Published online 19 November 2007 *Nucleic Acids Research, 2008, Vol. 36, Database issue* ***D165–D169***
doi:10.1093/nar/gkm1012

miRNAMap 2.0: genomic maps of microRNAs in metazoan genomes

Sheng-Da Hsu[1], Chia-Huei Chu[1], Ann-Ping Tsou[2], Shu-Jen Chen[3], Hua-Chien Chen[3], Paul Wei-Che Hsu[1], Yung-Hao Wong[1], Yi-Hsuan Chen[3], Gian-Hung Chen[3] and Hsien-Da Huang[1,4,5,*]

[1]Institute of Bioinformatics, National Chiao Tung University, Hsin-Chu 300, [2]Institute of Biotechnology in Medicine, National Yang-Ming University, Taipei 112, [3]Molecular Medicine Research Center, Chang Gung University, Tao-Yuan 333, [4]Department of Biological Science and Technology, National Chiao Tung University, Hsin-Chu 300 and [5]Core Facility for Structural Bioinformatics, National Chiao Tung University, Hsin-Chu 300, Taiwan

Received September 23, 2007; Revised October 22, 2007; Accepted October 24, 2007

ABSTRACT

MicroRNAs (miRNAs) are small non-coding RNA molecules that can negatively regulate gene expression and thus control numerous cellular mechanisms. This work develops a resource, miRNAMap 2.0, for collecting experimentally verified microRNAs and experimentally verified miRNA target genes in human, mouse, rat and other metazoan genomes. Three computational tools, miRanda, RNAhybrid and TargetScan, were employed to identify miRNA targets in 3′-UTR of genes as well as the known miRNA targets. Various criteria for filtering the putative miRNA targets are applied to reduce the false positive prediction rate of miRNA target sites. Additionally, miRNA expression profiles can provide valuable clues on the characteristics of miRNAs, including tissue specificity and differential expression in cancer/normal cell. Therefore, quantitative polymerase chain reaction experiments were performed to monitor the expression profiles of 224 human miRNAs in 18 major normal tissues in human. The negative correlation between the miRNA expression profile and the expression profiles of its target genes typically helps to elucidate the regulatory functions of the miRNA. The interface is also redesigned and enhanced. The miRNAMap 2.0 is now available at http://miRNAMap.mbc.nctu.edu.tw/.

INTRODUCTION

MicroRNAs (miRNAs) are small non-coding RNA molecules that can negatively regulate gene expression by hybridizing to the 3′-untranslated regions (3′-UTR) of the target gene. miRNAs are important in cell development, cell death, cell proliferation, fat metabolism, hematopoiesis and nervous system patterning in animals, as well as stress responses, and leaf and flower development in plants (1–4).

Numerous miRNAs and miRNA targets have been discovered and experimentally confirmed over the last few years. miRBase (5), which is the most comprehensive database of experimentally validated miRNAs across many genomes, provides integrated interfaces for presenting information on miRNA and computationally predicted miRNA targets. DIANA TarBase (6) collects experimentally validated miRNA targets in eight species. It contains a total of 750 miRNA target sites in 550 target genes. miRGen (7) collects positional relationships between miRNAs and genomic sets and miRNA targets according to combinations of widely used target prediction programs. Recently, many biologists have been paying much more attention to the functions of miRNAs in biological systems. Several miRNA target prediction tools were have been developed, such as miRanda (8), TargetScan (9) and RNAhybrid (10), for determining the energetically favored hybridization sites of small RNA to large RNAs. Lu *et al.* (11) developed an miRNA microarray to measure the expression profiles of all known miRNA in various normal tissues and tumors.

This work develops a resource, miRNAMap 2.0, to collect experimentally verified microRNAs and miRNA target genes in human, mouse, rat and other metazoan genomes. Three computational tools—miRanda, RNAhybrid and TargetScan—were utilized to identify miRNA targets in 3′-UTR of genes as well as known miRNA targets. Three criteria are applied for filtering the putative miRNA target sites to retain the more probable

*To whom correspondence should be addressed: Tel: +886 3 5712121, Ext. 56952; Fax: 886 3 5729288; Email: bryan@mail.nctu.edu.tw
Present address:
Hsien-Da, Huang, Ph.D. Department of Biological Science and Technology, Institute of Bioinformatics, National Chiao Tung University, Hsin-Chu 300, Taiwan, ROC

Table 1. Enhancements and new features of miRNAMap 2.0

Features	miRNAMap 1.0	miRNAMap 2.0
Known miRNAs	MiRBase (version 6.0)	MiRBase (version 9.2)
Supported species	Human, mouse, rat and dog	Two insects, nine vertebrates and one worm
Experimental miRNA targets	Taken from the literature	Taken from TarBase (6) and the literature
miRNA expression profiling	Lu *et al*. (11): miRNA profiling in human	Lu *et al*. (11): miRNA profiling in human Q-PCR miRNA profiling in human
Expression profiles of miRNA targets	–	NCBI-GEO-GDS596 (76 human tissues) (18)
miRNA target prediction tools	miRanda	miRanda, RNAhybrid and TargetScan
Criteria for filtering predicted miRNA targets	–	Criterion 1: *predicted by at least two tools* Criterion 2: *target genes contain multiple sites* Criterion 3: *target site is accessible*
Accessible region of miRNA target sites	–	Sfold (17)
Tissue specificity of human miRNAs	–	Q-PCR miRNA profiling (18 human tissues)

miRNA target sites to reduce the rate of false positive predictions of miRNA target sites. In particular, the RNA accessibilities of the identified miRNA target site were examined, providing information on understanding the miRNA/target relationship.

The miRNA expression profiles offer valuable clues on the properties of the miRNAs, such as tissue specificity and differential expression in cancer/normal cell. Accordingly, (quantitative polymerase chain reaction) experiments were conducted to monitor the expression profiles of 224 human miRNAs in 18 major normal tissues in humans. The negative correlation between the miRNA expression profiles and the expression profiles of their target genes helps to elucidate the regulatory functions of the miRNA. Finally, both textual and graphical web interfaces were redesigned and enhanced to facilitate the retrieval of data from the miRNAMap.

The main contribution of this work is the extension of miRNAMap version 1.0. The new version supports more genomes of known miRNAs and known miRNA targets. Prediction of miRNA genes is eliminated to examine miRNA/target relationship. Three criteria were applied to reduce the rate of false positive prediction of miRNA targets. The analysis of the correlation between the miRNA expression profile and the expression profiles of the target genes is useful to evaluate the possibility of the miRNA/miRNA targets relationship.

IMPROVEMENTS

Table 1 presents the major differences between the previous version and miRNAMap 2.0. The major enhancements and new features of miRNAMap 2.0 are described subsequently. miRNAMap 2.0 collects the known miRNAs in metazoan genomes, including two insects, nine vertebrates and one worm. miRNAMap 2.0 supports, apart from human, mouse, rat and dog, other metazoan genomes, others such as chicken, fruit fly, worm, zebrafish, fugu, frog (*Xenopus tropicalis*), malarial mosquito and opossum. Twelve genomes contain 2241 known miRNAs, which were obtained from miRBase (release 9.2, May 2007). The experimentally verified miRNA targets were obtained from DIANA TarBase and by surveying the literature. The numbers of experimental miRNA targets extracted from the DIANA TarBase and by surveying the literature are 346 and 29, respectively. Table 2 presents statistics for each genome.

Table 2. The data statistics

Species	Number of known miRNAs	Number of target sites	Number of miRNA target sites after filtering		
			Criterion 1	Criterion 2	Criterion 3
C. elegans	132	5751	378	31	2771
Mosquito	38	3168	122	14	1240
Fruit fly	78	10 180	0	101	5088
Human	475	6750	1717	781	4282
Mouse	377	6763	1266	438	3681
Dog	6	7017	418	106	3404
Rat	234	5087	865	181	2539
Chicken	149	1021	140	9	474
Zebrafish	337	1393	155	32	685
Opossum	107	2616	204	34	1161
Fugu	131	246	7	0	15
Frog	177	2646	658	91	1334
Total	2241	52 638	5930	1818	26 674

Identification of miRNA target

This version of miRNAMap incorporates three previously developed computational tools, such as miRanda (8), TargetScan (9) and RNAhybrid (10), to identify miRNA target sites within the conserved regions of 3′-UTR of genes in 12 metazoan genomes. The conserved regions were extracted from the UCSC Genome Browser Most Conserved Regions (12). The MFE threshold of the miRNA and target duplex was −12 kcal/mol and the miRanda score was specified as 120. Therefore, the miRNA targets whose MFEs are smaller than −12 kcal/mol and whose miRanda score exceeds 120 are identified and compiled in the miRNAMap database. The predictive parameters of TargetScan and RNAhybrid were set as default values. Each miRNA target prediction tool yields a set of candidate miRNA target sites. However, some of these candidates may be false positive predictions. This work presents three criteria for eliminating false positives and retaining better candidate miRNA target sites.

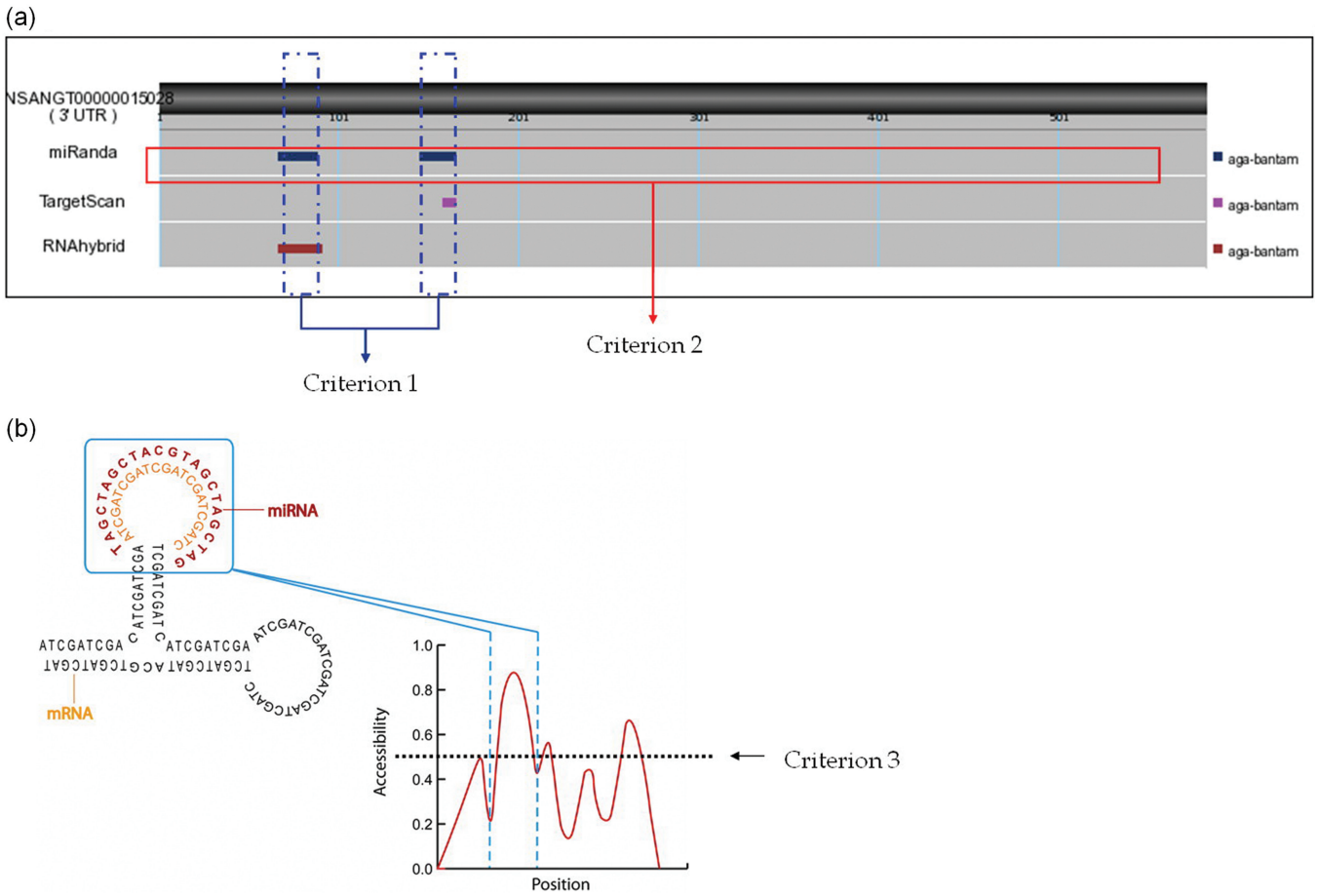

Figure 1. Criteria for identifying miRNA targets in miRNAMap. (**a**) Criterion 1 selects potential miRNA target sites, which are predicted using at least two tools; criterion 2 selects the target gene that contains multiple target sites; (**b**) Criterion 3 selects the accessible miRNA target sites (15).

Criterion 1: target sites must be predicted by at least two tools. The three tools, miRanda, RNAhybrid and TargetScan, were applied separately to identify miRNA targets within the conserved regions of 3′-UTR in all metazoan genomes. This criterion retains putative miRNA targets that have been predicted by at least two tools for a miRNA, as displayed in Figure 1(a).

Criterion 2: target gene contains multiple target sites. Previous investigations have suggested that one gene can contain several miRNA target sites, bound by multiple distinct miRNAs or single miRNA. For example, six let-7 miRNA target sites were discovered in lin-4 (13–15), and miR-33, miR-124, miR-277 and miR-312 target the eye development gene seven-up (svp) in *Drosophila* (14). Hence, this criterion retains the miRNA target sites and the corresponding gene, which contains multiple target sites, as shown in Figure 1(a).

Criterion 3: target site must be located in accessible regions. The conventional target prediction tools exploit the complementarity between the miRNA and its target sequence, the conservation of the target sites, and the kinetics and thermodynamics of the miRNA/target duplex. Although these properties importantly determine the miRNA target sites, the sequence context that surrounds the miRNA target sites influences the binding affinities of the miRNA/target duplex. Robins *et al*. (16) hypothesized that single-strand miRNAs can only bind to stretches of free mRNA for potential target sites. Long *et al*. (15) posited the accessible model of miRNA target sites for predicting miRNA targets and successfully interpreted the published data on the *in vivo* activity of C. *elegans* reporter genes that contain modified lin-41 3′-UTR sequences.

This work exploits the aforementioned concept to filter out the false positive predictions, such that miRNAs hybridize to the target sites that are located in the accessible regions, are more likely to be real, as presented in Figure 1(b). The accessibility of RNA sequences is determined by Sfold (17).

Expression profiling of microRNAs and target genes

As stated above, the computational tools for identifying miRNA target sites are developed based on the complementarities of miRNA and target sites, and of the kinetics and the thermodynamics of miRNA/target duplex, the combinatorial properties of miRNA target sites, and the accessibility of the target sites. Additionally, the expression profiles of miRNAs are useful in elucidating the roles of miRNAs in complex complicated biological systems.

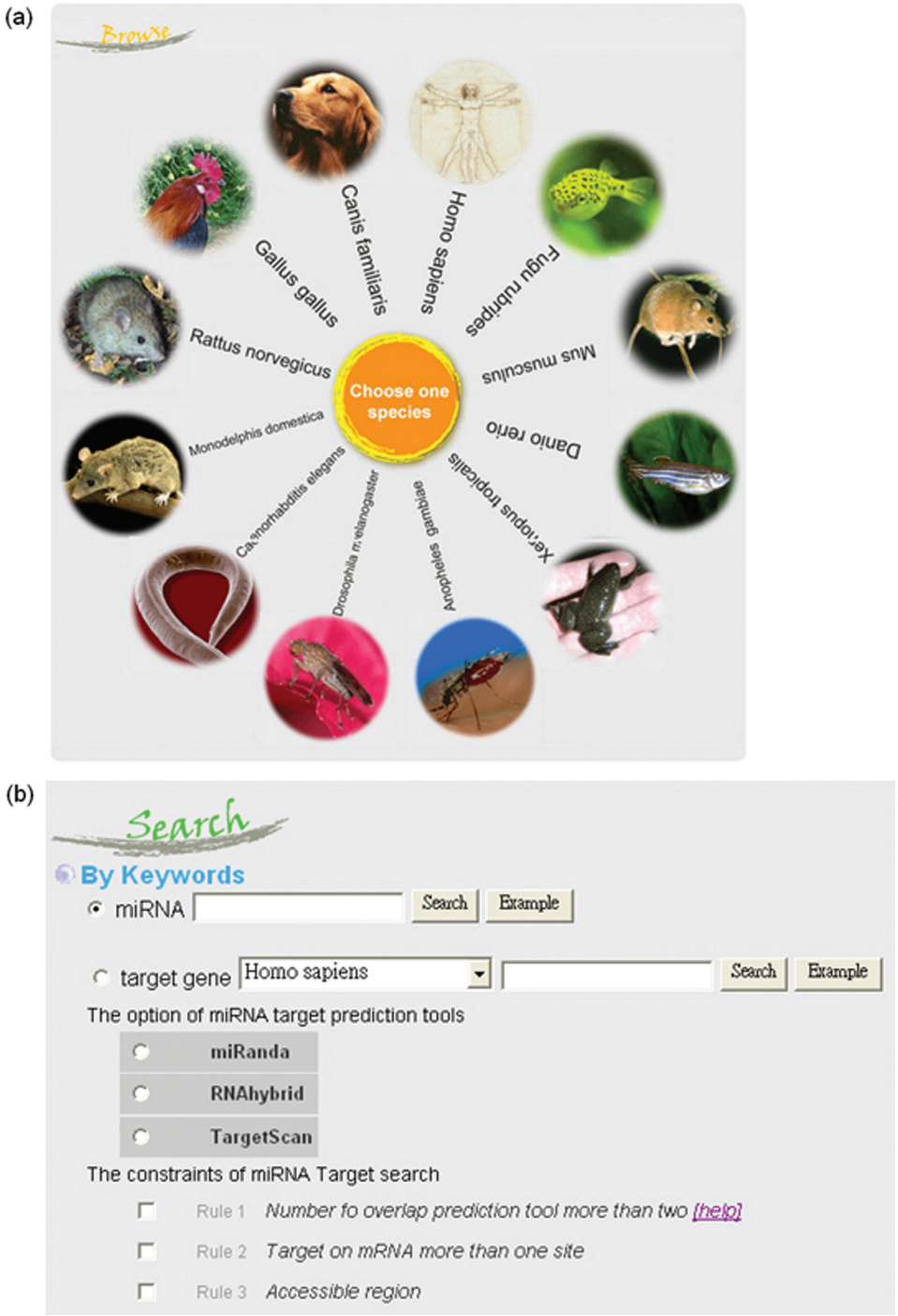

Figure 2. The miRNAMap 2.0 enhanced web interface.

In this work, two data sets of miRNA expression profiles, which were obtained by different experimental methods, Q-PCR and miRNA-bead array (11), are integrated.

The expression level of 224 human miRNAs in 18 major normal tissues in humans was detected by using a real-time PCR-based 220-plex miRNA expression profiling method to determine the tissue-specificity of human miRNAs. The detailed experimental protocol is described in Supplementary Material available online The expression levels of miRNAs are currently provided in the miRNAMap to examine the tissue-specificity of human miRNAs. Another data set was generated by Lu *et al.* (11), who employed a bead-based flow cytometric miRNA expression profiling method to present a systematic expression analysis of 217 mammalian miRNAs from 334 human samples. GDS596 (GEO accession), which is the gene expression profiles of coding genes in 79 human tissues (18), was obtained from NCBI GEO (19).

The expression profile of miRNA and the expression profiles of its target genes are typically negatively correlated since the miRNA downregulates its target gene. For each miRNA target site that is associated with a miRNA and a target gene (coding gene), the Pearson correlation coefficient is computed from the miRNA expression profile and the target gene expression profile, to elucidate the described phenomenon in experimental expression data in humans. Thirteen overlapping human tissues exist between the Q-PCR data set of the miRNA expression profiles and the GDS596 data set of the expression profiles of the target genes.

ENHANCED INTERFACE

Various query interfaces and graphical visualization pages were re-designed and implemented to facilitate access to data and further analyses to support research on miRNA. As presented in Figure 2, the miRNAMap has two modes for browsing miRNA information—the miRNA genes browser and the miRNA target browser. Search functions allow advanced users to access data concerning miRNA and its target genes.

The interface supports two new features, which are the filtering of putative miRNA target sites by three computational tools and the correlation analysis between miRNA expression profiles and the expression profiles of target genes. Users can select any one of the three criteria to filter the putative miRNA target sites in the database. The miRNA expression profiles in the interface reveal the tissue-specificity of the miRNA expression. In particular, the correlation between the miRNA expression profiles and the expression profile of the target genes of 13 major tissues in humans can help users to select miRNA and its target genes for experimental confirmation.

CONCLUSION

miRNAMap 2.0 is an integrated resource for explicating regulatory functions of miRNAs. Various criteria were developed herein for filter out predicted miRNA target sites and elucidating the negative correlation between miRNA expression and its target gene expression, especially in humans, helping to elucidate the miRNA/target relationship.

Future works should evaluate the effectiveness of the filtering criteria using experimental miRNA targets. The disease/cancer-related expression profiles in humans will be considered and integrated into the database. The miRNA expression profiles and the expression profiles of the miRNA target genes in other genomes, such mouse and rat, will be incorporated into the resource.

AVAILABILITY

The miRNAMap database will be continuously maintained and updated. The database is now freely available at http://miRNAMap.mbc.nctu.edu.tw/.

ACKNOWLEDGEMENTS

The authors would like to thank the National Science Council of the Republic of China for financially supporting this research under Contract No. NSC 96-3112-B-009-002. Special thanks for the financially supports from National Research Program For Genomic Medicine (NRPGM), Taiwan. This work was also partially supported by MOE ATU. Funding to pay the Open Access publication charges for this article was provided by National Science Council of the Republic of China.

Conflict of interest statement. None declared.

REFERENCES

1. Kim,V.N. and Nam,J.W. (2006) Genomics of microRNA. *Trends Genet*., **22**, 165–173.
2. Wienholds,E. and Plasterk,R.H. (2005) MicroRNA function in animal development. *FEBS Lett*., **579**, 5911–5922.
3. Bartel,D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.
4. Jover-Gil,S., Candela,H. and Ponce,MR. (2005) Plant microRNAs and development. *Int. J. Dev. Biol*., **49**, 733–744.
5. Griffiths-Jones,S., Grocock,R.J., van Dongen,S., Bateman,A. and Enright,A.J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res*., **34**, D140–D144.
6. Sethupathy,P., Corda,B. and Hatzigeorgiou,AG. (2006) TarBase: A comprehensive database of experimentally supported animal microRNA targets. *RNA*, **12**, 192–197.
7. Megraw,M., Sethupathy,P., Corda,B. and Hatzigeorgiou,A.G. (2007) miRGen: a database for the study of animal microRNA genomic organization and function. *Nucleic Acids Res*., **35**, D149–D155.
8. John,B., Enright,AJ., Aravin,A., Tuschl,T., Sander,C. and Marks,D.S. (2004) Human MicroRNA targets. *PLoS Biol*., **2**, e363.
9. Lewis,B.P., Shih,I.H., Jones-Rhoades,M.W., Bartel,D.P. and Burge,C.B. (2003) Prediction of mammalian microRNA targets. *Cell*, **115**, 787–798.
10. Kruger,J. and Rehmsmeier,M. (2006) RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res*., **34**, W451–W454.
11. Lu,J., Getz,G., Miska,E.A., Alvarez-Saavedra,E., Lamb,J., Peck,D., Sweet-Cordero,A., Ebert,B.L., Mark,R.H. *et al*. (2005) MicroRNA expression profiles classify human cancers. *Nature*, **435**, 834–838.
12. Kuhn,R.M., Karolchik,D., Zweig,AS., Trumbower,H., Thomas,D.J., Thakkapallayil,A., Sugnet,C.W., Stanke,M., Smith,K.E. *et al*. (2007) The UCSC genome browser database: update 2007. *Nucleic Acids Res*., **35**, D668–D673.
13. Vella,M.C., Choi,E.Y., Lin,S.Y., Reinert,K. and Slack,F.J. (2004) The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3′UTR. *Genes Dev*., **18**, 132–137.
14. Enright,A.J., John,B., Gaul,U., Tuschl,T., Sander,C. and Marks,D.S. (2003) MicroRNA targets in *Drosophila*. *Genome Biol*., **5**, R1.
15. Long,D., Lee,R., Williams,P., Chan,C.Y., Ambros,V. and Ding,Y. (2007) Potent effect of target structure on microRNA function. *Nat. Struct. Mol. Biol*., **14**, 287–294.
16. Robins,H., Li,Y. and Padgett,R.W. (2005) Incorporating structure to predict microRNA targets. *Proc. Natl Acad. Sci. USA*, **102**, 4006–4009.
17. Ding,Y. and Lawrence,C.E. (2003) A statistical sampling algorithm for RNA secondary structure prediction. *Nucleic Acids Res*., **31**, 7280–7301.
18. Su,A.I., Wiltshire,T., Batalov,S., Lapp,H., Ching,K.A., Block,D., Zhang,J., Soden,R., Hayakawa,M. *et al*. (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl Acad. Sci. USA*, **101**, 6062–6067.
19. Barrett,T. and Edgar,R. (2006) Gene expression omnibus: microarray data storage, submission, retrieval, and analysis. *Methods Enzymol*., **411**, 352–369.