GENINV: GENERATIVE INVERSE LEARNING FOR CELL PERTURBATION PREDICTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Predicting how drugs alter gene expression across cellular contexts can guide the selection of treatments for experimental screening. Large single-cell perturbation screens provide a basis for learning these responses, but measure separate control and treated populations without observing individual cells before and after treatment. Recent models address this constraint by learning transitions between populations. Their forward objective directly supervises the treated distribution, yet does not require the predictor to recover the untreated context represented in those observations. This suggests a complementary learning signal: predicting controls from observed treated cells. We introduce **GenInv**, a generative inverse learning approach that jointly supervises forward response prediction and reverse control prediction through a shared network. Both tasks use the same experimental measurements, and inference retains the original forward procedure. On five evaluation cell lines in State-Tahoe-Filtered, GenInv improves all four primary response metrics over a matched forward-only model. Mean response correlation (Pearson-Δ) increases from 0.922 to 0.931, and differential-expression log-fold-change Spearman correlation increases from 0.612 to 0.665. Label interventions show that reverse prediction retains an expression-level benefit under shuffled inverse labels, while correct conditioning further improves differential-expression recovery. These results establish generative inverse supervision as an effective training principle in this setting, extending the predictive value of existing perturbation measurements. Code will be available at `https://anonymous.4open.science/r/geninv/`.

1 INTRODUCTION

Drug-induced transcriptional responses depend on the cellular context in which a perturbation is applied. Predicting this dependence can help prioritize treatments and cell types for experimental investigation. Single-cell RNA sequencing provides a detailed molecular readout by representing each cell through the abundance of its RNA transcripts. Perturbation atlases such as Tahoe-100M (Zhang et al., 2025) collect these expression profiles across many treatments and cell lines. Their scale makes it possible to learn response patterns shared across contexts while accounting for the variation associated with each treated population.

The destructive nature of sequencing imposes a central modeling constraint. An experiment measures separate control and treated populations, rather than the expression profiles of the same cells before and after treatment. Response prediction must therefore relate distributions without access to cell-level correspondences. Latent-space models, optimal transport, and more recent set-based and diffusion models provide different approaches to this problem (Lotfollahi et al., 2019; 2023; Bunne et al., 2023; Adduri et al., 2025; Yuan et al., 2026). In particular, the State transition model conditions on a sample of control cells and a treatment label to predict a treated population. A distributional loss compares this prediction with the measured treated cells, allowing supervision without assigning individual source cells to individual targets (Adduri et al., 2025).

Forward supervision aligns with the prediction task, but constrains only one direction of the observed relationship. Treated cells also contain information about their cellular context, and the matched control population provides a direct target for recovering it. This supports an auxiliary task that predicts controls from treated cells through the same network. Recovering the untreated distri-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

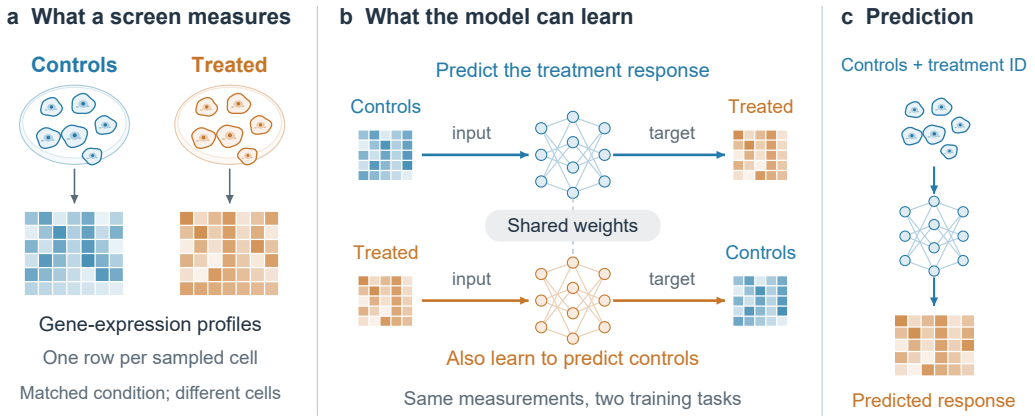

Figure 1: **Generative inverse supervision from paired cell populations.** **(a)** Control and treated populations are measured separately, with one expression-matrix row per cell. **(b)** GenInv jointly supervises treatment-response prediction and inverse control prediction using a shared model and observed targets in both directions. **(c)** Inference uses control cells and the treatment label through the forward predictor. Cell icons and matrix colors are schematic; arrows denote population-level prediction tasks rather than individual-cell trajectories.

bution may encourage representations that preserve cellular context alongside treatment-associated changes. Both observed distributions can therefore serve as supervision targets for a common predictor.

This observation motivates joint forward and reverse population prediction (Figure 1), which we term GenInv, or generative inverse learning. Each direction conditions on an observed source population and the treatment label, and matches the corresponding observed target distribution. A shared network couples the tasks; separate treatment encoders accommodate their distinct conditioning roles. Both objectives contribute during training, while inference retains only forward prediction. Inverse prediction estimates a control distribution statistically, without assuming biological reversibility or recovery of an individual cell's pretreatment state.

We evaluate GenInv using the State transition model on State-Tahoe-Filtered under matched data splits and 200k-step budgets. Inverse supervision improves response correlation, differential-expression (DE) direction agreement, DE effect correlation, and significant-gene recall across five cell lines. Mean DE log-fold-change Spearman increases from 0.612 to 0.665, an 8.6% relative improvement. Label interventions distinguish reverse prediction from correct inverse conditioning: shuffling retains much of the expression-level benefit but reduces DE recovery.

Our contributions comprise: *1)* a generative inverse objective that extracts bidirectional supervision from existing population pairs, *2)* a matched evaluation establishing gains across five cell lines, and *3)* a diagnostic analysis of inverse conditioning. The rest of this paper is organized as follows: Section 2 reviews related work, Section 3 presents the architecture and objective, and Section 4 describes the experimental protocol and evaluates predictive performance and inverse conditioning. Section 5 summarizes the findings and future directions.

2 Related Work

Prediction of cellular perturbation responses. Early perturbation models demonstrated transfer of treatment effects between cellular contexts through learned representations. scGen combines variational modeling with latent arithmetic (Lotfollahi et al., 2019), while the compositional perturbation autoencoder (CPA) separates basal state, perturbation, and other covariate contributions (Lotfollahi et al., 2023). For genetic interventions, GEARS uses relationships between genes to predict responses to unobserved multigene perturbations (Roohani et al., 2023). Population-level

methods explicitly address the absence of paired source and target cells. CellOT learns a transport map between their distributions (Bunne et al., 2023).

State learns transitions between cell sets (Adduri et al., 2025), drawing on attention and permutation-aware set modeling (Vaswani et al., 2017; Lee et al., 2019; Zaheer et al., 2017); PerturbDiff uses a diffusion process over distributions (Yuan et al., 2026). These advances expand the class of population transitions that models can represent. GenInv addresses a complementary aspect of the learning problem: the supervision extracted from each observed transition.

Cellular representations and response modeling. A cell's expression profile reflects both biological state and measurement-associated variation. scVI uses a deep generative model to represent this variation (Lopez et al., 2018). Geneformer, scGPT, and scFoundation learn cell or gene representations from large expression corpora (Theodoris et al., 2023; Cui et al., 2024; Hao et al., 2024). These representations support a range of transcriptomic tasks, while perturbation prediction additionally requires learning treatment-associated transitions. We study this transition-learning problem using highly variable genes (HVGs), a subset selected to retain informative expression variation. Our backbone is the State transition model trained from scratch in HVG space; the separately pretrained cell encoder in the full State system is not used in this evaluation.

Bidirectional supervision and distribution matching. Learning across related prediction directions is well established. Dual learning couples translation in two languages (He et al., 2016; Xia et al., 2017), and back-translation uses a reverse model to generate additional training examples (Sennrich et al., 2016). CycleGAN and DualGAN connect unpaired domains through a cycle-consistency objective (Zhu et al., 2017; Yi et al., 2017). Cell perturbation experiments provide condition-matched populations despite the absence of cell-level pairs. Both observed distributions can therefore serve as direct supervision targets, without composing predictions in a reconstruction cycle. Energy distances and maximum mean discrepancy provide established discrepancies for comparing population samples (Gretton et al., 2012; Székely & Rizzo, 2013; Sejdinovic et al., 2013; Bińkowski et al., 2018). GenInv applies the State distributional loss in both directions. Label interventions assess the contribution of treatment conditioning, while exploratory discriminative objectives examine the potential for shortcut learning (Geirhos et al., 2020).

3 METHOD

3.1 SHARED POPULATION PREDICTION ARCHITECTURE

Training uses samples from control and perturbed populations in the same cellular context. Let $X_{\mathrm{ctrl}} = \{x_i^c\}_{i=1}^n$ and $X_{\mathrm{pert}} = \{x_j^p\}_{j=1}^m$ denote these samples, and let p identify the perturbation. Each vector records expression for $d = 2000$ HVGs. The populations may differ in sample size, and individual cells are unpaired. The predictive objective is to estimate the treated expression distribution conditional on the control sample and p.

We write the State transition network as $F_\theta(X, e)$, where X is the source population, e is a perturbation embedding, and θ contains the parameters shared between directions. The forward predictor is

$$f(X_{\mathrm{ctrl}}, p) = F_\theta\big(X_{\mathrm{ctrl}}, E_{\mathrm{fwd}}(p)\big). \tag{1}$$

The network projects each cell's expression profile into a token. The encoder E_{fwd} maps the one-hot treatment label to an embedding, which is added to every token. A set transformer models interactions among cell tokens, conditioning each prediction on the sampled population as well as the source cell. Writing the cell projector as P, the transformer as T, and the shared linear output head as W, the prediction path in Figure 2 is

$$F_\theta(X, e) = \mathrm{ReLU}(W\,[T(P(X) + e) + P(X)]), \tag{2}$$

where e is broadcast across cell tokens. The residual combines each projected source cell with its population-conditioned transformer representation before the output head. The final ReLU enforces nonnegative expression predictions. Both directions use this shared computation.

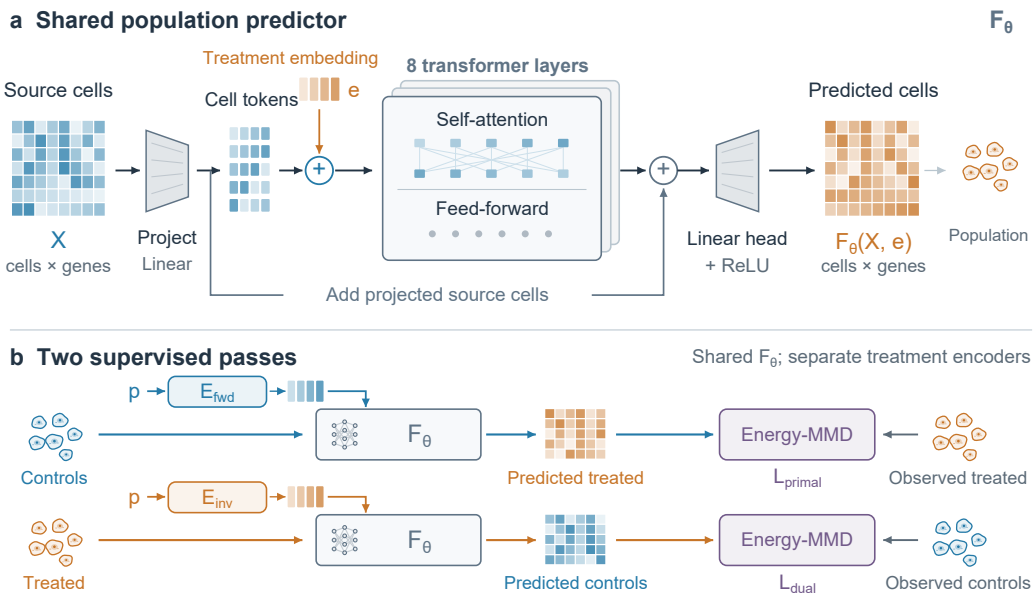

Figure 2: **Shared architecture for bidirectional population modeling. (a)** Projected expression profiles and treatment embeddings form the input to a shared set transformer. A residual connection adds projected source tokens before the linear output head and ReLU. **(b)** Forward and inverse passes share the prediction network but use independent treatment encoders. Each pass receives an observed source population and is supervised against the corresponding observed target using energy-MMD. Matrix colors and token interactions are schematic representations of the computation.

3.2 Generative inverse supervision

The reverse task exchanges the roles of the two observed populations. We use the same perturbation identity p, now embedded by a separate inverse encoder,

$$g(X_{\text{pert}}, p) = F_\theta\big(X_{\text{pert}}, E_{\text{inv}}(p)\big). \tag{3}$$

The inverse task takes the observed treated sample as input and the observed control sample as its target. The label identifies the original treatment, while a separate encoder allows its representation to adapt to control prediction. We initialize E_{inv} from E_{fwd} and optimize the two independently. The cell projector, transformer, and output head receive gradients from both objectives.

We supervise unpaired populations using State's energy-MMD loss, denoted $D_E(A, B)$. This distributional discrepancy compares samples without specifying a correspondence between individual cells. The forward and inverse losses are

$$\mathcal{L}_{\text{primal}} = D_E\big(f(X_{\text{ctrl}}, p), X_{\text{pert}}\big), \tag{4}$$

$$\mathcal{L}_{\text{dual}} = D_E\big(g(X_{\text{pert}}, p), X_{\text{ctrl}}\big). \tag{5}$$

Both observed targets are available in the same training batch. The joint objective is

$$\mathcal{L} = \mathcal{L}_{\text{primal}} + \lambda_{\text{dual}}\mathcal{L}_{\text{dual}}, \qquad \lambda_{\text{dual}} = 1. \tag{6}$$

We use a unit inverse-loss weight throughout the reported experiments. Each pass starts from an observed population and receives direct distributional supervision. Inverse training therefore requires no composition with a forward prediction. The objective includes neither a cycle-consistency term nor an output-variance regularizer.

Parameter sharing transfers the inverse training signal to the forward predictor. The network must model treated populations both as prediction targets and as inputs for recovering controls, potentially preserving information about cellular context across the two tasks. A competing explanation is that inverse prediction primarily recovers common control distributions with limited dependence on p. The label interventions in Section 4.4 evaluate the contribution of correct treatment conditioning.

3.3 Implementation and inference

We use the State transition architecture with an eight-layer LLaMA-style transformer (Touvron et al., 2023), 12 attention heads, a hidden dimension of 768, and bidirectional attention. Each training set contains 256 cells. Single-layer linear perturbation encoders map 1138-dimensional one-hot vectors to the hidden space. GenInv adds only the inverse perturbation encoder, comprising 874,752 parameters, approximately 0.8% of the baseline model's stored parameters. This encoder is removed for inference. Appendix G provides the complete architecture, optimization settings, and data-partition details.

Each training step evaluates both directions on the same batch, sums their losses, and updates the shared and task-specific parameters jointly. Inference evaluates f once using the supplied controls and treatment label. The auxiliary objective adds a training pass while preserving the forward computation used for prediction. Parameter counts quantify the additional model capacity; a controlled comparison of end-to-end training time is unavailable.

4 Experiment Evaluation

4.1 Experiment Settings

Datasets. We use State-Tahoe-Filtered, the 50-cell-line corpus derived from Tahoe-100M (Zhang et al., 2025), with 100.65M stored cells and a supplied 2000-dimensional log-normalized HVG representation. We follow the State few-shot split: designated treatment conditions in C32, HOP62, HepG2-C3A, Hs 766T, and PANC-1 are reserved for validation and testing. Training includes the other 45 cell lines and the remaining conditions in these five lines. Evaluation therefore measures prediction of held-out cell-line–treatment combinations. At inference, each model receives control cells and the treatment label; observed treated cells for a test condition are used exclusively for evaluation.

Models. Our primary baseline, denoted *State-ST*, is a forward-only replication of the State transition model in the HVG space. State-ST and GenInv use the same architecture for forward prediction, data split, optimizer settings, and 200k-step budget. GenInv adds the inverse encoder and auxiliary objective. Each reported neural configuration uses one training run with seed 42; differences across cell lines or conditions are therefore not estimates of variation across training seeds. We train from scratch on one A100-80GB GPU per run using Adam, a learning rate of 3×10^{-4}, no weight decay, and batches of 64 cell sets. These optimizer settings are recorded in the checkpoints. Validation and checkpointing occur every 4k steps. We report the final 200k-step comparison and use 100k-step checkpoints for the shuffled-label analysis.

Metrics. We evaluate both overall expression responses and the recovery of differentially expressed genes. Pearson-Δ compares predicted and observed shifts in mean expression, each centered on its corresponding control mean. Differential-expression (DE) analysis uses Wilcoxon tests with Benjamini–Hochberg correction at FDR 0.05 (Benjamini & Hochberg, 1995). DE direction agreement evaluates the signs of observed significant changes; log-fold-change (LFC) Spearman correlation evaluates their relative magnitudes; significant-gene recall evaluates recovery of the responding set. Each predicted test population contains the same number of cells as its observed target. We use a memory-efficient implementation of Cell-Eval-style metrics. Scores are averaged within each cell line and then equally across the five lines; condition-level plots compare paired prediction cases. Complete metric definitions and per-cell-line results are provided in Appendices F and B.

These metrics distinguish effect estimation from significance detection. Direction agreement and LFC correlation are evaluated on genes significant in the observed data, regardless of their predicted significance. Recall measures the fraction of this observed set also identified as significant by the model. DE AUROC additionally evaluates the ranking induced by $-\log_{10}(q_{\mathrm{pred}})$, where q_{pred} is the adjusted p-value in the predicted data, against observed significance labels. Consequently, accurate response directions need not imply accurate selection of significant genes, and high recall can coexist with excessive positive calls. Joint evaluation of effect agreement, recall, and ranking performance separates these aspects of response prediction.

Table 1: **Perturbation-response accuracy with generative inverse supervision.** State-ST and GenInv use the same forward architecture, data split, and 200k-step budget. Entries average conditions within each cell line; mean rows weight the five lines equally. PD: Pearson-Δ; direction: DE direction agreement; LFC: DE LFC Spearman; recall: significant-gene recall. Higher values indicate better performance. Bold and underline denote the higher and lower score in each pair, ranked before rounding.

Cell line	Model	PD	Direction	LFC	Recall
C32	State-ST	0.917	0.899	0.605	0.713
	GenInv	**0.927**	**0.917**	**0.665**	**0.735**
HOP62	State-ST	0.950	0.914	0.685	0.759
	GenInv	**0.954**	**0.929**	**0.722**	**0.783**
HepG2-C3A	State-ST	0.908	0.877	0.571	0.637
	GenInv	**0.919**	**0.901**	**0.638**	**0.668**
Hs 766T	State-ST	0.906	0.895	0.572	0.680
	GenInv	**0.919**	**0.918**	**0.642**	**0.709**
PANC-1	State-ST	0.929	0.906	0.627	0.717
	GenInv	**0.936**	**0.921**	**0.656**	**0.748**
Mean	State-ST	0.922	0.898	0.612	0.701
	GenInv	**0.931**	**0.917**	**0.665**	**0.729**

4.2 Inverse supervision improves forward response prediction

GenInv improves all four primary metrics in every evaluation cell line at 200k steps (Table 1). The most pronounced improvement concerns gene-level effect recovery: mean DE LFC Spearman increases from 0.612 to 0.665. Direction agreement increases from 0.898 to 0.917, indicating more accurate signs of treatment-associated expression changes. Pearson-Δ increases from 0.922 to 0.931, extending the benefit to population-level response estimates. These gains arise under matched forward architectures, data splits, and training-step budgets.

Condition-level comparisons show that the aggregate improvement extends across a majority of evaluated responses. Among the 3673 pairs in Figure 3, GenInv improves Pearson-Δ in 79.6% and DE AUROC in 71.1%. Positive binned means span a range of baseline scores, including conditions with high forward-only accuracy. The benefit of inverse supervision is therefore distributed across conditions rather than concentrated in a small subset of large improvements.

The strongest evidence for inverse supervision concerns response correlation and DE effect recovery. Cosine discrimination is nearly saturated for both models (0.9975 for State-ST and 0.9973 for GenInv), whereas gene-overlap metrics exhibit mixed changes across cell lines. These results locate the principal benefit in estimating treatment-associated expression changes, with less uniform improvement in the exact membership of the predicted responding gene set.

4.3 Inverse supervision sustains gains during continued training

At 100k steps, GenInv exceeds State-ST in Pearson-Δ, direction agreement, and DE LFC correlation on all five lines. Continued training increases its mean Pearson-Δ advantage from 0.0047 to 0.0092: State-ST decreases from 0.9275 to 0.9220, while GenInv changes from 0.9322 to 0.9312. The DE LFC advantage also increases, from 0.0361 to approximately 0.052. The improvement is therefore sustained across the two evaluated checkpoints and becomes more pronounced for response correlation and gene-level effect recovery.

The smaller decline in response correlation is consistent with a regularizing effect. Retained validation segments and sparse monitoring records also favor GenInv, with a released State run providing context. The test comparisons use matched checkpoints from the locally trained models.

4.4 Correct inverse conditioning improves gene-level effect recovery

A reverse predictor could recover a cell line's common control distribution with limited dependence on treatment identity. We evaluate this possibility by replacing the treatment label at inference. On

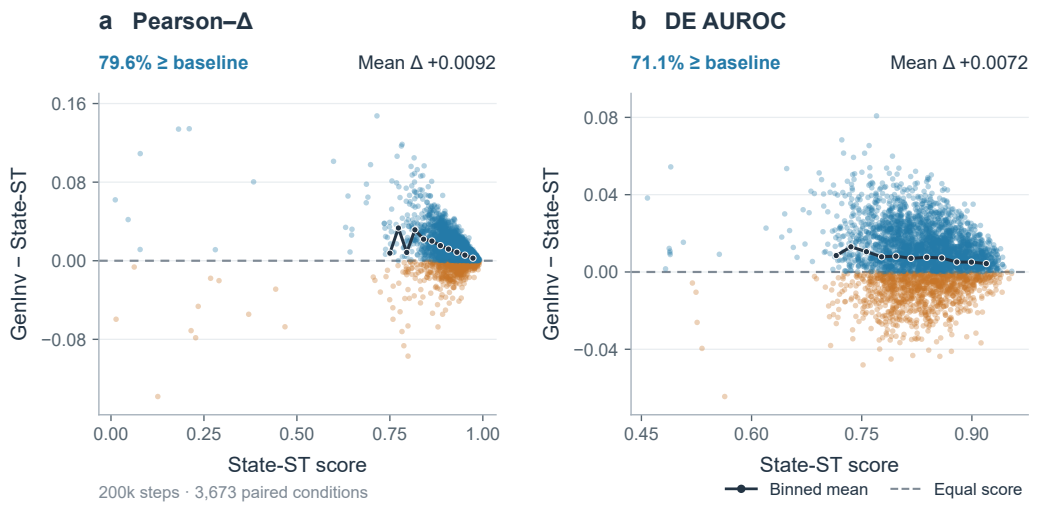

Figure 3: **Condition-level gains in response correlation and DE discrimination.** At 200k steps, GenInv improves **(a)** Pearson-Δ in 79.6% and **(b)** DE AUROC in 71.1% of 3673 cell-line–perturbation pairs; mean gains are 0.0092 and 0.0072. Axes report State-ST scores and paired differences (GenInv minus State-ST). Blue/amber denotes positive/negative differences, dark curves denote binned means, and dashed lines denote zero difference. All eligible pairs are shown; neither metric has ties.

Table 2: **Contribution of inverse-label conditioning to forward prediction.** Scores are equal-weight means over five evaluation cell lines at 100k steps. The shuffled-label control permutes inverse labels while retaining correct forward labels. Both inverse objectives improve Pearson-Δ over State-ST; correct conditioning yields the highest DE direction agreement and LFC correlation. Higher values indicate better performance. Bold and underline identify the best and second-best scores, ranked before rounding.

Setting	Pearson-Δ	DE AUROC	DE direction	DE LFC
State-ST	0.928	0.834	0.903	0.617
GenInv, shuffled	**0.932**	**0.839**	<u>0.908</u>	<u>0.633</u>
GenInv, correct	<u>0.932</u>	<u>0.837</u>	**0.914**	**0.653**

C32 and PANC-1, the intervention reduces forward response correlation from approximately 0.83 to 0.34–0.38. Reverse reconstruction is substantially less sensitive: correct and substituted labels yield 0.871 versus 0.868 and 0.865 versus 0.865 in the corresponding probes. These results indicate weaker dependence on treatment identity during reverse prediction than during forward prediction.

Inference-time sensitivity and the contribution of conditioning during training are distinct properties. To evaluate the latter, we train a separate *shuffled-inverse-label* model, retaining correct forward labels and the same training recipe while permuting labels in the inverse branch. Table 2 compares this control with GenInv and State-ST under the same 100k-step budget.

The shuffled-label model retains the expression-level benefit: its Pearson-Δ is 0.9323, compared with 0.9322 for correct labels and 0.9275 for State-ST. Correct labels nevertheless improve DE direction agreement by 0.0061 and DE LFC correlation by 0.0201 over shuffled labels. Relative to the total GenInv improvement over State-ST at this checkpoint, shuffled labels recover approximately 46% of the direction gain and 44% of the LFC gain. Correct labels match or exceed shuffled labels in 68.3% of conditions for direction agreement (including 7.1% ties) and improve LFC correlation in 65.6% (Figure 4).

Control prediction retains a benefit under shuffled conditioning, while correct labels further improve DE recovery. This distinguishes the empirical contribution of the auxiliary task from that of treat-

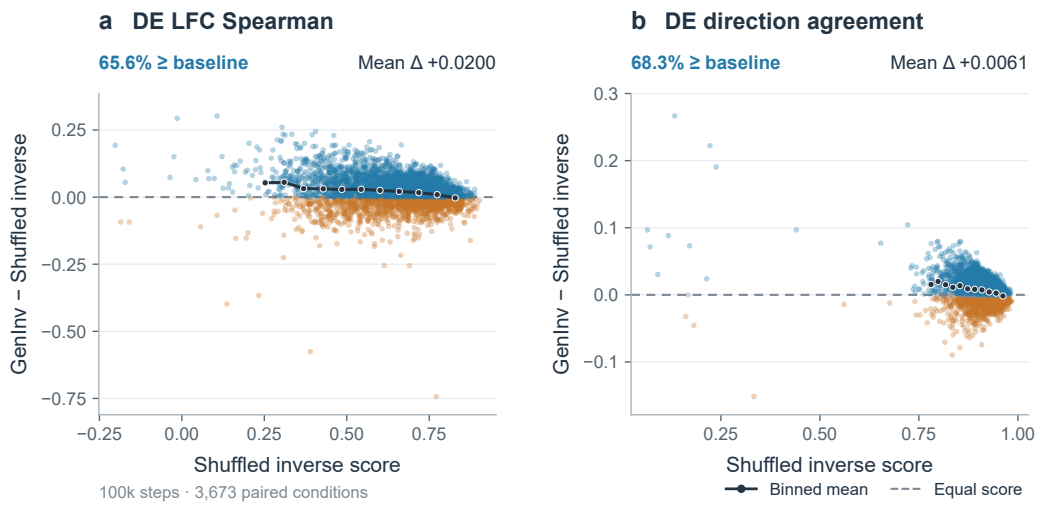

Figure 4: **Gene-level effect recovery with correct inverse conditioning.** Correct versus shuffled inverse labels at 100k steps on 3673 cell-line–perturbation pairs with at least five observed significant genes. Correct conditioning improves **(a)** DE LFC Spearman in 65.6% of pairs and matches or exceeds **(b)** direction agreement in 68.3% (7.1% ties). Axes report shuffled-label scores and paired differences. Blue/amber denotes positive/negative differences; dark curves denote binned means. All eligible pairs are shown.

ment identity. The gain ratios are descriptive rather than a causal decomposition, because shuffling also changes the learned parameters.

4.5 Gene-selection accuracy and response breadth

We examine the relationship between response accuracy and significant-gene selection using three structured linear predictors, an additive response, a ridge-regularized scalar slope with gene-specific intercepts, and a diagonal map with a separate slope and intercept for each gene and perturbation, together with a control-mean predictor that outputs the mean control expression as a no-response lower bound (Table 3). The models are fitted on 45 cell lines and scored with the same evaluation harness. Their fitting data exclude the five evaluation lines, whereas the neural models follow the few-shot protocol. This comparison characterizes prediction behavior across model classes; State-ST remains the matched baseline for evaluating inverse supervision.

The control-mean predictor attains near-zero Pearson-Δ (0.022) and chance AUROC (0.500), confirming the no-response lower bound. Gene-specific flexibility improves the linear models' Pearson-Δ from 0.428 to 0.615. The three generative VAE baselines (scGen, CPA, chemCPA) attain lower Pearson-Δ (0.23–0.30) and predict nearly all genes as significant with near-chance DE AUROC ($\sim$0.55). The neural predictors achieve higher scores; because their training coverage differs, this comparison does not isolate the contribution of nonlinear modeling.

High recall is not sufficient to establish accurate DE recovery: the diagonal model achieves the highest recall (0.823) while identifying 1605 of 2000 genes as significant, compared with an observed mean of 425, and its DE AUROC of 0.499 indicates near-chance discrimination. GenInv also predicts more significant genes than observed, but combines higher recall than State-ST with improvements in AUROC and effect correlation.

DE count Spearman correlates significant-gene counts across conditions within each cell line; its five-line mean improves from 0.8966 to 0.9300 at 200k steps. State-ST and GenInv predict 567.3 and 597.0 significant genes on average, compared with 425.1 observed, so GenInv better captures variation in response breadth while both models overestimate absolute counts.

Table 3: **Response recovery and DE selection across linear, generative, and neural predictors.** Five-cell-line means under a common scoring procedure. Linear and generative VAE models are fitted on the 45 training lines; the State-based neural models follow the official few-shot protocol and are evaluated at 200k steps. Bold and underline rank methods within each training group. #DE denotes the mean number of predicted significant genes, compared with approximately 425 observed. Recall measures sensitivity, while AUROC assesses discrimination between responding and nonresponding genes.

Model	PD	Direction	LFC	Recall	AUROC	#DE
Control mean	0.022	0.501	0.003	0.000	0.500	0
Additive	0.428	0.668	0.261	0.571	0.494	1168
Ridge	0.536	0.777	0.374	0.711	0.505	1380
Diagonal	0.615	0.804	0.436	0.823	0.499	1605
scGen	0.281	0.655	0.227	**0.953**	0.550	1752
CPA	0.303	0.649	0.240	0.950	0.552	1740
chemCPA	0.226	0.625	0.200	0.952	0.559	1721
State-ST	0.922	0.898	0.612	0.701	0.833	567
GenInv (this work)	**0.931**	**0.917**	**0.665**	0.729	**0.840**	597

4.6 Exploratory evaluation of alternative auxiliary objectives

Generative inverse supervision constrains predicted expression distributions. We also explored classifiers that recover treatment labels from transformer-state differences or predicted expression deltas, including a reinforced output-level variant. Classification saturated early without improving response prediction. Because the label is already supplied to the predictor, these objectives can exploit information inherited from the input. This behavior is consistent with shortcut learning (Geirhos et al., 2020) and motivates distributional supervision that directly evaluates predicted expression.

Exploratory cycle-consistency and disjoint-supervision variants did not improve the selected design. The former adds composed predictions; the latter removes one direction of supervision for each assigned perturbation. Controls for an earlier variance-anchored variant attributed its gain to the variance regularizer, with no additional benefit from its dual component. GenInv excludes this regularizer. These historical experiments inform the design but are not a matched 200k-step ablation. Supplementary condition-level analyses, validation records, and exploratory settings are consolidated in Appendices C and E.

5 Conclusion and future work

GenInv demonstrates that observed control and treated populations can jointly supervise a shared perturbation-response model. The auxiliary control-prediction task improves all four primary response metrics across five evaluation cell lines while preserving the forward inference procedure. Label interventions distinguish two empirical effects: reverse population prediction retains a benefit under shuffled inverse labels, and correct treatment conditioning further improves gene-level response recovery.

Control populations thus provide supervision as well as a reference state: predicting their distribution from treated observations improves estimates of treatment-induced changes. The evidence covers one benchmark, one backbone, and one seed per configuration, and generative inverse supervision improves prediction by making additional use of existing data. This principle motivates two extensions: conditioning on cellular context could distinguish transferable from memorized control distributions, and time-series measurements could supervise transitions beyond a single comparison.

AI USE STATEMENT

Generative AI tools assisted manuscript revision, reference and implementation checks, and polishing of the figures. The experimental results reported here originate from the existing experimental record; this manuscript revision did not generate any new experiments. The authors are responsible for checking the final scientific content, references, and presentation.

ETHICS STATEMENT

This study uses the publicly released Tahoe-100M-derived data described in Section 4 and introduces no new biological experiments or participant recruitment. The evaluation concerns transcriptional responses in cultured cancer cell lines. Predictions for primary tissues, patients, or clinical treatment decisions are outside the validated scope. Reporting DE selection errors alongside response correlations helps distinguish plausible population predictions from reliable identification of individual responsive genes.

REPRODUCIBILITY STATEMENT

Section 3 specifies the forward and inverse objectives and parameter sharing, and Section 4 describes the training comparison and evaluation setting. Appendix G lists the architecture and optimization settings; Appendices A and B report the complete numerical results. Appendix F describes the evaluation metrics. An anonymous repository is available for reproduction purposes.

REFERENCES

Abhinav K. Adduri, Dhruv Gautam, Beatrice Bevilacqua, Alishba Imran, Rohan Shah, Mohsen Naghipourfar, et al. Predicting cellular responses to perturbation across diverse contexts with State. *bioRxiv*, pp. 2025.06.26.661135, 2025.

Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1):289–300, 1995.

Mikołaj Bińkowski, Dougal J. Sutherland, Michael Arbel, and Arthur Gretton. Demystifying MMD GANs. In *International Conference on Learning Representations (ICLR)*, 2018.

Charlotte Bunne, Stefan G. Stark, Gabriele Gut, Jacobo Sarabia del Castillo, et al. Learning single-cell perturbation responses using neural optimal transport. *Nature Methods*, 20:1759–1768, 2023.

Haotian Cui, Chloe Wang, Hassaan Maan, Kuan Pang, Fengning Luo, Nan Duan, and Bo Wang. scGPT: Toward building a foundation model for single-cell multi-omics using generative AI. *Nature Methods*, 21:1470–1480, 2024.

Robert Geirhos, Jörn-Henrik Jacobsen, Claudio Michaelis, Richard Zemel, Wieland Brendel, Matthias Bethge, and Felix A. Wichmann. Shortcut learning in deep neural networks. *Nature Machine Intelligence*, 2(11):665–673, 2020.

Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13:723–773, 2012.

Minsheng Hao, Jing Gong, Xin Zeng, Chiming Liu, Yucheng Guo, Xingyi Cheng, et al. Large-scale foundation model on single-cell transcriptomics. *Nature Methods*, 21:1481–1491, 2024.

Di He, Yingce Xia, Tao Qin, Liwei Wang, Nenghai Yu, Tie-Yan Liu, and Wei-Ying Ma. Dual learning for machine translation. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2016.

Juho Lee, Yoonho Lee, Jungtaek Kim, Adam Kosiorek, Seungjin Choi, and Yee Whye Teh. Set Transformer: A framework for attention-based permutation-invariant neural networks. In *International Conference on Machine Learning (ICML)*, pp. 3744–3753, 2019.

Romain Lopez, Jeffrey Regier, Michael B. Cole, Michael I. Jordan, and Nir Yosef. Deep generative modeling for single-cell transcriptomics. *Nature Methods*, 15:1053–1058, 2018.

Mohammad Lotfollahi, F. Alexander Wolf, and Fabian J. Theis. scGen predicts single-cell perturbation responses. *Nature Methods*, 2019.

Mohammad Lotfollahi, Anna Klimovskaia Susmelj, Carlo De Donno, Leon Hetzel, Yuge Ji, Ignacio L Ibarra, et al. Predicting cellular responses to complex perturbations in high-throughput screens. *Molecular Systems Biology*, 2023.

Yusuf Roohani, Kexin Huang, and Jure Leskovec. Predicting transcriptional outcomes of novel multigene perturbations with GEARS. *Nature Biotechnology*, 2023.

Dino Sejdinovic, Bharath Sriperumbudur, Arthur Gretton, and Kenji Fukumizu. Equivalence of distance-based and RKHS-based statistics in hypothesis testing. *The Annals of Statistics*, 41(5): 2263–2291, 2013.

Rico Sennrich, Barry Haddow, and Alexandra Birch. Improving neural machine translation models with monolingual data. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 86–96, 2016.

Gábor J. Székely and Maria L. Rizzo. Energy statistics: A class of statistics based on distances. *Journal of Statistical Planning and Inference*, 143(8):1249–1272, 2013.

Christina V. Theodoris, Ling Xiao, Anant Chopra, Mark D. Chaffin, et al. Transfer learning enables predictions in network biology. *Nature*, 618:616–624, 2023.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, et al. LLaMA: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

Yingce Xia, Tao Qin, Wei Chen, Jiang Bian, Nenghai Yu, and Tie-Yan Liu. Dual supervised learning. In *International Conference on Machine Learning (ICML)*, pp. 3789–3798, 2017.

Zili Yi, Hao Zhang, Ping Tan, and Minglun Gong. DualGAN: Unsupervised dual learning for image-to-image translation. In *IEEE International Conference on Computer Vision (ICCV)*, pp. 2849–2857, 2017.

Xinyu Yuan, Xixian Liu, Ya Shi Zhang, Zuobai Zhang, Hongyu Guo, and Jian Tang. PerturbDiff: Functional diffusion for single-cell perturbation modeling. *arXiv preprint arXiv:2602.19685*, 2026.

Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Ruslan Salakhutdinov, and Alexander Smola. Deep Sets. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

Jesse Zhang, Airol A Ubas, Richard de Borja, Valentine Svensson, Nicole Thomas, Neha Thakar, et al. Tahoe-100M: A giga-scale single-cell perturbation atlas for context-dependent gene function and cellular modeling. *bioRxiv*, pp. 2025.02.20.639398, 2025.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *IEEE International Conference on Computer Vision (ICCV)*, 2017.

A COMPLETE REPORTED RESULTS

This appendix reports the complete numerical results at four-decimal precision. State-ST denotes the forward-only replication. Scores are from one training seed per reported neural configuration; no seed dispersion is available. Table 4 gives the 100k-step evaluation, and Table 5 summarizes the broader 200k-step metric suite. Means use the aggregation described in Section **??**.

Table 4: **Forward response prediction after 100k training steps.** PD: Pearson-Δ; dir: DE direction agreement; lfc: DE LFC Spearman on observed significant genes; recall: significant-gene recall. Scores average conditions within each cell line, with equal-weight means across lines. Higher values indicate better performance, and bold marks the larger score within each pair. GenInv improves PD, direction agreement, and LFC correlation on all five lines; recall improves on three.

Cell line	Model	pd ↑	dir ↑	lfc ↑	recall ↑
C32	State-ST	0.9210	0.8993	0.5863	0.7345
	GenInv	**0.9263**	**0.9142**	**0.6498**	**0.7415**
HOP62	State-ST	0.9520	0.9186	0.6894	0.7776
	GenInv	**0.9543**	**0.9295**	**0.7219**	**0.7905**
HepG2-C3A	State-ST	0.9113	0.8850	0.5982	**0.6844**
	GenInv	**0.9194**	**0.8950**	**0.6201**	0.6809
Hs 766T	State-ST	0.9188	0.9026	0.5818	**0.6978**
	GenInv	**0.9237**	**0.9139**	**0.6215**	0.6977
PANC-1	State-ST	0.9343	0.9093	0.6306	0.7422
	GenInv	**0.9371**	**0.9184**	**0.6536**	**0.7563**
Mean	State-ST	0.9275	0.9030	0.6173	0.7273
	GenInv	**0.9322**	**0.9142**	**0.6534**	**0.7334**

Table 5: **Response agreement and gene-selection performance at 200k steps.** Equal-weight means over five evaluation cell lines. Arrows indicate the preferred metric direction; bold marks the better score. DE count Spearman measures agreement in the number of significant genes across conditions, whereas DE LFC Spearman measures gene-level effect agreement.

Metric	State-ST	GenInv
Pearson-Δ ↑	0.9220	**0.9312**
MSE-delta ↓	5.0e-4	**4.0e-4**
Discrimination (cosine) ↑	**0.9975**	0.9973
DE direction match ↑	0.8982	**0.9171**
DE LFC Spearman (sig.) ↑	0.6121	**0.6646**
DE sig. gene recall ↑	0.7012	**0.7285**
DE count Spearman ↑	0.8966	**0.9300**
PR AUC ↑	0.6580	**0.6651**
ROC AUC ↑	0.8330	**0.8402**
Overlap@100 ↑	0.2190	**0.2230**
Precision@100 ↑	0.2191	**0.2232**

B PER-CELL-LINE METRICS

Tables 6 and 7 report the complete metric suite. The four primary metrics improve in all five lines at 200k steps. Overlap and precision exhibit mixed changes across lines and cutoffs, while discrimination scores approach their upper bound. Both models predict more significant genes than observed; recall is therefore interpreted jointly with ranking performance and effect-size agreement.

Table 6: **Cell-line-specific performance of the forward-only State-ST baseline.** Evaluation at 200k steps using the metric definitions in Appendix F. Arrows indicate the preferred direction; DE #sig reports the mean number of significant genes per condition. Count Spearman is computed across conditions within each cell line.

Metric	C32	HOP62	HepG2-C3A	Hs 766T	PANC-1
Pearson-Δ ↑	0.9174	0.9495	0.9081	0.9063	0.9286
MSE ↓	0.0008	0.0004	0.0006	0.0006	0.0002
MAE ↓	0.0082	0.0061	0.0090	0.0073	0.0044
MSE-Δ ↓	0.0007	0.0004	0.0005	0.0006	0.0002
MAE-Δ ↓	0.0078	0.0057	0.0086	0.0070	0.0040
Discrim. (L1) ↑	0.9833	0.9960	0.9906	0.9944	0.9960
Discrim. (L2) ↑	0.9841	0.9969	0.9936	0.9951	0.9960
Discrim. (cosine) ↑	0.9937	0.9986	0.9981	0.9984	0.9986
Overlap@50 ↑	0.1126	0.1495	0.1150	0.1089	0.1254
Overlap@100 ↑	0.2138	0.2491	0.2247	0.1929	0.2146
Overlap@200 ↑	0.3626	0.3808	0.3574	0.3179	0.3423
Overlap@500 ↑	0.5112	0.5338	0.4338	0.4380	0.5000
Overlap@N ↑	0.5304	0.5659	0.4373	0.4515	0.5404
Precision@50 ↑	0.1126	0.1494	0.1150	0.1089	0.1254
Precision@100 ↑	0.2138	0.2490	0.2248	0.1930	0.2146
Precision@200 ↑	0.3628	0.3809	0.3578	0.3189	0.3427
Precision@500 ↑	0.5132	0.5372	0.4389	0.4424	0.5098
Precision@N ↑	0.5312	0.5623	0.4441	0.4551	0.5528
DE direction match ↑	0.8988	0.9139	0.8771	0.8949	0.9064
DE LFC Spearman (sig.) ↑	0.6050	0.6854	0.5711	0.5723	0.6269
DE sig. gene recall ↑	0.7132	0.7592	0.6368	0.6796	0.7172
DE count Spearman ↑	0.8925	0.8985	0.8528	0.8998	0.9392
DE #sig (real)	426.1	536.5	281.4	375.0	506.5
DE #sig (pred)	556.0	702.1	400.6	544.5	633.5
PR AUC ↑	0.6837	0.7402	0.5701	0.6028	0.6931
ROC AUC ↑	0.8424	0.8510	0.8167	0.8168	0.8382

Table 7: **Cell-line-specific performance with generative inverse supervision.** GenInv at 200k steps, evaluated on the same conditions with the same metric definitions and aggregation as State-ST in Table 6. Arrows indicate the preferred direction; DE #sig denotes the mean significant-gene count per condition.

Metric	C32	HOP62	HepG2-C3A	Hs 766T	PANC-1
Pearson-Δ ↑	0.9274	0.9543	0.9192	0.9187	0.9362
MSE ↓	0.0007	0.0003	0.0005	0.0005	0.0002
MAE ↓	0.0078	0.0057	0.0085	0.0067	0.0041
MSE-Δ ↓	0.0006	0.0003	0.0005	0.0005	0.0001
MAE-Δ ↓	0.0073	0.0054	0.0080	0.0066	0.0038
Discrim. (L1) ↑	0.9828	0.9958	0.9916	0.9936	0.9966
Discrim. (L2) ↑	0.9828	0.9962	0.9940	0.9946	0.9965
Discrim. (cosine) ↑	0.9928	0.9985	0.9977	0.9984	0.9988
Overlap@50 ↑	0.1231	0.1601	0.1141	0.1019	0.1208
Overlap@100 ↑	0.2309	0.2704	0.2202	0.1818	0.2120
Overlap@200 ↑	0.3799	0.4026	0.3531	0.2960	0.3363
Overlap@500 ↑	0.5228	0.5471	0.4259	0.4076	0.4838
Overlap@N ↑	0.5404	0.5771	0.4288	0.4202	0.5211
Precision@50 ↑	0.1230	0.1601	0.1141	0.1019	0.1208
Precision@100 ↑	0.2308	0.2703	0.2205	0.1822	0.2121
Precision@200 ↑	0.3800	0.4026	0.3545	0.3009	0.3374
Precision@500 ↑	0.5227	0.5501	0.4366	0.4283	0.4985
Precision@N ↑	0.5346	0.5702	0.4433	0.4430	0.5426
DE direction match ↑	0.9167	0.9289	0.9013	0.9182	0.9205
DE LFC Spearman (sig.) ↑	0.6650	0.7216	0.6383	0.6417	0.6562
DE sig. gene recall ↑	0.7349	0.7832	0.6681	0.7089	0.7475
DE count Spearman ↑	0.9247	0.9335	0.8920	0.9445	0.9555
DE #sig (real)	426.1	536.5	281.4	375.0	506.5
DE #sig (pred)	574.5	722.1	426.1	586.4	675.7
PR AUC ↑	0.6928	0.7494	0.5775	0.6027	0.7029
ROC AUC ↑	0.8501	0.8602	0.8258	0.8221	0.8430

C CONDITION-LEVEL AND TRAINING DIAGNOSTICS

This section reports condition-level DE comparisons and validation trajectories. Figure 3 requires both Pearson-Δ and DE AUROC; Hs 766T treated with 0.5 μM encorafenib lacks a DE record and is excluded from that plot, but remains in the table's Pearson-Δ aggregate. The 100k DE plots use the intersection of conditions with at least five observed significant genes and finite scores for State-ST, GenInv, and the shuffled-label control. Each scatter plot shows all eligible pairs. Dark curves connect mean differences in 11 equal-width baseline-score bins between the first and 99th percentiles; bins with at most five points are omitted from the curve, not from the scatter.

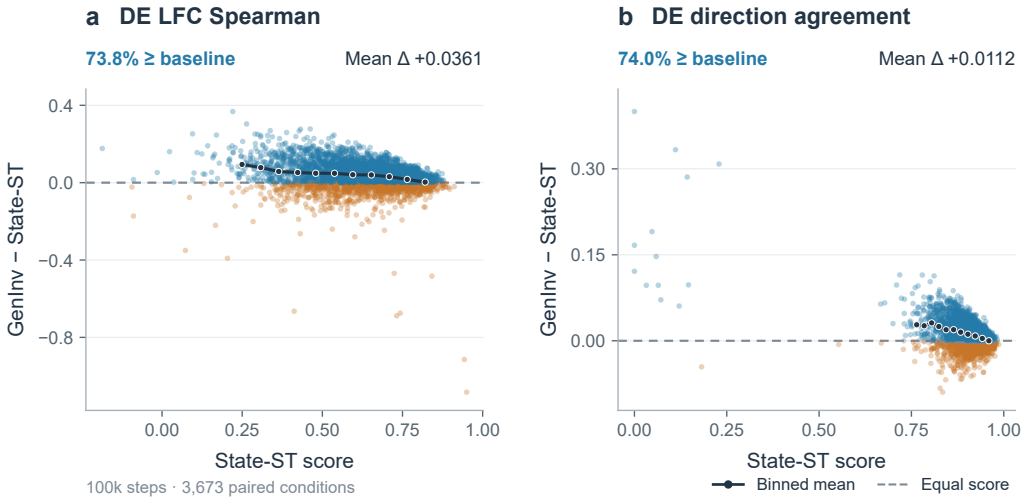

Figure 5: **Gene-level response gains over forward-only supervision.** GenInv versus State-ST at 100k steps on 3673 cell-line–perturbation pairs. GenInv improves **(a)** DE LFC Spearman in 73.8% of pairs (mean gain 0.0361) and matches or exceeds **(b)** direction agreement in 74.0% (6.4% ties; mean gain 0.0112). Blue/amber denotes positive/negative paired differences; dark curves denote binned means. All eligible prediction cases are shown. Each model represents one training run.

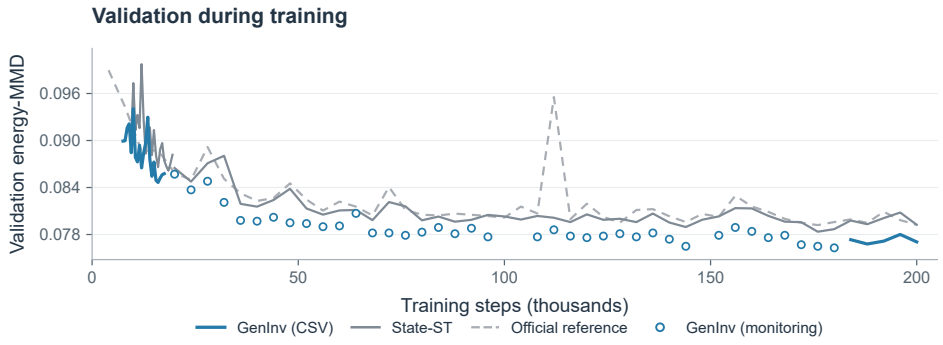

Figure 6: **Validation dynamics under forward-only and inverse-supervised training.** Energy-MMD loss for State-ST, GenInv, and the released State reference. Lines connect observations within retained CSV segments; open circles show intermediate GenInv monitoring records. Missing segments are not interpolated, and trajectories are unsmoothed. Final recorded losses are 0.0771 for GenInv and 0.0792 for the released reference. Each trajectory represents one run; State-ST is the matched baseline, while the released model provides contextual validation evidence.

D INVERSE-LABEL CONTROLS

The training-time control shuffles labels in the inverse branch while retaining correct labels for the forward task. Table 8 gives the four-decimal values underlying Table 2. At 100k steps, the shuffled-label model has Pearson-Δ 0.9323 versus 0.9322 for GenInv. Their near equality is a descriptive observation, not a statistical equivalence test. Correct labels yield larger DE direction and LFC scores.

Table 8: **Forward prediction under correct and shuffled inverse conditioning.** Equal-weight means over five cell lines at 100k steps, reported at four-decimal precision. The shuffled-label control retains correct forward labels. Higher values indicate better performance; the results correspond to Table 2.

Model	Pearson-Δ	DE AUROC	DE direction	DE LFC
State-ST	0.9275	0.8340	0.9030	0.6173
GenInv, shuffled labels	0.9323	0.8390	0.9081	0.6333
GenInv, correct labels	0.9322	0.8371	0.9142	0.6534

The separate test-time probe changes the label of an already trained model. We observe a forward correlation decrease from approximately 0.83 to 0.34–0.38 on C32 and PANC-1. Reverse reconstruction correlation changes from 0.871 to 0.868 and from 0.865 to 0.865. These are diagnostic reconstruction probes and should not be substituted for the five-line forward benchmark scores.

E EXPLORATORY ALTERNATIVES AND REPLICATION CONTEXT

Cycle consistency. An optional round-trip loss was explored,

$$\begin{aligned}\mathcal{L}_{\text{cycle}} = \tfrac{1}{2} D_E\big(f(g(X_{\text{pert}}, p), p), X_{\text{pert}}\big) \\ + \tfrac{1}{2} D_E\big(g(f(X_{\text{ctrl}}, p), p), X_{\text{ctrl}}\big).\end{aligned} \tag{7}$$

The reported GenInv model fixes its weight to zero. Exploratory runs showed no measurable improvement from adding this term, with increased computation from the composed passes. These runs are not a matched final-checkpoint comparison.

Disjoint supervision by perturbation. An exploratory variant assigned 20% of perturbations to inverse-only training and 80% to forward-only training. This variant performed worse than using both tasks on the same sampled data. This comparison changes the allocation of direct supervision, so it is distinct from shuffling the inverse labels while keeping all forward supervision.

Discriminative auxiliary tasks. Three variants predicted the perturbation label from transformer hidden-state differences, output expression deltas, or a reinforced output-level variant. These classifiers saturated rapidly without downstream benefit. One output-level variant had a reported 15% win rate at 20k steps in the earlier five-line setting. Because this historical comparison is not the full-corpus 200k-step experiment, we retain it as exploratory evidence rather than a directly comparable final ablation.

Variance anchoring. An earlier variant included an output-variance regularizer. An initial attribution exercise associated 84.6% of its gain with variance anchoring; subsequent controlled comparisons found that the dual component supplied no benefit in that configuration. The final GenInv objective excludes variance anchoring. This earlier configuration is separate from the shared generative inverse objective evaluated in the main comparison.

Training-corpus scale. An earlier five-cell-line training run on approximately 13M cells reached validation loss near 0.17 and Pearson-Δ between 0.43 and 0.50. The main experiments instead use the training portion of the full 50-line corpus. The forward-only run has validation loss 0.0981 at 4k steps, close to the released run's 0.099. Changing the corpus changes both scale and context coverage, so this history does not isolate either factor's causal contribution.

Published cross-model scores. For context, PerturbDiff reports Tahoe results (Yuan et al., 2026): Pearson-Δ of 0.686 (scratch) and 0.648 (fine-tuned), versus 0.535 for its State comparison; direction agreement of 0.734 and 0.641, versus 0.665; and DE LFC Spearman of 0.445 and 0.691, versus 0.644. The released State checkpoint obtains approximately 0.89 Pearson-Δ under our Cell-Eval evaluation. These results use different condition subsets and processing protocols. They are retained as context and are not ranked against GenInv or used to calculate improvements. The direct controlled comparison is the locally trained State-ST baseline.

Table 9: **Response estimation and DE selection for structured linear and generative baselines.** Four-decimal scores corresponding to Table 3, averaged equally across five evaluation cell lines. Linear and generative VAE models are fitted on the other 45 lines; neural models follow the official few-shot protocol and are evaluated at 200k steps. The mean predicted significant-gene count is reported alongside recall and DE AUROC to characterize overcalling.

Method	pd ↑	dir ↑	lfc ↑	recall ↑	AUROC ↑	#DE sig. pred
Control mean	0.0224	0.5012	0.0031	0.0000	0.5000	0
Additive response	0.4282	0.6681	0.2605	0.5706	0.4935	1168
Diagonal linear	0.6149	0.8035	0.4364	0.8227	0.4992	1605
Scalar-slope ridge	0.5362	0.7768	0.3738	0.7114	0.5051	1380
scGen	0.2805	0.6551	0.2271	0.9530	0.5500	1752
CPA	0.3034	0.6485	0.2396	0.9499	0.5521	1740
chemCPA	0.2260	0.6247	0.2004	0.9517	0.5591	1721
State-ST	0.9220	0.8982	0.6121	0.7012	0.8330	567
GenInv	0.9312	0.9171	0.6646	0.7285	0.8402	597

F METRIC DEFINITIONS AND INTERPRETATION

The evaluation uses the experiment's memory-efficient expression and DE metric scripts. The definitions below specify their behavior, including control references and gene support. Let $\bar{p}$ and $\bar{r}$ denote predicted and observed perturbed means, and let $\bar{c}_{\mathrm{pred}}$ and $\bar{c}_{\mathrm{real}}$ denote predicted and observed control means. Expression values are clipped to $[0, 14]$ before aggregation. Scores are computed per condition, averaged within each cell line, and then averaged equally across the five lines.

Expression agreement. Pearson-Δ is the Pearson correlation between $\bar{p}-\bar{c}_{\mathrm{pred}}$ and $\bar{r}-\bar{c}_{\mathrm{real}}$. MSE and MAE measure errors in mean expression; MSE-Δ and MAE-Δ measure errors in these response vectors. The predicted and observed control means are computed separately. This distinction matters when a predictor changes the control distribution.

Condition discrimination. Discrimination scores compare each predicted response with the observed responses from all candidate conditions using L_1, L_2, or cosine distance. For K conditions and a zero-based rank r of the matching condition, the score is $1 - r/K$. High discrimination scores and accurate gene-level effects are different properties; a nearly saturated discrimination score can leave substantial differences in DE recovery.

Differential expression. The DE pipeline uses `pdex` in reference mode with the DMSO control, log1p input, Wilcoxon rank-sum testing, and Benjamini–Hochberg correction. A gene is significant when its adjusted p-value is strictly below 0.05. Direction agreement joins observed significant genes to the full predicted gene table and compares the signs of their LFCs; predicted significance is not required. LFC Spearman also uses observed significant genes, filling absent predicted LFCs with zero. Undefined correlations are omitted from the condition mean. Recall is the intersection of observed and predicted significant sets divided by the observed set size, with zero assigned when the latter is empty.

DE PR AUC is computed as average precision. ROC AUC uses the same binary labels, defined by observed significance, and ranking scores $-\log_{10}(q_{\mathrm{pred}})$, with predicted adjusted p-values clipped

to $[10^{-10}, 1]$ and missing values set to one. Conditions containing only one label class are omitted. The metric `de_spearman_sig`, labeled *DE count Spearman* in our tables, correlates the numbers of significant genes across conditions. It measures agreement in response breadth, not correlation of gene-level effect sizes.

For overlap and precision, significant genes are ranked by absolute LFC. Let R_a and P_a denote the first a genes in the observed and predicted ranked lists, respectively. With $a = \min(k, |R|)$, overlap@k is $|R_a \cap P_a|/a$. Precision@k instead uses $a = \min(k, |P|)$ in both lists and in the denominator. At cutoff N, these denominators become $|R|$ and $|P|$, respectively. Empty denominators yield zero. These are the implemented conventions for $k \in \{50, 100, 200, 500\}$.

Positive-call counts. The per-line observed mean counts are 426.1, 536.5, 281.4, 375.0, and 506.5, averaging 425.1. Corresponding predicted counts average 567.3 for State-ST and 597.0 for GenInv. The diagonal baseline's reported count of 1605 is approximately 3.8 times the observed mean. High recall accompanied by many more positive calls and near-chance AUROC indicates overcalling rather than accurate gene selection. These diagnostics describe prediction quality; individual cells are not treated as independent replicates for model-to-model significance testing.

G Training and implementation details

Table 10 gives settings checked against the model code, saved configurations, and checkpoint metadata. Inputs use precomputed log-normalized HVG vectors (`obsm/X_hvg`); no separately pretrained State embedding model is used. The predictive layers are trained from scratch. The inverse perturbation encoder is copied from the forward encoder at initialization and subsequently optimized independently.

Table 10: **Architecture and optimization settings for the matched neural comparison.** State-ST and GenInv share the forward architecture and optimization configuration. GenInv additionally uses an inverse perturbation encoder and inverse loss. Stored parameter counts include auxiliary and inactive modules.

Component	Value
Transformer	8 layers, 12 heads, hidden 768, intermediate 3072
Attention	bidirectional; attention dropout 0; no rotary embeddings
Cell set length	256 cells
Gene space	2000 HVGs (`obsm/X_hvg`, log-normalized)
Perturbation encoding	one-hot (1138 conditions) $\rightarrow$ linear $\rightarrow$ 768
Cell projector / output head	linear 2000 $\rightarrow$ 768 / linear 768 $\rightarrow$ 2000 + ReLU
Inverse encoder (GenInv only)	independent copy; +874,752 parameters
Stored parameters	108.69M (State-ST) / 109.57M (GenInv)
Objective	energy-MMD; GenInv adds $\lambda_{\text{dual}} = 1.0$ dual term
Optimizer	Adam, lr 3×10^{-4}, weight decay 0
Adam moments / epsilon	$(0.9, 0.999)$ / 10^{-8}
Batch size	64 cell sets (=16,384 cells)
Steps	200,000; val/ckpt every 4,000
Gradient clipping	10.0
Hardware	1$\times$ NVIDIA A100-80GB per run

Prediction layers and auxiliary modules. Equation 2 gives the evaluated HVG prediction path. A separate stored gene decoder, with widths $2000 \rightarrow 1024 \rightarrow 1024 \rightarrow 512 \rightarrow 2000$ and dropout 0.1, is not the output returned as `preds` for these evaluations. The input and output projections and perturbation encoders each have one linear layer. The transformer's configured attention and hidden dropout values are zero; the model-level dropout value of 0.1 should not be confused with attention dropout.

Training schedule and runtime. The recipe uses gradient clipping at norm 10, seed 42, 200k steps, and validation and checkpointing every 4k steps. State-ST and GenInv use 24 data workers;

the shuffled-label run uses 20. Although the YAML files contain a weight-decay field of 5×10^{-4}, the optimizer constructor does not consume it: the saved optimizer states confirm Adam with zero weight decay. GenInv adds a second training pass over the already loaded populations, while inference retains one forward task. Interrupted logs and short-window throughput measurements do not provide a controlled end-to-end runtime comparison, so we make no quantitative training-cost claim.

Checkpoint reporting. The main response comparison uses the fixed 200k-step endpoint for both models; the shuffled-label comparison uses 100k-step checkpoints. Their step counts are verified from saved checkpoint metadata. Intermediate checkpoints and per-condition observations from the same run are not independent training replicates. In the shuffled-label implementation, a random permutation reassigns perturbation labels across cell sets in a training batch; fixed points and repeated labels are possible.

Few-shot partition. The supplied on-disk corpus contains 100,648,790 cells in 50 HDF5 files. The split file designates validation and test treatment labels within five lines; all remaining labels are assigned to training. Table 11 counts labels actually present in those files. The training column includes the DMSO control. These are corpus-level label counts before dataloader filtering and control reuse, rather than counts of independently evaluated conditions.

Table 11: **Condition-label partition within the five evaluation cell lines.** Counts refer to labels present in the stored corpus under the official few-shot split. Training includes the DMSO control and all remaining labels within these lines, together with the other 45 cell lines. Validation and test labels are held out within each listed line.

Cell line	Training (including control)	Validation	Test
C32	305	56	777
HOP62	304	56	778
HepG2-C3A	304	56	778
Hs 766T	304	56	778
PANC-1	304	56	777