Analysis of bHLH genes from foxtail millet (*Setaria italica*) and their potential relevance to drought stress

Pengfei Wang[1‡]*, **Haili Wang**[2,3‡], **Yongmei Wang**[1], **Fengshan Ren**[1]*, **Wei Liu**[2]*

1 Shandong Academy of Grape, Jinan, PR China, **2** Biotechnology Research Center, Shandong Academy of Agricultural Sciences, Jinan, PR China, **3** Shandong Academy of Pesticide Sciences, Jinan, PR China

‡ Co-first authors
* fengqiaoyouzi@126.com (PW); rensd65@163.com (FR); wheiliu@163.com (WL)

Abstract

Foxtail millet is very a drought-tolerant crop. Basic helix–loop–helix (bHLH) transcription factors are involved in many drought-stress responses, but foxtail millet bHLH genes have been scarcely examined. We identified 149 foxtail millet bHLH genes in a genome-wide analysis and performed Swiss-Prot, GO, and KEGG pathway analyses for these genes. Phylogenetic analyses placed the genes into 25 clades, with some remaining orphans. We identified homologs based on gene trees and Swiss-Prot annotation. We also inferred that some homologs underwent positive selection in foxtail millet ancestors, and selected motifs differed among homologs. Expression of eight foxtail millet bHLH genes varied with drought stress. One of these genes was localized to a QTL that contributes to drought tolerance in foxtail millet. We also perform a *cis*-acting regulatory element analysis on foxtail millet bHLH genes and some drought-induced genes. Foxtail millet bHLH genes were inferred to have a possible key role in drought tolerance. This study clarifies both the function of foxtail millet bHLH genes and drought tolerance in foxtail millet.

Introduction

The basic helix–loop–helix (bHLH) transcription factor family is a large gene super-family found in plant and animal genomes [1], and its members play very key roles in a wide range of metabolic, physiological, and developmental processes [2–5]. bHLH family members have many different functions [6], and they each contain a core bHLH domain of approximately 60 amino acids, including a basic region (at the N-terminus) and a HLH region [7–8]. bHLH proteins can interact with each other and form homo-dimers or hetero-dimers that are promoted by the bHLH domains [1, 9]. As a core transcription factor domain, the bHLH domain is involved in DNA binding [10], with bHLH domains or bHLH proteins binding to E-box (5′-CANNTG-3′) and G-box (5′-CACGTG-3′) *cis* elements and regulating gene expression [4, 11].

Only a small number of plant bHLH transcription factors have been characterized functionally, far fewer than have been characterized in animals [6]. A previous study showed that

bHLH transcription factors can act as transcriptional activators or repressors and are involved in the regulation of fruit dehiscence, anther and epidermal cell development, hormone signalling, and other similar processes in plants [12]. The plant bHLH protein PIF3 is a direct phytochrome reaction partner in the photoreceptor's signalling network [4] and is involved in controlling the expression of light-regulated genes [13]. Some bHLH transcription factors can interact with MYB transcription factors and WD40 or WDR proteins to form a MYB–bHLH–WD40 (MBW) complex or MYB–bHLH–WDR (MBW) complexes, which can activate anthocyanin biosynthesis genes, resulting in anthocyanin pigment accumulation and fiber development in plants [14–17].

Some functions of unknown bHLH transcription factors as well as some new functions of known bHLH transcription factors have been gradually identified in different plant species. In the medicinal plant *Catharanthus roseus*, the bHLH transcription factor BIS2 is essential for monoterpenoid indole alkaloid production [18]. In *Salvia miltiorrhiza*, bHLH transcription factors are related to tanshinone biosynthesis [19]. *Arabidopsis* bHLH129 appears to regulate root elongation [20], while *Arabidopsis* bHLH109 is associated with somatic embryo induction [21]. Additionally, the *Arabidopsis* bHLH transcription factor PIF4 plays a major role in integrating multiple signals to regulate growth [22]. Research has shown that grasses can use an alternatively wired bHLH transcription factor network to establish stomatal identity [23], further enriching our understanding of plant bHLH transcription factors.

Some plant bHLH transcription factors have also been recently reported to be related to responses to abiotic stresses such as drought and cold. For example, Feng et al. recently found that a novel tomato bHLH transcription factor, SlICE1a, could confer cold, osmotic-stress, and salt tolerance to plants [24]. Similarly, *Eleusine coracana* bHLH57 transcription factors are related to tolerance of drought, salt, and oxidative stresses [25]. bHLH122 plays an important role in drought and osmotic-stress resistance in *Arabidopsis* [26], where it regulates the expression of genes involved in abiotic stress tolerance [27]. In sheep grass (*Leymus chinensis*), many bHLH transcription factor family members were identified via RNA-seq to be responsive to drought stress [28]. Drought stress could affect plant growth, agricultural yields, and survival. Plants have evolved highly complex reactions to drought stress, and many genes are involved in drought stress [29, 30]. Plant bHLH genes are likely very important in responses to drought stress [29, 30]. Foxtail millet has been proposed as a new model organism for functional genomics studies of the Panicoideae and has the potential to become a new model organism for the study of drought stress responses because of its outstanding tolerance to drought stress [29–31].

We identified the foxtail millet bHLH transcription factors in a genome-wide survey and studied the expression of bHLH genes in foxtail millet in various tissues under drought stress conditions. Our purpose was to identify foxtail millet bHLH transcription factor family members, find candidates that may be relevant to drought stress, and improve the current understanding of drought tolerance mechanisms in foxtail millet.

Material and methods

Data collection and identification of bHLH genes

Whole genome sequences of foxtail millet (*Setaria italica*) were obtained from the 2012 Foxtail Millet Database (http://foxtailmillet.genomics.org.cn/page/species/index.jsp) [32]. The bHLH domain is conserved within bHLH proteins, and the HMM ID of the bHLH domain is (PF00010) in the pfam database (http://pfam.xfam.org/). The amino acid sequences of HMMs were used as queries to identify all possible candidate bHLH protein sequences in the foxtail millet genome database using BLASTP ($E < 0.001$). SMART online software (http://smart.

embl-heidelberg.de/) was used to identify integrated bHLH domains in putative foxtail millet bHLH proteins. Candidate proteins without integrated bHLH domains were discarded.

Swiss-Prot, GO and KEGG pathway annotation

We performed Swiss-Prot function annotation analysis based on the UniProtKB/Swiss-Prot database (http://www.uniprot.org/), GO function annotation analysis based on the GO database (http://geneontology.org/page/go-database), and KEGG pathway annotation analysis based on the KEGG database (http://www.kegg.jp/kegg/ko.html).

Phylogenetic analysis

We aligned the foxtail millet bHLH protein sequences using Clustal Omega online software (http://www.ebi.ac.uk/Tools/msa/clustalo) and constructed neighbor-joining (NJ) trees using MEGA 6.0 with the aligned foxtail millet bHLH protein sequences. Support for inferred evolutionary relationships was calculated from 1000 bootstrap samples [33].

Conserved motif analysis of foxtail millet bHLH protein sequences

We conducted a conserved Motif Analysis of foxtail millet bHLH protein sequences using the Multiple Em for Motif Elicitation (MEME) suite 4.11.1 software (http://meme.nbcr.net/meme/) [34] with the following parameter settings: output motifs, 20; minimum motif width, 6; maximum motif width, 300 [31].

Selection pressure analysis

The codeml portion of the phylogenetic analysis maximum likelihood (PAML) program (version 4.7 software) [35] was used to infer potential selective pressures. A comparison of site models M0–M3 was used to determine which kinds of selective pressure the genes underwent, and a M7–M8 comparison was used to identify sites shaped by positive selection [33, 36].

Identification of foxtail millet bHLH genes within drought tolerance QTLs

QTLs for drought tolerance were identified from previous research by Qie et al [37], and the physical locations of foxtail millet bHLH genes were collected from the foxtail millet genome database (http://foxtailmillet.genomics.org.cn/page/species/index.jsp). The bHLH genes that overlapped with QTLs were inferred to be the genes located within each QTL.

Cis-acting regulatory element analysis

Plantcare (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) was used to analyse the *cis*-acting regulatory elements of bHLH genes [33].

Plant material, stress treatments and RNA isolation

To induce drought conditions, 14-day-old foxtail millet cv. 'Yugu1' shoots were grown under a 20% polyethylene glycol 6000 (PEG 6000) treatment [29] for 0, 0.5, 6, and 12 h; the 0-h treatment was the control (CK) treatment, while the other treatments simulated droughts of various lengths. The 14-day-old foxtail millet shoots were also grown under a 100 mm/L ABA treatment [38] for 0, 0.5, 6, and 12 h. RNA was isolated using the CTAB method, and we performed reverse transcription according to a previously described protocol [33].

Gene expression analysis

Quantitative RT-PCR (qRT-PCR) analysis was conducted as previously described [31]. Three replicates were carried out in this study and *t*-tests were used to analyze significance. The qRT-PCR primers are provided in S1 Table. A heat map was generated based on RPKM values using Multiexperiment View software. All the PRKM values or RNA-seq data were based on RNA data hosted by the foxtail millet genome database (http://foxtailmillet.genomics.org.cn/page/species/index.jsp) [32]. RPKM values less than 0.3 were considered unexpressed genes in this study [39].

Results and discussion

Identification, annotation, and phylogenetic analysis of foxtail millet bHLH genes

The amino acid sequences of bHLHs were extracted from the foxtail millet genome database (http://foxtailmillet.genomics.org.cn/page/species/index.jsp) using BLASTP with amino acid sequences of bHLH domains (Pfam: PF00010) as queries. We identified 149 bHLH family members distributed among all nine chromosomes. We assayed their annotated functions based on the UniProtKB/Swiss-Prot database (http://www.uniprot.org/). All of these bHLHs were annotated based on the best-hit proteins (S2 Table). In a previous study, the function of some bHLHs from *Arabidopsis* had also been reported [12]. Swiss-Prot functional annotation revealed that most homologs of these *Arabidopsis* bHLHs can be found in foxtail millet excluding some members, including NAI1 (ER body formation); RHD6 and RSL1 (root hair formation); LHW (root development); PRE1, PRE2, PRE3, PRE4, and PRE5 (gibberellin signalling transduction); KDR (light signal transduction); and some orphans.

We also found some functional annotations of bHLH genes that were not identified by a previous study of *Arabidopsis* bHLHs [12], including LAX_ORYSJ transcription factor LAX PANICLE, WIT1_ARATH WPP domain-interacting tail-anchored protein 1, AIB_ARATH transcription factor ABA-INDUCIBLE bHLH-TYPE, MGP_ARATH Zinc finger protein MAGPIE, PP425_ARATH Pentatricopeptide repeat-containing protein, BH032_ARATH transcription factor AIG1, and Anthocyanin regulatory Lc protein (S2 Table).

We used the full-length amino acid sequences of the 149 foxtail millet bHLHs for phylogenetic analysis, in which clades with relatively high bootstrap support ($\geq$50) were considered. The phylogenetic tree revealed 25 clades (clades 1–25) in the foxtail millet bHLH family and some orphans (Fig 1). The identified orphan genes were consistent with previous findings by Feller et al., as were the divisions of the clades [12].

The potential function of foxtail millet bHLH genes was studied using gene ontology (GO) annotation analysis, and these bHLH genes were classified into three categories based on GO annotation, cellular component, molecular function, and biological process. Cellular component contains four terms, including nucleus (GO, 0005634), nucleosome (GO, 0000786), intracellular (GO, 0005622), and ribosome (GO, 0005840). Molecular function contained five terms, including amino acid binding (GO, 0016597), DNA binding (GO, 0003677), zinc ion binding (GO, 0008270), structural constituent of ribosome (GO: 000373), and transcription regulator activity (GO, 0030528). Biological process contained three terms, including metabolic process (GO, 0008152), regulation transcription (GO, 0045449), and spermatogenesis (GO 0007283; S3 Table).

Nine foxtail millet homologs of the PIF subfamily members were also found based on Swiss-Prot annotation, and analysis of the KEGG pathway annotation showed that all foxtail millet homologs of PIF subfamily members could be mapped to a KEGG pathway.
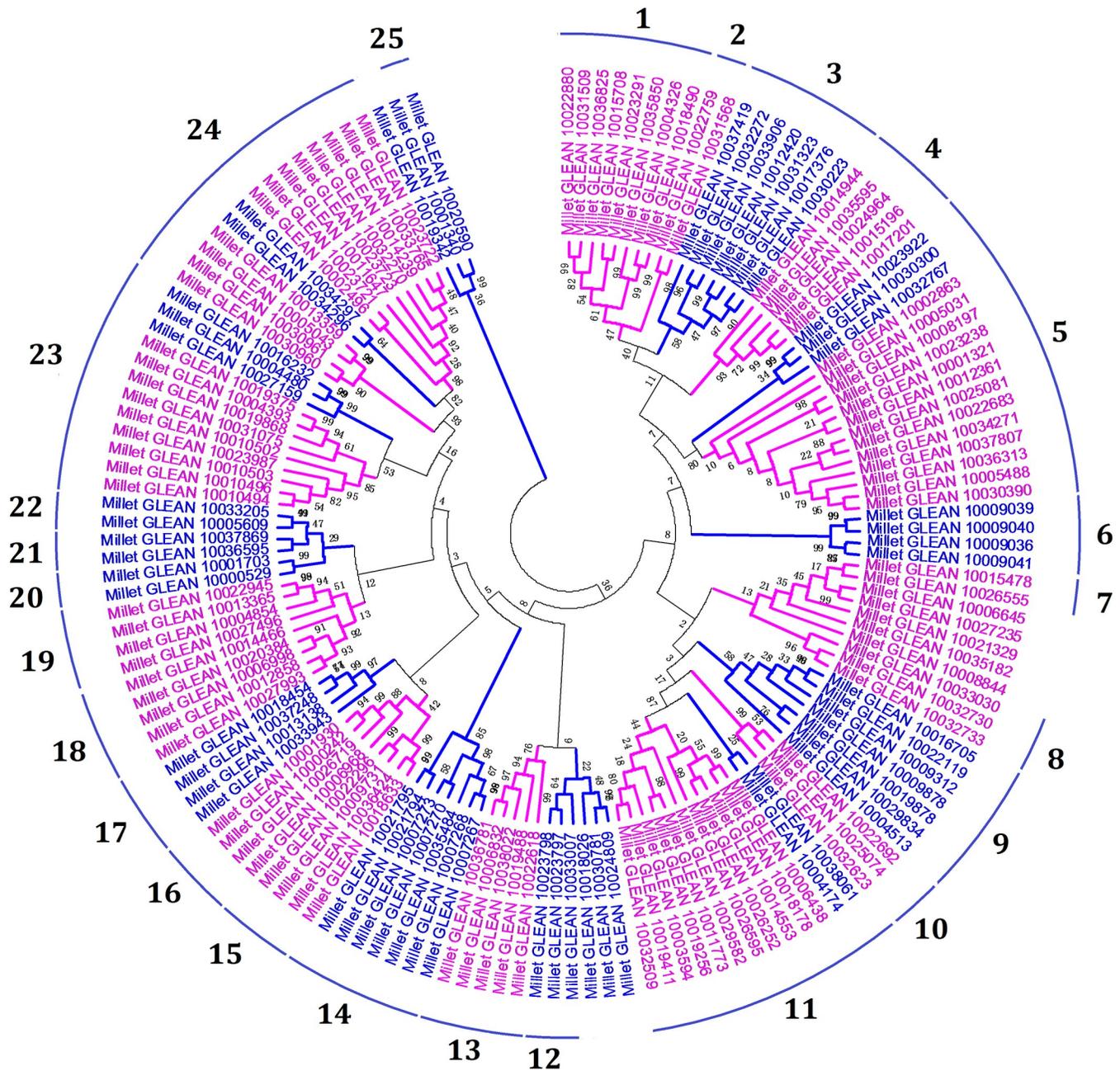
Fig 1. Neighbor-joining phylogenic analysis of bHLHs in foxtail millet. The numbers on the discontinuous cycle represents the names of clades (clades 1–25).

Millet_GLEAN_10012420 (Swiss-Prot ID, PIF1), Millet_GLEAN_10031323 (PIF5), Millet_-GLEAN_10017376 (PIF1), Millet_GLEAN_10030223 (PIF1), and Millet_GLEAN_10009039 (PIF1) were in clade 3 (Fig 1). These genes were mapped to the plant hormone signal transduction pathway (Ko04075) and their KEGG annotation is PIF4 (K16189; S1 Fig). Millet_-GLEAN_10031568 (Swiss-Prot ID, PIF3) proteins were in clade 2. These could be mapped onto the circadian rhythm-plant pathway (ko04712), and their KEGG annotation is PIF3 (K12126; S2 Fig). Millet_GLEAN_10009040 (Swiss-Prot ID, PIF1), Millet_GLEAN_10009036

(Swiss-Prot ID, PIF1), and Millet_GLEAN_10009041 (Swiss-Prot ID, PIF1) were in clade 6. These genes were mapped to the plant hormone signal transduction (Ko04075) and circadian rhythm-plant (ko04712) pathways, with corresponding KEGG annotations of PIF4 (K16189) and PIF3 (K12126).

Selection pressure and motif analysis of foxtail millet bHLH genes

The bHLH genes that were placed into the same clades and had the same annotation categories were considered homologs. We wanted to know if the genes in one homologous group are functionally redundant or functionally divergent, so we performed a selection pressure analysis of some homologous groups. Molecular signatures of selection were categorized as purifying, positive, and neutral. The d_n/d_s value (ω) can provide a measurement for changes in selective pressures. Values of ω that are equal to, less than, or greater than one indicate neutral evolution, purifying selection, or positive selection on the target genes, respectively [40].

Purifying selection may generate genes with conserved functions or pseudogenization, while neofunctionalization or subfunctionalization is less likely [41]. Signatures of positive selection may indicate adaptive evolution, gene function losses, and pseudogenization [42–44]. Positive selection can also lead to new functions of genes [40, 45–46].

Some homologous groups underwent positive selection, such as the ICE1 group (including Millet_GLEAN_10036424 and Millet_GLEAN_10018633; M0 model, ω (d_n/d_s) = 1.51). The M8vsM7 model showed there were many positive selection sites in the ICE group (BEB analysis, LTR, $P < 0.1$). There was also positive selection in the MYC2 group (including Millet_GLEAN_1001 0494, Millet_GLEAN_10010503, Millet_GLEAN_10023987, and Millet_GLEAN_10031075; M0 model, ω = 1.85), as shown by the M8 versus M7 comparison (BEB analysis, LTR, $P < 0.1$). As such, we tentatively suggested functional divergence may have occurred in some homologous groups. Moreover, some homologous groups underwent purifying selection, such as the UNE (ω = 0.21), bHLH82 (ω = 0.45) and bHLH35 (ω = 0.26) groups (Table 1).

Motif divergence was observed in many homologous groups, such as bHLH82, FIT, bHLH35, BIM2, bHLH51, and myc2 groups. Different motifs may indicate different functions or functional divergence [31]. Motifs of some homologs were in agreement, such as the ILR3, bHLH30, and UNE12 groups (Fig 2). We analyzed the PI, grand average of hydropathicity, instability index, nuclear localization signals, and transmembrane domains of some homologous groups and found functional divergence may also exist in the homologs containing the same motifs. For example, some bHLH30 and ILR3 members contained nuclear localization signals but some did not (S4 Table).

Previous research has indicated that some bHLH genes are duplicated [2, 6]. Duplicated genes are the raw material for the evolution of new biological functions and thus play crucial roles in adaption [47].

Expression profile of foxtail millet bHLH genes

The expression profiles of each identified foxtail millet bHLH gene were analysed among several tissues: root, leaf, stem, and spica. The expression levels of foxtail millet bHLH genes in the four tissues based on the previous RNA-seq data (http://foxtailmillet.genomics.org.cn/page/species/index.jsp) and the expression level were captured as RPKM values. Most of these genes were expressed in at least one tissue, and only 20 genes (14.7%) were not expressed in the other three tissues (Fig 3A and S5 Table). According to RPKM values, Millet_GLEAN_100 29834, Millet_GLEAN_10037807, Millet_GLEAN_10010494, Millet_GLEAN_10006 968, Millet_GLEAN_10018454, Millet_GLEAN_10022618, Millet_GLEAN_10016705, Millet_GLEAN_10001930, Millet_GLEAN_10005609, Millet_GLEAN_10019878,

Table 1. Selection pressure of partail homolougus group.

Gene ID	Group ID	ω (dN/dS) of group	Positive site
Millet_GLEAN_10015708	UNE12_ARATH Transcription factor UNE12	0.21398	18
Millet_GLEAN_10023291			
Millet_GLEAN_10035850	BH082_ARATH Transcription factor bHLH82	0.44577	9
Millet_GLEAN_10004326			
Millet_GLEAN_10018490			
Millet_GLEAN_10031509			
Millet_GLEAN_10010494	RAP1_ARATH Transcription factor MYC2	1.85238	35
Millet_GLEAN_10010503			
Millet_GLEAN_10023987			
Millet_GLEAN_10031075			
Millet_GLEAN_10007273	BH035_ARATH Transcription factor bHLH35	0.25858	34
Millet_GLEAN_10007270			
Millet_GLEAN_10035484			
Millet_GLEAN_10007268			
Millet_GLEAN_10007267			
Millet_GLEAN_10018178	BH112_ARATH Transcription factor bHLH112	1.50748	>100
Millet_GLEAN_10026252			
Millet_GLEAN_10036424	ICE1_ARATH Transcription factor ICE1	1.16554	>100
Millet_GLEAN_10018633			

https://doi.org/10.1371/journal.pone.0207344.t001

Millet_GLEAN_10023721, Millet_GLEAN_10023987, and Millet_GLEAN_10027159 were only expressed in spica tissue and not expressed in the other three tissues. Millet_GLEAN_10006645, Millet_GLEAN_10014239, Millet_GLEAN_10021795, Millet_GLEAN_10023722, Millet_GLEAN_10023723, Millet_GLEAN_10000529, Millet_GLEAN_10021329, Millet_-GLEAN_10033765, and Millet_GLEAN_10034296 were only expressed in root tissue and not expressed in the other three tissues. In contrast, just one gene, Millet_GLEAN_10010503, was only expressed in leaf tissue and not expressed in the other three tissues. No gene was only expressed in stem tissue and not expressed in the other three tissues. The expression of Millet_GLEAN_1002038 (Swiss-Prot ID, ILR3_ARATH transcription factor ILR3) was highest wherever it was expressed in leaf, stem, spica, and root. Its homologs are involved in metal homeostasis, auxin-conjugate metabolism, and salicylic-dependent defence signalling responses in plants [12, 48–49].

In total, 116 foxtail millet bHLH genes were only expressed in root tissue, 77 were only expressed in leaf tissue, 115 were only expressed in spica tissue, and 72 were only expressed in stem tissue. Just 61 genes were expressed in all four tissues (Fig 3B). In contrast, 72 foxtail millet bHLH genes were not expressed in leaf tissue, 33 genes were not expressed in root tissue, 34 genes were not expressed in spica tissue, and 77 were not expressed in stem tissue (Fig 3B and S5 Table). This suggests that foxtail millet bHLH genes are biased towards expression in root and spica tissue.

Some foxtail millet bHLHs are related to drought stress

Because foxtail millet is remarkably tolerant to drought stress, it has substantial potential to become a new model organism for understanding this trait, which will become even more vital as climate change continues [29]. Previous studies have shown that some plant bHLH genes are involved in tolerance to drought stress [24]. To understand which foxtail millet bHLH members are involved in tolerance to drought stress, candidate genes that are related to
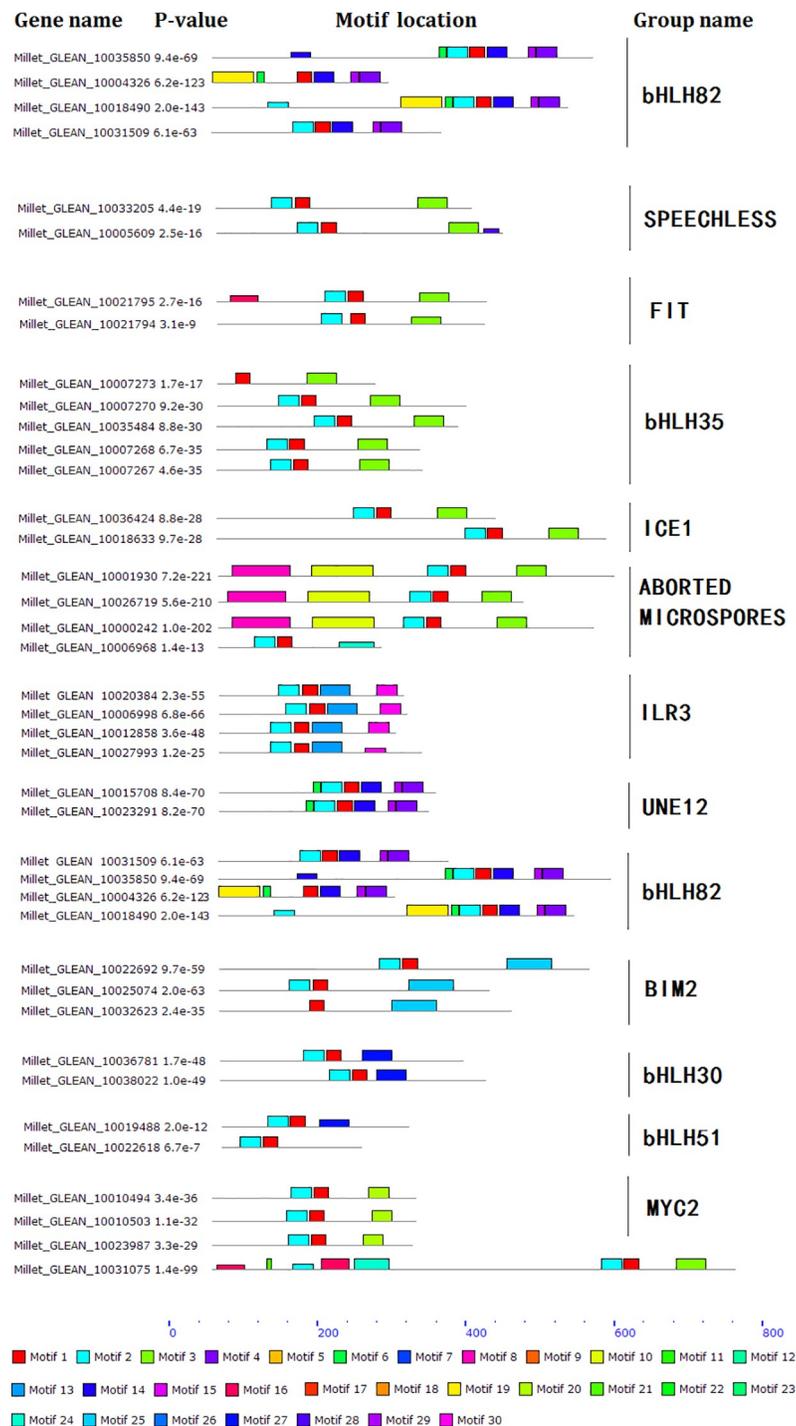
Fig 2. Motifs identified from some bHLH homologous groups in foxtail millet. The names of homologous groups were labeled on right and the name of each gene were labeled on left. Color boxes represent the motifs respectively.

drought stress tolerance are first required. In our study, we mainly focused on three kinds of foxtail millet bHLH genes: class A, genes located in QTLs that contribute to drought tolerance; class B, genes whose homologs in other plants were reported to be involved in drought tolerance; and class C, genes that respond to drought stress in foxtail millet.

A

B

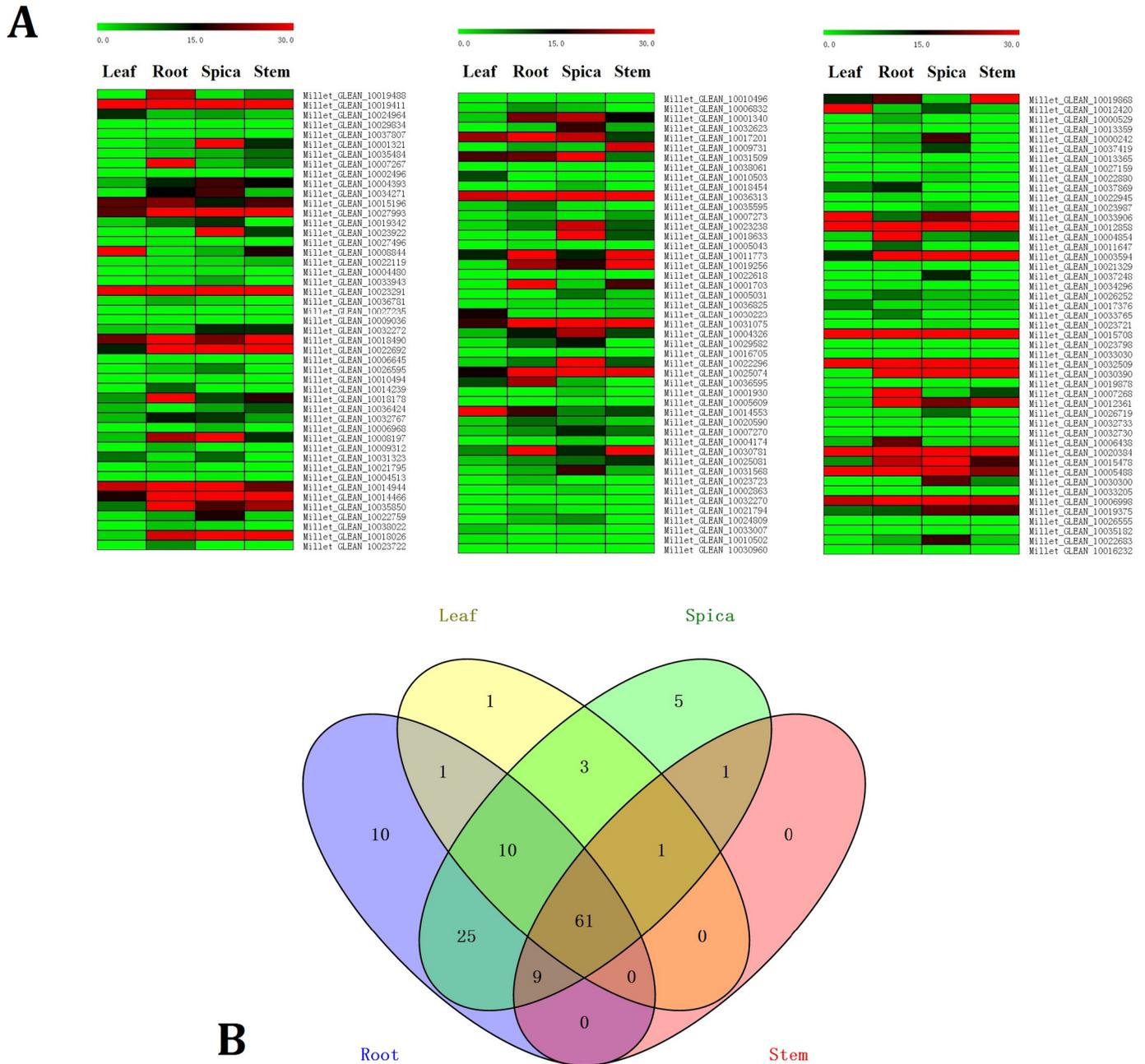**Fig 3. Expression profiles of bHLH genes in different tissues of foxtail millet.** A represents the heatmaps of bHLH genes expressions in leaf, root, spica and stem of foxtail millet. B represents The Venn diagram of expressed bHLH genes in four tissues.

Six QTLs (LOD > 2.5) for drought tolerance have been identified in foxtail millet, including QGSI_D_7A, QCLD_D_1A, QLRND_D_7A, QCLD_D_1B, QCLR_D_6A, and QSR_D_1A [37]. In this study, we determined that only the QTLs QGSI_D_7A (chr7, 33,221,000–27,196, 000), QCLD_D_1A (chr1, 29,834,000–32947000), and QLRND_D_7A (chr7, 30,571,000–21, 648,000) contain bHLH genes in foxtail millet. QTL QCLD_D_1A is related to coleoptile length decreases in foxtail millet [37]. This QTL was estimated to contribute 7% of the observed phenotypic variance [37], and it also contained the bHLH genes Millet_GLEAN_10023797

(Swiss-Prot ID, bHLH95), Millet_GLEAN_10023798 (Swiss-Prot ID, bHLH95), and Millet_GLEAN_10035595 (Swiss-Prot ID, bHLH128). QTL QLRND_D_7A was related to a lateral root number decrease of foxtail millet [37]. It was estimated to contribute 10% of the observed phenotypic variance [37], and it contained the bHLH genes Millet_GLEA N _1 00 02496 (Swiss-Prot ID, bHLH25) and Millet_GLEAN_10029582 (Swiss-Prot ID, bHLH113). QTL QGSI_D_7A was related to the germination stress tolerance index [37]. It was estimated to contribute 14% of the phenotypic variance [37], and it contained the bHLH genes Millet_GLEAN_10016232 (Swiss-Prot ID, ARLC_MAIZE Anthocyanin regulatory Lc protein), Millet_GLEAN_10037248 (Swiss-Prot ID, bHLH91), and Millet_GLEAN_10002496 (Swiss-Prot ID, bHLH25, which is also contained by the QTL QLRND_D_7A) (Table 2).

Currently, some plant genes, including bHLH57 [25], bHLH 122 [26], bHLH112 [27], ABA-Inducible bHLH or MYC2 [43–44], and ICE [24, 50–51], have been reported to be involved in drought tolerance. In our study, we were unable to identify the homologs of bHLH122 and bHLH57 in foxtail millet, but we did find the homologs of bHLH112, ABA-Inducible bHLH or MYC2, and ICE based on the Swiss-Prot annotation, including Millet_GLEAN_10026252 and Millet_GLEAN_10018178 (Swiss-Prot ID, bHLH112).

By referring to previous RNA-seq data [29], we found eight foxtail millet bHLH genes were involved in the response to drought stress (i.e., the 20% PEG 6000 treatment). Most class A and B genes did not respond to drought conditions, except for Millet_GLEAN_10035595 (Swiss-Prot ID, bHLH128; the function of bHLH 128 is unknown). The other seven genes that did respond were Millet_GLEAN_10023721 (Swiss-Prot ID, bHLH25), Millet_GLEAN_10 007270 (Swiss-Prot ID, bHLH35), Millet_GLEAN_10008844 (Swiss-Prot ID, UNE10; UNE10 is involved in the fertilization process), Millet_GLEAN_10005488 (Swiss-Prot ID, factor bHL H49), Millet_GLEAN_10036595 (Swiss-Prot ID, ORG2; ORG2 is involved in Iron homeostasis), Millet_GLEAN_10007267 (Swiss-Prot ID, bHLH35), and Millet_GLEAN_10030390 (Swiss-Prot ID, bHLH49). Excluding Millet_GLEAN_10008844 and Millet_GLEAN_100 36595, the function of homologs from other plant species of the other six genes were unknown. Additionally, the identified function of homologs from other plant species, Millet_GLEAN_10 008844 and Millet_GLEAN_10036595, are not thought to be involved in tolerance to drought stress, as shown by a previous study [12].

We analyzed the expressions of the genes that respond to drought stress using qRT-PCR. The 14-day foxtail millet shoots were subjected to 20% PEG 6000 for 0, 0.5 h, 6 h and 12 h treatment. The expression levels of the eight genes treated for 6 h and 12 h under PEG were significantly changed, and they all showed similar variation trends with that of the Qi's RNA-seq data excluding Millet_GLEAN_10036595 (Fig 4A) [29].

Table 2. Foxtail millet bHLH genes in QTL for drought tolerance.

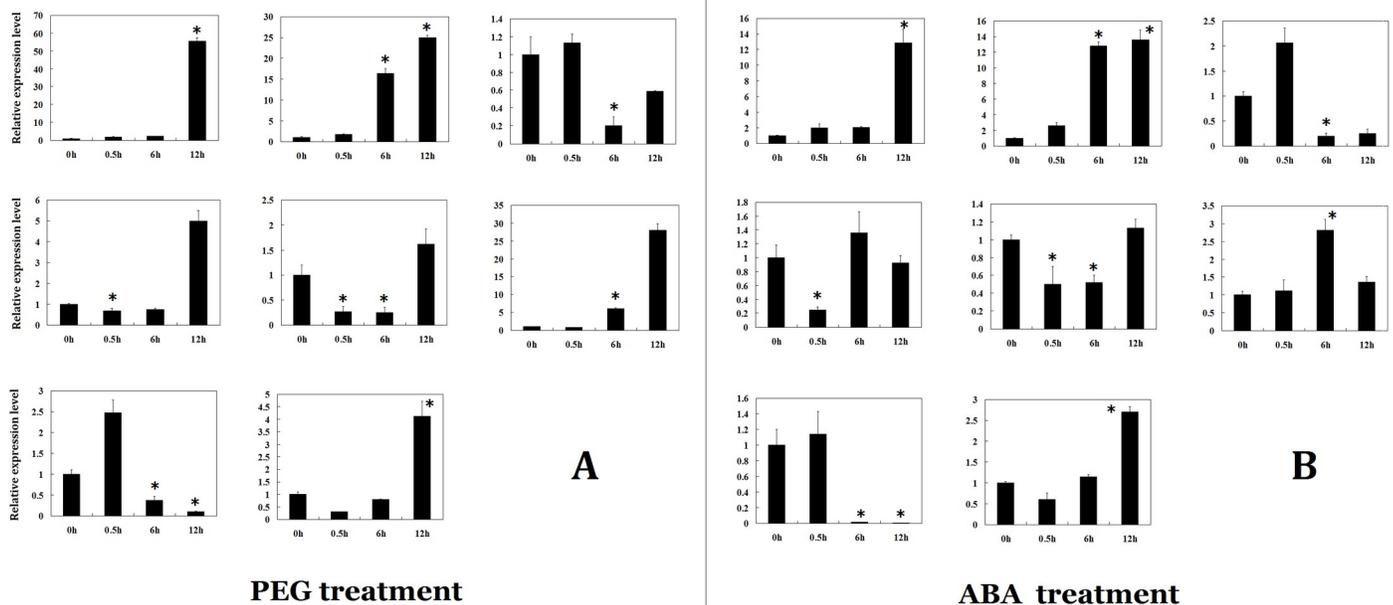| QTL name	Chr.	Marker interval	Percentage of the variance explained	bHLH genes in QTL
QTL QCLD_D_1A	1	p88-p16	7%	Millet_GLEAN_10023797
				Millet_GLEAN_10023798
				Millet_GLEAN_10035595
QTL QLRND_D_7A	7	si136-si119	10%	Millet_GLEAN_10002496
				Millet_GLEAN_10029582
QTL QGSI_D_7A	7	si256- sims1409	14%	Millet_GLEAN_10016232
				Millet_GLEAN_10037248
				Millet_GLEAN_10002496

Fig 4. Expression patterns of eight bHLH genes under PEG or ABA treatment. A represents the relative expression levels of eight bHLH genes under PEG treatment. B represents the relative expression levels of bHLH genes under ABA treatment. Y axis represents the relative expressed level. X axis represents the time-points of treatment. "*" represent significantly changed comparing with that of CK (0h) ($p < 0.05$).

Many candidate genes among the foxtail millet bHLH members have been linked to drought stress tolerance in other species. Accordingly, bHLH genes are likely to play an important role in drought stress tolerance in fox millet, though the homologs in other species of some candidates were not determined to be involved in tolerance to drought stress by previous studies, such as the homologs of bHLH128, bHLH35, bHLH25, and ORG2. We guessed these candidates (homologs of bHLH128, bHLH35, bHLH25 and ORG2) may have evolved different function from their homologs in foxtail millet or other species, and their functions may have contributed to drought stress tolerance, though perhaps they were not assayed clearly. Our inference is based on the analysis of sequence motifs and a molecular signature of selection. Divergence in motifs and positive selection between homologs or duplication pairs suggested functional divergence may have occurred in these gene trees while genes with de novo functions were created through expansion of the foxtail millet bHLH family.

The role of bHLH in foxtail millet drought resistance

ABA-dependent and -independent signalling pathways appear to be involved in drought stress tolerance. However, previous studies ignored a direct link between AREB and bHLH (such as ICE and Myc) in tolerance to drought stress tolerance. Rather, bHLH was only considered part of the DREB/CBF pathway, which may be involved in tolerance to drought and cold stress [52]. However, analysis of candidate promoters showed that most contained AREB elements (*cis*-acting elements involved in abscisic acid responsiveness; S6 Table). This research showed these candidates may be regulated by AREB. The 14-day foxtail millet shoots were subjected to exogenous ABA treatments of 0, 0.5, 7, and 12 h. The changes in expression of the eight genes under the ABA treatment were similar to the changes in expression of the eight genes under the PEG treatment (i.e., drought stress) over the same treatment times (Fig 4B). This indicated

that these foxtail millet bHLHs genes may be regulated by ABA and that drought may regulate these genes through the ABA-dependent signalling pathway. Analysis of promoters revealed that most promoters of the candidate genes we identified contain MBS elements (MYB binding sites involved in drought inducibility; S6 Table) and some contain G-box element. Accordingly, MYB genes may regulate bHLH genes, and bHLHs may in turn regulate drought responses.

Additionally, we comprehensively identified the promoters of 40 drought-responsive genes in common between foxtail millet and some monocot and dicot species [29, 53–55] including the drought-responsive marker gene COR47 [29], which contains E-box and G-box elements (S7 Table). These genes may accordingly be regulated by bHLH genes. Thus, we propose the following hypothesis. When foxtail millet is under drought stress, some foxtail millet bHLH genes may be regulated by ABA-dependent signalling pathways, including AREB, MYB, or bHLH transcript factors; it is these genes that could affect downstream genes that are directly involved in drought stress responses.

Supporting information

S1 Fig. Roles of bHLHs in foxtail millet hormone signal transduction pathway. Red boxes represents the location of foxtail millet bHLHs.
(TIF)

S2 Fig. Roles of bHLHs in foxtail millet circadian rhythm pathway. Red boxes represents the location of foxtail millet bHLHs.
(TIF)

S1 Table. qRT-PCR primers for bHLH genes in foxtail millet.
(XLS)

S2 Table. bHLH genes and function annotation in foxtail millet.
(XLS)

S3 Table. GO annotation of bHLH genes in foxtail millet.
(XLS)

S4 Table. Physicochemical properties of some homologous groups.
(XLSX)

S5 Table. RPKM value of bHLH genes in foxtail millet tissues.
(XLS)

S6 Table. Cis-acting elements of candidate bHLH promoters in foxtail millet.
(XLS)

S7 Table. Cis-acting elements of drought-responsive gene promoters in foxtail millet.
(XLS)

Author Contributions

Data curation: Pengfei Wang, Haili Wang.

Funding acquisition: Pengfei Wang, Fengshan Ren.

Resources: Wei Liu.

Software: Pengfei Wang, Haili Wang, Yongmei Wang.

References

1. Zhang X, Luo HM, Xu ZC, Zhu YJ, Ji AJ, Song JY, et al. Genome-wide characterisation and analysis of bHLH transcription factors related to tanshinone biosynthesis in Salvia miltiorrhiza. Sci Rep. 2015; 5 (11244). Epub 2015/7/15. https://doi.org/:10.1038 PMID: 26174967; PubMed Central PMCID: PMCPM C4502395.

2. Sun H, Fan HJ, Ling HQ. Genome-wide identification and characterization of the bHLH gene family in tomato. BMC Genomics. 2015; 16(9). Epub 2015/2/22. https://doi.org/10.1186/s12864-014-1209-2 PMID: 25612924; PubMed Central PMCID: PMCPMC4312455.

3. Ledent V, Vervoort M. The basic helix-loop-helix protein family: comparative genomics and phyloge- netic analysis. Genome Res. 2001; 11:754–770. Epub 2001/3/1. https://doi.org/10.1101/gr.177001 PMID: 11337472; PubMed Central PMCID: PMCPMC311049.

4. Toledo-Ortiz G, Huq E, Quail PH. The Arabidopsis basic/helix-loop-helix factor family. The Plant Cell. 2003; 15(8):1749–1770. Epub 2003/8. https://doi.org/10.1105/tpc.013839 PMID: 12897250 PubMed Central PMCID: PMCPMC167167.

5. Sonnenfeld MJ, Delvecchio C, Sun X. Analysis of the transcriptional activation domain of the Drosophila tango bHLH-PAS transcription factor. Dev Genes Evol. 2005; 215(5): 221–229. Epub 2005/4/8. https:// doi.org/10.1007/s00427-004-0462-9 PMID: 15818484.

6. Wang J, Hu Z, Zhao T, Yang Y, Chen T, Yang M, et al. Genome-wide analysis of bHLH transcription fac- tor and involvement in the infection by yellow leaf curl virus in tomato (Solanum lycopersicum). BMC Genomics. 2015; 16(1): 39. Epub 2015/2/5. https://doi.org/10.1186/s12864-015-1249-2 PMID: 25652024; PubMed Central PMCID: PMCPMC4333901.

7. Murre C, McCaw PS, Baltimore D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell. 1989; 56(5): 777–83. Epub 1989/3/10. https://doi.org/10.1016/0092-8674(89)90682-X PMID: 2493990.

8. Jones S. An overview of the basic helix-loop-helix proteins. Genome Biol. 2004; 5(6):226. Epub 2004/ 5/28. https://doi.org/10.1186/gb-2004-5-6-226 PMID: 15186484; PubMed Central PMCID: PMCPMC463060.

9. Massari ME, Murre C. Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. Mol Cell Biol. 2000; 20(2):429–40. Epub 2000/1. PMID: 10611221; PubMed Central PMCID: PMCPMC85097.

10. Kong Q, Pattanaik S, Feller A, Werkman JR, Chai CL, Wang YQ, et al. Regulatory switch enforced by basic helixloop- helix and ACT-domain mediated dimerizations of the maize transcription factor R. PNAS. 2012; 109 (30) E209 1–E2097. Epub 2012/7/24. https://doi.org/10.1073/pnas.1205513109.

11. Li XX, Duan XP, Jiang HX, Sun YJ, Tang YP, Yuan Z, et al. Genome wide analysis of basic/helix-loop- helix TF family in rice and Arabidopsis. Plant Physiol. 2006; 141:1167–1184. Epub 2006/8. https://doi. org/10.1104/pp.106.080580 PMID: 16896230; PubMed Central PMCID: PMCPMC1533929.

12. Feller A, Machemer K, Braun EL, Grotewold E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. Plant J. 2011; 66(1):94–116. Epub 2011/4. https://doi.org/10.1111/j. 1365-313X.2010.04459.x PMID: 21443626.

13. Martinez-Garcia JF, Huq E, Quail PH. Direct targeting of light signals to a promoter element-bound tran- scription factor. Science. 288(5467):859–63. Epub 2000/5/5. https://doi.org/10.1126/science.288. 5467.859 PMID: 10797009

14. Albert NW, Davies KM, Schwinn KE. Gene regulation networks generate diverse pigmentation patterns in plants. Plant Signal Behav. 2014; 9(9):e29526. Epub 2014/6/13. https://doi.org/10.4161/psb.29526 PMID: 25763693; PubMed Central PMCID: PMCPMC4205132.

15. Hichri I, Heppel SC, Pillet J, Léon C, Czemmel S, Delrot S, et al. The basic helix-loop-helix transcription factor MYC1 is involved in the regulation of the flavonoid biosynthesis pathway in grapevine. Mol Plant. 2010; 3(3):509–23. Epub 2010/1/29 https://doi.org/10.1093/mp/ssp118 PMID: 20118183.

16. Matus JT, Poupin MJ, Cañón P, Bordeu E, Alcalde JA, Arce-Johnson P. Isolation of WDR and bHLH genes related to flavonoid synthesis in grapevine (Vitis vinifera L.). Plant Mol Biol. 2010; 72(6):607–20. Epub 2010/1/ 29. Ht tp s: // d oi.org/10.1007/s11103-010-9597-4 https://doi.org/10.1007/s11103-010- 9597-4 PMID: 20112051.

17. Shangguan XX, Yang CQ, Zhang XF, Wang LJ. Functional characterization of a basic helix-loop-helix (bHLH) transcription factor GhDEL65 from cotton (Gossypium hirsutum). Physiol Plant. 2016; 158 (2):200–12. Epub 2016/6/6. https://doi.org/10.1111/ppl.12450 PMID: 27080593.

18. Van Moerkercke A, Steensma P, Gariboldi I, Espoz J, Purnama PC, Schweizer F, et al. The basic helix-loop-helix transcription factor BIS2 is essential for monoterpenoid indole alkaloid production in the medicinal plant Catharanthus roseus. Plant J. 2016; 88(1):3–12. Epub 2016/8/13. https://doi.org/10.1111/tpj.13230 PMID: 27342401.

19. Zhang X, Luo H, Xu Z, Zhu Y, Ji A, Song J, et al. Genome-wide characterisation and analysis of bHLH transcription factors related to tanshinone biosynthesis in Salvia miltiorrhiza. Sci Rep. 2015; 5: 11244. Epub 2015/7/15. https://doi.org/10.1038/srep11244. https://doi.org/10.1038/srep11244 PMID: 26174967; PubMed Central PMCID: PMC PMC4502395.

20. Tian H, Guo H, Dai X, Cheng Y, Zheng K, Wang X, et al. An ABA down-regulated bHLH transcription repressor gene, bHLH129 regulates root elongation and ABA response when overexpressed in Arabidopsis. Sci Rep. 2015; 5: 17587. Epub 2015/12/2. https://doi.org/10.1038/srep17587. https://doi.org/10.1038/srep17587 PMID: 26625868; PubMed Central PMCID: PMCPMC4667245.

21. Nowak K, Gaj MD. Stress-related function of bHLH109 in somatic embryo induction in Arabidopsis. J Plant Physiol. 2016; 193:119–126. Epub 2016/3/2. https://doi.org/10.1016/j.jplph.2016.02.012 PMID: 26973252.

22. Choi H, Oh E. PIF4 Integrates Multiple Environmental and Hormonal Signals for Plant Growth Regulation in Arabidopsis. Mol Cells. 2016; 39(8):587–93. Epub 2016/7/19. https://doi.org/10.14348/molcells.2016.0126 PMID: 27432188; Pub Med Central PMCID: PMCPMC4990750.

23. Raissig MT, Abrash E, Bettadapur A, Vogel JP, Bergmann DC. Grasses use an alternatively wired bHLH transcription factor network to establish stomatal identity. PNAS. 2016; 113(29):8326–31. Epub 2016/7/5. https://doi.org/10.1073/pnas.1606728113 PMID: 27382177; PubMed Central PMCID: PMC PMC 496 11 63.

24. Feng HL, Ma NN, Meng X, Zhang S, Wang JR, Chai S, et al. A novel tomato MYC-type ICE1-like transcription factor, SlICE1a, confers cold, osmotic and salt tolerance in transgenic tobacco. Plant Physiol Biochem. 2013; 73:309–320. Epub 2013/10/18. https://doi.org/10.1016/j.plaphy.2013.09.014 PMID: 24184451.

25. Babitha KC, Vemanna RS, Nataraja KN, Udayakumar M. Overexpression of EcbHLH57 Transcription Factor from Eleusine coracana L. in Tobacco Confers Tolerance to Salt, Oxidative and Drought Stress. PLoS One. 2015; 10(9): e01 370 98. https://doi.org/10.1371/journal.pone.0137098 PMID: 26366726; Pub Med Central PMCID: PMCPMC4569372.

26. Liu W, Tai H, Li S, Gao W, Zhao M, Xie C et al. bHLH122 is important for drought and osmotic stress resistance in Arabidopsis and in the repression of ABA catabolism. New Phytol. 2014; 201(4):1192–1204. Epub 2013/11/21. https://doi.org/10.1111/nph.12607 https://doi.org/10.1111/nph.12607 PMID: 24261563.

27. Liu Y, Ji X, Nie X, Qu M, Zheng L, Tan Z, et al. Arabidopsis AtbHLH112 regulates the expression of genes involved in abiotic stress tolerance by binding to their E-box and GCG-box motifs. New Phytol. 2015; 207(3):692–709. Epub 2015/4/1. https://doi.org/10.1111/nph.13387 PMID: 25827016.

28. Zhao P, Liu P, Yuan G, Jia J, Li X, Qi D, et al. New Insights on Drought Stress Response by Global Investigation of Gene Expression Changes in Sheepgrass (Leymus chinensis). Front Plant Sci. 2016; 7: 954. Epub 2016/6/30. https://doi.org/10.3389/fpls.2016.00954 PMID: 27446180; PubMed Central PMCID: PM C PMC4928129.

29. Qi X, Xie S, Liu Y, Yi F, Yu J. Genome-wide annotation of genes and noncoding RNAs of foxtail millet in response to simulated drought stress by deep sequencing. Plant Mol Biol. 2013; 83(4–5):459–73. Epub 2013/7/17. https://doi.org/10.1007/s11103-013-0104-6 PMID: 23860794.

30. Tang S, Li L, Wang Y, Chen Q, Zhang W, Jia G, et al. Genotype-specific physiological and transcriptomic responses to drought stress in Setaria italica (an emerging model for Panicoideae grasses). Sci Rep. 2017; 7(1):10009. Epub 2017/8/30. https://doi.org/10.1038/s41598-017-08854-6 PMID: 28855520 PMCID: PMC5577110.

31. Liu X, Tang S, Jia G, Schnable JC, Su H, Tang C, et al. The C-terminal motif of SiAGO1b is required for the regulation of growth, development and stress responses in foxtail millet (Setaria italica (L.) P. Beauv). J Exp Bot. 2016; 67(11):3237–3249. Epub 2016/4/4. https://doi.org/10.1093/jxb/erw135 PMID: 27045099 PMCID: PMC4892719.

32. Zhang G, Liu X, Quan Z, Cheng S, Xu X, Pan S, et al. Genome sequence of foxtail millet (Setaria italica) provides insights into grass evolution and biofuel potential. Nat Biotechnol. 2012; 30(6):549–54. Epub 2012/5/13. https://doi.org/10.1038/nbt.2195 PMID: 22580950.

33. Wang PF, Gao C, Bian XT, Zhao SZ, Zhao CZ, Xia H, et al. Genome-Wide Identification and Comparative Analysis of Cytosine-5 DNA Methyltransferase and Demethylase Families in Wild and Cultivated Peanut. Front Plant Sci. 2016; 7: 7. Epub 2016/2/3. https://doi.org/10.3389/fpls.2016.00007 PMID: 26870046; PubMed Central PMCID: PMCPMC4737905.

34. Bailey TL, Williams N, Misleh C, Li WW. MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res. 2006; 34(suppl-2): W369–73. Epub 2016/7/1. https://doi.org/10.1093/nar/gkl198 PMID: 16845028; PubMed Central PMCID: PMCPMC1538909.

35. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 2007; 24(8):1586–91. Epub 2007/5/4. https://doi.org/10.1093/molbev/msm088 PMID: 17483113.

36. Yang Z, Nielsen R, Goldman N, Pedersen AM. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics. 2000; 155(1):431–4 9. Epub 2000/5. PMID: 10790415; PubMed Central PMCID: PMC1461088.

37. Qie L, Jia G, Zhang W, Schnable J, Shang Z, Li W, et al. Mapping of quantitative trait locus (QTLs) that contribute to germination and early seedling drought tolerance in the interspecific cross Setaria italica×-Setaria viridis. PLoS One. 2014; 9(7):e101868. Epub 2014/7/17. https://doi.org/10.1371/journal.pone.0101868 PMID: 25033201; PubMed Central PMCID: PMCPMC4102488

38. Muthamilarasan M, Bonthala VS, Khandelwal R, Jaishankar J, Shweta S, Nawaz K, et al. Global analysis of WRKY transcription factor superfamily in Setaria identifies potential candidates involved in abiotic stress signaling. Front Plant Sci. 2015; 6:910. Epub 2015/10/26. https://doi.org/10.3389/fpls.2015.00910 PMID: 26635818; PubMed Central PMCID:PMCID: PMC4654423

39. Cui P, Lin Q, Ding F, Hu S, Yu J. The transcript-centric mutations in human genomes. Genomics Proteomics Bioinformatics. 2012; 10(1):11–22. Epub 2012/2. https://doi.org/10.1016/S1672-0229(11)60029-6 PMID: 22449397; PubMed Central PMCID: PMCPMC5054492

40. Wang P, Li C, Li C, Zhao C, Xia H, Zhao S, et al. Identification and expression dynamics of three WUSCHEL related homeobox 13 (WOX13) genes in peanut. Dev Genes Evol. 2015; 225(4):221–33. Epub 2015/6/27. https://doi.org/10.1007/s00427-015-0506-3 PMID: 26115849

41. Zhang JZ. Evolution by gene duplication: an update. Trends Ecol. Evol. 2003; 18(6): 292–298. https://doi.org/10.1016/S0169-5347(03)00033-8

42. Zhao H, Yang JR, Xu H, Zhang J. Pseudogenization of the umami taste receptor gene Tas1r1 in the giant panda coincided with its dietary switch to bamboo. Molecular biology and evolution. 2010; 27 (12):2669–73. Epub 2010/6/23. https://doi.org/10.1093/molbev/msq153 PMID: 20573776; PubMed Central PMCID: PMCPMC3108379

43. Yu L, Wang XY, Jin W, Luan PT, Ting N, Zhang YP. Adaptive evolution of additional colobines. Molecular biology and evolution. 2010; 27(1):121–31. Epub 2010/1/1. https://doi.org/10.1093/molbev/msp216 PMID: 19748931.

44. Zhao H, Rossiter SJ, Teeling EC, Li C, Cotton JA, Zhang S. The evolution of color vision in nocturnal mammals. PNAS. 2009; 106(22):8980–5. Epub 2009/5/26. https://doi.org/10.1073/pnas.0813201106 PMID: 19470491; PubMed Central PMCID: PMCPMC2690009.

45. Soskine M, Tawfik DS. Mutational effects and the evolution of new protein functions. Nature reviews Genetics. 2010; 11(8):572–82. Epub 2010/8/1. https://doi.org/10.1038/nrg2808 PMID: 20634811.

46. Song H, Wang P, Lin JY, Zhao C, Bi Y, Wang X. Genome-Wide Identification and Characterization of WRKY Gene Family in Peanut. Front Plant Sci. 2016; 7: 534. Epub 2016/4/26. https://doi.org/10.3389/fpls.2016.00534 PMID: 27200012; PubMed Central PMCID: PMCPMC4845656.

47. Nei M, Rooney AP. Concerted and birth-and-death evolution of multigene families. Annu Rev Genet. 2005; 39:121–152. Epub 2005/12/15. https://doi.org/10.1146/annurev.genet.39.073003.112240 PMID: 16285855; PubMed Central PMCID: PMCPMC1464479.

48. Rampey RA, AW, Hobbs BN, Tierney MP, Lahner B, Salt DE, et al. An Arabidopsis basic helix-loop-helix leucine zipper protein modulates metal homeostasis and auxin conjugate responsiveness. Genetics. 2006; 174(4):1841–57. Epub 2006/10/8. https://doi.org/10.1534/genetics.106.061044 PMID: 17028341; PubMed Central PMCID: PMC PMC1698629

49. Aparicio F, Pallás V. The coat protein of Alfalfa mosaic virus interacts and interferes with the transcriptional activity of the bHLH transcription factor ILR3 promoting salicylic-dependent defense signaling response. Mol Plant Pathol. 2016; 18(2):173–186. Epub 2016/2/29. https://doi.org/10.1111/mpp.12388 PMID: 26929142

50. Chen L, Chen Y, Jiang J, Chen S, Chen F, Guan Z, et al. The constitutive expression of Chrysanthemum dichrum ICE1 in Chrysanthemum grandiflorum improves the level of low temperature, salinity and drought tolerance. Plant Cell Rep. 2012; 31(9):1747–58. Epub 2012/5/30. https://doi.org/10.1007/s00299-012-1288-y PMID: 22645020.

51. Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. ctivators in abscisic acid signaling. Plant Cell. 2003; 15(1):63–78. Epub 2003/1. https://doi.org/10.1105/tpc.006130 PMID: 12509522; PubMed Central PMCID: PMCPMC143451.

52. Huang GT, Ma SL, Bai LP, Zhang L, Ma H, Jia P, et al. Signal transduction during cold, salt, and drought stresses in plants. Mol Biol Rep. 2012; 39(2):969–87. Epub 2011/5/15. https://doi.org/10.1007/s11033-011-0823-1 PMID: 21573796.

53. Matsui A, Ishida J, Morosawa T, Mochizuki Y, Kaminuma E, Endo TA, et al. Arabidopsis transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. Plant Cell Physiol. 2008; 49(8):1135–49. Epub 2008/7/13. https://doi.org/10.1093/pcp/pcn101 PMID: 18625610.

54. Zheng J, Fu JJ, Guo MY, Huai JL, Liu YJ, Jian M, et al. Genome-wide transcriptome analysis of two maize inbred lines under drought stress. Plant Mol Biol. 2010; 72(4–5): 407–421. Epub 2010/3. https://doi.org/10.1007/s11103-009-9579-6 PMID: 19953304.

55. Dugas DV, Monaco MK, Oslon A, Klein RR, Kumari S, Ware D, et al. Functional annotation of the transcriptome of Sorghum bicolor in response to osmotic stress and abscisic acid. BMC Genomics. 2011; 12:514. Epub 2011/10/18. https://doi.org/10.1186/1471-2164-12-514 PMID: 22008187; PubMed Central PMC ID: P M CPMC3219791.