A lack of distinct cellular identities in single cell data: revisiting Waddington's landscape

Breanne Sparta[1,2], Timothy Hamilton[1,2], and Eric J. Deeds[1,2]

***Short Abstract* — A prevailing interpretation of Waddington's landscape is that cells with similar physiologies exist within a shared basin of attraction and exhibit similar gene expression patterns. To test this hypothesis, we used graph theory to characterize the distribution of cells in epigenetic space. Within a variety of single-cell omics datasets, we find that cell types exist in the same regions of epigenetic space, with a density distribution that is approximately power law. These findings are inconsistent with the idea that cell types are produced by attractors, and encourage us to consider alternative hypotheses for how epigenetic changes maintain physiological stability.**

***Keywords* — Cell types, dynamical attractors, single-cell heterogeneity, scale-free networks, graph theory**

I. Introduction

OUR current understanding for the molecular basis of development rests upon concepts from dynamical systems theory [1]. These concepts are schematized in Waddington's epigenetic landscape: here, cell differentiation proceeds as movement through epigenetic constraints. These constraints are generated by the underlying gene regulatory networks, which produce stable attractors and unstable states, buffer the differentiation trajectory from environmental and intracellular perturbations, and ultimately ensure the production of distinct cell fates. These attractors thus provide a structural explanation for how development can be robust to noise and how the gene expression patterns of a cell can be stable.

The prevailing prediction of Waddington's landscape is that groups of physiologically similar cells should cluster together in gene expression space. This concept is intuitive; cells with similar physiologies should be similar in mRNA levels, protein composition, and have similar functions. Drawing from this assumption, single-cell omics studies almost universally apply clustering algorithms to identify physiologically similar groups of cells, based on similarities in gene expression patterns alone [2]. These studies are often employed to characterize changes in cell-type specific gene expression during development, or in response to external tissue perturbations [2].

While single-cell omics technologies have revealed high levels of heterogeneity in the gene expression states of cells, the structure of transcriptional variation in multicellular tissues and organisms has yet to be characterized directly [3]. Using epigenetic data from a variety of single-cell omics experiments, we characterized the distribution of cells, and evaluated whether cells of similar physiologies exist within locally dense, distinct regions of epigenetic space. To do this, we developed a graph theoretical approach that uses "ε networks", to consider the type and number of cells that are within a certain distance threshold of each cell in the dataset. We find that none of the epigenetic data we analyzed is consistent with the structure predicted by Waddington's landscape.

II. Results

Rather than observing some degree of separation between cell types in epigenetic space, we find that cells of distinct lineages occupy the same region of space. This territorial overlap remained consistent regardless of: the subset of genes used (whether using feature selection methods or supervised approaches), the measurement modality (whether scRNA-seq, scATAC-seq, MERFISH, or CITE-seq, etc), whether transformations were applied (including various normalization techniques and principal component analysis), and regardless of the type of biological sample measured (including post-mitotic brain tissues).

Further, we find that the density distribution of cells is approximately power-law, with most cells existing in low-density regions, very far from other cells. This highly heterogeneous density distribution is unexpected, as it is not consistent with the distributions we would expect to find in the neighborhood of an attractor. Again, this observation is universal in single-cell data on epigenetic state of multicellular organisms, regardless of the tissue, organism, or measurement technique employed.

III. Conclusion

The absence of cell-type clusters in epigenetic space presents both practical and theoretical challenges for the era of single-cell biology. Overall, these findings challenge the idea that cell fates are produced by attractor states of gene regulatory networks, and encourage us to reexamine our assumptions about the molecular basis of cellular differentiation.

References

[1] Goldberg AD, Allis CD, Bernstein E (2007) Epigenetics: A Landscape Takes Shape. *Cell.* **128**, 635-638.

[2] Nguyen QH, et al. (2018) A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Front. Cell Dev. Biol. Med.* **6**.

[3] Buettner F, et al. (2015) Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat. Biotechnol.* **33**, 155-160.

[1]Institute for Quantitative and Computational Biosciences, University of California Los Angeles, E-mail: BSparta@ucla.edu,

[2]Department for Integrated Biology and Physiology, UCLA.