

Research Article

miRNA-Disease Association Prediction with Collaborative Matrix Factorization

Zhen Shen,¹ You-Hua Zhang,² Kyungsook Han,³ Asoke K. Nandi,⁴
Barry Honig,⁵ and De-Shuang Huang¹

¹*Institute of Machine Learning and Systems Biology, School of Electronics and Information Engineering, Tongji University, Shanghai 201804, China*

²*School of Information and Computer, Anhui Agricultural University, Changjiang West Road 130, Hefei, Anhui, China*

³*Department of Computer Science and Engineering, Inha University, Incheon, Republic of Korea*

⁴*Department of Electronic and Computer Engineering, Brunel University London, Uxbridge UB8 3PH, UK*

⁵*Center for Computational Biology and Bioinformatics, Columbia University, 1130 St. Nicholas Avenue, Room 815, New York, NY 10032, USA*

Correspondence should be addressed to De-Shuang Huang; dshuang@tongji.edu.cn

Received 31 March 2017; Accepted 2 May 2017; Published 28 September 2017

Academic Editor: Fang-Xiang Wu

Copyright © 2017 Zhen Shen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As one of the factors in the noncoding RNA family, microRNAs (miRNAs) are involved in the development and progression of various complex diseases. Experimental identification of miRNA-disease association is expensive and time-consuming. Therefore, it is necessary to design efficient algorithms to identify novel miRNA-disease association. In this paper, we developed the computational method of Collaborative Matrix Factorization for miRNA-Disease Association prediction (CMFMDA) to identify potential miRNA-disease associations by integrating miRNA functional similarity, disease semantic similarity, and experimentally verified miRNA-disease associations. Experiments verified that CMFMDA achieves intended purpose and application values with its short consuming-time and high prediction accuracy. In addition, we used CMFMDA on Esophageal Neoplasms and Kidney Neoplasms to reveal their potential related miRNAs. As a result, 84% and 82% of top 50 predicted miRNA-disease pairs for these two diseases were confirmed by experiment. Not only this, but also CMFMDA could be applied to new diseases and new miRNAs without any known associations, which overcome the defects of many previous computational methods.

1. Introduction

MicroRNAs (miRNAs) are a class of short noncoding RNAs (19~25 nt), which normally regulate gene expression and protein production by targeting messenger RNAs (mRNAs) at the posttranscriptional level [1–9]. Since the first two miRNA lin-4 and let-7 were found in 1993 and 2000 [10, 11], thousands of miRNAs have been detected in eukaryotic organisms ranging from nematodes to humans. The latest version of miRBase contains 26845 entries and more than 2000 miRNAs have been detected in human [12–14]. With the development of bioinformatics and the progress of miRNA-related projects, researches are gradually focused on the function of miRNAs. Existing studies have shown that miRNAs are involved in many important biological processes

[15, 16], like cell differentiation [17], proliferation [18], signal transduction [19], viral infection [20], and so on. Therefore, it is easy to find that miRNAs have close relationship with various human complex diseases [12, 21–26]. For example, researchers found that mir-433 is upregulated in gastric carcinoma by regulating the expression of GRB2, which is a known tumour-associated protein [27]. Mir-126 can not only function as an inhibitor to suppress the growth of colorectal cancer cells by its overexpression, but also can help to differentiate between malignant and normal colorectal tissue [28]. Besides, the change of mir-17~92 miRNA cluster expression has close relationship with kidney cyst growth in polycystic kidney disease [29]. Considering the close relationship between miRNA and disease, we should try all means to excavate all latent associations between miRNA

and disease and to facilitate the diagnose, prevention, and treatment human complex disease [30–33]. However, using experimental methods to identify miRNA-disease association is expensive and time-consuming. As the miRNA-related theories are becoming more and more common, such as the prediction model about miRNA and disease, the function of miRNA in biological processes, and signaling pathways, new therapies are urgently needed for the treatment of complex disease; it is necessary to develop powerful computational methods to reveal potential miRNA-disease associations [12, 15, 20, 34–40].

Previous studies had shown that functionally similar miRNAs always appear in similar diseases; therefore many computational models were proposed to identify novel miRNA-disease associations [13, 41–46]. For example, Jiang et al. [31] analyzed and improved disease-gene prediction model, introduced the principle of hypergeometric distribution and how to use it, and discussed its application in prediction model and its actual effect. In order to realize the prediction function of the improved model, they used different types of dataset including miRNA functional similarity data, disease phenotype similarity data, and the known human disease-miRNA association data. Therefore, the prediction accuracy of this method is greatly impacted by miRNA neighbor information and miRNA-target interaction prediction. Chen et al. [47] reported a new method HGIMDA to identify novel miRNA-disease association by using heterogeneous graph inference. This algorithm can get better prediction accuracy by integrating known miRNA-disease associations, miRNA functional similarity, disease semantic similarity, and Gaussian interaction profile kernel similarity for diseases and miRNAs. In addition, HGIMDA could be applied for new diseases and new miRNAs which do not have any known association. Li et al. [48] proposed the computational model Matrix completion for MiRNA-disease association prediction (MCMMDA) to predict miRNA-disease associations. This model only uses known miRNA-disease associations and achieved better prediction performance. The limitation of MCMMDA is that it could not be applied for new diseases and new miRNAs which do not have any known association. You et al. [49] developed model Path-Based MiRNA-Disease Association Prediction (PBMMDA) to predict miRNA-disease associations by integrating known human miRNA-disease associations, miRNA functional similarity, disease semantic similarity, and Gaussian interaction profile kernel similarity for miRNAs and diseases. Depth-first search algorithm was used in this model to identify novel miRNA-disease associations. Benefiting from effective algorithm and reliable biological datasets, PBMMDA has better prediction performance. Furthermore, Xu et al. [50] introduced an approach to identify disease-related miRNAs by the miRNA target-dysregulated network (MTDN). Furthermore, in order to distinguish and identify disease-related miRNAs from candidate, a SVM classifier based on radial basis function and the lib SVM package had been proposed. Researches have shown that miRNAs can functionally interact with environmental factors (EFs) to affect and determine human complex disease. Chen [51] proposed model miREFRWR to predict the association between disease and miRNA-EF

interactions. Random walks theory was applied on miRNA similarity network and EF similarity network. In addition, drug chemical structure similarity, miRNA function similarity, and networked-based similarity were also used in miREFRWR. Based on these biological datasets and efficient calculation method, miREFRWR could be an effective tool in computational biology. What is more, Chen et al. [52] also proposed a computational model RKNNMDA to predict the potential associations between miRNA and disease. Four biological datasets, experimentally verified human miRNA-disease associations, miRNA functional similarity, disease semantic similarity, and Gaussian interaction profile kernel similarity for miRNAs and diseases were integrated into RKNNMDA. It can be found that the prediction accuracy of RKNNMDA is excellent. Moreover, RKNNMDA could be applied for new diseases which do not have any known related miRNA information.

Generally speaking, current prediction model on miRNA-disease association is still demonstrating some shortcomings. For example, unreliable datasets have a great influence on the accuracy of prediction model, such as miRNA-target interactions and disease-genes associations. In addition, for miRNAs and diseases which do not have any known associations, we cannot use some of the existing models to predict its relevant information. In other words, we need to design and develop a new effective computational model. According to the assumption that functionally similar miRNAs always appear in similar diseases, we introduce the model of Collaborative Matrix Factorization for MiRNA-Disease Association prediction (CMFMDA) to reveal novel miRNA-disease association by integrating experimentally validated miRNA-disease associations, miRNA functional similarity information, and disease semantic similarity information. For CMFMDA, we can obtain its test results with three different ways: 5-fold CV, Local LOOCV, and global LOOCV. The AUCs of these three methods are 0.8697, 0.8318, and 0.8841, respectively, which suggest that CMFMDA is a reliable and efficient prediction model. And then, we use two cases: Esophageal Neoplasms and Kidney Neoplasms, to evaluate the performance of CMFMDA. In both of these two important diseases, 42 and 41 out of top 50 predicted miRNA-disease associations were confirmed by recent experimental literatures, respectively. In addition, experiments show that CMFMDA can be applied for diseases and miRNAs without any known association.

2. Materials and Methods

2.1. Human miRNA-Disease Associations. We obtained information about the associations between miRNA and disease from HMDD, including 5430 experimentally confirmed human miRNA-diseases associations about 383 diseases and 495 miRNAs. Adjacency matrix A is proposed to describe the association between miRNA and disease. If miRNA $m(i)$ is associated with disease $d(j)$, the entity $A(m(i), d(j))$ is 1, otherwise 0. Furthermore, we declared two variables nm and nd to represent the number of miRNAs and diseases investigated in this paper, respectively.

2.2. MiRNA Functional Similarity. Base on the assumption that miRNAs with similarity functions are regarded to be involved in similar diseases, Wang et al. [42] present a method to calculate the miRNA functional similarity score. We downloaded miRNA functional similarity scores from <http://www.cuilab.cn/files/images/cuilab/misim.zip> and constructed matrix SM to represent the miRNA function similarity network, where the entity $SM(m(i), m(j))$ represents the functional similarity score between miRNA $m(i)$ and $m(j)$.

2.3. Disease Semantic Similarity. In this paper, disease can be described as a Directed Acyclic Graph (DAG) and $DAG(D) = (D, T(D), E(D))$ was used to describe disease D , where $T(D)$ is the node set including all ancestor nodes of D and D itself and $E(D)$ is the corresponding links set including the direct edges from parent nodes to child nodes. The semantic value of disease D in $DAG(D)$ is defined as follows:

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

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

where Δ is the semantic contribution factor. For disease D , the contribution of itself to the semantic value of disease D is 1. However, with the growth of the distance between D and other disease, the contribution will fall. Therefore, disease terms in the same layer would have the same contribution to the semantic value of disease D .

If there is much in common between two diseases in DAG, their semantic similarity will become larger. Therefore, the semantic similarity between diseases $d(i)$ and $d(j)$ can be defined as follows:

$$SD(d(i), d(j)) = \frac{\sum_{t \in T(i) \cap T(j)} (D1_i(t) + D1_j(t))}{DV1(i) + DV1(j)}, \quad (2)$$

where SD is the disease semantic similarity matrix.

2.4. CMFMDA. In this study, we developed the computational model of Collaborative Matrix Factorization for MiRNA-Disease Association prediction (CMFMDA) to predict novel miRNA-disease associations [53]. The flow of CMFMDA is shown in Figure 1.

In the first step in Figure 1, we will get the final miRNA similarity matrix SM and diseases similarity matrix SD by integrating miRNA functional similarity network, disease semantic similarity network, and experimentally verified miRNA-disease associations.

Then, we use WKNKN [54] to estimate the association probability for these unknown cases based on their known neighbors.

Thirdly, Collaborative Matrix Factorization was used to obtain the final prediction F . This step contains three parts:

- (1) For the input matrix Y , this step adopts singular value decomposition to get the initial value of A and B .

$$[U, S, V] = SVD(Y, k),$$

$$A = US_k^{1/2}, \quad (3)$$

$$B = VS_k^{1/2}.$$

- (2) We use L to represent the objection function and use a_i and b_j to represent the i th and j th row vectors of A and B . Two alternative update rules (one for updating matrix A and one for updating matrix B) were derived by setting $\partial L / \partial A = 0$ and $\partial L / \partial B = 0$. According to alternating least squares, these two update rules are run alternately until convergence.

$$\min_{A, B} \|Y - AB^T\|_F^2 + \lambda_l (\|A\|_F^2 + \|B\|_F^2) + \lambda_m \|SM - AA^T\|_F^2 + \lambda_d \|SD - BB^T\|_F^2$$

$$A = (YB + \lambda_d S_d A) (B^T B + \lambda_l I_k + \lambda_d A^T A)^{-1} \quad (4)$$

$$B = (Y^T A + \lambda_m S_m B) (A^T A + \lambda_l I_k + \lambda_m B^T B)^{-1}.$$

Finally, the predicted matrix for miRNA-disease associations is then obtained by multiplying A and B .

3. Results

3.1. Performance Evaluation. Based on the known miRNA-disease associations obtained from HMDD database [55], the predictive performance of CMFMDA is evaluated through two ways: Local and global LOOCV. Not only that, three computational models: WBSMDA [4], RLSMDA [12], and NCPMDA [56], were introduced to compare the prediction performance with CMFMDA. To obtain relevant miRNA information for the chosen disease d , all association related to disease d was left out, and the rest of the associations serve as a training set to get prediction association by CMFMDA. For cross-validation, the difference between local LOOCV and global LOOCV is that all diseases would be investigated simultaneously or not. Furthermore, Receiver-Operating Characteristics (ROC) were used to express the difference between true positive rate (TPR, sensitivity) and false positive rate (FPR, $1 - \text{specificity}$) at different thresholds. In this case, sensitivity indicates that the percentage of the test miRNA-disease association which obtained ranks higher than the given threshold. Meanwhile, specificity indicates the percentage of miRNA-disease associations below the threshold. What is more, Area under the ROC curve (AUC) could be calculated to demonstrate the prediction performance of CMFMDA. $AUC = 1$ showed that the model has perfect prediction ability; $AUC = 0.5$ indicates random prediction ability.

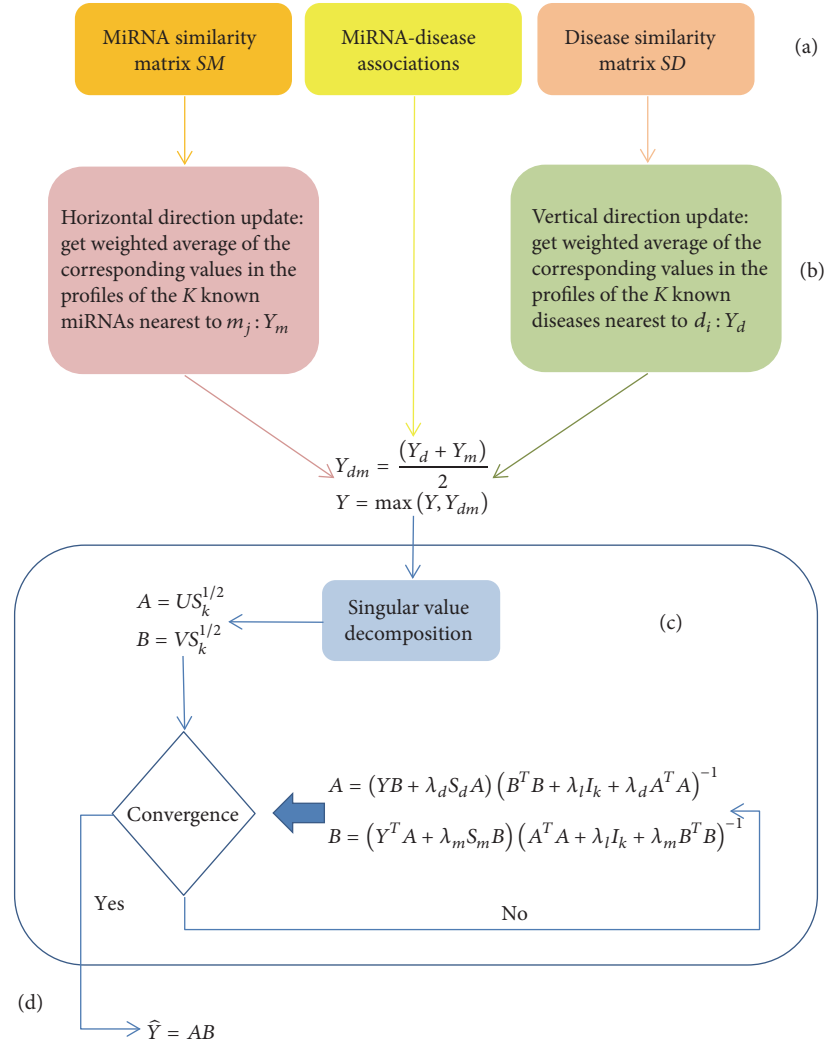


FIGURE 1: Flowchart of potential miRNA-disease associations prediction based on CMFMDA.

To illustrate the performance of CMFMDA, we compare it with the existed computational model: NCPMDA, RLSMDA, and WBSMDA. The comparison result has been shown in Figure 2. As a result, these four models obtained AUCs of 0.8841, 0.8630, 0.8501, and 0.7799 in the global LOOCV, respectively. For local LOOCV, these four models obtained AUCs of 0.8318, 0.8198, 0.8068, and 0.7213, respectively. In general, CMFMDA has not only high prediction performance, but also better ability to identify novel miRNA-disease association.

3.2. Case Studies. All diseases in this paper have been investigated by CMFMDA to predict some novel miRNAs which have association with the disease. Here, two case studies, Esophageal Neoplasms and Kidney Neoplasms, were proposed to demonstrate the prediction performance of CMFMDA. In addition, we use two important miRNA-disease association databases to validate the prediction results: miR2Disease [57] and dbDEMC [58]. A final note about validation datasets is that only the associations which were absent from the HMDD database would be used. In

other words, validation datasets have no correlation with the datasets which have been used for prediction.

Esophageal Neoplasms is a serious disease in digestive system, which leads to high death rate [59–61]. Early diagnosis and treatment is essential for improving patient's survival [62, 63]. Here, we use CMFMDA to identify potential miRNAs associated with Esophageal Neoplasms. As a result, 9 out of the top 10 and 42 out of the top 50 predicted related miRNAs were experimentally confirmed to be associated with Esophageal Neoplasms (See Table 1). For example, mir-133b can inhibit the cell growth and invasion of esophageal squamous cell carcinoma (ESCC) by targeting Fascin homolog 1 [59]. The expression level of mir-335 is an independent prognostic factor in ESCC, which might be a potential valuable biomarker for ESCC [64].

As a common urologic malignancy, the morbidity and mortality of Kidney Neoplasm have been shown to rise gradually [65–68]. Renal cell carcinoma (RCC) can be divided into several different types of cancer [69–71], including chromophobe RCC (CHRCC), collecting duct carcinoma (CDC),

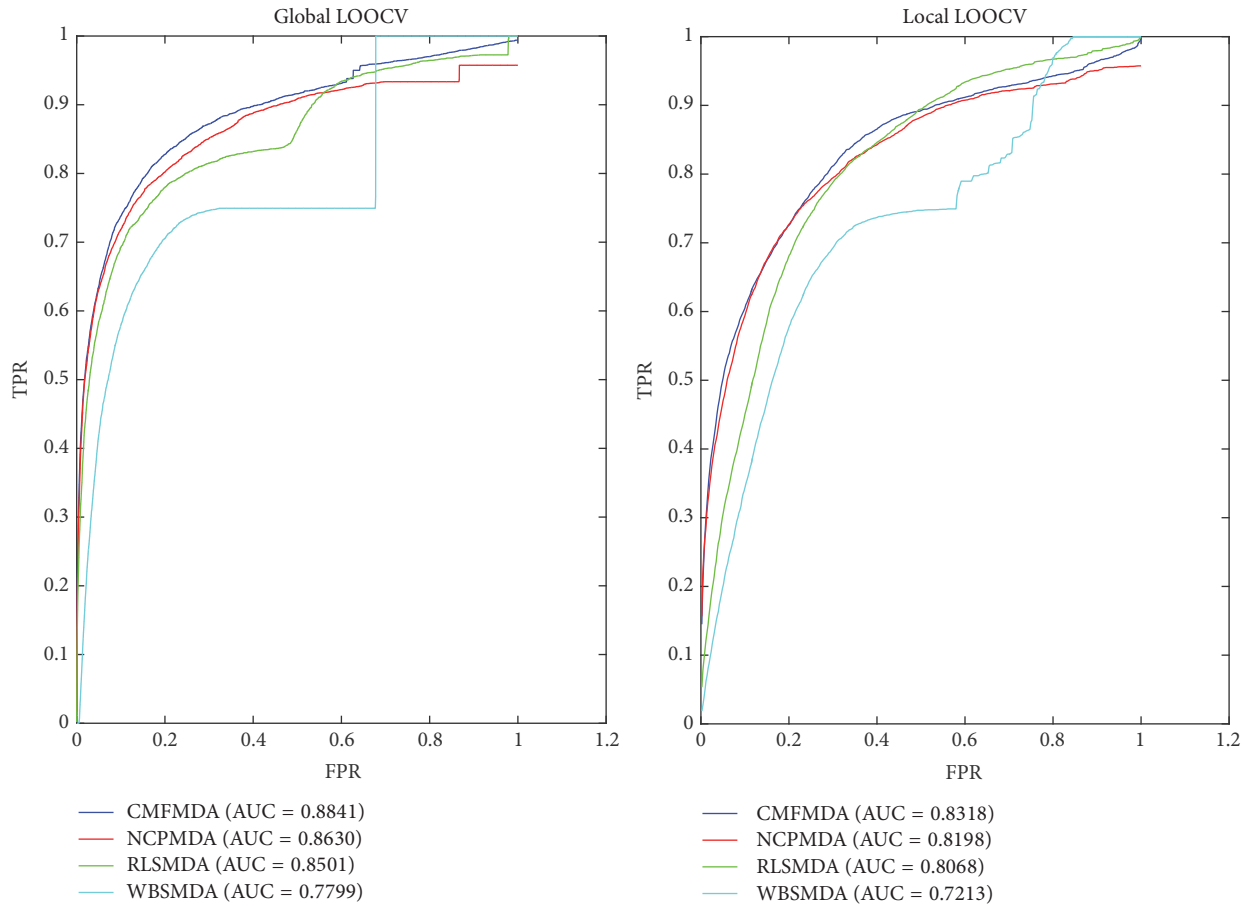


FIGURE 2: Performance comparisons between CMFMMDA and three state-of-the-art disease-miRNA association prediction models (NCPMDA, RLSMDA, and WBSMDA) in terms of ROC curve and AUC based on local and global LOOCV, respectively. As a result, CMFMMDA achieved AUCs of 0.8841 and 0.8318 in the global and local LOOCV, significantly outperforming all the previous classical models.

clear cell RCC (CCRCC), and papillary RCC (PRCC). Previous studies have shown that miRNAs play a significant part in Kidney Neoplasm [72–74]. In this paper, CMFMMDA was employed to identify potential miRNAs associated with Kidney Neoplasms. As a result, 9 out of the top-10 candidates and 41 out of the top-50 candidates of Kidney Neoplasm related miRNAs were confirmed by dbDEMC and miR2Ddisease (See Table 2). For example, the serum level of mi-210 may be used as a novel noninvasive biomarker for the detection of CCRCC [75]. Experiment results demonstrate that mir-9 expression is correlative not only with the development of CCRCC, but also with the development of metastatic recurrence [76].

The results of cross-validation and independent case studies show that CMFMMDA can satisfy the needs to identify potential miRNA-disease associations. Furthermore, all diseases in HMDD have been investigated by CMFMMDA to predict potential miRNAs (See Supplementary Table 1 in Supplementary Material available online at <https://doi.org/10.1155/2017/2498957>). We hope that potential disease-miRNA association predicted by CMFMMDA could be confirmed by further biological experiments.

4. Discussion

According to the assumption that functionally similar miRNAs are often associated with similar diseases, we proposed the computational model of Collaborative Matrix Factorization for MiRNA-Disease Association prediction (CMFMMDA) to identify potential miRNA-disease associations by integrating miRNA functional similarity, disease semantic similarity, and experimentally verified miRNA-disease associations. We compare CMFMMDA with the existing computational model: NCPMDA, RLSMDA, and WBSMDA, and concluded that CMFMMDA has better prediction performance from these four models' obtained AUCs in the global LOOCV or local LOOCV, respectively. There are some reasons for the reliable performance of CMFMMDA. Firstly, several types of experimentally confirmed biological datasets are used in CMFMMDA, including known miRNA-disease associations, miRNA functional similarity network, and disease semantic similarity network, which help improve the prediction performance and reduce variance. Then, CMFMMDA can work not only for known miRNA-disease association, but also for diseases and miRNAs without any known association. Finally,

TABLE 1: We implemented CMFMDA to predict potential Esophageal Neoplasms-related miRNAs. As a result, 9 out of the top 10 and 42 out of the top 50 predicted Esophageal Neoplasms related miRNAs were confirmed based on miR2Disease and dbDEMOC (1st column: top 1–25; 2nd column: top 26–50).

miRNA	Evidence	miRNA	Evidence
hsa-mir-142	dbDEMOC	hsa-mir-30c	dbDEMOC
hsa-mir-1	dbDEMOC	hsa-mir-212	Unconfirmed
hsa-mir-16	dbDEMOC	hsa-mir-424	dbDEMOC
hsa-mir-127	dbDEMOC	hsa-mir-429	dbDEMOC
hsa-mir-497	dbDEMOC	hsa-mir-498	dbDEMOC
hsa-mir-200b	dbDEMOC	hsa-mir-340	Unconfirmed
hsa-mir-376c	Unconfirmed	hsa-mir-222	dbDEMOC
hsa-mir-148b	dbDEMOC	hsa-mir-146b	dbDEMOC
hsa-mir-335	dbDEMOC	hsa-mir-195	dbDEMOC
hsa-mir-93	dbDEMOC	hsa-mir-10b	dbDEMOC
hsa-mir-125b	dbDEMOC	hsa-mir-218	Unconfirmed
hsa-mir-18a	dbDEMOC	hsa-mir-181a	dbDEMOC
hsa-mir-17	dbDEMOC	hsa-mir-137	dbDEMOC
hsa-mir-30a	dbDEMOC	hsa-let-7e	dbDEMOC
hsa-mir-133b	dbDEMOC	hsa-mir-181b	dbDEMOC
hsa-mir-135a	dbDEMOC	hsa-mir-106a	dbDEMOC
hsa-mir-107	dbDEMOC, miR2Disease	hsa-mir-19b	dbDEMOC
hsa-mir-224	dbDEMOC	hsa-mir-95	dbDEMOC
hsa-mir-199b	dbDEMOC	hsa-mir-122	Unconfirmed
hsa-mir-221	dbDEMOC	hsa-mir-152	dbDEMOC
hsa-mir-18b	dbDEMOC	hsa-mir-370	dbDEMOC
hsa-mir-191	dbDEMOC	hsa-mir-30d	dbDEMOC
hsa-let-7g	dbDEMOC	hsa-mir-15b	dbDEMOC
hsa-let-7f	Unconfirmed	hsa-mir-629	Unconfirmed
hsa-mir-494	dbDEMOC	hsa-mir-204	Unconfirmed

as a global prediction model, CMFMDA could be used to predict all disease-related miRNA at the same time.

Although CMFMDA has better prediction performance, the limitation still exists in it and needs to be improved in the future. Firstly, CMFMDA may cause bias to miRNAs with more known associated diseases. Secondly, the known miRNA-disease associations with experimental evidences are still insufficient. The prediction performance of CMFMDA will be improved by integrating more reliable biological information [77–86]. Finally, how to more reasonably extract and integrate information from biological datasets should be investigated in the future.

5. Conclusions

Research has shown that the abnormal expression of miRNA plays a crucial role in the occurrence and development of human complex diseases. The in-depth study and analysis of diseases-related miRNA could help find new biomarker and therapies and then improve the survival rate of patients.

TABLE 2: We implemented CMFMDA to prioritize candidate miRNAs for Kidney Neoplasms based on known associations in the HMDD database. As a result, 9 out of the top 10 and 41 out of the top 50 predicted Kidney Neoplasms related miRNAs were confirmed based on miR2Disease and dbDEMOC (1st column: top 1–25; 2nd column: top 26–50).

miRNA	Evidence	miRNA	Evidence
hsa-mir-429	dbDEMOC	hsa-mir-34c	dbDEMOC
hsa-mir-200b	dbDEMOC, miR2Disease	hsa-mir-10b	dbDEMOC
hsa-mir-200a	dbDEMOC	hsa-mir-9	dbDEMOC
hsa-mir-210	dbDEMOC, miR2Disease	hsa-mir-199b	dbDEMOC
hsa-mir-203	dbDEMOC	hsa-mir-99b	dbDEMOC
hsa-mir-218	dbDEMOC	hsa-mir-126	dbDEMOC, miR2Disease
hsa-mir-127	dbDEMOC	hsa-mir-224	dbDEMOC
hsa-mir-155	dbDEMOC	hsa-mir-195	dbDEMOC
hsa-mir-205	dbDEMOC, miR2Disease	hsa-mir-145	dbDEMOC
hsa-mir-196a	dbDEMOC	hsa-mir-7	dbDEMOC, miR2Disease
hsa-mir-143	dbDEMOC	hsa-mir-142	Unconfirmed
hsa-mir-135a	Unconfirmed	hsa-mir-139	dbDEMOC
hsa-mir-92a	Unconfirmed	hsa-mir-16	dbDEMOC
hsa-mir-19a	dbDEMOC	hsa-mir-183	dbDEMOC
hsa-mir-29c	dbDEMOC, miR2Disease	hsa-mir-296	Unconfirmed
hsa-mir-146a	dbDEMOC	hsa-mir-367	Unconfirmed
hsa-mir-99a	dbDEMOC	hsa-mir-15b	dbDEMOC
hsa-mir-101	dbDEMOC, miR2Disease	hsa-mir-204	dbDEMOC
hsa-mir-199a	dbDEMOC, miR2Disease	hsa-mir-196b	dbDEMOC
hsa-mir-27a	dbDEMOC, miR2Disease	hsa-mir-339	dbDEMOC
hsa-mir-100	dbDEMOC	hsa-mir-26a	dbDEMOC, miR2Disease
hsa-mir-452	dbDEMOC	hsa-mir-302c	Unconfirmed
hsa-mir-34b	dbDEMOC	hsa-mir-302b	Unconfirmed
hsa-mir-93	dbDEMOC	hsa-mir-107	dbDEMOC
hsa-mir-34a	dbDEMOC	hsa-mir-133b	Unconfirmed

Therefore, it is necessary to develop more effective computational models to identify potential miRNA-disease associations. In this paper, we presented a computational model CMFMDA to identify novel miRNA-disease associations. Except for disease semantic similarity and miRNA functional similarity, CMFMDA also uses known miRNA-disease associations to predict miRNA-disease associations. LOOCV was chosen to evaluate the predict performance of CMFMDA. The results of LOOCV and case studies show that CMFMDA has better prediction performance than other models. In other words, as an effective tool, CMFMDA can be used not only to predict potential miRNA-disease associations, but

also to identify new biomarker that gave new direction for diagnosis and treatment of human complex disease.

Disclosure

De-Shuang Huang is the corresponding author of this paper. Professor A. K. Nandi is a Distinguished Visiting Professor at Tongji University, Shanghai, China.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Acknowledgments

This work was supported by the grants of the National Science Foundation of China, nos. 61520106006, 61732012, 31571364, 61532008, U1611265, 61672382, 61772370, 61402334, 61472173, and 61472282, and China Postdoctoral Science Foundation [Grant nos. 2015M580352, 2017M611619, and 2016M601646].

References

- [1] V. Ambros, "microRNAs: tiny regulators with great potential," *Cell*, vol. 107, no. 7, pp. 823–826, 2001.
- [2] V. Ambros, "The functions of animal microRNAs," *Nature*, vol. 431, no. 7006, pp. 350–355, 2004.
- [3] D. P. Bartel, "MicroRNAs: genomics, biogenesis, mechanism, and function," *Cell*, vol. 116, no. 2, pp. 281–297, 2004.
- [4] X. Chen, C. C. Yan, X. Zhang et al., "WBSMDA: within and between score for MiRNA-disease association prediction," *Scientific Reports*, vol. 6, Article ID 21106, 2016.
- [5] R. Ibrahim, N. A. Yousri, M. A. Ismail, and N. M. El-Makky, "MiRNA and gene expression based cancer classification using self-learning and co-training approaches," in *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine, IEEE BIBM 2013*, pp. 495–498, Shanghai, China, December 2013.
- [6] Y. Katayama, M. Maeda, K. Miyaguchi et al., "Identification of pathogenesis-related microRNAs in hepatocellular carcinoma by expression profiling," *Oncology Letters*, vol. 4, no. 4, pp. 817–823, 2012.
- [7] G. Meister and T. Tuschl, "Mechanisms of gene silencing by double-stranded RNA," *Nature*, vol. 431, no. 7006, pp. 343–349, 2004.
- [8] D. S. Huang and H.-J. Yu, "Normalized feature vectors: a novel alignment-free sequence comparison method based on the numbers of adjacent amino acids," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 10, no. 2, pp. 457–467, 2013.
- [9] S.-P. Deng and D. S. Huang, "SFAPS: an R package for structure/function analysis of protein sequences based on informational spectrum method," *Methods*, vol. 69, no. 3, pp. 207–212, 2014.
- [10] R. C. Lee, R. L. Feinbaum, and V. Ambros, "The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*," *Cell*, vol. 75, no. 5, pp. 843–854, 1993.
- [11] B. J. Reinhart, F. J. Slack, M. Basson et al., "The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*," *Nature*, vol. 403, no. 6772, pp. 901–906, 2000.
- [12] X. Chen and G.-Y. Yan, "Semi-supervised learning for potential human microRNA-disease associations inference," *Scientific Reports*, vol. 4, article 5501, 2014.
- [13] S. Bandyopadhyay, R. Mitra, U. Maulik, and M. Q. Zhang, "Development of the human cancer microRNA network," *Silence*, vol. 1, article 6, 2010.
- [14] A. Kozomara and S. Griffiths-Jones, "miRBase: annotating high confidence microRNAs using deep sequencing data," *Nucleic Acids Research*, vol. 42, pp. D68–D73, 2013.
- [15] X. Chen, C. Clarence Yan, X. Zhang et al., "RBMMDA: predicting multiple types of disease-microRNA associations," *Scientific Reports*, vol. 5, Article ID 13877, 2015.
- [16] D. P. Bartel, "MicroRNAs: target recognition and regulatory functions," *Cell*, vol. 136, no. 2, pp. 215–233, 2009.
- [17] E. A. Miska, "How microRNAs control cell division, differentiation and death," *Current Opinion in Genetics and Development*, vol. 15, no. 5, pp. 563–568, 2005.
- [18] A. M. Cheng, M. W. Byrom, J. Shelton, and L. P. Ford, "Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis," *Nucleic Acids Research*, vol. 33, no. 4, pp. 1290–1297, 2005.
- [19] Q. Cui, Z. Yu, E. O. Purisima, and E. Wang, "Principles of microRNA regulation of a human cellular signaling network," *Molecular Systems Biology*, vol. 2, article 46, 2006.
- [20] X. Chen, M. X. Liu, and G. Y. Yan, "RWRMDA: predicting novel human microRNA-disease associations," *Molecular BioSystems*, vol. 8, no. 10, pp. 2792–2798, 2012.
- [21] J. Li, Y. Liu, X. Xin et al., "Evidence for positive selection on a number of microRNA regulatory interactions during recent human evolution," *PLoS Genetics*, vol. 8, no. 3, Article ID e1002578, 2012.
- [22] K. Chen and N. Rajewsky, "Natural selection on human microRNA binding sites inferred from SNP data," *Nature Genetics*, vol. 38, no. 12, pp. 1452–1456, 2006.
- [23] M. A. Saunders, H. Liang, and W.-H. Li, "Human polymorphism at microRNAs and microRNA target sites," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 9, pp. 3300–3305, 2007.
- [24] P. Sethupathy and F. S. Collins, "MicroRNA target site polymorphisms and human disease," *Trends in Genetics*, vol. 24, no. 10, pp. 489–497, 2008.
- [25] S.-P. Deng, L. Zhu, and D. S. Huang, "Predicting hub genes associated with cervical cancer through gene co-expression networks," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 13, no. 1, pp. 27–35, 2016.
- [26] S.-P. Deng, L. Zhu, and D. S. Huang, "Mining the bladder cancer-associated genes by an integrated strategy for the construction and analysis of differential co-expression networks," *BMC Genomics*, vol. 16, no. 3, supplement 3, article S4, 2015.
- [27] H. Luo, H. Zhang, Z. Zhang et al., "Down-regulated miR-9 and miR-433 in human gastric carcinoma," *Journal of Experimental and Clinical Cancer Research*, vol. 28, article 82, 2009.
- [28] X.-M. Li, A.-M. Wang, J. Zhang, and H. Yi, "Down-regulation of miR-126 expression in colorectal cancer and its clinical significance," *Medical Oncology*, vol. 28, no. 4, pp. 1054–1057, 2011.
- [29] V. Patel, D. Williams, S. Hajarnis et al., "MiR-17~92 miRNA cluster promotes kidney cyst growth in polycystic kidney

- disease,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 26, pp. 10765–10770, 2013.
- [30] G. A. Calin and C. M. Croce, “MicroRNA signatures in human cancers,” *Nature Reviews Cancer*, vol. 6, no. 11, pp. 857–866, 2006.
- [31] Q. Jiang, Y. Hao, G. Wang et al., “Prioritization of disease microRNAs through a human phenome-microRNAome network,” *BMC Systems Biology*, vol. 4, supplement 1, article S2, 2010.
- [32] Q. Jiang, G. Wang, S. Jin, Y. Li, and Y. Wang, “Predicting human microRNA-disease associations based on support vector machine,” *International Journal of Data Mining and Bioinformatics*, vol. 8, no. 3, pp. 282–293, 2013.
- [33] C.-H. Zheng, D. S. Huang, L. Zhang, and X.-Z. Kong, “Tumor clustering using nonnegative matrix factorization with gene selection,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 4, pp. 599–607, 2009.
- [34] X. Chen, “Predicting lncRNA-disease associations and constructing lncRNA functional similarity network based on the information of miRNA,” *Scientific Reports*, vol. 5, Article ID 13186, 2015.
- [35] X. Chen, C. C. Yan, C. Luo, W. Ji, Y. Zhang, and Q. Dai, “Constructing lncRNA functional similarity network based on lncRNA-disease associations and disease semantic similarity,” *Scientific Reports*, vol. 5, Article ID 11338, 2015.
- [36] X. Chen and G.-Y. Yan, “Novel human lncRNA-disease association inference based on lncRNA expression profiles,” *Bioinformatics*, vol. 29, no. 20, pp. 2617–2624, 2013.
- [37] X. Chen, M.-X. Liu, Q.-H. Cui, and G.-Y. Yan, “Prediction of disease-related interactions between microRNAs and environmental factors based on a semi-supervised classifier,” *PLoS ONE*, vol. 7, no. 8, Article ID e43425, 2012.
- [38] X. Chen, “KATZLDA: KATZ measure for the lncRNA-disease association prediction,” *Scientific Reports*, vol. 5, Article ID 16840, 2015.
- [39] C.-H. Zheng, L. Zhang, V. T.-Y. Ng, S. C.-K. Shiu, and D. S. Huang, “Molecular pattern discovery based on penalized matrix decomposition,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 8, no. 6, pp. 1592–1603, 2011.
- [40] L. Zhu, S.-P. Deng, and D. S. Huang, “A two-stage geometric method for pruning unreliable links in protein-protein networks,” *IEEE Transactions on Nanobioscience*, vol. 14, no. 5, pp. 528–534, 2015.
- [41] M. Lu, Q. Zhang, M. Deng et al., “An analysis of human microRNA and disease associations,” *PLoS ONE*, vol. 3, no. 10, Article ID e3420, 2008.
- [42] D. Wang, J. Wang, M. Lu, F. Song, and Q. Cui, “Inferring the human microRNA functional similarity and functional network based on microRNA-associated diseases,” *Bioinformatics*, vol. 26, no. 13, pp. 1644–1650, 2010.
- [43] K. Goh, M. E. Cusick, D. Valle, B. Childs, M. Vidal, and A. Barabási, “The human disease network,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 21, pp. 8685–8690, 2007.
- [44] C. Pasquier and J. Gardès, “Prediction of miRNA-disease associations with a vector space model,” *Scientific Reports*, vol. 6, Article ID 27036, 2016.
- [45] T. D. Le, J. Zhang, L. Liu, and J. Li, “Computational methods for identifying miRNA sponge interactions,” *Briefings in Bioinformatics*, p. bbw042, 2016.
- [46] D. S. Huang and C. H. Zheng, “Independent component analysis-based penalized discriminant method for tumor classification using gene expression data,” *Bioinformatics*, vol. 22, no. 15, pp. 1855–1862, 2006.
- [47] X. Chen, C. C. Yan, X. Zhang, Z.-H. You, Y.-A. Huang, and G.-Y. Yan, “HGIMDA: Heterogeneous graph inference for miRNA-disease association prediction,” *Oncotarget*, vol. 7, no. 40, pp. 65257–65269, 2016.
- [48] J.-Q. Li, Z.-H. Rong, X. Chen, G.-Y. Yan, and Z.-H. You, “MCMDA: matrix completion for MiRNA-disease association prediction,” *Oncotarget*, vol. 8, pp. 21187–21199, 2017.
- [49] Z.-H. You, Z.-A. Huang, Z. Zhu et al., “PBMDA: a novel and effective path-based computational model for miRNA-disease association prediction,” *PLOS Computational Biology*, vol. 13, no. 3, Article ID e1005455, 2017.
- [50] J. Xu, C.-X. Li, J.-Y. Lv et al., “Prioritizing candidate disease miRNAs by topological features in the miRNA target-dysregulated network: case study of prostate cancer,” *Molecular Cancer Therapeutics*, vol. 10, no. 10, pp. 1857–1866, 2011.
- [51] X. Chen, “miREFRWR: a novel disease-related microRNA-environmental factor interactions prediction method,” *Molecular BioSystems*, vol. 12, no. 2, pp. 624–633, 2016.
- [52] X. Chen, Q.-F. Wu, and G.-Y. Yan, “RKNNMDA: ranking-based KNN for MiRNA-disease association prediction,” *RNA Biology*, pp. 1–11, 2017.
- [53] X. Zheng, H. Ding, H. Mamitsuka, and S. Zhu, “Collaborative matrix factorization with multiple similarities for predicting drug-target interactions,” in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1025–1033, Chicago, Ill, USA, August 2013.
- [54] A. Ezzat, P. Zhao, M. Wu, X. Li, and C. Kwok, “Drug-target interaction prediction with graph regularized matrix factorization,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 14, no. 3, pp. 646–656, 2016.
- [55] Y. Li, C. Qiu, J. Tu et al., “HMDD v2.0: a database for experimentally supported human microRNA and disease associations,” *Nucleic Acids Research*, vol. 42, pp. D1070–D1074, 2014.
- [56] C. Gu, B. Liao, X. Li, and K. Li, “Network consistency projection for human miRNA-disease associations inference,” *Scientific Reports*, vol. 6, Article ID 36054, 2016.
- [57] Q. Jiang, Y. Wang, Y. Hao et al., “miR2Disease: a manually curated database for microRNA deregulation in human disease,” *Nucleic Acids Research*, vol. 37, supplement 1, pp. D98–D104, 2009.
- [58] Z. Yang, F. Ren, C. Liu et al., “dbDEMC: a database of differentially expressed miRNAs in human cancers,” *BMC Genomics*, vol. 11, supplement 4, article S5, 2010.
- [59] M. Kano, N. Seki, N. Kikkawa et al., “MiR-145, miR-133a and miR-133b: tumor-suppressive miRNAs target FSCN1 in esophageal squamous cell carcinoma,” *International Journal of Cancer*, vol. 127, no. 12, pp. 2804–2814, 2010.
- [60] P. C. Enzinger and R. J. Mayer, “Esophageal cancer,” *The New England Journal of Medicine*, vol. 349, no. 23, pp. 2241–2252, 2003.
- [61] B. He, B. Yin, B. Wang, Z. Xia, C. Chen, and J. Tang, “microRNAs in esophageal cancer (Review),” *Molecular Medicine Reports*, vol. 6, no. 3, pp. 459–465, 2012.
- [62] Z. Xie, G. Chen, X. Zhang et al., “Salivary microRNAs as promising biomarkers for detection of esophageal cancer,” *PLoS ONE*, vol. 8, no. 4, Article ID e57502, 2013.

- [63] J. Wan, W. Wu, Y. Che, N. Kang, and R. Zhang, "Insights into the potential use of microRNAs as a novel class of biomarkers in esophageal cancer," *Diseases of the Esophagus*, vol. 29, no. 5, pp. 412–420, 2016.
- [64] B.-J. Zhang, H.-Y. Gong, F. Zheng, D.-J. Liu, and H.-X. Liu, "Up-regulation of miR-335 predicts a favorable prognosis in esophageal squamous cell carcinoma," *International Journal of Clinical and Experimental Pathology*, vol. 7, no. 9, pp. 6213–6218, 2014.
- [65] A. Jemal, R. Siegel, E. Ward et al., "Cancer statistics, 2006," *CA: A Cancer Journal for Clinicians*, vol. 56, no. 2, pp. 106–130, 2006.
- [66] N. M. A. White, T. T. Bao, J. Grigull et al., "MiRNA profiling for clear cell renal cell carcinoma: biomarker discovery and identification of potential controls and consequences of miRNA dysregulation," *The Journal of Urology*, vol. 186, no. 3, pp. 1077–1083, 2011.
- [67] N. M. A. White, H. W. Z. Khella, J. Grigull et al., "MiRNA profiling in metastatic renal cell carcinoma reveals a tumour-suppressor effect for miR-215," *British Journal of Cancer*, vol. 105, no. 11, pp. 1741–1749, 2011.
- [68] R. Siegel, D. Naishadham, and A. Jemal, "Cancer statistics, 2012," *CA: A Cancer Journal for Clinicians*, vol. 62, no. 1, pp. 10–29, 2012.
- [69] W. M. Linehan, "Genetic basis of kidney cancer: role of genomics for the development of disease-based therapeutics," *Genome Research*, vol. 22, no. 11, pp. 2089–2100, 2012.
- [70] W. M. Linehan, M. M. Walther, and B. Zbar, "The genetic basis of cancer of the kidney," *The Journal of Urology*, vol. 170, no. 6, part 1, pp. 2163–2172, 2003.
- [71] W. M. Linehan and B. Zbar, "Focus on kidney cancer," *Cancer Cell*, vol. 6, no. 3, pp. 223–228, 2004.
- [72] U. Senanayake, S. Das, P. Vesely et al., "miR-192, miR-194, miR-215, miR-200c and miR-141 are downregulated and their common target ACVR2B is strongly expressed in renal childhood neoplasms," *Carcinogenesis*, vol. 33, no. 5, pp. 1014–1021, 2012.
- [73] H. Hidaka, N. Seki, H. Yoshino et al., "Tumor suppressive microRNA-1285 regulates novel molecular targets: aberrant expression and functional significance in renal cell carcinoma," *Oncotarget*, vol. 3, pp. 44–57, 2012.
- [74] M. Redova, A. Poprach, J. Nekvindova et al., "Circulating miR-378 and miR-451 in serum are potential biomarkers for renal cell carcinoma," *Journal of Translational Medicine*, vol. 10, article 55, 2012.
- [75] A. Zhao, G. Li, M. Péoc'h, C. Genin, and M. Gigante, "Serum miR-210 as a novel biomarker for molecular diagnosis of clear cell renal cell carcinoma," *Experimental and Molecular Pathology*, vol. 94, no. 1, pp. 115–120, 2013.
- [76] M. A. T. Hildebrandt, J. Gu, J. Lin et al., "Hsa-miR-9 methylation status is associated with cancer development and metastatic recurrence in patients with clear cell renal cell carcinoma," *Oncogene*, vol. 29, no. 42, pp. 5724–5728, 2010.
- [77] Y. Liu, D. A. Tennant, Z. Zhu, J. K. Heath, X. Yao, and S. He, "DiME: A scalable disease module identification algorithm with application to glioma progression," *PLoS ONE*, vol. 9, no. 2, Article ID e86693, 2014.
- [78] L. Wong, Z.-H. You, Z. Ming, J. Li, X. Chen, and Y.-A. Huang, "Detection of interactions between proteins through rotation forest and local phase quantization descriptors," *International Journal of Molecular Sciences*, vol. 17, no. 1, p. 21, 2015.
- [79] Z. H. You, L. Zhu, C. H. Zheng, H. Yu, S. Deng, and Z. Ji, "Prediction of protein-protein interactions from amino acid sequences using a novel multi-scale continuous and discontinuous feature set," *BMC Bioinformatics*, vol. 15, supplement 15, p. S9, 2014.
- [80] L. Zhu, W.-L. Guo, S.-P. Deng, and D. S. Huang, "ChIP-PIT: enhancing the analysis of ChIP-Seq data using convex-relaxed pair-wise interaction tensor decomposition," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 13, no. 1, pp. 55–63, 2016.
- [81] D. S. Huang, L. Zhang, K. Han, S. Deng, K. Yang, and H. Zhang, "Prediction of protein-protein interactions based on protein-protein correlation using least squares regression," *Current Protein and Peptide Science*, vol. 15, no. 6, pp. 553–560, 2014.
- [82] L. Zhu, Z. You, D. S. Huang, and B. Wang, "t-LSE: a novel robust geometric approach for modeling protein-protein interaction networks," *PLoS ONE*, vol. 8, no. 4, Article ID e58368, 2013.
- [83] D. S. Huang and J.-X. Du, "A constructive hybrid structure optimization methodology for radial basis probabilistic neural networks," *IEEE Transactions on Neural Networks*, vol. 19, no. 12, pp. 2099–2115, 2008.
- [84] D. S. Huang, *Systematic Theory of Neural Networks for Pattern Recognition*, Publishing House of Electronic Industry of China, Beijing, China, 1996.
- [85] D. S. Huang and W. Jiang, "A general CPL-AdS methodology for fixing dynamic parameters in dual environments," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, no. 5, pp. 1489–1500, 2012.
- [86] D. S. Huang, "Radial basis probabilistic neural networks: model and application," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 13, no. 7, pp. 1083–1101, 1999.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

