OXFORD

Genome analysis

ShinyArchR.UiO: user-friendly,integrative and open-source tool for visualization of single-cell ATAC-seq data using ArchR

Ankush Sharma [1,2,3,4,5,*,†], Akshay Akshay[6], Marie Rogne[3,5] and Ragnhild Eskeland [3,4,5,*]

[1]Department of Biosciences, University of Oslo, 0316 Oslo, Norway, [2]Department of Informatics, Institute of Basic Medical Sciences, University of Oslo, 0316 Oslo, Norway, [3]Department of Molecular Medicine, Institute of Basic Medical Sciences, University of Oslo, 0317 Oslo, Norway, [4]PharmaTox Strategic Research Initiative, Faculty of Mathematics and Natural Sciences, University of Oslo, 0316 Oslo, Norway, [5]Centre for Cancer Cell Reprogramming, Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, 0372 Oslo, Norway and [6]Urology Research Laboratory, Department for BioMedical Research DBMR, University of Bern, 3012 Bern, Switzerland

*To whom correspondence should be addressed.

†Present addresses: Department of Cancer Immunology, Institute for Cancer Research, Oslo University Hospital, Norway and KG Jebsen Centre for B-Cell Malignancies, Institute for Clinical Medicine, University of 0379 Oslo, Norway.

Associate Editor: Can Alkan

Abstract

Motivation: Mapping of chromatin accessibility landscapes in single-cells and the integration with gene expression enables a better understanding of gene regulatory mechanisms defining cell identities and cell-fate determination in development and disease. Generally, raw data generated from single-cell Assay for Transposase-Accessible Chromatin sequencing (scATAC-seq) are deposited in repositories that are generally inaccessible due to lack of in-depth knowledge of computational programming.

Results: We have developed ShinyArchR.UiO, an R-based shiny app, that facilitates scATAC-seq data accessibility and visualization in a user-friendly, interactive and open-source web interface. ShinyArchR.UiO is an application that can streamline collaborative efforts for interpretation of massive chromatin accessibility datasets and allow for open access data sharing for wider audiences.

Availability and implementation: https://Github.com/EskelandLab/ShinyArchRUiO and a demo server with a hematopoietic tutorial dataset https://cancell.medisin.uio.no/ShinyArchR.UiO

Contact: Ankush.Sharma@medisin.uio.no or Ragnhild.Eskeland@medisin.uio.no

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Assay for Transposase-Accessible Chromatin with high throughput sequencing (ATAC-seq) is a powerful method for the assessment of genome-wide chromatin accessibility (Buenrostro *et al.*, 2013) and can be explored at a single-cell resolution (Buenrostro *et al.*, 2015; Cusanovich *et al.*, 2015). A variety of software packages exist, but no uniform analysis has been developed for scATAC-seq data (Baek and Lee, 2020). Analysis of Regulatory Chromatin in R (ArchR) is a comprehensive package for scATAC-seq analysis (Granja *et al.*, 2021) that allows for integration with single-cell RNA-seq using Seurat (Satija *et al.*, 2015; Stuart *et al.*, 2019). Tools for sharing

scATAC-seq data are currently very scarce, and there are no open-source tools based on the analysis performed using the ArchR software toolkit. Here, we present ShinyArchR.UiO, a tool written in R, that adapts preprocessed ArchR data for online visualization and sharing.

2 Workflow and outputs

We have made the source code for ShinyArchR.UiO open to the public and the code can easily be customized for other scATAC-seq applications. The plots generated can be saved as high-quality figures in png, tiff or pdf format.
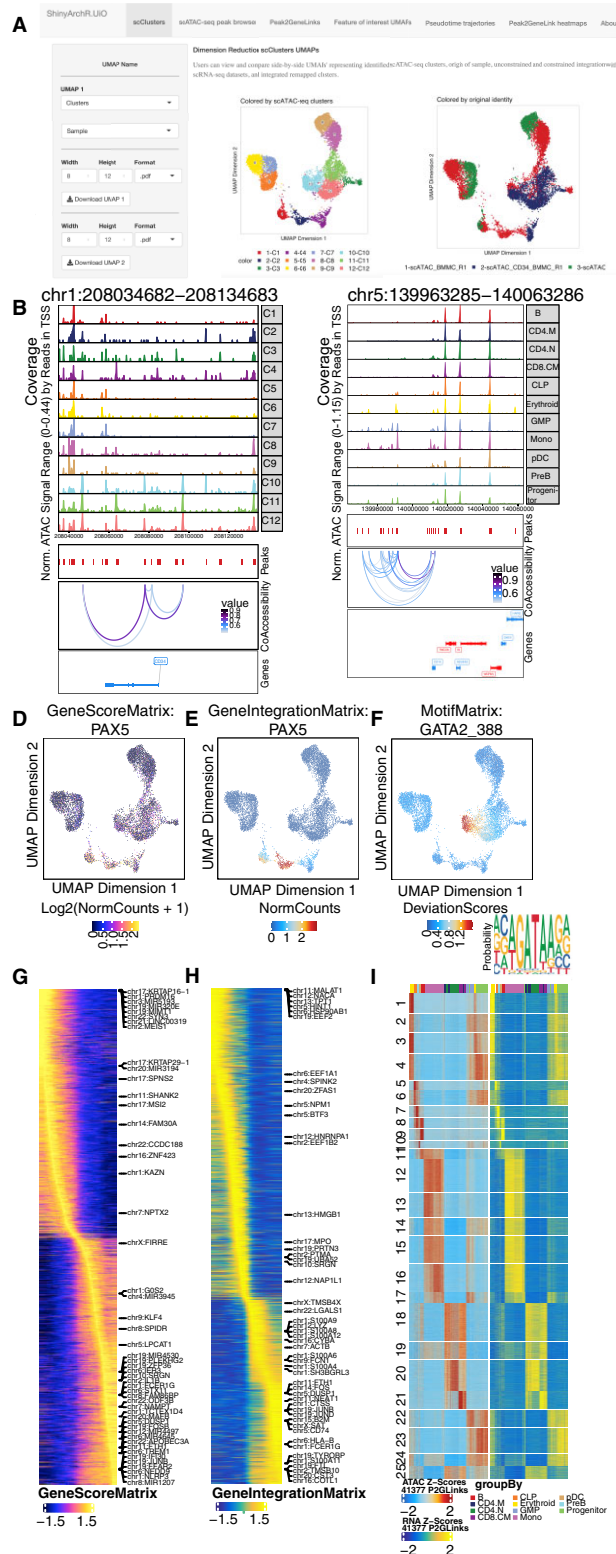
Fig. 1. ShinyArchR.UiO have different web interfaces where the user can compute and plot scATAC-seq data. Users can select and explore: a) Five different UMAPs side-by-side. b) PlotBrowser of scATAC-seq clusters with peaks and co-accessibility. c) Peak2GeneLinks browser with peaks and co-accessibility. Side-by-side feature comparisons of d) GeneScoreMatrix, e)GeneIntegrationMatrix and f) MotifMatrix with motifseq logo and downloadable list of motif positions. Trajectory heatmaps including g) GeneScoreMatrix, and h) GeneIntegrationMatrix. i) Peak2GeneLink heatmaps of gene scores (left) and gene expression (right). Color bars above each heatmap represent the cell clusters as detailed below

3 Visualization of scATAC-seq data

We have used a down sampled tutorial dataset (Granja *et al.*, 2021) to illustrate the features of ShinyArchR.UiO which can be explored on our server (Fig. 1, Supporting Figs S1 and S2).

3.1 Key features of ShinyArchR.UiO

The user can: (i) Perform multidimensional reduction UMAP plot of original samples, scATAC-seq clusters and clusters from cross-platform linkage integration with scRNA-seq generated by ArchR. (ii) Explore plot browser tracks grouped by original samples or scATAC-seq clusters with an adjustable window of 250 KB distance from the center of selected gene of interest with corresponding peaks and coaccessibility. (iii) View coaccessibility of integrated peaks and tracks in constrained remap clusters based on Peak2GeneLink analysis. (iv) Compare single-cell enrichment based on GeneScore, GeneIntegration or Motif Matrix UMAPs for selected genes side-by-side, generate a motif seqlogo for selected transcription factors and download genome-wide list of motif positions. (v) Visualize four different ArchR defined pseudotime trajectory heatmaps. (vi) Explore Peak2GeneLink heatmap on scATAC-seq and scRNA-seq modality.

3.2 Basic usage

ShinyArchR.UiO uses preprocessed ArchR HDF5 formatted objects stored on disk and generates the ShinyArchR.UiO for visualization of the scATAC-seq metadata (Fig. 1). This process can be executed in two steps: (i) Download or git clone https://github.com/EskelandLab/ShinyArchRUiO. (ii) Provide a path to saved arrow file folders (Saved-ArchRproject) obtained from ArchR analysis in *global. R* and name given to trajectory analysis in the getTrajectory function of ArchR, see SI (2) for more details.

Running ShinyArchR.UiO from command line:

```
R -e
"shiny::runApp('~/ShinyArchR.UiO'
,launch.browser = TRUE)"
```

Or open *app. R* or *global. R file*, press *Run App* button on R Graphical User Interface.

3.3 Advanced usage

ShinyArchR.UiO can be run locally or be made available to be hosted on an open-source shiny-server, offering features such as apps behind firewalls (RStudio Team). We have made an open-source R-based tool for visualization of scATAC-seq data including integration with complementary scRNA-seq for a broader range of users than experienced computer scientists. The ShinyArchR.UiO web interface allows the user to easily explore chromatin opening across the genome at a single-cell level for improved insight into complex biological data.

Author contributions

A.S. and A.A. developed the code with scientific input and tests performed by M.R. and R.E. R.E. and A.S. wrote the manuscript with input from M.R.

Funding

Conflict of interest: none declared.

References

Baek,S. and Lee,I. (2020) Single-cell ATAC sequencing analysis: from data preprocessing to hypothesis generation. *Comput. Struct. Biotechnol. J.*, **18**, 1429–1439.

Buenrostro,J.D. *et al.* (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods*, **10**, 1213–1218.

Buenrostro,J.D. *et al.* (2015) Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*, **523**, 486–490.

Cusanovich,D.A. *et al.* (2015) Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science*, **348**, 910–914.

Granja,J.M. *et al.* (2021) ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat. Genet.*, **53**, 403–411.

RStudio Team (2020). RStudio: integrated development for *R*. *RStudio*, PBC, Boston, MA. http://www.rstudio.com/.

Satija,R. *et al.* (2015) Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.*, **33**, 495–502.

Stuart,T. *et al.* (2019) Comprehensive integration of single-cell data. *Cell*, **177**, 1888–1902.e21.