Biotechnol Lett (2012) 34:737–745
DOI 10.1007/s10529-011-0808-0

ORIGINAL RESEARCH PAPER

Mining of miRNAs and potential targets from gene oriented clusters of transcripts sequences of the anti-malarial plant, *Artemisia annua*

Álvaro L. Pérez-Quintero · Gaurav Sablok · Tatiana V. Tatarinova · Ana Conesa · Jimmy Kuo · Camilo López

Received: 1 July 2011 / Accepted: 21 November 2011 / Published online: 9 December 2011

Abstract miRNAs involved in the biosynthesis of artemisinin, an anti-malarial compound form the plant *Artemisia annua*, have been identified using computational approaches to find conserved pre-miRNAs in available *A. annua* UniGene collections. Eleven pre-miRNAs were found from nine families. Targets predicted for these miRNAs were mainly transcription factors for conserved miRNAs. No target genes involved in artemisinin biosynthesis were found. However, miR390 was predicted to target a gene involved in the trichome development, which is the site of synthesis of artemisinin and could be a candidate for genetic transformation aiming to increase the content of artemisinin. Phylogenetic analyses were carried out to determinate the relation between *A. annua* and other plant pre-miRNAs: the pre-miRNA-based phylogenetic trees failed to correspond to known phylogenies, suggesting that pre-miRNA primary sequences may be too variable to accurately predict phylogenetic relations.

Keywords Anti-malarial plant · *Artemisia annua* · miRNAs

Álvaro L. Pérez-Quintero, Gaurav Sablok and Tatiana V. Tatarinova contributed equally to this work.

Á. L. Pérez-Quintero · C. López
Departamento de Biología, Universidad Nacional de Colombia, Bogotá, Bogota, DC, Colombia

G. Sablok (✉)
Key Laboratory of Horticultural Plant Biology (MOE), Huazhong Agricultural University, Wuhan 430070, China
e-mail: sablokg@gmail.com

T. V. Tatarinova
Division of Mathematics and Statistics, University of Glamorgan, Pontypridd CF37 1DL, UK

A. Conesa
Bioinformatics and Genomics Department, Centro de Investigación Príncipe FelipeAvda, Autopista Saler 16, 46012 Valencia, Spain

J. Kuo
Department of Planning and Research, National Museum of Marine Biology and Aquarium, Pingtung 944, Taiwan

Introduction

Artemisinin is widely used in anti-malarial therapy. However, the relatively low content of artemisinin in cultivated types of *Artemisia annua* has been a limiting factor for its commercialization. The principal challenge is, in consequence, to increase the production of artemisinin which could be done using biotechnological strategies and taking into account that the biosynthesis pathway of artemisinin has been partially elucidated (Lin et al. 2011; Olofsson et al. 2011). Tetraploid *A. annua* plants have been produced with increased artemisinin yields using colchicine. This increase could be correlated with the up-regulation in the expression of some genes coding for

important enzymes involved in the artemisinin biosynthesis pathway (Lin et al. 2011). In addition, the expression of 14 genes involved in the terpene metabolism, which can affect the production of artemisinin, was studied by QRT-PCR in different *A. annua* tissues (Olofsson et al. 2011). This identified some genes having positive and negative effects on the biosynthesis of artemisinin. Deep sequencing has also been applied to identify genes and markers for fast-track breeding and transcriptome analysis, revealing extensive genetic variation with as many as nine linkage groups defined (Graham et al. 2010). Quantitative trait loci (QTL) have also been identified to account for key traits controlling artemisinin synthesis and will provide a novel tool to increase the content of artemisinin (Bouwmeester et al. 1999).

Clearly a better knowledge of the molecular mechanisms involved in the artemisinin biosynthesis and their regulation will provide better strategies to develop new varieties with a higher content of this compound. The different biosynthetic pathways in plants are not only controlled by positive regulators but also there are negative ways to control the production of particular metabolic compounds. Rydén et al. (2010) identified Red1 as an enzyme negatively regulating the biosynthesis of artemisinin, although the real role of this enzyme has been recently questioned (Olofsson et al. 2011).

MicroRNAs are important regulators in various stages of plant development as well as in response to different abiotic or biotic stresses (Jones-Rhoades et al. 2006; Kidner and Martienssen 2005). The regulation of metabolite biosynthesis is also dependent on the regulation mediated by miRNAs (Robert-Seilaniantz et al. 2010). A UniGene set is defined as a group of transcript sequences that, based on sequence similarity, originate from the same gene or expressed pseudo gene (Gupta and Rustgi 2004). UniGene represents an organized array of sequences of well-characterized genes as well as hundreds of thousands novel ESTs sequences (Parida et al. 2006).

With the aim of identifying possible miRNA-mediated regulation in the artemisinin biosynthesis pathway, we have now identified *A. annua* miRNAs based on the available UniGene set. The knowledge of *A. annua* miRNAs constitutes an important tool for studying different developmental miRNA-mediated processes in the plant.

Materials and methods

Homology based searching of microRNA sequences and target prediction

A set of known miRNAs was downloaded from miRBase (version 15.0, April 2010) comprising a total of 2,222 mature miRNA sequences from 29 plant species (Griffiths-Jones et al. 2008). A BLASTN search was made against the set of 88,174 UniGenes of *A. annua* obtained from the UniGene repository (NCBI: http://www.ncbi.nlm.nih.gov/sites/entrez?db=UniGene) (Altschul et al., 1990; Benson et al. 2007). Regions from *A. annua* were considered to contain potential conserved miRNAs when the alignments with the miRBase set had (i) less than two mismatches, (ii) no gaps, (iii) maximum *e* value of 0.001 (Meyers et al. 2008; Zhang et al. 2009).

MicroRNA precursor prediction

A region comprising 150 nucleotides upstream and 150 nucleotides downstream from the predicted conserved microRNA was extracted from the *A. annua* UniGenes using *fastacmd* (Altschul et al. 1990). The RNA folding of these regions was analyzed using *MFold*, *RNA 3.0* (Zuker 2003). A region was considered a true miRNA precursor when they formed a hairpin-like structure with less than six mismatches between the mature region and the complementary strand, a maximum of three asymmetric bulges (with a maximum length of three bases) in the whole structure, and low minimum folding energy (MFE) less than −30 kcal/mol (Meyers et al. 2008; Ambros et al. 2003). The precursor region was the delimited to the region forming the low-energy hairpin. The significance of the MFE (*P* value) was calculated to discard effects due to dinucleotide frequencies, for this 1,000 random sequences were generated for each potential precursor with the same base composition, dinucleotide frequency and length, using *uShuffle* the random sequences were then folded using *MFold*, *RNA 3.0* (Zuker 2003, Jiang et al. 2008). The *P* value was calculated as the proportion of structure with a MFE equal or lower than the predicted precursor's (Freyhult et al. 2005). True microRNA precursors must have *P* value lower than 0.05.

MicroRNA target prediction and UniGene annotations

Potential targets for the mature miRNAs with identified precursors in *A. annua* were predicted by screening against the *A. annua* UniGenes using a modified version of the software *miRanda* (v. September 2008) with parameters described elsewhere (Enright et al. 2003; Pérez-Quintero et al. 2010). The criteria to consider a sequence as a putative miRNA target were: four or fewer mismatches overall, only one or none mismatches in the 5′ region of the miRNA (positions 1–12), no more than two consecutives mismatches in positions 13–21, no mismatches in positions 10 and 11. Additionally, the miRNA: target pair should have low free-energy of bonding (less than −20 kcal/mol). These criteria are based on experimental work and have been extensively used for miRNA target prediction in plant (Schwab et al. 2005; Zhang et al. 2006). The best targets for each microRNA with a *miRanda* score over 70 were kept and annotated (a total of 51 targets). Target miRNA UniGene sequences were functionally annotated with Gene Ontology (GO) terms by the Blast2GO (B2G) tool (Conesa et al. 2005; Conesa and Götz 2008; Götz et al. 2008).

Phylogenetic analyses

Sequences for complete pre-miRNAs from all available plants were obtained from miRBase (version 15.0, April 2010) (Griffiths-Jones et al. 2008). Sequences were aligned using ClustalW (Larkin et al. 2007) and T-coffee (Notredame 2010). Genetic distances (neighbor joining method) and trees were made using ClustalW2 with default parameters (Larkin et al. 2007).

Results and discussion

Gene-oriented clusters of transcripts sequences of *A. annua* were analyzed to identify miRNAs; 138 regions from the *A. annua* UniGenes were found containing the potentially conserved miRNAs. RNA folding of these regions was analyzed resulting in a total of 11 true miRNA precursors from nine families. Information about these precursors is summarized in Table 1 and their respective folded structures are shown (Fig. 1). The regions that matched conserved miRNAs but did not formed the characteristic precursor folding may correspond to miRNA targets, true miRNAs derived from non-canonic precursors, or simply short regions resembling miRNAs. A low representation of miRNA precursors in UniGenes is expected, since many of these precursors are not sequenced in cDNA libraries. The frequency of miRNAs in *A.annua* corresponds with the previous frequencies where a miRNA precursor is found each 10000 ESTs (Xie et al. 2007; Zhang et al. 2006) (Table 2).

According to the plant miRNA database PMRD, the mean number of miRNAs per genome is 75 and the median is 2 (Zhang et al. 2010). The differences in number of miRNA per genome are mainly attributed to different stages of sequencing and annotation of

Table 1 *A. annua* UniGenes regions containing conserved microRNA precursors

miRNA Family	Unigene	Strand	Precursor region	miRNA region	MFE (kcal/mol)	*P* value	Precursor length	miRNA length
156	gnl|UG|Aan#S55206502	−	153–235	215–234	−32.2	0	82	19
162	gnl|UG|Aan#S55209873	+	5–110	79–99	−45.3	0	105	20
166	gnl|UG|Aan#S55143299	−	226–319	284–302	−42.6	0	93	18
167	gnl|UG|Aan#S55097120	+	11–156	119–140	−60.8	0	145	21
168	gnl|UG|Aan#S55198463	−	562–667	626–646	−56.9	0	105	20
172	gnl|UG|Aan#S55152608	+	91–204	167–187	−47.8	0	113	20
390	gnl|UG|Aan#S55096165	−	127–234	195–215	−48.1	0	107	20
393a	gnl|UG|Aan#S41909146	+	397–494	442–463	−49	0	97	21
393b	gnl|UG|Aan#S55186564	+	100–250	130–151	−54.8	0	150	21
393c	gnl|UG|Aan#S55190344	+	55–150	69–90	−47.3	0	95	21
397	gnl|UG|Aan#S55145898	+	1–130	37–56	−46.1	0	129	19

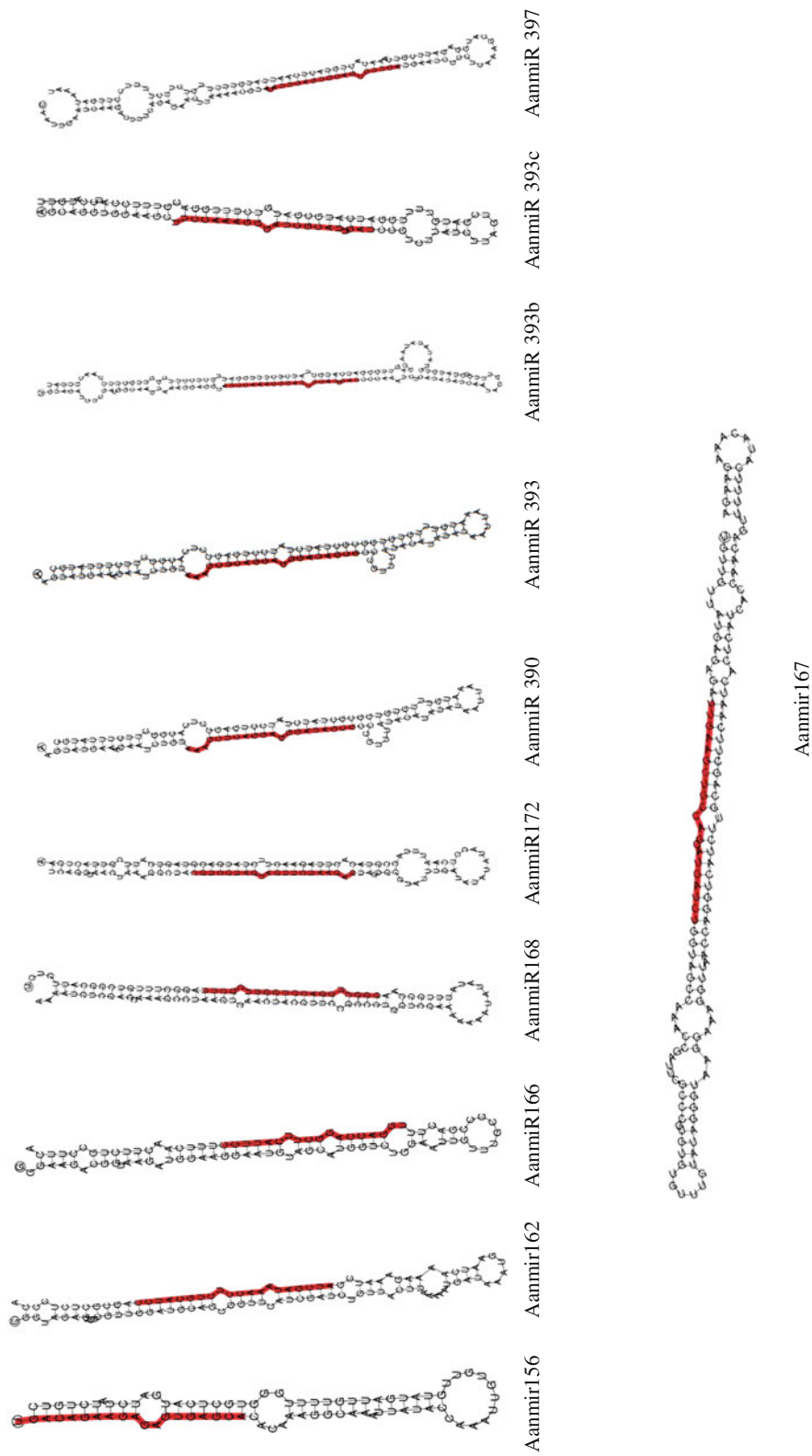

Fig. 1 Folded pre-miRNA structures for identified miRNA families in *Artemisia annua*

Table 2 Predicted miRNA targets for identified miRNAs in *A. annua* UniGene set

Family	Target scaffold	miRNA cleavage region		Energy (kcal/mol)	Description
156	Aan#S55004256	14	32	−36.42	LIGULELESS1 protein, putative
156	Aan#S41952363	775	793	−32.5	Teosinteglume architecture 1
156	Aan#S41953101	343	361	−32.5	Unknown
156	Aan#S41970295	409	427	−32.5	Unknown
156	Aan#S55012000	365	383	−32.5	Transcription factor squamosa promoter binding
156	Aan#S55019143	523	541	−32.5	Squamosa promoter-binding
156	Aan#S55019731	322	340	−32.5	spl3
156	Aan#S55037614	114	132	−32.5	Unknown
156	Aan#S55096756	198	216	−32.5	Unknown
156	Aan#S55105151	285	303	−32.5	Squamosa-promoter binding protein
156	Aan#S55216017	36	54	−32.5	Unknown
156	Aan#S55171092	24	42	−32.07	Unknown
156	Aan#S55031563	553	571	−31.59	Squamosa promoter-binding
156	Aan#S55207571	429	447	−31.59	Squamosa promoter-binding
156	Aan#S55130450	35	53	−27.45	Unknown
162	Aan#S55042603	207	227	−23.62	Arginine-trna-proteintransferase 1 homolog
162	Aan#S55160847	205	223	−23.04	Unknown
166	Aan#S55100423	250	267	−36.78	Unknown
166	Aan#S55033188	184	201	−36.44	Homeoboxleucine-zipper protein
166	Aan#S55196989	777	794	−30.39	Protein
167	Aan#S55102869	88	108	−27.1	Unknown
167	Aan#S55202306	151	170	−26.82	26s proteasome non-atpase regulatory subunit 13
167	Aan#S55188900	45	65	−26.35	gtp binding
167	Aan#S55058255	45	65	−23.7	Clathrin heavy
167	Aan#S55177736	14	35	−23.19	Unknown
167	Aan#S55024523	288	308	−23.12	Unknown
168	Aan#S55033730	194	213	−27.19	Unknown
172	Aan#S41934784	310	329	−35.69	Unknown
172	Aan#S41934785	600	619	−35.69	ap2 domain-containing transcription factor
172	Aan#S55194368	207	226	−26.11	Unknown
172	Aan#S54989082	290	309	−24.88	Unknown
172	Aan#S55184380	434	454	−24.22	Protein
172	Aan#S55114559	125	144	−23.21	Protein
390	Aan#S55192778	178	197	−37.44	Unknown
390	Aan#S55083216	173	192	−35.37	Unknown
390	Aan#S55035787	201	221	−34.43	Unknown
390	Aan#S55021073	390	409	−32.01	Unknown
390	Aan#S55123496	104	123	−24.61	bwf1-like protein
390	Aan#S55056498	226	245	−24.54	Protein
390	Aan#S55136478	227	246	−24.54	Protein
390	Aan#S55207716	315	334	−24.54	Protein
393	Aan#S41943802	486	506	−37.54	Transport inhibitor response 1
393	Aan#S41963250	617	637	−37.54	Transport inhibitor response 1

Table 2 continued

Family	Target scaffold	miRNA cleavage region		Energy (kcal/mol)	Description
393	Aan#S55147011	84	104	−37.54	Transport inhibitor response 1
393	Aan#S55097866	101	122	−32.07	Aquaporin
393	Aan#S55167746	205	226	−30.05	Unknown
393	Aan#S41955858	60	80	−25.82	dnaj heat shock
393	Aan#S55131804	344	364	−25.45	Glycosylhydrolase family 18 protein
397	Aan#S55021978	145	163	−33.61	Diphenoloxidase
397	Aan#S55062318	39	57	−29.6	Unknown
397	Aan#S41968868	206	224	−26.56	Unknown

Table 3 Genome sizes, numbers of genes and miRNAs for six plant species (from NCBI and PRMD)

Organism	miRNAs (PMRD)	miRNAs (miRBase)	Genome size (MB)	UniGenes	Number of genes
P. trichocarpa	2780	234	486	15,583	42,966
O. sativa	2641	491	430	41,317	30,697
A. thaliana	1441	232	63.4	30,67	33,500
P. patens	282	229	500	18,889	36,006
Z. mays	269	170	2500	98,696	26,037
G. max	228	203	975	36,295	5,059

different plant genomes. Using information from PMRD, miRBase and NCBI, we summarized miRNA data for six well-annotated plant species (genome sizes, numbers of genes and UniGenes, and number of miRNA) (Table 3). Differences in number of miRNA for various species can be explained by (i) differences in genome sizes and genome organization, (ii) the extent of sequencing studies aimed to study miRNAs in different plants (Chuang and Jones 2007). The differences between the miRNAs reported in the data bases are explained because miRBase includes only miRNAs with wet-lab experimental confirmation (Northern Blot, qPCR) while PMRD does not (Griffiths-Jones et al. 2008; Zhang et al. 2010).

Unsurprisingly the miRNA families found in *A. annua* are mainly families with great conservation across plant lineages. Among the families found, miR156, miR166 and miR390 are conserved in all land plants (Embriophyta), miR397 is conserved in all vascular plants (Spermatophyta) and miR162, miR167, miR168, miR172 and miR393 are conserved in all angiopserms (monocots and dicots) (Cuperus et al. 2011). Many targets (57%) correspond to unknown proteins, which may be due mainly to the incomplete stage of *A. annua* genome sequencing and annotation, and 12 (23%) correspond to transcription factors, which is usual for conserved plant miRNAs (Jones-Rhoades et al. 2006). Squamosa promoter-binding transcription factors were identified as targets for miR156 and also transport inhibitor response 1 (TIR1) transcription factors as targets for miR393, which has also been identified and confirmed experimentally in *A. thaliana* and other plants, indicating that the miRNA-mediated mechanisms and not only the miRNAs are conserved across plants (Jones-Rhoades et al. 2006; Garcia 2008; Yamasaki et al. 2009; Ruegger et al. 1998; Jones-Rhoades and Bartel 2004). The predominance of conserved miRNAs among the *A. annua* UniGene collections may be due to the fact that conserved miRNAs are usually more expressed than novel or phylogenetically restricted miRNAs and are thus more easily synthesized and sequenced in cDNA libraries (Cuperus et al. 2011).

Most pre-miRNAs were identified in contigs constructed form one cDNA (singletons) and only miR390

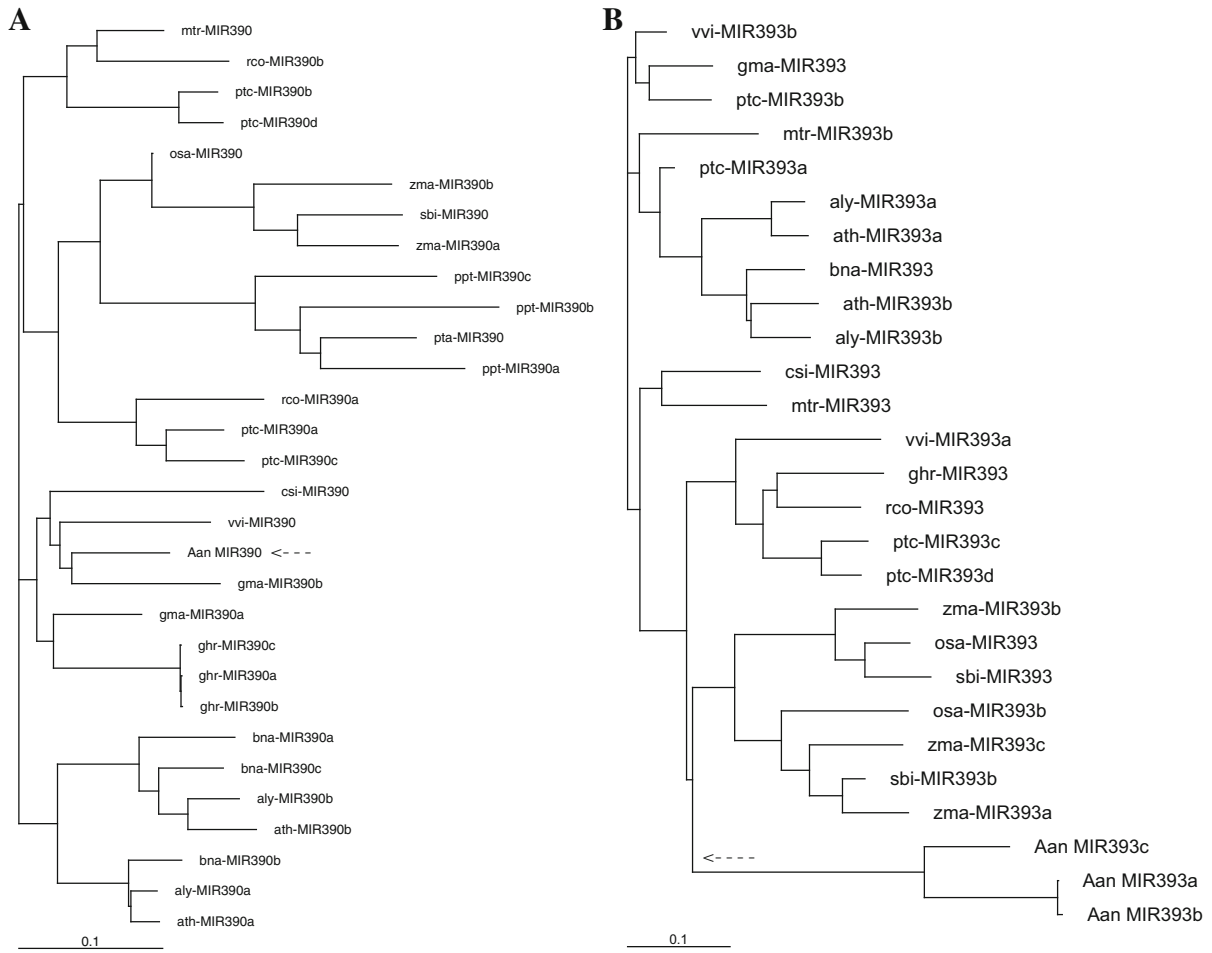

Fig. 2 Phylogenetic (neighbor-joining) trees made using pre-miRNA sequences available from all plants. **a** Phylogenetic tree for miR390. **b** Phylogenetic tree for family miR393. Multiple alignments were made using Clustal W. *Artemisia annua* in the tree is shown with an arrow. Plant names *Aan* = *Artemisia annua, aly* = *Arabidopsis lyrata, ath* = *Arabidopsis thaliana, bdi* = *Brachypodium distachyon, bna* = *Brassica napus, csi* = *Citrus sinensis, gma* = *Glycine max, ghe* = *Gossypium herbacium, ghi* = *Gossypium hirsutum, gra* = *Gossypium raimondii, mtr* = *Medicago truncatula, osa* = *Oryza sativa, ppt* = *Physcomitrella patens, pta* = *Pinus taeda, ptc* = *Populus trichocarpa, rco* = *Ricinus communis, sbi* = *Sorghum bicolor, vvi* = *Vitis vinifera, zma* = *Zea mays*

was identified in three different cDNA libraries, indicating that it might be more transcriptionally active than other miRNAs. In other plants miR390 is known to be involved in auxin signaling in various developmental stages via the generation of trans-acting small RNAs (Axtell et al. 2007; Yoon et al. 2010). In this work miR390 was found in libraries coming from young and mature leaf trichomes as well as meristems. Importantly, miR390 is also involved in trichome patterning and development, with mutants showing precocious production of trichomes (Garcia 2008; Adenot et al. 2006), which makes it an interesting candidate for *A. annua* transformation aiming to produce more artemisinin in trichomes.

Families miR156, miR162, miR168 and miR393 were found in libraries coming from flower bud meristems which may be due to this tissue's over representation in *A. annua* UniGene (Graham et al. 2010; Wang et al. 2009), although a role in flower development was reported for some of these families (Gandikota et al. 2007). Additionally, miR166 and miR397 families were found in cotyledon tissues, family miR167 in young leaf trichomes and family miR172 in mature leaves.

With the aims of assessing miRNA-mediated regulation of artemisinin biosynthesis we searched specifically for miRNA targets of identified miRNAs among annotated genes in the artemisinin biosynthesis pathway as identified (Lin et al. 2011; Olofsson et al. 2011). However no miRNA targets were found among these genes, which may be due to the miRNAs being conserved and having stable evolutionary functions. If miRNA-mediated regulation of artemisinin synthesis exists this may be exerted primarily by novel or clade-specific miRNAs which cannot be identified at the current status of the *A. annua* transcriptome.

Finally, we further explored the phylogenetic analyses to assess the conservation of primary pre-miRNA sequences across plant lineages with reference to *A. annua*. Phylogenetic trees were made for each miRNA family identified in *A. annua* including all plant pre-miRNAs available for each family in miRBase (Griffiths-Jones et al. 2008). *A. annua* is an Asteraceae, and currently there are no miRNAs identified in other members of these families in miRBase (Griffiths-Jones et al. 2008), the closest relatives to *A. annua* with identified miRNAs are *Solanum lycopersicum* and *Rehmannia glutinosa*. It was expected that at the sequence level *A. annua* pre-miRNAs would be similar to those of other plants within the clade; however, different results were found for each family. Phylogenetic trees for miR162, miR165, miR167, miR168 and miR390 grouped *A. annua* with other eudicots as expected, mainly *Glycine max, Vitis vinifera, Citrus sinensis* and *Medicago truncatula*. Phylogenetic depiction of miR390 is shown as an example of this group pattern (Fig. 2a). The trees for miR172, miR393 and miR397 grouped *A. annua* with plants in the Poaceae family including *Oryza sativa, Zea mays, Sorghum bicolor*. Phylogenetic analysis of miR393 tree presents an example of this grouping and the miR166 tree grouped *A. annua* closest to the moss *Physcomitrella patens* (Fig. 2b). Similar results were obtained doing alignments with ClustalW or T-Coffee. This shows that primary sequence of pre-miRNAs in most cases is not a good indicator of phylogenetic relations between plants.

Conclusions

The computational identification of microRNAs and their possible targets have been made using the UniGene dataset (88,174) of *A. annua*. The identified targets probably play important roles in plant development and signaling. We further stress that the utilization of the pre-miRNAs sequence for phylogenetic inference is not a universal factor for inference of phylogenetic relationship.

Acknowledgments Gaurav Sablok thanks Key Laboratory of Horticultural Plant Biology (MOE), Huazhong Agricultural University. Tatiana Tatarinova would like to thank the University of Glamorgan's Research Investment Scheme for supporting this project and Dr Owain Kerton for editing. Financial support for Camilo López and Alvaro Perez comes from Dirección de Investigaciones sede Bogota (Universidad Nacional), Colciencias, and Ministerio de Agricultura de Colombia. This work has been partially funded by Spanish MICINN grant BIO2009-10799.

References

Adenot X, Elmayan T, Lauressergues D, Boutet S, Bouché N, Gasciolli V, Vaucheret H (2006) DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. Current Biol 16:927–932

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X, Dreyfuss G, Eddy SR, Griffiths-Jones S, Marshall M, Matzke M, Ruvkun G, Tuschl T (2003) A uniform system for microRNA annotation. RNA 9:277–279

Axtell MJ, Snyder JA, Bartel DP (2007) Common functions for diverse small RNAs of land plants. Plant Cell 19:1750–1769

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281–297

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2007) GenBank. Nucleic Acids Res 35:D21–D25

Bouwmeester HJ, Wallaart TE, Janssen MH, Van LB, Jansen BJ, Posthumus MA, Schmidt CO, De Kraker JW, König WA, Franssen MC (1999) Amorpha-4, 11-diene synthase catalyses the first probable step in artemisinin biosynthesis. Phytochemistry 52:843–854

Chuang JC, Jones PA (2007) Epigenetics and microRNAs. Pediatr Res 61:24R–29R

Conesa A, Götz S (2008) Blast2GO: A comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics 2008:619832

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676

Cuperus JT, Fahlgren N, Carrington JC (2011) Evolution and Functional Diversification of MIRNA Genes. Plant Cell 23:431–442

Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS (2003) MicroRNA targets in Drosophila. Genome Biol 5:R1

Freyhult E, Prusis P, Lapinsh M, Wikberg JE, Moulton V, Gustafsson MG (2005) Unbiased descriptor and parameter

selection confirms the potential of proteochemometric modelling. BMC Bioinformatics 6:50

Gandikota M, Birkenbihl RP, Höhmann S, Cardon GH, Saedler H, Huijser P (2007) The miRNA156/157 recognition element in the 3′ UTR of the *Arabidopsis* SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. Plant J 49:683–693

Garcia D (2008) A miRacle in plant development: role of microRNAs in cell differentiation and patterning. Semin Cell Dev Biol 19:586–595

Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A (2008) High-throughput functional annotation with the Blast2GO suite. Nucleic Acids Res 36:3420–3435

Graham IA, Besser K, Blumer S, Branigan CA, Czechowski T, Elias L, Guterman I, Harvey D et al. (2010) The genetic map of *Artemisia annua* L. identifies loci affecting yield of the antimalarial drug Artemisinin. Science 327:328–331

Griffiths-Jones S, Saini HK, Van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. Nucleic Acids Res 36:D154–D158

Gupta PK, Rustgi S (2004) Molecular markers from the transcribed/expressed region of the genome in higher plants. Funct Integr Genomics 4:139–162

Jiang M, Anderson J, Gillespie J, Mayne M (2008) uShuffle: a useful tool for shuffling biological sequences while preserving the k-let counts. BMC Bioinformatics 9:192

Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell. 14:787–799

Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAS and their regulatory roles in plants. Annu Rev Plant Biol 57:19–53

Kidner CA, Martienssen RA (2005) The developmental role of microRNA in plants. Curr Opin Plant Biol 8:38–44

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948

Lin X, Zhou Y, Zhang J, Lu X, Zhang F, Shen Q, Wu S, Chen Y, Wang T, Tang K (2011) Enhancement of artemisinin content in tetraploid *Artemisia annua* plants by modulating the expression of genes in artemisinin biosynthetic pathway. Biotechnol Appl Biochem 58:50–57

Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D, Bowman JL, Cao X, Carrington JC et al (2008) Criteria for annotation of plant MicroRNAs. Plant Cell. 20:3186–3190

Notredame C (2010) Computing Multiple Sequence/Structure Alignments with the T-Coffee Package. Curr Protoc Bioinformatics Chapter 3: pp 1–25

Olofsson L, Engström A, Lundgren A, Brodelius PE (2011) Relative expression of genes of terpene metabolism in different tissues of *Artemisia annua* L. BMC Plant Biol 11:45

Parida SK, Anand Raj Kumar K, Dalal V, Singh NK, Mohapatra T (2006) UniGene derived microsatellite markers for the cereal genomes. Theor Appl Genet 112:808–817

Pérez-Quintero AL, Neme R, Zapata A, López C (2010) Plant microRNAs and their role in defense against viruses: a bioinformatics approach. BMC Plant Biol 10:138

Robert-Seilaniantz A, Bari R, Jones JDG (2010) A biotic or abiotic stress? In: Pareek A, Sopory SK, Bohnert HJ, Govindjee (eds) Abiotic stress adaptation in plants: physiological. Molecular and Genomic Foundation, Springer, Dordrecht, pp 113–116

Ruegger M, Dewey E, Gray WM, Hobbie L, Turner J, Estelle M (1998) The TIR1 protein of Arabidopsis functions in auxin response and is related to human SKP2 and yeast grr1p. Genes Dev 12:198–207

Rydén AM, Ruyter-Spira C, Litjens R, Takahashi S, Quax W, Osada H, Bouwmeester H, Kayser O (2010) Molecular cloning and characterization of a broad substrate terpenoidoxidoreductase from *Artemisia annua*. Plant Cell Physiol 51:1219–1228

Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D (2005) Specific effects of microRNAs on the plant transcriptome. Dev Cell 8:517–527

Wang W, Wang Y, Zhang Q, Qi Y, Guo D (2009) Global characterization of *Artemisia annua* glandular trichome transcriptome using 454 pyrosequencing. BMC Genomics 10:465

Xie FL, Huang SQ, Guo K, Xiang AL, Zhu YY, Nie L, Yang ZM (2007) Computational identification of novel microRNAs and targets in Brassicanapus. FEBS Lett 581:1464–1474

Yamasaki H, Hayashi M, Fukazawa M, Kobayashi Y, Shikanai T (2009) SQUAMOSA Promoter Binding Protein-Like7 Is a Central Regulator for Copper Homeostasis in *Arabidopsis*. Plant Cell 1:347–361

Yoon EK, Yang JH, Lim J, Kim SH, Kim SK, Lee WS (2010) Auxin regulation of the microRNA390-dependent transacting small interfering RNA pathway in Arabidopsis lateral root development. Nucleic Acid Res 38:1382–1391

Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA (2006a) Conservation and divergence of plant microRNA genes. Plant J 46:243–259

Zhang B, Pan X, Wang Q, Cobb GP, Anderson TA (2006b) Computational identification of microRNAs and their targets. Comput Biol Chem 30:395–407

Zhang J, Xu Y, Huan Q, Chong K (2009) Deep sequencing of *Brachypodium* small RNAs at the global genome level identifies microRNAs involved in cold stress response. BMC Genomics 10:449

Zhang Z, Yu J, Li D, Zhang Z, Liu F, Zhou X, Wang T, Ling Y, Su Z (2010) PMRD: plant microRNA database. Nucleic Acids Res 38:D806–D813

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31:3406–3415