



# Prediction of Potential Small Molecule-Associated MicroRNAs Using Graphlet Interaction

Na-Na Guan<sup>1</sup>, Ya-Zhou Sun<sup>1</sup>, Zhong Ming<sup>1,2</sup>, Jian-Qiang Li<sup>1\*</sup> and Xing Chen<sup>3\*</sup>

<sup>1</sup> College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China, <sup>2</sup> National Engineering Laboratory for Big Data System Computing Technology, Shenzhen University, Shenzhen, China, <sup>3</sup> School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, China

## OPEN ACCESS

### Edited by:

Zhi-Liang Ji,  
Xiamen University, China

### Reviewed by:

Quan Zou,  
Tianjin University, China  
Feixiong Cheng,  
Cleveland Clinic, United States  
Francesco Pappalardo,  
Università degli Studi di Catania, Italy

### \*Correspondence:

Jian-Qiang Li  
lijq@szu.edu.cn  
Xing Chen  
xingchen@amss.ac.cn;  
xingchen@cumt.edu.cn

### Specialty section:

This article was submitted to  
Translational Pharmacology,  
a section of the journal  
Frontiers in Pharmacology

**Received:** 08 May 2018

**Accepted:** 24 September 2018

**Published:** 15 October 2018

### Citation:

Guan N-N, Sun Y-Z, Ming Z, Li J-Q  
and Chen X (2018) Prediction  
of Potential Small  
Molecule-Associated MicroRNAs  
Using Graphlet Interaction.  
*Front. Pharmacol.* 9:1152.  
doi: 10.3389/fphar.2018.01152

MicroRNAs (miRNAs) have been proved to be targeted by the small molecules recently, which made using small molecules to target miRNAs become a possible therapy for human diseases. Therefore, it is very meaningful to investigate the relationships between small molecules and miRNAs, which is still yet in the newly-developing stage. In this paper, we presented a prediction model of Graphlet Interaction based inference for Small Molecule-MiRNA Association prediction (GISMMA) by combining small molecule similarity network, miRNA similarity network and known small molecule-miRNA association network. This model described the complex relationship between two small molecules or between two miRNAs using graphlet interaction which consists of 28 isomers. The association score between a small molecule and a miRNA was calculated based on counting the numbers of graphlet interaction throughout the small molecule similarity network and the miRNA similarity network, respectively. Global and two types of local leave-one-out cross validation (LOOCV) as well as five-fold cross validation were implemented in two datasets to evaluate GISMMA. For Dataset 1, the AUCs are 0.9291 for global LOOCV, 0.9505, and 0.7702 for two local LOOCVs,  $0.9263 \pm 0.0026$  for five-fold cross validation; for Dataset 2, the AUCs are 0.8203, 0.8640, 0.6591, and  $0.8554 \pm 0.0063$ , in turn. In case study for small molecules, 5-Fluorouracil, 17 $\beta$ -Estradiol and 5-Aza-2'-deoxycytidine, the numbers of top 50 miRNAs predicted by GISMMA and validated to be related to these three small molecules by experimental literatures are in turn 30, 29, and 25. Based on the results from cross validations and case studies, it is easy to realize the excellent performance of GISMMA.

**Keywords:** small molecule, microRNA, association prediction, graphlet interaction, similarity calculation

## INTRODUCTION

MicroRNAs (miRNAs) are a family of small non-coding RNAs, having about 22 nucleotides in length, which regulate gene expression at a post-transcriptional level (Ambros, 2003). The first miRNA was discovered over 30 years ago in the *Caenorhabditis elegans*. Subsequently, thousands of miRNAs have been discovered in many organisms, and there are currently 2588 annotated miRNAs in the human genome (Kozomara and Griffiths-Jones, 2014). MiRNAs can simultaneously regulate the expression of hundreds of genes due to the fact that their nucleotide pairing by

complementarity is imperfect (He and Hannon, 2004). In this manner, they play a critical role in a variety of crucial processes such as tissue development, morphogenesis, apoptosis, signal transduction pathways, etc., (Esquela-Kerscher and Slack, 2006; Spizzo et al., 2009; Wang and Lee, 2009). This additionally implicates them in an array of disease associated processes. The development of large-expression screens has been proven useful in identifying novel miRNAs involved in diseases, which could potentially become an attractive therapeutic target (Monroig and Calin, 2013; Chen et al., 2017a, 2018a,b,c; Matsui and Corey, 2017).

Regulation of miRNAs by small molecules is an efficient mean to modulate endogenous miRNA function and to treat miRNA-related diseases (Xia et al., 2015). Small molecules have been thoroughly used with clinical applications for numerous diseases (Zhang et al., 2009). However, drug discovery and development are currently an extremely long process, which takes approximately 10–15 years (Monroig Pdel et al., 2015). Also, drug production results in an incredible economic burden and patients end up having to pay exaggerated prices for their treatments (Chen et al., 2015; Monroig Pdel et al., 2015). The use of chemical compounds that are already FDA approved to treat a specific disease would accelerate the process of completing toxicological studies and clinical trials in order to apply them to other diseases. It would shorten both money expenses and time consuming processes.

As miRNAs have been associated with many diseases (Chen et al., 2017b), the development of small-molecule drugs targeting specific miRNAs seems to be a promising approach to meet the challenge (Monroig Pdel et al., 2015). Small molecule may modulate the expression of miRNAs by either activating or repressing their transcription (Xia et al., 2015). Transcriptional inhibitors were identified by completing a small molecule screen in which a 3' UTR complementary to miR-21 was inserted into a luciferase mRNA reporter (Gumireddy et al., 2008). This study identified a type of diazobenzene as miR-21 transcriptional inhibitors (Gumireddy et al., 2008). Small molecules were also discovered to modulate transcription of miR-122, a highly expressed and liver-specific miRNA whose aberrant expression is associated with hepatocellular carcinoma (Thomas and Deiters, 2013). Two small molecules that inhibit transcription and another small molecule that promotes transcription of pri-miR-122 were identified using a luciferase reporter system (Thomas and Deiters, 2013). The examples above show that miRNA expression can be altered with small molecules, providing promise to expand miRNAs from diagnostic signatures of disease to therapeutic targets. Therefore, the prediction of associations between small molecules and miRNAs could promote the drug repurposing for miRNA-related diseases. Besides, since the regulation of miRNA expression can be caused by targeting miRNAs directly (Zhang et al., 2010) or by targeting the relative proteins (Lim et al., 2016), identifying the small molecule-miRNA associations would be conducive to the drug discovery. However, experimental methods to study the small molecular-miRNA association are expensive and time-consuming, which makes it urgent to develop computational approaches to provide reliable predictions that can give some guidance to experiments.

Recently, several computational models have been proposed to investigate the relations between small molecules and miRNAs. For example, Jiang et al. (2012) proposed a high-throughput method to investigate the biological connections between small molecules and miRNAs in 23 human cancers based on transcriptional responses, which was the first model to systematically study the associations between bioactive small molecules and miRNAs. They constructed a complex Small molecule and MiRNA Network (SMirN) for each cancer and explored the molecular and functional features for small molecule modules, as well as miRNA modules for each cancer type. Each module of small molecular was linked to a miRNA, and each module of miRNA was connected with one small molecular. One of the advantages of this method is that it does not need to know the information of small molecule structure or miRNA structure in advance. However, the reliability of the approach was limited due to the small data of transcriptional response to genome-wide miRNA perturbations. Furthermore, Meng et al. (2014) built a bioactive Small molecule and miRNA association Network in Alzheimer's Disease (SMiRN-AD) through comparing the gene expression profiles after bioactive small molecule treating with the AD-related miRNA (ADM) regulating expressions, to get the scores of associations between small molecules and ADMs. Besides, the positive and negative associations were identified to investigate the biological insights of the SimRN-AD. Recently, Wang et al. (2016) developed another method to identify small molecule-miRNA associations based on their functional similarity. They searched the functional link of each small molecule-miRNA pair by calculating Gene Ontology enrichment after identifying differentially expressed genes for small molecules and miRNAs. Compared with previous models based on transcriptional responses, this method is more repeatable by using functional associations. Additionally, Lv et al. (2015) presented a novel computational model to predict potential associations between small molecules and miRNAs. They implemented the random walk with restart algorithm on an comprehensive network, which was established by combining small molecule similarity, miRNA similarity, as well as known small molecule-miRNA associations. Especially, this model can predict the novel related miRNAs for small molecules without any known associated miRNAs. However, it has too many adjustable parameters that need to be affirmed. Moreover, Li et al. (2016) developed a network based framework called predictive Small Molecule-miRNA Network-Based Inference (SMiR-NBI), to investigate the underlying regulations of anticancer drugs on miRNAs. This model constructed a heterogeneous network that was composed of drugs, miRNAs and genes to conduct a network based algorithm. It is mentionable that the accuracy of this method is quite high even it only depended on the network topology information. However, SMiR-NBI could not be applied to prediction of isolated miRNAs that have no interlinked small molecules. Besides, it failed to predict potential miRNAs associated with small molecules that had different dose-responses, due to lack of known data.

So far, the number of computational models is still not satisfying for the prediction of novel associations between small molecules and miRNAs. Moreover, there are still some limitations

existing in the previous models. In order to predict potential small molecule-miRNA associations more effectively and reliably, in this paper, we presented the Graphlet Interaction based inference for Small Molecule-miRNA Association prediction (GISMMA). In this model, the similarity of small molecules and the similarity of miRNAs were combined with known associations between small molecules and miRNAs in two different datasets, which were labeled with Dataset 1 and Dataset 2. In Dataset 1, only a fraction of small molecules and miRNAs were involved in known small molecule-miRNA associations, whereas in Dataset 2 all small molecules and miRNAs were implicated in known small molecule-miRNA associations. Based on the measuring of graphlet interaction between any two nodes on the network of small molecules and on the network of miRNAs, respectively, we can compute the correlation scores of small molecule-miRNA pairs. We have implemented leave-one-out cross validation (LOOCV) and five-fold cross validation to evaluate the performance of GISMMA. The AUCs of global LOOCV are 0.9291 and 0.8203 for Dataset 1 and Dataset 2, respectively; the AUCs of local LOOCV by ranking the small molecules for each fixed miRNA are, respectively 0.9505 and 0.8640 for the two datasets; the AUCs of local LOOCV by ranking the miRNAs for each fixed small molecule are, respectively 0.7702 and 0.6591 for the two datasets. And the average AUCs and standard deviations of five-fold cross validations are  $0.9263 \pm 0.0026$  and  $0.8088 \pm 0.0044$  for the two datasets, respectively. In case study, small molecule was set as a new one by turning all known related miRNAs into unknown ones. GISMMA was then applied to predicting latent related miRNAs for each small molecule based on the Dataset 1. For the small molecules, 5-Fluorouracil, 17 $\beta$ -Estradiol and 5-Aza-2'-deoxycytidine, there were in turn 30, 29, and 25 out of top 50 predicted miRNAs, which were validated to be associated with these three small molecules by experimental literatures, respectively. The results both in cross validations and case studies have suggested that GISMMA is a powerful and reliable model to predict novel associations between small molecules and miRNAs.

## MATERIALS AND METHODS

### Small Molecule-miRNA Associations

In this paper, we obtained the known small molecule-miRNA associations from SM2miR (Version 1) (Liu et al., 2013). The total number of known associations is 664. For comparison of model performance on different datasets, we have constructed two datasets. Dataset 1 consists of 831 small molecules extracted and integrated from SM2miR, DrugBank (Knox et al., 2011) and PubChem (Wang et al., 2009), and 541 miRNAs that were collected from SM2miR, HMDD (Lu et al., 2008), miR2Disease (Jiang et al., 2009) and PhenomiR (Jiang et al., 2009; Ruepp et al., 2010). In Dataset 1, there are only 39 small molecules and 286 miRNAs implicated in the 664 known associations, while 792 small molecules and 255 miRNAs are completely new ones without any known associations. Dataset 2 is only composed of those 39 small molecules and 286 miRNAs, which are involved in the known associations. Based on the known data, an adjacency

matrix  $A$  was constructed to represent the relations between small molecules and miRNAs, in which  $A(i, j)$  was set to be 1 if there is an association between small molecule  $s(i)$  and miRNA  $m(j)$ , 0 otherwise.

### Small Molecule Similarity

In this paper, according to the method proposed in (Lv et al., 2015), the small molecule similarity was calculated by integrating four usual small molecule similarities which were side effect based similarity that was computed by Jaccard score using small molecule side effect dataset (Gottlieb et al., 2011), functional consistency based similarity that was obtained by comparing the function of small molecule target genes (Lv et al., 2012), chemical structure based similarity that was calculated with the method of chemical structure comparison between any two small molecules (Hattori et al., 2003), and indication phenotype based similarity that was constructed through identifying phenotype similarity between small molecule related diseases (Gottlieb et al., 2011). Therefore, the integrated similarity of small molecules can be computed with the following formula:

$$SS = \frac{\beta_1 S_S^D + \beta_2 S_S^T + \beta_3 S_S^C + \beta_4 S_S^S}{\sum_{i=1}^4 \beta_i} \quad (1)$$

where,  $S_S^D$ ,  $S_S^T$ ,  $S_S^C$ , and  $S_S^S$  denote the four different similarity types, respectively, i.e., indication phenotype based similarity, functional consistency based similarity, chemical structure based similarity and side effect based similarity, and  $\beta_i$  ( $i = 1, 2, 3, 4$ ) are the weights used to balance the different similarity contributions, whose default values were all set as 1.

### MiRNA Similarity

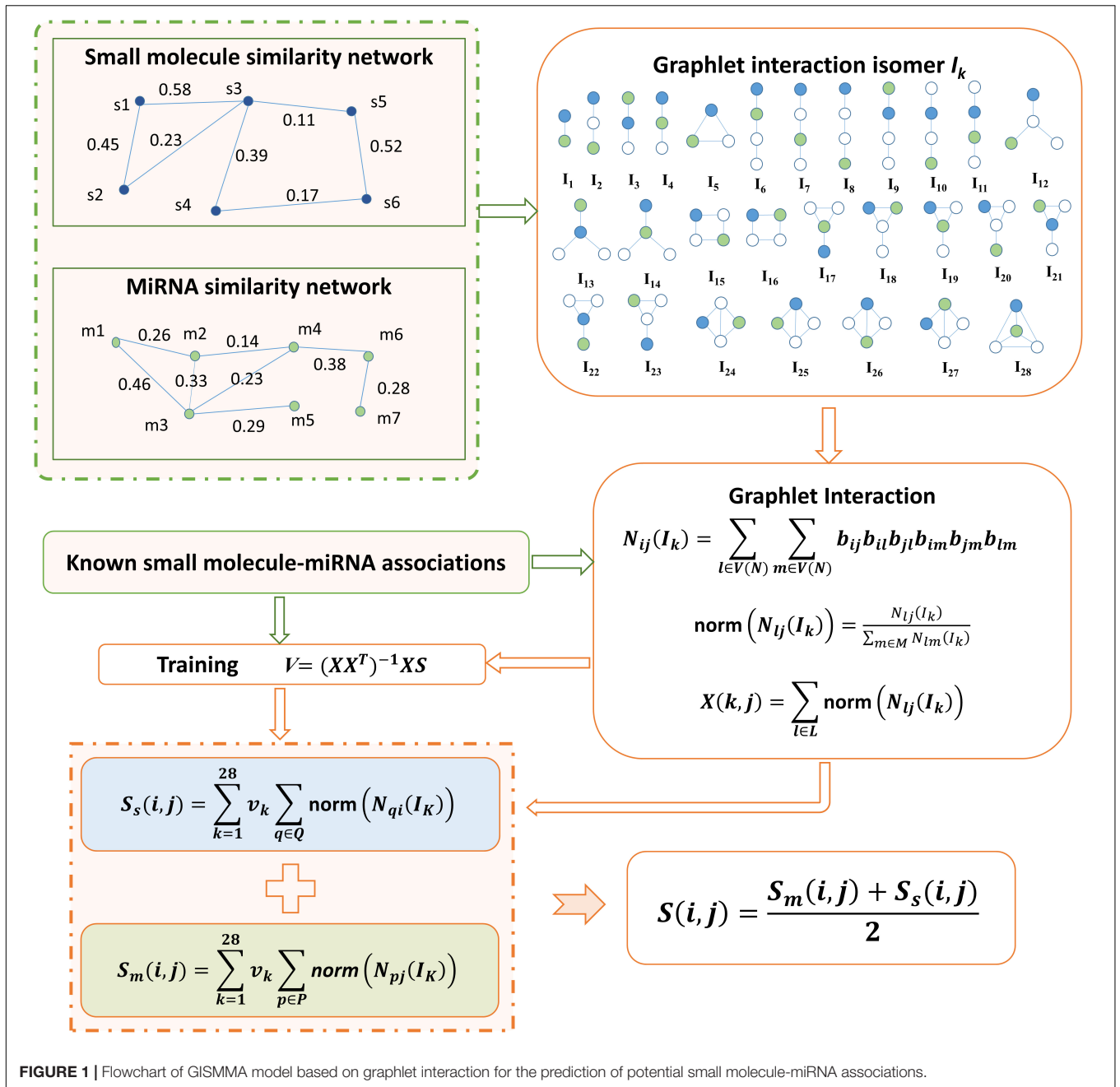
The miRNA similarity we used in this paper was established using the method in (Lv et al., 2015), by combining functional consistency based similarity that was calculated by comparing the function of miRNA target genes (Lv et al., 2012) and indication phenotype based similarity that was computed by measuring phenotype similarity between diseases associated with miRNAs (Gottlieb et al., 2011). Similarly, to reduce the bias of each similarity measurement, the integrated similarity of miRNAs was defined as follows:

$$SM = \frac{\alpha_1 S_M^D + \alpha_2 S_M^T}{\sum_{j=1}^2 \alpha_j} \quad (2)$$

where,  $S_M^D$  is the indication phenotype based similarity and  $S_M^T$  represents the functional consistency based similarity, and  $\alpha_j$  ( $j = 1, 2$ ) are the weights of each similarity measurement, which were both set as 1.

### GISMMA

In this study, by integrating small molecule similarity, miRNA similarity and known associations between small molecules and miRNAs, we developed a graphlet interaction based method to predict the potential associations between small molecules and miRNAs, which is motivated by the study of Wang et al. (2014). Prediction code of our model is available



at: <https://github.com/AnnaGuan/GISMMA/tree/AnnaGuan-patch-1>. The concept of graphlet interaction is traced to the definition in (Wang et al., 2014), which describes the relationship between any two nodes in a graphlet that is a type of subgraph in a large network. As was done in (Wang et al., 2014), in GISMMA only those graphlets that have 1 to 4 nodes were used, based on which 28 graphlet interaction isomers were constructed, denoted by labels  $I_1$  to  $I_{28}$  in **Figure 1**. The graphlet interaction isomer depends on the positions of the two involved nodes, which means that the graphlet interaction between two nodes have two different set of isomers. Through counting the number of each isomer, we can represent the graphlet interaction between any

two nodes in a network with a vector that contains 28 numbers (Przulj, 2007; Wang et al., 2014).

We have created a network  $NS$  to represent the small molecule similarity and a network  $NM$  to represent the miRNA similarity, where each node in the network denotes a small molecule or a miRNA. The edge with similarity value as its weight exists to link any two nodes that have similarity. The associations between small molecules and miRNAs were investigated in the two similarity networks  $NS$  and  $NM$ , respectively.

In the miRNA network  $NM$ , the number of isomer  $I_k$  for graphlet interaction from node  $m(i)$  to node  $m(j)$  can be

calculated as follows (Wang et al., 2014):

$$N_{ij}(I_k) = \sum_{l \in V(NM)} \sum_{m \in V(NM)} b_{ij} b_{il} b_{jl} b_{im} b_{jm} b_{lm} \quad (3)$$

where  $V(NM)$  denotes the node set of all nodes in network  $NM$ ,  $l$  and  $m$  are two nodes different with node  $m(i)$  and  $m(j)$ , and  $b$  is defined as:

$$b_{st} = \begin{cases} a_{st} & s \text{ and } t \text{ has a link in } I_k \\ 1 - a_{st} & s \text{ and } t \text{ has no link in } I_k \end{cases} \quad (4)$$

where,  $a_{st}$  is the edge weight assigned with the similarity value of  $m(s)$  and  $m(t)$ . Especially,  $a_{st}$  is 0 when nodes  $m(s)$  and  $m(t)$  have no connection. Then we normalized the graphlet interaction as follows:

$$\text{norm}(N_{ij}(I_k)) = \frac{N_{ij}(I_k)}{\sum_{m \in M} N_{im}(I_k)} \quad (5)$$

where  $M$  contains all other nodes but  $m(i)$ . Based on the normalized form in equation (5), we can compute the association score of a small molecule-miRNA pair as follows:

$$S_m(i, j) = \sum_{k=1}^{28} v_k \sum_{p \in P(i)} \text{norm}(N_{pj}(I_k)) \quad (6)$$

where  $i$  denotes a small molecule  $s(i)$  and  $j$  denotes a miRNA  $m(j)$ ,  $v_k$  is the weight of the  $k$ th isomer,  $P(i)$  is the set of miRNAs with known associations with small molecule  $s(i)$ . By defining the summation of norm in equation (6) as following:

$$X_m(k, j) = \sum_{p \in P(i)} \text{norm}(N_{pj}(I_k)) \quad (7)$$

we can modify equation (6) into the matrix form as following:

$$S_m = X_m^T V_m \quad (8)$$

The weight coefficients  $V_m$  can be learnt from known associations by performing a simple linear regression (Wang et al., 2014), which is given as following:

$$V_m = (X_m X_m^T)^{-1} X_m S_m \quad (9)$$

We computed the number of graphlet interaction isomer between two small molecules in the similar way as described in equations (3–5). Then the association score between small molecule  $s(i)$  and miRNA  $m(j)$  can be calculated in the small molecule network  $NS$  as follows:

$$S_s(i, j) = \sum_{k=1}^{28} v_k \sum_{q \in Q(j)} \text{norm}(N_{qi}(I_k)) \quad (10)$$

where  $Q(j)$  is the set of small molecules that have known associations with miRNA  $m(j)$ . Also, the term of summation of norm in equation (10) can be defined with the matrix:

$$X_s(k, j) = \sum_{q \in Q(j)} \text{norm}(N_{qi}(I_k)) \quad (11)$$

Thus equation (10) was rewritten as:  $S_s = X_s^T V_s$ , and the undetermined matrix  $V_s$  can be obtained by training the model with known association scores:

$$V_s = (X_s X_s^T)^{-1} X_s S_s \quad (12)$$

Finally, we calculated the association score between small molecule  $s(i)$  and miRNA  $m(j)$  by combining the scores from  $NM$  and  $NS$  in a simple average form as following:

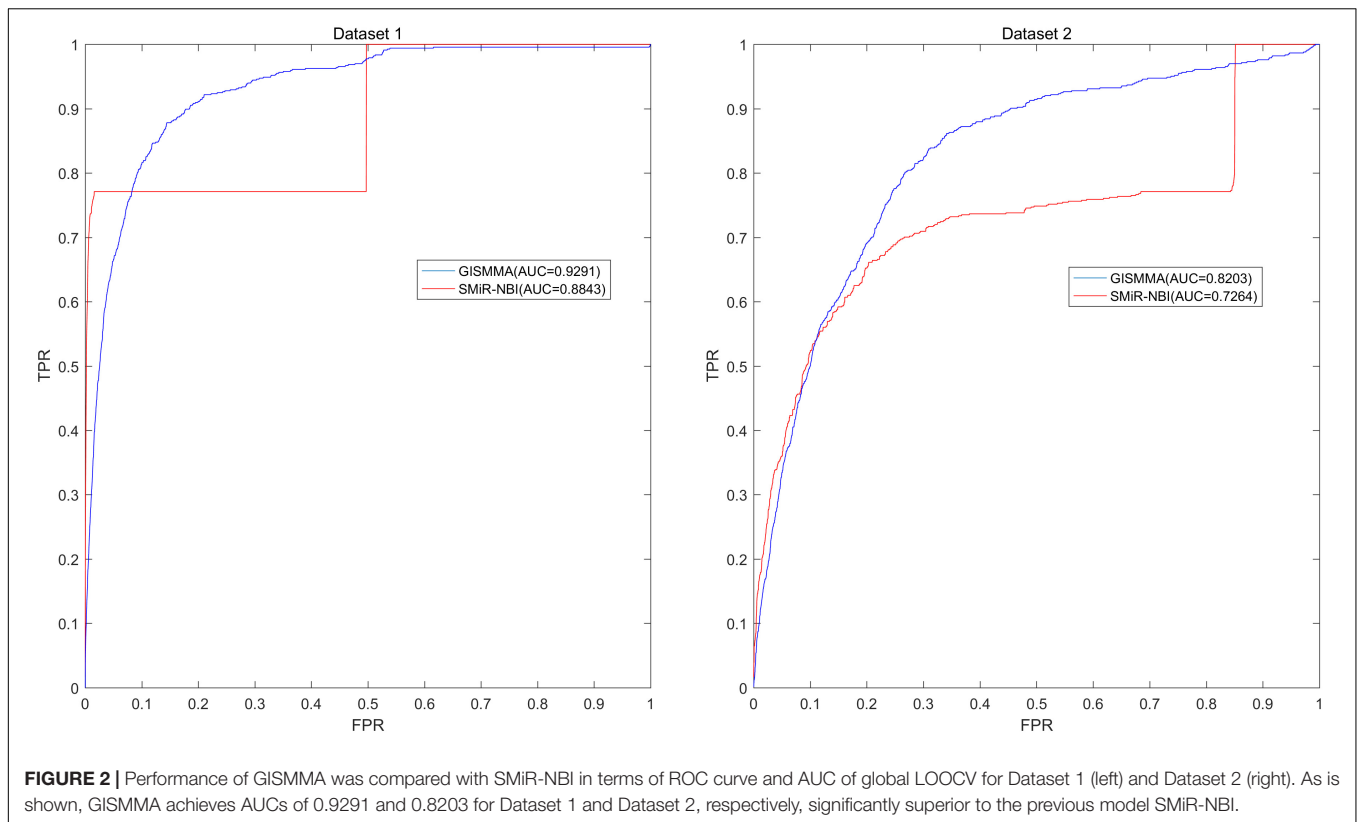
$$S(i, j) = \frac{S_m(i, j) + S_s(i, j)}{2} \quad (13)$$

## RESULTS

### Performance Evaluation

In this work, two commonly used methods, LOOCV and five-fold cross validation, were implemented to evaluate the performance of GISMMA based on Dataset 1 and Dataset 2, respectively. The LOOCV has three different types including global LOOCV, local LOOCV of ranking small molecules for fixed miRNA and local LOOCV of ranking miRNAs for fixed small molecule. Each confirmed association we collected was taken as the test sample one by one and the rest of known associations were considered as the training samples in LOOCV. Candidate samples in global LOOCV consist of all the small molecule-miRNA pairs that have no known associations. In the case of local, we only consider those small molecules that do not relate to the fixed miRNA or those miRNAs unconnected to the fixed small molecule in the test sample as candidates. The scores as association probabilities were computed using the GISMMA method for both test sample and all candidate samples. Then we ranked them for the corresponding type of LOOCV. The five-fold cross validation was performed in the following steps. Firstly, all the known small molecule-miRNA associations were randomly split into five parts with equal size. Secondly, the five parts take turns to act as the test sample set one after another and the other four parts as the training sample sets; similarly, all small molecule-miRNA pairs that have no known associations play the roles of candidate samples. Thirdly, the test samples as well as the candidate samples were endowed with association scores by GISMMA. Finally, each test sample was picked out in turn to be compared with candidate samples according to their scores. The model was considered to be successfully predict the test sample only when its rank exceeded the given rank threshold.

Based on the ranking, the receiver operating characteristic (ROC) curves were used to illustrate the results of the three types of LOOCV described above, in which the abscissa axis is true positive rate (TPR, sensitivity) and the ordinate axis represents false positive rate (FPR, 1-specificity) for different thresholds given in advance. The sensitivity means the ratio that the positive samples rank above the given threshold, while the specificity is defined as the percentage of candidate samples whose ranks are below the set threshold. The area



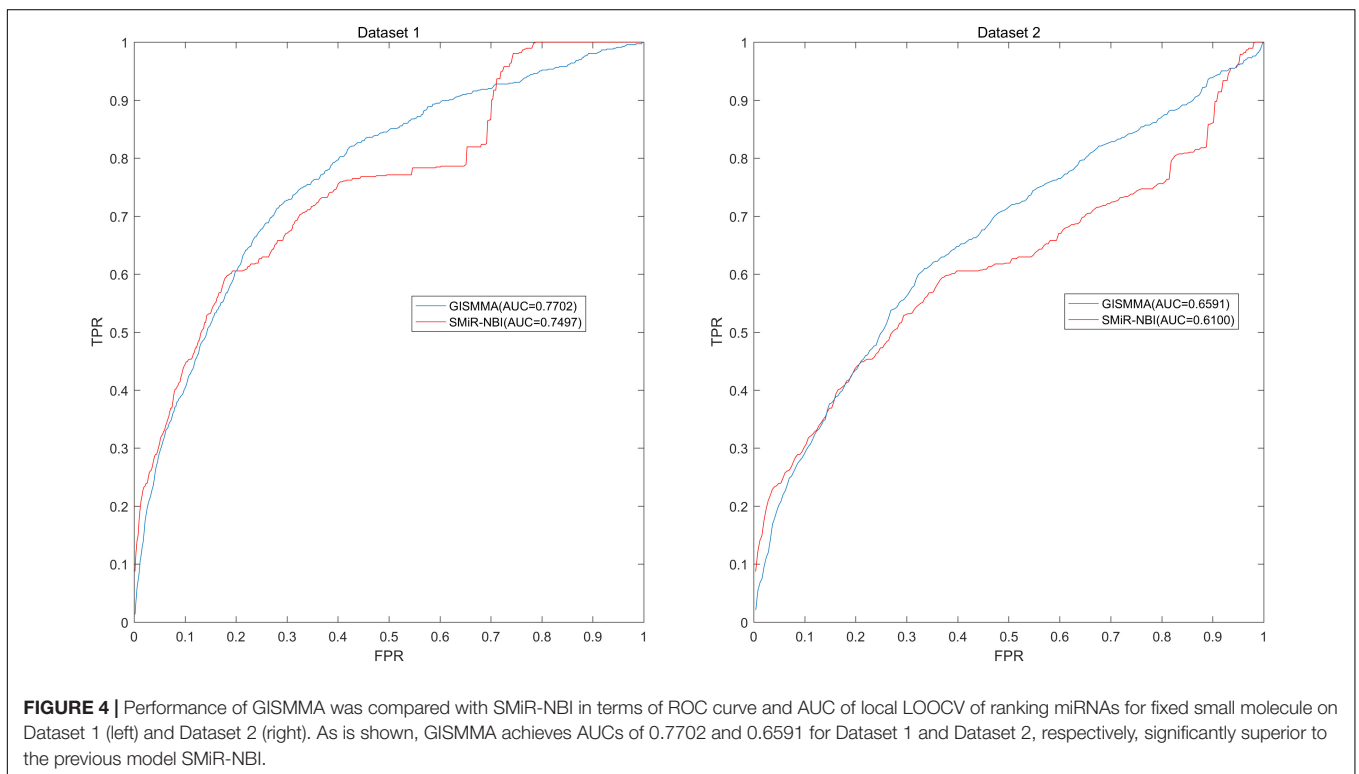
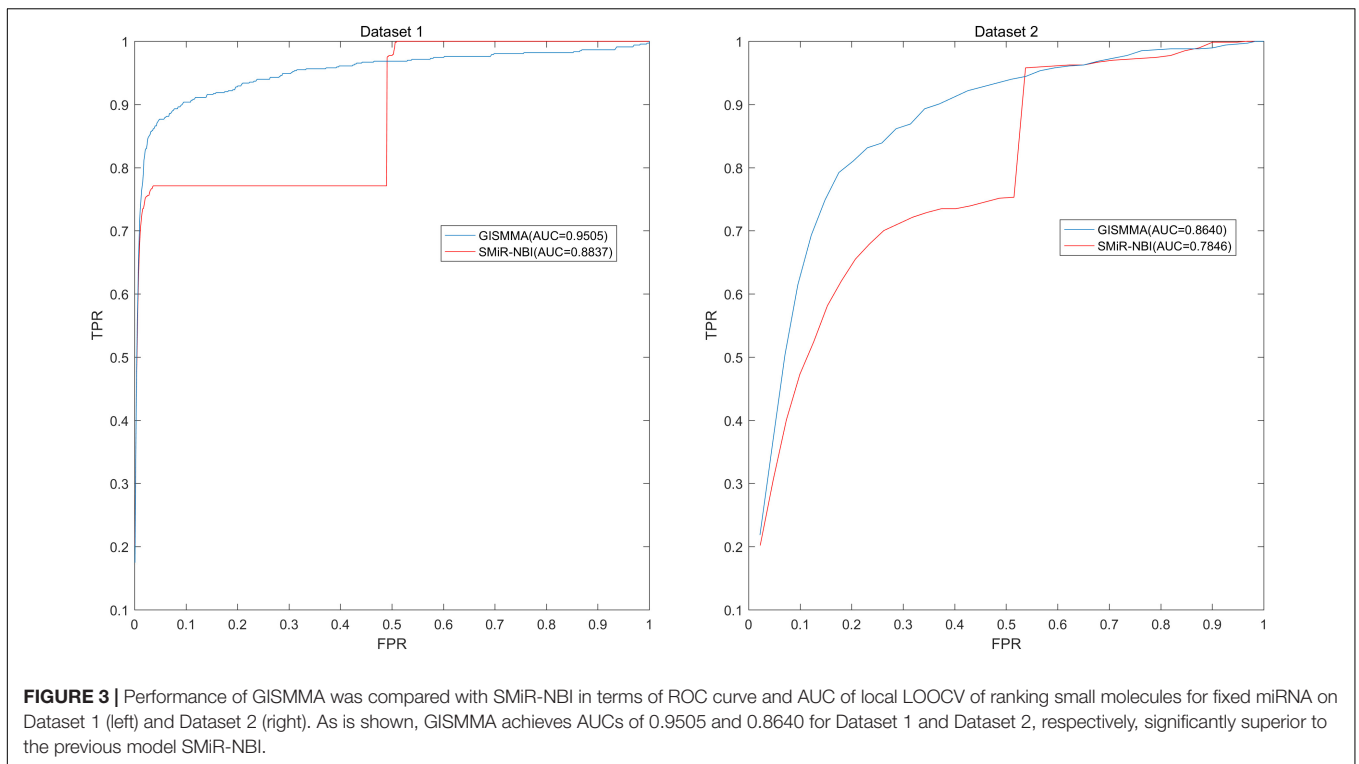
**FIGURE 2 |** Performance of GISMMA was compared with SMiR-NBI in terms of ROC curve and AUC of global LOOCV for Dataset 1 (left) and Dataset 2 (right). As is shown, GISMMA achieves AUCs of 0.9291 and 0.8203 for Dataset 1 and Dataset 2, respectively, significantly superior to the previous model SMiR-NBI.

under the ROC curve (AUC) was correspondingly calculated to estimate the reliability of the GISMMA. When the model correctly predicts all test samples,  $AUC = 1$ ; but if the model has a random prediction,  $AUC = 0.5$ . To make comparison with previous method, we implemented SMiR-NBI (Li et al., 2016) for global and two types of local LOOCVs, 5-fold cross validation based on the same datasets. The global AUCs of GISMMA for Dataset 1 and Dataset 2 are 0.9291 and 0.8203, respectively, which are shown in **Figure 2** in comparison with previous model SMiR-NBI whose results are 0.8843 and 0.7264, respectively. In the case of local LOOCV of ranking small molecules for fixed miRNA, the AUCs of GISMMA for Dataset 1 and Dataset 2 are 0.9505 and 0.8640, respectively, compared with 0.8837 and 0.7846 of SMiR-NBI, which can be seen in **Figure 3**. The results of local LOOCV of ranking miRNAs for fixed small molecule are shown in **Figure 4**, from which we can see that the AUCs of GISMMA and SMiR-NBI are 0.7702, 0.7497 for Dataset 1, and 0.6591, 0.6100 for Dataset 2, respectively. Besides, in five-fold cross validation, the average AUCs with standard deviations of GISMMA and SMiR-NBI are  $0.9263 \pm 0.0026$ ,  $0.8554 \pm 0.0063$  for Dataset 1, and  $0.8088 \pm 0.0044$ ,  $0.7104 \pm 0.0087$  for Dataset 2. The **Table 1** lists the comparison of GISMMA and SMiR-NBI for all AUC results of the four types of cross validations on two datasets. We can make a conclusion from the comparisons that the novel method proposed in this work is more reliable and more effective in predicting potential associations between small molecules and miRNAs.

## Case Study

Based on the known database and published references in PubMed database, we studied three common small molecules to further evaluate the predictive ability of GISMMA, in which the small molecule in study was set as a new one by taking away its known associations. We ulteriorly observed the number of the experimentally verified miRNAs in the top 50 ones predicted to be related to the three small molecules, respectively.

The small molecular 5-Fluorouracil (5-FU) is a widely used chemotherapeutic drug in colorectal cancer (Windle et al., 1987). For a long time, the 5-FU-induced cytotoxic effects were thought to result exclusively from its impact on DNA metabolism (Andreuccetti et al., 1996; Airley, 2009). However, several evidences indicated that the cytotoxic effect of 5-FU also results from its capacity to alter RNA metabolism and mRNA expression (Longley et al., 2003). Exposure to 5-FU promotes a profound transcriptional reprogramming leading to modification of mRNA and miRNAs expression profiles that contributes in modifying cell fate (Hernandez-Vargas et al., 2006; Rossi et al., 2007; Shah et al., 2011). After implementing GISMMA, we got the total ranking of potential miRNAs associated with 5-FU. As the result shown, among the top 10 and 50 potential 5-FU-related miRNAs, there were 8 and 30 miRNAs confirmed by experiments, respectively (See **Table 2**). For instance, miR-21 and miR-23a were predicted as the first and fifth candidates for 5-FU, respectively, which were significantly down regulated in comparison between 5-FU treated and control samples in miRNA microarray analysis of 5-FU treated MCF-7 cells (Shah et al.,



2011). Besides, miR-24-1, the third candidate in the ranking list, showed a significantly down regulation in HCT-8 colon cancer cell after exposure to 5-FU (Zhou et al., 2010). In addition, MiR-27b that ranked the fourth in the prediction list of 5-FU was

found to be consistently up regulated in human colon cancer cells HC.21 following exposure to 5-FU *in vitro* (Rossi et al., 2007).

The small molecular 17 $\beta$ -Estradiol (E2) is the principal intracellular human estrogen that exerts important effects on

**TABLE 1** | The comparison results between GISMMA and SMiR-NBI on AUC values of four cross validations based on two datasets.

DATASET	MODEL	GLOBAL LOOCV	LOCAL LOOCV (fix miRNA)	LOCAL LOOCV (fix SM)	5-FOLD CV
Dataset 1	GISMMA	0.9291	0.9505	0.7702	0.9263 ± 0.0026
	SMiR-NBI	0.8843	0.8837	0.7497	0.8554 ± 0.0063
Dataset 2	GISMMA	0.8203	0.8640	0.6591	0.8088 ± 0.0044
	SMiR-NBI	0.7264	0.7846	0.6100	0.7104 ± 0.0087

the reproductive as well as many other organ systems in both men and women (Simpson and Santen, 2015). The analogs of estradiol exhibit significant anticancer activity against human breast cancer cell lines (Sathish Kumar et al., 2014). Estrogens have associations with cancer in target tissues, which is because they have a phenolic ring structure in common with the carcinogenic hydrocarbons (Ryan, 1982). After implementing GISMMA, we got the total ranking of the E2-associated miRNAs. As the result shown, among the top 10 and 50 potential E2-related miRNAs, there were 5 and 29 miRNAs confirmed by experiments, respectively (See **Table 3**). For example, miR-21, miR-27b, and miR-23a dominated in turn the first, fourth, and fifth places of the ranking list predicted for E2, which were all down regulated after treatment of MCF-7 cells with E2 (Bhat-Nakshatri et al., 2009; Tilghman et al., 2012). Besides, E2 showed a capacity to

down regulate the expression level of miR-21 in breast cancer cells (Selcuklu et al., 2012).

The small molecular 5-Aza-2'-deoxycytidine (5-Aza-CdR) is a nucleoside analog inhibitor of DNA methyltransferase (DNMT). It has been used to reverse methylation and reactivate the expression of silenced genes (Patra and Bettuzzi, 2009). 5-Aza-CdR is able to suppress the growth of various tumors *in vitro*, animal models, and clinical trials including prostate cancer (Hurtubise and Momparler, 2004; Issa et al., 2004; McCabe et al., 2006). We performed GISMMA on 5-Aza-CdR, and got the total ranking of the predicted miRNAs. As the result shown, among the top 10 and 50 potential 5-Aza-CdR related miRNAs, there were 7 and 25 miRNA-5-Aza-CdR associations confirmed by experiments (See **Table 4**). For example, in the ranking list of miRNAs predicted for 5-Aza-CdR, miR-21, and miR-27b were

**TABLE 2** | Top 50 miRNAs associated with 5-Fluorouracil were predicted by GISMMA based on Dataset 1.

miRNA	Evidence	miRNA	Evidence
hsa-mir-21	26198104	hsa-mir-22	25449431
hsa-mir-324	unconfirmed	hsa-mir-409	unconfirmed
hsa-mir-24-1	26198104	hsa-mir-337	unconfirmed
hsa-mir-27b	26198104	hsa-let-7a-3	26198104
hsa-mir-23a	26198104	hsa-let-7a-2	26198104
hsa-mir-638	26198104	hsa-mir-155	28347920
hsa-mir-27a	26198104	hsa-mir-181b-2	unconfirmed
hsa-let-7b	25789066	hsa-mir-181b-1	unconfirmed
hsa-mir-181a-1	unconfirmed	hsa-mir-15b	26198104
hsa-mir-126	26062749	hsa-let-7i	unconfirmed
hsa-mir-125b-2	unconfirmed	hsa-mir-320a	26198104
hsa-mir-125b-1	unconfirmed	hsa-mir-26a-2	unconfirmed
hsa-mir-124-3	unconfirmed	hsa-mir-328	unconfirmed
hsa-mir-124-2	unconfirmed	hsa-mir-16-2	26198104
hsa-mir-124-1	unconfirmed	hsa-let-7e	26198104
hsa-let-7a-1	26198104	hsa-mir-34b	unconfirmed
hsa-mir-181a-2	24462870	hsa-mir-145	24447928
hsa-mir-24-2	26198104	hsa-mir-200b	26198104
hsa-mir-17	26198104	hsa-let-7c	25951903
hsa-mir-26a-1	unconfirmed	hsa-mir-874	27221209
hsa-mir-16-1	26198104	hsa-mir-650	unconfirmed
hsa-mir-518c	unconfirmed	hsa-mir-501	26198104
hsa-mir-99b	unconfirmed	hsa-mir-500a	unconfirmed
hsa-mir-18a	26198104	hsa-mir-1226	26198104
hsa-mir-663a	26198104	hsa-mir-200c	26198104

The top 1-25 miRNAs are shown in the first column while the top 26-50 in the second. As a result, 8 and 30 out of top 10 and top 50 were confirmed by the known experimental literatures, respectively.

**TABLE 3** | Top 50 miRNAs associated with 17β-Estradiol were predicted by GISMMA based on Dataset 1.

miRNA	Evidence	miRNA	Evidence
hsa-mir-21	26198104	hsa-mir-222	24601884
hsa-mir-324	unconfirmed	hsa-mir-31	23143558
hsa-mir-24-1	unconfirmed	hsa-mir-125a	21914226
hsa-mir-27b	26198104	hsa-mir-663a	26198104
hsa-mir-23a	26198104	hsa-mir-22	24715036
hsa-mir-638	26198104	hsa-mir-132	26282993
hsa-mir-27a	26198104	hsa-mir-501	unconfirmed
hsa-mir-181a-1	unconfirmed	hsa-mir-1226	unconfirmed
hsa-mir-24-2	unconfirmed	hsa-mir-328	unconfirmed
hsa-mir-125b-2	unconfirmed	hsa-mir-155	23568502
hsa-mir-125b-1	unconfirmed	hsa-let-7a-3	26198104
hsa-mir-16-1	unconfirmed	hsa-let-7a-2	26198104
hsa-mir-124-3	26198104	hsa-mir-181b-2	unconfirmed
hsa-mir-124-2	26198104	hsa-mir-181b-1	unconfirmed
hsa-mir-124-1	26198104	hsa-mir-26a-2	unconfirmed
hsa-mir-18a	24245576	hsa-mir-15b	26198104
hsa-let-7b	26198104	hsa-mir-20a	21914226
hsa-mir-181a-2	unconfirmed	hsa-mir-29a	22334722
hsa-let-7a-1	26198104	hsa-mir-19a	unconfirmed
hsa-mir-17	26198104	hsa-mir-200b	26198104
hsa-mir-126	26198104	hsa-mir-221	21057537
hsa-mir-26a-1	unconfirmed	hsa-mir-518c	26198104
hsa-mir-320a	27965096	hsa-mir-194-2	unconfirmed
hsa-mir-16-2	unconfirmed	hsa-mir-181d	unconfirmed
hsa-mir-99b	unconfirmed	hsa-mir-197	unconfirmed

The top 1-25 miRNAs are shown in the first column while the top 26-50 in the second. As a result, 5 and 29 out of top 10 and top 50 were confirmed by the known databases or experimental literatures, respectively.



**TABLE 4 |** Top 50 miRNAs associated with 5-Aza-2'-deoxycytidine were predicted by GISMMA based on Dataset 1.

miRNA	Evidence	miRNA	Evidence
hsa-mir-21	26198104	hsa-mir-518c	unconfirmed
hsa-mir-324	unconfirmed	hsa-mir-200b	23626803
hsa-mir-23a	unconfirmed	hsa-let-7d	26802971
hsa-mir-24-1	26198104	hsa-mir-501	unconfirmed
hsa-mir-27b	26198104	hsa-mir-1226	unconfirmed
hsa-mir-27a	26198104	hsa-mir-200c	23626803
hsa-mir-638	26198104	hsa-mir-99b	unconfirmed
hsa-let-7a-1	unconfirmed	hsa-mir-181a-2	26198104
hsa-mir-124-3	23200812	hsa-let-7e	22053057
hsa-mir-124-2	23200812	hsa-mir-132	unconfirmed
hsa-mir-124-1	unconfirmed	hsa-mir-203a	26577858
hsa-let-7b	26708866	hsa-mir-409	unconfirmed
hsa-mir-18a	unconfirmed	hsa-mir-337	unconfirmed
hsa-mir-24-2	26198104	hsa-mir-1915	unconfirmed
hsa-mir-17	26198104	hsa-mir-128-2	unconfirmed
hsa-mir-181a-1	26198104	hsa-mir-128-1	unconfirmed
hsa-mir-663a	unconfirmed	hsa-mir-320a	26198104
hsa-let-7a-3	26227220	hsa-mir-181b-2	unconfirmed
hsa-let-7a-2	unconfirmed	hsa-mir-181b-1	unconfirmed
hsa-mir-126	26198104	hsa-mir-222	unconfirmed
hsa-mir-26a-1	unconfirmed	hsa-mir-26a-2	unconfirmed
hsa-mir-15b	unconfirmed	hsa-mir-328	unconfirmed
hsa-mir-16-1	26198104	hsa-mir-16-2	26198104
hsa-mir-125b-2	26198104	hsa-mir-29a	26198104
hsa-mir-125b-1	26198104	hsa-let-7c	unconfirmed

The top 1–25 miRNAs are shown in the first column while the top 26–50 in the second. As a result, 7 and 25 out of top 10 and top 50 were confirmed by the known databases or experimental literatures, respectively.

ranked in the first and fifth position, respectively, both of which showed significant down regulation after 5-Aza-CdR treatment in breast cancer cells (Radpour et al., 2011). Moreover, miR-24-1 was the fourth miRNA predicted to be associated with 5-Aza-CdR. Microarray analysis showed miR-24-1 were up regulated upon 5-Aza-CdR therapy in pancreatic cancer PANC-1 cells compared to control cells (Lee et al., 2009).

The whole prediction list of all candidate small molecule-miRNA pairs in Dataset 1 was provided in **Supplementary Table 1**, which was ranked in a descending order according to the association scores resulted from GISMMA. It is hoped that the ranked list can be useful in guiding biological experiments, and can be verified by more experimental results in the future.

## DISCUSSION

This paper presented a graphlet interaction based method GISMMA to infer the potential associations between small molecules and miRNAs by combining small molecule similarity, miRNA similarity and known associations between small molecules and miRNAs. In GISMMA, we used a similarity network to represent the small molecules and used another similarity network to represent the miRNAs. An edge with a

weight of the similarity value between two nodes was plotted when there was similarity between the two nodes, otherwise not. We utilized graphlet interaction to measure the complex relationship between two nodes in the network, where the graphlet is defined as a type of non-isomorphic subgraph (Wang et al., 2014). Then, we counted each graphlet interaction isomer in a special pattern from the node having known associations to the node which does not have known associations. Therefore, we obtained a vector to describe the graphlet interaction between the two nodes. The correlation score between a small molecule and a miRNA can be computed through summing the weighted graphlet interaction isomers, where the weights can be learnt from the known associations. The performance of GISMMA on predicting novel small molecule-miRNA associations was evaluated with four validation approaches that were global and two types of local LOOCV, as well as five-fold cross validation. The cross validation results were compared between GISMMA and SMiR-NBI, which showed the superior performance of GISMMA over SMiR-NBI. Besides, the ROC curves of SMiR-NBI are some unusual in **Figures 2, 3**, which may be attribute to that SMiR-NBI could not predict associated miRNAs (small molecules) for new small molecules (miRNAs). When ranking the test small molecule-miRNA pair with those candidate pairs for SMiR-NBI, we assigned fixed rank to those pairs that contain new small molecules (miRNAs) with an average number, which may cause the presence of line segments in the ROC curve. We have implemented cross validations on two datasets with different sizes. The results showed that GISMMA performed better on Dataset 1 than on Dataset 2, which could be resulted from two factors. The one is the more similarity information in Dataset 1. The other is that Dataset 1 contains those small molecules and miRNAs without any known associations, which often get lower association scores and lower rankings than the test sample. This could also make the AUCs higher. And we further executed case study for three small molecules using Dataset 1. The numbers of miRNAs that were validated to be related to these three small molecules by experimental literatures are in turn 30, 29 and 25 in top 50 miRNAs predicted by GISMMA. Via cross validations together with case study, we can see that GISMMA is well-performed and reliable in predicting new associations between small molecules and miRNAs. Furthermore, a list of all predicted small molecule-miRNA associations was provided, which would be favorable for the development of miRNA-targeted therapy and drug reposition. In detail, for a specific small molecule, we focused on the predicted miRNAs that are most possibly associated with this small molecule. These miRNAs might be related to some diseases that were not confirmed to be treated by this small molecule. Through regulating the expressions of these miRNAs, this small molecule could be used for the treatment of these diseases. Therefore, we believed that the prediction results of this work could offer some guidance for the experiment of drug reposition to some extent.

The outstanding performance of GISMMA can be attributed to several factors. Firstly, we mapped the similarity between small molecules and similarity between miRNAs into two networks, in which the similarity values were fully exploited to investigate the complex relationship between two nodes by measuring their

graphlet interaction. Secondly, in GISMMA, not only direct but also indirect links were considered between the nodes in the counting of graphlet interaction isomers. Finally, the GISMMA is a bipartite method which combines miRNA network with small molecule network. It can be used to predict miRNAs associated with new small molecules without any known related miRNAs, as well as to predict small molecules associated with new miRNAs without any known related small molecules, because it computes the association score by combining the result calculated in the small molecule network with that in the miRNA network.

However, GISMMA still has some limitations. For example, the lack of the known association data, especially the presence of many new small molecules or new miRNAs that have no known associations, affected the performance to a large extent. It can be expected that the model will obtain better performance when more experimental datasets are produced in the future. Besides, the simple algorithm of averaging the scores from two networks to compute the final association score may cause bias to those pairs that can be predicted only in one network. Furthermore, GISMMA considered 4 nodes at most within a graphlet, which hindered it to contain more similarity information from more distant nodes. Finally, this model cannot be applied to the prediction of the association in which the small molecule and the miRNA are both new. We anticipate that more network-based methods could be developed to improve the prediction of novel small molecule-miRNA association. For example, Petri nets based models have been proved to be a useful tool for many prediction problems, inspired by the work in (Russo et al., 2017), we could construct algorithm using

Petri nets for the inference of potential small molecule-miRNA association.

## AUTHOR CONTRIBUTIONS

N-NG implemented the experiments, analyzed the result, and wrote the paper. Y-ZS analyzed the result and wrote the paper. XC conceived the project, developed the prediction method, designed the experiments, analyzed the result, and revised the paper. ZM analyzed the result. J-QL analyzed the result and revised the paper. All authors read and approved the final manuscript.

## FUNDING

XC was supported by National Natural Science Foundation of China under grant nos. 61772531 and 11631014. JL was supported by National Natural Science Foundation of China under grant nos. U1713212 and 61572330, Natural Science foundation of Guangdong Province under grant no. 2014A030313554, and Technology Planning Project from Guangdong Province under grant no. 2014B010118005.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphar.2018.01152/full#supplementary-material>

## REFERENCES

- Airley, R. (2009). Cancer chemotherapy: basic science to the clinic. *ChemMedChem* 14, 2123–2124.
- Ambros, V. (2003). MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. *Cell* 113, 673–676. doi: 10.1016/S0092-8674(03)00428-8
- Andreuccetti, M., Allegrini, G., Antonuzzo, A., Malvaldi, G., Conte, P. F., Danesi, R., et al. (1996). Azidothymidine in combination with 5-fluorouracil in human colorectal cell lines: in vitro synergistic cytotoxicity and DNA-induced strand-breaks. *Eur. J. Cancer* 32A, 1219–1226. doi: 10.1016/0959-8049(96)00018-4
- Bhat-Nakshatri, P., Wang, G., Collins, N. R., Thomson, M. J., Geistlinger, T. R., Carroll, J. S., et al. (2009). Estradiol-regulated microRNAs control estradiol response in breast cancer cells. *Nucleic Acids Res.* 37, 4850–4861. doi: 10.1093/nar/gkp500
- Chen, X., Sun, Y.-Z., Zhang, D.-H., Li, J.-Q., Yan, G.-Y., An, J.-Y., et al. (2017a). NRDTD: a database for clinically or experimentally supported non-coding RNAs and drug targets associations. *Database* 2017:bax057. doi: 10.1093/database/bax057
- Chen, X., Xie, D., Zhao, Q., and You, Z.-H. (2017b). MicroRNAs and complex diseases: from experimental results to computational models. *Brief. Bioinform.* doi: 10.1093/bib/bbx130 [Epub ahead of print].
- Chen, X., Wang, L., Qu, J., Guan, N.-N., and Li, J.-Q. (2018a). Predicting miRNA-disease association based on inductive matrix completion. *Bioinformatics* doi: 10.1093/bioinformatics/bty503 [Epub ahead of print].
- Chen, X., Xie, D., Wang, L., Zhao, Q., You, Z.-H., and Liu, H. (2018b). BNPMDA: bipartite network projection for MiRNA-disease association prediction. *Bioinformatics* 34, 3178–3186. doi: 10.1093/bioinformatics/bty333
- Chen, X., Yin, J., Qu, J., and Huang, L. (2018c). MDHGI: matrix decomposition and heterogeneous graph inference for miRNA-disease association prediction. *PLoS Comput. Biol.* 14:e1006418. doi: 10.1371/journal.pcbi.1006418
- Chen, X., Yan, C. C., Zhang, X., Zhang, X., Dai, F., Yin, J., et al. (2015). Drug-target interaction prediction: databases, web servers and computational models. *Brief. Bioinform.* 17, 696–712. doi: 10.1093/bib/bbv066
- Esquela-Kerscher, A., and Slack, F. J. (2006). Oncomirs - microRNAs with a role in cancer. *Nat. Rev. Cancer* 6, 259–269. doi: 10.1038/nrc1840
- Gottlieb, A., Stein, G. Y., Ruppin, E., and Sharan, R. (2011). PREDICT: a method for inferring novel drug indications with application to personalized medicine. *Mol. Syst. Biol.* 7:496. doi: 10.1038/msb.2011.26
- Gumireddy, K., Young, D. D., Xiong, X., Hogenesch, J. B., Huang, Q., and Deiters, A. (2008). Small-molecule inhibitors of microRNA miR-21 function. *Angew. Chem. Int. Ed. Engl.* 47, 7482–7484. doi: 10.1002/anie.200801555
- Hattori, M., Okuno, Y., Goto, S., and Kanehisa, M. (2003). Development of a chemical structure comparison method for integrated analysis of chemical and genomic information in the metabolic pathways. *J. Am. Chem. Soc.* 125, 11853–11865. doi: 10.1021/ja036030u
- He, L., and Hannon, G. J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* 5, 522–531. doi: 10.1038/nrg1379
- Hernandez-Vargas, H., Ballestar, E., Carmona-Saez, P., Von Kobbe, C., Banon-Rodriguez, I., Esteller, M., et al. (2006). Transcriptional profiling of MCF7 breast cancer cells in response to 5-Fluorouracil: relationship with cell cycle changes and apoptosis, and identification of novel targets of p53. *Int. J. Cancer* 119, 1164–1175.
- Hurtubise, A., and Momparler, R. L. (2004). Evaluation of antineoplastic action of 5-aza-2'-deoxycytidine (Dacogen) and docetaxel (Taxotere) on human breast, lung and prostate carcinoma cell lines. *Anticancer Drugs* 15, 161–167. doi: 10.1002/ijc.21938

- Issa, J. P., Garcia-Manero, G., Giles, F. J., Mannari, R., Thomas, D., Faderl, S., et al. (2004). Phase 1 study of low-dose prolonged exposure schedules of the hypomethylating agent 5-aza-2'-deoxycytidine (decitabine) in hematopoietic malignancies. *Blood* 103, 1635–1640. doi: 10.1182/blood-2003-03-0687
- Jiang, Q., Wang, Y., Hao, Y., Juan, L., Teng, M., Zhang, X., et al. (2009). miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Res.* 37, D98–D104. doi: 10.1093/nar/gkn714
- Jiang, W., Chen, X., Liao, M., Li, W., Lian, B., Wang, L., et al. (2012). Identification of links between small molecules and miRNAs in human cancers based on transcriptional responses. *Sci. Rep.* 2:282. doi: 10.1038/srep00282
- Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., et al. (2011). DrugBank 3.0: a comprehensive resource for 'Omics' research on drugs. *Nucleic Acids Res.* 39, D1035–D1041. doi: 10.1093/nar/gkq1126
- Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 42, D68–D73. doi: 10.1093/nar/gkt1181
- Lee, K. H., Lotterman, C., Karikari, C., Omura, N., Feldmann, G., Habbe, N., et al. (2009). Epigenetic silencing of MicroRNA miR-107 regulates cyclin-dependent kinase 6 expression in pancreatic cancer. *Pancreatology* 9, 293–301. doi: 10.1159/000186051
- Li, J., Lei, K., Wu, Z., Li, W., Liu, G., Liu, J., et al. (2016). Network-based identification of microRNAs as potential pharmacogenomic biomarkers for anticancer drugs. *Oncotarget* 7, 45584–45596. doi: 10.18632/oncotarget.10052
- Lim, D., Byun, W. G., Koo, J. Y., Park, H., and Park, S. B. (2016). Discovery of a small-molecule inhibitor of protein-microRNA interaction using binding assay with a site-specifically labeled Lin28. *J. Am. Chem. Soc.* doi: 10.1021/jacs.6b06965 [Epub ahead of print].
- Liu, X., Wang, S., Meng, F., Wang, J., Zhang, Y., Dai, E., et al. (2013). SM2miR: a database of the experimentally validated small molecules' effects on microRNA expression. *Bioinformatics* 29, 409–411. doi: 10.1093/bioinformatics/bts698
- Longley, D. B., Harkin, D. P., and Johnston, P. G. (2003). 5-fluorouracil: mechanisms of action and clinical strategies. *Nat. Rev. Cancer* 3, 330–338. doi: 10.1038/nrc1074
- Lu, M., Zhang, Q., Min, D., Jing, M., Guo, Y., Wei, G., et al. (2008). An analysis of human MicroRNA and disease associations. *PLoS One* 3:e3420. doi: 10.1371/journal.pone.0003420
- Lv, S., Li, Y., Wang, Q., Ning, S., Huang, T., Wang, P., et al. (2012). A novel method to quantify gene set functional association based on gene ontology. *J. R. Soc. Interface* 9, 1063–1072. doi: 10.1098/rsif.2011.0551
- Lv, Y., Wang, S., Meng, F., Yang, L., Wang, Z., Wang, J., et al. (2015). Identifying novel associations between small molecules and miRNAs based on integrated molecular networks. *Bioinformatics* 31, 3638–3644. doi: 10.1093/bioinformatics/btv417
- Matsui, M., and Corey, D. R. (2017). Non-coding RNAs as drug targets. *Nat. Rev. Drug Discov.* 16, 167–179. doi: 10.1038/nrd.2016.117
- McCabe, M. T., Low, J. A., Daignault, S., Imperiale, M. J., Wojno, K. J., and Day, M. L. (2006). Inhibition of DNA methyltransferase activity prevents tumorigenesis in a mouse model of prostate cancer. *Cancer Res.* 66, 385–392. doi: 10.1158/0008-5472.CAN-05-2020
- Meng, F., Dai, E., Yu, X., Zhang, Y., Chen, X., Liu, X., et al. (2014). Constructing and characterizing a bioactive small molecule and microRNA association network for Alzheimer's disease. *J. R. Soc. Interface* 11:20131057. doi: 10.1098/rsif.2013.1057
- Monroig, P. D., and Calin, G. A. (2013). MicroRNA and Epigenetics: diagnostic and therapeutic opportunities. *Curr. Pathobiol. Rep.* 1, 43–52. doi: 10.1007/s40139-013-0008-9
- Monroig Pdel, C., Chen, L., Zhang, S., and Calin, G. A. (2015). Small molecule compounds targeting miRNAs for cancer therapy. *Adv. Drug Deliv. Rev.* 81, 104–116. doi: 10.1016/j.addr.2014.09.002
- Patra, S. K., and Bettuzzi, S. (2009). Epigenetic DNA-(cytosine-5-carbon) modifications: 5-aza-2'-deoxycytidine and DNA-demethylation. *Biochemistry* 74, 613–619.
- Przulj, N. (2007). Biological network comparison using graphlet degree distribution. *Bioinformatics* 23, e177–e183. doi: 10.1093/bioinformatics/btl301
- Radpour, R., Berekati, Z., Kohler, C., Schumacher, M. M., Grussenmeyer, T., Jenoe, P., et al. (2011). Integrated epigenetics of human breast cancer: synoptic investigation of targeted genes, microRNAs and proteins upon demethylation treatment. *PLoS One* 6:e27355. doi: 10.1371/journal.pone.0027355
- Rossi, L., Bonmassar, E., and Faraoni, I. (2007). Modification of miR gene expression pattern in human colon cancer cells following exposure to 5-fluorouracil in vitro. *Pharmacol. Res.* 56, 248–253. doi: 10.1016/j.phrs.2007.07.001
- Ruepp, A., Kowarsch, A., Schmid, D., Buggenthin, F., Brauner, B., Dunger, I., et al. (2010). PhenoMiR: a knowledgebase for microRNA expression in diseases and biological processes. *Genome Biol.* 11:R6. doi: 10.1186/gb-2010-11-1-r6
- Russo, G., Pennisi, M., Boscarino, R., and Pappalardo, F. (2017). Continuous Petri Nets and microRNA analysis in melanoma. *IEEE/ACM Trans. Comput. Biol. Bioinform.* doi: 10.1109/TCBB.2017.2733529 [Epub ahead of print].
- Ryan, K. J. (1982). Biochemistry of aromatase: significance to female reproductive physiology. *Cancer Res.* 42, 3342s–3344s.
- Sathish Kumar, B., Kumar, A., Singh, J., Hasanain, M., Singh, A., Fatima, K., et al. (2014). Synthesis of 2-alkoxy and 2-benzyloxy analogues of estradiol as anti-breast cancer agents through microtubule stabilization. *Eur. J. Med. Chem.* 86, 740–751. doi: 10.1016/j.ejmech.2014.09.033
- Selcuklu, S. D., Donoghue, M. T., Kerin, M. J., and Spillane, C. (2012). Regulatory interplay between miR-21, JAG1 and 17beta-estradiol (E2) in breast cancer cells. *Biochem. Biophys. Res. Commun.* 423, 234–239. doi: 10.1016/j.bbrc.2012.05.074
- Shah, M. Y., Pan, X., Fix, L. N., Farwell, M. A., and Zhang, B. (2011). 5-Fluorouracil drug alters the microRNA expression profiles in MCF-7 breast cancer cells. *J. Cell. Physiol.* 226, 1868–1878. doi: 10.1002/jcp.22517
- Simpson, E., and Santen, R. J. (2015). Celebrating 75 years of oestradiol. *J. Mol. Endocrinol.* 55, T1–T20. doi: 10.1530/JME-15-0128
- Spizzo, R., Nicoloso, M. S., Croce, C. M., and Calin, G. A. (2009). SnapShot: microRNAs in cancer. *Cell* 137, 586.e1–586.e1. doi: 10.1016/j.cell.2009.04.040
- Thomas, M., and Deiters, A. (2013). MicroRNA miR-122 as a therapeutic target for oligonucleotides and small molecules. *Curr. Med. Chem.* 20, 3629–3640. doi: 10.2174/0929867311320290009
- Tilghman, S. L., Bratton, M. R., Segar, H. C., Martin, E. C., Rhodes, L. V., Li, M., et al. (2012). Endocrine disruptor regulation of microRNA expression in breast carcinoma cells. *PLoS One* 7:e32754. doi: 10.1371/journal.pone.0032754
- Wang, J., Meng, F., Dai, E., Yang, F., Wang, S., Chen, X., et al. (2016). Identification of associations between small molecule drugs and miRNAs based on functional similarity. *Oncotarget* 7, 38658–38669. doi: 10.18632/oncotarget.9577
- Wang, X. D., Huang, J. L., Yang, L., Wei, D. Q., Qi, Y. X., and Jiang, Z. L. (2014). Identification of human disease genes from interactome network using graphlet interaction. *PLoS One* 9:e86142. doi: 10.1371/journal.pone.0086142
- Wang, Y., and Lee, C. G. (2009). MicroRNA and cancer—focus on apoptosis. *J. Cell Mol. Med.* 13, 12–23. doi: 10.1111/j.1582-4934.2008.00510.x
- Wang, Y., Xiao, J., Suzek, T. O., Zhang, J., Wang, J., and Bryant, S. H. (2009). PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res.* 37, W623–W633. doi: 10.1093/nar/gkp456
- Windle, R., Bell, P. R., and Shaw, D. (1987). Five year results of a randomized trial of adjuvant 5-fluorouracil and levamisole in colorectal cancer. *Br. J. Surg.* 74, 569–572. doi: 10.1002/bjs.1800740707
- Xia, T., Li, J., Cheng, H., Zhang, C., and Zhang, Y. (2015). Small-molecule regulators of MicroRNAs in biomedicine. *Drug Dev. Res.* 76, 375–381. doi: 10.1002/ddr.21271
- Zhang, J., Yang, P. L., and Gray, N. S. (2009). Targeting cancer with small molecule kinase inhibitors. *Nat. Rev. Cancer* 9, 28–39. doi: 10.1038/nrc2559
- Zhang, S., Chen, L., Jung, E. J., and Calin, G. A. (2010). Targeting microRNAs with small molecules: from dream to reality. *Clin. Pharmacol. Ther.* 87, 754–758. doi: 10.1038/clpt.2010.46
- Zhou, J., Zhou, Y., Yin, B., Hao, W., Zhao, L., Ju, W., et al. (2010). 5-Fluorouracil and oxaliplatin modify the expression profiles of microRNAs in human colon cancer cells in vitro. *Oncol. Rep.* 23, 121–128.

**Conflict of Interest Statement:** 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.

Copyright © 2018 Guan, Sun, Ming, Li 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.