



# QIMCMDA: MiRNA-Disease Association Prediction by q-Kernel Information and Matrix Completion

Lin Wang<sup>1</sup>, Yaguang Chen<sup>1</sup>, Naiqian Zhang<sup>1</sup>, Wei Chen<sup>1</sup>, Yusen Zhang<sup>1\*</sup> and Rui Gao<sup>2\*</sup>

<sup>1</sup> School of Mathematics and Statistics, Shandong University, Jinan, China, <sup>2</sup> School of Control Science and Engineering, Shandong University, Jinan, China

## OPEN ACCESS

### Edited by:

Tao Huang,  
Shanghai Institute for Biological  
Sciences (CAS), China

### Reviewed by:

Qi Zhao,  
University of Science and Technology  
Liaoning, China  
Pingjian Ding,  
University of South China, China  
Cheng Liang,  
Shandong Normal University, China

### \*Correspondence:

Yusen Zhang  
zhangys@sdu.edu.cn  
Rui Gao  
gaorui@sdu.edu.cn

### Specialty section:

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

**Received:** 14 August 2020

**Accepted:** 21 September 2020

**Published:** 22 October 2020

### Citation:

Wang L, Chen Y, Zhang N,  
Chen W, Zhang Y and Gao R (2020)  
QIMCMDA: MiRNA-Disease  
Association Prediction by q-Kernel  
Information and Matrix Completion.  
*Front. Genet.* 11:594796.  
doi: 10.3389/fgene.2020.594796

Studies have shown that microRNAs (miRNAs) are closely associated with many human diseases, but we have not yet fully understand the role and potential molecular mechanisms of miRNAs in the process of disease development. However, ordinary biological experiments often require higher costs, and computational methods can be used to quickly and effectively predict the potential miRNA-disease association effect at a lower cost, and can be used as a useful reference for experimental methods. For miRNA-disease association prediction, we have proposed a new method called Matrix completion algorithm based on q-kernel information (QIMCMDA). We use fivefold cross-validation and leave-one-out cross-validation to prove the effectiveness of QIMCMDA. LOOCV shows that AUC can reach 0.9235, and its performance is significantly better than other commonly used technologies. In addition, we applied QIMCMDA to case studies of three human diseases, and the results show that our method performs well in inferring potential interaction between miRNAs and diseases. It is expected that QIMCMDA will become an excellent supplement in the field of biomedical research in the future.

**Keywords:** microRNA-disease interaction, association prediction, heterogeneous omics data, q-kernel neighborhood similarity, matrix factorization

## INTRODUCTION

MicroRNAs (miRNAs) are a type of single-stranded small non-coding RNA (~22 nt) that play an important role in gene regression by interfering with post-transcriptional regulation (Filipowicz et al., 2008; Bartel, 2009). Lee et al. (1993) discovered the first miRNA lin-4 in *Caenorhabditis elegans*, and since then, 1000s of currently annotated miRNAs have been found in various species from plants, animals to viruses (Jopling et al., 2005; Kozomara and Griffiths-Jones, 2011). More and more evidence have shown that miRNA is an important component in cells and may play an important role in a variety of biological processes including cell growth (Ambros, 2003), immune response (Taganov et al., 2006), cell proliferation and differentiation (Chen et al., 2004, 2006), cell development, cell cycle regulation (Carleton et al., 2007), inflammation (Urbich et al., 2008), apoptosis (Petrocca et al., 2008), and stress response (Leung and Sharp, 2010). Many studies have shown that miRNA abnormalities are associated with various human diseases, such as cancer, Alzheimer's disease, and diabetes (Iorio et al., 2005; Nunez-Iglesias et al., 2010; Catto et al., 2011; Guay et al., 2011; Farazi et al., 2013). For example, there is evidence that MicroRNA-155 regulates colon

cancer cell proliferation, cell cycle, apoptosis, migration, and targets CBL (Yu et al., 2017). miR-21 negatively regulates Pcd4 and inhibits TPA-induced tumor transformation (Asangani et al., 2008). MicroRNA-494 has become a major epigenetic regulator in aggressive human hepatocellular carcinoma neoplasms (Chuang et al., 2005). miR-146a is a tumor suppressor that inhibits NF- $\kappa$ B activity related to the promotion and inhibition of tumor growth (Li et al., 2014b). This makes miRNAs increasingly recognized as key regulators in gene expression (Niu et al., 2019). Finding the association of miRNA-disease is an important field of biomedicine. It not only helps humans understand the mechanism of diseases, but also helps the discovery, prognosis, diagnosis, treatment, and prevention of human complex diseases (Calin and Croce, 2006; Tricoli and Jacobson, 2007; Cho, 2010; Jiang et al., 2010).

However, the identification of miRNA-disease associations using traditional biological methods is often costly (Chen et al., 2018). Therefore, the use of mathematical and computational tools to predict potential miRNA-disease associations based on various experimentally validated association datasets is a hot issue. Through the integration and collection of data from a large number of biological experiments, there are now multiple databases related to miRNA-disease relationships such as HMDD and dbDEMC (Lu et al., 2008; Yang et al., 2010; Li et al., 2014a). In recent years, a large number of miRNA-disease association prediction methods have been proposed. For instance, Chen and Yan (2014) proposed a regularized least squares model (RLSMDA) to predict miRNA-disease associations. This model is a semi-supervised model that learns in the miRNA space and disease space respectively, and then combines to get the final prediction score. However, it should be pointed out that the parameter selection of this model is more difficult, and the combined form of the two spatial scores can be improved in the end. Xu et al. (2011) proposed a method based on support vector machine (SVM) to predict the interaction between miRNA and the disease. However, the current database rarely provides data for non-cancer miRNAs. Therefore, the main problem of the model is the lack of negative samples, which will make the supervised learning model unsuitable for the prediction of large-scale disease-miRNA interactions. Obtaining large numbers of negatively associated samples is still difficult (Guan et al., 2020). Chen et al. (2012) adopted restart random walk (RWRMDA) to predict the potential miRNA-disease interaction, which restarted the known miRNA-disease interaction network, using random walks on miRNA functional similarity network to predict potential miRNA-disease interaction. However, this method is not applicable to the prediction of new diseases that are not related to any miRNA. Chen (2018) introduced the induction matrix completion model (IMCMDA) for the prediction of miRNA disease association based on the known miRNA-disease association matrix, miRNA functional similarity and disease semantic similarity matrix. However, this method is too sensitive to the noise in the data, which affects its performance. Chen et al. (2016b) introduced the model of Within and Between Score for MiRNA-Disease Association prediction (WBSMDA)

by a combination of integrated similarity and known miRNA-disease associations. Chen et al. (2018) introduced the MiRNA-disease association prediction (TLHNMDA) model based on three-layer heterogeneous network inference, which integrates multi-level data about miRNA, disease, lncRNA and their associated information into three layers heterogeneous network to determine the relationship between miRNA and disease Potential biological connection. Zhao et al. (2018) proposed a novel computational model of Symmetric Non-negative Matrix Factorization for MiRNA-Disease Association prediction (SNMFMDA) to reveal the relation of miRNA-disease pairs. Compared to the direct use of the integrated similarity in previous computational models, the integrated similarity needs to be interpolated by symmetric non-negative matrix factorization (SymNMF) before application in SNMFMDA. Jihwan Ha et al. (2020) present IMIPMF, a novel method for predicting miRNA-disease associations using probabilistic matrix factorization (PMF), which is a machine learning technique that is widely used in recommender systems. Zhu et al. (2020) proposed a new computational model based on biased heat conduction for MiRNA-Disease Association prediction (BHCMDA), which can achieve the AUC of 0.8890 in LOOCV.

We hope to use a simple and effective method for prediction. Here, we proposed a new matrix completion algorithm based on the q-kernel function to predict new miRNA disease associations (QIMCMDA). This model used miRNA q-kernel similarity, disease q-kernel similarity, known miRNA disease associations, and miRNA functional similarity. A matrix decomposition algorithm based on KL divergence was used to complement missing miRNA-disease associations. Here we used the receiver operating characteristic (ROC) curve as an evaluation index to evaluate the effectiveness of QIMCMDA. For known miRNA-disease associations downloaded from HMDD V2.0, the relevant data was cross-validated using the method of leave-one-out cross-validation (LOOCV) and fivefold cross-validation, and compared with the four previous classic methods (TLHNMDA, WBSMDA, RLSMDA, and IMCMDA). In addition, case studies were conducted on three common human diseases (Breast Neoplasms, Carcinoma Hepatocellular, Colon Neoplasms). All candidate miRNAs for these three diseases were ranked according to the predicted scores of QIMCMDA. Then the top 50 predicted miRNAs of these three diseases were verified in dbDEMC and HMDD 3.2 respectively. As a result, 46, 45, and 48 of the top 50 potentially relevant miRNAs for the three diseases were confirmed. These results indicated the effectiveness of QIMCMDA in predicting potential miRNA-disease associations.

## MATERIALS AND METHODS

### Human MiRNA-Disease Associations

In this study, we used human disease-miRNA associations in the HMDD v2.0 database, the dataset contains 383 diseases, 495 miRNAs, and 5430 high-quality experimentally verified human miRNA-diseases associations (Chen et al., 2018). We defined the

adjacency matrix  $A \in R^{nd*nm}$  as follows:

$$A(d(i), m(j)) = \begin{cases} 1 & \text{diseased } (i) \text{ has association with miRNA } m(j) \\ 0 & \text{diseased } (i) \text{ has no association with miRNA } m(j) \end{cases} \quad (1)$$

### MiRNA Functional Similarity

MiRNA functional similarity score was calculated by Wang et al. (2010) based on the hypothesis that similarly functional miRNAs tend to be associated with diseases with similar phenotypes. Thanks to their work, we obtained from <http://www.cuilab.cn/files/images/cuilab/misim.zip> downloaded the data. We constructed a matrix  $FS$ , where the matrix  $FS(m(i), m(j))$  represents the functional similarity between miRNAs  $m(i)$  and  $m(j)$ .

### Disease Semantic Similarity

#### Disease Semantic Similarity 1

A Directed Acyclic Graph ( $DAG$ ) was constructed to describe a disease based on the MeSH descriptors downloaded from the National Library of Medicine (Lipscomb, 2000). The  $DAG$  of disease  $D$  included not only the ancestor nodes of  $D$  and  $D$  itself but also the direct edges from parent nodes to child nodes. The semantic score of disease  $D$  could be defined by the following equation:

$$DV1(D) = \sum_{d \in T(D)} D1_D(d) \quad (2)$$

we defined the contribution score of disease  $d$  in  $DAG(D)$  to the disease  $D$  by:

$$\begin{cases} D1_D(d) = 1 & \text{if } d = D \\ D1_D(d) = \max \{ \Delta * D1_D(d') \mid d' \in \text{children of } d \} & \text{if } d \neq D \end{cases} \quad (3)$$

$\Delta$  is the semantic contribution factor. The contribution score of disease is decreased as the distance between  $D$  and other diseases increases. Based on the assumption that two diseases with larger shared area of their  $DAGs$  may have greater similarity score, the semantic similarity score between disease  $d(i)$  and disease  $d(j)$  could be defined by the following equation:

$$SS1(d(i), d(j)) = \frac{\sum_{t \in T(d(i)) \cap T(d(j))} (D1_{d(i)}(t) + D1_{d(j)}(t))}{DV1(d(i)) + DV1(d(j))} \quad (4)$$

#### Disease Semantic Similarity 2

From above formula (3), it is easy to see that the diseases in the same layer of  $DAG(D)$  will make the same contribution to the semantic value of  $D$ . Moreover, for diseases in the same layer of  $DAG(D)$ , it is reasonable to assume that the diseases appeared in fewer  $DAGs$  will be more specific than those diseases appeared in more  $DAGs$ . Hence, to protrude the contribution of these more specific diseases, the contribution of the node  $d$  in  $T(D)$  to the semantic value of the disease  $D$  could be obtained according to the following formula as well (Chen, 2018):

$$D2_D(d) = -\log \left[ \frac{\text{the number of DAGs containing } d}{\text{the number of diseases}} \right] \quad (5)$$

Based on the above formula, the semantic value of the disease  $D$  could be obtained according to the following formula as well:

$$DV2(D) = \sum_{d \in T(D)} D2_D(d) \quad (6)$$

Hence, the semantic similarity between two diseases  $d_i$  and  $d_j$  could be obtained according to the following formula as well:

$$SS1(d(i), d(j)) = \frac{\sum_{t \in T(d(i)) \cap T(d(j))} (D2_{d(i)}(t) + D2_{d(j)}(t))}{DV2(d(i)) + DV2(d(j))} \quad (7)$$

### q-Kernel Similarity

Many contributions indicate that the performance of kernel-based learning algorithms largely depends on the choice of kernel (Chapelle et al., 2002; Lanckriet et al., 2002; Nogayama et al., 2003). Boughorbel also proved through experiments that in some applications, kernels with only positive conditions may be better than most classical kernels (Boujemaa et al., 2005). Based on this theory, Zhang et al. (2019) designed a variety of q-Kernel Functions, such as Non-Linear q-Kernel, Gaussian q-Kernel, Laplacian q-Kernel, Rational Quadratic q-Kernel, Multiquadric q-Kernel, Inverse Multiquadric q-Kernel, Wave q-Kernel, and so on. A q-analog is a mathematical expression parameterized by a quantity  $q$  that generalizes a known expression and reduces to the known expression. Therefore, after a long period of trial, we have chosen the inverse quadratic square  $q$  kernel function as the main method for calculating similarity.

Here we introduce a q-Kernel function (inverse multiquadric q-Kernel) and construct a q-Kernel similarity. Based on the assumption that similar miRNAs are more likely to exhibit interactions with similar diseases and vice versa. The q-Kernel similarity is used to calculate the kernel similarity of miRNA and disease, respectively, based on known miRNA- diseases. The value range of the two parameters  $c$  and  $q$  of the function is between 0 and 1.

$$H_q(x, y) = \frac{1}{1-q} \left( q^{-\frac{1}{c}} - q^{-\frac{1}{\sqrt{|x-y|^2+c^2}}} \right) \quad (8)$$

### Similarity Calculation of miRNA Based on q-Kernel

In previous work, we obtained a similarity network between two miRNAs. But the integrity of this network is only 0.2058, and too many missing values make it impossible for us to use this network directly. Here, the q-kernel function is used to complete the matrix. First, the obtained q-kernel distance needs to be normalized and scaled to [0,1], because the similarity network value of the previous miRNA is between [0,1]. Then we used the  $1-H_q$  to convert the kernel distance into the similarity and a q-kernel similarity network of miRNA is obtained, which is called QM. The similarity of MiRNA is constructed as follows:

$$S_m(m(i), m(j)) = \begin{cases} \omega FS(m(i), m(j)) + (1-\omega)QM(m(i), m(j)) & \text{miRNA } m(i) \text{ and } m(j) \text{ has similarity} \\ QM(m(i), m(j)) & \text{otherwise} \end{cases} \quad (9)$$

The  $\omega$  is a weighting parameter defined as limiting the effect of FS and QM on miRNA similarity. Set  $\omega$  to 0.01 through training. The greater similarity between miRNAs, the more similar the miRNAs are.

### Network Similarity Calculation for Diseases Based on q-Kernel

We used the same method as the miRNA similarity network to build the disease similarity network QD. Then integrated QD with disease semantic similarities SS1 and SS2:

$$S_d(d(i), d(j)) = \begin{cases} \omega SS(d(i), d(j)) + (1 - \omega) QD(d(i), d(j)) & d(i) \text{ and } d(j) \\ & \text{has similarity} \\ QD(d(i), d(j)) & \text{otherwise} \end{cases} \quad (10)$$

$$SS(d(i), d(j)) = \frac{SS1(d(i), d(j)) + SS2(d(i), d(j))}{2} \quad (11)$$

We set the parameter values of c and q through training, that is,  $c = 0.1$  and  $q = 0.6$ . Finally, we obtained two kernel similarity matrices,  $S_m$  and  $S_d$ .

### Matrix Completion

After integrated various known data and similarity calculations of q-kernel, we can obtain human miRNA-disease correlation matrix A (Matrix density is 0.028), disease similarity matrix  $S_d$ , miRNA similarity matrix  $S_m$ . Our goal is to deduce undiscovered miRNA-disease associations based on this known information. Here we use  $S_d \in R^{nd \times nd}$  as the feature matrix of  $nd$  diseases, and  $S_m \in R^{nm \times nm}$  as the feature matrix for miRNAs.  $S_d(i)$  denote the feature vector of disease  $d(i)$ , and  $S_m(j)$  denote the feature vector of miRNA  $m(j)$ . The main idea of QIMCMDA is to complement the two feature matrices  $S_d$  and  $S_m$  by the similarity of the q-kernel, and then supplement the missing elements under the restriction of the association matrix A to obtain the potential associations. Finally, the recovery matrix Z is obtained, and the form of Z is  $Z = S_d W H^T S_m$ . where  $W \in R^{nd \times r}$  and  $H \in R^{r \times nm}$ ,  $r$  is the desired rank which is equal to  $\min(\text{rank}(W), \text{rank}(H))$ . The parameter  $r$  mainly affects the convergence speed of the algorithm, and has little effect on the results. The matrices  $W$  and  $H$  can be obtained as a solution to the following optimization problems.

$$\min_{W, H} \phi = \sum_{i=1}^{nd} \sum_{j=1}^{nm} (A_{ij} \ln \frac{A_{ij}}{S_d * W * H * S_m} - A_{ij} + (S_d * W * H * S_m)_{ij})$$

$s.t. W \geq 0, H \geq 0$  (13)

$W$  and  $H$  were set to random dense matrices, and then the alternating gradient descent method is used to update iterations  $W$  and  $H$ .

$$W \leftarrow \frac{W * \left[ \frac{S_d * A}{(S_d * W * H * S_m)} * S_m * H' \right]}{S_d * \text{ONES} * S_m * H'} \quad (14)$$

$$H \leftarrow \frac{H * \left[ W' * S_d * \left( \frac{A * S_m}{S_d * W * H * S_m} \right) \right]}{W' * S_d * \text{ONES} * S_m} \quad (15)$$

Through the alternating gradient descent algorithm,  $W$  and  $H$  will stabilize and stop the iteration after reaching the maximum number of iterations. Here, the maximum number of iterations

TABLE 1 | Notations.

Symbol	Description
$nm$	number of miRNAs
$nd$	number of diseases
$A \in R^{nd \times nm}$	miRNA-diseases associations matrix
$S_m \in R^{nm \times nm}$	miRNA similarity matrix
$S_d \in R^{nd \times nd}$	disease similarity matrix
$W \in R^{nd \times r}$	alternating iteration matrix in matrix factorization
$H \in R^{r \times nm}$	alternating iteration matrix in matrix factorization

is set to 100. ONES is a matrix, all its elements are 1. It is used to multiply two matrixes of different ranks. We can use  $W$  and  $H$  to calculate the predicted score between disease  $d(i)$  and miRNA  $m(j)$  by the following formula (Symbol meaning can refer to Table 1).

$$\text{Score}(d(i), m(j)) = S_d(i) W H S_m(j) \quad (16)$$

The specific implementation process of QIMCMDA is shown in Figure 1.

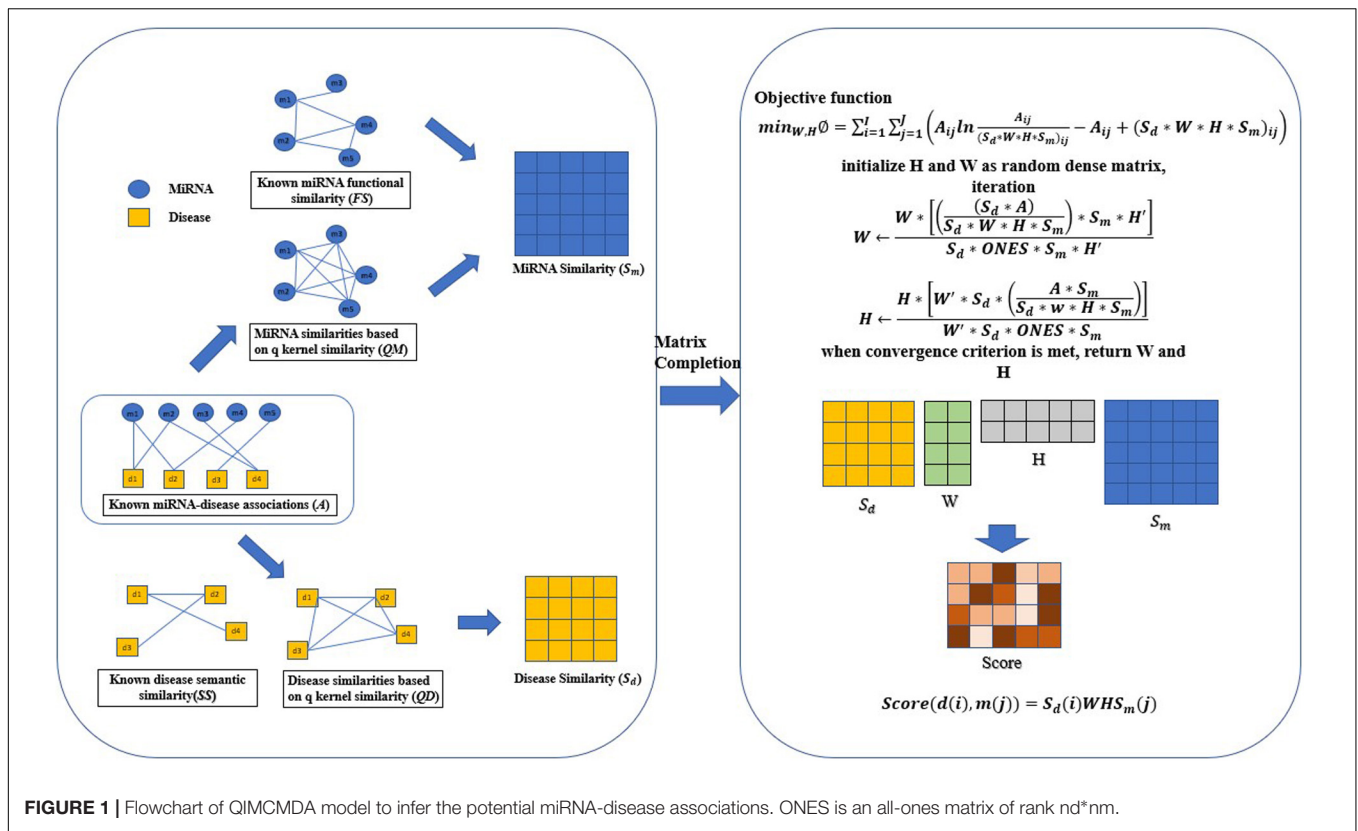
## RESULTS

We used 5,430 miRNA-disease associations from HMDD v2.0 as the gold standard dataset, and we used LOOCV and fivefold CV to test the effectiveness of QIMCMDA. In addition, QIMCMDA will be compared with four other methods IMCMDA (Chen, 2018), RLSMDA (Chen and Yan, 2014), TLHNMDA (Chen et al., 2018), WBSMDA (Chen et al., 2016b) to evaluate the predictive ability of QIMCMDA (see Table 2). In the framework of the LOOCV evaluation, 5430 miRNA-disease associations in the data set are considered as test samples one by one, the other remaining samples are considered as training samples, and samples with unknown associations are considered as candidate samples. Through the calculation of the model, we can obtain the prediction score, and then rank and record according to the prediction score. The process of fivefold CV is similar to LOOCV. The miRNA-disease association of the golden data set was randomly divided into five groups, one of which was selected as the test set in turn, and the rest as the training set. Candidate sample settings are the same as LOOCV. Then rank and record the predicted scores for each test sample. Figure 2 shows a comparison of the prediction performance based on the overall AUC value of LOOCV. As a result of LOOCV, the AUC of QIMCMDA is 0.9235, and the AUC values obtained by IMCMDA, RLSMDA, TLHNMDA and WBSMDA are 0.8378, 0.8193, 0.8795, 0.8010, respectively. For fivefold QIMCMDA, IMCMDA, RLSMDA, TLHNMDA and WBSMDA 10 times were performed, and the average AUC and standard deviation were recorded as  $0.9170 \pm 0.0006$ ,  $0.8311 \pm 0.0006$ ,  $0.7814 \pm 0.0020$ ,  $0.8735 \pm 0.0010$ ,  $0.7980 \pm 0.0009$ , respectively (see Figure 3).

### Parameter Analysis

There are several hyper-parameters in QIMCMDA that need to be tuned, i.e., c, q, w, k. We use a random search strategy





**FIGURE 1** | Flowchart of QIMCMDA model to infer the potential miRNA-disease associations. ONES is an all-ones matrix of rank  $nd * nm$ .

**TABLE 2** | Under the fivefold CV and LOOCV verification framework, the performance of QIMCMDA and other benchmark methods.

Methods	LOOCV	Fivefold CV
QIMCMDA	0.9235	0.9170 ± 0.0006
IMCMDA	0.8378	0.8311 ± 0.0006
RLSMDA	0.8193	0.7814 ± 0.0020
TLHNMDA	0.8795	0.8735 ± 0.0010
WBSMDA	0.8010	0.7980 ± 0.0009

to select hyper-parameters from fixed ranges (Zhang et al., 2020).  $c$  and  $q$  are parameters for adjusting the q-Kernel function. In this study, the value of  $c$  is selected from  $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$ , and the value of  $q$  is selected from  $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$ .  $q$  can't be equal to 1.  $\omega$  is the weight parameter used to integrate similarity. Here,  $\omega$  is selected from  $\{0.01, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.8, 1\}$ . Next, we show the influence of the these parameters under the fivefold CV.

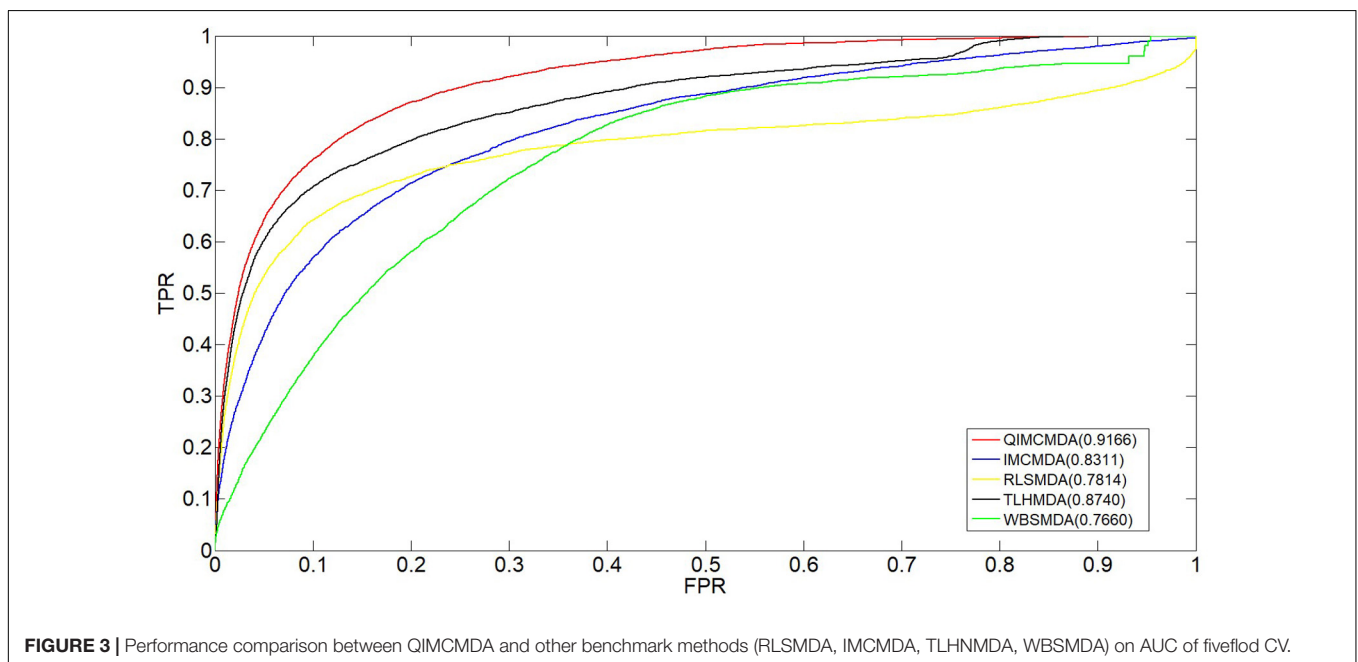
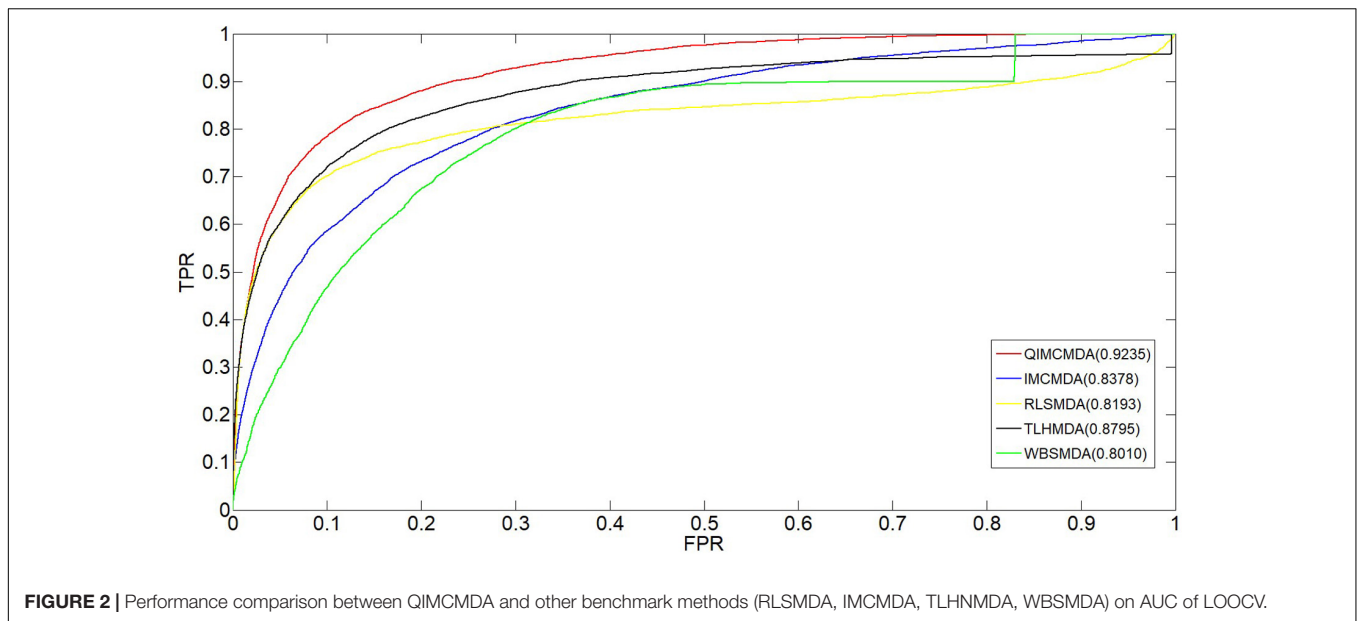
The  $k$  is a potential feature size. In our test, the impact of this variable is actually very small, but we still decided to use PCA to calculate the cumulative contribution rate to obtain the most appropriate  $k$  value. This method is in the paper by Wang et al. (2017). It has been well-verified. In this article, the cumulative contribution rate of 95% is used to select the PC, and the final  $k$  is set 114.

$\omega$  is a weight parameter used to integrate the similarity matrix.

**Figure 4** shows the effect of changes in  $\omega$  on AUC when other parameters are fixed. When  $\omega = 0.01$ , AUC takes the maximum value. When  $c = 0.1$ ,  $q = 0.6$ , the model can achieve the best effect (see **Figure 5**).

### Case Study

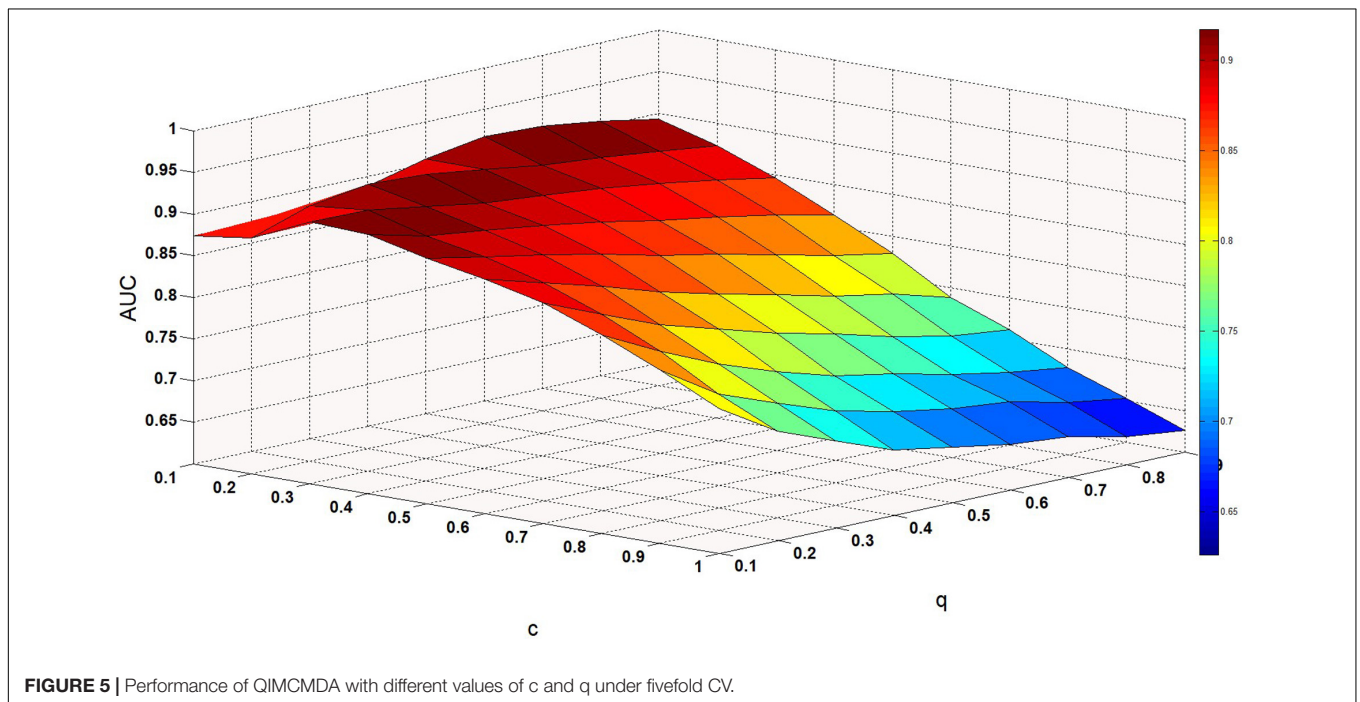
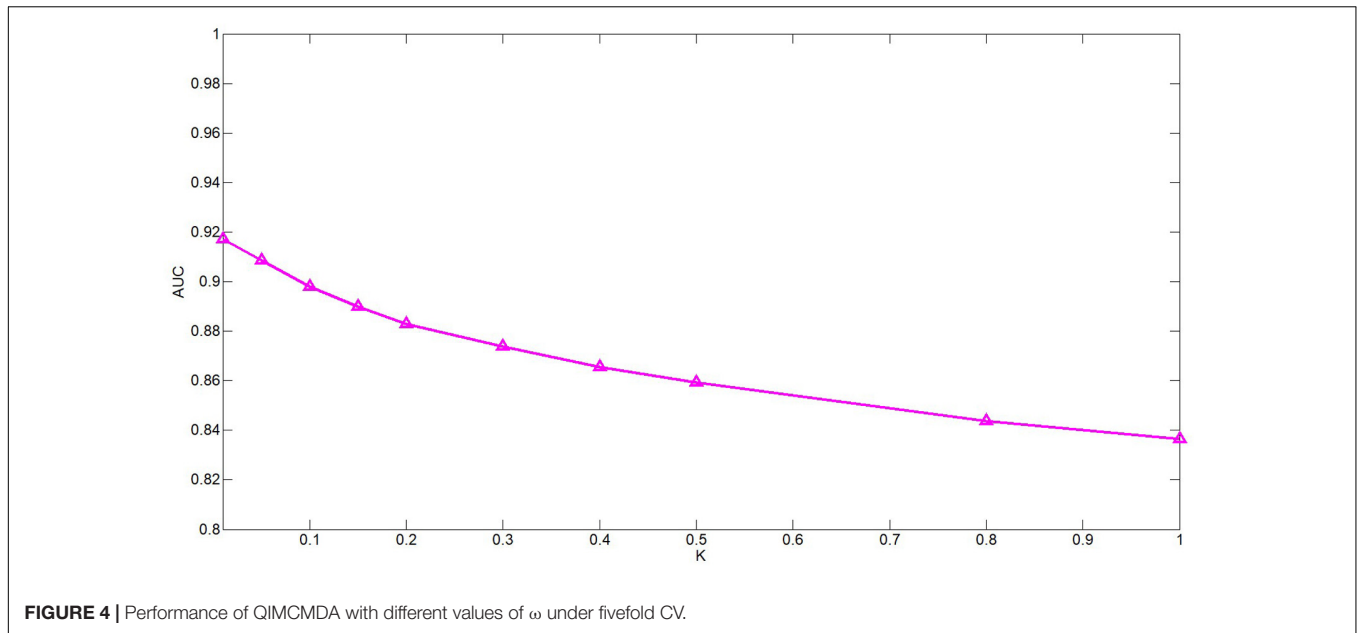
In this article, we used case studies to further demonstrate the effectiveness of QIMCMDA. We performed case studies on three diseases: Breast Neoplasms, Carcinoma Hepatocellular, and Colon Neoplasms. These diseases were selected in our case study because they all have high incidence and insignificant early symptoms. In addition, they have been considered as case studies in many previous publications (Guan et al., 2020). Our case study used HMDD v2.0 as the training database for QIMCMDA. HMDD 3.2 and dbDEMC (Lu et al., 2008; Yang et al., 2010; Li et al., 2014a) serve as validation databases to confirm the predicted potential associations. Compared with the previous 2.0 version, the 3.2 version contains more than double the association between human diseases and miRNAs, the classification of evidence is more clear, and there is a clear third-party annotation for each association. The differentially expressed miRNA database (dbDEMC) in human cancer is a comprehensive database microRNA (miRNA) designed to store and display differentially expressed human cancers detected by high-throughput methods. The database collected a total of 209 newly released data sets from Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA). The current version contains data from 436 biological experiments, including 2224 differentially expressed miRNAs in



36 cancer types. We only perform ranking verification on candidate miRNAs of interest, so training samples are not in the final result. In other words, the miRNA disease associations obtained from the predicted list do not overlap with the known 5430 associations.

Breast Neoplasms is one of the most common malignancies in women. With more than 2 million new cases worldwide each year, it ranks second among the world's major cancer types (Jemal et al., 2017). More than half of these cases occurred in industrialized countries (Parkin et al., 2005). It was one of the leading causes of death among women aged 20–59 (Siegel et al., 2015). With the development of biological technology, researchers have found more miRNAs related to Breast Neoplasms. Our results are supported by

third-party annotations in two databases, HMDD3.2 and dbDEMC. For example, miR-150 and miR-372 can promote the proliferation and growth of Breast Neoplasms cells by targeting the pro-apoptotic purinergic P2X7 receptor and LATS2 respectively (Huang et al., 2017; Cheng et al., 2018). MicroRNA-130a targets RAB5A to inhibit the proliferation, invasion and migration of Breast Neoplasms cells (Pan et al., 2015). miR-494 targets CXCR4 through the Wnt/ $\beta$ -catenin signaling pathway, thereby inhibiting Breast Neoplasms progression *in vitro* (Song et al., 2015). The increased miR-451 expression may negatively regulate Bcl-2 mRNA and protein expression, which in turn affects caspase 3 protein expression and accelerates Breast Neoplasms cell apoptosis (Gu et al., 2015). MiR-449a inhibits cell migration and invasion in Breast Neoplasms by targeting PLAGL2



(Wang et al., 2018). We selected the top 50 in the results and verified them with two databases, HMDD 3.2 and dbDEM. It was found that 10 of the first 10 predictions and 46 miRNAs of the first 50 predictions were verified (see **Table 3**).

Hepatocellular carcinoma (HCC), one of the most common malignancies worldwide (Yegin et al., 2016), was also the main cause of cancer in men under 60 in China (Chen et al., 2016a). MiRNAs have important roles in the treatment of HCC and have been corroborated. For example, related *in vitro* experiments have further confirmed the anti-tumor effect of miR-132 in HCC (Liu et al., 2015; Zhang et al., 2016).

The newly identified miR-429-CRKL axis represents a new potential therapeutic target for HCC therapy (Guo et al., 2018). MicroRNA-23b inhibits epithelial-mesenchymal transition (EMT) and metastasis of Hepatocellular Carcinoma by targeting Pyk2 (Cao et al., 2017). MicroRNA-494 is a major epigenetic regulator of microRNAs for multiple invasion inhibitors by targeting 10 11 translocation 1 in aggressive human Hepatocellular Carcinoma (Chuang et al., 2005). MicroRNA-340 inhibits the proliferation and invasion of Hepatocellular Carcinoma cells by targeting JAK1 (Yuan et al., 2017). Therefore, 10 of the top 10 predicted miRNAs

**TABLE 3 |** Prediction results of the top 50 predicted Breast Neoplasms-related miRNAs based on known associations in HMDD V2.0.

miRNA	Evidence	miRNA	Evidence
hsa-mir-151	HMDD3.2	hsa-mir-663	dbDEMC
hsa-mir-30e	HMDD3.2	hsa-mir-382	dbDEMC
hsa-mir-92b	HMDD3.2	hsa-mir-494	HMDD3.2
hsa-mir-451	HMDD3.2	hsa-mir-575	HMDD3.2
hsa-mir-130a	HMDD3.2	hsa-mir-658	dbDEMC
hsa-mir-192	HMDD3.2	hsa-mir-181d	dbDEMC
hsa-mir-98	HMDD3.2	hsa-mir-376a	HMDD3.2
hsa-mir-372	HMDD3.2	hsa-mir-211	dbDEMC
hsa-mir-32	HMDD3.2	hsa-mir-484	HMDD3.2
hsa-mir-106a	HMDD3.2	hsa-mir-455	Unconfirmed
hsa-mir-130b	HMDD3.2	hsa-mir-432	dbDEMC
hsa-mir-99b	dbDEMC	hsa-mir-381	HMDD3.2
hsa-mir-95	dbDEMC	hsa-mir-99a	HMDD3.2
hsa-mir-28	dbDEMC	hsa-mir-154	dbDEMC
hsa-mir-150	HMDD3.2	hsa-mir-523	dbDEMC
hsa-mir-186	dbDEMC	hsa-mir-526b	HMDD3.2
hsa-mir-15b	HMDD3.2	hsa-mir-507	Unconfirmed
hsa-mir-142	HMDD3.2	hsa-mir-525	Unconfirmed
hsa-mir-449b	dbDEMC	hsa-mir-660	HMDD3.2
hsa-mir-198	dbDEMC	hsa-mir-181c	HMDD3.2
hsa-mir-196b	HMDD3.2	hsa-mir-300	dbDEMC
hsa-mir-491	HMDD3.2	hsa-mir-297	dbDEMC
hsa-mir-449a	HMDD3.2	hsa-mir-136	dbDEMC
hsa-mir-424	HMDD3.2	hsa-mir-331	HMDD3.2
hsa-mir-212	HMDD3.2	hsa-mir-512	Unconfirmed

**TABLE 4 |** Prediction results of the top 50 predicted Carcinoma Hepatocellular-related miRNAs based on known associations in HMDD V2.0.

miRNA	Evidence	miRNA	Evidence
hsa-mir-132	HMDD3.2	hsa-mir-516a	unconfirmed
hsa-mir-429	HMDD3.2	hsa-mir-663	dbDEMC
hsa-mir-34b	HMDD3.2	hsa-mir-340	HMDD3.2
hsa-mir-151	HMDD3.2	hsa-mir-28	dbDEMC
hsa-mir-30e	HMDD3.2	hsa-mir-186	HMDD3.2
hsa-mir-367	HMDD3.2	hsa-mir-575	HMDD3.2
hsa-mir-339	dbDEMC	hsa-mir-658	dbDEMC
hsa-mir-9	HMDD3.2	hsa-mir-452	HMDD3.2
hsa-mir-215	HMDD3.2	hsa-mir-193b	HMDD3.2
hsa-mir-451	HMDD3.2	hsa-mir-196b	dbDEMC
hsa-mir-194	HMDD3.2	hsa-mir-494	HMDD3.2
hsa-mir-302a	dbDEMC	hsa-mir-449a	HMDD3.2
hsa-mir-32	HMDD3.2	hsa-mir-424	HMDD3.2
hsa-mir-204	HMDD3.2	hsa-mir-520c	HMDD3.2
hsa-mir-135b	HMDD3.2	hsa-mir-382	unconfirmed
hsa-mir-95	HMDD3.2	hsa-mir-301b	dbDEMC
hsa-mir-488	dbDEMC	hsa-mir-510	unconfirmed
hsa-mir-302d	HMDD3.2	hsa-mir-376c	unconfirmed
hsa-mir-23b	HMDD3.2	hsa-mir-455	HMDD3.2
hsa-mir-133a	HMDD3.2	hsa-mir-206	HMDD3.2
hsa-mir-299	HMDD3.2	hsa-mir-137	HMDD3.2
hsa-mir-143	HMDD3.2	hsa-mir-211	HMDD3.2
hsa-mir-153	HMDD3.2	hsa-mir-154	HMDD3.2
hsa-mir-516b	Unconfirmed	hsa-mir-27b	HMDD3.2
hsa-mir-383	dbDEMC	hsa-mir-523	dbDEMC

**TABLE 5 |** Prediction results of the top 50 predicted Colon Neoplasms-related miRNAs based on known associations in HMDD V2.0.

miRNA	Evidence	miRNA	Evidence
hsa-mir-143	HMDD3.2	hsa-mir-200b	HMDD3.2
hsa-mir-106b	HMDD3.2	hsa-mir-24	HMDD3.2
hsa-mir-21	HMDD3.2	hsa-mir-1	HMDD3.2
hsa-mir-128	HMDD3.2	hsa-mir-205	HMDD3.2
hsa-mir-18a	HMDD3.2	hsa-mir-29b	HMDD3.2
hsa-mir-9	dbDEMC	hsa-let-7b	HMDD3.2
hsa-mir-155	HMDD3.2	hsa-mir-31	HMDD3.2
hsa-mir-181a	HMDD3.2	hsa-mir-223	HMDD3.2
hsa-mir-494	unconfirmed	hsa-let-7c	HMDD3.2
hsa-mir-483	HMDD3.2	hsa-mir-15a	HMDD3.2
hsa-let-7a	HMDD3.2	hsa-mir-200c	HMDD3.2
hsa-mir-125b	HMDD3.2	hsa-mir-222	HMDD3.2
hsa-mir-146a	HMDD3.2	hsa-mir-199a	HMDD3.2
hsa-mir-34a	HMDD3.2	hsa-mir-30b	HMDD3.2
hsa-mir-210	HMDD3.2	hsa-mir-141	HMDD3.2
hsa-mir-16	HMDD3.2	hsa-mir-200a	HMDD3.2
hsa-mir-146b	dbDEMC	hsa-let-7e	HMDD3.2
hsa-mir-221	HMDD3.2	hsa-mir-196a	HMDD3.2
hsa-mir-93	HMDD3.2	hsa-mir-142	HMDD3.2
hsa-mir-92a	HMDD3.2	hsa-let-7f	HMDD3.2
hsa-mir-20b	dbDEMC	hsa-mir-34c	Unconfirmed
hsa-mir-19a	HMDD3.2	hsa-let-7i	HMDD3.2
hsa-mir-29a	HMDD3.2	hsa-let-7d	HMDD3.2
hsa-mir-18b	HMDD3.2	hsa-let-7g	HMDD3.2

and 45 of the top 50 predicted miRNAs were confirmed by experimental literature from the dbDEMC and HMDD3.2 (see Table 4).

Colon Neoplasms are the most common type of gastrointestinal cancer (Jemal et al., 2011; Ogata-Kawata et al., 2014). Siegel et al. (2018), there were 97,220 new cases in the United States alone, and approximately 50,630 patients died. A variety of miRNAs have been experimentally confirmed to be associated with colon neoplasms. For example, MicroRNA-155 regulates Colon Neoplasms cell proliferation, cell cycle, apoptosis, migration and targets CBL (Yu et al., 2017). MicroRNA-21 induces stem cells by down-regulating transforming growth factor beta receptor 2 (TGFbetaR2) in Colon Neoplasms cells (Yu et al., 2012). Let-7 is also involved in the development of Colon Neoplasms (Williams, 2008). MicroRNA-221 promotes Colon Neoplasms cell proliferation *in vitro* (Sun et al., 2011). MicroRNA-34a inhibits the migration and invasion of Colon Neoplasms cells by targeting Fra-1 (Wu et al., 2012). Verification of dbDEMC and HMDD3.2 confirmed 10 of the first 10 predictions and 48 miRNAs of the first 50 predictions (see Table 5).

## DISCUSSION

Research on the potential prediction of miRNA-disease associations will help us to understand the pathogenesis and treatment of the disease more deeply. Especially for cancer, targeted therapy by regulating miRNA may be a breakthrough point for future treatment.



In this paper, we developed an algorithm for miRNA-disease association prediction (QIMCMDA), which mainly introduced the q-kernel function to complete the similarity information required. The QIMCMDA model is based on the known miRNA disease association and miRNA functional similarity network. First, calculated and completed the miRNA similarity network and the disease similarity network using the q-kernel function. Then used the matrix decomposition method to calculate the prediction score for each sample, and finally sort the scores. The AUC of QIMCMDA based on LOOCV is 0.9235, showing better performance than previous methods. In addition, experimental literature has confirmed the validity of potential miRNA-disease association predictions for three major human diseases: Breast Neoplasms, Carcinoma Hepatocellular, Colon Neoplasms).

The reasons for the reliable performance of QIMCMDA are as follows: the key advantage of QIMCMDA is that it utilizes the functional similarity of known miRNAs in combination with q-kernel similarity as features of diseases and miRNAs to complete the association of missing miRNAs and diseases. And the use of alternating gradient descent algorithm to search for the optimal solution can ensure the reliability of disease feature vectors and miRNA feature vectors. In addition, the overall complexity of our method from the construction of the network to the final prediction score calculation is low, and the operation is simple and easy to reproduce. QIMCMDA has a short running time and is suitable for large-scale data research. It is a simple and effective method. Finally, QIMCMDA is a semi-supervised model that does not require negative samples, reducing the difficulty of model construction. Compared with methods that require a large number of negative samples, our method has some advantages. However, QIMCMDA currently has some limitations. First of all, there are inevitable noises and outliers in the known materials we use. Second, QIMCMDA used the KL divergence as an error function, which is unstable

due to noise and outliers. With the development of the times, database construction will become more and more perfect. As the number of associated data increases, our predictions will become more accurate. In addition, for miRNA or disease without any known associations, our method may be less effective, because the calculation of q-kernel is mainly based on known associations. In the future, we can use a large amount of biological data to further increase the reliability and practicability of the model prediction. And our method can be practiced in other fields such as the interaction between microorganisms and diseases or the interaction between drugs and targets.

## DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

## AUTHOR CONTRIBUTIONS

LW and YZ conceived the study. LW, YZ, and YC developed the prediction method and designed the experiments. LW analyzed the result and wrote the manuscript. NZ and WC optimized the flow chart and manuscript structure. All authors reviewed and improved the manuscript.

## FUNDING

This work has been supported by the National Natural Science Foundation of China (under Grant Nos. 61877064, U1806202, and 61533011).

## REFERENCES

- Ambros, V. (2003). MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. *Cell* 113, 673–676. doi: 10.1016/S0092-8674(03)00428-8
- Asangani, I. A., Rasheed, S. A. K., Nikolova, D. A., Leupold, J. H., Colburn, N. H., Post, S., et al. (2008). MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdc4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene* 27, 2128–2136. doi: 10.1038/sj.onc.1210856
- Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233. doi: 10.1016/j.cell.2009.01.002
- Boujemaa, N., Tarel, J., and Boughorbel, S. (2005). “Conditionally positive definite kernels for svm based image recognition,” in *IEEE International Conference on Multimedia and Expo(ICME)*, Amsterdam, 113–116. doi: 10.1109/ICME.2005.1521373
- Calin, G. A., and Croce, C. M. (2006). MicroRNA signatures in human cancers. *Nat. Rev. Cancer* 6, 857–866. doi: 10.1038/nrc1997
- Cao, J., Liu, J. K., Long, J. Y., Fu, J., Huang, L., Li, J., et al. (2017). MicroRNA-23b suppresses epithelial-mesenchymal transition (EMT) and metastasis in hepatocellular carcinoma via targeting Pyk2. *Biomed. Pharmacother.* 17:30 doi: 10.1016/j.biopha.2017.02.030
- Carleton, M., Cleary, M. A., and Linsley, P. S. (2007). MicroRNAs and cell cycle regulation. *Cell Cycle* 6, 2127–2132. doi: 10.4161/cc.6.17.4641
- Catto, J. W. F., Alcaraz, A., Bjartell, A. S., White, R. D. V., Evans, C. P., Fussell, S., et al. (2011). MicroRNA in prostate, bladder, and kidney Cancer: a systematic review. *Eur. Urol.* 59, 671–681. doi: 10.1016/j.eururo.2011.01.044
- Chapelle, O., Vapnik, V., Bousquet, O., and Mukherjee, S. (2002). Choosing multiple parameters for support vector machines. *Mach. Learn.* 46, 131–159. doi: 10.1023/A:1012450327387
- Chen, C. Z., Li, L., Lodish, H. F., and Bartel, D. P. (2004). MicroRNAs modulate hematopoietic lineage differentiation. *Science* 303:903. doi: 10.1126/science.1091903
- Chen, J. F., Mandel, E. M., Thomson, J. M., Wu, Q., and Wang, D. Z. (2006). The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nature* 38:1725. doi: 10.1038/ng1725
- Chen, W., Zheng, R., Baade, P. D., Zhang, S., Zeng, H., Bray, F., et al. (2016a). Cancer statistics in China, 2015. *CA Cancer J. Clin.* 66, 115–132. doi: 10.3322/caac.21338
- Chen, X. (2018). IMCMDA: predicting miRNA-disease association based on inductive matrix completion. *Bioinformatics* 34, 4256–4265. doi: 10.1093/bioinformatics/bty503
- Chen, X., Liu, M., and Yan, G. (2012). RWRMDA: predicting novel human microRNA-disease associations. *Mol. BioSyst.* 8, 2792–2798. doi: 10.1039/c2mb25180a
- Chen, X., Qu, J., and Yin, J. (2018). TLHNMDA: triple layer heterogeneous network based inference for MiRNA-Disease association prediction. *Front. Genet.* 18:234. doi: 10.3389/fgene.2018.00234
- Chen, X., Yan, C. C., Zhang, X., You, Z. H., Deng, L. X., Liu, Y., et al. (2016b). WBSMDA: within and between score for MiRNA-disease association prediction. *Sci. Rep.* 6:21106. doi: 10.1038/srep21106
- Chen, X., and Yan, G. Y. (2014). Semi-supervised learning for potential human microRNA-disease associations inference. *Sci. Rep.* 4:5501. doi: 10.1038/srep05501

- Cheng, X. Y., Chen, J. Q., and Huang, Z. (2018). MiR-372 promotes breast cancer cell proliferation by directly targeting LATS2. *Exp. Therap. Med.* 15:5761. doi: 10.3892/etm.2018.5761
- Cho, W. C. (2010). MicroRNAs: potential biomarkers for cancer diagnosis, prognosis and targets for therapy. *Int. J. Biochem. Cell Biol.* 42, 1273–1281. doi: 10.1016/j.biocel.2009.12.014
- Chuang, K. H., Whitney-Miller, C. L., Chu, C.-Y., Zhou, Z., Dokus, K. M., Schmit, S., et al. (2005). MicroRNA-494 is a master epigenetic regulator of multiple invasion-suppressor microRNAs by targeting ten eleven translocation 1 in invasive human hepatocellular carcinoma neoplasms. *Hepatology* 62:27816. doi: 10.1002/hep.27816
- Farazi, T. A., Hoell, J. I., Morozov, P., and Tuschl, T. (2013). MicroRNAs in human cancer. *Adv. Exp. Med. Biol.* 774, 1–20. doi: 10.1007/978-94-007-5590-1\_1
- Filipowicz, W., Bhattacharyya, S. N., and Sonenberg, N. (2008). Mechanisms of posttranscriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.* 9, 102–114. doi: 10.1038/nrg2290
- Gu, X., Li, J. Y., Guo, J., Li, P. S., and Zhang, W. H. (2015). Influence of MiR-451 on drug resistances of paclitaxel-resistant breast cancer cell line. *Med. Sci. Monit.* 21:894475. doi: 10.12659/MSM.894475
- Guan, N. N., Wang, C. C., Zhang, L., Huang, L., Li, J. Q., and Piao, X. (2020). In silico prediction of potential miRNA–disease association using an integrative bioinformatics approach based on kernel fusion. *J. Cell Mol. Med.* 24, 573–587. doi: 10.1111/jcmm.14765
- Guay, C., Roggli, E., Nesca, V., Jacovetti, C., and Regazzi, R. (2011). Diabetes mellitus, a microRNA-related disease? *Transl. Res.* 157, 253–264. doi: 10.1016/j.trsl.2011.01.009
- Guo, C. M., Zhao, D. T., Zhang, Q. L., Liu, S. Q., and Sun, M. S. (2018). MiR-429 suppresses tumor migration and invasion by targeting CRKL in hepatocellular carcinoma via inhibiting Raf/MEK/ERK pathway and epithelial-mesenchymal transition. *Sci. Rep.* 18:8 doi: 10.1038/s41598-018-20258-8
- Ha, J., Park, C. H., Park, C. Y., and Park, S. (2020). IMIPMF: inferring miRNA-disease interactions using probabilistic matrix factorization. *J. Biomed. Inform.* 102:103358. doi: 10.1016/j.jbi.2019.103358
- Huang, S. Y., Chen, Y. S., Wu, W., Ouyang, N. Y., Chen, J. N., Li, H. Y., et al. (2017). MiR-150 promotes human breast cancer growth and malignant behavior by targeting the pro-apoptotic purinergic P2X7 receptor. *PLoS One* 8:707. doi: 10.1371/journal.pone.0080707
- Iorio, M. V., Ferracin, M., Liu, C. G., Veronese, A., Spizzo, R., Sabbioni, S., et al. (2005). MicroRNA gene expression deregulation in human breast cancer. *Cancer Res.* 65:70657070. doi: 10.1158/0008-5472.CAN-05-1783
- Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., Forman, D., et al. (2011). Global cancer statistics. *CA Cancer J. Clin.* 61, 69–90. doi: 10.3322/caac.20107
- Jemal, A., Ward, E. M., Johnson, C. J., Cronin, K. A., Ma, J., Ryerson, B., et al. (2017). Annual report to the nation on the status of cancer, 1975–2014, featuring survival. *J. Natl. Cancer Inst.* 17:30. doi: 10.1093/jnci/djx030
- Jiang, Q. H., Hao, Y. Y., Wang, G., Juan, L. R., Zhang, T. J., Teng, M. X., et al. (2010). Prioritization of disease microRNAs through a human phenome-microRNAome network. *BioMed. Central* 4:S2. doi: 10.1186/1752-0509-4-S1-S2
- Jopling, C. L., Yi, M., 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
- Kozomara, A., and Griffiths-Jones, S. (2011). miRbase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* 39, D152–D157. doi: 10.1093/nar/gkq1027
- Lanckriet, G. R. G., Christianini, N., Bartlett, P. L., Ghaoui, L. E., and Jordan, M. I. (2002). “Learning the kernel matrix with semi-definite programming,” in *Nineteenth International Conference on Machine Learning*, Sydney, 323–330. doi: 10.1023/B:JODS.0000012018.62090.a7
- 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* 75:90529. doi: 10.1016/0092-8674(93)90529-Y
- Leung, A. K., and Sharp, P. A. (2010). MicroRNA functions in stress responses. *Mol. Cell.* 40, 205–215. doi: 10.1016/j.molcel.2010.09.027
- Li, Y., Qiu, C., Tu, J., Geng, B., Yang, J., Jiang, T., et al. (2014a). HMDD v2.0: a database for experimentally supported human microRNA and disease associations. *Nucleic Acids Res.* 42, D1070–D1074. doi: 10.1093/nar/gkt1023
- Li, Y., Zhang, Z., Mao, Y., Jin, M., Jing, F., Ye, Z., et al. (2014b). A genetic variant in MiR146a modifies digestive system Cancer risk: a meta-analysis. *Asian Pac. J. Cancer Prev.* 15, 145–150. doi: 10.7314/APJCP.2014.15.1.145
- Lipscomb, C. E. (2000). Medical subject headings (MeSH). *Bull. Med. Libr. Assoc.* 88, 265–266. doi: 10.0000/PMID10928714
- Liu, K., Li, X. L., Cao, Y. C., Ge, Y. Y., Wang, J. M., and Shi, B. (2015). MiR-132 inhibits cell proliferation, invasion and migration of hepatocellular carcinoma by targeting PIK3R3. *Int. J. Oncol.* 15:3112 doi: 10.3892/ijo.2015.3112
- Lu, M., Zhang, Q. P., Deng, M., Miao, J., Guo, Y. H., Gao, W., et al. (2008). An analysis of human MicroRNA and disease associations. *PLoS One* 3:3420. doi: 10.1371/journal.pone.0003420
- Niu, Y. W., Wang, G. H., Yan, G. Y., and Chen, X. (2019). Integrating random walk and binary regression to identify novel miRNA-disease association. *BMC Bioinformatics* 20:59. doi: 10.1186/s12859-019-2640-9
- Nogayama, T., Takahashi, H., and Muramatsu, M. (2003). Generalization of kernel pca and automatic parameter tuning. *Techn. Report Ieice Prmu* 103, 43–48.
- Nunez-Iglesias, J., Liu, C. C., Morgan, T. E., Finch, C. E., and Zhou, X. J. (2010). Joint genome-wide profiling of miRNA and mRNA expression in Alzheimer’s disease cortex reveals altered miRNA regulation. *PLoS One* 5:8898. doi: 10.1371/journal.pone.0008898
- Ogata-Kawata, H., Izumiya, M., Kurioka, D., Honma, Y., Yamada, Y., Furuta, K., et al. (2014). Circulating exosomal microRNAs as biomarkers of colon cancer. *PLoS One* 14:921. doi: 10.1371/journal.pone.0092921
- Pan, Y. Q., Wang, R. J., Zhang, F. W., Chen, Y. L., Lv, Q. F., Long, G., et al. (2015). MicroRNA-130a inhibits cell proliferation, invasion and migration in human breast cancer by targeting the RAB5A. *Int. J. Clin. Exp. Pathol.* 8, 384–393.
- Parkin, D. M., Bray, F., Ferlay, J., and Pisani, P. (2005). Global cancer statistics, 2002. *CA Cancer J. Clin.* 55, 74–108. doi: 10.3322/canjclin.55.2.74
- Petrocca, F., Visone, R., Onelli, M. R., Shah, M. H., Nicoloso, M. S., Martino, I. D., et al. (2008). E2F1-regulated microRNAs impair TGF $\beta$ -dependent cell-cycle arrest and apoptosis in gastric cancer. *Cancer Cell* 13, 272–286. doi: 10.1016/j.ccr.2008.02.013
- Siegel, R. L., Kimberly D M., and Jemal, A. (2018). Cancer statistics, 2018. *CA Cancer J. Clin.* 68, 7–30. doi: 10.3322/caac.21442
- Siegel, R. L., Miller, K. D., and Jemal, A. (2015). Cancer statistics, 2015. *CA Cancer J. Clin.* 65, 5–29. doi: 10.3322/caac.21208
- Song, L. Q., Liu, D., Wang, B. F., He, J. J., Zhang, Q. Q., Dai, Z. J., et al. (2015). MiR-494 suppresses the progression of breast cancer in vitro by targeting CXCR4 through the Wnt/ $\beta$ -catenin signaling pathway. *Oncol. Rep.* 34:3965. doi: 10.3892/or.2015.3965
- Sun, K., Wang, W., Lei, S. T., Wu, C. T., and Li, G. X. (2011). MicroRNA-221 promotes colon carcinoma cell proliferation in vitro by inhibiting CDKN1C/p57 expression. *J. South. Med. Univ.* 11:2011. doi: 10.1038/cmi.2011.4
- Taganov, K. D., Boldin, M. P., Chang, K. J., and Baltimore, D. (2006). NF- $\kappa$ B-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12481–12486. doi: 10.1073/pnas.0605298103
- Tricoli, J. V., and Jacobson, J. W. (2007). MicroRNA: potential for Cancer detection, diagnosis, and prognosis. *Cancer Res.* 67, 4553–4555. doi: 10.1158/0008-5472.CAN-07-0563
- Urbich, C., Kuehbach, A., and Dimmeler, S. (2008). Role of microRNAs in vascular diseases, inflammation, and angiogenesis. *Cardiovasc Res.* 79, 581–588. doi: 10.1093/cvr/cvn156
- Wang, D., Wang, J., Lu, M., and Song, F. (2010). Qinghua cui. inferring the human microRNA functional similarity and functional network based on microRNA-associated diseases. *Bioinformatics* 26:241. doi: 10.1093/bioinformatics/btq241
- Wang, H. L., Xiao, Y., Wu, L., and Ma, D. C. (2018). Comprehensive circular RNA profiling reveals the regulatory role of the circRNA-000911/miR-449a pathway in breast carcinogenesis. *Int. J. Oncol.* 18:4265. doi: 10.3892/ijo.2018.4265
- Wang, W. J., Chen, X., Jiao, P. F., and Jin, D. (2017). Similarity-based regularized latent feature model for link prediction in bipartite networks. *Sci. Rep.* 7:9. doi: 10.1038/s41598-017-17157-9
- Williams, A. E. (2008). Functional aspects of animal microRNAs. *Cell. Mol. Life Sci.* 8:9. doi: 10.1007/s00018-007-7355-9
- Wu, J. M., Wu, G., Lv, L., Ren, Y. F., Zhang, X. J., Xue, Y. F., et al. (2012). MicroRNA-34a inhibits migration and invasion of colon cancer cells via targeting to Fra-1. *Carcinogenesis* 12:304. doi: 10.1093/carcin/bgr304

- Xu, J., Li, C. X., Lv, J. Y., Li, Y. S., Huan, R., Xiao, Y., et al. (2011). Prioritizing candidate disease miRNAs by topological features in the miRNA target-dysregulated network: case study of prostate cancer. *Mol. Cancer Therap.* 10:55. doi: 10.1158/1535-7163.MCT-11-0055
- 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 Genomic* 10:S5. doi: 10.1186/1471-2164-11-S4-S5
- Yegin, E. G., Oymaci, E., Karatay, E., and Coker, A. (2016). Progress in surgical and nonsurgical approaches for hepatocellular carcinoma treatment. *Hepatobiliary Pancreat Dis. Int.* 15, 234–256. doi: 10.1016/S1499-3872(16)60097-8
- Yu, H., Xu, W. L., Gong, F. C., Chi, B. R., Chen, J. Y., and Zhou, L. (2017). MicroRNA-155 regulates the proliferation, cell cycle, apoptosis and migration of colon cancer cells and targets CBL. *Exp. Therap. Med.* 14:5085. doi: 10.3892/etm.2017.5085
- Yu, Y. J., Kanwar, S. S., Patel, B., Ohta, P. S., Nautiyal, J., Sarkar, F. H., et al. (2012). MicroRNA-21 induces stemness by downregulating transforming growth factor beta receptor 2 (TGF $\beta$ R2) in colon cancer cells. *Carcinogenesis* 12:246. doi: 10.1093/carcin/bgr246
- Yuan, J. Y., Ji, H. X., Xiao, F., Lin, Z. P., Zhao, X.J., Wang, Z. C., et al. (2017). MicroRNA-340 inhibits the proliferation and invasion of hepatocellular carcinoma cells by targeting JAK1. *Biochem. Biophys. Res. Commun.* 17:102. doi: 10.1016/j.bbrc.2016.12.102
- Zhang, X., Tang, W., Chen, G., Ren, F. H., Liang, H. W., Dang, Y. W., et al. (2016). An encapsulation of gene signatures for hepatocellular carcinoma, MicroRNA-132 predicted target genes and the corresponding overlaps. *PLoS One* 16:e0159498 doi: 10.1371/journal.pone.0159498
- Zhang, Y. S., Pang, D. L., Wang, J. H., and Zhang, J. L. (2019). *qkerntool: Q-Kernel-Based and Conditionally Negative Definite Kernel-Based Machine Learning Tools*. Available online at: <https://cran.r-project.org/package=qkerntool> (accessed April 13, 2019).
- Zhang, Z. C., Zhang, X. F., Wu, M., Ou-Yang, L., Zhao, X. M., and Li, X. L. (2020). A graph regularized generalized matrix factorization model for predicting links in biomedical bipartite networks. *Bioinformatics (Oxf Engl)* 36:157. doi: 10.1093/bioinformatics/btaa157
- Zhao, Y., Chen, X., and Yin, J. (2018). A Novel computational method for the identification of potential miRNA-disease association based on symmetric non-negative matrix factorization and kronecker regularized least square. *Front. Genet.* 9:324. doi: 10.3389/fgene.2018.00324
- Zhu, X., Wang, X., Zhao, H., Pei, T., Kuang, L., and Wang, L. (2020). BHCMDA: a new biased heat conduction based method for potential MiRNA-disease association prediction. *Front. Genet.* 11:384. doi: 10.3389/fgene.2020.00384

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

Copyright © 2020 Wang, Chen, Zhang, Chen, Zhang and Gao. 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.