

Original Research Paper

KRLSMDA: Identifying Human miRNA–Disease Association Based on Similarity and Kronecker Regularized Least Squares Method

Lingzhi Zhu, Weimin Gao*, Guixiang Li, Bufan Ge, Bohan Zhang and Yulu Cao

School of Computer and Information Science, Hunan Institute of Technology, Hengyang 421008, China

Article history

Received: 27-01-2021

Revised: 10-03-2021

Accepted: 12-03-2021

Corresponding Author:

Weimin Gao

School of Computer and Information Science, Hunan Institute of Technology, Hengyang, 421002, China
Email: gwmhy@163.com

Abstract: A growing number of studies have suggested that miRNAs (microRNAs) have associations with human diseases, the design and discovery of drug. But so far, we do not yet fully understand the molecular mechanism of miRNAs in the development of human diseases. Predicting miRNA-disease associations is helpful for understanding the molecular mechanism of miRNAs in the development of human diseases. However, wet-lab experiments are time-consuming and need higher costs to discover miRNA-disease associations. Some computational methods are proposed for predicting miRNA-disease associations, but the prediction performance of these methods needs to be further improved. In this study, we propose a new computational model (KRLSMDA) based on similarity and the Kronecker Regularized Least Squares algorithm. In KRLSMDA, the miRNA functional similarity, the miRNA sequence similarity and the Gaussian Interaction Profile (GIP) kernel similarity are integrated into the comprehensive miRNA similarity. Then we compute the disease semantic similarity, disease functional similarity and the GIP kernel similarity to construct the comprehensive disease similarity based on the disease semantic information, the disease functional information and known miRNA-disease associations, respectively. Finally, the kronecker regularized least squares algorithm is used to predict hidden miRNA-disease associations. The experimental results show that KRLSMDA has achieved the average Area Under the Curve (AUC) values of 0.9181 ± 0.032 and 0.9267 ± 0.022 in 5-fold Cross-Validation (5CV) and 10-fold Cross-Validation (10CV), respectively, which demonstrates KRLSMDA is superior to four competing models. We expect KRLSMDA to be a supplement in the field of biomedical research in the future.

Keywords: MiRNA-Disease Association, Similarity, Gaussian Interaction Profile Kernel, Kronecker Regularized Least Squares

Introduction

MicroRNAs (miRNAs) are short, single-stranded non-coding RNAs (~22 nt) that can regulate gene expression by base pair binding to the 3' Untranslated Region (UTR) of their messenger RNA (mRNA) (Llave *et al.*, 2002; Eulalio *et al.*, 2008). Up to now, miRNAs have been found to be involved in a series of key life processes, including cell growth, differentiation, death and apoptosis (Zhu *et al.*, 2016; Fernando *et al.*, 2012; Lize *et al.*, 2010). A growing number of evidences show that miRNA is directly and closely associated with human diseases, especially complex diseases like cancer

(Akhtar *et al.*, 2016; Calin and Croce, 2006). MiRNA can regulate the expression of disease genes (Ambros, 2004). The abnormality, dysregulation and dysfunction of miRNAs may cause some human diseases (Ha and Kim, 2014). For example, miR-15a and miR-16-1 are down-regulated in chronic lymphocytic leukemia (Lu *et al.*, 2005), prostate cancer (Bonci *et al.*, 2008) and pituitary adenoma (Bottoni *et al.*, 2005) and are induced by the tumor suppressor p53 (He *et al.*, 2007; Chang *et al.*, 2007) and the transcriptional activity of these miRNAs can regulate the p53 response. Furthermore, the expressions of miRNA miR-29b-1, miR-29a, miR-29b-2 and miR-29c in lung cancer (Yanaihara *et al.*, 2006),

breast cancer (Wang *et al.*, 2008), cholangiocarcinoma (Mott *et al.*, 2007), lymphoma (Zhao *et al.*, 2010) and liver cancer (Xiong *et al.*, 2010; Agsalda-Garcia *et al.*, 2020) were down-regulated and could be negatively regulated by oncoprotein Myc (Cimmino *et al.*, 2005). Mir-377-3p is up-regulated in patients with Multiple Sclerosis (Khorasgani *et al.*, 2019). Therefore, identifying hidden miRNA-disease associations can be helpful for probing the molecular mechanism of miRNAs in the development of human diseases and designing appropriate and effective treatments (Xu *et al.*, 2013). But predicting hidden miRNA-disease associations by biological experiments is time-consuming and expensive. This problem can be effectively resolved by computational methods as effective complements of predicting hidden miRNA-disease associations. The successful application of these methods to predict hidden miRNA-disease associations are based on the biological assumption that similar miRNAs are more likely to be associated with similar diseases and vice versa. However, there are some disadvantages in these methods. On the one hand, the prediction performance of these methods needs to be further improved. On the other hand, local information is applied to some methods (Li *et al.*, 2019). It will result in high false positives.

In our work, a new computational model (KRLSMDA) is proposed to predict miRNA-disease associations based on the Regularized Least Squares algorithm of Kronecker product kernel. Based on the miRNA functional information, miRNA sequence information and known miRNA-disease associations, KRLSMDA computes the miRNA functional similarity, the miRNA sequence similarity and the Gaussian Interaction Profile (GIP) kernel similarity to construct a comprehensive miRNA similarity matrix by the linear weighted method. Then two disease semantic similarities (DS_{Mesh} and DS_{DO}) are computed by the representation of direct acyclic graph based on the Mesh database and the Disease Ontology database. We further compute the disease functional similarity and the GIP kernel similarity based on the disease functional information and known miRNA-disease associations, respectively. In order to get a comprehensive similarity matrix, we integrate the disease functional similarity, the Gaussian Interaction Profile (GIP) kernel similarity and two disease semantic similarities (DS_{Mesh} and DS_{DO}). We can compute the kernel of the miRNA-disease pair via the kronecker product of kernel of the miRNA-disease pair. Finally, the kronecker product kernel-based regularized least squares algorithm is applied for predicting the associations scores of miRNA-disease pairs.

To confirm the prediction ability of KRLSMDA, we introduce 5-fold Cross-Validation (5CV) and 10-fold cross-validation (10CV). In 5-fold Cross-Validation (5CV), the AUC value of KRLSMDA is 0.9181 ± 0.032 , while the

AUC values of NetLapRLS, KATZ, NBI and SNMF are 0.8952 ± 0.0011 , 0.8988 ± 0.00062 , 0.8737 ± 0.0015 and 0.8303 ± 0.0038 , respectively. In 10-fold cross-validation (10CV), KRLSMDA achieves an AUC value of 0.9267 ± 0.022 , while NetLapRLS, KATZ, NBI and SNMF have 0.8981 ± 0.0052 , 0.9016 ± 0.00065 , 0.8762 ± 0.00062 and 0.8406 ± 0.002 in 10CV, respectively. 5CV and 10CV computational experiment results show that KRLSMDA is consistently superior to the four methods (NetLapRLS, KATZ, NBI and SNMF). KRLSMDA is effective to predict hidden miRNA-disease associations.

Materials and Methods

Materials

In our study, the known miRNA-disease associations are obtained from the HMDD V1.0 database (Jiang *et al.*, 2009). We sort and standard these downloaded data and obtain 1395 curated miRNA-disease associations, 271 miRNAs and 137 diseases. Let Y be an adjacency matrix with nm rows and nd columns. If there is a known interaction y_{ij} between miRNA mi and disease dj , the value of y_{ij} is 1, otherwise is 0. Therefore the benchmark dataset consists of 1395 curated miRNA-disease associations and 35732 unknown miRNA-disease associations. This benchmark dataset \mathbb{S} is represented as follows:

$$\mathbb{S} = \mathbb{S}^+ \cup \mathbb{S}^- \quad (1)$$

where, \mathbb{S}^+ , \cup and \mathbb{S}^- are 1395 miRNA-disease associations, a union of the sets and 35732 unknown miRNA-disease associations, respectively.

Methods

The Similarity of miRNAs

For miRNAs, we compute three similarities of miRNAs, including the miRNA GIP kernel similarity, the miRNA sequence similarity and the miRNA functional similarity.

GIP kernel has been widely used to compute the similarity between biological entities in biological networks (Van Laarhoven *et al.*, 2011; Zhu *et al.*, 2021). According to the biological assumption that similar miRNA sin cline to interact with similar diseases, the GIP kernel similarity between miRNA m_i and miRNA m_j is also calculated based on the known miRNA-disease associations:

$$KM_{GIP}(m_i, m_j) = \exp(-\gamma_m \|y_{m_i} - y_{m_j}\|^2) \quad (2)$$

$$\gamma_m = \gamma'_m / \left(\frac{1}{nm} \sum_{i=1}^{nm} \|y_{m_i}\|^2 \right), \quad (3)$$

where, y_{mi} and y_{mj} denote the interaction profiles of disease m_i to disease m_j , respectively. γ_m regulates the normalized kernel bandwidth by the original bandwidth γ'_m .

These miRNA nucleotide sequences are downloaded from miR Base database (Kozomara and Griffiths-Jones, 2014). Based on these miRNA nucleotide sequences, we can use Emboss-Needle tool (McWilliam *et al.*, 2013) compute the miRNA sequence similarity $MS_{seq}(m_i, m_j)$ between miRNA m_i and miRNA m_j .

The miRNA functional information is also downloaded from MISIM. We also use the misim method (Wang *et al.*, 2010) to compute the miRNA functional similarity based on the miRNA functional information.

As shown above, three miRNA similarity matrices are computed. To minimize the effect of adjusting too many parameters on the performance of our model, the final miRNA similarity is computed based on three miRNA similarity matrices by the linear weighted method:

$$SM = \frac{KM_{GIP} + MS_{seq} + MS_{fun}}{3} \quad (4)$$

The Similarity of Diseases

For diseases, we compute the disease functional similarity and the disease semantic similarity. Based on the assumption that similar diseases incline to associate with genes (Cheng *et al.*, 2014), we can use the gene function similarity (Lee *et al.*, 2011) to compute the disease functional similarity. Let DS_{Fun} denote the disease functional similarity between disease d_i and disease d_j , then DS_{Fun} can be computed as:

$$DS_{Fun}(d_i, d_j) = \frac{\sum_{1 \leq p \leq m} GS_{G_j}(g_{ip}) + \sum_{1 \leq k \leq n} GS_{G_i}(g_{jk})}{m+n}, \quad (5)$$

in which G_i and G_j are the set of genes which is related to disease d_i and disease d_j , respectively:

$$GS_{G_j}(g_{ip}) = \underset{1 \leq k \leq n}{\text{maximum}}(GF(g_{ip}, g_{jk})), \quad (6)$$

where, $GS_{G_j}(g_{ip})$ is the functional similarity between gene g_{ip} and gene set G_j and $GF(g_{ip}, g_{jk})$ is the gene functional similarity between gene g_{ip} and gene g_{jk} :

$$GF_{Fun}(g_{ip}, g_{jk}) = \begin{cases} 1, & (p = k) \\ LLS(g_{ip}, g_{jk}), & (p, k \in \text{HumanNet}) \\ 0, & (p, k \notin \text{HumanNet}) \end{cases} \quad (7)$$

in which $LLS(g_{ip}, g_{jk})$ is the score of log likelihood between gene g_{ip} and gene g_{jk} .

The Mesh database (Schaefer *et al.*, 2013) and the Disease Ontology database (Kibbe *et al.*, 2015) can be used to compute two disease semantic similarities: DS_{Mesh} and DS_{DO} . For the first semantic similarity of diseases, we use DAG to describe the Mesh database. For disease d_i , $T(d_i)$ is a set of disease d_i and its ancestor and $E(d_i)$ is a edge set. The semantic contribution of disease d_i to disease d_i in the DAG as:

$$\begin{cases} C_{d_i}(d_i) = 1, & (d_i = d_i) \\ C_{d_i}(d_i) = \max\{0.5 * C_{d_i}(d') \mid d' \in (d_i)\}, & (d_i \neq d_i) \end{cases} \quad (8)$$

The disease semantic value $CV(d_i)$ is computed as:

$$CV(d_i) = \sum C_{d_i}(d_i) \quad (9)$$

The disease semantic similarity DS_{Mesh} can be defined as:

$$DS_{Mesh}(d_i, d_j) = \frac{\sum_{d_i \in T(d_i) \cap T(d_j)} (C_{d_i}(d_i) + C_{d_j}(d_j))}{CV(d_i) + CV(d_j)}, \quad (10)$$

where, $C_{d_i}(d_i)$ is the semantic contribution of disease d_i to disease d_i in the DAG.

For the second semantic similarity of diseases, we also use DAG to describe the Disease Ontology database. The disease semantic similarity DS_{DO} can be defined as:

$$DS_{DO}(d_i, d_j) = \frac{|G_i| * |G_j|}{|G_c|^2}, \quad (11)$$

in which, $|G_i|$ and $|G_j|$ are the set of genes which is related to disease d_i and disease d_j , respectively. $|G_c|$ is the number of genes in G_c . G_c is the gene set associated with D_c , which denote the nearest common ancestor of disease d_i and disease d_j in the DAG of the Disease Ontology.

Based on the known miRNA-disease associations, the GIP kernel similarity between disease d_i and disease d_j is calculated as:

$$KD_{GIP}(d_i, d_j) = \exp(-\gamma_d \|y_{d_i} - y_{d_j}\|^2) \quad (12)$$

$$\gamma_d = \gamma'_d / \left(\frac{1}{nd} \sum_{i=1}^{nd} \|y_{d_i}\|^2 \right), \quad (13)$$

where, y_{d_i} and y_{d_j} denote the interaction profiles of disease d_i to disease d_j , respectively. γ_d regulates the normalized kernel bandwidth by the original bandwidth γ'_d .

As shown above, four disease similarity matrices are computed. We compute the final miRNA similarity based on these four disease similarity matrices by the linear weighted method:

$$SD = \frac{DS_{Fum} + DS_{Mesh} + DS_{DO} + KD_{GIP}}{4} \quad (14)$$

KRLSMDA for Predicting miRNA-Disease Associations

As a machine learning-based computational method, the kronecker regularized least squares algorithm has been widely applied in many fields. Driven by the kronecker regularized least squares algorithm of successful applications, we propose a computational method (KRLSMDA) to predict hidden miRNA-disease associations based on the kronecker regularized least squares algorithm. In KRLSMDA, the miRNA-disease associations matrix can be defined as:

$$vec(\tilde{Y})^T = K(K + \sigma I)^{-1} vec(Y_{New}^T) \quad (15)$$

where, \tilde{Y} is the association scores of miRNAs and diseases, vec is the vector symbol, K is a nuclear matrix of miRNA-disease and σ is the regularization parameter. In our study, we set parameter σ to be 1.

Based on the Kronecker product $K_m \otimes K_d$, the kernel K can be computed as:

$$K((m_i, d_j), (m_p, d_k)) = K_m(m_i, m_p) K_d(d_j, d_k) \quad (16)$$

where, K_m is a miRNA similarity matrix and K_d is a disease similarity matrix. To obtain the predictive matrix, we can use a method of the matrix eigenvalue decomposition to calculate the inverse of an $nmnd \times nmnd$ matrix. The kernel K can be defined as:

$$K = K_m \otimes K_d = \underset{\vee}{\vee} \wedge \underset{\vee^T}{\vee^T}, \underset{\vee}{\vee} = \underset{\vee_m}{\vee_m} \otimes \underset{\vee_d}{\vee_d}, \wedge = \wedge_m \otimes \wedge_d \quad (17)$$

In which $\underset{\vee_m}{\vee_m}$ and $\underset{\vee_d}{\vee_d}$ are the unitary of the eigenvectors of the miRNA similarity matrix K_m and the disease similarity matrix K_d , respectively. \otimes denotes the Kronecker product, $\underset{\vee_m}{\vee_m}$ is the eigenvalue diagonal matrices of the miRNA similarity matrix K_m and $\underset{\vee_d}{\vee_d}$ is the eigenvalue diagonal matrices of the disease similarity matrix K_d . The decomposition process of the eigenvalue diagonal matrices is computed as:

$$K_m = \underset{\vee_m}{\vee_m} \wedge_m \underset{\vee_m^T}{\vee_m^T} \quad (18)$$

$$K_d = \underset{\vee_d}{\vee_d} \wedge_d \underset{\vee_d^T}{\vee_d^T} \quad (19)$$

So, we can compute the predictive miRNA-disease associations' matrixes below:

$$\tilde{Y} = \underset{\vee_m}{\vee_m} Z^T \underset{\vee_d^T}{\vee_d^T} \quad (20)$$

$$vec(Z) = (\wedge_m \otimes \wedge_d)(\wedge_m \otimes \wedge_d + \sigma I)^{-1} vec(\underset{\vee_d}{\vee_d} Y^T \underset{\vee_m}{\vee_m}) \quad (21)$$

Results and Discussions

Performance Evaluation

For 5CV, known miRNA-disease associations S^+ are randomly divided into 5 exclusive subsets as below:

$$S^+ = S_1^+ \cup S_2^+ \cup \dots \cup S_5^+ \quad (22)$$

With:

$$\emptyset = S_1^+ \cap S_2^+ \cap \dots \cap S_5^+ \quad (23)$$

$$|S_1^+| \approx |S_2^+| \approx \dots \approx |S_5^+| \quad (24)$$

in which, \cup , \cap and \emptyset are symbols of union, intersection and the empty set, respectively. Each subset (e.g., S_1^+) in turn, acts as a test set and the remaining four subsets as the training set. S_5^+ denotes the fifth subset. 5CV is performed 10times, with the average of the predicted results as the final result.

Similar to the 5CV method, S^+ are randomly also divided into 10 exclusive subsets with equal size as:

$$S^+ = S_1^+ \cup S_2^+ \cup \dots \cup S_{10}^+ \quad (25)$$

With:

$$\emptyset = S_1^+ \cap S_2^+ \cap \dots \cap S_{10}^+ \quad (26)$$

$$|S_1^+| \approx |S_2^+| \approx \dots \approx |S_{10}^+| \quad (27)$$

Each subset (e.g., S_1^+) in turn, acts as a test set and the remaining four subsets as the training set. 10CV is also performed 10 times, with the average of the predicted results as the final result.

Comparison with Different Methods

To evaluate the prediction performance of KRLSMDA, 10-fold CV and LOOCV are applied. The performance of KRLSMDA is compared with other four methods, namely KATZ (Chen *et al.*, 2017; Qu *et al.*, 2017), NetLapRLS (Xia *et al.*, 2010), NBI (Cheng *et al.*, 2014) and SNMF.

As Fig. 1 shows, the AUC value of KRLSMDA is 0.9181 ± 0.032 , while the AUC value of NetLapRLS, KATZ, NBI and SNMF are 0.8952 ± 0.0011 , 0.8988 ± 0.00062 , 0.8737 ± 0.0015 and 0.8303 ± 0.0038 in the 5CV, respectively. This result shows that KRLSMDA obtains the better performance over other four methods.

In 10CV, we also compared KRLSMDA with other four prediction models. As we can see from Fig. 2, KRLSMDA achieves an AUC value of 0.9267 ± 0.022 , while NetLapRLS, KATZ, NBI and SNMF have 0.8981 ± 0.0052 , 0.9016 ± 0.00065 , 0.8762 ± 0.00062 and 0.8406 ± 0.002 in 10CV, respectively.

As shown in the Figs. 1 and 2, obviously, KRLSMDA is better than the other four methods in 5CV and 10CV.

Parameter Analysis

In order to analyze the robustness of KRLSMDA, we quantify the effects of different values of three parameter (σ , γ'_m and γ'_d) on the prediction performance of KRLSMDA in 5CV and 10CV, respectively. Parameter

σ is a regularization parameter in KRLSMDA. The original bandwidth γ'_m and γ'_d are used to regulate the normalized kernel bandwidth.

In Fig. 3, KRLSMDA can obtain AUCs of 0.5541 ± 0.014 , 0.7465 ± 0.0048 , 0.8634 ± 0.047 , 0.9099 ± 0.0033 and 0.9181 ± 0.0032 in 5CV, when $\sigma = 0.2, 0.4, 0.6, 0.8$ and 1.0 , respectively. It is clear that the AUC trend of KRLSMDA increases from 0.5541 ± 0.014 to 0.9181 ± 0.0032 in 5CV when σ increases from 0.2 to 1.0 , respectively.

As shown in Fig. 4, KRLSMDA also describes an increasing AUC trend of KRLSMDA from 0.566 ± 0.015 to 0.9267 ± 0.022 in 10CV, when σ increases from 0.2 to 1.0 , respectively. It is obvious for KRLSMDA to make a better performance when the value of parameter σ is 1.0 . So σ is set to be 1.0 in this study.

Considering that the effect of γ'_m is similar to the effect of γ'_d , we use γ denote γ'_m and γ'_d . Figure 5 shows that AUC values of KRLSMDA are 0.8897 ± 0.0041 , 0.8936 ± 0.04 , 0.9181 ± 0.032 , 0.9055 ± 0.043 and 0.9073 ± 0.045 in 5CV, when γ are $2^{-2}, 2^{-1}, 2^0, 2^1$ and 2^2 , respectively.

Figure 6 shows that KRLSMDA can obtain AUCs of 0.8856 ± 0.0026 , 0.9047 ± 0.033 , 0.9267 ± 0.022 , 0.9244 ± 0.024 and 0.9126 ± 0.034 in 10CV, when γ increases from 2^{-2} to 2^2 .

It is obvious for KRLSMDA to make a better performance when γ is equal to 2^1 .

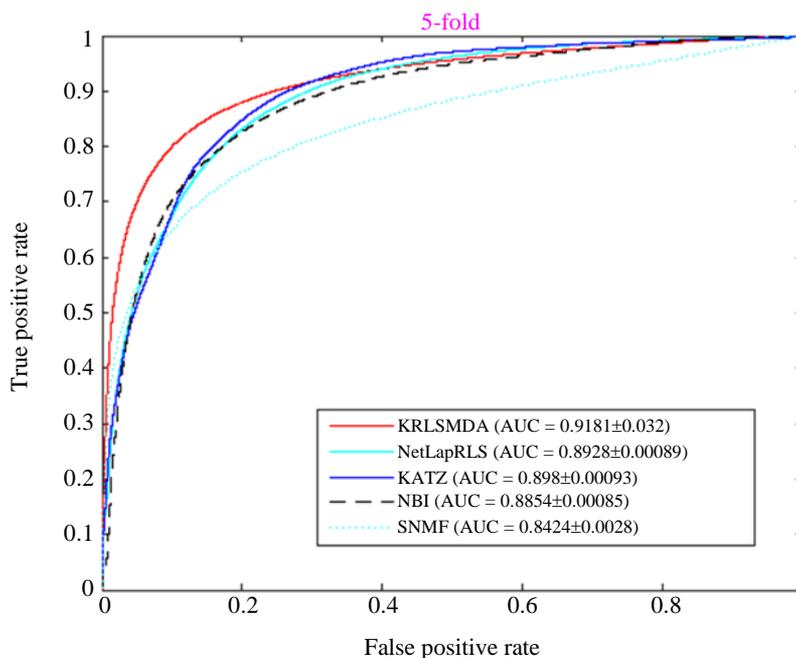


Fig. 1: The prediction performance of different methods in 5CV

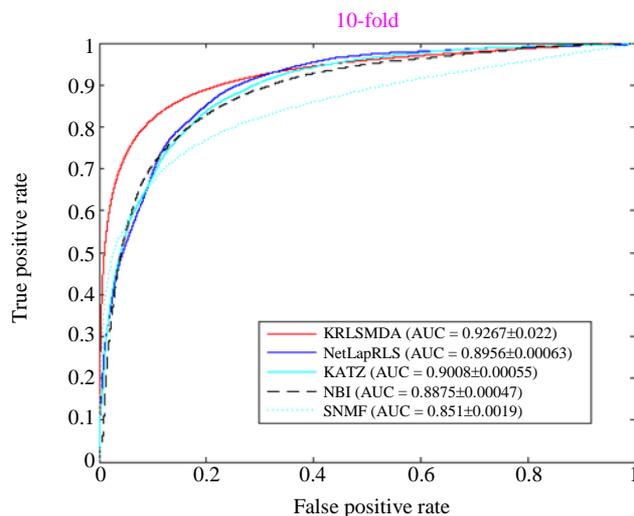


Fig. 2: The prediction performance of different methods in 10CV

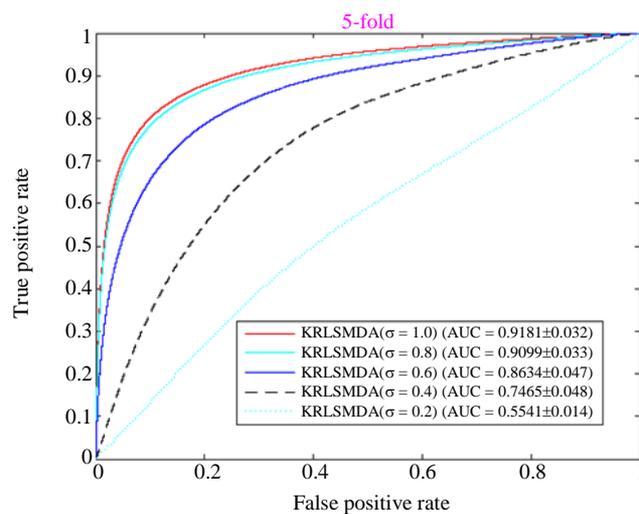


Fig. 3: The prediction performance of KRLSMDA in 5CV when σ ranging from 0.2 to 1.0

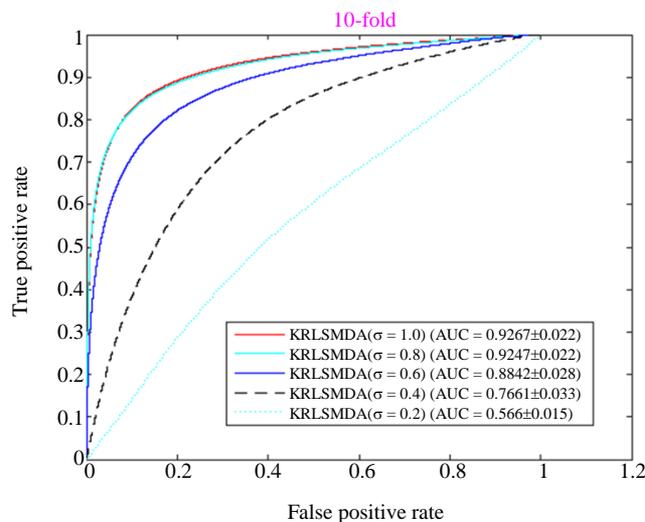


Fig. 4: The prediction performance of KRLSMDA in 10CV when σ ranging from 0.2 to 1.0

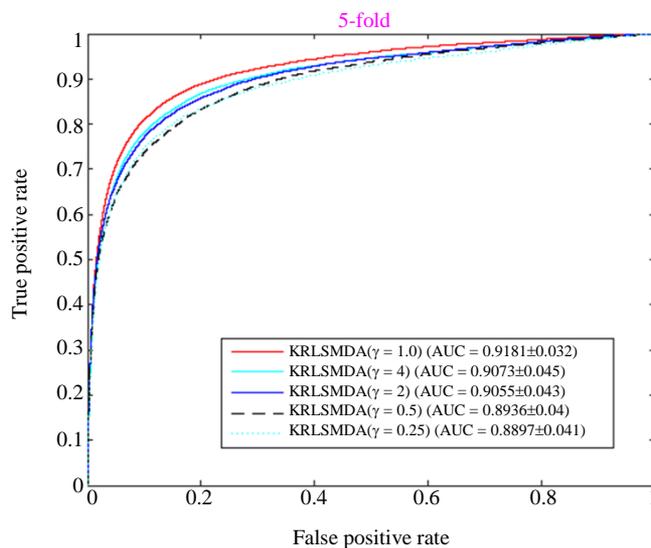


Fig. 5: The prediction performance of KRLSMDA in 5CV when σ ranging from 2^{-2} to 2^2

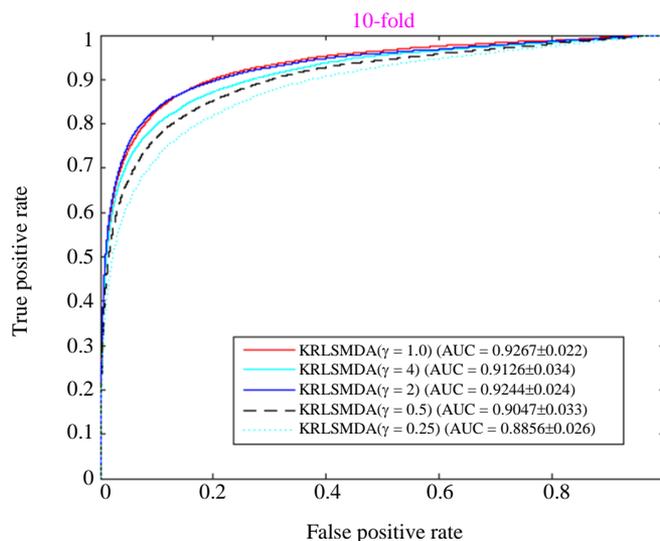


Fig. 6: The prediction performance of KRLSMDA in 10CV when σ ranging from 2^{-2} to 2^2

Conclusion

The KRLSMDA method is proposed to predict miRNA-disease associations based on the Kronecker Regularized Least Squares algorithm. In KRLSMDA, we compute the miRNA functional similarity and miRNA sequence similarity based on the miRNA functional information and miRNA sequence information. Then the disease semantic similarity and disease functional similarity are computed by the disease semantic and functional information, respectively. Finally, we apply kronecker product kernel-based regularized least squares algorithm to predict hidden miRNA-disease associations. The experimental results show that KRLSMDA is effective to predict potential miRNA-disease associations.

However, the limitations of our method are also discussed as follows: (i) Some relevant biological information should be integrated; (ii) some deep learning methods should be used to enhance the prediction performance of KRLSMDA in the future.

Acknowledgement

This work is supported in part by the Scientific Research Foundation of Hunan Provincial Education Department (No.18B46, 20C0561), the Natural Science Foundation of Hunan Province (No.2018JJ2084), the Hengyang Civic Science and Technology Program (2019jh010657, 202010031491), the Teaching Reform

Research Foundation of Hunan Provincial Education Department (Xiang Jiao Tong [2019] 291-1023), Hunan Provincial Innovation and Entrepreneurship Training Program for Undergraduate (S202011528006).

Author's Contributions

Lingzhi Zhu: Designed and performed the experiments, work.

Weimin Gao and Guixiang Li: Participated to collect the materials related to the experiment.

Lingzhi Zhu and Bohan Zhang: Wrote the paper.

Weimin Gao, Bufan Ge and Yulu Cao: Revised the manuscript.

Ethics

The authors declare their responsibility for any ethical issues that may arise after the publication of this manuscript.

Conflict of Interest

The authors declare that they have no competing interests. The corresponding author affirms that all of the authors have read and approved the manuscript.

References

- Agsalda-Garcia, M., Shieh, T., Souza, R., Kamada, N., Loi, N., Oda, R., ... & Shiramizu, B. (2020). Raman-Enhanced Spectroscopy (RESpect) probe for childhood non-Hodgkin lymphoma. *SciMedicine Journal*, 2(1), 1-7. <https://doi.org/10.28991/SciMedJ-2020-0201-1>
- Akhtar, M. M., Micolucci, L., Islam, M. S., Olivieri, F., & Procopio, A. D. (2016). Bioinformatic tools for microRNA dissection. *Nucleic Acids Research*, 44(1), 24-44. <https://doi.org/10.1093/nar/gkv1221>
- Ambros, V. (2004). The functions of animal microRNAs. *Nature*, 431(7006), 350-355. <https://doi.org/10.1038/nature02871>
- Bonci, D., Coppola, V., Musumeci, M., Addario, A., Giuffrida, R., Memeo, L., ... & De Maria, R. (2008). The miR-15a-miR-16-1 cluster controls prostate cancer by targeting multiple oncogenic activities. *Nature Medicine*, 14(11), 1271-1277. <https://doi.org/10.1038/nm.1880>
- Bottoni, A., Piccin, D., Tagliati, F., Luchin, A., Zatelli, M. C., & Degli Uberti, E. C. (2005). miR-15a and miR-16-1 down-regulation in pituitary adenomas. *Journal of Cellular Physiology*, 204(1), 280-285. <https://doi.org/10.1002/jcp.20282>
- Calin, G. A., & Croce, C. M. (2006). MicroRNA signatures in human cancers. *Nature Reviews Cancer*, 6(11), 857-866. <https://doi.org/10.1038/nrc1997>
- Chang, T. C., Wentzel, E. A., Kent, O. A., Ramachandran, K., Mullendore, M., Lee, K. H., ... & Mendell, J. T. (2007). Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. *Molecular Cell*, 26(5), 745-752. <https://doi.org/10.1016/j.molcel.2007.05.010>
- Chen, X., Huang, Y. A., You, Z. H., Yan, G. Y., & Wang, X. S. (2017). A novel approach based on KATZ measure to predict associations of human microbiota with non-infectious diseases. *Bioinformatics*, 33(5), 733-739. <https://doi.org/10.1093/bioinformatics/btw715>
- Cheng, L., Li, J., Ju, P., Peng, J., & Wang, Y. (2014). SemFunSim: a new method for measuring disease similarity by integrating semantic and gene functional association. *PloS One*, 9(6), e99415. <https://doi.org/10.1371/journal.pone.0099415>
- Cimmino, A., Calin, G. A., Fabbri, M., Iorio, M. V., Ferracin, M., Shimizu, M., ... & Croce, C. M. (2005). miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proceedings of the National Academy of Sciences*, 102(39), 13944-13949. <https://doi.org/10.1073/pnas>
- Eulalio, A., Huntzinger, E., & Izaurralde, E. (2008). Getting to the root of miRNA-mediated gene silencing. *Cell*, 132(1), 9-14. <https://doi.org/10.1016/j.cell.2007.12.024>
- Fernando, T. R., Rodriguez-Malave, N. I., & Rao, D. S. (2012). MicroRNAs in B cell development and malignancy. *Journal of Hematology & Oncology*, 5(1), 1-10. <https://doi.org/10.1186/1756-8722-5-7>
- Ha, M., & Kim, V. N. (2014). Regulation of microRNA biogenesis. *Nature reviews Molecular Cell Biology*, 15(8), 509-524. <https://doi.org/10.1038/nrm3838>
- He, L., He, X., & Lim, L. P. (2007). A microRNA component of the p53 tumour suppressor network. *Nature*, 447 (7148), 1130-1134. <https://doi.org/10.1038/nature05939>
- Jiang, Q., Wang, Y., Hao, Y., Juan, L., Teng, M., Zhang, X., ... & Liu, Y. (2009). miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Research*, 37(suppl_1), D98-D104. <https://doi.org/10.1093/nar/gkn714>
- Khorasgani, M. A., Nejad, P. M., Bashfi, M. M. M., & Hedayati, M. (2019). Evaluation of mir-377-3p Expression in Patients with Multiple Sclerosis. *SciMedicine Journal*, 1(2), 48-54. <https://doi.org/10.28991/SciMedJ-2019-0102-2>

- Kibbe, W. A., Arze, C., Felix, V., Mitraga, E., Bolton, E., Fu, G., ... & Schriml, L. M. (2015). Disease Ontology 2015 update: An expanded and updated database of human diseases for linking biomedical knowledge through disease data. *Nucleic Acids Research*, 43(D1), D1071-D1078. <https://doi.org/10.1093/nar/gku1011>
- Kozomara, A., & Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research*, 42(D1), D68-D73. <https://doi.org/10.1093/nar/gkt1181>
- Lee, I., Blom, U. M., Wang, P. I., Shim, J. E., & Marcotte, E. M. (2011). Prioritizing candidate disease genes by network-based boosting of genome-wide association data. *Genome Research*, 21(7), 1109-1121. <https://doi.org/10.1101/gr.118992.110>
- Li, J., Wen, G., Gan, J., Zhang, L., & Zhang, S. (2019). Sparse nonlinear feature selection algorithm via local structure learning. *Emerging Science Journal*, 3(2), 115-129. <https://doi.org/10.28991/esj-2019-01175>
- Lize, M., Pilarski, S., & Dobbstein, M. (2010). E2F1-inducible microRNA 449a/b suppresses cell proliferation and promotes apoptosis. *Cell Death & Differentiation*, 17(3), 452-458. <https://doi.org/10.1038/cdd.2009.188>
- Llave, C., Xie, Z., Kasschau, K. D., & Carrington, J. C. (2002). Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science*, 297(5589), 2053-2056. <https://doi.org/10.1126/science.1076311>
- Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., ... & Golub, T. R. (2005). MicroRNA expression profiles classify human cancers. *Nature*, 435(7043), 834-838. <https://doi.org/10.1038/nature03702>
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., ... & Lopez, R. (2013). Analysis tool web services from the EMBL-EBI. *Nucleic Acids Research*, 41(W1), W597-W600. <https://doi.org/10.1093/nar/gkt376>
- Mott, J. L., Kobayashi, S., Bronk, S. F., & Gores, G. J. (2007). mir-29 regulates Mcl-1 protein expression and apoptosis. *Oncogene*, 26(42), 6133-6140. <https://doi.org/10.1038/sj.onc.1210436>
- Qu, Y., Zhang, H., Liang, C., & Dong, X. (2017). KATZMDA: prediction of miRNA-disease associations based on KATZ model. *IEEE Access*, 6, 3943-3950. <https://doi.org/10.1109/ACCESS.2017.2754409>
- Schaefer, M. H., Lopes, T. J., Mah, N., Shoemaker, J. E., Matsuoka, Y., Fontaine, J. F., ... & Andrade-Navarro, M. A. (2013). Adding protein context to the human protein-protein interaction network to reveal meaningful interactions. *PLoS Comput Biol*, 9(1), e1002860. <https://doi.org/10.1371/journal.pcbi.1002860>
- Van Laarhoven, T., Nabuurs, S. B., & Marchiori, E. (2011). Gaussian interaction profile kernels for predicting drug-target interaction. *Bioinformatics*, 27(21), 3036-3043. <https://doi.org/10.1093/bioinformatics/btr500>
- Wang, D., Wang, J., Lu, M., Song, F., & Cui, Q. (2010). Inferring the human microRNA functional similarity and functional network based on microRNA-associated diseases. *Bioinformatics*, 26(13), 1644-1650. <https://doi.org/10.1093/bioinformatics/btq241>
- Wang, H., Garzon, R., Sun, H., Ladner, K. J., Singh, R., Dahlman, J., ... & Guttridge, D. C. (2008). NF- κ B-YY1-miR-29 regulatory circuitry in skeletal myogenesis and rhabdomyosarcoma. *Cancer cell*, 14(5), 369-381. <https://doi.org/10.1016/j.ccr.2008.10.006>
- Xia, Z., Wu, L. Y., Zhou, X., & Wong, S. T. (2010, September). Semi-supervised drug-protein interaction prediction from heterogeneous biological spaces. In *BMC Systems Biology* (Vol. 4, No. 2, pp. 1-16). BioMed Central. <https://doi.org/10.1186/1752-0509-4-S2-S6>
- Xiong, Y., Fang, J. H., Yun, J. P., Yang, J., Zhang, Y., Jia, W. H., & Zhuang, S. M. (2010). Effects of microRNA-29 on apoptosis, tumorigenicity and prognosis of hepatocellular carcinoma. *Hepatology*, 51(3), 836-845. <https://doi.org/10.1002/hep.23380>
- Xu, G., Zhang, Y., Wei, J., Jia, W., Ge, Z., Zhang, Z., & Liu, X. (2013). MicroRNA-21 promotes hepatocellular carcinoma HepG2 cell proliferation through repression of mitogen-activated protein kinase-kinase 3. *BMC cancer*, 13(1), 1-9. <https://doi.org/10.1186/1471-2407-13-469>
- Yanaihara, N., Caplen, N., Bowman, E., Seike, M., Kumamoto, K., Yi, M., ... & Harris, C. C. (2006). Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer cell*, 9(3), 189-198. <https://doi.org/10.1016/j.ccr.2006.01.025>
- Zhao, J. J., Lin, J., Lwin, T., Yang, H., Guo, J., Kong, W., ... & Cheng, J. Q. (2010). microRNA expression profile and identification of miR-29 as a prognostic marker and pathogenetic factor by targeting CDK6 in mantle cell lymphoma. *Blood, The Journal of the American Society of Hematology*, 115(13), 2630-2639. <https://doi.org/10.1182/blood-2009-09-243147>
- Zhu, L., Yan, C., & Duan, G. (2021). Prediction of Virus-Receptor Interactions Based on Improving Similarities. *Journal of Computational Biology*. <https://doi.org/10.1089/cmb.2020.0544>
- Zhu, L., Zhao, J., Wang, J., Hu, C., Peng, J., Luo, R., ... & Cheng, G. (2016). MicroRNAs are involved in the regulation of ovary development in the pathogenic blood fluke *Schistosoma japonicum*. *PLoS Pathogens*, 12(2), e1005423. <https://doi.org/10.1038/nature03702>