Send Orders for Reprints to reprints@benthamscience.net

624 *Current Genomics,* 2020, *21,* 624-644

RESEARCH ARTICLE

Genome-wide Identification, Evolution and Expression Analysis of Basic Helix-loop-helix (bHLH) Gene Family in Barley (*Hordeum vulgare L.*)

Qinglin Ke[1,#], Wenjing Tao[1,#], Tingting Li[1], Wenqiu Pan[2], Xiaoyun Chen[1], Xiaoyu Wu[1], Xiaojun Nie[2] and Licao Cui[1,2,*]

[1]*College of Bioscience and Engineering, Jiangxi Agricultural University, Nanchang 330045, Jiangxi, China;* [2]*State Key Laboratory of Crop Stress Biology in Arid Areas and College of Agronomy, Northwest A&F University, Yangling 712100, Shaanxi, China*

ARTICLE HISTORY

Received: June 15, 2020
Revised: September 17, 2020
Accepted: October 05, 2020

DOI:
10.2174/1389202921999201102165537

Abstract: ***Background*:** The basic helix-loop-helix (bHLH) transcription factor is one of the most important gene families in plants, playing a key role in diverse metabolic, physiological, and developmental processes. Although it has been well characterized in many plants, the significance of the bHLH family in barley is not well understood at present.

***Methods*:** Through a genome-wide search against the updated barley reference genome, the genomic organization, evolution and expression of the bHLH family in barley were systematically analyzed.

***Results*:** We identified 141 bHLHs in the barley genome (*HvbHLHs*) and further classified them into 24 subfamilies based on phylogenetic analysis. It was found that *HvbHLHs* in the same subfamily shared a similar conserved motif composition and exon-intron structures. Chromosome distribution and gene duplication analysis revealed that segmental duplication mainly contributed to the expansion of *HvbHLHs* and the duplicated genes were subjected to strong purifying selection. Furthermore, expression analysis revealed that *HvbHLHs* were widely expressed in different tissues and also involved in response to diverse abiotic stresses. The co-expression network was further analyzed to underpin the regulatory function of *HvbHLHs*. Finally, 25 genes were selected for qRT-PCR validation, the expression profiles of *HvbHLHs* showed diverse patterns, demonstrating their potential roles in relation to stress tolerance regulation.

***Conclusion*:** This study reported the genome organization, evolutionary characteristics and expression profile of the bHLH family in barley, which not only provide the targets for further functional analysis, but also facilitate better understanding of the regulatory network bHLH genes involved in stress tolerance in barley.

Keywords: Barley, bHLH family, expression profile, qRT-PCR, transcription factor, helix-loop-helix.

1. INTRODUCTION

Transcription factors (TFs) play essential roles in regulating diverse physiological and developmental processes in eukaryotes through interacting with cis-element to activate or repress the expression of downstream functional genes [1, 2]. As the second-largest TF family in plants, the basic helix-loop-helix (bHLH) family is composed of approximately 60 amino acids with two functional domains, namely a basic conserved motif and a HLH motif [3, 4]. The basic motif consists of approximately 17 amino acids and possesses the typical six basic residues at the N-terminal, which functions to bind to the E-box (5'-CANNTG-3') or the variant G-box (5'-CACGTG-3') element [5]. The HLH motif is composed of two amphipathic α helices separated by a loop region with variable size, and the loop region could participate in the formation of homodimeric or heterodimeric complexes [6, 7], which contribute to the specific binding of bHLH transcription factors on their target genes [5].

Up to now, the bHLH gene family has been extensively identified in many plant species. It is reported that 167 bHLH genes were identified in *Arabidopsis thaliana* [2], 167 in rice (*Oryza sativa)* [7], 146 in *Brachypodium distachyon* [8], 571 in wheat (*Triticum aestivum*) [9], 208 in maize (*Zea mays*) [10] and 602 in *Brassica napus* [11]. The bHLH genes are widely involved in diverse signal transductions and metabolic pathways. For example, the bHLH transcription factor ROOT HAIR DEFECTIVE SIX-LIKE4 (*RSL4*) in *Arabidopsis* stimulates root hair formation [12]; three duplicated transcription factors *AtbHLH010*, *AtbHLH089* and *At-*

*Address correspondence to this author at the College of Bioscience and Engineering, Jiangxi Agricultural University, Nanchang 330045, Jiangxi, China; and State Key Laboratory of Crop Stress Biology in Arid Areas and College of Agronomy, Northwest A&F University, Yangling 712100, Shaanxi, China; Tel: 086-791-83813459; Fax: 086-791-83813459;
E-mail: cuilicao@jxau.edu.cn

[#] *These authors contributed equally to this work.*

 ©2020 Bentham Science Publishers

bHLH091 mediate anther development in *Arabidopsis* [13]; the *Arabidopsis* bHLH transcription factor *CFLAP1* (*CFL1* associated protein 1) and its homolog *CFLAP2* regulate cuticle development [14]; the major quantitative trait locus An-1 in wild rice, which encodes a basic helix-loop-helix protein, regulates awn development, grain size, and grain number in rice [15]. Meanwhile, many bHLH genes respond to various types of stress, such as salt, drought and cold stresses. For instance, *AtbHLH112* in *Arabidopsis* responds to drought and salt signal transduction through proline biosynthesis and reactive oxygen species (*ROS*) scavenging pathways [16]; *OsbHLH006* and *OsbHLH148* in rice play a transcriptional role in drought stress through the jasmonic acid signaling pathway [17, 18]; the bHLH transcription factor PHYTOCHROME-INTERACTING FACTOR 4 (*PIF4*) in *Arabidopsis* regulates stomatal lineage initiation under high-temperature stress [19]; alternatively, *OsbHLH1* in rice, which is independent of ABA, responds to cold stress [20]. In addition, some bHLH genes are also involved in plant photomorphogenesis [21, 22] and phytohormone signaling, such as brassinosteroid [23], abscisic acid [24, 25] and jasmonic acid [26, 27].

Barley (*Hordeum vulgare L*) is one of the earliest domesticated crops approximately 10,000 years ago, and nowadays, ranks as the fourth most important cereal crop worldwide in terms of cultivated areas and total yield [28]. It has been well-studied in genetics, cytology and genomics and therefore, is considered as a research model for the large-genome small grain temperate cereals. Recently, the completion of its reference genome laid the foundation to identify the gene family at the genome-wide level. Here, we comprehensively identified 141 bHLHs in barley (*HvbHLHs*) based on a genome-wide search method. Then, the phylogenetic relationships, gene duplication, cis-regulatory elements, conserved motifs and gene structure as well as co-expression networks of these *HvbHLH*s were also investigated. Finally, the expression patterns of *HvbHLH*s were detected based on a total of 148 RNA-seq samples, and the expression of 25 genes was further validated by qRT-PCR analysis. This study will not only shed light on the functional study of *HvbHLHs*, but also contribute to better understanding of the regulatory roles and evolutionary mechanism of the bHLH gene family in barley and beyond.

2. MATERIALS AND METHODS

2.1. Identification of bHLH Gene Family in Barley

The protein sequences of the updated barley reference genome of genotype Morex [28] were obtained from the IPK database (http://webblast.ipk-gatersleben.de/barley_ibsc/) and used as the local protein database. The bHLH proteins of *Arabidopsis* and rice were downloaded from the TAIR (https://www.arabidopsis.org/) and RGAP (http://rice.plantbiology.msu.edu/) databases respectively, and then used as the query to search against the local protein database by BLASP tool with the identity more than 50% and e-value of 1e-5 as a threshold. Then, the Hidden Markov Model (HMM) of bHLH domain (PF00010) was used as a query to search against the barley local protein database using the HMMER 3.0 tool. The results from these two methods were integrated and explained by manual editing to remove redundancy. The remaining proteins were further verified using Plant Transcription Factor Database (PlantTFDB) v.4.0 (http://planttfdb.cbi.pku.edu.cn/prediction.php), HMMER v3.2.1 (https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan) and PFAM (http://pfam.xfam.org/search/sequence) database. Only the proteins having the complete bHLH domain were considered as putative *HvbHLHs*.

Finally, the ProtParam tool (https://web.expasy.org/protparam/) was used to estimate the molecular weight (MW), theoretical isoelectric point (pI), instability index (II), aliphatic index and Gravy value. The subcellular localization was predicted based on the Plant-mPLoc online tools (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/). To confirm the existence of these *HvbHLHs*, the expressed sequence tags (ESTs) were downloaded from the NCBI database and a BLASTN search was performed with a cut-off e-value of 1e-5 and identity of 80%.

2.2. Multiple Sequence Alignment and Phylogenetic Analysis

Multiple sequence alignment of the *HvbHLH*s was performed using ClustalX v1.83 with default parameters. To visualize the conserved motifs, the sequence logo of the bHLH domain was generated using WEB LOGO tool (http://weblogo.berkeley.edu/logo.cgi). An unrooted neighbor-joining (NJ) tree was constructed using MEGA v7.0 based on the conserved bHLH domain sequences with the following parameters: 1000 times bootstrap, Poisson model and 95% partial deletion.

2.3. Gene Structure and Conserved Motif Analysis

The intron-exon structure was obtained based on the gene annotation file of the barley genome and visualized using the online Gene Structure Display Server (http://gsds.cbi.pku.edu.cn/). The conserved motifs were identified using MEME software (http://meme-suite.org/tools/meme) with the following parameters: maximum number of motifs with 10, optimum motif width with 8 to 50, other settings with default. The predicted motifs were then submitted to the InterPro database to determine their putative functions (http://www.ebi.ac.uk/interpro/search/sequence/). The upstream 1.5kb sequences from the transcription initiation site of the *HvbHLHs* were extracted from the barley genome, and then submitted to the PlantCARE database to detect the *cis-acting elements.* (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

2.4. Gene Duplication and Molecular Evolution Analysis

A duplication event was defined based on the criteria described by Chen *et al.*, including (a) the shorter alignment sequence covered > 70% of the longer gene in length; (b) the aligned region had an identity > 70%; and (c) only one duplication event was counted for the tightly linked genes [29]. In order to reveal the evolutionary effect of bHLH genes during barley domestication, the bHLH genes in wild barley were also identified using the same methods as barley domestication with all protein sequences of wild barley (http://db.ncgr.ac.cn/wild_barley/) as local protein database. The orthologous bHLH gene pairs between wild and culti-

vated barley were identified using inParanoid v4.1 [30]. Moreover, the synteny relationships between barley and other species, including *Arabidopsis thaliana*, *Brachypodium distachyon*, *Oryza sativa*, *Solanum lycopersicum*, *Solanum tuberosum*, and *Vitis vinifera,* were detected using the MCScanX software [31]. The Ka (non-synonymous substitution)/Ks (synonymous substitution) ratio was calculated to estimate the evolutionary effect using the codeml program embedded in the PAML v1.3.1 package [32]. The formula $T=Ks/2\lambda$ was employed to estimate the divergence time between these gene pairs, where λ refers to the mutation rate showing 6.5×10^{-9} synonymous substitutions per site per year. Circos v0.67 tool was used to visualize the orthologous genes pairs.

2.5. Expression Profiles and Regulatory Network Analysis

A total of 148 RNA-seq samples, of various developmental stages and tissues as well as diverse biotic and abiotic stresses, were downloaded from the NCBI Sequence Read Archive (SRA) database and then were used to investigate the expression patterns of these *HvbHLHs*. The accession number and sample information of RNA-seq are listed in Table **S1**. Then, HISAT2 v2.1.0 and StringTie v1.3.5 pipeline were used to calculate the FPKM (fragments per kilobase of transcript per million fragments mapped) value and the differentially expressed genes were identified using the R package Ballgown with the following parameters: fold change ≥ 2, FDR(false discovery rate) ≤ 0.05, and the absolute ratio of $\log2 \geq 1$ [33]. The log2 transformed FPKM value was displayed using the R package heatmap. In order to identify the candidate miRNA targets, the transcript of *HvbHLHs* was submitted to psRNATarget (http://plantgrn.noble.org/psRNATarget/) to search against all barley miRNAs in the miRbase. Finally, the WGCNA package in the R software was employed to run the co-expression networks using the 148 RNA-seq samples. The genes linked to the bHLHs with the top 5% weighted parameter obtained from WGCNA were selected for further analysis. The functions of co-expressed genes were determined using BLAST against the *Arabidopsis* protein database. The network was visualized using the Cytoscape v3.8.0.

2.6. Validation the Expression of *HvbHLH*s by qRT-PCR Analysis

Seeds of barley cv. Morex were hydroponically cultured in the growth chamber under controlled conditions (23 ± 1 °C, 16-h light/8-h dark cycle). And the seedlings at the three-leaf stage were used for stress treatment. For salt stress, the plants were incubated in 150 mM NaCl solution for 0, 6, 12 and 24 hours. For drought stress, they were incubated into 19.2% (w/v) PEG-6000 solution for 0, 6, 12 and 24 hours. For cold and heat stresses, they were kept at 4 °C for 0, 6, 12 and 24 hours, and at 42 °C for 0, 1, 6 hours, respectively. Seedlings under the normal condition at the same time point were used as the control. Leaves of all samples were collected from three to five plants at each time point with three biological replications. All the plant materials were promptly frozen in liquid nitrogen and stored at -80 °C for RNA extraction. Total RNA isolation, cDNA synthesis and qRT-PCR reaction were performed as described previously [34]. Then, 25 candidates *HvbHLHs* were selected to design primers for qRT-PCR analysis. *HvACTIN2* (GenBank accession no. AY145451.1) was used as the internal reference and all the primers used in this study were listed in Table **S2**. The TB-Green® Premix Ex Taq™ II kit (Takara, Dalian, China) was used for qRT-PCR amplification using the QuantStudio™ Real-Time PCR system (ThermoFisher, USA). The thermal cycling condition was 95 °C temperature for 30 s followed by 40 cycles of 95 °C for 3s, and at 60 °C for 30 s. Three technological replications were applied and the expression profile was calculated using the $2^{-\Delta\Delta CT}$ method [35].

3. RESULTS AND DISCUSSION

3.1. Identification of bHLH Gene Family in Barley

Using the method as described in Materials and Methods, a total of 141 bHLH encoding genes were identified in barley (Table **1**), accounting for 0.35% of the total annotated barley genes, which was similar to that of rice (0.44%), while slightly less than *Brachypodium* (0.55%) and *Arabidopsis* (0.65%) (Table **S3**) [7, 8]. Since there is no standard nomenclature and description, these barley bHLH proteins were named *HvbHLH1* to *HvbHLH141* according to their physical positions on the chromosomes. As shown in Table **1** and Fig. **S1**, these *HvbHLHs* were unevenly distributed across the 7 chromosomes. In detail, 26 *HvbHLHs* were found on chromosome 5, ranking the most abundant one, followed by chromosome 7 (25) and chromosome 4 (22), whereas only 10 genes were mapped on chromosome 1. The physical and chemical properties of the putative genes were then characterized. The size of them ranged from 36 to 887 amino acids, molecular weight ranged from 4.22 to 95.97 kDa and theoretical isoelectric point ranged from 4.49 to 11.9, which are similar to those of *Brachypodium* [8]. The value of hydropathicity ranged from -1.247 (*HvbHLH62*) to 0.057 (*HvbHLH55*), with an average of -0.49. Except for *HvbHLH55*, other *HvbHLHs* displayed the negative gravity values, representing the hydrophilic characteristics of barley bHLH genes. Subcellular location prediction showed that most (90.07%) of the *HvbHLHs* were localized in the nucleus, which was similar to the previous study on *Brassica napus* [11]. To further validate the existence of these *HvbHLHs*, we performed BLAST search against the barley ESTs. Results showed that 116 out of 141 (82.3%) *HvbHLHs* were supported by EST hits, of which *HvbHLH53,* possessed the most hits with 27 ESTs. At the same time, 25 *HvbHLHs* had no EST support, mainly due to their lower expression level and tissue or stage-specific expression.

3.2. Conserved Motif Analysis and DNA Binding Ability of *HvbHLHs*

To evaluate the sequence conservation of these *HvbHLHs*, the conserved bHLH domain sequences were aligned to identify conserved sites. Totally, 5 amino acids in the basic region, 7 in the HLH region, 3 in the loop region and 7 in the second helix region were found to be highly conserved with the consensus sequence ratio more than 50% (Fig. **1**, Table **S4**). Among them, the residues Arg-16, Arg-17, Leu-27, Pro-32, and Leu-59 showed identical composition in all of the 141 *HvbHLHs*, which was consistent with

Table 1. Characteristics of bHLH transcription factor gene family in barley.

HvbHLHs	Barley Gene ID	Protein Length (aa)	Isoelectric Point (pI)	Molecular Weight (kDa)	EST Validation	Subcellular Location	*Arabidopsis thaliana* Ortholog	*Oryza sativa* Ortholog
HvbHLH 1	HOR-VU0Hr1G012790	287	558	3123436	-	Nucleus	-	-
HvbHLH 2	HOR-VU0Hr1G014300	330	561	3567917	1	Nucleus	-	-
HvbHLH 3	HOR-VU0Hr1G019640	186	793	2031412	-	Nucleus	-	OsbHLH122
HvbHLH 4	HOR-VU0Hr1G020200	440	558	4754577	5	Nucleus	AtbHLH049	OsbHLH086
HvbHLH 5	HOR-VU0Hr1G021220	60	1067	658106	6	Chloroplast	-	-
HvbHLH 6	HOR-VU0Hr1G023050	164	776	1823667	-	Nucleus	-	-
HvbHLH 7	HOR-VU0Hr1G025420	164	649	1831167	-	Nucleus	-	-
HvbHLH 8	HOR-VU1Hr1G017900	493	629	5311099	14	Nucleus	AtbHLH008	OsbHLH106
HvbHLH 9	HOR-VU1Hr1G020370	380	568	4144222	-	Nucleus	-	-
HvbHLH 10	HOR-VU1Hr1G024280	212	88	2308352	3	Nucleus	-	-
HvbHLH 11	HOR-VU1Hr1G040810	60	1012	666184	8	Chloroplast	-	-
HvbHLH 12	HOR-VU1Hr1G043780	60	102	672488	7	Chloroplast	-	-
HvbHLH 13	HOR-VU1Hr1G050560	684	606	7390273	17	Nucleus	AtbHLH006	OsbHLH009
HvbHLH 14	HOR-VU1Hr1G054260	326	848	3518579	9	Nucleus	AtbHLH016	OsbHLH113
HvbHLH 15	HOR-VU1Hr1G071330	419	514	4531638	2	Nucleus	AtbHLH097	OsbHLH051
HvbHLH 16	HOR-VU1Hr1G072810	238	77	2591731	18	Nucleus	-	OsbHLH058
HvbHLH 17	HOR-VU1Hr1G079450	400	717	4133974	8	Nucleus	-	OsbHLH119
HvbHLH 18	HOR-VU2Hr1G013450	393	79	421995	7	Nucleus	AtbHLH047	OsbHLH063
HvbHLH 19	HOR-VU2Hr1G020990	238	775	2592061	-	Nucleus	AtbHLH045	OsbHLH055
HvbHLH 20	HOR-VU2Hr1G034320	36	998	422302	7	-	-	-
HvbHLH 21	HOR-VU2Hr1G035190	264	501	286704	2	Nucleus	-	OsbHLH128

(Table 1) contd….

HvbHLHs	Barley Gene ID	Protein Length (aa)	Isoelectric Point (pI)	Molecular Weight (kDa)	EST Validation	Subcellular Location	*Arabidopsis thaliana* Ortholog	*Oryza sativa* Ortholog
HvbHLH 22	HOR-VU2Hr1G044230	294	861	3242503	10	Nucleus	AtbHLH105	OsbHLH057
HvbHLH 23	HOR-VU2Hr1G060680	362	599	3875149	3	Nucleus	-	OsbHLH104
HvbHLH 24	HOR-VU2Hr1G066000	194	449	2171025	1	Nucleus	-	-
HvbHLH 25	HOR-VU2Hr1G066100	305	501	3364493	6	Nucleus	-	OsbHLH006
HvbHLH 26	HOR-VU2Hr1G068960	306	-	-	1	Nucleus	AtbHLH029	OsbHLH156
HvbHLH 27	HOR-VU2Hr1G073240	164	631	1850639	6	Nucleus	-	-
HvbHLH 28	HOR-VU2Hr1G096810	250	10	2727246	-	Nucleus	-	OsbHLH012
HvbHLH 29	HOR-VU2Hr1G104040	147	91	1680628	10	Nucleus	-	OsbHLH101
HvbHLH 30	HOR-VU2Hr1G106030	75	1139	902536	1	Nucleus	-	-
HvbHLH 31	HOR-VU2Hr1G114070	253	599	2775054	3	Nucleus	AtbHLH100	-
HvbHLH 32	HOR-VU2Hr1G115240	97	608	1041362	1	Nucleus	AtbHLH135	OsbHLH154
HvbHLH 33	HOR-VU3Hr1G000170	253	942	27832	-	Nucleus	-	-
HvbHLH 34	HOR-VU3Hr1G000180	211	946	2334342	3	Nucleus	AtbHLH162	-
HvbHLH 35	HOR-VU3Hr1G000280	210	92	2301307	4	Nucleus	AtbHLH162	-
HvbHLH 36	HOR-VU3Hr1G000830	275	543	2930171	1	Nucleus	AtbHLH086	OsbHLH125
HvbHLH 37	HOR-VU3Hr1G002620	230	958	2551222	-	Nucleus	-	-
HvbHLH 38	HOR-VU3Hr1G018680	268	89	2899176	4	Nucleus	AtbHLH051	OsbHLH035
HvbHLH 39	HOR-VU3Hr1G027630	98	8.98	10.61165	-	Nucleus	-	-
HvbHLH 40	HOR-VU3Hr1G034540	284	9.82	31.34548	12	Nucleus	AtbHLH093	-
HvbHLH 41	HOR-VU3Hr1G046430	59	10.16	6.65267	11	Nucleus	-	-
HvbHLH 42	HOR-VU3Hr1G048770	482	8.27	50.13659	2	Nucleus	-	OsbHLH117

(Table 1) contd….

HvbHLHs	Barley Gene ID	Protein Length (aa)	Isoelectric Point (pI)	Molecular Weight (kDa)	EST Validation	Subcellular Location	*Arabidopsis thaliana* Ortholog	*Oryza sativa* Ortholog
HvbHLH 43	HOR-VU3Hr1G057280	767	9.51	83.53867	13	Nucleus	-	OsbHLH008
HvbHLH 44	HOR-VU3Hr1G057780	310	6.01	33.84592	1	Nucleus	-	-
HvbHLH 45	HOR-VU3Hr1G062240	60	10.37	6.67089	9	Chloroplast	-	-
HvbHLH 46	HOR-VU3Hr1G066390	465	7.77	49.81545	4	Nucleus	-	OsbHLH010
HvbHLH 47	HOR-VU3Hr1G066600	405	6.06	41.38465	6	Nucleus	-	OsbHLH118
HvbHLH 48	HOR-VU3Hr1G087710	210	5.88	22.86064	1	Nucleus	-	OsbHLH123
HvbHLH 49	HOR-VU3Hr1G089090	378	7.29	39.87731	19	Nucleus	AtbHLH128	OsbHLH109
HvbHLH 50	HOR-VU3Hr1G096460	231	5.09	23.54851	4	Nucleus	-	-
HvbHLH 51	HOR-VU3Hr1G108660	232	6.63	25.95128	16	Nucleus	-	-
HvbHLH 52	HOR-VU3Hr1G108670	227	8.27	25.21689	25	Nucleus	-	-
HvbHLH 53	HOR-VU3Hr1G108680	246	6.5	27.00253	27	Nucleus	-	OsbHLH056
HvbHLH 54	HOR-VU4Hr1G003210	365	5.5	38.87177	6	Nucleus	-	OsbHLH020
HvbHLH 55	HOR-VU4Hr1G003340	269	9.28	29.29683	7	Nucleus	-	-
HvbHLH 56	HOR-VU4Hr1G009440	331	5.51	35.87252	3	Nucleus	AtbHLH025	OsbHLH018
HvbHLH 57	HOR-VU4Hr1G009970	316	5.88	34.79796	2	Nucleus	-	OsbHLH084
HvbHLH 58	HOR-VU4Hr1G013720	313	4.64	34.03353	6	Nucleus	-	OsbHLH131
HvbHLH 59	HOR-VU4Hr1G020740	887	5.99	95.97193	18	Nucleus	-	OsbHLH151
HvbHLH 60	HOR-VU4Hr1G032690	309	11.84	34.08783	8	Nucleus	-	-
HvbHLH 61	HOR-VU4Hr1G039390	173	9.33	19.03418	-	Nucleus	-	-
HvbHLH 62	HOR-VU4Hr1G046830	43	11.51	4.82556	7	-	-	-
HvbHLH 63	HOR-VU4Hr1G052500	205	11.12	22.85349	8	Nucleus	-	-
HvbHLH 64	HOR-VU4Hr1G061760	402	5.52	42.92323	-	Nucleus	AtbHLH030	OsbHLH040

(Table 1) contd….

HvbHLHs	Barley Gene ID	Protein Length (aa)	Isoelectric Point (pI)	Molecular Weight (kDa)	EST Validation	Subcellular Location	*Arabidopsis thaliana* Ortholog	*Oryza sativa* Ortholog
HvbHLH 65	HOR-VU4Hr1G065640	313	5.96	34.18329	-	Nucleus	-	OsbHLH019
HvbHLH 66	HOR-VU4Hr1G069820	328	5.74	35.50931	9	Nucleus	AtbHLH085	OsbHLH129
HvbHLH 67	HOR-VU4Hr1G075340	85	6.58	9.63389	4	Nucleus	AtbHLH136	OsbHLH153
HvbHLH 68	HOR-VU4Hr1G078270	224	9.21	24.70528	4	Nucleus	-	-
HvbHLH 69	HOR-VU4Hr1G078880	346	6.17	36.89555	1	Nucleus	-	-
HvbHLH 70	HOR-VU4Hr1G078910	344	6.41	37.3124	-	Nucleus	-	-
HvbHLH 71	HOR-VU4Hr1G080890	355	5.17	38.55531	6	Nucleus	-	OsbHLH003
HvbHLH 72	HOR-VU4Hr1G084850	298	8.43	31.4009	-	Nucleus	-	OsbHLH052
HvbHLH 73	HOR-VU4Hr1G087590	257	5.14	27.7684	12	Nucleus	-	-
HvbHLH 74	HOR-VU4Hr1G087610	251	4.98	27.97863	3	Nucleus	-	-
HvbHLH 75	HOR-VU4Hr1G090880	50	10.38	5.53851	7	Chloroplast	-	-
HvbHLH 76	HOR-VU5Hr1G002090	341	6.42	35.58769	9	Nucleus	-	OsbHLH021
HvbHLH 77	HOR-VU5Hr1G011780	341	5.65	37.14074	6	Nucleus	-	OsbHLH103
HvbHLH 78	HOR-VU5Hr1G018100	312	6.21	33.78668	15	Nucleus	-	OsbHLH130
HvbHLH 79	HOR-VU5Hr1G023670	202	-	-	11	Nucleus	-	-
HvbHLH 80	HOR-VU5Hr1G031400	60	10.72	6.61183	7	Chloroplast	-	-
HvbHLH 81	HOR-VU5Hr1G040090	760	6.94	82.40031	6	Nucleus	AtbHLH156	OsbHLH150
HvbHLH 82	HOR-VU5Hr1G053800	36	9.52	4.25305	8	-	-	-
HvbHLH 83	HOR-VU5Hr1G057460	541	9.36	61.68851	5	Chloroplast	-	-
HvbHLH 84	HOR-VU5Hr1G059290	287	5.58	31.23436	-	Nucleus	-	-
HvbHLH 85	HOR-VU5Hr1G065450	383	9.37	39.80757	7	Nucleus	-	OsbHLH100

(Table 1) contd….

HvbHLHs	Barley Gene ID	Protein Length (aa)	Isoelectric Point (pI)	Molecular Weight (kDa)	EST Validation	Subcellular Location	*Arabidopsis thaliana* Ortholog	*Oryza sativa* Ortholog
HvbHLH 86	HOR-VU5Hr1G066530	367	5.38	37.86235	6	Nucleus	-	OsbHLH039
HvbHLH 87	HOR-VU5Hr1G068110	238	5.96	25.45977	2	Nucleus	-	OsbHLH120
HvbHLH 88	HOR-VU5Hr1G069580	210	9.78	22.7207	-	Nucleus	-	OsbHLH043
HvbHLH 89	HOR-VU5Hr1G070000	411	6.2	43.54592	1	Nucleus	AtbHLH096	OsbHLH046
HvbHLH 90	HOR-VU5Hr1G070510	427	6.38	44.84281	5	Nucleus	AtbHLH049	OsbHLH085
HvbHLH 91	HOR-VU5Hr1G070800	515	6.98	54.97378	15	Nucleus	-	OsbHLH032
HvbHLH 92	HOR-VU5Hr1G075430	258	7.63	28.29626	3	Nucleus	-	-
HvbHLH 93	HOR-VU5Hr1G077390	279	5.94	30.66625	7	Nucleus	AtbHLH044	OsbHLH081
HvbHLH 94	HOR-VU5Hr1G079900	270	-	-	-	Nucleus	-	-
HvbHLH 95	HOR-VU5Hr1G080040	268	6	28.90752	-	Nucleus	-	OsbHLH167
HvbHLH 96	HOR-VU5Hr1G080080	210	11.9	22.81786	-	Nucleus	-	-
HvbHLH 97	HOR-VU5Hr1G093310	233	9.44	25.75102	11	Nucleus	-	-
HvbHLH 98	HOR-VU5Hr1G093960	288	6.52	31.5373	4	Nucleus	AtbHLH092	OsbHLH148
HvbHLH 99	HOR-VU5Hr1G097520	292	8.93	31.56547	9	Nucleus	AtbHLH139	OsbHLH134
HvbHLH 100	HOR-VU5Hr1G106310	312	8.73	33.55686	9	Nucleus	AtbHLH007	OsbHLH098
HvbHLH 101	HOR-VU5Hr1G107520	175	-	-	3	Nucleus	-	-
HvbHLH 102	HOR-VU6Hr1G002140	60	10.37	6.62881	8	Chloroplast	-	-
HvbHLH 103	HOR-VU6Hr1G012730	558	4.63	59.01517	3	Nucleus	AtbHLH021	OsbHLH005
HvbHLH 104	HOR-VU6Hr1G012760	558	4.63	59.01517	3	Nucleus	AtbHLH021	OsbHLH005
HvbHLH 105	HOR-VU6Hr1G020520	297	-	-	5	Nucleus	AtbHLH095	OsbHLH146
HvbHLH 106	HOR-VU6Hr1G020710	59	10.7	6.72385	7	Chloroplast	-	-
HvbHLH 107	HOR-VU6Hr1G039830	631	8.11	66.77104	1	Nucleus	AtbHLH041	OsbHLH029

(Table 1) contd….

HvbHLHs	Barley Gene ID	Protein Length (aa)	Isoelectric Point (pI)	Molecular Weight (kDa)	EST Validation	Subcellular Location	*Arabidopsis thaliana* Ortholog	*Oryza sativa* Ortholog
HvbHLH 108	HOR-VU6Hr1G054910	306	7.21	32.17456	4	Nucleus	-	OsbHLH110
HvbHLH 109	HOR-VU6Hr1G067000	140	9.5	15.36348	1	Nucleus	-	-
HvbHLH 110	HOR-VU6Hr1G068110	98	9.77	10.80136	1	Chloroplast	-	-
HvbHLH 111	HOR-VU6Hr1G068980	353	8	37.25795	6	Nucleus	AtbHLH058	OsbHLH079
HvbHLH 112	HOR-VU6Hr1G069690	114	9.72	12.2922	-	Nucleus	-	-
HvbHLH 113	HOR-VU6Hr1G072300	407	8.89	44.23082	1	Nucleus	-	OsbHLH034
HvbHLH 114	HOR-VU6Hr1G081120	318	6.7	33.18008	2	Nucleus	-	-
HvbHLH 115	HOR-VU6Hr1G085500	193	5.72	19.86918	3	Nucleus	AtbHLH082	OsbHLH115
HvbHLH 116	HOR-VU6Hr1G088020	227	7.14	25.1175	1	Nucleus	-	-
HvbHLH 117	HOR-VU7Hr1G024790	251	-	-	-	Nucleus	-	OsbHLH142
HvbHLH 118	HOR-VU7Hr1G026560	338	6.54	36.56862	14	Nucleus	-	OsbHLH108
HvbHLH 119	HOR-VU7Hr1G026690	60	10.12	6.7149	7	Chloroplast	-	-
HvbHLH 120	HOR-VU7Hr1G030250	130	10.11	13.3803	1	Nucleus	-	-
HvbHLH 121	HOR-VU7Hr1G032420	481	4.79	51.47423	12	Nucleus	-	OsbHLH096
HvbHLH 122	HOR-VU7Hr1G038060	86	-	-	2	Nucleus	-	-
HvbHLH 123	HOR-VU7Hr1G046580	331	5.71	35.20252	-	Nucleus	-	OsbHLH127
HvbHLH 124	HOR-VU7Hr1G047180	251	10.31	28.63651	7	Nucleus	-	-
HvbHLH 125	HOR-VU7Hr1G050530	269	-	-	11	Nucleus	AtbHLH130	OsbHLH112
HvbHLH 126	HOR-VU7Hr1G052820	182	9.98	20.4143	11	Nucleus	-	-
HvbHLH 127	HOR-VU7Hr1G052870	383	6.03	41.19198	4	Nucleus	AtbHLH078	OsbHLH091
HvbHLH 128	HOR-VU7Hr1G054880	222	9.2	24.19122	2	Nucleus	AtbHLH037	OsbHLH121

(Table 1) contd….

HvbHLHs	Barley Gene ID	Protein Length (aa)	Isoelectric Point (pI)	Molecular Weight (kDa)	EST Validation	Subcellular Location	*Arabidopsis thaliana* Ortholog	*Oryza sativa* Ortholog
HvbHLH 129	HOR-VU7Hr1G055180	293	6.14	32.12877	2	Nucleus	AtbHLH137	OsbHLH080
HvbHLH 130	HOR-VU7Hr1G062400	147	10.71	16.12702	8	Nucleus	-	-
HvbHLH 131	HOR-VU7Hr1G073310	165	7.15	18.53204	18	Nucleus	-	OsbHLH060
HvbHLH 132	HOR-VU7Hr1G074490	524	5.21	54.2347	5	Nucleus	AtbHLH116	OsbHLH002
HvbHLH 133	HOR-VU7Hr1G077270	279	6.95	30.85244	2	Nucleus	AtbHLH043	OsbHLH124
HvbHLH 134	HOR-VU7Hr1G079250	365	7.08	38.96541	-	Nucleus	AtbHLH098	OsbHLH054
HvbHLH 135	HOR-VU7Hr1G083760	275	5.13	29.89411	1	Nucleus	-	-
HvbHLH 136	HOR-VU7Hr1G087150	525	8.35	55.65763	-	Nucleus	-	OsbHLH030
HvbHLH 137	HOR-VU7Hr1G090520	60	10.12	6.6168	8	Chloroplast	-	-
HvbHLH 138	HOR-VU7Hr1G094690	247	11.03	27.62075	4	Nucleus	AtbHLH060	OsbHLH095
HvbHLH 139	HOR-VU7Hr1G118740	313	5.07	33.87134	10	Nucleus	-	OsbHLH155
HvbHLH 140	HOR-VU7Hr1G118860	70	11.21	7.90544	7	Nucleus	-	-
HvbHLH 141	HOR-VU7Hr1G120420	296	7.22	32.07814	2	Nucleus	-	-

the previous results in *Brachypodium* [8]. It is reported that His-9, Glu-13, Arg-16 and Arg-17 in the basic region of the bHLH domain mainly contribute to DNA binding [36], and Leu-27 and Leu-59 in the first and second helix region are essential for the dimerization [37]. Obviously, the conserved ratio of Lys-36 (81%) within the loop region of barley was significantly higher than that of *Brachypodium* [8], rice and *Arabidopsis* [7].

Based on the criteria developed by Massari [6], the *HvbHLH* proteins were further categorized into two major groups, composing 139 DNA-binding and 2 non-DNA-binding proteins based on 17 N-terminal amino acids within the bHLH domain (Table **S5**). In accordance with the presence or absence of amino acid residues Glu-13 and Arg-16 in the basic region, which were essential for binding E-box [37], the *HvbHLHs* were further divided into 132 putative E-box-binding and 7 non-E-box-binding proteins. The E-box-binding proteins can be further divided into G-box-binding protein (67) and non-G-box-binding protein (65) based on whether His-9, Glu-13 and Arg-17 are found or not in the basic region, which was necessary for recognition of the G-box [38].

3.3. Gene Structure and Conserved Motif Analysis

To understand the phylogenetic relationships of these *HvbHLHs*, an unrooted neighbor-joining tree was generated based on the alignment of 308 bHLH domain sequences, including 141 from barley and 167 from *Arabidopsis*. According to the criteria for bHLH classifications [2, 11], these bHLH proteins were categorized into 34 subfamilies and the *HvbHLHs* were divided into 24 subfamilies, of which two subfamilies (S33, S34) were newly identified in our study (Fig. **2** and **S2**). The number of proteins belonging to different subfamilies varied significantly, of which, the S1 subfamily possessed 18 *HvbHLHs*, whereas S7 only contained one protein (*HvbHLH98*), representing the largest and the smallest group, respectively. The G-box binding proteins were mainly grouped within the subfamilies S2, S5, S7, S10, S11, S12, S13, S14, S24, S25, S26 and S33, whereas the non-G-box binding proteins were mostly clustered into subfamilies S1, S4, S9, S12, S17, S23, S27, S28, S30, S31 and S34. Three subfamilies (S16, S17 and S27) were mainly composed of the non-E-box binding proteins and two subfamilies (S12 and S24) were mainly composed of the non-DNA binding proteins.

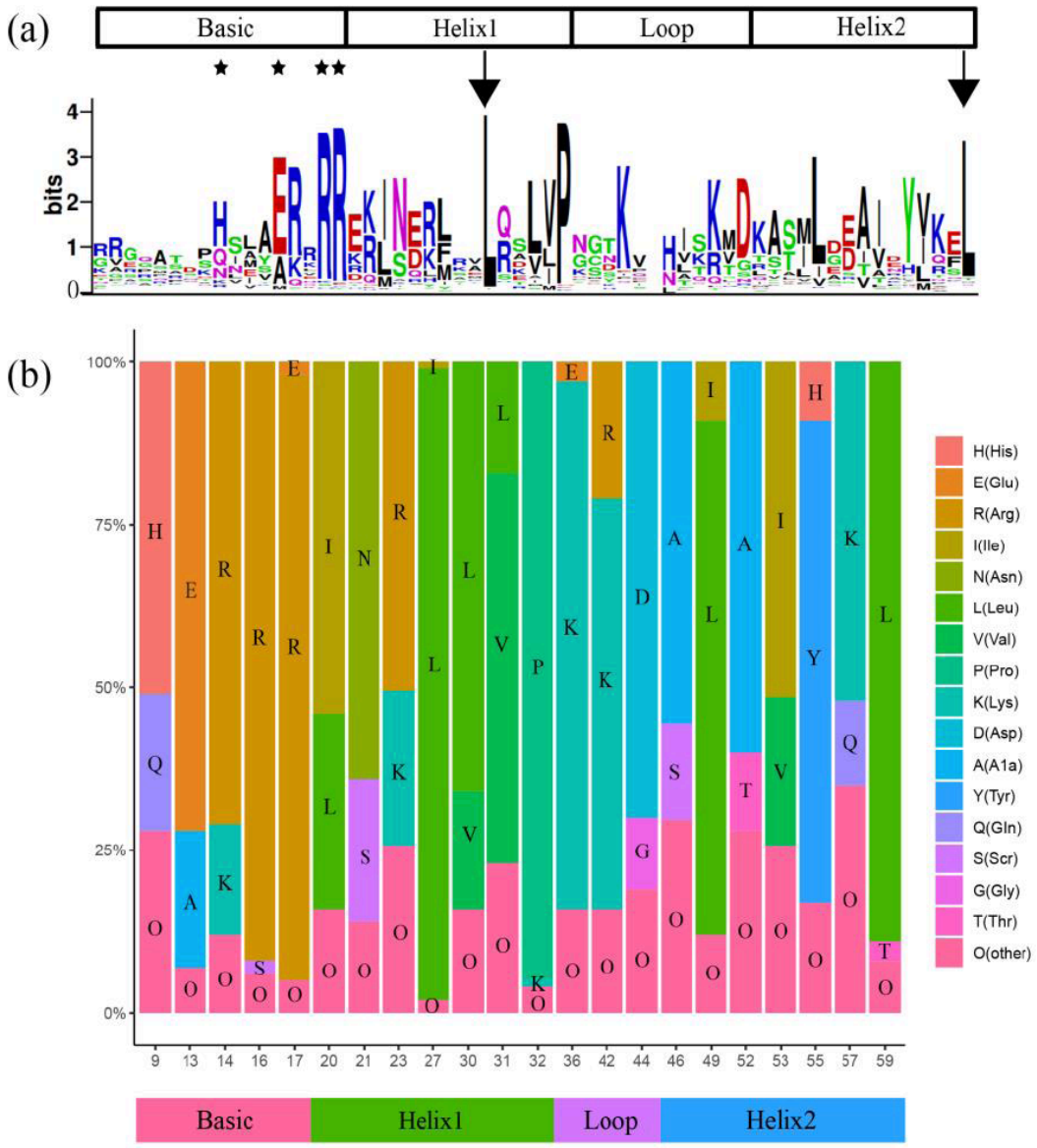

Fig. (1). Sequence characteristics and conserved amino acids organization of the *HvbHLHs*. (**a**): Sequence logo of the *HvbHLH* domain by WEBLOGO. The H9, E13 and R16 amino acids in the basic domain are important for DNA binding as indicated by stars. Amino acids are vital for dimerization of the helix-loop-helix domain as indicated by arrows. (**b**): Distribution of amino acids in the bHLH consensus motif among *HvbHLHs*. The numbers of horizontal ordinate refer to the positions of the residues in the alignments of the studies. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

The exon-intron structure variation could provide important evolutionary clues for gene family [39]. The exon-intron organizations of *HvbHLHs* were also analyzed. The exon numbers ranged from 1 to 12 and exon length varied from 4 bp to 2,053 bp with an average of 219.2 bp. The variations of intron length were significantly higher than those of exon, which ranged from 50bp to 16,093bp with the average of 471.4bp. It is reported that intron gain or loss was influenced by the selection pressure during plant evolution, and intron evolved faster than exon as it can accumulate more mutations with no selective pressure [40]. Notably, the members within the same subfamily in the phylogenetic tree were found to share similar exon-intron structures. Genes within the subfamily S34 showed almost the same gene structure and abundance of intron, which was the most conserved group. We also observed that *HvbHLHs* within the subfamily S2 tended to be intron-less, while subfamily S14 possessed the most abundant introns with more than 9 introns per gene, which had more complicated gene structure patterns than other groups. The various exon-intron structures indicated that the genetic divergence might have occurred in *HvbHLHs* during the formation and evolution of a barley genome.

Additionally, the conserved protein motifs in *HvbHLHs* were identified by the MEME search tool. A total of ten conserved motifs were predicted (Fig. **3**). Among them, motif 3 and 4 were predicted to be located at the C- and N-terminal, whereas other motifs were found to be located at the functional motif regions. Motif 1 and motif 2, located at the bHLH domains, were shared by 134 and 113 proteins,

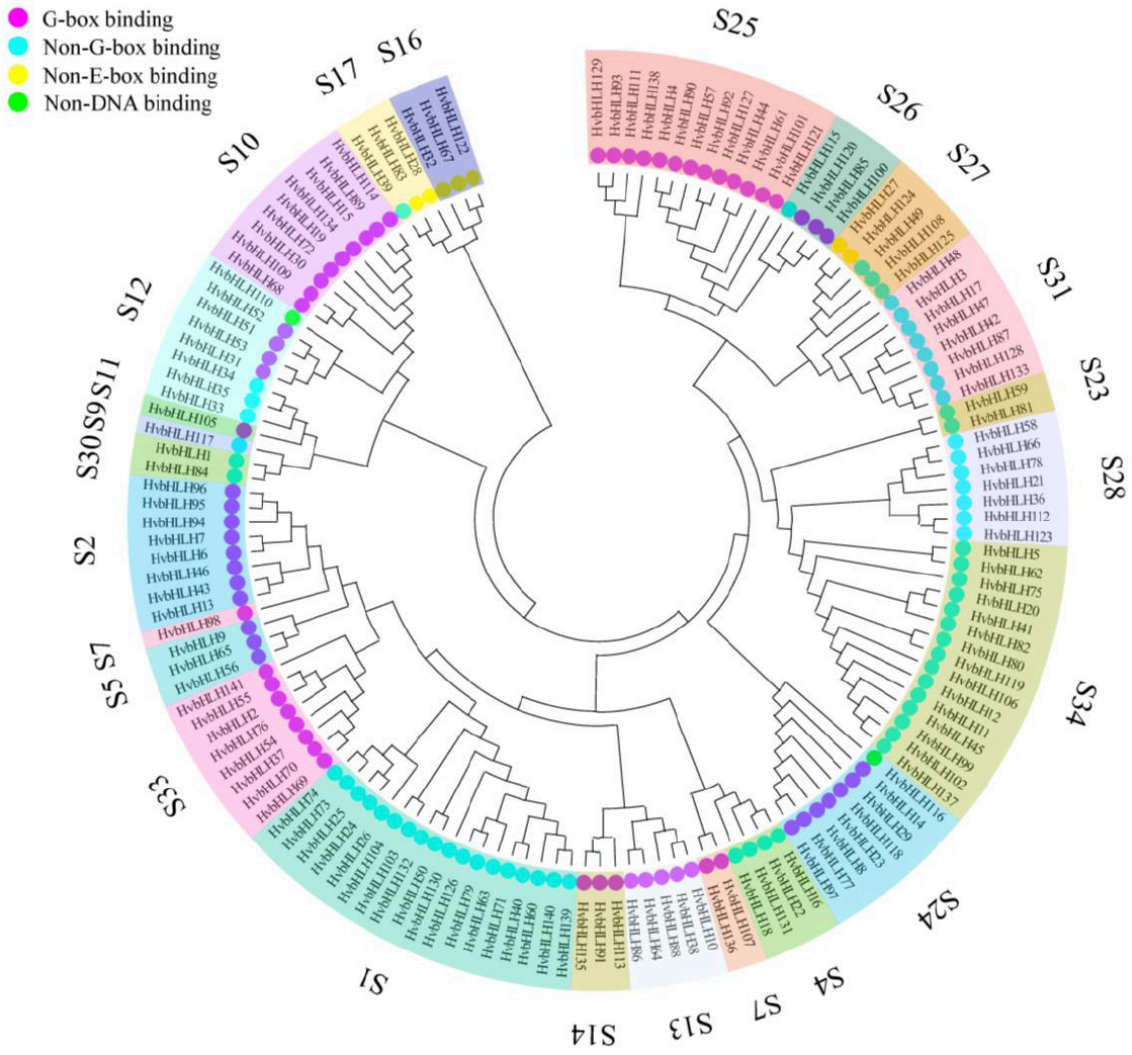

Fig. (2). Phylogenetic tree of *HvbHLHs* based the conserved bHLH domain sequences using the neighbor-joining method. The 24 phylogenetic subfamilies were marked with different colors. The different color on the left side of *HvbHLH* represents the predicted DNA-binding activity of each protein: G-box in purple, Non-G-box in blue, Non-E-box in yellow, Non-DNA in green. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

respectively. Motif 7 was detected in 10 proteins within the subfamily S25. Motif 9 and motif 10 were unique to subfamilies S2 and S25, respectively. Additionally, motif 5 and motif 6 were only found in the subfamily S1.Obviously, the bHLH proteins clustered into the same phylogenetic group displayed the same motif organization, indicating that more similar evolutionary history and biological function were shared by the members within the same subfamilies [41].

3.4. Cis-element and miRNA Target Sites Prediction

In order to investigate the potential regulatory functions of the *HvbHLHs*, the upstream 1.5kb sequences from the transcription initiation sites were extracted to predict the cis-element*s*. Results showed that a total of 40 functional cis-regulatory elements were identified. The ABRE (ABA-responsive, 122 genes), CGTCA-motif (MeJA-responsive, 122 genes), TGACG-motif (MeJA-responsive, 122 genes) and G-box (light-responsive, 123 genes) were found in most of the *HvbHLHs*, while MSA-like domain involved in cell cycle regulation, and the AC-II domain, regulating xylem expression and phloem repression, were only found in 3 and 4 genes, respectively. In addition, we also observed a number of organogenesis-related (including cell cycle, xylem, meristem, endosperm and circadian), hormone-responsive (*e.g.* gibberellin, auxin, salicylic acid and ethylene), light-responsive as well as biotic and abiotic stress-related (*e.g.* wound, low-temperature, drought and anaerobic) cis-element*s* in the promoter regions of *HvbHLHs* (Table **S6**), suggesting the potential functions to regulate diverse stress adaption and signal transduction processes in barley.

MicroRNAs (miRNAs) are a class of endogenous, non-coding RNAs with the size of 21 to 25 nucleotides, which play crucial roles in post-transcriptional regulation of gene expression by guiding target mRNAs' cleavage or translational inhibition [42]. In this study, the putative miRNAs targeted *HvbHLHs* were also analyzed. Results showed that a total of 19 miRNA-bHLH interaction pairs, referring to 15 *HvbHLHs* targeted by 12 miRNAs, were detected (Table **S7**). Most miRNAs silenced the expression of *HvbHLHs* through the transcript cleavage, while *HvbHLH47* and

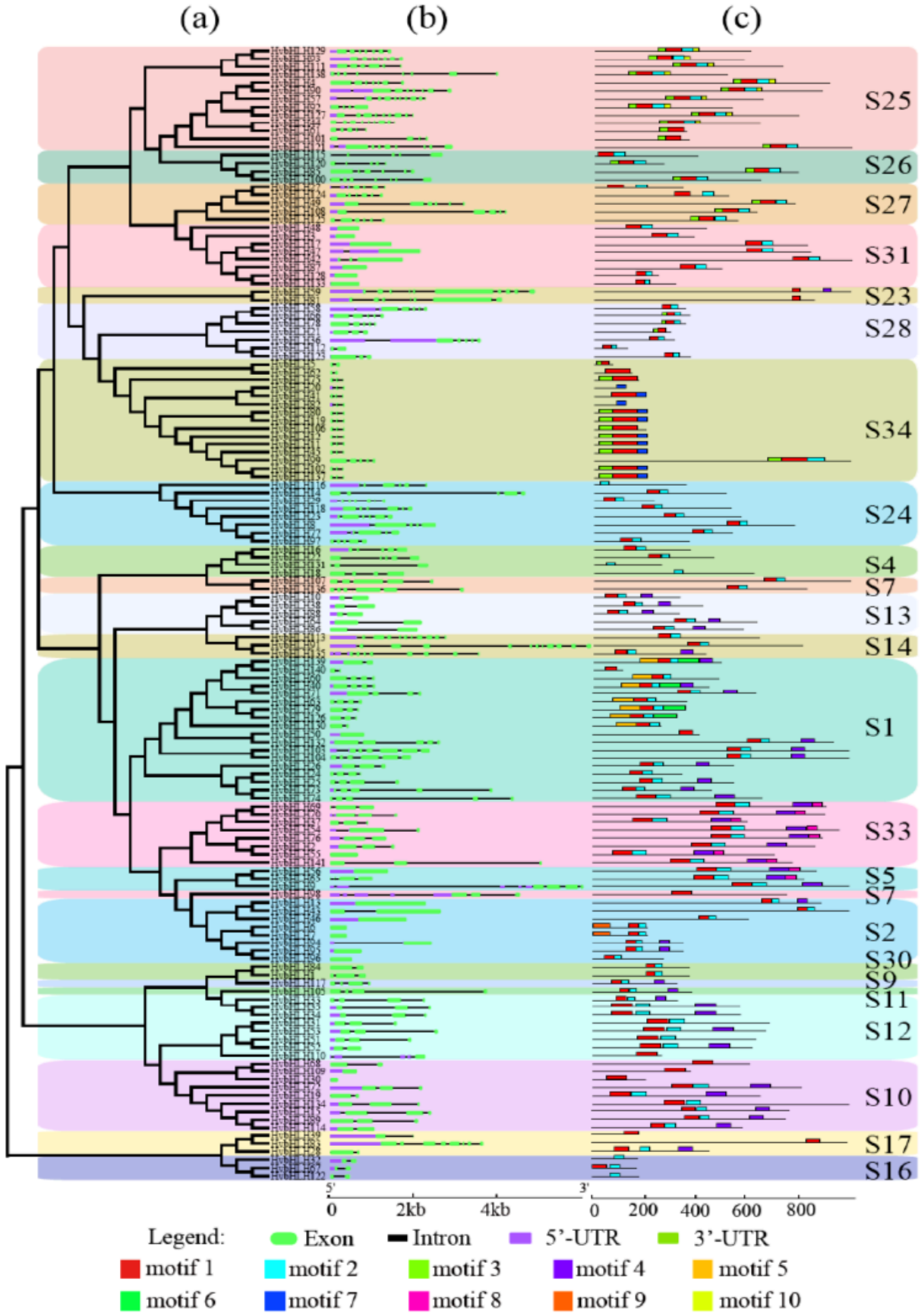

Fig. (3). The gene structure and motif composition of *HvbHLHs*. (**a**): Phylogenetic tree. (**b**): Gene structure. (**c**): Conserved motif analysis. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

HvbHLH26 were predicted to be inhibited by the translation of miRNAs. Furthermore, all the miRNA target sites were found within the CDS region but outside the bHLH domain, of which 10 and 9 pairs were predicted to be targeted upstream and downstream of bHLH domain, respectively. These results indicated that miRNA could be involved in regulating the expression of *HvbHLHs* and further study on the miRNA-mediated post-transcriptional regulation network of *HvbHLHs* will help to understand the roles of *HvbHLHs* in growth and development as well as stress responses of barley.

3.5. Gene Duplication and Synteny Analysis

To get insight into the mechanism of the expansion of *HvbHLHs*, the duplication events were further investigated through a genome-wide synteny analysis. A previous study revealed that some tightly packed bHLH members are most likely to be derived from tandem repeat events in rice and *Arabidopsis* [7]. In this study, a total of 30 gene pairs composed of 46 genes were identified as segmental duplications, whereas no tandem duplication was found, suggesting that segmental duplication was responsible for the expansion of *HvbHLHs* (Fig. **4**). The segmental duplications were mainly found at chromosome 3, 5 and 7. Intriguingly, all of the 15 genes in the subfamily S34 were subjected to segmental duplications, suggesting that there was specific evolution pressure to drive the segmental duplication of these genes. In order to examine the selection effect, the substitution rate of non-synonymous (Ka) *versus* synonymous (Ks) was calculated (Table **S8**), wherein Ka/Ks < 1 means purifying selection, Ka/Ks = 1 means neutral selection and Ka/Ks > 1 means positive selection [43]. The Ka/Ks ratio for duplicated genes of *HvbHLHs* ranged from 0 to 0.4049, with an average of 0.09, indicating that they were subjected to strong purifying selection pressure during the expansion process.

The Ka/Ks of the orthologous genes between the wild and cultivated barley was also calculated to determine the evolutionary relationships of *HvbHLHs*. Using the same method, a total of 142 bHLH genes were identified in wild barley (Table **S9**) and named as *HsbHLH1* to *HsbHLH142*. Through the orthologous analysis (Fig. **S3**), a total of 108 orthologous gene pairs were found (Table **S10**). Crops domestication was the result of a positive selective process, leading to wild species that are exposed to new selective environments associated with human demand [44, 45]. Human artificial selection for crops produced a suite of shared

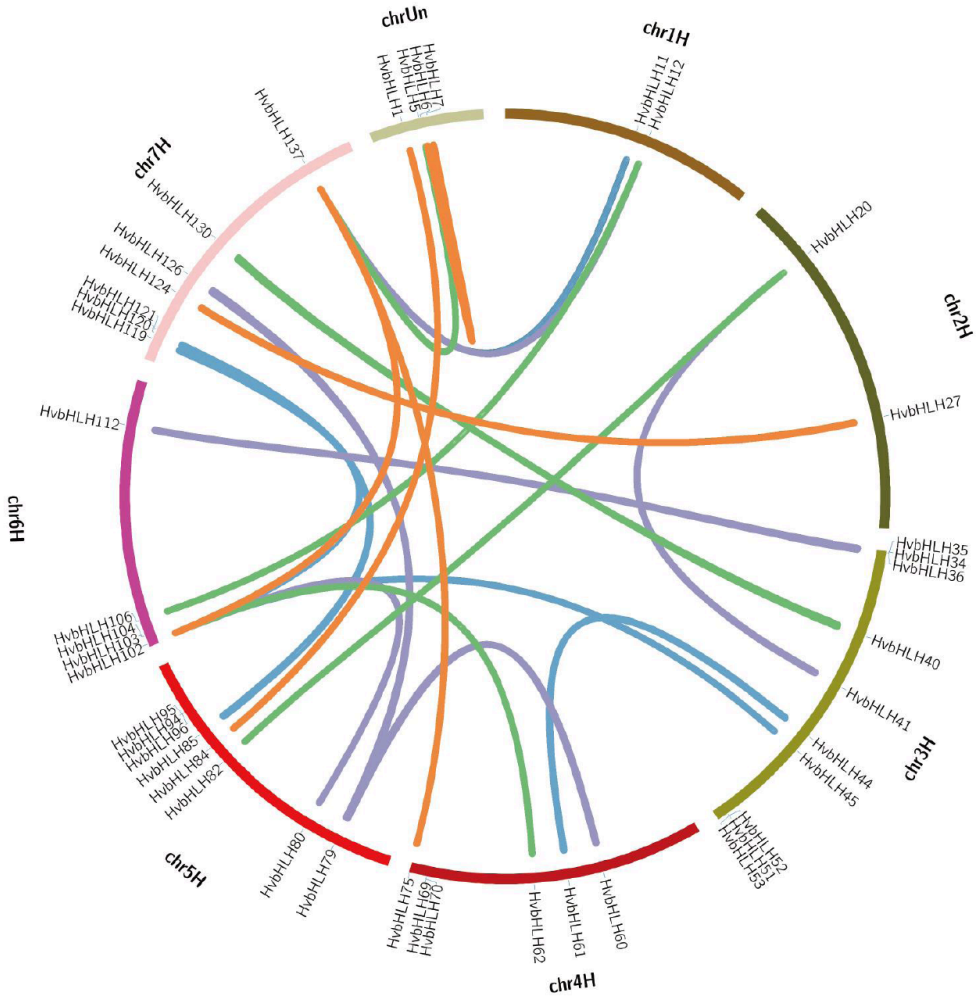

Fig. (4). Genomic distribution of bHLH genes and the gene duplication in barley. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

phenotypic changes in grain crops referred to as 'the domestication syndrome' with various genes [46]. The Ka/Ks ratio between wild and domesticated barley was calculated and the value of 12 *HvbHLH-HsbHLH* gene pairs was found to be more than 1, which was considered to be positive selected genes. Based on the orthologous analysis of *Arabidopsis*, these genes were found to be involved in root initiation [47], cell elongation [48], light signal transduction [49, 50], and cold response [51], which reflected the barley domestication process. The Ka/Ks values of the remaining 96 pairs were found to be lower than 1, indicating that these genes have undergone intense purifying selection during domestication.

Finally, the syntenic relationships with the other six plant species were detected. The whole genome-wide orthologous analysis resulted in 44, 87, 78, 41, 46, and 46 syntenic relationships with *Arabidopsis thaliana*, *Brachypodium distachyon*, *Oryza sativa*, *Solanum lycopersicum*, *Solanum tuberosum*, and *Vitis vinifera,* respectively (Table **S11** and Fig. **S4**). The average value of Ka/Ks was 0.2083 between barley and *Brachypodium*, followed by rice (0.1884), suggesting that these *HvbHLHs* were subjected to extensive purifying selection. It is noteworthy that most of the *HvbHLHs* showed biasness on specific chromosomes of the other six species, indicating that chromosomal rearrangement events like duplication and inversion might predominantly shape the composition and organization of bHLH genes in these genomes.

3.6. Expression Profiles of these *HvbHLHs*

The Spatio-temporal expression profiles of these *HvbHLHs* were investigated at 16 various developmental stages and tissues using publically available RNA-seq samples (Fig. **5**). A total of 67 *HvbHLHs* expressed in at least one stage or tissue with the FPKM value more than 1.0. Among them, 5 genes (*HvbHLH4*, *HvbHLH13*, *HvbHLH16*, *HvbHLH22* and *HvbHLH43*) showed relatively high expression at all of the tested stages and tissues, suggesting that they played indispensable roles in barley growth and development. Remarkably, *HvbHLH22* was stably highly expressed in different organs, suggesting *HvbHLH22* may be considered as a housekeeping gene. The tissue- and stage-specific bHLHs were also identified. *HvbHLH29* and *HvbHLH123* showed preferential expression in embryo while *HvbHLH105* and *HvbHLH72* were found to be mainly expressed in bracts. The ortholog of *HvbHLH105* in *Arabidopsis* is *AtRGE1*. A previous study showed that *AtRGE1* was highly expressed in the endosperm at the heart stage of embryo development and played an important role in controlling embryo growth [52]. Moreover, *HvbHLH48* exhibited the highest expression in the young inflorescences while with extremely low expression levels in other detected samples. Interestingly, its orthologous gene in rice, Lax-panicle (*OsLAX*) also displayed a specific expression in inflorescence, which determined the inflorescence architecture by controlling rachis-branch and spikelet development in rice. It indicated that *HvbHLH48* might share similar functions in barley [53]. Additionally, *HvbHLH134* was also found to be specially expressed in inflorescence, while its ortholog in *Arabidopsis*, SPEECHLESS (*AtSPCH*), is reported to be indispensable for the asymmetric divisions of stomatal lineage, indicating that *HvbHLH134* may have different functions in barley compared to *AtSPCH* [54]. Most of the *HvbHLHs* had diverse tissue-specific expression patterns, highlighting the extensive involvement of *HvbHLH*s in regulating various organs' development.

In order to get preliminary information about the roles of *HvbHLHs* in response to abiotic stresses, the expression patterns of *HvbHLHs* under cold, salt and metal poisoning conditions were also analyzed. Results showed that 49 genes were expressed under cold stress (Fig. **6a**), of which 13 and 7 genes were significantly up-regulated and down-regulated, respectively. *HvbHLH74* showed about 19.7 fold higher expressions under cold stress compared to control. Its orthologous gene in rice, *OsPIL12* had the ability to interact with the biological clock component *OsPRR1* to control circadian rhythms, suggesting this gene may have different functions in barley [55].

The expression profile of bHLH genes under salt stress was further investigated (Fig. **6b**). In response to the high concentration of salt, 12, 10 and 12 genes were identified as up-regulated genes in meristematic root, elongation zone and meristematic zone, respectively, of which, the expression level of *HvbHLH46* was more than 25.22 times higher at the elongation zone and *HvbHLH54* showed 25.07-fold higher expression level at the meristematic zone under salt stress. It is noteworthy that *HvbHLH86* was up-regulated at all tested samples, including meristematic root (6.62 times), elongation zone (3.79 times) and meristematic zones (5.57 times). Previous studies indicated that the knockout lines of *AtbHLH106* (orthologous to *HvbHLH86*) showed more sensitivity to NaCl and over-expression of *AtbHLH106* could enhance the NaCl tolerance in *Arabidopsis* [56]. In addition, we identified 10, 14 and 9 genes to be down-regulated at the meristematic zone , elongation zone and root of meristem. Among them, *HvbHLH59* showed 8.62 and 6.55 times lower expressions at the meristematic zone and elongation zone under salt treatment compared to that of control, suggesting that these genes may act as the negative regulators in response to stress salt stress in barley. Increasing evidence highlights the potential roles of bHLH transcription factors that could be functional as the negative regulators in various biological processes. For example, bHLH transcription factor Paclobutrazol Resistance 6 (*PRE6*) is a transcription repressor that negatively regulates auxin responses in *Arabidopsis* through the direct interaction with Auxin Response Factors 5 (*ARF5*) and *ARF8* [57]. Tanabe *et al.*, also reported the transcription factor *bHLH11*, a negative regulator of Fe homeostasis in plants [58]. Additionally, the bHLH transcription factor *BEE* was found to be a redundant negative regulator involving in *Arabidopsis* drought and salinity tolerance [59]. In summary, *HvbHLH59* may also play a role of a negative regulator to regulate salt stress response in barley.

Finally, the expression patterns of these genes under metal poisoning stress were determined (Fig. **6c**). Results showed a total of 11, 14 and 20 up-regulated *HvbHLHs* that were identified under zinc, copper and cadmium ion stresses, and 17, 22 and 18 down-regulated genes were also found, respectively. It is noteworthy that *HvbHLH26* was differentially expressed under all the 3 stress conditions. Homology analysis revealed that this gene is required for iron deficiency to regulate iron uptake in rice (*OsFIT1*) [60]. Besides, a previous study showed that *OsIRO2* (orthologous to

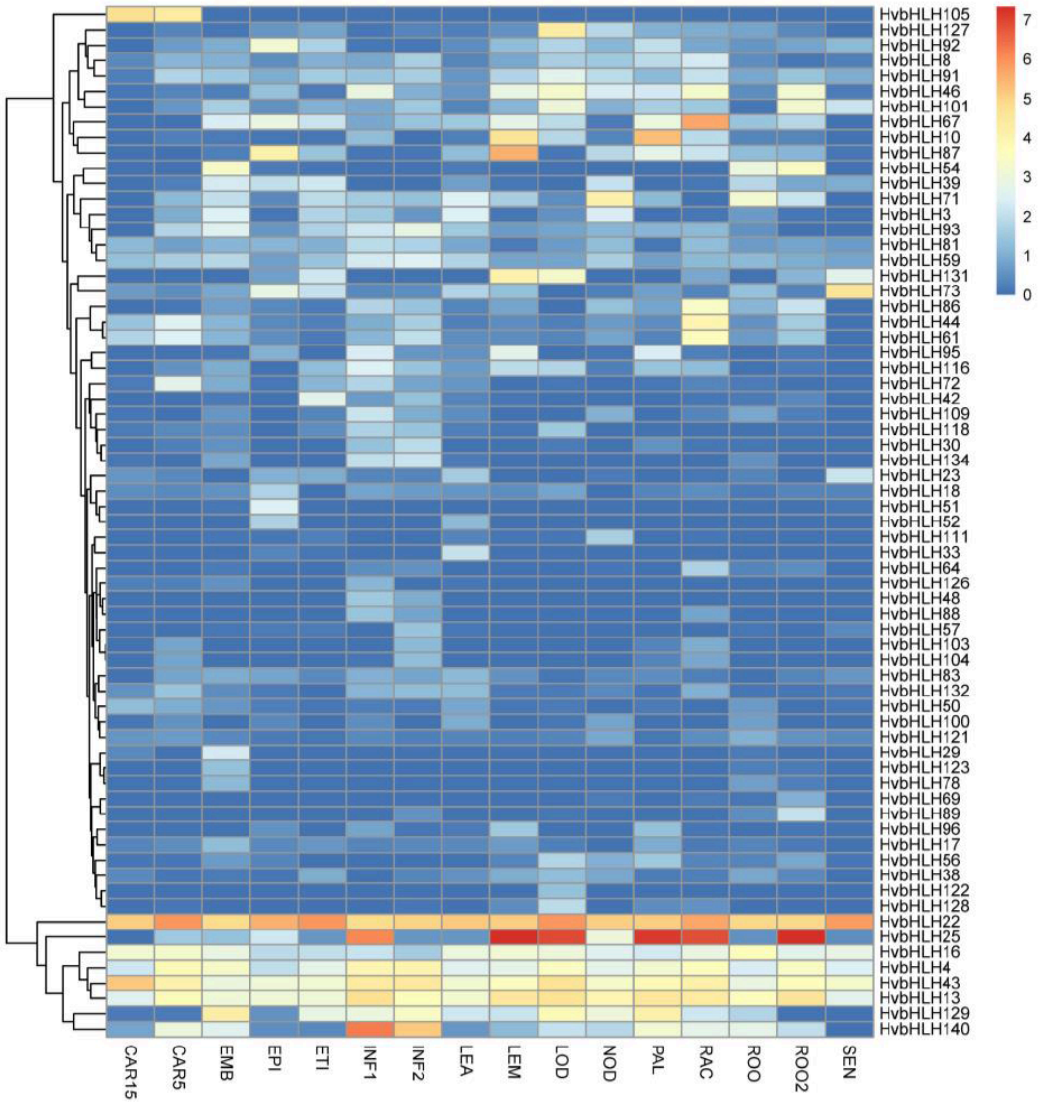

Fig. (5). Hierarchical clustering of expression profiles of *HvbHLHs* across different stages. CAR15: bracts removed grains at 5DPA; CAR5: bracts removed grains at 5DPA; EMB: embryos dissected from 4d-old germinating grains; EPI: epidermis with 4 weeks old; ETI: etiolated from 10-day old seedling; INF1: young inflorescences with 5 mm; INF2: young inflorescences with 1–1.5 cm; LEA: shoot with the size of 10 cm from the seedlings; LEM: lemma with 6 weeks after anthesis; LOD: lodicule with 6 weeks after anthesis; NOD: developing tillers at six-leaf stage; PAL: 6-week old palea; RAC: rachis with 5 weeks after anthesis; ROO2: root from 4-week old seedlings; ROO: Roots from the seedlings at 10 cm shoot stage; SEN: senescing leaf. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

HvbHLH53) was an essential regulator involved in iron uptake under Fe-deficient conditions and it was responsible for maintaining cellular zinc availability in rice [61-63]. In our study, the expression patterns of *HvbHLH53* under zinc and cadmium were 14.11 and 19.17 times significantly higher than that of control.

3.7. Co-expression Network Analysis between bHLHs and Other Genes in Barley

To reveal the regulatory function of bHLHs in relation to the other barley genes, the co-expression network was constructed based on the available 148 RNA-seq samples (Fig. **7** and Table **S12**). The top 5% related gene nodes were filtered based on the WGCNA weighted values and only the correlations among the selected nodes were observed. The co-expression network resulted in a total of 1,324 branches composed of 28 *HvbHLHs* and 279 other genes. Among them, 89 (31.90%) genes were co-expressed with *HvbHLH46*, suggesting the central regulatory role of *HvbHLH46* in the co-expression network. Six *HvbHLHs* (*HvbHLH36, HvbHLH58, HvbHLH78, HvbHLH85, HvbHLH112* and *HvbHLH120*) were co-expressed with *EXP7*. Previous studies reported that *EXPA7* played crucial roles in the regulation of root hair elongation and growth in plants [64]. Other five *HvbHLHs* (*HvbHLH16, HvbHLH23, HvbHLH46, HvbHLH97, HvbHLH116* and *HvbHLH118*) were predicted to be co-expressed with *RGL2*, which has a clear function in the regulation of flower development, ovule number and fertility in *Arabidopsis* [65]. It is noteworthy that *HvbHLH127* was co-expressed with *ARF8*. In Arabidopsis, the module *AtPRE6* (ortholog of *HvbHLH27)-AtARF8*

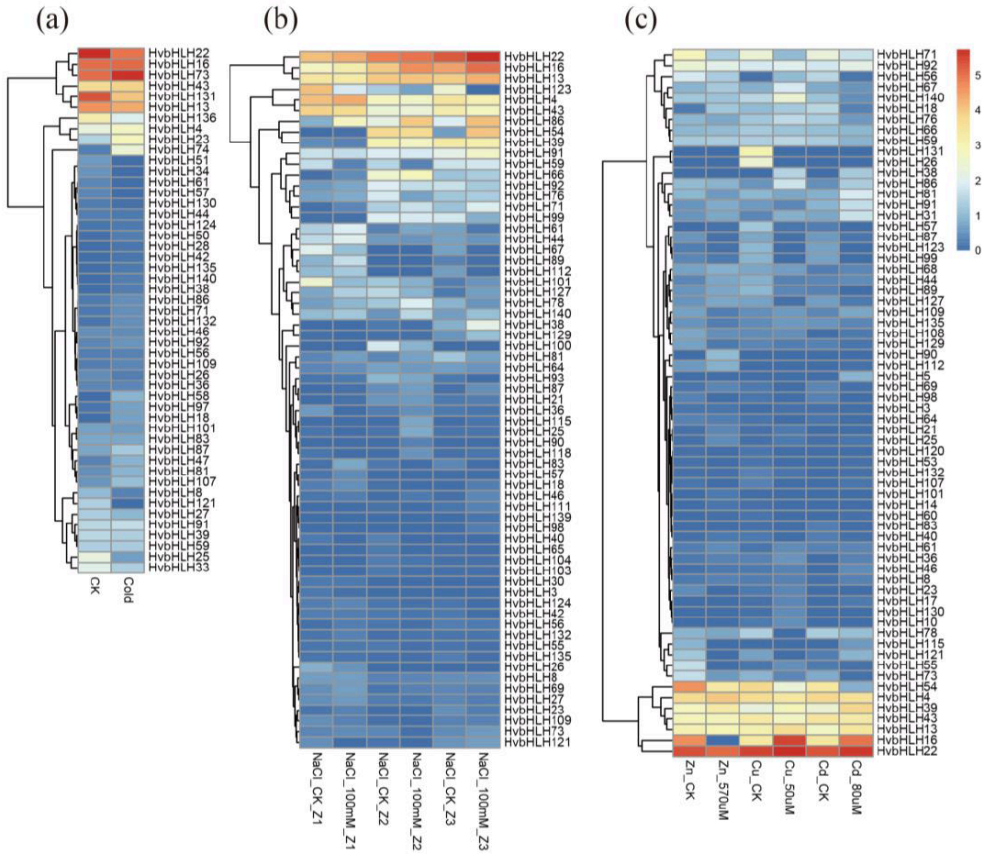

Fig. (6). Hierarchical clustering of expression profiles of barley *HvbHLHs* under five stressed conditions. (**a**): Cold stress; (**b**): Salt stress; (**c**): Zinc, copper and cadmium stress. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

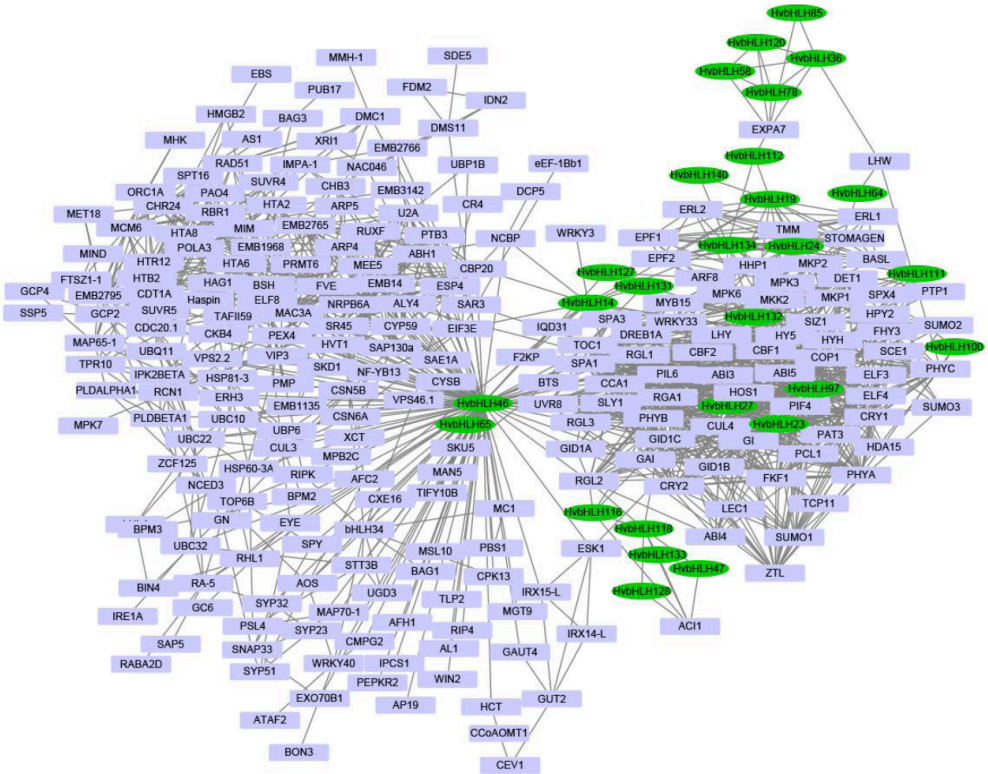

Fig. (7). The co-expression network analysis of *HvbHLHs* with other barley genes. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

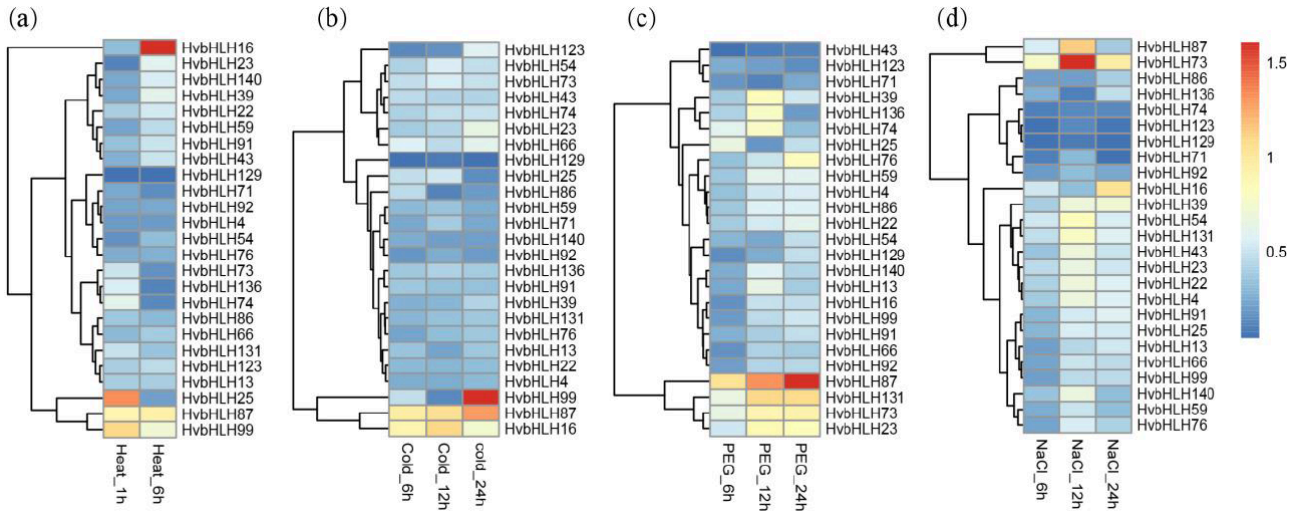

Fig. (8). The expression levels of the 25 *HvbHLHs* under diverse abiotic stresses were investigated using qRT-PCR method. The relative expression levels under: (**a**): Heat treatment; (**b**): Cold treatment; (**c**): Drought treatment and (**d**): Salt treatment. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

has been validated to regulate auxin responses [57], suggesting that *HvbHLH127* could function on *HvARF8* to regulate auxin responses in barley. Remarkably, the gibberellin-insensitive gene *GAI* was predicted to be co-expressed with *HvbHLH14, HvbHLH23, HvbHLH46, HvbHLH97* and *HvbHLH116*. In *Arabidopsis*, a light-labile basic helix-loop-helix protein *PIL5* could increase the expression of *GAI* through directly binding to the promoter region of *GAI* while it did not regulate gibberellin metabolic-related genes in darkness [66]. Moreover, several light responsive motifs such as G-box, GATA-motif, GT1-motif, Sp1, TCCC-motif and TCT-motif were abundantly found within the promoter regions of these genes. Thus, we postulated that these *HvbHLHs* might be functional as the candidate genes in relation to photosynthesis signaling and circadian rhythm management in barley.

3.8. Validation of the Expression of *HvbHLHs* in Response to Abiotic Stresses by qRT-PCR Analysis

To validate the expression of these *HvbHLHs*, 25 stress-responsive candidates revealed by RNA-seq analysis were selected to investigate their expression levels under diverse abiotic stresses by qRT-PCR approach (Fig. **8**). Results showed 11 *HvbHLHs* were up-regulated and 2 *HvbHLHs* were down-regulated under drought stress at all of the 3-time points (6, 12 and 24h). *HvbHLH71* and *HvbHLH129* were up-regulated at 24h, whereas *HvbHLH136* down-regulated at 24h. In addition, the expression patterns seem to link with the cis-elements in the promoter sequences. *HvbHLH39*, *HvbHLH86* and *HvbHLH87* were significantly highly expressed under drought stress, and several MYS elements (MYB binding site involved in drought-induced) were found in the promoter regions of these genes.

Under cold treatment, a total of 7 *HvbHLHs* were up-regulated and 11 *HvbHLHs* were down-regulated at 6, 12 and 24h treatment. The LTR cis-acting element that is involved in low-temperature responsiveness within the promoter regions was also predicted among these differentially expressed genes, suggesting that the cis-elements may have potential regulatory roles in gene-specific expressions. There were 2 *HvbHLHs* (*HvbHLH23* and *HvbHLH136*) that were found to be down-regulated at 6h, but up-regulated at 12h and 24h. Remarkably, the expression levels of *HvbHLH99* and *HvbHLH87* were significantly increased with 20.24 and 10.57 times than control at 24h.

Under heat treatment, only 5 up-regulated and 7 down-regulated genes were detected at 1h and 6h. Notably, the expression of *HvbHLH16* was significantly 21.04 times higher than that of control. Under salt treatment, only four genes (*HvbHLH23, HvbHLH54*, *HvbHLH73* and *HvbHLH131*) were up-regulated at all the points, whereas five genes (*HvbHLH71*, *HvbHLH74*, *HvbHLH86*, *HvbHLH92* and *HvbHLH129*) were found to be down-regulated. In addition, *HvbHLH87* showed a relatively higher expression at 6, 12 and 24h. In addition, a series of *HvbHLHs* were found to be related to multiple stresses, including *HvbHLH99* induced by cold and heat, *HvbHLH73* by salt and drought, *HvbHLH16* by cold, salt and drought, as well as *HvbHLH87* by cold, heat, salt and drought, which could be used as invaluable gene resources for enhancement of barley stress tolerance in future.

Finally, most of the qRT-PCR results were highly consistent with that of RNA-seq. For example, *HvbHLH23*, *HvbHLH73*, *HvbHLH74* and *HvbHLH87* were up-regulated while *HvbHLH131* and *HvbHLH136* were down-regulated in both RNA-seq and qRT-PCR when exposed to low temperature. Additionally, both RNA-seq and qRT-PCR showed that *HvbHLH23* and *HvbHLH66* were up-regulated and *HvbHLH123* and *HvbHLH129* were down-regulated under salt stress. However, some inconsistent results were also observed. RNA-seq data revealed that *HvbHLH86* was significantly up-regulated at three zones of barley root under salt stress, while qRT-PCR showed *HvbHLH86* to be down-regulated at 6, 12 and 24h. We proposed that the inconsistent results were probably caused by different genotypes used, suggesting that some *HvbHLHs* have a tissue-specific expression. The *HvbHLHs* mediated a complicated mechanism to regulate stress response in barley.

CONCLUSION

In this study, we identified a total of 141 *HvbHLHs* and categorized them into 24 subfamilies, as supported by the phylogenetic analysis. Members within the same subfamily shared the similar protein composition and gene structure. Segmental duplications were found to contribute significantly to the expansion of these *HvbHLHs*. Expression analysis revealed that *HvbHLHs* were widely expressed in 16 different tissues and also involved in response to diverse abiotic stresses. Furthermore, 28 *HvbHLHs* were found to be co-expressed with 279 functional genes, which underpinned the regulatory roles of *HvbHLHs*. Finally, the expression of 25 stress-responsive *HvbHLH*s was validated by the qRT-PCR method and some candidates were obtained for further functional study. Taken together, this study provides a solid foundation for further evolutionary and functional investigations of *HvbHLHs* in the future.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The data supporting the findings of the article is available within the article and its supplementary material.

FUNDING

This research was funded by the Jiangxi Natural Science Foundation (Grant No. 20202BAB215002), the Science and Technology Research Project of Jiangxi Provincial Department of Education (Grant No. GJJ180241) and the Open Project Program of the State Key Laboratory of Crop Stress Biology in Arid Areas, Northwest A&F University (Grant No. CSBAA2019001). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

We are grateful to the three anonymous reviewers for their insightful comments. We also thank Dr. Ruimin Li for his constructive comments on data analysis, and we thank the High-Performance Computing Center of Northwest A&F University for providing computational resources in this work.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's website along with the published article.

REFERENCES

[1] Ledent, V.; Vervoort, M. The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis. *Genome Res.*, **2001**, *11*(5), 754-770. https://doi.org/10.1101/gr.177001 PMID:11337472

[2] Carretero-Paulet, L.; Galstyan, A.; Roig-Villanova, I.; Martínez-García, J.F.; Bilbao-Castro, J. R.; Robertson, D.L. Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in Arabidopsis, poplar, rice, moss, and algae. *Plant Physiol.*, **2010**, *153*(3), 1398-1412. https://doi.org/10.1104/pp.110.153593 PMID:20472752

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

[4] Murre, C.; McCaw, P.S.; Baltimore, D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell*, **1989**, *56*(5), 777-783. https://doi.org/10.1016/0092-8674(89)90682-x PMID:2493990

[5] Atchley, W.R.; Terhalle, W.; Dress, A. Positional dependence, cliques, and predictive motifs in the bHLH protein domain. *J. Mol. Evol.*, **1999**, *48*(5), 501-516. https://doi.org/10.1007/pl00006494 PMID:10198117

[6] Massari, M. E.; Murre, C. Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol. Cell Biol.*, **2000**, *20*(2), 429-440. https://doi.org/10.1128/mcb.20.2.429-440.2000 PMID:10611221

[7] Li, X.; Duan, X.; Jiang, H.; Sun, Y.; Tang, Y.; Yuan, Z.; Guo, J.; Liang, W.; Chen, L.; Yin, J. Genome-wide analysis of basic/helix-loop-helix transcription factor family in rice and Arabidopsis. *Plant Physiol.*, **2006**, *141*(4), 1167-1184. https://doi.org/10.1104/pp.106.080580 PMID:16896230

[8] Niu, X.; Guan, Y.; Chen, S.; Li, H. Genome-wide analysis of basic helix-loop-helix (bHLH) transcription factors in *Brachypodium distachyon*. *BMC Genomics*, **2017**, *18*(1), 619. https://doi.org/10.1186/s12864-017-4044-4 PMID:28810832

[9] Wei, K.; Chen, H. Comparative functional genomics analysis of bHLH gene family in rice, maize and wheat. *BMC Plant Biol.*, **2018**, *18*(1), 309. https://doi.org/10.1186/s12870-018-1529-5 PMID:30497403

[10] Zhang, T.; Lv, W.; Zhang, H.; Ma, L.; Li, P.; Ge, L.; Li, G. Genome-wide analysis of the basic Helix-Loop-Helix (bHLH) transcription factor family in maize. *BMC Plant Biol*., **2018**, *18*(1), 235. https://doi.org/10.1186/s12870-018-1441-z PMID:30326829

[11] Ke, Y.Z.; Wu, Y.W.; Zhou, H.J.; Chen, P.; Wang, M.M.; Liu, M.M.; Li, P.F.; Yang, J.; Li, J.N.; Du, H. Genome-wide survey of the bHLH super gene family in *Brassica napus*. *BMC Plant Biol*., **2020**, *20*(1), 115. https://doi.org/10.1186/s12870-020-2315-8 PMID:32171243

[12] Hwang, Y.; Choi, H.S. Tracheophytes contain conserved orthologs of a basic helix-loop-helix transcription factor that modulate ROOT HAIR SPECIFIC genes. *Plant Cell,* **2017**, *29*(1), 39-53. https://doi.org/10.1105/tpc.16.00732 PMID:28087829

[13] Zhu, E.; You, C.; Wang, S.; Cui, J.; Niu, B.; Wang, Y.; Qi, J.; Ma, H.; Chang, F. The DYT1-interacting proteins bHLH010, bHLH089 and bHLH091 are redundantly required for Arabidopsis anther development and transcriptome. *Plant J.,* **2015**, *83*(6), 976-990. https://doi.org/10.1111/tpj.12942 PMID:26216374

[14] Li, S.; Wang, X.; He, S.; Li, J.; Huang, Q.; Imaizumi, T.; Qu, L.; Qin, G.; Qu, L.J.; Gu, H. CFLAP1 and CFLAP2 are two bHLH transcription factors participating in synergistic regulation of AtCFL1-mediated cuticle development in arabidopsis. *PLoS Genet.*, **2016**, *12*(1), e1005744. https://doi.org/10.1371/journal.pgen.1005744 PMID:26745719

[15] Luo, J.; Liu, H.; Zhou, T.; Gu, B.; Huang, X.; Shangguan, Y.; Zhu, J.; Li, Y.; Zhao, Y.; Wang, Y.. An-1 encodes a basic helix-loop-helix protein that regulates awn development, grain size, and grain number in rice. *Plant Cell*, **2013**, *25*(9), 3360-3376. https://doi.org/10.1105/tpc.113.113589 PMID:24076974

[16] Liu, Y.; Ji, X.; Nie, X.; Qu, M.; Zheng, L.; Tan, Z.; Zhao, H.; Huo, L.; Liu, S.; Zhang, B. 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.
https://doi.org/10.1111/nph.13387 PMID:25827016

[17] Seo, J.S.; Joo, J.; Kim, M.J.; Kim, Y.K.; Nahm, B.H.; Song, S.I.; Cheong, J.J.; Lee, J.S.; Kim, J.K.; Choi, Y.D. OsbHLH148, a basic helix-loop-helix protein, interacts with OsJAZ proteins in a jasmonate signaling pathway leading to drought tolerance in rice. *Plant J.,* **2011**, *65*(6), 907-921.
https://doi.org/10.1111/j.1365-313X.2010.04477.x PMID:21332845

[18] Kiribuchi, K.; Sugimori, M.; Takeda, M.; Otani, T.; Okada, K.; Onodera, H.; Ugaki, M.; Tanaka, Y.; Tomiyama-Akimoto, C.; Yamaguchi, T.. RERJ1, a jasmonic acid-responsive gene from rice, encodes a basic helix-loop-helix protein. *Biochem. Biophys. Res. Commun.*, **2004**, *325*(3), 857-863.
https://doi.org/10.1016/j.bbrc.2004.10.126 PMID:15541369

[19] Lau, O.S.; Song, Z.; Zhou, Z.; Davies, K.A.; Chang, J.; Yang, X.; Wang, S.; Lucyshyn, D.; Tay, I.H.Z.; Wigge, P.A.. Direct control of SPEECHLESS by PIF4 in the high-temperature response of stomatal development. *Curr. Biol.,* **2018**, *28*(8), 1273-1280.e1273.
https://doi.org/10.1016/j.cub.2018.02.054 PMID:29628371

[20] Wang, Y.J.; Zhang, Z.G.; He, X.J.; Zhou, H.L.; Wen, Y.X.; Dai, J.X.; Zhang, J.S.; Chen, S.Y. A rice transcription factor OsbHLH1 is involved in cold stress response. *Theor. Appl. Genet.,* **2003**, *107*(8), 1402-1409.
https://doi.org/10.1007/s00122-003-1378-x PMID:12920519

[21] Martín, G.; Rovira, A.; Veciana, N.; Soy, J.; Toledo-Ortiz, G.; Gommers, C.M.M.; Boix, M.; Henriques, R.; Minguet, E.G.; Alabadí, D. Circadian waves of transcriptional repression shape PIF-regulated photoperiod-responsive growth in Arabidopsis. *Curr. Biol.,* **2018**, *28*(2), 311-318.e315.
https://doi.org/10.1016/j.cub.2017.12.021 PMID:29337078

[22] Li, K.; Yu, R.; Fan, L.M.; Wei, N.; Chen, H. DELLA-mediated PIF degradation contributes to coordination of light and gibberellin signalling in Arabidopsis. *Nat. Comm.,* **2016**, *7*, 11868.
https://doi.org/10.1038/ncomms11868 PMID:27282989

[23] Chandler, J.W.; Cole, M.; Flier, A.; Werr, W. BIM1, a bHLH protein involved in brassinosteroid signalling, controls Arabidopsis embryonic patterning *via* interaction with DORNROSCHEN and DORNROSCHEN-LIKE. *Plant Mol. Biol.*, **2009**, *69*(1-2), 57-68.
https://doi.org/10.1007/s11103-008-9405-6 PMID:18830673

[24] Tian, H.; Guo, H.; Dai, X.; Cheng, Y.; Zheng, K.; Wang, X.; Wang, S. An ABA down-regulated bHLH transcription repressor gene, bHLH129 regulates root elongation and ABA response when overexpressed in Arabidopsis. *Sci. Rep.*, **2015**, *5*, 17587.
https://doi.org/10.1038/srep17587 PMID:26625868

[25] Takahashi, Y.; Kinoshita, T.; Matsumoto, M.; Shimazaki, K. Inhibition of the Arabidopsis bHLH transcription factor by monomerization through abscisic acid-induced phosphorylation. *Plant J.,* **2016**, *87*(6), 559-567.
https://doi.org/10.1111/tpj.13217 PMID:27227462

[26] Wang, H.; Li, Y.; Pan, J.; Lou, D.; Hu, Y.; Yu, D. The bHLH transcription factors MYC2, MYC3, and MYC4 are required for jasmonate-mediated inhibition of flowering in arabidopsis. *Mol. Plant*, **2017**, *10*(11), 1461-1464.
https://doi.org/10.1016/j.molp.2017.08.007 PMID:28827172

[27] Cui, Y.; Chen, C.L.; Cui, M.; Zhou, W.J.; Wu, H.L.; Ling, H.Q. Four IVa bHLH transcription factors are novel interactors of FIT and mediate JA inhibition of iron uptake in arabidopsis. *Mol. Plant*, **2018**, *11*(9), 1166-1183.
https://doi.org/10.1016/j.molp.2018.06.005 PMID:29960107

[28] Mascher, M.; Gundlach, H.; Himmelbach, A.; Beier, S.; Twardziok, S. O.; Wicker, T.; Radchuk, V.; Dockter, C.; Hedley, P. E.; Russell, J. A chromosome conformation capture ordered sequence of the barley genome. *Nature*, **2017**, *544*(7651), 427-433.
https://doi.org/10.1038/nature22043 PMID:28447635

[29] Chen, L.; Hu, W.; Tan, S.; Wang, M.; Ma, Z.; Zhou, S.; Deng, X.; Zhang, Y.; Huang, C.; Yang, G. Genome-wide identification and analysis of MAPK and MAPKK gene families in *Brachypodium distachyon*. *PLoS One*, **2012**, *7*(10), e46744.
https://doi.org/10.1371/journal.pone.0046744 PMID:23082129

[30] Ostlund, G.; Schmitt, T.; Forslund, K.; Köstler, T.; Messina, D.N.; Roopra, S.; Frings, O.; Sonnhammer, E.L. InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Res.*, **2010**, *38*(Database issue), D196-203.
https://doi.org/10.1093/nar/gkp931 PMID:19892828

[31] Wang, Y.; Tang, H.; Debarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T. H.; Jin, H.; Marler, B.; Guo, H. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.*, **2012**, *40*(7), e49.
https://doi.org/10.1093/nar/gkr1293 PMID:22217600

[32] Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.*, **2007**, *24*(8), 1586-1591.
https://doi.org/10.1093/molbev/msm088 PMID:17483113

[33] Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. **2016**, *11*(9), 1650-1667.
https://doi.org/10.1038/nprot.2016.095 PMID:27560171

[34] Yue, W.; Nie, X.; Cui, L.; Zhi, Y.; Zhang, T.; Du, X.; Song, W. Genome-wide sequence and expressional analysis of autophagy Gene family in bread wheat (*Triticum aestivum* L.). *J. Plant Physiol.*, **2018**, *229*, 7-21.
https://doi.org/10.1016/j.jplph.2018.06.012 PMID:30025220

[35] Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **2001**, *25*(4), 402-408.
https://doi.org/10.1006/meth.2001.1262 PMID:11846609

[36] Atchley, W.R.; Fitch, W.M. A natural classification of the basic helix-loop-helix class of transcription factors. *Proc. Natl. Acad. Sci. U S A*, **1997**, *94*(10), 5172-5176.
https://doi.org/10.1073/pnas.94.10.5172 PMID:9144210

[37] Ferré-D'Amaré, A.R.; Pognonec, P.; Roeder, R.G.; Burley, S.K. Structure and function of the b/HLH/Z domain of USF. *Embo J.*, **1994**, *13*(1), 180-189.
PMID:8306960

[38] Shimizu, T.; Toumoto, A.; Ihara, K.; Shimizu, M.; Kyogoku, Y.; Ogawa, N.; Oshima, Y.; Hakoshima, T. Crystal structure of PHO_4 bHLH domain-DNA complex: flanking base recognition. *Embo J.*, **1997**, *16*(15), 4689-4697.
https://doi.org/10.1093/emboj/16.15.4689 PMID:9303313

[39] Xu, G.; Guo, C.; Shan, H.; Kong, H. Divergence of duplicate genes in exon-intron structure. *Proc. Natl. Acad. Sci. U S A*, **2012**, *109*(4), 1187-1192.
https://doi.org/10.1073/pnas.1109047109 PMID:22232673

[40] Pyle, A.M. The tertiary structure of group II introns: implications for biological function and evolution. *Crit. Rev. Biochem. Mol. Biol.*, **2010**, *45*(3), 215-232.
https://doi.org/10.3109/10409231003796523 PMID:20446804

[41] Toledo-Ortiz, G.; Huq, E.; Quail, P.H. The Arabidopsis basic/helix-loop-helix transcription factor family. *Plant Cell*, **2003**, *15*(8), 1749-1770.
https://doi.org/10.1105/tpc.013839 PMID:12897250

[42] Bartel, D.P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **2004**, *116*(2), 281-297.
https://doi.org/10.1016/s0092-8674(04)00045-5 PMID:14744438

[43] Lynch, M.; Conery, J.S. The evolutionary fate and consequences of duplicate genes. *Science*, **2000**, *290*(5494), 1151-1155.
https://doi.org/10.1126/science.290.5494.1151 PMID:11073452

[44] Purugganan, M.D. Evolutionary insights into the nature of plant domestication. *Curr. Biol.*, **2019**, *29*(14), R705-R714.
https://doi.org/10.1016/j.cub.2019.05.053 PMID:31336092

[45] Purugganan, M.D.; Fuller, D.Q. The nature of selection during plant domestication. *Nature*, **2009**, *457*(7231), 843-848.
https://doi.org/10.1038/nature07895

[46] Lai, X.; Yan, L.; Lu, Y.; Schnable, J.C. Largely unlinked gene sets targeted by selection for domestication syndrome phenotypes in maize and sorghum. *Plant J.,* **2018**, *93*(5), 843-855.
https://doi.org/10.1111/tpj.13806 PMID:29265526

[47] Schlereth, A.; Möller, B.; Liu, W.; Kientz, M.; Flipse, J.; Rademacher, E.H.; Schmid, M.; Jürgens, G.; Weijers, D. MONOPTEROS controls embryonic root initiation by regulating a mobile transcription factor. *Nature*, **2010**, *464*(7290), 913-916.
https://doi.org/10.1038/nature08836 PMID:20220754

[48] Zhang, L.Y.; Bai, M.Y.; Wu, J.; Zhu, J.Y.; Wang, H.; Zhang, Z.; Wang, W.; Sun, Y.; Zhao, J.; Sun, X. Antagonistic HLH/bHLH transcription factors mediate brassinosteroid regulation of cell elongation and plant development in rice and Arabidopsis. *Plant Cell*, **2009**, *21*(12), 3767-3780.
https://doi.org/10.1105/tpc.109.070441 PMID:20009022

[49] Liu, Y.; Li, X.; Li, K.; Liu, H.; Lin, C. Multiple bHLH proteins form heterodimers to mediate CRY2-dependent regulation of flowering-time in Arabidopsis. *PLoS Genet.*, **2013**, *9*(10), e1003861.

https://doi.org/10.1371/journal.pgen.1003861 PMID:24130508

[50] Castelain, M.; Le Hir, R.; Bellini, C. The non-DNA-binding bHLH transcription factor PRE3/bHLH135/ATBS1/TMO7 is involved in the regulation of light signaling pathway in Arabidopsis. *Physiologia Plantarum*, **2012**, *145*(3), 450-460. https://doi.org/10.1111/j.1399-3054.2012.01600.x PMID:22339648

[51] Petridis, A.; Döll, S.; Nichelmann, L.; Bilger, W.; Mock, H.P. *Arabidopsis thaliana* G2-LIKE FLAVONOID REGULATOR and BRASSINOSTEROID ENHANCED EXPRESSION1 are low-temperature regulators of flavonoid accumulation. *New Phytol.*, **2016**, *211*(3), 912-925. https://doi.org/10.1111/nph.13986 PMID:27125220

[52] Kondou, Y.; Nakazawa, M.; Kawashima, M.; Ichikawa, T.; Yoshizumi, T.; Suzuki, K.; Ishikawa, A.; Koshi, T.; Matsui, R.; Muto, S. RETARDED GROWTH OF EMBRYO1, a new basic helix-loop-helix protein, expresses in endosperm to control embryo growth. *Plant Physiol.*, **2008**, *147*(4), 1924-1935. https://doi.org/10.1104/pp.108.118364 PMID:18567831

[53] Komatsu, M.; Maekawa, M.; Shimamoto, K.; Kyozuka, J. The LAX1 and FRIZZY PANICLE 2 genes determine the inflorescence architecture of rice by controlling rachis-branch and spikelet development. *Develop. Biol.*, **2001**, *231*(2), 364-373. https://doi.org/10.1006/dbio.2000.9988 PMID:11237465

[54] de Marcos, A.; Houbaert, A. A mutation in the bHLH domain of the SPCH transcription factor uncovers a BR-dependent mechanism for stomatal development. **2017**, *174*(2), 823-842. https://doi.org/10.1104/pp.17.00615 PMID:28507175

[55] Nakamura, Y.; Kato, T.; Yamashino, T.; Murakami, M.; Mizuno, T. Characterization of a set of phytochrome-interacting factor-like bHLH proteins in *Oryza sativa*. *Biosci. Biotechnol. Biochem.*, **2007**, *71*(5), 1183-1191. https://doi.org/10.1271/bbb.60643 PMID:17485859

[56] Ahmad, A.; Niwa, Y.; Goto, S.; Ogawa, T.; Shimizu, M.; Suzuki, A.; Kobayashi, K.; Kobayashi, H. bHLH106 integrates functions of multiple genes through their g-box to confer salt tolerance on arabidopsis. *PLoS One*, **2015**, *10*(5), e0126872. https://doi.org/10.1371/journal.pone.0126872 PMID:25978450

[57] Zheng, K.; Wang, Y.; Zhang, N.; Jia, Q.; Wang, X.; Hou, C.; Chen, J. G.; Wang, S. Involvement of PACLOBUTRAZOL RESISTANCE6/KIDARI, an atypical bHLH transcription factor, in auxin responses in Arabidopsis. *Front. Plant Sci.*, **2017**, *8*, 1813. https://doi.org/10.3389/fpls.2017.01813 PMID:29114256

[58] Tanabe, N.; Noshi, M.; Mori, D.; Nozawa, K.; Tamoi, M.; Shigeoka, S. The basic helix-loop-helix transcription factor, bHLH11 functions in the iron-uptake system in *Arabidopsis thaliana*. *J. Plant Res.*, **2019**, *132*(1), 93-105. https://doi.org/10.1007/s10265-018-1068-z PMID:30417276

[59] Moreno, J.E.; Moreno-Piovano, G.; Chan, R.L. The antagonistic basic helix-loop-helix partners BEE and IBH1 contribute to control plant tolerance to abiotic stress. *Plant Sci.*, **2018**, *271*, 143-150. https://doi.org/10.1016/j.plantsci.2018.03.024 PMID:29650152

[60] Colangelo, E.P.; Guerinot, M.L. The essential basic helix-loop-helix protein FIT1 is required for the iron deficiency response. *Plant Cell*, **2004**, *16*(12), 3400-3412. https://doi.org/10.1105/tpc.104.024315 PMID:15539473

[61] Masuda, H.; Aung, M. S.; Kobayashi, T.; Hamada, T.; Nishizawa, N. K. Enhancement of iron acquisition in rice by the mugineic acid synthase gene with ferric iron reductase gene and OsIRO2 confers tolerance in submerged and nonsubmerged calcareous soils. *Front. Plant Sci.*, **2019**, *10*, 1179. https://doi.org/10.3389/fpls.2019.01179 PMID:31681346

[62] Ogo, Y.; Itai, R.N.; Kobayashi, T.; Aung, M.S.; Nakanishi, H.; Nishizawa, N.K. OsIRO2 is responsible for iron utilization in rice and improves growth and yield in calcareous soil. *Plant Mol. Biol.*, **2011**, *75*(6), 593-605. https://doi.org/10.1007/s11103-011-9752-6 PMID:21331630

[63] Ogo, Y.; Itai, R.N.; Nakanishi, H.; Kobayashi, T.; Takahashi, M.; Mori, S.; Nishizawa, N.K. The rice bHLH protein OsIRO2 is an essential regulator of the genes involved in Fe uptake under Fe-deficient conditions. *Plant J.,* **2007**, *51*(3), 366-377. https://doi.org/10.1111/j.1365-313X.2007.03149.x PMID:17559517

[64] ZhiMing, Y.; Bo, K.; XiaoWei, H.; ShaoLei, L.; YouHuang, B.; WoNa, D.; Ming, C.; Hyung-Taeg, C.; Ping, W. Root hair-specific expansins modulate root hair elongation in rice. *Plant J.,* **2011**, *66*(5), 725-734. https://doi.org/10.1111/j.1365-313X.2011.04533.x PMID:21309868

[65] Gómez, M.D.; Fuster-Almunia, C.; Ocaña-Cuesta, J.; Alonso, J.M.; Pérez-Amador, M.A. RGL2 controls flower development, ovule number and fertility in Arabidopsis. *Plant Science: Int. J. Exp. Plant Biol.*, **2019**, 281, 82-92. https://doi.org/10.1016/j.plantsci.2019.01.014 PMID:30824064

[66] Oh, E.; Yamaguchi, S.; Hu, J.; Yusuke, J.; Jung, B.; Paik, I.; Lee, H. S.; Sun, T.P.; Kamiya, Y.; Choi, G. PIL5, a phytochrome-interacting bHLH protein, regulates gibberellin responsiveness by binding directly to the GAI and RGA promoters in Arabidopsis seeds. *Plant Cell*, **2007**, 19,(4), 1192-1208. https://doi.org/10.1105/tpc.107.050153 PMID:17449805