**Explainable variational autoencoders for automatic annotation of hematopoietic stem and progenitor cells from scRNA-seq**

Maja Błażejewicz, Jagiellonian University

*maja.blazejewicz*@*student.uj.edu.pl***Keywords**: scRNA-seq, hematopoietic stem cells, deep learning, variational autoencoder, explainable AI, cell type annotation

1. **Introduction**

Single-cell RNA sequencing (scRNA-seq) has become a transformative tool for understanding cellular heterogeneity, yet accurate cell type annotation remains a major challenge [1]. This is particularly evident in the hematopoietic system, where hematopoietic stem cells (HSCs) and progenitors represent a transcriptional continuum rather than discrete, easily separable populations. Traditional approaches have relied on surface markers such as CD34, which remain fundamental for defining HSCs. However, their expression is not strictly specific or stable, being influenced by environmental cues, cell cycle status, or metabolic changes.

1. **Description of the problem**

The limitations of classical marker-based annotation highlight the need for more robust computational strategies. scRNA-seq datasets capture high-dimensional gene expression profiles, offering opportunities to uncover subtle differences between closely related cell states. In hematopoiesis, precise discrimination between long-term repopulating HSCs, short-term progenitors, and lineage-committed cells is essential for understanding differentiation trajectories, aging, and hematological disorders [2]. Misclassification can distort biological conclusions, for example, by underestimating the frequency of highly regenerative cells or overestimating progenitor populations. The ability to identify subtypes with high resolution is therefore not only fundamental for basic research, but also has implications for transplantation, gene therapy, and clinical diagnostics. Additionally, it is crucial to understand how variability in gene expression translates into functional differences between closely related cell types and states, as this variability often underlies the trajectories of hematopoietic differentiation and disease progression.

1. **Related work**

Advances in machine learning, and deep learning specifically, have opened new possibilities for cell type classification in scRNA-seq. Methods such as random forest, gradient boosting, neural networks, and autoencoders can integrate information across thousands of genes, capturing relationships invisible to marker-based strategies. Among them, variational autoencoders (VAEs) provide a particularly powerful framework for modeling complex and continuous transcriptomic landscapes. Previous studies have demonstrated that autoencoder-based approaches can effectively reconstruct cellular trajectories and detect subtle transcriptional programs, extending beyond the resolution of classical clustering methods [3].

1. **Solution to the problem**

The use of VAEs combined with an explainable AI (XAI) framework enables both accurate annotation and biological interpretability. By attributing gene-level importance scores to specific latent dimensions, XAI provides insights into the molecular features that drive differences between HSC subpopulations and along differentiation pathways. This interpretability bridges the gap between computational predictions and biological mechanisms, helping to identify transcription factors, regulatory programs, or stress-response genes that shape hematopoietic fate. Applied to scRNA-seq data, such models allow the mapping of trajectories with improved resolution while highlighting the gene expression signatures that underpin lineage priming and functional diversity within HSCs [4].

1. **Conclusions and future work**

Integrating VAEs with explainability methods offers a powerful approach to refine cell type annotation in scRNA-seq, particularly for complex and heterogeneous systems such as hematopoiesis. Beyond automating and improving classification, this strategy provides mechanistic insights into gene programs defining stemness, differentiation, and disease states.

Future work will focus on validating the results across independent clinical datasets, extending the models to incorporate multimodal measurements, and deepening the biological interpretation of explainability outputs to better link computational findings with hematopoietic biology and disease mechanisms, further enhancing their relevance for translational hematology.

**References**

1. Pei G, Yan F, Simon LM, Dai Y, Jia P, Zhao Z. *deCS: A Tool for Systematic Cell Type Annotations of Single-cell RNA Sequencing Data among Human Tissues*. Genomics Proteomics Bioinformatics. 2023 Apr;21(2):370-384. <https://doi.org/10.1016/j.gpb.2022.04.001>
2. Skulimowska, I., Sosniak, J., Gonka, M., Szade, A., Jozkowicz, A. and Szade, K. (2022), *The biology of hematopoietic stem cells and its clinical implications*. FEBS J, 289: 7740-7759. <https://doi.org/10.1111/febs.16192>
3. Tran, D., Nguyen, H., Tran, B. et al. *Fast and precise single-cell data analysis using a hierarchical autoencoder*. Nat Commun 12, 1029 (2021). <https://doi.org/10.1038/s41467-021-21312-2>
4. Sun SJ, Aguirre-Gamboa R, de Bree LCJ, Sanz J, Dumaine A, van der Velden WJFM, Joosten LAB, Khader S, Divangahi M, Netea MG, Barreiro LB., *BCG vaccination alters the epigenetic landscape of progenitor cells in human bone marrow to influence innate immune responses*. Immunity. 2024 Sep 10;57(9):2095-2107.e8. <https://doi.org/10.1016/j.immuni.2024.07.021>