Characterization of the basic helix-loop-helix gene family and its tissue-differential expression in response to salt stress in poplar

Kai Zhao [1] , Shuxuan Li [1] , Wenjing Yao [1, 2] , Boru Zhou [1] , Renhua Li [1] , Tingbo Jiang [Corresp. 1]

[1] State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin, China

[2] Northeast Institute of Geography and Agroecology, Chinese Academy of Science, Harbin, China

Corresponding Author: Tingbo Jiang
Email address: tbjiang@yahoo.com

The basic helix-loop-helix transcription factor gene family is one of the largest gene families and extensively involved in plant growth, development, and stress responses. However, limited studies are available on the gene family in poplar. In this study, we focused on 202 bHLH genes, exploring their DNA and protein sequences and physicochemical properties. According to their protein sequence similarities, we classified the genes into 25 groups with specific motif structures. In order to explore their expressions, we performed gene expression profiling using RNA-Seq and identified 19 genes that display tissue-differential expression patterns without treatment. Furthermore, we also performed gene expression profiling under salt stress. We found 74 differentially expressed genes, which are responsive to the treatment. Eighteen of the 19 genes correspond well to the differentially expressed genes. We validated the results using quantitative real-time PCR. This study lays the foundation for future studies on gene cloning, transgenes, and biological mechanisms.

1 Characterization of the basic helix-loop-helix gene family and its tissue-differential expression in

2 response to salt stress in poplar

3 Kai Zhao[1], Shuxuan Li[1], Wenjing Yao[1, 2], Boru Zhou[1], Renhua Li[1], Tingbo Jiang[1]

4 [1]State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, 51 Hexing
5 Road, Harbin 150040, China
6
7 [2]Northeast Institute of Geography and Agroecology, Chinese Academy of Science, 138 Haping
8 Road, Harbin 150081, China
9
10 Corresponding Author:
11 Tingbo Jiang[1]
12 Email address: tbjiang@yahoo.com

14 **Abstract**

16 The basic helix-loop-helix transcription factor gene family is one of the largest gene families and
17 extensively involved in plant growth, development, and stress responses. However, limited
18 studies are available on the gene family in poplar. In this study, we focused on 202 bHLH genes,
19 exploring their DNA and protein sequences and physicochemical properties. According to their
20 protein sequence similarities, we classified the genes into 25 groups with specific motif
21 structures. In order to explore their expressions, we performed gene expression profiling using
22 RNA-Seq and identified 19 genes that display tissue-differential expression patterns without
23 treatment. Furthermore, we also performed gene expression profiling under salt stress. We found
24 74 differentially expressed genes, which are responsive to the treatment. Eighteen of the 19
25 genes correspond well to the differentially expressed genes. We validated the results using
26 quantitative real-time PCR. This study lays the foundation for future studies on gene cloning,
27 transgenes, and biological mechanisms.

29 **Introduction**

30

31 The bHLH transcription factor gene family is widely existed in eukaryotes and plays an
32 important role in plant growth and development. However, limited studies are available in plants,
33 especially in poplar. Since the first bHLH protein structure was analyzed in 1989 (Murre et al.
34 1989), more bHLH proteins have been identified. The bHLH gene family is named for its highly
35 conserved bHLH domain, which consists of about 60 amino acids, including the basic region
36 distributed at the N-terminus and the HLH region distributed at the C-terminus of the
37 polypeptide chain (Feller et al. 2011). The basic region contains about 15 amino acids and plays
38 an important role in target DNA recognizing and binding (Li et al. 2006). The HLH region,
39 ranging from 40 to 50 amino acids in length, consists of two amphiphilic alpha helices and a loop

40 structure with uncertain length and sequence. The two amphiphilic alpha helices can form
41 homodimer or heterodimeric to interact with other bHLH proteins (Ellenberger et al. 1994). The
42 bHLH transcription factors can identify the element called E-box (5'-CANNTG-3 '), with the
43 most common one is G-Box (5'-CACGTG-3') (Li et al. 2006). Studies have shown that the
44 nucleotides in the flank of the external of core element also impact specific binding (Mart et al.
45 2000). Previous studies indicated that the known bHLH proteins in animals can be divided into
46 six groups of A-F (Atchley & Fitch 1997). In plants, however, many of the identified bHLH
47 proteins belong to group B (Buck & Atchley 2003), the most members of which are
48 characterized by binding to G-box (Li et al. 2006).
49
50 The bHLH transcription factors have various functions in plant growth and development. The
51 bHLH genes *SPT* and *ALC* impact the development of pistils (Groszmann et al. 2010; Heisler et
52 al. 2001), and *PIL5* affects the germination of seeds in *Arabidopsis thaliana* (Oh et al. 2004).
53 The bHLH gene *RGE1* in *Arabidopsis thaliana* plays an important role in controlling the growth
54 of embryos (Kondou et al. 2008), and the bHLH gene *SPT* is associated with final leaf size
55 (Ichihashi et al. 2010). Rice bHLH gene *LAX* is the main regulator of leaf meristem (Komatsu et
56 al. 2003), and rice bHLH gene *bHLH142* is involved in the development process of pollens (Ko
57 et al. 2014). The bHLH genes also regulate plant response to various abiotic stresses, such as
58 drought (Abe et al. 1997; Seo et al. 2011), salinity (Jiang et al. 2009; Zhou et al. 2009), cold
59 (Chinnusamy et al. 2003; Wang et al. 2003), high temperature (Koini et al. 2009), aluminum and
60 iron deficiency (Kumari et al. 2008; Ling et al. 2002; Zhang et al. 2015), and low phosphorus
61 stress (Yi et al. 2005). In addition, the bHLH genes play an important role in regulating multiple
62 signal transduction pathways and impacting biosynthesis (Carretero-Paulet et al. 2010).
63
64 Poplar is an economically important tree in northern China, where salinity is one of the major
65 abiotic stresses that limit poplar survival and growth. Since poplar is susceptible to salt, drought,
66 cold, heavy metals, and other abiotic stresses, it is critical to identify important transcription
67 factor gene families. In this study, we focus on 202 bHLH genes in poplar, in order to investigate
68 their structures and functions, especially tissue-differential gene expression patterns in response
69 to salt stress.
70
71 **Materials and methods**
72
73 **Plant culture and stress treatment**
74
75 We cut the twigs from the di-haploid *Populus simonii* × *Populus nigra* growing in the
76 greenhouse. The twigs were cultivated in the same beakers with water in order to obtain new
77 branches and roots. The conditions of culture include 60-70% relative humidity, 14-h light / 10-h
78 dark cycles, and average temperature of 25 °C. After two months, we selected 20 plants with
79 similar growth status, which were at random divided into two equal groups with two biological
80 replicates. One of the groups was then treated with 150 mM NaCl for 24 hours. The other was

81 used as a control. For gene expression profiling, we collected root, stem, and leaf samples from
82 each group. The 12 samples from each group were frozen in liquid nitrogen immediately, and
83 stored in a refrigerator at -80 °C.
84
85 **Gene expression analysis based on RNA – Seq**
86
87 In order to explore tissue-differential expression patterns within the bHLH gene family in
88 response to salinity, we sent the 12 samples to the GENEWIZ Company
89 (https://www.genewiz.com/) for RNA-Seq using the Illumina HiSeq2500 platform. The
90 sequencing generated pair-end reads with 150 bases. Construction of the RNA libraries was
91 described in our previous study (Yao et al. 2016). Data processing methods for sequence reads
92 are as follows: first, we used the Trimmomatic software (Bolger et al. 2014) to remove the
93 adaptor sequences and low quality sequences from the original sequencing data; second, the high
94 quality reads were mapped to the reference genome using the STAR software (Dobin et al. 2013)
95 with default parameters. Finally, gene expression levels were calculated as fragments per kilo-
96 bases per million mapped reads (FPKM) applying RSEM software (Li & Dewey 2011).
97
98 We identified differentially expressed genes (DEGs) by contrasting the treatment group with the
99 control, using DESeq package (Anders & Huber 2010) from Bioconductor
100 (http://www.bioconductor.org/). We used two thresholds for DEGs selection, that is, fold change
101 >= 2 and padj <= 0.05 (p value adjusted for multiple testing at the false discovery rate of 0.05).
102
103 **Phylogenetic relationship and Physicochemical properties of the bHLH gene family**
104
105 In order to investigate the bHLH gene family in poplar, we obtained their amino acid sequences
106 from the PlantTFDB database (Jin et al. 2017). Then the multiple sequence alignment was
107 carried out by use of ClustalX 1.83, and the phylogenetic trees were constructed with MEGA
108 5.05 (Tamura et al. 2011; Thompson et al. 1997). The parameters are as follows: the neighbor-
109 joining (NJ) method, 1000 iterations of bootstrap resample and poisson model. Since we are
110 interested in general sequences similarities, we then classified the proteins into grand groups. We
111 used bootstrap resample to support the classification. A threshold of the bootstrap score 50 is
112 applied, which means 50 percent of the chances proteins can be found in the same group.
113
114 In addition, we constructed an evolutionary tree, using the best model selected by the SMS
115 software and maximum likelihood (ML) method(Lefort et al. 2017). We obtained the DNA
116 binding domain sequences of poplar bHLH family from the PlantTFDB database, Pfam database
117 and SMART (Finn et al. 2016; Jin et al. 2017; Letunic & Bork 2018; Letunic et al. 2015),
118 followed by performing multiple sequence alignment. The DNA binding domain obtained can be
119 visualized by WebLogo (Crooks et al. 2004; Schneider & Stephens 1990). The physicochemical
120 properties of proteins, including length, molecular weight, theoretical isoelectric point, aliphatic
121 index, and grand average of hydropathicity, were predicted by ProtParam (Gasteiger et al. 2005).

Protein sequence motif analysis

Protein sequence motifs were identified by the MEME method (Bailey et al. 2009), and then the motif logos were obtained using the TBtools (https://github.com/CJ-Chen/TBtools). Annotations of motifs were obtained from Pfam database and InterProScan (Finn et al. 2017; Finn et al. 2016; Jones et al. 2014).

Gene structure analysis

We also downloaded the gene sequences from Phytozome (Goodstein et al. 2012; Tuskan et al. 2006). Structures of the bHLH gene family members were derived using Gene structure Display Server Program (Hu et al. 2015).

Subcellular location and Gene coordinate analysis

Subcellular localization of bHLH proteins was predicted by WoLF PSORT (Horton et al. 2007). Gene coordinate was identified by PopGenIE v3.0 database (Sjodin et al. 2009).

Verification by quantitative real-time PCR

In order to validate our results from RNA-Seq, we performed quantitative real-time PCR with 30 DEGs selected at random. Actin is used as a reference gene (Regier & Frey 2010). The primer sequences are listed in Table S1. The experimental procedures were described in our previous study (Yao et al. 2016). The expression level of each gene was calculated as relative to corresponding gene expression in leaf with no treatment.

Results
Identification of bHLH gene family and analysis of its physicochemical properties

We obtained 202 bHLH gene members in poplar from the PlantTFDB 4.0 (Jin et al. 2017). The gene family shares the DNA-binding domain. We named them *PtrbHLH1* to *PtrbHLH202* (Table S2). The aligned DNA binding domain sequences of poplar bHLH family are showed in Table S3, and the DNA binding domain alignment logo with the length of 60 amino acids is displayed in Fig. 1. Evidence from physicochemical properties indicted that lengths and molecular weights of the 202 corresponding proteins vary substantially (Table S2). The average length is 339.4 amino acids, ranging from 60 to 742. The mean of molecular weights is 37,634.23 Da (6,917.34-79,784.41 Da). The theoretical isoelectric points of these proteins are between 4.63 and 9.92. The aliphatic indexes are in the range of 32.67-108.33, which indicate that thermal stability of the proteins varies substantially. The grand average of hydropathicity runs the gamut of -1.885 and -0.06, indicating that these proteins are hydrophilic.

163
164 **Phylogenetic trees and protein sequence motif analysis**
165
166 In order to explore relationships of the bHLH gene family members in poplar, we constructed a
167 dendrogram and a NJ-phylogenetic tree with their protein sequences (Fig. 2, S1). The proteins
168 can be divided into 25 groups (A-Y). The realibility of the results is supported by the bootstrap
169 resample test (Fig. S1). The average size of the groups has approximately eight members,
170 ranging from two to twenty-six. To validate our phylogenetic tree, we also constructed a ML-
171 phylogenetic tree with the best model of JTT+G+F (Fig. S2). All of groups have a good
172 consistence on the ML-phylogenetic tree except group J and V with only six members.
173
174 As expected, members in each group, especially closely related proteins, have the same or
175 similar conserved motifs. We obtained 15 conserved protein sequence motifs based on the
176 MEME (Bailey et al. 2009). It is shown that motifs 1, 2, 3, 4, 6, 7 were annotated to the bHLH
177 domain, by use of Pfam and InterProScan (Finn et al. 2017; Finn et al. 2016; Jones et al. 2014).
178 Motifs 2 and 4 share the E-box/N-box specific site. Motif 5 is the ACT domain. Motif 8 is
179 achaete-scute transcription factor-related. Motif 9 is the bHLH-MYC and R2R3-MYB
180 transcription factors N-terminal. The others have no annotation (Table S4).
181
182 Based on MEME (Bailey et al. 2009), all the bHLH family genes contain the motif which was
183 annotated as the bHLH domain, except *bHLH202* without any motif (Fig. S3). Group A, B, C, E
184 mainly contain motif 1 and motif 3; thirteen groups largely share motif 2; group O harbors motif
185 4; seven groups mainly have motif 6 and motif 7; and several proteins in multiple groups contain
186 motif 1, 2 or 3. Regarding other motifs, motif 5 is mainly distributed in group K-R; motif 8 in
187 group T; motif 9, 11, 12, 13 in group Q; motif 10 in group B and Y; motif 14 in group I; motif 15
188 in group K, N, O, P, Q. In addition, group F only contains motif 2, group U only harbors motif 6
189 and 7, and group Y only has motif 2 and 10. Motif 4 only occurs in group O.
190
191 **Gene structure at the DNA level**
192
193 In order to compare similarities of the gene structure in each group, we characterized introns and
194 exons of each gene (Fig. S3). In general, the 202 genes differ substantially by length and the
195 number of introns and exons. The majority of the bHLH gene family members have multiple
196 introns, except for a few genes without a single intron (Fig. S3).
197
198 Within each of the 25 protein groups shown in the Fig. 2, multiple subgroups occur based on
199 gene structure. For example, the first four members in group A share the similar number of
200 introns and exons, as well as the length of the exons. Similar situation is observed for the last
201 four members in group C (Fig. S3).
202
203 **Subcellular localization and Gene coordinate**

204

205 In order to characterize cellular distribution of the bHLH proteins, we predicted the localization
206 of these proteins using the WoLF PSORT procedure (Horton et al. 2007). The results show that
207 the 181 bHLH proteins were predicted to locate in the nucleus, six proteins in the cytoplasm, ten
208 proteins in the chloroplast, three proteins in the mitochondria, and one protein in the Golgi
209 apparatus. In addition, one protein was predicted to be double localization with shuttling between
210 the cytoplasm and the nucleus (Table S5).

211

212 In order to examine chromosomal distribution of the bHLH gene family, we applied the
213 PopGenIE v3.0 database (Sjodin et al. 2009). The results indicate that the bHLH gene family
214 members are unevenly distributed on all 19 chromosomes. Chromosome 2 harbors a maximum
215 number of 22 genes; in contrast, each of chromosomes 16, 17, and 18 contains a minimum of
216 only six genes (Table S5). And the distribution is independent on the size of the chromosome.
217 For example, chromosomes 1, the largest one, harbors 18 genes. Chromosomes 9, the minimal
218 one, contains nine genes.

219

220 **Tissue-differential gene expression without salt treatment**

221

222 The sequencing results are given in Table S6. The sequencing depths are approximately tenfold.
223 Total reads for each sample range from 31.8 to 38.3 million. The total mapped reads run the
224 gamut of 27.3 (77.3054%) and 36.4 (95.044%) million. The pearson correlation coefficients of
225 uniquely mapped reads between biological replicates are from 0.961 to 0.994, which are highly
226 significant (p-values < 2.2e-16).

227

228 To explore tissue-differential expression of the gene family without salt treatment, we first
229 compared three sets of the tissue pairs, that is, leaf-root, leaf-stem, and root- stem (Data S1).
230 Then we contrasted between the tissue pairs and retrieved genes that are shared. Results are
231 shown in Fig. 3A-D.

232

233 In the comparison of the leaf–root pair, we identified 73 genes whose expression levels differ
234 significantly. Similarly, we found 45 and 42 genes in the root-stem and leave-stem pairs,
235 respectively (Data S1). We then focused on inter-pair comparisons. Regarding the root-leaf and
236 root-stem pairs, forty genes are shared (Fig. 3B), indicating the gene expressions are different in
237 roots relative to both leaves and stems. Similarly, thirty-four genes are shared in the leaf-root and
238 leaf-stem pairs, followed by 22 genes in the stem-leaf and stem-root pairs (Fig. 3A, 3C). Among
239 the three sets of shared parts, nineteen genes are shared (Fig. 3D and Table S7).

240

241 A heatmap regarding the 19 gene expressions across the three tissues is given in Fig. 4. As
242 expected, there are distinct gene expression patterns for the 19 genes. In general, the 19 genes
243 can be classified into three clusters that are tissue differential. Cluster 1 genes are highly
244 expressed in leaves, lowly expressed in roots, and moderately expressed in stems. Cluster 2

245   genes are greatly expressed in stems, and lowly expressed in roots and leaves. Cluster 3 genes
246   are highly, moderately, and lowly expressed in roots, stems, and leaves. Cluster 1 genes display
247   an opposite pattern to cluster 3 genes.

248

249   In order to explore the direction of differential expression in different comparisons,we
250   classified the tissue-differential genes into up- and down-regulated ones(Fig. S4). We found that
251   most of the tissue-differential genes in leaves have similar trends, for instance, there are
252   respective 15 genes consistently down-regulated and 17 genes agreeably up-regulated in
253   comparison between the leaf-root and leaf-stem tissue pairs. In contrast, only two genes display
254   inconsistency. Similar results are observed in the comparison between the root-leaf and root-stem
255   tissue pairs. However, there is no such trend in the stems.

256

257   **Differential expression genes in response to salinity**

258

259   Screen for salinity responsive members in the bHLH gene family is critical for mechanistic
260   understanding of gene regulation of stress tolerance. In the leaves treated with salt stress, we
261   identified 26 DEGs in the family. Eight of the genes are up-regulated and 18 are down-regulated
262   in treatment samples compared to controls. In roots, forty-seven DEGs were found, with 24 up-
263   regulated and 23 down-regulated. In stems, nineteen DEGs were obtained, with nine up-
264   regulated and 10 down-regulated (Data S2). The number of up-regulated DEGs in roots is larger
265   than the down-regulated DEGs, which is opposite to the DEGs in leaves or stems. Among the
266   URGs, variation of gene expression is remarkably greater in roots compared to that in the other
267   two tissues. In contrast, among the DRGs, leaf tissue displays the widest gene expression
268   variation (Fig. S5).

269

270   The expression degrees of the DEGs in each tissue are shown in Fig. S6. The majority of the
271   DEGs are concentrated at 2-8 fold changes. It is only in roots that six genes with over-expressing
272   of >16 fold changes. In contrast, 11 genes that are down-regulated with >16 fold changes are
273   mainly in leaf tissue.

274

275   **Tissue-differential gene expression in response to salinity**

276

277   In order to observe the distribution of DEGs in different tissues, we drew the Venn Diagram (Fig.
278   3E). There are 74 DEGs that are responsive to salt stress in at least one tissue. There exist 16
279   (21.6%) of the DEGs in any two of the three tissue combinations, with nine (12.2%) in the leaf-
280   stem pair, seven (9.5%) in the root-stem pair, and four (5.4%) in the leaf-root pair. Only 2 (2.7%)
281   DEGs are shared across the three tissues. There are 38 (51.4%) DEGs that are specifically
282   occurred in roots, followed by 15(20.3%) in leaves, and 5 (6.8%) in stems (Fig. 3E). The gene
283   list is given in Table S8. Number of respective DRGs and URGs are shown in Fig. 3F and 3G.
284   Across the three tissues, the majority of DRGs and URGs are in the root tissue.

285

286 Two heatmap plots regarding the expression of 74 DEGs across the three tissues are displayed in
287 Fig. 5. In general, the URGs, which are classified into four clusters, display distinct patterns that
288 contrast between the tissues (Fig. 5A). For example, the top four genes in cluster 1 are highly
289 expressed in roots and stems, but lowly expressed in leaves. Conversely, all the genes in cluster 3
290 display over expression in roots, but low expression in other two tissues. The DRGs, which are
291 classified into four clusters, exhibit similar patterns (Fig. 5B). For example, all the genes in
292 cluster 1 are highly expressed in leaves and stems, but lowly expressed in roots. In contrast,
293 cluster 4 genes display an opposite pattern compared to the cluster 1 genes.
294
295 To connect the tissue-differential genes without salt treatment with the DEGs in response to
296 salinity, we mapped the 19 genes onto the heatmaps of the DEGs (Fig. 5). In general, there is a
297 good correspondence between the matched genes, regarding tissue difference. For example, the
298 cluster 1 genes from Fig. 4 and highlighted in blue in Fig. 5 are highly or moderately expressed
299 in leaves and stems.
300
301 **Validation of the DEGs by RT-qPCR**
302
303 In order to validate the DEGs in response to salt stress, which are identified by RNA-Seq, we
304 selected at random 30 genes for RT-qPCR. Results turn out to be that the trend of the relative
305 gene expression is comparable to that from RNA-Seq (Fig. 6). For example, the salt-responsive
306 DEGs in leaf, such as *Potri.002G054100.1* and *Potri.002G248500.1*, are up-regulated with over
307 two fold changes in both RNA-Seq and RT-qPCR. In contrast, DEGs in leaf, such as
308 *Potri.012G055700.1* and *Potri.009G117300.1*, are down-regulated with over eight fold changes.
309
310 Congruent results are observed for DEGs in stems. For example, the salt-responsive DEGs in
311 stems, such as *Potri.005G158100.1*, *Potri.002G172100.1* and *Potri.008G165700.1*, are over-
312 expressed with over two fold changes in both experiments. *Potri.012G055700.1* is down-
313 regulated with over seven fold changes.
314
315 Similar results are obtained for DEGs in roots. For example, the salt-responsive DEGs in roots,
316 such as *Potri.012G104900.1* and *Potri.004G156000.1*, are up-regulated over two fold changes in
317 both experiments. In contrast, *Potri.002G054100.1* is down-regulated over 3.5 fold changes.
318
319 **Discussion**
320
321 The bHLH gene family is one of the largest families in plants. In previous studies, the bHLH
322 genes in Arabidopsis, poplar, rice, moss and algae had been updated and classified in 2010,and
323 found 183 bHLH genes in *Populus trichocarpa* using Joint Genome Institute (JGI) Ptri version
324 1.1 (Carretero-Paulet et al. 2010). However, the gene ID used in the previous study is
325 inconvenient to be found in the updated database. And the gene annotations in the updated
326 database are also more complete. In this study, we obtained 202 poplar bHLH genes from the

327   PlantTFDB database which identifies transcription factors applying Joint Genome Institute (JGI)
328   Ptri version 3.0 (Jin et al. 2017). And we have characterized the gene and protein structures of
329   the 202 bHLH gene family members. According to the bootstrap values in the NJ-phylogenetic
330   tree, the gene family can be classified into 25 distinct groups. And the groups correspond well to
331   the ML-phylogenetic tree. Each group harbors specific sequence motifs that are corresponding to
332   different protein domains. In addition, each group shares similar intron and exon structures.
333   These results of motif and gene structure analysis are similar to the study of the poplar trihelix
334   family (Wang et al. 2016). The previous study found that the domains outside the bHLH domain
335   may have diverse molecular functions in Arabidopsis (Toledo-Ortiz et al. 2003). For example,
336   the PAS domain of PIF3 plays a major role in the interaction with phytochrome B (Zhu et al.
337   2000). The poplar bHLH transcription factors also contain many different motifs outside the
338   bHLH domain, which suggests that poplar bHLH gene family members have different functions.
339   Since the functional studies of bHLH transcription factors are limited in poplar and there are
340   many amino acid motifs whose functions are unknown in this family, the functional diversity of
341   poplar bHLH transcription factors need to be explored in the further study.
342
343   For the subcellular localization prediction, most of the bHLH proteins were predicted to locate in
344   the nucleus except for 21 members locating in other organelles. We speculate that they may play
345   a role in these organelles and enter the nucleus by interacting with other transcription factors or
346   under stress conditions. And we found that nine of the 21 members inducible by salt stress.
347
348   Due to lack of studies on poplar regarding the gene family, we focused on investigating tissue-
349   differential gene expression. We have identified 19 genes that demonstrate tissue-differential
350   gene expression patterns. The 19 genes can be classified into three distinct clusters.
351
352   To better understand functions of the 19 genes, we then annotated the genes onto the genome of
353   *Arabidopsis thaliana* using Phytozome (Table S9) (Goodstein et al. 2012; Lamesch et al. 2012).
354   The best-hit homologous gene of *Potri.002G108400.1* is *AT5G65640*, which promotes flowering
355   under short-day condition (Sharma et al. 2016). *AT2G24260*, the homologous gene of
356   *Potri.006G186600.1*, plays a role in root development (Lin et al. 2015). *AT1G73830*, the
357   homologous gene of *Potri.012G055700.1*, is required for normal growth of *Arabidopsis thaliana*
358   (Friedrichsen et al. 2002). *AT2G18300*, the homologous gene of *Potri.007G023600.1*, is
359   associated with *Arabidopsis thaliana* growth and immune antagonism (Malinovsky et al. 2014a).
360   These studies suggest that the genes are important in the growth and development of poplar.
361
362   Based on the analysis of the expression levels of bHLH genes under salt stress, we have
363   identified 74 DEGs that are responsive to the stress in at least one tissue. There exist 16 (21.6%)
364   DEGs in the three tissue pairs, with four (5.4%) DEGs in the leaf-root pair, nine (12.2%) in the
365   leaf-stem pair, and seven (9.5%) in the root-stem pair. Only 2 (2.7%) DEGs are shared across the
366   three tissues, suggesting that different tissues may have diversified mechanisms in the regulation
367   of response of salinity in poplar. There are 38 (51.4%) DEGs that are specifically occurred in

368   roots, followed by 15(20.3%) in leaves and 5 (6.8%) in stems (Fig.4e). It is clear that the root
369   and leaf tissues play a more significant role in salt stress responses, compared to stem tissue. The
370   enrichment of DEGs in root is associated with the fact that root is a primary stress perception and
371   response organ.
372
373   By annotating the sequences of the DEGs onto the genome of *Arabidopsis thaliana*, we found
374   that homologous genes in Arabidopsis are associated with plant growth, development and stress
375   response (Table S9) (Goodstein et al. 2012; Lamesch et al. 2012). For example, *AT2G18300*, the
376   best-hit homologous gene of *Potri.005G121900.1*, is related to the growth and immunity of
377   Arabidopsis (Malinovsky et al. 2014b). *AT1G59640*, the best-hit homologous gene of
378   *Potri.008G190800.1*, plays a part in Arabidopsis petal growth (Varaud et al. 2011). *AT3G47640*,
379   the best-hit homologous gene of *Potri.015G142700.1*, regulates response to iron deficiency in
380   Arabidopsis roots (Long et al. 2010). *AT1G51140*, the best-hit homologous gene of
381   *Potri.003G207200.1*, impacts stomatal opening (Takahashi et al. 2013). In this study, the
382   *Potri.003G207200.1* gene was down-regulated with > 32 fold changes in poplar roots challenged
383   with salinity, compared to the control (Data S2). Since the stomata plays an important role in
384   regulating water balance under stress, we speculate that *Potri.003G207200.1* gene expresses in
385   roots may regulate signal transduction pathways that are related to stomatal opening and closing
386   in response to the stress.
387
388   Then, we mapped these genes into the gene family groups, and found that the genes, which show
389   tissue-differential expression patterns, distribute in nine groups unequally, and salt stress
390   response genes widely distribute in 20 groups. It should be noted that many gene pairs among
391   these genes are in the closest position of the phylogenetic tree. It also demonstrates that the genes,
392   which have closer genetic relationship, may have similar functions.
393
394   It is interesting to compare the genes that have tissue-differential expression before and after salt
395   treatment. Before the treatment, we identified 19 significant genes, with 11 over expressed in
396   leaf, followed by five in roots and three in stems (Fig. 4). After the treatment, we found 74
397   significant DEGs in response to salinity (Fig. 3E). The DEGs specific to root, leaf, and stem are
398   38, 15, and 5, respectively (Fig. 3E).
399
400   It is noteworthy that the 18 of the 19 bHLH genes showing tissue-differential expression without
401   treatment correspond well to those DEGs that are responsive to salinity (Fig. 5). We then
402   selected at random 30 DEGs for RT-qPCR for validation. The results from RNA-Seq and RT-
403   qPCR are congruent (Fig. 6).
404
405   **Conclusion**
406
407   In this study, we focus on the 202 bHLH gene family members in poplar, starting from analysis
408   of their physicochemical properties, evolutionary relationship, and gene structures, followed by

409   tissue differential gene expression and differential expression in response to salinity. Around
410   one-third of the genes are found to play a significant role in regulating salinity response. The
411   majority of the DEGs display gene expression patterns in a tissue-differential fashion. This study
412   lays the foundation for future work in gene cloning, transgenes, and biological mechanisms.
413

Reference

415   Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, and Shinozaki K. 1997. Role of Arabidopsis MYC and
416   MYB homologs in drought-and abscisic acid-regulated gene expression. *The Plant Cell* 9:1859-1868.
417   Anders S, and Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol* 11:R106.
418   10.1186/gb-2010-11-10-r106
419   Atchley WR, and Fitch WM. 1997. A natural classification of the basic helix–loop–helix class of transcription factors.
420   *Proceedings of the National Academy of Sciences* 94:5172-5176.
421   Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, and Noble WS. 2009. MEME SUITE:
422   tools for motif discovery and searching. *Nucleic Acids Res* 37:W202-208. 10.1093/nar/gkp335
423   Bolger AM, Lohse M, and Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data.
424   *Bioinformatics* 30:2114.
425   Buck MJ, and Atchley WR. 2003. Phylogenetic analysis of plant basic helix-loop-helix proteins. *J Mol Evol* 56:742-
426   750. 10.1007/s00239-002-2449-3
427   Carretero-Paulet L, Galstyan A, Roig-Villanova I, Martinez-Garcia JF, Bilbao-Castro JR, and Robertson DL. 2010.
428   Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in
429   Arabidopsis, poplar, rice, moss, and algae. *Plant Physiol* 153:1398-1412. 10.1104/pp.110.153593
430   Chinnusamy V, Ohta M, Kanrar S, Lee BH, Hong X, Agarwal M, and Zhu JK. 2003. ICE1: a regulator of cold-induced
431   transcriptome and freezing tolerance in Arabidopsis. *Genes Dev* 17:1043-1054. 10.1101/gad.1077503
432   Crooks GE, Hon G, Chandonia JM, and Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res*
433   14:1188-1190. 10.1101/gr.849004
434   Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR. 2013. STAR:
435   ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15.
436   Ellenberger T, Fass D, Arnaud M, and Harrison SC. 1994. Crystal structure of transcription factor E47: E-box
437   recognition by a basic region helix-loop-helix dimer. *Genes & development* 8:970-980.
438   Feller A, Machemer K, Braun EL, and Grotewold E. 2011. Evolutionary and comparative analysis of MYB and bHLH
439   plant transcription factors. *Plant J* 66:94-116. 10.1111/j.1365-313X.2010.04459.x
440   Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztanyi Z, El-Gebali S, Fraser M,
441   Gough J, Haft D, Holliday GL, Huang H, Huang X, Letunic I, Lopez R, Lu S, Marchler-Bauer A, Mi H, Mistry J,
442   Natale DA, Necci M, Nuka G, Orengo CA, Park Y, Pesseat S, Piovesan D, Potter SC, Rawlings ND, Redaschi N,
443   Richardson L, Rivoire C, Sangrador-Vegas A, Sigrist C, Sillitoe I, Smithers B, Squizzato S, Sutton G, Thanki N,
444   Thomas PD, Tosatto SC, Wu CH, Xenarios I, Yeh LS, Young SY, and Mitchell AL. 2017. InterPro in 2017-
445   beyond protein family and domain annotations. *Nucleic Acids Res* 45:D190-D199. 10.1093/nar/gkw1107
446   Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A,
447   Salazar GA, Tate J, and Bateman A. 2016. The Pfam protein families database: towards a more sustainable
448   future. *Nucleic Acids Res* 44:D279-285. 10.1093/nar/gkv1344
449   Friedrichsen DM, Nemhauser J, Muramitsu T, Maloof JN, Alonso J, Ecker JR, Furuya M, and Chory J. 2002. Three

450   redundant brassinosteroid early response genes encode putative bHLH transcription factors required for
451   normal growth. *Genetics* 162:1445-1456.
452 Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, and Bairoch A. 2005. Protein Identification
453   and Analysis Tools on the ExPASy Server. In: Walker JM, ed. *The Proteomics Protocols Handbook*. Totowa,
454   NJ: Humana Press, 571-607.
455 Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, and
456   Rokhsar DS. 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res*
457   40:D1178-1186. 10.1093/nar/gkr944
458 Groszmann M, Bylstra Y, Lampugnani ER, and Smyth DR. 2010. Regulation of tissue-specific expression of SPATULA,
459   a bHLH gene involved in carpel development, seedling germination, and lateral organ growth in
460   Arabidopsis. *J Exp Bot* 61:1495-1508. 10.1093/jxb/erq015
461 Heisler MG, Atkinson A, Bylstra YH, Walsh R, and Smyth DR. 2001. SPATULA, a gene that controls development of
462   carpel margin tissues in Arabidopsis, encodes a bHLH protein. *Development* 128:1089-1098.
463 Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, and Nakai K. 2007. WoLF PSORT: protein
464   localization predictor. *Nucleic Acids Res* 35:W585-587. 10.1093/nar/gkm259
465 Hu B, Jin J, Guo AY, Zhang H, Luo J, and Gao G. 2015. GSDS 2.0: an upgraded gene feature visualization server.
466   *Bioinformatics* 31:1296-1297. 10.1093/bioinformatics/btu817
467 Ichihashi Y, Horiguchi G, Gleissberg S, and Tsukaya H. 2010. The bHLH transcription factor SPATULA controls final
468   leaf size in Arabidopsis thaliana. *Plant Cell Physiol* 51:252-261. 10.1093/pcp/pcp184
469 Jiang Y, Yang B, and Deyholos MK. 2009. Functional characterization of the Arabidopsis bHLH92 transcription factor
470   in abiotic stress. *Mol Genet Genomics* 282:503-516. 10.1007/s00438-009-0481-3
471 Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, and Gao G. 2017. PlantTFDB 4.0: toward a central hub for
472   transcription factors and regulatory interactions in plants. *Nucleic Acids Res* 45:D1040-D1045.
473   10.1093/nar/gkw982
474 Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S,
475   Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, and Hunter S. 2014. InterProScan 5:
476   genome-scale protein function classification. *Bioinformatics* 30:1236-1240.
477   10.1093/bioinformatics/btu031
478 Ko SS, Li MJ, Sun-Ben Ku M, Ho YC, Lin YJ, Chuang MH, Hsing HX, Lien YC, Yang HT, Chang HC, and Chan MT. 2014.
479   The bHLH142 Transcription Factor Coordinates with TDR1 to Modulate the Expression of EAT1 and
480   Regulate Pollen Development in Rice. *Plant Cell* 26:2486-2504. 10.1105/tpc.114.126292
481 Koini MA, Alvey L, Allen T, Tilley CA, Harberd NP, Whitelam GC, and Franklin KA. 2009. High temperature-mediated
482   adaptations in plant architecture require the bHLH transcription factor PIF4. *Curr Biol* 19:408-413.
483   10.1016/j.cub.2009.01.046
484 Komatsu K, Maekawa M, Ujiie S, Satake Y, Furutani I, Okamoto H, Shimamoto K, and Kyozuka J. 2003. LAX and SPA:
485   major regulators of shoot branching in rice. *Proc Natl Acad Sci U S A* 100:11765-11770.
486   10.1073/pnas.1932414100
487 Kondou Y, Nakazawa M, Kawashima M, Ichikawa T, Yoshizumi T, Suzuki K, Ishikawa A, Koshi T, Matsui R, Muto S,
488   and Matsui M. 2008. RETARDED GROWTH OF EMBRYO1, a new basic helix-loop-helix protein, expresses in
489   endosperm to control embryo growth. *Plant Physiol* 147:1924-1935. 10.1104/pp.108.118364
490 Kumari M, Taylor GJ, and Deyholos MK. 2008. Transcriptomic responses to aluminum stress in roots of Arabidopsis

491   thaliana. *Mol Genet Genomics* 279:339-357. 10.1007/s00438-007-0316-z

492   Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-

493   Hernandez M, Karthikeyan AS, Lee CH, Nelson WD, Ploetz L, Singh S, Wensel A, and Huala E. 2012. The

494   Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res*

495   40:D1202-1210. 10.1093/nar/gkr1090

496   Lefort V, Longueville J-E, and Gascuel O. 2017. SMS: Smart Model Selection in PhyML. *Mol Biol Evol* 34:2422-2424.

497   10.1093/molbev/msx149

498   Letunic I, and Bork P. 2018. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res*

499   46:D493-D496. 10.1093/nar/gkx922

500   Letunic I, Doerks T, and Bork P. 2015. SMART: recent updates, new developments and status in 2015. *Nucleic Acids*

501   *Res* 43:D257-260. 10.1093/nar/gku949

502   Li B, and Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference

503   genome. *Bmc Bioinformatics* 12:323.

504   Li X, Duan X, Jiang H, Sun Y, Tang Y, Yuan Z, Guo J, Liang W, Chen L, Yin J, Ma H, Wang J, and Zhang D. 2006.

505   Genome-wide analysis of basic/helix-loop-helix transcription factor family in rice and Arabidopsis. *Plant*

506   *Physiol* 141:1167-1184. 10.1104/pp.106.080580

507   Lin Q, Ohashi Y, Kato M, Tsuge T, Gu H, Qu L-J, and Aoyama T. 2015. GLABRA2 directly suppresses basic helix-loop-

508   helix transcription factor genes with diverse functions in root hair development. *The Plant Cell* 27:2894-

509   2906.

510   Ling HQ, Bauer P, Bereczky Z, Keller B, and Ganal M. 2002. The tomato fer gene encoding a bHLH protein controls

511   iron-uptake responses in roots. *Proc Natl Acad Sci U S A* 99:13938-13943. 10.1073/pnas.212448699

512   Long TA, Tsukagoshi H, Busch W, Lahner B, Salt DE, and Benfey PN. 2010. The bHLH transcription factor POPEYE

513   regulates response to iron deficiency in Arabidopsis roots. *Plant Cell* 22:2219-2236.

514   10.1105/tpc.110.074096

515   Malinovsky FG, Batoux M, Schwessinger B, Youn JH, Stransfeld L, Win J, Kim S-K, and Zipfel C. 2014a. Antagonistic

516   regulation of growth and immunity by the Arabidopsis basic helix-loop-helix transcription factor homolog

517   of brassinosteroid enhanced expression2 interacting with increased leaf inclination1 binding bHLH1. *Plant*

518   *Physiol* 164:1443-1455.

519   Malinovsky FG, Batoux M, Schwessinger B, Youn JH, Stransfeld L, Win J, Kim SK, and Zipfel C. 2014b. Antagonistic

520   regulation of growth and immunity by the Arabidopsis basic helix-loop-helix transcription factor homolog

521   of brassinosteroid enhanced expression2 interacting with increased leaf inclination1 binding bHLH1. *Plant*

522   *Physiol* 164:1443-1455. 10.1104/pp.113.234625

523   Mart, iacute, nez G, iacute, and a JF. 2000. Direct Targeting of Light Signals to a Promoter Element-Bound

524   Transcription Factor. *Science* 288:859-863. 10.1126/science.288.5467.859

525   Murre C, McCaw PS, and Baltimore D. 1989. A new DNA binding and dimerization motif in immunoglobulin

526   enhancer binding, daughterless, MyoD, and myc proteins. *Cell* 56:777-783. 10.1016/0092-8674(89)90682-

527   x

528   Oh E, Kim J, Park E, Kim JI, Kang C, and Choi G. 2004. PIL5, a phytochrome-interacting basic helix-loop-helix protein,

529   is a key negative regulator of seed germination in Arabidopsis thaliana. *Plant Cell* 16:3045-3058.

530   10.1105/tpc.104.025163

531   Regier N, and Frey B. 2010. Experimental comparison of relative RT-qPCR quantification approaches for gene

532   expression studies in poplar. *BMC Mol Biol* 11:57. 10.1186/1471-2199-11-57

533 Schneider TD, and Stephens RM. 1990. Sequence logos: a new way to display consensus sequences. *Nucleic Acids*

534   *Res* 18:6097-6100. 10.1093/nar/18.20.6097

535 Seo JS, Joo J, Kim MJ, Kim YK, Nahm BH, Song SI, Cheong JJ, Lee JS, Kim JK, and Choi YD. 2011. OsbHLH148, a basic

536   helix-loop-helix protein, interacts with OsJAZ proteins in a jasmonate signaling pathway leading to

537   drought tolerance in rice. *Plant J* 65:907-921. 10.1111/j.1365-313X.2010.04477.x

538 Sharma N, Xin R, Kim D-H, Sung S, Lange T, and Huq E. 2016. NO FLOWERING IN SHORT DAY (NFL) is a bHLH

539   transcription factor that promotes flowering specifically under short-day conditions in Arabidopsis.

540   *Development* 143:682-690.

541 Sjodin A, Street NR, Sandberg G, Gustafsson P, and Jansson S. 2009. The Populus Genome Integrative Explorer

542   (PopGenIE): a new resource for exploring the Populus genome. *New Phytol* 182:1013-1025.

543   10.1111/j.1469-8137.2009.02807.x

544 Takahashi Y, Ebisu Y, Kinoshita T, Doi M, Okuma E, Murata Y, and Shimazaki K. 2013. bHLH transcription factors

545   that facilitate K(+) uptake during stomatal opening are repressed by abscisic acid through phosphorylation.

546   *Sci Signal* 6:ra48. 10.1126/scisignal.2003760

547 Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. 2011. MEGA5: molecular evolutionary genetics

548   analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol*

549   *Evol* 28:2731-2739. 10.1093/molbev/msr121

550 Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, and Higgins DG. 1997. The CLUSTAL_X windows interface:

551   flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*

552   25:4876-4882.

553 Toledo-Ortiz G, Huq E, and Quail PH. 2003. The Arabidopsis Basic/Helix-Loop-Helix Transcription Factor Family. *The*

554   *Plant Cell* 15:1749-1770. 10.1105/tpc.013839

555 Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, and Salamov

556   A. 2006. The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). *Science* 313:1596-1604.

557 Varaud E, Brioudes F, Szecsi J, Leroux J, Brown S, Perrot-Rechenmann C, and Bendahmane M. 2011. AUXIN

558   RESPONSE FACTOR8 regulates Arabidopsis petal growth by interacting with the bHLH transcription factor

559   BIGPETALp. *Plant Cell* 23:973-983. 10.1105/tpc.110.081653

560 Wang YJ, Zhang ZG, He XJ, Zhou HL, Wen YX, Dai JX, Zhang JS, and Chen SY. 2003. A rice transcription factor

561   OsbHLH1 is involved in cold stress response. *Theor Appl Genet* 107:1402-1409. 10.1007/s00122-003-1378-

562   x

563 Wang Z, Liu Q, Wang H, Zhang H, Xu X, Li C, and Yang C. 2016. Comprehensive analysis of trihelix genes and their

564   expression under biotic and abiotic stresses inPopulus trichocarpa. *Scientific Reports* 6:36274.

565 Yao W, Wang S, Zhou B, and Jiang T. 2016. Transgenic poplar overexpressing the endogenous transcription factor

566   ERF76 gene improves salinity tolerance. *Tree Physiol* 36:896-908. 10.1093/treephys/tpw004

567 Yi K, Wu Z, Zhou J, Du L, Guo L, Wu Y, and Wu P. 2005. OsPTF1, a novel transcription factor involved in tolerance to

568   phosphate starvation in rice. *Plant Physiol* 138:2087-2096. 10.1104/pp.105.063115

569 Zhang J, Liu B, Li M, Feng D, Jin H, Wang P, Liu J, Xiong F, Wang J, and Wang HB. 2015. The bHLH transcription

570   factor bHLH104 interacts with IAA-LEUCINE RESISTANT3 and modulates iron homeostasis in Arabidopsis.

571   *Plant Cell* 27:787-805. 10.1105/tpc.114.132704

572 Zhou J, Li F, Wang JL, Ma Y, Chong K, and Xu YY. 2009. Basic helix-loop-helix transcription factor from wild rice

573   (OrbHLH2) improves tolerance to salt- and osmotic stress in Arabidopsis. *J Plant Physiol* 166:1296-1306.

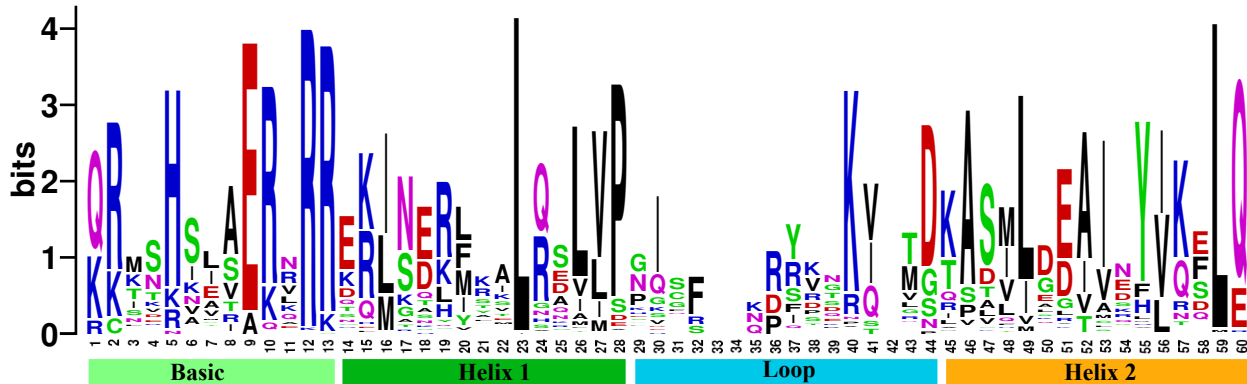
574   10.1016/j.jplph.2009.02.007

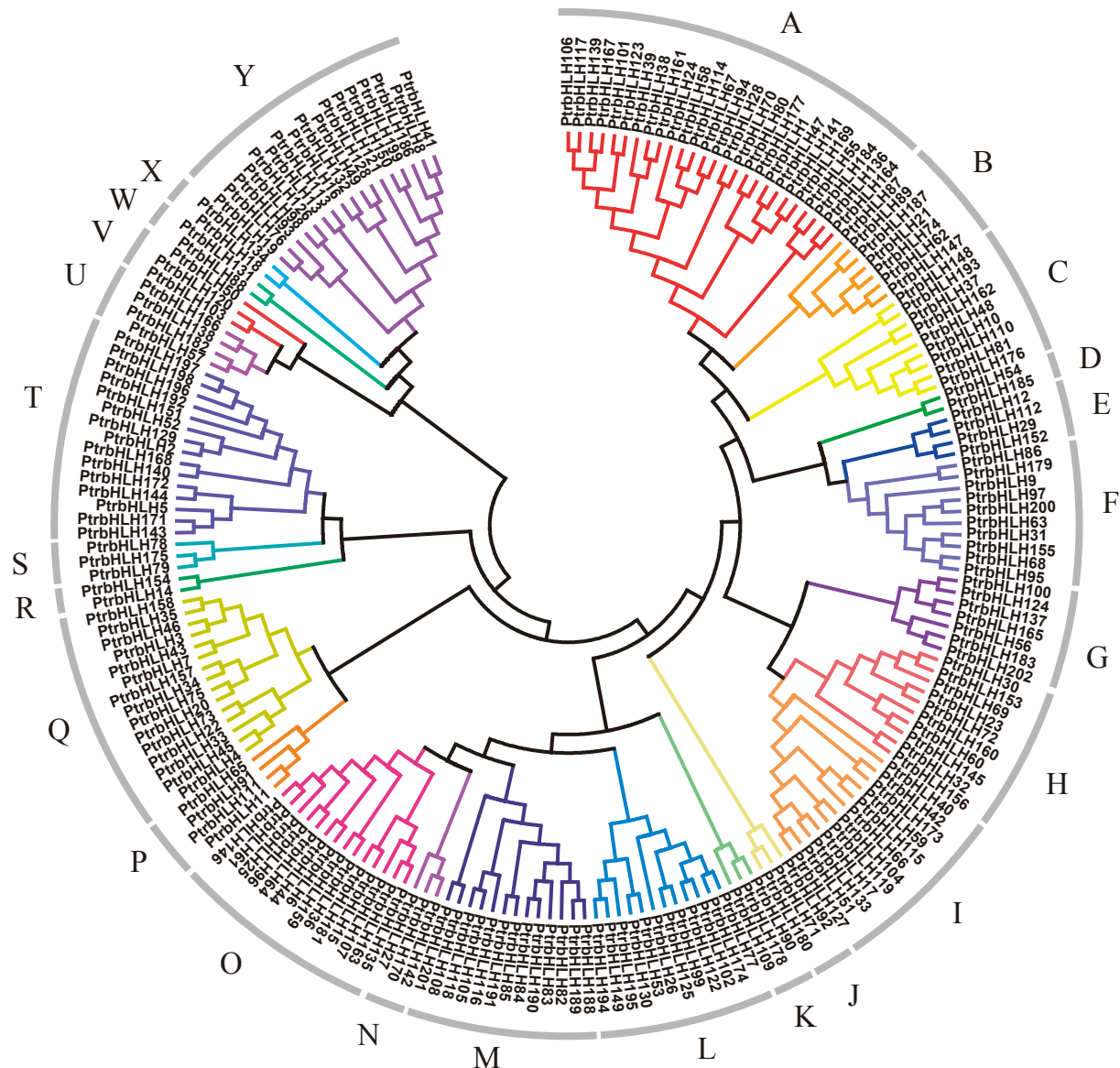
575   Zhu Y, Tepperman JM, Fairchild CD, and Quail PH. 2000. Phytochrome B binds with greater apparent affinity than

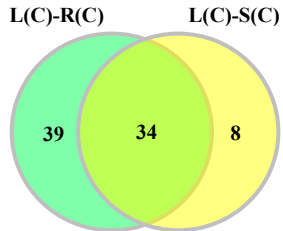
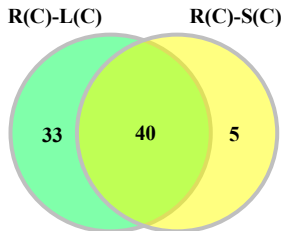
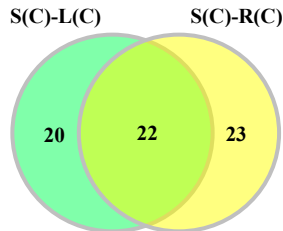
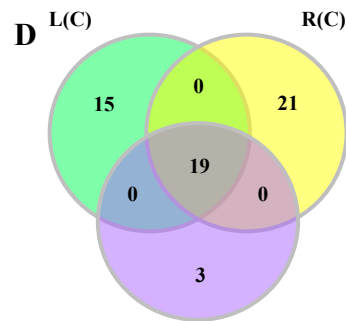
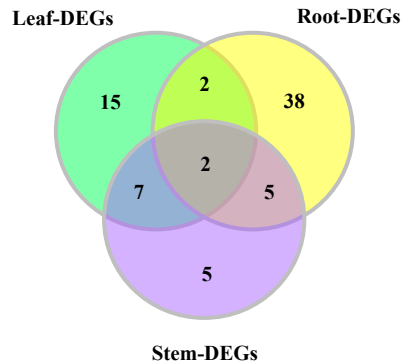
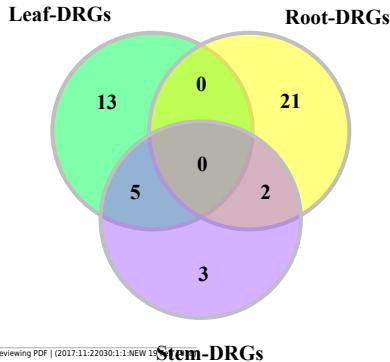
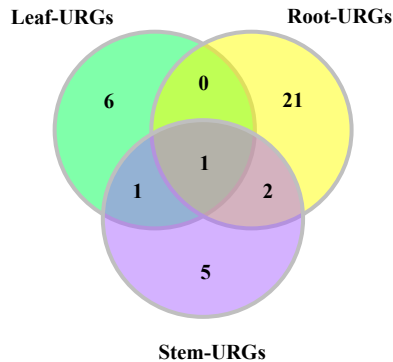
576   phytochrome A to the basic helix-loop-helix factor PIF3 in a reaction requiring the PAS domain of PIF3.

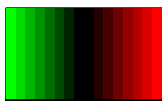
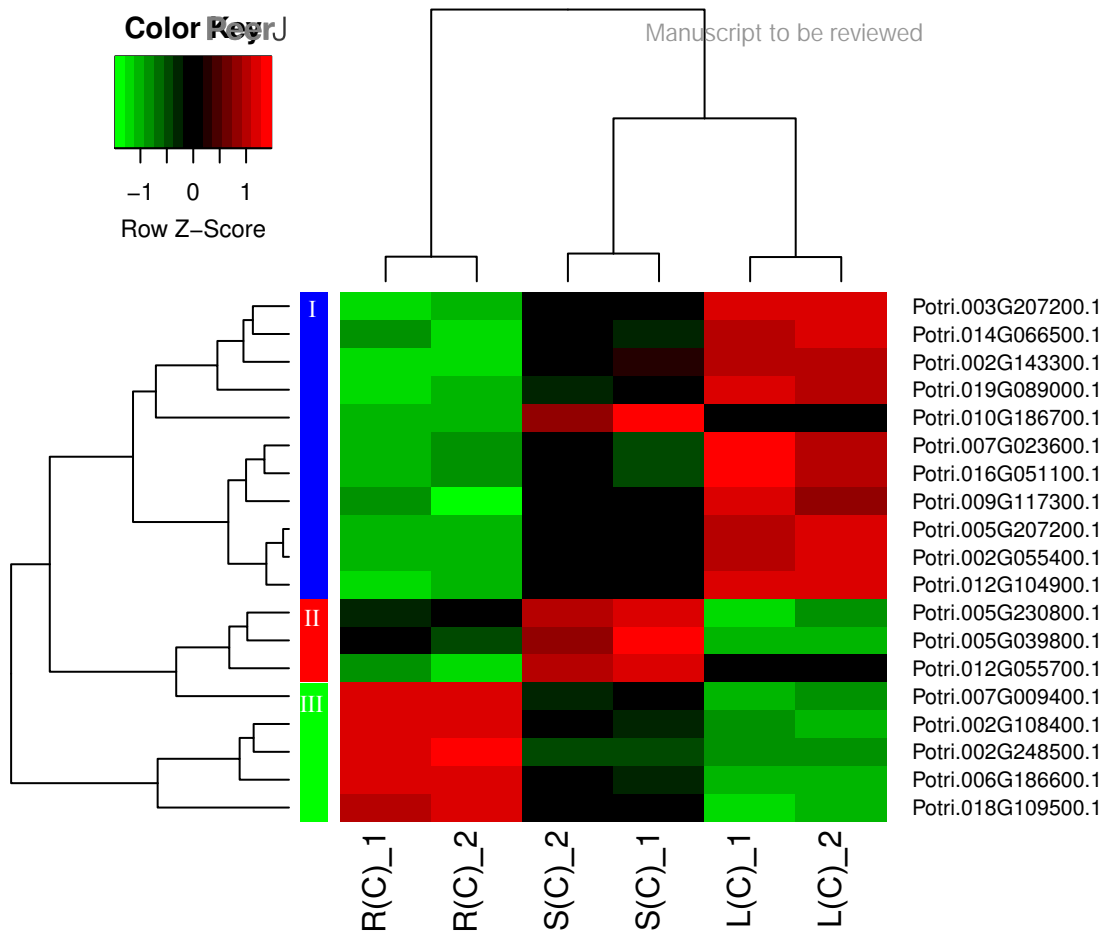
577   *Proc Natl Acad Sci U S A* 97:13419-13424. 10.1073/pnas.230433797

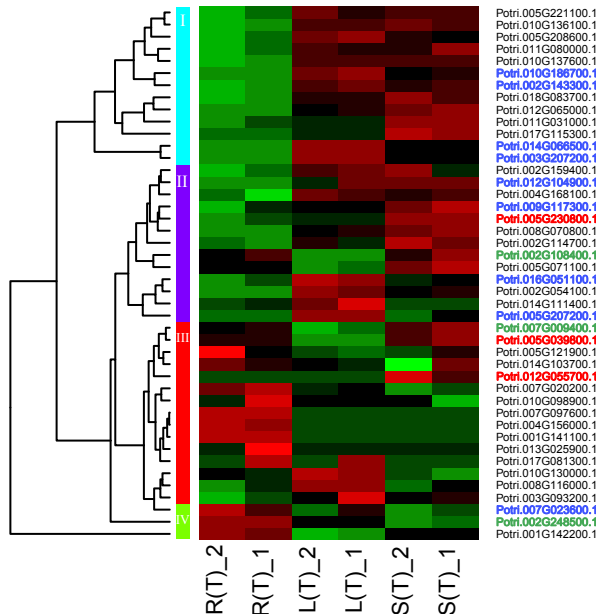
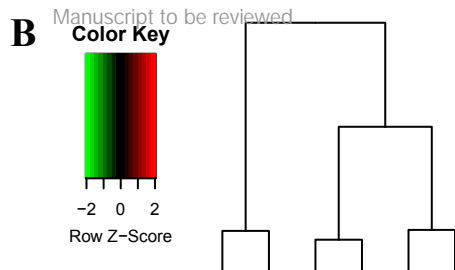
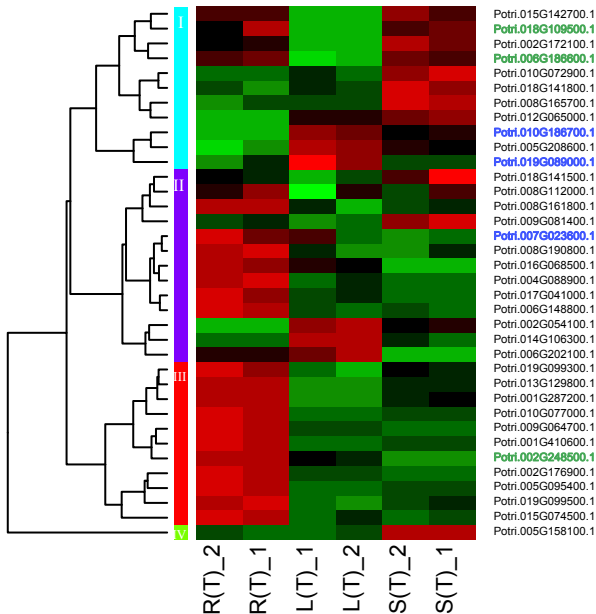
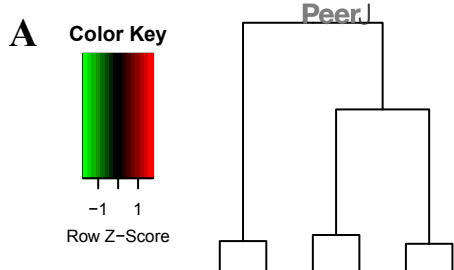
578

The DNA binding domain alignment logo of poplar bHLH family.

Dendrogram of bHLH gene family proteins.

Each color represents a special group.

Venn diagrams of tissue-differential bHLH genes and DEGs challenged with salinity.

(A-C) Number of genes displaying distinct and shared expression in tissue-pairs without salt treatment. The shared genes express differently in two tissue-pairs. (A) Comparison between leaf-root and leaf-stem pairs. (B) Comparison between root-leaf and root-stem pairs. (C) Comparison between stem-leaf and stem-root pairs. (D) Comparison between the three shared parts. We extracted the shared genes from each of the comparisons mentioned above, and compared them. (E-G) Number of DEGs, DRGs, or URGs in response to salt stress in each tissue.

A heatmap of genes displaying tissue-differential expression without treatment.

We focused on the 19 genes that have tissue-differential expression. The gene expression values are square-root transformed FPKM (fragments per kilo-bases per million mapped reads). We used the Z-score as standardization method for each gene. We used dist (Distance Matrix Computation) function with euclidean method to compute and return the distance matrix and hclust (Hierarchical Clustering) function with complete method to perform a hierarchical cluster analysis. The colorful vertical bars on the left side denote gene cluster 1-3 orderly.

Two heatmaps of URGs (A) and DRGs (B).

The gene expression values are square-root transformed FPKM (fragments per kilo-bases per million mapped reads). We used the Z-score as standardization method for each gene. We used dist (Distance Matrix Computation) function with euclidean method to compute and return the distance matrix and hclust (Hierarchical Clustering) function with complete method to perform a hierarchical cluster analysis. The colorful vertical bars on the left side denote gene cluster 1-4 orderly. The genes, which are highlighted in blue, red, and green, belong to the cluster1-3 in Fig. 4 orderly.

Barplots of DEGs relative expression levels based on RT-qPCR and RNA-Seq.

The expression levels of each gene were calculated as relative to corresponding gene expression in leaf with no treatment. Error bars represent SD of biologic replicates.