



PDF Download
3627673.3679576.pdf
27 February 2026
Total Citations: 2
Total Downloads: 680

 Latest updates: <https://dl.acm.org/doi/10.1145/3627673.3679576>

RESEARCH-ARTICLE

scACT: Accurate Cross-modality Translation via Cycle-consistent Training from Unpaired Single-cell Data

SIWEI XU, University of California, Irvine, Irvine, CA, United States

JUNHAO LIU, University of California, Irvine, Irvine, CA, United States

JING ZHANG, University of California, Irvine, Irvine, CA, United States

Open Access Support provided by:

University of California, Irvine

Published: 21 October 2024

Citation in BibTeX format

CIKM '24: The 33rd ACM International Conference on Information and Knowledge Management
October 21 - 25, 2024
ID, Boise, USA

Conference Sponsors:
SIGIR

scACT: Accurate Cross-modality Translation via Cycle-consistent Training from Unpaired Single-cell Data

Siwei Xu*
s.xu@uci.edu
University of California, Irvine
Irvine, California, USA

Junhao Liu*
junhao.liu@uci.edu
University of California, Irvine
Irvine, California, USA

Jing Zhang†
zhang.jing@uci.edu
University of California, Irvine
Irvine, California, USA

Abstract

Single-cell sequencing technologies have revolutionized genomics by enabling the simultaneous profiling of various molecular modalities within individual cells. Their integration, especially cross-modality translation, offers deep insights into cellular regulatory mechanisms. Many methods have been developed for cross-modality translation, but their reliance on scarce high-quality co-assay data limits their applicability. Addressing this, we introduce scACT, a deep generative model designed to extract cross-modality biological insights from unpaired single-cell data. scACT tackles three major challenges: aligning unpaired multi-modal data via adversarial training, facilitating cross-modality translation without prior knowledge via cycle-consistent training, and enabling interpretable regulatory interconnections explorations via in-silico perturbations. To test its performance, we applied scACT on diverse single-cell datasets and found it outperformed existing methods in all three tasks. Finally, we have developed scACT as an individual open-source software package to advance single-cell omics data processing and analysis within the research community.

CCS Concepts

• **Applied computing** → **Computational biology**; • **Computing methodologies** → **Knowledge representation and reasoning**.

Keywords

Multimodal Integration, Cross-modality Translation, Single-cell RNA-seq, Single-cell ATAC-seq, Model Interpretability, Cycle-consistent Adversarial Training

ACM Reference Format:

Siwei Xu, Junhao Liu, and Jing Zhang. 2024. scACT: Accurate Cross-modality Translation via Cycle-consistent Training from Unpaired Single-cell Data. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM '24)*, October 21–25, 2024, Boise, ID, USA. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3627673.3679576>

1 Introduction

The recent advancements in single-cell sequencing technologies have significantly transformed genomics, allowing for the detailed

*Both authors contributed equally to this work.

†Corresponding author.



This work is licensed under a Creative Commons Attribution International 4.0 License.

CIKM '24, October 21–25, 2024, Boise, ID, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0436-9/24/10
<https://doi.org/10.1145/3627673.3679576>

measurement of molecular information from individual cells [15, 18]. This development has been pivotal in shifting our understanding of biological complexity and diversity at the cellular level [4, 7, 9, 26]. Moreover, a versatile array of single-cell tools have been routinely used to separately explore diverse molecular aspects, such as genetics, epigenetics, transcriptomics, and proteomics, enabling a more nuanced understanding of cellular functions in disease and paving the way to more effective therapeutic strategies [3, 19, 31]. However, a notable limitation of standard single-cell technologies is their ability to measure only one type of molecular information from each cell [14, 20]. This constraint presents significant challenges in fully deciphering the intricate interactions among genomic, transcriptomic, epigenetic, and proteomic factors in a comprehensive, multi-dimensional manner.

To tackle this challenge, various computational methods have been introduced to integrate multi-modal single-cell data sequenced separately without shared barcodes. For instance, several methods, encompassing both linear and nonlinear approaches, aim to project the data into a unified latent space to facilitate easier visualization and comparison across different modalities [5, 8, 22]. Subsequently, newer methods have emerged to address issues like distribution shift and partial feature overlap during the integration process, thereby enhancing the integration's practical applicability and quality [32, 34]. Despite these advancements, a major limitation of these methods is their focus predominantly on the alignment of multiple modalities, rather than on actual cross-modality mapping. This limitation restricts a deeper understanding of the complex interactions between different molecular layers within individual cells.

In recent years, the development of single-cell multi-omics technologies has enabled the simultaneous acquisition of molecular data from multiple modalities within the same cell. This capability for paired measurements offers unparalleled opportunities to explore cellular heterogeneity through multi-dimensional information and characterizes the regulatory interplay among these modalities. Consequently, these technological advancements have spurred the creation of novel computational methods, such as Babel, which facilitate direct cross-modality translation, yielding fresh biological insights [32]. However, these methods rely heavily on joint sequencing technologies, which are not yet broadly accessible due to their more complex library preparation protocols and higher sequencing costs. As a result, fully leveraging the extensive existing single-cell atlases, which often contain data from separate modalities, remains challenging when it comes to unraveling the intricate interplay within individual cells.

Here, we propose scACT, a deep generative model to gain cross-modality biological insights from unpaired single-cell data for solve three key problems: (i) robust **alignment** of *unpaired* single-cell

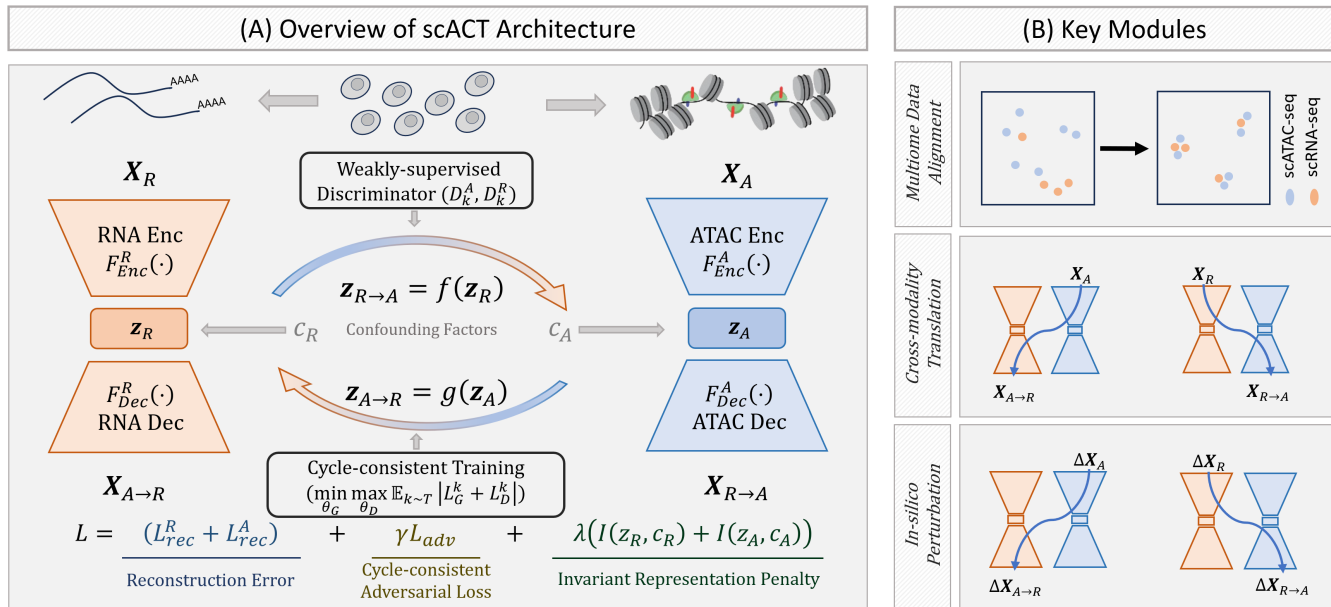


Figure 1: Overview, schematics, and use cases of scACT. (A) scACT’s workflow starts from the two encoder-decoder structure regularized to remove confounding factors, coupled by the two cross-modality transformation function $f(\cdot), g(\cdot)$, plus the weakly-supervised discriminators (D_k^A, D_k^R) and the cycle-consistent training mechanism as part of the loss functions. In the illustration, scACT uses scRNA-seq and scATAC-seq from the same tissue as input; however, these two modalities do not need to have a one-to-one correspondence (no co-assay data required). (B) Key modules of scACT with multiome data alignment (top), cross-modality translation in both directions (center), and in-silico perturbation (bottom).

multi-modal data, (ii) accurate **cross-modality translation** without prior knowledge, and (iii) interpretable **in-silico perturbations** to dissect regulatory relationships across modalities. To test its performance, we applied it on diverse single-cell datasets and found it outperformed existing methods in all three tasks. We have developed scACT as an individual open-source software package¹ to enable the scientific community for efficient processing and analysis of single-cell multiome data. Given the exponential growth of the amount of single-cell sequencing data, we anticipate that scACT will deepen our understanding of cross-modality interplay at the individual cell level and establish further downstream analysis as well as in-silico experiments to address tough biological problems with ease.

2 Methods

2.1 Method overview

As shown in Fig. 1, scACT is a deep generative model with three key modules, including data alignment, cross-modality translation, and regulatory relationship inference via in-silico perturbations. Specifically, it first learns modality-specific cell embeddings to overcome the sparsity of single cell data. Then it uses cycle-consistent training to allow robust cross-modality mapping, followed by accurate cross-modality translation via the decoder structures. To further boost its performance, we added adversarial training in a weakly supervised fashion to avoid possible feature distribution

shift. As a result, no paired data is needed throughout scACT’s entire learning process.

scACT is a general framework that can handle various types of single cell data, including but not limited to single-cell transcriptomics, epigenetics, proteomics. In the following sections, we used the example of scRNA-seq and scATAC-seq as examples.

2.2 Module 1 – multi-modal data alignment

Deep autoencoder for single-modality learning. First, we applied two deep autoencoders for scRNA-seq and scATAC-seq. Given the log-normalized scRNA-seq cell-by-gene matrix $X_R \in \mathbb{R}^{m \times q}$ with m cells and q genes and the scATAC-seq cell-by-peak matrix $X_A \in \mathbb{R}^{n \times p}$ with n cells and p peaks, we constructed two encoders f_{Enc}^R and f_{Enc}^A with trainable parameters W_{Enc}^R and W_{Enc}^A to calculate modality-specific cell embeddings $z_R^{(i)} \in \mathbb{R}^k$ and $z_A^{(j)} \in \mathbb{R}^k$ of dimension k (see Suppl. Table 1 for detailed notations) as shown in Eq. 1 (see supplementary methods for details).

$$\begin{aligned} z_R^{(i)} &= F_{Enc}^R(W_{Enc}^R, x_R^{(i)}) \\ z_A^{(j)} &= F_{Enc}^A(W_{Enc}^A, x_A^{(j)}) \end{aligned} \quad (1)$$

Similarly, decoders f_{Dec}^R and f_{Dec}^A were created with trainable parameters W_{Dec}^R and W_{Dec}^A to reconstruct the input vectors $x_{A \rightarrow R}^{(i)}$

¹<https://github.com/aicb-ZhangLabs/scACT>

and $x_{R \rightarrow R}^{(j)}$ in Eq. 2.

$$\begin{aligned} x_{A \rightarrow A}^{(i)} &= F_{Dec}^R(\mathbf{W}_{Dec}^R z_{A \rightarrow A}^{(i)}) \\ x_{R \rightarrow R}^{(j)} &= F_{Dec}^A(\mathbf{W}_{Dec}^A z_{R \rightarrow R}^{(j)}) \end{aligned} \quad (2)$$

In our architecture, we followed biological features of ATAC-seq by first limiting hidden layer connections by chromosome with 128, 64, and 32 units in hidden layers, leading to a 20-unit dimensional reduction for the latent space. This setup, adopted in both encoder (f_{Enc}^R) and decoder (f_{Dec}^R) phases, captured the chromosomal intricacies effectively, enabling enhanced pattern recognition and accurate data reconstruction.

We adopted the mean squared error (MSE) and the binary cross-entropy loss (BCE) as the reconstruction loss functions for the scRNA-seq and scATAC-seq modalities. Combining them and using the reconstructed matrices $\mathbf{X}_{R \rightarrow R}$ and $\mathbf{X}_{A \rightarrow A}$, we obtain Eq. 3, the overall reconstruction loss \mathcal{L}_{Rec} .

$$\mathcal{L}_{Rec} = \text{MSE}(\mathbf{X}_R, \mathbf{X}_{R \rightarrow R}) + \text{BCE}(\mathbf{X}_A, \mathbf{X}_{A \rightarrow A}) \quad (3)$$

To address the high dimensionality and sparsity of scATAC-seq data, we followed our previous approaches [6, 33] by removing inter-chromosomal connections in \mathbf{W}_{Enc}^A and \mathbf{W}_{Dec}^A .

Invariant representation learning to correct batch effect.

Given the confounding factors $c_R^{(i)} \in \mathbb{R}^a$ and $c_A^{(j)} \in \mathbb{R}^b$ (e.g., sequencing depth & batch) for each cell, we added \mathcal{L}_{Con} to minimize the mutual information between the confounding factors and the single-modality embeddings via Eq. 4.

$$\mathcal{L}_{Con} = \sum_i I(z_R^{(i)}, c_R^{(i)}) + \sum_j I(z_A^{(j)}, c_A^{(j)}) \quad (4)$$

Specifically, we approximated the mutual information using its upper bound

$$\begin{aligned} I(z, c) &\leq \mathbb{E}_{x, c \sim q(x, c)} [D_{KL}(F_{Enc}(x) \| \sum_{i'} F_{Enc}(x))] \\ &\quad - \mathbb{E}_{z_R \sim q_\phi(z|x)} [\log F_{Dec}(z)] \end{aligned} \quad (5)$$

Encoder-mapping-decoder structure for cross-modality translation. Following previous works [17, 36], we added two multi-layer perceptrons (MLPs) as the cross-modality mapping functions $f(\cdot)$ from RNA to ATAC and g from ATAC to RNA with trainable parameters \mathbf{W}_{RA} and \mathbf{W}_{AR} to generate the mapped embedding $z_{R \rightarrow A}^{(i)}$ and $z_{A \rightarrow R}^{(j)}$, respectively as shown in Eq. 6.

$$\begin{aligned} z_{R \rightarrow A}^{(i)} &= f(\mathbf{W}_{RA}, z_R^{(i)}) \\ z_{A \rightarrow R}^{(j)} &= g(\mathbf{W}_{AR}, z_A^{(j)}) \end{aligned} \quad (6)$$

Weakly-supervised adversarial training to overcome feature distribution shift. Assuming that the cell type annotations for both modalities were already known, we could use them as a weakly-supervised criteria throughout the training process of the translation. Therefore, we added an adversarial training scheme in scACT in the end. To ensure robustness in the generated modality-agnostic latent space z , we adopted a generative adversarial training mechanism [1, 11] with a discriminator using cell type annotations. More specifically, let the cell type annotation for each cell be

$\mathbf{C} = \{c_1, c_2, \dots, c_n\}$ where $c_i \in \{1, 2, \dots, k\}$ denoting k cell types. Then, we created networks D_A^t and D_R^t with regard to each cell type t for the ATAC and RNA embeddings z_A and z_R to predict whether the embedding (encoded or cross-generated) belongs to the cell type t or not. We calculated the log-likelihood and used the evidence lower-bound (ELBO) as the adversarial training loss shown in Eq. 7,

$$\begin{aligned} \mathcal{L}_D^k &= \mathbb{E}_{x \sim C^k} [\log D_A^k(z_A)] + \mathbb{E}_{y \sim C^k} [\log(1 - D_A^k(z_{R \rightarrow A}))] \\ &\quad + \mathbb{E}_{y \sim C^k} [\log D_R^k(z_R)] + \mathbb{E}_{x \sim C^k} [\log(1 - D_R^k(z_{A \rightarrow R}))] \end{aligned} \quad (7)$$

where C^k is the set of cells within cell type k .

Cycle-consistent training for robust reconstruction. In order to learn a robust mapping, we adopted a cycle-consistency loss using the output from the discriminators and the transformation functions so that they could be trained to simultaneously fool the discriminator and keep the cross-modality mapping consistent [36]. For example, for the RNA modality, we calculate the distance (MSE) between the embedding $z_R^{(i)}$ and its cycle-translated embedding $g(f(z_R^{(i)}))$. Then, we combine the negative log probability of the translated embedding $z_{R \rightarrow A} = f(z_R)$ being in the same cell type k . Similarly, we used BCE for the same procedure for the ATAC part. Combining both modalities, we obtained the cycle-consistency loss using Eq. 8.

$$\begin{aligned} \mathcal{L}_G^k &= \mathbb{E}_{x \sim C^k} [-\log D_R^k(z_{A \rightarrow R}) + \text{BCE}(f(g(z_A)), z_A)] \\ &\quad + \mathbb{E}_{y \sim C^k} [-\log D_A^k(z_{R \rightarrow A}) + \text{MSE}(g(f(z_R)), z_R)] \end{aligned} \quad (8)$$

Overall objective function. Therefore, the overall generative adversarial training process could be summarized in the following loss function:

$$\mathcal{L}_{Adv} = \min_{\theta_G} \max_{\theta_D} \mathbb{E}_{k \sim T} [\mathcal{L}_G^k + \mathcal{L}_D^k] \quad (9)$$

where θ_G were the trainable parameters of encoders f_{Enc}^r, f_{Enc}^a and the cross-mapping layers f, g , θ_{Dis} collected parameters of all T pairs of discriminators D_A^k, D_R^k . Then the overall cross-modality mapping objective is denoted in Eq. 10.

$$\mathcal{L} = \mathcal{L}_{Rec} + \gamma \mathcal{L}_{Adv} + \lambda \mathcal{L}_{Con} \quad (10)$$

Within the loss, γ and λ are hyperparameters to weight the adversarial and confounding loss.

2.3 Module 2 – cross-modality translation

In order to translate to the RNA modality from ATAC, given the input vector $x_A^{(i)}$, scACT generated the mapped RNA embedding $z_{A \rightarrow R}^{(i)} = g(F_{Enc}^A(\mathbf{W}_{Enc}^A, x_A^{(i)}))$ (Eq. 1,6). Then, the RNA decoder generated the translation $x_{A \rightarrow R}^{(i)} = F_{Dec}^R(\mathbf{W}_{Dec}^R, z_{A \rightarrow R}^{(i)})$ (Eq. 2). Translation from RNA to ATAC followed similar workflow and is summarized in Eq. 11 (see supplementary methods for details).

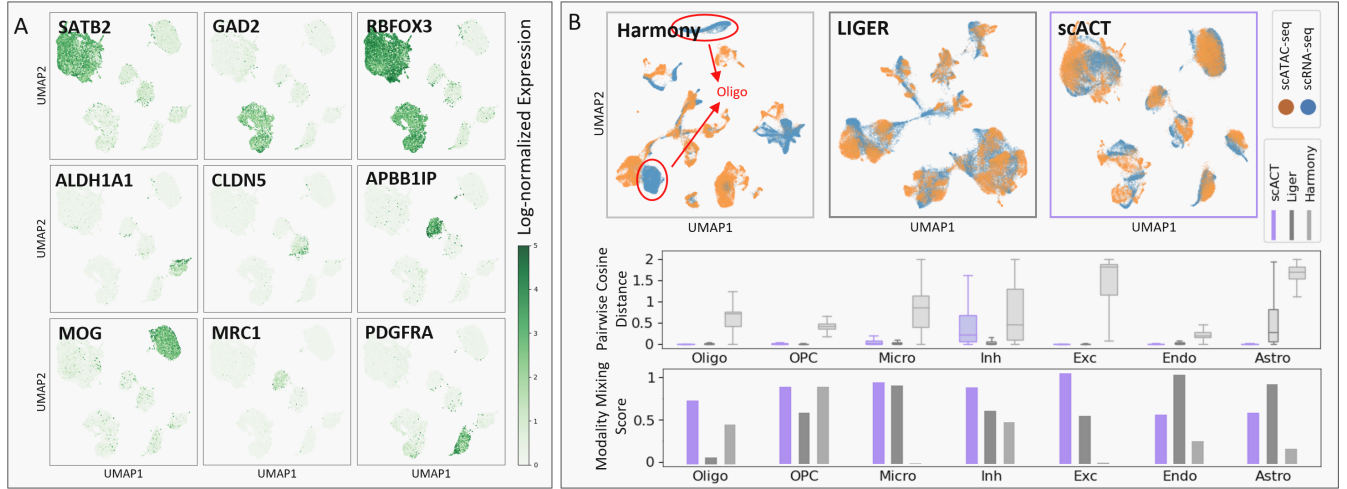


Figure 2: Multiome alignment performance benchmarks. (A) scACT generates better joint UMAP embeddings than Harmony or LIGER. (B) scACT’s joint embedding colored by actual gene expressions of notable marker genes show cell type consistency preserved in the embedding. (C) Pairwise cosine distance distributions within cell types show the closeness of cells within the same cell type in the generated embedding by scACT. (D) scACT mixed different modalities reasonably well.

$$\begin{aligned}
 z_R^{(i)} &= F_{Enc}^R(W_{Enc}^R, x_A^{(i)}) \\
 z_{R \rightarrow A}^{(i)} &= f(z_R^{(i)}) \\
 p_{R \rightarrow A}^{(i)} &= F_{Dec}^A(W_{Dec}^A, z_{R \rightarrow A}^{(i)}) \\
 x_{R \rightarrow A}^{(i)} &\sim \text{Bernoulli}(p_{R \rightarrow A}^{(i)})
 \end{aligned} \tag{11}$$

Note that a bernoulli sampling step was required to generate the binarized translation to the ATAC modality.

2.4 Module 3 – in-silico perturbations

Based on scACT’s flexible encoder-mapping-decoder framework, we can easily conduct in-silico perturbations to obtain key regulatory insights in a cell-specific manner. Specifically, we used the concept of Integrated Gradient [29]. Given a binary input $x_A^{(i)}$ from the ATAC modality $x_A^{(i)}$ and its baseline value $\tilde{x}_A^{(i)}$ where the baseline value would be 0 for the most cases, indicating inaccessible chromatin, the integrated gradient would be:

$$\begin{aligned}
 \text{IntegratedGradient}(\mathbf{x}_A^{(i)}) &= (\mathbf{x}_A^{(i)} - \tilde{\mathbf{x}}_A^{(i)}) \\
 &\times \int_{\alpha=0}^1 \frac{\partial F(\tilde{\mathbf{x}}_A^{(i)} + \alpha(\mathbf{x}_A^{(i)} - \tilde{\mathbf{x}}_A^{(i)}))}{\partial \mathbf{x}_A^{(i)}} d\alpha
 \end{aligned} \tag{12}$$

We used this integrated gradient score as the feature importance scores with regard to a gene to find downstream analysis and discover novel regulating elements.

2.5 Training details

Since the two key hyperparameters in scACT are γ and λ , we adopted a grid searching approach to fine tune them by conducting a uniform search from 0.1 to 2 for both hyperparameters. Then, the best-performing model with regard to the converged overall loss

was selected. After fixing the key hyperparameters, we conducted another search on the learning rate and the batch size used in the Adam optimizer we adopted in the process. We also iteratively searched from 0.00001 to 0.001 and from 16 to 256, and selected the best combination based on convergence speed and GPU memory available to us. We iteratively updated the autoencoders and the cycle-consistency units. Specifically, we prioritized on autoencoder training for feature extraction before focusing on the training of transformation functions and discriminators, and opted for a naive GAN over Wasserstein GAN based on its superior validation performance. This iterative training strategy ensures a balance between computational efficiency and model efficacy, enabling competitive outcomes without compromising on training duration or resource consumption. Detailed training steps were illustrated in the supplementary section.

2.6 Datasets

We included the publicly available human peripheral blood mononuclear cells (PBMC) 10k dataset and brain prefrontal cortex (PFC) dataset with parallel scRNA-seq and scATAC-seq sequencing to test our model. During the training process, we ignored the shared barcode to mimic the unpaired single-cell data, while we use the linking information as ground truth in our evaluation processes. For all modalities, we first extracted counts using CellRanger-arc (version 2.0.2) with hg38 and default parameters (see detailed pre-processing steps in supplementary methods).

scRNA-seq pre-processing. We filtered out cells with insufficient reads (< 200) or possible multiplets [10] and kept the top 3,000 highly variable genes to form the scRNA-seq matrix using Pegasus (version 1.7.1) and DoubletDetection (version 4.2). We then conducted log-normalization on the entire matrix to yield the final matrix for training. We also conducted LEIDEN clustering [30]

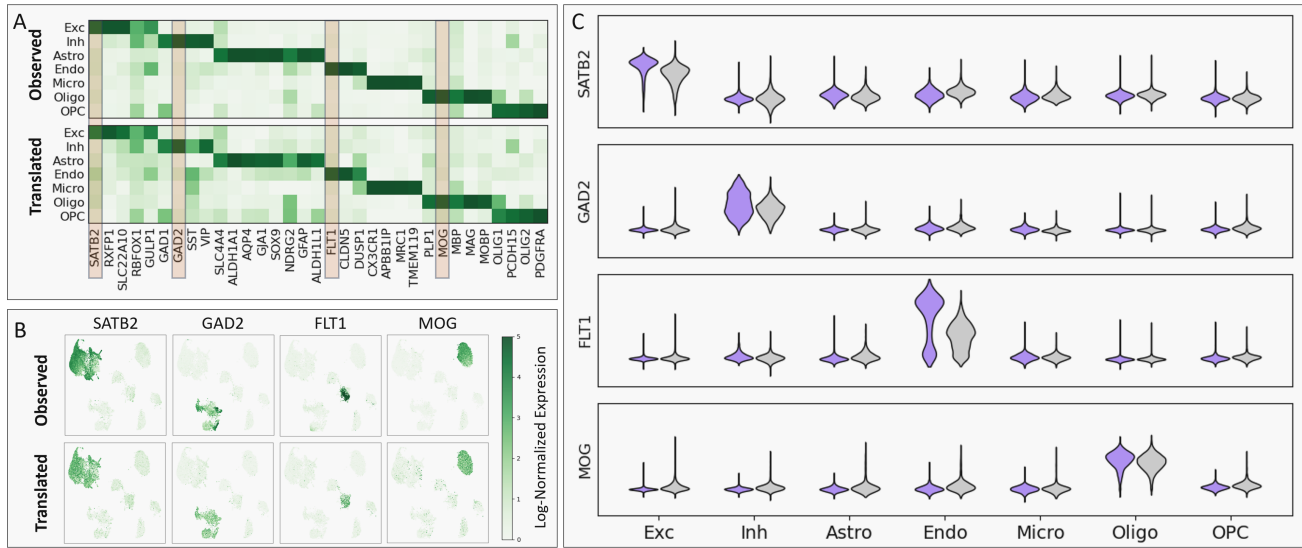


Figure 3: scRNA-seq prediction performance. (A) Distribution of predicted marker gene expression vs. ground truth across cell types showed great consistency and cell type specificity. (B) UMAP of actual scRNA-seq colored by actual and predicted gene expression values for the marker genes showed close resemblance. (C) normalized mean actual vs. predicted expression values captured cell type-specific variations.

based on the PCA results (20 dimensions) and annotated their cell types using previously-studied marker genes.

scATAC-seq pre-processing. Similarly, we filtered out cells with insufficient TSS enrichment (< 2.0), poor sequencing depths (< 1000), or possible multiplets using ArchR (version 1.0.1) and default parameters otherwise [12]. The peaks were called using Macs2 (version 2.2.9.1) [35] and underwent TF-IDF algorithm [25] such that only the most informative 100,000 peaks were kept. Finally, the binarized matrix was used in the training and evaluation process. To generate the cell type annotation for the discriminator, LEIDEN clustering based on the LSI results (top 20 dimension) were also used and the gene activity scores.

Summary of data used in this study. Previously, we downloaded deeply sequenced scATAC-seq and scRNA-seq co-assayed data from three frozen PFC tissues (GSE216270). After strict QC, we kept 27,414 barcodes from 5 samples with seven major cell types (excitatory, inhibitory, astrocytes, endothelial, microglia, oligodendrocytes, and OPC) using the set of marker genes proposed by Lake et al. [21].

We also use the publicly available human peripheral blood mononuclear cells (PBMC) 10k dataset from the 10x Genomics website. We kept 11,582 cells from 28 cell types as suggested on the website. Cell type annotations were adopted from the Signac vignettes [28].

2.7 Performance Benchmarking and Evaluation metrics

We thoroughly benchmark scACT’s performance on all three tasks, including alignment, cross-modality translation, and perturbation, against state-of-the-art methods on diverse datasets.

Qualitative and quantitative assessment of data alignment. We first visualize the latent space via joint-embedding UMAP [24].

Given the embeddings from RNA and ATAC modality on the same space, we created UMAP using the umap-learn package (version 0.5.5) with cosine distance, 13 neighbors, and default parameters otherwise. It was then colored by (i) the modality and (ii) the gene expression value of certain marker genes to show global homogeneity. The same UMAP generation procedure was done for Harmony and LIGER as well for comparison.

Also, to quantitatively measure the alignment, we adopted 2 metrics for global homogeneity and 2 metrics for local homogeneity. For global homogeneity, we calculated the silhouette score and the adjusted rand index (ARI) between the generated embedding and the ground truth annotation using scikit-learn (version 1.4). For local homogeneity, we measured pairwise cosine distance and modality mixing scores. We calculated a distribution of pairwise cosine distances for any two cells within the same cell type using the embeddings generated and conducted T-test. We also calculated the modality mixing score to measure how well the two modalities were mingled together in the generated joint embedding. Given the integrated latent embedding and for each cell type k , we randomly choose one cell c . Assuming that among the 1,000 closest cells, c_R cells were from RNA and c_A were from ATAC, then the modality mixing score m_k could be expressed in Eq. 13.

$$m_k = 1 - \left| \frac{c_A}{c_A + c_R} - 0.5 \right| \quad (13)$$

In the perfect case where half of the neighboring cells were from one modality, the modality mixing score would be 1.

ATAC-to-RNA Translation. We first visualize the generated scRNA-seq data by overlaying the expression values of previously-studied marker genes [21] on the generated UMAP. Also, we plotted a set of cell type-specific distributions for some marker genes to

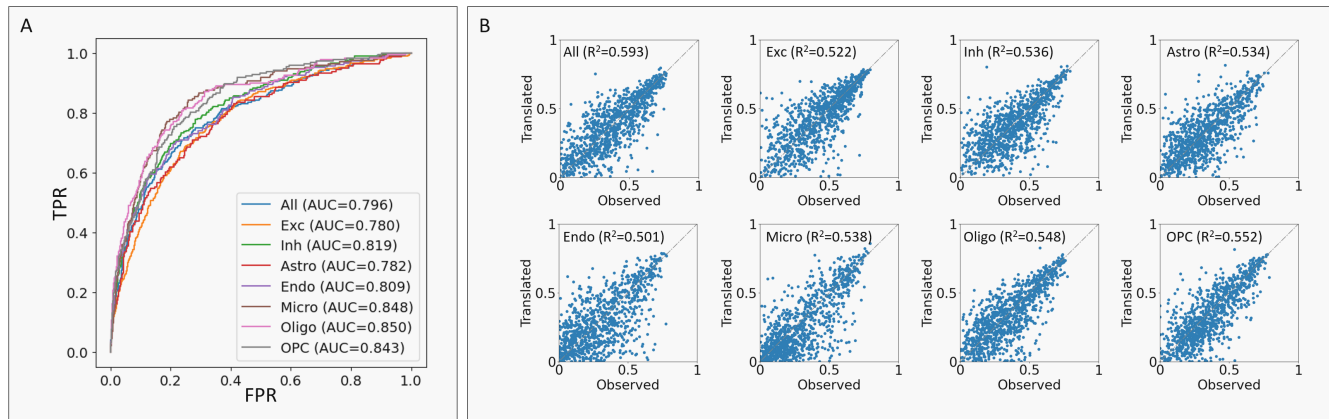


Figure 4: scATAC-seq prediction performance. (A) ROC curves captioned with AUCROC curves for both the entire data and each cell type. (B) Correlations between the observed and translated chromatin accessibility scores captioned with R^2 scores for both the entire data and each cell type.

evaluate the difference of the predicted gene expressions across cell types. We also plotted a heatmap showing the normalized gene expression values for more marker genes across all cell types.

RNA-to-ATAC translation. Since chromatin accessibility directly controls promoter-enhancer interactions, which in turns impacts motif enrichment and hence gene expression, we calculated activation scores for certain cell type-specific motifs. We here define a motif activation score,

$$a_n = \frac{\sum_{p \in P'} \lambda_p p}{|P'|} \quad (14)$$

to be a weighted average of accessibility of all peaks to the upstream of 750k base pairs of the transcription start site (TSS) and plotted a heatmap showing the activation scores for motifs of marker genes. In addition, we plot a cell type-specific motif footprint using the predicted scATAC-seq information for well-studied motifs and compare it to the ground truth.

In-silico perturbations. We plot the top most important peaks along with the target gene to show the capability of the model capturing important promoter-enhancer interactions of a particular marker gene (*SATB2*). To validate the found peak-to-gene interactions with regard to cell types, we plotted relevant gene expression profiles around the gene as well as the calculated peak-to-gene linkages with the most-important peaks found by scACT.

3 Results

3.1 scACT improves both local and global homogeneity when aligning different modalities

In order to evaluate scACT extensively, we adopted various metrics on its first module, multi-modal alignment on diverse datasets. As an example, we extracted a cohort of five human prefrontal cortex brain samples with 27,414 cells with both scRNA-seq and scATAC-seq and conducted model training and evaluation with the cell-level correspondence removed. Then, the trained model was evaluated

from both global and local homogeneity perspectives against other baseline models, Harmony and LIGER. Through this process, a wide spectrum of metrics, such as UMAP, silhouette score, cosine distance, and modality mixing score were used. In general, we found that scACT outperformed existing methods significantly and yielded biologically-interpretable results.

First, for global homogeneity assessment, UMAP and silhouette scores were employed as metrics. scACT showed superior global homogeneity by generating joint UMAP representations with more meaningful and biologically-interpretable clusters (Fig. 2A). Neuronal and non-neuronal marker gene expressions overlayed on scACT's joint embedding showed significant enrichment agreed with their cell type annotation (mean T-test p-values < 0.001) and were consistent with previous studies [21]. Also, any marker gene only illuminated one cluster, meaning that the UMAP representation was created based on cell types, not other confounding factors, such as input modalities, batch effects, or sequencing depths.

The UMAP colored by modality not only confirmed scACT's global homogeneity, but also proved its great local homogeneity (Fig. 2B). Harmony's embedding exhibited significant separation between modalities, lacking consistency. For instance, in the UMAP, scRNA-seq clusters, representing oligodendrocyte cells, were distant from corresponding clusters in scATAC-seq, which included endothelial and microglia cells (Suppl. Fig. 1). This discrepancy likely stems from neglecting batch effects, emphasizing experimental variations over intercellular differences. LIGER demonstrated improved cell type consistency but displayed teardrop-shaped clusters in the UMAP, suggesting susceptibility to technical factors like sequence depth. In contrast, scACT achieved better results, maintaining consistent cell type mapping and forming round or oval clusters, indicating minimal impact from confounding factors. The silhouette score further validated this, with scACT achieving a higher score (0.612) than Harmony (-0.011) and LIGER (0.510). Additional metrics (ARI and AMI) also showed clear advantages of scACT (Suppl. Table 2).

To further evaluate local homogeneity, we tested the closeness of cells within the same cell type. scACT outperformed its counterparts with a significantly lower distribution of pairwise cosine distance across all cell types (mean F-test p-value=0.0), especially in the astrocytes cells (0.026 vs. 0.426 and 0.576). Moreover, a high modality mixing score (F-test p-value < 0.05) was observed for all cell types except endothelial and astrocyte cells. In particular, in the excitatory cells, scACT was able to mix the two modalities much better than the baseline methods (0.973 vs. 0.526 and 0.081). We suspect that the less optimal performance could be partially attributed to the imbalance of training data and lack of samples, but with more abundant data, we believed that scACT could deliver its local homogeneity to all cell types.

3.2 scACT accurately translates the ATAC modality into the RNA modality

Moving one-step further, we evaluated scACT’s module 2: scRNA-seq translation performance by examining marker gene enrichments and expression similarity across cell types and observed largely preserved cell type-specific signatures. First, we selected a curated list of marker genes according to previous study [21] and compared the mean expressions between cell types and between the observed and translated cohort (Fig. 3A). We found that scACT managed to preserve similar expression patterns with the observed ground truth (mean $R^2=0.914$), but also provided strong enrichment (mean normalized value: enriched=0.894 vs. non-enriched=0.122, mean F-test p-value < 0.001, see Suppl. Table 3 for details) with the cell type enriched consistent with previous findings [2].

Taking a closer look, we focused on four key marker genes—*SATB2* for excitatory cells, *GAD2* for inhibitory cells, *FLT1* for endothelial cells, and *MOG* for oligodendrocyte cells as they were the well-known marker genes commonly recognized. The UMAP of them agreed with our previous findings, with the translated expression highlighting the mentioned cell types, plus a high correlation between the observed and the translated ($R^2 = 0.633, 0.603, 0.460, 0.684$, Fig. 3B). We noticed that the correlation of *FLT1* fell behind other marker genes, and suspected that the reason being the lack of training samples in that specific cell type and possible remaining multiplets (upper-left of the endothelial cluster), but overall, it yielded decent performance.

The distribution across cell types of observed versus translated profiles further verified the accuracy and correctness of scACT (Fig. 3C). A significant enrichment of these marker genes across their respective cell types was shown across all selected marker genes (mean F-test p-value < 0.0001), confirming scACT’s ability to accurately capture cell type-specific features. Plus, the distributions themselves from scACT also strictly followed the observed truth, indicating a robust learning capability. From both the distributional statistics to the actual data of each individual cell, we demonstrated its accurate translation capability from scATAC-seq to scRNA-seq.

3.3 scACT imputes meaningful and interpretable scATAC-seq signals

Besides the robust scRNA-seq translation capability, we also assessed scACT’s scATAC-seq translation performance from the RNA modality. From a computational point of view, we can treat the

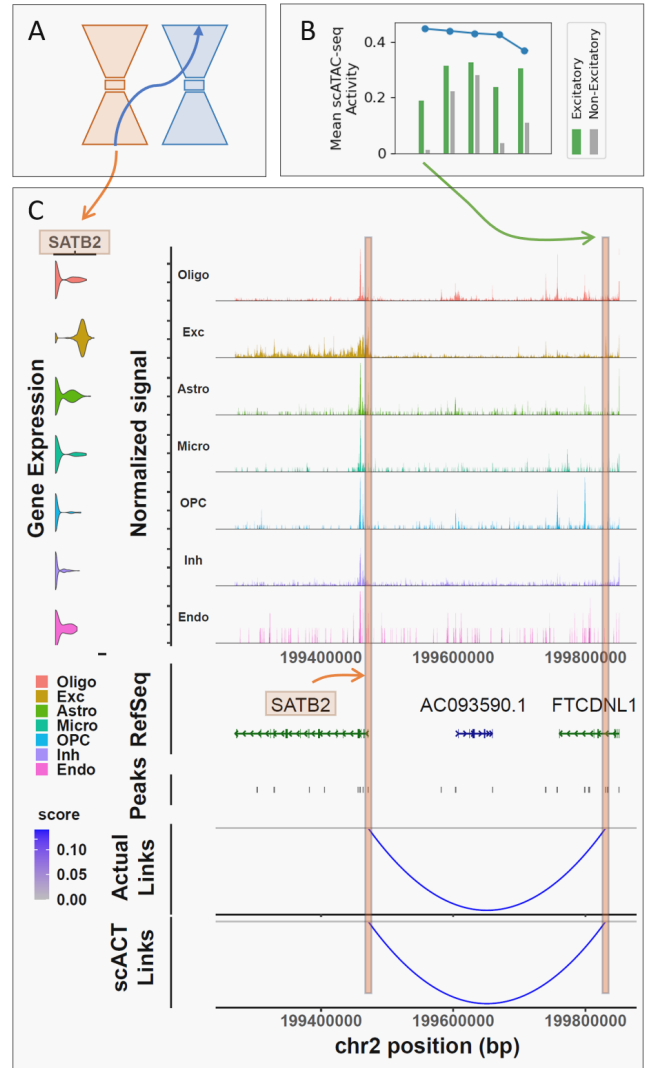


Figure 5: In-silico perturbation results with scACT. (A) General schematics of the module 3: perturbation framework of scACT. (B) Feature importance score (blue) plus the mean ground truth scATAC-seq activity (green and gray) for the top five peaks associated with *SATB2*. (C) Cell type-specific gene expression distribution, Tn5 insertion intensity, peak annotation, and the peak-to-gene linkages from both the ground truth and the scACT-translated scATAC-seq data.

translation task as a binary classification problem where for each cell, we predict the probability of each peak being accessible. In this way, we adopted the widely-used ROC curves and AUCROC scores as a key evaluation metric (Fig. 4A). Overall, scACT exhibited high accuracy, reflected in the area under the ROC curve (AUCROC=0.796). Although we found that our AUCROC was lower than other reported methods [32], we believed the reason being the different tissue of study (lymphoblastoid cells vs. brain cells) and the quality of samples used (curated cell lines vs. post-mortem

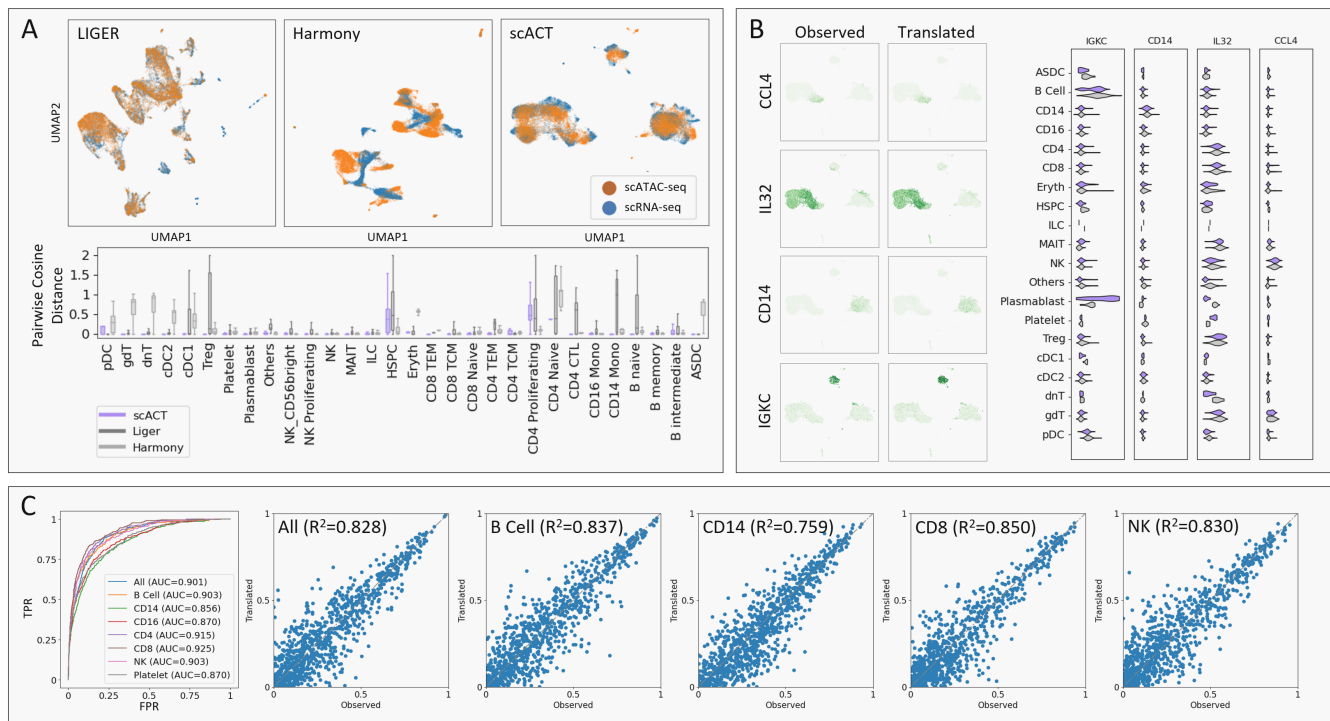


Figure 6: Evaluation of scACT on human PBMC multiome data. (A) Evaluation on multi-modal alignment with UMAP representations of LIGER, Harmony, and scACT (top) and comparison of pairwise cosine distances across cell types (bottom). (B) scRNA-seq translation performance with observed and predicted marker gene expressions (left) and the per-cell type distribution (right). (C) scATAC-seq translation benchmark with ROC curves captioned with AUCROC curves (left) and correlations between the observed and translated chromatin accessibility scores (right).

tissues). We also believed that the cell-to-cell correspondence required in the other model helped largely. However, in the real world, co-assay data were at a scarce and the quality of sample varied dramatically, and thus scACT could be a better candidate providing meaningful translation. Moreover, subsetting the problem into cell type-specific clusters, we observed consistent and reasonable performance (AUCROC=0.780-0.850).

Additionally, viewing the translation problem as a regression problem in the cell type level, we extracted chromatin accessibility scores from both observed and translated data, allowing us to evaluate the correlation between them. scACT demonstrated excellent correlation ($R^2=0.593$) across all data and maintained consistent performance within each cell type ($R^2=0.501-0.548$). These results jointly highlighted scACT’s proficiency in translating scATAC-seq profiles, showcasing both accurate classification and preservation of chromatin accessibility patterns.

3.4 scACT enables in-silico perturbations to identify enhancer-gene linkages

To evaluate scACT’s in-silico perturbation capability, we focused on *SATB2*, a marker gene for excitatory cells (Fig. 5A). Feature importance scores were computed using scACT’s module 3, and the mean chromatin accessibility of excitatory and non-excitatory cells plotted alongside their respective feature importance scores for

the top 5 peaks (Fig. 5B). Notably, T-tests demonstrated a statistically significant enrichment on the top peaks between these two cell groups ($p\text{-value} < 0.001$) and especially in the most important peak (activity=0.186 vs. 0.013, $p\text{-value}=0.0$), indicating the effectiveness of scACT in simulating perturbations that lead to discernible chromatin accessibility changes associated with the excitatory cell marker *SATB2*.

In addition to the test statistics across cell types, we validated the results from scACT and explored peak-to-gene linkages in both the observed and translated scATAC-seq data, focusing on the most contributing peak identified by scACT in the single-cell level (Fig. 5C). This approach allowed us to elucidate regulatory connections associated with the perturbation of *SATB2*. We found enrichment of signal for excitatory neurons on the exons and especially around the transcription start site (TSS), consistent with previous studies [16]. To our expectation, a strong enrichment was also observed for excitatory neurons on the peak identified by scACT. Further validation using the identification of peak-to-gene linkages in both observed and translated chromatin accessibility profiles (span of 100k bp; pearson test $p\text{-value} < 0.05$) supported scACT’s capability to not only simulate perturbations accurately, but also reveal potential regulatory mechanisms, providing valuable insights into the functional consequences with little need to conduct biological experiments.

3.5 scACT provides accurate alignment and translation in human PBMC

Since scACT is designed to handle various organisms and tissues, we expanded our study to evaluate its performance in the human PBMCs, which has a broader, more continuous spectrum of cell types compared to the brain prefrontal cortex data. This complexity poses additional challenges for computational models. Through extensive benchmarking, we observed that scACT maintained its robust multi-modal alignment and translation capabilities, outperforming baseline models.

Firstly, we scrutinized scACT's module 1, focusing on multi-ome alignment (Fig. 6A) and demonstrated its substantial ability to create joint embeddings. Qualitatively, scACT demonstrated global homogeneity by generating UMAP representations with more meaningful clusters. Compared to Harmony, both LIGER and scACT generated embeddings invariant to input modalities by integrating the two modalities. Moreover, scACT surpassed the baselines in cell grouping when benchmarked against annotated cell types (silhouette score: scACT=0.128, Harmony=0.125, LIGER=-0.076). Additionally, scACT's embedding exhibited an oval-shaped structure, indicating resistance to sequencing depths and batch effects, unlike Harmony and LIGER which generated clusters with sharp edges. Quantitatively, scACT showed significantly closer clusters by cell types, especially in smaller cell types like dnT and cDC1 (mean F-test p-value < 0.001).

Next, a comprehensive inspection of scACT's scRNA-seq translation module revealed its improved accuracy (Fig. 6B). Analysis of four marker genes (*CCL4* for NK cells, *IL32* for CD8 cells, *CD14* for CD14 cells, and *IGKC* for B cells) indicated high specificity with respect to the represented cell types (T-test p-value < 0.001), consistent with previous studies [13, 23, 27]. There was a high correlation between observed and translated matrices in most cases ($R^2 = 0.433, 0.523, 0.245, 0.095$). For the *IGKC* gene, a potential cause for weaker performance was hypothesized to be the decreased number of cells in the B cell group and further subtypes differentiation, but in general scACT gave great performance. Zooming into cell type levels of the four marker genes revealed a clear preservation of cell type-specific signatures in expression distributions between observed and translated. Compared to out-groups, cell types corresponding to marker genes showed a significant increase in translated profiles (T-test p-value < 0.05). Overall, scACT successfully generalized its scRNA-seq translation capability to PBMC tissue.

We also benchmarked scACT's scATAC-seq translation performance on the PBMC dataset (Fig. 6C). It effectively reconstructed chromatin accessibility from gene expressions, formulated as binary classification problems evaluated through ROC curves. Overall, scACT demonstrated high accuracy (AUCROC=0.901). When subset to larger or smaller cell type-specific clusters, it performed reasonably well (AUCROC=0.856-0.925). From a biological perspective, we sought the correlation between observed and translated chromatin accessibility per cell and per peak. scACT exhibited excellent performance with an overall R^2 value of 0.828 across all data and consistent performance in each cell type ($R^2=0.759-0.850$). In multiple evaluations, scACT showcased considerable translation performance from transcription profile to chromatin accessibility at the single-cell resolution.

4 Conclusion and Discussion

Understanding multi-modal cellular heterogeneity is crucial in genomics, and single-cell sequencing data provide great opportunities. However, deeper understanding of the hidden biology requires concurrent analysis of multiple modalities while paired co-assay data are scarce. Plus, the analysis of unpaired single-cell multi-modal data presents significant challenges. In this context, we introduced scACT, a comprehensive model designed to address these challenges by excelling in multi-modal alignment, cross-modality translation, and in-silico perturbation. Our model is particularly adept at handling unpaired single-cell multi-modal datasets, offering a powerful solution to unravel complex biological interactions.

Extensive benchmarking on diverse datasets, with a focus on brain prefrontal cortex data, demonstrated scACT's superior multi-ome alignment performance compared to existing methods such as Harmony and LIGER, in preserving both global and local homogeneity in the generated joint embeddings. Plus, we showcased scACT's proficiency in cross-modality translation in two directions: from scATAC-seq into scRNA-seq with multiple gene-level analyses and from scRNA-seq to scATAC-seq using discriminative and regressive metrics. Focusing on *SATB2* as a case study, scACT accurately conducted in-silico perturbations, as evidenced by statistically significant enrichment in chromatin accessibility changes associated with excitatory cells, aiding in understanding regulatory mechanisms as well as reducing the reliance on labor-intensive and time-consuming biological experiments. Moreover, scACT demonstrated its applicability not only in brain cells, but also in other diverse tissues, such as human PBMCs, maintaining accuracy and consistency in all its modules.

In summary, scACT offers a comprehensive solution for analyzing unpaired single-cell multi-modal data, providing advancements in multi-modal alignment, cross-modality translation, and in-silico perturbation using a weakly-supervised approach with cycle-consistency. The model's robust performance across diverse datasets and tissues highlights its potential impact on advancing our understanding of complex biological systems. Currently, much more single-modality data are available with relatively low cost and high quality, whereas the emerging joint sequencing technology needs time and effort to mature. We believe that scACT can help the community in robust and accurate mapping, translation, perturbation, and further downstream analysis more effectively using more cohorts of data with higher quality to unlock more intriguing biological mechanisms and regulatory insights of different tissues and organisms.

Acknowledgments

We thank the Yale Center for Research Computing and the UCI ICS Computing Support for guidance and use of the research computing infrastructure. This work was funded in part by the National Institutes of Health grant R01HG012572 and R01NS128523.

References

- [1] Martin Arjovsky, Soumith Chintala, and Léon Bottou. 2017. Wasserstein GAN. <https://doi.org/10.48550/ARXIV.1701.07875>
- [2] Aritra Bhattacharjee, Mohamed Nadhir Djekidel, Renchao Chen, Wenqiang Chen, Luis M. Tuesta, and Yi Zhang. 2019. Cell type-specific transcriptional programs in mouse prefrontal cortex during adolescence and addiction. *Nature Communications* 10, 1 (Sept. 2019). <https://doi.org/10.1038/s41467-019-12054-3>

- [3] A. P. Boyle, S. Davis, H. P. Shulha, P. Meltzer, E. H. Margulies, Z. Weng, T. S. Furey, and G. E. Crawford. 2008. High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132, 2 (2008), 311–22. <https://doi.org/10.1016/j.cell.2007.12.014>
- [4] Jason D. Buenrostro, Beijing Wu, Ulrike M. Litzenburger, Dave Ruff, Michael L. Gonzales, Michael P. Snyder, Howard Y. Chang, and William J. Greenleaf. 2015. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 523, 7561 (June 2015), 486–490. <https://doi.org/10.1038/nature14590>
- [5] Kai Cao, Xiangqi Bai, Yiguang Hong, and Lin Wan. 2020. Unsupervised topological alignment for single-cell multi-omics integration. *Bioinformatics* 36, Supplement 1 (July 2020), i48–i56. <https://doi.org/10.1093/bioinformatics/btaa443>
- [6] Yingxin Cao, Laiyi Fu, Jie Wu, Qinke Peng, Qing Nie, Jing Zhang, and Xiaohui Xie. 2022. Integrated analysis of multimodal single-cell data with structural similarity. *Nucleic Acids Research* 50, 21 (Sept. 2022), e121–e121. <https://doi.org/10.1093/nar/gkac781>
- [7] X. Chen, R. J. Miragaia, K. N. Natarajan, and S. A. Teichmann. 2018. A rapid and robust method for single cell chromatin accessibility profiling. *Nat Commun* 9, 1 (2018), 5345. <https://doi.org/10.1038/s41467-018-07771-0>
- [8] M. Colomé-Tatché and F.J. Theis. 2018. Statistical single cell multi-omics integration. *Current Opinion in Systems Biology* 7 (Feb. 2018), 54–59. <https://doi.org/10.1016/j.coisb.2018.01.003>
- [9] D. A. Cusanovich, R. Daza, A. Adey, H. A. Pliner, L. Christiansen, K. L. Gunderson, F. J. Steemers, C. Trapnell, and J. Shendure. 2015. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* 348, 6237 (2015), 910–4. <https://doi.org/10.1126/science.aab1601>
- [10] Adam Gayoso and Jonathan Shor. 2022. JonathanShor/DoubletDetection: doublet-detection v4.2. <https://doi.org/10.5281/ZENODO.2678041>
- [11] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Networks. <https://doi.org/10.48550/ARXIV.1406.2661>
- [12] Jeffrey M. Granja, M. Ryan Corces, Sarah E. Pierce, S. Tansu Bagdatli, Hani Choudhry, Howard Y. Chang, and William J. Greenleaf. 2021. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nature Genetics* 53, 3 (Feb. 2021), 403–411. <https://doi.org/10.1038/s41588-021-00790-6>
- [13] Thomas Gruber, Mirela Kremenovic, Hassan Sadozai, Nives Rombini, Lukas Baeriswyl, Fabienne Maibach, Robert L. Modlin, Michel Gilliet, Diego von Werdt, Robert E. Hunger, S. Morteza Seyed Jafari, Giulia Parisi, Gabriel Abril-Rodriguez, Antoni Ribas, and Mirjam Schenk. 2020. IL-32γ potentiates tumor immunity in melanoma. *JCI Insight* 5, 18 (Sept. 2020). <https://doi.org/10.1172/jci.insight.138772>
- [14] Laleh Haghverdi, Aaron T L Lun, Michael D Morgan, and John C Marioni. 2018. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nature Biotechnology* 36, 5 (April 2018), 421–427. <https://doi.org/10.1038/nbt.4091>
- [15] Ashrafal Haque, Jessica Engel, Sarah A. Teichmann, and Tapio Lonnberg. 2017. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Medicine* 9, 1 (Aug. 2017). <https://doi.org/10.1186/s13073-017-0467-4>
- [16] Ying Huang, Ning-Ning Song, Wei Lan, Ling Hu, Chang-Jun Su, Yu-Qiang Ding, and Lei Zhang. 2013. Expression of Transcription Factor <sc>S</sc> atb2 in Adult Mouse Brain. *The Anatomical Record* 296, 3 (Feb. 2013), 452–461. <https://doi.org/10.1002/ar.22656>
- [17] Yangbangan Jiang, Qianqian Xu, Zhiyong Yang, Xiaochun Cao, and Qingming Huang. 2019. DM2C: Deep Mixed-Modal Clustering. In *Advances in Neural Information Processing Systems*. 5880–5890.
- [18] Dragomirka Jovic, Xue Liang, Hua Zeng, Lin Lin, Fengping Xu, and Yonglun Luo. 2022. Single-cell RNA sequencing technologies and applications: A brief overview. *Clinical and Translational Medicine* 12, 3 (March 2022). <https://doi.org/10.1002/ctm2.694>
- [19] S. L. Klemm, Z. Shipony, and W. J. Greenleaf. 2019. Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet* 20, 4 (2019), 207–220. <https://doi.org/10.1038/s41576-018-0089-8>
- [20] David Lahnemann, Johannes Koster, Ewa Szczurek, Davis J. McCarthy, Stephanie C. Hicks, Mark D. Robinson, Catalina A. Vallejos, Kieran R. Campbell, Niko Beerenwinkel, Ahmed Mahfouz, Luca Pinello, Pavel Skums, Alexandros Stamatakis, Camille Stephan-Otto Attolini, Samuel Aparicio, Jasmijn Baaijens, Marleen Balvert, Buys de Barbanson, Antonio Cappuccio, Giacomo Corleone, Bas E. Duttil, Maria Florescu, Victor Guryev, Rens Holmer, Katharina Jahn, Thamar Jessurun Lobo, Emma M. Keizer, Indu Khatri, Szymon M. Kielbasa, Jan O. Korbel, Alexey M. Kozlov, Tzu-Hao Kuo, Boudewijn P.F. Lelieveldt, Ion I. Mandoiu, John C. Marioni, Tobias Marschall, Felix Molder, Amir Niknejad, Alicja Rączkowska, Marcel Reinders, Jeroen de Ridder, Antoine-Emmanuel Saliba, Antonios Somarakis, Oliver Stegle, Fabian J. Theis, Huan Yang, Alex Zelikovsky, Alice C. McHardy, Benjamin J. Raphael, Sohrab P. Shah, and Alexander Schonhuth. 2020. Eleven grand challenges in single-cell data science. *Genome Biology* 21, 1 (Feb. 2020). <https://doi.org/10.1186/s13059-020-1926-6>
- [21] Blue B. Lake, Rizi Ai, Gwendolyn E. Kaeser, Neeraj S. Salathia, Yun C. Yung, Rui Liu, Andre Wildberg, Derek Gao, Ho-Lim Fung, Song Chen, Raakhee Vijayaraghavan, Julian Wong, Allison Chen, Xiaoyan Sheng, Fiona Kaper, Richard Shen, Mostafa Ronaghi, Jian-Bing Fan, Wei Wang, Jerold Chun, and Kun Zhang. 2016. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* 352, 6293 (June 2016), 1586–1590. <https://doi.org/10.1126/science.aaf1204>
- [22] Jialin Liu, Chao Gao, Joshua Sodicoff, Velina Kozareva, Evan Z. Macosko, and Joshua D. Welch. 2020. Jointly defining cell types from multiple single-cell datasets using LIGER. *Nature Protocols* 15, 11 (Oct. 2020), 3632–3662. <https://doi.org/10.1038/s41596-020-0391-8>
- [23] Azzam A. Maghazachi. 2010. *Role of Chemokines in the Biology of Natural Killer Cells*. Springer Berlin Heidelberg, 37–58. https://doi.org/10.1007/82_2010_20
- [24] Leland McInnes, John Healy, Nathaniel Saul, and Lukas Grobberger. 2018. UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software* 3, 29 (Sept. 2018), 861. <https://doi.org/10.21105/joss.00861>
- [25] Juan Ramos. 1999. Using TF-IDF to Determine Word Relevance in Document Queries.
- [26] A. T. Satpathy, J. M. Granja, K. E. Yost, Y. Qi, F. Meschi, G. P. McDermott, B. N. Olsen, M. R. Mumbach, S. E. Pierce, M. R. Corces, P. Shah, J. C. Bell, D. Jhutti, C. M. Nemecek, J. Wang, L. Wang, Y. Yin, P. G. Giresi, A. L. S. Chang, G. X. Y. Zheng, W. J. Greenleaf, and H. Y. Chang. 2019. Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nat Biotechnol* 37, 8 (2019), 925–936. <https://doi.org/10.1038/s41587-019-0206-z>
- [27] Marcus Schmidt, Patrick Micke, Mathias Gehrmann, and Jan G. Hengstler. 2012. Immunoglobulin kappa chain as an immunologic biomarker of prognosis and chemotherapy response in solid tumors. *OncImmunology* 1, 7 (Oct. 2012), 1156–1158. <https://doi.org/10.4161/onci.21653>
- [28] Tim Stuart and Rahul Satija. 2019. Integrative single-cell analysis. *Nature Reviews Genetics* 20, 5 (Jan. 2019), 257–272. <https://doi.org/10.1038/s41576-019-0093-7>
- [29] Mukund Sundararajan, Ankur Taly, and Qiqi Yan. 2017. Axiomatic Attribution for Deep Networks. <https://doi.org/10.48550/ARXIV.1703.01365>
- [30] V. A. Traag, L. Waltman, and N. J. van Eck. 2019. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports* 9, 1 (March 2019). <https://doi.org/10.1038/s41598-019-41695-z>
- [31] Maria Tsompana and Michael J Buck. 2014. Chromatin accessibility: a window into the genome. *Epigenetics & Chromatin* 7, 1 (Nov. 2014). <https://doi.org/10.1186/1756-8935-7-33>
- [32] Kevin E. Wu, Kathryn E. Yost, Howard Y. Chang, and James Zou. 2021. BABEL enables cross-modality translation between multiomic profiles at single-cell resolution. *Proceedings of the National Academy of Sciences* 118, 15 (April 2021). <https://doi.org/10.1073/pnas.2023070118>
- [33] Siwei Xu, Mario Skarica, Ahyeon Hwang, Yi Dai, Cheyu Lee, Matthew J. Girgenti, and Jing Zhang. 2022. Translator: A Transfer Learning Approach to Facilitate Single-Cell ATAC-Seq Data Analysis from Reference Dataset. *Journal of Computational Biology* 29, 7 (July 2022), 619–633. <https://doi.org/10.1089/cmb.2021.0596>
- [34] Ran Zhang, Laetitia Meng-papaxanthos, Jean-philippe Vert, and William Stafford Noble. 2022. Multimodal Single-Cell Translation and Alignment with Semi-Supervised Learning. *Journal of Computational Biology* 29, 11 (Nov. 2022), 1198–1212. <https://doi.org/10.1089/cmb.2022.0264>
- [35] Yong Zhang, Tao Liu, Clifford A Meyer, Jérôme Eeckhoutte, David S Johnson, Bradley E Bernstein, Chad Nusbaum, Richard M Myers, Myles Brown, Wei Li, and X Shirley Liu. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* 9, 9 (Sept. 2008). <https://doi.org/10.1186/gb-2008-9-9-r137>
- [36] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. 2017. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. <https://doi.org/10.48550/ARXIV.1703.10593>