**Table S1: Identifiers of the data used in this study**

|  |  |  |
| --- | --- | --- |
| data | SOURCE | IDENTIFIER |
| BiaseKolodz.YanGoolamPollenDengKleinZeiselRcaTreutleinBaronUsoskin | Biase et al., 2014Kolodziejczyk et al., 2015Yang et al., 2013Goolam et al., 2016Pollen et al., 2014Deng et al., 2014Klein et al., 2015Zeisel et al., 2015Li et al., 2017Treutlein et al., 2014Baron et al., 2016Usoskin et al., 2015 | GEO: GSE57249E-MTAB-2600GEO: GSE36552E-MTAB-3321SRP041736GEO: GSE45719GEO: GSE65525GEO: GSE60361GEO: GSE81861GEO: GSE52583GEO: GSE84133GEO: GSE59739 |

**Table S2: the results of scSO in different ARI threshold**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ARI threshold | scSO | SC3 | Scanpy | Seurat | SIMLR |
| 0.80 | 7 | 3 | 1 | 1 | 1 |
| 0.75 | 9 | 5 | 1 | 1 | 1 |
| 0.70 | 9 | 5 | 4 | 2 | 1 |
| 0.65 | 10 | 5 | 5 | 3 | 2 |
| 0.60 | 10 | 6 | 5 | 4 | 2 |