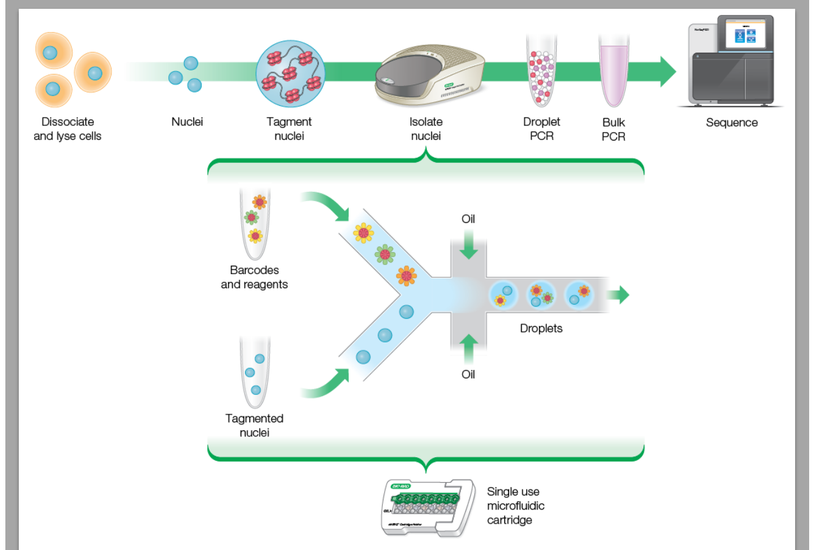
SureCell® ATAC-Seq Library Prep Kit: A New Solution for Single-Cell ATAC-Seq using Bio-Rad's Droplet Digital Technology

Ronald Lebofsky, Zachary Burkett, Michelle Delanoy, Pengchin Chen, Stephen Jones, Marwan Adada, Meiye Wu, Sherman Cheng, Seth Stadick, Jeffrey Bowie, Sankar Narayanaswamy, Felix Acevedo-Ruiz, Jason Moreira, Mark St Mary, Dipika Gurnani, Alanni Grogg, Alfonso Zabala, Adam Corner, Francisco Bizouarn, Sylvie Das Dores, YiFeng Luo, Nakry Sisowath, Teng Shen, Grace Lee, Yansu Li, Alicia Zonis, Phi Tran, Bin Zhang, Preeti Pattamatta, Josh Mopas, Duc Do, Douglas Greiner, Carolyn Reifsnyder

SureCell ATAC-Seq Library Prep Solution

Single-Cell ATAC-Seq (Assay for Transposase-Accessible Chromatin using sequencing) with high-cell throughput provides a novel method for mapping genome-wide chromatin accessibility for thousands of individual single cells. The simple one day workflow provides easy access to epigenetic information including transcriptional dynamics from cell to cell. Four samples can be processed per ddSEQ cartridge per run, providing a scalable solution.



Nuclei isolated from single cells are tagged using Tn5 transposase enzyme at regions of open chromatin in bulk. Amplification from tagged sites within individual droplets generates a library of fragments representative of the original open chromatin on a per cell basis. Barcoding within individual droplets enables the reconstruction of single cell data from the barcode sequences after sequencing.

Bio-Rad ATAC-Seq Analysis Toolkit Pipeline

A comprehensive bioinformatics pipeline has also been developed to facilitate rapid secondary and tertiary analysis of scATAC-Seq data. The bioinformatics pipeline is based on a series of tools performing stepwise analysis including: FASTQ file cleanup, BWA alignment to a reference genome, bead to cell filtering and deconvolution and finally summary statistics. Summary statistics include peak calling, fragments of reads in peaks and transcription start site enrichment.

The data generated by the ATAC-Seq Analysis Toolkit can be used for 3° and 4°analysis of data enabling:

- Estimation of gain and loss of chromatin accessibility
- within peaks
- Clustering of scATAC-Seq profiles by tSNE
- Characterization of sequence motifs associated with gene expression
- Identification of cis- and trans-acting elements that are the source of diverse cellular phenotypes

Dynamic cell range of 100s to 1000s of cells *per sample*, low doublet rates, and highly sensitive.

Dynamic Range data	N=42 low cell samples	N=28 high cell samples
Average number of cells pipetted into cartridge (measured)	537	6947
Average number of cells passing filter *per sample* (NGS data)	397	4408
Lower 95% confidence interval	376	4081
Cell utilization (from cells pipetted into cartridge)	73.9%	63.5%

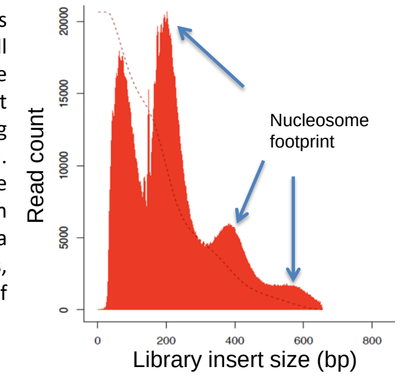
An equal number of mouse NIH-3T3 and human K562 cells were mixed and processed through the SureCell scATAC-Seq workflow. 42 low samples and 28 high samples were run, respectively. Cell throughput is 400 to 4000 with cell utilization ranging from 60-75% or greater. At ~5000 cells throughput, the majority of droplet barcodes were unique to either a single mouse or human cell (below left) Crosstalk, or the percentage of cell doublets for this sample was calculated as 3.76%. Using the 70 sample data set, at 50K reads / cell, >20K unique genomic fragments per cell were detected irrespective of cell throughput (below right)

ATAC-Seq data using fresh GM12878 cells: Peaks, fragments within peaks and transcription start sites.

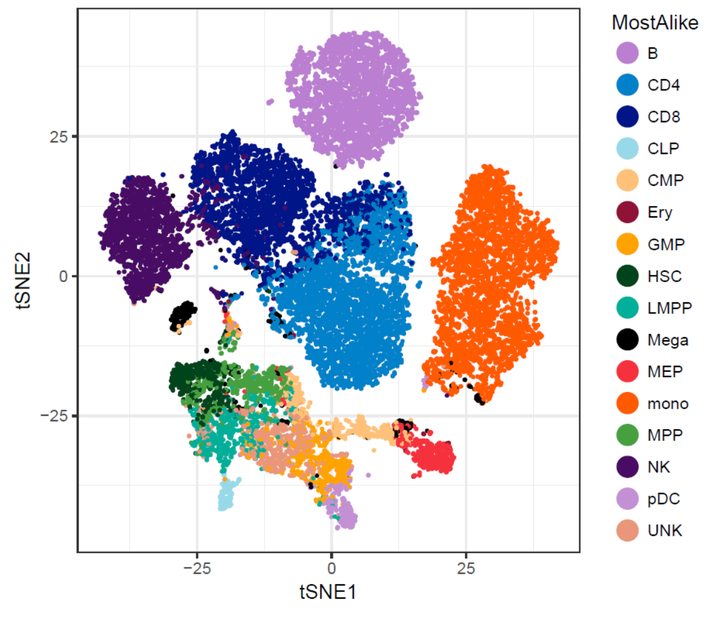
After peak calling and calculation of reads within peaks, scATAC Seq data can be used to generate Transcription Start Site enrichment plots. The top plot is a heat map of the number of unique fragments and their position relative to annotated transcription start sites, denoted as the "0" position, for both mouse and human reference genomes. The colored scale indicates a color code for the number of unique fragments. The total TSS/fragment distribution is summarized in the lower curved graph to form an aggregate distribution of reads centered on the TSS and extended 2 kb in both directions, after normalization. The value at the center of the distribution represents the TSS enrichment score for this sample (>40). These data are indicative of high signal to noise scATAC-Seq data.

ATAC-Seq data from Blood Cells

From a single sample well, frozen PBMCs were processed through the SureCell scATAC-Seq workflow to obtain 5032 single cell data points . The read length fragment distribution exhibits a nucleosome banding pattern that is a hallmark of ATAC data. The first peak represents the nucleosome free fraction. The second, third, and forth peaks represent fragments that span a single, two or three nucleosomes, respectively. The cumulative fraction of reads is plotted with the red dotted line.



30, 000 blood and bone marrow cells were analyzed using SureCell scATAC-Seq. tSNEs were generated and cell data were labeled according to the cell type most correlated with bulk datasets (provided by the Buenrostro lab, Broad-MIT/Harvard). Thus, using Bio-Rad chromatin accessibility data, we can identify and enumerate cell types from complex cellular mixes.



Acknowledgements

F. Duarte, C. Lareau, V. Kartha, J. Buenrostro, Broad Institute, Harvard-MIT

References

1: Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013 Dec;10(12):1213-8. doi: 10.1038/nmeth.2688.

2: Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, Chang HY, Greenleaf WJ. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature. 2015 Jul 23;523(7561):486-90. doi: 10.1038/nature14590.

3: Buenrostro JD, Corces MR, Lareau CA, Wu B, Schep AN, Aryee MJ, Majeti R, Chang HY, Greenleaf WJ. Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation. Cell. 2018 May 31;173(6):1535-1548.e16. doi: 10.1016/j.cell.2018.03.074..

4. Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B, Kathiria A, Cho SW, Mumbach MR, Carter AC, Kasowski M, Orloff LA, Risca VI, Kundaje A, Khavari PA, Montine TJ, Greenleaf WJ, Chang HY. Nat Methods. 2017 Oct;14(10):959-962. doi: 10.1038/nmeth.4396. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues.