BMC Medical Informatics and Decision Making

RESEARCH

Open Access

MiRNA-disease association prediction via hypergraph learning based on high-dimensionality features

Yu-Tian Wang[1†], Qing-Wen Wu[1†], Zhen Gao[1], Jian-Cheng Ni[1*] and Chun-Hou Zheng[2,3*]

From Fifteenth International Conference on Intelligent Computing (ICIC 2019) Nanchang, China. 3-6 August 2019

Abstract

Background: MicroRNAs (miRNAs) have been confirmed to have close relationship with various human complex diseases. The identification of disease-related miRNAs provides great insights into the underlying pathogenesis of diseases. However, it is still a big challenge to identify which miRNAs are related to diseases. As experimental methods are in general expensive and time-consuming, it is important to develop efficient computational models to discover potential miRNA-disease associations.

Methods: This study presents a novel prediction method called HFHLMDA, which is based on high-dimensionality features and hypergraph learning, to reveal the association between diseases and miRNAs. Firstly, the miRNA functional similarity and the disease semantic similarity are integrated to form an informative high-dimensionality feature vector. Then, a hypergraph is constructed by the K-Nearest-Neighbor (KNN) method, in which each miRNA-disease pair and its *k* most relevant neighbors are linked as one hyperedge to represent the complex relationships among miRNA-disease pairs. Finally, the hypergraph learning model is designed to learn the projection matrix which is used to calculate uncertain miRNA-disease association score.

Result: Compared with four state-of-the-art computational models, HFHLMDA achieved best results of 92.09% and 91.87% in leave-one-out cross validation and fivefold cross validation, respectively. Moreover, in case studies on Esophageal neoplasms, Hepatocellular Carcinoma, Breast Neoplasms, 90%, 98%, and 96% of the top 50 predictions have been manually confirmed by previous experimental studies.

Conclusion: MiRNAs have complex connections with many human diseases. In this study, we proposed a novel computational model to predict the underlying miRNA-disease associations. All results show that the proposed method is effective for miRNA–disease association predication.

Keywords: MicroRNA, Disease, MiRNA-disease association, K-nearest-neighbor, Hypergraph learning

Background

MicroRNAs (miRNAs) are endogenous non-coding single-stranded RNA molecules that play important roles in eukaryotic gene expression through posttranscriptional regulation [1–3]. Functional studies indicate that miRNA plays a significant role in manifold biological

*Correspondence: nijch@163.com; zhengch99@126.com
†Yu-Tian Wang and Qing-Wen Wu contributed equally to this work
[1] School of Software, Qufu Normal University, Qufu, China
[2] School of Computer Science and Technology, Anhui University, Hefei, China
Full list of author information is available at the end of the article

processes, such as cell proliferation, stem cell maintenance, immune responses and so on [4–6]. Dysregulation of miRNA expression and function is reported in various diseases including cancer, metabolic disorders as well as neurological disorders [7]. Therefore, identifying disease-related miRNAs is important to treat, diagnose, and prevent human complex diseases [8, 9].

Generally, researchers use biological experimental methods such as quantitative reverse transcription, microarray analysis, or deep sequencing of small RNAs to explore miRNAs that are differentially expressed in a disease state. For example, Pan et al. used microarray analysis and found that miR-130a-3p, miR-424-5p, miR-574-5p, and miR-146a presented significant difference between tuberculous meningitis and healthy controls [10]. However, experimental identification of disease-related miRNAs by existing techniques is expensive and time-consuming. So, based on vast amount of biological data about miRNAs, researchers have developed computational methods for predicting miRNA-disease associations [11–21], which can select most promising miRNAs for further analysis and hence decrease the number of the experiments.

For predicting disease-related miRNAs, many methods are based on a credible assumption that functionally similar miRNAs tend to have associations with phenotypically similar diseases and vice versa. Xiao et al. proposed a method called GRNMF, which based on graph regularized non-negative matrix factorization from the similarity and association perspective of miRNAs and diseases to discover potential associations [22]. Liu et al. proposed the method for predicting miRNA–disease associations by performing random walks on heterogeneous omics data [23]. You et al. presented the prediction model of PBMDA by constructing a heterogeneous graph consisting of three interlinked sub-graphs, and performing a depth-first search algorithm on the heterogeneous network to infer disease-related miRNAs [24]. PBMDA integrated different types of heterogeneous biological datasets, so it can be applied to the new diseases/miRNAs without known associated miRNAs/diseases. Subsequently, Chen et al. proposed a novel method based on Hybrid Approach for MiRNA-Disease Association prediction (HAMDA) [25]. They considered network structure, information propagation, and node attribution, and used the hybrid graph-based recommendation algorithm to uncover disease-related miRNAs. In addition, Chen et al. devised a computational approach by Graphlet Interaction to predict disease-related miRNAs (GIMDA) [26]. In this method, graphlet interaction was utilized to analyze the complex relationships between two nodes in a graph. However, HAMDA and GIMDA are not applicable to predicting a new association between a new miRNA and a new disease. Furthermore, Chen et al. developed a method of Graph Regression for MiRNA-Disease Association prediction (GRMDA) [27]. The graph regression was synchronously performed in three latent spaces, by using Singular Value Decomposition (SVD) and Partial Least-Squares (PLS) to extract important related attributes and filter the noise. But it is difficulties to choosing parameters in SVD and PLS. Lately, Jiang et al. implemented a improved collaborative filtering-based method to infer miRNA-disease associations (ICFMDA) [28]. They improved collaborative filtering algorithm by combining the similarity matrices, and defined significance SIG between pairs of diseases or miRNAs to predict disease-related miRNAs even new diseases without known association.

In addition, several computational models used machine learning to uncover the association between miRNAs and diseases. Xu et al. introduced an approach based on the miRNA target–dysregulated network (MTDN) to prioritize novel disease miRNAs [29]. They applied Support vector machine classifier to miRNAs in the MTDN. However, negative samples required by the classifier are difficult to obtain. To overcome this limitation, Chen et al. introduced a semi-supervised method named RLSMDA [30]. It is developed under the framework of regularized least squares and can predict new miRNAs for diseases which do not have any known related miRNAs. Similarly, Luo et al. developed another semi-supervised method named KRLSM based on Kronecker regularized least squares [31]. KRLSM integrated different omics data, combined the disease and miRNA space, and used the semi-supervised classifier of regularized least squares to predict disease-related miRNAs. However, this approach involves multiple parameters and establishing the optimal parameter values remains a challenging problem. Chen et al. designed a method based on restricted Boltzmann machine for predicting miRNA-disease associations [32]. This approach can also predict association types of miRNA-disease pairs, but can not applicable to a new disease with no known associated miRNAs. Furthermore, Chen et al. developed an effective method called HGIMDA [33]. HGIMDA calculated the disease-miRNA association possibility by investigating all the 3-length paths in the constructed heterogeneous graph. Recently, Chen et al. utilized Extreme Gradient Boosting Machine to uncover disease-related miRNAs and named EGBMMDA [34]. In this method, based on statistical measures, graph theoretical, and matrix factorization, they constructed an informative feature vector for each miRNA-disease pair and used a decision tree model to predict disease-related miRNAs.

Although existing methods have made great contributions to uncover disease-related miRNAs, there are still some limitations that could be improved. For example, many methods are difficult to extract the deep feature representation of the multiple kinds of data. In this study, we propose a novel prediction method via hypergraph learning based on high-dimensionality features and refer to it as HFHLMDA. Hypergraph learning, which can capture the high-order relationships of samples, has been widely used in clustering, classification and information retrieval tasks. In a hypergraph, an edge connects more than two vertices, thus it can well encode the relationship among more than two vertices. We construct high-dimensionality feature vectors for all the miRNA-disease pairs, and utilize K-Nearest-Neighbor (KNN) method to form a hypergraph to predict potential miRNA-disease association. To demonstrate the effectiveness of our method, we apply Leave-one-out cross validation (LOOCV) and fivefold cross validation to measure the prediction performance. We compare our method with four state-of-the-art methods and the results indicate that our method can achieve better performance. In addition, case studies of three common diseases are implemented to further verify the reliability and robustness of HFHLMDA.

Methods

Human MiRNA-disease associations network

The human miRNA-disease associations used in this work come from the HMDDv2.0 [35], which contains 5430 experimentally associations between 495 miRNAs and 383 diseases. Technically, we use an adjacency matrix A with 495 (nm) rows and 383 (nd) columns to clearly describe the relation of each miRNA-disease pairs. The element $A(m(\mathrm{i}), d(j))$ is equal to 1 if miRNA $m(i)$ is verified to be associated with disease $d(j)$, and 0 otherwise. Finally, 5430 entries of matrix A are assigned 1, the rest ones are assigned 0. Our goal is to confirm the uncertain associations between miRNAs and diseases.

MiRNA similarity matrix

Wang et al. developed a method named MISIM for calculating the function similarity scores of miRNA [36]. Here, we directly downloaded the miRNA functional similarity scores from http://www.cuilab.cn/files/images/cuilab/misim.zip. Then, an adjacency matrix SM with 495 rows and 495 columns is built to denote the similarity of miRNAs, in which the larger the $SM(m(i), m(j))$ is, the more similar $m(i)$ and $m(j)$ are.

However, SM has the problem of sparsity. Sparse matrix is difficult to provide more effective information, which will seriously affect the prediction performance of the computational model. So we calculate the Gaussian interaction profile kernel similarity of miRNAs [37]. Specifically, a binary vector $BV(m(i))$, i.e. the ith row of matrix A, is recorded as the interaction profiles of miRNA $m(i)$ for representing the associations between $m(i)$ itself and each disease. All known miRNA-disease associations in matrix A will be used to calculate similarity, two miRNAs would likely have greater similarities if they share more disease associations. Thus, the Gaussian interaction profile kernel similarity $GKM(m(i), m(j))$ of miRNA $m(i)$ and miRNA $m(j)$ is defined as

$$GKM\left(m(i), m(j)\right) = \exp(-\gamma_m||BV(m(i)) - BV\left(m(j)\right)||^2) \tag{1}$$

where γ_m is a parameter used to control the kernel bandwidth, which is set as

$$\gamma_m = \frac{1}{\frac{1}{nm}\sum_{i=1}^{nm}||BV(m(i))||^2} \tag{2}$$

By integrating SM and GKM, a new complete miRNA similarity matrix SM can be obtained as

$$SM\left(m(i), m(j)\right) = \begin{cases} GKM\left(m(i), m(j)\right) & ifSM\left(m(i), m(j)\right) = 0 \\ \frac{SM(m(i),m(j)) + GKM(m(i),m(j))}{2} & otherwise \end{cases} \tag{3}$$

Disease similarity matrix

The association between different diseases can be represented by a directed acyclic graph (DAG), which consists of some nodes and links. Each node represents a disease while a link represents the association of two diseases. For a given disease D, DAG = (D, T_D, E_D), where T_D represents its ancestor nodes and itself while E_D is the set of corresponding edges. The contribution values of disease $d(t)$ to the semantic value of disease $d(i)$ can be calculated as follows:

$$D_{d(i)}(d(t)) = -log\left(\frac{the\ number\ of\ DAGs\ includingd(t)}{the\ number\ of\ diseases}\right) \tag{4}$$

$$DV(d(i)) = \sum_{d(t)\in D(d(i))} D_{d(i)}(d(t)) \tag{5}$$

where $D(d(i))$ is the node set in DAG($d(i)$) including node $d(i)$ itself. Therefore, the semantic similarity between disease $d(i)$ and $d(j)$ can be defined as follows:

$$SD\left(d(i),d\left(j\right)\right)=\frac{\sum_{d(t)\in D(d(i))\cap D(d(j))}\left(D_{d(i)}(d(t))+D_{d(j)}(d(t))\right)}{DV(d(i))+DV\left(d\left(j\right)\right)} \tag{6}$$

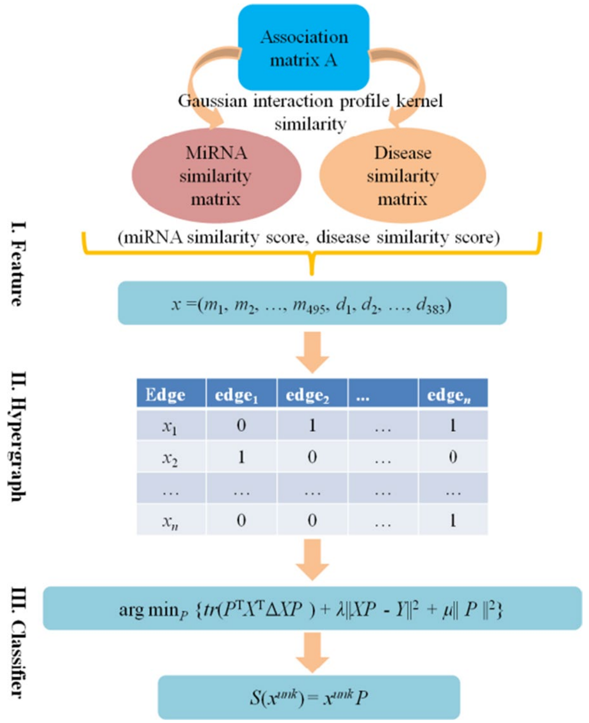
Similarly, we also calculate the Gaussian interaction profile kernel similarity *GKD* for diseases by the follow formulas

$$GKD\left(d(i),d\left(j\right)\right)\ =\exp(-\gamma_d||BV(d(i))\ -BV\left(d\left(j\right)\right)||^2) \tag{7}$$

$$\gamma_d=\frac{1}{\frac{1}{nd}\sum_{i=1}^{nd}||BV(d(i))||^2} \tag{8}$$

where $BV(d(i))$ and $BV(d(j))$ denote the ith column and the j-th column of A. At last, the disease similarity matrix SD is obtained by

$$SD\left(d(i),d\left(j\right)\right)\ =\begin{cases}GKD\left(d(i),d\left(j\right)\right) & if\ SD\left(d(i),d\left(j\right)\right)=0\\ \frac{GKD(d(i),d(j))+SD(d(i),d(j))}{2} & otherwise\end{cases} \tag{9}$$


Association matrix A
Gaussian interaction profile kernel similarity
MiRNA similarity matrix
Disease similarity matrix
(miRNA similarity score, disease similarity score)
I. Feature
$x=(m_1, m_2, \ldots, m_{495}, d_1, d_2, \ldots, d_{383})$
II. Hypergraph
III. Classifier
$\arg\min_P\{tr(P^TX^T\Delta XP\ )+\lambda\|XP-Y\|^2+\mu\|\ P\ \|^2\}$
$S(x^{unk})=x^{unk}P$

Edge	edge$_1$	edge$_2$	...	edge$_n$
x_1	0	1	…	1
x_2	1	0	…	0
…	…	…	…	…
x_n	0	0	…	1

Fig. 1 Flowchart of potential miRNA–disease association prediction based on HFHLMDA

HFHLMDA

The HFHLMDA model can be separated into three steps (see Fig. 1). First, feature factor construction, in which a feature factor x for each miRNA-disease pair consisting of corresponding rows of *SM* and *SD*. Second, hypergraph construction, where a hypergraph G is constructed to formulate the relationship between these feature vectors. Third, hypergraph learning, to learn the projection matrix P, which map the original feature x to the relevance score $S=x{\cdot}P$, and thus it can be used to predict the association for the unknown miRNA-disease pair x^{unk}.

Feature factor construction

According to the biological observation that miRNAs with more functional similarity tend to be more associated with similar diseases and vice versa, so the topologic information of miRNA/disease similarity network can be used to construct feature factor directly.

For each miRNA, there are 495 similarity scores. We use similarity scores as features to represent each miRNA by a 495-dimensional feature vector. For example, we represent miRNA $m(i)$ by a feature vector, $SM(m(i))=(m_1, m_2, \ldots, m_{495})$, where $SM(m(i))$ is the ith row vector of SM and represents the similarities between $m(i)$ and all the miRNAs.

For each disease, we can obtain a 383-dimensonal feature vector in a similar way to miRNA, $SD(d(j))=(d_1, d_2, \ldots, d_{383})$, where $SD(d(j))$ is the jth row of matrix SD. Therefore, each miRNA-disease pair can be described by an 878-dimensional vector $x=(SM(m(i)), SD(d(j)))$. Furthermore, we consider $(SM(m(i)), SD(d(j)))$ as a positive sample if miRNA $m(i)$ is associated with disease $d(j)$, otherwise as a negative sample. To construct the balanced dataset, the training set have 5,430 positive samples, and an equal number of samples were randomly selected as negative training examples from the pool of unknown associations. It is possible to use unconfirmed miRNA-disease pairs with association as negative samples, from the perspective of probability, because the miRNA-disease pairs we selected as negative samples account for only $5430 \div (495 \times 383) \approx 2.86\%$ of all miRNA-disease pairs, which is negligible [38].

Hypergraph construction

Firstly, we briefly introduce the hypergraph learning theory. As a generalization of graph, hypergraph represents the structure of data via measuring the similarity between groups of points. Different from a simple graph, an edge in a hypergraph can connect three or more vertices, it can model high-order relations between their vertices by hyperedges, whose influence can be assessed by properly estimating their weights. Obviously, modeling the high-order relationship among objects can improve the predicting performance significantly. Moreover, the quality of the hypergraph structure plays an important role for data modeling. A well constructed hypergraph structure can represent the data correlation accurately, and leading to better performance.

A hypergraph is defined as $G = (V, E, w)$, where V is a set of vertices, E is a set of hyperedges and each hyperedge e is given a positive weight $w(e)$. The hypergraph G can be denoted by a $|V| \times |E|$ incidence matrix H, in which each entry is defined by

$$h(v, e) = \begin{cases} 1 \ \mathit{if} \ v \in e \\ 0 \ \mathit{if} \ v \notin e \end{cases} \tag{10}$$

The degree of vertex $v \in V$ and hyperedge $e \in E$ can be respectively represented as:

$$d(v) = \sum_{e \in E} w(e) h(v, e) \tag{11}$$

$$\delta(e) = \sum_{v \in V} h(v, e) \tag{12}$$

Accordingly, denote Dv and De as two diagonal matrices of the vertex degrees and the hyperedge degrees, respectively.

Zhou et al. proposed a regularization framework on hypergraph [39], which is defined as

$$\arg\min_f \{\lambda R_{\mathrm{emp}}(f) + \Omega(f)\} \tag{13}$$

where f is the to-be-learned function, $\Omega(f)$ is a regularizer on the hypergraph, $R_{\mathrm{emp}}(f)$ is an empirical loss, and $\lambda > 0$ is the tradeoff parameter. Usually, the empirical loss $R_{\mathrm{emp}}(f)$ is defined as

$$R_{\mathrm{emp}}(f) = ||f - Y||^2 \tag{14}$$

where Y is the label matrix of samples. The regularizer on the hypergraph is defined by

$$\Omega(f) = \frac{1}{2} \sum_{e \in E} \sum_{u,v \in V} \frac{w(e)}{\delta(e)} \left(\frac{f(u)}{\sqrt{d(u)}} - \frac{f(v)}{\sqrt{d(v)}} \right) \tag{15}$$

Let $\Theta = D_v{}^{-(1/2)} H W D_{\mathrm{e}}{}^{-1} H^{\mathrm{T}} D_v{}^{-(1/2)}$**,** the normalized cost function can be written as

$$\Omega(f) = f^{\mathrm{T}} \Delta f \tag{16}$$

where $\Delta = I - \Theta$, which is a positive semi-definite matrix.

In this study, given a set of training samples $\{x_i \,|\, i = 1, \ldots, n\} \in \mathbb{R}^{878}$, the data matrix $X = [x_1, \ldots, x_i, \ldots, x_n]^{\mathrm{T}} \in \mathbb{R}^{n \times 878}$ contains n samples in its rows, the corresponding labels matrix $Y = [y_1, \ldots, y_2, \ldots, y_l] \in \mathbb{R}^{n \times l}$, y_i is the label vector of the i-th class. A miRNA-disease pairs hypergraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ is constructed, and its hyperedge is generated based on the KNN algorithm. Concretely, for each vertex v, we search its corresponding k nearest neighbors, and use these nearest neighbors to form a hyperedge $e(v)$. We initialize k as 15 here empirically. An illustration on the hyperedge generation process is shown in Fig. 2. Moreover, the diagonal matrix $\mathcal{W}$ denote the weights of

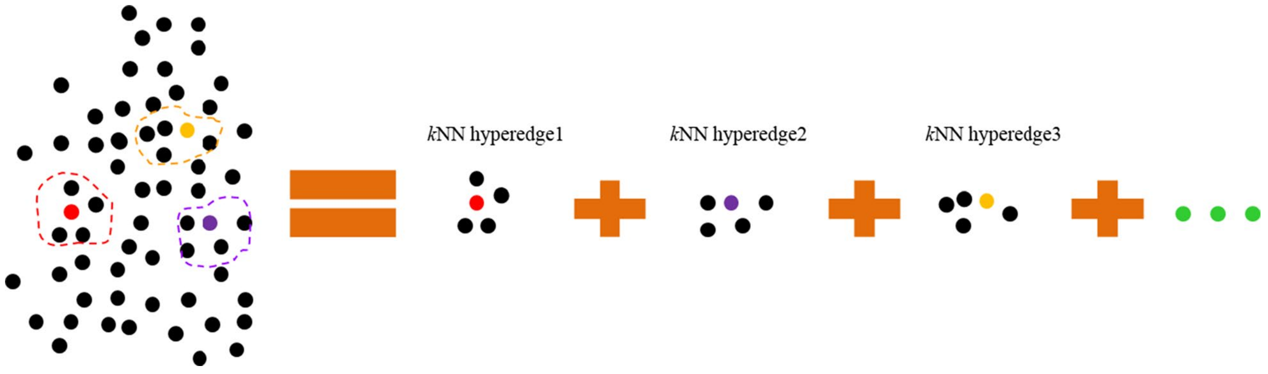


Fig. 2 Intuitive illustration of KNN hyperedge generation

the hyperedges. All the hyperedges are initialized with an equal weight, *e.g.*, $w(e)=1/n_e$, where n_e is the number of hyperedges.

Hypergraph learning

The hypergraph learning targets on learning a regularized projection to discriminate different categories. According to Zhang et al. introduction [40], the cost function F for learning the projection matrix P can be formulated as:

$$F = \{\Omega(P) + \lambda R_{emp}(P) + \mu\Phi(P)\} \tag{17}$$

where λ and μ are positive parameters, and we empirically set them as 10^1,10^0 respectively, which can achieve the best performance. Specifically, hypergraph Laplacian regularizer $\Omega(P)$ is calculated as

$$\begin{aligned}\Omega(P) &= \frac{1}{2}\sum_{k=1}^{l}\sum_{e\in E}\sum_{u,v\in V}\frac{W(e)H(u,e)H(v,e)}{\delta(e)}\left(\frac{(XP)(u,k)}{\sqrt{d(u)}}-\frac{(XP)(v,k)}{\sqrt{d(v)}}\right)^2\\ &= tr(P^TX^T\Delta XP)\end{aligned} \tag{18}$$

where function $tr(\cdot)$ returns the trace of matrix. The empirical loss term $R_{emp}(P)$ is defined as

$$R_{emp}(P) = ||XP - Y||^2 \tag{19}$$

$\Phi(P)$ is a l_2 norm regularizer to avoid over-fitting for P, which is defined as:

$$\Phi(P) = ||P||^2 \tag{20}$$

Consequently, Eq. (17) can be reformed as:

$$\arg\min_P\left\{tr\left(P^\mathrm{T}X^\mathrm{T}\Delta XP\right) + \lambda||XP - Y\Big|\Big|^2 + \mu\Big|\Big|P||^2\right\} \tag{21}$$

Such problem is a typical Least Square problem which can be efficiently solved, its solution is as follows:

$$P = \lambda(X^T\Delta X + \lambda X^\mathrm{T}X + \mu I)^{-1}X^\mathrm{T}Y \tag{22}$$

where I is an identity matrix. Based on the learned P, the relevance score of the unknown miRNA-disease pair x^{unk} can be obtained by

$$\mathrm{S}\left(x^{unk}\right) = x^{unk\cdot}P \tag{23}$$

Results

Effect of parameters on the performance of HFHLMDA

In this work, we used KNN algorithm to generate hyperedge, one parameters k was included, which represent the number of nearest neighbors of miRNA or disease. In the hypergraph learning section of the Methods, we defined two parameters, namely, λ and μ to balance the items in Eq. (17), the values of λ and μ ranged from 10^{-2}, 10^{-1}, 10^0, 10^1 to 10^2. We conducted a series of experiments on the above parameters to acquire the effects of these parameters. The experimental results are shown in Figs. 3 and 4. In Fig. 3, we can see that regardless of how k change, the AUC of fivefold cross validation keep around 0.9187. Thus, for efficiency, we set $k=15$. Furthermore, Fig. 4 describes the prediction performances of HFHLMDA with different values of λ and μ. We can see that HFHLMDA obtains the best prediction performance when λ is set to be 10^1 and μ is set to be 10^0.

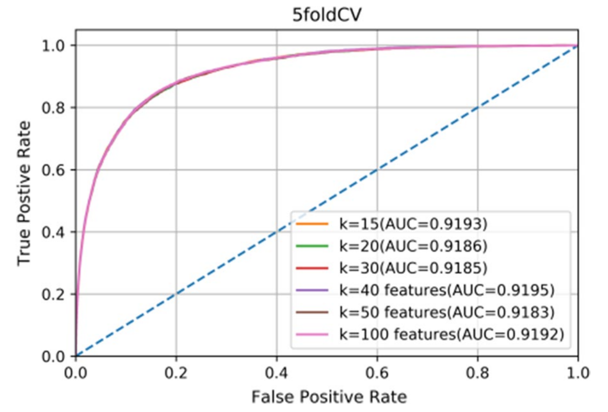


Fig. 3 ROC curve and AUC with different values for parameter k

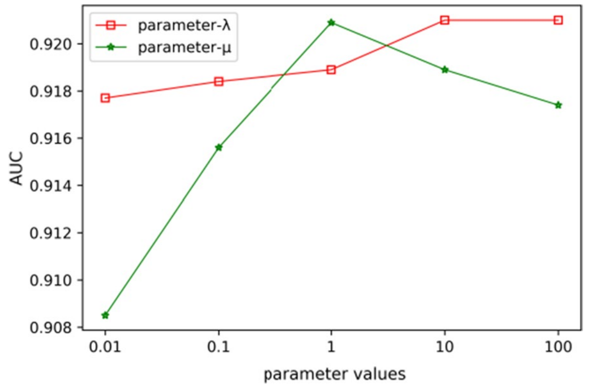


Fig. 4 AUC with different values for parameters λ and μ

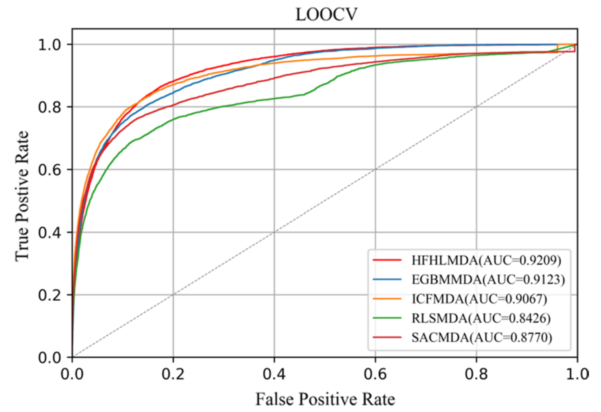


Fig. 5 Performance comparisons between HFHLMDA and four classical models in terms of ROC curve and AUC based on LOOCV

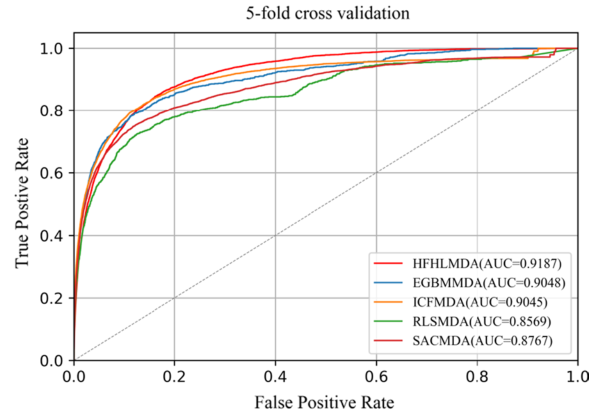


Fig. 6 Performance comparisons between HFHLMDA and four classical models in terms of ROC curve and AUC based on fivefold cross validation

Performance evaluation

Based on the known miRNA–disease associations in HMDDv2.0 database, two validation schemas were used to evaluate the performance of HFHLMDA: LOOCV and fivefold cross validation. We selected four classical computational methods: EGBMMDA [34], ICFMDA [28], RLSMDA [30], and SACMDA [41] to compete with HFHLMDA in cross validation. Specifically, LOOCV selected a known miRNA-disease association in turn as a test sample, and the rest of the associations were considered as training samples. All unknown associations were used as candidate samples. Considering that the Gaussian interaction profile kernel similarity depend on known miRNA-disease associations, the corresponding value of a test sample in matrix A should be set to 0. The predicted score for the test sample was ranked relative to the scores for candidate samples and, each ranking will take turns as a threshold in each fold, if test ranking was above a given threshold, we obtained a successful prediction made by the model. By changing the threshold, we could calculate the corresponding true positive rate (TPR) and false positive rate (FPR). Furthermore, receiver-operating characteristics (ROC) curve could be drawn according to TPR against FPR. The areas under the ROC curve (AUC) was used to evaluate the whole prediction performance. Figure 5 shows the global LOOCV ROC curves for HFHLMDA and other methods. HFHLMDA, EGBMMDA, ICFMDA, RLSMDA and SACMDA obtained AUCs of 0.9209, 0.9123, 0.9067, 0.8426 and 0.8770, respectively. HFHLMDA achieved the better prediction performance.

As for fivefold cross validation, in order to make the validation more accurate, we repeated fivefold cross validation procedure 100 times. The average AUC values of the five methods (HFHLMDA, EGBMMDA, ICFMDA, RLSMDA, SACMDA) were 0.9187(±0.0009), 0.9048(±0.0012), 0.9045(±0.0008), 0.8569(±0.0020) and 0.8767(±0.0011), respectively (see Fig. 6). In summary, under the same dataset, our model outperformed other competitive methods.

Case studies

Case studies were conducted to further verify the capability of HFHLMDA to predict miRNA-disease associations. We implemented three different kinds of case studies in this study. In the first case study, we conducted HFHLMDA to predict potential disease-miRNA associations taking advantages of known diseases-miRNAs associations included in HMDD v2.0 database. Subsequently, top 50 miRNAs for the investigated disease ranked according to their predicted scores were verified using another two well-known miRNA-disease association databases of dbDEMC [42] and miR2Disease [43]. In the second case study, we simulated the situation where HFHLMDA was conducted for disease without known miRNA associations. More concretely, we removed the known miRNA associations of the disease of interest, after which HFHLMDA was implemented according newly obtained association records. The prediction results were also verified by other databases. The final case study investigated the robustness of HFHLMDA prediction performance. We evaluated the model with a smaller and earlier version HMDDv1.0 database [44].

Esophageal cancer (EC) is one of the most common cancers worldwide, and its 5-year survival rate is about 20% [45]. Study indicate that miR-130b plays an oncogenic role in esophageal squamous cell carcinoma cells by repressing phosphatase and tensin homolog expression

Table 1 The top 50 predicted miRNAs associated with esophageal cancer

miRNA	Evidence	miRNA	Evidence
hsa-mir-221	dbDEMC	hsa-mir-9	dbDEMC
hsa-mir-125b	dbDEMC	hsa-mir-24	dbDEMC
hsa-mir-29a	dbDEMC	hsa-mir-132	dbDEMC
hsa-mir-206	dbDEMC	hsa-mir-224	dbDEMC
hsa-mir-17	dbDEMC	hsa-mir-23a	dbDEMC
hsa-mir-16	dbDEMC	hsa-let-7d	dbDEMC
hsa-mir-29b	dbDEMC	hsa-mir-195	dbDEMC
hsa-mir-222	dbDEMC	hsa-mir-335	dbDEMC
hsa-mir-1	dbDEMC	hsa-mir-124	dbDEMC
hsa-mir-146b	dbDEMC	hsa-mir-93	dbDEMC
hsa-mir-182	dbDEMC	hsa-mir-106a	dbDEMC
hsa-mir-122	Unconfirmed	hsa-mir-140	dbDEMC
hsa-mir-181a	dbDEMC	hsa-mir-30a	dbDEMC
hsa-mir-18a	dbDEMC	hsa-mir-184	Unconfirmed
hsa-mir-106b	dbDEMC	hsa-mir-429	dbDEMC
hsa-let-7e	dbDEMC	hsa-let-7i	dbDEMC
hsa-mir-200b	dbDEMC	hsa-let-7f	Unconfirmed
hsa-mir-20b	dbDEMC	hsa-mir-134	dbDEMC
hsa-mir-19b	dbDEMC	hsa-mir-27b	dbDEMC
hsa-mir-133b	dbDEMC	hsa-mir-23b	dbDEMC
hsa-let-7 g	dbDEMC	hsa-mir-152	dbDEMC
hsa-mir-181b	dbDEMC	hsa-mir-96	dbDEMC
hsa-mir-15b	dbDEMC	hsa-mir-193b	dbDEMC
hsa-mir-103a	Unconfirmed	hsa-mir-138	Unconfirmed
hsa-mir-142	dbDEMC	hsa-mir-125a	dbDEMC

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

and Akt phosphorylation [46]. Therefore, specific and sensitive biomarkers for diagnosis and targeted therapy of EC are urgently needed. As the first type of case study, 10 out of top 10, 28 out of top 30, 45 out of top 50 predicted esophageal neoplasms related miRNAs were confirmed by dbDEMC (See Table 1).

Hepatocellular carcinoma (HC) is a complex polygenetic disease ascribed to the interactions between genetic predisposition and environmental factors [47]. The discovery of vital target for genetic therapy are of great clinical significance to the improvement of the comprehensive effect of HC. For example, miR-122, let-7 family, and miR-101 are down-regulated in HC, suggesting that it is a potential tumor suppressor of HC. miR-221 and miR-222 are up-regulated in HCC and may act as oncogenic miRNAs in hepatocarcinogenesis [48]. We took hepatocellular carcinoma as the second kind of case study. Finally, 49 out of top 50 miRNAs were experimentally confirmed by HMDD v2.0, dbDEMC and miR2Disease (See Table 2).

Breast Neoplasms is the most common malignancy in women, accounting more than 40,000 deaths each year [49]. Data have shown that the number of affected people is climbing, and a forecast deemed that there will be nearly 3.2 million new patients per year by 2050 [50]. In breast cancer, approximately one-fifth of metastatic patients survive 5 years [51]. Researchers have found that many miRNAs are associated with breast neoplasms by clinical experiments, such as mir-155 and mir-21, both of which can lead to Breast Neoplasms tumorigenesis or metastasis [52]. We took breast neoplasms as the last kind of case study, in which we got the prediction with HFHLMDA using HMDDv1.0 database. Then, we verified the predicted potential breast neoplasms related miRNAs in other databases. At last, 48 out of top 50 miRNAs were experimentally confirmed by HMDD v2.0, dbDEMC and miR2Disease (See Table 3).

The aforementioned case studies indicate that HFHLMDA has good prediction performance. HFHLMDA can efficiently predict disease-related miRNAs based on known miRNA-disease associations, disease semantic similarity and miRNA functional similarity, and a disease without known associations also can be predicted.

Discussion

In this work, we developed a new computational model based on hypergraph learning to predict potential miRNA-disease associations. Several important factors contribute to the excellent performance of our model. First, high-dimensionality features. Based on a credible assumption that functionally similar miRNAs tend to have associations with phenotypically similar diseases. We use the miRNAs or diseases similarity scores directly as a feature factor, with a dimension of up to 878, which contains all similar information about miRNAs or diseases. Second, hypergraph is suitable to represent local group information and the high-order relationship of data, and can completely represent the complex relationships among miRNA-disease pairs. Different from the simple-graph learning methods consider only the pairwise relationship between two samples, and they ignore the relationship in a higher-order, hypergraph learning aims to get the relationship between several samples in a higher order. Hypergraph learning is a kind of graph clustering algorithm, the process of graph clustering is actually the optimization of graph partition. The purpose of optimization is to reduce the similarity between sub-graphs and increase the similarity within sub-graphs. Hypergraph-based models have proven to be beneficial for a variety of classification/clustering tasks, and we think it can also be applied to different fields of bioinformatics, such as drug-disease associations [53], miRNA–drug interactions [54].

Table 2 The top 50 predicted miRNAs associated with hepatocellular carcinoma

miRNA	Evidence	miRNA	Evidence
hsa-mir-21	HMDD; miR2disease	hsa-let-7b	HMDD; miR2disease
hsa-mir-155	HMDD; miR2disease; dbDEMC	hsa-mir-122	HMDD; miR2disease; dbDEMC
hsa-mir-146a	HMDD; miR2disease; dbDEMC	hsa-mir-18a	HMDD; miR2disease; dbDEMC
hsa-mir-221	HMDD; miR2disease; dbDEMC	hsa-mir-106b	HMDD; miR2disease; dbDEMC
hsa-mir-125b	HMDD; miR2disease	hsa-mir-200a	HMDD; miR2disease; dbDEMC
hsa-mir-145	HMDD; miR2disease; dbDEMC	hsa-mir-223	HMDD; miR2disease
hsa-mir-29a	HMDD; dbDEMC	hsa-mir-150	HMDD; miR2disease; dbDEMC
hsa-mir-206	Unconfirmed	hsa-mir-19b	HMDD; miR2disease
hsa-mir-17	HMDD; miR2disease	hsa-mir-29c	HMDD; dbDEMC
hsa-mir-16	HMDD; miR2disease; dbDEMC	hsa-mir-143	miR2disease; dbDEMC
hsa-mir-29b	HMDD; dbDEMC	hsa-let-7g	HMDD; miR2disease
hsa-mir-199a	HMDD; miR2disease; dbDEMC	hsa-mir-200b	HMDD; miR2disease
hsa-mir-214	HMDD; miR2disease; dbDEMC	hsa-mir-210	HMDD; dbDEMC
hsa-mir-20a	HMDD; miR2disease; dbDEMC	hsa-mir-126	HMDD; miR2disease; dbDEMC
hsa-mir-92a	HMDD; miR2disease	hsa-mir-20b	HMDD; dbDEMC
hsa-mir-222	HMDD; miR2disease; dbDEMC	hsa-let-7c	HMDD; miR2disease; dbDEMC
hsa-mir-1	HMDD; miR2disease	hsa-mir-34c	HMDD
hsa-mir-34a	HMDD; miR2disease; dbDEMC	hsa-mir-141	HMDD; miR2disease
hsa-mir-15a	HMDD; miR2disease; dbDEMC	hsa-mir-133b	HMDD
hsa-mir-182	HMDD; miR2disease	hsa-mir-224	HMDD; miR2disease; dbDEMC
hsa-mir-146b	HMDD	hsa-mir-15b	HMDD; dbDEMC
hsa-mir-181a	HMDD; miR2disease; dbDEMC	hsa-mir-133a	miR2disease
hsa-mir-499a	HMDD	hsa-mir-181b	HMDD; miR2disease; dbDEMC
hsa-let-7e	HMDD; miR2disease; dbDEMC	hsa-mir-200c	HMDD
hsa-let-7a	HMDD; miR2disease; dbDEMC	hsa-mir-19a	HMDD; miR2disease; dbDEMC

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

Despite the practicability and efficiency of HFHLMDA, there still has some limitations. Since our method is based on machine learning techniques, negative samples are required during the training process. However, experimentally confirmed negative samples are difficult to obtain. To resolve this issue, we have randomly selected a subset of unknown miRNA–disease associations as negative instances. In addition, in our method, after the hypergraph has been constructed, it never changes during the learning process, leading to a static hypergraph structure learning mechanism. However, it is uneasy to guarantee that the generated hypergraph structure is optimal and suitable for all applications. In future work, it is necessary to investigate the hypergraph structure optimization, leading to a dynamic hypergraph structure learning scheme.

Conclusion

Increasing evidence indicates that aberrant expression of miRNAs is closely related to the occurrence and development of human complex diseases. Understanding the underlying mechanisms of miRNAs in diseases is becoming an urgent problem worldwide. Compared with traditional methods, the computational model developed for processing heterogeneous biological big data is more efficient and convenient. To predict potentially disease-related miRNAs, we proposed a hypergraph learning

Table 3 The top 50 predicted miRNAs associated with breast neoplasms

miRNA	Evidence	miRNA	Evidence
hsa-mir-16	HMDD; dbDEMC	hsa-mir-148a	HMDD; miR2Disease; dbDEMC
hsa-mir-150	dbDEMC	hsa-mir-101	HMDD; miR2Disease; dbDEMC
hsa-mir-96	HMDD; miR2Disease; dbDEMC	hsa-mir-29c	HMDD; miR2Disease; dbDEMC
hsa-mir-223	HMDD; dbDEMC	hsa-mir-342	HMDD; miR2Disease; dbDEMC
hsa-mir-92a	HMDD	hsa-mir-372	dbDEMC
hsa-let-7 g	HMDD; dbDEMC	hsa-mir-30e	Unconfirmed
hsa-let-7e	HMDD; dbDEMC	hsa-mir-92b	dbDEMC
hsa-let-7i	HMDD; miR2Disease; dbDEMC	hsa-mir-193b	HMDD; miR2Disease; dbDEMC
hsa-mir-373	HMDD; miR2Disease; dbDEMC	hsa-mir-27a	HMDD; miR2Disease; dbDEMC
hsa-let-7b	HMDD; dbDEMC	hsa-mir-130a	dbDEMC
hsa-mir-199b	HMDD; dbDEMC	hsa-mir-195	HMDD; miR2Disease; dbDEMC
hsa-mir-183	HMDD; dbDEMC	hsa-mir-184	dbDEMC
hsa-mir-106a	dbDEMC	hsa-mir-126	HMDD; miR2Disease; dbDEMC
hsa-mir-23b	HMDD; dbDEMC	hsa-mir-152	HMDD; miR2Disease; dbDEMC
hsa-let-7c	HMDD; dbDEMC	hsa-mir-196b	dbDEMC
hsa-mir-32	dbDEMC	hsa-mir-660	dbDEMC
hsa-mir-15b	dbDEMC	hsa-mir-454	Unconfirmed
hsa-mir-191	HMDD; miR2Disease; dbDEMC	hsa-mir-224	HMDD; dbDEMC
hsa-mir-212	dbDEMC	hsa-mir-26a	HMDD; miR2Disease; dbDEMC
hsa-mir-24	HMDD; dbDEMC	hsa-mir-26b	HMDD; dbDEMC
hsa-mir-99b	dbDEMC	hsa-mir-203	HMDD; miR2Disease; dbDEMC
hsa-mir-181a	HMDD; miR2Disease; dbDEMC	hsa-mir-198	dbDEMC
hsa-mir-137	HMDD; dbDEMC	hsa-mir-142	HMDD; dbDEMC
hsa-mir-31	HMDD; miR2Disease; dbDEMC	hsa-mir-208b	HMDD; miR2Disease; dbDEMC
hsa-mir-32	dbDEMC	hsa-mir-95	HMDD; dbDEMC

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

method called HFHLMDA. Both cross-validation and case studies had proved the effectiveness of HFHLMDA in predicting potential miRNA-disease associations.

Abbreviations

KNN: K-Nearest-Neighbor; HMDD: Human microRNA disease database; dbDEMC: Database of differentially expressed miRNAs in human cancers; ROC: Receiver operating characteristics; AUC: The area under the ROC curve; LOOCV: Leave-one-out cross validation.

Acknowledgments

We would like to thank the editor and referees for the thoughtful and insightful comments.

About this supplement

This article has been published as part of BMC Medical Informatics and Decision Making Volume 21 Supplement 1, 2021: Proceedings of the 2019 International Conference on Intelligent Computing (ICIC 2019): medical informatics and decision making. The full contents of the supplement are available online at https://bmcmedinformdecismak.biomedcentral.com/articles/supplements/volume-21-supplement-1.

Authors' contributions

JCN and CHZ supervised the entire project. YTW and QWW conceptualized and designed the study. YTW and ZG undertook data collection. YTW, QWW and ZG performed the data analysis. QWW drafted the initial version. JCN and CHZ revised the manuscript iteratively for important intellectual content. All authors edited the paper and gave final approval for the version to be published. All authors read and approved the final manuscript.

Funding

Publication costs are funded by the National Natural Science Foundation of China (Nos. 61873001, U19A2064, 61872220, 61672037, 61861146002, 11701318 and 61732012), the Key Project of Anhui Provincial Education Department (No. KJ2017ZD01), and the Xinjiang Autonomous Region University Research Program (XJEDU2019Y002). The funding bodies did not play any role in the design of the study or collection, analysis and interpretation of data or in writing the manuscript.

Availability of data and materials

The datasets used during this study is provided by Li et al. [35]. Please download the data from http://www.cuilab.cn/hmdd/ or contact the authors for data requests.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1] School of Software, Qufu Normal University, Qufu, China. [2] School of Computer Science and Technology, Anhui University, Hefei, China. [3] College of Mathematics and System Science, Xinjiang University, Urumqi, China.

Received: 1 November 2020 Accepted: 9 November 2020
Published online: 20 April 2021

References

1. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell. 2009;136:215–33.
2. Kye MJ, Gonçalves ICG. The role of miRNA in motor neuron disease. Front Cell Neurosci. 2014;8:15.
3. Adams BD, Kasinski AL, Slack FJ. Aberrant regulation and function of microRNAs in cancer. Curr Biol. 2014;24(16):R762–76.
4. Cheng AM, Byrom MW, Shelton J, Ford LP. Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. Nucleic Acids Res. 2005;33(4):1290–7.
5. Karp X, Ambros V. Encountering microRNAs in cell fate signaling. Science. 2005;310(5752):1288–9.
6. Shivdasani RA. MicroRNAs: regulators of gene expression and cell differentiation. Blood. 2006;108(12):3646–53.
7. Sayed D, Abdellatif M. MicroRNAs in development and disease. Physiol Rev. 2011;91(3):827–87.
8. Tricoli JV, Jacobson JW. MicroRNA: potential for cancer detection, diagnosis, and prognosis. Cancer Res. 2007;67(10):4553–5.
9. Cho WCS. MicroRNAs: potential biomarkers for cancer diagnosis, prognosis and targets for therapy. Int J Biochem Cell Biol. 2010;42(8):1273–81.
10. Pan LP, Liu F, Zhang JL, et al. Genome-wide miRNA analysis identifies potential biomarkers in distinguishing tuberculous and viral meningitis. Front Cell Infect Microbiol. 2019;9:323.
11. Jiang QH, Hao YY, Wang GH, Juan LR, Zhang TJ, et al. Prioritization of disease microRNAs through a human phenome-microRNAome network. BMC Syst Biol. 2010;4(Suppl 1):S2.
12. Chen X, Liu MX, Yan GY. RWRMDA: predicting novel human microRNA-disease associations. Mol BioSyst. 2012;8(10):2792–8.
13. Shi HB, Xu J, Zhang GD, Xu LD, Li CQ, Wang L, et al. Walking the interactome to identify human miRNA-disease associations through the functional link between miRNA targets and disease genes. BMC Syst Biol. 2013;7(1):101.
14. Zhao XM, Liu KQ, Zhu G, et al. Identifying cancer-related microRNAs based on gene expression data. Bioinformatics. 2015;31(8):1226–34.
15. Qin GM, Li RY, Zhao XM. Identifying disease associated miRNAs based on protein domains. IEEE/ACM Trans Comput Biol Bioinform. 2016;13(6):1027–35.
16. Chen X, Huang L. LRSSLMDA: Laplacian regularized sparse subspace learning for MiRNA-disease association prediction. PLoS Comput Biol. 2017;13(12):e1005912.
17. Chen X, Xie D, Wang L, Zhao Q, You ZH, Liu H. BNPMDA: bipartite network projection for MiRNA-disease association prediction. Bioinformatics. 2018;34(18):3178–86.
18. Chen X, Wang L, Qu J, Guan NN, Li JQ. Predicting miRNA-disease association based on inductive matrix completion. Bioinformatics. 2018;34(24):4256–65.
19. Chen X, Yin J, Qu J, Huang L. MDHGI: matrix decomposition and heterogeneous graph inference for miRNA-disease association prediction. PLoS Comput Biol. 2018;14(8):e1006418.
20. Chen X, Zhu CC, Yin J. Ensemble of decision tree reveals potential miRNA-disease associations. PLoS Comput Biol. 2019;15(7):e1007209.
21. Chen X, Xie D, Zhao Q, You ZH. MicroRNAs and complex diseases: from experimental results to computational models. Brief Bioinform. 2019;20(2):515–39.
22. Xiao Q, Luo JW, Liang C, Cai J, Ding PJ. A graph regularized non-negative matrix factorization method for identifying microRNA-disease associations. Bioinformatics. 2018;34(2):239–48.
23. Liu Y, Zeng X, He Z, Zou Q. Inferring microRNA-disease associations by random walk on a heterogeneous network with multiple data sources. IEEE/ACM Trans Comput Boil Bioinform. 2017;14(4):905–15.
24. You ZH, Huang ZA, Zhu Z, et al. PBMDA: a novel and effective path-based computational model for miRNA-disease association prediction. PLoS Comput Biol. 2017;13(3):e1005455.
25. Chen X, Niu YW, Wang GH, et al. HAMDA: hybrid approach for MiRNA-disease association prediction. J Biomed Inform. 2017;76:50–8.
26. Chen X, Guan NN, Li JQ, et al. GIMDA: graphlet interaction-based MiRNA-disease association prediction. J Cell Mol Med. 2018;22(3):1548–61.
27. Chen X, Yang JR, Guan NN, Li JQ. GRMDA: graph regression for MiRNA-disease association prediction. Front Physiol. 2018;9:92.
28. Jiang YD, Liu BT, Yu LH, et al. Predict MiRNA-disease association with collaborative filtering. Neuroinformatics. 2018;16:363–72.
29. Xu J, Li CX, Lv JY, et al. Prioritizing candidate disease miRNAs by topological features in the miRNA target-dysregulated network: case study of prostate cancer. Mol Cancer Ther. 2011;10(10):1857–66.
30. Chen X, Yan GY. Semi-supervised learning for potential human microRNA-disease associations inference. Sci Rep. 2014;4:5501.
31. Luo JW, Xiao Q, Liang C, Ding PJ. Predicting microRNA-disease associations using Kronecker regularized least squares based on heterogeneous omics data. IEEE Access. 2017;5:2503–13.
32. Chen X, Yan CC, Zhang X, Li Z, Deng L, et al. RBMMMDA: predicting multiple types of disease-microRNA associations. Sci Rep. 2015;5:13877.
33. Chen X, Yan CC, Zhang X, You ZH, et al. HGIMDA: heterogeneous graph inference for miRNA-disease association prediction. Oncotarget. 2016;7(40):65257–69.
34. Chen X, Huang L, Xie D, Zhao Q. EGBMMDA: extreme gradient boosting machine for MiRNA-disease association prediction. Cell Death Dis. 2018;9(1):3.
35. Li Y, Qiu CX, Tu J, Geng B, Yang JC, et al. HMDD v2.0: a database for experimentally supported human microRNA and disease associations. Nucleic Acids Res. 2014;42:D1070–4.
36. Wang D, Wang J, Lu M, Song F, Cui QH. Inferring the human microRNA functional similarity and functional network based on microRNA-associated diseases. Bioinformatics. 2010;26(13):1644–50.
37. Van Laarhoven T, Nabuurs SB, Marchiori E. Gaussian interaction profile kernels for predicting drug-target interaction. Bioinformatics. 2011;27(21):3036–43.
38. Wang L, You ZH, Huang YA, Huang DS, Chan KCC. An efficient approach based on multi-sources information to predict CircRNA-disease associations using deep convoltional neural network. Bioinformatics. 2020;36(13):4038–46.
39. Zhou DY, Huang JY, Schlkopf B. Learning with hypergraphs: clustering, classification, and embedding. Adv Neural Inf Process Syst. 2006;19:1601–8.
40. Zhang ZZ, Liu HJ, Zhao XB, Ji RR, Gao Y. Inductive multi-hypergraph learning and its application on view-based 3D object classification. IEEE Trans Image Process. 2018;27(12):5957–68.
41. Shao BY, Liu BT, Yan CG. SACMDA: MiRNA-disease association prediction with short acyclic connections in heterogeneous graph. Neuroinformatics. 2018;16:373–82.
42. Yang Z, Ren F, Liu CN, He SM, Sun G, Gao Q, et al. dbDEMC: a database of differentially expressed miRNAs in human cancers. BMC Genomics. 2010;11(Suppl 4):S5.
43. Jiang QH, Wang YD, Hao YY, Juan LR, Teng MX, et al. miR2Disease: a manually curated database for microRNA deregulation in human disease. Nucleic Acids Res. 2009;37(1):D98–104.
44. Lu M, Zhang QP, Deng M, Miao J, Guo YH, et al. An analysis of human microRNA and disease associations. PLoS ONE. 2008;3(10):e3420.
45. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin. 2015;65:87–108.
46. Yu T, Cao R, Li S, et al. MiR-130b plays an oncogenic role by repressing PTEN expression in esophageal squamous cell carcinoma cells. BMC Cancer. 2015;15:29.
47. Nie X, Liu Y, Chen WD, Wang YD. Interplay of miRNAs and canonical Wnt signaling pathway in hepatocellular carcinoma. Front Pharmacol. 2018;9:657.
48. Saito Y, Suzuki H, Matsuura M, Sato A, Kasai Y, et al. MicroRNAs in hepatobiliary and pancreatic cancers. Front Gene. 2011;2:66.
49. Desantis CE, Fedewa SA, et al. Breast cancer statistics, 2015: convergence of incidence ratesbetween black and white women. CA Cancer J Clin. 2016;66(1):31–42.

50. Gomella LG. Prostate cancer statistics: anything you want them to be. Can J Urol. 2017;24(1):8603–4.
51. Lee JH, Zhao XM, Yoon I, et al. Integrative analysis of mutational and transcriptional profiles reveals driver mutations of metastatic breast cancers. Cell Discov. 2016;2:16025.
52. Feber A, Xi L, Luketich JD, et al. MicroRNA expression profiles of esophageal cancer. J Thorac Cardiovasc Surg. 2008;135(2):255–60.
53. Yang K, Zhao X, Waxman D, Zhao XM. Predicting drug-disease associations with heterogeneous network embedding. Chaos. 2019;29(12):123109.
54. Xie WB, Yan H, Zhao XM. EmDL: extracting miRNA–drug interactions from literature. IEEE/ACM Trans Comput Biol Bioinform. 2019;16(5):1722–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.