BioTechnologia vol. 102 (2) • pp. 179–195 • 2021
Journal of Biotechnology, Computational Biology and Bionanotechnology RESEARCH PAPERS

http://doi.org/10.5114/bta.2021.106525

Identification of putative miRNAs from expressed sequence tags of *Gnetum gnemon* L. and their cross-kingdom targets

DEBASISH B. KRISHNATREYA [1,2], DIPANJALI RAY [1], POOJA M. BARUAH [1], BHASKAR DOWARAH [1,3], KUNTALA S. BORDOLOI [1], HEENA AGARWAL [1,4], NIRAJ AGARWALA [1]*

[1] Department of Botany, Gauhati University, Jalukbari, Guwahati Assam, India

[2] Department of Botany, Suren Das College, Kamrup, Assam, India

[3] Department of Botany, Bahona College, Jorhat, Assam, India

[4] Department of Botany, Pandit Deendayal Upadhyaya Adarsha Mahavidyalaya, Biswanath, Assam, India

Abstract

Wild edible plants are often found to be rich sources of nutrients and medicinally beneficial compounds with pharmacological activities. *Gnetum gnemon* is a nutritionally important plant and a popular food source in parts of Assam and North-East India. Various microRNAs (miRNAs) have been recently identified in many plants; however, there are no records of identification of miRNAs in any species of *Gnetum*. The prediction of miRNA-target associations in *G. gnemon* is an important step to facilitate functional genomics studies in this species. In the present study, all known miRNAs from plants available in public domain were used to search for the conserved *G. gnemon* miRNA homologues in publicly accessible expressed sequence tags (ESTs) in NCBI database. An aggregate of 20 new potential miRNAs belonging to two diverse miRNA families (miR399 and miR5021) were identified through a homology-based search by following stringent filtering criteria. To investigate the potential cross-kingdom effects of the identified miRNAs, we further identified the putative target genes of *G. gnemon* miRNAs in human transcriptome and analyzed them against the NCBI non-redundant protein database. The KEGG analysis of the target genes indicated that these genes were involved in different metabolic pathways such as caffeine metabolism, drug metabolism, and nitrotoluene degradation. The target genes of *G. gnemon* miRNAs in humans were found to be associated with various disorders of both hereditary and non-hereditary origin. These results could help to shed new light on understanding of miRNA-mRNAs functional networks in this species and its potential use as a small RNA-based therapy against some human diseases.

Key words: microRNAs, expressed sequence tags, cross-kingdom, *Gnetum gnemon*

Introduction

Gnetum is a genus belonging to the group of gymnosperms, and it grows in diverse forms and habitats such as tropical evergreen trees, shrubs, or lianas. Due to the association of the fossils of some *Gnetum* species with those of extinct scorpionflies, they have been proposed as one of the first groups of plants to be insect-pollinated (Ren et al., 2009). A unique species of *Gnetum*, namely *Gnetum gnemon*, resembles the angiosperms in terms of its vegetative structures (Bhat and Binti, 2014). Nutritionally, *G. gnemon* is very rich in minerals and proteins, which indicates the potential role of this species in providing healthy sustenance in poor provincial regions that have limited access to proteinaceous foods (Bhat and Binti, 2014). It is also known to be rich in antioxidant, tyrosinase inhibitory, antimicrobial, anti-aging, and various bioactive compounds (Barua et al., 2015; Agarwal et al., 2020). The edible parts of *G. gnemon* include young leaves, inflorescences, and tender tips, which are used as vegetables in Assam, mostly by the Karbi tribe. Seeds are often used as a snack food ingredient by processing them into sundried, flat cakes

* Corresponding author: Department of Botany, Gauhati University, Gopinath Bordoloi Nagar, Jalukbari, Guwahati Assam, 781014, India; e-mail: niraj_botany@gauhati.ac.in

called emping, which are then fried in oil before consumption (Barua et al., 2015). Among other edible parts, tender leaves are one of the most commonly traded vegetables in ethnic markets in North-East India. Leaves of *G. gnemon* are cooked with "pholo" (alkaline water) along with dried fish, which is served as a delicacy; the leaves are also eaten fried or cooked with lentil (Terangpi et al., 2013). "Melinjo" (common name for *Gnetum*) seeds have high nutritive value and are consumed as a major food item by the Indonesian population as an additional ingredient in soups (viz., "Sayur Asem") or boiled with spices and vegetables (Narayanan et al., 2015; Ahmad et al., 2018). As a foodstuff, Melinjo seeds have been found to have no toxic effects even when consumed for a long period of time (Tatefuji et al., 2014).

MicroRNAs (miRNAs) are 21-24 nucleotide long, highly conserved, non-coding small RNAs originating from a self-complimentary stem-loop precursor sequence, and they mediate transcriptional or post-transcriptional gene regulation (Nithin et al., 2015). These short RNA sequences also play a pivotal role in manifold facets of plant development and adaptive potentiality of plants to biotic and abiotic stresses (Jones-Rhoades et al., 2006; Vishwakarma and Jadeja, 2013; Bordoloi et al., 2021; Krishnatreya et al., 2021). miRNAs from plants were first recognized in early 2002 in *Arabidopsis thaliana* (Reinhart et al., 2002), following which thousands of plant miRNAs were identified in different plants and submitted to databases such as the miRBase Sequence database (http://www.mirbase.org) (Griffiths-Jones et al., 2006).

The identification of miRNAs and their target genes from plants has been widely investigated in the last few years (Pirro et al., 2019; Xu et al., 2020; Panda et al., 2014). To identify miRNAs and their target genes and to perform functional analysis, various strategies are usually used, such as genetic screening (Xun et al., 2019), construction of a small RNA library, or computational approach-based identification from ESTs (Krishnatreya et al., 2020). The computational biology-based approach serves as a useful asset to identify conserved miRNAs in organisms for which complete genome sequences are not yet available. This approach is also used to study the conservation and evolution of such miRNAs among different species (Vishwakarma et al., 2013). Identification of miRNAs from the published ESTs has been shown to be effective for revealing new miRNAs from plants that are socio-economically important; lacking significant genomic resources (Lu et al., 2010). Most of the mature miRNAs in plants are evolutionarily conserved from species to species within the plant kingdom (You et al., 2017). This conserved nature of plant miRNAs has greatly enhanced the identification of conserved miRNAs using EST sequences. There are, however, no reports of identification of miRNAs from *G. gnemon* till date. In the present study, we identified conserved miRNAs in *G. gnemon* ESTs available from the NCBI database. These identified miRNAs of *G. gnemon* may even arbitrate the cross-kingdom miRNA-mediated human gene target networks that are potentially regulated through the target gene cleavage, based on the significant target specificity between the identified miRNAs of *G. gnemon* and human gene transcripts.

Materials and methods

Reference set of miRNAs and sequence database

To identify the potential of conserved miRNAs, the sequences of 38,589 plant miRNAs previously identified from different plant species were downloaded from the miRBase database (http://www.mirbase.org, release 22.1, March 2018). These sequences were used as query sequences to perform BLAST analysis against the *G. gnemon* ESTs. Publicly available 10,725 ESTs (as of August 1, 2019) of *G. gnemon* were downloaded from the NCBI (https://www.ncbi.nlm.nih.gov/). A local database for standalone BLAST was constructed for the *G. gnemon* ESTs by using a locally installed NCBI-Blast+ v-2.2.26 command line application (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/) (Altschul et al., 1997). Non-redundant (NR) protein sequences were used from the NR protein database of NCBI (ftp://ftp.ncbi.nlm.nih.gov/blast/db/).

Identification of putative miRNAs

Sequence identification of precursor miRNAs was performed using a comparative sequence-based homologue search and a secondary structure analysis. Mature miRNA sequences were used for BLASTn search as they are evolutionarily more conserved than their precursor sequences. miRNA sequences were used as queries and subjected to NCBI-Blast+ against the *G. gnemon* EST database. A maximum of two allowed mismatches of threshold *e*-value of < 0.001 and word-size value of 7 were set for Blast+ analysis. After removing

redundant sequences, the ESTs with matched hits were subjected to a Blastx analysis with the NR protein database, and the non-protein coding sequences were retained for further analysis of RNA secondary structure using the Zuker folding algorithm by Mfold v3.5 (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form) (Zuker, 2003). The following parameters were used in defining the sequences as miRNA homologs: 1) the sequence should fold into an appropriate stem-loop secondary structure, 2) the miRNA should exist in one arm of the hairpin structure, 3) the mature miRNA and its complementary miRNA* sequence should not have more than 5 mismatches, and 4) the value of the Minimal Folding free Energy Index (MFEI) of precursor miRNA structures should not be less than 0.5 and should have a high Minimal Folding free Energy (MFE) value. MFE is the negative equivalent of the ΔG value (Zhang et al., 2006). The MFEI value was calculated using the following formula proposed by Zhang et al. (2006):

$$AMFE^{*} = (MFE \times 100) / \text{length of precursor}$$

$$MFEI = AMFE / (GC)\,\%$$

$$MFEI = [(MFE/\text{length of the pre-miRNA}) \times 100] / (GC)\%$$

* Adjusted Minimal Folding free Energy

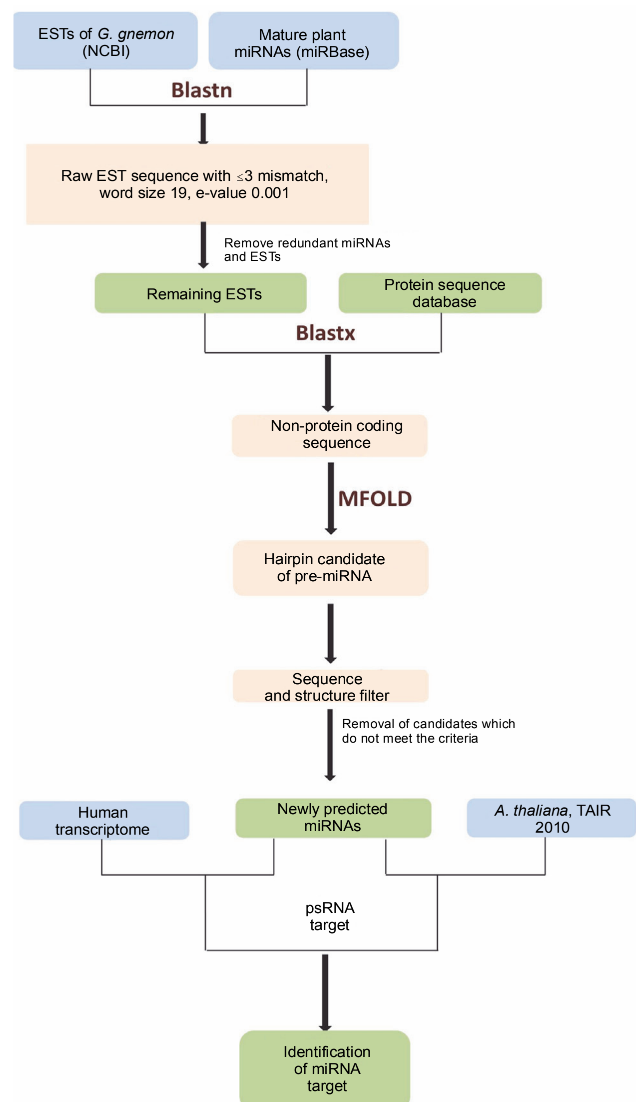

Fig. 1. Computational pipeline for the identification of putative miRNAs and their target genes.

Prediction of putative target genes

The perfect or near-perfect complementarity of miRNAs and their target mRNAs in plants has greatly simplified the identification of miRNA targets. This strategy is applied to search the targets of identified miRNAs by a homology algorithm. In the present study, *A. thaliana* transcript, TAIR version 10 available at psRNATarget (http://plantgrn.noble.org/psRNATarget/) was used as a reference for finding the targets of the candidate miRNAs as no genomic references of *Gnetum* species are available in public domain. To examine the possible cross-kingdom interaction exerted by miRNAs of *G. gnemon* on human genes, a target analysis was performed against the human transcriptome available at psRNATarget. A plant small RNA Target Analysis Server, namely psRNATarget, was used for predicting the targets of the newly identified miRNAs by using Schema V2 (2017 Release) with the maximum expectation value threshold as 3 and the remaining values set as default (Dai et al., 2018). A maximum of two mismatches were allowed in the complementary region of target genes with the miRNAs, whereas the mismatch inhibition was maintained at the 10th and 11th nucleotide positions along the aligned region. Figure 1 shows the pipeline followed for miRNA identification and target gene prediction. Information regarding the disease association and the function of human target genes was collected from GeneCards – a human gene database (https:// www.genecards.org/) and MalaCards – a human disease database (https://www.malacards.org/).

Gene Ontology, KEGG pathway, and phylogenetic analysis

Annotations of the target genes for both plant and cross-kingdom human targets were carried out using a Blastx analysis with a threshold *e*-value of 10^{-3} against

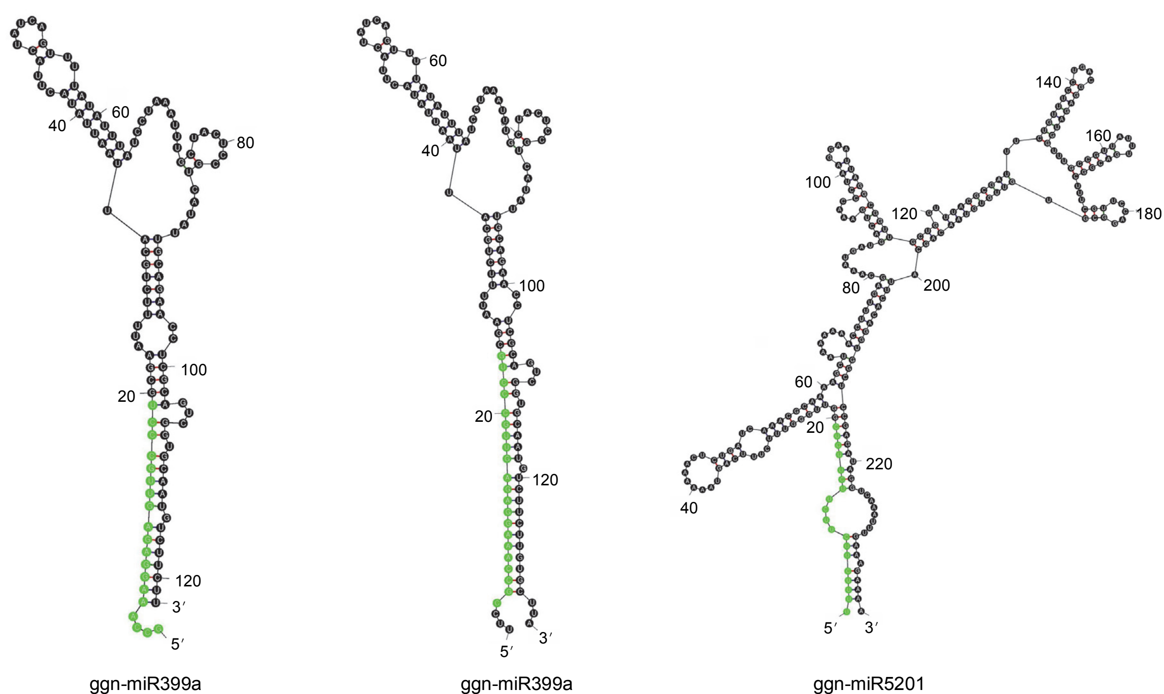

Fig. 2. Stem-loop precursors of the identified miRNAs in *G. gnemon*

Table 1. Identified putative miRNAs of *G. gnemon* from ESTs

Accession no	miRNAs	Mature miRNA sequence	PL*	ML**	(C + G)%	MFE	AMFE	MFEI
CB081553.1	sly-miR399	GCCAAAGGAGAGUUGCCCU	122	19	38.5	24.52	20.09	0.52
	tcc-miR399a	CGCCAAAGGAGAGUUGCCCUG	133	21	39.1	27.62	20.76	0.53
	ath-miR399b	GCCAAAGGAGAGUUGCCCUG	133	20	39.1	27.62	20.76	0.53
	csi-miR399b-3p	GCCAAAGGAGAGUUGCCCU	122	19	38.5	24.52	20.09	0.52
	pab-miR399c	CGCCAAAGGAGAGUUGCCCUG	133	21	39.1	27.62	20.76	0.53
	aly-miR399c-3p	GCCAAAGGAGAGUUGCCCUG	133	20	39.1	27.62	20.76	0.53
	mdm-miR399d	GCCAAAGGAGAGUUGCCCU	122	19	38.5	24.52	20.09	0.52
	csi-miR399d-3p	GCCAAAGGAGAGUUGCCCUG	133	20	39.1	27.62	20.76	0.53
	cme-miR399e	GCCAAAGGAGAGUUGCCCU	122	19	38.5	24.52	20.09	0.52
	zma-miR399e-3p	GCCAAAGGAGAGUUGCCCUG	133	20	39.1	27.62	20.76	0.53
	mes-miR399f	GCCAAAGGAGAGUUGCCCUG	133	20	38.5	27.62	20.76	0.53
	gma-miR399h	GCCAAAGGAGAGUUGCCCUG	133	20	39.1	27.62	20.76	0.53
	vvi-miR399i	CGCCAAAGGAGAGUUGCCCUG	133	21	39.1	27.62	20.76	0.53
	zma-miR399i-3p	GCCAAAGGAGAGUUGCCCUG	133	20	39.1	27.62	20.76	0.53
	osa-miR399j	GCCAAAGGAGAGUUGCCCU	122	19	38.5	24.52	20.09	0.52
	zma-miR399j-3p	GCCAAAGGAGAGUUGCCCUG	133	20	39.1	27.62	20.76	0.53
	mdm-miR399k	GCCAAAGGAGAGUUGCCCU	122	19	38.5	24.52	20.09	0.52
	mtr-miR399l	GCCAAAGGAGAGUUGCCCUG	133	20	39.1	27.62	20.76	0.53
	mtr-miR399p	GCCAAAGGAGAGUUGCCCUG	133	20	39.1	27.62	20.76	0.53
EX943654.1	ath-miR5021	UUUUCUUCUUCUUCUUCUC	240	19	41.6	58.05	24.18	0.58

* precursor length; ** mature miRNA length

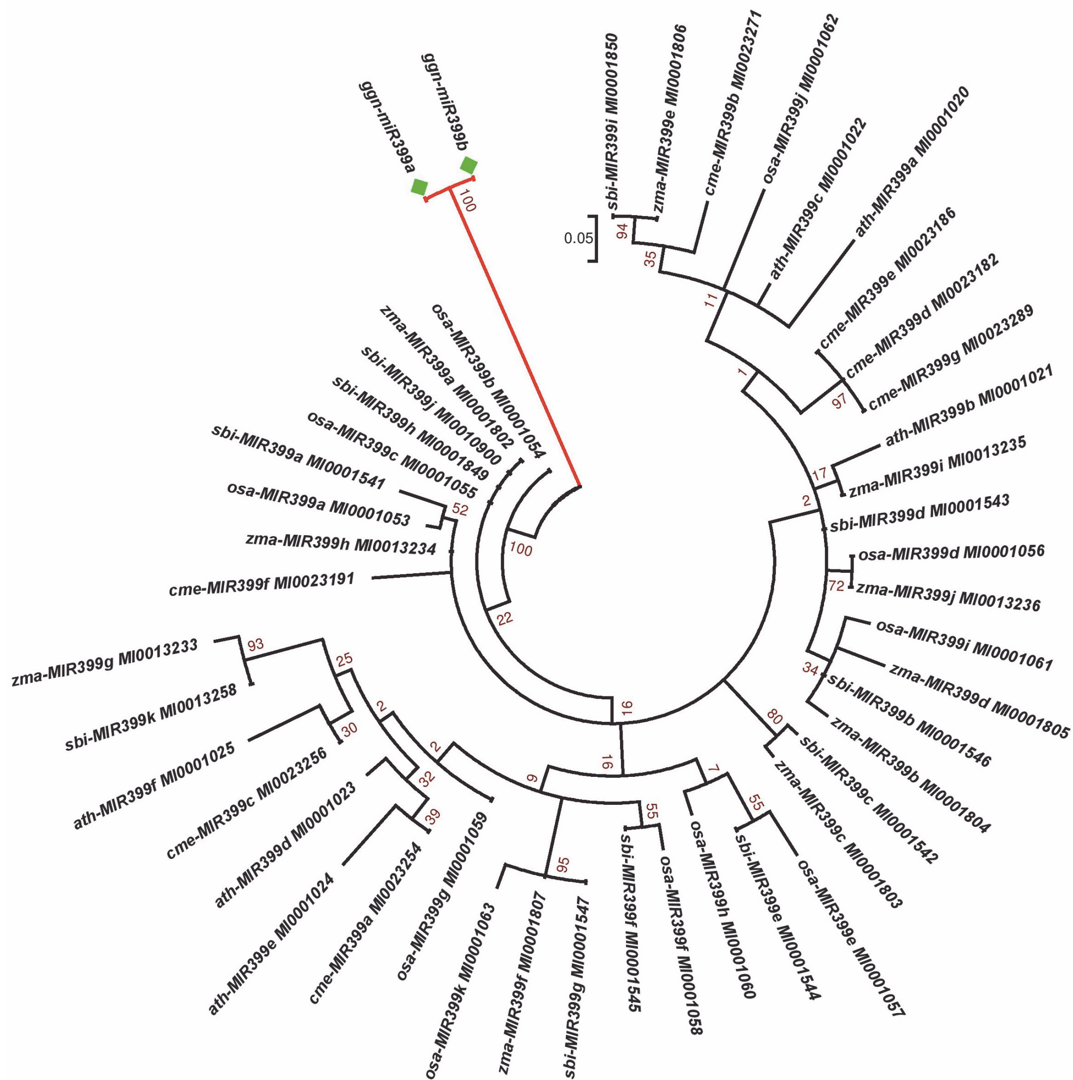

Fig. 3. Phylogenetic tree constructed for the miR399 family by using the maximum likelihood method with 500 bootstrap replicates; ggn-miR399a and ggn-miR399b form a separate cluster

the NCBI NR protein database. Blast2go version 5.2 (https://www.blast2go.com/blast2go-pro/) was used for the gene ontology and the KEGG pathway analysis of the annotated target genes in order to assess the traits that may be affected by the expression of the identified miRNAs (Conesa and Götz, 2008). The phylogenetic tree generated from the identified miRNA precursors was constructed using MEGA7. The precursor sequences of the family members of the miRNAs belonging to other plant species were downloaded from the miRBase and collated with the *G. gnemon* miRNA precursors. A multiple sequence alignment was performed using the MUSCLE algorithm, and the phylogenetic tree was generated using the Maximum likelihood method.

Results

miRNA identification and characterization

From a total of 10,725 published ESTs of *G. gnemon*, 144 ESTs showed homology with the previously deposited sequences of miRNAs in miRBase 22.1, belonging to various miRNA families. These sequences were obtained after following the criteria of Axtell and Meyers (2018) for plant miRNA annotation, wherein only those miRNAs of ≥19 nucleotides in length are filtered and retained. Further removal of redundancies in miRNAs and the ESTs yielded 142 potential miRNA sequences. A Blastx analysis of these ESTs against the NCBI NR database led to the identification of 56 sequences as non-coding

Table 2. Predicted target genes of *G. gnemon* miRNAs against *Arabidopsis thaliana*

miRNA_Acc.	Target_Acc.	Expect	Target_Desc.	Inhibition
ggn-miR5021	AT5G04180.1	0.5	alpha carbonic anhydrase 3	cleavage
ggn-miR5021	AT5G65630.1	0.5	global transcription factor group E7	cleavage
ggn-miR5021	AT3G61430.1	0.5	plasma membrane intrinsic protein 1A	cleavage
ggn-miR5021	AT1G18540.1	0.5	ribosomal protein L6 family protein	cleavage
ggn-miR5021	AT1G76580.1	0.5	squamosa promoter-binding protein-like (SBP domain) transcription factor family protein	cleavage
ggn-miR5021	AT5G56750.1	0.5	N-MYC downregulated-like 1	cleavage
ggn-miR5021	AT2G30320.1	0.5	pseudouridine synthase family protein	cleavage
ggn-miR5021	AT3G04470.1	0.5	ankyrin repeat family protein	cleavage
ggn-miR5021	AT3G18620.1	0.5	DHHC-type zinc finger family protein	cleavage
ggn-miR5021	AT5G67190.1	0.5	DREB and EAR motif protein 2	cleavage
ggn-miR5021	AT3G55820.1	0.5	fasciclin-like arabinogalactan family protein	cleavage
ggn-miR5021	AT4G21240.1	0.5	F-box and associated interaction domains-containing protein	cleavage
ggn-miR5021	AT1G12490.1	0.5	F-box associated ubiquitination effector family protein	cleavage
ggn-miR5021	AT4G39753.1	0.5	galactose oxidase/kelch repeat superfamily protein	cleavage
ggn-miR5021	AT4G38710.1	0.5	glycine-rich protein	cleavage
ggn-miR5021	AT4G26190.1	0.5	haloacid dehalogenase-like hydrolase (HAD) superfamily protein	cleavage
ggn-miR5021	AT1G03080.1	0.5	kinase interacting (KIP1-like) family protein	cleavage
ggn-miR5021	AT4G36260.1	0.5	lateral root primordium (LRP) protein-related	cleavage
ggn-miR5021	AT2G23290.1	0.5	MYB domain protein 70	cleavage
ggn-miR5021	AT5G27120.1	0.5	NOP56-like pre RNA processing ribonucleoprotein	cleavage
ggn-miR5021	AT4G13710.1	0.5	pectin lyase-like superfamily protein	cleavage
ggn-miR5021	AT3G05170.1	0.5	phosphoglycerate mutase family protein	cleavage
ggn-miR5021	AT4G31820.1	0.5	phototropic-responsive NPH3 family protein	cleavage
ggn-miR5021	AT5G07820.1	0.5	plant calmodulin-binding protein	cleavage
ggn-miR5021	AT4G32650.1	0.5	potassium channel	cleavage
ggn-miR5021	AT3G50730.1	0.5	protein kinase superfamily protein	cleavage
ggn-miR5021	AT3G52990.1	0.5	pyruvate kinase family protein	cleavage
ggn-miR5021	AT3G62560.1	0.5	ras-related small GTP-binding family protein	cleavage
ggn-miR5021	AT1G74180.1	0.5	receptor like protein 14	cleavage
ggn-miR5021	AT5G08710.1	0.5	regulator of chromosome condensation (RCC1) family protein	cleavage
ggn-miR5021	AT1G53860.1	0.5	remorin family protein	cleavage
ggn-miR5021	AT5G43420.1	0.5	RING/U-box superfamily protein	cleavage
ggn-miR5021	AT1G70200.1	0.5	RNA-binding (RRM/RBD/RNP motifs) family protein	cleavage
ggn-miR5021	AT3G56570.1	0.5	SET domain-containing protein	cleavage
ggn-miR5021	AT5G60120.1	0.5	target of early activation tagged (EAT) 2	cleavage
ggn-miR5021	AT1G71840.1	0.5	transducin family protein / WD-40 repeat family protein	cleavage
ggn-miR5021	AT1G26440.1	0.5	ureide permease 5	cleavage
ggn-miR5021	AT4G27240.1	0.5	zinc finger (C2H2 type) family protein	cleavage
ggn-miR5021	AT3G14980.1	0.5	acyl-CoA N-acyltransferase with RING/FYVE/PHD-type zinc finger protein	cleavage

ggn-miR5021	AT3G04620.1	0.5	alba DNA/RNA-binding protein	cleavage
ggn-miR5021	AT4G21320.1	0.5	aldolase-type TIM barrel family protein	cleavage
ggn-miR5021	AT4G17150.1	0.5	alpha/beta-Hydrolases superfamily protein	cleavage
ggn-miR5021	AT4G33980.1	0.5	cold regulated gene 27	cleavage
ggn-miR5021	AT4G38410.1	0.5	dehydrin family protein	cleavage
ggn-miR5021	AT5G55040.1	0.5	DNA-binding bromodomain-containing protein	cleavage
ggn-miR5021	AT2G28090.1	0.5	heavy metal transport/detoxification superfamily protein	cleavage
ggn-miR5021	AT2G22430.1	0.5	homeobox protein 6	cleavage
ggn-miR5021	AT4G28890.1	0.5	RING/U-box superfamily protein	cleavage
ggn-miR5021	AT2G18510.1	0.5	RNA-binding (RRM/RBD/RNP motifs) family protein	cleavage
ggn-miR5021	AT2G38610.1	0.5	RNA-binding KH domain-containing protein	cleavage
ggn-miR5021	AT5G23340.1	0.5	RNI-like superfamily protein	cleavage
ggn-miR5021	AT5G42850.1	0.5	thioredoxin superfamily protein	cleavage
ggn-miR5021	AT3G22380.1	0.5	time for Coffee	cleavage
ggn-miR5021	AT2G11030.1	0.5	transposable element gene	cleavage
ggn-miR5021	AT5G58620.1	0.5	zinc finger (CCCH-type) family protein	cleavage
ggn-miR5021	AT2G02480.1	1	AAA-type ATPase family protein	cleavage
ggn-miR5021	AT1G12930.1	1	ARM repeat superfamily protein	cleavage
ggn-miR5021	AT5G57180.1	1	chloroplast import apparatus 2	cleavage
ggn-miR5021	AT3G42670.1	1	chromatin remodeling 38	cleavage
ggn-miR5021	AT5G65900.1	1	DEA(D/H)-box RNA helicase family protein	cleavage
ggn-miR5021	AT5G45000.1	1	disease resistance protein (TIR-NBS-LRR class) family	cleavage
ggn-miR5021	AT2G38090.1	1	duplicated homeodomain-like superfamily protein	cleavage
ggn-miR5021	AT5G62620.1	1	galactosyltransferase family protein	cleavage
ggn-miR5021	AT5G05930.1	1	guanylyl cyclase 1	cleavage
ggn-miR5021	AT4G36180.1	1	leucine-rich receptor-like protein kinase family protein	cleavage
ggn-miR5021	AT2G44180.1	1	methionine aminopeptidase 2A	cleavage
ggn-miR5021	AT1G13590.1	1	phytosulfokine 1 precursor	cleavage
ggn-miR5021	AT1G54130.1	1	RELA/SPOT homolog 3	cleavage
ggn-miR5021	AT1G43770.1	1	RING/FYVE/PHD zinc finger superfamily protein	cleavage
ggn-miR5021	AT1G78000.1	1	sulfate transporter 1;2	cleavage
ggn-miR5021	AT1G13420.1	1	sulfotransferase 4B	cleavage
ggn-miR5021	AT4G28080.1	1	tetratricopeptide repeat (TPR)-like superfamily protein	cleavage
ggn-miR5021	AT2G26240.1	1	transmembrane proteins 14C	cleavage
ggn-miR5021	AT5G52380.1	1	VASCULAR-RELATED NAC-DOMAIN 6	cleavage
ggn-miR5021	AT1G78310.1	1	VQ motif-containing protein	cleavage
ggn-miR5021	AT4G29950.1	1	Ypt/Rab-GAP domain of gyp1p superfamily protein	cleavage
ggn-miR5021	AT3G43590.1	1	zinc knuckle (CCHC-type) family protein	cleavage
ggn-miR5021	AT2G44080.1	1	ARGOS-like	cleavage
ggn-miR5021	AT5G40890.1	1	chloride channel A	cleavage
ggn-miR5021	AT4G23750.1	1	cytokinin response factor 2	cleavage
ggn-miR5021	AT4G29920.1	1	double Clp-N motif-containing P-loop nucleoside triphosphate hydrolases superfamily protein	cleavage

ggn-miR5021	AT1G21390.1	1	embryo defective 2170	cleavage
ggn-miR5021	AT1G09650.1	1	F-box and associated interaction domains-containing protein	cleavage
ggn-miR5021	AT1G22300.1	1	general regulatory factor 10	cleavage
ggn-miR5021	AT4G01210.1	1	glycosyl transferase family 1 protein	cleavage
ggn-miR5021	AT2G37650.1	1	GRAS family transcription factor	cleavage
ggn-miR5021	AT2G37470.1	1	histone superfamily protein	cleavage
ggn-miR5021	AT1G80580.1	1	integrase-type DNA-binding superfamily protein	cleavage
ggn-miR5021	AT1G68790.1	1	little nuclei3	cleavage
ggn-miR5021	AT4G19380.1	1	long-chain fatty alcohol dehydrogenase family protein	cleavage
ggn-miR5021	AT1G26270.1	1	phosphatidylinositol 3- and 4-kinase family protein	cleavage
ggn-miR5021	AT1G66660.1	1	protein with RING/U-box and TRAF-like domains	cleavage
ggn-miR5021	AT1G32361.1	1	RING/U-box superfamily protein	cleavage
ggn-miR5021	AT4G02050.1	1	sugar transporter protein 7	cleavage
ggn-miR399	AT2G33770.1	2	phosphate 2	cleavage
ggn-miR399	AT3G54700.1	2	phosphate transporter 1;7	cleavage
ggn-miR399	AT4G00170.1	2.5	plant VAMP (vesicle-associated membrane protein) family protein	cleavage
ggn-miR399	AT4G05591.1	3	transposable element gene	cleavage
ggn-miR399	AT1G21270.1	3	wall-associated kinase 2	cleavage
ggn-miR399	AT1G31540.1	3	disease resistance protein (TIR-NBS-LRR class) family	cleavage

sequences, which did not show homology with any protein coding gene. These 56 sequences were retained for further analysis of the secondary structure of their precursors.

miRNA secondary structure

The 56 potential non-coding sequences were subjected to a structural validation analysis using Mfold v3.5 to predict the miRNA secondary structure. The miRNAs meeting the following criteria were considered valid for further analysis of their targets: 1) a valid stem-loop hairpin precursor, 2) the presence of complementary miRNA* sequence in the precursor with less than 6 mismatches, and 3) the MFEI value greater than 0.5. Twenty such conserved miRNAs were identified as belonging to two miRNA families. Nineteen members of miR399 were identified from various plant species, while miR5021 had only a single member from *A. thaliana*. On the basis of the stem-loop precursors, the identified miRNAs for *G. gnemon* were classified to be represented by two members of ggn-miR399 (ggn-miR399a and ggn-miR399b) and one member of ggn-miR5021 (Fig. 2). The ΔG values ranged from −58.05 to −24.52 kcal/mol (Table 1). It is often considered that the lower the value of ΔG, the higher is the thermodynamic stability of the miRNA precursor (Bonnet et al., 2004). A lower value of ΔG corresponds to higher MFE and MFEI values as MFE is equivalent to $-\Delta G$ (Zhang et al., 2006). The miRNA characterization indicates that the precursor length of miRNAs varied from 121 to 240 bases and the mature miRNA lengths ranged from 19 to 21 nucleotides (Table 1).

Phylogenetic analysis of precursor sequences

A phylogenetic analysis was performed to understand the relationship of the identified miRNAs in *G. gnemon* with other plant species available in the miRNA database in order to identify sequences belonging to the same miRNA families (Fig. 3). No miRNAs for *G. gnemon* have been reported earlier in any public domain. Hence, a phylogenetic tree generated using miRNA precursors will be helpful to understand the evolutionary and conserved nature of the identified miRNAs. Members of the miR399 family have been identified in many plant species, and precursor sequences of miR399 were retrieved from miRBase for *A. thaliana*, *Oryza sativa*, *Zea mays*, *Sorghum bicolor*, and *Cucumis melo* (Wang et al., 2020; Gonzalez-Ibeas et al., 2011; Zhang et al., 2011). Since

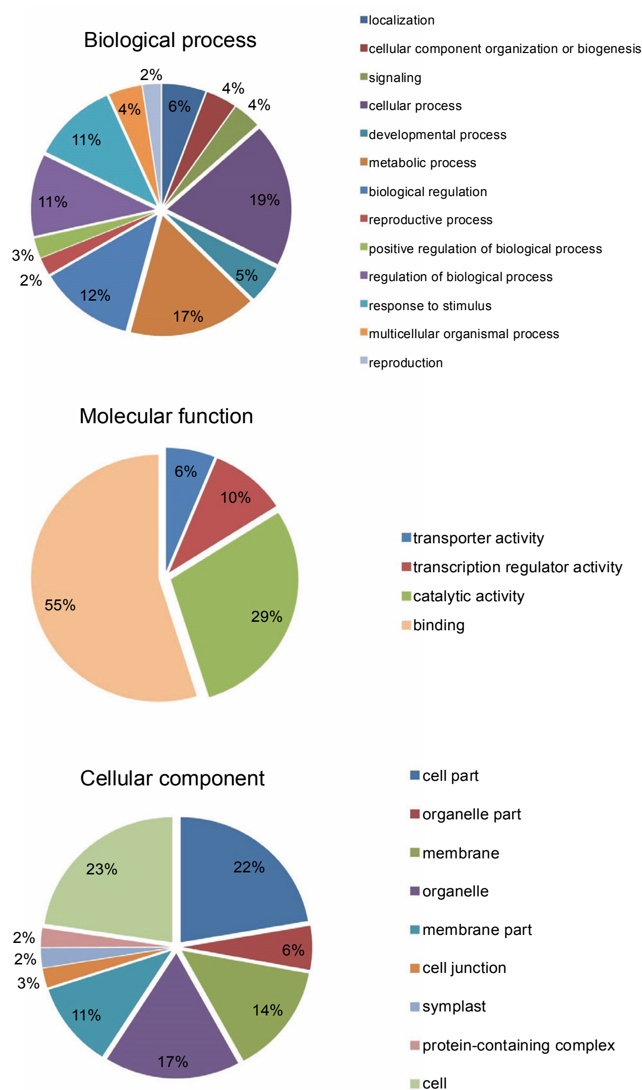

Fig. 4. GO reports of the identified target genes showing the percentage of sequences representing each class in three different categories, namely cellular component, biological processes, and molecular Function

the pre-miR5021 has been deposited only for *A. thaliana* in miRBase, the phylogenetic analysis was performed only for members of the miR399 family by using the maximum likelihood method. The precursor sequences of the identified miRNA families of *G. gnemon* formed a distinct cluster from other plant species (Fig. 3).

Prediction and annotation of target genes

miRNAs regulate the expression of specific genes by binding to mRNA transcripts based on perfect or near perfect complementarity to promote RNA degradation, inhibit translation or both (O'Brein et al., 2018). To identify genes plausibly recognized by the potential miRNAs, psRNATarget was used for searching target genes against the *A. thaliana* transcriptome acquired from TAIR10 and using the default parameters specified in SchemaV2 (2017). After removal of genes with unknown functions, a total of 100 genes (listed in Table 2) were identified as target genes of 20 identified miRNAs. To further realize the regulatory potential of the identified miRNAs, the target genes were subjected to the Gene Ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis by using Blast2GO v5.2. The identified target genes were found to participate in 13 distinct biological processes and 4 molecular functions (Fig. 4). Among the predicted targets, 10% geneswere sequence-specific transcription factors (TFs), 29% genes were involved in various catalytic functions, and 55% genes acted as sequence-specific or metal-ion binding factors. The two main categories represented among the biological processes were cellular processes and metabolic processes (19% and 17% genes, respectively). A pathway analysis of the target genes using the KEGG online database demonstrated the participation of the identified miRNAs in 14 vital metabolic networks such as purine metabolism, thiamine metabolism, nitrogen metabolism, glycolysis, and caffeine metabolism (Fig. 5). GO reports for each target gene can be accessed from Supplementary Table S1.

Cross-kingdom target gene prediction

Provided that the human diet is primarily dependent on plants, researchers have already proposed significant potential effects of plant small RNAs after food consumption on human gene expression (Vaucheret and Chupeau, 2011; Li et al., 2018; Sanchita et al., 2018). For better understanding of the potential biological functions of the newly identified *G. gnemon* miRNAs in human transcriptome, putative target genes were searched using the psRNATarget with default parameters against the *Homo sapiens* transcripts, Human Genome Sequencing project. A total of 49 probable targets were identified for the 20 identified *G. gnemon* miRNAs after eliminating genes with unknown or hypothetical functions. All the listed targets exhibited mRNA cleavage mode of regulation of gene expression by the predicted miRNAs (Supplementary Table S2). GO analysis results suggested that *G. gnemon* miRNAs could be involved in the regulation of 13 broadly defined biological processes (such as regulation of biological processes, cell signaling, and cellular processes) and 2 basic molecular functions

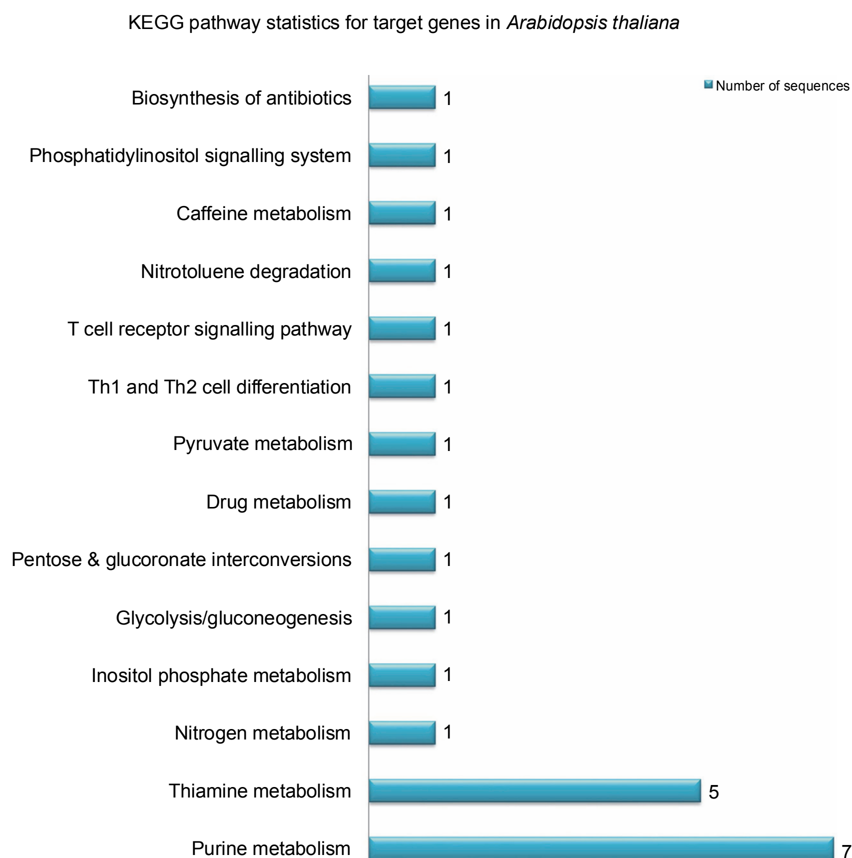

Fig. 5. KEGG pathway statistics showing the number of identified plant target genes belonging to the respective metabolic pathways

(catalytic activity and binding) in humans. The target genes were also found to be part of 11 different types of cellular components (such as cell organelles, cell membrane, and synapse) (Fig. 6). The KEGG pathway analysis of the target genes demonstrated the participation of the identified miRNAs in 11 different biological networks (Fig. 7). These networks are involved in various important pathways such as drug metabolism, caffeine metabolism, and nitrotoluene degradation.

Discussion

Screening and annotation of small RNAs help to interpret the critical roles played by such genetic modulators in the regulation of gene expression and their associated biological processes. Before this study, no comprehensive work was performed on the identification of putative miRNAs from ESTs of any species of *Gnetum.* This genus is often regarded as a very important plant group due to its significant taxonomic position as a connecting link between two diverse plant forms, namely gymnosperms and angiosperms (Bhat and Binti, 2014). In the present study, we considered all the important criteria such as the MFEI values, mismatch inhibition, and sequence length, which have been used for miRNA identification in other plant species. The MFEI values of the 20 identified miRNAs in our work were just above 0.5, which is in accordance with that reported for some miRNAs identified in other plant ESTs (Wang et al., 2012; Gupta et al., 2015; Patel et al., 2019; Zakeel et al., 2019). The G + C% of the miR399 members ranged from 37.5% to 38.8%, whereas it was found to be more than 41% for miR5021 mature miRNA sequence.

The TFs of plants have been found to be one of the notable targets of *G. gnemon* miRNAs. ggn-miR5021 and ggn-miR399 have shown a possible cleavage of a range of TFs such as SPL (SQUAMOSA Promoter-Binding Protein-Like) family of TFs, MYB (myeloblastosis) domain protein, and GRAS (named after the first three members: Gibberellic-acid Insensitive (GAI), Repressor of GAI, and SCARECROW). TFs are regarded as master

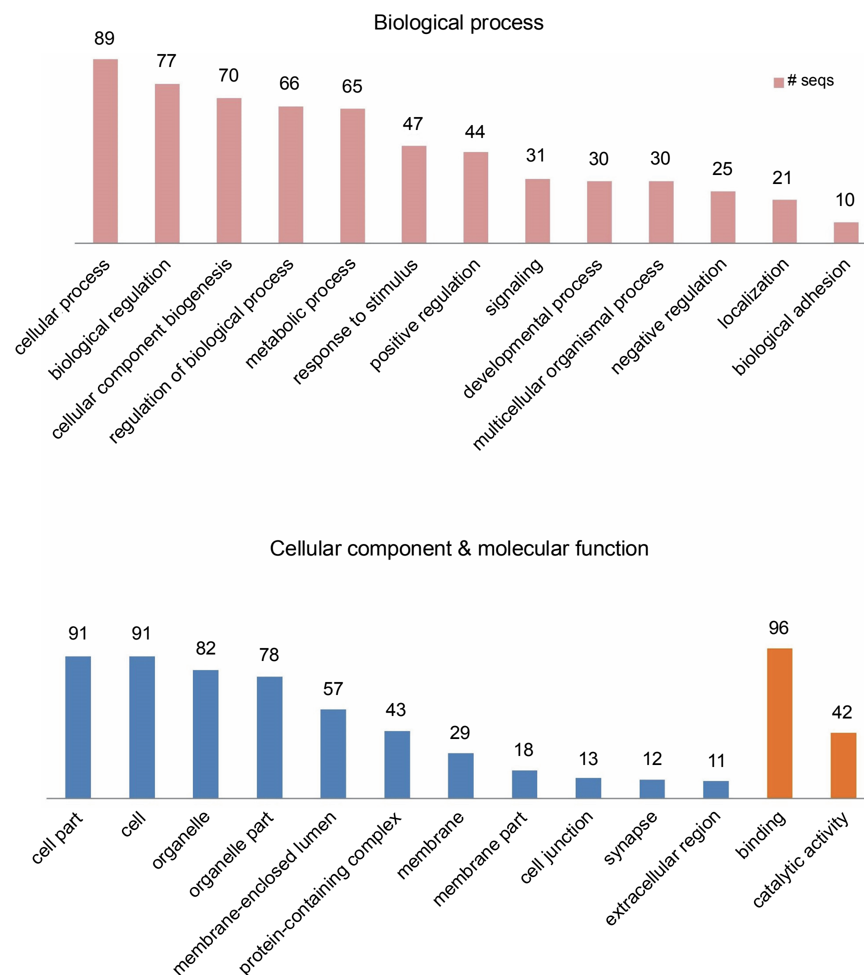

Fig. 6. GO reports of the identified target genes in humans; the coloured bars represent the number of gene sequences belonging to each category i.e, biological process (red), cellular component (blue) and molecular function (orange)

regulators of gene expression in eukaryotes and are responsible for promoting growth and development in plants (Kaufmann and Airoldi, 2018). SPL genes are known to encode plant-specific TFs that play important roles in phase transition, flower and fruit development, plant architecture, gibberellin signaling, sporogenesis, and response to copper and fungal toxins (Stief et al., 2014). SPL gene association with the miRNAs miR156 and miR172 has been found to play a critical role during the juvenile to adult leaf development in *A. thaliana* (Chen et al., 2010). Members of the MYB DNA-binding domain superfamily protein are involved in many important biochemical and physiological processes in plants (Du et al., 2009). The family of plant-specific MYB genes has immensely contributed to the evolution of physiological or developmental processes specific to plants, especially those involved in responses to fluctuating biotic or abiotic environments (Ambawat et al., 2013). Studies in several plants have revealed that GRAS family TFs serve diverse functions and are involved in various physiological processes. For instance, GRAS family members have been found to regulate GA signaling and the root and shoot radial pattern formation in *Arabidopsis* (Lu et al., 2020). The formation of nodules is regulated by two GRAS proteins in *Medicago trunculata* – a leguminous plant (Wang et al., 2018). Our results also show that ggn-miR5021 may also target several genes that encode various F-box proteins. These proteins are known as components of the SCF ubiquitin-ligase complexes (Skp I, Cullin, and an F-box protein), in which they bind to substrates for ubiquitin-mediated proteolysis (Kipreos and Pagano, 2000). Protein ubiquitination is considered to be a very critical post-translational modification process that is used by eukaryotes for the regulation of various types of cellular processes (Dye and Schulman, 2007). The target gene prediction ana-

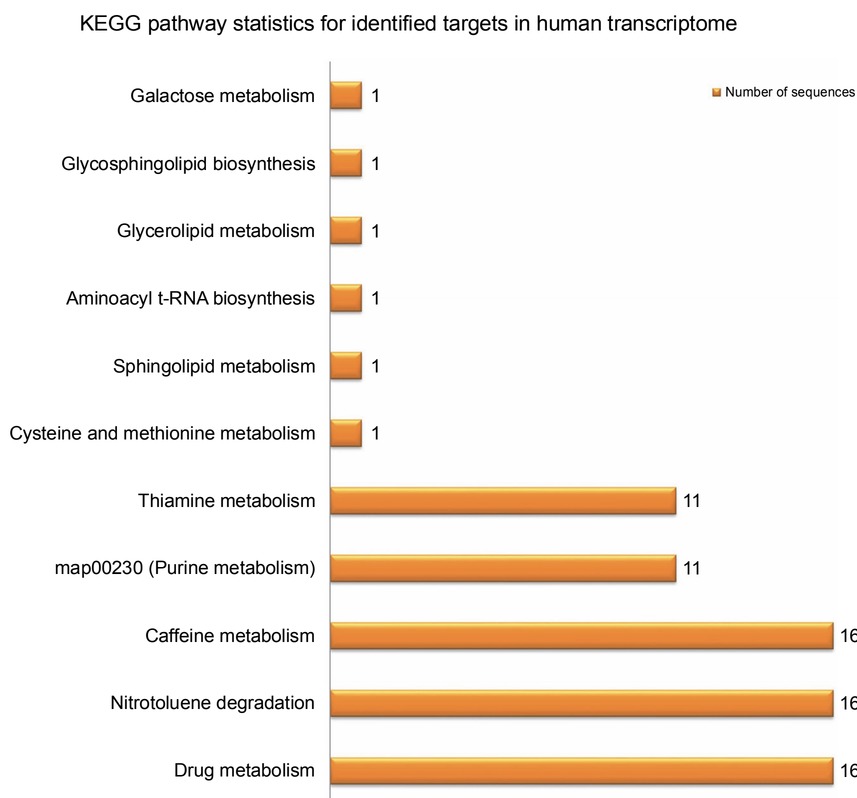

Fig. 7. KEGG pathway results for the identified human target genes of *G. gnemon* miRNAs

lysis also reveals another intriguing possibility that ggn-miR5021 and ggn-miR399 may also target three important defense-related genes, namely Leucine-rich repeat receptor kinases (LRR-RKs), TIR-NBS-LRR protein family, and Time for Coffee (TIC) – a circadian clock regulator involved in time-of-day-related regulation of jasmonate signaling (Shin et al., 2012).

miRNAs have also been shown to have beneficial effects on human health. In fact, cross-kingdom delivery of plant miRNAs into the human alimentary canal can be considered as a mode of herbal therapy against many diseases (Sanchita et al., 2018). *G. gnemon* is also a popular medicinal plant known for its therapeutic properties in treating human ailments, and it is consumed as food among various tribes (Barua et al., 2015; Agarwal et al., 2020). To analyze the potential effectiveness of the exogenous application of miRNAs derived from this edible plant in humans, we analyzed the target specificity of the identified miRNAs against human genes. Among the identified cross-kingdom targets in human transcriptome, 70% genes were found to be DNA-binding, metal ion-binding, or protein-binding factors, and 30% genes possessed various catalytic functions. In the GO analysis, the two main categories represented among the biological processes are biological pathway regulation and biogenesis of cellular components (13% and 11% genes, respectively). Only one TF – a bHLH TF, namely ASCL1 – was found to be a probable target of miR5021. The "achaete-scute complex-like" (ASCL) family or "achaete scute family basic helix-loop-helix transcription factor" is critical for the proper development of the nervous system in humans, and its dysregulation is proposed to play a key role in psychiatric and neurological disorders (Wang et al., 2017). *G. gnemon* is used as an edible plant in some North-Eastern parts of India and in some other countries including Indonesia and Thailand. It is often considered as an alternative for protein-rich foods because of the abundance of essential amino acids (Ota et al., 2013). *Gnetum* seeds are known to be rich in stilbenoids – a class of phenolic compounds – including trans-resveratrol, gnetins, and gnemonosides (Kato et al., 2009). Human studies have revealed that trans-resveratrol is beneficial in the management of diabetes and cardiovascular diseases (Brasnyó et al.,

Table 3. Disease-associated target genes of *G. gnemon* miRNAs in the human transcriptome

Target gene	Gene description and function	Associated disorders
MYBPC2	myosin binding protein C; located in striated muscle bands, modifies the activity of actin-activated myosin ATPase and may modulate muscle contraction	hypertrophic cardiomyopathy, distal arthrogryposis
C1RL1-AS1	complement C1r Sub-component like (C1RL) antisense RNA 1, long non-coding RNA	ovary adenocarcinoma, breast adenocarcinoma
FCN1	ficolin 1; pattern recognition receptor in innate immunity	autoimmune diseases, infectious and rheumatic disorders
OTOF	otoferlin; calcium ion sensor involved in Ca(2+)-triggered synaptic vesicle-plasma membrane fusion	branchiootic syndrome, deafness
MRAS	muscle RAS oncogene homolog; weakly activates the MAP kinase pathway	hypertension, cerebral arterial disease, vertebral artery occlusion
RPGR	*Retinitis pigmentosa* GTPase regulator; potentially involved in vesicle transport and ciliogenesis.	*Retinitis pigmentosa*
TERF2IP	telomeric repeat-binding factor 2-interacting protein; acts both as a regulator of telomere function and as a transcription regulator	*Cutaneous melanoma*
EPHA1-AS1	EPH receptor A1 antisense RNA 1, long non-coding RNA	*Placenta previa*, serous carcinoma of the ovary
KLF6	Kruppel-like factor 6; binds a GC box motif; could play a role in B-cell growth and development	prostate cancer, gastric cancer
MPZL2	myelin protein zero-like 2; mediates homophilic cell-cell adhesion	neurosensory deafness
PTCH1	patched 1; associates with the smoothened protein (SMO) to transduce the hedgehog's protein signal	basal cell carcinoma
CCL16	C-C motif chemokine ligand 16; chemotactic activity for lymphocytes and monocytes but not neutrophils	corneal dystrophy
ASCL1	achaete-scute family bHLH transcription factor 1; transcription factor that plays a key role in neuronal differentiation	central hypoventilation syndrome, pulmonary neuroendocrine tumor
BRD3	bromodomain containing 3; chromatin reader that recognizes and binds hyperacetylated chromatin	foodborne and infant botulism
FERMT3	fermitin family member 3; plays a central role in cell adhesion in hematopoietic cells	leukocyte adhesion deficiency, kindler syndrome
GRIA3	glutamate ionotropic receptor AMPA type subunit 3; receptor for glutamate that functions as a ligand-gated ion channel in the central nervous system	depression, mental retardation
INCENP	inner centromere protein; component of the chromosomal passenger complex (CPC) that acts as a key regulator of mitosis	pregnancy-related cholestasis, nephronophthisis
KAT6B	lysine acetyltransferase 6B; histone acetyltransferase – involved in the regulation of transcription	genitopatellar syndrome, mental retardation
NEXN	nexilin F-actin binding protein; regulating cell migration through association with the actin cytoskeleton	cardiomyopathy
LRP2	LDL receptor-related protein 2; endocytosis of high-density lipoproteins	dent disease, cystinosis, membranous nephropathy

2011, Wong et al., 2011). Local tribes attribute many other medicinal properties to the extracts of seeds and leaves of *G. gnemon*. Evidence confirming the inhibition of viral infection and other diseases in humans, such as cancer and rheumatoid arthritis, has been reported for exogenous plant-based miRNA application (Sanchita et al., 2018). To ascertain any possible medicinal roles that can be attributed to the miRNAs identified in *G. gnemon*, we assessed the possible target genes that may be related to any human disorders. Table 3 lists the

potential target genes of *G. gnemon* miRNAs predicted in human transcriptome that are associated with some disorders or ailments in humans and their basic functions. Several such genes were identified amongst the putative targets of ggn-5021 and ggn-399, such as *Retinitis pigmentosa* GTPase regulator (RPGR), LRP2, BCL2, KLF6, Muscle RAS oncogene homolog, and Otoferlin (OTOF) (Table 3). These genes are associated with various human disorders such as different forms of cancer, deafness, and various disorders of genetic and non-genetic origin (Safran et al., 2010; Rappaport et al., 2013). A noticeable feature found in the human target analysis is that some long non-coding RNAs can also act as potential targets of the identified miRNAs, such as EPHA1-AS1 and C1RL1-AS1 (Table 3). This type of activity suggests that a cleavage-mediated degradation of lncRNAs as target gene mimics might help to enhance the expression of the respective genes (Karlik et al., 2019). Studies related to miRNA-lncRNA-mRNA association modules in various organisms have been conducted as a recent purview in the field of biological research that needs a significant amount of evaluation and substantiation (Zhang et al., 2019). The antitumor activity of "melinjo" seeds has also been studied in human tumor models by Narayanan et al. (2015). The potential application miRNA-based drug delivery has already been studied in tumors, cancers, and osteoporotic fractures (Sun et al., 2019; Anthiya et al., 2018; Si et al., 2019). The identification of cross-kingdom target genes provides a future prospect and direction for exploring plausible medicinal characteristics that can be imparted by plant miRNAs to humans.

Conclusions

The ESTs of *G. gnemon* were used to identify miRNAs and to predict their target genes in *A. thaliana* and humans by using a computational approach. Twenty putative miRNAs, representing two miRNA families, were predicted from the ESTs of *G. gnemon*. A total of 100 target genes were predicted, representing various TFs, catalytic enzymes, and cofactors in *A. thaliana*. A cross-kingdom target analysis in the human transcriptome revealed various genes involved in downstream signaling pathways in diverse functional processes and different organs. Moreover, enlisting the probable target genes of *Gnetum* miRNAs in the human transcriptome will help to provide a basis for investigating the regulatory and medicinal roles that may be played by the miRNAs of this edible plant consumed as food in significant quantities in some tribal areas. Conclusively, these findings may further enhance our understanding of how the plant miRNAs regulate various metabolic and catalytic processes and help to recognize the therapeutic and nutritional potential of the plant. However, the availability of transcriptome sequences of *G. gnemon*, their miRNA expression analysis, and validation against the transcriptome sequences will be substantial benchmarks to confirm such possibilities.

Acknowledgments

The authors would like to acknowledge the funds received from the Department of Science and Technology (DST), Government of India, for the FIST project granted to the Department of Botany, Gauhati University, Guwahati. DBK would like to thank the DST, Government of India, for the INSPIRE fellowship grant (IF180472). KSB would like to thank the Council of Scientific and Industrial Research for providing CSIR-Junior Research Fellowship grant (09/059(0064)/2018-EMR-I).

Availability of data and materials

A total of 38 589 previously identified mature microRNAs from different plants were retrieved from the miRBase database (http://www.mirbase.org/) (release 22.1). These sequences were defined as the query sequence set and used for identifying miRNAs in *G. gnemon* Expressed Sequence Tags (ESTs). Publicly available 10 725 ESTs (as of August 1, 2019) of the species were downloaded from National Centre for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/). Non-redundant protein sequences were used from the NR protein database of NCBI (ftp://ftp.ncbi.nlm.nih.gov/blast/db/).

Funding

This work was supported by the funds from DST-SERB-funded ECR Award Project (ECR 000710/2017) granted to NA by the Government of India.

References

Agarwal H., Dowarah B., Baruah P.M., Bordoloi K.S., Krishnatreya D.B., Agarwala N. (2020) *Endophytes from Gnetum gnemon L. can protect seedlings against the infection of phytopathogenic bacterium Ralstonia solanacearum as well as promote plant growth in tomato*. Microbiol. Res. 238: 126503. https://doi.org/10.1016/j.micres.2020.126503

Ahmad I., Andriyani D., Gunawan C., Fauziah N.D., Sutriyo S., Munim A. (2018) *Effects of food processing on resveratrol and total phenolic content in Melinjo (Gnetum gnemon L.) seeds.* Pharmacogn J. 10(6): 1096–1100. http://dx.doi.org/10.5530/pj.2018.6.186

Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. (1997) *Gapped BLAST and PSI-*

BLAST: a new generation of protein database search programs. Nucl. Acids Res. 25(17): 3389–3402. https://doi.org/10.1093/nar/25.17.3389

Ambawat S., Sharma P., Yadav N.R., Yadav R.C. (2013) *MYB transcription factor genes as regulators for plant responses: an overview*. Physiol. Mol. Biol. Plants 19(3): 307–321. https://dx.doi.org/10.1007%2Fs12298-013-0179-1

Anthiya S., Griveau A., Loussouarn C., Baril P., Garnett M., Issartel J.P., Garcion E. (2018) *MicroRNA-based drugs for brain tumors*. Trends Cancer. 4: 222–238. https://doi.org/10.1016/j.trecan.2017.12.008

Axtell M.J., Meyers B.C. (2018) *Revisiting criteria for plant miRNA annotation in the era of big data*. Plant Cell. 30(2): 272–284. https://doi.org/10.1105/tpc.17.00851

Barua C.C., Haloi P., Barua I.C. (2015) *Gnetum gnemon Linn.: a comprehensive review on its biological, pharmacological and pharmacognostical potentials*. Int. J. Pharmacog. Phytochem. Res. 7(3): 531–539.

Bhat R., Binti Y.N. (2014) *Evaluating melinjau (Gnetum gnemon L.) seed flour quality as a base for development of novel food products and food formulations*. Food Chem. 156: 42–49. https://doi.org/10.1016/j.foodchem.2014.01.063

Bonnet E., Wuyts J., Rouzé P., Van de Peer Y. (2004) *Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences*. Bioinformatics 20: 2911–2917. https://doi.org/10.1093/bioinformatics/bth374

Bordoloi K.S., Agarwala N. (2021) *MicroRNAs in plant insect interaction and insect pest control*. Plant Gene. 25: 100271 (in press) https://doi.org/10.1016/j.plgene.2021. 100271

Brasnyó P., Molnár G.A., Mohás M., Marko L., Laczy B., Cseh J., Mikolas E., Szijarto I.A., Merei A., Halmai R., et al. (2011) *Resveratrol improves insulin sensitivity, reduces oxidative stress and activates the Akt pathway in type 2 diabetic patients*. Br. J. Nutr. 106(3): 383–389. https://doi.org/10.1017/S0007114511000316

Chen X., Zhang Z., Liu D., Zhang K., Li A., Mao L. (2010) *Squamosa promoter-binding protein-like transcription factors: star players for plant growth and development*. J. Integr. Plant Biol. 52: 946–951. https://doi.org/10.1111/j.1744-7909.2010.00987.x

Conesa A., Götz S. (2008) *Blast2GO: A comprehensive suite for functional analysis in plant genomics*. Int. J. Plant Genomics 2008: 619832. https://dx.doi.org/10.1155%2F2008%2F619832

Dai X., Zhuang Z., Zhao P.X. (2018) *psRNATarget: a plant small RNA target analysis server (2017 release)*. Nucl. Acids Res. 46(W1): W49–W54. https://doi.org/10.1093/nar/gky316

Du H., Zhang L., Liu L., Tang X.-F., Yang W.-J., Wu Y.-M., Huang Y.-B., Tang Y.-X. (2009) *Biochemical and molecular characterization of plant MYB transcription factor family*. Biochemistry (Moscow) 74(1): 1–11. https://doi.org/10.1134/s0006297909010015

Dye B.T., Schulman B.A. (2007) *Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins*. Annu. Rev. Biophys. Biomol. Struct. 36: 131–150. https://doi.org/10.1146/annurev.biophys.36.040306. 132820

Gonzalez-Ibeas D., Blanca J., Donaire L., Saladie M., Mascarell-Creus A., Cano-Delgado A., Garcia-Mas J., Llave C., Aranda M.A. (2011) *Analysis of the melon (Cucumis melo) small RNAome by high-throughput pyrosequencing*. BMC Genomics. 12: 393. https://dx.doi.org/10.1186%2F1471-2164-12-393

Griffiths-Jones S., Grocock R.J., van Dongen S., Bateman A., Enright A.J. (2006) *miRBase: microRNA sequences, targets and gene nomenclature*. Nucl. Acids Res. 34: D140–144. https://doi.org/10.1093/nar/gkj112

Gupta H., Tiwari T., Patel M., Mehta A., Ghosh A. (2015) *An approach to identify the novel miRNA encoded from H. Annuus EST sequences*. Genom Data 6: 139–144. https://dx.doi.org/10.1016%2Fj.gdata.2015.09.005

Jones-Rhoades M.W., Bartel D.P., Bartel B. (2006) *MicroRNAS and their regulatory roles in plants*. Annu. Rev. Plant Biol. 57: 19–53. https://doi.org/10.1146/annurev.arplant.57. 032905.105218

Karlik E., Ari S., Gozukirmizi N. (2019) *LncRNAs: genetic and epigenetic effects in plants*. Biotech. Biotech. Equip. 33(1): 429–439. https://doi.org/10.1080/13102818.2019. 1581085

Kato E., Tokunaga Y., Sakan F. (2009) *Stilbenoids isolated from the seeds of Melinjo (Gnetum gnemon L.) and their biological activity*. J. Agric. Food Chem. 57: 2544–2549. https://doi.org/10.1021/jf803077p

Kaufmann K., Airoldi C.A. (2018) *Master regulatory transcription factors in plant development: a blooming perspective*. Meth. Mol. Biol. 1830: 3–22. https://doi.org/10.1007/978-1-4939-8657-6_1

Kipreos E.T., Pagano M. (2000) *The F-box protein family*. Genome Biol. 1(5): reviews3002.1. https://dx.doi.org/10.1186%2Fgb-2000-1-5-reviews3002

Krishnatreya D.B., Baruah P.M., Dowarah B., Bordoloi K.S., Agarwal H., Agarwala N. (2020) *Mining of miRNAs from EST data in Dendrobium nobile*. Bioinformation 16(3): 245–255. https://doi.org/10.6026/97320630016245

Krishnatreya D.B., Agarwala N., Gill S.S., Bandyopadhyay T. (2021) *Understanding the role of miRNAs for improvement of tea quality and stress tolerance*. J. Biotech. 328: 34–46. https://doi.org/10.1016/j.jbiotec.2020.12.019

Li Z., Xu R., Li N. (2018) *MicroRNAs from plants to animals, do they define a new messenger for communication?* Nutr. Metab. 15: 68. https://doi.org/10.1186/s12986-018-0305-8

Lu Y., Yang X. (2010) *Computational identification of novel microRNAs and their targets in Vigna unguiculata*. Comp. Funct. Genomics. 2010: 128297. https://dx.doi.org/10.1155%2F2010%2F128297

Lu X., Liu W., Xiang C., Li X., Wang Q., Wang T., Liu Z., Zhang J., Gao L., Zhang W. (2020) *Genome-wide characterization of GRAS family and their potential roles in cold tolerance of cucumber (Cucumis sativus L.)*. Int. J. Mol. Sci. 21(11): 3857. https://dx.doi.org/10.3390% 2Fijms21113857

Narayanan N.K., Kunimasa K., Yamori Y., Mori M., Mori H., Nakamura K., Miller G., Manne U., Tiwari A.K., Naraya-

nan B. (2015) *Antitumor activity of melinjo (Gnetum gnemon L.) seed extract in human and murine tumor models in vitro and in a colon-26 tumor-bearing mouse model in vivo*. Cancer Med. 4(11): 1767–1780. https://dx.doi.org/10.1002%2Fcam4.520

Nithin N., Patwa P., Thomas A., Bahadur R.P., Basak J. (2015) *Computational prediction of miRNAs and their targets in Phaseolus vulgaris using simple sequence repeat signatures*. BMC Plant Biol. 15: 140. https://doi.org/10.1186/s12870-015-0516-3

O'Brein J., Hayder H., Zayed Y., Peng C. (2018) *Overview of microRNA biogenesis, mechanisms of actions, and circulation*. Front. Endocrinol. 9: 402. https://doi.org/10.3389/fendo.2018.00402

Ota H., Akishita M., Tani H., Tatefuji T., Ogawa S., Iijima K., Eto M., Shirasawa T., Ouchi Y. (2013) *Trans-resveratrol in Gnetum gnemon protects against oxidative-stress-induced endothelial senescence*. J. Nat. Prod. 76: 1242–1247. https://doi.org/10.1021/np300841v

Panda D., Dehury B., Sahu J., Barooah M., Sen P., Modi M.K. (2014) *Computational identi cation and characterization of conserved miRNAs and their target genes in garlic (Allium sativum L.) expressed sequence tags*. Gene 537: 333–342. https://doi.org/10.1016/j.gene.2014.01.010

Patel M., Mangukia N., Jha N., Gadhavi H., Shah K., Patel S., Mankad A., Pandya H., Rawal R. (2019) *Computational identification of miRNA and their cross kingdom targets from expressed sequence tags of Ocimum basilicum*. Mol. Biol. Rep. 46(3): 2979–2995. https://doi.org/10.1007/s11033-019-04759-x

Pirro S., Matic I., Guidi A., Zanella L., Gismondi A., Cicconi R., Bernardini R., Colizzi V., Canini A., Mattei M., Galgani A. (2019) *Identification of microRNAs and relative target genes in Moringa oleifera leaf and callus*. Sci. Rep. 9: 15145. https://doi.org/10.1038/s41598-019-51100-4

Rappaport N., Nativ N., Stelzer G., Twik M., Guan-Golan Y., Stein T.I., Bahir I., Belinky F., Morrey C.P., Safran M., et al. (2013) *MalaCards: an integrated compendium for diseases and their annotation*. Database (Oxford). 2013: bat018. https://dx.doi.org/10.1093%2Fdatabase%2Fbat018

Reinhart B.J., Weinstein E.G., Rhoades M.W., Bartel B., Bartel D.P. (2002) *MicroRNAs in plants [published correction appears in Genes Dev. 2002 Sep 1; 16(17): 2313*]. Genes Dev. 16(13): 1616–1626. https://dx.doi.org/10.1101%2Fgad.1004402

Ren D., Labandeira C.C., Santiago-Blay J.A., Rasnitsyn A., Shih C.K., Bashkuev A., Logan M.A., Hotton C.L., Dilcher D. (2009) *Probable pollination mode before angiosperms: Eurasian, long-proboscid scorpionflies*. Science 326(5954): 840–847. https://doi.org/10.1126%2Fscience.1178338

Safran M., Dalah I., Alexander J., Rosen N., Stein T.I., Shmoish M., Nativ N., Bahir I.Doniger T., Krug H., et al. (2010) *GeneCards Version 3: the human gene integrator*. Database (Oxford) 2010: baq020. https://dx.doi.org/10.1093%2Fdatabase%2Fbaq020

Sanchita, Trivedi R., Asif M.H., Trivedi P.K. (2018) *Dietary plant miRNAs as an augmented therapy: cross-kingdom gene regulation*. RNA Biol. 15(12): 1433–1439. https://doi.org/10.1080/15476286.2018.1551693

Shin J., Heidrich K., Sanchez-Villarreal A., Parker J.E., Davis S.J. (2012) *Time for coffee represses accumulation of the MYC2 transcription factor to provide time-of-day regulation of jasmonate signaling in Arabidopsis*. Plant Cell. 24(6): 2470–2482. https://dx.doi.org/10.1105%2Ftpc.111.095430

Si W., Shen J., Zheng H., Fan W. (2019) *The role and mechanisms of action of microRNAs in cancer drug resistance*. Clin. Epigenet. 11(1): 25. https://dx.doi.org/10.1186%2Fs13148-018-0587-8

Stief A., Altmann S., Hoffmann K., Pant B.D., Scheible W.R., Baurle I. (2014) *Arabidopsis miR156 regulates tolerance to recurring environmental stress through SPL transcription factors*. Plant Cell. 26: 1792–1807. https://doi.org/10.1105/tpc.114.123851

Sun X., Guo Q., Wei W., Robertson S., Yuan Y., Luo X. (2019) *Current progress on microRNA-based gene delivery in the treatment of osteoporosis and osteoporotic fracture*. Int. J. Endocrinol. 2019: 6782653. https://dx.doi.org/10.1155%2F2019%2F6782653

Tatefuji T., Yanagihara M., Fukushima S., Hashimoto K. (2014) *Safety assessment of melinjo (Gnetum gnemon L.) seed extract: acute and subchronic toxicity studies*. Food Chem. Toxicol. 67: 230–235. https://doi.org/10.1016/j.fct.2014.02.030

Terangpi R., Engtipi U., Teron R. (2013) *Utilization of less known plants Gnetum gnemon L. and Rhynchotechum ellipticum (Dietr.) A. DC. among the Karbis, North-east India*. J. Sci. Inn. Res. 2(5): 943–949.

Vaucheret H., Chupeau Y. (2011) *Ingested plant miRNAs regulate gene expression in animals*. Cell Res. 22: 3–5. https://doi.org/10.1038/cr.2011.164

Vishwakarma N.P., Jadeja V.J. (2013) *Identification of miRNA encoded by Jatropha curcas from EST and GSS*. Plant Signal Behav. 8(2): e23152. https://dx.doi.org/10.4161%2Fpsb.23152

Wang R., Fang Y.N., Wu X.M., Qing M., Li C.C., Xie K.D., Deng X.X., Guo W.W. (2020) *The miR399-CsUBC24 module regulates reproductive development and male fertility in citrus.* Plant Physiol. 183(4): 1681–1695. https://doi.org/10.1104/pp.20.00129

Wang C.Y., Shahi P., Huang J.T., Phan N.N., Sun Z., Lin Y-C., Lai M-D., Werb Z. (2017) *Systematic analysis of the achaete-scute complex-like gene signature in clinical cancer patients*. Mol. Clin. Oncol. 6(1): 7–18. https://dx.doi.org/10.3892%2Fmco.2016.1094

Wang M., Wang Q., Wang B. (2012) *Identification and characterization of microRNAs in Asiatic cotton (Gossypium arboreum L.)*. PLoS One 7(4): e33696. https://dx.doi.org/10.1371%2Fjournal.pone.0033696

Wang Y.X., Liu Z.W., Wu Z.J., Li H., Wang W.L., Cui X., Zhuang J. (2018) *Genome-wide identification and expression analysis of GRAS family transcription factors in tea plant (Camellia sinensis)*. Sci. Rep. 8(1): 3949. https://dx.doi.org/10.1038%2Fs41598-018-22275-z

Wong R.H.X., Howe P.R.C., Buckley J.D., Coates A.M., Kunz I., Berry N.M. (2011) *Acute resveratrol supplementation improves flow-mediated dilatation in overweight/obese individuals with mildly elevated blood pressure*. Nutr. Metab. Carbiovasc. Dis. 21: 851–856. https://doi.org/10.1016/j.numecd.2010.03.003

Xu X., Chen X., Chen Y., Zhang Q., Su L., Chen X., Chen Y., Zhang Z., Lin Y., Lai Z. (2020) *Genome-wide identification of miRNAs and their targets during early somatic embryogenesis in Dimocarpus longan Lour*. Sci. Rep. 10: 4626. https://doi.org/10.1038/s41598-020-60946-y

Xun Y., Tang Y., Hu L., Xiao H., Long S., Gong M., Wei C., Wei K., Xiang S. (2019) *Purification and identification of miRNA target sites in genome using DNA affinity precipitation*. Front. Genet. 10: 778. https://doi.org/10.3389/fgene.2019.00778

You C., Cui J., Wang H., Qi X., Kuo L., Ma H., Gao L., Mo B., Chen X. (2017) *Conservation and divergence of small RNA pathways and microRNAs in land plants*. Genome Biol. 18: 158. https://doi.org/10.1186/s13059-017-1291-2

Zakeel M.C.M., Safeena M.I.S., Komathy T. (2019) *In silico identification of microRNAs and their target genes in watermelon (Citrullus lanatus)*. Sci. Hortic. 252: 55–60. https://doi.org/10.1016/j.scienta.2019.02.012

Zhang B.H., Pan X.P., Cox S.B., Cobb G.P., Anderson T.A. (2006) *Evidence that miRNAs are different from other RNAs*. Cell Mol. Life Sci. 63(2): 246–254. https://doi.org/10.1007/s00018-005-5467-7

Zhang P., Meng J., Luan Y., Liu C. (2019) *Plant miRNA-lncRNA interaction prediction with the ensemble of CNN and IndRNN*. Interdiscip. Sci. Comput. Life Sci. 12: 82–89. https://doi.org/10.1007/s12539-019-00351-w

Zhang L., Zheng Y., Jagadeeswaran G., Li Y., Gowdu K., Sunkar R. (2011) *Identification and temporal expression analysis of conserved and novel microRNAs in Sorghum*. Genomics 98(6): 460–468. https://doi.org/10.1016/j.ygeno.2011.08.005

Zuker M. (2003) *Mfold web server for nucleic acid folding and hybridization prediction*. Nucl. Acids Res. 31(13): 3406–3415. https://doi.org/10.1093/nar/gkg595