The Pharma Innovation Journal 2017; 6(9): 543-548

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277- 7695
ISSN (P): 2349-8242
NAAS Rating 2017: 5.03
TPI 2017; 6(9): 543-548

www.thepharmajournal.com
Received: 17-07-2017
Accepted: 18-08-2017

Drushti H. Bhatt
Department of Botany, Bioinformatics and Climate Change Impacts Management, University School of Sciences, Gujarat University, Ahmedabad, Gujarat, India

Neha Jha
Department of Botany, Bioinformatics and Climate Change Impacts Management, University School of Sciences, Gujarat University, Ahmedabad, Gujarat, India

Kaid Johar SR
Department of Zoology, Gujarat University, Ahmedabad, GJ 380009, India

Himanshu A Pandya
Department of Botany, Bioinformatics and Climate Change Impacts Management, University School of Sciences, Gujarat University, Ahmedabad, Gujarat, India

Correspondence
Himanshu A Pandya
Department of Botany, Bioinformatics and Climate Change Impacts Management, University School of Sciences, Gujarat University, Ahmedabad, Gujarat, India

In silico exploration of miRNA from EST data of avocado and predicting its cross-kingdom effects on human

Drushti H. Bhatt, Neha Jha, Kaid Johar SR and Himanshu A Pandya

Abstract
The Mature MicroRNAs (miRNAs) are small non-coding RNAs having 21~ nt long, which involves in major role in translation repression process or mRNA cleavage by targeting mRNA. In this study miRNAs *Persea americana* (Avocado) are identified along with their target gene, a total 16620 ESTs are taken into consideration and aligned with 325 mature miRNA of *Arabidopsis thaliana* retrieved from miRBase by using BLASTn tool. The 234 putative miRNAs were found to hybridize with human transcripts with psRNA Targets for predicting functional targets. The energy calculation and functional annotation of genes has been performed to find out potency of targets and designing of networks between set of human miRNAs and predicted genes has been carried out to find sequence similarities especially in "Seed" site of miRNAs. The predicted targets involve in many biological process like metabolic process and cellular process. This study may helpful to open gate for cross-species regulation and provide new insight for future in-vitro and in-vivo experiments.

Keywords: Silico exploration, miRNA, EST data, cross-kingdom

Introduction
MicroRNAs are small non-coding regulatory RNA ranging in length from 21 to 25 nucleotides. Biogenesis of miRNA is performed by action of Dicer and RNase on the transcripts which leads to formation of hair pin structure called precursors miRNA (Pre-miRNA). Pre-miRNAs are processed further to form mature miRNAs [1]. The only difference between plant and animal miRNA biogenesis is in the steps of nuclear processing and export. As in animals where pre-miRNA is cleaved by two different enzymes, first inside and second outside the nucleus, both cleavages of the plant miRNA is performed by a Dicer homolog, called Dicer-like1 (DL1). DL1 is only located in the nucleus of plant cells, which indicates that both reactions take place inside the nucleus. Before plant miRNA:miRNA* duplexes are transported out of the nucleus, its 3' overhangs are methylated by a RNA methyl transferase called Hua-Enhancer1 (HEN1). The duplex is then transported out of the nucleus to the cytoplasm by a protein called Hasty (HST), an exportin 5 homolog, where they disassemble and the mature miRNA is incorporated into the RNA Inducing Silencing Complex (RISC) [2]. The miRNAs recognizes the complementary target site in 3` untranslated region(UTR) of messenger RNA (mRNA) to regulate genes at post transcriptional level [1]. The miRNAs are present in all kind of living organisms like human, plant and viruses etc and are involve in control of many fundamental cellular and physiological processes such as proliferation, stem cell maintenance, metabolic and signalling pathways, cell death, cellular differentiation, tissue development, etc [3].

Avocado is a fruit avocado tree known to human from at least 10,000 years. Though the avocado plant is native to South Central Mexico, it is now grown at many parts of world and consumed throughout the world. The health benefits of avocado include protection from cardiovascular diseases and diabetes, treating osteoarthritis, reduces the risk of cancer, liver damage and Vitamin K deficiency-related bleeding. It not only helps in maintaining blood sugar and has antioxidant properties but also increases circulation, boost cognitive abilities and build stronger bones. Avocados are also good source of monounsaturated fatty acids, nutrients, vitamins, minerals, calcium, iron, magnesium, potassium, copper, manganese, phosphorus, zinc, etc. All above compounds of avocado has been well studied but miRNAs of avocado and their effect on the human body in terms of cross kingdom gene expression regulation is not studied.

In the present study, we have identified miRNAs of avocado from the possible information available online and have predicted potential and functional miRNAs from the EST data with the help of various bioinformatic tools.

Result and Discussions

Creation of dataset and local alignment with reference sequences

Expressed sequence tags (ESTs) are short DNA sequences (200–500 nucleotides) generated by sequencing the 5′ and/or 3′ ends of cDNAs that are subsequently clustered and counted [4].

The identification of miRNA through EST data has following advantages (1) No need to have special software and we can be used to discover miRNAs from any species if their EST are already registered. (2) As EST are originated from transcribed sequences, its analysis provided evidence for miRNA expression directly [5].They may be used to identify gene transcripts and are instrumental in gene discovery and gene sequence determination. There are 16620 EST sequences available in dbEST (NCBI) till date and 325 matured miRNAs of *A. thaliana* have been retrieved from the miRBase. ESTs and miRNA of *A. thaliana* were locally aligned with BLASTn tool where we have found 302 miRNA of avocado. Filtration of these sequences was performed using parameters like word size 7 (Seven), number of threads 3 (Three) and maximum target sequences 6 (Six) [Table 1].

Identification of potential miRNA and their targets

After filtration, we have achieved 234 unique miRNAs of avocado and made result set through Bedtools2 for further investigations. The human targets of predicted miRNA have been identified by using bioinformatics tool named psRNA Target (Updated version). Out of 234 unique miRNA only 79 miRNA were found which have potential to bind with human genes [Table 2]. We have found 160 repetitive human targets which were subjected to calculation of hybridizing energy by IntaRNA [Table 2] and were analysed for inter-gene interactions by STRING tool [Figure 1].

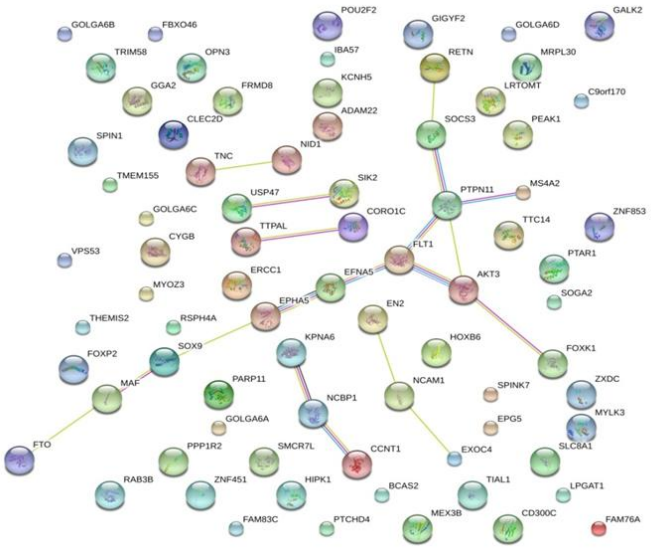

Fig 1: Avocado microRNA target Protein - Protein interaction derived using STRING Tool

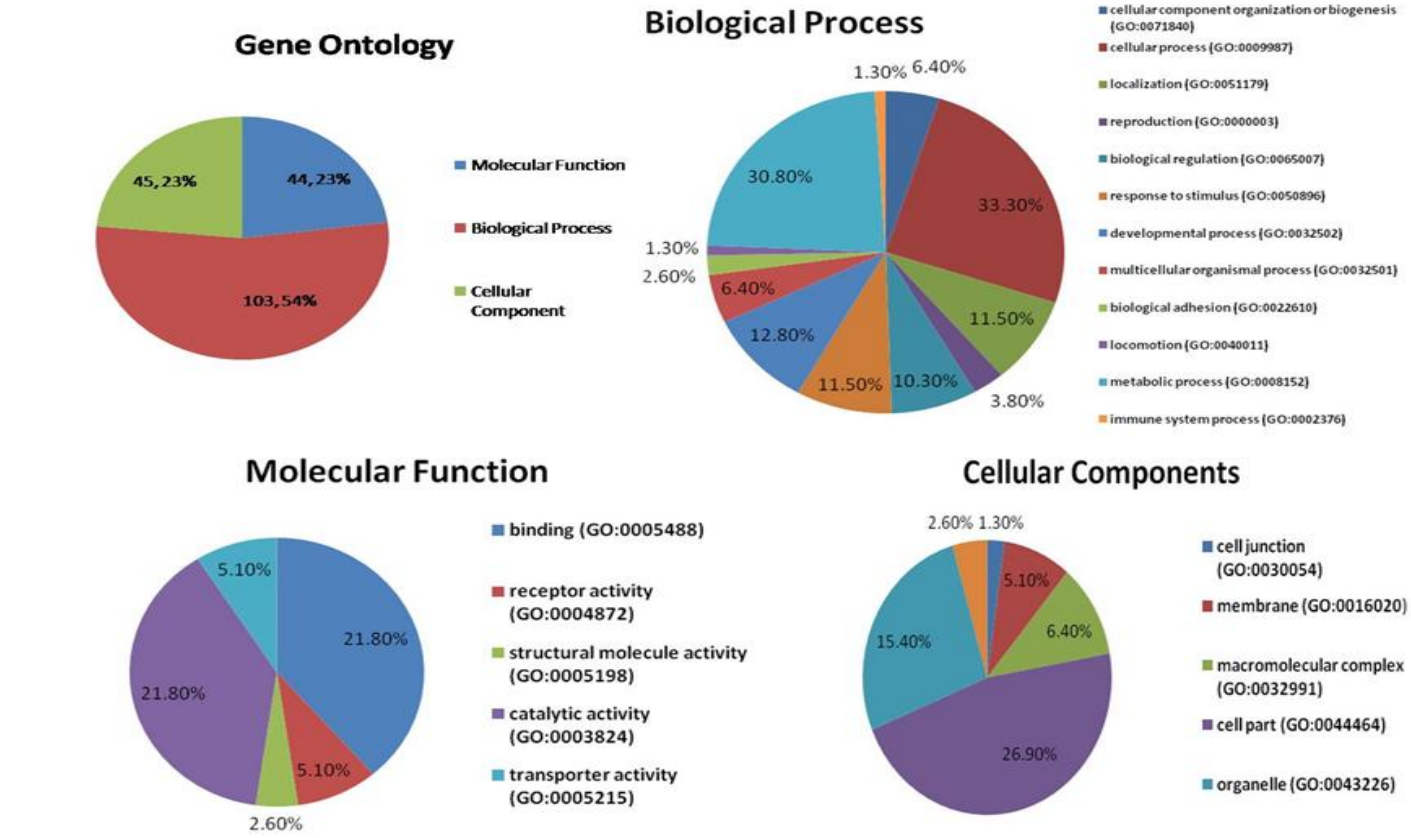

Fig 2: List of comprehensive functional analysis of targets with respect to biological process, cellular components and molecular function. In gene ontology pie chart (upper extreme left), the first value indicates the functional hits and second represents the overall percentage

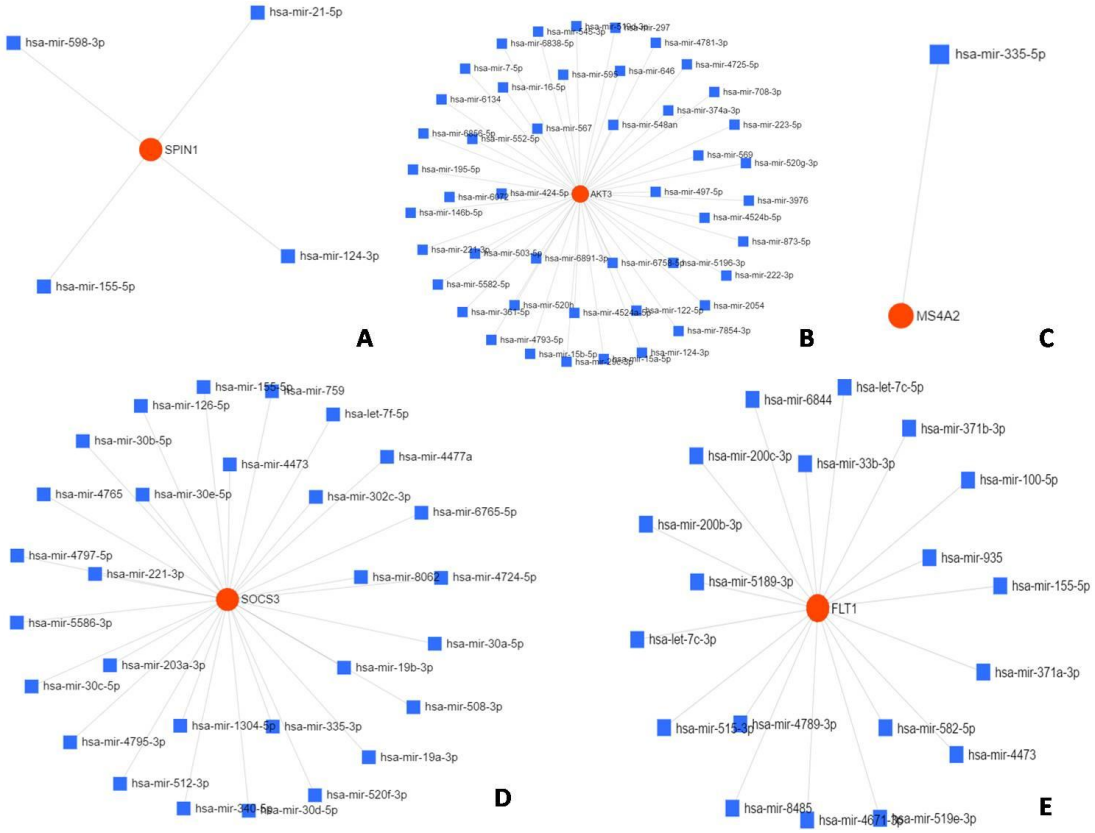

Fig 3: Network displays the reported human miRNAs with target genes. (A) SPIN1 (B) AKT3 (C) MS4A2 (D) SOSC3 and (E) FLT1

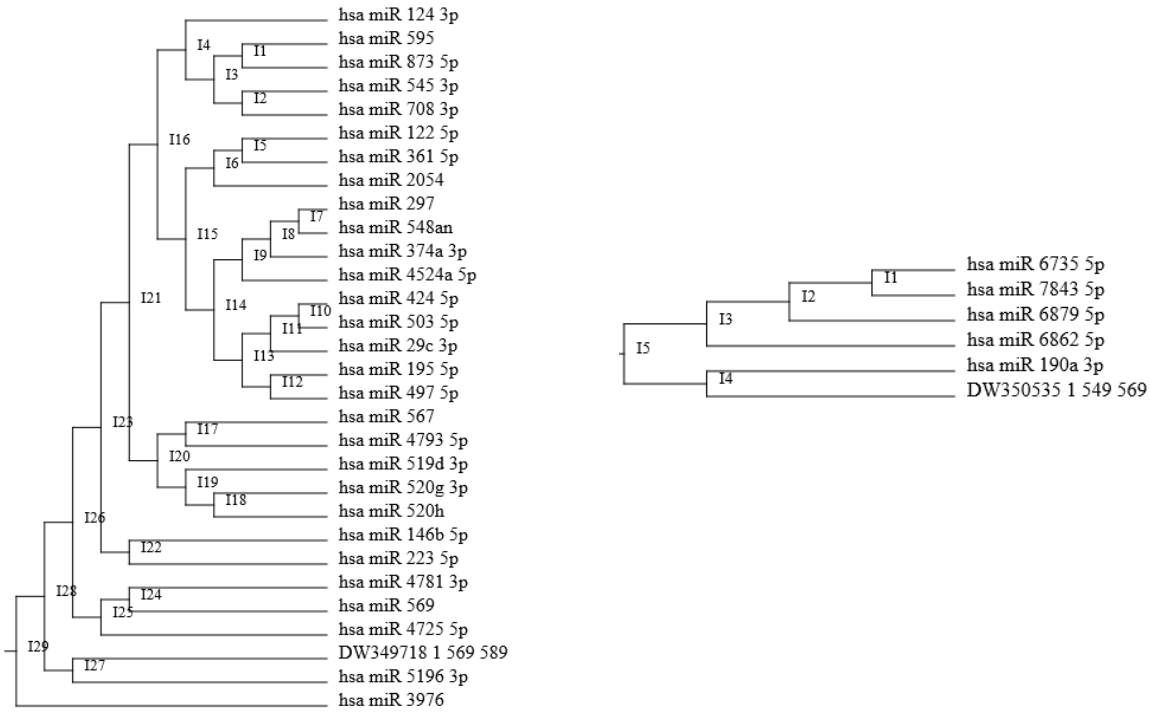

Fig 4: Multiple sequence alignment of human miRNAs along with predicted miRNAs of avocado.

Tab 1: Statistics of study

Number of Expressed Sequence Tags (dbEST)	**16620**
***Arabidopsis thaliana*miRNA Sequences (miRBase)**	**325**
Total miRNAs detected from EST (BLASTn)	**302**
Unique miRNAs Sequences (After Filtrations)	**234**
Functional miRNAs	**79**
Number of human targets (psRNATarget)	**160**
Interactive Targets (STRING)	**12**

Tab 2: List of predicted miRNAs of avocado and its human targets with hybridization score in Kcal/mol.

miRNAs (More than 18 nt)	Targets	Target-miRNA Binding	Energy (Kcal/mol)	Target_start	Target_end
CB819818.1:158-177	SOCS3	5' 3' CAGUUUUUUGUUUUUGUU |||||||||||||:||| GUCAAAAAACAAAAGCAA 3' 5'	-19.52410	1269	1287
CB818725.1:464-483	FLT1	5' 3' UCAUCAUCAUCAUCAUCAU ||||||||||||||||||| AGUAGUAGUAGUAGUAGUA 3' 5'	-27.95030	13	31
	CCNT1	5'-UCUA 3' AUCAUCAUCAUCAUCAU |||||||||||||||| UAGUAGUAGUAGUAGUA 3'-AG 5'	-24.54940	824	844
DW349718.1:569-589	AKT3	5' 3' GAUGGUGAUGAGGAUGAUG ||||:||||||:||||||| CUACUACUACUUCUACUAC 3' 5'	-29.48360	4545	4564
	RETN	5' 3' UGAUGAUGAUG GAUGAUG ||||||||||| ||||||| ACUACUACUAC UUACUAC 3' 5'	-26.39260	86	105
CV044668.1:299-318	MS4A2	5' 3' AUUUUUAGGCUUUGAUUU |:|||:||||||||||| UGAAAGUCCGAAACUAAA 3' 5'	-21.31840	1315	1333
DW350535.1:549-569	NCBP1	5' 3' UGA GAUGAUG GAUGAAG ||| |||||| ||||||| ACU CUACUAC UUACUUC 3' 5'	-22.03570	1529	1548
DW349718.1:572-592	SPIN1	5'- 3' UGAUGGUGAUGAUGAGAUG |||||:||||||||| || ACUACUACUACUACUUAC 3' 5'	-24.81150	2190	2209

Tab 3: List of avocado microRNA targets with chromosomal position, disease association and gene card ID.

Target keyword	Full name	Chromosomal Position	Diseases Associate with	GCID
FLT1	Fms Related Tyrosine Kinase 1	13q12.3	Anal Canal Squamous Cell Carcinoma and Eclampsia	GC13M028300
AKT3	AKT Serine/Threonine Kinase 3	1q44	Megalencephaly-Polymicrogyria-Polydactyly-Hydrocephalus Syndrome 2 and Mpph Syndrome.	GC01M243488
SOCS3	Suppressor Of Cytokine Signaling 3	17q25.3	Atopic Dermatitis and Overnutrition	GC17M078356
RETN	Resisti	19p13.2	Diabetes Mellitus, Noninsulin-Dependent and Acquired Generalized Lipodystrophy	GC19P007669
MS4A2	Membrane Spanning4-Domains A2	11q12.1	Atopy and Hand Dermatosis	GC11P060107
CCNT1	Cyclin T1	12q13.12	Hiv-1 and Ipex Syndrome	GC12M048688

Hybridizing energy calculation and functional annotation of targets

After getting 160 human targets, we have studied all the targets thoroughly and selected those targets which have interaction amongst them. Amongst them FLT1, PTPN11, MS4A2, SOCS3, RETN, AKT3, FOXK1, EFNA5, EPHA5, SOX9, MAF and FTO were interacting with each other [Figure 1]. Our idea was to select those targets which are involved in certain diseases and their expressions reflect the negative impact on the human body. Hence, we have selected SOCS3, FLT1, AKT3, MS4A2, NCBP1, RETN1 and SPIN1 for further investigation [Table 3]. Hybridization energy has calculated with IntaRNA (RNA-RNA interaction tool) and most of all selected targets have energy near -20Kcal/mol which is favourable for RNA interfaces. AKT3 and FLT1 required lesser energy (-29.48Kcal/mol and -27.97Kcal/mol respectively) compared to other targets [Table 2].

We have annotated genes through Gene Ontology and categorised them into molecular functions, biological processes and cellular components. During gene ontology study we have observed that most of targets lied into biological process especially metabolic process (30.80%) and cellular process (33.30%) [Figure 2].The network of human miRNAs for every predicted target genes has been created by miRNet software. We have observed has-miR-21-5p, has-miR-598-3p, has-miR-155-5p and has-miR-124-3p which are associated with and regulate SPIN1.Only one human miRNA, has-miR355-5p has been reported as regulator of MS4A2 gene. Therefore multiple sequence alignment of this particular gene was not been possible [Figure 3].Alignment of these lists of human miRNAs and predicted miRNAs of avocado has been performed for sequence similarity search by LocARNA and created dendrogram which showed that has-miR-5196-3p has very similar sequence with predicted miRNA of avocado (DW349718 1 569 589) which targets AKT3 and in same way has-miR-190a-3p having similarity with DW350535 1 549 569 which target NCBP1 [Figure 4].

Conclusion

EST analysis provides information for designing more experiments to understand miRNAs. The present study shows that predicted miRNAs of *Persea americana* (avocado) were found to target various genes which are involve in biological processes. The list of predicted miRNAs of avocado might be used as potential agents for treatment of diseases associated with respective targets. According to our analysis, we suggest that the ingestion of these miRNAs may have a functional impact to consuming organisms in a cross-kingdom way. Our findings may also be useful for discovering cross-species regulatory mechanism in further study.

Computational Methods

The EST sequences of *Persea americana* (avocado) has been retrieved from dbEST (NCBI) [6]. The mature miRNAs of *Arabidopsis thaliana* have been used to set reference miRNAs for prediction. We have first created database of matured miRNAs of *Arabidopsis thaliana*. The ESTs of avocado was aligned using BLASTn software with reference set of mature miRNAs of *A. thaliana* [7]. After alignment, the unique miRNAs has been achieved based on filtration criteria (E-vlaue, bit score and mismatches etc) for predicting functional miRNAs. These predicted miRNAs further hybridized with 3'-UTR of human transcripts using psRNA Targets [8]. This is to select predicted targets for further investigation. Interactions of the target proteins with other proteins have been performed using STRING tool [9]. We have also created networks of predicted targets with human miRNAs, which are already reported using miRNet software [10]. These human miRNAs were further aligned with predicted miRNA of avocado to find out sequence similarities especially in "Seed" site through LocARNA tool [11]. The functions of targets have been studied by using Gene Ontology program [12].

References

1. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. Review. PubMed PMID: 14744438. 2004; 116(2):281-97.
2. Millar AA, Waterhouse PM. Plant and animal microRNAs: similarities and differences. FunctIntegr Genomics. Epub 2005 May 5. Review. PubMed PMID: 15875226. 2005; 5(3):129-35.
3. Carthew RW, Sontheimer EJ. Origins and Mechanisms of

miRNAs and siRNAs. Cell. 2009; 136(4):642-55. doi: 10.1016/j.cell.2009.01.035. Review. PubMed PMID:19239886; PubMed Central PMCID: PMC2675692

4. Mathur EJ, Adams MW, Callen WN, Cline JM. The DNA polymerase gene from the hyperthermophilic marine archaebacterium, Pyrococcus furiosus, shows sequence homology with alpha-like DNA polymerases. Nucleic Acids Res. PubMed PMID: 1762925; PubMed Central PMCID: PMC329336. 1991; 19(24):6952.
5. Zhu QW, Luo YP. Identification of miRNAs and their targets in tea (*Camellia sinensis*). J Zhejiang UnivSci B. 2013; 14(10):916-23. doi:10.1631/jzus.B1300006. PubMed PMID: 24101208; PubMed Central PMCID: PMC3796643.
6. Boguski MS, Lowe TM, Tolstoshev CM. dbEST-database for "expressed sequence tags". Nat Genet. 1993; 4(4):332-3. PubMed PMID: 8401577.
7. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res. 2006; 34(Database issue):D140-4. PubMed PMID: 16381832; PubMed Central PMCID: PMC1347474.
8. Dai X, Zhao PX. psRNA Target: a plant small RNA target analysis server. Nucleic Acids Res. 2011 Jul;39(Web Server issue):W155-9. doi:10.1093/nar/gkr319. Epub. 2011. PubMed PMID: 21622958; PubMed Central PMCID: PMC3125753.
9. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P *et al*. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res. 2011; 39(Database issue):D561-8. doi: 10.1093/nar/gkq973. Epub 2010 Nov 2. PubMed PMID: 21045058; PubMed Central PMCID: PMC3013807.
10. Fan Y, Siklenka K, Arora SK, Ribeiro P, Kimmins S, Xia J. miRNet - dissecting miRNA-target interactions and functional associations through network-based visual analysis. Nucleic Acids Res. 2016 Jul 8;44(W1):W135-41. doi: 10.1093/nar/gkw288. Epub 2016 Apr 21. PubMed PMID: 27105848; PubMed Central PMCID: PMC4987881.
11. http://rna.informatik.uni-freiburg.de/LocARNA/Input.jsp
12. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM *et al*. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000; 25(1):25-9. PubMed PMID: 10802651; PubMed Central PMCID: PMC3037419.