Network Analysis of Single-cell Transcriptome

Dissertation Mid-Term Evaluation

ADVISOR
Malay Bhattacharyya

BY
Pradip Das (CS2115)

ISI Kolkata , 21st Feb 2023

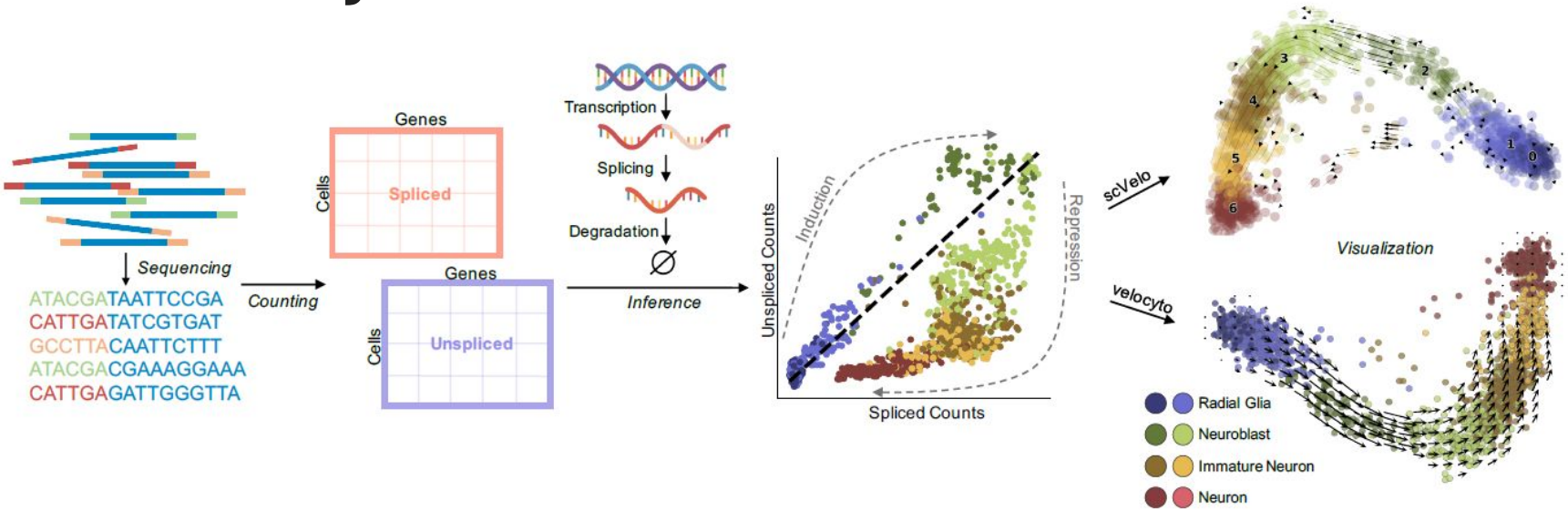
Abstract

- Single cell RNA sequencing (ScRNA-seq) method have enabled us studying gene expression at the level of individual cells.
- This data can be useful to reconstruct the dynamics of interactions that play crucial role in an organism.
- Unfortunately there is a gap in understanding how to integrate data by aligning cells with time point and learning continuous trajectories.
- Here we aim to perform network reconstruction with the approach of multi-omics for exploratory analysis.

Background

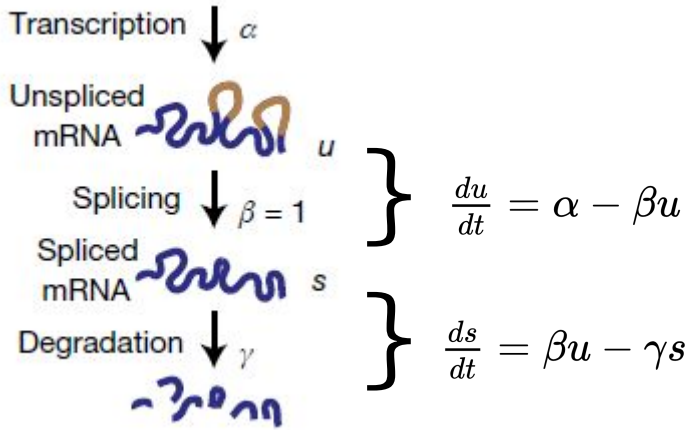
- RNA abundance is a powerful indicator of the state of individual cells.
- Single-cell RNA sequencing can reveal RNA abundance with high quantitative accuracy, sensitivity and throughput
- This approach captures only a static snapshot at a point in time.
- RNA velocity—the time derivative of the gene expression state—can be directly estimated by distinguishing between unspliced and spliced mRNAs in common single-cell RNA sequencing protocols.
- RNA velocity is a high-dimensional vector that predicts the future state of individual cells.

RNA velocity estimation



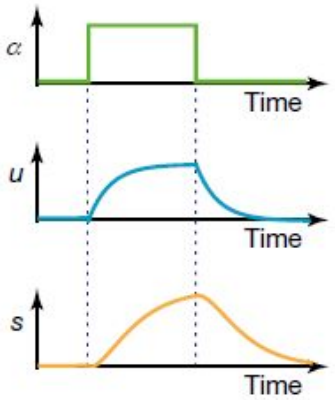
Initial processing of sequencing reads produces spliced and unspliced counts for every cell, across all genes. Inference procedures, implemented in velocyto and scVelo,fit a model of transcription, and predict cell-level velocities.

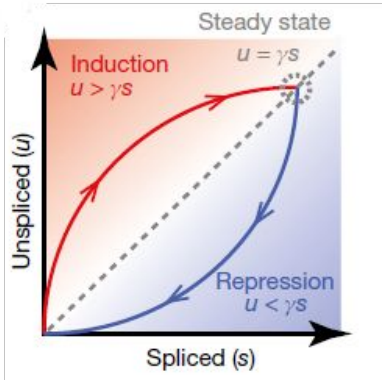
RNA velocity estimation



$$\frac{du}{dt} = \alpha - \beta u$$

$$\frac{ds}{dt} = \beta u - \gamma s$$





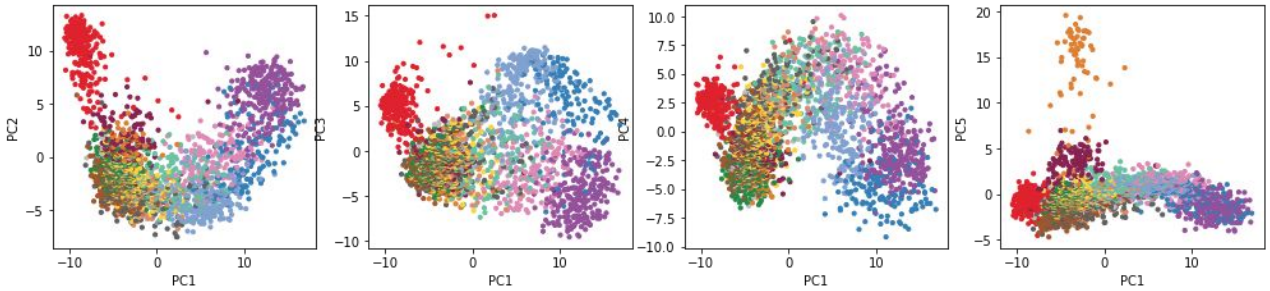
a. Model of transcriptional dynamics, capturing transcription (α), splicing (β) and degradation (γ) rates involved in production of unspliced (u) and spliced (s) mRNA products.
b. Solution of the model in b as a function of time, showing unspliced and spliced mRNA dynamics in response to step changes in α.
c. Steady states for different values of transcription rates α fall on the diagonal given by slope γ (dashed line). Levels of unspliced mRNA above or below this proportion indicate increasing (red shading) or decreasing (blue shading) expression of a gene, respectively

Visualization of RNA velocity

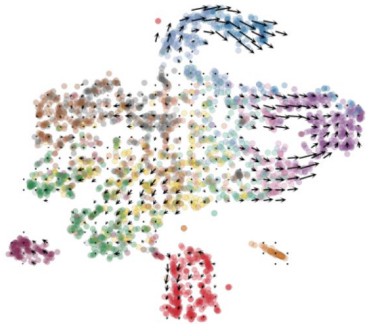
- A variety of techniques can be used to visualize the velocity estimates in low dimensions.
- The observed and extrapolated cell states can be jointly embedded in a common low-dimensional space (using PCA).
- Alternatively, velocities can be projected onto existing low-dimensional embeddings, such as t-distributed stochastic neighbour embedding (t-SNE), on the basis of the similarity of the extrapolated state to other cells in the local neighbourhood.

Visualization of RNA velocity

PCA:

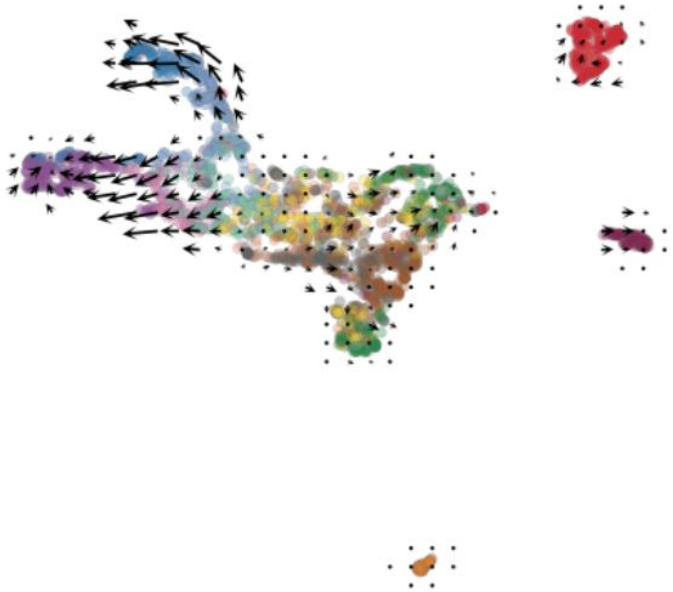


t-SNE:



Our Approach

- To visualize the velocity estimates in low dimensions I used UMAP.



Our Approach

Mean of each cluster:

	UMAP		t-SNE	
	X	Y	X	Y
cell_name				
Endocrine	7.920300	12.386894	-5.923522	11.712447
Enterocyte.Immature.Distal	0.932987	9.163415	7.632138	-7.193831
Enterocyte.Immature.Proximal	0.019843	6.075967	-2.561881	-9.155263
Enterocyte.Mature.Distal	-1.512860	10.624992	7.544077	-13.222531
Enterocyte.Mature.Proximal	-3.142745	6.529403	-2.544090	-15.326616
Enterocyte.Progenitor	3.799403	5.390419	-2.540660	-2.064715
Enterocyte.Progenitor.Early	5.058754	5.400520	-6.801756	0.268642
Enterocyte.Progenitor.Late	2.044651	5.867680	-1.118137	-4.904982
Goblet	17.844820	12.191573	3.391587	15.187263
Paneth	6.943798	4.551712	1.552190	5.141168
Stem	7.234573	4.231500	3.438310	6.247867
TA	6.624190	5.079043	-6.481460	4.739302
TA.Early	5.200008	4.830892	0.401255	1.747746
Tuft	8.971432	-3.583772	12.728454	1.936137

Our Approach

Standard deviation of each cluster:

	UMAP		t-SNE	
	X	Y	X	Y
cell_name				
Endocrine	2.705864	4.000463	4.136861	4.951844
Enterocyte.Immature.Distal	0.759679	1.708573	4.290772	1.952017
Enterocyte.Immature.Proximal	0.790331	0.863974	3.143520	1.609597
Enterocyte.Mature.Distal	1.184959	1.796865	4.035805	2.359452
Enterocyte.Mature.Proximal	1.089266	0.649063	2.206679	2.282305
Enterocyte.Progenitor	1.277071	1.305742	5.337868	2.568934
Enterocyte.Progenitor.Early	1.844621	1.134345	5.542981	3.947807
Enterocyte.Progenitor.Late	1.421981	1.257838	4.160703	2.795687
Goblet	1.874857	0.947054	2.537438	2.374235
Paneth	2.591504	2.849477	4.915822	4.367827
Stem	1.056861	2.953109	4.440063	2.432091
TA	1.237432	1.389774	5.169622	2.864294
TA.Early	1.816618	2.033362	4.122751	3.636204
Tuft	2.032072	5.922331	5.716447	2.176255

Future Plan

RNA Acceleration

- Second order derivative $\frac{d^2 s}{dt^2}$ represent the RNA Acceleration.

$$\frac{d^2 s}{dt^2} = \beta\alpha - \beta(\beta + \gamma)u + \gamma^2 s$$

- To determine RNA Acceleration we have to estimate α.

  ➔ $\alpha = u$

  ➔ α varies according to a trigonometric function (Oscillatory gene expression).

  ➔ Estimate α using time series model on gene expression.

Reference

- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A. and Fan, J., 2018. RNA velocity of single cells. *Nature*, *560*(7719), pp.494-498.

- Bar-Joseph, Z., Gitter, A. and Simon, I., 2012. Studying and modelling dynamic biological processes using time-series gene expression data. *Nature Reviews Genetics*, *13*(8), pp.552-564.


thank
you