



# TLHNMDA: Triple Layer Heterogeneous Network Based Inference for MiRNA-Disease Association Prediction

Xing Chen\*, Jia Qu and Jun Yin

School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, China

## OPEN ACCESS

### Edited by:

Xiaogang Wu,  
Institute for Systems Biology,  
United States

### Reviewed by:

Chi Zhang,  
University of Nebraska-Lincoln,  
United States

Georges Nemer,  
American University of Beirut,  
Lebanon

Edwin Wang,  
University of Calgary, Canada  
Zhaohui Steve Qin,  
Emory University, United States

### \*Correspondence:

Xing Chen  
xingchen@amss.ac.cn

### Specialty section:

This article was submitted to  
Systems Biology,  
a section of the journal  
Frontiers in Genetics

**Received:** 10 October 2017

**Accepted:** 12 June 2018

**Published:** 03 July 2018

### Citation:

Chen X, Qu J and Yin J (2018)  
TLHNMDA: Triple Layer  
Heterogeneous Network Based  
Inference for MiRNA-Disease  
Association Prediction.  
Front. Genet. 9:234.  
doi: 10.3389/fgene.2018.00234

In recent years, microRNAs (miRNAs) have been confirmed to be involved in many important biological processes and associated with various kinds of human complex diseases. Therefore, predicting potential associations between miRNAs and diseases with the huge number of verified heterogeneous biological datasets will provide a new perspective for disease therapy. In this article, we developed a novel computational model of Triple Layer Heterogeneous Network based inference for MiRNA-Disease Association prediction (TLHNMDA) by using the experimentally verified miRNA-disease associations, miRNA-long noncoding RNA (lncRNA) interactions, miRNA function similarity information, disease semantic similarity information and Gaussian interaction profile kernel similarity for lncRNAs into an triple layer heterogeneous network to predict new miRNA-disease associations. As a result, the AUCs of TLHNMDA are 0.8795 and  $0.8795 \pm 0.0010$  based on leave-one-out cross validation (LOOCV) and 5-fold cross validation, respectively. Furthermore, TLHNMDA was implemented on three complex human diseases to evaluate predictive ability. As a result, 84% (kidney neoplasms), 78% (lymphoma) and 76% (prostate neoplasms) of top 50 predicted miRNAs for the three complex diseases can be verified by biological experiments. In addition, based on the HMDD v1.0 database, 98% of top 50 potential esophageal neoplasms-associated miRNAs were confirmed by experimental reports. It is expected that TLHNMDA could be a useful model to predict potential miRNA-disease associations with high prediction accuracy and stability.

**Keywords:** microRNA, disease, association prediction, computational prediction model, triple layer heterogeneous network

## INTRODUCTION

According to the central law of molecular biology, genetic information was found to be stored in protein-coding genes (Crick et al., 1961). Recent studies have revealed that up to 70% of the human genome is transcribed into RNA, whereas protein-coding genes only make up less than 2% of total genome (Djebali et al., 2012). The majority of the human genome is made up of non-coding RNAs (ncRNAs) (Derrien et al., 2012). Based on whether transcript lengths are larger than 200 nucleotides, ncRNAs can be further divided into small ncRNA and long ncRNA (lncRNA) (Kapranov et al., 2007; Guttman et al., 2013). MicroRNAs (miRNAs) are endogenous non-coding RNAs (~22 nt) that bind to the 3'-untranslated regions (3'-UTRs)

of their target RNAs (mRNAs) and control the expression of gene (Ganju et al., 2017). MiRNAs could also serve as positive regulators (Jopling et al., 2005; Vasudevan et al., 2007). Sufficient evidences indicated that thousands of miRNAs have associations with many critical biological processes (Lu et al., 2008), such as cell proliferation (Cheng et al., 2005), development (Karp and Ambros, 2005), metabolism (Alshalalfa and Alhajj, 2013), aging (Bartel, 2009), transduction (Cui et al., 2006), viral infection (Miska, 2005), and so on. Some researchers also founded that allogeneic T cell responses are regulated by miRNAs (Sun et al., 2013). It also has been shown that by attenuating shared miRNAs, competing endogenous RNAs (ceRNAs) could crosstalk and regulate each other, which is essential for regulating many biological functions (Yuan et al., 2016). Moreover, miRNA34s might be key effectors of p53 tumor-suppressor function, and their inactivation might contribute to certain cancers (Bommer et al., 2007). Recently, experiments further showed that special class of 5'-capped pre-miRNAs have been identified in both *C. elegans* and mouse, this promotes the understanding of the transcriptional regulation of miRNA genes themselves (Chen et al., 2017a). Therefore, it is no wonder that miRNAs are closely connected with diverse human cancer types, including breast neoplasms, lung neoplasms, colon neoplasms, kidney neoplasms, lymphoma, etc. (Pasquier and Gardès, 2016). For example, studies have implicated that miR-16-1 and miR-15a could cause chromosomal translocations in patients with chronic lymphocytic leukemia (CLL) (Calin et al., 2002). Experiments further shown that miRNAs may be a new target for the molecular targeted therapy of various cancers (Guzzi et al., 2015; Chen et al., 2017b). Thus, the identification of disease-associated miRNAs can provided a new viewpoint with the respect to the diagnosis, prevention and treatment of human complex diseases in the field of medicine (Chen, 2016). However, using the traditional biological methods to identify miRNA-disease associations is usually time-consuming and expensive. Therefore, more and more scholars have focused on developing efficient computational models to predict potential miRNA-disease associations by integrating various experimentally validated datasets. Database HMDD and miR2Disease (Jiang et al., 2009; Li et al., 2014c) have been constructed to collect the associations between human miRNAs and diseases based on previous biological experiments.

According to the assumption that functionally similar miRNAs tend to be associated with phenotypically similar diseases (Lu et al., 2008; Bandyopadhyay et al., 2010), several computational approaches have been established to infer the new miRNA-disease associations. Mork et al. (2014) introduced a computational model, named miRPD. They identified potential miRNA-disease associations by systematic combination of known miRNA-protein associations with known protein-disease associations. Shi et al. (2013) established a computational framework on the basis of the assumption that miRNAs whose target genes are associated with specific diseases are more possible to be related to these diseases. They constructed protein-protein interaction (PPI) networks and implemented random walk on the network to calculate the probability scores

of each miRNA-disease pair. Xu et al. (2011) introduced an approach to infer novel human miRNA-disease associations by combining computational target prediction with expression profiles of miRNA and mRNA in tumor and nontumor tissues. In the model, the probability scores of each miRNA-disease pair could be converted into the functional similarity calculation between miRNA targets and known diseases-related genes. More importantly, the model could be a useful tool for miRNA-disease association prediction without relying on the known miRNA-disease associations. Jiang et al. (2010) proposed a computational model on the basic of hypergeometric distribution to predict new disease-associated miRNA by systematic integration of miRNA functional similarity network, disease phenotype similarity network, and experimentally verified disease-miRNA association network. However, less than 40 percent of the molecular for human disease is known and the dataset of miRNA-target interactions used in the above studies were not highly accurate, which may limit the application of the method mentioned above.

Researchers have also proposed other methods without relying on the dataset of miRNA-target interactions. For example, Chen et al. (2012b) developed the method of Random Walk with Restart for MiRNA-Disease Association (RWRMDA) to identify new disease-associated miRNAs by applying a similarity-based RWR on miRNA functional similarity network. Xuan et al. (2015) proposed the method of MIRNAs associated with Diseases Prediction (MIDP) to predict new miRNAs candidates using random walk. In which they built a miRNA network derived from miRNA-associated diseases by integration of the nodes similarities, nodes prior information and their local topological structure. Then, the potential association between a disease and a miRNA could be inferred until the iterative walking process on the network converged. Xuan et al. (2013) further proposed an effective computational approach of HDMP by comprehensive integration of miRNA functional similarity and the distribution of miRNAs associated with the disease in the  $k$  most similar neighbors to obtain scores of new miRNAs-disease associations. Li et al. (2017) developed Matrix Completion for MiRNA-Disease Association prediction (MCMDA), a reliable computational method in which they updated scores of each pair using matrix completion algorithm. The model is of high efficiency to update the low-rank miRNA-disease association matrix. Chen and Yan (2014) reported a method named Regularized Least Squares for MiRNA-Disease Association prediction (RLSMDA) on the basis of miRNA functional similarity, disease semantic similarity and known human miRNA-disease associations using a semi-supervised classifier. Recently, Chen et al. (2016a) introduced the model of Within and Between Score for MiRNA-Disease Association prediction (WBSMDA) by combination of integrated similarity and known miRNA-disease associations. The model built two prediction functions from the perspective of disease and miRNA according to the idea that functionally similar miRNAs tend to be associated with similar diseases, and combined them to calculate the association probability of each miRNA-disease pair. Chen et al. (2016b) further developed Heterogeneous Graph Inference for MiRNA-Disease Association prediction

(HGIMDA), a new approach in which they constructed a heterogeneous network on the basis of miRNA functional similarity, disease semantic similarity, known miRNA-disease associations and an iterative update equation that propagates information across the heterogeneous network were established to infer new disease-associated miRNAs. A deep ensemble miRNA-disease association prediction (DeepMDA) framework was also introduced by Fu and Peng (2017) to identify potential miRNA-disease associations using a three-layer neural network classifier based on high-level features extracted from miRNA and disease similarity. Moreover, some other computing models for the identification of miRNA-disease associations were also gradually proposed, such as Liu et al. (2016) predicted miRNA-disease associations by implementing random walk on a heterogeneous network with multiple data sources. Zou et al. (2015) introduced two computational methods of KATZ and CATAPULT to make prediction for miRNA-disease pairs based on social network analysis methods. Pallez et al. (2017) presented a predictive approach named MiRAI using an evolutionary tuned latent semantic analysis. Pasquier and Gardès (2016) make prediction for miRNA-disease associations with a vector space model.

As mentioned above, an integration strategy may provide more comprehensive and accurate information to predict disease-related miRNAs. Actually, miRNA dysregulation is related to many human diseases through many factors, including, for example, miRNA-mRNA interactions, miRNA-lncRNA interactions, miRNA-protein interactions and so on. The miRNAs involved in genes, coding RNAs, and proteins have been used widely in other computational model for the identification of miRNA-disease associations (Shi et al., 2013) (Mork et al., 2014). In this paper, considering many experimentally verified miRNA-lncRNA interactions have been confirmed by recent biological experiments (Li et al., 2014a), we introduced the model of Triple Layer Heterogeneous Network based inference for MiRNA-Disease Association prediction (TLHNMDA) to identify the potential biological links between miRNAs and diseases by integrating multi-level data regarding miRNAs, diseases, lncRNAs and their association information into a triple layer heterogeneous network. We implemented leave-one out cross validation (LOOCV) and 5-fold cross validation on the TLHNMDA to evaluate its performance. The AUCs of LOOCV were respectively 0.8795, and the model obtained the average AUC of  $0.8795 \pm 0.0010$  on 5-fold cross validation. Then, case studies of kidney neoplasms, prostate neoplasms and lymphoma were implemented to assess the independent prediction performance of the model. As a result, 42, 38, and 39 out of top 50 potential miRNAs for these three important diseases were confirmed in dbDEMOC (Yang et al., 2010) and miR2Disease (Jiang et al., 2009) database, respectively. We further tested TLHNMDA on the database HMDD v1.0 (Lu et al., 2008) to see whether the TLHNMDA still performs well. Taking esophageal neoplasms as an example, as a result, 49 of the top 50 esophageal neoplasms-associated miRNAs were verified by experimental reports. It has proved that TLHNMDA is reliable and effective in predicting potential disease-associated miRNAs.

## MATERIALS AND METHODS

### Human miRNA-Disease Association

In this paper, the known dataset of human miRNA-disease associations were downloaded from HMDD v2.0 database. The dataset contains 383 diseases, 495 miRNAs and 5430 high-quality experimentally verified human miRNA-diseases associations. Furthermore, an adjacency matrix  $A$  was established to denote known miRNAs-disease associations. The row of the matrix represents the disease, and the column represents the miRNAs. We used the variables  $nm$  and  $nd$  to represent the number of miRNAs and diseases in the dataset, respectively. The value of  $A(d(i), m(j))$  is 1 when miRNA  $m(i)$  is associated with disease  $d(j)$ , otherwise 0.

### miRNA-lncRNA Interactions

The dataset of miRNA-lncRNA interactions can be obtained from starBase v2.0 database (Li et al., 2014a), which provided the most comprehensive experimentally confirmed miRNA-lncRNA interactions. The dataset consists of 10112 known miRNA-lncRNA interactions about 132 miRNAs and 1114 lncRNAs. In addition, the known lncRNAs-related miRNAs that do not appear in the dataset of known miRNA-disease associations mentioned above is deleted. As a result, 9088 miRNA-lncRNA interactions were obtained. We also constructed an adjacency matrix  $B$  to represent known miRNA-lncRNA interactions. The row of the  $B$  represents the miRNAs, and the column represents the lncRNAs. The variable  $nl$  represents the number of lncRNA in the dataset. If miRNA  $m(i)$  is interacted with lncRNA  $l(j)$ , the value of  $B(m(i), l(j))$  in the  $B$  is 1, otherwise 0.

### miRNA Functional Similarity

Wang et al. (2010) introduced a computational method of miRNA functional similarity between a miRNA pair ( $m_i$  and  $m_j$ ). The whole process of the computational method can be divided into four steps. First, we need to identify the diseases set  $D(m_i)$  (diseases related to  $m_i$ ) and  $D(m_j)$  (diseases related to  $m_j$ ) for miRNA  $m_i$  and  $m_j$ , respectively. Second, in both sets, the semantic values of all diseases are calculated according to the corresponding DAG. Third, the semantic similarity for each disease pairs between  $D(m_i)$  and  $D(m_j)$  can be computed by consideration of their semantic value. In the last step, the functional similarity between  $m_i$  and  $m_j$  is calculated in the light of the semantic similarity obtained in step three. From <http://www.cuilab.cn/files/images/cuilab/misim.zip>, miRNA functional similarity probability scores can be downloaded. Similarly, we built matrix  $FS$  to stand for the miRNA functional similarity matrix, where  $FS(m(i), m(j))$  is the functional similarity probability score between miRNA  $m(i)$  and  $m(j)$ .

### Disease Semantic Similarity Model 1

Each disease can be described as a Directed Acyclic Graph (DAG). For example, disease  $D$  can be denoted as  $DAG(D) = (D, T(D), E(D))$ , where  $T(D)$  is a set of node  $D$  itself and its ancestor nodes,  $E(D)$  stands for the edges between parent and child nodes (Wang et al., 2010). Therefore, the semantic value of disease  $D$

could be calculated as follows:

$$DV1(D) = \sum_{d \in T(D)} D_{D1}(d) \quad (1)$$

where

$$\begin{cases} D_{D1}(d) = 1 & \text{if } d = D \\ D_{D1}(d) = \max \left\{ \Delta^* D_{D1}(d') \mid d' \in \text{children of } d \right\} & \text{if } d \neq D \end{cases} \quad (2)$$

$\Delta$  is the semantic contribution factor. For disease  $D$ , the contribution of itself to the semantic value of disease  $D$  is 1. If the distance between  $D$  and  $d$  increases, the semantic contribution value of disease  $d$  to the  $D$  will decrease. Thus, if diseases in the same layer, they would have the same contribution to the semantic value of disease  $D$ . The value of semantic similarity in disease semantic similarity model 1 between disease  $d(i)$  and  $d(j)$  can be defined as follows:

$$SS1(d(i), d(j)) = \frac{\sum_{t \in T(d(i)) \cap T(d(j))} (D_{d(i)1}(t) + D_{d(j)1}(t))}{DV1(d(i)) + DV1(d(j))} \quad (3)$$

## Disease Semantic Similarity Model 2

In the disease semantic similarity model 2, considering different disease terms in the same layer of DAG( $D$ ) may appear in different numbers of disease DAGs, disease with more specific which appears in less disease DAGs should contribute to the semantic similarity of disease  $D$  at a higher contribution level. Therefore, the contribution of disease  $d$  to the semantic value of disease  $D$  can be calculated as follows:

$$SM(m(i), m(j)) = \begin{cases} \frac{KM(m(i), m(j)) + FS(m(i), m(j))}{KM(m(i), m(j))^2} & m(i) \text{ and } m(j) \text{ has functional similarity} \\ KM(m(i), m(j)) & \text{otherwise} \end{cases} \quad (12)$$

$$SD(d(i), d(v)) = \begin{cases} \frac{KD(d(u), d(v)) + SS(d(u), d(v))}{KD(d(u), d(v))^2} & d(u) \text{ and } d(v) \text{ has similarity} \\ KD(d(u), d(v)) & \text{otherwise} \end{cases} \quad (13)$$

$$D_{D2}(d) = -\log \left[ \frac{\text{the number of DAGs including } d}{\text{the number of diseases}} \right] \quad (4)$$

In disease semantic similarity model 2, the value of semantic similarity between  $d(i)$  and  $d(j)$  can be defined as follows:

$$SS2(d(i), d(j)) = \frac{\sum_{t \in T(d(i)) \cap T(d(j))} (D_{d(i)2}(t) + D_{d(j)2}(t))}{DV2(d(i)) + DV2(d(j))} \quad (5)$$

## Gaussian Interaction Profile Kernel Similarity

Gaussian interaction profile kernel similarity for diseases can be defined based on the known miRNA-disease associations dataset by considering the assumption that similar diseases tend to be

related with more common miRNAs. In this paper, the binary vector  $IP(d(u))$  is the  $u$ th row of matrix  $A$ , which was used to indicate the interaction profiles between disease  $d(u)$  and each miRNA. Therefore, the value of Gaussian interaction profile kernel similarity between diseases  $d(u)$  and  $d(v)$  is defined as follows.

$$KD(d(u), d(v)) = \exp(-\gamma_d \|IP(d(u)) - IP(d(v))\|^2) \quad (6)$$

where parameter  $\gamma_d$  is used to control the kernel bandwidth, which can be obtained from the normalization of a new bandwidth  $\gamma'_d$  by the average number of associated miRNAs for all the diseases.

$$\gamma_d = \frac{\gamma'_d}{\left(\frac{1}{nd} \sum_{n=1}^{nd} \|IP(d(u))\|^2\right)} \quad (7)$$

Similarly, we defined the value of Gaussian interaction profile kernel similarity between miRNA  $m(i)$  and  $m(j)$  as follows:

$$KM(m(i), m(j)) = \exp(-\gamma_m \|IP(m(i)) - IP(m(j))\|^2) \quad (8)$$

$$\gamma_m = \frac{\gamma'_m}{\left(\frac{1}{nm} \sum_{n=1}^{nm} \|IP(m(i))\|^2\right)} \quad (9)$$

Gaussian interaction profile kernel similarity for lncRNA  $l(i)$  and  $l(j)$  can also be calculated as follows:

$$KL(l(i), l(j)) = \exp(-\gamma_l \|IP(l(i)) - IP(l(j))\|^2) \quad (10)$$

$$\gamma_l = \frac{\gamma'_l}{\left(\frac{1}{nl} \sum_{n=1}^{nl} \|IP(l(i))\|^2\right)} \quad (11)$$

## Integrated Similarity for miRNAs and Diseases

Here, integrated miRNA similarity matrix  $SM$  are defined on the basis of miRNA functional similarity and Gaussian interaction profile kernel similarity for miRNAs. Integrated disease similarity matrix  $SD$  are constructed according to disease semantic similarity and Gaussian interaction profile kernel similarity for diseases.

where

$$SS = \frac{SS1 + SS2}{2} \quad (14)$$

## TLHMADA

According to the guilt-by-association principle (Barabási et al., 2011), new miRNA-disease associations can be inferred through

existing associations between similar miRNAs and similar diseases, Likewise, novel miRNA-lncRNA interactions can be inferred through existing associations between similar miRNA and lncRNA (see **Figure 1**). We infer new miRNA-lncRNA associations in the newly proposed triple layer heterogeneous network by using an information flow-based method. New disease-lncRNA association matrix  $W_{dl}^{new}$  could be constructed as follows:

$$W_{dl}^{new} = W_{dm} \times SM \times W_{ml} \quad (15)$$

As shown in the above formula, we can identify potential disease-lncRNA associations on the basis of miRNA-disease associations  $W_{dm}$ , miRNA-lncRNA interactions  $W_{ml}$  as well as integrated similarity for miRNAs  $SM$  according to the equation. Once the associations between diseases and lncRNAs are established. New association  $W_{dm}^{new}$  between diseases and miRNAs can be defined by considering these associations:

$$W_{dm}^{new} = W_{dl} \times KL \times W_{ml}^T \quad (16)$$

Equation (16) is potentially more powerful in capturing miRNA-disease associations by incorporating lncRNA information into miRNA-disease prediction. As a by-product from the model, we can also obtain a new interaction between each miRNA and lncRNA pair by incorporating miRNA-disease associations  $W_{dm}$ , disease-lncRNA associations  $W_{dl}$  as well as integrated similarity

for miRNAs  $SD$ . New association  $W_{ml}^{new}$  between miRNAs and lncRNAs can be defined as follows:

$$W_{ml}^{new} = W_{dm}^T \times SD \times W_{dl} \quad (17)$$

where the superscript T indicates the transpose of the corresponding matrix.

We treat  $W_{dl}$  as a temporary value, and replace  $W_{dl}$  in the two Equations (16, 17) using the Equation (15), respectively.

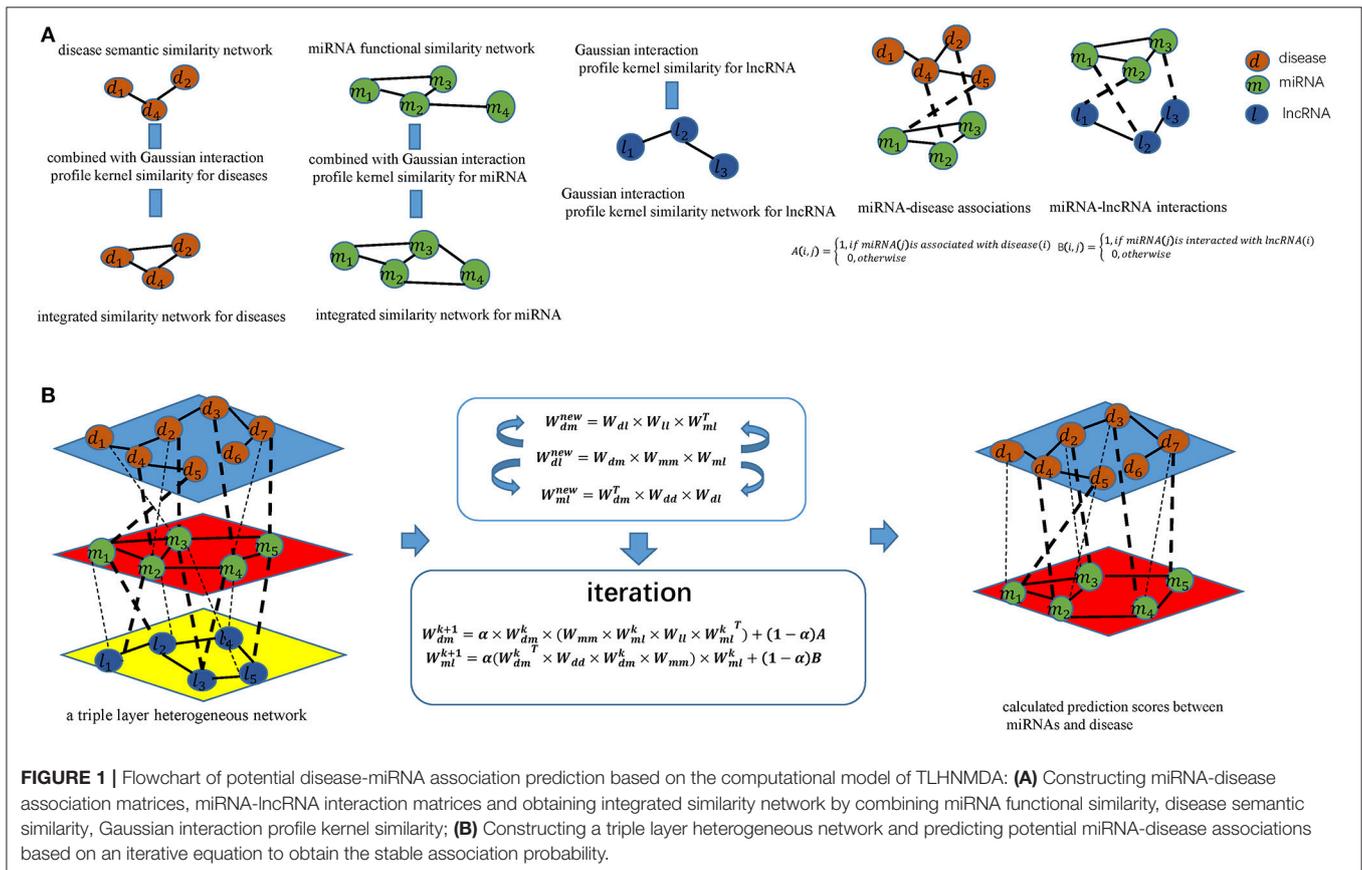
$$W_{dm}^{new} = W_{dm} \times SM \times W_{ml} \times KL \times W_{ml}^T \quad (18)$$

$$W_{ml}^{new} = W_{dm}^T \times SD \times W_{dm} \times SM \times W_{ml} \quad (19)$$

Once the new miRNA-disease associations  $W_{dm}^{new}$  and new miRNA-lncRNA interactions  $W_{ml}^{new}$  were obtained, we established iterative updating procedure based on Equations (18, 19). The final computational model can be written as follows:

$$W_{dm}^{k+1} = \alpha \times W_{dm}^k \times (SM \times W_{ml}^k \times KL \times W_{ml}^{kT}) + (1 - \alpha) A \quad (20)$$

$$W_{ml}^{k+1} = \alpha (W_{dm}^{kT} \times SD \times W_{dm}^k \times SM) \times W_{ml}^k + (1 - \alpha) B \quad (21)$$



**FIGURE 1 |** Flowchart of potential disease-miRNA association prediction based on the computational model of TLHNMDA: **(A)** Constructing miRNA-disease association matrices, miRNA-lncRNA interaction matrices and obtaining integrated similarity network by combining miRNA functional similarity, disease semantic similarity, Gaussian interaction profile kernel similarity; **(B)** Constructing a triple layer heterogeneous network and predicting potential miRNA-disease associations based on an iterative equation to obtain the stable association probability.

Here  $\alpha$  a decay factor in the range of (0,1).  $A$  and  $B$  represents the initial disease-miRNA associations and miRNA-lncRNA interactions, respectively.  $W_{dm}^k$  and  $W_{ml}^k$  would be converge with proper normalization utilizing Equations (24, 25), respectively (Wang et al., 2013c) (the proof can be found in the Supplementary Materials).

$$IM = SM \times W_{ml}^k \times KL \times W_{ml}^{kT} \quad (22)$$

$$ID = W_{dm}^{kT} \times SD \times W_{dm}^k \times SM \quad (23)$$

$$IM(m(i), m(j)) = \frac{IM(m(i), m(j))}{\sqrt{\sum_{l=1}^{nm} IM(m(i), m(l))} \sqrt{\sum_{l=1}^{nm} IM(m(j), m(l))}} \quad (24)$$

$$ID(d(i), d(j)) = \frac{ID(d(i), d(j))}{\sqrt{\sum_{l=1}^{nd} ID(d(i), d(l))} \sqrt{\sum_{l=1}^{nd} ID(d(j), d(l))}} \quad (25)$$

After some steps, the iteration will be stable after some steps (the change in value between  $W_{dm}^{k+1}$  and  $W_{dm}^k$  measured by L1 norm is less than a given cutoff, the cutoff in this paper was  $10^{-6}$ ).

The three-layer model is proposed by incorporating miRNA-lncRNA information into miRNA-disease association prediction based on miRNA dysregulation is associated with many human complex diseases may through miRNA-lncRNA interactions. It can be seen from the two iterative algorithms, once new association between miRNA and disease is estimated, it can be used to update other miRNA-disease associations and miRNA-lncRNA interactions. Similarly, once new association between miRNA and lncRNA is estimated, it can also be used to update other miRNA-disease associations and miRNA-lncRNA interactions. Therefore, the layer between miRNA and disease and the layer between miRNA and lncRNA play the same important role in the triple layer heterogeneous network to propagate information for the identification of potential miRNA-disease associations and miRNA-lncRNA interactions simultaneously. In order to make the two constructed iterative equations to work effectively, known miRNA-disease associations and known miRNA-lncRNA interactions as weights were added to the inferred equations because the initial links deserve more credibility. At last,  $W_{dm}^{new}$  and  $W_{ml}^{new}$  were expected to converge, which means that the propagation of information would be stable at the end.

## RESULTS

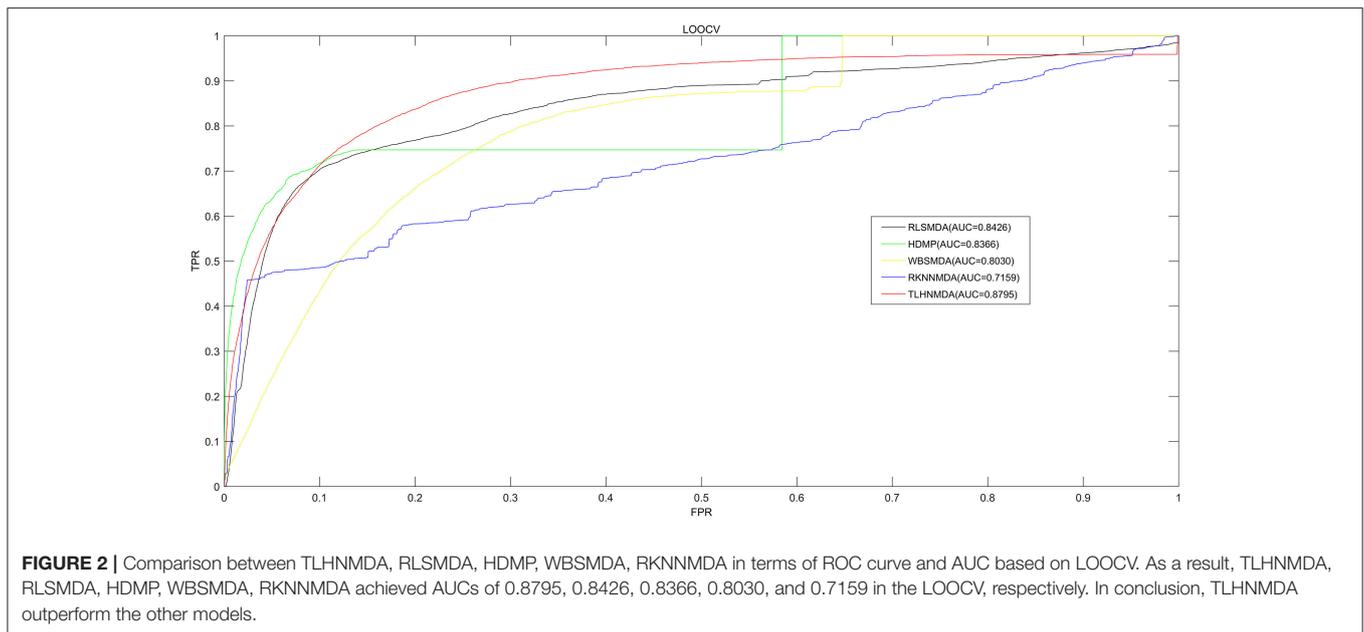
### Performance Evaluation

We implemented LOOCV as well as 5-fold cross validation on the basis of the experimentally verified miRNA-disease associations in HMDD v2.0 database (Li et al., 2014c) to evaluate the prediction performance of TLHNMDA. Moreover, TLHNMDA

were compared with four previous classical computational methods: RLSMDA (Chen et al., 2012b), HDMP (Xuan et al., 2013), WBSMDA (Xu et al., 2011), RKNNMDA (Chen et al., 2017c). In the framework of LOOCV evaluation, each known association of miRNA-disease pair in the database was considered as test samples in turn, the other known miRNA-disease associations were considered as training samples, the miRNA-disease pairs with no known verified associations were regarded as candidate samples. After TLHNMDA was implemented, we would obtain the scores of the test samples and the scores of the candidate samples, and then the score of the test sample was compared with the scores of all the candidate samples in LOOCV. While in 5-fold cross validation, the experimentally verified miRNA-disease associations were evenly divided into five disjoint parts. One part was selected as test samples and the other four parts were regarded as training samples in each time. Similarly, the miRNA-disease pairs without known association evidences were regarded as candidate samples. Then, the score of each test sample was compared with the scores of all the candidate samples. It is worth noting that the above process was repeated 100 times, we would get 100 rankings for all miRNA and disease pairs. It is worth noting that almost all the models for the prediction of miRNA-disease associations according to the assumption that miRNAs with similar functions tend to be related to phenotypically similar diseases were proposed based on the LOOCV and 5-fold cross validation (Mork et al., 2014; Xuan et al., 2015; You et al., 2017; Zhong et al., 2017). At last, we drew Receiver Operating Characteristics (ROC) curve using true positive rate (TPR, sensitivity) against the false positive rate (FPR, 1-specificity) at different thresholds evaluate the performance of TLHNMDA clearly. Sensitivity refers to the percentage of the positive miRNA-disease associations whose score ranks are higher than the preset threshold, while specificity refers to the percentage of negative miRNA-disease pairs with ranks lower than the threshold. Then, the value of Area under the ROC curve (AUC) could be calculated to evaluate the prediction performance of the model. If the value of AUC is 1, it tells us the approach possesses perfect prediction performance; if the value of AUC is 0.5, it stands for the method possesses random prediction performance. For LOOCV, TLHNMDA, RLSMDA, HDMP, WBSMDA, RKNNMDA obtained AUCs of 0.8795, 0.8426, 0.8366, 0.8030 and 0.7159, respectively (see **Figure 2**). For 5-fold, TLHNMDA, RLSMDA, HDMP, WBSMDA, RKNNMDA obtained the average AUCs and corresponding standard deviations of  $0.8795 \pm 0.0010$ ,  $0.8569 \pm 0.0020$ ,  $0.8342 \pm 0.0010$ ,  $0.8185 \pm 0.0009$ , and  $0.6723 \pm 0.0027$ , respectively.

### Case Studies

Here, to evaluate the prediction accuracy of TLHNMDA, case studies were implemented on kidney neoplasms, lymphoma and prostate neoplasms. In the model, the 5430 known miRNA-disease associations in HMDD v2.0 were utilized as the training set. All candidate miRNAs for each interested disease were ranked in accordance with their predicted scores. After that, the top 50 predicted miRNAs were picked out and verified in other two important miRNA-disease association databases (i.e., dbDEMC and miR2Disease). Furthermore, the results



showed that 232 of the 5430 known miRNA-disease associations in HMDD v2.0 also existed in miR2Disease and 546 known associations also existed in dbDEMCC. It is noteworthy that there was no overlap between the training samples and the prediction lists. That is because only candidate miRNAs (miRNAs have any no known associations with interested disease in HMDD v2.0) for interested disease were ranked and verified in case studies. Accordingly, none of the top 50 predicted miRNAs existed in HMDD v2.0 and the verification of miRNAs in the prediction lists was completely independent of HMDD v2.0.

Kidney neoplasms, known as renal cancer, is a common health problem in cancer diseases (Manojlovi et al., 1986). The age of its incidence can be in all ages, particularly in the age between 50 and 70 years old (Nickerson et al., 2002). The most common symptoms of kidney neoplasms patients are pains in the lumbar and hematuria (Duque et al., 1998). Many existing treatments of kidney neoplasms are usually radiation therapy and chemotherapy drugs, which do not have much effect in the cure (Zbar et al., 2003). Up to now, lots of miRNAs have been reported to be associated with kidney neoplasms. For example, miRNA-192, miRNA-194, miRNA-215, miRNA-200c, and miRNA-141 were proved to be associated with renal childhood neoplasms (Senanayake et al., 2012). MiRNA-210 was reported to be upregulated in renal neoplasms (Eilertsen et al., 2014). Another miRNA named miRNA-23b could act as an oncogene and reducing the expression of miRNA-23b would be an effective way to inhibit the growth of kidney tumor, which might contribute to the treatment of renal neoplasms in medicine (Liu et al., 2010). In case studies, we implemented TLHNMDA on kidney neoplasms to predict the potential miRNA-disease associations. In short, 8 of the top 10 and 42 of the top 50 novel identified miRNAs associated with kidney neoplasms were validated by the two database deDEMCC and miR2Disease (see **Table 1**).

Lymphoma is the fastest growing human tumor (Chen et al., 2013), which is a group of blood cell tumors develop from lymphocytes (a type of white blood cell). The disease consists of two categories: Hodgkin lymphomas (HL) and the non-Hodgkin lymphomas (NHL) (Mcduffie et al., 2009). Many lymphoma-related miRNAs have been reported based on recent biological experiments. For example, the expression of miRNA-150 was confirmed to be a tumor suppressor in malignant lymphoma (Watanabe et al., 2011), which induces the differentiation of EBV-positive Burkitt lymphoma differentiation based on the modulation of c-Myb *in vitro* (Li et al., 2014a). In addition, miR-21 could regulate cell activity of proliferation, invasion, and apoptosis. Accordingly, it has a potential therapeutic application in lymphoma (Sekar et al., 2014). We implemented TLHNMDA on lymphoma to predict the top 10 and top 50 related miRNAs. Briefly speaking, 7 of top 10 and 39 of top 50 potential lymphoma-related miRNAs were verified in the deDEMCC and miR2Disease database (see **Table 2**).

Prostate neoplasms is the most common disease in men (Siegel et al., 2013). The malignant tumor originates from prostate in the epithelial cells (Gmyrek et al., 2001). In the recent years, many miRNAs have been verified to be related with prostate neoplasms base on accumulating researches. For instance, miR-141, miR-375, miR-21, miR-93, miR-106a, miR-874, miR-1207, and miR-26a were reported to upregulate in prostate neoplasms (Xiao et al., 2012; Chu et al., 2014; Dong et al., 2015). We also implemented TLHNMDA on prostate neoplasms to identify the related miRNAs. As a result, 7 of top 10 and 38 of top 50 potential Prostate neoplasms-miRNAs were confirmed in the deDEMCC and miR2Disease database (see **Table 3**).

Moreover, we further implemented TLHNMDA on the known miRNA-disease associations in HMDD v1.0 database (Lu et al., 2008) to see whether the approach worked properly on a different dataset. Consequently, the predicted scores for

**TABLE 1** | Prediction of the top 50 predicted miRNAs associated with kidney neoplasms based on known associations in HMDD v2.0 database.

miRNA	Evidence	miRNA	Evidence
hsa-mir-16	dbDEMC	hsa-mir-20a	dbDEMC miR2Disease
hsa-mir-15b	dbDEMC	hsa-mir-539	unconfirmed
hsa-mir-195	dbDEMC	hsa-mir-26a	dbDEMC miR2Disease
hsa-mir-424	dbDEMC miR2Disease	hsa-mir-27b	dbDEMC
hsa-mir-497	dbDEMC	hsa-mir-34a	dbDEMC
hsa-mir-103a	unconfirmed	hsa-mir-17	miR2Disease
hsa-mir-485	unconfirmed	hsa-mir-29b	dbDEMC miR2Disease
hsa-mir-23a	dbDEMC	hsa-mir-125b	unconfirmed
hsa-mir-214	dbDEMC miR2Disease	hsa-mir-143	dbDEMC
hsa-mir-155	dbDEMC	hsa-mir-128	dbDEMC
hsa-mir-107	dbDEMC	hsa-mir-320a	unconfirmed
hsa-mir-590	unconfirmed	hsa-mir-708	unconfirmed
hsa-mir-19a	dbDEMC	hsa-mir-124	dbDEMC
hsa-mir-125a	dbDEMC	hsa-mir-149	dbDEMC
hsa-mir-142	unconfirmed	hsa-mir-199a	dbDEMC miR2Disease
hsa-mir-19b	dbDEMC miR2Disease	hsa-mir-34c	dbDEMC
hsa-mir-138	dbDEMC	hsa-mir-181a	dbDEMC
hsa-mir-26b	dbDEMC	hsa-mir-152	dbDEMC
hsa-mir-150	dbDEMC miR2Disease	hsa-mir-106a	dbDEMC miR2Disease
hsa-mir-29c	dbDEMC miR2Disease	hsa-mir-18a	dbDEMC
hsa-mir-370	dbDEMC	hsa-mir-181b	dbDEMC
hsa-mir-31	dbDEMC	hsa-mir-193a	dbDEMC
hsa-mir-185	dbDEMC miR2Disease	hsa-mir-7	dbDEMC miR2Disease
hsa-mir-24	dbDEMC	hsa-mir-122	dbDEMC miR2Disease
hsa-mir-29a	dbDEMC miR2Disease	hsa-mir-106b	dbDEMC miR2Disease

The first column records top 1–25 related miRNAs. The second column records the top 26–50 related miRNAs.

**TABLE 2** | Prediction of the top 50 predicted miRNAs associated with lymphoma based on known associations in HMDD v2.0 database.

miRNA	Evidence	miRNA	Evidence
hsa-mir-15b	dbDEMC	hsa-mir-199a	dbDEMC
hsa-mir-195	dbDEMC	hsa-mir-34c	unconfirmed
hsa-mir-424	dbDEMC	hsa-mir-152	dbDEMC
hsa-mir-497	dbDEMC	hsa-mir-106a	dbDEMC miR2Disease
hsa-mir-103a	unconfirmed	hsa-mir-181b	dbDEMC
hsa-mir-485	unconfirmed	hsa-mir-193a	unconfirmed
hsa-mir-23a	dbDEMC	hsa-mir-7	dbDEMC
hsa-mir-214	dbDEMC	hsa-mir-106b	dbDEMC
hsa-mir-107	dbDEMC	hsa-mir-22	dbDEMC
hsa-mir-590	unconfirmed	hsa-mir-27a	dbDEMC
hsa-mir-142	unconfirmed	hsa-mir-144	unconfirmed
hsa-mir-26b	dbDEMC	hsa-mir-326	dbDEMC
hsa-mir-370	unconfirmed	hsa-mir-93	dbDEMC
hsa-mir-31	dbDEMC	hsa-mir-186	dbDEMC
hsa-mir-185	dbDEMC	hsa-mir-30a	dbDEMC
hsa-mir-23b	dbDEMC	hsa-mir-148a	dbDEMC
hsa-mir-29a	dbDEMC	hsa-mir-182	dbDEMC
hsa-mir-27b	dbDEMC	hsa-mir-199b	dbDEMC
hsa-mir-34a	dbDEMC	hsa-mir-145	dbDEMC miR2Disease
hsa-mir-29b	dbDEMC	hsa-mir-328	dbDEMC miR2Disease
hsa-mir-125b	unconfirmed	hsa-mir-330	dbDEMC
hsa-mir-143	dbDEMC miR2Disease	hsa-mir-421	unconfirmed
hsa-mir-128	dbDEMC	hsa-mir-1	dbDEMC
hsa-mir-320a	unconfirmed	hsa-mir-181c	dbDEMC
hsa-mir-149	dbDEMC miR2Disease	hsa-mir-141	dbDEMC

The first column records top 1–25 related miRNAs. The second column records the top 26–50 related miRNAs.

candidate miRNAs showed that 10 of top 10 and 49 of top 50 potential esophageal neoplasms-associated miRNAs were verified by three databases (see **Table 4**). Lastly, we list the potential miRNAs related to all the human diseases and the association scores of the entire ranking results obtained by the computational model of TLHNMDA (see Supplementary Table 1).

## DISCUSSION

Although progress has been made in the discovery of miRNA, the role of miRNAs in physiologic and pathophysiologic processes is just emerging. MiRNAs as governors of gene expression during cardiovascular development and disease have associations with many critical biological processes (Liu and Olson, 2010). Identification of miRNAs expressed in specific cardiac cell types may provide us with new diagnostic, prognostic, and therapeutic targets for many forms of cardiovascular disease (Cordes and Srivastava, 2009). Furthermore, aberrant expression of miRNAs has also been involved in various neurological disorders (NDs) of the central nervous system such as alzheimer disease, parkinson's disease, huntington disease, amyotrophic lateral sclerosis, schizophrenia and autism. If dysregulated miRNAs are found in patients with NDs, this may also be a biomarker

for the earlier diagnosis and monitoring of disease progression. Identifying the role of miRNAs in normal cellular processes is critical in the development of new therapeutic strategies for NDs (Kamal et al., 2015). Therefore, predicting disease-associated miRNAs is important for the understanding of disease pathogenesis and treatment of a variety of clinically important disease. In this paper, according to the hypothesis that functional similar miRNAs and lncRNAs are likely to be associated with similar diseases. We introduced a novel model, named TLHNMDA, which constructed a triple layer heterogeneous network by systematic combination of miRNA functional similarity, disease semantic similarity, Gaussian interaction profile kernel similarity, known miRNA-disease associations and miRNA-lncRNA interactions to identify new disease-associated miRNAs. In the model, an iterative updating algorithm that propagates information across the network was proposed based on the triple layer heterogeneous graph to obtain final prediction scores between diseases and miRNAs. The experimental results from LOOCV and 5-fold cross validation have demonstrated that TLHNMDA outperforms other four computational methods. What's more, case studies of four human diseases: kidney neoplasms, lymphoma, prostate neoplasms and esophageal neoplasms were implemented and the results were verified by the

**TABLE 3** | Prediction of the top 50 predicted miRNAs associated with prostate neoplasms based on known associations in HMDD v2.0 database.

miRNA	Evidence	miRNA	Evidence
hsa-mir-15a	dbDEMC miR2Disease	hsa-mir-24	dbDEMC miR2Disease
hsa-mir-16	dbDEMC miR2Disease	hsa-mir-29a	dbDEMC miR2Disease
hsa-mir-15b	dbDEMC	hsa-mir-539	unconfirmed
hsa-mir-195	dbDEMC miR2Disease	hsa-mir-20a	miR2Disease
hsa-mir-424	unconfirmed	hsa-mir-26a	dbDEMC miR2Disease
hsa-mir-497	miR2Disease	hsa-mir-34a	dbDEMC miR2Disease
hsa-mir-103a	unconfirmed	hsa-mir-27b	dbDEMC miR2Disease
hsa-mir-485	unconfirmed	hsa-mir-29b	dbDEMC miR2Disease
hsa-mir-23a	dbDEMC miR2Disease	hsa-mir-17	miR2Disease
hsa-mir-214	dbDEMC miR2Disease	hsa-mir-143	dbDEMC miR2Disease
hsa-mir-155	dbDEMC	hsa-mir-128	dbDEMC
hsa-mir-107	unconfirmed	hsa-mir-320a	unconfirmed
hsa-mir-590	unconfirmed	hsa-mir-708	unconfirmed
hsa-mir-19a	dbDEMC	hsa-mir-124	dbDEMC
hsa-mir-125a	dbDEMC miR2Disease	hsa-mir-149	dbDEMC miR2Disease
hsa-mir-142	unconfirmed	hsa-mir-199a	dbDEMC miR2Disease
hsa-mir-19b	dbDEMC miR2Disease	hsa-mir-34c	dbDEMC
hsa-mir-138	dbDEMC	hsa-mir-181a	dbDEMC miR2Disease
hsa-mir-26b	dbDEMC miR2Disease	hsa-mir-152	dbDEMC
hsa-mir-150	dbDEMC	hsa-mir-18a	unconfirmed
hsa-mir-370	miR2Disease	hsa-mir-21	dbDEMC miR2Disease
hsa-mir-29c	dbDEMC	hsa-mir-106a	dbDEMC miR2Disease
hsa-mir-31	dbDEMC miR2Disease	hsa-mir-181b	dbDEMC miR2Disease
hsa-mir-185	unconfirmed	hsa-mir-193a	unconfirmed
hsa-mir-23b	dbDEMC miR2Disease	hsa-mir-7	dbDEMC

The first column records top 1–25 related miRNAs. The second column records the top 26–50 related miRNAs.

**TABLE 4** | Prediction of the top 50 predicted miRNAs associated with esophageal neoplasms based on HMDD v1.0 database.

miRNA	Evidence	miRNA	Evidence
hsa-mir-15a	dbDEMC and HMDD	hsa-mir-143	dbDEMC and HMDD
hsa-mir-16	dbDEMC	hsa-mir-29a	dbDEMC
hsa-mir-15b	dbDEMC	hsa-mir-125b	dbDEMC
hsa-mir-195	dbDEMC	hsa-mir-29b	dbDEMC
hsa-mir-424	dbDEMC	hsa-mir-181b	dbDEMC
hsa-mir-497	dbDEMC	hsa-mir-34a	dbDEMC HMDD
hsa-mir-214	dbDEMC HMDD	hsa-mir-106a	dbDEMC
hsa-mir-107	dbDEMC miR2Disease	hsa-mir-106b	dbDEMC
hsa-mir-155	dbDEMC HMDD	hsa-mir-199a	dbDEMC HMDD
hsa-mir-19a	dbDEMC HMDD	hsa-mir-330	dbDEMC
hsa-mir-19b	dbDEMC	hsa-mir-20b	dbDEMC
hsa-mir-125a	dbDEMC	hsa-mir-26a	dbDEMC HMDD
hsa-mir-185	dbDEMC	hsa-mir-1	dbDEMC
hsa-mir-20a	dbDEMC HMDD	hsa-mir-181a	dbDEMC
hsa-mir-24	dbDEMC	hsa-mir-186	dbDEMC
hsa-mir-17	dbDEMC	hsa-mir-141	dbDEMC HMDD
hsa-mir-23a	dbDEMC	hsa-mir-93	dbDEMC
hsa-mir-26b	dbDEMC	hsa-mir-421	dbDEMC
hsa-mir-539	unconfirmed	hsa-mir-222	dbDEMC
hsa-mir-150	dbDEMC HMDD	hsa-mir-28	dbDEMC HMDD
hsa-mir-23b	dbDEMC	hsa-mir-145	dbDEMC HMDD
hsa-mir-29c	dbDEMC HMDD	hsa-mir-92a	HMDD
hsa-mir-370	dbDEMC	hsa-mir-22	dbDEMC HMDD
hsa-mir-142	dbDEMC	hsa-mir-199b	dbDEMC
hsa-mir-18a	dbDEMC	hsa-mir-34c	dbDEMC HMDD

The first column records top 1–25 related miRNAs. The second column records the top 26–50 related miRNAs.

experimental literatures in dbDEMC and miR2Disease database. We can see that the TLHNMDA turns out to be more reliable and effective in inferring the potential miRNA–disease associations than the previous computational models. Therefore, our model could be an effective and useful computational model to predict new miRNA-disease associations. Biomedical researchers could use TLHNMDA to computationally identify the miRNAs that were potentially related to the investigated diseases.

TLHNMDA could obtain the valid performances due to the following several reasons. Firstly, TLHNMDA improved prediction accuracy and decrease the prediction bias by integration of several reliable types of biological datasets, including the accurate experimentally verified miRNA-disease associations, known miRNA-lncRNA interactions, miRNA functional similarity network, disease semantic similarity network and Gaussian interaction profile kernel similarity. Secondly, the model captured new miRNAs-diseases associations using global network similarity information, it has an advantage over the local network similarity information model to capture miRNA-disease associations. Finally, TLHNMDA is an iterative algorithm to update predicted scores based on global network similarity information until the state is in convergence, which promote the effective prediction of TLHNMDA. However,

several limitations also exist in the TLHNMDA, for example, TLHNMDA cannot predict the new miRNAs associated with the new diseases without any known miRNA-disease associations. Besides, there is no powerful methods to find optimal parameters of TLHNMDA. The selection of parameters in the iterative algorithm is based on past experiences which can't guarantee the model with best state in the implementation process. Finally, the number of miRNA-disease associations and miRNA-lncRNA interactions, confirmed by biological experiments, is still insufficient. Therefore, in the future research, we can have a try to propose a new model by integrating more available biological datasets.

It is noteworthy that there exist many other types of data can also be used to predict miRNA-disease associations, for example, miRNA-mRNA interactions (Li et al., 2014b), miRNA-protein interactions (Shi et al., 2016), miRNA-environmental factors interactions (Chen et al., 2012a), and so on. Considering some existing methods have taken advantage of different datasets to identify miRNA-disease associations, which makes direct comparison of their performance and the performance of the proposed method is not realistic. For example, two model proposed by Pallez et al. (2017) and Pasquier and Gardès (2016) were based on the dataset of miRNA-disease associations,

miRNA-neighbor associations, miRNA-target associations, miRNA-word associations and miRNA-family associations. The model proposed by Mork et al. (2014) was based on the dataset of miRNA-protein associations and protein-disease associations to predict potential miRNA-disease associations. The model introduced by Shi et al. (2013) for the identification of miRNA-disease associations was based on disease-gene association, protein-protein interaction, miRNA-target associations. Moreover, Liu et al. (2016) proposed a new computational to predict unobserved miRNA-disease associations based on disease functional similarity, disease semantic similarity and miRNA similarity. It is worth noting that miRNA similarity in the model was calculated based on miRNA-lncRNA interactions. In addition to datasets, there are different ways in defining relationships among nodes of the same type. For example, in the DeepMDA proposed by Fu and Peng (2017), Gaussian interaction profile kernel similarity for disease was calculated by using three association matrices, the miRNA-disease association matrix, the lncRNA-disease association matrix, and the gene-disease association matrix. MiRNA similarity used in KATZ and CATAPULT introduced by Zou et al. (2015) was calculated by text mining analysis of their phenotype descriptions in the Online Mendelian Inheritance in Man (OMIM) database. Especially, the relative merits of using different measures are worth further study. Network analysis and modeling researches constructed by diverse data were also widely applied in other fields. Some studies modeled cancer cells by constructing and modeling networks for individual clones based on tumor genome sequencing (Wang et al., 2013a). Integrative network modeling has been applied in the modeling of drug resistance for personalized treatment (Wang et al., 2013b). Moreover, Hallmark-specific networks were modeled to better understand key cellular processes, which are involved in cancer development

and progression (Gao et al., 2016). The hallmarks of cancer are one of the most widely acknowledged organizing principles for research on cancer (Wang et al., 2015). Accumulating evidences indicated that there are some associations between cancer hallmarks and genes (Wang et al., 2015). For example, miR-16 obtained the highest score in the case study on kidney neoplasms and the second high score in the case study on prostate neoplasms. APP, ATG12, and ATF2 are the common targets for this miRNA and have been identified to be involved in hallmark of inflammation (Wang et al., 2015). In the future work, we plan to extend the model we proposed into new multi-layer prediction model, one extension is to add more diverse datasets of different types (other than the three discussed here) and more associations to the model, then construct the iterative updating algorithm to identify disease-associated miRNAs.

## AUTHOR CONTRIBUTIONS

XC conceived the project, developed the prediction method, designed the experiments, analyzed the result, and wrote the paper. JQ implemented the experiments, analyzed the result, and wrote the paper. YJ analyzed the result and revised the paper.

## FUNDING

XC was supported by National Natural Science Foundation of China under Grant Nos. 61772531 and 11631014.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2018.00234/full#supplementary-material>

## REFERENCES

- Alshalhafa, M., and Alhaji, R. (2013). Using context-specific effect of miRNAs to identify functional associations between miRNAs and gene signatures. *BMC Bioinform.* 14:S1. doi: 10.1186/1471-2105-14-S12-S1
- Bandyopadhyay, S., Mitra, R., Maulik, U., and Zhang, M. Q. (2010). Development of the human cancer microRNA network. *Silence* 1:6. doi: 10.1186/1758-907X-1-6
- Barabási, A. L., Gulbahce, N., and Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* 12, 56–68. doi: 10.1038/nrg2918
- Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233. doi: 10.1016/j.cell.2009.01.002
- Bommer, G. T., Gerin, I., Feng, Y., Kaczorowski, A. J., Kuick, R., Love, R. E., et al. (2007). p53-mediated activation of miRNA34 candidate tumor-suppressor genes. *Curr. Biol.* 17, 1298–1307. doi: 10.1016/j.cub.2007.06.068
- Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., et al. (2002). Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15524–15529. doi: 10.1073/pnas.242606799
- Chen, J., Zhu, D., and Sun, Y. (2017a). Cap-seq reveals complicated miRNA transcriptional mechanisms in *C. elegans* and mouse. *Quant. Biol.* 5, 352–367. doi: 10.1007/s40484-017-0123-4
- Chen, S., Wang, Z., Dai, X., Pan, J., Ge, J., Han, X., et al. (2013). Re-expression of microRNA-150 induces EBV-positive Burkitt lymphoma differentiation by modulating *c-Myb* *in vitro*. *Cancer Sci.* 104, 826–834. doi: 10.1111/cas.12156
- Chen, X. (2016). miREFRWR: a novel disease-related microRNA-environmental factor interactions prediction method. *Mol. Biosyst.* 12, 624–633. doi: 10.1039/C5MB00697J
- Chen, X., and Yan, G. Y. (2014). Semi-supervised learning for potential human microRNA-disease associations inference. *Sci. Reports* 4:5501. doi: 10.1038/srep05501
- Chen, X., Liu, M. X., and Yan, G. Y. (2012b). RWRMDA: predicting novel human microRNA-disease associations. *Mol. Biosyst.* 8, 2792–2798. doi: 10.1039/c2mb25180a
- Chen, X., Liu, M. X., Cui, Q. H., and Yan, G. Y. (2012a). Prediction of disease-related interactions between microRNAs and environmental factors based on a semi-supervised classifier. *PLoS ONE* 7:e43425. doi: 10.1371/journal.pone.0043425
- Chen, X., Sun, Y. Z., Zhang, D. H., Li, J. Q., Yan, G. Y., An, J. Y., et al. (2017b). NRDTD: a database for clinically or experimentally supported non-coding RNAs and drug targets associations. *Database (Oxford)* 2017:bax057. doi: 10.1093/database/bax057
- Chen, X., Wu, Q. F., and Yan, G. Y. (2017c). RKNMMDA: ranking-based KNN for MiRNA-disease association prediction. *RNA Biol.* 14, 952–962. doi: 10.1080/15476286.2017.1312226

- Chen, X., Yan, C. C., Zhang, X., You, Z. H., Deng, L., Liu, Y., et al. (2016a). WBSMDA: within and between score for miRNA-disease association prediction. *Sci. Reports* 6:21106. doi: 10.1038/srep21106
- Chen, X., Yan, C. C., Zhang, X., You, Z. H., Huang, Y. A., and Yan, G. Y. (2016b). HGIMDA: heterogeneous graph inference for miRNA-disease association prediction. *Oncotarget* 7, 65257–65269. doi: 10.18632/oncotarget.11251
- Cheng, A. M., Byrom, M. W., Shelton, J., and Ford, L. P. (2005). Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. *Nucleic Acids Res.* 33, 1290–1297. doi: 10.1093/nar/gki200
- Chu, M., Chang, Y., Li, P., Guo, Y., Zhang, K., and Gao, W. (2014). Androgen receptor is negatively correlated with the methylation-mediated transcriptional repression of miR-375 in human prostate cancer cells. *Oncol. Reports* 31, 34–40. doi: 10.3892/or.2013.2810
- Cordes, K. R., and Srivastava, D. (2009). MicroRNA regulation of cardiovascular development. *Circ. Res.* 104, 724–732. doi: 10.1161/CIRCRESAHA.108.192872
- Crick, F. H., Barnett, L., Brenner, S., and Watts-Tobin, R. J. (1961). General nature of the genetic code for proteins. *Nature* 192, 1227–1232. doi: 10.1038/1921227a0
- Cui, Q., Yu, Z., Purisima, E. O., and Wang, E. (2006). Principles of microRNA regulation of a human cellular signaling network. *Mol. Sys. Biol.* 2:46. doi: 10.1038/msb4100089
- Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 22, 1775–1789. doi: 10.1101/gr.132159.111
- Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., et al. (2012). Landscape of transcription in human cells. *Nature* 489, 101–108. doi: 10.1038/nature11233
- Dong, B., Shi, Z., Wang, J., Wu, J., Yang, Z., and Fang, K. (2015). IL-6 Inhibits the Targeted Modulation of PDCD4 by miR-21 in Prostate Cancer. *PLoS ONE* 10:e0134366. doi: 10.1371/journal.pone.0134366
- Duque, J. L., Loughlin, K. R., O'leary, M. P., Kumar, S., and Richie, J. P. (1998). Partial nephrectomy: alternative treatment for selected patients with renal cell carcinoma. *Urology* 52, 584–590. doi: 10.1016/S0090-4295(98)00380-X
- Eilertsen, M., Andersen, S., Al-Saad, S., Richardsen, E., Stenvold, H., Hald, S. M., et al. (2014). Positive prognostic impact of miR-210 in non-small cell lung cancer. *Lung Cancer* 83, 272–278. doi: 10.1016/j.lungcan.2013.11.005
- Fu, L., and Peng, Q. (2017). A deep ensemble model to predict miRNA-disease association. *Sci. Rep.* 7:14482. doi: 10.1038/s41598-017-15235-6
- Ganju, A., Khan, S., Hafeez, B. B., Behrman, S. W., Yallapu, M. M., Chauhan, S. C., et al. (2017). miRNA nanotherapeutics for cancer. *Drug Disc. Tod.* 22, 424–432. doi: 10.1016/j.drudis.2016.10.014
- Gao, S., Tibiche, C., Zou, J., Zaman, N., Trifiro, M., O'connor-Mccourt, M., et al. (2016). Identification and construction of combinatory cancer hallmark-based gene signature sets to predict recurrence and chemotherapy benefit in stage II colorectal cancer. *JAMA Oncol.* 2, 37–45. doi: 10.1001/jamaoncol.2015.3413
- Gmyrek, G. A., Walburg, M., Webb, C. P., Yu, H. M., You, X., Vaughan, E. D., et al. (2001). Normal and malignant prostate epithelial cells differ in their response to hepatocyte growth factor/scatter factor. *Am. J. Pathol.* 159, 579–590. doi: 10.1016/S0002-9440(10)61729-4
- Guttman, M., Russell, P., Ingolia, N. T., Weissman, J. S., and Lander, E. S. (2013). Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Cell* 154, 240–251. doi: 10.1016/j.cell.2013.06.009
- Guzzi, P. H., Di Martino M. T., Tagliaferri, P., Tassone, P., and Cannataro, M. (2015). Analysis of miRNA, mRNA, and TF interactions through network-based methods. *Eurasip. J. Bioinform. Syst. Biol.* 2015:4. doi: 10.1186/s13637-015-0023-8
- Jiang, Q., Hao, Y., Wang, G., Juan, L., Zhang, T., Teng, M., et al. (2010). Prioritization of disease microRNAs through a human phenome-microRNAome network. *BMC Syst. Biol.* 4 (Suppl. 1):S2. doi: 10.1186/1752-0509-4-S1-S2
- 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
- Jopling, C. L., Yi, M. K., Lancaster, A. M., Lemon, S. M., and Sarnow, P. (2005). Modulation of Hepatitis C Virus RNA Abundance by a Liver-Specific MicroRNA. *Science* 309, 1577–1581. doi: 10.1126/science.1113329
- Kamal, M. A., Mushtaq, G., and Greig, N. H. (2015). Current update on synopsis of miRNA dysregulation in neurological disorders. *CNS Neurol Disord Drug Targets* 14, 492–501. doi: 10.2174/1871527314666150225143637
- Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., et al. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 316, 1484–1488. doi: 10.1126/science.1138341
- Karp, X., and Ambros, V. (2005). Encountering microRNAs in cell fate signaling. *Science* 310, 1288–1289. doi: 10.1126/science.1121566
- Li, J. H., Liu, S., Zhou, H., Qu, L. H., and Yang, J. H. (2014a). starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* 42, D92–97. doi: 10.1093/nar/gkt1248
- Li, J. Q., Rong, Z. H., Chen, X., Yan, G. Y., and You, Z. H. (2017). MCMDA: Matrix completion for miRNA-disease association prediction. *Oncotarget* 8, 21187–21199. doi: 10.18632/oncotarget.15061
- Li, Y., Liang, C., Wong, K. C., Luo, J., and Zhang, Z. (2014b). Mirsynergy: detecting synergistic miRNA regulatory modules by overlapping neighbourhood expansion. *Bioinformatics* 30, 2627–2635. doi: 10.1093/bioinformatics/btu373
- Li, Y., Qiu, C., Tu, J., Geng, B., Yang, J., Jiang, T., et al. (2014c). HMDD v2.0: a database for experimentally supported human microRNA and disease associations. *Nucleic Acids Res.* 42, 1070–1074. doi: 10.1093/nar/gkt1023
- Liu, N., and Olson, E. N. (2010). MicroRNA regulatory networks in cardiovascular development. *Dev. Cell* 18, 510–525. doi: 10.1016/j.devcel.2010.03.010
- Liu, W., Zabornyk, O., Wang, H., Shiao, Y. H., Nickerson, M. L., Khalil, S., et al. (2010). miR-23b targets proline oxidase, a novel tumor suppressor protein in renal cancer. *Oncogene* 29, 4914–4924. doi: 10.1038/onc.2010.237
- Liu, Y., Zeng, X., He, Z., and Zou, Q. (2016). Inferring microRNA-disease associations by random walk on a heterogeneous network with multiple data sources. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 14, 905–915. doi: 10.1109/TCBB.2016.2550432
- Lu, M., Zhang, Q., Deng, M., Miao, J., Guo, Y., Gao, W., et al. (2008). An analysis of human microRNA and disease associations. *PLoS ONE* 3:e3420. doi: 10.1371/journal.pone.0003420
- Manojlović, D., Elaković, D., and Lj, M. (1986). [Therapeutic value of transcatheter embolization in malignant tumors of the renal parenchyma]. *Srpski Arhiv. Za Celokupno Lekarstvo* 114, 631–637.
- McDuffie, H. H., Pahwa, P., Karunanayake, C. P., Spinelli, J. J., and Dosman, J. A. (2009). Clustering of cancer among families of cases with Hodgkin Lymphoma (HL), Multiple Myeloma (MM), Non-Hodgkin's Lymphoma (NHL), Soft Tissue Sarcoma (STS) and control subjects. *BMC Cancer* 9:70. doi: 10.1186/1471-2407-9-70
- Miska, E. A. (2005). How microRNAs control cell division, differentiation and death. *Curr. Opin. Genet. Develop.* 15, 563–568. doi: 10.1016/j.gde.2005.08.005
- Mork, S., Pletscher-Frankild, S., Pallega, A., Gorodkin, J., and Jensen, L. J. (2014). Protein-driven inference of miRNA-disease associations. *Bioinformatics* 30, 392–397. doi: 10.1093/bioinformatics/btt677
- Nickerson, M. L., Warren, M. B., Toro, J. R., Matrosova, V., Glenn, G., Turner, M. L., et al. (2002). Mutations in a novel gene lead to kidney tumors, lung wall defects, and benign tumors of the hair follicle in patients with the Birt-Hogg-Dube syndrome. *Cancer Cell* 2, 157–164. doi: 10.1016/S1535-6108(02)00104-6
- Pallez, D., Gardès, J., and Pasquier, C. (2017). Prediction of miRNA-disease associations using an evolutionary tuned latent semantic analysis. *Sci. Rep.* 7:10548. doi: 10.1038/s41598-017-10065-y
- Pasquier, C., and Gardès, J. (2016). Prediction of miRNA-disease associations with a vector space model. *Sci. Rep.* 6:27036. doi: 10.1038/srep27036
- Sekar, D., Hairul Islam, V. I., Thiruganasambantham, K., and Saravanan, S. (2014). Relevance of miR-21 in HIV and non-HIV-related lymphomas. *Tum. Biol.* 35, 8387–8393. doi: 10.1007/s13277-014-2068-9
- Senanayake, U., Das, S., Vesely, P., Alzoughbi, W., Fröhlich, L. F., Chowdhury, P., et al. (2012). miR-192, miR-194, miR-215, miR-200c and miR-141 are downregulated and their common target ACVR2B is strongly expressed in renal childhood neoplasms. *Carcinogenesis* 33, 1014–1021. doi: 10.1093/carcin/bgs126
- Shi, H., Xu, J., Zhang, G., Xu, L., Li, C., Wang, L., et al. (2013). Walking the interactome to identify human miRNA-disease associations through the functional link between miRNA targets and disease genes. *BMC Syst. Biol.* 7:101. doi: 10.1186/1752-0509-7-101

- Shi, H., Zhang, G., Zhou, M., Cheng, L., Yang, H., Wang, J., et al. (2016). Integration of multiple genomic and phenotype data to infer novel miRNA-disease associations. *PLoS ONE* 11:e0148521. doi: 10.1371/journal.pone.0148521
- Siegel, R., Ms, D. N. M., and Phd, A. J. D. (2013). Cancer statistics, 2013. *Ca A Cancer J. Clin.* 63, 11–30. doi: 10.3322/caac.21166
- Sun, Y., Tawara, I., Zhao, M., Qin, Z. S., Toubai, T., Mathewson, N., et al. (2013). Allogeneic T cell responses are regulated by a specific miRNA-mRNA network. *J. Clin. Invest.* 123, 4739–4754. doi: 10.1172/JCI70013
- Vasudevan, S., Tong, Y., and Steitz, J. A. (2007). Switching from repression to activation: microRNAs can up-regulate translation. *Science* 318, 1931–1934. doi: 10.1126/science.1149460
- Wang, D., Wang, J., Lu, M., Song, F., and Cui, Q. (2010). Inferring the human microRNA functional similarity and functional network based on microRNA-associated diseases. *Bioinformatics* 26, 1644–1650. doi: 10.1093/bioinformatics/btq241
- Wang, E., Zaman, N., Mcgee, S., Milanese, J. S., Masoudi-Nejad, A., and O'connor-Mccourt, M. (2015). Predictive genomics: a cancer hallmark network framework for predicting tumor clinical phenotypes using genome sequencing data. *Semin. Cancer Biol.* 30, 4–12. doi: 10.1016/j.semcancer.2014.04.002
- Wang, E., Zou, J., Zaman, N., Beitel, L. K., Trifiro, M., and Paliouras, M. (2013a). Cancer systems biology in the genome sequencing era: part 1, dissecting and modeling of tumor clones and their networks. *Semin Cancer Biol.* 23, 279–285. doi: 10.1016/j.semcancer.2013.06.002
- Wang, E., Zou, J., Zaman, N., Beitel, L. K., Trifiro, M., and Paliouras, M. (2013b). Cancer systems biology in the genome sequencing era: part 2, evolutionary dynamics of tumor clonal networks and drug resistance. *Semin. Cancer Biol.* 23, 286–292. doi: 10.1016/j.semcancer.2013.06.001
- Wang, W., Yang, S., and Li, J. (2013c). Drug target predictions based on heterogeneous graph inference. *Pac. Symp. Biocomput.* 2013, 53–64. doi: 10.1142/9789814447973\_0006
- Watanabe, A., Tagawa, H., Yamashita, J., Teshima, K., Nara, M., Iwamoto, K., et al. (2011). The role of microRNA-150 as a tumor suppressor in malignant lymphoma. *Leukemia* 25, 1324–1334. doi: 10.1038/leu.2011.81
- Xiao, J., Gong, A. Y., Eischeid, A. N., Chen, D., Deng, C., Young, C. Y., et al. (2012). miR-141 modulates androgen receptor transcriptional activity in human prostate cancer cells through targeting the small heterodimer partner protein. *Prostate* 72, 1514–1522. doi: 10.1002/pros.22501
- Xu, J., Li, C. X., Lv, J. Y., Li, Y. S., Xiao, Y., Shao, T. T., et al. (2011). Prioritizing candidate disease miRNAs by topological features in the miRNA target-dysregulated network: case study of prostate cancer. *Mol. Cancer Ther.* 10, 1857–1866. doi: 10.1158/1535-7163.MCT-11-0055
- Xuan, P., Han, K., Guo, M., Guo, Y., Li, J., Ding, J., et al. (2013). Correction: prediction of microRNAs associated with human diseases based on weighted k most similar neighbors. *PLoS ONE* 8:e70204. doi: 10.1371/annotation/a076115e-dd8c-4da7-989d-c1174a8cd31e
- Xuan, P., Han, K., Guo, Y., Li, J., Li, X., Zhong, Y., et al. (2015). Prediction of potential disease-associated microRNAs based on random walk. *Bioinformatics* 31, 1805–1815. doi: 10.1093/bioinformatics/btv039
- Yang, Z., Ren, F., Liu, C., He, S., Sun, G., Gao, Q., et al. (2010). dbDEMOC: a database of differentially expressed miRNAs in human cancers. *BMC Genomics* 11:S5. doi: 10.1186/1471-2164-11-S4-S5
- You, Z. H., Huang, Z. A., Zhu, Z., Yan, G. Y., Li, Z. W., Wen, Z., et al. (2017). PBMDA: A novel and effective path-based computational model for miRNA-disease association prediction. *PLoS Comput. Biol.* 13:e1005455. doi: 10.1371/journal.pcbi.1005455
- Yuan, Y., Ren, X., Xie, Z., and Wang, X. (2016). A quantitative understanding of microRNA-mediated competing endogenous RNA regulation. *Quant. Biol.* 4, 47–57. doi: 10.1007/s40484-016-0062-5
- Zbar, B., Klausner, R., and Linehan, W. M. (2003). Studying cancer families to identify kidney cancer genes. *Ann. Rev. Med.* 54, 217–233. doi: 10.1146/annurev.med.54.101601.152514
- Zhong, Y., Xuan, P., Wang, X., Zhang, T., Li, J., Liu, Y., et al. (2017). A non-negative matrix factorization based method for predicting disease-associated miRNAs in miRNA-disease bilayer network. *Bioinformatics* 34, 267–277. doi: 10.1093/bioinformatics/btx546
- Zou, Q., Li, J., Hong, Q., Lin, Z., Wu, Y., Shi, H., et al. (2015). Prediction of microRNA-disease associations based on social network analysis methods. *Biomed. Res. Int.* 2015:810514. doi: 10.1155/2015/810514

**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 Chen, Qu and Yin. 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.