Combining Structure-Based Features and Conserved Data for MicroRNA Target Prediction by the Neural Network Method

Shih-Yi Chao and Jung-Hsien Chiang
Department of Computer Science and Information Engineering
National Cheng Kung University, Taiwan
(telephone: 886-6-2757575, e-mail: chaosy@cad.csie.ncku.edu.tw)

ABSTRACT

Most of MicroRNAs are thought to control post-transcriptional mechanisms by base pairing with MicroRNA recognition elements found in their messenger RNA (mRNA) targets. A new computational method we provide is to predict mRNA targets for human MicroRNAs. Combined structure-based features and conserved data across species, the overall results are 89.1% in sensitivity. This means, the about 30nt short sequences of nucleotides can be accepted by our system not only when they appear in the interior loops and bulges, but also in the G:U, A:U and G:C nucleotide pairs of RNA secondary structures. We also provide the computationally deriving that guides single MicroRNA for multiple target mRNAs recognition. Incorporation of computational procedures allows prediction of human MicroRNA containing multiple target mRNAs. The results suggest that by providing structure-based features and conserved data can improve the performance of predicting MicroRNA target mRNAs.

1: INTRODUCTION

MicroRNAs (miRNAs) are short non-coding RNAs that regulate gene expression in animals and plants. MicroRNAs directly take part in post-transcriptional regulation either by arresting the translation of messenger RNAs (mRNAs) or by the cleavage of mRNAs. [1][2] MicroRNAs are defined as single-stranded RNAs of about 19~25 nucleotides in length generated from endogenous transcripts that can form local hairpin structures. The importance of MicroRNAs for development is highlighted by the fact that they comprise approximately 1 % of genes in animals, and are often highly conserved across a wide range of species. [3][4] Their biological meaning has become more and more important, however, how they recognize and regulate target genes remains less well understood. Furthermore, mutations in proteins require for MicroRNA functions or biogenesis impairing animal development. [5] To date, functions have been assigned to only a few of human MicroRNA genes. Prediction of MicroRNA targets in Human provides an alternative approach to assign biological functions. This has been very effective in plants, where MicroRNAs and target mRNAs are often nearly perfectly complementary. [6] On the other hand, in animals, functional duplexes can be more variable in structure, which means they contain only short complementary sequences, interrupted by gaps and mismatches. Moreover, specific rules for functional MicroRNA-mRNA pairing that captures all known functional targets have not been devised. This is the major problem for computing searching strategies, which apply different assumptions about how to ideally identify functional sites. As a result, the number of predicted targets varies considerably with only limited overlap in the top-ranking targets, indicating that these approaches might only capture subsets of real targets or may include a high number of background matches. [7][8].

Several computational approaches have been implemented for prediction of MicroRNAs and target mRNAs pairings using methods based on primary sequence conservation or secondary structure alignment. [9][10] However, it is particularly difficult to identify the functions of short MicroRAN target sequences by only performing primary sequences alignment. That is, it may cause high false-positive rate while alignment of short primary nucleotides sequences. Because MicroRNA target motifs are conserved more in structure than in primary sequences [11], the computational detections of RNA secondary structures is quite a challenging problem. The RNA secondary structures (hairpin structures), or called stem-loop, usually is a lollipop-shaped structure formed when a single-stranded nucleic acid molecule loops back on itself, to form a complementary double helix (stem) topped by a loop, which is shown in Figure 1(b).

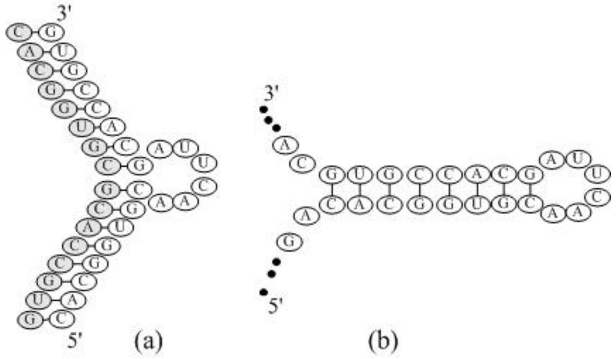


Figure 1. (a) General scheme of MicroRNA-mRNA pairing. The gray circles indicate the MicroRNA nucleotides, which are complementary to the sequences of the target mRNA (the white circles). (b) An example of the secondary structure alignment by RNAfold program.

The first contribution of this approach presented here is the combination of data from several sources. By integrating data from multiple species, we stabilize the learning process and construct a model that is more likely to be applicable to a variety of genomes. The second contribution is that, we not only consider the conserved primary sequences across species, but also take RNA secondary structures into account. That means, the about 30nt short sequences of nucleotides can be accepted by our system not only when they appear in the interior loop and bulges, but also in the G:U, A:U and G:C nucleotide pairs of RNA secondary structures. Therefore, we reduce the fault positive rates while predicting MicroRNA target mRNAs. The third contribution this approach provides is that, we computationally derive the guiding single MicroRNA for multiple target mRNAs recognition. Incorporation of computational procedures allows prediction of human MicroRNA containing multiple target mRNAs.

This article is organized as follows, in the beginning, we briefly introduce the materials and data sets that used by this approach. Secondarily, we give an introduction of predictions of RNA secondary structures. Next, we also explain the importance of searching homology genes and necessity of extracting structure-based features from RNA secondary structures. Finally, we exhibit experimental results and discuss the prediction results by some known biology experiments identified mRNA targets.

2: MATERIALS

2.1: The data set of experimentally defined MicroRNA – mRNA pairing (the training sets)

Some approaches, which are related to MicroRNA function studies, many MicroRNA target sites have been presented as putative ones based on the complementarity of sequences or other computational prediction methods without experimental verification of biological experiments. [12] However, these data may include some biological factual examples and non-factual (predicted) binding site sequences if using them as training sets. Hence, all the non-factual data sets should be excluded in order to improve the quality of training sets. As a result, we collect experimentally verified binding site sequences of mRNA sequences for several MicroRNAs over several animal species. We collect carefully MicroRNA-mRNA pairing sequences from the biological literatures [13][14][15][16][17]. Both *C.elegans* and *Drrosophila* MicroRNA sequences and their target mRNAs sequences are collected for the training sets.

2.2: Random sequences – the negative training sets

Random sequences are produced according to [9], which is the sampling of specified AUCG background frequencies of $P_A = 0.34$, $P_C = 0.19$, $P_G = 0.18$, and $P_U = 0.29$. These frequencies are consistent with the sequence composition of the *C. elegans* 3'UTRs of the target genes in the training set. Using randomly generated data as a negative training example is very dangerous because the random data might be quite different from actual data and might be hyper estimated. Therefore, we select negative training examples in the same criterion of the artificial (positive) training ones. That means, the random sequences are also predicted by RNAfold for producing the RNA secondary structures, so that we can also extract structure-based features from the negative training samples.

2.3: Using the RNAfold program to predict the RNA secondary structures

The RNA secondary structure includes external elements (non-paired bases), double strand or stacks (paired bases), a hairpin (a double strand and a loop), bulge loops, interior loops and multi-loops. RNA can also have tertiary interaction, such as pseudoknots [18]. In this approach, we focus on the secondary structures particularly. The RNAfold [19] programs, based on the rules of minimum free energy, are used for the RNA secondary structure predictions. Accordingly, we use RNAfold to generate the structures predictions, which is the prediction of RNA secondary structure between the target sequence in the mRNA 3'UTRs and the MicroRNAs. After predicting the RNA secondary structures, we are capable of extracting the structure-based features such as bulges, interior loops and so on.

2.4: Extraction of 3'UTR sequences in Human

Data are originally downloaded from NCBI web site at ftp://ftp.ncbi.nlm.nih.gov/gene/DATA. We extract 3'UTRs sequences, gene ids, gene official symbol names, gene alias names, accession numbers, and so on columns from the data set; then store in database for human species. This data is used for predicting the target mRNAs while giving a known human MicroRNA sequences. Since target mRNA sequences of MicroRNAs are conserved between species, we draw sequences from highly conserved regions, as defined by homology using the BLAST sequence alignment. [26] The Human sequences are pairwisely aligned to all the *C.elegans* and *Drrosophila* target mRNA sequences and all sequences identified homology to *C.elegans* and *Drrosophila* are annotated as "potential homologue" and stored in our database.

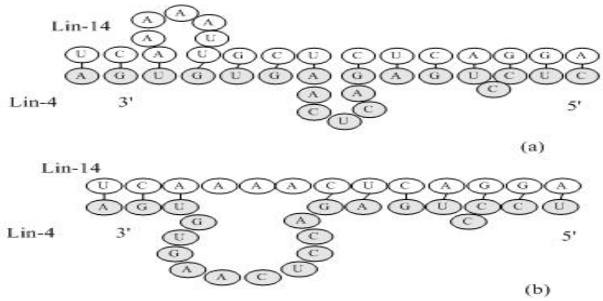


Figure 2. This figure demonstrate the MicroRNA lin-4 and its target mRNA (lin-14 3'UTR), which is modified from [13]. The major point that this figure manifests is to stand out the importance of structure-based features while predicting target mRNAs for MicroRNAs. With the same MicroRNAs nucleotide sequences, they have different shape of loops while recognizing target mRNAs. (a) The gray circles indicate the MicroRNA nucleotides, which are complementary to the sequences of the target lin-14 (the white circles). (b) Another binding target position in lin-14 3'UTR for MicroRNA lin-4, which forms another structural formations.

Table 1 The extracted features used by this approach

ID	The descriptions of features	Decoded value while inputting NN classifier
1.	AU match	1
2	GU match	2
3	GC match	3
4	mismatch	4
5	single nucleotide bulge	5
6	Non-single nucleotide bulge	6
7	gap	7
8	# of AU match at 5' part	#
9	# of AU match at 3' part	#
10	# of GU match at 5' part	#
11	# of GU match at 3' part	#
12	# of GC match at 5' part	#
13	# of GC match at 3' part	#
14	Total # of mismatch	#
15	Total # of gap	#
16	# of nucleotides within a non-single nucleotide bulge	#

3: METHODS

In this section, we briefly introduce the features that are extracted from known MicroRNA – mRNA pairings, and also make a description of how the Neural Network classifier is constructed and trained.

3.1: Extracted Features

According to [10], structure-based features are important features which show the shapes of loops or bulges and the mechanism of MicroRNA–mRNA pairing. A single nucleotide mutation could repress the MicroRNA function according to the various changing of nucleotides structures. Take Figure 2 as an example, the MicroRNA, lin-4, is confirmed by biological experiments that one of its target mRNAs is lin-14 3'UTR sequences. Lin-4 not only recognizes seven binding targets in gene lin-14 3'UTR, but also every type of bulge and loop in each recognized binding site varied [13]. In other words, structure-based features are somehow diverse while MicroRNA binding to different target mRNAs. We keep these biological characteristics while training our Neural Network classifier. For one input MicroRNA–mRNA pairing sample, each position has 7 conditions to select, which are consisted of AU match (indicate 1), GU match (indicate2), GC match (indicate 3), mismatch (indicate 4), single nucleotide bulge (indicate 5), non-single nucleotide bulge (indicate 6), and, finally, the gap (indicate 7). Furthermore, we also extract some quantity features, such as the number of AU match at 5' part, the number of AU match at 3' part, the number of GU match at 5' part, the number of GU match at 3'part, the number of GC match at 5' part, the number of GC match at 3'part, and the total number of mismatches. As listed in Table 1, the input values of ID 8 ~ 16 features are integers that depend on each one of the training samples. Hence, we use the symbol '#' to represent different values in training cases.

3.2: RBF Neural Network method as a learning algorithm

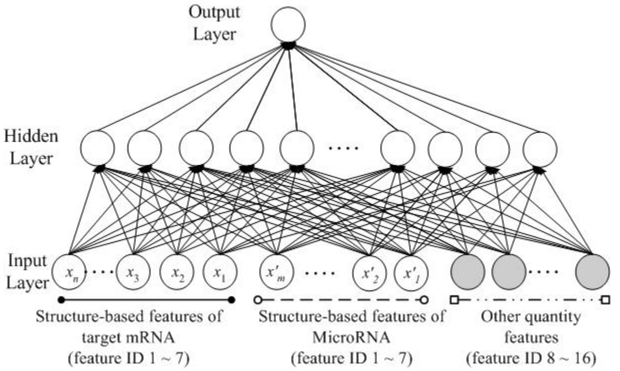
We use Radial Basis Function neural network (RBF) to learn the target mRNA discriminating rules from the training dataset. RBF networks have been extensively studied in the past [20] [21]. RBF consist of three layers, an input, a hidden and an output layer. The input layer corresponds to the input vector space and the output layer corresponds to the pattern classes. The whole architecture is consequently fixed by determining the hidden layer and the weights between the middle and the output layers. We demonstrate the RBF architecture used by this approach in Figure 3. Afterward, we briefly describe the input layer and hidden layer. Let X represents input layer vector, and G_i (where i =1, 2, …, n) represent neurons in hidden layer, which is a Gaussian kernel in the form:

$$G_i(X, \mu_i) = \exp(\frac{-1}{2\sigma_i^2} \| X - \mu_i \|^2) \quad (1)$$

where μ_i is a vector representing the center of the i kernel and σ_i^2 is the corresponding variance. The output layer implements a weighted sum of hidden-output units:

$$F(X) = \sum_{i=1}^{m} w_i G(X, \mu_i) - \theta \quad \text{for } i = 1,2,\ldots, m \quad (2)$$

where w_i is the output weight, and each corresponds individually to the connection between a hidden neuron and an output neuron. Later we use gradient descent to determine the weights of the network. At last, the vector θ represents biases. Additionally, the outputs of RBF are "labels represent true or false" that describe whether the input mRNAs sequences are binding targets or not by given a MicroRNA.



Output Layer
Hidden Layer
Input Layer
x_n
x_3
x_2
x_1
x'_m
x'_2
x'_1
Structure-based features of target mRNA (feature ID 1 ~ 7)
Structure-based features of MicroRNA (feature ID 1 ~ 7)
Other quantity features (feature ID 8 ~ 16)

Figure 3. The RBF architecture used by this approach. There are three kinds of input neurons, one is structure-based features of input MicroRNA, another is structure-based features of input target mRNA, and the other is quantity features.

As Figure 3 illustrated, there are three kinds of input neurons. The first kind of neuron represents the structure-based features of input MicroRNA sequences, identified by dotted line, started from 5' to 3' part nucleotides (indicated by *x'*$_i$, where *i* = 1 to *m*) . Each position (nucleotide) has 7 conditions to select to represent the pairing status. The second kind of input neuron represents the structure-based features of input target mRNA sequences, start from 3' to 5' part, which are indicated by x_j, (where *j* = 1 to *n*). The last kind of neuron represents the quantity features that described in Table 1 (ID 8 ~ 16), respectively.

Take Figure 2 (b) as an example, the input vector for MicroRNA Lin-4 is"1 3 3 5 1 3 1 3 6 6 6 6 6 6 6 6 6 6 3 1 3 1", which is started form the 5' part. The persisted "6" indicates the structure of non-single nucleotide bulge, and the element 5 represents the single nucleotide bulge in Lin-4. As for the target mRNA, Lin-14, the input vector is "3 1 3 7 1 3 1 3 4 4 4 7 7 7 7 7 7 7 3 1 3", from 3' to 5' part, and the persisted "7" indicates gap positions. The remained input neurons are fed by the values of feature ID 8 ~ 16. The number of hidden neurons is the same as the number of input neurons. There is only one neuron in output layer, which outputs the value 0 (false) or 1 (true).

3.3: Searching multiple target mRNAs for single MicroRNA in Human 3'UTRs

Most of the targets identified by [13][14][15][16] contain multiple target mRNAs for the same MicroRNA or are regulated by more than one MicroRNA. The targets reported for *Drosophila* MicroRNAs also contain multiple target mRNAs. [22] However, the searching procedures guiding single MicroRNA for multiple target mRNAs interactions have not been investigated. Therefore, predictions of MicroRNA targets containing multiple mRNAs are lacking. Here we describe computationally procedures that guide single MicroRNA for multiple target mRNAs recognition. Incorporation of homology sequences in computational procedures allows prediction of human MicroRNA containing multiple target mRNAs. One of the target mRNAs for MicroRNA let-7b is found in the 3'UTR of the human mRNA that code for the human homolog of the *C. elegans* LIN-28 protein, a putative RNA-binding protein [23]. Thus, we collect human 3'UTRs sequences that homologize the *C.elegans* and *Drosophila* target mRNAs sequences. With these homologous sequences, given a known MicroRNA, the trained RBF classifier is able to decide the input sequences are whether one of the target mRNAs of the given MicroRNA or not.

3.4: System Procedures

We illustrate the prediction procedures in Figure 4. As we described in section 2.4, we have the collected human 3'UTRs sequences that are homolog of the *C.elegans* and *Drosophila* target mRANs sequences, and input to the prediction procedures one time. Given a known MicroRNA sequences, if there are more than three nucleotide pairs, such as AU, GU, GC matches and can be tolerant of single nucleotide bulge interrupted, then we call the RBF classifier to judge whether the mRNA is one of the targets or not. If there are less than three nucleotide pairs in the previous step, then we shift the position with three nucleotides length to check if any nucleotide pairs exist. This sub-procedure will not stop until the end of the mRNA sequences. In other words, given one known MicroRNA and one mRNA sequence, the system may predict more than one binding sites within the same mRNA in different positions. We derive computational procedures that guide single MicroRNA for multiple target recognition positions within one mRNA. Besides, we also consider the rules that guide single MicroRNA for multiple target mRNAs interactions. For instance, if there are *n* known MicroRNAs, *m* mRNA sequences, and *k* recognized positions in each mRNA, the system will run for *n* × *m* × *k* times.

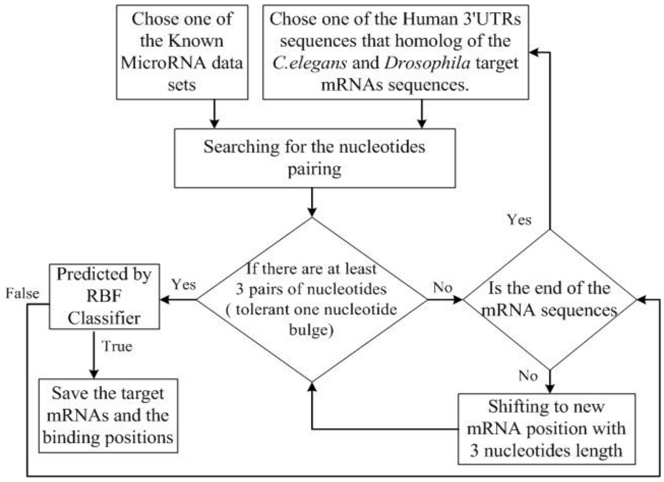

Chose one of the Known MicroRNA data sets
Chose one of the Human 3'UTRs sequences that homolog of the *C.elegans* and *Drosophila* target mRNAs sequences.
Searching for the nucleotides pairing
If there are at least 3 pairs of nucleotides (tolerant one nucleotide bulge)
Yes
No
Is the end of the mRNA sequences
Yes
No
Predicted by RBF Classifier
False
True
Save the target mRNAs and the binding positions
Shifting to new mRNA position with 3 nucleotides length

Figure 4. The computational procedures for prediction of the target mRNAs while giving MicroRNA.

4: EXPERIMENTAL RESULTS

We investigate the targets discovery in 3'UTR regions first. Using the approach we provided in previous section, we discover several conserved mRNA targets while giving a known MicroRNA. Next we discuss the performance of the RBF classifier, and the effects of bulges, interior loops or G:U pairings.

4.1: The predicted targets of Human 3'UTRs

Because the target of 3'UTRs for MicroRNAs have not been well studied, we can not compare the discovered targets to the known targets that confirmed by biological experiments. Therefore, we provide GO [27] terms for genes that are targets predicted by our system. Gene Ontologies were assigned to target human mRNA genes according to the NCBI database. As we can see in Table 2, most target mRNAs for MicroRNA let-7b and let-7e are annotated as "protein binding". Moreover, it is biologically confirmed that MicroRNA target genes contain binding sites with G:U base pairs or single nucleotide bulges [24][25]. As a result, it is reasonable to train the RBF classifier by structure-based features. It is also interesting that, the predicted target genes such as PARVB, GIPC1, ARMC4, and FANCD2, are identified as two binding positions that MicroRNA let-7b may recognize. The results prove that performing computational procedures to guide single MicroRNA for multiple target mRNAs recognition is practicable.

4.1: Performance of the RBF classifier

The 10-fold cross-validation is known to create a good estimate of the predictive accuracy of classification methods. In this approach, we use 10-fold cross-validation for accuracy estimation. Figure 5 shows the performance of the RBF classifier according to the training dataset. The performance is represented in statistical measures: sensitivity and specificity. Where,

The Sensitivity = # of True Positives / (# of True Positives + # of False Positives).

The Specificity = # of True Negatives / (# of True Negatives + # of False Positives).

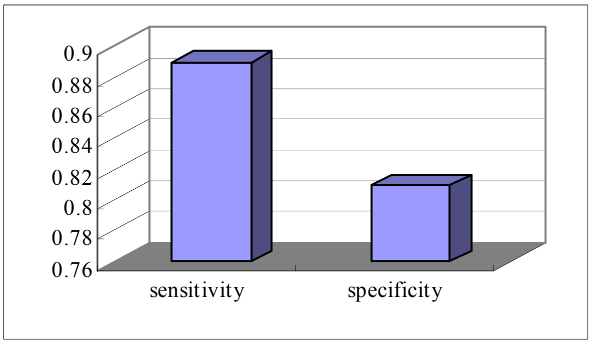


Figure5. The performance of the RBF classifier according to the training dataset

5: CONCLUSIONS

In this approach, we propose a RBF classifier to predict MicroRNA target mRNA with structure-based features and conserved sequences across species. We input extracted structure-based features such as single nucleotide bulges, G:U, A:U, and G:C pairings, interior loops (the non-single nucleotide bulges) to represent the functional mechanism of the MicroRNA-mRNA pairing. A RBF classifier, inputted by structure-based features, performs well on training data and predicting results. In other words, the about 30nt short sequences of nucleotides can be accepted by our system not only when they appear in the interior loop, bulges, but also in the G:U, A:U or G:C pairs of RNA secondary structures. Therefore, we reduce the fault positive rates while predicting MicroRNA target mRNAs. Also, we computationally derive that how to guide single MicroRNA for multiple target mRNAs recognition. Incorporation of computational procedures allows prediction of human MicroRNA containing multiple target mRNAs. The results provided by our system, imply a sample step towards a comprehensive inventory of human mRNA 3'UTRs targets that play a major role of understanding post-transcriptional mechanism, or cellular mechanism in disease and health. With the analysis of focusing on the conserved targets of genes, it should be a possibility of predicting a more complete catalogues of mRNA targets.

REFERENCES

[1]V. Ambros, "The functions of animal microRNAs", Nature, vol. 431, pp. 350-355, 2004.

[2]J. C. Carrington, and V. Ambros, "Role of microRNAs in plant and animal development", Science, vol. 301, pp. 336-338, 2003.

[3]A.E. Pasquinelli, B.J. Reinhart, F. Slack, M.Q. Martindale, M.I. Kuroda, et al., "Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA", Nature, vol. 408, pp. 86-89, 2000.

[4]L.P. Lim, N.C. Lau, E.G. Weinstein, A. Abdelhakim, S. Yekta, et al., "The MicroRNAS of Caenorhabditis elegans", Genes and Development, vol. 17, pp. 991-1008, 2003.

[5]A. Grishok, A. E. Pasquinelli, D. Conte, N. Li, S. Parrish, et al., “Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing”, Cell, vol.106, pp. 23-34, 2001.

[6]M.W. Rhoades, B.J. Reinhart, L.P. Lim, Burge, C.B. Bartel, et al., “Prediction of plant microRNA targets”, Cell, vol.110, pp.513-520, 2002.

[7]B.P. Lewis, I.H. Shih, M.W. Jones-Rhoades, D.P. Bartel, C.B. Burge, “Prediction of mammalian MicroRNA targets”, Cell, vol.115, pp.787-797, 2003.

[8]A.J. Enright, J. Bino, U. Gaul, T. Tuschl, C. Sander, et al., “MicroRNA targets in Drosophila”, Genome Biol, vol.5, 2005.

[9]N. Rajewsky and N. D. Socci, “Computational identification of MicroRNA targets”, Developmental Biology, vol. 267, pp. 529-535, 2004.

[10]S. K. Kim, J. W. Nam, W. J. Lee, and B.T. Zhang, “A Kernel Method for MicroRNA Target Prediction Using Sensible Data and Position-Based Features”, IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB 2005), pp. 46-52, 2005.

[11]E. R. Sean, “Non-coding RNA genes and the modern RNA world”, Nauret Reviews Genetics, vol. 2, pp. 919-929, 2001.

[12]S. Griffiths-Jones, J.G. Russell, S. V. Dongen, A. Bateman, A. J. Enright, “miRBase: microRNA sequences, targets and gene nomenclature”, Nucleic Acids Research, Database Issue, pp. D140-D144, 2006.

[13]D. Banerjee, and F. Slack, “Control of developmental timing by small temporal RNAs: a paradigm for RNA-mediated regulation of gene expression”, BioEssays, vol.24, pp. 119-129, 2002.

[14]Lin, S.Y., Johnson, S.M., Abraham, M., Vella, M.C., Pasquinclli, A., Gamberi, C., Gottlieb, E., Slack, F.J., “The C.elegans hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target”, Dev. Cell, vol. 4, pp. 639-650.

[15]M.C. Vella, K. Reinert, and F.J. Slack, “Architecture of a validated microRNA::target interaction”, Chem. Biol., vol. 11, pp. 1619-1623, 2004.

[16]A. Stark, J. Brennecke, R.B. Russell, and S. M. Cohen, “Identification of Drosophila MicroRNA targets”, PLoSBiol., vol. 1, pp.60, 2003.

[17]J.Brennecke, A. Stark, R.B. Russell, and S.M. Cohen, “Principles of microRNA-target recognition”, PLoSBiol., vol. 3, pp. 85, 2005.

[18]G. Pavesi, G. Mauri, and G. Pesole, “Predicting Conserved Hairpin Motifs in Unaligned RNA Sequences”, in Proceedings: 15th IEEE International Conference on Tools with artificial Intelligence, 2003.

[19]I. L. Hofacker et al., “Vienna RNA secondary structure server“, Nucleic Acids Research, vol. 31, pp. 3429-3431, 2003.

[20]J. Moody and C.J. Darken "Fast Learning in Networks of Locally-Tuned Processing Units" *Neural Computation 1*, pp. 281-294, 1989.

[21]T. Poggio and F. Girosi "Networks for Approximation and Learning" *Proceedings of the IEEE*, Vol. 78, pp. 1481-1497, 1990.

[22]A. J. Enright, B. John, U. Gaul, T. Tuschl, C. Sander and D.S. Marks, “MicroRNA targets in Drosophila”, Genome Biol., vol. 5, 2003.

[23]E.G. Moss and L. Tand, “Conservation of the hetero-chronic regulator Lin-28, its developmental expression and MicroRNA complementary sites”, Dev Biol., vol. 258, pp. 432-442, 2003.

[24]B. J. Reinhart, F.J. Slack, M. Basson, A.E. Pasquinelli, J.C. Bettinger, et al., “The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*”, Nature, vol. 40.3, pp. 901-906.

[25]J. Brennecke, D.R. Hipfner, A. Stack R.B. Russell, S.M. Cohen, “Bantam encodes a developmentally regulated MicroRNA that controls cell proliferation and regulates the pro-apoptotic gene hid in *Drosophila*”, Cell, vol. 113, pp. 25-36.

[26]Kent, W.J., “The BLAST-Like alignment tool”, *Genome Res.*, vol.12, pp. 996-1006, 2002.

[27] Gene Ontology Consortium, 1999-2006. http://www.geneontology.org

Table 2. The predicted target mRNA

GENE NAME	Accession NUM	GO terms	Homology	MicroRNA
PARVB	NM_013327	✓ protein binding ✓ cytoskeleton ✓ cell adhesion	C. elegans	Let-7b
GIPC1	NM_005716	✓ receptor binding ✓ protein binding ✓ membrane fraction ✓ soluble fraction ✓ cytosol G-protein coupled ✓ receptor protein signaling ✓ pathway membrane	C. elegans	Let-7b
ARMC4	NM_018076	N/A	Drosophila	Let-7b
FANCD2	NM_001018115	✓ protein binding ✓ Nucleus ✓ Chromosome ✓ DNA repair ✓ Cell cycle	Drosophila	Let-7b
TDO2	NM_005651	✓ Function iron ion binding ✓ tryptophan metabolism ✓ oxidoreductase activity neurotransmitter metabolism ✓ metal ion binding	C. elegans	Let-7b
GABBR1	NM_001470	✓ gamma-aminobutyric acid signaling pathway ✓ negative regulation of adenylate cyclase activity ✓ osteoblast differentiation ✓ cytoplasm ✓ integral to plasma membrane ✓ GABA-B receptor activity	C. elegans	Let-7e
RTKN	NM_033046	✓ nucleotide binding ✓ GTPase inhibitor activity ✓ protein binding ✓ GTP binding ✓ intracellular ✓ apoptosis ✓ signal transduction ✓ Rho protein signal transduction ✓ GTP-Rho binding ✓ regulation of anti-apoptosis	C. elegans	Let-7e
BECN1	NM_003766	✓ autophagy ✓ anti-apoptosis ✓ cellular defense response ✓ response to virus	C. elegans	miR-23b