Life Science Informatics Publications
Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences
Journal Home page http://www.rjlbpcs.com/
RJLBPCS
ISSN 2454-6348

Original Research Article **DOI - 10.26479/2016.0202.02**

MICRORNA TARGET GENE PREDICTION AND VALIDATION IN DIFFERENT GENOMES

Shibsankar Das[1*], Debabrata Mandal[2]

[1]Department of Mathematics, Uluberia College, Uluberia, West Bengal, India-711315

[2]Department of Computer Science, Tamralipta Mahavidyalaya, Tamluk, West Bengal, India- 721636

ABSTRACT : MicroRNAs (miRNAs) are small non-coding RNAs that regulate gene expression by binding to complementary sites of 3' untranslated region of target genes in plants and animals. Although interactions between miRNAs and their targets are complex and less understood, many miRNA target prediction algorithms exist but only a limited number of these were experimentally validated. Here, we propose a new computational method, microTarget for identifying miRNA – target interactions based on complementarity score and thermodynamic duplex stability. We validate our algorithm with experimental results in four different genomes and compare performance with other computational methods, miRanda and Diana-micro T. Statistical tests like signal-to-noise ratio, z score are performed and microTarget shows good performance and validation than other miRNA-target interactions methods.

KEYWORDS: miRNA target interaction, gene regulation, target prediction algorithm, target validation, signal-to-noise ratio, z score.

***Corresponding Author: DR. SHIBSANKAR DAS** Ph. D.
Department of Mathematics, Uluberia College, Uluberia, West Bengal, India.
*Email Address: ssdas80@gmail.com

© 2016 Life Science Informatics Publication All rights reserved
Peer review under responsibility of Life Science Informatics Publications

1. INTRODUCTION

MicroRNAs (miRNAs) are class of small, non-coding RNA product that plays important roles in a variety of normal and diseased biological processes. They target messenger RNAs (mRNAs) either by repressing mRNA translation or mediating mRNA degradation for regulating the expression of target genes in animals and plants [1-4]. There are several discovered miRNAs available in animal and plants, but only fraction of these miRNAs function are known to us. As a result, pathway and biological function of genes affected by miRNAs in various diseases like breast cancer [5], lung cancer [6], prostate cancer [7], colon cancer [8], remains largely unknown. Therefore, main goal of current miRNAs study is to understand the cellular functions of miRNAs and first step is the identifying target genes by miRNAs. Even though experimental based prediction algorithms of miRNAs and their potential targets have been developed but due to lengthy, laborious and economically unfavorable, computational target prediction methods tied with high-throughput experimental result can deliver valuable and more efficient target prediction methods. Different miRNA-target prediction algorithms predict targets with different techniques and criteria including base pairing, target accessibility, differing in the position and localization of Watson-Crick pairings and mismatches and evolutionary conservation of target site [6]. The widely used miRNA target prediction methods are miRanda [9], DIANA-micro T [10] etc. miRanda was first developed by Enright et al. [9] in 2003 to predict miRNA targets genes in *D. melanogaster*. The algorithm of this method is based on three features: comparison of miRNA complementarity of $3'$ UTR regions, free energies of RNA-RNA duplexes, and conservation of target sites in related genomes. Diana-microT [10] approach utilizes an extended miRNA seed region which is shifted across the $3'$ UTR of the target gene to search for alignments with 7, 8 or 9 nt Watson-Crick base-pairing. miRanda and DIANA-micro T have high Signal-to-noise ratio (SNR) which is a ratio between a total of predicted targets by single miRNA in searched $3'$ UTR and a total of predicted targets by artificial miRNA with randomized sequence in searched $3'$ UTR. Brennecke et al. [11] showed experimentally in *D. melanogaster* and *D. pseudoobscura* that strong complementary in $5'$ end of miRNA required to confer regulation and sites with weaker $5'$ complementarity require compensatory pairing to the $3'$ end of the miRNA in order to function. They also showed that miRNA with more than one G=U base-pair or bulge or mismatch in seed region could lose its activity almost completely. Xiaowei Wang [12] found that 6-mer or 7-mer with perfect or one G=U base pairing in positions 2-10 counted from miRNA $5'$ end, have good enrichment ratio in CLASH data. Betel et al. [13] showed that additional Watson-Crick Pairing at positions 12–17 enhances miRNA Targeting. We have incorporated the results of Brennecke et al [11], Xiaowei Wang [12]

© 2016 Life Science Informatics Publication All rights reserved

and Betel et al. [13] in our proposed miRNA-target prediction algorithm to improve false positives. In this study, we have proposed a new computational method to detect miRNA targets. Most miRNA-target prediction algorithms predict first targets in different genome and then try to validate experimental results. Here we first try to validate with experimental results and compare validation results with miRanda and Diana-micro T. We have performed various measures like signal-to-noise ratio, z score to access statistical significance of our algorithm and compared results with experimentally validated evidence in *D. melanogaster, Danio rerio* , *Gallus gallus* and *Arabidopsis thaliana* genomes.

2. MATERIAL AND METHODS

miRNA sequences

Experimentally validated miRNA sequences of one insect, *D. melanogaster* (Fruit Fly), two vertebrates, *Danio rerio* (Zebrafish) , *Gallus gallus* (Chicken) and one plant, *Arabidopsis thaliana* (Thale cress) were downloaded from MirTarBase database [14]. We have collected 45 miRNA sequences of *D. melanogaster* and 39, 31 and 40 miRNA sequences of *Danio rerio* , *Gallus gallus* and *Arabidopsis thaliana* respectively.

3' UTR sequences

MirTarBase database [15] validates 86 target genes of *D. melanogaster* and target genes of *Danio rerio* , *Gallus gallus* and *Arabidopsis thaliana* to 98, 34 and 68 respectively. We have downloaded 3' UTR of target genes of *D. melanogaster, Danio rerio* , *Gallus gallus* and *Arabidopsis thaliana* from UTRdb [15]. We have found only 80, 90, 22 and 50 3'UTR of target genes of *Danio rerio* , *Gallus gallus* and *Arabidopsis thaliana* respectively and used to validate miRNA-target duplex of miRanda, Diana-Micro T and our microTarget algorithm. For this study, we have downloaded miRanda program from and developed Diana-micro T program from the algorithm described in supplementary material by Kiriakidou et al. [10]. The thresholds for miRanda used for possible target are: complementarity score ≥ 80 and energy of the duplex structure ≤ -14 kcal/mol as described in [9]. Energy cut-off used for Diana-Micro T program is -8 kcal/mol instead of using -20 kcal/mol because we could not find good number of results using original energy cut-off -20 kcal/mol used by Kiriakidou et al. [10]. We have proposed our miRNA target prediction algorithm microTarget as follows.

microTarget algorithm

microTarget algorithm is similar to miRanda algorithm [9], however instead of using empirical rules. It uses similar complementarity parameters as miRanda algorithm [9] at individual alignment positions: +5 for G≡C, +5 for A=U, +2 for G=U and -3 for all other nucleotide pairs. The algorithm

© 2016 Life Science Informatics Publication All rights reserved

uses affine penalties for gap-opening (-8) and gap-extension (-1). Complementarity scores (positive and negative values) at the first eleven positions are multiplied by a scaling factor (here it is set at 2.0). The following five rules are applied with positions counted starting at the 5' end of the miRNA:

(1) There must be 6 to 8 base pairs between positions 1 to 10.

(2) Seed region with 8 base pairs and starting from position 1, may have up to two G=U base-pairs or one bulge (either of the miRNA or of the 3'UTR) or single non-G=U mismatch in between seed region (i.e. from positions 2-7).

(3) Seed region with 7 base pairs and starting from positions 1-4, may have one G=U base-pair or one bulge (either of the miRNA or of the 3'UTR) or single non-G=U mismatch in between seed region.

(4) Seed region with 6 base pairs and starting from positions 2-5, may have only one G=U base-pair in between seed region.

(5) If G=U base pair or bulge or mismatch are used in seed region and seed region starting from positions either 3-4 or 4-5, there must be at least 4 base pairs (including G=U base-pairs) from positions 12 to 3' UTR end of miRNA.

Using these parameters and rules, complementarity score between a miRNA sequence and 3' UTR sequence is optimized using dynamic programming and summed over all aligned positions. All non-overlapping hybridization alignments in decreasing order of complementarity score down to some cutoff value (default value 80) are found. In order to calculate free energies of the RNA: RNA duplexes, we use folding routines from the Vienna RNA secondary structure programming library (RNAlib) [16]. The thresholds used for possible target are: complementarity score ≥ 80 and energy of the duplex structure ≤ -10 kcal/mol. Each possible target site between a miRNA and a UTR sequence is then scored according to the total energy and total score of all possible targets sites between those two sequences. The top ten ranked genes are selected as its candidate target genes of each miRNA. A target gene binding by multiple miRNAsis selected by the miRNAs that assign the highest scoring and lowest free energy of the miRNA target duplex to each potential site so that different miRNA target sites cannot overlap.

Randomized test

We perform similar randomized test as Enright et al. [9] did in miRanda. Randomized miRNA sequence was constructed by accumulating 100 sets of all miRNAs for each genome of *D. melanogaster, Danio rerio* , *Gallus gallus* and *Arabidopsis thaliana* respectively. Each random miRNA is generated by retaining the compositional characteristics of the miRNA. Each of these

© 2016 Life Science Informatics Publication All rights reserved

sets of randomized miRNAs was independently searched against all 3'UTR of target genes for each genome of *D. Melanogaster, Danio rerio* , *Gallus gallus* and *Arabidopsis thaliana* downloaded from UTRdb [15]. Results and counts are then averaged over all 100 random sets, and compared with the results of the actual miRNA. The Z-scores for each algorithm are generated from the actual miRNA counts, averaged random miRNA counts and their standard deviations for each genome of *D. melanogaster, Danio rerio* , *Gallus gallus* and *Arabidopsis thaliana.*

3. RESULTS AND DISCUSSION:

Validation results of miRanda, Diana-micro T and microTarget

In *D. melanogaster,* 45 miRNAs predict 47 target genes with 57 miRNA-target interactions by miRanda algorithm and 42 target genes with 53 miRNA-target interactions by Diana-Micro T algorithm, whereas our microTarget algorithm predicts 64 target genes with 82 miRNA-target interactions. miRanda and Diana-micro T validate 35 target genes with 42 miRNA-target interactions and 37 target genes with 44 miRNA-target interactions in *Danio rerio.* microTarget validates 54 target genes with 69 miRNA-target interactions in *Danio rerio.* miRanda validates 13 target genes with 15 miRNA-target interactions and 7 target genes with 9 miRNA-target interactions in *Gallus gallus* and *Arabidopsis thaliana* respectively. Diana-micro T validates 9 target genes with 10 miRNA-target interactions and 10 target genes with 11 miRNA-target interactions in *Gallus gallus* and *Arabidopsis thaliana* respectively. In *Gallus gallus* and *Arabidopsis thaliana,* microTarget validates 16 target genes with 21 miRNA-target interactions and 14 target genes with 18 miRNA-target interactions respectively. Figure- 1 shows that microTarget performs better than miRanda and Diana-micro T in all four genomes.

Signal - to - noise ratio analysis

Signal-to-noise ratio is defined as the ratio between average no. of predicted targets by single miRNA in searched 3'UTR and average no. of predicted targets by artificial miRNA with randomized sequence in searched 3'UTR. We perform randomized test of miRanda, Diana-micro T and microTarget by constructing 100 sets of all miRNAs for each genome of *D. melanogaster, Danio rerio* , *Gallus gallus* and *Arabidopsis thaliana* respectively.

© 2016 Life Science Informatics Publication All rights reserved

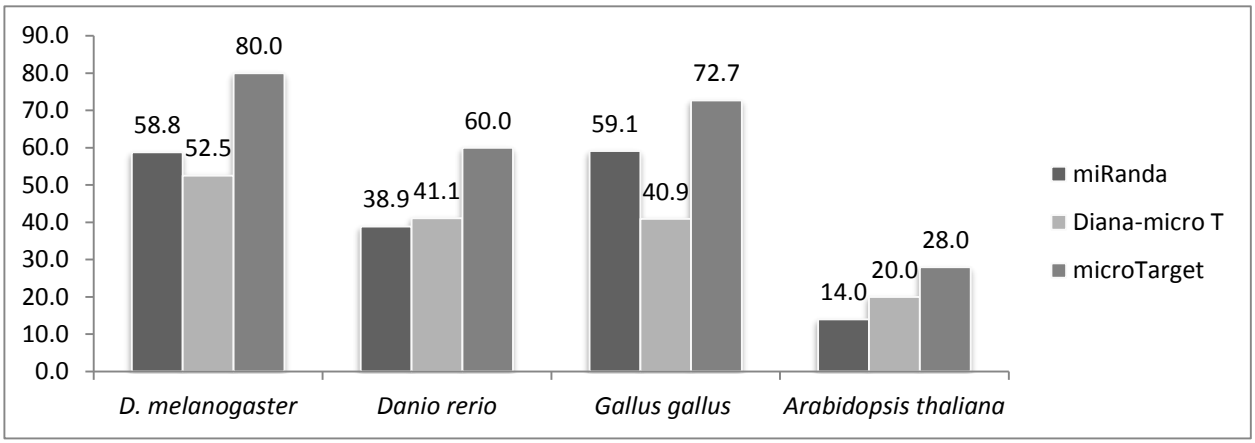

Figure 1 Genome wise percentage of validation of target genes by miRand, Diana-micro T and microTarget

Figure -2 shows the signal-to-noise ratio of miRanda, Diana-micro T and microTarget in all four genomes. Signal-to-noise ratios of microTarget are greater than 2.0 in three genomes and 1.9 in one genome which shows that rate of false positive of microTarget is less than miRanda and Diana-micro T. We have also calculated the z score of miRanda , Diana- micro T and microTarget in all four genomes (shown in Table 1). These results indicate that microTarget is significantly predicted miRNA-target interactions than miRanda and Diana- micro T.

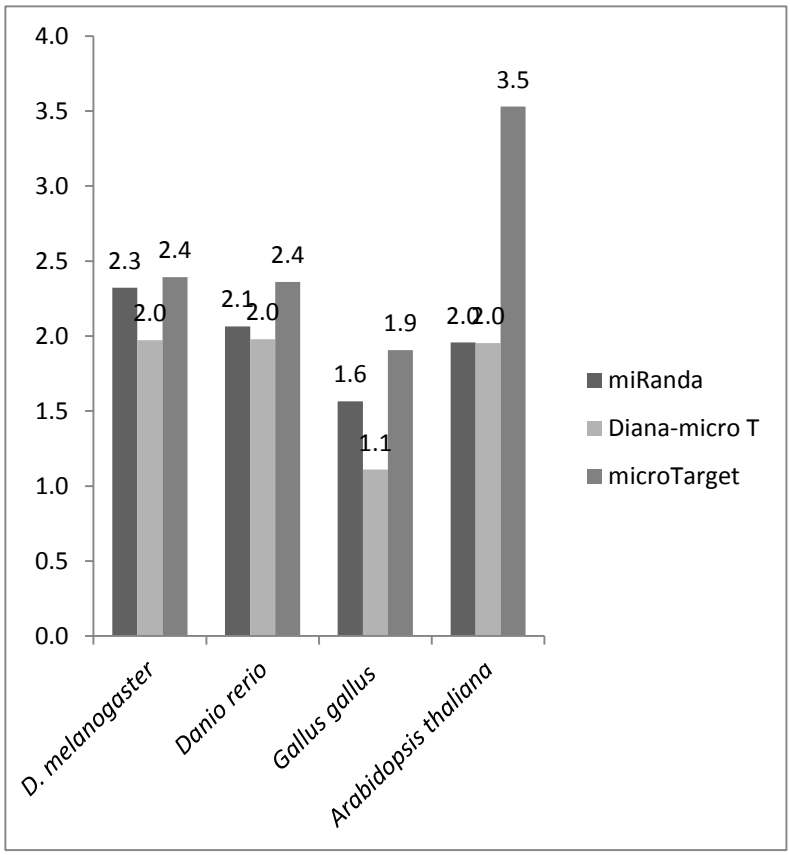

Figure 2 Signal-to-noise ratio of miRanda, Diana-micro T and microTarget

© 2016 Life Science Informatics Publication All rights reserved

Table 1: Genome wise comparison of real versus randomized miRNAs of miRanda, Diana-micro T and microTarget algorithm

		Average no. of predicted targer gene per miRNA	Average no. of predicted targer gene per randomized miRNA	Standard Deviation	Z Score
D. melanogaster	miRanda	1.27	0.55	0.08	9.00
	Diana-micro T	1.16	0.59	0.08	7.13
	microTarget	1.82	0.76	0.10	10.60
Daniorerio	miRanda	1.08	0.52	0.09	6.22
	Diana-micro T	1.13	0.57	0.1	5.60
	microTarget	1.77	0.75	0.05	20.40
Gallus gallus	miRanda	0.48	0.31	0.06	2.83
	Diana-micro T	0.32	0.29	0.07	0.43
	microTarget	0.68	0.36	0.07	4.57
Arabidopsis thaliana	miRanda	0.23	0.12	0.02	5.50
	Diana-micro T	0.25	0.13	0.05	2.40
	microTarget	0.45	0.13	0.05	6.40

Strong / weak experimentally validated evidence

We have downloaded miRNA-target interactions supported by strong or weak evidences from MirTarBase database [15]. Strong evidences are supported by Luciferase reporter assay [17], GFP Reporter assay [18][19],Western blot [20] and RT-Pcr [21]. Evidence supported by pSILAC [22], Microarray [23] and Northern blot [24][25] are considered as Weak evidences. Genes which are supported by both Strong and Weak evidences are categorized as Strong evidence to avoid the use of multiple copies of same gene. In *Arabidopsis thaliana, 8* out of 13 target genes supported by Stong or Weak evidences are predicted by mocroTarget and miRanda and Diana-micro T predict 5 target genes each. List of genes of *Arabidopsis thaliana* genome are shown in Table 2. miRanda and Diana-micro T predict 10 and 6 target genes respectively out of 19 target genes supported by Strong/Weak evidences of *Gallus gallus* genome. microTarget predicts 13 target genes (list shown in Table 3).

© 2016 Life Science Informatics Publication All rights reserved

Table 2: Predicted target genes list supported by Strong/Weak evidences of *Arabidopsis thaliana* Genome; '+' means predicted by the algorithm and '-' means not predicted by the algorithm

Target Gene	Entrez Gene ID	Evidance Type	Function	microTarget	miRanda	Diana -micro T
AP2	829845	Strong	transcription factor	-	-	-
CSD1	837405	Strong	Copper/Zinc Superoxide Dismutase 1	-	-	-
CSD2	817365	Strong	Copper/Zinc Superoxide Dismutase 2	-	-	-
MYB33	830497	Strong	DNA binding / transcription factor	-	-	-
NF-YA1	831124	Weak	transcription factor	+	+	-
NF-YA2	819738	Weak	transcription factor	+	+	+
SPL2	834345	Weak	DNA binding / transcription factor	+	-	-
SPL3	817948	Strong	DNA binding / transcription factor	+	+	+
SPL3	817948	Weak	DNA binding / transcription factor	+	+	+
SPL4	841749	Weak	DNA binding / transcription factor	-	-	-
SPL5	820758	Weak	DNA binding / transcription factor	+	-	+
SPL9	818820	Weak	transcription factor	+	-	-
TOE2	836134	Weak	DNA binding / transcription factor	+	+	+

In *D. melanogaster,* microTarget predicts 57 out of 73 target genes supported by Strong/Weak evidences whereas miRanda and Diana-micro T predicts 42 and 38 target genes respectively. All the genes predicted by three algorithms are shown in Table S1 (Supplementary Data). 31 out of 82 target genes are validated by Diana-micro T in *Danio rerio* and 34 target genes are validated by miRanda. microTarget validates 50 target genes in *Danio rerio* (list of target genes shown in Table S2 (Supplementary Data)). These results (in Figure 3) show that microTarget validates strong/weak experimental results better than miRanda and Diana-mocro T.

© 2016 Life Science Informatics Publication All rights reserved

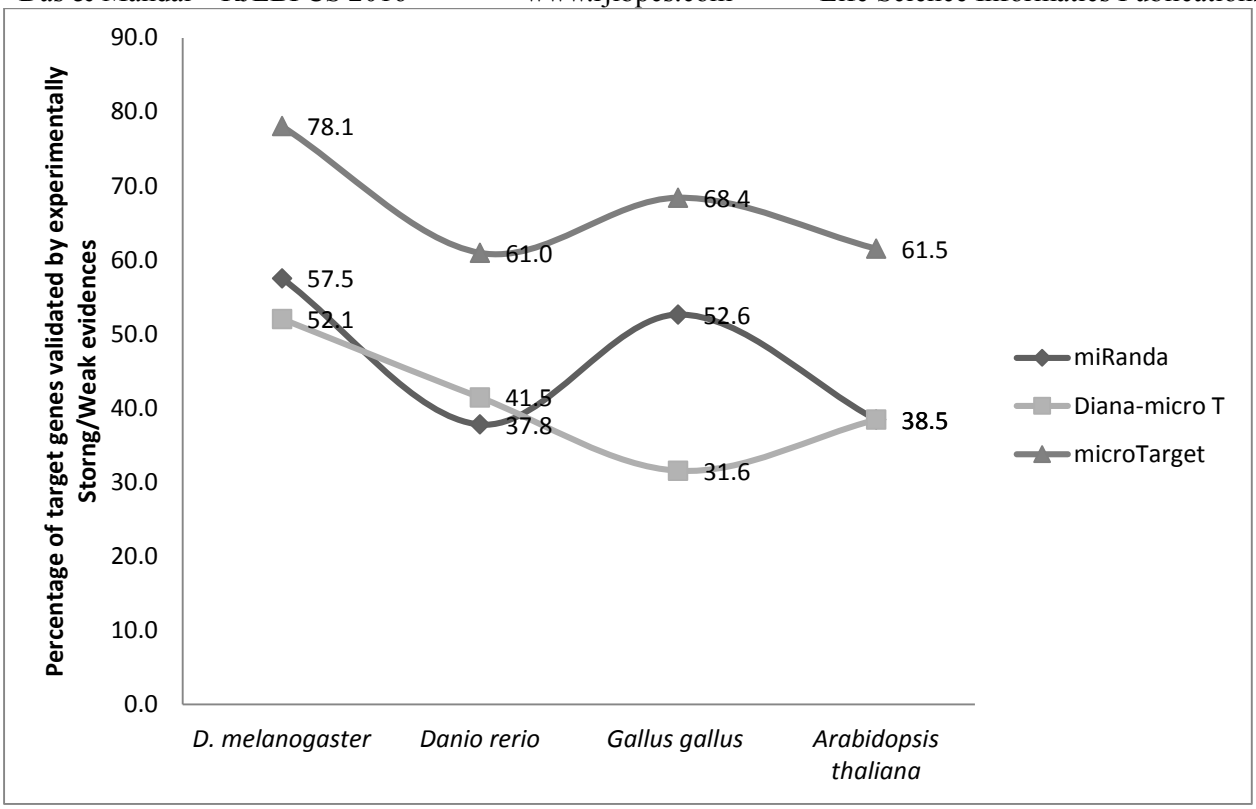

Figure 3 Genome wise performance of miRand, Diana-micro T and microTarget

4. CONCLUSION

We have presented here new miRNA-target prediction algorithm microTarget. We have tried to validate our algorithm using experimentally validated results downloaded from MirTarBase database [14]. Results show that microTarget is competitive with miRanda and Diana-micro T. Evaluation by a variety of statistical measures like signal–to-noise ratio, z score are performed and we have shown that microTarget peroforms better results than miRanad and Diana-micro T in *D. melanogaster*, *Danio rerio* , *Gallus gallus* and *Arabidopsis thaliana* genomes. Although microTarget incorporates many features important for microRNA-target interactions, all potential aspects of target specificity are not included in the model. New results from high-throughput microRNA target identification experiments may give some light to improve the success rate of target prediction algorithms and help to unravel the biology of regulation by miRNA-mRNA interaction.

ACKNOWLEDGEMENT

Financial support by the University Grants Commission, India, sanction No. F.PSW-081/13-14 (ERO) ID No. WC2-152, is gratefully acknowledged

© 2016 Life Science Informatics Publication All rights reserved

Table 3: Predicted target genes list supported by Strong/Weak evidences of *Gallus gallus* Genome; '+' means predicted by the algorithm and '-' means not predicted by the algorithm

Target Gene	Entrez Gene ID	Evidance Type	Function	microTarget	miRanda	Diana micro T
ATF2	395727	Strong	cyclic AMP-dependent transcription factor ATF-2	+	+	-
CREB1	395099	Strong	cyclic AMP-responsive element-binding protein 1	+	+	+
CTNNB1	395964	Strong	catenin beta-1	+	+	+
E2F8	422978	Strong	similar to E2F family member 8	+	+	-
ERBB4	395671	Strong	receptor tyrosine-protein kinase erbB-4	+	+	-
EZH2	420784	Strong	similar to Enhancer of zeste homolog 2	-	-	-
FHL2	418726	Strong	similar to DRAL gene product	+	+	+
FN1	396133	Strong	similar to fibronectin 1 isoform 1 preproprotein	-	-	-
GPI	415783	Strong	glucose-6-phosphate isomerase	-	-	-
HAND2	395813	Strong	heart- and neural crest derivatives-expressed protein 2	-	-	+
HDAC4	374207	Strong	histone deacetylase 4	-	-	-
ITGB1	374058	Strong	integrin beta-1 precursor	+	+	+
KLF11	421934	Strong	Krueppel-like factor 11	+	-	-
LAMC1	424442	Strong	laminin, gamma 1 (formerly LAMB2)	+	+	+
MDM2	395609	Strong	similar to MDM2	+	+	-
NOS2	395807	Strong	nitric oxide synthase, inducible	+	+	-
SOX9	374148	Strong	transcription factor SOX-9	+	-	-
SPRED1	423292	Strong	hypothetical protein	+	-	-
YAP1	396171	Strong	yorkie homolog	-	-	-

© 2016 Life Science Informatics Publication All rights reserved

REFERENCES

1. Pasquinelli AE. MicroRNAs: deviants no longer. Trends Genet, 2002; 18:171-173.
2. Ambros, V. MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. Cell, 2003; 113: 673-676.
3. Carrington JC, Ambros V. Role of microRNAs in plant and animal development. Science, 2003; 301: 336-338.
4. Lin SY, Johnson SM, Abraham M, Vella MC, Pasquinelli A, Gamberi C, Gottlieb E, Slack FJ. The C. elegans hunch back Homolog, hbl-1, controls temporal patterning and is a probable MicroRNA target. Dev Cell, 2003; 4:639-650.
5. Ridley M. Genome: The Autobiography of a Species in 23 Chapters, 1st ed.; Topeka Bindery: New York, 2006.
6. Bartel D. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell, 2004; 116: 281–97.
7. Zhao T, Li G, Mi S, Hannon GJ, Wang XJ, Qi YA. complex system of small RNAs in the unicellular green alga Chlamydomonas reinhardtii. Genes Dev, 2007; 21(10):1190–2003.
8. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet, 2004; 5 (7): 522–531.
9. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. Genome Biol., 2004; 5 (1): R1-14.
10. Kiriakidou M, Nelson PT, Kouranov A, Fitziev P, Bouyioukos C, Mourelatos Z, Hatzigeorgiou A. A combined computational-experimental approach predicts human microRNA targets. Genes Dev., 2004; 18(10):1165-1178.
11. Brennecke J, Stark A, Russell RB, Cohen SM. Principles of microRNA–target recognition. PLoS Biol, 2005; 3 (3): e85.
12. Xiaowei Wang. Composition of seed sequence is a major determinant of microRNA targeting patterns. Bioinformatics, 2014; 30 (10):1377–1383.
13. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell. , 2007; 27(1): 91-105.
14. Chou CH, Chang NW, Shrestha S, Hsu SD, Lin YL, Lee WH, Yang CD, Hong HC, Wei TY, Tu SJ, Tsai TR, Ho SY, Jian TY, Wu HY, Chen PR, Lin NC, Huang HT, Yang TL, Pai CY, Tai CS, Chen WL, Huang CY, Liu CC, Weng SL, Liao KW, Hsu WL, Huang HD. miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. Nucleic Acids Res. 2016; 44 (D1):239-47
15. Giorgio G, Antonio T, Flavio, L, Flavio M, Sabino L, Sandro B, Vincenzo AG, David SH,

© 2016 Life Science Informatics Publication All rights reserved

Giulio P, Ernesto P, Graziano P. UTRdb and UTRsite (RELEASE 2010): a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. Nucleic Acids Res. 2010; 38 (Database issue): D75–D80.

16. Wuchty S, Fontana W, Hofacker IL, Schuster P. Complete suboptimal folding of RNA and the stability of secondary structures. Biopolymers, 1999; 49:145-165.

17. Wood KV. The chemistry of bioluminescent re- porter assays. Promega Notes, 1998; 65*:*14–20.

18. Koo J, Kim Y, Kim J, Yeom M, Lee IC, Nam HG. A GUS/Luciferase Fusion Reporter for Plant Gene Trapping and for Assay of Promoter Activity with Luciferin-Dependent Control of the Reporter Protein Stability. Plant and Cell Physiology, 2007; 48 (8):1121–1131.

19. Nordgren IK, Tavassoli A. A bidirectional fluorescent two-hybrid system for monitoring protein-protein interactions. Molecular BioSystems, 2014; 10 (3):485–90.

20. Burnette WN. Western blotting: electrophoretic transfer of proteins from sodium dodecyl sulfate—polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Analytical Biochemistry ,1981; 112 (2):195–203.

20. Freeman WM, Walker SJ, Vrana KE . Quantitative RT-PCR: pitfalls and potential. BioTechniques, 1999; 26 (1):112–22, 124–5.

21. Schwanhäusser B, Manfred G, Gunnar Di, Matthias S. Global analysis of cellular protein translation by pulsed SILAC. Proteomics , 2009; 9 (1):205–9.

22. Adomas A, Heller G, Olson A, Osborne J, Karlsson M, Nahalkova J, Van ZL, Sederoff R, Stenlid J, Finlay R, Asiegbu FO. Comparative analysis of transcript abundance in Pinussylvestris after challenge with a saprotrophic, pathogenic or mutualistic fungus. Tree Physiol., 2008; 28 (6):885–897.

23. Collins ML, Irvine B, Tyner D, Fine E, Zayati C, Chang C, Horn T, Ahle D, Detmer J, Shen LP, Kolberg J, Bushnell S, Urdea MS, Ho DD. A branched DNA signal amplification assay for quantification of nucleic acid targets below 100 molecules/ml. *N*ucleic Acids Research, 1997; 25 (15):2979–2984.

24. Alberts B, Johnson A, Lewis JRM, Roberts K, Walter P. In: Molecular Biology of the Cell; 5th ed.; Taylor & Francis Group, Garland Science: NY, 2002; pp 538–539.

25. Kevil CG, Walsh L, Laroux FS, Kalogeris T, Grisham MB, Alexander JS. An Improved Rapid Northern Protocol. Biochem. and Biophys. Research Comm., 1997; 238: 277–279

© 2016 Life Science Informatics Publication All rights reserved

Supplementary Tables

Table S1: Predicted target genes list supported by Strong/Weak evidences of *D. melanogaster* Genome; '+' means predicted by the algorithm and '-' means not predicted by the algorithm

Target Gene	Entrez Gene ID	Evidence Type	Function	microTarget	miRanda	Diana-micro T
ab	34560	Strong	abrupt, isoform A	+	+	+
abd-A	42037	Strong	abdominal A, isoform A	+	+	+
Abd-B	47763	Strong	abdominal B, isoform A	+	+	-
ana	35913	Strong	anachronism, isoform A	-	+	-
aop	33392	Strong	anterior open, isoform A	+	+	+
arm	31151	Strong	armadillo, isoform A	-	-	+
Axn	43565	Strong	axin, isoform A	-	-	-
bap	42537	Strong	bagpipe	-	-	-
BobA	50281	Strong	brother of bearded A	+	+	+
boss	43146	Strong	bride of sevenless	+	+	+
brd	247926	Strong		+	+	+
Bx	32846	Strong	beadex, isoform A	+	+	+
Cbl	38961	Weak	Cbl, isoform A	-	-	-
CG11377	33166	Strong	CG11377	+	+	-
CG13060	39802	Strong	CG13060	-	-	+
CG13380	40029	Strong	CG13380	+	+	+
CG17065	33026	Strong	CG17065, isoform A	+	-	-
CG18542	41176	Strong	CG18542	+	+	-
CG31121	42975	Strong	CG31121, isoform A	+	+	-
CG32767	31430	Strong	CG32767, isoform C	+	+	+
CG7158	40410	Strong	CG7158	+	-	-
CG8420	41097	Strong	CG8420	-	-	-
Chd64	38490	Strong	Chd64	+	-	+
cos	35653	Strong	costa	+	+	+
Cpr56F	37299	Strong	cuticular protein 56F	+	+	+
CrebA	39682	Strong	Cyclic-AMP response element binding protein A, isoform A	+	+	+
crim	39321	Strong		+	+	-

© 2016 Life Science Informatics Publication All rights reserved

Target Gene	Entrez Gene ID	Evidence Type	Function	microTarget	miRanda	Diana-micro T
CRMP	40675	Strong	collapsin response mediator protein, isoform A	+	-	+
dg	5657005	Strong		-	-	-
Dip1	53580	Strong	dorsal interacting protein 1	-	-	-
Dl	42313	Strong	delta, isoform A	-	+	-
dpr10	39180	Strong	dpr10, isoform A	+	+	+
e	42521	Strong	ebony	-	-	-
EcR	35540	Strong	ecdysone receptor, isoform A	+	+	-
Eip74EF	39962	Strong	Ecdysone-induced protein 74EF, isoform A	+	+	+
ena	37201	Strong	enabled, isoform A	-	-	-
esg	34903	Strong	escargot	-	-	-
ex	33218	Strong	expanded	+	+	+
Fit1	38488	Strong	fermitin 1, isoform A	+	-	-
fng	40314	Strong	fringe	+	-	-
fu	32855	Strong	fused	+	-	-
gli	251204	Strong		+	+	+
grim	40014	Strong	grim	+	+	+
Gug	46156	Strong	grunge, isoform A	+	-	-
h	38995	Strong	hairy, isoform A	+	+	-
IA-2	33277	Strong		-	-	-
Ice	43514	Strong	Ice	-	-	-
IP3K2	32285	Strong	inositol 1,4,5-triphosphate kinase 2, isoform A	+	+	+
Jafrac2	53577	Strong	thioredoxin peroxidase 2, isoform A	+	-	-
l(1)MZ4	249418	Strong		-	-	-
Mef2	36032	Strong	myocyte enhancer factor 2, isoform A	+	+	-
mei-P26	45775	Strong	mei-P26	-	-	-
Msr-110	38629	Strong	Msr-110, isoform A	+	-	-

© 2016 Life Science Informatics Publication All rights reserved

Target Gene	Entrez Gene ID	Evidence Type	Function	microTarget	miRanda	Diana-micro T
Nedd4	39958	Strong	Nedd4, isoform I	+	-	-
nerfin-1	44786	Strong	nervous fingers 1	+	+	+
Notum	39751	Strong	notum, isoform A	+	-	+
os	32813	Strong	outstretched, isoform A	+	-	-
pan	43769	Strong	pangolin, isoform A	+	+	-
rpr	40015	Strong	reaper	+	+	+
rt	39297	Strong	rotated abdomen	+	+	+
sens	45328	Strong	senseless	+	+	+
sha	36213	Strong		-	-	-
sinu	46187	Strong	sinuous, isoform A	-	-	-
skl	40016	Strong	sickle	+	+	+
smo	33196	Strong	smoothened	+	+	+
SP555	53471	Strong	SP555, isoform A	+	+	+
Su(z)12	48071	Strong	Su(z)12, isoform A	+	+	+
sug	36424	Strong	sugarbabe	-	-	-
Tom	39619	Strong	twin of m4	+	+	+
ttk	48317	Strong	tramtrack, isoform A	+	+	+
tup	35147	Strong	tailup, isoform A	+	+	+
tutl	46015	Strong	turtle, isoform A	+	-	-
Ubx	42034	Strong	ultrabithorax, isoform A	+	-	+
ush	33225	Strong	u-shaped	+	-	+
W	40009	Strong	wrinkled	+	-	-
wls	39259	Strong	wntless, isoform A	+	+	+
yellow-c	34879	Strong	yellow-c	+	+	+

© 2016 Life Science Informatics Publication All rights reserved

Table S2: Predicted target genes list supported by Strong/Weak evidences of *Danio rerio* Genome; '+' means predicted by the algorithm and '-' means not predicted by the algorithm

Target Gene	Entrez Gene ID	Evidence Type	Function	microTarget	miRanda	Diana- micro T
actb1	57934	Strong	actin, beta 1	+	+	+
add3a	556762	Strong	gamma-adducin	+	-	+
alas2	64607	Strong	5-aminolevulinate synthase, erythroid-specific, mitochondrial precursor	+	-	-
alg9	393598	Strong	alpha-1,2-mannosyltransferase ALG9	+	+	+
ap2m1b	394001	Strong	AP-2 complex subunit mu-1-B	-	-	+
arcn1b	399661	Strong	archain 1b	-	-	+
arl6ip5b	436790	Strong	ADP-ribosylation-like factor 6 interacting protein 5	+	+	-
arpc4l	445305	Strong	actin related protein 2/3 complex, subunit 4, like	+	-	-
ascl1a	30466	Strong	achaete-scute homolog 1a	+	+	-
atp6v0d1	322811	Strong	V-type proton ATPase subunit d 1	+	-	-
atp6v1ba	359839	Strong	V-type proton ATPase subunit B, kidney isoform	+	-	+
atp6v1e1b	192335	Strong	ATPase, H+ transporting, lysosomal, V1 subunit E1b	-	-	-
cav1	323695	Strong	caveolin-1 isoform a	+	-	+
cmyb	30519	Strong	transcriptional activator Myb	+	+	-
cnn2	406658	Strong	calponin-2	+	+	+
cnn3a	326931	Strong	calponin 3, acidic a	+	-	-
cnpy1	402819	Strong	protein canopy-1 precursor	-	-	-
copz1	57970	Strong	coatomer subunit zeta-1	+	-	+
cxcl12a	352944	Strong	chemokine (C-X-C motif) ligand 12a (stromal cell-derived factor 1)	+	-	-
cxcr7b	100000764	Strong	chemokine (C-X-C motif) receptor 7b	-	-	+
ddx18	321127	Strong	ATP-dependent RNA helicase DDX18	-	-	-
disp2	405793	Strong	protein dispatched homolog 2	-	-	-

© 2016 Life Science Informatics Publication All rights reserved

Target Gene	Entrez Gene ID	Evidence Type	Function	microTarget	miRanda	Diana- micro T
dkk3	100038765	Strong	dickkopf-related protein 3	-	-	-
dll4	563920	Strong	delta-like protein 4	+	-	-
dpm1	445202	Strong	dolichol-phosphate mannosyltransferase	+	-	-
fgf8a	30538	Strong	fibroblast growth factor 8 a	+	-	+
fgfr1a	30705	Strong	basic fibroblast growth factor receptor 1-A	-	-	-
fzd4	30363	Strong	frizzled class receptor 4	+	+	+
gata2a	30480	Strong	GATA-binding protein 2a	+	-	-
gata6	58076	Strong	transcription factor GATA-6	+	+	-
gsk3b	30654	Strong	glycogen synthase kinase-3 beta	-	+	-
gstm	324366	Strong	glutathione S-transferase M	+	-	-
hand2	58150	Strong	heart- and neural crest derivatives-expressed protein 2	-	-	+
her5	30285	Strong	hairy-related 5	+	+	+
her9	140613	Strong	hairy-related 9	-	-	-
homer1b	436769	Strong	homer scaffolding protein 1b	-	-	-
hoxb3a	30339	Strong	homeobox protein Hox-B3a	-	-	-
hspd1	282676	Strong	60 kDa heat shock protein, mitochondrial	+	+	+
idh1	100006589	Strong	isocitrate dehydrogenase [NADP] cytoplasmic	+	-	-
jag1b	140423	Strong	protein jagged-1b precursor	+	+	-
jun	335916	Strong	transcription factor AP-1	-	-	-
klf3	117603	Strong	Krueppel-like factor 3	+	+	-
klf4b	65238	Strong		-	-	-
klfd	30104	Strong	Kruppel-like factor d	-	-	-
lef1	30701	Strong	lymphoid enhancer-binding factor 1	-	+	-
lfng	30158	Strong	beta-1,3-N-acetylglucosaminyltransferase lunatic fringe	+	+	-
lft2	30146	Strong	lefty2	+	+	+
lin28a	394066	Strong	protein lin-28 homolog A	+	+	+
lmo2	30332	Strong	rhombotin-2	+	-	+

© 2016 Life Science Informatics Publication All rights reserved

Target Gene	Entrez Gene ID	Evidence Type	Function	microTarget	miRanda	Diana- micro T
meis1	170446	Strong	homeobox protein Meis1	+	-	-
mknk2b	373121	Strong	MAP kinase-interacting serine/threonine kinase 2b	+	+	-
myca	30686	Strong	transcriptional regulator Myc-A	-	-	-
mycb	393141	Strong	transcriptional regulator Myc-B	+	+	+
nanos1	322903	Strong	nanos homolog 1	-	+	-
ndr1	799352	Strong	nodal-related 1	-	-	-
nme2b.1	30083	Strong	nucleoside diphosphate kinase A	-	-	+
nog3	30173	Strong	noggin 3	-	-	-
pard3	403050	Strong	partitioning defective 3 homolog	+	-	-
parn	791461	Strong	poly(A)-specific ribonuclease PARN	+	+	+
pax6b	60639	Strong	paired box gene 6b	-	-	-
pdgfra	386856	Strong	alpha-type platelet-derived growth factor receptor	+	+	+
pdlim1	550568	Strong	PDZ and LIM domain protein 1	+	+	+
pfn2l	321383	Strong	profilin-2	+	-	+
pik3r2	404211	Strong	phosphoinositide-3-kinase, regulatory subunit 2 (beta)	-	-	-
ppp1cab	327301	Strong	protein phosphatase 1, catalytic subunit, alpha-like	-	-	-
prkci	117507	Strong	protein kinase C iota type	-	-	-
rab13	373105	Strong	ras-related protein Rab-13	+	-	-
sdr16c5b	406799	Strong	short chain dehydrogenase/reductase family 16C, member 5b	+	+	+
slc9a3r1	327272	Strong	solute carrier family 9 (sodium/hydrogen exchanger), isoform 3 regulator 1	-	-	-
smarca5	282615	Strong	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5	+	-	+
smarcb1b	30731	Strong	SWI/SNF related, matrix associated, actin dependent regulator of chromatin,	-	-	-

© 2016 Life Science Informatics Publication All rights reserved

Target Gene	Entrez Gene ID	Evidence Type	Function	microTarget	miRanda	Diana- micro T
			subfamily b, member 1b			
smarcd1	323115	Strong	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 1	+	+	-
smo	30225	Strong	smoothened homolog	-	-	-
snx5	406797	Strong	sorting nexin-5	+	-	-
spred1	406513	Strong	sprouty-related, EVH1 domain-containing protein 1	+	+	+
spry2	335098	Strong	protein sprouty homolog 2	-	-	+
spryd7b	393498	Strong	SPRY domain containing 7b	+	+	+
sufu	100001615	Strong	suppressor of fused homolog	+	+	-
tdrd7	406379	Strong	tudor domain containing 7 isoform 1	+	-	+
tp53	30590	Strong	cellular tumor antigen p53	-	+	+
tp63	260407	Strong	tumor protein 63 isoform gamma	-	-	-
tpm3	373076	Strong	tropomyosin 3 isoform 1	+	+	+
trim71	561754	Strong	tripartite motif-containing 71	+	-	+
ttk	317763	Strong	dual specificity protein kinase Ttk	-	-	-
vasa	30263	Strong	probable ATP-dependent RNA helicase DDX4	-	-	-
vcam1	561971	Strong	vascular cell adhesion protein 1	-	-	-
vcana	116993	Strong	chondroitin sulfate proteoglycan 2	-	+	+
vegfaa	30682	Strong	vascular endothelial growth factor Aa	-	-	-
wnt2	30127	Strong	protein Wnt-2 precursor	-	-	-
zeb1b	114425	Strong	zinc finger E-box binding homeobox 1b	-	-	-

© 2016 Life Science Informatics Publication All rights reserved