Havens sciATAC-seq Summer Project Summary

Spellman Lab

August 25th 2017

Contents

[AIM: 1](#_Toc490229705)

[Perspective on Analysis: 1](#_Toc490229706)

[Project Steps Outline: 1](#_Toc490229707)

[Current state of project: 2](#_Toc490229708)

[Next Steps: 2](#_Toc490229709)

[Selected References: 2](#_Toc490229710)

[Single Cell ATAC: 2](#_Toc490229711)

[Perspectives: 3](#_Toc490229712)

[Tools/Pipelines: 3](#_Toc490229713)

[Other: 3](#_Toc490229714)

AIM:

The overall aim of this experiment is to identify characteristics of chromatin state which are associated with metastatic ability in breast cancer. We do this by comparing chromatin accessibility (measured with sci-ATAC-seq) in primary, low burden, and high burden tumors.

Perspective on Analysis:

sci-ATAC-seq shares many of the changes faced in scRNA-seq but also has some of its own unique issues. Most importantly because we are sequencing DNA from 1 cell, it is only possible to get 0, 1, or 2 unique reads from a single location. This leads to very space data, and traditional analysis tools are not designed to handle this type of data. For example, using MDS, cells with poor quality reads cluster with other cells with poor quality reads, rather than similar cells.

It has been shown in different data sets that aggregating reads from single cells into ‘1 sample’ has similar results to traditional bulk ATAC-seq data. Thus, validating the method of sci-ATAC. Note, aggregating single cells allows us to use analyses similar to those used analyzing bulk data, however, this averages out the heterogeneity that we can see within the single cells. To look at heterogeneity of chromatin accessibility we need different approaches.

Project Steps Outline:

1. Get tumor samples
2. Express and purify Tn5 transposase, and tagment.
3. Sort cells, fragment with Tn5, sort cells, PCR adding index
4. Sequence aggregated cells
5. Analyze and compare accessibility

Current state of project:

Tumor samples have been taken, frozen, and are ready to be processed (this is my understanding, I could be wrong).

Purified the Tn5 protein expressed by bacteria from Adey Lab. We are currently working on verifying protein with mass spec (sent out, waiting on results) and with a western blot (need appropriate antibody). This work is being done with Dmitri Rozanov‎.

Have a preliminary outline for the computational analysis pipeline in ATACnotes (v 1.4) with some supplementary scripts.

Next Steps:

To complete the steps outlined above.

I recommend looking into Cicero, it is not currently available online but is being developed in the Trapnell lab, collaborating with Adey. Cicero will be available on github pending their publication.

Identify cells in sample (eg. tumor vs normal tissue) based on their chromatin state. This may be interesting to explore based on the makers used to distinguish samples in RNAseq.

Selected References:

**Single Cell ATAC:**

Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y. and Greenleaf, W.J., 2015. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature, 523(7561), p.486.

Single cell ATAC (not combinatorial indexing)

Cusanovich, D.A., Daza, R., Adey, A., Pliner, H.A., Christiansen, L., Gunderson, K.L., Steemers, F.J., Trapnell, C. and Shendure, J., 2015. Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. Science, 348(6237), pp.910-914.

Original single cell combinatorial indexing paper, useful supplementary

Cusanovich, D.A., Reddington, J.P., Garfield, D.A., Daza, R., Marco-Ferreres, R., Christiansen, L., Qiu, X., Steemers, F., Trapnell, C., Shendure, J. and Furlong, E.E., 2017. The cis-regulatory dynamics of embryonic development at single cell resolution. bioRxiv, p.166066.

Analysis of fly development using sci-ATAC-seq using genomic windows

Pliner, H., Packer, J., McFaline-Figueroa, J., Cusanovich, D., Daza, R., Srivatsan, S., Qiu, X., Jackson, D., Minkina, A., Adey, A. and Steemers, F., 2017. Chromatin accessibility dynamics of myogenesis at single cell resolution. bioRxiv, p.155473.

Introduction to cicero, uses monocle

**Perspectives:**

Martelotto, L.G., Ng, C.K., Piscuoglio, S., Weigelt, B. and Reis-Filho, J.S., 2014. Breast cancer intra-tumor heterogeneity. Breast Cancer Research, 16(3), p.210.

General perspective on intra-tumor heterogeneity, including scRNAseq and clinical implications

Trapnell, C., 2015. Defining cell types and states with single-cell genomics. Genome research, 25(10), pp.1491-1498.

Perspective on the value and some challenges related to single-cell sequencing

**Tools/Pipelines:**

Ji, Z., Zhou, W. and Ji, H., 2017. Single-cell regulome data analysis by SCRAT. Bioinformatics, p.btx315.

(Web/R GUI) for assigning motifs and dimensionality reduction of peak/motif interactions

Schep, A.N., Wu, B., Buenrostro, J.D. and Greenleaf, W.J., 2017. chromVAR: Inferring transcription factor variation from single-cell epigenomic data. bioRxiv, p.110346.

(R package) for annotated and analyzing motif/peak interactions

Trapnell, C., Cacchiarelli, D. and Qiu, X., 2017. Monocle: Cell counting, differential expression, and trajectory analysis for single-cell RNA-Seq experiments.

(R package) Monocle pseudotime analysis look at <http://cole-trapnell-lab.github.io/monocle-release/> for information on usage

Zhao, C., Huo, X. and Zhang, Y., 2017. Dr. seq2: A Quality Control And Analysis Pipeline For Parallel Single Cell Transcriptome And Epigenome Data. bioRxiv, p.143271.

(Python package) single cell sequencing QC and unsupervised clustering pipeline

Zhicheng Ji and Hongkai Ji. TSCAN: Pseudo-time reconstruction and evaluation in single-cell RNA-seq analysis. (2016) Nucleic Acids Research, 44(13):e117.

(Web/R GUI) for pseudotime analysis and differential expression analysis intended for scRNA

**Other:**

Buenrostro, J.D., Wu, B., Chang, H.Y. and Greenleaf, W.J., 2015. ATAC‐seq: A Method for Assaying Chromatin Accessibility Genome‐Wide. Current protocols in molecular biology, pp.21-29.

Bulk ATAC-seq protocol

Cao, J., Packer, J.S., Ramani, V., Cusanovich, D.A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S.N., Steemers, F.J. and Adey, A., 2017. Comprehensive single cell transcriptional profiling of a multicellular organism by combinatorial indexing. bioRxiv, p.104844.

Compare to sci-RNA-seq

Litzenburger, U.M., Buenrostro, J.D., Wu, B., Shen, Y., Sheffield, N.C., Kathiria, A., Greenleaf, W.J. and Chang, H.Y., 2017. Single-cell epigenomic variability reveals functional cancer heterogeneity. Genome biology, 18(1), p.15.

Uses ‘Single-cell chromatin accessibility’ data (Buenrostro 2015) in analysis of heterogeneity

Picelli, S., Björklund, Å.K., Reinius, B., Sagasser, S., Winberg, G. and Sandberg, R., 2014. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. Genome research, 24(12), pp.2033-2040.

Contains protocol used for Tn5 tagmentation

Rotem, A., Ram, O., Shoresh, N., Sperling, R.A., Goren, A., Weitz, D.A. and Bernstein, B.E., 2015. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nature biotechnology, 33(11), p.1165.

Sc-ChiP-seq as a comparison of methods and expected results

Vitak, S.A., Torkenczy, K.A., Rosenkrantz, J.L., Fields, A.J., Christiansen, L., Wong, M.H., Carbone, L., Steemers, F.J. and Adey, A., 2017. Sequencing thousands of single-cell genomes with combinatorial indexing. Nature Methods, 14(3), pp.302-308.

Suggest looking at the pre-processing scripts from this paper, subclonal tumor analysis from single cell genome (not ATAC) sequencing

Vitak, S.A., Torkenczy, K.A., Rosenkrantz, J.L., Fields, A.J., Christiansen, L., Wong, M.H., Carbone, L., Steemers, F.J. and Adey, A., 2016. Construction of thousands of single cell genome sequencing libraries using combinatorial indexing. bioRxiv, p.065482.

Preprint for above

Zhou, W., Ji, Z. and Ji, H., 2016. Global Prediction of Chromatin Accessibility Using RNA-seq from Small Number of Cells. bioRxiv, p.035816.

Computational work detecting chromatin accessibility in varying sized aggregations of scATAC