scRNAseq and Optimal Transport

February the Twelfth, 2025

nature biotechnology

Article

https://doi.org/10.1038/s41587-024-02186-3

Gene trajectory inference for single-cell data by optimal transport metrics

Received: 19 December 2022

Accepted: 26 February 2024

Published online: 05 April 2024

Tublished billine. 05 April 202

Check for updates

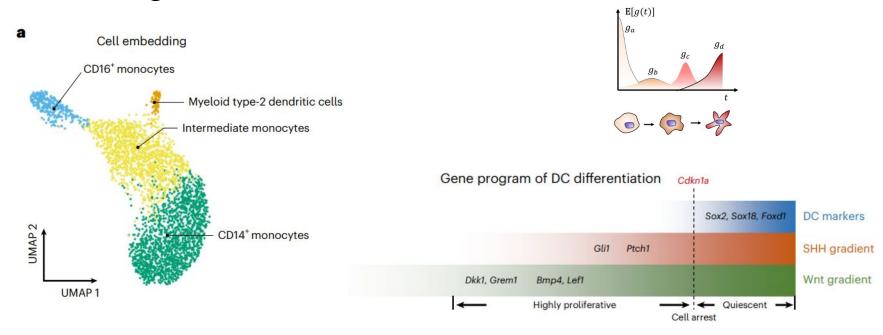
Rihao Qu^{1,2,3,11}, Xiuyuan Cheng^{4,11}, Esen Sefik ® ³, Jay S. Stanley III⁵, Boris Landa⁵, Francesco Strino ® ⁶, Sarah Platt^{2,7}, James Garritano⁵, Ian D. Odell^{3,7}, Ronald Coifman^{5,8,9}, Richard A. Flavell ® ³,10,12</sup>, Peggy Myung ® ²,7,12 & Yuval Kluger ® ¹,2,5,12 ⊠

Why Should I Care?

- Practical Use-Case of OT
- Can be Generalised to Other Data Types
- Interplay between Dynamical Systems and Data Analysis

(Case Study in Examining Shapes of Data)

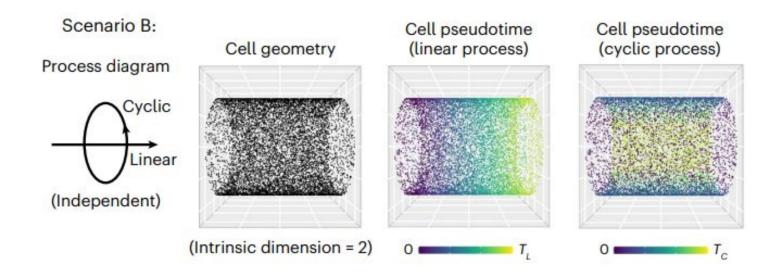
Uncovering the "Idea" in a dataset.



A simple linear process

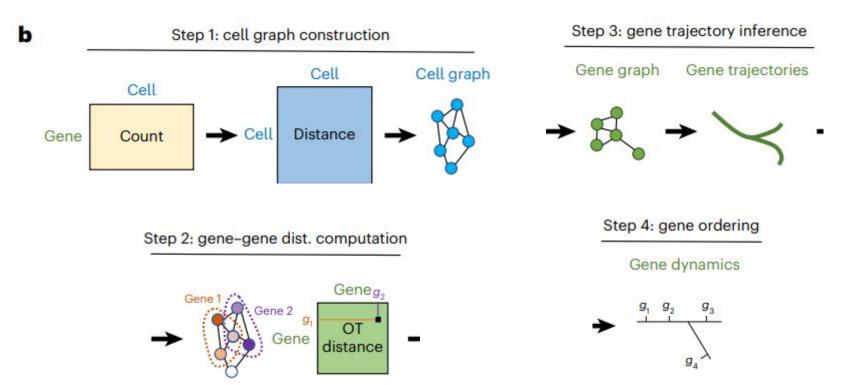
What can modelling a dataset actually *tell us* about the world (as opposed to simulating it)

The Geometry of Dynamical Processes

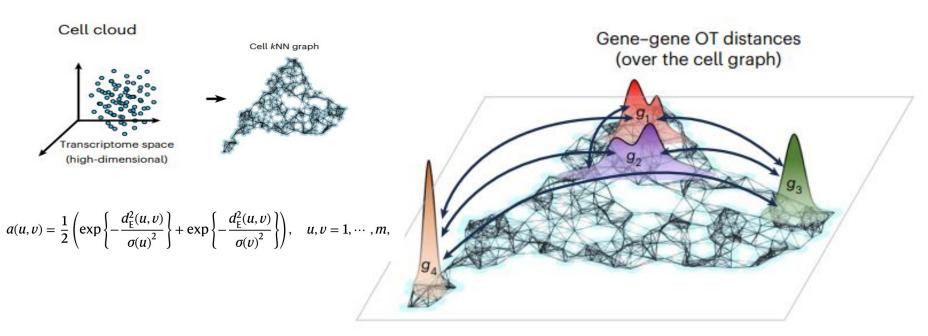


Datasets are shaped by the dynamics that generate them. How can data science reconstruct those dynamics?

Overview



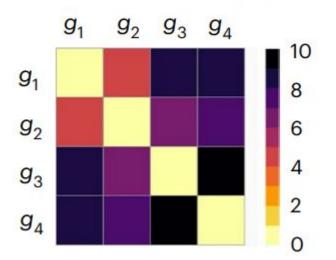
Summarising Features with OT



Summarising Features with OT

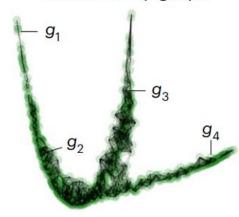
e

Gene-gene OT distance matr (submatrix example)



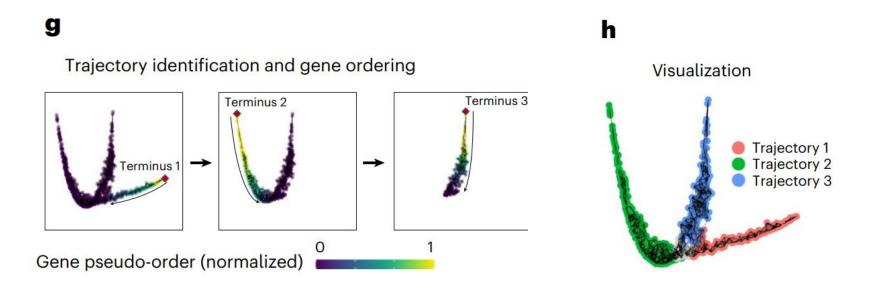


Gene affinity graph



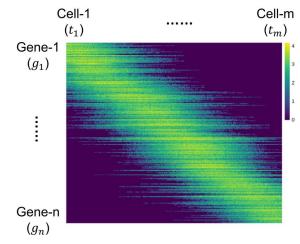
$$A_{i,j} = \frac{1}{2} \left(\exp \left\{ -\frac{\delta^{(p)}(\rho_i, \rho_j)^2}{(\sigma_i)^2} \right\} + \exp \left\{ -\frac{\delta^{(p)}(\rho_i, \rho_j)^2}{(\sigma_j)^2} \right\} \right)$$

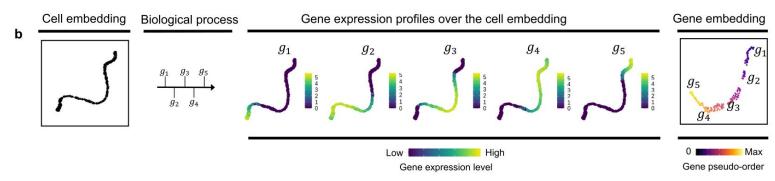
Identifying Gene Programs From OT Geometry



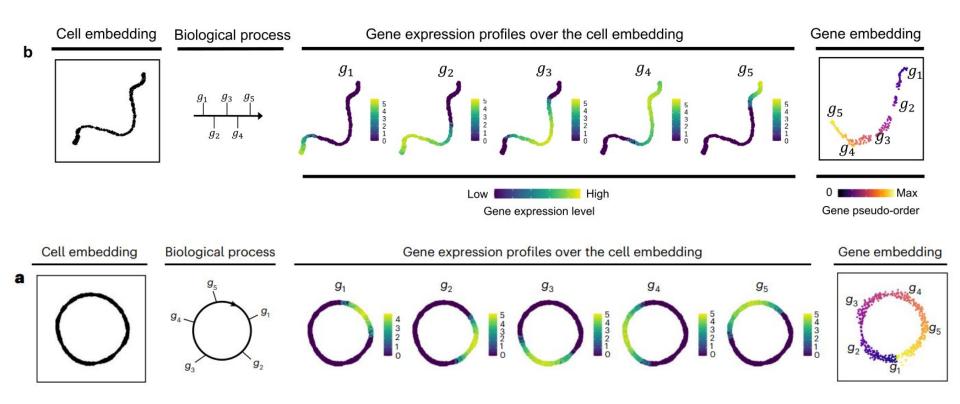
Termini were identified by maximal distance from origin in spectral embedding

A Simplified Differentiation Dataset



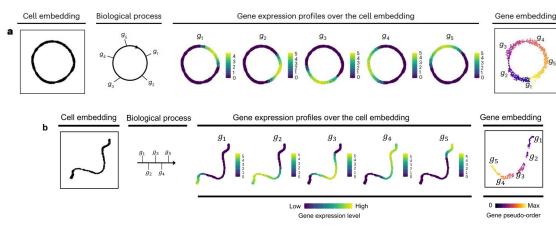


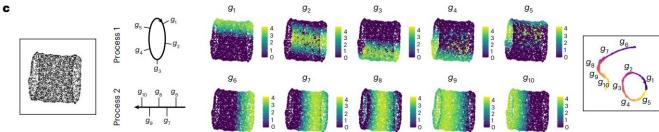
Gene OT Networks for Simple Datasets



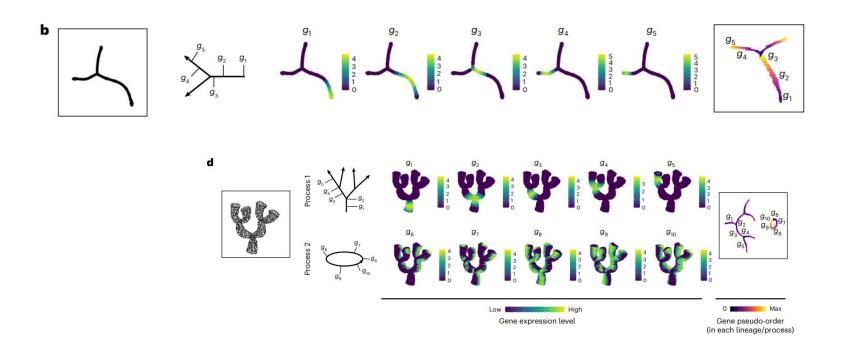
Analysing Multiple Processes

The gene OT network has different components/communities relating to different processes.

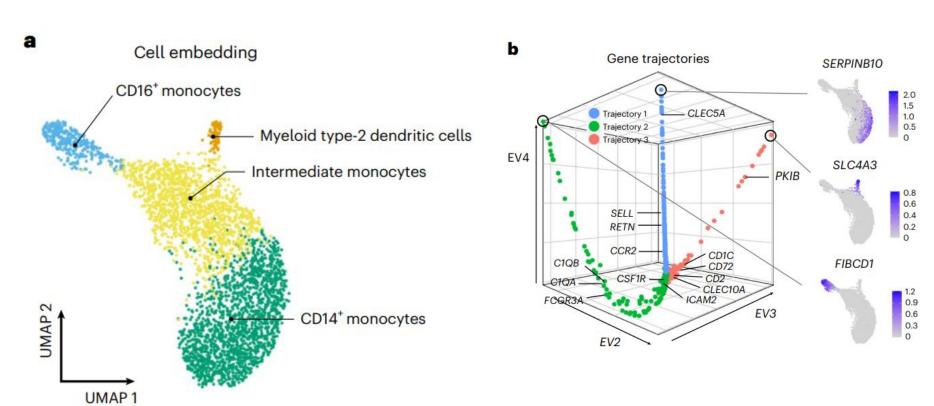




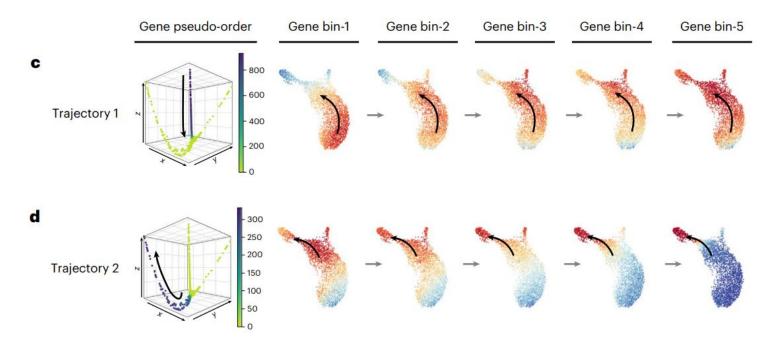
Same Principle, Branched Differentiation



Identifying Gene Programs From OT Geometry

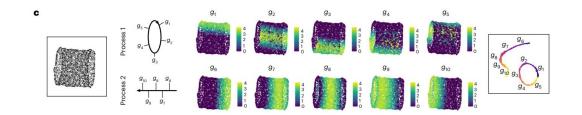


Identifying Gene Programs From OT Geometry

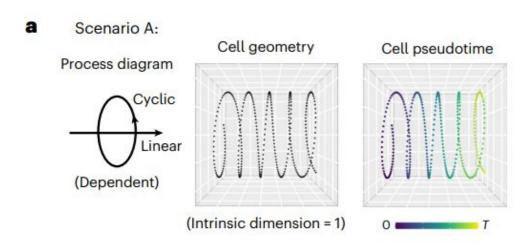


Visualise the Gene Trajectories on UMAP

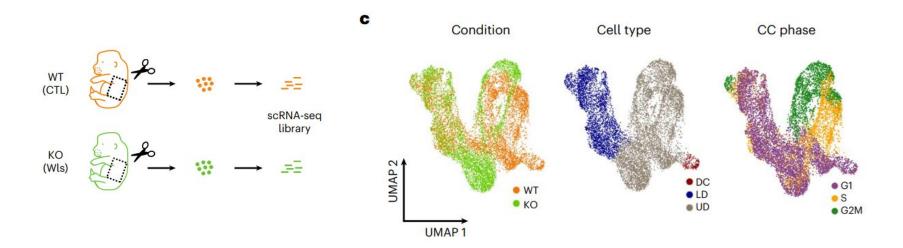
A Problem with the Method



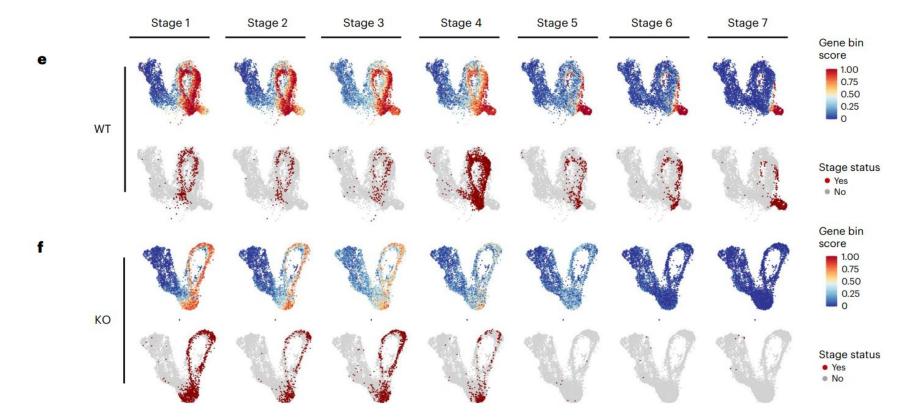
But What if...



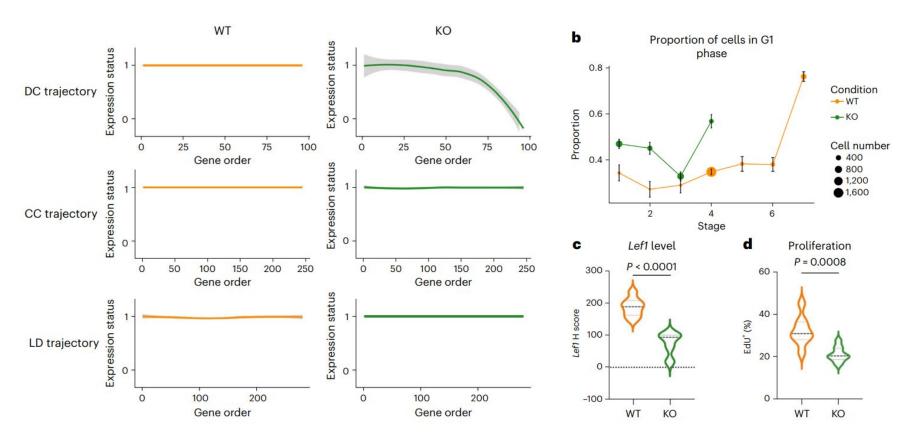
Differentiation of Dermal Condensates



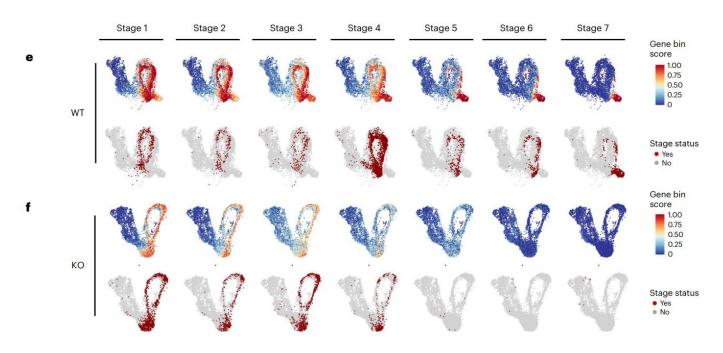
Differentiation of Dermal Condensates



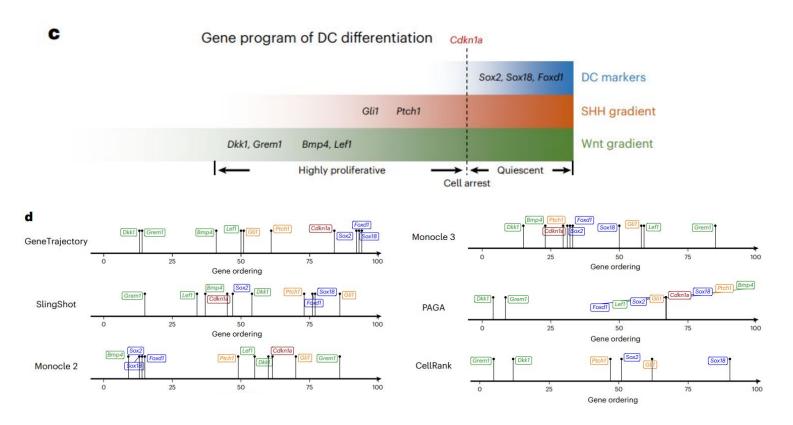
Using Gene Trajectories to Decipher Mechanisms



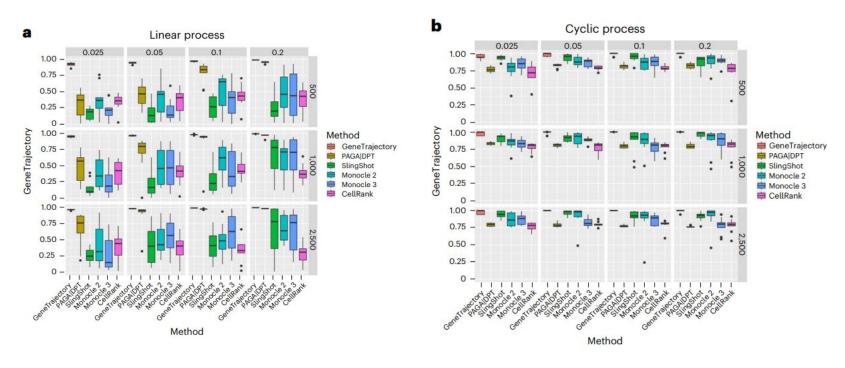
Using Gene Trajectories to Decipher Mechanisms



Comparison to Other Methods



Comparison to Other Methods

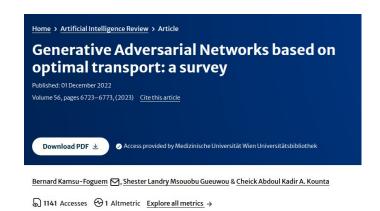


How good was the recovery of the order in which genes were expressed?

What's Next?

Optimal Transport for Generative Models Disentangled Recurrent Wasserstein Autoencoder (ICLR 2021) Wasserstein-2 Generative Networks (ICLR 2021) Learning Generative Models across Incomparable Spaces (ICML 2019) | Code Sliced-Wasserstein Flows: Nonparametric Generative Modeling via Optimal Transport and Diffusions (ICML 2019) | Code Sliced Wasserstein Generative Models (CVPR 2019) OT-GAN: Improving GANs Using Optimal Transport (ICLR 2018) | Code Learning Generative Models with Sinkhorn Divergences (AISTATS 2018) WGAN-GP: Improved Training of Wasserstein GANs (NeurIPS 2017) WGAN: Wasserstein Generative Adversarial Networks (NeurIPS 2017)

https://github.com/kilianFatras/awesome-optima <u>I-transport?tab=readme-ov-file#generative-models</u>



https://link.springer.com/article/ 10.1007/s10462-022-10342-x?u tm_source=chatgpt.com