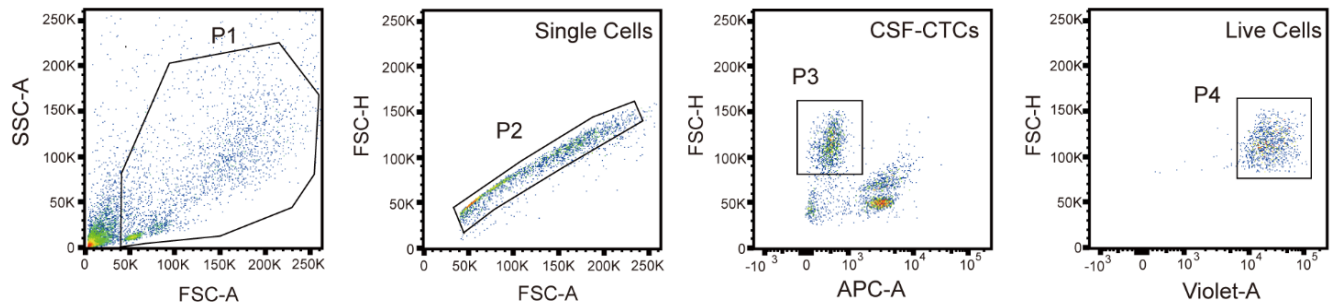


Supplemental Figure 1

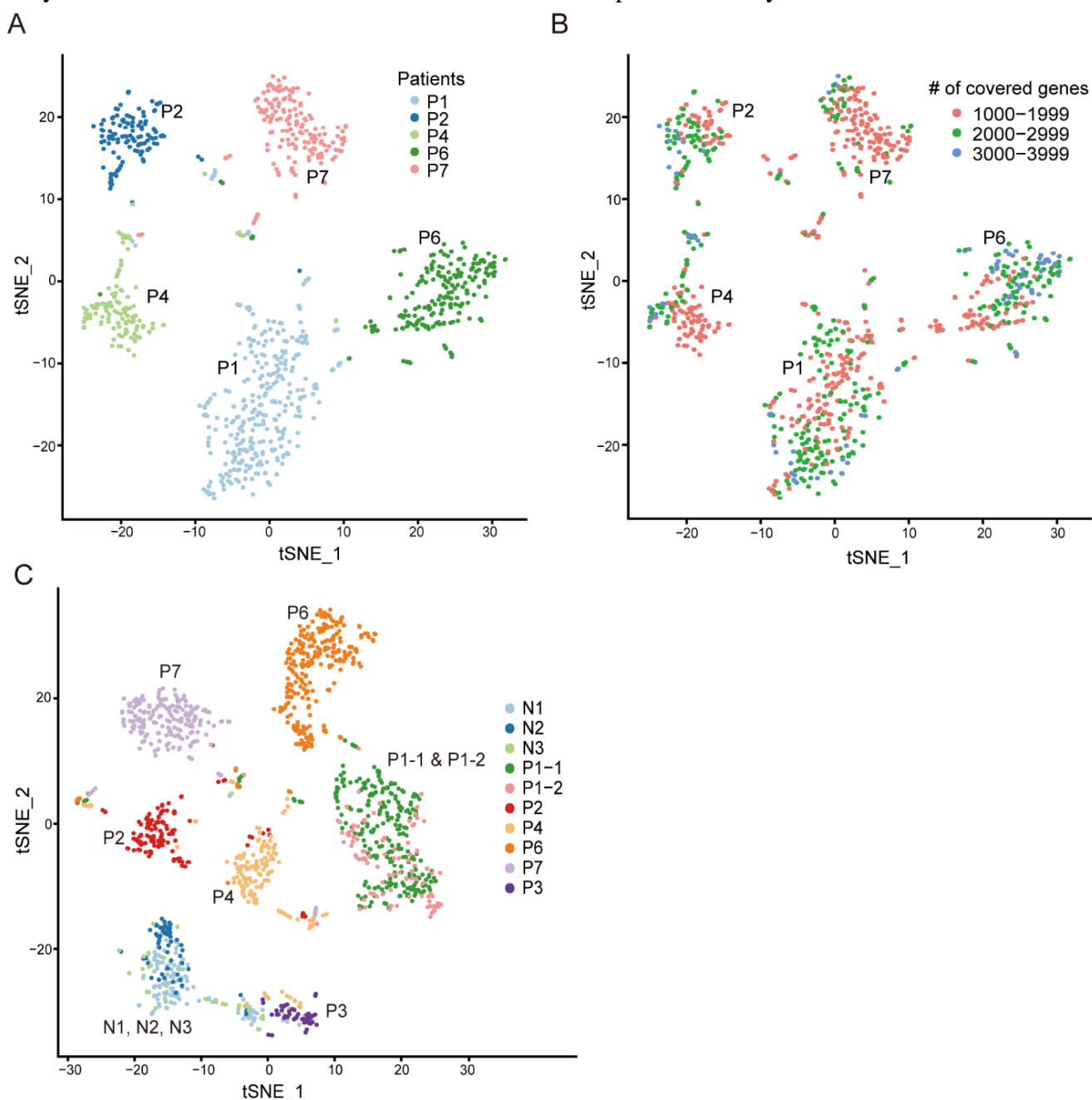
Fluorescence-activated cell sorting (FACS) of CSF-CTCs from LUAD-LM patient samples.

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.



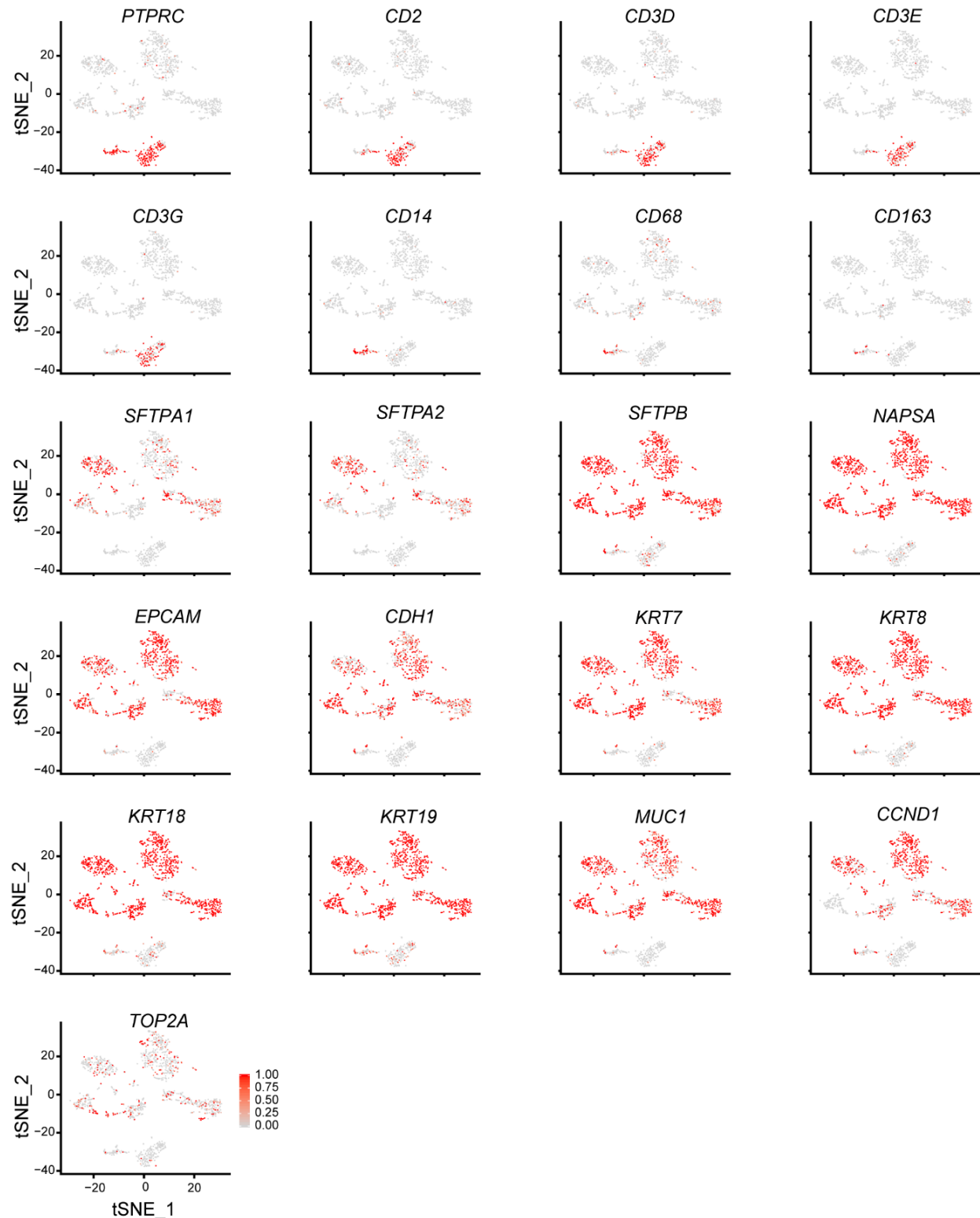
Supplemental Figure 2

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 Supplemental Figure 2A. The clustering pattern is not driven by the number of covered genes. (C) 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. Candidate CTCs were manually selected based on cell morphology only. Cell identity analysis revealed that all cells collected from P3 CSF sample are monocytes.



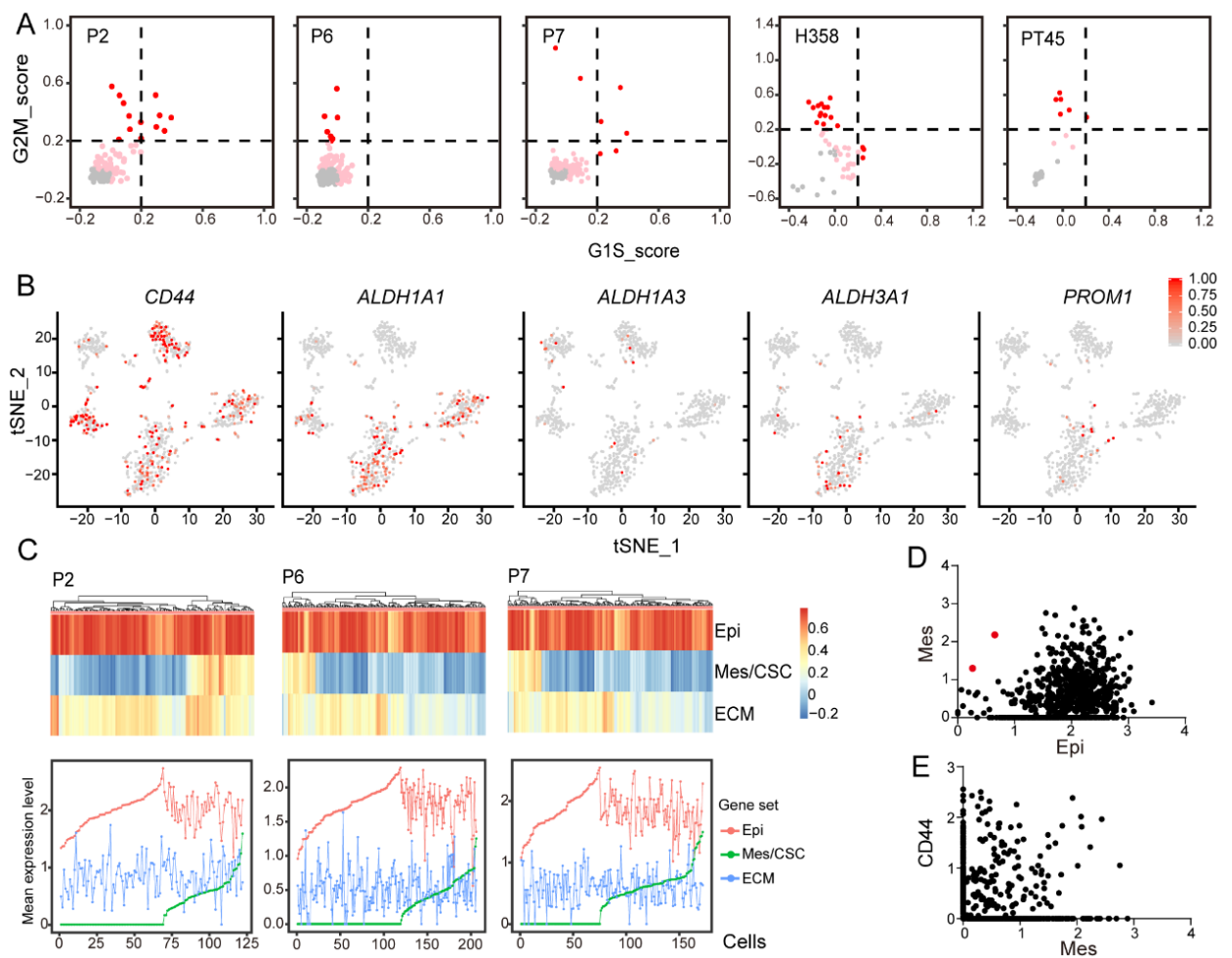
Supplementary Figure 3

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*.



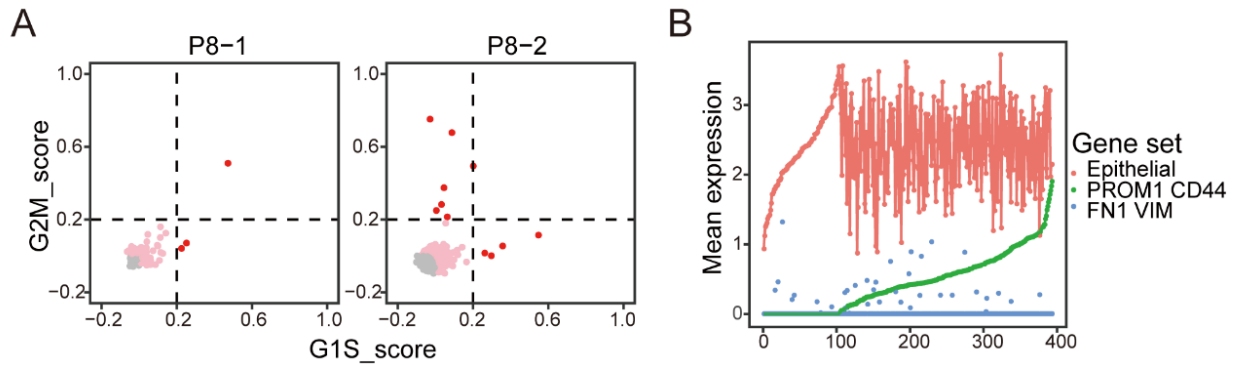
Supplementary Figure 4

Gene expression profile of cell cycle genes, CSC genes and genes related to partial EMT in CSF-CTCs. (A) Estimation of the cell cycle state of individual CTCs (dots) in P2, P6, P7, H358 cell line and PT45 based on the relative expression of G1/S (x axis) and G2/M (y axis) gene sets. Cells are colored as in Figure 4A. (B) Feature plots demonstrating the expression of CSC markers (*PROM1*, *CD44*, *ALDH1A1*, *ALDH1A3* and *ALDH3A1*) on the t-SNE plot (Supplementary Figure 2A). Scaled expression levels are depicted using a red gradient (grey stands for lack of expression). (C) Unsupervised clustering based on GSVA score (*top*) and average normalized expression level (*bottom*) of epithelial gene set, mesenchymal/CSC gene set and extracellular matrix gene set in CSF-CTCs from patient P2, P6 and P7. (D) 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. (E) Scatterplot of average expression value of mesenchymal genes on the x-axis and *CD44* on the y-axis for every LUAD-LM CSF-CTCs. The gene sets used in Extended Data Fig. 4c-e are the same as in Figure 4C.



Supplementary Figure 5

Gene expression profile of genes related to cell cycle, partial EMT and CSC in P8 CSF-CTCs. (A) Cell cycle state of individual CSF-CTCs (dots) inferred from relative expression of G1/S (x-axis) and G2/M (y-axis) gene sets in P8-1 and P8-2 samples. (B) Average normalized expression level (y-axis) of epithelial genes (*EPCAM*, *KRT18*, *KRT19*, *KRT7*, *KRT8*, *MUC1*), mesenchymal genes (*VIM*, *FN1*) and CSC genes (*PROM1*, *CD44*) in P8 cells (x-axis).



Supplementary Table 1. 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+ CD45-	168
						1/7/2019 (P1-2)		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+ CD45	480
						6/21/2019 (P8-1)		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.

Supplementary Table 2. 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

Supplementary Table 3. Summary of sequenced reads and detected genes of CSF samples in patients, normal patients, blood T cell and blood-B cell samples.

Sample ID	# of sequencing reads	# of filtered reads*	# of uniquely mapped reads	Average # of uniquely mapped reads per cell	Total genes detected	# of genes whose counts>1000
P1-1	58380064	15206906	6734715	40088	11655	436
P1-2	237321692	42055297	16346993	45408	13813	1412
P2	147219734	33865056	12336410	64252	12893	652
P4	350537537	15235749	6579460	22845	13518	1203
P6	302888314	123455881	47199598	98332	14262	1692
P7	94506564	24521678	7850335	27258	12170	371
P3	89980565	6222803	2340870	24384	9243	233
P8-1	297958378	74043518	35977243	74953	13340	2012
P8-2	329258208	73341499	27047520	70436	13340	1332
N1	125071290	27453376	6834185	23730	13592	308
N2	224306614	49201160	5036090	26230	10944	236
N3	105595344	12538866	3107257	32367	10101	90
Blood-T	91515818	22523435	9798201	58323	7522	428
Blood-B	104966890	25259710	11184723	66576	9084	484

*: *filtered reads*: demultiplex combined sequenced reads via barcodes; removed poly-T bases in forward reads using manual scripts; for each read, extracted barcode and expand read name with it; remove low-quality bases and reads using *Trim galore!* setting `-q=25 --phred33 --stringency=3`. *mapping parameters*: reference genome is hg38; default parameters of STAR alignment.

Supplementary Table 4. 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.

Supplementary Table 5. List of 25 top candidate patient P8 cluster defining genes.

gene	average log₂ (FC)	adjusted <i>P</i>- value	pct in P8	pct in other patients	pct diff
<i>PIP</i>	7.496	0.000	0.945	0.050	0.895
<i>ANKRD30A</i>	3.423	0.000	0.794	0.007	0.787
<i>SERHL2</i>	2.550	0.000	0.699	0.034	0.665
<i>PNMT</i>	2.620	0.000	0.606	0.010	0.596
<i>SRARP</i>	1.415	0.000	0.554	0.000	0.554
<i>UGT2B11</i>	2.201	0.000	0.400	0.004	0.396
<i>MEIS1</i>	1.281	0.000	0.400	0.025	0.375
<i>TRPS1</i>	0.779	0.000	0.386	0.025	0.361
<i>CKMT1B</i>	0.749	0.000	0.400	0.042	0.358
<i>KMO</i>	0.688	0.000	0.335	0.015	0.320
<i>MUC5B</i>	1.478	0.000	0.313	0.026	0.287
<i>UGT2B28</i>	0.784	0.000	0.279	0.000	0.279
<i>CYP4Z1</i>	0.705	0.000	0.281	0.002	0.279
<i>TFF3</i>	2.027	0.000	0.309	0.031	0.278
<i>KRT81</i>	2.123	0.000	0.287	0.009	0.278
<i>HSD17B2</i>	0.587	0.000	0.212	0.006	0.206
<i>HPX</i>	0.603	0.000	0.216	0.021	0.195
<i>CCDC160</i>	0.635	0.000	0.196	0.003	0.193
<i>MOV10L1</i>	0.585	0.000	0.194	0.020	0.174
<i>MUC5AC</i>	0.928	0.000	0.174	0.003	0.171
<i>ZCCHC2</i>	0.906	0.000	0.172	0.031	0.141
<i>EFHD1</i>	0.700	0.000	0.139	0.025	0.114
<i>DCD</i>	3.152	0.000	0.111	0.000	0.111
<i>IGKC</i>	2.774	0.000	0.121	0.041	0.080
<i>R3HCC1L</i>	0.651	0.019	0.111	0.049	0.062

Patient P8 cluster defining genes were selected compared to other CSF samples (Figure 5i) by fold change > 1.5, pct in P8 ≤ 0.05 and adjusted *P*-value < 0.05; pct in P8: the percentage of cells where the gene is detected in P8 CSF samples; pct in other patients: the percentage of cells where the gene is detected in other CSF samples; pct diff: the pct difference in P8 compared to other CSF samples (P8 minus others).

P8 cluster defining genes with pct dff > 0.2 are shown in Fig. 6b.

Supplementary Table 6. List of patient P8 cluster defining genes with significant biased expression during tumor progression between P8-1 and P8-2 samples

gene	avg log ₂ FC	adjusted P-value	pct in P8	pct in others	pct in P8-1	pct in P8-2	adjusted P-value	avg log ₂ FC	pct diff
<i>GLYATL2</i>	1.219	0.000	0.196	0.011	0.486	0.082	0.000	1.708	0.404
<i>KLK5</i>	0.853	0.000	0.109	0.001	0.297	0.031	0.000	1.657	0.266
<i>ZG16B</i>	1.435	0.000	0.426	0.003	0.662	0.386	0.000	1.093	0.276
<i>S100A8</i>	2.488	0.000	0.293	0.017	0.568	0.171	0.000	0.639	0.397
<i>S100A7</i>	2.382	0.000	0.214	0.004	0.493	0.109	0.000	0.830	0.384
<i>MSMB</i>	0.855	0.000	0.188	0.005	0.372	0.126	0.000	1.021	0.246
<i>C15orf48</i>	0.603	0.000	0.188	0.023	0.385	0.113	0.000	0.915	0.272
<i>CALML5</i>	0.792	0.000	0.158	0.004	0.351	0.085	0.000	1.090	0.266
<i>TCN1</i>	0.787	0.000	0.115	0.003	0.264	0.055	0.000	1.102	0.209
<i>SPINK8</i>	3.245	0.000	0.703	0.002	0.486	0.918	0.000	-2.022	-0.432
<i>OLFM4</i>	1.706	0.000	0.481	0.000	0.122	0.741	0.000	-2.002	-0.619
<i>CRYAB</i>	2.418	0.000	0.626	0.030	0.486	0.792	0.000	-0.939	-0.306
<i>TFAP2B</i>	1.613	0.000	0.576	0.001	0.466	0.727	0.002	-0.449	-0.261
<i>KCNMA1</i>	0.859	0.000	0.337	0.001	0.182	0.468	0.000	-0.546	-0.286
<i>TFF1</i>	1.561	0.000	0.267	0.001	0.149	0.372	0.009	-0.602	-0.223

Patient P8 cluster defining genes were selected compared to other CSF clusters (Figure 5i) by fold change > 1.5, pct in P8 ≤ 0.05 and adjusted P-value < 0.05; pct in P8: the percentage of cells where the gene is detected in P8 CSF samples; pct in others: the percentage of cells where the gene is detected in other CSF samples.

The 15 biased expression genes in P8-1 or P8-2 of P8 cluster defining genes were selected by pct in P8-1 or P8-2 < 0.5, |pct diff| > 0.2 and adjust P-value; pct in P8-1, the percentage of cells where the gene is detected in P8-1 CTCs; pct in P8-2, the percentage of cells where the gene is detected in P8-2 CTCs; pct diff, the different percentage of cells where the gene is detected in P8-1 compared to P8-2 (P8-1 minus P8-2).

P8 cluster defining genes with expression biased in P8-1 CTCs (9 genes) are shown in Fig. 6d, and with expression biased in P8-2 CTCs (6 genes) are shown in Fig. 6e.