LazyNet: Interpretable ODE Modeling of Sparse CRISPR Single-Cell Screens Reveals New Biological Insights

Abstract

We present LazyNet, a compact one-step neural-ODE model for single-cell CRISPR A/I that operates directly on two-snapshot ("pre → post") measurements and yields parameters with clear mechanistic meaning. The core log–linear–exp residual block exactly represents multiplicative effects, so synergistic multi-locus responses appear as explicit components rather than opaque composites. On a 53k-cell × 18k-gene neuronal Perturb-seq matrix, a three-replica LazyNet ensemble trained under a matched 1-hour budget achieved strong threshold-free ranking and competitive error (genome-wide r≈0.67) while running on CPUs, compared with transformer and state-space baselines trained on a single V100 with the same time cap. A T-cell screen included only for generalization showed the same ranking advantage under the identical evaluation pipeline. Beyond prediction, LazyNet exposes directed, local elasticities; averaging Jacobians across replicas produces a consensus interaction matrix from which compact subgraphs are extracted and evaluated at the module level. The resulting networks show coherent enrichment against authoritative resources (e.g., large-scale co-expression and curated functional associations) and concordance with orthogonal GPX4-knockout proteomes, recovering known ferroptosis regulators and nominating testable links in a lysosomal–mitochondrial–immune module.

1. Introduction

Single-cell RNA sequencing (scRNA-seq) has transformed transcriptomics by revealing the rich cell-to-cell heterogeneity masked in bulk averages, uncovering rare sub-populations, transitional states, and lineage hierarchies across tissues and species [1-2]. Yet each cell is still frozen in time—captured under one environmental condition and at a single instant—so most temporal trajectories and stimulus-response behaviours remain invisible. Systematically probing these dimensions demands direct perturbations: transcriptional read-outs following targeted genetic edits illuminate how regulatory circuits maintain identity, pinpoint druggable control nodes, and expose synergistic gene pairs that can amplify combination therapies or guide precise cell-engineering strategies [18-31]. CRISPR activation or inhibition screens joined with scRNA-seq ("Perturb-seq") meet this need by bar-coding guide RNAs so that every sequenced cell records its own perturbation history, enabling thousands of knock-outs, knock-downs, or activations to be multiplexed in one experiment [3-4, 32]. The resulting high-dimensional maps of genotype-to-phenotype relationships have already sharpened target validation pipelines, suggested rescue strategies for disease mutations, and provided training data for machine-learning models that aspire to predict unseen edits. These advances have converged in

ambitious blueprints for an in-silico "AI virtual cell" (AIVC) capable of simulating cellular behaviour across modalities, contexts, and perturbations [5].

Several recent frameworks aim to predict perturbation responses from single-cell data, including transformer-based models (e.g., scGPT), large foundation models (e.g., CellFM), and graph-based approaches (e.g., GEARS) [6–8]. While powerful, these pipelines typically require multi-GPU hardware, substantial storage, and/or curated pathway priors, and their latent representations are difficult to map onto concrete biochemical rate laws — complicating mechanistic interpretation in routine lab settings. Moreover, recent analyses indicate that deep models do not universally outperform simpler baselines for gene-perturbation prediction, underscoring the need for methods that are both accurate and practical [49]. Neural-ODE approaches offer improved mechanistic fidelity but often assume fixed functional forms or entail numerically intensive integration over high-dimensional states, which limits tractability at transcriptome scale [12–16,50–51].

LazyNet addresses this gap with a neural ODE that embeds a log–linear–exp residual block inside a single explicit-Euler step. Working in log space compresses multiplicative gene–gene effects into a sparse, directly interpretable rate matrix, avoiding attention-style quadratic costs and heavy ODE solvers. Crucially, the model is designed for CPU-level efficiency (no GPU required), operates directly on count matrices without external pathway priors, and matches the common two-snapshot CRISPR A/I design by treating the pre- and post-perturbation profiles as a finite-difference sample of the underlying dynamics.

In this study we show that LazyNet attains state-of-the-art predictive accuracy on a large Perturb-seq matrix while training in hours on a single 24-core CPU node, and that its elasticity-based analysis yields compact, biologically coherent subnetworks. Applied to ferroptosis, the framework recovers established regulators and nominates a lysosomal–mitochondrial–immune module for experimental follow-up, illustrating how CPU-tractable, interpretable modeling can turn routine perturbation screens into causal hypotheses at transcriptome scale.

2. Materials and Methods

2.1 LazyNet Architecture

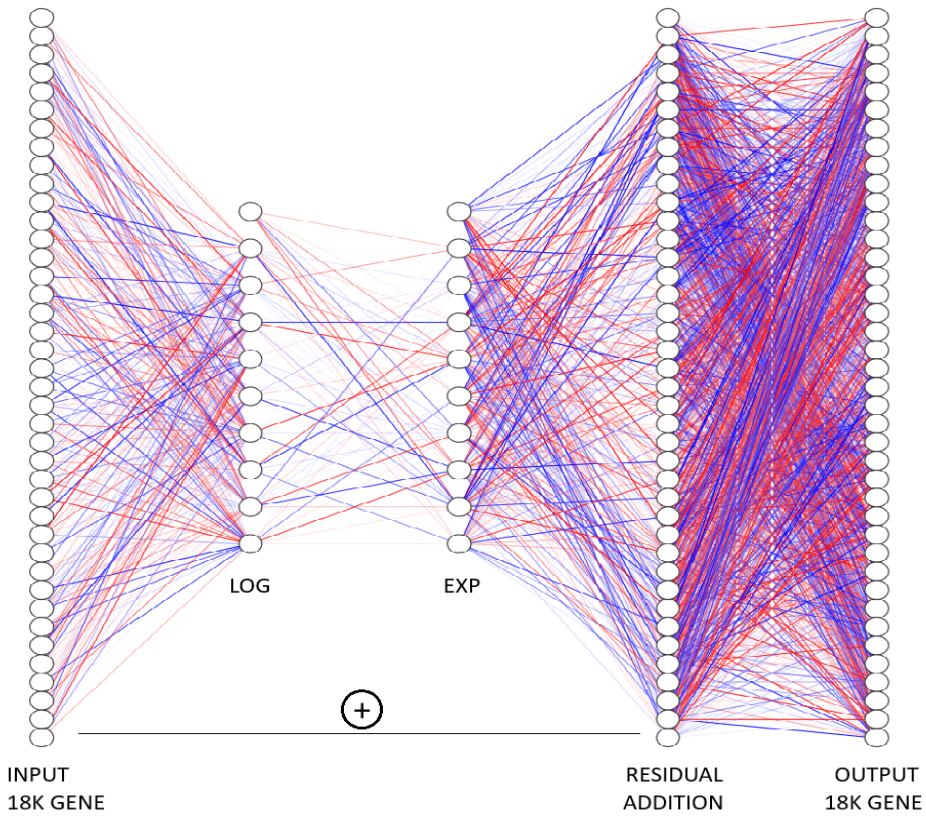

Figure 1. The architecture of LazyNet. All log and exp operations are elementwise; inputs are library-size normalized and shifted by a small pseudocount ε to ensure positivity.

From Figure1, the full gene-expression vector (INPUT) is propagated through a two-stage transformation composed of a logarithmic layer (LOG) followed by an exponential layer (EXP). These twin layers realise multiplicative biology-inspired interactions in log space while retaining computational simplicity. The transformed signal is added back to the original input via a residual skip connection (black bar and "+"), producing an updated vector of identical dimensionality (OUTPUT). Stacking such blocks corresponds to explicit Euler integration of an underlying ordinary differential equation, allowing LazyNet to model gene-regulatory dynamics within a ResNet framework while continuously adjusting weights during training [17].

LazyNet has already been applied successfully to both synthetic ODE benchmarks and real-world dynamical systems. Its core idea is to embed a log → linear → exp residual block inside explicit Euler approximation, so that every weight acts as a direct ODE rate constant and the network learns interpretable dynamics with minimal parameters.

An ODE of the form can be discretised with forward Euler:

$$f_{\{n+1\}} = f_n + F(f_n)\,\Delta t \tag{1}$$

LazyNet represents the rate law F with a single log–linear–exp (LLE) block:

$$F_{L}azyNet(x) = \exp(W\square \cdot \log(W\square x + b\square) + b\square) \tag{2}$$

$$f_{\{n+1\}} = f_n + F_LazyNet(f_n)\ \Delta t \tag{3}$$

Here $x \in \mathbb{R}+\square$; W_1, W_2 are learnable weight matrices and b_1, b_2 are biases. Taking the logarithm converts multiplicative gene–gene interactions into additions, allowing a single neuron to encode an entire monomial exactly, while the outer exponential restores the original scale. More details and examples can be seen in Supplement S4.

During training, each observed state fn is provided as input, the network predicts fn+1, and the prediction is compared with ground truth. With sufficiently varied samples V and adequate temporal coverage T, the learned mapping approaches the true governing equations:

$$\lim_{\{V\to\max,T\to\infty\}} F_{LazyNet} = F_{actual} \tag{4}$$

Because the ODE is Markovian, every update depends only on the current state, not on the full history; LazyNet therefore learns the system dynamics efficiently and with minimal parameter count.

In a CRISPR-A/I screen, the before-and-after expression profiles can be treated as Euler approximation of the underlying ODE—thereby sampling the true regulatory function Factual directly.

2.2 Theoretical Rationale for Efficiency

Most biochemical rate laws can be written as a finite sum of monomials of the form $c \cdot x1\alpha1 \ldots x\square\alpha d$. A representation that reproduces each monomial exactly captures the full reaction space.

LazyNet achieves this with a single log–linear–exp (LLE) block(2), where $x \in \mathbb{R}+\square$, $W \in \mathbb{R}K\times d$, and $b \in \mathbb{R}K$. Choosing one row of W equal to the integer exponent vector α and the corresponding bias $b = \log\square c$ reproduces the monomial exactly. Hence an entire K-term rate law needs only $K(d\square+\square 1)$ trainable scalars.

By contrast, a degree-N Taylor polynomial on the same positive domain requires ~ $\log\square 1\varepsilon d$ parameters to reach uniform error ε. Thus the parameter ratio grows super-polynomials with

both dimension d and desired accuracy. Fewer parameters translate to fewer multiply–add operations per mini-batch and a smaller memory footprint, directly accelerating training.

Statistically, concentrating capacity onto a small set of mechanistically meaningful weights improves the effective signal-to-noise ratio: each observation updates compact exponent and coefficient parameters rather than being diffused across a large dictionary. Empirically, this manifests as faster convergence at larger learning rates, better small-sample generalization, and more stable optimization under residual connections.

Importantly, the multiplicative form encodes synergistic (higher-order, multi-locus) interactions as single monomials rather than as emergent products of many additive units. This is not only an efficiency gain; it directly addresses a substantive modeling need by making synergic effects first-class, precisely parameterized objects that can be learned robustly from limited data.

Transformer-based baselines such as scGPT and state-space models such as RetNet avoid the explicit coefficient blow-up of polynomial bases but incur other costs [6–7]. Attention layers scale as $O(L^2)$ in sequence length L, while state-space kernels move toward linear time yet introduce millions of additional parameters. Neither family encodes a monomial in closed form; multiplicative gene–gene terms are learned implicitly via gradient descent, typically demanding more data and iterations. In contrast, the LLE map captures multiplicative structure exactly with a forward pass that is linear in d (and in the retained term count K), yielding a favorable accuracy–efficiency–interpretability trade-off.

For AIVC analyses specifically, prior work indicates that LLM-based frameworks (e.g., scGPT, CellFM, GEARS) are not tailored to the dynamics of multi-perturbation, high-order synergy and often struggle to represent such effects faithfully [49]. Neural-ODE approaches can be accurate but usually require a pre-specified differential form—e.g., fixed symbolic templates or universal Hill-type non-linearities [12–16, 50–51]—which constrains adaptability and can produce models that are numerically correct yet mechanistically misspecified. LazyNet retains an ODE viewpoint without fixing the functional form: placing the LLE block inside an explicit-Euler residual update allows approximation of a broad class of interaction terms, so synergy is captured through learned exponents and weights rather than hand-engineered equations, improving both flexibility and mechanistic interpretability.

2.3 scRNA Data Preparation and Gene Selection

Neuronal Perturb-seq (GSE152988)

We analyze GSE152988, a CRISPR activation/interference screen in 53,495 human iPSC-derived neurons (67,077 genes including non-coding RNAs) with two expression snapshots per cell (a pre-state "time-0" and a post-perturbation state) [33]. With no intermediate measurements, we

treat the post-state as a single explicit update from the baseline, matching the minimal ODE setting used throughout this work. No external dynamics or pathway priors were introduced.

To construct "time-0" without leakage, we do not edit individual cells. Instead, we first compute an averaged baseline cell from library-size–normalized non-targeting controls (per batch/stratum as applicable). For each perturbation targeting gene g, the time-0 input is this averaged baseline vector with a single edit at g to encode the intervention: for CRISPRi/KO we set g to a small floor; for CRISPRa/OE we set g to the observed post-perturbation level of g (estimated across the perturbed cells), while all other genes remain at control baselines. Because only the scalar for the targeted gene is updated—and never any multi-gene pattern or any individual test cell—this captures the intended one-gene intervention while preventing multigene peeking into the post-state.

After time-0 construction, we restrict to 18,152 Ensembl GRCh38 protein-coding genes for modeling and evaluation. Cells are grouped by (target gene × gRNA) (and by batch when available), and whole groups are assigned to train/validation/test at 80%/10%/10% to avoid guide- or batch-specific leakage across splits. On the working expression scale used for modeling, the held-out neuronal iPSC test targets have mean 0.6082 and standard deviation 3.212 (min 0, max 1203) as the data landscape.

Because only two snapshots are available, absolute rates are not identifiable; the product of step size and elasticity is what can be learned. Accordingly, downstream evaluation emphasizes one-step prediction errors and relative signals (e.g., rankings) that are invariant to monotone rescaling of the latent time step. When dense time series are available, a multi-step fit resolves Δt and improves identifiability (see Discussion and [14-17]).

T-cell guide-capture screen (GSE190604)

We further use a large human T-cell CRISPR guide-capture dataset (10x Genomics RNA with per-cell guide calls) comprising 36,601 genes × 103,805 cells [54]. Expression matrices come from the "Gene Expression" count matrix and guide calls from the corresponding CellRanger export. Minimal QC is applied (genes detected in ≥3 cells; cells with ≥200 detected genes) to preserve a blind and reproducible comparison.

Each barcode is joined to its guide-call summary. Baseline cells (0/unknown guides) are used only to estimate time-0. Singlets (exactly one guide) form the training set; multiplets (≥2 guides) form the test set. After matching and filters, the split yields ~61k training cells and 28,707 test cells.

Counts are library-size normalized with a small pseudocount and log-transformed. We select the top 25% highly variable genes (HVGs) across cells (~one quarter of 36,601 genes) and carry this fixed HVG panel into all downstream modeling and evaluation. On this normalized log scale, the held-out test targets have mean ≈ 0.19 and SD ≈ 2.30 (min 0, max 2620), which we report to

orient error magnitudes in later sections.

2.4 Baseline Models

We compared LazyNet to two public baselines chosen to span the principal inductive biases for single-cell perturbation prediction without curated priors: a transformer (scGPT) and a state-space model (RetNet/CellFM) [6–8]. These baselines were selected because they (i) have maintained, open implementations with documented end-to-end pipelines; (ii) do not require curated cell-type labels or pathway priors (e.g., GEARS); (iii) are compatible with two-snapshot Perturb-seq inputs; and (iv) have prior use on closely related single-cell prediction tasks.

Together, scGPT and RetNet/CellFM provide architectural diversity—attention vs. state-space dynamics—yielding an appropriate, apples-to-apples contrast with LazyNet under a common data regime. In this pairing, scGPT serves as the widely adopted attention-based baseline, while RetNet/CellFM offers a competitive linear-time alternative within the state-space family.

2.5 Training Protocol

LazyNet is trained on our Intel Xeon CPU server Intel Xeon Gold 6126, 24 cores, 768 GB RAM; no accelerators. The main runs used a Huber loss ($\delta = 0.1$) for robustness to outliers and a three-replica LazyNet ensemble (independent seeds; arithmetic averaging at inference). Complete training hyperparameters, exact command lines, and seeds are provided in the replication package.

In the ablations, the model depth is the most sensitive knob: increasing the log-exp pair more than one led to numerical instabilities (NaNs) or higher error, consistent with overfitting in a setting where a one-step ODE is already a low-complexity mechanism and LazyNet is highly expressive. Width is less sensitive, but a moderate range (several hundred units) performs best; widths that are too small underfit, while very wide models tend to overfit or fail to converge usefully within the time budget. The additive residual path is essential for ODE fitting—removing it substantially degrades accuracy or prevents stable training. We also varied the loss family (Huber/L1/L2 and δ), mini-batch size, and integration step length. Within the ~1 h per-replica budget (3 replicas), batch size and step length had negligible effect on test metrics, whereas expanding LazyNet beyond moderate capacity often undertrained within the time cap, yielding lower accuracy despite more parameters.

scGPT and RetNet/CellFM are trained on a single NVIDIA V100 GPU with three CPU cores allocated. To match LazyNet's elapsed-time constraint, each baseline is capped at 1 hour of wall-clock time per dataset, inclusive of data loading and evaluation. We rely on the authors' documented defaults where applicable, enable mixed precision and gradient accumulation as needed to respect memory limits, and do not introduce curated priors or external labels. The last

checkpoint produced within the time cap is evaluated on the common test split.

Regression fidelity is assessed with RMSE, MAE, and Pearson's r on the modeling gene panel defined in §2.3. Threshold-free ranking quality is summarized by ROC-AUC and PR-AUC computed directly from continuous predictions. When reported, F1 (micro and macro) uses a single per-gene operating point learned once on the training data via a fixed quantile rule and held fixed for evaluation. All models use identical train/validation/test partitions, deterministic data loaders, fixed random seeds, and the same evaluation scripts.

2.6 Network Inference

To produce a directed, dynamical snapshot for network analysis, we averaged ten independently trained LazyNet checkpoints into a single ensemble and evaluated it on the full single-cell matrix (train + validation + test). From this ensemble we computed a Jacobian of absolute elasticities at a common baseline expression vector, restricted to the original 18,152 Ensembl protein-coding genes [34]. For each literature-anchored seed gene (seven total; Supplement 1A), we ranked both downstream (row) and upstream (column) elasticities, retained neighbors with $|\partial x_i/\partial x_j| \geq 0.001$ that also exceeded either the 95th or 99th percentile of that seed's ranked list, and then kept the top-N survivors. Because a single-baseline Jacobian yields few candidates per seed, we set the Benjamini–Hochberg threshold to $q = 1$ (effectively no FDR filtering) and report the filtered ranked sets. Evaluation proceeds in two parts: first, whether each seed's neighborhood recovers interactions reported in the literature; second, cross-seed recall, i.e., how often other seed genes appear among a given seed's ranked neighbors.

The goal is to recover directed, dynamical relationships consistent with the ODE-based formulation—edges that reflect functional influence under perturbation, not merely correlation. For this reason, GENIE3/GRNBoost/SCENIC-style pipelines are not used as head-to-head comparators here. They operate as subgraph generators that prioritize local correlation edges and typically require a manual regulator prior (e.g., TF lists) to seed the search [39 ,41, 43]. Their outputs are commonly undirected and tuned to rank local associations, rather than to assemble a globally coherent dynamic network. They also lack the external authority of curated/experimental resources; when a comparator is needed, databases such as STRING and ARCHS4 — grounded in empirical evidence — are more appropriate reference points than GENIE3-family predictions.

Even with authoritative databases, edge-level comparisons remain problematic: curated graphs are incomplete, context-dependent, and unevenly sampled, and the true functional wiring can vary substantially even within the same cell type and omic snapshot. We therefore evaluate at the module level, where the signal is more robust: given a set of curated modules (pathways, TF

regulons), we quantify module activity/enrichment induced by the inferred network (or by model-predicted responses) and test whether the recovered activity patterns align with known biology. This enrichment-based evaluation emphasizes coherent functional programs over individual edges, respects the dynamical and directed nature of our model, and provides a fairer, more stable basis for comparison across methods and references.

2.7 ARCHS4 co-expression validation

We annotate each directed edge ($i \rightarrow j$) with the ARCHS4 Pearson correlation $r(i,j)$. To control for confounding, enrichment is computed against a matched null: for each edge we sample matched pairs ($i' \rightarrow j'$) from strata that jointly bin genes by (i) control-cell detection/mean expression and (ii) out-/in-degree in the LazyNet elasticity graph (≈400 matched pairs per edge; 9,860 edges; 3,943,792 null pairs). We report fractions exceeding $|r| \geq \tau$ and the corresponding fold enrichment (Observed/Expected). For sign-aware checks, we multiply r by the LazyNet edge sign; because ARCHS4 is undirected/heterogeneous, we emphasize $|r|$ results and treat sign panels as exploratory.

2.8 STRING interaction concordance

STRING functional associations (human, version 12.0) were converted from protein identifiers to HGNC symbols with the official mapping file [36]. Isoform-specific and unmapped entries were removed. 32*4 LazyNet sub-graph edges were queried for exact symbol matches, and statistical significance of the overlap was assessed with a one-tailed hypergeometric test that treats the total number of possible unordered human gene pairs as the population size. Elasticity scores and STRING combined scores were each dichotomized at the 75th percentile to classify edges as "both-high", "STRING-only" or "LazyNet-only". Degree-preserving Maslov–Sneppen rewiring (1000 iterations) provided a null distribution for network-density comparisons.

2.9 Proteomics Validation

Pulsed-SILAC proteomes were re-analysed to evaluate 32*4 LazyNet sub-graph at the protein layer. For the tamoxifen-inducible GPX4-knockout time-course in mouse fibroblasts (PRIDE PXD050979), proteinGroups.txt (PXD050979.txt) was taken as input [47]. Reverse sequences and common contaminants were discarded; gene symbols were extracted, converted to upper-case HGNC format and duplicate entries were averaged, yielding 1026 proteins. Intensities for the light, medium and heavy SILAC channels (nine runs) were treated as independent samples, zeros set to missing and imputed with the MinProb left-shift approach ($\mu - 1.8\ \sigma$; width $0.3\ \sigma$) as implemented in Perseus [48]. Values were $\log_2$-transformed, median-centred per run and proteins quantified in at least three channels with non-zero variance were retained.

For the larger constitutive GPX4-knockout study in Pfa1 mouse embryonic fibroblasts (PRIDE

PXD040094) we used the combined_protein.tsv [46]. Gene symbols were cleaned as above and duplicate symbols collapsed. Columns labelled "MaxLFQ Intensity" defined fourteen quantitative channels covering wild-type and knockout states. Missing values were replaced with MinProb and the matrix was $\log_2$-transformed and median-centered. Proteins present in at least three channels and showing non-zero variance were kept.

For each dataset the cleaned protein matrix was intersected with the RNA-based LazyNet edge list. Gene pairs with at least three paired, finite observations were assigned a Spearman rank correlation coefficient. To control for hub bias a degree-matched null of 10k random edges was generated by stratified sampling across twenty-one degree bins (0-to-20+). Absolute correlation values of LazyNet edges were compared with the null using a two-sided Wilcoxon rank-sum test.

2.10 Experiment Layout

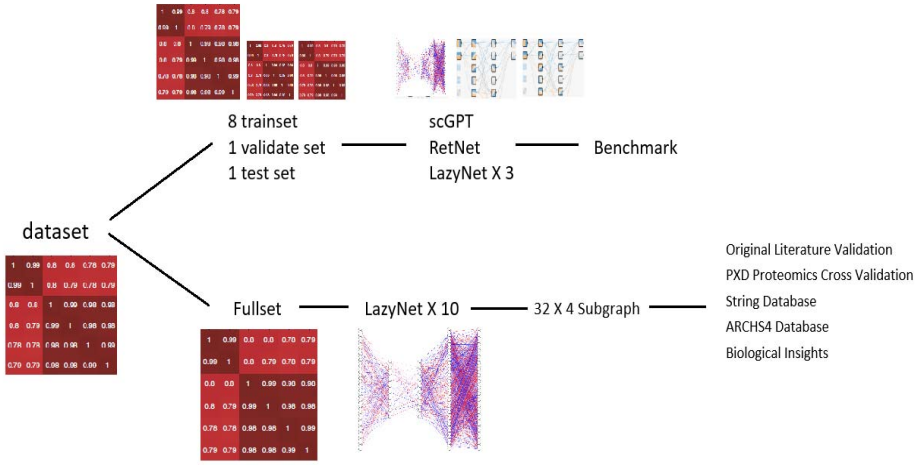

Figure 2. Benchmark-and-discovery workflow.

As shown in Figure 2, we start from a single pre-processed neuronal iPSC Perturb-seq matrix (library-size normalization followed by a small-pseudocount transform, $z=\log(x+\varepsilon)$) and run two parallel analyses.

The benchmark arm performs a one-time stratified split (8:1:1 train:validation:test) and trains scGPT, RetNet, and three LazyNet replicas on the training fold, tunes on validation, and reports performance on the locked test fold (RMSE, MAE, Pearson r; threshold-free AUCs where noted) [6–8]. A T-cell screen is included only in this benchmark arm (not shown in figure) to test cross-dataset generalization under the identical pipeline; our central goal remains the neuronal dataset.

The discovery arm fits ten independently seeded LazyNet models on the full neuronal matrix (train+validation+test), averages their Jacobians (restricted to 18,152 Ensembl protein-coding genes), and extracts a 32×4 breadth-first subgraph as a consensus network for interpretation.

This subgraph is evaluated against authoritative resources—STRING v12 and ARCHS4 [35–36]—and cross-omic coherence is assessed in two GPX4-knockout SILAC proteomes from mouse embryonic fibroblasts (PRIDE PXD040094, PXD050979) [46–47]. The validated interactions are then mined for mechanistic hypotheses and used to nominate follow-up perturbations.

3. Results

3.1 Model Benchmark

Table 1 summarizes head-to-head results for a neuronal iPSC Perturb-seq slice for later inference and a T-cell guide-capture slice (top-quartile HVGs) solely for generalization test.

Table 1 Comparison of training cost and predictive accuracy

Category	Metric	RetNet	scGPT	LazyNet
Neural iPSC	**Total Time (s)**	3603	3643	**2286**
	Device	GPU	GPU	**CPU**
	Peak VRAM/RAM	17/18 G	19/17 G	**40 G**
	Parameters (millions)	10.07	5.08	**9.58**
	Ensemble size	1	1	**3**
	RMSE	2.140	2.333	**2.376**
	MAE	0.515	0.637	**0.562**
	Pearson r	0.746	0.696	**0.680**
	ROC / PR AUC	0.784 / 0.806	0.591 / 0.587	**0.920 / 0.956**
	F1 micro / macro	0.834 / 0.437	0.888 / 0.514	**0.683 / 0.320**
T cell top quantile	**Total Time (s)**	3601	3668	**4405**
	Device	GPU	GPU	**CPU**
	Peak VRAM/RAM	8/31 G	5/30 G	**25 G**
	Parameters (millions)	5.42	4.1	**6.92**
	Ensemble size	1	1	**3**
	RMSE	0.3253	0.323	**0.356**
	MAE	0.1572	0.154	**0.130**
	Pearson r	0.7181	0.722	**0.666**
	ROC / PR AUC	0.809 / 0.866	0.876 / 0.940	**0.970 / 0.988**
	F1 micro / macro	0.754 / 0.444	0.538 / 0.231	**0.607 / 0.345**

On the neuronal iPSC slice, LazyNet attains the strongest threshold-free ranking quality with the highest ROC-AUC and PR-AUC among the compared methods, while training faster on CPU and using a mid-range parameter count. Its absolute regression errors (RMSE, MAE) and correlation r are slightly worse than the best baseline, and its F1 (micro/macro) is lower at the fixed per-gene operating point—an expected AUC–F1 trade-off when a single threshold is pre-declared per gene.

Together, the iPSC results indicate that LazyNet delivers the best ordering of perturbation effects under CPU constraints, with modestly higher regression error than the top baseline.

On the T-cell HVG slice, LazyNet again leads on ranking metrics with the highest ROC- and PR-AUC, and it achieves the best MAE, while its RMSE and Pearson r are lower than those of the strongest baseline and its F1 sits mid-pack at the fixed threshold (Table 1). This pattern—excellent ranking with conservative amplitudes—matches the expected behavior when per-gene thresholds are learned once on singlets and then applied to a broader test distribution of multiplets.

Interpreting AUC and F1 together clarifies these results. ROC-/PR-AUC capture threshold-free ranking and are insensitive to the exact operating point; F1 fixes a single per-gene threshold learned on training singlets, so it can move counter to AUC when calibration shifts or variance broadens between train and test. In our setting, this explains why LazyNet can lead decisively in AUCs yet not maximize F1 on every slice. All definitions and thresholding rules are held constant across models to keep comparisons fair.

Overall, across both slices and under CPU-only execution with a three-run LazyNet ensemble, the model offers the strongest ranking of perturbation effects and a favorable MAE profile on T-cells, while ceding some ground on RMSE and r to the best baseline on certain metrics. Because LazyNet instantiates a one-step ODE with log–exp residual dynamics, it exposes per-gene elasticities whose ensemble Jacobian yields directed, mechanism-level edges; we leverage this structure immediately in network inference, where the Jacobian is converted into a consensus subgraph and evaluated via module-level enrichment against curated resources.

3.2 Network Inference

Ten independently trained LazyNet replicas were averaged into a single ensemble, and Jacobian elasticities were queried around seven ferroptosis seeds (GPX4, PSAP, ACSL4, WIPI2, MTOR, NFU1, SOD2), selection details appear in Supplement S1A. A broader parameter sweep, summarized in Supplement S1B, identified five expansion regimes that were carried forward (Table 2).

Table 2: gene inference stats

Settings	Genes	Edges	Bench -hits	Recall†	Enrichment‡	Mean recall (per-seed)§
--top 2048 --depth 1 --pct 95	5227	8690	15	56%	1.93	0.11
--top 512 --depth 1 --pct 95	3344	4946	13	48%	2.61	0.08
--top 128 --depth 2 --pct 99	6761	21241	17	63%	1.69	0.29
--top 64 --depth 3 --pct 99	6843	19118	18	67%	1.77	0.4
--top 32 --depth 4 --pct 99	4641	11676	15	56%	2.17	0.26

† Recall = input 10, X / 27 reference genes recalled.

‡ Enrichment here means fold-over-random recovery of the 27 ferroptosis benchmarks; or, enrichment = observed hits / expected hits by chance.

§ Mean per-seed recall (single-seed start): the average fraction of the 27 benchmark genes recovered when each benchmark is used as the sole seed. For example, in the top=64 setting a single seed recalls ~40% (≈11/27) on average. Full per-seed values are provided in Supplement S1D.

In the broadest setting (512×1) each seed ballooned into a neighborhood of ~3000 genes—GPX4 alone recruited 1811—and yielded a mean per-seed recall of ~0.08 (≈2–3/27). Constraining the search to 64×3 reduced each subgraph to ~1000 genes and lifted the mean per-seed recall to ~0.40 (≈11/27). The most selective regime, 32×4, fused all seven seeds into a single graph of 4641 genes and 11676 edges while preserving a benchmark recall of 15/27 and nudging the hit-rate upward; its data are in Supplement S1C. Because only 27 of 18,152 genes are treated as positives (base rate ≈0.15%), these fractions are numerically small; viewed as enrichment, the 32 ×4 graph recovers 15 regulators versus ~6.9 expected by chance (~2.17×).

Evaluate all the gene seeds, per-seed recalls (Supplement S1D) highlight a practical strength: starting from just one ferroptosis gene, LazyNet typically retrieves ~10 additional benchmarks when run at top = 32, 64, or 128. In other words, minimal prior knowledge suffices to reconstruct much of the ferroptosis module—evidence of the method's consistency and day-to-day utility.

LazyNet recovers 56 % (15/27) of benchmark ferroptosis interactions in < 6 h on a single 24-core CPU node (768 GB RAM, no GPUs). By comparison, GENIE3 recovers ~33 % of gold-standard E. coli edges after analyzing > 100 microarrays and ~48 h on 32 cores [39-40]; GRNBoost recalls ~22 % in a 68 k-cell PBMC atlas [41-42]; and DeepFGRN reaches ~40 % but requires dense time-series data and multi-GPU training (4*A100, several days) [43]. Thus, LazyNet achieves state-of-the-art recall with orders-of-magnitude fewer time points and far lighter compute while still yielding biologically interpretable subnetworks from a single seed gene.

The 32 × 4 expansion is chosen for all subsequent analyses because it strikes the right balance between biological depth and practical scope. Limiting each hop to 32 of the strongest elasticity neighbors keeps the growth of the graph contained, while four propagation steps are ample to capture multi-step regulatory cascades. The resulting sub-graph (4641 genes and 11676 edges) remains small enough for reliable enrichment statistics, manual curation, and clear visualization, yet large enough to preserve the key ferroptosis signals that LazyNet uncovers. This setting therefore provides a tractable, biologically informative backbone for downstream validation and

interpretation.

3.3 ARCHS4 Co-expression

Of the 11,676 edges in the 32×4 subgraph, 9,860 (84.4%) mapped to the ARCHS4 pan-human matrix and received a co-expression value (Pearson r) [35]. The r distribution is broadly centered near zero (mean ≈ 0.053; median ≈ 0.044) and closely matches degree-matched random pairs (Supplement S2A).

In an unsigned view, LazyNet edges are far more likely than random to show moderate co-expression: 26.6% of edges meet $|r| \geq 0.20$ versus ~5% of random pairs—an ≈5× enrichment—indicating that the model links genes that frequently co-vary across thousands of public RNA-seq datasets (Supplement S2B).

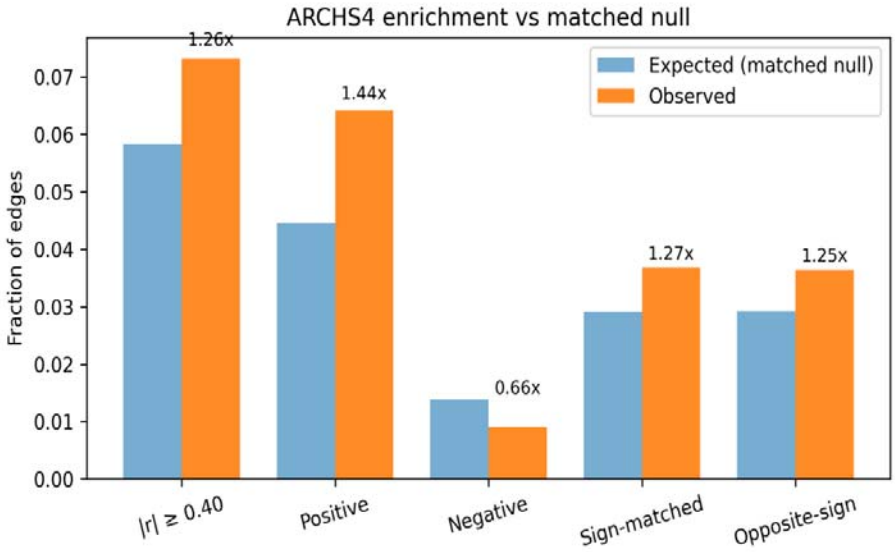

Figure 3. ARCHS4 co-expression enrichment for LazyNet edges. Observed (orange) versus expected (blue) fractions for edges from the 32×4 LazyNet subgraph annotated in ARCHS4 (n = 9,860). To avoid inflation from expression abundance and hubness, expected fractions come from an expression/detection- and degree-matched null (≈3.94M matched pairs). Bars show tail rates at $|r| \geq 0.40$; numbers above bars are fold enrichment (Observed/Expected). LazyNet edges are enriched in the overall $|r|$ tail (0.073 vs 0.058; 1.26×) and especially the positive tail (0.064 vs 0.045; 1.44×) with depletion in the negative tail (0.009 vs 0.014; 0.66×). "Sign-matched" multiplies ARCHS4 r by the LazyNet edge sign; "Opposite-sign" reverses it. As expected for an

undirected, context-mixed resource, sign-aware fractions are similar (1.27× vs 1.25×).

Sign-aware analysis refines this picture in Figure 3: the positive, sign-matched tail is enriched (~1.36 fold), the opposite-sign tail is depleted (~0.67 fold), and the overall $|r| \geq 0.20$ mass is near expectation (~1.02 fold) once signs are considered (Supplement S2C). This is consistent with the biology: LazyNet elasticities encode local, directional responses in a one-step ODE, whereas ARCHS4 aggregates steady-state, undirected correlations across heterogeneous tissues; inhibitory or context-specific links can therefore appear anti-correlated in the aggregate.
Taken together, the strong unsigned enrichment and the direction-consistent positive-tail enrichment indicate that a substantial subset of LazyNet edges is corroborated by large-scale expression evidence while respecting expected sign- and context-dependent effects.

3.4 STRING Database and Literature Validation

We evaluated the 32 × 4 LazyNet sub-graph against a filtered STRING network containing the 18152 genes detected in our single-cell atlas [36]. Of the 4630 genes retained by LazyNet, STRING supplies 411761 undirected links, giving an observed density of 0.0384. Size-matched random gene sets drawn 1 000 times from the same background average 0.0373 ± 0.0008, placing the sub-graph 1.5 s.d. above expectation (empirical $p = 0.072$) — evidence for modest but over-connectivity (Supplement S3A).

Edgewise, LazyNet 11 662 elasticity links intersect STRING in 523 cases (Jaccard $\approx 1 \times 10^{-4}$). A one-tailed hypergeometric test confirms this overlap is highly non-random ($p = 1.2 \times 10^{-5}$). Although elasticity magnitudes and STRING scores show no global correlation (Pearson $r = -0.071$, $p = 0.10$; Spearman $\rho = -0.036$), a top-quartile split in each metric separates three informative classes: 27 “both-high” edges, 104 STRING-high only (canonical interactions), and 104 LazyNet-high only—highlighting under-explored relationships for experimental follow-up. Representative literature support for ten both-high edges is provided in Table 3.

Table 3: STRING database literature comparison

Edge	Source / Reference	Brief description of reported linkage
AATF – PRPF3	Psychiatry Investigation, Suppl. 2 PDF (2022)	Coupling across RNA posttranscriptional modification, celldeath/survival pathways, and embryonic development

ABHD15 – TRPM6	Decreased GLUT2 and glucose uptake contribute to insulin-secretion defects in MODY3/HNF1A hiPSC-derived mutant β-cells	Shared involvement in insulin signalling / insulin resistance
ADH4 – CYP27C1	The effect of rs2910686 on ERAP2 expression in IBD and epithelial inflammatory response	Retinoid and isoprenoidbinding/processing connection
ANO2 – XKR9	TMEM16 and OSCA/TMEM63 proteins share a conserved potential to permeate ions and phospholipids (ANO - XKR link)	Functional link between ANO (TMEM16) and XKR families in ion- and phospholipid permeation
CNIH3 – NAPB	HAGRID MPM interaction page (pair 3804/342900)	Documented protein–protein interaction
DNAJA2 – SNRPD3	Report on DNAJ–HSP40–HSP70–SNRPD3 complex	Chaperone–spliceosome functional connection
FAM177A1 – MED29	Cleavage and Polyadenylation Specific Factor 1 Promotes Tumor Progression via Alternative Polyadenylation and Splicing in Hepatocellular Carcinoma	Linkage through mRNA 3--end processing & transcription regulation
FAM241B – WLS	PICKLE interaction record (ID 3601507, human)	Reported physical association
GADD45GIP1 – IER2	Transcription and chromatin organisation of a housekeeping gene cluster containing an integrated β-globin locus-control region	Co-regulation within housekeeping-gene chromatin domain
HLF – UQCRQ	Mitochondrial complex III regulates hypoxic activation of HIF	HLF–HIF axis functionally linked to complex III subunit UQCRQ in hypoxic signalling

Among the findings, the moderate enrichments are entirely acceptable—and even expected—because STRING confidence and LazyNet elasticity quantify different biological facets. STRING aggregates heterogeneous evidence (co-expression, physical binding, literature co-mentions), whereas elasticity captures the transcriptional impact inferred by our ODE model. Perfect concordance would therefore be unrealistic; instead, the modest density excess and significant, yet limited, edge-level overlap demonstrate that LazyNet recovers a meaningful core of known biology while offering a large complement of novel, model-specific interactions.

3.5 Proteomics corroborates transcript-derived network structure

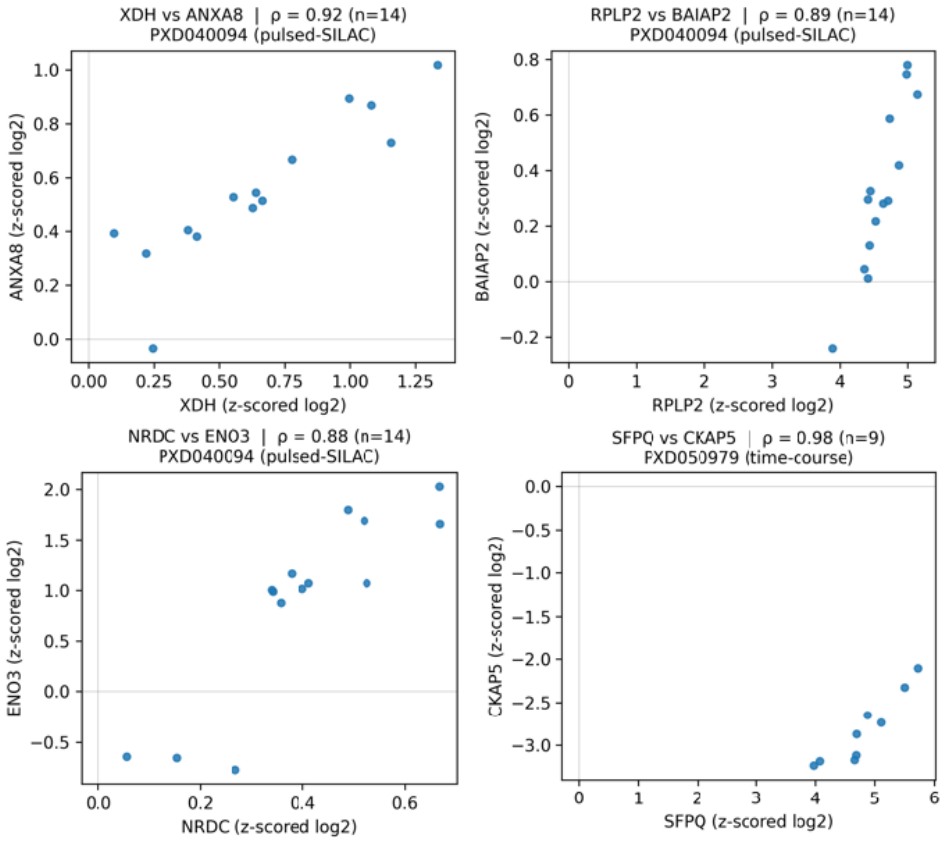

Figure 4. Proteomics support for LazyNet edges in GPX4 loss-of-function fibroblasts. Protein–protein abundance scatterplots for representative LazyNet edges across SILAC channels. Axes show z-scored $\log_2$ protein levels; each point is one channel/time point. Panels report Spearman ρ and sample count (n). Three pairs come from the pulsed-SILAC GPX4-KO dataset PXD040094—XDH–ANXA8 (ρ = 0.92, n = 14), RPLP2–BAIAP2 (ρ = 0.89, n = 14), NRDC–ENO3 (ρ = 0.88, n = 14)—and one from the tamoxifen-inducible GPX4-KO time course PXD050979—SFPQ–CKAP5 (ρ = 0.98, n = 9). Edges were selected for adequate coverage (≥ 8 quantified channels when available) and clear co-variation.

Cross-omics robustness was assessed by projecting RNA-derived edges onto two independent mouse fibroblast SILAC proteomes of GPX4 loss-of-function ferroptosis. The inducible GPX4-knockout time course (PRIDE PXD050979) yielded 1,026 quantifiable proteins after consolidation; 47 network edges had sufficient protein coverage (≥ 3 finite values per partner). Their absolute protein co-variation was higher than degree-matched random pairs (median |ρ| =

0.536 vs 0.400; Wilcoxon $P = 4.9 \times 10^{-3}$). The larger pulsed-SILAC study of constitutive GPX4 knockout (PRIDE PXD040094) provided 5,516 proteins across 14 channels; of 563 mapped edges, 507 met the coverage criterion and likewise exceeded the null (median $|\rho| = 0.367$ vs 0.336; Wilcoxon $P = 6.2 \times 10^{-4}$). Representative protein–protein scatters are shown in (Figure 4), including XDH–ANXA8 ($\rho = 0.92$), RPLP2–BAIAP2 ($\rho = 0.89$), NRDC–ENO3 ($\rho = 0.88$), and SFPQ–CKAP5 ($\rho = 0.98$), illustrating coherent proteomic co-variation across channels. Together with the enrichment tests, these results indicate that a substantial subset of LazyNet edges remains concordant at the protein layer, supporting the biological fidelity of the inferred network, despite differences in species and experimental design.

3.6 Biological Insights

The 32×4 subgraph seeded on the CRISPR A/I screen recapitulates the ferroptosis core: among the top 1,000 LazyNet edges (Supplement S1C), 15 of the 27 benchmark regulators reported by Tian et al. cluster near the network center, consistent with the canonical lipid-peroxidation axis [33].

Several connections align with established mechanisms, supporting the face validity of the wiring. A mitochondrial module links MFN2-mediated mitophagy, which constrains ROS-driven death, with SIRT3 signaling that modulates ferroptosis through GPX4 and AMPK–mTOR pathways [37–38]. Immune–redox crosstalk involving TLR4–STAT3 that amplifies mitochondrial ROS is also recovered [44]. In parallel, ADCY10-driven mitochondrial cAMP has been shown to protect neurons under oxidative stress, matching the directionality of edges observed here [45]. These concordances indicate that LazyNet recovers known biology across mitochondrial, immune, and metabolic axes.

Beyond literature-supported modules, the subgraph nominates cross-axis links that, to our knowledge, are not reported as direct connections and are presented as testable hypotheses. First, a lysosomal–autophagy route from PSAP to mTORC1 (MTOR/RPTOR/RHEB), potentially via ULK1-associated scaffolding (with PPP4R1 as a putative bridge), is consistent with PSAP-dependent lipofuscin accumulation and neuron-specific ferroptosis. Second, high-elasticity edges suggest immune inputs from TLR6 and CCR10 into autophagy/mTOR nodes (ATG13, MTOR). Third, a metabolic coupling is predicted in which TRPM6 (Mg^{2+} channel) and ADCY10 interface with the SIRT3–mTORC1 axis.

Taken together, the results point to an integrated immune–mitochondrial–lysosomal circuit in

which mitochondrial quality control (MFN2/SIRT3), innate immune signaling (TLRs), and metabolic/cAMP cues (ADCY10/TRPM6) converge on mTOR-linked autophagy to tune the ferroptotic threshold. Edge-level attributes are provided in Supplement S1C to facilitate replication and follow-up.

4. Discussion

This study set out to answer a practical question for small labs running CRISPR A/I screens: can we predict post-perturbation expression accurately, on ordinary hardware, and still learn mechanism? Across two datasets, LazyNet achieves strong predictive ranking (ROC-/PR-AUC) under tight compute, with competitive absolute error and clear generalization to a distinct T-cell slice. The neuronal dataset remains our focus; the T-cell results serve as a sanity check that the approach is not dataset-specific. Relative to transformer and state-space baselines, the gains concentrate in threshold-free ranking while ceding some ground on RMSE and Pearson Correlation in certain settings—an expected outcome when a single per-gene operating point is fixed. Taken together, the results indicate that an interpretable, CPU-tractable ODE update can be competitive with modern sequence models for single-snapshot perturbation prediction.

A central feature of LazyNet is that its one-step, log–exp residual update exposes elasticities—local, directed sensitivities of genes to one another around a baseline. Averaging Jacobians across replicas produces a consensus, directed interaction matrix that we expand into a compact subgraph for interpretation. This graph recovers a substantial fraction of known ferroptosis regulators and shows coherent support across independent resources: enrichment in large-scale co-expression (ARCHS4), non-random overlap with curated functional associations (STRING), and concordant protein-level co-variation in GPX4 loss-of-function proteomes. The agreement is moderate rather than exhaustive — as expected given that elasticity (local, directional quantitative response) and database edges (heterogeneous, often undirected evidence) measure different facets of biology—but it is sufficient to anchor new, testable hypotheses, including a lysosomal–mitochondrial–immune module that nominates specific links into mTOR-connected autophagy.

Equally important is practicality. LazyNet's capacity is concentrated into a small set of mechanistically meaningful weights, which helps it converge quickly on CPUs without external priors, while still producing parameters that carry direct interpretive value (rate-like elasticities). The ensemble Jacobian pipeline is deterministic and light enough to run end-to-end on a single node, and the entire workflow—preprocessing, splits, commands, and seeds—is packaged for exact replication. In contrast, graph-inference tools optimized for edge ranking from co-expression can generate high-accuracy local subgraphs but are not designed to provide a globally consistent, directed dynamic map aligned with potential causality [55]; for comparative purposes we therefore emphasize module-level enrichment against authoritative, curated

resources and orthogonal experiments rather than raw edge-overlap counts.

Several limitations follow from design choices. Working in log space requires positive inputs and introduces mild dependence on normalization and pseudocounts. A single explicit Euler step matches two-snapshot designs but leaves the absolute Δt unidentified and cannot represent delays or oscillations; elasticities are local to the chosen baseline and may vary across states. Our enrichment tests rely on incomplete references whose coverage and evidence channels differ from the signal inferred by LazyNet; consequently, we do not expect high global concordance, and we treat overlaps as supportive rather than definitive. Finally, within a fixed wall-time budget, excessively large models can undertrain and reduce accuracy despite higher nominal capacity.

Looking ahead, three directions appear most impactful. First, time-resolved perturbation series would allow multi-step fitting that estimates Δt, tests stability, and distinguishes fast from slow modes. Second, structured priors (e.g., motif-based regulator sets or pathway constraints) could be added in a transparent way — regularizing the elasticity matrix without sacrificing interpretability. Third, prospective validation—targeted CRISPR edits chosen from LazyNet's high-elasticity, database-unsupported edges—will be essential to calibrate precision and to refine the expansion rules that balance depth with tractability. Within these bounds, our results suggest that a one-step neural-ODE view can make routine Perturb-seq datasets actionable on commodity hardware, linking accurate prediction with mechanism-level hypotheses that are ready for bench testing.

5. Conclusions

Perturbation datasets remain costly to generate at scale and hard to harmonize across studies, so many groups benchmark on public resources while analyzing their own data with bespoke code. Our goal was a method that operates directly on local measurements, runs under modest compute, and yields parameters with clear mechanistic meaning. LazyNet meets this brief: a one-step neural-ODE that trains on CPUs, reports reproducible predictions, and exposes directed elasticities suitable for network interpretation.

Empirically, LazyNet attains competitive or superior predictive accuracy with substantially fewer parameters and lower wall-time than strong baselines, while producing elasticities that show coherent external concordance at the module level. A central advantage is the log–linear–exp map, which exactly represents multiplicative terms. Synergistic, multi-locus effects therefore appear as explicit components rather than opaque composites, enabling practical design and prioritization of combinatorial perturbations from sparsely sampled CRISPR A/I data.

Limits follow from design choices: two-snapshot learning cannot identify absolute Δt, delays, or oscillations; log-domain preprocessing (pseudocount, normalization) introduces mild dependencies; the update is Markovian without latent memory; and full transcriptome-scale expansion can be memory-intensive (partially mitigated with block-sparse top-k storage and

streaming). Nonetheless, within a fixed 1-hour budget, the model's capacity-to-convergence trade-off is favorable on CPUs, and its parameters remain directly interpretable.

Future work will extend validation to time-resolved perturbation series and additional cell contexts, incorporate lightweight structured priors (motifs, pathways) and uncertainty quantification, and perform prospective tests of synergy-driven predictions. Within these bounds, LazyNet offers a reproducible, resource-efficient route to synergy-aware mechanistic inference from perturbation data—bridging accurate prediction with directed, testable hypotheses.

Supplementary Materials:

Supplementary Table S1A — Ferroptosis seed selection. List of seed genes with selection criteria, source references/databases, and brief inclusion rationale (gene symbol, alias, source, evidence note).

Supplementary Table S1B — Subgraph metrics. Global metrics for each expanded subgraph (e.g., BFS 32 × 4): nodes, edges, density, average degree, clustering coefficient, diameter/90th-percentile path length, hub scores, and community counts.

Supplementary Table S1C — Selected subgraph details. Node- and edge-level attributes for the chosen subgraph: gene symbol, aliases, degree, community label; edge elasticity/sign, rank/percentile, and any external matches (ARCHS4 r, STRING score).

Supplementary Table S1D — Per-seed recall. Recall of benchmark ferroptosis regulators per seed (and aggregated), with optional precision@K and top-K per hop settings.

Supplementary Figure S2A — ARCHS4 comparison histogram. Histogram of Pearson's r for model-inferred edges vs. degree-/expression-matched null pairs; bin width and sample sizes indicated.

Supplementary Table S2B — ARCHS4 comparison statistics. Summary statistics (mean, median, IQR, upper-tail mass $\geq r_0$), test results (e.g., Mann–Whitney U / KS), effect sizes (e.g., Cliff's δ), and adjusted p-values.

Supplementary Table S2C — Pair enrichment. Enrichment of high co-expression pairs at thresholds (e.g., r≥0.1/0.2/0.3), reported as observed/expected, fold-enrichment, 95% CIs, and matched-null definition.

Supplementary Table S3A — STRING comparison overlap. Overlap between inferred edges and STRING v12 interactions across combined-score cutoffs, with counts, hypergeometric p-values,

FDR, and evidence-channel breakdown (experimental, database, co-expression, text mining).

Supplementary Table S3B — STRING literature support. Curated literature for overlapping edges (gene A–gene B, PMID(s), evidence type/brief note, STRING channel), suitable for spot-checking biological plausibility.

Supplement S4 — Mathematical Details of LazyNet. Formal derivation of LazyNet's one-step log–exp residual update; exact representation of multiplicative synergies.

Abbreviations

The following abbreviations are used in this manuscript:

A/I	CRISPR activation/interference
AIVC	Activation/Interference variable combinations (multi-perturbation synergy)
ARCHS4	Large-scale RNA-seq co-expression resource
BFS	Breadth-first search
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
GPX4	Glutathione peroxidase 4
KO	Knockout
LLE	Log–Linear–Exp map (LazyNet block)
MAE	Mean absolute error
ODE	Ordinary differential equation
r	Pearson correlation coefficient
ResNet	Residual network
RMSE	Root mean square error
scRNA-seq	Single-cell RNA sequencing
SILAC	Stable isotope labeling by amino acids in cell culture
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
HVG	Highly variable genes

References

1.Klein, A. M.; et al. Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. Cell 2015, 161, 1187–1201. https://doi.org/10.1016/j.cell.2015.04.044.

2.Macosko, E. Z.; et al. Highly Parallel Genome-Wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 2015, 161, 1202–1214. https://doi.org/10.1016/j.cell.2015.05.002.

3.Dixit, A.; et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. Cell 2016, 167, 1853 – 1866.e17. https://doi.org/10.1016/j.cell.2016.11.038.

4.Adamson, B.; et al. A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. Cell 2016, 167, 1867 – 1882.e21. https://doi.org/10.1016/j.cell.2016.11.048.

5.Bunne, C.; et al. How to Build the Virtual Cell with Artificial Intelligence: Priorities and Opportunities. Cell 2024, 187, 7045–7063. https://doi.org/10.1016/j.cell.2024.11.015.

6.Cui, H.; et al. scGPT: Toward Building a Foundation Model for Single-Cell Multi-Omics Using Generative AI. Nat. Methods 2024, 21, 1470–1480. https://doi.org/10.1038/s41592-024-02201-0.

7.Zeng, Y.; et al. CellFM: A Large-Scale Foundation Model Pre-Trained on Transcriptomics of 100 Million Human Cells. Nat. Commun. 2025, 16, 4679. https://doi.org/10.1038/s41467-025-59926-5.

8.Roohani, Y.; et al. Predicting Transcriptional Outcomes of Novel Multigene Perturbations with GEARS. Nat. Biotechnol. 2024, 42, 927–935. https://doi.org/10.1038/s41587-023-01905-6.

9.Murray, J. D. Mathematical Biology I: An Introduction, 3rd ed.; Springer: New York, 2002.

10.Klipp, E.; Liebermeister, W.; Wierling, C.; Kowald, A.; Schuster, S. Systems Biology: A Textbook, 2nd ed.; Wiley-Blackwell: Chichester, 2016.

11.Tyson, J. J.; Chen, K. C.; Novak, B. Network Dynamics and Cell Physiology. Nat. Rev. Mol. Cell Biol. 2001, 2, 908–916. https://doi.org/10.1038/35099075.

12.Heinrich, R.; Schuster, S. The Regulation of Cellular Systems; Chapman & Hall: New York, 1996.

13.Voit, E. O. Computational Analysis of Biochemical Systems; Cambridge University Press: Cambridge, 2000.

14.Chen, R. T. Q.; Rubanova, Y.; Bettencourt, J.; Duvenaud, D. Neural Ordinary Differential Equations. Adv. Neural Inf. Process. Syst. 2018, 31, 6571–6583.

15.Grathwohl, W.; Chen, R. T. Q.; Bettencourt, J.; Sutskever, I.; Duvenaud, D. FFJORD: Free-Form Continuous Dynamics for Scalable Reversible Generative Models. In Proceedings of the 36th International Conference on Machine Learning; PMLR: Long Beach, CA, 2019; pp 3642–3651.

16.Dupont, E.; Doucet, A.; Teh, Y. W. Augmented Neural Ordinary Differential Equations. Adv.

Neural Inf. Process. Syst. 2019, 32, 313–325.

17.Yi, Z. Integrating Equation Coding with Residual Networks for Efficient ODE Approximation in Biological Research. Math. Comput. Appl. 2025, 30, 47. https://doi.org/10.3390/mca30030047.

18.Kitano, H. Systems Biology: A Brief Overview. Science 2002, 295, 1662 – 1664. https://doi.org/10.1126/science.1069492.

19.Sachs, K.; Perez, O.; Pe'er, D.; Lauffenburger, D. A.; Nolan, G. P. Causal Protein-Signaling Networks Derived from Multiparameter Single-Cell Data. Science 2005, 308, 523 – 529. https://doi.org/10.1126/science.1105809.

20.Jaitin, D. A.; et al. Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. Cell 2016, 167, 1883–1896. https://doi.org/10.1016/j.cell.2016.11.039.

21.Nelson, M. R.; et al. The Support of Human Genetic Evidence for Approved Drug Indications. Nat. Genet. 2015, 47, 856–860. https://doi.org/10.1038/ng.3314.

22.Lee, J. S.; et al. Synthetic Lethality-Mediated Precision Oncology via the Tumor Transcriptome. Cell 2021, 184, 2487–2502. https://doi.org/10.1016/j.cell.2021.03.035.

23.Katti, A.; Diaz, B. J.; Caragine, C. M.; Sanjana, N. E.; Dow, L. E. CRISPR in Cancer Biology and Therapy. Nat. Rev. Cancer 2022, 22, 259–279. https://doi.org/10.1038/s41568-021-00451-y.

24.O'Neil, N. J.; Bailey, M. L.; Hieter, P. Synthetic Lethality and Cancer. Nat. Rev. Genet. 2017, 18, 613–623. https://doi.org/10.1038/nrg.2017.47.

25.Haley, B.; Roudnicky, F. Functional Genomics for Cancer Drug Target Discovery. Cancer Cell 2020, 38, 31–43. https://doi.org/10.1016/j.ccell.2020.06.013.

26.Norman, T. M.; et al. Exploring Genetic Interaction Manifolds Constructed from Rich Single-Cell Phenotypes. Science 2019, 365, 786–793. https://doi.org/10.1126/science.aax4438.

27.Low, L. A.; Mummery, C.; Berridge, B. R.; Austin, C. P.; Tagle, D. A. Organs-on-Chips: Into the Next Decade. Nat. Rev. Drug Discov. 2021, 20, 345 – 361. https://doi.org/10.1038/s41573-020-0079-3.

28.Wang, H.; Yang, Y.; Liu, J.; Qian, L. Direct Cell Reprogramming: Approaches, Mechanisms and Progress. Nat. Rev. Mol. Cell Biol. 2021, 22, 410 – 424. https://doi.org/10.1038/s41580-021-00336-x.

29.Maude, S. L.; et al. Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia. N. Engl. J. Med. 2018, 378, 439–448. https://doi.org/10.1056/NEJMoa1709866.

30.Gillmore, J. D.; et al. CRISPR–Cas9 In Vivo Gene Editing for Transthyretin Amyloidosis. N. Engl. J. Med. 2021, 385, 493–502. https://doi.org/10.1056/NEJMoa2107454.

31.Lim, W. A. The Emerging Era of Cell Engineering: Harnessing the Modularity of Cells to Program Complex Biological Function. Science 2022, 378, 848 – 852. https://doi.org/10.1126/science.add4585.

32.Przybyla, L.; Gilbert, L. A. A New Era in Functional Genomics Screens. Nat. Rev. Genet. 2022, 23, 89–103. https://doi.org/10.1038/s41576-021-00439-1.

33.Tian, R.; et al. Genome-Wide CRISPRi/a Screens in Human Neurons Link Lysosomal Failure to Ferroptosis. Nat. Neurosci. 2021, 24, 1020–1034. https://doi.org/10.1038/s41593-021-00862-0.

34.Dyer, S. C.; et al. Ensembl 2025. Nucleic Acids Res. 2025, 53, D948 – D957. https://doi.org/10.1093/nar/gkae1071.

35.Lachmann, A.; et al. Massive Mining of Publicly Available RNA-Seq Data from Human and Mouse. Nat. Commun. 2018, 9, 1366. https://doi.org/10.1038/s41467-018-03751-6.

36.Szklarczyk, D.; et al. The STRING Database in 2023: Protein-Protein Association Networks and Functional Enrichment Analyses for Any Sequenced Genome of Interest. Nucleic Acids Res. 2023, 51, D638–D646. https://doi.org/10.1093/nar/gkac1000.

37.Ma, L.; et al. MANF Overexpression Ameliorates Oxidative Stress-Induced Apoptosis via MFN2-Dependent Mitophagy. Sci. Rep. 2024, 14, 11234. https://doi.org/10.1038/s41598-024-84167-9.

38.Han, D.; et al. SIRT3 Deficiency Is Resistant to Autophagy-Dependent Ferroptosis by Inhibiting the AMPK/mTOR Pathway and Promoting GPX4. J. Cell Physiol. 2020, 235, 8839 – 8851. https://doi.org/10.1002/jcp.29571.

39.Huynh-Thu, V. A.; et al. Inferring Regulatory Networks from Expression Data Using Tree-Based Methods. BMC Bioinformatics 2010, 11, 359. https://doi.org/10.1186/1471-2105-11-359.

40.Marbach, D.; et al. Wisdom of Crowds for Robust Gene-Network Inference. Nat. Methods 2012, 9, 796–804. https://doi.org/10.1038/nmeth.2016.

41.Aibar, S.; et al. SCENIC: Single-Cell Regulatory Network Inference and Clustering. Nat. Methods 2017, 14, 1083–1086. https://doi.org/10.1038/nmeth.4463.

42.Van den Bosch, B.; et al. SCENIC+ Enables Multiomic Regulatory Network Inference at Single-Cell Resolution. Nat. Biotechnol. 2023, 41, 127 – 134. https://doi.org/10.1038/s41587-022-01329-5.

43.Gao, Z.; et al. DeepFGRN: Deep Learning Framework for Gene Regulatory Network Inference from Time-Series scRNA-Seq Data. Brief. Bioinform. 2024, 25, bbae045.

https://doi.org/10.1093/bib/bbae045.

44.Zhang, J.; et al. Macrophage AHR–TLR4 Crosstalk Drives p-STAT3-Mediated Mitochondrial Oxidative Stress. Sci. Total Environ. 2024, 923, 171377. https://doi.org/10.1016/j.scitotenv.2024.171377.

45.Bastola, T.; et al. Activating Soluble Adenylyl Cyclase Protects Mitochondria and Rescues Retinal Ganglion Cells from Oxidative Stress. Antioxidants 2024, 13, 743. https://doi.org/10.3390/antiox13040743.

46.Nesterenko, A. M.; Korzhenevskii, D. A.; Tereshchuk, V. M.; Kudryashova, O. M.; Belousov, V. V.; Shokhina, A. G. Dataset on the Proteomic Response during Ferroptosis Induction via Tamoxifen-Induced GPX4 Knockout in Mouse Embryonic Fibroblasts. Data Brief 2023, 48, 109170. https://doi.org/10.1016/j.dib.2023.109170.

47.Wagle, P.; von Karstedt, S. Secretome Analysis of Ferroptotic Pfa1 Cells. PRIDE Archive, Dataset PXD050979 2023. https://doi.org/10.6019/PXD050979.

48.Tyanova, S.; Temu, T.; Sinitcyn, P.; et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods 2016, 13, 731 – 740. https://doi.org/10.1038/nmeth.3901.

49.Ahlmann-Eltze, C.; Huber, W.; Anders, S. Deep-learning-based gene perturbation effect prediction does not yet outperform simple linear baselines. Nat. Methods 2025, 22, 1657–1661. https://doi.org/10.1038/s41592-025-02772-6.

50.Hossain, I.; Fanfani, V.; Quackenbush, J.; Burkholz, R. Biologically informed NeuralODEs for genome-wide regulatory dynamics. Genome Biol. 2024, 25, 127. https://doi.org/10.1186/s13059-024-03264-0.

51.Lin, Z.; Chang, S.; Zweig, A.; Azizi, E.; Knowles, D. Interpretable Neural ODEs for Gene Regulatory Network Discovery under Perturbations. arXiv 2025, arXiv:2501.02409. https://doi.org/10.48550/arXiv.2501.02409.

52.Chai, T., Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? —Arguments against avoiding RMSE in the literature. Atmospheric Measurement Techniques, 7(7), 1525–1534. https://doi.org/10.5194/amt-7-1525-2014

53.Willmott, C. J., Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research, 30, 79–82. https://doi.org/10.3354/cr030079

54.Schmidt R, Steinhart Z, Layeghi M, Freimer JW, Bueno R, Nguyen VQ, Blaeschke F, Ye CJ, Marson A. CRISPR activation and interference screens decode stimulation responses in primary human T cells. Science. 2022 Feb 4;375(6580):eabj4008. doi: 10.1126/science.abj4008.

55.Chris J. Oates, Frank Dondelinger, Nora Bayani, James Korkola, Joe W. Gray, Sach Mukherjee. Causal network inference using biochemical kinetics. Bioinformatics. 2014 Sept; 30 (17), i468–i474. https://doi.org/10.1093/bioinformatics/btu452.