**Template of the Method Section Using scATACpipe**

The single-cell ATAC-seq analysis was performed with scATACpipe (Hu et al., 2022). Preprocessing was performed using (**pick one from the list below**):

1. default:

For sequence quality assessment, the default preprocessing used FastQC (Andrews 2010). After that, barcodes were added to read names with the sinto (Stuart) package. Adapter sequences were removed with Cutadapt (Martin 2011). A custom bash script and SAMtools (Li, Handsaker et al. 2009) were used to prepare a valid reference genome file. After that, read mapping was performed using BWA (Li and Durbin 2009). BAM files were further deduplicated and tagged with sinto and custom Python scripts. Whitelist barcodes were determined using SeqKit (Shen, Le et al. 2016) and custom Python code. Barcodes were further corrected with Pheniqs (Galanti, Shasha et al. 2021) when specified, and valid cell barcodes were determined using a custom R script. Finally, BAM quality control was performed with Qualimap (Okonechnikov, Conesa et al. 2016).

 2. 10xgenomics:

As part of the 10xgenomics preprocessing, the cellranger-atac software suite (Satpathy, Granja et al. 2019) was used, and the valid cells and fragments were obtained with a custom Python script. For the preparation of a valid reference genome and annotation file, SAMtools (Li, Handsaker et al. 2009), gffread (Pertea and Pertea 2020), and a custom Python script were used.

3.    chromap:

In the chromap preprocessing, Chromap software was used (Zhang, Song et al. 2021), and a custom R script was written to retrieve valid barcodes. Cells with valid barcodes were obtained with a Python script. SAMtools (Li, Handsaker et al. 2009), gffread (Pertea and Pertea 2020), and custom Python scripts were used to prepare the reference genome and annotation file.

For downstream analysis, ArchR (Granja, Corces et al. 2021) was primarily used. If "amulet" was selected, the doublet scores were calculated using AMULET (Thibodeau, Eroglu et al. 2021). When starting with fragment files, SAMtools (Li, Handsaker et al. 2009), gffread (Pertea and Pertea 2020), and custom Python script were used to prepare for a valid reference genome and annotation file. Specifically, a custom ArchR::create\_ArchR\_genomeannotaion() and ArchR::create\_ArchR\_geneannotation() or ArchR::create\_ArchR\_geneannotation\_WO\_OrgDb() function were used to create custom ArchR annotation files if user-supplied genomes were not in hg19, hg38, mm9, or mm10.

A customized version of MultiQC (Ewels, Magnusson et al. 2016) was used to create the final report.

**References:**

# Hu, K., et al. (2022) “scATACpipe: A nextflow pipeline for comprehensive and reproducible analyses of single cell ATAC-seq data” Front. Cell Dev. Biol., 27 September 2022

Andrews, S. (2010). "FastQC: A Quality Control Tool for High Throughput Sequence Data." online.

Ewels, P., M. Magnusson, S. Lundin and M. Kaller (2016). "MultiQC: summarize analysis results for multiple tools and samples in a single report." Bioinformatics **32**(19): 3047-3048.

Galanti, L., D. Shasha and K. C. Gunsalus (2021). "Pheniqs 2.0: accurate, high-performance Bayesian decoding and confidence estimation for combinatorial barcode indexing." BMC Bioinformatics **22**(1): 359.

Granja, J. M., M. R. Corces, S. E. Pierce, S. T. Bagdatli, H. Choudhry, H. Y. Chang and W. J. Greenleaf (2021). "ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis." Nat Genet **53**(3): 403-411.

Li, H. and R. Durbin (2009). "Fast and accurate short read alignment with Burrows-Wheeler transform." Bioinformatics **25**(14): 1754-1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing (2009). "The Sequence Alignment/Map format and SAMtools." Bioinformatics **25**(16): 2078-2079.

Martin, M. (2011). "Cutadapt removes adapter sequences from high-throughput sequencing reads." EMBnet.Journal **17(1), 10**.

Okonechnikov, K., A. Conesa and F. Garcia-Alcalde (2016). "Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data." Bioinformatics **32**(2): 292-294.

Pertea, G. and M. Pertea (2020). "GFF Utilities: GffRead and GffCompare." F1000Res **9**.

Satpathy, A. T., J. M. Granja, K. E. Yost, Y. Qi, F. Meschi, G. P. McDermott, B. N. Olsen, M. R. Mumbach, S. E. Pierce, M. R. Corces, P. Shah, J. C. Bell, D. Jhutty, C. M. Nemec, J. Wang, L. Wang, Y. Yin, P. G. Giresi, A. L. S. Chang, G. X. Y. Zheng, W. J. Greenleaf and H. Y. Chang (2019). "Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion." Nat Biotechnol **37**(8): 925-936.

Shen, W., S. Le, Y. Li and F. Hu (2016). "SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation." PLoS One **11**(10): e0163962.

Stuart, T. "Sinto: single-cell analysis tools."

Thibodeau, A., A. Eroglu, C. S. McGinnis, N. Lawlor, D. Nehar-Belaid, R. Kursawe, R. Marches, D. N. Conrad, G. A. Kuchel, Z. J. Gartner, J. Banchereau, M. L. Stitzel, A. E. Cicek and D. Ucar (2021). "AMULET: a novel read count-based method for effective multiplet detection from single nucleus ATAC-seq data." Genome Biol **22**(1): 252.

Zhang, H., L. Song, X. Wang, H. Cheng, C. Wang, C. A. Meyer, T. Liu, M. Tang, S. Aluru, F. Yue, X. S. Liu and H. Li (2021). "Fast alignment and preprocessing of chromatin profiles with Chromap." Nat Commun **12**(1): 6566.