Genome-wide identification of CBL family and expression analysis of *CBLs* in response to potassium deficiency in cotton

Tingting Lu [1, 2] , Gaofeng Zhang [1] , Lirong Sun [1] , Ji Wang [1] , Fushun Hao [Corresp. 1]

[1] College of Life Sciences, Henan University, State Key Laboratory of Cotton Biology, Henan Key Laboratory of Plant Stress Biology, Kaifeng, Henan, China
[2] Henan University of Animal Husbandry and Economy, College of Pharmacetical Engineering, Zhengzhou, Henan, China

Corresponding Author: Fushun Hao
Email address: haofsh@henu.edu.cn

Calcineurin B-like (CBL) proteins, as calcium sensors, play pivotal roles in plant responses to diverse abiotic stresses and in growth and development through interaction with CBL-interacting protein kinases (CIPKs). However, knowledge about functions and evolution of CBLs in *Gossypium* plants is scarce. Here, we conducted a genome-wide survey and identified 13, 13 and 22 CBL genes in the progenitor diploid *Gossypium arboreum* and *Gossypium raimondii*, and the cultivated allotetraploid *Gossypium hirsutum*, respectively. Analysis of physical properties, chromosomal locations, conserved domains and phylogeny indicated rather conserved nature of CBLs among the three *Gossypium* species. Moreover, these CBLs have closer genetic evolutionary relationship with the CBLs from cocoa than with those from other plants. Most CBL genes underwent evolution under purifying selection in the 3 *Gossypium* plants. Additionally,nearly all *G. hirsutum* CBL (GhCBL) genes were expressed in the root, stem, leaf, flower and fiber. Many *GhCBLs* were preferentially expressed in the flower while several *GhCBLs* were mainly expressed in roots. Expression patterns of GhCBL genes in response to potassium deficiency were also studied. The expression of most *GhCBLs* were moderately induced in roots after treatments with low-potassium stress. Yeast two-hybrid experiments indicated that GhCBL1-2, GhCBL1-3, GhCBL4-4, GhCBL8, GhCBL9 and GhCBL10-3 interacted with GhCIPK23, respectively. Our results provided a comprehensive view of the *CBLs* and valuable information for researchers to further investigate the roles and functional mechanisms of the CBLs in *Gossypium*.

1 **Title page**

2 **Title:**

3 **Genome-wide identification of CBL family and expression analysis of *CBLs* in response to**

4 **potassium deficiency in cotton**

5 **Authors:**

6 **Tingting Lu[1,2]*, Gaofeng Zhang[1]*, Lirong Sun[1], Ji Wang[1] and Fu-Shun Hao[1]**

7 **[1]State Key Laboratory of Cotton Biology, Henan Key Laboratory of Plant Stress Biology,**

8 **College of Life Sciences, Henan University, Kaifeng 475004, China**

9 **[2]Henan University of Animal Husbandry and Economy, Zhengzhou 450011, China**

10 *** These authors contributed equally to this work**

11 **Corresponding author**

12 **Fu-Shun Hao, haofsh@henu.edu.cn**

13 **Genome-wide identification of CBL family and expression analysis of *CBLs* in response to**

14 **potassium deficiency in cotton**

15 **Tingting Lu[1,2*], Gaofeng Zhang[1*], Lirong Sun[1], Ji Wang[1] and Fu-Shun Hao[1]**

16 **1 State Key Laboratory of Cotton Biology, Henan Key Laboratory of Plant Stress Biology,**

17 **College of Life Sciences, Henan University, Kaifeng 475004, China**

18 **2 Henan University of Animal Husbandry and Economy, Zhengzhou 450011, China**

19 **ABSTRACT**

20 Calcineurin B-like (CBL) proteins, as calcium sensors, play pivotal roles in plant responses to

21 diverse abiotic stresses and in growth and development through interaction with CBL-interacting

22 protein kinases (CIPKs). However, knowledge about functions and evolution of CBLs in

23 *Gossypium* plants is scarce. Here, we conducted a genome-wide survey and identified 13, 13 and

24 22 CBL genes in the progenitor diploid *Gossypium arboreum* and *Gossypium raimondii*, and the

25 cultivated allotetraploid *Gossypium hirsutum*, respectively. Analysis of physical properties,

26 chromosomal locations, conserved domains and phylogeny indicated rather conserved nature of

27 CBLs among the three *Gossypium* species. Moreover, these CBLs have closer genetic

28 evolutionary relationship with the CBLs from cocoa than with those from other plants. Most

29 CBL genes underwent evolution under purifying selection in the 3 *Gossypium* plants.

30 Additionally, nearly all *G. hirsutum* CBL (GhCBL) genes were expressed in the root, stem, leaf,

31 flower and fiber. Many *GhCBLs* were preferentially expressed in the flower while several

32 *GhCBLs* were mainly expressed in roots. Expression patterns of GhCBL genes in response to

33 potassium deficiency were also studied. The expression of most *GhCBLs* were moderately

34 induced in roots after treatments with low-potassium stress. Yeast two-hybrid experiments

35 indicated that GhCBL1-2, GhCBL1-3, GhCBL4-4, GhCBL8, GhCBL9 and GhCBL10-3

36 interacted with GhCIPK23, respectively. Our results provided a comprehensive view of the *CBLs*

37 and valuable information for researchers to further investigate the roles and functional

38 mechanisms of the CBLs in *Gossypium*.

39   **Keywords** *Gossypium*; calcineurin B-like proteins (CBLs); gene family; phylogeny; gene

40   expression

INTRODUCTION

42   Calcium ion (Ca^{2+}) plays pivotal roles in mediating and regulating many fundamental growth

43   and developmental processes and in response to various environmental stimuli (Luan, 2009;

44   Kudla et al., 2010; Sarwat et al., 2013). The Ca^{2+} signals are primarily perceived by some Ca^{2+}

45   sensors including Ca^{2+} dependent protein kinases, calmodulins and calcineurin B-like proteins

46   (CBLs), and then are transmitted by these sensors to downstream targets to initiate diverse

47   cellular responses (Luan, 2009; Kudla, et al., 2010; Sarwat et al., 2013).

48   CBLs are proteins sharing sequence similarity with the B subunit of calcineurin B in yeast and

49   neuronal calcium sensors in animals (Kudla et al., 1999). Each CBL has at least three EF

50   domains and Ca^{2+}-binding sites (Mohanta et al., 2015; Mao et al., 2016). CBLs relay Ca^{2+} signals

51   through interaction with and activation of the CBL-interacting protein kinases (CIPKs).

52   Moreover, CBL-CIPK has been demonstrated to serve as an essential signaling network

53   regulating plant responses to multiple abiotic stresses such as salinity, K^+ deficiency, excess of

54   Mg^{2+} and drought (Sanyal et al., 2015; Thoday-Kennedy et al., 2015; Mao et al., 2016). It also

55   modulates growth and development, absorption and/or transport of nitrate, ammonium and iron,

56   sustaining of H^+ homeostasis, and transduction of reactive oxygen species signals in plants

57   (Sanyal et al., 2015; Thoday-Kennedy et al., 2015; Mao et al., 2016).

58   In *Arabidopsis*, 10 genes (*CBL1-10*) encoding CBL proteins have been found (Kolukisaoglu et

59   al., 2004). *CBL1* and *CBL9* were reported to positively regulate the uptake and transport of K^+,

60   NO_3^-, NH_4^+, aluminum and iron, and the promotion of stomatal opening (Li et al., 2006; Xu et al.,

61   2006; Ho et al., 2009; Mao et al., 2016; Tian et al., 2016; Ligaba-Osena et al., 2017; Straub et al.,

62   2017). *CBL1* and *CBL9* also affect abscisic acid (ABA)-induced stomatal closure and ROS

63   signaling (Pandey et al., 2004; Cheong et al., 2007; Drerup et al., 2013). *CBL2* plays a negative

64   role in the activation of plasma membrane (PM) H^+-ATPase (Fuglsang et al., 2007). Moreover,

65 *CBL2* and *CBL3* are cooperatively implicated in sequestering Mg^{2+} and modulation of pollen

66 germination and tube growth (Steinhorst et al., 2015; Tang et al., 2015). *CBL3* are also engaged

67 in K^+ distribution and translocation (Liu et al., 2013). *CBL4* was proven to be a crucial regulator

68 for excluding Na^+ and translocation of AKT2 (*Arabidopsis* K^+ transporter 2) from endoplasmic

69 reticulum to PM (Held et al., 2011). *CBL10* is involved in enhancing salt tolerance, stimulating

70 K^+ absorption, and modulating GTPase activity (Kim et al., 2007; Ren et al., 2013; Cho et al.,

71 2016). In cotton (*Gossypium hirsutum*), *GhCBL2* and *GhCBL3* appear to modulate fiber

72 elongation (Gao et al., 2008). Many *CBLs* in other plant species also play important parts in

73 regulating the responses to various abiotic stress as well as growth and development (Li et al.,

74 2014a; Thoday-Kennedy et al., 2015).

75 In recent years, multiple CBL gene families have been identified at genome-wide levels in rice,

76 maize, wheat and other plants (Kolukisaoglu et al., 2004; Zhang et al., 2014; Sun et al., 2015; Li

77 et al., 2016; Zhang et al., 2016). Some conserved domains such as EF-hands, myristoylation and

78 palmitoylation sites were discovered in CBLs (Kolukisaoglu et al., 2004; Mohanta et al., 2015).

79 The expression patterns of many *CBL* genes were also investigated in different tissues and in

80 response to various abiotic stresses in plants (Mohanta et al., 2015; Zhang et al., 2016). These

81 findings lay the foundation for people to further explore the functional mechanisms of CBLs in

82 plants. However, to date, knowledge about genomics and evolutionary information of CBLs in

83 *Gossypium* is limited.

84 Cotton is an essential tetraploid fiber crop that supplies lint for the textile industry worldwide.

85 It is considered to descend from an ancestral combination of two diploid most similar to modern

86 A (for example *Gossypium arboretum*) and D genome species (*Gossypium raimondii*) (Wendel

87 et al., 2010).

88 Cotton growth and development are severely threatened by diverse abiotic stresses such as

89 drought, salinity and potassium starvation (Allen, 2000). Therefore, enhancing stress tolerance of

90 cotton cultivars is one of most important strategies for us to improve their productivity and

91 quality. Potassium is a vital macronutrient for plants, especially for cotton. Potassium shortage in

92 soil seriously affects the yield and quality of cotton (Oosterhuis et al., 2013). Moreover, it has

93 been demonstrated that K^+ uptake is controlled by CBLs through interacting with CIPK23 in

94 *Arabidopsis* and rice under potassium deficiency (Li et al., 2014a; Mao et al., 2016). Research is

95 needed to determine which and how CBLs modulate K^+ absorption in cotton. In this report,

96 genome-wide and comprehensive analyses of the CBL family in *G. arboreum*, *G. raimondii* and

97 *G. hirsutum* were conducted. The expression patterns of *GhCBLs* were monitored in tissues and

98 in response to potassium deficiency in cotton. These analyses will provide a basis for further

99 investigation of the functions of CBLs in *Gossypium*.

100 **MATERIALS AND METHODS**

101 **Identification of CBL family in *Gossypium***

102 The genome sequences of *G. arboreum* (BGI-CGB v2.0 assembly genome), *G. raimondii* (JGI

103 assembly v2.0 data.) and *G. hirsutum* (NAU-NBI v1.1 assembly genome) were downloaded from

104 the CottonGen database (www.cottongen.org), respectively. The protein sequences of 10

105 *Arabidopsis* CBLs were applied as queries to search the three genomes using BLAST-2.4.0

106 software (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST) with default parameters

107 (E-value $< e^{-10}$). EF-hand domains, the typical CBL domains, were analyzed within the candidate

108 CBLs one by one using online software SMART (http://smart.embl-heidelberg.de/). The CBL

109 motifs were also queried against the Pfam databases (Finn et al., 2010). The putative CBLs with

110 questionable annotations (i.e. having a typical CBL domain but low E-value or low coverage of a

111 domain) were manually reanalyzed.

112 **Analysis of *Gossypium CBLs* family**

113 The properties of the *Gossypium* CBL proteins were analyzed using online tools ExPaSy

114 (http://web.expasy.org/protparam/). The subcellular localizations of the CBLs were examined in

115 the website http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc/. The locations of the *CBLs* in

116 chromosomes   were   assessed   by   MapInspect   software

117 (http://www.softsea.com/review/MapInspect.html). Structures of the CBLs were determined by

118 GSDS (http://gsds.cbi.pku.edu.cn/). The conserved domains in the CBLs were affirmed by

119 SMART (http://smart.embl-heidelberg.de). The sequence logo of myristoylation motif in the

120 CBLs were generated by MEME program (http://meme-suite.org/tools/meme).

121 **Analyses of synteny and Ka/Ks ratio**

122 The homologous gene pairs among the *Gossypium* CBLs were searched by the MCScanx

123 software (http://chibba.pgml.uga.edu/mcscan2/). The gene collinearity results were obtained by

124 CIRCOS program (http://www.circos.ca/). The ratio of Ka (nonsynonymous substitution rate) to

125 Ks (synonymous substitution rate) of the CBL genes were estimated by PAML program

126 (http://abacus.gene.ucl.ac.uk/software/paml.html).

127 **Phylogenetic analysis of CBLs**

128 The CBL data were downloaded from the websites for various plant species including

129 *Arabidopsis thaliana* (http://www.arabidopsis.org/), *Oryza sativa* (http://rapdb.dna.affrc.go.jp),

130 *Vitis vinifera* (http://www.genoscope.cns.fr/spip/Vitis-vinifera-e.html), *Populus trichocarpa*

131 (http://www.phytozome.net/poplar), *Glycine max* (http://www.phytozome.net/soybean),

132 *Theobroma cacao* (http://cocoagendb.cirad.fr), *Carica papaya* (http://asgpb.mhpcc.hawaii.edu)

133 and castor bean (http://castorbean.jcvi.org). The full-length amino acid sequences of CBL

134 proteins were aligned using Clustal W software through pairwise and multiple alignment with

135 default parameters (Larkin et al, 2007). Then, phylogenetic trees were generated based on the

136 alignment results using the neighbor joining method (Neighbor-Joining, NJ) and 1,000 bootstrap

137 trials with the MEGA 5.0 software (http://www.megasoftware.net/).

138 **Expression analysis of *GhCBL* genes in tissues and in response to potassium deficiency**

139 For measuring the expression of the *GhCBLs* in tissues, samples of roots, stems and leaves were

140 collected from 20-day-old *G. hirsutum* TM-1 plants normally grown in soil containing 1:1 (v:v)

141 peat:vermiculite in a growth chamber (day/night temperature cycle of 28°C/26°C, 14 h light/10 h

142 dark, and about 50% relative humidity). Flowers were isolated in the morning at the first day of

143 anthesis from cotton grown in the field. The fibers at elongation stage were obtained from the

144 ovules (23 days post anthesis). For monitoring the expression of *GhCBLs* in responding to

145   potassium deprivation, cotton plants grew in clean small pebbles (watered by liquid 1/2 MS

146   medium) (Murashige and Skoog, 1962) in the growth chamber described above for 3 weeks.

147   Then, the plants were watered with K^+-lacking liquid 1/2 MS medium (KNO_3 was replaced by

148   NH_4NO_3 and KH_2PO_4 was replaced by $NH_4H_2PO_4$) for 0 h, 6 h, 2 d and 5 d, respectively.

149   Meanwhile, some K^+-starved seedlings for 5 d were resupplied with K^+ (watered with K^+-

150   contained 1/2 MS medium) for 3 h. The cotton roots were collected, immediately frozen in liquid

151   nitrogen and stored at -70℃. Total RNA of samples was extracted using RNA Pure Plant Kit's

152   protocol (TIANGEN Company). The purity of RNA was examined using a Nanodrop2000

153   nucleic acid analyzer. The A260/280 ratio for each RNA sample was about 2.0. Then, total

154   cDNA was synthesized using M-MLV reserve transcriptase synthesis system (Promega, USA)

155   following the instructions in the Promega kit

156   (https://tools.thermofisher.com/content/sfs/manuals/superscriptIII_man.pdf.)

157   Quantitative real-time RT-PCR (qRT-PCR) experiments were performed using the cDNA,

158   SYBR Green Master mix, the specific primers of *GhCBL* genes (Table 1), and an ABI 7500 real-

159   time PCR system. *GhUBQ7* was used as the internal control. At least three biological replicates

160   were carried out.

161   Table 1   Gene primers used for quantitative real-time RT-PCR experiments

Genes	AGI number	Forward primers (5′-3′)	Reverse primers (5′-3′)
GhUBQ7	Gh_A11G0969	GAAGGCATTCCACCTGACCAAC	CTTGACCTTCTTCTTCTTGTGCTTG
GhCBL1-1	Gh_A11G0257	GAGCGTAACGAGGTCAAGCAAA	CTTCCCGTCCTGATTAATGTCC
GhCBL1-2	Gh_D11G0276	TTTTGTTCGAGCACTCAATGTTT	TTGCCTCAATCGTTTCATCAG
GhCBL1-3	Gh_A03G0043	GACATTCTTGGAAGCCGATA	CTGAGGTATGGGAGGGTCAT
GhCBL1-4	Gh_D09G1875	AGAGTAATGACCCTCCCATACCTAA	CGAGCGAGTATTCTCCGACAA
GhCBL1-5	Gh_A09G1766	GGATGCCGACACTAACCAGG	TCCAACAACGTAGCGGCC
GhCBL3-1	Gh_A01G0740	AGTTGCTCGTGCTCTCTCTGT	ATCATCTGAAAGGTTCATGCCA
GhCBL3-2	Gh_D01G0760	GCAAGAGAGACCGTTTT	AATCTTATCGTCAATGGG

		TAGTG	CG
GhCBL3-3	Gh_A13G1099	GGGCTGATTAACAAGGA GGAGT	ACAGAAAGAGCACGAGC AAACT
GhCBL3-4	Gh_D13G1364	ATGGGCTGATTAACAAG GAGGAG	GACAGAAAGAGCACGAG CGAAC
GhCBL3-5	Gh_A04G0051	GCGGTGATAGATGACGG ACT	GACAGAGAGAGCACGAG CAA
GhCBL3-6	Gh_D05G3682	TACACGCTTCCGACCCT ATT	ATCAATGAGCCCGTCGTA AC
GhCBL4-1	Gh_A11G0126	ACGGCTAGTGAAGTAGA ATCCC	CGAACAAATCAAAAACCC TGTC
GhCBL4-2	Gh_D11G0140	TTCTTGCTGCTGAAACAC CT	CGAACAAATCAAAAACCC TG
GhCBL4-3	Gh_A12G2144	TAAGCGTCTTTCATCCCA AC	TGATTCACCAAGCAGAGC CA
GhCBL4-4	Gh_A09G1696	AACTTAGACACAAGGCT GGGTATG	GAGGTTCTGCTTATTGCTG TTTTT
GhCBL4-5	Gh_D12G2320	CCTGAGGAGGTCAAGGA GATG	AAATTGGGTTGCGAGCTA CAAA
GhCBL9	Gh_D08G1764	GACATTCTTGGATGCCG ACA	ACGCAGCAACCTCGTCTA CT
GhCBL10-1	Gh_A06G0800	AGTCTCACAGTGGCGGC A	TTCATTGGCAAGACGGGT AA
GhCBL10-2	Gh_D06G0922	GTCGCGAGAAATGCCGT TAT	ATTCTCGCCGTATGGAGT TTG
GhCBL10-3	Gh_A05G0335	CTGAAATGAATTTGTCC GATGAC	ACTGGAAATAGTAGTTCA TCACGGA
GhCBL10-4	Gh_D05G0440	TCTGGAATGAATTTGTC GGATG	CTGGAAATAGGAGTTCTT CACGG

162   **Yeast two-hybrid (Y2H) analysis**

163   The full-length CDS sequences of *GhCBLs* and *GhCIPK23* genes were amplified, sequenced and

164   cloned into pGBKT7 and pGADT7 vectors, respectively, using primers listed in Table 2. The

165   plasmids were then transformed into yeast strain AH109 according to the method described in

166   page 18-21 in Yeast Protocols Handbook (Clontech,

167   http://www.clontech.com/xxclt_searchResults.jsp). The cotransformants were plated on non-

168   selective SD/-Leu/-Trp (synthetic dropout medium without Leu and Trp) solid medium and

169   selective SD/-Leu/-Trp/-His/-Ade solid medium. The medium was prepared by ourselves. The

170   concentrations of each component for SD/-Leu/-Trp medium are as follows: L-isoleucine 300

171   mg/L, L-valine 1.5 g/L, adenine 200 mg/L, L-arginine 200 mg/L, L-lysine 300 mg/L, L-

172   methionine 200 mg/L, L-phenylalanine 500 mg/L, L-threonine 2 g/L, L-tyrosine 300 mg/L, L-

173   histidine 200 mg/L, uracil 200 mg/L, yeast nitrogen base without amino acids 6.7 g/L, glucose 20

174   g/L. Serial 1:10 dilutions of the cotransformants were made in water, and 2 μl of the dilution was

175   dropped to generate one spot. Plates were incubated at 30 °C for 3-4 d. 5-bromo-4-chloro-3-

176   indoxyl-α-D-galactopyranoside (X-α-Gal) staining assay was carried out following the

177   instruction (the Clontech protocol, page 26).

178   Table 2   Gene primers used for yeast two-hybrid experiments

Genes	AGI number	Forward primers (5′-3′)	Reverse primers (5′-3′)
GhCBL1-2(BD)	Gh_D11G0276	CCGGAATTCATGGGCTGCTTTCAATCT	CGCGGATCCTGTGGCAACCTCATCA
GhCBL1-3(BD)	Gh_A03G0043	CCGGAATTCATGGGTTGCTTTCATTCT	CGCGGATCCAGTGGCAACTTCATCTAC
GhCBL1-4(BD)	Gh_D09G1875	CGCGGATCCATGGGCTGCTTGCAATGTAAA	GCACTGCAGTATGCCATTCGCCGAGCGAGT
GhCBL1-5(BD)	Gh_A09G1766	ATAGGATCCATGGGCTGCTTGCAATGTA	GCACTGCAGGTATAACATCGGTATTATGTACCT
GhCBL3-2(BD)	Gh_D01G0760	CGCGGATCCATGTTGCAGTGCATAGAC	GCACTGCAGTGTATCATCAACTTGAGAGTGGAAAA
GhCBL3-4(BD)	Gh_D13G1364	CGCGGATCCATGGGAATTTGTTGTTTT	GCACTGCAGTTTGCCACCCATATTCAACT
GhCBL4-1(BD)	Gh_A11G0126	CGCGGATCCATGAAATGGTGTTTTCAAACT	GCACTGCAGATCTCCATTGACGGAGACGCT
GhCBL4-3(BD)	Gh_A12G2144	CGCGGATCCATGGGTTGTTTTTGCTTG	GCACTGCAGCTTATTCCCAACGATTTCAGCT
GhCBL4-4(BD)	Gh_A09G1696	CGCGGATCCATGGGCTGCTTTTGCTTG	GCACTGCAGGTTTTTTCTCAATTCTTCACTGGT
GhCBL8(BD)	Gh_D09G1801	CGCGGATCCATGGGCTGCTTTTGCTTGAAGAA	GCACTGCAGATTCTTCACTGGTTGCTGCAAATCTGAGAC
GhCBL9(BD)	Gh_D08G1764	CCGGAATTCATGGGCTGCTTTCATTCT	CGCGGATCCCGCAGCAACCTCGTCTA
GhCBL10-3(BD)	Gh_A05G0335	CGCGGATCCATGGATTCAACTAGCAAAACC	GCACTGCAGCCGGAGATAGGAAAGGGCCAA
GhCIPK23(AD)	Gh_A06G1219	CCGGAATTCATGGCGAATCGCACTAGT	CGCGGATCCACCATCCTTTTCTTCCAC

179   **RESULTS**

180 **Genome-wide identification of the CBL family in two progenitor diploid and the tetraploid**

181 **cotton species**

182 The CBL genes in *Gossypium* were identified using the homologous alignment method. A total

183 of 13, 13, and 22 CBL genes were respectively detected in A genome (*G. arboretum*), D genome

184 (*G. raimondii*) and A_tD_t genome (*G. hirsutum*) using 10 *Arabidopsis* CBL protein sequences as

185 queries (Table 3). Further, the CBL candidate genes in *Gossypium* were confirmed by domain

186 analysis programs of Pfam and SMART. The CBL family members were named according to

187 their orthologous similarity to the 10 *Arabidopsis* CBL proteins (Mohanta et al., 2015). In

188 general, the *CBLs* in *G. arboretum*, *G. raimondii* and *G. hirsutum* were named *GaCBLs, GrCBLs*

189 and *GhCBLs*, respectively.

190 Most CBLs had very similar physical properties in the 3 *Gossypium* plants (Table 3). The open

191 reading frame (ORF) lengths of the CBL genes ranged from 570 bp to 882 bp except that of

192 *GhCBL3-6*, whose ORF length was 3981 bp. The GaCBL and GrCBL proteins contained 199-

193 279 and 209-253 amino acids (AA), respectively, while GhCBLs were composed of 189-293 AA

194 except GhCBL3-6, which consisted of 1326 AA. The molecular weights (MWs) of GaCBLs

195 varied from 23.25 kDa (GaCBL10-1) to 32.43 kDa (GaCBL10-2), and of GrCBLs ranged from

196 23.25 kDa (GrCBL3-3) to 29.26 kDa (GrCBL10-1). The sizes of GhCBLs were 21.64 kDa

197 (GhCBL3-4) to 33.56 kDa (GhCBL10-1) with an exception of GhCBL3-6 (150.21 kDa). The

198 theoretical isoelectric point (pI) is small for overwhelming majority of the CBLs, ranging from

199 4.65 (GaCBL9) to 5.64 (GhCBL4-5). By contrast, pI of GhCBL3-6 was 8.05 (Table 3).

200 Putative subcellular localizations of the *Gossypium* CBL proteins were also analyzed. It was

201 predicted that all of CBLs were located in cell membrane except that GhCBL3-6 was in the

202 nucleus (Table 3). The quite different characteristics of GhCBL3-6 from other members suggest

203 that GhCBL3-6 likely play a special role in cotton.

204 Table 3 The CBL family genes in *Gossypium*

Gene name	Gene ID	pI	MW (kDa)	Hydro-philicity	Predicted subcellular localization	amino acid residues	coding sequence

GaCBL1-1	Cotton_A_16036	4.74	24.33	-0.163	Cell membrane	213	642
GaCBL1-2	Cotton_A_16034	4.74	24.33	-0.163	Cell membrane	213	642
GaCBL1-3	Cotton_A_16590	5.06	25.39	-0.216	Cell membrane	221	666
GaCBL1-4	Cotton_A_09151	4.72	24.39	-0.142	Cell membrane	213	642
GaCBL2	Cotton_A_07469	4.78	25.94	-0.2	Cell membrane	226	681
GaCBL3-1	Cotton_A_06492	4.77	25.98	-0.189	Cell membrane	226	681
GaCBL3-2	Cotton_A_02147	5.08	27.68	-0.314	Cell membrane	240	723
GaCBL4-1	Cotton_A_02388	4.81	24.88	-0.13	Cell membrane	220	663
GaCBL4-2	Cotton_A_13237	4.97	24.47	-0.173	Cell membrane	215	648
GaCBL8	Cotton_A_08153	4.89	23.48	-0.134	Cell membrane	205	618
GaCBL9	Cotton_A_13238	4.65	24.22	-0.141	Cell membrane	210	633
GaCBL10-1	Cotton_A_14000	4.55	23.25	-0.175	Cell membrane	199	600
GaCBL10-2	Cotton_A_34841	4.82	32.43	-0.028	Cell membrane	279	840
GrCBL1-1	Gorai.007G030300	4.72	24.38	-0.143	Cell membrane	213	642
GrCBL1-2	Gorai.003G178700	4.71	24.45	0.075	Cell membrane	214	645
GrCBL1-3	Gorai.004G191400	4.67	23.86	0.016	Cell membrane	209	630
GrCBL1-4	Gorai.006G214700	4.99	25.39	-0.226	Cell membrane	221	666
GrCBL3-1	Gorai.013G150400	4.79	25.96	-0.208	Cell membrane	226	681
GrCBL3-2	Gorai.002G102900	4.77	25.98	-0.189	Cell membrane	226	681
GrCBL3-3	Gorai.009G450400	4.84	23.25	-0.21	Cell membrane	226	681
GrCBL4-1	Gorai.007G015400	4.78	24.91	-0.193	Cell membrane	233	702
GrCBL4-2	Gorai.006G207100	4.98	25.26	-0.161	Cell membrane	221	666
GrCBL4-3	Gorai.008G255900	5.11	24.02	-0.161	Cell membrane	211	636
GrCBL9	Gorai.008G255800	4.66	24.58	-0.139	Cell membrane	213	642
GrCBL10-1	Gorai.010G101400	4.74	29.26	-0.096	Cell membrane	252	759
GrCBL10-2	Gorai.009G045600	4.83	29.23	-0.095	Cell membrane	253	762
GhCBL1-1	Gh_A11G0257	4.72	24.44	-0.148	Cell membrane	213	642
GhCBL1-2	Gh_D11G0276	4.79	24.38	-0.145	Cell membrane	213	642
GhCBL1-3	Gh_A03G0043	4.98	22.76	-0.163	Cell membrane	199	600
GhCBL1-4	Gh_D09G1875	5.06	25.69	-0.194	Cell membrane	224	675
GhCBL1-5	Gh_A09G1766	5.51	23.23	-0.165	Cell membrane	200	603
GhCBL3-1	Gh_A01G0740	4.77	25.98	-0.189	Cell membrane	226	681
GhCBL3-2	Gh_D01G0760	4.77	25.99	-0.189	Cell membrane	226	681
GhCBL3-3	Gh_A13G1099	4.84	23.25	-0.21	Cell membrane	202	609
GhCBL3-4	Gh_D13G1364	4.98	21.64	-0.205	Cell membrane	189	570
GhCBL3-5	Gh_A04G0051	5.14	21.76	-0.274	Cell membrane	189	570
GhCBL3-6	Gh_D05G3682	8.05	150.21	-0.284	Nucleus	1326	3981

GhCBL4-1	Gh_A11G0126	4.77	23.01	-0.059	Cell membrane	201	606
GhCBL4-2	Gh_D11G0140	4.82	24.97	-0.185	Cell membrane	220	663
GhCBL4-3	Gh_A12G2144	4.97	24.5	-0.175	Cell membrane	215	648
GhCBL4-4	Gh_A09G1696	5.27	28.4	-0.184	Cell membrane	248	747
GhCBL4-5	Gh_D12G2320	5.64	25.06	0.023	Cell membrane	218	657
GhCBL8	Gh_D09G1801	4.85	24.74	-0.177	Cell membrane	217	654
GhCBL9	Gh_D08G1764	4.74	23.8	-0.032	Cell membrane	209	630
GhCBL10-1	Gh_A06G0800	5.18	33.56	-0.143	Cell membrane	293	882
GhCBL10-2	Gh_D06G0922	4.95	30.41	-0.159	Cell membrane	265	798
GhCBL10-3	Gh_A05G0335	5.16	30.48	-0.114	Cell membrane	262	789
GhCBL10-4	Gh_D05G0440	5.01	30.25	-0.08	Cell membrane	262	789

205 **Distribution of the *Gossypium* CBL family members in the whole genome**

206 Chromosomal distributions of the *CBL* genes were examined in *Gossypium*. In general, the *CBLs*

207 were unevenly distributed among the *Gossypium* chromosomes. Thirteen *GaCBLs* were

208 distributed on 7 chromosomes. Among them, 3 *GaCBLs* were located on each of Gachr07 and

209 Gachr11 chromosomes. Two *GaCBLs* were situated in each of Gachr06 and Gachr13, and 1

210 *GaCBL* was on Gachr01, Gachr08 and Gachr09, respectively (Fig.1). Thirteen *GrCBL* genes

211 were identified on 9 chromosomes. Each of the 4 chromosomes Grchr06, Grchr07, Grchr08 and

212 Grchr09 owned 2 genes, and other chromosomes (Grchr02, Grchr03, Grchr04, Grchr10, Grchr13)

213 individually contained 1 gene (Fig.1). Likewise, 22 *GhCBL* family members were mapped onto

214 17 chromosomes. Each of the 5 chromosomes Ghchr09, Ghchr11, Ghchr19, Ghchr21 and

215 Ghchr23 had 2 *CBL* members, and other chromosomes individually carried 1 *CBL* member

216 (Fig.1). We observed the phenomena of 2 CBL genes joining together in a chromosome. For

217 instance, GaCBL4-2 and GaCBL9 were mapped within 16.0 Mb in Gachr06, and GrCBL4-3 and

218 GrCBL9 were mapped within 53.8 Mb in Grchr08. These findings suggest that tandem

219 duplication play a role in generating these genes during evolution.
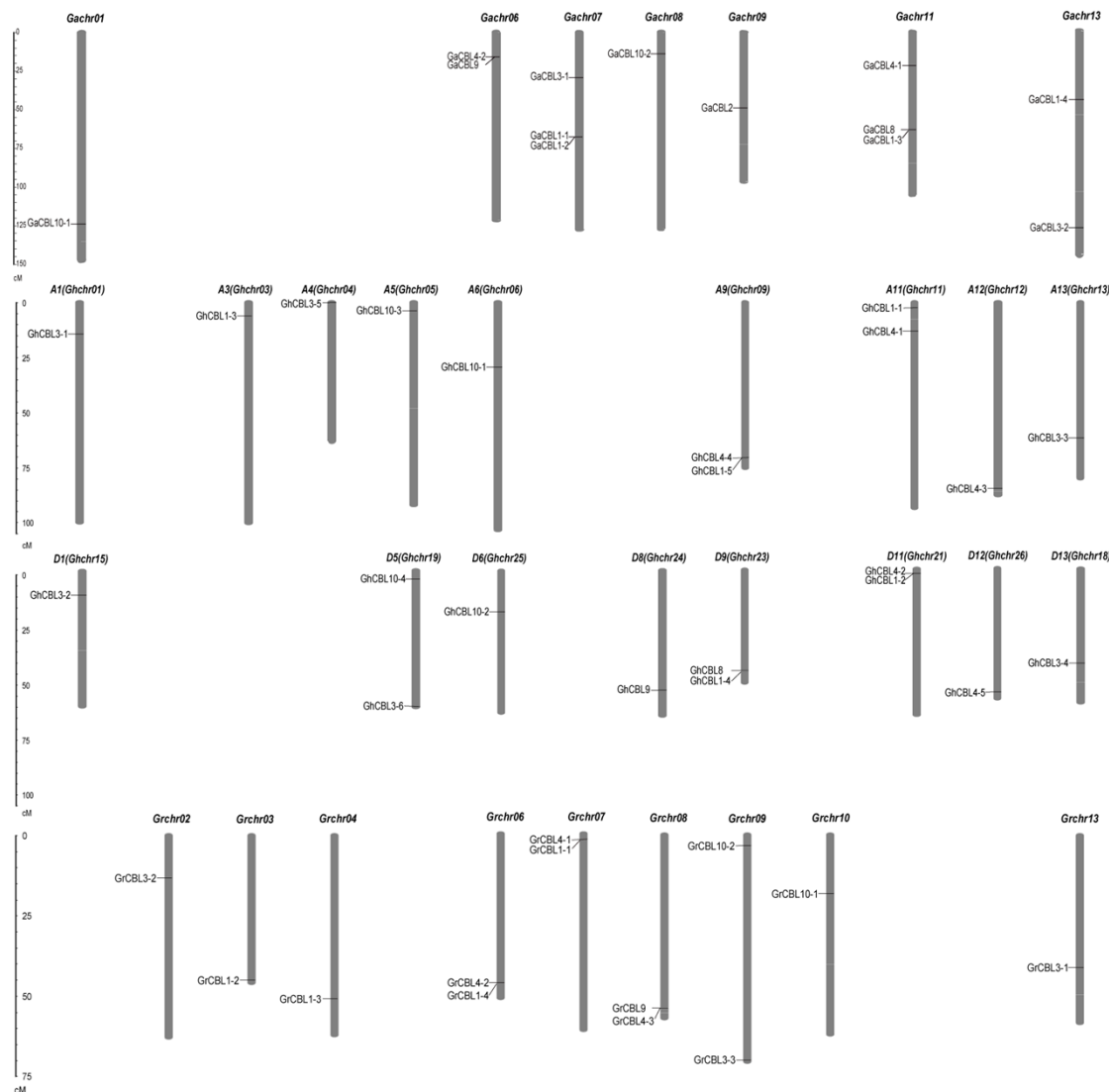
Fig.1 Distributions of the *CBL* family genes on chromosomes in *Gossypium*

The *GaCBLs*, *GrCBLs* and *GhCBLs* are from *G. arboreum*, *G. raimondii* and *G. hirsutum*, respectively.

Phylogenetic analysis and structural properties of *CBL* genes in *Gossypium*

To determine the sequence similarity relationship of the CBLs among *G. arboreum*, *G. raimondii*, and *G. hirsutum*, the phylogenetic tree for the 48 CBLs was constructed. The CBLs can be classified into four families (I to Ⅳ) (Fig.2a). Family I consisted of 12 CBLs (3 GaCBLs, 3 GrCBLs and 6 GhCBLs). The members in family Ⅱ were 8 CBLs (2 GaCBLs, 2 GrCBLs and 4 GhCBLs). Family Ⅲ contained 14 CBLs (4 GaCBLs, 4 GrCBLs and 6 GhCBLs). Family Ⅳ had

226   14 CBLs (4 GaCBLs, 4 GrCBLs and 6 GhCBLs) (Fig.2a).

227   The structure of a protein is closely related to its functions in cells. We therefore identified the

228   intron-exon structures of the *CBL* genes in *Gossypium* by mapping the cDNA sequences onto

229   their genomic sequences. Most of *GaCBLs* and *GrCBLs* owned 8 exons except that *GaCBL3-2*,

230   *GrCBL10-1*, *GrCBL10-2* had 9 and *GaCBL9*, *GrCBL1-2* had 7. The majority of *GhCBLs* carried

231   7-11 exons, but *GhCBL4-4* had 3 exons and *GhCBL3-6* had 22 exons (Fig. 2a).

232   The putative domains in the *Gossypium* CBL proteins were also investigated. EF-hand motifs,

233   which bind to Ca^{2+} ions to transfer calcium signals, were observed in all CBL members. Each

234   CBL proteins had 3 EF-hand motifs except for GaCBL9, which contained 2 such motifs (Fig.

235   2A). Furthermore, a conserved myristoylation motif (MGCXXS/T) was detected in the N-

236   terminal regions of 11 CBL proteins. These proteins included 4 GaCBLs, 2 GrCBLs and 5

237   GhCBLs (Fig. 2B, C). A conserved palmitoylation site with N-terminal Cys residue at third,

238   fourth, fifth or sixth position in amino acid sequence also existed in many cotton CBL members.

239   The two sites are important in the attachment of a protein to membrane (Mohanta et al., 2015).

Fig. 2 Analysis of phylogenetic relationship, gene architecture and conserved domains of CBLs in *Gossypium*

(A) The phylogenetic tree, exon-intron architecture and EF-hand domains of CBLs in *G. arboreum*, *G. raimondii* and *G. hirsutum*. The four major subfamilies are numbered I to IV. The color boxes indicate exons, and the color lines indicate introns; (B) The logo of the myristoylation motif, the capital letters stand for the amino acids, the higher the letter, the higher

the conservation; (C) Multiple sequences containing the myristoylation motif in *Gossypium* CBLs.

Synteny analysis of CBL genes in *Gossypium*

To investigate the genetic origins and evolution of the CBLs in *Gossypium*, the homologous gene pairs among the *CBLs* from *G. arboretum*, *G. raimondii* and *G. hirsutum* were monitored, and the collinear analysis was carried out. The results revealed that 10 homologous gene pairs existed between *G. arboreum* and *G. hirsutum*, and 11 homologous gene pairs were found between *G. raimondii* and *G. Hirsutum* (Fig. 3A). Using the same method, 7 homologous gene pairs were observed between *G. arboreum* and *G. raimondii*. They were distributed on 5 chromosomes in *G. arboreum* and 5 chromosomes in *G. raimondii*, respectively (Fig. 3B). Moreover, 212 homologous gene pairs (both based on orthology and paralogy) were found among the CBLs from the 3 *Gossypium* species (Table S 1). These results imply that many cotton CBL genes may have evolved through segmental duplication.
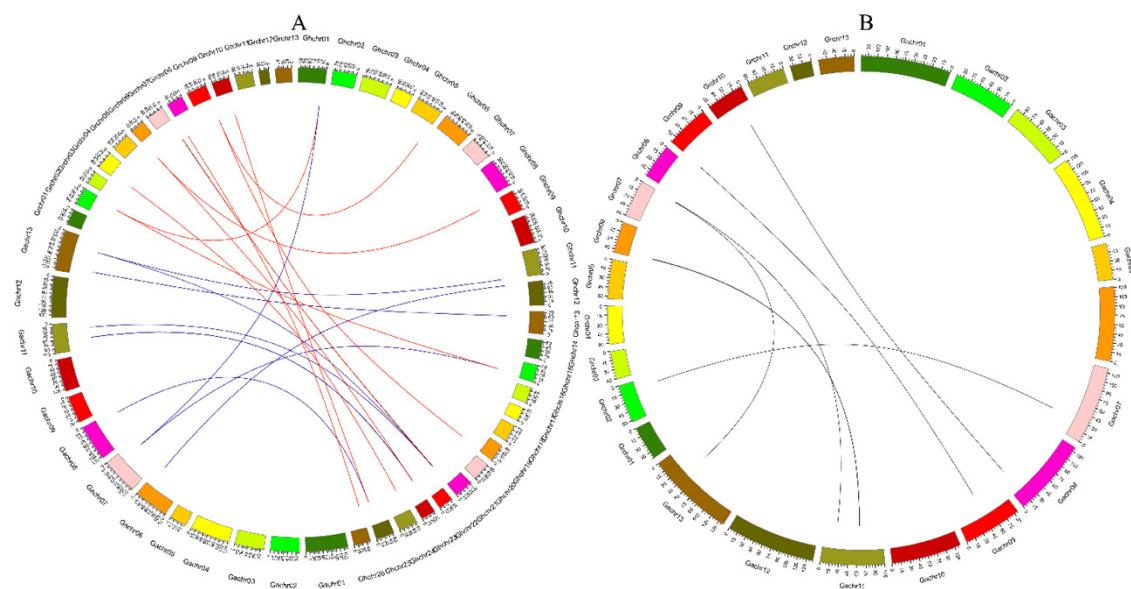
Fig. 3 Genome-wide synteny analysis of *Gossypium* CBL genes

(A) Synteny analysis between *G. hirsutum* and two diploid species *G. arboreum* and *G. raimondii*. Blue lines link gene pairs between *G. arboreum* and *G. hirsutum*, and red lines link gene pairs between *G. raimondii* and *G. hirsutum*; (B) Synteny analysis between *G. arboreum* and *G. raimondi*.

251 **Analysis of Ka/Ks values of the CBLs**

252 To better understand the divergence of the *Gossypium CBL* genes after polyploidization, the

253 value Ka and Ks and their ratio (Ka/Ks) were evaluated for the homologous gene pairs among *G.*

254 *arboreum, G. raimondii* and *G. hirsutum* (Fig. 4,Table S 2). The results showed that the Ka/Ks

255 values among most of the homologous genes were less than 1, indicating they evolved under the

256 purifying selection effect. Only GhCBL10-2/GrCBL10-1 has a Ka/Ks ratio more than 1, hinting

257 that the gene pair may have been generated via the directional selection.
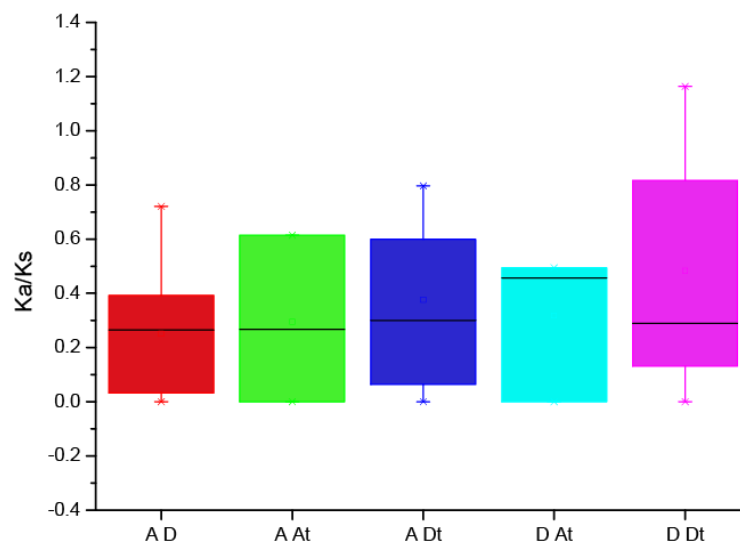
Fig 4. The Ka/Ks values of the CBL homologous genes between the A genome, D genome and

subgenomes of *G. hirsutum* (A_tD_t)

Phylogenetic relationship of CBLs in *Gossypium* and other plant species

258 To gain insight into the evolutionary relationships among GaCBLs, GrCBLs, GhCBLs and CBLs

259 of other plant species, we constructed a phylogenetic tree. Full-length amino acid sequences of

260 126 predicted CBL proteins were obtained from *G. arboretum, G. raimondii, G. hirsutum, A.*

261 *thaliana, C. papaya, G. max, V. vinifera, T. cacao, P. trichocarpa, R. communis* and *O. sativa.*

262 Phylogenetic trees were generated using the neighbor-joining method and MEGA 5.0 software.

263 The CBLs family was divided into thirteen subfamilies according to the topology of the

264 phylogenetic tree (Fig.5). As expected, the three *Gossypium* CBLs commonly clustered closely

265 in a subfamily. Most of them belonged to subfamily two, eight and thirteen. We found that the

266 CBL members from different dicotyledon species and rice always clustered in a subfamily,

267 suggesting that the CBLs shared an ancestral sequence before the divergence of eudicots and

268 monocots or convergent evolution events for these CBLs might have occurred in adaptations to

269 drastic changes in the environment. Moreover, the CBLs from *Gossypium* plants often clustered

270 together with those from *T. cacao* (Fig.5). These results are expected because both *Gossypium*

271 and *T. cacao* are in the *Malvaceae* family.
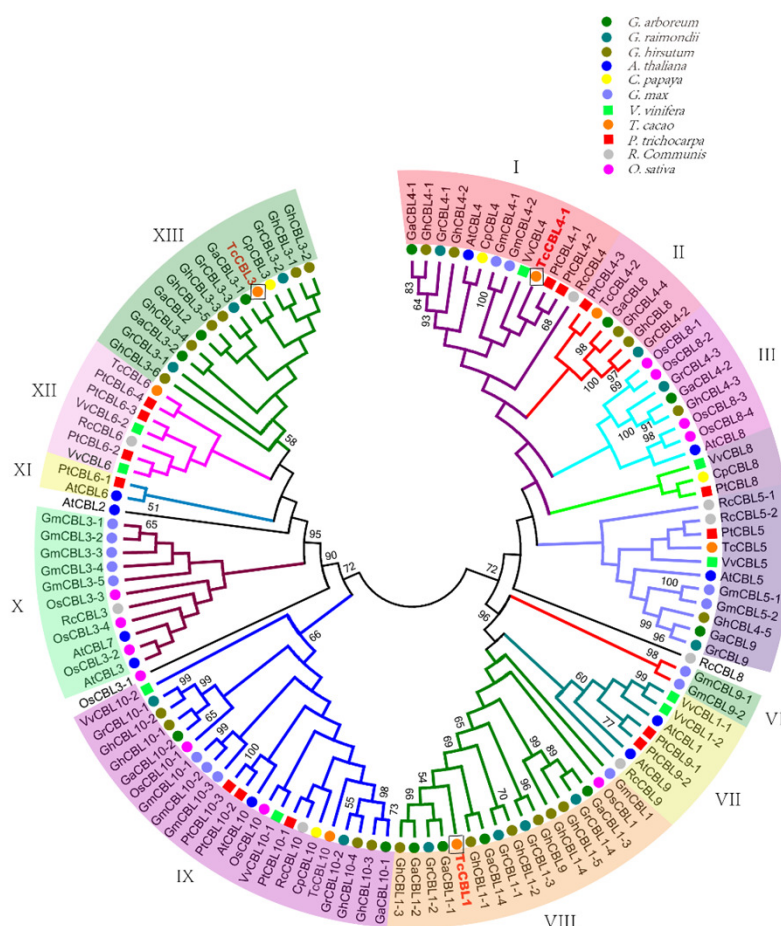
Fig. 5 Phylogenetic tree of CBLs in *Gossypium* and other plant species

The plants in the square frame indicated that the CBL genes outside of *Gossypium* have the

closest evolutionary relationship with *Gossypium CBLs*.

Annotation analysis of GhCBLs

Putative functions of GhCBLs were analyzed using KOG (EuKaryotic orthologous groups (KOG) database (ftp://ftp.ncbi.nih.gov/pub/COG/KOG). Only the information on GhCBL3-6 was obtained. It was predicted that GhCBL3-6 played roles in modulation of RNA processing and modification, signal transduction, and coenzyme transport and metabolism. Gene ontology (GO) database for the 22 GhCBLs was also assessed. The result showed that these GhCBL members were capable of binding calcium ion, like those of other plant species. These analyses indicate that GhCBLs and other CBLs are of great importance in Ca^{2+} signal transduction in plants.

Expression analysis of *GhCBL* genes in tissues

The expression patterns of all the 22 *GhCBL* genes in tissues were monitored by qRT-PCR. We found that most genes were highly expressed in flowers except that *GhCBL4-3*, *GhCBL4-4*, and *GhCBL8* were dominantly expressed in roots and *GhCBL3-6* strongly expressed in leaves. Moreover, the transcripts of *GhCBL1-1*, *GhCBL1-4*, *GhCBL1-5*, *GhCBL3-4*, *GhCBL3-5*, *GhCBL3-6* and *GhCBL9* were relatively abundant in fiber, and those of *GhCBL4-3* were also numerous in flowers (Fig. 6). These results suggest that *GhCBL4-3*, *GhCBL4-4* and *GhCBL8* may mainly function in roots, *GhCBL3-6* mainly functions in leaves and other genes may chiefly act in flowers. *GhCBL1-1*, *GhCBL1-4*, *GhCBL1-5*, *GhCBL3-4*, *GhCBL3-5*, *GhCBL3-6* and *GhCBL9* also probably play a part in fiber development in cotton.
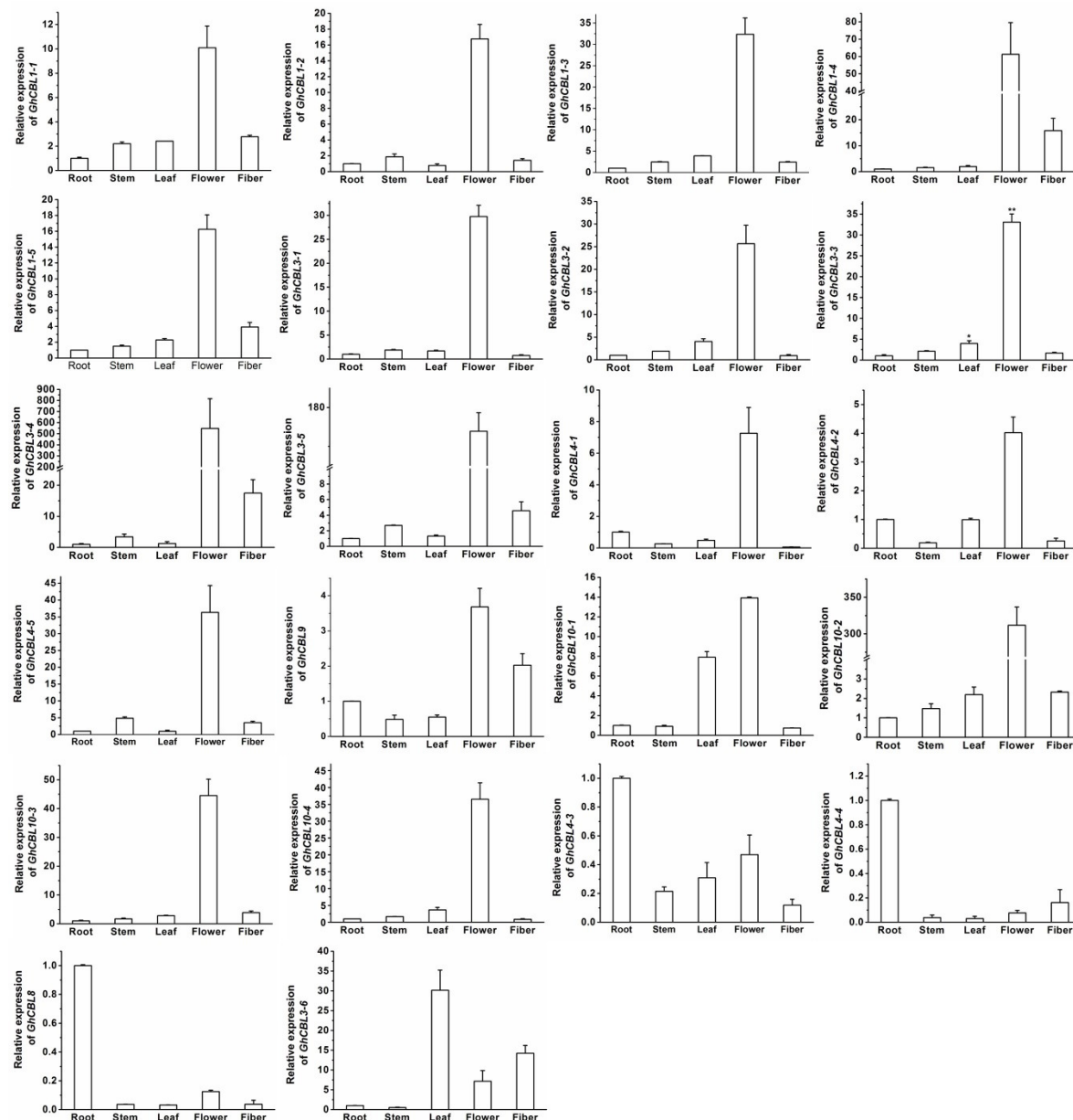
Fig. 6 Expression of 22 *GhCBL* genes in tissues of cotton

The relative expression of genes was calculated from 3 independent replicates. The expression

value of the gene in roots was set as 1. The vertical bars represent the standard error.

290 **Expression patterns of *GhCBLs* in responding to potassium deficiency**

291 CBLs have been addressed to play key roles in response to K⁺ deprivation in *Arabidopsis* and

292 rice (Li et al., 2014a; Mao et al., 2016). Accordingly, we measured the expression patterns of the

293 22 *GhCBL* genes in response to potassium deficiency. As a whole, potassium deficiency

294   moderately altered the expression levels of *GhCBL* genes (Fig. 7). Under potassium deficiency,

295   the transcripts of many genes were reduced at 6 h, but increased at 2 d and/or 5 d. These gene

296   included *GhCBL3-1*, *GhCBL3-2*, *GhCBL3-3*, *GhCBL3-4*, *GhCBL4-4*, and *GhCBL10-3*. The

297   expression levels of *GhCBL3-5*, *GhCBL3-6*, *GhCBL4-3*, *GhCBL4-5*, *GhCBL8* and *GhCBL9* were

298   decreased while those of other genes were unchanged after shortage of potassium (Fig. 7). The

299   effects of K^+ resupply on the abundances of *GhCBL* transcripts were also investigated.

300   Compared with 5 d of low-K^+ treatments, 3 h of K^+ refeeding clearly resulted in decreases in the

301   expression of many genes such as *GhCBL1-3*, *GhCBL1-5*, *GhCBL3-2*, *GhCBL3-3*, *GhCBL3-4*,

302   *GhCBL10-1* and *GhCBL10-3*. However, K^+ resupply increased the expression of *GhCBL4-1*. The

303   transcriptional levels of other genes did not significantly alter upon K^+ resupply (Fig. 7). These

304   results suggest that a number of GhCBLs may play roles in response to potassium starvation in

305   cotton.
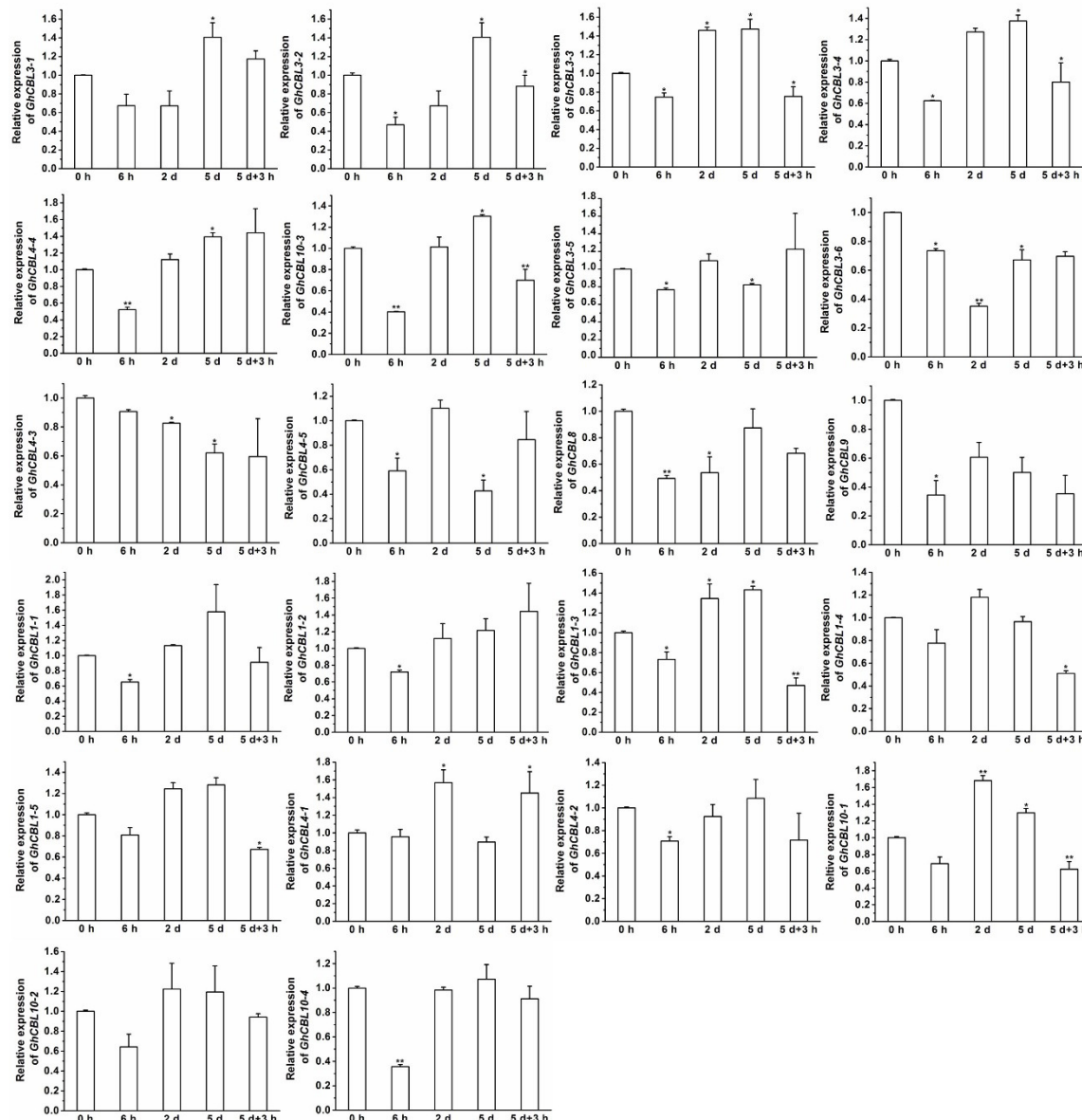
Fig. 7 Expression of 22 *GhCBL* genes under potassium deprivation

306   The relative expression of *GhCBLs* was examined under potassium deficiency or resupply for

307   indicated period of time. The expression value of the gene at 0 h was set as 1. The vertical bars

308   mean the standard error. Statistical analyses were conducted by student's *t* test to assess the

309   differences between the samples at 0 h and those at 6 h, 2 d, or 5 d as well as between the

310   samples at 5 d and those upon resupplying potassium for 3 h (5 d+3 h). The single and double

311   asterisks means that the differences are significant ($P \leqslant 0.05$) and extremely significant

312   ($P \leqslant 0.01$), respectively.

313   **Several GhCBLs can interact with GhCIPK23 *in vitro***

314   To examine whether GhCBLs interact with GhCIPK23, yeast two-hybrid experiments were

315   performed and total of 12 GhCBLs were measured. Among them, GhCBL1-2, GhCBL1-3,

316   GhCBL4-4, GhCBL8, GhCBL9 and GhCBL10-3 were observed to interact with GhCIPK23.

317   Furthermore, GhCBL1-2 and GhCBL9, the respective homologues of *Arabidopsis* CBL1 and

318   CBL9, displayed more strong interactive signals with GhCIPK23 in yeast, suggesting that

319   GhCBL1-2 and GhCBL9 may directly regulate GhCIPK23 in cotton.
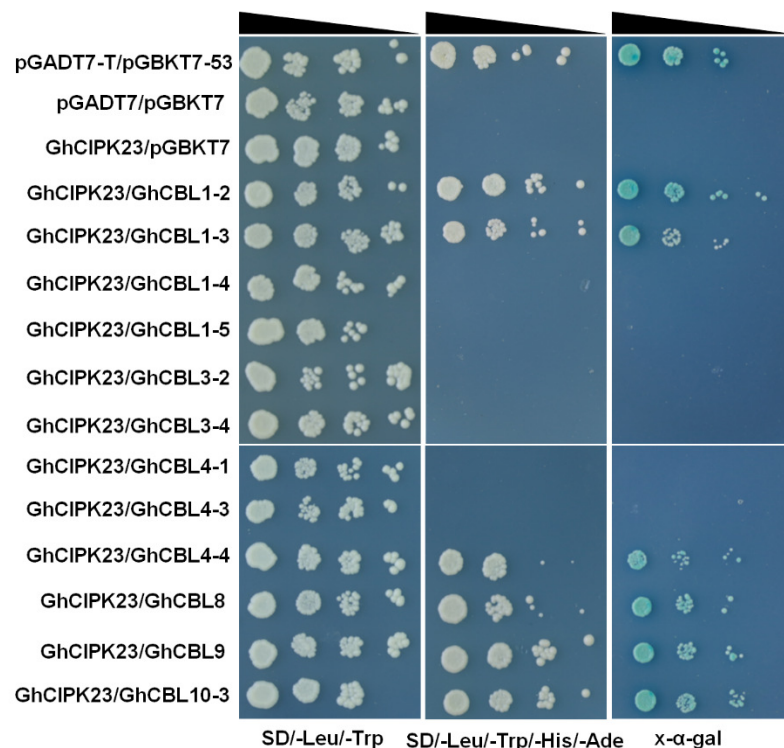
Fig. 8 Yeast two-hybrid analysis of interactions between GhCBLs and GhCIPK23

The yeast cells containing the indicated plasmids were grown on the non-selective SD/-Leu/-Trp

solid medium and selective SD/-Leu/-Trp/-His/-Ade solid medium, followed by X-α-Gal staining.

The reduced cell densities in the dilution series are shown by narrowing triangles when

proceeding from left to right. The first row represents a positive control, the 2th and 3th rows

represent two negative controls.

320   **DISCUSSION**

321   In the present study, we identified 13, 13 and 22 *CBL* genes in *G. arboreum*, *G. raimondii* and *G.*

322   *hirsutum* genomes, respectively (Table 3). Among the 22 *GhCBL* genes, 11 and 11 were

323   assigned to the A_t and D_t subgenome, respectively. They were similar to the number of *CBL*s

324   found in *G. arboreum* and *G. raimondii*, respectively. We detected that 8 *GaCBLs* and 9 *GrCBLs*

325   were homologous genes of *GhCBLs*. However, homologues of 5 GaCBLs and 4 GrCBLs were

326   not discovered in the genome of *G. hirsutum*. These findings indicate that the 8 *GaCBLs* and 9

327   *GrCBLs* have been maintained in *G. hirsutum* after polyploidization event, while the 5 *GaCBLs*

328   and 4 *GrCBLs* diverged from their orthologs in *G. hirsutum* during evolution. Moreover, we

329   observed 5 *GhCBLs* (*GhCBL1-3*, *GhCBL3-5*, *GhCBL4-1*, *GhCBL4-4*, *GhCBL10-1*) in A_t

330   subgenome and 2 *GhCBLs* (*GhCBL3-4*, *GhCBL3-6*) in D_t had no homologues in A genome of *G.*

331   *arboreum* and D genome of *G. raimondii*, respectively. It is conceivable because selection

332   pressures in diploids per loci are different than in the allotetraploid. Relaxed selection allows for

333   development of novel and new functional alleles, but may also accumulate non functional, both

334   at a higher rate possible that within the diploids. *G. arboreum* originates in the Africa/Arabia

335   while *G. raimondii* and *G. hirsutum* originate in the Americas (Wendel et al., 2010). They are

336   distributed in quite different places during evolution. Moreover, *G. arboreum* and *G. hirsutum* are

337   two domasticated species (Wendel et al., 2010). Hence, geographic separation of the three

338   species, and human selection may be essential for the diversity of the CBLs in *Gossypium*.

339   The physical properties of most GaCBLs and GrCBLs were similar to those of GhCBLs

340   (Table 3), suggesting that the functions of the CBLs from the three cotton species remained

341   highly conserved during evolution. The majority of *Gossypium* CBLs was predicted to localize in

342   the membrane, just like many CBLs in *Arabidopsis* and rice. In *Arabidopsis*, CBL1 and CBL9

343   were described to localize in the PM. CBL2, CBL3 and CBL6 localize in tonoplast whereas

344   CBL10 is in both PM and tonoplast (Mao et al., 2016). Rice CBL1 is also present in PM. The

345   localizations of the CBLs should be consistent with their primary roles of sensing and

346   transferring Ca^{2+} signals in *Gossypium*. However, GhCBL3-6 was predicted to be nuclear. Its

347   roles are unknown at present. Experimental characterization of GhCBL3-6 might shed light on

348   some novel functions of it. GhCBL3-6 also gives obvious proof of the evolutionary advantage of

349   being tetraploid. It may be a product of significant human intervention because nothing like it

350   was seen in either diploid.

351   Analysis of gene distributions on chromosomes showed that most homologues of *GaCBLs* and

352   *GrCBLs* in *G. hirsutum* were present in their corresponding A_t and D_t homologous chromosomes,

353   respectively. These findings indicate that *GhCBLs* originate from DNA polyploidization.

354   However, some *GhCBLs* homologues of *GaCBLs* and *GrCBLs* did not appear on their

355   corresponding A_t or D_t chromosomes, suggesting that complex exchange events of chromosome

356   segments occurred in *G. hirsutum* during evolution. Additionally, separated (e.g. *GaCBL4-1* and

357   *GaCBL4-2*; *GrCBL1-1* and *GrCBL1-2*) and jointed (*GaCBL4-2* and *GaCBL9*) distributions of

358   the *Gossypium* CBL homologous genes in chromosomes in combination with the colinearity

359   results of these genes (Fig. 1; Fig. 3) imply that both segmental duplication and tandem

360   duplication are essential for the generation of cotton CBLs during genetic evolution.The number

361   of introns in coding region of most CBL genes in *Gossypium* was six or seven, very similar to

362   that in CBLs genes in *Arabidopsis*, rice, maize, wheat, canola and eggplant (Kolukisaoglu et al.,

363   2004; zhang et al., 2014; Sun et al., 2015; Li et al., 2016; Zhang et al., 2016), reflecting the rather

364   conserved structure of CBL genes in different species. Moreover, nearly all of the *Gossypium*

365   CBLs shared three conserved EF hand domains with other higher plants (Fig. 2). In addition,

366   many *CBLs* from *Gossypium* contained the myristoylated and palmitoylated sites, which may

367   facilitate the targeting of CBL-CIPK complex to membrane. These features are also similar to

368   those in *Arabidopsis*, rice and other plants (Kolukisaoglu et al., 2004; Mohanta et al., 2015). The

369   conserved structure of these CBL family members in different plants might reflect a very similar

370   mode of action and/or conserved interaction with their target protein CIPKs (Mohanta et al.,

371   2015).

372   Measurement of the ratio of Ka to Ks indicated that majority of *Gossypium* CBL homologous

373   genes have undergone purifying selection whereas *GhCBL10-2*/*GrCBL10-1* has experienced

374   directional selection after polyploidization (Fig. 4). These results suggest that most *GhCBLs* have

375 very high similarity in gene sequences and highly conserved functions to their orthologs from *G.*

376 *arboretum* and *G. raimondii* during evolution. By contrast, a large divergence between

377 *GhCBL10-2* and *GrCBL10-1G* has happened. *GhCBL10-2* may have evolved some novel

378 functions through natural selection and human selection.

379 Phylogenetic analysis results revealed that the CBLs in *Gossypium* have closer relationship

380 with those in cocoa than in other plants tested (Fig. 5). These findings strongly suggest that the

381 cotton species may have a more recent common ancestor with cacao relative to other plant

382 species, in line with the results of other gene families in *Gossypium* (Li et al., 2014b; Li et al.,

383 2016). It may justify using CBL as another evolutionary model in plants because it showed

384 highest similarity with another taxon from the same family and may help to narrow down the

385 most vital or evolutionarily conserved or ancient sequences in *Gossypium*.

386 Expression analysis results showed that almost all of the *GhCBL* genes were expressed in

387 various tissues including the root, stem, leaf, flower and fiber. Of note, most genes were

388 dominantly expressed in the flower and fiber (Fig. 6), hinting that these genes may play

389 important roles in the reproductive development in cotton. *G. hirsutum* is a highly domesticated

390 plant for its seed fiber, which is developed from the flower. Preferential expression of many

391 *GhCBLs* in flowers and fibers suggests that human selection markedly affects the genetic

392 variation and expression profiles of *GhCBLs*. Besides, the expression levels of *GhCBL4-3*

393 *GhCBL4-4* and *GhCBL8* in roots were clearly higher than those of other genes. These data imply

394 that the three genes may function in modulation of ion transport or acclimation to diverse abiotic

395 stresses in roots. Their detailed actions and mechanisms will be examined in the future.

396   The expression of 22 *GhCBLs* in responding to potassium starvation was determined. The

397 transcription of most genes was moderately promoted at 2 d and/or 5 d post low-potassium

398 treatments (Fig. 7), indicating multiple GhCBL genes likely regulate cotton response to

399 potassium deprivation. Strikingly, in *Arabidopsis*, the expression of *CBL1* and *CBL9* was

400 reported to be stable, and the transcripts of *CBL10* in roots were moderately decreased under

401 low-potassium conditions (Cheong et al., 2007; Ren et al., 2013). These results imply that

402 constitutive expression of some *CBL* genes may be enough for transmitting Ca^{2+} signals to

403 downstream targets in response to potassium deficiency in plants. Thus, those *GhCBLs* that were

404 not induced by low-potassium stress also likely play a part in adaptations to potassium

405 deprivation in cotton. However, which sequences and how GhCBLs regulate potassium starved

406 responses remains to be investigated in the future.

407   CIPK23 has been observed to function in diverse cellular processes in *Arabidopsis* (Mao et

408 al., 2016). In this study, 6 out of 12 GhCBLs could interact with GhCIPK23 in yeast (Fig. 8),

409 indicating that different GhCBL members may interact with and modulate GhCIPK23 in various

410 growth and/or stress responses in cotton. The cotton homologues of *Arabidopsis* CBL1 and

411 CBL9 suggest that GhCBL1 and GhCBL9 probably play similar roles to CBL1 and CBL9 in

412 cotton.

413 **ACKNOWLEDGEMENTS**

414 This work was supported by the Science and Technology Development Program of He'nan in

415 China (162102110005) and Foundation of He'nan Educational Committee of China (15A210018,

416 17A180018 and 14B180029).

417 **REFERENCES**

418 **Allen RD. 2010.** Opportunities for engineering abiotic stress tolerance in cotton plants. *Cotton*.

419   Berlin:Springer Verlag, 127–160.

420 **Cheong YH, Pandey GK, Grant JJ, Batistic O, Li L, Kim BG, Lee SC, Kudla J, Luan S.**

421   **2007.** Two calcineurin B-like calcium sensors, interacting with protein kinase CIPK23,

422   regulate leaf transpiration and root potassium uptake in *Arabidopsis*. *The Plant Journal*

423   **52:**223–239 DOI 10.1111/j.1365-313X.2007.03236.x.

424 **Cho JH, Lee JH, Park YK, Choi MN, Kim KN. 2016.** Calcineurin B-like protein CBL10

425   directly interacts with TOC34 (Translocon of the Outer membrane of the Chloroplasts) and

426   decreases its GTPase activity in *Arabidopsis*. *Frontiers in Plant Science* **7:**1911 DOI

427   10.3389/fpls.2016.01911.

428 **Drerup MM, Schlücking K, Hashimoto K, Manishankar P, Steinhorst L, Kuchitsu K,**

429   **Kudla J. 2013.** The calcineurin B-like calcium sensors CBL1 and CBL9 together with their

430   interacting protein kinase CIPK26 regulate the *Arabidopsis* NADPH oxidase RBOHF.

431   *Molecular Plant* **6:**559–569 DOI 10.1093/mp/sst009.

432   **Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P,**

433   **Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A. 2010.** The pfam

434   protein families database. *Nucleic Acids Research* **38:**D211-22 DOI 10.1093/nar/gkp985.

435   **Fuglsang AT, Guo Y, Cuin TA, Qiu Q, Song C, Kristiansen KA, Bych K, Schulz A, Shabala**

436   **S, Schumaker KS, Palmgren MG, Zhu JK. 2007.** Arabidopsis protein kinase PKS5

437   inhibits the plasma membrane H⁺-ATPase by preventing interaction with 14-3-3 protein.

438   *Plant Cell* **19:**1617–1634 DOI 10.1105/tpc.105.035626.

439   **Gao P, Zhao PM, Wang J, Wang HY, Du XM, Wang GL, Xia GX. 2008.** Co-expression and

440   preferential interaction between two calcineurin B-like proteins and a CBL-interacting

441   protein kinase from cotton. *Plant Physiology and Biochemistry* **46:**935–940 DOI

442   10.1016/j.plaphy.2008.05.001.

443   **Held K, Pascaud F, Eckert C, Gajdanowicz P, Hashimoto K, Corratgé-Faillie C, Offenborn**

444   **JN, Lacombe B, Dreyer I, Thibaud JB. 2011.** Calcium-dependent modulation and plasma

445   membrane targeting of the AKT2 potassium channel by the CBL4/CIPK6 calcium

446   sensor/protein kinase complex. *Cell Research* **21:**1116–1130 DOI 10.1038/cr.2011.50.

447   **Ho CH, Lin SH, Hu HC, Tsay YF. 2009.** CHL1 functions as a nitrate sensor in plants. *Cell*

448   **138:**1184–1194 DOI 10.1016/j.cell.2009.07.004.

449   **Kim BG, Waadt R, Cheong YH, Pandey GK, Dominguez-Solis JR, Schültke S, Lee SC,**

450   **Kudla J, Luan S. 2007.** The calcium sensor CBL10 mediates salt tolerance by regulating

451   ion homeostasis in *Arabidopsis*. *Plant Journal* **52:**473–484 DOI 10.1111/j.1365-

452   313X.2007.03249.x.

453   **Kolukisaoglu Ü, Weinl S, Blazevic D, Batistic O, Kudla J. 2004.** Calcium sensors and their

454   interacting protein kinases: Genomics of the *Arabidopsis* and Rice CBL-CIPK signaling

455   networks. *Plant Physiology* **134:**43–58 DOI 10.1104/pp.103.033068.

456   **Kudla J, Batistič O, Hashimoto K. 2010.** Calcium signals: the lead currency of plant

457   information processing. *The Plant Cell* **22:**541–563 DOI 10.1105/tpc.109.072686.

458   **Kudla J, Xu Q, Harter K, Gruissem W, Luan S. 1999.** Genes for calcineurin B-like proteins in

459   *Arabidopsis* are differentially regulated by stress signals. *Proceedings of the National*

460   *Acadamy of Sciences of the United States of America* **96:**4718–4723.

461   **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H,**

462   **Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG.**

463   **2007.** Clustal W and Clustal X version 2.0. *Bioinformatics* **23:**2947–2948 DOI

464   10.1093/bioinformatics/btm404.

465   **Li F, Fan G, Wang K, Sun F, Yuan Y, Song G, Li Q, Ma Z, Lu C, Zou C, Chen W, Liang X,**

466   **Shang H, Liu W, Shi C, Xiao G, Gou C, Ye W, Xu X, Zhang X, Wei H, Li Z, Zhang G,**

467   **Wang J, Liu K, Kohel RJ, Percy RG, Yu JZ, Zhu YX, Wang J, Yu S. 2014b.** Genome

468   sequence of the cultivated cotton *Gossypium arboreum*. *Nature Genetics* **46:**567–572 DOI

469   10.1038/ng.2987.

470   **Li J, Jiang MM, Ren L, Liu Y, Chen HY. 2016.** Identification and characterization of CBL and

471   CIPK gene families in eggplant (*Solanum melongena* L.). *Molecular Genetics and*

472   *Genomics* **291:**1769–1781 DOI 10.1007/s00438-016-1218-8.

473   **Li J, Long Y, Qi GN, Li J, Xu ZJ, Wu WH, Wang Y. 2014a.** The Os-AKT1 channel is critical

474   for K^+ uptake in rice roots and is modulated by the rice CBL1-CIPK23 complex. *The Plant*

475   *Cell* **26:**3387–3402 DOI 10.1105/tpc.114.123455.

476   **Li L, Kim BG, Cheong YH, Pandey GK, Luan S. 2006.** A Ca^{2+} signaling pathway regulates a

477   K^+ channel for low-K response in *Arabidopsis*. *Proceedings of the National Acadamy of*

478   *Sciences of the United States of America* **103:**12625–12630 DOI 10.1073/pnas.0605129103.

479   **Li W, Shang H, Ge Q, Zou C, Cai J, Wang D, Fan S, Zhang Z, Deng X, Tan Y, Song W, Li**

480   **P, Jamshed M, Lu Q, Gong W, Li J, Shi Y, Chen T, Gong J, Liu A, Yuan Y. 2016.**

481   Genome-wide identification, phylogeny, and expression analysis of pectin methylesterases

482   reveal their major role in cotton fiber development. *BMC Genomics* **17:**1000 DOI

483   10.1186/s12864-016-3365-z.

484   **Ligaba-Osena A, Fei Z, Liu J, Xu Y, Shaff J, Lee SC, Luan S, Kudla J, Kochian L, Piñeros**

485   **M. 2017.** Loss-of-function mutation of the calcium sensor CBL1 increases aluminum

486   sensitivity in *Arabidopsis*. *New Phytologist* **214**:830–841 DOI 10.1111/nph.14420.

487   **Liu LL, Ren HM, Chen LQ, Wang Y, Wu WH. 2013.** A protein kinase, calcineurin B-like

488   protein-interacting protein kinase9, interacts with calcium sensor calcineurin B-like protein3

489   and regulates potassium homeostasis under low-potassium stress in *Arabidopsis*. *Plant*

490   *Physiology* **161**:266–277 DOI 10.1104/pp.112.206896.

491   **Luan S. 2009.** The CBL-CIPK network in plant calcium signaling. *Trends in Plant Science*

492   **14**:37–42 DOI 10.1016/j.tplants.2008.10.005.

493   **Mao J, Manik SMN, Shi S, Chao J, Jin Y, Wang Q, Liu H. 2016.** Mechanisms   and

494   physiological roles of the CBL-CIPK networking system in *Arabidopsis thaliana*. *Genes*

495   **7**:62 DOI 10.3390/genes7090062.

496   **Mohanta TK, Mohanta N, Mohanta YK, Parida P, Bae H. 2015.** Genome-wide identification

497   of calcineurin B-Like (CBL) gene family of plants reveals novel conserved motifs and

498   evolutionary aspects in calcium signaling events. *Plant Biology* **15**:189 DOI

499   10.1186/s12870-015-0543-0.

500   **Murashige T, Skoog F. 1962.** A revised medium for rapid growth and bio-assays with tobacco

501   tissue culture. *Physiologia Plantarum* **15**: 473－497

502   **Oosterhuis DM, Loka DA, Raper TB. 2013.** Potassium and stress alleviation: Physiological

503   functions and management of cotton. *Journal of Plant Nutrition and Soil Science* **176**:331–

504   343 DOI 10.1002/jpln.201200414.

505   **Pandey GK, Cheong YH, Kim KN, Grant JJ, Li L, Hung W, D'Angelo C, Weinl S, Kudla J,**

506   **Luan S. 2004.** The calcium sensor calcineurin B-like 9 modulates abscisic acid sensitivity

507   and biosynthesis in *Arabidopsis*. *The Plant Cell* **16**:1912–1924 DOI 10.1105/tpc.021311.

508   **Ren XL, Qi GN, Feng HQ, Zhao S, Zhao SS, Wang Y, Wu WH. 2013.** Calcineurin B-like

509   protein CBL10 directly interacts with AKT1 and modulates K^+ homeostasis in *Arabidopsis*.

510   *The Plant Journal* **74:**258–266 DOI 10.1111/tpj.12123.

511   **Sanyal SK, Pandey A, Pandey GK. 2015.** The CBL–CIPK signaling module in plants: a

512   mechanistic perspective. *Physiologia Plantarum* **155:**89–108 DOI 10.1111/ppl.12344.

513   **Sarwat M, Ahmad P, Nabi G, Hu XY. 2013.** Ca^{2+} signals: the versatile decoders of

514   environmental cues. *Critical Reviews in Biotechnology* **33:**97–109 DOI

515   10.3109/07388551.2012.672398.

516   **Steinhorst L, Mähs A, Ischebeck T, Zhang C, Zhang X, Arendt S, Schültke S, Heilmann I,**

517   **Kudla J. 2015.** Vacuolar CBL-CIPK12 Ca^{2+}-sensor-kinase complexes are required for

518   polarized pollen tube growth. *Current Biology* **25:**1475–1482 DOI

519   10.1016/j.cub.2015.03.053.

520   **Straub T, Ludewig U, Neuhäuser B. 2017.** The kinase CIPK23 inhibits ammonium transport in

521   *Arabidopsis thaliana*. *The Plant Cell* **29:**409–422 DOI 10.1105/tpc.16.00806.

522   **Sun T, Wang Y, Wang M, Li T, Zhou Y, Wang X, Wei S, He G, Yang G. 2015.** Identification

523   and comprehensive analyses of the CBL, and CIPK, gene families in wheat (*Triticum*

524   *aestivum* L). *Plant Biology* **15:**269 DOI 10.1186/s12870-015-0657-4.

525   **Tang RJ, Zhao FG, Garcia VJ, Kleist TJ, Yang L, Zhang HX, Luan S. 2015.** Tonoplast

526   CBL-CIPK calcium signaling network regulates magnesium homeostasis in *Arabidopsis*.

527   *Proceedings of the National Acadamy of Sciences of the United States of America*

528   **112:**3134–3139 DOI 10.1073/pnas.142094412.

529   **Thoday-Kennedy EL, Jacobs AK, Roy SJ. 2015.** The role of the CBL-CIPK calcium signaling

530   network in regulating ion transport in response to abiotic stress. *Plant Growth Regulation*

531   **76:**3–12 DOI 10.1007/s10725-015-0034-1.

532   **Tian QY, Zhang XX, Yang A, Wang TZ, Zhang WH. 2016.** CIPK23 is involved in iron

533   acquisition of *Arabidopsis* by affecting ferric chelate reductase activity. *Plant Science*

534   **246:**70–79 DOI 10.1016/j.plantsci.2016.01.010.

535   **Wendel JF, Brubaker CL, Seelanan T. 2010.** The origin and evolution of *Gossypium*. Pp 1-18

536   in: **Stewart JM, Oosterhuis D, Heitholt JJ, Mauney JR** (Eds.), *Physiology of Cotton*,

537 Berlin: Springer, Netherlands.

538 **Xu J, Li HD, Chen LQ, Wang Y, Liu LL, He L, Wu WH. 2006.** A protein kinase, interacting

539 with two calcineurin B-like proteins, regulates K$^+$ transporter AKT1 in *Arabidopsis*. *Cell*

540 **125:**1347–1360.

541 **Zhang F, Li L, Jiao Z, Chen Y, Liu H, Chen X, Fu J, Wang G, Zheng J. 2016.**

542 Characterization of the calcineurin B-Like (CBL) gene family in maize and functional

543 analysis of ZmCBL9 under abscisic acid and abiotic stress treatments. *Plant Science*

544 **253:**118–129 DOI 10.1016/j.plantsci.2016.09.011.

545 **Zhang H, Yang B, Liu W-Z, Li H, Wang L, Wang B, Deng M, Liang W, Deyholos MK,**

546 **Jiang YQ. 2014.** Identification and characterization of CBL and CIPK gene families in

547 canola (*Brassica napus* L.). *BMC Plant Biology* **14:**1 DOI 10.1186/1471-2229-14-8.