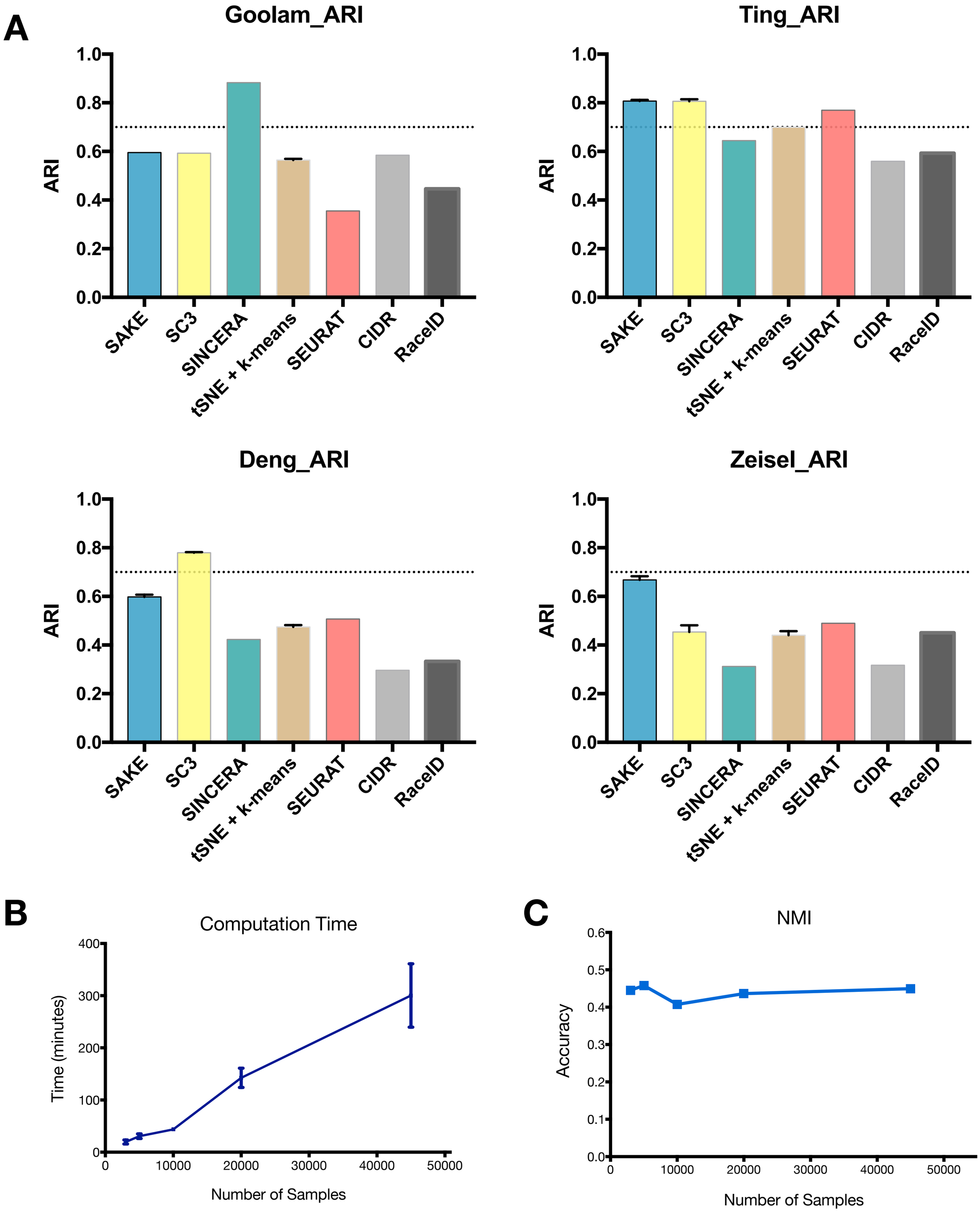

Supplemental Fig. S1 | Robustness for scRNA-Seq software tools/pipelines. A) Cluster memberships from each method were compared to the published cluster assignments from the references, as described in Figure 2. Adjusted Rand Index (ARI) was calculated and used to compare performances across different methods, as an alternate measure to NMI. ARI scores were generally similar to the NMI results given in Fig 2, though ARI can be more sensitive to the size of the clusters identified. In general, no single method outperforms all the other methods across all datasets. We suggest running several different methods to get a consensus result, which is strengthened by the algorithmic independence of each of these methods. **B)** SAKE computation time on published large data sets, where ~68,000 single cells were sorted based on surface markers of known immune cell types, so that the identity of each cell was defined (Zheng et al. 2017). We randomly selected subsets of these 68,000 cells to test the performance of SAKE on data sets of increasing size. The computational time required to run SAKE is given in minutes. **C)** The performance on this dataset is not correlated with dataset size, as measured by NMI scores against the cell identity based on surface markers.