README: 5' scRNA-seq datasets

<https://doi.org/10.5061/dryad.gtht76hv9> (opens in new window)

Description of the data and file structure

5scCTSSbed\_All.zip: There are 102 files containing count data for analyzing transcription start site (TSS) signals.

md5sum\_CTSS\_files.txt: A text file containing the md5sum of the files included in 5scCTSSbed\_All.zip.

20230928\_TC\_anno\_Gene\_rank\_bed6.bed: This is a bed6 file for robust TSS peaks generated using ReapTEC (with a cutoff of log2CPM ≥ 2 in at least one cell cluster across 136 cell clusters of CD4+ T cells in this study). Robust TSS peaks were named according to the nearest known transcript and numbered in ascending order from upstream to downstream of the transcript.

CD4bulk\_5sc\_TCver\_TCRremoved.removeMono.harmonyintegration\_FigS14.rds: This is a Seurat object used in Supplementary figure S14 (Oguchi et al. Science). The data from 5′ scRNA-seq of bulk CD4+ T cells were integrated using Harmony.

Code/Software

https://github.com/MurakawaLab/ReapTEC