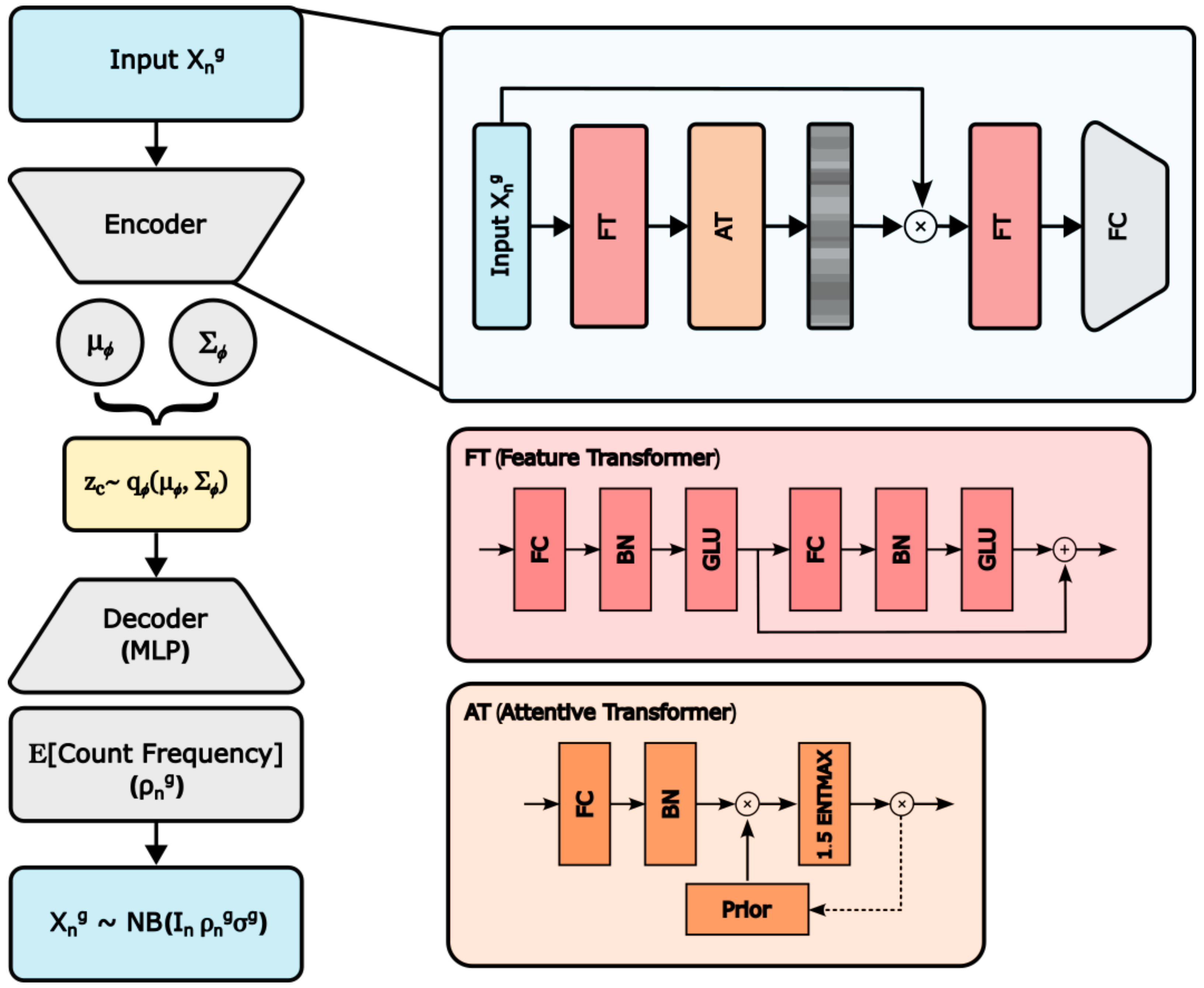
ALLEN INSTITUTE for BRAIN SCIENCE

Leveraging Transformer Architectures to Learn Biologically Meaningful Cellular Representations

Aditi Chandrashekar[1], Rohan Gala[2], and Mariano Gabitto[2]
1. California Institute of Technology, Pasadena, CA, 2. Allen Institute for Brain Sciences, Seattle, WA

Introduction

- **We propose TabVI**: A probabilistic transformer model for **single-cell genomics**.
- Adapts transformer architectures to gene functionality, which is hierarchical and modular, unlike the sequential structure of natural language.
- Improves latent embedding learning. Validated on cell type annotation and integration benchmarks.
- Robust to dataset scaling using interpretable, sample-specific feature attention.
- Excels where large-scale foundation models are less effective in single-cell analysis.



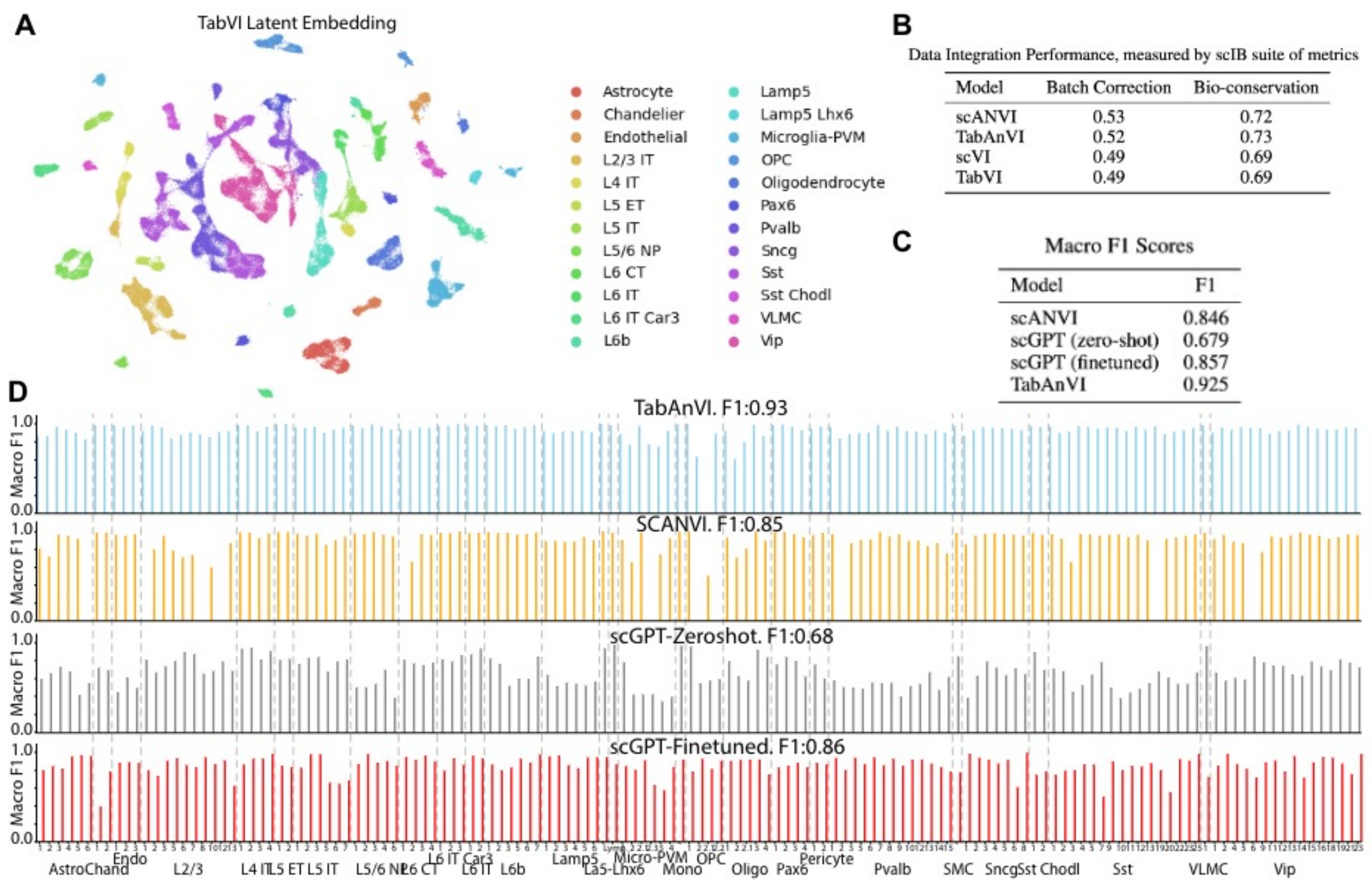
Architecture combines the latent space interpretability of VAEs with a sample efficient tabular feature transformer.

References

[1] Cui H, Wang C, Maan H, Pang K, Luo F, Duan N, and Wang B. scgpt: toward building a foundation model for single-cell multi-omics using generative ai. Nature Methods, 21(8):1470–1480, 2024.
[2] Diederik P Kingma and Max Welling. Auto-encoding variational bayes, 2022.
[3] Romain Lopez, Jeffrey Regier, Michael B. Cole, Michael I. Jordan, and Nir Yosef. Deep generative modeling for single-cell transcriptomics. Nature Methods, 15(12):1053–1058, November 2018.
[4] Sercan O. Arik and Tomas Pfister. Tabnet: Attentive interpretable tabular learning, 2020.
[5] Chenling Xu, Romain Lopez, Edouard Mehlman, Jeffrey Regier, Michael I. Jordan, and Nir Yosef. Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. January 2019.
[6] Felix Fischer, David S Fischer, Roman Mukhin, Andrey Isaev, Evan Biederstedt, Alexandra-Chloé Villani, and Fabian J Theis. sctab: Scaling cross-tissue single-cell annotation models. Nature Communications, 15(1):6611, 2024.

Results

Benchmarking across integration and annotation tasks.



B

Data Integration Performance, measured by scIB suite of metrics

Model	Batch Correction	Bio-conservation
scANVI	0.53	0.72
TabAnVI	0.52	0.73
scVI	0.49	0.69
TabVI	0.49	0.69

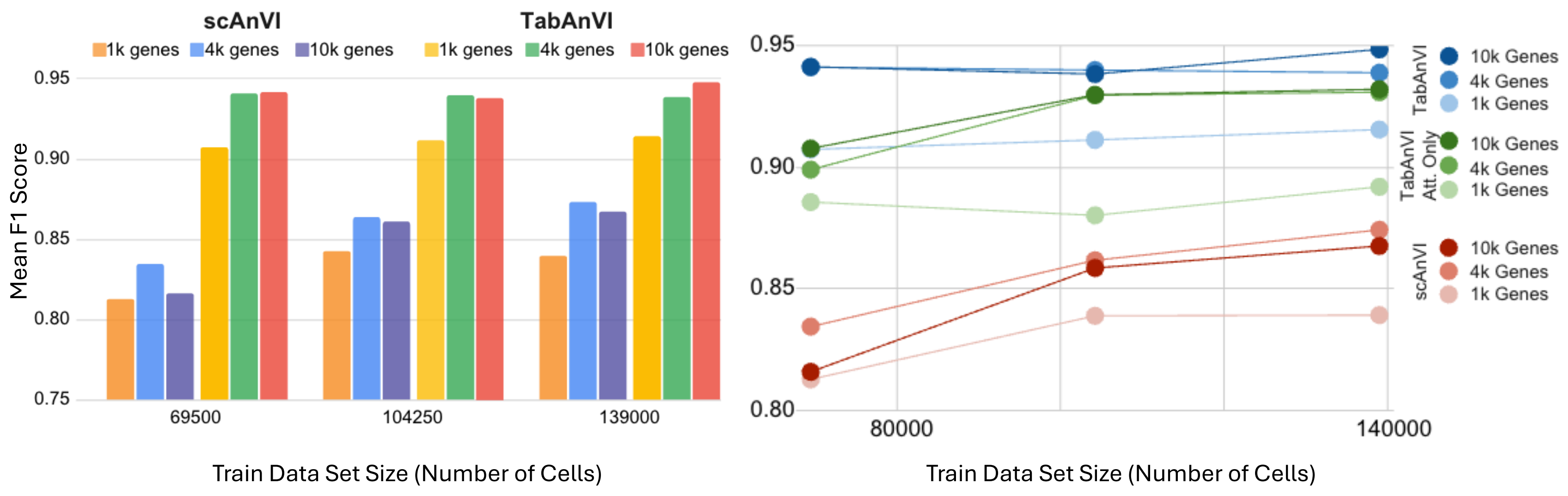
C

Macro F1 Scores

Model	F1
scANVI	0.846
scGPT (zero-shot)	0.679
scGPT (finetuned)	0.857
TabAnVI	0.925

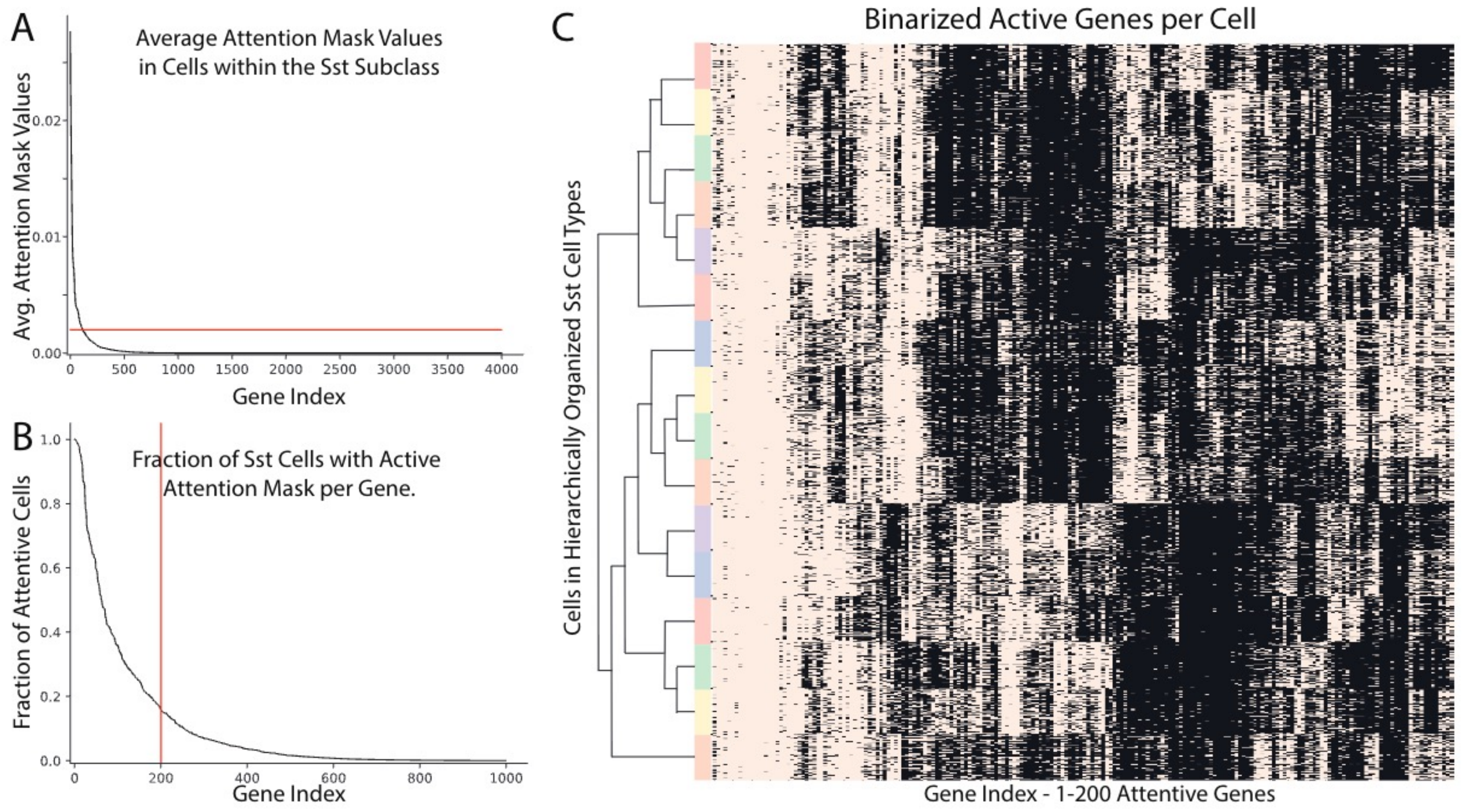
- TabVI's latent representations of human middle temporal gyrus cells originating from SEA-AD donors spanning the entire spectrum of AD (A).
- TabVI achieves higher overall classification performance compared to scAnVI [5] or scGPT [1] representations (B), while maintaining scVI's [3] representation capabilities (C).
- TabVI achieves high classification performance across cell types (D).

Scalability performance evaluated on subsets of features and observations.

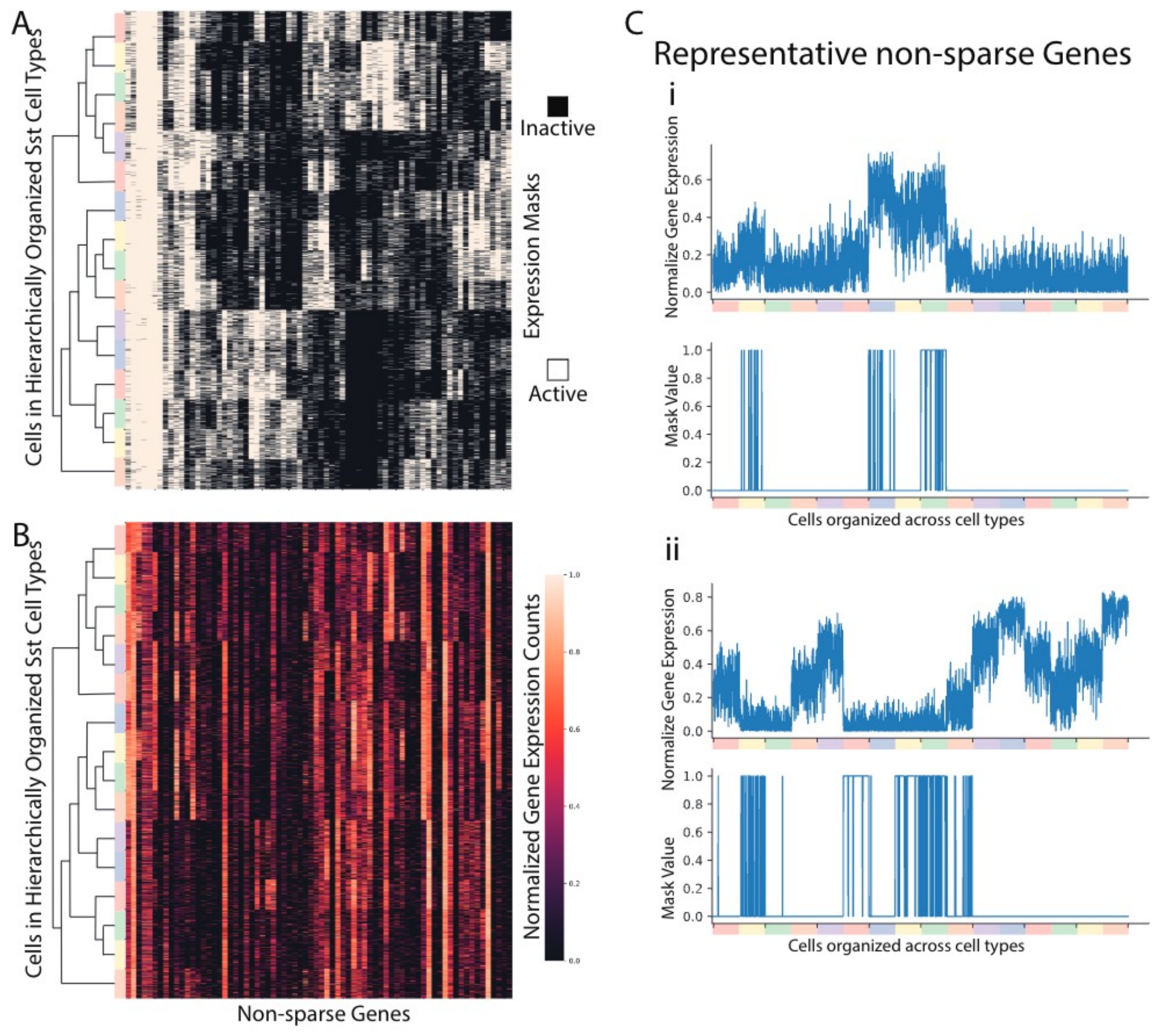


- We compare macro F1 annotation metrics from TabVI against scAnVI [5] when training data consists of fewer genes, fewer cells, or both.
- TabVI performance consistently improves when more features are introduced in the input data. *(left)*
- Each component of TabVI is necessary to produce peak performance (we show scAnVI [5], a model with attentive masking only, and TabVI, the model with both attentive masking and a tabular transformer)

Cell Type-specific Feature Selection



A. Threshold for "attended to" genes.
B. Fraction of Sst cells in which each gene is attended to (y-axis).
C. Heatmap of binarized attention mask values, arranged hierarchically by genes and stratified by cell types. Color denotes cell types within the Sst subclass.



A. Attention mask of non-sparsely expressed genes.
B. Normalized expression values corresponding to non-sparsely expressed genes. Cells and genes organized as in **(A)**.
C. Expression and mask values for representative, non-sparse genes.

- **The attention mechanism captures cell type-specific and combinatorial attention patterns.**
- These patterns are amplified when we select against sparsity.

Acknowledgements

I would like to thank my mentor, Dr. Mariano Gabitto for his invaluable support and mentorship this summer, Dr. Rohan Gala for his guidance, and the Allen Institute for Brain Science for hosting me.
Thank you to Arthur Rock for generously supporting my SURF fellowship.