Genome-Wide Analysis of the Phospholipase D Family in Five Cotton Species, and Potential Role of GhPLD2 in Fiber Development and Anther Dehiscence

Changkai Ma†, Qian Zhang‡, Jiaoyan Lv*, Kaikai Qiao†, Shuli Fan*, Qifeng Ma* and Chaojun Zhang*

†State Key Laboratory of Cotton Biology, Institute of Cotton Research of CAAS, Key Laboratory of Cotton Genetic Improvement, Ministry of Agriculture, Anyang, China, ‡Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University, Zhengzhou, China

Phospholipase D (PLD) and its hydrolysis product phosphatidic acid play an important role in the regulation of several cellular processes, including root growth, pollen tube elongation, and microtubule reorganization. Here, we systematically identified and analyzed the membership, characterization, and evolutionary relationship of PLD genes in five species of cotton. The results of the transcriptomic analysis suggested that the evaluated PLD genes showed high expression levels in anther tissue and during the fiber initiation and elongation periods. Quantitative real-time polymerase chain reaction showed differential expression of GhPLD genes in the anthers of photoperiod sensitive male sterility mutant 5 (psm5). Previous research on multiple stable quantitative trait loci also suggests the role of PLD genes in the fiber development. Further analyses showed that GhPLD2 protein is localized to the plasma membrane. The virus-induced gene silencing of GhPLD2 in cotton seedlings repressed its expression by 40–70%, which led to a reduction in reactive oxygen species (ROS) levels, 22% anther indehiscence, and disrupted fiber initiation and elongation. Thus, we inferred that GhPLD2 may promote ROS production, which, in turn, may regulate anther dehiscence and fiber development.

Keywords: phospholipase D, cotton fiber, anther dehiscence, male sterility, GhPLD2

INTRODUCTION

Globally, cotton (Gossypium spp.) is among the most important cash crops for textiles, fine chemical raw materials, and strategically important supplies. Among the 50 cotton species, four are cultivated species, namely Gossypium hirsutum L. [(AD), 2n = 4x = 52], G. arboreum L. [(AD), 2n = 4x = 52], G. barbadense L. (A1, 2n = 4x = 26), and G. herbaceum L. (A1, 2n = 4x = 26; Wendel and Grover, 2015). Among them, upland cotton (G. hirsutum L.) accounts for more than 95% of cotton production owing to its excellent yield potential and adaptability. Examining the male sterile lines provides an important theoretical basis for understanding the breeding
of hybrid cotton, the yield of which is about 15% higher than that of conventional cotton, and has significantly improved fiber quality.

Successful anther dehiscence, or maturation of stamens, depends on several processes, including dilation of the endothecium layer, deposition of fibrous bands in endothecal cells, degradation of septal cells, and opening of the stomium (Keijzer, 1987). The process of cotton fiber initiation involves cell wall loosening due to increased turgor pressure in selected epidermal cells. Subsequently, the cotton fiber undergoes: (a) polarity elongation; (b) transitional wall thickening and primary wall remodeling; and (c) secondary wall thickening (Haigler et al., 2012). Cell division is a vital process for anther dehiscence and fiber development, and the previous evidence suggests that reactive oxygen species (ROS) homeostasis is critical for the maintenance of normal cell division in the root-tip cells of *Triticum turgidum* and *Arabidopsis thaliana* (Livanos et al., 2012b). ANOTHER DEHISCENCE REPRESSOR (ADR) suppresses ROS accumulation and endothecium thickening negatively regulates anther dehiscence by suppressing ROS accumulation and endothecium thickening in *A. thaliana* (Dai et al., 2019). Thus, phospholipase D (PLD), as an activator of ROS production, may play an important role in this process (Livanos et al., 2012a).

Phospholipase D is a lipid hydrolase of the phospholipase superfamily and is widely found in prokaryotes and eukaryotes. Based on the distinct protein domain identities found in the N-terminal with specific catalytic properties, the PLD family can be divided into the C2-PLD, PX/PH-PLD, and SP-PLD subfamilies. The C2 domain is a Ca<sup>2+</sup>- and phospholipid-binding domain; the PX and PH domains consist of the pheox homologue and the pleckstrin homologue, respectively, which are two distinct phosphoinositide-interacting structural folds; SP is a signal peptide (Wang, 2005; Li et al., 2007). Depending on gene structures, sequence similarities, conservative domains, and biochemical properties, PLDs in plants can be further classified into six classes in plants: α, β/γ, δ, ε, ζ, and φ. Among these, five classes (α, β/γ, δ, ε, ζ) have been identified in *Arabidopsis*, and PLD<sub>α</sub> is found in rice (Qin and Wang, 2002; Li et al., 2007). The phospholipid-hydrolyzing activities of PLD<sub>α</sub>, PLD<sub>β</sub>, PLD<sub>γ</sub>, PLD<sub>δ</sub>, and PLD<sub>ε</sub> are calcium-dependent (C2-PLD), while that of PLD<sub>ζ</sub> is calcium-independent (PX/PH-PLD). Unlike the C2 domain or PX and PH domains at the N-terminus, the PLD<sub>φ</sub> harbors a signal peptide and is therefore designated as SP-PLD.

Currently, PLD genes have been identified in various plant species, such as *A. thaliana* (Eliås et al., 2002; Qin and Wang, 2002), *Oryza sativa* L. (Li et al., 2007), *Populus alba* L., *Vitis vinifera* L. (Liu et al., 2010), *Glycine max* L. (Zhao et al., 2012), *Malus domestica* B., *Fragaia ananassa* Duchesne, *Prunus salicina* L. (Du et al., 2013), *Prunus persica* L. (Wan et al., 2019), *Zea mays* L. (Chen et al., 2017), *Brassica napus* L. (Lu et al., 2019), *G. arboretum* (Tang et al., 2016b), *G. hirsutum*, and *G. raimondii* (Tang et al., 2016a). In this study, we analyzed the PLD gene family in *G. barbadense* and *G. herbaceum* for the first time and re-analyzed them with new genomic data of *G. hirsutum*, *G. arboretum*, and *G. raimondii*. We also examined a newer male-sterile cotton mutant, *photoperiod sensitive male sterility mutant 5* (*psm5*) to investigate the potential function of the PLD gene in anther development. *psm5* is a filial generation of *psm4* and W10 having defective anther dehiscence, to discover the potential function of the PLD gene in anther development (Zhang, 2018). The functions of the candidate gene *GhPLD2* in fiber development and anther dehiscence were also investigated.

**MATERIALS AND METHODS**

**Plant Materials**

Upland cotton (*G. hirsutum cv. W10*) and *psm5*, generated by self-pollination of hybrid F<sub>1</sub> of *psm4* was crossed with W10 (Zhang, 2018), were cultivated under 16 h light/8 h darkness in a culture room at approximately 26–28°C. Samples of W10 and *psm5* anther were collected during two sampling periods: the binuclear period (flower bud length: 10 mm) and a day before the flowering (flower bud length: 30 mm). All of these samples were immediately frozen in liquid nitrogen and thereafter stored at −80°C until RNA extraction. Each sample comprised pooled material from 20 cotton plants. Three independent biological experiments were performed for each sample.

**Identification of PLD Gene Family in Cotton**

The PLD sequences of five *Gossypium* species were obtained from CottonFGD (Zhu et al., 2017)<sup>4</sup> and COTTONGEN (Yu et al., 2014)<sup>5</sup>. A hidden Markov model atlas of PLDc (PF00614) and PLD<sub>c</sub>_2 (PF13091) was used to find the PLD proteins in the *Gossypium* spp. database. Twelve AtPLD protein sequences, identified from 'TAIR',<sup>3</sup> were used as queries in the cotton genome database with the Basic Local Alignment Search Tool. In addition, the number of members of PLD gene family in five cotton species was checked by BLAST analysis of the reference genome sequence itself. Pfam,<sup>4</sup> and the simple modular architecture research tool,<sup>5</sup> and the National Center for Biotechnology Information conservative domain database<sup>6</sup> were used to verify the presence of two complete HKD domains in the putative PLD candidates. Proteins with incomplete domains were then manually removed. The theoretical molecular weight (Mw) and isoelectric point (pI) were computed by

---

<sup>1</sup>http://www.cottonfgd.org/  
<sup>2</sup>https://www.cottongen.org/  
<sup>3</sup>http://www.arabidopsis.org/  
<sup>4</sup>http://pfam.xfam.org/  
<sup>5</sup>http://smart.embl-heidelberg.de/  
<sup>6</sup>https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
ExPASy. The presence of a signal peptide was predicted using the CPHmodels 3.2 Server.

Bioinformatics Analyses
Multiple sequence alignment of the PLD protein sequences in *Gossypium* spp. and *Arabidopsis* genomes was performed with the E-INS-I algorithm of MAFFT v. 7.205 (Katoh and Standley, 2013). The best-fitting substitution model was estimated using AIC in jModelTest v. 2.1.2 (Darriba et al., 2012). Maximum likelihood (ML) and Bayesian inference (BI) were performed at the CIPRES Science Gateway (Miller et al., 2011) using RAxML 8.1.12 (Stamatakis et al., 2008) and MrBayes 3.2.6 (Ronquist et al., 2012), respectively. ML analysis included 1,000 bootstrap replicates under the LG + I + G model. BI analysis was made using the GTR + I + G model, analyses were run for 5 × 10^6 generations and Markov chains were sampled every 100 generations and 25% of the converged runs were regarded as burn-in.

The Multiple Collinearity Scan toolkit was used to ascertain the collinearity relation between the PLD members. The program KaKs_Calculator was used to calculate the nonsynonymous substitution rate (Ka) and synonymous substitution rate (Ks) of paralogous PLD genes. The model averaging method was used to calculate the values of Ka and Ks.

Co-localization of PLDs With Stable QTLs for Fiber Development
Quantitative trait loci (QTLs) and their corresponding molecular markers were investigated using COTTONGEN (see Footnote 2). The QTLs for fiber development-related traits were identified (Yu et al., 2014). To acquire the information on the physical position, first, the sequence of each marker was first retrieved from COTTONGEN. Second, the molecular marker sequences were aligned with the HAU genome of upland cotton using BLAST in the CottonFGD website (Zhu et al., 2017). Using Mapchart software, we constructed a diagram to display the location of PLD genes on chromosomes along with their adjacent loci and QTLs (Voorrips, 2002). As the cotton LD decay was approximately 0.80 Mb, gene locations pinpointing to the QTL or within 500 kb of a QTL were considered as anchored genes in QTL analysis.

Expression of *GhPLD* Genes
To investigate the potential biological functions of PLDs, the expression patterns of diverse *PLD* genes in upland cotton were examined based on the RNA-seq data from the NCBI Sequence Read Archive. The heat map of tissue-specific expression was plotted using the TBtools software (Chen et al., 2020a). Total RNA for quantitative real-time polymerase chain reaction (qPCR) analysis was extracted from cotton anthers using the RNAprep pure plant kit according to the manufacturer’s instructions (TIANGEN, Beijing, China). The PrimeScript® RT reagent kit (TaKaRa, Dalian, China) was used for the synthesis of cDNAs by reverse transcription to synthesize the cDNAs. The primer sequences for qPCR are listed in Supplementary Table S3. The qPCR data were processed with the 2^−ΔΔCT method (Livak and Schmittgen, 2001). Three biological replicates were performed, and mean ± SD values were calculated accordingly.

Subcellular Localization of *GhPLD2* Protein
The coding sequence of *GhPLD2* was cloned and ligated into the pCAMBIA-2300-35S-eGFP vector. Then, the fusion vector was transformed into *Agrobacterium tumefaciens* strain GV3101 and transiently expressed in tobacco (*Nicotiana benthamiana*) leaves. The injected tobacco leaves were cultured for 36–48 h. The eGFP fluorescence in the leaves was observed using a Dmi8 inverted microscope (Leica, Wetzlar, Germany), upon labeling the cell membranes using the membrane dye FM4-64 (4 mM; AAT Bioquest).

Virus-Induced *GhPLD2* Gene Silencing in Cotton
The recombinant vector pCLCrV::GhPLD2 was constructed with the pCLCrV vector and a 200bp fragment of *GhPLD2*. Then, the vectors containing the recombinant vector (pCLCrV::GhPLD), the positive control vector (pCLCrV::PDS), and empty vector (pCLCrVA) were separately transformed into *A. tumefaciens* strain GV3101, respectively. The transformed bacteria were collected by centrifugation and resuspended in osmotic buffer (10 mM MES, 10 mM MgCl_2, and 0.2 mM acetosyringone) with a final concentration of OD_600 = 1.5. Then, *A. tumefaciens* with pCLCrVB and *A. tumefaciens* with the pCLCrV::GhPLD2, pCLCrVA, and pCLCrV::PDS plasmids, respectively, were mixed in equal amounts and incubated at 28°C for 3 h; they were injected into cotyledons of W10 seedlings. Subsequently, all seedlings were transferred to a greenhouse under conditions of 16 h/8 h (light/dark) cycle at 25°C, for culture. Total RNA from the anthers of the gene-silenced cotton plants was extracted and processed for differences in gene expression analysis.

Determination of the Phospholipase D Activity and ROS Levels
The anthers of cotton were harvested and weighed. A specified amount of phosphate-buffered saline (pH = 7.4) was added to the samples and homogenized by grinders. The samples were centrifuged for 20 min at 14,000 rpm; solid material was removed, and the supernatant was carefully collected. PLD activity was determined using the Plant Phospholipase D ELISA Kit (MEIMIAN) according to the manufacturer’s instructions. ROS production rate was measured by Mitochondrial ROS production rate detection Kit (Comin, Suzhou, China), and ROS levels were assayed using the Plant ROS ELISA Kit (Nanjing JianCheng Bioengineering Institute, Nanjing, China) according to the manufacturer’s instructions. Three biological replicates were performed; the data for the corresponding anthers from at least two flowers were recorded from each identically treated plant.
In vitro Culture of Cotton Ovules Treated With n-Butanol and Tert-Butanol and Quantitation of Cotton Fiber

The ovaries from flowers of W10 plants, 1 day post-anthesis (DPA), were soaked in 75% ethanol for 2 min, and washed with 100% ethanol for 2 min. They were then dried in a sterile petri dish. Subsequently, the ovules were collected from the ovaries and were placed in a liquid BT medium with 5 μM indole-3-acetic acid and 0.5 μM gibberellic acid. Next, 0.2% n-butanol (1-But), tert-butanol (3-But), and sterile water (control) were added, respectively, for each treatment; they were cultivated in the dark at 30°C for 20 days. Fiber length and total fiber units (TFUs) were measured according to the methods described in previous studies (Beasley and Ting, 1974; Gong et al., 2014).

RESULTS

Molecular Structure Analysis of PLD Genes in Gossypium

In this study, 40, 40, 21, 21, and 20 PLD genes were identified in G. hirsutum [(AD), 2n=4x=52], G. barbadense [(AD), 2n=4x=52], G. herbaceum (A1, 2n=4x=26), G. arboreum (A2, 2n=4x=26), and G. raimondii (D3, 2n=2x=26), respectively. A total of 142 PLD proteins were noted according to their homologous positions. Gene names, gene IDs, chromosomal locations, amino acid amounts, molecular weights, pI, subcellular localization, and signal peptide predictions are summarized in Supplementary Table S1. To evaluate the evolutionary relationships between the PLD genes in Gossypium, a phylogenetic tree based on maximum likelihood was constructed with PLD proteins from A. thaliana and the five cotton species. The phylogenetic analysis showed that 154 PLDs were divided into six monophyletic clades: α, β/γ, δ, ε, ζ, and ϕ (Figure 1). Among the six clades, the largest group was δ comprising 36 members including 10 GhPLDs, 10 GbPLDs, five GhePLDs, five GaPLDs, five GrPLDs, and one AtPLDδ; group ε was the smallest comprising eight members, including two GhPLDs, two GbPLDs, one GhePLDs, one GaPLDs, one GrPLDs, and one AtPLDε. The details of the group compositions are listed in Supplementary Table S1. The predicted pI values of the PLDs ranged from 5.32 to 9.38. All of the PLDs in the ϕ clade had the signal peptide.

Patterns of intron-position correspondence between widely diverged eukaryotic species have provided valuable information into the evolutionary relationships. We analyzed the exon-intron structures of different PLDs in A. thaliana and five cotton species (Supplementary Figure S1). The intron numbers and lengths highly varied among PLD members. But most PLD members within the same subgroups shared similar intron/exon numbers. The result suggested PLD member are conserved in different species during evolution.

Evolution of PLD Genes in Gossypium

To decipher the evolutionary relationships of PLD genes, interspecific synten was analyzed for the five cotton species (Supplementary Figure S2). Eighteen collinear gene pairs were obtained for the D-subgenomes of two allotetraploid cotton species and G. raimondii. The collinearity between the A-subgenomes of the two allotetraploid cotton species and G. herbaceum and G. arboreum were also analyzed. We found that there were 16 collinear gene pairs between two allotetraploid cotton species and G. herbaceum; there were 17 collinear gene pairs between the two allotetraploid cotton species and G. arboreum. Furthermore, there were 18 collinear gene pairs between G. hirsutum and G. barbadense for both the A-subgenomes and D-subgenomes (Supplementary Figure S2). These results suggested high collinearity in allotetraploid cotton species and their corresponding A- and D-subgenomes in diploid cotton; many PLD genes in G. raimondii and the D-subgenomes of the two allotetraploid cotton species were evolutorily closely related.

To examine the duplication events of PLD genes in cotton, the tandem, and segmental duplications were identified. A total of 28, 28, 4, 3, and 4 segmental gene pairs were identified in G. hirsutum, G. barbadense, G. herbaceum, G. arboreum, and G. raimondii, respectively; four pairs were ascribed to tandem duplication (containing 2, 2, and 1 tandem gene pairs in G. hirsutum, G. barbadense, and G. arboreum, respectively; Supplementary Table S2). Ka/Ks ratios were calculated to measure the selection pressure on these duplicated gene pairs. Among the five cotton species, 70 of 71 duplicated pairs of PLD genes had Ka/Ks ratios <1. This implied that the PLD gene family had undergone intense purifying selection pressure in the course of evolution.

Tissue-Level Expression of GhPLD Genes

To further understand the latent developmental roles of PLD genes in G. hirsutum, we analyzed the published transcriptomic data for the expression of PLD genes in diverse cotton tissues (Supplementary Figure S3). The results showed that all GhPLD genes had diverse expression patterns in tissues based on the different developmental stages. Two GhPLDα genes (GhPLD13 and GhPLD33) showed higher expression levels in the pistil, bract, ovule, and fiber, while other GhPLDα isotypes were expressed at lower levels in these tissues. Most GhPLDδ genes including GhPLD2, GhPLD5, GhPLD19, GhPLD22, GhPLD24, and GhPLD39 showed selectively higher expression in anther and filament and analogously also had high expression in fiber during initiation (−3 to 0 DPA; Supplementary Figure S3). These results indicated that these GhPLDδ isotypes may be involved in anther and fiber development.

Differential Expression Between psm5 and W10

Based on the analysis of transcriptomic sequencing data in our laboratory (unpublished data), we found that some members of the PLD gene family had significant differential expression between psm5 and W10 (Supplementary Figure S4A); nine of these genes were significantly differentially expressed in anthers of psm5 on the day before flowering (B), while the transcripts of only two GhPLD genes (GhPLD12, GhPLD18) had a significant
difference in the binuclear period (A), as compared to the WT (Supplementary Figure S4A and Supplementary Table S4). We speculated that GhPLD genes (GhPLD2, GhPLD7, GhPLD9, GhPLD15, GhPLD18, GhPLD22, GhPLD29, GhPLD36, GhPLD38) may play an essential role in the later stages of anther development.

To confirm the differential expression of GhPLD genes in the psm5, we analyzed and compared the levels of expression of 20 randomly selected GhPLD genes in the psm5 and W10 (WT) using quantitative real-time PCR (qPCR). The results of qPCR indicated that the expression of 10 genes was significantly down-regulated in anthers of psm5 in the B period as compared to WT; the expression of two genes was significantly up-regulated, while the transcripts of most GhPLDs showed no significant differences in the A period (Supplementary Figure S5). Consistent with transcriptomic data, we validated the potential role of
the nine GhPLD genes (GhPLD2, GhPLD7, GhPLD9, GhPLD15, GhPLD18, GhPLD22, GhPLD29, GhPLD36, and GhPLD38) in the regulation of anther dehiscence in cotton.

**Co-localization of GhPLDs With QTL for Fiber Development**

Fiber quality traits and yield are quantitative and controlled by numerous QTLs (Jiang et al., 1998). The potential functions of GhPLDs were further examined by mapping the GhPLDs onto the previously reported QTL related to fiber length (FL), fiber elongation (FE), and fiber strength (FS). Four QTL including qFL-Chr5-5 (Ma et al., 2017), qFL.DjIM-RIL_ch08.h11, qFL.GG-F2_ch10.cz081, and qFS.GG-RIL_ch21.cz12 were found to co-localize with GhPLD6 (PLDα), GhPLD12 (PLDβ/γ), GhPLD14 (PLDζ), and GhPLD36 (PLDβ/γ), respectively. Four PLD genes were mapped in the proximity of the QTLs for fiber length, fiber elongation, and fiber strength (Figure 2). Thus, we provided evidence for the potential function of GhPLD genes in the fiber development.

**GhPLD2 Protein Is Localized to the Plasma Membrane**

Based on the significantly higher expression of the GhPLD2 gene in anther tissue and its most significant differential expression between the psm5 and W10, we further examined its function during anther dehiscence. The subcellular localization of the GhPLD2 protein was determined using a construct that transiently expressed GhPLD2-enhanced green fluorescent fusion protein (GhPLD2::eGFP) in tobacco leaves. Fluorescence microscopy showed that the empty vector constructs expressed eGFP protein which was localized in the nucleus, cytoplasm, and cell membrane, while the fluorescence signal from the expression of the GhPLD2::eGFP fusion vector was specifically localized to the plasma membrane (Figure 3). We further

![FIGURE 2](image-url) | Distribution of co-localized GhPLDs on the A05 (A), A08 (B), A10 (C), and D11 (D) chromosomes of G. hirsutum. The scale indicates the position of PLD genes and quantitative trait locus (QTL)-linked markers in the chromosome. The unit is megabase (Mb). QTLs related to fiber length (FL), fiber elongation (FE), and fiber strength (FS) are shown. GhPLD genes that co-localize with QTLs of different fiber traits are indicated in red.
validated these results using the lipid dye FM4-64 which labels the cell membranes (Supplementary Figure S6). Therefore, GhPLD2 protein was a membrane-localized protein.

**GhPLD2 May Mediate in Anther Dehiscence**

Male sterility is a hotspot in plants research. In previous studies, the PLD genes were shown to have possible involvement in the male sterility by regulating anther dehiscence (Qiu et al., 2007). To verify this possibility for the PLD gene in cotton, virus-induced gene silencing (VIGS) assay was used to silence GhPLD2 expression. Due to GhPLD2 downregulation, anther indehiscence occurred in CLCrV::GhPLD2 plants (Figure 4A-N). To further determine the GhPLD2 gene silencing efficiency, qPCR was used to assess the quantitative expression of GhPLD2 in both the CLCrV::GhPLD2 and CLCrV::00 cotton plants. Notably, GhPLD2 expression was significantly reduced in CLCrV::GhPLD2 plants as compared to the CLCrV::00 control plants (Figure 4E). PLD activity in the CLCrV::GhPLD2 plants was also significantly downregulated (Figure 4J). In addition, ROS levels in the anther were also significantly reduced along with a decrease in phospholipase activity (Figure 4O), consistent with late anther development in *psm5* as compared to W10 (Supplementary Figure S4B). These findings suggested that the GhPLD2 gene may regulate the development of anther dehiscence through its effects on PLD activity and ROS levels.

**GhPLDs May Regulate Fiber Initiation**

Phospholipase D function has been shown to underlie fiber elongation (Tang and Liu, 2017). To confirm the likely involvement of GhPLD2 in the regulation of fiber initiation, the protodermal cells of the cotton ovule on the day of anthesis were observed using a scanning electron microscope. We found that ballooning out of protodermal cells was delayed in CLCrV::GhPLD2 plants (Figures 5A,B). Subsequently, mature fiber length was measured in CLCrV::GhPLD2 and CLCrV::00 plants (Figures 5C,J); the mature fiber length of CLCrV::GhPLD2 was significantly lower than that for the CLCrV::00 plant. These results suggested that GhPLDs may play a novel role in fiber initiation in addition to regulation of fiber strength (Tang and Liu, 2017).

Phospholipase D regulates plant growth and development through the generation of PA as a lipid messenger. To evaluate the relationship between PLD, microtubules, and cotton fiber development, the phenotypic observation and measurement of the fiber length and TFUs using *in vitro* cotton ovule cultures were performed at 20 DPA (Figures 5D–I, K, L). The results suggested that the length and TFUs of fiber were disposed by 1-butanol, a specific inhibitor of PA; there was no difference with 3-butanol treatment, as it cannot activate PLD. These results further showed that GhPLD genes may regulate microtubule stability by altering fiber initiation and elongation.

**DISCUSSION**

**PLD Family in Cotton Expanded During Evolution**

We found that the number of PLD genes varied among different cotton species, with 40, 40, 21, 21, and 20 genes expressed in *G. hirsutum*, *G. barbadense*, *G. herbaceum*, *G. arboreum*, and *G. raimondii*, respectively. The variability of member of PLD gene in *G. arboreum* was compared with previous reports...
(Tang et al., 2016b), and this phenomenon may be attributed to the difference reference genomes. We have used the latest genome assemblies for analyses of the PLD family in *G. arboreum* genomes (Huang et al., 2020). Therefore, the new databases undoubtedly facilitated accuracy in the identification of PLD gene family in cotton.

Polyplody is a feature of genome evolution in allotetraploid cotton that doubled the size of the PLD family due to whole-genome duplication (WGD) and segmental duplication. Segmental duplication events result in a wide dispersion of gene copies (BaumGaten et al., 2003). For example, a total of 12 PLD genes one PLDδ gene are found in *Arabidopsis*, while the number of PLDδ genes in cotton is much higher. There are ten PLDδ in allotetraploid cotton species, and five PLDδ genes in diploid cotton species. Compared with *A. thaliana*, we found that the expansion of the cotton δ branch in the cotton genome was higher. This phenomenon occurs in other species including *G. max* (Zhao et al., 2012), *B. napus* (Lu et al., 2019), *P. alba*, and *V. vinifera* (Liu et al., 2010). Correspondingly, there is several segmental gene pairs in the δ clade of the PLD family in cotton (Supplementary Table S2). Gene duplication may have played an important role in the evolution of phenotypic

![FIGURE 4](image-url)

FIGURE 4 | Silencing of GhPLD2-induced anther dehiscence defect and reduction in reactive oxygen species (ROS) content in upland cotton. (A–D) The cotton plant, flower, and anthers of CLCrV::00 are shown, respectively. (E–I) The cotton plant, flower, and anthers of CLCrV::PDS are shown, respectively. (K–N) The cotton plant, flower, and anthers of CLCrV::GhPLD2 are shown, respectively. (E) Expression levels of GhPLD2 in CLCrV::00 and CLCrV::GhPLD2 plants. GhHistone 3 was used as the reference gene for normalization. (J) Phospholipase D activity in CLCrV::00 and CLCrV::GhPLD2 plants. (O) ROS levels in CLCrV::00 and CLCrV::GhPLD2 plants. Error bars indicate the estimate of standard deviation in three independent replicates. Statistical analyses of experimental data were performed by t-tests. Asterisks showed that the value of p indicates significant (*p < 0.05) or very significant differences (**p < 0.01) as compared to that of the control. (A,F,K) scale bar = 8 cm; (B,C,G,H,L,M) scale bar = 1 cm; (D,I,N) scale bar = 1.5 mm.
novelty within plants. This tendency indicates the specific evolution of the PLDδ branch and expansion of PLDδ isoforms which is necessary for adaptive evolution of plants. This also indicates that segmental duplication might have contributed to the expansion of the cotton PLD gene family. Similarly, the occurrence of tandem replication events has facilitated the expansion of the gene family (Cannon et al., 2004; Leister, 2004). Here, four important tandem duplications were found only in *G. hirsutum*, *G. barbadense*, and *G. arboreum*, which indeed supported the above point of view. In conclusion, our results confirmed that segmental replication and WGD were the main driving forces for the amplification of the PLD gene family in cotton.

**GhPLD Genes May Mediate Anther Dehiscence in Cotton**

Male sterility is a common phenomenon in flowering plant species, occurring in several crops (Liu et al., 2014; Li et al., 2016; Chen et al., 2020b; Saxena et al., 2020). Flowering plants require normal anther dehiscence to ensure timely pollen release at fruit-setting stages. The tissue-specific expression analysis of PLD genes showed that most PLDs, especially *GhPLD*δs, had higher expression in the anther and filament (Supplementary Figure S4), consistent with previous studies (Tang et al., 2016a). Furthermore, a recent study also suggests that the PLD signaling pathway is mostly enriched in the late developmental stage of the anther in cotton (Chen et al., 2018). A new photosensitive cotton mutant, *psm5*, exhibits male sterility resulting from the absence of anther dehiscence in long-day conditions. qPCR analyses were performed, and the results independently confirmed the differential expression of PLD genes in the late stage of anther development in *psm5* as compared to the WT. Among these differentially expressed genes, *GhPLDs* belonging to the δ clade accounted for the majority proportion. *GhPLD13* (α) and *GhPLD39* (δ) were upregulated in *psm5*, while the 10 other *GhPLDs* were downregulated (Figure 5). The results indicated that *GhPLDδs* may be involved in anther dehiscence, and may underlie male infertility in *psm5*. In this study, *GhPLD2* was chosen for subsequent functional analyses due to its high expression in anther tissue and significant differential expression in *psm5*. Silencing the gene expression of *GhPLD2* by VIGS further validated that *GhPLD2* was indeed involved in anther dehiscence and was a positive regulator. Although the efficiency of silencing was only about 40–70%, and the flowers with non-dehiscent anthers accounted for approximately 22%. This implied that even a partial loss of *GhPLD2* gene function could have a severe impact on plant development.

**FIGURE 5 |** Inhibition of PLD activity affects fiber elongation, and GhPLD2 is involved in fiber initiation and elongation. (A–C) Silencing GhPLD2 inhibited fiber initiation (A, ovule of blank control, B, ovule of VIGS plant) and significantly reduced fiber length (C,J) scale bar = 1 mm, scale bar = 10 μm. (D–I) Seed and fiber morphological characteristics in the *in vitro* culture after treatment with 1-But (F,I), 3-But (E,H), and sterile water (CK; D,G); scale bar = 1 cm. (K) Fiber length in 1-But and 3-But treated cultures. (L) Total fiber units (TFUs) in 1-But and 3-But treated cultures. Error bars show the standard deviation estimated using three independent experimental repeats. Statistical analyses of experimental data were performed using t-tests. Asterisks show that the values of *p* indicate very significant (**p < 0.01) differences as compared to the control.
Furthermore, the ROS levels in anther tissues significantly decreased along with the reduction in PLD activity in CLCrV::GhPLD2 cotton plants. ROS production is catalyzed by NADPH oxidase (Livanos et al., 2012a); AtPLDα1 and phosphatidic acid (PA) increase the activity of NADPH oxidase and ROS production (Zhang et al., 2009). Based on a comprehensive analysis of these results, we speculated that GhPLD2 regulates NADPH oxidase activity to promote ROS production, and the silencing of GhPLD2 in CLCrV::GhPLD2 may alter ROS accumulation to induced anther dehiscence.

**GhPLD Genes May Control Fiber Growth in Cotton**

In this study, the identification of QTLs associated with GhPLDαs, GhPLDβ/γs, and GhPLDζs provided strong evidence for the role of PLD genes in the fiber development. A previous study reported PLDα protein expression in an elongating cotton fiber sample (Yang et al., 2008). The fiber elongation role of GhPLDα1 in the α branch has been reported for upland cotton (Tang and Liu, 2017). Here, tissue-specific expression analysis suggests that GhPLD2, GhPLD5, GhPLD19, GhPLD24, and GhPLD38 genes, belonging to the δ branch, could not only regulate fiber elongation but may also be involved in fiber initiation.

Phospholipase D regulates plant growth and development through the generation of a lipid messenger, PA. 1-butanol, a specific inhibitor of PLD-dependent production of the signaling molecule PA, disrupts the organization of interphase cortical microtubules (Gardiner et al., 2003); 3-butanol can be used as a control for any non-specific butanol effect as it has no role in PLD activation (Munnik et al., 1995). The results of in vitro cotton ovules cultures showed that the length and TFUs of fiber dealt with 1-butanol were significantly inhibited, and this may be attributed to the effects of 1-butanol on the conformation of microtubules in fibroblasts. PLD serves as a dynamic membrane linker for microtubules, and its activation induces microtubule depolymerization and rearrangement, which help cells growth adjustment, division, and stimulus-response requirements owing to its different conformations (Whittaker and Triplett, 1999; Dhoukhshe and Gadella, 2003; Li et al., 2004; Zhang et al., 2017). Taken together, the results of the VIGS experiment, it was suggested that GhPLD genes may regulate microtubule stability, and thus impact both fiber initial and elongation (Supplementary Figure S7).

**CONCLUSION**

In conclusion, a total of 40, 40, 21, 21, and 20 PLD genes were identified in G. hirsutum, G. barbadense, G. herbaceum, G. arboreum, and G. raimondii, respectively. Their phylogenetic classifications, gene structures, motifs, and gene duplications were analyzed. The PLD genes in the same evolutionary branches were found to share similar motifs and gene structures. Our analysis indicated that WGD and segmental duplication events were the main driving forces underlying PLD gene amplification and that purifying selection played a major role in the evolution of the PLD gene family.

Based on the transcriptomic analysis, higher expression of GhPLDδ genes including GhPLD2, GhPLD5, GhPLD19, GhPLD24, and GhPLD39 was found during the fiber initiation (~3 to 3 DPA) and elongation (5–20 DPA) periods. VIGS experiments showed that GhPLD2 had a positive impact on fiber initiation and elongation. qPCR results indicated that differential expression of some GhPLD genes may underlie male sterility in cotton. Silencing of the GhPLD2 gene assays indicated that GhPLD2 may be involved in ROS production by NADPH oxidase, and thus may positively regulate anther dehiscence. In conclusion, our study provides functional insights into the roles of the PLD gene family in male sterility and fiber development in cotton.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

**AUTHOR CONTRIBUTIONS**

CM: conceptualization, methodology, software, and writing. QZ: methodology and data curation. JL and KQ: software and visualization. SF: visualization and investigation. QM and CZ: reviewing and editing. All authors contributed to the article and approved the submitted version.

**FUNDING**

This work was supported by the Major Research Plan of the National Natural Science Foundation of China (Grant no. 31690093), the National Natural Science Foundation of China (Grant no. 31701474), and Agricultural Science and Technology Innovation ProGrm of Chinese Academy of Agricultural Sciences.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.728025/full#supplementary-material
diploid cotton species (G. herbaceum and G. arboreum). (B) Collinear analysis for D-genome of allotetraploid cotton species and G. raimondii. Orange lines depict the collinearity of the PLDs between the genomes of two allotetraploid cotton species; green lines depict the collinearity of the PLDs for the genomes of allotetraploid cotton and diploid cotton; gray lines in the background depict the collinear blocks within the different genomes, while other colored lines highlight the syntenic PLD gene pairs.

**Supplementary Figure S3** | The expression patterns of the GnPPLDs according to a common database in different tissues and organs and during fiber development. (A) The expression patterns of the GnPPLDs in different tissues including anther, filament, pistil, bract, sepal, petal, torus, root, leaf, and stem. (B) The expression profiles of the GnPPLDs at −3, 0, 1, 3, 5, 10, 15, 20, and 25 day post-anthesis (DPA) ovules and 10, 15, 20, and 25 DPA fibers.

**Supplementary Figure S4** | (A) The expression profiles of GnPPLD genes in W10 (WT) and psm5 based on RNA-seq data. Transcriptome data are summarized in **Supplementary Table S4**. Data are presented as log2 fold change. (B) Mitochondrial reactive oxygen species (ROS) production rate in anthers of psm5 and W10. The first stage A is a binuclear phase of the microspore, the early flowering stage (flower bud length, ~10 mm). The second stage B is 1 day before flowering. Each experiment was performed with three biological repeats. The values are presented as the mean ± SDs of the three replicates. Value of ρ was calculated based on the Student’s t-tests.

**Supplementary Figure S5** | Differential expression profiles of the GnPPLDs in W10 (WT) and psm5. The first stage A is a binuclear period of the microspore, the early flowering stage (flower bud length, ~10 mm). The second stage B is 1 day before flowering. The values are presented as the mean ± SDs of three replicates. Values of ρ were calculated using the Student’s t-tests. ** denotes significance at p < 0.01 as compared to the WT. The details of the primer sequences are listed in **Supplementary Table S3**.

**Supplementary Figure S6** | Re-validation of subcellular localization of GhPLD2 in tobacco (Nicotiana benthamiana) leaves. The results in **Figure 3** were re-confirmed using the lipid dye FM4-64 (Marker) which labels the cell membranes. Scale bar = 50 μm.

**Supplementary Figure S7** | Model of GnPPLD2 regulating anther dehiscence and fiber development in cotton. To visually present the function of the GnPPLD2 gene, a model depicting the mechanisms underlying GnPPLD2 gene regulation of anther dehiscence, fiber initiation, and fiber elongation was constructed. GnPPLD2 positively promotes ROS production to maintain normal anther dehiscence and fiber development.

**References**

BaumGaten, A., Cannon, S., Spangler, R., and May, G. (2003). Genome-level evolution of resistance genes in Arabidopsis thaliana. Genetics 165, 309–319. doi: 10.1093/genetics/165.1.309

Beasley, C., and Ting, I. P. (1974). Effects of plant growth substances on in vitro fiber development from unfertilized cotton ovules. Am. J. Bot. 61, 188–194. doi: 10.1002/1529-8547(197406)61:3<197::AID-AMJ2>3.0.CO;2-E

Cannon, S. B., Mitra, A., BaumGaten, A., Young, N. D., and May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. BMC Plant Biol. 4:10. doi: 10.1186/1471-2229-4-10

Chen, L., Cao, B., Han, N., Tao, Y., Zhou, S. F., Li, W. C., et al. (2017). Phospholipase D family and its expression in response to abiotic stress in maize. Plant Growth Regul. 81, 197–207. doi: 10.1007/s10725-016-0197-4

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020a). TBoeasl-an inteGrtive toolkit developed for interactive analyses of big biological data. Mol. Plant 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, J., Su, P., Chen, P., Li, Q., Yuan, X., and Liu, Z. (2018). Insights into the cotton anther development through association analysis of transcriptomic and small RNA sequencing. BMC Plant Biol. 18:154. doi: 10.1186/s12870-017-1367-4

Chen, J., Xu, Y., Fei, K., Wang, R., He, J., Fu, L., et al. (2020b). Physiological mechanism underlying the effect of high temperature during anthesis on spikelet-opening of photo-thermo-sensitive genic male sterile rice lines. Sci. Rep. 10:2210. doi: 10.1038/s41598-020-59183-0

Dai, S. Y., Hsu, W. H., and Yang, C. H. (2019). The gene ANOTHER DEHISCENCE REPRESSOR (ADR) controls male fertility by suppressing the ROS accumulation and anther wall thickening in Arabidopsis. Sci. Rep. 9:5112. doi: 10.1038/s41598-019-4382-x

Darriba, D., Taboada, G. L., Doallo, R., and Posada, D. (2012). jModelTest 2: a powerful single-cell model for cell wall and cellulose research. Front. Plant Sci. 3:104. doi: 10.3389/fpls.2012.00104

Gardiner, J., Collings, D., Harper, J., and Marc, J. (2003). The effects of the phospholipase D-antagonist 1-butanol on seedling development and microtubule organisation in Arabidopsis. Plant Cell Physiol. 44, 687–696. doi: 10.1093/pcp/pcg095

Hong, G., Wu, Z., Percy, R. G., Bai, M. Z., Li, Y., Frelichowski, J. E., et al. (2020). Genome sequence of Gossypium herbaceum and genome updates of Gossypium arboresum and Gossypium hirsutum provide insights into cotton A-genome evolution. Nat. Genet. 52, 516–524. doi: 10.1038/s41588-020-0606-7

Jiang, C. X., Wright, R. J., El-Zik, K. M., and Patersonet, A. H. (1998). Polyploid formation created unique avenues for response to selection in Gossypium (cotton). Proc. Natl. Acad. Sci. U. S. A. 95, 4419–4424. doi: 10.1073/pnas.95.8.4419

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. doi: 10.1093/molbev/mso100

Keijzer, C. (1987). The processes of anther dehiscence and pollen dispersal: I. The opening mechanism of longitudinally dehiscent anthers. New Phytol. 105, 487–498. doi: 10.1111/j.1469-8137.1987.tb00886.x

Leister, D. (2004). Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. Trends Genet. 20, 116–122. doi: 10.1016/j.tig.2004.01.007

Li, W., Li, M., Zhang, W., Welti, R., and Wang, X. (2004). The plasma membrane–bound phospholipase D8 enhances freezing tolerance in Arabidopsis thaliana. Nat. Biotechnol. 22, 427–433. doi: 10.1038/nb949

Li, G., Lin, F., and Xue, H. W. (2007). Genome-wide analysis of the phospholipase D gene family in three Rosaceae species. J. Sertica. 153, 13–21. doi: 10.1016/j.j.sclenta.2013.01.002

Eliáš, M., Potocký, M., Cvrčková, F., and Žársky, V. (2002). Molecular diversity of phospholipase D in angiosperms. BMC Genomics 3:2. doi: 10.1186/1471-2164-3-2

Ma et al. Genome-Wide Analysis of PLD Genes

Frontiers in Plant Science | www.frontiersin.org

October 2021 | Volume 12 | Article 728025

11
Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using realtime quantitative PCR and the 2^{−ΔΔCT} method. Methods 25, 402–408. doi: 10.1016/S1046-8023(01)01262
Livano, P., Apostolakos, P., and Galatis, B. (2012a). Plant cell division: ROS homeostasis is required. Plant Signal. Behav. 7, 771–778. doi: 10.4161/psb.20530
Livano, P., Galatis, B., Quader, H., and Apostolakos, P. (2012b). Disturbance of reactive oxygen species homeostasis induces atypical tubulin polymer formation and affects mitosis in root-tip cells of Triticum turgidum and Arabidopsis thaliana. Cytoskeleton 69, 1–21. doi: 10.1002/cm.20538
Lu, S., Fadalla, T., Tang, S., Li, L., Ali, U., Li, Q., et al. (2019). Genome-wide analysis of phospholipase D gene family and profiling of phospholipids under abiotic stresses in Brassica napus. Plant Cell Physiol. 60, 1556–1566. doi: 10.1093/pcp/pcz071
Ma, L., Zhao, Y., Wang, Y., Shang, L., and Hua, J. (2017). QTLs analysis and validation for fiber quality traits using maternal backcross population in upland cotton. Front. Plant Sci. 8:2168. doi: 10.3389/fpls.2017.02168
Miller, M. A., Pfeiffer, W., and Schwartz, T. (2011). The CIPRES science gateway: a platform for phylogenetic analyses. Proceedings of the 2011 TeraGrid Conference: Extreme Digital Discovery 1, 1–8. doi: 10.1145/2016741.2016785
Munnik, T., Artisz, S. A., de Vrie, T., and Musgrave, A. (1995). G protein activation stimulates phospholipase D signaling in plants. Plant Cell 7, 2197–2210. doi: 10.1007/bf02370162
Qin, C., and Wang, X. (2002). The Arabidopsis phospholipase D family. Characterization of a calcium-independent and phosphatidylcholine-selective PLDζ1 with distinct regulatory domains. Plant Physiol. 128, 1057–1068. doi: 10.1104/pp.010928
Qiu, X., Gao, Q., Wang, D., Hao, Y., Wang, Y., and Jiao, J. (2007). Preliminary study on morphologic and genetics of male sterile mutant in transformed wheat with anti—PLDy gene. Mai Lei Zuo Wu Xue Bao 27, 207–210.
Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542. doi: 10.1093/sysbio/sys029
Saxena, S., Sahu, S., Kaila, T., Nigam, D., Chaudhuri, P. K., Rao, A., et al. (2020). Transcriptome profiling of differentially expressed genes in cytoplasmic male-sterile line and its fertility restorer line in pigeon pea (Cajanus cajan L.). BMC Plant Biol. 20:74. doi: 10.1186/s12870-020-2284-y
Stamatakis, A., Hoover, P., and Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML web servers. Syst. Biol. 57, 758–771. doi: 10.1080/10635150802429642
Tang, K., Dong, C. J., and Liu, J. Y. (2016a). Genome-wide comparative analysis of the phospholipase D gene families among allotetraploid cotton and its diploid progenitors. PLoS One 11:e0156281. doi: 10.1371/journal.pone.0156281
Tang, K., Dong, C. J., and Liu, J. Y. (2016b). Genome-wide analysis and expression profiling of the phospholipase D gene family in Gossypium arboreum. Sci. China Life Sci. 59, 130–141. doi: 10.1007/s11427-015-4916-2
Tang, K., and Liu, J. Y. (2017). Molecular characterization of GhPLDδ1 and its relationship with secondary cell wall thickening in cotton fibers. Acta Biochim. Biophys. Sin. 49, 33–43. doi: 10.1093/abbs/gmw113
Voorrips, R. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. J. Hered. 93, 73–78. doi: 10.1093/hered/93.1.77

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Ma, Zhang, Lv, Qiao, Fan, Ma and Zhang, This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.