

Genome-wide identification of the Dicer-like family in cotton and analysis of the DCL expression modulation in response to biotic stress in two contrasting commercial cultivars

Marianna de Oliveira Moura

Universidade Federal do Rio de Janeiro Instituto de Microbiologia

Anna Karoline S Fausto

Universidade Federal do Rio de Janeiro

Amanda Fanelli

Universidade de Sao Paulo Escola de Engenharia de Lorena

Fernanda A de F Guedes

Universidade Federal do Rio de Janeiro

Tatiane da F Silva

Universidade de Sao Paulo Escola de Engenharia de Lorena

Elisson Romanel

Universidade de Sao Paulo Escola de Engenharia de Lorena

Maite F S Vaslin (✉ maite@micro.ufrj.br)

Universidade Federal do Rio de Janeiro <https://orcid.org/0000-0002-7252-7050>

Research article

Keywords: cotton, *G. hirsutum*, DCL, Dicer-like, CLRDV, biotic stress

Posted Date: November 13th, 2019

DOI: <https://doi.org/10.21203/rs.2.15897/v2>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published on November 15th, 2019. See the published version at <https://doi.org/10.1186/s12870-019-2112-4>.

Abstract

Background

Dicer-like proteins (DCLs) are fundamental players in RNA-silencing mechanisms acting in gene regulation via miRNAs and in antiviral protection in plants, in addition to being related to other biotic and abiotic stresses. Despite being identified in some crops, cotton DCLs weren't characterized until now. Here we characterize the DCLs of three cotton species and analyzed their expression profiles during biotic stress.

Results

We identified 11 DCLs in the allotetraploid cotton *Gossypium hirsutum* and 7 and 6 in the diploid *G. arboreum* and *G. raimondii*, respectively. Among some DCL duplication observed in these genomes, we observe the presence of an extra DCL3 in the three cotton species, not found in others eudicots until now. All the DCL types identified by in silico analysis in the allotetraploid cotton genome were able to generate transcripts, as observed by gene expression analysis in distinct tissues. Based on the importance of DCLs for plant virus defense, responses of cotton DCLs to virus infection and/or herbivore attack using two commercial cotton cultivars (cv.), one susceptible (Fibermax 966) and another resistant (Delta Opal) to polerovirus CLRDV infection were analyzed. Both cvs. responded differently to virus infection. At the initial stages, 24dpi, the resistant cv. showed strong induction of DCL2a and b, while the susceptible cv. showed a down-regulation of these genes, where DCL4 expression was highly induced. Herbivore attack did not induce contrasting profiles between cotton DCLs transcripts in either cotton cv.

Conclusions

The allotetraploid cotton *G. hirsutum* has almost all DCLs found in their diploid relative and duplication of DCL2 and DCL3 were found in the three species. All the four classes of DCL respond to aphid attack and virus infection in *G. hirsutum*, however, remarkable differences in DCL initial responses against the virus itself may be responsible for the virus susceptible and/or resistant phenotype of the contrasting cotton cv. studied.

Background

DICER proteins represent an ubiquitous class of RNase III enzymes that are vital for the establishment of RNA interference (RNAi). DCLs recognize and cut a long ds-RNA substrate to release 21–24-nt small RNA duplexes, which have 2-nt 3' overhangs at each end [1–3]. These small RNAs are important riboregulators in fungi, plants and animals, negatively regulating the expression of specific target genes by base-pairing. The function of small RNAs is largely described in plants associated with both development and responses to abiotic and biotic stress [4–6]. Plant small RNAs are classically separated into short-interfering RNAs (siRNAs) and microRNAs (miRNAs). Different classes of siRNAs have been previously described as natural-antisense siRNAs (natsiRNA) and trans-acting siRNAs (tasiRNAs), which, as miRNAs, are involved in post-transcriptional gene regulation by degrading their mRNA targets or inhibiting

their translation. Heterochromatin-associated siRNAs (hcsiRNA) are associated with chromatin modifications and transcriptional repression of their target DNA *loci*. The long non-translated RNA (lsiRNA) and the viral small RNA (vsiRNA) were generated by DICER cutting of intermediate viral genome structures during viruses infection. Once produced, small RNAs are incorporated into ARGONAUTE (AGO)-containing RNA-induced silencing complexes (RISCs) to confer sequence specificity in the silencing of RNA information [3,5,7,8]. In plants and fungi, cellular RNA-dependent RNA polymerase (RDR) acts to convert aberrant RNAs to dsRNA, leading to small RNA amplification and more intensive RNA silencing [3,9,10].

DICERs, or DICER-like (DCLs), as these proteins are called in plants, present six domains: DEAD, Helicase-C and DUF283 domains at the N-terminus, a Piwi/Argonaute/Zwille (PAZ) domain in the middle and a dual RNase III domain followed by one or two dsRNA-binding domains in the C-terminal half [11,12]. In lower eukaryotes, one or more of these domains may be absent [4]. The helicase domain serves to recruit co-factor regulatory proteins [13–15]. ATP hydrolysis is used to achieve progressive cleavage of the long dsRNA substrate. The DEAD domain acts as an ATP-binding domain [16]. The PAZ domain in turn mediates the recognition of the dsRNA substrate terminus, and the distance between PAZ and the RNase III catalytic center determines the size of the small RNAs produced [17,18]. Each of the two RNase III domains cuts one of the dsRNA strands, leaving a characteristic 2-nt overhang at the 3'-end of the product [19–21], and the C-terminal dsRNA-binding domains (dsRBD or DSRM) act as a protein–protein interaction interface and nuclear localization signals, in addition to having dsRNA-binding functions. dsRBD mediates the discrimination of different RNA substrates and subsequent incorporation of effector complexes [22–24]. Nevertheless, multiple dsRBD may be able to act in combination by recognizing secondary structures of specific RNAs [13].

In *Arabidopsis*, 4 genes (*DCL1–4*) encoding DCL proteins have been found [12,25]. Each of the four AtDCLs is involved in the biogenesis of specific small RNA species. However, they may play redundant and hierarchical roles in the production of distinct sRNAs [26]. AtDCL1 is responsible for miRNA biogenesis. AtDCL2 generates 22-nt siRNAs and viral siRNAs (vsiRNAs) from endogenous inverted-repeats, integrated viruses, transgenes and RDR-amplified virus dsRNAs [26–28]. AtDCL3 is associated with transcriptional silencing, producing 24-nt length siRNAs involved in the establishment and maintenance of the heterochromatin state through RNA-dependent DNA methylation and histone modification [29,30]. These 24 siRNAs are mostly generated from repetitive DNA *loci* and transposons, as well as retrotransposon insertion of DNA regions. AtDCL4 responds to the biogenesis of primary vsiRNAs that has 21-nt and are generated by dicing of viral genomes replicative intermediates. This DCL is also responsible for the production of phased siRNAs, ta-siRNAs and virus-activated siRNAs (vasiRNAs, a new class of endogenous RDR1-dependent siRNAs induced by virus infection) [31–33]. DCL4 is very important in viral resistance, initiating virus silencing in primary infected cells [33]. Primary vsiRNAs may further initiate secondary siRNA production, under the action of AtDCL2 and AtDCL4, and other genes [25,34,35]. So, DCL2 and 4 are key control virus replication levels even in susceptible plants. The production of the secondary vsiRNAs enables the increase of the amplitude of viral defense, making RNAi crucial for anti-virus defense in plants.

In contrast to the number of reports about DCL characterization and functions in *Arabidopsis*, only a few reports have characterized them in other plant species. In general, these reports show the presence of 4 classes of DCLs, with more than one member for each class. Seven *DCLs* have been identified in the tomato *Solanum lycopersicum* [36], eight in rice [37], five in maize [38], five in poplar [12,39], four in grapevine [40], eight in *Brassica napus* [41], three in *Archis duranensis*, and four in chickpea, pigeonpea, *A. ipaensis* [42] and four in pepper [43]. Thus far, cotton DCLs have only been identified in the *G. raimondii* genome [44], but they have not been investigated by comparative evolutionary and expression analyses in cotton species.

Cotton crop is important to the economies of more than 30 countries and represent the principal source of natural fiber in the textile industry worldly. Besides cotton fiber, cotton oil is widely used for human consumption. The evolution of the *Gossypium* genus is marked by two important events. First, a divergence occurred 5–10 million years ago (MYA), when A and D diploid genomes form two separated branches. Subsequently, around 1–2 MYA, allopolyploidization events occurred by interspecific hybridization between A and B-genome ancestors resembling *G. arboreum* and *G. raimondii*, respectively, generating new *Gossypium* species, as the cultivated *G. hirsutum* (AADD) between others [44,45]. As a consequence of the allopolyploidization, thousands of genes were duplicated and showed different expression levels, explaining the drastically phytomorphology changes observed in the allopolyploid cotton, compared to the diploid species *G. arboreum* and *G. raimondii* [44].

Here, we identify and characterize the DCL family of the allotetraploid *G. hirsutum* and two diploids, *G. arboreum* and *G. raimondii*, cotton species. As observed earlier in *G. raimondii* [44], an extra *DCL3* is also present in *G. arboreum* and *G. hirsutum*, where it is currently expressed. Due to the importance of DCL in viral defense, we studied the expression patterns of *G. hirsutum* DCLs in two commercial cotton cultivars (cv.) during biotic stress induced by CLRVDV (*Cotton leafroll dwarf virus*) infection and/or herbivore attack by *Aphis gossypii*. Comparison of a CLRVDV-resistant and susceptible cotton accession or cv. during infection showed that the modulation of *DCL* expression, especially *DCL2a*, *2b*, and *DCL4*, might explain the contrasting viral responses exhibited by each accession. Interestingly, both cotton *DCL3* were also differentially modulated during viral infection.

Results

Identification of cotton Dicer-like gene family and chromosomal localization

Using *Arabidopsis* and rice *DCLs* as query sequences, we identified the *Dicer-like* genes of two diploid cottons, *G. raimondii* and *G. arboreum*, and of the allotetraploid *G. hirsutum* acc. TM-1 were identified by searching the cotton genome database [46–48]. Based on these queries and Blast tool analyses, we identified 6 genes encoding DCL proteins in *G. raimondii*, 7 in *G. arboreum* and 11 in the allotetraploid *G. hirsutum*, respectively (Table 1).

DCL genes were named according to the closest orthologs in *A. thaliana*. Among 4 *DCL* genes in *A. thaliana*, *AtDCL1*, *AtDCL2*, *AtDCL3*, and *AtDCL4* have orthologs in cotton. *AtDCL1* has only one ortholog in

G. raimondii and *G. arboreum*, and two in *G. hirsutum*. *AtDCL2* has 2 orthologs in *G. raimondii*, 3 in *G. arboreum*, and 4 in *G. hirsutum*. *AtDCL3* has 2 orthologs in the two diploid genomes and 3 in *G. hirsutum*. All three cotton species present two *DCL3*. It seems that this duplication could be relevant for cotton evolution as it is present in the three species studied here; however, we cannot exclude the possibility that it is neutral. *AtDCL4* has 2 orthologs in tetraploid cotton, but only 1 ortholog in diploid cottons. Cotton *DCLs* distinct paralogs were coded as a and b, according to their order on the homologous chromosomes.

Detailed information for these genes is listed in Table 1, including the chromosome location, ORF and protein lengths, molecular weight and theoretical isoelectric point. The newly identified *Dicer-like* loci showed coding potentials of 1209 to 2009 amino acid polypeptides, with predicted molecular weights (MW) of 132.99 to 220.99 kDa, respectively. A very small *DCL* putative polyprotein was identified in *G. arboreum* for *DCL2a* (Cotton_A_34031), which were 770 amino acids in length (84.7 kDa). It represents an extra *DCL2a*, that is present only in the *G. arboreum* genome.

The physical location of the three cotton species *DCL* genes is shown in Figure 1. A total of eleven *GhDCL* genes were distributed on 10 *G. hirsutum* chromosomes. All the chromosomes (A01, A04, A05, A06, A07, D01, D06, D07, and D13) contained a single representative of *GhDCL*, with the exception of chromosome D05, which contained *GhDCL2a* and *GhDCL4*. Curiously, almost all the *GhDCLs* were correlated with chromosomes inherited from each parental-related species. *G. raimondii* presented 6 *DCL* genes distributed in 5 chromosomes, and *G. arboreum* 7 had *DCLs* distributed in 5 chromosomes and in a scaffold region (scaffold3086), that had not yet been incorporated into the physical map of chromosomes. *G. arboreum* presented two *DCL2a* located in very close proximity on chromosome 10, separated by approximately 9 kb. These two *DCL2as* (named herein as *GaDCL2a31* and *32*) were highly similar, sharing 93.4% identity at the amino acid level. The predicted *GaDCL2a32* was shorter than its orthologs and seemed to have lost nucleotides/amino acids at the 5' extremity/N-terminus.

The intron/exon (I/E) distribution as well as intron numbers, are shown in Table 1 and Figure 2. *GaDCL1* (Cotton_A_14097) showed the longest ORF of 6027 bp and coding potential for a polypeptide of 2009 amino acids. The maximum number of introns, 25, was found in *DCL3a* and *b* of *G. raimondii*. *DCL4* of all three distinct species showed a very similar intron/exon distribution. The same could be observed for *DCL3a* of *G. arboreum*, *G. raimondii*, and *G. hirsutum* from subgenome D. *GhDCL3a* from subgenome A, however, showed a distinct pattern at the 5' end of the gene. It seemed to have lost part of the nucleotides from this region. A very similar distribution of introns and exons was also observed for *DCL2b* of *G. raimondii* and *G. hirsutum* subgenomes A and D. *GaDCL2b*, on the other hand, showed a very different distribution of I/E.

Cotton DCL phylogenetic and domain composition analysis

The phylogenetic relationships of the amino acid sequence of the three cotton species (*G. raimondii*, *G. arboreum*, and *G. hirsutum*) were used to construct a neighbor-joining (N-J) phylogenetic tree, using MEGA 7.0 software (Figure 2). These analyses indicated that the *DCL* genes clustered in two separate clades, one composed of *DCL1*, *2*, and *4* and the other of *DCL3s*. The two *GhDCL* subfamilies were further divided

into many subclades that clearly grouped the *DCLs* from the parental/ancestral diploid species with each corresponding subgenome in the allotetraploid cotton. The extra *DCL3*, named herein as *DCL3b*, was present in all three species, but the allotetraploid cotton presented only one member of this gene on the genome, showing the closest relationship to *GrDCL3b*. We can hypothesize that the *DCL3b* gene, that was acquired from *G. arboreum* during the *G. hirsutum* evolution, had been lost. Alternately, it may have been acquired independently by the three species after allotetraploid hybridization. The two *DCL2a* from *G. arboreum* formed a group together with *GhDCL2a* from subgenome A, showing that they probably duplicated after tetraploid hybridization. The constructed tree suggested a high level of sequence conservation for *DCL* sequences in the diploid species and the allotetraploid *G. hirsutum* during evolution.

SMART was used to identify the DCL domains in all Dicer-like genes from the three species (Figure 2). All DCLs contained a DEAD, helicase-C, DUF283 and PAZ domain, excluding GaDCL2a31 and GaDCL2b, which did not present any DEAD or helicase-C domains. In fact, GaDCL2a31 was a truncate DCL because it also did not present a DUF283 domain. Two RNase III domains were present in all DCLs, and at least one DSRM domain was present in 15 out of the 22 cotton DCL proteins. All 7 cotton DCL3s had a dsRB domain, which is a characteristic of DCL3 plant proteins. Interestingly, GhDCL1 from subgenome D seemed to have lost a DSRM domain, while DCL1 of subgenome A maintained 2 DSRB domains.

A comparison of cotton DCL proteins with those from Arabidopsis, rice, poplar, grapevine and *Medicago* sp. revealed that cotton DCL1s shared a common ancestor with grapevine and poplar DCL1s (Figure 3). DCL2a and 2b from cotton, however, were more related to poplar DCL2s. Grapevine and poplar DCL3s showed the closest relationship to cotton DCL3 from all species analyzed. DCL3 duplication in rice seemed to occur over a long time before cotton DCL3 duplication, but after dicot/monocot divergence. The duplication of DCL3 in cotton probably occurred before cotton allotetraploid hybridization and therefore more than 1.2 MYA. DCL4 seemed to be the most conserved DCL among the eudicots analyzed herein.

DCL gene expression profiles in different G. hirsutum organs

To evaluate if the *in silico*-identified *DCLs* in cotton were able to generate transcripts, we collected tissue samples from root, leaf, stem and flower from greenhouse-grown cotton of two distinct commercial *G. hirsutum* cv., Fibermax 966 (FM966) and Delta Opal (DO), at 60 days post germination (60 dpG) and analyzed gene expression by quantitative real-time PCR (RT-qPCR). These two cultivars were especially selected because they show contrasting phenotypes against an important worldwide distributed cotton virus disease, the cotton blue disease (CBD). Fibermax is susceptible to CBD, while Delta Opal is resistant.

As shown in Figure 4, transcripts of all 6 *DCLs* types were identified in *G. hirsutum* plants. *DCL1* was expressed in all analyzed tissues from both cv. FM and DO, as well as *DCL2a*, *DCL3a*, and *DCL4*. *DCL2b* was not detected in leaves from the DO cv. and flowers of cv. FM, while *DCL3b* was almost undetectable in stems from FM and leaves from both cultivars. *DCL4*, which is essential for intracellular antiviral silencing, was expressed at the same levels in FM and DO plants, as well as *DCL2a*. However, *DCL2b*

showed slightly contrasting basal expression levels between them. The extra *DCL3* identified in cotton seemed to be important, especially in flower and root tissues from healthy plants.

DCL gene expression is modulated in response to herbivore attack and virus infection

Plant DCLs initiate the RNAi innate defense system against invading viruses because they recognize and process incoming viral and transposon nucleic acids into small siRNAs of 21, 22, and 24 nts. Thus, we were interested in shedding some light on how cotton *DCL* expression is modulated during RNA viral infection using virus-resistant and virus-susceptible contrasting cotton cvs.: Delta Opal and FM966, respectively. *Cotton leafroll dwarf virus* (CLRVDV) (genus, *Polerovirus*; family, *Luteoviridae*), which is transmitted only by an aphid vector (*Aphis gossypii*), is the causal agent of cotton blue disease [49]. CLRVDV is phloem-restricted, and its genome consists of a single-strand, positive-sense, non-polyadenylated RNA (5.8 kb) containing six open reading frames (ORFs) [50].

As CLRVDV is only transmitted by its aphid vector, another important point is to understand the aphid vector component in the DCL modulation. Consequently, we evaluated cotton DCL expression patterns during aphid herbivore attack, mediated by *Aphis gossypii* and/or CLRVDV infection.

To analyze the influence of herbivore attack on DCL expression, 30 day cotton plants were inoculated in the greenhouse with virus-free aphids. Aphid were restricted to one basal leave (inoculated leave) per plant and twenty-four hours after inoculation, the aphids were eliminated by insecticide application. The expression levels of all cotton DCLs were evaluated in young systemic leaves (3–4 leaves above the inoculated leaves) at 24 hpi, 5, 15, and 25 dpi after contact with the aphids (Figure 5). For CLRVDV infection, a similar biological approach was applied using viruliferous aphids harboring CLRVDV.

In general, all *DCLs* showed an induction of their expression levels in the virus-resistant DO plants after aphid contact (Figure 5). When these plants were subjected to the aphid and virus simultaneously, in the case of viruliferous aphid contact, DO *DCL* levels showed an otherwise repression pattern. In the virus-susceptible cv, the inverse was observed (Figure 6).

Interestingly, it was observed that after the first 24 hpi a systemic modulation of all *DCLs* mRNAs occurred in both cotton cvs., meaning that aphid feed in inoculated leaves is inducing a systemic response. For FM plants, this systemic response is more pronounced than for DO plants. At 24 hpi, FM plants showed strong down regulation of the three cotton DCLs involved in virus defense, *DCL2a*, *2b* and *4*, with reductions of approximately 5, 10 and 15-fold, respectively. DO plants, in contrast, showed a systemic upregulation of these DCLs. This contrasting modulation of DCLs by aphid feeding may predispose the CBD susceptible FM cv. to be more vulnerable against virus infection, as they have less *DCL2* and *4* accumulation in the incoming virus tissues. Whereas for DO plants, when the virus is trying to spread from the local infection site, it faces a strong antiviral silencing pre-activated in systemic tissues that present 3–5 higher levels of *DCL2a*, *2b* and *4* than healthy plant tissues. The presence of previously systemic accumulation of these DCL induced distally by aphid feeding can help these plants to block the virus infection cycle establishment in new cells far from local infection site.

DCL1 mRNA expression was also modulated, been induced in systemic leaves of DO cv. 24 hpi and 5 days after aphid contact. At 15 and 25 dpi, *DCL1* expression was reduced in both FM and DO cvs. in comparison to healthy control plants (Figure 5).

DCL3a expression was induced by aphid feeding in both DO and FM cvs. at all time points analyzed. Important exceptions were noted for FM at 24 hpi and for DO at 25 dpi, at which time this DCL was downregulated. In contrast, the cotton extra *DCL3* (*DCL3b*) showed strong repression during the first 24 hpi to 5 dpi in both cvs.

When DO and FM plants were infected with viruliferous aphids (Figure 6), cotton *DCLs* showed a distinct expression pattern in systemic leaves, compared with those inoculated with virus-free aphids, showing that the presence of virus also modulate *DCL* expression (Figure 6 and Figure S2 - Additional file 2). The fold change analysis between FM CLRDV-infected and FM mock plants (virus-free aphid inoculum) showed that the presence of virus induced an additional systemically down-regulation of almost all their *DCLs* during the initial stages of virus infection (24 hpi), with the exception of *DCL2a*. Five days later, however, the levels of all FM *DCLs* markedly increased, showing almost 60, 10, 47, 17, 28 and 28-fold change variations for *DCL1*, *DCL2a*, *DCL2b*, *DCL3a*, *DCL3b*, and *DCL4*, respectively. In contrast, the *DCL* transcript levels from DO plants at all time points were lightly induced or repressed, maintaining levels that were very similar to uninfected mock plants. Even with the strong DCL modulation observed in the susceptible plant FM, CLRDV could be easily detected in systemic leaves from 24 hpi, showing that even with such extensive efforts of the antiviral machinery, the virus was replicating and spreading throughout the plant (Figure 6B). This strong modulation of *DCL* expression was not observed in DO plants, in which *DCLs* were only slightly induced or reduced up to 25 dpi, and the virus was not detected in any systemic leaves between 24 hpi and 25 dpi (Figure 6B). Absence of virus accumulation in DO plants may explain why DO DCLs were almost no induced at these leaves.

Deep sequencing of viral small RNA from CLRDV FM-infected plants, previously performed by our group at this same time point on systemic leaves, showed that the most abundant viral siRNAs were 22 nt [51], highlighting the importance of DCL2 in combating the spread and/or replication of the virus in aerial parts far from inoculated leaves. These efforts are insufficient for the blockade of virus infection. Both DCL2s probably participate in the processing/dicing of viral dsRNA for the generation of these second viral siRNAs of 22 nts. However, as *DCL2b* levels are more than 4 times higher than *DCL2a* levels at the systemic infected leaf cells 5 dpi, we can hypothesize that *DCL2b* is most important for 22-nt viral siRNA generation.

Modulation of DCL2 and DCL4 expression by virus/aphid infection at local infection sites

It has been already shown for Arabidopsis that DCL4 is an essential component of intracellular antiviral silencing, whereas both DCL4 and DCL2 are necessary for the inhibition of systemic infection. Our results indicated that *DCL4* was almost not modulated by CLRDV infection in aerial parts of virus-resistant cotton during infection, while *DCL2a* and *2b* were downregulated (Figure 6). In susceptible plants, however, *DCL4* was strongly downregulated systemically at the beginning of infection. Thus, the next step

was to examine how these *DCLs* were expressed at the infection site. Consequently, we collected samples from inoculated leaves 24 h after virus infection and analyzed *DCL* mRNA expression profiles by RT-qPCR (Figure 7).

Aphid contact induced *DCL2a* and *DCL4* expression at similar levels in both FM and DO cvs. (Figure 7A). *DCL2b* in turn was 6 times more induced in DO than in FM plants. Local aphid attack modulation responses seems to be the same in the resistance/susceptibility phenotype for *DCL4* as both plants showed similar level of induction of this DCL at the inoculated leaves.

The presence of the virus, in contrast, induces stronger modulations of almost all DCLs. At virus infection sites, *DCL2a*, *DCL2b*, and *DCL3b* were strongly induced (approximately 250, 70, and 100-fold change, respectively) in the resistant cv. at 24 hpi (Figure 7B). The strong *DCL2a* and *2b* induction seemed to be relevant for virus resistance because systemic spread of the virus was completely inhibited in DO plants (Figure 6B). Virus replication is inhibited even in the virus inoculated leaves, as observed by high sensitivity nested RT-PCR for CLRDV detection (Figure 7C). The susceptible cv. FM did not respond to virus presence in the same way as the DO cv., as only a very slight induction of *DCL2a* and *b* was observed at 24 hpi in inoculated leaves. In these plants virus accumulation is observed in systemic and inoculated leaves since 24 hpi (Figure 6B and 7C, respectively).

An induction of FM *DCL4* 6x higher than that of DO *DCL4* was also observed. This finding indicated that both susceptible and resistant plants produced more *DCL4* to combat virus invasion; however, this overexpression alone did not seem to be sufficient to inhibit the spread of infection, since FM plants that produce more *DCL4* than DO are completely susceptible. So, the two cotton DCL2s seem to be very important to avoid virus dissemination and accumulation at inoculation sites while DCL4 may have a secondary paper. However, we cannot say with these results that the strong induction of DCL2 is the responsible for DO CLRDV resistance phenotype as unrelated resistance mechanism may be acting also. Curiously, *DCL3a* and *3b* were highly induced in the FM cv. (11 and 260-fold greater expression than in mock-infected plants and more than 2-fold in DO plants, respectively), while only *DCL3b* was induced in DO cv. (approximately 100-fold greater expression than in the mock). These results showed that *DCL3b* seemed to be important during the initial virus defense activation.

Taken together, our results suggested that the contrasting CLRDV susceptibility phenotype demonstrated by FM and DO plants might be related to their distinct DCL modulation mediated both by aphid feed and virus infection.

Corroborating the importance of cotton *DCL2* in the polerovirus infection, a profile of the vsRNAs produced by DCL dicing was obtained for CLRDV and another polerovirus infected cotton plants by deep sequencing. As observed in Figure 7D, in the plants infected with both CLRDV and Cotton anthocyanosis virus (CAV), 22-nt vsRNAs accumulated in major levels than 21-nt vsRNAs showing the relevant paper of cotton DCL2s in the virus defense.

Biotic and abiotic stress-responsive cis-acting regulatory elements are present in the promoters of upland cotton DCL genes.

Promoter sequences 1.5 kb upstream of the translation start of all *G. hirsutum* *DCL* genes were obtained from the cotton genome project to attempt to understand how cotton *DCLs* are modulated by both virus infection and herbivore attack. Transcriptional responsive *cis*-elements of *DCL* gene promoters were analyzed using PlantCare. Analysis of the promoter region of all 11 upland cotton *DCL* genes revealed the presence of various biotic and abiotic stress-responsive *cis*-acting regulatory elements, including the TCA-element, ERE and ABRE. Light stress-responsive elements were relatively the most abundant in the promoters of the upland cotton *DCL* genes, specifically Box1, Box4 and GT1-motif (Figure 8 and Additional file 3 Figure S3), indicating that all *DCL* proteins might have an important functional role in light stress responses. All *DCL* promoters displayed the development of *cis* elements, especially HD-Zip1 and 2, and almost all the elements that are responsive to drought (especially MBS), salicylic acid (TCA) and other biotic stresses (as WUN, box-W1, ELI-box3, TC-rich repeats, Box S, JERE, and GCC boxes), revealing possible mechanisms mediated by almost all *DCLs* in drought tolerance and biotic stress responses in the upland cotton *G. hirsutum*. Surprisingly, *GhDCL4A* did not show typical biotic stress or ethylene-responsive *cis* elements, but they were present in *GhDCL4D*. In general, there were significant differences in the average proportions of the promoter elements detected within the different *DCL* gene families, as well as between the same *DCLs* originating from a distinct parental diploid cotton (Figure 8). However, abiotic stress elements presented the highest average proportions in all *DCLs*. Phytohormone-responsive elements, especially those associated with ethylene and gibberellin, were found in *DCL 1*, *DCL2a* and *2b*, *DCL3b*, and *DCL4*, while those correlated to Me-jasmonate were predominant in *DCL2b*, *DCL3*, and *DCL4-A*. A large number of enhancer elements were found in all *DCLs* (Additional file 3 Figure S3), suggesting that all the *DCLs* from the two subgenomes were able to generate transcripts.

Discussion

In the present study, we first characterized the *DCLs* of 3 cotton species (the commercial allotetraploid *G. hirsutum* and its parental-like *G. arboreum* and *G. raimondii*) at the genomic level. Our results showed that cotton carried extra copies of *DCL2* and *DCL3*. The extra *DCL2* has been previously reported in poplar, rice and maize between other plants. Our phylogenetic analysis corroborated other findings, showing that *DCL2* duplication seemed to have evolved independently in mono- and eudicots, after their genetic separation. The *DCL3* duplication shown here is a unique event in eudicots [12]. The extra *DCL3*, denoted herein as *DCL3b*, showed a complete *DCL3* structure in the allotetraploid *G. hirsutum* and in the diploids *G. arboreum* and *G. raimondii*. Furthermore, quantitative RT-PCR showed that it was expressed in distinct cotton tissues/organs such as flowers and roots. The importance of this extra *DCL3* for cotton must be studied, but we can hypothesize that it may be associated with the large number of transposons detected in the cotton genome. *G. raimondii*, *G. arboreum* and *G. hirsutum* genome account approximately to 53, 67 and 62% of retrotransposons, respectively [44]. Curiously, in allotetraploid cotton, we found only one set of *DCL3b*, which probably derived from its *G. raimondii*-like parental species. *DCL3b* from the parental *G. arboreum* seemed to have been missed during evolution.

Analyzing expression of these cotton extra DCL transcripts during virus infection and herbivore attack, we could observe that they seem to be functional and are modulated in both stresses. However, a further characterization of all the cotton DCL2s and DCL3s in heterologous system will be necessary to understand the paper of these DCLs extra copies in RNA silencing machinery.

Previous studies from our group have shown that CLRDV infection in *G. hirsutum* impacts small RNA accumulation both quantitatively and qualitatively [53,54]. Our deep sequencing identified several cotton miRNAs that were down- or upregulated during CLRDV infection of the susceptible FiberMax 966 *G. hirsutum* cv. [52]. All these miRNA families participate in distinct and overlapping networks, such as plant disease resistance (miR393, 403 and 472), leaf development (miR159, 164, 319 and 393), flowering time (miR172, 156 and 157), stress nutrition (miR395, 397, 398 and 408), and miRNA pathways (miR403) [53]. In addition to the global alteration of Gh-miRNAs during CLRDV infection, we observed a decrease in the frequency of 24-nt sRNA in the infected plant with an enrichment of *gypsy*- and *copia*-like retrotransposon transcripts [52]. In contrast, the CLRDV viral siRNA profile during infection revealed a greater abundance of the 22-nt vsRNAs [54]. Thus, cotton DCLs are anticipated to play an important role in the mediation of CLRDV:cotton interactions during this compatible infection.

Here, the analysis of the expression of the 6 cotton *DCLs* in two contrasting virus-susceptible commercial cotton cvs., during two types of biotic stress, revealed that almost all DCLs were modulated during these stresses. The CLRDV-susceptible cv. FM966 showed a strong downregulation of *DCL2a*, *2b*, *3a*, *3b*, and *4* in the aerial part of the plant during the initial hours after aphid contact. This downregulation was maintained until 25 dpi for *DCL2b*, *DCL2b*, and *DCL4* at the majority of the analyzed time points. For the CLRDV-resistant cv. Delta Opal, this downregulation was not observed, except for *DCL3b*, which was repressed almost 15-fold compared with the mock-infected control plants at 24 hpi (Figure 5). So, sensing of aphid sucking is able to distinctly modulate *DCL* expression in aerial parts of these two cvs. that were not in contact with the aphid. However, at the inoculated leaves (Figure 7) *DCL2a* and *DCL4* modulation by herbivore attack was very similar in both cotton cvs. Thus, virus resistance and/or susceptibility did not seem to be a consequence of differential responses to aphid attack at the initial site of aphid contact but was more likely a consequence of the virus versus cv. interaction. Bozorov et al. (2012) [55] studied the effects of *N. attenuata* *DCL* silencing on the performance of *Manduca sexta* larvae. They found that *NaDCL2*, *NaDCL3*, and *NaDCL4* are involved in the regulation of a number of different genes, leading to signaling and defense metabolite activation in response to herbivore attack. Thus, like *N. attenuata* *DCLs*, cotton *DCLs* are apparently important in anti-herbivore defense activation, as they are clearly modulated locally and especially systemically by aphid sucking.

Antiviral activity of DCLs was already intensively reported [56–59, among others]. Although DCL4 is the major producer of 21-nt antiviral siRNA, DCL2 is essential for secondary 22-nt siRNA-mediated transitive virus silencing and can compensate the loss of DCL4 in *dcl4* mutants [60–63]. During potyvirus *Zucchini yellow mosaic virus* (ZYMV) infection of *N. benthamiana*, for example, DCL4 has been observed to restrict the systemic movement of the virus [64]. The authors observed an enhancement of systemic virus accumulation in double *dcl2* and *4* and, particularly, in the triple knockdown mutant *dcl2/3/4*.

Examination of the modulation of *DCL* expression during virus infection by CLRDV showed that the virus induced strong systemic repression of *DCL 1, 2b, 3b, and 4* in the susceptible cv. FM966 during the initial hours after infection. This drastic repression was not observed in the resistant cv., in which the expression of all these *DCLs* remained almost similar to that of mock infected plants. Investigation of *DCL* expression modulation at infection sites showed that even FM *DCL4* was expressed 6x higher than DO *DCL4*, and this high induction of *DCL4* was insufficient for virus restriction to the initially infected cells. In contrast, both FM *DCL2a* and *b* were almost not induced in the inoculated leaves. Others authors have reported previously that *DCL2* and *DCL4* may have different antiviral activities in inoculated and systemic leaves [33,56,59]. In these studies it was observed that *DCL4* alone responded to the antiviral silencing in inoculated leaves. In systemic leaves, however, both *DCL2* and *DCL4* are necessary to trigger virus silencing. It has been also previously demonstrated that *DCL4* can downregulate *DCL2* expression [56]. FM *DCL4* overexpression may impair *DCL2a* and *b* expression at the inoculation site in the susceptible cotton cv., allowing the cell-to-cell spread of the virus and establishment of systemic infection due to the noneffective intercellular antiviral activity of *DCL2* at these sites. In DO-resistant plants, the high levels of *DCL2a* and *b* induced by the presence of virus may efficiently activate intercellular antiviral silencing, resulting in restriction of the virus to the first cell-autonomous or inoculated cells during the initial hours after infection and dramatically reducing virus replication and aborting infection spread completely.

DCL2 acts at intercellular VIGS, generating the second front line against viral infection in virus-neighboring recipient cells [56]. CLRDV, similar to all poleroviruses, is able to replicate only in companion cells of the phloem, and the presence of *DCL2* at high levels in the neighborhood of companion cells may impose strong antiviral silencing mediated by the 22-nt virus small RNAs, blocking the spread of the virus in subsequent tissues in resistant DO plants.

Donaire et al. (2009) [65] analyzed the vsiRNAs derived from nine viruses belonging to eight different genera. The predominant class of vsiRNAs were 21 nt for most viruses, with the exception of tombusvirus *Cymbidium ring spot virus*-infected plants that accumulated higher levels of 22-nt vsiRNAs. Recently, *Brassica yellows virus* (BrYV) vsiRNA accumulation was analyzed in single-infected *N. benthamiana* plants, and the 22-nt class represented the dominant class of BrYV vsiRNAs, followed by the 21-nt class [66]. These results are in agreement with our previous findings [51,54] and the present studies in which 22-nt vsiRNAs were the most abundant class of vsiRNAs in two other virus species in the polerovirus genus. The accumulation of 22-nt vsiRNAs appears to be characteristic of polerovirus infections, highlighting the important role of *DCL2* in antiviral defense against these viruses, which may exceed that of *DCL4*.

In addition to classical *DCL4* and *DCL2* antiviral functions, several authors have reported the importance of *DCL3* in virus defense via the production of high levels of 24-nt siRNAs from invading viruses [26,67,68], guiding AGO family members to cleave viral RNA or recruit cellular machinery to methylate viral DNA. In *A. thaliana*, *DCL3* are also responsible for the directed cleavage of positive-sense RNA viruses, including *Cucumber mosaic virus* (a cucumovirus), *Oilseed rape mosaic virus* (a tobamovirus) and *Turnip crinkle virus* (a carmovirus) [26,69]. Thus, *DCL3* also plays an important role in virus silencing.

In CLRDV-infected cotton, *GhDCL3a* and *b* were induced on infected leaves (with high levels of expression), as well as systemic leaves, suggesting a possible contribution of these DCLs in viral defense.

It should take in account, however, that our results do not exclude that other CLRDV-resistance mechanisms not directly related to DCL2/4 modulation should be active in DO plants. These resistance mechanisms may turn down virus replication consequentially decreasing virus imposed stress levels in these plants. One way to deeply check the importance of the DCL modulation observed here by itself in the CLRDV resistance in DO plants would be induce the silencing of each cotton DCL genes individually and/or in combinations in DO cotton plants by VIGS and CLRDV-infect plants the plants further.

Conclusions

We identified *G. raimondii*, *G. arboreum* and *G. hirsutum* cotton DCLs and found duplication of DCL2 and DCL3. Additionally, we found that all *G. hirsutum* DCLs are involved in polerovirus infection responses, especially the DCL2 and DCL3 paralogs and DCL4, and that modulation of the expression of these DCLs by viruses may be crucial for virus susceptibility and/or resistance response.

Methods

Identification of the DCL family in Gossypium

The genome sequences of three cotton species, *G. arboreum* (BGI-CGB v2.0 assembly genome), *G. raimondii* (JGI assembly v2.0 data) and *G. hirsutum* acc. TM-1 (NAU-NBI v1.1 assembly genome), were downloaded from the Phytozone (<http://www.phytozone.net/>) and CottonGen (<https://www.cottongen.org/>) databases. Dicer-like protein sequence data were obtained for *A. thaliana* and *O. sativa* from the General Feature Format (GFF) in *Arabidopsis* Information Resource (TAIR release 10, <http://www.arabidopsis.org>) and from the Rice Genome Annotation Project Database (RGAP release 7, <http://rice.plantbiology.msu.edu/index.shtml>). Gene names and IDs are listed in Additional File 1 and in Additional File 2 Table S1).

The physico-chemical properties of cotton DCL proteins were predicted using the ExPASy Compute pI/Mw tool (http://au.expasy.org/tools/pi_tool.html; Bjellqvist et al. 1993).

Chromosomal location analysis and phylogenetic tree construction

The locations of the DCLs in chromosomes were assessed using MapInspect software (<http://www.softsea.com/review/MapInspect.html>) using start and end position of each open read frame obtained from the genome database..

A phylogenetic tree was constructed using MUSCLE (*Multiple Sequence Comparison by Log-Expectation*) alignment and the neighbor-Joining (NJ) method in MEGA 7.0 software [70], with the 1000-replicate

bootstrap test. A keyword search of the Phytozome v12.1 database (<https://phytozome.jgi.doe.gov/pz>) and National Center for Biotechnology Information database - NCBI (<https://www.ncbi.nlm.nih.gov>) was further performed to obtain the DCL genes in different plant organisms.

Intron–exon and domain analysis of the DCL family

The Gene Structure Display Server (<http://gsds.cbi.pku.edu.cn/index.php>) was used to analyze the intron–exon structure by comparing the CDS of cotton *DCL* genes with their corresponding genomic sequences [45,71]. The conserved domains in the DCLs were identified by SMART (<http://smart.embl-heidelberg.de>).

Plant materials

Two upland cotton *G. hirsutum* cultivars were used for the *DCL* expression assays and viral infection: acc. FiberMax966 (FM966) (Aventis Crop Science, Australia) and acc. Delta Opal (DO) (Delta and Pine Land Co., United States). Seeds were kindly provided by IMA, Instituto Matogrossense do Algodão, Primavera do Leste, Mato Grosso state, Brazil. Plantlets were grown under greenhouse conditions at 28 +/- 2°C as previously described [49].

For cotton organ *DCL* expression, samples of leaves, shoots, flowers, and roots from independent plants at 30 days after germination (dag) were collected from each cotton acc. FM and DO. Samples were immediately frozen in liquid N₂ and stored at –80°C until RNA extraction.

Plant aphid inoculation and virus infection

Cotton (*Gossypium hirsutum*) plants of cultivars acc. FM966 (susceptible to Cotton blue disease) and acc. DO (resistant to Cotton blue disease) grown in the greenhouse and at the 30 dag stage were infected with *Cotton leafroll dwarf virus* (CLRDV, polerovirus, *Luteoviridae* family) by viruliferous aphids (*Aphis gossypii* Glover) as described previously [52]. Full-developed leaves of plants with the same age from both cultivars were inoculated with non-viruliferous aphids (virus free aphids). Aphids were restricted at the inoculation site by double-side adhesive tapes (3MM Co.) and killed 24 hours after inoculation with insecticide. Aphids were restricted to the inoculation sites at the inoculated leaves by surrounding the inoculation site with a double face tap.

Systemic leaves, localized 3–4 leaves above the inoculated leaf, were collected at 24 hours post-infection (hpi) and 5, 15 and 25 days post-infection (dpi) with CLRDV aphids or non-viruliferous aphids. Leaves from the same position of uninfected plants were used as uninoculated controls. Leaves from 3–5 independent inoculated plants composed each biological replicate. Samples were stored at –80°C until RNA isolation and expression analysis.

All the samples recovered from aphid-free inoculated, CLRDV infected and uninoculated were assayed for CLRDV detection by a Nested RT-PCR assay that amplifies viral coat protein sequence following [51]. CLRDV susceptible infected plants were used as control.

Real-Time Quantitative RT-PCR (RT-qPCR)

Total RNA was isolated using the Invisorb Spin Plant RNA Kit (Invitex Molecular Co., Germany). The quality and concentration of each RNA samples was determined by agarose gel electrophoresis and using a NanoDrop 2000 (Thermo Fisher Scientific Co.) spectrophotometer. Only RNA samples with a 260/280 nm ratio between 1.8–2.1 and 260/230 ratio ≥ 2.0 were used. The c-DNA first-strand **reverse transcription** was conducted with the Revertaid First Strand cDNA, Synthesis Kit (Thermo Fisher Scientific Co.) in accordance with the manufacturer's instructions. One microgram of total RNA from each sample was treated with DNase I (Fermentas Co.), and cDNAs were synthesized by adding 100 μM of oligo dT24V primer. *G. hirsutum* DCL reverse and forward primers were designed using NCBI primer-BLAST (Additional File 2 Table S2). Maxima SYBR Green/ROX qPCR Master Mix (Thermo Fisher Scientific Co.) was used to perform RT-qPCR in a 7500 Fast Real-Time PCR System (Applied Biosystems), in accordance with the manufacturer's instructions. The cycling conditions were 10 min at 95°C for initial denaturation, followed by 40 cycles of denaturation at 95°C for 15 s and annealing/extension at 60°C for 30 s. Cotton *GhmiR390* and *GhPP2A* (catalytic subunit of phosphatase 2A) were used as reference genes for the analysis of *DCL* expression in different organs, the results were obtained by ΔCT method and the $2^{-\Delta\Delta\text{CT}}$ method was used to analyze *DCL* expression during viral infection and/or aphid inoculation [73,74]. Cotton *GhmiR390* and *GhPP2A* (PROTEIN PHOSPHATASE 2), which were previously identified as the best reference genes for CLRDV infected cotton studies [72], were used to normalize cDNA expression levels. Reactions were prepared in a total volume of 20 μL , which contained 10 μL of SYBR green master mix, 2 μL of cDNA template, 6 μL of ddH₂O, and 2 μL of each primer to make a final concentration of 10 μM . Three biological replicates were performed, and three technical replicates were assayed per cDNA sample.

In order to show DCL expression in each organ, a heat-map of the *DCLs* expression pattern was established by complete clustering method analysis using Euclidean distance and calculated in the R software environment. The DCL Cts results are shown at Additional File 2 Table S3.

Sequencing of CAV sviRNAs by deep sequencing

Leaves of cotton infected with Cotton anthocyanosis virus (CAV) obtained in Mato Grosso state, Brazil, were used for total RNA extraction and small RNA purification and sequencing by Illumina platform following procedures described by [54].

Statistical analysis

DCL relative expression levels determined for the different samples under herbivore attack and/or virus infection were compared with these controls: untreated plants (control) x aphid inoculated plants for herbivore expression analysis, and mock (virus-free aphids inoculated plants) x CLRDV-aphid inoculated plants for viral infection using the parametric one-way ANOVA test at $P \leq 0.05$ and $P \leq 0.01$. For checking if the different relative expression levels of DCLs, from FM and DO, are statistical significant T-student test were used.

Promoter region cis-acting element analysis

For the identification of cis-regulatory elements present in *G. hirsutum* DCLs, upstream sequences of 1500 nucleotides from ATG star codon were downloaded from the CottonGen website (<https://www.cottongen.org>) and cis elements predicted with PlantCARE software (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>).

Abbreviations

CLR DV: Cotton leafroll dwarf virus DCL: Dicer-like protein

DO: Delta Opal cotton cultivar *FM*: Fibermax966 cotton cultivar

GhDCL: Dicer-like gene or protein from *Gossypium hirsutum*

GrDCL: Dicer-like gene or protein from *Gossypium raimondii*

GaDCL: Dicer-like gene or protein from *Gossypium arboreum*

RNAi: RNA interference

vsRNA: viral small interfering RNAs

Declarations

Ethics approval and consent to participate

Not applicable.

Consent to publish

Not applicable.

Availability of data and materials

All data generated or analysed during this study are included in this published article and in its supplementary information files.

Competing interests

None of the authors have any competing interests.

Funding

This research was financially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/Brazil), Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro

(FAPERJ/Brazil) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP/Brazil). MOM and AKSF received grants from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/Brazil) for master and PhD development, respectively. None of the funding bodies was involved in the design of the study, the collection, analysis, and interpretation of the data, nor in the writing of the manuscript, which was made entirely by the authors.

Authors' contributions

MFSV and ER conceived and designed the experiments. MOM and AKSF performed the experiments. TFS, MOM, AKSF, AF and ER carried out bioinformatics analyses and analyzed data. FAFG performed transcriptome analysis. MOM and MVFS wrote the paper. All authors read and approved the final manuscript.

Acknowledgements

We thank Sarah M. Nardeli for help and grateful discussions.

References

1. Bartel DP: MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. 2004, 116: 281-297. doi: 10.1016/S0092-8674(04)00045-5.
2. Henderson IR, Zhang X, Lu C, Johnson L, Meyers BC, Green PJ, Jacobsen SE: Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nat Genet*. 2006, 38: 721-725. doi: 10.1038/ng1804
3. Borges F, Martienssen RA: The expanding world of small RNAs in plants. *Nat Rev Mol Cell Biol*. 2015, 16(12):727-741. doi: 10.1038/nrm4085.
4. Liu QP, Feng Y, Zhu ZJ: Dicer-like (DCL) proteins in plants. *Functional & Integrative Genomics*. 2009, 9(3):277-286. doi: 10.1007/s10142-009-0111-5.
5. Aliyari R, Ding SW: RNA-based viral immunity initiated by the Dicer family of host immune receptors. *Immunol Rev*. 2009, 227: 176-188. doi: 10.1111/j.1600-065X.2008.00722.x.
6. Khraiweh B, Zhu JK, Zhu J: Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim Biophys Acta*. 2012, 1819:137-48. doi: 10.1016/j.bbagr.2011.05.001
7. Hall TM: Structure and function of argonaute proteins. *Structure*. 2005, 13(10):1403-1408. doi: 10.1016/j.str.2005.08.00
8. Rivas FV, Tolia NH, Song JJ, Aragon JP, Liu J, Hannon GJ, Joshua-Tor L: Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat Struct Mol Biol*. 2005, 12(4):340-349. doi: 10.1038/nsmb918

9. Dalmay T, Hamilton A, Rudd S, Angell S, Baulcombe DC: An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell*. 2000, 101:543-553.
10. Wang X, Wu Q, Ito T, Cillo F, Li WX, Chen X, Yu JL, Ding SW: RNAi-mediated viral immunity requires amplification of virus-derived siRNAs in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. USA*. 2010, 107, 484-489. doi: 10.1073/pnas.0904086107
11. Bernstein E, Caudy AA, Hammond SM, Hannon GJ: Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*. 2001, 409:363-366. doi: 10.1038/35053110
12. Margis R, Fusaro AF, Smith NA, Curtin SJ, Watson JM, Finnegan EJ, Waterhouse PM: The evolution and diversification of Dicers in plants. *FEBS Lett*. 2006, 580:2442-2450. doi: 10.1016/j.febslet.2006.03.072.
13. Lee Y, Hur I, Park SY, Kim YK, Suh MR, Kim VN: The role of PACT in the RNA silencing pathway. *EMBO J*. 2006, 25:522-532. doi: 10.1038/sj.emboj.7600942
14. Ye XC, Paroo Z, Liu QH: Functional anatomy of the Drosophila MicroRNA-generating enzyme. *J Biol Chem*. 2007, 282:28373-28378. doi: 10.1074/jbc.M705208200.
15. Welker NC, Pavelec DM, Nix DA, Duchaine TF, Kennedy S, Bass BL: Dicer's helicase domain is required for accumulation of some, but not all, *C. elegans* endogenous siRNAs. *RNA*. 2010, 16(5): 893-903. doi: 10.1261/rna.2122010.
16. Cenik ES, Fukunaga R, Lu G, Dutcher R, Wang YM, Hall TMT, Zamore PD: Phosphate and R2D2 restrict the substrate specificity of Dicer-2, an ATP-driven ribonuclease. *Mol Cell*. 2011, 42:172-184. doi: 10.1016/j.molcel.2011.03.002.
17. MacRae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA: Structural basis for double-stranded RNA processing by Dicer. *Science*. 2006, 311:195-198. doi: 10.1126/science.1121638.
18. MacRae IJ, Doudna JA: Ribonuclease revisited: structural insights into ribonuclease III family enzymes. *Curr Opin Struct Biol*. 2007, 17:138-145. doi: 10.1016/j.sbi.2006.12.002.
19. Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T: Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*. 2001, 411:494-498. doi: 10.1038/35078107.
20. Takeshita D, Zenno S, Lee WC, Nagata K, Saigo K, Tanokura M: Homodimeric structure and double-stranded RNA cleavage activity of the c-terminal RNase III domain of human Dicer. *J Mol Biol*. 2007, 374:106-120. doi: 10.1016/j.jmb.2007.08.069.
21. Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W: Single Processing Center Models for Human Dicer and Bacterial RNase III. *Cell*. 2004, 118:57-68. doi: 10.1016/j.cell.2004.06.017.

22. Doyle M, Badertscher L, Jaskiewicz L, Güttinger S, Jurado S, Hugenschmidt T, Kutay U, Filipowicz W: The double-stranded RNA binding domain of human Dicer functions as a nuclear localization signal. *RNA*. 2013, 19:1238–1252. doi: 10.1261/rna.039255.113.
23. Wostenberg C, Lary JW, Sahu D, Acevedo R, Quarles KA, Cole JL, Showalter SA: The role of human Dicer-dsRBD in processing small regulatory RNAs. *PLoS ONE*. 2012, 7 (12): e51829. doi: 10.1371/journal.pone.0051829.
24. Kini HK, Walton SP: In vitro binding of single-stranded RNA by human Dicer. *FEBS Lett*. 2007, 581:5611–5616. doi: 10.1016/j.febslet.2007.11.010.
25. Moissiard G, Parizotto EA, Himber C, Voinnet O: Transitivity in Arabidopsis can be primed, requires the redundant action of the antiviral Dicer-like 4 and Dicer-like 2, and is compromised by viral-encoded suppressor proteins. *RNA*. 2007, 13:1268–1278. doi: 10.1261/rna.541307.
26. Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O: Hierarchical action and inhibition of plant DICER-like proteins in antiviral defence. *Science*. 2006, 313: 68–71. doi: 10.1126/science.1128214.
27. Kurihara Y, Watanabe Y: *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. *Proc. Natl Acad. Sci*. 2004, 101:12753–12758. doi: 10.1073/pnas.0403115101.
28. Mlotshwa S, Pruss GJ, Peragine A, Endres MW, Li J, Chen X, Poethig RS, Bowman LH, Vance V: DICER-LIKE2 plays a primary role in transitive silencing of transgenes in *Arabidopsis*. *PLOS One*. 2008, 3:e1755. doi: 10.1371/journal.pone.0001755.
29. Pontes O, Li CF, Costa Nunes P, Haag J, Ream T, Vitins A, Jacobsen SE, Pikaard CS: The *Arabidopsis* chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. *Cell*. 2006, 126:79–92. doi: 10.1016/j.cell.2006.05.031.
30. Xie ZX, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC: Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol*. 2004, 2:642–652. doi: 10.1371/journal.pbio.0020104.
31. Xie Z, Allen E, Wilken A, Carrington JC: DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci USA*. 2005, 102:12984–12989. doi: 10.1073/pnas.0506426102.
32. Cao M, Dua P, Wang X, Yua Y-Q, Qiu Y-H, Lia W, Gal-Onc A, Zhou C, Lib, Y, Ding S-W: Virus infection triggers widespread silencing of host genes by a distinct class of endogenous siRNAs in *Arabidopsis*. *PNAS*. 2014, 111: 40:14613–14618. doi: 10.1073/pnas.1407131111.
33. Dunoyer P, Himber C, Voinnet O: DICER-LIKE 4 is required for RNA interference and produces the 21-nucleotide small interfering RNA component of the plant cell-to-cell silencing signal. *Nat Genet*. 2005, 37:1356–1360. doi: 10.1038/ng1675.

34. Wang XB, Jovel J, Udornporn P, Wang Y, Wu Q, Li WX, Gascioli V, Vaucheret H, Ding SW: The 21-nucleotide, but not 22- nucleotide, viral secondary small interfering RNAs direct potent antiviral defense by two cooperative argonautes in *Arabidopsis thaliana*. *Plant Cell*. 2011, 23:1625–1638. doi: 10.1105/tpc.110.082305.
35. Parent J-S, Bouteiller N, Elmayan T, Vaucheret H: Respective contributions of *Arabidopsis* DCL2 and DCL4 to RNA silencing. *Plant J*. 2015, 81:223-232. doi: 10.1111/tpj.12720.
36. Bai M, Yang GS, Chen WT, Mao ZC, Kang HX, Chen GH, Yang YH, Xie BY: Genome-wide identification of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families and their expression analyses in response to viral infection and abiotic stresses in *Solanum lycopersicum*. *Gene*. 2012, 501:52–62. doi: 10.1016/j.gene.2012.02.009
37. Kapoor M, Arora R, Lama T, Nijhawan A, Khurana JP, Tyagi AK, Kapoor S: Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA Polymerase gene families and their expression analysis during reproductive development and stress in rice. *BMC Genomics*. 2008, 9:451. doi: 10.1186/1471-2164-9-451.
38. Qian Y, Cheng Y, Cheng X, Jiang H, Zhu S, Cheng B: Identification and characterization of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in maize. *Plant Cell*. 2011, 30:1347–1363. doi: 10.1007/s00299-011-1046-6.
39. Zhao K, Zhao H, Chen Z, Feng L, Ren J, Cai R, Xiang Y: The Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in *Populus trichocarpa*: gene structure, gene expression, phylogenetic analysis and evolution. *Journal of Genetics*. 2015, 94: 317-321.
40. Zhao H, Zhao K, Wang J, Chen X, Chen Z, Cai R, Xiang, Y: Comprehensive analysis of Dicer-like, Argonaute, and RNA-dependent RNA polymerase gene families in grapevine (*Vitis vinifera*). *J. Plant Growth Regul*. 2015, 34:108–121. doi: 10.1007/s00344-014-9448-7.
41. Cao JY, Xu YP, Li W, Li SS, Rahman H, Cai XZ: Genome-wide identification of Dicer-like, Argonaute, and RNA-dependent RNA polymerase gene families in *Brassica* species and functional analyses of their *Arabidopsis* homologs in resistance to *Sclerotinia sclerotiorum*. *Front. Plant Sci*. 2016, 7:1614. doi: 10.3389/fpls.2016.01614.
42. Garg V, Agarwal G, Pazhamala LT, Nayak SN, Kudapa H, Khan AW, Doddamani D, Sharma M, Kishor PB, Varshney RK: Genome-Wide Identification, Characterization, and Expression Analysis of Small RNA Biogenesis Purveyors Reveal Their Role in Regulation of Biotic Stress Responses in Three Legume Crops. *Frontiers in Plant Science*. 2017, 8:488. doi: 10.3389/fpls.2017.00488.
43. Qin L, Mo N, Muhammad T, Liang Y: Genome-Wide Analysis of DCL, AGO, and RDR Gene Families in Pepper (*Capsicum Annuum*) *Int. J. Mol. Sci*. 2018, 19: e1038. doi:

10.3390/ijms19041038.

44. Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J, Yoo M-J, Byers R, Vaslin MF and other authors: Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. *Nature* 2012, 492:423-427. doi: 10.1038/nature11798.
45. Li F, Fan G; Lu C., Xiao G, Zou C, Kohel R.J, Ma Z, Shang, H, Ma, X, Wu J, Liang, X, Huang G, Percy RG, Liu K, Yang W, Chen W, Du X, Shi C, Yuan Y, Ye W, other authors: Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nature*. 2015, 3: 524-530. doi: 10.1038/nbt.3208.
46. Wang K, Wang Z, Li F, Ye W, Wang J, Song G, Yue Z, Cong L, Shang H, Zhu S, Zou C, Li Q, Yuan Y, Lu C, Wei H, and other authors: The draft genome of a diploid cotton *Gossypium raimondii*. *Nat Genet*. 2012, 44:1098-1103. doi: 10.1038/ng.2371.
47. Li F, Fan G, Wang K, Sun F, Yuan Y, Song G, Li Q, Ma Z, Lu C, Zou C, Chen W, Liang X, Shang H, Liu W, Shi C, Xiao G, Gou C, Ye W, and other authors: Genome sequence of the cultivated cotton *Gossypium arboreum*. *Nat Genet*. 2014, 46:567-72. doi: 10.1038/ng.2987.
48. Zhang T., Hu Y, Jiang W, Fang L, Guan X, Chen J, Zhang J, Saski CA, Scheffler BE, Stelly DM, Hulse-Kemp AM, Wan Q, Liu B, Liu C, Wang S, Pan M, Wang Y, Wang D, Ye W, Chang L, other authors: Sequencing of allotetraploid cotton (*Gossypium hirsutum* acc. TM-1) provides a resource for fiber improvement. *Nat Biotechnol*. 2015, 33:531-537. doi: 10.1038/nbt.3207.
49. Correa RL, Silva TF, Simões-Araujo JL, Barroso PA, Vidal MS, Vaslin MF: Molecular characterization of a virus from the family *Luteoviridae* associated with cotton blue disease. *Arch Virol*. 2005, 150:1357-1367. doi: 10.1007/s00705-004-0475-8.
50. Distefano AJ, Bonacic Kresic I, Hopp HE: The complete genome sequence of a virus associated with cotton blue disease, cotton leafroll dwarf virus, confirms that it is a new member of the genus *Polerovirus*. *Archives of Virology* 2010, 155:1849-1854. doi: 10.1007/s00705-010-0764-3.
51. Silva TF, Romanel EAC, Andrade RRS, Farinelli L, Osteras M, Deluen C, Correa RL, Scharago CEG, Vaslin MFS: Profile of small interfering RNAs from cotton plants infected with the polerovirus *Cotton leafroll dwarf virus*. *BMC Molecular Biology*. 2011, 12:1471-2199. doi: 10.1186/1471-2199-12-40.
52. Romanel E, Silva TF, Corrêa RL, Farinelli L, Hawkins JS, Schrago CE, Vaslin MF: Global alteration of microRNAs and transposon-derived small RNAs in cotton (*Gossypium hirsutum*) during *Cotton leafroll dwarf polerovirus* (CLRDV) infection. *Plant Mol Biol*. 2012, 80(4-5):443-60. doi: 10.1007/s11103-012-9959-1.
53. Mallory AC, Vaucheret H: Functions of microRNAs and related small RNAs in plants. *Nat Genet*. 2006, 38:S31-S36. doi: 10.1038/ng1791.

54. Andrade RSR, Vaslin MFS: Search Small RNA: a graphical interface tool for the assemblage of viral genomes using small RNA libraries data. *Virology Journal*. 2014, 7:11-45. doi: 10.1186/1743-422X-11-45.
55. Bozorov TA, Pandey SP, Dinh ST, Kim SG, Heinrich M, Gase K, Baldwin IT: DICER-like proteins and their role in plant-herbivore interactions in *Nicotiana* *Journal of Integrative Plant Biology*. 2012, 54: 189–206. doi: 10.1111/j.1744-7909.2012.01104.x.
56. Qin C, Li B, Fan Y, Zhang X, Yu Z, Ryabov E, Zhao M, Wang H, Shi N, Zhang P, Jackson S, other authors: Roles of Dicer-Like Proteins 2 and 4 in Intra- and Intercellular Antiviral Silencing *Plant Physiology*. 2017, 174:1067–1081. doi: 10.1104/pp.17.00475.
57. Diaz-Pendon JA, Li F, Li W-X, Ding S-W: Suppression of Antiviral Silencing by Cucumber Mosaic Virus 2b Protein in *Arabidopsis* is Associated with Drastically Reduced Accumulation of Three Classes of Viral Small Interfering RNAs. *Plant Cell Online*. 2007, 19:2053-2063. doi: 10.1105/tpc.106.047449.
58. Zhang X, Zhang X, Singh J, Li D, Qu F: Temperature-dependent survival of *turnip crinkle virus-infected Arabidopsis* plants relies on an RNA silencing-based defense that requires DCL2, AGO2, and HEN1. *J. Virol*. 2012, 86:6847–6854. doi: 10.1128/JVI.00497-12.
59. Garcia-Ruiz H, Takeda A, Chapman EJ, Sullivan CM, Fahlgren N, Brempelis KJ, Carrington JC: Arabidopsis RNA-dependent RNA polymerases and dicer-like proteins in antiviral defence and small interfering RNA biogenesis during *Turnip mosaic virus* *Plant Cell*. 2010, 22:481–496. doi: 10.1105/tpc.109.073056.
60. Gascioli V, Mallory AC, Bartel DP, Vaucheret H: Partially redundant functions of *Arabidopsis* DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. *Curr. Biol*. 2005, 15: 1494–1500. doi: 10.1016/j.cub.2005.07.024
61. Schauer SE, Jacobsen SE, Meinke DW, Ray A: DICER-LIKE1: Blind men and elephants in *Arabidopsis* *Trends Plant Sci*. 2002, 7: 487-491. doi: 10.1016/S1360-1385(02)02355-5.
62. Bouche N, Laressergues D, Gascioli V, Vaucheret H: An antagonistic function for *Arabidopsis* DCL2 in development and a new function for DCL4 in generating viral siRNAs. *EMBO J*. 2006, 25: 3347–3356. doi: 10.1038/sj.emboj.7601217.
63. Qu F, Ye X, Morris TJ: *Arabidopsis* DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. *Proc. Natl Acad. Sci. USA*. 2008, 105, 14732–14737. doi: 10.1073/pnas.0805760105.
64. Cordero T, Cerdán L, Carbonell A, Katsarou K, Kalantidis K, Daròs JA: Dicer-Like 4 Is Involved in Restricting the Systemic Movement of *Zucchini yellow mosaic virus* in *Nicotiana benthamiana*. *Mol Plant Microbe Interact*. 2017, 30:63-71. doi: 10.1094/MPMI-11-16-0239-R.

65. Donaire L, Wang Y, Gonzalez-Ibeas D, Mayer KF, Aranda MA, Llave C: Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes. *Virology*. 2009, 392:203–214. doi: 10.1016/j.virol.2009.07.005.
66. Zhou C-J, Zhang X-Y, Liu S-Y, Wang Y, Li D-W, Yu J-L, Han C-G: Synergistic infection of BrYV and PEMV 2 increases the accumulations of both BrYV and BrYV-derived siRNAs in *Nicotiana benthamiana*. *Sci Rep*. 2017, 7: 45132. doi: 10.1038/srep45132.
67. Fusaro AF, Matthew L, Smith NA, Curtin SJ, Dedic-Hagan J, other authors: RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. *EMBO Rep*. 2006, 7: 1168–1175. doi: 10.1038/sj.embor.7400837.
68. Llave C: Virus-derived small interfering RNAs at the core of plant-virus interactions. *Trends in Plant Science*. 2010, 15 (12):701-707. doi: 10.1016/j.tplants.2010.09.001.
69. Blevins T, Rajeswaran R, Shivaprasad PV, Beknazariants D, Si-Ammour A, Park H S, Vazquez F, Robertson D, Meins F, other authors: Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Res*. 2006, 34: 6233–6246. doi: 10.1093/nar/gkl886.
70. Kumar S, Stecher G, Tamura K: MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol*. 2016, 33(7):1870-4. doi: 10.1093/molbev/msw054.
71. Hu G, Koh J, Yoo MJ, Chen S, Wendel JF Gene-expression novelty in allopolyploid cotton: a proteomic perspective. *Genetics*. 2015, (1):91-104. doi: 10.1534/genetics.115.174367.
72. Fausto AKS, Silva T F, Romanel E, Vaslin M: MicroRNAs as reference genes for quantitative PCR in cotton. *PLoS One*. 2017, 12: e0174722. doi: 10.1371/journal.pone.0174722.
73. Livak, KJ, Schmittgen, TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCt} method. *Methods*. 2001, 25:402–408. doi: 10.1006/meth.2001.1262.
74. Schmittgen TD, Livak KJ: Analyzing real-time PCR data by the comparative Ct method. *Nat Protoc*. 2008, 3(6):1101-1108.

Table 1

Table 1: The cotton DCL genes and properties of the deduced proteins

Serial no.	Gene name	Accession numbers	ORF (bp)	Introns no.	Genomic location	Protein		
						Length (a.a.)	MW (kDa)	pI
<i>Gossypium hirsutum</i>								
1	GhDCL1A	Gh_A07G1222	5868	18	A07:27651487-27660850	1955	215.05	5.69
2	GhDCL1D	Gh_D07G1326	3849	14	D07:21226922-21232919	1289	141.79	5.89
3	GhDCL2aA	Gh_A04G0311	4005	20	A04:7175960-7232649	1334	146.74	6.07
4	GhDCL2aD	Gh_D05G3343	4188	21	D05:54148674-54156601	1395	153.45	6.23
5	GhDCL2bA	Gh_A01G1100	4245	20	A01:42143646-42153237	1414	155.54	6.26
6	GhDCL2bD	Gh_D01G1133	4179	21	D01:24614805-24624231	1392	153.12	6.21
7	GhDCL3aA	A06:22197055-22207732*	4481	24	A06:22197055-22207732	1209	132.99	6.13
8	GhDCL3aD	Gh_D06G0845	4839	24	D06:15420427-15431049	1612	177.32	5.68
9	GhDCL3bD	Gh_D13G2027	4329	23	D13:55748697-55756048	1442	158.62	5.62
10	GhDCL4A	Gh_A05G0400	4926	23	A05:4495791-4511254	1648	181.28	6.33
11	GhDCL4D	Gh_D05G0516	4968	24	D05:4159682-4172501	1655	182.05	6.38
<i>Gossypium arboreum</i>								
12	GaDCL1	Cotton_A_14097	6027	19	Chr12:122863850-122873212	2009	220.99	5.78
13	GaDCL2a	Cotton_A_34031	3533	7	Chr10:48146458-48149990	770	84.70	5.50
14	GaDCL2a	Cotton_A_34032	3723	19	Chr10:48129898-4813760	1240	136.40	5.92
15	GaDCL2b	scaffold3086	4207	11	scaffold3086:32023-41620	803	88.33	6.35
16	GaDCL3a	Cotton_A_31241	4980	24	Chr7:97079503-97090015	1659	182.49	5.90
17	GaDCL3b	Cotton_A_03190	4941	22	Chr13:83564342-83577790	1647	181.17	5.79
18	GaDCL4	Cotton_A_01271	4842	24	Chr4:32633913-32646902	1613	177.43	6.12
<i>Gossypium raimondii</i>								
19	GrDCL1	Gorai_001G150400_1	5853	19	Chr01:20836533-20846178	1950	214.5	5.67
20	GrDCL2a	Gorai_009G374900	4188	23	Chr09:50970390-50979491	1395	153.45	6.20
21	GrDCL2b	Gorai_002G143000	4182	23	Chr02:25828154-25839065	1393	153.23	6.25
22	GrDCL3a	Gorai_010G093100_1	5028	25	Chr10:15093632-15104933	1675	184.25	5.77
23	GrDCL3b	Gorai_013G221400_1	4389	25	Chr13:54135530-54142923	1462	160.82	5.79
24	GrDCL4	Gorai_009G053300_1	4965	24	Chr09:3872981-3886515	1654	181.94	6.29

* gene not annotated. The letters A and D after DCL name in *G. hirsutum* indicate the sub-genome A and D where they were identified, respectively. Cotton DCLs parologs were coded as "a" and "b" according to their chromosome position. ORF: open reading frame length; aa: amino acid; MW: molecular weight; pI: theoretical isoelectric point.

Additional Files Legend

Additional File 1: Cotton DCL c-DNA and protein sequences

Additional File 2: Table S1: Gene name and gene ID of DCLs in *Arabidopsis*, *Medicago*, rice, *Populus*, *Physcomitrella* and grapevine. Table S2: List of primers and amplicon characteristics of DCL genes. Table S3: DCL expression in organs. Figure S1: DCL amplification melting curves. Figure S2: Expression pattern of *G. hirsutum* acc. FM DCL genes based on transcriptome sequencing data.

Additional File 3: Figure S3 - Cis-elements distribution in DCL promoters with enhancer elements. Figure S4: Promoter sequences of *Gossypium hirsutum* DCLs.

Additional File 4: Positions of all cis elements found in *G. hirsutum* DCLs promoters

Figures

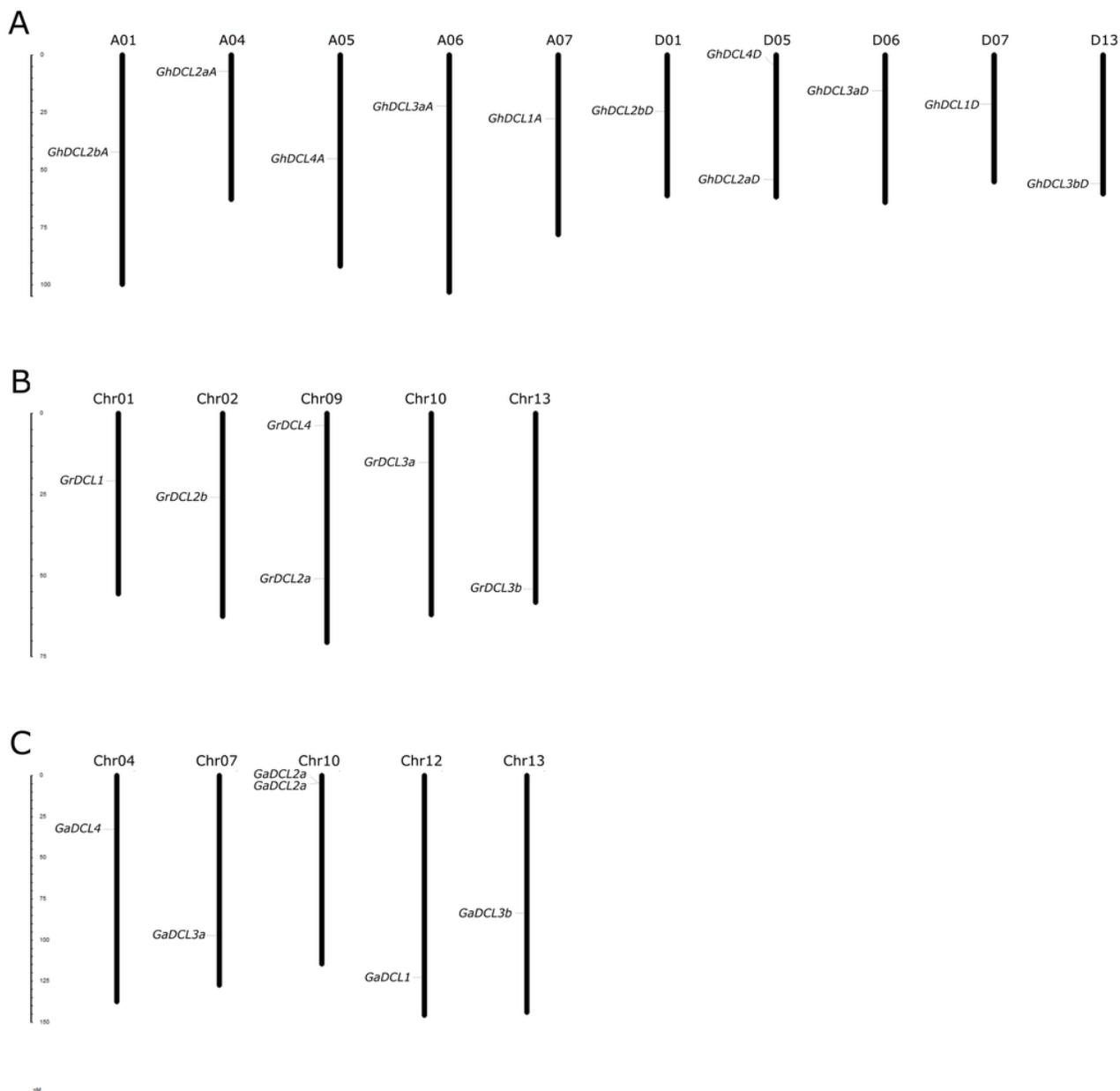


Figure 1

Chromosomal distributions of DCL genes in *Gossypium*. *G. hirsutum* DCLs (GhDCLs) are distributed between chromosomes (Chr) A01, A04-07, D01, D05-07 and D13; *G. raimondii* DCLs (GrDCLs) between chromosomes 01, 02, 09, 10 and 13, and *G. arboreum* DCLs (GaDCLs) between chromosomes 04, 07, 10, 12 and 13, respectively.

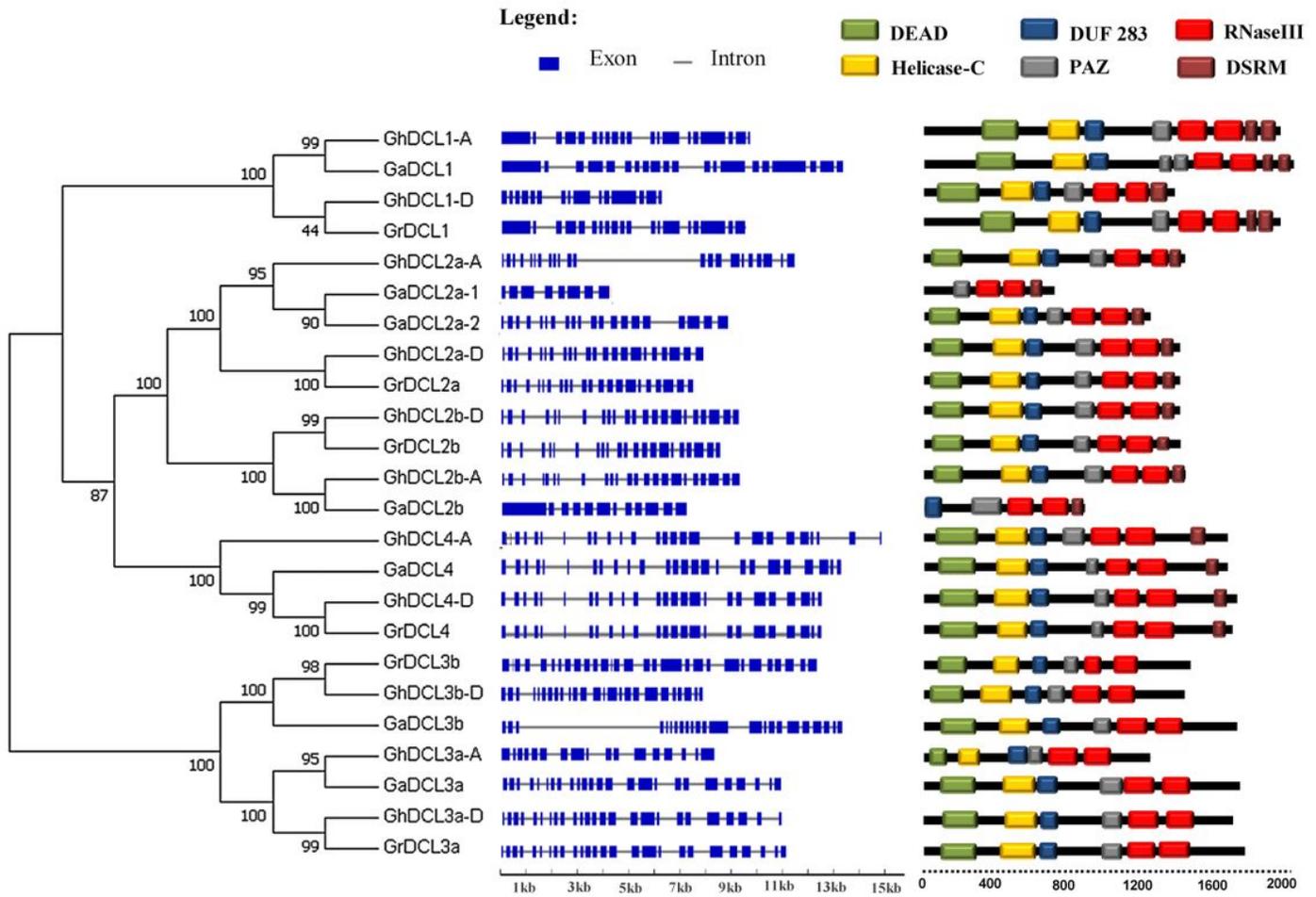


Figure 2

Phylogenetic tree, gene structure and domain compositions of cotton DCLs. The phylogenetic tree was constructed using MEGA 7.0. Exon/intron structures of DCL genes are represented by blue boxes and gray lines, respectively. Protein domain analysis are represented by different colors. Sequence used are listed at Additional File 1.

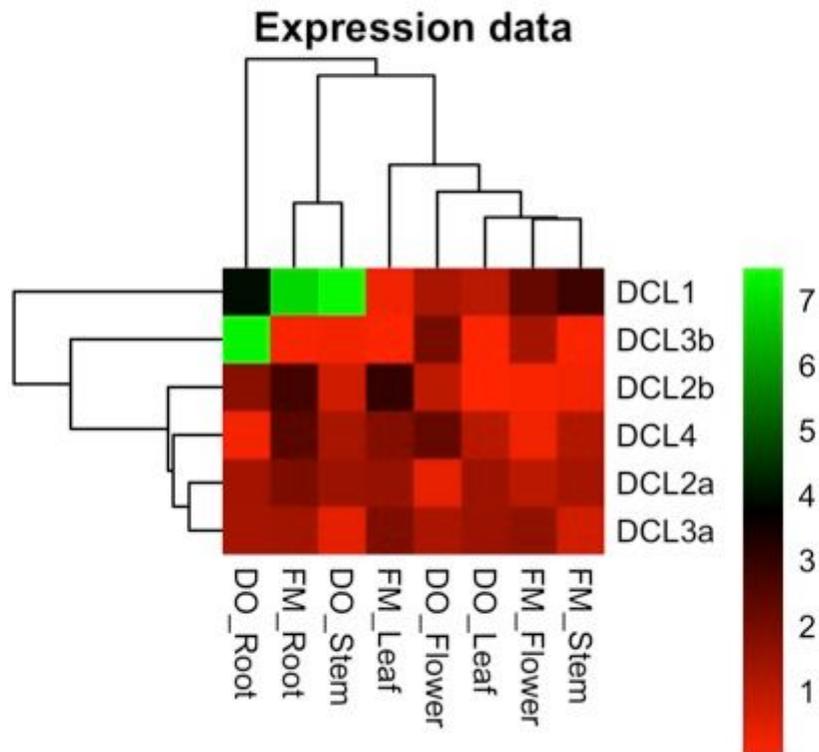


Figure 4

Real-time PCR expression profiling of Dicer-like genes in different cotton tissues. DCL expressions levels of all cotton DCLs were analyzed in stem, flower, leaf and root of two commercial cotton cv. Fibermax 966 (FM) and Delta Opal (DO) at 60 dag by qRT-PCR. The heat map was generated by a log transformation of the real-time PCR data presented as $-\Delta\Delta Ct$ ($Ct_{DCL} - Ct_{GhmiR390/GhPP2A}$). The expression values are detailed at Additional File 2 Table S3.

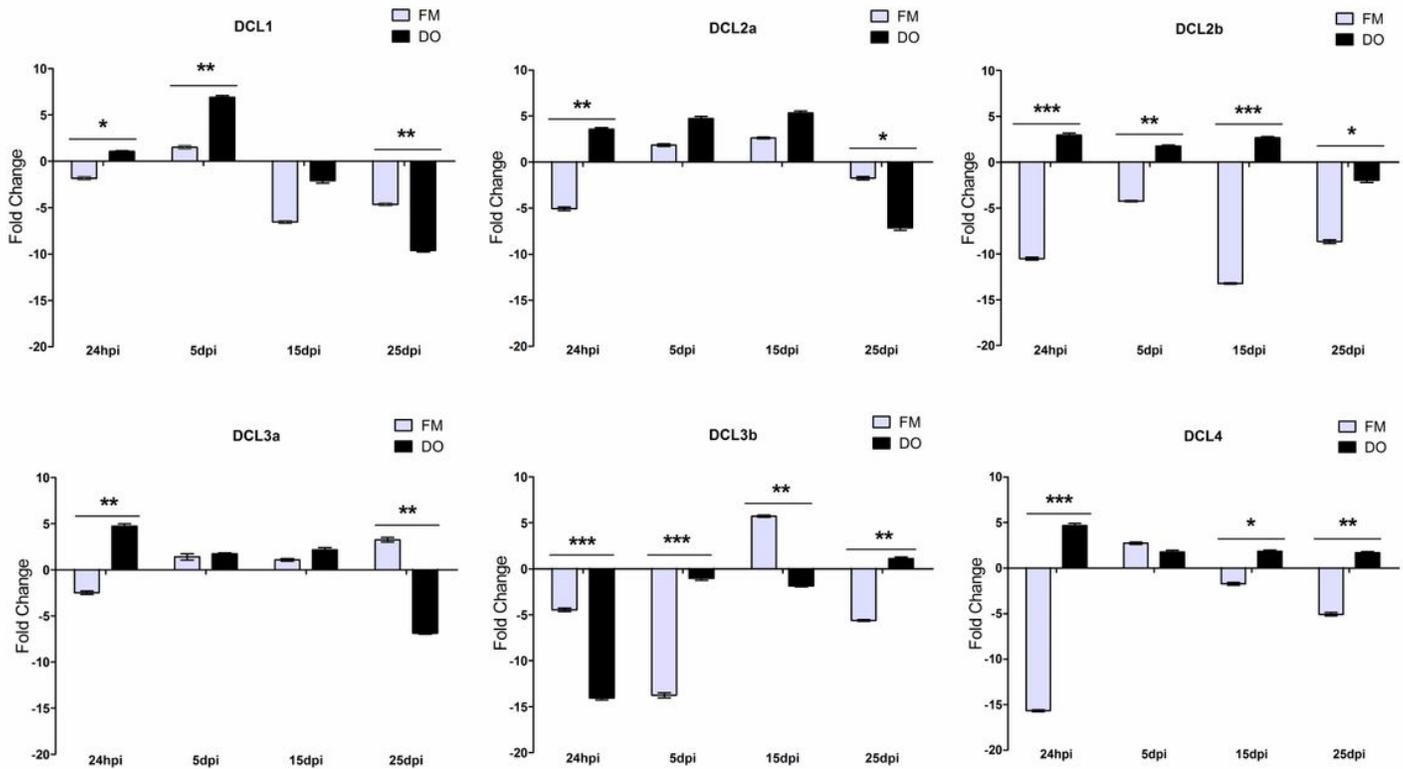
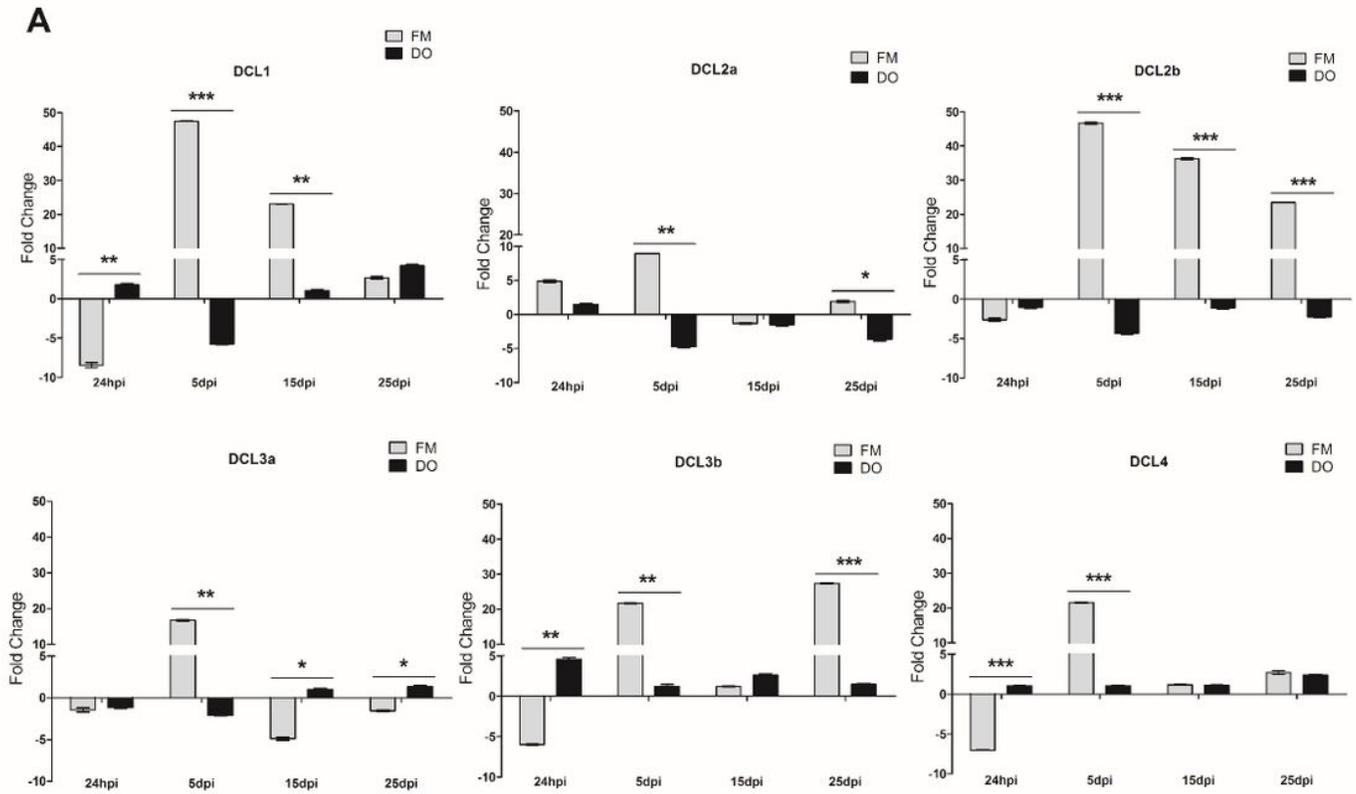


Figure 5

Expression profile of cotton DCL analysis by RT-qPCR during herbivore attack. Fibermax 966 and Delta Opal cotton plants were submitted to contact with 10 aphids for 24 hours. Systemic leaves were collected at 24 hpi and 5, 15 and 25 dpi for evaluation of each DCL level over time. The fold change was calculated using the $2^{-\Delta\Delta Ct}$ method as described by [76]. The standard deviation is indicated by the error bars, and "*" indicates significant differences, calculated using the GraphPad Prism program. * P < 0.1 ** P < 0.01 *** P < 0.001. hpi - hours post-aphid inoculation, dpi - days post-aphid inoculation.



B Systemic leaves

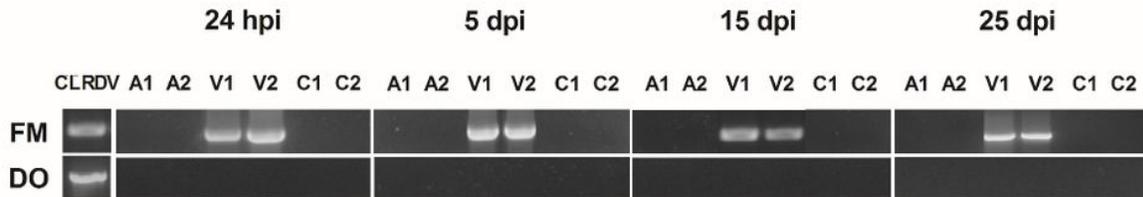


Figure 6

Expression profiles of cotton DCLs analyzed by RT-qPCR after CLRDRV virus infection. Cotton plants from the cv. Fibermax that is susceptible to CLRDRV infection and Delta Opal that is resistant to CLRDRV infection, were inoculated with CLRDRV vectors. Vectors were eliminated 24 hours after inoculation. A - Systemic leaves of Fibermax 966 and Delta Opal plants were collected at 24 hpi and 5, 15 and 25 dpi for the evaluation of each DCL level over time. The fold change was calculated using the $2^{-\Delta\Delta Ct}$ method as described by [75]. The standard deviation is indicated by the error bars, and "*" indicates significant differences, calculated using the GraphPad Prism program. * P < 0.1 ** P < 0.01 *** P < 0.001. hpi - hours

post-virus infection, dpi - days post-virus infection. B – A Nested RT-PCR for CLRDV capsid protein (CP) gene amplification was performed in all systemic leaves to check virus presence in virus-free aphid inoculated (A), CLRDV infected (V) and uninoculated plants (C). The number 1 and 2 represent independent pools of treated plants. FM – CLRDV susceptible plant and DO – CLRDV resistant plant.

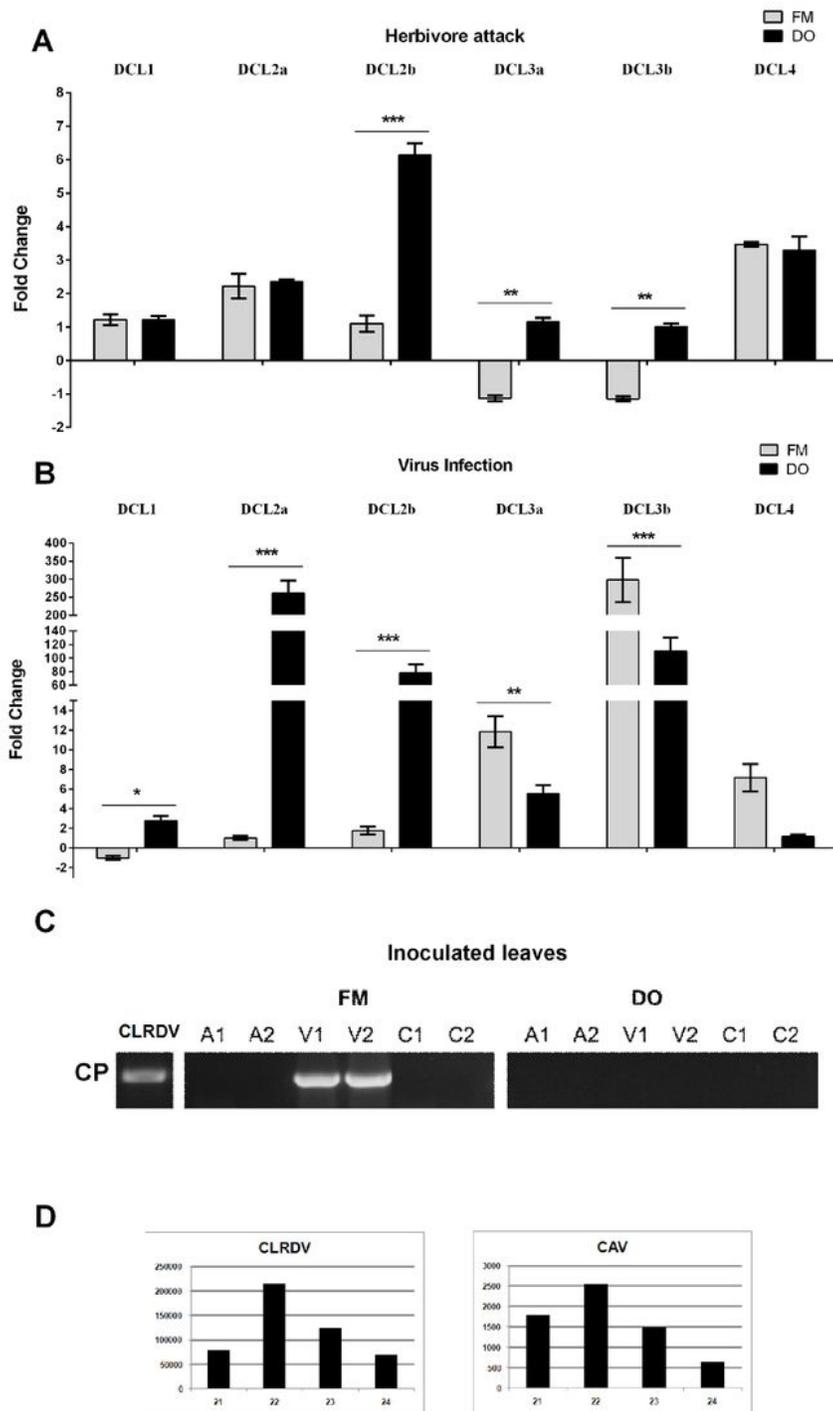


Figure 7

Profiles of cotton DCL expression at inoculation sites. A - DCL expression profiles in aphid-inoculated leaves at 24 hpi. B – Cotton DCL expression profiles at 24 hpi with CLRDV. C – Identification of CLRDV replication in inoculated leaves. Virus replication was assayed by CLRDV coat protein detection using nested RT-PCR in aphid inoculated (A), CLRDV infected and uninoculated plants (C). D – Profiles of viral small RNAs from cotton plants infected with distinct species of polerovirus. CLRDV sviRNA profile was obtained previously by [54] and is shown here just for illustrative purpose. Numbers 1 and 2 represent independent pools of treated plants. FM – CLRDV susceptible plant and DO – CLRDV resistant plant. CAV – Cotton anthocyanosis virus.

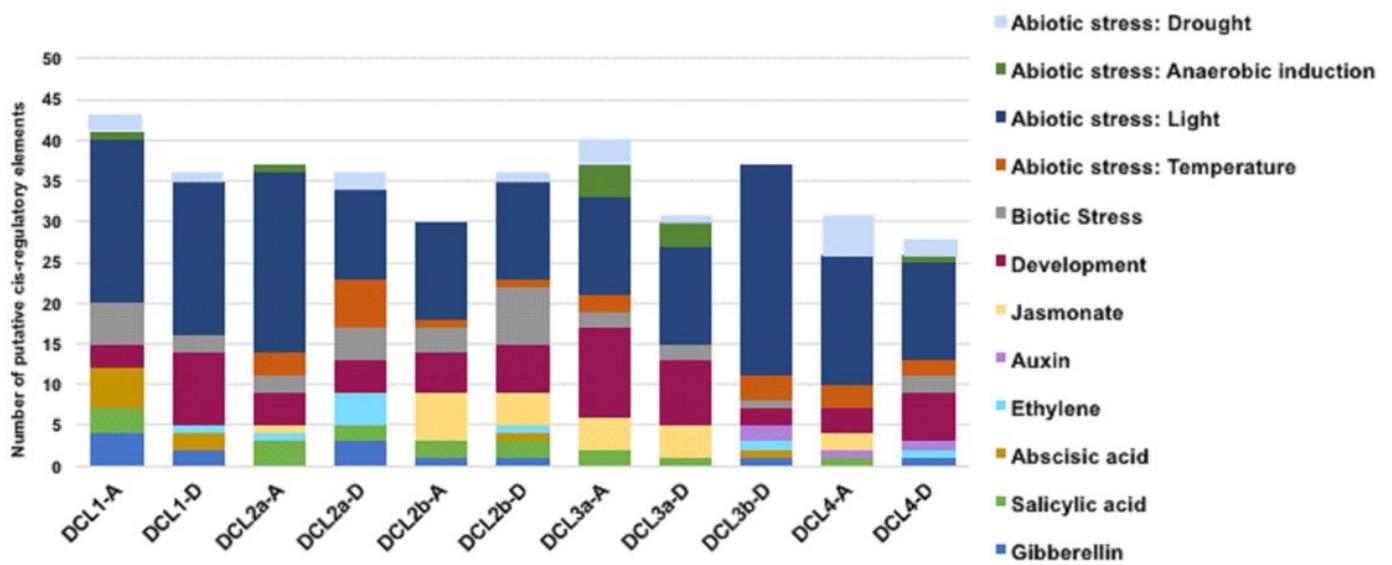


Figure 8

Analysis of *G. hirsutum* DCL promoter sequences. The 1.5-kb upstream promoter sequence of the transcription start site of each DCL gene was retrieved and analyzed for the presence of putative cis-elements. Different cis-elements with the same or similar functions are shown in the same color. Enhancer cis elements are shown in Additional File 3 Figure S3.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFile2.docx](#)
- [Additionalfile3.docx](#)
- [Additionalfile1.docx](#)
- [Additionalfile4.xlsx](#)