

Table S5. Pan-cancer single-cell transcriptome data

Dataset Name	Species	Cells	Platform	PMID
AEL_GSE142213	Human	3994	10x Genomics	32330454
ALL_GSE132509	Human	37936	10x Genomics	32415257
ALL_GSE153697	Human	2904	10x Genomics	33558546
ALL_GSE154109	Human	10800	10x Genomics	32692727
AML_GSE116256	Human	38348	Smart-seq2	30827681
AML_GSE135851	Human	1794	10x Genomics	32603599
AML_GSE154109	Human	9623	10x Genomics	32692727
BCC_GSE123813	Human	52884	10x Genomics	31359002
BCC_GSE141526	Human	56116	10x Genomics	33033234
BRCA_EMTAB8107	Human	33043	10x Genomics	32561858
BRCA_GSE136206	Mouse	27532	10x Genomics	31730857
BRCA_GSE143423	Human	4375	10x Genomics	
BRCA_GSE148673	Human	10359	10x Genomics	33462507
BRCA_GSE150660	Human	10605	10x Genomics	32675368
BRCA_GSE161529	Human	332168	10x Genomics	33950524
BRCA_GSE176078	Human	89471	10x Genomics	34493872
BRCA_SRP114962	Human	2472	SNRS	29681456
CESC_GSE168652	Human	22998	10x Genomics	33996252
CHOL_GSE125449	Human	5761	10x Genomics	31588021
CHOL_GSE138709	Human	33990	10x Genomics	32505533
CHOL_GSE142784	Human	10535	10x Genomics	32505533
CLL_GSE132065	Human	47695	10x Genomics	32393363
CLL_GSE152469	Human	13017	CITE-seq	33298182
CLL_GSE159417	Human	4763	10x Genomics	34362918
CRC_EMTAB8107	Human	23176	10x Genomics	32561858
CRC_GSE146771	Human	43817	10x Genomics	32302573
CRC_GSE166555	Human	66050	10x Genomics	34409732
DLBC_GSE175510	Human	1412	Smart-seq2	34471864
ESCA_GSE160269	Human	208658	10x Genomics	34489433
GIST_GSE162115	Human	35308	10x Genomics	33393143
Glioma_GSE131928	Human	13553	10x Genomics	31327527
Glioma_GSE138794	Human	18458	10x Genomics	31554641
Glioma_GSE139448	Human	12152	10x Genomics	32004492
Glioma_GSE141383	Human	10502	Microwell-seq	34011400
Glioma_GSE141460	Human	6385	Smart-seq2	32663469
Glioma_GSE141982	Human	5263	10x Genomics	32105316
Glioma_GSE148842	Human	111397	Microwell	33975634
Glioma_GSE84465	Human	3533	Smart-seq2	29091775
HB_GSE180665	Human	46486	10x Genomics	34497364
HNSC_GSE103322	Human	5902	Smart-seq2	29198524
KICH_GSE159115	Human	2850	10x Genomics	34099557
KIRC_GSE159115	Human	27669	10x Genomics	34099557
KIRC_GSE171306	Human	11427	10x Genomics	34168986
LIHC_GSE125449	Human	3834	10x Genomics	31588021
LIHC_GSE146115	Human	3134	C1	33531041
LIHC_GSE146409	Human	2916	MARS-seq	33332768
LIHC_GSE166635	Human	22631	10x Genomics	33619115
LSCC_GSE150321	Human	16869	10x Genomics	32638385
MB_GSE119926	Human	7745	Smart-seq2	31341285
MB_GSE155446	Human	37445	10x Genomics	34077540
MCC_GSE117988	Human	10134	10x Genomics	30250229
MCC_GSE118056	Human	11024	10x Genomics	30250229
MF_GSE165623	Human	9753	10x Genomics	33968070
MF_GSE173820	Human	3926	10x Genomics	34018188
MM_GSE117156	Human	24918	MARS-seq	30523328
MM_GSE143317	Human	4882	10x Genomics	33619368

MM_GSE161801	Human	204045	10x Genomics	34845188
MM_GSE175385	Human	79277	10x Genomics	34133718
NET_GSE140312	Human	3158	10x Genomics	32054662
NPC_GSE150430	Human	45959	10x Genomics	32686767
NPC_GSE162025	Human	176440	10x Genomics	33531485
NSCLC_EMTAB6149	Human	40218	10x Genomics	29988129
NSCLC_GSE117570	Human	11453	10x Genomics	31033233
NSCLC_GSE127465	Human	31179	Smart-seq2	30979687
NSCLC_GSE143423	Human	12193	10x Genomics	
NSCLC_GSE148071	Human	82267	GEXSCOPETM	33953163
NSCLC_GSE150660	Human	5407	10x Genomics	32675368
OS_GSE162454	Human	46544	10x Genomics	34367994
OSCC_GSE172577	Human	57503	10x Genomics	34044317
OV_EMTAB8107	Human	24781	10x Genomics	32561858
OV_GSE118828	Human	1909	Smart-seq2	30383866
OV_GSE130000	Human	11215	Drop-seq	34992217
OV_GSE147082	Human	9796	Drop-seq	34038734
OV_GSE151214	Human	59446	10x Genomics	33852846
OV_GSE154600	Human	42583	10x Genomics	32747365
OV_GSE158722	Human	96846	10x Genomics	34031395
PAAD_CRA001160	Human	57443	10x Genomics	31273297
PAAD_GSE111672	Human	6122	inDrop	31932730
PAAD_GSE141017	Human	4934	10x Genomics	33649045
PAAD_GSE148673	Human	6196	10x Genomics	33462507
PAAD_GSE154778	Human	14953	10x Genomics	32988401
PAAD_GSE162708	Human	22133	10x Genomics	34671197
PAAD_GSE165399	Human	9569	GEXSCOPETM	34326696
PPB_GSE163678	Human	12239	10x Genomics	
PRAD_GSE137829	Human	8640	10x Genomics	33328604
PRAD_GSE141445	Human	33441	10x Genomics	33420488
PRAD_GSE143791	Human	36850	10x Genomics	34719426
PRAD_GSE176031	Human	18807	Drop-seq	35013146
SCLC_GSE150766	Human	1924	10x Genomics	34016693
SKCM_GSE115978	Human	7186	Smart-seq2	30388455
SKCM_GSE72056	Human	4645	Smart-seq2	27124452
SS_GSE131309	Human	7109	Smart-seq2	33495604
STAD_GSE134520	Human	41554	10x Genomics	31067475
THCA_GSE148673	Human	19407	10x Genomics	33462507
UVM_GSE138433	Human	12682	10x Genomics	33462406
UVM_GSE139829	Human	103703	10x Genomics	31980621
UVM_GSE160883	Human	12977	inDrop	34518527