



Predicting Cellular Reprogramming Trajectories with Deep Generative Models: A Case Study with Epithelial-Mesenchymal Transition (EMT)

Samia Islam^{1,3}, Sudin Bhattacharya^{2,3}

¹Department of Computer Science and Engineering

²Departments of Biomedical Engineering and Pharmacology & Toxicology

³Institute for Quantitative Health Science and Engineering
Michigan State University

BHATTACHARYA LAB @
MSU

Introduction

- Understanding cellular plasticity through trajectory inference is crucial for uncovering dynamic biological processes.
- We propose a computational framework using Variational Autoencoder (VAE) to synthesize cell state transitions between distinct cell types.
- As an initial validation, we apply this method to the well-characterized epithelial to mesenchymal Transition, a key process in development and cancer metastasis [1].



- Epithelial cells are tightly connected, polarized cells that form protective barriers and are involved in absorption and secretion.
- Mesenchymal cells are loosely organized, motile cells that play critical roles in migration, wound healing, and tissue remodeling.
- The Epithelial-Mesenchymal transition (EMT) is a reversible biological process where epithelial cells lose their cell-cell adhesion properties and gain mesenchymal features.

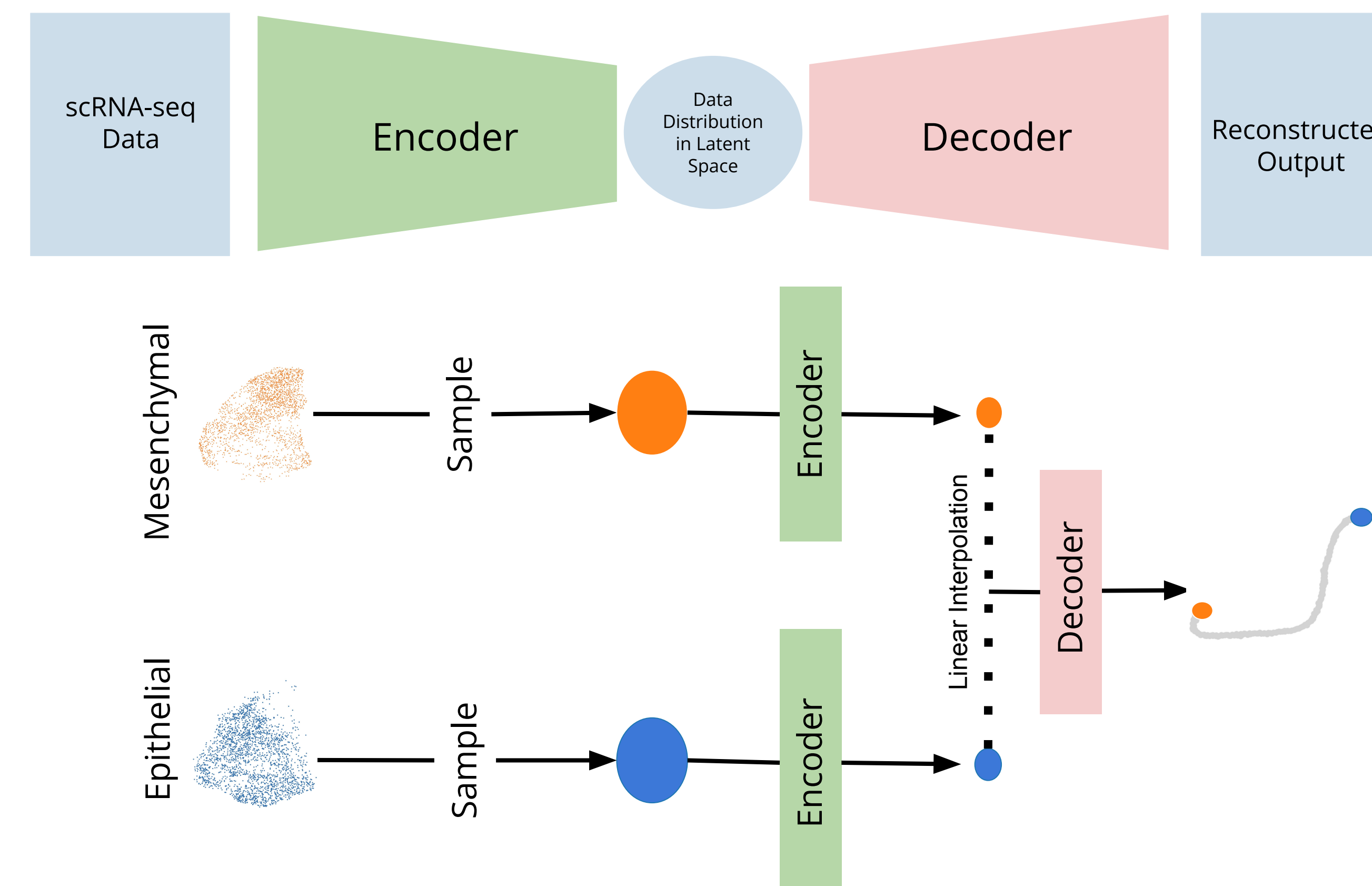
Goal and Approach

- **Simulate a Epithelial→Mesenchymal Transition:** We generate a synthetic trajectory, pinpoint potential key regulatory genes that drive the transition.
- **Validate with Prior Knowledge:** By analyzing our synthetic trajectory, we can see if we correctly recover known dynamics for EMT Transition.

Methodology

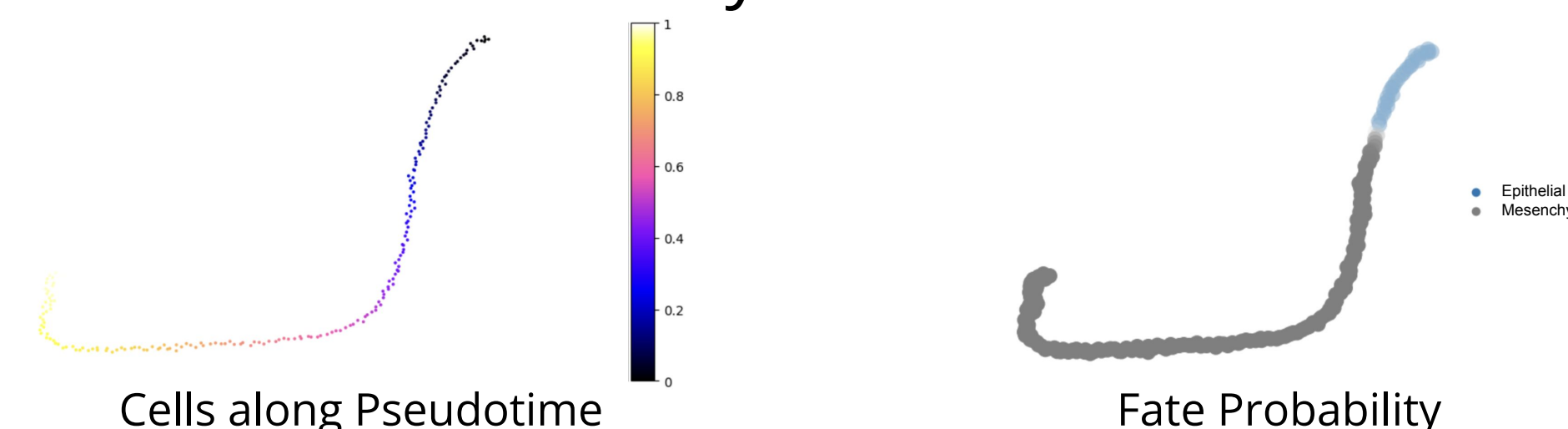
➤ Synthetic Cell Trajectory Generation:

- A Variational Autoencoder (VAE) [2] was trained with scRNA-seq data
- The encoder of the VAE gave the latent representation of the cells
- Two cells in latent space were linearly interpolated
- The decoder of the VAE projected the cells back to the original dimension



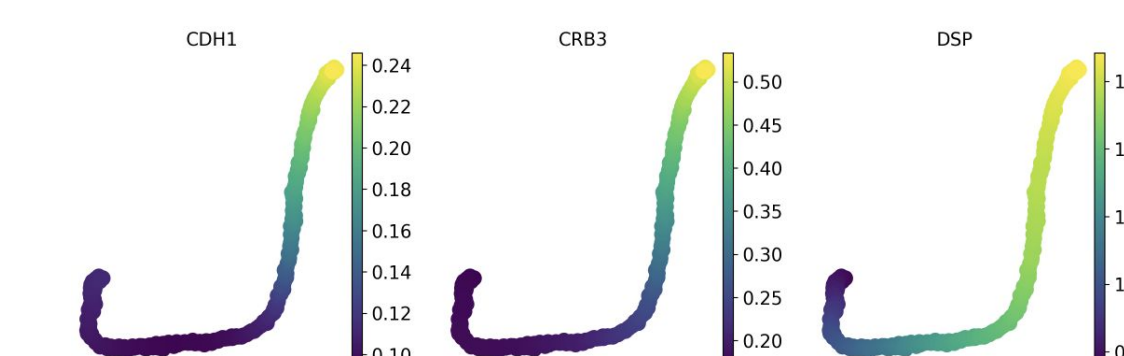
➤ Analyzing the synthetic transition:

- The cells are sorted along a pseudotime trajectory and treated the first point as "initial state" (Epithelial cell) and the last point as "terminal fate" (Mesenchymal cell)
- CellRank [3] computed the "fate probability" of becoming the Mesenchymal end state for each intermediate cell, 0% at the pure Epithelial start to 100% at the Mesenchymal end.



Results

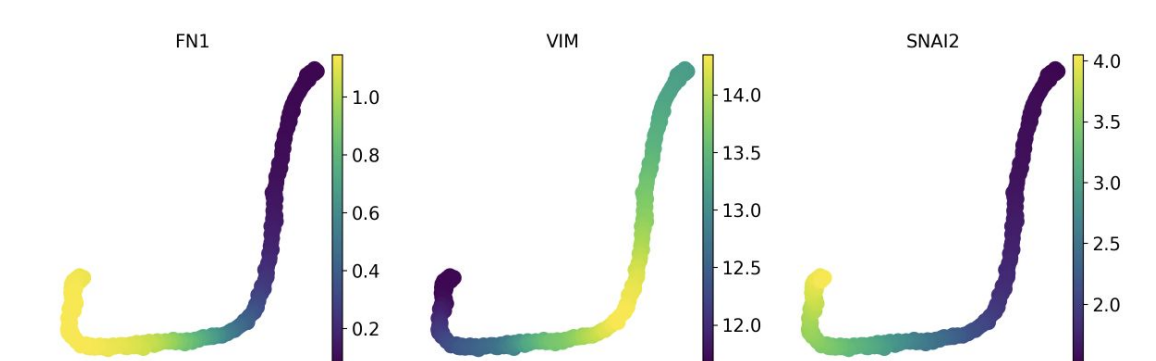
➤ Epithelial Markers:



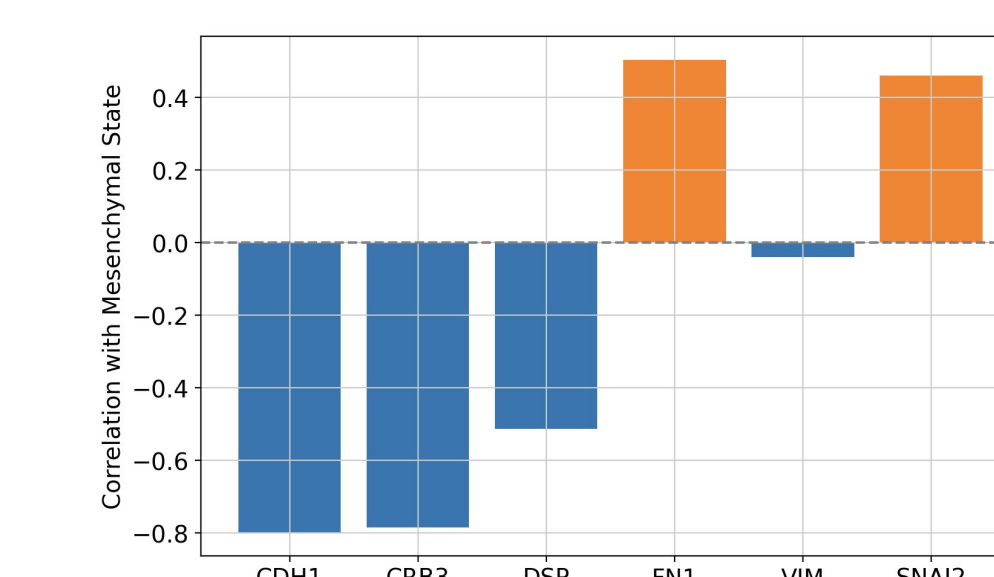
Gene expression trends goes down, which is expected during EMT.

➤ Mesenchymal Markers:

- FN1 and SNAI2 confirms expected late EMT upregulation.
- VIM shows non-monotonic expression: suggests that EMT is not purely linear

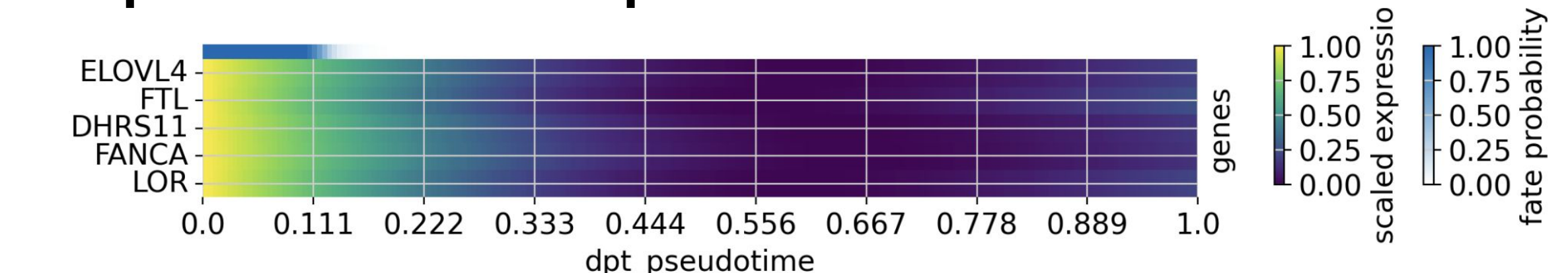


➤ Gene Expression Correlation with Mesenchymal State

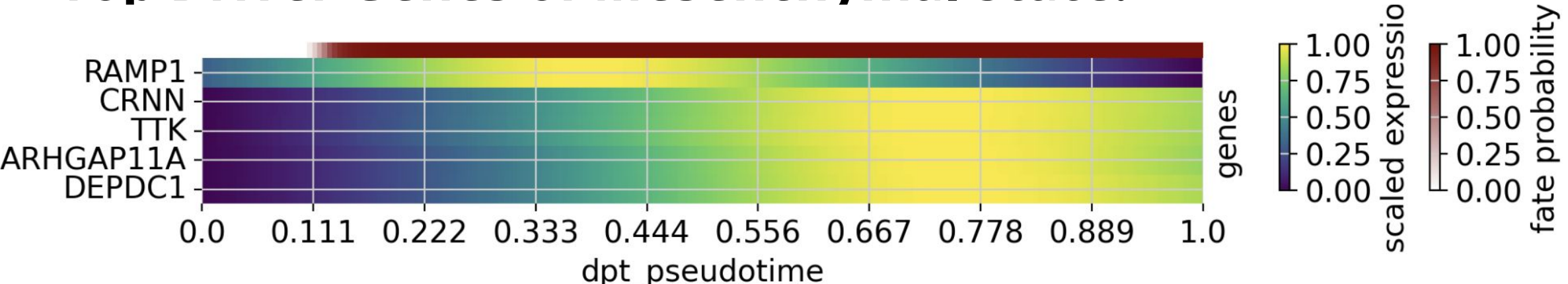


- CDH1, CRB3, DSP show strong negative correlations: consistent with EMT progression
- FN1 and SNAI2 exhibit positive correlations: consistent with EMT progression
- VIM shows a weak correlation close to zero: possibly peaking mid-trajectory

➤ Top Driver Genes of Epithelial State:



➤ Top Driver Genes of Mesenchymal State:



References

1. McFaline-Figueroa, José L., et al. "A pooled single-cell genetic screen identifies regulatory checkpoints in the continuum of the epithelial-to-mesenchymal transition." *Nature genetics* 51.9 (2019): 1389-1398.
2. Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes." *Dec. 2013*,
3. Lange, Marius, et al. "CellRank for directed single-cell fate mapping." *Nature methods* 19.2 (2022): 159-170.