Benchmarking Single-Cell Trajectory Inference Algorithms

Satwik Nimmagadda
Supervised by Dr. Hamim Zafar
IIT Kanpur

Abstract

Understanding the developmental pathways of cell types from common progenitors is crucial in developmental biology and disease research. Single-cell RNA sequencing enables such understanding by measuring gene expression at the level of individual cells. In this project, we evaluated several popular trajectory inference methods that try to reconstruct cell development pathways from the gene expression data. Using four real biological datasets and four evaluation metrics, we compared methods like VIA, Palantir, MARGARET, and Slingshot. Our findings highlight key differences in how each method performs depending on the metrics and datasets.

Introduction

Single-cell RNA sequencing examines gene expression levels at the resolution of individual cells within a sample and enables characterization of distinct cell types and their development pathways. Trajectory inference in scRNA-seq aims to reconstruct the progression of cells through these pathways. A wide range of methods have been developed for this purpose, based on diverse algorithmic approaches such as Markov chains, metric learning, and graph-based techniques like minimum spanning trees. Each method performs differently depending on the dataset's characteristics, such as its underlying trajectory topology.

In this project, we compared several popular trajectory inference methods, including VIA, Palantir, MARGARET, and Slingshot, across four single-cell datasets. The evaluation is based on four metrics (HIM, F1_branches, correlation, featureimp_wcor) that capture different aspects of performance (preservation of graph topology, cell to branch assignments, cell orderings and biological relevance reflected through gene expression). We also analyzed how these metric values vary across the datasets to understand the consistency of the metrics under different biological conditions.

1. Datasets

We evaluate the methods on four real-world biological datasets:

- **Placenta trophoblast differentiation**[10] — A disconnected trajectory with two roots branching into three lineages.
- **Mouse Cell Atlas (MCA)**[10] — Contains two disconnected components, one linear and one multifurcating.
- **Oligodendrocyte differentiation**[11] — A multifurcating topology with one root and six terminal states.
- **Planaria parenchyme differentiation**[12] — A multifurcating topology with four direct branches.

2. Methods

The seven methods evaluated are:

- **VIA**[2]: Uses lazy teleporting random walks on the kNN graph. It handles complex topologies including disconnected and cyclic graphs.
- **Margaret**[3]: Uses deep metric learning to construct an embedding and finds connectivity strengths between clusters.
- **Palantir**[4]: Combines diffusion maps with an absorbing Markov chain to compute pseudotime and fate probabilities.
- **PAGA**[5]: Builds a cluster-level abstraction of the kNN graph, preserving coarse global topology.
- **Slingshot**[6]: Fits principal curves on an MST over clusters, handles tree-like trajectories.
- **DPT**[7]: Computes pseudotime using diffusion distances, assumes a bifurcating topology.
- **SCORPIUS**[8]: Uses multidimensional scaling and principal curves, assumes a linear progression.

3. Evaluation Metrics

We use four metrics to evaluate performance:

- **HIM Distance**: Combines Hamming and Ipsen–Mikhailov distances to measure topological similarity to ground truth.

- **F1_branches**: F1 score on branch assignment, using precision and recall of cell to branch assignments.
- **Correlation**: Pearson correlation between inferred and true geodesic distances.
- **FeatureImp_wcor**: Weighted correlation of feature importance scores, reflecting gene relevance recovery.

All metrics range from 0 to 1, with higher values indicating better performance.

Results

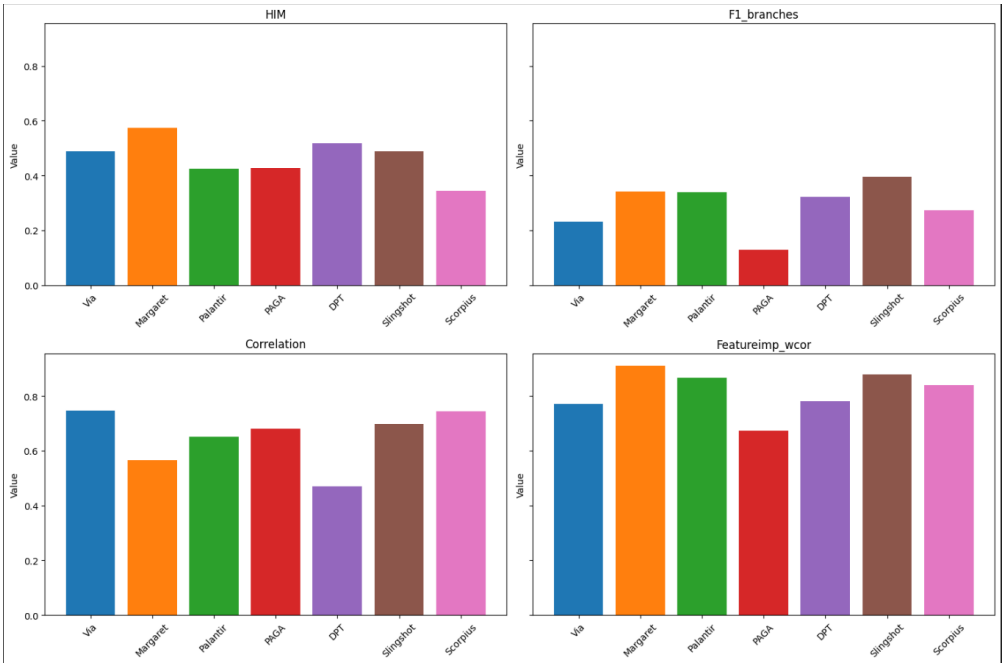

Figure 1: Metric scores for each method (averaged across all datasets)

1. **Metric-wise Performance of Methods:.**
   - **Top methods in each aspect:**
     - Topology (HIM): MARGARET
     - Branch assignment (F1_branches): Slingshot
     - Cell ordering (Correlation): VIA
     - Biological relevance (FeatureImp_wcor): MARGARET
   - All metrics highlight different top performers, indicating that no single metric alone is sufficient and a combined overall score is needed.

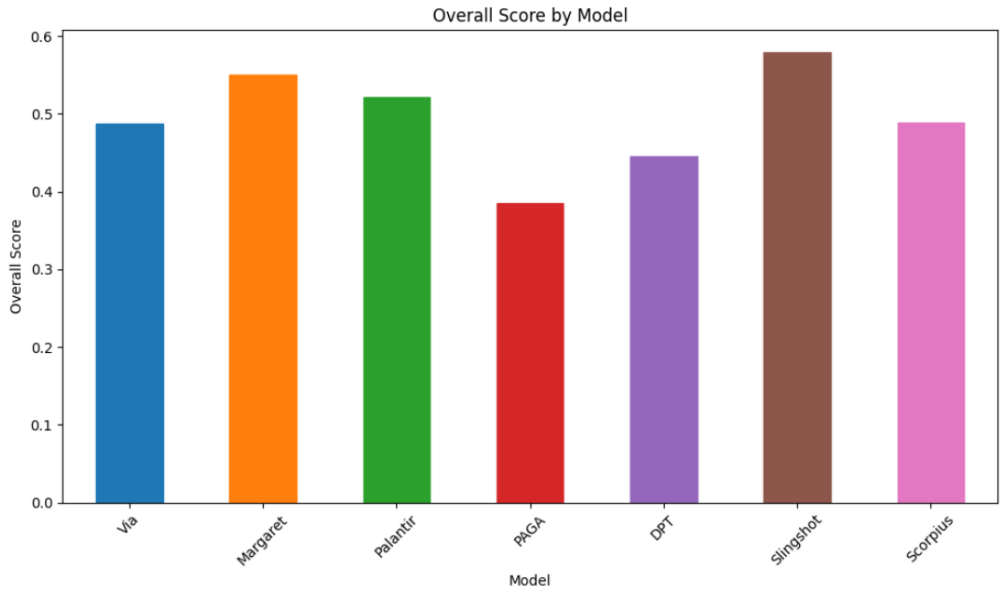

Figure 2: Overall score for each method (geometric mean of HIM, F1 branches, Correlation, FeatureImp wcor), averaged across all datasets.

2. Overall Score Performance of Methods.

$$\text{Overall score}_m \;=\; \left(\text{HIM}_m \times \text{F1\_branches}_m \times \text{Correlation}_m \times \text{FeatureImp\_wcor}_m\right)^{1/4}$$

Method	**Overall Score**
Slingshot	0.58
MARGARET	0.55
Palantir	0.52
VIA	0.49
SCORPIUS	0.48
DPT	0.44
PAGA	0.38

Table 1: Overall scores of the methods (averaged across all datasets).

3. Average Overall Score per Dataset.

- No single dataset consistently outperforms the others across all metrics.
- This indicates the benchmarking is *methods-agnostic*, with no systematic bias towards the selection of the methods.

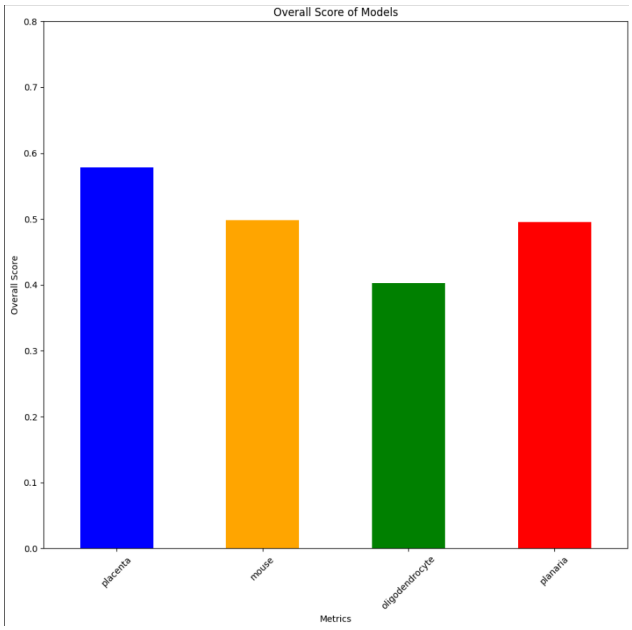

Figure 3: Average overall score (arithmetic mean of method-wise overall scores) for each dataset

Conclusion

- No single method dominates across all datasets and metrics.
- **Slingshot**, **MARGARET**, and **Palantir** are the top performers when averaged across datasets.
- The four metrics are not strongly correlated and capture distinct aspects of trajectory inference—topology, branch accuracy, cell ordering, and biological relevance.
- Benchmarking across diverse datasets confirms the generalizability of conclusions and highlights the need to match methods to dataset characteristics.
- **Future work:** Develop a consensus framework that integrates the strengths of multiple methods for more robust and accurate trajectory reconstruction.

References

[1] Saelens, W., Cannoodt, R., Todorov, H. *et al.* A comparison of single-cell trajectory inference methods. *Nat. Biotechnol.* **37**, 547–554 (2019).

[2] Stassen, S. V., Yip, G. G. K., Wong, K. K. Y. *et al.* Generalized and scalable trajectory inference in single-cell omics data with VIA. *Nat. Commun.* **12**, 5528 (2021).

[3] Pandey, K., Zafar, H. Inference of cell state transitions and cell fate plasticity from single-cell with MARGARET. *Nucleic Acids Res.* **50**, e86 (2022).

[4] Setty, M., Kiseliovas, V., Levine, J. *et al.* Characterization of cell fate probabilities in single-cell data with Palantir. *Nat. Biotechnol.* **37**, 451–460 (2019).

[5] Wolf, F. A., Hamey, F. K., Plass, M. *et al.* PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 59 (2019).

[6] Street, K., Risso, D., Fletcher, R. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).

[7] Laleh Haghverdi, Maren Büttner, F. Alexander Wolf, Florian Buettner, Fabian J. Theis. Diffusion pseudotime robustly reconstructs lineage branching. bioRxiv 041384 (2016).

[8] Robrecht Cannoodt, Wouter Saelens, Dorine Sichien, Simon Tavernier, Sophie Janssens, Martin Guilliams, Bart Lambrecht, Katleen De Preter, Yvan Saeys. SCORPIUS improves trajectory inference and identifies novel modules in dendritic cell development. bioRxiv 079509 (2016).

[9] Cannoodt, R., Saelens, W., Sichien, D. *et al.* Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics. *Science* **360**, eaaq1723 (2018).

[10] Han, X., Wang, R., Zhou, Y. *et al.* Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell* **172**, 1091–1107.e17 (2018).

[11] Marques, S., Zeisel, A., Codeluppi, S. *et al.* Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system. *Science* **352**, 1326–1329 (2016).

[12] Plass, M., Solana, J., Wolf, F. A. *et al.* Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics. *Science* **360**, eaaq1723 (2018).