

Trajectory analysis

Ken Lau, Assistant Professor of Cell and Developmental Biology

ken.s.lau@vanderbilt.edu

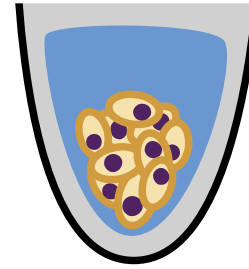
CQS Summer Academy (8/15/2018)

<http://www.mc.vanderbilt.edu/vumcdept/cellbio/laulab/index.html>

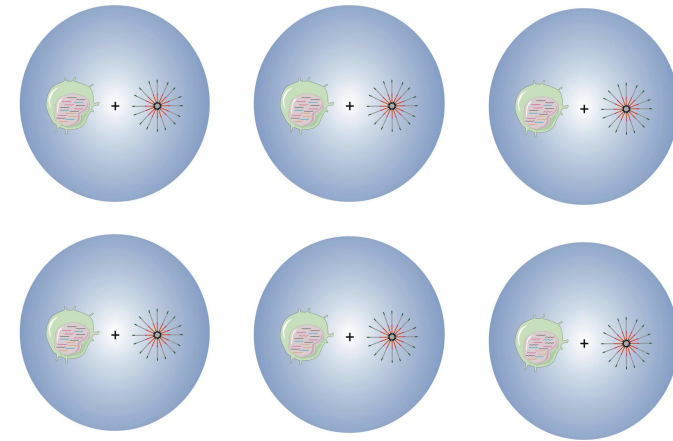
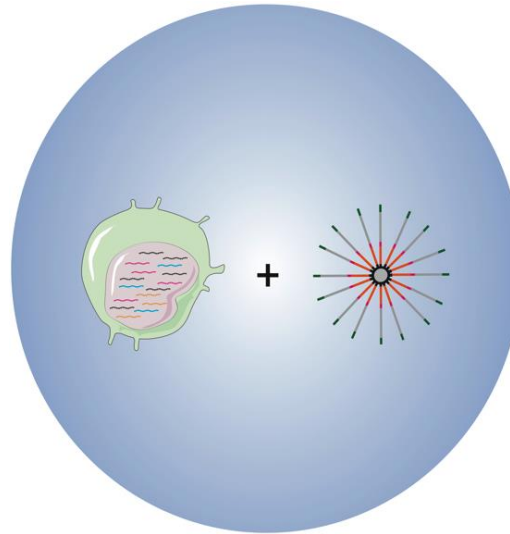
Twitter: @KenLauLab



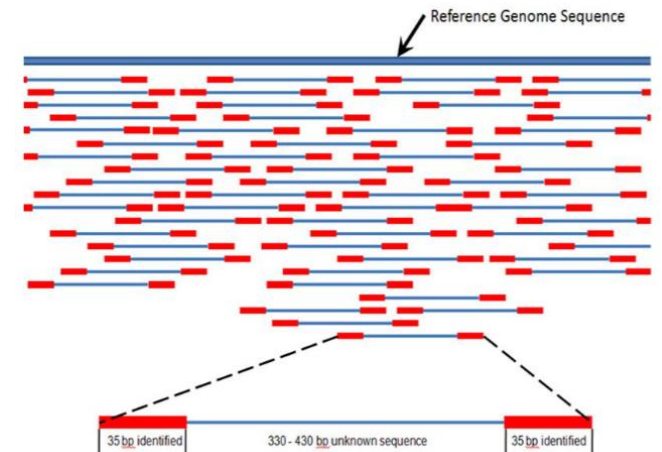
Single cell suspension***



**Single-cell encapsulation/
Library preparation**



**Sequencing and alignment
(Bioinformatics I)**



scRNA-seq processing steps

1. Cell identification from barcodes – doublet and low quality cell discrimination

2. Normalize and transform

3. Feature selection or data imputation

4. Batch correction

5. Dimension reduction

6. Clustering

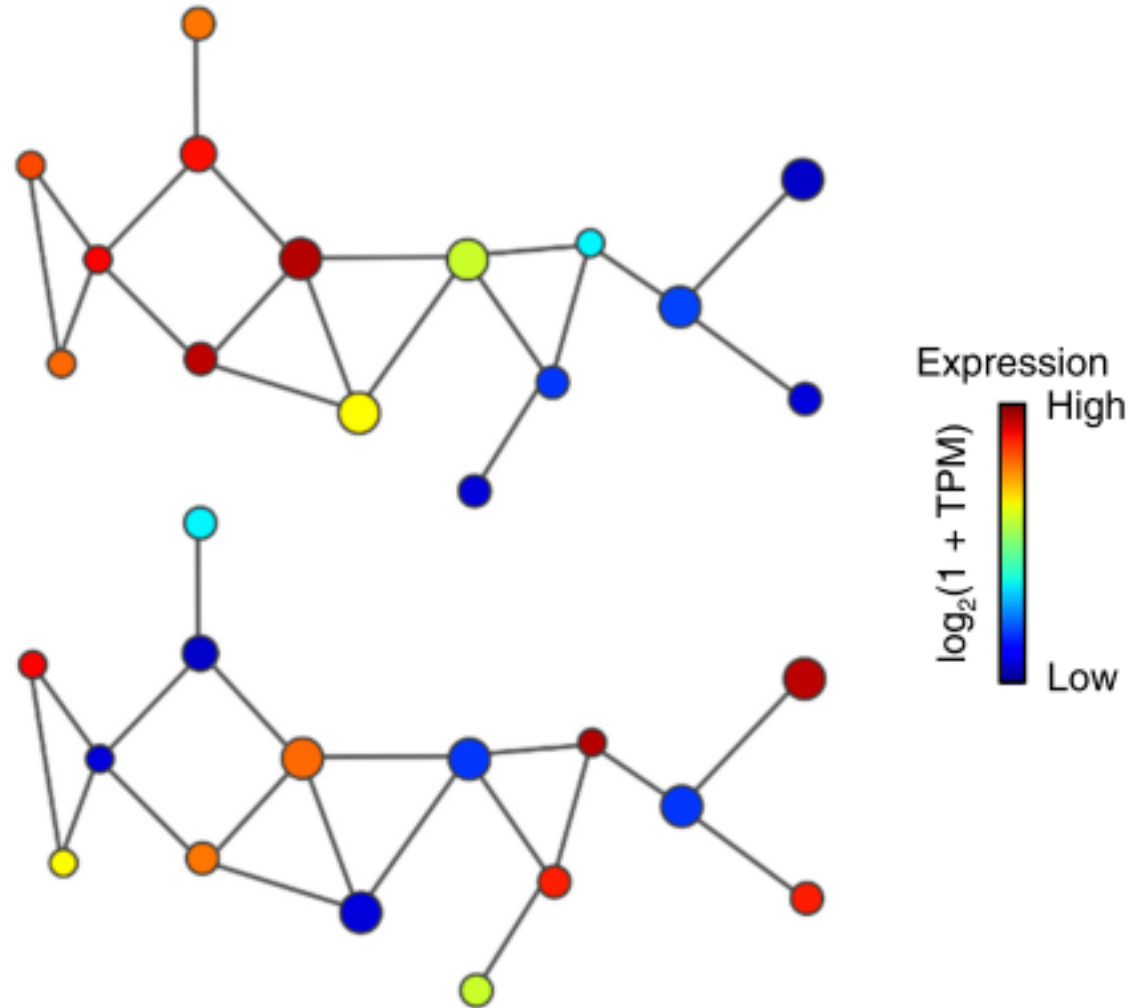
7. Trajectory analysis

8. Data alignment

9. Differential expressed genes analysis

knn

Neighborhood variance feature selection



Allows for selection of genes that change monotonically over all cells

Trajectory analysis

The pancreatic landscape is different between the adult and embryonic stages

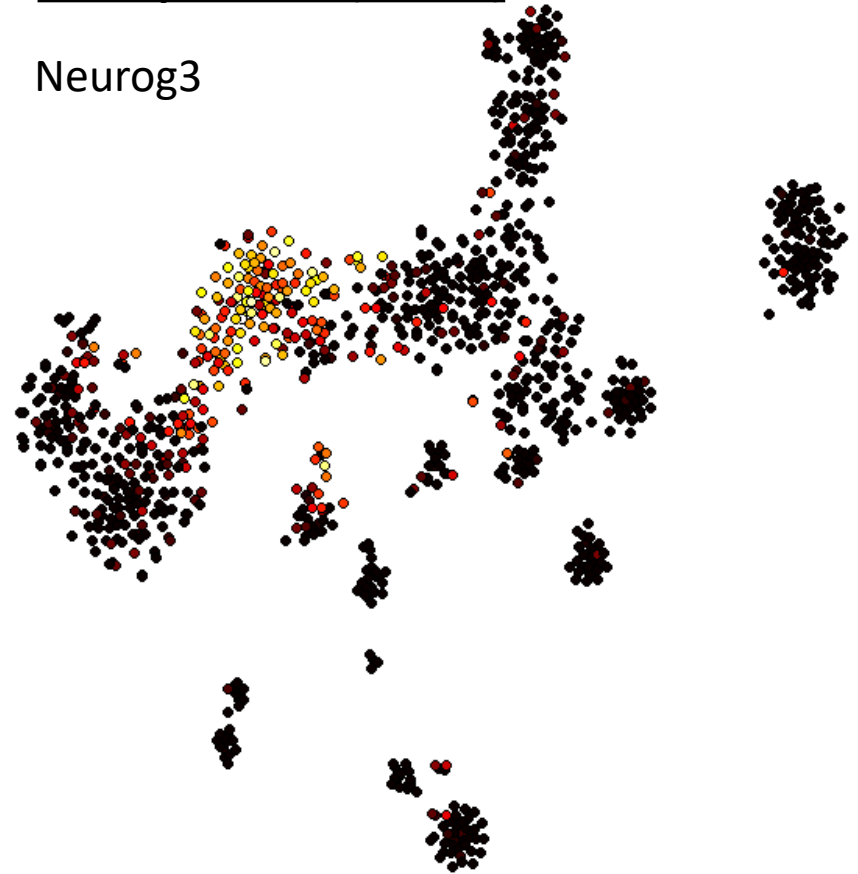
Adult pancreas (Magnuson lab)

Neurog3

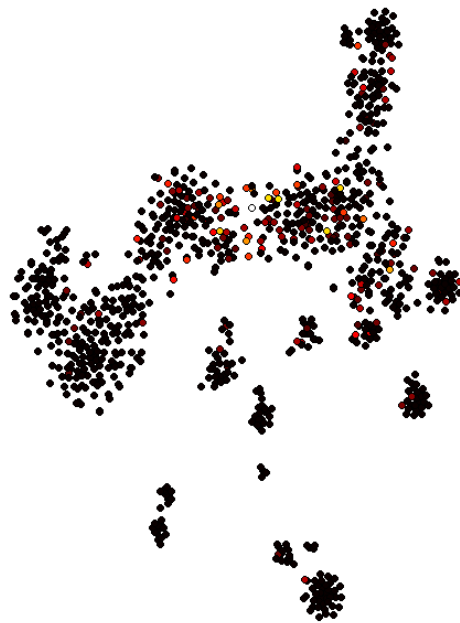


E14.5 pancreas (Gu lab)

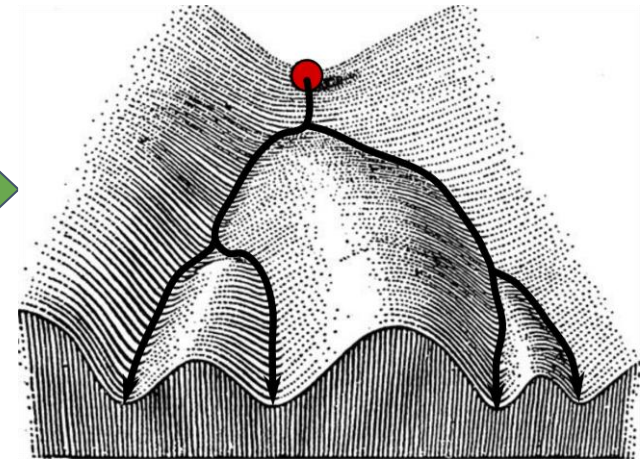
Neurog3



Single-cell data continuum represented as transitional trajectories



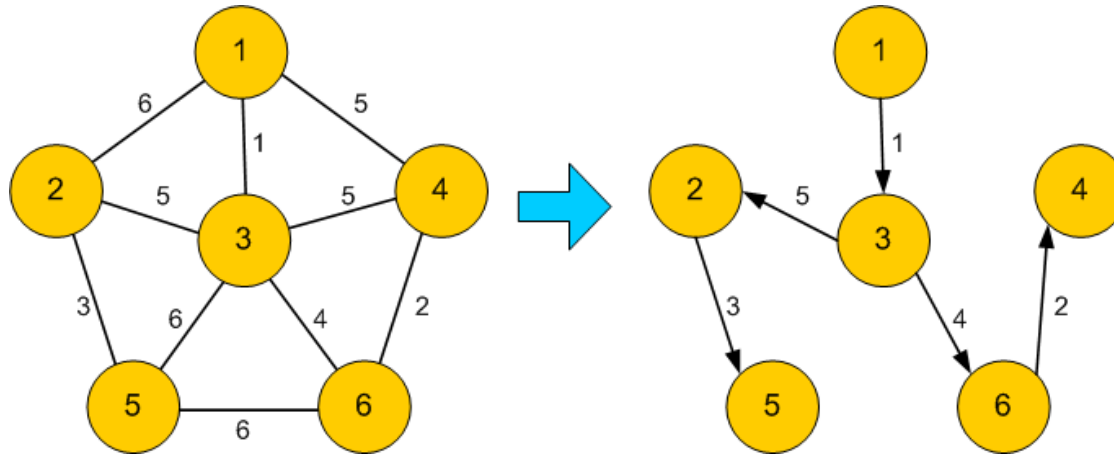
(Herring, Banerjee et al.,
2017, Cell Systems)



Waddington's
Landscape

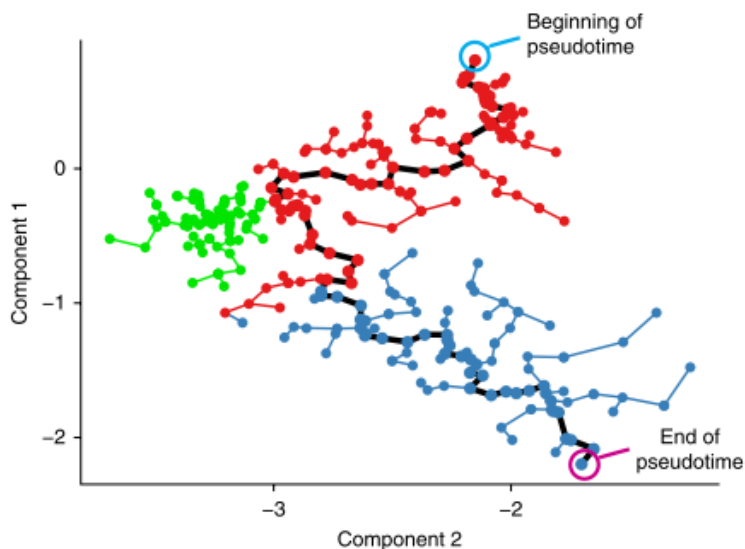
Minimum spanning tree algorithms

$$1 + 5 + 3 + 4 + 2 = 15 = \text{Tree Score}$$



doi:10.1038/nbt.2859

● Proliferating cell ● Differentiating myoblast ● Interstitial mesenchymal cell

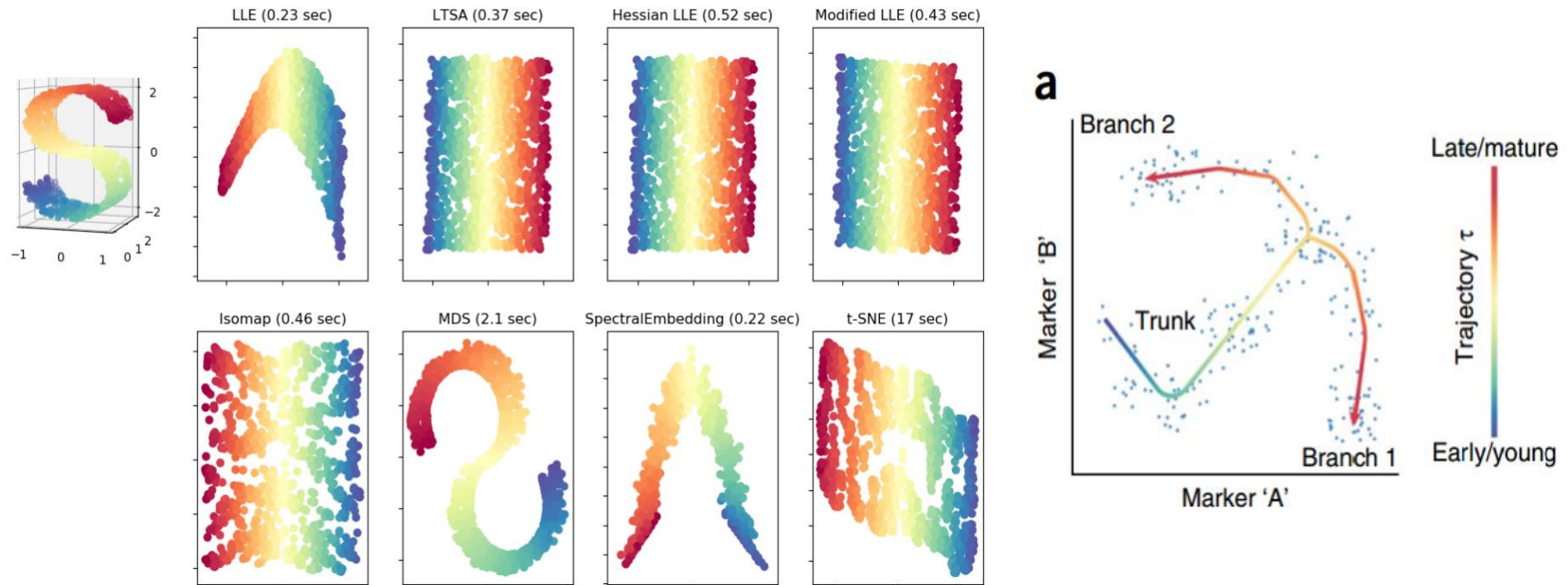


- Monocle (Trapnell *et al.*, 2014)
- TSCAN (Ji and Ji, 2016)
- Waterfall (Shin *et al.*, 2015)
- SPADE (Qiu *et al.*, 2011)
- Slingshot (Street *et al.* BioRxiv)
- ... and many others

Problem: Not robust. Different results at different runs with the same data.

Non-linear embedding algorithms

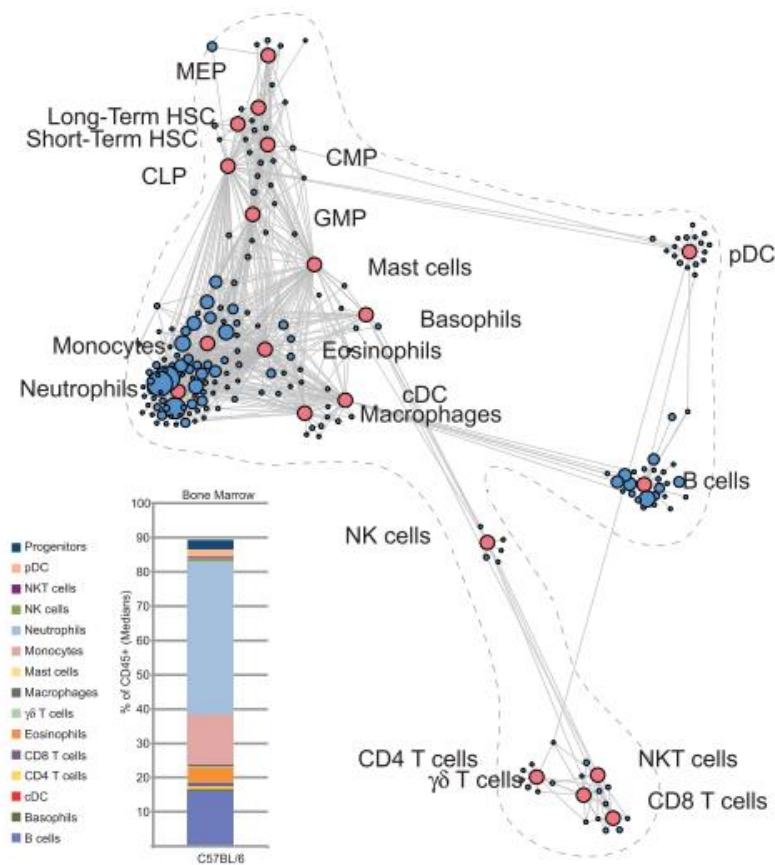
Non-linear embedding



- SLICER (Welch *et al.*, 2016) - LLE
- Diffusion Map Pseudotime (Haghverdi *et al.*, 2015) – Diffusion Map
- Wishbone (Setty *et al.*, 2016) – Diffusion Map plus network
- scTDA (Rizvi *et al.*, 2017) – MDS, Mapper

Problem: Stretching and compressing of data a tradeoff between increased resolution and noise. Thus, limited to simple structures.

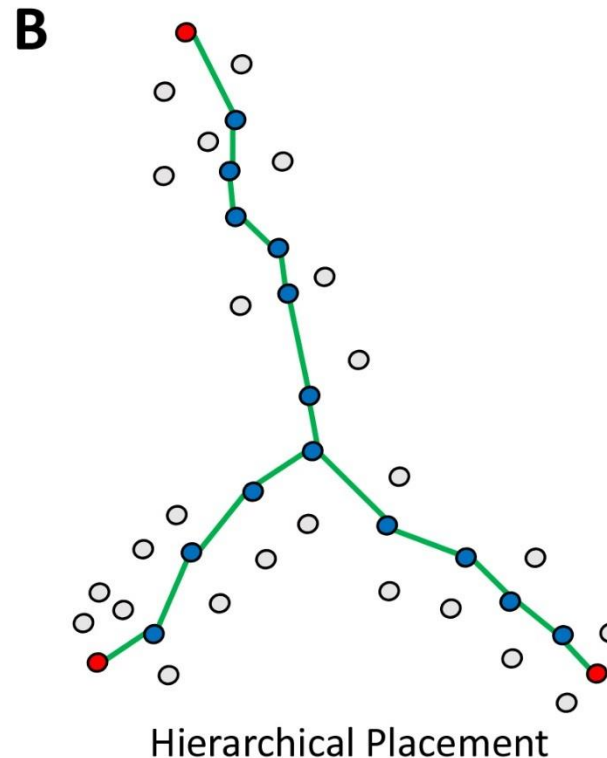
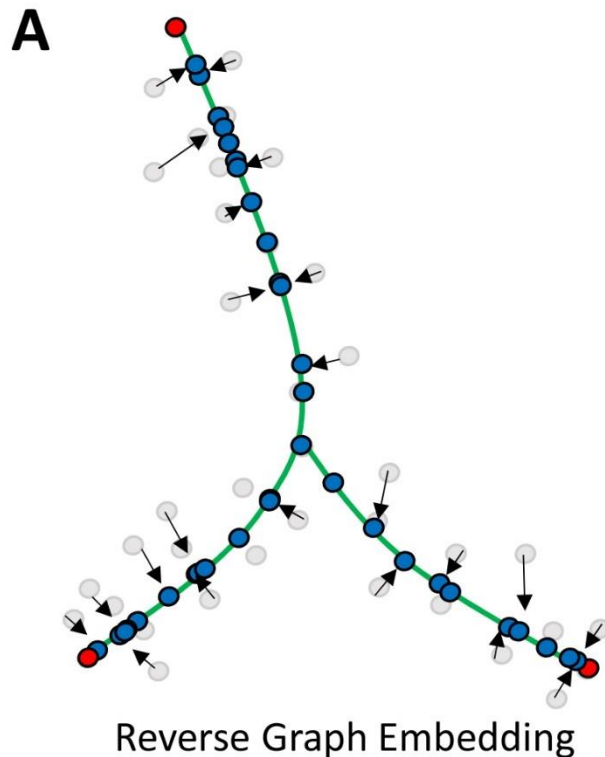
Force directed algorithms



- FLOW-Map (Eli Zunder)
- Scaffold (Garry Nolan)
- SPRING (Allon Klein)

Problem: Complex structures difficult to interpret; visualization only

Graph methods



- Monocle2 (Trapnell, 2017)
- p-Create (Lau, 2017)
- CellRouter (Collins and Daley, 2018)
- PAGA (Wolf and Theis, 2018)
- URD (Regev and Schier, 2018)

cmgh CELLULAR AND MOLECULAR
GASTROENTEROLOGY AND HEPATOLOGY

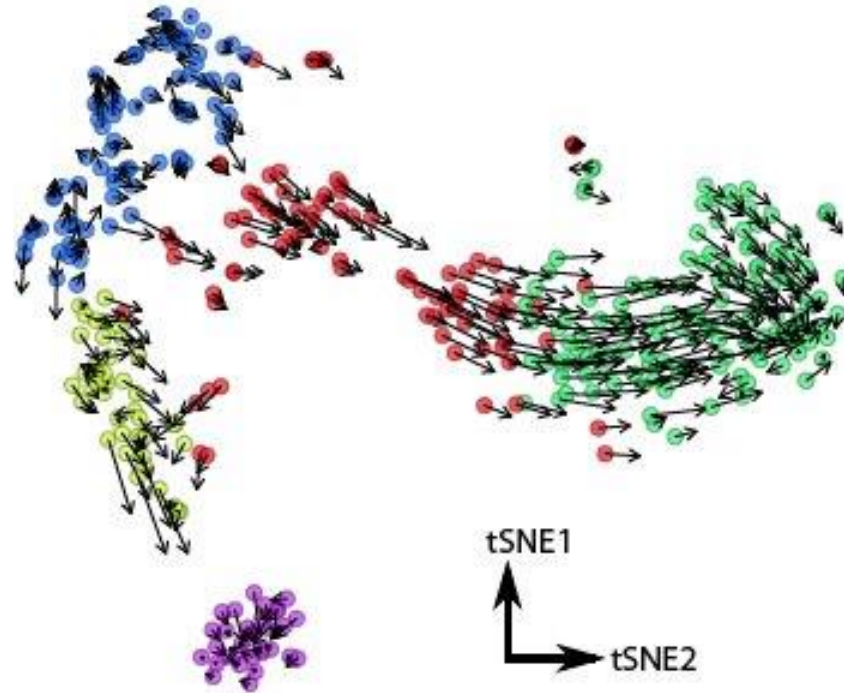
REVIEW

**Single-Cell Computational Strategies for Lineage Reconstruction
in Tissue Systems**

Charles A. Herring,^{1,2} Bob Chen,^{1,3} Eliot T. McKinley,^{1,4} and Ken S. Lau^{1,2,3}

¹Epithelial Biology Center, ²Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee; ³Program in Chemical and Physical Biology, ⁴Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, Tennessee

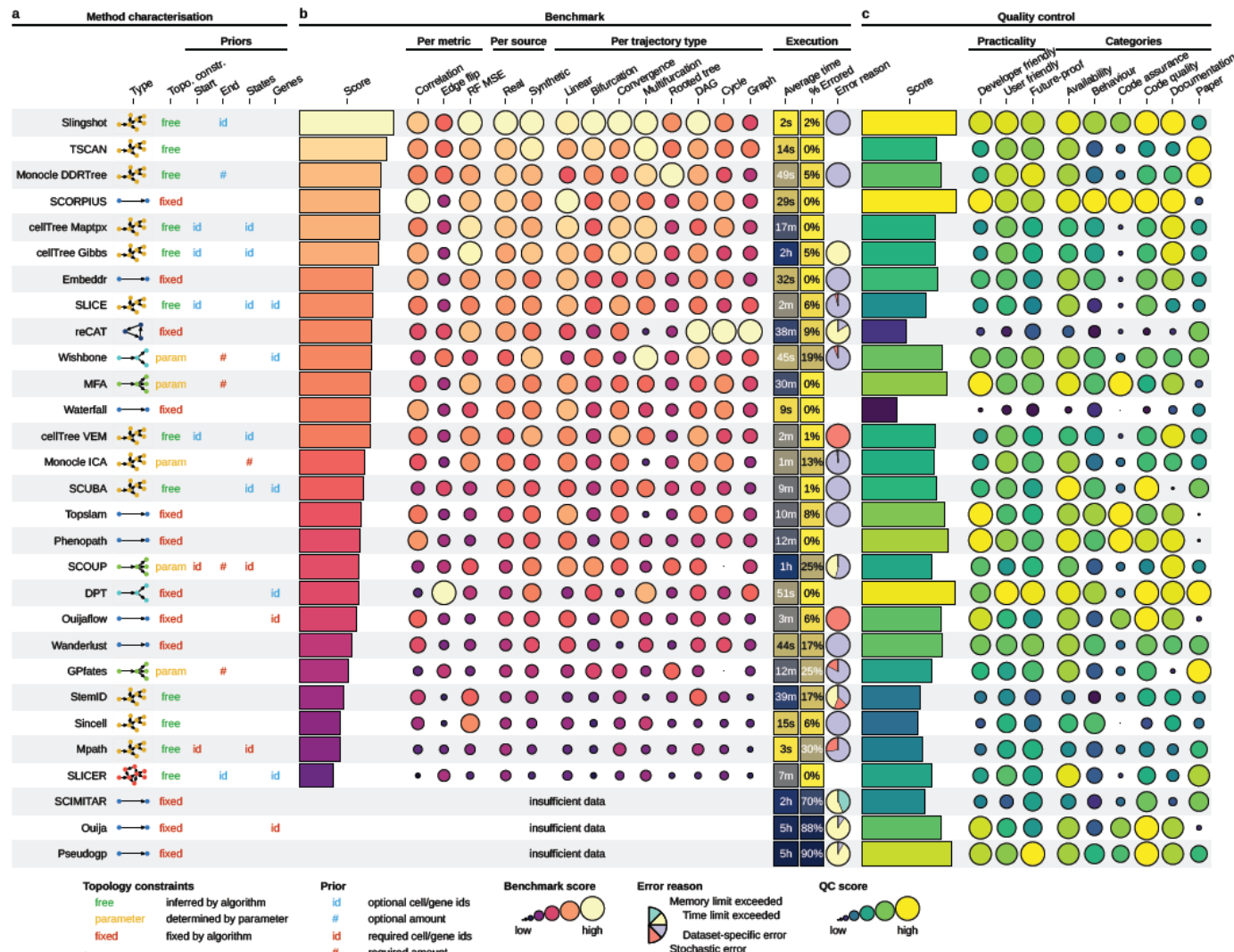
Methods that cheat (include prior/extra knowledge)



- Velocityto (Kharchenko - splicing)
- SLICE (Yan Xu – gene ontology grouping entropy)
- scdiff (Ding et al. – TF→gene expression relationships)
- SCENT (Teschendorff et al. - protein-protein interaction node entropy)

Problem: Depends on how good is the prior knowledge

Single-cell data analysis is a field in its infancy





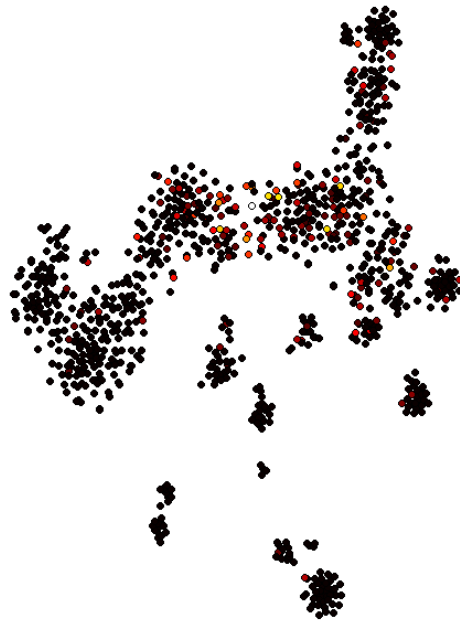
The p-Creode algorithm

Unsupervised Trajectory Analysis of Single-Cell RNA-Seq and Imaging Data Reveals Alternative Tuft Cell Origins in the Gut

Charles A. Herring,^{1,2} Amrita Banerjee,^{1,3} Eliot T. McKinley,^{1,4} Alan J. Simmons,^{1,3} Jie Ping,^{5,6} Joseph T. Roland,¹ Jeffrey L. Franklin,^{1,3} Qi Liu,^{5,6} Michael J. Gerdes,⁷ Robert J. Coffey,^{1,3,4,8} and Ken S. Lau^{1,2,3,6,9,*}

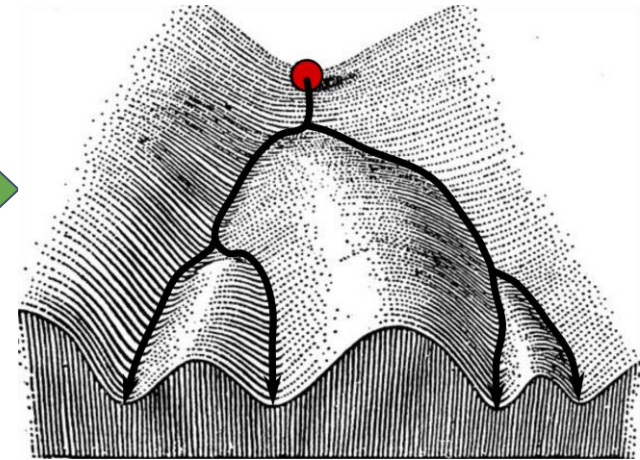
<https://github.com/KenLauLab/pCreode>

Single-cell data continuum represented as transitional trajectories



p-Create

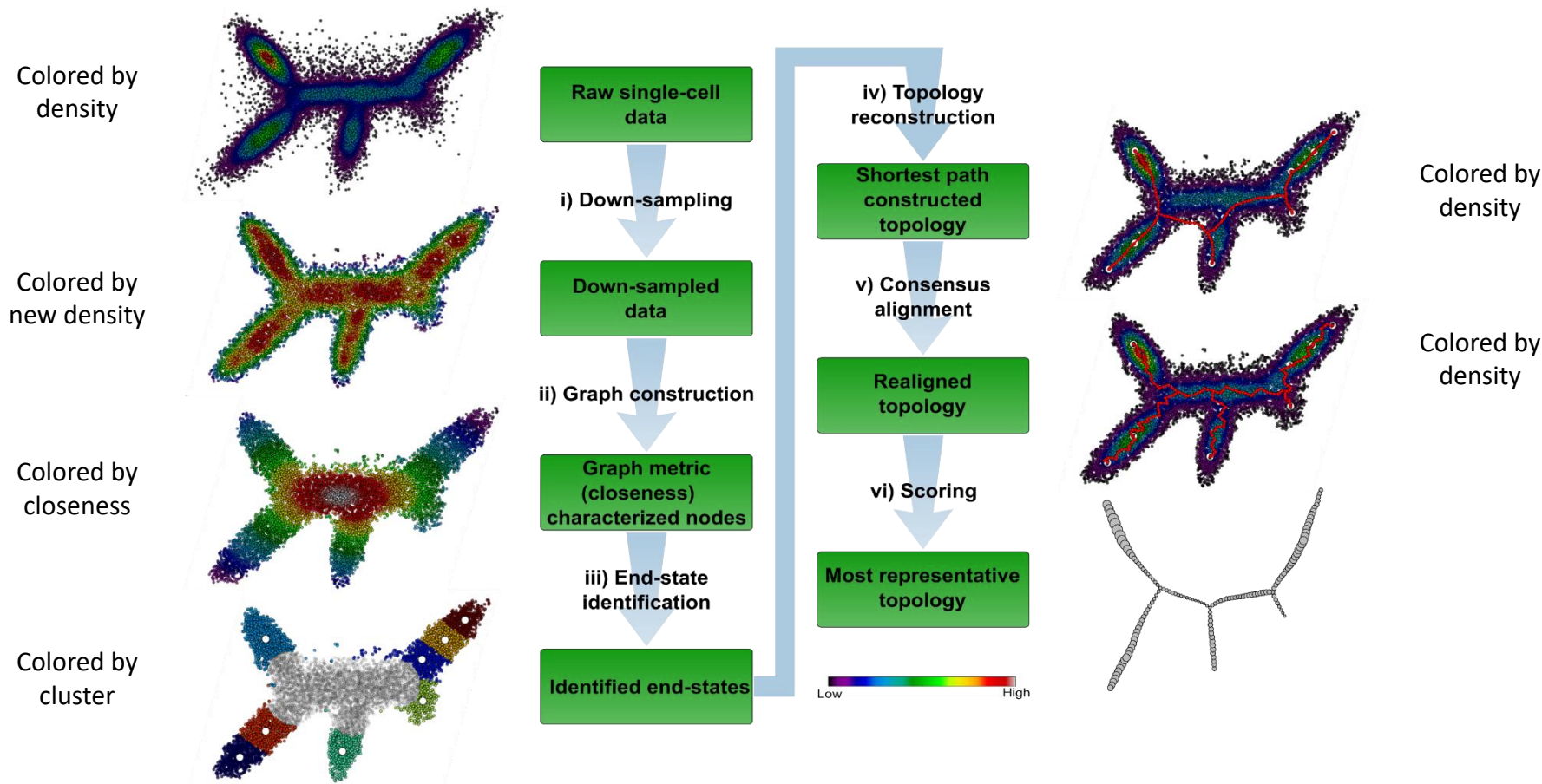
(Herring, Banerjee et al.,
2018, Cell Systems)



Waddington's
Landscape

putative-Creode Overview

The term creode was coined by C.H. Waddington, combining the Greek words for “necessary” and “path”



p-Creode characteristics

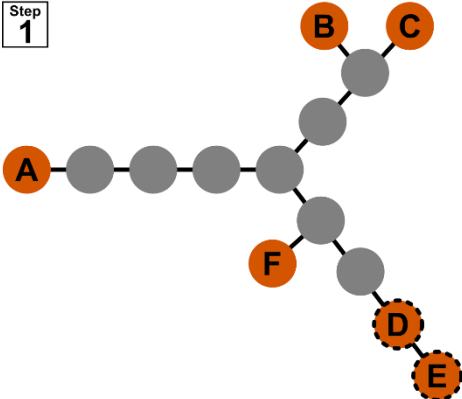
- Map complex multi-branching topologies
- Results reproducible when run on well-regulated processes
- Ensemble approach for deconvolving dysregulated transition processes

Three major novel developments of p-Creode

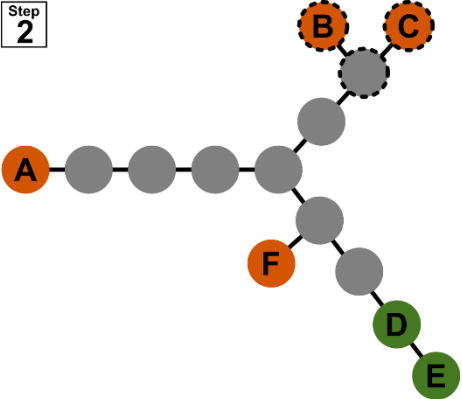
- Unsupervised end state identification
- Hierarchical placement algorithm (coupled to consensus alignment)
- Unique scoring method that leverages an ensemble approach to statistically assess the quality of a result

p-Creode's hierarchical placement strategy

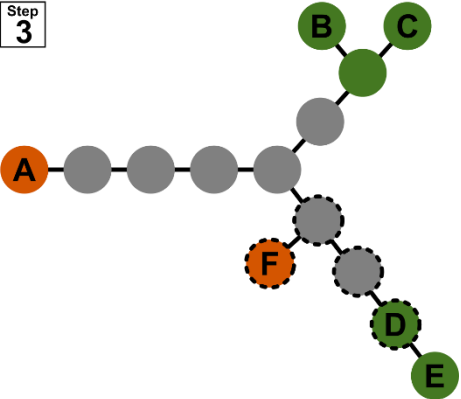
Step 1



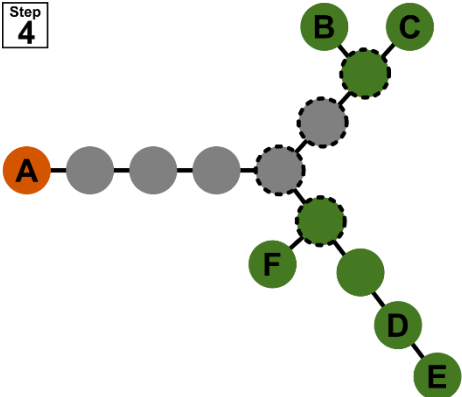
Step 2



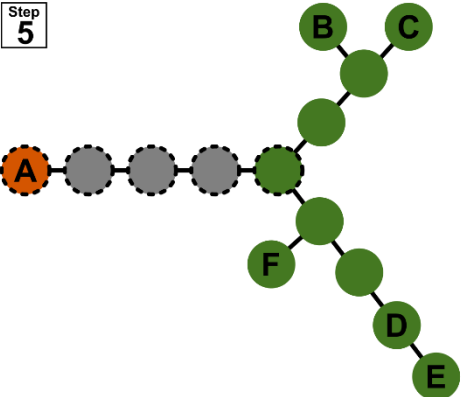
Step 3



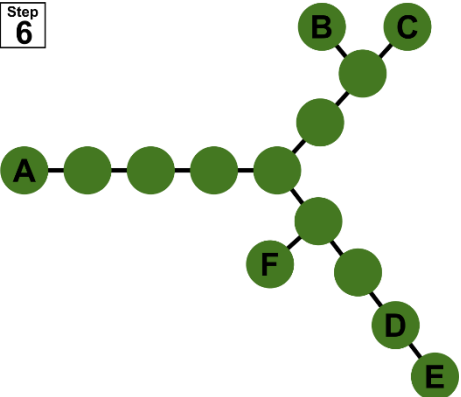
Step 4




Step 5

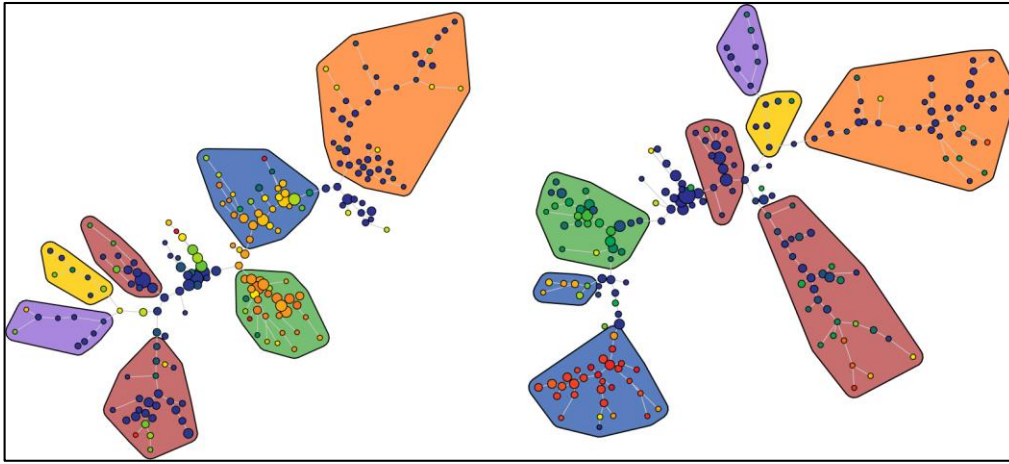


Step 6

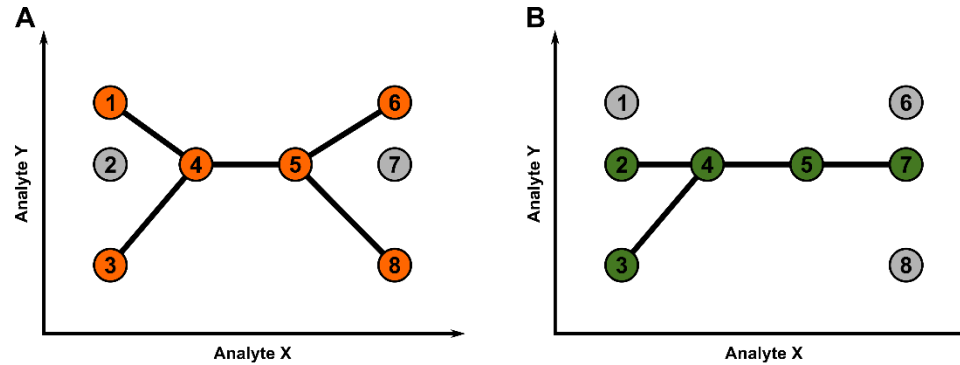


Unconnected End-states  Unconnected Nodes  Connected Nodes  Shortest Path Nodes 

p-Creode scoring method



All graphs are different, some more than other



Breaking New Ground in the Landscape of Single-Cell Analysis

Kenji Kamimoto^{1,2,3} and Samantha A. Morris^{1,2,3,*}

¹Department of Developmental Biology

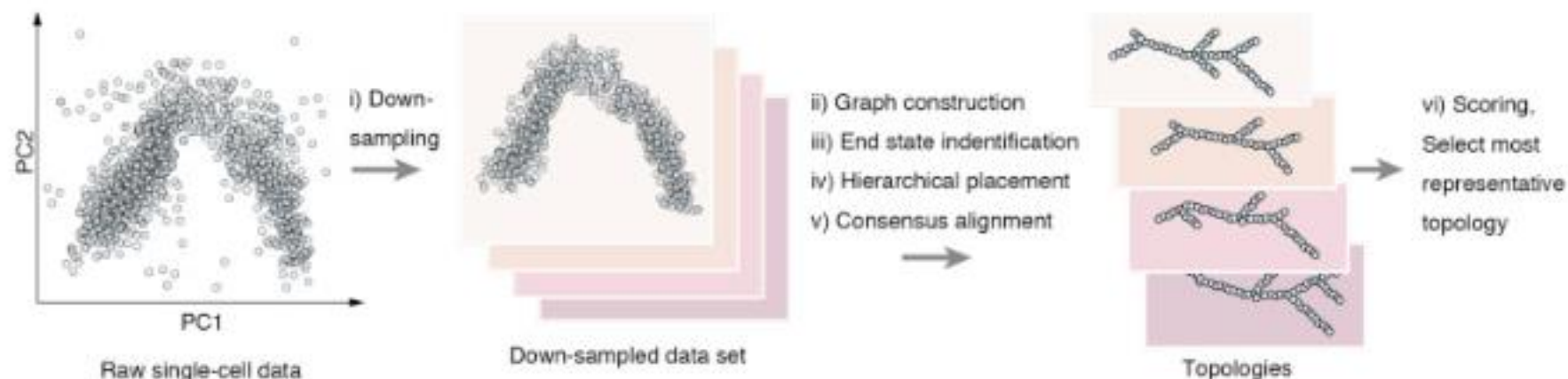
²Department of Genetics

³Center of Regenerative Medicine

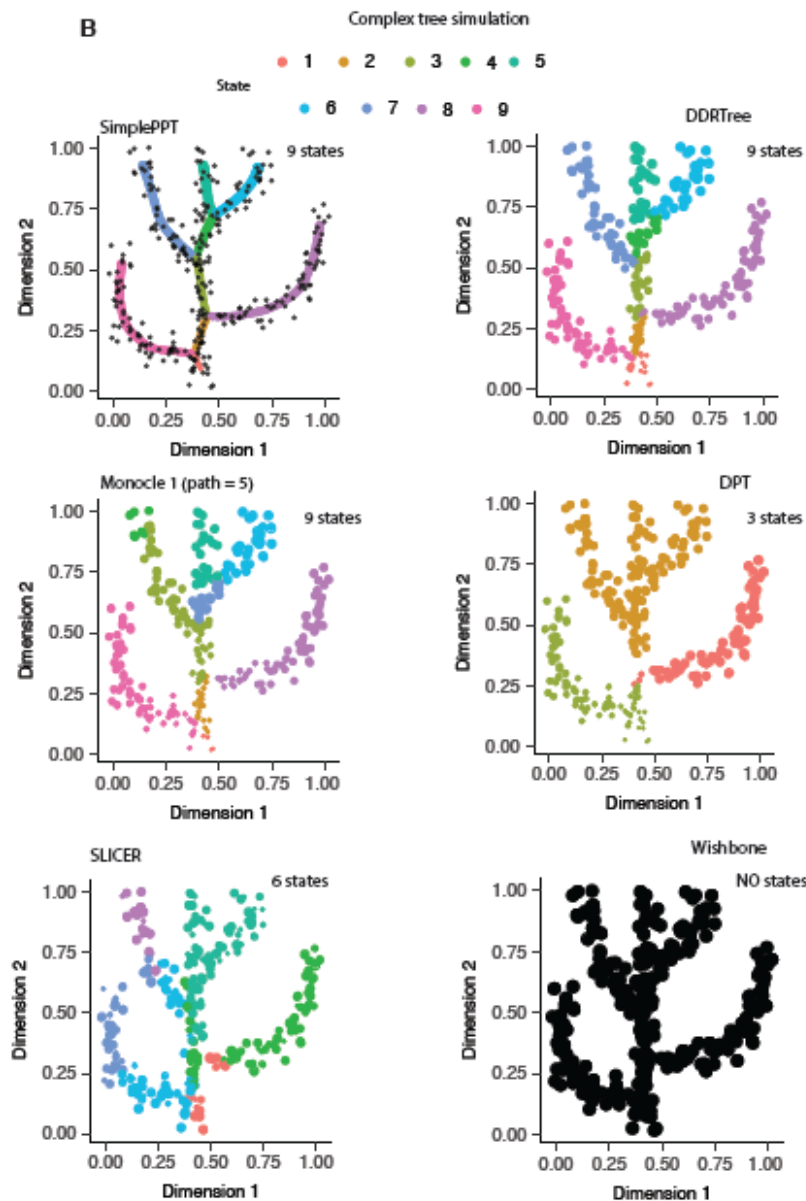
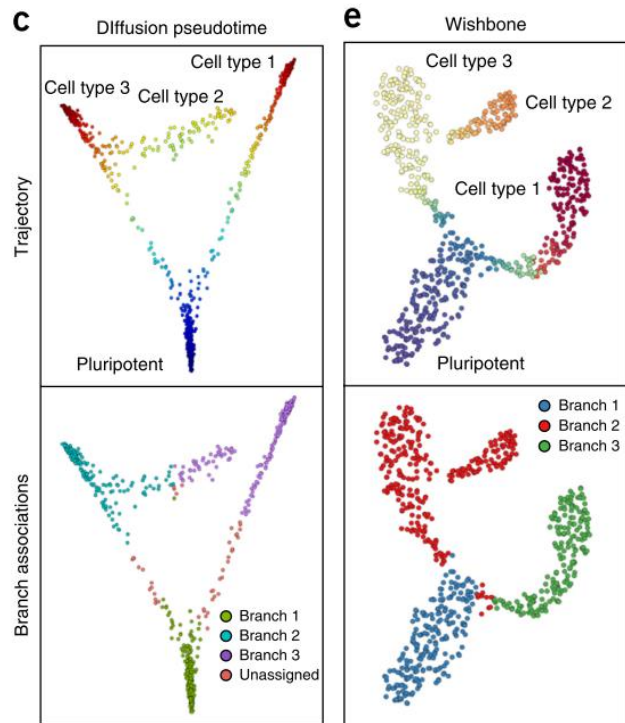
Washington University School of Medicine in St. Louis. 660 S. Euclid Avenue, Campus Box 8103, St. Louis, MO 63110, USA

*Correspondence: s.morris@wustl.edu

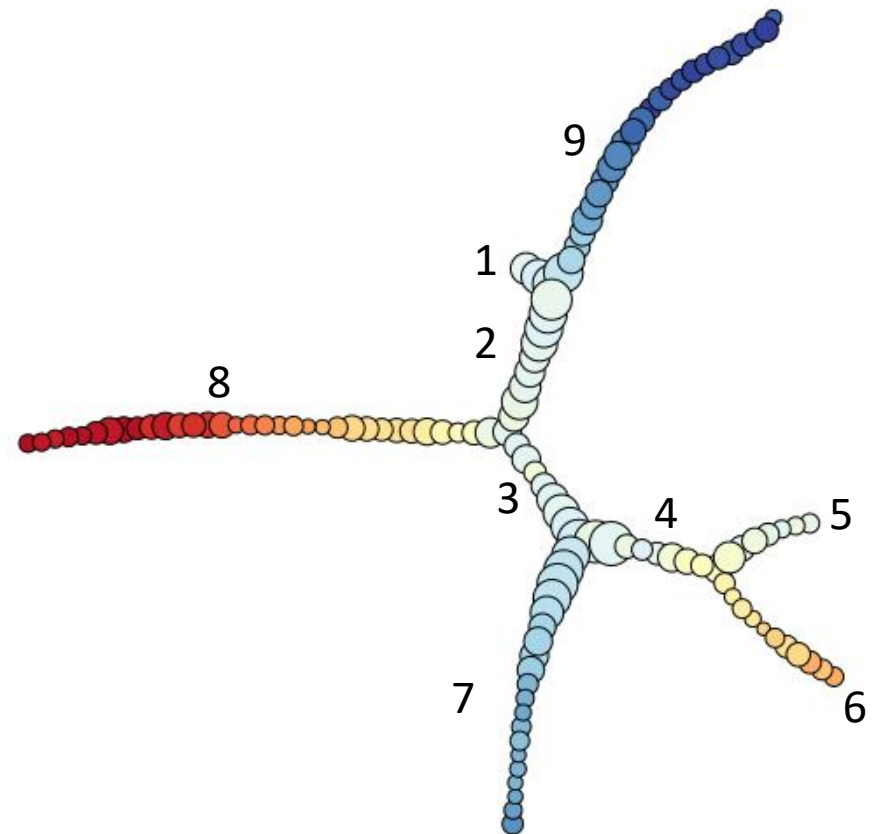
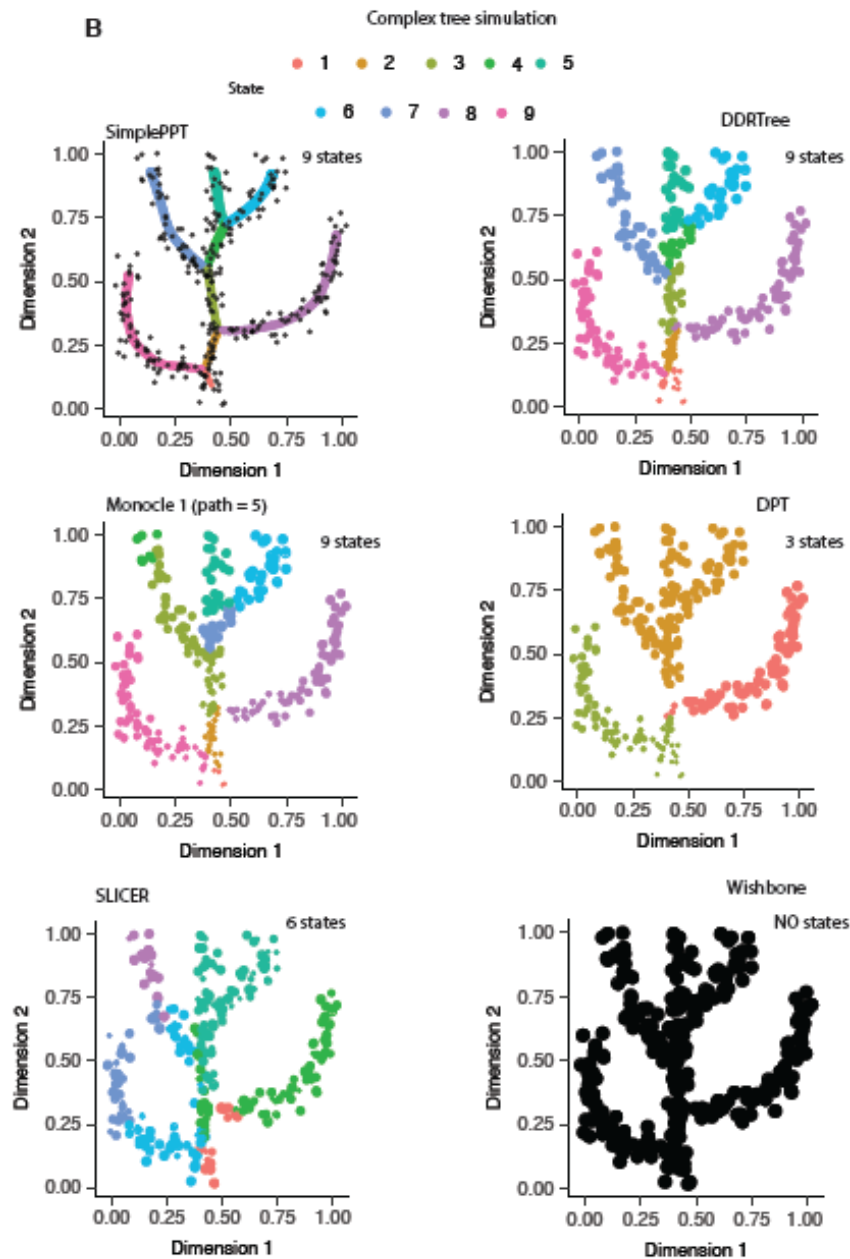
<https://doi.org/10.1016/j.cels.2017.12.015>



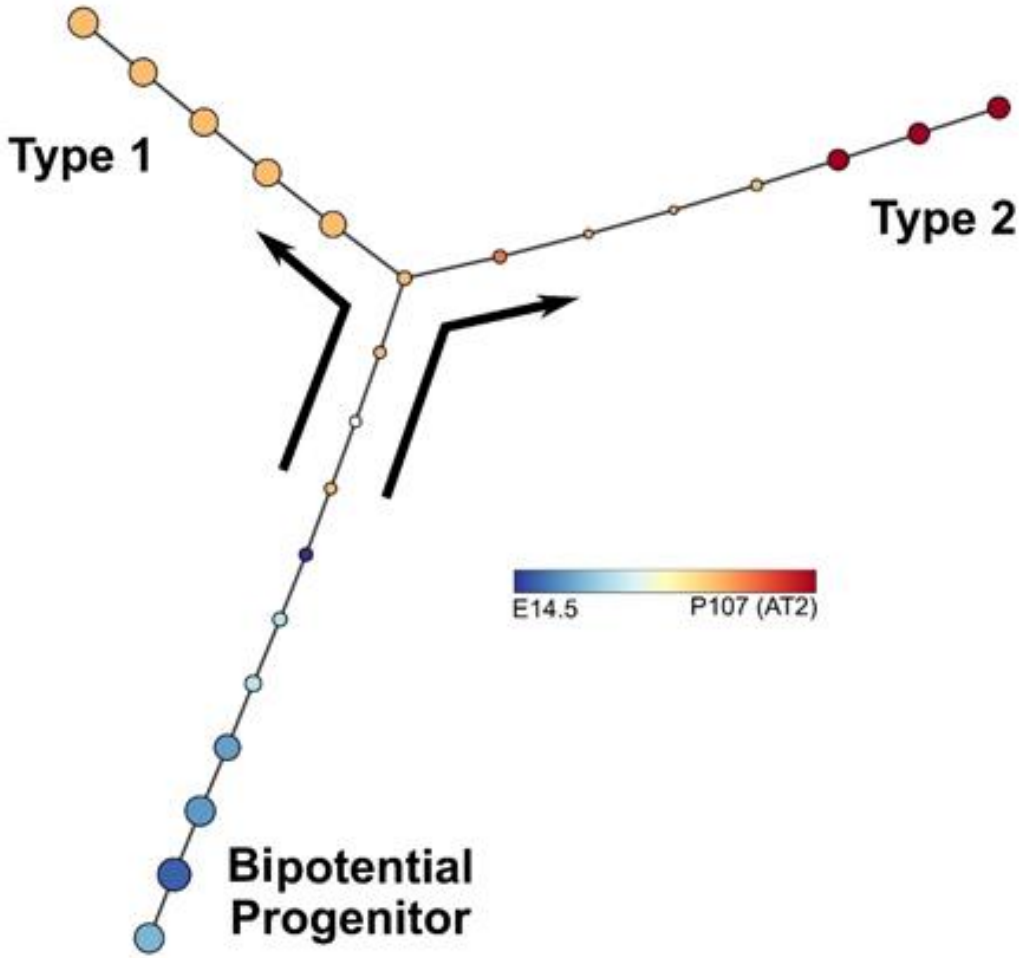
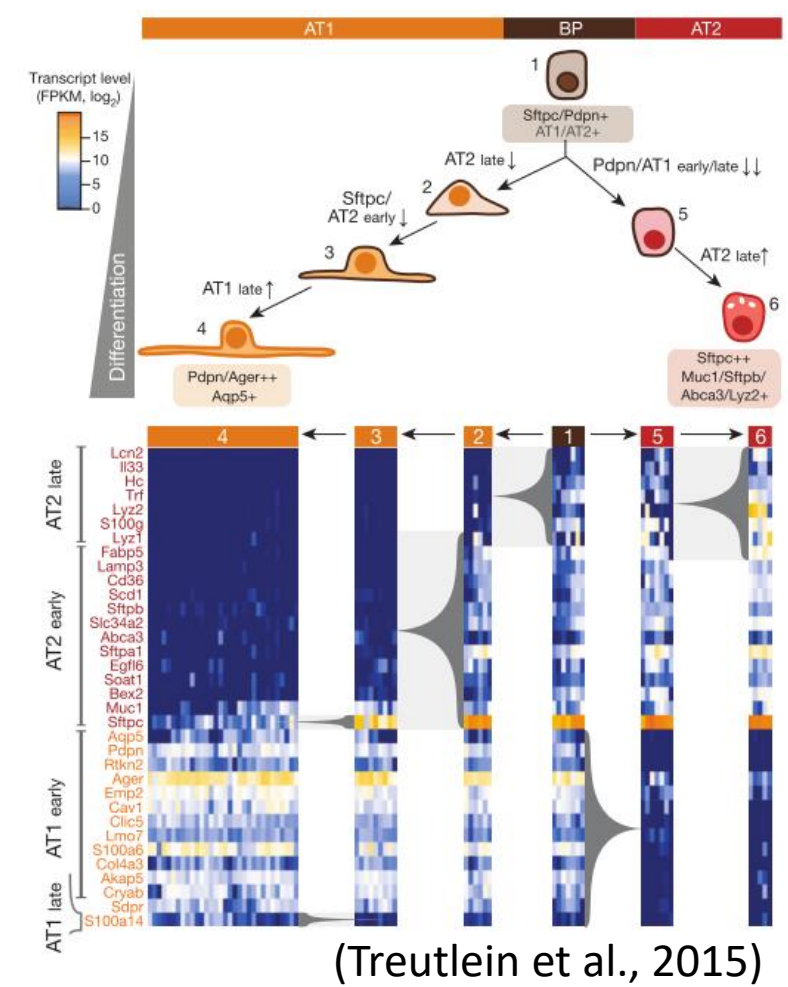
Some examples of algorithm comparisons



p-Creode application on other people's synthetic data



Developmental time of alveolar differentiation reconstructed by p- Creode applied to Fluidigm C1 sc-RNA-seq data

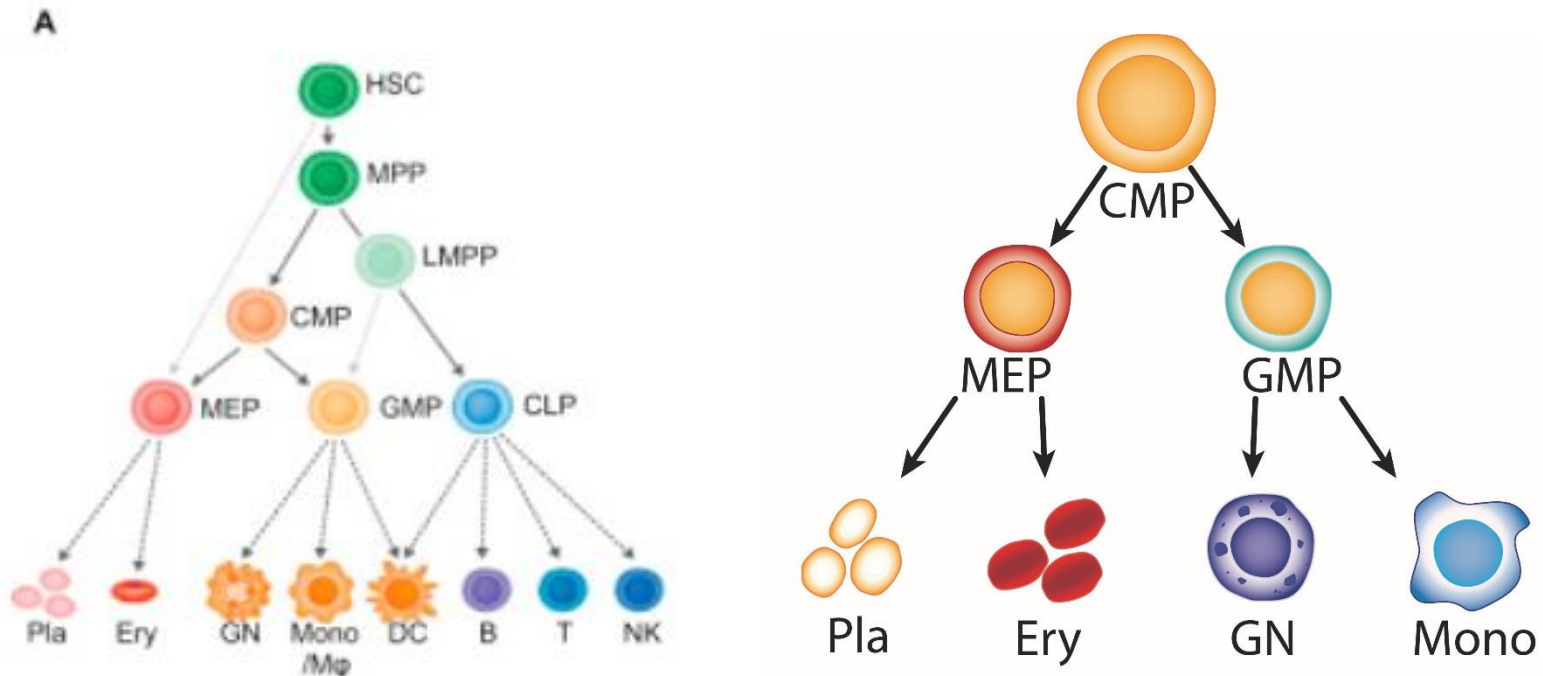


p-Creode applied to multi-branching myeloid differentiation

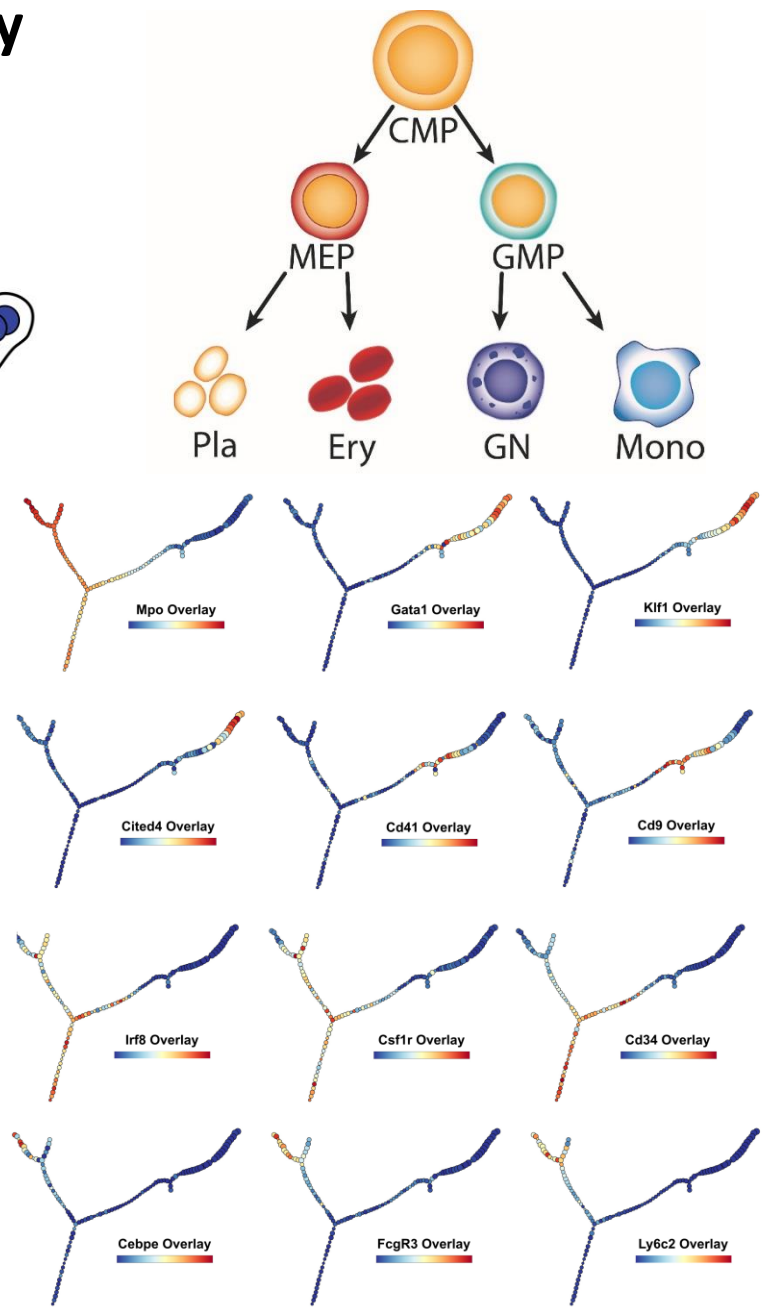
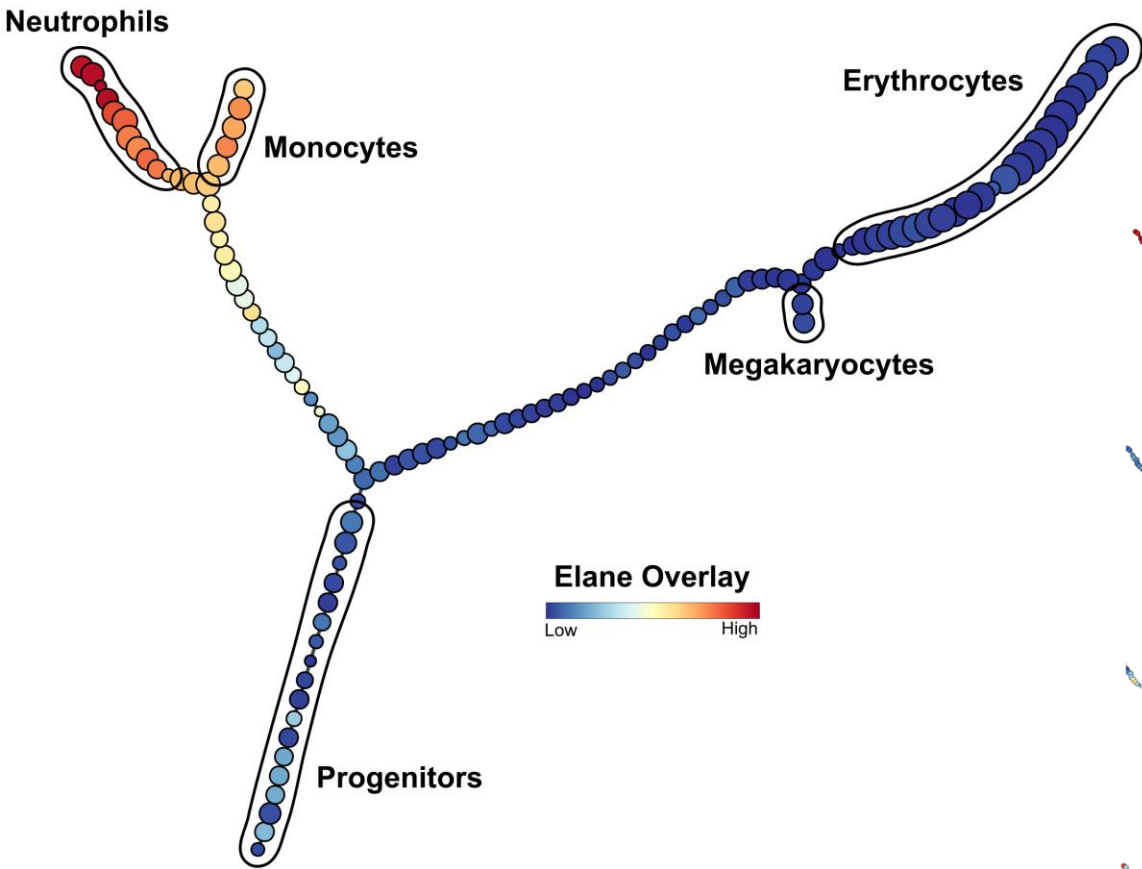
Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors

Franziska Paul,^{1,9} Ya'ara Arkin,^{2,9} Amir Giladi,^{1,9} Diego Adhemar Jaitin,¹ Ephraim Kenigsberg,² Hadas Keren-Shaul,¹ Deborah Winter,¹ David Lara-Astiaso,¹ Meital Gury,¹ Assaf Weiner,¹ Eyal David,¹ Nadav Cohen,² Felicia Kathrine Bratt Lauridsen,^{3,4,5} Simon Haas,⁶ Andreas Schlitzer,^{7,8} Alexander Mildner,¹ Florent Ginhoux,⁷ Steffen Jung,¹ Andreas Trumpp,⁶ Bo Torben Porse,^{3,4,5} Amos Tanay,^{2,10,*} and Ido Amit^{1,10,*}

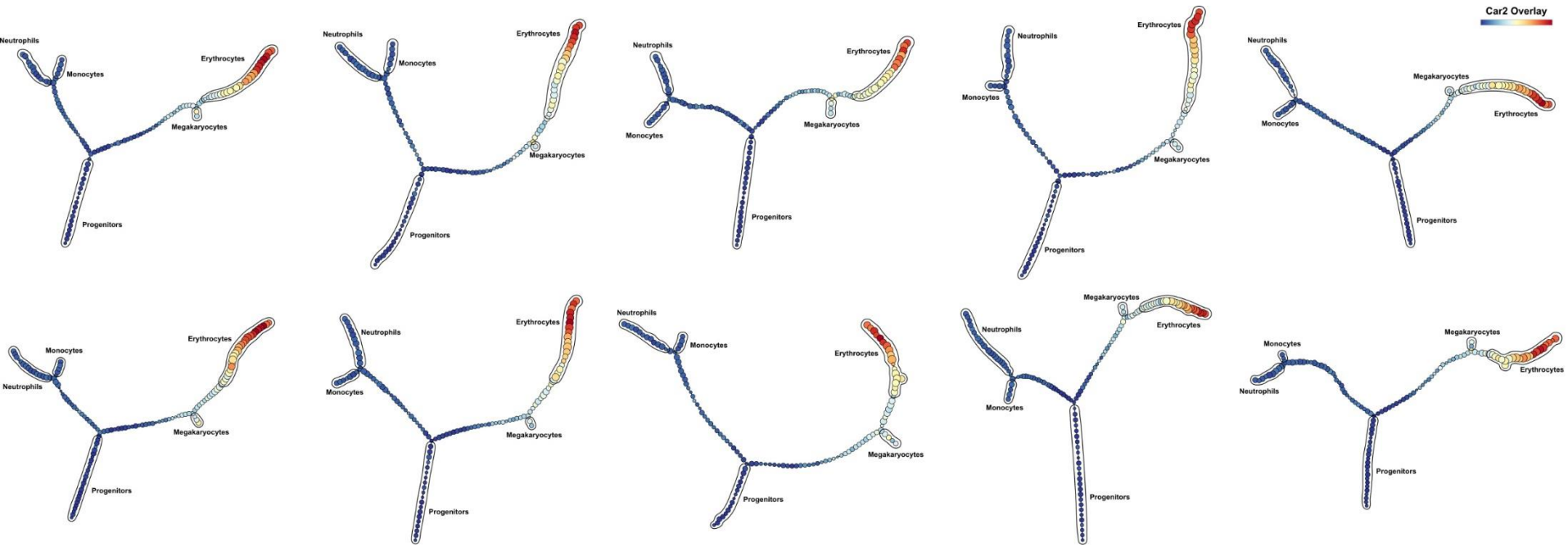
Plate-based MARS-seq - >4000 cells



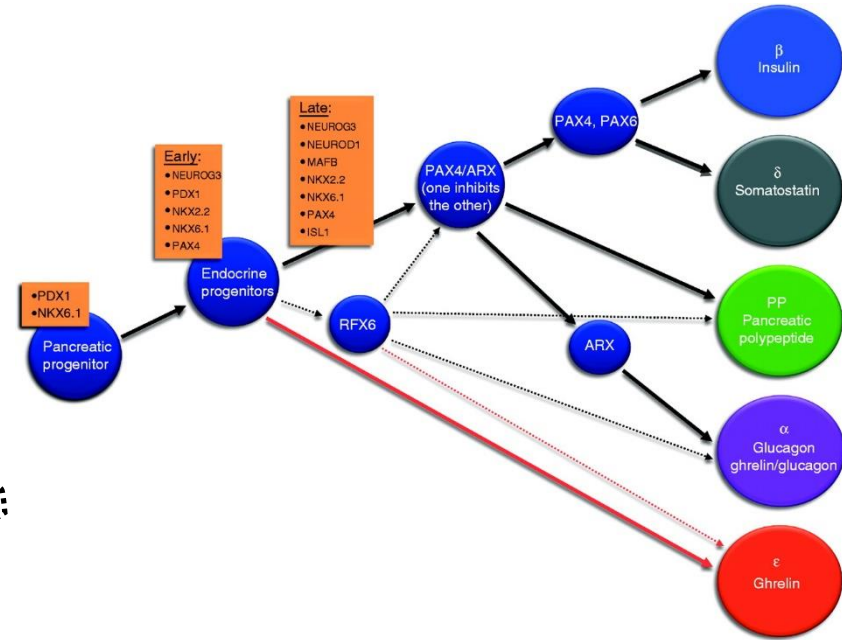
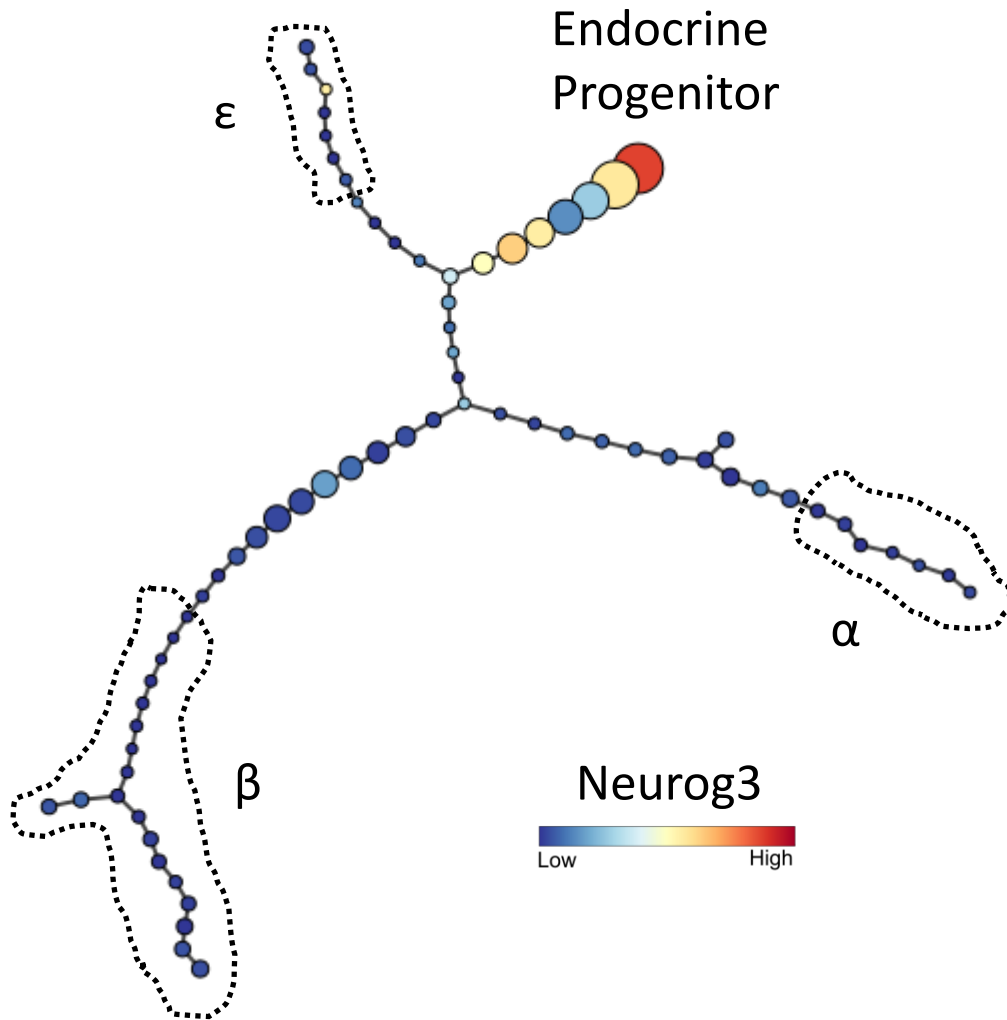
p-Creode reconstructs complex hierarchy



Robustness of p-Creode demonstrated on complex trajectories



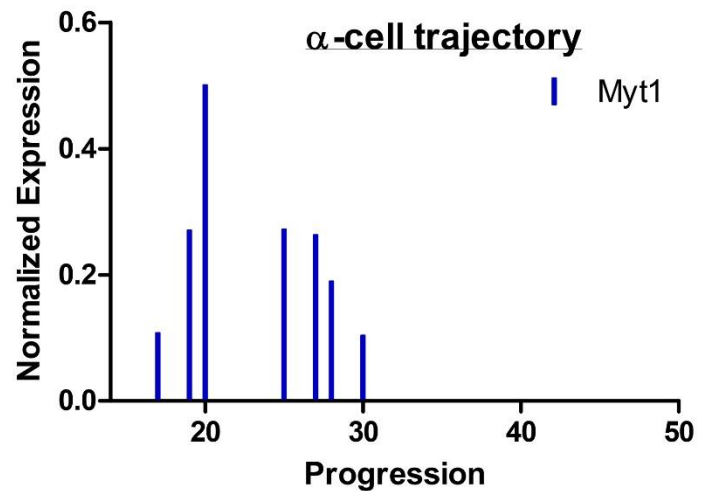
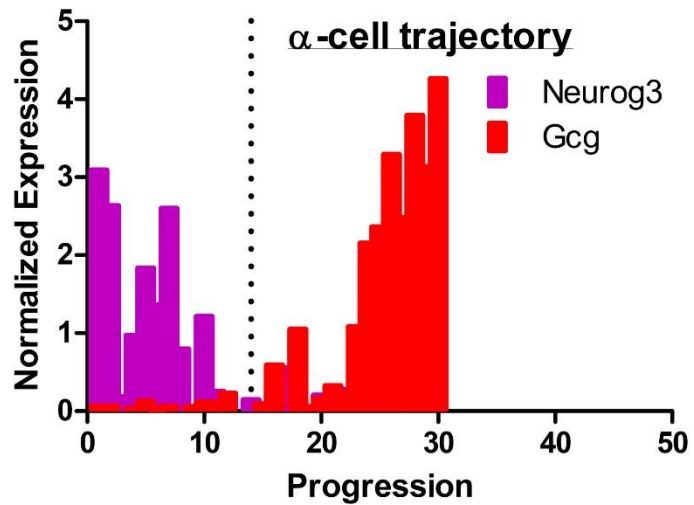
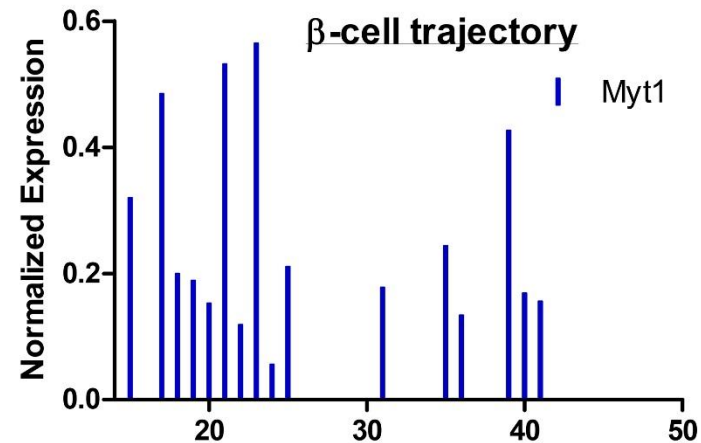
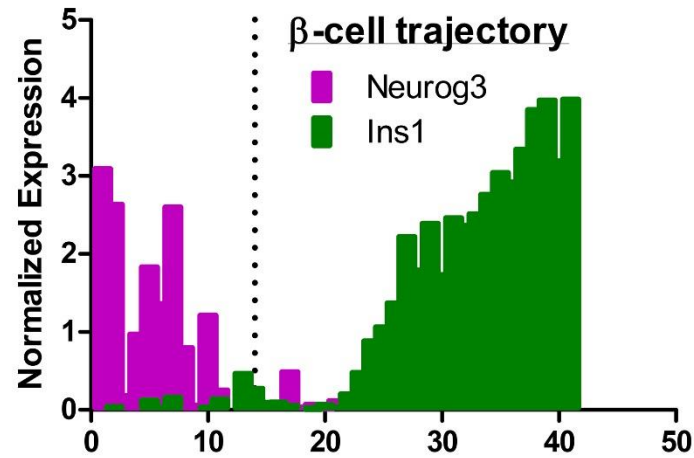
p-Creode generates epsilon, beta, alpha cell trajectories for E14.5 pancreas



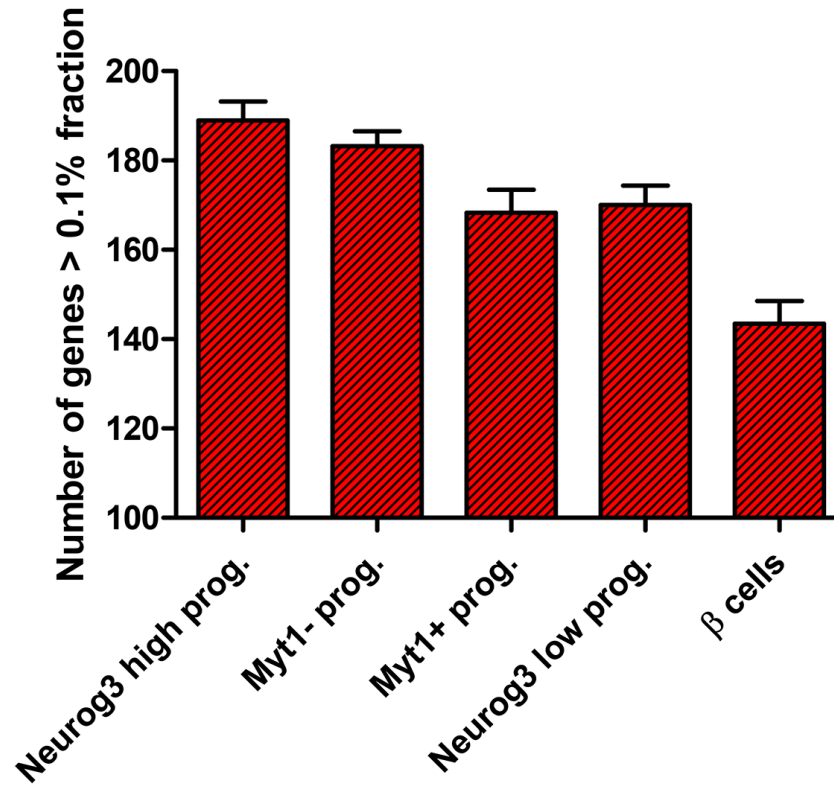
(Granata et al., J Mol Endocrinol, 2010)

(Liu et al., *Developmental Cell*, in revision)

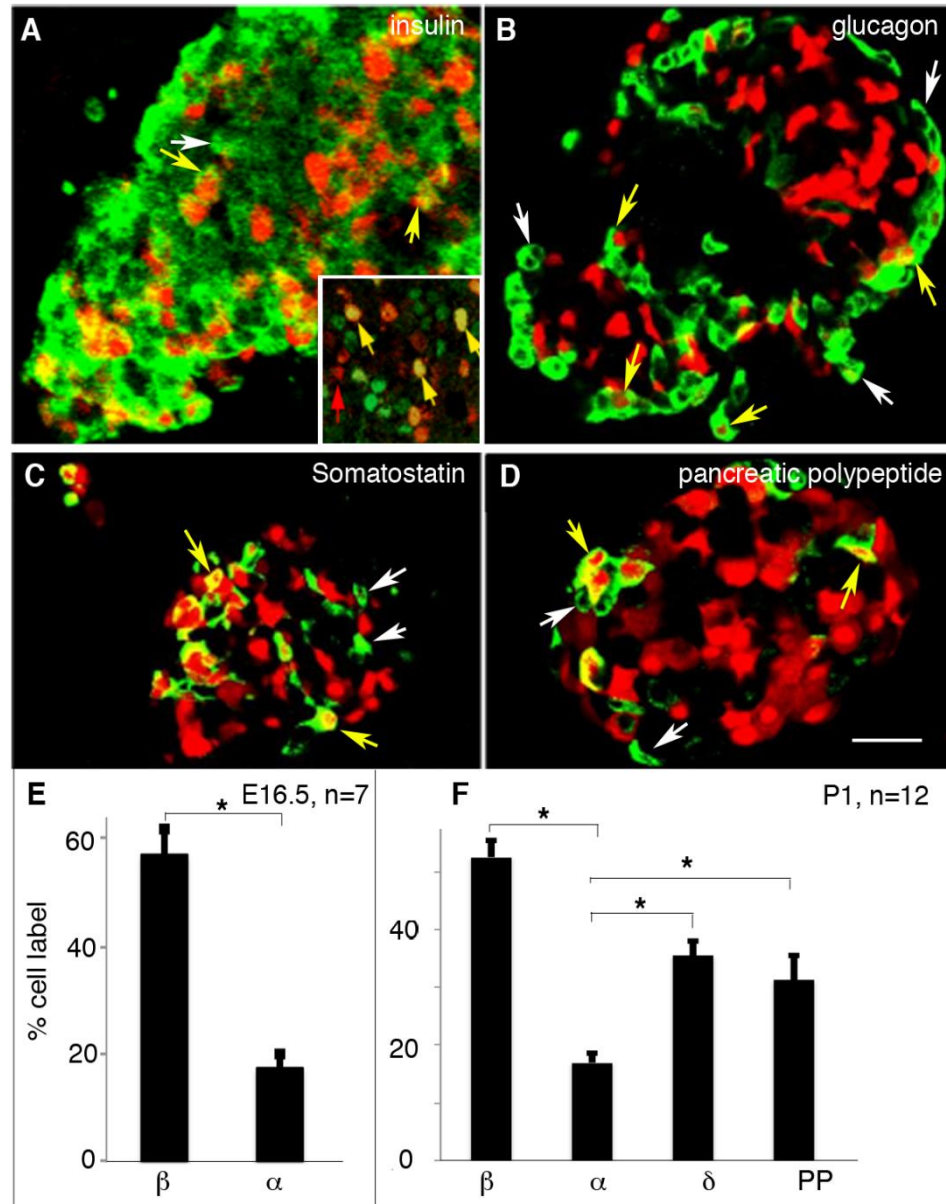
Trajectory analysis demonstrates the beta cell lineage bias of Myt1 expression



Trajectory analysis demonstrates the beta cell lineage bias of Myt1 expression



Ngn3/Myt1 lineage tracing is biased towards beta cells compared to alpha cells



Intestinal tuft cell biology

- ~1.0% epithelial cells in mouse intestine
- Express chemoreceptors (TRPM5, etc.)
- Characteristic “tufted” morphology due to apical acetylated tubulin bundle

Tuft cells, taste-chemosensory cells, orchestrate parasite type 2 immunity in the gut

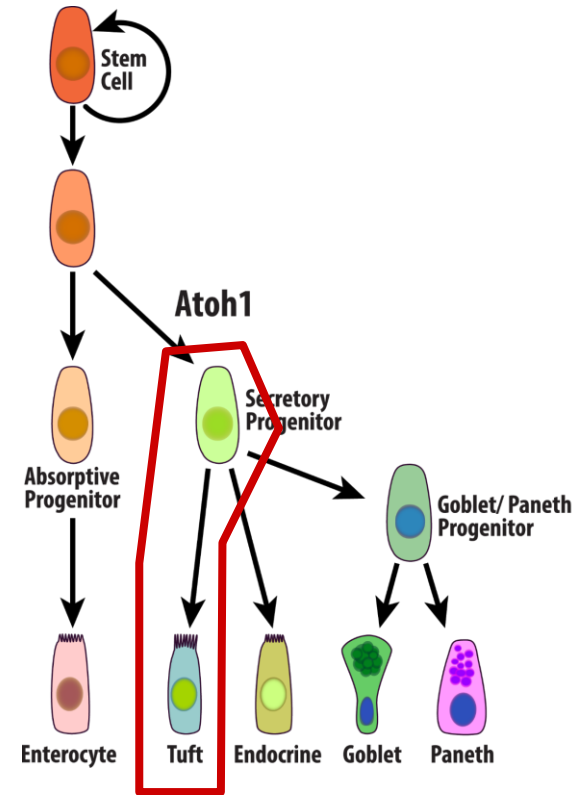
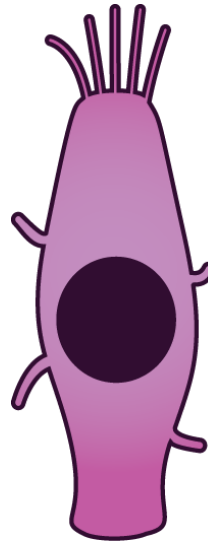
Michael R. Howitt,¹ Sydney Lavoie,¹ Monia Michaud,¹ Arthur M. Blum,² Sara V. Tran,³ Joel V. Weinstock,² Carey Ann Gallini,¹ Kevin Redding,³ Robert F. Margolskee,³ Lisa C. Osborne,^{4*} David Artis,⁴ Wendy S. Garrett^{1,5,6†}

Intestinal epithelial tuft cells initiate type 2 mucosal immunity to helminth parasites

François Gerbe^{1,2,3}, Emmanuelle Sidot^{1,2,3}, Danielle J. Smyth^{4†}, Makoto Ohmoto⁵, Ichiro Matsumoto⁵, Valérie Dardalhon^{3,6}, Pierre Cesses^{1,2,3}, Laure Garnier^{1,2,3}, Marie Pouzolles^{3,6}, Bénédicte Brulin^{1,2,3}, Marco Bruschi^{1,2,3}, Yvonne Harnusch⁴, Valérie S. Zimmermann^{3,6}, Naomi Taylor^{3,6}, Rick M. Maizels^{4†} & Philippe Jay^{1,2,3}

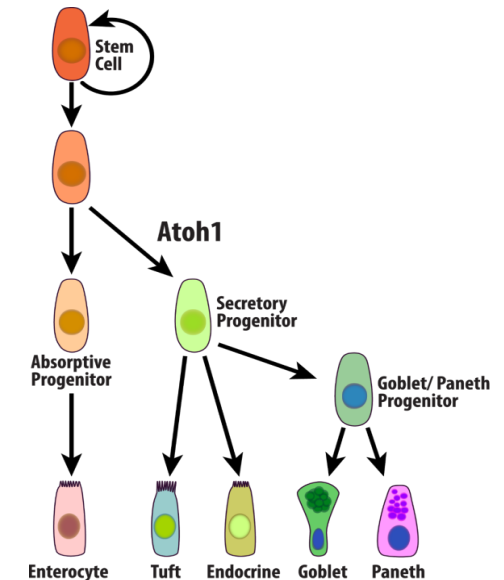
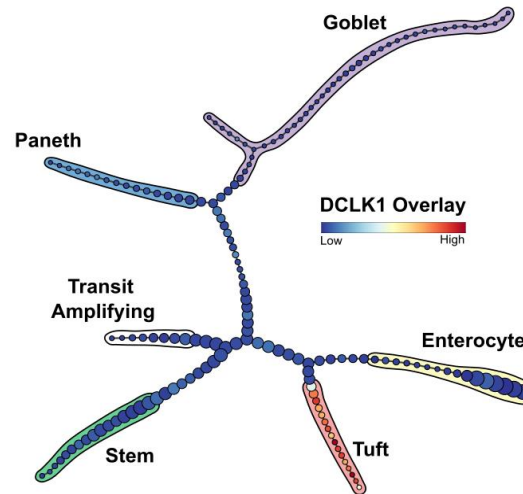
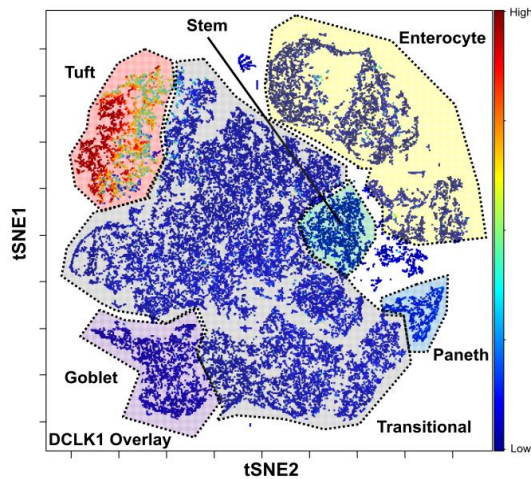
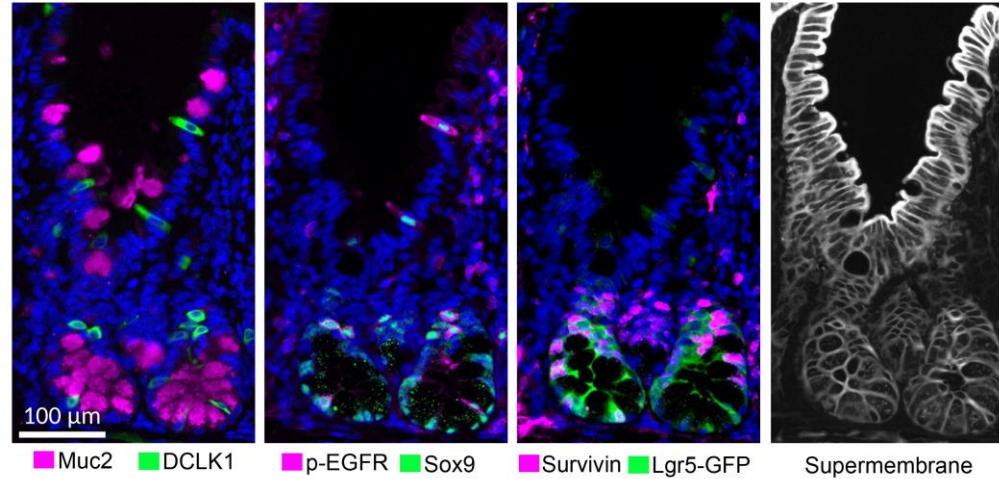
Tuft-cell-derived IL-25 regulates an intestinal ILC2-epithelial response circuit

Jakob von Moltke¹, Ming Ji^{1,2}, Hong-Erh Liang¹ & Richard M. Locksley^{1,2,3}

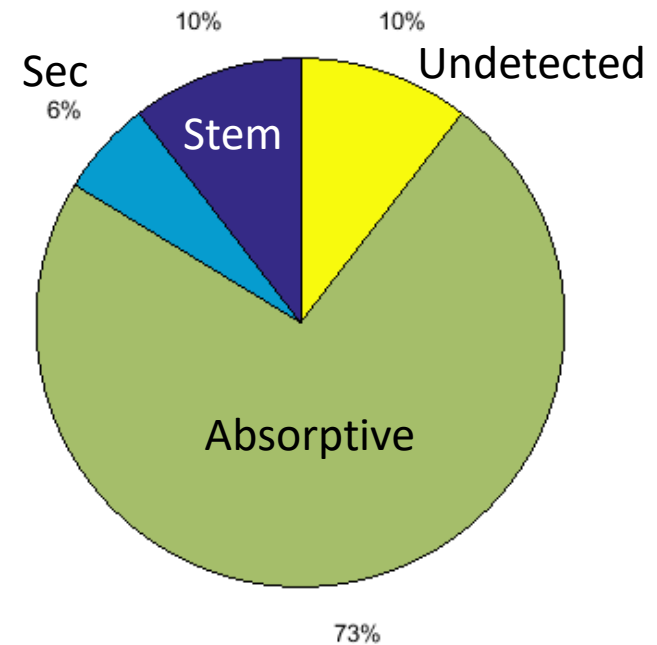
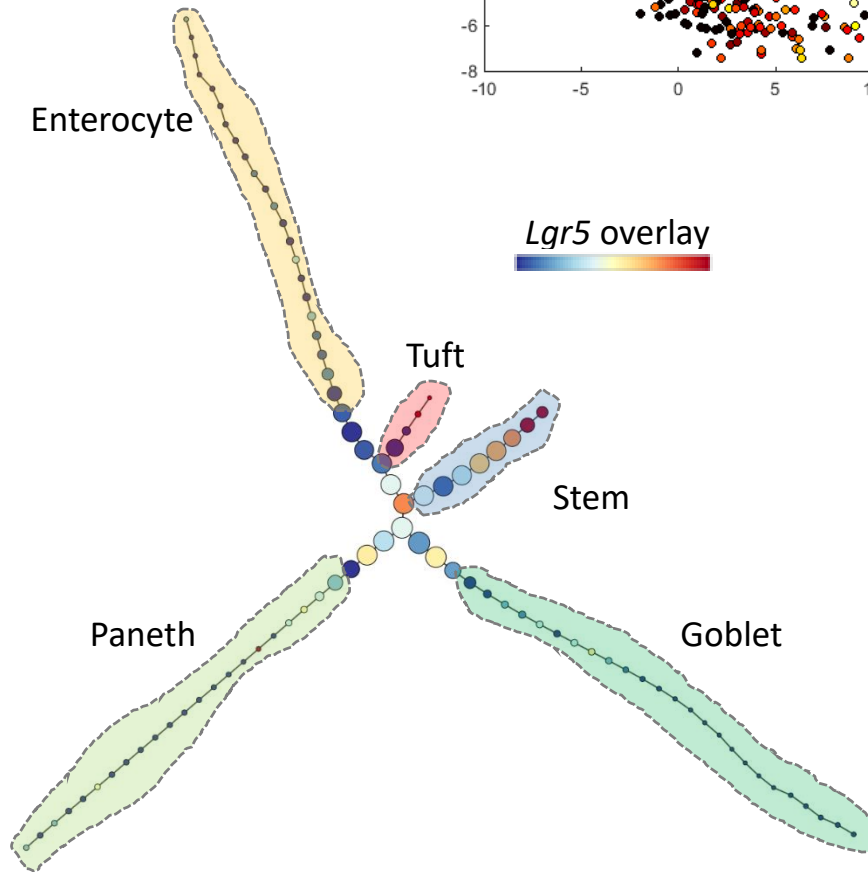
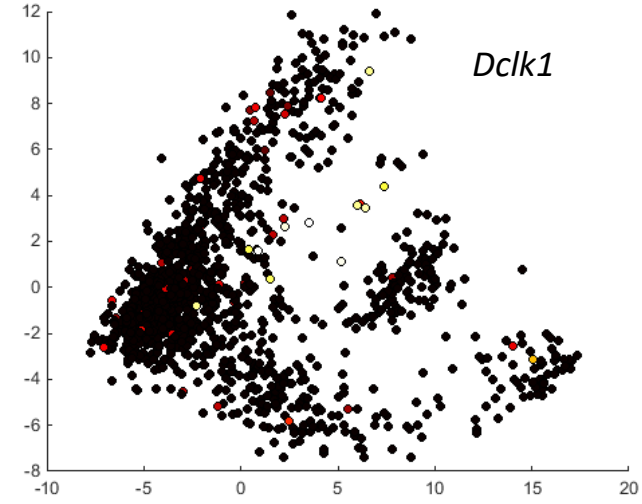
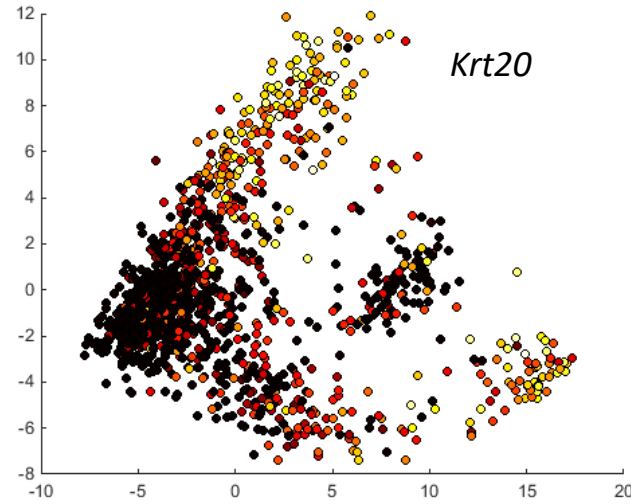


(Gerbe et al., 2011)

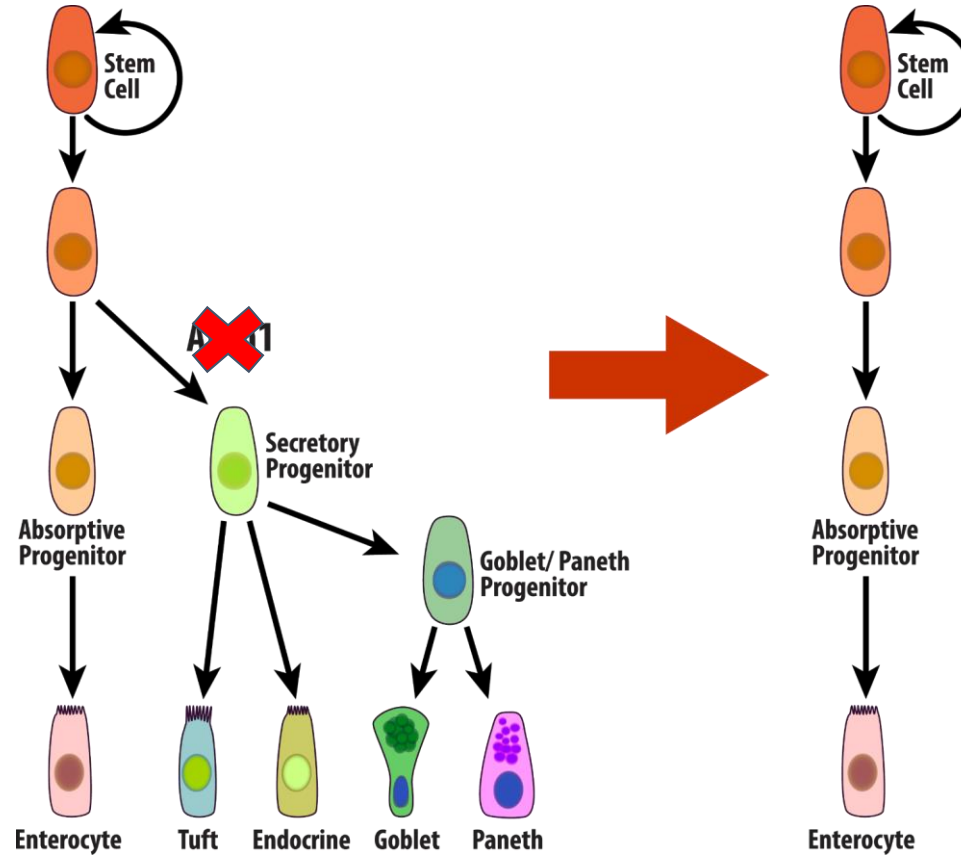
p-creode suggests tuft cells are specified outside of secretory lineage in the small intestine



p-Create suggests tuft cells are specified outside of secretory lineage in the small intestine

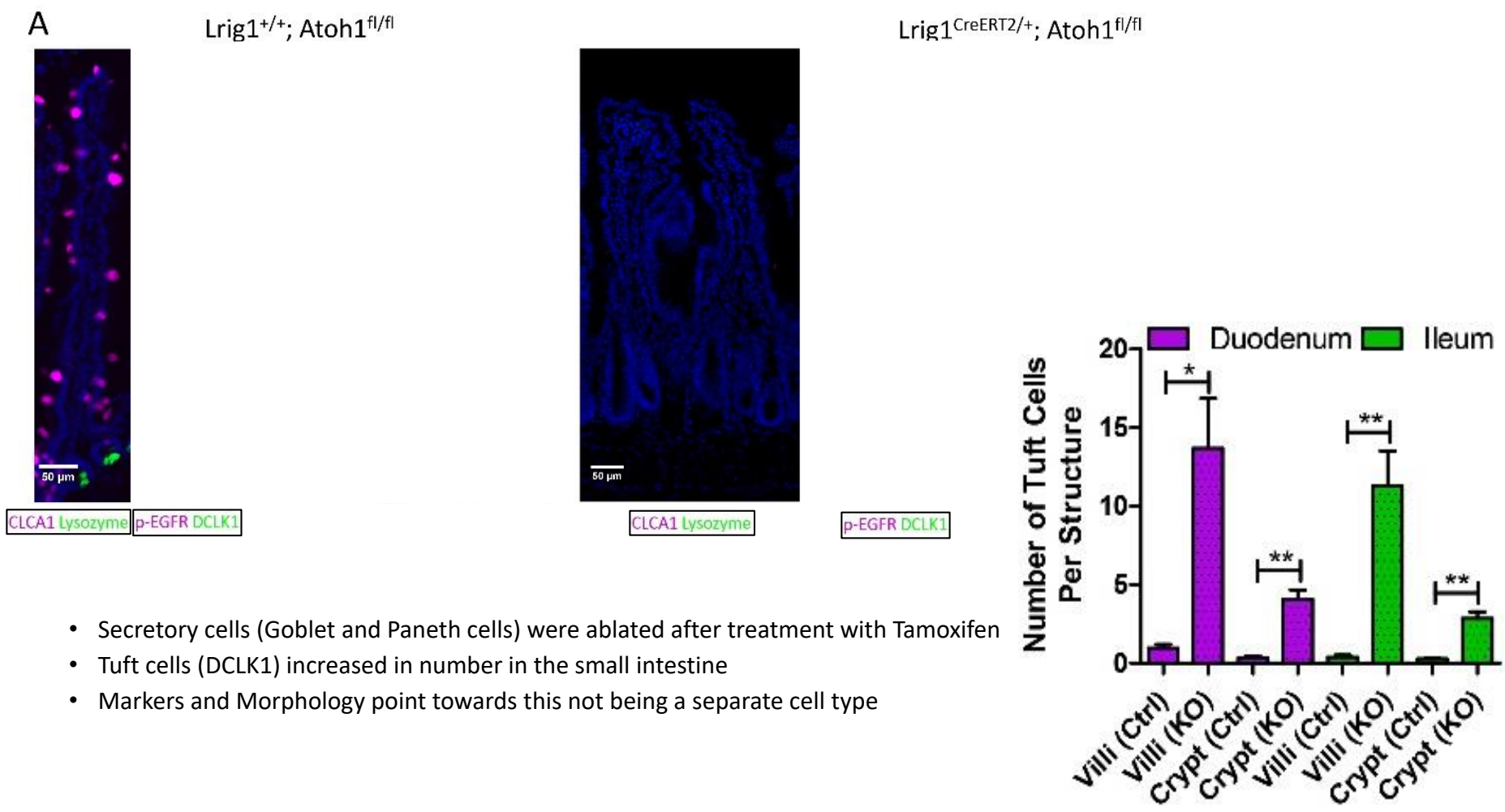


The Atoh1 transcription factor is an early determinant of secretory fate



Epithelial specific Atoh1 KO suppresses secretory cells but increases tuft cells in the small intestine

Small Intestine



Epithelial specific Atoh1 KO suppresses secretory cells but increases tuft cells in the small intestine

Small Intestine

