

Multiview Joint Learning-Based Method for Identifying Small-Molecule-Associated MiRNAs by Integrating Pharmacological, Genomics, and Network Knowledge

Cong Shen, Jiawei Luo,* Zihan Lai, and Pingjian Ding



Cite This: *J. Chem. Inf. Model.* 2020, 60, 4085–4097



Read Online

ACCESS |



Metrics & More

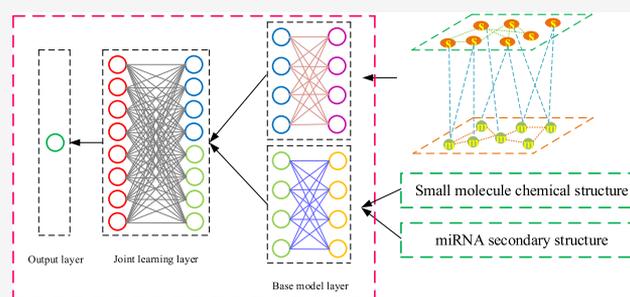


Article Recommendations



Supporting Information

ABSTRACT: The emergence of a large amount of pharmacological, genomic, and network knowledge data provides new challenges and opportunities for drug discovery and development. Identification of real small-molecule drug (SM)–miRNA associations is not only important in the development of effective drug repositioning but also crucial in providing a better understanding of the mechanisms by which small-molecule drugs achieve the purpose of treating diseases by regulating miRNA expression. However, challenges remain in accurately determining potential associations between small molecules and miRNAs using information from multiomics data. In this study, we adopted a novel framework called SMAJL to improve the prediction of small molecule–miRNA associations with joint learning. First, we use enhancing matrix completions to obtain the network knowledge of small molecule–miRNA associations. Then, we extract the information of small-molecule fingerprints and miRNA sequences into feature vectors to obtain small-molecule structure and miRNA sequence information. Finally, we incorporate small-molecule structure information, miRNA sequence data, and heterogeneous network knowledge into a joint learning model based on a Restricted Boltzmann Machine (RBM) to predict association scores. To validate the effectiveness of our method, the SMAJL model is compared with four state-of-the-art methods in terms of 5-fold cross-validation. The results demonstrate that the AUC and AUPRC of the SMAJL are obviously superior to those of other comparison methods. The SMAJL model also achieved great results in terms of robustness and case studies, further demonstrating its strong predictive power.



1. INTRODUCTION

The efforts of characterizing and predicting drug–target associations have been inspired by drug polypharmacology.¹ Where available, the targeted drugs may continue to inform clinical trials, drug discovery, and efforts to overcome drug resistance through deeply understanding the mechanistic action. Thus, it is important to guide future drug innovation and development by maintaining an up-to-date and accurate map of FDA-approved drugs and their efficacy targets.² Previous studies have shown that the precursors and mature miRNAs could be targeted by small molecular drugs (abbreviation small molecule or SM);^{3–5} the secondary structures of miRNA, which mainly include stem loops and bulges, are targets to which small molecules can be applied.⁶ More critically, cumulative studies have identified that many diseases are associated with abnormal expression of miRNAs.^{7–9} Therefore, it is feasible to discover miRNAs that exhibit specific associations with small molecules for disease therapeutic purposes.¹⁰

In the past decade, wet lab experiments have been used to verify associations between small molecules and miRNAs. However, these methods are time consuming, costly, and

laborious.¹¹ Thus, computational means are needed to complement wet lab experiments. Although wet lab experiments still remain the gold standard, one can broaden the research scope of wet lab experiments using computational techniques.^{12–14} Computational prediction of small molecule–miRNA associations has become a critical step in drug discovery or repositioning, aiming to identify potential novel small molecular drugs or new miRNAs for existing small molecular drugs. Computational methods can efficiently discover potential small-molecule-associated miRNA candidates to guide experimental validation and thus significantly reduce the cost and time required for drug repositioning or discovery.¹⁵

Recently, many computational approaches have been developed to reveal novel or potential small-molecule-

Received: March 10, 2020

Published: July 10, 2020



associated miRNAs. One major category of computational methods is feature extraction-based approaches for calculating the association score between small molecules and miRNAs. A key assumption of these methods is that similar small molecules may share similar miRNAs and vice versa. Based on this assumption, the computational process of small molecule–miRNA association prediction can be formulated as a binary classification task, which aims to verify whether an association between a small molecule and miRNA is present. For example, Jamal et al.¹⁶ used Naïve Bayes and Random Forest to generate computational models of biological activity of small molecules, providing the first comprehensive analysis of the miRNA–small molecule modulator prediction model. Xie et al.¹⁷ proposed a text mining method called EmDL, which determined known associations between small molecular drugs and miRNAs from the literature. This approach first extracted features by calculating the distance between entities on the sentences containing them. Then, a support vector machine was used to calculate prediction scores. Finally, using the EmDL method to obtain known small molecule–miRNA associations, the author constructed a database named MTD.¹⁸ In addition to extracting features from the literature, extracting features from the network is also a viable method. Wang et al. developed a calculation framework named RFSMMA, which was based on Random Forest to predict small molecule–miRNA associations. The similarity between small molecules and miRNAs was extracted as features to implement machine learning.¹⁹

Graph mining approaches provide a multiview perspective and diverse information for verifying potential small molecule–miRNA associations. Several studies indicated that graph-mining-based computational methods could achieve good prediction performance. For instance, Jiang et al. identified the biological interactions between small molecules and miRNAs in 23 different cancers using a novel high-throughput method and systematically analyzed the properties of small molecule–miRNA associations by constructing the association network for each cancer.¹⁰ Lv et al.²⁰ performed the Random Walk with Restart algorithm to predict the associations between small molecules and miRNAs. MiRNAs can be used as small molecular targets because small molecules can affect miRNA expression. Based on this notion, Meng et al.²¹ computed the similarity of the transcriptional response to verify small molecule–miRNA associations by analyzing 39 miRNA-perturbed gene expression profiles. To increase feasibility and provide high efficiency and accuracy for comprehensively evaluating associations between small molecules and miRNAs on a large scale, Li et al.²² proposed a network-based miRNA pharmacogenomic framework. Qu et al.²³ proposed a HeteSim-based inference model for small molecule–miRNA association prediction. Guan et al.²⁴ presented a graphlet interaction-based inference method for small molecule-associated miRNAs prediction. Wang et al.²⁵ identified small molecule–miRNA associations based on cross-layer dependency inference on multilayered networks. To fully consider the various information from heterogeneous networks, Qu et al.²⁶ developed a triple-layer heterogeneous network-based model by integrating small molecule–miRNA associations and miRNA–disease associations to uncover potential small molecule–miRNA associations. Zhao et al.²⁷ presented a framework called SNMFSMMA to predict potential small molecule–microRNA association using symmetric non-negative matrix factorization and Kronecker

regularized least squares. In addition, some ideas from the perspective of network sparsity²⁸ or the statistical viewpoint²⁹ might be useful to improve the prediction performance of small molecule–miRNA associations.

Whether using feature extraction approaches to predict small molecules and miRNA associations or mining potential relationships from a graph mining perspective, the results are analyzed from a single perspective. A computational model based on a single perspective may provide prediction results with poor robustness and low accuracy. Accumulated studies^{30–32} have shown that the incorporation of multisource (multimodal) information is broadly embraced in fields of drug discovery, chemogenomics, etc. For example, Luo et al.¹⁵ proposed a calculation-based methodology named DTINet to predict potential drug–target associations based on a constructed heterogeneous network with integrated diverse information, which includes four types of drug-related nodes and six types of edges. Zeng et al.³³ used an arbitrary-order proximity embedded deep forest approach (AOPEDF) to predict novel DTIs. In AOPEDF, the authors constructed a large heterogeneous network incorporating 15 networks covering network profiles and phenotypic, genomic, and chemical information on drugs, diseases, and proteins/targets. If we integrate multisource information in the prediction of the association between small molecules and miRNA, it may improve the prediction performance. Thus, we integrate structural information of small molecules, sequence information of miRNAs, and association network information to predict small molecule-associated miRNAs, which potentially provides a new perspective. Previous studies found that matrix factorization achieved good performance in extracting network information for association prediction. For example, Pliakos et al.³⁴ used neighborhood regularized logistic matrix factorization methods to predict DTI on the reconstructed network and achieved good performance. Ammad-Ud-Din et al.³⁵ utilized multitask matrix factorization to predict drug responses and verified the powerful prediction ability of the model through experiments. Meanwhile, the Restricted Boltzmann Machine (RBM) model has broad application scenarios in both feature enhancement and model prediction. For example, the dgMDL model³⁶ used a deep belief network composed of three layers of restricted Boltzmann machines to filter genes and disease features, and then predicted their associations. The IMTRBM method³⁷ enhanced the accuracy of previous miRNA–target association prediction results based on Restricted Boltzmann Machines. On the basis of the above studies, we plan to first extract network features in a manner similar to matrix completion, and then filter the features of multisource information through restricted Boltzmann machines. Finally, we predict the association of small molecules with miRNAs based on the filtered features.

In this study, we present a multiview joint learning-based computational framework named SMAJL to predict small molecule–miRNA associations by integrating small molecule structural information, miRNA sequence data, and heterogeneous network knowledge. The small molecular structural features and miRNA sequences are obtained through RDKit (<http://www.rdkit.org/>) and Pre-in-One.³⁸ Meanwhile, the network feature is extracted by enhancing matrix completion. In our method, to verify the necessity of each stage, we compare the performance of SMAJL with its variant model. To better identify the superiority of our method, the AUC and AUPRC of SMAJL and other state-of-the-art methods are

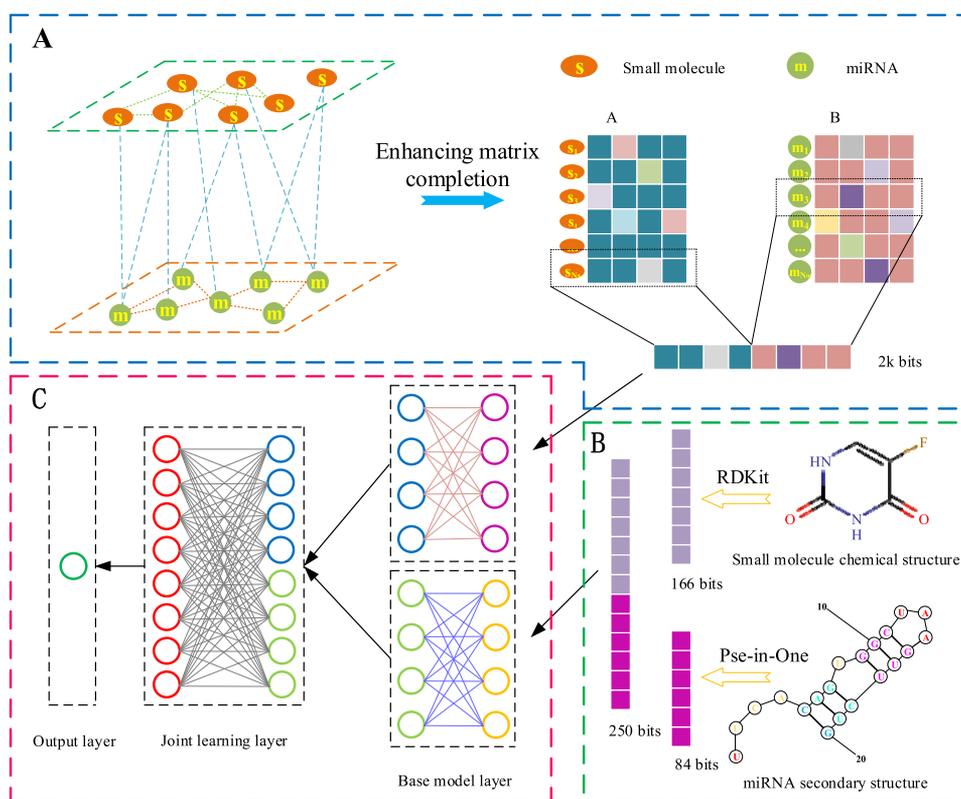


Figure 1. Overview of our proposed SMAJL framework. (A) SMAJL first integrates similarity network and association network to construct a heterogeneous network, and applies enhancing matrix completion to obtain a low-dimensional vector representation of the features describing the topological properties for each node. (B) SMAJL then uses RDKit to extract features of the small-molecule chemical structure and utilizes Pse-in-One to obtain features of the miRNA secondary structure. After that, (C) SMAJL incorporates network features, small-molecule chemical structural features, and the miRNA secondary structure as the input of the multiview joint learning model to predict potential small molecule–miRNA associations.

compared to illustrate the good prediction performance of SMAJL. In addition, we remove some known small molecule–miRNA associations to predict potential small molecule-associated miRNAs to indicate the robustness of SMAJL. Finally, case studies are used to identify de novo small molecule–miRNA associations through relevant literature to further demonstrate the strong predictive power of the SMAJL model.

2. METHODS

2.1. Overview. In this section, we propose a framework named SMAJL to predict small molecule–miRNA associations. Key to our algorithm is the multiview joint learning-based method for feature aggregation, which can consider pharmacological, genomics, and network knowledge simultaneously. Figure 1 provides an overview of the developed SMAJL framework, which includes network feature extraction (Figure 1A), molecular structure feature extraction (Figure 1B), and a joint learning model (Figure 1C). The process of the SMAJL framework is described in detail below.

2.2. Network Representation Learning with Enhancing Matrix Completion. **2.2.1. Small-Molecule Clinical Similarity.** Anatomical Therapeutic Chemical (ATC) classification system codes are widely used to calculate the clinical similarities of small-molecule drug pairs.^{30,39,40} The Drug-Bank⁴¹ provides numerous ATC codes for small molecules used in this work. We use the ATC codes to define the l level similarity (S_s^l) of small molecules A and B as follows

$$S_s^l(A, B) = \frac{ATC_l(A) \cap ATC_l(B)}{ATC_l(A) \cup ATC_l(B)} \quad (1)$$

where ATC_l indicates all ATC codes at the l level. The clinical similarity of small molecules A and B should be defined by averaging all levels of similarity as noted below

$$S_s(A, B) = \frac{\sum_{l=1}^k S_s^l(A, B)}{k} \quad (2)$$

where k indicates the five levels of the ATC code. The first three layers of the ATC code represent anatomic classification, therapeutic classification, and pharmacological classification, which cover most of the information of the ATC code.⁴² Therefore, we use the first three layers of the ATC code to calculate the clinical similarity of small molecules, which is commonly used to facilitate the calculation of the similarity of small molecular drugs.⁴⁰

2.2.2. miRNA Functional Similarity. In this work, we use experimentally verified miRNA–gene interactions to calculate miRNA functional similarity based on the calculation steps of Xiao’s method.⁴³ First, we download the gene functional interaction network from HumanNet,⁴⁴ which provides associated log-likelihood scores (LLS) that represent the strength of the interaction between two genes. Next, the min–max normalization is used to normalize LLS to calculate gene similarity. Finally, miRNA functional similarity can be obtained by integrating miRNA–gene associations and gene similarities, which is calculated based on the BMA method⁴⁵ as follows

$$S_m(m_i, m_j) = \frac{\sum_{g \in G_i} S(g, G_j) + \sum_{g \in G_j} S(g, G_i)}{|G_i| + |G_j|} \quad (3)$$

where G_i and G_j indicate the gene sets associated with m_i and m_j , respectively, $|l|$ denotes the number of genes in sets, and $S(g, G)$ represents the similarity of gene g and gene set G .

2.2.3. Enhancing Matrix Completion. Matrix completion is an effective technique and has been widely used for data representation.^{46–48} This method aims to recover the missing values in the original matrix that has only a partially observed entry set. Matrix factorization is a common method used for matrix completion. Matrix factorization refers to using A and B to approximate the matrix Y . Then, the elements of AB^T can be used to estimate the value of the element in the corresponding invisible position in Y , and AB^T can be regarded as the factorization of Y . In this study, the small molecule–miRNA association matrix $Y \in \mathbb{R}^{N_s \times N_m}$ could be completed by two matrices $A \in \mathbb{R}^{N_s \times k}$ and $B \in \mathbb{R}^{N_m \times k}$ ($k \ll \min(N_s, N_m)$), and $Y \approx AB^T$. Thus, the process of small-molecule-associated miRNA identification is mathematically defined as the following objective function

$$\min_{A,B} \|Y - AB^T\|_F^2 + \delta(\|A\|_F^2 + \|B\|_F^2), \quad \text{st } A \geq 0, B \geq 0 \quad (4)$$

where $\|\cdot\|$ indicates the Frobenius norm and δ represents the contribution of the regularization term. To consider more biological information from heterogeneous networks, the similarities of small molecules and miRNAs are utilized as prior knowledge to constrain the objective function of matrix completion. We reconstruct the objective function as follows

$$\min_{A,B} \|Y - AB^T\|_F^2 + \alpha \text{Tr}(A^T L_s A) + \beta \text{Tr}(B^T L_m B) + \delta(\|A\|_F^2 + \|B\|_F^2), \quad \text{st } A \geq 0, B \geq 0 \quad (5)$$

where $\text{Tr}(\cdot)$ represents the trace of a matrix, $L_s = D_s - S_s$ and $L_m = D_m - S_m$ are the graph Laplacian matrices of S_s and S_m ,⁴⁹ respectively, S_s and S_m indicate the similarities of small molecules and miRNAs, D_s and D_m are diagonal matrices for which entries are column (or row) sums of S_s and S_m , respectively, and α and β are the regularization coefficients.

In this study, ATC codes are used to calculate the clinical similarities between two small molecules. As described in the section on the clinical similarity of small molecules, it is not difficult to obtain the ATC codes and compute the correct results. A large number of accumulated studies have demonstrated that the similarity information can be directly used as the neighbor regularization term of the main optimization model to improve performance in the matrix factorization model. For example, NRLMF⁵⁰ is a neighborhood regularized logistic matrix factorization model that predicts drug–target interactions; the entire model is optimized by directly using drug similarity and target protein similarity as neighbor regularization terms of logistic matrix factorization. GRGMF⁵¹ is used to identify potential links in biomedical bipartite networks by graph regularized generalized matrix factorization, which utilizes similarity information as Laplacian regularized terms to enforce nodes with high similarities to have similar representations in latent space. However, unlike these methods, the SMAJL model uses ATC codes to calculate the similarity of small molecules, but not all ATC codes of small molecules can be obtained. Thus, the

similarity between small molecules without ATC codes and other small molecules cannot be obtained using the calculation method of clinical similarity. Thus, the small-molecule clinical similarity matrix will be sparse and cannot truly represent the similarity relationship of small molecules. To solve this problem, we reconstruct the similarity matrix S_s^* to take the place of the original small-molecule similarity matrix S_s and define a new indicator matrix W . The objective function can be rewritten as follows

$$\min_{A,B} \|Y - AB^T\|_F^2 + \alpha \text{Tr}(A^T L_s A) + \beta \text{Tr}(B^T L_m B) + \gamma \|W \odot (S_s - S_s^*)\|_F^2 + \delta(\|A\|_F^2 + \|B\|_F^2 + \|S_s^*\|_F^2), \quad \text{st } A \geq 0, B \geq 0, S_s^* \geq 0 \quad (6)$$

where $S_s^* \in \mathbb{R}^{N_s \times N_s}$ represents the reconstructed small-molecule similarity matrix for which the initial value is a random value between 0 and 1 and it is the same size as matrix S_s . In addition, $W \in \mathbb{R}^{N_s \times N_s}$ represents the indicator matrix. If the i th small molecule and the j th small molecule both have ATC codes, then $W_{ij} = 1$, otherwise 0. Here, $\|W \odot (S_s - S_s^*)\|_F^2$ indicates the enhancing term of the matrix completion and γ represents the contribution of the enhancing term. Since the matrix S_s^* is used to represent more real small-molecule similarities, the graph Laplacian matrices L_s should be redefined as $L_s = D_s - S_s^*$, where D_s is a diagonal matrix for which entries are the column (or row) sums of S_s^* .

To solve the optimization problem in eq 6, we decompose the optimization problem into several subproblems. The whole optimization process is similar to that in refs 52–54, which requires updating some variables iteratively while other variables are fixed. Thus, each subproblem converges to its local minima. Then, we could obtain iteratively updated formulas as follows

$$a_{ik} \leftarrow a_{ik} \frac{YB + \alpha S_s^* A}{AB^T B + \alpha D_s A + \delta A} \quad (7)$$

$$b_{jk} \leftarrow b_{jk} \frac{Y^T A + \beta S_m B}{BA^T A + \beta D_m B + \delta B} \quad (8)$$

$$s_{ii'} \leftarrow s_{ii'} \frac{\alpha AA^T + 2\gamma W \odot S_s}{\alpha \zeta(A) + 2\gamma W \odot S_s^* + 2\delta S_s^*} \quad (9)$$

where

$$\zeta(A) = \begin{bmatrix} \sum_{i=1}^m a_{1i}^2 & \sum_{i=1}^m a_{1i}^2 & \dots & \sum_{i=1}^m a_{1i}^2 \\ \sum_{i=1}^m a_{2i}^2 & \sum_{i=1}^m a_{2i}^2 & \dots & \sum_{i=1}^m a_{2i}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^m a_{ni}^2 & \sum_{i=1}^m a_{ni}^2 & \dots & \sum_{i=1}^m a_{ni}^2 \end{bmatrix} \quad (10)$$

In eqs 7–9, a_{ik} , b_{jk} , and $s_{ii'}$ represent the element values of matrices A , B , and S_s^* , respectively. The non-negative matrices A and B are updated through eqs 7–9 until convergence. Then, we obtain the small-molecule representation data A and miRNA representation data B . Finally, matrices A and B are fed

into the RBM model as the small molecules and miRNAs network feature vectors for the next prediction. The detailed process of objective function optimization is shown in the [Supporting Information](#).

2.3. Feature Engineering of Small-Molecule Structure and miRNA Sequence. The chemical structure of small molecules contains chemical properties; thus, the transformation of the chemical structure of small molecules into feature vectors and its application to the machine learning model will provide a new perspective for the prediction of small-molecule-associated miRNAs. First, we download the chemical structure information (SMILES format) from the DrugBank⁴¹ and calculate the MACCS fingerprints of each small molecule using RDKit (<http://www.rdkit.org/>). Then, each small molecule is expressed as a vector of length 166 for the next computation.

Just as there are chemical properties in the chemical structure of small molecules, there are biological properties in miRNA sequences. We download miRNA sequence data from miRBase.⁵⁵ The miRNA sequence is composed of different arrangements of four bases: A, U, G, and C. We calculate the frequency of each base that appears in the sequence and obtain a vector of length 4. The biological properties of an miRNA sequence are determined by the different arrangements of four bases. We also consider the combination of the two bases and three bases and calculate the occurrence probability of each combination. Then, a vector of length 16 is obtained by combining the two bases with each other, and a vector of length 64 could be obtained by combining the three bases with each other. We connect the three vectors obtained, and each miRNA sequence can be expressed as a vector of length 84. The process of miRNA feature extraction can be realized by Pse-in-One.³⁸

2.4. Base Model RBM. The Restricted Boltzmann Machine (RBM)^{37,56} is a random neural network (i.e., when the neural node of the network is activated, there will be random behavior). It consists of a layer of visible layers and a layer of hidden layers. Neurons in the same layer are independent of each other, and neurons are interconnected (bidirectional) between different network layers. When the network is trained and used, the information flows in both directions, and the weights in both directions are the same. However, the biases are different (the number of offset values is the same as the number of neurons), and the structure of the restricted Boltzmann machine is presented in [Figure 2](#).

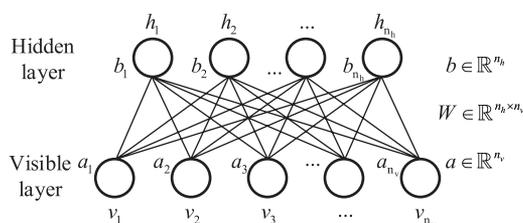


Figure 2. Schematic of a Restricted Boltzmann Machine (RBM).

In the RBM, assume that $v = (v_1, v_2, \dots, v_{n_v})$ and $h = (h_1, h_2, \dots, h_{n_h})$ represent the state vectors of the visible layer and hidden layer, $a \in \mathbb{R}^{n_v}$ and $b \in \mathbb{R}^{n_h}$ are biases, and $W \in \mathbb{R}^{n_h \times n_v}$ indicates the weight matrix. RBM is an energy-based probability distribution model. The current energy function of the RBM can be expressed as follows

$$E(v, h) = -a^T v - b^T h - h^T W v \quad (11)$$

With the energy function, the joint probability distribution of (v, h) is described as follows

$$P(v, h) = \frac{1}{Z} e^{-E(v, h)} \quad (12)$$

where $Z = \sum_{v, h} e^{-E(v, h)}$ is known as the partition function. Due to difficulties in processing the partition function Z , we use the maximum likelihood gradient to approximate it. First, the conditional distribution from the joint distribution is derived as follows

$$P(h|v) = \frac{P(h, v)}{P(v)} = \frac{1}{P(v)} \frac{1}{Z} \exp\{a^T v + b^T h + h^T W v\} \quad (13)$$

It is easy to obtain the probability that the j th node in the hidden layer is 1 on the basis of the given visible layer v

$$\begin{aligned} P(h_j = 1|v) &= \frac{P(h_j = 1|v)}{P(h_j = 1|v) + P(h_j = 0|v)} \\ &= \frac{1}{1 + \exp\{-(b_j + W_{:,j} v)\}} = \text{sigmoid}(b_j + W_{:,j} v) \end{aligned} \quad (14)$$

This method is equivalent to using the sigmoid activation function, so we can write a complete conditional distribution about the hidden layer

$$P(v_j = 1|h) = \text{sigmoid}(a_j + W_{:,j} h_j) \quad (15)$$

These equations are important for iterative updates between visible and hidden layers when training RBM models, but it is difficult to calculate the expectations over all possible configurations of input data. To solve this problem, Hinton et al.,⁵⁷ Tieleman et al.,⁵⁸ and Cho et al.⁵⁹ developed several sampling techniques to estimate the expectations with a fixed number of samples. For simplicity, we choose contrast divergence (CD), and method details are referenced in ref 57.

2.5. Joint Learning Model. Joint learning was originally proposed for natural language-processing tasks (for example, entity recognition and entity standardization joint learning as well as word segmentation and part of speech tagging joint learning).^{60,61} In this study, joint learning is used to learn cross-modality features from original features, including network features and molecular structure features. As shown in [Figure 1C](#), the process of joint learning includes three steps, and each step uses the RBM model. The RBM is used to first enhance the network information representation, the representation of the features connected by small-molecule structural features and miRNA sequence features and then to strengthen the connection of the first two outputs. In fact, RBM is an unsupervised method for enhancing feature representation, and the training process is the same in every phase. Thus, the parameters are trained based on the objective function

$$\min \|R - \hat{R}\|_F^2 \quad (16)$$

where \hat{R} represents the feature vector output by the RBM and R is the input feature vector. First, small-molecule and miRNA network features can be obtained by enhancing matrix completion, and the connected small-molecule and miRNA features are used as input to the RBM model to obtain an

Table 1. Details of Multi-Type Data

data type	database	description
small molecule–miRNA associations	SM2miR ⁶²	228 small molecules, 794 miRNAs, and 3743 associations
MiRNA–gene associations	miRTarbase ⁶³	794 miRNA and 196565 associations
small-molecule ATC codes	DrugBank ⁴¹	112 small molecules and their ATC codes for calculating clinical similarity
small-molecule SMILES	DrugBank ⁴¹	228 small molecules and their SMILES for extracting features.
miRNA sequence	miRBase ⁵⁵	794 miRNA and their sequence for extracting features.

output of equal length. Then, the output and input of RBM are used to train the parameters in the objective function (eq 16). After the RBM is stabilized, the output result will be used for the input of the joint learning layer in the next stage. Second, we obtain the chemical structure features of small molecules and the secondary structure features of miRNA through RDKit (<http://www.rdkit.org/>) and Pre-in-One.³⁸ Similar to the first step, the connected small-molecule and miRNA features are used as RBM inputs to obtain an output of equal length. Similarly, the output and input of RBM are used to train the parameters in the objective function (eq 16). After the RBM is stabilized, the output result will be used for the input of the joint learning layer in the next stage. Finally, after training two RBM submodels, we obtain the feature vectors h_{net} and v_{seq} which contain network knowledge and structural information (including small-molecule structure and miRNA structure). Thus, we connect these features as follows

$$w_{\text{net_seq}} = \text{concat}((1 - \mu)h_{\text{net}}, \mu v_{\text{seq}}) \quad (17)$$

where concat is a concatenation function and μ is used to control the significance of network knowledge and structural information. The connected vector $w_{\text{net_seq}}$ is used as the input to train the joint RBM. To avoid the difference in the scale of the element values of different vectors, we used the two vectors after the max–min normalization as input in the next concatenation. The computational method of max–min normalization is reported below

$$V_{\text{max-min}} = \frac{V - V_{\text{min}}}{V_{\text{max}} - V_{\text{min}}} \quad (18)$$

where V is the vector that needs to be regularized, and V_{max} and V_{min} represent the maximum and minimum values of the element in vector V , respectively.

The joint learning model is trained in an unsupervised manner, and the resulting model can be further analyzed using many approaches. In this study, we add an output layer with a logistic regression model to identify the relationship score of each small molecule–miRNA pair being associated using the multiview joint learning-based method learned by the multi-RBM.

3. RESULTS

3.1. Data Sources. The small molecule–miRNA association data were downloaded from SM2miR.⁶² After removing the same associations and merging the same mature miRNAs, we obtained 228 small molecules, 794 miRNAs, and 3743 associations. The miRTarbase⁶³ provides data, including 196 565 miRNA–gene associations for 794 miRNAs, to calculate miRNA functional similarity. To compute small-molecule similarity, we downloaded ATC codes from DrugBank.⁴¹ Among the 228 small molecules, 112 small molecules have ATC codes, which are subsequently used to calculate the clinical similarity of small molecules. We

downloaded chemical structure information (SMILES format) from DrugBank⁴¹ and miRNA sequence data from miRBase.⁵⁵ Given that each small molecule has chemical structure information (SMILES format) and each miRNA has sequence data, there are 228 small molecule features and 794 sequence features. Table 1 shows the details of multi-type data.

3.2. Experimental Setup. We evaluated the performance of our model for small molecule–miRNA association prediction with 5-fold cross-validation and repeated it 10 times. In this study, network representation vectors are obtained by enhancing matrix completion based on the known small molecule–miRNA association data set. Thus, for each 5-fold cross-validation, the network representation vector of small molecules and miRNAs needs to be rederived by enhancing matrix completion to avoid the information of the test set in the training set. We use the area under receiver operating characteristic (ROC) curve (AUC) to evaluate the performance of the SMAJL model. The ROC curve can be plotted by the true-positive rate (TPR) and false-positive rate (FPR), and we calculate TPR and FPR using $\text{TPR} = [\text{TP}/(\text{TP} + \text{FN})]$ and $\text{FPR} = [\text{FP}/(\text{FP} + \text{TN})]$ at different thresholds, respectively. The recall is calculated by $\text{recall} = [\text{TP}/(\text{TP} + \text{FN})]$ to evaluate algorithm performance, where TP, FP, TN, and FN represent the true-positive, false-positive, true-negative, and false-negative rates, respectively. The AUPRC is another important evaluation metric, which is the area under the precision/recall curve. We calculate precision and recall through $\text{precision} = [\text{TP}/(\text{TP} + \text{FP})]$ and $\text{recall} = [\text{TP}/(\text{TP} + \text{FN})]$ at different thresholds. The output of the SMAJL model is the prediction score of associations between small molecules and miRNAs using the logistic regression model. To balance the number of positive samples and negative samples, we set the same negative sample size as the positive sample to train the logistic regression model, and the negative sample is randomly chosen from all uncertain small molecule–miRNA association data sets.

3.3. Overall Performance. Figure 3 shows the average AUC obtained with various variant models from different views of the overall framework. The model_Net represents the vector that connects small-molecule network representation and miRNA network representation as the input of the logistic regression model directly. The model_NetRBM indicates that the network representation vector is passed as an input to the logistic regression model after passing through the RBM model. The model_Seq represents the vector that connects small-molecule structure features and miRNA sequence features as the input of the logistic regression model directly. The model_SeqRBM indicates that the structure representation vector is passed as an input to the logistic regression model after passing through the RBM model. The model_SeqNet represents the vector that connects the output vector of the model_NetRBM and model_SeqRBM as the input of the logistic regression model directly. As shown in Figure 3, after passing through the RBM model, the models

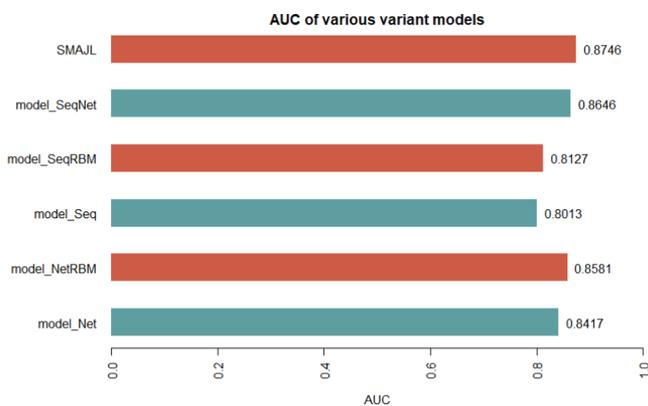


Figure 3. AUC of SMAJL and different variant models.

achieved better performance than other models. After incorporating network representation, the models obtained better small molecule and miRNA structure feature results compared with the single model. This experimental result potentially has two explanations. First, the RBM model is used to filter the original feature, which could remove various noise data of feature vectors to improve prediction performance. Second, the joint learning model incorporates multisource information, including heterogeneous network information, small-molecule structural information, and miRNA sequence data, to improve prediction results. Thus, the multiview joint learning-based method is a meaningful method to identify small-molecule-associated miRNAs.

3.4. Selection of Hyperparameters. In the SMAJL model, several hyperparameters are crucial for the overall performance of the prediction. For enhanced matrix completion, the hyperparameters α , β , γ , and δ represent the contribution of the different regularization terms. We used the grid search method to search α , β , γ , and δ of the SMAJL model from $\{0.01, 0.1, 1, 10\}$ to discover the best parameter combination. The representation vector dimension k of the small molecule and miRNA in the enhancing matrix completion is another important parameter; we select this parameter from $\{80, 90, 100, 110, 120, 130, 140\}$. The parameter μ represents the significance of network knowledge

and structural information, which is selected from $\{0.001, 0.005, 0.01, 0.05, 0.1, 0.5\}$. To avoid the influence of small molecular structure information and miRNA sequence information, we used the RBM model to filter the representation obtained by the enhancement matrix and then added the representation to the logistic model to select the optimal parameters. For hyperparameters α , β , γ , and δ , as shown in Tables S1–S4, we obtain the best performance when the combination of parameters is $\alpha = 10$, $\beta = 0.01$, $\gamma = 0.1$, and $\delta = 0.1$. For hyperparameters k and μ , Figure S1 shows that the SMAJL model achieves optimal results when $k = 120$, and Figure S2 indicates that better performance is obtained based on $\mu = 0.01$.

3.5. Comparison with State-Of-the-Art Models. We compare the performance of SMAJL with six other methods [i.e., NRLMF,⁵⁰ SMiR_NBI,²² SMANMF,⁵² FSMMA,¹⁹ HSSMMA,²³ RWR²⁰] in the task of identifying small molecule–miRNA associations. The detailed introduction and parameter settings of the comparison method are reported in the Supporting Information. The SMAJL model and the comparison methods use the same data in this study.

We use 5-fold cross-validation to evaluate the performance of SMAJL with the other six methods. All known small molecule–miRNA associations are randomly divided into five parts with the same size. Each part separately serves as the test set, and the remaining parts serve as the training set. The experimental results are shown in Figures 4 and 5. The ROC in Figure 4a indicates that the SMAJL method obtained an AUC value of 0.8746. For comparison, NRLMF, SMiR_NBI, SMANMF, RFSMMA, HSSMMA, and RWR achieved AUC values of 0.8552, 0.8378, 0.8326, 0.8128, 0.7703, and 0.7727, respectively. Figure 4b shows the precision/recall curve of the seven models, and the AUPRCs of seven models are 0.0498, 0.0419, 0.0341, 0.0319, 0.0155, 0.0230, and 0.0176, respectively. This result indicates that joint learning is a feasible model to reveal small molecule–miRNA associations. Furthermore, the AUCs and AUPRCs of SMMART and other methods with different runs were compared using paired *t*-tests via 5-fold cross-validation. As shown in Table 2, the *p*-values were less than 0.05, suggesting that the differences between AUCs and AUPRCs were statistically significant. However, it is

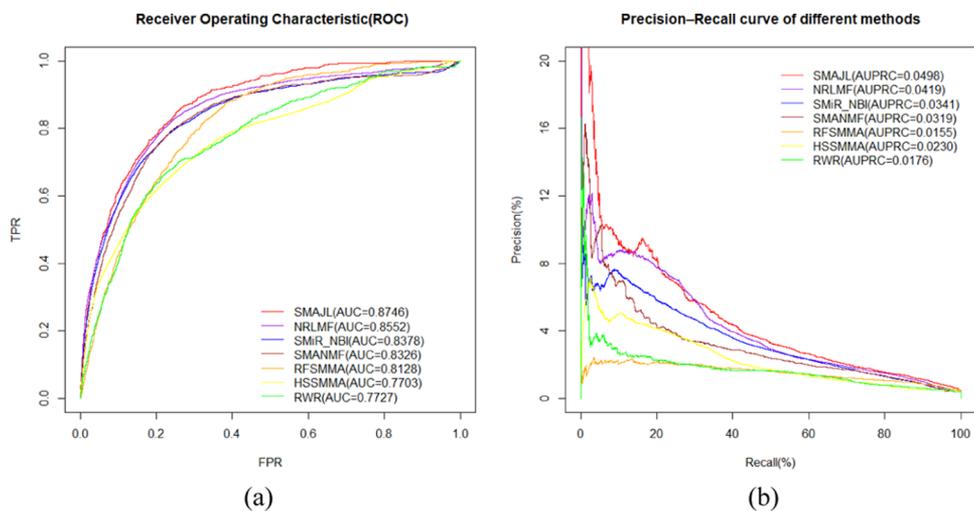


Figure 4. Comparison of SMAJL with other models. (a) ROC curve and AUC value of the SMAJL model and comparison methods. (b) Precision/recall curve and AUPRC value of the SMAJL model and comparison methods.

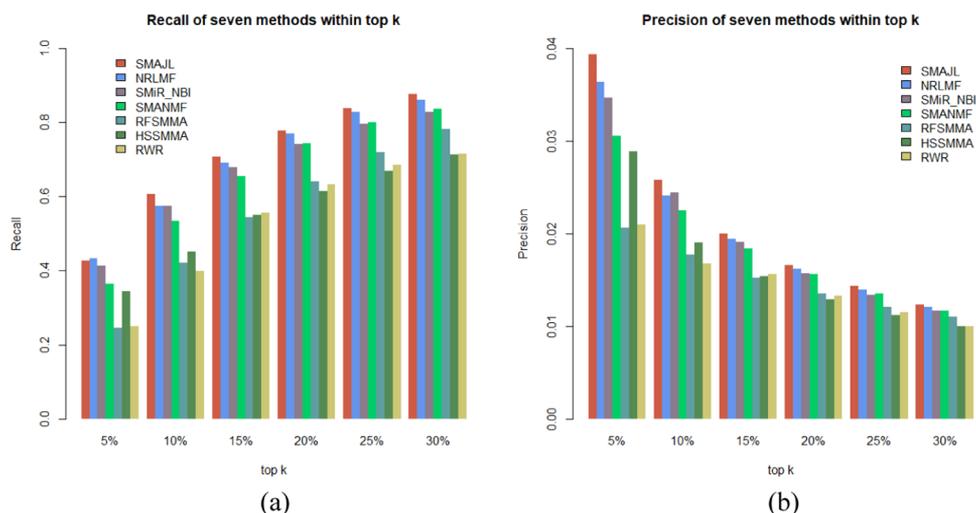


Figure 5. Comparison of SMAJL with other models. (a) Recall of the SMAJL model and comparison methods. (b) Precision of the SMAJL model and comparison methods.

Table 2. *P*-values Obtained Through Paired *t*-Test of the AUCs and AUPRCs of SMAJL and Other Compared Methods for 10 Runs

	<i>p</i> -value					
	NRLMF	SMANMF	SMiR_NBI	RFSMMA	HSSMMA	RWR
AUCs	2.74×10^{-4}	2.252×10^{-7}	2.51×10^{-8}	4.129×10^{-9}	3.604×10^{-11}	4.694×10^{-12}
AUPRCs	1.318×10^{-4}	2.305×10^{-5}	1.248×10^{-5}	1.47×10^{-8}	4.819×10^{-8}	1.002×10^{-5}

Table 3. AUC of SMAJL and Comparison Methods for the 13 Small Molecules

small molecule name	AUC						
	SMAJL	NRLMF	SMiR_NBI	SMANMF	RFSMMA	HSSMMA	RWR
glucose	0.7094	0.7050	0.6281	0.6651	0.6260	0.6585	0.7027
perfluorooctane sulfonate	0.6825	0.6717	0.6001	0.5950	0.6818	0.5833	0.6490
reversine	0.6981	0.6335	0.6078	0.5901	0.5399	0.4848	0.6327
formaldehyde	0.8647	0.8535	0.8036	0.8178	0.7295	0.7551	0.8238
17 β -estradiol	0.8916	0.8474	0.8341	0.8266	0.7691	0.6921	0.8284
dexamethasone	0.8936	0.8418	0.7970	0.8107	0.7020	0.7198	0.8025
diethylstilbestrol	0.7905	0.6778	0.6643	0.6796	0.6098	0.6410	0.6740
trichostatin A	0.8013	0.7962	0.7138	0.7772	0.6351	0.7610	0.7905
gemcitabine	0.8717	0.7922	0.7371	0.7693	0.6426	0.5715	0.7369
isoproterenol	0.8601	0.8514	0.8297	0.7992	0.7037	0.7572	0.8083
5-aza-2'-deoxycytidine	0.9310	0.9274	0.9278	0.9083	0.7525	0.7506	0.9119
cisplatin	0.7485	0.6752	0.6543	0.6537	0.6302	0.6541	0.7303
comfrey	0.9383	0.9097	0.8622	0.8466	0.8467	0.6577	0.8410

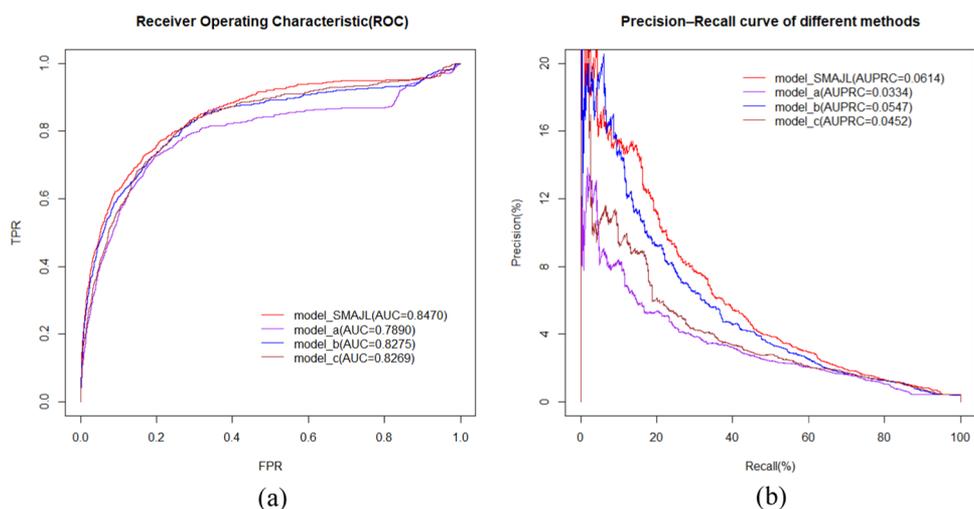
important to guide wet experiments through the top-ranked small-molecule-associated miRNAs obtained based on the computational model. Thus, we used the recall and precision to verify the performance of SMAJL and other methods within the top *k*. Figure 5a,b show the recall and precision obtained by applying the 5-fold cross-validation, respectively. From the top 5% to the top 30%, the SMAJL model achieved better recall and precision than other comparison methods. To illustrate the predictive power of the SMAJL model on a single small molecule, 5-fold cross-validation is used to obtain small molecules. There are 28 small molecules with greater than 40 small-molecule-associated miRNAs. We choose 40 to prevent sparseness. As shown in Table 3, we select the 13 small molecules most relevant to human life and health. The remaining small molecules are listed in Table S5. These results confirm that the joint learning-based SMAJL model is an

effective prediction model for discovering small molecule–miRNA associations. The superiority of the performance of the SMAJL model may be the incorporation of network knowledge, small-molecule structural information, and miRNA sequence information. In addition, enhancement matrix completion considers the sparsity of small-molecule clinical similarity, which may be another reason for the excellent performance of the SMAJL model.

3.6. Data Perturbation Analysis. To analyze the performance of the SMAJL model with different ratios of positive and negative samples, we removed 5% known associations between small molecules and miRNAs randomly. The hyperparameters of the SMAJL model were changeless, and we still used the 5-fold cross-validation to identify the performance. Table 4 shows that the SMAJL model achieves greater accuracy (AUC = 0.8622 and AUPRC = 0.0426) than

Table 4. Comparison of SMAJL with Other Methods with Various Evaluation Metrics After Removing 5% of Known Associations

		methods						
		SMAJL	NRLMF	SMiR_NBI	SMANMF	RFSMMA	HSSMMA	RWR
AUC		0.8622	0.8533	0.8357	0.8334	0.8014	0.7653	0.7490
AUPRC		0.0426	0.0378	0.0323	0.0305	0.0139	0.0211	0.0148
recall	5%	0.4191	0.4191	0.4191	0.3615	0.2489	0.3390	0.2406
	10%	0.5823	0.5710	0.5682	0.5432	0.4135	0.4388	0.3730
	15%	0.7004	0.6709	0.6976	0.6501	0.5316	0.5457	0.4947
	20%	0.7609	0.7496	0.7511	0.7401	0.6259	0.6118	0.5922
	25%	0.8200	0.8059	0.7947	0.7969	0.7103	0.6582	0.6444
	30%	0.8650	0.8509	0.8411	0.8385	0.7820	0.7103	0.6858
precision	5%	0.0341	0.0334	0.0320	0.0288	0.0199	0.0277	0.0199
	10%	0.0232	0.0228	0.0226	0.0217	0.0165	0.0185	0.0163
	15%	0.0183	0.0178	0.0178	0.0173	0.0141	0.0144	0.0143
	20%	0.0151	0.0150	0.0147	0.0148	0.0125	0.0124	0.0119
	25%	0.0129	0.0129	0.0127	0.0127	0.0113	0.0109	0.0103
	30%	0.0113	0.0113	0.0111	0.0112	0.0104	0.0096	0.0092

**Figure 6.** Comparison of model_SMAJL with other models. (a) Recall of model_SMAJL and comparison methods. (b) Precision of model_SMAJL and comparison methods.

the other six methods: NRLMF (AUC = 0.8533 and AUPRC = 0.0378), SMiR_NBI (AUC = 0.8357 and AUPRC = 0.0323), SMANMF (AUC = 0.8334 and AUPRC = 0.0305), RFSMMA (AUC = 0.8014 and AUPRC = 0.0139), HSSMMA (AUC = 0.7653 and AUPRC = 0.0211), and RWR (AUC = 0.749 and AUPRC = 0.0148). For recall and precision of various methods, the SMAJL model also obtained good performance within the top 5% or other top performers in Table 4. The results of various evaluation metrics demonstrate that the SMAJL model exhibits better performance than other methods for different ratios of positive and negative sample data. There are two main reasons why SMAJL achieves better performance. First, the combination of multisource information is beneficial to improve model stability. Second, enhancing matrix completion for data sparseness is beneficial to increase model stability.

3.7. Ablation Study. In the Anatomical Therapeutic Chemical (ATC) classification system, the active substances are divided into different groups according to the organ or system on which they act and their therapeutic, pharmacological, and chemical properties. Drugs are classified into groups at five different levels. A significant amount of

therapeutic, pharmacological, and chemical information of a drug is contained in five different layers.⁴² Thus, clinical similarities of drug pairs derived from the drug ATC classification systems codes have been commonly used to predict new drug targets^{32,39} and drug synergy.^{40,64} To quantify the benefits of ATC codes in the SMAJL model, we set three variant models called model_a, model_b, and model_c. Among them, model_a is the model generated after removing clinically similar small molecules, model_b is the model generated after replacing clinical similarity with functional similarity, and model_c indicates that the sparse clinical similarity model is used directly without increasing the indicator matrix. We call the enhancement matrix completion of the SMAJL model in this article model_SMAJL. Here, $Y^* = AB^T$ is used to represent the small molecule–miRNA potential association score. To reduce the effect of small-molecule structures and miRNA sequences, the enhanced matrix completion-based model was employed to predict small-molecule-associated miRNAs to quantify the benefits of ATC codes. A detailed introduction of the three variant models is provided in the Supporting Information. Figure 6 shows the AUCs and AUPRCs of model_SMAJL and three variant

models. We found that model_SMAJL exhibited high accuracy (AUC = 0.847 and AUPRC = 0.0614) in 5-fold cross-validation, outperforming those of several variant methods: model_a (AUC = 0.789 and AUPRC = 0.0334), model_b (AUC = 0.8275 and AUPRC = 0.0847), and model_c (AUC = 0.8269 and AUPRC = 0.0452) (Figure 6a,b). Table 5 shows

Table 5. Recall and Precision of Model_SMAJL and Comparison Methods Within Top k

		methods of ablation study			
		model_SMAJL	model_a	model_b	model_c
recall	5%	0.4786	0.3904	0.4666	0.4011
	10%	0.6230	0.5401	0.6003	0.5615
	15%	0.6939	0.6497	0.6644	0.6751
	20%	0.7540	0.7246	0.7299	0.7326
	25%	0.7981	0.7594	0.7834	0.7901
precision	30%	0.8369	0.7955	0.8275	0.8329
	5%	0.0402	0.0328	0.0392	0.0337
	10%	0.0262	0.0227	0.0252	0.0236
	15%	0.0194	0.0182	0.0186	0.0189
	20%	0.0158	0.0152	0.0153	0.0154
	25%	0.0134	0.0128	0.0132	0.0133
	30%	0.0117	0.0111	0.0116	0.0117

the recall and precision obtained by applying the 5-fold cross-validation, respectively. Using the top 5% to top 30%, the model_SMAJL achieved better recall and precision than other comparison models. The overall performance of model_SMAJL is better than that of model_a, indicating that the ATC code improves model performance. In addition, model_SMAJL is better than model_b in various evaluation metrics, demonstrating that the clinical similarity of small molecules based on ATC codes provides more effective information to the model than the functional similarity of small molecules. Finally, the AUC value of model_SMAJL is increased compared with model_c, verifying that the indicator matrix W added in eq 6 alleviates the sparsity in the clinical similarity matrix.

3.8. Prediction Analysis of Gemcitabine and Dexamethasone. To further demonstrate the performance of SMAJL in predicting potential associations between small molecules and miRNAs, case studies are conducted using the small molecules gemcitabine and dexamethasone. Here, we make predictions using all of the known association data, and the unknown association information is used for the candidate validation set. For each of these two small molecules, we

collect the top 10 candidates by prioritizing the candidate small molecule–miRNA association prediction scores. We use the PMID number in PubMed to indicate that the literature can confirm the association between small molecules and miRNA. “Unconfirmed” indicates that the association has not been confirmed.

Gemcitabine is a pyrimidine antitumor drug with the same mechanism of action as cytarabine, and its main metabolite is incorporated into DNA in cells, mainly in the G1/S phase. Dexamethasone, a corticosteroid that prevents the release of substances in the body that cause inflammation, is used to treat different inflammatory conditions, such as allergic disorders and skin conditions. When studying the pharmacology of gemcitabine and dexamethasone, it is useful to assess the association of gemcitabine and dexamethasone with miRNA. Seven and five candidate miRNAs for gemcitabine and dexamethasone are identified and supported by direct evidence in Table 6, respectively. For example, studies^{65,66} indicate that miR-29a functions as a potent autophagy inhibitor and decreases cancer cell invasion by sensitizing them to gemcitabine, and the inhibition of miR-29b attenuates atrophy induced by dexamethasone treatment. All of the case studies indicate that SMAJL is indeed capable of predicting potential small-molecule-associated miRNAs.

4. CONCLUSIONS AND DISCUSSION

Accumulated studies have demonstrated that miRNAs play important roles in various human diseases and could be regarded as potential targets of small molecular drugs. Thus, it is meaningful for drug discovery to discover potential small-molecule-associated miRNAs. In this study, we used the multiview joint learning model (SMAJL) to address the problem of predicting the associations between small molecules and miRNAs through integrating pharmacological, genomics, and network knowledge. In SMAJL, an enhancing matrix completion method was used to obtain network knowledge, and two tools (RDKit and Pse-in-One) were used to extract pharmacological information of small molecules and genomics knowledge of miRNAs, respectively. Finally, we used a multiview joint learning-based method to predict the association score of small molecules and miRNAs.

Our experimental results showed that the SMAJL model outperformed state-of-the-art models in 5-fold cross-validation by directly modeling the incorporation prediction model. We also predicted the association of small molecules with miRNA by deleting 5% of known correlation data. The results showed that SMAJL achieves better results in data perturbation,

Table 6. Top 10 Potential miRNA Candidates Detected by SMAJL Based on Pubmed for the Three Selected Small Molecules

small molecule	no. of miRNAs confirmed	top 10 ranked predictions					
		rank	miRNAs	evidence	rank	miRNAs	evidence
gemcitabine	7	1	mir-29b	30 915 884	6	mir-34a	unconfirmed
		2	mir-125b	26 606 261	7	mir-30c	27 506 865
		3	mir-27b	2 5184 537	8	mir-25	24 040 438
		4	mir-200c	unconfirmed	9	mir-182	25 833 690
		5	mir-29a	27 626 694	10	mir-9	unconfirmed
dexamethasone	5	1	mir-21	31 412 983	6	mir-182	24 871 856
		2	mir-29b	28 541 289	7	mir-9	29 492 899
		3	mir-125b	unconfirmed	8	mir-199a	unconfirmed
		4	mir-34a	28 918 747	9	mir-10a	unconfirmed
		5	mir-30c	unconfirmed	10	mir-203	26 748 295

demonstrating that the SMAJL model has better robustness. We use an ablation study to verify the effect of clinical similarity based on ATC codes, and experimental results show that ATC codes significantly improve the SMAJL model. Finally, we validated our predictions from the published literature. Seven and five of the top 10 predicted miRNAs for gemcitabine and dexamethasone, respectively, were validated. The results demonstrated the feasibility of our method for predicting the potential association between small molecules and miRNAs. There are three main reasons why the SMAJL model can achieve good results. First, in the enhancement matrix completion section, we considered the sparsity of small-molecule clinical similarity and used enhancing terms to solve this problem. Second, the robustness of the model was enhanced by the fusion of the structural information of small molecules and the sequence information of miRNAs. Simultaneously, this method offered more useful information for target prediction of small molecules. Third, the network feature information and the structural feature information were collaboratively predicted using the RBM-based multiview joint learning model. The noise in the feature vector can be reduced to improve model accuracy.

In conclusion, the SMAJL model proposed here incorporates multisource information using multiview joint learning-based models for small molecule–miRNA prediction and proved its effectiveness. In future work, integrating more useful association information and relevant feature knowledge from other databases and literature may provide an expandable application perspective of our model. In addition, the predicted results may have better guiding value for wet experiments if a predictive model of the association between small molecules and miRNAs can be constructed for a specific cancer cell line.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.0c00244>.

Detail process of objective function optimization of enhancing matrix completion; Baseline methods; ablation study; AUC of different parameters combination when $\alpha = 0.01$ (Table S1); AUC of SMAJL sub-model with different subspace dimension k (Figure S1); AUC of SMAJL with different significance μ of network knowledge and structural information (Figure S2) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jiawei Luo – College of Computer Science and Electronic Engineering, Hunan University, Changsha 410083, China; Email: luojiawei@hnu.edu.cn

Authors

Cong Shen – College of Computer Science and Electronic Engineering, Hunan University, Changsha 410083, China; orcid.org/0000-0001-8505-6406

Zihan Lai – College of Computer Science and Electronic Engineering, Hunan University, Changsha 410083, China

Pingjian Ding – School of Computer Science, University of South China, Hengyang 421001, China; orcid.org/0000-0002-2613-2496

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jcim.0c00244>

Funding

This work has been supported by the National Natural Science Foundation of China (Grant no. 61873089, 61572180), Hunan Provincial Innovation Foundation for Postgraduate (Grant No. CX20200436) and Scientific Research Startup Foundation of University of South China (Grant no. 190XQD096)

Notes

The authors declare no competing financial interest.

<https://github.com/CS-BIO/SMAJL>.

■ REFERENCES

- (1) Keiser, M. J.; Setola, V.; Irwin, J. J.; Laggner, C.; Abbas, A. I.; Hufeisen, S. J.; Jensen, N. H.; Kuijer, M. B.; Matos, R. C.; Tran, T. B. Predicting new molecular targets for known drugs. *Nature* **2009**, *462*, 175–181.
- (2) Santos, R.; Ursu, O.; Gaulton, A.; Bento, A. P.; Donadi, R. S.; Bologa, C. G.; Karlsson, A.; Al-Lazikani, B.; Hersey, A.; Oprea, T. I. A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discovery* **2017**, *16*, 19–34.
- (3) Bose, D.; Jayaraj, G.; Suryawanshi, H.; Agarwala, P.; Pore, S. K.; Banerjee, R.; Maiti, S. The tuberculosis drug streptomycin as a potential cancer therapeutic: inhibition of miR-21 function by directly targeting its precursor. *Angew. Chem., Int. Ed.* **2012**, *51*, 1019–1023.
- (4) Srinivasan, S.; Selvan, S. T.; Archunan, G.; Gulyas, B.; Padmanabhan, P. MicroRNAs—the next generation therapeutic targets in human diseases. *Theranostics* **2013**, *3*, 930–942.
- (5) Hesse, M.; Arenz, C. miRNAs as Novel Therapeutic Targets and Diagnostic Biomarkers for Parkinson's Disease: A Patent Evaluation of Expert Opinion on Therapeutic Patents. WO Patent WO20140186502014; pp 1271–1276.
- (6) Zhang, S.; Chen, L.; Jung, E. J.; Calin, G. A. Targeting microRNAs with small molecules: from dream to reality. *Clin. Pharmacol. Ther.* **2010**, *87*, 754–758.
- (7) Farh, K. K.-H.; Grimson, A.; Jan, C.; Lewis, B. P.; Johnston, W. K.; Lim, L. P.; Burge, C. B.; Bartel, D. P. The widespread impact of mammalian microRNAs on mRNA repression and evolution. *Science* **2005**, *310*, 1817–1821.
- (8) Migliore, C.; Giordano, S. MiRNAs as new master players. *Cell Cycle* **2009**, *8*, 2185–2186.
- (9) Pan, C.; Luo, J.; Zhang, J. Computational identification of RNA-Seq based miRNA-mediated prognostic modules in cancer. *IEEE J. Biomed. Health Inf.* **2019**, *24*, 626–633.
- (10) Jiang, W.; Chen, X.; Liao, M.; Li, W.; Lian, B.; Wang, L.; Meng, F.; Liu, X.; Chen, X.; Jin, Y. Identification of links between small molecules and miRNAs in human cancers based on transcriptional responses. *Sci. Rep.* **2012**, *2*, No. 282.
- (11) Ding, P.; Ouyang, W.; Luo, J.; Kwok, C.-K. Heterogeneous information network and its application to human health and disease. *Briefings Bioinf.* **2019**, *21*, 1327–1346.
- (12) Mongia, A.; Jain, V.; Chouzenoux, E.; Majumdar, A. In *Deep Latent Factor Model for Predicting Drug Target Interactions*. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); IEEE, 2019; pp 1254–1258.
- (13) Pan, C.; Luo, J.; Zhang, J.; Li, X. BiModule: biclique modularity strategy for identifying transcription factor and microRNA co-regulatory modules. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2019**, *17*, 321–326.
- (14) Zeng, X.; Zhong, Y.; Lin, W.; Zou, Q. Predicting disease-associated circular RNAs using deep forests combined with positive-unlabeled learning methods. *Briefings Bioinf.* **2019**, *21*, 1425–1436.
- (15) Luo, Y.; Zhao, X.; Zhou, J.; Yang, J.; Zhang, Y.; Kuang, W.; Peng, J.; Chen, L.; Zeng, J. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. *Nat. Commun.* **2017**, *8*, No. 573.

- (16) Jamal, S.; Periwal, V.; Scaria, V. Computational analysis and predictive modeling of small molecule modulators of microRNA. *J. Cheminf.* **2012**, *4*, No. 16.
- (17) Xie, W.; Yan, H.; Zhao, X.-M. EmDL: Extracting miRNA-Drug Interactions from Literature. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2017**, *16*, 1722–1728.
- (18) Chen, X.; Xie, W.-B.; Xiao, P.-P.; Zhao, X.-M.; Yan, H. mTD: A database of microRNAs affecting therapeutic effects of drugs. *J. Genet. Genomics* **2017**, *44*, 269–271.
- (19) Wang, C.-C.; Chen, X.; Qu, J.; Sun, Y.-Z.; Li, J.-Q. modeling, RFSMMA: a new computational model to identify and prioritize potential small molecule-miRNA associations. *J. Chem. Inf.* **2019**, *59*, 1668–1679.
- (20) Lv, Y.; Wang, S.; Meng, F.; Yang, L.; Wang, Z.; Wang, J.; Chen, X.; Jiang, W.; Li, Y.; Li, X. Identifying novel associations between small molecules and miRNAs based on integrated molecular networks. *Bioinformatics* **2015**, *31*, 3638–3644.
- (21) Meng, F.; Wang, J.; Dai, E.; Yang, F.; Chen, X.; Wang, S.; Yu, X.; Liu, D.; Jiang, W. Psmir: a database of potential associations between small molecules and miRNAs. *Sci. Rep.* **2016**, *6*, No. 19264.
- (22) Li, J.; Lei, K.; Wu, Z.; Li, W.; Liu, G.; Liu, J.; Cheng, F.; Tang, Y. Network-based identification of microRNAs as potential pharmacogenomic biomarkers for anticancer drugs. *Oncotarget* **2016**, *7*, 45584–45596.
- (23) Qu, J.; Chen, X.; Sun, Y.-Z.; Zhao, Y.; Cai, S.-B.; Ming, Z.; You, Z.-H.; Li, J.-Q. In Silico Prediction of Small Molecule-miRNA Associations Based on the HeteSim Algorithm. *Mol. Ther.–Nucleic Acids* **2019**, *14*, 274–286.
- (24) Guan, N.-N.; Sun, Y.-Z.; Ming, Z.; Li, J.-Q.; Chen, X. Prediction of potential small molecule-associated microRNAs using graphlet interaction. *Front. Pharmacol.* **2018**, *9*, 1152–1163.
- (25) Wang, C.-C.; Chen, X. A Unified Framework for the Prediction of Small Molecule–MicroRNA Association Based on Cross-Layer Dependency Inference on Multilayered Networks. *J. Chem. Inf. Model.* **2019**, *59*, 5281–5293.
- (26) Qu, J.; Chen, X.; Sun, Y.-Z.; Li, J.-Q.; Ming, Z. Inferring potential small molecule–miRNA association based on triple layer heterogeneous network. *J. Cheminf.* **2018**, *10*, 1–14.
- (27) Zhao, Y.; Chen, X.; Yin, J.; Qu, J. SNMFSSMA: using symmetric nonnegative matrix factorization and Kronecker regularized least squares to predict potential small molecule-microRNA association. *RNA Biol.* **2019**, *17*, 281–291.
- (28) Yin, J.; Chen, X.; Wang, C.-C.; Zhao, Y.; Sun, Y.-Z. Prediction of small molecule-microRNA associations by sparse learning and heterogeneous graph inference. *Mol. Pharmaceutics* **2019**, *16*, 3157–3166.
- (29) Velagapudi, S. P.; Gallo, S. M.; Disney, M. D. Sequence-based design of bioactive small molecules that target precursor microRNAs. *Nat. Chem. Biol.* **2014**, *10*, 291–297.
- (30) Ding, P.; Yin, R.; Luo, J.; Kwok, C. K. Ensemble Prediction of Synergistic Drug Combinations Incorporating Biological, Chemical, Pharmacological and Network Knowledge. *IEEE J. Biomed. Health Inf.* **2018**, *23*, 1336–1345.
- (31) Xie, W.; Luo, J.; Pan, C.; Liu, Y. SG-LSTM-FRAME: a computational frame using sequence and geometrical information via LSTM to predict miRNA-gene associations. *Briefings Bioinf.* **2020**, No. bbaa022.
- (32) Zeng, X.; Zhu, S.; Lu, W.; Liu, Z.; Huang, J.; Zhou, Y.; Fang, J.; Huang, Y.; Guo, H.; Li, L. Target identification among known drugs by deep learning from heterogeneous networks. *Chem. Sci.* **2020**, *11*, 1775–1797.
- (33) Zeng, X.; Siyi, Z.; Yuan, H.; Pengyue, Z.; Lang, L.; Jing, L.; Frank, H. L.; J. L. S.; Ruth, N.; Feixiong, C. Network-based prediction of drug–target interactions using an arbitrary-order proximity embedded deep forest. *Bioinformatics* **2020**, *36*, 2805–2812.
- (34) Pliakos, K.; Vens, C. Drug-target interaction prediction with tree-ensemble learning and output space reconstruction. *BMC Bioinf.* **2020**, *21*, 49.
- (35) Ammad-Ud-Din, M.; Khan, S. A.; Malani, D.; Murumagi, A.; Kallioniemi, O.; Aittokallio, T.; Kaski, S. Drug response prediction by inferring pathway-response associations with kernelized Bayesian matrix factorization. *Bioinformatics* **2016**, *32*, 455–463.
- (36) Luo, P.; Yuanyuan, L.; Li-Ping, T.; Fang-Xiang, W. Enhancing the prediction of disease–gene associations with multimodal deep learning. *Bioinformatics* **2019**, *35*, 3735–3742.
- (37) Liu, Y.; Luo, J.; Ding, P. Inferring MicroRNA Targets Based on Restricted Boltzmann Machines. *IEEE J. Biomed. Health Inf.* **2019**, *23*, 427–436.
- (38) Liu, B.; Liu, F.; Wang, X.; Chen, J.; Fang, L.; Chou, K.-C. Pse-in-One: a web server for generating various modes of pseudo components of DNA, RNA, and protein sequences. *Nucleic Acids Res.* **2015**, *43*, W65–W71.
- (39) Cheng, F.; Li, W.; Wu, Z.; Wang, X.; Zhang, C.; Li, J.; Liu, G.; Tang, Y. modeling, Prediction of polypharmacological profiles of drugs by the integration of chemical, side effect, and therapeutic space. *J. Chem. Inf.* **2013**, *53*, 753–762.
- (40) Cheng, F.; Kovács, I. A.; Barabási, A.-L. Network-based prediction of drug combinations. *Nat. Commun.* **2019**, *10*, No. 1197.
- (41) Law, V.; Knox, C.; Djoumbou, Y.; Jewison, T.; Guo, A. C.; Liu, Y.; Maciejewski, A.; Arndt, D.; Wilson, M.; Neveu, V. DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res.* **2013**, *42*, D1091–D1097.
- (42) Cheng, X.; Zhao, S.-G.; Xiao, X.; Chou, K.-C. iATC-mISF: a multi-label classifier for predicting the classes of anatomical therapeutic chemicals. *Bioinformatics* **2017**, *33*, 341–346.
- (43) Xiao, Q.; Luo, J.; Liang, C.; Cai, J.; Ding, P. A graph regularized non-negative matrix factorization method for identifying microRNA-disease associations. *Bioinformatics* **2017**, *34*, 239–248.
- (44) Lee, I.; Blom, U. M.; Wang, P. I.; Shim, J. E.; Marcotte, E. M. Prioritizing candidate disease genes by network-based boosting of genome-wide association data. *Genome Res.* **2011**, *21*, 1109–1121.
- (45) Wang, J. Z.; Du, Z.; Payattakool, R.; Yu, P. S.; Chen, C.-F. A new method to measure the semantic similarity of GO terms. *Bioinformatics* **2007**, *23*, 1274–1281.
- (46) Lu, C.; Yang, M.; Luo, F.; Wu, F.-X.; Li, M.; Pan, Y.; Li, Y.; Wang, J. Prediction of lncRNA–disease associations based on inductive matrix completion. *Bioinformatics* **2018**, *34*, 3357–3364.
- (47) Sun, R.; Luo, Z.-Q. Guaranteed matrix completion via non-convex factorization. *IEEE Trans. Inf. Theory* **2016**, *62*, 6535–6579.
- (48) Chen, X.; Wang, L.; Qu, J.; Guan, N.-N.; Li, J.-Q. Predicting miRNA–disease association based on inductive matrix completion. *Bioinformatics* **2018**, *34*, 4256–4265.
- (49) Liu, X.; Zhai, D.; Zhao, D.; Zhai, G.; Gao, W. Progressive image denoising through hybrid graph Laplacian regularization: A unified framework. *IEEE Trans. Image Process.* **2014**, *23*, 1491–1503.
- (50) Liu, Y.; Wu, M.; Miao, C.; Zhao, P.; Li, X. Neighborhood Regularized Logistic Matrix Factorization for Drug-Target Interaction Prediction. *PLoS Comput. Biol.* **2016**, No. e1004760.
- (51) Zhang, Z.; Zhang, X.; Wu, M.; Ouyang, L.; Zhao, X.; Li, X. A Graph Regularized Generalized Matrix Factorization Model for Predicting Links in Biomedical Bipartite Networks. *Bioinformatics* **2020**, *36*, 3474–3481.
- (52) Luo, J.; Shen, C.; Lai, Z.; Cai, J.; Ding, P. Incorporating Clinical, Chemical and Biological Information for Predicting Small Molecule-microRNA Associations based on Non-negative Matrix Factorization. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2020**, *1*.
- (53) Xuan, P.; Cao, Y.; Zhang, T.; Wang, X.; Pan, S.; Shen, T. Drug repositioning through integration of prior knowledge and projections of drugs and diseases. *Bioinformatics* **2019**, *35*, 4108–4119.
- (54) Xuan, P.; Shen, T.; Wang, X.; Zhang, T.; Zhang, W. Inferring disease-associated microRNAs in heterogeneous networks with node attributes. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2018**, *17*, 1019–1031.
- (55) Kozomara, A.; Birgaoanu, M.; Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* **2018**, *47*, D155–D162.

(56) Srivastava, N.; Salakhutdinov, R. R. Multimodal learning with deep boltzmann machines. *Adv. Neural Inf. Process. Syst.* **2012**, *2012*, 2222–2230.

(57) Hinton, G. E. Training products of experts by minimizing contrastive divergence. *Neural Comput.* **2002**, *14*, 1771–1800.

(58) Tieleman, T. In *Training Restricted Boltzmann Machines Using Approximations to the Likelihood Gradient*. Proceedings of the 25th International Conference on Machine Learning; ACM, 2008; pp 1064–1071.

(59) Cho, K.; Raiko, T.; Ilin, A. In *Parallel Tempering is Efficient for Learning Restricted Boltzmann Machines*. The 2010 International Joint Conference on Neural Networks (ijcnn); Citeseer, 2010; pp 1–8.

(60) Zheng, S.; Hao, Y.; Lu, D.; Bao, H.; Xu, J.; Hao, H.; Xu, B. Joint entity and relation extraction based on a hybrid neural network. *Neurocomputing* **2017**, *257*, 59–66.

(61) Miwa, M.; Bansal, M. In *End-to-End Relation Extraction Using LSTMs on Sequences and Tree Structures*. Meeting of the Association for Computational Linguistics, 2016; Vol. 1, pp 1105–1116.

(62) Liu, X.; Wang, S.; Meng, F.; Wang, J.; Zhang, Y.; Dai, E.; Yu, X.; Li, X.; Jiang, W. SM2miR: a database of the experimentally validated small molecules' effects on microRNA expression. *Bioinformatics* **2012**, *29*, 409–411.

(63) Chou, C.-H.; Chang, N.-W.; Shrestha, S.; Hsu, S.-D.; Lin, Y.-L.; Lee, W.-H.; Yang, C.-D.; Hong, H.-C.; Wei, T.-Y.; Tu, S.-J. miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res.* **2015**, *44*, D239–D247.

(64) Ding, P.; Shen, C.; Lai, Z.; Liang, C.; Li, G.; Luo, J. Incorporating Multisource Knowledge To Predict Drug Synergy Based on Graph Co-regularization. *J. Chem. Inf. Model.* **2020**, *60*, 37–46.

(65) Kwon, J. J.; Willy, J. A.; Quirin, K. A.; Wek, R. C.; Korc, M.; Yin, X.-M.; Kota, J. Novel role of miR-29a in pancreatic cancer autophagy and its therapeutic potential. *Oncotarget* **2016**, *7*, 71635–71650.

(66) Li, J.; Chan, M. C.; Yu, Y.; Bei, Y.; Chen, P.; Zhou, Q.; Cheng, L.; Chen, L.; Ziegler, O.; Rowe, G. C. miR-29b contributes to multiple types of muscle atrophy. *Nat. Commun.* **2017**, *8*, No. 15201.