snATAC-seq data processing workflow

The paired end short read sequence data (illumina) was processed using scripting (bash and R) on Iowa State University high performance computing platform in R 4.1.1 or R 4.0.5. Porcine genome reference and gff3 file were downloaded from ensembl 102 and used to generate the config to create a reference package using cellranger-atac mkref function of Cell Ranger ATAC (V.1.2.0) developed by 10x Genomics. Then the base call files (BCLs) were demultiplexed using cellranger-atac function to produce the FASTQ files. For each library, the single cell accessibility counts matrix was generated using the customized reference package by cellranger-atac count command. The nuclei doublets were removed using ArchR (1.0.1)(Granja et al., 2021). The quality control, datasets integration, batch effect correction, normalization, clustering, cell type annotation of snATAC-seq were performed using Signac (1.4.0)(Stuart et al., 2021).

References:

Granja, J. M., Corces, M. R., Pierce, S. E., Bagdatli, S. T., Choudhry, H., Chang, H. Y., et al. (2021). ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. Nat Genet 53, 403–411. doi: 10.1038/s41588-021-00790-6.

Stuart, T., Srivastava, A., Madad, S., Lareau, C. A., and Satija, R. (2021). Single-cell chromatin state analysis with Signac. Nat Methods 18, 1333–1341. doi: 10.1038/s41592-021-01282-5.

The Iowa State University SOP Written by Pengxin Yang