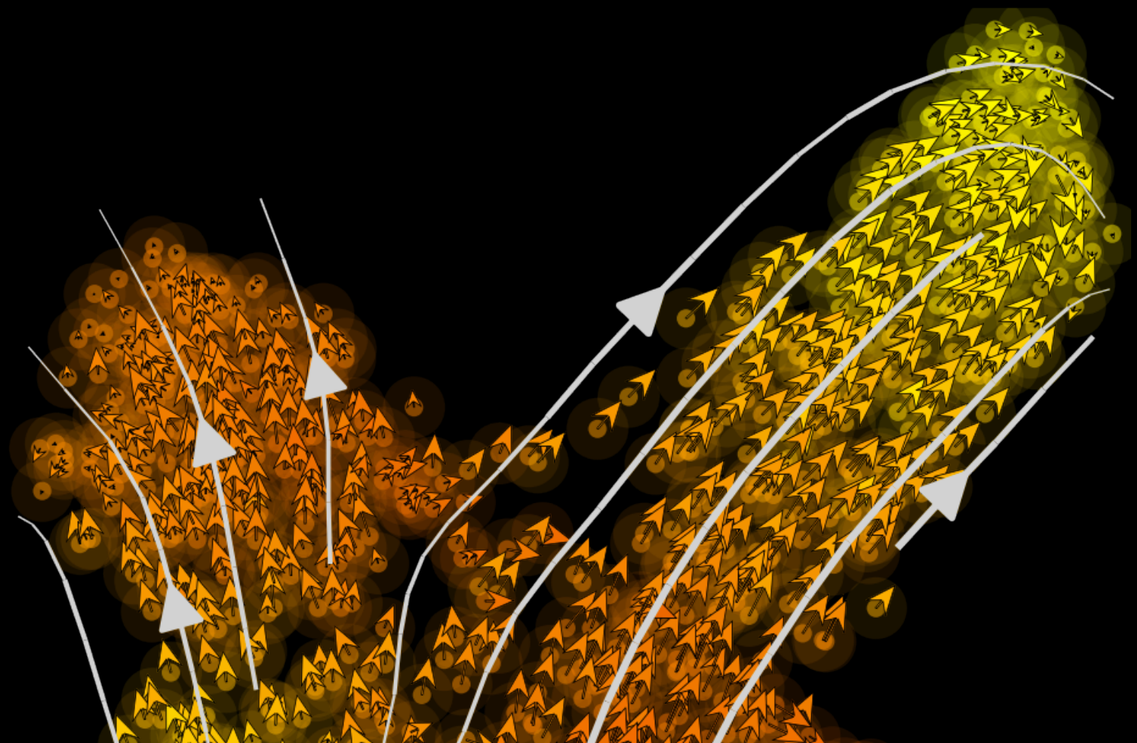
NBIS/SIB single-cell school
RNA velocity

Aug 30, 2021

Volker Bergen

vbergen@cellarity.com



MSc Mathematics & MBA Finance

(2011-2017)

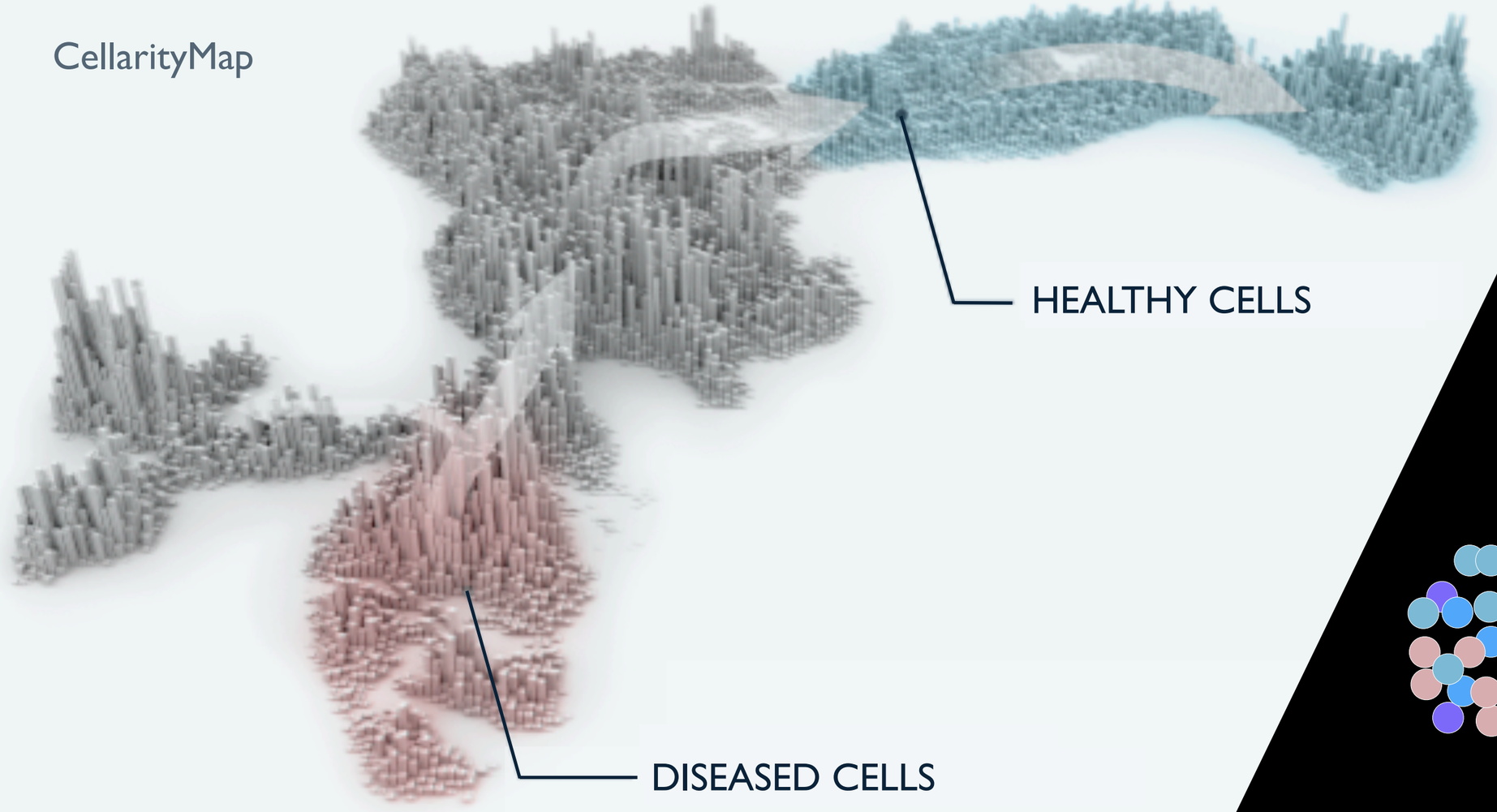
PhD Computational Biology

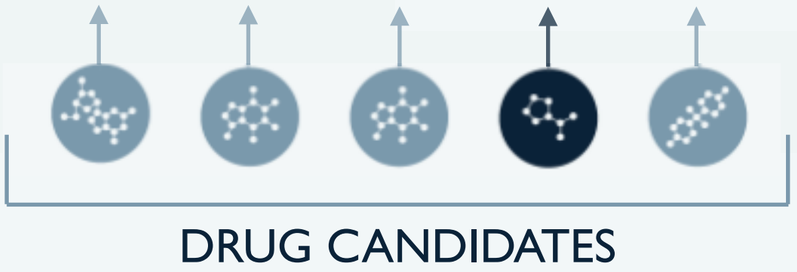
(2017-2020)

ML Scientist & Strategy

(present)

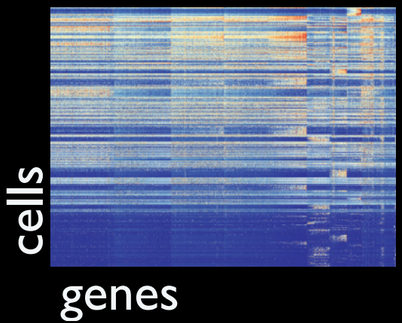
CellarityMap

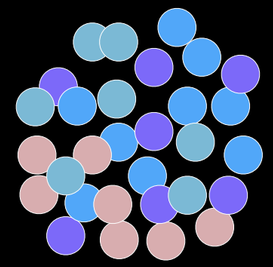


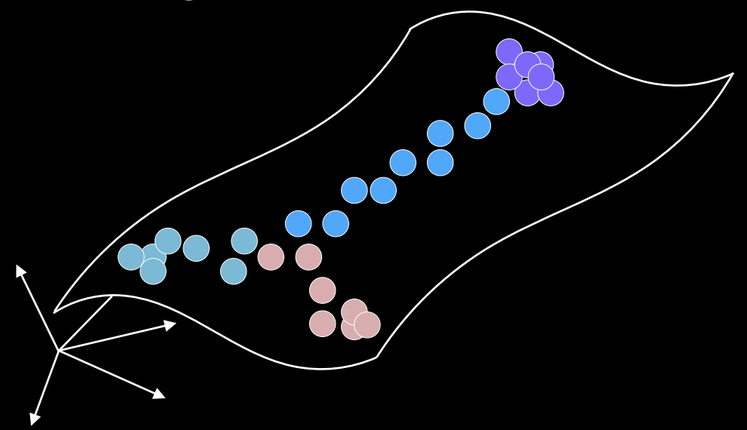


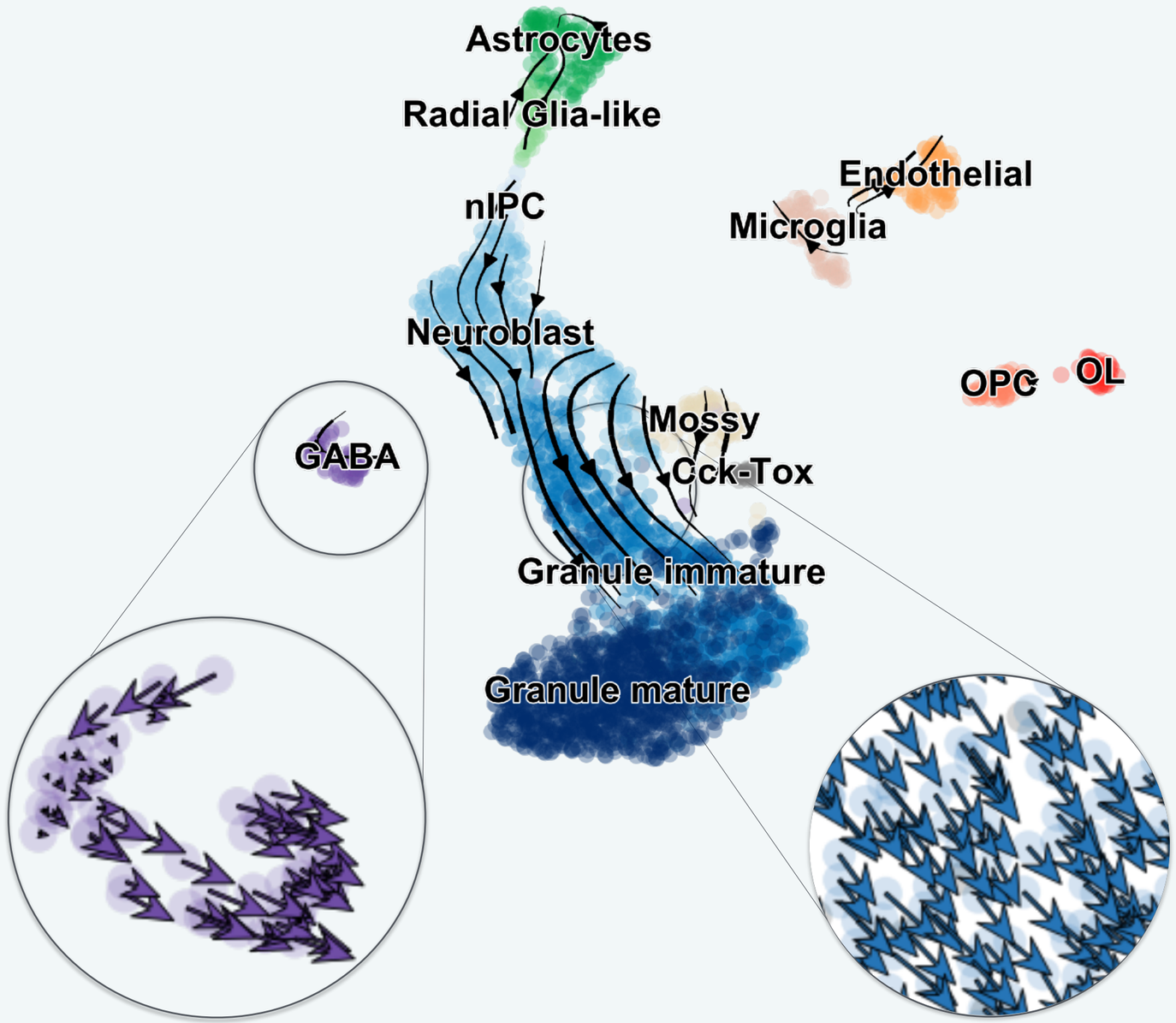
scanpy

Wolf et al. (2018)









RNA velocity

La Manno *et al.* (Nature 2018)

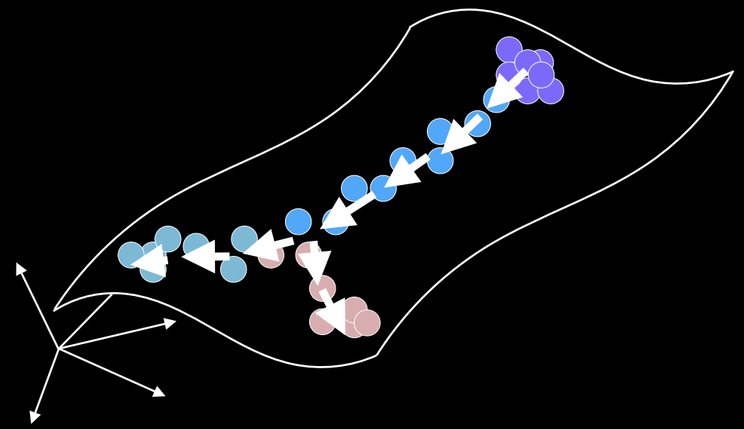
Bergen *et al.* (Nature Biotech 2020)

Reviews

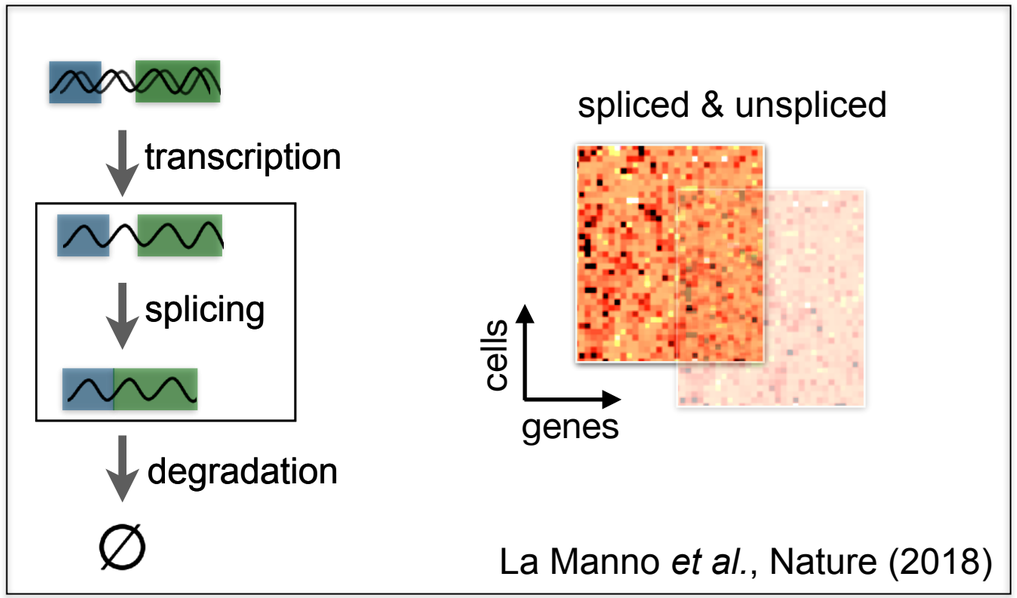
Lederer & La Manno (Nat. Biotech 2020)

Bergen *et al.* (MSB, 2021)

Modeling cellular dynamics with
RNA Velocity

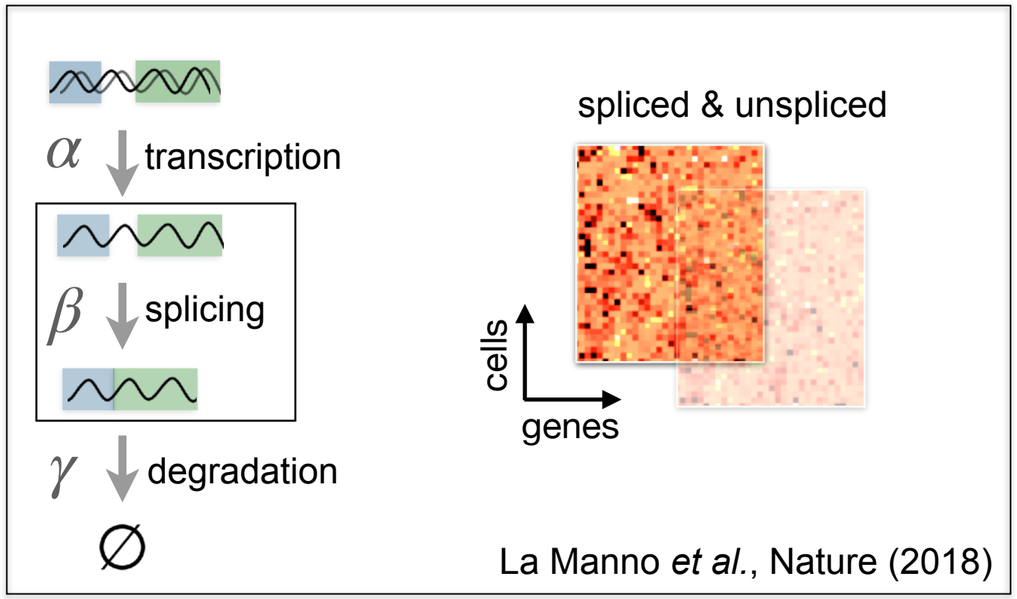


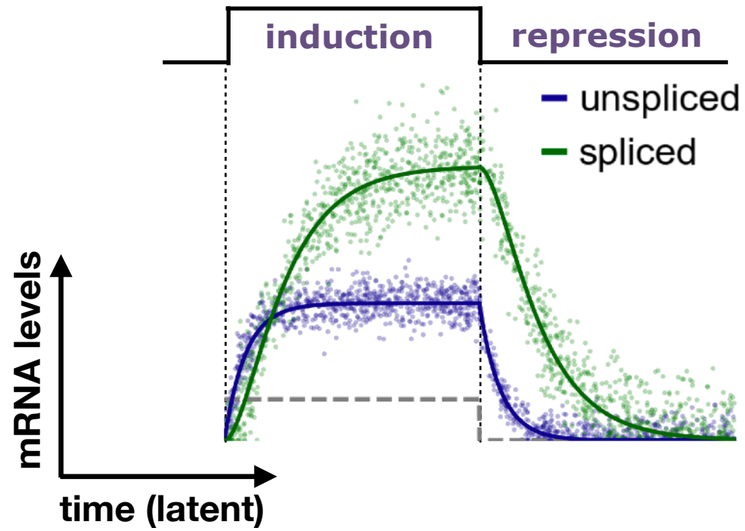
Concept of RNA velocity

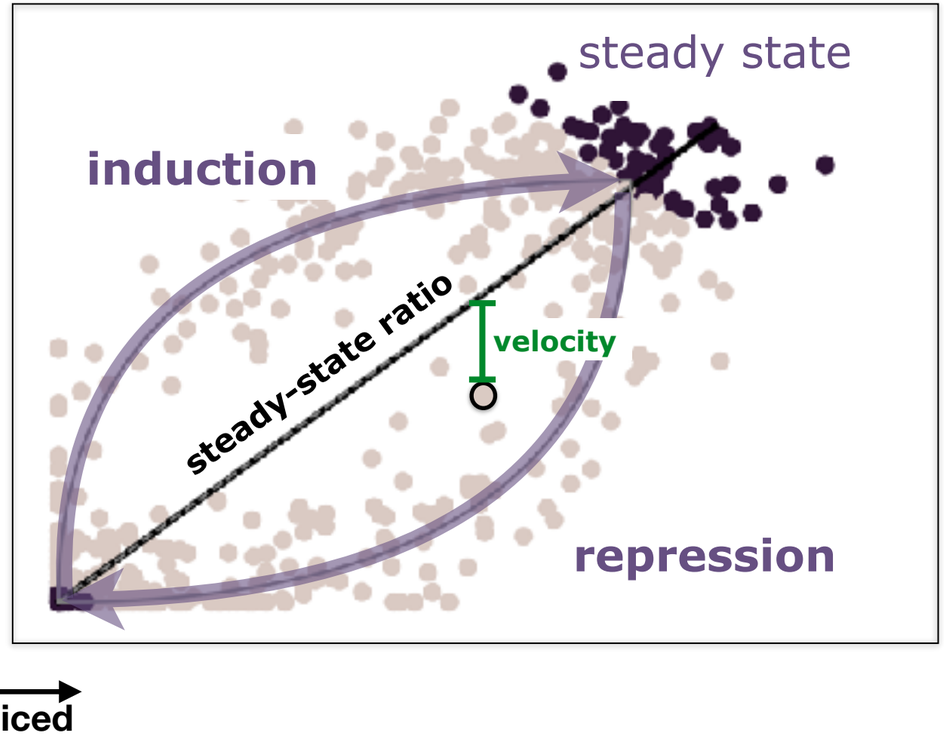


La Manno *et al.*, Nature (2018)

Concept of RNA velocity

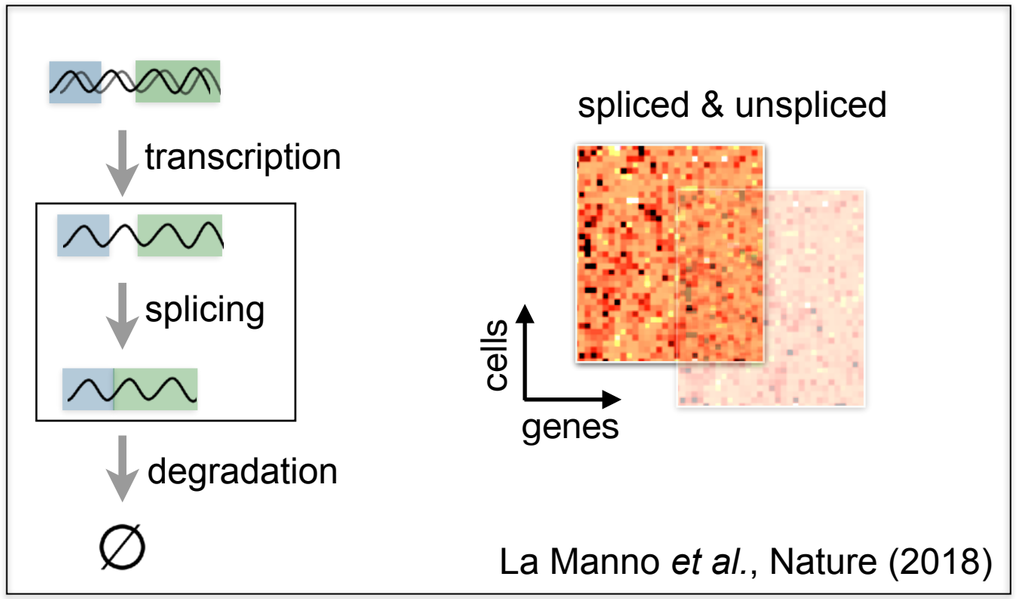


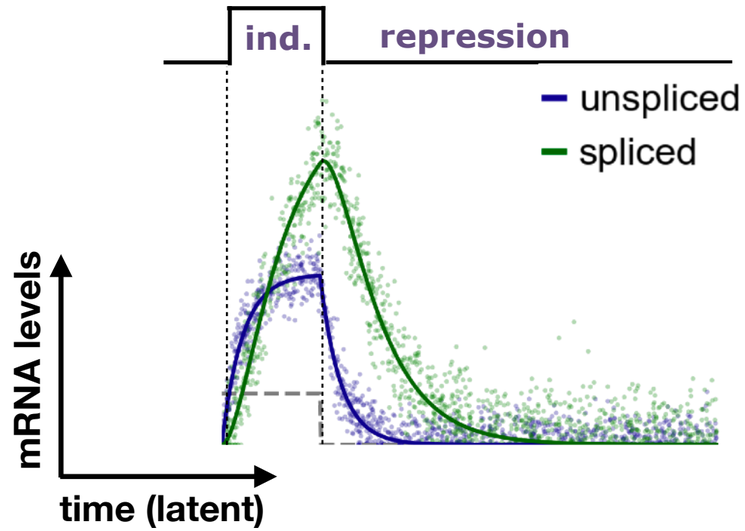


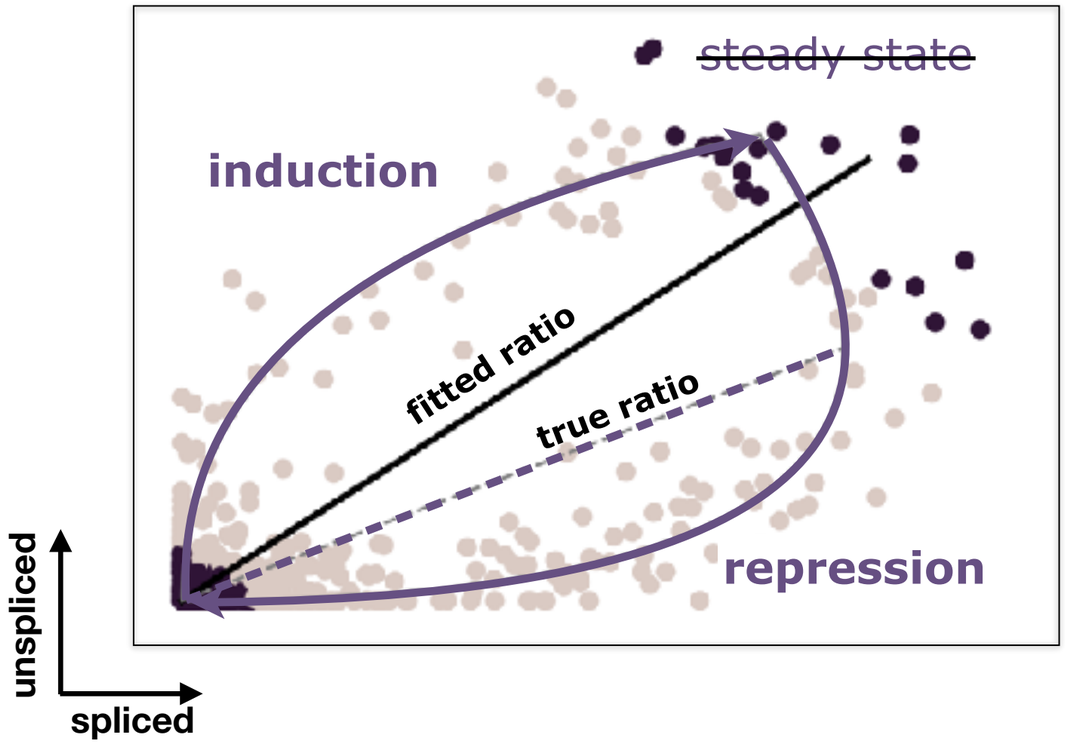


$$\frac{du(t)}{dt} = \alpha - \beta u(t), \qquad \frac{ds(t)}{dt} = \beta u(t) - \gamma s(t)$$

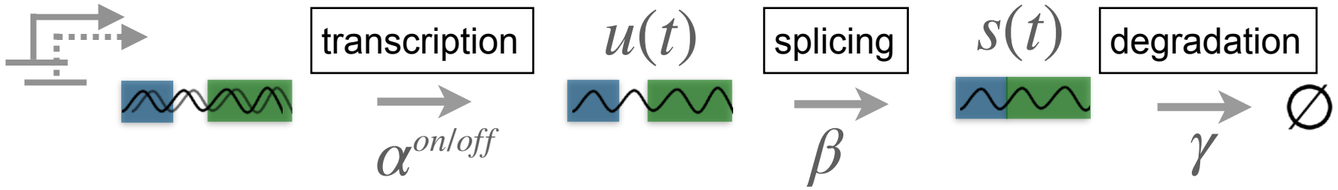
Concept of RNA velocity







RNA velocity generalized through dynamical modeling



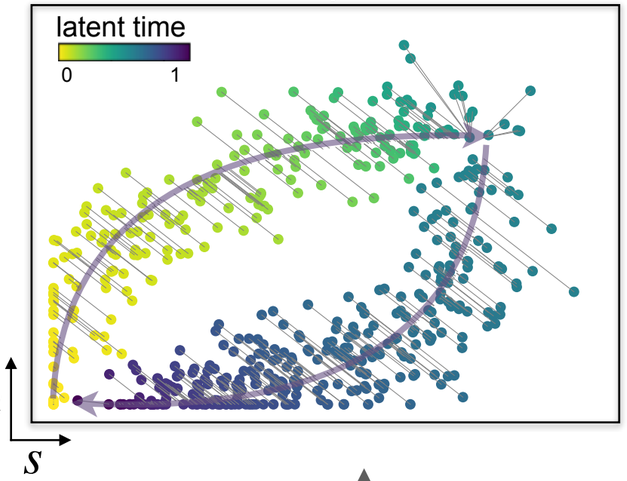
$$u(t) = u_0 e^{-\beta\tau} + \frac{\alpha}{\beta}(1 - e^{-\beta\tau})$$

$$s(t) = s_0 e^{-\gamma\tau} + \frac{\alpha}{\gamma}(1 - e^{-\gamma\tau}) + \frac{\alpha - \beta u_0}{\gamma - \beta}(e^{-\gamma\tau} - e^{-\beta\tau}) \qquad \tau = t - t_0$$

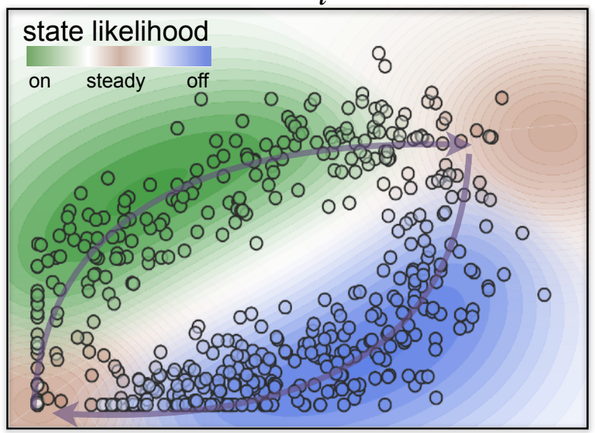
parameters of **reaction rates** $\theta = (\alpha^{off}, \alpha^{on}, \beta, \gamma)$

cell-specific **latent variables** (switch, time, state) $\eta_i = (t_0^{(i)}, t_i, k_i)$

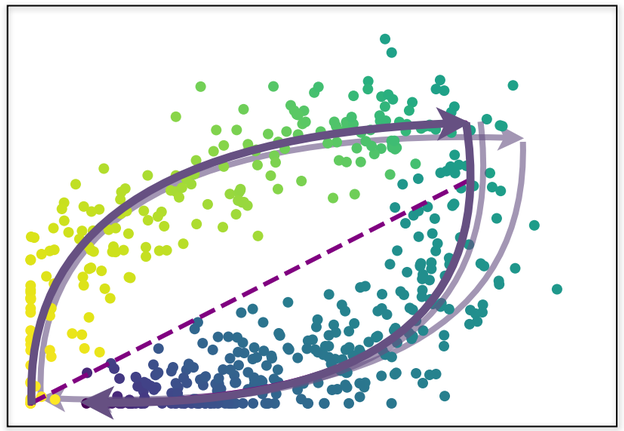
time assignment t_i



state assignment k_i

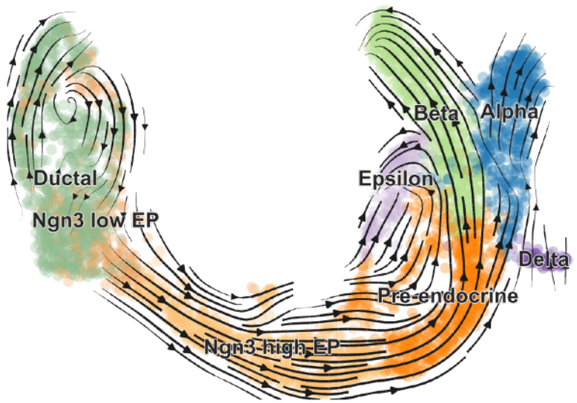


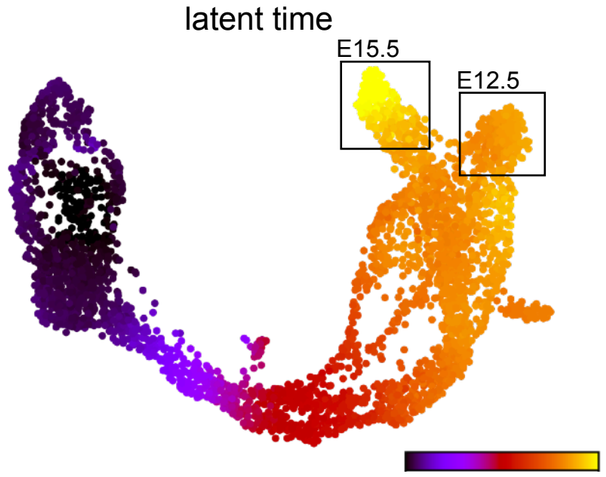
parameter update $\theta = (\alpha^{on/off}, \beta, \gamma)$

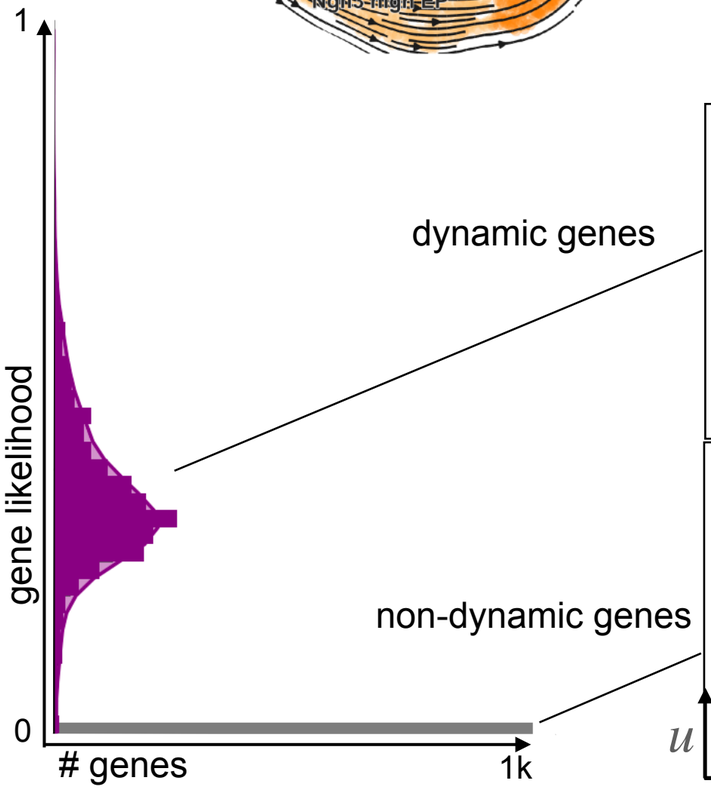


inference by maximizing joint likelihood $P\left((u_i, s_i) \,|\, (\theta, t_i, k_i)\right)$

Applications of RNA velocity

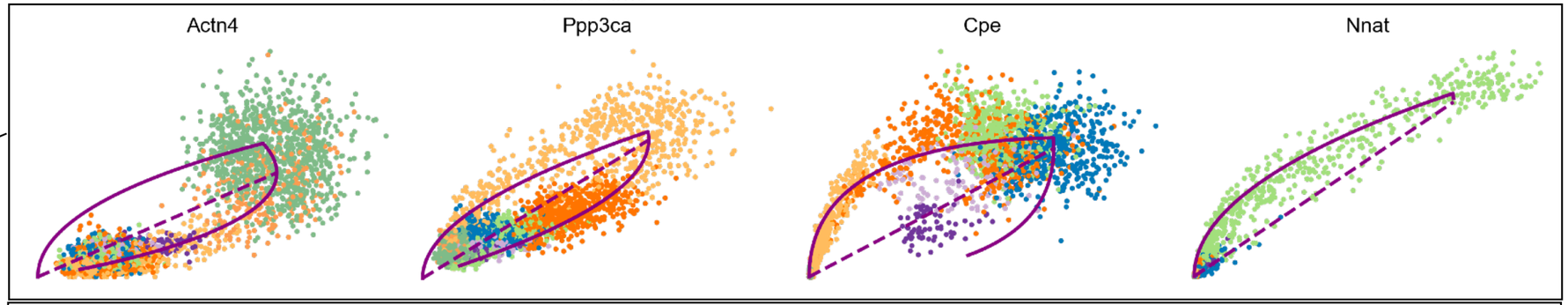


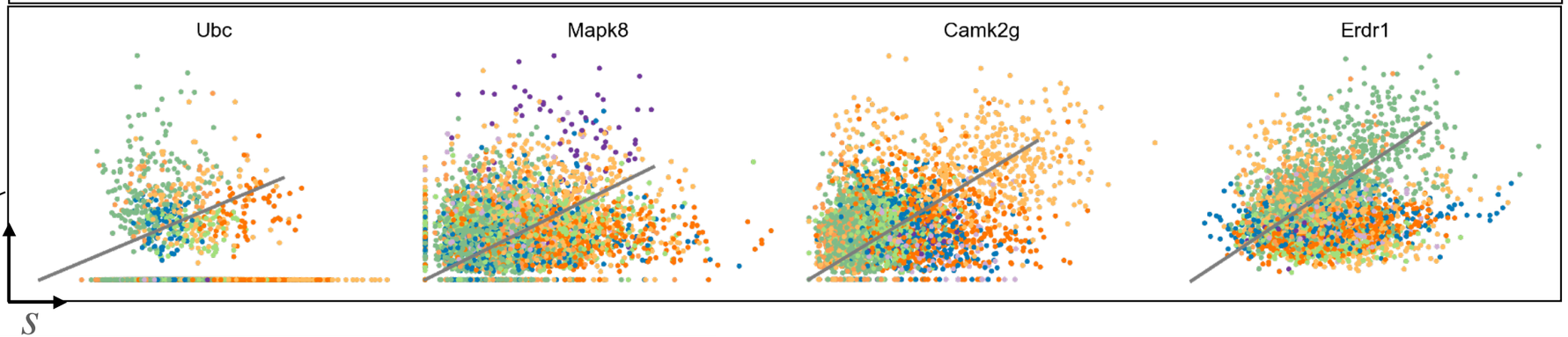




dynamic genes

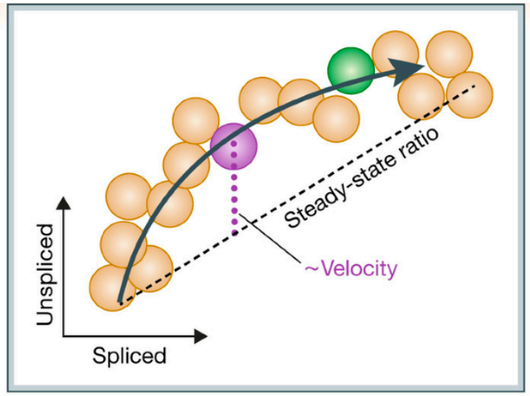
non-dynamic genes

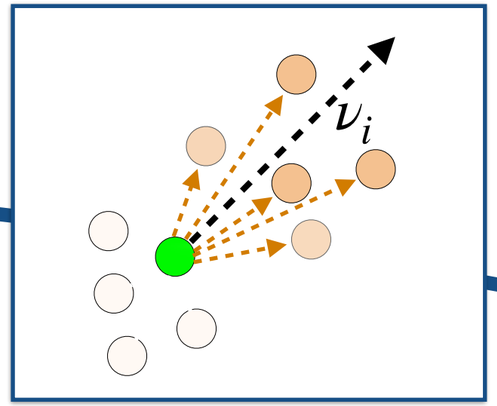




RNA velocity in two dimensions

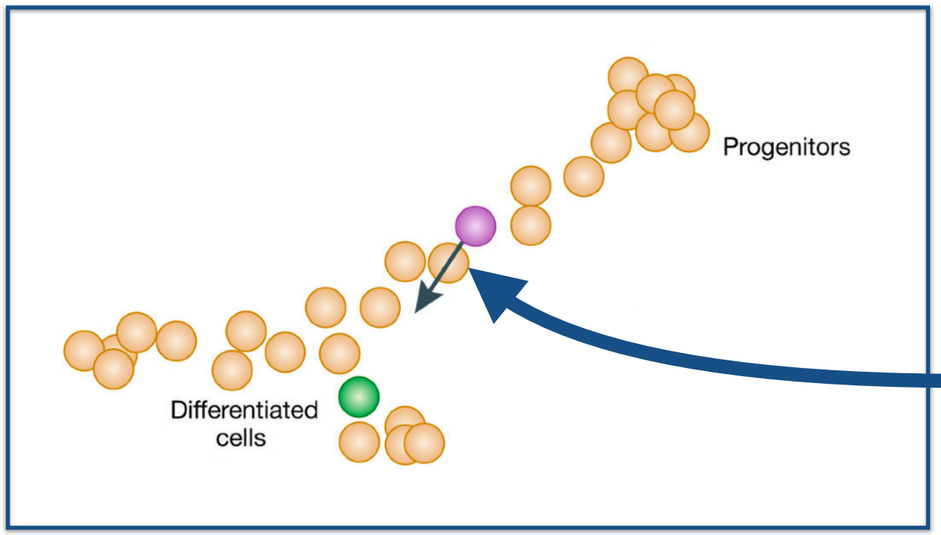
Compute cell-to-cell transition probabilities by how much the transition correlates with the velocity vector (high-dim)



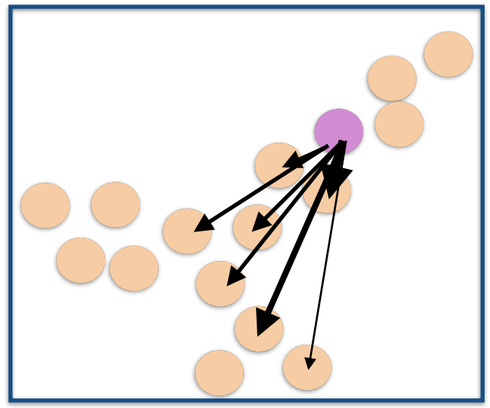


$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & \dots & P_{1,j} & \dots & P_{1,S} \\ P_{2,1} & P_{2,2} & \dots & P_{2,j} & \dots & P_{2,S} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ P_{i,1} & P_{i,2} & \dots & P_{i,j} & \dots & P_{i,S} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ P_{S,1} & P_{S,2} & \dots & P_{S,j} & \dots & P_{S,S} \end{bmatrix}.$$

$$P_{ij} = e^{\rho(\delta_{ij},\nu_i)\ /\ \sigma_i^2}$$



UMAP / TSNE

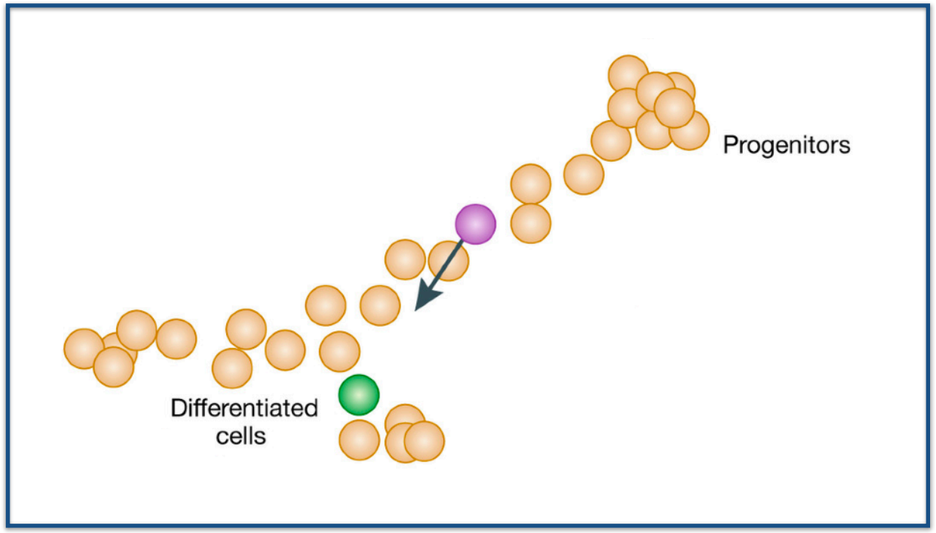


Given the transitions probabilities, compute the expected transition in the lower-dim embedding

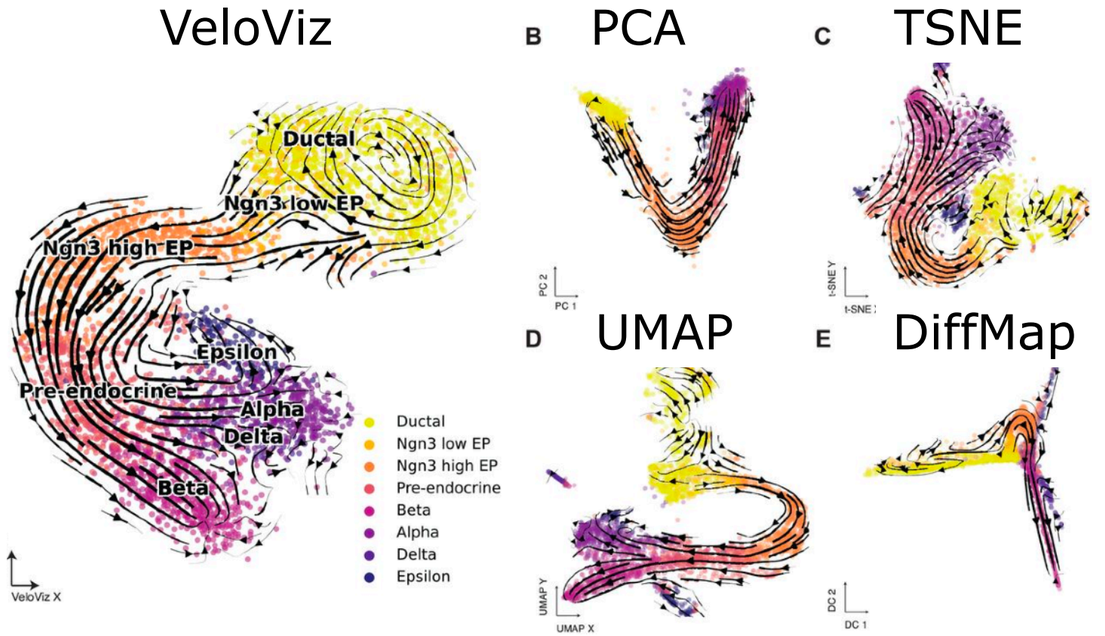
RNA velocity - finding a good representation

Topic 1

- What embedding specification best represents the high-dim vector field?
- Can we find other ways to project the data (e.g., parametric UMAP)?



UMAP / TSNE

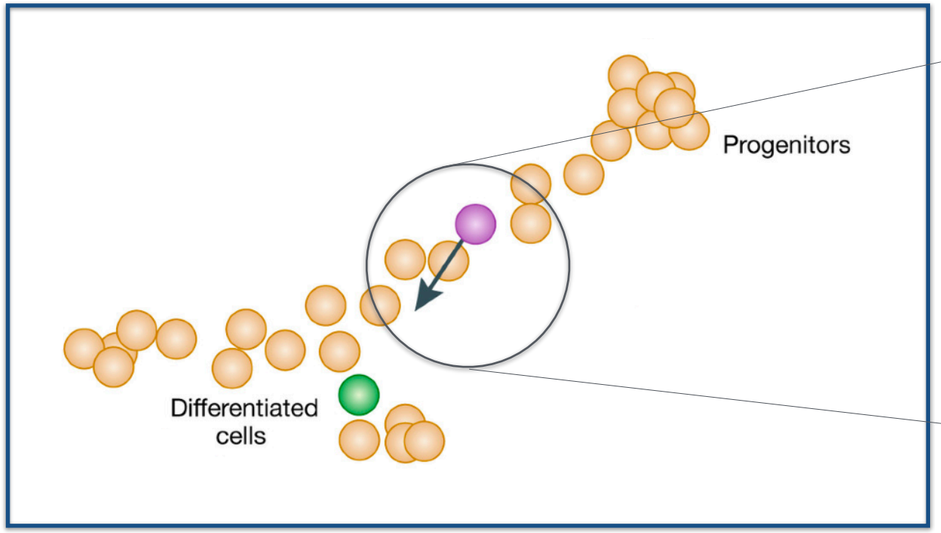


Atta *et al.* (2021, BioRxiv)

RNA velocity - identifying relevant genes

Topic 2

- What genes are driving the projected arrows in the low-dim manifold?
- Can we systematically identify genes that are important in a particular compartment?



UMAP / TSNE

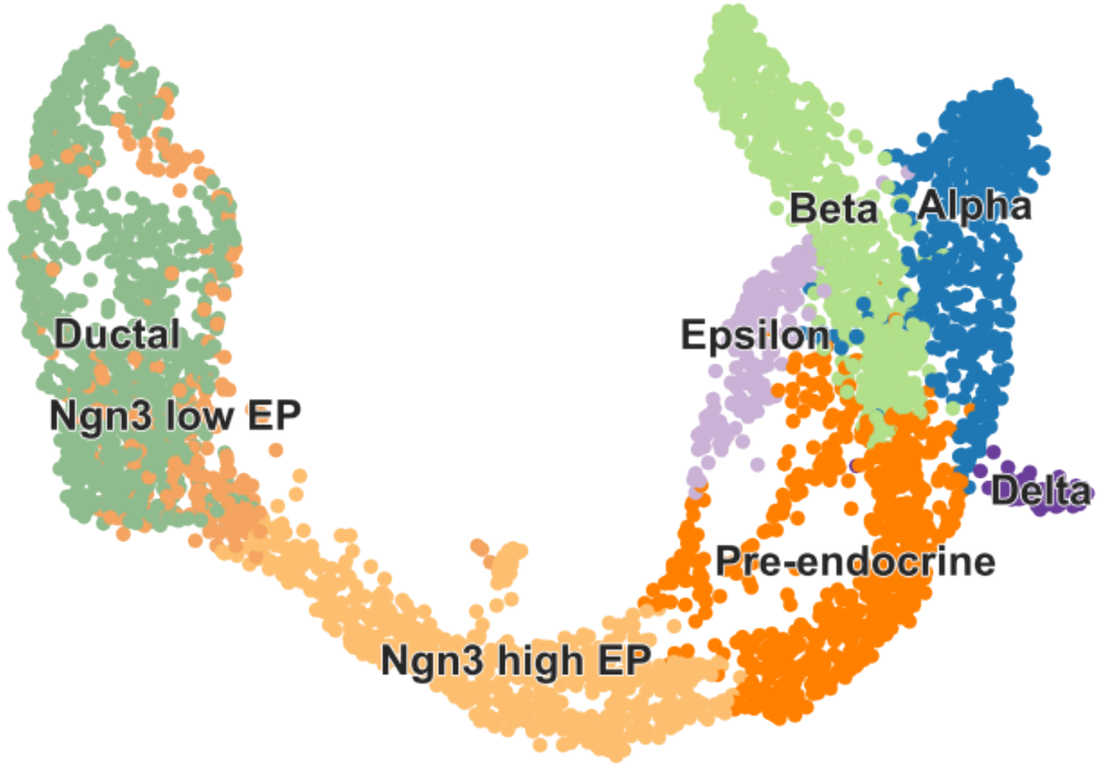
Cell type A

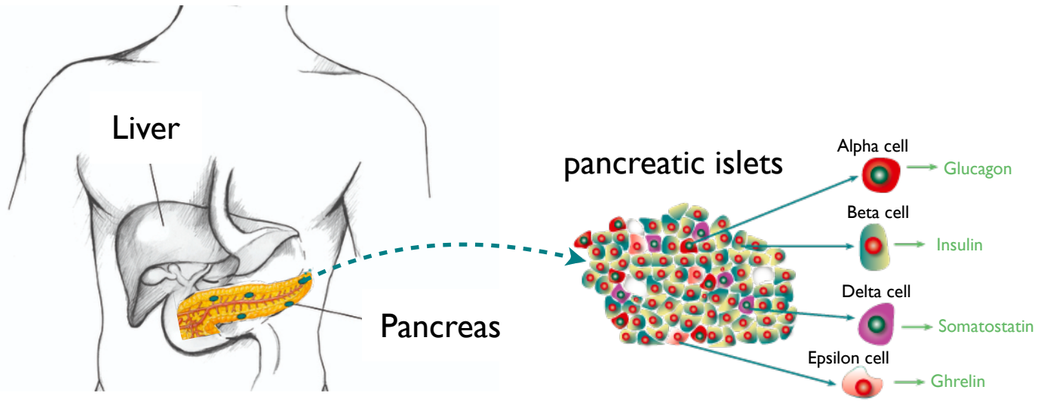
Genes ranked	some metric
Gene XX	*0.85*
Gene XY	*0.72*
….	
Gene XZ	*0.02*

Cell type B

Genes ranked	some metric
Gene AA	*0.91*
Gene AB	*0.83*
….	
Gene AZ	*0.05*

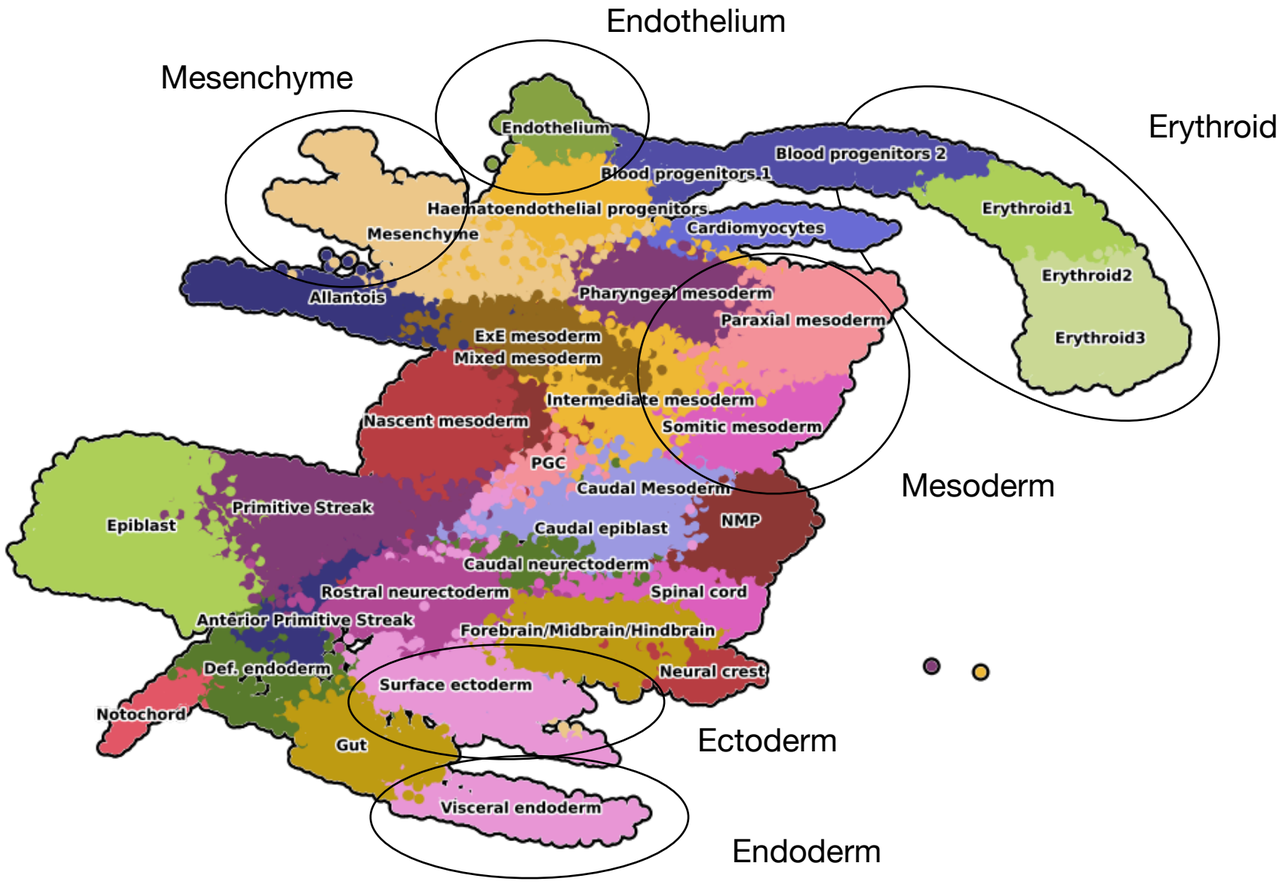
Datasets - pancreatic endocrinogenesis

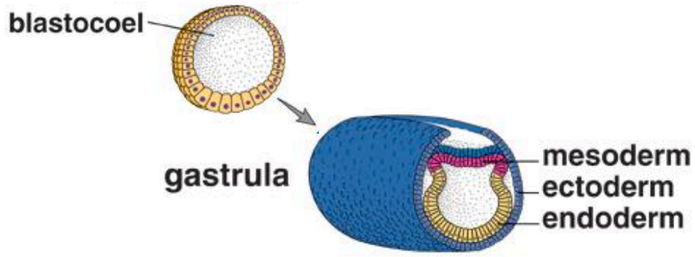




Bastidas-Ponce *et al.,* Development (2019)
Bergen *et al.,* Nat. Biotech (2020)

Datasets - gastrulation and early organogenesis





Pijuan-Sala *et al.,* Nature (2019)

The scvelo API

```
import scvelo as scv
```

Read the data

```
adata = scv.datasets.pancreas()
```

Preprocessing

```
scv.pp.filter_and_normalize(adata, n_top_genes=2000)
scv.pp.moments(adata)
```

Velocity estimation

```
scv.tl.velocity(adata)
```

Velocity projection

```
scv.tl.velocity_embedding(adata, basis='umap')
```

Visualization

```
scv.pl.velocity_embedding_stream(adata, basis='umap')
```