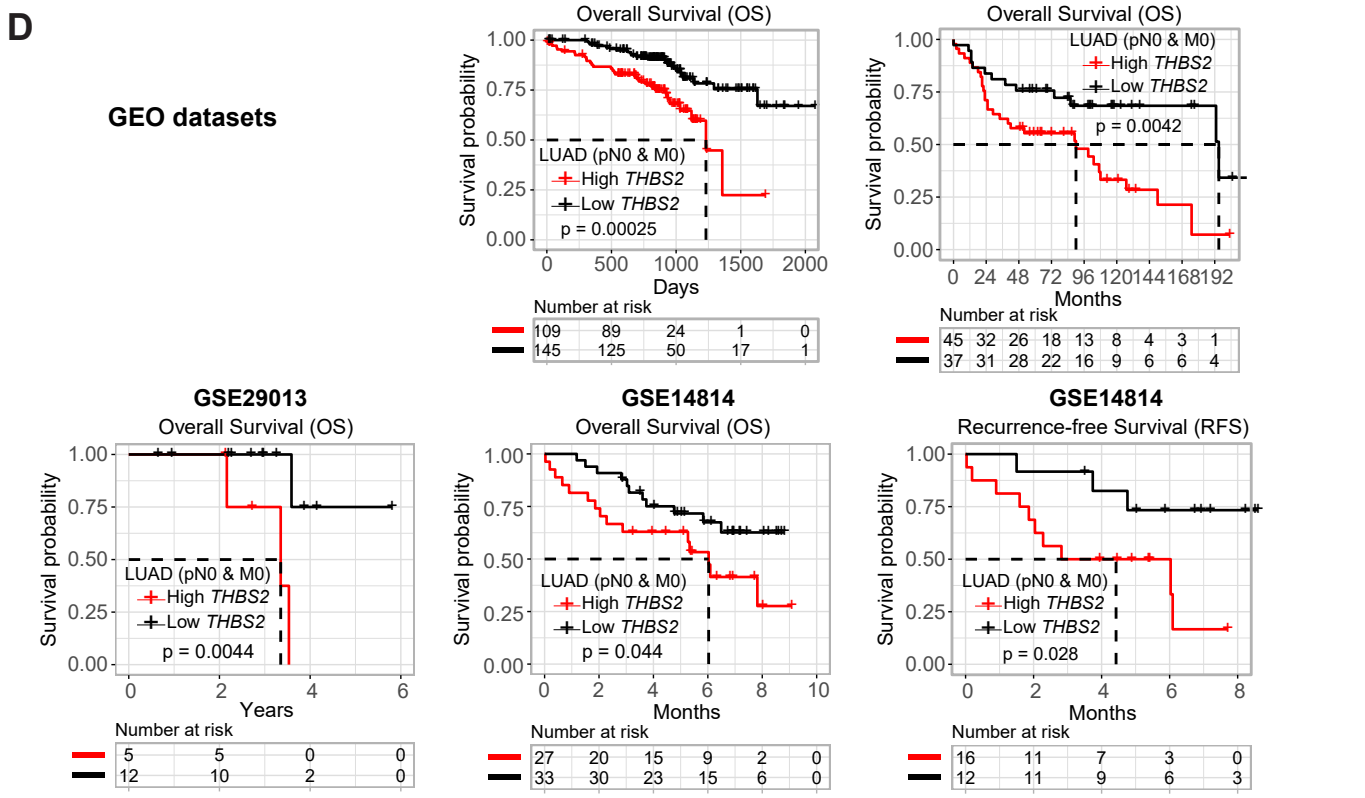
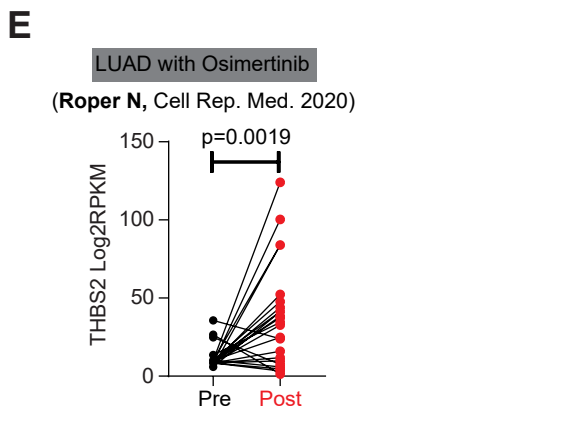
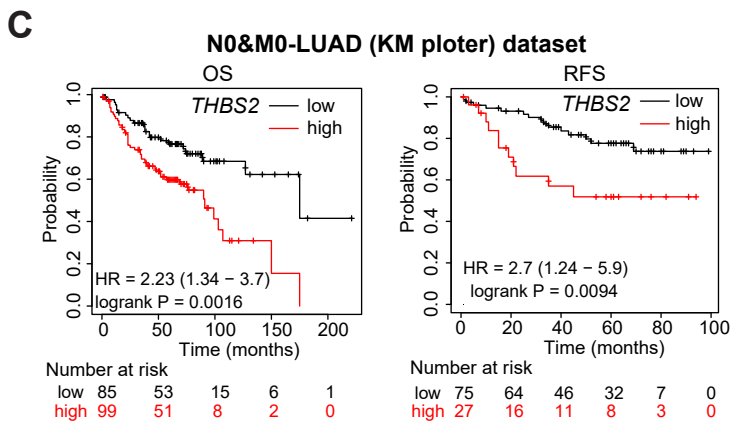
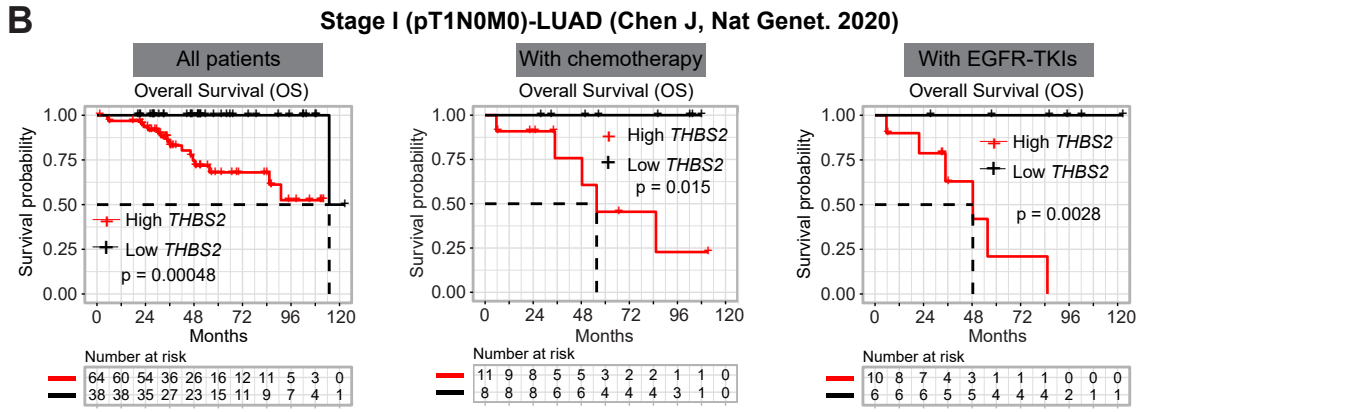
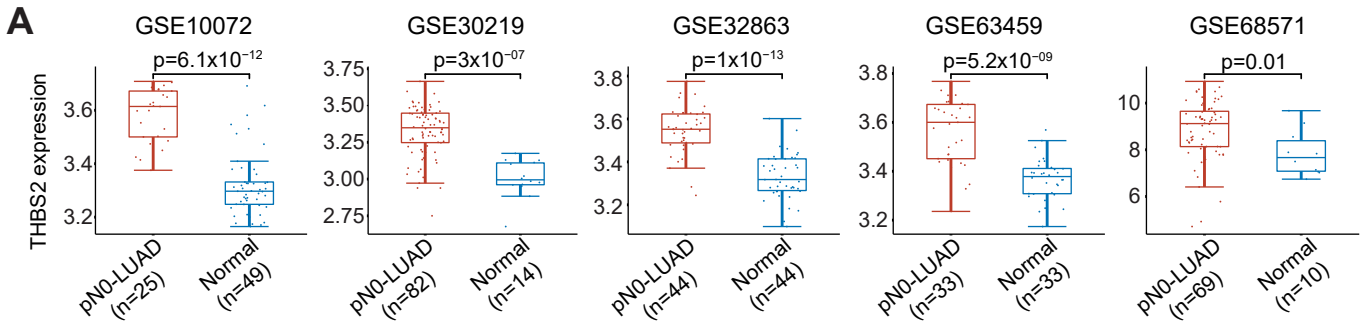


Figure S4. Cross-validation of THBS2 as a prognostic biomarker in early pN0-stage LUAD

A, THBS2 gene expression across pN0M0-stage LUAD and adjacent normal lung tissue from multiple public cohorts. B-D, The association of THBS2 gene expression with the survival of pN0M0-stage LUAD across multiple public cohorts. In B, data were downloaded from Chen J, Nat Genet. 2020; in C, data were downloaded from the Kaplan-Meier plotter portal (Kaplan-Meier plotter [Lung] (kmplot.com)); in D, data were downloaded from GEO datasets (<https://www.ncbi.nlm.nih.gov/geo/>). Of note, in B, pN0M0-stage LUAD patients who received adjuvant chemotherapy (middle) or EGFR-TKI targeted therapy (right) were specifically investigated. In addition, of the 5 datasets shown in Fig. S4A, there were only transcriptomic data of early-stage LUAD and matched normal lung but without the survival data in GSE10072 and GSE32863. Although the survival data in GSE63459 were available, the cancer-related death only occurred in 5 patients in the entire cohort, leading to the fact that there are no sufficient endpoint events for analysis. For GSE68571, the survival data were provided, but our analysis (not shown) suggested a dramatical survival difference between the THBS2-high vs. THBS2-low pN0-stage LUAD, although the significance was not reached, which is largely due to the short follow-up in this study cohort because the median survival in the two subgroups was not reached.

E, Differentially THBS2 expression (log2RPKM [Reads Per Kilobase per Million mapped reads]) in *EGFR*-mutant LUAD tumors with pre- and post-osimertinib (a third-generation of epidermal growth factor receptor (EGFR)-tyrosine kinase inhibitor (TKI)) treatment. Transcriptomic data were downloaded from Roper N, Cell Rep. Med. 2020. P-value was calculated by paired student's t-test.

Figure 5

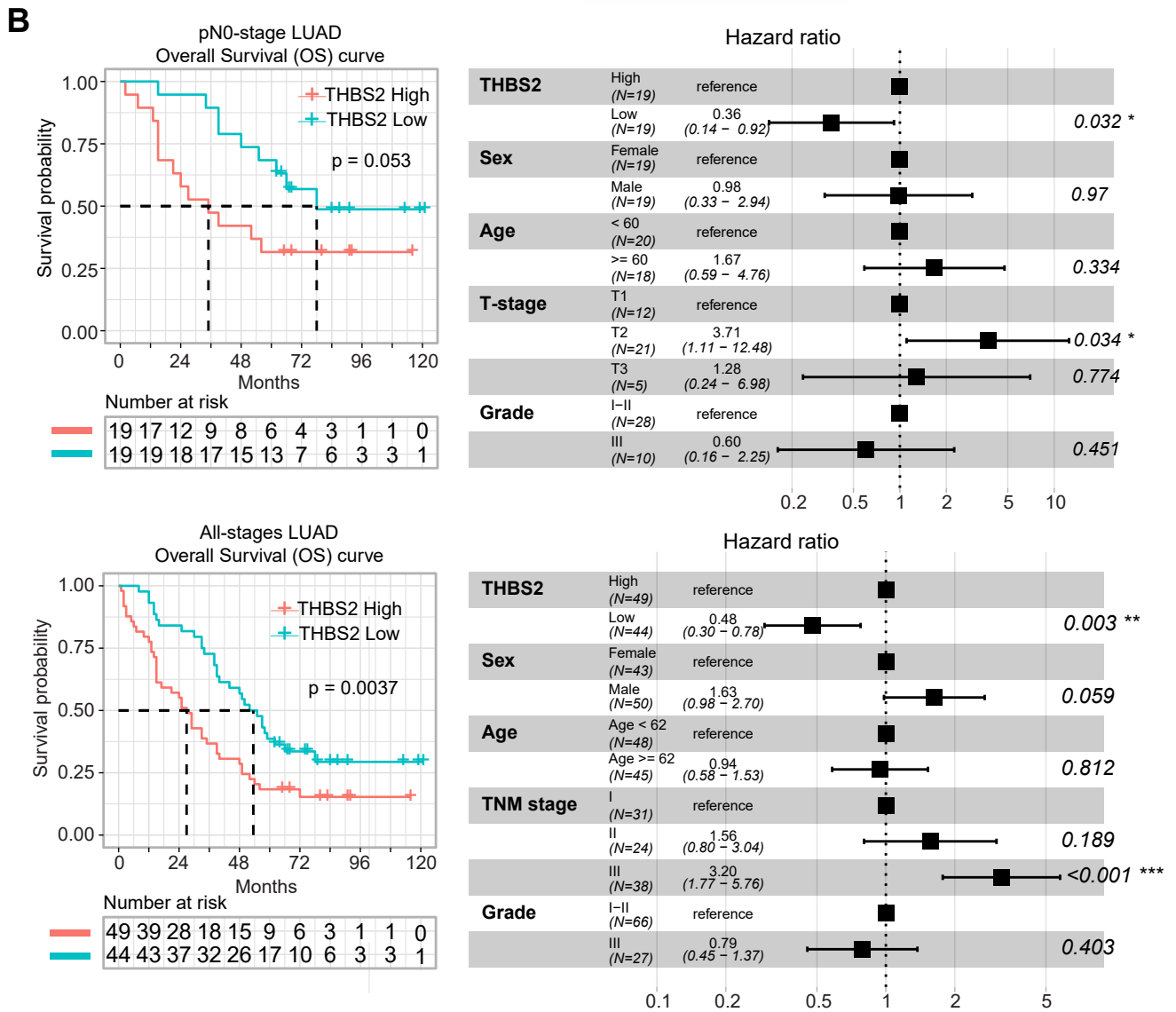
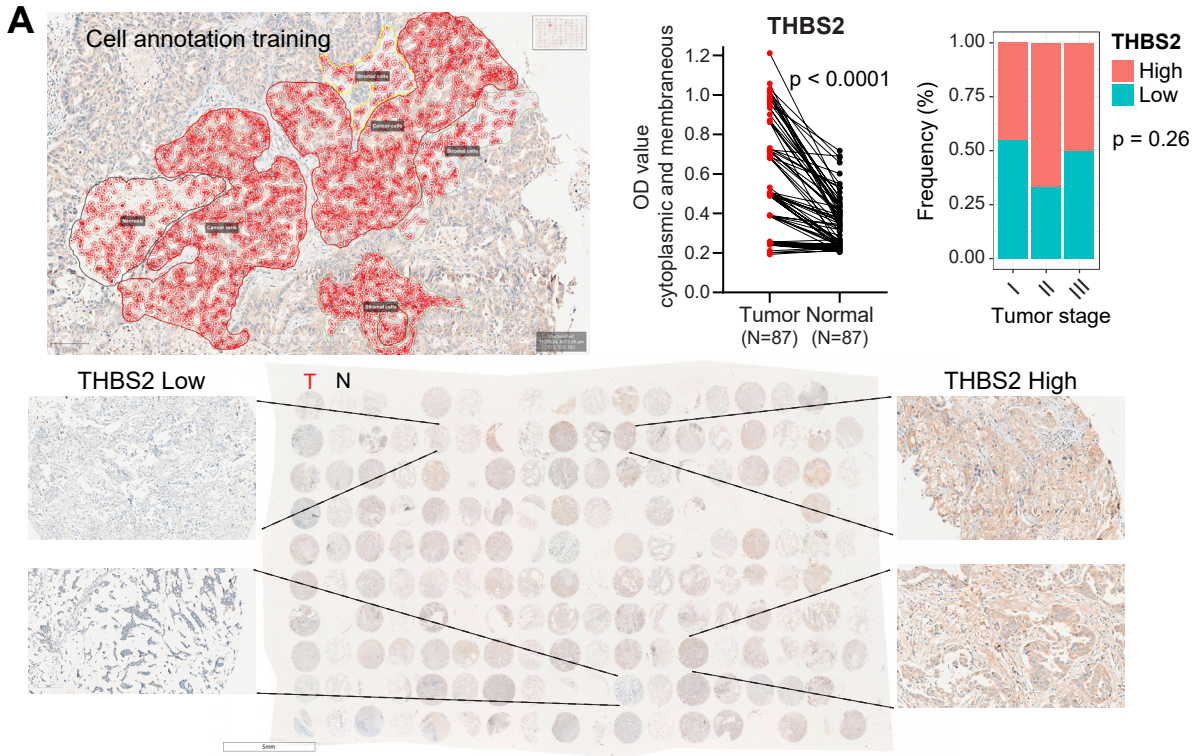


Figure S5. Internal LUAD cohort validating THBS2 as a prognostic biomarker in early pN0-stage LUAD

A, B, Our internal treatment-naïve LUAD cohort (tissue microarray data; N = 93) showing the association of THBS2 protein levels with OS after surgery. A, Data quantifications were performed at the single-cell level using Qupath software. B, Univariate (left) and multivariate (right, forest plot) analyses showing the association of THBS2 expression with OS in patients with pN0-stage LUAD (upper panel) and the entire cohort (lower panel).

Figure S6

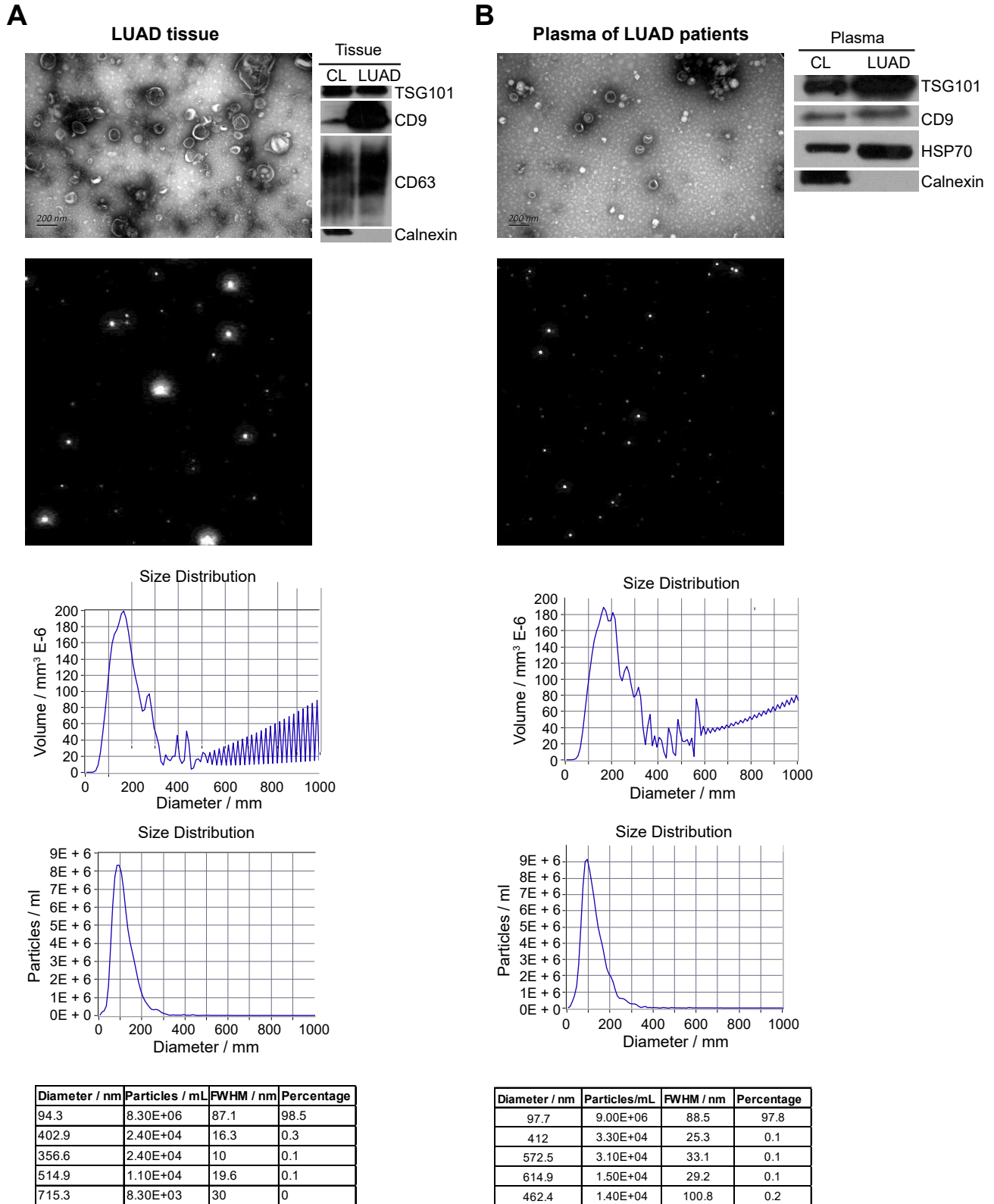


Fig. S6. Exosome isolation, purification and characterization (related to Figure 4)

A, B, Lung tissue- (A) or plasma (B)-derived exosomes were characterized by transmission electron microscopy (morphology; upper left panel in A, B), and Nanoparticle Tracking Analysis (the size; lower panels in A, B), and western blots using three positive exosomal markers (CD9, TSG101, CD63, or HSP70) and one negative marker (Calnexin) (upper right panel in A, B). CL: cell lysate, used as a positive quality control. Scale bar: 200 nm.

Figure S7

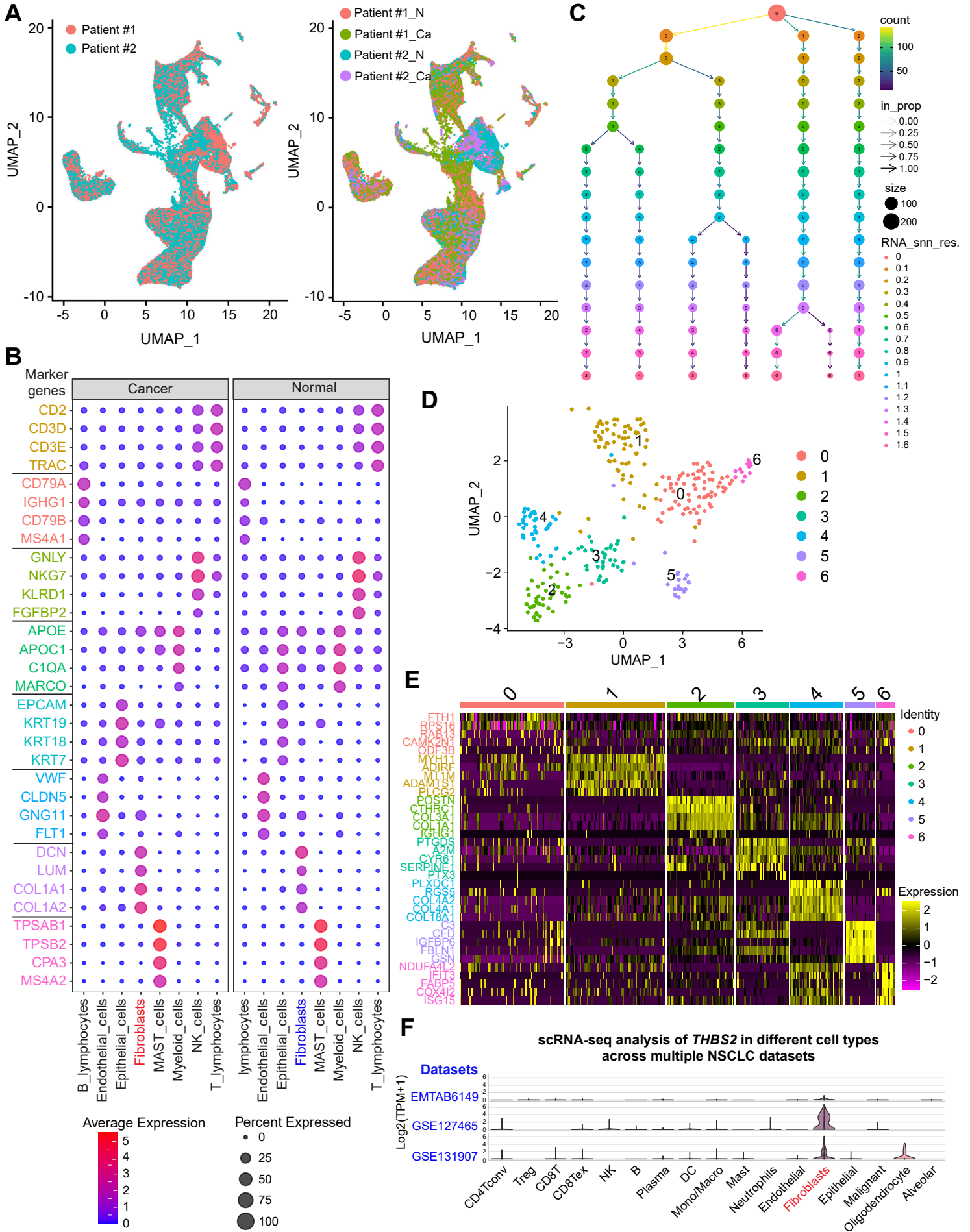


Figure S7. THBS2 defines a unique CAF subset

A, UMAP plots showing patients' origin (left) and samples' origin (right). Related to Figure 5B.

B, Marker genes used to define different cluster identities of whole cell populations. Related to Figure 5B. C, Implementing the Clustree algorithm to determine the optimal cluster separation of CAFs (cancer-associated fibroblasts). R package "Clustree" was used. Related to Figure 5D-G.

D, UMAP plot showing that CAFs could be divided into 7 subclusters, based on C.

E, Heatmap showing the top 5 marker genes of CAFs.

F, Data mining of public NSCLC (non-small cell lung cancer) scRNA-seq data showing THBS2 expression in distinct cell types across different NSCLC datasets. Data were directly obtained from the TISCH portal, a comprehensive web resource enabling interactive single-cell transcriptome visualization of the tumor microenvironment (<http://tisch.comp-genomics.org/home/>). Of note, THBS2 is consistently expressed by CAFs across these datasets, although CAFs were not sub-grouped here.

Figure S8

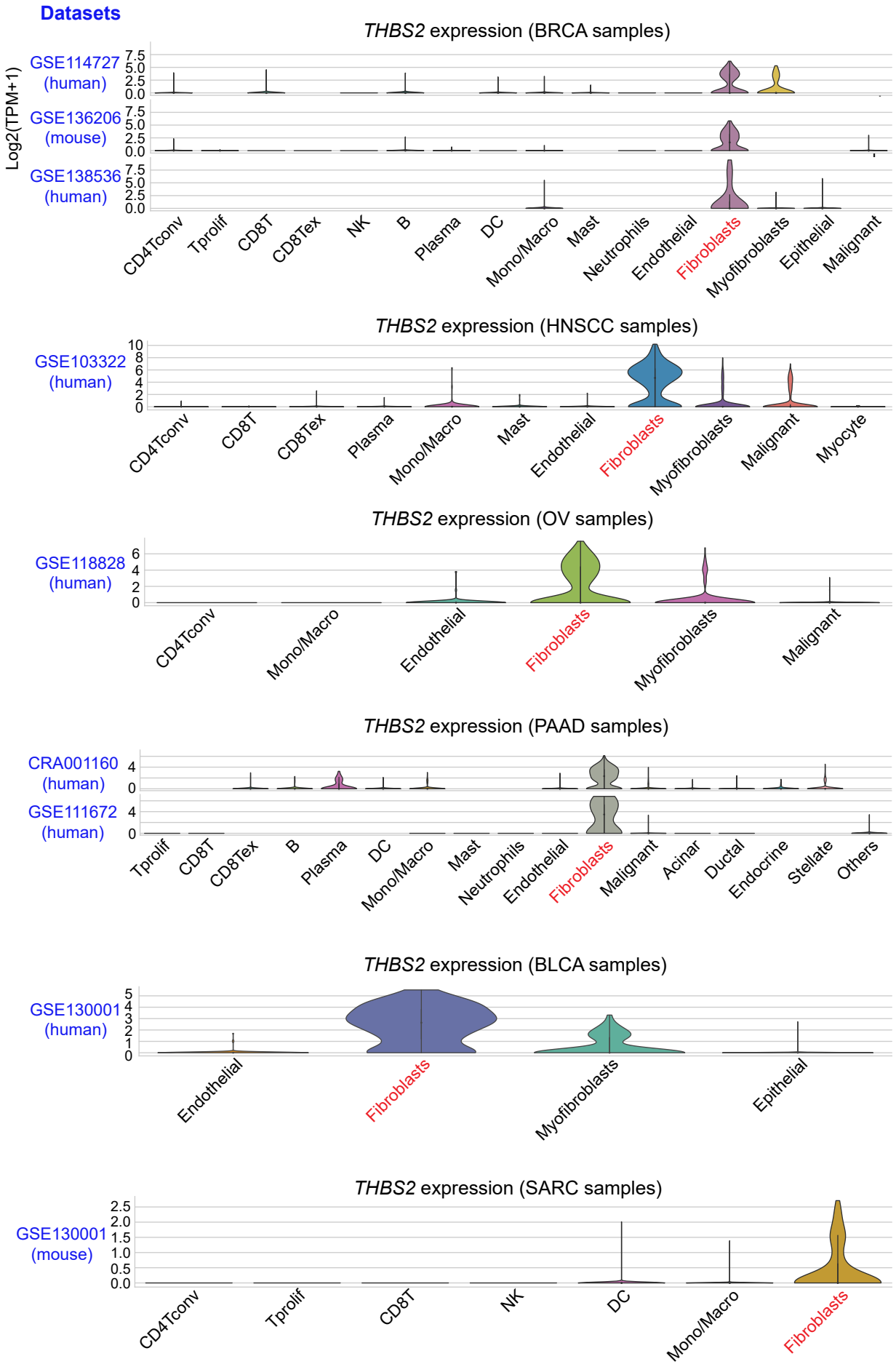
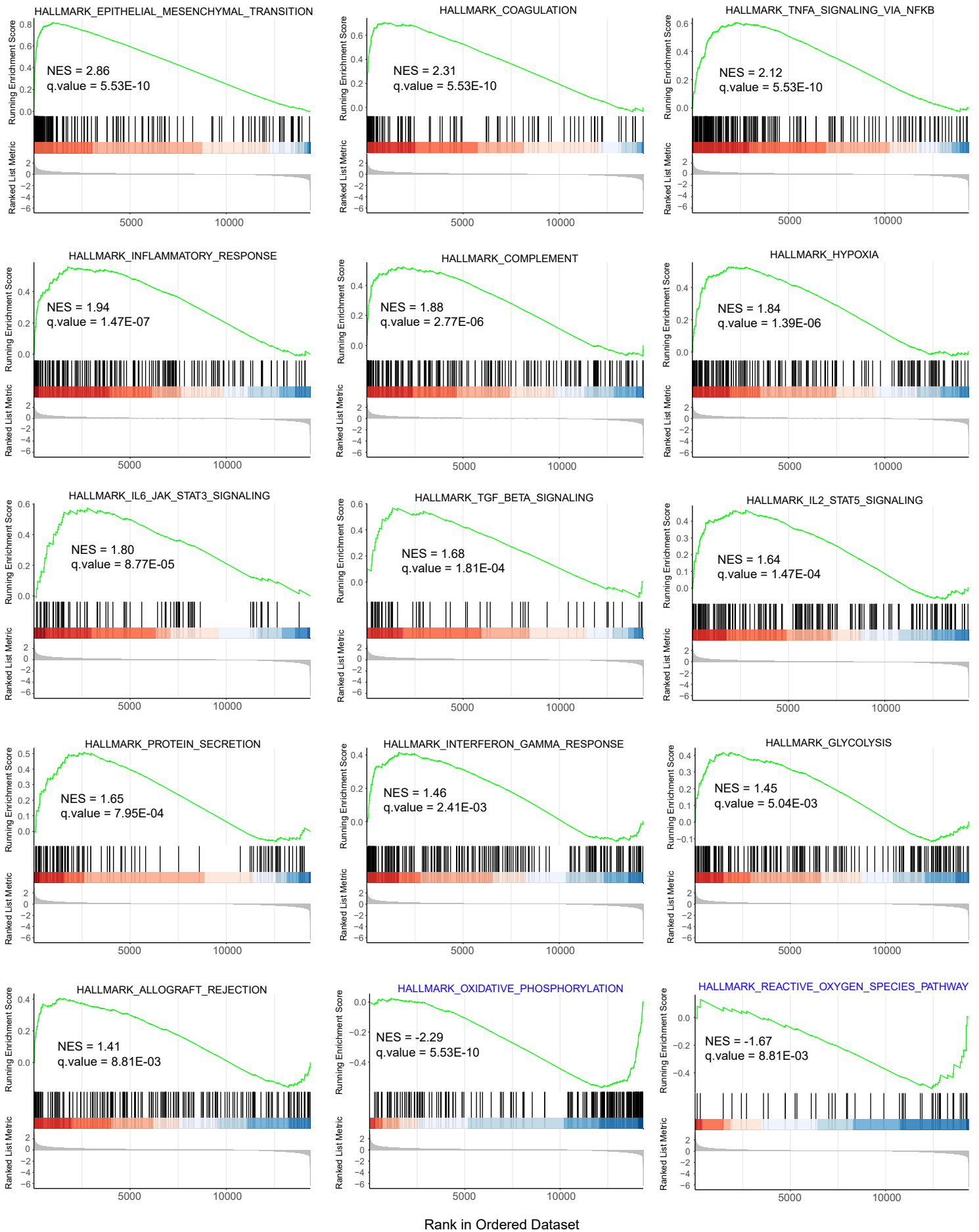


Figure S8. Multiple scRNA-seq analyses reveal THBS2 expression predominantly derived from subsets of CAFs

Data mining of public scRNA-seq data showing THBS2 expression in distinct cellular clusters across several cancer types that are enriched for stromal compartments. Data were obtained from the TISCH portal, a comprehensive web resource enabling interactive single-cell transcriptome visualization of the tumor microenvironment.

Figure S9



Rank in Ordered Dataset

Figure S9. Pathway enrichment analysis of THBS2+ vs. THBS2– CAFs

Gene Set Enrichment Analysis (GSEA) of THBS2+ CAFs compared to THBS2– CAFs based on the scRNA-seq data of this study (related to Figure 5D).

Figure S10

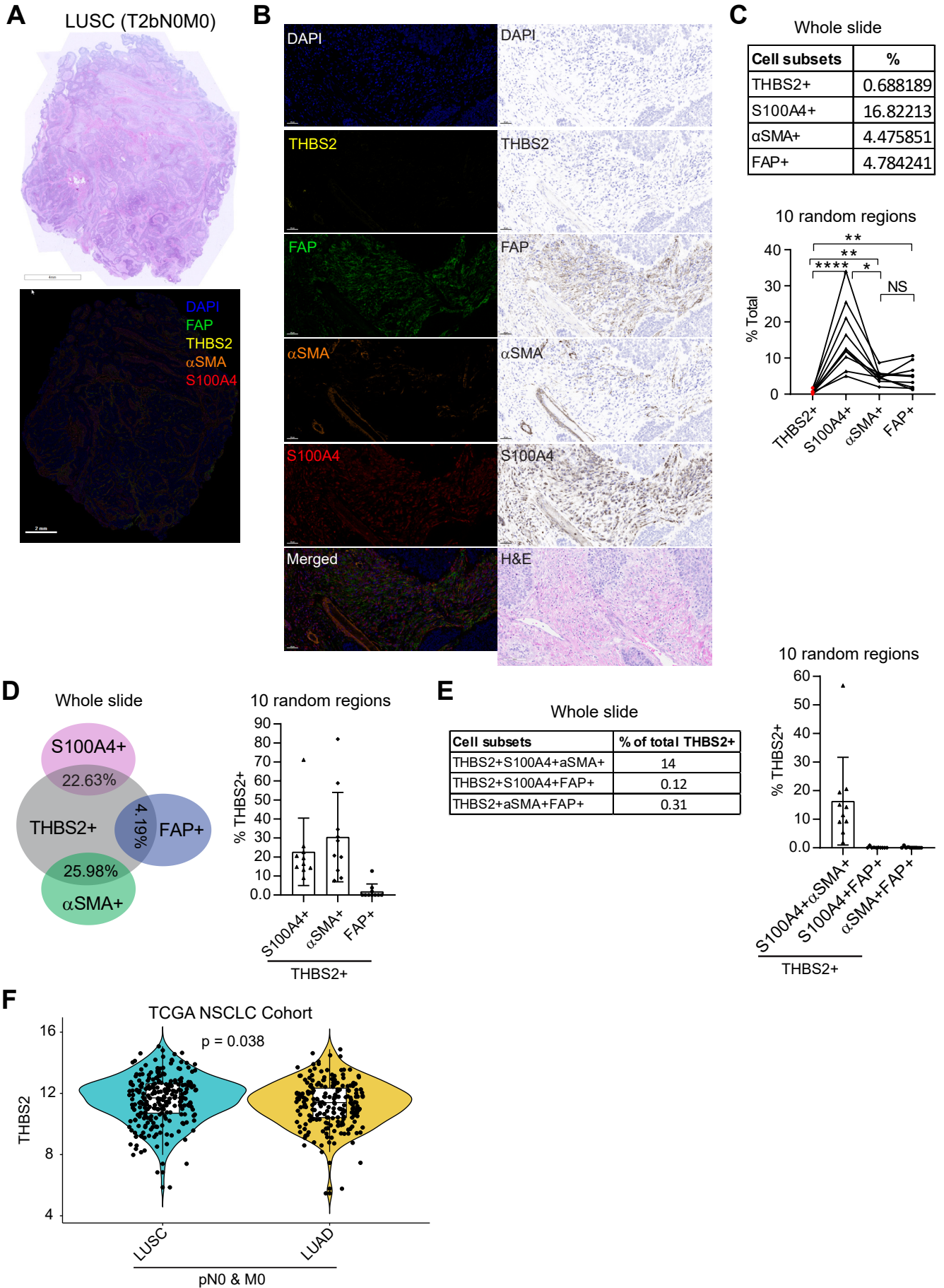


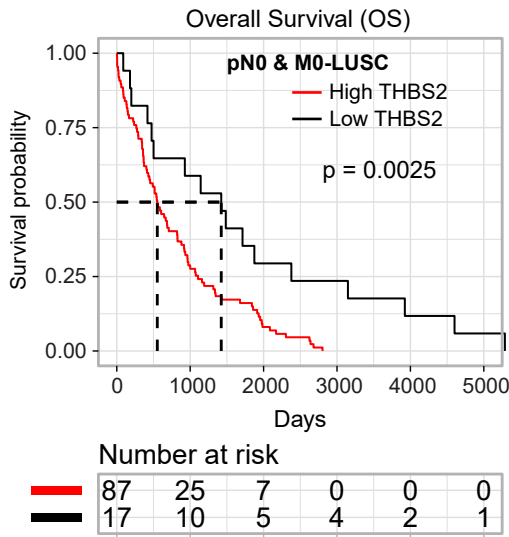
Figure S10. Multiplexed IHC shows the spatial association between THBS2 and CAF-associated markers in LUSC
A, B, THBS2 co-staining with four classical CAF markers (α SMA, FAP, S100A4 [FSP1]) in a LUSC (lung squamous cell carcinoma) (pT2bN0M0) sample. The image acquisition of all markers occurred simultaneously. Panel A shows the whole slide scan; Scale bar: 2 mm.; panel B shows the representative regions (200x); Scale bar: 50 μ m..

C-E, Individual or selected combinations of markers (whole slide and 10 randomly selected regions) were quantified and are shown. * $p < 0.05$; ** $p < 0.01$; **** $p < 0.0001$ by paired ANOVA test.

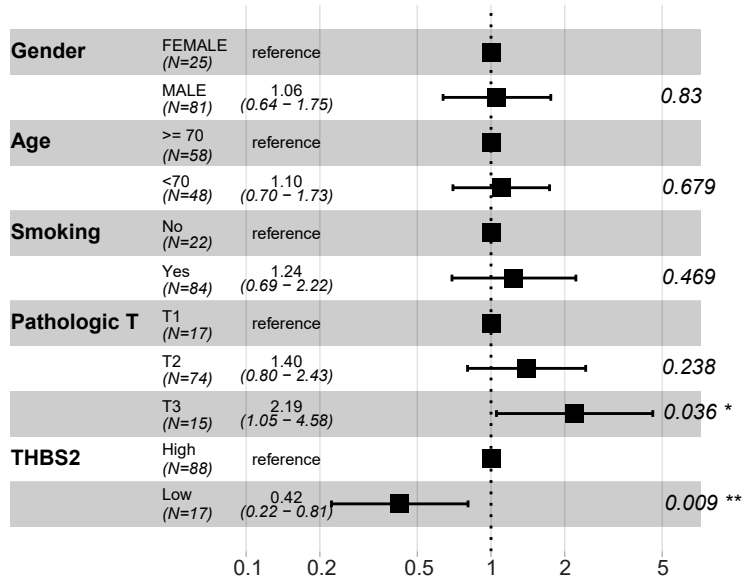
F, Comparative analysis of THBS2 expression between TCGA pN0-stage LUSC and LUAD. P-value was calculated by a two-side student's t-test.

Figure S11

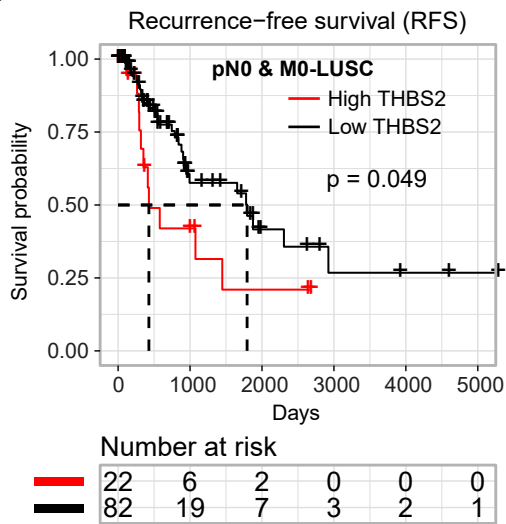
A



Hazard ratio



B



Hazard ratio

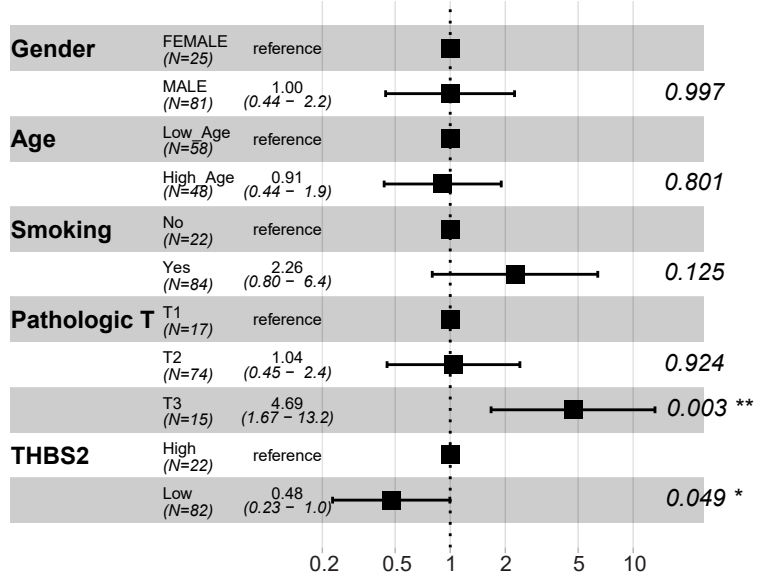


Figure S11. THBS2 expression predicts early pN0-stage LUSC

A, B, Univariate (left) and multivariate (right, forest plot) analyses showing the association of THBS2 expression with OS (A) and RFS (B) in patients with pN0M0-stage LUSC (lung squamous cell carcinoma) after surgery.

Figure S12

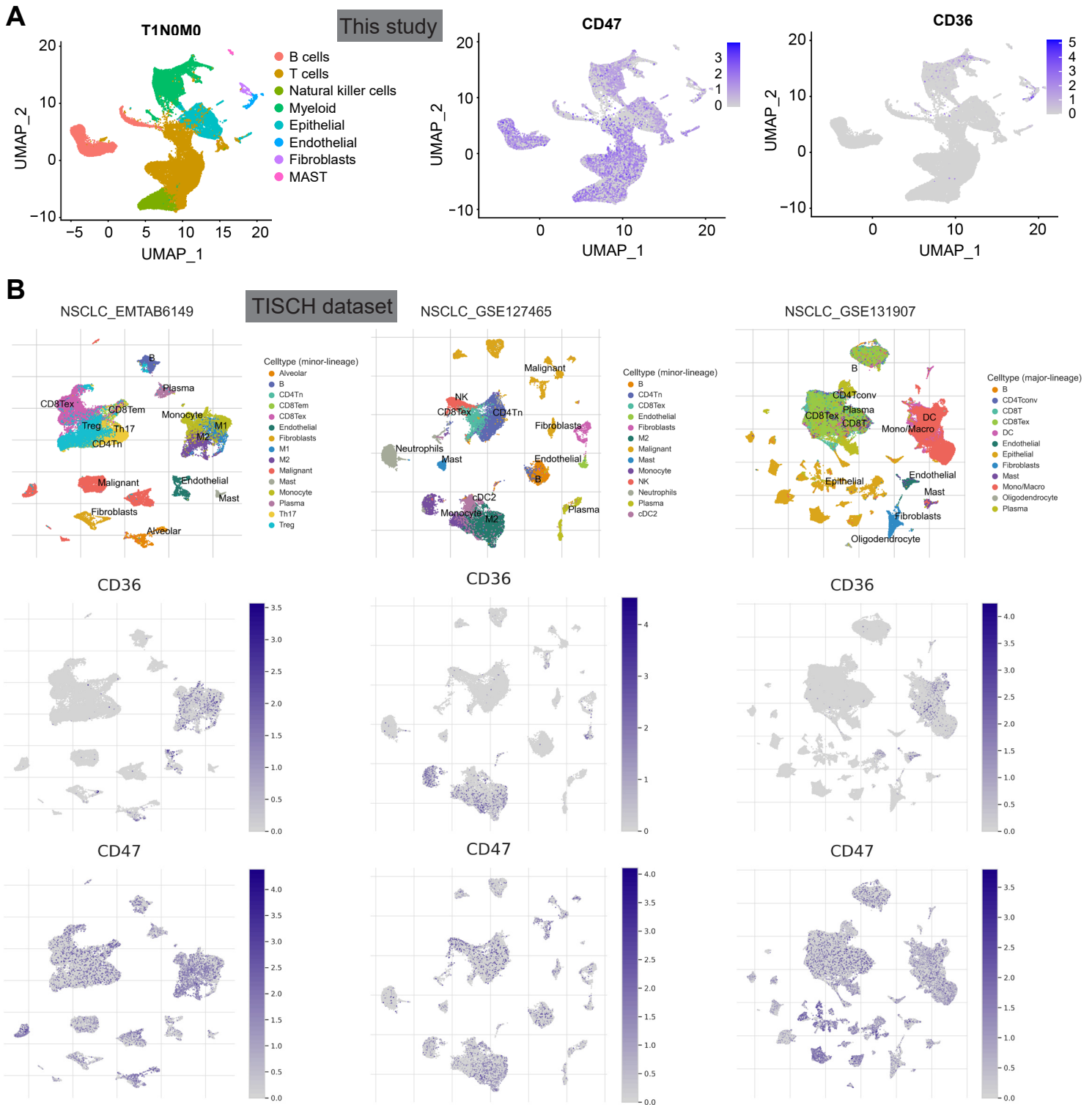


Figure S12. Expression of CD36 and CD47 across cellular subsets of lung tumors

A, B, scRNA-seq showing the gene expression of THBS2 in different cell clusters of lung tumor from this study (LUAD, pT1N0M0; panel A) and public datasets (TISCH portal; panel B).

Figure S13

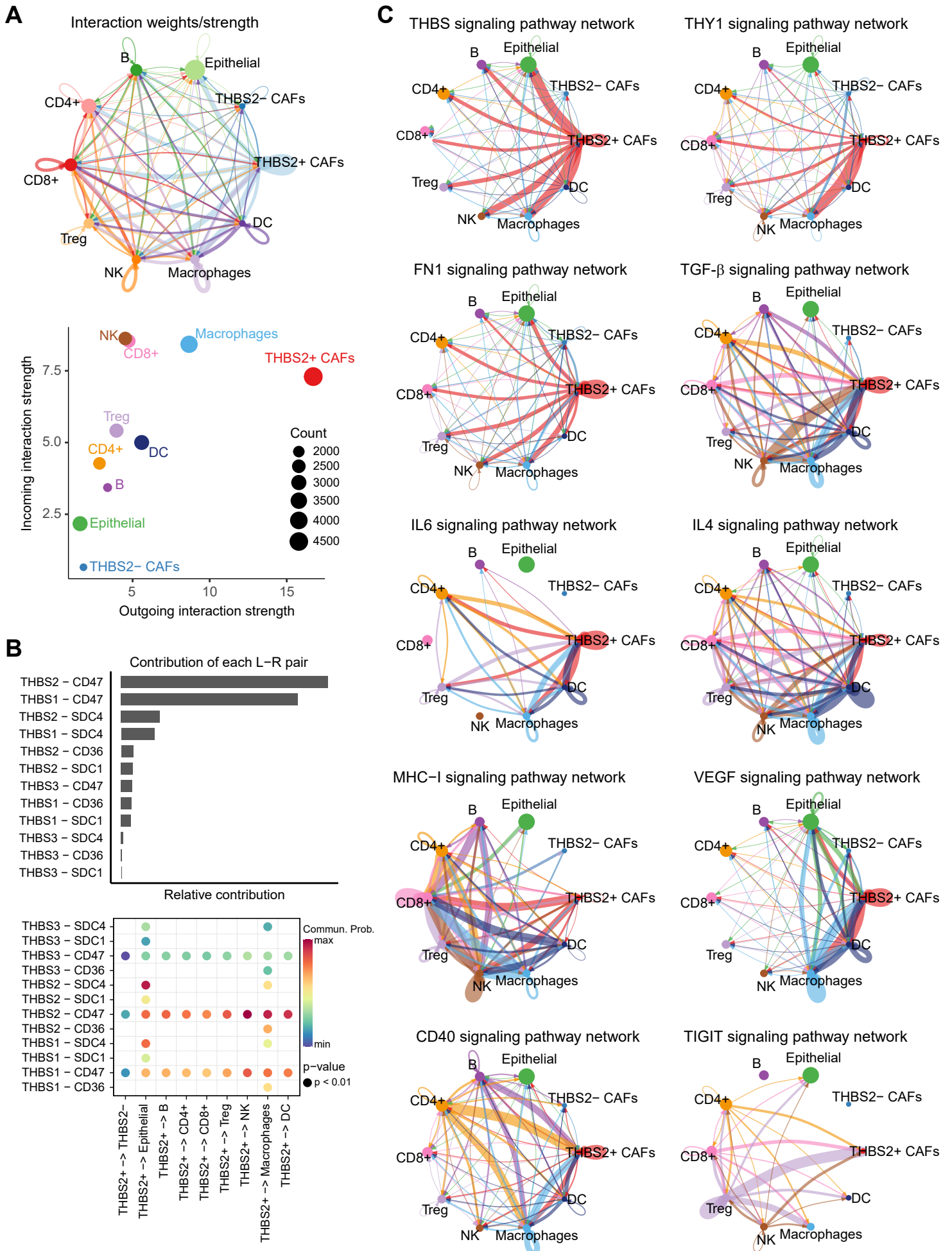


Figure S13. Intercellular communication analysis of THBS2+ CAFs and other cells in early-stage lung cancer

A, Circle plot (top) showing global communications among the major cell types within the microecosystem of lung cancer (pT1N0M0). The edges are directional and the thickness of them corresponds to the signal strength of interactions, colored according to cell type. The size of the node is related to the number of cells in each cell type. Dot plot (bottom) showing the global incoming and outgoing interaction strength of each cell type, indicating the dominant senders and receivers. Data were based on our scRNA-seq of two pT1N0M0 lung cancer samples (related to Figure 4) using CellChat package in R software. Of note, THBS2+ CAFs have more profound interaction weights with both tumor and immune cells compared with THBS2- CAFs.

B, Bar plot (upper panel) showing the relative contribution of each ligand-receptor pair to the THBS signaling pathway. Dot plot (lower panel) showing the potential interaction weights (ligand-receptor pairs) from THBS+ cancer-associated fibroblasts to other cell types.

C, Top enriched intercellular communication networks constructed based on the interactions among ligands, receptors and their cofactors.

Figure S14

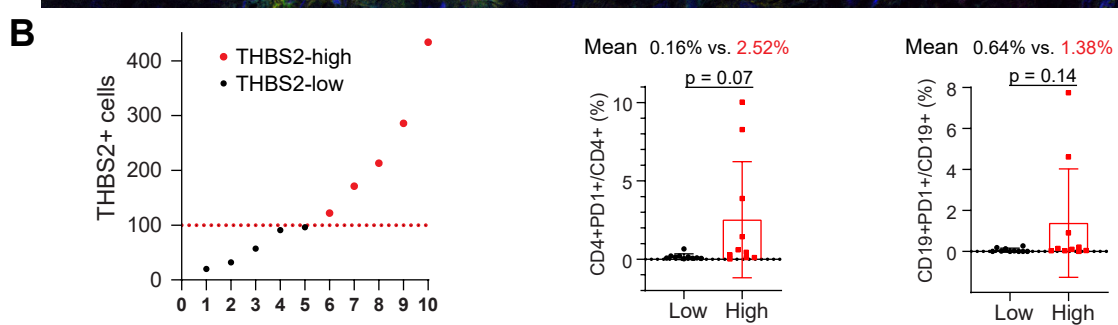
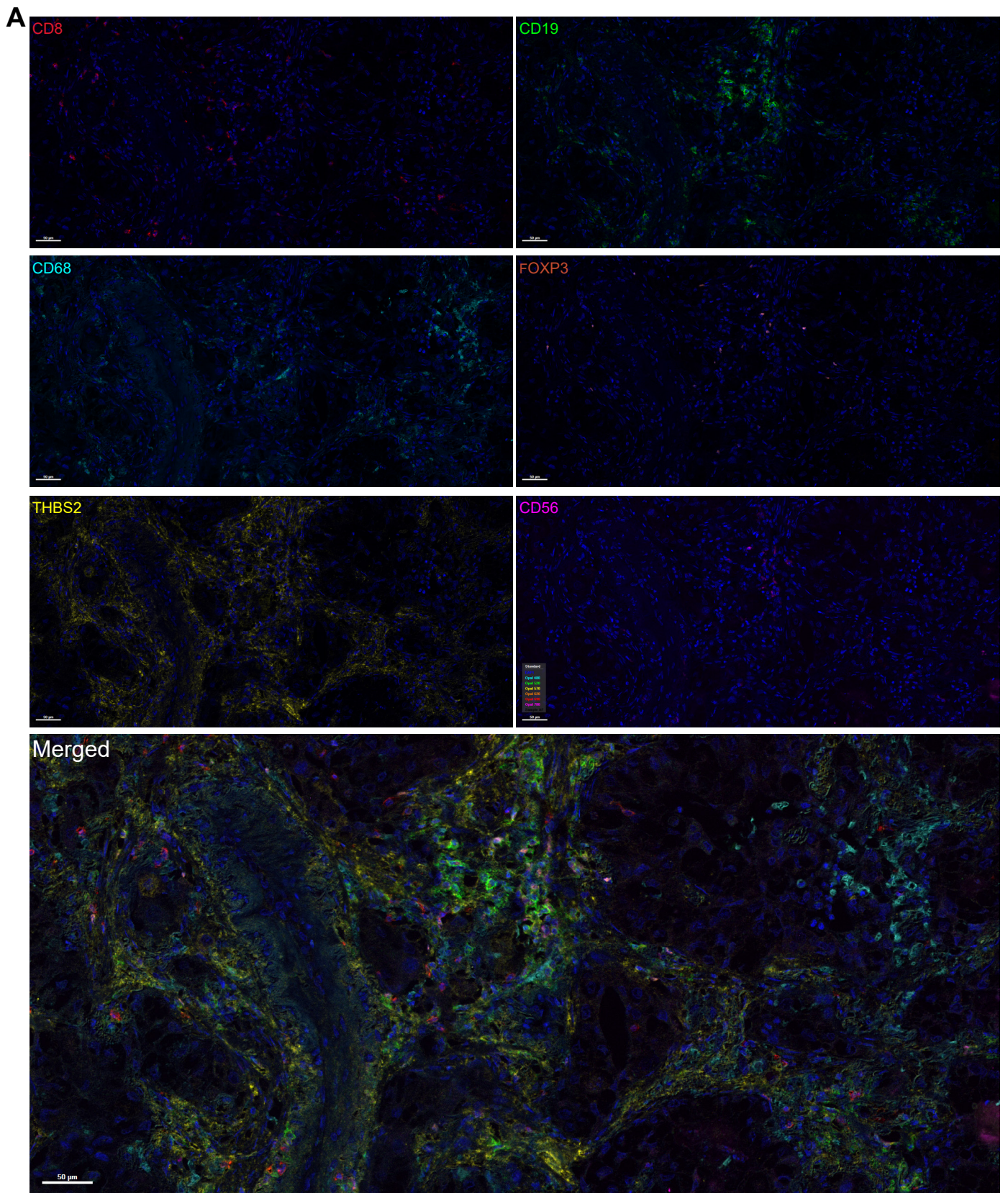


Figure S14. High THBS2 expression is associated with elevated exhaustion markers in immune infiltrates

A, THBS2 co-staining with the indicated immune markers in a resected LUAD (lung adenocarcinoma) (pT2bN0M0) sample. The image acquisition of all markers occurred simultaneously with the representative regions were shown (200x). Scale bar: 50 μ m.

B, Ratio of CD4+PD1+ to total CD4+ T cells and CD19+PD1+ to total CD19+ T cells in 10 randomly (5 tumoral and 5 stromal regions) selected regions of a LUAD tumor related to Figure 8. p-value was calculated using a two-side student's t-test.

Figure S15

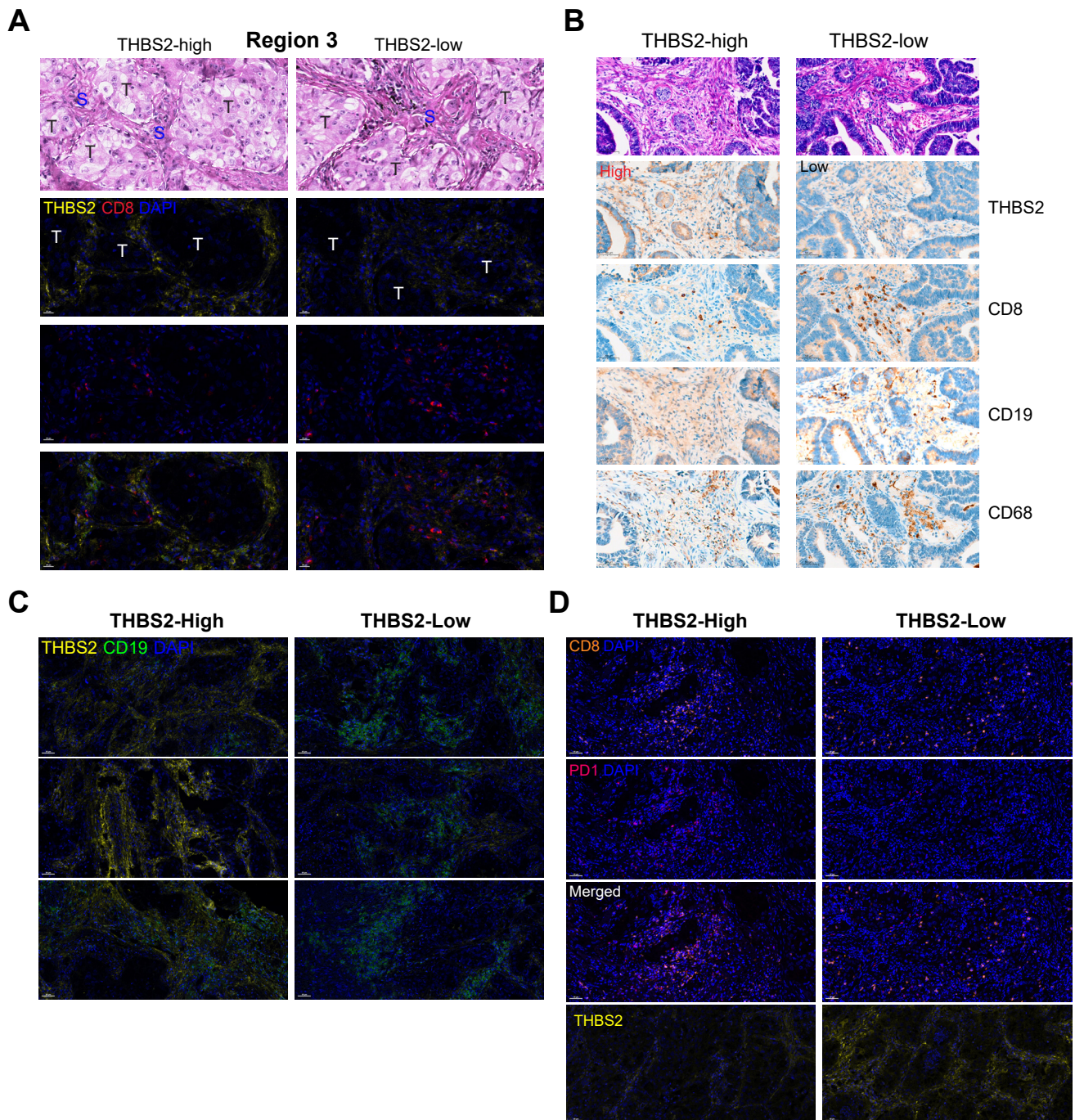


Figure S15. High THBS2 expression is associated with dysregulated immune cell infiltrates.

A, Representative images showing the THBS2 co-staining with CD8 T lymphocytes in another regions of one LUAD sample (related to Figure 8A, B). Scale bar: 50 μ m.

B, Representative figures showing the difference in immune cell infiltrates between THBS2-high and -low groups (related to Figure 8C). Scale bar: 50 μ m.

C, THBS2 co-staining with the CD19 B lymphocytes in a resected LUAD (lung adenocarcinoma) (pT2bN0M0) sample. The image acquisition of all markers occurred simultaneously with the representative regions were shown (200x; panel A, B). Scale bar: 50 μ m.

D, CD8, PD1 expression between between THBS2-high and -low groups (related to Figure 7H). Scale bar: 50 μ m.

Figure S16

NSCLC patients with anti-PD1/PD-L1 (GSE135222)

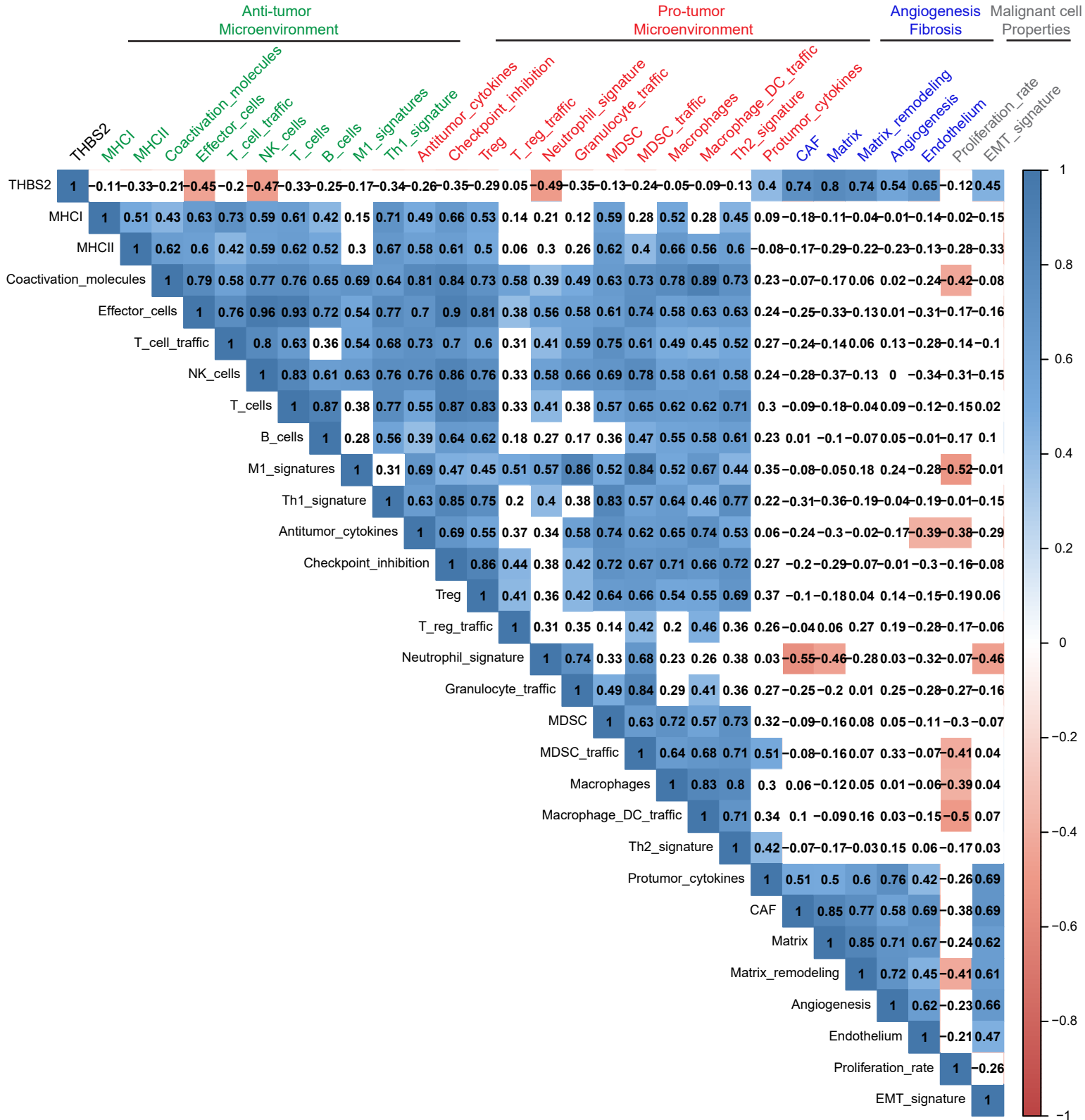


Figure S16. THBS2 is associated with an aberrant tumor immune microenvironment

The correlation matrix showing the correlation between THBS2 expression and the tumor microenvironment signature across NSCLC patients treated with immunotherapy. Data were downloaded and reanalyzed from Bagaev A et al. Cancer Cell. 2021.

Figure S17

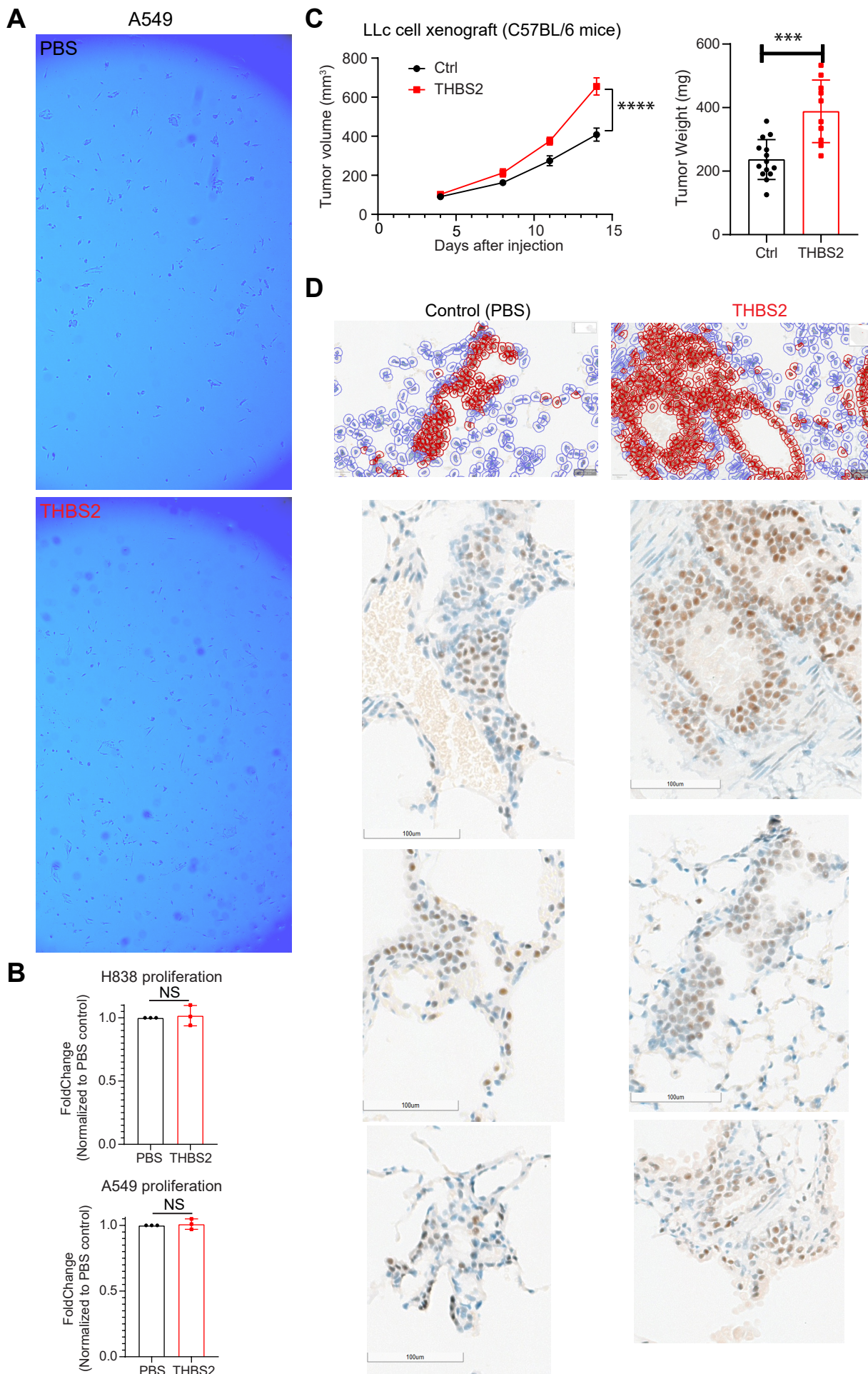


Figure S17. THBS2 promoted tumor micro-metastasis in A549 LUAD xenografts

A, Representative images (10x) of A549 transwell assays were shown.

B, The effect of THBS2 recombinant protein on the proliferation of LUAD cells (upper: H838; lower: A549). NS: not significant, by two-sided student's t-test.

C, *In vivo* LUAD (mouse LLC cell line) xenografts showing the effect of THBS2 on subcutaneous tumor growth in immune-competent C57BL/6 mouse models. **** $p < 0.0001$ by two-way ANOVA test (tumor volume). *** $p < 0.001$ by two-sided student's t-test.

D, Immunohistochemistry-based quantifications (related to the lower panel in Figure 7C); Qupath software (see the methods) was used to identify and quantify the KU80 positive cells. In B, another three representative IHC images showing the micro-metastasis to the lung in subcutaneous A549 LUAD xenografts. Human specific KU80 antibodies were used to detect human-derived A549 LUAD cells. Scale bar: 100 μm .