

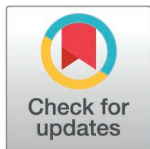
RESEARCH ARTICLE

MultiPert: An adversarial alignment and dual attention framework for single-cell multi-omics perturbation prediction

Mengyuan Zhao^{1,2}, Xinyue Tang³, Jiawei Li⁴, Cheng Liang⁵, Jijun Tang^{2,3*}, Fei Guo^{6*}

1 Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, **2** University of Chinese Academy of Sciences, Beijing, China, **3** Shenzhen University of Advanced Technology, Shenzhen, China, **4** College of Intelligence and Computing, Tianjin University, Tianjin, China, **5** Shandong Normal University, Jinan, China, **6** School of Computer Science and Engineering, Central South University, Changsha, China

* jj.tang@siat.ac.cn (JT); guofei@csu.edu.cn (FG)



OPEN ACCESS

Citation: Zhao M, Tang X, Li J, Liang C, Tang J, Guo F (2026) MultiPert: An adversarial alignment and dual attention framework for single-cell multi-omics perturbation prediction. PLoS Comput Biol 22(3): e1014054. <https://doi.org/10.1371/journal.pcbi.1014054>

Editor: Zhaoyuan Fang, Zhejiang University, CHINA

Received: September 26, 2025

Accepted: February 24, 2026

Published: March 11, 2026

Copyright: © 2026 Zhao et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All data used in this study is publicly available. The THP-1 dataset was downloaded from Zenodo (<https://zenodo.org/records/7041849>). The kidney dataset was downloaded from Gene Expression Omnibus (GEO) with accession number GSE213957. The transcriptome–epigenome

Abstract

Precise prediction of perturbation responses is essential in systems biology research, as it plays a pivotal role in characterizing cellular identities and elucidating the regulatory mechanisms of biological pathways. Existing perturbation-responses prediction approaches are predominantly confined to single-modality transcriptomic data, limiting their capacity to capture cross-layer molecular effects. Here, we present MultiPert, a deep learning framework specifically designed for predicting perturbation responses in single-cell multi-omics data. MultiPert employs modality-specific encoders with dedicated pretraining, integrates perturbation through a dual-attention mechanism, and achieves cross-modal alignment via adversarial training. Benchmarking on human THP-1 and kidney multi-omics datasets demonstrates that MultiPert reliably predicts both perturbed gene expression and protein abundance profiles, achieving superior accuracy and stability compared to state-of-the-art strategies. MultiPert generalizes to unseen perturbations and uncovers regulatory mechanisms of immune checkpoint molecules based on perturbed proteomic predictions. In addition, enrichment analyses of perturbed transcriptomic predictions reveal immune-related pathways. By providing an integrated and interpretable framework, MultiPert expands the scope of perturbation modeling at the multi-omics level, thereby offering a robust methodological foundation for comprehensive research into pathogenesis and drug discovery.

Author summary

In systems biology research, accurately predicting how cells respond to perturbations—such as gene knockout or drug intervention—is crucial for understanding cell identities and the regulatory mechanisms of biological pathways.

dataset and epigenome–proteome dataset were downloaded from Gene Expression Omnibus (GEO) with accession number GSE156478.

Funding: This work was supported by the National Natural Science Foundation of China (Grants No. U24A20257 to J.T., No. 62322215 and No. 62532017 to F.G.), Shenzhen Science and Technology Program (Grants No. JCYJ20241202130212016 and No. KQTD20200820113106007 to J.T., and No. JCYJ20230807140709020 to M.Z.), Natural Science Foundation of Hunan Province (Grants No. 2026JJ30018 to F.G.). This study was also supported in part by the High-Performance Computing Center of Central South University. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

However, most existing methods only support scRNA-seq data and cannot capture perturbation effects from single-cell multi-omics data. To address this limitation, we developed MultiPert to predict perturbation responses while integrating single-cell multi-omics data. It uses dedicated encoders for different molecular layers to capture unique biological signals, aligns multi-omics data through adversarial training, and fuses perturbation information via a dual attention mechanism. Experiments on different tissues show that MultiPert outperforms existing methods in predicting gene expression and protein abundance. It can also predict unseen perturbations and uncover the regulatory mechanisms. We hope this work provides a more comprehensive tool for studying disease pathogenesis and drug discovery, making multi-omics-level perturbation research more accessible.

Introduction

Precise prediction of perturbation responses holds unparalleled significance in systems biology research. It serves not only as an essential approach to functionally characterize genes but also as a core method to elucidate regulatory mechanisms of biological pathways [1–3]. In disease pathogenesis studies, perturbation prediction enables *in silico* simulation of how disease-causing mutations or pharmacological interventions alter cellular states, thereby providing novel perspectives on the molecular basis of disease initiation and progression [4–6]. At single-cell resolution, cells exhibit heterogeneity in their responses to perturbations, which correlates with drug resistance and cell fate determination. Conventional experimental approaches (e.g., CRISPR or RNAi) face inherent limitations in cost and throughput, restricting their capacity to interrogate complex combinatorial perturbations or dynamic processes [7–11]. Computational approaches leveraging single-cell data have thus emerged as pivotal tools for overcoming experimental barriers and accelerating perturbation studies [12–18]. By learning associations between perturbations and cellular states, these methods enable systematic modeling of cellular responses to perturbations, thereby advancing fundamental research including gene function and biological pathways, and empowering translational applications such as drug target discovery and personalized therapeutic design.

The proliferation of single-cell RNA sequencing (scRNA-seq) datasets has accelerated development of perturbation-response prediction tools including scGen, scPRAM, CPA, CoupleVAE, GEARS, and scGPT. scGen [17] employs variational autoencoders (VAE) with latent space arithmetic to predict transcriptomic responses generatively. CoupleVAE [18] employs two coupled VAEs that capture complex perturbation-induced state transitions through nonlinear mutual transformations in latent space. CPA [19] integrates covariates describing perturbations and cell types to linearly combine these with cellular embedding states, predicting responses under novel covariate combinations. scPRAM [20] aligns pre- and post-perturbation cell states using optimal transport theory coupled with attention mechanisms. Diverging

from VAE-based approaches, GEARS [16] incorporates prior knowledge through graph neural networks to capture higher-order gene interactions, excelling in complex perturbation scenarios. As a generative pre-trained foundation model in single-cell analysis, scGPT [21] employs stacked Transformer layers to support various downstream applications such as perturbation-response prediction. Critically, all these methods operate exclusively on scRNA-seq data and are constrained by single-modal information bottlenecks, making them inappropriate for the proliferating single-cell multi-omics perturbation-response prediction tasks.

Rapid experimental innovations are accelerating the generation of single-cell multi-omics perturbation datasets [1,22–25]. For instance, Papalexi et al. [23,26] employed ECCITE-seq, an enhanced CRISPR-compatible CITE-seq protocol, to integrate pooled CRISPR screening with simultaneous single-cell mRNA and surface protein measurements, constructing a multimodal perturbation atlas spanning transcriptome and proteome. While transcriptomes capture intermediate gene expression states, proteomes serve as direct functional executors whose abundance and modifications dictate cellular phenotypes [27,28]. Consequently, single-cell multi-omics perturbation data reveal cascade effects of perturbations across gene expression and protein abundance. Integrating these complementary modalities enables precise reconstruction of molecular regulatory networks and uncovers cross-layer dependencies. The expanding landscape of such datasets now creates a pressing need for models natively designed for multimodal perturbation-response prediction [29–31].

Developing predictive models for multi-omics perturbation-response data presents unique computational challenges. First, modality alignment is essential because transcriptome and proteome capture cellular states at distinct molecular layers, and effective cross-modal integration is prerequisite for holistically portraying perturbation responses. Second, modality-specific feature engineering is required due to fundamental differences in data characteristics, demanding tailored computational strategies for each datatype. Furthermore, models must quantify omics-specific perturbation effects—where identical perturbations may elicit divergent transcriptomic versus proteomic responses—and reconstruct multimodal response profiles across biological layers.

To address these challenges, we developed MultiPert, a deep learning framework for predicting perturbation-responses single-cell multi-omics profiles. MultiPert implements differentiated data processing pipelines for transcriptome and proteome, leveraging modality-specific encoders with pretraining to capture distinct biological signals. It integrates prior knowledge-derived perturbation embeddings and during integrative training, achieves cross-modal alignment via adversarial networks and applies perturbations through a dual-attention mechanism. Extensive benchmarking on human THP-1 and kidney datasets demonstrates MultiPert's superior performance in out-of-sample cell prediction and unseen perturbation inference. Through multi-omics joint analysis, MultiPert identifies response patterns of differentially expressed genes, reveals regulatory relationships of immune checkpoint molecules, and functionally enriches associated molecular pathways.

Materials and methods

Overview of the MultiPert model

MultiPert is a deep learning framework specialized for predicting perturbation responses of single-cell multi-omics data. Given control single-cell multi-omics profiles and target perturbations (Fig 1A), it predicts perturbed multi-omics states. MultiPert comprises four stages: preprocessing, pretraining, perturbation embedding extraction, and integrative training. The workflow begins with modality-specific preprocessing to generate normalized multi-omics profiles (Fig 1B). These profiles subsequently initialize modality-specific encoders through pretraining: a zero-inflated negative binomial (ZINB) variational autoencoder for transcriptomics and a standard autoencoder for proteomics, with their encoder parameters transferred to the integrative training phase (Fig 1C). Perturbation embeddings are extracted from a Gene Ontology (GO)-based knowledge graph, providing prior biological knowledge that enables prediction of unseen perturbations (Fig 1D).

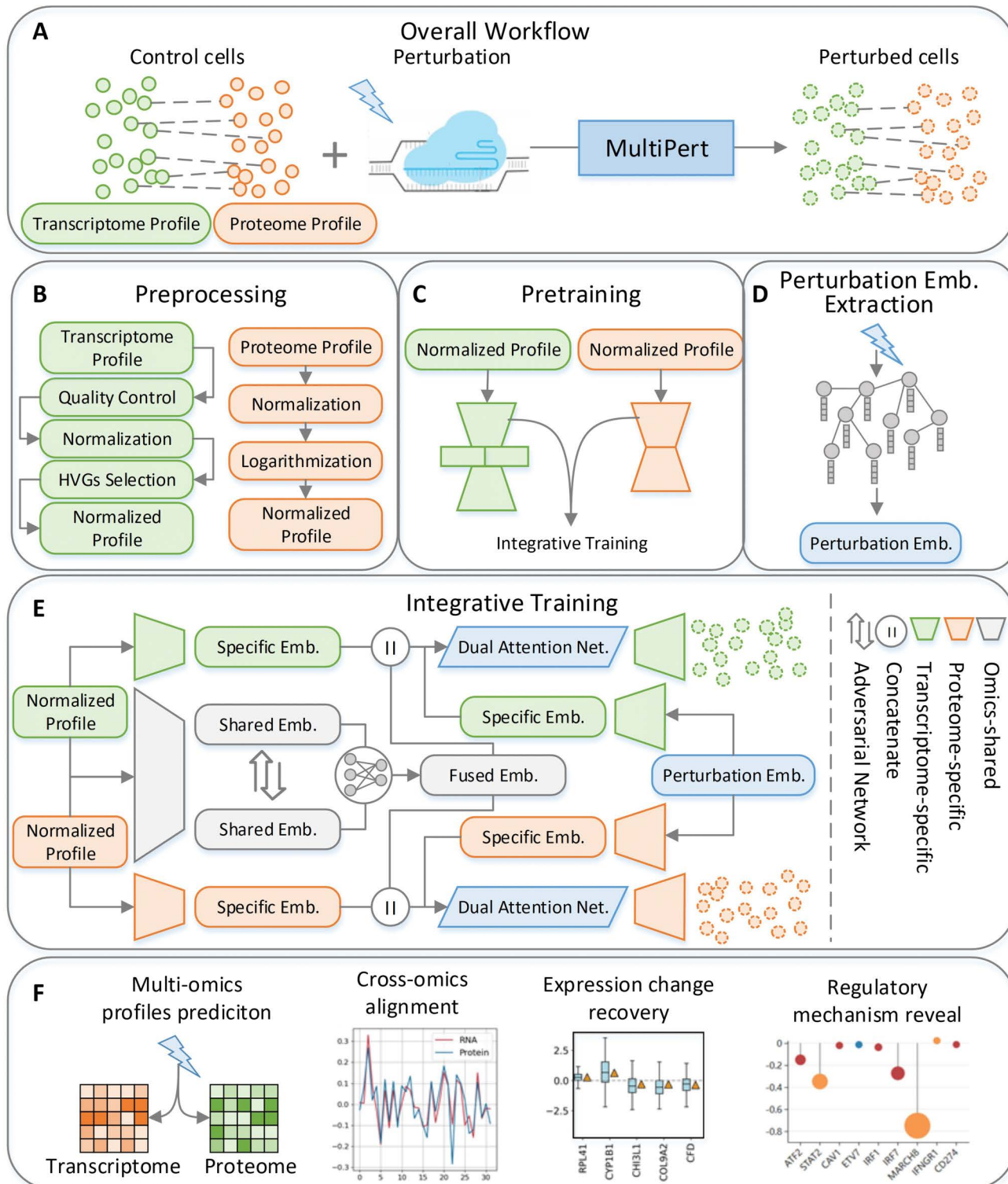


Fig 1. Overview of the MultiPert model. (A) Problem formulation: given control multi-omics profiles and applied perturbation, predict the perturbed multi-omics profiles. Image adapted from Scott and Morris (Wikimedia Commons), licensed under CC BY 4.0. (B) MultiPert preprocesses the data of each modality using the corresponding customary strategies. (C) MultiPert pretrains the encoders in a modality-specific manner. For transcriptome-specific encoder, a Variational Autoencoder (VAE) is employed based on the normalized transcriptomic profiles. For proteome-specific encoder, an

Autoencoder (AE) is employed based on the normalized proteomics profiles. MultiPert then performs integrative training based on the pretrained parameters. **(D)** MultiPert extracts perturbation embeddings from a Gene Ontology (GO)-derived knowledge graph. **(E)** MultiPert performs integrated training using normalized multi-omics profiles and perturbation embeddings as input, mainly comprising modules such as omics-specific encoders, shared encoders, adversarial networks, and dual attention networks. **(F)** MultiPert supports a series of tasks while predicting multi-omics perturbation responses.

<https://doi.org/10.1371/journal.pcbi.1014054.g001>

During integrative training (Fig 1E), normalized multi-omics profiles and perturbation embeddings are processed through core architecture consisting of: modality-specific encoders, a shared encoder, an adversarial network, and modality-specific decoders. To capture distinct biological representations, where transcriptome reflects transcriptional activity and proteome represents functional gene products, modality-specific encoders extract unique cellular features. Concurrently, the shared encoder learns cross-modal information from both omics. Shared embeddings from transcriptome and proteome are aligned via the adversarial network and fused through a multilayer perceptron. Perturbation embeddings are processed by modality-specific encoders to model layer-specific molecular responses. Modality-specific cell embeddings and fused embeddings are concatenated, then integrated with specific perturbation embeddings through a dual attention network to derive perturbed cell states. These states are ultimately decoded by modality-specific decoders to generate the perturbed multi-omics profiles. While predicting multi-omics perturbation responses, MultiPert also supports further analytical tasks, including cross-omics alignment, identifying the perturbation-induced expression changes, and revealing regulatory mechanisms (Fig 1F).

Preprocessing for multi-omics profiles

Raw transcriptomic reads typically exhibit high sparsity and dimensionality, hindering effective learning of cellular expression patterns. We therefore implemented standardized preprocessing pipelines. For transcriptome, genes expressed in fewer than 10 cells were excluded to reduce noise and focus on widely expressed genes. Counts per cell were normalized to 10,000 total reads, subjected to log1p transformation ($\log(\text{count} + 1)$), and the top 5,000 highly variable genes (HVGs) were selected to form normalized profiles. For proteome, cellular counts were normalized to 10,000 total reads followed by identical log1p transformation to generate normalized profiles.

Pretraining of modality-specific autoencoders

Pretraining of modality-specific encoders was conducted to learn latent representations that capture the intrinsic statistical properties of single-cell RNA sequencing (scRNA-seq) and antibody-derived tag (ADT) data, providing a robust initialization for subsequent integrative modeling. This step focused on encoding modality-specific biological variation using control cells, ensuring that baseline patterns—such as gene expression dynamics in RNA and protein abundance relationships in ADT—were preserved.

For scRNA-seq data, which is characterized by sparsity and overdispersion, a zero-inflated negative binomial variational autoencoder (ZINB-VAE) was employed by MultiPert [32]. The encoder maps normalized gene expression profiles $x \in \mathbb{R}_{\geq 0}^{d_{\text{RNA}}}$ (where d_{RNA} is the number of genes) to a latent distribution $q(z|x) = \mathcal{N}(\mu_z, \text{diag}(\sigma_z^2))$, where $\mu_z \in \mathbb{R}^{32}$ and $\sigma_z^2 \in \mathbb{R}_{>0}^{32}$ represent the mean and variance of the 32-dimensional latent space, respectively. Latent samples z are generated via a reparameterization trick as $z = \mu_z + \epsilon \odot \sigma_z$ (with $\epsilon \sim \mathcal{N}(0, I)$ and $\odot$ denotes element-wise multiplication), allowing gradient propagation through the stochastic sampling step during training. The encoder architecture is formulated as follows:

$$h = \text{ReLU}(\text{Linear}(x)) \quad (1)$$

$$\mu_z = \text{Linear}(h) \quad (2)$$

$$\log(\sigma_z^2) = \text{Linear}(h) \quad (3)$$

The decoder of the ZINB-VAE parameterizes the conditional distribution $p(x|z)$ as a zero-inflated negative binomial (ZINB) distribution. This distribution is characterized by three parameters derived from z : mean expression $\mu \in \mathbb{R}_{>0}^{d_{\text{RNA}}}$, dispersion $\theta \in \mathbb{R}_{>0}^{d_{\text{RNA}}}$, and dropout probability $\pi \in [0, 1]^{d_{\text{RNA}}}$. The decoder architecture is formulated as follows:

$$h' = \text{ReLU}(\text{Linear}(z)) \quad (4)$$

$$\mu = \text{Linear}(h') \quad (5)$$

$$\theta = \exp(\text{Linear}(h')) \quad (6)$$

$$\pi = \sigma(\text{Linear}(h')) \quad (7)$$

where $\sigma(\cdot)$ denotes the sigmoid function. The *Linear* function represents a single-layer linear network, but the input and output dimensions differ in [Eqs 1–7](#).

The ZINB-VAE is trained by minimizing a composite loss function that combines the negative log-likelihood of the data under the ZINB model and the Kullback-Leibler (KL) divergence. Formally, this loss is:

$$\mathcal{L}_{\text{RNA}} = \mathbb{E}_{q(z|x)} [-\log p(x|z)] + \text{KL}(q(z|x)||p(z)) \quad (8)$$

$$p(x_i|z) = \begin{cases} \pi_i + (1 - \pi_i) p_{\text{NB}}(0|\mu_i, \theta_i), & x_i = 0 \\ (1 - \pi_i) p_{\text{NB}}(x_i|\mu_i, \theta_i), & \text{otherwise} \end{cases} \quad (9)$$

where p_{NB} is the probability mass function of the negative binomial distribution parameterized by mean μ_i and dispersion θ_i .

For ADT data, a typical autoencoder was used. This model includes an encoder mapping protein abundance $y \in \mathbb{R}_{\geq 0}^{d_{\text{ADT}}}$ (where d_{ADT} is the number of proteins) to a 32-dimensional latent space $z \in \mathbb{R}^{32}$, and a decoder reconstructing y from z . Training minimized the mean squared error (MSE) between input and reconstruction:

$$\mathcal{L}_{\text{ADT}} = \frac{1}{d_{\text{ADT}}} \|y - \hat{y}\|_2^2 \quad (10)$$

where $\hat{y}$ denotes the reconstructed protein abundance. All models were trained using the Adam optimizer for 10 epochs, with pretrained weights saved to initialize modality-specific encoders in subsequent integrative training. This step ensured that modality-specific biological signals were preserved in the latent representations, providing a stable foundation for cross-omics integration.

Perturbation embedding extraction

MultiPert constructs unique perturbation embeddings from the Gene Ontology (GO) knowledge graph [\[33\]](#). This approach associates discrete gene perturbations with the system functional network, helping the model understand the biological

significance of perturbations. Additionally, the topological structure of graph supports the transfer of known perturbations to unknown ones through functional similarity, enhancing the model's prediction ability for unseen perturbations.

A bipartite graph connecting genes to GO terms was first established. Gene-gene functional similarity was computed using the Jaccard index of shared GO terms, with top-K ($K=20$) most similar genes forming edges in the perturbation graph G_{pert} . Learnable embeddings initialized for all perturbations were refined via a graph neural network (GNN) that aggregated neighbor representations. Given GEARS' prior validation of GO knowledge graph for perturbation prediction, we directly adopted their pre-trained embeddings to initialize MultiPert's perturbation representations [16].

Integrative training

The integrative training aims to combine modality-specific biological patterns, cross-modal shared features, and perturbation information to predict perturbation-induced changes in RNA and ADT expression. This process integrated multiple components to model both baseline cellular states and the effects of perturbations.

The training system comprised two core components: a generator and a discriminator. The generator processes normalized multi-omics profiles through a hierarchical encoder-decoder architecture. Modality-specific encoders which are initialized from pretrained weights, extract unique biological signatures from transcriptomic and proteomic inputs, generating specific cellular embeddings (z_c^t, z_c^p). Simultaneously, a shared encoder projects both modalities into a common latent space to enable cross-modal alignment (z_d^t, z_d^p). c and d represent cell embeddings generated by the modality-specific encoder and shared encoder, respectively. z_d^t and z_d^p are fused into a unified embedding (z_f) via a fusion network, formulated as:

$$z_f = W_2 \cdot \text{ReLU}(W_1 \cdot [z_d^t | z_d^p] + b_1) + b_2 \quad (11)$$

where W_1 and W_2 are weight matrices, b_1 and b_2 are bias vectors, and $[|]$ denotes concatenation. Separately, perturbation embeddings (p) are transformed into modality-adapted forms through dedicated encoders (p_c^t, p_c^p). Then, a dual-attention mechanism integrates cell embeddings and perturbation embeddings to derive the perturbed cellular states.

The dual-attention mechanism integrates cell embeddings and perturbation embeddings through two sequential attention steps. First, cross-attention integrates information from them using multi-head attention and produces an intermediate representation $attn$, where *MultiHeadAttn* splits the inputs into 4 heads, computes scaled dot-product attention for each, and concatenates the results along the feature dimension. Second, channel attention learns adaptive weights for each feature dimension by aggregating global statistics via average pooling. The above process is formalized as follows:

$$attn = \text{MultiHeadAttn}(Query, Key, Value) \quad (12)$$

$$w = \sigma(W_4 \cdot \text{ReLU}(W_3 \cdot \text{avgpool}(attn))) \quad (13)$$

$$z = attn \odot w \quad (14)$$

where $\odot$ denotes element-wise multiplication. For transcriptome, $Query = [z_c^t | z_f]$, Key and $Value$ refer to p_c^t , z in Formula 14 represents z^t . For proteome, $Query = [z_c^p | z_f]$, Key and $Value$ refer to p_c^p , z in Formula 14 represents z^p . Finally, modality-specific decoders reconstruct perturbation-induced gene expression from the perturbed cellular states corresponding to transcriptome (z^t), and reconstruct perturbation-induced protein abundance from the perturbed cellular states corresponding to proteome (z^p). Among them, the shared encoder, perturbation encoder, and decoder are all composed of a two-layer linear network with ReLU activation. The discriminator is a network tasked with distinguishing the modality origin of shared

features, encouraging the generator to learn modality-agnostic representations and align modalities. It took shared features as input and output a probability score via:

$$D(z_d) = \sigma(W_2^{disc} \cdot (W_1^{disc} \cdot z_d + b_1^{disc}) + b_2^{disc}) \quad (15)$$

Training proceeded in epochs, with alternating updates to the generator and discriminator, followed by validation to monitor performance. For each batch of control data (x^t , x^p), perturbed data (y^t , y^p), and perturbation embeddings (p), the generator processed inputs through its components to produce $\hat{y}^t$, $\hat{y}^p$.

Loss calculation combined reconstruction and adversarial components to balance predictive accuracy and cross-modal alignment. The reconstruction loss is computed as the sum of mean squared errors (MSE) between predicted and observed perturbed multi-omics profiles:

$$\mathcal{L}_{recon} = \|\hat{y}^t - y^t\|_2^2 + \|\hat{y}^p - y^p\|_2^2 \quad (16)$$

The discriminator is trained by binary cross-entropy (BCE) to classify shared features as transcriptome-derived (label = 1) or proteome-derived (label = 0):

$$\mathcal{L}_{disc} = \text{BCE}(D(z_d^t), 1) + \text{BCE}(D(z_d^p), 0) \quad (17)$$

Conversely, the generator is trained to confuse the discriminator, with loss:

$$\mathcal{L}_{conf} = \text{BCE}(D(z_d^t), 0.5) + \text{BCE}(D(z_d^p), 0.5) \quad (18)$$

The total generator loss combines these components:

$$\mathcal{L}_{gen} = \mathcal{L}_{recon} + \mathcal{L}_{conf} \quad (19)$$

where $\mathcal{L}_{recon}$ concentrates on enhancing the model's ability for recovering the perturbed multi-omics profiles, and $\mathcal{L}_{conf}$ emphasizes multi-modality alignment and multi-omics data integration.

The adversarial training employed a two-phase update cycle to harmonize discriminator-generator dynamics. During the discriminator phase, all generator parameters were frozen while the discriminator minimized its objective ($\mathcal{L}_{disc}$), refining its capacity to distinguish the data source of shared embeddings. Subsequently, the generator phase froze discriminator parameters and minimized the composite loss ($\mathcal{L}_{gen}$). The discriminator and generator adopt asymmetric Adam optimization, and set the learning rate to 0.002 and 0.001 respectively. The training ran for up to 1000 epochs, with an early stopping mechanism based on validation loss. The model is implemented in PyTorch 2.0, running on Python 3.9 with CUDA 11.7 on an NVIDIA A100 GPU for training and evaluation.

Datasets

To comprehensively evaluate the multi-omics perturbation-response prediction performance of MultiPert in different biological contexts, we collected gene perturbation datasets obtained from different tissues and techniques. The THP-1 dataset was generated by Papalexi et al. [23] using ECCITE-seq technology in the acute myeloid leukemia (AML) cell line THP-1, providing simultaneous transcriptomic and proteomic measurements. The dataset comprises 8,984 cells, profiling 16,826 genes and 4 proteins, and includes 10 single-gene knockout perturbation experiments. The perturbation conditions containing less than 10 cells were removed, so the following evaluations involved 9 perturbations. Furthermore, the kidney dataset was generated via CaRPool-seq technology [34], which co-profiles transcriptomic and proteomic responses to 7

perturbations in human epithelial embryonic kidney cells. This dataset comprises 8,802 cells, profiling 20,639 genes and 7 proteins. The transcriptome–epigenome dataset was generated by 10x Genomics Multiome and an epigenome–proteome dataset was generated by ASAP-seq technology, where perturbation was induced by 16-hour IL2 and anti-CD3/CD28 incubation [35].

Evaluation metrics

To comprehensively evaluate the performance of MultiPert for single-cell multi-omics perturbation-response prediction, we employed complementary quantitative metrics. The Mean Squared Error (MSE) measures the average squared difference between predicted and ground truth expression values, providing a fundamental assessment of reconstruction fidelity:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (20)$$

where y_i represents the true expression value, $\hat{y}_i$ denotes the predicted value, and n is the number of cells. The Pearson Correlation Coefficient (PCC) quantifies the linear relationship between predicted and true expression profiles across cell populations:

$$PCC = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (21)$$

Results

Performance comparison with state-of-the-art methods

Owing to the absence of dedicated multi-omics methods, we benchmarked MultiPert against representative scRNA-seq-based perturbation prediction models, including the VAE-based model (scGen [17]), compositional perturbation autoencoder-based model (CPA [19]), optimal transport-based model (scPRAM [20]), and pretrained large-scale single-cell model (scGPT [21]). For each perturbation in the THP-1 dataset, all cells were randomly partitioned into training, validation, and test sets at a 3:1:1 ratio. For baseline methods dispensing with validation sets, cells were randomly split into training and test sets at a 4:1 ratio. Results shown in figures are derived from model predictions on test sets. Since baseline methods lack multi-omics data integration capabilities, transcriptomic and proteomic data were processed as separate datasets to generate corresponding predictions. The scGPT framework is restricted to perturbation inference based on scRNA-seq data and thus cannot model proteomic perturbation experiments.

For transcriptome, Fig 2A (first row) displays MSE and PCC values between predicted and true expression across all genes. CoupleVAE, CPA, and MultiPert demonstrate superior performance compared to other methods. Specifically, MultiPert attains the lowest median MSE (0.09), and concurrently achieves the highest median PCC (0.78). Given that differentially expressed genes (DEGs) serve as primary mediators of a perturbation's biological impact, their accurate prediction determines a model's capacity to capture critical expression changes [36,37]. We thus leveraged Wilcoxon rank-sum tests in Scanpy [38] to identify DEGs under distinct perturbations and spotlighted the methods' performance for the top 50 DEGs (Fig 2A second row). Compared to all-gene analyses, all methods exhibit increased MSE values which confirms heightened sensitivity of DEGs to perturbations relative to other genes. In terms of PCC, both scGPT and MultiPert demonstrate greater stability than other methods, though scGPT yields a low median PCC value. Against benchmarks, MultiPert achieves the lowest median MSE (0.37) and highest median PCC (0.88), representing 7% and 2% improvements over

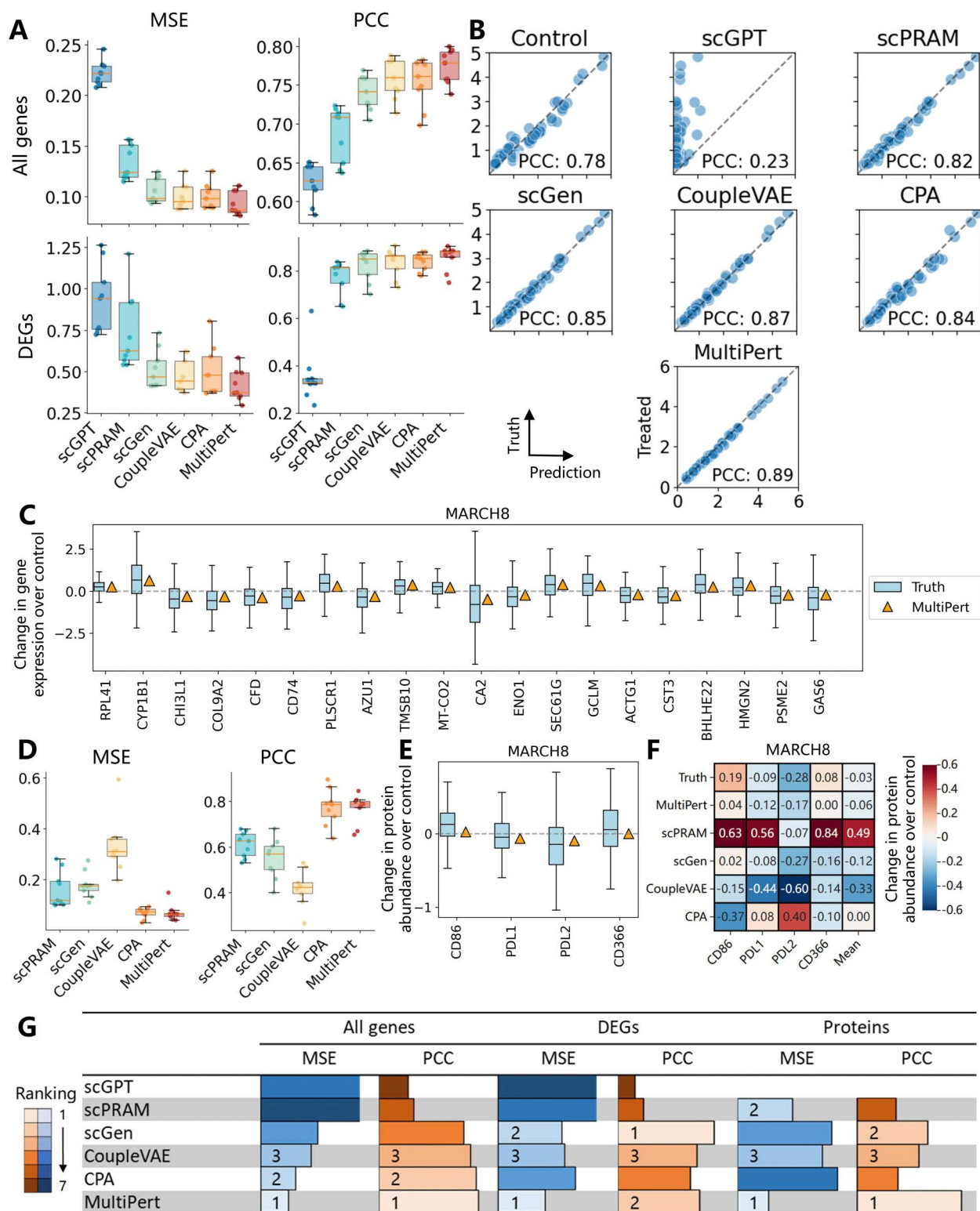


Fig 2. Comprehensive evaluation for single-cell perturbation prediction. (A) Box plots of MSE and PCC values for all genes and top 50 DEGs. Scatters represent different perturbations. (B) Comparison of scatter plots for predicted and true gene expression values under MARCH8 perturbation. Each point represents the average expression level of a gene across all cells. The horizontal axis shows the prediction results of different methods, and

the vertical axis represents the true average expression level of genes. **(C)** Analysis of gene expression changes following perturbation of MARCH8. The plot compare predicted versus true expression changes for the top 20 DEGs. The horizontal dashed line indicates the null effect baseline. **(D)** Box plots of MSE and PCC values for all proteins. Scatters represent different perturbations. **(E)** Analysis of protein abundance changes following perturbation of MARCH8. **(F)** The heatmap of change in protein abundance over control. Red and blue indicate MARCH8 perturbation activates or suppresses protein expression, respectively. Each value represents the difference between the predicted or observed perturbed protein abundance and the protein abundance in control samples. Mean represents the average value of all proteins. **(G)** Performance of methods on kidney dataset. The top three methods for each metric are highlighted. Color indicates ranking, and bar length represents the magnitude of the metric.

<https://doi.org/10.1371/journal.pcbi.1014054.g002>

the runners-up. Collectively, these results indicate that MultiPert not only accurately reconstructs global gene expression patterns but also effectively captures perturbation-induced differential expression features. While overall performance was benchmarked across perturbations, we further dissected method behavior under identical perturbations. Illustrating the MARCH8 perturbation (gene knockout of MARCH8), [Fig 2B](#) visualizes expression changes in DEGs by jointly plotting control and perturbed cells. MultiPert achieves alignment between predictions and true data along the diagonal (PCC=0.89), representing an 11% improvement over control samples and outperforming all comparators. Subsequently, [Fig 2C](#) visualizes directional expression shifts for the top 20 DEGs, confirming that MultiPert accurately predicts perturbation-induced directions in gene expression. The results corresponding to other perturbations are shown in [S1](#) and [S2 Figs](#).

For proteome, MultiPert achieves optimal performance with the lowest median MSE (0.07) and highest median PCC (0.79) among all methods ([Fig 2D](#)). It also demonstrated superior stability, exhibiting the narrowest interquartile ranges (IQR: 0.02 for MSE, 0.08 for PCC). Full IQR comparisons across all benchmarks are detailed in [S3 Fig](#). Specific to individual perturbations, [Fig 2E](#) and [S4 Fig](#) illustrate the ability of MultiPert to precisely reveal the direction of protein abundance changes. [Fig 2F](#) further compares the ability of MultiPert and other methods to reflect changes in protein abundance. While scGen and MultiPert yield predictions closest to the truth, scGen incorrectly predicts the upregulation of CD366 as a downregulation. When considering the average effect across all proteins, MultiPert correctly reflects the overall downregulation trend observed in the ground truth, with a magnitude that most closely matches the experimental measurements. This indicates that MultiPert not only recovers protein-level trends more accurately but also outperforms other methods in overall evaluation.

To evaluate MultiPert across tissues and technologies, we benchmarked it on the kidney dataset generated via CarPool-seq technology [\[34\]](#). Mean performance metrics are summarized in [Fig 2G](#) (detailed values: [S5 Fig](#)). MultiPert achieved top performance in all metrics except PCC in DEGs, with a proteomic PCC of 0.98. We further benchmarked MultiPert on two additional human peripheral blood mononuclear cells (PBMC) multi-omics datasets from Mimitou et al. [\[35\]](#), including a transcriptome–epigenome dataset generated with 10x Genomics Multiome and an epigenome–proteome dataset generated with ASAP-seq technology, where perturbation was induced by 16-hour IL2 and anti-CD3/CD28 incubation. The results are summarized in [S1](#) and [S2 Tables](#). These results demonstrate MultiPert’s ability to integrate multi-omics data and its broad applicability across tissues in perturbation prediction.

Interpretability analysis of MultiPert’s modular components

We assessed the efficacy of MultiPert’s adversarial network, specific encoders, and shared encoder using the THP-1 dataset. [Fig 3A](#) (first panel) visualizes baseline UMAP embeddings of raw expression profiles, where transcriptomic and proteomic cells showed distinct spatial distributions with limited overlap. The second panel presents omics-specific embeddings from MultiPert, revealing mutually exclusive clustering that verifies the encoders’ capacity to isolate molecular-layer signatures. The third panel depicts omics-shared embeddings, demonstrating an intermingled distribution that confirms the adversarial network and shared encoder successfully aligned cross-modality features for multi-omics data integration.

We further analyzed training dynamics of omics-shared embeddings ([Fig 3B](#)). Population-level embeddings—averaged across all transcriptomic and proteomic cells—revealed pronounced modality divergence early in training. As epochs

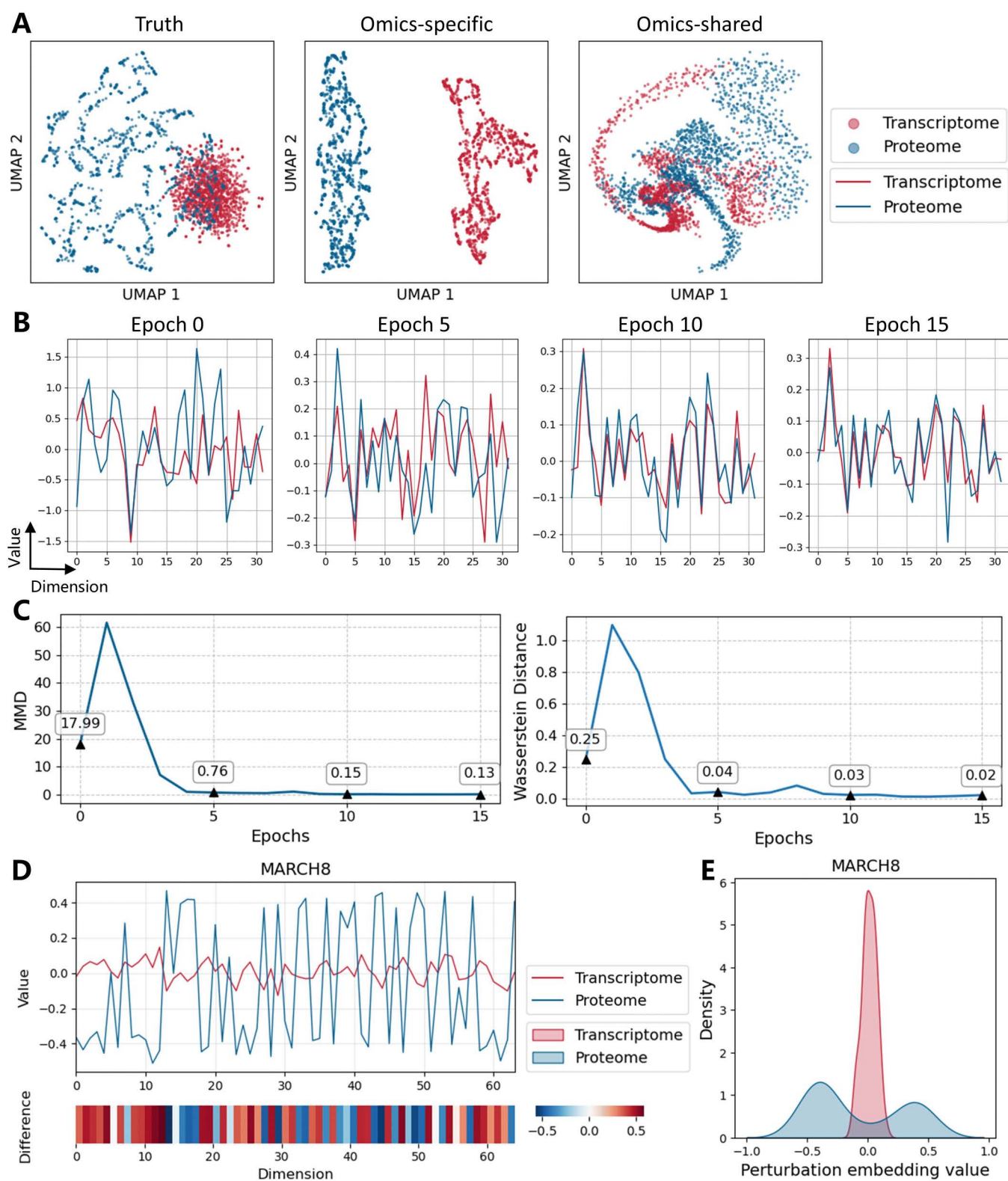


Fig 3. Interpretability analysis of omics embedding and perturbation embedding in MultiPert. (A) Scatter plots of cell from transcriptome and proteome. The subgraphs are obtained via the UMAP algorithm based on true expression data, omics-specific embedding, and omics-shared embedding,

respectively. Each point corresponds to a cell. **(B)** Visualization of omics-specific embedding under epochs 0, 5, 10 and 15. The red and blue lines represent transcriptomic-specific embedding and proteomic-specific embedding, respectively. **(C)** Line graphs of MMD and Wasserstein distance values along epochs. The specific values corresponding to epochs 0, 5, 10, and 15 are marked above the line. **(D)** Visualization of omics-specific perturbation embedding under MARCH8 perturbation. The heat map displays the differences in transcriptome-specific and proteome-specific perturbation embedding in each dimension. **(E)** The kernel density estimation plot of omics-specific perturbation embedding. The x-axis denotes the one-dimensional perturbation embedding value, and the y-axis indicates probability density.

<https://doi.org/10.1371/journal.pcbi.1014054.g003>

progressed, the embeddings progressively converged, achieving alignment by epoch 15. Both Maximum Mean Discrepancy (MMD) [39] and Wasserstein distance [40] are commonly used metrics to measure the difference in distributions. MMD employs a linear kernel function to compute the mean distance between two probability distributions in the reproducing kernel Hilbert space, while Wasserstein distance quantifies the minimum cost required to transform one distribution to another. Quantitative validation in Fig 3C used MMD and Wasserstein Distance, both exhibiting rapid initial decline followed by a plateau after epoch 5. The rapid decrease of both MMD and Wasserstein distance in early epochs reflects the initial strong distributional discrepancy between transcriptomic and proteomic embeddings, which is quickly reduced by the shared encoder under adversarial alignment. As training proceeds, the embeddings from different modalities become progressively aligned in the shared latent space, resulting in diminishing improvements and a plateau in both metrics. Such a trajectory clearly captures cross-modality feature alignment during multi-omics data integration, enabled by MultiPert's shared encoder and adversarial network architecture.

MultiPert generates omics-specific perturbation embeddings via specific encoders, addressing potential cross-omics differential perturbation responses. Using the MARCH8 perturbation as a representative case, we extracted transcriptome- and proteome-specific perturbation embeddings for test cells, averaging population-level embeddings for visualization (Fig 3D and 3E). Line plots reveal dimension-wise divergences between embeddings, which are further emphasized by heat-map comparisons. Kernel Density Estimation (KDE) curves highlight global distributional differences. The x-axis values correspond to the one-dimensional omics-specific perturbation embeddings, which are real-valued and can take both positive and negative values. These results demonstrate that specific encoders map the same perturbation to distinct latent spaces, optimizing their fusion with transcriptomic or proteomic states.

Ablation experiments of MultiPert's modules

The mechanism for combining perturbation with cell embeddings in latent space is essential for inferring cellular response states. We investigated four combining strategies: addition (MultiPert-A), multiplication (MultiPert-M), concatenation (MultiPert-C), and a cross-attention mechanism (MultiPert-T; cell embeddings as Query, perturbation embeddings as Key/Value). On the THP-1 dataset, MultiPert consistently surpassed all variants in transcriptomic predictions (Fig 4A and 4B), achieving lowest MSE and highest PCC across all perturbations. MultiPert-T ranked second, followed by MultiPert-A, -M, and -C—indicating that additive, multiplicative, and concatenative operations fail to capture complex perturbation responses. Proteomic predictions validated these findings with concordant results (Fig 4C).

MultiPert incorporates two pivotal architectural innovations: an adversarial network for cross-omics feature alignment and specific encoders generating omics-specific perturbation embeddings. To evaluate their contributions, we created ablation variants—MultiPert-WA (adversarial network removed) and MultiPert-WS (specific encoders replaced with a shared encoder)—and benchmarked them on the THP-1 dataset. For transcriptomic predictions (Fig 4D and 4E), MultiPert-WA outperformed MultiPert-WS, indicating that specific encoders better capture post-perturbation gene expression dynamics. Conversely, in proteomics (Fig 4F), MultiPert-WS surpassed MultiPert-WA, revealing the adversarial network's critical role in optimizing protein abundance modeling. Crucially, the full MultiPert architecture achieved superior MSE and PCC metrics over both variants, collectively demonstrating the synergistic necessity of co-designing these components for robust multi-omics perturbation prediction.

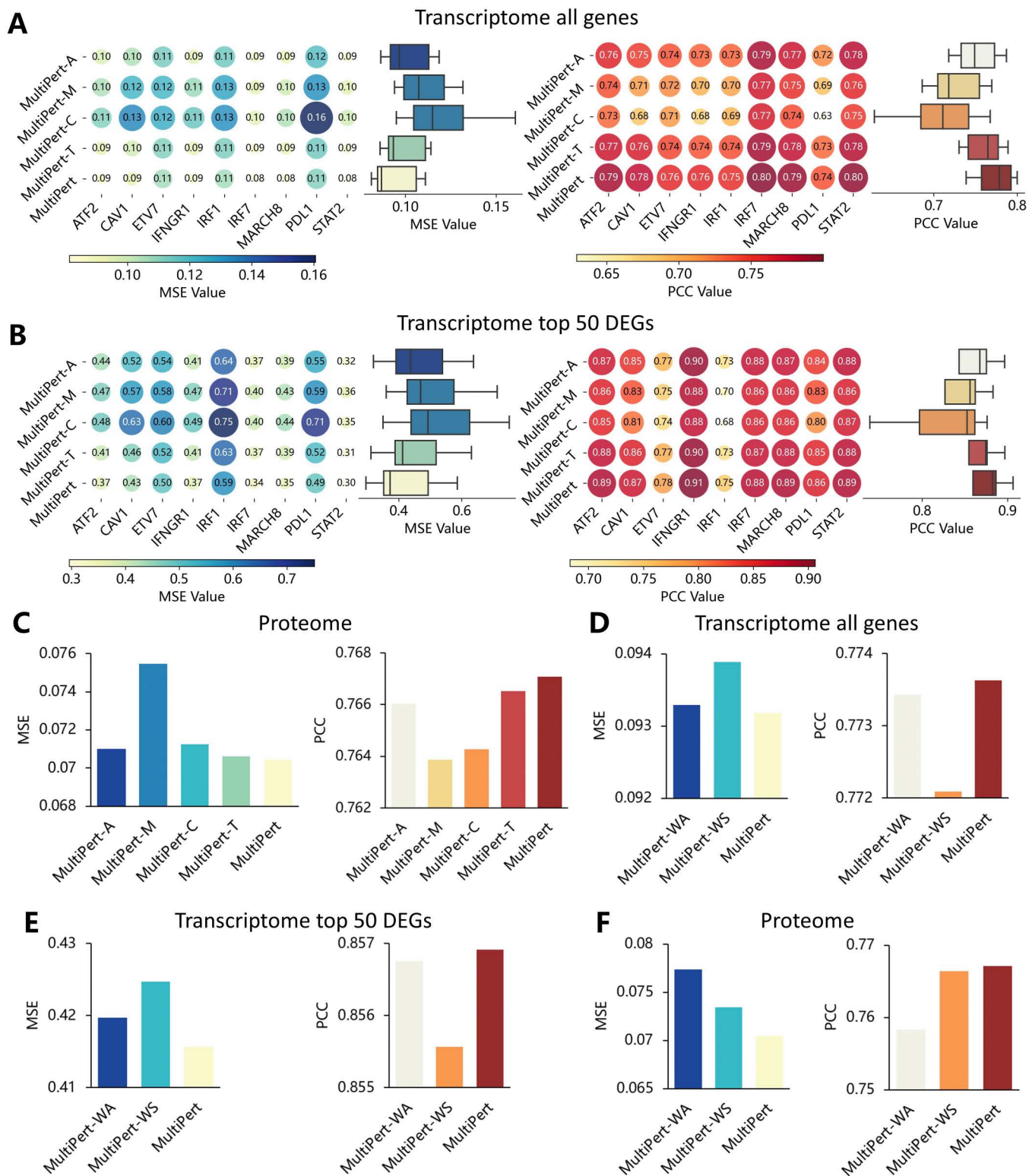


Fig 4. Performance comparison of MultiPert and multiple variants. (A, B) The MSE and PCC values of MultiPert and its combination variants obtained on transcriptome for all genes and top 50 DEGs, respectively. The heat map shows the specific metrics for each framework in each perturbation, while the box plot shows the overall metric distribution for all perturbations. (C) The MSE and PCC values of MultiPert and its combination variants

obtained on proteome for all proteins. Bar plot is drawn based on the average value across all perturbations. **(D, E)** The MSE and PCC values of MultiPert and its module variants obtained on transcriptome for all genes and top 50 DEGs, respectively. **(F)** The MSE and PCC values of MultiPert and its module variants obtained on proteome for all proteins.

<https://doi.org/10.1371/journal.pcbi.1014054.g004>

Generalization of MultiPert to unseen perturbation prediction

Existing single-cell methods excel at predicting perturbation responses for cells out of sample, lacking generalizability to unseen perturbations. Learning the relationship between cellular responses and perturbations from limited datasets, then extrapolating to novel perturbation scenarios, is critical for drug development and therapeutic target discovery. To evaluate this capability, we employed a leave-one-out cross-validation strategy across nine perturbations in the THP-1 dataset, with GEARS—a method specifically designed for novel perturbation prediction—as the baseline.

As shown in [Fig 5A](#) and [5B](#), MultiPert consistently achieved lower MSE and higher PCC than GEARS across all perturbations, attributable to its integration of both transcriptomic and proteomic data. Since GEARS does not support proteomic perturbation prediction, [Fig 5C](#) demonstrates the PCC values of MultiPert on proteome, which are comparable to its transcriptomic performance. These results establish MultiPert as a robust framework for multi-omics perturbation prediction while enabling reliable inference on unseen perturbations.

Uncovering regulatory mechanisms by MultiPert

As an inhibitory immune checkpoint molecule, PDL1 plays a pivotal role in balancing immune activation and suppression. Investigating its interactions is critical for understanding cancer immune evasion and developing immunotherapies [\[23,41,42\]](#). Leveraging MultiPert's predictions, [Fig 5D](#) visualizes perturbation-induced changes in PDL1 protein abundance relative to control cells. These changes reveal specific gene regulatory effects on PDL1 expression. Notably, MultiPert predicts that ATF2 knockout would lead to decreased PDL1 abundance, implying that ATF2 may act as a positive regulator of PDL1. This prediction is consistent with previous findings demonstrating ATF2's direct binding to the PDL1 promoter [\[43\]](#). Similarly, MultiPert predicts a reduction in PDL1 levels upon IRF7 knockout, corroborating its known role as a transcriptional activator [\[44\]](#). Among nine single-gene knockout perturbations, predicted regulatory relationships between five genes with PDL1 were validated by literatures, while three genes were reported to associate with PDL1 expression [\[45–49\]](#). These results demonstrate MultiPert's capability to effectively uncover regulatory interactions, thereby facilitating mechanistic studies of immune checkpoint interactions.

Among all perturbations, MARCH8 knockout induced the most pronounced change in PDL1 abundance. For transcriptomic predictions under MARCH8 perturbation, [Fig 5F](#) compares expression changes of the top 20 DEGs across ground truth, GEARS, and MultiPert predictions. MultiPert more accurately captured true expression changes than GEARS, particularly for genes including CYP1B1, CFD, CA2, and GCLM. Subsequent GO enrichment analysis of DEGs ($Qvalue < 0.01$) revealed the top 10 significantly enriched terms ([Fig 5E](#)), all associated with immune response, inflammatory regulation, and cellular defense mechanisms—providing insights into pathogenic mechanisms of acute myeloid leukemia and broader immunobiological processes.

Discussion

In this study, we developed MultiPert, the pioneer deep learning framework specifically designed for predicting single-cell multi-omics perturbation responses. Unlike existing methods that are limited to scRNA-seq profiles, MultiPert integrates transcriptomic and proteomic measurements, thereby overcoming the single-modality bottleneck that constrains current perturbation-response prediction. Through modality-specific encoders, adversarial alignment, and a dual-attention perturbation mechanism, MultiPert achieves both high accuracy and strong generalizability.

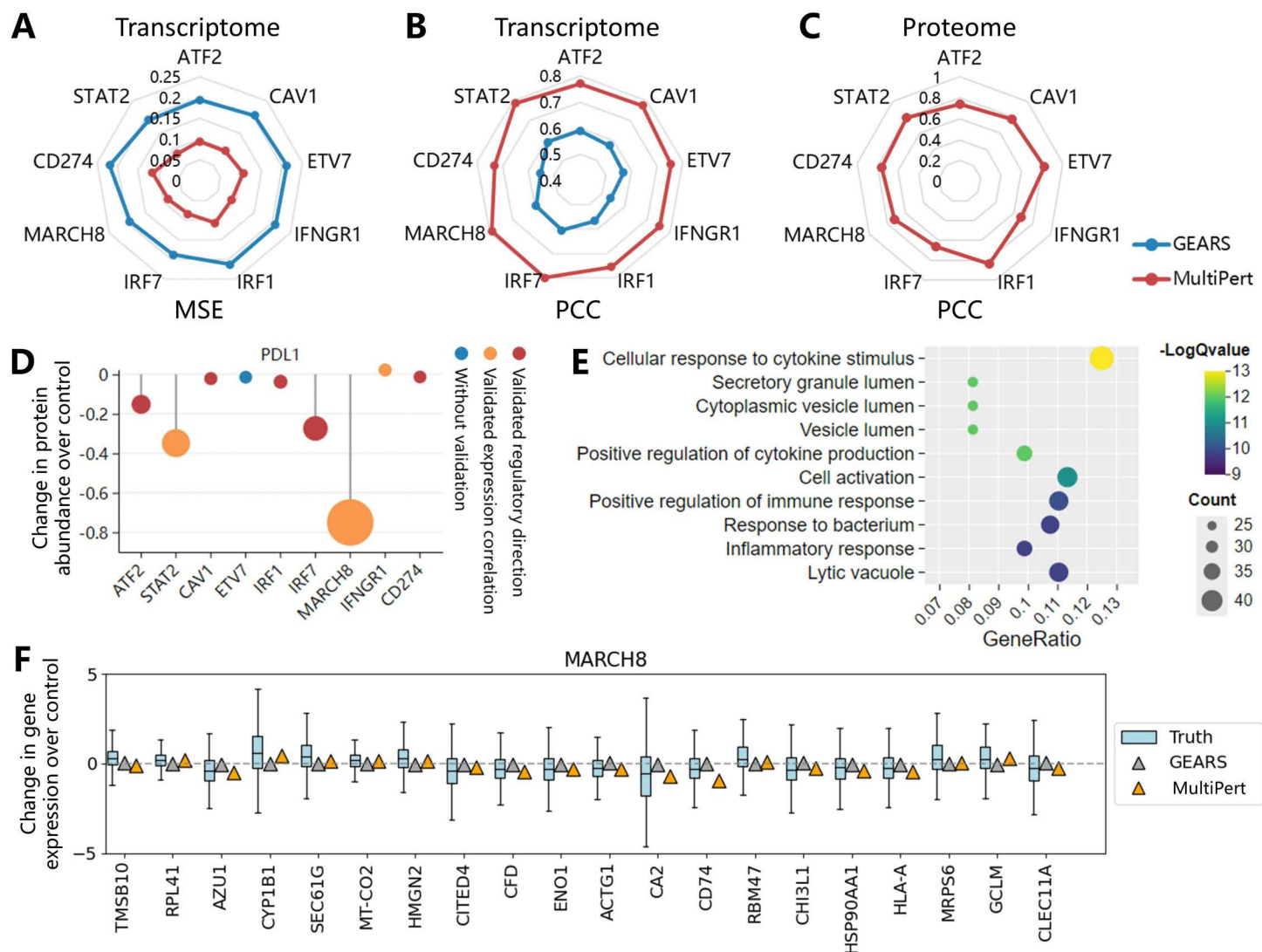


Fig 5. Perturbation analysis combined with multi-omics data reveals the biology mechanisms. (A-C) MSE and PCC values obtained by GEARs and MultiPert corresponding to transcriptome, and PCC values corresponding to proteome. (D) The lollipop plot of MultiPert-predicted PDL1 abundance changes following different perturbations. The direction of lollipops reflects the regulatory effect (enhance or inhibit) of perturbations on PDL1 abundance. Red and orange indicate that the regulatory effect or expression correlation has been verified by literatures. Blue indicates no literature support. The radius of dots is positively correlated with the predicted protein abundance changes. (E) GO enrichment results of DEGs ($Q_{value} < 0.01$) under MARCH8 perturbation. GeneRatio = Count/N, where Count is the number of genes belonging to GO term in DEGs, and N is the number of DEGs. The color of bubbles represents the significance of GO term, and the size represents the number of genes enriched in GO item. (F) Analysis of gene expression changes following perturbation of MARCH8. The plot compares the prediction from GEARs and MultiPert versus true gene expression for the top 20 DEGs. The horizontal dashed line indicates the null effect baseline.

<https://doi.org/10.1371/journal.pcbi.1014054.g005>

The benchmarking across THP-1 and kidney datasets highlights several key advantages of MultiPert. First, it consistently outperforms leading methods, which demonstrates its ability to jointly model cross-layer perturbation effects. Second, MultiPert can generalize to unseen perturbations, a capability essential for therapeutic discovery where experimental screening is inherently incomplete. Third, the framework provides biological interpretability, with analysis of gene and protein level responses revealing regulatory mechanisms of immune checkpoint molecules consistent with existing

experimental evidence. Such interpretability not only validates the model but also enables generation of novel biological hypotheses. MultiPert's architecture further reveals the synergistic role of its components. The adversarial network ensures effective modality alignment, while modality-specific encoders capture unique biological signatures that cannot be represented by a shared encoder alone. Ablation experiments confirmed that the integration of these modules is critical for robust multi-omics perturbation modeling.

Looking forward, extending MultiPert to incorporate additional modalities, such as epigenomics or spatial transcriptomics, will further enhance its ability to fully capture cellular states. Moreover, the modular design of MultiPert makes it naturally compatible with emerging foundation models and large-scale pre-trained representations in computational biology, which could be leveraged to augment modality-specific encoders or represent perturbation signals. As multi-omics datasets continue to expand, MultiPert provides a methodological foundation for advancing systems-level understanding of perturbation biology and accelerating discovery in precision medicine.

Supporting information

S1 Fig. Comparison of gene expression changes direction between true and MultiPert predictions following perturbation of ATF2, STAT2, CAV1 and ETV7.

(PDF)

S2 Fig. Comparison of gene expression changes direction between true and MultiPert predictions following perturbation of IRF1, IRF7, IFNGR1 and CD274.

(PDF)

S3 Fig. Comparison of IQR for all methods on proteomics prediction. (A) IQR results calculated based on MSE metric. (B) IQR results calculated based on PCC metric.

(PDF)

S4 Fig. Comparison of protein abundance changes direction between true and MultiPert predictions following different perturbations.

(PDF)

S5 Fig. Summary of metrics for methods on the kidney dataset. Color indicates ranking.

(PDF)

S1 Table. The performance of different methods on epigenome–proteome dataset.

(XLSX)

S2 Table. The performance of different methods on transcriptome–epigenome dataset.

(XLSX)

Code availability

The MultiPert algorithm is implemented in Python and is available on GitHub: <https://github.com/MengyuanZhao/MultiPert>.

Author contributions

Conceptualization: Mengyuan Zhao.

Data curation: Mengyuan Zhao, Xinyue Tang.

Formal analysis: Mengyuan Zhao.

Funding acquisition: Jijun Tang, Fei Guo.

Methodology: Mengyuan Zhao, Jiawei Li.

Software: Mengyuan Zhao.

Supervision: Fei Guo.

Validation: Mengyuan Zhao, Xinyue Tang.

Visualization: Mengyuan Zhao.

Writing – original draft: Mengyuan Zhao.

Writing – review & editing: Mengyuan Zhao, Cheng Liang, Jijun Tang, Fei Guo.

References

1. Peidli S, Green TD, Shen C, Gross T, Min J, Garda S, et al. scPerturb: harmonized single-cell perturbation data. *Nat Methods*. 2024;21(3):531–40. <https://doi.org/10.1038/s41592-023-02144-y> PMID: [38279009](#)
2. Ji Y, Lotfollahi M, Wolf FA, Theis FJ. Machine learning for perturbational single-cell omics. *Cell Syst*. 2021;12(6):522–37. <https://doi.org/10.1016/j.cels.2021.05.016> PMID: [34139164](#)
3. Liu B, Jing Z, Zhang X, Chen Y, Mao S, Kaundal R, et al. Large-scale multiplexed mosaic CRISPR perturbation in the whole organism. *Cell*. 2022;185(16):3008–3024.e16. <https://doi.org/10.1016/j.cell.2022.06.039> PMID: [35870449](#)
4. Song B, Liu D, Dai W, McMyn NF, Wang Q, Yang D, et al. Decoding heterogeneous single-cell perturbation responses. *Nat Cell Biol*. 2025;27(3):493–504. <https://doi.org/10.1038/s41556-025-01626-9> PMID: [40011559](#)
5. Monfort-Lanzas P, Rungger K, Madersbacher L, Hackl H. Machine learning to dissect perturbations in complex cellular systems. *Comput Struct Biotechnol J*. 2025;27:832–42. <https://doi.org/10.1016/j.csbj.2025.02.028> PMID: [40103613](#)
6. Chandrasekaran SN, Alix E, Arevalo J, Borowa A, Byrne PJ, Charles WG, et al. Morphological map of under- and overexpression of genes in human cells. *Nat Methods*. 2025;22(8):1742–52. <https://doi.org/10.1038/s41592-025-02753-9> PMID: [40775081](#)
7. Pacesa M, Pelea O, Jinek M. Past, present, and future of CRISPR genome editing technologies. *Cell*. 2024;187(5):1076–100. <https://doi.org/10.1016/j.cell.2024.01.042> PMID: [38428389](#)
8. Torres-Ruiz R, Rodriguez-Perales S. CRISPR-Cas9 technology: applications and human disease modelling. *Brief Funct Genomics*. 2017;16(1):4–12. <https://doi.org/10.1093/bfgp/eww025> PMID: [27345434](#)
9. Cheng J, Lin G, Wang T, Wang Y, Guo W, Liao J, et al. Massively parallel CRISPR-based genetic perturbation screening at single-cell resolution. *Adv Sci (Weinh)*. 2023;10(4):e2204484. <https://doi.org/10.1002/advs.202204484> PMID: [36504444](#)
10. Zhang H, Li X, Song D, Yukselen O, Nanda S, Kucukural A, et al. Worm Perturb-Seq: massively parallel whole-animal RNAi and RNA-seq. *Nat Commun*. 2025;16(1):4785. <https://doi.org/10.1038/s41467-025-60154-0> PMID: [40404656](#)
11. Dirmeier S, Dächert C, van Hemert M, Tas A, Ogando NS, van Kuppeveld F, et al. Host factor prioritization for pan-viral genetic perturbation screens using random intercept models and network propagation. *PLoS Comput Biol*. 2020;16(2):e1007587. <https://doi.org/10.1371/journal.pcbi.1007587> PMID: [32040506](#)
12. Xu Y, Fleming S, Tegtmeyer M, McCarroll SA, Babadi M. Explainable modeling of single-cell perturbation data using attention and sparse dictionary learning. *Cell Syst*. 2025;16(4):101245. <https://doi.org/10.1016/j.cels.2025.101245> PMID: [40187352](#)
13. Gao Y, Wei Z, Dong K, Chen K, Yang J, Chuai G, et al. Toward subtask-decomposition-based learning and benchmarking for predicting genetic perturbation outcomes and beyond. *Nat Comput Sci*. 2024;4(10):773–85. <https://doi.org/10.1038/s43588-024-00698-1> PMID: [39333790](#)
14. Yu H, Qian W, Song Y, Welch JD. PerturbNet predicts single-cell responses to unseen chemical and genetic perturbations. *Mol Syst Biol*. 2025;21(8):960–82. <https://doi.org/10.1038/s44320-025-00131-3> PMID: [40640612](#)
15. Xing H, Yau C. GPerturb: Gaussian process modelling of single-cell perturbation data. *Nat Commun*. 2025;16(1):5423. <https://doi.org/10.1038/s41467-025-61165-7> PMID: [40593897](#)
16. Roohani Y, Huang K, Leskovec J. Predicting transcriptional outcomes of novel multigene perturbations with GEARS. *Nat Biotechnol*. 2024;42(6):927–35. <https://doi.org/10.1038/s41587-023-01905-6> PMID: [37592036](#)
17. Lotfollahi M, Wolf FA, Theis FJ. scGen predicts single-cell perturbation responses. *Nat Methods*. 2019;16(8):715–21. <https://doi.org/10.1038/s41592-019-0494-8> PMID: [31363220](#)
18. Wu Y, Liu J, Xiao Y, Zhang S, Li L. CoupleVAE: coupled variational autoencoders for predicting perturbational single-cell RNA sequencing data. *Brief Bioinform*. 2025;26(2):bbaf126. <https://doi.org/10.1093/bib/bbaf126> PMID: [40178283](#)
19. Lotfollahi M, Klimovskaia Susmelj A, De Donno C, Hetzel L, Ji Y, Ibarra IL, et al. Predicting cellular responses to complex perturbations in high-throughput screens. *Mol Syst Biol*. 2023;19(6):e11517. <https://doi.org/10.15252/msb.202211517> PMID: [37154091](#)

20. Jiang Q, Chen S, Chen X, Jiang R. scPRAM accurately predicts single-cell gene expression perturbation response based on attention mechanism. *Bioinformatics*. 2024;40(5). <https://doi.org/10.1093/bioinformatics/btae265>
21. Cui H, Wang C, Maan H, Pang K, Luo F, Duan N, et al. scGPT: toward building a foundation model for single-cell multi-omics using generative AI. *Nat Methods*. 2024;21(8):1470–80. <https://doi.org/10.1038/s41592-024-02201-0> PMID: 38409223
22. Gavrilidis GI, Vasileiou V, Orfanou A, Ishaque N, Psomopoulos F. A mini-review on perturbation modelling across single-cell omic modalities. *Comput Struct Biotechnol J*. 2024;23:1886–96. <https://doi.org/10.1016/j.csbj.2024.04.058> PMID: 38721585
23. Papalexi E, Mimitou EP, Butler AW, Foster S, Bracken B, Mauck WM 3rd, et al. Characterizing the molecular regulation of inhibitory immune checkpoints with multimodal single-cell screens. *Nat Genet*. 2021;53(3):322–31. <https://doi.org/10.1038/s41588-021-00778-2> PMID: 33649593
24. Mimitou EP, Lareau CA, Chen KY, Zorretto-Fernandes AL, Hao Y, Takeshima Y, et al. Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nat Biotechnol*. 2021;39(10):1246–58. <https://doi.org/10.1038/s41587-021-00927-2> PMID: 34083792
25. Dhainaut M, Rose SA, Akturk G, Wroblewska A, Nielsen SR, Park ES, et al. Spatial CRISPR genomics identifies regulators of the tumor microenvironment. *Cell*. 2022;185(7):1223–1239.e20. <https://doi.org/10.1016/j.cell.2022.02.015> PMID: 35290801
26. Mimitou EP, Cheng A, Montalbano A, Hao S, Stoeckius M, Legut M, et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat Methods*. 2019;16(5):409–12. <https://doi.org/10.1038/s41592-019-0392-0> PMID: 31011186
27. Ghazalpour A, Bennett B, Petyuk VA, Orozco L, Hagopian R, Mungrue IN, et al. Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genet*. 2011;7(6):e1001393. <https://doi.org/10.1371/journal.pgen.1001393> PMID: 21695224
28. Kumar D, Bansal G, Narang A, Basak T, Abbas T, Dash D. Integrating transcriptome and proteome profiling: strategies and applications. *Proteomics*. 2016;16(19):2533–44. <https://doi.org/10.1002/pmic.201600140> PMID: 27343053
29. Heumos L, Schaar AC, Lance C, Litinetskaya A, Drost F, Zappia L, et al. Best practices for single-cell analysis across modalities. *Nat Rev Genet*. 2023;24(8):550–72. <https://doi.org/10.1038/s41576-023-00586-w> PMID: 37002403
30. Stanojevic S, Li Y, Ristivojevic A, Garmire LX. Computational methods for single-cell multi-omics integration and alignment. *Genomics Proteomics Bioinformatics*. 2022;20(5):836–49. <https://doi.org/10.1016/j.gpb.2022.11.013> PMID: 36581065
31. Argelaguet R, Cuomo ASE, Stegle O, Marioni JC. Computational principles and challenges in single-cell data integration. *Nat Biotechnol*. 2021;39(10):1202–15. <https://doi.org/10.1038/s41587-021-00895-7> PMID: 33941931
32. Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell transcriptomics. *Nat Methods*. 2018;15(12):1053–8. <https://doi.org/10.1038/s41592-018-0229-2> PMID: 30504886
33. Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, et al. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res*. 2004;32(Database issue):D258–61. <https://doi.org/10.1093/nar/gkh036> PMID: 14681407
34. Wessels H-H, Méndez-Mancilla A, Hao Y, Papalexi E, Mauck WM 3rd, Lu L, et al. Efficient combinatorial targeting of RNA transcripts in single cells with Cas13 RNA Perturb-seq. *Nat Methods*. 2023;20(1):86–94. <https://doi.org/10.1038/s41592-022-01705-x> PMID: 36550277
35. Mimitou EP, Lareau CA, Chen KY, Zorretto-Fernandes AL, Hao Y, Takeshima Y, et al. Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nat Biotechnol*. 2021;39(10):1246–58. <https://doi.org/10.1038/s41587-021-00927-2> PMID: 34083792
36. Chakraborty C, Bhattacharya M, Alshammari A, Albekairi TH. Blueprint of differentially expressed genes reveals the dynamic gene expression landscape and the gender biases in long COVID. *J Infect Public Health*. 2024;17(5):748–66. <https://doi.org/10.1016/j.jiph.2024.02.018> PMID: 38518681
37. Chen Y, Zou Q, Chen Q, Wang S, Du Q, Mai Q, et al. Methylation-related differentially expressed genes as potential prognostic biomarkers for cervical cancer. *Heliyon*. 2024;10(17):e36240. <https://doi.org/10.1016/j.heliyon.2024.e36240> PMID: 39263148
38. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol*. 2018;19(1):15. <https://doi.org/10.1186/s13059-017-1382-0> PMID: 29409532
39. Gretton A, Borgwardt KM, Rasch MJ, Schoelkopf B, Smola A. A Kernel two-sample test. *J Mach Learn Res*. 2012;13:723–73.
40. Panaretos VM, Zemel Y. Statistical aspects of Wasserstein distances. *Annual Review of Statistics and Its Application*. 2019;6(1):405–31. <https://doi.org/10.1146/annurev-statistics-030718-104938>
41. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer*. 2012;12(4):252–64. <https://doi.org/10.1038/nrc3239> PMID: 22437870
42. Wang X, Teng F, Kong L, Yu J. PD-L1 expression in human cancers and its association with clinical outcomes. *Onco Targets Ther*. 2016;9:5023–39. <https://doi.org/10.2147/OTT.S105862> PMID: 27574444
43. Mao P, Feng W, Zhang Z, Huang C, Zhou S, Zhao Z, et al. Cyclic adenosine monophosphate potentiates immune checkpoint blockade therapy in acute myeloid leukemia. *Clin Transl Med*. 2023;13(11):e1489. <https://doi.org/10.1002/ctm2.1489> PMID: 37997561
44. Lai Q, Wang H, Li A, Xu Y, Tang L, Chen Q, et al. Decitabine improve the efficiency of anti-PD-1 therapy via activating the response to IFN/PD-L1 signal of lung cancer cells. *Oncogene*. 2018;37(17):2302–12. <https://doi.org/10.1038/s41388-018-0125-3> PMID: 29422611
45. Garcia-Diaz A, Shin DS, Moreno BH, Saco J, Escuin-Ordinas H, Rodriguez GA, et al. Interferon Receptor Signaling Pathways Regulating PD-L1 and PD-L2 Expression. *Cell Rep*. 2017;19(6):1189–201. <https://doi.org/10.1016/j.celrep.2017.04.031> PMID: 28494868
46. Wu Z, Wang Z, Hua Z, Ji Y, Ye Q, Zhang H, et al. Prognostic signature and immunotherapeutic relevance of Focal adhesion signaling pathway-related genes in osteosarcoma. *Heliyon*. 2024;10(21):e38523. <https://doi.org/10.1016/j.heliyon.2024.e38523> PMID: 39524888

47. Xu Y, Zhang D, Ji J, Zhang L. Ubiquitin ligase MARCH8 promotes the malignant progression of hepatocellular carcinoma through PTEN ubiquitination and degradation. *Mol Carcinog.* 2023;62(7):1062–72. <https://doi.org/10.1002/mc.23546> PMID: [37098835](#)
48. Zhao T, Li Y, Zhang J, Zhang B. PD-L1 expression increased by IFN- γ via JAK2-STAT1 signaling and predicts a poor survival in colorectal cancer. *Oncol Lett.* 2020;20(2):1127–34. <https://doi.org/10.3892/ol.2020.11647> PMID: [32724352](#)
49. Chen Q, Zhuang S, Hong Y, Yang L, Guo P, Mo P, et al. Demethylase JMJD2D induces PD-L1 expression to promote colorectal cancer immune escape by enhancing IFNGR1-STAT3-IRF1 signaling. *Oncogene.* 2022;41(10):1421–33. <https://doi.org/10.1038/s41388-021-02173-x> PMID: [35027670](#)