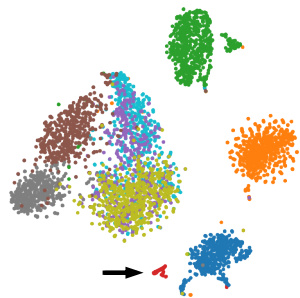


a SeuratV3 + sqrt(CPMedian),
PCA



b Pearson res. ($\theta = 100$),
PCA



c kNN classification in 50 dim.,
macro F1 score

gene selection	log(CPM+1), PCA	log(CPMedian+1), PCA	sqrt(CPMedian), PCA	log(CPMedian+1), PCA	Deviance res. ($\theta = 100$), std., PCA	Pearson res. ($\theta = 100$), PCA	Pearson res. ($\theta = 10$), PCA	GLM-PCA (NB, $\hat{\theta}$ fit)	GLM-PCA (NB, $\theta = 100$)	GLM-PCA (Poisson, $\theta = \infty$)	
	0.69	0.75	0.75	0.72	0.75	0.74	0.75	0.74	0.70	0.71	Deviance res. ($\theta = 100$)
	0.69	0.75	0.75	0.72	0.75	0.75	0.75	0.71	0.70	0.71	sqrt(CPMedian)
	0.57	0.65	0.65	0.67	0.76	0.71	0.75	0.65	0.57	0.56	Seurat
	0.63	0.65	0.66	0.76	0.79	0.79	0.80	0.68	0.61	0.60	SeuratV3
	0.63	0.65	0.66	0.79	0.74	0.79	0.79	0.78	0.65	0.63	Pearson res. ($\theta = 10$)
	0.64	0.72	0.74	0.83	0.74	0.83	0.84	0.77	0.68	0.67	Pearson res. ($\theta = 100$)
	0.70	0.75	0.75	0.84	0.75	0.83	0.84	0.68	0.68	0.66	no gene selection