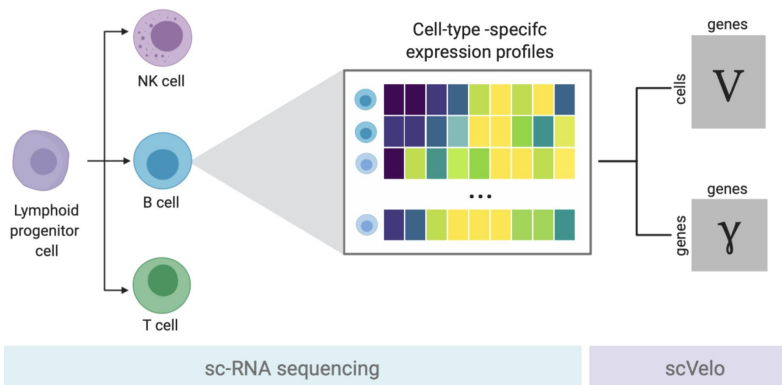


scMomentum: Inference of Cell-Type-Specific Regulatory Networks and Developmental Landscapes

Larisa Morales-Soto^{1,*}, Juan P. Bernal-Tamayo¹, Robert Lehmann¹, Subash Balsamy¹, Xabier Martinez-de-Morentin², Amaia Vilas-Zornoza³, Patxi San-Martin³, Felipe Prosper⁴, David Gomez-Cabrero^{1,2,5}, Narsis Kiani⁶, Jesper Tegner^{1,*}

Introduction

How to infer Gene Regulatory Networks (GRNs) and their changes during cell differentiation is still fundamentally unresolved. Here, we present **scMomentum**, a **data-driven** approach for **cell-type-specific GRN inference** that uses RNA expression, velocity to infer **strength, direction, and effect** for each regulatory interaction. scMomentum then uses those GRNs to reconstruct developmental landscapes from **single-cell RNA-seq data**.

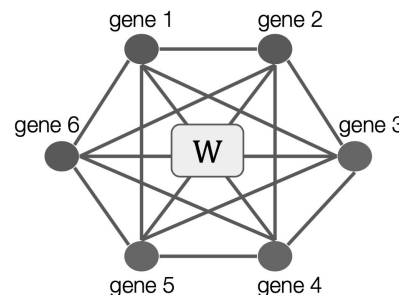


Methods

GRN inference - using a linear system that links the effect of a set of informative genes within each cell type, as follows:

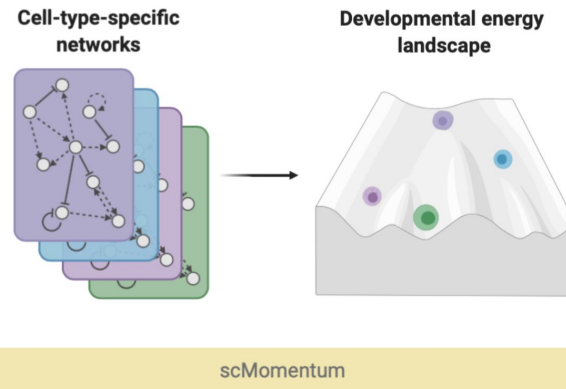
$$W = X^+V + \gamma$$

This matrix can be interpreted as a neural network or **Hopfield network**.



Estimations of the **energy** of the system used to reconstruct the **developmental landscape** are then obtained following Hopfield's method [1].

$$H(S) = -\frac{1}{2} \sum_{i=1}^N \sum_{j \neq i}^N S_i W_{ij} S_j$$

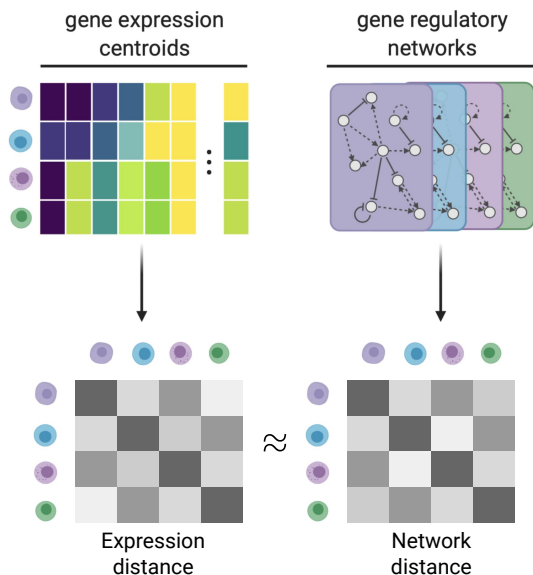


¹ Biological and Environmental Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal, **Saudi Arabia**. ² Mucosal and Salivary Biology Division, King's College London Dental Institute, London, **United Kingdom**. ³ Navarrabiomed, Complejo Hospitalario de Navarra, Universidad Pública de Navarra, Navarra, **Spain**. ⁴ Centro de Investigación Médica Aplicada, Navarra, **Spain**. ⁵ Clínica Universidad de Navarra, Navarra, **Spain**. ⁶ Department of Oncology and pathology, Karolinska Institute, Stockholm, **Sweden**.

* Corresponding authors: jesper.tegner@kaust.edu.sa, larisa.soto@kaust.edu.sa

Methods

Preservation of cluster distances



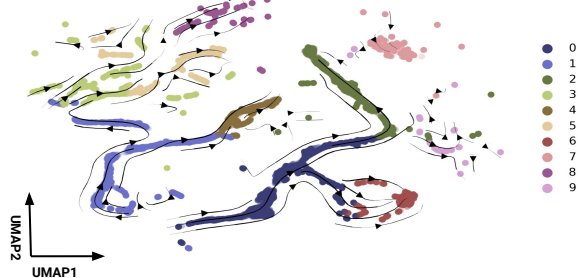
Targeted network perturbations

Removing genes with important network properties (largest eigenvalues) should have a large effect on several network properties such as distance preservation.

Results

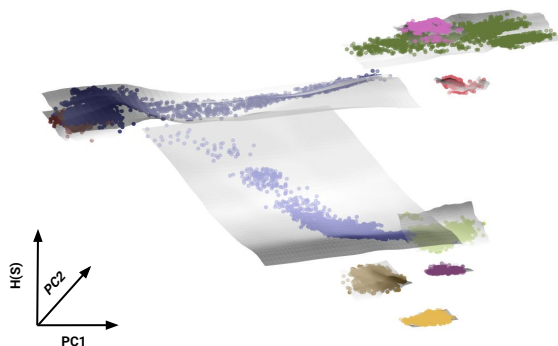
1

We used simulated data (20K cells, 500 genes) generated with DYNGEN [2] and obtained the estimations of RNA velocity and kinetic rates from scVelo [3].



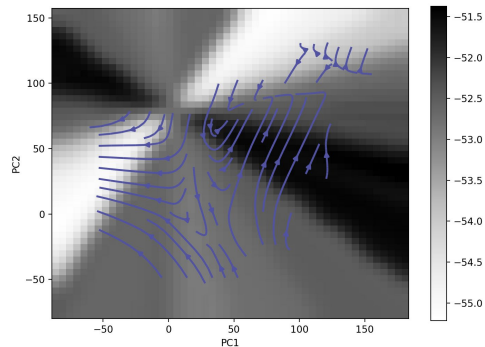
2

The energy landscapes derived from our networks successfully recapitulate the expected cellular progressions.



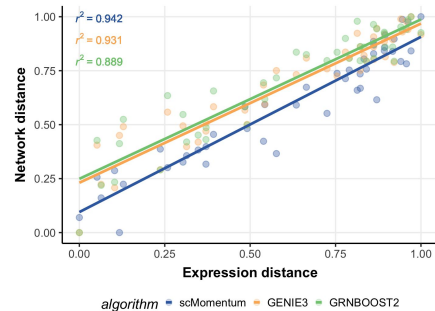
3

Cell velocities follow the expected trajectory in the landscape.



4

scMomentum preserves cluster distances better than two methods [4,5] in a simulated dataset.



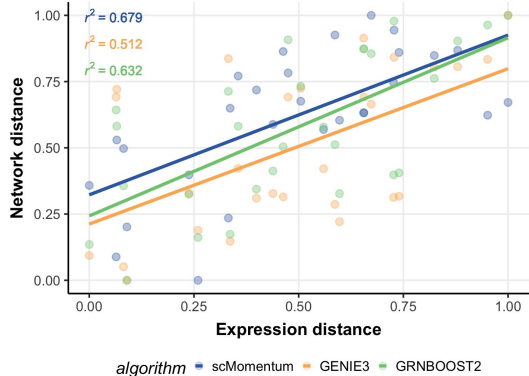
Results

1 Distances are preserved in real datasets

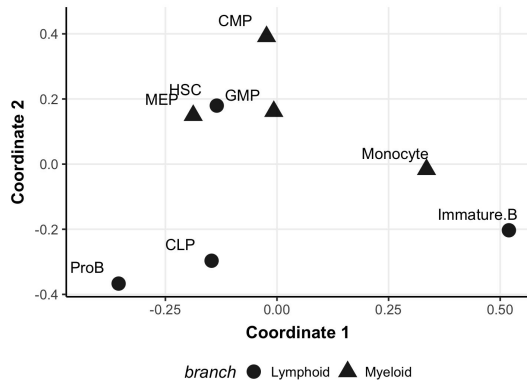
Dataset	# cells	# clusters	r^2	Ref
hFB18	1720	4	0.94	6
hPB20	15,094	8	0.55	7
hED19	4,805	5	0.64	8
mSC19	81,933	19	0.70	9
mBA18	97,181	14	0.44	10

* All correlations had a p -value < 0.05

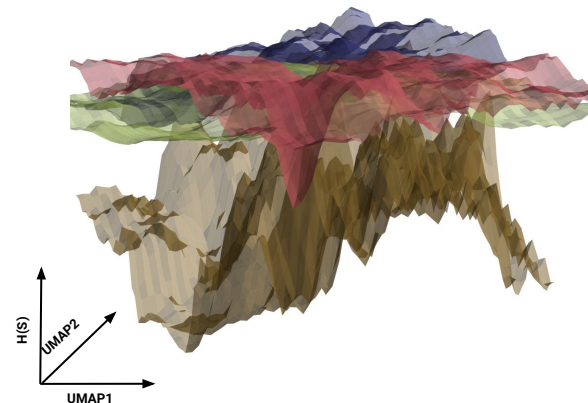
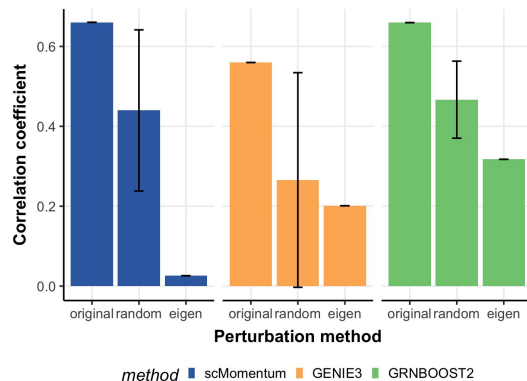
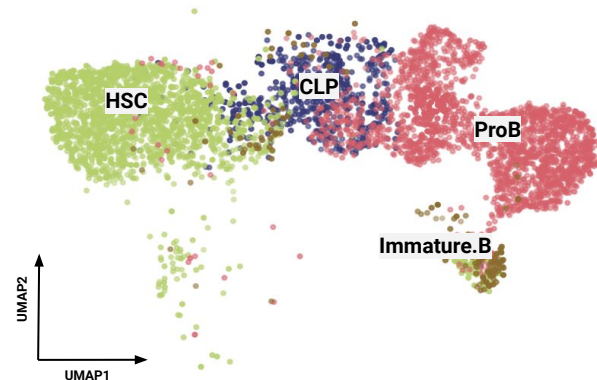
2 scMomentum is better at preserving distances in a real dataset than existing methods.



3 A case study on human hematopoiesis revealed insights on the dynamic properties of the networks.



4 Developmental energy landscape of the Lymphoid lineage

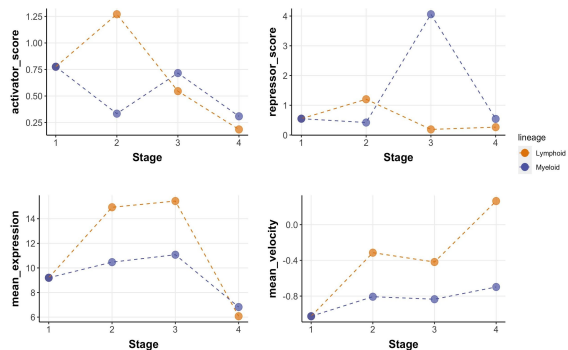


Results

5

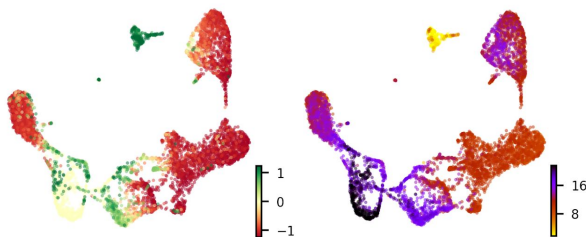
Our networks provide insights on the regulatory changes of individual genes during differentiation

PTMA



velocity

expression



Conclusions

The similarity of the inferred networks captures the **similarities among different cell-types** better than existing methods

The inferred weights permit the reconstruction of **Waddington landscape** analogue

We can discover genes with dynamic properties, providing insights on possible **reprogramming factors**

First cell-type-specific GRN inference method recovering interaction **direction, weight, and sign** in a **data-driven** manner scaling to **large datasets**

References

1. Hopfield, John J. "Neural Networks and Physical Systems with Emergent Collective Computational Abilities." *Feynman and Computation*, pp. 7–19., doi:10.1201/9780429500459-2
2. Cannoodt, R., Saelens, W., Deconinck, L., & Saeyns, Y. (2020). Dyngen: A multi-modal simulator for spearheading new single-cell omics analyses. doi:10.1101/2020.02.06.936971
3. Bergen, V., Lange, M., Peidli, S., Wolf, F. A., & Theis, F. J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nature Biotechnology*. doi:10.1038/s41587-020-0591-3
4. Huynh-Thu, V. A., Irrthum, A., Wehenkel, L., & Geurts, P. (2010). Inferring Regulatory Networks from Expression Data Using Tree-Based Methods. *PLoS ONE*, 5(9). doi:10.1371/journal.pone.0012776
5. Moerman, Thomas, et al. "GRNBoost2 And Arboreto: Efficient and Scalable Inference of Gene Regulatory Networks." *Bioinformatics*, vol. 35, no. 12, 2018, pp. 2159–2161., doi:10.1093/bioinformatics/bty916.
6. Manno, G. L., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., . . . Kharchenko, P. V. (2018). RNA velocity of single cells. *Nature*, 560(7719), 494-498. doi:10.1038/s41586-018-0414-6
7. Single Cell Gene Expression Datasets. (n.d.). Retrieved from <https://support.10xgenomics.com/single-cell-gene-expression/datasets>
8. Zeng, Y., He, J., Bai, Z., Li, Z., Gong, Y., Liu, C., . . . Liu, B. (2019). Tracing the first hematopoietic stem cell generation in human embryo by single-cell RNA sequencing. *Cell Research*, 29(11), 881-894. doi:10.1038/s41422-019-0228-6
9. Delle, J., Rayon, T., Melchionda, M., Edwards, A., Briscoe, J., & Sagner, A. (2019). Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord. *Development*, 146(12). doi:10.1242/dev.173807
10. Zeisel, A., Hochgerner, H., Lönnerberg, P., Johnsson, A., Memic, F., Zwan, J. V., . . . Linnarsson, S. (2018). Molecular Architecture of the Mouse Nervous System. *Cell*, 174(4). doi:10.1016/j.cell.2018.06.021