Velocity inference of single cells from gene regulatory networks

Daniel A. Ramirez[1,2] and Mingyang Lu[2]

***Short Abstract* — We developed a new computational method that couples regulatory information with dynamical system behaviors to predict cellular state transitions at single-cell resolution based on a set of regulatory interactions. Our approach allows parameter-free and equation-free investigation of state transition trajectories directly from single cell gene expression data due to the fact that a regulator's gene activity precedes its target's gene expression. Using this algorithm, we correctly identified the directionality of single cell state transitions of several simple circuit motifs and identified phenotypic basins of attraction, bifurcations, and oscillatory patterns in single cell gene expression data.**

***Keywords* — Gene regulatory network, single cell transcriptomics, systems biology**

I. Introduction

CELLULAR state transitions and their governing regulatory mechanisms remain one of the fundamental questions in many areas of biological study. Whereas dynamical features of biological systems have only been obtainable from time-series data collection, computational advances now permit us to learn the same information from static snapshots. RNA velocity, for example, achieves this by relating quantities of unspliced and spliced transcripts to infer rates of transcription [1]. We hypothesize that a gene regulatory network (GRN) might provide a similarly exploitable memory effect: if regulator expression precedes target expression, we can use a GRN to infer the likelihood of one gene expression state leading to another.

Here, we develop a method to infer gene expression velocity of single cells based on a GRN with a parameter-free and equation-free approach. Building on a previous study where we investigated state transitions in the cell cycle [2], here we generate vectors at single-cell resolution for several simulated [3] and experimental [4,5] networks.

II. Methods

To compute the likelihood of one cell transiting toward another, we previously developed a metric called cross-cell correlation (CCC)[2]. The CCC from one cellular state to another represents the correlation of regulator activity in the first state with the activity of corresponding targets in the second state. As such, CCC measures the extent to which the latter cellular state represents a reasonable time evolution of the former state, given the topology provided.

In this study, we greatly extend the approach to allow the inference of directionality of gene expression state transitions at single cell level. To achieve that, we first compute, for every cell, the CCC to all cells within a specified distance in gene expression space. Then, we obtain a gene expression vector towards increasing CCC using multiple linear regression. Once a vector is inferred for each cell, we perform vector field smoothing to improve the interpretability of the results.

III. Results and Discussion

We apply the method to several simulated datasets for common circuit motifs, producing vector maps that identify oscillatory, multi-stable, and bifurcating behaviors. By omitting information on certain interactions in the ground truth network, we can also identify the functional role of a specific edge in the overall system behavior. We also demonstrate that CCC accurately captures the cellular state transitions for circuits driven by an external signal.

In multiple experimental single cell gene expression datasets, we show that with a well-validated gene regulatory network, our approach can recapitulate known cellular state transition properties of biological systems. In particular, the algorithm identifies oscillatory behaviors in a cell cycle dataset, bifurcation during hematopoiesis, and a multi-state transition during neurogenesis.

IV. Conclusion

Due to its parameter-free nature, we expect our method to be useful in both understanding fundamental dynamical behaviors of GRN motifs and in gaining insight into regulatory mechanisms of complex biological processes, like those involved in tumorigenesis and stem cell development.

References

[1] La Manno G, et al. (2018) RNA velocity of single cells. *Nature* **560**, 494-498. https://doi.org/10.1038/s41586-018-0414-6.

[2] Katebi A, Kohar V, Lu M (2020) Random Parametric Perturbations of Gene Regulatory Circuit Uncover State Transitions in Cell Cycle. *IScience* **23**, 101150. https://doi.org/10.1016/j.isci.2020.101150.

[3] Huang B, et al. (2017) Interrogating the topological robustness of gene regulatory circuits by randomization. *PLOS Computational Biology* **13**, e1005456. https://doi.org/10.1371/journal.pcbi.1005456

[4] Mojtahedi M et al. (2016) Cell Fate Decision as High-Dimensional Critical State Transition. *PLOS Biology* **14**, e2000640.

[5] Kohar V, Lu M (2018) Role of noise and parametric variation in the dynamics of gene regulatory circuits. *Npj Systems Biology and Applications* **4**, 40. https://doi.org/10.1038/s41540-018-0076-x.

Acknowledgements: This work was funded by NIH grant R35GM128717.

[1]College of Health Solutions, Arizona State University, Tempe, AZ. E-mail: darami14@asu.edu

[2]Department of Bioengineering, Northeastern University, Boston, MA. E-mail: m.lu@northeastern.edu