

Table S1 Applied data and software in this study

	Source	Website
Datasets		
Raw genomic and clinical data	Genomic Data Commons Data Portal	https://portal.gdc.cancer.gov/
Arranged clinical data	Hoadley <i>et al.</i> 2018	https://gdc.cancer.gov/about-data/publications/pancanatlas
Single cell RNA-seq for human early embryo	Yan <i>et al.</i> 2013	http://single-cell.clst.riken.jp/fpkm_tables/GSE36552_fpkm.table.gz
mRNA signature index (mRNAsi)	Malta <i>et al.</i> 2018	https://gdc.cancer.gov/about-data/publications/PanCanStemness-2018
Software and Algorithms		
R-3.6.2	The R Project for Statistical Computing	https://www.r-project.org/
Rstudio 1.1.423	Rstudio Company	https://www.rstudio.com/
Rtsne 0.15	Krijthe <i>et al.</i> 2018	https://cran.r-project.org/web/packages/Rtsne/index.html
ggplot2 2.2.1	Wickham, 2009	https://cran.r-project.org/web/packages/ggplot2/index.html
pheatmap 1.0.8	Kolde <i>et al.</i> 2018	https://cran.r-project.org/web/packages/pheatmap/index.html
survminer 0.4.3	Kassambara <i>et al.</i> 2018	https://cran.rstudio.com/web/packages/survminer/index.html
clusterProfiler 3.6.0	Yu <i>et al.</i> 2017	http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html
calling mRNAsi	Malta <i>et al.</i> 2018	http://tcgabiolinks.fmrp.usp.br/PanCanStem/
CIBERSORT	Gentles <i>et al.</i> 2015	https://precog.stanford.edu/

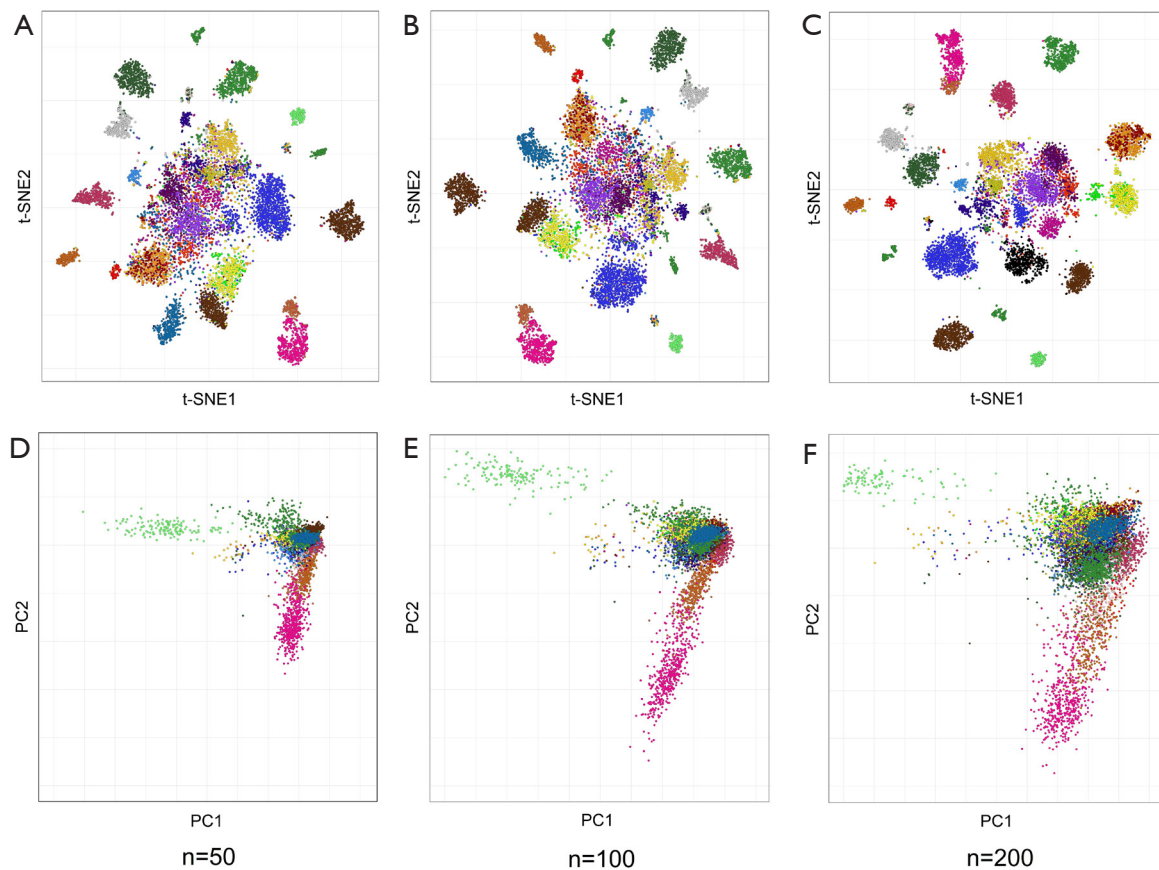


Figure S1 Comparison of dimensionality reduction effect using t-SNE and PCA. (A-C) Visualization using t-SNE by TTSGs size of 50, 100 and 200. (D-F) Visualization using PCA by TTSGs size of 50, 100 and 200. One color represents a tumor type. t-SNE, t-distributed Stochastic Neighbor Embedding; PCA, principal component analysis; TTSGs, tumor tissue-specific gene sets.