Genome-wide characterization and expression analysis of *PP2CA* family members in response to ABA and osmotic stress in *Gossypium*

Tingting Lu [Equal first author, 1, 2] , Gaofeng Zhang [Equal first author, 1] , **Yibin Wang** [1] , **Shibin He** [1] , **Lirong Sun** [Corresp., 1] , **Fushun Hao** [Corresp. 1]

[1] State Key Laboratory of Cotton Biology, Henan Key Laboratory of Plant Stress Biology, School of Life Sciences, Henan University, Kaifeng, Henan, China

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

Corresponding Authors: Lirong Sun, Fushun Hao
Email address: sunlr9208@henu.edu.cn, haofsh@henu.edu.cn

Clade A type 2C protein phosphatases (PP2CAs), as central regulators of abscisic acid (ABA) signaling, negative control growth, development and responses to multiple stresses in plants. PP2CA gene families have been characterized at genome-wide levels in several diploid plants like *Arabidopsis* and rice. However, the information about genome organization, phylogenesis and putative functions of PP2CAs in *Gossypium* is lacking. Here, PP2CA family members were comprehensively analyzed in four *Gossypium* species including the diploid progenitor *G. arboretum*, *G. raimondii* and the tetraploid *G. hirsutum* and *G. barbadense*, and 14, 13, 27 and 23 PP2CA genes were identified in the genomic sequences of these plants, respectively. Analysis results showed that most *Gossypium* PP2CAs were highly conserved in chromosomal locations, structures and phylogeny among the four cotton species. Segmental duplication might play important roles in the formation of the *PP2CAs*, and most PP2CAs may be under purifying selection in *Gossypium* during evolution. The majority of the *PP2CAs* were expressed specifically in diverse tissues, and highly expressed in flowers in *G. hirsutum*. The *GhPP2CAs* displayed diverse expression patterns in responding to ABA and osmotic stress. Yeast-two hybrid assays revealed that many GhPP2CAs were capable of interaction with the cotton ABA receptor s pyrabactin resistance1/PYR1-like/regulatory components of ABA receptors (PYR1/PYL/RCAR) GhPYL2-2D (Gh_D08G2587), GhPYL6-2A (Gh_A06G1418) and GhPYL9-2A (Gh_A11G0870) in the presence and/or absence of ABA. These results gave a comprehensive view of the *Gossypium* PP2CAs, and are valuable for further studying the functions of PP2CAs in *Gossypium*.

1 **Title page**

2 **Title:**

3 **Genome-wide characterization and expression analysis of PP2CA family members in**

4 **response to ABA and osmotic stress in *Gossypium***

6 **Authors:**

7 **Tingting Lu[1,2][†], Gaofeng Zhang[1][†], Yibin Wang[1], Shibin He[1], Lirong Sun[1][*] and Fushun**

8 **Hao[1][*]**

9 **[1]State Key Laboratory of Cotton Biology, Henan Key Laboratory of Plant Stress Biology,**

10 **School of Life Sciences, Henan University, Kaifeng 475004, China**

11 **[2]Henan University of Animal Husbandry and Economy, Zhengzhou 450046, China**

13 **[†]These authors contributed equally to this work**

14 **Tingting Lu: 4147505@qq.com**

15 **Gaofeng Zhang: 2604035350@qq.com**

16 **Yibin Wang: 1532231014@qq.com**

17 **Shibin He: Hsb108@163.com**

18 **[*]Corresponding author**

19 **Lirong Sun: sunlr9208@henu.edu.cn**

20 **Fushun Hao: haofsh@henu.edu.cn**

21 **Tel. 86-371-23881387; Fax 86-371-23881387**

22 Genome-wide characterization and expression analysis of *PP2CA* family members in response to

23 ABA and osmotic stress in *Gossypium*

25 Tingting Lu[1,2†], Gaofeng Zhang[1†], Yibin Wang[1], Shibin He[1], Lirong Sun[1*] and Fushun Hao[1*]

26 [1]State Key Laboratory of Cotton Biology, Henan Key Laboratory of Plant Stress Biology,

27 College of Life Sciences, Henan University, Kaifeng 475004, China

28 [2]Henan University of Animal Husbandry and Economy, Zhengzhou 450046, China

ABSTRACT

31 Clade A type 2C protein phosphatases (PP2CAs), as central regulators of abscisic acid (ABA)

32 signaling, negatively control growth, development and responses to multiple stresses in plants.

33 PP2CA gene families have been characterized at genome-wide levels in several diploid plants like

34 *Arabidopsis* and rice. However, the information about genome organization, phylogenesis and

35 putative functions of PP2CAs in *Gossypium* is lacking. Here, PP2CA family members were

36 comprehensively analyzed in four *Gossypium* species including the diploid progenitor species *G.

37 arboretum*, *G. raimondii* and the tetraploid species *G. hirsutum* and *G. barbadense*, and 14, 13, 27

38 and 23 PP2CA genes were identified in the genomic sequences of these plants, respectively.

39 Analysis results showed that most *Gossypium* PP2CAs were highly conserved in chromosomal

40 locations, structures and phylogeny among the four cotton species. Segmental duplication might

41 play important roles in the formation of the *PP2CAs*, and most *PP2CAs* may be under purifying

42 selection in *Gossypium* during evolution. The majority of the *PP2CAs* were expressed specifically

43 in diverse tissues, and highly expressed in flowers in *G. hirsutum*. The *GhPP2CAs* displayed

44 diverse expression patterns in responding to ABA and osmotic stress. Yeast-two hybrid assays

45 revealed that many GhPP2CAs were capable of interaction with the cotton suggested ABA

46 receptors pyrabactin resistance1/PYR1-like/regulatory components of ABA receptors

47 (PYR1/PYL/RCAR) GhPYL2-2D (Gh_D08G2587), GhPYL6-2A (Gh_A06G1418) and

48 GhPYL9-2A (Gh_A11G0870) in the presence and/or absence of ABA. These results gave a

49 comprehensive view of the *Gossypium* PP2CAs, and are valuable for further studying the functions

50 of these PP2CAs.

51 **Keywords** *Gossypium*; Clade A type 2C protein phosphatases (PP2CAs); gene family;
52 phylogeny; protein interaction

53

54 **INTRODUCTION**

55 Protein phosphorylation and dephosphorylation, as two central mechanisms of cellular signal

56 transduction, play pivotal roles in many biological processes including growth, development and

57 adaptations to various environmental stimuli in plants (Schweighofer et al., 2004). They are

58 catalyzed by protein kinases and phosphatases, respectively. Phosphatases are generally

59 categorized into serine/threonine (Ser/Thr) phosphatases and tyrosine (Tyr) phosphatases

60 according to the different amino acid residues they dephosphorylate. Based on biochemical and

61 pharmacological properties, Ser/Thr phosphatases can be further classified into three large

62 families: phosphoprotein phosphatases (PPs), phosphoprotein metallophosphatases and aspartate-

63 based protein phosphatases (Schweighofer et al., 2004; Kerk et al., 2007; Fuchs et al., 2013;

64   Singh et al., 2015). The PPs includes PP1, PP2A, PP2B, PP4, PP5, PP6 and PP7, the

65   phosphoprotein metallophosphatases consist of Mg^{2+}/Mn^{2+}-dependent type 2C protein

66   phosphatases (PP2Cs) and other Mg^{2+}-dependent phosphatases (Schweighofer et al., 2004; Singh

67   et al., 2010; Fuchs et al., 2013; Singh et al., 2015). PP2Cs, which play key roles in

68   dephosphorylation events in plants, belong to a large subfamily, and can be further divided into

69   11 clades (A–K) in *Arabidopsis* and rice (Singh et al., 2010) and 12 clades (A-L) in

70   *Brachypodium distachyon* (Cao et al., 2016). Among these, clade A proteins of PP2Cs (PP2CAs)

71   are the ones of well-studied PP2Cs in *Arabidopsis*, and they have been shown to have important

72   roles in controlling abscisic acid (ABA) signaling, and negatively regulate plant growth,

73   development and response to various biotic and abiotic stresses in plants (Tähtiharju et al., 2001;

74   Fuchs et al., 2013; Singh et al., 2015). In *Arabidopsis* genome, 9 *PP2CA* members have been

75   identified. They are *ABI1* (*ABA insensitive 1*), *ABI2*, *HAB1* (*Homology to ABI1*), *HAB2*, *AHG1*

76   (*ABA hypersensitive germination 1*), *HAI1* (*Highly ABA-induced PP2C1*), *HAI2*, *HAI3* and

77   *AHG3*/*AtPP2CA* (Fuchs et al., 2013). These genes, particularly *ABI1*, *ABI2* and *AHG3/PP2CA*

78   alone or cooperatively control ABA-mediated transpiration, stomatal closure, seed germination

79   and root growth, and are involved in the regulation of many abiotic stress responses like drought,

80   high salinity, cold, heat and potassium deprivation (Schweighofer et al., 2004; Rubio et al., 2009;

81   Singh et al., 2015). Some *PP2CAs* also play important roles in responses to pathogen attack

82   (Schweighofer et al., 2004; Singh et al., 2015). *PP2CAs* are functionally redundant, and their

83   expression is upregulated by high concentrations of ABA (Rubio et al., 2009; Singh et al., 2015).

84   Moreover, PP2CAs physically interact with numerous cytosolic and nuclear localized proteins

85   such as AtHB6 (Homeobox protein 6), CIPK8 (Calcineurin B–like protein-interacting protein

86   kinase 8), CIPK24, and SnRK2s (Sucrose nonfermenting 1-related protein kinases subfamily 2

87   proteins) (Ohta et al., 2003; Fuchs et al., 2013; Singh et al., 2015). SnRK2s exert central and

88   positive roles in ABA signal cascade in plants (Fujii et al., 2009a; Fujita et al., 2009).

89   Recently, ABA receptors pyrabactin resistance1/PYR1-like/regulatory components of ABA

90   receptors (PYR1/PYL/RCAR) (named PYLs for simplicity) have been found (Ma et al., 2009;

91   Park et al., 2009). This is a breathtaking discovery in plants. PP2CAs were identified as co-

92   receptors, specifically interact with PYLs and control ABA signaling. In the presence of ABA,

93   ABA binds to PYLs, further interacts with and inhibits the activities of PP2CAs; thereby

94   releasing and activating SnRK2s. SnRK2s subsequently regulate multiple downstream

95   transcriptional factors and other proteins to trigger ABA responses (Fujii et al., 2009b; Geiger et

96   al., 2009; Lee et al., 2009).

97   Currently, PP2C gene families including *PP2CAs* have been analyzed at genome-wide levels

98   in *Arabidopsis*, rice, maize and *Brachypodium distachyon* (Xue et al., 2008; Wei and Pan, 2014;

99   Cao et al., 2016). The domain structure of PP2CAs was also studied (Schweighofer et al., 2004).

100   Moreover, the expression patterns of *PP2CAs* have been examined in response to ABA and

101   multiple stresses in *Arabidopsis*, rice, maize and *Brachypodium distachyon* (Xue et al., 2008;

102   Wei and Pan, 2014; Zhang et al., 2017a). However, knowledge about the genomic information

103   and expression profiles of PP2CAs in cotton is unknown to date.

104   Cotton is the most important fiber crop, which provides the spinnable lint for the textile

105   industry in the world. The yield and quality of cotton are adversely affected by many abiotic

106   stresses such as drought and high salinity, which are governed by ABA signaling (Hauser et al.,

107   2011; Liang et al., 2017; Ullah et al., 2017). Therefore, it is essential for us to uncover the

108   functional mechanisms of PP2CAs in ABA signal transduction pathway in cotton. Here, we

109   carried out a genome-wide identification of PP2CA gene family in diploid *G. arboretum* (A2)

110   and *G. raimondii* (D5), and their decendant tetraploid species *G. hirsutum* (AD1) and *G.*

111   *barbadense* (AD2). The evolutionary relationships of these PP2CAs were analyzed. Changes in

112   the transcriptional levels of the *PP2CAs* were also investigated in diverse tissues and in response

113   to ABA and osmotic stress in *G. hirsutum*. Furthermore, the interactions between *G. hirsutum*

114   PP2CAs and several GhPYLs were detected by the yeast-two hybrid method. These results may

115   be valuable for further functional characterization of cotton PP2CAs in ABA signaling in the

116   future.

117

118   **MATERIALS AND METHODS**

119   **Analysis of the PP2C family in four *Gossypium* species**

120   To explore all the members of the PP2C family in *Gossypium*, the protein sequences of 80

121   AtPP2Cs were initially applied as queries to search against the databases of *G. arboreum* (BGI-

122   CGB v2.0 assembly genome), *G. raimondii* (JGI assembly v2.0 data.), *G. hirsutum* (NAU-NBI

123   v1.1   assembly   genome)   (www.cottongen.org)   and   *G.   barbadense*

124   (http://database.chgc.sh.cn/cotton/index.html), respectively, using the BLAST program with

125   default setting (E-value$<e^{-10}$) (Camacho et al., 2009). After removing the redundant sequences

126   from the data set, the putative *Gossypium* PP2Cs were then characterized using the PP2C model

127   (PF00481) (http://pfam.xfam.org/) by the Hmmer software (http://hmmer.org/), and the proteins

128   without a PP2C catalytic domain were deleted. The molecular weight (MW) and isoelectric point

129   (pI)   of   the   *Gossypium*   PP2Cs   were   predicted   by   the   online   tool   ExPaSy

130   (http://web.expasy.org/protparam/), which can give various physico-chemical properties of a

131   protein based on its amino acid sequence (the extinction coefficient and the absorbance of a native

132   protein in water at 280 nm were used).   The composition and position of exons and introns of the

133   PP2CAs were obtained from the website (https://www.cottongen.org/) and characterized by the

134   Gene Structure Display Server (GSDS) tools (http://gsds.cbi.pku.edu.cn/) (Hu et al., 2015). The

135   conserved   domains   of   PP2CAs   were   validated   in   NCBI

136   (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) using the automatic mode (Marchler-

137   Bauer et al., 2017). The MEME program (meme-suite.org/tools/meme) was applied to determine

138   the motifs of PP2CAs in *Gossypium* ('any number of repetitions' to be distributed in sequences

139   was set). The locations of *Gossypium PP2CAs* in chromosomes were assessed using the

140   MapInspect   software   (http://www.mybiosoftware.com/mapinspect-compare-display-linkage-

141   maps.html).

142   **Analysis of synteny and *Ka*/*Ks* of *PP2CAs***

143   The homologous regions of *PP2CAs* in *Gossypium* were identified by the MCScanx software

144   (http://chibba.pgml.uga.edu/mcscan2/), and syntenic blocks were determined by the CIRCOS

145   program (http://www.circos.ca/). The syntenic maps of the *PP2CAs* were obtained using the

146   circos-0.69±3 software with default parameters (http://www.circos.ca/). Some genes located

147   within the same or adjacent intergenic region were regarded as tandem duplications. The

148   nucleotide substitution parameters Ka (non-synonymous) and Ks (synonymous) were assessed by

149   the PAML program (http://abacus.gene.ucl.ac.uk/software/paml.html). Then, the ratio of Ka/Ks

150   was calculated. $Ka/Ks<1$ means purifying selection; $Ka/Ks=1$ indicates neutral selection, while

151     $Ka/Ks>1$ represents positive selection (Hurst, 2002).

152   **Phylogenetic analysis of *PP2CAs***

153   The PP2CA databases were downloaded for *Arabidopsis thaliana* (http://www.arabidopsis.org/),

154   *Theobroma cacao* (http://cocoagendb.cirad.fr), *Ricinus communis* (http://castorbean.jcvi.org),

155   *Populus trichocarpa* (http://www.phytozome.net/poplar), *Glycine max*

156   (http://www.phytozome.net/soybean), *Brachypodium distachyon*

157   (http://plants.ensembl.org/Brachypodium_distachyon/Info/Index), *Oryza sativa*

158   (http://rapdb.dna.affrc.go.jp), and the four *Gossypium* species mentioned above. The amino acid

159   sequences of PP2CAs were aligned by the MUSCLE software (Edgar et al., 2004), and a

160   phylogenetic tree of the PP2CAs was generated using the IQ-TREE server

161   (http://www.iqtree.org/) following the maximum likelihood (ML) method (Nguyen et al., 2015;

162   Trifinopoulos et al, 2016). The best-fitting model was chosen using the ModelFinder

163   (Kalyaanamoorthy et al. 2017) and the support was assessed using the ultrafast bootstrap (Minh

164   et al., 2013). Evolutionary tree was visualized using the FigTree v1.4.4 software (available from

165   https://github.com/rambaut/figtree/releases).

166 **Measurements of *GhPP2CAs* expression in tissues and in response to ABA or osmotic stress**

167 For measuring the expression of *GhPP2CAs* in tissues in each experiment, seeds of *G. hirsutum*

168 L. acc. Texas Marker-1 (TM-1) were sown in pots containing the mixed nutrient soil (rich

169 soil:vermiculite = 2:1, v/v) in a growth chamber. After 21 d, about 2 g samples of roots, stems or

170 leaves were collected from 10 plants. About 20 flowers were collected 1 d post anthesis, and

171 about 5 g fibers were obtained from ovules 23 d post anthesis of cotton plants grown in the

172 fields. For monitoring the expression of *GhPP2CAs* after ABA treatment or under osmotic stress,

173 TM-1 seeds were germinated and planted in liquid 1/2 MS medium (Murashige and Skoog,

174 1962) in a growth chamber (the medium was aerated. day/night temperature cycle of 28°C/26°C,

175 14 h light/10 h dark, and about 50% relative humidity). Three weeks later, the plants were

176 sprayed with 100 μM ABA or treated with 10% PEG6000 (dissolved in medium) for 0, 3, 6, 12

177 and 24 h, respectively. Then, about 2 g roots were sampled, frozen in liquid nitrogen and stored

178 at -70℃. Total RNA was extracted from some of the samples and cDNA was generated

179 according to the method described previously (Ma et al., 2012; Zhang et al., 2017b).

180 Quantitative real-time RT-PCR (qRT-PCR) experiments were constructed in an ABI 7500

181 real-time PCR amplifier using the cDNA, SYBR Green Master mix, the specific primers of

182 *GhPP2CA* genes (Table S1). *GhUBQ7* was used as an internal control (Lu et al., 2017).

183 Experiments were independently repeated three times. The interval between two repeated

184 experiments was 7-10 d.

185 **Monitoring protein interaction by yeast two-hybrid method**

186 The CDS sequences of *GhPYLs* (*GhPYL2-2D*, *GhPYL6-2A*, and *GhPYL9-2A*) and *GhPP2CAs*

187   were amplified, and cloned into pGADT7 and pGBKT7 vectors respectively using gene specific

188   primers (Table S2). After sequencing, the fused vectors were transformed into AH109. The

189   cotransformants were plated on non-selective SD/-Leu/-Trp solid medium and selective SD/-

190   Leu/-Trp/-His/-Ade solid medium as described previously (Lu et al., 2017; Zhang et al., 2017b).

191

192   **RESULTS**

193   **Genome-wide analysis of *PP2CAs* in four *Gossypium* species**

194   To identify the putative PP2CA family members in *Gossypium*, the amino acid sequences of 80

195   *Arabidopsis* PP2Cs (Xue et al., 2008) were used to survey the *Gossypium* dababases. Putative

196   *PP2Cs* were assigned to a total of 114, 116, 239 and 232 genomic sequences that were retrieved

197   from *G. arboretum*, *G. raimondii*, *G. hirsutum* and *G. barbadense*, respectively. They were

198   individually denominated as GaPP2Cs, GrPP2Cs, GhPP2Cs and GbPP2Cs (Table S3). According

199   to the phylogenetic relationships of PP2Cs between *Gossypium* and *Arabidopsis*, the *Gossypium*

200   PP2Cs could be clustered into 12 clades (A-L) (Fig. S1-S4). The PP2CAs possessed 14

201   GaPP2CAs, 13 GrPP2CAs, 27 GhPP2CAs and 23 GbPP2CAs, respectively. They were named

202   individually according to their gene identifiers (Table 1). In this report, we focused on the PP2CA

203   family members in the four *Gossypium* species.

204   It was found that the predicted coded amino acid lengths of *Gossypium* PP2CAs ranged from

205   118 to 593, with an average of 420. These PP2CAs had molecular weights of 12.8 kDa

206   (GhPP2CA27) to 66 kDa (GaPP2CA6). The mean theoretical pIs of PP2CAs was approximately

207   5.9 with a minimum of 4.65 (GrPP2CA5) and a maximum of 8.74 (GhPP2CA22) (Table 1).

208 **Phylogenetic and structural analysis of PP2CAs in *Gossypium***

209 In order to understand the evolutionary relationship among GaPP2CAs, GrPP2CAs, GhPP2CAs

210 and GbPP2CAs, we conducted a phylogenetic tree using the protein sequences of the *Gossypium*

211 PP2CAs (Fig. 1A). As expected, most of GaPP2CAs were individually clustered closely with their

212 corresponding orthologs of GhPP2CAs (GhPP2CA1-13) and GbPP2CAs (GbPP2CA1-10) in A

213 genomes, and a majority of GrPP2CAs individually clustered closely with their homologs of

214 GhPP2CAs (GhPP2CA14-27) and GbPP2CAs (GbPP2CA11-23) in D genomes. Noteworthily,

215 GaPP2CA10 clustered together with GaPP2CA11, and a similar case occurred between

216 GbPP2CA21/GbPP2CA22 and GbPP2CA16/GbPP2CA23. Moreover, homologues of 14

217 GaPP2CAs and of 13 GrPP2CAs were found in the *G. hirsutum* At and Dt subgenomes,

218 respectively; and homologs of 10 GaPP2CAs (except GaPP2CA1, GaPP2CA3, GaPP2CA4 and

219 GaPP2CA14) and 11 GrPP2CAs (except GrPP2CA6 and GrPP2CA9) were detected in the *G.*

220 *barbadense* At' or Dt' subgenomes, respectively. Additionally, three pairs of paralogues with high

221 sequence similarity in GbPP2CAs including GbPP2CA8/GbPP2CA10,

222 GbPP2CA16/GbPP2CA23, GbPP2CA21/GbPP2CA22 were clustered together. They were

223 seemingly derived from GaPP2CA8, GrPP2CA7 and GrPP2CA1, respectively.

224   Most *PP2CAs* had 3-4 exons except that *GaPP2CA6, GrPP2CA9, GhPP2CA22, GbPP2CA13*

225 possessed 5 exons, and *GhPP2CA27* had 2 exons. Among the PP2CA genes, *GaPP2CA6,*

226 *GrPP2CA9, GhPP2CA9* and *GhPP2CA22* individually had a longer intron sequence than other

227 genes did (Fig. 1B). These results indicate that the exon/intron structures of the *Gossypium* PP2CA

228 genes were highly conserved.

229   The motif compositions of the PP2CA proteins were analyzed in the four *Gossypium* species.

230   Twenty putative motifs named motif 1 to motif 20 were identified. Among those, motif 1, 2, 3, 4,

231   5, 6 and 7 existed in every cluster and the majority of the PP2CA members. Moreover, most

232   orthologous PP2CA proteins in the four *Gossypium* plants had the same or very similar

233   compositions and distributions of motifs, suggesting that the PP2CA members in the same

234   cluster likely share similar functions (Fig. 1C).

235   **Chromosomal distributions of *PP2CAs* in *Gossypium***

236   To determine the putative evolutionary relationships of the *Gossypium* PP2CA genes, the positions

237   of the genes on chromosomes were analyzed. We found that the distributions of these *PP2CAs*

238   were uneven. The 14 *GaPP2CAs*, 13 *GrPP2CAs*, 27 *GhPP2CAs* and 23 *GbPP2CAs* were

239   distributed on 8, 9, 16 and 14 chromosomes, respectively. Most of the chromosomes contained

240   one PP2CA gene. By contrast, some chromosomes individually had two PP2CA genes. They were

241   D11 and D13 in *G. raimondii*, At10, At13, Dt05, Dt10 and Dt13 in *G. hirsutum*, and Dt'07 and

242   Dt'13 in *G. barbadense*. Besides, each of D09 and Dt'05 owned 3 PP2CAs. A13, At05 and At'05

243   separately possessed 4 *PP2CAs*. In contrast, *GaPP2CA5, GaPP2CA8, GaPP2CA14, GhPP2CA18,*

244   *GhPP2CA20, GhPP2CA27, GbPP2CA2* and *GbPP2CA12* were located on scaffolds, in which

245   contigs were not spliced into any chromosome in genomic mapping.

246   We compared the positions of the orthologs of *GaPP2CAs*, *GrPP2CAs* and *GhPP2CAs* or

247   *GbPP2CAs* in chromosomes. As expected, most homologs of *GaPP2CAs* and *GrPP2CAs* in *G.*

248   *hirsutum* were located in their corresponding At subgenomes and Dt subgenomes. The similar

249   situation also occurred in *G. barbadense*. However, homologous genes of *GaPP2CAs* and

250 *GrPP2CAs* were barely located in their corresponding homoeologous chromosomes and collinear

251 loci in *G. hirsutum* and *G. barbadense* (Fig. 2). For example, GaPP2CA7 was found in

252 chromosome A09, however, its ortholog GhPP2CA3 and GbPP2CA9 were present in chromosome

253 At05 and At'05, respectively. These results imply that specific, unique and complex variation

254 events in *PP2CA*-contained homoeologous chromosomes may happen within each of the two

255 diploid and tetraploid species during genetic evolution.

256 **Synteny analysis of *PP2CA* genes**

257 During evolutionary processes, tandem and segmental duplications contribute to expanding gene

258 family in plants (Cannon et al., 2004). We examined the duplication relationship of the *PP2CAs*

259 among *G. arboretum*, *G. raimondii* and *G. hirsutum* (the related database for *G. barbadense* was

260 lacking). It was found that *GaPP2CA10* and *GaPP2CA11* joined together, and *GbPP2CA16* and

261 *GbPP2CA23* clustered together in the chromosome. There are less than 5 genes between each pair

262 of the genes, suggesting that the two pairs of genes are tandemly duplicated.

263 The synteny relationship of gene pairs was also explored among *GaPP2CAs*, *GrPP2CAs* and

264 *GhPP2CAs*. A total of 136 homologous gene pairs were observed in 133 collinearity blocks. Most

265 of the blocks had one gene pair. Some blocks owned two gene pairs (*GrPP2CA9/GhPP2CA22,*

266 *GrPP2CA10/GhPP2CA23*) between chromosomal D11 and Dt10. Another block harbored three

267 gene pairs (*GrPP2CA5/GhPP2CA2, GrPP2CA6/GhPP2CA3, GrPP2CA7/GhPP2CA4*) between

268 chromosomal D09 and At05 (Fig. 3). These findings imply that segmental duplication plays major

269 roles in generating *PP2CAs* during evolution in *Gossypium*.

270 **Analysis of *Ka/Ks* values of *PP2CAs***

271   To further understand the evolution processes among *Gossypium* PP2CAs, the effects of selection

272   on duplication of *PP2CA* genes were determined. The non-synonymous (*Ka*) and synonymous

273   (*Ks*) substitutions, and *Ka*/*Ks* values were calculated for the homologous gene pairs among

274   *GaPP2CAs*, *GrPP2CAs* and *GhPP2CAs*. The mean values of *Ka*/*Ks* for these gene pairs between

275   species Ga/Gh, Gr/Ga, Gr/Gh, Ga/Ga, Gr/Gr, Gh/Gh were 0.22, 0.21, 0.22, 0.21, 0.19 and 0.21,

276   respectively. All of them were less than 1, indicating that the formation of these genes were mainly

277   under purifying selection during evolution. The *Ka*/*Ks* ratios for the two gene pairs

278   *GrPP2CA11*/*GhPP2CA15* and *GrPP2CA3*/*GhPP2CA21* were higher than 1, suggesting that the

279   two gene pairs were generated under positive selection and the selection likely has effects to

280   change these genes during evolution (Fig. 4).

281   **Phylogenetic analysis of PP2CAs in *Gossypium* and other plants**

282   We constructed a phylogenetic tree of PP2CA proteins in *G. arboretum*, *G. raimondii*, *G. hirsutum*,

283   *G. barbadense*, *A. thaliana*, *T. cacao*, *R. communis*, *P. trichocarpa*, *G. max, B. distachyon* and *O.*

284   *sativa* using the maximum likelihood (ML) method, and analyzed the evolutionary relationships

285   of these PP2CAs. It was found that the PP2CAs included both dicotyledonous and

286   monocotyledonous members (Fig. 5). This suggests that these PP2CAs formed before the

287   divergence of eudicots and monocots and are in general highly conserved. Indeed, the PP2CAs

288   from the eudicots *Gossypium*, cacao, poplar, castor, soybean and *Arabidopsis* clustered more

289   closely, and those of the monocots rice and *distachyon* clustered together. Moreover, many

290   PP2CAs from *Gossypium* clustered more closely with those from cacao than from poplar, castor,

291   soybean and *Arabidopsis* (Fig. 5), indicating that PP2CAs of *Gossypium* had closer relationship

292   with those of cacao than those of other plants. As expected, PP2CAs in the four *Gossypium* species

293   always clustered together, in line with their homologous evolutionary relationships (Fig. 5).

294   **Expression patterns of *GhPP2CA* genes in different tissues**

295   The transcript abundances of 27 *GhPP2CA* in various tissues were measured by qRT-PCR to

296   determine the putative functions of the PP2CAs in cotton. The results showed that all of the

297   *GhPP2CAs* were highly expressed in flowers. *GhPP2CA11* and *GhPP2CA27* were also

298   preferentially expressed in roots. Moreover, the transcriptional levels of *GhPP2CA3*, *11*, *13*, *27*

299   were high in fibers. The transcripts of *GhPP2CA4*, *16*, *22* were abundant in stems. These results

300   imply that most cotton PP2CA members may function in reproductive development, and some

301   PP2CAs also play roles in some specific tissues like roots, fibers and stems (Fig. 6).

302   **Transcriptional changes of *GhPP2CAs* in responses to ABA and osmotic stress**

303   To gain insight into the roles of *GhPP2CAs* in ABA signaling, transcriptional abundances of

304   *GhPP2CAs* in roots were detected after treatments with 100 μM ABA or 10% PEG6000 for

305   indicated periods of time. We observed that the transcriptional levels of some *GhPP2CA* genes

306   such as *GhPP2CA5*, *11*, *18*, *20*, *25*, *27* continually increased with the extension of ABA

307   treatment time. In contrast, the expression levels of some members including *GhPP2CA2-4*, *8*,

308   *10*, *24* had decreasing trend. The expression levels of some genes were decreased at 3 h or 6 h

309   but increased at 12 h or 24 h. These genes included *GhPP2CA1*, *16*, *17*, *19*, *26*. The expression

310   levels of several genes increased at 3 h or 6 h but decreased at 12 h or 24 h. These genes were

311   *GhPP2CA6*, *7*, *9*, *12-15*, *21-23* (Fig. 7). Treatment of cotton seedlings with PEG6000 also

312   altered the expression of most GhPP2CA genes (Fig. 8). The majority of *GhPP2CAs* were

313   upregulated after treatments with PEG for a short time period, and downregulated afterwards.

314   For example, the transcriptional levels of *GhPP2CA8* and *GhPP2CA21* were prominently

315   enhanced at 3 h, and then reduced at 6 h, 12 h and 24 h, while those of *GhPP2CA5*, *6*, *17*, *23*, *25*,

316   *26* were pronouncedly increased at 6 h and decreased at 12 h and 24 h. By contrast, the

317   expression of some genes was significantly elevated at 12 or 24 h post PEG treatment. These

318   genes included *GhPP2CA1*, *3*, *4*, *7*, *9*, *12-14*, *18*, *20*, *24*, *27*. The overall trend of expression of

319   *GhPP2CA10* and *GhPP2CA15* was increased while that of *GhPP2CA2*, *11*, *16* was decreased

320   under osmotic stress (Fig. 8). Together, these data suggest that *GhPP2CAs* exhibit diverse

321   expression patterns in responses to ABA and osmotic stress.

322   **Many GhPP2CAs interact with GhPYL2-2D, GhPYL6-2A and GhPYL9-2A**

323   PP2CAs have been documented to interact with ABA receptor PYLs in ABA signal pathway

324   (Ma et al., 2009; Park et al; 2009). Accordingly, we investigated the interactions between

325   GhPP2CAs and GhYPLs in the absence or presence of ABA by yeast-two hybrid method. A total

326   of 11 *GhPP2CAs* were cloned, and three *GhPYLs GhPYL2-2D* (Gh_D08G2587), *GhPYL6-2A*

327   (Gh_A06G1418) and *GhPYL9-2A* (Gh_A11G0870) were randomly selected and cloned. These

328   genes were fused into yeast vectors, and yeast-two hybrid experiments were performed. In the

329   absence of ABA, GhPP2CA2 and GhPP2CA25 respectively interacted with GhPYL2-2D while

330   multiple GhPP2CAs like GhPP2CA2, 3, 6, 10, 13, 15, 18, 19, 25 individually interplayed with

331   GhPYL2-2D in the presence of ABA (Fig. 9). In contrast, several GhPP2CAs could respectively

332   interact with GhPYL6-2A or GhPYL9-2A either with or without ABA. These GhPP2CAs

333   included GhPP2CA2, 6, 10, 13, 15, 18, 19, 24, 25. Besides, GhPP2CA3 interact with GhPYL9-

334   2A but not with GhPYL6-2A either in the presence or absence of ABA (Fig.10, Fig. 11). These

335   results imply that GhPP2CAs differentially interact with GhPYLs in responding to ABA in

336   cotton.

337

338   **DISCUSSION**

339   PP2CAs are central components of ABA signal transduction pathway, and negatively control

340   ABA and stress responses in plants (Fuchs et al., 2013; Singh et al., 2015). They have been

341   identified in several plants including *Arabidopsis*, rice, maize and *Brachypodium distachyon* in

342   recent years (Xue et al., 2008; Wei and Pan, 2014; Cao et al., 2016). However, phylogenesis and

343   putative functions of PP2CAs in *Gossypium* remain elusive. In the present study, 14, 13, 27 and

344   23 PP2CA genes were characterized in genomes of *G. arboretum*, *G. raimondii*, *G. hirsutum* and

345   *G. barbadense*, respectively (Table 1). Compared to the number of PP2CAs in *Arabidopsis* (9),

346   rice (10), maize (16) and *Brachypodium distachyon* (8), that in *G. hirsutum* and *G. barbadense*

347   was great (Xue et al., 2008; Wei and Pan, 2014; Cao et al., 2016). This suggests that more

348   complex and elaborate ABA signaling mechanisms modulated by PP2CAs may exist in the

349   upland and island cotton species. Conceivably, the high number of PP2CAs of the two species is

350   related to their tetrapoid nature. The two plants retain most PP2CA homologs of both *G.*

351   *arboreum* and *G. raimondii* but not a copy of either one progenitor during evolution. This may

352   be due to long-term human selection with these two tetraploid cotton species for higher yields,

353   growth in hotter and drier regions, day neutral flowering, and adaptation to agronomic areas far

354   outside their original habitats. These altered characteristics may be associated with more PP2CA

355   proteins and complex ABA signal mechanisms in the cultivated cotton plants than in wild plants

356   and in the greater opportunities to accumulate sequences, amenable to mutation and selection, in

357   a tetraploid genome than in a diploid genome.

358   We noticed that 27 GhPP2CAs and 23 GbPP2CAs individually had their corresponding

359   orthologs in *G. arboretum* or *G. raimondii* (Fig. 1), indicating that those PP2CAs in *G. hirsutum*

360   and *G. barbadense* are ancestrally related to those PP2CAs in the two diploid species.

361   Additionally, no orthologous genes of *GaPP2CA1, 3, 4, 14, GrPP2CA6* and *GrPP2CA9* were

362   observed in *G. barbadense* (Fig. 1). This hints that these genes are possibly lost, or these genes

363   arose after the tetraploid species appeared and separated from the diploid species during the

364   evolutionary processes.

365   The structures and the numbers of introns and exons in *PP2CAs* were similar among the 4

366   *Gossypium* species as well as *Arabidopsis*, rice, maize and *Brachypodium distachyon* (Xue et al.,

367   2008; Wei and Pan, 2014; Cao et al., 2016), suggesting that the *PP2CAs* undergo conserved

368   evolutionary processes even after the divergence of monocotyledons and dicotyledons.

369   Colinearity results showed that 136 homologous gene pairs existed among *GaPP2CAs*,

370   *GrPP2CAs* and *GhPP2CAs* (Fig. 3), indicating that PP2CA genes expand primarily through

371   segmental duplication of DNA. Segmental duplicates may be more often maintained through

372   subsequent gene subfunctionalization compared to tandem duplicates (Lynch and Conery, 2000).

373   Accordingly, these PP2CAs probably had diverse functions in *Gossypium*. Moreover, in

374   agreement with our results, *Arabidopsis* phosphatase family genes also showed segmental

375   duplication (Cannon et al., 2004), suggesting the evolutionary mechanism of PP2CAs may be

376   conserved in plants. The mean value of Ka/Ks for a majority of *PP2CA* homologous gene pairs

377   was about 0.2, significantly less than 1 (Fig. 4). This hints that most mutations occurred in the

378   genomic sequences of PP2CAs in *G. arboretum*, *G. raimondii* and *G. hirsutum* during evolution

379   were detrimental for plant survival, or the mutations-caused traits were not required for man.

380   Thus these mutated genes were gradually eliminated, and only those we found were kept during

381   the long time selection.

382   Phylogenetic results showed that the PP2CA members from monocotyledonous plants

383   clustered together, and similar results occurred in dicotyledonous PP2CAs (Fig. 5). This suggests

384   that great changes in DNA sequences of the *PP2CAs* have taken place after isolation of

385   monocotyledons and dicotyledons although these genes shared a common ancestor. PP2CAs in

386   *Gossypium* always clustered together with those in *T. cacao* rather than with those in *A. thaliana*,

387   *R. communis*, *P. trichocarpa*, *G. max*, *B. distachyon* and *O. sativa* (Fig. 5), pointing to the closer

388   evolutionary relationship of *Gossypium* with *T. cacao*. That is, most of homologous PP2CA

389   members in *Gossypium* and *cacao* were generated before separation of the two genera from the

390   common ancestor. The common PP2CAs across all plants show that they still have core

391   functions essential to basic plant survival and functions. Differences in PP2CAs that follow

392   differentiation of different genera and even species show that they are still diverse and can

393   accommodate functions specific to the survival of species and even in response to selection by

394   man. 'Housekeeping' PP2CAs could probably be subtracted from the picture to illuminate the

395   more unique ones to better understand functions of individual PP2CAs and their roles in specific

396   species, traits, or even agronomic performance of specific cultivars.

397   Transcript abundance analysis indicated that the majority of the *GhPP2CAs* was

398   predominantly expressed in flowers (Fig. 6), suggesting that GhPP2CA-mediated ABA signaling

399   may be of great importance in flower development of cotton. High expression of *GhPP2CAs* in

400   flowers was likely due to the importance of timing flowering to environmental conditions of

401   native *Gossypium* plants. Because evolution of some species is tied to long-term human selection

402   of cotton plants with high yields of fibers and good adaptations to hot and dry growth conditions.

403   Cotton yields are closely associated with flowering in agronomic conditions created by man that

404   often are much different from the natural habitats of wild or ancestral *Gosssypium* species (e.g.

405   cultivation only in summers of termperate-tropical latitudes instead of perennial growth in

406   tropical latitudes closer to the equator). Drought and hot stresses should limit flower

407   development. PP2CAs are negative regulators of the adverse stresses; and therefore, may

408   facilitate flowering of cotton in these newer enviroments. The expression of most *GhPP2CAs*

409   was upregulated after treatment with ABA or PEG6000 (Fig. 7, Fig. 8), in good agreement with

410   the results from *AtPP2CAs*, *OsPP2CAs* and *BdPP2CAs* (Xue et al. 2008; Cao et al. 2016). These

411   findings imply that PP2CAs are essential for plant response to ABA and osmotic stress.

412   The interactions between 11 GhPP2CAs and 3 GhPYLs were examined. The results revealed

413   that most GhPP2CAs can individually interact with the three GhPYLs in the absence or presence

414   of ABA (Fig. 9-11). GhPYLs are homologs of AtPYLs and some GhPYLs have been suggested

415   to be functional ABA receptors (Liang et al., 2017; Zhang et al., 2017b). These data indicate that

416   a large number of GhPP2CAs may play roles via interactions with GhPYLs in ABA-dependent

417   or ABA-independent manner in cotton. The detailed mechanisms of GhPP2CAs in ABA

418 signaling will be further explored in the future.

419 **CONCLUSIONS**

420 In total, 14, 13, 27 and 23 PP2CA genes were characterized from *G. arboretum*, *G. raimondii*, *G.*

421 *hirsutum* and *G. barbadense*, respectively. These genes shared high similarity in chromosomal

422 locations, structures and phylogeny among the species. Most of them might be under purifying

423 selection during evolution. Moreover, *PP2CAs* displayed specific expression patterns in tissues

424 and diverse expression profiles in response to ABA and osmotic stress in *G. hirsutum*. Yeast-two

425 hybrid experiments indicated that most GhPP2CAs interacted with GhPYL2-2D, GhPYL6-2A

426 and GhPYL9-2A with or without ABA. These findings provide essential information for in-

427 depth investigations of the functions of PP2CAs in *Gossypium* in the future.

428

429 **ADDITIONAL INFORMATION AND DECLARATIONS**

430 **Funding**

431 This work was supported by Foundation of Program for young backbone teachers in universities

432 of Henan Province [2016GGJS-024], the scientific research and innovation team in Henan

433 University of Animal Husbandry and Economy (2018KYTD18), and the "111" Project. The

434 funders had no role in study design, data collection and analysis, decision to publish, or

435 preparation of the manuscript.

436 **Grant Disclosures**

437 The following grant information was disclosed by the authors:

438 Program for young backbone teachers in universities of Henan Province: 2016GGJS-024.

439   Foundation of scientific research and innovation team in Henan University of Animal Husbandry

440   and Economy: (2018KYTD18).

441   "111" Project.

Competing Interests

443   The authors declare there are no competing interests.

Author Contributions

445   • Tingting Lu and Gaofeng Zhang analyzed the data and performed the experiments,

446   • Yibin Wang and Shibin He contributed to prepare various reagents/materials/analysis tools.

447   • Lirong Sun and Fushun Hao conceived and designed the experiments, and wrote the

448   manuscript.

449

450   **REFERENCES**

451   **Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden LT. 2009.**

452   Blast+: architecture and applications. *BMC Bioinformatics* **10:** 421.

453   **Cannon SB, Mitra A, Baumgarten A, Young ND, May G. 2004.** The roles of segmental and

454   tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*.

455   BMC *Plant Biol* **4**:10.

456   **Cao J, Jiang M, Li P, Chu Z. 2016.** Genome-wide identification and evolutionary analyses of

457   the PP2C gene family with their expression profiling in response to multiple stresses in

458   *Brachypodium distachyon*. BMC *Genomics* **17**:175.

459   **Edgar RC. 2004.** Muscle: multiple sequence alignment with high accuracy and high throughput.

460   *Nucleic Acids Res* 32:1792–1797.

461   **Fuchs S, Grill E, Meskiene I, Schweighofer A. 2013.** Type 2C protein phosphatases in plants.

462   *FEBS Journal* 280:681–693.

463   **Fujii H, Zhu JK. 2009a.** Arabidopsis mutant deficient in 3 abscisic acid-activated protein

464   kinases reveals critical roles in growth, reproduction, and stress. *Proceedings of the*

465   *National Academy of Sciences* **106:**8380–8385.

466   **Fujii H, Chinnusamy V, Rodrigues A, Rubio S, Antoni R, Park SY, Cutler SR, Sheen J,**

467   **Rodriguez PL, Zhu JK. 2009b.** In vitro reconstitution of an abscisic acid signalling

468   pathway. *Nature* **462**:660–664.

469   **Fujita Y , Nakashima K , Yoshida T , Katagiri T, Kidokoro S, Kanamori N, Umezawa T,**

470   **Fujita M, Maruyama K, Ishiyama K, Kobayashi M, Nakasone S, Yamada K, Ito T,**

471   **Shinozaki K, Yamaguchi-Shinozake K. 2009.** Three SnRK2 protein kinases are the main

472   positive regulators of abscisic acid signaling in response to water stress in *Arabidopsis*.

473   *Plant and Cell Physiology* **50:**2123–2132.

474   **Geiger D , Scherzer S , Mumm P , Stange A, Marten I, Bauer H, Ache P, Matschi S, Liese**

475   **A, AI-Rasheid KA, Romeis T, Hedrich R. 2009.** Activity of guard cell anion channel

476   SLAC1 is controlled by drought-stress signaling kinase-phosphatase pair. *Proceedings of*

477   *the National Academy of Sciences* **106:**21425–21430.

478   **Hauser F, Waadt R, Schroeder J. 2011.** Evolution of Abscisic Acid Synthesis and Signaling

479   Mechanisms. *Current Biology* **21:**346–355.

480   **Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. 2015.** GSDS 2.0: an upgraded gene feature

481   visualization server. *Bioinformatics* **31:** 1296–1297.

482   **Hurst LD. 2002.** The Ka/Ks ratio: diagnosing the form of sequence evolution. *Cell* 18:486.

483   **Kalyaanamoorthy S, Minh BQ, Wong TKF, vib Haeseker A, Jermiin LS. 2017.**

484   ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*

485   **14**:587-589.

486   **Kerk D, Templeton G, Moorhead GBG. 2007.** Evolutionary radiation pattern of novel protein

487   phosphatases revealed by analysis of protein data from the completely sequenced genomes

488   of humans, green algae, and higher plants. *Plant Physiology* **146**:351–367.

489

490   **Lee SC, Lan W, Buchanan BB, Luan S. 2009.** A protein kinase-phosphatase pair interacts with

491   an ion channel to regulate ABA signaling in plant guard cells. *Proceedings of the National*

492   *Academy of Sciences* **106**:21419–21424.

493   **Liang C, Liu Y, Li Y, Meng Z, Yan R, Zhu T, Wang Y, Kang S, Abid MA, Malik W, Sun**

494   **G, Guo S, Zhang R. 2017.** Activation of ABA receptors gene GhPYL9-11A is positively

495   correlated with cotton drought tolerance in transgenic Arabidopsis. *Frontiers in Plant*

496   *Science* **8**:1453.

497   **Lu TT, Zhang GF, Sun LR, Wang J, Hao FS. 2017.** Genome-wide identification of CBL

498   family and expression analysis of CBLs in response to potassium deficiency in cotton.

499   *PeerJ* **5**:e3653.

500   **Lynch M, Conery JS. 2000.** The evolutionary fate and consequences of duplicate genes.

501   *Science* **290**:1151–1155.

502 **Ma LY, Zhang H, Sun LR, Jiao YH, Zhang GZ, Miao C, Hao FS. 2012.** NADPH oxidase

503 AtrbohD and AtrbohF function in ROS-dependent regulation of Na$^+$/K$^+$ homeostasis in

504 *Arabidopsis* under salt stress. *Journal of Experimental Botany* **63**:305–317.

505 **Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK,**

506 **Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki**

507 **N, Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, Bryant SH. 2017.**

508 CDD/SPARCLE: functional classification of proteins via subfamily domain architectures.

509 *Nucleic Acids Research* **45**: D200-D203.

510 **Ma Y, Szostkiewicz I, Korte A, Moes A, Yang Y, Christmann A, Grill E. 2009.** Regulators of

511 PP2C phosphatase activity function as abscisic acid sensors. *Science* **324**:1064–1068.

512 **Minh BQ, Nguyen MA, von Haeseler A. 2013.** Ultrafast approximation for phylogenetic

513 bootstrap. *Molecular& Biology& Evolution* **30**:1188–1195.

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

515 tissue culture. *Physiologia Plantarum* **15**:473–497.

516 **Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. 2015.** IQ-TREE: a fast and effective

517 stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology*

518 *and Evolution* **32**:268–274.

519 **Ohta M, Guo Y, Halfter U, Zhu J. 2003.** A novel domain in the protein kinase SOS2 mediates

520 interaction with the protein phosphatase 2C ABI2. *Proceedings of the National Academy of*

521 *Science* **100**:11771–11776.**Park SY, Fung P, Nishimura N, Jensen DR, Fujii H, Zhao Y,**

522 **Lumba S, Santiago J, Rodrigues A, Chow TF, Alfred SE, Bonetta D, Finkelstein R,**

523   **Provart NJ,Desveaux D, Rodriguez PL, McCourt P, Zhu JK, Schroeder JI, Volkman**

524   **BF, Cutler SR. 2009.** Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL

525   family of START proteins. *Science* **324**:1068–1071.

526   **Rubio S, Rodrigues A, Saez A, Dizon MB, Galle A, Kim TH, Santiago J, Flexas J,**

527   **Schroeder JI, Rodriguez PL. 2009.** Triple loss of function of protein phosphatases type 2C

528   leads to partial constitutive response to endogenous abscisic acid. *Plant Physiol* **150:**1345–

529   1355.

530   **Schweighofer A, Hirt H, Meskiene I. 2004.** Plant PP2C phosphatases: emerging functions in

531   stress signaling. *Trends in Plant Science* **9**:236–243.

532   **Singh A, Giri J, Kapoor S, Tyagi AK, Pandey GK. 2010.** Protein phosphatase complement in

533   rice: genome-wide identification and transcriptional analysis under abiotic stress conditions

534   and reproductive development. *BMC Genomics* **11**:435.

535   **Singh A, Pandey A, Srivastava AK, Tran LS, Pandey GK. 2015.** Plant protein phosphatases

536   2C: from genomic diversity to functional multiplicity and importance in stress management.

537   *Critical Reviews in Biotechnology* **36**:1023–1035.

538   **Tähtiharju S, Palva T. 2001.** Antisense inhibition of protein phosphatase 2C accelerates cold

539   acclimation in *Arabidopsis thaliana*. *The Plant Journal* **26:** 461–470.

540   **Trifinopoulos J, Nguyen LT, Von Haeseler A, Minh BQ. 2016.** W-IQ-TREE: a fast online

541   phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research* **44**:232–235.

542   **Ullah A, Sun H, Yang X, Zhang X. 2017.** Drought coping strategies in cotton: increased crop

543   per drop. *Plant Biotechnology Journal* 15: 271–284.

544   **Wei K, Pan S. 2014.** Maize protein phosphatase gene family: identification and molecular

545   characterization. *BMC Genomics* **15**:773.

546   **Xue T, Wang D, Zhang S, Ehlting J, Ni F, Jakab S, Zheng C, Zhong Y. 2008.** Genome-wide

547   and expression analysis of protein phosphatase 2C in rice and *Arabidopsis*. *BMC Genomics*

548   **9**:550.

549   **Zhang F, Wei Q, Shi J, Jin X, He Y, Zhang Y, Luo Q, Wang Y, Chang J, Yang G, He G.**

550   **2017a.** *Brachypodium distachyon* BdPP2CA6 interacts with BdPYLs and BdSnRK2 and

551   positively regulates salt tolerance in transgenic *Arabidopsis. Frontiers in Plant Science* **8**:264.

552   **Zhang GF, Lu TT, Miao WW, Sun LR, Tian M, Wang J, Hao FS. 2017b.** Genome-wide

553   identification of ABA receptor PYL family and expression analysis of PYLs in response to

554   ABA and osmotic stress in *Gossypium. PeerJ* **5**:e4126.

Figure 1

Phylogenetic relationships, gene structures and conserved motifs of PP2CA genes in *Gossypium* .

(A) The phylogenetic tree was generated by the maximum likelihood (ML) method, with ultrafast bootstrap. (B) Exon/intron architectures of *Gossypium* PP2C genes. The black boxes indicate exons, and the black lines represent introns. The sizes of exons and introns were determined by the scale at the bottom. (C) Distributions of conserved motifs. The motifs are displayed by 20 different color boxes.

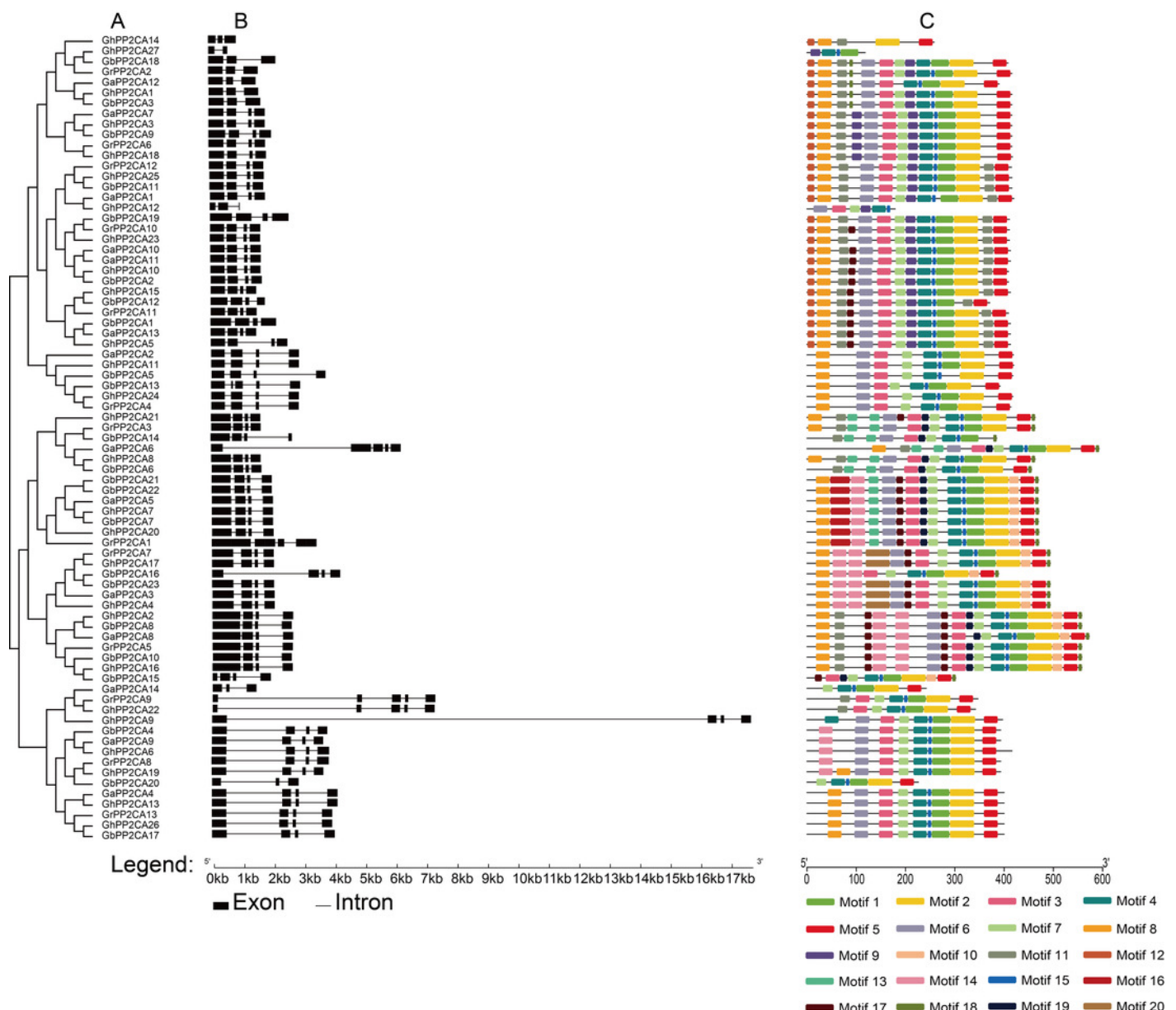



Figure 2

Positions of *Gossypium* PP2CA genes on chromosomes.

GaPP2CAs , *GrPP2CA s GhPP2CA s*, and *GbPP2CAs* were from *G. arboreum*, *G. raimondii* , *G. hirsutum*, and *G. barbadense* , respectively.

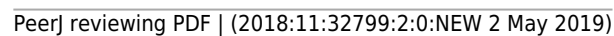



Figure 3

Genome-wide synteny results of PP2CA genes from *G.arboreum*, *G. raimondii* and *G. hirsutum*.

Green lines linked gene pairs between *G. arboreum* and *G. hirsutum*, red lines connected gene pairs between *G. raimondii* and *G. hirsutum*, blue lines bridged gene pairs between *G. arboreum* and *G. raimondii*, yellow lines joined gene pairs within individual species in *G. arboretum*, *G. raimondii* and *G. hirsutum.*

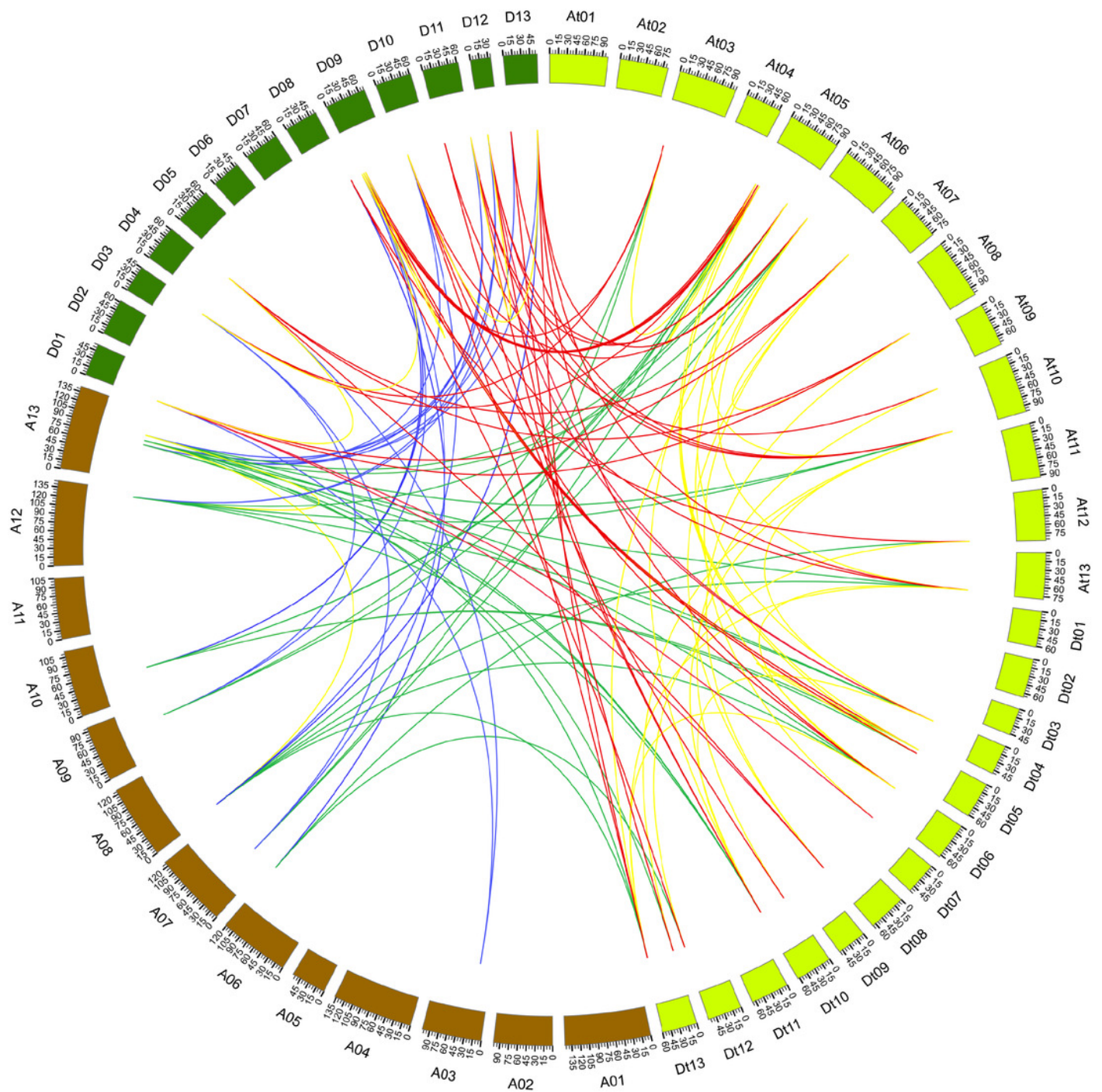



Figure 4

The *Ka/Ks* values of the homologous PP2CA gene pairs in *Gossypium*.

(Ga), (Gr) and (Gh) represented *G. arboretum*, *G. raimondii*, and *G. hirsutum* , respectively.

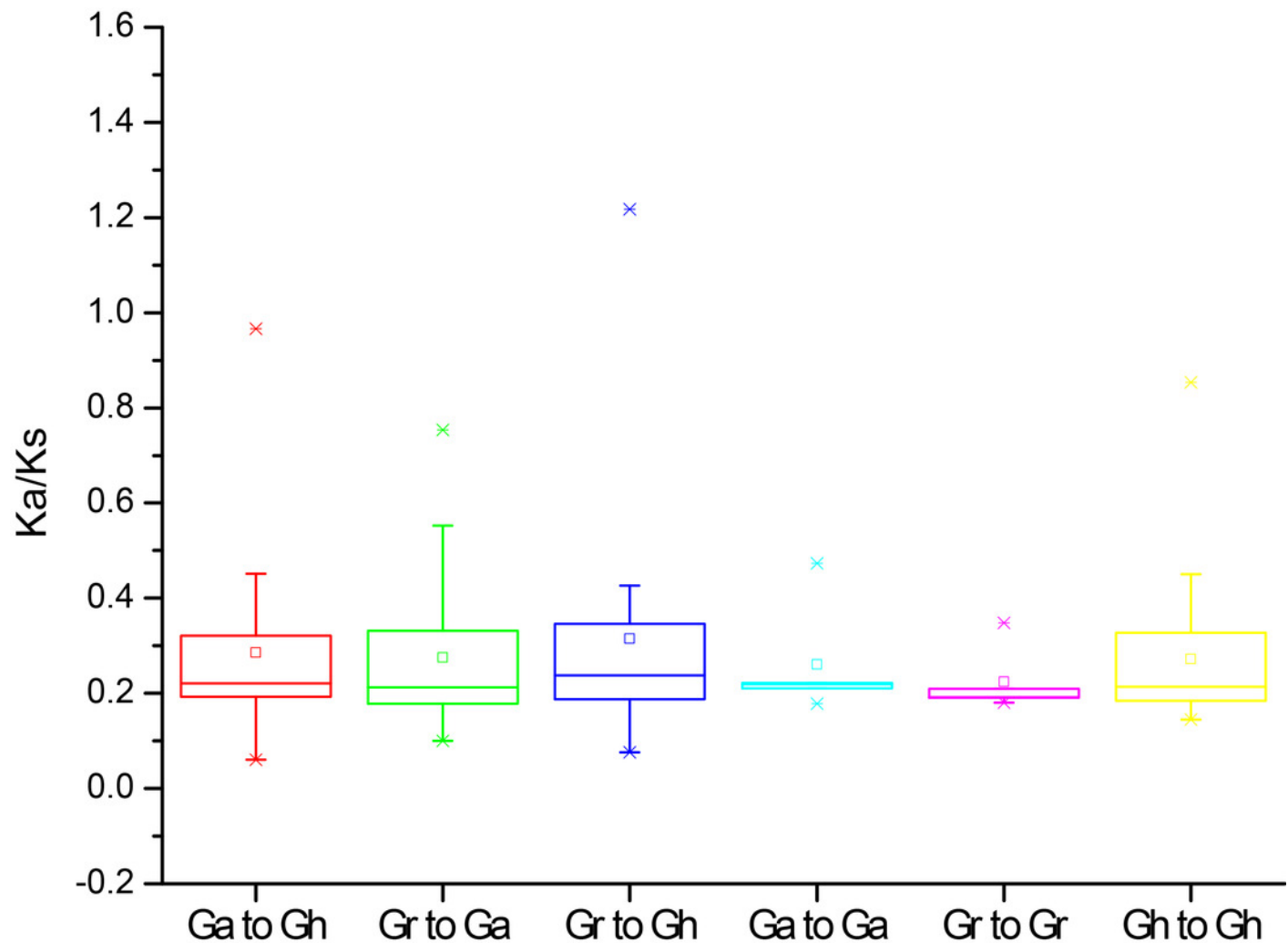



Figure 5

Phylogenetic tree of PP2CAs in *Gossypium* and other plants.

The multiple alignment was performed by MUSCLE program. IQTREE was used to create the maximum likelihood with Dayhoff model. The phylogenetic relationship of PP2CAs in *G. arboretum*, *G. raimondii*, *G. hirsutum*, *G. barbadense*, *A. thaliana*, *T. cacao*, *R. communis*, *P. trichocarpa*, *G. max*, *B. distachyon* and *O. sativa* were analyzed.

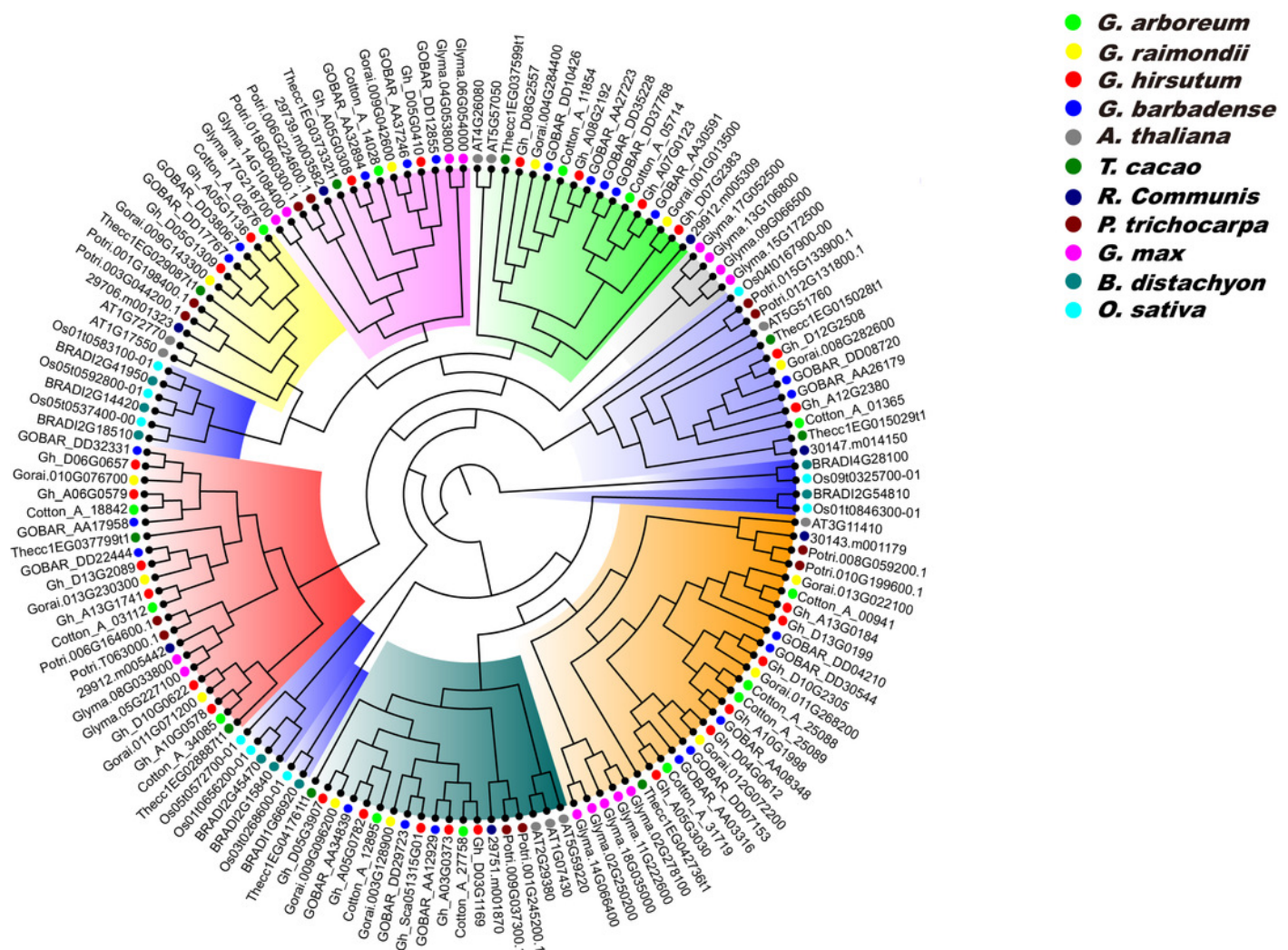



Figure 6

Transcription levels of *GhPP2CAs* in cotton tissues.

The genes preferentially expressed in flowers (A-Z) and fibers (AA). Gene *GhUBQ7* was used as the internal control. The expression level of the gene in roots was set as 1. The data were mean±SE. Statistical analyses were carried out by student's *t*-test to determine the differences in gene abundances between roots and other tissues.

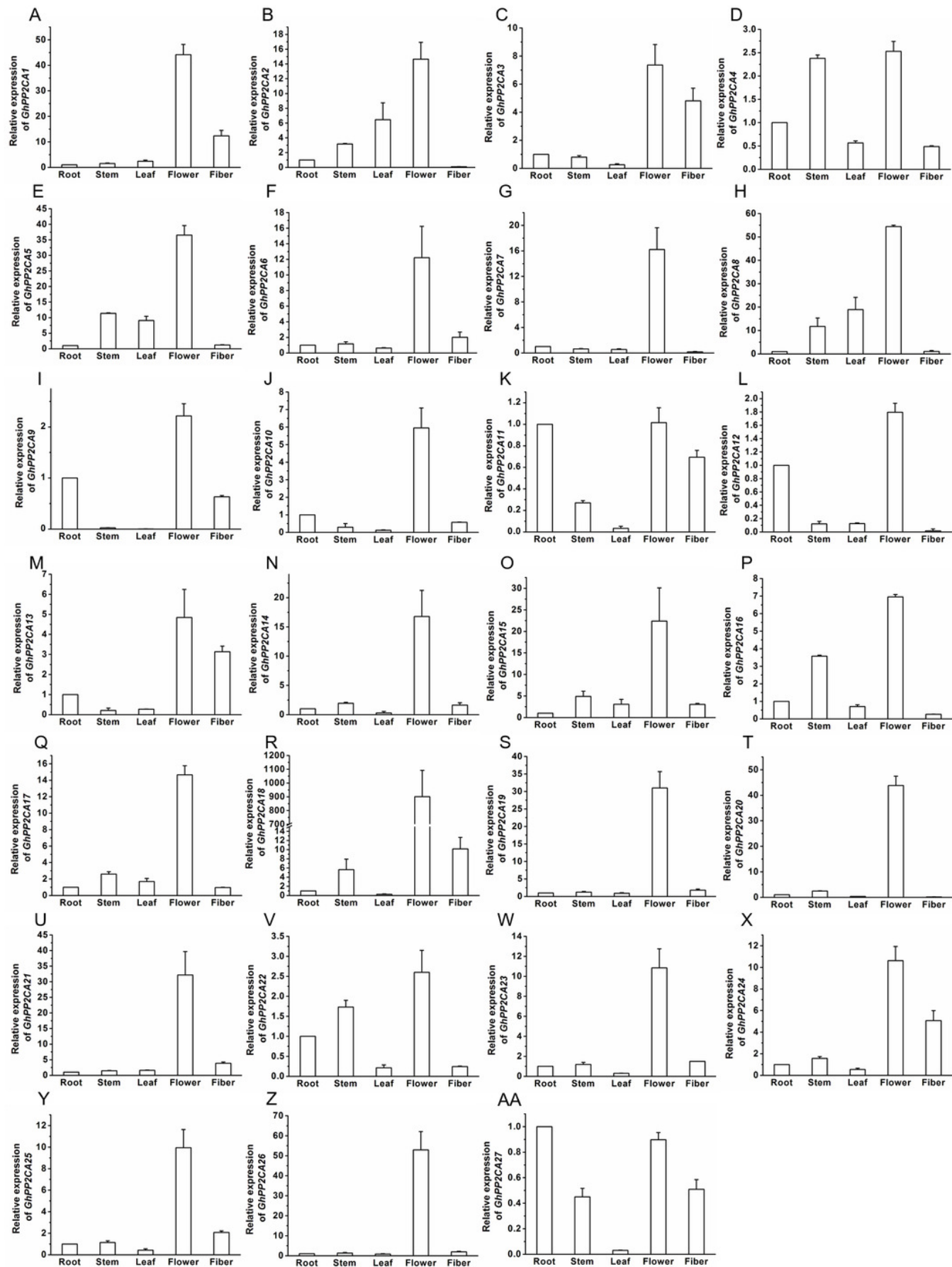



Figure 7

Expression of *GhPP2CA* genes in response to ABA.

The relative expression of *GhPP2CAs* was monitored after treatments with 100 μM ABA for indicated period of time. The expression levels of the genes were prominently reduced at 3 h or 6 h but elevated at 12 h or 24 h (A-E), and were increased at 3 h or 6 h but decreased at 12 h or 24 h (F-O). The transcription abundances of the genes were generally enhanced (P-U), and diminished (V-AA), respectively, with the prolongation of treatment time. Cotton gene *GhUBQ7* was applied as the internal control. The gene expression value at 0 h was set as 1. The values were mean±SE. Statistical analyses were carried out by student's *t*-test to determine the differences in gene abundances between 0 h and other treatment times.

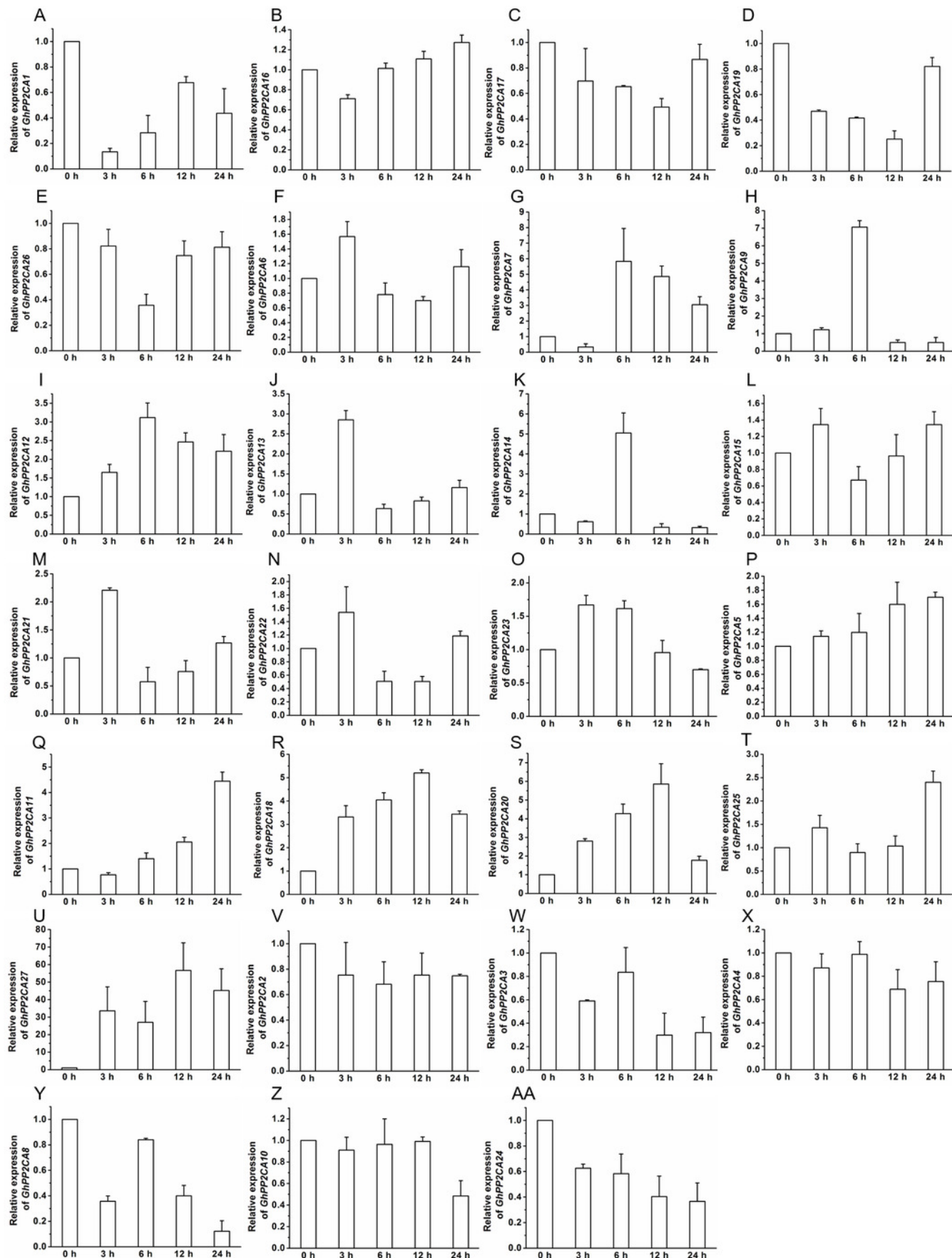



Figure 8

Expression of *GhPP2CA* genes in response to osmotic stress.

The relative expression of *GhPP2CAs* was determined after exposure upon 10% PEG6000 for indicated time . The transcriptional levels of the genes were clearly decreased at 3 h or 6 h but increased at 12 h or 24 h (A-B), and rose at 3 h or 6 h but dropped down at 12 h or 24 h (C-L), the expression of the genes in (M-X) was elevated, and that in (Y-AA) was reduced with the extension of treatment time. *GhUBQ7* was used as the internal control. The gene expression value at 0 h was set as 1. The values were mean±SE. Statistical analyses were carried out by student's *t*-test to determine the differences in gene abundances between 0 h and other treatment times.

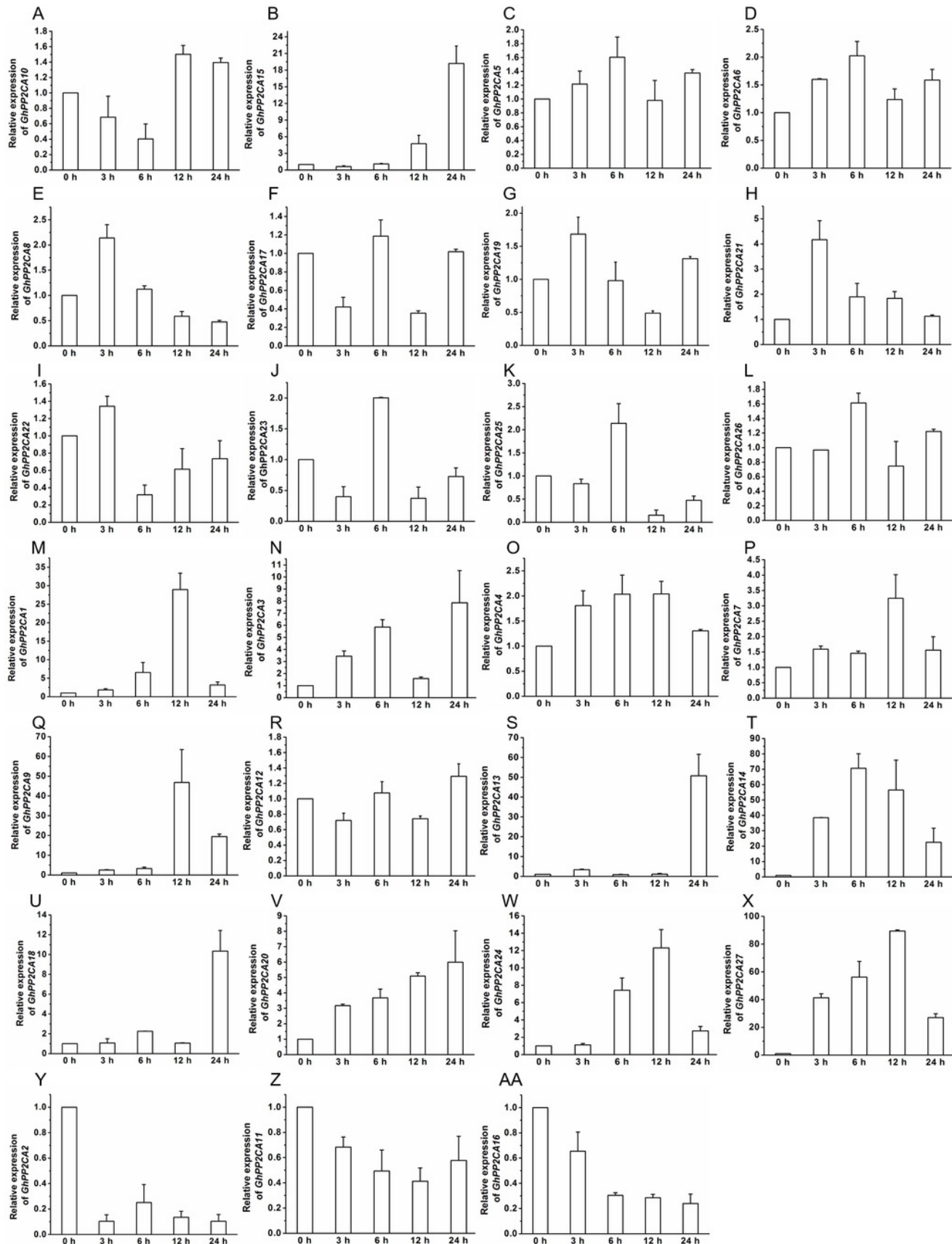



Figure 9

Analysis of interactions between GhPP2CAs and GhPYL2-2D by yeast two-hybrid method.

Narrowing triangles indicate the reduced cell densities in the dilution series from left to right.

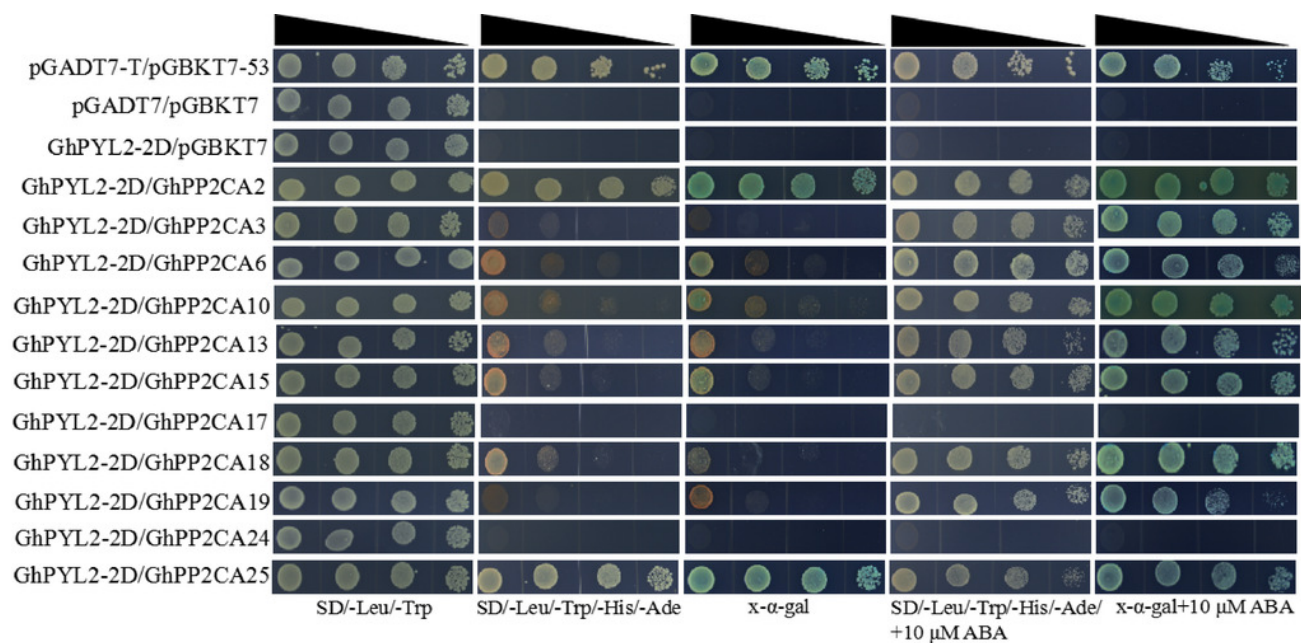



Figure 10

Assay of interactions between GhPP2CAs and GhPYL6-2A using yeast two-hybrid technology.

Narrowing triangles show the reduced cell densities in the dilution series from left to right.

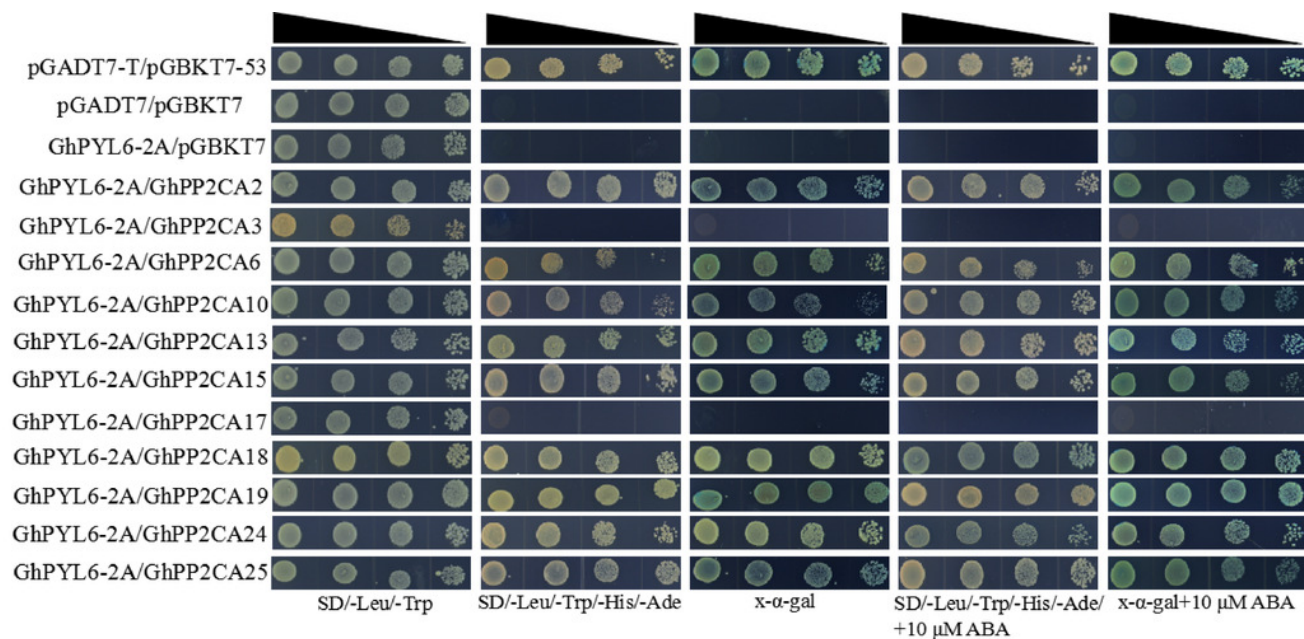
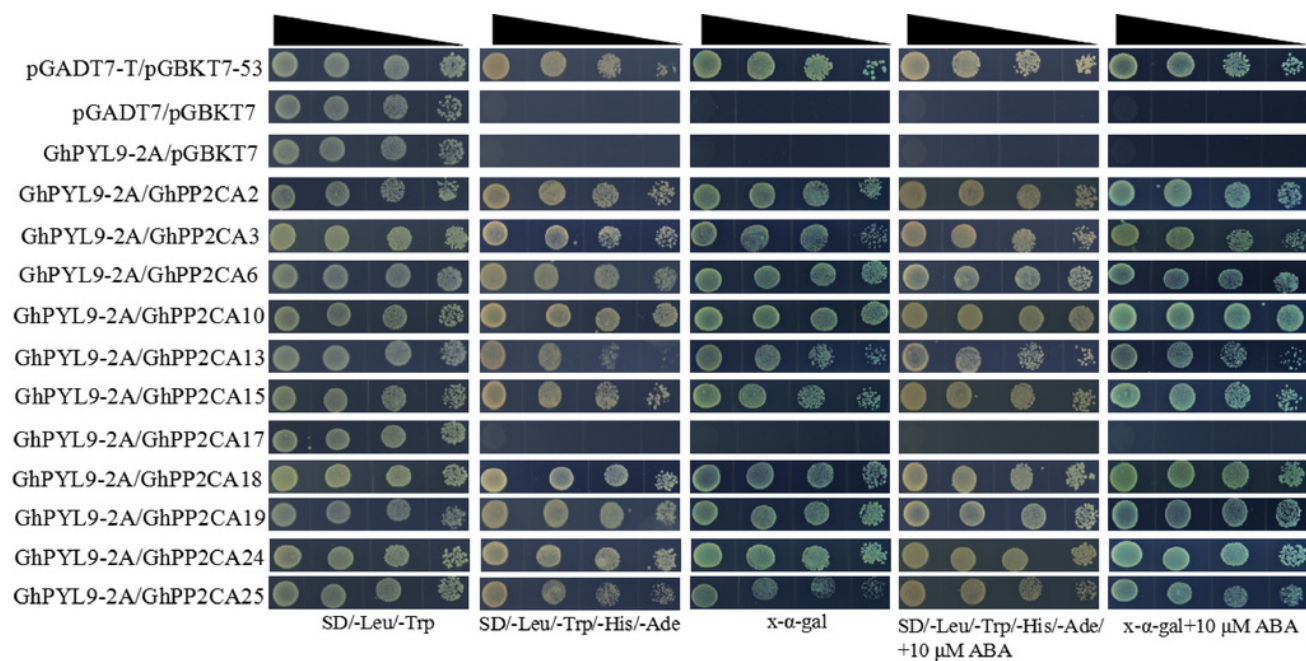



Figure 11

Determination of interactions between GhPP2CAs and GhPYL9-2A by yeast two-hybrid method.

Narrowing triangles reveal the reduced cell densities in the dilution series from left to right.

PP2CA family genes in *Gossypium.*

Gene identifier	Gene name	Size (aa)	Mass (kDa)	pI
Cotton_A_00941	GaPP2CA1	420	46.9	6.61
Cotton_A_01365	GaPP2CA2	419	45.9	6.39
Cotton_A_02676	GaPP2CA3	494	53.4	5
Cotton_A_03112	GaPP2CA4	400	43.4	5.12
Cotton_A_05714	GaPP2CA5	470	51.8	5.82
Cotton_A_11854	GaPP2CA6	593	66	5.51
Cotton_A_12895	GaPP2CA7	416	45.4	6.07
Cotton_A_14028	GaPP2CA8	573	62.4	4.69
Cotton_A_18842	GaPP2CA9	393	43	5.25
Cotton_A_25088	GaPP2CA10	413	45.1	5.15
Cotton_A_25089	GaPP2CA11	413	45.1	5.15
Cotton_A_27758	GaPP2CA12	390	43.3	8.34
Cotton_A_31719	GaPP2CA13	413	45.4	5.55
Cotton_A_34085	GaPP2CA14	242	26.6	5.73
Gh_A03G0373	GhPP2CA1	416	45.9	8.09
Gh_A05G0308	GhPP2CA2	558	60.9	4.69
Gh_A05G0782	GhPP2CA3	416	45.4	6.07
Gh_A05G1136	GhPP2CA4	494	53.5	5.09
Gh_A05G3030	GhPP2CA5	413	45.4	5.52
Gh_A06G0579	GhPP2CA6	416	45.9	5.64
Gh_A07G0123	GhPP2CA7	471	52	5.58
Gh_A08G2192	GhPP2CA8	463	51.4	7.56
Gh_A10G0578	GhPP2CA9	397	42.8	5.76
Gh_A10G1998	GhPP2CA10	409	44.6	5.1
Gh_A12G2380	GhPP2CA11	420	46	6.27
Gh_A13G0184	GhPP2CA12	179	20	5.71
Gh_A13G1741	GhPP2CA13	400	43.3	5.12
Gh_D03G1169	GhPP2CA14	258	28.1	8.6
Gh_D04G0612	GhPP2CA15	413	45.5	5.61
Gh_D05G0410	GhPP2CA16	558	60.8	4.69
Gh_D05G1309	GhPP2CA17	494	53.5	5
Gh_D05G3907	GhPP2CA18	417	45.4	5.83
Gh_D06G0657	GhPP2CA19	393	43	5.49
Gh_D07G2383	GhPP2CA20	471	51.9	5.91
Gh_D08G2557	GhPP2CA21	463	51.3	6.48
Gh_D10G0622	GhPP2CA22	342	38.1	8.74
Gh_D10G2305	GhPP2CA23	411	45	5.05
Gh_D12G2508	GhPP2CA24	418	45.8	6.42
Gh_D13G0199	GhPP2CA25	416	46.2	6.53

Gh_D13G2089	GhPP2CA26	400	43.4	5.17
Gh_Sca051315G01	GhPP2CA27	118	12.8	4.93
GOBAR_AA03316	GbPP2CA1	413	45.4	5.52
GOBAR_AA08348	GbPP2CA2	409	44.6	5.1
GOBAR_AA12929	GbPP2CA3	416	45.8	8.07
GOBAR_AA17958	GbPP2CA4	393	43	5.31
GOBAR_AA26179	GbPP2CA5	418	45.9	6.39
GOBAR_AA27223	GbPP2CA6	456	50.6	7.16
GOBAR_AA30591	GbPP2CA7	470	51.8	5.58
GOBAR_AA32894	GbPP2CA8	558	60.9	4.67
GOBAR_AA34839	GbPP2CA9	416	45.4	6.07
GOBAR_AA37246	GbPP2CA10	558	60.8	4.69
GOBAR_DD04210	GbPP2CA11	416	46.1	6.76
GOBAR_DD07153	GbPP2CA12	371	40.7	6.07
GOBAR_DD08720	GbPP2CA13	392	43.2	7.2
GOBAR_DD10426	GbPP2CA14	385	42.4	6.64
GOBAR_DD12855	GbPP2CA15	302	33.4	6.46
GOBAR_DD17767	GbPP2CA16	389	42	4.96
GOBAR_DD22444	GbPP2CA17	400	43.3	5.11
GOBAR_DD29723	GbPP2CA18	408	44.9	8.52
GOBAR_DD30544	GbPP2CA19	411	45	5.05
GOBAR_DD32331	GbPP2CA20	226	25.2	5.58
GOBAR_DD35228	GbPP2CA21	470	51.8	5.82
GOBAR_DD37768	GbPP2CA22	470	51.8	5.82
GOBAR_DD38067	GbPP2CA23	494	53.5	5
Gorai.001G013500	GrPP2CA1	471	51.9	5.91
Gorai.003G128900	GrPP2CA2	416	45.9	8.09
Gorai.004G284400	GrPP2CA3	463	51.2	6.34
Gorai.008G282600	GrPP2CA4	414	45.4	6.3
Gorai.009G042600	GrPP2CA5	558	60.8	4.65
Gorai.009G096200	GrPP2CA6	416	45.4	5.73
Gorai.009G143300	GrPP2CA7	494	53.4	5.04
Gorai.010G076700	GrPP2CA8	393	43	5.48
Gorai.011G071200	GrPP2CA9	347	38.8	8.43
Gorai.011G268200	GrPP2CA10	411	44.9	5
Gorai.012G072200	GrPP2CA11	409	44.9	5.7
Gorai.013G022100	GrPP2CA12	415	46.1	6.2
Gorai.013G230300	GrPP2CA13	400	43.3	5