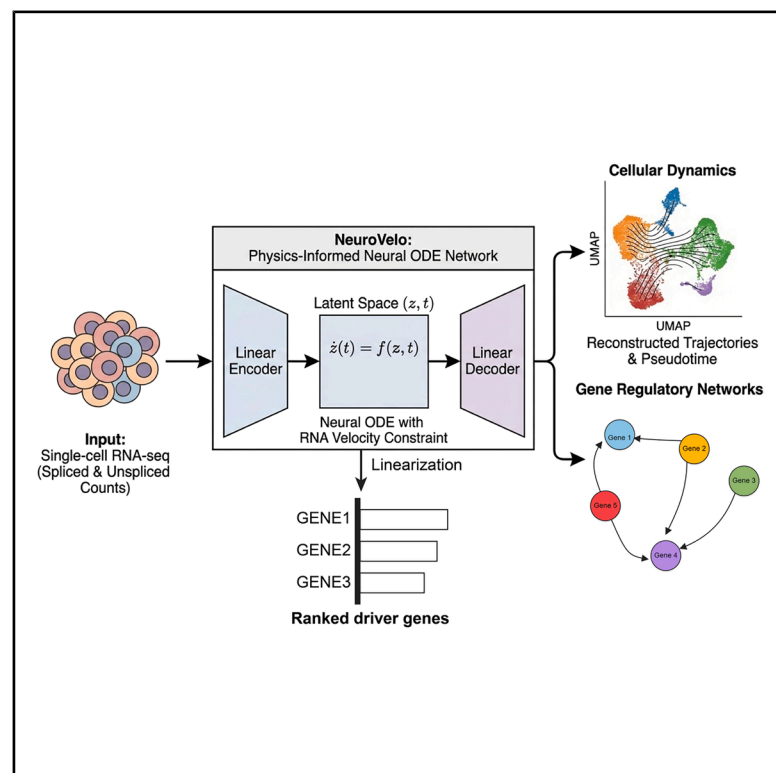


Interpretable learning of temporal cellular dynamics from single-cell data

Graphical abstract



Authors

Idris Kouadri Boudjelthia, Salvatore Milite, Nour El Kazwini, Yuanhua Huang, Andrea Sottoriva, Guido Sanguinetti

Correspondence

gsanguin@sissa.it

In brief

Kouadri et al. introduce NeuroVelo, a physics-informed neural ODE model that infers cell-state trajectories and interpretable gene regulatory mechanisms from static scRNA-seq. Across five datasets, NeuroVelo matches or outperforms RNA velocity baselines and recovers ChIP-seq-supported regulatory networks.

Highlights

- NeuroVelo couples RNA velocity with neural ODEs in an interpretable latent space
- Physics-informed loss improves velocity-field accuracy across five scRNA-seq datasets
- Latent dynamics decoding infers GRNs consistent with ChIP-seq-derived interactions



Article

Interpretable learning of temporal cellular dynamics from single-cell data

Idris Kouadri Boudjelthia,^{1,2} Salvatore Milite,³ Nour El Kazwini,¹ Yuanhua Huang,⁴ Andrea Sottoriva,³ and Guido Sanguinetti^{1,5,6,*}

¹Theoretical and Scientific Data Science, International School for Advanced Studies, Trieste, Italy

²Abdus Salam International Centre for Theoretical Physics, Trieste, Italy

³Computational Biology Research Centre, Human Technopole, Milan, Italy

⁴School of Biomedical Sciences, University of Hong Kong, Hong Kong, China

⁵School of Informatics, University of Edinburgh, Edinburgh, UK

⁶Lead contact

*Correspondence: gsanguin@sissa.it

<https://doi.org/10.1016/j.crmeth.2026.101342>

MOTIVATION Inferring temporal dynamics from static single-cell transcriptomic data remains challenging. While RNA velocity methods provide intuitive cell-level dynamics, they rely on restrictive assumptions and offer limited mechanistic interpretability at the gene regulatory level. Existing approaches typically decouple trajectory inference from gene-regulatory-network reconstruction, losing biological insight. This motivates the need for models that jointly learn cellular dynamics and interpretable regulatory mechanisms directly from data.

SUMMARY

Reconstructing temporal cellular dynamics from static single-cell transcriptomics remains a major challenge. Methods based on RNA velocity are useful, but interpreting their results to learn new biology remains difficult, and their predictive power is limited. Here, we propose NeuroVelo, a method that couples learning of an optimal linear projection with non-linear neural ordinary differential equations. Using dynamical systems theory in the optimized latent space, NeuroVelo can at the same time determine cellular transitions and identify gene interactions that drive the observed temporal dynamics of gene expression. We benchmark NeuroVelo against several state-of-the-art methods using single-cell datasets, demonstrating that NeuroVelo simultaneously reconstructs correct cell-type transitions and identifies gene-regulatory networks that drive cell fate directly from the data.

INTRODUCTION

Single-cell RNA sequencing (scRNA-seq) technologies have transformed our understanding of cellular diversity and heterogeneity,¹ yet the destructive nature of the measurement process poses fundamental limits to their ability to capture temporal biological dynamics. Inferring dynamic information from static snapshots of scRNA-seq has been a major focus of computational research in the last decade. While most early efforts used advanced machine learning techniques to order cells along a pseudo-time trajectory,^{2–5} more recently, the concept of RNA velocity⁶ offered a more mechanistic, biophysically grounded approach to solve the problem.

RNA velocity leverages the surprisingly abundant pre-mRNA reads present in many scRNA-seq datasets to parameterize a simple first-order kinetic model of RNA production and splicing. The relative abundance of spliced and unspliced

reads in a given cell can then be used to deduce whether a gene is being transcriptionally activated or repressed. This idea enables researchers to capture a shadow of cellular dynamics even in static datasets. Recent years have witnessed a flourishing of RNA velocity approaches,^{7–12} permitting a simultaneous inference of pseudo-time and RNA velocity, and coupling the velocity equations with non-linear transformations that can better model the complexity of the data. Nevertheless, both the biochemical foundations and the biological interpretation of RNA velocity have been questioned.^{13,14} In particular, while RNA velocity enables an intuitive grasp of cellular dynamics at the level of individual cells, much less can be said at the level of the gene regulatory networks (GRNs), which underpin such dynamics, which, in general, can only be obtained by combining RNA velocity analyses with post-processing with bespoke GRN inference methods.¹⁵



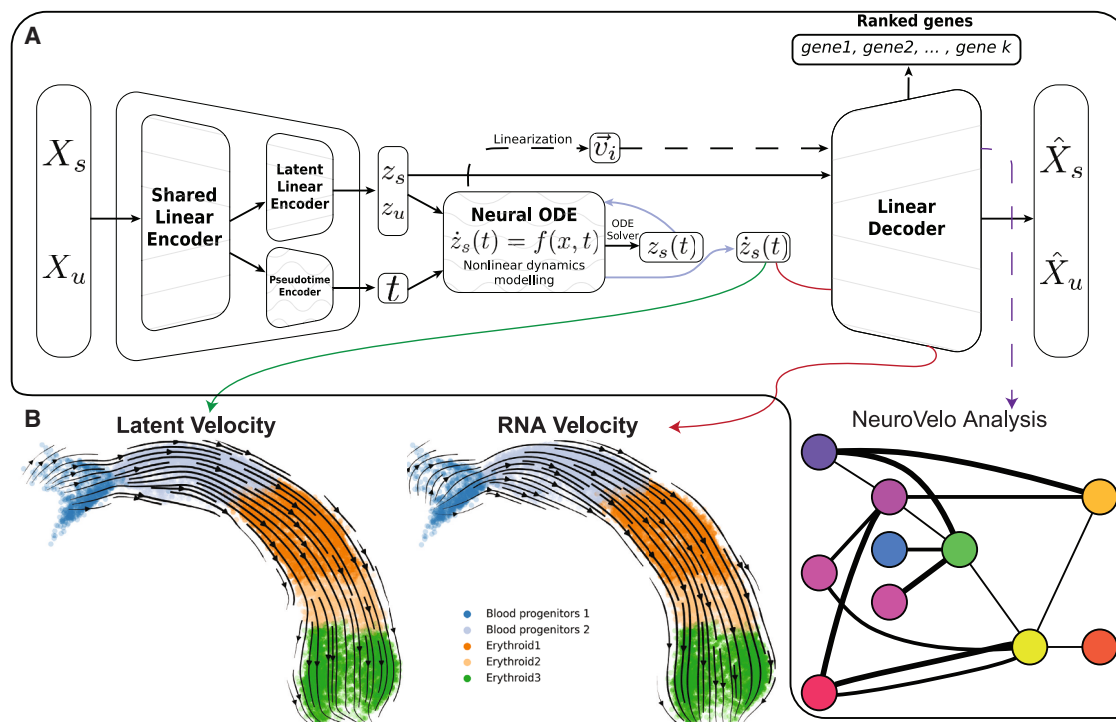


Figure 1. Schematic representation of the NeuroVelo model and its output

(A) Schematic representation of the NeuroVelo model: two autoencoders, a linear one to define the phase space and a non-linear 1D one to define pseudotime, are coupled in latent space through a non-linear neural ODE. RNA velocity constraints are added as an additional penalty in the loss function. (B) Example output of velocity field in the latent space and in the gene space on the mouse gastrulation dataset projected to the uniform manifold approximation and projection (UMAP)-based embedding.

Here, we propose NeuroVelo, a new approach that bridges the cellular and gene scales, simultaneously inferring temporal cellular dynamics and major gene interactions that drive such dynamics from scRNA-seq data. NeuroVelo is an approach grounded in dynamic system theory that combines ideas from neural ordinary differential equations (ODEs)^{16,17} and RNA velocity in a physics-informed neural network architecture. We test extensively NeuroVelo on five scRNA-seq datasets and compare its performance against six recently published methods. On all datasets, NeuroVelo achieved state-of-the-art performance under a variety of metrics. Additionally, we show that NeuroVelo results can be further expanded to recover highly accurate regulatory networks and biological processes underpinning the dynamics in a natural way, thus providing a bridge between RNA velocity and gene regulatory network inference.

RESULTS

The NeuroVelo model

We first present a high-level overview of our method, focusing on its main features and innovative aspects. A detailed mathematical description of the model is provided in the STAR Methods section. Figure 1A illustrates the basic concept of NeuroVelo. scRNA-seq data consisting of both spliced and unspliced reads are encoded to a low-dimensional space through two dimen-

sionality reduction channels: one is a non-linear 1D encoder learning a pseudo-time coordinate associated with each cell, while the second is a linear projection to an *effective phase space* for the system. The observed single-cell transcriptomes are modeled as discrete-time snapshots of a continuous time trajectory happening on this latent phase space.

Cellular dynamics are defined through an autonomous system of differential equations where the rate of change of the state variables (right-hand side of the ODE) is parametrized by a neural network (a neural ODE). The resulting dynamics can be extremely rich: since neural networks are universal function approximators, a neural ODE with a sufficiently large network may give rise to any first-order dynamics. To ground biophysically these dynamics and avoid overfitting, we impose the RNA velocity principle as a further penalty in the loss function in the spirit of physics-informed neural networks. This introduces a trade-off in the fitting procedure: the parameters of the neural ODE and of the projections must strike a balance between providing an accurate interpolation of the observed transcriptomes and maintaining a balance of predicted spliced/unspliced reads consistent with the RNA velocity principle. Notice that this constraint is applied locally to each cell, thus removing limiting assumptions on global transcriptional dynamics, which are often a major problem for RNA velocity approaches. Because the effective phase space is defined via a linear projection, the RNA velocity constraint can be applied *exactly* in the reduced dimensional space.

Table 1. Metrics table for different methods and datasets

	Erythroid mouse		3435_NT_T1	3726_NT_T1
	Mean CDBir	Mean ICCoh	Spearman correlation	Spearman correlation
NeuroVelo	0.8304	0.9999	0.88	0.86
UniTVelo	0.832575	0.98776	0.88	0.77
ScTour	0.801175	0.38038	0.90	0.84
LatentVelo	0.84415	0.97412	0.92	0.88
DeepVelo	0.41515	0.92614	0.45	0.85
ScVelo	−0.3519	0.83798	0.12	−0.043
PhyloVelo	–	–	0.83	0.86

	Human bone marrow		Pancreatic endocrinogenesis	
	Mean CDBir	Mean ICCoh	Mean CDBir	Mean ICCoh
NeuroVelo	0.81943	0.996017	0.603533	0.9997
UniTVelo	0.804933	0.97743	0.614	0.9866
ScTour	0.786033	0.47849	0.5898	0.67406
LatentVelo	0.7664	0.98141	0.5854	0.9763
DeepVelo	−0.804933	0.93207	0.5674	0.8854
ScVelo	−0.839467	0.91819	0.6131	0.8698
PhyloVelo	–	–	–	–

Metrics table for different methods and datasets. For method that allows different random seeds (NeuroVelo, ScTour, and LatentVelo) the table shows the best performance of running 10 seeds. More on stability of velocity is shown in Figure S1.

The choice of defining the effective phase space as a linear subspace of gene space, while may limit the expressivity of the model, brings additional significant advantages: it avoids confounding the non-linear dynamics in the effective phase space with the effects of a non-linear projection, leading to more robust results (see Figure S1). Most importantly, a major benefit of the linearity of the phase space is that standard techniques for the analysis of low-dimensional non-linear dynamical systems can be applied to the trained model. The resulting insights, thanks to the linearity of the encoder/decoder model, can readily be translated in terms of genes and regulatory interactions.

The main tool we use to interpret the regulatory interactions underlying the dynamics is the Jacobian matrix of the system. The Jacobian is obtained by taking first derivatives of the neural network with respect to the state variables, and it tells us how small changes in state variables affect their rates of change. The leading eigenvectors, corresponding to the largest eigenvalues in magnitude, highlight the direction in phase space that contributes most to the system's dynamics. In principle, both the Jacobian and its eigenvectors are time dependent. By decoding the Jacobian matrix in gene space, we can obtain estimates of how a gene affects the transcription rate of another gene at a particular time, hence providing a method to infer time-varying gene regulatory networks. Decoding the leading eigenvectors of the Jacobian into gene space provides ranked lists of important genes that can then be used for enrichment analyses and model validation.

Datasets used, competing methods, and evaluation metrics

To test NeuroVelo, we utilize a total of five recent scRNA-seq datasets:

- A dataset of erythroid cell differentiation¹⁸ during mouse gastrulation, comprising 9,815 cells quantifying a median of 2,791 non-zero expressed genes per cell.
- A human bone marrow hematopoiesis dataset,⁵ widely used in RNA velocity studies, which quantifies a median of 1,190 non-zero expressed genes in 5,780 individual cells.
- A pancreatic endocrinogenesis¹⁹ dataset measuring a median of 2,447 non-zero expressed genes in 3,696 cells.
- Two smaller CRISPR-based lineage-barcoded mouse cancer datasets,²⁰ measuring gene expression in 754 and 1,082 cells at a higher depth (3,490 and 2,352 median non-zero expressed genes, respectively).

To contextualize the performance of NeuroVelo, we analyzed each dataset with a number of recently proposed methods to reconstruct cellular dynamics. Perhaps the most widely used RNA velocity method is scVelo,⁷ which remains the main benchmark. Very recently, multiple methods extended the RNA velocity framework in several directions: we choose here UniTVelo,⁹ which focuses on improving the pseudotime estimation, and latentVelo²¹ and deepVelo,⁸ which couple RNA velocity with non-linear dimensionality reduction. We also use scTour,¹⁷ which does not use RNA velocity but directly fits a neural ODE to the data, as a splicing-independent comparison, and Dynamo,²² which was developed to tackle metabolic labeled scRNA-seq and uses concepts from dynamical systems theory to model RNA processing. Dynamo relies on similar concepts as NeuroVelo, but uses the (real) time measured by metabolic labeling, while NeuroVelo relies on the inferred pseudotime and, therefore, can be generally applicable also for standard scRNA-seq data. On the lineage-barcoded data, we use PhyloVelo,²⁰ a specialized method that exploits barcode information to obtain robust estimates of pseudotime. As performance metrics, we use the cross-boundary direction (CDBir) and the in-cluster coherence (ICCoh), two widely used metrics that summarize the consistency of the inferred dynamics with known biological dynamics (in datasets where cell-type annotation is present). For all methods, we report the best performance across several random restarts, when a default initialization is not provided. Table 1 provides a comprehensive assessment of all methods on all datasets with respect to the inferred RNA velocity fields.

To benchmark the ability of NeuroVelo to reconstruct GRNs, we compare its performance to GENIE3²³ and GRNboost,²⁴ two machine learning-based methods that were highlighted as state of the art in a recent review.²⁵ Assessing the performance of GRN methods on real data is notoriously difficult, due to the lack of gold standards (and indeed the difficulty in associating precise molecular semantics to the interactions summarized by a network edge). To overcome this hurdle, we assembled a database of chromatin immunoprecipitation sequencing (ChIP-seq) tracks from ChIP-Atlas,²⁶ a number of transcription factors in a number of cell types that are relevant in gastrulation and hematopoiesis, the two most studied processes we consider. A list of all datasets

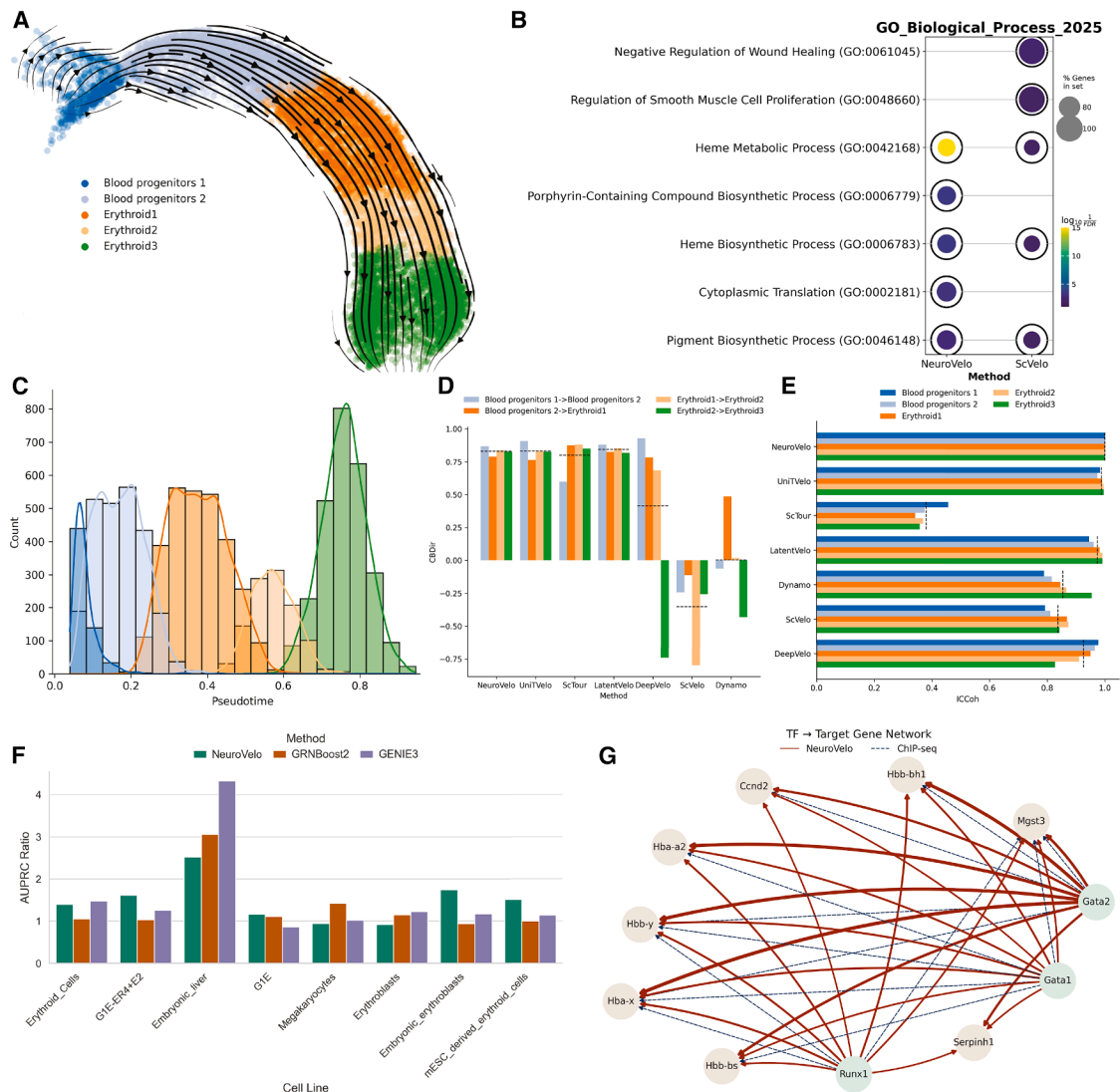


Figure 2. NeuroVelo analysis and benchmark on mouse gastrulation dataset

(A) Velocity field in gene space on the mouse gastrulation dataset projected to the UMAP-based embedding.
(B) Comparison on gene set enrichment analysis of the gene influencing the dynamics.
(C) Histogram of inferred pseudotime.
(D) Comparison of the cross-boundary direction.
(E) Comparison of the inter-cluster coherence.
(F) Comparison of AUPRC ratio for the gene regulatory network inference.
(G) Gene regulatory network inferred by NeuroVelo along with connections from ChIP-seq data.

used with accession codes is provided in [Tables S1](#) and [S2](#). This enables us to obtain a ground-truth network for each cell type, where the various methods can be assessed using standard metrics such as area under the precision-recall curve (AUPRC). See STAR Methods for details of how this evaluation is performed.

Reconstructing dynamical mechanisms from scRNA-seq data

We start by assessing NeuroVelo performance on the three standard (non-barcoded) scRNA-seq, focusing on the ability to reconstruct lineage differentiation and GRNs.

An analysis of the erythroid lineage in mouse gastrulation is shown in [Figure 2](#). NeuroVelo effectively captures the single-branched nature of the process, providing the right cell-type transitions of the dynamics both in latent space and in gene space ([Figure 2A](#)) and a very good sequential ordering of the different cell types ([Figure 2C](#)). The quality of the inferred dynamics is further supported by [Figures 2D](#) and [2E](#), which plot histograms of the evaluation metrics for each cell type, clearly showing NeuroVelo to provide state-of-the-art performance on every cell type.

[Figures 2F](#) and [2G](#) instead concern the interpretation of the decoded dynamics. The barplots in [Figure 2F](#) show the

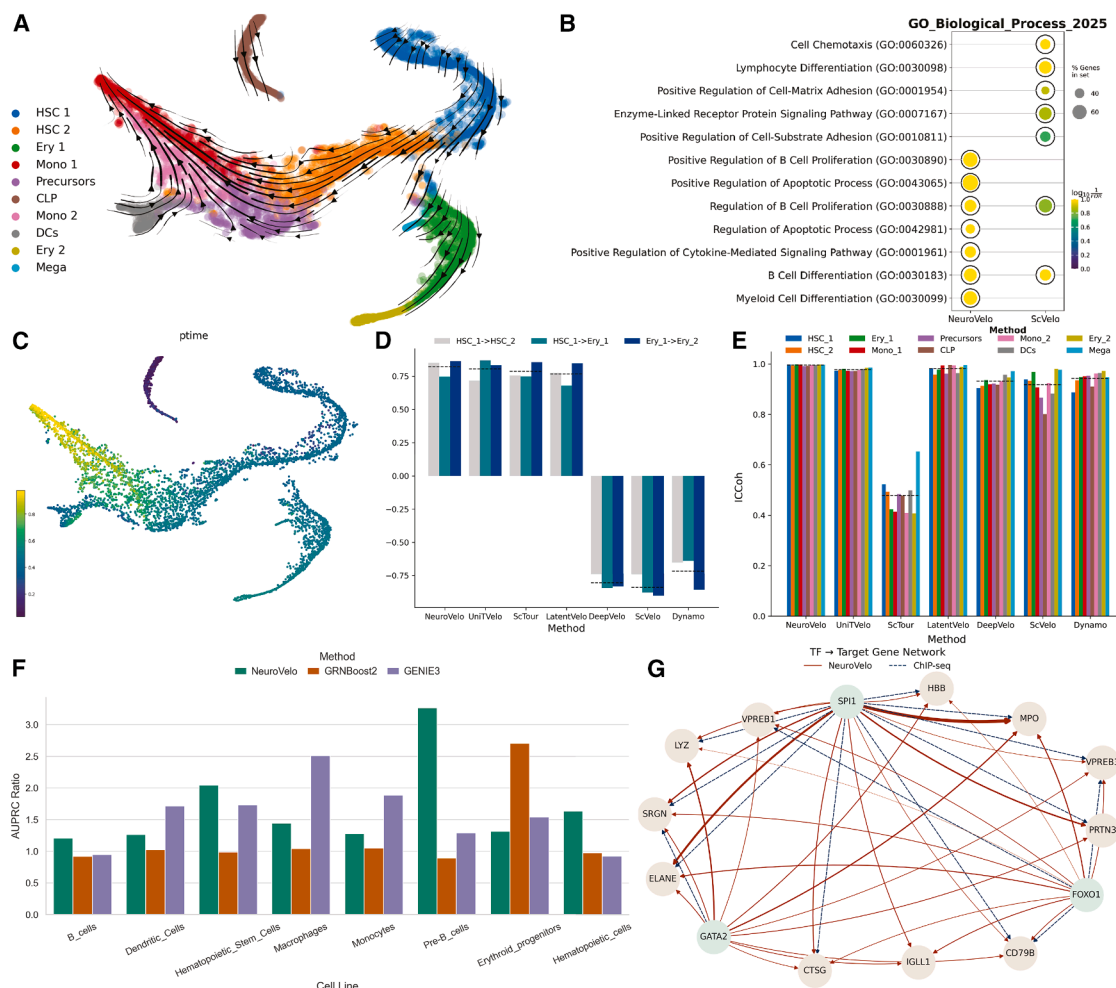


Figure 3. NeuroVel analysis and benchmark on the human bone marrow dataset

- (A) UMAP of NeuroVel velocity field on the human bone marrow.
 (B) Comparison on gene set enrichment analysis of the gene influencing the dynamics.
 (C) UMAP of NeuroVel pseudotime on the human bone marrow dataset.
 (D) Comparison of the cross-boundary direction.
 (E) Comparison of the inter-cluster coherence.
 (F) Comparison of AUPRC ratio for the gene regulatory network inference.
 (G) Gene regulatory network inferred by NeuroVel along with connections from ChIP-seq data.

normalized AUPRC for NeuroVel, Genie3, and GRNboost for eight cell types that are associated with gastrulation (the normalization is such that the random predictor would return a score of one). NeuroVel returns the best performance in half of the cases, with a particularly marked improvement for two of the most relevant cell lines, embryonic erythroblasts and mESC-derived erythroid cells (mouse embryonic stem cells). It should be noticed that NeuroVel is agnostic with respect to which genes are transcription factors, while both Genie3 and GRNboost require as input the list of transcription factor genes, thus significantly narrowing the search space for their models. Figure 2B instead shows an enrichment analysis for NeuroVel (using the ranked list of genes provided by the eigenvectors of the Jacobian) and for scVel, using the list of driver genes in its dynamical model setting. The two analyses broadly overlap

(see also Figure S3 for a comparison of the two gene rankings), although neuroVel returns larger gene sets with a slightly higher significance level.

A similar picture emerges from the analysis of the human bone marrow dataset in Figure 3. NeuroVel provides the right cell-type transitions and a correct pseudotemporal ordering (Figures 3A and 3C), mirrored in state-of-the-art performance under both metrics (Figures 3D and 3E). An enrichment analysis returns several highly enriched pathways related to immune cell proliferation and differentiation (Figure 3B), while in this case, the Gene Ontology analysis was not able to identify meaningful biological signals from the scVel driver genes set.

Figure 3F again shows the agreement of the inferred networks with the ChIP-seq-derived gold standard in a number of relevant cell types. NeuroVel achieves best performance in four of eight

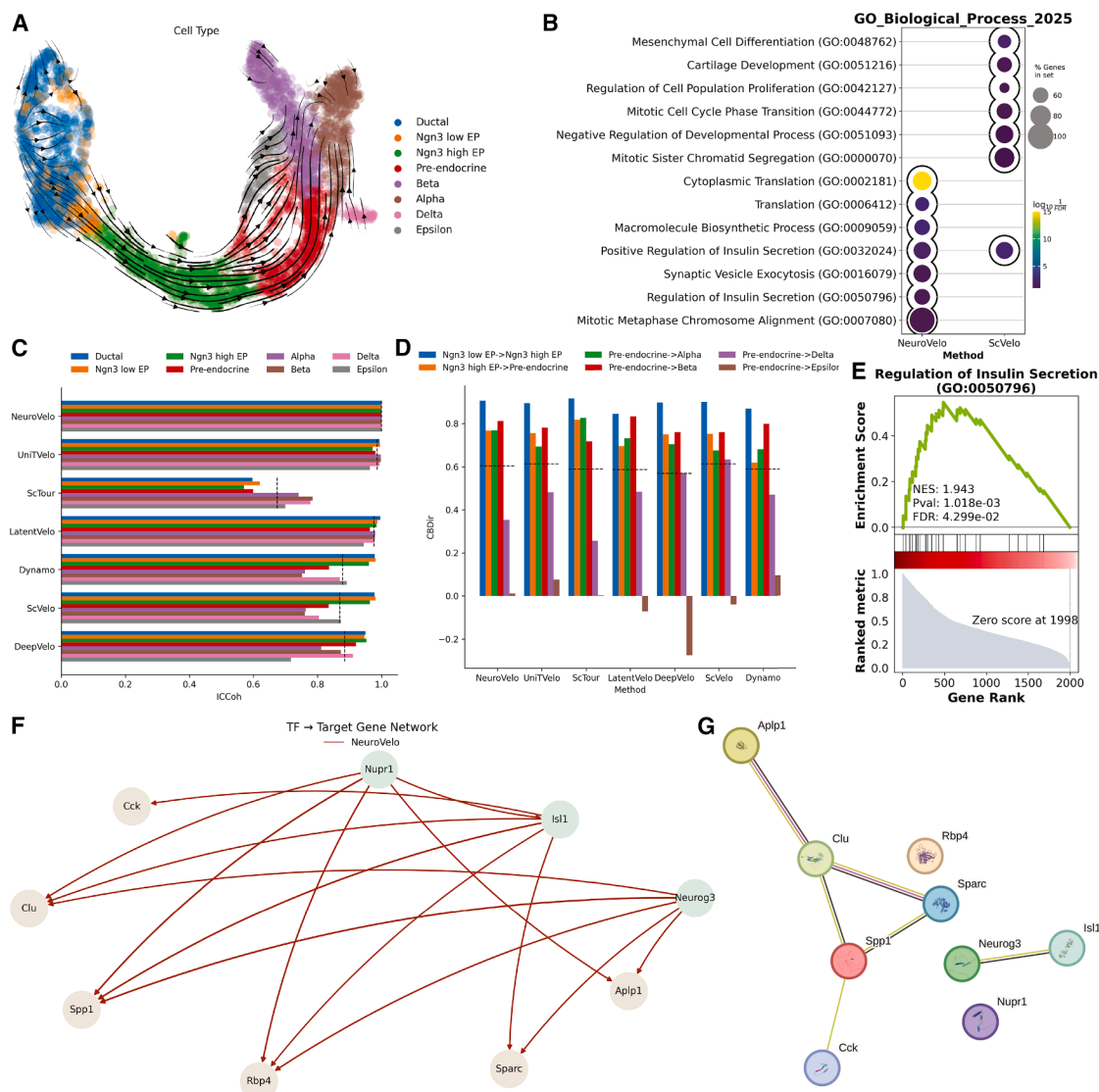


Figure 4. NeuroVelo analysis and benchmark on pancreatic endocrinogenesis dataset

(A) UMAP of NeuroVelo velocity field on pancreatic endocrinogenesis dataset.
(B) Comparison of gene set enrichment analysis of the gene influencing the dynamics.
(C) Comparison of the in-cluster coherence.
(D) Comparison of the cross-boundary direction. (E) Gene set enrichment analysis of insulin secretion pathway.
(F and G) (F) Gene regulatory network inferred by NeuroVelo (G) Gene network structured using string-db.

cell types, including the most relevant cell types (hematopoietic cells and pre-B cells). Interestingly, in this well-annotated dataset, we can also linearize about different points in latent space, producing cell-type specific networks, which highlight the varying temporal dynamics in the process (see Videos S1 and S2).

As a third example of neuroVelo analysis on standard scRNA-seq, we examined the pancreatic endocrinogenesis dataset. Results are shown in Figure 4; once again, neuroVelo provides an excellent visualization (Figure 4A) and performs best or among the best for all cell types under the metrics considered (Figures 4C and 4D). An enrichment analysis (Figures 4B and 4E) highlights highly relevant pathways connected with pancreas

function with very high normalized enrichment score. An analysis of the interactions predicted again highlights some interactions already detailed in the literature (Figures 4F and 4G); however, it is noticeable that in this less-well-studied scenario, fewer interactions can be derived from the literature, underscoring the potential for data-driven approaches like neuroVelo to uncover novel biological hypotheses.

Taken together, these results demonstrate that neuroVelo achieves state of the art or is competitive with the best RNA velocity methods on a variety of datasets. Additionally, neuroVelo is able to predict (possibly time-varying) interaction mechanisms, which can explain the observed transcriptome dynamics and

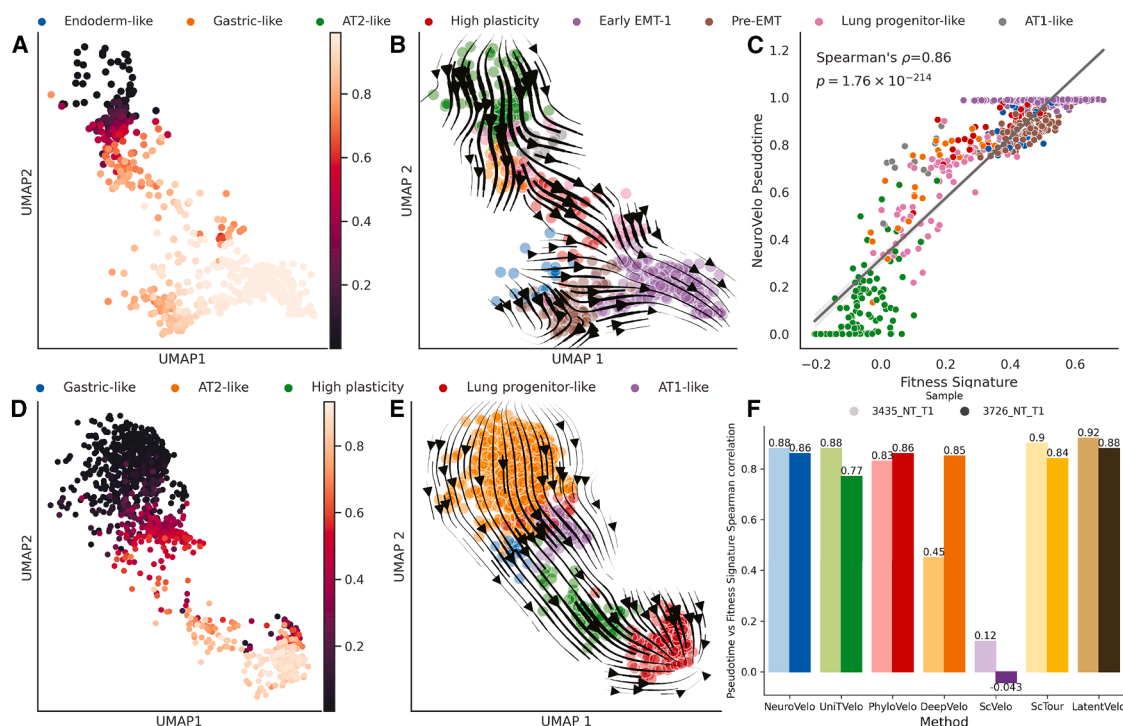


Figure 5. NeuroVelo pseudotime and benchmark on mouse cancer dataset

- (A) UMAP of NeuroVelo-inferred pseudotime on the sample 3726_NT_T1 of the mouse cancer dataset.
 (B) UMAP of NeuroVelo velocity field on the sample 3726_NT_T1 of the mouse cancer dataset.
 (C) Scatterplot of the neuroVelo pseudotime against the fitness signature (inferred from barcoding information).
 (D) UMAP of NeuroVelo-inferred pseudotime on the sample 3435_NT_T1 of the mouse cancer dataset.
 (E) UMAP of NeuroVelo velocity field on the sample 3435_NT_T1 of the mouse cancer dataset.
 (F) Comparison on pseudotime and fitness signature Spearman correlation for 2 different samples.

which frequently recapitulate experimentally observed interactions in well-characterized systems.

Inference of fitness in lineage-barcoded data

As a second set of benchmark data, we consider two lineage-barcoded datasets that were recently published by Yang et al.²⁷ This technology combines scRNA-seq with barcoding to determine simultaneously the fitness of a cell population and its transcriptomic profile. Wang et al.²⁰ also developed a specialized “velocity” method called PhyloVelo, which does not use splicing information but relies on the barcoding to infer cellular dynamics.

Results of the analyses on these two datasets are shown in Figure 5, which again show a very good visualization on both datasets (Figures 5A and 5B and 5D and 5E). To evaluate quantitatively the results, we focus on correlating the inferred pseudotime with the fitness signature, which is a specific gene module whose expression is associated with a cell’s proliferating fitness, and it is estimated from the phylogenetic tree (see Yang et al.²⁷ for details on how fitness signatures are computed).

Figure 5C shows the scatterplot of fitness values versus inferred pseudotime for one of the datasets, while Figure 5F shows a comparison of all methods (including PhyloVelo) in terms of Spearman correlation with fitness. This shows that most methods perform extremely well (in fact, comparably to PhyloVelo, which explicitly uses the barcoding information),

with the exception of scVelo and DeepVelo, with the latter method performing well only on one of the datasets.

This analysis demonstrates that utilizing splicing information can provide an effective way to reconstruct cellular ordering, which validates against barcoding information. Once again, neuroVelo emerges as one of the strongest performers on this task also.

DISCUSSION

NeuroVelo innovates over other RNA velocity methods by bringing a new dimension of functional interpretability at the level of gene regulation: while most existing methods focus on defining a cell-level velocity field, depicting the evolution of individual cells, neuroVelo provides a data-driven mechanistic model that can be interpreted in terms of (time-varying) regulatory networks. Results on all dataset demonstrated not only that neuroVelo is among the best performing tools in terms of standard RNA velocity metrics but also that the mechanisms it captures are biologically meaningful and extend known interactions from the literature. Importantly, a quantitative assessment of the reconstructed GRNs against experimental ChIP-seq data showed NeuroVelo to perform competitively with bespoke GRN inference methods, which were supplied the list of transcription factor genes, a notable advantage that neuroVelo evidently could reconstruct independently from data.

Methodologically, *neuroVelo* attempts to strike a balance between complexity and interpretability. It does so by limiting its non-linearity to the dynamical system in the low-dimensional latent space and in the definition of pseudotime, which enables both a very direct interpretation in terms of genes and the imposition of the RNA velocity constraint as physics-informed loss term. An additional benefit of the relative simplicity of *neuroVelo* is its robustness.

Defining a semi-mechanistic model for single-cell dynamics opens a number of exciting opportunities to expand the scope of *neuroVelo*. First of all, principled ways to incorporate known mechanisms might be devised using ideas from Bayesian statistics. Second, other data modalities, such as chromatin accessibility, could be incorporated when available, thus providing a mechanistic extension to current methods for multi-omics analysis.^{28–30} Finally, different types of underlying dynamics might be considered, for example, to accommodate mutational events, which might lead to non-smooth dynamics, to incorporate spatial correlation in spatial transcriptomic data.

Limitations of the study

NeuroVelo defines the effective phase space through a linear projection of gene expression. While this design choice improves interpretability and robustness, it may limit the model's capacity to capture complex biological processes. In addition, as with other neural network-based approaches, model training can converge to different local minima depending on random initialization. Although we mitigate this effect by aggregating results across multiple trained models, some variability in the inferred dynamics and gene rankings remains unavoidable.

Furthermore, the evaluation of inferred gene regulatory networks relies on ChIP-seq-derived interactions as a proxy for ground truth. Such data provide an incomplete and context-dependent representation of functional regulatory relationships and do not capture all mechanisms influencing gene regulation; hence the GRN benchmark is limited.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Guido Sanguinetti (gsanguin@sissa.it).

Materials availability

This study did not generate new materials.

Data and code availability

- The mouse gastrulation dataset, human bone marrow dataset, and pancreas dataset are available in the *ScVelo* package.⁷ The mouse lung cancer dataset was requested from Wang et al.²⁰ All other data reported in this paper will be shared by the lead contact upon request.
- The source code for *NeuroVelo* is available on GitHub, <https://github.com/idriscb/NeuroVelo>, and Zenodo under DOI <https://doi.org/10.5281/zenodo.18358806>. Along with tutorial notebooks on how to run the method for different datasets, it contains a post-analysis tutorial for extracting the driver genes and how to construct gene regulatory networks from the trained model.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

The authors thank all members of the lab for their feedback and support, especially Federico Caretti for the useful discussion regarding the manuscript and the software. G.S. and N.K. are supported by Associazione Italiana per la Ricerca contro il Cancro - AIRC (27631). A.S. is supported by Associazione Italiana per la Ricerca contro il Cancro - AIRC (28961) and by ERC Consolidator Award (101125077).

AUTHOR CONTRIBUTIONS

I.K.B. and G.S. devised the research with input from S.M., A.S., and Y.H.; I.K.B. implemented the methods with input from S.M. and N.E.K.; all authors analyzed the results; I.K.B., A.S., and G.S. wrote the manuscript; and all authors reviewed the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **METHOD DETAILS**
 - RNA velocity
 - Model architecture
 - Encoder/decoder architecture
 - Pseudotime encoding
 - Neural ODE
 - Linearisation and interpretation
 - GRN construction
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - GRN evaluation: ChIP-seq data processing
 - Model training
 - Parameters selection

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.crmeth.2026.101342>.

Received: October 30, 2025

Revised: December 17, 2025

Accepted: February 9, 2026

Published: March 23, 2026

REFERENCES

1. Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., et al. (2017). The human cell atlas. *eLife* 6, e27041.
2. Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* 32, 381–386.
3. Haghverdi, L., Büttner, M., Wolf, F.A., Büttner, F., and Theis, F.J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. *Nat. Methods* 13, 845–848.
4. Campbell, K.R., and Yau, C. (2018). Uncovering pseudotemporal trajectories with covariates from single cell and bulk expression data. *Nat. Commun.* 9, 2442.

5. Setty, M., Kisieliovas, V., Levine, J., Gayoso, A., Mazutis, L., and Pe'er, D. (2019). Characterization of cell fate probabilities in single-cell data with palantir. *Nat. Biotechnol.* 37, 451–460.
6. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastrioti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). Rna velocity of single cells. *Nature* 560, 494–498.
7. Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing rna velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* 38, 1408–1414.
8. Chen, Z., King, W.C., Hwang, A., Gerstein, M., and Zhang, J. (2022). Deep-velo: Single-cell transcriptomic deep velocity field learning with neural ordinary differential equations. *Sci. Adv.* 8, eabq3745.
9. Gao, M., Qiao, C., and Huang, Y. (2022). Unitvelo: temporally unified rna velocity reinforces single-cell trajectory inference. *Nat. Commun.* 13, 6586.
10. Gayoso, A., Weiler, P., Lotfollahi, M., Klein, D., Hong, J., Streets, A., Theis, F.J., and Yosef, N. (2023). Deep generative modeling of transcriptional dynamics for rna velocity analysis in single cells. *Nat. Methods* 21, 50–59.
11. Li, S., Zhang, P., Chen, W., Ye, L., Brannan, K.W., Le, N.T., Abe, J.I., Cooke, J.P., and Wang, G. (2023). A relay velocity model infers cell-dependent rna velocity. *Nat. Biotechnol.* 42, 99–108.
12. Qiao, C., and Huang, Y. (2021). Representation learning of rna velocity reveals robust cell transitions. *Proc. Natl. Acad. Sci. USA* 118, e2105859118.
13. Barile, M., Imaz-Rosshandler, I., Inzani, I., Ghazanfar, S., Nichols, J., Marioni, J.C., Guibentif, C., and Göttgens, B. (2021). Coordinated changes in gene expression kinetics underlie both mouse and human erythroid maturation. *Genome Biol.* 22, 197.
14. Gorin, G., Fang, M., Chari, T., and Pachter, L. (2022). Rna velocity unraveled. *PLoS Comput. Biol.* 18, e1010492.
15. Singh, R., Wu, A.P., Mudide, A., and Berger, B. (2024). Causal gene regulatory analysis with RNA velocity reveals an interplay between slow and fast transcription factors. *Cell Syst.* 15, 462–474.e5.
16. Chen, R.T., Rubanova, Y., Bettencourt, J., and Duvenaud, D.K. (2018). Neural ordinary differential equations. *arXiv*, Preprint at. <https://doi.org/10.48550/arXiv.1806.07366>.
17. Li, Q. (2023). scout: a deep learning architecture for robust inference and accurate prediction of cellular dynamics. *Genome Biol.* 24, 149.
18. Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaaid, W., Calero-Nieto, F.J., Mulas, C., Ibarra-Soria, X., Tyser, R.C.V., Ho, D.L.L., et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* 566, 490–495. <https://doi.org/10.1038/s41586-019-0933-9>.
19. Bastidas-Ponce, A., Tritschler, S., Dony, L., Scheibner, K., Tarquis-Medina, M., Salinno, C., Schirge, S., Burtscher, I., Böttcher, A., Theis, F.J., et al. (2019). Comprehensive single cell mrna profiling reveals a detailed roadmap for pancreatic endocrinogenesis. *Development* 146, dev173849. <https://doi.org/10.1242/dev.173849>.
20. Wang, K., Hou, L., Wang, X., Zhai, X., Lu, Z., Zi, Z., Zhai, W., He, X., Curtis, C., Zhou, D., and Hu, Z. (2023). Phylovelo enhances transcriptomic velocity field mapping using monotonically expressed genes. *Nat. Biotechnol.* 42, 778–789.
21. Farrell, S., Mani, M., and Goyal, S. (2023). Inferring single-cell transcriptomic dynamics with structured latent gene expression dynamics. *Cell Rep. Methods* 3, 100581. <https://doi.org/10.1016/j.crmeth.2023.100581>.
22. Qiu, X., Zhang, Y., Martin-Rufino, J.D., Weng, C., Hosseinzadeh, S., Yang, D., Pogson, A.N., Hein, M.Y., Hoi Joseph Min, K., Wang, L., et al. (2022). Mapping transcriptomic vector fields of single cells. *Cell* 185, 690–711.e45.
23. Huynh-Thu, V.A., Irlthum, A., Wehenkel, L., and Geurts, P. (2010). Inferring regulatory networks from expression data using tree-based methods. *PLoS One* 5, e12776. <https://doi.org/10.1371/journal.pone.0012776>.
24. Moerman, T., Aibar Santos, S., Bravo González-Blas, C., Simm, J., Moreau, Y., Aerts, J., and Aerts, S. (2018). Grnboost2 and arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics* 35, 2159–2161. <https://doi.org/10.1093/bioinformatics/bty916>.
25. Pratapa, A., Jaliha, A.P., Law, J.N., Bharadwaj, A., and Murali, T.M. (2020). Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. *Nat. Methods* 17, 147–154. <https://doi.org/10.1038/s41592-019-0690-6>.
26. Zou, Z., Ohta, T., and Oki, S. (2024). Chip-atlas 3.0: a data-mining suite to explore chromosome architecture together with large-scale regulome data. *Nucleic Acids Res.* 52, W45–W53. <https://doi.org/10.1093/nar/gkae358>.
27. Yang, D., Jones, M.G., Naranjo, S., Rideout, W.M., Min, K.H.J., Ho, R., Wu, W., Replogle, J.M., Page, J.L., Quinn, J.J., et al. (2022). Lineage tracing reveals the phylogenetics, plasticity, and paths of tumor evolution. *Cell* 185, 1905–1923.e25.
28. Li, C., Virgilio, M.C., Collins, K.L., and Welch, J.D. (2023). Multi-omic single-cell velocity models epigenome–transcriptome interactions and improves cell fate prediction. *Nat. Biotechnol.* 41, 387–398. <https://doi.org/10.1038/s41587-022-01476-y>.
29. Kazwini, N.E., and Sanguinetti, G. (2024). Share-topic: Bayesian interpretable modeling of single-cell multi-omic data. *Genome Biol.* 25, 55. <https://doi.org/10.1186/s13059-024-03180-3>.
30. El Kazwini, N., Gao, M., Boudjelthia, I.K., Cai, F., Huang, Y., and Sanguinetti, G. (2024). Crak-velo: Chromatin accessibility kinetics integration improves rna velocity estimation. Preprint at bioRxiv. <https://doi.org/10.1101/2024.09.12.612736>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Mouse gastrulation	Pijuan-Sala et al. 2019	E-MTAB-6967
Human bone marrow	Setty et al. 2018	ERP120467
Pancreatic endocrinogenesis	Bastidas-Ponce et al. 2019	GSE132188
KP-Tracer mice Raw data	Yang et al. 2022	PRJNA803321
Software and algorithms		
NeuroVelo	This manuscript	https://github.com/idriskb/NeuroVelo https://doi.org/10.5281/zenodo.18358806
ScVelo (0.2.5)	Bergen et al. 2020	https://github.com/theislab/scvelo
scanpy (1.9.4)	Wolf et al. 2018	https://github.com/scverse/scanpy
torch (2.0.0)	PyTorch Foundation	https://pytorch.org/
torchdiffq (0.2.3)	Chen et al. 2018	https://github.com/rtqichen/torchdiffq

METHOD DETAILS

RNA velocity

ScRNA-seq captures a static snapshot of gene expression at a single time point, hence modeling cellular dynamics from single-cell RNA-seq data is not straightforward. La Manno et al.⁶ leverage the presence of unspliced pre-mRNA to model transcriptional dynamics from single cell data. The coexistence of unspliced and spliced transcripts enables the inference of RNA velocity by modeling the underlying splicing kinetics. A simple model was assumed by⁶ describing the evolution of unspliced and spliced mRNA molecules over time given by the system of Equation 1

$$\begin{aligned}\dot{u} &= \alpha - \beta u(t) \\ \dot{s} &= \beta u(t) - \gamma s(t)\end{aligned}\quad (\text{Equation 1})$$

$\dot{u}$ and $\dot{s}$ are time derivatives of the abundance of unspliced and spliced RNA (also known as RNA velocity), respectively. α is the transcription rate, β is the splicing rate, and γ is the degradation rate of the spliced RNA. By estimating parameter β and γ , one can make RNA velocity estimate $\dot{s}$, which is a vector in gene space that gives the infinitesimal change in abundance of spliced RNA amount (hence the analogy with the concept of velocity from classical mechanics).

Model architecture

NeuroVelo is a neural network architecture that consists of a linear auto-encoder and neural ODEs architecture in the latent representation of the data. The main idea of NeuroVelo is to construct cell dynamics from scRNA-seq data in the latent space.

Encoder/decoder architecture

We used spliced and unspliced reads of scRNA-seq data with g number of genes, and embedded them using the same 2 layers linear encoder function $E: \mathbb{R}^g \rightarrow \mathbb{R}^l$, into a latent space $\mathbf{z} = (\mathbf{z}_s, \mathbf{z}_u)$ of dimension l .

$$\begin{aligned}\mathbf{z}_s &= W_e^{(2)}(W_e^{(1)}X_s + b_e^{(1)}) + b_e^{(2)} \\ \mathbf{z}_u &= W_e^{(2)}(W_e^{(1)}X_u + b_e^{(1)}) + b_e^{(2)}\end{aligned}$$

This step is mainly for dimensionality reduction and feature extraction of the highly dimensional gene space to a much smaller latent space (50 by default).

The encoder latent representation $\mathbf{z}$ is decoded with a 2 layers linear decoder network $D: \mathbb{R}^l \rightarrow \mathbb{R}^g$ to create a reconstruction of the original spliced and unspliced inputs $(\hat{X}_s, \hat{X}_u)$ learning a function that maps from latent space to gene space.

$$\begin{aligned}\hat{X}_s &= W_d^{(2)}(W_d^{(1)}\mathbf{z}_s + b_d^{(1)}) + b_d^{(2)} \\ \hat{X}_u &= W_d^{(2)}(W_d^{(1)}\mathbf{z}_u + b_d^{(1)}) + b_d^{(2)}\end{aligned}$$

Pseudotime encoding

The output of the first layer of the encoder is also used to estimate the pseudotime t of the cells, such that:

$$t = S(W_t(W_e^{(1)}X_s + b_e^{(1)}) + b_t)$$

Where $S(x)$ is the sigmoid function. Both pseudotime and latent representation of the cell are used to train nonlinear neural ODEs.

Even though the encoder and decoder consist of two linear layers each, we note that a composition of linear transformations is mathematically equivalent to a single linear mapping. This implies that the two layer network does not increase capacity compared to a one-layer network. However, in practice we found that the two-layer parameterization provides training stability. Moreover, the use of two linear layers for the encoder serves an additional purpose. The first layer is shared between the latent encoder and the (nonlinear) pseudotime encoder, enforcing a common feature representation extracted from the high-dimensional gene space. Hence, this shared intermediate layer couples the latent representation and pseudotime inference which improves stability.

Neural ODE

Once we project the cells into a lower-dimensional space, we use a set of neural ODEs to describe the cellular dynamics in this latent space

$$\frac{dz_s^{(d)}(t)}{dt} = f^{(d)}(z_s^{(d)}(t), t), \quad (\text{Equation 2})$$

where $f^{(d)}: \mathbb{R}^l \rightarrow \mathbb{R}^l$ is a two layers neural network that takes $z_s \in \mathbb{R}^l$ (with $l = 50$ by default) as input. The network uses the Exponential Linear Unit (ELU) function as the nonlinear activation function, and has a hidden layer dimension of 100. This neural ODE describes the dynamics of cells from particular sample d . Using the nonlinear ODEs in the latent space helps to capture the essential and complex dynamics of cells in this space.

Notice that, since we are learning both the pseudotime and the generator function of the dynamics (the f function), one could reverse simultaneously the pseudotime and the sign of f without altering the trajectories. This ambiguity may be resolved either by prior knowledge or by inverting both signs *a posteriori*.

Linearisation and interpretation

For any dynamical process governed by a nonlinear ODE, we can understand the local behavior of the system by linearizing the ODE around a given point. In this case we choose as the point about which to linearise the average of the encoded state of (a group of) cells; the linearized system becomes:

$$\dot{\mathbf{z}}_s(t) = \mathbf{J}\mathbf{z}_s(t), \mathbf{J} = \begin{bmatrix} \frac{\partial f_1}{\partial z_1} & \cdots & \frac{\partial f_1}{\partial z_l} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_l}{\partial z_1} & \cdots & \frac{\partial f_l}{\partial z_l} \end{bmatrix} \quad (\text{Equation 3})$$

where the matrix of partial derivatives $\mathbf{J}$ is the Jacobian matrix. We perform *eigenanalysis* on the Jacobian matrix. The eigenvalues give the important directions of the correspondent eigenvectors in the latent space, then we use the decoder function to map these important directions in the latent space to the gene space. The output of this decoder is a list of ranked genes based on their importance in a particular direction.

The main issues related to using neural networks are interpretability and robustness, the former is not a problem because NeuroVelo uses both linear embedding and projection, thus the interpretation of the latent space into the gene space is straightforward. The robustness part is more of a concern because a different initialization leads to a different local minima which makes biological interpretation uncertain. Here we propose a geometrical and statistical approach that gives a solid gene list by using many trained models.

The first step is to find the set of eigenvectors from other models that align the most with an eigenvector of a current model $\vec{v}_i^{(m)}$, the alignment is based on the absolute value of the cosine similarity. This set of the most aligned eigenvectors is given by the following equation:

$$U(\vec{v}_i^{(m)}) = \left\{ \max_j \left| S_c(\vec{v}_i^{(m)}, \vec{v}_j^{(m')}) \right|, \text{for } m' \neq m \right\}$$

where S_c is the cosine similarity function and m' is the index of the trained model.

We take the ranked genes of the aligned eigenvectors, and we average the ranks gene-wise. The output is a list of genes ranked based on the average rank across the aligned eigenvectors from different trained models. This indicates that the genes that are at the top of the list hold significant importance for the dynamics learned by various models and vice versa. A set of different analyses can be performed on these average-ranked genes for validation.

GRN construction

As shown in Equation 3, the non-linear latent dynamics are approximated by the linearized system $\dot{\mathbf{z}}_s(t) = \mathbf{J}\mathbf{z}_s(t)$.

This $\mathbf{J}$, is obtained by linearizing the latent dynamical system around an average of group of cells, this could be all the cells -as done for benchmarking against other GRN methods- or by cell-type or cell pseudotime ordering -as done for the time dependent GRN in supplementary material-.

By mapping the latent representation to the gene space, we can show that this system can be written as:

$$\dot{\mathbf{X}}_s(t) = \hat{\mathbf{J}}\mathbf{X}_s(t) + \mathbf{C}$$

where $\mathbf{C}$ is a bias term that does not depend on $\mathbf{X}$. We call $\hat{\mathbf{J}} = \mathbf{W}_d\mathbf{J}\mathbf{W}_e$ the decoded Jacobian, $\mathbf{W}_d$ and $\mathbf{W}_e$ are the weight of the decoder and encoder networks respectively.

The matrix $\hat{\mathbf{J}} \in \mathbb{R}^{g \times g}$ (where g is the number of genes) captures the effective interactions between gene expression levels-i.e., the element $\hat{J}_{ij}$ tell us how a change in the expression of gene j influences the rate of change of gene i . A positive $\hat{J}_{ij}$ suggests that gene j activates gene i , while a negative value suggests that gene j represses gene i .

To construct a graph from the matrix $\hat{\mathbf{J}}$ we take a directed graph $G=(V,E)$ approach, such that each node v_i correspond to a gene i and a directed edge $(v_j \rightarrow v_i)$ is weighted based on the magnitude of $\hat{J}_{ij}$.

QUANTIFICATION AND STATISTICAL ANALYSIS

GRN evaluation: ChIP-seq data processing

The main challenge with benchmarking GRN is to establish the ground truth network to compare it with, and taking into consideration that gene regulatory network depends on the cell-type/cell-line. We create a *gold standard* gene regulatory network using chip-seq data from ChIP-Atlas,²⁶ and we take the MACS2 score as a signal of the transcription factor binding to a target gene. We look at potential genes where the transcription factor binds at $\pm 10k$ distance from TSS of that gene.

To compare the inferred GRN by the 3 methods with the gold standard one, we take a sub-graph containing only TFs-Target genes pairs that exist in the ground truth. For each method, the inferred network provides a continuous score for every TF-target interaction (feature importance for GRNBoost and GENIE3, interaction strength for NeuroVelo), and these scores are used directly to rank predicted interactions. We construct a reference binary vector in which each TF-target pair is labeled as positive if it appears in the gold standard graph and negative otherwise. The corresponding predicted scores from each method form a continuous valued prediction vector.

A precision-recall curve is then generated by successively thresholding the predicted scores and treating each threshold as defining binary classifier, i.e., edges with scores above the threshold are considered predicted positives. At each threshold, we compute precision and recall relative to the reference binary vector. and we compute the area under precision recall curve (AUPRC) using the trapezoidal rule. We use different cell lines and different threshold on the MACS2 score to evaluate the AUPRC, and we report AUPRC ratio, which is AUPRC/Baseline(Random Guess).

For the case of NeuroVelo and because we are taking MACS2 score as ground truth, we cannot benchmark for the type of interaction -whether it is activation or repression-, thus we take the magnitude as weighted graph edge.

Model training

The tasks learned by neural networks are based on what loss function the network is optimizing. To ensure that the cellular dynamics learned by the network have biological meaning, we propose a loss function that puts biophysical constraints on the data and does not require any parameter fine-tuning.

The first part of the loss is the usual mean squared error (MSE) of autoencoders network, we want our spliced and unspliced inputs to match their reconstructed counterparts from the decoder and so the first loss part is a sum of MSE of spliced and unspliced reads.

The second part of the loss is what puts a biophysical constraint on the neural ODEs in the same fashion of physics-informed neural networks. The idea is to make the derivative of latent spliced reads $\dot{\mathbf{z}}_s^{(d)}(t)$ align with the splicing dynamics from the RNA velocity but in latent representation ($e^{\beta}\mathbf{z}_u - e^{\gamma}\mathbf{z}_s$), There is an exponential of β (and γ) because they are analogy of splicing and degradation rate and they must be positive.

The final loss is written as follows:

$$L = \text{MSE}(\mathbf{X}_s, \hat{\mathbf{X}}_s) + \text{MSE}(\mathbf{X}_u, \hat{\mathbf{X}}_u) + \sum_{d=1}^D \text{MSE}(\dot{\mathbf{z}}_s^{(d)}(\mathbf{z}_s^{(d)}(t), t), e^{\beta}\mathbf{z}_u^{(d)} - e^{\gamma}\mathbf{z}_s^{(d)}) + \text{MSE}(\mathbf{z}_s(t), \mathbf{z}_s) \quad (\text{Equation 4})$$

Where the sum over d is the sum over specific group of cells or treatments. $\dot{\mathbf{z}}_s^{(d)}$ is the function that describes the splicing dynamics of a sample d in the latent space. $\mathbf{z}_s^{(d)}$ and $\mathbf{z}_u^{(d)}$ are the projected spliced and unspliced reads of the sample d . Notice that the third term in the loss enforces the RNA-velocity constraint, so that the role of the transcription rate is modeled flexibly and learned directly from the data. The last term in the loss aim to align the latent dynamics with the latent representation, due to the linearity of the model, this is

equivalent to $\text{MSE}(\hat{X}_s(t), X_s)$ where $\hat{X}_s(t)$ is the decoded latent dynamics $z_s(t)$.

In this way, we are able to investigate complex and specific cellular dynamics by learning splicing dynamics constrained nonlinear ODEs in a reduced space and without the need to fine-tune parameters on the level of the loss function as well.

Parameters selection

RNA velocity estimation can be very sensitive to the processing done such as constructing kNN graph as shown in¹⁴ to use “moments” instead of raw reads. In addition, using neural network to estimate RNA velocity increases the number of parameter combinations, such as learning rate, optimizer, batch size and network size (dimensions). Moreover, for the same set of parameters, different initializations can be stuck in different local minima, leading to convergence in any case, but to different velocity estimates.

In address this issue, we propose a metric, *Velocity Alignment* which uses the same cosine similarity measure as *ICCoH* but is applied to velocities estimated using models trained with different random seeds.

$$VA^{(n,m)} = \frac{1}{N} \text{diag}(S_c(\mathbf{V}^{(n)}, \mathbf{V}^{(m)}))$$

$V^{(n)}$ and $V^{(m)}$ are RNA velocity estimated from different trained model (n) and (m) and N is the total number of cells. The diag takes the diagonal elements of the cosine similarity measure, meaning that we are measuring how the same cell velocity from 2 different models is aligned.

The idea is to select the set of parameters with the highest cosine similarity. This approach ensures that even if the model converges to a local minimum, the velocity estimated from different seeds is aligned.

It is important to point that this metric should be computed for models with different random initializations but for the same set of other parameters.