



Inferring Latent Disease-IncRNA Associations by Label-Propagation Algorithm and Random Projection on a Heterogeneous Network

Min Chen^{*†}, Yingwei Deng[†], Ang Li and Yan Tan

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

Long noncoding RNA (lncRNA), a type of more than 200 nucleotides non-coding RNA, is related to various complex diseases. To precisely identify the potential lncRNA–disease association is important to understand the disease pathogenesis, to develop new drugs, and to design individualized diagnosis and treatment methods for different human diseases. Compared with the complexity and high cost of biological experiments, computational methods can quickly and effectively predict potential lncRNA–disease associations. Thus, it is a promising avenue to develop computational methods for lncRNA–disease prediction. However, owing to the low prediction accuracy of state-of-the-art methods, it is vastly challenging to accurately and effectively identify lncRNA–disease at present. This article proposed an integrated method called LPARP, which is based on label-propagation algorithm and random projection to address the issue. Specifically, the label-propagation algorithm is initially used to obtain the estimated scores of lncRNA–disease associations, and then random projections are used to accurately predict disease-related lncRNAs. The empirical experiments showed that LPARP achieved good prediction on three gold datasets, which is superior to existing state-of-the-art prediction methods. It can also be used to predict isolated diseases and new lncRNAs. Case studies of bladder cancer, esophageal squamous-cell carcinoma, and colorectal cancer further prove the reliability of the method. The proposed LPARP algorithm can predict the potential lncRNA–disease interactions stably and effectively with fewer data. LPARP can be used as an effective and reliable tool for biomedical research.

Keywords: disease similarity, lncRNA similarity, space projection, computational prediction model, label-propagation algorithm

INTRODUCTION

Long noncoding RNAs (lncRNAs) are more than 200 nucleotides long and lacks protein-coding RNAs (Peng et al., 2019). Studies have shown that lncRNAs are closely related to biological processes such as chromatin modification, transcription, translation, splicing, and epigenetic regulation (Wang and Chang, 2011; Wapinski and Chang, 2011; Song et al., 2014; Sun et al., 2017; Tian et al., 2021; Peng et al., 2021). The abnormal function of lncRNAs can reportedly lead to abnormal cell behavior, and lncRNAs are related to the occurrence and development of many human diseases. For example,

OPEN ACCESS

Edited by:

Jijun Tang,
University of South Carolina,
United States

Reviewed by:

Cheng Liang,
Shandong Normal University, China
Nana Guan,
Guizhou University of Finance and
Economics, China

*Correspondence:

Min Chen
chenmin@hnit.edu.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Statistical Genetics and Methodology,
a section of the journal
Frontiers in Genetics

Received: 07 December 2021

Accepted: 18 January 2022

Published: 04 February 2022

Citation:

Chen M, Deng Y, Li A and Tan Y (2022)
Inferring Latent Disease-IncRNA
Associations by Label-Propagation
Algorithm and Random Projection on a
Heterogeneous Network.
Front. Genet. 13:798632.
doi: 10.3389/fgene.2022.798632

Wang et al. [5] found that lncRNA PVT1 promotes the progression of melanoma through endogenous sponge cell miR-26b, and Cai et al. (2018) found that BCAR4 can activate the GLI2 signaling pathway in prostate cancer. The specific secondary structure of lncRNAs and its ability to control gene expression also render it an ideal target for drug development (Chen ZJ. et al., 2016; Tripathi et al., 2018; Xu et al., 2019). Our current understanding of the role of lncRNAs in disease is far from complete, so further understanding the relationship between lncRNAs and diseases is significant. However, experimentally identifying the association between lncRNAs and diseases through biotechnology is expensive and laborious. Increased attention is being paid to predicting the association between lncRNAs and diseases by computational prediction method.

Many researches predicted the associations between lncRNAs and diseases based on known information about lncRNA–disease associations, disease–disease similarity information, and lncRNA–lncRNA similarity information. Based on the hypothesis that similar diseases may be related to lncRNAs with similar functions, many researches used information such as lncRNA–disease association network, disease-similarity network, and lncRNA similarity network to realize the association prediction between lncRNAs and diseases through random-walk algorithm. For example, Sun et al. (2014) constructed a random-walk model RWRlncD based on global network, but this method cannot be used to predict isolated diseases (diseases without any lncRNA associated with it). Chen X. et al. (2016) proposed an improved prediction model with restart random-walk algorithm (RWR), IRWRLDA. Yu et al. constructed a prediction model based on double random walk (Yu et al., 2017). Li et al. (2019c) developed an improved local random-walk prediction model, LRWHLDA. Fan et al. (2019) combined positive-point mutual information with multiple heterogeneous information and then implemented RWR to construct an lncRNA–disease correlation prediction model IDHI–MIRW, Li et al. (2019b) constructed an lncRNA–disease-associated prediction model, TCSRWRLD, by using node information called as target convergence set combined with random-walk algorithm, but the prediction accuracy of these methods is not very high.

Chen (2015a) applied KATZ index to lncRNA–disease association prediction, and this model can infer potential lncRNAs without known related diseases. Ping et al. (2018) used known lncRNA–disease associations to construct a binary network and then predicted the lncRNA–disease association based on its strict power-law distribution. According to the path length in the lncRNA–disease heterogeneous network, Xiao et al. (2018) predicted the probability of lncRNA–disease association. Liu et al. (2019) constructed a weighted network based on the resource-allocation strategy of unequal allocation and unbiased consistency and then applied the label-propagation algorithm to predict the lncRNA–disease association. However, the prediction results of these methods may be biased toward lncRNAs with more known related diseases and diseases with more known related lncRNAs.

With the rapid development of machine-learning technology, many researches used machine-learning methods to predict potential lncRNA–disease associations and miRNA–disease-associated prediction (Liang et al., 2019). For example, Yu et al. (2018); Yu et al. (2019) proposed two prediction models based on the Naïve Bayes classifier to infer potential lncRNA–disease associations. Guo et al. (2019b) used autoencoder neural network and Rotating Forest to predict the associations between lncRNAs and diseases. Liang et al. (2021) identified cancer subtype by using graph autoencoders. Chen et al. (2018); Chen X. et al. (2019) predicted miRNA–disease association by using the decision-tree model. Zhao et al. (2019) predicted miRNA–disease association by using adaptive boosting. Chen et al. (2017a) predicted miRNA–disease association by using support vector machine combined with k-nearest neighbor method. In this type of machine-learning prediction model, the main disadvantage is that negative samples are required as a training set. Given that negative samples are usually difficult to obtain, their prediction performance is significantly affected. Many semi-supervised methods are attracting attention. Xuan Z. et al. (2019). developed a probabilistic matrix-factorization model based on semi-supervised learning methods to identify potential associations between lncRNAs and diseases. Laplacian regularized least squares obtained wide application in the area of bioinformatics (Shen et al., 2021). By fusing the semantic similarity and cosine similarity of disease, lncRNA expression similarity, and cosine similarity. Lan et al. (2020) denoised lncRNA feature information and disease feature information with an automatic encoder. They then predicted lncRNA–disease association by using matrix-decomposition algorithm. Xie et al. (2019) predicted the association between lncRNAs and diseases by Laplacian regularized least squares. Chen et al. (Chen and Yan, 2013) developed a model LRLSLDA that uses Laplacian regularized least squares to identify the associations between lncRNAs and diseases. Later, on the basis of LRLSLDA, Chen et al. (2015) proposed a new lncRNA–disease association prediction model, LRLSLDA–LNCSSIM. Huang et al. (2016) used the topological feature of a directed acyclic graph of disease-similarity network to propose another improved model ILNCSIM. None of these semi-supervised methods require negative samples to train the model, but the problem of how to select parameters more reasonably has not been resolved.

In recent years, deep learning has attracted increased attention from artificial intelligence communities (Lihong et al., 2021; Zhou et al., 2021a; Lihong et al. (2021) and Zhou et al. (2021b) developed two deep learning-based models, deep Learning framework with Dual-net Neural Architecture and multiple-layer deep model based on gradient boosting decision trees, to predict possible lncRNA–protein interactions. Xuan et al. proposed a series of lncRNA–disease association prediction models based on convolutional neural networks, including CNNLDA (Xuan et al., 2019a), GCNLDA (Xuan et al., 2019c), CNNDLP (Xuan et al., 2019d), and LDAPred (Xuan et al., 2019b). Wu et al. (2020) also predicted the potential association between lncRNA and disease by using a graph-convolutional network. Lan et al. (2021) denoised heterogeneous data through principal component analysis. They then extracted features by graph-attention network, ultimately

predicting the potential association between lncRNA and disease by using multilayer perceptron. These methods have good performance in lncRNA–disease association prediction, but the parameters of these models are relatively difficult to determine.

Various biological information from different sources can help us understand the relationships between lncRNAs and diseases more comprehensively (Chen et al., 2017b; Fu et al., 2018) (Peng et al., 2017). For example, Liu et al. (2014) integrated the human lncRNA expression profile, human gene-expression profile, and other data to predict lncRNA–disease association. This method can achieve lncRNA–disease association prediction without knowing lncRNA–disease association. Chen Q. et al. (2019) used support-vector machine (SVM) to implement lncRNA–disease association prediction by integrating lncRNA–gene interaction, lncRNA–disease association, and disease semantic similarity. Lu et al. (2018) integrated known lncRNA–disease interactions, disease–gene interactions, and gene–gene interactions and used the inductive matrix-completion method to identify the associations between lncRNAs and diseases. Ding et al. (2018) combined gene–disease and lncRNA–disease association information and established a lncRNA–disease association prediction model, TPGLD, based on a lncRNA–disease–gene tripartite network. Wang Y. et al. (2019) pre-set weights for various association matrices between genes, lncRNAs, and diseases, decomposed these matrices into low-rank matrices, and developed a weighted-matrix decomposition lncRNA–disease association prediction model WMFLDA. Chen (2015b) predicted lncRNA–disease association through the integration of lncRNA–miRNA interaction and miRNA–disease correlation. Zhang et al. (2019) developed a prediction model based on DeepWalk through the integration of miRNA–disease, lncRNA–disease, and miRNA–lncRNA correlation. Zhou et al. (2015) realized the random-walk algorithm on the heterogeneous network composed of the known lncRNA–disease-related network, miRNA-related lncRNA crosstalk network, and disease-similarity network and proposed a prediction model, RWRHLD. Wang et al. (2016) used the known lncRNA–miRNA crosstalk to develop a sequence-based lncRNA–disease association prediction model, LncDisease. However, owing to the high false negatives and positives in the prediction of miRNA–lncRNA interaction, the performance of LncDisease is limited.

Zhao et al. (2015) integrated genome, transcriptome, and rule set data and then used the naïve Bayesian classifier to predict the lncRNA–cancer association. Lan et al. (2016) integrated information such as lncRNA sequence information, disease–gene associations, and GO annotations and identified new lncRNA–disease associations through bagging SVM. Fu et al. (2018) used different biological data sources of lncRNAs, miRNAs, genes, disease, and drugs for prediction, decomposed the correlation matrix into different biological entities, and reconstructed the lncRNA–disease correlation matrix through matrix decomposition. However, the method does not deal with the noise of the original features, so the prediction performance is not high. Sumathipala et al. (2019) integrated integrin disease, protein–lncRNA, and protein–protein correlation and used the

network-diffusion method to predict lncRNA–disease association. Zhang et al. (2018) used lncRNA similarity, protein–protein interactions, and disease similarity to construct a composite network and then used flow-propagation algorithm for prediction. Guo et al. (2019a) constructed a molecular-association network based on the known association among diseases, proteins, miRNA, lncRNA, and drugs and then used random-forest classifier to infer the association between any two of them. The above studies can help elucidate cellular processes and complex pathogenesis at the molecular level to a certain extent, but the use of multiple biological data sources may introduce noise and irrelevant information, leading to increased false-positive rates.

In the present study, we proposed an lncRNA–disease association prediction method called LPARP, which is based on a label-propagation algorithm and random projection. LPARP uses the semantic similarity of diseases, functional similarity of lncRNAs, and known information on lncRNA–disease association and then predicts them through label-propagation algorithms and random projections. Experimental results showed that LPARP is superior to several existing classic methods in predicting candidate lncRNAs. Case studies on bladder cancer, esophageal squamous-cell carcinoma, and colorectal cancer show that LPARP can effectively identify potential diseases associated with lncRNAs.

MATERIALS AND METHODS

Materials

LncRNA–Disease Association Network

Known experiments supporting lncRNA–disease-related data are from the lncRNADisease database (Chen et al., 2012). We obtain three datasets of lncRNA–disease-related data from different versions of the database. From the 2014 version, 352 pairs of lncRNA–disease-related data are obtained, covering 156 lncRNAs and 190 diseases (dataset1); from the 2015 version, 621 pairs of associations are obtained, covering 285 lncRNAs and 226 diseases (dataset2); from the 2017 version, 1,695 pairs of associations are obtained, including 828 lncRNAs and 314 diseases (dataset3). For convenience, a Boolean matrix $LD = (ld_{ij})_{n_l \times n_d}$ is used to represent the association between lncRNAs and diseases. If a known association exist between lncRNA l_i and disease d_j , then $LD(i, j) = 1$. Otherwise, $LD(i, j) = 0$. n_l and n_d are used to represent the number of diseases and lncRNA, respectively.

Disease Semantic Similarity

Many researchers used disease semantic similarity data to describe the similarity between diseases. In this method, the disease is represented as a directed acyclic graph (DAG), and then the similarity between the diseases is calculated based on the DAG. The detailed calculation process can be found in literature (Wang et al., 2010). This method is used to calculate the semantic similarity between diseases, as represented by matrix DD .

LncRNA Functional Similarity

Considering that lncRNAs with similar functions are often associated with similar diseases, we calculate the functional similarity between diseases based on the semantic similarity of diseases. This type of method is used in many lncRNA–disease associations (Chen et al., 2020; Zhang et al., 2020; Zhang et al., 2021). It will not be introduced in detail here. The matrix LL is used to represent the functional similarity of lncRNA.

Disease (LncRNA) Gaussian Interaction-Profile Kernel Similarity

Many zeros exist in the disease semantic similarity matrix DD and the lncRNA functional similarity matrix LL , so we further introduce the Gaussian interaction-profile kernel similarity (van Laarhoven et al., 2011) to improve this shortcoming. The Gaussian interaction-profile kernel similarity is also based on the assumption that lncRNAs with similar functions are often associated with diseases with similar phenotypes. The Gaussian interaction-profile kernel similarity between lncRNAs is defined as follows:

$$GL(i, j) = \exp\left(-\gamma_l \|lp(l_i) - lp(l_j)\|^2\right) \quad (1)$$

$lp(l_i)$ indicates the number of diseases associated with lncRNA l_i , and γ_l is the width of the nuclear spectrum, defined as follows:

$$\gamma_l = \frac{1}{\frac{1}{nl} \sum_{i=1}^{nl} \|lp(l_i)\|^2} \quad (2)$$

Similarly, we can obtain the similarity of Gaussian nuclear spectrum between diseases:

$$GD(i, j) = \exp\left(-\gamma_d \|lp(d_i) - lp(d_j)\|^2\right) \quad (3)$$

$lp(d_i)$ is the number of lncRNAs associated with disease d_i , and γ_d is the width of the nuclear spectrum, defined as follows:

$$\gamma_d = \frac{1}{\frac{1}{nd} \sum_{i=1}^{nd} \|lp(d_i)\|^2} \quad (4)$$

Integrated Disease Similarity and LncRNA Similarity

Next, lncRNA functional similarity and lncRNA Gaussian interaction-profile kernel similarity are integrated to construct lncRNA similarity.

If the functional similarity between lncRNA node l_i and lncRNA node l_j is 0, the similarity between l_i and l_j is taken as the lncRNA Gaussian interaction-profile kernel similarity value between l_i and l_j . Otherwise, the value is the functional similarity LL between l_i and l_j , and the formula is as follows:

$$LL^f(i, j) = \begin{cases} LL(i, j), & \text{if } LL(i, j) \neq 0 \\ GL(i, j), & \text{otherwise} \end{cases} \quad (5)$$

In the same way, the semantic similarity between diseases and the Gaussian interaction-profile kernel similarity between diseases are used to construct the similarity between diseases.

$$DD^f(i, j) = \begin{cases} DD(i, j), & \text{if } DD(i, j) \neq 0 \\ GD(i, j), & \text{otherwise} \end{cases} \quad (6)$$

LDAI-ISPS Workflow Model

The algorithm is divided into three steps. In step 1, Integrated disease similarity is constructed by using semantic similarity between diseases and the Gaussian interaction-profile kernel similarity between diseases, and integrated lncRNA similarity is constructed by using functional similarity between lncRNAs and Gaussian interaction profile kernel similarity between lncRNAs. In step 2, the label-propagation algorithm is used to obtain the estimated score of lncRNA–disease association. In step 3, random projections are used to obtain precise scores of lncRNA–disease associations. (Figure 1.).

Estimated Score of LncRNA–Disease Association

First, the label-propagation algorithm in the lncRNA network is implemented. During the label-propagation process, each point retains the information of its neighbors and receives its initial label information. The iterative equation can be written as follows (Wang and Zhang, 2008):

$$F_L(t+1) = \alpha * LL^* * F_L(t) + (1-\alpha)LD \quad (7)$$

In the above formula, t represents the time step, $F_L(t)$ represents the result of the t th iteration in the label-propagation algorithm, and LD is the known lncRNA–disease association matrix, which represents the initial matrix. $\alpha \in [0,1]$ is a hyper-parameter used to balance the ratio between the information from its neighbors and its initial label information. LL^* is the normalized matrix of the integrated lncRNA similarity network LL^f , whose calculation method is as follows:

$$LL^*(i, j) = LL^f(i, j) / \left(\sum_{i=1}^{nl} LL^f_{i,j}(i, j) + \sum_{j=1}^{nl} LL^f_{i,j}(i, j) \right) \quad (8)$$

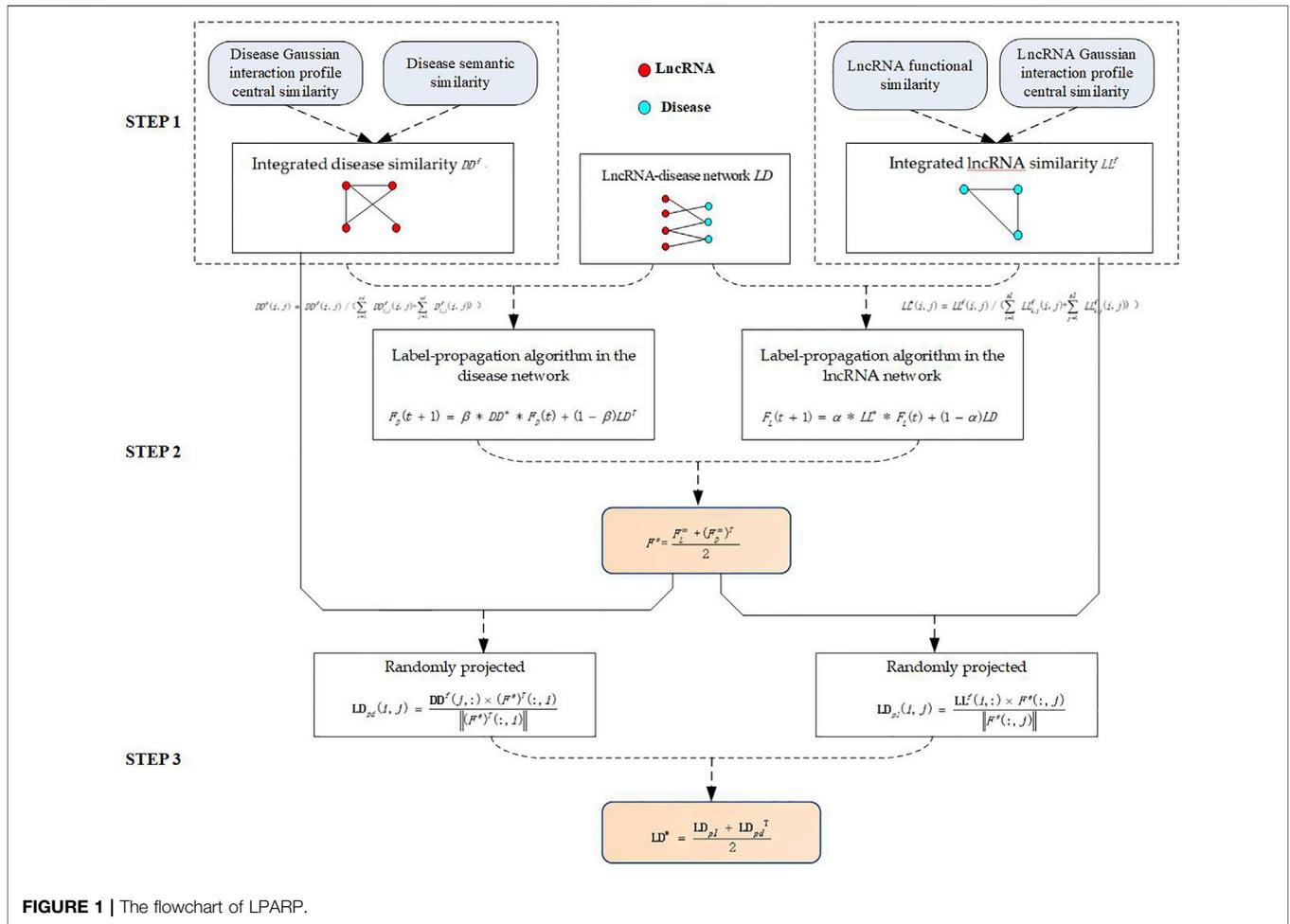
After finite iterations, the probability space reaches a stable state F_L^∞ ($|F_L(t+1) - F_L(t)| < 10^{-6}$) to stop the iteration.

Then, the iterative equation of the label-propagation algorithm in the disease is implemented as follows:

$$F_D(t+1) = \beta * DD^* * F_D(t) + (1-\beta)LD^T \quad (9)$$

LD^T is the transposed matrix of LD . Let $F_D(0) = LD^T$. $\beta \in [0,1]$ is a hyper-parameter used to control the rate of retaining information from neighbors and DD^* is the normalized matrix of the integrated disease-similarity network DD^f , then the calculation method is as follows:

$$DD^*(i, j) = DD^f(i, j) / \left(\sum_{i=1}^{nd} DD^f_{i,j}(i, j) + \sum_{j=1}^{nd} DD^f_{i,j}(i, j) \right) \quad (10)$$



The prediction result of the label-propagation algorithm used in the disease network is represented by matrix F_D^∞ (the iteration was stopped when $|F_D(t+1) - F_D(t)| < 10^{-6}$).

Finally, the median values of the prediction results F_L^∞ and F_D^∞ of the implementation of the label-propagation algorithm in the two networks are used as the estimated scores of the lncRNA–disease association:

$$F^e = \frac{F_L^\infty + (F_D^\infty)^T}{2} \tag{11}$$

Accurate Score of lncRNA–Disease Association

First, the integrated lncRNA similarity matrix LL^f is randomly projected in the lncRNA–disease association prediction score matrix F^e :

$$LD_{pl}(i, j) = \frac{LL^f(i, :) \times F^e(:, j)}{\|F^e(:, j)\|} \tag{12}$$

In the above formula, $\|F^e(:, j)\|$ is the 2-norm of $F^e(:, j)$.

Then, the integrated disease similarity matrix DD^f is randomly projected into the lncRNA–disease association estimation-score transposition matrix $(F^e)^T$:

$$LD_{pd}(i, j) = \frac{DD^f(j, :) \times (F^e)^T (:, i)}{\|(F^e)^T (:, i)\|} \tag{13}$$

Finally, LD_{pl} and LD_{pd} are synthesized to obtain the final prediction score.

$$LD^* = \frac{LD_{pl} + LD_{pd}^T}{2} \tag{14}$$

RESULTS

Parameter Selection Method

In the process of label propagation, each node retains the information of its neighbors and receives its initial label information. In **formula 7** (**formula 9**), the parameter $\alpha \in [0,1]$ ($\beta \in [0,1]$) is used to control the rate of retaining information from neighbors, and $1-\alpha(1-\beta)$ means the probability of receiving its initial tag information. For simplicity, we set parameter α and β to the same size. When the parameter value changes from 0 to 1, leave-one-out cross-validation (LOOCV) is implemented on the three data sets to identify the optimum parameters. In LOOCV, we use a known lncRNA–disease association as a test sample and the remaining lncRNA–disease

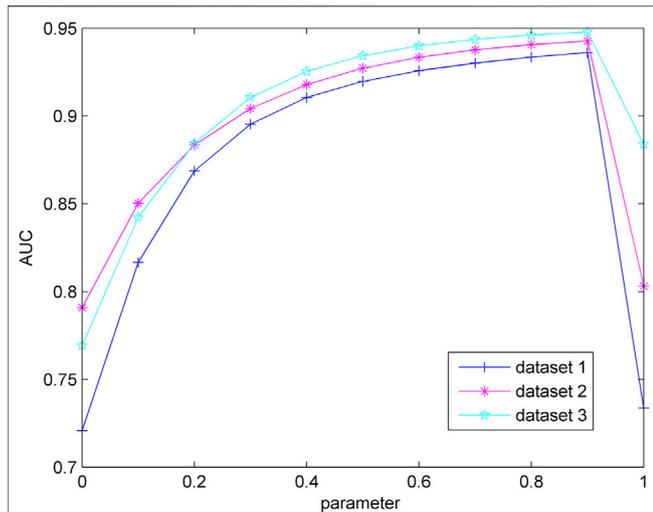


FIGURE 2 | Influence of parameter variation on model prediction accuracy.

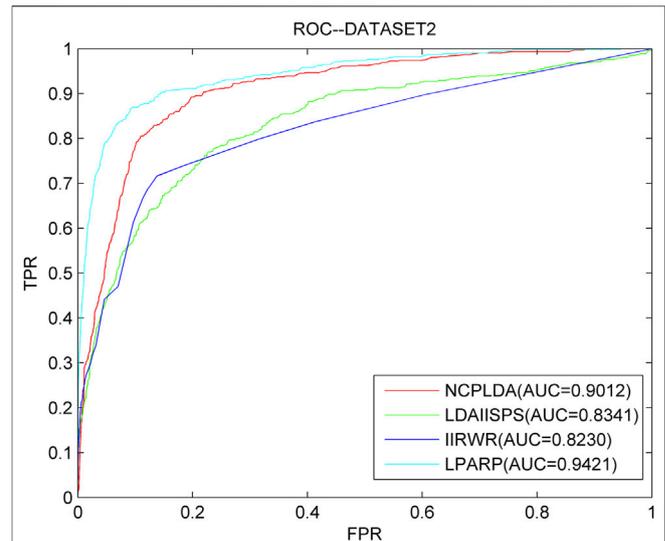


FIGURE 4 | The ROC curves and AUC values of LPARP compared with other methods on the dataset 2.

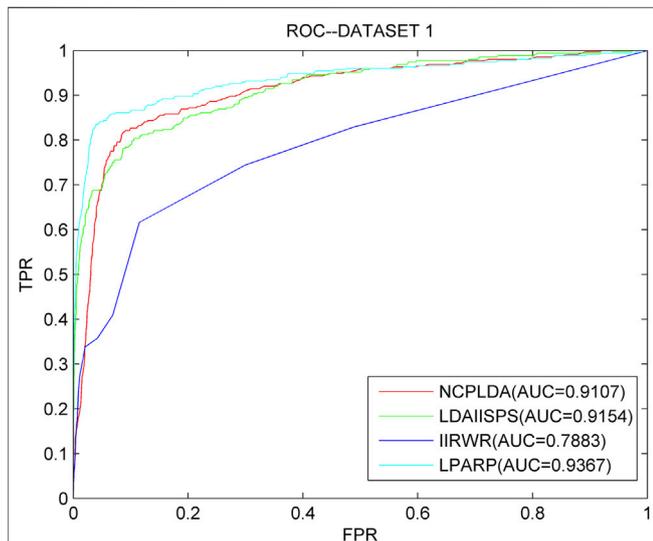


FIGURE 3 | The ROC curves and AUC values of LPARP compared with other methods on the dataset 1.

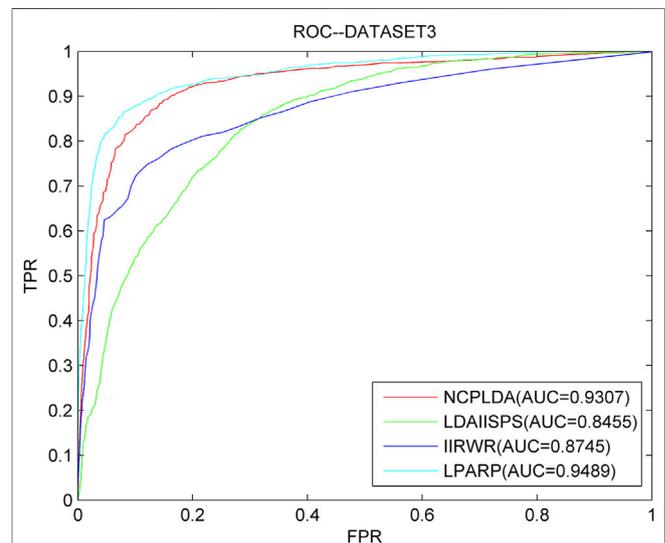


FIGURE 5 | The ROC curves and AUC values of LPARP compared with other methods on the dataset 3.

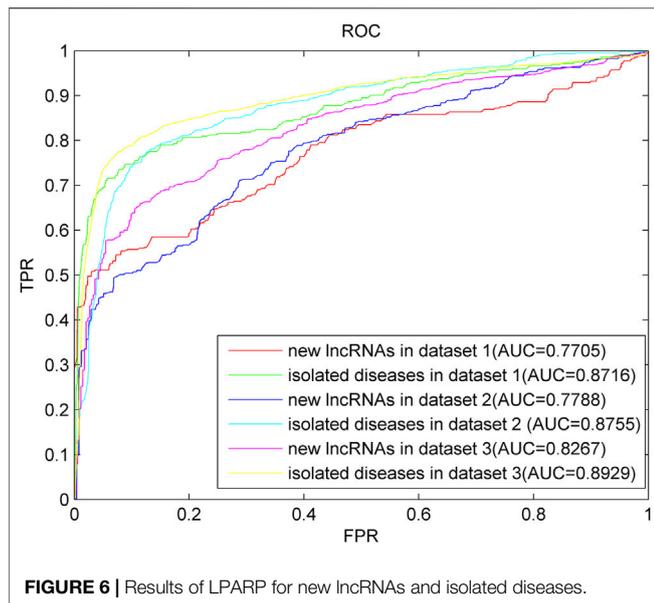
associations as a training sample each time. After the model is trained, the true positive rate (TPR) and false-positive rate (FPR) are calculated to draw the receiver operating characteristic (ROC) curve according to the TPR and FPR under different thresholds. The area under the ROC curve (AUC) is used to evaluate the performance of the model, and a larger AUC value means better prediction performance. The ROC curve and AUC value of each parameter are listed in **Figure 2**. In the three different data sets, LPARP has the largest AUC value when the parameter is 0.9. Therefore, we set the parameters to 0.9 on the three different data sets.

Comparison With Other Methods

As we know, NCPLDA (Li G. et al., 2019), IIRWR (Wang L. et al., 2019), and LDAI-ISPS (Zhang et al., 2020) are excellent calculation

methods currently used to predict the association of lncRNA diseases. The data used by these three methods is the same as ours. Here, we compare LPARP with them. The comparison results of implementing LOOCV on the three datasets are shown in **Figures 3–5**.

The AUC values of NCPLDA, IIRWR, LDAI-ISPS, and LPARP on dataset 1 are 0.9107, 0.7883, 0.9154, and 0.9367, respectively, and the AUC values on dataset 2 are 0.9383, 0.9012, 0.8230, 0.8341, and 0.9421, respectively. The AUC values on dataset 3 are 0.9307, 0.8745, 0.8455, and 0.9489, respectively. Obviously, on three different datasets, the prediction performance of LPARP is significantly better than those of NCPLDA, IIRWR, and LDAI-ISPS.



Prediction for New lncRNAs and Isolated Diseases

With the continuous improvement in lncRNA-recognition technology, more lncRNAs are being unearthed continuously, and most of them have unknown relationships with diseases. We call them new lncRNAs. Isolated diseases refer to diseases without any known relation with lncRNAs. The association prediction of new lncRNAs and isolated diseases helps scientists understand the molecular mechanism of diseases and can help diagnose and treat diseases.

To simulate new lncRNAs, when a certain lncRNA is used as the test sample, all associations between the lncRNA and the diseases are removed. In the experiment, we select each lncRNA as the test sample and all associated information with other lncRNA as the training sample until all lncRNAs are tested as the prediction sample. A similar method is used to verify the prediction effect of LPARP on isolated diseases. For the prediction of new lncRNAs, the AUC on data sets 1, 2, and 3 are 0.7705, 0.7788, and 0.8267, respectively. For the prediction of isolated diseases, the AUC on data sets 1, 2, and 3 are 0.8716, 0.8755, and 0.8929, and the curves are shown in **Figure 6**. These results indicate that LPARP has a good predictive effect.

CASE STUDY

To further evaluate the actual effect of LPARP, the three human diseases including bladder cancer, esophageal squamous-cell carcinoma, and colorectal cancer are selected for the case analysis. The association of dataset 2 is extracted from the lncRNADisease database established in 2015. This database was selected for training, later it was verified in the 2017 lncRNADisease database, which is dataset 3, and the latest related literature.

First, all experimentally verified associations are taken as training samples, and the lncRNA–disease associations that have not been experimentally verified are were taken as

candidate associations. For a specific disease, the candidate lncRNAs are sorted according to their prediction scores. For the three diseases bladder cancer, esophageal squamous-cell carcinoma, and colorectal cancer, the top five associations of lncRNA are predicted, as shown in **Table 1**.

Bladder cancer is the ninth most common cancer in the world, and more than 60% of all bladder cancer cases occur in less developed areas of the world (Antoni et al., 2017). **Table 1** shows that three of the first five predicted lncRNAs have found supporting evidence in the 2017 version of the lncRNADisease database. MEG3 and PVT1 have not been verified by the lncRNADisease database, but we have manually excavated recent biomedical literature and find them and bladder cancer-related evidence. For example, Fan et al. (2020) found that MEG3 can control the progression of bladder cancer through PI3K/AKT/mTOR pathway regulation. Tian et al. (2019) found that PVT1 can regulate the growth, migration, and invasion of bladder cancer through mir31/CDK1.

Esophageal squamous-cell carcinoma accounts for about 90% of 456,000 cases of esophageal cancer each year (Abnet et al., 2018). The predicted top five lncRNAs are MALAT1, MEG3, BCYRN1, UCA1, and LSINCT5, among which MALAT1 and MEG3 are found to be associated with esophageal squamous-cell carcinoma in lncRNADisease in 2017. Through literature search, UCA1 and LSICT5 are found to be related to esophageal squamous-cell carcinoma. Although we have not manually excavated recent literature to prove that BCYRN1 is related to esophageal squamous-cell carcinoma, we believe that scientists will find the evidence that BCYRN1 is associated with esophageal squamous-cell carcinoma in the future.

Colorectal cancer is the third most common cancer among men and the second most cancer among women (Favoriti et al., 2016). Among the predicted five lncRNAs, three are verified by lncRNADisease database, but MINA and EPB41L4A-AS1 do not show any association with colorectal cancer in the lncRNADisease database. However, Bin et al. (Bin et al., 2020) found in 2020 that EPB41L4a AS1 acts as an oncogene by regulating the Rho/ROCK pathway of colorectal cancer. All of the above literatures were published after the 2017 edition of the lncRNADisease was updated, which confirms the reliability of our method.

To further verify the predictive effect of LPARP on isolated diseases, we select bladder cancer, esophageal squamous-cell carcinoma, and colorectal cancer in dataset2 for case study. The difference between them is that for any kind of disease prediction, to simulate an isolated disease, when training the model, all associations of the disease are removed. The prediction results of the three diseases are shown in **Table 2**. For Esophageal squamous-cell carcinoma and colorectal cancer, the top five predicted lncRNAs have supporting evidence in the latest lncRNADisease database. For bladder cancer, three lncRNAs have supporting evidence, and MEG3 and PVT1 have not been verified by the lncRNADisease database. When conducting case analysis of common diseases, these two lncRNAs are also considered to be closely related to bladder cancer. Recently, many scientists have proven that they are related to bladder cancer.

TABLE 1 | The top 5 lncRNA candidates predicted for bladder cancer, esophageal squamous-cell carcinoma, and colorectal cancer.

disease	lncRNA name	Evidences	Rank
bladder cancer	HOTAIR	LncRNADisease	1
bladder cancer	MALAT1	LncRNADisease	2
bladder cancer	MEG3	Fan et al. (2020)	3
bladder cancer	PVT1	Tian et al. (2019)	4
bladder cancer	GAS5	LncRNADisease	5
Esophageal squamous cell carcinoma	MALAT1	LncRNADisease	1
Esophageal squamous cell carcinoma	MEG3	LncRNADisease	2
Esophageal squamous cell carcinoma	BCYRN1	Unconfirmed	3
Esophageal squamous cell carcinoma	UCA1	Kang et al. (2018)	4
Esophageal squamous cell carcinoma	LSINCT5	Jing et al. (2019)	5
colorectal cancer	MEG3	LncRNADisease	1
colorectal cancer	H19	LncRNADisease	2
colorectal cancer	MINA	Unconfirmed	3
colorectal cancer	UCA1	LncRNADisease	4
colorectal cancer	EPB41L4A-AS1	Bin et al. (2020)	5

TABLE 2 | The top 5 novel disease-correlated lncRNA candidates predicted for bladder cancer, esophageal squamous-cell carcinoma, and colorectal cancer.

disease	lncRNA name	Evidences	RANK
bladder cancer	HOTAIR	LncRNADisease	1
bladder cancer	MALAT1	LncRNADisease	2
bladder cancer	H19	LncRNADisease	3
bladder cancer	MEG3	Fan et al. (2020)	4
bladder cancer	PVT1	Tian et al. (2019)	5
Esophageal squamous cell carcinoma	HOTAIR	LncRNADisease	1
Esophageal squamous cell carcinoma	MALAT1	LncRNADisease	2
Esophageal squamous cell carcinoma	H19	LncRNADisease	3
Esophageal squamous cell carcinoma	MEG3	LncRNADisease	4
Esophageal squamous cell carcinoma	PVT1	LncRNADisease	5
colorectal cancer	HOTAIR	LncRNADisease	1
colorectal cancer	MALAT1	LncRNADisease	2
colorectal cancer	H19	LncRNADisease	3
colorectal cancer	MEG3	LncRNADisease	4
colorectal cancer	PVT1	LncRNADisease	5

DISCUSSION

This study shows how to combine lncRNA similarity, disease similarity, and known lncRNA–disease interactions to predict new lncRNA–drug interactions. A new integration method of label-propagation algorithm and random-projection algorithm (LAPRP) is proposed. After evaluating three different datasets, we find that compared with other state-of-the-art methods, LAPRP improves performance effectively and can predict isolated diseases and new lncRNAs. Two types of case studies are carried out on three human diseases: bladder cancer, esophageal squamous-cell carcinoma, and colorectal cancer. The first category is general disease prediction. Among the predicted top five lncRNAs, all five lncRNAs related to bladder cancer, four related to esophageal squamous-cell carcinoma, and four related to colorectal cancer have verified to be the latest confirmation of database or latest literature. The second category is the prediction of isolated diseases. The top five lncRNAs predicted to be related to the three diseases have been confirmed by the latest database or the latest literature.

Comparative experiments and case studies show that LAPRP has high prediction accuracy and does not require negative samples. It can be used to predict isolated diseases and new lncRNAs. LAPRP is a useful supplement to experimental methods.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

Conceptualization, YD and MC; Data curation, AL; Formal analysis, AL; Funding acquisition, YD, MC, and AL; Methodology, YD and MC; Software, YT; Validation, YT and YD; Writing—original draft, YD and MC; Writing—review and editing, YD and MC.

REFERENCES

- Abnet, C. C., Arnold, M., and Wei, W.-Q. (2018). Epidemiology of Esophageal Squamous Cell Carcinoma. *Gastroenterology* 154, 360–373. doi:10.1053/j.gastro.2017.08.023
- Antoni, S., Ferlay, J., Soerjomataram, I., Znaor, A., Jemal, A., and Bray, F. (2017). Bladder Cancer Incidence and Mortality: A Global Overview and Recent Trends. *Eur. Urol.* 71, 96–108. doi:10.1016/j.eururo.2016.06.010
- Bin, J., Nie, S., Tang, Z., Kang, A., Fu, Z., Hu, Y., et al. (2020). Long Noncoding RNA EPB41L4A-AS1 Functions as an Oncogene by Regulating the Rho/ROCK Pathway in Colorectal Cancer. *J. Cel Physiol* 236, 523–535. doi:10.1002/jcp.29880
- Cai, Z., Wu, Y., Li, Y., Ren, J., and Wang, L. (2018). BCAR4 Activates GLI2 Signaling in Prostate Cancer to Contribute to Castration Resistance. *Aging* 10, 3702–3712. doi:10.18632/aging.101664
- Chen, G., Wang, Z., Wang, D., Qiu, C., Liu, M., Chen, X., et al. (2012). LncRNADisease: a Database for Long-Non-Coding RNA-Associated Diseases. *Nucleic Acids Res.* 41, D983–D986. doi:10.1093/nar/gks1099
- Chen, M., Peng, Y., Li, A., Deng, Y., and Li, Z. (2020). A Novel LncRNA-Disease Association Prediction Model Using Laplacian Regularized Least Squares and Space Projection-Federated Method. *IEEE Access*, 1. doi:10.1109/access.2020.3002588
- Chen, Q., Lai, D., Lan, W., Wu, X., Chen, B., Chen, Y. P., et al. (2019a). ILDMSE: Inferring Associations between Long Non-coding RNA and Disease Based on Multi-Similarity Fusion. *Ieeelacm Trans. Comput. Biol. Bioinform.* 1. doi:10.1109/TCBB.2019.2936476
- Chen, X., Yan, C. C., Zhang, X., and You, Z. H. (2017b). Long Non-coding RNAs and Complex Diseases: from Experimental Results to Computational Models. *Brief Bioinform* 18, 558–576. doi:10.1093/bib/bbw060
- Chen, X., Clarence Yan, C., Luo, C., Ji, W., Zhang, Y., and Dai, Q. (2015). Constructing LncRNA Functional Similarity Network Based on LncRNA-Disease Associations and Disease Semantic Similarity. *Sci. Rep.* 5, 11338. doi:10.1038/srep11338
- Chen, X., Huang, L., Xie, D., and Zhao, Q. (2018). EGBMMDA: Extreme Gradient Boosting Machine for MiRNA-Disease Association Prediction. *Cell Death Dis* 9, 3. doi:10.1038/s41419-017-0003-x
- Chen, X. (2015a). KATZLDA: KATZ Measure for the LncRNA-Disease Association Prediction. *Sci. Rep.* 5, 16840. doi:10.1038/srep16840
- Chen, X. (2015b). Predicting LncRNA-Disease Associations and Constructing LncRNA Functional Similarity Network Based on the Information of miRNA. *Sci. Rep.* 5, 13186. doi:10.1038/srep13186
- Chen, X., Wu, Q.-F., and Yan, G.-Y. (2017a). RKNMMDA: Ranking-Based KNN for MiRNA-Disease Association Prediction. *RNA Biol.* 14, 952–962. doi:10.1080/15476286.2017.1312226
- Chen, X., and Yan, G.-Y. (2013). Novel Human LncRNA-Disease Association Inference Based on LncRNA Expression Profiles. *Bioinformatics* 29, 2617–2624. doi:10.1093/bioinformatics/btt426
- Chen, X., You, Z.-H., Yan, G.-Y., and Gong, D.-W. (2016a). IRWRLDA: Improved Random Walk with Restart for LncRNA-Disease Association Prediction. *Oncotarget* 7, 57919–57931. doi:10.18632/oncotarget.11141
- Chen, X., Zhu, C.-C., and Yin, J. (2019b). Ensemble of Decision Tree Reveals Potential miRNA-Disease Associations. *Plos Comput. Biol.* 15, e1007209. doi:10.1371/journal.pcbi.1007209
- Chen, Z. J., Zhang, Z., Xie, B. B., and Zhang, H. Y. (2016b). Clinical Significance of Up-Regulated LncRNA NEAT1 in Prognosis of Ovarian Cancer. *Eur. Rev. Med. Pharmacol. Sci.* 20, 3373–3377.
- Ding, L., Wang, M., Sun, D., and Li, A. (2018). TPGLDA: Novel Prediction of Associations between lncRNAs and Diseases via LncRNA-Disease-Genes Tripartite Graph. *Sci. Rep.* 8, 1065. doi:10.1038/s41598-018-19357-3
- Fan, X.-N., Zhang, S.-W., Zhang, S.-Y., Zhu, K., and Lu, S. (2019). Prediction of LncRNA-Disease Associations by Integrating Diverse Heterogeneous Information Sources with RWR Algorithm and Positive Pointwise Mutual Information. *BMC Bioinformatics* 20, 87. doi:10.1186/s12859-019-2675-y
- Fan, X., Huang, H., Ji, Z., and Mao, Q. (2020). Long Non-coding RNA MEG3 Functions as a Competing Endogenous RNA of miR-93 to Regulate Bladder Cancer Progression via PI3K/AKT/mTOR Pathway. *Transl Cancer Res. TCR* 9, 1678–1688. doi:10.21037/tcr.2020.01.70
- Favoriti, P., Carbone, G., Greco, M., Pirozzi, F., Pirozzi, R. E. M., and Corcione, F. (2016). Worldwide burden of Colorectal Cancer: a Review. *Updates Surg.* 68, 7–11. doi:10.1007/s13304-016-0359-y
- Fei Wang, F., and Changshui Zhang, C. (2008). Label Propagation through Linear Neighborhoods. *IEEE Trans. Knowl. Data Eng.* 20, 55–67. doi:10.1109/tkde.2007.190672
- Fu, G., Wang, J., Domeniconi, C., and Yu, G. (2018). Matrix Factorization-Based Data Fusion for the Prediction of LncRNA–Disease Associations. *Bioinformatics* 34, 1529–1537. doi:10.1093/bioinformatics/btx794
- Guo, Z.-H., Yi, H.-C., and You, Z.-H. (2019a). Construction and Comprehensive Analysis of a Molecular Association Network via LncRNA miRNA Diseases Drug Protein Graph. *Cells* 8.
- Guo, Z.-H., You, Z.-H., Wang, Y.-B., Yi, H.-C., and Chen, Z.-H. (2019b). A Learning-Based Method for LncRNA-Disease Association Identification Combining Similarity Information and Rotation Forest. *iScience* 19, 786–795. doi:10.1016/j.isci.2019.08.030
- Huang, Y.-A., Chen, X., You, Z.-H., Huang, D.-S., and Chan, K. C. C. (2016). ILNCSIM: Improved LncRNA Functional Similarity Calculation Model. *Oncotarget* 7, 25902–25914. doi:10.18632/oncotarget.8296
- Jing, L., Lin, J., Zhao, Y., Liu, G. J., Liu, Y. B., Feng, L., et al. (2019). Long Noncoding RNA LSINCT5 Is Upregulated and Promotes the Progression of Esophageal Squamous Cell Carcinoma. *Eur. Rev. Med. Pharmacol. Sci.* 23, 5195–5205. doi:10.26355/eurrev_201906_18184
- Kang, K., Huang, Y. H., Li, H. P., and Guo, S. M. (2018). Expression of UCA1 and MALAT1 Long-Chain Non-coding RNAs in Esophageal Squamous Cell Carcinoma Tissues Is Predictive of Patient Prognosis. *Arch. Med. Sci.* 14, 752–759. doi:10.5114/aoms.2018.73713
- Lan, W., Li, M., Zhao, K., Liu, J., Wu, F. X., Pan, Y., et al. (2016). LDAP: a Web Server for LncRNA-Disease Association Prediction. *Bioinformatics* 33, 458–460. doi:10.1093/bioinformatics/btw639
- Lan, W., Lai, D., Chen, Q., Wu, X., Chen, B., Liu, J., et al. (2020). LDICDL: LncRNA-Disease Association Identification Based on Collaborative Deep Learning. *IEEE/ACM transactions on computational biology and bioinformatics.*
- Lan, W., Wu, X., Chen, Q., Peng, W., Wang, J., and Chen, Y. (2021). GANLDA: Graph Attention Network for LncRNA-Disease Associations Prediction. *Neurocomputing.*
- Li, G., Luo, J., Liang, C., Xiao, Q., Ding, P., and Zhang, Y. (2019a). Prediction of LncRNA-Disease Associations Based on Network Consistency Projection. *IEEE Access* 7, 58849–58856. doi:10.1109/access.2019.2914533
- Li, J., Li, X., Feng, X., Wang, B., Zhao, B., and Wang, L. (2019b). A Novel Target Convergence Set Based Random Walk with Restart for Prediction of Potential LncRNA-Disease Associations. *BMC Bioinformatics* 20, 626. doi:10.1186/s12859-019-3216-4
- Li, J., Zhao, H., Xuan, Z., Yu, J., Feng, X., Liao, B., et al. (2019c). A Novel Approach for Potential Human LncRNA-Disease Association Prediction Based on Local Random Walk. *IEEE/ACM Trans. Comput. Biol. Bioinformatics.*
- Liang, C., Yu, S., and Luo, J. (2019). Adaptive Multi-View Multi-Label Learning for Identifying Disease-Associated Candidate miRNAs. *Plos Comput. Biol.* 15, e1006931. doi:10.1371/journal.pcbi.1006931
- Liang, C., Shang, M., and Luo, J. (2021). Cancer Subtype Identification by Consensus Guided Graph Autoencoders. *Bioinformatics.* doi:10.1093/bioinformatics/btab535
- Lihong, P., Wang, C., Tian, X., Zhou, L., and Li, K. (2021). Finding LncRNA-Protein Interactions Based on Deep Learning with Dual-Net Neural Architecture. *IEEE/ACM transactions on computational biology and bioinformatics.*
- Liu, M.-X., Chen, X., Chen, G., Cui, Q.-H., and Yan, G.-Y. (2014). A Computational Framework to Infer Human Disease-Associated Long Noncoding RNAs. *Plos one* 9, e84408. doi:10.1371/journal.pone.0084408
- Liu, Y., Feng, X., Zhao, H., Xuan, Z., and Wang, L. (2019). A Novel Network-Based Computational Model for Prediction of Potential LncRNA-Disease Association. *Ijms* 20, 1549. doi:10.3390/ijms20071549
- Lu, C., Yang, M., Luo, F., Wu, F.-X., Li, M., Pan, Y., et al. (2018). Prediction of LncRNA-Disease Associations Based on Inductive Matrix

- Completion. *Bioinformatics* 34, 3357–3364. doi:10.1093/bioinformatics/bty327
- Peng, L., Liu, F., Yang, J., Liu, X., Meng, Y., Deng, X., et al. (2019). Probing lncRNA-Protein Interactions: Data Repositories, Models, and Algorithms. *Front. Genet.* 10, 1346. doi:10.3389/fgene.2019.01346
- Peng, L., Liao, B., Zhu, W., Li, Z., and Li, K. (2017). Predicting Drug-Target Interactions with Multi-Information Fusion. *IEEE J. Biomed. Health Inform.* 21, 561–572. doi:10.1109/jbhi.2015.2513200
- Peng, L., Yuan, R., Shen, L., Gao, P., and Zhou, L. (2021). "LPI-EnEDT: An Ensemble Framework with Extra Tree and Decision Tree Classifiers for Imbalanced lncRNA-Protein Interaction Data Classification".
- Ping, P., Wang, L., Kuang, L., Ye, S., Iqbal, M. F. B., and Pei, T. (2018). A Novel Method for lncRNA-Disease Association Prediction Based on an lncRNA-Disease Association Network. *Ieee/acm Trans. Comput. Biol. Bioinform* 16, 688–693. doi:10.1109/TCBB.2018.2827373
- Shen, L., Liu, F., Huang, L., Liu, G., Zhou, L., and Peng, L. (2021). VDA-RWLRLS: An Anti-SARS-CoV-2 Drug Prioritizing Framework Combining an Unbalanced Bi-random Walk and Laplacian Regularized Least Squares. *Comput. Biol. Med.* 140, 105119. doi:10.1016/j.combiomed.2021.105119
- Song, X., Cao, G., Jing, L., Lin, S., Wang, X., Zhang, J., et al. (2014). Analysing the Relationship between lnc RNA and Protein-coding Gene and the Role of lnc RNA as ce RNA in Pulmonary Fibrosis. *J. Cel. Mol. Med.* 18, 991–1003. doi:10.1111/jcmm.12243
- Sumathipala, M., Maiorino, E., Weiss, S. T., and Sharma, A. (2019). Network Diffusion Approach to Predict lncRNA Disease Associations Using Multi-type Biological Networks: LION. *Front. Physiol.* 10, 888. doi:10.3389/fphys.2019.00888
- Sun, J., Shi, H., Wang, Z., Zhang, C., Liu, L., Wang, L., et al. (2014). Inferring Novel lncRNA-Disease Associations Based on a Random Walk Model of a lncRNA Functional Similarity Network. *Mol. Biosyst.* 10, 2074–2081. doi:10.1039/c3mb70608g
- Sun, Y.-Z., Zhang, D.-H., Ming, Z., Li, J.-Q., and Chen, X. (20172017). DLREFD: a Database Providing Associations of Long Non-coding RNAs, Environmental Factors and Phenotypes. *Database*.
- Tian, X., Shen, L., Wang, Z., Zhou, L., and Peng, L. (2021). A Novel lncRNA Protein Interaction Prediction Method Based on Deep forest with cascade forest Structure. *Scientific Rep.* 11. doi:10.1038/s41598-021-98277-1
- Tian, Z., Cao, S., Li, C., Xu, M., Wei, H., Yang, H., et al. (2019). lncRNA PVT1 Regulates Growth, Migration, and Invasion of Bladder Cancer by miR-31/CDK1. *J. Cel Physiol* 234, 4799–4811. doi:10.1002/jcp.27279
- Tripathi, M. K., Duxtater, K., Keramatnia, F., Zacheaus, C., Yallapu, M. M., Jaggi, M., et al. (2018). Role of lncRNAs in Ovarian Cancer: Defining New Biomarkers for Therapeutic Purposes. *Drug Discov. Today*. doi:10.1016/j.drudis.2018.04.010
- Van Laarhoven, T., Nabuurs, S. B., and Marchiori, E. (2011). Gaussian Interaction Profile Kernels for Predicting Drug-Target Interaction. *Bioinformatics* 27, 3036–3043. doi:10.1093/bioinformatics/btr500
- Wang, D., Wang, J., Lu, M., Song, F., and Cui, Q. (2010). Inferring the Human microRNA Functional Similarity and Functional Network Based on microRNA-Associated Diseases. *Bioinformatics* 26, 1644–1650. doi:10.1093/bioinformatics/btq241
- Wang, J., Ma, R., Ma, W., Chen, J., Yang, J., Xi, Y., et al. (2016). lncDisease: a Sequence Based Bioinformatics Tool for Predicting lncRNA-Disease Associations. *Nucleic Acids Res.* 44, e90. doi:10.1093/nar/gkw093
- Wang, K. C., and Chang, H. Y. (2011). Molecular Mechanisms of Long Noncoding RNAs. *Mol. Cel.* 43, 904–914. doi:10.1016/j.molcel.2011.08.018
- Wang, L., Xiao, Y., Li, J., Feng, X., Li, Q., and Yang, J. (2019a). IIRWR: Internal Inclined Random Walk with Restart for lncRNA-Disease Association Prediction. *IEEE Access* 7, 54034–54041. doi:10.1109/access.2019.2912945
- Wang, Y., Yu, G., Wang, J.-Y., Fu, G., Guo, M., and Domeniconi, C. (2019b). Weighted Matrix Factorization on Multi-Relational Data for lncRNA-Disease Association Prediction. *Methods*.
- Wapinski, O., and Chang, H. Y. (2011). Long Noncoding RNAs and Human Disease. *Trends Cel Biol.* 21, 354–361. doi:10.1016/j.tcb.2011.04.001
- Wu, X., Lan, W., Chen, Q., Dong, Y., Liu, J., and Peng, W. (2020). Inferring lncRNA-Disease Associations Based on Graph Autoencoder Matrix Completion. *Comput. Biol. Chem.* 87, 107282. doi:10.1016/j.compbiolchem.2020.107282
- Xiao, X., Zhu, W., Liao, B., Xu, J., Gu, C., Ji, B., et al. (2018). BPLDDA: Predicting lncRNA-Disease Associations Based on Simple Paths with Limited Lengths in a Heterogeneous Network. *Front. Genet.* 9, 411. doi:10.3389/fgene.2018.00411
- Xie, G., Meng, T., Luo, Y., and Liu, Z. (2019). SKF-LDA: Similarity Kernel Fusion for Predicting lncRNA-Disease Association. *Mol. Ther. - Nucleic Acids* 18, 45–55. doi:10.1016/j.omtn.2019.07.022
- Xu, Y., Jiang, T., Wang, C., and Wang, F. (2019). Sinomenine Hydrochloride Exerts Antitumor Outcome in Ovarian Cancer Cells by Inhibition of Long Non-coding RNA HOST2 Expression. *Artif. Cell Nanomedicine, Biotechnol.* 47, 4131–4138. doi:10.1080/21691401.2019.1687496
- Xuan, P., Cao, Y., Zhang, T., Kong, R., and Zhang, Z. (2019a). Dual Convolutional Neural Networks with Attention Mechanisms Based Method for Predicting Disease-Related lncRNA Genes. *Front. Genet.* 10, 416. doi:10.3389/fgene.2019.00416
- Xuan, P., Jia, L., Zhang, T., Sheng, N., Li, X., and Li, J. (2019b). LDAPred: A Method Based on Information Flow Propagation and a Convolutional Neural Network for the Prediction of Disease-Associated lncRNAs. *Ijms* 20, 4458. doi:10.3390/ijms20184458
- Xuan, P., Pan, S., Zhang, T., Liu, Y., and Sun, H. (2019c). Graph Convolutional Network and Convolutional Neural Network Based Method for Predicting lncRNA-Disease Associations. *Cells* 8. doi:10.3390/cells8091012
- Xuan, P., Sheng, N., Zhang, T., Liu, Y., and Guo, Y. (2019d). CNNLDP: A Method Based on Convolutional Autoencoder and Convolutional Neural Network with Adjacent Edge Attention for Predicting lncRNA-Disease Associations. *Ijms* 20, 4260. doi:10.3390/ijms20174260
- Xuan, Z., Li, J., Yu, J., Feng, X., Zhao, B., and Wang, L. (2019e). A Probabilistic Matrix Factorization Method for Identifying lncRNA-Disease Associations. *Genes* 10, 126. doi:10.3390/genes10020126
- Yu, G., Fu, G., Lu, C., Ren, Y., and Wang, J. (2017). BRWLDA: Bi-random Walks for Predicting lncRNA-Disease Associations. *Oncotarget* 8, 60429–60446. doi:10.18632/oncotarget.19588
- Yu, J., Ping, P., Wang, L., Kuang, L., Li, X., and Wu, Z. (2018). A Novel Probability Model for lncRNA-Disease Association Prediction Based on the Naive Bayesian Classifier. *Genes* 9, 345. doi:10.3390/genes9070345
- Yu, J., Xuan, Z., Feng, X., Zou, Q., and Wang, L. (2019). A Novel Collaborative Filtering Model for lncRNA-Disease Association Prediction Based on the Naive Bayesian Classifier. *BMC Bioinformatics* 20, 396. doi:10.1186/s12859-019-2985-0
- Zhang, H., Liang, Y., Peng, C., Han, S., Du, W., and Li, Y. (2019). *Predicting lncRNA-Disease Associations Using Network Topological Similarity Based on Deep Mining Heterogeneous Networks*. New York, NY: Mathematical biosciences, 108229.
- Zhang, J., Zhang, Z., Chen, Z., and Deng, L. (2018). Integrating Multiple Heterogeneous Networks for Novel lncRNA-Disease Association Inference. *IEEE/ACM Trans. Comput. Biol. Bioinform.*, 1.
- Zhang, Y., Chen, M., Li, A., Cheng, X., Jin, H., and Liu, Y. (2020). LDAI-ISPS: lncRNA-Disease Associations Inference Based on Integrated Space Projection Scores. *Ijms* 21, 1508. doi:10.3390/ijms21041508
- Zhang, Y., Chen, M., Xie, X., Shen, X., and Wang, Y. (2021). Two-Stage Inference for lncRNA-Disease Associations Based on Diverse Heterogeneous Information Sources. *IEEE Access* 9, 16103–16113. doi:10.1109/access.2021.3053030
- Zhao, T., Xu, J., Liu, L., Bai, J., Xu, C., Xiao, Y., et al. (2015). Identification of Cancer-Related lncRNAs through Integrating Genome, Regulome and Transcriptome Features. *Mol. Biosyst.* 11, 126–136. doi:10.1039/c4mb00478g
- Zhao, Y., Chen, X., and Yin, J. (2019). Adaptive Boosting-Based Computational Model for Predicting Potential miRNA-Disease Associations. *Bioinformatics* 35, 4730–4738. doi:10.1093/bioinformatics/btz297
- Zhou, L., Duan, Q., Tian, X., Xu, H., Tang, J., and Peng, L. (2021a). LPI-HyADBS: a Hybrid Framework for lncRNA-Protein Interaction Prediction Integrating

- Feature Selection and Classification. *BMC Bioinformatics* 22, 568. doi:10.1186/s12859-021-04485-x
- Zhou, L., Wang, Z., Tian, X., and Peng, L. (2021b). LPI-deepGBDT: a Multiple-Layer Deep Framework Based on Gradient Boosting Decision Trees for lncRNA Protein Interaction Identification. *BMC Bioinformatics* 22. doi:10.1186/s12859-021-04399-8
- Zhou, M., Wang, X., Li, J., Hao, D., Wang, Z., Shi, H., et al. (2015). Prioritizing Candidate Disease-Related Long Non-coding RNAs by Walking on the Heterogeneous lncRNA and Disease Network. *Mol. Biosyst.* 11, 760–769. doi:10.1039/c4mb00511b

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Chen, Deng, Li and Tan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.