

Figure S1. Drop-out rates comparison on Drop-seq and Fluidigm data.

(A-B) Cumulative density plot of the dropout rate among commonly detected genes from STAR and Kallisto. The x-axis represents the dropout rate = 1-(number of cells detected with gene X)/(total number of cells). The y-axis represents the cumulative density.



Figure S2. Comparison of STAR and Kallisto results that had RNA-FISH measurements.

Scatterplots of Gini coefficients for Kallisto vs. STAR for the 26 genes with RNA FISH measurements, on Drop-seq (A) and Fluidigm (B) platforms. Pearson correlation coefficients and p-values are shown.



Figure S3. Comparison of Bowtie to STAR and Kallisto for Drop-seq data

(A) The distribution of the log-transformed gene counts for three methods. (B) Comparison of expressed gene counts between Bowtie and STAR. Pearson correlation = 0.94. (C) Comparison of expressed gene counts between Bowtie and Kallisto. Pearson correlation = 0.81. (D-F) Scatterplots of Gini coefficients for Bowtie vs. FISH, STAR and Kallisto protocols for 26 selected genes. Missing genes from each method are listed below the plots. (G-H) Density plots of the total number of detected genes within each cell for Drop-seq. The x-axis represents the log 10-transformed total estimated gene counts in each sample, and the y-axis represents the density.



Figure S4. The proportion of cell types identified by STARsolo and Kallisto.

(A) The proportion of cell types in the Human PBMC dataset. STARsolo has nine cell types.Kallisto has eight cell types and one unassigned. (B) The proportion of cell types in the mouse cortex 1 dataset. Both STARsolo and Kallisto result in six uniquely clustered cell types, and one endothelial + pericyte mixed cell type.



SUPP FIG. Kallisto only using cDNA as the index

Figure S5. Clustering results on PBMC 3K data, using Kallisto with cDNA index.

(A) Unsupervised clustering only results in seven clusters for Kallisto with the cDNA-built index. Due to the lack of marker gene information, cell type annotation could not be applied to each cluster. (B) Feature plot of known marker genes from Kallisto bus generated results. Note that PPBP and FCER1A marker genes are not detected due to zero expression among all the cells.

Supplementary Table

Table S1. Alignment/Pseudoalignment rate comparison. The percentage of reads recovered

by STAR and Kallisto on four different datasets across three platforms.

Platforms/ Datasets	STAR	Kallisto
Dropseq	62.40%	35.11%
Fluidigm	66.57%	34.03%
10x Human PMBC	89.50%	61.89%
10x Mouse Cortex I	81.61%	74.90%

Table S2. Cell type comparison between STARsolo and Cell Ranger.

Cell labels resulting from Cell Ranger count matrix and STARsolo count matrix are listed. The "Accuracy" column refers to the percentage of each cell type in STARsolo result that is labeled the same as the output from Cell Ranger. The overall accuracy refers to the percentage of "reference" labels among whole cell populations in STARsolo result.

	Matched cell type		Mismatched cell			
Cell type from Cellranger	Cell type from STAR	# cells	Cell type from STAR	# cells	Accurac y	
Naive CD4 T	Naive CD4 T	500	Memory CD4 T	196	96 71.74%	
	Naive CD4 1		CD8 T	1		
CD 14 + Mono	CD 14 + Mono	473	FCGR3A + Mono	5	98.95%	
Memory CD4 T	Momory CD4 T	168	Naive CD4 T	11	-97.10%	
	Memory CD4 1	400	CD8 T	3		
В	В	344	/	/	100.00%	
CD8 T		251	NK	9	92.96%	
	CD8 T		Memory CD4 T	5		
			Naive CD4 T	5		
FCGR3A + Mono	A + Mono FCGR3A + Mono 153		CD 14 + Mono	7	95.68%	
NK	NK	153	CD8 T	1	99.35%	
DC	DC	32	/	/	100.00%	
Platelet	Platelet	12	CD 14 + Mono	1	92.31%	

Overall Accuracy 90.73%

Table S3. Detection of marker genes in single nuclei Mouse Cortex 1 data.

The Checkmark means that the gene is detected. Shaded cell with "X" means the gene is not detected.

	Genes	Slc17a7	Neurod6	Neurod2	Tbr1	Satb2	Cbln2	Slc17a6	
Excitatory Neuron	STAR	\checkmark	\checkmark	V	\checkmark	\checkmark	\checkmark	\checkmark	
	Kallisto	\checkmark							
Inhibitory Neuron	Genes	Gad1	Gad2	DIx1	DIx2	Slc32a1	Erbb4		
	STAR	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark		
	Kallisto	\checkmark	\checkmark	Х	\checkmark	\checkmark	\checkmark		
Astrocytes	Genes	Slc1a3	Aqp4	Gja1	F3	Aldoc	Fgfr3		
	STAR	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark		
	Kallisto	\checkmark	Х	\checkmark	\checkmark	\checkmark	V		
Oligodendrocyte	Genes	Mog	Mbp	Mobp	Plp1				
	STAR	\checkmark	\checkmark	\checkmark	\checkmark				
	Kallisto	\checkmark	Х	\checkmark	\checkmark				
Oligodendrocyte precursor	Genes	Pdgfra	Cspg4	Rlbp1	C1ql1				
	STAR	\checkmark	\checkmark	\checkmark	\checkmark)			
	Kallisto	\checkmark	\checkmark	\checkmark	Х				
Microglia	Genes	Csf1r	C1qa	Aif1	Tmem119	Ctss	C1qb	Tyrobp	Laptm5
	STAR	\checkmark	\checkmark	\checkmark	\checkmark	V	\checkmark	\checkmark	\checkmark
	Kallisto	Х	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	V
Endothelial	Genes	Flt1	ld1	Foxf2	Foxq1	Lef1	Cldn5		
	STAR	\checkmark	\checkmark	V	\checkmark	\checkmark	\checkmark		
	Kallisto	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark		
Pericyte	Genes	Vtn	Pdgfrb	Atp13a5	Rgs5	Kcnj8	Des		
	STAR	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	Х		
	Kallisto	\checkmark	\checkmark	V	\checkmark	\checkmark	Х		