

Research Article

BP Neural Network Could Help Improve Pre-miRNA Identification in Various Species

Limin Jiang,^{1,2} Jingjun Zhang,² Ping Xuan,³ and Quan Zou^{1,4}

¹*School of Computer Science and Technology, Tianjin University, Tianjin 300350, China*

²*School of Information and Electrical Engineering, Hebei University of Engineering, Handan 056038, China*

³*School of Computer Science and Technology, Heilongjiang University, Harbin 150080, China*

⁴*State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin 300074, China*

Correspondence should be addressed to Ping Xuan; 2004058@hlju.edu.cn and Quan Zou; zouquan@tju.edu.cn

Received 17 May 2016; Revised 5 July 2016; Accepted 17 July 2016

Academic Editor: Xing Chen

Copyright © 2016 Limin Jiang 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.

MicroRNAs (miRNAs) are a set of short (21–24 nt) noncoding RNAs that play significant regulatory roles in cells. In the past few years, research on miRNA-related problems has become a hot field of bioinformatics because of miRNAs' essential biological function. miRNA-related bioinformatics analysis is beneficial in several aspects, including the functions of miRNAs and other genes, the regulatory network between miRNAs and their target mRNAs, and even biological evolution. Distinguishing miRNA precursors from other hairpin-like sequences is important and is an essential procedure in detecting novel microRNAs. In this study, we employed backpropagation (BP) neural network together with 98-dimensional novel features for microRNA precursor identification. Results show that the precision and recall of our method are 95.53% and 96.67%, respectively. Results further demonstrate that the total prediction accuracy of our method is nearly 13.17% greater than the state-of-the-art microRNA precursor prediction software tools.

1. Introduction

MicroRNAs are some of the most important noncoding RNA genes with rather short length. They regulate the expression of whole organism genes at the posttranscriptional level [1]. miRNA is widely involved in the metabolic activity of the body as well as in many important life processes, including cell proliferation and apoptosis, cell differentiation, growth and development of plants and animals, and organ formation [2–4]. Recently, several studies have shown that microRNAs are related to several cancers [5–7] and other diseases [8–10]. Caligiuri et al. [11] proposed that methods and compositions involving miRNAs are useful for the treatment of various diseases and cancers. Some miRNAs are regarded as potential therapeutic targets for various diseases [12]. Recently, the target gene (cancer gene) drugs, which developed in accordance with the theory on miRNA's gene silencing, have been used for incurable disease that has become a threat to human health problems for years [13]. In addition, the viral genome can encode a large number of miRNAs by

itself. Through combination with target genes and coding by viruses or host cell, these miRNAs can lead to immune escape or antiviral effect against the host cell. Therefore, the accurate prediction of miRNA and its target genes, as well as the correct understanding of miRNA mechanism, has important practical significance in medical treatments. Thus, the research on novel miRNA identification is rather essential.

Feature selection mainly dominated the performance of the prediction model in the machine learning process [14–20]. In addition, effective features can represent the characteristics of the entire sequence data, which enables easy-to-build better prediction model. To represent the microRNA precursors, Xue et al. [21] proposed 32D novel triplet features, which involved secondary structure information. Jiang et al. [22] found that random rearrangement of the sequence could help obtain significant free-energy features. However, the free-energy computation for many random rearrangement sequences is very time consuming. Wei et al. [23] combined Xue et al.'s features and triplet nucleotide frequency to 98D

features and obtained good performance result in human pre-miRNA identification. However, more features would not mean better performance because of some irrelevant and redundant features in the high dimensional or ultra-high dimensional feature set. The purpose of feature selection is to eliminate the irrelevant and redundant features of the feature set. In addition, the training time could be reduced effectively by the feature selection optimization [24]. Some studies focus on developing computational predictors by incorporating the sequence-order or structure-order effects [25, 26]. Several works indicated that proper features could improve the prediction performance of classification in a certain extent. For example, Wang et al. [27] employed the feature selection techniques to optimize the features in miR-SF. They proved that an optimized feature subset could improve the prediction performance. In addition, some popular recently proposed multiobjective optimization evolutionary algorithms can also be used as a possibly promising feature selection approach [28, 29].

Another factor that affects the performance of machine learning prediction method is the classifier algorithm. The selection of different classifiers often leads to the difference of classification results. Several different classifiers and strategies were employed for miRNA identification. Bayesian classifier algorithm was tested for predicting miRNA across different species in 2006 [30]. The method also utilized the multiple species of miRNA sequences and structural features. It proved that miRNA genes could be detected effectively in large scale of different species genomes.

MiPred classifier was tested for predicting miRNA in 2007 [22]. The method utilized random forest classifier algorithm. The prediction accuracy of MiPred is 10% higher than that of Triplet-SVM; the sensitivity and specificity of MiPred can reach to 95.09% and 98.21%. CSHMM classifier was also used for mining miRNA sequences from the genome [31], which utilized the Markov model. Overall, the accuracy of machine learning algorithm was up to 90%. The machine learning method is more accurate than the other methods.

In this study, we chose backpropagation neural network as the classifier. It has three advantages, including better generalization performance, faster learning speed, and good learning ability.

2. miRNA Identification with BP Neural Network

2.1. Pre-miRNA Features

2.1.1. *n*-Gram Frequency. Some studies showed that the local primary sequence is crucial to the pre-miRNA sequence [32]. Thus, the *n*-gram frequency is often applied for the feature map in the selection of the primary sequence feature [33, 34]. However, no good methods are still available for tuning the value of *n*. In general, we choose *n* by comparing the effect of *n*-gram frequency with different *n*-values. In our feature set, we select the different *n* values ($n = 2, 3, 4$) for comparison. The different frequency characteristics have almost the same effect on the classifier. Thus, consider that its base and

adjacent base have practical biological significance. We chose *n* as 3. A total of 64 ($64 = 4^3$)-dimensional frequency features were calculated.

2.1.2. Triple Structure Sequence. In addition to high specificity of the primary sequence features, the secondary structure sequence of pre-miRNA is also a contributing factor. To analyze the contribution of the secondary structure, the secondary structure prediction software RNAfold is used to calculate the potential structures. In the secondary structure, each nucleotide of the sequence corresponds to two states, matching and nonmatching: record matching as “(” or “)” and nonmatching as “.” In the structure, three character groups are considered as a unit, and every “)” is replaced as “(.” Thus, 8 ($8 = 2^3$) different combinations are available as a unit, including “(((,” “((,” “(.”,” “(.”,” “(.”,” “(.”,” “(.”,” and “(.”.”

To characterize pre-miRNA sequence better, the first nucleotide of the corresponding subsequence was added to the front of each structure unit. This provides 32 ($32 = 4 \times 8$) different combinations, that is, “A(((,” “U(((,” ..., “G(.”,” “G(.”.” For a sequence, the occurrence frequency of each combination is determined and coded into the 32D feature vector as the input of the classifier. This calculated 32D triple structure sequence feature is used to train the SVM classifier; the inclusion of the SVM classifier significantly improved the classification ability of pre-miRNA sequences [21].

2.1.3. Energy Characteristics. The real pre-miRNA sequences are generally more stable and show a lower minimum of free energy (MFE) than the randomly generated pre-miRNA. Therefore, energy characterization is often used to describe the structure pre-miRNA sequence as an aspect of feature extraction of the pre-miRNA sequence. To do this, the MFE value is obtained by using RNAfold to calculate the structure.

2.1.4. Structural Diversity Characteristics. The potential for nucleotide pairing in the sequence is a significant characteristic that can also be used to describe the pre-miRNA sequence. This includes both traditional Watson-Crick nucleotide pairing (A–U pairing and C–G pairing) and also other forms of nucleotide pairing, such as the G–U pairing that can occur in the loop of RNA hairpin structures. We included possible G–U pairing in our description of base pairing.

To summarize, we extracted 98 features for the input of the neural network, including 64-dimensional *n*-gram frequency characteristics, 32-dimensional triple structure sequence characteristics, one-dimensional energy feature, and one-dimensional structural diversity characteristics.

2.2. Fixing the Number of Nodes in the Hidden Layer. In general, to select the number of nodes in the hidden layer in changing the BP neural network structure is difficult. Technically, a hidden layer could facilitate operation. However, too many hidden layers can reduce the operation rate.

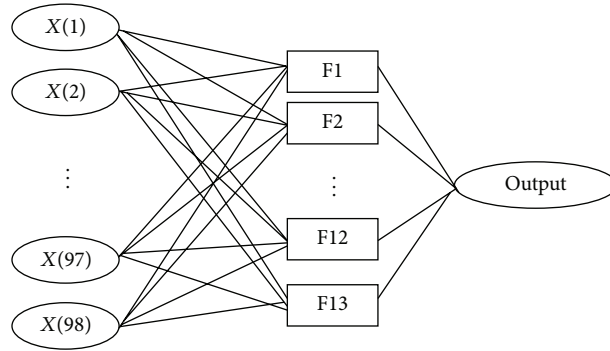


FIGURE 1: Topology structure of the BP neural network.

TABLE 1: Corresponding training results with different numbers of nodes in the hidden layers.

Hidden layers	Training times	Training errors	Hidden layers	Training times	Training errors
11	43	$9.57718e - 005$	12	39	$9.88418e - 005$
13	17	$9.42136e - 005$	14	65	$9.92537e - 005$
15	34	$9.88206e - 005$	16	74	$8.38658e - 005$
17	48	$7.82527e - 005$	18	157	$6.63468e - 005$
19	7	$9.46711e - 005$	20	47	$9.3627e - 005$

Currently, no theoretical methods are available to fix the number of nodes in the hidden layer. However, the number generally depends on the empirical formula, as calculated in

$$\begin{aligned}
 M &= \sqrt{N + L} + \alpha, \\
 M &= \log_2 N, \\
 M &= \sqrt{NL},
 \end{aligned}
 \tag{1}$$

where M represents the neuron number of the hidden layers, N is the neuron number of the input layers, L is the neuron number of the output layers, and α is a constant between 1 and 10.

In this study, $N = 98$ and $L = 1$. Therefore, (1) can be used for any values between 11 and 20. A comprehensive analysis of the training results with different numbers of nodes in the hidden layer was performed with the error set to 0.0001. A total of 621 samples were used to train the network, and one sample was used to test the network. The results are shown in Table 1.

From the data shown in Table 1, the increased number of nodes in the hidden layer did not result in better convergence. Additionally, the increased number of nodes increased the network parameters and greatly increased the amount of calculation of the classifier. Thus, keeping 13 nodes in the hidden layers required relatively less training times and less error and still produced relatively good training effects.

2.3. Fixing the Number of Nodes in the Output Layer. Two kinds of output exist, positive and negative, which are represented as 1 for a positive sample and 0 for a negative sample. The topology structure of this prediction method based on BP neural network is shown in Figure 1.

2.4. Selecting Training and Test Model Samples. The collection and organization of training samples are often limited by the objective conditions. Appropriate numbers of training samples are required to achieve sufficient precision. Therefore, it refers to the rule of experience:

$$P = (5 \sim 10) \times P_w,
 \tag{2}$$

where P represents the numbers of training samples and P_w is the total of network connection weight equal to the sum of nodes of the input and hidden layers. In this study, 2236 samples were used for training.

The data set used for the pre-miRNAs was downloaded from <http://bioinf.sce.carleton.ca/SMIRP> [35], and these data include negative and positive samples for *Arabidopsis lyrata*. The FASTA file was converted to ARFF file using a jar package written by Java converting the reference index to numerical form. We randomly selected real pre-miRNAs and pseudo pre-miRNAs to evaluate our algorithm.

2.5. Error Evaluation Steps Based on BP. The structure of the intelligent diagnosis model contains three layers of 98-13-1. First, we set the nodes of the input, output, and hidden layers as N , M , and L , respectively. Assuming the training sample set $\{\xi^p, Y\} \subset R^N \times R^L$, the weight matrix between the input and hidden layers can be written as $V = (v_{mn})_{M \times N}$, where $V_m = (v_{m1}, v_{m1}, \dots, v_{mN})^T \in R^N$ and $m = 1, 2, \dots, M$. We assume the connection weight matrix between the hidden and output layers as $W = (w_{lm})_{L \times M}$, where $W_l = (v_{l1}, v_{l1}, \dots, v_{lm})^T \in R^M$, $l = 1, 2, \dots, L$. Then, respectively, take g and f as the activation function of each node of the hidden and output layers. To simplify the derivation, we use the vector function $G(X)$ for $X = (x_1, x_2, \dots, x_m)^T \in R^M$, where

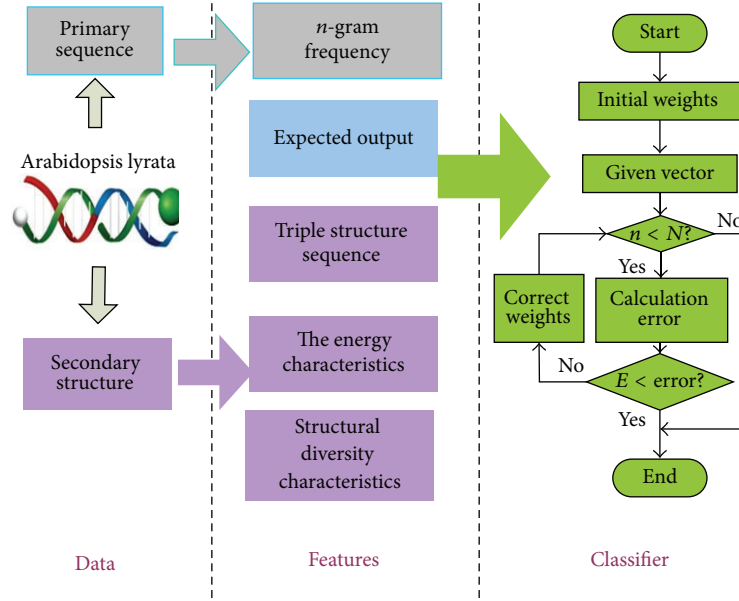


FIGURE 2: Process flow of model generation and training.

$G(X) = (g(x_1), g(x_2), \dots, g(x_m))^T \in R^M$. After input of the sample $\xi^p \in R^N$, the actual output can be calculated by

$$\zeta_i^p = f(w_i \cdot G(V_i \xi^p)). \quad (3)$$

The error function is defined in

$$E(W, V) = \frac{1}{2} \sum_{p=1}^P \sum_{l=1}^L (O^p - f(w_l \cdot G(V_l \xi^p)))^2. \quad (4)$$

Objectively, the target of BP training is to compute the W and V to minimize the solution of the error function $E(W, V)$. With this, a combination of gradient descent, common, and simple derivatives was used. To simplify the derivation process, we derive

$$f_{pl}(x) = \frac{1}{2} (O_l^p - f(x))^2. \quad (5)$$

Then, the error function can be written as

$$E(W, V) = \sum_{p=1}^P \sum_{l=1}^L f_{pl}(w_l \cdot G(V_m \xi^p)). \quad (6)$$

The corresponding gradient function of W and V can then be expressed as

$$\begin{aligned} E_{wl}(W, V) &= \sum_{p=1}^P f_{pl}^l(w_l \cdot G(V \xi^p)) G(V \xi^p), \\ E_{vm}(W, V) &= \sum_{p=1}^P \sum_{l=1}^L f_{pl}^l(w_l \cdot G(V \xi^p)) w_{lm} \dot{g}(v_m \cdot \xi^p) \xi^p. \end{aligned} \quad (7)$$

TABLE 2: Basic parameters of the classifier based on BP neural network.

Setting items	The value set
The learning rate	0.1
Error bounds	0.0001
The number of iterations	1000
Transfer function of hidden layer nodes	Tansig
Transfer function of output nodes	Purelin
The training function	Trainlm

For arbitrary initial values of $W_0 \in R^{L \times M}$ and $V_0 \in R^{M \times N}$, gradient descent rules to modify the weight of the BP learning algorithm are applied in

$$\begin{aligned} W_l^{n+1} &= W_l^n + \Delta W_l^n, \\ \Delta W_l^n &= -\eta_n E_{wl}(W^n, V^n), \\ V_m^{n+1} &= V_m^n + \Delta V_m^n, \\ \Delta V_m^n &= -\eta_n E_{vm}(W^n, V^n), \end{aligned} \quad (8)$$

where η_n represents the learning efficiency. ΔW_l^n is the partial derivative of the error function relative to W . ΔV_m^n is the partial derivative of the error function relative to V .

2.6. Selection of Training Functions and Related Parameters. The above analysis allows fixing of the BP neural network structure. Table 2 shows the chosen training functions and the relevant parameters.

This condition allows establishment of a complete classifier based on BP neural network structure. The model generation and training are summarized in Figure 2.

TABLE 3: Measurements for the classification problems.

Actual result	Classification result	
	Forecast result	
	P	N
P	TP	FN
N	FP	TN

2.7. Measurement. The use of pattern recognition and machine learning methods can be used as a two-way classification problem. Four kinds of prediction results are presented in Table 3.

The four kinds of prediction results are true positive (TP), the number of positive cases that were correctly predicted; false positive (FP), the number of positive cases represented by error prediction; true negative (TN), the number of counter negative examples that were correctly predicted; and false negative (FN), the number of negative cases represented by error prediction.

Many evaluation indicators can be used for the classification results. First, the accuracy rate (ACC) is the ratio of the correctly predicted cases for the entire data set. Precision and recall can also be used as evaluation indicators in tests of pattern recognition models. Precision is expressed as the ratio of the correctly predicted values for the entire positive data set and recall reflects the number correctly judged as positive examples in the positive example test set [36]. The above three indicators are expressed in

$$\begin{aligned} \text{ACC} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}, \\ \text{precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}. \end{aligned} \quad (9)$$

Additionally, sensitivity and specificity parameters may be used to evaluate the function of the model. Sensitivity record (SE) is the same as the recall and specificity record (SP) calculated in accordance with

$$\text{SP} = \frac{\text{TN}}{\text{TN} + \text{FP}}. \quad (10)$$

A challenge may be presented if the positive and negative test sets are unbalanced in the study of biological information. In most cases, the number of positive samples is far less than the number of negative samples. In a few cases, the number of positive samples may be much larger than the number of negative samples. We can easily obtain ACC-SP when the number of positive samples is greater than the negative samples. In this case, the classifier only reflects the classification effect of the negative samples and is unable to accurately express the prediction effect of the classifier on the

TABLE 4: Comparison of classification results based on different feature sets.

Features	SP (%)	SE (%)	Gm (%)	ACC (%)
B	67.89	68.25	68.07	68.00
C	92.74	76.42	84.19	88.03
A + B	91.79	90.41	91.10	91.31
A + C	94.03	80.85	87.19	89.67
B + C	96.12	85.21	90.50	92.49
A + B + C	96.33	86.51	91.29	93.42

Notes: A: energy feature and structural diversity; B: 32-dimensional triad structure characteristic; C: 64-dimensional n -gram frequency characteristics.

entire test data set. To solve this problem, researchers typically use the geometric mean (Gm) as described in

$$\text{Gm} = \sqrt{\text{SE} \cdot \text{SP}}. \quad (11)$$

Matthew's correlation coefficient (MCC) [16, 21, 37, 38] can provide more equitable response forecast ability when a large difference exists between the number of positive samples and the number of negative samples. MCC can be expressed as

$$\begin{aligned} \text{MCC} \\ &= \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP} + \text{FP}) \times (\text{TN} + \text{FN}) \times (\text{TP} + \text{FN}) \times (\text{TN} + \text{FP})}}. \end{aligned} \quad (12)$$

Currently, studies on miRNA commonly use one or more of these above evaluation indices. In this work, we estimate the overall performance of the classifier by analysis of ACC, SE, SP, Gm, and MCC.

3. Results and Discussion

3.1. Analysis of Feature Set Performance. To select a better feature set for classification, we needed to determine the effect of different feature subsets on the performance of the classifier. To do this, we used the BP neural network method with the same training set (553 positive samples and 1150 samples) to test different feature sets, with the results shown in Table 4.

From Table 4, we learn that the accuracy of the entire feature sets can be as high as 93.42%. This result indicates that our feature set is more effective for processing of a more complex structure or sequence diversity. Considering that the feature sets used here are not very large and each feature subset is highly independent, reducing the dimension of the feature vector is no longer needed.

3.2. Performances of BP. V -fold cross-validation with moderate computational complexity is widely used for model selection. The selection of V is important because V not only determines the number of samples but also determines the computational complexity. Usually, a value of V between 5 and 10 is selected based on experience. Statistical

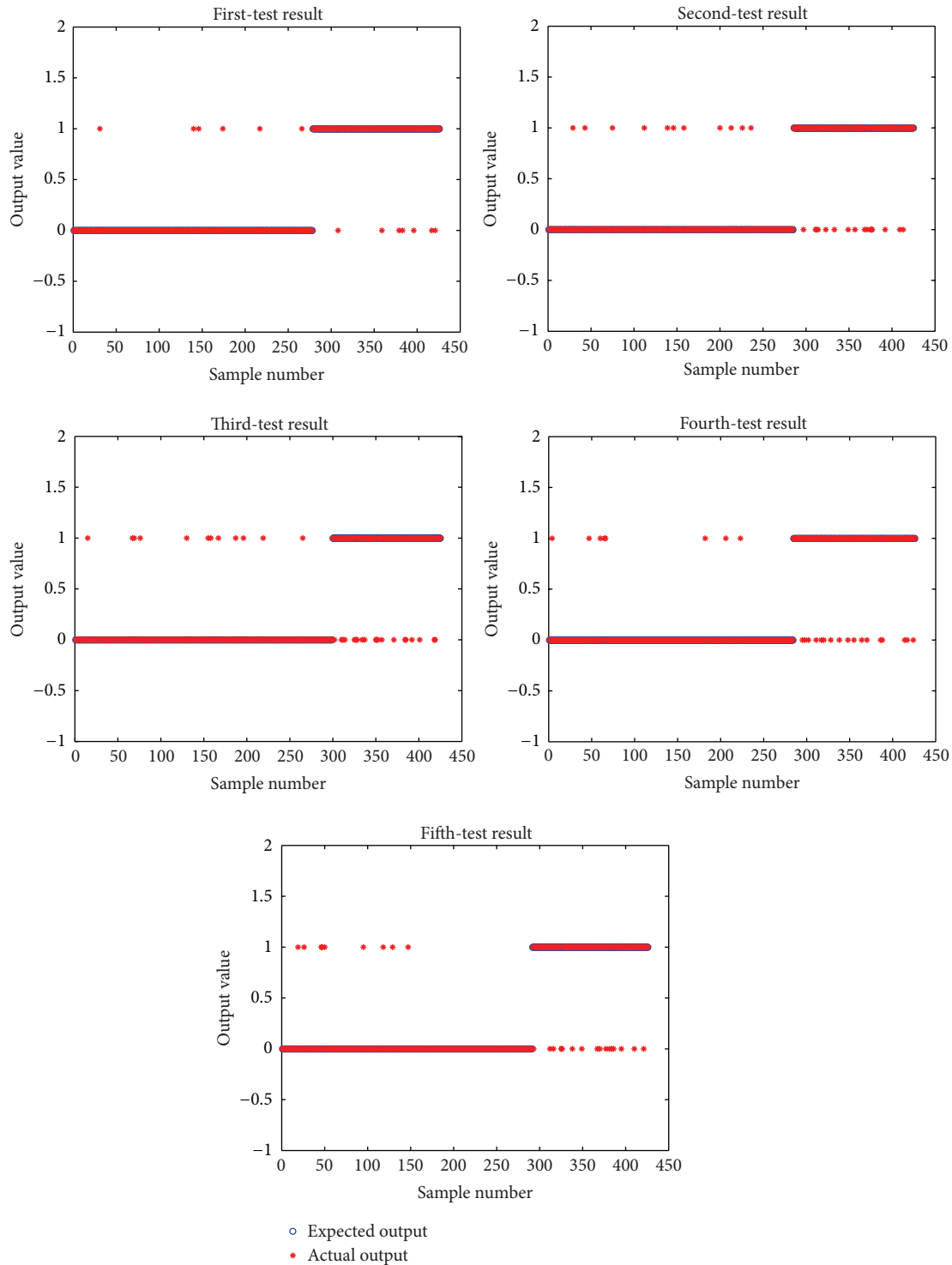


FIGURE 3: Different test results for varying sample quantities.

performance shows little improvement when V selection is greater than 10. Again, computational complexity must be considered; thus a value between 5 and 10 is best [32].

We divided the samples into two cases for training and testing. In the first one, a large difference was observed between the number of positive and negative samples: 518 positive samples and 1078 negative samples as the training

set and 166 positive samples and 366 negative samples for the test set. The second case included equal numbers of positive and negative samples: 552 positive samples and 552 negative samples as the training set and 138 positive samples and 138 negative samples for the test set. These training and testing were repeated five times. The testing performance is shown in Figures 3 and 4.

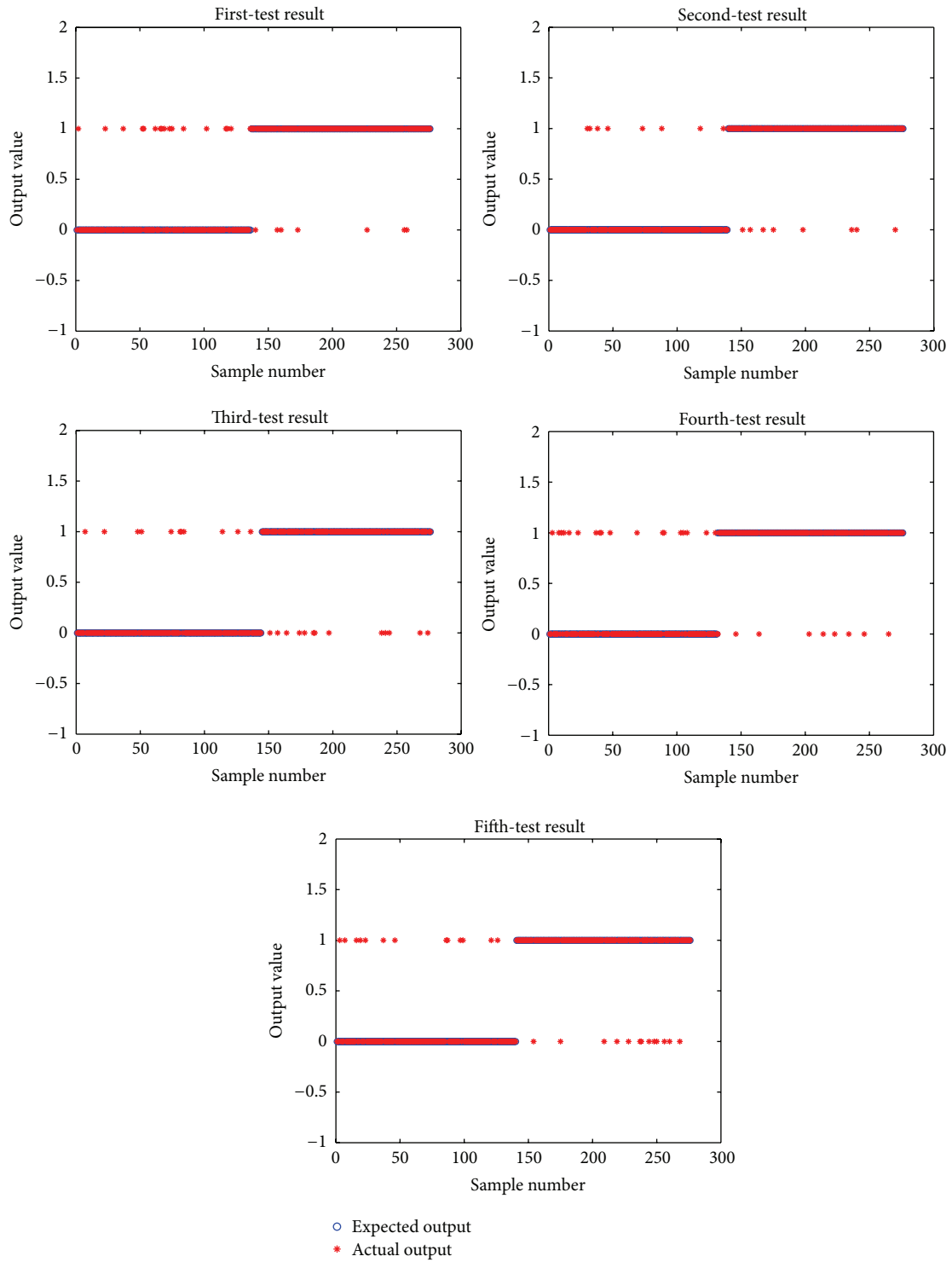


FIGURE 4: Different test results for same sample quantity.

From comparison of the data in Figures 3 and 4, no significant difference was observed between the actual output and the expected output of each test. As described above, the evaluation of the reference index is shown in Table 5.

From the data presented in Table 5, the number of samples affects the accuracy and recall rate of the positive samples. In particular, the precision and recall rate of the negative samples decreased with the decrease in the number of negative samples in the training set. This result indicates that

TABLE 5: Evaluation of the reference index.

	Training sample	Test sample	Output sample	Correct sample	Precision (%)	Recall (%)	Gm (%)
D							
Positive	553	138	128	124	96.0	90.0	93.43
Negative	1150	287	296	282	95.38	98.19	
E							
Positive	552	138	136	128	94.10	92.82	93.98
Negative	552	138	140	130	92.87	94.12	

Note: D: sample set has different numbers of positive and negative samples; E: the sample set has equal numbers of positive and negative samples; correct sample: the number of correct predictions.

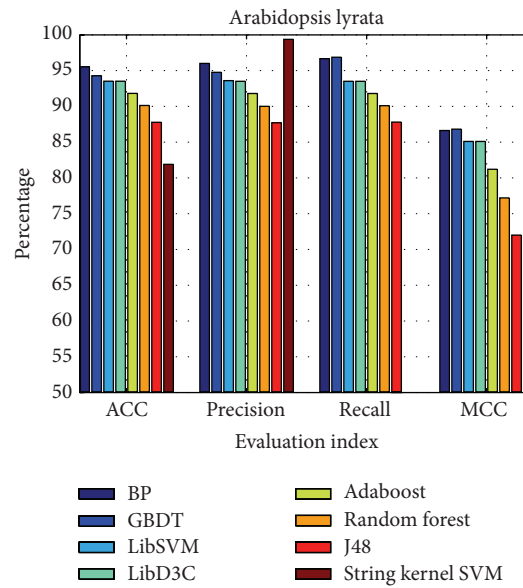


FIGURE 5: Comparison results of different models.

the more the samples in the training process, the better the classification effect of the classifier. At the same time, the precision and recall rate of the number of positive samples were affected. With the number of negative samples in the training set increased, the number of correct predictions increased by four and the number of error predictions was reduced by eight. This result shows that the precision and recall rate of the positive samples decreased with the increase in the number of the negative samples.

3.3. Comparison with Other Methods. The performance of our method was compared with other methods: J48, random forest, LibD3C [39], Adaboost, string kernel SVM [40], LibSVM, and GBDT, which were classified on the same data set. The data set contains 691 real pre-miRNAs and 1437 pseudo pre-miRNAs. As shown in Table 6 and Figure 5, the results demonstrate that the total prediction accuracy of our method is 13.64% greater than the string kernel SVM model and nearly 2% greater than the LibD3C and LibSVM models. The overall performance of the models as measured by MCC was in the following order: GBDT (0.8682), BP (0.8662), LibSVM (0.8510), LibD3C (0.8510), Adaboost (0.8120), random forest (0.7720), J48 (0.7200), and string kernel SVM (0.6002).

TABLE 6: Comparison of the BP with alternative models.

	ACC	Precision	Recall	MCC
BP	95.53%	96.00%	96.67%	0.8662
GBDT	94.27%	94.76%	96.87%	0.8682
LibSVM	93.52%	93.60%	93.50%	0.8510
LibD3C	93.52%	93.50%	93.50%	0.8510
Adaboost	91.82%	91.80%	91.80%	0.8120
Random forest	90.13%	90.00%	90.10%	0.7720
J48	87.78%	87.70%	87.80%	0.7200
String kernel SVM	81.89%	99.37%	46.31%	0.6002

Thus, we conclude that the BP method allows improved recognition accuracy.

3.4. Performance on Different Species. To demonstrate the validity and the universal applicability of the BP method, we analyzed six other species: *Anolis carolinensis*, *Arabidopsis thaliana*, *Drosophila melanogaster*, *Drosophila pseudoobscura*, Epstein-Barr virus, and *Xenopus tropicalis*. The results shown

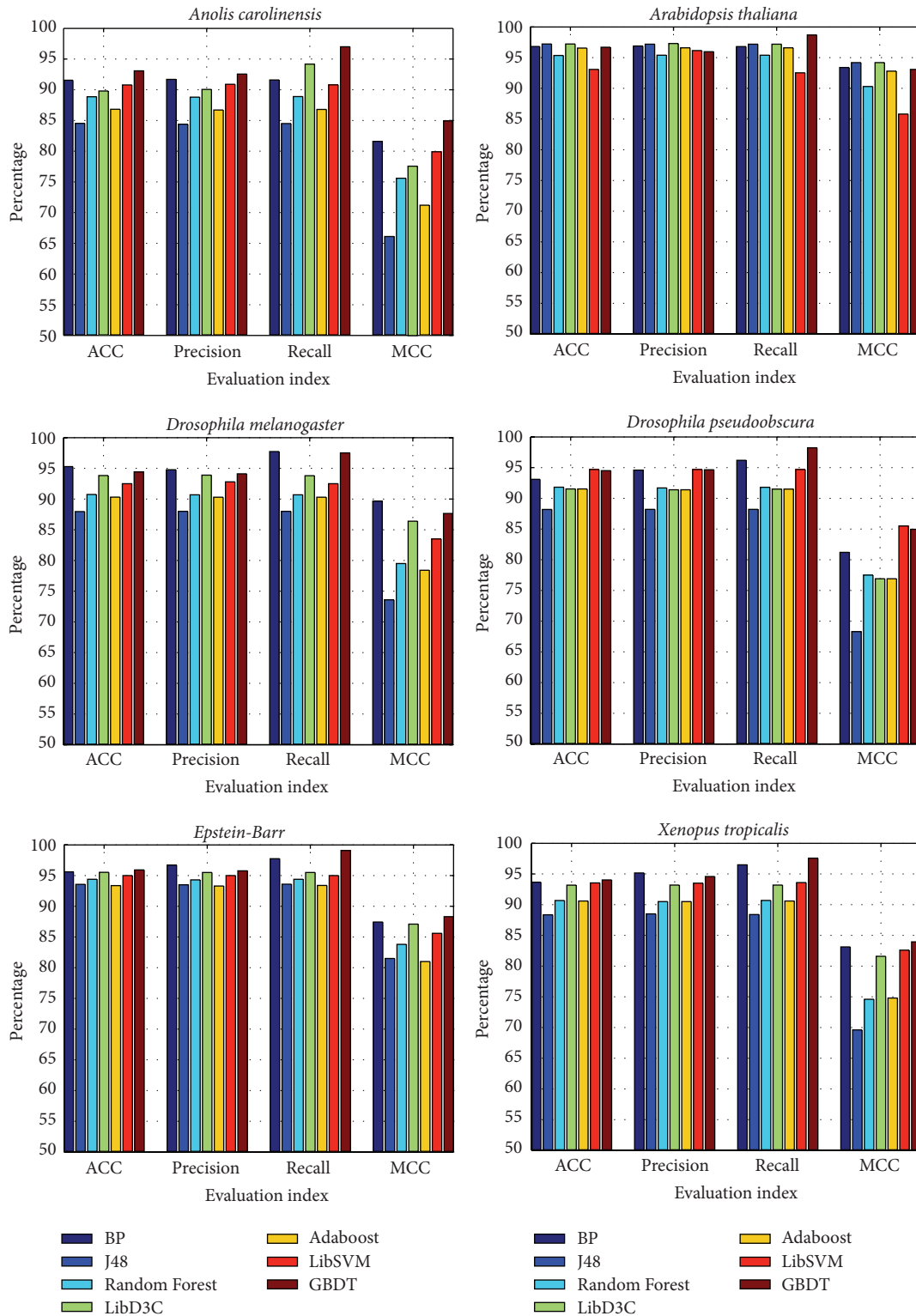


FIGURE 6: Test comparison results for six different species.

in Figure 6 indicate that the accuracy of the GBDT is better than BP method in some situations, but the BP method has been achieved fairly good results in terms of ACC, precision, recall, and MCC.

4. Conclusions

Identification of miRNAs is the first step toward understanding their biological characteristics. Many approaches

have been proposed to predict pre-miRNAs in recent years. However, feature extraction in these methods can result in information redundancy. To overcome this drawback, a BP neural network algorithm together with optimal 98D features was employed for this analysis. We compare our method with the existing methods of J48, random forest, LibD3C, Adaboost, GBDT, string kernel SVM, and LibSVM, which were trained on the same training data set. The results demonstrate that the total prediction accuracy of our method is 13.17% greater than the string kernel SVM model and nearly 2% greater than LibD3C and LibSVM.

After the identification step, functional analysis is also important for miRNA research. If human miRNA and diseases were focused on, two main approaches would be employed to predict the relationship. The first one is the statistical comparison analysis for the miRNA or isomiR expression [41]. The second one is the network analysis and prediction for miRNA-disease relationship [42–45]. Several advanced machine learning, network techniques, and bio-inspired models can be utilized on this problem, including random forest [46], semisupervised learning [47], HeteSim Scores [48], spiking neural P systems [49–52], and membrane computing *ENREF_51* [53–57]. Functional analysis of the novel detected miRNAs would be our future works.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The work was supported by the Natural Science Foundation of China (no. 61370010 and no. 61302139) and the State Key Laboratory of Medicinal Chemical Biology in China.

References

- [1] D. P. Bartel, “microRNAs: genomics, biogenesis, mechanism, and function,” *Cell*, vol. 116, no. 2, pp. 281–297, 2004.
- [2] D. Wu, Y. Huang, J. Kang et al., “ncRDeathDB: a comprehensive bioinformatics resource for deciphering network organization of the ncRNA-mediated cell death system,” *Autophagy*, vol. 11, no. 10, pp. 1917–1926, 2015.
- [3] Y. Huang, N. Liu, J. P. Wang et al., “Regulatory long non-coding RNA and its functions,” *Journal of Physiology and Biochemistry*, vol. 68, no. 4, pp. 611–618, 2012.
- [4] X. Zhang, D. Wu, L. Chen et al., “RAID: a comprehensive resource for human RNA-associated (RNA–RNA/RNA–protein) interaction,” *RNA*, vol. 20, no. 7, pp. 989–993, 2014.
- [5] S. Hua, W. Yun, Z. Zhiqiang, and Q. Zou, “A discussion of micrnas in cancers,” *Current Bioinformatics*, vol. 9, no. 5, pp. 453–462, 2014.
- [6] Q. Wang, L. Wei, X. Guan, Y. Wu, Q. Zou, and Z. Ji, “Briefing in family characteristics of microRNAs and their applications in cancer research,” *Biochimica et Biophysica Acta—Proteins and Proteomics*, vol. 1844, no. 1, pp. 191–197, 2014.
- [7] C. Yang, D. Wu, L. Gao et al., “Competing endogenous RNA networks in human cancer: hypothesis, validation, and perspectives,” *Oncotarget*, vol. 7, no. 12, pp. 13479–13490, 2016.
- [8] X. Zeng, X. Zhang, and Q. Zou, “Integrative approaches for predicting microRNA function and prioritizing disease-related microRNA using biological interaction networks,” *Briefings in Bioinformatics*, vol. 17, no. 2, pp. 193–203, 2016.
- [9] Q. Zou, J. Li, Q. Hong et al., “Prediction of microRNA-disease associations based on social network analysis methods,” *BioMed Research International*, vol. 2015, Article ID 810514, 9 pages, 2015.
- [10] Y. Wang, L. Chen, B. Chen et al., “Mammalian ncRNA-disease repository: a global view of ncRNA-mediated disease network,” *Cell Death and Disease*, vol. 4, no. 8, article e765, 2013.
- [11] M. A. Caligiuri, J. Yu, S. He, and R. Trott, “Activation of Innate Immunity by miRNA for Cancer and Infection Treatment,” United States Patent, 2016.
- [12] H. Zhou, X. Ge, and X. Xue, “microRNAs regulation and its role as biomarkers in diseases,” *Oncology and Translational Medicine*, vol. 2, no. 1, pp. 39–46, 2016.
- [13] P. S. Kelly, C. Gallagher, M. Clynes, and N. Barron, “Conserved microRNA function as a basis for Chinese hamster ovary cell engineering,” *Biotechnology Letters*, vol. 37, no. 4, pp. 787–798, 2015.
- [14] Q. Zou, J. Zeng, L. Cao, and R. Ji, “A novel features ranking metric with application to scalable visual and bioinformatics data classification,” *Neurocomputing*, vol. 173, pp. 346–354, 2016.
- [15] W.-X. Liu, E.-Z. Deng, W. Chen, and H. Lin, “Identifying the subfamilies of voltage-gated potassium channels using feature selection technique,” *International Journal of Molecular Sciences*, vol. 15, no. 7, pp. 12940–12951, 2014.
- [16] H. Tang, W. Chen, and H. Lin, “Identification of immunoglobulins using Chou’s pseudo amino acid composition with feature selection technique,” *Molecular BioSystems*, vol. 12, no. 4, pp. 1269–1275, 2016.
- [17] P.-P. Zhu, W.-C. Li, Z.-J. Zhong et al., “Predicting the subcellular localization of mycobacterial proteins by incorporating the optimal tripeptides into the general form of pseudo amino acid composition,” *Molecular BioSystems*, vol. 11, no. 2, pp. 558–563, 2015.
- [18] H. Ding, P.-M. Feng, W. Chen, and H. Lin, “Identification of bacteriophage virion proteins by the ANOVA feature selection and analysis,” *Molecular BioSystems*, vol. 10, no. 8, pp. 2229–2235, 2014.
- [19] H. Ding, S.-H. Guo, E.-Z. Deng et al., “Prediction of Golgi-resident protein types by using feature selection technique,” *Chemometrics and Intelligent Laboratory Systems*, vol. 124, pp. 9–13, 2013.
- [20] H. Ding, H. Lin, W. Chen et al., “Prediction of protein structural classes based on feature selection technique,” *Interdisciplinary Sciences: Computational Life Sciences*, vol. 6, no. 3, pp. 235–240, 2014.
- [21] C. Xue, F. Li, T. He, G.-P. Liu, Y. Li, and X. Zhang, “Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine,” *BMC Bioinformatics*, vol. 6, article 310, 2005.
- [22] P. Jiang, H. Wu, W. Wang, W. Ma, X. Sun, and Z. Lu, “MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features,” *Nucleic Acids Research*, vol. 35, no. 2, pp. W339–W344, 2007.
- [23] L. Wei, M. Liao, Y. Gao, R. Ji, Z. He, and Q. Zou, “Improved and promising identification of human microRNAs by incorporating a high-quality negative set,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 11, no. 1, pp. 192–201, 2014.

- [24] Y. Wang, X. Chen, W. Jiang et al., "Predicting human microRNA precursors based on an optimized feature subset generated by GA-SVM," *Genomics*, vol. 98, no. 2, pp. 73–78, 2011.
- [25] B. Liu, L. Fang, F. Liu, X. Wang, J. Chen, and K.-C. Chou, "Identification of real microRNA precursors with a pseudo structure status composition approach," *PLoS ONE*, vol. 10, no. 3, article e0121501, 2015.
- [26] B. Liu, L. Fang, J. Chen, F. Liu, and X. Wang, "MiRNA-dis: MicroRNA precursor identification based on distance structure status pairs," *Molecular BioSystems*, vol. 11, no. 4, pp. 1194–1204, 2015.
- [27] X. Wang, J. D. Laurie, T. Liu, J. Wentz, and X. S. Liu, "Computational dissection of Arabidopsis smRNAome leads to discovery of novel microRNAs and short interfering RNAs associated with transcription start sites," *Genomics*, vol. 97, no. 4, pp. 235–243, 2011.
- [28] X. Zhang, Y. Tian, R. Cheng, and Y. Jin, "An efficient approach to nondominated sorting for evolutionary multiobjective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 2, pp. 201–213, 2015.
- [29] X. Zhang, Y. Tian, and Y. Jin, "A knee point-driven evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 19, no. 6, pp. 761–776, 2014.
- [30] M. Yousef, M. Nebozhyn, H. Shatkay, S. Kanterakis, L. C. Showe, and M. K. Showe, "Combining multi-species genomic data for microRNA identification using a Naïve Bayes classifier," *Bioinformatics*, vol. 22, no. 11, pp. 1325–1334, 2006.
- [31] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Statistics Surveys*, vol. 4, pp. 40–79, 2010.
- [32] E. Bonnet, J. Wuyts, P. Rouzé, and Y. Van de Peer, "Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences," *Bioinformatics*, vol. 20, no. 17, pp. 2911–2917, 2004.
- [33] H. Liu and L. Wong, "Data mining tools for biological sequences," *Journal of Bioinformatics and Computational Biology*, vol. 1, no. 1, pp. 139–167, 2003.
- [34] B. Liu, F. Liu, X. Wang, J. Chen, L. Fang, and K. Chou, "Pse-in-One: a web server for generating various modes of pseudo components of DNA, RNA, and protein sequences," *Nucleic Acids Research*, vol. 43, no. 1, pp. W65–W71, 2015.
- [35] R. J. Peace, K. K. Biggar, K. B. Storey, and J. R. Green, "A framework for improving microRNA prediction in non-human genomes," *Nucleic Acids Research*, vol. 43, no. 20, article e138, 2015.
- [36] S. Yang, S. Cai, F. Zheng et al., "Representation of fluctuation features in pathological knee joint vibroarthrographic signals using kernel density modeling method," *Medical Engineering and Physics*, vol. 36, no. 10, pp. 1305–1311, 2014.
- [37] R. Wang, Y. Xu, and B. Liu, "Recombination spot identification Based on gapped k-mers," *Scientific Reports*, vol. 6, Article ID 23934, 2016.
- [38] Y. Wu, P. Chen, X. Luo et al., "Quantification of knee vibroarthrographic signal irregularity associated with patellofemoral joint cartilage pathology based on entropy and envelope amplitude measures," *Computer Methods and Programs in Biomedicine*, vol. 130, pp. 1–12, 2016.
- [39] C. Lin, W. Chen, C. Qiu, Y. Wu, S. Krishnan, and Q. Zou, "LibD3C: ensemble classifiers with a clustering and dynamic selection strategy," *Neurocomputing*, vol. 123, pp. 424–435, 2014.
- [40] M. Ghandi, D. Lee, M. Mohammad-Noori, and M. A. Beer, "Enhanced regulatory sequence prediction using gapped k-mer features," *PLoS Computational Biology*, vol. 10, no. 7, Article ID e1003711, 2014.
- [41] L. Guo, J. Yu, T. Liang, and Q. Zou, "miR-isomiRExp: a web-server for the analysis of expression of miRNA at the miRNA/isomiR levels," *Scientific Reports*, vol. 6, Article ID 23700, 2016.
- [42] 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.
- [43] X. Chen, C. Clarence Yan, X. Zhang et al., "RBMMMDA: predicting multiple types of disease-microRNA associations," *Scientific Reports*, vol. 5, article 13877, 2015.
- [44] 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.
- [45] X. Chen, "miREFRWR: a novel disease-related microRNA-environmental factor interactions prediction method," *Molecular BioSystems*, vol. 12, no. 2, pp. 624–633, 2016.
- [46] Y. Liu, X. Zeng, Z. He, and Q. Zou, "Inferring microRNA-disease associations by random walk on a heterogeneous network with multiple data sources," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2016.
- [47] X. Chen and G.-Y. Yan, "Semi-supervised learning for potential human microRNA-disease associations inference," *Scientific Reports*, vol. 4, article 5501, 2014.
- [48] X. Zeng, Y. Liao, Y. Liu, and Q. Zou, "Prediction and validation of disease genes using HeteSim Scores," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2016.
- [49] T. Song, J. Xu, and L. Pan, "On the universality and non-universality of spiking neural P systems with rules on synapses," *IEEE Transactions on NanoBioscience*, vol. 14, no. 8, pp. 960–966, 2015.
- [50] X. Zhang, X. Zeng, B. Luo, and L. Pan, "On some classes of sequential spiking neural P systems," *Neural Computation*, vol. 26, no. 5, pp. 974–997, 2014.
- [51] T. Song and L. Pan, "Spiking neural P systems with request rules," *Neurocomputing*, vol. 193, pp. 193–200, 2016.
- [52] X. Wang, T. Song, F. Gong, and P. Zheng, "On the computational power of spiking neural P systems with self-organization," *Scientific Reports*, vol. 6, Article ID 27624, 2016.
- [53] X. Zeng, L. Xu, X. Liu, and L. Pan, "On languages generated by spiking neural P systems with weights," *Information Sciences*, vol. 278, pp. 423–433, 2014.
- [54] X. Zhang, L. Pan, and A. Păun, "On the universality of axon P systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 11, pp. 2816–2829, 2015.
- [55] X. Chen, M. J. Pérez-Jiménez, L. Valencia-Cabrera, B. Wang, and X. Zeng, "Computing with viruses," *Theoretical Computer Science*, vol. 623, pp. 146–159, 2016.
- [56] T. Wu, Z. Zhang, G. Păun, and L. Pan, "Cell-like spiking neural P systems," *Theoretical Computer Science*, vol. 623, pp. 180–189, 2016.
- [57] X. Zhang, Y. Liu, B. Luo, and L. Pan, "Computational power of tissue P systems for generating control languages," *Information Sciences*, vol. 278, pp. 285–297, 2014.