Identifying Meta-Clonal Trajectories in Lineage Tracing Datasets

Arya Kaul, Huidong Chen, Samantha Morris, Luca Pinello

Background

As single-cell profiling evolves, novel lineage tracing technologies that simultaneously track clonal relationships and transcriptional states are becoming increasingly refined. Lineage tracing tags individual cells with unique genetic barcodes to follow their evolution over time by sequencing the cells at frequent intervals. As these techniques grow in accuracy and precision, developing unbiased analytical tools to explore and characterize recurrent trajectory patterns is critical. Here we present the first such tool, Megatron.

Methods

Megatron is an interactive tool to enable users to explore clonal trajectories from single cell data. Currently, 2 distance metrics have been implemented. The first utilizes the Wasserstein metric to compute the distance between each set of clones in each dimension. The second constructs a nearest neighbors graph and computes the ratio of neighbors belonging to the same clone or different clones. Both methods show promise when applied to the CellTagging dataset; however, further work is being done to validate these findings. [1]

Results

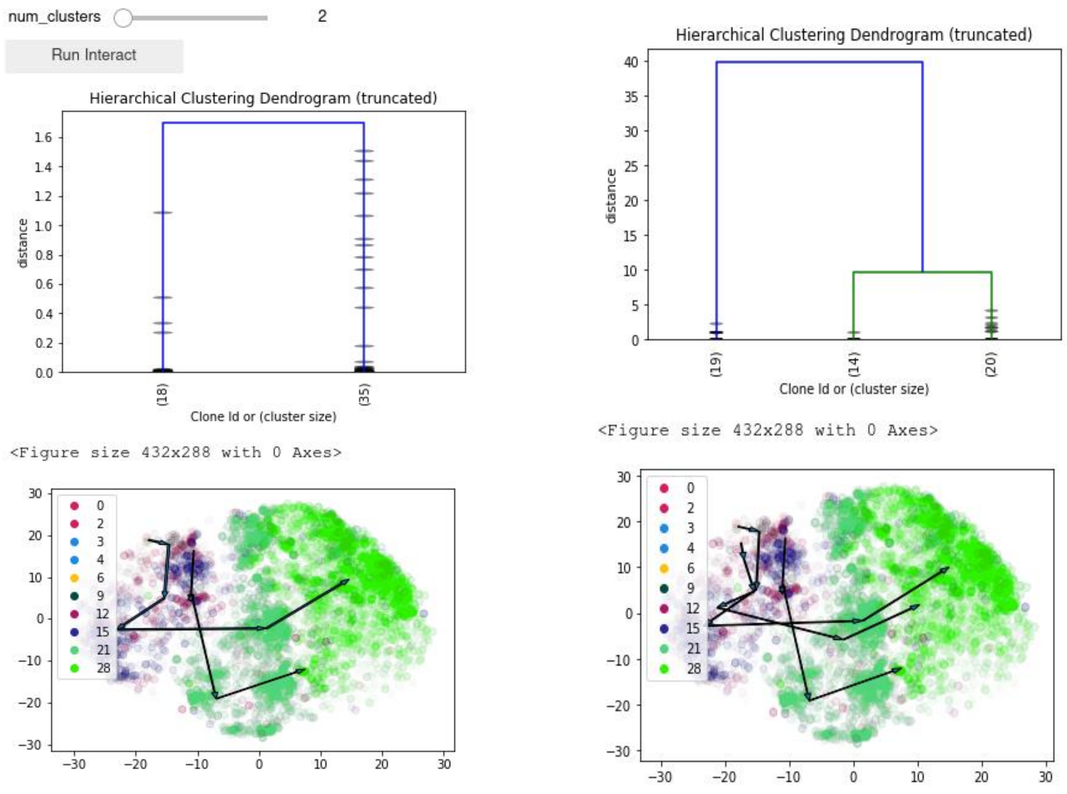

Figure 1 – Megatron analysis of the CellTagging dataset. Clonal trajectories are computed using the Wasserstein derived metric on the left and the nearest-neighbors graph on the right. Dendrogram values in parentheses refer to number of clonal populations in each branch. Bottom visualization is the tSNE representation of each cell in gene expression space with the meta-clonal trajectories overlaid on top.

Significance

Though several lineage tracing technologies have been developed, no computational tool is available to fully explore the clonal trajectories to uncover emerging patterns. Not only is Megatron technology-agnostic, but also our work introduces and showcases the concept and utility of 'meta-clones' i.e. a group of distinct clones that follow similar trajectories in a given time period. We anticipate Megatron will be valued by biological researchers interested in exploring clonal trajectories present in their data and uncovering larger meta-clones that are driving cell fate. We are currently applying the same technique to the Klein ‘20 dataset analyzing hematopoiesis.[2] Future work will investigate which distance metric is most generalizable, which space one can best apply these metrics (UMAP, tSNE, expression), and identify the biological mechanisms driving these changes.

Acknowledgements

- Luca Pinello for advice and guidance
- Huidong Chen for technical help
- Samantha Morris and the Morris lab for datasets and discussion

References

1. Biddy, B. A., Kong, W., Kamimoto, K., Guo, C., Waye, S. E., Sun, T., & Morris, S. A. (2018). Single-cell mapping of lineage and identity in direct reprogramming. Nature, 564(7735), 219–224. https://doi.org/10.1038/s41586-018-0744-4
2. Weinreb, C., Rodriguez-Fraticelli, A., Camargo, F. D., & Klein, A. M. (2020). Lineage tracing on transcriptional landscapes links state to fate during differentiation. Science, 367(6479), eaaw3381. https://doi.org/10.1126/science.aaw3381