Available online at www.sciencedirect.com
ScienceDirect
Plant Physiology and Biochemistry 46 (2008) 739–751
Plant Physiology and Biochemistry
www.elsevier.com/locate/plaphy

Research article

Identification of micro-RNAs in cotton

Muhammad Younas Khan Barozai [a,b,*], Muhammad Irfan [b], Rizwan Yousaf [b], Imran Ali [b], Uzma Qaisar [b], Asma Maqbool [b], Muzna Zahoor [b], Bushra Rashid [b], Tayyab Hussnain [b], Sheikh Riazuddin [b]

[a] *Botany Department, University of Baluchistan, Quetta, Pakistan*
[b] *Centre of Excellence in Molecular Biology, 87 West Canal Bank Road, Thoker Niaz Baig, Lahore, Pakistan*

Received 28 April 2008
Available online 28 May 2008

Abstract

The plant genome has conserved small non-coding microRNAs (miRNAs) genes about 20–24 nucleotides long. They play a vital role in the gene regulation at various stages of plant life. Their conserved nature among the various organisms not only suggests their early evolution in eukaryotes but also makes them a good source of new miRNA discovery by homology search using bioinformatics tools. A systematic search approach was used for interspecies orthologues of miRNA precursors, from known sequences of *Gossypium* in GenBank. The study resulted in 22 miRNAs belonging to 13 families. We found 7 miRNA families (miR160, 164, 827, 829, 836, 845 and 865) for the first time in cotton. All 22 miRNA precursors form stable minimum free energy (mfe) stem loop structure as their orthologues form in *Arabidopsis* and the mature miRNAs reside in the stem portion of the stem loop structure. Fifteen miRNAs belong to the world's most commercial fiber producing upland cotton (*Gossypium hirsutum*), five are from *Gossypium raimondii* and one each is from *Gossypium herbaceum* and *Gossypium arboreum*. Their targets consist of transcription factors, cell division regulating proteins and virus response gene. The discovery of 22 miRNAs will be helpful in future for detection of precise function of each miRNA at a particular stage in life cycle of cotton.

Keywords: Cotton; Micro RNAs; Post-transcriptional gene silencing; Homology search

1. Introduction

Micro RNA genes that encode miRNAs reside in a specific genomic region. The non-coding RNA such as miRNA, transfer RNA (tRNA) and others constitute 3% out of the total 5% of the functional genome [1]. miRNAs are endogenous, non-coding, small RNAs about 20–24 nucleotides long [2] and are conserved in plants and animals [3,4]. They have a crucial role in post-transcriptional gene regulation [5,6]. Mature miRNAs are produced by a chain of reaction with the help of enzymes [7]. Primary transcripts of mature miRNAs (pri-miRNAs) fold into a stable stem loop structure forming miRNA precursor (pre-miRNA). The loop of pre-miRNA is cleaved, producing a short double-stranded RNA (dsRNA); a single strand of the dsRNA acts as mature miRNA [8]. The processing occurs in nucleus and is processed by a special RNaseIII-like endonuclease, Drosha and Dicer in animals [9] and Dicer-like enzyme (DCL) in plants [10], that also predominantly incorporate the mature miRNA into the RNA induced silencing complex (RISC) [8]. The RISC complex negatively regulates gene expression either by inhibiting translation

Abbreviations: ath, *Arabidopsis thaliana*; BLAST, basic local alignment search tool; ch_ratio, core hairpin ratio; DCL, dicer-like enzyme; dsRNA, double-stranded RNA; ESTs, expressed sequence tags; ghr, *Gossypium hirsutum*; mRNA, messenger RNA; MIR, micro-RNA; miRNAs, microRNAs; mfe, minimum free energy; pre-miRNAs, miRNAs precursor; NCBI, National Center for Biotechnology Information; osa, *Oryza sativa*; Ptc, *Populus trichocarpa*; Ptn, position of terminal nucleotide; Pfn, position of the first nucleotide; pri-miRNAs, primary transcripts of mature miRNAs; rRNA, ribosomal RNA; RISC, RNA induced silencing complex; tRNA, transfer RNA; UTRs, untranslated regions.

* Corresponding author. Centre of Excellence in Molecular Biology, 87 West Canal Bank Road, Thoker Niaz Baig, Lahore, Pakistan. Tel.: +92 (0)33 781 7319.
E-mail address: barozaikhan@gmail.com (M.Y. Khan Barozai).

0981-9428/$ - see front matter © 2008 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.plaphy.2008.05.009

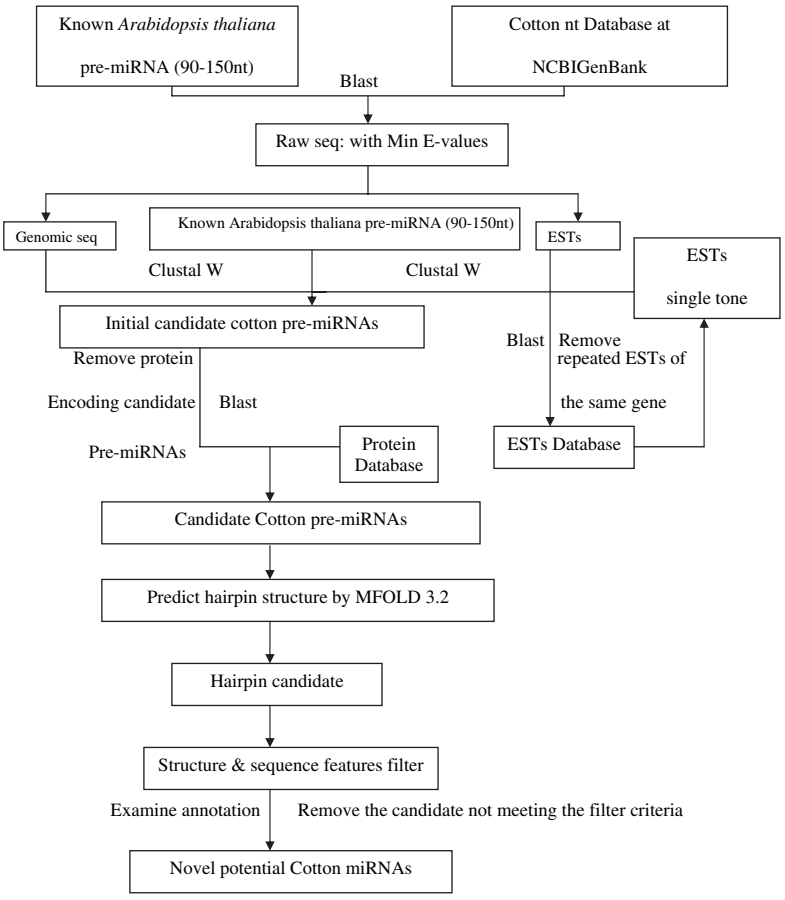

Fig. 1. Schematic representation of the cotton pre-miRNA search procedure used to identify homologues of known *Arabidopsis thaliana* pre-miRNAs.

elongation or by triggering messenger RNA (mRNA) destruction on the basis of the degree of complementarity of miRNA within its target [11,12]. Most animal-mRNA targets have many weak miRNA complementary sites, so miRNA imperfectly complement these sites and suppress gene expression [6,13]. The plant mRNA targets have single and perfect or near perfect miRNA complementary sites, so corresponding miRNAs perfectly complement these sites and trigger the mRNA degradation [14]. Usually the 3′ untranslated regions (UTRs) of the mRNA targets contain the miRNA complementary sites [6].

The first finding of miRNA in *lin-4* and *let-7* mutants of *Caenorhabditis elegans* [15,16] has helped in the discovery of miRNAs in plants [17] and animal-species [18].

The miRNAs perform versatile functions in plant and animals, e.g. in development [19], organ morphogenesis [14,19], signaling pathway [20], transgene repression [21], abiotic stresses [22,23], disease progression [24], and parasitism for the host cell invasion by viruses [25].

Most miRNAs are conserved in animals and plants and from animals to plants [4,17,26], suggesting early development of the miRNAs in the evolution. The conserved nature also indicates their conserved function in different organisms. The conserved nature of these miRNAs becomes a powerful strategy for identification of new orthologues by homology

Table 1
Raw sequences accession numbers, nature, maximum score of homology with known *Arabidopsis thaliana* pre-miRNAs and E-values

Arabidopsis thaliana pre-miRNA	Raw sequences			
	Accession #	Nature of nucleotide	Max score	*E*-value
ath-MIR156e	BV679360	DNA	72	5e-12
ath-MIR157a	CO076888	EST	62	4e-09
ath-MIR157a	CO082782	EST	62	4e-09
ath-miR160b	BQ401391	EST	40	0.001
ath-miR160b	CO095743.1	EST	40	0.001
ath-miR164c	DR461140	EST	40	0.001
ath-MIR166e	DQ908436	DNA	38	0.006
ath-miR171a	CO129257	EST	42	4e-04
ath-MIR390a	DW238152	EST	64	1e-09
ath-MIR390a	DW518163	EST	64	1e-09
ath-MIR399f	AY632360	DNA	44	0.001
ath-MIR399f	AY632359	DNA	44	0.001
ath-miR827	DW476866	EST	40	0.001
ath-miR829	DX525305	DNA	40	0.002
ath-miR829	DX531919	DNA	40	0.002
ath-miR829	DX53349	DNA	40	0.002
ath-miR829	DX541807	DNA	40	0.002
ath-miR829	DX545537	DNA	40	0.002
ath-miR829	DX546324	DNA	40	0.002
ath-miR836	DX561980	DNA	38	0.009
ath-miR845a	AC197184	DNA	38	0.006
ath-miR865-5p	DX543587	DNA	38	0.006