



## OPEN ACCESS

## EDITED BY

Xiaoqing Ru,  
Hebei University of Engineering, China

## REVIEWED BY

Ju Xiang,  
Changsha University of Science and  
Technology, China  
XianFang Tang,  
Wuhan Textile University, China

## \*CORRESPONDENCE

Mingming Qi,  
✉ webqmm1974@163.com  
Wei Zhang,  
✉ zw6676@163.com  
Yifan Chen,  
✉ cyf176@hnu.edu.cn

RECEIVED 22 May 2025

ACCEPTED 16 June 2025

PUBLISHED 07 July 2025

## CITATION

Li Z, Qi M, Huang J, Zhang W, Tan X and Chen Y  
(2025) Geometry-enhanced graph neural  
networks accelerate circRNA therapeutic  
target discovery.  
*Front. Genet.* 16:1633391.  
doi: 10.3389/fgene.2025.1633391

## COPYRIGHT

© 2025 Li, Qi, Huang, Zhang, Tan and Chen. This  
is an open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# Geometry-enhanced graph neural networks accelerate circRNA therapeutic target discovery

Zhen Li<sup>1</sup>, Mingming Qi<sup>2\*</sup>, Juyuan Huang<sup>3</sup>, Wei Zhang<sup>3\*</sup>, Xu Tan<sup>1</sup>  
and Yifan Chen<sup>4\*</sup>

<sup>1</sup>School of Artificial Intelligence, Shenzhen Institute of Information Technology, Shenzhen, China,

<sup>2</sup>School of Data Science and Artificial Intelligence, Wenzhou University of Technology, Wenzhou, China,

<sup>3</sup>Department of Gynecology, Zhongnan Hospital of Wuhan University, Wuhan, China, <sup>4</sup>School of  
Information Engineering, Changsha Medical University, Changsha, China

Circular RNAs (circRNAs) play pivotal roles in various biological processes and disease progression, particularly in modulating drug responses and resistance mechanisms. Accurate prediction of circRNA-drug associations (CDAs) is essential for biomarker discovery and the advancement of therapeutic strategies. Although several computational approaches have been proposed for identifying novel circRNA therapeutic targets, their performance is often limited by inadequate modeling of higher-order geometric information within circRNA-drug interaction networks. To overcome these challenges, we propose G2CDA, a geometric graph representation learning framework specifically designed to enhance the identification of CDAs and facilitate therapeutic target discovery. G2CDA introduces torsion-based geometric encoding into the message propagation process of the circRNA-drug network. For each potential association, we construct local simplicial complexes, extract their geometric features, and integrate these features as adaptive weights during message propagation and aggregation. This design promotes a richer understanding of local topological structures, thereby improving the robustness and expressiveness of learned circRNA and drug representations. Extensive benchmark evaluations on public datasets demonstrate that G2CDA outperforms state-of-the-art CDA prediction models, particularly in identifying novel associations. Case studies further confirm its effectiveness by uncovering potential drug interactions with the ALDH3A2 and ANXA2 biomarkers. Collectively, G2CDA provides a robust and interpretable framework for accelerating circRNA-based therapeutic target discovery and streamlining drug development pipelines. Our code are archived in: <https://github.com/lizhen5000/G2CDA>.

## KEYWORDS

drug development, circRNA therapeutic targets, geometric graph representation, biomarker discovery, circRNA-drug network

## Introduction

Circular RNAs (circRNAs) are highly stable endogenous non-coding RNAs resistant to nuclease degradation due to their covalently closed loop structure (Geng et al., 2020). Numerous studies indicate that circRNAs play crucial roles in various biological processes, such as transcriptional regulation, miRNA sponge effects, and protein interaction modulation (Zang et al., 2020). In disease research, particularly in cancer, circRNA aberrant expression is closely linked to tumorigenesis, progression, and drug resistance (Xu et al., 2020). Mounting evidence indicates that circRNAs play a crucial role in drug response and resistance mechanisms (Xie et al., 2020; Ding et al., 2021a; Wang et al., 2023; Liu et al., 2018). Accurate circRNA-drug association (CDA) identification is vital for discovering novel therapeutic targets, uncovering drug response mechanisms, and supporting personalized therapy. However, experimental CDA identification is often time-consuming, costly, and inefficient. Thus, developing efficient computational methods is essential to expedite circRNA therapeutic target and drug discovery.

Current deep learning methods have shown remarkable performance in target identification and drug discovery, further advancing CDA prediction. For instance, Deng et al. introduced GATECDA, a graph attention auto-encoder framework for predicting circRNA-drug sensitivity associations (Deng et al., 2022). Yang et al. developed MNGACDA, a graph auto-encoder network that integrates multiview techniques and node-level attention (Yang and Chen, 2023). This model combines multiple information sources from circRNAs and drugs to form a multimodal network, uses multiview techniques to learn low-dimensional embedded representations of circRNAs and drugs, and predicts their association scores through an inner product decoder. Luo et al. presented DPMGCDA, which combines a dual perspective learning mechanism with a pathway masking graph autoencoder to enhance the modeling of circRNA-drug sensitivity relationships (Luo and Deng, 2024). Additionally, Huang et al. proposed DeepHeteroCDA, a computational framework based on a multi-scale heterogeneous network structure and graph attention mechanism, designed to more comprehensively capture the complex sensitivity relationships between circRNAs and drugs (Huang et al., 2025).

In recent years, various innovative drug discovery models have emerged, offering new methodological support and references for circRNA therapeutic target discovery and associated drug development. For example, Zhou et al. proposed the JDASA-MRD model, which combines deep autoencoders and subgraph augmentation to infer microbial responses to drugs (Zhou et al., 2024a). This model captures higher-order neighbor relationships, constrains local message propagation, and enhances drug and microbial representations with remarkable performance. Additionally, Zhou et al. integrated self-supervised strategies and masking mechanisms to explore miRNA responses to small molecule drugs (Zhou et al., 2024b). They also noted that existing drug discovery models often overlook indirectly connected drug-target pairs and proposed a novel model from a global-local perspective (Zhou et al., 2024c). Furthermore, Wei et al. employed multisource prompting with large model technology for efficient drug repurposing (Wei et al., 2024a). Notably, Wei et al. used integrated deep learning to accurately identify unknown drug-

target interactions and validated their findings through experiments (Wei et al., 2024b). Although these models do not directly address circRNA therapeutic targets, they provide valuable guidance and references.

Deep learning models like Graph Neural Networks (GNNs) have shown strong performance in CDA prediction tasks, but they also present certain limitations. First, conventional GNNs often rely on shallow feature extraction and struggle to capture higher-order neighbor information, as deeper message propagation may lead to the well-known “oversmoothing” problem. However, higher-order relationships in the circRNA-drug interaction map frequently encode crucial regulatory mechanisms. Second, many existing models depend on complex, task-specific feature engineering pipelines, which can hinder their scalability and generalizability. Moreover, these models tend to overlook the structural semantics of the circRNA-drug graph—namely, the local geometric and topological relationships among CDAs—thereby limiting their ability to model intricate interaction patterns. Recent advances in geometric deep learning have demonstrated that integrating geometric and topological priors, such as curvature, torsion, and simplicial structures, into GNN architectures can enhance representation power in various scientific domains, including protein structure modeling, drug discovery, and biological network analysis. In particular, torsion-based geometric features, which reflect how elements twist or bend in a local topological space, have been shown to encode important structural signals in molecular graphs and manifolds. Despite this progress, their application in the context of circRNA-drug association modeling remains largely unexplored. To address the aforementioned issues, this study proposes a novel graph representation learning framework that explicitly incorporates geometric information to identify potential CDAs. By embedding an analytic torsion technique into the message propagation process over the circRNA-drug graph, our method introduces a new form of local geometric encoding. For each CDA, the framework constructs a local simplicial complex, computes its torsion value, and uses it as an adaptive weight to modulate the message-passing process. This mechanism enhances the model’s capacity to perceive and utilize higher-order geometric and topological cues within the interaction graph, thereby improving its ability to accurately capture potential sensitivity associations between circRNAs and drugs. As a result, our approach contributes to more reliable circRNA therapeutic target discovery and accelerates drug development. Our contributions are summarized as follows:

- 1) We propose G2CDA, a novel geometric graph representation learning framework for circRNA-drug association (CDA) prediction. Extensive benchmarking demonstrates its strong ability to identify novel CDAs, offering a promising tool to accelerate circRNA-based therapeutic target discovery and drug development.
- 2) A key innovation of G2CDA is the integration of analytic torsion into the prediction pipeline. By constructing local simplicial complexes and computing torsion values, we transform high-order geometric structures into adaptive, learnable weights—enabling the model to better capture complex circRNA-drug interaction patterns.

- 3) The computed torsion values are further leveraged as core weights in the message propagation mechanism of the graph neural network. This torsion-guided propagation enhances the model's focus on geometrically significant substructures, improving representation learning for both circRNAs and drugs.
- 4) We validate the effectiveness of G2CDA through comprehensive experiments on multiple public datasets, where it consistently outperforms state-of-the-art methods. Case studies demonstrate its potential for translational applications, including the identification of candidate drugs targeting key biomarkers such as ALDH3A2 and ANXA2.

## Materials and methods

This study focuses on identifying potential circRNA–drug associations (CDAs) to facilitate circRNA therapeutic target discovery and drug development. To this end, we propose G2CDA, a novel graph neural network framework that integrates geometric information to capture higher-order interactions between circRNAs and drugs for more accurate CDA prediction. A key innovation of G2CDA is the introduction of the geometric quantity “torsion,” which intuitively reflects the local bending or twisting within the graph structure. By calculating torsion through local simplicial complexes, the model effectively quantifies subtle structural perturbations in the interaction network, analogous to how regulatory influences propagate and modulate biological networks. This approach converts local geometric information into learnable weights, enhancing the model's ability to capture complex and nuanced relationships between circRNAs and drugs. Furthermore, during message propagation, the torsion information guides both the direction and magnitude of message flow independently of node features, enabling the model to focus on critical structural regions that are most relevant for accurate CDA prediction. Together, these innovations contribute to improved predictive performance. The following sections provide detailed descriptions of the datasets, methodologies, and theoretical principles underpinning this approach.

## Data preparation

This study evaluated the proposed model and comparative models using a publicly available dataset from previous work (Huang et al., 2025). The data, sourced from the circRic database (Vromman et al., 2021), identifies and annotates circRNA molecules across approximately 1,000 human cancer cell lines. The dataset includes experimental data on circRNA responses to drugs, aiming to provide a high-confidence circRNA resource for cancer research and to explore their expression patterns across different cancer types. After screening and removing low-confidence CDAs, 4314 CDAs remained, involving 271 circRNAs and 218 drugs. A circRNA–drug graph was constructed for message propagation. Additionally, 4314 unknown circRNA–drug pairs were randomly selected as negative samples. This balanced sampling approach ensures training process stability.

## Methods

### Model overview

As shown in Figure 1, the G2CDA model workflow comprises five key modules: constructing the circRNA–drug graph, extracting CDA-centered subgraphs, calculating CDA torsion, integrating geometric message propagation, and training and inference. Module (A) constructs the circRNA–drug graph using known CDAs from the dataset and initial drug/circRNA representations. Module (B) extracts all CDA-centered subgraphs. Module (C) calculates CDA torsion via 1-simplicial complex. Module (D) fuses these representations to predict drug–circRNA pair scores through an MLP. Module (E) converts CDA torsion into weights for message propagation to refine drug and circRNA representations.

### Problem description

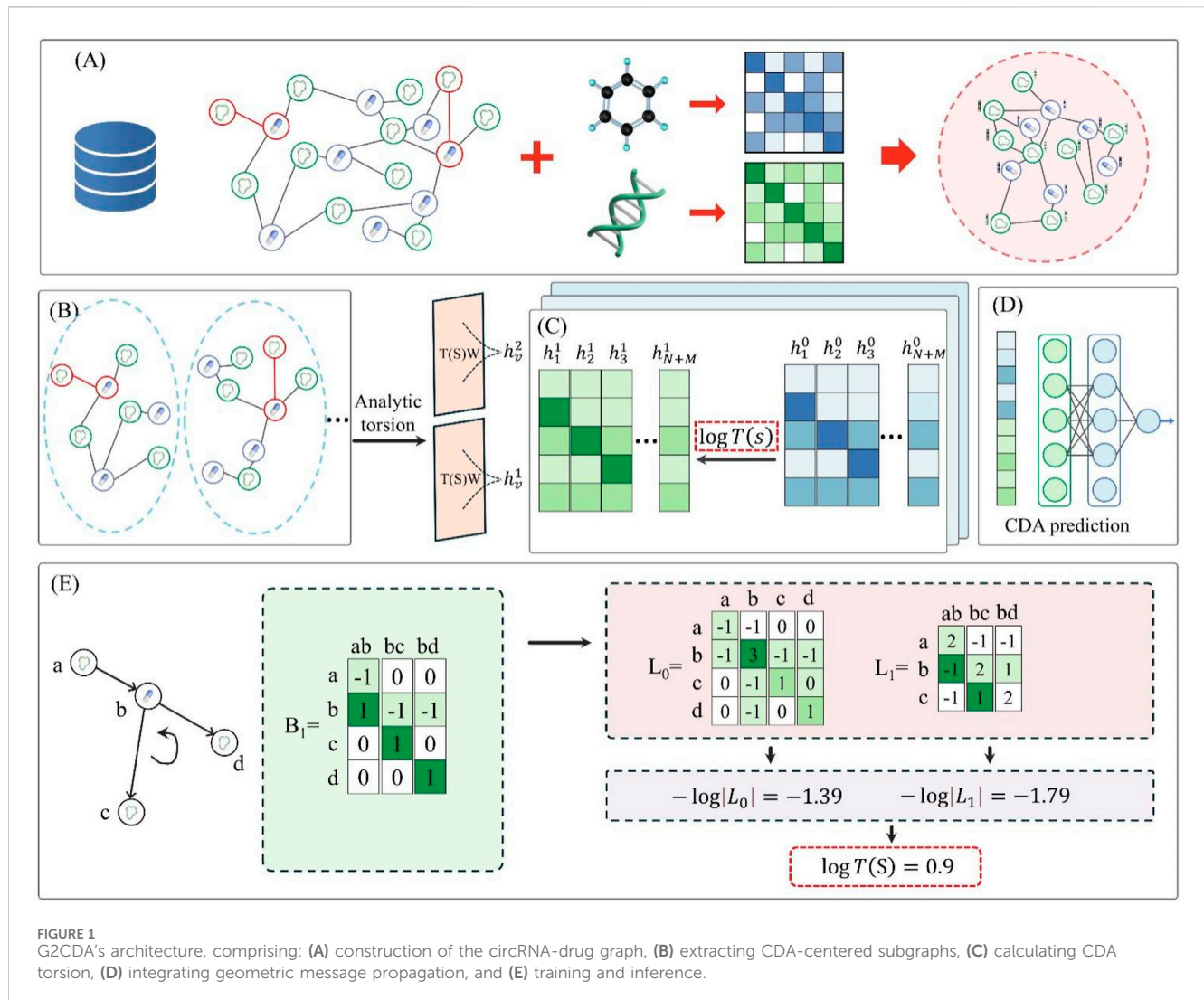
This study first constructs a circRNA–drug graph using known CDAs. Specifically, the graph  $G$  is defined as  $G = \langle V_C, V_D, E \rangle$ , where  $V_C, V_D$  denote the set of circRNAs and drugs nodes, respectively, and  $E$  represents the set of observed CDAs.  $A \in \mathbb{R}^{N \times M}$  is the adjacency matrix of graph  $G$ , where  $N$  and  $M$  indicate the number of circRNAs and drugs, respectively.  $A_{uv}$  represents the relationship between circRNA  $u$  and drug  $v$ , where  $A_{uv} = 1$  indicates an association and  $A_{uv} = 0$  indicates no association. The initial representation of a circRNA is defined as  $H^0 = h_1^0, h_2^0, h_3^0, \dots, h_N^0$ , and that of a drug as  $H^0 = h_{N+1}^0, h_{N+2}^0, h_{N+3}^0, \dots, h_{N+M}^0$ . These initial representations can be derived from various sources, such as k-mer frequency characteristics of circRNA sequences and SMILES features of drugs.

Secondly, the study introduces the “analytic torsion” theory, focusing on the local geometric information of CDAs and transforming it into learnable weights to guide message propagation and enhance key region recognition. Specifically,  $S$  represents the simplicial complex, with  $S_{uv}$  denoting the simplicial complex of edge  $A_{uv}$  (the edge between circRNA  $u$  and drug  $v$ ). Each  $A_{uv} = 1$  has a corresponding  $S_{uv}$ , and  $T(S_{uv})$  calculates the resolved torque for  $S_{uv}$ . Additionally,  $Z_g$  is the  $g$ -th Hodge Laplacian, with  $|Z_g|$  representing its determinant. If  $Z_g$  has zero eigenvalues, the determinant is the product of all non-zero eigenvalues.  $B_g$  is the boundary matrix corresponding to  $Z_g$ , and  $\gamma_g(p)$  is the regularization function for  $Z_g$  eigenvalues.

The study aims to enhance message propagation by leveraging the local geometric higher-order information of each CDA based on known CDAs using the “analytic torsion” theory, enabling accurate identification of potential CDAs from unknown circRNA–drug pairs.

### Graph neural network (GNN)

GNN effectively resolves structural information and has been widely applied in bioinformatics, including molecular interaction networks and molecular structure analysis (Cai et al., 2024; Ma et al., 2024; Wang et al., 2024a; Wang et al., 2024b; Wang et al.,



2024c; Zhang et al., 2024). This research employs the GNN model to model known CDAs, perform message propagation in the circRNA-drug graph, and output circRNA and drug representations. Specifically, the GNN model's message propagation involves aggregating and updating information from circRNA and drug nodes. Let  $V = V_C \cup V_D$ ,  $v_a \in V$ , the aggregation process for circRNAs (or drugs) is defined as Equation 1:

$$h_a^i = f\left(h_a^{i-1}, \sum_{b \in N(a)} h_b^{i-1}, A, \alpha\right) \quad (1)$$

where  $h_a^i$  represents the representation of node  $a$  at layer  $i$ , and  $f$  represents the aggregation function.  $N(a)$  is the set of neighbors of node  $a$ , and  $\alpha$  regulates the weights of the neighbors relative to their own information. The GNN model then updates the circRNA (or drug) representation using an activation function and iterates for a set number of times. This study tested the GCN (Kipf and Welling, 2016a), GAT (Veličković et al., 2017), and GIN (Xu et al., 2018) models, incorporating "analytic torque" into CDA prediction tasks.

## Analytic torsion

A simplicial complex  $S$  is defined as a collection of multiple simplices. From the total set of circRNAs and drugs,  $g+1$  nodes are selected to form a  $g$ -simplex  $\pi^g$ , denoted as  $v_0, v_1, \dots, v_g$ . For any  $\pi^g$ , its non-empty subset is defined as a face, which is the boundary of  $\pi^g$  when the dimension is  $g-1$ . Geometrically, a  $g$ -simplex represents the convex hull formed by  $g+1$  nodes. According to this theory, 0-, 1-, 2-, and 3-simplices correspond to vertices, edges, triangles, and tetrahedrons, respectively.

Given an oriented  $S$  with nodes in a predetermined order  $v_0, v_1, \dots, v_g$ , a  $g$ -chain is defined as the sum of  $g$ -simplices in  $S$ , denoted as  $d = \sum_c \pi_c^g$ . The set of all  $g$ -chains in  $S$  forms the  $g$ -chain group  $D_g$ . A chain complex is formed by integrating a series of chain groups and their corresponding boundary operations:  $0 \leftarrow \rho_0 D_0 \leftarrow \rho_1 D_1 \leftarrow \rho_2 D_2 \leftarrow \dots \leftarrow \rho_g D_g \leftarrow \dots$ , where  $\rho_g: D_g \rightarrow D_{g-1}$ . The boundary matrix  $B_g$  is derived from  $\rho_g$ . The  $g$ -th step is defined as  $\rho_g \pi^g = \sum_{j=0}^g (-1)^j [v_0, v_1, \dots, \hat{v}_j, \dots, v_g]$ , where  $\hat{v}_j$  indicates the removal of  $v_j$ , and  $\rho_{g-1} \rho_g = 0$ . Let the adjoint operator of  $\rho_g$  be  $\delta^g$ . The  $g$ -th Hodge Laplacian component of  $S$  is calculated as



$\Delta_g = \delta^g \rho_g + \delta^{g+1} \rho_{g+1}$ . Therefore, the  $g$ -th Hodge Laplacian matrix of  $S$  is computed as  $Z_g = B_g^T B_g + B_{g+1} B_{g+1}^T$ .

Therefore, in  $S$ , the Hodge Laplacian corresponding to the 1-simplex is  $Z_0 = B_1 B_1^T$ . Then, the highest-order Hodge Laplacian corresponding to the  $m$ -simplex is calculated as:  $Z_m = B_m^T B_m$ . The computation of  $Z_g$  depends on the simplex relations. For example, when  $g = 0$ ,  $Z_0$  can be calculated as Equation 2:

$$Z_0(u, v) = \begin{cases} d(\pi_u^0), & \text{if } u = v; \\ -1, & \text{if } u \neq v \text{ and } \pi_u^0 \cap \pi_v^0; \\ 0, & \text{if } u \neq v \text{ and } \pi_u^0 \cap \pi_v^0; \end{cases} \quad (2)$$

where  $\text{degree}(\pi_u^0)$  denotes the degree of  $\pi_u^0$ ,  $\pi_u^0 \cap \pi_v^0$  indicates that simplices  $\pi_u^0$  and  $\pi_v^0$  are upper adjacent, and  $\pi_u^0 \cap \pi_v^0$  denotes non-upper adjacency. Furthermore, the value of  $Z_g(u, v)$  can be inferred based on geometric information such as upper adjacency, lower adjacency, and directional consistency between two  $g$ -simplices (Shen et al., 2025). Specifically, on the circRNA-drug graph, this can be treated as known geometric information corresponding to CDAs.

## Analytical torsion calculation

$Z_g$  is a positive semi-definite matrix with eigenvalues  $\{\theta_j\}$ . Its regularization function  $\gamma_g(p)$  is computed as  $\gamma_g(p) = \sum_{\theta_j > 0} \frac{1}{\theta_j^p}$ , where  $p$  is the regularization factor. Then, the analytical torsion for  $S$  can be calculated in logarithmic form as follows:

$$\log T(S) = \frac{1}{2} \sum_{g=0}^{N+M} (-1)^g g \gamma_g^{(1)}(0) \quad (3)$$

where  $\gamma_g^{(1)}(0)$  represents the first derivative of  $\gamma_g(p)$  evaluated at  $p = 0$ . Assuming the determinant of  $Z_g$  is the product of its non-zero eigenvalues  $|Z_g| = \prod_{\theta_j > 0} \theta_j$ , then  $\gamma_g^{(1)}(0) = -\log |Z_g|$ . Equation 3 can thus be further transformed as follows:

$$\log T(S) = \frac{1}{2} \sum_{g=0}^{N+M} (-1)^{g+2} g \log |Z_g| \quad (4)$$

For the 1-dimensional simplicial complex  $S$  of the circRNA-drug graph, its analytic torsion  $T(S)$  is calculated as  $T(S) = |Z_1|^{\frac{1}{2}}$ , where  $Z_1 = B_1^T B_1$ . For the 2-dimensional simplicial complex, the analytic torsion is computed as  $T(S) = \frac{|Z_1|^{\frac{1}{2}}}{|Z_2|^{\frac{1}{2}}}$ . The calculation of analytic torsion depends on the simplicial complex, and the specific process is illustrated in Figure 1E.

## Analytic torsion on CircRNA-Drug network

This study incorporates analytic torsion calculations into message propagation on the circRNA-drug graph, based on previous work (Reidemeister, 1935). Within the GNN model, analytic torsion is transformed into CDA weights for message aggregation. Equation 4 is formalized as:

$$h_a^i = \text{sigmod} \left( \sum_{b \in N(a) \cup a} \frac{1}{\sqrt{d(a)} \sqrt{d(b)}} |\log T(S_{a,b})| W h_b^{i-1} \right) \quad (5)$$

where  $\text{sigmod}()$  is the activation function,  $N(a)$  represents the neighbors of node  $a$ , and  $d()$  indicates node degree.  $h_b^0$  corresponds

to the initial representation of the  $b$ -th node on the circRNA-drug graph, defined as the  $b$ -th row of matrix  $A$ .

In Equation 5,  $S_{a,b}$  refers to the local simplicial complex constructed from CDA  $\langle a, b \rangle$ . To determine the local simplicial complex, one common approach is to extract a subgraph from direct neighbors. This subgraph includes nodes  $a, b$ , all their neighbors, and all edges (CDAs) between circRNAs and drugs within this group. A second-order subgraph can be derived from the second-order neighbors of circRNAs (or drugs). For a 2-dimensional 2-simplex complex  $S_{a,b}$ , a triangle should be constructed from three nodes. The model can incorporate higher-order neighbors to enhance local geometric information. If any circRNA and drug within the subgraph are associated,  $g+1$  entities can form a  $g$ -simplex. The computation of  $S_{a,b}$  involves the dimension  $Q$  of the simplicial complex and the subgraph order  $P$ . When  $Q$  and  $P$  are both 1, the message propagation process on the circRNA-drug graph is illustrated in Figure 1E.

## CDA prediction

This study integrates GNN models (e.g., GCN, GAT, GIN) for torque analysis on the circRNA-drug graph. Once the models complete the preset number of message propagation processes, the final representations of circRNAs and drugs are extracted. The study aims to calculate the probability of edges (associations) between circRNAs and drugs based on these representations. Let the final representations of circRNA  $u$  and drug  $v$  be  $h_u$  and  $h_v$ , respectively. The edge probability for the  $\langle u, v \rangle$  pair is predicted as Equation 6:

$$\hat{q}_{xy} = \text{MLP}(\| (h_u + h_v, h_u \odot h_v, h_u \| h_v) \|) \quad (6)$$

where MLP denotes a multilayer perceptron,  $\|$  represents concatenation, and  $\odot$  denotes the Hadamard product. The study employs the binary cross-entropy (BCE) function to compute the final loss.

## Results

### Performance comparison

This study systematically evaluates the G2CDA model's performance in CDA prediction through comprehensive comparison experiments. The experimental design includes multiple benchmark model comparisons. For CDA prediction, the study selects several state-of-the-art models: GATECDA (Deng et al., 2022), MNGACDA (Yang and Chen, 2023), DPMGCDA (Luo and Deng, 2024), and DeepHeteroCDA (Huang et al., 2025). These models represent the current best practices in CDA prediction. Additionally, other GNN-based interaction prediction models are included, such as variational graph auto-encoders (VGAE) (Kipf and Welling, 2016b), Multi-view variational graph auto-encoder with matrix factorization (VGAMF) (Ding et al., 2021b), graph convolutional network based GCNMDA (Long et al., 2020), and local attention graph convolutional network (LAGCN) (Yu et al., 2021). Although these models have not been previously applied to CDA prediction, they

**TABLE 1** Comparison results between G2CDA and other advanced models (%).

Methods	AUC	AUPR	F1-score	Accuracy	Recall
SVM	86.48	85.47	80.49	79.28	85.50
RF	88.81	88.85	82.04	81.65	83.83
KNN	86.42	87.60	79.26	79.01	80.20
XGBoost	89.97	89.96	82.94	82.52	84.94
AdaBoost	88.52	88.88	82.07	81.55	84.43
VGAE	86.28	87.30	79.88	78.92	82.27
VGAMF	87.40	86.62	81.76	81.04	84.37
GCNMDA	87.78	87.62	81.98	81.19	84.28
GATECDA	88.46	89.29	81.94	81.68	83.16
LAGCN	89.82	90.23	82.85	82.61	84.03
MNGACDA	90.98	91.50	84.13	83.79	85.92
DPMGCDA	90.15	91.73	84.08	84.24	83.25
DeepHeteroCDA	92.33	92.93	85.61	85.20	88.07
MLP	93.53	93.09	86.74	86.59	86.63
G2CDA	94.91	94.48	87.06	87.70	83.03

have demonstrated success in other association prediction tasks. Furthermore, five classical machine learning algorithms are selected as benchmark references: Support Vector Machines (SVM), Random Forest (RF), K Nearest Neighbors (KNN), Extreme Gradient Boosting (XGBoost), Adaptive Boosting (AdaBoost) and Multilayer Perceptron (MLP). These algorithms, widely used in supervised learning, provide a performance comparison at the traditional method level. The multi-dimensional comparison system ensures the evaluation results are both comprehensive and reliable by covering both domain-specific and generalized models. Following established methodologies (Wang et al., 2024d; Wei et al., 2024c; Feng et al., 2025; Fu et al., 2025; Chen et al., 2024), we adopted five key metrics: AUC (Area Under the Curve), AUPR (Area Under the Precision-Recall Curve), Accuracy, Recall, and F1-score.

According to Table 1, the G2CDA model outperforms all comparative methods across all evaluation metrics, achieving an AUC of 94.91%, AUPR of 94.48%, F1-score of 87.06%, accuracy of 87.70%, and recall of 83.03%. This demonstrates its superior prediction ability and robustness. Traditional machine learning methods like SVM, RF, KNN, XGBoost, and AdaBoost struggle to effectively model graph structures and complex nonlinear relationships, resulting in inferior overall performance. While generic GNN models such as VGAE, VGAMF, GCNMDA, and LAGCN can capture some structural information, they lack specific optimization for CDA prediction, leading to a significant performance gap. Among task-specific models, DeepHeteroCDA, MNGACDA, and DPMGCDA show relatively strong performance, particularly DeepHeteroCDA with an AUC of 92.33%, but they still underperform the G2CDA model across key metrics. Overall, the G2CDA model excels not only in prediction accuracy but also in

handling sample imbalance, structural modeling, and multimodal information fusion, confirming its effectiveness for CDA prediction tasks.

## Performance evaluation

In Table 2, the five - fold cross - validation results show that the G2CDA model has excellent stability and generalization in CDA prediction. The model's average AUC is 94.58% ( $\pm 0.53\%$ ), and average AUPR is 94.02% ( $\pm 0.43\%$ ), indicating its strong discriminative ability comparable to top existing models. Notably, the F1 score ( $85.97\% \pm 1.26\%$ ) and accuracy ( $86.67\% \pm 1.15\%$ ) show fluctuations of less than 1.5 standard deviations, demonstrating the model's robustness to data distribution changes. In Fold 1 and Fold 3, the AUC reaches 94.91% and 94.93%, respectively, with the F1 score surpassing 87%. The recall rate, while fluctuating between 79.69% and 83.03%, remains consistently above 80%, highlighting the model's effectiveness in retaining positive sample features via its torsion - weight - guided information dissemination mechanism. These findings confirm that the local - simple - complex - shape - based torsion counting strategy enhances the model's geometric interaction feature capture ability, and the non - original - feature - dependent attention propagation mechanism offers significant advantages for modeling complex biological relationships. This work provides a novel methodological framework for CDA prediction.

## Parameter experiments

This study examines various factors, including GNN encoder types, GNN layer numbers, hidden layer dimensionalities, output layer dimensionalities and node initial representations, necessitating multiple parameter experiments to verify the G2CDA model's stability and establish parameter setting guidelines.

Different GNN models were analyzed for their information transfer and feature extraction capabilities. Three representative GNN variants were experimented with to adapt to diverse graph structures and task requirements. GCN, leveraging Laplacian smoothing, suits homogeneous graphs. GIN, with multilayer perceptron-like expressiveness, handles complex heterogeneous graphs. GAT introduces attention mechanisms to enhance sensitivity to key connections. Comparative experiments among these encoders confirmed their performance and adaptability under different graph encoding approaches.

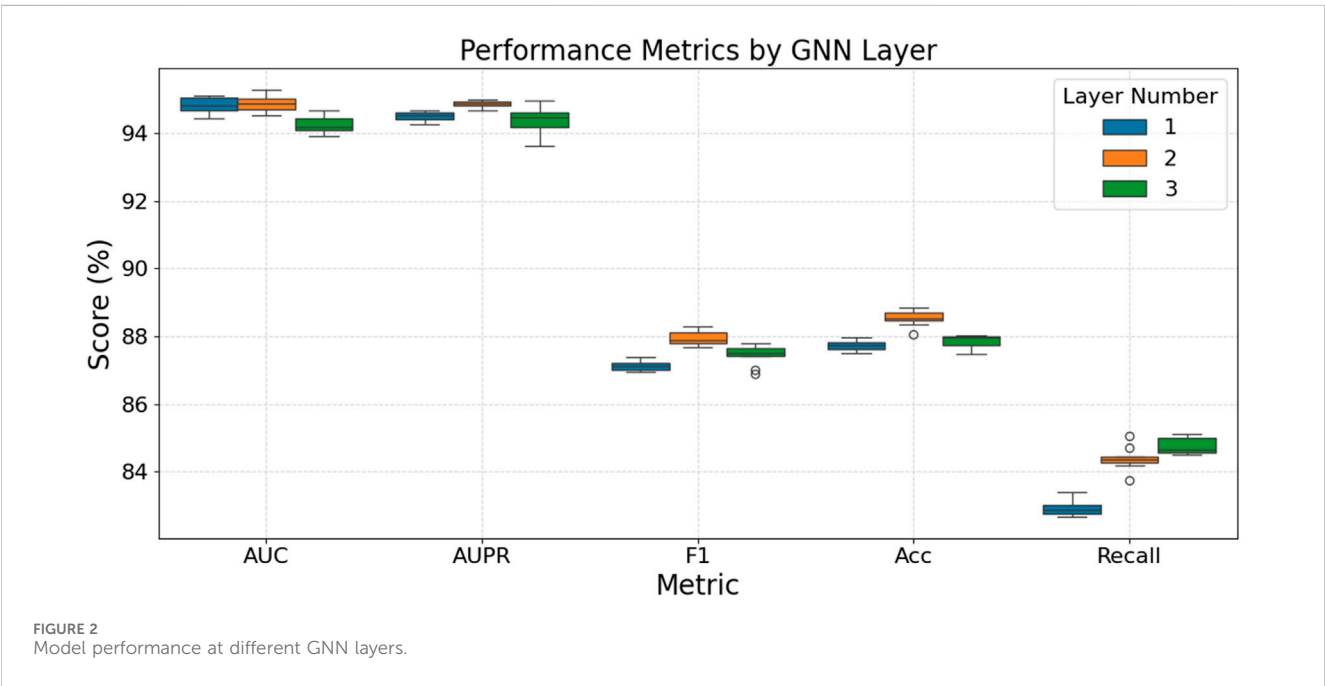
Table 3 evaluates their performance across metrics like AUC, AUPR, F1-score, Accuracy, and Recall. GCN showed the best overall performance with an AUC of 94.91%, AUPR of 94.48%, F1 score of 87.06%, and accuracy of 87.70%. GIN had slightly lower accuracy (86.72%), AUC (94.45%), and AUPR (94.06%), with an F1 score of 85.71% and a significantly lower Recall (79.58%), indicating weaker positive sample recall. GAT achieved the highest Recall (88.05%), suggesting its attention mechanism aids positive sample recognition, but had lower AUC (93.90%) and AUPR (93.41%), showing shortcomings in overall discriminative ability and stability. In summary, GCN balanced all five metrics best, maintaining high accuracy and F1 scores while ensuring good recall and

TABLE 2 Results of 5-fold cross validation of G2CDA model (%).

Folds/metrics	AUC	AUPR	F1-score	Accuracy	Recall
1	94.91	94.48	87.06	87.7	83.03
2	93.74	93.63	84.33	85.09	79.69
3	94.93	94.49	87.11	87.76	82.90
4	93.93	93.55	84.81	85.68	79.95
5	94.38	93.95	86.53	87.11	82.60
Average	94.58 ± 0.53	94.02 ± 0.43	85.97 ± 1.26	86.67 ± 1.15	81.63 ± 1.67

TABLE 3 Results of G2CDA with different GNN encoders (%).

Models/metrics	AUC	AUPR	F1-score	Accuracy	Recall
GCN	94.91	94.48	87.06	87.70	83.03
GIN	94.45	94.06	85.71	86.72	79.58
GAT	93.90	93.41	87.15	86.98	88.05



generalization, making it the optimal encoder choice. GAT’s high recall also offers a feasible alternative for applications requiring high positive sample sensitivity.

Figure 2 assesses the impact of different GNN layer depths on model performance. Overall, a 2-layer GNN architecture demonstrates the optimal balance and stability across these metrics, achieving the highest median scores for AUPR, F1-score, and Accuracy, while also exhibiting strong and stable performance in Recall. In contrast, although a 1-layer GNN performs best on the AUC metric, it shows a notable deficiency in Recall and is generally not the optimal choice for F1-score and Accuracy. Increasing the GNN layers to three does not yield significant performance improvements across most metrics; instead, it may lead to a

slight decrease in some indicators (such as AUC) or exhibit greater volatility and instability in metrics like F1-score and Recall, even though its median Recall is marginally higher than that of the 2-layer model. Thus, the 2-layer architecture optimally balances performance, supporting the “shallow geometry perception” design concept.

Figure 3 examines how different GNN hidden layer dimensionalities affect model performance. As depicted, for the AUC and AUPR metrics, all tested hidden dimensions exhibit very similar high performance, with scores generally concentrated between 94% and 95%, indicating that hidden dimension size has a minimal effect on these two metrics, although dimensions 64, 128, and 256 show slightly better stability for AUPR. However, for F1-

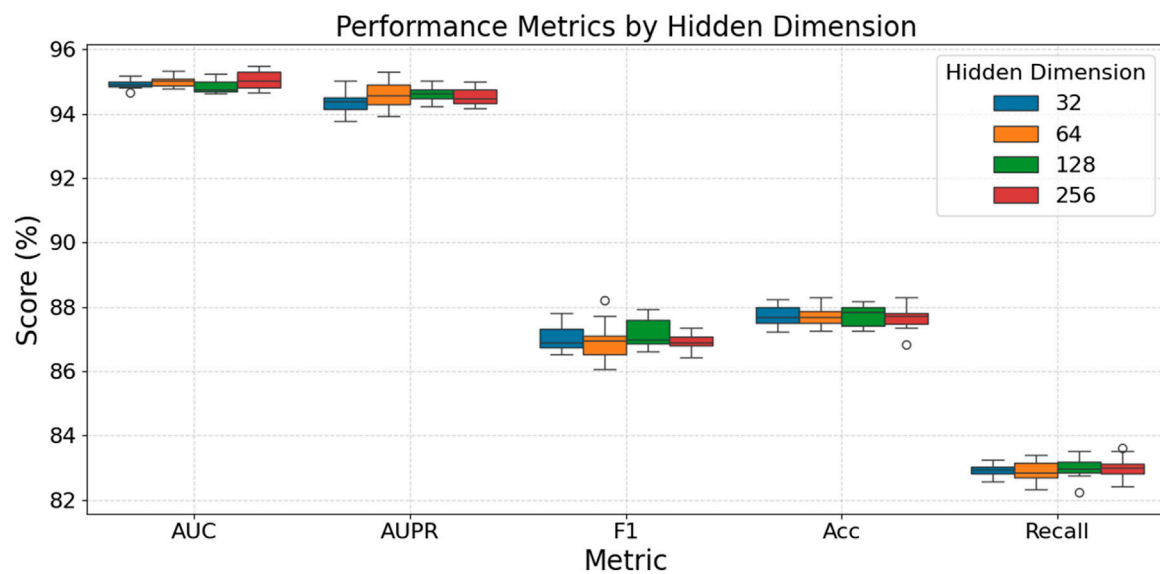


FIGURE 3  
Model performance at different hidden layer dimensions.

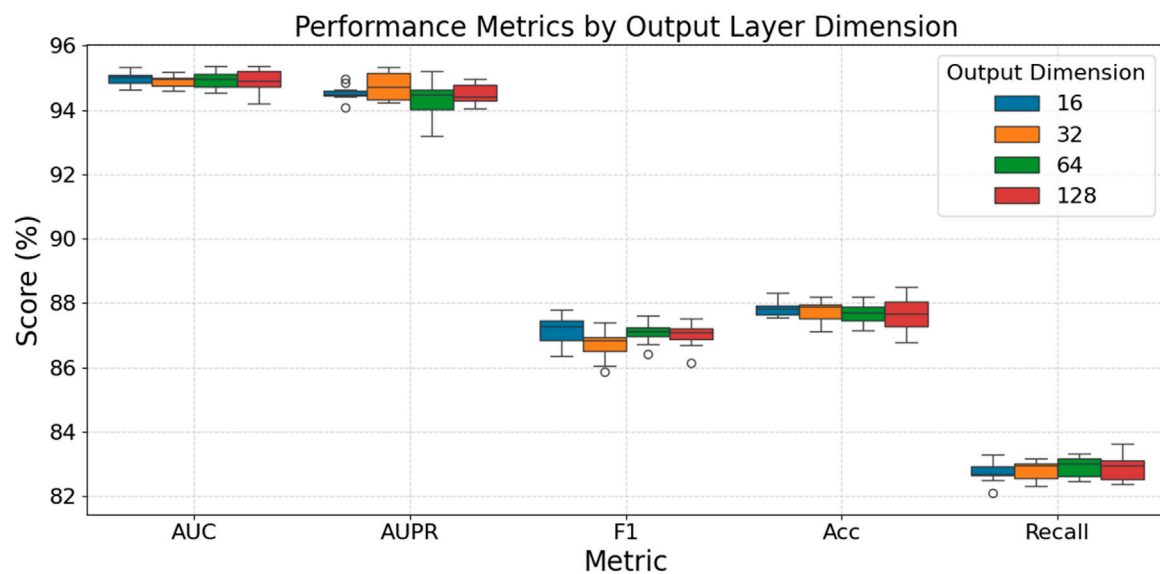


FIGURE 4  
Model performance at different output layer dimensions.

score and Accuracy, a hidden dimension of 128 appears to provide the optimal performance, with median F1-scores around 87.5% and median Accuracy scores approaching 87.8%–87.9%. Regarding Recall, while there is a slight upward trend in median scores with increasing dimensionality (around 83.1% for the 256 dimension), this is accompanied by greater volatility and the appearance of some outliers, particularly at the 256 dimension. In summary, while larger hidden dimensions do not significantly alter AUC or AUPR performance, a mid-range dimension like 128 seems to offer the best balance for F1-score and Accuracy, whereas dimensions 64 and

128 also provide good and stable performance for Recall, avoiding the increased variance potentially introduced by the largest dimension.

Figure 4 evaluates how varying the output layer dimension impacts model performance. For AUC, Accuracy, and Recall, the choice of output dimension within this range appears to have a minimal impact, with all tested dimensions yielding very similar and stable high scores; for instance, AUC scores consistently hover around 94.8%–95%, and Accuracy scores are tightly grouped around 87.6%–87.8%. In the case of AUPR, dimensions 32 and



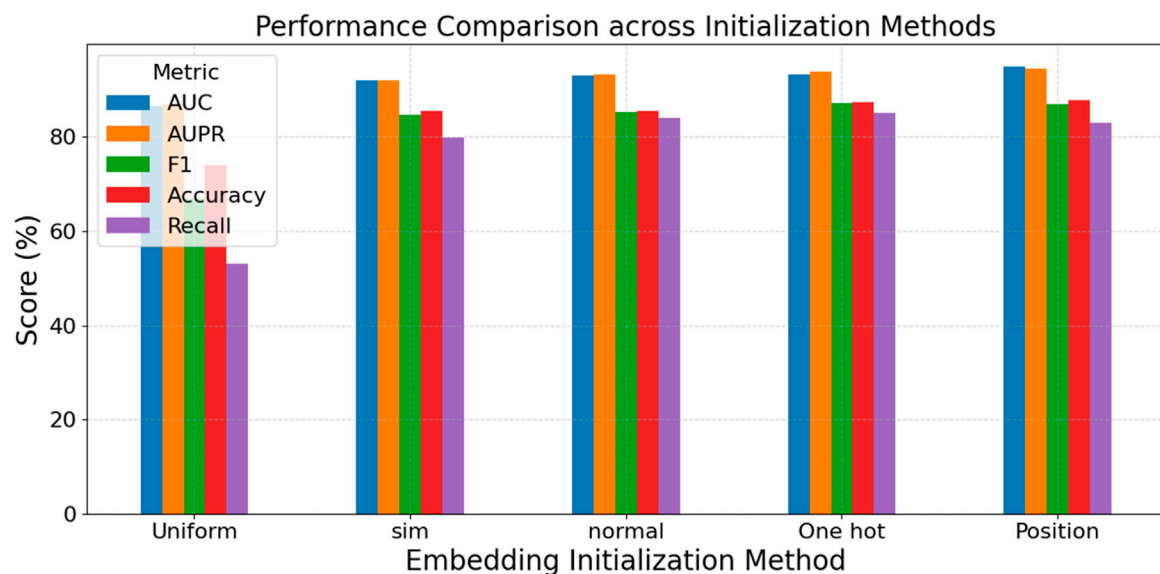


FIGURE 5  
Performance comparison across initialization methods.

128 demonstrate slightly more stable and marginally higher median performance (around 94.5%–94.8%) compared to dimensions 16 and 64. For the F1-score, an output dimension of 32 results in a slightly lower median and increased variability, while dimensions 16, 64, and 128 offer comparable and slightly better median performance (around 87.0%–87.2%). Overall, while most metrics show little sensitivity to the output layer dimension, dimensions such as 32 or 128 might provide minor advantages in AUPR stability, and dimensions other than 32 could be preferable for F1-score, but generally, the model's performance is robust across these output dimension variations.

In the CDA modeling task, circRNA and drug features play a crucial role in the GNN's ability to understand semantic and graph structures. To evaluate the impact of different feature representations on model performance, we use various node feature construction methods in our experiments. These include Uniform and Normal distributions for random initialization, which are suitable for scenarios with no prior knowledge and mainly test the model's structural learning ability. One-hot encoding assigns unique identities to nodes, distinguishing them without providing semantic information. Sim features, based on the circRNA-drug similarity matrix, leverage domain-specific prior knowledge to enhance the model's perception of actual associations. CircRNA similarity is quantified using two metrics: sequence-based similarity and GIP (Gaussian interaction profile) kernel similarity (Kipf and Welling, 2016a; Veličković et al., 2017). Similarly, drug similarity is evaluated from two perspectives: structural similarity and GIP kernel similarity. The comprehensive similarity matrices for circRNAs and drugs are obtained by merging the respective similarity matrices. Position features use the adjacency matrix to represent nodes' positions in the graph structure, emphasizing the role of structural connectivity relationships.

The experimental results in Figure 5 demonstrate the crucial role of node features in GNN models. Position features (using the adjacency matrix) achieve the best overall performance, with the

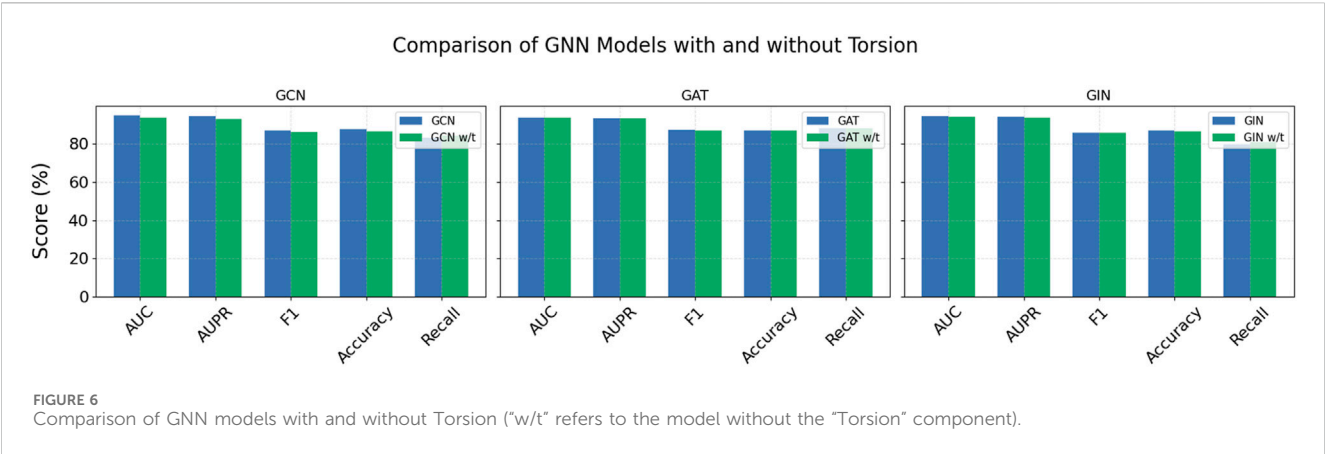
highest AUC (94.91%), AUPR (94.48%), and Accuracy (87.70%), indicating that structural information significantly impacts model accuracy and stability. One-hot encoding and Normal initialization also show strong performance, particularly in F1 and Recall. The F1 score of One-hot reaches 87.17%, suggesting that clearly distinguishing node identities aids the model's discriminative ability. In contrast, Uniform initialization performs the weakest across all metrics, especially with a Recall of only 53.09%, indicating a lack of effective structural or semantic information that limits the model's learning ability. Sim features, based on biomolecular similarity, perform better than random initialization but not as well as Position features, indicating that domain knowledge can enhance the model's representation and generalization capabilities.

In summary, incorporating structural information (e.g., adjacency matrix) and semantic similarity (e.g., Sim features) is vital for improving graph neural network performance. Simple random initialization struggles to achieve excellent performance without prior knowledge. These findings provide important guidance for future model design and feature construction.

## Ablation experiments

The resolving torsion module, a core innovative component of our model, significantly enhances the representation of graph structures. It captures higher-order node relationships by modeling local graph regions with simplicial complexes. We further integrate topological torsion values into edge weights, strengthening the model's sensitivity to local geometric and topological features. This method boosts the graph neural network's modeling dimensions and improves its expressiveness and discriminative power in complex graph data.

To validate this model's contribution, we designed an ablation study. By removing torsion weights or replacing them with simple



edge similarity metrics, we observed performance changes across multiple evaluation metrics. Results in Figure 6 show that the torsion module substantially improves overall model performance. For example, in GCN models, AUC increases from 93.64% to 94.91%, AUPR from 93.11% to 94.48%, and F1 from 86.29% to 87.06%. These improvements highlight how torsion-derived topological information enhances the model's discriminative and generalization abilities. Although GAT and GIN models show minor changes in Recall after incorporating torsion, their F1 scores and accuracy improve, indicating torsion weights play a key role in capturing local structural differences and enhancing feature expression.

In summary, the torsion module is crucial for improving the model's ability to understand and utilize higher-order graph relationships, making it a key factor in enhancing model performance.

Case analysis

To validate the constructed model's real - world biomedical applicability, a case study was conducted. circRNAs were selected from databases, and their potential drug interactions were predicted using the trained model. This assessed the model's accuracy in identifying circRNA - drug associations and its utility in studying disease mechanisms and screening new drugs.

ALDH3A2 (aldehyde dehydrogenase 3 family member A2), a key human enzyme encoding fatty aldehyde dehydrogenase, is mainly found in peroxisomes and the endoplasmic reticulum. It oxidizes medium - and long - chain aliphatic aldehydes to fatty acids, playing a crucial role in lipid metabolism and cellular antioxidant responses (Rizzo and Carney, 2005). Highly expressed in liver, skin, and brain tissues, it is vital for skin barrier function and nervous system performance. ALDH3A2 loss - of - function or mutation can cause Sjögren - Larsson syndrome, an autosomal recessive disorder with symptoms like ichthyosis, mental retardation, and spastic paralysis. Using the trained model, drugs related to ALDH3A2 were predicted, and the top 20 candidates with the highest model evaluation scores are listed in Table 4.

ANXA2 (Annexin A2), a calcium-dependent phospholipid-binding protein, is widely expressed in various human tissues, particularly in epithelial and vascular endothelial cells (Wang

TABLE 4 Top 20 predicted drugs with potential associations with ALDH3A2.

Drugs	Evidence	Drugs	Evidence
SNX-2112	Confirmed	YK 4-279	Confirmed
MP470	Confirmed	BX-795	Confirmed
CP724714	Confirmed	GSK690693	Confirmed
Axitinib	Confirmed	CH5424802	Confirmed
CCT018159	Confirmed	Navitoclax	Confirmed
AMG-706	Confirmed	TPCA-1	Confirmed
KIN001-055	Confirmed	BHG712	Confirmed
Cetuximab	Confirmed	XMD8-92	Confirmed
UNC0638	Confirmed	Temozolomide	Unconfirmed
Olaparib	Confirmed	Embelin	Unconfirmed

et al., 2019). Located in the cytoplasm, inner cell membrane, and extracellular environment, it participates in processes like membrane transport, cytoskeletal remodeling, and fibrinolytic system regulation. ANXA2 also acts as a co-receptor for tissue fibrinogen activator and fibrinogen, influencing fibrinogenesis and thrombolysis. Its abnormal expression is linked to diseases, especially cancers, where it is associated with tumor invasion, metastasis, and drug resistance, making it a potential tumor marker and therapeutic target. Additionally, ANXA2 plays roles in viral infections, autoimmune diseases, and inflammatory responses, highlighting its broad biomedical and clinical significance. Using the trained model, we predicted drugs related to ANXA2, and the top 20 candidates based on model evaluation scores are listed in Table 5.

ASPH (Aspartate Beta-Hydroxylase) is a protein encoded by the human ASPH gene and belongs to the 2 - ketoglutarate - dependent dioxygenase family (Hou et al., 2018). It plays a role in various biological processes, particularly in cell migration and cancer. Using the trained model, we predicted drugs related to ASPH. The top 20 candidates with the highest model evaluation scores are listed in Table 6.

CALD1 (Caldesmon 1), encoded by the CALD1 gene, is a human-encoded regulatory protein related to actin and is

TABLE 5 Top 20 predicted drugs with potential associations with ANXA2.

Drugs	Evidence	Drugs	Evidence
Camptothecin	Confirmed	Midostaurin	Confirmed
Bortezomib	Confirmed	CP724714	Confirmed
Tipifarnib	Confirmed	CHIR-99021	Confirmed
Salubrinol	Confirmed	YM201636	Confirmed
Tamoxifen	Confirmed	Afatinib	Confirmed
Vinblastine	Confirmed	GSK429286A	Confirmed
KIN001-055	Confirmed	GSK690693	Confirmed
Embelin	Confirmed	Ruxolitinib	Confirmed
Docetaxel	Confirmed	Cetuximab	Confirmed
Nilotinib	Confirmed	Gemcitabine	Unconfirmed

TABLE 6 Top 20 predicted drugs with potential associations with ASPH.

Drugs	Evidence	Drugs	Evidence
Elesclomol	Confirmed	Zibotentan	Confirmed
Temsirolimus	Confirmed	Gemcitabine	Confirmed
KIN001-055	Confirmed	TAK-715	Confirmed
Parthenolide	Confirmed	Embelin	Confirmed
CCT018159	Confirmed	Vinblastine	Confirmed
Temozolomide	Confirmed	Tipifarnib	Confirmed
Tamoxifen	Confirmed	Dasatinib	Confirmed
Salubrinol	Confirmed	CEP-701	Confirmed
AS605240	Confirmed	Midostaurin	Confirmed
BIRB 0796	Confirmed	Bicalutamide	Unconfirmed

expressed in various tissues, particularly functional in smooth muscle and non-myocytes (Cheng et al., 2021). It plays key roles in cytoskeletal remodeling, cell migration, and muscle contraction, making it crucial in multiple biological processes, especially in cancer. We used our trained model to predict drugs related to CALD1, and the top 20 drug candidates based on model evaluation scores are listed in Table 7.

The experimental results regarding the four specific circRNAs showed that most drug associations predicted by the model were validated in existing databases, confirming the model’s accuracy and reliability in identifying potential circRNA-drug relationships. These findings not only enhance the model’s applicability to real-world data but also validate its potential biomedical applications, particularly in circRNA function research and drug discovery.

## Conclusion

To address the computational demands of the proposed geometric modeling approach, an empirical evaluation was performed to quantify

TABLE 7 Top 20 predicted drugs with potential associations with CALD1.

Drugs	Evidence	Drugs	Evidence
AS605240	Confirmed	Nilotinib	Confirmed
Camptothecin	Confirmed	KIN001-055	Confirmed
Bortezomib	Confirmed	Docetaxel	Confirmed
CP724714	Confirmed	selumetinib	Confirmed
S-Trityl-L-cysteine	Confirmed	Salubrinol	Confirmed
17-AAG	Confirmed	Epothilone B	Confirmed
Tipifarnib	Confirmed	Axitinib	Confirmed
CI-1040	Confirmed	Temsirolimus	Confirmed
PD-0325901	Confirmed	PD-173074	Unconfirmed
Ruxolitinib	Confirmed	Embelin	Unconfirmed

the time necessitated by local structural extraction and the computation of torsion-based geometric features. The results indicate that the initial phase, involving the extraction of 1-hop, 2-hop, and 3-hop subgraph structures for each circRNA-drug pair, is relatively efficient, requiring approximately 0.326 s per pair. However, the subsequent phase of geometric feature computation, which includes the construction of simplicial complexes and the calculation of torsion features, demands a significantly greater processing time, averaging approximately 905 s. This substantial increase in computational duration underscores the considerable overhead associated with the incorporation of higher-order geometric information. While this computational cost is considered acceptable for datasets of moderate size, particularly given the concomitant improvements in predictive performance, it may pose scalability challenges when the model is applied to large-scale or real-time biomedical knowledge graphs. Consequently, future research will focus on exploring more efficient approximation techniques or parallelization strategies to reduce the time required for geometric encoding.

circRNA plays a crucial regulatory role in the cellular microenvironment, with its aberrant expression frequently linked to disease progression, including cancer. Additionally, circRNA is widely involved in regulating cellular drug resistance, making the study of circRNA–drug associations (CDAs) vital for therapeutic target discovery and drug development. While existing graph neural network-based methods have advanced CDA prediction, they often overlook important geometric information inherent in the local structures around known CDAs, which limits their predictive power. In this study, we propose a novel graph representation learning framework that integrates geometric information through a resolved torsion technique. By identifying each CDA’s local simplicial complex and computing its torsion value as an edge weight during message propagation, our model effectively captures higher-order geometric interactions within the circRNA–drug network. This innovation enables enhanced perception of subtle structural perturbations, thereby improving the accuracy of unknown CDA identification beyond existing methods. Despite these advancements, the model has limitations. For example, the current approach may face challenges in scalability when applied to extremely large or heterogeneous datasets, and the reliance on local simplicial

complexes assumes sufficient data quality and completeness. Future work will focus on extending the model's applicability to broader datasets, integrating multi-omics information, and exploring adaptive geometric encoding strategies to further enhance predictive performance. Overall, our findings demonstrate that incorporating torsion-based geometric information is a promising direction for advancing circRNA–drug association prediction, with potential to accelerate therapeutic target discovery and drug development.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

## Author contributions

ZL: Methodology, Writing – original draft. MQ: Formal Analysis, Methodology, Writing – review and editing. JH: Formal Analysis, Writing – review and editing. WZ: Supervision, Writing – review and editing. XT: Supervision, Writing – review and editing. YC: Methodology, Supervision, Writing – review and editing.

## Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The work was supported

by the Major Scientific and Technological Innovation Project of Wenzhou (No. ZG2024013) and Doctoral Initiation Projects of Shenzhen Institute of Information Technology (No. SZIIT2025KJ058).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. We disclose the limited application of Generative AI exclusively for linguistic refinement, with strict exclusion from research methodology, analytical logic, or experimental design.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Cai, L., He, Y., Fu, X., Zhuo, L., Zou, Q., and Yao, X. (2024). AEGNN-M: A 3D graph-spatial Co-Representation model for molecular property prediction. *IEEE J. Biomed. Health. Inform.* 29, 1726–1734. doi:10.1109/JBHI.2024.3368608
- Chen, Y., Wang, J., Zou, Q., Niu, M., Ding, Y., Song, J., et al. (2024). DrugDAGT: a dual-attention graph transformer with contrastive learning improves drug-drug interaction prediction. *BMC Biol.* 22 (1), 233. doi:10.1186/s12915-024-02030-9
- Cheng, Q., Tang, A., Wang, Z., Fang, N., Zhang, Z., Zhang, L., et al. (2021). CALD1 modulates gliomas progression via facilitating tumor angiogenesis. *Cancers* 13 (11), 2705. doi:10.3390/cancers13112705
- Deng, L., Liu, Z., Qian, Y., and Zhang, J. (2022). Predicting circRNA-drug sensitivity associations via graph attention auto-encoder. *BMC Bioinformatics* 23 (1), 160. doi:10.1186/s12859-022-04694-y
- Ding, C., Yi, X., Chen, X., Wu, Z., You, H., Chen, X., et al. (2021a). Warburg effect-promoted exosomal circ\_0072083 releasing up-regulates NANO expression through multiple pathways and enhances temozolomide resistance in glioma. *J. Exp. Clin. Cancer. Res.* 40 (1), 164. doi:10.1186/s13046-021-01942-6
- Ding, Y., Lei, X., Liao, B., and Wu, F.-X. (2021b). Predicting miRNA-disease associations based on multi-view variational graph auto-encoder with matrix factorization. *IEEE J. Biomed. Health. Inform.* 26 (1), 446–457. doi:10.1109/JBHI.2021.3088342
- Feng, L., Fu, X., Du, Z., Guo, Y., Zhuo, L., Yang, Y., et al. (2025). MultiCTox: empowering accurate cardiotoxicity prediction through adaptive multimodal learning. *J. Chem. Inf. Model.* 65 (7), 3517–3528. doi:10.1021/acs.jcim.5c00022
- Fu, X., Du, Z., Chen, Y., Chen, H., Zhuo, L., Lu, A., et al. (2025). DrugKANs: a paradigm to enhance drug-target interaction prediction with KANs. *IEEE J. Biomed. Health. Inform.* 1–12. doi:10.1109/JBHI.2025.3566931
- Geng, X., Jia, Y., Zhang, Y., Shi, L., Li, Q., Zang, A., et al. (2020). Circular RNA: biogenesis, degradation, functions and potential roles in mediating resistance to anticarcinogens. *Epigenomics* 12 (3), 267–283. doi:10.2217/epi-2019-0295
- Hou, G., Xu, B., Bi, Y., Wu, C., Ru, B., Sun, B., et al. (2018). Recent advances in research on aspartate  $\beta$ -hydroxylase (ASPH) in pancreatic cancer: a brief update. *Bosnian J. Basic Med. Sci.* 18 (4), 297–304. doi:10.17305/bjbm.2018.3539
- Huang, Z., Chen, K., Xiao, X., Fan, Z., Zhang, Y., and Deng, L. (2025). DeepHeteroCDA: circRNA-drug sensitivity associations prediction via multi-scale heterogeneous network and graph attention mechanism. *Brief Bioinform.* 26 (2), bbaf159. doi:10.1093/bib/bbaf159
- Kipf, T. N., and Welling, M. (2016a). Semi-supervised classification with graph convolutional networks. *arXiv Preprint. arXiv:1609.02907*. Available online at: <https://openreview.net/forum?id=SJU4ayYgl>.
- Kipf, T. N., and Welling, M. (2016b). Variational graph auto-encoders. *arXiv Preprint. arXiv:1611.07308*.
- Liu, Y., Dong, Y., Zhao, L., Su, L., and Luo, J. (2018). Circular RNA-MTO1 suppresses breast cancer cell viability and reverses monastrol resistance through regulating the TRAF4/Eg5 axis. *Int. J. Oncol.* 53 (4), 1752–1762. doi:10.3892/ijo.2018.4485
- Long, Y., Wu, M., Kwok, C. K., Luo, J., and Li, X. (2020). Predicting human microbe–drug associations via graph convolutional network with conditional random field. *Bioinformatics* 36 (19), 4918–4927. doi:10.1093/bioinformatics/btaa598
- Luo, Y., and Deng, L. (2024). DPMGCDA: deciphering circRNA-Drug sensitivity associations with dual perspective learning and path-masked graph autoencoder. *J. Chem. Inf. Model.* 64 (10), 4359–4372. doi:10.1021/acs.jcim.4c00573
- Ma, X., Fu, X., Wang, T., Zhuo, L., and Zou, Q. (2024). GraphADT: empowering interpretable predictions of acute dermal toxicity with multi-view graph pooling and structure remapping. *Bioinformatics* 40 (7), btae438. doi:10.1093/bioinformatics/btae438
- Reidemeister, K. (1935). Homotopieringe und linsenräume. *Abh. aus dem Math. Semin. Univ. Hambg.* 11, 102–109. doi:10.1007/BF02940717
- Rizzo, W. B., and Carney, G. (2005). Sjögren-larsson syndrome: diversity of mutations and polymorphisms in the fatty aldehyde dehydrogenase gene (ALDH3A2). *Hum. Mutat.* 26 (1), 1–10. doi:10.1002/humu.20181

- Shen, C., Liu, X., Luo, J., and Xia, K. (2025). Torsion graph neural networks. *IEEE Trans. Pattern. Anal. Mach. Intell.* 47, 2946–2956. doi:10.1109/TPAMI.2025.3528449
- Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., and Bengio, Y. (2017). Graph attention networks. *arXiv Preprint*. arXiv:1710.10903.
- Vromman, M., Vandesompele, J., and Volders, P.-J. (2021). Closing the circle: current state and perspectives of circular RNA databases. *Brief Bioinform.* 22 (1), 288–297. doi:10.1093/bib/bbz175
- Wang, R., Wang, T., Zhuo, L., Wei, J., Fu, X., Zou, Q., et al. (2024d). Diff-AMP: tailored designed antimicrobial peptide framework with all-in-one generation, identification, prediction and optimization. *Brief Bioinform.* 25 (2), bbac078. doi:10.1093/bib/bbae078
- Wang, T., Du, Z., Zhuo, L., Fu, X., Zou, Q., and Yao, X. M. C. B. (2024c). MultiCBlo: enhancing predictions of compound-induced inhibition of cardiac ion channels with advanced multimodal learning. *Int. J. Biol. Macromol.* 276, 133825. doi:10.1016/j.ijbiomac.2024.133825
- Wang, T., Li, Z., Zhuo, L., Chen, Y., Fu, X., and Zou, Q. (2024a). MS-BACL: enhancing metabolic stability prediction through bond graph augmentation and contrastive learning. *Brief Bioinform.* 25 (3), bbae127. doi:10.1093/bib/bbae127
- Wang, T., Wang, Z., Niu, R., and Wang, L. (2019). Crucial role of Anxa2 in cancer progression: highlights on its novel regulatory mechanism. *Cancer Biol. Med.* 16 (4), 671–687. doi:10.20892/j.issn.2095-3941.2019.0228
- Wang, T., Zhuo, L., Chen, Y., Fu, X., Zeng, X., and Zou, Q. (2024b). ECD-CDGI: an efficient energy-constrained diffusion model for cancer driver gene identification. *PLoS Comput. Biol.* 20 (8), e1012400. doi:10.1371/journal.pcbi.1012400
- Wang, X., Wang, H., Jiang, H., Qiao, L., and Guo, C. (2023). Circular RNAcirc\_0076305 promotes cisplatin (DDP) resistance of non-small cell lung cancer cells by regulating ABCC1 through miR-186-5p. *Cancer. Biother. Radiopharm.* 38 (5), 293–304. doi:10.1089/cbr.2020.4153
- Wei, J., Wang, L., Zhou, Z., Zhuo, L., Zeng, X., Fu, X., et al. (2024c). BloodPatrol: revolutionizing blood cancer diagnosis - advanced real-time detection leveraging deep learning and cloud technologies. *IEEE J. Biomed. Health. Inform.*, 1–11. doi:10.1109/JBHI.2024.3496294
- Wei, J., Zhu, Y., Zhuo, L., Liu, Y., Fu, X., and Li, F. (2024b). Efficient deep model ensemble framework for drug-target interaction prediction. *J. Phys. Chem. Lett.* 15 (30), 7681–7693. doi:10.1021/acs.jpclett.4c01509
- Wei, J., Zhuo, L., Fu, X., Zeng, X., Wang, L., Zou, Q., et al. (2024a). DrugReAlign: a multisource prompt framework for drug repurposing based on large language models. *BMC Biol.* 22 (1), 226. doi:10.1186/s12915-024-02028-3
- Xie, F., Xiao, X., Tao, D., Huang, C., Wang, L., Liu, F., et al. (2020). circNR3C1 suppresses bladder cancer progression through acting as an endogenous blocker of BRD4/C-myc complex. *Mol. Ther. Nucleic. Acids.* 22, 510–519. doi:10.1016/j.omtn.2020.09.016
- Xu, K., Hu, W., Leskovec, J., and Jegelka, S. (2018). How powerful are graph neural networks? *arXiv Prepr.* arXiv:1810.00826.
- Xu, T., Wang, M., Jiang, L., Ma, L., Wan, L., Chen, Q., et al. (2020). CircRNAs in anticancer drug resistance: recent advances and future potential. *Mol. Cancer* 19, 127–20. doi:10.1186/s12943-020-01240-3
- Yang, B., and Chen, H. (2023). Predicting circRNA-drug sensitivity associations by learning multimodal networks using graph auto-encoders and attention mechanism. *Brief Bioinform.* 24 (1), bbac596. doi:10.1093/bib/bbac596
- Yu, Z., Huang, F., Zhao, X., Xiao, W., and Zhang, W. (2021). Predicting drug–disease associations through layer attention graph convolutional network. *Brief Bioinform.* 22 (4), bbaa243. doi:10.1093/bib/bbaa243
- Zang, J., Lu, D., and Xu, A. (2020). The interaction of circRNAs and RNA binding proteins: an important part of circRNA maintenance and function. *J. Neurosci. Res.* 98 (1), 87–97. doi:10.1002/jnr.24356
- Zhang, X., Wang, H., Du, Z., Zhuo, L., Fu, X., Cao, D., et al. (2024). CardiOT: towards interpretable drug cardiotoxicity prediction using optimal transport and Kolmogorov-arnold networks. *IEEE J. Biomed. Health. Inform.* 29, 1759–1770. doi:10.1109/JBHI.2024.3510297
- Zhou, Z., Liao, Q., Wei, J., Zhuo, L., Wu, X., Fu, X., et al. (2024c). Revisiting drug–protein interaction prediction: a novel global–local perspective. *Bioinformatics* 40 (5), btae271. doi:10.1093/bioinformatics/btae271
- Zhou, Z., Zhuo, L., Fu, X., Lv, J., Zou, Q., and Qi, R. (2024b). Joint masking and self-supervised strategies for inferring small molecule–miRNA associations. *Mol. Ther. Nucleic. Acids.* 35 (1), 102103. doi:10.1016/j.omtn.2023.102103
- Zhou, Z., Zhuo, L., Fu, X., and Zou, Q. (2024a). Joint deep autoencoder and subgraph augmentation for inferring microbial responses to drugs. *Brief Bioinform.* 25 (1), bbad483. doi:10.1093/bib/bbad483