Additional file 1: Supplementary Figures for iVAE: An Interpretable Representation Learning Framework Enhances Clustering Performance for Single-Cell Data

Zeyu Fu[1*†], Chunlin Chen[2†], Song Wang[1], Junping Wang[1*], Shilei Chen[1*]

[1*]State Key Laboratory of Trauma and Chemical Poisoning, Institute of Combined Injury, Chongqing Engineering Research Center for Nanomedicine, College of Preventive Medicine, Army Medical University, Chongqing, 400038, China.
[2]Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, 510080, China.

*Corresponding author(s). E-mail(s): fuzeyu99@126.com; wangjunping@tmmu.edu.cn; chen.shilei@foxmail.com;
Contributing authors: chenchlin3@mail2.sysu.edu.cn; swang1981@tmmu.edu.cn;
†These authors contributed equally to this work.

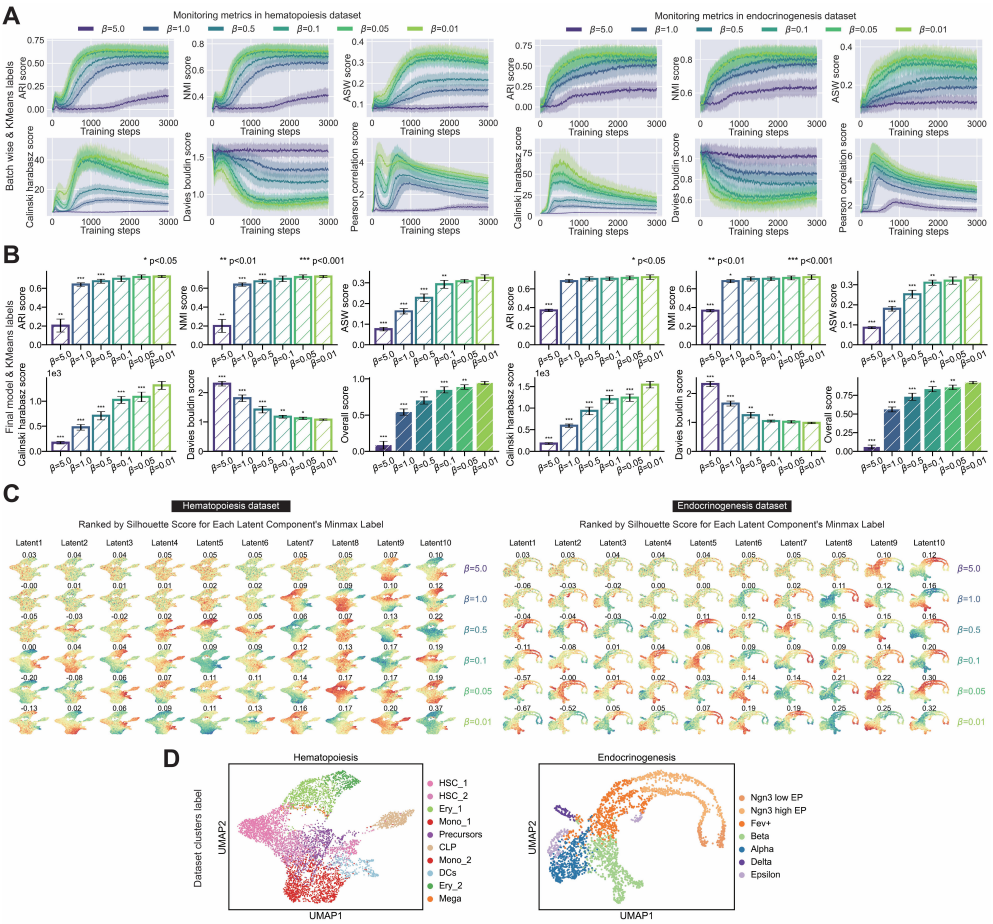


Fig. S1 Comparative evaluation of β-VAE across different β values (5.0, 1.0, 0.5, 0.1, 0.05, 0.01) on Hematopoiesis and Endocrinogenesis datasets with 10 technical replicates. **(A)** Performance monitoring across β values using 5 unsupervised clustering metrics: Adjusted Rand Index (ARI), Normalized Mutual Information (NMI), Average Silhouette Width (ASW), Calinski-Harabasz Index (CH), and Davies-Bouldin Index (DB), along with summed absolute Pearson correlation coefficients between latent components (lower DB values indicate better performance). **(B)** Final model evaluation comparing performance across β values using the same 5 unsupervised metrics and composite overall scores. Statistical significance was assessed using Bonferroni correction (5 comparisons) with paired t-tests for normally distributed metrics or Wilcoxon signed-rank tests for metrics violating normality assumptions, as determined by Shapiro-Wilk normality testing. **(C)** Intensity distribution visualization of latent components from the final model encoding projected onto a two-dimensional UMAP embedding space. **(D)** Cell type distribution patterns visualized in the corresponding two-dimensional embedding space.

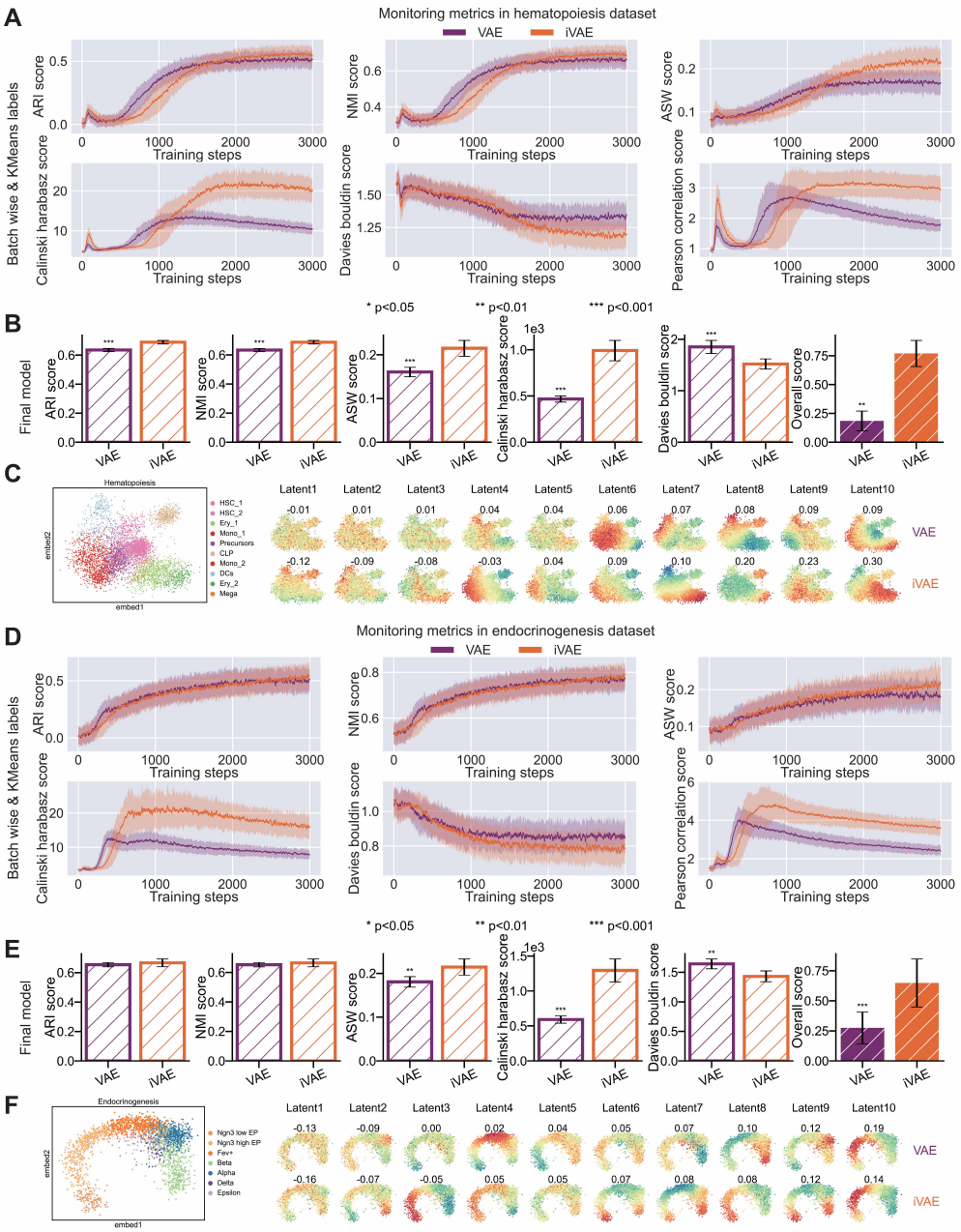


Fig. S2 Comparative evaluation of improved iVAE versus standard VAE on Hematopoiesis and Endocrinogenesis datasets across 10 technical replicates. **(A, D)** Performance monitoring comparison using 5 unsupervised clustering metrics: Adjusted Rand Index (ARI), Normalized Mutual Information (NMI), Average Silhouette Width (ASW), Calinski-Harabasz Index (CH), and Davies-Bouldin Index (DB), along with summed absolute Pearson correlation coefficients between latent components for Hematopoiesis (A) and Endocrinogenesis (D) datasets respectively (lower DB values indicate better performance). **(B, E)** Final model evaluation comparing iVAE and VAE performance using the same 5 unsupervised metrics and composite overall scores for Hematopoiesis (B) and Endocrinogenesis (E) datasets. Statistical significance was assessed using paired t-tests for normally distributed metrics or Wilcoxon signed-rank tests for metrics violating normality assumptions, as determined by Shapiro-Wilk normality testing. **(C, F)** Cell type distribution patterns and intensity visualization of latent components encoded by iVAE projected onto two-dimensional UMAP embedding spaces for Hematopoiesis (C) and Endocrinogenesis (F) datasets.

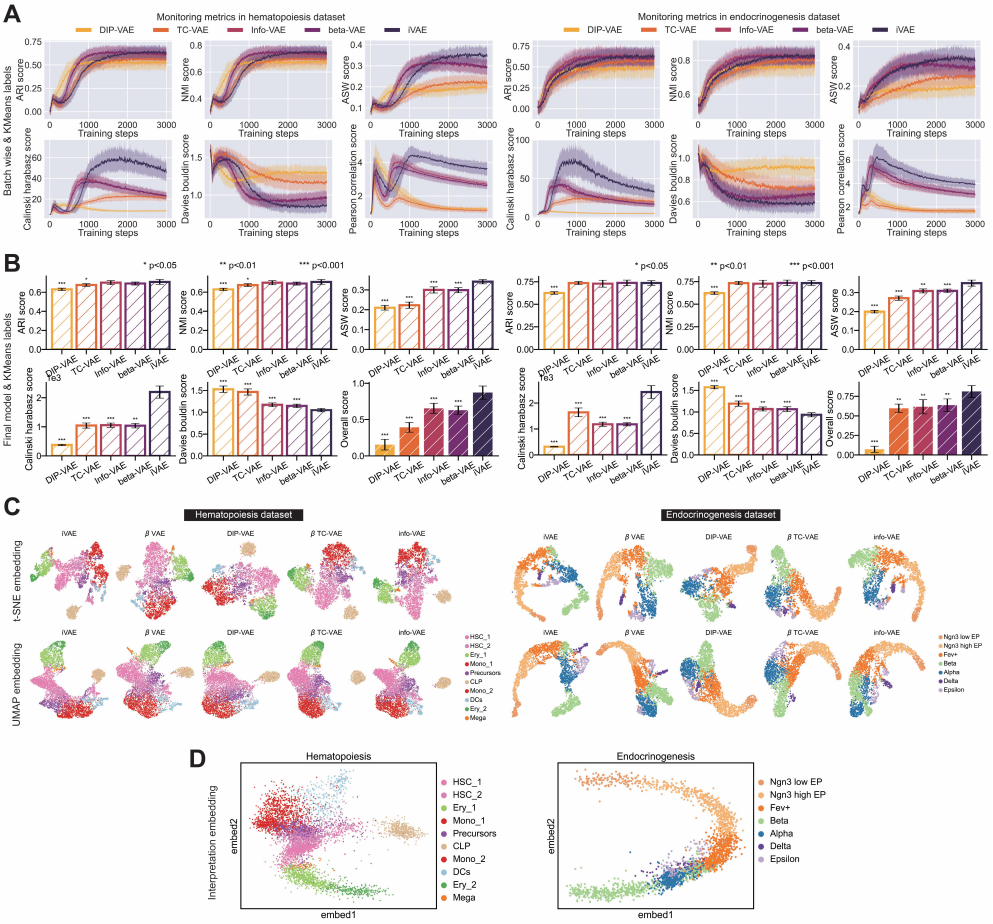


Fig. S3 Comparative evaluation of iVAE against VAE variants on Hematopoiesis and Endocrinogenesis datasets across 10 technical replicates. **(A)** Performance monitoring comparison of 5 methodologies (DIP-VAE, TC-VAE, info-VAE, β-VAE, and iVAE) across both datasets using 5 unsupervised clustering metrics: Adjusted Rand Index (ARI), Normalized Mutual Information (NMI), Average Silhouette Width (ASW), Calinski-Harabasz Index (CH), and Davies-Bouldin Index (DB), along with summed absolute Pearson correlation coefficients between latent components (lower DB values indicate better performance). **(B)** Final model evaluation comparing the 5 methodologies across both datasets using the same 5 unsupervised metrics and composite overall scores. Statistical significance was assessed using Bonferroni correction (4 comparisons) with paired t-tests for normally distributed metrics or Wilcoxon signed-rank tests for metrics violating normality assumptions, as determined by Shapiro-Wilk normality testing. **(C)** Intensity visualization of latent components from final model encodings projected onto two-dimensional UMAP embedding spaces for all 5 methodologies across both datasets, demonstrating learned representation patterns. **(D)** Cell type distribution patterns visualized in the corresponding two-dimensional embedding spaces.

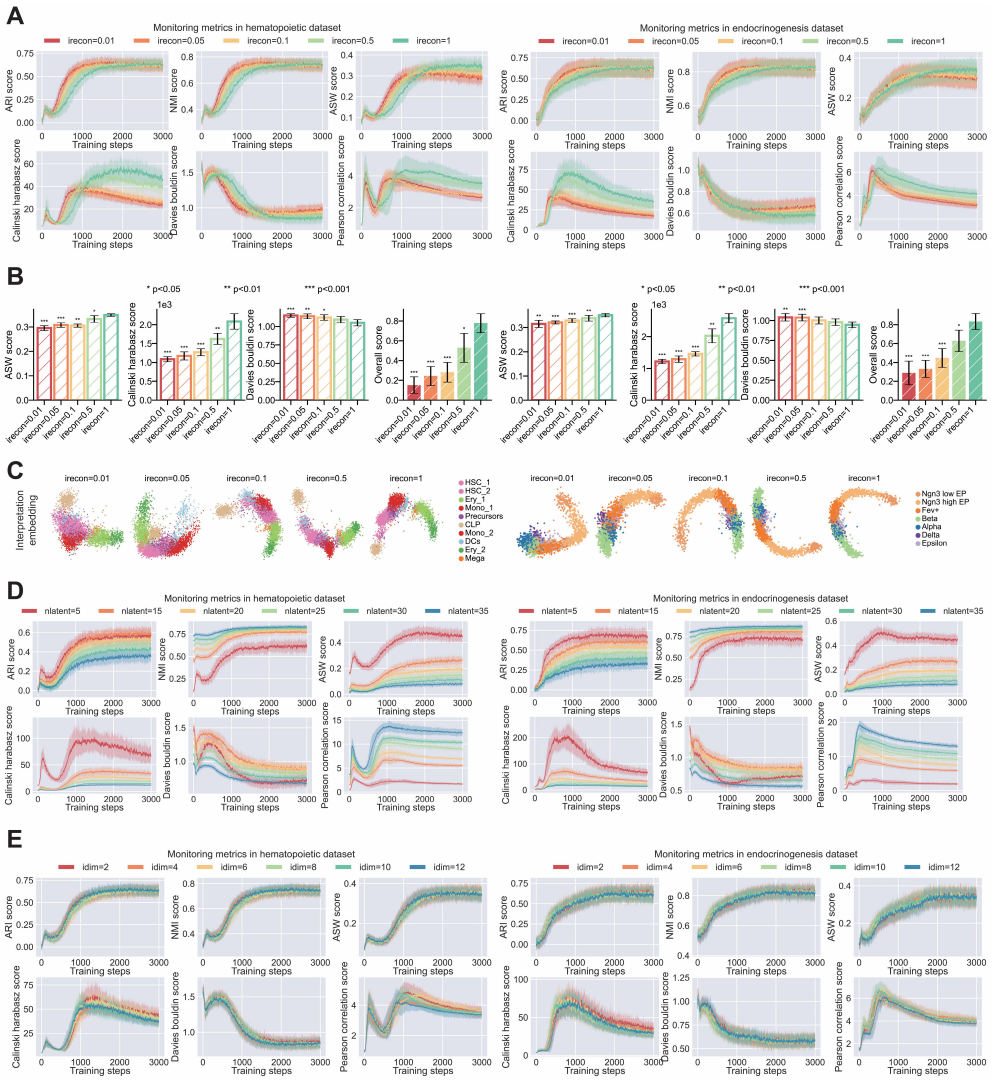


Fig. S4 Analysis of iVAE performance across hyperparameter gradients on Hematopoiesis and Endocrinogenesis datasets. **(A)** Performance monitoring of iVAE across a sequence of irecon regularization values (10 technical replicates) using 5 unsupervised clustering metrics: Adjusted Rand Index (ARI), Normalized Mutual Information (NMI), Average Silhouette Width (ASW), Calinski-Harabasz Index (CH), and Davies-Bouldin Index (DB), along with summed absolute Pearson correlation coefficients between latent components (lower DB values indicate better performance). **(B)** Final model evaluation of iVAE across the same sequence of irecon regularization values (10 technical replicates) using the 5 unsupervised metrics and composite overall scores. Statistical significance was assessed using Bonferroni correction (4 comparisons) with paired t-tests for normally distributed metrics or Wilcoxon signed-rank tests for metrics violating normality assumptions, as determined by Shapiro-Wilk normality testing. **(C)** Two-dimensional UMAP interpretation embeddings visualizing the effect of varying n_{latent} (latent space dimensionality) values on learned representations. **(D)** Performance monitoring of iVAE across a sequence of irecon regularization values with 5 technical replicates. **(E)** Performance monitoring of iVAE across a sequence of i_{dim} (interpretation dimension) values with 5 technical replicates.

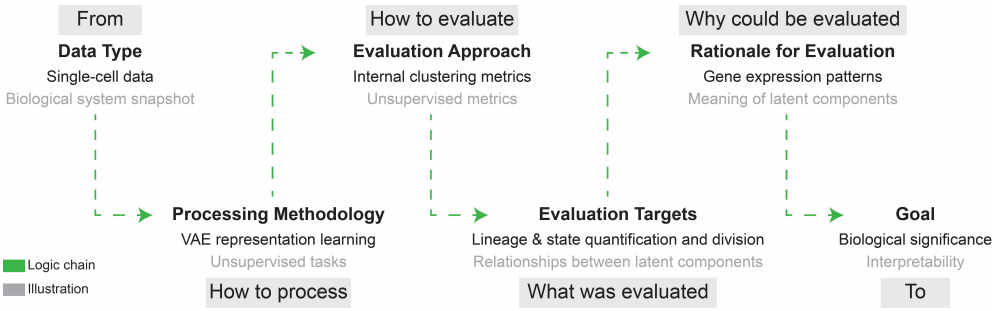


Fig. S5 Overall logic flow of this study. Schematic overview illustrating the logical framework underlying the evaluation of iVAE. The flow begins with the unique characteristics of single-cell transcriptomic data and proceeds through VAE-based representation learning as the processing methodology. Internal clustering metrics are employed as significant evaluation approaches to quantify the structure of the latent space in an unsupervised manner. Evaluation targets focus on the delineation of cell lineages and states based on relationships among latent components. The rationale for such evaluation is rooted in gene expression patterns, which underpin the biological meaning interpretable from the latent space. Ultimately, this logic chain aims to ensure that the learned representations promote biological significance and interpretability in downstream analyses.

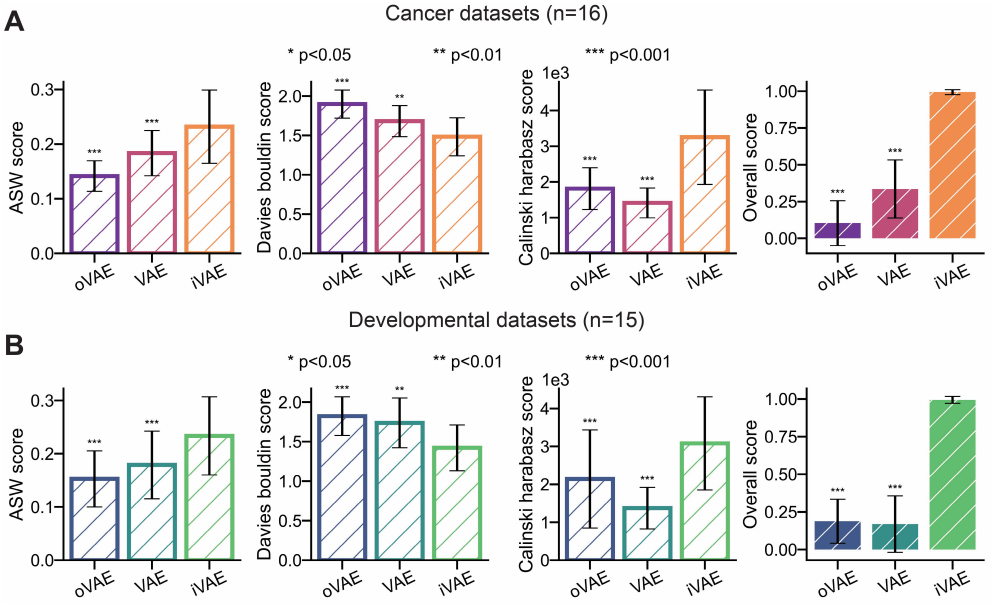


Fig. S6 Structural evaluation of iVAE. The oVAE denotes the iVAE architecture without the forward component. **(A)** Performance comparison on 16 cancer-derived single-cell datasets, with statistical significance indicated by asterisks (*). Statistical significance was assessed using Bonferroni correction (2 comparisons) with paired t-tests for normally distributed metrics or Wilcoxon signed-rank tests for metrics violating normality assumptions, as determined by Shapiro-Wilk normality testing. **(B)** Performance comparison on 15 developmental single-cell datasets.