Comprehensive characterization and gene expression patterns of *LBD* gene family in *Gossypium*

Jingwen Yu¹ · Qianwen Xie¹ · Cheng Li¹ · Yating Dong¹ · Shuijin Zhu¹ · Jinhong Chen¹

Received: 8 December 2019 / Accepted: 13 February 2020 / Published online: 17 March 2020
© The Author(s) 2020

Abstract

Main conclusion A comprehensive account of the *LBD* gene family of *Gossypium* was provided in this work. Expression analysis and functional characterization revealed that *LBD* genes might play different roles in *G. hirsutum* and *G. barbadense*.

Abstract The Lateral Organ Boundaries Domain (LBD) proteins comprise a plant-specific transcription factor family, which plays crucial roles in physiological processes of plant growth, development, and stress tolerance. In the present work, a systematical analysis of *LBD* gene family from two allotetraploid cotton species, *G. hirsutum* and *G. barbadense*, together with their genomic donor species, *G. arboreum* and *G. raimondii*, was conducted. There were 131, 128, 62, and 68 *LBD*s identified in *G. hirsutum*, *G. barbadense*, *G. arboreum* and *G. raimondii*, respectively. The LBD proteins could be classified into two main classes, class I and class II, based on the structure of their lateral organ boundaries domain and traits of phylogenetic tree, and class I was further divided into five subgroups. The gene structure and motif composition analyses conducted in both *G. hirsutum* and *G. barbadense* revealed that *LBD* genes kept relatively conserved within the subfamilies. Synteny analysis suggested that segmental duplication acted as an important mechanism in expansion of the cotton *LBD* gene family. Cis-element analysis predicated the possible functions of *LBD* genes. Public RNA-seq data were investigated to analyze the expression patterns of cotton *LBD* genes in various tissues as well as gene expression under abiotic stress treatments. Furthermore, RT-qPCR results found that *GhLBD*s had various expression regulation under MeJA treatments. Expression analysis indicated the differential functions of cotton *LBD* genes in response to abiotic stress and hormones.

Keywords Genome-wide analysis · Lateral root · *LBD* gene family · Methyl jasmonate

Abbreviations

LBD Lateral organ boundaries domain
LOB Lateral organ boundary
MeJA Methyl jasmonate

Introduction

The Lateral Organ Boundaries Domain (LBD) proteins are a family of plant-specific transcription factors with a characteristic N-terminal Lateral Organ Boundary (LOB) domain (Iwakawa et al. 2002). The *LOB* gene was first identified in *Arabidopsis thaliana* based on the expression pattern of an enhancer trap insertion and was found to be expressed at the boundaries of lateral organs during vegetative and reproductive plant development (Shuai et al. 2002). Since *LBD* genes are found only in plants, it is implied that this gene family may regulate plant-specific growth and development processes (Shuai et al. 2002). The *LBD* gene family of *A. thaliana* can be divided into two classes according to the structure of the LOB domain. Class I has a completely conserved CX₂CX₆CX₃C zinc finger-like motif with a potential DNA-binding ability. *LBD* genes of class I also process an invariant glycine containing GAS block and a

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00425-020-03364-8>) contains supplementary material, which is available to authorized users.

✉ Shuijin Zhu
shjzhu@zju.edu.cn
✉ Jinhong Chen
jinhongchen@zju.edu.cn

¹ Zhejiang Key Laboratory of Crop Germplasm, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China

leucine zipper-like structure (LX6LX3LX6L), which allows the formation of coiled-coil protein interactions, whereas class II only contain a conserved zinc finger-like motif (Landschulz et al. 1988; Iwakawa et al. 2002; Shuai et al. 2002; Majer and Hochholdinger 2011). The majority of LBD proteins belongs to class I. Class II LBD proteins have an incomplete, probably not functional, leucine zipper that cannot form a coiled-coil structure (Majer and Hochholdinger 2011). Recently, the crystal structure of the LOB domain in an LBD protein from wheat was crystallized and determined, which reveals that the previously mentioned C-block in fact forms a C4-type zinc finger. Together with GAS motif and the perpendicular conformation between $\alpha 4$ and $\alpha 5$, the zinc finger constitutes an interaction network that precisely determines the spatial configuration of DNA recognition and binding site (Chen et al. 2019).

The lateral *LBD* gene family was originally characterized as a group of genes typically expressed by these boundary forming cells at the base of emerging lateral organs, indicating a potential role in organ separation and lateral organ development (Iwakawa et al. 2002). Some other studies also revealed additional pivotal roles of the LBD protein family in pollen development and plant regeneration, photomorphogenesis, pathogen response, and nitrogen metabolism (Thatcher et al. 2012b).

LBD genes of class I are mostly involved in plant development such as lateral organ development (Majer and Hochholdinger 2011; Xu et al. 2016), and in auxin signal transduction cascade that leads to the formation of lateral roots (Liu et al. 2005; Lee et al. 2015). In contrast, genes of class II seem to be involved in metabolism, particularly as repressors of anthocyanin synthesis and N availability signals in the plant (Rubin et al. 2009). Expression profiles of *LBD* family members in *Arabidopsis thaliana* revealed the induction of almost all *LBD* genes in class II by multiple pathogens, suggesting functions of *LBD* genes in plant defense responses (Thatcher et al. 2012a).

Many previous studies have confirmed that LBD proteins play an essential role in the regulation of the growth and development of a variety of plants. *AthLBDs* are found to be involved in various tissues development such as leaf development (Semiarti et al. 2001) and lateral root initiation (Goh et al. 2019; Cho et al. 2019; Lee et al. 2019a), as well as signaling transduction such as cytokinin (Naito et al. 2007) and gibberellin pathways (Zentella et al. 2007). To be specific, *AthLBD16* is involved in instigating the migration of nuclei and in the asymmetric division of lateral root founder cells to promote lateral root initiation (Goh et al. 2012). *AthLBD29* is involved in aspects of auxin signaling that inhibits fiber wall thickening in *Arabidopsis* stems (Lee et al. 2019b). *AtLBD20* is a *Fusarium oxysporum* susceptibility gene that appears to regulate components of jasmonic acid (JA) signalling required for full elicitation of *F. oxysporum*- and

JA-dependent responses (Feng et al. 2012). Besides *Arabidopsis*, *LBD* genes have also been studied in many other species, such as fruit, grass, and wood species. *OsIG1* is involved in the flower organ number and gametogenesis in rice (Zhang et al. 2015). Additionally, the MdLBD13 protein can inhibit anthocyanin synthesis and nitrogen uptake in apple (Li et al. 2017). Genome-wide analysis of *LBD* gene family have been studied in plants like *Camellia sinensis* (Zhang et al. 2019), *Solanum tuberosum* (Liu et al. 2019), *Eucalyptus grandis* (Lu et al. 2018), *Brassica rapa* (Huang et al. 2018), *grapevine* (Grimplet et al. 2017), and soybean (Yang et al. 2017). However, the function of the *LBD* genes in *Gossypium* remains largely unexplored.

Cotton is not only a worldwide leading textile fiber crop, but also a crop that is of significance for edible oil and other production (Zhang et al. 2008). *G. hirsutum* and *G. barbadense* are the two most widely cultivated cotton species, both of them are allotetraploids and formed by inter-genomic hybridization of *G. raimondii* and *G. arboreum*, which are the putative donor species for the D and A sub-genomic groups of tetraploid cotton species, respectively (Kadir 1976; Grover et al. 2007). In the current study, we performed a genome-wide screening for *LBD* genes in cotton, based on data gathered from recent whole-genome sequencing results. Here, we used various in silico approaches to identify and characterize *LBD* genes in cotton, and then focused on the characterization and phylogenetic relationships between *G. hirsutum* and *G. barbadense*. The *LBD* gene structures, conserved domains, synteny, as well as Cis-acting elements were explored in this systematical analysis. Moreover, the tissue-specific expression patterns and the transcriptional responses of *GhLBDs* to hormone and abiotic stress were studied in this article. These findings may provide inspirations for breeding of cotton with high stress resistance.

Materials and methods

Sequence retrieval and analysis

Two diploid cottons, *Gossypium raimondii* (Paterson et al. 2012) and *G. arboreum* (Du et al. 2018), and two allotetraploid cottons, *G. hirsutum* and *G. barbadense* (Hu et al. 2019) were analyzed. The genome sequence and annotation information of *Gossypium* species were obtained from Cottongen (Yu et al. 2014) (<https://www.cottongen.org>). The *LBD* genes of *Arabidopsis* were retrieved from TAIR 10 (<https://www.Arabidopsis.org>). Both local blast and HMMER were used to identify the *LBD* sequences in cotton. Putative protein sequences were determined by searching for the LOB domain (PF03195.14) downloaded from Pfam protein family database (<https://pfam.xfam.org>) using Hmmssearch program (E value as the default).

Besides, local blast was also applied to identify the putative sequences with the retrieved *AthLBDs*. Then, the preliminary sequences to be analyzed in cotton were obtained considering both results above. After that, candidate *LBDs* were further filtered based on their conserved domains using online software SMART (<https://smart.embl-heidelberg.de>) and Pfam (<https://pfam.sanger.ac.uk>). All these genes identified were renamed according to their genomic locations. Finally, the acquired sequences of cotton were submitted to ExPASy (<https://web.expasy.org/protparam>) to calculate the physicochemical parameters like molecular weights (MW) and theoretical isoelectric points (pI).

Phylogenetic tree construction, gene structure and motif composition analysis

Multiple sequence alignment of all these *LBD* proteins were performed using ClustalX (Larkin et al. 2007) with default parameters. A phylogenetic tree of deduced *LBD* proteins in the four *Gossypium* species and *A.thaliana* was constructed using the Maximum Likelihood method of MEGA7 (Kumar et al. 2016) with 1000 bootstrap replicates based on the JTT matrix-based model. The phylogenetic tree was subsequently decorated using EvolView (<https://www.evolgenius.info/evolview>). The exon–intron structure was analyzed based on the full-length genome sequence and the corresponding coding sequences of *LBD* genes in cotton. To identify conserved motifs in the identified sequences, the MEME program ([https://meme.nbcrl.net/meme/](https://meme.nbcrl.net/meme/intro.html) intro.html) was used to identify conserved motifs. The parameters were employed as the following descriptions: the maximum number of motifs, 15; and the optimum width of each motif, between 6 and 100 residues. The characteristics of *LBD* gene structure with motif composition were visualized by the TBtools.

Chromosomal mapping and synteny analysis of *LBD* genes in allotetraploid cotton

LBD genes were mapped on chromosomes using Mapchart software. The Multiple Collinearity Scan toolkit (MCS-canX) was used to determine *LBD* gene synteny and collinearity. Then, the Circos (Krzywinski et al. 2009) software was applied to express the syntenic relationship of the duplicated genes. To assess the selection pressure for each gene pair, the nonsynonymous substitutions rate (K_a) and synonymous substitution rate (K_s) were calculated using the KaKs_Calculator (Wang et al. 2010). Generally, $K_a/K_s > 1$ means positive selection, $K_a/K_s = 1$ indicates neutral selection, and $K_a/K_s < 1$ signifies negative or purifying selection (Hurst 2002).

Analysis of Cis-acting element in promoters of *GhLBDs* and *GbLBDs*

The upstream sequences (1.5 kb) of the *LBD* sequences were retrieved from the cotton genome sequence based on gene location.

Then, the retrieved promoter sequences were submitted to PlantCARE (<https://bioinformatics.psb.ugent.be/webtools/plantcare/html>) (Lescot et al. 2002). There were 12 Cis-elements related to phytohormone responses and stress were investigated, such as ethylene-responsive element (ERE), antioxidant response element (ARE), MeJA-responsiveness (CGTCA-motif & TGACG-motif), ABA-responsive element (ABRE), salicylic acid responsiveness (TCA-element), gibberellin-responsive element (P-box), W-box, low-temperature responsiveness (LTR), defense and stress responsiveness (TC-rich repeats), wound-responsive element (WUN-motif), and MYB binding site involved in drought-inducibility (MBS).

Expression patterns in different tissues

The Illumina RNA-seq data were downloaded at the Sequence Read Archive under accession number PRJNA490626 which were provided by Hu et al.(2019). SRA data of 10 tissues (root, stem, leaf, anther, bract, filament, petal, sepal, torus and pistil) as well as expression data of abiotic stress treatments in both *G. hirsutum* and *G. barbadense* were analyzed in this study. The SRA data were converted to fastq data with the SRA Toolkit, the clean data were mapped to the index genomes by TopHat2 program (Kim et al. 2013). The expression levels of the annotated genes were calculated with stringtie program (Pertea et al. 2015). Thereafter, the gene expression levels were visualized with the homogenized method involving $\log_2(\text{TPM} + 1)$ and $\log_2(\text{fold change})$ using the pheatmap program (<https://CRAN.R-project.org/package=pheatmap>).

Plant materials and treatments

Gossypium hirsutum cv. TM-1 was used, which is a standard genetic line of upland cotton (obtained from USDA-ARS, College Station, Texas, USA). Cotton seeds were sown in pots containing the mixed nutrient soil under a 16-h light/8-h dark cycle at 26 °C for 2 weeks. Then, all plants were transferred to Hoagland solution for several days. The seedlings at three-leaf stage were subjected to 100 μM MeJA. The roots of treated plants were harvested at 0, 3, 6, and 12 h. All these samples were immediately frozen in liquid nitrogen and kept at -80 °C for subsequent analysis.

RNA extraction and quantitative real-time (RT-qPCR) analysis

The total RNA of cotton roots were extracted using the RNAprep Pure Plant Kit (Tiangen, Beijing, China). The quantity and quality were determined by a NanoDrop 2000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). The cDNA was reverse transcribed with the ReverTra Ace qPCR RT Kit (with gDNA remover) (Toyobo, Osaka, Japan). All the operational procedures followed the manufacturer's protocols. The qPCR analysis was performed using a LightCycler FastStart DNA Master SYBR Green I kit (Roche, Basel, Switzerland) on a LightCycler96 Real-Time PCR detection system (Roche). For internal reference gene, *UBQ7* was used, which was stably expressed in cotton plants and not affected by treatments and genotypes. All RT-qPCR analyses were performed with three biological replicates and each analysis was carried out in three technical RT-qPCR replicates. Statistical analysis was conducted with biological replicates with mean values of three technique replicates. $2^{-\Delta CT}$ method was used to further study the data from real-time PCR amplification (Livak and Schmittgen 2001). The primers were designed using Primer-BLAST (Ye et al. 2012), listed in Supplementary Table S1.

Results

Identification of *LBD* genes in cotton

To evaluate the *LBD* gene family in cotton, an HMMER alignment search on the proteome of the *G. hirsutum*, *G. barbadense*, *G. raimondii*, and *G. arboreum*, and genome with the hidden Markov model (HMM) profile of the lateral organ boundaries domain (PF3195) was performed. Then, all these retrieved sequences were verified by the Pfam and SMART. The molecular weight was also considered to identify the *LBD* genes. Sequences without a typical LOB domain or with a molecular weight outside of the 11–60-kDa range were excluded. Finally, the confirmed *LBD* genes in *G. raimondii* (GrLBDs), *G. arboreum* (GaLBDs), *G. hirsutum* (GhLBDs), and *G. barbadense* (GbLBDs) amounted to be 68, 62, 131, and 128, respectively (Supplementary Table S2). All these genes were named based on their chromosomal locations. The *LBD* genes in tetraploid cotton (*G. hirsutum* and *G. barbadense*) were almost twice the number of diploid cotton (*G. raimondii* and *G. arboreum*). Therefore, it was proven again that allotetraploid cotton was formed by inter-genomic hybridization of A-genome diploids and D-genome diploids, and followed by chromosome doubling (Paterson et al. 2012).

Gene characteristics, including the size of the protein sequence (amino acid), the protein molecular weight

(MW), isoelectric point (pI), and the subcellular localization were listed in Supplementary Table S2. It was observed that the length of the GbLBD proteins ranged from 106 to 521 amino acids, and GhLBD was from 137 to 524 amino acids. The average length of GhLBDs was much longer than that of GbLBDs. The average MW for GhLBDs was 23,591.72 Da, and that of GbLBDs, GaLBDs, and GrLBDs were 23,710.96 Da, 23,407.13 Da, and 23,056.92 Da, respectively. The PI value of LBDs in four cotton species above showed similar statistical results (about 7.3). The predicted subcellular location of LBD proteins suggested that nuclear was the only area where cotton LBDs implements their biological functions as most of the LBDs in cotton were located in the nuclear region except for the NONE results.

Phylogenetic analysis of the cotton *LBD* gene family

To detect the evolutionary relationships of LBD proteins among four cotton species, an unrooted phylogenetic Maximum Likelihood (ML) tree was constructed using full-length amino acid sequences of these identified genes. The phylogenetic analysis (Fig. 1) revealed that the LBDs could be divided into two main clades corresponding to class I and II of *Arabidopsis* as defined by Mangeon et al. (2012). There were 330 and 59 *LBD* genes involved in class I and II of the fore mentioned cotton species. Therefore, the quantity of *LBD* genes in class I was fifth as big as that in class II of each cotton species, approximately. Class II consisted of 20, 19, 10, and 10 *LBD* genes separately in *G. hirsutum*, *G. barbadense*, *G. raimondii*, and *G. arboreum*. According to previous studies of traditional classification and distribution of LBDs (Coudert et al. 2013), class I could be further grouped into five sub-classes (1A–1E) based on tree topology and branch lengths. Each of the subgroups contained proteins from both diploid and allotetraploid cotton species. Subgroup 1B was the smallest in class I with 8, 7, 4, and, 4 genes in *G. hirsutum*, *G. barbadense*, *G. raimondii*, and *G. arboreum*, respectively. While the subgroup 1C was the largest with 31, 31, 16, and 15 *LBD* genes in the same cotton species. The *LBD* genes in tetraploid cotton were almost twice as many as diploid cotton in each subgroup. Each clade comprised members of all the species analyzed, which suggested that *LBDs* were homologous with those in *Arabidopsis*. According to previous studies (Majer and Hochholdinger 2011), many *LBD* genes in *Arabidopsis* (class I) were involved in lateral organ development, including *AthLBD13*, *AthLBD3*, *AthLBD12*, *AthLBD6*, *AthLBD36* and so on. Genes such as *AthLBD16*, *AthLBD18*, and *AthLBD29* participated in the auxin signal transduction cascade that lead to the formation of lateral roots in *Arabidopsis*. However, class II *LBD* genes, *AthLBD37*, *AthLBD38* and *AthLBD39* were

Fig. 1 Molecular phylogenetic analysis of 389 LBD proteins from *G. hirsutum*, *G. barbadense*, *G. arboreum*, and *G. raimondii*. MEGA 7.0 was used to build the Maximum Likelihood method (ML) tree with

1000 bootstrap replicates. GhLDBs, GbLDBs, GaLDBs, GrLDBs, and AthLDBs were represented by stars in red, green, yellow, black, and white, respectively

involved in metabolism. *AtLBD37* was down-regulated by gibberellin and up-regulated by DELLA proteins (Majer and Hochholdinger 2011). The phylogenetic analysis indicated that the homologs in cotton might have similar biological functions as *AthLDBs*.

Gene structure and motif composition analysis

To further explore the possible evolutionary relationship of LBD proteins in allotetraploid cotton (*G. hirsutum* and *G. barbadense*), phylogenetic trees of both GhLDBs and GbLDBs

Fig. 2 Phylogenetic relationships, architecture of conserved protein motifs and gene structure in *LBD* genes of *G. hirsutum*. **a** The phylogenetic tree was constructed based on the full-length sequences of *GhLBD* using MEGA 7 software. **b** The motif composition of *LBD*

proteins. The 15 motifs were displayed in different colored boxes. The sequence information for each motif was provided in Supplementary Table S3. **c** Exon–intron structure of *GhLBDs*. Green boxes indicated exons; black lines indicated introns

were constructed (Figs. 2a, 3a). Conserved motifs in these proteins were identified using MEME software. Here, 15 conserved motifs (named motif 1 to motif 15) were identified in both *GhLBDs* and *GbLBDs*, and the results were represented in schematic diagrams (Fig. 2b, Fig. 3b). The sequence information for each motif of *G. hirsutum* and *G. barbadense* was provided in Supplementary Table S3. *LBD* members within the same clade were usually found to share a similar motif composition, which was conspicuous in class II. All the proteins belonged to class II only contained motif 3 in *G. barbadense*. While in *G. hirsutum*, motif 3, motif 13 and motif 15 were existed in *LBD* proteins of class II. As for the *LBD* proteins of class I, the number of motifs was two to four in *G. hirsutum* and two to five in *G. barbadense*. The different compositions of the motifs may indicate functional diversity. The subgroups of *GbLBDs* in class I showed similar motif arrangements, revealing that the protein architecture was conserved within a specific subfamily. The functions of most of these conserved motifs remain to be elucidated. Then, comparative analysis of intron–exon structures was performed

by the alignment of genomic DNA with full-length cDNA (Figs. 2c, 3c). The gene structure patterns were similar in the two cotton species. *GhLBDs* and *GbLBDs* possessed 1–13 exons and could be divided into four categories based on exon number. Therefore, in *G. hirsutum*, 27 (20.6%) *LBD* genes had 1 single exons, 93 (71%) *LBD* genes possessed 2 exons, 10 (7.6%) *LBD* genes owned 3 exons and only 1 (0.8%) *LBD* gene had more than 3 exons. In *G. barbadense*, 93 genes (73%) only had 1 exon, 25 genes (20%) had 2 exons, 9 genes (7%) possessed 3 exons, and none had more than 3 exons. Notably, the intron–exon structures were consistent with the alignment clusters of *GbLBDs*, indicating that exon–intron structure and the phylogenetic relationship among *GhLBDs* were highly correlated.

Chromosomal distribution and synteny analysis of *GbLBD* genes

Based on the public information of *G. hirsutum* and *G. barbadense* genome, the location of genes and the length of

Fig. 3 Phylogenetic relationships, architecture of conserved protein motifs and gene structure in *LBD* genes of *G. barbadense*. **a** The phylogenetic tree was constructed based on the full-length sequences of *GbLBD* using MEGA 7 software. **b** The motif composition of *LBD*

proteins. The 15 motifs were displayed in different colored boxes. The sequence information for each motif was provided in Supplementary Table S3. **c** Exon–intron structure of *GbLBDs*. Green boxes indicated exons; black lines indicated introns

chromosomes were available to analyze the chromosomal distribution (Fig. S1, Fig. S2). High similarity was found in the chromosomal distribution patterns of these two cotton species. The *LBD* genes were unevenly arranged on chromosomes. However, no *GbLBDs* were found on chromosome A02 and A04 in both *G. hirsutum* and *G. barbadense*. There were 131 *GhLBDs* distributed throughout the 24 chromosomes comprising 66 located on the At subgenome and 65 located on the Dt subgenome. In case of *G. barbadense*, 63 *LBD* genes were located on the At subgenome and 65 located on the Dt subgenome. The majority of *GhLBDs* and *GbLBDs* was located on the proximate or the distal ends of the chromosomes. In the two allotetraploid cotton species,

the homologous chromosome pair A11 and D11 separately contained the largest gene clusters in At and Dt sub-genome. The chromosome A01 and D01 had the minimal *LBD* gene numbers in its sub-genomes. There were no positive correlations between the length of chromosome and the *LBD* gene number.

According to previous reports, gene duplications are considered to be one of the primary driving forces in the evolution of genomes and genetic systems (Moore and Purugganan 2003), and segmental and tandem duplications have been suggested to represent two of the main causes of gene family expansion in plants (Cannon et al. 2004). Segmental duplications multiple genes through polyploidy followed by

chromosome rearrangements (Yu et al. 2005). It occurs most frequently in plants, because most plants are diploidized polyploids and retain numerous duplicated chromosomal blocks within their genomes (Cannon et al. 2004). Tandem duplications were characterized as multiple members of one family occurring within the same intergenic region or in neighboring intergenic regions (Ramamoorthy et al. 2008). According to a previous study, a chromosome region containing two or more genes within 200 kb could be defined as a gene cluster (Holub 2001). In the present study, blast and MCScanX were used to further investigate the expanded mechanism of *LBD* gene family by conducting genome synteny analysis (Fig. 4). Eight *GhLBDs*, *GhLBD22/23*, *GhLBD29/30*, *GhLBD88/89*, and *GhLBD95/96* were clustered into three tandem duplication event regions on chromosome A06, A08, D06, and D08, respectively. However, for *GbLBDs*, there were three tandem duplication events, *GbLBD28/29*, *GbLBD85/86*, and *GbLBD92/93*, on chromosome A08, D06, and D08, respectively. While segmental duplication accounted for a large part of the *GhLBDs* and *GbLBDs*, indicating that compared to tandem duplication, the segmental duplication played a predominant driving force in *LBD* gene family evolution. In addition, the segmental duplication events were mainly occurred among different subgenomes. To decipher the evolutionary relationships of *LBD* genes, intergenomic synteny analysis was further performed among different cotton species (Fig. 5). There

were 60, 58, 49, and 63 collinear gene pairs between *G. hirsutum* and *G. arboreum*, *G. barbadense* and *G. arboreum*, *G. hirsutum* and *G. raimondii*, and *G. barbadense* and *G. raimondii*, respectively. Therefore, it is illustrated that *LBD* genes in *G. raimondii* and *G. barbadense* possess closer evolutionary relationships. Thus, the current results provided evidences that *LBD* genes may have experienced some genomic rearrangements during polyploidization. To better understand the evolutionary constraints acting on *LBD* genes functional divergence after gene duplication, non-synonymous to synonymous substitution ratio (K_a/K_s) for each pair of duplicated *LBD* genes were calculated (Supplementary Table S4). There were 15 duplicated *GhLBDs* pairs predicted to undergo positive selection with the $K_a/K_s < 1$. As for the *GbLBDs*, a majority of K_a/K_s ratio was more than 1 except for 7 pairs of genes ($K_a/K_s < 1$), suggesting that the *LBD* gene family in *G. hirsutum* and *G. barbadense* might have experienced strong purifying selective pressure during evolution.

Cis-element analysis of *LBD* genes in allotetraploid cotton

In this current study, Cis-elements are responsive to corresponding elicitors to regulate the expression of genes. In this study, a 1.5-kb region upstream from the start codon of each *LBD* gene was selected to investigate putative

Fig. 4 **a** Schematic representations for the chromosomal distribution and interchromosomal relationships of *LBD* genes in *G. hirsutum* **b** Schematic representations for the chromosomal distribution and inter-

chromosomal relationships of *LBD* genes in *G. barbadense*. Gray lines indicated all synteny blocks in the cotton genome, and red lines indicated duplicated *LBD* gene pairs

Fig. 5 **a** Synteny analyses of *LBD* genes between A genome (*G. arboreum*) and At subgenome (*G. hirsutum* and *G. barbadense*). **b** Synteny analyses of *LBD* genes between D genome (*G. raimondii*) and Dt subgenome of two allotetraploid cotton species. Gray

lines in the background indicated the collinear blocks within different genomes, while other colored lines highlighted the syntenic *LBD* gene pairs

Cis-elements involved in the mediation of gene expression using the PlantCARE server. Some cis-elements predicted to be involved in phytohormone responses and stress were found in *GhLBDs* and *GbLBDs* (Fig. S3). In *G. hirsutum*, 95 *GhLBD* gene promoters processed the ethylene-responsive element (ERE), and 90 *GhLBD* genes had at least 1 antioxidant response element (ARE). There were 87 *GhLBDs* processed the elements involved in the MeJA-responsiveness (CGTCA-motif and TGACG-motif), 71 *GhLBDs* had ABA-responsive element (ABRE), 53 *GhLBDs* had the cis-acting element involved in salicylic acid responsiveness (TCA-element), and 35 *GhLBDs* had gibberellin-responsive element (P-box). The putative Cis-elements exposed to stress were less as compared to hormones. W-box is a cis-acting element recognized by *WRKY* transcription factors, involved in pathogen related responses (Eulgem et al. 1999). Some of these W-boxes were also responsive to ethylene (Miao et al.

2004). 73 *GhLBDs* processed the W-box. The remaining elements related to stress, like cis-acting element involved in low-temperature responsiveness (LTR), defense and stress responsiveness (TC-rich repeats), wound-responsive element (WUN-motif), and MYB-binding site involved in drought inducibility (MBS), were existed in about 40 *GhLBD* genes. However, as compared to *G. hirsutum*, fewer cis-elements were found in the promoter regions of *GbLBDs*. That was 90 contained ARE, 88 possessed ERE, 67 had ABRE, 57 had W-box, and 55 possessed TCA-element. Many of the *GbLBDs* also harbored cis-elements such as TGACG-motif (41 *GbLBDs*), CGTCA-motif (40 *GbLBDs*), TC-rich repeats (38 *GbLBDs*), WUN-motif (37 *GbLBDs*), MBS (37 *GbLBDs*), and LTR and P-box (35 *GbLBDs*).

These results suggested that *LBD* genes might play an important role in the phytohormone signal transduction and stress response in both *G. hirsutum* and *G. barbadense*.

Tissue-specific expression patterns of *GhLBDs* and *GbLBDs*

To gain further insights into the potential developmental roles of *LBD* genes in allotetraploid cotton, the published transcriptomic data (Hu et al. 2019) were investigated to analyze the expression of cotton *LBD* genes in various tissues, including vegetative (root, stem, and leaf) and reproductive (anther, bract, filament, petal, sepal, torus and pistil) tissues. According to the heat map (Fig. 6), those expression characteristics of *GhLBDs* and *GbLBDs* in different tissues were quite various. In *G. hirsutum*, a large part of *GhLBDs* witnessed low expression levels in leaf, while the majority showed high transcriptional levels in stem and root, indicating that they had a conserved functional role in root and stem development. Genes ranging from *GhLBD87* to *GhLBD106* were highly expressed in both root and stem, and among them only one gene, *GhLBD47*, was in class II. The situation was more harmonious in anther, bract, filament, petal, sepal, torus, and pistil, where the expression level of *GhLBDs* was similar among these reproductive tissues, except for the anther and filament with a higher expression level. It was confirmed with the previous finding that *LBD* protein family

played pivotal roles in pollen development (Thatcher et al. 2012b). Moreover, a majority of *GhLBDs* predominantly expressed in the vegetative tissues might express with a relatively low level in those reproductive tissues and vice versa. For *LBD* genes in *G. barbadense*, the situation was quite similar. *GbLBDs* were also most abundant in root and stem, followed by these reproductive tissues and leaf.

Expression profiles of cotton *LBD* genes under abiotic stress and hormone treatment

The *LBD* gene expression profiles under abiotic stress were checked to investigate their potential functions in *G. hirsutum* and *G. barbadense* using transcriptome data (Hu et al. 2019). According to the results of *GhLBDs* expression patterns (Fig. 7a), more than half of them were induced by PEG treatment and a small part of them were down-regulated by PEG. These genes responding to PEG also suggested that the influences may happen immediately, and 12 h and 24 h past treatment may play a more important role in this process. Treatments such as hot, cold and salt played insignificant effects on *GhLBDs* expression with a small amount of genes differently expressed. Among these genes, *GhLBD123* was

Fig. 6 **a** Expression patterns of *GhLBDs* in different tissues. **b** Expression patterns of *GbLBDs* in different tissues. Scale bars represent the log₂ transformations of the TPM values

Fig. 7 **a** Expression profiles of *GhLBDs* under abiotic stress treatments. **b** Expression profiles of *GbLBDs* under abiotic stress treatments. Scale bars represent the $\log_2$ transformations of the foldchange values

up-regulated by heat, and reached a peak at 24 h past treatment, *GhLBD125* was induced by cold and the expression level reached a peak at 12 h past treatment. Nevertheless, the situation was quite different in *G. barbadense* (Fig. 7b). Great majority of *GbLBDs* did not seem to express differently under the abiotic stresses, even under PEG treatment. MeJA, belonging to a class of cyclopentanone compounds, is a naturally and ubiquitously occurring phytohormone involved in signal transduction pathway and plant response to environmental stressors (Santino et al. 2013), and *LBD* genes played roles in response to MeJA treatment (Thatcher et al. 2012b). The present results showed that a majority of *GhLBDs* processed the cis-elements involved in the MeJA-responsiveness, so these *GhLBD* genes might have functions in response to MeJA treatment. To verify these predictions, the expression changes of *GhLBDs* were monitored by RT-qPCR at different time (0, 3, 6, and 12 h) under MeJA treatment. Finally, 22 genes covering almost all the sub-families were picked up to be analyzed through RT-qPCR. Nine genes, *GhLBD01*, *GhLBD07*, *GhLBD12*, *GhLBD14*, *GhLBD38*, *GhLBD54*, *GhLBD85*, *GhLBD91*, *GhLBD106*, and *GhLBD123*, with higher relative expression level were analyzed (Fig. 8). Almost all these nine genes reached the highest expression level at 12 h except for *GhLBD91*,

which increased to the maximum level at 6 h. Additionally, *GhLBD123* showed strongest up-regulation compared to the other eight *GhLBDs*.

Discussion

LBD genes comprise a plant-specific transcription factor family, which had been proven to have essential roles in the regulation of lateral organ development and metabolic processes, such as anthocyanin and nitrogen metabolism in higher plants (Iwakawa et al. 2002; Shuai et al. 2002; Majer and Hochholdinger 2011). Comprehensive analyses of this gene family had been conducted in many species such as *Arabidopsis* (Liu et al. 2005), *Eucalyptus grandis* (Lu et al. 2018), *Brassica rapa* (Huang et al. 2018), grapevine (Grimplet et al. 2017), soybean (Yang et al. 2017), and *Camellia sinensis* (Teng et al. 2018). However, no systematic study of *LBD* genes has been conducted in allotetraploid cotton. The *de-novo* assembly of the genomes by integrating data from Illumina PCR-free short-read sequencing, 10× Genomics sequencing, Hi-C, and optical and super-dense genetic maps have resulted in substantial improvements to the contiguity and accuracy of assembly, with a notable improvement

Fig. 8 Relative expressions of 9 *GhLBD* genes under MeJA treatment based on RT-qPCR. Data were normalized to *UBQ7* gene. The values are means $\pm$ standard deviation (SD) of three replicates. The details of the primer sequences were listed in Supplementary Table S1

in the assembly of centromeres (Hu et al. 2019). In recent years, higher quality *de-novo*-assembled genomes of the two allotetraploid species (Hu et al. 2019) and *G. arboreum* (Du et al. 2018) were published, which will enable a new era of genome-wide analyses of cotton gene families. In the current study, 68, 62, 131, and 128 candidate *LBD* genes were systematically identified in *G. raimondii*, *G. arboreum*, *G. hirsutum*, and *G. barbadense*, respectively. The number of *LBD* genes in the two tetraploid species was almost the sum of those in the two ancestral diploid progenitors, which provided evidence that allotetraploid cottons may have appeared through hybridization and subsequent polyploidization events between the A- and D-subgenome progenitors (Lynch and Conery 2000). The characteristic LOB

domain comprises a C-block containing four cysteine with spacing (CX2CX6CX3C) required for DNA-binding activity, a Gly-Ala-Ser (GAS) block and a leucine zipper-like coiled-coil motif (LX6LX3LX6L) responsible for protein dimerization (Shuai et al. 2002; Majer and Hochholdinger 2011). The *LBD* genes in *Gossypium* could be divided into two main subfamilies, class I and class II, according to the structure of LOB domain. Most of *LBD* genes belong to class I. 20 (15.3%), 19 (14.8%), 10 (16.1%), and 10 (14.7%) *LBD* genes were classified as class II in *G. hirsutum*, *G. barbadense*, *G. arboreum*, and *G. raimondii*, respectively. Members of class I could be grouped into five sub-classes (1A–1E) based on tree topology and branch lengths. The intron–exon structures and motif compositions of *GhLBDs*

and *GbLBDs* were consistent with the alignment clusters of these *LBD* genes. Finally, many similarities could be found between *GhLBDs* and *GbLBDs* in the gene structure, which indicated that *LBD* genes in the two allotetraploid cotton species might serve close functions, and they might have the same origin.

Gene duplication generates functional divergence, which is essential for environmental adaptability and speciation (Conant and Wolfe 2008). During evolution, duplicated gene pairs can experience functional divergence resulting in neo-functionalization, sub-functionalization, or non-functionalization (Prince and Pickett 2002). Gene duplication is the most commonly evaluated mechanism for gene family expansion (Cannon et al. 2004). It is widely accepted that the homologous genes created by a duplication inside a genome might have occurred through various mechanisms, including (i) single gene duplications (tandem repeats, clustering, short range duplications, and transposition); (ii) reverse transcript insertion, (iii) duplications of chromosomal blocks (translocations); (iv) duplications of single chromosomes (aneuploidy); or (v) duplication of the whole genome (polyploidy) (Prince and Pickett 2002). To reveal the expansion mechanism of the *LBD* gene family, intragenomic and intergenomic duplication data of these cotton species were filtered by MCScanX. According to the current intragenomic analysis results, only several tandem duplication events existed in both *G. hirsutum* and *G. barbadense*, which revealed that segmental duplication played a predominant driving force in *LBD* gene family evolution. In both *G. hirsutum* and *G. barbadense*, the number of homologous *LBD* genes in At and Dt subgenome was identical, thus we can infer that the translocations and reverse transcript insertion rarely occurred in the generation of cotton *LBD* gene family. The fates of duplicated genes are diverse (Charon et al. 2012). First, when the original function is maintained in both copies, this leads to functional redundancy. Another possibility is that both copies of a gene share the ancestral function, gradually developing partially different functions, the so-called sub-functionalization. In addition there is also a possibility, although rare, is that one of the two copies evolves, acquiring a new function. Finally, intermediate evolutionary stages between sub- and neo-functionalization in which only genes critical for plant growth and development are retained. The expression level of many duplicated *GhLBDs* and *GbLBDs* was quite similar (Fig. 6), which suggested that the duplicated *LBD* genes in *G. hirsutum* and *G. barbadense* may lead to functional redundancy mainly.

Synteny analysis in different species is a way to provide insights to their evolution and relationship. In our study, the synteny analysis of *LBD* genes between diploid and allotetraploid cotton was conducted (Fig. 5). Only two pairs of *GhLBD* and *GbLBD* genes were within the intraspecific

rearrangement blocks between *G. hirsutum* and *G. raimondii*, as well as *G. barbadense* and *G. raimondii*, which were far less than that located in rearrangement blocks between *G. hirsutum* and *G. arboreum*, or *G. barbadense* and *G. arboreum*. The result was consistent with the conclusion that the A-derived subgenomes have apparently been more active than D-derived subgenomes during evolution (Yang et al. 2019). To illuminate the divergence after gene duplication, the synonymous substitution rate (*Ks*) and nonsynonymous substitution rate (*Ka*) of duplicated *LBD* genes from *G. hirsutum* and *G. barbadense* were calculated. The current results suggested that *GhLBDs* and *GbLBDs* were under strong purifying selection, and the *LBD* genes from *G. hirsutum* had experienced stronger positive selection than that from *GbLBDs*.

LBD proteins orchestrate a lot of plant developmental processes and respond to environmental stimuli, especially play the essential roles in the regulation of lateral organ development and metabolic processes. Several *LBD* genes, such as *LBD16*, *LBD13*, *LBD18*, and *LBD29* have been proven to play important roles in regulating lateral root development in *Arabidopsis* (Cannon et al. 2004; Goh et al. 2019; Cho et al. 2019; Lee et al. 2019a). *LBD18* controls lateral root development via inhibition of *LBD18* DNA-binding activity (Lee et al. 2013). Thus, it is reasonable to predicate that the *LBD* genes may play roles in plant resistance as the root architecture plasticity is an important feature in both abiotic and biotic responses. Generally speaking, the main root may be affected by the overgrowth of lateral roots, which may reduce the ability of plant to extract water and nutrients. Too many lateral roots spread in the soil will increase the possibility of being attacked by pathogens. Some previous work has proven these predications. *LBD1* in *Medicago* controls root architecture under salt stress via the inhibition of primary root growth and maintenance of lateral root emergence (Ariel et al. 2010). *AtLBD20*, a root-specific *LBD* gene in *Arabidopsis*, is a negative regulator of susceptibility to the root-infecting fungal pathogen *Fusarium oxysporum* (Thatcher et al. 2012b). *LBD* genes are crucial molecular targets for plant pathogen invasion due to their transcriptional regulation of a range of downstream genes (Xu et al. 2016).

It has been demonstrated that *LBD* proteins make differences in response to phytohormone signaling. In *Arabidopsis*, *LBD40* was reported to be down-regulated by gibberellin (Zentella et al. 2007), *LBD20* was proposed playing a role in JA-response (Thatcher et al. 2012a). In *Camellia sinensis*, *LBD* genes had differential functions in response to the treatments with gibberellic acid (GA3), salicylic acid (SA), MeJA, ABA, and ethephon (Eth) (Teng et al. 2018). Considering that *LBD* genes are conservative in structure and function in different species (Wang et al. 2013; Yang et al. 2016), and the results of Cis-acting

element in promoters of *GhLBDs* and *GbLBDs* led us to draw the predication that the *LBDs* in cotton were involved in phytohormone responses and stress resistance. In this study, expression profiles of *GhLBDs* and *GbLBDs* responding to abiotic stress were studied using the published transcriptome data. More *LBD* genes in upland cotton expressed differently after the hot, cold, salt, and PEG treatments. Furthermore, expression levels of *GhLBD* were widely induced under PEG, which was hard to find in *G. barbadense*. Therefore, it indicated that *GhLBD* genes might be involved in cotton adapting to drought especially, while *GbLBDs* might play an unimportant role in improving cotton abiotic stress tolerance. Several *GhLBDs* were selected by the expression level in different tissues (Fig. 6) as well as the results of Cis-element. According to the current study, *GhLBDs* in response to MeJA showed different expression regulation, seven genes witnessed up-regulated at 12 h. Especially, for the *GhLBD123*, the expression level owned huge increase (Fig. 8). These results demonstrated that *GhLBD* genes had differential functions in response to hormone like MeJA and water-deficit stress. The present work not only established a foundation for the further functional analyses of *LBD* genes in cotton, but can also provide valuable information for potential candidate genes involved in plant growth, development, and stress tolerance.

In conclusion, comparisons of gene structure, chromosomal distribution, phylogenetic relationship, synteny characteristic as well as cis-element traits between *GhLBDs* and *GbLBDs* were performed in the present study, which largely enriched our knowledge of the cotton *LBD* gene family. In addition, expression and functional characterization indicated the involvement of *GhLBDs* in abiotic stresses responses, especially in drought condition. It also indicated the differential functions of *GhLBD* genes in response to hormones like MeJA. These findings could provide valuable information for subsequent elucidation roles of *LBD* genes.

Author contribution statement SZ and JY designed the research. JY and QX conducted experiments. CL and JY performed the data analyses. JY wrote the manuscript. All authors read and approved the manuscript. SZ and JC contributed substantially to revisions.

Acknowledgements We are grateful to Dr. Zhang (Zhejiang University, China) for assistance with our greenhouse work. The research work was funded by the National Key Technology R&D program of China (2016YFD0101404), Agriculture Research System of China (CARS-18-25), Jiangsu Collaborative Innovation Center for Modern Crop Production.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ariel FD, Diet A, Crespi M, Chan RL (2010) The LOB-like transcription factor Mt LBD1 controls *Medicago truncatula* root architecture under salt stress. *Plant Signal Behav* 5(12):1666–1668. <https://doi.org/10.4161/psb.5.12.14020>
- Cannon SB, Mitra A, Baumgarten A, Young ND, 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. <https://doi.org/10.1186/1471-2229-4-10>
- Charon C, Bruggeman Q, Thareau V, Henry Y (2012) Gene duplication within the Green Lineage: the case of *TEL* genes. *J Exp Bot* 63(14):5061–5077. <https://doi.org/10.1093/jxb/ers181>
- Chen WF, Wei XB, Rety S, Huang LY, Liu NN, Dou SX, Xi XG (2019) Structural analysis reveals a “molecular calipers” mechanism for a lateral organ boundaries domain transcription factor protein from wheat. *J Biol Chem* 294(1):142–156. <https://doi.org/10.1074/jbc.RA118.003956>
- Cho C, Jeon E, Pandey SK, Ha SH, Kim J (2019) *LBD13* positively regulates lateral root formation in *Arabidopsis*. *Planta* 249(4):1251–1258. <https://doi.org/10.1007/s00425-018-03087-x>
- Conant GC, Wolfe KH (2008) Turning a hobby into a job: How duplicated genes find new functions. *Nature Rev Genet* 9(12):938–950. <https://doi.org/10.1038/nrg2482>
- Coudert Y, Dievart A, Droc G, Gantet P (2013) ASL/LBD phylogeny suggests that genetic mechanisms of root initiation downstream of auxin are distinct in Lycophytes and Euphyllophytes. *Mol Biol Evol* 30(3):569–572. <https://doi.org/10.1093/molbev/mss250>
- Du X, Huang G, He S, Yang Z, Sun G, Ma X, Li N, Zhang X, Sun J, Liu M, Jia Y, Pan Z, Gong W, Liu Z, Zhu H, Ma L, Liu F, Yang D, Wang F, Fan W, Gong Q, Peng Z, Wang L, Wang X, Xu S, Shang H, Lu C, Zheng H, Huang S, Lin T, Zhu Y, Li F (2018) Resequencing of 243 diploid cotton accessions based on an updated A genome identifies the genetic basis of key agronomic traits. *Nat Genet* 50(6):796–802. <https://doi.org/10.1038/s41588-018-0116-x>
- Eulgem T, Rushton PJ, Schmelzer E, Hahlbrock K, Somssich IE (1999) Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. *EMBO J* 18(17):4689–4699. <https://doi.org/10.1093/emboj/18.17.4689>
- Feng Z, Zhu J, Du X, Cui X (2012) Effects of three auxin-inducible LBD members on lateral root formation in *Arabidopsis thaliana*. *Planta* 236(4):1227–1237. <https://doi.org/10.1007/s00425-012-1673-3>
- Goh T, Kasahara H, Mimura T, Kamiya Y, Fukaki H (2012) Multiple AUX/IAA-ARF modules regulate lateral root formation: the role of *Arabidopsis* SHY2/IAA3-mediated auxin signalling. *Phil Trans Royal Soc B-Biol Sci* 367(1595):1461–1468. <https://doi.org/10.1098/rstb.2011.0232>

- Goh T, Toyokura K, Yamaguchi N, Okamoto Y, Uehara T, Kaneko S, Takebayashi Y, Kasahara H, Ikeyama Y, Okushima Y, Nakajima K, Mimura T, Tasaka M, Fukaki H (2019) Lateral root initiation requires the sequential induction of transcription factors LBD16 and PUCHI in *Arabidopsis thaliana*. *New Phytol* 224(2):749–760. <https://doi.org/10.1111/nph.16065>
- Grimplet J, Pimentel D, Agudelo-Romero P, Miguel Martinez-Zapater J, Fortes AM (2017) The *Lateral Organ Boundaries Domain* gene family in grapevine: genome-wide characterization and expression analyses during developmental processes and stress responses. *Sci Rep* 7:15968. <https://doi.org/10.1038/s41598-017-16240-5>
- Grover CE, Kim H, Wing RA, Paterson AH, Wendel JF (2007) Microcolinearity and genome evolution in the *AdhA* region of diploid and polyploid cotton (*Gossypium*). *Plant J* 50(6):995–1006. <https://doi.org/10.1111/j.1365-3113X.2007.03102.x>
- Holub EB (2001) The arms race is ancient history in *Arabidopsis*, the wildflower. *Nature Rev Genet* 2(7):516–527. <https://doi.org/10.1038/35080508>
- Hu Y, Chen J, Fang L, Zhang Z, Ma W, Niu Y, Ju L, Deng J, Zhao T et al (2019) *G. barbadense* and *G. hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. *Nature Genet* 51(4):739–748. <https://doi.org/10.1038/s41588-019-0371-5>
- Huang X, Liu G, Zhang W (2018) Genome-wide analysis of LBD (Lateral Organ Boundaries Domain) gene family in *Brassica rapa*. *Brazilian Arch Biol Technol*. <https://doi.org/10.1590/1678-4324-2018180049>
- Hurst LD (2002) The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet* 18(9):486–487. [https://doi.org/10.1016/s0168-9525\(02\)02722-1](https://doi.org/10.1016/s0168-9525(02)02722-1)
- Iwakawa H, Ueno Y, Semiarti E, Onouchi H, Kojima S, Tsukaya H, Hasebe M, Soma T, Ikezaki M, Machida C, Machida Y (2002) The *ASYMMETRIC LEAVES2* gene of *Arabidopsis thaliana*, required for formation of a symmetric flat leaf lamina, encodes a member of a novel family of proteins characterized by cysteine repeats and a leucine zipper. *Plant Cell Physiol* 43(5):467–478. <https://doi.org/10.1093/pcp/pcf077>
- Kadir ZBA (1976) DNA evolution in genus *Gossypium*. *Chromosoma* 56(1):85–94
- Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14:R36. <https://doi.org/10.1186/gb-2013-14-4-r36>
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: An information aesthetic for comparative genomics. *Genome Res* 19(9):1639–1645. <https://doi.org/10.1101/gr.092759.109>
- Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. *Mol Biol Evol* 33(7):1870–1874. <https://doi.org/10.1093/molbev/msw054>
- Landschulz WH, Johnson PF, McKnight SL (1988) The leucine zipper - A hypothetical structure common to a new class of DNA-binding proteins. *Science* 240(4860):1759–1764. <https://doi.org/10.1126/science.3289117>
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and clustal X version 2.0. *Bioinformatics* 23(21):2947–2948. <https://doi.org/10.1093/bioinformatics/btm0404>
- Lee HW, Kim M-J, Park MY, Han K-H, Kim J (2013) The conserved proline residue in the LOB domain of LBD18 is critical for DNA-binding and biological function. *Mol Plant* 6(5):1722–1725. <https://doi.org/10.1093/mp/sst037>
- Lee HW, Cho C, Kim J (2015) *Lateral Organ Boundaries Domain16* and *18* act downstream of the AUXIN1 and LIKE-AUXIN3 auxin influx carriers to control lateral root development in *Arabidopsis*. *Plant Physiol* 168(4):1792–1806. <https://doi.org/10.1104/pp.15.00578>
- Lee HW, Cho C, Pandey SK, Park Y, Kim M-J, Kim J (2019a) *LBD16* and *LBD18* acting downstream of *ARF7* and *ARF19* are involved in adventitious root formation in *Arabidopsis*. *BMC Plant Biol* 19:46. <https://doi.org/10.1186/s12870-019-1659-4>
- Lee KH, Du Q, Zhuo CL, Qi LY, Wang HZ (2019b) *LBD29*-involved auxin signaling represses nac master regulators and fiber wall biosynthesis. *Plant Physiol* 181(2):595–608. <https://doi.org/10.1104/pp.19.00148>
- Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res* 30(1):325–327. <https://doi.org/10.1093/nar/30.1.325>
- Li HH, Liu X, An JP, Hao YJ, Wang XF, You CX (2017) Cloning and elucidation of the functional role of apple *MdLBD13* in anthocyanin biosynthesis and nitrate assimilation. *Plant Cell Tissue Organ Cult* 130(1):47–59. <https://doi.org/10.1007/s11240-017-1203-x>
- Liu HJ, Wang SF, Yu XB, Yu J, He XW, Zhang SL, Shou HX, Wu P (2005) ARL1, a LOB-domain protein required for adventitious root formation in rice. *Plant J* 43(1):47–56. <https://doi.org/10.1111/j.1365-3113X.2005.02434.x>
- Liu HZ, Cao MX, Chen XL, Ye MH, Zhao P, Nan YY, Li W, Zhang C, Kong LS, Kong NN, Yang CH, Chen Y, Wang DD, Chen Q (2019) Genome-wide analysis of the lateral organ boundaries domain (*LBD*) gene family in *Solanum tuberosum*. *Int J Mol Sci* 20(21):23. <https://doi.org/10.3390/ijms20215360>
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* 25(4):402–408. <https://doi.org/10.1006/meth.2001.1262>
- Lu Q, Shao F, Macmillan C, Wilson IW, van der Merwe K, Hussey SG, Myburg AA, Dong X, Qiu D (2018) Genomewide analysis of the lateral organ boundaries domain gene family in *Eucalyptus grandis* reveals members that differentially impact secondary growth. *Plant Biotechnol J* 16(1):124–136. <https://doi.org/10.1111/pbi.12754>
- Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* 290(5494):1151–1155. <https://doi.org/10.1126/science.290.5494.1151>
- Majer C, Hochholdinger F (2011) Defining the boundaries: structure and function of LOB domain proteins. *Trends Plant Sci* 16(1):47–52. <https://doi.org/10.1016/j.tplants.2010.09.009>
- Mangeon A, Lin W-c, Springer PS (2012) Functional divergence in the *Arabidopsis* LOB-domain gene family. *Plant Signal Behav* 7(12):1544–1547. <https://doi.org/10.4161/psb.22320>
- Miao Y, Laun T, Zimmermann P, Zentgraf U (2004) Targets of the WRKY53 transcription factor and its role during leaf senescence in *Arabidopsis*. *Plant Mol Biol* 55(6):853–867. <https://doi.org/10.1007/s11103-005-2142-1>
- Moore RC, Purugganan MD (2003) The early stages of duplicate gene evolution. *Proc Natl Acad Sci USA* 100(26):15682–15687. <https://doi.org/10.1073/pnas.2535513100>
- Naito T, Yamashino T, Kiba T, Koizumi N, Kojima M, Sakakibara H, Mizuno T (2007) A link between cytokinin and *ASL9* (*Asymmetric Leaves 2 Like 9*) that belongs to the *AS2/LOB* (*Lateral Organ Boundaries*) family genes in *Arabidopsis thaliana*. *Biosci Biotechnol Biochem* 71(5):1269–1278. <https://doi.org/10.1271/bbb.60681>
- Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D et al (2012) Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. *Nature* 492(7429):423–427. <https://doi.org/10.1038/nature11798>

- Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL (2015) StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnol* 33(3):290–295. <https://doi.org/10.1038/nbt.3122>
- Prince VE, Pickett FB (2002) Splitting pairs: The diverging fates of duplicated genes. *Nature Rev Genet* 3(11):827–837. <https://doi.org/10.1038/nrg928>
- Ramamoorthy R, Jiang S-Y, Kumar N, Venkatesh PN, Ramachandran S (2008) A comprehensive transcriptional profiling of the *WRKY* gene family in rice under various abiotic and phytohormone treatments. *Plant Cell Physiol* 49(6):865–879. <https://doi.org/10.1093/pcp/pcn061>
- Rubin G, Tohge T, Matsuda F, Saito K, Scheible W-R (2009) Members of the *LBD* family of transcription factors repress anthocyanin synthesis and affect additional nitrogen responses in *Arabidopsis*. *Plant Cell* 21(11):3567–3584. <https://doi.org/10.1105/tpc.109.067041>
- Santino A, Taurino M, De Domenico S, Bonsegna S, Poltronieri P, Pastor V, Flors V (2013) Jasmonate signaling in plant development and defense response to multiple (a)biotic stresses. *Plant Cell Rep* 32(7):1085–1098. <https://doi.org/10.1007/s00299-013-1441-2>
- Semiarti E, Ueno Y, Tsukaya H, Iwakawa H, Machida C, Machida Y (2001) The *ASYMMETRIC LEAVES2* gene of *Arabidopsis thaliana* regulates formation of a symmetric lamina, establishment of venation and repression of meristem-related homeobox genes in leaves. *Development* 128(10):1771–1783
- Shuai B, Reynaga-Pena CG, Springer PS (2002) The *Lateral Organ Boundaries* gene defines a novel, plant-specific gene family. *Plant Physiol* 129(2):747–761. <https://doi.org/10.1104/pp.010926>
- Teng R-M, Wang Y-X, Wang W-L, Li H, Shen W, Zhuang J (2018) Genome-wide identification, classification and expression pattern of *LBD* gene family in *Camellia sinensis*. *Biotechnol Biotechnol Equipment* 32(6):1387–1397. <https://doi.org/10.1080/13102818.2018.1521303>
- Thatcher LF, Kazan K, Manners JM (2012a) Lateral organ boundaries domain transcription factors: new roles in plant defense. *Plant Signal Behav* 7(12):1702–1704. <https://doi.org/10.4161/psb.22097>
- Thatcher LF, Powell JJ, Aitken EAB, Kazan K, Manners JM (2012b) The lateral organ boundaries domain transcription factor *LBD20* functions in fusarium wilt susceptibility and jasmonate signaling in *Arabidopsis*. *Plant Physiol* 160(1):407–418. <https://doi.org/10.1104/pp.112.199067>
- Wang D, Zhang Y, Zhang Z, Zhu J, Yu J (2010) KaKs_Calculator 2.0: A toolkit incorporating gamma-series methods and sliding window strategies. *Genom Proteom Bioinform* 8(1):77–80. [https://doi.org/10.1016/s1672-0229\(10\)60008-3](https://doi.org/10.1016/s1672-0229(10)60008-3)
- Wang X, Zhang S, Su L, Liu X, Hao Y (2013) A genome-wide analysis of the *LBD* (LATERAL ORGAN BOUNDARIES Domain) gene family in *Malus domestica* with a functional characterization of *MdLBD11*. *PLoS ONE* 8(2):e57044. <https://doi.org/10.1371/journal.pone.0057044>
- Xu C, Luo F, Hochholdinger F (2016) *LOB* domain proteins: Beyond Lateral Organ Boundaries. *Trends Plant Sci* 21(2):159–167. <https://doi.org/10.1016/j.tplants.2015.10.010>
- Yang T, Fang GY, He H, Chen J (2016) Genome-wide identification, evolutionary analysis and expression profiles of *Lateral Organ Boundaries Domain* Gene family in *Lotus japonicus* and *Medicago truncatula*. *PLoS ONE* 11(8):e0161901. <https://doi.org/10.1371/journal.pone.0161901>
- Yang H, Shi G, Du H, Wang H, Zhang Z, Hu D, Wang J, Huang F, Yu D (2017) Genome-wide analysis of soybean Lateral Organ Boundaries Domain-containing genes: a functional investigation of *GmLBD12*. *Plant Genome*. <https://doi.org/10.3835/plantgenom2016.07.0058>
- Yang Z, Ge X, Yang Z, Qin W, Sun G, Wang Z, Li Z, Liu J, Wu J, Wang Y, Lu L, Wang P, Mo H, Zhang X, Li F (2019) Extensive intraspecific gene order and gene structural variations in upland cotton cultivars. *Nat Commun* 10(1):2989. <https://doi.org/10.1038/s41467-019-10820-x>
- Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL (2012) Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinform* 13:134. <https://doi.org/10.1186/1471-2105-13-134>
- Yu J, Wang J, Lin W, Li SG, Li H, Zhou J, Ni PX, Dong W, Hu SN, Zeng CQ et al (2005) The genomes of *Oryza sativa*: A history of duplications. *Plos Biol* 3(2):266–281. <https://doi.org/10.1371/journal.pbio.0030038>
- Yu J, Jung S, Cheng C-H, Ficklin SP, Lee T, Zheng P, Jones D, Percy RG, Main D (2014) CottonGen: a genomics, genetics and breeding database for cotton research. *Nucleic Acids Res* 42(D1):D1229–D1236. <https://doi.org/10.1093/nar/gkt1064>
- Zentella R, Zhang Z-L, Park M, Thomas SG, Endo A, Murase K, Fleet CM, Jikumaru Y, Nambara E, Kamiya Y, Sun T-P (2007) Global analysis of *DELLA* direct targets in early gibberellin signaling in *Arabidopsis*. *Plant Cell* 19(10):3037–3057. <https://doi.org/10.1105/tpc.107.054999>
- Zhang H-B, Li Y, Wang B, Chee PW (2008) Recent advances in cotton genomics. *Int J Plant Genom* 2008:742304–742304. <https://doi.org/10.1155/2008/742304>
- Zhang JR, Tang W, Huang YL, Niu XL, Zhao Y, Han Y, Liu YS (2015) Down-regulation of a *LBD*-like gene, *OsIG1*, leads to occurrence of unusual double ovules and developmental abnormalities of various floral organs and megagametophyte in rice. *J Exp Bot* 66(1):99–112. <https://doi.org/10.1093/jxb/eru396>
- Zhang XY, He YQ, He WD, Su H, Wang YF, Hong GJ, Xu P (2019) Structural and functional insights into the *LBD* family involved in abiotic stress and flavonoid syntheses in *Camellia sinensis*. *Sci Rep* 9:14. <https://doi.org/10.1038/s41598-019-52027-6>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.