Genome-Wide Identification and Analysis of LBD Gene Family in Gossypium Hirsutum

Dianyu Liu[1,a], Chuiwen Qian[1,*]

[1]National Engineering Research Center of Genetic Medicine, Guangdong Provincial biotechnology drug & Engineering Technology Research Center, Guangdong Province Key Laboratory of Bioengineering Medicine, Department of Cell Biology, College of Life Science and Technology, Jinan University, Guangzhou 510632, China.

[a]513792654@qq.com, *69866880@qq.com

Abstract

LBD (lateral organ boundaries domain) protein is a kind of plant-specific transcription factor, which plays a key role in the regulation of plant growth and development, nutrition metabolism as well as response to adversity stress. However, at the genome level, study of GhLBD gene family has not been reported yet. In this study, by using bioinformatics methods, a total of 136 LBD genes were identified in Gossypium hirsutum. genome, and their molecular weight and isoelectric point were predicted. These genes were distributed on 24 chromosomes. The structure of this gene family was simple whose introns number was less than 2. According to the analysis of phylogenyetic relationship, GhLBD genes could be classified into 2 classes (Class and Class). This study might lay the foundation for verifying the function of LBD genes in Gossypium hirsutum.

Keywords

Gossypium hirsutum, LBD gene family, Phylogeny analysis, Bioinformatics.

1. Introduction

The Lateral organ boundary Domain (LBD) gene family is a specific transcription factor in plants and plays an important role in the development of lateral organs [1] and the nutrient absorption and metabolism of nitrogen. The N terminal of LBD gene contains a conserved LOB domain, and according to the difference of LOB domain [2], the LBD gene family can be divided into two types: Class I contains a conserved domain similar to zinc finger, CX2CX6CX3C[3], and a leucine zipper similar motif, LX6LX3LX6L, while class II contains only CX2CX6CX3C.

At present, the function of LBD gene has been verified in many species. Among them, AtLBD16, AtLBD18, AtLBD29 plays an important role in arabidopsis adventata root and lateral root development [4-6], and ARF7 and ARF19 can activate gene transcription by directly binding to autin response elements on the AtLBD16 gene promoter to form lateral roots [6]. AtLBD10 plays an important role in the development of Arabidopsis pollen [7]. The microsporogenesis and pollen development of Arabidopsis are related to AtLBD27 [8-9]. AtLBD37, AtLBD38 and AtLBD39 are involved in anthocyanin biosynthesis and nitrogen metabolism in Arabidopsis thaliana [10]. AtLBD30 and AtLBD36 play important roles in leaf formation and embryogenesis [11]. In addition to the function of LBD gene in model plant Arabidopsis tharabidopsis, there have also been reports on the function of LBD family gene in other plants. For example, apple MdLBD13 inhibits the synthesis of anthocyanin[12]. The number of female gametes in rice organs is regulated by OslG1[13].

With the development of sequencing technology, several species have been sequenced, providing a basis for the identification of the LBD gene family. In rice, pepper and Arabidopsis, 35, 45 and 43 related genes of LBD gene families were identified [1,14,15]. The genome sequencing of cotton, an important economic and oil crop in the world, has been completed. Although the identification and functional studies of LBD gene family have been carried out in many species, there are few reports on LBD gene family in upland cotton. Based on the recently published tM-1 genomic sequence of

the upland cotton genetic standard, this study made a comprehensive analysis of the physical and chemical properties, chromosome distribution, gene structure, and phylogenetic relationship of LBD of Gossypium hirsutum, laying a foundation for subsequent studies on the function of LBD gene in cotton.

2. Materials and Methods

2.1 Identification of LBD gene family members in Gossypium hirsutum

In Nanjing Agricultural University, state Key Laboratory of Crop Genetics and Germplasm Enhancement (http://mascotton.njau.edu.cn/Data/Genome_sequence.htm) to download the upland cotton tetraploid standard genome data of TM-1. Download the LBD seed file PF03195 using the Pfam (https://pfam.xfam.org/) online database. Protein sequences containing LOB conserved domains were identified by HMMSearch using HMMER3.0 software. Finally, all LBD protein sequences were submitted to Phmmer Search online Database and NCBI CDD (Conserved Domain Database) for verification, and finally all LBD gene family members of upland cotton were obtained. The online tool ExPASy Proteomics [16] (http://www.expasy.org) was used to predict the biological information of amino acid sequence of members of LBD family in Gossypium hirsutum, and the online tool Softberry (http://linux1.softberry.com) was used to analyze and predict the subcellular localization of LBD family.

2.2 Multiple sequence alignment and conserved domain characteristics of LBD protein in Gossypium hirsutum

By using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) online multiple sequence alignment tools of upland cotton LBD proteins in multiple sequence alignment, using the Jalview [17] (https://www.jalview.org/) software to extract conservative sequence, finally using WebLogo [18] online software display upland cotton LBD conservative area conservative of amino acids.

2.3 Phylogenetic tree construction and gene structure analysis

The LBD protein sequence of upland cotton was compared with that of the classified LBD protein sequence of Arabidopsis thaliana, and the phylogenetic tree was constructed based on the comparison results of Neighbour-Joining Method with MEGA-X [19] software (Bootstrap was set as 1000 times). The structure of LBD gene family was visualized using online software tool GSDS2.0 [20] (http://gsds.gaolab.org/). The online software MEME [21] was used to analyze the LBD conservative motif in Gossypium hirsutum, and the parameter was set as the maximum number of motif discoveries to be 10. Finally, Tbtools [22] was used to visualize the analysis results of MEME.

2.4 Cis-acting element analysis of LBD gene in Gossypium hirsutum

Extract LBD genes transcription start site ATG upstream 2000 bp sequence, the Plant CARE (http:// bioinformatics.psb.ugent.be/webtools/plantcare/html/) online database retrieval, analysis of cis-acting element. Finally visualize with Tbtools [22].

3. Results

3.1 Identification of LBD gene family members in Gossypium hirsutum

Based on the latest TM-1 protein sequence of Gossypium hirsutum, HMMER3.0 software was used to compare and search PF03195 files, and the candidate protein sequences were obtained. The conserved domain was verified by PHmmer Search and NCBI CDD database, and 136 LBD genes were finally identified in the Gossypium hirsutum TM-1 genome. LBD genes were named according to their positions on chromosomes (Table 1). By analyzing the physical and chemical properties of the LBD gene family of Gossypium hirsutum, it was found that the length of LBD protein was 60~ 304AA.The number of exons was 1 (24 genes), 2(105 genes) and 3(7 genes), indicating that LBD gene in upland cotton mainly existed in the form of 1 exon and 2 exons. The relative molecular weight of protein was 6.71 kD ~3.31 kD. The isoelectric point of Gh_D10G2300 is the smallest, and the value is 4.75. Gh_Sca115125G01 and Gh_Sca179406G01 have a maximum isoelectric point of 9.51 and an average of 7.35. Subcellular localization detection showed that most of THE LBD was located

in the nucleus and the rest were located outside the cell. These results indicate that the diversity of LBD protein sequence may be related to the adaptation of different functions.

Table 1 Analysis of physicochemical properties of LBD family members in the Gossypium hirsutum

Gene	Chromosome	Protein length/aa	Length of CDS/bp	pI	Molecular weight	Subcellular localization
Gh_A01G0756	A01	175	528	7.65	19291.82	Nuclear
Gh_A01G1555	A01	221	666	9.1	24485.27	Extracellular
Gh_A03G0198	A03	225	678	8.84	24317.61	Nuclear
Gh_A03G0566	A03	294	885	5.51	33294.12	Extracellular
Gh_A03G1288	A03	231	696	8.67	25426.98	Extracellular
Gh_A03G1613	A03	216	651	5.88	22488.31	Extracellular
Gh_A03G2052	scaffold511_A03	176	531	6.35	19596.25	Nuclear
Gh_A05G0943	A05	155	468	8.24	17267.58	Nuclear
Gh_A05G1015	A05	151	456	8.85	17000.27	Nuclear
Gh_A05G1201	A05	171	516	8.57	1850.58	Nuclear
Gh_A05G1333	A05	278	837	8.91	31126.72	Extracellular
Gh_A05G1814	A05	234	705	8.91	26013.44	Extracellular
Gh_A05G2399	A05	228	687	8.76	24800.46	Nuclear
Gh_A05G3105	A05	173	522	6.28	18943.13	Nuclear
Gh_A05G3659	scaffold1166_A05	208	627	5.85	22440.53	Nuclear
Gh_A05G3778	scaffold1219_A05	138	416	5.21	14682.52	Nuclear
Gh_A06G1223	A06	170	513	7.59	18544.25	Nuclear
Gh_A06G1421	A06	185	558	6.78	20572.46	Nuclear
Gh_A06G1674	A06	170	513	8.4	19449.51	Nuclear
Gh_A06G1675	A06	164	495	8.91	18958.41	Nuclear
Gh_A06G1676	A06	166	501	8.1	19143.87	Nuclear
Gh_A06G1941	scaffold1295_A06	204	615	6.29	22269.46	Nuclear
Gh_A07G1059	A07	202	609	5.9	2205.15	Nuclear
Gh_A07G1922	A07	211	636	6.48	2308.37	Nuclear
Gh_A07G1923	A07	226	681	5.52	24665.24	Nuclear
Gh_A07G2309	scaffold1917_A07	270	813	7.57	29881.33	Extracellular
Gh_A08G0081	A08	202	609	5.15	21847.86	Nuclear
Gh_A08G0556	A08	230	693	6.42	25054.92	Extracellular
Gh_A08G0557	A08	204	615	7.59	22500.51	Nuclear
Gh_A08G1109	A08	167	504	8.37	18740.57	Nuclear
Gh_A08G1200	A08	253	762	8.14	25910.29	Nuclear
Gh_A08G1804	A08	171	516	6.7	1892.96	Nuclear
Gh_A08G2362	scaffold1933_A08	156	471	8.32	17032.36	Nuclear
Gh_A08G2363	scaffold1933_A08	216	651	5.69	2338.51	Nuclear
Gh_A09G0283	A09	160	483	5.43	17808.99	Nuclear
Gh_A09G0469	A09	240	723	8.88	25734.55	Extracellular
Gh_A09G0988	A09	193	582	5.69	2111.22	Nuclear
Gh_A10G0436	A10	171	516	8.24	185.25	Nuclear
Gh_A10G0834	A10	303	912	6.89	33119.32	Nuclear
Gh_A10G1992	A10	217	654	4.82	24102.28	Nuclear
Gh_A10G2146	A10	173	522	5.51	19019.61	Nuclear
Gh_A10G2153	A10	161	486	9.18	1810.94	Nuclear
Gh_A10G2252	scaffold2456_A10	246	741	5.54	28316.51	Extracellular
Gh_A11G0447	A11	224	675	8.74	24222.69	Extracellular
Gh_A11G0562	A11	226	681	8.39	25278.71	Extracellular
Gh_A11G0617	A11	163	492	7.03	18277.66	Nuclear
Gh_A11G0741	A11	186	561	8.57	20457.23	Nuclear
Gh_A11G0891	A11	171	516	5.4	19463.46	Nuclear
Gh_A11G1186	A11	201	606	8.17	22206.27	Nuclear
Gh_A11G1187	A11	234	705	6.17	25679.53	Extracellular
Gh_A11G1760	A11	206	621	8.3	2279.76	Nuclear
Gh_A11G1963	A11	232	699	9.1	2551.41	Extracellular
Gh_A11G2675	A11	252	759	7.24	2767.42	Extracellular
Gh_A11G2723	A11	227	684	8.21	2462.46	Nuclear
Gh_A11G2724	A11	209	630	7.32	22809.59	Nuclear
Gh_A11G3014	scaffold2728_A11	297	894	5.81	33567.34	Extracellular
Gh_A12G0495	A12	302	909	7.49	32953.49	Extracellular
Gh_A12G0643	A12	296	891	6.18	3313.54	Nuclear
Gh_A12G0696	A12	232	699	8.72	2548.09	Extracellular
Gh_A12G2355	A12	137	414	9.19	15605.82	Nuclear
Gh_A12G2596	scaffold3366_A12	227	684	8.5	24023.51	Nuclear
Gh_A12G2597	scaffold3366_A12	206	621	7.55	22315.28	Nuclear
Gh_A13G0015	A13	211	636	5.17	2294.82	Nuclear
Gh_A13G0040	A13	213	642	8.45	23676.13	Nuclear
Gh_A13G0132	A13	172	519	6.99	193.16	Nuclear
Gh_D01G1806	D01	221	666	9.1	24426.24	Extracellular
Gh_D02G1728	D02	231	696	8.67	2544.11	Extracellular
Gh_D02G2027	D02	216	651	5.09	22595.41	Extracellular

Gh_D03G0845	D03	296	891	5.51	33527.35	Extracellular
Gh_D03G1361	D03	176	531	6.02	1958.72	Nuclear
Gh_D03G1383	D03	223	672	8.61	24098.36	Nuclear
Gh_D04G0127	D04	214	645	9.07	23376.74	Nuclear
Gh_D04G0535	D04	173	522	6.28	18953.17	Nuclear
Gh_D04G0807	D04	205	618	5.02	22241.28	Nuclear
Gh_D04G1971	scaffold3962_D04	205	618	5.02	22241.28	Nuclear
Gh_D05G1028	D05	155	468	8.24	17340.68	Nuclear
Gh_D05G1133	D05	151	456	8.95	1689.49	Nuclear
Gh_D05G1378	D05	171	516	8.58	1851.49	Nuclear
Gh_D05G1502	D05	277	834	8.35	31029.62	Extracellular
Gh_D05G2010	D05	234	705	8.75	26073.49	Extracellular
Gh_D05G2053	D05	153	465	8.26	16675.69	Nuclear
Gh_D05G2664	D05	228	687	8.76	24791.37	Nuclear
Gh_D06G1540	D06	168	507	8.19	1832.52	Nuclear
Gh_D06G1770	D06	185	558	6.48	20633.54	Nuclear
Gh_D06G1825	D06	127	384	9.38	13934.25	Nuclear
Gh_D06G2049	D06	170	513	8.1	19450.58	Nuclear
Gh_D06G2050	D06	142	429	8.99	16205.83	Extracellular
Gh_D06G2304	scaffold4090_D06	204	615	5.26	21909.83	Extracellular
Gh_D07G0184	D07	270	813	6.94	29951.46	Extracellular
Gh_D07G1151	D07	203	612	6.28	22201.25	Nuclear
Gh_D07G2145	D07	211	636	6.42	2303.02	Nuclear
Gh_D07G2146	D07	247	744	5.62	27078.94	Extracellular
Gh_D08G0122	D08	203	612	5.6	22075.14	Nuclear
Gh_D08G0649	D08	226	681	6.42	24646.54	Nuclear
Gh_D08G0650	D08	229	690	6.58	251.2	Nuclear
Gh_D08G0651	D08	204	615	6.64	22528.48	Nuclear
Gh_D08G1390	D08	167	504	7.73	18749.55	Nuclear
Gh_D08G1484	D08	252	759	8.14	25843.16	Nuclear
Gh_D08G2036	D08	154	465	8.32	16763.13	Nuclear
Gh_D08G2037	D08	216	651	5.69	233.51	Nuclear
Gh_D08G2158	D08	173	522	6.7	19103.76	Nuclear
Gh_D09G0285	D09	160	483	5.43	17793.99	Nuclear
Gh_D09G0477	D09	239	720	8.77	25596.34	Extracellular
Gh_D09G1008	D09	193	582	5.69	2110.01	Nuclear
Gh_D10G0452	D10	171	516	8.25	1855.31	Nuclear
Gh_D10G0555	D10	246	741	5.63	28321.53	Extracellular
Gh_D10G0922	D10	304	915	6.89	33194.28	Nuclear
Gh_D10G2300	D10	217	654	4.75	24178.33	Nuclear
Gh_D10G2393	D10	173	522	5.51	19003.61	Nuclear
Gh_D10G2400	D10	168	507	9.2	1889.71	Nuclear
Gh_D11G0311	D11	295	888	5.87	33285.15	Extracellular
Gh_D11G0518	D11	217	654	8.36	23468.81	Extracellular
Gh_D11G0645	D11	226	681	8.39	25272.75	Extracellular
Gh_D11G0704	D11	163	492	7.03	18227.64	Nuclear
Gh_D11G1038	D11	171	516	5.92	19478.58	Nuclear
Gh_D11G1343	D11	201	606	8.17	2222.43	Nuclear
Gh_D11G1344	D11	234	705	6.17	25612.39	Extracellular
Gh_D11G2012	D11	232	699	8.87	2553.84	Extracellular
Gh_D11G3036	D11	252	759	7.24	2764.72	Extracellular
Gh_D11G3076	D11	225	678	8.51	24342.33	Nuclear
Gh_D12G0503	D12	302	909	6.48	32921.34	Extracellular
Gh_D12G0701	D12	232	699	8.51	2543.14	Extracellular
Gh_D12G1018	D12	206	621	6.57	22314.25	Nuclear
Gh_D12G1019	D12	229	690	7.64	24034.45	Nuclear
Gh_D12G2492	D12	137	414	8.97	15591.79	Nuclear
Gh_D13G0031	D13	209	630	5.14	22487.67	Nuclear
Gh_D13G0055	D13	176	531	6.5	19210.87	Nuclear
Gh_D13G0153	D13	172	519	6.51	19291.93	Nuclear
Gh_Sca034902G01	scaffold34902	114	342	9.14	11807.63	Nuclear
Gh_Sca036239G01	scaffold36239	63	189	8.82	711.41	Nuclear
Gh_Sca082187G01	scaffold82187	102	309	8.78	11105.89	Nuclear
Gh_Sca106972G01	scaffold106972	96	288	8.33	10681.28	Nuclear
Gh_Sca115125G01	scaffold115125	60	180	9.51	6709.88	Nuclear
Gh_Sca118136G01	scaffold118136	90	270	7.58	9922.43	Nuclear
Gh_Sca179406G01	scaffold179406	73	219	9.51	8227.38	Nuclear
Gh_Sca196076G01	scaffold196076	61	183	9.12	6663.71	Nuclear

3.2 Multiple sequence alignment and conserved domain characteristics of LBD protein in Gossypium hirsutum

Multi-sequence alignment of Gossypium hirsutum LBD protein was performed using Clustal Omega online multi-sequence alignment tool to analyze its conserved domain.Typical LOB domains were found at the N-terminal of Gossypium hirsutum LBD (Fig. 1).

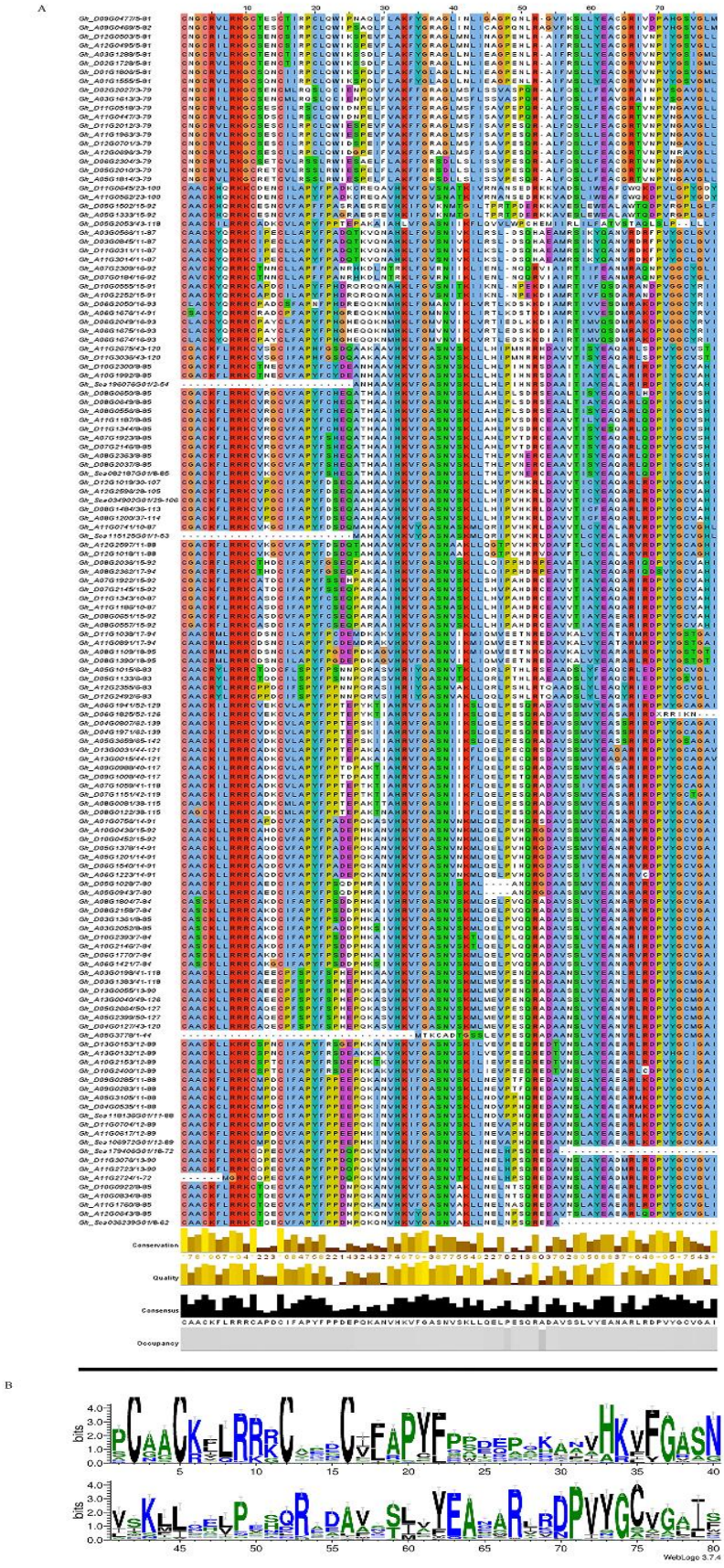

Fig.1.The conservation domain analysis of LBD proteins in G. hirsutum

A: Extraction of conserved domains of LBD proteins.

B: The conserved domains showed by WebLogo.

3.3 Chromosome distribution of LBD gene family members in Gossypium hirsutum

The MapGeneChrom (http://mg2c.IASK.in/mg2C_v2.0/) online tool was used to display the chromosome distribution information of the members of the Gossypium hirsutum LBD family according to the gene location information (Figure 2). Fifty-four LBD genes were distributed in subgroup A, 58 genes were distributed in subgroup D, and 24 LBD genes were matched on the Scaffold (Table 1). Members of the LBD gene family in Gossypium hirsutum were not uniform on chromosomes, and the results showed that they were most distributed on A11 and D11 chromosomes, with 11 and 10 LBD genes respectively. LBD gene was the least distributed on Chromosome D01, with only 1 gene.

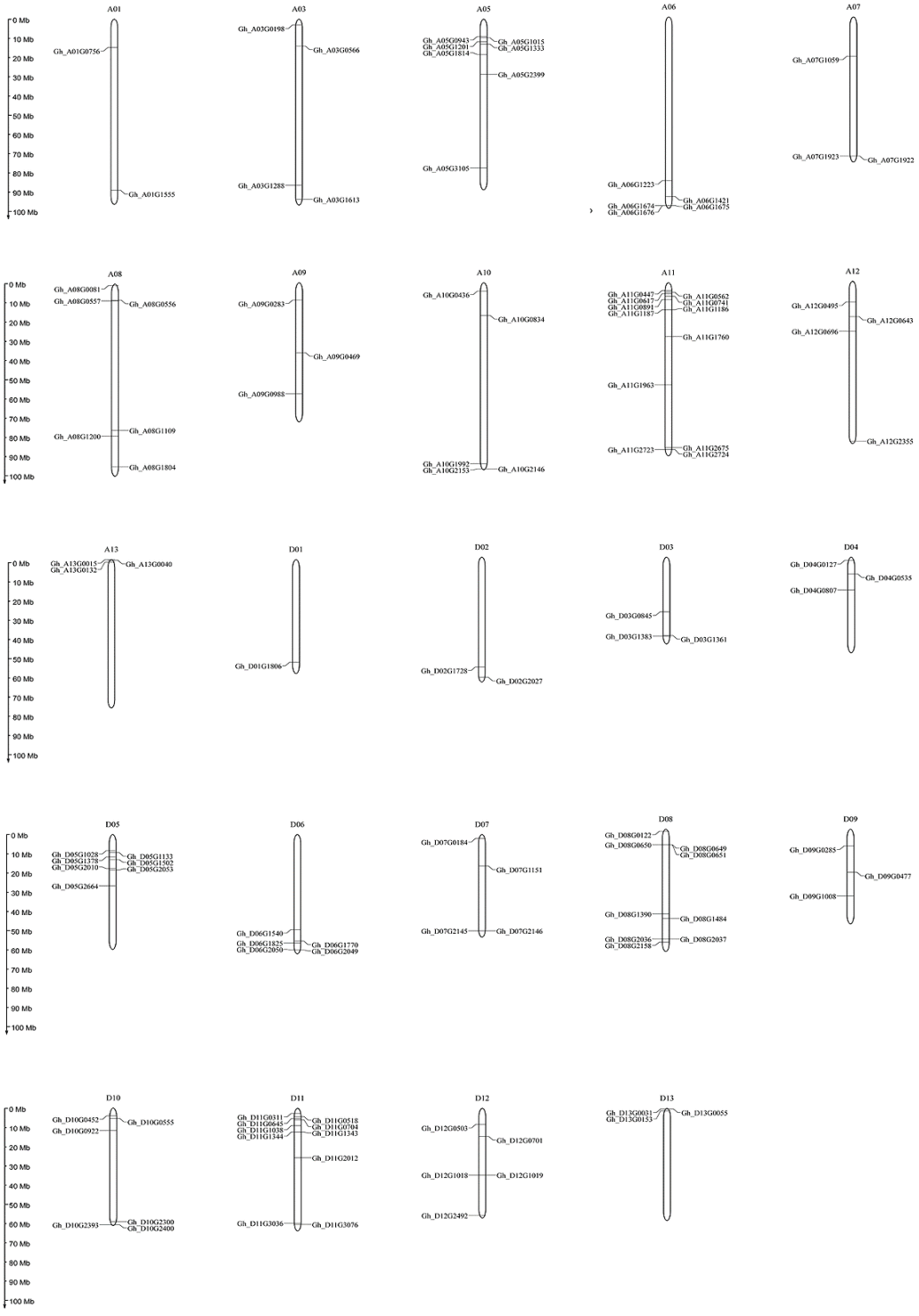

Fig.2. ChromosomaI distribution of LBD genes in G. hirsutum

3.4 Phylogenetic analysis of LBD gene in Gossypium hirsutum

In order to understand the evolutionary relationship within the LBD gene family of Gossypium hirsutum, 136 LBD protein sequences of Gossypium hirsutum and 38 LBD protein sequences of Arabidopsis thaliana were constructed by adjacency method. The LBD proteins of upland cotton were classified according to the classification relationship of Arabidopsis LBD family proteins (Fig.3). The analysis results showed that 136 LBD protein sequences could be divided into Class I and Class II. Among them, a small number of upland cotton LBD proteins (19) were classified into Class II, while the remaining homologous gene members' LBD proteins (117) and the division were classified into Class I.

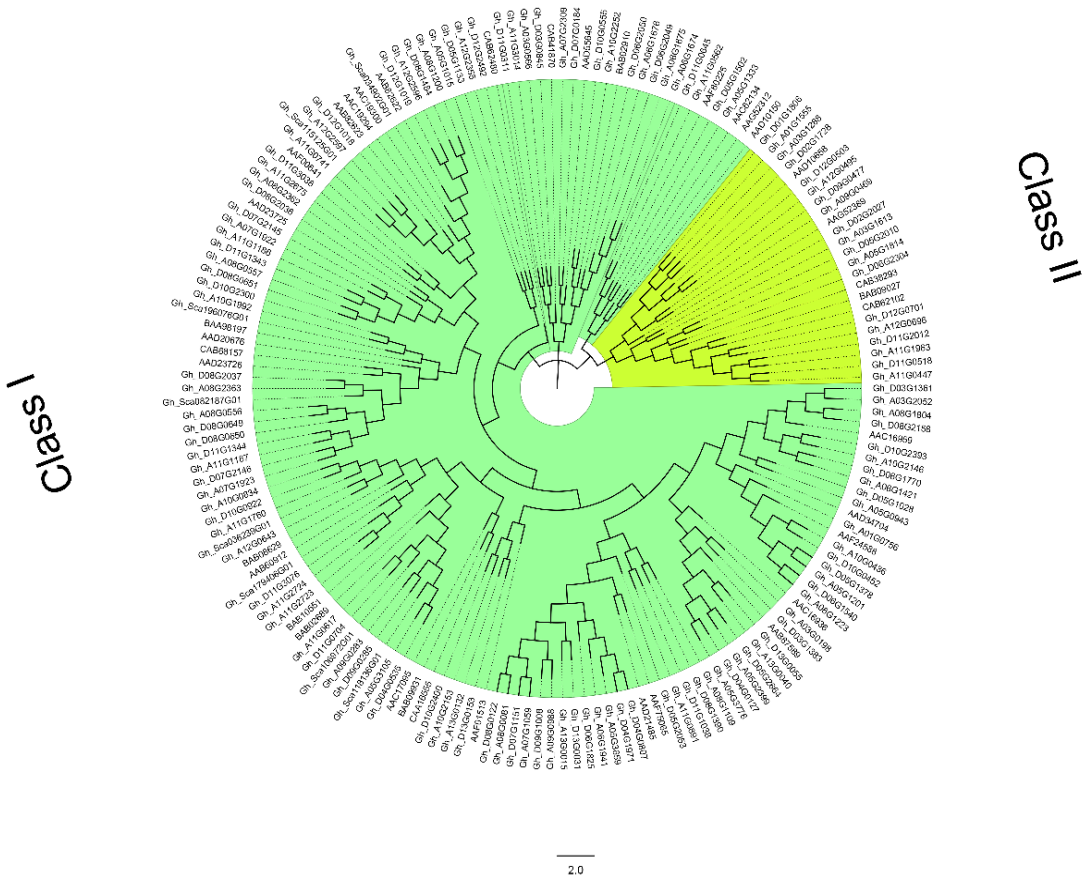

Fig.3. Phylogenetic tree of Dof protein sequences in G. hirsutum and Arabidopsis thaliana

3.5 LBD gene family structure in Gossypium hirsutum

In order to understand the structure of LBD gene in upland cotton, the intron/exon and the conserved motif of the coding product of LBD gene were compared according to the phylogenetic relationship (Fig.4). It was found that most members of the LBD gene family of Gossypium hirsutum did not contain or only contained 1 intron(Fig.4C). The conserved amino acid sequence of cotton LBD was analyzed by using MEME online software, and 10 conserved motifs were found (Fig.4B). Motif1 motif represents the conserved domain of LBD protein, which is distributed in all LBD and has the largest number, followed by motif 3(Fig.4B).The similarities and differences between the gene structure and the conserved motif reflect the relative conservatism of the LBD gene family during the long evolutionary process and the diversity generated for adaptation to the environment.

3.6 Cis-acting element analysis

Plant CARE online software was used to analyze cis-acting elements in the promoter region of 2000bp upstream of the transcription start site of LBD gene in upland cotton, and it was found that

all LBD genes contained AT-box. In addition to these conservative elements, there are also four types of cis-acting elements in the LBD gene promoter region of upland cotton :(1) light regulatory elements, including Box 4, G-box, TCCC-motif, and AE-box.(2) Plant growth and development regulatory elements, including cat-box and circadian, are mainly involved in the regulation of meristems.(3) Plant hormone response elements mainly included the methyl jasmonate response element TGACG-motif and CGTCA-motif, Abscisic acid response element (ABRE), gibberellin response element P-box, etc.(4) Adversity stress response elements, including low-temperature response element (LTR), drought response element (MBS), Anaerobic response element (ARE). These analysis results indicated that LBD gene family members of upland cotton play an important role in plant growth and development, plant hormone response and different stress.

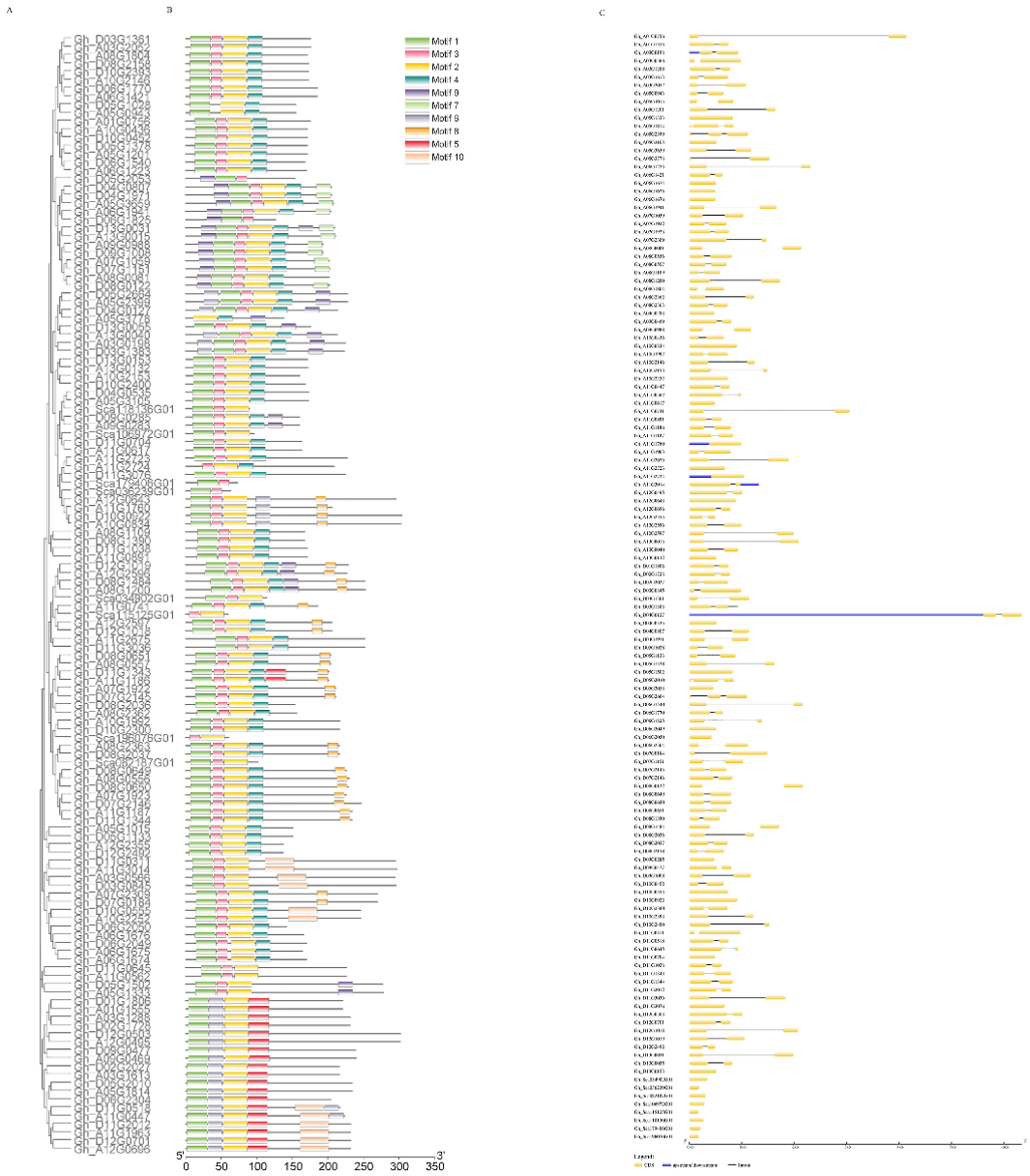

Fig.4. Phylogenetic tree, conserved motifs and gene structures of LBD protein sequences in G.hirsutum

A: the phylogenetic tree of LBD protein sequences classifications in G.hirsutum.

B: motif structures, conserved motifs are numbered and indicated in colored boxes.

C: structures of the LBD genes in G.hirsutum. Bars: exons; black lines: introns; blue: upstream/downstream..

Fig.5. Cis-acting elements of the Gossypium hirsutum LBD protein family

4. Discussion

LBD is a family transcription factor unique to plants. Since the first LBD gene was discovered in Arabidopsis thaliana in 2002, many LBD genes have been cloned and their functions verified in model plants Arabidopsis thaliana, rice and maize. Previous studies mainly focused on the growth and development of plant tissues and organs, such as the development of roots, stems, leaves, embryos, flowers and inflorescence, etc. However, there has been no reported study on LBD gene in uplands so far.

In this study, through genome-wide bioinformatics analysis of Gossypium hirsutum, a total of 136 members of the LBD protein family of 2 types of upland cotton were identified. LBD is a kind of plant evolutionary conservative family. LBD gene has a simple structure, a small number of exons (1-3), and stable structure. Because the number of the upland cotton gene family is more than that of the 36 genes of Arabidopsis thaliana.it indicates that the LBD gene family has been expanded due to gene replication events during the evolution. It was found that 136 Gossypium hirsutum LBD genes were unevenly distributed on 24 chromosomes through chromosome localization.

LBD family genes play an important role in plant growth and development, so revealing the function of LBD genes in plants will be the focus of future research. This study to identify 136 upland cotton

LBD genes in the chromosome location, physical and chemical properties, evolutionary relationships, intron-exon structure, conservative motif and LOB domain characteristics analysis of structure, the function of land cotton LBD genes for the further research in the future laid a certain foundation and is expected to offer wood fiber traits such as improved some genetic information.

References

[1] Shuai, B., Reynaga-Pena, C. G. & Springer, P. S. The lateral organ boundaries gene defines a novel, plant-specific gene family. Plant Physiol 129, 747-761, doi:10.1104/pp.010926 (2002).

[2] Albinsky, D. et al. Metabolomic screening applied to rice FOX Arabidopsis lines leads to the identification of a gene-changing nitrogen metabolism. Mol Plant 3, 125-142, doi:10.1093/ mp/ ssp069 (2010).

[3] Matsumura, Y., Iwakawa, H., Machida, Y. & Machida, C. Characterization of genes in the ASYMMETRIC LEAVES2/LATERAL ORGAN BOUNDARIES (AS2/LOB) family in Arabidopsis thaliana, and functional and molecular comparisons between AS2 and other family members. Plant Journal 58, 525-537, doi:10.1111/j.1365-313X.2009. 03797.x (2009).

[4] Fan, M., Xu, C., Xu, K. & Hu, Y. LATERAL ORGAN BOUNDARIES DOMAIN transcription factors direct callus formation in Arabidopsis regeneration. Cell Res 22, 1169-1180, doi:10.1038/cr.2012.63 (2012).

[5] Lee, H. W., Kim, N. Y., Lee, D. J. & Kim, J. LBD18/ASL20 regulates lateral root formation in combination with LBD16/ASL18 downstream of ARF7 and ARF19 in Arabidopsis. Plant Physiol 151, 1377-1389, doi:10.1104/pp.109.143685 (2009).

[6] Okushima, Y., Fukaki, H., Onoda, M., Theologis, A. & Tasaka, M. ARF7 and ARF19 regulate lateral root formation via direct activation of LBD/ASL genes in Arabidopsis. Plant Cell 19, 118-130, doi:10.1105/tpc.106.047761 (2007).

[7] Kim, M. J., Kim, M. & Kim, J. Combinatorial interactions between LBD10 and LBD27 are essential for male gametophyte development in Arabidopsis. Plant Signal Behav 10, e1044193, doi:10.1080/15592324.2015.1044193 (2015).

[8] Kim, M. J., Kim, M. & Kim, J. LATERAL ORGAN BOUNDARIES DOMAIN (LBD) 10 interacts with SIDECAR POLLEN/LBD27 to control pollen development in Arabidopsis (vol 81, pg 794, 2015). Plant Signaling & Behavior 10, doi:ARTN e104419310.1080/ 15592324.2015. 1044193 (2015).

[9] Kim, M., Kim, M. J., Pandey, S. & Kim, J. Expression and Protein Interaction Analyses Reveal Combinatorial Interactions of LBD Transcription Factors During Arabidopsis Pollen Development. Plant Cell Physiol 57, 2291-2299, doi:10.1093/pcp/pcw145 (2016).

[10] Rubin, G., Tohge, T., Matsuda, F., Saito, K. & Scheible, W. R. Members of the LBD family of transcription factors repress anthocyanin synthesis and affect additional nitrogen responses in Arabidopsis. Plant Cell 21, 3567-3584, doi:10.1105/tpc.109.067041 (2009).

[11] Rast, M. I. & Simon, R. Arabidopsis JAGGED LATERAL ORGANS acts with ASYMMETRIC LEAVES2 to coordinate KNOX and PIN expression in shoot and root meristems. Plant Cell 24, 2917-2933, doi:10.1105/tpc.112.099978 (2012).

[12] Li, H. H. et al. Cloning and elucidation of the functional role of apple MdLBD13 in anthocyanin biosynthesis and nitrate assimilation. Plant Cell Tiss Org 130, 47-59, doi:10.1007/s11240-017-1203-x (2017).

[13] Zhang, J. et al. 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, 99-112, doi:10.1093/jxb/eru396 (2015).

[14] Majer, C. & Hochholdinger, F. Defining the boundaries: structure and function of LOB domain proteins. Trends Plant Sci 16, 47-52, doi:10.1016/ j.tplants.2010.09.009 (2011).

[15] Yang, Y., Yu, X. B. & Wu, P. Comparison and evolution analysis of two rice subspecies LATERAL ORGAN BOUNDARIES domain gene family and their evolutionary characterization from Arabidopsis. Mol Phylogenet Evol 39, 248-262, doi:10.1016/ j.ympev. 2005.09.016 (2006).

[16] Artimo, P. et al. ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 40, W597-603, doi:10.1093/nar/gks400 (2012).

[17] Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189-1191, doi:10.1093/ bioinformatics/btp033 (2009).

[18] Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. Genome Res 14, 1188-1190, doi:10.1101/gr.849004 (2004).

[19] Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35, 1547-1549, doi:10.1093/molbev/msy096 (2018).

[20] Hu, B. et al. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31, 1296-1297, doi:10.1093/bioinformatics/btu817 (2015).

[21] Bailey, T. L. et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37, W202-208, doi:10.1093/nar/gkp335 (2009).

[22] Chen, C. et al. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant 13, 1194-1202, doi: 10.1016/j.molp.2020.06.009 (2020).