Comparative miRNAome Analysis Revealed Numerous Conserved and Novel Drought Responsive miRNAs in Cotton (Gossypium spp.)

Boopathi M.N. 1, 2, Sathish S. 1, Kavitha P. 1, Dachinamoorthy P. 1, Ravikesavan R. 3

1 Department of Plant Molecular Biology and Bioinformatics, Center for Plant Molecular Biology and Biotechnology, India
2 Department of Fruit Crops, Horticultural College and Research Institute, TNAU, Periyakulam, 625604, India
3 Department of Millets, CPBG, Tamil Nadu Agricultural University (TNAU), Coimbatore, 641003, India

Corresponding Author, nmbboopathi@tnau.ac.in

Cotton Genomics and Genetics, 2016, Vol. 7, No. 2 doi: 10.5376/cgg.2016.07.0002

Abstract Negative regulations of gene expression by endogenous, non-coding miRNAs have been shown to play important role in abiotic stress responses in plants. However, limited knowledge is available on water stress responsive miRNAs in cotton. In this study, differentially expressed drought responsive miRNAs and their targets were identified under natural field conditions through high-throughput small RNA sequencing by comparing leaf samples of drought tolerant Gossypium hirsutum cv. KC3 and drought susceptible G. barbadense cv. Suvin. Totally four small RNA libraries were constructed and sequenced by employing ion proton™ technology. Altogether, there were 39 407 089 reads with a mean read length of 22 nt. In KC3, there were 5 138 unique miRNA reads that were differentially expressed with at least two folds under water stressed conditions. In contrast, Suvin have shown 8 469 unique miRNA reads that were differentially expressed with minimum of two folds under water stressed conditions. Comparison of miRNAs expressed under water stressed conditions between KC3 and Suvin, have resulted 7 494 miRNA reads and interestingly majority of them were down regulated with at least two folds. Besides identifying large numbers of novel miRNAs, several abiotic stress responsive conserved miRNAs were also noticed. Of particular interest were miR750, miR2, miR14, miR276, miR279, miR-bantam and miR5176 that were highly down regulated in KC3 under water stress conditions. Strikingly, miR2 and miR-bantam were previously shown to target pre-apoptotic genes in biological systems. Further, the identified miRNAs were also targeting different classes of dehydrogenases, protein kinases and transcription factors. Our results revealed for the first time that there were large numbers of water stress related miRNAs that might be sequentially and/or complexly involved in gene regulation that confers drought tolerance in cotton under field conditions and they have enormous potential in elucidating the molecular mechanism of miRNA based gene regulation and more importantly in genetic improvement of drought tolerance in cotton. Further, this is the first report on experimentally identifying miRNAs in G. barbadense.

Keywords Cotton; Drought tolerance; MicroRNAs; Next generation sequencing; Transcriptomics

Introduction Diverse types of ribonucleic acids (RNA) with various functions such as coding, noncoding/non-messenger, structural, enzymatic and regulatory functions were reported so far. Among them, RNAs with regulatory activities are termed as micro-RNAs (miRNAs). They are about 22-24 nucleotide (nt) in length and found in all the living organisms including crop plants (Ambros, 2004). As negative regulators of plant gene expression, miRNAs can function by means of minimum two ways: 1) miRNAs base pair with mRNA targets by precise or nearly precise complementarity and cleave / destruct the target mRNA through the RNA interference machinery and 2) they inhibit protein synthesis through an unknown mechanism (Tang et al., 2003). Some known miRNA genic loci form clusters in the genome and these miRNA clusters and the miRNAs in a given cluster are often expressed simultaneously (Wei et al., 2013). Confirmations for the role of miRNA in cell growth and proliferation, developmental timing and phase switch such as vegetative to reproductive growth and response to biotic and abiotic stresses are increasing endlessly and now it is largely deliberated that their regulatory role is more persistent than what was assumed earlier.
Cotton (Gossypium spp.) is a highly preferred natural textile fiber because of its economic value and comfort and cotton seed oil is a significant food source for human and livestock. Gossypium consists of at least 45 diploid and 5 allotetraploid species; however, only four species viz., G. hirsutum, G. barbadense, G. arboreum and G. herbaceum (described in the order of highest to lowest global acreage), have been domesticated for their abundant seed trichomes and provide the foundation for the global textile industry. India has the largest area dedicated to cotton production worldwide; however, the cotton fiber production is largely limited by water stress since more than 60% of this area is under rainfed (Boopathi and Pathmanaban, 2012). Genetic improvement of cotton with drought resistance will be a viable and cost effective option to mitigate such losses. Though significant advancement has been documented, the progress in this direction is not up to the expected level due to lack of knowledge on drought resistance mechanisms. Although progress has been made in unraveling the complex stress response mechanisms, particularly in the identification of water stress responsive protein-coding genes, the key factors such as miRNAs that regulate the expression of these genes remain unclear. Such gap severely hampers the precise manipulation of drought resistance in cotton. Even though there is no direct proof that miRNAs are the crucial regulatory elements in development of drought tolerance in cotton, we postulate that differential expression of miRNAs might be involved in such complex process. As a first step towards the understanding of their regulatory mechanisms and networks of target genes in drought tolerance in cotton, differential expression of miRNAs between drought tolerant (Gossypium hirsutum cv. KC3) and drought susceptible (G. barbadense cv. Suvin) cotton lines were compared in the present study under water stressed and irrigated conditions in the field using a deep sequencing approach.

In general, miRNAs are being identified in plants by forward genetics (whereby miRNAs are isolated from an organism showing abnormal phenotypic characteristics), reverse genetics (whereby a specific gene is knocked out or over expressed), cDNA sequencing and computational prediction approaches (Ambros, 2004). In the early days, the identification of cotton miRNAs depended largely on computational approaches based on the conserved characteristics of miRNAs (Qiu et al., 2007; Zhang et al., 2007; Barozai et al., 2007), which was accomplished by searching the cotton genomic sequences and/or ESTs that were homologous to known mature miRNAs or pre-miRNAs. Later, after complete sequencing of G. raimondii (Wang et al., 2012) and G. arboreum (Li et al., 2014) it was proposed that there were totally 348 and 431 miRNA sequences, respectively in these two species.

The first experimental validation and exploration of cotton miRNAs by cloning and sequencing identified only three cotton miRNAs (Abdurakhmonov et al., 2008). However, recently deep sequencing has emerged as an effective approach for large-scale miRNA discovery. Small RNA libraries prepared from cotton ovules and seedlings have been analysed by high-throughput sequencing (Kwak et al., 2009; Ruan et al., 2009). Pang et al. (2009) reported 27 conserved and four novel miRNA families from G. hirsutum by sequencing the small RNAs of leaves and ovules. Along with data from microarrays and Northern blots, they examined the changing pattern of certain miRNAs and their targets. Similar comparative miRNAome analysis revealed seven fiber initiation-related miRNAs expressed in cotton ovules besides 36 novel miRNAs and two conserved miRNAs (Wang et al., 2012). Targets of these miRNAs were experimentally validated and they described complex regulatory networks that coordinate fiber initiation responses. Yin et al. (2012) reported genome-wide profiling of miRNAs and other small non-coding RNAs in the Verticillium dahliae inoculated cotton roots. Small RNAs and their targets were also identified during cotton somatic embryogenesis through high-throughput small RNA and degradome sequencing (Yang et al., 2013). Zhang et al. (2013) have identified 93 miRNAs, including 63 novel miRNAs that are involved in fiber initiation and seed development. Similarly, 257 novel miRNAs were reported by Xue et al. (2013) in elongating cotton fiber cells. In another comparative miRNA expression study, Wei et al. (2013) identified sixteen conserved miRNA families during anther development in genetic male sterile and wild type cotton. Chen et al. (2013) has studied genetic variation of miRNAs and their target genes and they were genetically mapped in cotton.

Despite of these progresses, only 378 mature miRNA sequences from cotton (Gossypium spp.), were described in miRBase (http://www.mirbase.org/cgi-bin/browse.pl; Release 21; June, 2014). As several thousands of miRNAs were identified in plants, this number is extremely lesser than the expected number of miRNAs in cotton and they
are remain to be recognized. Further, among the reported miRNAs, only very few that were specific to drought and other abiotic stress has been reported in cotton using computational (Boopathi and Pathmanaban, 2012) or experimental (Yin et al., 2011; Wang et al., 2013) approaches, even though such abiotic stress specific miRNAs have been expressively established in Arabidopsis, rice and other crops (Sunkar et al., 2008). Identification of miRNAs in cotton, in response to water stress, certainly have important implications for gene regulation under drought and such knowledge can impressively enhance the drought tolerance breeding efficiency besides contributing significantly to the goal of having a complete profile of miRNAs in cotton. Besides, knowledge on miRNA-guided stress regulatory networks would provide new tools for the genetic improvement of other abiotic stress tolerances. In addition, all the miRNA studies employed G. hirsutum, G. arboreum, G. herbaceum and G. raimondii. However, miRNA expressions in G. barbadense, which has superior fiber quality traits that suit to modern textile mills, remain to be explored. Hence, in view of the importance of miRNAs in regulation of water stress resistance in cotton, this study was conducted with the following objective: to capture and characterize the differentially expressed miRNAs in drought tolerant and susceptible cotton cultivars using next generation sequencing platform.

1 Results

1.1 Small RNA libraries

Totally four different small RNA libraries were constructed from the leaf tissues of KC3 and Suvin collected from two extreme water regimes and intensely sequenced using ion proton™ technology. After the deletion of low quality reads and several kinds of contaminant tags 9 826 785 clean reads from leaf tissues of KC3-C were obtained (Table 1). Similarly, clean reads of 8 672 895 from KC3-WS; 8 627 325 from Suvin-C and 8 017 500 from Suvin-WS were obtained (Table 1) and they were used for analysis of differentially expressed small RNAs. Altogether, there were 39 407 089 reads found in all the four small RNA libraries with a mean read length of 22 nt.

Table 1 Output details of the small RNA reads of four libraries

| Name of the library | Bases       | Reads     | Mean read length (nt) |
|---------------------|-------------|-----------|-----------------------|
| KC3-C               | 192,062,293 | 9,826,785 | 19                    |
| KC3-WS              | 192,360,514 | 8,672,895 | 22                    |
| Suvin-C             | 179,836,027 | 8,627,325 | 20                    |
| Suvin-WS            | 181,349,766 | 8,017,500 | 22                    |
| Total number of reads | 39,407,089 |           | 22                    |

The sequence reads of the each library was normalized to RPKM and each reads that were differentially expressed were recorded in RPKM (Table 2 and Figure 1). Small RNA reads with a size of 27 nt were found to be present in significantly greater number in all the four comparisons followed by 28 or 29 nt reads (Figure 1). Small RNAs in the size range of 24-33 nt that were differentially expressed in drought susceptible cotton cultivar, Suvin, under two different water regimes were significantly higher than the other three comparisons. However, predominantly the size range of 20-23 nt were differentially expressed between KC3-WS and Suvin-WS (Figure 1). Totally 8 469 differentially expressed distinct reads were found between Suvin-C and Suvin-WS followed by 7 494 differentially expressed distinct reads between KC3-WS and Suvin-WS (Table 2). ESTs from NCBI and the genome sequence of G. raimondii (Wang et al., 2012), were used as reference sequences to predict the miRNA precursors since no allopolyploid cotton genome sequence was available at the time of this analysis. Nevertheless, to a maximum of 15.35% of the small RNA reads obtained from this study were mapped with G. raimondii genome sequences (Table 2). The distinct reads that were identified as rRNA, snRNA, snoRNA, tRNA, tmRNA during the analysis (number of reads in each small RNA class is given in Table 2) were separated and the remaining reads (which constitutes the majority of the reads; Table 2) were used further.

1.2 Differentially expressed conserved miRNA families

RPKM reads of each miRNA were used to notify their abundance and differential expression in two different contrasting cotton cultivars under two different water regimes. Simple blast analysis with matured miRNA
sequences deposited at miRBase has identified several conserved miRNAs (Supplementary Table 1). For effective comparison of results, only miRNAs that have shown more than two fold change in RPKM either in up or down regulation were considered and discussed hereunder.

Table 2 Statistics of differentially expressed small RNA sequence reads in two different cotton accessions under two different water regimes

| Description   | Differentially expressed reads in |
|---------------|----------------------------------|
|               | KC3-C Vs. KC3-WS | KC3-C Vs. Suvin-C | KC3-WS Vs. Suvin-WS | Suvin-C Vs. Suvin-WS |
| Total number^a| 5138              | 5341              | 7494 | 8469 |
| Match genome^b| 789 (15.35%)      | 759 (14.21%)      | 932 (12.43%) | 1250 (14.75%) |
| rRNA          | 4                 | 6                 | 6    | 7    |
| snRNA         | 1                 | 1                 | 5    | 3    |
| snoRNA        | 11                | 10                | 15   | 10   |
| tRNA          | 20                | 19                | 28   | 25   |
| tmRNA         | 0                 | 0                 | 0    | 1    |

Note: ^a Total number of distinct reads found in each comparison; ^b number of reads that match with G. raimondii genome

Figure 1 Sequence length distribution of small RNAs that were differentially expressed in two different cotton accessions under two different water regimes

Seventy one conserved miRNAs were found as differentially expressed between the miRNA libraries prepared from KC3-C and KC3-WS (Supplementary Table 1). Among them, 35 (belonging to 14 miRNA families) were up regulated and 36 (belonging to 21 miRNA families) were down regulated. Interestingly, members of three of the miRNA families (viz., miR166, miR167 and miR399) were both up and down regulated (Supplementary Table 1). Among the up-regulated miRNAs, miR398 family was highly induced under water stress in KC3 (38.880 folds), followed by miR397 (13.256 folds) and miR10 (8.242 folds). On the other hand, expression of miR166 was found to be highly down regulated under water stress in KC3 (54.840 folds), followed by miR479 (46.919 folds) and miR2111 (29.654) (Supplementary Table 1). Interestingly, several miRNA families were found to exhibit more than 10 fold change in their expression level in KC3-WS (for example, miR397 and miR398 were more than 10 fold up regulated and miR160, miR164, miR166, miR390, miR399, miR479, miR2111 and miR2796 were more than 10 fold down regulated; Supplementary Table 1). In contrast, some of the miRNA family had lowest fold change (miR535 was in 2.031 fold up regulation and miR160 was in 2.060 fold down regulation).

When comparing the miRNAs that were differentially expressed between the miRNA libraries prepared from drought tolerant, KC3 and drought susceptible, Suvin under water stress in the natural field conditions, totally 142 conserved miRNAs that belongs to 29 conserved families were noticed (Supplementary Table 1). It was found
that only two miRNAs of two different miRNA families were up regulated and 132 miRNAs of 23 miRNA families were down regulated in Suvin under water stressed conditions (Figure 2). In contrast, four miRNA families (viz., miR166, miR167, miR169 and miR482) were both up and down regulated (Supplementary Table 1; Fig. 2). Thirty five members from the single miR396 family were found to be down regulated during this analysis. Interestingly, there was a wide range of fold change (between 2.033 and 196.813) under down regulated miRNAs between these two libraries. miR750 was the most down regulated miRNA (196.813 folds) followed by miR2 (159.578 folds) and miR14 (76.243 folds). Strikingly, several miRNA families and their members have exhibited higher order of down fold expression (for example, miR279, miR276, hme-bantam, miR5176, miR482, miR3050, miR162, miR7499 and miR396 had 58.068, 57.625, 35.018, 33.688, 15.957, 15.514, 15.071, 11.525 and 11.456 fold down regulation, respectively; Supplementary Table 1). Similarly, under up regulated miRNAs, the fold change was also shown significant variation (between 2.085 and 6.050). Highest induction of miR169 (6.050 folds) was noticed in Suvin under water stress followed by miR166 (4.949 folds) (Supplementary Table 1).

Figure 2 Conserved miRNA families that were differentially expressed between KC3-WS and Suvin-WS

Supplementary Table 1 Conserved miRNA families that were differentially expressed in two cotton accessions under two different water regimes. (Please refer the attached Excel sheets 1-4) Comparison of miRNAs that were differentially expressed between KC3-C and Suvin-C identified 90 conserved miRNAs that belongs to 36 families (Supplementary Table 1). There were 14 miRNAs belonging to 11 miRNA families were up regulated, 44 miRNAs belonging to 18 miRNA families were down regulated and seven miRNA families were both up and down regulated (Supplementary Table 1). Remarkably, pronounced fold change (between 2.022 and 3473.119) under down regulated miRNAs was identified in this situation. In the same way, extensive variation in the fold change (between 2.024 and 38.453) under up regulated miRNAs was also noticed. The highest induction was observed for miR399 (38.453 folds) followed by miR166 (9.173 folds) and miR160 (7.819 folds) in Suvin under irrigated conditions and highest down regulation (perhaps, the highest down regulated miRNA that was found in this report) was seen for miR2118 (which recorded 3473.119 folds) followed by miR167 (75.310 folds) and miR482 (17.611 folds) (Supplementary Table 1).

Totally 73 conserved miRNAs derived from 31 miRNA families were identified as differentially expressed in drought susceptible, Suvin under two different water regimes (Suvin-C Vs. Suvin-WS) (Supplementary Table 1). Majority of the miRNAs (59 miRNAs belonging to 24 miRNA families) were down regulated. However, three miRNAs belonging to three different miRNA families were up regulated and four miRNA families (viz., miR167, miR169, miR482 and miR535) were both up and down regulated (Supplementary Table 1). Significant variation of fold change in up regulated miRNAs [between 2.163 (miR181) and 5.286 (miR169)] was noticed. On the other hand, under down regulated miRNA groups, there was a vast variation in fold change [i.e., between 2.011
(miR167) and 56.326 (miR14). It was further noticed that miRNA families, hme-bantam, miR2, miR14, miR156, miR160, miR166, miR279, miR2111, miR3027, and miR3050 were immensely downregulated and one member of miR169 family was significantly up-regulated in Suvin-WS (Supplementary Table 1).

### 1.3 Identification of novel miRNAs

In addition to conserved miRNAs, deep sequencing had also helped to identify new miRNAs in cotton that have not yet reported in miRBase. Since, the sequencing of the *G. hirsutum* and *G. barbadense* genome is incomplete and information on complete genome-wide cotton small RNA population is unknown, accurate identification of non-conserved miRNA in cotton is a difficult task. In order to ensure the truthful identification of miRNA and considering of the complicated nature of plant small RNA, a series of strict filtrations was used to enhance reliability of the result presented here. ESTs from NCBI and the genome sequence of *G. raimondii* (Wang et al., 2012) were used as reference sequences to predict several numbers of miRNA precursors. However, with the availability of genomic sequencing data from *G. hirsutum* and *G. barbadense* in near future, there is a greater potential to find several additional miRNAs of *Gossypium* with this same data sets.

Though several hundreds of new miRNA loci were identified as eligible candidates, following a BLASTn search and hairpin structure prediction, only 59 putative unique miRNAs were detected in all the four small RNA libraries (Table 3). The RNA structures were predicted by mFold software and manually checked according to the criteria of Meyers et al. (2008). The lowest minimum free energy (MFE) among all hairpin structures of the novel miRNAs precursors was −43.9 kcal/mol (Table 3). All precursors of novel miRNAs had regular hairpin structures and two of these (Ghi-miR(contig-4796) and Ghi-miR(contig-13785) are presented in Figure 3. GC content of the identified new miRNA precursor sequences was ranged from 23-71 percent (Table 3).

| Contig ID   | Small RNA sequence identified in this study                  | Length | MFE  | GC(%) |
|------------|----------------------------------------------------------------|--------|------|-------|
| Contig_5142| UUCUGAGAGGCUGCCACAUACUAUCA                                 |        | -74.7| 46    |
| Contig_4796| GCCUCCAAAGAGAAUACAUAGCACCAUGGUAUU                         |        | -121.1| 46    |
| Contig_8004| UGCCAAAGCGGGAAGCGAAG                                       |        | -66.0| 50    |
| Contig_16568|UUUCCAUCAUAUUAUUCGCAUG                                       |        | -104.7| 35    |
| Contig_13785|UUCACAAAGCGGGAAGCACCUUG                                      |        | -80.9| 50    |
| Contig_15645|UUUCCAGUCAAGCAGAAG                                       |        | -65.4| 41    |
| Contig_6913|AGAAUCAAGUUAUCAUAGAUGAAGACG                               |        | -46.1| 33    |
| Contig_16002|UAGUGUCCAGCAACACAAACC                                      |        | -46.7| 46    |
| Contig_14218|AGAUCGAGUAAUGGGUGUGUAAUGUAAGUAAAC                       |        | -53.1| 39    |
| Contig_18822|GGGAUCUUUAAGAACUAAGACG                                       |        | -43.9| 40    |
| Contig_17043|AGUGUCAUCUUCUGCUAGCAUAAA                                    |        | -73.4| 42    |
| Contig_16448|GAAGCCGUGAGGAGGAGAGU                                       |        | -63.6| 68    |
| Contig_17400|UCGUCUAGCGCUCGUCACUAUAG                                   |        | -71.3| 54    |
| Contig_15122|UGGAAGGGUUGAGGAGGAGUAG                                     |        | -63.7| 46    |
| Contig_9778|GUAAGGGAACCUAAGAUAUCAUUCCAC                                |        | -77.2| 47    |
| Contig_16528|GUAAGGGAUCUAGAUAUCAA                                         |        | -67.9| 35    |
| Contig_15303|CCACGAGCCAGUUAAAGUAAUCC                                     |        | -91.3| 48    |
| Contig_14086|AUCGCCGUGUUCACUCUGUUAU                                      |        | -63.6| 40    |
| Contig_20897|GGGCACCUCACUCAUGGCGAAG                                      |        | -69.8| 64    |
| Contig_13516|UGAGGGUCAACAGCAGAC                                       |        | -94.2| 59    |
| Contig_22213|UUAGGGUCCUUCAUUAAU                                         |        | -79.4| 24    |
| Contig_20063|UCGUGCGGAGGGAGAGCGUAG                                      |        | -74.9| 64    |
| Contig_21500|UUCCAGUCACAGAGUUGU                                          |        | -61.2| 43    |
| Contig_19382|AUUCCAGUAGUAUUUCCAGAGGCU                                    |        | -67.8| 46    |
Comparison of differentially expressed miRNAs between the small RNA libraries derived from KC3 under water stress and Suvin under water stress revealed 36 novel miRNAs. All the miRNAs (except nine miRNAs) were down regulated in drought susceptible Suvin under water stress. Among them, contig_18155 had shown highest down regulation (33.688 fold) and contig_4359 had highest up regulation (11.054 fold) in Suvin-WS (Table 4). Contig_16568 accumulated abundantly with more than 17 RPKM. Contig_1094, Contig_15570, and Contig_5490 were also detected in relatively high levels with more than 10 RPKM in these two libraries. Interestingly, the majority of the novel miRNAs was detected in more than 3.5 RPKM (Table 4).

| Contig ID       | Small RNA sequence identified in this study | Length | MFE * | GC(%) # |
|-----------------|---------------------------------------------|--------|-------|---------|
| Contig_15835    | GAAGAGAUAUCUUCUACUGU                       | 21     | -48.3 | 38      |
| Contig_1094     | GACACGAACCGUGUUGCUAGCAGCC                  | 30     | -100.8| 57      |
| Contig_8575     | GUCACACACGGUCUAGACACAGCC                   | 25     | -68.6 | 60      |
| Contig_19062    | UCACAGGGAUCAAAAUUGGA                      | 21     | -92.9 | 43      |
| Contig_15122    | UGGAAAGGUUGGGAGGAAUGA                      | 22     | -63.7 | 46      |
| Contig_9778     | GUAAGGGAACCUAGAUAUCUACCCAGC               | 30     | -77.2 | 47      |
| Contig_16528    | GUAAGGGAUCUAGACUAUAA                     | 23     | -67.9 | 35      |
| Contig_18252    | UAUAACAGGAAUGUGGAAUU                      | 21     | -69.0 | 29      |
| Contig_18054    | UAGGACCUACAGCUCUCCAGGAG                   | 23     | -82.5 | 48      |
| Contig_16564    | GACCCCUUGUUGGAGAGUCC                    | 21     | -67.4 | 52      |
| Contig_15875    | UUACAGUCCGAGAUCUCCCGUACCU                 | 25     | -84.2 | 52      |
| Contig_13633    | UUGUUGUUGUCGCGCGUGAGACGUA                | 27     | -77.7 | 56      |
| Contig_17693    | UUAUGGCUUUCUCAUAAAGU                     | 22     | -79.4 | 23      |
| Contig_11169    | UUGACACUAACGUCAACCCAGGA                  | 24     | -83.9 | 46      |
| Contig_21397    | UUGACUGCGCAGAAUCUCCCGUACCU               | 24     | -84.2 | 54      |
| Contig_18244    | CAUUCUCCAGGAUUUUCCAGAGG                 | 22     | -67.8 | 50      |
| Contig_21146    | AGUCCCUUCAAUUCUCAAC                    | 21     | -78.0 | 33      |
| Contig_18155    | AUUAAUAGUGUCUGCCUGCAACUU                  | 23     | -47.8 | 35      |
| Contig_15570    | UUGUUCGCCGCUGGAGACGCG                  | 22     | -76.4 | 64      |
| Contig_15717    | GGCUGUUGUUGUAUCCCGGCGAAGA               | 22     | -78.5 | 55      |
| Contig_20599    | UGAAUACACAAUAAUUCUAGUGCAUCACA            | 30     | -84.0 | 27      |
| Contig_21324    | UCCACAGGUAUCUUGGCAAGUGAU                | 22     | -75.4 | 36      |
| Contig_20427    | UGGCAUGGAGUUGUUGGCGCCACG               | 22     | -72.3 | 59      |
| Contig_21332    | AACAUUGGCAUAUAUAUGUGGAAAU               | 26     | -93.7 | 31      |
| Contig_18964    | UCACAGGGAUCAAAUUGGGGA                  | 21     | -92.9 | 43      |
| Contig_7495     | UCCAGAAAGAUAUCUUAGACCAUGGGAUAU         | 30     | -117.0| 40      |
| Contig_5490     | GGAAGUGGGGUUGGCAAGAAAGAUA            | 25     | -68.1 | 56      |
| Contig_17372    | AUCACCGGUGUACUUGUCUAAUC                | 25     | -63.6 | 40      |
| Contig_15837    | UUUCUAAAGCCUCUAGAGAUAAGUUA           | 24     | -71.9 | 29      |
| Contig_16924    | AGCGACGGAUAUCGAAUUCAUAUGUGUGUC       | 32     | -65.2 | 38      |
| Contig_12125    | UGAGGGAUCAGACUGACCCUC                | 22     | -94.4 | 59      |
| Contig_20315    | UCACGCGACCCUGCGAGCUGG               | 21     | -114.3| 71      |
| Contig_4359     | UGACACUGGAGAUCUGGAAGAAGAUGG         | 26     | -127.7| 46      |
| Contig_14297    | GUAAGGGAACCUAGAUAUC                      | 21     | -47.6 | 43      |
| Contig_16284    | UGAGGGAACCCACUGCCUGG                  | 22     | -66.8 | 55      |

Note: * MFE: minimum folding energy identified through mFold; GC content (%) found in Corresponding precursor sequence identified from G. raimondii genome.
Figure 3 Secondary hairpin structures predicted for putative novel miRNAs. The highlighted nucleotides are the sequence of putative novel miRNA found in this study and the remaining are the sequences of *G. raimondii*.

### 1.4 Targets of miRNAs identified in this study

It is known that miRNAs mediate plant abiotic stress responses through regulating their target genes and thus serving as key players in the gene regulation networks. The targets of miRNAs were predicted by the web tool psRNATarget (Dai and Zhao, 2011) using the *Gossypium* (cotton) DFCI Gene Index (CGI) Release 11 (http://compbio.dfci.harvard.edu/index.html) as the sequence library for the targets search. The predicted targets of some of the drought-responsive miRNAs, identified in this study, were found to be involved in diverse cellular processes including development, transcription, protein degradation, detoxification, nutrient status, and cross-adaptation to various abiotic stresses (data not shown). Besides, miRNAs reported here have also shown (elsewhere) to target several genes that are associated with abiotic stress resistance.

### 2 Discussion and Conclusion

A comprehensive understanding of miRNAs and their targets in cotton genome presents a necessary milestone in our understanding of gene regulation under water stress, which is an important abiotic stress that limits the cotton lint production and fiber quality, worldwide. This is especially true in some cases where knowledge of the miRNA
sequence alone is sufficient for the identification of the physiologically relevant target genes (e.g., hme-bantam by Brennecke et al., 2003).

This study has helped to identify several drought responsive candidate miRNAs that has potential in revealing the role of small RNA in drought tolerance in crop plants and cotton in particular. Among the several differentially expressed conserved miRNAs, those that were found between KC3-WS and Suvin-WS has more pronounced role in deciphering the molecular mechanism of drought resistance in cotton. The drought responsive miRNAs identified in this study were also found to be associated with drought or other abiotic stresses elsewhere and they were described hereunder.

The conserved miRNA, hme-bantam with 23 nt length has found to be highly down regulated (35.018 fold change in RPKM reads; Supplementary Table 1) in the drought susceptible parent, Suvin. Bantam microRNA in Heliconius melpomene was predicted independently by two groups using computational approaches (Brennecke et al., 2003; Lai et al., 2003). Northern blotting confirmed that hme-bantam with a 23 nt sequence was the most commonly expressed. Interestingly, bantam microRNA simultaneously stimulates cell proliferation and prevents apoptosis and targets the 3’ UTR of the pro-apoptotic gene hid in Drosophila. In addition to bantam miRNA, miR2 family have also been shown to regulate the expression of the proapoptotic genes such as reaper, grim and skl and to limit apoptosis in Drosophila (Zhang and Cohan, 2013). Two members of miR2 was identified in this study and both were highly down regulated (159.578 and 34.575 folds; Supplementary Table 1) in the drought susceptible, Suvin. Drosophila melanogaster presents perhaps the most important model organism for understanding the basic principles and molecular mechanisms of animal development. Similarly, Drosophila genetics has played an important role in understanding the functional roles of several animal and plant miRNAs.

### Table 4 Putative novel miRNAs identified in this study that are differentially expressed between KC3-WS and Suvin-WS

| S. No | Contig number | Mature sequence | Reads (RPKM)* | Fold Change Δ |
|-------|---------------|----------------|---------------|---------------|
| 1.    | Contig_4796   | GCCUCCAGAAGAUCAUUGCAGCAUGGUAUAU | 7.29372 5.46745 | 3.546 down |
| 2.    | Contig_8004   | UGCCAAACACGGAAGCGAAAG | 6.7624 5.60969 | 2.227 down |
| 3.    | Contig_16568  | UUUCCAUCAUAAUUAUCCGCAUG | 17.41518 14.09761 | 9.969 down |
| 4.    | Contig_14218  | AGAUGGCAGUUGGGUUGUGUAGAAUAGCUAC | 7.27723 4.866 | 5.319 down |
| 5.    | Contig_15122  | UGGAAGGGGUUGGAGGAGAUUGA | 6.61113 8.02857 | 2.666 up |
| 6.    | Contig_16528  | GUAGGAGAAGAUCAGAUAAUA | 13.88788 12.15502 | 3.323 down |
| 7.    | Contig_15303  | CCAACGACGAAGUAUUGUCC | 7.44543 3.37123 | 16.844 down |
| 8.    | Contig_14086  | AUCAACCGUGUACUCGUCUAAUC | 8.45815 8.85407 | 1.315 up |
| 9.    | Contig_20063  | UCUGUGCGAGGGGAGAUGGCUG | 5.51943 3.37123 | 4.432 down |
| 10.   | Contig_15835  | GAAGAGAUCACUCAUACUGU | 6.58521 6.10724 | 1.392 down |
| 11.   | Contig_1094   | GACACGAACACGUGUGUGCUACACC | 9.99151 11.22683 | 2.354 up |
| 12.   | Contig_8575   | GUCACACAGCGUCAGACACGC | 7.51461 7.53694 | 1.015 up |
| 13.   | Contig_16564  | GAACCCUUUGUGGAGAUGG | 8.33933 5.15551 | 9.087 down |
| 14.   | Contig_13633  | UUGUGUUGUUCGGCGUGUGAGCAGGUA | 7.42923 5.866 | 2.955 down |
| 15.   | Contig_17693  | UUAUAGGUCCUUCUAAUAAAGU | 9.24686 7.03592 | 4.629 down |
| 16.   | Contig_11693  | UAUGACCAACAGCUACUACGGGAGA | 8.47363 5.32543 | 8.865 down |
| 17.   | Contig_21397  | UUGCAGUCGAGAUCACCGUACC | 8.73666 7.54782 | 2.279 down |
| 18.   | Contig_18244  | CAUUCAGUAGUUAUCCAGAGG | 8.59916 6.67335 | 3.799 down |
| 19.   | Contig_21146  | AGUUCUUCAAAUUUCUUCAC | 10.00731 6.25504 | 13.475 down |
| 20.   | Contig_18155  | AUUAAGUGAGUGCUGGCAACACU | 8.87607 3.80187 | 33.688 down |
| 21.   | Contig_15570  | UUGUGUUGCCCGGUGGACGACG | 10.88209 12.21473 | 2.518 up |
| S. No | Contig number | Mature sequence                  | KC3-WS RPKM | Suvin-WS RPKM | Fold Change |
|-------|---------------|----------------------------------|-------------|---------------|-------------|
| 22.   | Contig_15717  | GGCUGUGGGUUGAUUCCGGCAAGA         | 8.49962     | 6.45096       | 4.137 down  |
| 23.   | Contig_20599  | UGAAACAAACAUUUCAUUGCUCAUACA      | 6.82977     | 4.41854       | 5.319 down  |
| 24.   | Contig_21324  | UCAAAUUACUUGGCAAGUGAU            | 7.86219     | 6.18793       | 3.191 down  |
| 25.   | Contig_20427  | UGGCAAGUGUGGCGCCACG              | 7.9402      | 5.45096       | 5.614 down  |
| 26.   | Contig_21332  | AACAUGGCGAAUAUGAUGGAAAU          | 6.62118     | 7.94692       | 2.506 up    |
| 27.   | Contig_18969  | UCACGGAUCAAAAUUGGGGA             | 8.84684     | 6.25504       | 6.028 down  |
| 28.   | Contig_7495   | UCCAGAAGAUACAUUGCAUUGGAAU         | 7.56674     | 5.74047       | 3.546 down  |
| 29.   | Contig_5490   | GGAGUGGGGUGCGAGGAAAGCUAAUC        | 11.78397    | 13.88125      | 2.479 up    |
| 30.   | Contig_17372  | AUCAACCGUGUACUCUGUCUAUCC         | 9.36582     | 8.07389       | 2.448 down  |
| 31.   | Contig_15837  | UUUCUAAAGCUCUAGAUAACGUUA         | 10.22495    | 8.49536       | 3.166 down  |
| 32.   | Contig_16924  | AGGCGUGUACGAGAAUCAUAUGUGUGUC     | 5.95905     | 3.32543       | 2.605 down  |
| 33.   | Contig_12125  | UGAGGGUCAAGAGUCUCCUC             | 9.0142      | 10.64736      | 3.101 up    |
| 34.   | Contig_20315  | UCAGCAGCGACUCCUGCGAGUG           | 7.66627     | 3.93311       | 13.298 down |
| 35.   | Contig_4359   | UGGACUUGGAUACGUGAAGAUGUG         | 8.09512     | 11.56163      | 11.054 up   |
| 36.   | Contig_14297  | GUAAGGGAGAGCUCCUAGAGAUA          | 9.34344     | 6.51808       | 7.092 down  |

Note: *- RPKM: reads per kilo base million; $^5$ - Differential expression in fold change (up or down regulation in Suvin-WS)

Continuously, high level expression of miR2 in root tissues were reported in Mangrove (Khraiwesh et al., 2013) and more specifically differential expression of miR2 under drought was noticed in maize (Xu et al., 2010). Another report in cucumber has also shown that miR2 targeted a gene coding a tetrameric peptide repeat-containing protein (Martinez et al., 2011) that usually interacts with heat shock proteins (Yin et al., 2011). Thus, it can be strongly hypothesised that vigorous up regulation of hme-bantam and miR2 in the drought tolerant KC3 accelerates anti-apoptotic activity in the cells during water stress and provide useful leads to promote research that confirms that these miRNAs endorse drought tolerance in cotton. Similarly several other conserved miRNAs identified in this study were found to be linked to water or other abiotic stresses and cellular processes.

For example, miR14 was 76.243 fold down regulated in Suvin under water stress in this study (Supplementary Table 1). The targets for miR14 include cell-death effectors other than hid (Xu et al., 2003) and mutants of miR14 were shown to be sensitive to high salinity (Leung and Sharp, 2010). In foxtail millet, miR14 has reported to target GRAS transcription factor that has been involved in development and other processes (Yi et al., 2013). This study has identified that there were nine members of small RNAs (maximum of 4.058 fold down regulated) under water stressed conditions in Suvin and all of them have shown > 90% percent identity with miR156 family. Squamosa Promoter Binding like (SPL) proteins are negatively regulated by the miR156, whose cellular levels are higher in younger plants and progressively decrease as the plant ages and especially during flowering (Wu et al., 2009; Yamaguchi et al., 2009). Among the miRNA-target that were studied in cotton, miR156-SPL2, shown significant regulation relationship in roots and leaves under salinity stress (Wang et al., 2013). Up regulation of miR156 in rice shoot and root tissues in response to heat (Sailaja et al., 2014) and drought (Zhou et al., 2010) stresses was also reported as it found in KC3-WS in this study. Further, high-throughput sequencing and expression analysis of microRNAs and their targets under salt stress in Caragana intermedia has shown that the miR156 family exhibited the highest abundance among all of the miRNA families (Zhu et al., 2013).

More than 8 fold up regulation of miR159 was noticed in KC3-WS when compared with Suvin-WS (Supplementary Table 1). Ding et al. (2013) reviewed that miR159 was induced by ABA and drought treatments and miR159 play a key role in ABA response by directing the degradation of MYB mRNAs. Bertolina et al. (2013) addressed the role of miR159 in reprogramming leaf growth during drought stress in Brachypodium. 
distachyon. Shuai et al. (2013) shown that the expression patterns and RT-qPCR results of drought-responsive miRNAs in *Populus* were consistent and both indicating that miR159 members were up-regulated after drought treatment. Similar kind of up regulation of miR159 under drought was also reported in legumes (Mantri et al., 2013), finger millet (Nageshbabu et al., 2013) and in cotton (Wang et al., 2013).

Eight members of up regulated miR162 with maximum elevated expression of 15.071 folds (Supplementary Table 1) was found in water stressed KC3 leaves in this study. Similar kind of up regulation of miR162 under drought was reported in cotton (Wang et al., 2013) and *Brachypodium* (Bertolinia et al., 2013). Up regulation of miR162 under salt stress was has also been reported in roots of *Vigna unguiculata* (Paul et al., 2011). In cotton, Ruan et al. (2009) have shown that miR162 target zinc finger protein and miR162 was also found to be key regulator of dicer-like protein (DCL) that are essential for miRNA or small RNA biogenesis (Liu et al., 2009). Khraiswesh et al. (2012) have shown that both up regulation of miR162 in salt-shocked maize roots and down regulation of miR162 under mechanical stress.

Totally 20 members of miR166 family was found that were both up (two members; with maximum fold change of 4.949) and down (18 members; with maximum fold change of 5.305) in Suvin under water stress (Supplementary Table 1). Both up regulation (in *Medicago truncatula* plants subjected to progressive water deficit; Trindade et al., 2010) and down regulation (in *Triticum dicoccoides* in response to shock drought stress; Kantar et al., 2010) of miR166 was reported earlier. It has shown that miR166 was chiefly expressed under abiotic stresses in cassava and has basic-leucine zipper (bZIP) transcription factor as target (Ballén-Taborda et al., 2013) which is shown to be involved in ABA and abiotic stress signalling in several plant species (Nicolas et al., 2014). Clear role for up and down regulation of miR166 was not yet deciphered; hence, additional studies on miR166 under drought stress should shed more information on precise molecular details on stress resistance mechanisms in plants.

Down regulation of 5 members of miR167 was noticed in Suvin-WS (Supplementary Table 1). Microarray analysis showed that miR167 was up regulated by drought stress in Arabidopsis and miR167 has been reported to play major roles in drought and ABA response in plants (Liu et al., 2008). Similar to more than 5 fold down regulation of miR169 in KC3-WS, down regulation of miR169 has also been noticed in drought resistant, soybean (Ni et al., 2013). It has experimentally validated that *GmNFYA3* (a gene encoding Nuclear factor Y (NF-Y), which is a heterotrimeric transcription factor), was target gene of miR169 and demonstrated that *GmNFYA3* is a positive regulator of plant tolerance to drought stress (Ni et al., 2013). In addition, Xu et al. (2014) evidently demonstrated that miR169 family members might be involved in flowering time regulation under abiotic stresses in plants including drought. Plants often flower earlier under water stress conditions and early flowering is considered as one of the key drought resistance component traits in plants (Komashita et al., 2008). The stress-activated transition to flowering enhances the chance of a plant population surviving under the threatening water stress.

All the four members of miR171 were up regulated in KC3-WS (Supplementary Table 1). Induced expression of miR171 was shown in barley under dehydration stress and the identified targets for miR171 was scarecrow like 6 (SCRL6). It encodes a transcription factor that is involved in diverse plant developmental processes such as leaf or root (Kantar et al., 2010).

Elevated up regulation of miR276 (57.625 fold; Supplementary Table 1) was found in this study in KC3-WS and to the best of our knowledge, there is no report in plant that specifies the role of miR276 under water stress. However, it has been reported in insects and shown to have key role in larval development and age dependent behavioural changes (Hori et al., 2010; Kozomara and Griffiths-Jones, 2011).

Yet another miRNA family that has high level of induction in KC3-WS were two members of miR279 (maximum of 58.068 fold changes; Supplementary Table 1). Up regulated expression of miR279 was mainly found in roots and it is expressed at lower levels in leaves, and scarcely expressed in stems and booting panicles in the common wild rice, *O. rufipogon* (Chen et al., 2013). However, their role in drought tolerance in plant is yet to be determined at molecular level.
In our previous study in cotton, we have computationally identified miR319 and established through *in silico* analysis that miR319 might be involved in abiotic stress resistance (Boopathi and Pathmanaban, 2012). Interestingly, miR319 was found to be induced 4.939 folds in KC3-WS (Supplementary Table 1). Transgenic plants overexpressing miR319 displayed morphological changes and exhibited enhanced drought and salt tolerance associated with increased leaf wax content and water retention but reduced sodium uptake (Zhou et al., 2013). As one of the first experimentally characterized and most conserved miRNA families, the miR319 targets Teosinte Branched / Cycloidea / Proliferating cell factors (TCP) genes, which encode plant-specific transcription factors sharing a conserved TCP domain with a basic helix-loop-helix structure. The TCP family is known to be largely involved in plant development such as the control of cell proliferation in leaf morphogenesis (Axtell and Bowman, 2008; Zhou et al., 2013).

It is known that miR393 was commonly up regulated during drought stress in Arabidopsis, rice and sugarcane (Ding et al., 2013). The target of miR393 encodes *TIR1* (*transport inhibitor response 1*), an auxin receptor in Arabidopsis. The TIR1 enzyme is a positive regulator of auxin signalling by promoting the degradation of Aux/IAA proteins through ubiquitination. Thus, increased levels of miR393 would down regulate auxin signalling and may reduce plant growth under drought stress (Ding et al., 2013). It was also shown in this study that miR393 was down regulated more than two folds in KC3-WS; Supplementary Table 1) which might favour the growth of KC3 under drought. Similarly, more than two fold down regulation of miR394 was noticed in KC3-WS. Downregulated miR394 was noticed in Populus, whose predicted targets are annotated as dehydration-responsive protein and F-box proteins, which were reported to play significant roles in the abiotic stress-response pathway (Shuai et al., 2013).

High induction of 36 members of miR396 (maximum of 11.456 folds) was displayed in KC3-WS in this study (Supplementary Table 1). miR396 family was found to have maximum number of members in this study and interestingly, all were up regulated in KC3-WS. Up regulation of miR396 under water stress was also noticed in Arabidopsis and tobacco (Liu et al., 2009; Yang and Yu, 2009). miR396 has been shown to target six growth regulating factor (GRF) transcription factors with roles in the coordination of cell division and differentiation during leaf development in Arabidopsis (Wang et al., 2011). Transgenic Arabidopsis plants overexpressing miR396 displayed narrow-leaf phenotypes because of reduction in cell number, achieved through repression of the expression of GRF genes and were more tolerant to drought than wild-type plants, likely because of the lower stomatal density (Liu et al., 2009). Overexpression of miR396 also conferred increased tolerance to drought stress in tobacco (Yang and Yu, 2009). These observations suggest that miR396 plays important roles not only in leaf development but also in drought tolerance in plants (Ding et al., 2013).

Three members of miR399 were up regulated (maximum of 7.150 fold) in KC3-WS (Supplementary Table 1). miR399 was found as one of the largest family in plants and it appears to target four sites in the 5′ UTR of a putative ubiquitin conjugating enzyme E2 (UBC; Sunkar and Zhu, 2004). It was found that there is no substantial regulation of miR399 by salt, drought, or cold stress. Low K or low N did not induce the expression of miR399. But miR399 was highly induced (whereas the target UBC mRNA was reduced) by low-phosphate (Pi) stress (Lu and Huang, 2008; Mantri et al., 2013).

Interestingly, 15 members of miR482 were up regulated (to the maximum tune of 15.957 folds) in KC3 under water limited conditions (however, one member of miR482 was down regulated (2.085 fold); Supplementary Table 1). Up regulation of miR482 under drought was reported in finger millet (Nageshbabu et al., 2013). Four UDP-glucosyltransferases (UDPGs) were found to be targeted by miR482 (Shuai et al., 2013). The UDPGs are enzymes that attach a sugar molecule to a specific acceptor in plants. As in Arabidopsis, the UDPG is a key regulator of stress adaption through auxin IBA (Tognetti et al., 2010) and plays a role in fine-tuning nitrogen assimilation in cassava (Kannangara et al., 2011). miR482 has also shown to target resistance gene receptor kinases (Nageshbabu et al., 2013; Shuai et al., 2013) and involved in disease resistance in cotton (Zhu et al., 2013) and soybean nodulation (Nageshbabu et al., 2013). The ability of drought stress to increase and decrease the
abundance of same members of miR482 highlights the need for more in-depth and detailed characterization of stress-responsive miRNAs in plants, particularly in cotton.

Up regulation of all the 10 members of miR535 family were reported in this study in KC3-WS (Supplementary Table 1). Similarly, miR535 was the most abundant miRNA in drought-exposed peach (Eldem et al., 2012). It has been suggested that stress-regulated miRNAs might be regulated by stress-related transcription factors. Interestingly, the promoters of miR535 in Arabidopsis contained all of the stress-related cis-elements investigated, including the drought-responsive element, low temperature-induced element, MYC-binding site, and MYB-binding site (Abdel-Ghany and Pilon, 2008).

The most abundant conserved miRNA family that has highly induced in drought tolerant cotton cultivar KC3-WS is miR750: there was 196.813 orders of magnitude of changes when compared with drought susceptible, Suvin. It has been shown that expression of miR750 was the most abundant miRNA after infecting silkworm with Bombyx mori cytoplasmic polyhedrosis virus and has six target genes including kinases (Wu et al., 2013). Nevertheless, the role of miR750 in plant abiotic stress resistance is yet to be determined. Similarly, high accumulation of several conserved miRNA families such as miR3027, miR3050, miR5176, miR7496, miR7499 and miR7500 (20.390, 15.514, 33.688, 6.660, 11.525 and 7.535 folds, respectively) were found in KC3-WS. However, their role in drought or abiotic stress resistance in crop plants have not yet been reported. Thus, this study has identified several drought responsive highly induced conserved miRNAs in cotton and they provided strong lead to specifically characterize such miRNA roles in drought tolerance in plants, particularly in cotton.

Two members of miR2118 were up regulated in KC3-WS with maximum of 7.806 folds (Supplementary Table 1). Similar kind of accumulation of miR2118 upon drought and ABA treatments in legumes were reported and showed that they may target regulatory elements for cellular processes (Arenas-Huertero et al., 2009). In our previous study in cotton, we have computationally identified miR2118 and established through in silico analysis that miR2118 might be involved in abiotic stress resistance (Boopathi and Pathmanaban, 2012). In addition, significant level of expression of miR2118 was reported in drought tolerant cowpea genotype (Barrera-Figueroa et al., 2011) and it was also up regulated in response to drought stress in Medicago truncatula (Wang et al., 2011). Further, it has been shown that miR2118 targets TIR-NBS-LRR domain protein, which usually expressed in response to multiple abiotic stresses such as drought, cold, salinity and ABA (Wang et al., 2011; Mantri et al., 2013).

More than six fold up regulation of two members of miR2949 and two fold induction of miR3476 were displayed in KC3-WS (Supplementary Table 1). Though the role of miR2949 and miR3476 in drought tolerance has not been reported in plants, high accumulation of miR2949 and miR3476 in cotton was predicted as positive regulators of the process of fiber initiation (Wang et al., 2012) and somatic embryogenesis (Yang et al., 2013).

The increased rates of miRNA detection afforded by deep-sequencing technologies provide challenges to the level of confidence required to annotate a sequence as a miRNA. A typical RNA deep-sequencing experiment will identify millions of short sequences. Increased coverage results in detection of sequences of ever-lower abundance. It therefore becomes more and more challenging to distinguish true miRNAs from fragments of other transcripts, other short RNAs and spurious transcription. Guidelines for microRNA annotation were established in 2003 (Ambros et al., 2003), requiring evidence of expression of a ~22 nt sequence (for example, cloning, sequencing or northern blot), together with evidence for a microRNA precursor structure (predicted stem–loop flanking the mature sequence). Updated annotation criteria were recently suggested to distinguish microRNAs from other classes of short RNAs in plants (Meyers et al., 2008). These standards have proved extremely powerful in maintaining a clean data set of miRNA sequences for the community. The eukaryotic genome also contains millions of predicted hairpins, so a flanking stem–loop structure should be considered necessary but not sufficient to annotate a sequence as a miRNA. If poorly analysed, a single data set thus has the potential to generate a large number of suspicious annotations, thus overwhelming the real miRNAs. However, correct interpretation of RNA
deep-sequencing data provides several additional signals to help distinguish miRNAs from other sequences (Kozomara and Griffiths-Jones, 2011).

The lowest minimum free energy (MFE) among all hairpin structures of the novel miRNAs precursors was −43.9 kcal/mol (Table 4), which is significantly lower than the threshold of −30 kcal/mol reported in a previous study [25]. Among the 36 novel miRNAs identified in this report, none of the novel miRNA had homologs in miRBase Release 19. Hence, we assumed that they were Gossypium-specific, or restricted to closely related species. Further, there is a chance to get more novel or conserved miRNAs if new genomic resources are developed in near future as discussed above.

In conclusion, our previous computation study (Boopathi and Pathmanaban, 2012) has shown that conserved miRNAs were shown to be linked with abiotic stress resistance. Here we provided experimental proof for those and other miRNAs for their roles in drought tolerance in cotton. During water stress responsive transcript profiling, it is important to target the different tissues and their precise stage besides the dynamics of the stress (such as time, duration and intensity), for isolation of RNA. Thus, it will be more useful if the ultimate economically important tissues, the flower, boll and fiber, are used during sampling instead of leaves that was used in this study. Further, instead of using genotypes with different genetic backgrounds, near-isogenic lines, which differ only in the target trait, would be ideal genetic material which ensures that differentially expressed genes are linked to the drought resistance and not to the genetic background. Our laboratory has already started to work on these lines. The number of new miRNAs identified in this study is relatively low when compared with other crops. This is mainly due to lack of genome sequence in G. hirsutum and G. barbadense. Release of their genome sequences in the near future may help to identify more number of novel miRNAs in cotton in response to drought stress using this data.

3 Materials and Methods

3.1 Plant materials

Two contrasting accessions of Gossypium, which are considered as ideal genetic materials for studying drought tolerance mechanism in cotton, were used in this study. One was short duration upland cotton, G. hirsutum L. var. KC 3 (hereafter referred as KC3) that has good performance under rainfed tract of Tamil Nadu, India and provide assured minimum yield with relatively good quality fiber even under severe water stress. Another line was the long duration G. barbadense L. var. Suvin (hereafter referred as Suvin) which provide poor yield with inferior fiber qualities under drought stress (Boopathi et al., 2011). These two lines have also shown significant dissimilarity at molecular level (Thiagu et al., 2011) and hence they were selected for development of different types of mapping population at this laboratory. Therefore, the differentially expressed miRNAs identified in this study could also be genetically mapped, since there is no genome sequence currently available in G. hirsutum or G. Barbadense.

3.2 Experimental conditions in the field

Field trial was conducted under upland condition at TNAU, Coimbatore, India during the regular cotton cultivating season, kharif (July - December, 2012). Both the cotton lines were grown in a single row by following 75 x 40 cm spacing between and within rows, respectively with two replications and evaluated under two water regimes: irrigated (control or non-stress) and water limited (stress). A buffer area of 3.0 m wide and 0.75 m deep along the length of the experimental plot divided the control and stress plots. All the standard agronomic practices were equally applied to both the treatments. Both the lines were surface irrigated as per the regular cultivation practices, except for water limited regime. Water stress was imposed by withholding irrigation when the plants were in peak vegetative phase to stress plots i.e., from 30 days after sowing (DAS) and there were continuous rain free period of 20 days. The leaf samples were collected from both the accessions and treatments on 50th DAS with two replications and stored in Shrimpex lysis buffer (Shrimpex Biotech Services Private Limited, Chennai, India).

3.3 Small RNA library construction and sequencing
Small RNA library construction and sequencing were done at Shrimpex Biotech Services Private Limited, Chennai, India. Briefly, the miRNA was isolated according to the manufacturer’s protocol and concentration was measured using Nanodrop Spectrophotometer. To avoid DNA contamination, the miRNA was treated with DNase @ 0.5 unit per μg of miRNA (the reaction mixture contains, 0.5 unit of DNase, 10X DNase buffer and the miRNA template to a final volume of 30 μL) and incubated at 37°C for 30 minutes. The reaction was stopped by incubating at 65°C for 10 minutes with 1 μL of 50 mM EDTA. The miRNA was purified from DNase using Shrimpex spin column concentration module according to the manufacturer’s protocol. The quality of the obtained miRNA before and after DNase treatment was analyzed using Agilent 2100 small RNA kit.

The miRNA library was prepared using Ion Total-RNA Seq kit v2 and the miRNAs were enriched using magnetic bead clean up module for small RNA enrichment. In the next step, Ion RNA-Seq adapter was ligated to the enriched miRNA to facilitate reverse transcription of first strand cDNA synthesis and further purification using magnetic bead clean up module. The purified cDNA were amplified using a barcoded forward primer and a common reverse primer and quantified using Agilent Bioanalyzer 2100 DNA 1000 kit. Smear analysis was performed to determine the size distribution range of the four samples to confirm that the size of the four libraries was in the range of 94–114 bp.

The molar concentration of each library was determined and diluted to the same molar concentration. Equal volumes of each barcoded library of the replicated samples were mixed to prepare a pool of barcoded libraries and were amplified using emulsion PCR (ePCR). In ePCR, the clonal bead populations are generated in water-in-oil microreactors. After ePCR, the templates were denatured and bead enrichment step was performed to separate the beads with templates from non-template beads. The enriched beads were then loaded onto the Ion P1 Chip and sequenced.

3.4 Sequence analysis for miRNA identification

Sequences were trimmed to eliminate low-quality reads and other contaminants from the adaptor tags. All low quality reads were removed such as reads with 3’ and 5’ adapter contaminations, those without insert tags, and those with poly-A sequences. The remaining high-quality sequences were trimmed of their adapter sequences. All high-quality sequences were considered significant and number of reads in each nucleotide class was classified. Each read were annotated through BLAST searching for identification of rRNA, scRNA, snoRNA, snRNA and tRNA using ‘Rfam’ (ftp://ftp.sanger.ac.uk/pub/databases/Rfam/) and those matched reads were avoided for further analysis.

Initially, a simple blast search of resulted small RNA sequences identified in both the replications (in the size range of 20 - 33 nt) with miRBase was done to categorize their family classification based on conservation (version 19 (August, 2012), http://microrna.sanger.ac.uk). Sequences that were identical or related to known miRNA sequences (with maximum of single nucleotide substitution) were considered as known or conserved miRNAs and others were considered as putative novel miRNAs.

3.5 Identification of novel miRNAs

Considering the complicated structure of miRNA gene and its biogenesis, a series of strict filters were used during the new miRNA identification process to enhance reliability of the result presented here. Currently, no allopolyploid cotton genome sequence was available; therefore ESTs from NCBI and the genome sequence of G. raimondii (Wang et al., 2012), the closest living relative of the progenitor D-genome donor of allotetraploid cottons (G. hirsutum and G. barbadense) were used as reference sequences to predict the miRNA precursors. The 20-33 nt reads that were assumed as putative novel miRNA were mapped to cotton nucleotide sequences available at NCBI and D genome sequence (http://cgp.genomics.org.cn/page/species/index.jsp). Approximately 100 nucleotides flanking both side of the query sequence were obtained and the characteristic secondary hairpin structure of each miRNA precursors was predicted using the miRNA prediction software, mFold 3.2 (Zuker, 2003). During this process the following characteristics were considered to affirm novel cotton miRNAs: i) formation of flawless secondary structure (including formation of a perfect stem-loop structure, the read sequence
of this study sat at one arm of the stem and asymmetric bulges was minimal in size (one base at most) and infrequent (typically one or less) within the miRNA/miRNA* duplex) and ii) the minimum free energy should be lower than the threshold of ~30 kcal/mol (Bonnet et al., 2004).

3.6 Differentially expressed miRNAs
Differential expression analysis of both conserved and novel miRNAs between the two contrasting genotypes under two different water regimes were identified as described below. The miRNA reads of the each libraries were assembled de novo into contigs and normalized to reads per kilobase per million (RPKM: Mortazavi et al., 2008) and RPKM was used as a statistical value for the analysis of differences in expression among the samples. The increased or decreased level of expression (up- or down- regulation) of miRNA in each sample under given treatment were represented in RPKM fold change (as described in Rada-Iglesias et al., 2011) and minimum of two fold change in RPKM were considered further. Such analysis was helped to identify major (thus insignificant miRNAs were not included in this report) and differentially expressed miRNAs in leaf samples between i) KC3 irrigated control (hereafter mentioned as KC3-C) and KC3 water stressed (hereafter mentioned as KC3-WS) ii) KC3-WS and Suvin water stressed (hereafter mentioned as Suvin-WS) iii) Suvin irrigated control (hereafter mentioned as Suvin-C) and Suvin-WS and iv) KC3-C and Suvin-C.

3.7 Identification of miRNA’s targets
The targets of miRNAs was predicted by the web tool psRNATarget (Dai and Zhao, 2011) using the Gossypium (cotton) DFCI Gene Index (CGI) Release 11 (http://compbio.dfci.harvard.edu/index.html) as the sequence library for the targets search. No more than three mismatches between miRNAs and targets were allowed, especially no mismatches were allowed within the maximum expectation region.

Acknowledgements
The authors gratefully acknowledge DBT, Government of India for the financial assistance.

References
Abdel-Ghany S.E., and Pilon M., 2008, MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in Arabidopsis, Journal of Biology and Chemistry, 283(23): 15932-15945
http://dx.doi.org/10.1074/jbc.M801406200

Abdurakhmonov I.Y., Devor E.J., Buriev Z.T., Huang L., Makamov A., Shermatov S.E., Bozorov T., Kushman F.N., Mavljonov G.T., and Abdukarimov A., 2008, Small RNA regulation of ovule development in the cotton plant, G. hirsutum L, BMC Plant Biology, 8(1): 1-12
http://dx.doi.org/10.1186/1727-4295-8-93

Ambros V., Bartel D.P., Bartel D.P., Burge C.B., Carrington J.C., Chen X., Dreyfuss G., Eddy S.R., Griffiths-Jones S., Marshall M., Matzke M., Ravkun G., and Tuschi T., 2003, A uniform system for microRNA annotation, RNA, 9(3): 277-279
http://dx.doi.org/10.1016/S0961-8524(02)00309-3

Ambros V., 2004, The functions of animal microRNAs, Nature, 431(7006):350-355
http://dx.doi.org/10.1038/nature02871

Arenas-Huertero C., Perez B., Rabanal F., Blanco-Melo D., De la Rosa C., Estrada-Navarrete G., Sanchez F., Covarrubias A.A., and Reyes J.L., 2009, Conserved and novel miRNAs in the legume Phaseolus vulgaris in response to stress, Plant Molecular Biology, 70(4): 385-401
http://dx.doi.org/10.1007/s11103-009-9480-3

Axtell M.J., and Bowman J.L., 2008, Evolution of plant microRNAs and their targets, Trends Plant Science, 13(7): 343-349
http://dx.doi.org/10.1016/j.tplants.2008.03.009

Ballén-Taborda C., Plata G., Ayling S., Rodriguez-Zafrana F., Becerra Lopez-Lavalle L., Duimana J., and Tohme J., 2013, Identification of Cassava MicroRNAs under Abiotic Stress, International Journal of Genomics, 22(2013): 103
http://dx.doi.org/10.1155/2013/857986

Baroza M.Y.K., Irfan M., Yousaf R., Ali I., Qaisar U., Maqbool A.L., Zaheer M., Rashid B., Hussain T., and Riazuddin S., 2007, Identification of micro-RNAs in cotton, Plant Physiology Biochemistry, 46(8-9): 739-751
http://dx.doi.org/10.1016/j.plaphy.2008.05.009

Barrera-Figueroa B.E., Gao L., Diop N.N., Wu Z., Ehlers J.D., Roberts P.A., Close T.J., Zhu J.K., and Liu R., 2011, Identification and comparative analysis of drought-associated microRNAs in two cowpea genotypes, BMC Plant Biology, 11: 127
http://dx.doi.org/10.1186/1728-4202-11-127

Bertolini E., Verelst W., Horner D.S., Gianfrancesca L., Piccolo V., Inzé D., Pe M.E., and Mica E., 2013, Addressing the Role of microRNAs in Reprogramming Leaf Growth during Drought Stress in Brachypodium distachyon, Molecular Plant, 6(2): 423-443
http://dx.doi.org/10.1093/mp/sss160
Bonnet E., Wuyts J., Rouze P., and Van de Peer Y., 2004, Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences, Bioinformatics, 20(17): 2911-2917

http://dx.doi.org/10.1093/bioinformatics/bth374

Boopathi N.M., Thayagu K., Urbi B., et al., 2011, Marker-Assisted Breeding as Next-Generation Strategy for Genetic Improvement of Productivity and Quality: Can It Be Realized in Cotton?, International Journal of Plant Genomics, 2011(5550): 670104-670104

http://dx.doi.org/10.1155/2011/670104

Boopathi N.M., and Pathanamthan R., 2012, Additional insights into the adaptation of cotton plants under abiotic stresses by in silico analysis of conserved miRNAs in cotton expressed sequence tag database (dbEST), African Journal Biotechnology, 11: 14054-14063

Brennecke J., Hjörner D.R., Stark A., Russell R.B., and Cohen S.M., 2003, Bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene bid in Drosophila, Cell, 113(1): 25-36

http://dx.doi.org/10.1016/S0092-8674(03)00231-9

Chen X.M., Gao W.H., Zhang J.F., Zhang X.L., and Lin Z.X., 2013, Linkage mapping and expression analysis of miRNAs and their target genes during fiber development in cotton, BMC Genomics, 14(14): 234-239

http://dx.doi.org/10.1186/1471-2164-14-706

Chen X.Z., Li F.L., Yang S.N., Dong Y.B., Yuan Q.H., Wang F., Li W.M., Jiang Y., Jia S.R., and Pei X.W., 2013, Identification and expression profile of two UDP-glucosyltransferases, UGT85K4 and UGT85K5, catalyzing the last step in cyanogenic glucoside biosynthesis in cassava, Plant Journal, 68(2): 287-301

http://dx.doi.org/10.1007/s00126-012-1645-7

http://dx.doi.org/10.1007/s00126-012-1645-7

Dai X.B., and Zhao P.X., 2011, psRNATarget: a plant small RNA target analysis server

http://dx.doi.org/10.1016/j.nbt.2011.05.001

Ding Y., Tao Y., and Zhu C., 2013, Emerging roles of microRNAs in the mediation of drought stress response in plants, Journal of Experimental Botany, 64(11): 3077-3086

http://dx.doi.org/10.1093/jxb/ert116

Eldem V., Akçay U.C., Ozsuner E., Bakır Y., Uranbey S., and Unver T., 2012, Genome-wide identification of miRNAs responsive to drought in peach (Prunus persica) by high-throughput deep sequencing, PLoS One, 7(12): e50298

http://dx.doi.org/10.1371/journal.pone.0050298

Hori S., Kaneko K., Saito T.H., Takeuchi H., and Kubo T., 2011, Expression of two microRNAs, sme-mir276 and -1000, in the adult honeybee (Apis mellifera) brain, Apidologie, 42(1): 89-102

Kamoshita A., Baba R.C., Boopathi N.M., and Fukai S., 2008, Phenotypic and genotypic analysis of drought-resistance traits for development of rice cultivars adapted to rainfed environments, Field Crops Research, 109(1-3): 1-23

http://dx.doi.org/10.1016/j.fcr.2008.06.010

Kammangara R., Motawia M.S., Hansen N.K.K., Paquette S.M., Olsen C.E., and Jorgensen K., 2011, Characterization and expression profile of two UDP-glucosyltransferases, UGT85K4 and UGT85K5, catalyzing the last step in cyanogenic glucoside biosynthesis in cassava, Plant Journal, 68(2): 287-301

http://dx.doi.org/10.1111/j.1365-313X.2011.04695.x

Kantar M., Unver T., and Budak H., 2010, Regulation of barley miRNAs upon dehydration stress correlated with target gene expression, Functional & Integrative Genomics, 10: 493-507

http://dx.doi.org/10.1007/s10142-010-0181-4

Khraiwesh B., Pugalenthi G., and Fedorov N.V., 2013, Identification and analysis of red sea mangrove (Avicennia marina) microRNAs by high-throughput sequencing and their association with stress responses, PLoS One, 8(8): 1221-1221

http://dx.doi.org/10.1371/journal.pone.0060774

Khraiwesh B., Zhu J.K., and Zhu J., 2012, Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants, Biochimica et Biophysica Acta Gene Regulatory Mechanisms, 1819(2): 137-148

http://dx.doi.org/10.1016/j.bbagrm.2011.05.001

Kozomara A., and Griffiths-Jones S., 2011, miRBase: integrating microRNA annotation and deep-sequecing data, Nucleic Acids Research, 39: 152-157

http://dx.doi.org/10.1093/nar/gkr1027

Kwak P.B., Wang Q.Q., Chen X.S., Qiu C.X., and Yang Z.M., 2009, Enrichment of a set of microRNAs during the cotton fiber development, BMC Genomics, 10: 457

http://dx.doi.org/10.1186/1471-2164-10-457

Lai E.C., Tomanacak P., Williams R.W., and Rubin G.M., 2003, Computational identification of Drosophila microRNA genes, Genome Biology, 4(7): 54-56

http://dx.doi.org/10.1186/gb-2003-4-7-r42

Leung A.K.L., and Sharp P.A., 2010, MicroRNA Functions in Stress Responses, Molecular Cell, 40(2): 205–215

http://dx.doi.org/10.1016/j.molcel.2010.09.027

Li F.G., Fan G.Y., Wang K.B., Sun F.M., Yuan Y.L., Song G.L., Li Q., Ma Z.Y., Lu C.R., Zou C.S., Chen W.B., Liang X.M., Sang H.H., Liu W.Q., Shi C.C., Xiao G.H., Gou C.Y., Ye W.W., Xu X., Zhang X.Y., Wei H.L., Li Z.F., Zhang G.Y., Wang J.Y., Liu K., Kohel R.J., Percy R.G., Yu J.Z., Zhu Y.X., Wang J., and Yu S.X., 2014, Genome sequence of the cultivated cotton Gossypium arboreum, Nature Genetics, 46(6): 567-574

http://dx.doi.org/10.1038/ng.2987

Liu H.H., Tian X., Li Y.J., Wu C.A., and Zheng C.C., 2008, Microarray-based analysis of stress-regulated microRNAs in Arabidopsis thaliana, RNA, 14(5): 836-843

http://dx.doi.org/10.1371/journal.pone.0060774

http://dx.doi.org/10.1515/jbb.2011.05.001

Kozomara A., and Griffiths-Jones S., 2011, miRBase: integrating microRNA annotation and deep-sequecing data, Nucleic Acids Research, 39: 152-157

http://dx.doi.org/10.1093/nar/gkr1027

Kwak P.B., Wang Q.Q., Chen X.S., Qiu C.X., and Yang Z.M., 2009, Enrichment of a set of microRNAs during the cotton fiber development, BMC Genomics, 10: 457

http://dx.doi.org/10.1186/1471-2164-10-457

Lai E.C., Tomanacak P., Williams R.W., and Rubin G.M., 2003, Computational identification of Drosophila microRNA genes, Genome Biology, 4(7): 54-56

http://dx.doi.org/10.1186/gb-2003-4-7-r42

Leung A.K.L., and Sharp P.A., 2010, MicroRNA Functions in Stress Responses, Molecular Cell, 40(2): 205–215

http://dx.doi.org/10.1016/j.molcel.2010.09.027

Li F.G., Fan G.Y., Wang K.B., Sun F.M., Yuan Y.L., Song G.L., Li Q., Ma Z.Y., Lu C.R., Zou C.S., Chen W.B., Liang X.M., Sang H.H., Liu W.Q., Shi C.C., Xiao G.H., Gou C.Y., Ye W.W., Xu X., Zhang X.Y., Wei H.L., Li Z.F., Zhang G.Y., Wang J.Y., Liu K., Kohel R.J., Percy R.G., Yu J.Z., Zhu Y.X., Wang J., and Yu S.X., 2014, Genome sequence of the cultivated cotton Gossypium arboreum, Nature Genetics, 46(6): 567-574

http://dx.doi.org/10.1038/ng.2987

Liu H.H., Tian X., Li Y.J., Wu C.A., and Zheng C.C., 2008, Microarray-based analysis of stress-regulated microRNAs in Arabidopsis thaliana, RNA, 14(5): 836-843

http://dx.doi.org/10.1371/journal.pone.0060774

http://dx.doi.org/10.1515/jbb.2011.05.001
Lu Q., Feng Y., and Zhu Z., 2009, Dicer-like (DCL) proteins in plants, Functional & Integrative Genomics, 9(3): 277-286
http://dx.doi.org/10.1007/s10142-009-0111-5

Lu X.Y., and Huang X.L., 2008, Plant miRNAs and abiotic stress responses, Biochemical Biophysical Research Communication, 368: 458-462
http://dx.doi.org/10.1016/j.bbrc.2008.02.007

Mani N., Basker N., Ford R., Pang E., and Pardeshi V., 2013, The role of micro-homologous acids in legumes with a focus on abiotic stress response, The Plant Genome, 6(3): 841-856
http://dx.doi.org/10.3835/plantgenome2013.05.0013

Martínez G., Fortmenn J., Llave C., Pallás V., and Gómez G., 2011, High-Throughput Sequencing, Characterization and Detection of New and Conserved Cucumber miRNAs, PLoS One, 6(5): e19523
http://dx.doi.org/10.1371/journal.pone.0019523

Meyers B.C., Austell M.J., Bartel D.P., Baulcombe D., Bowman J.L., Cao X., Carrington J.C., Chen X., and Green P.J., 2008, Criteria for annotation of plant microRNAs, Plant Cell, 20(12): 3186-3190
http://dx.doi.org/10.1105/tpc.108.064311

Mortazavi A., Williams B.A., McCue K., Schaeffer L., and Wold B., 2008, Mapping and quantifying mammalian transcriptomes by rna-seq. Nature Methods, 5: 621-628
http://dx.doi.org/10.1038/nmeth.1226

Nageshbabu R., Jyothi M.N., Sharadamma N., Rai D.V., and Devaraj V.R., 2013, Expression of miRNAs confers enhanced tolerance to drought and salt stress in Finger millet (Eleusine coracana). Journal of Stress Physiology and Biochemistry, 9(3)

Ni Z., Hu Z., Jiang Q., and Zhang H., 2013, GmNFYA3, a target gene of miR169, is a positive regulator of plant tolerance to drought stress, Plant Molecular Biology, 82(1-2): 113-129
http://dx.doi.org/10.1007/s11103-013-0404-5

Nicolas P., Lecourieux D., Kappel C., Cluzet S., Cramer G., Delrot S., and Lecourieux F., 2014, The Basic Leucine Zipper Transcription Factor Abscisic Acid Response Element-Binding Factor 2: Is an Important Transcriptional Regulator of Abscisic Acid-Dependent Grape Berry Ripening Processes, Plant Physiology, 164(1): 365-83
http://dx.doi.org/10.1104/pp.113.231977

Pang M., Woodward A.W., Agarwal V., Guan X., Hu M., Ramachandran V., Chen X., Tripplett B.A., Stelly D.M., and Chen Z.J., 2009, Genome-wide analysis reveals rapid and dynamic changes in miRNA and siRNA sequence and expression during ovule and fiber development in allotetraploid cotton (Gossypium hirsutum L.). Genome Biology, 10(11): 100-105
http://dx.doi.org/10.1186/gb-2009-10-11-r122

Paul S., Kundu A., and Pal A., 2011, Identification and validation of conserved microRNAs along with their differential expression in roots of Vigna unguiculata grown under salt stress, Plant Cell, Tissue and Organ Culture, 105: 233-242
http://dx.doi.org/10.1007/s12229-010-9857-7

Qu C.X., Xie F.L., Zhu Y.Y., Guo K., Huang S.Q., Nie L., and Yang Z.M., 2007, Computational identification of microRNAs and their targets in Gossypium hirsutum expressed sequence tags, Gene, 395: 49-61
http://dx.doi.org/10.1016/j.gene.2007.01.034

Rada-Islasias A., Bajpai R., Swigut T., Brugmann S.A., Flynn R.A., Wysocka J., 2011, A unique chromatin signature uncovers early developmental enhancers in humans, Nature, 470: 279-283
http://dx.doi.org/10.1038/nature09692

Ruan M.B., Zhao Y.T., Meng Z.H., Wang X.J., and Yang W.C., 2009, Conserved miRNA analysis in Gossypium hirsutum through small RNA sequencing, Genomics, 94: 263-268
http://dx.doi.org/10.1016/j.ygeno.2009.07.002

Sailaja et al., 2014, Prediction and Expression Analysis of miRNAs Associated with Heat Stress in Oryza sativa, Rice Science, 21(1): 3-12
http://dx.doi.org/10.1016/S1672-6308(13)60164-X

Shah A., Zhang Z., Yin W., and Xia X., 2013, Identification of drought-responsive and novel Populus trichocarpa microRNAs by high-throughput sequencing and their targets using degradome analysis, BMC Genomics, 14(3): 433-440
http://dx.doi.org/10.1186/1471-2164-14-233

Sunkar R., and Zhu J.K., 2004, Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis, The Plant Cell Online, 16: 2001-2019
http://dx.doi.org/10.1105/tpc.104.022830

Sunkar R., Chimmsamy V., Zhu J., and Zhu J.K., 2008, Small RNAs as big players in plant abiotic stress responses and nutrient deprivation, Trends Plant Science, 12: 355-364
http://dx.doi.org/10.1016/j.tplants.2013.5.002

Tang G., Reinhardt B.J., Bartel D.P., and Zamore P.D., 2003, A biochemical framework for RNA silencing in plants, Genes Development, 17: 49-63
http://dx.doi.org/10.1101/gad.1048103

Thiyagu K., Boopathi N.M., Nadarajan N., Gopikrishnan A., Selvakumar P., Santoshkumar M., and Ravikesavan R., 2011, Sampling and exploitation of genetic variation exist in locally adapted accessions using phenotypic and molecular markers for genetic improvement of cotton, Gene Conserve, 10: 129-153
http://dx.doi.org/10.1016/j.gene.2007.01.034

http://dx.doi.org/10.1101/gad.1048103
Tognetti V.B., et al., 2010, Perturbation of indole-3-butyric acid homeostasis by the UDP-Glucosyltransferase UGT74E2 modulates Arabidopsis architecture and water stress tolerance, Plant Cell, 22: 2660-2679
http://dx.doi.org/10.1105/tpc.109.071316

Trindade I., Capitão C., Dalmany T., Fevereiro M.P., and dos Santos D.M., 2010, miR398 and miR408 are up-regulated in response to water deficit in Medicago truncatula, Planta, 231: 705-716
http://dx.doi.org/10.1007/s00425-009-1078-0

Wang K.B., Wang Z.W., Li F.G., Ye W.W., Wang J.Y., Song G.L., Yue Z., Cong L., Shang H.H., Zhu S.L., Zou C.S., Li Q., Yuan Y.L., Lu C.R., Wei H.L., Gou C.Y., Zheng Z.Q., Yin Y., Zhang X.Y., Liu K., Wang B., Song C., Shi N., Kohel R.J., Percy R.G., Yu J.Z., Zhu Y.X., Wang J., and Yu S.X., 2012, The draft genome of a diploid cotton Gossypium raimondii, Nature Genetic, 44(10): 1098-1103
http://dx.doi.org/10.1038/ng.2371

Wang M., Wang Q., and Zhang B., 2013, Response of miRNAs and their targets to salt and drought stresses in cotton (Gossypium hirsutum L.), Gene, 530: 26-32
http://dx.doi.org/10.1016/j.gene.2013.08.009

Wang T., Chen L., Zhao M., Tian Q., and Zhang W.H., 2011, Identification of drought-responsive microRNAs in Medicago truncatula by genome-wide high throughput sequencing, BMC Genomics, 12: 367
http://dx.doi.org/10.1186/1471-2164-12-367

Wang Z.M., Xue W., Dong C.J., Jin L.G., Bian S.M., Wang C., Wu X.Y., Liu J.Y., 2012, A comparative miRNAome analysis reveals seven fiber initiation-related and 36 novel miRNAs in developing cotton ovules, Molecular Plant, 5: 889-900
http://dx.doi.org/10.1093/mp/sor094

Wei M., Wei H., Wu M., Song M., Zhang J., Yu J., Fan S., and Yu S., 2013, Comparative expression profiling of miRNA during anther development in genetic male sterile and wild type cotton, BMC Plant Biology, 13(1): 1275-1296
http://dx.doi.org/10.1186/1471-2229-13-66

Wu G., Park M.Y., Conway S.R., Wang J.W., Weigel D., and Poethig R.S., 2009, The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis, Cell, 138: 750-759
http://dx.doi.org/10.1016/j.cell.2009.06.031

Xu C., Xue W., Dong C.J., Jin L.G., Bian S.M., Wang C., Wu X.Y., Liu J.Y., 2012, Comparative expression profiling of miRNA during anther development in genetic male sterile and wild type cotton, BMC Plant Biology, 13(1): 1275-1296
http://dx.doi.org/10.1186/1471-2229-13-66

Wu P., et al., 2013, Involvement of MicroRNAs in Infection of Silkworm with Bombyx mori Cytoplasmic Polyhedrosis Virus (BmCPV), PLoS One, 8: e68209
http://dx.doi.org/10.1371/journal.pone.0068209

Xu C., Yang R.F., Li W.C., and Fu F.L., 2010, Identification of 21 microRNAs in maize and their differential expression under drought stress, African Journal of Biotechnology, 9(30): 4741-4753

Xu et al., 2014, Stress-induced early flowering is mediated by miR169 in Arabidopsis thaliana, Journal of Experimental Botany, 65: 89-101
http://dx.doi.org/10.1038/jxbert335

Xu P., Vernooy S.Y., Guo M., and Hay B.A., 2003, The Drosophila microRNA mir-14 suppresses cell death and is required for normal fat metabolism, Current Biology, 13: 790-795
http://dx.doi.org/10.1016/S0960-9822(03)00250-1

Xue W., Wang Z., Du M., Liu Y., and Liu J.Y., 2013, Genome-wide analysis of small RNAs reveals eight fiber elongation-related and 257 novel microRNAs in elongating cotton fiber cells, BMC Genomics, 14(1): 1-14
http://dx.doi.org/10.1186/1471-2164-14-629

Yamaguchi A., Wu M.F., Yang L., Wu G., Poethig R.S., and Wagner D., 2009, The microRNA-regulated SBP-Box transcription factor SPL3 is a direct upstream activator of LEAFY, FRUITFULL, and APETALA1, Developmental Cell, 17: 268-278
http://dx.doi.org/10.1016/j.devcel.2009.06.007

Yang F., and Yu D., 2009, Overexpression of Arabidopsis miR396 enhances drought tolerance in transgenic tobacco plants, Acta Botanica Yunnanica, 31: 421-426
http://dx.doi.org/10.1016/j.ajxb.2009.09.044

Yang X., Wang L., Yuan D., Lindsey K., and Zhang X., 2013, Small RNA and degradome sequence reveal complex mRNA regulation during cotton somatic embryogenesis, Journal Experimental Botany, 64: 1521-1536
http://dx.doi.org/10.1093/jxb/ert1013

Yi F., Xue S., Liu Y., Qi X., and Yu J., 2013, Genome-wide characterization of microRNA in foxtail millet (Setaria italica), BMC Plant Biology, 13(1): 1-15
http://dx.doi.org/10.1186/1471-2229-13-212

Yin Z., Li Y., Yu J., Liu Y., Li C., Han X., and Shen F., 2011, Difference in miRNA expression profiles between two cotton cultivars with distinct salt sensitivity, Molecular Biology Reporter, 39(4): 4961-4970
http://dx.doi.org/10.1007/s11342-011-1392-2

Yin Z., Li Y., Han X., and Shen F., 2012, Genome-wide profiling of miRNAs and other small non-coding RNAs in the Verticillium dahliae inoculated cotton roots, PLoS One, 7: e35765
http://dx.doi.org/10.1371/journal.pone.0035765

Zhang B., Wang Q., Wang K., Pan X., Liu F., Guo T., Cobb G.P., and Anderson T.A., 2007, Identification of cotton microRNAs and their targets, Gene, 397: 26-37
http://dx.doi.org/10.1016/j.gene.2007.03.020

http://cgg.biopublisher.ca
Zhang H., Wan Q., Ye W.X., Lv Y.D., Wu H.T., and Zhang T.Z., 2013, Genome-Wide Analysis of Small RNA and Novel MicroRNA Discovery during Fiber and Seed Initial Development in *Gossypium hirsutum* L, PLoS ONE, 8(7): e69743
http://dx.doi.org/10.1371/journal.pone.0069743

Zhang W., and Cohan S.M., 2013, The Hippo pathway acts via p53 and microRNAs to control proliferation and proapoptotic gene expression during tissue growth, Biology Open, 2: 822-828
http://dx.doi.org/10.1242/bio.20134317

Zhou L.G., Liu Y.H., Liu Z.C., Kong D.Y., Duan M., and Luo L.J., 2010, Genome-wide identification and analysis of drought-responsive microRNAs in *Oryza sativa*. J Experimental Botany, 61: 4157-168
http://dx.doi.org/10.1093/jxb/erq237

Zhou M., Li D., Li Z., Hu Q., Yang C., Zhu L., and Luo H., 2013, Constitutive expression of a miR319 gene alters plant development and enhances salt and drought tolerance in transgenic creeping bentgrass, Plant Physiology, 161: 1375-1391
http://dx.doi.org/10.1104/pp.112.208702

Zhu Q.H., Fan L., Liu Y., Xu H., Llewellyn D., and Wilson I., 2013, miR482 Regulation of NBS-LRR Defense Genes during Fungal Pathogen Infection in Cotton, PLoS On, 8: e84390
http://dx.doi.org/10.1371/journal.pone.0084390

Zuker M., 2003, Mfold web server for nucleic acid folding and hybridization prediction, Nucleic Acids Research, 31: 3406-3415
http://dx.doi.org/10.1093/nar/gkg595
Supplementary Table 1 Conserved miRNA families that were differentially expressed in two cotton accessions under two different water regimes

| Small RNA sequence identified in this study | Corresponding conserved miRNA family (obtained from miRBase) | Reads (RPKM)*  | Differential expression in fold change# |
|--------------------------------------------|-------------------------------------------------------------|----------------|----------------------------------------|
| UGAGAUCAUUGUGAAAGCUAAUU                   | hme-bantam                                                  | 9.90631        | 4.80369 | 34.358 down |
| UAUCACAGCCAGCUUUGAUGAGCG                  | api-miR-2b                                                   | 8.24759        | 3.6834 | 23.656 down |
| UCAGUCUUUUCUCUCUCCUAU                     | hme-miR-14                                                  | 10.75067       | 4.93494 | 56.326 down |
| GCUCUCUAUGCUCCUGUAUCA                     | stu-miR156d-3p                                              | 7.62466        | 4.86782 | 6.759 down |
| GCUCUCUAUGCCUCUGUAUCA                     | stu-miR156d-3p                                              | 7.74014        | 3.86782 | 14.644 down |
| GCUCUCUAUGCUUCUGUCAAC                     | stu-miR156d-3p                                              | 9.89768        | 7.95529 | 3.843 down |
| UUGGCAGAAGAUAUGAGAGAC                    | lus-miR156h                                                 | 8.42874        | 7.10486 | 2.503 down |
| GCUCUCUAAGCUUCUGUCAUCAC                   | aly-miR157b-3p                                              | 13.87493       | 11.58505 | 4.890 down |
| AGCUCUUGGCUAUGGAUCC                     | gma-miR159d                                                 | 9.49913        | 6.5199  | 7.885 down |
| GAGCUCUGAGCUUGAUGGAAUCC                   | gma-miR159d                                                 | 12.88205       | 11.63601 | 2.371 down |
| GCGUAUGAGGAGCACAAGCAUAU                  | stu-miR160a-3p                                              | 11.29709       | 9.8451  | 2.735 down |
| UCGCCUGCCCCUGCAUGCCAA                    | nta-miR160d                                                 | 9.39725        | 4.86782 | 23.093 down |
| UCCCUGUCUCUGUAGUCCA                      | cpa-miR160f-5p                                              | 8.63038        | 6.74229 | 3.701 down |
| UCCCUGGCCUCUCCUGUACCCA                   | cpa-miR160f-5p                                              | 9.10681        | 5.5199  | 12.016 down |
| UCCCUGGCCUCUCCUGUACCA                    | cpa-miR160f-5p                                              | 10.17329       | 6.74472 | 10.767 down |
| UGCCAGUCUCUCUCUCUACCA                    | ppt-miR160i                                                 | 10.37992       | 7.45279 | 7.605 down |
| UGGAGGCAGCGGUCAUCAGAU                   | csi-miR162-5p                                               | 10.74014       | 9.03775 | 3.254 down |
| UGCACGUGUCUCUCCUCCUCAAA                 | cpa-miR164e                                                 | 9.46742        | 6.61105 | 7.241 down |
| CACUGUCUCCCUCCUCCUCCAC                 | zma-miR164g-3p                                               | 12.65505       | 10.51794 | 4.398 down |
| GAAAUGUGUUCUGCUCUCUGAGGA               | stu-miR166a-5p                                               | 12.80125       | 10.52603 | 4.840 down |
| GAAAUGUGUUGUCUGACCAGGGG              | bdi-miR166b-5p                                               | 11.49913       | 8.68982 | 7.009 down |
| GAAAUGUGUUUGGGUCUCUAGGGG             | stu-miR166c-5p                                               | 9.59867        | 5.5199  | 16.897 down |
| GAAAUGUGUGUGUUGGCGUGGUCUGGUC         | stu-miR166c-5p                                               | 11.28845       | 7.26312 | 16.283 down |
| UCGGACCAGCUCAUUCACC               | bdi-miR166h-3p                                               | 11.43773       | 9.86567 | 2.973 down |
| UCUCGGACCAGGCCUCAUUCCCS                 | gma-miR166k                                                 | 11.97557       | 10.38866 | 3.004 down |
Continuing Supplementary Table 1

| Small RNA sequence identified in this study | Corresponding conserved miRNA family (obtained from miRBase) | Reads (RPKM)* | Differential expression in fold change# |
|--------------------------------------------|---------------------------------------------------------------|---------------|----------------------------------------|
| UCUCGGGCCAGGCUAUUCCCS                      | gma-miR166k                                                  | 10.79575      | 9.01315 3.440 down                     |
| UCUCUGACCAGGCUAUUCCCS                      | gma-miR166k                                                  | 7.25377       | 5.52638 3.311 down                     |
| GCGGACCAGGCUAUUCCCS                        | lus-miR166k                                                  | 5.73355       | 3.96476 3.407 down                     |
| UCGGACCAAGCUAUUCCCCS                       | lus-miR166k                                                  | 9.31627       | 7.63538 3.206 down                     |
| UCGGACUAGGCUAUUCCCS                        | lus-miR166k                                                  | 9.27674       | 8.10486 2.253 down                     |
| UCAGACCAGGCUAUUCCCS                        | lus-miR166k                                                  | 6.97728       | 4.98397 3.981 down                     |
| UCGGACCUGCGCUAUUCCCS                       | lus-miR166k                                                  | 11.55803      | 10.30998 2.375 down                    |
| UCGGACCAAGGCUACUCCCS                       | lus-miR166k                                                  | 12.04166      | 10.43305 3.049 down                    |
| UCGGACCAAGGAUCAUUCCCS                      | ppt-miR166m                                                  | 11.44909      | 9.97635 2.775 down                     |
| UCGGACCAAGGCUAUUCCCG                       | ptc-miR166p                                                  | 10.89145      | 9.00533 3.696 down                     |
| UCGGACCAAGGCUCAUUCCCS                      | ptc-miR166p                                                  | 11.33563      | 10.12476 2.314 down                    |
| UGAAGCUGCCAGCAUGAUCUUGA                    | cpa-miR167c                                                  | 13.30254      | 12.07014 2.349 down                    |
| UGAAGCUGCCAGCAUGAUCUGG                    | nta-miR167c                                                  | 14.76131      | 13.62149 2.203 down                    |
| UGAAGCUGCCACAGAUGAUCGG                    | cpa-miR167d                                                  | 10.66188      | 8.93551 3.330 down                     |
| UGAAGCUGCCACAGAUGAUGAG                    | cpa-miR167d                                                  | 15.99529      | 14.31078 3.214 down                    |
| UGAAGCUUCCAGAUCUGAUWG                     | cpa-miR167d                                                  | 8.28763       | 9.45279 2.242 up                       |
| GGAUCAUGUGGUAGCUUCACC                     | stu-miR167d-3p                                               | 10.80725      | 8.18286 6.166 down                     |
| AUCAGAUAGAAGUGCCAGCAUGA                   | gma-miR167i                                                  | 10.43008      | 9.42213 2.011 down                     |
| GGCAAGUGUUGUCGCAGCUACAUU                  | aly-miR169a-3p                                               | 9.648         | 7.80369 3.590 down                     |
| UAGCCAAGAAUGACUGCAGCCUCGC                | bdi-miR169d                                                  | 9.53155       | 11.93391 2.486 up                      |
| UAAUGGCCUGUUGUCACUCAGA                   | stu-miR171c-5p                                               | 9.27674       | 7.5199 3.379 down                      |
| AUGAGAAUCUUGAUGAUCUCAGCUC                 | aau-miR172                                                  | 7.31471       | 6.00533 2.478 down                     |
| ACAUUCAGUGCUGCUGGUGGGU                   | hhi-miR-181b                                                | 8.0397        | 9.15323 2.163 up                       |
| UGACUAGAUCACACACAUCCA                     | hme-miR-279a                                               | 11.79459      | 6.45279 40.554 down                    |
| AGAGCUUUCUUCAGUCCACUCU                   | mtr-miR319d-5p                                              | 7.94659       | 6.67518 2.413 down                     |
**Continuing Supplementary Table 1**

| Small RNA sequence identified in this study | Corresponding conserved miRNA family (obtained from miRBase) | Reads (RPKM)* | Differential expression in fold change# |
|--------------------------------------------|---------------------------------------------------------------|----------------|---------------------------------------|
| UAAGCUCAGGAGGGAUAGCGCC                     | cpa-miR390b                                                   | 16.02649       | 2.524 down                           |
| AUCAUGCGAUCCCUCGGAAU                       | mdm-miR393f                                                   | 8.49913        | 2.916 up                             |
| UUCCACAGCUUCCUGAAUG                       | cpa-miR396                                                   | 9.15121        | 0.80486 up                           |
| UUCCACAGCUUUCACUGAAUG                     | cpa-miR396                                                   | 10.28763       | 2.952 down                           |
| UCAUUGAGUGCAACGUUGAAUG                     | ppe-miR397                                                   | 12.01996       | 2.323 down                           |
| UGUGUUCAGGGUACCCCUU                       | cme-miR398b                                                  | 10.01371       | 8.448 down                           |
| UCUUUCUACUCUCCAAUCC                       | pde-miR482a                                                  | 13.73974       | 2.578 down                           |
| UCUUUCUAUCUCCUCACUCC                     | pde-miR482a                                                  | 10.27674       | 8.4585 up                            |
| UCUUUCUCUACUCCUCGCUCC                    | pde-miR482a                                                  | 9.01013        | 2.066 down                           |
| UCUUUCUACUCCCACUCCAUCC                   | pde-miR482a                                                  | 7.63251        | 2.065 down                           |
| UCUUUCUACUCCUCCACUCC                     | pde-miR482a                                                  | 8.66363        | 2.254 down                           |
| UCUUUCUACUCCCACUCCUUCC                   | pde-miR482a                                                  | 10.84457       | 4.960 up                             |
| UCUUUCUACUCCUCCACUCCUUCC                 | pde-miR482a                                                  | 14.37677       | 2.926 down                           |
| UGACAGCGAGAGAGACGACG                     | mes-miR535a                                                  | 9.15121        | 7.84183 down                         |
| UGACAAAGAGAGAAUCAGC                     | mes-miR535a                                                  | 8.01371        | 9.3612 up                            |
| UUAUUGUCUGUUCGACCUA                       | mdm-miR858                                                   | 10.4106        | 8.39437 down                         |
| UUUCGUUGCUUGUUCGACCCUU                   | ath-miR858b                                                  | 13.79095       | 11.76369 down                        |
| GUCCUUGGGGUGCAGAUUACC                      | lja-miR2111-3p                                              | 11.13773       | 7.45279 up                           |
| ACUUUGACUGGUAUUUGCGA                      | ghr-miR2949c                                                 | 12.90947       | 11.15326 down                        |
| CCAGUCUUGCAAUUUAUCCAU                     | api-miR-3027                                                | 9.83355        | 4.80369 down                         |
| UGGAUCUUGUAAACUCGCCUU                   | api-miR-3050                                                | 9.36789        | 8.40369 down                         |
| UGAAACUGUUGUUGUGCUGCU                   | ghr-miR3476-5p                                              | 12.49091       | 10.91222 down                        |
| UUCGAAACCAUCCUCCUCU                        | ghr-miR7505                                                 | 13.66715       | 14.90432 up                          |

* RPKM: Reads per kilo base per million  
# fold change in up or down regulation