Cell Reports Methods

CellPress
OPEN ACCESS

Spotlight

Profiling intermediate cell states in high resolution

Adam L. MacLean[1,*]
[1]Department of Quantitative and Computational Biology, University of Southern California, Los Angeles, CA 90089, USA
*Correspondence: macleana@usc.edu
https://doi.org/10.1016/j.crmeth.2022.100204

Kong et al. present Capybara, a computational method to identify cell states from single-cell gene expression data. Notably, Capybara can identify intermediate cell states and cell state transitions, offering biologists new means with which to interrogate the states and fates of cells.

To understand cell fate is to understand what cells are and where they are going. It is a question of fundamental importance in developmental and cell biology, for it underlies the rules that govern organ growth, responses to pathogens, and cancer. The power we have with which to address these questions of cellular identities has expanded enormously over the last decade, thanks in large part to the rise of single-cell genomics technologies. Yet, the curse of dimensionality looms always nearby. Whereas before, the identity of a cell may have been defined by a few morphological features or a handful of markers measurable by flow cytometry, characterizing the state of a cell from thousands of noisy measurements of noisy gene expression genome-wide requires more sophisticated tools.

Capybara (Kong et al., 2022) is a new tool to ID cells, a giant rodent bouncer, if you will. Established methods are able to identify cell fates when these are well-defined states, such as previously known cell types (Herman et al., 2018; Setty et al., 2019); the task becomes more difficult when cells acquire less well-characterized cell states. These include hybrid or intermediate cell states, which can arise during development or differentiation, or as a result of perturbations, e.g., pathological insults or cellular reprogramming. Intermediate cell states may be transitional, which is characteristic of a dynamic cell fate change, or may represent cells of mixed-lineage state (Figure 1).

Using single-cell gene expression as input, Capybara employs quadratic programming (a method for nonlinear optimization) to assign cell identities sequentially: first to broad tissue-level categories and then to more specific cell types. Capybara builds upon the successes of probabilistic methods for cell type assignment (Zhang et al., 2019) to assign cells not only to distinct cell types but also to hybrid identities. Benchmarking against data generated *in silico* or with clonal barcodes provide support that Capybara is accurately capturing the identities of both well-characterized and hybrid cell states during hematopoiesis.

Cell reprogramming presents greater hurdles for cell identification due to an amplification in the number of cells with intermediate-state identities and even the creation of new "dead-end" cell states (Biddy et al., 2018). In application to fibroblast-to-cardiomyocyte reprogramming, Capybara distinguishes both well-defined cardiac and off-target cell states. Using a method for transition scoring based on the proximity of cells to hybrid states, Capybara identifies higher transition rates in the first two days of reprogramming than at the initial or the final time points. Furthermore, Kong et al. (2022) show that these transition scores are correlated with graph connectivity and RNA velocity, lending support to the notion that they capture local changes in cell state space. It would also be of interest to compare the transition scores as defined here with those based upon entropy and concepts from statistical physics (Teschendorff and Feinberg, 2021). There are limitations to the extent to which dynamic processes can be quantified by single-cell sequencing overall (Weinreb et al., 2018), in part due to the snapshot nature of the data, although new technologies and methodologies may help to overcome some of these challenges (Lange et al., 2022; Qiu et al., 2022).

A crucial assumption underlying the use of quadratic programming to assign cell identities is that the cell state space is continuous. However, it is possible that hybrid cell states can be produced from both continuous and discontinuous transitions in the cell state space (Moris et al., 2016). This becomes particularly relevant in the case of multistability, as predicted by theory (Mojtahedi et al., 2016) and observed experimentally (Schuijers et al., 2015; Wang et al., 2022). Transitions on bistable cell state landscapes can be sharp, discontinuous, and lead to short-term increases in cell state heterogeneity. Whether or not discontinuous cell state transitions violate the assumptions required for quadratic programming likely depends on their magnitude: if small enough, this may not be a barrier to use in practice.

Cell states are defined and maintained by a tug-and-pull of internal and external cell signals. The effects of external signaling on cell fate determination are oft-neglected, yet can exert key control over the processes of fate determination (Rommelfanger and MacLean, 2021). Currently, Capybara does not incorporate the effects of cell-cell interactions into its methods. Recently, the inclusion of dynamic information via pseudotemporal ordering has improved the inference of cell-cell interactions (Li et al., 2022). It will be interesting to see in future work whether the incorporation of cell-cell interaction information can also assist

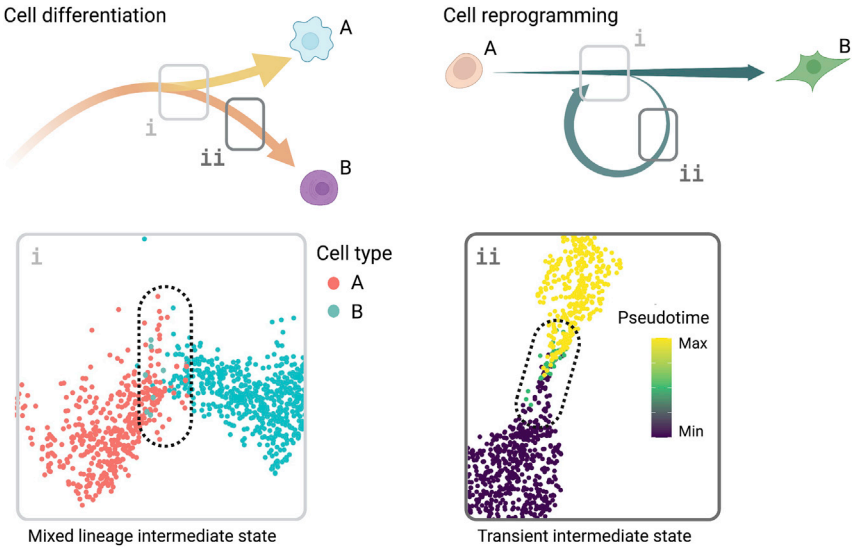

Figure 1. Examples of cell state transitions

Cell differentiation and cell reprogramming give rise to different distributions of cell fates that Capybara can distinguish. In the case of differentiation, two cell fates (A and B) are produced. In the case of reprogramming, cell fate B can be reprogrammed into cell fate A. Under either type of cell state transition, intermediate cell states can arise either as mixed-lineage states or as transient states.

here with the task of cell state identification.

Identifying cell fates and states is a fickle business. In many cases it has transpired that the more data we gather, the less we know with certitude. Kong et al.(2022) offer new means to work with this uncertainty rather than fight against it. The ability to identify hybrid cell states from single-cell data and track them (over time, pseudotime, or experimental conditions) strengthens our understanding of the cell states present in a particular dataset as well as of the processes underlying cell state transitions more generally. This, in turn, helps us move toward a fuller picture of the dynamic identities of cells.

ACKNOWLEDGMENTS

A.L.M. is grateful to Lingyun Xiong for assistance with illustration. This work was supported by the National Institutes of Health (R35GM143019).

DECLARATION OF INTERESTS

The author declares no competing interests.

REFERENCES

Biddy, B.A., Kong, W., Kamimoto, K., Guo, C., Waye, S.E., Sun, T., and Morris, S.A. (2018). Single-cell mapping of lineage and identity in direct reprogramming. Nature *564*, 219–224. https://doi.org/10.1038/s41586-018-0744-4.

Herman, J.S., Sagar, and Grün, D. (2018). FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. Nat. Methods *15*, 379–386. https://doi.org/10.1038/nmeth.4662.

Kong, W., Fu, Y.C., Holloway, E.M., Garipler, G., Yang, X., Mazzoni, E.O., and Morris, S.A. (2022). Capybara: A computational tool to measure cell identity and fate transitions. Cell Stem Cell. https://doi.org/10.1016/j.stem.2022.03.001.

Lange, M., Bergen, V., Klein, M., Setty, M., Reuter, B., Bakhti, M., Lickert, H., Ansari, M., Schniering, J., Schiller, H.B., et al. (2022). CellRank for directed single-cell fate mapping. Nat. Methods *19*, 159–170. https://doi.org/10.1038/s41592-021-01346-6.

Li, D., Velazquez, J.J., Ding, J., Hislop, J., Ebrahimkhani, M.R., and Bar-Joseph, Z. (2022). TraSig: inferring cell-cell interactions from pseudotime ordering of scRNA-Seq data. Genome Biol. *23*, 73. https://doi.org/10.1186/s13059-022-02629-7.

Mojtahedi, M., Skupin, A., Zhou, J., Castaño, I.G., Leong-Quong, R.Y.Y., Chang, H., Trachana, K., Giuliani, A., and Huang, S. (2016). Cell fate Decision as High-Dimensional Critical state transition. Plos Biol. *14*, e2000640. https://doi.org/10.1371/journal.pbio.2000640.

Moris, N., Pina, C., and Arias, A.M. (2016). Transition states and cell fate decisions in epigenetic landscapes. Nat. Rev. Genet. *17*, 693–703. https://doi.org/10.1038/nrg.2016.98.

Qiu, X., Zhang, Y., Martin-Rufino, J.D., Weng, C., Hosseinzadeh, S., Yang, D., Pogson, A.N., Hein, M.Y., Hoi (Joseph) Min, K., Hoi (Joseph) Min, K., Wang, L., Grody, E.I., Shurtleff, M.J., Yuan, R., Xu, S., Ma, Y., Replogle, J.M., Lander, E.S., Darmanis, S., Bahar, I., Sankaran, V.G., Xing, J., and Weissman, J.S. (2022). Mapping transcriptomic vector fields of single cells. Cell *185*, 690–711.e45. e45. https://doi.org/10.1016/j.cell.2021.12.045.

Rommelfanger, M.K., and MacLean, A.L. (2021). A single-cell resolved cell-cell communication model explains lineage commitment in hematopoiesis. Development *148*, dev199779. https://doi.org/10.1242/dev.199779.

Schuijers, J., Junker, J.P., Mokry, M., Hatzis, P., Koo, B.-K., Sasselli, V., van der Flier, L.G., Cuppen, E., van Oudenaarden, A., and Clevers, H. (2015). Ascl2 Acts as an R-spondin/Wnt-Responsive Switch to control Stemness in Intestinal Crypts. Cell Stem Cell *16*, 158–170. https://doi.org/10.1016/j.stem.2014.12.006.

Setty, M., Kiseliovas, V., Levine, J., Gayoso, A., Mazutis, L., and Pe'er, D. (2019). Characterization of cell fate probabilities in single-cell data with Palantir. Nat. Biotechnol. *37*, 451–460. https://doi.org/10.1038/s41587-019-0068-4.

Teschendorff, A.E., and Feinberg, A.P. (2021). Statistical mechanics meets single-cell biology. Nat. Rev. Genet. *22*, 459–476. https://doi.org/10.1038/s41576-021-00341-z.

Wang, W., Poe, D., Yang, Y., Hyatt, T., and Xing, J. (2022). Epithelial-to-mesenchymal transition proceeds through directional destabilization of multidimensional attractor. eLife *11*, e74866. https://doi.org/10.7554/eLife.74866.

Weinreb, C., Wolock, S., Tusi, B.K., Socolovsky, M., and Klein, A.M. (2018). Fundamental limits on dynamic inference from single-cell snapshots. Proc. Natl. Acad. Sci. *115*, 201714723. https://doi.org/10.1073/pnas.1714723115.

Zhang, A.W., O'Flanagan, C., Chavez, E.A., Lim, J.L.P., Ceglia, N., McPherson, A., Wiens, M., Walters, P., Chan, T., Hewitson, B., Lai, D., Mottok, A., Sarkozy, C., Chong, L., Aoki, T., Wang, X., Weng, A.P., McAlpine, J.N., Aparicio, S., Steidl, C., Campbell, K.R., and Shah, S.P. (2019). Probabilistic cell-type assignment of single-cell RNA-seq for tumor microenvironment profiling. Nat. Methods *16*, 1007–1015. https://doi.org/10.1038/s41592-019-0529-1.