Supplementary data

Integrative single-cell and multi-omics analysis identifies an SLC2A1-associated malignant epithelial cell state predicting prognosis and immunotherapy response in lung adenocarcinoma

Linqian Song1,2, Shiyun Xing1,2, Peijie Li1,2, Yunliang Cao1,2, Hu Ma1,2,*

1Department of Oncology, The Second Affiliated Hospital of Zunyi Medical University, Zunyi, Guizhou 563000, P. R. China;

2Key Laboratory for Cancer Prevention and Treatment of Guizhou Province, Zunyi, Guizhou 563000, P. R. China.

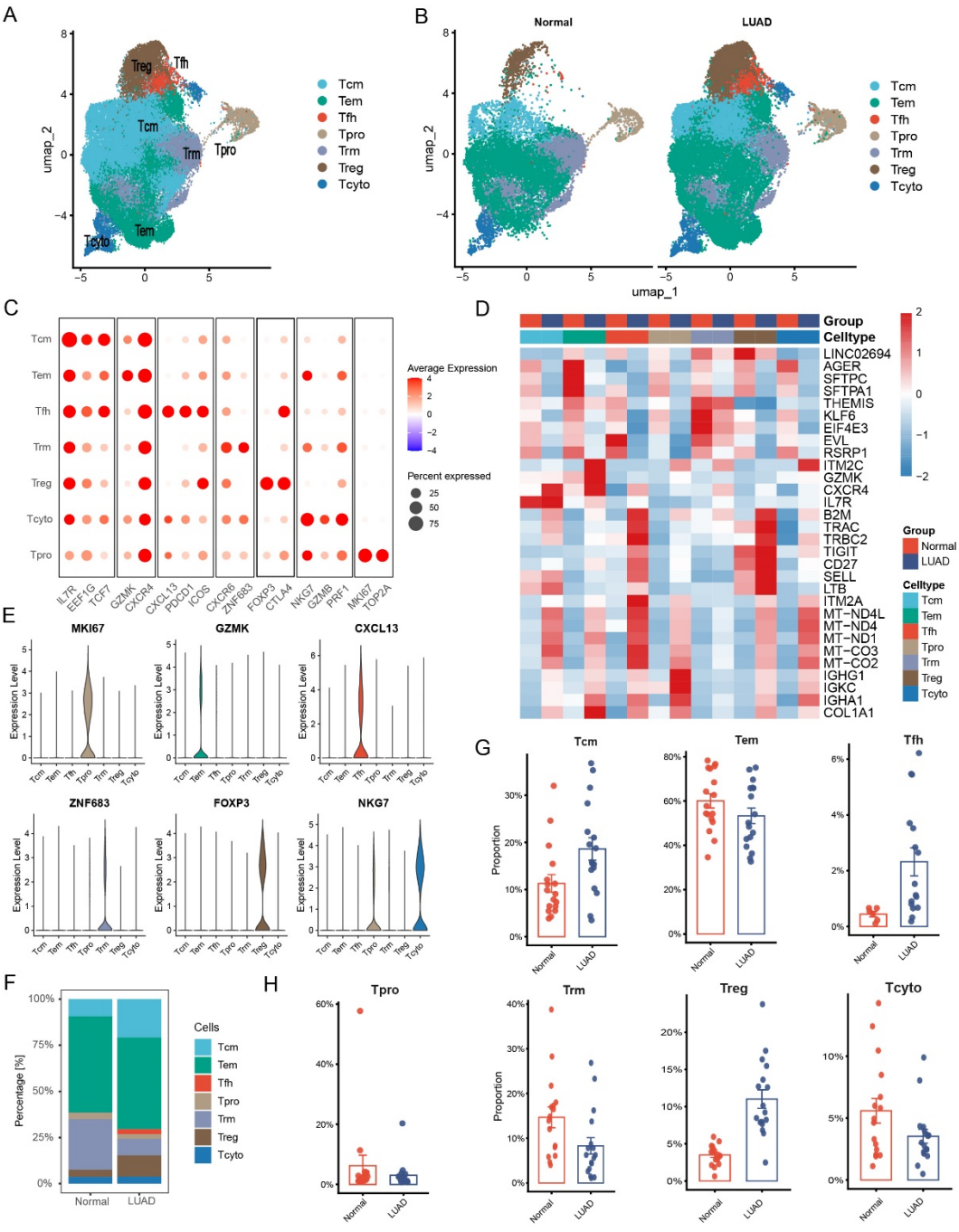
* Corresponding author

Correspondence to: Hu Ma, Email: mahuab@163.com; ORCID: 0000-0003-1654-1576.

Supplemental Tabels and Figures

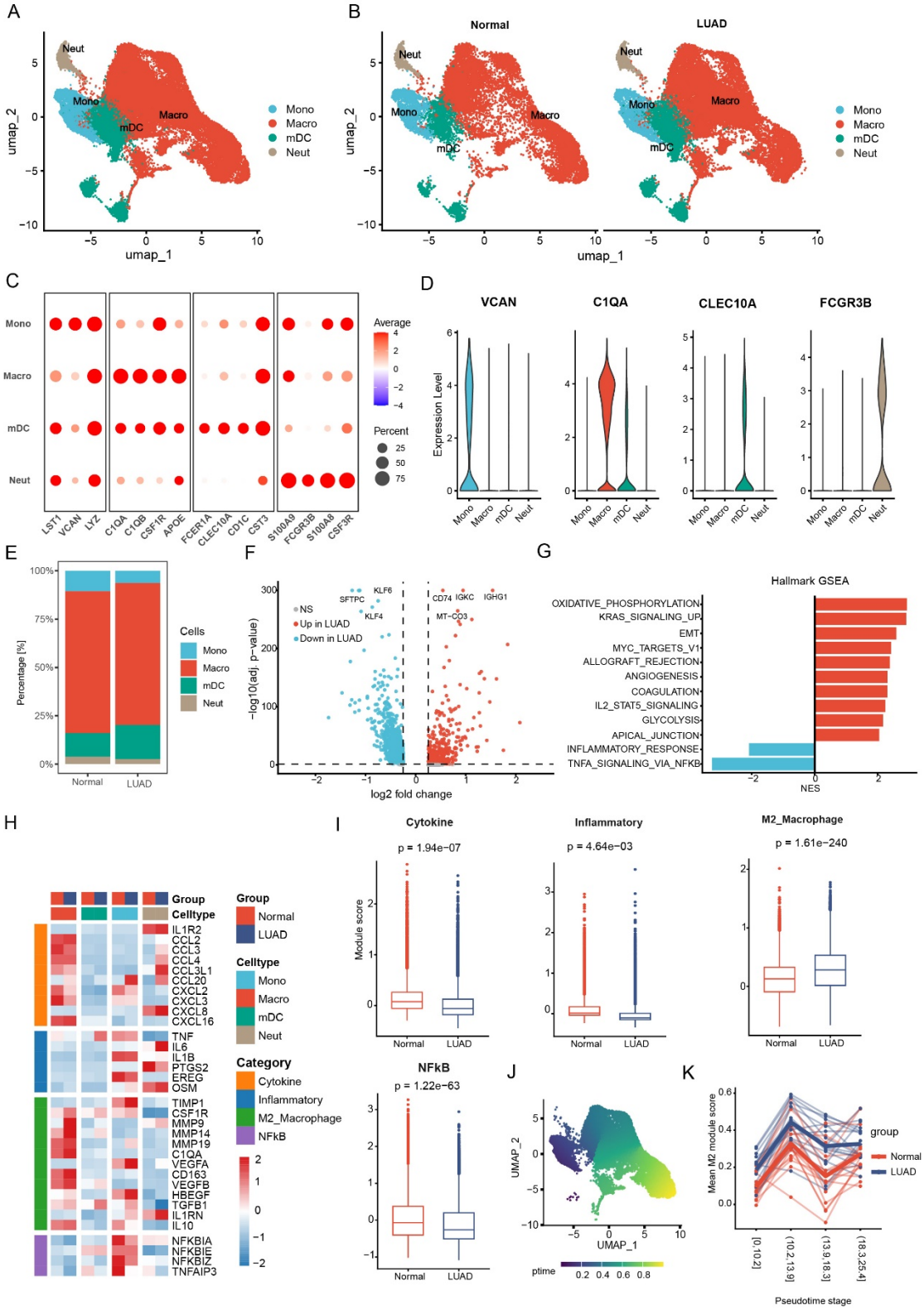
List of Supplementary Figures

List of Supplementary Tables



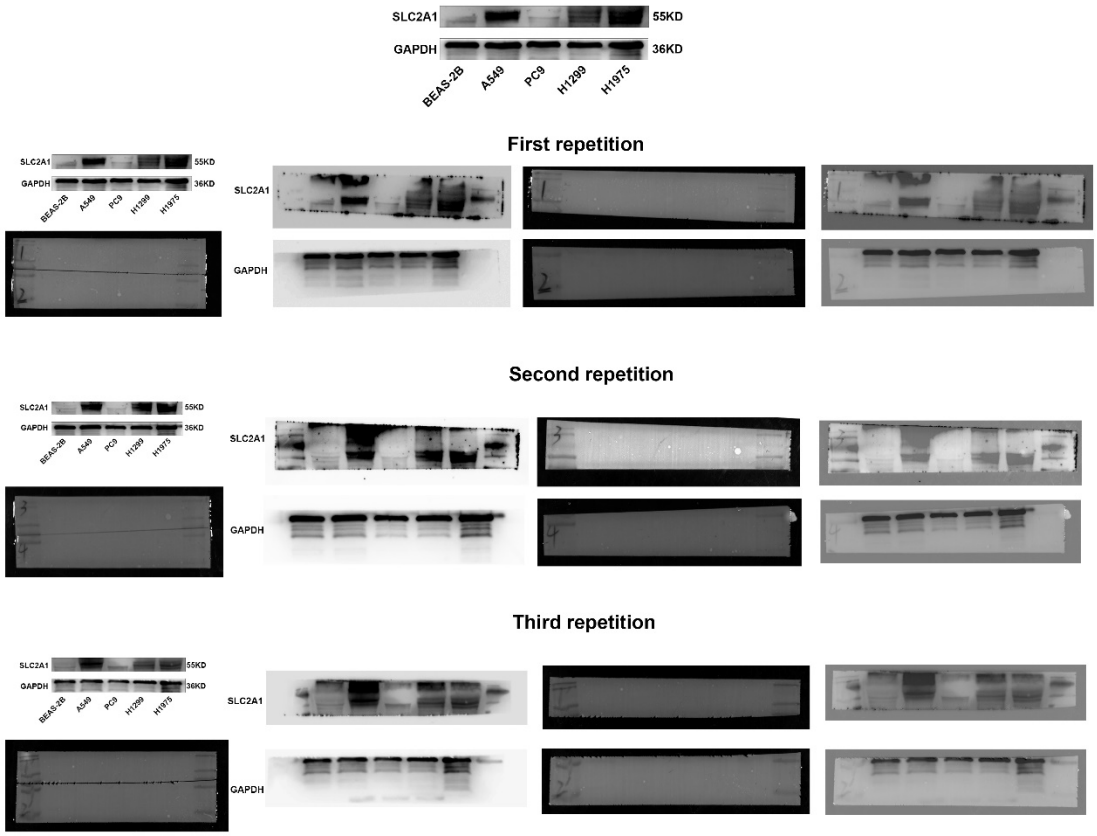
Supplementary Figure S1. Single-cell transcriptomic features of T cell subpopulations. (A) UMAP visualization of T cell subpopulations. (B) UMAP distribution of T cells in normal lung tissues (Normal) and lung adenocarcinoma tissues (LUAD). (C) Dot plot showing the expression of representative marker genes for each T cell subpopulation. (D) Heatmap showing gene expression features of different T cell subpopulations in the Normal and LUAD groups. (E) Violin plots showing the expression levels of six representative genes across T cell subpopulations. (F) Distribution of the proportions of T cell subpopulations in the Normal and LUAD groups. (G) The proportions of Tcm, Tem, and Tfh cells between the Normal and LUAD groups. (H) Comparison of the proportions of Tpro, Trm, Treg, and Tcyto cells between the

Normal and LUAD groups.



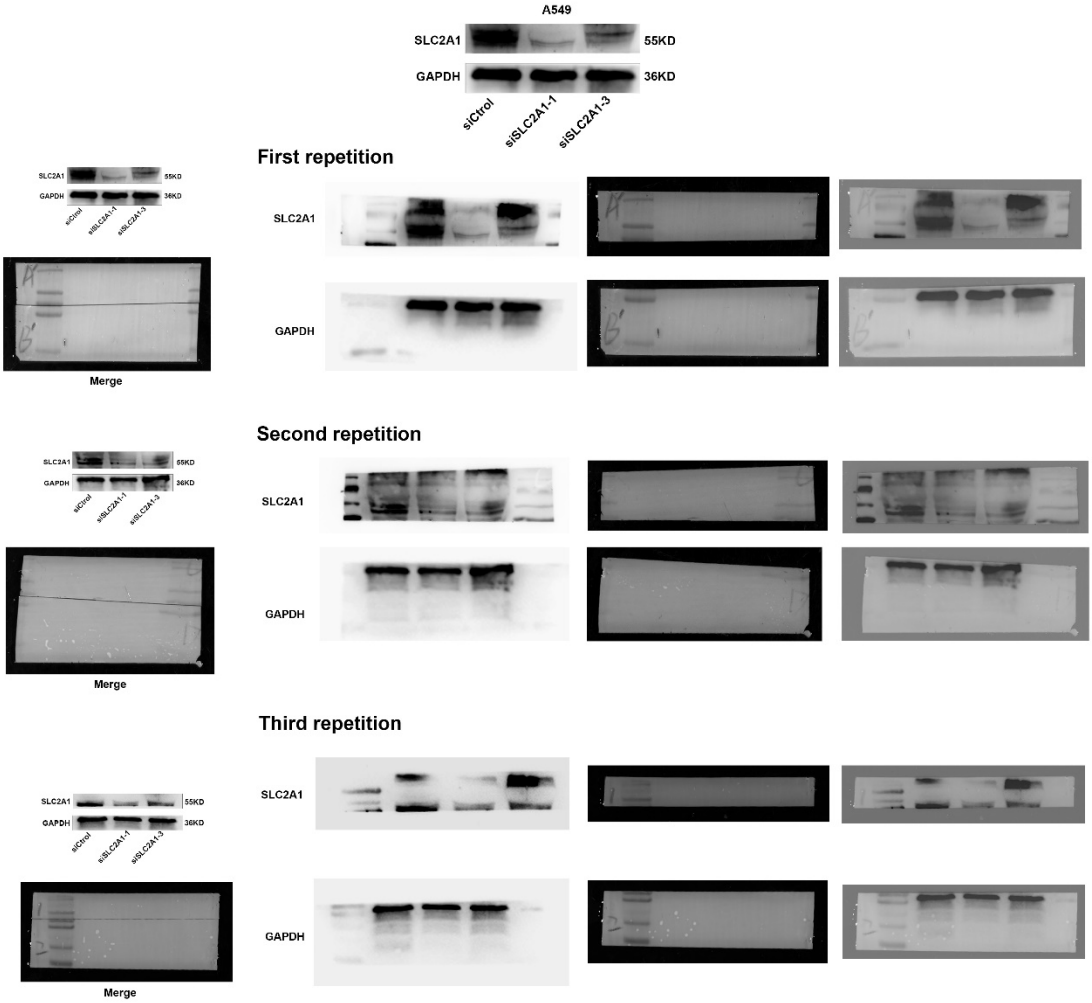
Supplementary Figure S2. Landscape features of myeloid cells. (A) UMAP visualization of myeloid cell subpopulations, including monocytes (Mono), macrophages (Macro), myeloid dendritic cells (mDC), and neutrophils (Neut). (B) UMAP distribution of myeloid cells in

normal lung tissues (Normal) and lung adenocarcinoma tissues (LUAD). (C) Dot plot showing the expression of representative marker genes for each myeloid cell subpopulation. (D) Violin plots showing the expression levels of representative marker genes across different myeloid cell subpopulations. (E) Proportions of myeloid cell subpopulations in the Normal and LUAD groups. (F) Volcano plot showing differentially expressed genes in myeloid cells between the Normal and LUAD groups. (G) Gene set enrichment analysis (GSEA) based on the Hallmark gene sets. (H) Heatmap showing the expression of module genes related to cytokine signaling, inflammatory response, M2 macrophages, and NF-κB signaling across different myeloid cell subpopulations in the Normal and LUAD groups. (I) Comparison of the scores of the four functional modules between the Normal and LUAD groups. (J) UMAP visualization of pseudotime analysis of myeloid cells. (K) Changes in the M2 macrophage module score across different pseudotime stages in the Normal and LUAD groups.



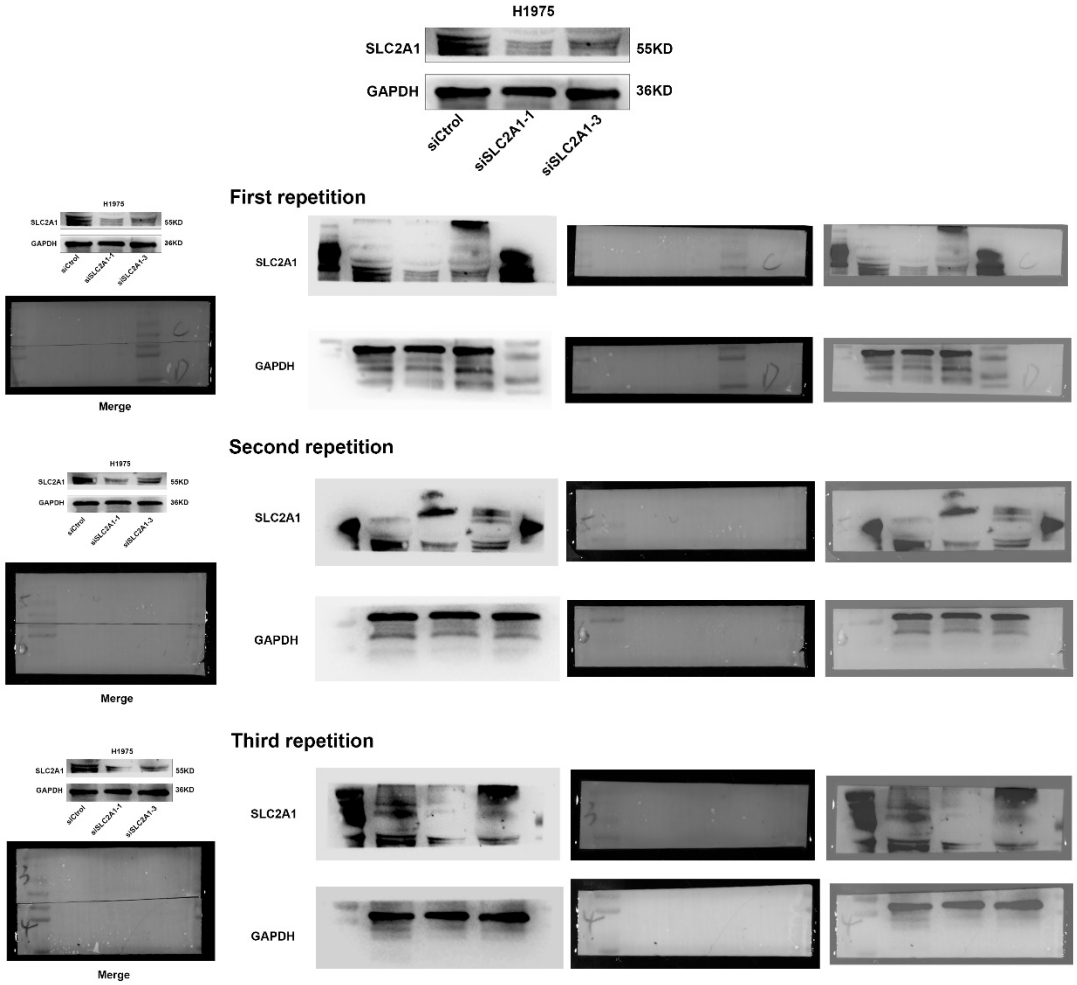
Supplementary Figure S3. Original uncropped Western blot images of SLC2A1 expression across lung cell lines.

Uncropped and full-length Western blot images corresponding to Figure 10B are shown. Protein expression of SLC2A1 (55 kDa) and GAPDH (36 kDa) was detected in normal bronchial epithelial cells (BEAS-2B) and lung adenocarcinoma cell lines (A549, PC9, H1299, and H1975). Three independent experimental repetitions are presented.



Supplementary Figure S4. Original uncropped Western blot images of SLC2A1 knockdown in A549 cells.

Uncropped and full-length Western blot images corresponding to Figure 10D are shown. A549 cells were transfected with control siRNA (siCtrl) or siRNAs targeting SLC2A1 (siSLC2A1-1 and siSLC2A1-3). Protein levels of SLC2A1 (55 kDa) and GAPDH (36 kDa) were detected. Three independent experimental repetitions are presented.



Supplementary Figure S5. Original uncropped Western blot images of SLC2A1 knockdown in H1975 cells.

Uncropped and full-length Western blot images corresponding to Figure 10D are shown. H1975 cells were transfected with control siRNA (siCtrl) or siRNAs targeting SLC2A1 (siSLC2A1-1 and siSLC2A1-3). Protein levels of SLC2A1 (55 kDa) and GAPDH (36 kDa) were detected. Three independent experimental repetitions are presented.

Supplementary Table S1. Canonical marker genes used for cell-type annotation in single-cell RNA sequencing analysis.

cellMarker	Marker Genes
T_cells	CD3E, TRAC, CD3D
B_cells	MS4A1, CD79A, CD19
Plasma	MZB1, JCHAIN, IGHA1
Endothelial	PECAM1, VWF
Fibroblast	COL1A1, COL1A2, DCN

Epithelial	EPCAM, KRT7, KRT8
Myeloid	LYZ, CD14, SPI1
Tcm(T_cells)	IL7R, EEF1G, TCF7
Tem(T_cells)	GZMK, CXCR4
Tfh(T_cells)	CXCL13, PDCD1, ICOS
Trm(T_cells)	CXCR6, ZNF683
Treg(T_cells)	FOXP3, CTLA4
Tcyto(T_cells)	NKG7, GZMB, PRF1
Tpro(T_cells)	MKI67, TOP2A
Mono(Myeloid)	LST1, VCAN, LYZ
Macro(Myeloid)	C1QA, C1QB, CSF1R, APOE
mDC(Myeloid)	FCER1A, CLEC10A, CD1C, CST3
Neut(Myeloid)	S100A9, FCGR3B, S100A8, CSF3R

Supplementary Table S1. Canonical marker genes used for cell-type annotation in single-cell RNA sequencing analysis.

This table lists the representative marker genes used to define major cell types, including T cells, B cells, plasma cells, myeloid cells, fibroblasts, endothelial cells, and epithelial cells, as well as subtype-specific markers for T-cell and myeloid cell subpopulations.

Supplementary Table S2. Sequences of siRNAs and primers used in this study.

	Forward (5'-3')	**Reverse (5'-3')**
	si-RNA	
NC-SLC2A1	UUCUCCGAACGUGUCACGUTT	ACGUGACACGUUCGGAGAATT
Si1-SLC2A1	CCTGCAGTTTGGCTACAACA	GTGGACCCATGTCTGGTTG
Si2-SLC2A1	GCUCUGUUGAUGAAGAUUATT	UAAUCUUCAUCAACAGAGCTT
Si3-SLC2A1	CCUUGUGCUUCAACAUCAATT	UUGAUGUUGAAGCACAAGGTT
	RT-PCR	
SLC2A1	TGGCATCAACGCTGTCTTCT	AACAGCGACACGACAGTGAA
GAPDH	GAAGGTGAAGGTCGGAGTC	GAAGATGGTGATGGGATTTC

Supplementary Table S2. Sequences of siRNAs and primers used in this study.

This table provides the sequences of small interfering RNAs (siRNAs) targeting SLC2A1 and the corresponding negative control, as well as primer sequences used for RT–qPCR analysis of SLC2A1 and GAPDH.

Supplementary Table S3. Summary of datasets used in this study.

Category	Dataset	Platform/Database	Data Source	Data Availability
Single-cell RNA sequencing	GSE308103	GEO	Gene Expression Omnibus (NCBI)	Publicly available
External validation cohort	GSE31210	GEO	GEO	Publicly available
External validation cohort	GSE50081	GEO	GEO	Publicly available
External validation cohort	GSE72094	GEO	GEO	Publicly available
Immunotherapy cohort	GSE135222	GEO	GEO	Publicly available
Bulk RNA-seq and clinical data	TCGA-LUAD	TCGA	The Cancer Genome Atlas	Publicly available (via data portal)
Spatial transcriptomics	E-MTAB-13530	ArrayExpress	EMBL-EBI ArrayExpress	Publicly available
Immunotherapy dataset	POPLAR	Roche / Data provider	Clinical trial data platform	Controlled access
Immunotherapy dataset	OAK	Roche / Data provider	Clinical trial data platform	Controlled access

Supplementary Table S3. Summary of datasets used in this study.

This table summarizes all datasets included in the analysis, including single-cell RNA sequencing data, bulk RNA-seq cohorts, spatial transcriptomics data, and immunotherapy-treated cohorts. Information on dataset accession numbers, platforms, data sources, and availability is provided.