frontiers
in Genetics

Check for
updates

Biased Random Walk With Restart on Multilayer Heterogeneous Networks for MiRNA–Disease Association Prediction

Jia Qu[1]*, Chun-Chun Wang[2], Shu-Bin Cai[3], Wen-Di Zhao[1], Xiao-Long Cheng[1] and Zhong Ming[3]*

[1] School of Computer Science and Artificial Intelligence & Aliyun School of Big Data, Changzhou University, Changzhou, China, [2] Information and Control Engineering, China University of Mining and Technology, Xuzhou, China, [3] College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China

Numerous experiments have proved that microRNAs (miRNAs) could be used as diagnostic biomarkers for many complex diseases. Thus, it is conceivable that predicting the unobserved associations between miRNAs and diseases is extremely significant for the medical field. Here, based on heterogeneous networks built on the information of known miRNA–disease associations, miRNA function similarity, disease semantic similarity, and Gaussian interaction profile kernel similarity for miRNAs and diseases, we developed a computing model of biased random walk with restart on multilayer heterogeneous networks for miRNA–disease association prediction (BRWRMHMDA) through enforcing degree-based biased random walk with restart (BRWR). Assessment results reflected that an AUC of 0.8310 was gained in local leave-one-out cross-validation (LOOCV), which proved the calculation algorithm's good performance. Besides, we carried out BRWRMHMDA to prioritize candidate miRNAs for esophageal neoplasms based on HMDD v2.0. We further prioritize candidate miRNAs for breast neoplasms based on HMDD v1.0. The local LOOCV results and performance analysis of the case study all showed that the proposed model has good and stable performance.

Keywords: microRNA, disease, association prediction, degree, biased random walk with restart

INTRODUCTION

MicroRNA (miRNA) is a noncoding single-stranded RNA with a length of about 22 nucleotides and pervasive in both animals and plants (Axtell et al., 2011). MiRNAs play their regulator role through binding to imperfect complementary sites within the 3′ untranslated regions (UTRs) of their messenger RNA (mRNA) targets (Reinhart et al., 2000; Ambros, 2004; Bartel, 2009). Nowadays, a large number of experimental studies have proved that miRNAs regulate multiple biological activities and per miRNA can regulate hundreds of gene targets (Lee et al., 1993; Pasquinelli and Ruvkun, 2002; Brennecke et al., 2003; Lin et al., 2003; Cheng et al., 2005; Karp and Ambros, 2005; Miska, 2005; Pillai et al., 2005; Cui et al., 2006; Lu et al., 2008; Bartel, 2009; Alshalalfa and Alhajj, 2013). Moreover, miRNAs have potential influences on almost all genetic pathways, and the upregulation and downregulation of miRNA expression in the human body are correlated to various complex diseases (Liu et al., 2008). It indicates that miRNAs have close

associations with many complex diseases, and miRNAs may be used as a tumor suppressor gene to treat cancer in clinical medicine (Cheng et al., 2005). For example, the abnormal expression of miR-21 could be conducive to the growth and spread of human hepatocellular cancer (HCC) via the regulation of phosphatase and tensin homolog (PTEN) expression and PTEN-dependent pathways (Meng et al., 2007). MiR-10b is expressed in metastatic breast cancer cells highly and has a positive regulatory effect on cell migration and invasion (Ma et al., 2007). Research further suggested that the overexpression of miR-17-92 in lung cancer could enhance cell proliferation (Hayashita et al., 2005). Moreover, the miRNA family of let-7 was reported to downregulate in lung cancers and regulate an oncogene of RAS, so the inhibition of let-7 may help in the treatment of the cancer (Johnson et al., 2005). Also, through targeting an antiapoptotic factor of B-cell lymphoma-2 (BCL2), miR-15 and miRNA-16 were proved to downregulate in chronic lymphocytic leukemias and induce apoptosis (Cimmino et al., 2005). Certainly, identification of potential miRNA–disease associations has become a very significant research goal in the field of biomedical research. Predicting potential miRNAs related to diseases would promote people's understanding of the pathogenesis of diseases at the molecular level and benefit for the diagnosis, treatment, and prevention of diseases. Recently, some reliable databases have been developed to store experimental verified miRNA–disease associations, such as HMDD v2.0 (Li et al., 2014), miR2Disease (Jiang et al., 2009), and dbDEMC (Yang et al., 2010). Using traditional experiment approach to identify potential miRNA-disease associations is usually complex, time consuming and expensive. It is an urgent need for scholars to develop calculation models to predict new miRNA–disease associations. We expect that miRNA–disease pairs with high scores could be selected for experimental verification, which would significantly reduce the time and cost of biological experiments.

Great progress has been made in developing calculation models for the potential miRNA–disease association prediction in recent years. These prediction models are usually proposed by the consideration of complex network-based or machine learning-based methods (Chen et al., 2019a). For the experimentally confirmed miRNA–disease associations that have been collected, a lot of calculation models were put forward for the identification of new miRNA–disease associations on the basis of the hypothesis that functionally similar miRNAs are often associated with phenotypically similar diseases (Perez-Iratxeta et al., 2002; Aerts et al., 2006). In 2013, human disease-related miRNA prediction (HDMP), an effective prediction algorithm based on weighted k most similar neighbors, was proposed by Xuan et al. (2015). In the model, functional similarity between each miRNA pair was calculated by combining the information of their related disease terms and disease phenotype similarity. Then the possibility of unobserved miRNA–disease pairs was predicted via the sum of subscores of miRNA's k neighbor. The subscore for a neighbor of a miRNA can be calculated based on the weight of the neighbor and the functional similarity between the neighbor and the miRNA. In 2014, based on known miRNA–disease associations, disease similarity, and

miRNA similarity, a global method of regularized least squares for miRNA–disease association (RLSMDA) was introduced by Chen and Yan (2014) to uncover novel associations between miRNAs and diseases under the framework of a semisupervised classifier. In 2015, based on the constructed miRNA functional network, another new model of miRNAs associated with diseases prediction (MIDP) was developed by Xuan et al. (2015) to prioritize candidate miRNAs for investigated diseases with known related miRNAs. In the model, for the marked nodes and unmarked nodes, transition matrices are different, and the transition weight of marked nodes was higher than that of unmarked nodes. Moreover, due to the fact that MIDP could not predict potential miRNAs (diseases) associated with new diseases (miRNAs) without any known related miRNAs (diseases), an extension approach of MIDPE was also proposed to predict potential miRNAs (diseases) associated with new diseases (miRNAs). Chen et al. (2017) published a model of ranking-based KNN for miRNA–disease association prediction (RKNNMDA), in which the KNN approach was employed to gain the k-nearest-neighbors of each miRNA and disease according to the collected data information. Then, based on the Hamming loss of per disease pair and miRNA pair, a support vector machine (SVM) ranking model was introduced to achieve scores of potential miRNA–disease associations. Furthermore, Chen and Huang (2017) presented a computational model named Laplacian regularized sparse subspace learning for miRNA–disease association prediction (LRSSLMDA), which projected miRNAs' feature and diseases' feature into a common subspace. Then, the local structures of the training data were obtained based on Laplacian regularization, and the final predicted scores would be obtained by carrying out the L1-norm constraint. Chen et al. (2018a) put forward a machine learning-based method of extreme gradient boosting machine for miRNA–disease association prediction (EGBMMDA), in which a feature vector for the miRNA–disease pair was established by merging three matrices of miRNA functional similarity, disease semantic similarity, and known miRNA–disease associations. Then, based on the characteristics and the gradient boosting framework, a regression tree was applied to obtained scores of potential miRNA–disease associations. In the same year, a computational model of ensemble learning and link prediction for miRNA–disease association prediction (ELLPMDA) was brought forward by Chen et al. (2018f); they inferred new miRNA–disease associations via the weight-based integration of three classified results gained from common neighbors, Jaccard index and Katz index. Also, from the angle of reducing the noise of the original collected known miRNA–disease association information, Chen et al. (2018e) further brought up a calculation method of matrix decomposition and heterogeneous graph inference for miRNA–disease association prediction (MDHGI). The sparse learning method was carried out firstly on the initial association information to reduce noise. Then, an iterative formula for propagating miRNA and disease information was established based on the built heterogeneous network to predict potential miRNA–disease associations. Besides, Chen et al. (2018c) presented a novel method of inductive matrix completion for miRNA–disease

association prediction (IMCMDA) through enforcing a low-rank inductive matrix completion approach on the collected datasets. Chen et al. (2018d) also developed a prediction model of bipartite network projection for miRNA–disease association prediction (BNPMDA). In the model, the bias ratings for miRNAs and diseases were built based on agglomerative hierarchical clustering. Then, through assigning transfer weights to resource allocation links between miRNAs and diseases according to the bias ratings, the bipartite network recommendation algorithm was implemented to predict the potential miRNA–disease associations. Chen et al. (2019b) put forward a machine learning-based method named ensemble of decision tree-based miRNA–disease association prediction (EDTMDA), which identifies potential disease–miRNA association by implementing ensemble learning based on decision trees (DTs) and dimensionality reduction based on principal component analysis (PCA). In recent years, Chen et al. (2021) further proposed the neighborhood constraint matrix completion for miRNA–disease association prediction (NCMCMDA), which combined the neighborhood constraint with matrix completion. The prediction problem in NCMCMDA can be transformed into an optimization problem, and a fast iterative shrinkage–thresholding algorithm was implemented to solve it.

Some scholars have also introduced some calculation models on the basis of various types of association networks, rather than limited to the miRNA–disease network. In 2014, through the analysis of miRNA–protein associations and protein–disease associations, Mork et al. (2014) developed a scoring scheme for the potential miRNA–disease association prediction. In 2016, through taking advantage of miRNA–disease associations, miRNA–neighbor associations, miRNA–target associations, miRNA–word associations, and miRNA–family associations, Pasquier and Gardes (2016) expressed the distribution information of miRNAs and diseases in a high-dimensional vector space and then inferred association scores between miRNAs and diseases according to their vector similarity. In 2017, based on the phenome–miRNA network constructed by known miRNA–disease associations, miRNA functional similarity, disease semantic similarity, and phenotypic similarity, a combinatorial prioritization algorithm was proposed by Yu et al. (2017) to predict potential miRNA–disease associations. In 2018, through constructing a three-layer heterogeneous network based on the integration of known miRNA–lncRNA interactions, miRNA–disease associations, miRNA similarity, disease similarity, and lncRNA similarity, Chen et al. (2018b) designed a method of triple-layer heterogeneous network-based inference for miRNA–disease association prediction (TLHNMDA) by establishing two information spreading iterative formulas.

In this manuscript, based on a multilayer heterogeneous network established by known miRNA–disease associations, disease semantic similarity, miRNA functional similarity, and Gaussian interaction profile kernel similarity for diseases and miRNAs, we put forward a calculating model of biased random walk with restart on multilayer heterogeneous networks for miRNA–disease association prediction (BRWRMHMDA). In the model, degree-based biased random walk with restart (BRWR) was implemented to predict potential miRNA–disease associations on the basis of the constructed multilayer heterogeneous network. For evaluating the property of the introduced calculation model, local leave-one-out cross-validation (LOOCV) was presented and the outcome showed that BRWRMHMDA possesses an AUC of 0.8310 in local LOOCV. In the case study, we not only employed BRWRMHMDA to infer candidate miRNAs for esophageal neoplasms in the light of known miRNA–disease associations extracted from HMDD v2.0 (Li et al., 2014) but also implemented the model to predict breast neoplasms-associated miRNAs on the basis of known miRNA–disease associations collected from HMDAD v1.0. From the result of LOOCV and the case study, we can be sure that BRWRMHMDA has better prediction ability, and BRWRMHMDA can be used to predict potential miRNA–disease associations.

MATERIALS AND METHODS

Human miRNA–Disease Association

The dataset of 5,430 experimentally verified associations between 383 diseases and 495 miRNAs came from the HMDD v2.0 database (Li et al., 2014). We used the variables nm and nd to refer to the number of diseases and miRNAs in the dataset, respectively. Afterward, an adjacency matrix A was established to indicate known miRNA–disease associations. If miRNA $m(i)$ is related to $d(j)$, the value of entity $A(i, j)$ would equal to 1, otherwise 0.

$$A(i, j) = \begin{cases} 1, & \text{if miRNA } m(j) \text{ is related to disease } d(i) \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

MiRNA Functional Similarity

Since functionally similar miRNAs are more likely to be associated with phenotypically similar diseases on the basis of the previous study (Wang et al., 2010), we got the information of miRNA functional similarity from http://www.cuilab.cn/files/images/cuilab/misim.zip. After that, we constructed a miRNA functional similarity matrix FS with the row and column of nm. It is remarkable that the value of entity $FS(i, j)$ refers to the similarity score between miRNA $m(i)$ and miRNA $m(j)$.

Disease Semantic Similarity Model 1

Each disease can be described as a directed acyclic graph (DAG) according to previous literature (Wang et al., 2010). For example, disease D can be described as $DAG = (D, T(D), E(D))$, where $T(D)$ refers to all disease nodes, and $E(D)$ indicates all edges that connect disease nodes based on $DAG(D)$. Inspired by previous work (Xuan et al., 2013), the contribution value of disease d in $DAG(D)$ to the semantic value of disease D can be defined as follows:

$$\begin{cases} D_D1(d) = 1 & \text{if } d = D \\ D_D1(d) = \max\{\Delta^* D_D1(d') | d' \in \text{children of } d\} & \text{if } d \neq D \end{cases} \quad (2)$$

where Δ is the semantic contribution decay factor, and the semantic value of disease D can be described as follows:

$$DV1(D) = \sum_{d \in T(D)} D_D1(d) \qquad (3)$$

Considering that two diseases would have greater similarity if they share larger part of their *DAGs*, we defined the semantic similarity between disease $d(i)$ and $d(j)$ in disease semantic similarity model 1 as follows:

$$SS1(d(i), d(j)) = \frac{\sum_{t \in T(d(i)) \cap T(d(j))} (D_{d(i)}1(t) + D_{d(j)}1(t))}{DV1(d(i)) + DV1(d(j))} \qquad (4)$$

Disease Semantic Similarity Model 2

Also inspired by previous work (Xuan et al., 2013), we also introduced disease semantic similarity model 2. For two diseases in the same layer of $DAG(D)$, if the first disease occurs more frequently in $DAG(D)$ than the second disease, the second disease would be regarded to be more specific to disease D. By the consideration of the idea that the contribution of different disease terms in the same layer of $DAG(D)$ may be the difference, the contribution of disease d in $DAG(D)$ to the semantic value of disease D could be described as follows:

$$D_D2(d) = -\log[\frac{the\ number\ of\ DAGs\ including\ d}{the\ number\ of\ disease}] \qquad (5)$$

The value of semantic similarity in disease semantic similarity model 2 between disease $d(i)$ and $d(j)$ could be calculated as follows:

$$SS2(d(i), d(j)) = \frac{\sum_{t \in T(d(i)) \cap T(d(j))} (D_{d(i)}2(t) + D_{d(j)}2(t))}{DV2(d(i)) + DV2(d(j))} \qquad (6)$$

where

$$DV2(D) = \sum_{d \in T(D)} D_D2(d) \qquad (7)$$

Gaussian Interaction Profile Kernel Similarity

The calculation of Gaussian interaction profile kernel similarity for diseases and miRNAs depends on the topologic information of known miRNA–disease associations (van Laarhoven et al., 2011). For diseases, we used a binary vector $IP(d(u))$ (i.e., the uth row of the adjacency matrix A) to indicate the interaction profiles of disease $d(u)$. Accordingly, the Gaussian interaction profile kernel similarity between diseases $d(u)$ and $d(v)$ can be described.

$$KD(d(u), d(v)) = \exp(-\gamma_d ||IP(d(u)) - IP(d(v))||^2) \qquad (8)$$

The parameter γ_d was used to regulate the kernel bandwidth and could be acquired via the normalization of a new bandwidth γ_d' by the average number of associated miRNAs for each disease.

$$\gamma_d = \gamma_d' / \left(\frac{1}{nd} \sum_{u=1}^{nd} ||IP(d(u))||^2\right) \qquad (9)$$

For miRNAs, the binary vector $IP(m(i))$ (i.e., the ith column of the adjacency matrix A) was introduced to indicate the interaction profiles of miRNA $m(i)$. At last, the Gaussian interaction profile kernel similarity between miRNA $m(i)$ and $m(j)$ can be constructed as follows:

$$KM(m(i), m(j)) = \exp(-\gamma_m ||IP(m(i)) - IP(m(j))||^2) \qquad (10)$$

$$\gamma_m = \gamma_m' / \left(\frac{1}{nm} \sum_{i=1}^{nm} ||IP(m(i))||^2\right) \qquad (11)$$

Integrated Similarity for miRNAs and Diseases

Based on past work (Chen et al., 2016), integrated similarity for a pair of diseases ($d(u)$ and $d(v)$) can be defined via the combination of disease semantic similarity and Gaussian interaction profile kernel similarity for diseases. The formula of integrated similarity for diseases is displayed as follows:

$$SD(d(u), d(v)) = \begin{cases} \frac{SS1(d(u) + d(v)) + SS2(d(u), d(v))}{2} & d(u)\ and\ d(v)\ has \\ & semantic\ similarity \\ KD(m(u), m(v)) & otherwise \end{cases} \qquad (12)$$

Also, the integrated similarity for a pair of miRNAs ($m(i)$ and $m(j)$) could be formed by taking miRNA functional similarity with Gaussian interaction profile kernel similarity for miRNA into account (Chen et al., 2016).

$$SM(m(i), m(j)) = \begin{cases} FS(m(i), m(j)) & m(i)\ and\ m(j)\ has\ functional\ similarity \\ KM(m(i), m(j)) & otherwise \end{cases} \qquad (13)$$

BRWRMHMDA

Via the integration of known miRNA–disease associations, disease semantic similarity, miRNA functional similarity, and Gaussian interaction profile kernel similarity for miRNAs and diseases, we put forward a calculating model of BRWRMHMDA based on the degree for the identification of potential miRNA–disease associations by enforcing BRWR on a constructed multilayer heterogeneous network according to previous work (Bonaventura et al., 2014) (see **Figure 1**).

In the model, based on the constructed multisource dataset, we used W_{dd}, W_{mm}, W_{dm} to represent the initial matrix of integrated disease similarity, integrated miRNA similarity, and known miRNA–disease associations, respectively. Then, the multilayer heterogeneous network was constructed and described as $W = \begin{bmatrix} W_{dd} & W_{dm} \\ W_{md} & W_{mm} \end{bmatrix}$. In BRWR, if we predicted potential miRNAs for disease $d(i)$, the disease $d(i)$ is the seed node in the disease network. If the miRNA $m(j)$ is associated with disease $d(i)$, miRNA $m(j)$ is the seed node for disease $d(i)$ in the miRNA network. If the miRNA $m(j)$ has no known association with disease $d(i)$, miRNA $m(j)$ is the candidate node for disease $d(i)$ in
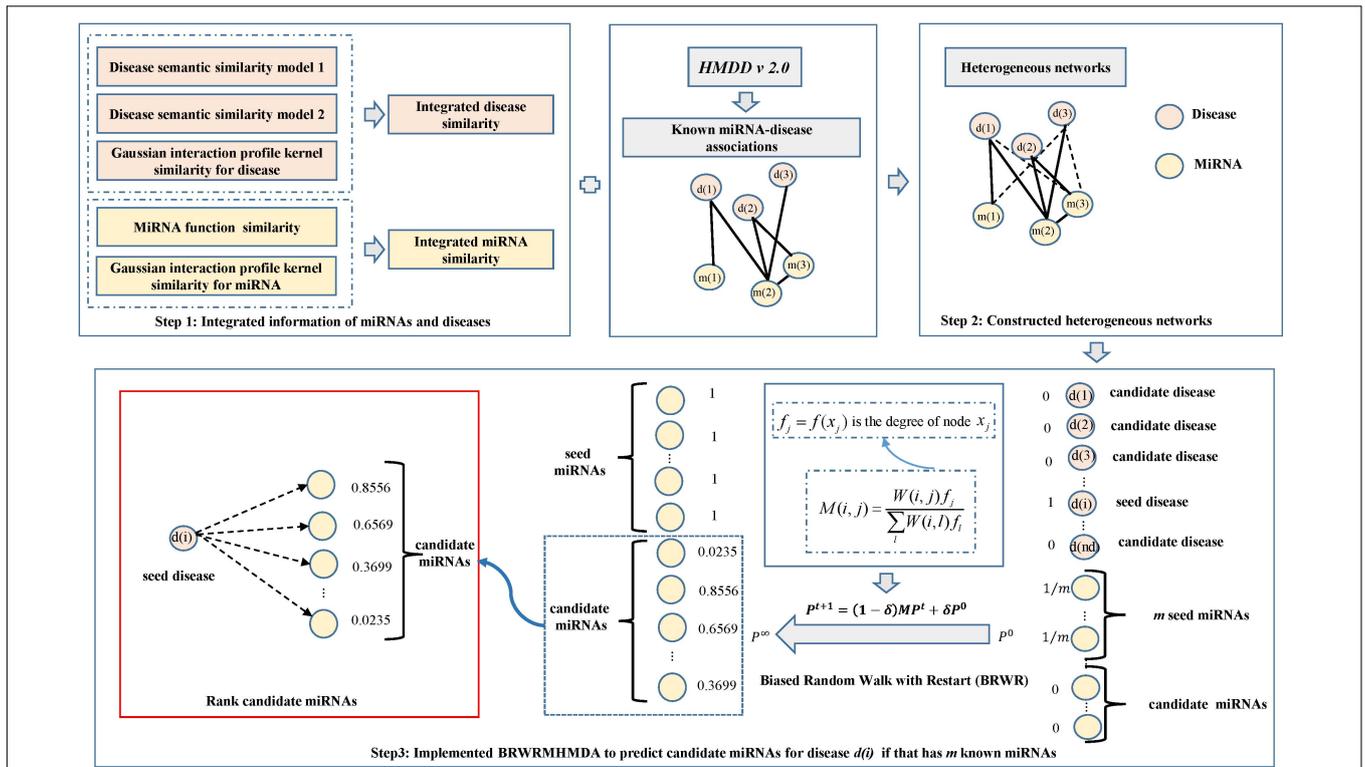
FIGURE 1 | Comparing to the other calculating algorithms (ELLPMDA, IMCMDA, EGBMMDA, MDHGI, TLHNMDA, MaxFlow, RLSMDA, HDMP, WBSMDA, MirAI, and MIDP) in terms of AUCs, BRWRMHMDA gained a better AUC value of 0.8310. It indicates that the proposed model is more suitable for the miRNA–disease association prediction.

the miRNA network. For predicting potential miRNAs for disease $d(i)$, the original probability vector v_0 of the miRNA network is computed through assigning equal probability to the seed nodes in the miRNA network with a total equal to 1. In the disease network, the probability value of 1 was assigned to $d(i)$, and the probability value of 0 was assigned to other diseases to form u_0, where the initial seed node probability $P^0 = \begin{bmatrix} \alpha * u_0 \\ (1-\alpha) * v_0 \end{bmatrix}$; α and $(1 - \alpha)$ refer to the weight of the disease network and the miRNA network, respectively.

Seed nodes at each step move to their immediate neighbors with a probability $(1 - \delta)$ or return to the seed nodes with a restart probability δ ($\delta \in (0, 1)$). P^0 was the initial probability vector, and P^{t+1} was a probability vector of node at time $t + 1$, which could be defined as follows:

$$P^{t+1} = (1 - \delta)MP^t + \delta P^0 \qquad (14)$$

where matrix $M = \begin{bmatrix} M_{dd} & M_{dm} \\ M_{md} & M_{mm} \end{bmatrix}$ is the transition matrix of our established network. In random walk with restart (RWR), the transition probability $M(i, j)$ of a walker from node i to node j can be described as follows:

$$M(i, j) = \frac{W(i, j)}{\sum_l W(i, l)} \qquad (15)$$

where $W(i, j)$ is the similarity between node i and node j. In this model, BRWR of degree biased random walk was proposed to identify potential miRNA–disease associations. Biases were usually considered to be related to graph topological properties. For example, a walk at node x_i selects it neighbors of x_j with a probability $f_j = f(x_j)$ relying on the node property x_j. Usually, the node property can be described as a function of the vertex properties (the network degree, closeness centrality, etc.) or the edge properties (multiplicity or shortest path), or the combination of them (Gomez-Gardenes and Latora, 2008). There are other related bias choice of maximal entropy (Burda et al., 2009). Thus, the transition probability of a walker from i to j in BRWR can be defined as

$$M(i, j) = \frac{W(i, j)f_j}{\sum_l W(i, l)f_l} \qquad (16)$$

Therefore, in the disease similarity network, the transition probability from vertex d_i to d_j can be defined as

$$M_{dd}(i, j) = p(d_j | d_i) =$$
$$\begin{cases} W_{dd}(i, j)f_j / \sum_j W_{dd}(i, j)f_j & \text{if } \sum_j W_{dm}(i, j) = 0 \\ (1 - \lambda)W_{dd}(i, j)f_j / \sum_j W_{dd}(i, j)f_j & \text{otherwise} \end{cases}$$
$$(17)$$

In the miRNA similarity network, the transition probability from m_i to m_j can be defined as

$$M_{mm}(i,j) = p(m_j|m_i) = \begin{cases} W_{mm}(i,j)f_j / \sum_j W_{mm}(i,j)f_j & if \sum_j W_{dm}(j,i) = 0 \\ (1-\lambda)W_{mm}(i,j)f_j / \sum_j W_{mm}(i,j)f_j & otherwise \end{cases}$$

$$(18)$$

In the miRNA–disease association network, the transition probability from vertex d_i to m_j can be defined as

$$M_{dm}(i,j) = p(m_j|d_i) = \begin{cases} \lambda W_{dm}(i,j)f_j / \sum_j W_{dm}(i,j)f_j & if \sum_j W_{dm}(i,j) \neq 0 \\ 0 & otherwise \end{cases}$$

$$(19)$$

The transition probability from vertex m_i to d_j can be defined as

$$M_{md}(i,j) = p(d_j|m_i) = \begin{cases} \lambda W_{dm}(j,i)f_j / \sum_j W_{dm}(j,i)f_j & if \sum_j W_{dm}(j,i) \neq 0 \\ 0 & otherwise \end{cases}$$

$$(20)$$

In this paper, we focus on the case of BRWR by considering the degree nodes. Therefore, $f_j = f(x_j)$ in the model is the degree of node x_j in the transition probability. The degree f_i of a disease node i is defined by computing the number of edges involved in the disease node. Therefore, in the disease similarity network, the degree of disease node j can be defined as $f_j = \sum_i W_{dd}(i,j)$. In the miRNA similarity network, the degree of miRNA node j can be defined as $f_j = \sum_i W_{mm}(i,j)$. In the transition probability matrix of the miRNA–disease association network, the degree of miRNA m_j can be described as $f_j = \sum_i W_{dm}(i,j)$. Also, in the transition probability matrix of the miRNA–disease association network, the degree of disease d_j can be described as $f_j = \sum_i W_{dm}(j,i)$. Therefore, based on BRWR of degree nodes, the potential miRNA–disease associations would be obtained.

RESULTS

Performance Evaluation

Since BRWR is a local calculating method, it cannot infer candidate miRNAs for all diseases simultaneously. Therefore, in order to analyze the performance of BRWRMHMDA, the proposed method has been extensively compared with some classic algorithms (ELLPMDA, IMCMDA, EGBMMDA, MDHGI, TLHNMDA, MaxFlow, RLSMDA, HDMP, WBSMDA, MirAI, and MIDP) based on the 5,430 known miRNA–disease associations from the HMDD v2.0 database (Li et al., 2014) via local LOOCV. In local LOOCV, each known miRNA–disease association was considered as a test sample in turn, and the rest of 5,429 known associations were treated as training samples. After enforcing BRWRMHMDA, the score of the test sample would be sorted with the scores of all unobserved pairs between miRNAs and the investigated disease. The proposed approach would be regarded as reliable if the test sample's ranking is higher than a set threshold. Then a receiver operating characteristics (ROC) curve with the true positive rate (TPR, sensitivity) versus the

false positive rate (FPR, 1-specificity) at various thresholds would be drawn. Sensitivity refers to the percentage of test samples ranked higher than the given threshold, and specificity refers to the percentage of candidates ranked lower than the threshold. Finally, the area under the ROC curve (AUC) was calculated to accurately evaluate the prediction ability of BRWRMHMDA. The value of the AUC is between 0 and 1, and the higher the value of the AUC, the better the prediction performance of the algorithm. If the value of the AUC is 0.5, the prediction performance of BRWRMHMDA is random. The final assessment results showed that BRWRMHMDA has better prediction performance with an AUC of 0.8310 than those of the other server classical algorithms of ELLPMDA (0.8181), IMCMDA (0.8034), EGBMMDA (0.8221), MDHGI (0.8240), TLHNMDA (0.7756), MaxFlow (0.7774), RLSMDA (0.6953), HDMP (0.7702), WBSMDA (0.8031), MirAI (0.6299), and MIDP (0.8196) (see **Figure 2**). Here, the AUC value of MirAI is lower than that reported in its literature (Pasquier and Gardes, 2016) because MirAI was proposed on the basis of a collaborative filtering algorithm affected by the data sparsity problem. Compared with the training set in the original literature, our dataset is relatively scarce. The training set in the original literature contains 83 diseases and at least 20 known related miRNAs for each disease, while our training set contains 383 diseases and most diseases-related miRNAs are rare.

Case Studies

In order to further analyze the performance of the algorithm effectively, we carried out two types of case studies. The first type of case studies is the prediction of potential miRNAs associated with esophageal neoplasms based on the known miRNA–disease association collect from HMDD v2.0. The second type of case studies is the prediction of potential miRNAs associated with breast neoplasms based on the known miRNA–disease association collect from HMDD v1.0.

Esophageal neoplasm is one of the most lethal cancers in the world; its main nature is highly invasive and of low survival rate (Domper Arnal et al., 2015). The disease contains two main histological types of squamous cell cancer and adenocarcinoma (Zhang et al., 2016). Malnutrition is a main risk factor for esophageal squamous cell carcinoma (ESCC), and obesity is the main risk factor for esophageal adenocarcinoma (Domper Arnal et al., 2015). Accordingly, looking for sensitive molecular biomarkers and individual treatment approach for early diagnosis of esophageal cancer has become the main clinical and basic research direction. Numerous studies suggested that miRNAs play an important role in diseases and can be a biomarker for esophageal neoplasms' treatment. For example, miR-506 was abnormally expressed in a variety of tumors and could be used as a prognostic biomarker for ESCC (Li et al., 2016). Besides, plasma miR-718 was reported to downregulate in ESCC patients and might be treated as a potential diagnostic marker for the disease (Sun et al., 2016). Here, we employed BRWRMHMDA to prioritize candidate miRNAs for esophageal neoplasms according to the dataset of 5,430 known miRNA–disease associations between 383 diseases and 495 miRNAs. As a result, of the first 50 miRNAs predicted

FIGURE 2 | Flowchart of BRWRMHMDA to prioritize candidate miRNAs for diseases. Through employing BRWR on the established heterogeneous networks, final scores p^∞ of candidate miRNAs predicted for each disease would be gotten after some steps.
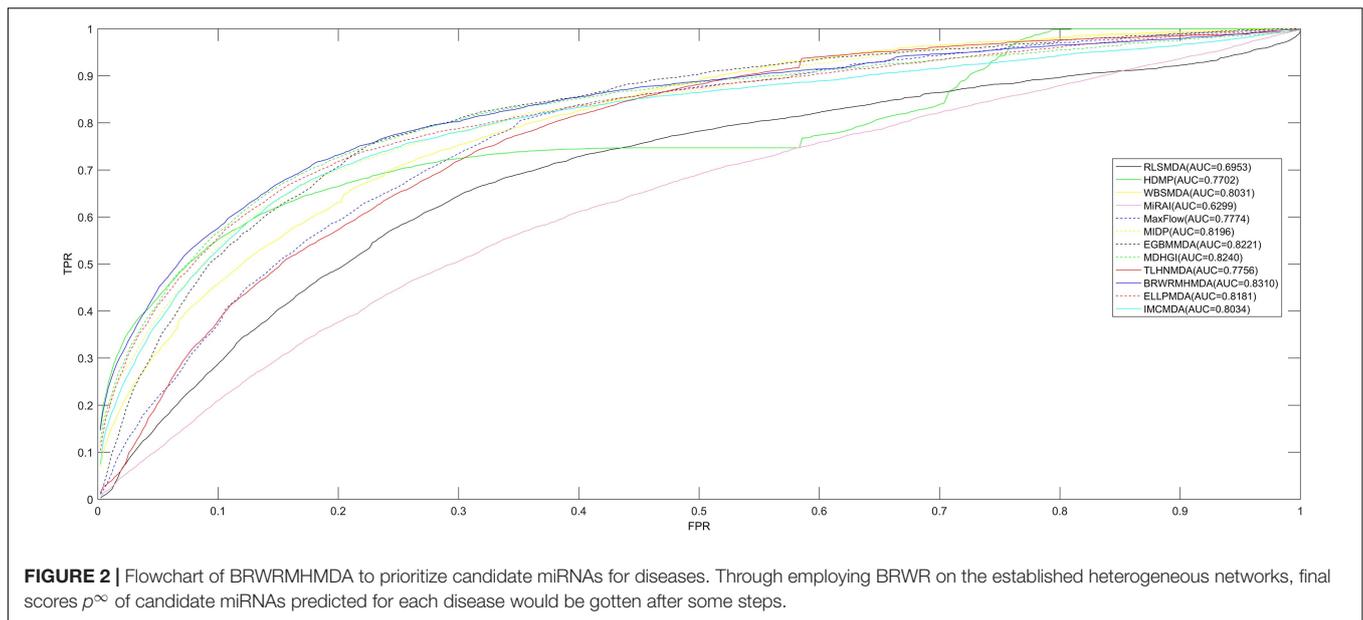
for esophageal neoplasms in the ranking, 49 miRNAs have been confirmed by the database of dbDEMC and miR2Disease (see **Table 1**). For example, the predicted association score between hsa-mir-125b and esophageal neoplasms is ranked first. Yu et al. (2020) have found that hsa-mir-125b suppresses cell proliferation and metastasis by targeting HAX-1 in ESCC, which proves that hsa-mir-125b is related to esophageal neoplasms. Moreover, the predicted association score between hsa-mir-200b and esophageal neoplasms is ranked second. Researchers have confirmed that hsa-mir-200b is downregulated in ESCC in the comparison of the respective adjacent benign tissues (Zhang et al., 2014). Therefore, hsa-mir-200b is associated with esophageal neoplasms.

Breast neoplasm is one of the three most common cancers for women (Siegel et al., 2018). In particular, metastatic breast cancer (MBC) is usually incurable, and about 5% of patients have metastases at diagnosis (Torre et al., 2015). With recent research, miR-10b sponge has been shown to effectively inhibit the growth of MDA-MB-231 and MCF-7 cells in breast cancer (Liang et al., 2016). In addition, miR-223 was demonstrated to function as a potential tumor marker for breast neoplasm through suppressing its protein expression of FOXO1 (Wei et al., 2017). Accordingly, identifying breast neoplasm-related miRNAs is meaningful, which could help the medical diagnosis and treatment for MBC (McGuire et al., 2015). Here, we enforced BRWRMHMDA to infer potential miRNAs related to breast neoplasms on the basis of 1,395 known miRNA–disease associations between 137 diseases and 271 miRNAs collected from HMDD v1.0. The results showed that 48 of the first 50 miRNAs predicted for breast neoplasms have been confirmed by the databases of dbDEMC, miR2Disease, and HMDD v2.0 (see **Table 2**). For example, hsa-let-7b was predicted to associate with breast neoplasms, and the predicted score is ranked second. It is worth noting that hsa-let-7b can significantly change oncogenic signaling in breast cancer

cells. Consequently, hsa-let-7b may have important roles in breast neoplasm progression and can be considered as potential targets for breast neoplasm therapy and diagnosis (Bozgeyik, 2020). Besides, hsa-mir-16 was predicted to be related to breast neoplasms, and the predicted score is ranked third. Haghi et al. indicated that has-mir-16 and has-mir-34a can collaborate in breast tumor suppression, which proved that hsa-mir-16 has association with breast neoplasms.

At last, we have released the whole prediction results via the implementation of BRWRMHMDA for all miRNA–disease pairs between 383 diseases and 495 miRNAs from HMDD v2.0 (see **Supplementary Table 1**).

DISCUSSION

Through integrating known miRNA–disease associations, disease semantic similarity, miRNA function similarity, and Gaussian interaction profile kernel similarity for miRNAs and diseases, BRWRMHMDA was employed in this manuscript to prioritize candidate miRNAs for diseases via the implementation of degree-based BRWR on the established networks. The assessment results of LOOCV showed that the developed algorithm outperforms the other 11 classic prediction algorithms in accuracy. We further enforced the proposed algorithm to infer candidate miRNAs for esophageal neoplasms in the light of known miRNA–disease associations extracted from HMDD v2.0 and infer candidate miRNAs for breast neoplasms in the light of known miRNA–disease associations extracted from HMDD v1.0. The results of the case study fully demonstrated the stability of this introduced algorithm. It is worth mentioning that our research group will keep on studying this issue in depth. Furthermore, we hope more external research groups would select potential associations with high prediction scores and verify them based on biological experiment in the future.

TABLE 1 | The implementation of BRWRMHMDA to prioritize candidate miRNAs for esophageal neoplasms based on experimentally confined miRNA–disease associations collected from HMDD v2.0 and 47 of the first 50 predicted miRNAs were confirmed.

miRNA	Evidence	miRNA	Evidence
hsa-mir-125b	dbDEMC	hsa-mir-429	dbDEMC
hsa-mir-200b	dbDEMC	hsa-mir-106a	dbDEMC
hsa-mir-18a	dbDEMC	hsa-mir-24	dbDEMC
hsa-mir-17	dbDEMC	hsa-mir-30c	dbDEMC
hsa-mir-221	dbDEMC	hsa-mir-218	unconfirmed
hsa-mir-19b	dbDEMC	hsa-mir-93	dbDEMC
hsa-mir-16	dbDEMC	hsa-mir-132	dbDEMC
hsa-mir-1	dbDEMC	hsa-mir-30a	dbDEMC
hsa-mir-222	dbDEMC	hsa-mir-127	dbDEMC
hsa-let-7i	dbDEMC	hsa-mir-195	dbDEMC
hsa-mir-29a	dbDEMC	hsa-mir-199b	dbDEMC
hsa-let-7e	dbDEMC	hsa-mir-10b	dbDEMC
hsa-let-7d	dbDEMC	hsa-mir-15b	dbDEMC
hsa-mir-29b	dbDEMC	hsa-mir-107	dbdemc and miR2Disease
hsa-let-7f	unconfirmed	hsa-mir-7	dbDEMC
hsa-mir-181b	dbDEMC	hsa-mir-224	dbDEMC
hsa-mir-181a	dbDEMC	hsa-mir-18b	dbDEMC
hsa-mir-125a	dbDEMC	hsa-mir-133b	dbDEMC
hsa-let-7g	dbDEMC	hsa-mir-335	dbDEMC
hsa-mir-9	dbDEMC	hsa-mir-194	dbdemc and miR2Disease
hsa-mir-146b	dbDEMC	hsa-mir-302b	dbDEMC
hsa-mir-106b	dbDEMC	hsa-mir-20b	dbDEMC
hsa-mir-182	dbDEMC	hsa-mir-124	dbDEMC
hsa-mir-142	dbDEMC	hsa-mir-373	dbdemc and miR2Disease
hsa-mir-122	unconfirmed	hsa-mir-191	dbDEMC

TABLE 2 | The implementation of BRWRMHMDA to prioritize candidate miRNAs for breast neoplasms based on experimentally confined miRNA–disease associations collected from HMDD v1.0 and 48 of the first 50 predicted miRNAs were confirmed.

miRNA	Evidence	miRNA	Evidence
hsa-let-7i	dbDEMC and miR2Disease and HMDD	hsa-mir-203	dbDEMC and miR2Disease and HMDD
hsa-let-7b	dbDEMC and HMDD	hsa-mir-32	dbDEMC
hsa-mir-16	dbDEMC and HMDD	hsa-mir-30e	unconfirmed
hsa-let-7e	dbDEMC and HMDD	hsa-mir-532	dbDEMC
hsa-let-7g	dbDEMC and HMDD	hsa-mir-335	dbDEMC and miR2Disease and HMDD
hsa-let-7c	dbDEMC and HMDD	hsa-mir-150	dbDEMC
hsa-mir-92a	HMDD	hsa-mir-199b	dbDEMC and HMDD
hsa-mir-126	dbDEMC and miR2Disease and HMDD	hsa-mir-99a	dbDEMC
hsa-mir-223	dbDEMC and HMDD	hsa-mir-98	dbDEMC and miR2Disease
hsa-mir-92b	dbDEMC	hsa-mir-142	unconfirmed
hsa-mir-373	dbDEMC and miR2Disease and HMDD	hsa-mir-128b	miR2Disease
hsa-mir-101	dbDEMC and miR2Disease and HMDD	hsa-mir-107	dbDEMC and HMDD
hsa-mir-191	dbDEMC and miR2Disease and HMDD	hsa-mir-224	dbDEMC and HMDD
hsa-mir-182	dbDEMC and miR2Disease and HMDD	hsa-mir-27a	dbDEMC and miR2Disease and HMDD
hsa-mir-99b	dbDEMC	hsa-mir-195	dbDEMC and miR2Disease and HMDD
hsa-mir-106a	dbDEMC	hsa-mir-124	dbDEMC and HMDD
hsa-mir-181a	dbDEMC and miR2Disease and HMDD	hsa-mir-30a	miR2Disease and HMDD
hsa-mir-29c	dbDEMC and miR2Disease and HMDD	hsa-mir-520b	dbDEMC and HMDD
hsa-mir-100	dbDEMC and HMDD	hsa-mir-95	dbDEMC
hsa-mir-18b	dbDEMC and HMDD	hsa-mir-23b	dbDEMC and HMDD
hsa-mir-372	dbDEMC	hsa-mir-491	dbDEMC
hsa-mir-24	dbDEMC and HMDD	hsa-mir-183	dbDEMC and HMDD
hsa-mir-130a	dbDEMC	hsa-mir-31	dbDEMC and miR2Disease and HMDD
hsa-mir-15b	dbDEMC	hsa-mir-192	dbDEMC
hsa-mir-196b	dbDEMC	hsa-mir-135a	dbDEMC and HMDD

Actually, the method's high accuracy in the miRNA–disease association predictions mainly rely on the following attractive properties. First, the training set of known miRNA–disease associations used in this manuscript was collected from a very reliable database of HMDD v2.0, and the several bioinformatics data (disease semantic similarity, miRNA function similarity, and Gaussian interaction profile kernel similarity for miRNAs and diseases) mentioned in the paper were accurately calculated and integrated. All the reliable biological information mentioned above would attribute to the accuracy of BRWRMHMDA. Second, compared with the machine learning-based methods that randomly select negative samples as the training set, the proposed algorithm only uses positive samples as the training set that would provide higher prediction value. At last, BRWRMHMDA, a degree-biased random walk, could fully take advantage of the information about node degree and improve the prediction accuracy. From the preceding discussion, it is no surprise that this algorithm is superior to other comparison algorithms and has good performance.

However, the proposed model still has some weaknesses and shortcomings. For example, despite the biological information collected here being reliable, the number of 5,430 experimentally verified miRNA–disease associations extracted from HMDD v2.0 is still far from enough. If more associations between miRNAs and diseases are validated, the prediction accuracy of the model would be higher. Moreover, except for the fact that miRNA similarity could be calculated via the consideration of miRNA functional similarity and Gaussian interaction profile kernel for

miRNAs, it could also be calculated based on other miRNA features. At the same time, disease similarity could also be calculated based on other disease features. Also, the model could not predict candidate miRNAs for new diseases that have no known related miRNAs. In addition, due to the fact that the proposed algorithm is a local ranking model, it could not infer candidate miRNAs for all diseases simultaneously.

Nowadays, more and more researchers are studying the regulatory interactions between ncRNA classes, as well as the associations between ncRNA and other biological entities including diseases, small molecules, etc. Prediction of ncRNA-related networks will greatly expand our understanding of ncRNA function and its regulatory network. Simultaneously, predictions including miRNA–lncRNA interactions, miRNA–circRNA interactions, drug–target interactions, small molecule–miRNA associations, and disease–lncRNA associations have made great progress. In the field of miRNA–disease association prediction, the number of known miRNA–disease associations is limited, which will affect the prediction performance of the model. In the future, integrating multisource biological data that was mentioned above to build a multilayer heterogeneous network based on machine learning-based method can effectively improve the prediction performance of the model.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

JQ implemented the experiments, analyzed the result, and wrote the manuscript. C-CW analyzed the result, revised the manuscript, and supervised the project. S-BC and ZM analyzed the result and revised the manuscript. W-DZ and X-LC contributed to the analysis of the data for the manuscript and revised the manuscript. All authors read and approved the final manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.720327/full#supplementary-material

REFERENCES

Aerts, S., Lambrechts, D., Maity, S., Van Loo, P., Coessens, B., De Smet, F., et al. (2006). Gene prioritization through genomic data fusion. *Nat. Biotechnol.* 24, 537–544. doi: 10.1038/nbt1203

Alshalalfa, M., and Alhajj, R. (2013). Using context-specific effect of miRNAs to identify functional associations between miRNAs and gene signatures. *BMC Bioinformatics* 14:S1. doi: 10.1186/1471-2105-14-S12-S1

Ambros, V. (2004). The functions of animal microRNAs. *Nature* 431, 350–355. doi: 10.1038/nature02871

Axtell, M. J., Westholm, J. O., and Lai, E. C. (2011). Vive la difference: biogenesis and evolution of microRNAs in plants and animals. *Genome Biol.* 12:221. doi: 10.1186/gb-2011-12-4-221

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233. doi: 10.1016/j.cell.2009.01.002

Bonaventura, M., Nicosia, V., and Latora, V. (2014). Characteristic times of biased random walks on complex networks. *Phys. Rev. E Stat. Nonlin. Soft. Matter. Phys.* 89:012803.

Bozgeyik, E. (2020). Bioinformatic analysis and in vitro validation of Let-7b and Let-7c in breast cancer. *Comput. Biol. Chem.* 84:107191. doi: 10.1016/j.compbiolchem.2019.107191

Brennecke, J., Hipfner, D. R., Stark, A., Russell, R. B., and Cohen, S. M. (2003). bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. *Cell* 113, 25–36. doi: 10.1016/s0092-8674(03)00231-9

Burda, Z., Duda, J., Luck, J. M., and Waclaw, B. (2009). Localization of the maximal entropy random walk. *Phys. Rev. Lett.* 102:160602.

Chen, X., and Huang, L. (2017). LRSSLMDA: laplacian regularized sparse subspace learning for MiRNA-disease association prediction. *PLoS Comput. Biol.* 13:e1005912. doi: 10.1371/journal.pcbi.1005912

Chen, X., and Yan, G. Y. (2014). Semi-supervised learning for potential human microRNA-disease associations inference. *Sci. Rep.* 4:5501.

Chen, X., Huang, L., Xie, D., and Zhao, Q. (2018a). EGBMMDA: extreme gradient boosting machine for MiRNA-disease association prediction. *Cell Death Dis.* 9:3.

Chen, X., Qu, J., and Yin, J. (2018b). TLHNMDA: triple layer heterogeneous network based inference for miRNA-disease association prediction. *Front. Genet.* 9:234. doi: 10.3389/fgene.2018.00234

Chen, X., Sun, L. G., and Zhao, Y. (2021). NCMCMDA: miRNA-disease association prediction through neighborhood constraint matrix completion. *Brief. Bioinform.* 22, 485–496. doi: 10.1093/bib/bbz159

Chen, X., Wang, L., Qu, J., Guan, N. N., and Li, J. Q. (2018c). Predicting miRNA-disease association based on inductive matrix completion. *Bioinformatics* 34, 4256–4265.

Chen, X., Wu, Q. F., and Yan, G. Y. (2017). RKNNMDA: ranking-based KNN for MiRNA-disease association prediction. *RNA Biol.* 14, 952–962. doi: 10.1080/15476286.2017.1312226

Chen, X., Xie, D., Wang, L., Zhao, Q., You, Z. H., and Liu, H. (2018d). BNPMDA: bipartite network projection for MiRNA-disease association prediction. *Bioinformatics* 34, 3178–3186. doi: 10.1093/bioinformatics/bty333

Chen, X., Xie, D., Zhao, Q., and You, Z. H. (2019a). MicroRNAs and complex diseases: from experimental results to computational models. *Brief. Bioinform.* 20, 515–539. doi: 10.1093/bib/bbx130

Chen, X., Yan, C. C., Zhang, X., You, Z. H., Deng, L., Liu, Y., et al. (2016). WBSMDA: within and Between score for MiRNA-disease association prediction. *Sci. Rep.* 6:21106.

Chen, X., Yin, J., Qu, J., and Huang, L. (2018e). 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., Zhou, Z., and Zhao, Y. (2018f). ELLPMDA: ensemble learning and link prediction for miRNA-disease association prediction. *RNA Biol.* 15, 807–818.

Chen, X., Zhu, C. C., and Yin, J. (2019b). Ensemble of decision tree reveals potential miRNA-disease associations. *PLoS Comput. Biol.* 15:e1007209. doi: 10.1371/journal.pcbi.1007209

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

Cimmino, A., Calin, G. A., Fabbri, M., Iorio, M. V., Ferracin, M., Shimizu, M., et al. (2005). miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc. Natl. Acad. Sci. U.S.A.* 102, 13944–13949. doi: 10.1073/pnas.0506654102

Cui, Q., Yu, Z., Purisima, E. O., and Wang, E. (2006). Principles of microRNA regulation of a human cellular signaling network. *Mol. Syst. Biol.* 2:46. doi: 10.1038/msb4100089

Domper Arnal, M. J., Ferrandez Arenas, A., and Lanas Arbeloa, A. (2015). Esophageal cancer: risk factors, screening and endoscopic treatment in Western

and Eastern countries. *World J. Gastroenterol.* 21, 7933–7943. doi: 10.3748/wjg. v21.i26.7933

Gomez-Gardenes, J., and Latora, V. (2008). Entropy rate of diffusion processes on complex networks. *Phys. Rev. E Stat. Nonlin. Soft. Matter. Phys.* 78:065102.

Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., et al. (2005). A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. *Cancer Res.* 65, 9628–9632. doi: 10.1158/0008-5472.can-05-2352

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.

Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., et al. (2005). RAS is regulated by the let-7 microRNA family. *Cell* 120, 635–647. doi: 10.1016/j.cell.2005.01.014

Karp, X., and Ambros, V. (2005). Encountering microRNAs in cell fate signaling. *Science* 310, 1288–1289. doi: 10.1126/science.1121566

Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75, 843–854. doi: 10.1016/0092-8674(93)90529-y

Li, S. P., Su, H. X., Zhao, D., and Guan, Q. L. (2016). Plasma miRNA-506 as a prognostic biomarker for esophageal squamous cell carcinoma. *Med. Sci. Monit.* 22, 2195–2201. doi: 10.12659/msm.899377

Li, Y., Qiu, C., Tu, J., Geng, B., Yang, J., Jiang, T., et al. (2014). HMDD v2.0: a database for experimentally supported human microRNA and disease associations. *Nucleic Acids Res.* 42, D1070–D1074.

Liang, A. L., Zhang, T. T., Zhou, N., Wu, C. Y., Lin, M. H., and Liu, Y. J. (2016). MiRNA-10b sponge: an anti-breast cancer study in vitro. *Oncol. Rep.* 35, 1950–1958. doi: 10.3892/or.2016.4596

Lin, S. Y., Johnson, S. M., Abraham, M., Vella, M. C., Pasquinelli, A., Gamberi, C., et al. (2003). The C elegans hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. *Dev. Cell* 4, 639–650. doi: 10.1016/s1534-5807(03)00124-2

Liu, Z., Sall, A., and Yang, D. (2008). MicroRNA: an emerging therapeutic target and intervention tool. *Int. J. Mol. Sci.* 9, 978–999. doi: 10.3390/ijms9060978

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

Ma, L., Teruya-Feldstein, J., and Weinberg, R. A. (2007). Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature* 449, 682–688. doi: 10.1038/nature06174

McGuire, A., Brown, J. A., and Kerin, M. J. (2015). Metastatic breast cancer: the potential of miRNA for diagnosis and treatment monitoring. *Cancer Metastasis Rev.* 34, 145–155. doi: 10.1007/s10555-015-9551-7

Meng, F., Henson, R., Wehbe-Janek, H., Ghoshal, K., Jacob, S. T., and Patel, T. (2007). MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer. *Gastroenterology* 133, 647–658. doi: 10.1053/j.gastro.2007.05.022

Miska, E. A. (2005). How microRNAs control cell division, differentiation and death. *Curr. Opin. Genet. Dev.* 15, 563–568. doi: 10.1016/j.gde.2005.08.005

Mork, S., Pletscher-Frankild, S., Palleja Caro, A., Gorodkin, J., and Jensen, L. J. (2014). Protein-driven inference of miRNA-disease associations. *Bioinformatics* 30, 392–397. doi: 10.1093/bioinformatics/btt677

Pasquier, C., and Gardes, J. (2016). Prediction of miRNA-disease associations with a vector space model. *Sci. Rep.* 6:27036.

Pasquinelli, A. E., and Ruvkun, G. (2002). Control of developmental timing by micrornas and their targets. *Annu. Rev. Cell Dev. Biol.* 18, 495–513. doi: 10.1146/annurev.cellbio.18.012502.105832

Perez-Iratxeta, C., Bork, P., and Andrade, M. A. (2002). Association of genes to genetically inherited diseases using data mining. *Nat. Genet.* 31, 316–319. doi: 10.1038/ng895

Pillai, R. S., Bhattacharyya, S. N., Artus, C. G., Zoller, T., Cougot, N., Basyuk, E., et al. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* 309, 1573–1576. doi: 10.1126/science.1115079

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., et al. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901–906. doi: 10.1038/35002607

Siegel, R. L., Miller, K. D., and Jemal, A. (2018). Cancer statistics, 2018. *CA Cancer J. Clin.* 68, 7–30. doi: 10.3322/caac.21442

Sun, L., Dong, S., Dong, C., Sun, K., Meng, W., Lv, P., et al. (2016). Predictive value of plasma miRNA-718 for esophageal squamous cell carcinoma. *Cancer Biomark.* 16, 265–273. doi: 10.3233/cbm-150564

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. *CA Cancer J. Clin.* 65, 87–108. doi: 10.3322/caac.21262

van Laarhoven, T., Nabuurs, S. B., and Marchiori, E. (2011). Gaussian interaction profile kernels for predicting drug-target interaction. *Bioinformatics* 27, 3036–3043. doi: 10.1093/bioinformatics/btr500

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

Wei, Y. T., Guo, D. W., Hou, X. Z., and Jiang, D. Q. (2017). miRNA-223 suppresses FOXO1 and functions as a potential tumor marker in breast cancer. *Cell Mol. Biol. (Noisy-le-grand)* 63, 113–118. doi: 10.14715/cmb/2017.63.5.21

Xuan, P., Han, K., Guo, M., Guo, Y., Li, J., Ding, J., et al. (2013). Prediction of microRNAs associated with human diseases based on weighted k most similar neighbors. *PLoS One* 8:e70204. doi: 10.1371/journal.pone.0070204

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). dbDEMC: a database of differentially expressed miRNAs in human cancers. *BMC Genomics* 11 Suppl 4:S5. doi: 10.1093/nar/gkw1079

Yu, H., Chen, X., and Lu, L. (2017). Large-scale prediction of microRNA-disease associations by combinatorial prioritization algorithm. *Sci. Rep.* 7:43792.

Yu, Z., Ni, F., Chen, Y., Zhang, J., Cai, J., and Shi, W. (2020). miR-125b suppresses cell proliferation and metastasis by targeting HAX-1 in esophageal squamous cell carcinoma. *Pathol. Res. Pract.* 216:152792. doi: 10.1016/j.prp.2019.152792

Zhang, H. F., Zhang, K., Liao, L. D., Li, L. Y., Du, Z. P., Wu, B. L., et al. (2014). miR-200b suppresses invasiveness and modulates the cytoskeletal and adhesive machinery in esophageal squamous cell carcinoma cells via targeting Kindlin-2. *Carcinogenesis* 35, 292–301. doi: 10.1093/carcin/bgt320

Zhang, L., Ma, J., Han, Y., Liu, J., Zhou, W., Hong, L., et al. (2016). Targeted therapy in esophageal cancer. *Expert. Rev. Gastroenterol. Hepatol.* 10, 595–604.