Journal of Zhejiang University-SCIENCE B 2026 27(10):1129-1139
www.jzus.zju.edu.cn; www.springer.com/journal/11585
E-mail: jzus_b@zju.edu.cn

Research Article
https://doi.org/10.1631/jzus.B2500424

Check for updates

Integrative graph deep learning with directional decoding for robust gene regulatory network inference from single-cell transcriptomics

Binhua TANG[1,2]✉, Yingying FENG[1], Xinyu GAO[1], Mengyao MAO[1], Yujia ZHANG[1]

[1]*Key Laboratory of Maritime Intelligent Cyberspace Technology (Ministry of Education of China), Hohai University, Nanjing 213200, China*
[2]*Shanghai Key Laboratory of Data Science, Fudan University, Shanghai 200438, China*

Abstract: While gene regulatory networks (GRNs) are fundamental to understanding complex cellular mechanisms, the accurate inference from single-cell RNA-sequencing (scRNA-seq) data is hindered by the inherent noise, transcriptional dropout events, and dynamic nature of gene regulation. Current computational approaches often overlook long-range regulatory interactions, lack explicit modeling of regulatory directionality, and suffer from limited generalizability. To address these limitations, we introduce scGRAIL, a supervised deep learning framework that combines inductive local-aware graph aggregation (LGA) with an asymmetric relation decoder to enable robust GRN inference. Within this framework, LGA performs multi-hop neighborhood aggregation on a transcription factor (TF)–target graph, jointly leveraging topological structure and gene expression to capture long-range regulatory dependencies. A node reconstruction autoencoder (NAE) imposes gene expression reconstruction constraints on the latent embeddings, thereby enhancing robustness to noise and data heterogeneity and improving generalization. An asymmetric difference mechanism is introduced during decoding for distinct modeling of TFs and target genes; by computing an ordered difference vector between their embeddings, the model explicitly captures the directionality and target specificity of TF–target regulatory interactions, facilitating more accurate identification of potential causal regulatory relationships. Comprehensive evaluations demonstrate the model's superiority: it achieves optimal performance on 86.4% of benchmark datasets (38/44), surpassing state-of-the-art methods. Notably, scGRAIL achieves a 4.25% increase in average area under the receiver operating characteristic curve (AUROC) across four real-world networks compared to the graph contrastive link prediction (GCLink) model, yielding a 17.64% AUROC improvement for the TF500+ Non-Specific network in the mouse hematopoietic stem cells with lymphoid-lineage (mHSC-L) dataset, demonstrating its resilience to noise and scalability for large-scale GRNs. These results establish scGRAIL as a transformative tool for uncovering gene regulatory logic in single-cell transcriptomics. Overall, scGRAIL enhances structural modeling and directional discrimination in single-cell GRN inference by using multi-hop graph aggregation to capture long-range regulatory dependencies, expression reconstruction constraints to improve cross-dataset robustness, and an asymmetric decoding mechanism to explicitly model the regulatory directionality between TFs and target genes.

Key words: Single-cell RNA sequencing (scRNA-seq); Gene regulatory network (GRN) inference; Local-aware graph aggregation (LGA); Node reconstruction autoencoder; Asymmetric relation decoder

1 Introduction

Gene regulatory networks (GRNs), comprising regulatory relationships among transcription factors (TFs) and target genes, serve as critical tools for elucidating biological processes and identifying molecular regulators and biomarkers in complex diseases. Advances in single-cell RNA-sequencing (scRNA-seq) technology have enabled the acquisition of high-dimensional omics data at single-cell resolution, revealing intercellular heterogeneity in gene expression and providing pivotal insights for reconstructing GRNs (Cuevas-Diaz Duran et al., 2024). However, the inherent high sparsity and noise in scRNA-seq datasets, primarily due to widespread "dropout" events, pose significant challenges for accurate GRN inference.

Several recently developed methods address these challenges, generally falling into two categories: unsupervised and supervised approaches. Unsupervised methods infer regulatory relationships between genes from expression data without relying on prior knowledge

✉ Binhua TANG, bh.tang@hhu.edu.cn
Binhua TANG, https://orcid.org/0000-0003-4744-5835

Received July 20, 2025; Revision accepted Jan. 4, 2026;
Crosschecked Aug. 13, 2026; Published online Sept. 22, 2026

© Zhejiang University Press 2026

of regulatory interactions. For instance, GEne Network Inference with Ensemble of trees (GENIE3), Gene Regulatory Network Boosting 2 (GRNBoost2), and SINgle CEll Regularized Inference using TImestamped Expression profileS (SINCERITIES) formulate GRN inference as a regression problem, employing statistical or machine learning techniques to infer GRNs (Huynh-Thu et al., 2010; Papili Gao et al., 2018; Moerman et al., 2019). SCODE infers regulatory matrices by fitting linear differential equations to expression dynamics along pseudotime trajectories (Matsumoto et al., 2017). BiXGBoost uses gradient-boosted decision trees to separately infer regulatory and regulated relationships and integrates bidirectional information to generate consistent gene regulatory scores (Zheng et al., 2019), and DeepSEM employs β-variational autoencoders for GRN inference (Shu et al., 2021). Given that most of these methods treat GRN inference as a regression or classification problem, they require substantial computational time, which limits their applicability to large-scale datasets. Furthermore, many unsupervised methods rely solely on scRNA-seq data, thereby limiting the capture of complex biological contexts and consequently affecting the accuracy of inferred results.

Conversely, supervised methods effectively address the limitations of unsupervised approaches by incorporating prior knowledge about gene interactions, organisms, or tissues. These methods not only improve inference accuracy but also adapt better to diverse biological scenarios. For instance, GENELink employs graph attention network (GAT) for GRN inference (Chen and Liu, 2022); GNNLink uses a graph convolutional network-based interaction graph encoder to refine gene features and capture interdependencies among nodes for GRN prediction (Mao et al., 2023); Deep learning model for Regulatory networks Inference among Genes (DeepRIG) adopts graph autoencoders to extract gene–gene interactions from scRNA-seq data (Wang et al., 2023); IGEGRNS leverages GraphSAGE for gene embedding and applies Top-k pooling to select key nodes for regulatory relationship prediction (Gan et al., 2024); the Graph Contrastive Link prediction (GCLink) model combines graph contrastive learning with GAT to capture global information and propagate it for GRN inference (Yu et al., 2025); scMGATGRN models node embeddings using view-level attention and reconstructs GRNs via dot-product scoring (Yuan et al., 2024). However, most of these methods focus solely on local relationships between individual TFs and target genes, treating each regulatory pair as an independent sample while ignoring contextual information from other regulatory interactions. Additionally, while they primarily capture correlations in gene expression, they fail to model regulatory directionality and causal mechanisms, resulting in predictions with limited biological interpretability.

To address these challenges, we propose scGRAIL, a structure-aware, directional deep-learning method that comprehensively models gene regulatory relationships. Based on the local-aware graph aggregation module, scGRAIL learns and integrates gene expression with topological structure embeddings by inductively aggregating local neighborhood information. The model performs multi-hop neighborhood aggregation on the TF–target graph, and by stacking multiple layers of local-aware graph aggregation (LGA), it expands the receptive field such that gene embeddings encode not only direct TF–target relationships but also long-range regulatory dependencies and higher-order structural information, including shared target genes and co-regulatory modules. A node reconstruction mechanism is integrated to preserve original expression features, thereby improving the model's generalizability and consistency, increasing its resilience to noise and data heterogeneity, and enhancing cross-dataset generalization capabilities via expression reconstruction constraints.

Furthermore, to account for the inherent directionality of regulatory relationships in GRNs, scGRAIL employs an asymmetric differential mechanism during the decoding phase, explicitly modeling regulatory directionality and target specificity to improve discriminative performance for regulatory edges. Its decoder employs distinct mappings for TFs and target genes and performs direction-sensitive scoring based on an ordered difference vector, enabling the interpretation of each predicted edge as the directed, relative regulatory effect of a specific TF on its target gene. The experimental results confirm scGRAIL's high accuracy in identifying TF–target regulatory relationships and inferring GRNs.

2 Materials and methods

2.1 Architecture of the proposed scGRAIL

Here, we present scGRAIL, a supervised deep learning framework that combines inductive node

embedding learning with an asymmetric relation decoder for link prediction (Hamilton et al., 2017). As shown in Fig. 1, scGRAIL comprises three main modules: data acquisition and preprocessing, gene node embedding learning, and asymmetric relation decoding. Then, the gene node embedding learning module aims to derive low-dimensional gene embeddings enriched with biological information. This module first employs LGA to iteratively aggregate information from gene nodes and their neighborhood nodes, using gene expression data and prior knowledge to generate low-dimensional node embeddings with enhanced semantic representation. Subsequently, a node reconstruction autoencoder performs effective feature extraction and representation learning, producing reconstructed data that closely resemble the original input.

The asymmetric relation decoder module calculates the embedding differences between TFs and their potential target genes. This enables the model to focus on directional regulatory relationships (how TFs regulate target genes) rather than merely describing associations, thereby learning biologically interpretable regulatory features. Finally, the differential feature vectors are processed by a multilayer perceptron (MLP) and scored using a sigmoid function.

2.2 Inference of gene node embeddings

2.2.1 Problem formulation

Let $\mathcal{G}=(\mathcal{V},\mathcal{E})$ represent a GRN, where $\mathcal{V}=\{v_1,\cdots,v_N\}$ denotes the node set of size $N=|\mathcal{V}|$ and $\mathcal{E}=\{e_1,\cdots,e_M\}$ represents the edge set of size M. The

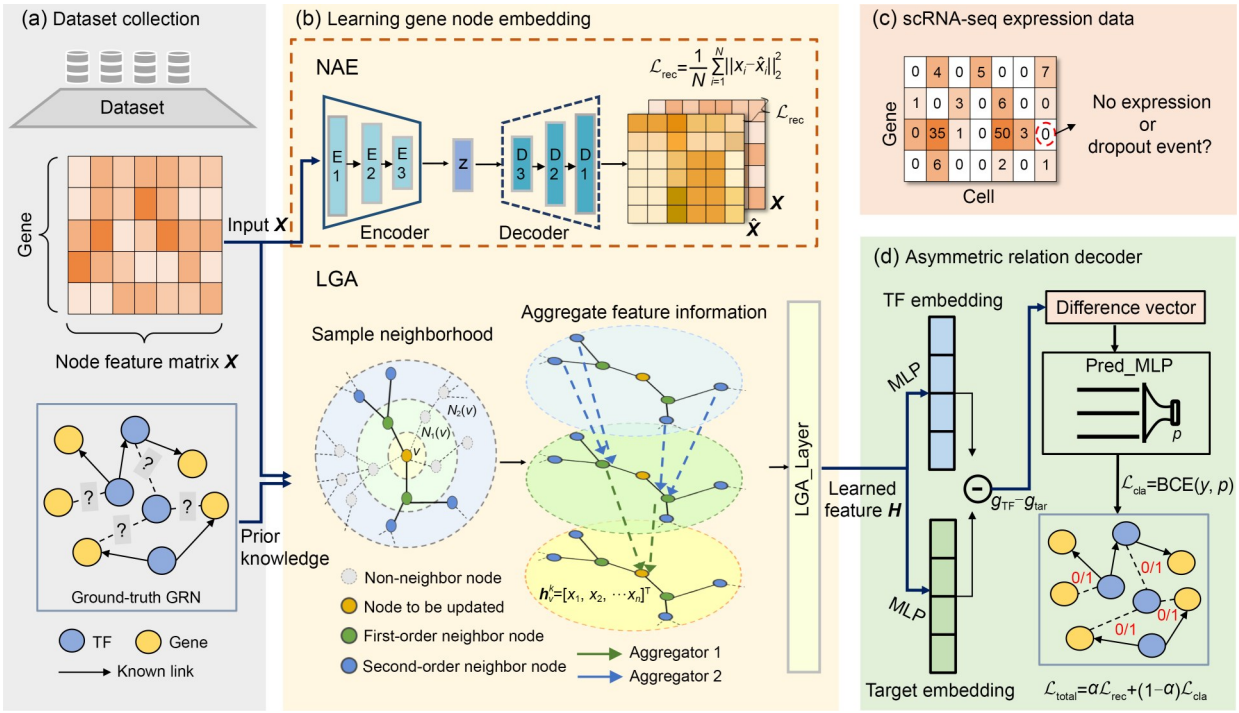

Fig. 1 scGRAIL framework. (a) Data acquisition and preprocessing module: scRNA-seq data are denoted as a node feature matrix X, and TF–target prior knowledge is encoded as directed regulatory edges and fed into the graph learning module. (b) Node embedding learning module, comprising the NAE and LGA: the autoencoder compresses X into a low-dimensional representation and preserves information, while LGA aggregates multi-hop neighborhood features on the TF–target graph to produce structure-aware embeddings. z represents latent layer; h_v^k corresponds to the updated feature representation of node v at the kth layer; $\mathcal{L}_{rec}$ denotes reconstruction loss between input matrix X and reconstructed matrix $\hat{X}$. (c) "Zero-expression" in the data: either from genuine absence of expression or from technical noise. (d) Asymmetric relation decoding module: embeddings of TFs and targets are obtained separately; an ordered difference vector is constructed and then passed through a prediction network (Pred_MLP) to output the edge probability. g_{TF} and g_{tar} represent embedding vectors of transcription factor (TF) and target gene, respectively; $\mathcal{L}_{cal}$ is binary cross-entropy (BCE) classification loss for regulatory-edge prediction, y is ground-truth label, and p is predicted probability; $\mathcal{L}_{total}$ is total loss of the model; α is the weight coefficient balancing two loss terms. scRNA-seq: single-cell RNA sequencing; LGA: local-aware graph aggregation; NAE: node reconstruction autoencoder; Pred_MLP: predictive multilayer perceptron.

corresponding adjacency matrix $\boldsymbol{A}\in\{0,1\}^{N\times N}$ satisfies A_{ij}=1 if $(v_i, v_j)\in\mathcal{E}$; otherwise, A_{ij}=0. Gene expression is denoted as $\boldsymbol{X}=[x_1,\cdots,x_N]\in\mathbb{R}^{N\times F}$, where N is the number of genes (equal to the number of nodes in $\mathcal{G}$ and F is the number of cells. TFs and target genes are treated as nodes in $\mathcal{G}$, while edges in $\mathcal{E}$ represent experimentally verified regulatory interactions between TFs and target genes. We formulate GRN inference as a link prediction problem, predicting potential edges between unconnected nodes by leveraging known node features and network topology (Liu et al., 2024).

2.2.2 Architecture of LGA

This study developed an LGA module to exploit higher-order neighborhood attributes in graph data and produce semantically rich node embeddings. As shown in Fig. 1b, LGA leverages inductive graph learning to establish localized receptive fields for nodes using a multi-level adjacency feature propagation method. By integrating topological relationships and expression features from contextual neighborhoods, it generates structurally aware embeddings, thereby establishing a more discriminative semantic foundation for subsequent regulatory relationship modeling.

For any node $v\in\{1,2,\cdots,N\}$ in the graph, let its neighbor set be $\mathcal{N}(v)$. Assuming that LGA stacks k layers with the initial feature representation of node v, denoted as $\boldsymbol{h}_v^0\in\mathbb{R}^F$, its aggregated neighborhood representation at the kth layer is defined as:

$$\boldsymbol{h}_{\mathcal{N}(v)}^k=\text{Aggregate}^k\left(\left\{\boldsymbol{h}_u^{k-1},u\in\mathcal{N}(v)\right\}\right),\tag{1}$$

where the feature aggregation function at the kth layer, denoted as Aggregate$^k(\cdot)$, employs a mean pooling operation defined as:

$$\text{Aggregate}^k=\frac{1}{|\mathcal{N}(v)|}\sum_{u\in\mathcal{N}(v)}\boldsymbol{h}_u^{k-1}.\tag{2}$$

The aggregated neighborhood representation is concatenated with the target node's features and transformed through a learnable linear layer with nonlinear activation to update the node representation:

$$\boldsymbol{h}_v^k=\sigma\left(\boldsymbol{W}^k\cdot\left(\text{Concat}\left(\boldsymbol{h}_v^{k-1},\boldsymbol{h}_{\mathcal{N}(v)}^k\right)+\boldsymbol{b}^k\right)\right),\tag{3}$$

where Concat$(\cdot)$ denotes vector concatenation; $\boldsymbol{W}^k$ and $\boldsymbol{b}^k$ represent the weight matrix and bias term at the kth layer, respectively; $\boldsymbol{h}_v^k$ and $\boldsymbol{h}_v^{k-1}$ correspond to the updated feature representation of node v at the kth and $(k-1)$th layers, respectively; $\sigma(\cdot)$ represents the nonlinear activation function.

2.2.3 Node reconstruction autoencoder module

We introduce a node reconstruction autoencoder (NAE), which learns low-dimensional representations that preserve key semantic and structural information by minimizing the node feature reconstruction error through an encoder–decoder framework. The NAE models node features using an encoder–decoder mechanism that recovers the original feature distribution, thereby revealing deep relational patterns and structural dependencies among nodes. By minimizing reconstruction error, the model learns latent representations that reflect nodes' semantic roles and distribution patterns, facilitating the capture of critical feature information within graph structures. The autoencoder maps the original node data $\boldsymbol{X}=[x_1,\cdots,x_N]\in\mathbb{R}^{N\times F}$ to a low-dimensional latent space through an encoder, yielding a latent representation $\boldsymbol{z}\in\mathbb{R}^{N\times Q}$ that preserves the key features and intrinsic structural information of the nodes. Assuming the encoder comprises L layers, the representation learned at the lth layer is given by:

$$\text{Enc}_l=\sigma\left(\boldsymbol{W}_l\text{Enc}_{l-1}+\boldsymbol{b}_l\right),\tag{4}$$

where $\text{Enc}_0=\boldsymbol{X}$, latent representation $\boldsymbol{z}=\text{Enc}_L$; $\boldsymbol{W}_l$ and $\boldsymbol{b}_l$ are the weight matrix and bias term of the lth layer, respectively.

The decoder reconstructs the original input from the low-dimensional encoded representation $\boldsymbol{z}$. For a decoder with T layers, the decoded representation at the tth layer is expressed as:

$$\text{Dec}_t=\sigma\left(\boldsymbol{W}_t\text{Dec}_{t-1}+\boldsymbol{b}_t\right),\tag{5}$$

where $\text{Dec}_0=\boldsymbol{z}$, the reconstructed node feature is $\hat{\boldsymbol{x}}=\text{Dec}_T$; $\boldsymbol{W}_t$ and $\boldsymbol{b}_t$ represent the weight matrix and bias term of the tth decoder layer, respectively.

2.3 Asymmetric relation decoding mechanism

After obtaining node embeddings, we employ an asymmetric relation decoder to perform direction-sensitive differential modeling of TF–target pairs and use a two-layer MLP to output regulatory strength scores for directed regulatory relationship prediction.

A joint loss function combining node reconstruction loss and binary cross-entropy loss was implemented to optimize the representational learning and predictive performance of our proposed model. We selected the area under the receiver operating characteristic curve (AUROC) and area under the precision–recall curve (AUPRC) as evaluation metrics to evaluate the performance of various GRN inference methods. The higher the AUROC and AUPRC scores, the better the performance of the corresponding method and the more accurate and reliable the GRN inference (Section S1).

3 Results

3.1 Experimental settings

To ensure fair and reproducible comparisons, all datasets and methods were trained under consistent experimental settings. All models were trained based on the predefined training/validation/test splits from the original benchmark implementation and evaluated using AUROC and AUPRC metrics.

3.2 Performance comparison with other methods on benchmark dataset

To evaluate the proposed model's effectiveness, we conducted comprehensive comparisons with 12 existing approaches using the benchmark networks and datasets. These methods included GRNBoost2 (Moerman et al., 2019), GNNLink (Mao et al., 2023), SCODE (Matsumoto et al., 2017), DeepSEM (Shu et al., 2021), GENELink (Chen and Liu, 2022), DeepRIG (Wang et al., 2023), IGEGRNS (Gan et al., 2024), GCLink (Yu et al., 2025), scMGATGRN (Yuan et al., 2024), STGRNS (Xu et al., 2023), DGRNS (Zhao et al., 2022), and GNE (KC et al., 2019). To ensure the accuracy and fairness of the experiments, all methods were implemented using their default parameter settings. We incorporated two direction-aware baseline models, GAEDGRN (Wei et al., 2025) and convolutional neural network for co-expression (CNNC) (Yuan and Bar-Joseph, 2019). The experimental results are shown in Fig. 2. scGRAIL achieved optimal prediction performance on 86.4% (38/44) of all evaluated datasets. Compared with GCLink, the second-best method, which performed optimally on 31.8% (14/44) of the datasets, scGRAIL demonstrated a maximum improvement of 4.25% in average AUROC across four types of gold-standard networks. The most significant AUROC enhancement (17.64%) was observed for the TF500+ gene scale in the Non-Specific class gold-standard network (mHSC-L dataset).

For Specific-type gold-standard networks, scGRAIL achieved the highest average AUROC (87.78% and 88.94%) and AUPRC (74.03% and 74.91%) values across both TF500+ and TF1000+ gene scales. Compared with the second-best method, scMGATGRN, scGRAIL improved AUROC by 0.95% and 1.29% and AUPRC by 2.55% and 2.30%, respectively. Furthermore, our method exhibited more stable performance across both AUROC and AUPRC, consistently outperforming unsupervised approaches. These results demonstrate that supervised methods possess distinct performance advantages for GRN inference tasks.

3.3 Ablation analysis on the proposed model structure

We performed an ablation study by systematically removing or replacing key components (NAE, LGA, and the asymmetric relation decoder) to verify their individual contributions to scGRAIL's GRN inference performance. In Fig. 3, scGRAIL achieved optimal predictive performance in AUROC across all seven datasets under the TF500+ gene scale for Specific-type gold-standard networks, with a maximum improvement of 5.2% (Section S2 and Table S1). Specifically, when employing either LGA or NAE alone for node feature learning, the models demonstrated reduced ability to capture consensus-enriched node embeddings, resulting in average AUROC decreases of 1.92% and 2.05%, respectively. Removing the asymmetric relation decoder (replaced by an MLP that symmetrically processes TF–target relationships) resulted in a 1.15% average AUROC reduction relative to scGRAIL, as this modification fails to account for the biologically inherent unidirectional nature of TF–target regulatory interactions. The substitution of NAE with either adversarial autoencoder (AAE) or variational autoencoder (VAE) led to inferior performance due to their dependence on learned data distributions, a limitation exacerbated by our constrained training set size, which compromised sample reconstruction quality.

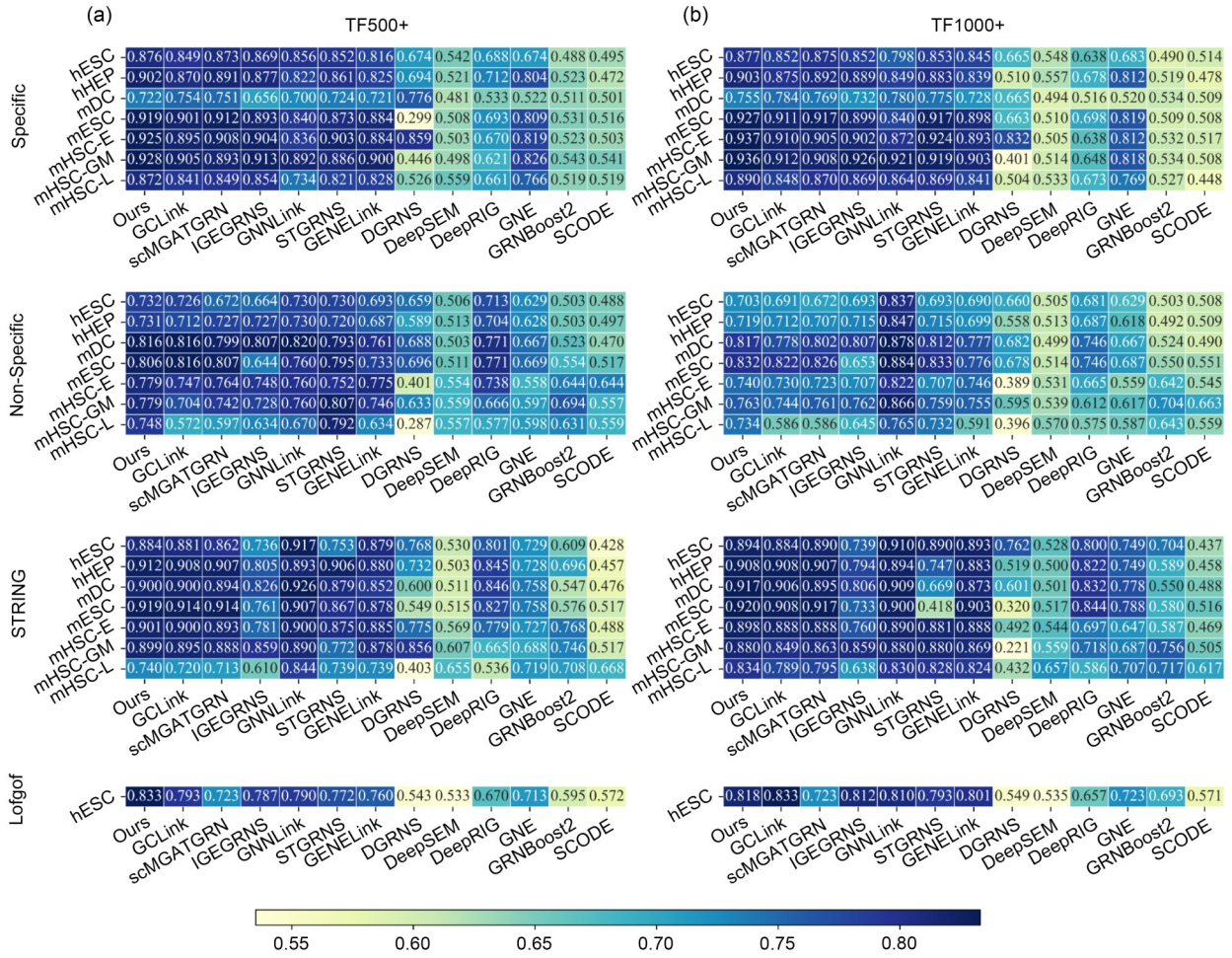

Fig. 2 Performance comparison of AUROC scores between the proposed method and 12 different methods on four types of gold-standard networks. (a) A comparative analysis of the AUROC performance of multiple computational approaches applied to gene-scale TF500+ datasets. The horizontal axis denotes the evaluated computational approaches; the vertical axis specifies distinct benchmark network datasets. Each grid cell contains the numerical AUROC value achieved by a specific method on a particular network dataset, with higher values indicating superior predictive performance. (b) The case of the gene-scale TF1000+ datasets. AUROC: area under the receiver operating characteristic curve; hESC: human embryonic stem cell; hHEP: human hepatocyte; mDC: mouse dendritic cell; mESC: mouse embryonic stem cell; mHSC-E: mouse hematopoietic stem cell erythroid; mHSC-GM: granulocyte-macrophage lineage mouse hematopoietic stem cell; mHSC-L: lymphoid lineage mouse hematopoietic stem cell.

3.4 Hyperparameter analysis of the proposed scGRAIL

We further conducted a systematic hyperparameter analysis across multiple datasets and two gold-standard networks—Non-Specific and search tool for the retrieval of interacting genes/proteins (STRING)—examining the effects of key settings such as the hidden dimensions of the NAE and LGA modules and the batch size on scGRAIL's performance. The performance of scGRAIL is influenced by multiple hyperparameters, including the output dimension of the learned node embeddings, the learning rate, the hidden layer dimensions (hidden_dim), and the batch size. To thoroughly evaluate the impact of key parameters on model performance, this section analyzes the effects of hidden layer dimensions in both NAE and LGA architectures, as well as batch size configuration, on GRN inference (Table S2).

3.5 Analysis of computational complexity

To further validate scGRAIL's computational efficiency, we analyzed the runtime of seven GRN inference methods across four benchmark networks (Specific, Non-Specific, STRING, and Lofgof) using

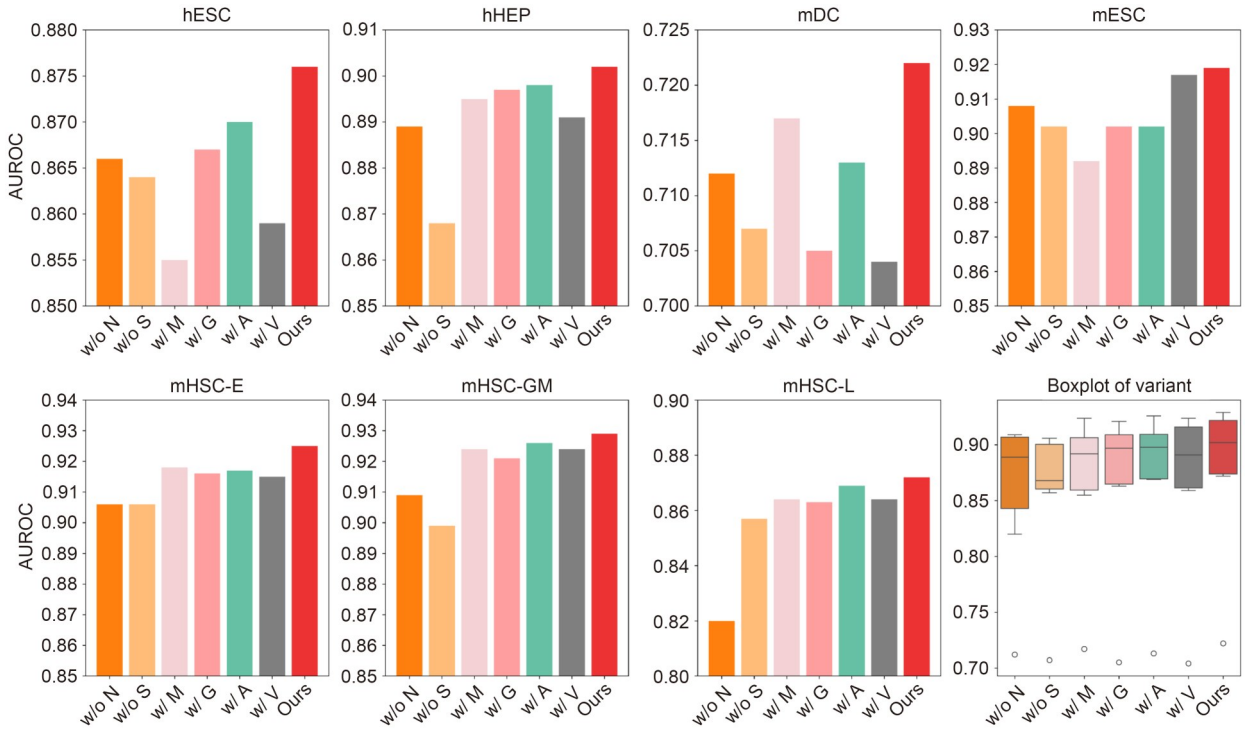

Fig. 3 Performance comparison of AUROC (range 0–1) metrics between scGRAIL and its six variants on Specific-type gold-standard network datasets with TF500+ gene scale. w/o N and w/o S are variants without dataset N and dataset S, respectively; w/ M, w/ G, w/ A, and w/ V denote variants trained using only the corresponding single dataset; Ours represents the full model trained on all datasets. hESC: human embryonic stem cell; hHEP: human hepatocyte; mDC: mouse dendritic cell; mESC: mouse embryonic stem cell; mHSC-E: mouse hematopoietic stem cell erythroid; mHSC-GM: granulocyte-macrophage lineage mouse hematopoietic stem cell; mHSC-L: lymphoid lineage mouse hematopoietic stem cell; AUROC: area under the receiver operating characteristic curve.

scRNA-seq datasets. As summarized in Table S3, scGRAIL achieved the shortest average runtime across all networks: 28.67 s (Specific), 3.88 s (Non-Specific), 5.84 s (STRING), and 7.79 s (Lofgof). The second-best method, GCLink, required substantially longer runtime (140.28, 19.61, 24.52, and 29.50 s, respectively). These results demonstrate that scGRAIL not only delivers superior prediction performance but also exhibits significantly higher computational efficiency, making it particularly suitable for large-scale GRN inference tasks.

3.6 Effects of single-cell data imputation

scRNA-seq data frequently include a high proportion of zero counts, which may be attributable to biological mechanisms or technical limitations in detection. The presence of excessive zeros can lead to biased results in GRN inference. We additionally evaluated the effect of single-cell data imputation (using DeepImpute) on GRN inference and found that imputing excessive zeros can improve scGRAIL's inference accuracy (Table S4).

3.7 Robustness study of the proposed model

To evaluate the robustness of different methods to noise perturbation, we introduced random noise at varying proportions to the input data to simulate experimental or measurement errors. We further assessed model robustness by injecting increasing levels of random noise into the input data and observed that scGRAIL consistently maintained the highest AUROC with the smallest performance variance across datasets, indicating superior stability under perturbations (Table S5).

3.8 Case study on breast cancer transcriptomics data

To validate scGRAIL's effectiveness on real biological data, we selected an authentic breast cancer transcriptome dataset and two hESC datasets of different scales from the Specific category, performing model training and regulatory prediction based on known TF–target gene regulatory relationships. In Fig. 4, we further conducted case studies on breast cancer and

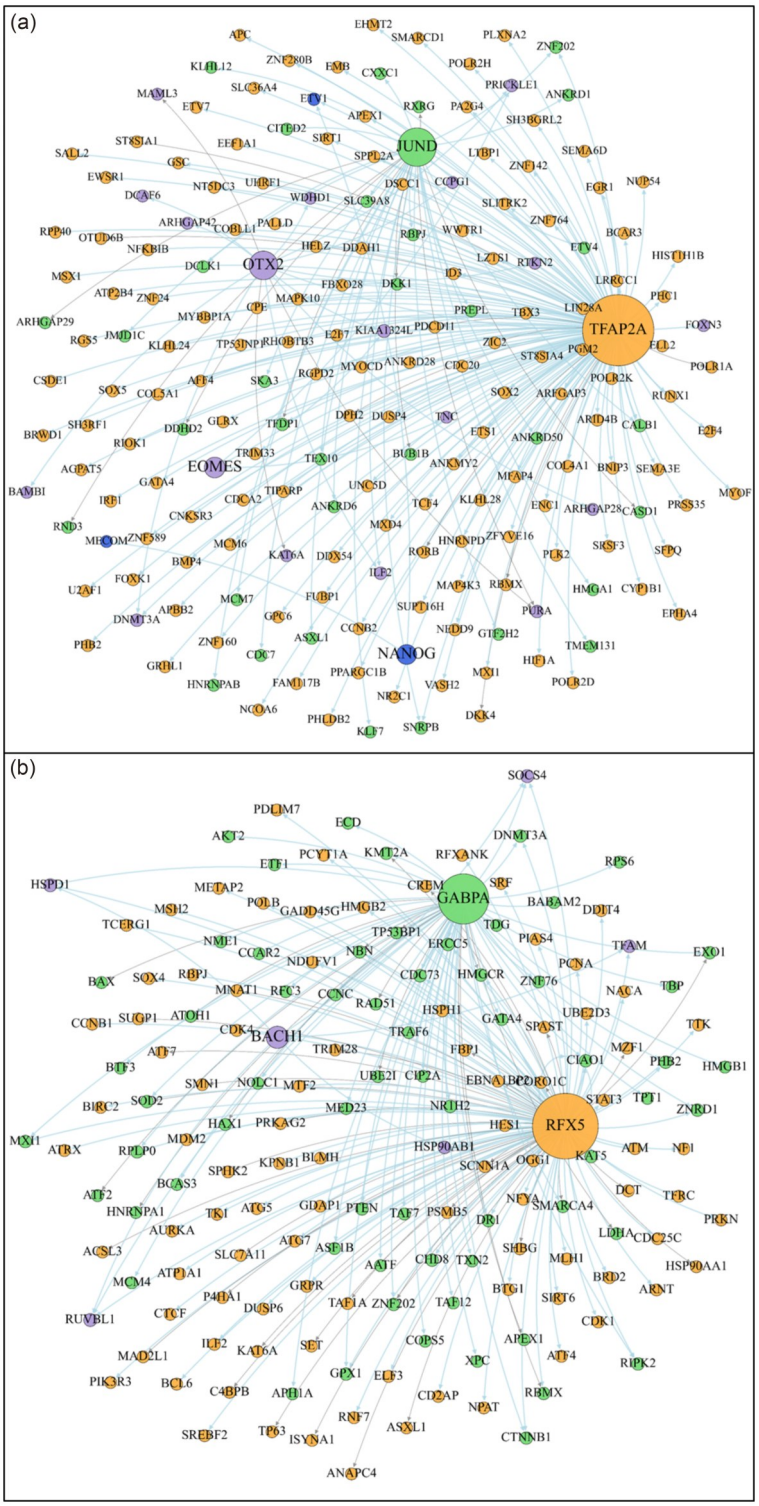

Fig. 4 Top 200 high-confidence TF–target pairs inferred by scGRAIL across multiple datasets. Edge colors indicate experimental support for regulatory relationships: blue (gold-standard) and gray (no prior annotation). (a) hESC dataset (TF500+) in the Specific category, highlighting key TFs (NANOG, EOMES, TFAP2A, JUND, and OTX2). (b) Breast cancer transcriptome results showing critical TFs (RFX5, GABPA, and BACH1). TF: transcription factor; hESC: human embryonic stem cell; NANOG: Nanog homeobox; EOMES: eomesodermin; TFAP2A: transcription factor AP-2α; JUND: JunD proto-oncogene, AP-1 transcription factor subunit; OTX2: orthodenticle homeobox 2; RFX5: regulatory factor X5; GABPA: GA-binding protein transcription factor subunit α; BACH1: BTB domain and CNC homolog 1.

hESC transcriptomic datasets to examine the biological plausibility of the inferred GRNs and the key TF–target regulators identified by scGRAIL. The visualizations, top-ranked regulatory pairs, and biological interpretations are provided in Section S3, Fig. S1, and Table S6.

4 Discussion and conclusions

The rapid development of scRNA-seq techniques has provided new opportunities to uncover GRNs between TFs and target genes. However, scRNA-seq data typically exhibit high dimensionality, sparsity, and significant noise, posing challenges for accurate GRN inference. Furthermore, existing methods often overlook global structural dependencies and fail to model regulatory directionality and causal mechanisms. To address these limitations, this study introduces scGRAIL, which incorporates structure-aware and directional modeling mechanisms to enhance the accuracy and biological interpretability of GRN inference.

The model first constructs a multi-level LGA module that integrates gene expression features with graph topology under a structural induction framework, enabling high-level structure-aware node embedding learning. Simultaneously, it introduces a node reconstruction mechanism to preserve original expression information, thereby enhancing the consistency and generalization capabilities of node representations. To strengthen the directional discrimination of regulatory edges, the model incorporates an asymmetric differential mechanism during the decoding phase, which explicitly models regulatory directionality based on vector differences between upstream and downstream node embeddings, thereby improving both the prediction accuracy and biological interpretability of regulatory edge inference. In this work, "biological interpretability" is primarily reflected in two aspects: first, high-confidence TF–target edges exhibit strong consistency with multiple sources of prior knowledge, including gold-standard networks, cell-type-specific regulatory networks, and STRING interactions; second, in datasets such as human embryonic stem cell (hESC) and breast cancer, the model highlights key TFs and their regulatory axes that are consistent with those reported in the literature. Comparative experiments with 12 state-of-the-art methods demonstrate that scGRAIL achieves optimal or near-optimal performance on most datasets, highlighting its superiority in identifying TF–target regulatory relationships.

Further structural ablation and hyperparameter analyses validate the rationality and effectiveness of the model's design and parameter settings, emphasizing the importance of integrating structural modeling with biological priors. We also investigated the impact of data imputation on model performance, revealing that scGRAIL achieves higher accuracy when inferring GRNs from imputed scRNA-seq data, suggesting that "zero-expression" may interfere with the capture of effective regulatory relationships. In terms of computational efficiency, although the graph sampling mechanism increases training time, scGRAIL remains highly efficient overall and exhibits strong scalability potential. Finally, case studies on multiple real biological datasets demonstrate that scGRAIL successfully identifies key TFs, such as transcription factor AP-2α (TFAP2A) and regulatory factor X5 (RFX5), that are supported by both standard networks and the existing literature. These findings further validate the model's biological interpretability and provide valuable insights into potential regulatory mechanisms that can guide subsequent experimental research.

Despite its strong performance, scGRAIL has certain limitations. Currently, the model relies on static prior graph structures and cannot dynamically adapt graph construction to data. In other words, the current implementation of scGRAIL reevaluates and refines edge weights on a high-confidence candidate TF–target prior network, and it is not yet capable of learning novel graph topology end-to-end from data at the whole-genome scale. Additionally, while the model accounts for directionality in decoding, its representation of causal relationships and state transitions (e.g., developmental trajectories) remains insufficient. Furthermore, scGRAIL currently treats cross-sectional single-cell expression profiles from hESC, human hepatocyte (hHEP), and other cell types as a set of mutually independent "snapshots," inferring static GRNs without explicitly incorporating lineage or pseudotime information to capture temporal regulatory dynamics during development. To mitigate this limitation, future efforts should focus on integrating more sophisticated models, such as causal discovery methods, pseudotime trajectories, and graph structure learning, to improve dynamic regulatory modeling and cross-tissue

generalization, and to accurately reflect temporal dynamics and lineage relationships among cells.

Data availability statement

The benchmark datasets are available from the Gene Expression Omnibus (GEO) database under accession numbers: GSE75748 (hESC), GSE81252 (hHEP), GSE48968 (mDC), GSE98664 (mESC), and GSE81682 (mHSC). The breast cancer dataset can be obtained from the China National Center for Bioinformation (CNCB) under project ID: GEND000024. The codes are available at https://github.com/gladex/scGRAIL.

Acknowledgments

The work was supported by the StarGuard Research Project, Guangzhou BonHeart Foundation, China (No. 2026-BH-KJHX-002).

Author contributions

Binhua TANG and Yingying FENG drafted and revised the manuscript; Yingying FENG, Xinyu GAO, Mengyao MAO, and Yujia ZHANG performed the coding and analysis; Binhua TANG led the project. All authors have read and approved the final manuscript, and therefore, have full access to all the data in the study and take responsibility for the integrity and security of the data.

Compliance with ethics guidelines

Binhua TANG, Yingying FENG, Xinyu GAO, Mengyao MAO, and Yujia ZHANG declare that they have no conflicts of interest.

This article does not contain any studies with human or animal subjects performed by any of the authors.

Declaration on the use of generative AI tools

No generative AI tool was used to generate scientific conclusions or to fabricate or alter data.

References

Chen GY, Liu ZP, 2022. Graph attention network for link prediction of gene regulations from single-cell RNA-sequencing data. *Bioinformatics*, 38(19):4522-4529. https://doi.org/10.1093/bioinformatics/btac559

Cuevas-Diaz Duran R, Wei HC, Wu JQ, 2024. Data normalization for addressing the challenges in the analysis of single-cell transcriptomic datasets. *BMC Genomics*, 25:444. https://doi.org/10.1186/s12864-024-10364-5

Gan YL, Yu JC, Xu GW, et al., 2024. Inferring gene regulatory networks from single-cell transcriptomics based on graph embedding. *Bioinformatics*, 40(5):btae291. https://doi.org/10.1093/bioinformatics/btae291

Hamilton WL, Ying R, Leskovec J, 2017. Inductive representation learning on large graphs. Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, p.1025-1035.

Huynh-Thu VA, Irrthum A, Wehenkel L, et al., 2010. Inferring regulatory networks from expression data using tree-based methods. *PLoS One*, 5(9):e12776. https://doi.org/10.1371/journal.pone.0012776

KC K, Li R, Cui F, et al., 2019. GNE: a deep learning framework for gene network inference by aggregating biological information. *BMC Syst Biol*, 13(Suppl 2):38. https://doi.org/10.1186/s12918-019-0694-y

Liu W, Teng ZJ, Li ZJ, et al., 2024. CVGAE: a self-supervised generative method for gene regulatory network inference using single-cell RNA sequencing data. *Interdiscip Sci*, 16(4):990-1004. https://doi.org/10.1007/s12539-024-00633-y

Mao G, Pang ZB, Zuo K, et al., 2023. Predicting gene regulatory links from single-cell RNA-seq data using graph neural networks. *Brief Bioinform*, 24(6):bbad414. https://doi.org/10.1093/bib/bbad414

Matsumoto H, Kiryu H, Furusawa C, et al., 2017. SCODE: an efficient regulatory network inference algorithm from single-cell RNA-seq during differentiation. *Bioinformatics*, 33(15):2314-2321. https://doi.org/10.1093/bioinformatics/btx194

Moerman T, Aibar Santos S, Bravo González-Blas C, et al., 2019. GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics*, 35(12):2159-2161. https://doi.org/10.1093/bioinformatics/bty916

Papili Gao N, Ud-Dean SMM, Gandrillon O, et al., 2018. SINCERITIES: inferring gene regulatory networks from time-stamped single cell transcriptional expression profiles. *Bioinformatics*, 34(2):258-266. https://doi.org/10.1093/bioinformatics/btx575

Shu HT, Zhou JT, Lian QY, et al., 2021. Modeling gene regulatory networks using neural network architectures. *Nat Comput Sci*, 1(7):491-501. https://doi.org/10.1038/s43588-021-00099-8

Wang JC, Chen YJ, Zou Q, 2023. Inferring gene regulatory network from single-cell transcriptomes with graph autoencoder model. *PLoS Genet*, 19(9):e1010942. https://doi.org/10.1371/journal.pgen.1010942

Wei PJ, Jin HW, Gao Z, et al., 2025. GAEDGRN: reconstruction of gene regulatory networks based on gravity-inspired graph autoencoders. *Brief Bioinform*, 26(3):bbaf232. https://doi.org/10.1093/bib/bbaf232

Xu J, Zhang AD, Liu F, et al., 2023. STGRNS: an interpretable transformer-based method for inferring gene regulatory networks from single-cell transcriptomic data. *Bioinformatics*, 39(4):btad165. https://doi.org/10.1093/bioinformatics/btad165

Yu WM, Lin ZR, Lan MF, et al., 2025. GCLink: a graph contrastive link prediction framework for gene regulatory network inference. *Bioinformatics*, 41(3):btaf074. https://doi.org/10.1093/bioinformatics/btaf074

Yuan L, Zhao L, Jiang YF, et al., 2024. scMGATGRN: a multiview graph attention network–based method for

inferring gene regulatory networks from single-cell transcriptomic data. *Brief Bioinform*, 25(6):bbae526. https://doi.org/10.1093/bib/bbae526

Yuan Y, Bar-Joseph Z, 2019. Deep learning for inferring gene relationships from single-cell expression data. *Proc Natl Acad Sci USA*, 116(52):27151-27158. https://doi.org/10.1073/pnas.1911536116

Zhao MY, He WY, Tang JJ, et al., 2022. A hybrid deep learning framework for gene regulatory network inference from single-cell transcriptomic data. *Brief Bioinform*, 23(2): bbab568. https://doi.org/10.1093/bib/bbab568

Zheng RQ, Li M, Chen X, et al., 2019. BiXGBoost: a scalable, flexible boosting-based method for reconstructing gene regulatory networks. *Bioinformatics*, 35(11):1893-1900. https://doi.org/10.1093/bioinformatics/bty908

Supplementary information

Sections S1–S3; Tables S1–S6; Fig. S1