Manifold-constrained RNA velocity estimation with VeloCycle
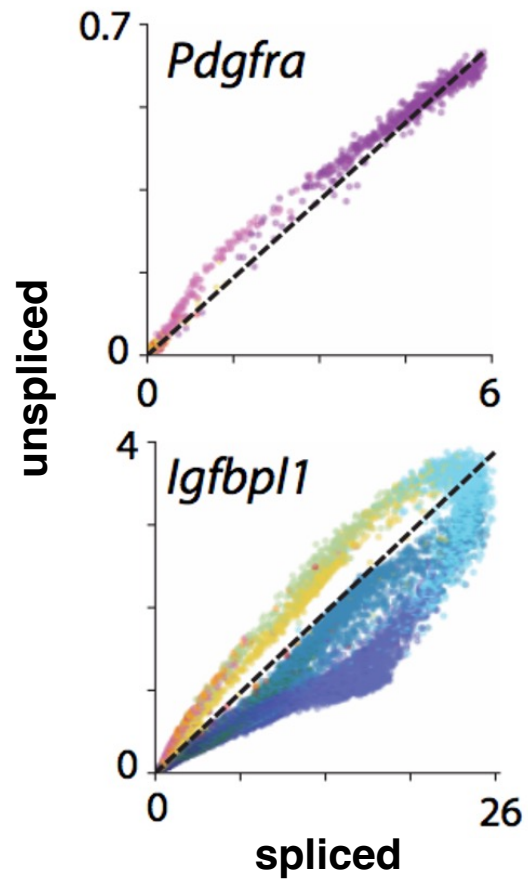
Single Cell Transcriptomics in Python

Alex Lederer

Presentation Overview

- VeloCycle and probabilistic modeling

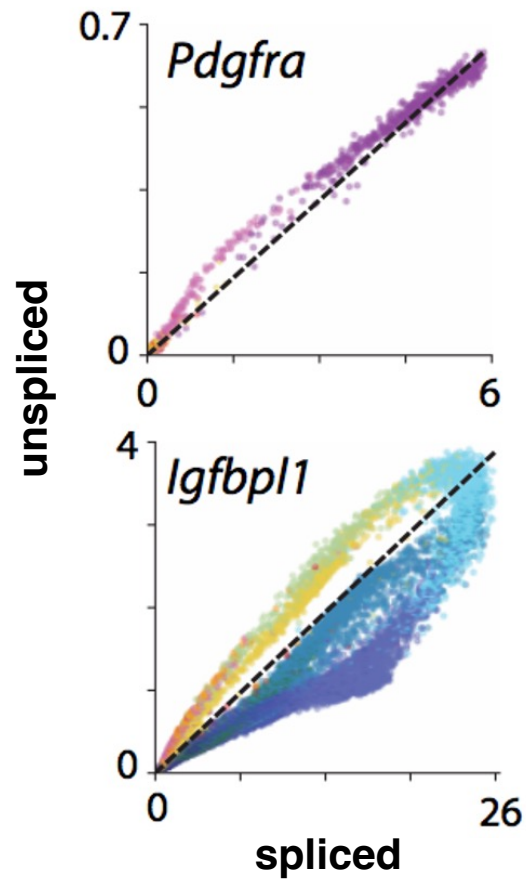- Exercise on using VeloCycle for cycling pancreatic ductal cells

The RNA velocity framework has several limitations
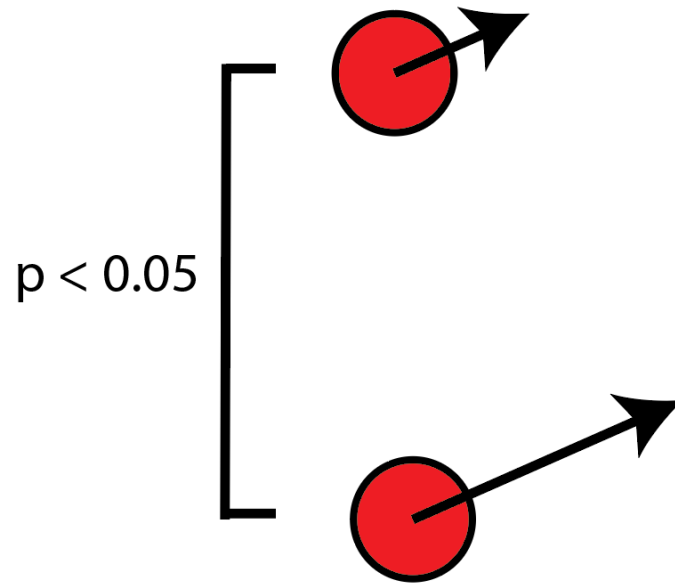
Necessity to run
smoothing methods

The RNA velocity framework has several limitations
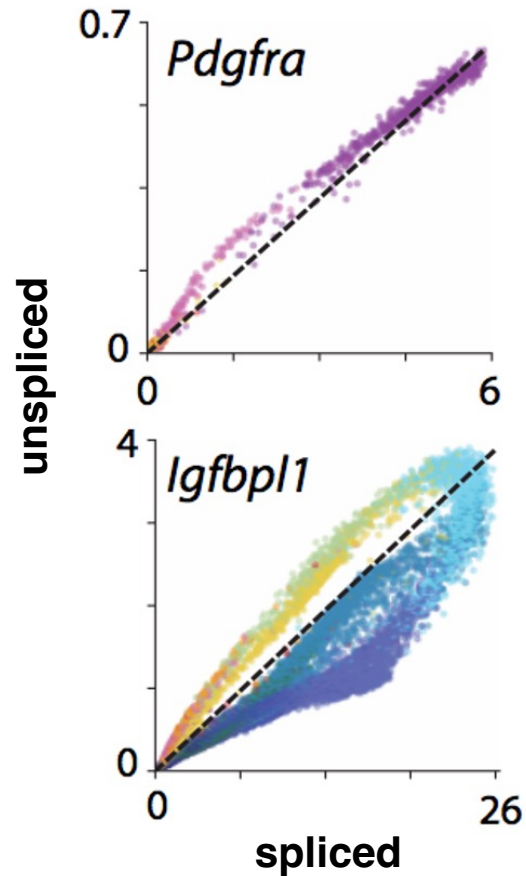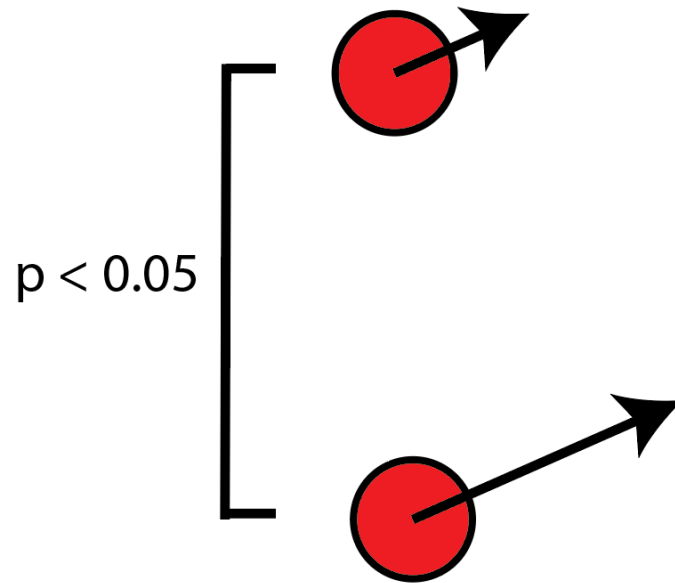
Necessity to run
smoothing methods

Inference on velocity
is currently not possible

$p < 0.05$

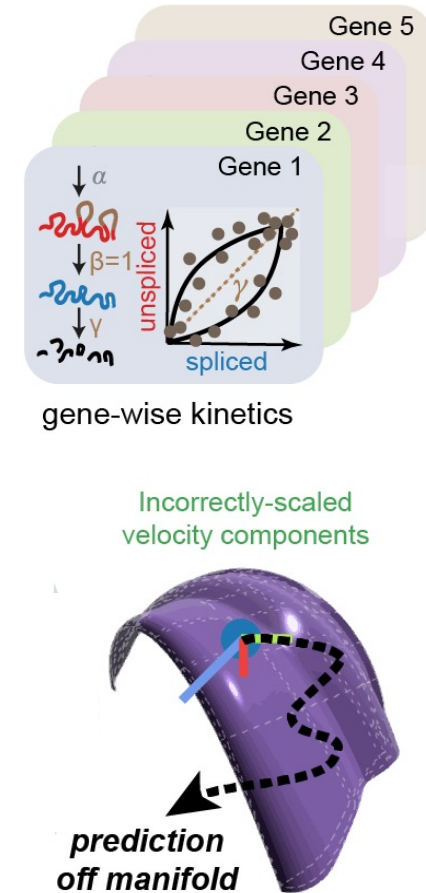# The RNA velocity framework has several limitations

Necessity to run smoothing methods

Inference on velocity is currently not possible

Gene-wise velocity estimates yield incorrectly scaled components

$p < 0.05$

gene-wise kinetics

Incorrectly-scaled velocity components

prediction off manifold

Incorrectly scaled gene-wise velocity is not consistent to the true gene expression manifold

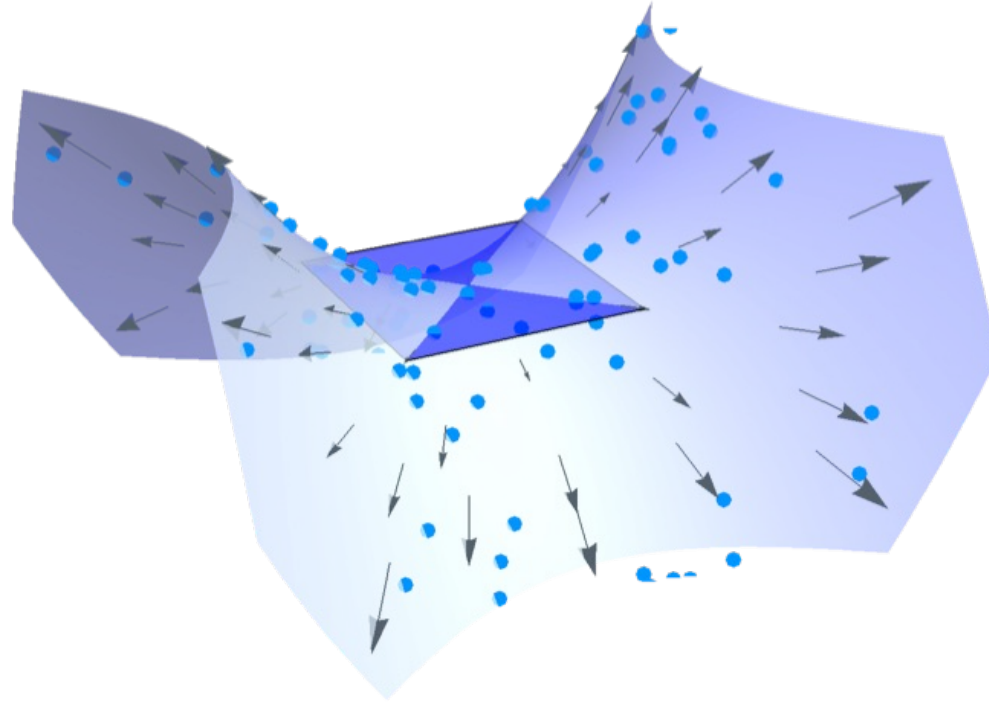# Incorrectly scaled gene-wise velocity is not consistent to the true gene expression manifold

Incorrectly scaled gene-wise velocity is not consistent to the true gene expression manifold

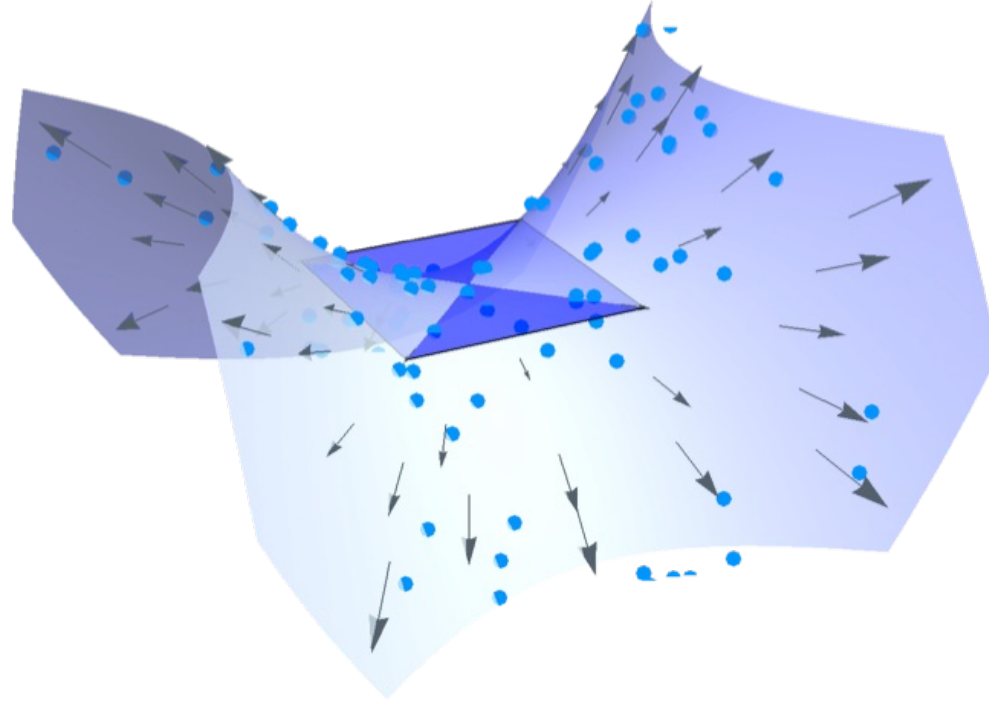*image credit: Gioele La Manno*
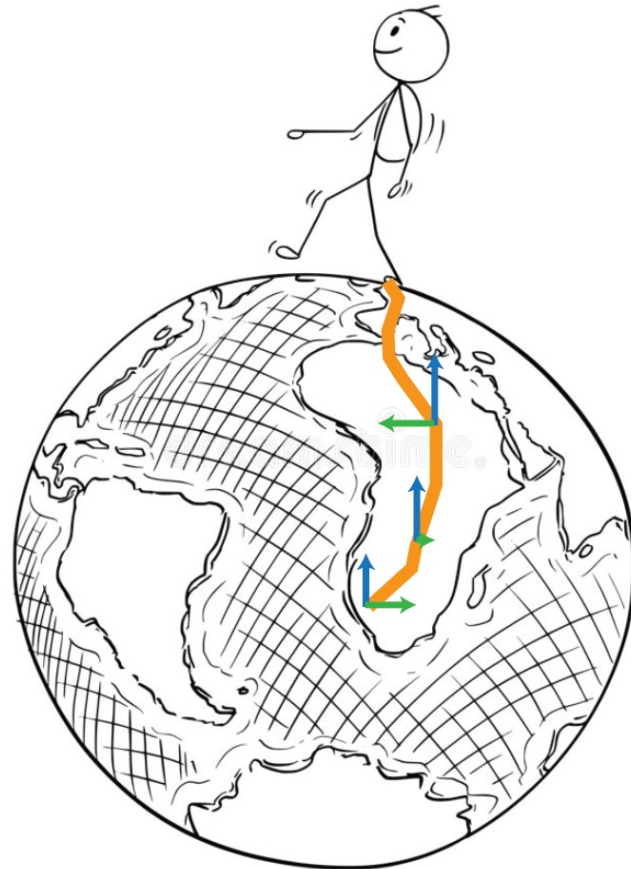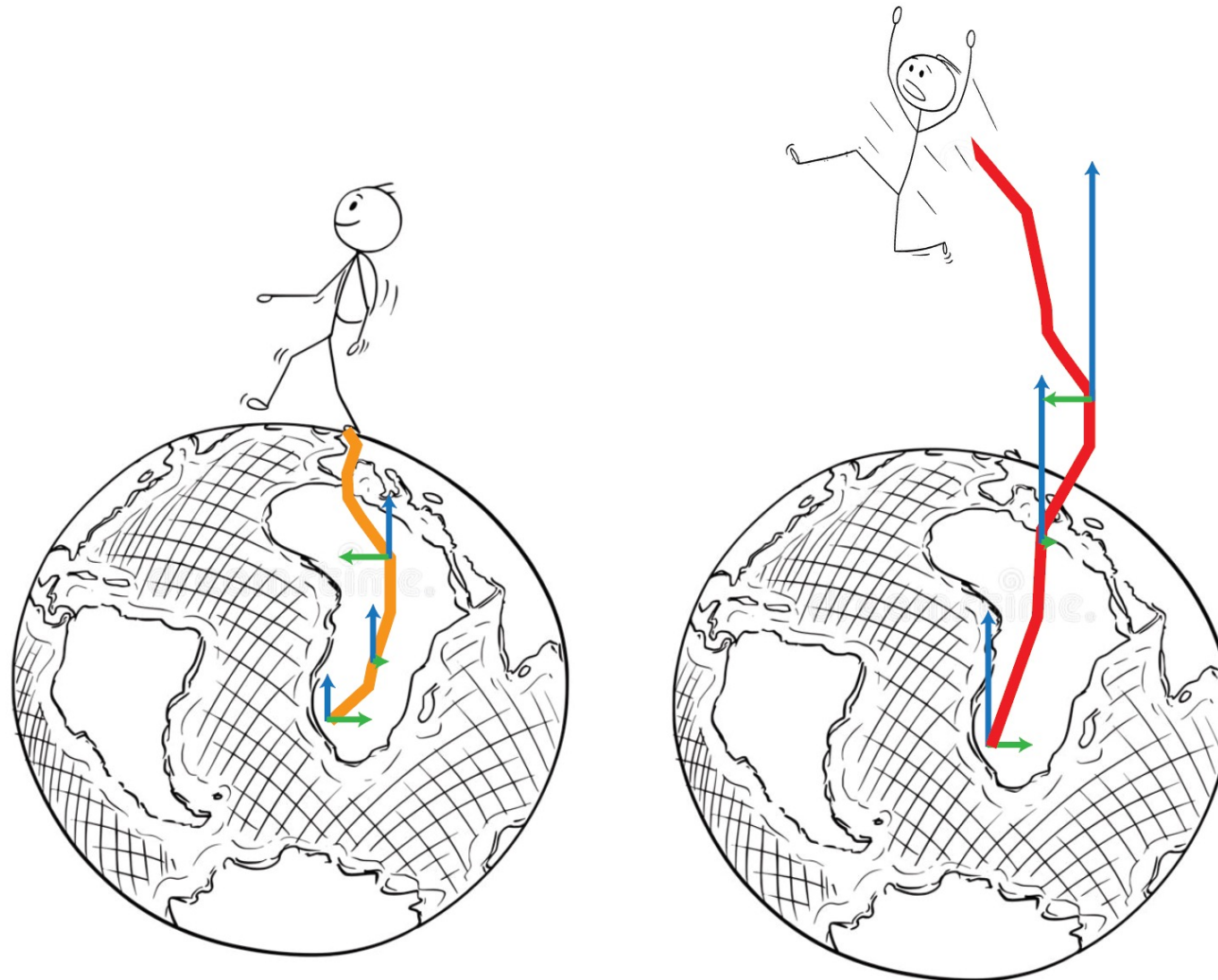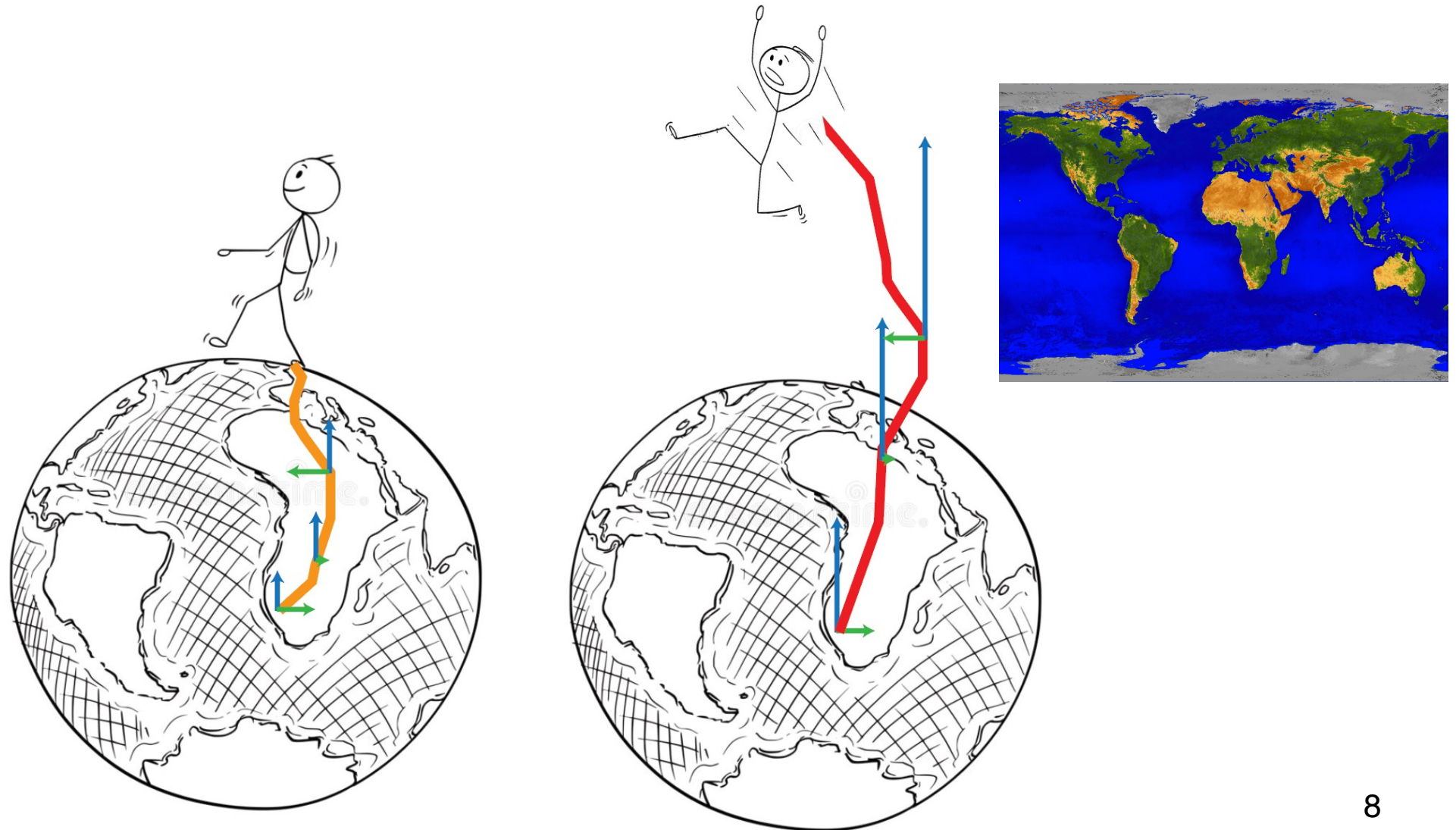
The true gene expression manifold is high-dimensional

The true gene expression manifold is high-dimensional


Can we write a system of differential equations linking the manifold and velocity?
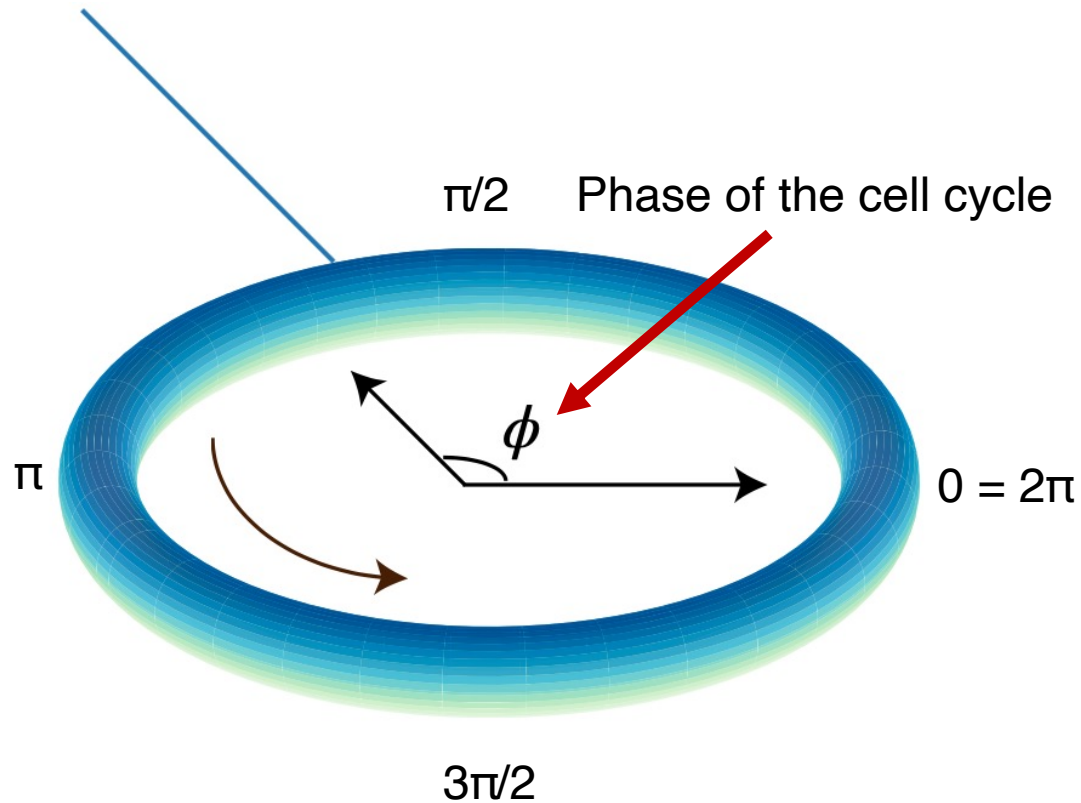
What is our true gene expression manifold?

RNA velocity can be reformulated to obtain a cell- and gene-independent cell cycle speed

1-dimensional circular manifold

π/2 Phase of the cell cycle

ϕ

π

0 = 2π

3π/2

RNA velocity can be reformulated to obtain a cell- and gene-independent cell cycle speed

1-dimensional circular manifold

π/2   Phase of the cell cycle

ϕ

π   0 = 2π

3π/2
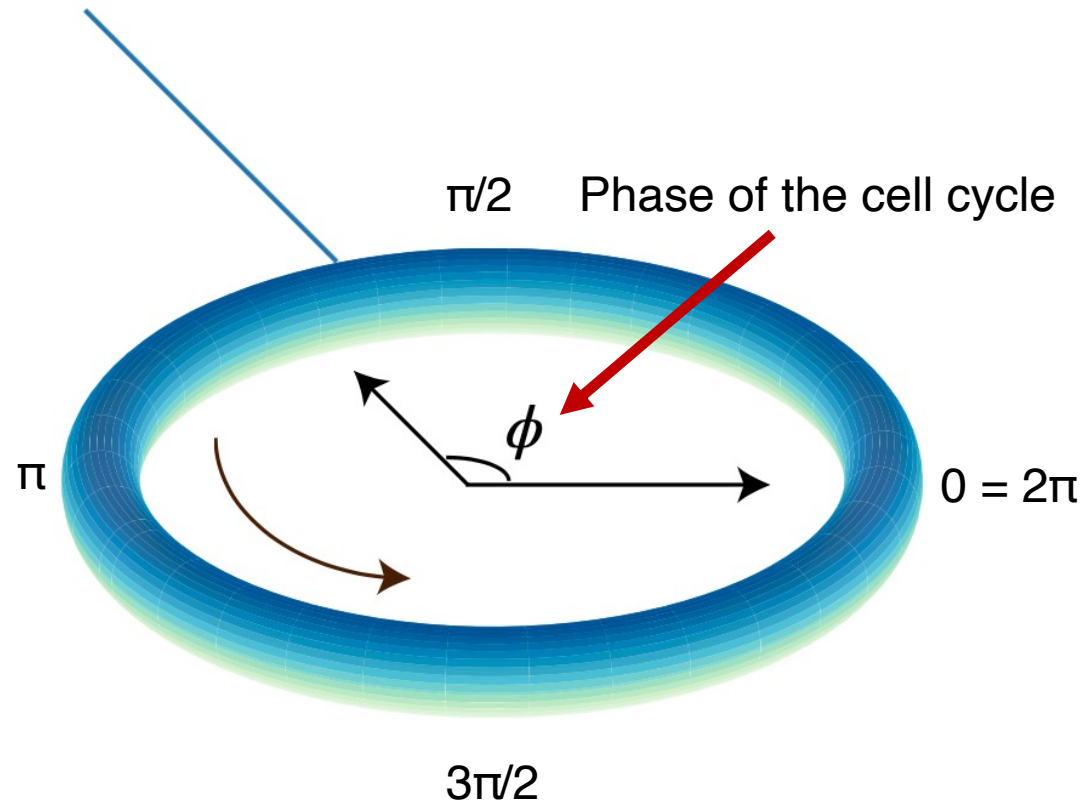
$$\frac{ds}{dt} = \beta u - \gamma s$$

becomes

$$\nabla_\phi s(\phi) \cdot \omega(\phi) = \beta u(\phi) - \gamma s(\phi)$$

Speed   Phase

12

Graphical representation of VeloCycle probabilistic model

- Each random variable is sampled from a pre-defined probability distribution
- **Bayesian inference** aims to learn the mean and variance of each variable's distribution
- Similar types of models have been developed for other single-cell challenges, including data integration and clustering.
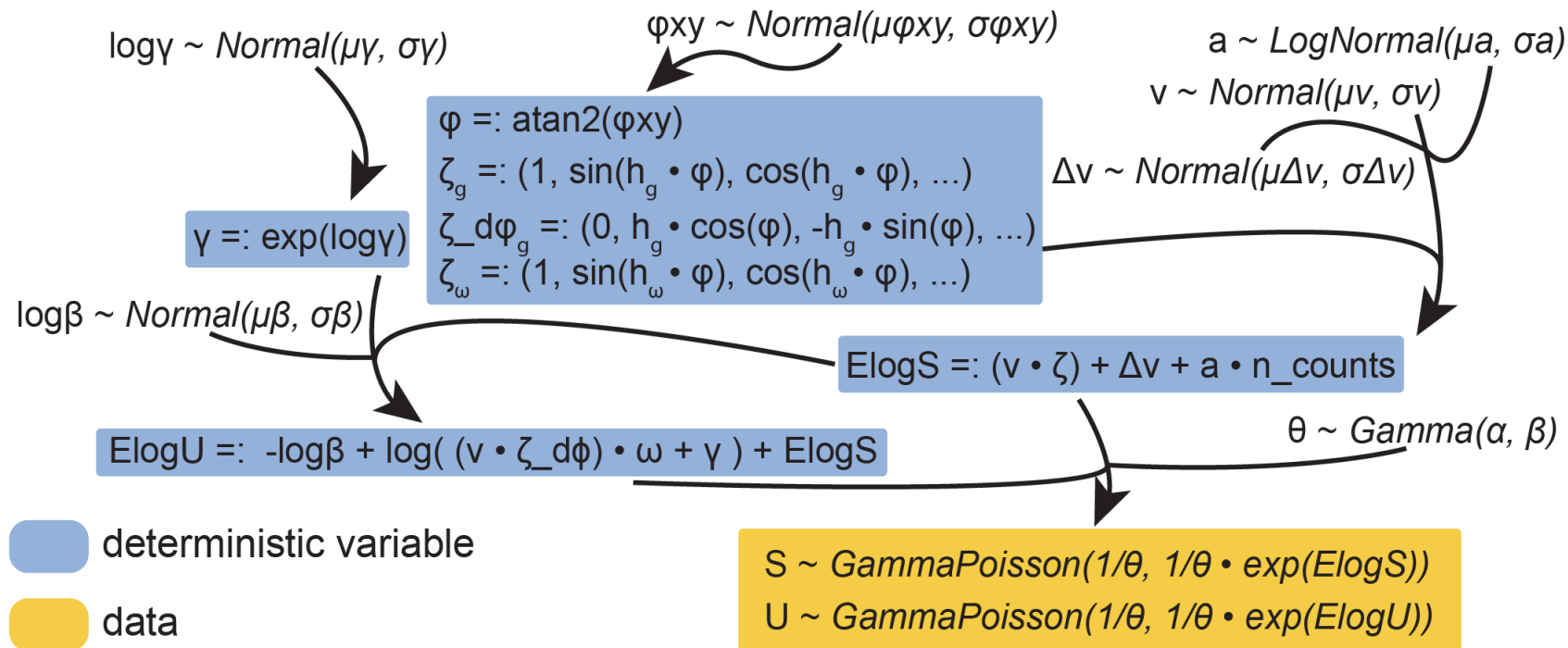
Graphical representation of VeloCycle probabilistic model
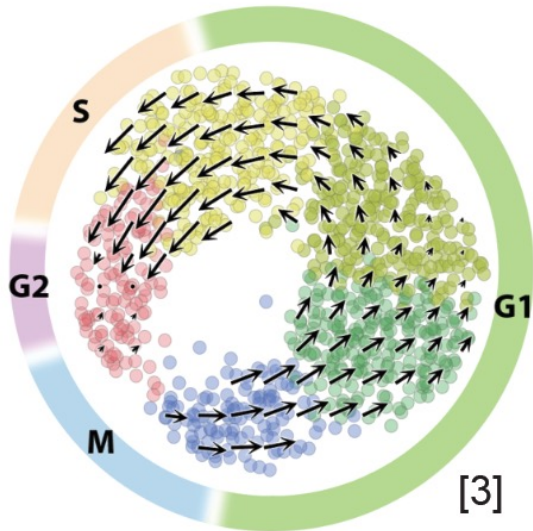

Simplified graphical model
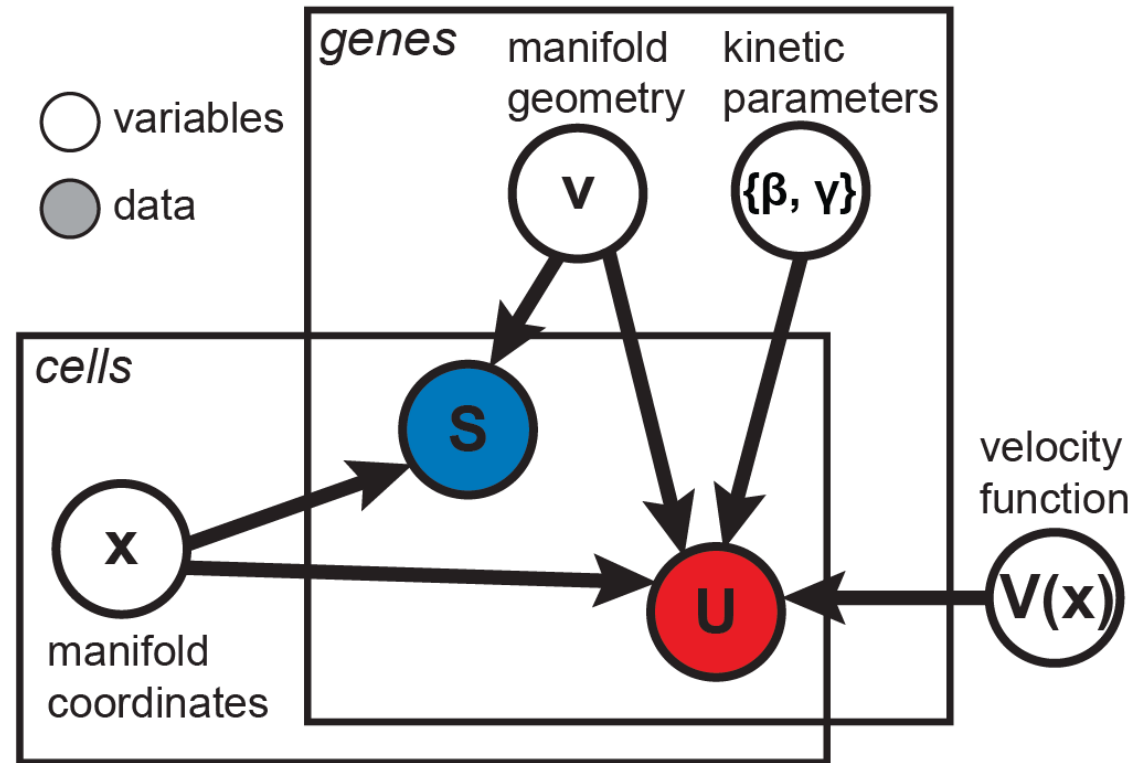
S ~ NegativeBinomial(s(x, v), θ)

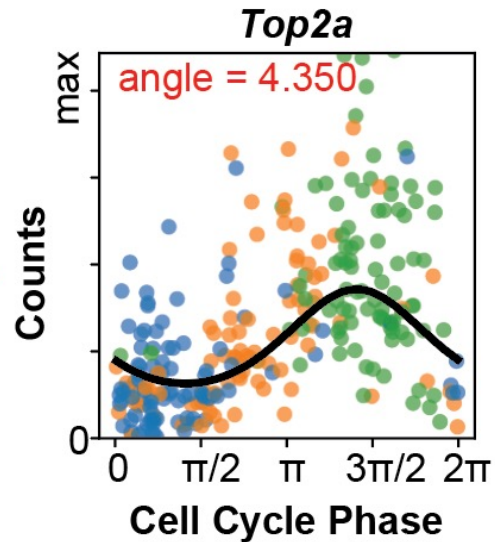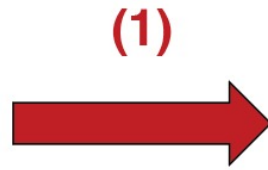U ~ NegativeBinomial(u(x, v, β, γ, V(·)), θ)

14

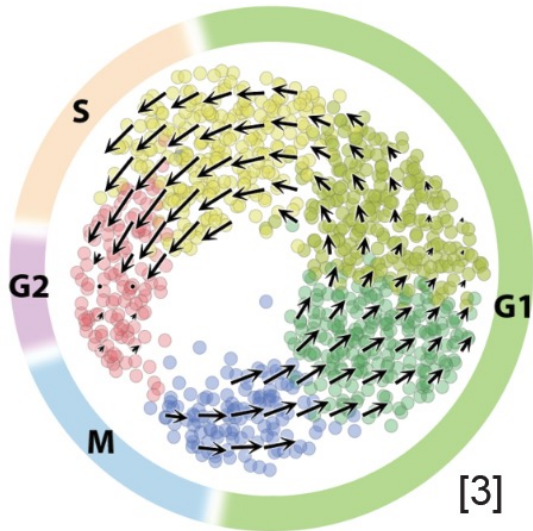# VeloCycle is a multivariate periodic velocity model
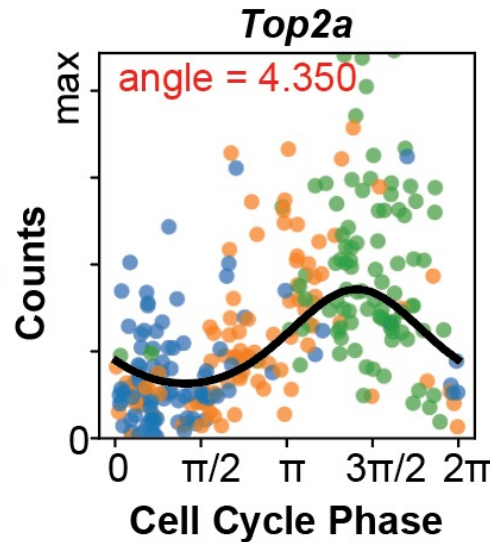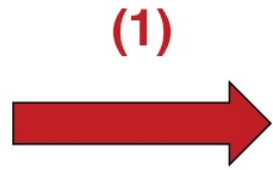

"manifold learning"

VeloCycle is a multivariate periodic velocity model

(1) Fit cell cycle phase and gene harmonics

(2) Fit cell cycle velocity and kinetic parameters

scRNA-seq dataset

[3]

(1)

Top2a

angle = 4.350

Counts

max

0

0 π/2 π 3π/2 2π

Cell Cycle Phase

FACS label
- G1
- S
- G2M

(2)

Scaled Velocity

0.7

0.6

0.5

0 π/2 π 3π/2 2π

Cell Cycle Phase
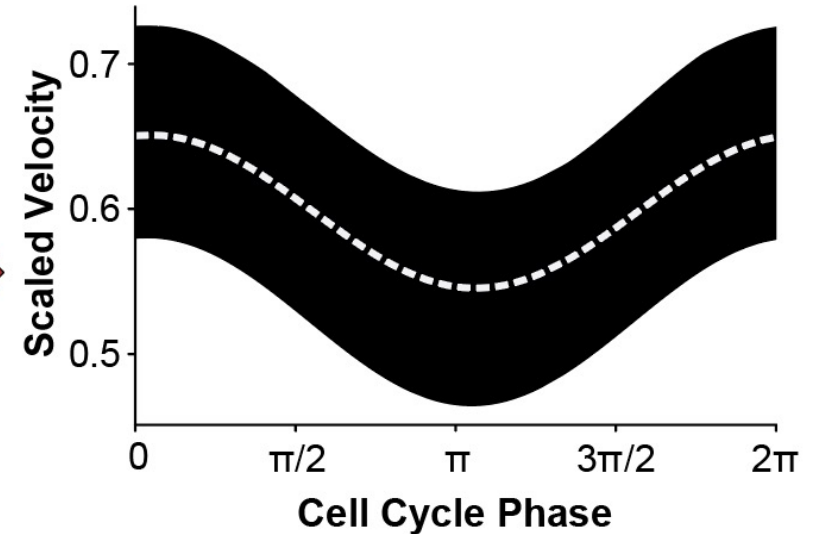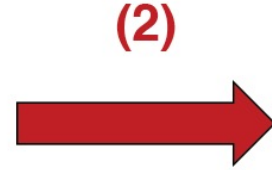
Convert velocity to hours time scale
Perform velocity statistical tests
Transfer learning of gene manifold

"manifold learning"

"velocity learning"

16

VeloCycle is a versatile framework for manifold-constrained cell cycle velocity estimation

Statistical testing
between conditions
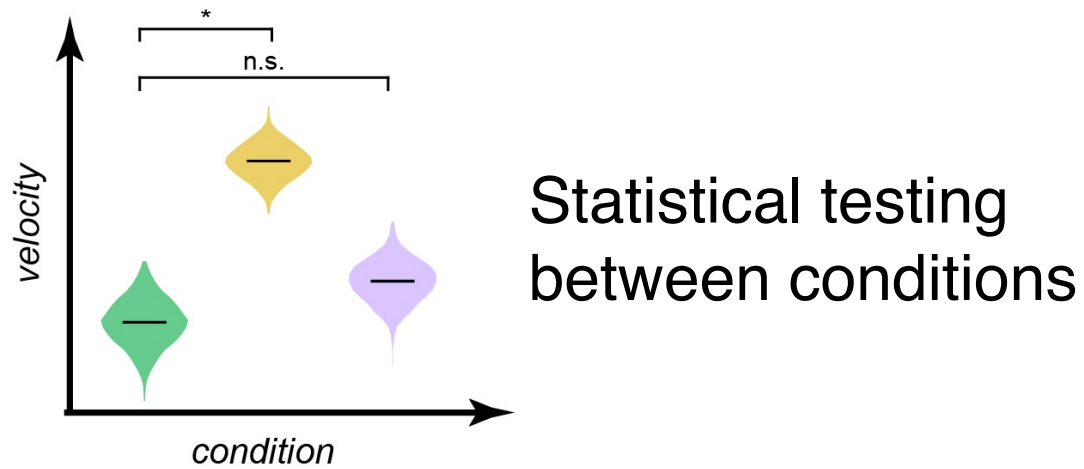
VeloCycle is a versatile framework for manifold-constrained cell cycle velocity estimation
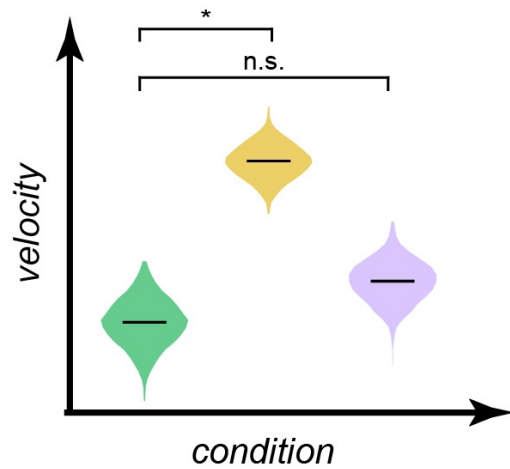
Statistical testing between conditions
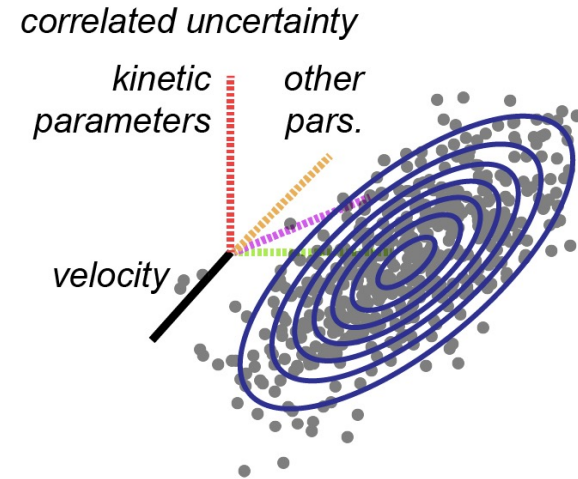
Structured posterior between velocity and kinetic parameters

VeloCycle is a versatile framework for manifold-constrained cell cycle velocity estimation

Statistical testing between conditions

Structured posterior between velocity and kinetic parameters

Experimental validation
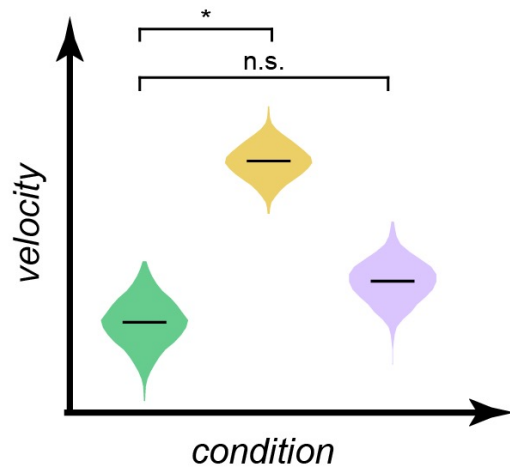
VeloCycle is a versatile framework for manifold-constrained cell cycle velocity estimation
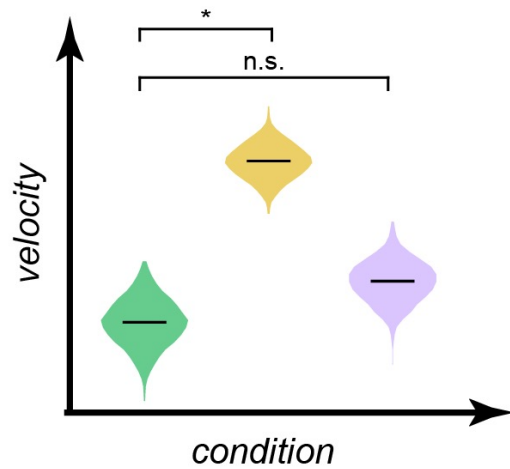
Statistical testing between conditions

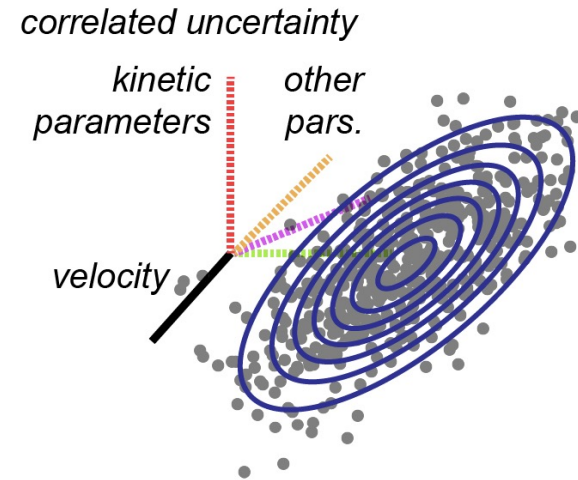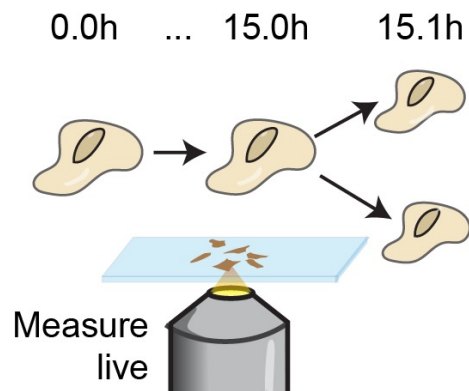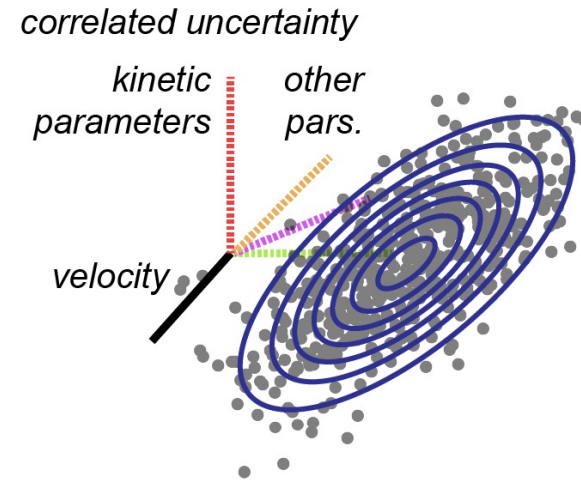Structured posterior between velocity and kinetic parameters
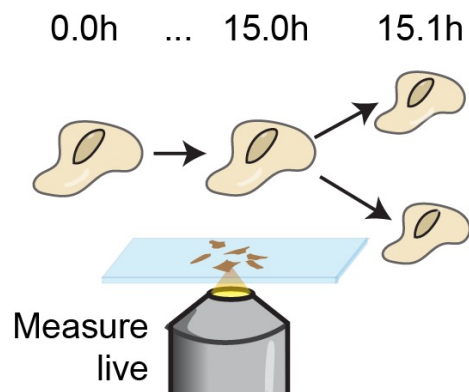
Experimental validation
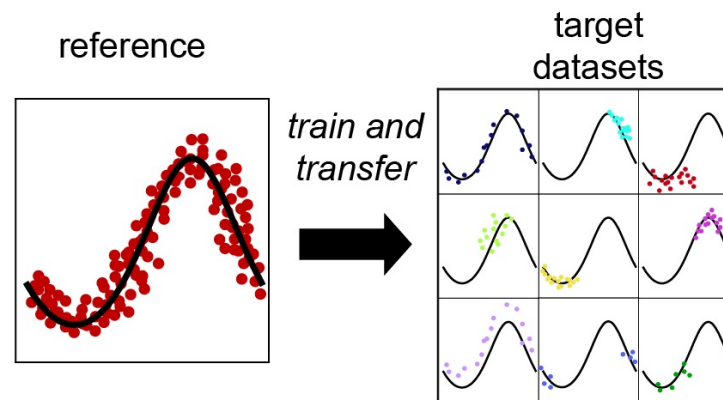
Transfer learning

Does manifold-learning recover accurate cell cycle phases?

Phase assignments can be validated with FACS-sorted mESCs
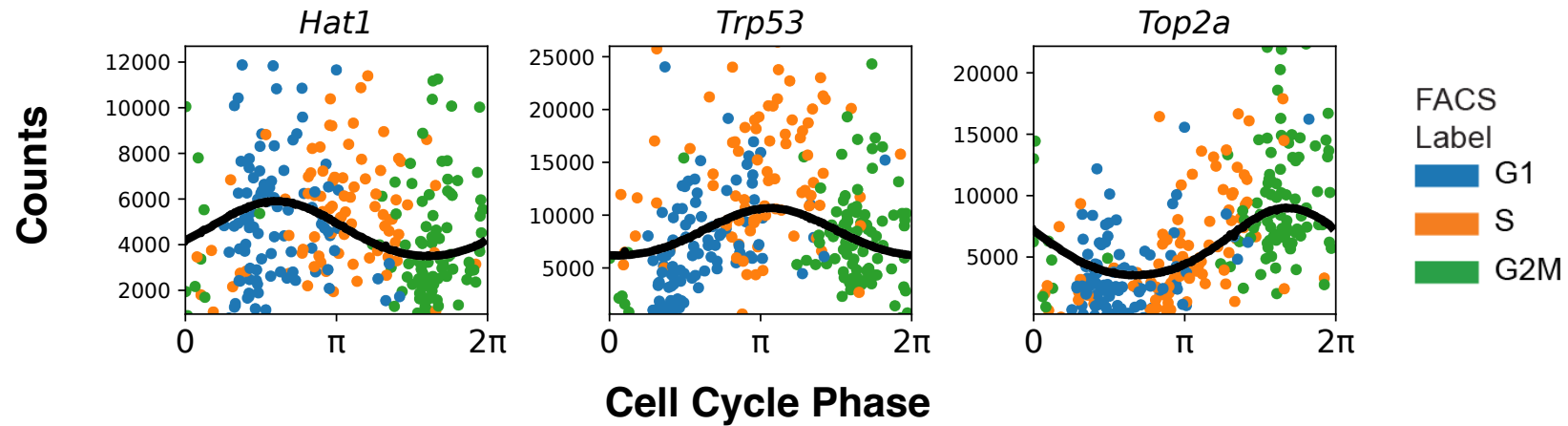
Buettner et al 2015

Phase assignments can be validated with FACS-sorted mESCs

Buettner et al 2015
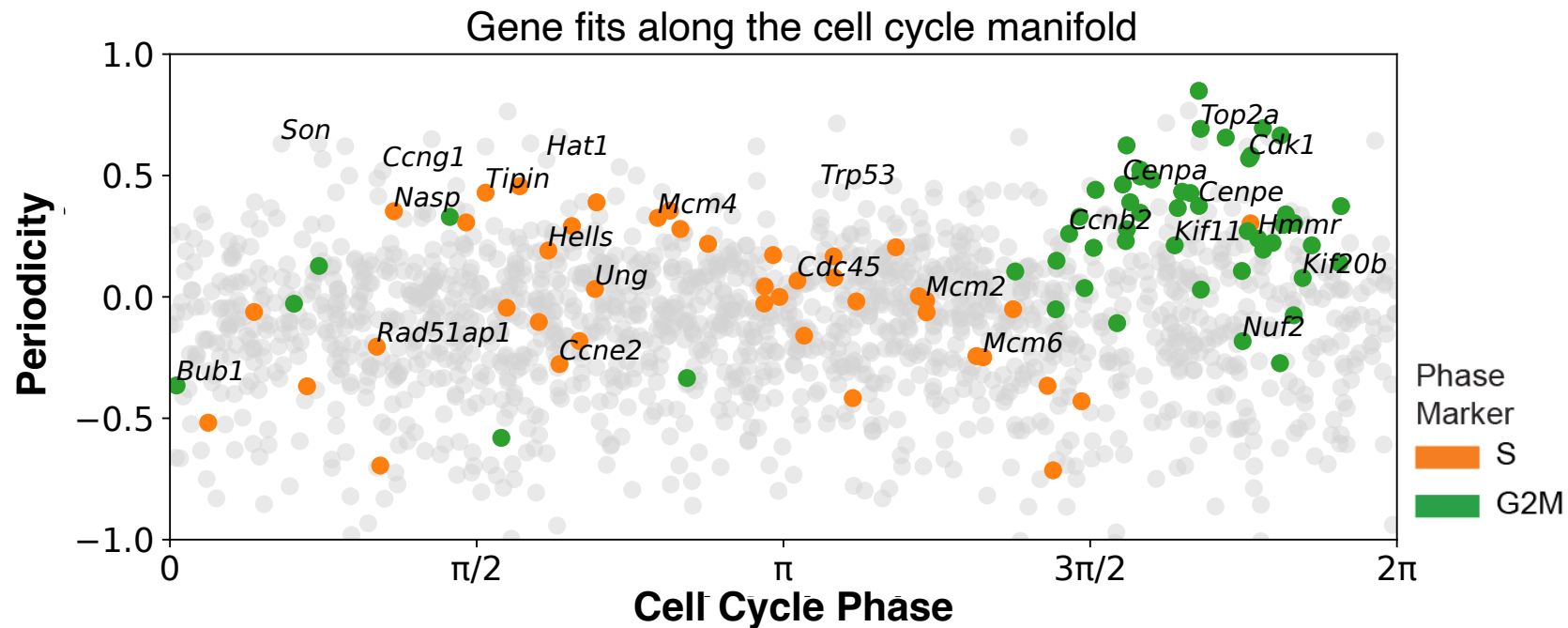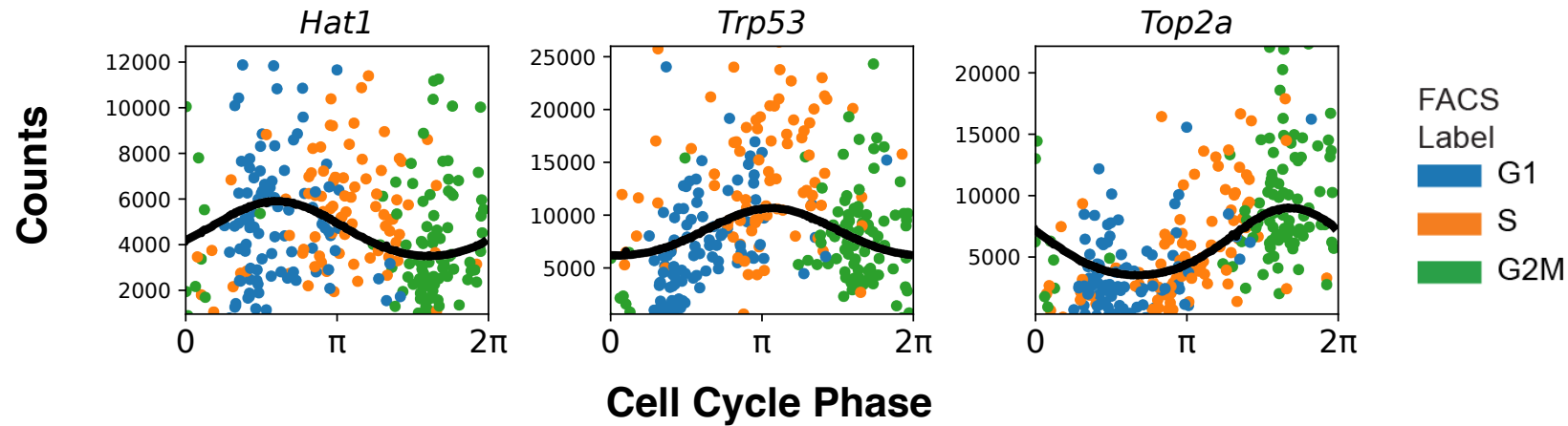
VeloCycle manifold-learning estimates accurate and robust phases
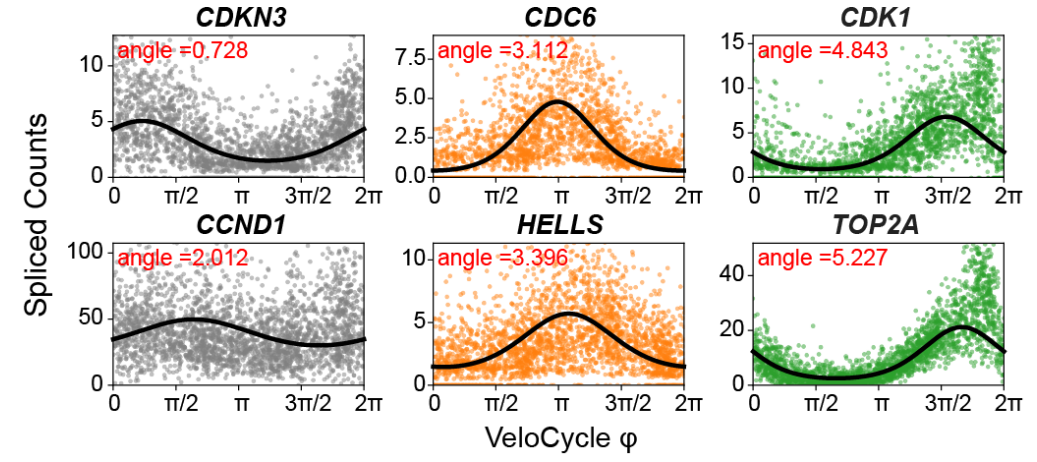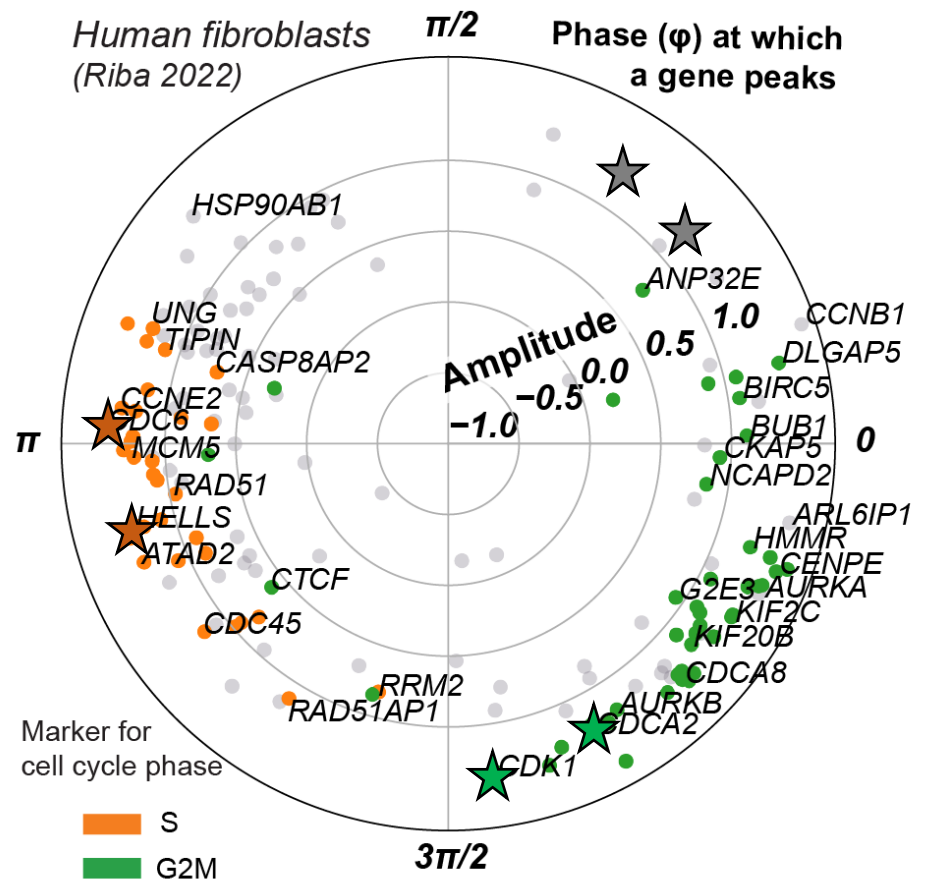
VeloCycle manifold-learning estimates accurate and robust phases
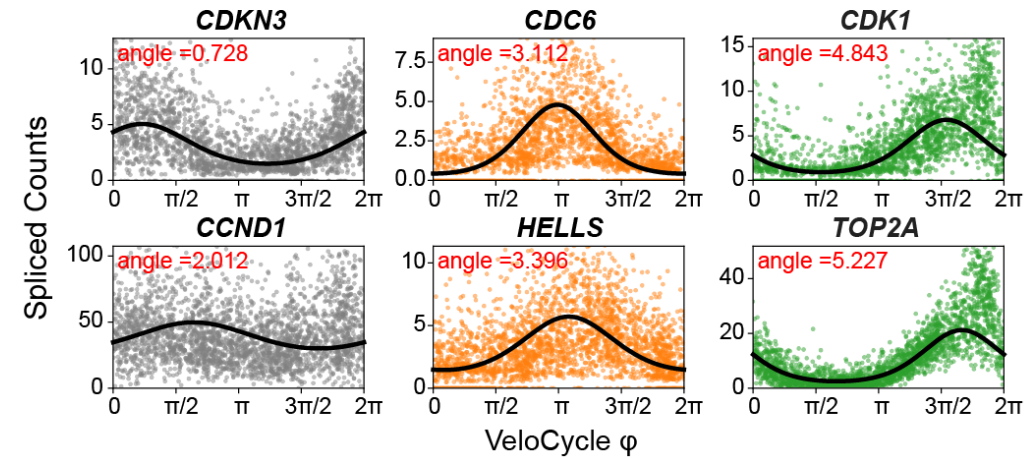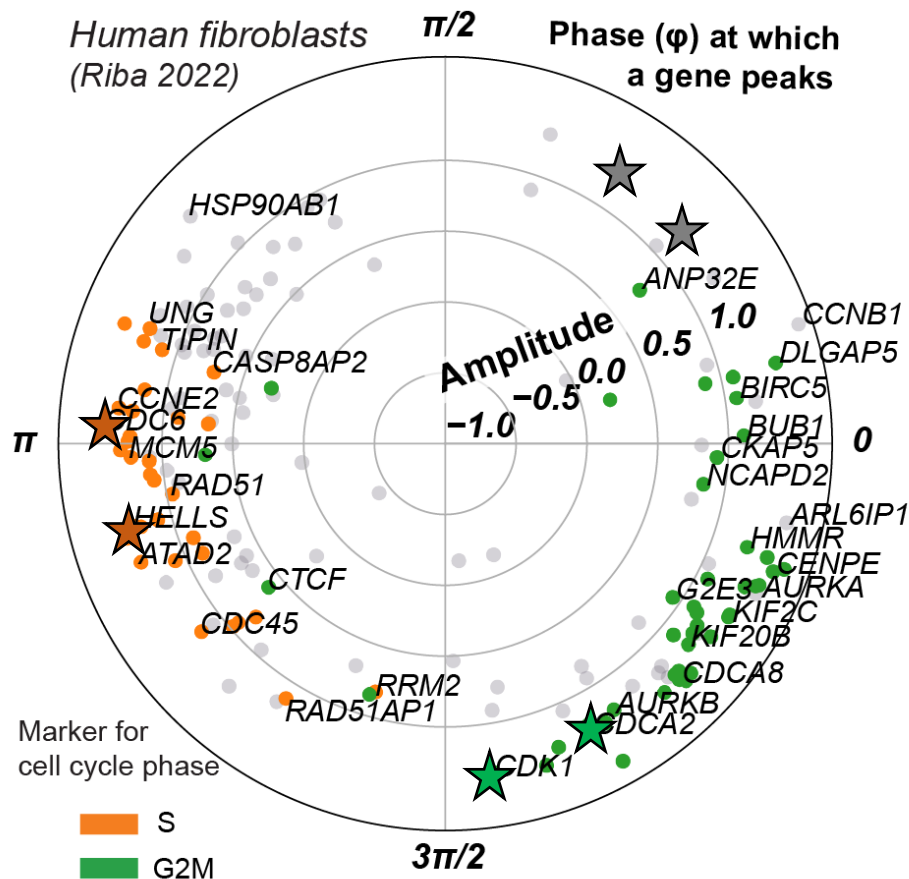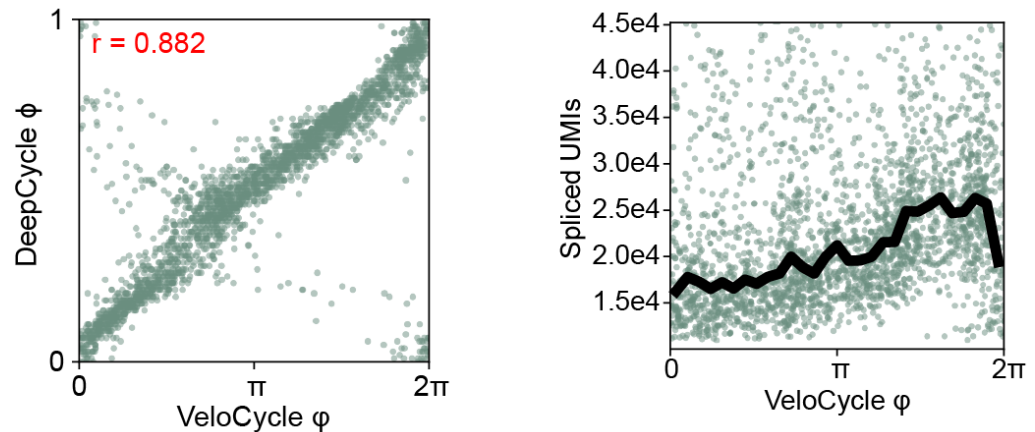
Can we experimentally validate our cell cycle speed estimates?

Time-lapse microscopy validates computationally inferred cell cycle speeds
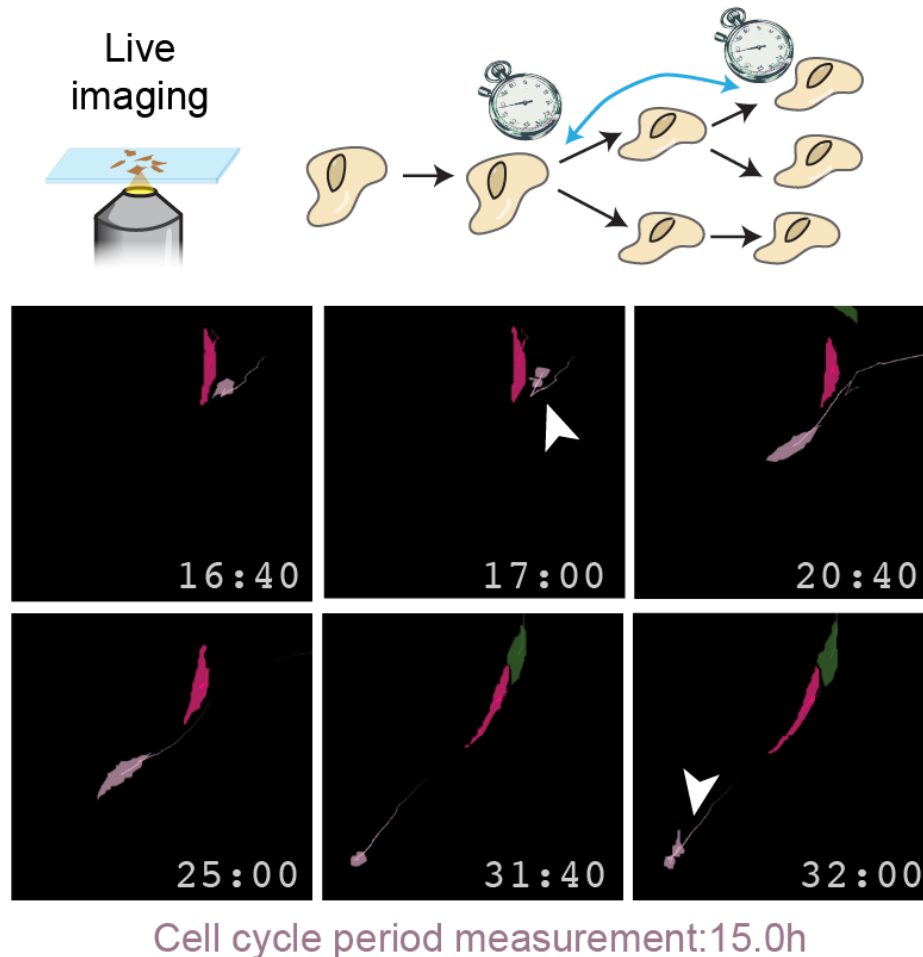
Cell cycle period measurement:15.0h

Dataset from: Capolupo et al. 2022

Time-lapse microscopy validates computationally inferred cell cycle speeds

Cell cycle period measurement:15.0h

Live-tracked cell cycle periods

Dataset from: Capolupo et al. 2022

Time-lapse microscopy validates computationally inferred cell cycle speeds

Live imaging

Live-tracked cell cycle periods

n=282

15.8h

Cell Cycle Length (hours)

16:40 17:00 20:40

25:00 31:40 32:00

Cell cycle period measurement:15.0h

VeloCycle speed estimate

16.5±2.1 h

VeloCycle Phase (φ)

Dataset from: Capolupo et al. 2022

29

Velocity estimates vary by phase and correspond to the unspliced-spliced expression delay

Velocity of Cells by Categorical Phase
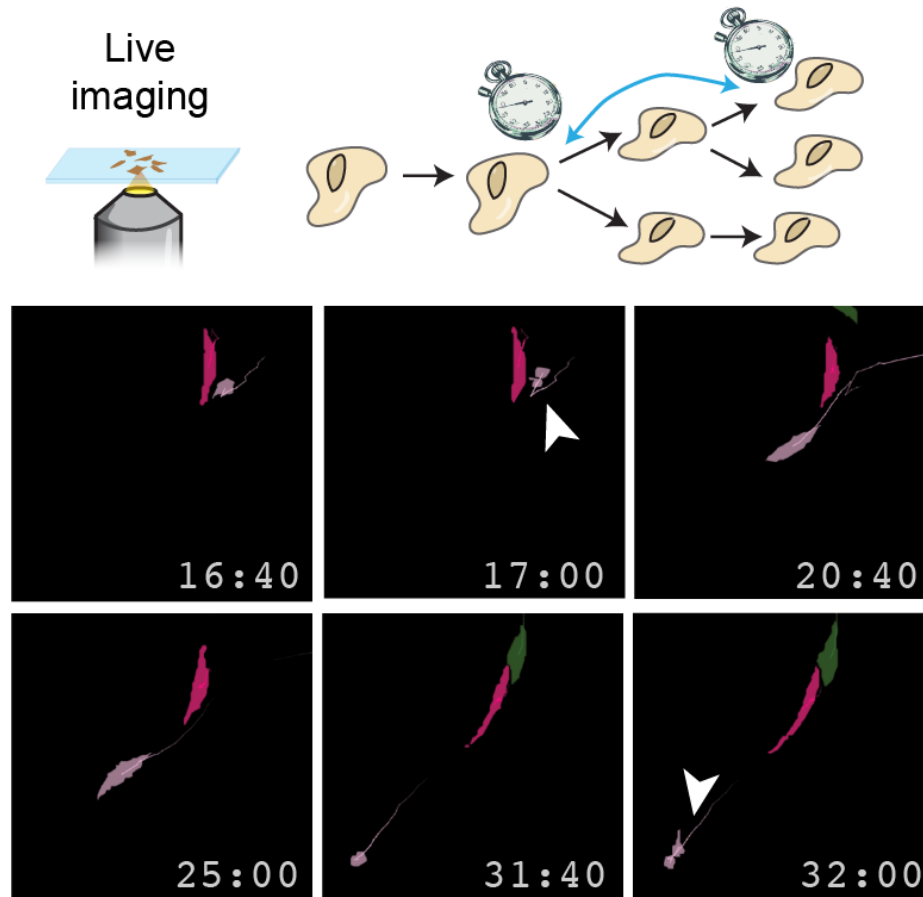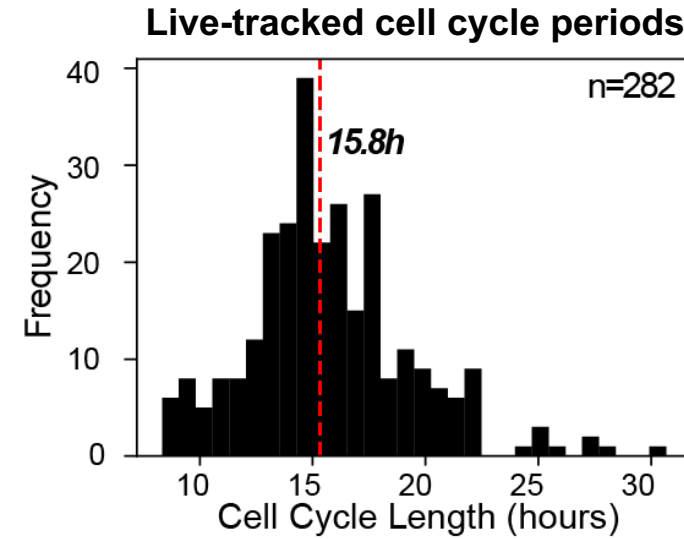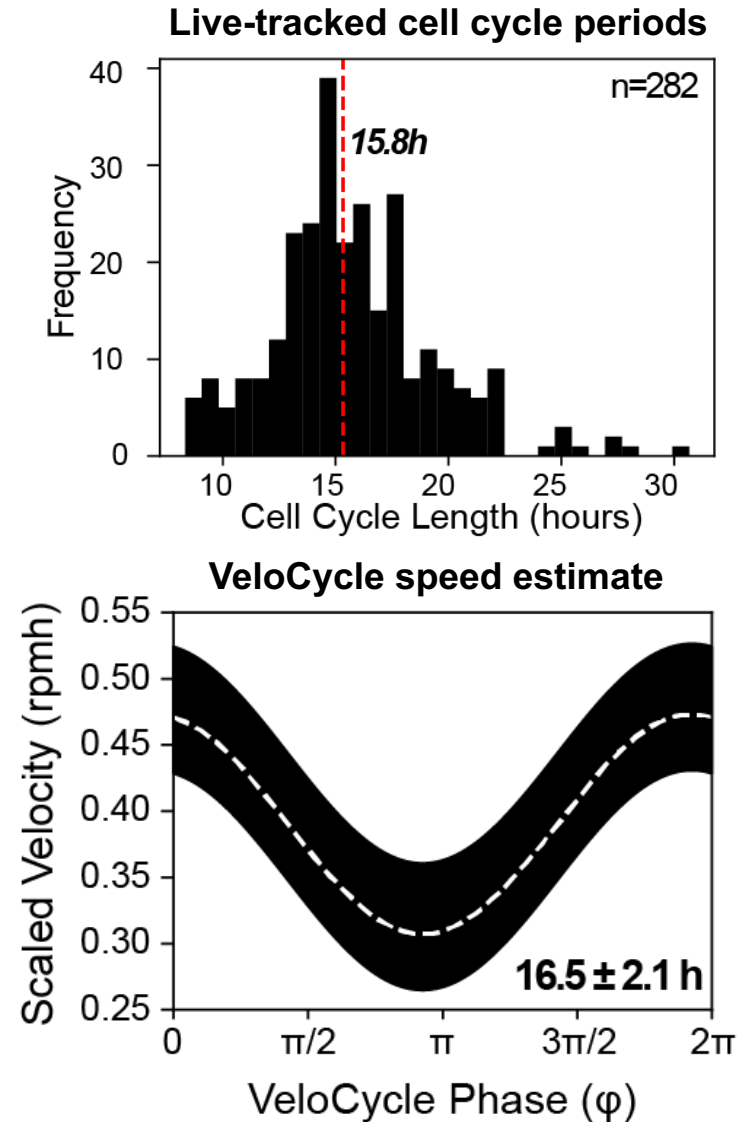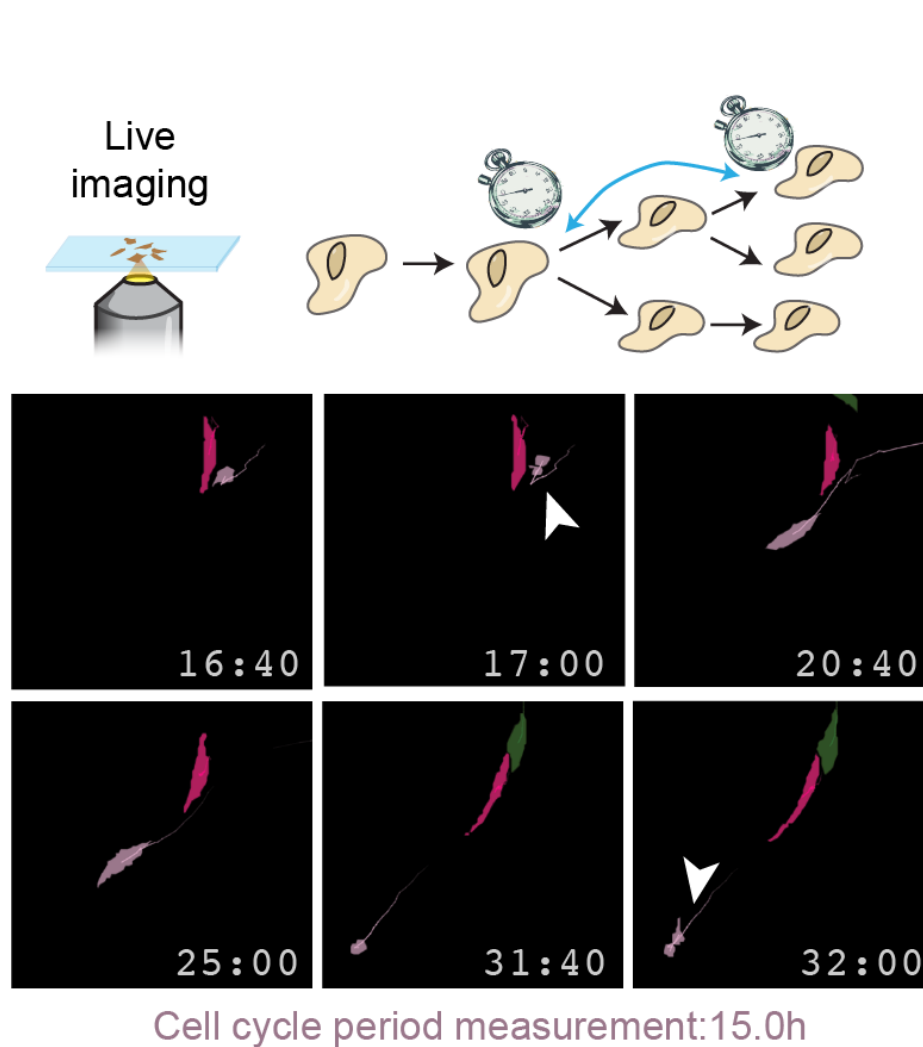
Dataset from: Capolupo et al. 2022

Velocity estimates vary by phase and correspond to the unspliced-spliced expression delay

Dataset from: Capolupo et al. 2022
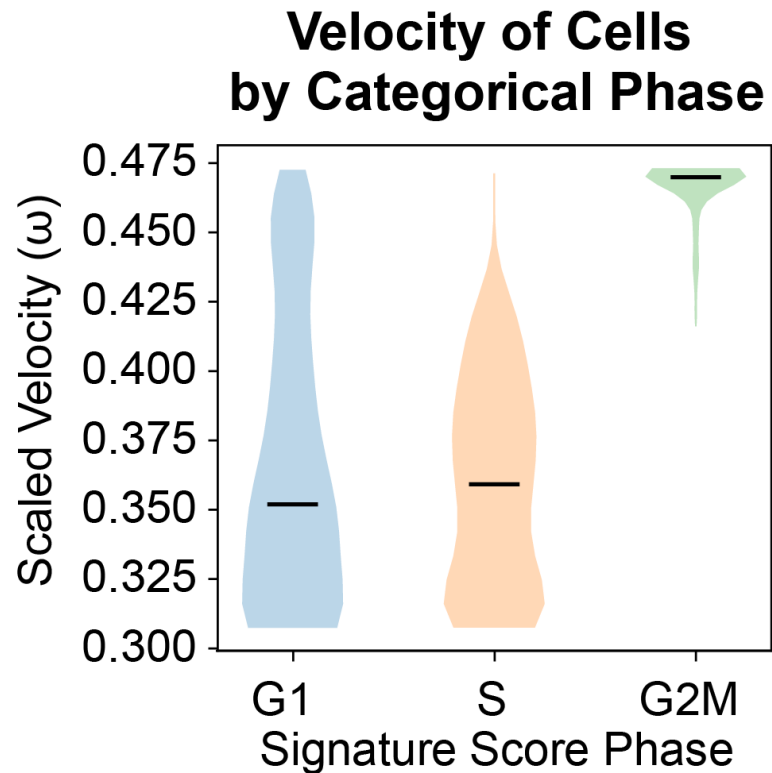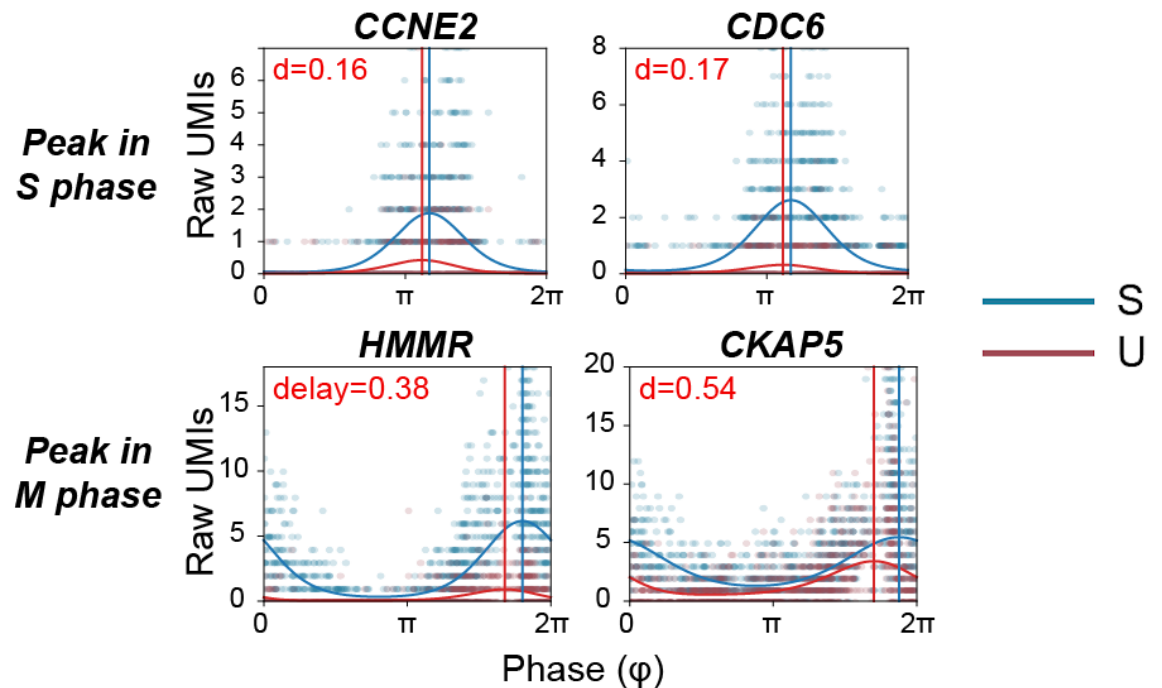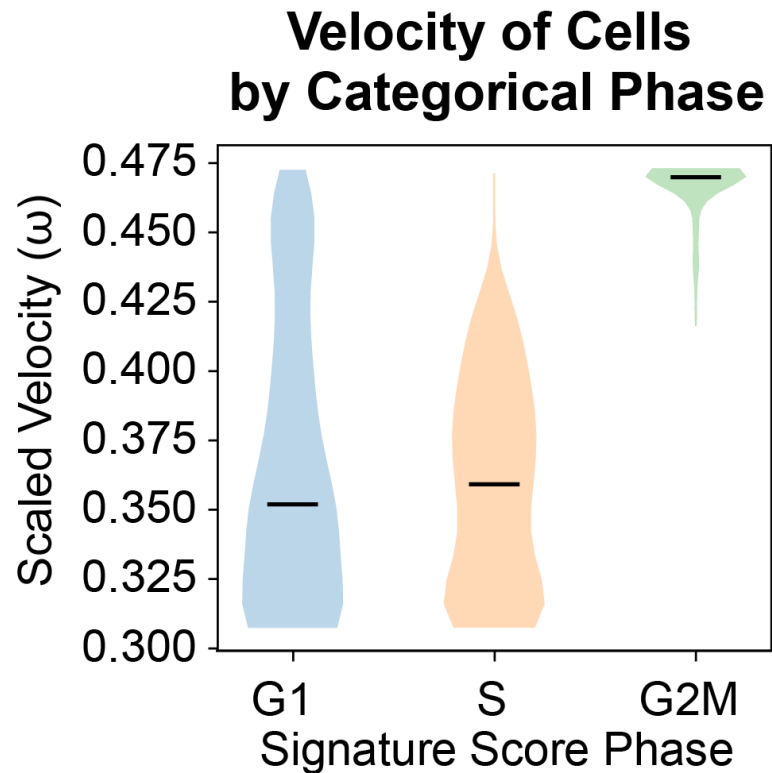
Does the rate of the cell cycle vary across different tissues and environmental contexts?

Credibility testing of cell cycle velocity before and after cell treatment with erlotinib

PC9 cancer cell line

D3 Treated (Erlotinib)

D0 Untreated

UMAP

Steady-state velocity analysis

D3 Treated (Erlotinib)

D0 Untreated

Dataset from: Aissa et al. 2021

33

Credibility testing of cell cycle velocity before and after cell treatment with erlotinib

PC9 cancer cell line

Steady-state velocity analysis

D3 Treated (Erlotinib)

D0 Untreated

D3 Treated (Erlotinib)

D0 Untreated

Does erlotinib treatment affect the cell cycle speed?

UMAP

Dataset from: Aissa et al. 2021

34

Credibility testing of cell cycle velocity before and after cell treatment with erlotinib

PC9 cancer cell line

D3 Treated (Erlotinib)

D0 Untreated

UMAP

Steady-state velocity analysis

D3 Treated (Erlotinib)

D0 Untreated

Does erlotinib treatment affect the cell cycle speed?

Is there a measurable non-zero velocity?

D0 Untreated

18.3 ± 0.5 h

G1
S
G2M

Scaled Velocity (rpmh)

VeloCycle Phase (φ)

Dataset from: Aissa et al. 2021

35

Credibility testing of cell cycle velocity before and after cell treatment with erlotinib
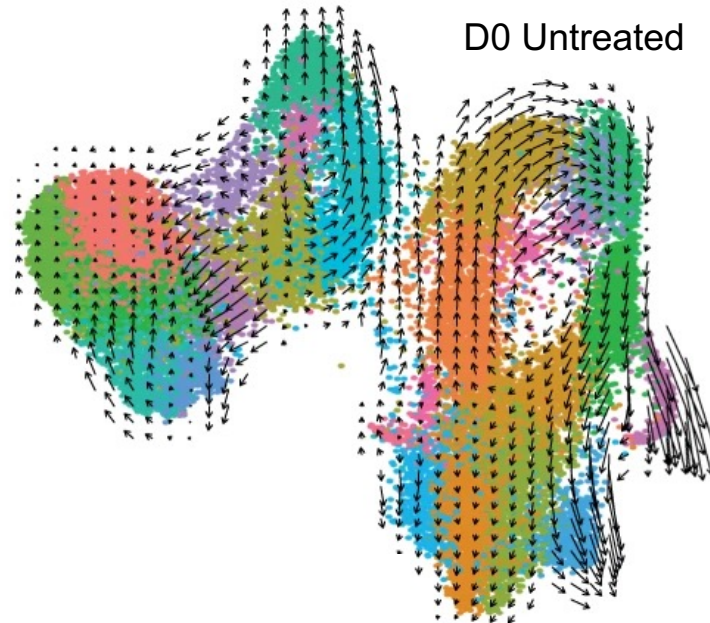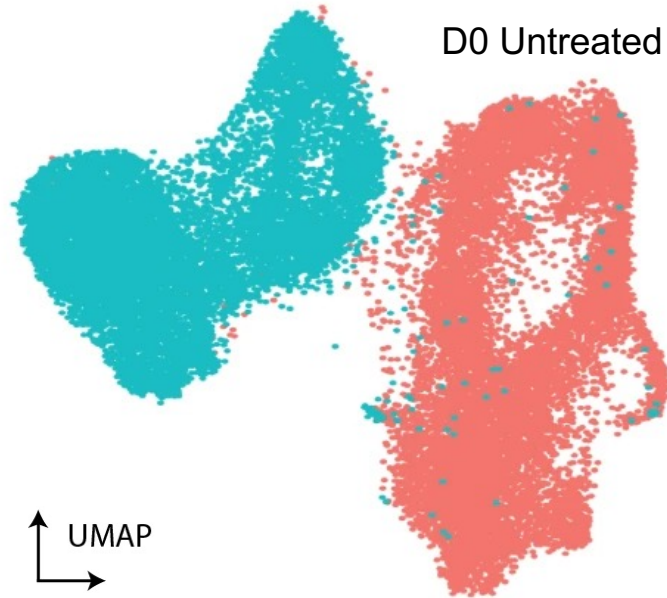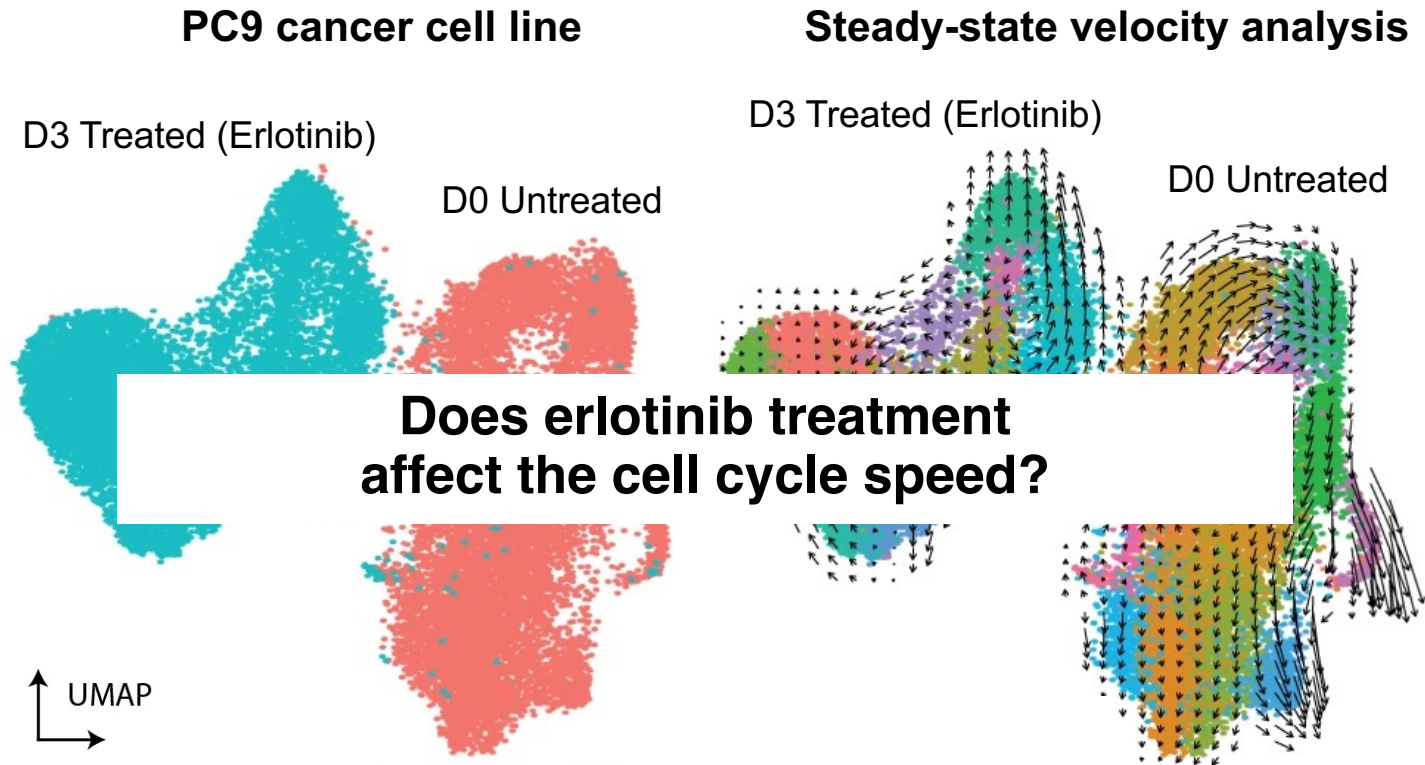

PC9 cancer cell line

D3 Treated (Erlotinib)

D0 Untreated

UMAP

Steady-state velocity analysis

D3 Treated (Erlotinib)

D0 Untreated

Does erlotinib treatment affect the cell cycle speed?

Is there a credible velocity difference?

18.3±0.5 h

21.0±0.5 h

G1
S
G2M

Scaled Velocity (rpmh)

VeloCycle Phase (φ)

Cell cycle speed in radial glial progenitors varies along a spatio-temporal axis in mouse development

La Manno et al 2021
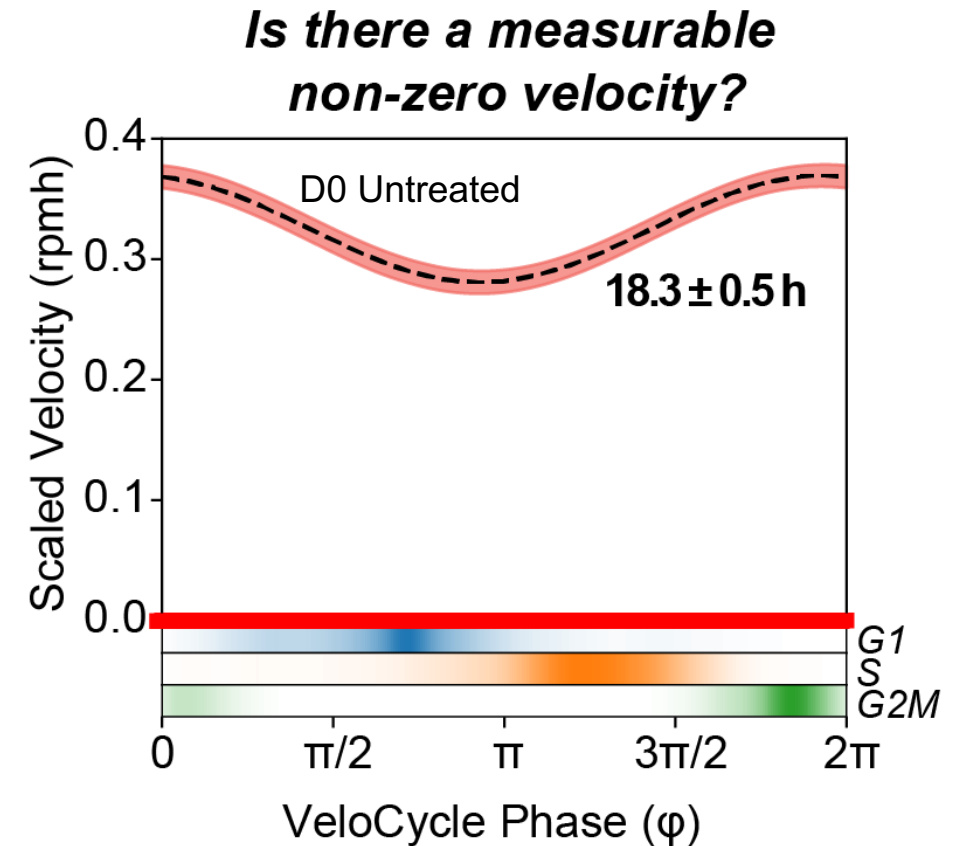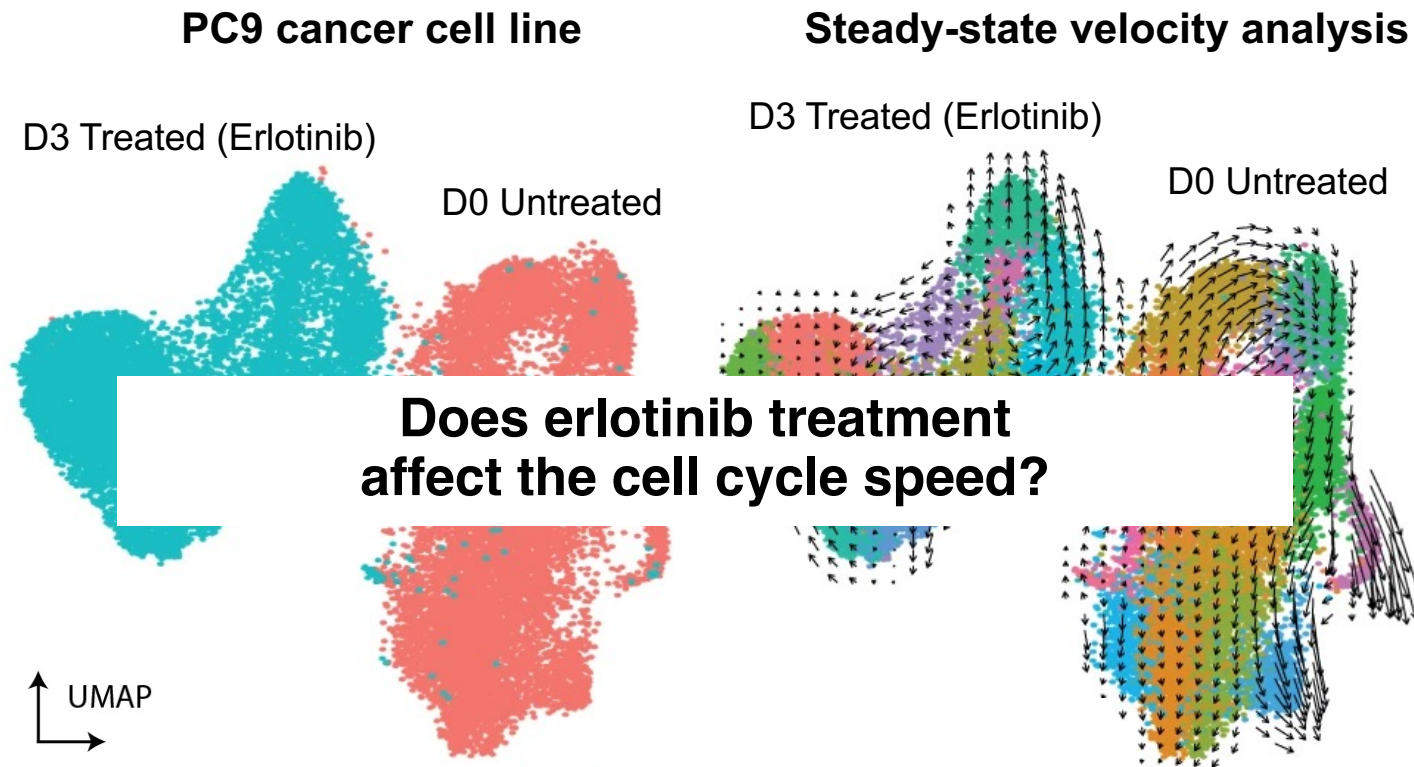
Cell cycle speed in radial glial progenitors varies along a spatio-temporal axis in mouse development
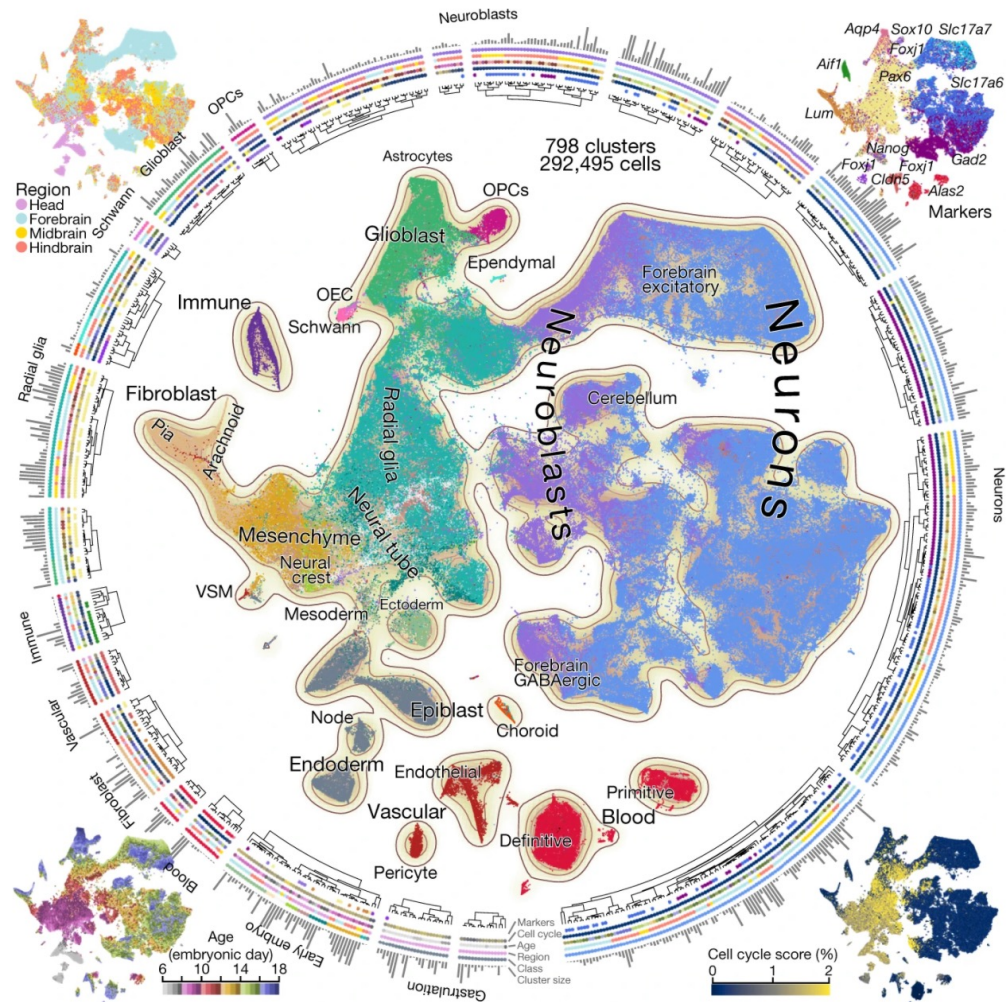
Does the cell cycle speed vary in brain progenitors from different spatial regions?

La Manno et al 2021

Cell cycle speed in radial glial progenitors varies along a spatio-temporal axis in mouse development

La Manno et al 2021

Cell cycle speed in radial glial progenitors varies along a spatio-temporal axis in mouse development

Spatial projection using HybISS and BoneFight
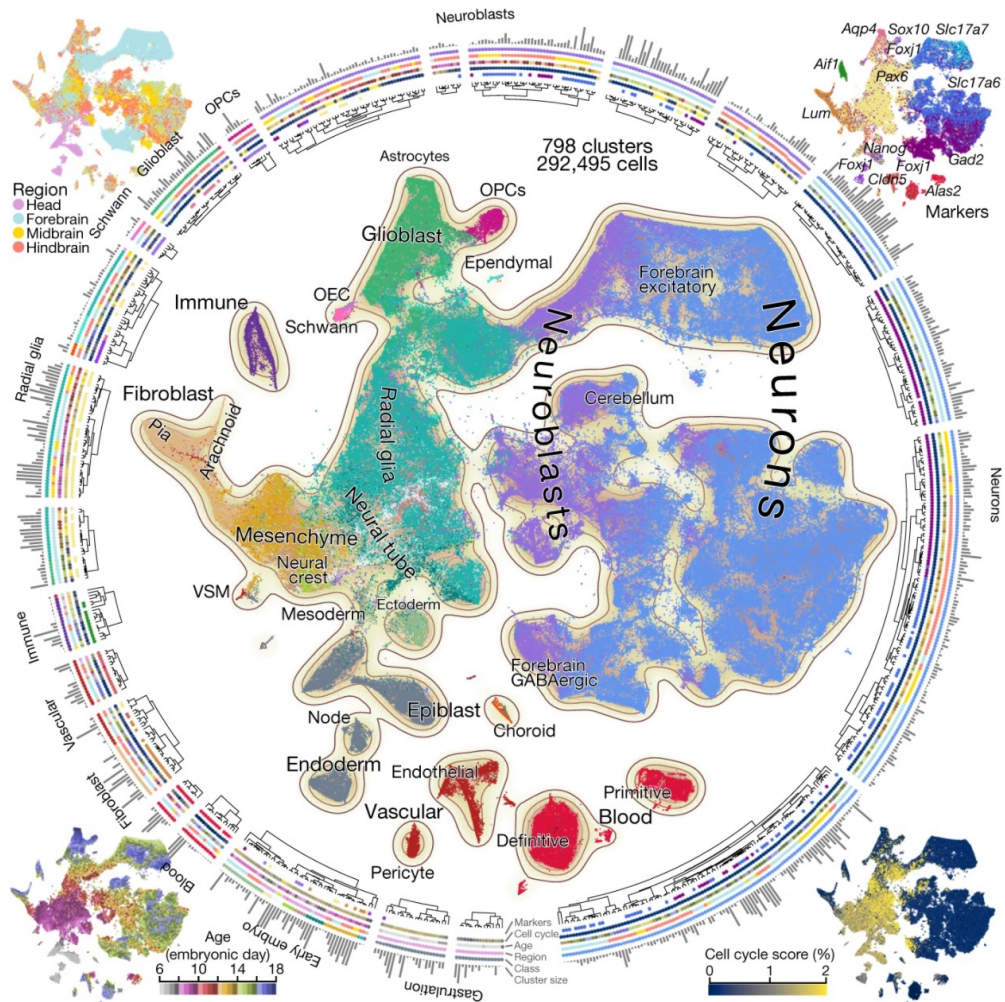
La Manno et al 2021

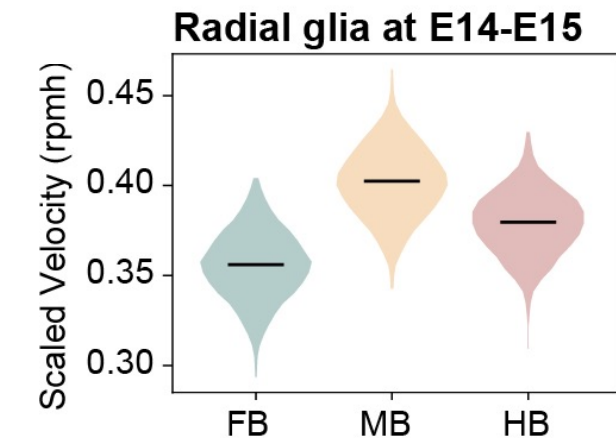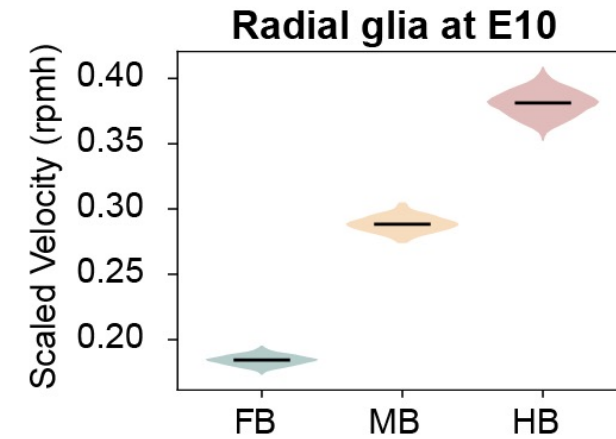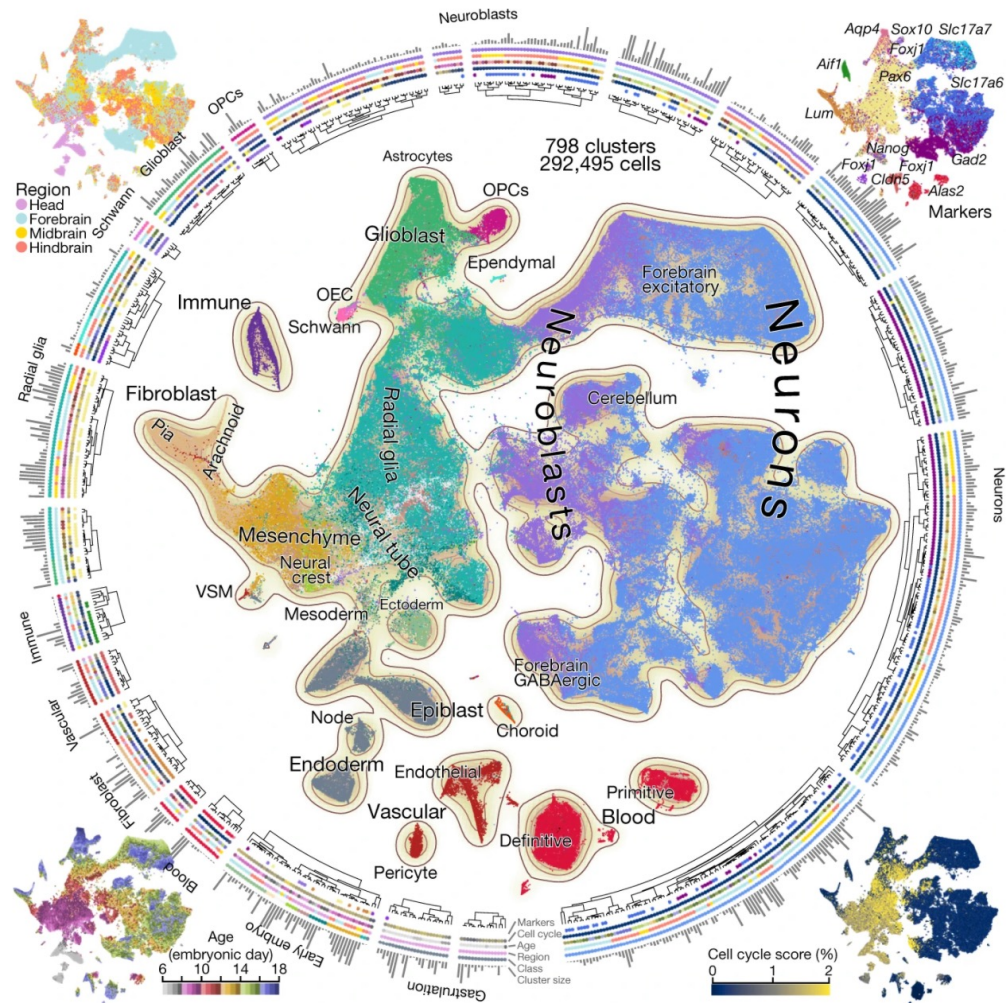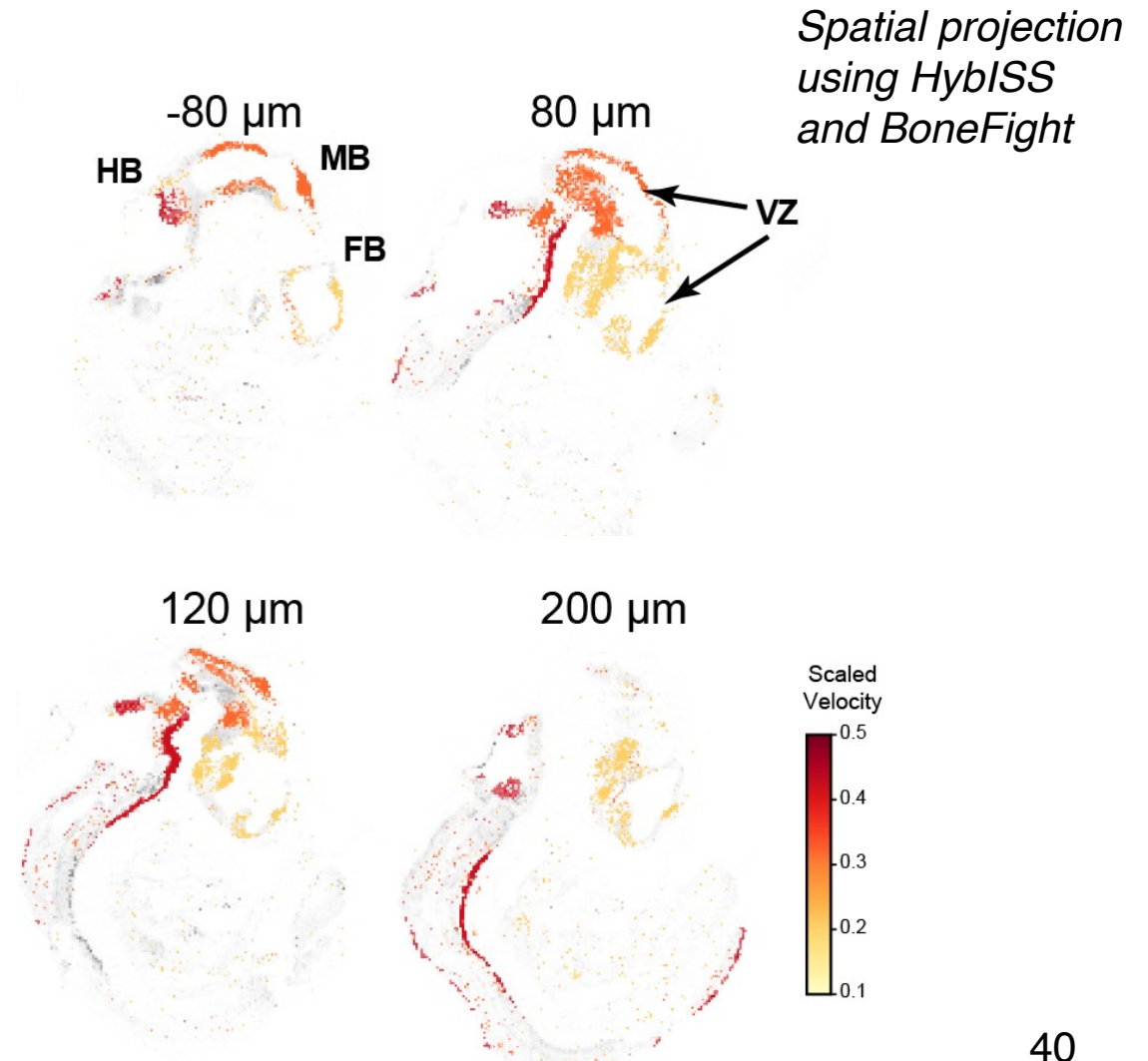# Genome-wide Perturb-seq constructs a comprehensive genotype-phenotype map
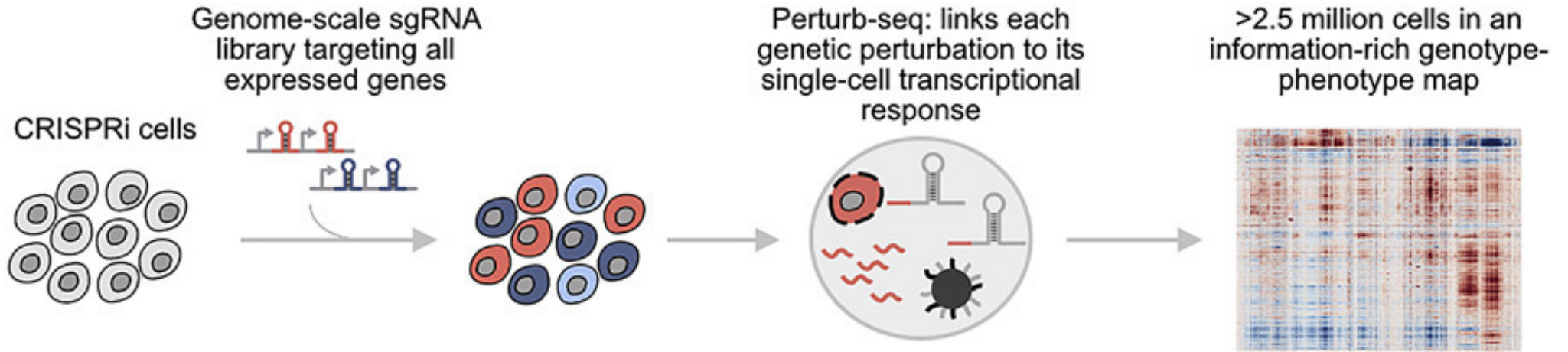
Replogle, Saunders, et al. 2022

Genome-wide Perturb-seq constructs a comprehensive genotype-phenotype map

Genome-scale sgRNA library targeting all expressed genes

CRISPRi cells

Perturb-seq: links each genetic perturbation to its single-cell transcriptional response

>2.5 million cells in an information-rich genotype-phenotype map

Replogle, Saunders, et al. 2022
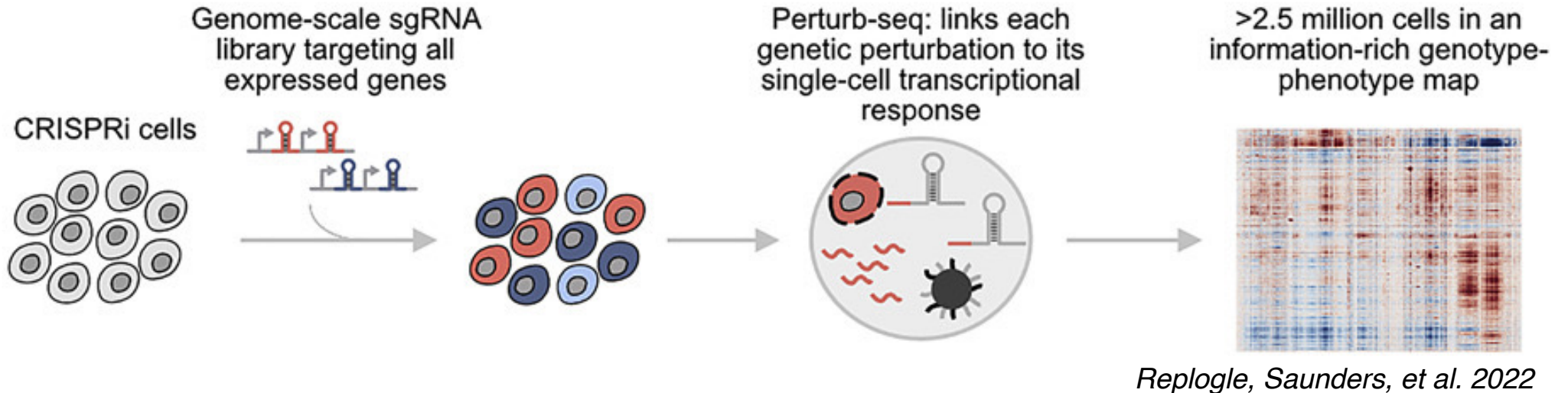
How do single-gene knockdowns affect the cell cycle speed?

Challenge: there are very few cells for each perturbation condition.

Cells containing a targeted knock-down of a cell cycle gene have a slower cell cycle velocity and more G1 cells

- Genome-wide Perturb-seq dataset in RPE1 cell line

Dataset from: Replogle, Saunders, et al. 2022

Cells containing a targeted knock-down of a cell cycle gene have a slower cell cycle velocity and more G1 cells

- Genome-wide Perturb-seq dataset in RPE1 cell line

Dataset from: Replogle, Saunders, et al. 2022
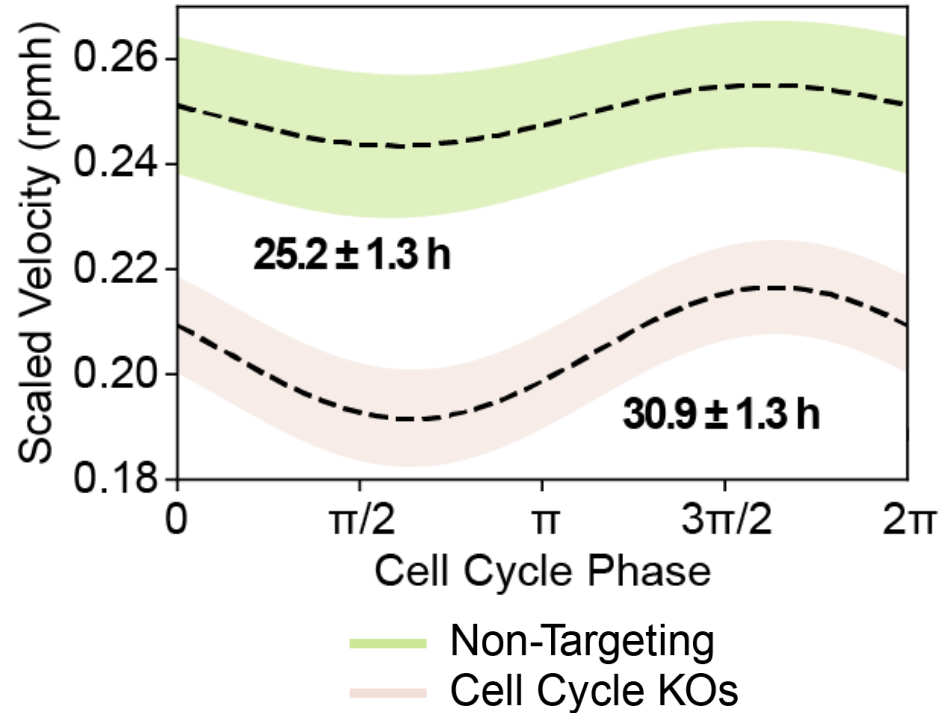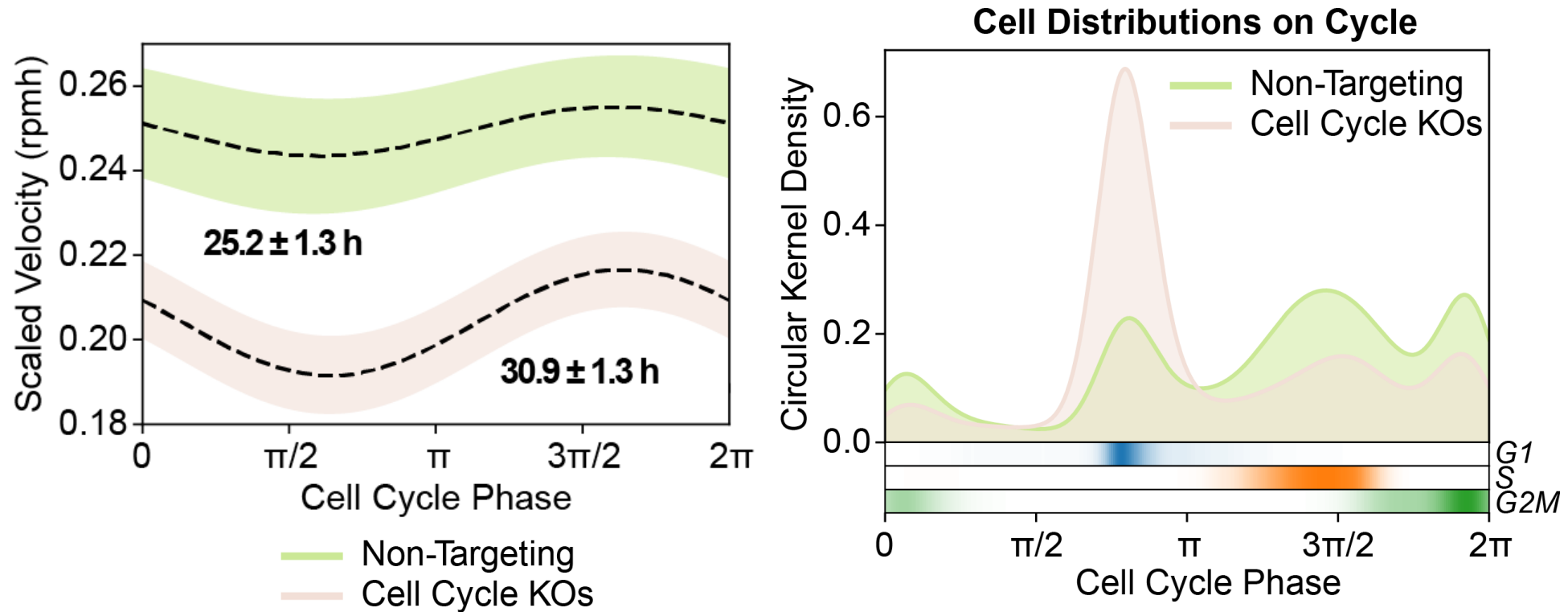
Transfer learning of gene harmonics enables study of the effects of gene knockdowns on velocity
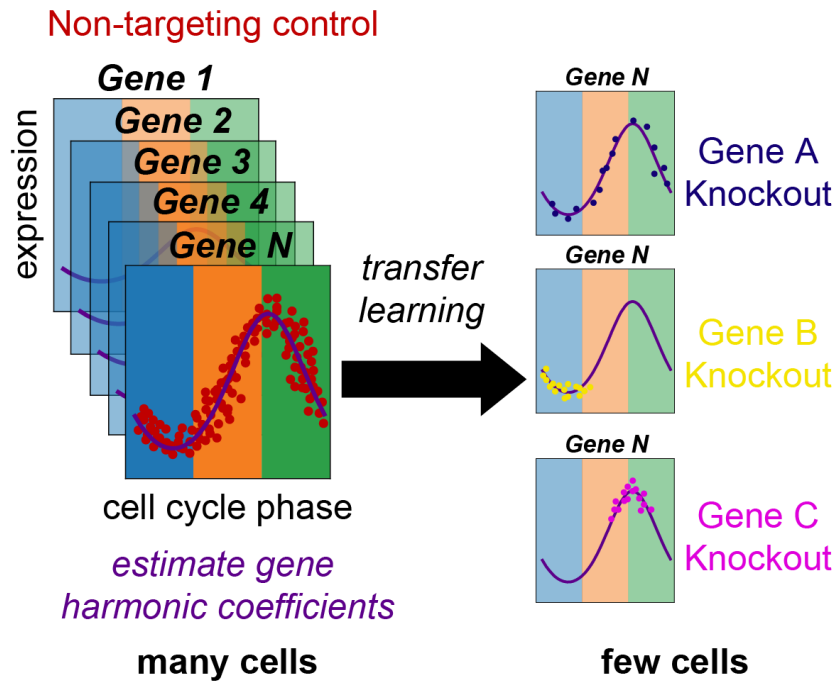
Transfer gene harmonics inferred on non-targeted cells to obtain phases of small batches of knockdown cells

Dataset from: Replogle, Saunders, et al. 2022

Perturb-seq conditions with altered speed are linked to DNA replication, translation initiation, and mRNA splicing
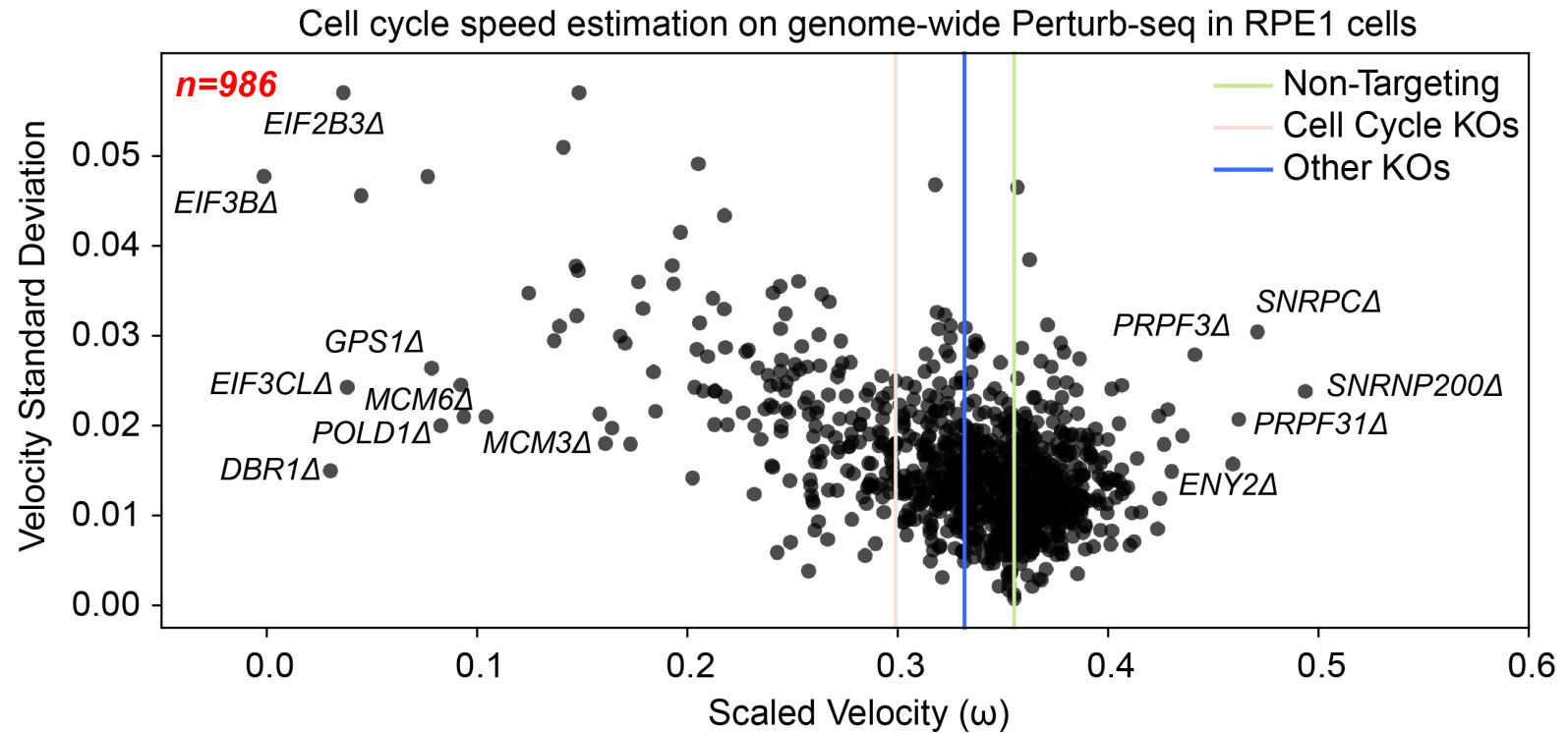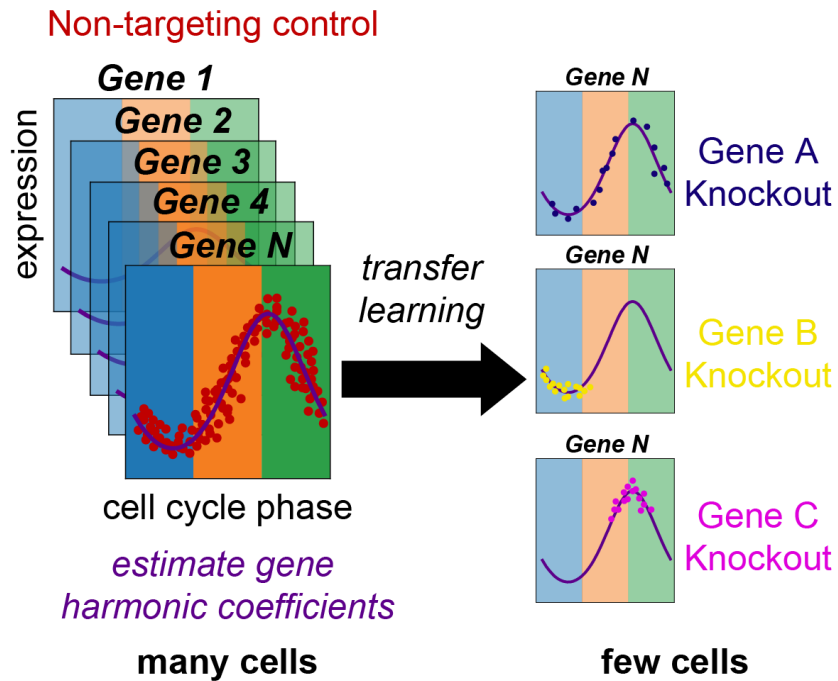
167,119 cells from 986 knockdown conditions

Takeaways

- RNA velocity vector fields are inconsistent with the gene expression manifold.

Takeaways

- RNA velocity vector fields are inconsistent with the gene expression manifold.

- VeloCycle is a generative Bayesian model that couples velocity field and manifold estimation into a unified framework for periodic manifolds.

Takeaways

- RNA velocity vector fields are inconsistent with the gene expression manifold.

- VeloCycle is a generative Bayesian model that couples velocity field and manifold estimation into a unified framework for periodic manifolds.

- VeloCycle can infer cell cycle periods on a real-time scale, validated by time-lapse microscopy imaging.

Takeaways

- RNA velocity vector fields are inconsistent with the gene expression manifold.

- VeloCycle is a generative Bayesian model that couples velocity field and manifold estimation into a unified framework for periodic manifolds.

- VeloCycle can infer cell cycle periods on a real-time scale, validated by time-lapse microscopy imaging.

- **VeloCycle can harness statistical testing and transfer learning in multiple *in vitro* and *in vivo* experimental contexts.**