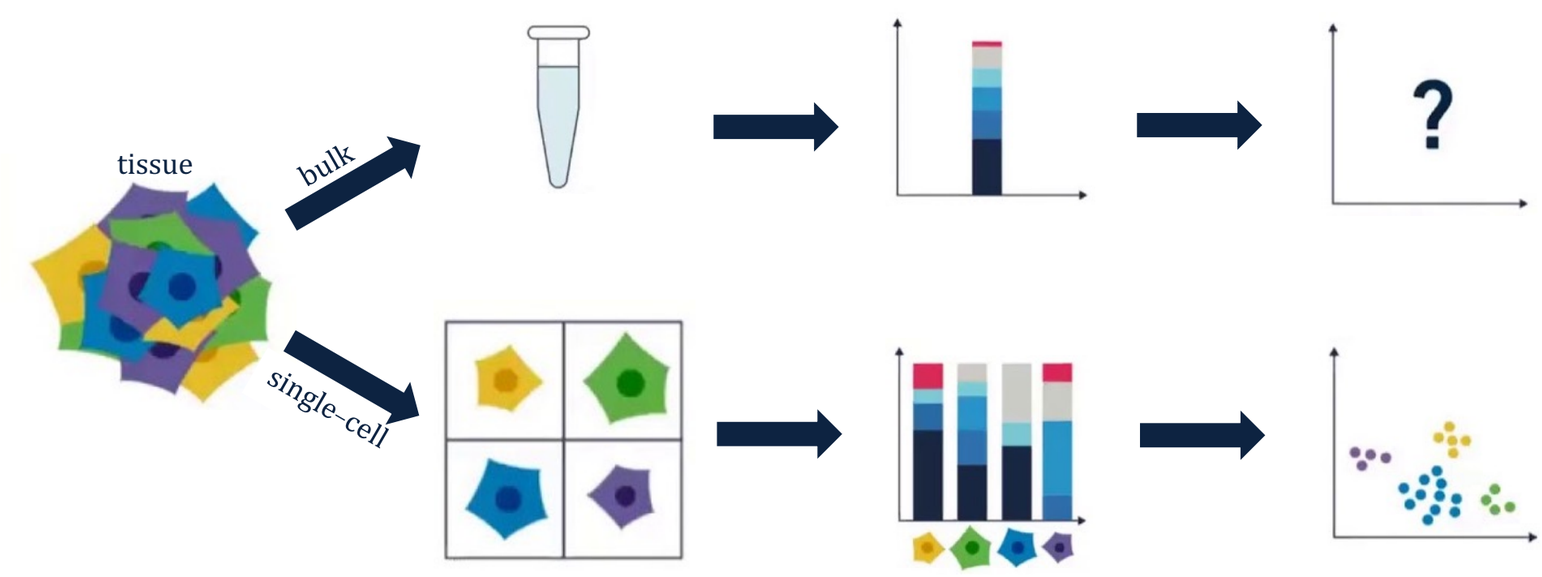
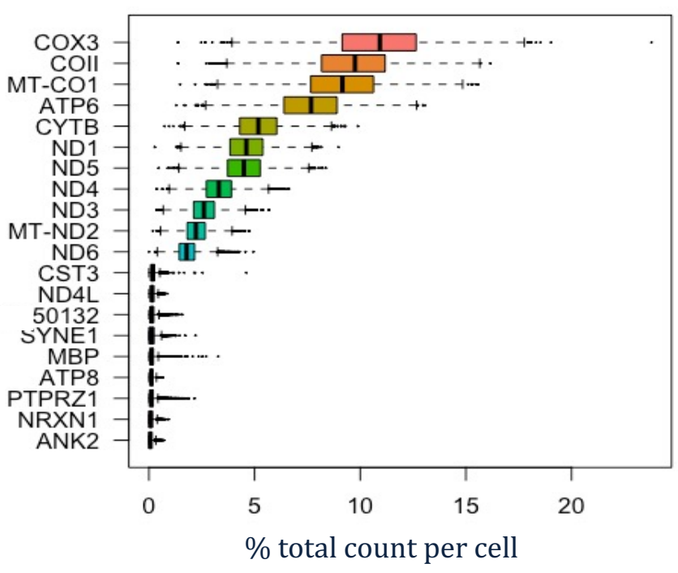
Data preparation and quality control of single–cell data from the hypothalamus in chickens

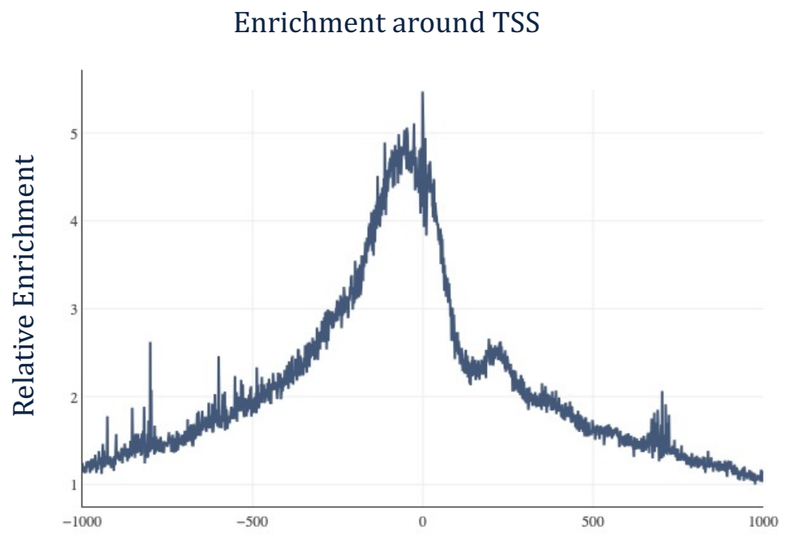
LINKÖPING UNIVERSITY

Introduction

Single–cell sequencing technologies are a recent invention that enable high resolution studies of different types of omics in organisms by generating data per individual cell. I examined data quality and preparation steps of single–cell data originating from hypothalamic tissue from chickens. These datasets will be used to study regulation of transcriptional kinetics.





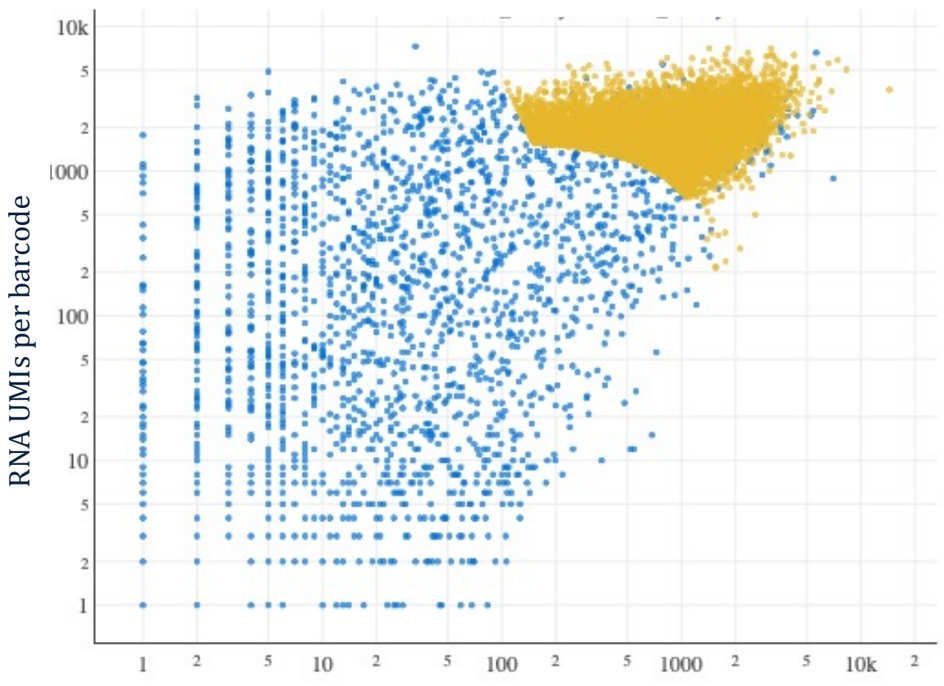


Conclusion

The proportion of median gene count per cell and total gene reflected the heterogeneity of the cells in the hypothalamus. As to data quality, I concluded that most of the initial metrics did not show clear issues. However, some of the cell counts indicated that several samples were overloaded, which created doublets and caused increased levels of ambient RNA.

Methods

For this project the hypothalamus was dissected on embryonic day 19 from 19th and 20th generation chicken embryos from an advanced inter–cross between White Leghorn and red junglefowl. I used the Chromium Next GEM Single Cell Multiome ATAC + Gene Expression protocol from 10x Genomics to create single–cell RNA–seq and ATAC–seq datasets. Cell Ranger ARC was used together with custom R scripts to generate initial summary metrics and insights on ambient RNA.



ATAC transposition events in peaks per barcode

20 000 selected cells

Doortje Theunissen

Supervisor: Dominic Wright