



# Computational denoising facilitates the exploratory analysis of single-cell RNA-Seq data from complex tissues

Florian Wagner, PhD<sup>1</sup> & Itai Yanai, PhD<sup>2</sup>

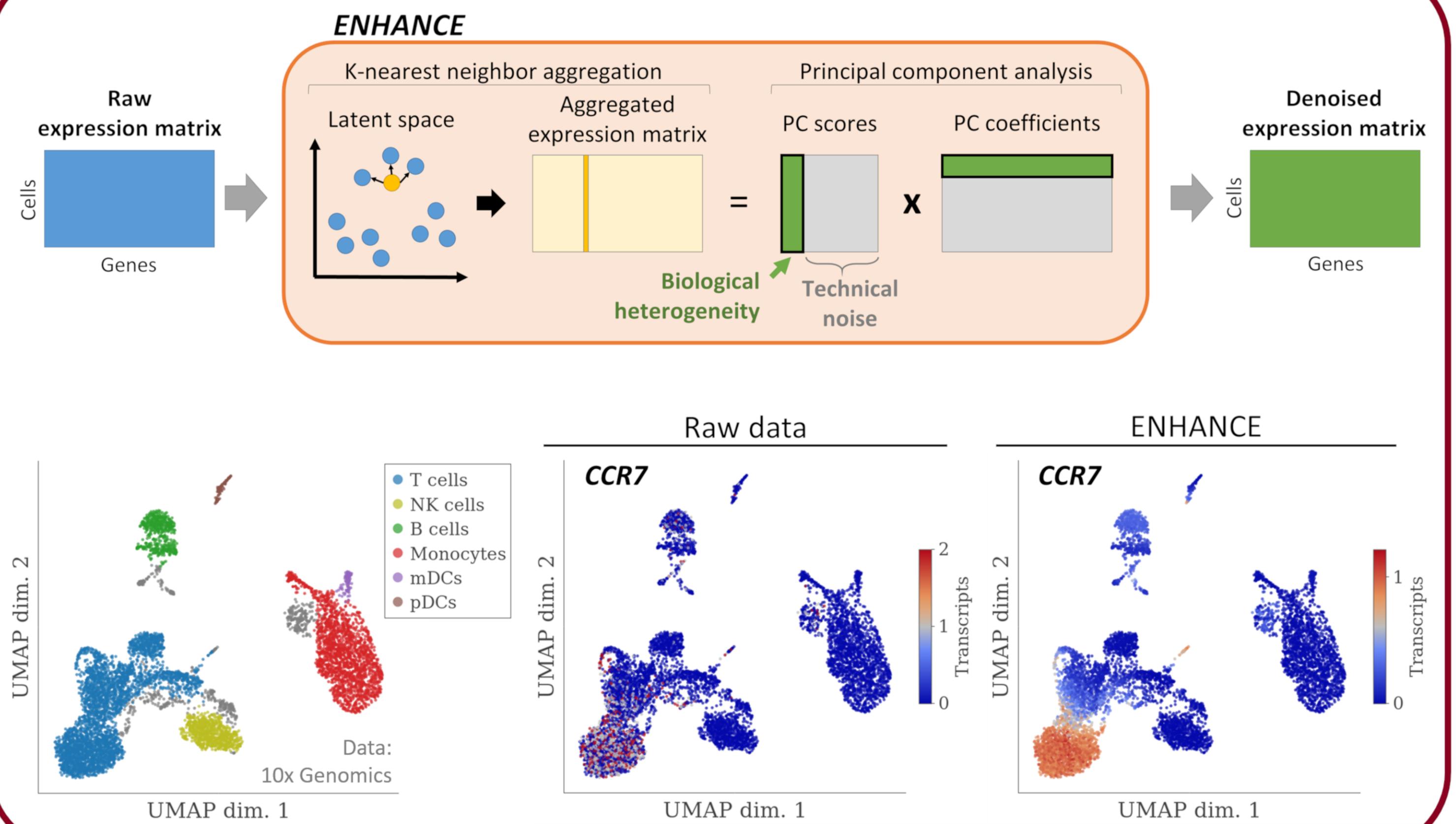
<sup>1</sup>Basu Lab, Section of Genetic Medicine & Pott Lab, Department of Human Genetics, University of Chicago

<sup>2</sup>Institute for Computational Medicine, NYU School of Medicine

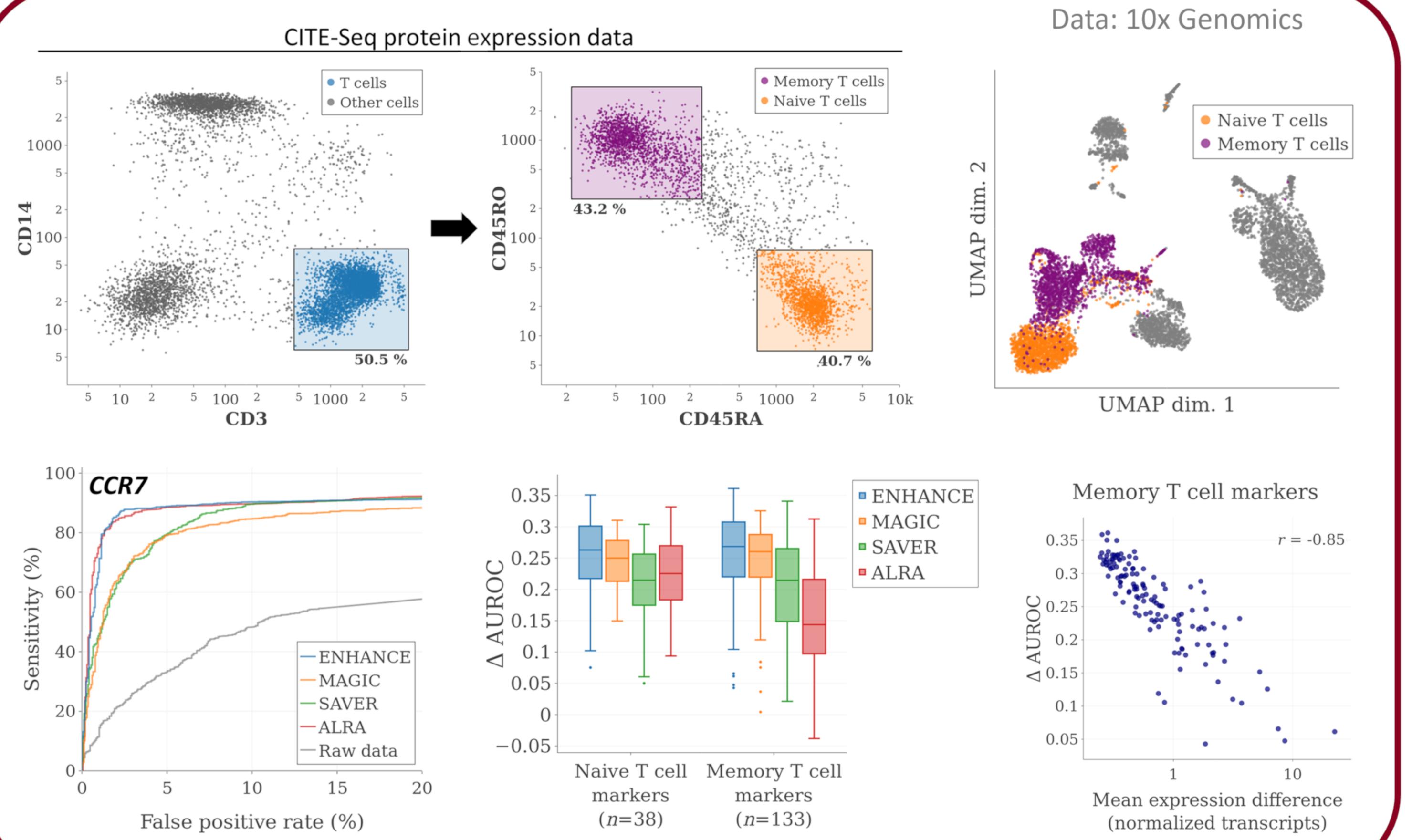
## Summary

Single-cell RNA-Seq data suffer from inherently high levels of technical noise, which poses challenges for exploratory data analysis. To study the expression patterns of individual genes, it is often necessary to first perform clustering, which then allows a comparison of average expression levels between clusters. However, identifying the cluster structure that best reflects the underlying biology is often non-trivial, and a clustering approach furthermore assumes discrete cell phenotypes. We propose *ENHANCE*, a generally applicable denoising algorithm that facilitates the exploratory analysis of single-cell data without clustering. While the reliable benchmarking of different denoising algorithms remains challenging, recent research suggests that it is possible to obtain an unbiased measure of denoising accuracy using a purely computational approach.

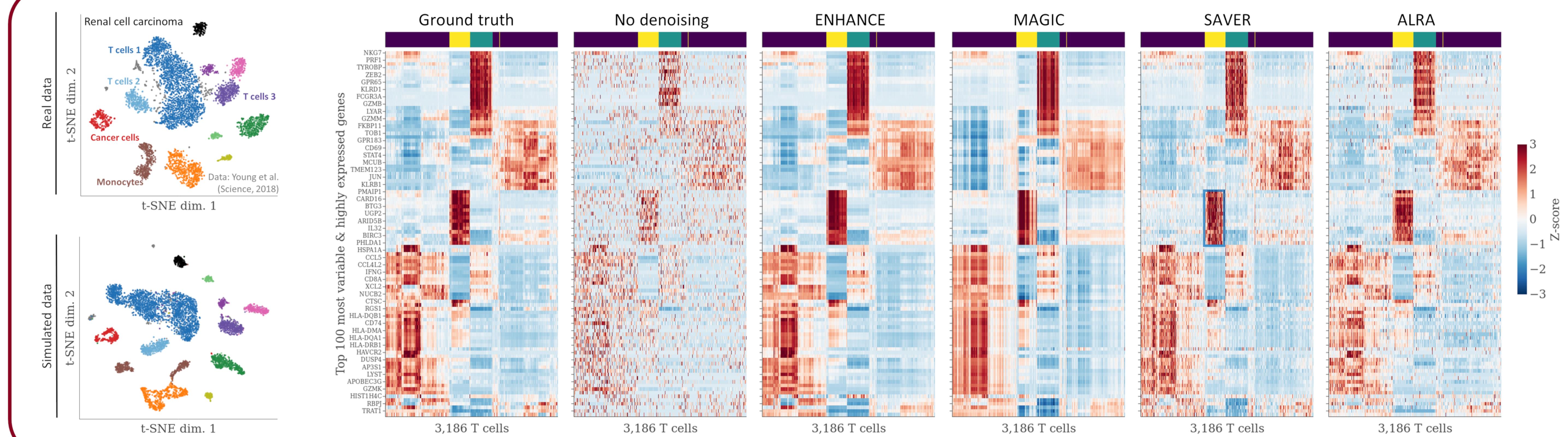
## Approach



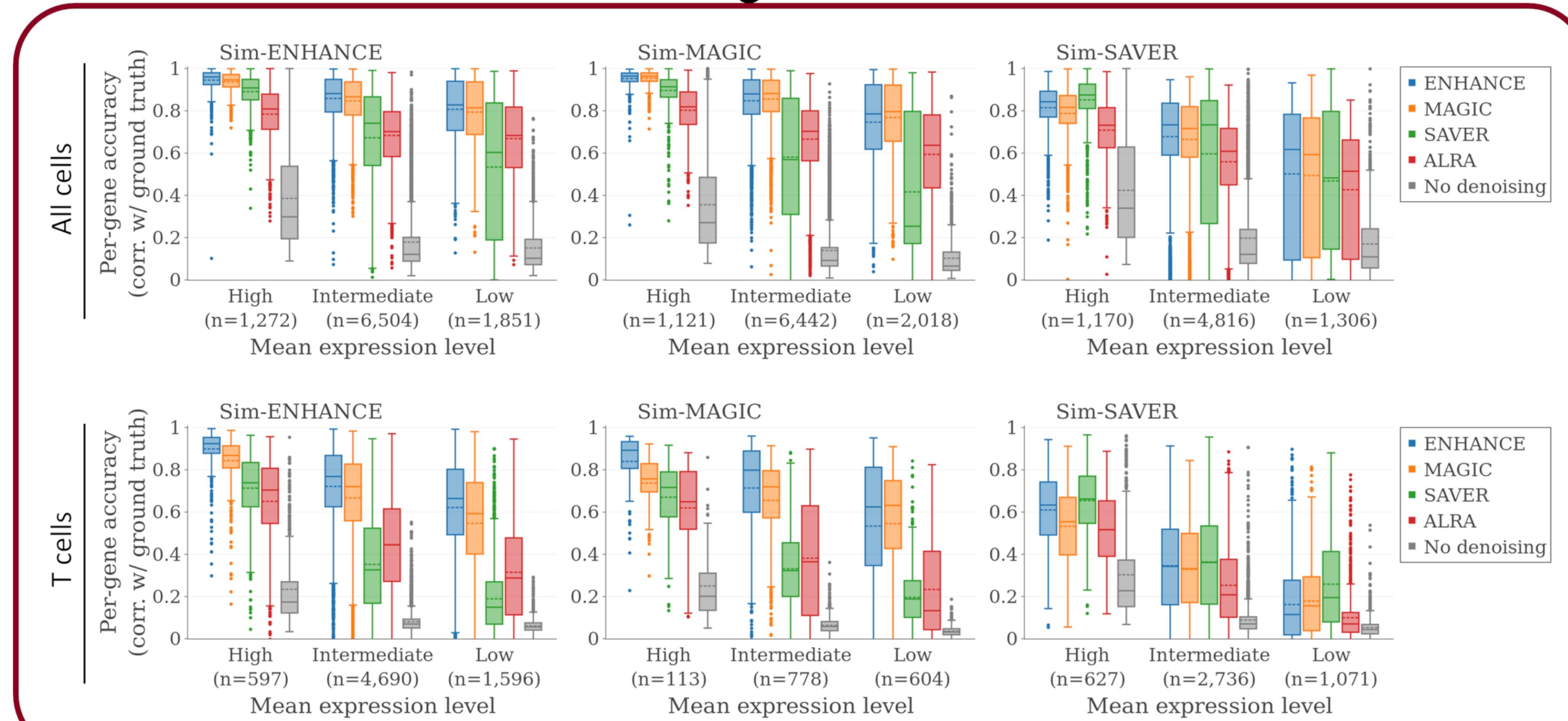
## Experimental validation using CITE-Seq data



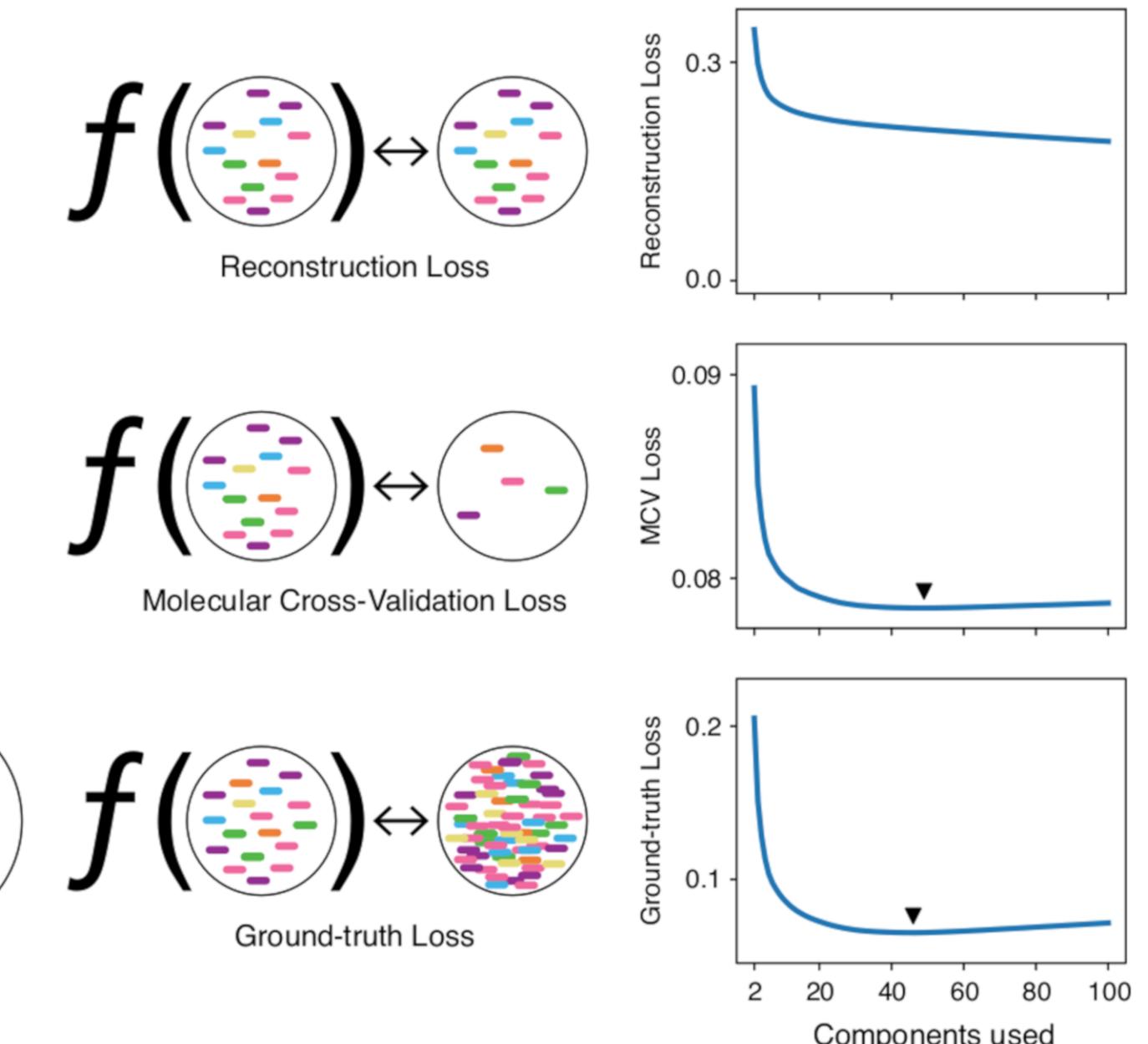
## Visual comparison of denoising methods using simulated renal cell carcinoma data



## Simulation-based benchmarking results

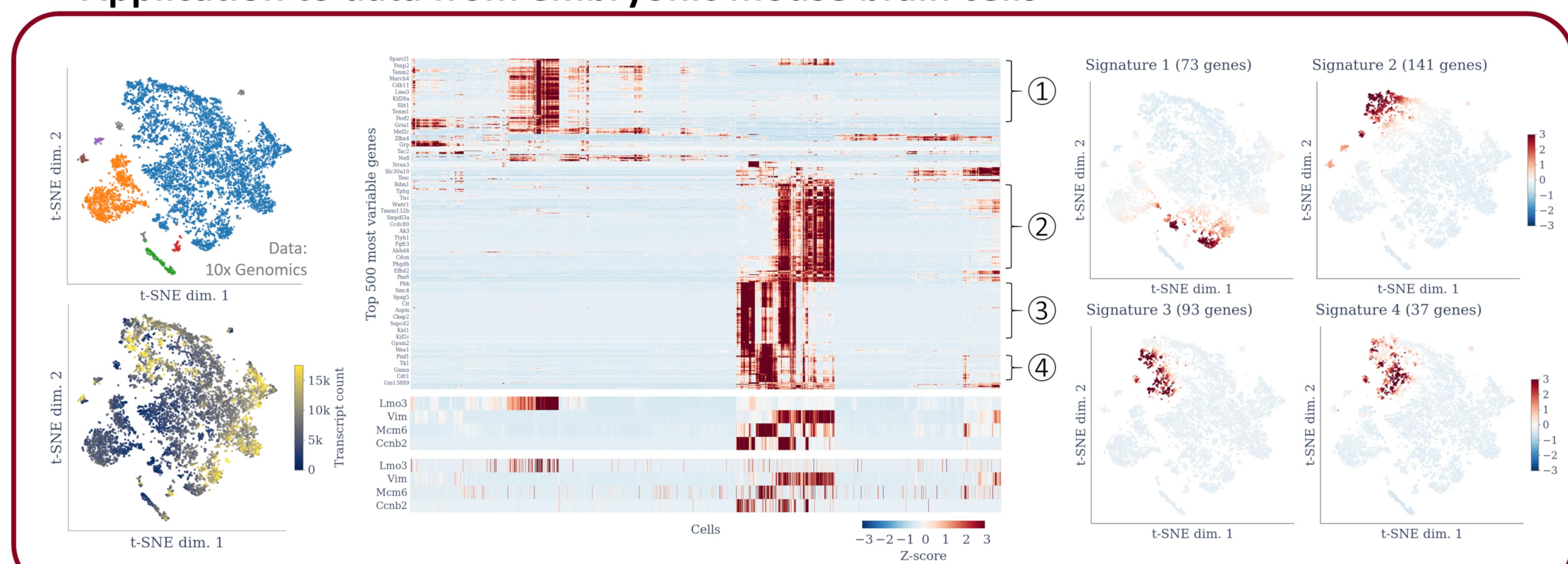


Goal: Simulation-free benchmarking using molecular cross-validation (Batson et al., 2019)



Adapted from Batson et al. (bioRxiv, 2019; CC-BY 4.0)

## Application to data from embryonic mouse brain cells



## References

- Wagner, Florian, Dalia Barkley, and Itai Yanai. "Accurate Denoising of Single-Cell RNA-Seq Data Using Unbiased Principal Component Analysis." *BioRxiv*, June 17, 2019, 655365.
- Batson, Joshua, Loïc Royer, and James Webber. "Molecular Cross-Validation for Single-Cell RNA-Seq." *BioRxiv*, September 30, 2019, 786269.
- Python/R code on GitHub: <https://github.com/yanailab/enhance>

Email: [florian.wagner@uchicago.edu](mailto:florian.wagner@uchicago.edu)  
Twitter: [@flo\\_compbio](https://twitter.com/flo_compbio)