VeloViz: RNA-velocity informed 2D embeddings for visualizing cellular trajectories

Lyla Atta[1,2,3], Arpan Sahoo[2,4v], Jean Fan[1,2,4]

[1]Department of Biomedical Engineering, Johns Hopkins University, Baltimore MD; [2]Center for Computational Biology, Johns Hopkins University, Baltimore MD;
[3]Medical Scientist Training Program, Johns Hopkins University School of Medicine, Baltimore MD; [4]Department of Computer Science, Johns Hopkins University, Baltimore MD

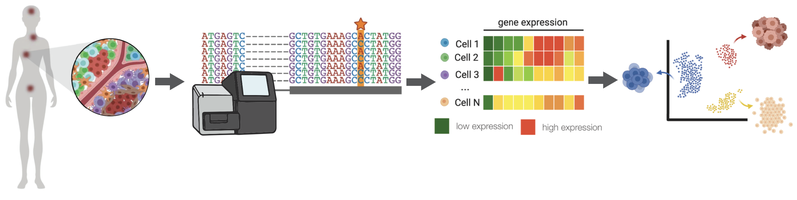
Abstract

Advances in high-throughput transcriptomic profiling technologies have enabled the genome-wide gene expression measurements in individual cells but can only provide a static snapshot of cell states because these protocols destroy cells. Recently developed RNA velocity analysis infers gene expression dynamics from snapshot measurements to predict the future state of a cell from single cell RNA-sequencing and imaging data. To interpret these cell state changes as part of underlying cellular trajectories, current approaches rely on visualization on 2D representations of the observed data derived from dimensional reduction methods like principal components analysis, t-distributed stochastic neighbor embedding, among others. However, these methods can yield different representations of the underlying trajectories, hindering the interpretation of cell state changes. To address this challenge, we developed VeloViz, which incorporates RNA velocity information to create 2D representations of single cell data. We use VeloViz to visualize cellular developmental trajectories in simulated data as well as single cell RNA-sequencing and spatial transcriptomic imaging data. We show that by taking into consideration the predicted future gene expression states from RNA velocity analysis, VeloViz can help ensure a more reliable representation of underlying cellular trajectories.

Preprint:
tinyurl.com/
velovizbiorxiv

Software:
jef.works/
veloviz

Motivation: understanding tissues at a single cell resolution



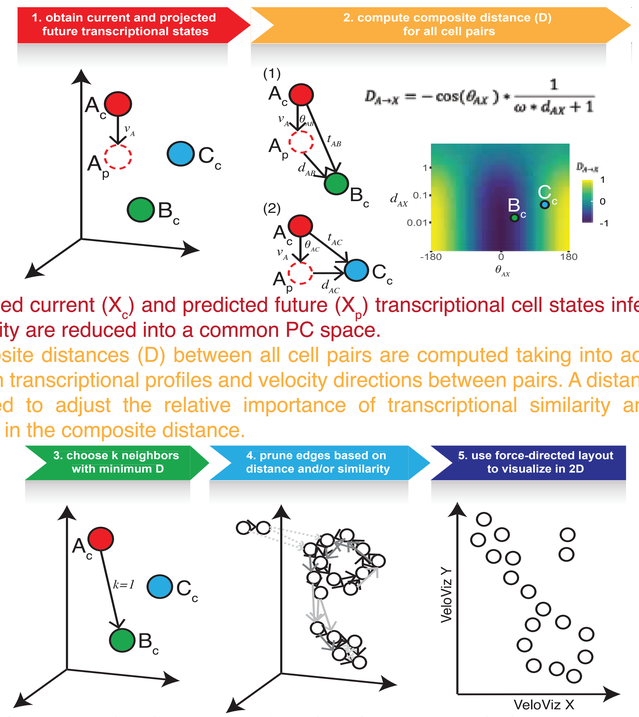
Single cell transcriptomics provides a snapshot of cells' current state, offering insight into cell and tissue heterogeneity at a given time point.

RNA velocity infers gene expression changes from transcriptomics data to predict future transcriptional states of single cells.

Current methods used to interpret single cell gene expression data rely on dimensional reduction (using PCA, t-SNE, UMAP, or diffusion map) to visualize cells in two or three dimensions, but only consider current transcriptional states.

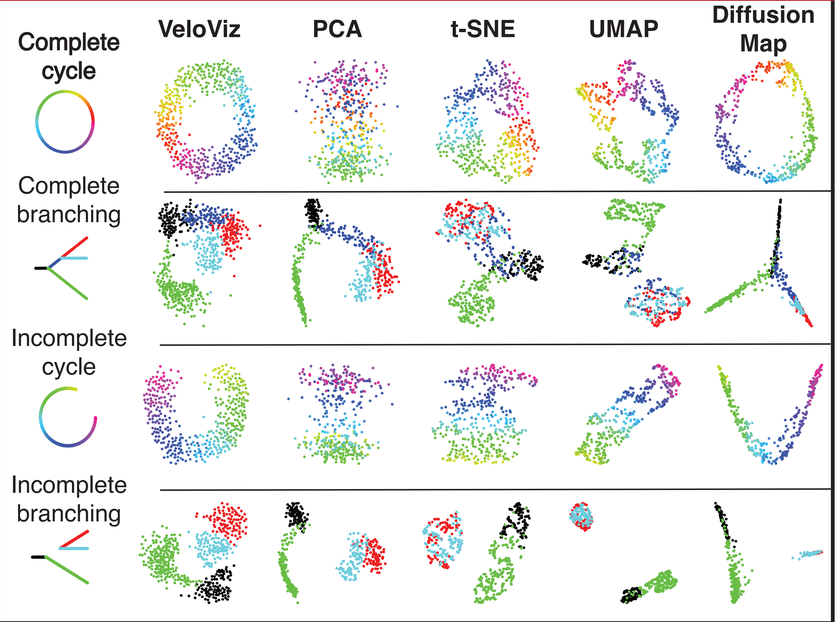
Taking into account predicted future transcriptional states from RNA velocity analysis when creating low dimensional representations can help ensure a more reliable representation of underlying cellular trajectories.

Method: Graph construction using velocity



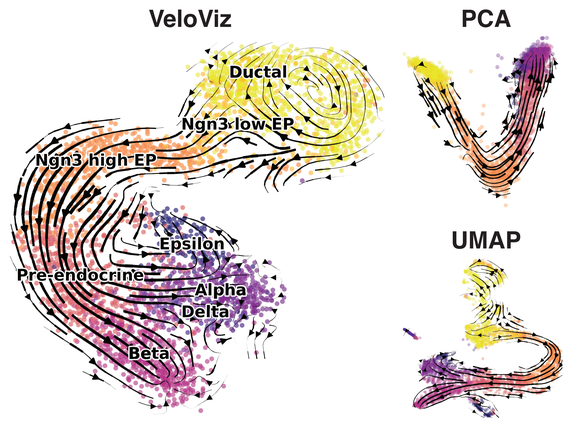
(1) Observed current (X_c) and predicted future (X_p) transcriptional cell states inferred from RNA velocity are reduced into a common PC space.
(2) Composite distances (D) between all cell pairs are computed taking into account the similarity in transcriptional profiles and velocity directions between pairs. A distance weight (ω) is used to adjust the relative importance of transcriptional similarity and cosine correlation in the composite distance.
(3) For each cell, graph edges are assigned to the k cells with the minimum composite distances to create a graph.
(4) Edges are removed (in grey, dashed) if they are above similarity or distance thresholds
(5) The graph is visualized as a 2D embedding using a force-directed graph layout.

Results: Visualizing simulated trajectories



Results: Visualizing pancreas development

Full trajectory

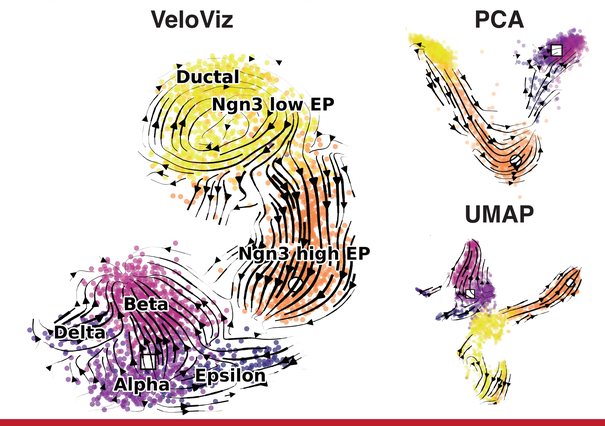


Data:
Mouse embryonic pancreas development single cell RNA sequencing.

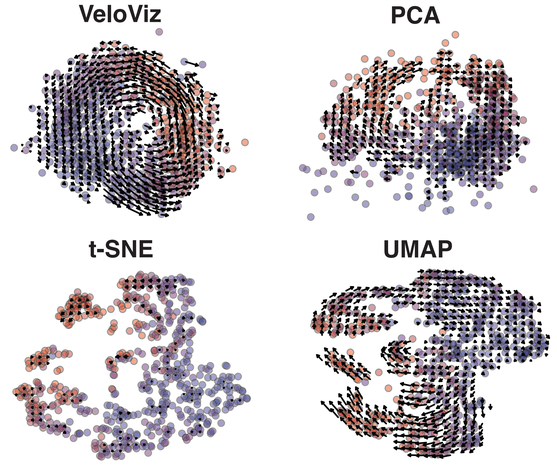
(Bergen *et. al.* Nat. Biotech. 2020; Bastidas-Ponce *et. al.* Development 2019)

Expected trajectory

Trajectory with missing intermediates



Results: Visualizing cell cycle



Data:
U-2 OS human osteosarcoma cell line muliplexed error-robust fluorescent in situ hybridization.

(Xia *et. al.* PNAS 2019)