

List of Supplemental Materials

Figure S1. Fluorescence-activated cell sorting (FACS) of CSF-CTCs from LUAD-LM patient samples.

CTCs lack CD45 (APC-A) expression and its FSC-H is larger than normal leukocytes. Calcein Blue AM (UltraViolet-A excitation) was used to select live cells.

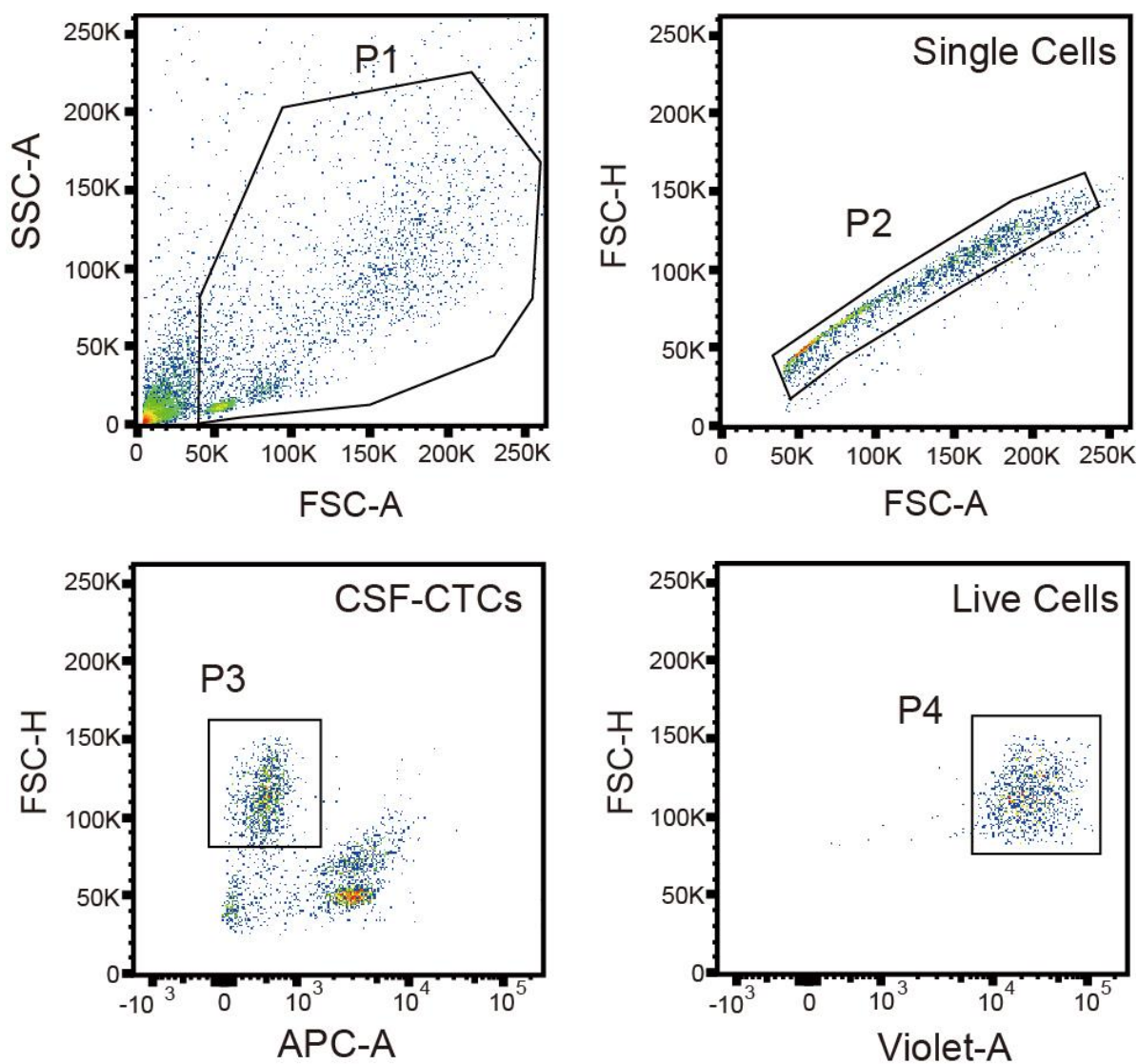


Figure S2. Single-cell expression profile of specific marker genes in control samples on the t-SNE feature plots.

Clustering plot is the same as in Figure 1A. Scaled expression was depicted using a red gradient. Grey denotes lack of expression. Leukocyte marker: *PTPRC*; T cell markers: *CD2*, *CD3D*, *CD3E*, *CD3G*; B cell markers: *CD19*, *MS4A1*, *CD17A*, *CD79B*, Monocyte markers: *CD14*, *CD68* and *CD163*.

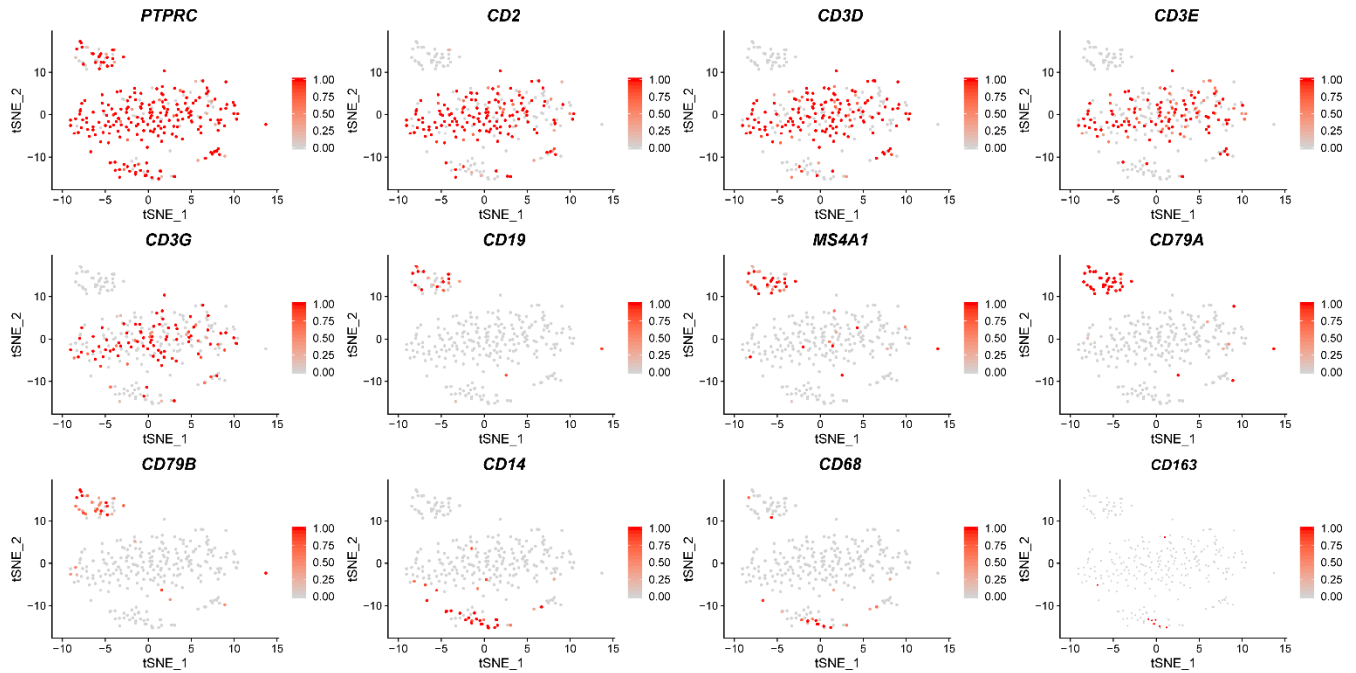


Figure S3. t-SNE plot of LUAD-LM patients CSF-CTCs single-cell transcriptome profiles.

(A). t-SNE plots of CSF-CTC clustering of five LUAD-LM CSF samples with a minimum number of covered genes of 1,000.

(B). Three covered genes number groups (1000~1999, 2000~2999, and 3000~3999) are highlighted for each cell in the t-SNE plot in Figure S3A.

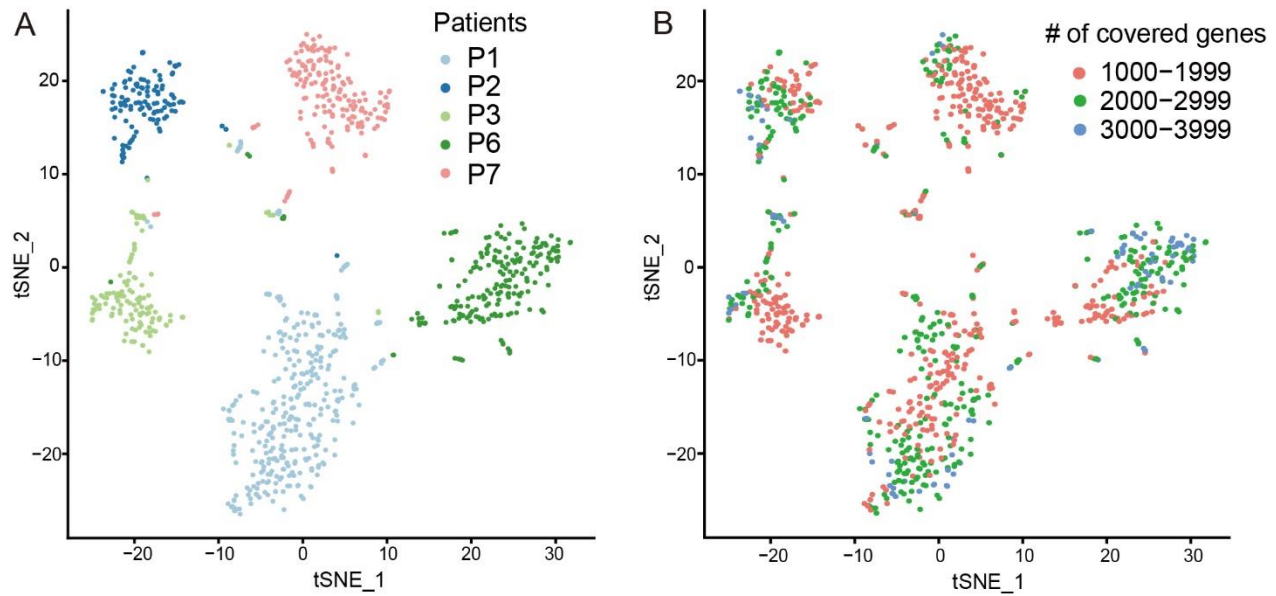


Figure S4. t-SNE plot of gene expression clustering of patient P3 with nine other CSF samples.

t-SNE clustering plot of patient P3 with normal CSF samples (N1, N2 and N3) and LUAD-LM CSF samples (P1, P2, P4, P6 and P7). LUAD-LM patient P3 CSF sample was processed following the protocol in Figure 1A without CD45 negative cell selection. Potential CTCs were manually selected based only on cell morphology. Cell identity analysis revealed that all cells collected from P3 CSF sample are monocytes.

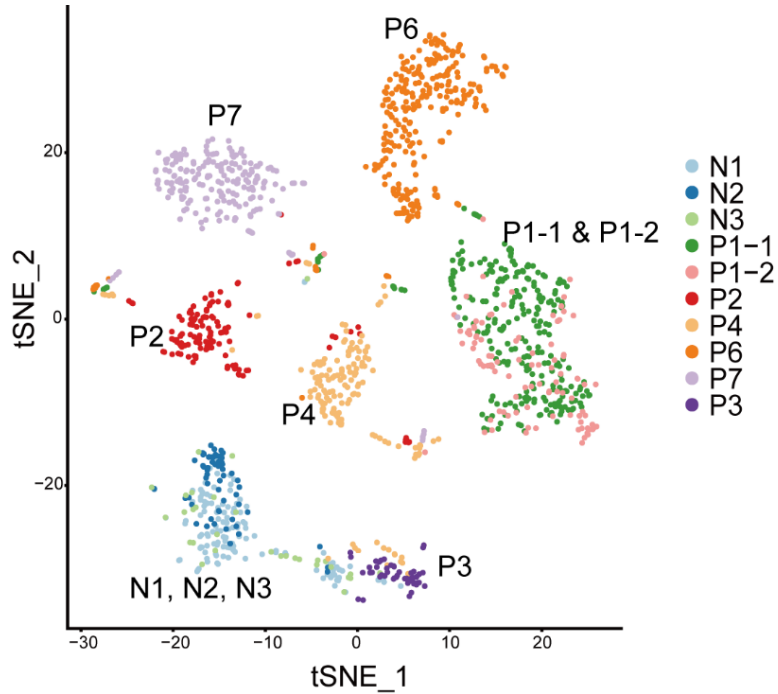


Figure S5. Characterization of the single-cell gene expression profile across all cells in CSF samples for selected marker genes.

Feature plots demonstrating the expression of the following selected genes on the t-SNE plot (Figure 2A). Scaled expression levels are depicted using a red gradient (grey denotes lack of expression).

Immune marker genes: *PTPRC*, *CD2*, *CD3D*, *CD3E*, *CD3G*, *CD14*, *CD68* and *CD163*. Lung associated genes: *SFTPA1*, *SFTPA2*, *SFTPB* and *NAPSA*. Epithelial marker genes: *EPCAM*, *CDH1*, *KRT7*, *KRT8*, *KRT17*, *KRT18*, *KRT19* and *MUC1*. Cell cycle associated genes: *CCND1* and *TOP2A*.

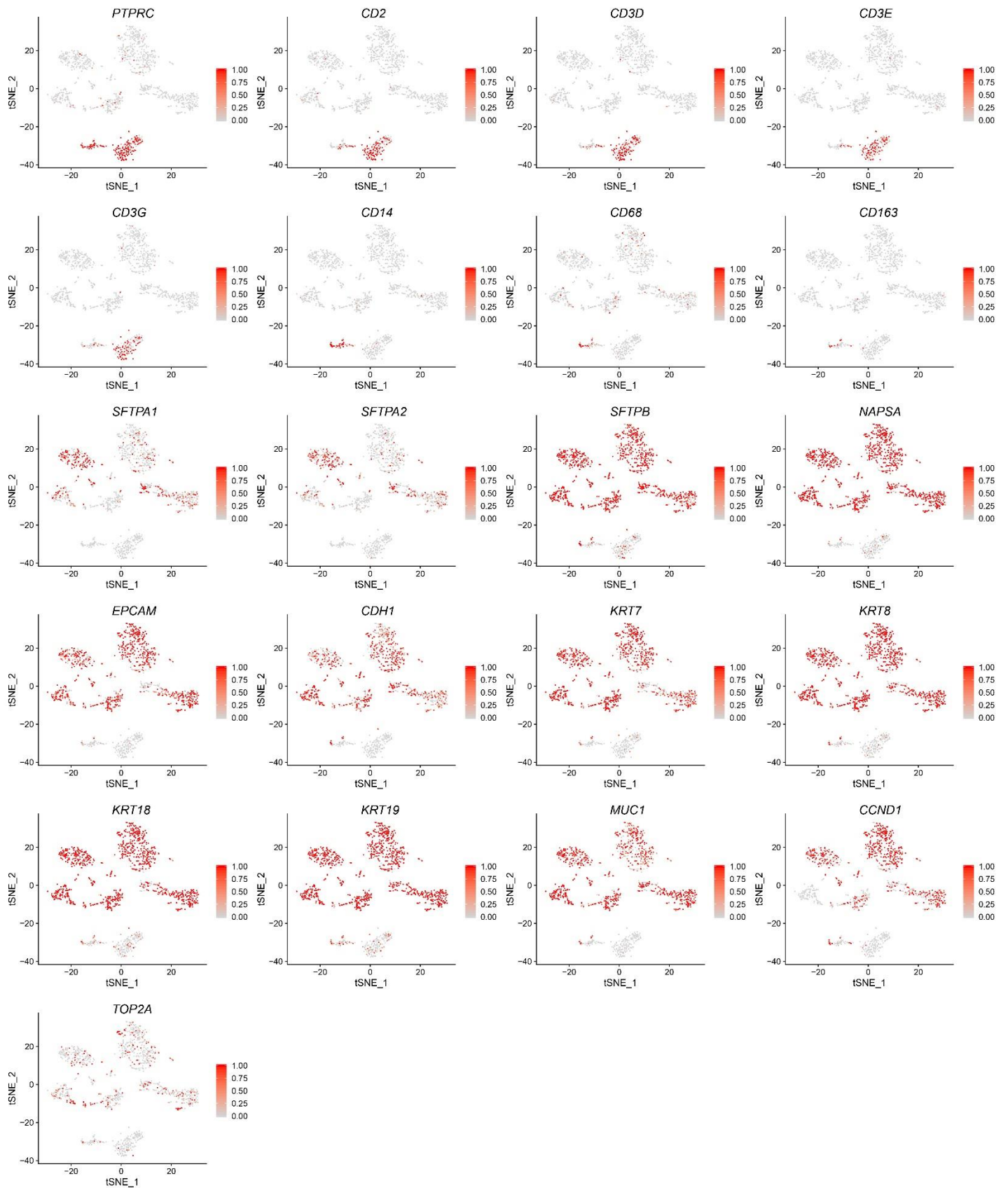


Figure S6. Inference of Cell cycle state in CSF-CTCs of LUAD-LM patients, H358 cell line and two PDX samples.

(A). Estimation of the cell cycle state of individual CTCs (dots) in P2, P6, P7, H358 cell line and two PDX samples (MBT15 and PT45) based on the relative expression of G1/S (x axis) and G2/M (y axis) gene sets. Cells are colored by their inferred cell cycle states: cycling cells (score > 0.2), red; intermediate (0 < score ≤ 0.2), pink; and noncycling cells (score ≤ 0), gray.

(B) Summary of cell cycle state proportions (y-axis) enriched in each sample (x-axis). The color code is consistent with in Figure S6A.

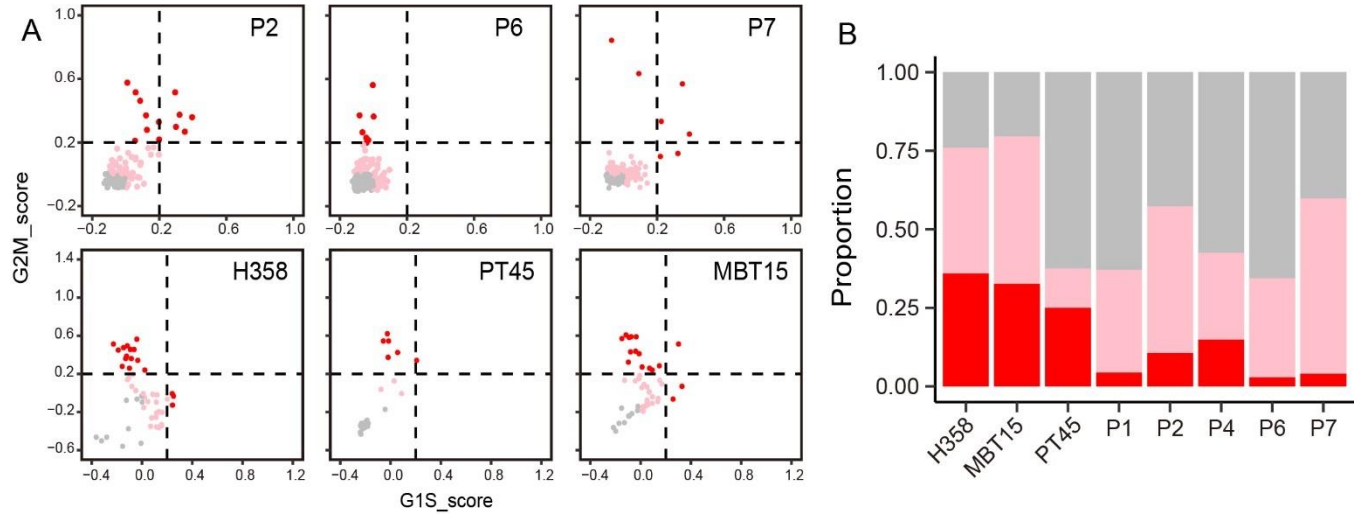


Figure S7. Single-cell gene expression profiles in CSF samples for cancer stem cell (CSC) biomarkers.

Feature plots demonstrating the expression of CSC markers (*PROM1*, *CD44*, *ALDH1A1*, *ALDH1A3* and *ALDH3A1*) on the t-SNE plot (Figure S3A). Scaled expression levels are depicted using a red gradient (grey stands for lack of expression).

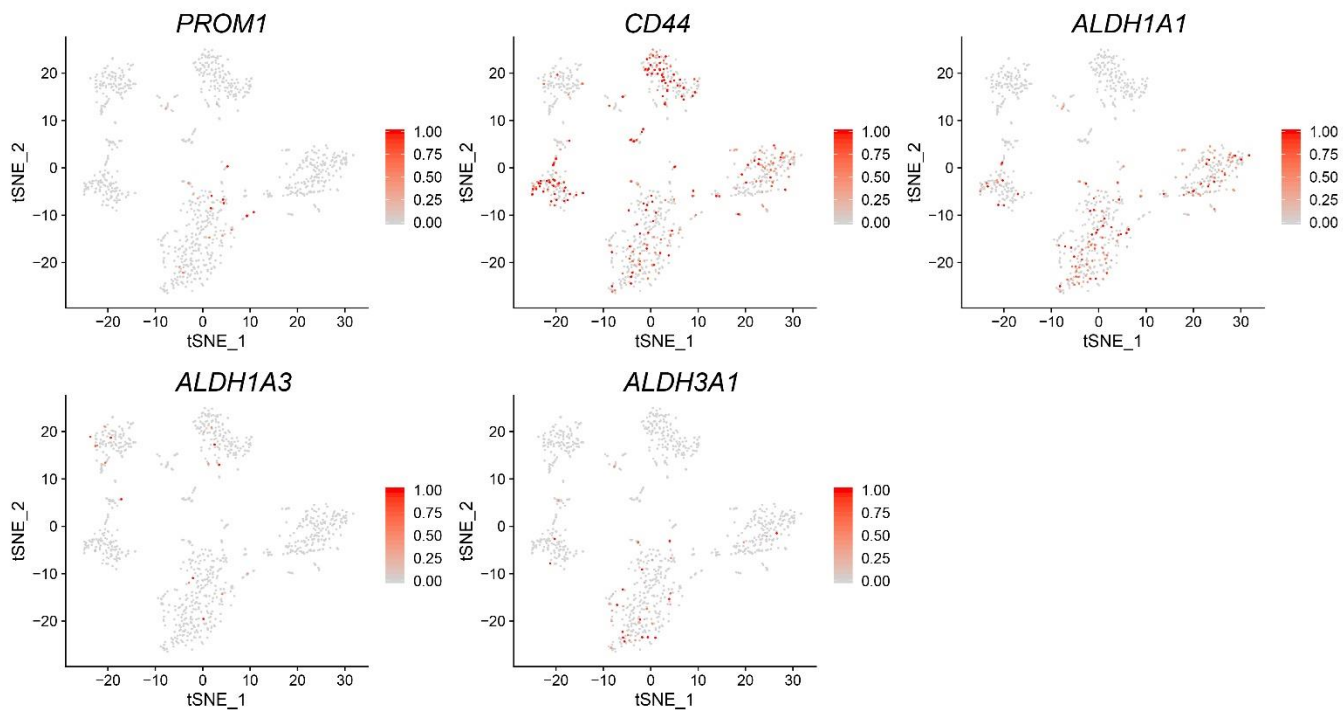


Figure S8. The epithelial, mesenchymal/CSC and extracellular matrix characteristics in CSF-CTCs from patient P2, P6 and P7.

Unsupervised clustering based on GSVA score (*top panels*) and average normalized expression level (*bottom panels*) of epithelial gene set, mesenchymal/CSC gene set and extracellular matrix gene set in CSF-CTCs from patient P2, P6 and P7. The gene sets are consistent with in Figure 4B.

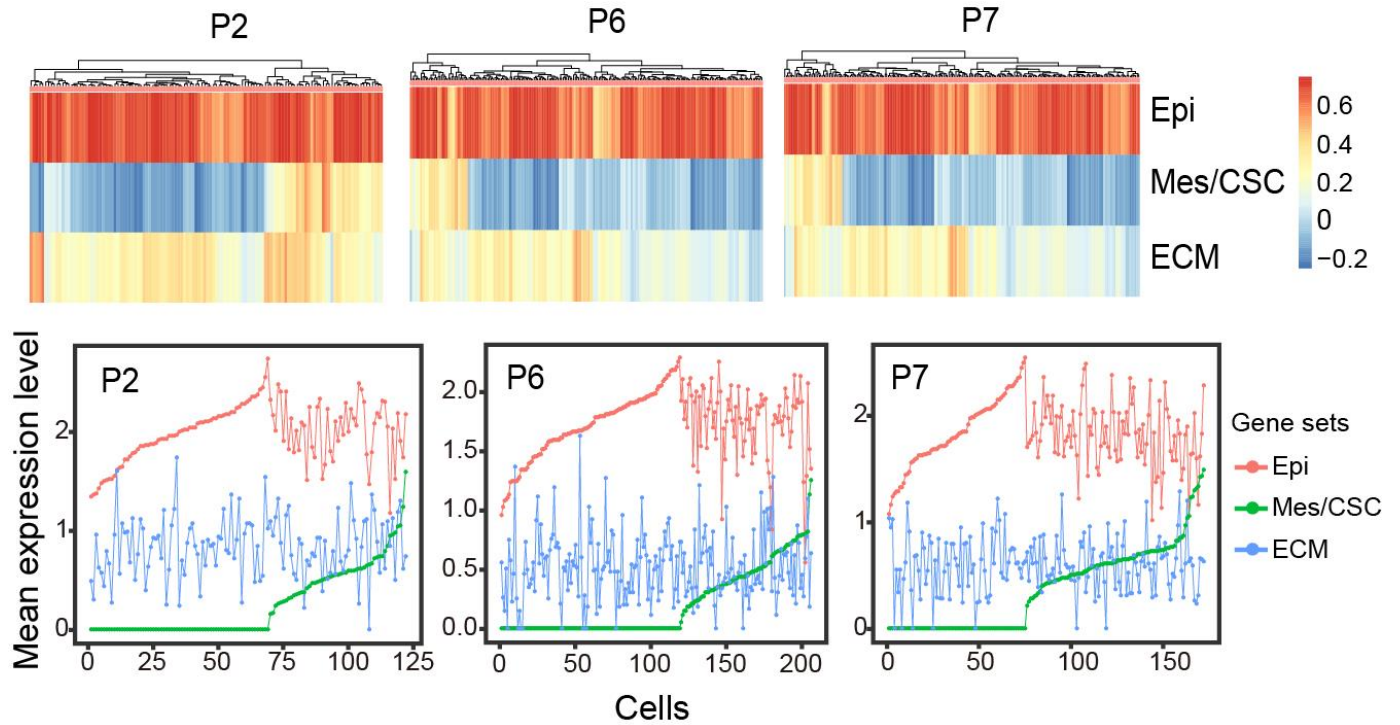


Figure S9. Heatmap of single-cell expression profile of cancer-testis antigens (CTAs) in CSF-CTCs from five LUAD-LM patients.

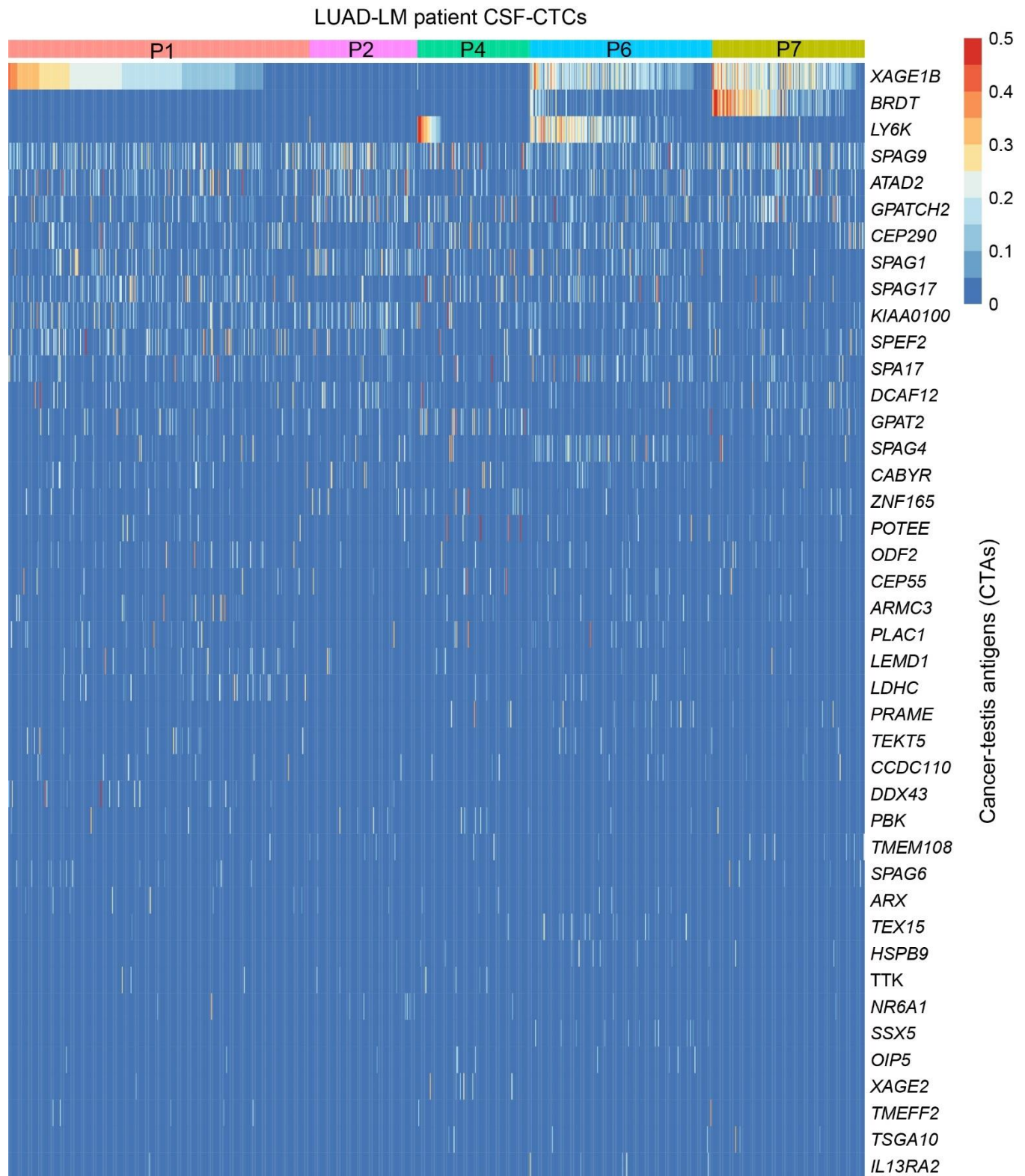


Figure S10. Clinic characteristics of CUP patient P8.

(A) Brain MRI results of patient P8, demonstrating leptomeningeal enhancement, especially in cerebellum (red arrows).

(B-I) H-E staining (B) and immunohistochemistry (IHC) of left lymph nodes biopsy for epithelial markers CK pan+, CK7+, CK8+, CK18+, CK19+, MUC1+ (C), lung cancer markers NAPSAs-, TTF1-, P63-, synaptophysin/SYP- (D), apocrine gland marker PIP/GCDFP15+ (E), proliferation marker MKI67/KI67+/- (F), gastrointestinal cancer marker VIL1/Villin- (G), prostate cancer KLK3/PAS- (H), and liver cancer marker GPC3-/Glypican- (I). "+": positive, "-": negative, "+/-": partly positive, scale bar represents 200 μ m.

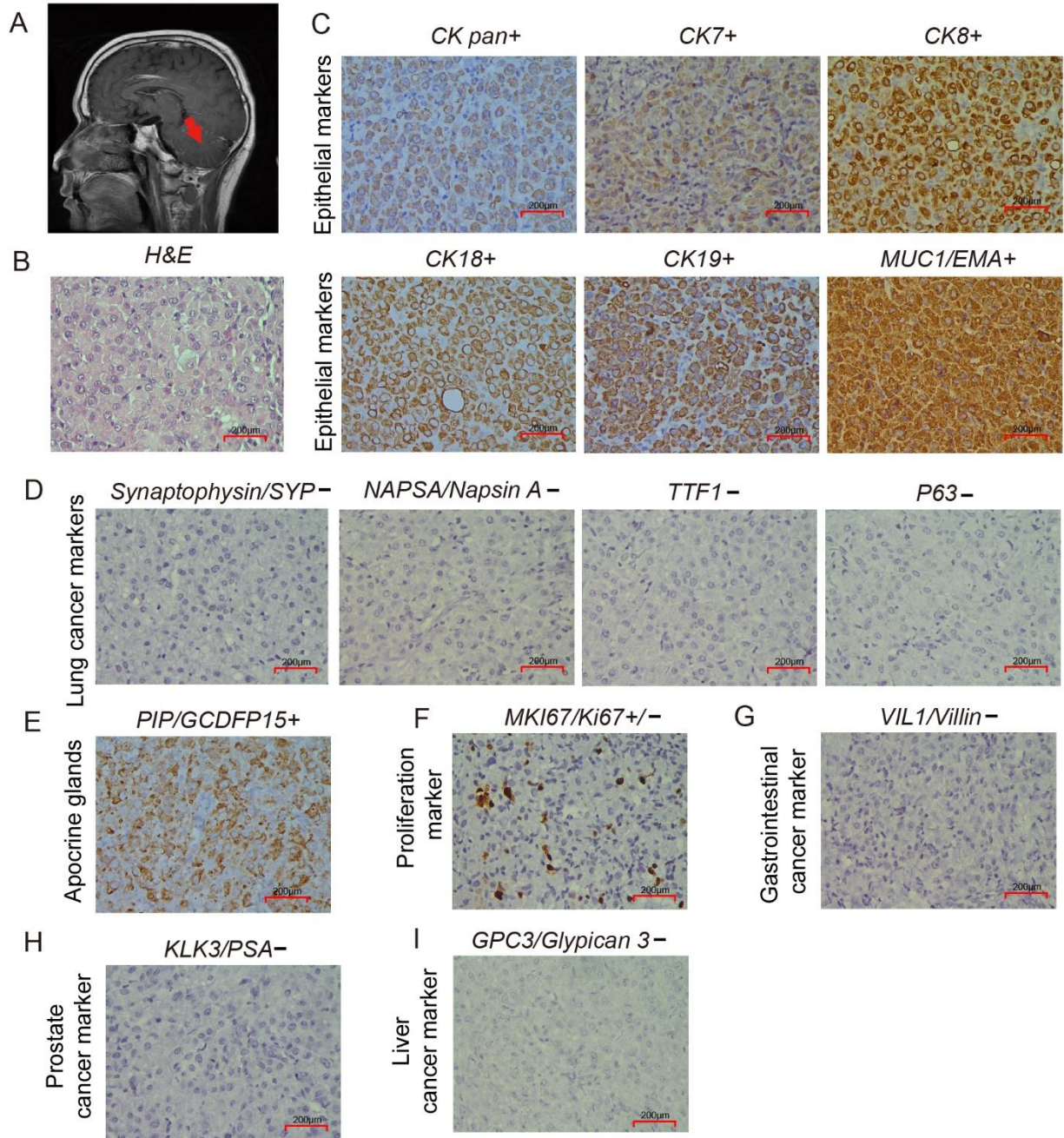


Figure S11. The gene expression profile of P8 CSF-CTCs compared to leukocytes and LUAD-LM CSF-CTCs.

(A-B) Volcano plot of significant differential gene expression (adjust P -value < 0.05) between P8 CSF-CTCs and leukocytes (A), and between P8 CSF-CTCs and LUAD-LM CSF-CTCs (B). Up-regulated (**up**) and down-regulated (**down**) genes are defined using a fold-change cutoff of 2. Gene names are labeled for selected genes, including important internal reference and marker genes. Candidate P8-specific genes (**specific**) are labeled in yellow and they are also included in Figure S12.

(C) Significantly enriched KEGG pathways in P8 CSF-CTCs compared to leukocytes by GSEA analysis (FDR: 0.05-0.20).

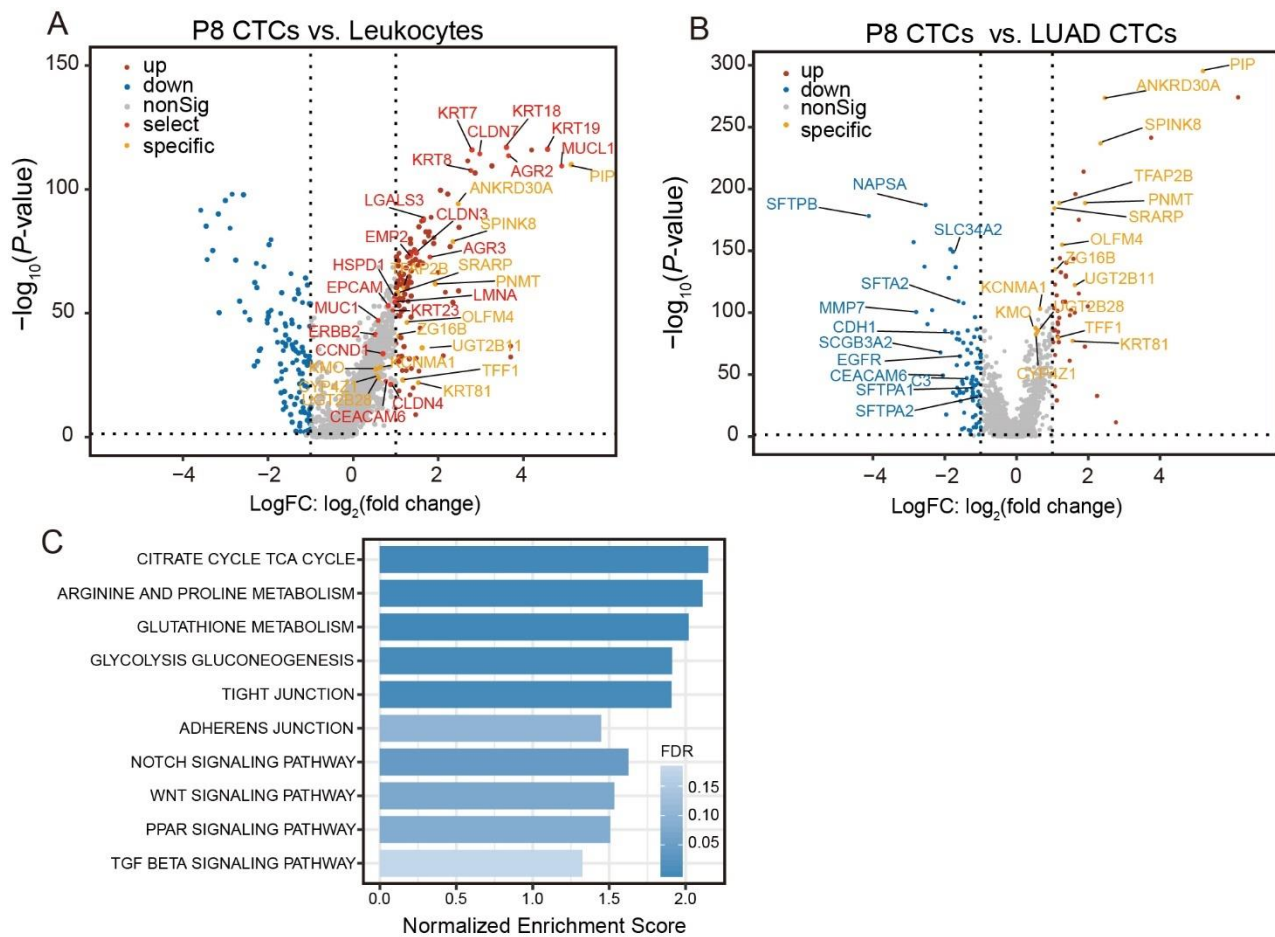


Figure S12. Characterization of the single-cell gene expression profile across all CSF cells for candidate specific genes of P8 CSF-CTCs.

Feature plots of the profile of genes with P8-specific expression on the t-SNE plot (Figure 5A). Scaled expression levels are depicted using a red gradient (grey denotes lack of expression).

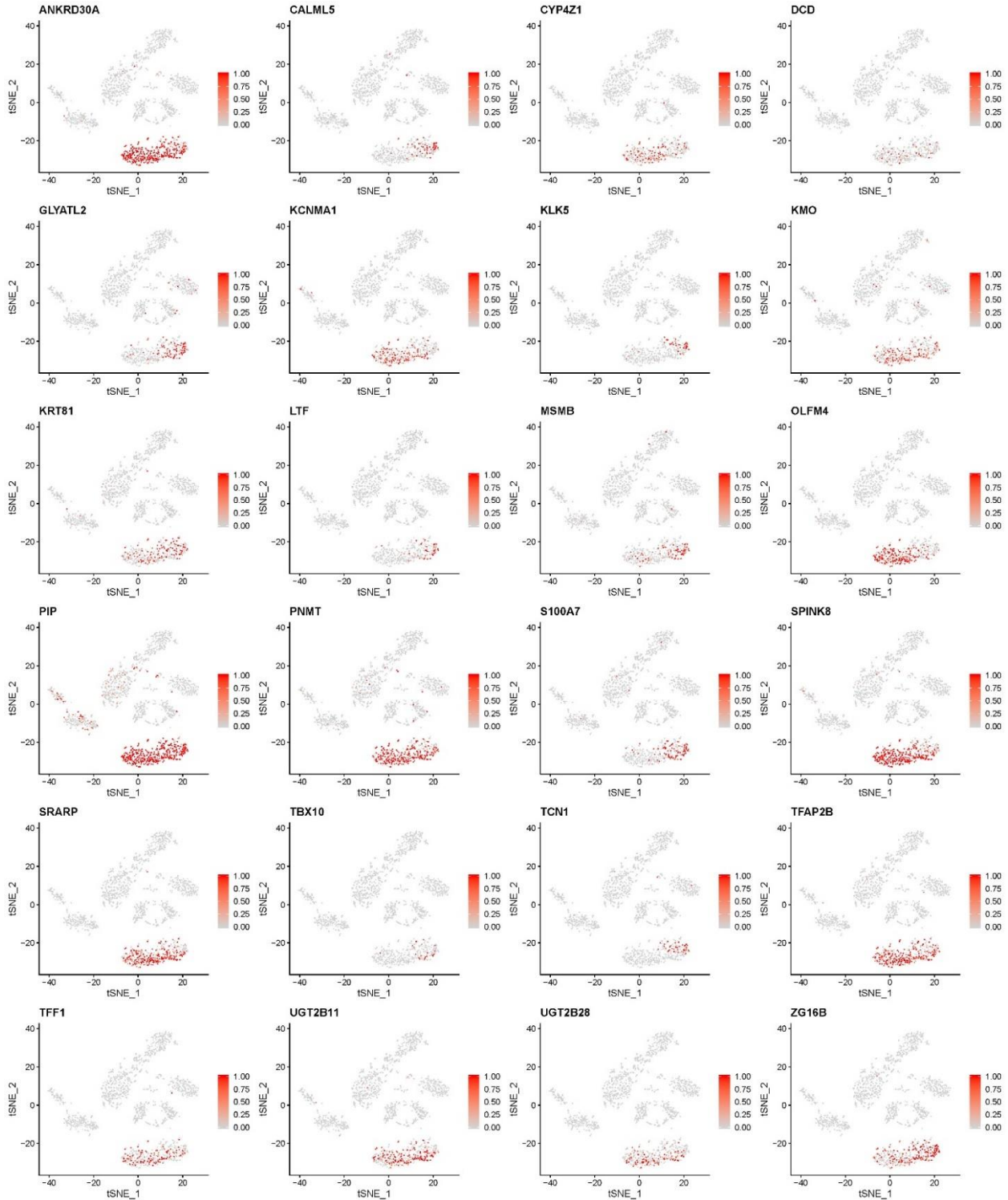


Figure S13. The expression of epithelial, mesenchymal and CSC (cancer stem cell) genes in LUAD-LM patient CSF-CTCs.

(A). Scatterplot of average expression value of epithelial genes on the x -axis and mesenchymal genes on the y -axis for every LUAD-LM CSF-CTCs. The red cells have high expression of mesenchymal genes with low expression of epithelial genes.

(B). Scatterplot of average expression value of mesenchymal genes on the x -axis and $CD44$ on the y -axis for every LUAD-LM CSF-CTCs.

Epithelial (Epi) genes: *CDH1*, *EPCAM*, *KRT18*, *KRT19*, *KRT7*, *KRT8* and *MUC1*; Mesenchymal (Mes) genes: *FNI* and *VIM*.

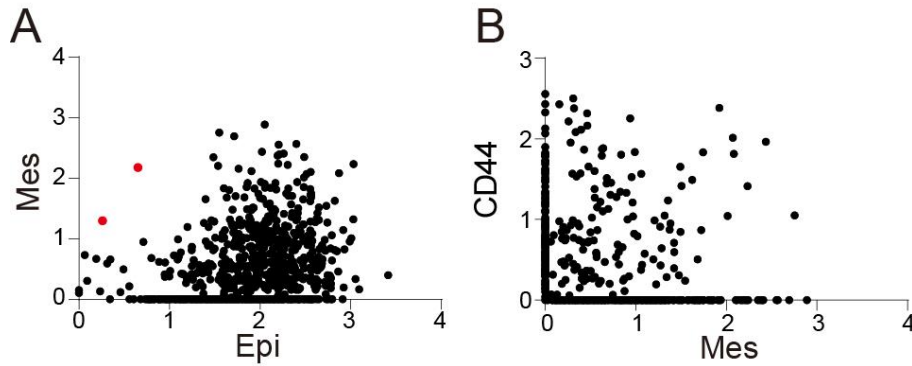


Table S1. Clinical information and cell selection conditions for CSF and blood samples.

Patient ID/info	Age	Gender	Cancer diagnosis time	LM diagnosis time	Metastatic tumor location in LM by MRI	Sample collection date	Markers for FACS sorting	Number of cells sequenced
P1 LUAD	36	Male	11/2017	01/2018	Left frontal lobe, Right parietal lobe	11/1/2018 (P1-1)	Calcein Blue AM+	168
						1/7/2019 (P1-2)	CD45-	360
P2 LUAD	62	Female	Unknown	12/2018	Bilateral temporal lobe	12/17/2018	Calcein Blue AM+ CD45-	192
P4 LUAD	65	Female	11/2017	02/2019	Cerebellum	3/1/2019	Calcein Blue AM+ CD45-	288
P6 LUAD	55	Male	12/2012	03/2019	Bilateral temporal lobe	3/8/2019	Calcein Blue AM+ CD45-	480
P7 LUAD	60	Male	07/2017	05/2019	/	5/10/2019	Calcein Blue AM+ CD45-	288
P3 LUAD	49	Male	02/2017	02/2018	Right temporal lobe	3/3/2019	Calcein Blue AM+	96
P8 CUP	49	Male	12/2017	05/2018	Cerebellum and Cerebral hemisphere	1/14/2019 (P8-1)	Calcein Blue AM+	480
						6/21/2019 (P8-1)	CD45	816
N1 control	40	Male	N/A	N/A	N/A	10/8/2018	Calcein Blue AM+	288
N2 control	35	Female	N/A	N/A	N/A	10/10/2018	Calcein Blue AM+	240
N3 control	39	Male	N/A	N/A	N/A	10/18/2018	Calcein Blue AM+	96
Blood-T cells	27	Female	N/A	N/A	N/A	1/25/2019	Calcein Blue AM+ CD45+ CD3+	168
Blood-B cells	27	Female	N/A	N/A	N/A	1/25/2019	Calcein Blue AM+ CD45+ CD19+	168

LUAD: lung adenocarcinoma; CUP: cancer of unknown primary site; LM: leptomeningeal metastases; FACS: fluorescence activating cell sorter; Calcein Blue AM: labeling dye for live cells selection; CD45: protein tyrosine phosphatase receptor type C, marker for leukocytes; CD19: marker for B cells; CD3: marker for T cells; N/A: Not applicable; “/”: no leptomeningeal enhancement shown by MRI.

Table S2. Summary of cell type identity of CSF samples in normal patients, blood T cell and blood-B cell samples.

Sample ID	Total filtered cells	T cells	B cells	Monocytes
N1	127	105		22
N2	53	51		2
N3	27	14		13
Blood-T	41	41		
Blood-B	41		41	
Total	289	211	41	37

Table S3. Summary of number of sequenced reads and detected genes of CSF samples in patients, normal patients, blood T cell and blood-B cell samples.

Sample ID	Age/Sex	Sequenced reads	Total genes detected	# of cells with genes counts > 1000
P1-1	36/Male	58,380,064	11,655	436
P1-2	36/Male	237,321,692	13,813	1,412
P2	62/Female	147,219,734	12,893	652
P4	65/Female	350,537,537	13,518	1,203
P6	55/Male	302,888,314	14,262	1,692
P7	60/Male	94,506,564	12,170	371
P3	49/Male	89,980,565	9,243	233
P8-1	49/Male	297,958,378	13,340	2,012
N1	40/Male	329,258,208	13,592	1,332
N2	35/Female	125,071,290	10,944	308
N3	39/Male	94,756,214	10,101	236
Blood-T	27/Female	105,595,344	7,522	90
Blood-B	27/Female	91,515,818	9,084	428

Table S4. Summary of disease relevant mutations detected in CSF cell free DNA (CSF-cfDNA) by next-generation sequencing.

Patient ID	Sample collection date	EGFR	TP53	ALK	KRAS(P21)
P1	20180122	p.Glu746_Ala750del (25.34%)	p.Trp53Ter (85.47%)	p.Arg1192Trp (82.66%)	p.Gln61Leu (0.13%), p.Gly138Arg (0.20%)
P2	N/A				
P4	20190307	p.Leu858Arg (16.09%)	p.Arg248Gln (44.72%)		
P6	N/A				
P7	20190510		p.Asp281Tyr (89.72%)		

p: protein; Numbers in brackets indicate mutation frequencies.