

BIOINFORMATION

Discovery at the interface of physical and biological sciences

open access

www.bioinformatics.net

Hypothesis

Volume 11(12)

Computational identification of putative miRNAs and their target genes in pathogenic amoeba *Naegleria fowleri*

Dyavegowda Padmashree & Narayanaswamy Ramachandra Swamy*

Department of Biochemistry, Central College Campus, Bangalore University, Bangalore, Karnataka, India; N Ramachandra Swamy – Email: dnrswamy@bub.ernet.in; Phone: 22961335; *Corresponding author

Received December 05, 2015; Revised December 11, 2015; Accepted December 12, 2015; Published December 31, 2015

Abstract:

Naegleria fowleri is a parasitic unicellular free living eukaryotic amoeba. The parasite spreads through contaminated water and causes primary amoebic meningoencephalitis (PAM). Therefore, it is of interest to understand its molecular pathogenesis. Hence, we analyzed the parasite genome for miRNAs (microRNAs) that are non-coding, single stranded RNA molecules. We identified 245 miRNAs using computational methods in *N. fowleri*, of which five miRNAs are conserved. The predicted miRNA targets were analyzed by using miRanda (software) and further studied the functions by subsequently annotating using AmiGo (a gene ontology web tool).

Keywords: miRNAs, Computational approach, *Naegleria fowleri*, PAM, Prediction**Background:**

The genus *Naegleria* is a free-living amoeba, most commonly found in warm fresh water and soil across the world [1, 2]. More than 40 species of *Naegleria* have been identified but only *Naegleria fowleri* (*N. fowleri*) acts as a pathogenic species in human [3, 4]. *N. fowleri* is a thermophilic organism primitively grown at 37°C and can survive with temperature upto 45°C [5, 6]. *N. fowleri* exists in three stages; trophozoite (amoeba), flagellate and cyst. It enters body through the nasal route along the olfactory nerve tracts and finally infects central nervous system causing the disease primary amoebic meningoencephalitis (PAM) [7, 8]. The mitochondrial genome and a 60-kb nuclear DNA segment are responsible for PAM [9]. Once the host is infected, the symptoms observed are severe headache, vomiting, stiff neck, nausea, fever, etc., and the host dies within one to two weeks [7, 4]. Previous reports have shown that, the presence of *liver hydrolysate* (LH) enhanced pathogenicity of *N. fowleri* [10]. Reports about RNA interference in protozoan parasites [11], motivated us to identify the putative miRNAs and their targets in this protozoan parasite.

microRNAs (miRNAs) are endogenous, single stranded, small regulatory RNA molecules which are non-coding and are ~20-25 nucleotides in length [12-14]. MiRNAs form complementary base pairs to target mRNA which controls the expression of gene by inducing degradation or repressing the translation of mRNA or chromatin modification [14]. In 1993 Victor Ambros discovered first miRNA, lin-4 in *Caenorhabditis elegans* [15]. Today thousands of miRNAs have been identified in various multicellular organisms, which include drosophila, plants, viruses, vertebrates, nematodes etc., [16-20]. Recent studies have reported the identification of miRNA in *Entamoeba histolytica* and *Chlamydomonas reinhardtii* [21, 22]. MiRNA is a key regulator in various cellular and biological processes, which have been experimentally demonstrated in cell cycle, apoptosis, developmental stages, differentiation, response to stress, and in various types of cancers [23-25].

Following three procedures are involved in identifying miRNA (a) Classic cloning methods; (b) Deep sequencing method and (c) Computational approach. Computational approach can be further divided into three types: (i) ab initio prediction based

on the sequence and structural features; (ii) comparative genomic strategy based on evolutionary conservation; and (iii) integrated approach. In this study, we used computational approach based on sequence and structural feature to identify miRNA in *N. fowleri* and their targets.

Methodology:

Collection of data

In this study, nucleotide sequences of DNA-contigs of *N. fowleri* were downloaded from NCBI (<http://www.ncbi.nlm.nih.gov/>). Each contigs containing DNA nucleotides were downloaded and stored in FASTA format which were used for further analyses. The following steps were followed to identify miRNA genes from *N. fowleri* (Figure 1).

Prediction of miRNA genes by computational methods

N. fowleri strain ATCC 30863, genome size 27.79Mb, consists of 574 DNA contigs, which were downloaded from NCBI (<http://www.ncbi.nlm.nih.gov/>). Individual contigs were subjected to scan by using software called Einverted EMBOSS which is available in the website <http://emboss.bioinformatics.nl/cgi-bin/emboss/einverted>. EMBOSS (the European Molecular Biology Open Software Suite) is an open source software analysis package. This package contains Einverted, which is used to find DNA inverted repeats in nucleotide sequences to form secondary structure of DNA. It is also used to identify the sequences that can form hairpin like structure. In our study, we used parameter condition as described; threshold score 30, match score 3, mismatch score 4, gap penalty 6, maximum repeat length 200. The resulted outputs from Einverted EMBOSS were further filtered by collecting the sequences which are greater than and equal to (\geq) 15 base pairs in complementary strand and the length was separated by 40 bases in second step, this avoided more bulges and mismatches in the secondary structures. Since, the precursor miRNA length varies [26], we collected the sequences of 60-120 nucleotides in length in third step. In the Fourth step, secondary structures of the miRNAs were predicted by using RNA-fold program with minimum folding free energy (MFE) -20kcal/mol . In the fifth step, the miRNA sequences were reduced to 30-60% by screening GC content. The results were subjected to screening by using BlastX program to remove the protein coding sequences and excluded them in sixth step. In next step, results for miRNA candidates were further narrowed by using miPred software (<http://www.bioinf.seu.edu.cn/miRNA/>) to distinguish the real miRNA from that of a pre-miRNA like hairpin and pseudo miRNA precursor in ninth step [27]. Further, we removed the homologous sequences and repetitive sequences using the program RepeatMasker version 4.0.5 [28]. Finally, mature miRNA sequences were identified using mature naïve bayes [29].

Identification of conserved miRNA

To identify the conserved miRNA candidates in *N. fowleri*, we retrieved all mature miRNA sequences from miRBase Release 21 (<http://www.mirbase.org/>) [30]. A BlastN search was performed for all mature miRNA sequences from miRBase against 245 precursor miRNA of *N. fowleri* with default parameters and e-value 0.001. The results of blastN were further analyzed by applying two conditions: (1) more than 90% identity between the sequences and (2) mismatches between the sequences not more than three bases Table 1 (see supplementary material).

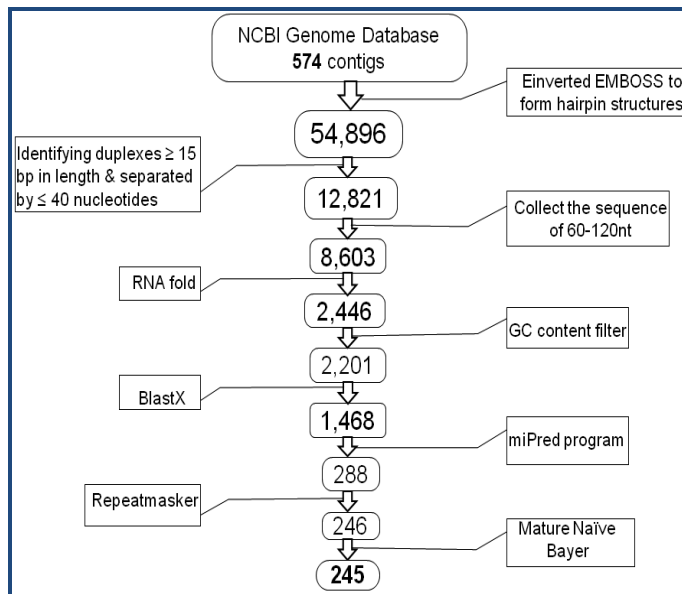


Figure 1: The flow chart followed for identification of miRNA in *N. fowleri*.

Prediction of miRNA targets

We used miRanda software version 3.3a released in 2010 to identify the target genes in *N. fowleri*. The following parameters were used in these studies; match with minimum threshold score of 120; target mRNA duplex with minimum folding free energy threshold -20kcal/mol ; gap opening penalty -8 ; gap extension penalty -2 ; scaling parameter 4 for complementary nucleotide match score. The output of results were analyzed by using following conditions as (1) Counting nucleotide position started from 5' end of miRNA; (2) the sequence from two to eight nucleotides having complete complementary region with respective mRNA (miRNA: mRNA) not more than two mismatch is allowed; (3) no mismatch from position 2 to 4 (5' end miRNA); (4) not more than one gap is allowed in the alignment; (5) not more than one G-U pairing allowed in the seed region (6) at least 18 nucleotide alignment between complementary region with respective mRNA (miRNA: mRNA).

Gene ontology of target genes

For better understanding of target gene functions and metabolic role in *N. fowleri*, the target genes were subjected to BLAST in AmiGo version 1.8 (<http://amigo1.geneontology.org/cgi-bin/amigo/go.cgi>), the annotated results of target genes based on sequence similarity against NCBI and SwissProt database. For each target gene best hit were selected. Gene ontology was classified into biological processes, molecular functions and cellular components with the GO terms at AmiGo.

Results & Discussion:

Computational prediction of miRNA in *N. fowleri*

Different computational approaches are used to identify miRNA in both animal and plants [16, 17, & 25]. In this study, we used existing methodology with slight modifications [21]. The collected DNA contigs of *N. fowleri* from NCBI were scanned in Einverted EMBOSS software. The resulted output hits of Einverted EMBOSS software is 54896, this contain multi-loop structures and gap formed in hairpin like structures. Einverted program inverted the sequences that can form

reverse complementary sequence to form hairpin loop like structure which contain mismatches and bulges in the stem loop part. To minimize the gaps and multi-loop structures, we applied the condition by writing the Java script to retrieve the duplexes greater than or equal to (\geq) 15 base pair in length and

separated by less than or equal to (\leq) 40 nucleotides, this would narrow the result to 12821. Since the miRNA length varies [31], we have collected the sequences of 60-120 nucleotides length which come to 8603.



Figure 2: Sequences alignment of novel miRNA of *N. fowleri* and the mature miRNA from miRBase: **A)** sequence form 5' stem of miRNA; **B)** sequence form 3' stem of miRNA; prd-Panagrellus redivivus, oan-Ornithorhynchus anatinus, ath-Arabidopsis thaliana, ssc-Sus scrofa, dre- Danio rerio, ssa-Salmo salar, eca-Equus caballus.

The minimal folding free energy (MFE) plays an important role to determine the secondary structure of RNA. We identified the

secondary structure of RNA by using the program called RNA fold, this software works based on algorithm called Vienna

RNA package [32]. Individual miRNA secondary structures were calculated by using MFE which is available in RNA fold software. miRNA sequences have lower folding free energy than that of shuffled sequences, this characteristic of miRNA allows the formation of stable secondary structure [33]. However MFE values also depend on length of the RNA. The free energy range considered in this study would be less than or equal to -20 to -40 kcal/mol. The following requirements should satisfy to select precursor miRNA candidates: (1) RNA sequence fold with appropriate structure to form hair pin like secondary structure; (2) the mature miRNA sequence is present in any one of the arm; (3) not more than five mismatches between predicted mature miRNA and the opposite miRNA* sequence in the hairpin structure; (4) no breakage or loop in mature miRNA and miRNA* sequences; (5) at least 16 base pair should be present in miRNA stem loop structure. Manually, we analyzed each miRNA candidate to satisfy the above criteria and were retained.

We further narrowed down the miRNA candidates by filtering the GC content. The overall GC content of *N. fowleri* genome is 37%, in our analyses we retained GC content of 30-60% which reduced the data to 2201. We ran BlastX 2.2.30 in order to remove the protein coding gene in *N. fowleri* which narrow down the data to 1468. Then, according to miPred software we characterized the pre-miRNA candidates to identify the real miRNA and to remove pseudo miRNA and not real precursor miRNA with the confidence greater than and equal to 70%. After performing miPred software out of 1468 only 288 miRNA were retained. However, few miRNA copies in *N. fowleri* genome sequences found to be repeated, two to four copies present in same contigs ID, are eliminated. Remaining miRNA data were further analyzed to remove repeated elements in the sequences by using the tool called RepeatMasker. The software searches for repetitive sequence by aligning input genomic sequence against rebase by performing alignment program with cross_match as search engine [28]. Dataset decreases to 246 miRNA candidates. Finally, MatureBayes tool (<http://mirna.imbb.forth.gr/MatureBayes.html>) is used to identify mature miRNA sequence in miRNA precursor [29]. Out of 246 miRNAs one miRNA was not able form mature sequence which reduced the data to 245 novel miRNAs.

Characterization of *N. fowleri* miRNA

Previous reports have shown that some of the features of pre-miRNA help us to identify conserved and non-conserved miRNAs; miRNA are conserved during the evolution from worm to human based on this study. We collected all the published mature miRNAs from miRBase [34] by performing BlastN search using BioEdit software [35] with lower e-value (0.001). Out of 245 five miRNAs were conserved. Gap in the alignment is due to insertion or deletion of the nucleotides. Alignments were performed using both the strand of 3' stem and 5' stem mature miRNA/miRNA* of all the putative novel miRNA sequences. Hits were collected based on the alignment having fewer than three mismatches between the sequences. Separated alignments are studied for the mature sequences in 5' and 3' stem (Figure 2). In our study, we were able to identify some of the features of precursor miRNA, such as: (a) MFE, range from -20 to -40.90 kcal/mol, respectively Table 2 (Available with authors) (b) GC content (c) minimal free energy of the thermodynamics ensemble (d) adjusted minimal free energy (AMFE). AMFE is calculated by (MEF/Length of RNA sequence) ×100; minimal free energy index (MFEI) calculated

by AMFE/ (G+C) % all these features are calculated for each miRNA candidates. MFEI is an important criterion to distinguish miRNA from other types of RNA (coding and non-coding). For majority of precursor miRNAs MFEI was more than 0.85 with an average of 0.97 for other types of RNA like tRNA (0.64), rRNA (0.59) and mRNA (0.65) [36]. Along with these, we also calculated the percentage of individual nucleotides A%, G%, C% and U% in the precursor miRNA candidates (Table 2 - available with authors). In miRNA the percentage of G and C is less than the A and U [36]. Previous studies, shows that U is the predominant nucleotide presented at first position of 5' in majority of mature miRNAs [26] which follows the same for *N. fowleri*. Out of 245 novel miRNA 44.25% have U nucleotide in first position, only 33.2%, 14.6% and 7.75% have first position of nucleotide at A, C and G, respectively. These suggest that position of nucleotide in miRNA play an important role in identifying the mature sequences and binding the site in target mRNA.

Target prediction

Prediction of miRNA in plants and animals follows similar features such as thermodynamics stability of pre-miRNA, hairpin loop structures, MFE, number of mismatches, bulge sizes, etc. Several computational approaches have been used to predict and to identify the miRNA and their targets in both plants and animals [16-18 21, & 25]. miRNA regulates the mRNA by forming complementary binding site. From the 5' end of miRNA the nucleotides position 2-8 is called "seed region". This region is the most important region to form complementary binding site to the target transcript mRNA [37]. However, in plants the binding between miRNA and its target mRNA happens at a perfect or very near perfect complementary binding site [38]. Whereas in most of the animals few nucleotides are unbound which makes complementary binding site of miRNA to its target less tight [15, 39]. This happen due to the presence of more mismatches and wobbles bases (i.e. G-U) between the complementary binding sites.

In our study to identify the target in *N. fowleri*, we used miRanda software version 3.3a. A set of parameters was followed as described in the Methodology. From the resulted outputs, we achieved only 30 target genes after screening the condition which is described in method. The identified target genes play a vital role in various biological activities especially in mitochondria as a component of the respiratory chain, oxidation-reduction process, electron transport chain and apoptosis (Table 1 in supplementary material). All these activities play a key role in cellular growth and development.

Target Gene annotation

To understand the miRNA target genes in *N. fowleri*, individual target genes sequences were analyzed using AmiGo Software version 1.8 (<http://amigo1.geneontology.org/cgi-bin/amigo/go.cgi>). The predicted result showed that most of the target genes are involved in oxidation-reduction process, dehydrogenase activity, electron transport chain in mitochondria etc., this suggests that most of the target genes involved in mitochondria biological processes Table 3 (see supplementary material).

Conclusion:

N. fowleri is a free living amoeba that acts as a human pathogen causing PAM. Therefore, it is of interest to understand its

molecular pathogenesis using miRNA. Hence, we report 245 predicted miRNAs from *N. fowleri*. Out of that five miRNAs show high homology with mature sequences in the miRBase. Predicted miRNA sequences shows that U is the predominant nucleotide present in the precursor and mature sequences at the 5' end; this is one of the features of miRNA. Gene annotation shows that target gene functions are mainly involved in mitochondrial regulation. This data provides insights to design experimental approach for understanding regulatory mechanism.

References:

- [1] De Jonckheer J & Voorde H, *Am J Trop Med Hyg.* 1977 **26**: 10 [PMID: 842770]
- [2] De Jonckheere JF, *Infect Genet Evol.* 2011 **11**: 1520 [PMID: 21843657]
- [3] De Jonckheere JF, *Protist.* 2004 **155**: 89 [PMID: 15144061]
- [4] Yoder JS *et al.* *Clin Infect Dis.* 2012 **55**: e79 [PMID: 22919000]
- [5] Kadlec V, *Folia Parasitol (Praha).* 1975 **22**: 317 [PMID: 4366]
- [6] Visvesvara GS *et al.* *FEMS Immunol Med Microbiol.* 2007 **50**:1 [PMID: 17428307]
- [7] Carter RF, *J Pathol.* 1970 **100**: 217 [PMID: 4989229]
- [8] Siddiqui R & Khan NA, *PLoS Negl Trop Dis.* 2014 **8**: e3017 [PMID: 25121759]
- [9] Herman EK *et al.* *J Eukaryot Microbiol.* 2013 **60**: 179 [PMID: 23360210]
- [10] Burri DC *et al.* *Microbiology.* 2012 **158**: 2652 [PMID: 22878396]
- [11] Ullu E *et al.* *Cell Microbiol.* 2004 **6**:509 [PMID: 15104593]
- [12] Ambros V, *Cell.* 2001 **107**: 823 [PMID: 11779458]
- [13] Shimoni Y *et al.* *Mol Syst Biol.* 2007 **3**: 138 [PMID: 17893699]
- [14] Bartel DP, *Cell.* 2004 **116**: 281 [PMID: 14744438]
- [15] Ambros V, *Nature.* 2004 **431**: 350 [PMID: 15372042]
- [16] Brown JR & Sanseau P, *Drug Discov Today.* 2005 **10**: 595 [PMID: 15837603]
- [17] Lai EC *et al.* *Genome Biol.* 2003 **4**: R42 [PMID: 12844358]
- [18] Adai A *et al.* *Genome Res.* 2005 **15**: 78 [PMID: 15632092]
- [19] Lim LP *et al.* *Science.* 2003 **299**: 1540 [PMID: 12624257]
- [20] Lim LP *et al.* *Genes Dev.* 2003 **17**: 991 [PMID: 12672692]
- [21] De S *et al.* *Exp Parasitol.* 2006 **113**: 239 [PMID: 16515787]
- [22] Zhao T *et al.* *Genes Dev.* 2007 **21**: 1190 [PMID: 17470535]
- [23] Hwang HW & Mendell JT, *Br J Cancer.* 2006 **94**: 776 [PMID: 16495913]
- [24] Croce CM & Calin GA, *Cell.* 2005 **122**: 6 [PMID: 16009126]
- [25] Jones-Rhoades MW & Bartel DP, *Mol Cell.* 2004 **14**: 787 [PMID: 15200956]
- [26] Zhang B *et al.* *Gene.* 2009 **443**: 100 [PMID: 19422892]
- [27] Jiang P *et al.* *Nucleic Acids Res.* 2007 **35**: W339 [PMID: 17553836]
- [28] Tarailo-Graovac M & Chen N, *Curr Protoc Bioinformatics.* 2009 **4**: 10 [PMID: 19274634]
- [29] Gkirtzou K *et al.* *PLoS One.* 2010 **5**: e11843 [PMID: 20700506]
- [30] Griffiths-Jones S *et al.* *Nucleic Acids Res.* 2008 **36**: D154 [PMID: 17991681]
- [31] Starega-Roslan J *et al.* *Nucleic Acids Res.* 2011 **39**: 257 [PMID: 20739353]
- [32] Hofacker IL, *Nucleic Acids Res.* 2003 **31**: 3429 [PMID: 12824340]
- [33] Bonnet E *et al.* *Bioinformatics.* 2004 **20**: 2911 [PMID: 15217813]
- [34] Griffiths-Jones S *et al.* *Nucleic Acids Res.* 2006 **34**: D140 [PMID: 16381832]
- [35] Hall TA, *Nucleic Acids Symposium Series.* 1999 **41**: 95
- [36] Zhang BH *et al.* *Cell Mol Life Sci.* 2006 **63**: 246 [PMID: 16395542]
- [37] Brennecke J *et al.* *PLoS Biol.* 2005 **3**: e85 [PMID: 15723116]
- [38] Rhoades MW *et al.* *Cell.* 2002 **110**: 513 [PMID: 12202040]
- [39] Rehmsmeier M *et al.* *RNA.* 2004 **10**: 1507 [PMID: 15383676]

Edited by P. Kanguane

Citation: Padmashree & Ramachandra Swamy, *Bioinformation* 11(12): 550-557 (2015)

License statement: This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License.

Supplementary material:

Table 1: List of putative miRNAs of *N. floweri* and their target genes. Query: gene sequence; Ref: miRNA sequence

miRNA Names	Location of miRNA	Targets	Gene Names	Binding complementary	Score	Region	Energy
miR-10	3'	cds_AFP72327	nad2	Query: 3' aaUAAUCGUUGACGUCGACAc 5' : :	131	863 to 883	-21.389 kCal/Mol
miR-10	5'	cds_AFP72302	sdh2	Ref: 5' gtATTGG-ATCTATGGTTGTGg 3' Query: 3' cggcGUCGACGGUUUA-ACGGUg 5' : :	138	330 to 352	-20.240 kCal/Mol
miR-12	5'	cds_AFP72327	nad2	Ref: 5' ttttCCGTTGCCACATATGCCAg 3' Query: 3' cgagAUAAGUAGU-UCGCAUAGu 5' : : : : :	150	1134 to 1156	-20.350 kCal/Mol
miR-23	3'	cds_AFP72295	atp1	Ref: 5' taagTATTTATCGTTGTGTATCa 3' Query: 3' cuUCUUCU-UGUUCUCAACCCu 5' :	140	1546 to 1568	-20.680 kCal/Mol
miR-33	3'	cds_AFP72294	rps7	Ref: 5' agAAAAGATTTAAGAGCTTGGGt 3' Query: 3' guCUCGCUAAGUUUGAGAGAAg 5' : : :	156	201 to 222	-27.059 kCal/Mol
miR-39	5'	cds_AFP72310	rps2	Ref: 5' caGAGCGATTAGAATTTTCTTt 3' Query: 3' uaAGCGAGGUAAGUAG-UAAAAc 5' : :	148	764 to 786	-20.129 kCal/Mol
miR-42	5'	cds_AFP72337	nad9	Ref: 5' taTCGTTTTTATCTTCGATTTTg 3' Query: 3' gucGUUCGUUAGUGAGUCUUUu 5' : : : : : :	151	504 to 525	-20.209 kCal/Mol
miR-51	3'	cds_AFP72306	cob	Ref: 5' cctTAGCCAGTCGTTTAGAAAa 3' Query: 3' ggaGGUCUUACGAGGCCCAAc 5' : : :	143	1223 to 1244	-21.139 kCal/Mol
miR-54	5'	cds_AFP72295	atp1	Ref: 5' ttaTCGGAATGGTTATGGGTTt 3' Query: 3' ucUCCUGGUCAGGUUCGUGGAg 5' : : :	144	406 to 427	-23.910 kCal/Mol
miR-83	5'	cds_AFP72320	rps14	Ref: 5' agAGGCCTGTAAGAGTACCTt 3' Query: 3' uuGUAAUUAUUGUUCUCCACa 5' :	171	189 to 209	-22.450 kCal/Mol
miR-98	5'	cds_AFP72295	atp1	Ref: 5' ttCATATAGACGA-AGAGGTGt 3' Query: 3' cuGAACAGCUUCUUACCAAUu 5' : : : : :	156	92 to 113	-20.360 kCal/Mol
miR-105	3'	cds_AFP72324	rps13	Ref: 5' taTTCGTAGGGGAAGTGGTTAa 3' Query: 3' acGAAUGGUCCUAGUGUACAGu 5' : : :	132	118 to 139	-20.690 kCal/Mol
miR-109	3'	cds_AFP72303	nad11	Ref: 5' gaTTTTTCAGGTCTTCATGTCa 3' Query: 3' cgaCUCUUUAUUGU-AGAGACAu 5' : : :	123	374 to 396	-20.709 kCal/Mol
miR-106	5'	cds_AFP72309	nad6	Ref: 5' ttaGAGAGATGAAACGTTCTGTa 3' Query: 3' cgAAGAAC-UAAAGUUAUGAGCu 5' 	144	376 to 398	-20.080 kCal/Mol
miR-115	5'	cds_AFP72309	nad6	Ref: 5' ttTCTAGTATTTCTTTACTCGa 3' Query: 3' ucUGAAAGGGUACUACGUAGu 5' : : :	120	533 to 554	-24.160 kCal/Mol
miR-126	5'	cds_AFP72332	nad7	Ref: 5' aaGGCTATTCCATGAGACATCa 3' Query: 3' uugUGUGGAGGUGUGAGAGCa 5' : :	143	1010 to 1031	-20.059 kCal/Mol
miR-129	3'	cds_AFP72298	orf164	Ref: 5' gtgAGACAGGAGTATTACTCGt 3' Query: 3' uuCUUCUCUCGUGGUGGUGGuc 5' : : :	140	46 to 67	-21.480 kCal/Mol
miR-136	5'	cds_AFP72335	rps4	Ref: 5' gaGTGGAGAGAATTATTACCAt 3' Query: 3' uuGCUCAGAAGGAGGAGAGUa 5' : : :	132	894 to 915	-21.320 kCal/Mol
miR-143	3'	cds_AFP72300	cox1	Ref: 5' aaCTAGTCTTTTTTCATCTCAc 3' Query: 3' ucgaCCU-UUCUUCUUUACCAAu 5' : : : :	126	678 to 700	-20.910 kCal/Mol
miR-145	3'	cds_AFP72338	cox3	Ref: 5' agtaGGAGGAGAGATGTGGTTt 3' Query: 3' ggGAACGUACUUCUAUCCGUAa 5' : : : :	152	467 to 488	-20.520 kCal/Mol
miR-149	3'	cds_AFP72332	nad7	Ref: 5' taTTAGCGTCAGGGTGGGCATg 3' Query: 3' ucUAGUAAGUAGUUUGCGUACu 5' :	147	54 to 74	-20.250 kCal/Mol
miR-166	5'	cds_AFP72332	nad7	Ref: 5' acAACATCCATC-TGCCGATGg 3' Query: 3' guUGUGACCUUAGUUUA-CACUu 5' : : :	132	86 to 108	-21.740 kCal/Mol
				Ref: 5' taATTTGGAATTAGATGGTGAA 3'			

miR-172	5'	cds_AFP72333	yejR	Query: 3' agAAGAUCCGAGGGAGAGAAAa 5' :	160	446 to 467	-21.889 kCal/Mol
miR-179	5'	cds_AFP72305	nad5	Ref: 5' gtTACTGGGCCAATTCTCTTTt 3' Query: 3' caCAAGAGGGAAUAAAUGUCCu 5' :	152	1161 to 1182	-21.450 kCal/Mol
miR-179	5'	cds_AFP72331	Orf504	Ref: 5' ggGTTTTCCTTTTTTTTCAGGt 3' Query: 3' cacAAGAGGGA-AUAAAUGUCCu 5' :	139	71 to 93	-21.330 kCal/Mol
miR-189	5'	cds_AFP72337	nad9	Ref: 5' acaTACTTCCTATGTATACAGGg 3' Query: 3' ccGUAACU-UUUUCUUAAGGGUu 5' 	160	417 to 439	-21.700 kCal/Mol
miR-205	3'	cds_AFP72295	atp1	Ref: 5' ccCATTGAGAAAAGATTTCCCA 3' Query: 3' ggGUAGAGGAGUAGUAGUAGUu 5' :	159	1076 to 1096	-21.490 kCal/Mol
miR-205	3'	cds_AFP72335	rps4	Ref: 5' gaCAAATTTTC-TTATCATCAc 3' Query: 3' ggGUAGAGGAGUAGUAGUAGUu 5' :	151	629 to 649	-21.570 kCal/Mol
miR-217	5'	cds_AFP72332	nad7	Ref: 5' ttCCTTTTTC-TCATTATCAg 3' Query: 3' uuGAAGUGCUCAGCU-UUUUACc 5' 	128	549 to 571	-21.000 kCal/Mol
miR-229	3'	cds_AFP72300	cox1	Ref: 5' ttCTTCTAGAGTGGATGAAATGg 3' Query: 3' aaACUCUCUUCUCCUCCUCCu 5' :	164	670 to 691	-26.420 kCal/Mol
				Ref: 5' taTGATCCAGTAGGAGGAGAg 3'			

Table 3: miRNA target genes annotations in detail.

miRNA Name	Target Genes	Molecular function	Biological process	Cellular component
miR-10	cds_AFP72327	NADH dehydrogenase (quinone) activity (GO:0050136)	ND	ND
miR-10	cds_AFP72302	electron carrier activity (GO:0009055), succinate dehydrogenase activity (GO:0000104)	Tri-carboxylic acid cycle (GO:0006099)	plasma membrane succinate dehydrogenase complex (GO:0045282)
miR-12	cds_AFP72327	ND	ND	ND
miR-23	cds_AFP72295	ND	ND	ND
miR-33	cds_AFP72294	ND	ND	ND
miR-39	cds_AFP72310	ND	ND	ND
miR-42	cds_AFP72337	ND	ND	ND
miR-51	cds_AFP72306	ND	ND	ND
miR-54	cds_AFP72295	proton-transporting ATP synthase activity, rotational mechanism (GO:0046933)	ATP synthesis coupled proton transport (GO:0015986)	plasma membrane proton-transporting ATP synthase complex, catalytic core F(1) (GO:0045262)
miR-83	cds_AFP72320	ND	ND	ND
miR-98	cds_AFP72295	ND	ND	ND
miR-105	cds_AFP72324	ND	ND	ND
miR-109	cds_AFP72303	iron-sulfur cluster binding (GO:0051536), NADH dehydrogenase (ubiquinone) activity (GO:0008137)	cellular respiration (GO:0045333)	mitochondrial respiratory chain complex I (GO:0005747)
miR-106	cds_AFP72309	ND	ND	ND
miR-115	cds_AFP72309	ND	ND	ND
miR-126	cds_AFP72332	ND	ND	mitochondrion (GO:0005739)
miR-129	cds_AFP72298	ND	ND	ND
miR-136	cds_AFP72335	ND	ND	ND
miR-143	cds_AFP72300	ND	ND	mitochondrion (GO:0005739)
miR-145	cds_AFP72338	ND	ND	ND
miR-149	cds_AFP72332	NAD binding (GO:0051287), NADH dehydrogenase (ubiquinone) activity (GO:0008137), NADH dehydrogenase activity (GO:0003954), oxidoreductase activity (GO:0016491), protein binding (GO:0005515), quinone binding (GO:0048038)	macropinocytosis (GO:0044351), oxidation-reduction process (GO:0055114)	mitochondrion (GO:0005739), respiratory chain (GO:0070469)
miR-166	cds_AFP72332	NADH dehydrogenase (ubiquinone) activity (GO:0008137)	ND	ND
miR-172	cds_AFP72333	catalytic activity (GO:0003824)	cytochrome complex assembly (GO:0017004), oxidation-reduction process (GO:0055114)	ND
miR-179	cds_AFP72305	ND	ND	mitochondrion (GO:0005739)

miR-179	cds_AFP72331	ND	ND	ND
miR-189	cds_AFP72337	ND	ND	mitochondrion (GO:0005739)
miR-205	cds_AFP72295	ATP binding (GO:0005524) hydrolase activity, acting on acid anhydrides, catalyzing transmembrane movement of substances (GO:0016820), nucleotide binding (GO:0000166), proton-transporting ATP synthase activity, rotational mechanism (GO:0046933), proton-transporting ATPase activity) rotational mechanism:GO:0046961	ATP biosynthetic process (GO:0006754), ATP hydrolysis coupled proton transport (GO:0015991), ATP metabolic process (GO:0046034), ATP synthesis coupled proton transport (GO:0015986), ion transport (GO:0006811), macropinocytosis(GO:0044351), proton transport (GO:0015992), transport (GO:0006810)	membrane (GO:0016020), mitochondrial inner membrane (GO:0005743), mitochondrion (GO:0005739), proton-transporting ATP synthase complex, catalytic core F(1) (GO:0045261), proton-transporting two-sector ATPase complex catalytic domain (GO:0033178)
miR-205	cds_AFP72335	ND	ND	ND
miR-217	cds_AFP72332	ND	ND	mitochondrion (GO:0005739)
miR-229	cds_AFP72300	ND	ND	mitochondrion (GO:0005739)

ND: No Biological Data Available