nature research

Corresponding author(s): Eileen Furlong & Jay Shendure

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Using our sci-ATAC-seq protocol for collecting chromatin accessibility data on individual cells at high throughput, we collected data on 384 wells of cells for each of three time points. This resulted in data from ~7,000 cells for each time point. As this was the first time this protocol was applied to Drosophila, it was not possible to a priori estimate the sufficient sample size beforehand. However, we note that this sample size met the criteria of both being experimentally manageable and yet still generating one of the larger single cell data sets to date.

2. Data exclusions

Describe any data exclusions.

There were two reasons for excluding data. (1) We excluded information from cell barcodes that had very few reads associated with them. Looking at the distribution of reads assigned to individual barcodes (Extended Data Figure 1), we noted a bimodal distribution that likely arose from the combination of a normally distributed population of barcodes that represent true cells and a second population near zero that represents background noise barcodes that likely do not represent individual cells. (2) We additionally excluded from analyses the bottom 10% of cells (in terms of feature coverage) to exclude any noisy, low-coverage cells.

3. Replication

Describe whether the experimental findings were reliably reproduced.

Because we collected data on many thousands of cells we have replicate measurements for cells from each major cell group. We do not present the results of replicate experiments.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Samples from each time point were processed separately and sequenced as a single pool. However, within a time point all cells were processed in parallel in a highly randomized fashion.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Researchers were not blind to the identity of the time points. However (as noted above), all cells within a time point were processed in parallel.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☒	☐ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	bcl2fastq v2.16, trimmomatic v0.32, bowtie2 v2.2.3, samtools v1.1, MACS v2.1.0, BEDTools v2.21.0, deepTools v2.6.0. R and python were used for additional data processing and plotting. The scripts necessary for the primary processing of the raw data will be made available on GitHub, upon request.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	No restrictions.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	Mef2 positive cells were stained with a rabbit polyclonal antibody generated at EMBL which was initially described and validated in PMID:16740481 . Elav positive cells were stained with a monoclonal mouse antibody against Drosophila Elav (Elav-9F8A9 from the Developmental Studies Hybridoma Bank) which was initially described in PMID:1716300.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	No eurkaryotic cell lines were used
b. Describe the method of cell line authentication used.	No eurkaryotic cell lines were used
c. Report whether the cell lines were tested for mycoplasma contamination.	No eurkaryotic cell lines were used
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No eurkaryotic cell lines were used

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

Mef2 antibodies were generated from rabbits at EMBL in accordance with European Law and EMBL ethical guidelines. Wild-type (CantonS strain) Drosophila melanogaster were reared and collected at EMBL in accordance with standard practice and the ethical standards of the European Research community. Embryo collections (mixed male and female) were performed at 2-4, 6-8, and 10-12 hours after egg laying.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

This study did not involve human participants.