

1 **Genome-wide analysis of GATA factors in moso bamboo (*Phyllostachys edulis*)**

2 **unveils that PeGATAs regulate shoot rapid-growth and rhizome development**

3 Taotao Wang¹, Yong Yang¹, Shuaitong Lou², Wei Wei¹, Zhixin Zhao³, Chentao Lin⁴,

4 Liuyin Ma^{2*}

5 ¹Basic Forestry and Proteomics Research Center, College of Forestry, Fujian

6 Provincial Key Laboratory of Haixia Applied Plant Systems Biology, Fujian

7 Agriculture and Forestry University, Fuzhou 350002, China

8 ²Fujian Provincial Key Laboratory of Plant functional Biology, College of Life

9 Sciences, Fuzhou 350002, China

10 ³College of Biopharmaceutical and Food Engineering, Shangluo University, Shangluo

11 726000, China

12 ⁴Department of Molecular, Cell and Developmental Biology, University of California,

13 Los Angeles, CA 90095, USA

14 *Corresponding author:

15 Liuyin Ma: Tel. +86 591 86392267; email: lma223@fafu.edu.cn

16

17

18 **ABSTRACT**

19 **Background**

20 Moso bamboo is well-known for its rapid-growth shoots and widespread rhizomes.

21 However, the regulatory genes of these two processes are largely unexplored.

22 GATA factors regulate many developmental processes, but its role in plant height

23 control and rhizome development remains unclear.

24 **Results**

25 Here, we found that bamboo GATA factors (PeGATAs) are involved in the

26 growth regulation of bamboo shoots and rhizomes. Bioinformatics and

27 evolutionary analysis showed that there are 31 PeGATA factors in bamboo, which

28 can be divided into three subfamilies. Light, hormone, and stress-related

29 *cis*-elements were found in the promoter region of the *PeGATA* genes. Gene

30 expression of 12 *PeGATA* genes was regulated by phytohormone-GA but there

31 was no correlation between auxin and *PeGATA* gene expression. More than 27

32 *PeGATA* genes were differentially expressed in different tissues of rhizomes, and

33 almost all *PeGATAs* have dynamic gene expression level during the rapid-growth

of bamboo shoots. These results indicate that *PeGATAs* regulate rhizome development and bamboo shoot growth partially via GA signaling pathway. In addition, *PeGATA26*, a rapid-growth negative regulatory candidate gene modulated by GA treatment, was overexpressed in Arabidopsis, and over-expression of *PeGATA26* significantly repressed Arabidopsis primary root length and plant height. The *PeGATA26* overexpressing lines were also resistant to exogenous GA treatment, further emphasizing that *PeGATA26* inhibits plant height from Arabidopsis to moso bamboo via GA signaling pathway.

Conclusions

Our results provide an insight into the function of GATA transcription factors in regulating shoot rapid-growth and rhizome development, and provide genetic resources for engineering plant height.

Running title: GATA family in moso bamboo

Keywords: GATA factors; gene family; rapid-growth; shoot; rhizome tissues; gibberellin; plant height; moso bamboo

49

50 **Background**

51 Moso bamboo is one of most-abundant non-timber forestry species and provides
52 important resources for food, architecture, papermaking and fiber [1]. More
53 importantly, moso bamboo is known for its explosive shoot growth rate, with a peak
54 growth rate of 1 meter per day [1]. The rapid-growth shoot is largely dependent on the
55 widespread rhizome system, which provide energy resources by absorbing from soil
56 and more importantly, transporting from other rhizome-connected adult bamboos [2].
57 Therefore, studying the development of shoots and rhizomes will help us understand
58 the rapid-growth regulation of bamboo and provide effective candidate genes for
59 genetic manipulation of crop and forestry species.

60 The GATA factors play important roles in many developmental processes by binding
61 to the consensus DNA sequence (A/T)GATA(A/G) to regulate gene expression at the
62 transcriptional level [3, 4]. The GATA factors have a highly conserved type IV zinc
63 finger DNA binding domain (CX₂CX₁₇₋₂₀CX₂C) and followed by a basic region [5-7].
64 In animals, GATA factors typically contain two zinc finger domains
65 (CX₂CX₁₇₋₂₀CX₂C), while only the C-terminal domain has DNA binding function [5].

66 Animal GATAs are involved in development, differentiation, and control of cell
67 proliferation [7]. However, the fungal GATA factors only contain a single zinc finger
68 domain that is highly similar to the C-terminal zinc finger domain of the animal
69 GATA factors [4, 8]. In plants, GATA factors contain CX₂CX₁₈CX₂C or
70 CX₂CX₂₀CX₂C zinc finger domain [9, 10]. Interestingly, most of plant GATA factors
71 have a single zinc finger domain, and very few of them also contain two zinc finger
72 domains [9-11].

73 In animals, GATA factors involve in cell differential and organ development.
74 Mutations in animal GATA factors cause severe developmental disorder diseases
75 including anemia, deafness, renal and cardiac defects [12]. Fungal GATA factors play
76 roles in nitrogen control, siderophore biosynthesis, light-regulated
77 photomorphogenesis and circadian regulation [4].

78 Plant GATA factors originates from the identification of GATA motifs in regulatory
79 regions of light and circadian clock responsive genes [13]. The first GATA factor
80 identified in plant is NTL1 from *Nicotiana tabacum* [14]. GATA factors have been
81 identified in many plant species, including Arabidopsis (29), rice (28), apple (35) and

82 soybean (64) [9, 10, 15]. Plant GATA factors are involved in many developmental
83 processes, including plant architecture [16], flowering development [17], hypocotyl
84 elongation [18] and seed germination [19]. Plant GATA factors employ several
85 underneath molecular mechanisms, such as modulate nitrogen metabolism [14, 20],
86 act as transcriptional regulator by either integrity of light and phytohormone signal
87 transduction [21, 22] or direct involvement in phytohormone signal transduction to
88 regulate plant growth [23].

89 Plant GATA factors regulate light signal transduction by combining with GATA
90 promoter of light related genes [24, 25]. *GATA2* (At2g45050) has also been identified
91 as a key transcriptional regulator of the integration of light and brassinosteroid
92 signaling pathways [22]. Recent evidences suggest that GATA factors are involved in
93 the regulation of plant hormone signal transduction. Two orthologous GATA-type
94 transcription factors- *GNC* and *CGA1/GNL* from *Arabidopsis thaliana* were identified
95 as GA-regulated genes [21, 23]. Loss-of-function mutants and overexpression lines of
96 *GNC* and *GNL* are functionally related to germination, greening, and flowering time
97 [17]. Chromatin immunoprecipitation (CHIP) results show that these two genes are

98 direct targets of PIF transcription factors, together with the fact that *gnc* and *gnl*
99 loss-of-function mutations suppress *ga1* phenotypes, supporting that GNC and GNL
100 are important repressors of GA signaling [21]. Another important phytohormone, auxin,
101 is also regulated by GNC and GNL through functioning downstream of ARF2 [23]. In
102 addition, the GATA factors are induced by cytokinin [26]. These results indicate that
103 GATA factors play crucial roles in plant development and phytohormone-mediated
104 growth. However, the role of GATA factors in rapid-growth and rhizome development
105 remains elusive.

106 Recently, large-scale transcriptome analysis has shown that light and phytohormones
107 may play important roles in the rapid-growth of bamboo [27-29]. In addition, a large
108 number of transcription factor families are involved in the abiotic stress response and
109 flower development have been studied in moso bamboo [30-32]. Although our group
110 has functionally characterized rapid-growth associated key gene-*PeGSK1*, the
111 rapid-growth associated transcription factor families are largely unexplored in moso
112 bamboo.

113 In this study, we performed genome-wide survey of GATA factors in moso bamboo. A
114 total of 31 GATA factors were identified in the moso bamboo genome. The
115 phylogenetic relationship, gene structure and conserved domains of moso bamboo were
116 systematically analyzed. The phytohormone-related *cis*-element and gene expression
117 of *PeGATAs* under GA and auxin treatment were also characterized. More importantly,
118 the gene expression of *PeGATAs* in different rhizome tissues and rapid-growth shoot
119 were detailed analyzed. In addition, one of growth related *PeGATA*-*PeGATA26* was
120 overexpressed in *Arabidopsis* to functional validate its role in regulating plant height.
121 Overall, our results provide information on the involvement of GATA factors in
122 rhizome tissue development and rapid-growth shoot.

123 **Results**

124 **Genome-wide characterization of GATA factors in moso bamboo**

125 To identify the GATA factors in moso bamboo, the bamboo reference genome was used
126 to scan the GATA factors using HMMER and blast tools
127 (<http://forestry.fafu.edu.cn/db/PhePacBio/download.php>) [33]. A total of 31 potential
128 GATA factors were identified in moso bamboo and named *PeGATA1* to *PeGATA31*

Table 1 GATA factors in moso bamboo

Name	Gene ID	Location	ORF length(bp)	Size (aa)	MW (kDa)	PI	Sub-family
PeGATA1	PH01000001G0820	603239-605821(- stand)	762	253	27.2	7.11	□
PeGATA2	PH01000036G1110	651571-655167(+ stand)	1212	403	42.4	5.2	□
PeGATA3	PH01000040G1560	1013911-1015300(- stand)	801	266	28.7	9.77	•
PeGATA4	PH01000114G0660	460195-464981(+ stand)	912	303	32.2	5.96	•
PeGATA5	PH01000157G0800	521887-523791(- stand)	654	217	23.2	6.15	□
PeGATA6	PH01000162G1360	945504-954318(+ stand)	1500	499	56.4	9.4	□
PeGATA7	PH01000232G0180	85809-87165(- stand)	420	139	15.6	9.23	•
PeGATA8	PH01000242G0460	296415-297844(- stand)	648	215	22.6	8.2	•
PeGATA9	PH01000263G0760	473691-475362(- stand)	1095	364	37.5	7.72	□
PeGATA10	PH01000284G0590	365850-367843(- stand)	1131	376	39.3	5.89	□
PeGATA11	PH01000417G1130	669097-674119(- stand)	948	315	35.8	8.87	□
PeGATA12	PH01000468G1050	681872-683301(+ stand)	663	220	22.7	5.78	•
PeGATA13	PH01000604G0620	351426-352613(+ stand)	399	132	14.6	9.36	•
PeGATA14	PH01000750G0690	435897-442169(+ stand)	777	258	28.1	8.18	•
PeGATA15	PH01000836G0660	444348-452795(- stand)	1413	470	51.6	8.57	•
PeGATA16	PH01000985G0260	141878-143100(+ stand)	402	133	14.8	9.87	•
PeGATA17	PH01001002G0190	175975-177694(+ stand)	831	276	29.4	8.98	•
PeGATA18	PH01001129G0380	296297-297943(+ stand)	699	232	26	9.16	•
PeGATA19	PH01001155G0480	343290-344746(- stand)	741	246	25.3	9.66	□
PeGATA20	PH01001253G0390	263024-267886(- stand)	516	171	18.9	9.95	•
PeGATA21	PH01001451G0450	270274-272851(- stand)	930	309	32.6	8.68	□
PeGATA22	PH01001557G0370	244289-248001(+ stand)	594	197	20.7	9.14	•
PeGATA23	PH01001584G0350	297120-303918(- stand)	1035	344	37.8	4.75	•
PeGATA24	PH01001907G0160	109783-111580(+ stand)	1017	338	36.2	9.26	•
PeGATA25	PH01002105G0190	151795-153666(+ stand)	1248	415	44	8.6	□
PeGATA26	PH01002473G0050	19960-21655(- stand)	1011	336	36.1	9.64	•
PeGATA27	PH01002681G0110	59436-60698(- stand)	690	229	24	8.49	•
PeGATA28	PH01002830G0260	174984-177126(- stand)	381	126	13.8	9.69	•
PeGATA29	PH01003365G0100	53196-55342(+ stand)	369	122	13.3	9.4	•
PeGATA30	PH01003433G0110	88789-92096(- stand)	1167	388	39.8	9.33	□
PeGATA31	PH01004789G0060	58264-61570(- stand)	993	330	35	6.06	□

bp: base pair, aa: amino acids, MW: molecular weight, PI: isoelectric point, kDa: kilodalton

129 based on the chromosomal location. The CDS and protein sequences of PeGATA genes
130 were listed in Additional file 1 and 2: Table S1 and S2. The detailed information of
131 these PeGATA factors including length of CDS, size of amino acid, molecular weight
132 (MW) of protein, gene location on chromosome and isoelectric point (PI) were listed in
133 Table 1.

134 The length of CDS ranges from 366 bp to 1,500 bp, and the length of proteins ranges
135 from 122 aa to 499 aa (Table 1). PeGATA29 is the smallest GATA protein with 122
136 amino acids, and the largest protein is PeGATA6 with 499 amino acids (Table S1). The
137 predicted molecular weight of 31 PeGATA proteins ranges from 13.3 kDa (PeGATA29)
138 to 56.4 kDa (PeGATA6) with an average size of 29.86 kDa (Table 1). The predicted PI
139 of 31 PeGATA factors are all below 10.0, and the minimal protein is PeGATA23 with
140 only 4.75 (Table 1).

141 To further investigate and characterize sequence conservation in the GATA proteins,
142 multiple sequence alignments were performed using the amino acid sequences of the
143 conserved GATA motifs in 31 PeGATAs (Fig. 1). Most bamboo GATA factors contain a
144 single zinc finger domain. However, unlike Arabidopsis, several bamboo GATA factors
145 contain multiple zinc finger domains (Fig. 2). Most of bamboo GATA factors contain
146 18 residues in the zinc finger loop (CX₂CX₁₈CX₂C), while five of them have 20
147 residues in the zinc finger loop (CX₂CX₂₀CX₂C) (Fig. 1). Interestingly, the gene
148 subfamily analysis revealed that all of these five PeGATA factors all belong to the Class
149 C type of the PeGATA family (Fig. 3). Similar to Arabidopsis and rice, moso bamboo

150 does not contain the animal- and fungal-type CX₂CX₁₇CX₂C zinc finger domains (Fig.

151 1). Notably, five PeGATA genes factors have a defective GATA zinc finger domain (Fig.

152 1). PeGATA1 lacks the first Cys residue (-SHC) and PeGATA30 lacks the last Cys

153 residue (CND-). Meanwhile, the GATA factors PeGATA14, PeGATA17 and

154 PeGATA18 have only partial GATA motif (SRLTPAMRRGPTGPRSLCNAC for

155 PeGATA14, CSDCNTTKTPLWRSGPCGPKAA for PeGATA17 and

156 CSDCNTTKTPLWRSGP for PeGATA18) (Fig. 1). The observation is similar to the

157 rice GATA factors as OsGATA24 also contains a partial GATA motif [9]. The results

158 indicated that the bamboo GATA factors have a highly conserved GATA motif,

159 especially compared to rice.

160 To further reveal the diversification of GATA genes in moso bamboo, putative

161 conserved functional domain and motifs were also predicted in the NCBI conserved

162 domain database and program MEME. Through MEME analysis, 10 motifs among the

163 different gene subfamilies is shown in Fig. 2 and the identified multilevel consensus

164 sequence for the motifs is shown in Additional file 3: Table S3. Motif 1 and 5 presented

165 in 29 PeGATA proteins and they were annotated as conserved GATA zinc finger

domain CX₂CX₁₈CX₂C and CX₂CX₂₀CX₂C, respectively (Fig. 2). Motif 5 was not found in PeGATA1, PeGATA14 and PeGATA23 by MEME (Fig. 2), which may be attributed to the zinc finger GATA subfamily domain corresponding to the conserved domain. Motif 2, 7 and 10 appeared nearly all members in subfamily I, and motif 4 and motif 6 only appeared in subfamily II (Fig. 2). Motif 3 was identified as the CCT domain and motif 9 was identified as TIFY domain (Fig. 2). These two domains were specific to subfamily III that was consistent with the classification by conserved domain as shown in Fig. 2. The identification of subfamily-specific motifs from bamboo GATA factors suggests that these motifs may contribute to the functional differences among different subfamilies.

Comparison analysis of the GATA subfamily among Arabidopsis, rice and moso bamboo

GATA factors in Arabidopsis and rice are classified in the clade A-D according to the residues of zinc fingers [9]. To determine the phylogenetic relationship among GATA genes in Arabidopsis, rice and moso bamboo, unrooted phylogenetic tree with 90 GATA factor sequences from all three species was constructed. The phylogenetic tree

analysis shows that all GATA factors have three major clades (Classes A, B and C) (Fig. 3). Among them, Class A is the largest clade and contains 38 members. In this clade, twelve bamboo GATA factors (PeGATA1/2/5/6/9/10/11/19/21/25/30/31) clustered with the Arabidopsis GATA factors AtGATA1, AtGATA2, and AtGATA4, which have been reported to be involved in light regulation of gene expression and photomorphogenesis [22, 34]. Class B formed the second largest clade containing 33 members and 13 bamboo GATA factors (PeGATA3/7/8/12/13/16/17/18/24/26/27/28/29) clustered with the Arabidopsis GATA factors AtGATA21 (GNC) and AtGATA22. These two GATA factors regulates phytohormone response, chlorophyll biosynthesis, starch production, plant architecture, and nitrogen metabolism [17, 21, 23, 34, 35]. In Class C, six bamboo GATA factors (PeGATA4/14/15/20/22/23) clustered with the Arabidopsis GATA factor AtGATA25 (ZIM, Zinc-finger protein expressed in Inflorescence Meristem) and shows the potential roles of hypocotyl and petiole elongation [18]. It is worth noting that no bamboo GATA factor is found in Class D, which explains that bamboo GATA factors may have different functions compared to Arabidopsis and rice.

Gene structure of bamboo GATA genes in moso bamboo

198 To determine the phylogenetic relationships among different members of the GATA
199 factors in moso bamboo, a phylogenetic analysis based on alignments of the 31
200 full-length GATA protein sequences was performed. As shown in Fig. 1 and Fig. 3, the
201 protein sequence alignment and neighbor-joining phylogenetic tree divides 31
202 *PeGATAs* into three clades according to the pattern of zinc finger domain or
203 homologous domains to the *Arabidopsis* and rice GATA factor families. The gene
204 structure of the *PeGATA* genes was shown in Fig. 4. The total exon numbers of
205 *PeGATAs* from each subfamily were calculated. Subfamily I comprised 12 members
206 with two or three exons except *PeGATA19* and *PeGATA6*. *PeGATA19* has only one
207 exon and *PeGATA6* has more than three exons with long introns. Subfamily II consists
208 of 13 members, and all of them contain two or three exons. Subfamily III was formed
209 by included 6 members with five to twelve exons (Fig. 4). The gene structure of GATA
210 factors is similar to that of rice [9]. Overall, the *PeGATA* genes contain exons ranging
211 from one to twelve in its CDS, and the gene structure is obviously different from each
212 other. The results indicated that the bamboo *GATA* genes have undergone significant
213 changes during its long evolutionary history.

214 Identification of hormone-related *cis*-elements in the promoter of the *PeGATA* 215 genes

216 To further explore the function and regulatory pattern of the *PeGATA* gene, the
217 PlantCARE database was used to scan the putative *cis*-elements inside the 1500 bp
218 upstream of transcription start site. We categorized *cis*-elements into four categories
219 based on their functions: light response elements, development, hormone and stress
220 associated *cis*-elements (Fig. 5). The predicted *cis*-elements in *PeGATA* genes were
221 closely related to the function of the GATA family in other plants [17, 19, 22, 23, 25].
222 Light responsive elements like G-box, GT1 and TCT were widely present in the
223 promoter of *PeGATA* genes, and the G-box element has been reported to be involved in
224 the regulation of chlorophyll II biosynthesis in Arabidopsis [36]. We also identified
225 several hormone-responsive *cis*-elements such as ABRE [37], CGTCA-motif, TGACG,
226 and TCA-elements (abscisic acid, MeJA and salicylic acid), and abiotic
227 stress-responsive elements including ARE, GC-motif, LTR and MBS. In addition,
228 tissue specific elements such as CAT-box, circadian responsive element and cell cycle
229 regulation elements like MSA-like were also found in the promoter of the *PeGATA*

230 genes, which may have function in the regulation of plant morphology, flowering and
231 growth [38]. Overall, *cis*-elements analysis indicated that bamboo GATA factors might
232 be involved in response to light and phytohormone to regulate growth.

233 Transcription factors are typically located in the nucleus and regulate transcription of
234 the target genes by binding to the *cis*-elements in their promoters. Consistent with our
235 hypothesis, subcellular localization assays in tobacco showed that randomly selected
236 bamboo *GATA* genes *PeGATA7*, 20, 26 and 28 were clearly localized in the nucleus
237 according to the GFP and DAPI stain signals (Fig. 6). Localization analysis revealed
238 that bamboo GATA factors could also act as transcription factors to regulate gene
239 expression.

240 **Dynamic gene expression pattern of *PeGATAs* in rhizome tissues**

241 The bamboo rhizome system can be divided into three groups: lateral buds, rhizome
242 tips, and new shoot tips [2]. The widespread rhizome system is essential for
243 rapid-growth of bamboo shoot through adopting and utilizing nutrients including
244 nitrate [39]. The GATA factors are also related to nitrogen metabolism in other species
245 [20], so we firstly checked if GATA genes expressing differentially in lateral buds,

rhizome tips, and new shoot tips. By analyze the RNA-seq data from our previous study [2], we showed significant differential expression pattern of *PeGATA* genes among different rhizome tissues (Fig. 7a, Additional file 4: Table S4). As shown in Fig. 7a, a total of 15 *PeGATA* genes (*PeGATA1*, 5, 9, 10, 14, 18, 19, 20, 21, 23, 25, 26, 27, 28 and 29) showed significantly higher expression in lateral buds than that from other two tissues. Five *PeGATA* genes (*PeGATA6*, 7, 8, 11 and 22) highly expressed in the new shoot tips, while reduced their expression in lateral buds. In the rhizome tips, seven *PeGATA* genes (*PeGATA2*, 3, 4, 15, 16, 24 and 30) have remarkable higher expression than other two tissues. In addition, *PeGATA12* and 31 expressed highly both in lateral buds and rhizome tips, while slightly expressed in new shoot tips. Overall, 29 of the 31 *PeGATA* genes showed differential expression in three bamboo rhizome tissues, suggesting that *PeGATA* factors may contribute to the growth regulation of rhizome.

Expression profile of *PeGATAs* in bamboo under the treatment of exogenous phytohormone

GATA factors are closely related to phytohormones to regulate Arabidopsis growth and development [21, 23], take together with the identification of phytohormone related

262 *cis*-elements in the promoter of the bamboo *GATA* genes (Fig. 5b), we rationally
263 hypothesized that the *PeGATA* genes are also tightly regulated by phytohormones. To
264 test our hypothesis, we performed gene expression analysis of the *PeGATA* genes under
265 GA and auxin treatment based on the RNA-seq data published in the previous studies
266 [28, 40]. A total of 12 *PeGATA* genes showed significant gene expression under GA
267 treatment (Fig. 7b, Additional file 5: Table S5). Among them, the expression of
268 *PeGATA7*, 9, 10, 19, 30 and 31 was increased in GA₃ (100 μM) treated seedlings
269 compared to that from untreated control (Fig. 7b). The largest difference was observed
270 in *PeGATA10* (increased by 3.22-fold after GA₃ treatment). In contrast, six genes
271 (*PeGATA1*, 17, 18, 24, 25 and 26) showed lower expression in GA₃-treated seedlings
272 than control seedlings (Fig. 7b). *PeGATA26* was the most down-regulated gene with a
273 56% expression level reduction, and followed by *PeGATA18* with a 54% decline. It is
274 worth noting that the other 19 *PeGATA* genes did not show significant expression
275 change under GA treatment. These results indicate that the gene expression of
276 *PeGATAs* is at least partially regulated by GA.

277 To test the relationship between *PeGATA* gene expression and auxin, we also analyzed
278 the gene expression pattern of *PeGATA* genes under NAA treatment (5 μ M NAA) in
279 bamboo seedlings (Fig. 7c, Additional file 6: Table S6). Interestingly, unlike the results
280 of GA treatment, only *PeGATA8* and *PeGATA9* showed significant gene expression
281 change under auxin treatment. Although the expression levels of some other genes
282 including *PeGATA1*, 2, 10, 16 and 22 were slightly changed, the gene expression of
283 most *PeGATA* genes did not change under auxin treatment (Fig. 7c), suggesting that
284 *PeGATAs* may not be affected by auxin. Overall, these results suggest that *PeGATAs*
285 are partially regulated by GA, but are not affected by auxin.

286 **Genes expression pattern of *PeGATAs* in the rapid-growth of bamboo shoots**

287 As the rapid-growth of bamboo shoots is largely determined by phytohormone and
288 nutrients [2, 27], and we have demonstrated that *PeGATAs* are differentially expressed
289 in rhizome tissues and under GA treatment (Fig. 7a, b), we hypothesized that *PeGATAs*
290 may also be involved in fast-growing bamboo shoots. To validate our hypothesis, the
291 expression profiles of the *PeGATAs* in the fast-growing bamboo shoots (0.15 m, 0.5 m,
292 1.6 m, 4.2 m and 9 m) were examined by qRT-PCR (Fig. 8). The results showed that

almost all GATA genes changed their gene expression in at least one of fast-growing stages. Among them, seven *PeGATA* genes (*PeGATA1*, 3, 9, 12, 14, 16, 26 and 31) continued to decrease their gene expression with the increase of shoot height (Fig. 8). The best example is *PeGATA9*, which showed over 30-fold expression reduction in 9-meter shoots compared to 0.15-meter shoots. The results indicate that these *PeGATAs* genes may be negatively correlated with shoot height. Another groups of *PeGATA* genes (*PeGATA5*, 6, 7, 8, 11, 15, 16, 20, 21, 22, 23, 24, 27, 28, 29 and 30) showed minimal gene expression at the middle growth stages. The results indicate that these *PeGATAs* play an important role in the negative regulation of shoot growth at the middle shoot development stages. Another sets of *PeGATA* genes (*PeGATA2*, 4, 10 and 19) increased their expression at early shoot developmental stages, and then reduce their expression along with the increase of shoot heights. Finally, three *PeGATA* genes (*PeGATA17*, 18 and 27) were increased their expression during early shoot developmental stages, then reduced their expression during middle developmental stages, and then increased their expression again during late shoot developmental stages. Interestingly, we did not find that any *PeGATA* genes continued to increase its

309 expression along with bamboo shoot development. Overall, our results indicate that

310 *PeGATAs* genes may be negatively correlated with rapid-growth of bamboo shoots.

311 **Over-expression of *PeGATA26* negatively regulates plant height in Arabidopsis**

312 To understand the function of *PeGATA* factors, we chose *PeGATA26* as an example to

313 verify the role of *PeGATA* factors in plant growth. *PeGATA26* showed higher gene

314 expression in the growth-inactive lateral buds than the growth-active rhizome tips and

315 new shoot tips (Fig. 7a). Moreover, *PeGATA26* have lower gene expression under GA

316 treatment (Fig. 7b), and its expression decreased along with rapid-growth of bamboo

317 shoots (Fig. 8). These results suggest that *PeGATA26* act as a negative regulator of plant

318 growth and height in moso bamboo. Therefore, we hypothesized that *PeGATA26* plays

319 a crucial role in regulating plant growth. In a previous study, we successfully

320 characterized one of the fast growth-suppressing genes *PeGSK1* by over-expressing it

321 into Arabidopsis [1]. Therefore, we used a similar strategy to verify the function of

322 *PeGATA26* in regulating plant growth by over-expressing it into Arabidopsis.

323 The homozygous T3 transgenic lines were used to analyze phenotype, and

324 over-expression of *PeGATA26* resulted in a significant growth retardation phenotype

(Fig. 9a). Gene expression of *PeGATA26* was successfully detected by qRT-PCR (Fig. 9b). The phenotypes between the two over-expressing lines were similar and the intensity of phenotype correlated with the expression of each transgenic lines (Fig. 9a). Therefore, *PeGATA26* over-expressing line 1 (*PeGATA26-ox1*) with a stronger phenotype was used for further detailed phenotypic analysis. Interestingly, *PeGATA26-ox1* showed a significant dwarf phenotype with a dramatic shorter inflorescence compared to the control (Fig. 9c), indicating that *PeGATA26* inhibits growth of plant height. Moreover, *PeGATA26* also repressed primary root growth (Fig. 9c). However, *PeGATA26* promoted *Arabidopsis* hypocotyl length (Fig. 9c). These results indicate that *PeGATA26* regulates plant growth in a tissue-specific manner: repressing cell growth in roots and inflorescences, while promoting cell growth in hypocotyls.

As *PeGATA26* was down-regulated under the GA treatment in bamboo seedlings (Fig. 7b), we subsequently analyzed whether *PeGATA26* was also regulated by GA in *Arabidopsis*. Interestingly, exogenous GA treatment did not recovery the dwarf phenotype of *PeGATA26-ox1* (Fig. 9d), indicating that *PeGATA26* is negatively

341 correlated with GA to regulate plant height in Arabidopsis. Similar to the results of its
342 Arabidopsis orthologous gene [21], the gene expression pattern of the GA signaling
343 pathway genes in PeGATA26-ox1 were also similar to those observed in the GA
344 biosynthesis mutants- *ga1* (Fig. 9e). The results support that PeGATA26 also repressed
345 GA signaling downstream of the DELLA protein. These results suggest that
346 PeGATA26 inhibits plant root and stem growth in Arabidopsis, take together with its
347 gene expression negatively correlated with the growth of rhizome tissues and shoot in
348 moso bamboo, we concluded that PeGATA26 is a negative growth regulator for plant
349 height control from Arabidopsis to moso bamboo.

350

351 Discussion

352 Moso bamboo is one of important non-timber forestry species with great value in
353 providing food and building materials [33]. Moreover, bamboo is known for its
354 fast-growing shoots and widespread rhizomes [2]. It has been reported that several gene
355 families are involved in flower development and abiotic stress [30-32, 41], the
356 rapid-growth associated transcription factors remain elusive. The genome sequences of

357 moso bamboo[33] and transcriptome studies [2, 27, 28, 40] provide important
358 platforms for the identification of rapid-growth shoot and rhizome development
359 associated gene families. The rapid-growth related genes could provide useful
360 information for genetic manipulation of plant height in future.

361 GATA factors have key functions in developmental control and response to the
362 environmental stresses [16, 17, 19]. In this study, we characterized 31 GATA factors
363 from the moso bamboo (Table 1), and PeGATA factors have highly conserved zinc
364 finger protein domains compared to Arabidopsis and rice GATA factors (Figs. 1-4).

365 More importantly, gene expression analysis of *PeGATAs* in different rhizome tissues
366 and fast-growing bamboo shoots showed that several *PeGATAs* had negative
367 expression patterns with bamboo shoots and rhizome growth (Figs. 7a, 8). Furthermore,
368 the gene expression of *PeGATAs* was partially dependent on phytohormone-GA in
369 bamboo (Figs. 7b). Moreover, overexpression of *PeGATA26* in Arabidopsis repressed
370 the growth of root and plant height in a GA dependent manner (Fig. 9). Overall, our
371 results indicate that *PeGATAs* are involved in regulating the growth of plant height
372 from Arabidopsis to moso bamboo probably through GA signaling pathway.

373 Bioinformatics analysis showed that there were 31 PeGATA factors in moso bamboo
374 (Fig. 1). The number of bamboo GATA factors was closer to other species, including
375 Arabidopsis (29), rice (28) and apple (35) [9, 10, 15]. Furthermore, most of PeGATA
376 factors have a conserved single CX₂CX₁₈₋₂₀CX₂C zinc finger domain that is highly
377 similar to that from Arabidopsis and rice (Fig. 1). In addition, the subfamily of I to III
378 from moso bamboo showed a highly evolutionary conservation compared to
379 Arabidopsis and rice (Fig. 3). These results indicate that most of the GATA factors from
380 moso bamboo are conserved compared to other species. However, unlike containing
381 only one zinc-finger domain in subfamily I GATA factors from Arabidopsis and rice [9],
382 PeGATA6 and PeGATA11 from the bamboo GATA subfamily I have two GATA-type
383 zinc finger domains (Fig. 1). Moreover, more protein domains from the bamboo GATA
384 subfamily III were identified compared to that from Arabidopsis and rice (Fig. 2).
385 Interestingly, a unique feature of the PeGATA factors is that they only have three
386 subfamilies compared to the four subfamilies of Arabidopsis and rice (Fig. 3). These
387 differences suggest that PeGATAs do have certain specificity compared to that from
388 Arabidopsis and rice. Future analysis the functions of GATA factors, including

389 AtGATA26, AtGATA27 and OsGATA30 from subfamily IV (Fig. 3), can help us reveal
390 why bamboo lacks these GATA factors.

391 The first GATA factor is identified according to the light and circadian clock related
392 *cis*-elements in its promoters [13]. Thus, the function of the GATA factors can be
393 predicted based on the identification of *cis*-elements from their promoter. In this study,
394 we found that the promoter of *PeGATAs* has many important *cis*-elements, including
395 light responsive element, cell cycle regulation and phytohormone responsive elements
396 (Fig. 5), which are closely related to the regulation of plant growth. Thus, *PeGATAs*
397 may be involved in regulating plant growth through these *cis*-elements to affect their
398 gene expression and further downstream genes.

399 The bamboo has a well-established rhizome system to develop new shoot tips and
400 widespread rhizome tips [2, 39]. However, the lateral buds of bamboo rhizomes are not
401 active and dominant in growth [2]. Therefore, identification of GATAs with different
402 expression patterns in these tissues will help us clarify the role of GATA factors in
403 rhizome development, which remains unclear. In this study, we found that 15 *PeGATA*
404 genes are highly expressed in lateral buds (Fig. 7a). Among them, *AtGATA2*

(orthologous gene of *PeGATA9*) has been reported to have a function to restrict cell division in the proliferation domain of Arabidopsis root meristem [42], and high expression of *PeGATA9* in lateral buds indicates that *PeGATAs* may also be involved in inhibiting the cell division in bamboo lateral buds (Fig. 7A). In contrast, the lower expression of *PeGATA9* in the actively growing new shoot tips and rhizome tips (Fig. 7A), suggesting a negative correlation between *PeGATA9* and cell growth in bamboo rhizome. It has been reported that *AtGATA22*, a orthologous gene of *PeGATA18*, is involved in response to cytokinin and negatively regulates root growth in Arabidopsis [43], we found that *PeGATA18* has higher gene expression in lateral buds (Fig. 7a), suggesting that *PeGATA18* may play a role in negative regulation of lateral buds cell growth. Similarly, the orthologous gene of *PeGATA26* also plays a negative role in elongation growth [21]. Overall, these results suggest that these 15 *PeGATAs* may contribute to negatively regulating the growth of lateral buds. Next, five *PeGATAs* were highly expressed in the new shoot tips than the other two tissues (Fig. 7a). Among them, the mutation of *GATA19* (orthologous gene of *PeGATA8*) in Arabidopsis causes meristem defects [44]. Here, the high expression of *PeGATA8* in new shoot tips (Fig.

7a), suggests that *PeGATA8* may also be involved in the regulation of shoot meristem development in bamboo. Furthermore, seven *PeGATAs* were highly expressed in the rhizome tips, indicating that they are involved in the growth of the rhizome tips (Fig. 7a). Overall, more than 87% of *PeGATAs* (27/31) was highly expressed in one of rhizome tissues (Fig. 7a). The results indicate that *PeGATAs* strongly participate in the regulation of rhizome growth. Once the transformation system is ready in future, functional characterization of these *PeGATAs* in moso bamboo will help us elucidate the exact role of *PeGATAs* in rhizome growth control.

The correlation between GATA factors and GA or auxin has been extensively studied in Arabidopsis [21, 23, 42, 45]. In this study, we found that gene expression of 12 *PeGATAs* changed under GA treatment, while only two *PeGATAs* responded to auxin treatment (Fig. 7b, c). In addition, motif analysis indicated that the promoter of *PeGATAs* has more GA-related *cis*-elements than auxin (Fig. 5b). Our results indicate that GA rather than auxin frequently regulates the expression of *PeGATAs* in moso bamboo.

436 Gene expression analysis showed that almost all *PeGATAs* have changed their
437 expression during the rapid-growth of bamboo shoots (Fig. 8). For example, *GATA2*
438 (orthologous gene of *PeGATA9*) limits cell division in root meristem of Arabidopsis
439 [42], and the expression of *PeGATA9* was down-regulated over 30 times in late
440 rapid-growth stage (9 m) than the early stage (0.15 m) (Fig. 8). The results indicate that
441 *PeGATA9* may also negatively regulate the rapid-growth of bamboo shoot.
442 Identification of many rapid-growth related *PeGATAs* indicates that *PeGATAs* are
443 involved in regulating the bamboo shoot. The rapid-growth of bamboo shoot is tightly
444 controlled by phytohormone [27]. Current studies reveals that ABA is the only negative
445 regulator of fast-growing shoots, while BR, auxin, GA and cytokinin antagonize with
446 ABA to promote rapid-growth of bamboo shoots [27]. Interestingly, all of the 12
447 GA-related *PeGATAs* showed differential expression in at least one of rapid-growth
448 stages (Figs. 7b, 8), suggesting that GA may regulate rapid-growth of bamboo shoots
449 via modulating gene expression of *PeGATAs*. To understand the function of GA-related
450 *PeGATAs* in plant height control, *PeGATA26* was selected to validate its role in
451 Arabidopsis growth (Fig. 8). Overexpression of *PeGATA26* in Arabidopsis resulted in

452 growth retardation phenotypes such as dwarfism and shorter primary root length, and

453 the *PeGATAs* over-expressed lines was resistant to GA treatment (Fig. 8). Overall, these

454 results further support that *PeGATAs* could regulate plant heights from Arabidopsis to

455 moso bamboo via GA signaling pathway.

456 **Conclusions**

457 With the explosive growth rates of bamboo shoots and widespread rhizomes, the

458 identification of key regulatory genes in the bamboo shoot and rhizome growth control

459 will provide important genetic resources for the genetic manipulation of plant height. In

460 this study, we characterized 31 GATA factors from moso bamboo. More importantly,

461 the gene expression of *PeGATAs* is closely related to the development of rhizome

462 tissues and rapid-growth of bamboo shoots. Moreover, the gene expression of

463 *PeGATAs* was partially regulated by the phytohormone-GA in bamboo. In addition,

464 functional characterization of *PeGATA26* in Arabidopsis provides insight into how

465 *PeGATAs* regulate plant height from Arabidopsis to bamboo via the GA signaling

466 pathway. However, we also noticed that GA regulates expression of only part of

467 *PeGATAs*. As ABA-related *cis*-elements are more widespread than GA, and ABA is the

only known negative regulatory hormones in the rapid-growth control of bamboo shoots, we cannot rule out that *PeGATAs* may also regulate plant height through ABA signaling pathway. In summary, our results provide certain evidence that GATA transcription factor regulate the development of rhizome tissues and the rapid-growth of bamboo shoots.

Methods

Identification of GATA factors in moso bamboo

To identify the GATA factors, the genome and protein sequences of moso bamboo were downloaded from BambooGDB database (<http://forestry.fafu.edu.cn/db/PhePacBio/download.php>) [33]. GATA protein sequences from Arabidopsis and rice were obtained from previous published data [9]. We performed multiple sequence blast and alignment with an expected value of 10. The HMMER profile of the GATA domain (PF00320) from Pfam (<http://pfam.xfam.org/>) was used to search the bamboo protein database with a threshold: e-values < 10⁻⁵ [46]. The bamboo GATA factors were determined with the criteria: it is present in both blast and HMMER motif analysis lists. The number of amino acid, molecular weights (MWs)

484 and isoelectric points (PI) of bamboo GATA factors were predicted by ProtParam
485 (<https://web.expasy.org/protparam/>).

486 **Phylogenetic tree, conserved domain, motif recognition and cis-elements analysis**

487 Multi-sequence alignment of the GATA protein sequences was carried out by ClustalX
488 [47], and phylogenetic tree was constructed using MEGA7 by the Neighbour-Joining
489 method (bootstrap analysis for 1000 replicates) [48]. Conserved domains were obtained
490 from NCBI (<https://www.ncbi.nlm.nih.gov/cdd>) [49] and motifs were analyzed using
491 MEME with default parameters (version 5.0.5, <http://meme-suite.org/tools/meme>) [50].
492 For cis-elements analysis, DNA sequences from 1.5-kb upstream region of each
493 *PeGATA* gene were used to scan any potential cis-element using the PlantCARE
494 database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) [51].

495 **Subcellular localization analysis**

496 To verify the location of PeGATAs, the full-length CDSs without stop codon from four
497 *PeGATA* genes were cloned into a modified pCambia3301 vector with C-terminal GFP
498 as described in our previous study [1]. The *ACTIN2::PeGATAs*-GFP and the
499 *ACTIN2::GFP* control constructs were then transiently transformed into tobaccos, and

500 GFP and DAPI fluorescence was observed using a microscope (20x, Zeiss,
501 LSM880). .

502 **Gene expression analysis**

503 To investigate gene expression levels of the PeGATA genes in different tissues or
504 hormone treatments, RNA-seq data was downloaded from Short Read Archive (SRA)
505 database for the lateral buds, rhizome tips and new shoot tips (SRP093919) [2], and
506 bamboo seedlings under GA and auxin treatment (SRP119416 and SRP109631) [28,
507 40], respectively. The pair-end reads were mapped to the moso bamboo reference
508 genome using tophat2, and differential expressed genes were detected by cufflinks with
509 defaults parameters [52].

510 **Plant materials and qRT-PCR analysis**

511 The moso bamboo shoots used in this study were collected in JianOu County (E118°28';
512 N27°00'), Fujian Province, China. The middle internode of different height of bamboo
513 shoots were sampled and stored in liquid nitrogen immediately.
514 qRT-PCR analysis was performed for each member of the GATA family genes during
515 the rapid-growth of bamboo shoots. Total RNA was extracted from the bamboo

516 samples using HiPure Plant RNA Mini Kit (Magen, R4151-02) and 1µg RNA was
517 taken for reverse transcription into cDNA using a commercial Kit (Monad,
518 RN05004M). Primers for qRT-PCR were designed on Primer3 (<http://primer3.ut.ee/>)
519 using the CDS of each PeGATA gene. qRT-PCR were performed using MonAmp™
520 ChemoHS qPCR Mix (Monad, RN04002M) in a 20 µl reaction. The following program
521 was used for qRT-PCR: 95 °C for 5 min; 40 cycles of 95 °C for 10 s, 60 °C for 10 s and
522 72 °C for 30 s.

523 **Ectopic expression analysis**

524 The *PeGATA26* was cloned and expressed in Arabidopsis exactly following the
525 procedures for the *PeGSK1* in our previous study[1]. The T3 generation seedlings were
526 used for phenotype analysis. The primary root length, plant height and hypocotyl length
527 were measured using ImageJ[1].

528 The primers used in this study were listed in Additional file 7: Table S7.

529 **Declarations**

530 ***Competing interests***

531 The authors declare that they have no competing interests.

532 ***Funding***

533 This work was supported by the National Natural Science Foundation of China (Nos.
534 31741025 and 31500258 to L.M), the Outstanding Youth Research Talents
535 Development Program in Fujian Province University to Liuyin Ma, the Outstanding
536 Youth Research Talents Program of Fujian Agriculture and Forestry University (No.
537 KXJQ17011 to L.M.), and the Scientific Research Foundation of the Graduate School
538 of Fujian Agriculture and Forestry University (to T.W.).

539 ***Author contributions***

540 T.W., C.L., and L.M. conceived the ideas. T.W. and W.W. performed the experiments.
541 T.W., Y.Y., S.L. and Z.Z. contributed to data analysis. T.W. and L.M. wrote the
542 manuscript.

543 **Figure legends**

544 **Fig. 1** Alignment of the amino acid sequences of bamboo GATA factors. The GATA
545 motifs and amino acid positions are marked with a box and an asterisk. The sequence
546 identities of GATA motifs are shown at the bottom.

547 **Fig. 2** Schematic diagram of conserved domain analysis in bamboo GATA proteins.

548 Each color represents a different motif.

549 **Fig. 3** Phylogenetic analysis of GATA factors in bamboo, rice and Arabidopsis. The
550 phylogenetic tree was made based on the amino acid sequences using MEGA7.0 by
551 the neighbor-joining method with 1000 bootstrap replicates. The tree shows four
552 major phylogenetic classes (Classes A to D) indicated by different colors.

553 □ **Fig. 4** Phylogenetic analysis and gene structure of bamboo GATA factors. (a)
554 Phylogenetic tree construction of the PeGATA factors based on the amino acid
555 sequences using MEGA 7.0. The tree showed three major phylogenetic subfamilies
556 (subfamilies I to •) , represented by different colored backgrounds. (b) CDS/UTR
557 structure of the *PeGATA* genes. The yellow and green boxes indicate exons and UTRs,
558 and the black lines represent introns. The size of exons and introns can be estimated
559 using the scale at the bottom.

560 **Fig. 5** *Cis*-elements analysis in the promoter of bamboo *GATA* genes. (a): Overview
561 of the main types of *cis*-elements identified from the 1.5-kb upstream sequence of the
562 bamboo *GATA* genes by the PLANTCARE database. (b): Hormone related
563 *cis*-elements were analyzed and each colored block with numbers represents the

564 number of *cis*-elements in the bamboo GATA promoter.

565 **Fig. 6** Subcellular localization analysis of bamboo GATA factors. The bamboo
566 GATA genes were cloned and constructed in a modified pCambia3301 vector with a
567 C-terminal GFP fusion. These vectors were transformed into tobacco, and the GFP
568 and DAPI signals were captured from the identical areas by microscopy (20x).

569 **Fig. 7** Expression profiles of bamboo GATA genes in different tissues and hormone
570 treatment. (a): The gene expression of 31 bamboo GATA genes in different rhizome
571 tissues was presented by heatmap. (b) and (c): The expression of bamboo GATA
572 genes in the seedlings under GA and auxin treatment. Expression values were
573 normalized and presented at the right side, and green represents lower expression
574 level and red indicates a higher expression level.

575 **Fig. 8** The expression level of GATA genes in bamboo shoots. The gene expression
576 values were detected by qRT-PCR. The Y-axis and X-axis indicate relative expression
577 level at different height of shoots. S1: 0.15 m shoots; S2: 0.5 m shoots; S3: 1.6 m
578 shoots; S4: 4.2 m shoots; S5: 9 m shoots.

579 **Fig. 9** Ectopic expression of *PeGATA26* inhibits the plant height of Arabidopsis. (a):

580 Overexpression of *PeGATA26* resulted in a dwarf phenotype in Arabidopsis. (b):

581 The expression level of *PeGATA26* were detected in both transgenic with the *ACTIN2*

582 gene as an internal control. (c): Phenotypic analysis of plant height, hypocotyl length

583 and primary root length in *PeGATA26* overexpressing line 1 compared to control.

584 They all have significant differences compared to the control by the t-test: $P \leq 0.0001$,

585 which was represented by four stars in the figure. (d): Overexpression of *PeGATA26*

586 repressed plant height of Arabidopsis and *PeGATA26-ox1* was resistant to exogenous

587 GA treatment. (e): The expression of GA signaling genes in the *PeGATA26-ox1* was

588 similar to that of GA biosynthesis mutants-*ga1*. The expression of GA signaling genes

589 was detected by qRT-PCR and the *ACTIN2* was used as the internal control.

590

591 SUPPORTING INFORMATION

592 Additional files

593 **Additional file 1: Table S1.** The coding region sequences of *PeGATA* genes.

594 **Additional file 2: Table S2.** The amino acid sequences of *PeGATA* factors.

595 **Additional file 3: Table S3.** List of protein motifs identified in *PeGATA* factors.

596 **Additional file 4: Table S4.** Gene expression of *PeGATA* genes in different rhizome
597 tissues.

598 **Additional file 5: Table S5.** Gene expression of *PeGATA* genes under GA treatment.

599 **Additional file 6: Table S6.** Gene expression of *PeGATA* genes under auxin
600 treatment.

601 **Additional file 7: Table S7.** List of primers used in this study.

602

603 **References**

- 604 1. Wang T, Li Q, Lou S, Yang Y, Peng L, Lin Z, Hu Q, Ma L: **GSK3/shaggy-like kinase 1**
605 **ubiquitously regulates cell growth from Arabidopsis to Moso bamboo**
606 **(Phyllostachys edulis)**. *Plant Sci* 2019, **283**:290-300.
- 607 2. Wang T, Wang H, Cai D, Gao Y, Zhang H, Wang Y, Lin C, Ma L, Gu L:
608 **Comprehensive profiling of rhizome-associated alternative splicing and**
609 **alternative polyadenylation in moso bamboo (Phyllostachys edulis)**. *Plant J* 2017,
610 **91**:684-699.
- 611 3. Lowry JA, Atchley WR: **Molecular evolution of the GATA family of transcription**
612 **factors: conservation within the DNA-binding domain**. *J Mol Evol* 2000,
613 **50**:103-115.
- 614 4. Scazzocchio C: **The fungal GATA factors**. *Curr Opin Microbiol* 2000, **3**:126-131.
- 615 5. Katsumura KR, Bresnick EH, Group GFM: **The GATA factor revolution in**
616 **hematology**. *Blood* 2017, **129**:2092-2102.
- 617 6. Patient RK, McGhee JD: **The GATA family (vertebrates and invertebrates)**. *Curr*
618 *Opin Genet Dev* 2002, **12**:416-422.
- 619 7. Crispino JD, Horwitz MS: **GATA factor mutations in hematologic disease**. *Blood*
620 2017, **129**:2103-2110.
- 621 8. Teakle GR, Gilmartin PM: **Two forms of type IV zinc-finger motif and their**

kingdom-specific distribution between the flora, fauna and fungi. *Trends Biochem Sci* 1998, **23**:100-102.

9. Reyes JC, Muro-Pastor MI, Florencio FJ: **The GATA family of transcription factors in Arabidopsis and rice.** *Plant Physiol* 2004, **134**:1718-1732.

10. Zhang C, Hou Y, Hao Q, Chen H, Chen L, Yuan S, Shan Z, Zhang X, Yang Z, Qiu D et al: **Genome-wide survey of the soybean GATA transcription factor gene family and expression analysis under low nitrogen stress.** *PLoS One* 2015, **10**:e0125174.

11. Zhang Z, Ren C, Zou L, Wang Y, Li S, Liang Z: **Characterization of the GATA gene family in Vitis vinifera: genome-wide analysis, expression profiles, and involvement in light and phytohormone response.** *Genome* 2018, **61**:713-723.

12. Tremblay M, Sanchez-Ferraz O, Bouchard M: **GATA transcription factors in development and disease.** *Development* 2018, **145**:dev164384.

13. Terzaghi WB, Cashmore AR: **Light-regulated transcription.** *Annual Review of Plant Physiology and Plant Molecular Biology* 1995, **46**:445-474.

14. Daniel-Vedele F, Caboche M: **A tobacco cDNA clone encoding a GATA-1 zinc finger protein homologous to regulators of nitrogen metabolism in fungi.** *Mol Gen Genet* 1993, **240**:365-373.

15. Chen H, Shao H, Li K, Zhang D, Fan S, Li Y, Han M: **Genome-wide identification, evolution, and expression analysis of GATA transcription factors in apple (Malus domestica Borkh.).** *Gene* 2017, **627**:460-472.

16. Hudson D, Guevara DR, Hand AJ, Xu Z, Hao L, Chen X, Zhu T, Bi Y-M, Rothstein SJ: **Rice cytokinin GATA transcription Factor1 regulates chloroplast development and plant architecture.** *Plant Physiol* 2013, **162**:132-144.

17. Richter R, Bastakis E, Schwechheimer C: **Cross-repressive interactions between SOC1 and the GATAs GNC and GNL/CGA1 in the control of greening, cold tolerance, and flowering time in Arabidopsis.** *Plant Physiol* 2013, **162**:1992-2004.

18. Shikata M, Matsuda Y, Ando K, Nishii A, Takemura M, Yokota A, Kohchi T: **Characterization of Arabidopsis ZIM, a member of a novel plant-specific GATA factor gene family.** *J Exp Bot* 2004, **55**:631-639.

19. Liu P-P, Koizuka N, Martin RC, Nonogaki H: **The BME3 (Blue Micropylar End 3) GATA zinc finger transcription factor is a positive regulator of Arabidopsis seed germination.** *Plant Journal* 2005, **44**:960-971.

- 655 20. Fu YH, Marzluf GA: **nit-2, the major nitrogen regulatory gene of *Neurospora***
656 ***crassa*, encodes a protein with a putative zinc finger DNA-binding domain.** *Mol*
657 *Cell Biol* 1990, **10**:1056-1065.
- 658 21. Richter R, Behringer C, Müller IK, Schwechheimer C: **The GATA-type transcription**
659 **factors GNC and GNL/CGA1 repress gibberellin signaling downstream from**
660 **DELLA proteins and PHYTOCHROME-INTERACTING FACTORS.** *Genes Dev* 2010,
661 **24**:2093-2104.
- 662 22. Luo X-M, Lin W-H, Zhu S, Zhu J-Y, Sun Y, Fan X-Y, Cheng M, Hao Y, Oh E, Tian M *et*
663 *al*: **Integration of light- and brassinosteroid-signaling pathways by a GATA**
664 **transcription factor in *Arabidopsis*.** *Dev Cell* 2010, **19**:872-883.
- 665 23. Richter R, Behringer C, Zourelidou M, Schwechheimer C: **Convergence of auxin and**
666 **gibberellin signaling on the regulation of the GATA transcription factors GNC**
667 **and GNL in *Arabidopsis thaliana*.** *Proc Natl Acad Sci U S A* 2013, **110**:13192-13197.
- 668 24. Buzby JS, Yamada T, Tobin EM: **A light-regulated DNA-binding activity interacts**
669 **with a conserved region of a *Lemna gibba* rbcS promoter.** *Plant Cell* 1990,
670 **2**:805-814.
- 671 25. Carre IA, Kay SA: **Multiple DNA-Protein Complexes at a Circadian-Regulated**
672 **Promoter Element.** *Plant Cell* 1995, **7**:2039-2051.
- 673 26. Naito T, Kiba T, Koizumi N, Yamashino T, Mizuno T: **Characterization of a unique**
674 **GATA family gene that responds to both light and cytokinin in *Arabidopsis***
675 ***thaliana*.** *Biosci Biotechnol Biochem* 2007, **71**:1557-1560.
- 676 27. Li L, Cheng Z, Ma Y, Bai Q, Li X, Cao Z, Wu Z, Gao J: **The association of hormone**
677 **signalling genes, transcription and changes in shoot anatomy during moso**
678 **bamboo growth.** *Plant Biotechnol J* 2018, **16**:72-85.
- 679 28. Zhang H, Wang H, Zhu Q, Gao Y, Wang H, Zhao L, Wang Y, Xi F, Wang W, Yang Y *et*
680 *al*: **Transcriptome characterization of moso bamboo (*Phyllostachys edulis*)**
681 **seedlings in response to exogenous gibberellin applications.** *BMC Plant Biol*
682 2018, **18**:125-125.
- 683 29. Zhao H, Lou Y, Sun H, Li L, Wang L, Dong L, Gao Z: **Transcriptome and**
684 **comparative gene expression analysis of *Phyllostachys edulis* in response to**
685 **high light.** *BMC Plant Biol* 2016, **16**:34-34.
- 686 30. Liu H-L, Wu M, Li F, Gao Y-M, Chen F, Xiang Y: **TCP Transcription Factors in Moso**
687 **Bamboo (*Phyllostachys edulis*): Genome-Wide Identification and Expression**

-
- 688 **Analysis.** *Front Plant Sci* 2018, **9**:1263-1263.
- 689 31. Cheng X, Xiong R, Yan H, Gao Y, Liu H, Wu M, Xiang Y: **The trihelix family of**
690 **transcription factors: functional and evolutionary analysis in Moso bamboo**
691 **(Phyllostachys edulis).** *BMC Plant Biol* 2019, **19**:154-154.
- 692 32. Hou D, Cheng Z, Xie L, Li X, Li J, Mu S, Gao J: **The R2R3MYB Gene Family in**
693 **Phyllostachys edulis: Genome-Wide Analysis and Identification of Stress or**
694 **Development-Related R2R3MYBs.** *Front Plant Sci* 2018, **9**:738.
- 695 33. Peng Z, Li L, Zhao Q, Feng Q, Gao Z, Lu H, Hu T, Yao N, Liu K, Li Y *et al*: **The draft**
696 **genome of the fast-growing non-timber forest species moso bamboo**
697 **(Phyllostachys heterocycla).** *Nat Genet* 2013, **45**:456-461.
- 698 34. Jeong MJ, Shih MC: **Interaction of a GATA factor with cis-acting elements**
699 **involved in light regulation of nuclear genes encoding chloroplast**
700 **glyceraldehyde-3-phosphate dehydrogenase in Arabidopsis.** *Biochem Biophys*
701 *Res Commun* 2003, **300**:555-562.
- 702 35. Chiang Y-H, Zubo YO, Tapken W, Kim HJ, Lavanway AM, Howard L, Pilon M, Kieber
703 JJ, Schaller GE: **Functional characterization of the GATA transcription factors**
704 **GNC and CGA1 reveals their key role in chloroplast development, growth, and**
705 **division in Arabidopsis.** *Plant Physiol* 2012, **160**:332-348.
- 706 36. Kobayashi K, Obayashi T, Masuda T: **Role of the G-box element in regulation of**
707 **chlorophyll biosynthesis in Arabidopsis roots.** *Plant Signal Behav* 2012,
708 **7**:922-926.
- 709 37. Yang X, Yang Y-N, Xue L-J, Zou M-J, Liu J-Y, Chen F, Xue H-W: **Rice ABI5-Like1**
710 **regulates abscisic acid and auxin responses by affecting the expression of**
711 **ABRE-containing genes.** *Plant Physiol* 2011, **156**:1397-1409.
- 712 38. Ito M, Iwase M, Kodama H, Lavis P, Komamine A, Nishihama R, Machida Y,
713 Watanabe A: **A novel cis-acting element in promoters of plant B-type cyclin**
714 **genes activates M phase-specific transcription.** *Plant Cell* 1998, **10**:331-341.
- 715 39. Song X, Peng C, Zhou G, Gu H, Li Q, Zhang C: **Dynamic allocation and transfer of**
716 **non-structural carbohydrates, a possible mechanism for the explosive growth**
717 **of Moso bamboo (Phyllostachys heterocycla).** *Sci Rep* 2016, **6**:srep25908.
- 718 40. Wang W, Gu L, Ye S, Zhang H, Cai C, Xiang M, Gao Y, Wang Q, Lin C, Zhu Q:
719 **Genome-wide analysis and transcriptomic profiling of the auxin biosynthesis,**
720 **transport and signaling family genes in moso bamboo (Phyllostachys**

heterocycla). *BMC Genomics* 2017, **18**:870-870.

41. Sun H, Wang S, Lou Y, Zhu C, Zhao H, Li Y, Li X, Gao Z: **Whole-Genome and Expression Analyses of Bamboo Aquaporin Genes Reveal Their Functions Involved in Maintaining Diurnal Water Balance in Bamboo Shoots.** *Cells* 2018, **7**:pii: E195.

42. Jiang K, Yung V, Chiba T, Feldman LJ: **Longitudinal patterning in roots: a GATA2-auxin interaction underlies and maintains the root transition domain.** *Planta* 2018, **247**:831-843.

43. Kollmer I, Werner T, Schmulling T: **Ectopic expression of different cytokinin-regulated transcription factor genes of Arabidopsis thaliana alters plant growth and development.** *J Plant Physiol* 2011, **168**:1320-1327.

44. Zhang X, Zhou Y, Ding L, Wu Z, Liu R, Meyerowitz EM: **Transcription repressor HANABA TARANU controls flower development by integrating the actions of multiple hormones, floral organ specification genes, and GATA3 family genes in Arabidopsis.** *Plant Cell* 2013, **25**:83-101.

45. Ravindran P, Verma V, Stamm P, Kumar PP: **A Novel RGL2-DOF6 Complex Contributes to Primary Seed Dormancy in Arabidopsis thaliana by Regulating a GATA Transcription Factor.** *Mol Plant* 2017, **10**:1307-1320.

46. Finn RD, Tate J, Mistry J, Coghill PC, Sammut SJ, Hotz H-R, Ceric G, Forslund K, Eddy SR, Sonnhammer ELL *et al*: **The Pfam protein families database.** *Nucleic Acids Res* 2008, **36**:D281-D288.

47. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools.** *Nucleic Acids Res* 1997, **25**:4876-4882.

48. Kumar S, Stecher G, Tamura K: **MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.** *Mol Biol Evol* 2016, **33**:1870-1874.

49. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR *et al*: **CDD/SPARCLE: functional classification of proteins via subfamily domain architectures.** *Nucleic Acids Res* 2017, **45**:D200-D203.

50. Bailey TL, Williams N, Misleh C, Li WW: **MEME: discovering and analyzing DNA and protein sequence motifs.** *Nucleic Acids Res* 2006, **34**:W369-W373.

51. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S: **PlantCARE, a database of plant cis-acting regulatory elements and a portal to**

754 **tools for in silico analysis of promoter sequences.** *Nucleic Acids Res* 2002,
755 **30**:325-327.

756 52. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley D, Pimentel H, Salzberg S, Rinn
757 J, Pachter L: **Differential gene and transcript expression analysis of RNA-seq**
758 **experiments with TopHat and Cufflinks.** *Nat Protoc* 2012, **7**:562-578.

759

Fig.1

PeGATA21	RFCSHC	GVQK..	TPQWR	AGPEG	AKTLC	NACSV	RYKSG	RLLPE	YRPAC	SPTFV	NALHS		55			
PeGATA25	RFCSHC	GVQK..	TPQWR	AGPEG	AKTLC	NACSV	RYKSG	RLLPE	YRPAC	SPTYV	SALHS		55			
PeGATA1	..SHC	GVQK..	TPQWR	AGPEG	AKTLC	NACSV	RYKSG	RLLPE	YRPAC	SPTFV	GSIHNSH	..	55			
PeGATA9	RFCTHC	ASEK..	TPQWR	TGHLG	PKTLC	NACSV	RFKSG	RMLPE	YRPAAS	SPTFV	LQHS		55			
PeGATA19	RFCTHC	GSEK..	TPQWR	TGHLG	AKTLC	NACSV	RFKSG	RMLPE	YRPAAS	SPTFV	LQHS		55			
PeGATA2	RFCTHC	QIEK..	TPQWR	AGHLG	PKTLC	NACSV	RYKSG	RLLPE	YRPAAS	SPTFV	PSIHS		55			
PeGATA5	RFCTHC	AVEE..	TPQWR	LGH	DGP	TLCNA	CSVR	FKSG	RLLPE	YRPAN	SPTFS	PLLHS		55		
PeGATA6	KRCTHC	MSYK..	TPQWR	AGHLG	PKTLC	NACSV	RFKSG	RLLPE	YRPAN	SPTFV	SYMHS		55			
PeGATA11	KRCTHC	MSYK..	TPQWR	AGHLG	PKTLC	NACSV	RFKSG	RLLPE	YRPAN	SPTFV	SYMHS		55			
PeGATA31	RFCLHC	ETDK..	TPQWR	TGFMG	PKTLC	NACSV	RYKSG	RLLVQ	YRPAAS	SPTFM	VSKHS		55			
PeGATA10	MFCLHC	ETDR..	TPQWR	TGFMG	PKTLC	NACSV	RYKSG	RLLVQ	YRPAAS	SPTFM	VSKHS		55			
PeGATA30	RFCLHC	ETDK..	TPQWR	TGFMG	PKTLC	NDAV	QVRV	RGAVQ	VGAAG	AVPAV	GEPDL	RR		55		
PeGATA17	RVCSDC	NNTK..	TFLWR	SGGPG	KAAEG	AA	DDGL	RG	GA	KVGT	PSDA	TAHPK	VKKE		55	
PeGATA26	RVCSDC	NNTK..	TFLWR	SGGPG	KSLC	NACG	IRQR	KA	RRAM	MAS	GASTE	GA	KVGT	PS	...	55
PeGATA18	RVCSDC	NNTK..	TFLWR	SGGPG	KVKLL	SRHPL	SLIF	MCSM	CNFAS	YRV	LQV	VT	LIP		55	
PeGATA24	RVCLDC	NNTK..	TFLWR	SGGPG	KSLC	NACG	IRQR	KA	RRAMA	AVTAAA	ANGGA	AGV	G		55	
PeGATA28	RSCVEC	RTTT..	TFMWR	GGGT	GPRSL	CNACG	IRYR	KK	RRQEL	GQDQ	KQP	QQR	GEAT	...	55	
PeGATA29	TSCVEC	GTTT..	TFMWR	GGGT	TRPS	SLCNA	CGIR	YR	KK	RRQEL	GLDQ	KQ	QQH	GEAT	...	55
PeGATA3	KACTDC	HTTK..	TFLWR	GGGPG	KSLC	NACG	IRYR	KK	RRREAL	GLDAGE	GAEC	QQ	QK	...	55	
PeGATA16	KACTDC	HTTK..	TFLWR	GGGPG	KSLC	NACG	IRYR	KK	RRREAL	GLDAGE	GAEC	QQ	QK	...	55	
PeGATA7	KACADCH	TTK..	TFLWR	GGGT	GPKSL	CNACG	IRYR	KK	RRQAL	GLDAT	ETEGA	EC	QQ	...	55	
PeGATA13	KACADCH	TTK..	TFLWR	GGGT	GPKSL	CNACG	IRYR	KK	RRQAL	GLEAAA	EGAE	QQ	QK	...	55	
PeGATA27	RFCA	NCGTTS..	TFLWR	NGRGP	KSLC	NACG	IRF	KKE	ERRAAAA	ESGGA	WCGY	SAQ		55		
PeGATA8	RFCA	NCDTTS..	TFLWR	NGRGP	KSLC	NACG	IRY	KKE	ERRAAAA	AVAPP	P	PQD	SGV	GY	...	55
PeGATA12	NACAN	CDTTS..	TFLWR	NGRGP	KSLC	NACG	IRY	KKE	ERRAAAA	AVAPT	ALPS	D	SGV		55	
PeGATA22	IFCQ	NCGTSEK	MTFAM	RRGPG	PRTLC	NACG	SLM	WANK	GT	LRSC	PKAN	VEAP	LVTI	...	55	
PeGATA23	LF	CQNCGTSEK	MTFAM	RRGPG	PRTLC	NACG	SLM	WANK	FLPY	FLW	AWR	KLFV	DEQ		55	
PeGATA4	SECH	HCGASAT	CTFMM	RRGPG	DGP	TLCNA	CSLM	WANK	ILV	LEAT	SRCH	HCGASAT		55		
PeGATA15	SECH	HCGISAT	LTFMM	RRGPG	DGP	MLCNA	CSLM	WANK	GMM	R	DLS	KAPT	APLR	VVP		55
PeGATA14		THSRL	TFAM	RRGPG	TGPR	SLCNA	CSL	KWANK	GT	LR	PLN	APK	VT	VQHPT	NLSKMC	54
PeGATA20	TH	CQNC	GISSRL	TFAM	RRGPG	TGPR	SLCNA	CSLM	WANK	GT	LR	PLN	APK	MTLQHPA		55

★ ★ GATA motif ★ ★

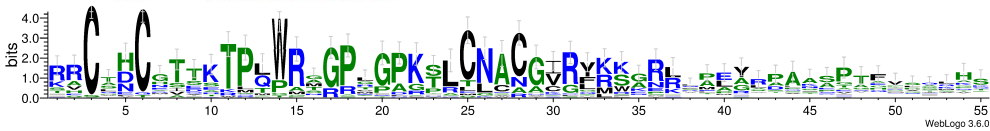


Fig.2

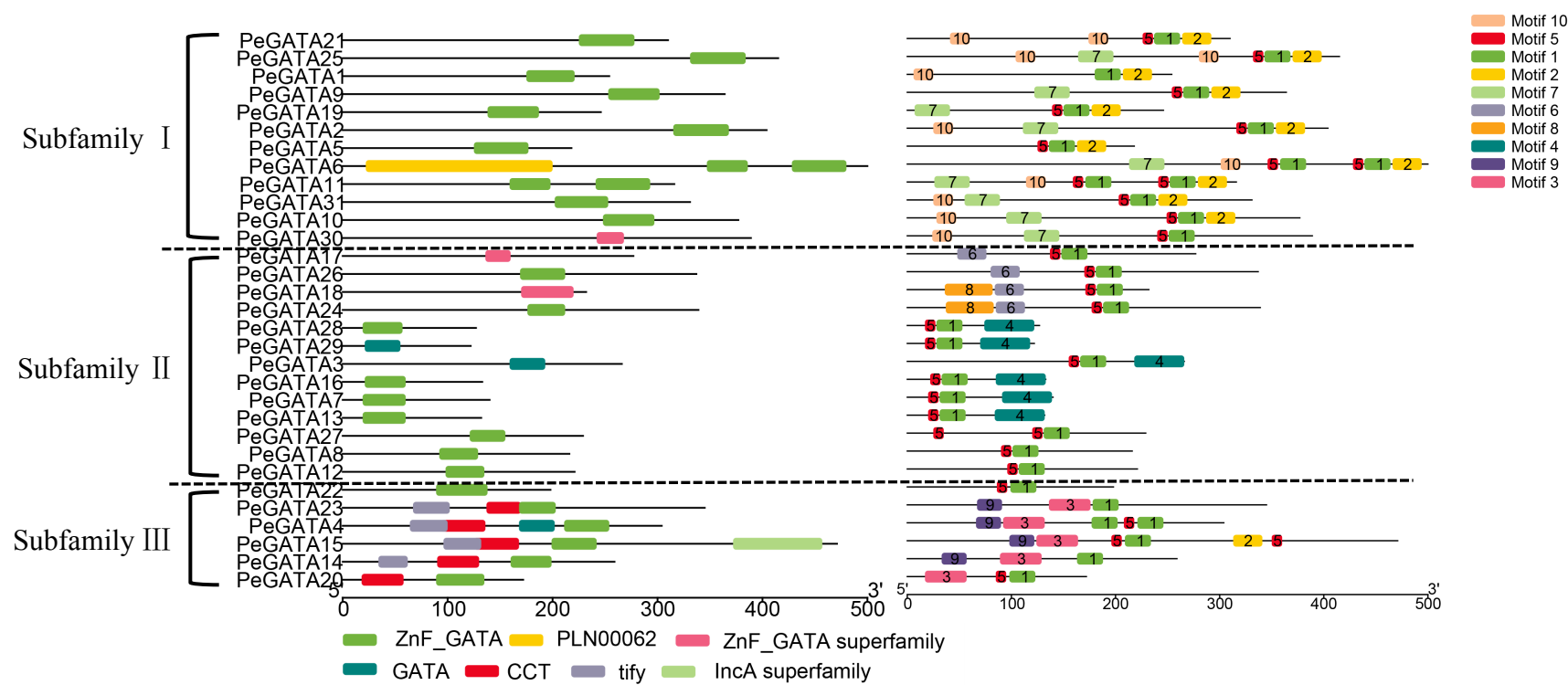


Fig.3

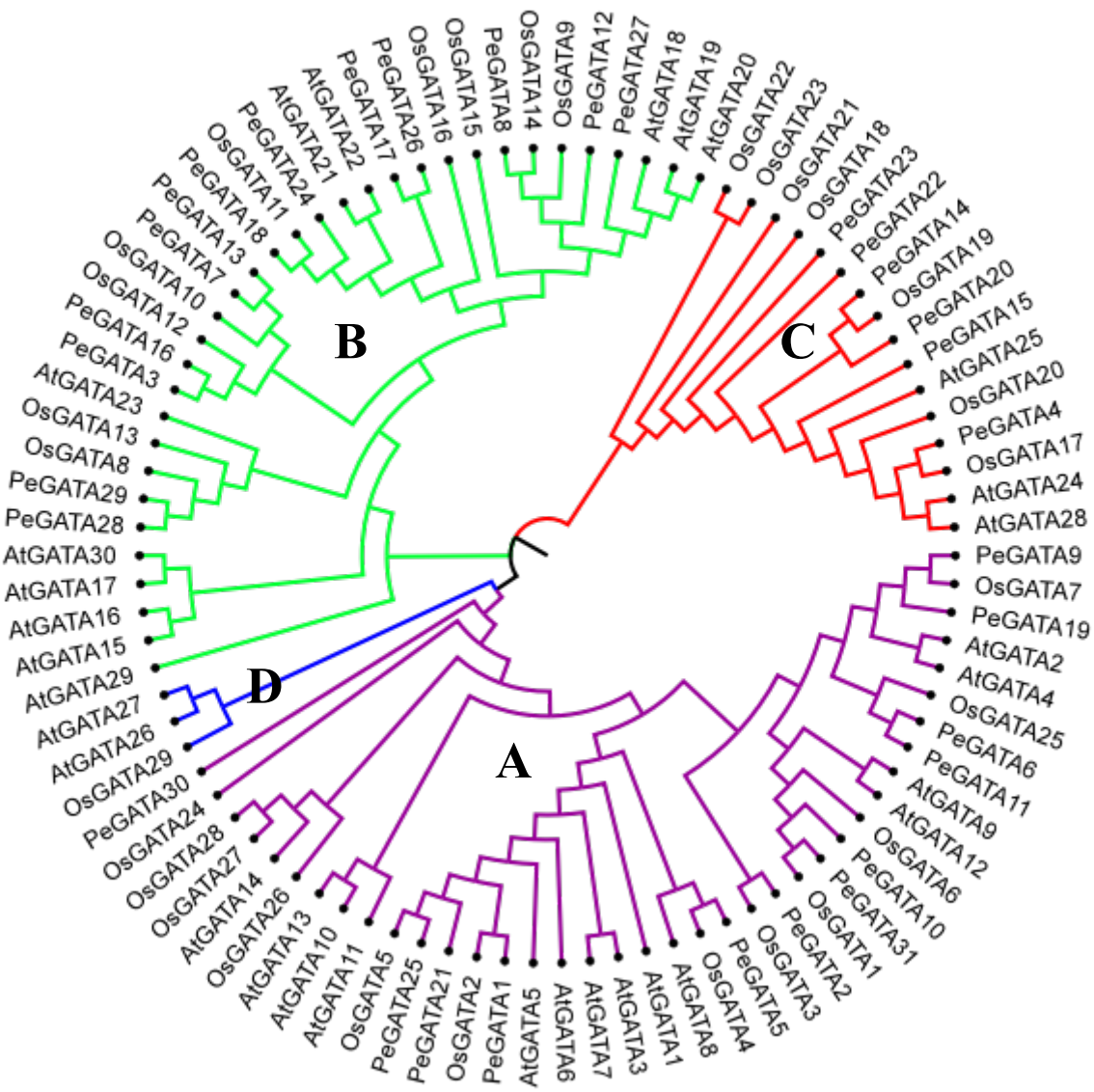


Fig.4

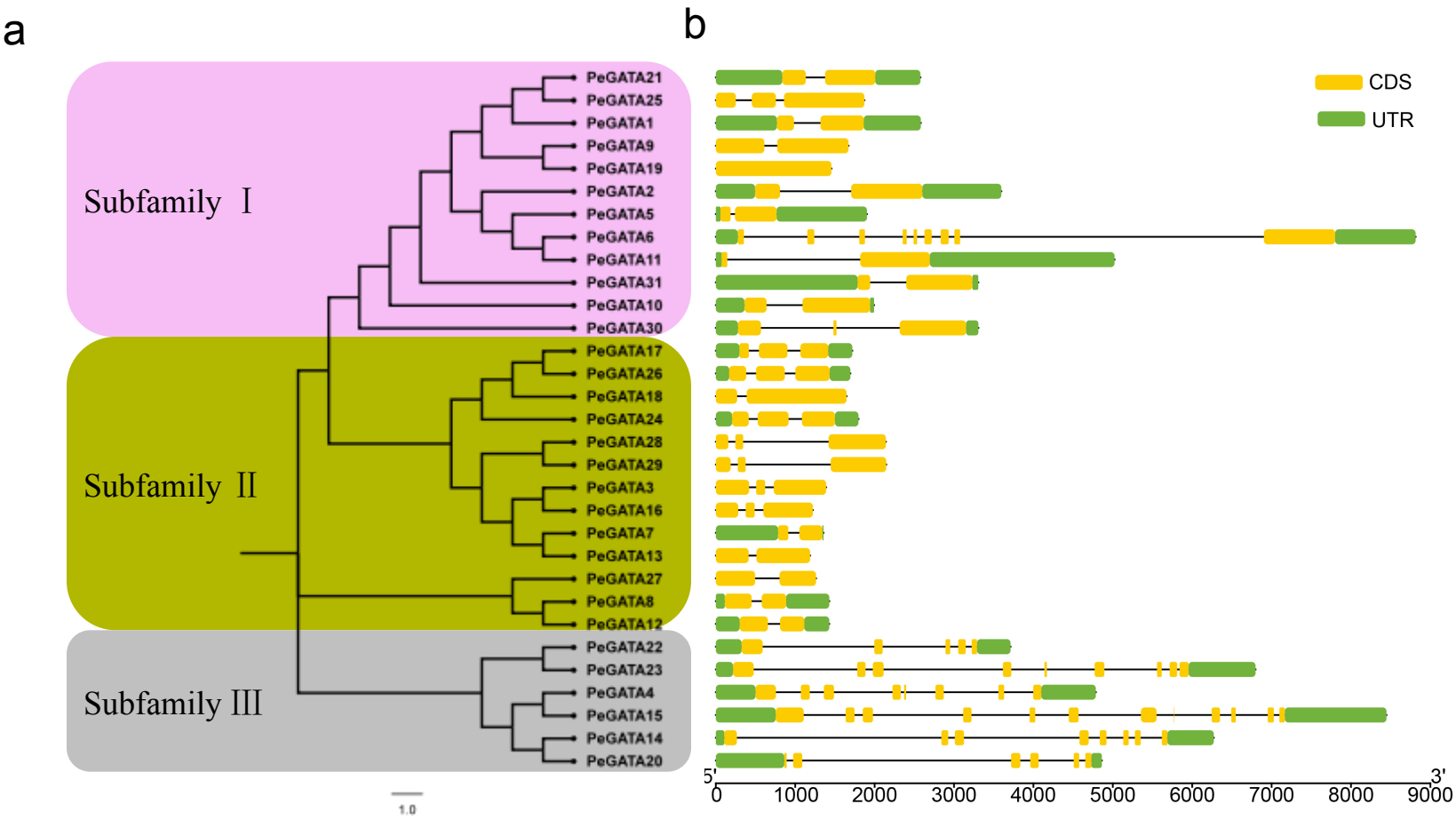


Fig.5

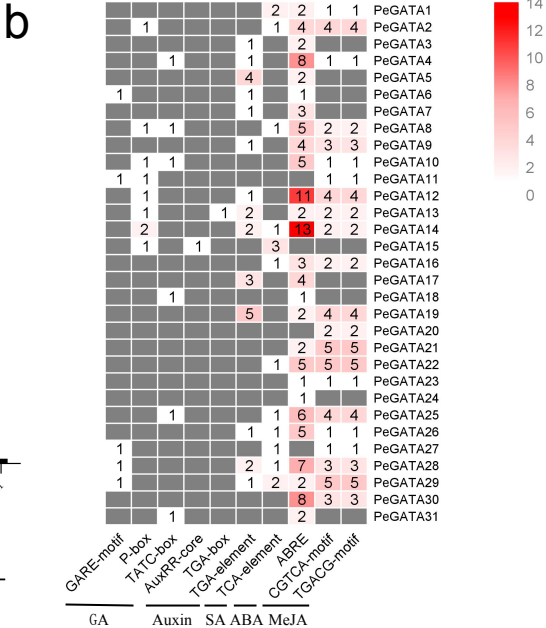
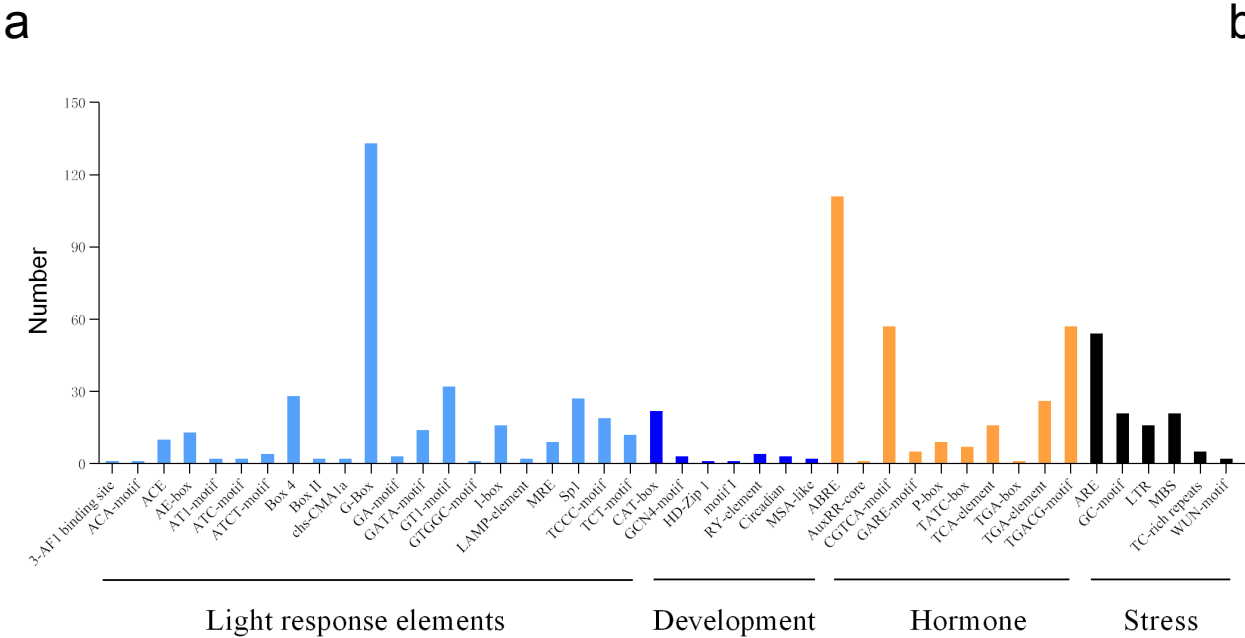


Fig.6

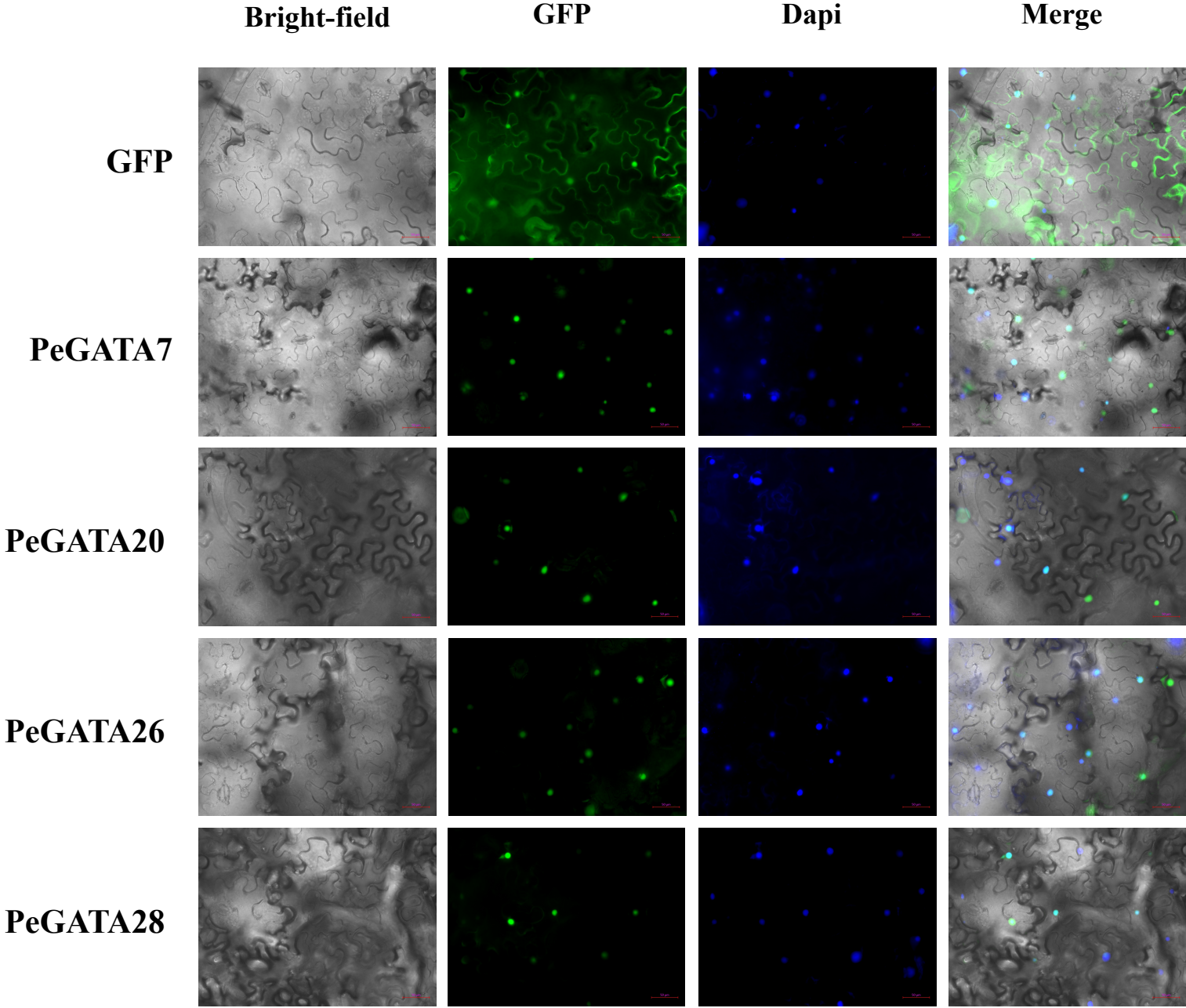
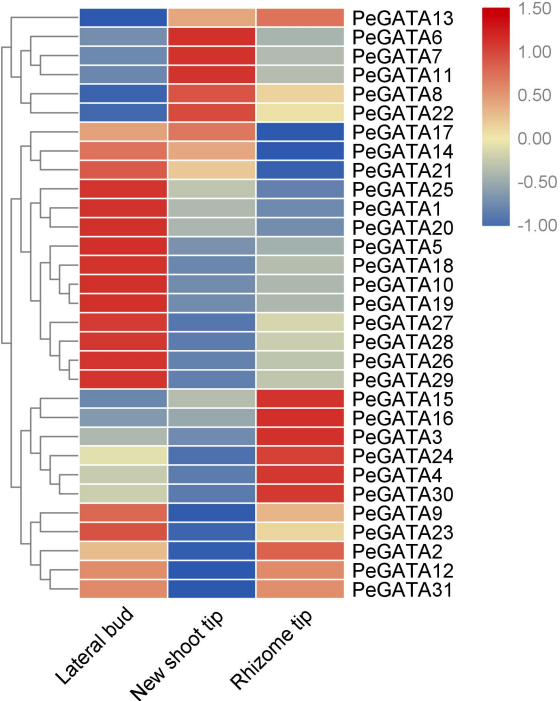
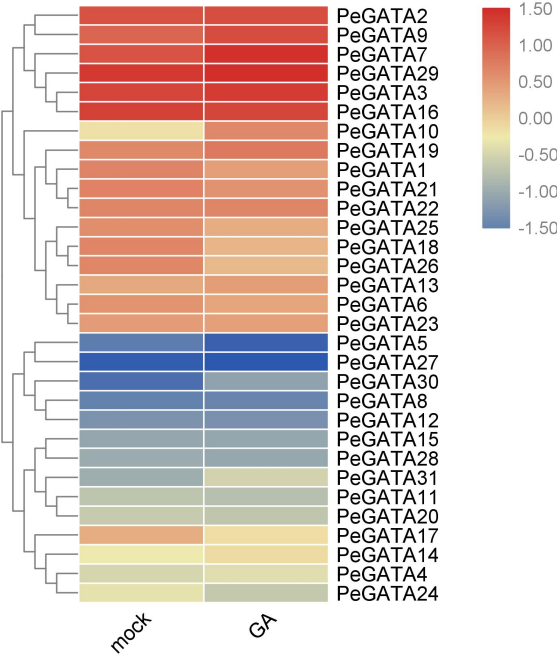


Fig.7

a



b



c

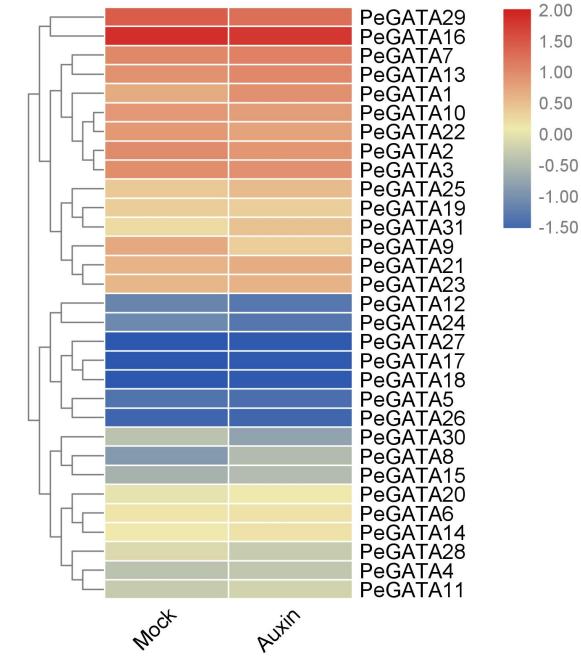


Fig.8

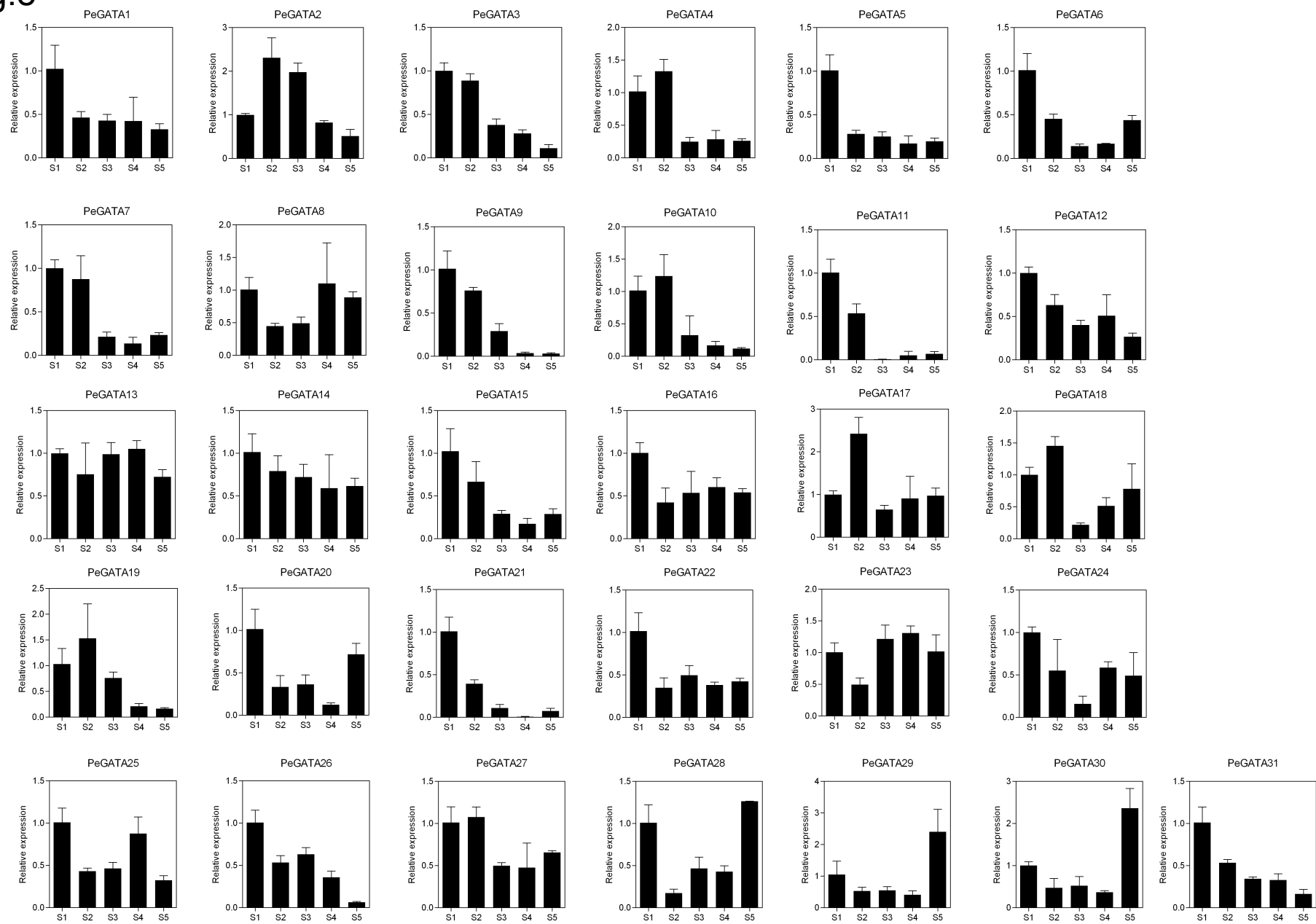
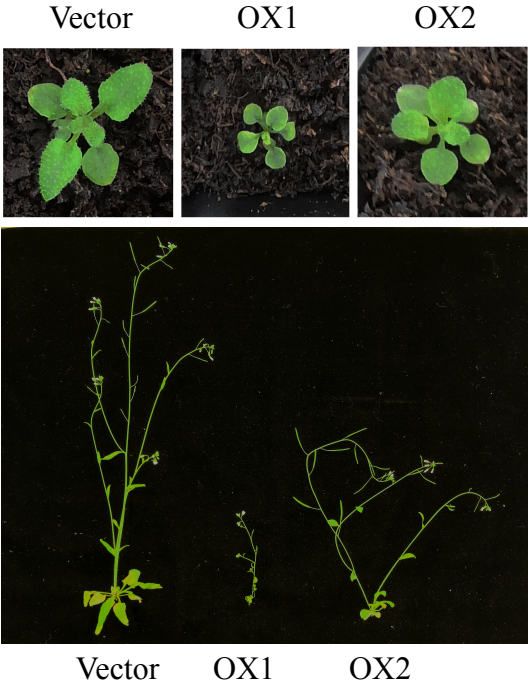
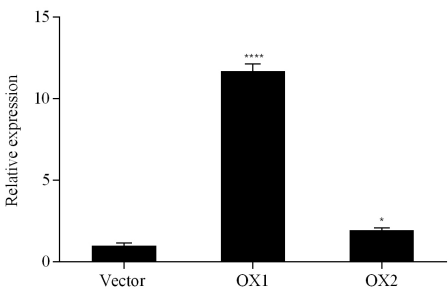


Fig.9

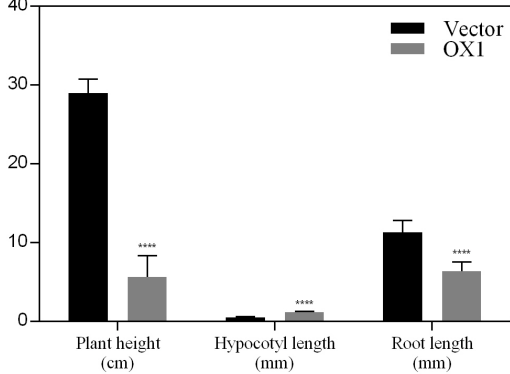
a



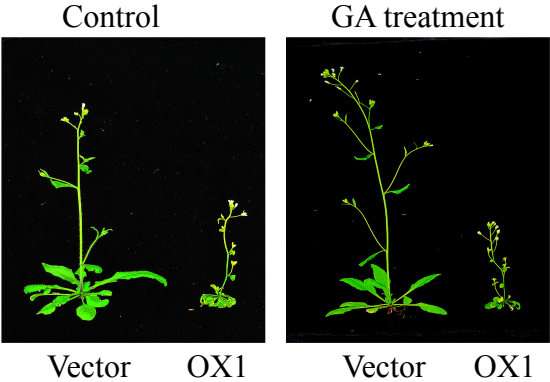
b



c



d



e

