Comprehensive Genome-Wide Identification and Expression Profiling of bHLH Transcription Factors in Areca catechu under Abiotic Stress

Akhtar Ali , Noor Muhammad Khan , Jiang Yi Qi , Guangzhen Zhou , Yinglang Wan *

Posted Date: 24 October 2024

doi: 10.20944/preprints202410.1903.v1

Keywords:

Areca catechu; bHLH gene family; Genome-wide analysis; Expression pattern; Abiotic stresses

Article

Comprehensive Genome-Wide Identification and Expression Profiling of bHLH Transcription Factors in *Areca catechu* under Abiotic Stress

Akhtar Ali [1], Noor Muhammad Khan [1], JiangYi Qi [1], Guangzhen Zhou [2] and Yinglang Wan [2,*]

[1] Hainan Key Laboratory for Sustainable Utilization of Tropical Bioresources, School of Tropical Agriculture and Forestry, Hainan University, Haikou, China

[2] The Ministry of Education Key Laboratory for Ecology of Tropical Islands, Key Laboratory of Tropical Animal and Plant Ecology of Hainan Province, College of Life Sciences, Hainan Normal University, Haikou, Hainan, China

* Correspondence: ylwan@hainanu.edu.cn

Abstract: The basic helix-loop-helix (*bHLH*) transcription factor family, is the second-largest superfamily in plants after MYB, which plays a significant role in the physiological processes of plant growth, tissue development, and environmental variation. However, the systematized study of the *bHLH* transcription factor family has not yet been conducted in *A. catechu*. Herein, we conducted genome genome-wide investigation of *AcbHLH* genes located on 16 chromosomes of *A. catechu*. A phylogenic tree was constructed to adjudge their homology of genes, and 24 subgroups were identified. Finally, we analyzed the dynamic changes of gene-expression levels of nine *AcbHLH* genes in response to drought and salt in leaves and roots. The expression patterns of 9 *AcbHLH* genes show differences in leaves and roots. Under stress conditions induced by salt and drought, *AcbHLH22*, *AcbHLH62* and *AcbHLH45* are significantly upregulated in both leaves and roots. In conclusion, this study will substantially contribute to the foundation for exploring the role of the *bHLH* superfamily in *A.catechu* in dealing with abiotic stresses.

Keywords: *Areca catechu*; *bHLH* gene family; genome-wide analysis; expression pattern; abiotic stresses

1. Introduction

Transcription factors (TFs) play a pivotal role in regulating plant growth and development by controlling gene expression [1,2]. They influence various cellular processes by interacting with other proteins involved in transcription [3]. The *bHLH* is the second-most abundant group of TFs family found in plants, animals, and microorganisms within this group of proteins [4]. The *bHLH* transcription factors domain consists of 50-60 amino acids in length and is divided into two functional regions, The N-terminal basic domain and the C-terminal (HLH) domain region [5]. The N- terminal basic region consists 10-15 bp amino acids and is responsible for binding to cis-elements E-box (CANNTC), while the C-terminal region consists of 40-50 bp amino acids and is responsible for the formation of homo and heterodimer protein complexes [6]. Moreover, *MYC*-like *bHLH* proteins possess MYB-interacting region and extra N-terminal MYC domain, enabling attachment to *bHLH* and R2R3-*MYB* domain proteins, respectively [7].

Many physiological and biochemical processes in plants have been linked to bHLH TFs, as shown by the extensive study conducted on these proteins. Key developmental processes such as seed coat differentiation [8], stomata differentiation [9], trichome/root hair creation [10], fruit development [11], and carpel edge development [12] re among those regulated by transcription. Certain *bHLH* TFs also have a role in premature seedling photo morphogenesis [12]. Additionally, growing body of research indicates the involvement of *bHLH* proteins in plant actions to a heterogeneity of abiotic stressors, for instance salinity, drought, low temperature, and mechanical injury [13–16]. In addition, bHLH proteins have been extensively studied in plants for their

 doi:10.20944/preprints202410.1903.v1

regulatory functions in the formation of secondary metabolites, including alkaloids [17], flavonoids [18], phenolic acids [19], and terpenoids [20].

The advancement of sequencing technologies facilitated the recognition and characterization of numerous families *(bHLH)* in plant genomes. Genome-wide analyses of *bHLH* family have been identified in various plants, including *Arabidopsis* [21] , rice [22], the tomato [23], maize [24], wheat [25], poplar (*Populus* sp.) [26], potato [27], and other plants based on genome sequences [24]. However, the bHLH transcription factors in Areca palms (*Areca catechu L.*) have not yet been identified, and it is unclear how they are expressed under abiotic stress conditions.

Areca catechu L., is a vital evergreen vascular monocot tree belonging to the *Areceae* family. *A. catechu*., possesses medicinal and economic significant importance, and extensively cultivated in China, India, Thailand, Indonesia, Malaysia, and Cambodia [28,29]. In China, areca nut is utilized as a component in traditional Chinese medicine. Due to its high arecoline, areca nut has rise as significant cash crop in Southeast Asia and Africa, as the fourth most addictive substance in the world after alcohol, caffeine, and nicotine. Additionally, the study of *bHLH* protein *A. catechu*, a rare monocot tree, not only carries substantial economic value but also holds theoretical importance in deciphering abiotic stress responses within this plant species [30]. Although, A comprehensive study of *AcbHLH* gene expression patterns within the *Areca* genome remains an understudied area within the scientific literature. In this research, we categorized 76 *AcbHLHs* genes from the *A. catechu* genome. The comprehensive analyses were conducted on their phylogenic relationships, sequence characteristics, gene structures, promoter sequences, and collinearity. Furthermore, investigated specific expression of these genes across various parts in response to drought and salt stresses. This work provides a base for future research on molecular mechanisms of *AcbHLH* genes in response to abiotic stresses in *A. catechu*.

2. Results

2.1. Identification of bHLH genes in A.catechu

The present study identified and characterized 76 *AcbHLH* genes in the *A. catechu* genome.

The gene names were assigned from *AcbHLH01* to *AcbHLH76* based on phylogenic connection with already known *A.thaliana bHLH* genes (162) (Table S3). Additionally, characterized these genes provide information on their molecular weight, isoelectric points, protein length, domain composition and subcellular localization (http://cello.life. nctu.edu.tw/) (Additional file 1: Table S1). Among 76 *AcbHLH* proteins, the amino acid lengths range from 149 amino acids (*AcbHLH17*) up to 685 amino acids (*AcbHLH59*), while the moderate range of amino acids were 363.079 kDa.The molecular mass of the proteins ranged from 16.28 kDa (*AcbHLH17*) to 76.32 kDa (*AcbHLH59*), whereas their molecular mass ranged from 16.28 kDa (*AcbHLH17*) to 76.32kDa (*AcbHLH59*). The calculated isoelectric point (pI) of *AcbHLH25* started from 4.99 to 10.3 (*AcbHLH57*), with a mean of 6.96 pI. of *AcbHLH* genes. The subcellular localization prediction result revealed that 70 *AcbHLHs* are present in the nucleus, cytoplasm 01, chloroplast 04, and mitochondria 01 (Additional file 1: Table S1).

2.2. Phylogenetic, Multiple Sequence Alignment, and classification of AcbHLH genes

The phylogenetic relationships among 76 *AcbHLH* and 162 *AtbHLH* proteins were inferred using the Neighbor-Joining (NJ) method with 1000 bootstrap replicates. The analysis was based on amino acid sequence datas (Figure 1; Additional file 1: Table S1). *bHLH* genes were classified into 24 groups based on topological structures. According to Pires and Gabriela's presented classification method and topological outline of the tree, the phylogenetic tree of 238 *bHLH* genes was classified into 24 groups(1-24)[31]. The absence of subgroups 01 and 23 in the *A.catechu bHLH* family specifies that it loss or undifferentiation during the Evolutionary development of *A. catechu*. Amongst the 24 subgroups of *A. catechu*, subgroup 15 had the most members (09 *AcbHLHs*, 11 *AtbHLHs*), while subgroups 3, 11, 12 and 22 had the least members (01 *AcbHLH*). The phylogenic tree shows a closed relationship between the *AcbHLHs* and *AtbHLHs* proteins with a bootstrap support of 80 or higher. The Synteny study of *A. catechu* and *A. thaliana*, *AcbHLHs* and *AtbHLHs* showed homologous, and indicated have

 doi:10.20944/preprints202410.1903.v1

similar functions. The *A. thaliana bHLH* domains and *A. catechu* family were selected and sequenced (Figure 2).

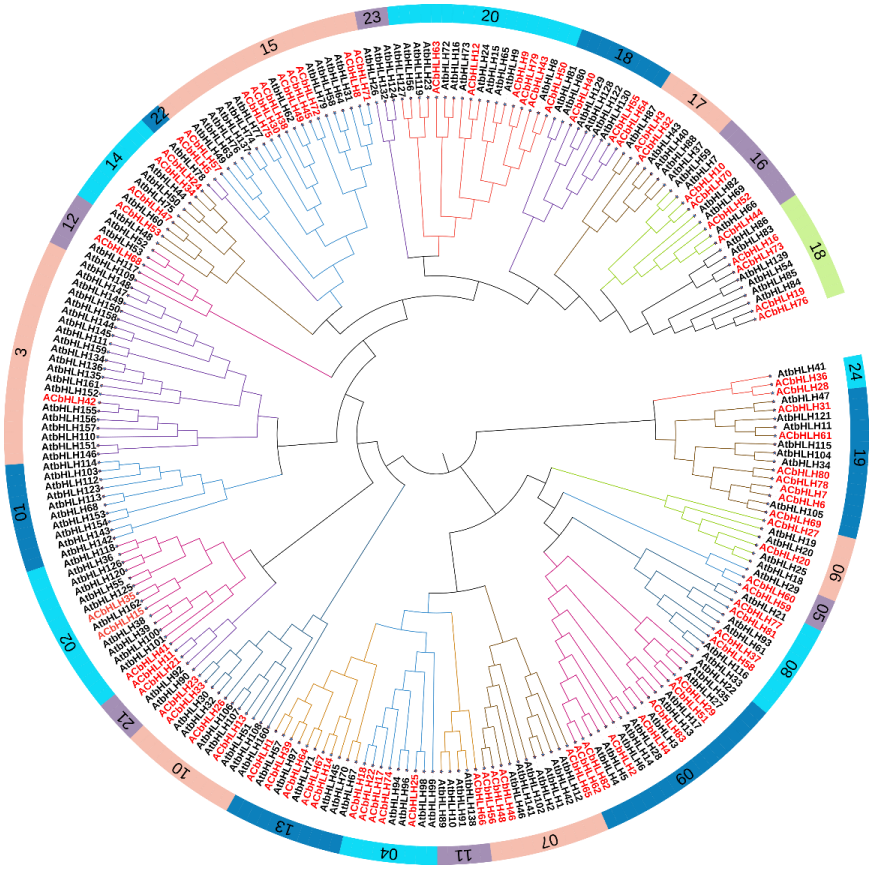

Figure 1. A phylogenic tree illustrates the relationship among *bHLH* domains of *A. catechu* and *A.thaliana*. The black colour presents (*AtbHLH) A. thaliana* and red represents *AcbHLH* of *A.catechu bHLH* protein.

In *A. catechu* mostly *bHLH* domains are about 53 amino acids, basic region consists of 17 amino acids, Helix region 15 and loop region 6 [32]. The conservative domains of the *A. catechu bHLH* family have shown significantly increased amino acid, especially in Helix region. The *bHLH* domain showed the greatest sequence variation in subgroups 15 and 11, In other plants, the same characteristic is true for *bHLH* protein, such as *A. thaliana* [32] and [33], *F.tataricum* [34] and *S. lycopersicum* [35].

 doi:10.20944/preprints202410.1903.v1

Figure 2. Multiple sequence alignments of *A. catechu* and 24 subgroups *A. thaliana*. *A. catechu* is devoid of subgroup 24. The location and boundaries of the *bHLH* domain are indicated at the top of each subgroup.

2.3. Analysis of gene structure and conserved domains of AcbHLH genes

To reveal the *AcbHLH* genes' intron-exon structures, we further compared 76 *AcbHLH* genes, ranking from 1 to 13, showed differences in exon-intron structures (Figures 3A, 3B; Tables S1). Out of the 76 *AcbHLH* genes studied, 2 genes (2.63%) possess a single exon, while the rest genes exhibit 2 or more exons in their structures. Two intron-free genes were identified within Subgroups 2 and 6. Among the remaining 69 genes, three intron patterns were most prevalent. *AcbHLH40* exhibited the highest number of introns (13), while genes in Group 3 contained either 0 introns or a single intron. Group 3 exhibited the greatest variant in the number of exons, ranged from a single intron in *AcbHLH22* to 18 introns in *AcbHLH40*. As revealed by comparative analysis Subgroup 3 among the analyzed *AcbHLH* genes exhibited the most diverse range of intron counts.

 doi:10.20944/preprints202410.1903.v1

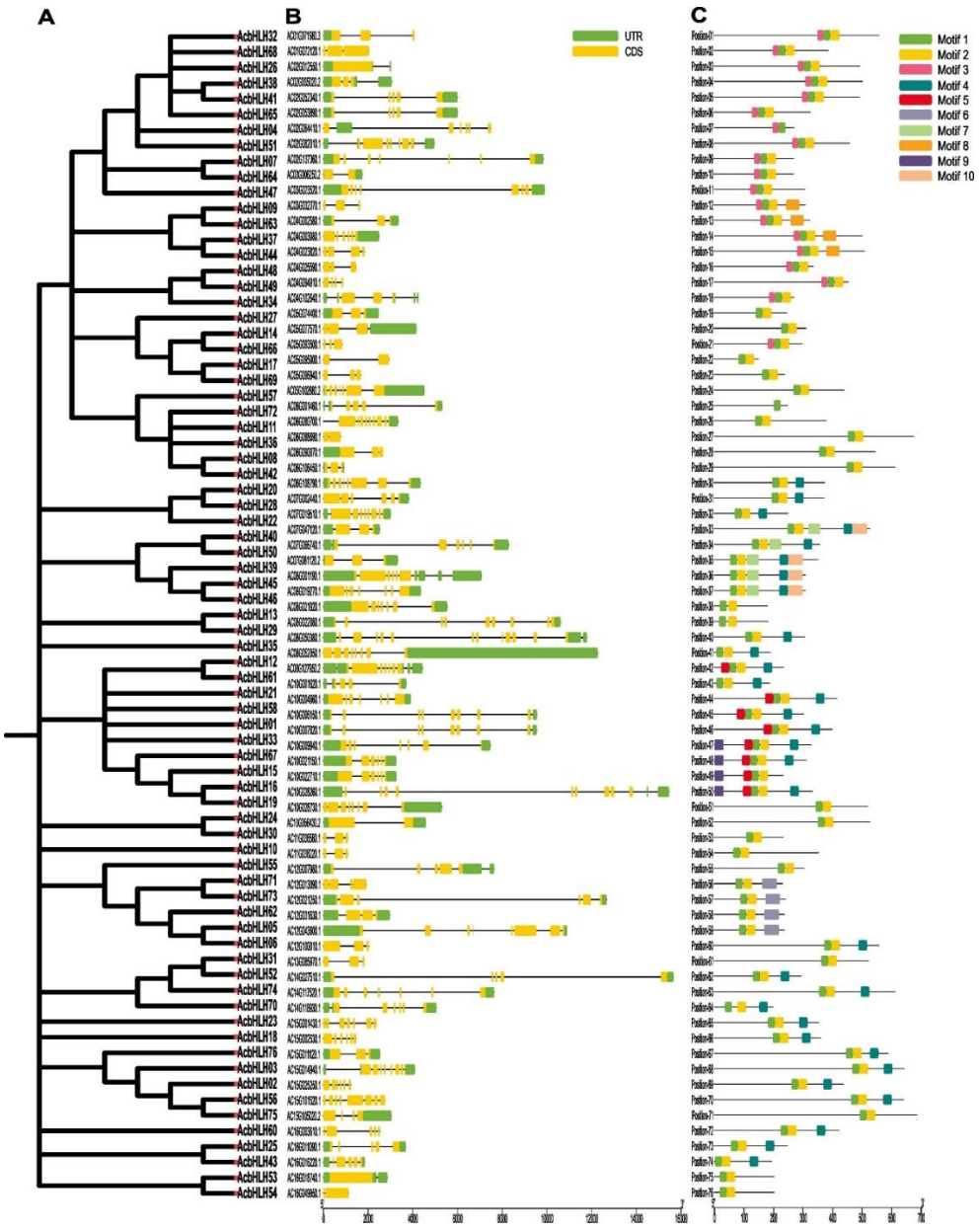

Figure 3. Comparative Analysis of *A. catechu AcbHLH* Gene Phylogeny, Structure, and Motifs. A) Phylogeny inferred based on NJ method with 1000 bootstrap replicates. B) Introns and exons are visually represented as yellow and black lines. C) Amino acid motifs (1-10) indicated by colored, relative protein lengths representing with black lines.

Distribution analysis of motifs among *AcbHLH* proteins revealed that motifs one, two and three were prevalent and abundant in all groups, while motifs 6, 8, 9 and 10 exhibited the least across the *AcbHLH* proteins. *AcbHLH* subgroups are characterized by distinct motif compositions. For instance, Motifs 1, 2, and 3 are commonly found in Subgroup 1 and 2 proteins. Subgroup 4 proteins, on the other hand, are characterized by the presence of motifs 1, 2, 5, and 9. Some motifs were notably prevalent within particular subgroups, such as motifs 8, 10 and 6 in subgroups 2, 3 and 7 respectively. Motifs 1 and 2 were mostly observed across multiple subgroups. Additional identified distinct patterns in the arrangement and relative positions of the motifs. For instance, motif 4 exclusively appears at the end of subgroups 3, 5 and 6, while a specific arrangement has been observed in motifs 1, 2 and 10.

 doi:10.20944/preprints202410.1903.v1

2.4. Characterization of cis-acting elements within AcbHLH promoter regions

To determine the functional activities and promoter regions of *AcbHLH* genes, spanning 2000 bp upstream coding sequence, were evaluated in the presence of potential cis-elements using PlantCARE (Figure 4). The promoter regions of *AcbHLH* genes consists a significant number of stress-related cis-elements, suggesting their crucial involvement in stress response pathways. These revealing indicate that the genes intricate in regulating the plant's response to environmental changes. A thorough investigation of these genes may reveal into the mechanisms employed by plants to adapt to their surrounding stimuli. Additionally, we identified cis-elements associated with abscisic acid (ABA) and salicylic acid (SA) within promoter regions *AcbHLH* genes convoluted in abiotic stress responses, which include intense temperature, wounding, and drought. These elements were found to be associated with several phytohormones, such as salicylic acid, gibberellin, methyl jasmonate (MeJA), auxin and ethylene (Figure 4B).

The heat map categorized into 4 discrete portion: growth, phytohormone regulation, light-responsiveness, and stresses, features of *AcbHLH* genes on the rows and cis-regulatory elements on the columns. The bar graph, with colours ranging from yellow to dark red, visually represents the quantitative values of elements within each row and column.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 October 2024 doi:10.20944/preprints202410.1903.v1

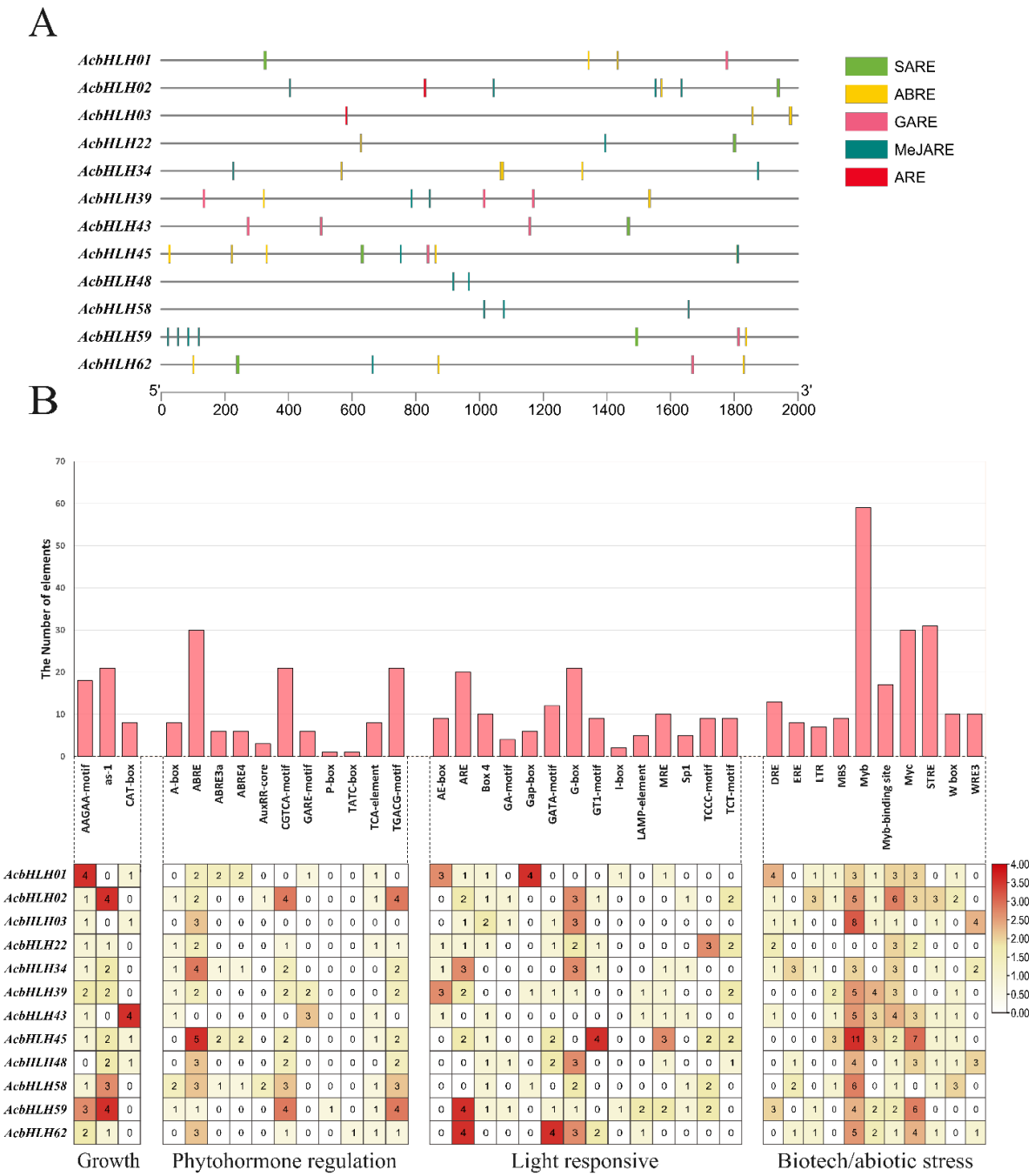

Figure 4. Cis-Regulatory Elements in the 2000 bp Upstream Region of AcbHLH Gene Promoters. Coloured rectangles visually depict the various cis-acting elements.

2.5. Chromosomal distribution and gene duplication of AcbHLH genes

The latest *A. catechu* genome database enabled the creation of a precise physical map that offers the specific positions of *AcbHLH* genes within the genome (Figure 5, Additional file 3: Table S3). The allocation of 76 *AcbHLH* genes across Chr 1 to 16 displayed uneven distribution patterns (Figure 4). The name of each *AcbHLH* gene was assigned based on its sequential physical location along *A. Catechu* Chr 1-10, from top-to-bottom. Chr10 possess greatest number of *AcbHLH* genes (10 genes, ~13.16%), followed by Chr2, 8 and 16 each contained (7 respectively, ~ 9.21%), while Chr13 contained the least (1, ~ 3.31%) and Chr9 contain Zero gene. Chr4, Chr5, Chr6 and Chr12 each contained 6 (~ 7.9%) *AcbHLH* genes. Chr7 and Chr15 each contained 5 (~ 6.6%) *AcbHLH* genes. Chr3, Chr14, contained 3 (~3.94%) and chr11 (2~ 2.6%) *AcbHLH* genes. In addition, we discussed the chromosome of *AcbHLH* gene-duplication instances. The chromosomal region spanning 30 Mb and 60 Mb

 doi:10.20944/preprints202410.1903.v1

exhibiting two homogenous genomic segments are clarified as a tandem duplication event [36]. On chromosomes 8 and 11, we observed 04 tandem duplication segments involving 2 genes *AcbHLH* (Figure 4). *AcbHLH38* and *AcbHLH41* on chr 8, *AcbHLH53* and *AcbHLH54* on chromosome 11, each had tandem repeat events, while Chr10 have 2 tandem duplication events involving 4 *AcbHLH* genes, (*AcbHLH45*, *AcbHLH46*, *AcbHLH48* and *AcbHLH49*). The gene involved in the tandem repeat occurrences originated exclusively within identical subfamilies. For instance, genes *AcbHLH38* and *AcHLH41* tandem repeat sequences, and clustered together in subfamilies 15 and 2, *AcbHLH53* and *AcbHLH54* were clustered together in subfamily 14 and 18, while *AcbHLH45*, *AcbHLH46*, *AcbHLH48* and *AcbHLH49* were clustered together in subfamily 15, 07, 07 and 15 respectively. (Figure 5B, Additional file 3: Table S3).

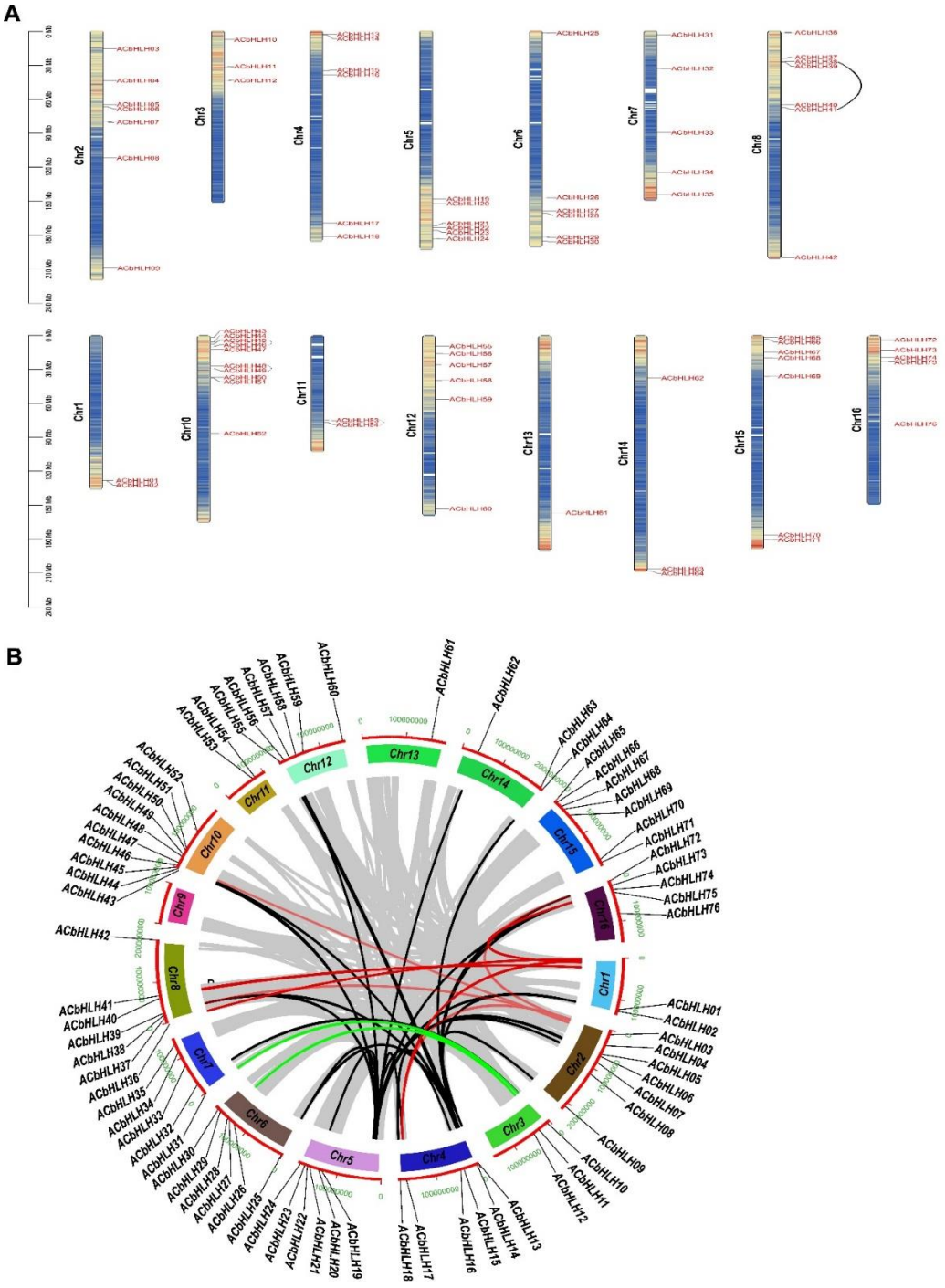

Figure 5. (A) Location of 76 *AcbHLH* genes across 16 *A. catechu* chromosomes. The left-hand scale indicates chromosomal length. (B) The schematic diagram represented the distribution of *A. catechu*

 doi:10.20944/preprints202410.1903.v1

chromosomes and interchromosomal interaction. Distinct coloured lines within the diagram represent gene pairs. Red lines indicate *AcbHLH* gene pairs, *A. catechu* are labeled outside the chromosome circles while chromosome numbers are indicated within.

Additionally, a total 23 segmental duplication identified within the *AcbHLH* gene family. As illustrated in Figure 5, the *AcbHLH* gene family contains 10 (13.15%) paralogs, indicating a common evolutionary origin of these *AcbHLH* members. The distribution of *AcbHLH* genes across the 16 linkage groups (LGs) of A. catechu was notably uneven (Figure 4). Some linkage groups, specifically LG10 and followed by LG8, exhibited a higher count of *AcbHLH* genes compared to other linkage groups. LG2 boasted the highest count of *AcbHLH* genes 14, while LG5 contained the fewest count of one *AcbHLH*. Further analysis of *bHLH* families, observed that many are significantly linked within their respective subfamilies, except *AcbHLH61* / 01 and *AcbHLH062* / 42. Within all identified *AcbHLH* genes, group 15 exhibited the highest count of linked genes, encompassing 9/76 genes.Additionally, group 15 have 9 genes, while groups 22, 12, 11 and 03 have only 1, and groups 23 and 01 have no genes (Additional file 4: Table S4). The presence study implies genes *AcbHLH* indicates gene duplication events, and have been a significant factor in the *AcbHLH* genes in *A. catechu*, resulting in emergence of novel functions and the expansion of the *AcbHLH* gene family.

2.6. Analysis of gene duplication events of AcbHLH genes

To investigate the possible driving force for the diversifications of *AcbHLH* gene family, the Dup-206 Gene finder used to examine gene duplication segments inclusive of DSD, WGD, TRD, TD and PD. The result revealed that there was high variation in the number of duplicated genes and the distribution of their protein. Furthermore, the largest number of duplicated genes at 23 in DSD, while the lowest (4) were observed in PD. Additionally, 11 duplicated genes were identified in WGD (Figure 6 and Table S9). To calculate the value of (Ks) and (Ka) substitute rate, and Ka/Ks ratio of the duplicated genes over five replication events, to provide insight into the selection pressure on *AcbHLH* genes. The Ka/Ks values less than 1,

Indicating that clear selection of these genes has been subjected, As shown in Figures 6B to F.

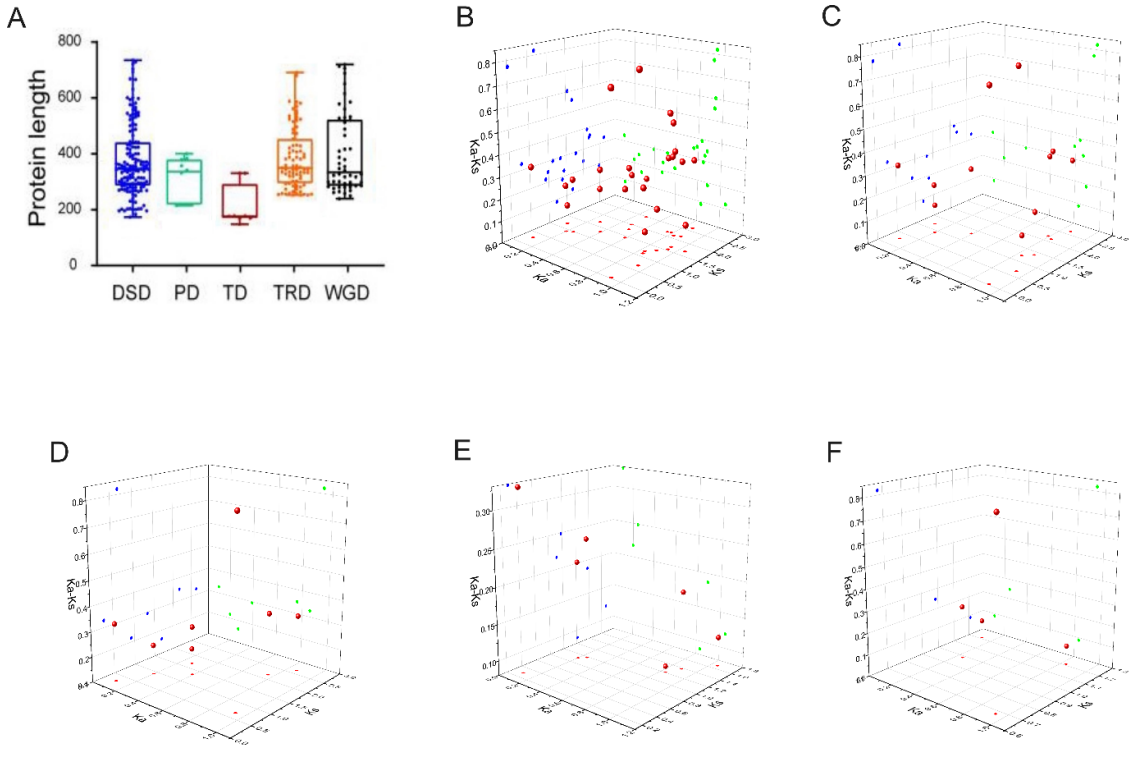

Figure 6. Gene *AcbHLH* duplication events. (A) Length distribution of AcbHLHs across 05 duplication events. (B-F) distribution of Ka, Ks, and Ka/Ks value in duplicated genes across five duplication

 doi:10.20944/preprints202410.1903.v1

events. (B-F) DSD, WGD, TRD, TD and PD event, respectively. Further details on the duplicated genes across the 5 duplication events are provided in table S9.

2.7. Synteny analysis of AcbHLH genes

To elucidate the phylogenic processes underlying *A. catechu bHLH* gene family, developed seven comparative synteny maps *A.catechu* association, encompassing 03 dicot, (*A. thaliana, V.vinifera* and *S.lycopersicum*) and four monocot (*B.distachyon, O.sativa, Z.mays* and *C.nucifera*) (Figure 7, Additional file 5: Table S5). Collinear analysis, 63 genes *AcbHLH*

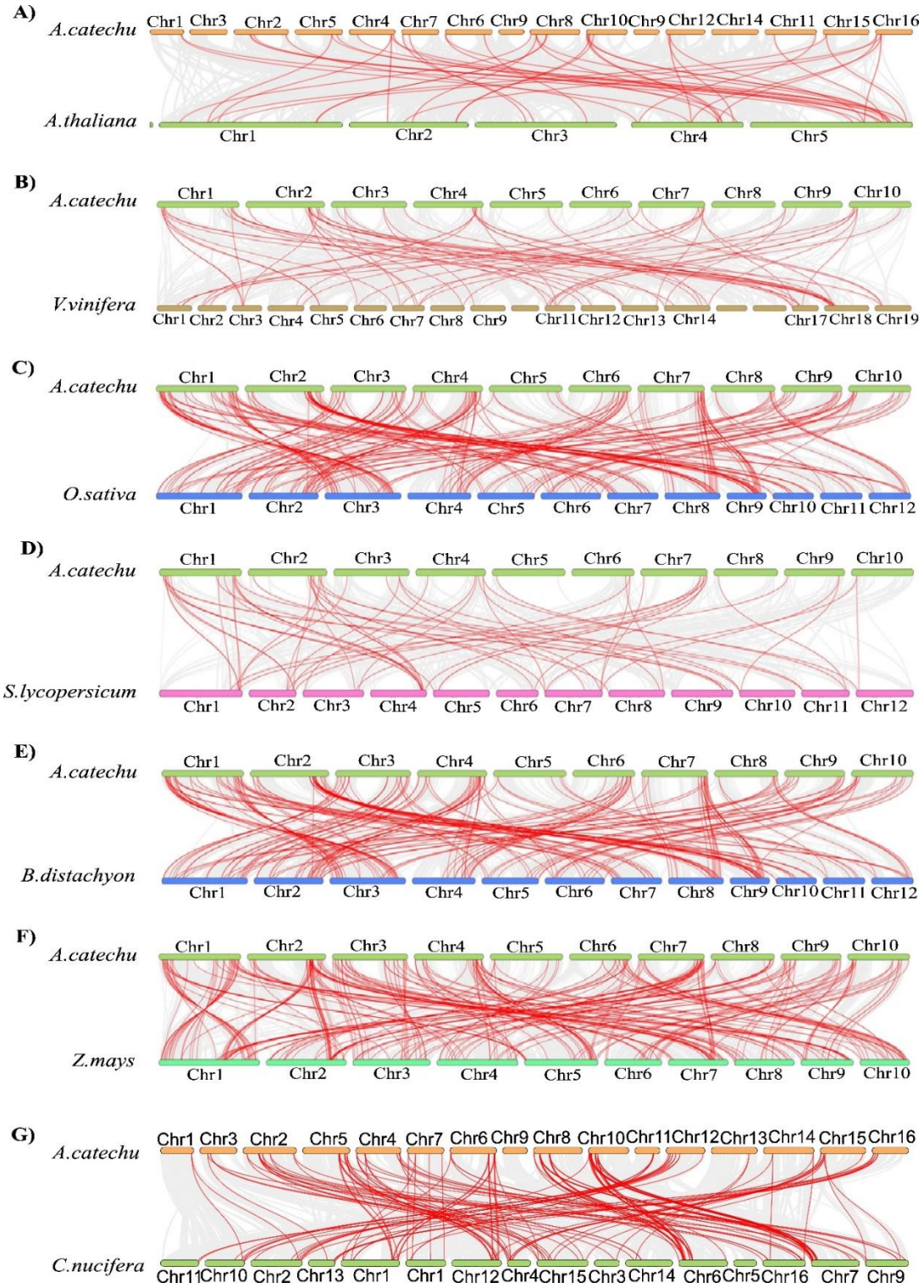

Figure 7. Comparative Synteny Analysis of *AcbHLH* Genes in *A. catechu*, 06 Representative Plant Species (*A.thaliana, V.vinifera, S.lycopersicum, B.distachyon, O.sativa* subsp. *indica, Z.mays* and

 doi:10.20944/preprints202410.1903.v1

C.nucifera). Gray lines show conserved syntenic blocks between *A. catechu* and other plant genomes. Red lines indicate *AcbHLH* pairs of gene are syntenic across species.

Were identified with the gene set of *A.thaliana* (41), *V.vinifera* (66), *S.lycopersicum* (59), *B. distachyon* (193), *O.sativa* (150), and *Z. mays* (170). The homologous logarithmic values for 6 species shown as *A. thaliana* (44), *V. vinifera* (66), *S. lycopersicum* (60), *B. distachyon* (194), *O. sativa* (207), *Z. mays* (273) and C. nucifera (160). A notable finding is that several *AcbHLH* are connected with at least three synthetic gene pairs, emphasising the link between *A. catechu* and *V. vinifera, AcbHLH43* and *AcbHLH64. AcbHLH* gene, along with *A. thaliana* (54.0%), *B. Distachyon* (53.73%), and *O. sativa* (55.56%) concerning 02 or more synthetic gene pairs, regarded as the presence of these genes within the homologous gene pairs, comprising over 50% of the total, underscores their importance in evolutionary pathways.

As expected some *AcbHLH* genes, especially *AcbHLH73, AcbHLH71, AcbHLH67* and *AcbHLH29*, entirely formed homologous genes with three dicotyledon. The fact that all these genes, such as *AcbHLH73, AcbHLH71, AcbHLH67 and AcbHLH29*, presence of homologous gene pairs with three typical dicotyledons indicated a possible evolutionary lineage leading to the development of dicotyledons. It was observed that sure. It was observed that specific *AcbHLH* genes homology with at least one gene in the five specified species, for instance, *AcbHLH37, AcbHLH30, AcbHLH21, AcbHLH73, AcbHLH17, AcbHLH42, AcbHLH3*, and *AcbHLH41*, This observation suggested that these might be crucial primordial gene either been lost or showed highly significant differentiation during *A. catechu* long-term evolutionary history.

2.8. The expression patterns of AcbHLH genes differ across tissues

The expression levels of *AcbHLH* genes in various tissues including male and female flowers, endosperm, pericarp, leaf, leaf vein, aerial roots, and underground roots were computed using the transcriptome data of *A. catechu* (Table S8). The expression pattern shows that *AcbHLH22* wer highly significantly expressed in all tissues and was followed by *AcbHLH07, AcbHLH31* and *AcbHLH44* in the pericarp, endosperm, male and female flower. The expression level of *AcbHLH70, AcbHLH74, AcbHLH59, AcbHLH66, AcbHLH40, AcbHLH39* and *AcbHLH13* were highly expressed in flowers while *AcbHLH74, AcbHLH59, AcbHLH52* and *AcbHLH37* were higher in endosperm. *AcbHLH02, AcbHLH13, AcbHLH26* and *AcbHLH37* were upregulating in pericarp while *AcbHLH57, AcbHLH40, AcbHLH39, AcbHLH13* and *AcbHLH08* were highly expressed in leaves. Furthermore, the expression level of *AcbHLH02, AcbHLH34, AcbHLH37, AcbHLH44* and *AcbHLH56* genes were highly expressed in roots. Notably, *AcbHLH04, AcbHLH12, AcbHLH17, AcbHLH24, AcbHLH30, AcbHLH33, AcbHLH53, AcbHLH54, AcbHLH61, AcbHLH66* and *AcbHLH69* expressed low in all tissues of *A.catechu* (Figure 8; Table S8).

 doi:10.20944/preprints202410.1903.v1

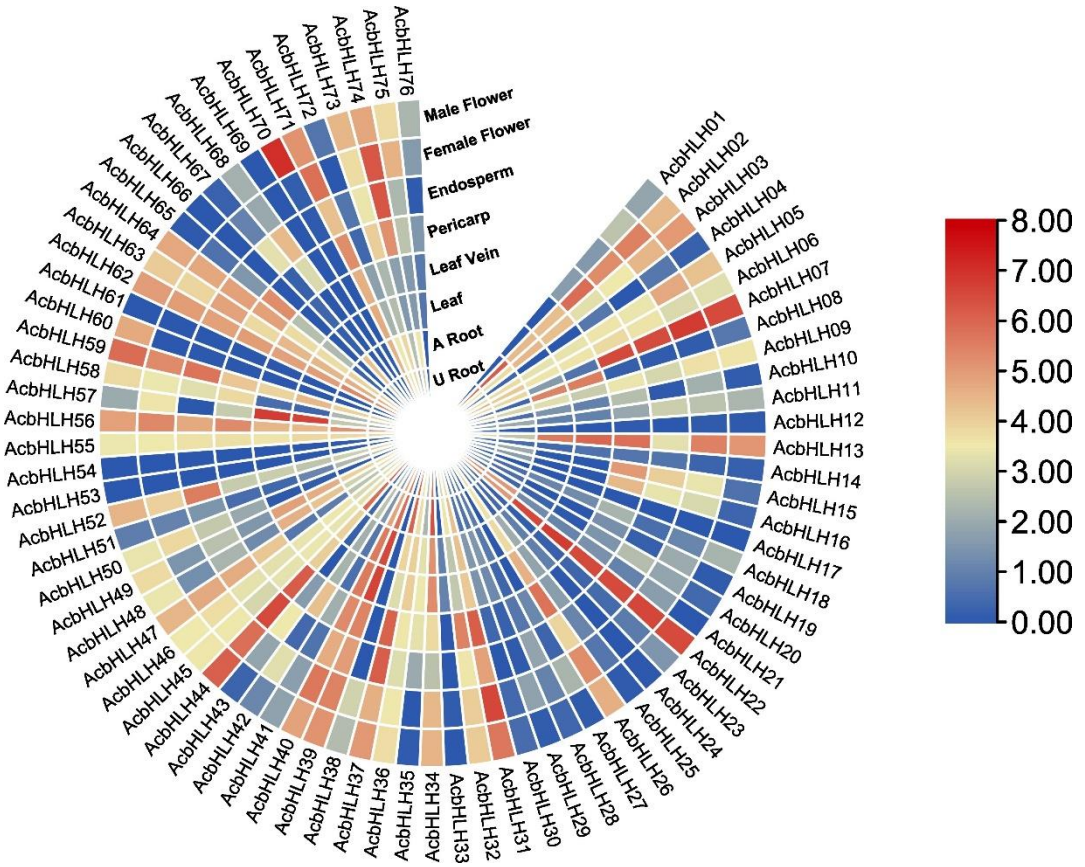

Figure 8. Heatmap of AcbHLH Gene Expression in Areca catechu. The red color represents higher log2FPKM values while blue represents the lover values.

2.9. Effects of different treatments on AcbHLH expression

To investigate the role of *AcbHLH* genes in response to abiotic stress treatments, Salt stress (NaCl) and Polyethylene Glycol stress (PEG) in roots and leaves, and evaluated the expression levels of 9 *AcbHLH* selected genes under each of these specific stress conditions utilizing (qRT-PCR) (Figure. 9). This result revealed a variety of responses to salt and PEG stress conditions. Expression of some specific genes activated or inhibited differently under different conditions, and showed significant stimuli in the initial stage of stress. *AcbHLH22* and *AcbHLH62* have elevated expression under salt and drought stresses. Under drought stress conditions specifically in roots, the expression of *AcbHLH59* was significantly upregulated, whereas under PEG stress, *AcbHLH59* was downregulated in leaves. *AcbHLH48* was downregulated in the roots under drought and NaCl stresses, PEG stress led to a significant upregulated of gene expression in the roots. Furthermore, we observed a correlation among all *AcbHLH* genes of roots and leaves (Figure 9E). We observed a substantial correlation between the expressions of different genes. Most of the genes possess significant positive correlations.

To illustrate, we demonstrated the four *AcbHLH22* with *AcbHLH39* and *AcbHLH43* with *AcbHLH45* genes as their highest expression level. These genes not only exhibited significant positive correlation but also showed significant positive correlations *AcbHLH22* with *AcbHLH43 and AcbHLH45*, while *AcbHLH45 with AcbHLH39 and AcbHLH43* expressed.

 doi:10.20944/preprints202410.1903.v1

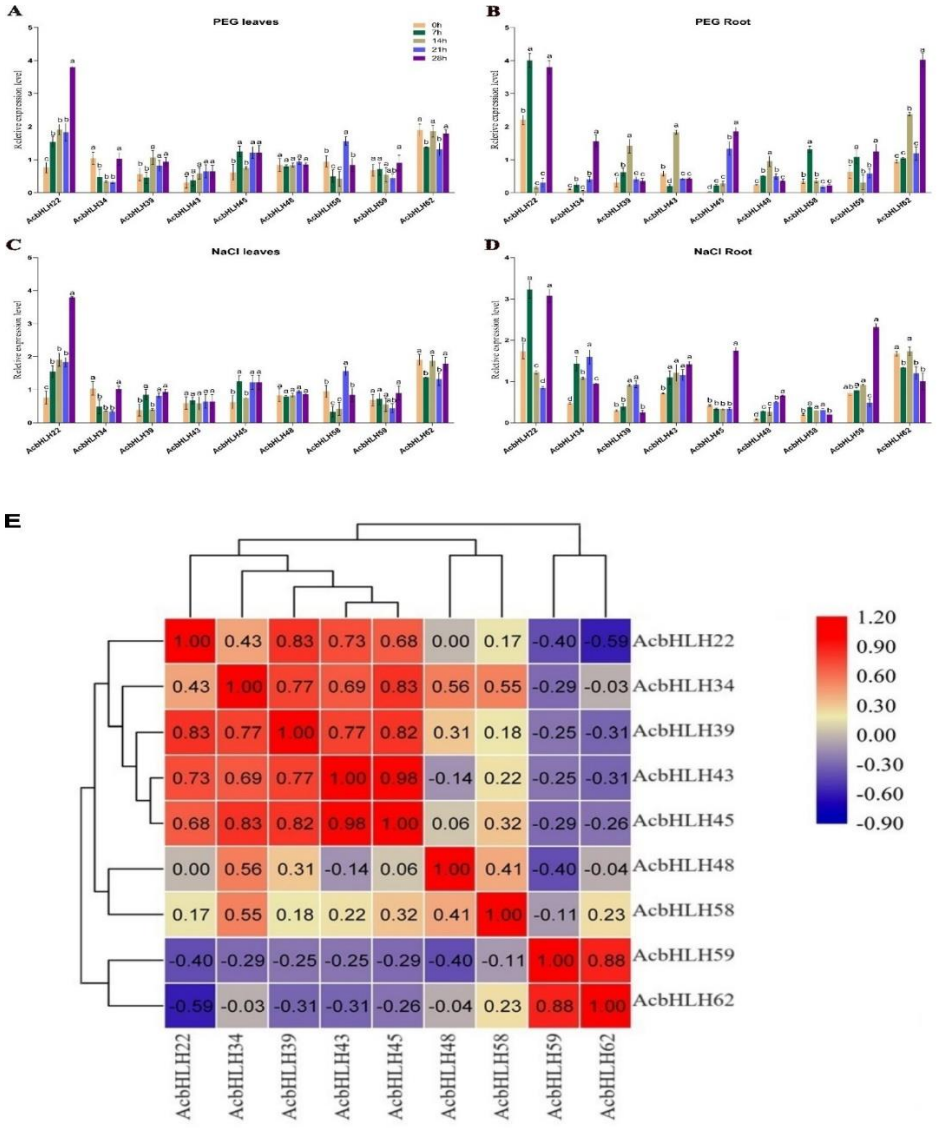

Figure 9. (A) The effects of abiotic stress (NaCl and PEG) detected by qRT-PCR on the expression of 09 *AcbHLH* genes in roots and leaves of young *A. catechu* seedlings (0, 7, 14, 21, and 28 hours). (B) Different letters indicate statistically significant differences between groups ($p < 0.05$, LSD).

3. Discussion

Transcription factors control several biological processes by controlling the expression of certain genes, such as growth, development, and stress responses in plants. Latest investigations shown that the *A. catechu* plant has a total of 31,406 protein-coding genes on 16 chromosomes. Among these genes, a significant portion are classified as transcription factors [29]. Only a small number of the transcription factor families, including the DOF [37], WRKY, and MYB families (paper under review), have been systematically investigated in *A. catechu*. However, the *bHLH* genes have not yet been characterized in *A. catechu*. Among eukaryotic TFs family *bHLH* is the second-most extensive family [38]. Previous research focusing on distribution of bHLH genes in different plants. In Arabidopsis, 162 members of *bHLH* have been divided into 12 sub-families based on phylogenic trees [21]. In rice, 167 members were divided into 22 sub-families [22]. In maize, 208 members were divided into 18 sub-families [24], while 261 in peanut [39], 202 in *Populus* [26], 152 in tomato [23], 159 in wheat [25] and 124 in potato [27] *bHLH* members were divided into 19, 25, 26, 19 and 15 sub-families respectively. In the current study, a total of 76 *bHLH* genes were identified in *A. Catechu* were divided into 24 sub-families based on phylogenic analysis, reflecting their evolutionary relationships and

 doi:10.20944/preprints202410.1903.v1

functional diversities. Variation in the *A. catechu bHLH* subgroup is primarily localized to the basic region, signifying its value in shaping *bHLH* domain [40]. The basic region plays pivotal role in determining activity of *bHLH* gene, allows the interaction of heterodimers or homodimers with other TFs [41]. The gene number is much less than other organisms of different familes. The ratio of AcbHLH genes to the total gene in A. catechu was about 0.57%, which is less than A. thaliana (0.59%), but more than in rice (0.44%), tomato 0.46% and poplar (0.40%) [37]. Though, while having a higher genome size of 2.7 Gb [29], *A.catechu* contains fewer *bHLH* genes than other plants. This shows that the species' evolutionary history influences the number of gene family members and genome size. Research has revealed that gene duplication events are essential in several gene families' rapid growth and expansion [42]. In the genome of *A. catechu*, four tandem duplications and 23 segmental duplication events were found in *AcbHLH* family. These results provide strong evidence that segmental and tandem duplication had a role in growth and divergence of *AcbHLH* gene. Previous research shows that segmental duplication has more substantial role in genome evolution than tandem duplication [38]. Gene functions within a gene family are frequently conserved across plant species, although not always possible. As a result, it is important to determine the orthologs across plant species using synteny analysis precisely. The result showed that AcbHLH 63 in the *Areca* genome exhibited significant synteny with *A. thaliana, V. vinifera, S. lycopersicum, B. distachyon, O. sativa*, and *Z.mays*. These result indicate that these genes may have played a crucial role in the evolutionary history of *A. catechu*, potentially undergoing loss or significant differentiation. The relationship between intron number and expression level in plants suggests that a more compact gene structure might facilitate rapid gene expression in response to stimulation [43]. In the current study, the intron-exon number of *AcbHLH* genes was found to range from 13 to 1 (Figure 3B), indicating evolutionary modifications, such as insertion or deletion or exon-gain or loss, in intron-exon structures of *AcbHLH* genes. The conserved motifs and gene structure investigation are crucial for phylogenic relationships among gene members. A significant proportion of *AcbHLH* genes within the same subfamily shared similar intron-exon, structures and motifs. Essentially *AcbHLH* proteins contained *bHLH* domains 1 and 2. Additionally, the composition of other motifs was distinctive, and they were conserved across subgroups. For instance, motif 4 is exclusively present at the end of subgroups 3, 5, and 6. Other plants have been observed to exhibit similar phenomena [25,44]. Understanding these motif and exon-intron distribution variations and their functional implications can contribute to a more comprehensive understanding of *AcbHLH* protein evolution and their diverse roles in biological processes. Cis-acting elements, which are substances that bind to trans-acting factors, play a crucial role in regulating the activity of target genes [45].These elements play an important role in molecular controlling genes, especially in response to stress expression [46]. The study predicted 76 cis-acting elements in the growth promoter region of *AcbHLH* genes.

AcbHLH22 in subgroup 04 of the *A. catechu* bHLH family showed significant upregulation during developmental growth and biological stress, representing 14.97% of the total fragment repeats within the *A. catechu bHLH* family. Therefore, it is speculated that the 18 genes within this subgroup might play a vital in *A. catechu*'s growth and development when compare with other species. Additionally, we selected 9 *AcbHLH* genes to comprehensively understand the responses of 9 *AcbHLH* genes to both abiotic stresses and various reproductive stages. The expression levels across genes showed significant differences, with some exhibiting two or more distinct levels of significant variations. For instance, *AcbHLH22* of subgroup 04 showed elevated expression under NaCl and drought stresses, and *AtbHLH92* under NaCl stress [47]. This study reveals that *AcbHLH* genes exhibit antagonistic effects under various stresses, with 11 genes upregulated significantly and under cold stresses, eight gene downregulated in *C. quinoa* leaves [33]. The result revealed expression of various genes *SibHLH* in *S. italica* was up/downregulated under abiotic stress [33]. This result showed that *AcbHLH22* gene was significantly upregulated under six stress conditions, a finding not observed in *A. thaliana*, warranting further investigation [21].

The *bHLH* gene, specifically *bHLH122* in *A. thaliana*, plays a crucial role in ABA signalling pathways, enhancing plant drought resistance by reducing ABA metabolism and catabolism [48]. Expression of *bHLH* genes is highly significantly upregulated in the roots and leaves of *C. quinoa*v

 doi:10.20944/preprints202410.1903.v1

under drought stress, possibly due to its involvement in the ABA signalling pathway, thus improving drought resistance. Previous research has indicated that *bHLH* is linked to leaf stomata development regulation [49,50]. Meanwhile, a higher proportion of *AcbHLH* genes are expressed in *A. catechu* leaves. Therefore, the observed upregulation of *AcbHLH* genes in *A. catechu* leaves in showed their potential involvement in both leaf development and drought tolerance. The observed gene expression patterns align with A. catechu strong drought resistance, emphasizing its adaptive molecular responses.

We identified the 9 upregulated *AcbHLH* genes and suggests their potential co-expression of multiple *AcbHLH* genes to influence plant physiological prosses. Furthermore, it is essential that functions of *AcbHLH* genes family in vegetative organs of *A. catechu*. We conducted the expression level 9 Ac*bHLH* genes in the vegetative parts of *A. catechu* at different developmental stages. In the vegetative stage, the *AcbHLH* genes of A. catechu also play crucial roles. For example, the gene *AcbHLH22* expressed significantly higher expression levels on the 28Th day in the leaves than in the roots. The gene expression during the growth development of the two genes *AcbHLH22* and *AcHLH62* showed significant changes. These two genes are likely to have a regulatory role in the vegetative development stage of *A. catechu*. The *bHLH* TFs *ZmbHLH180* and *ZmbHLH23* were shown to interact with synergistic expression in *Zea mays*, as revealed by yeast two-hybrid experiments [51]. Furthermore, A notable positive correlation was noticed among *AcbHLH45* and *AcbHLH42* in the gene correlation heatmap. We postulate synergistic interaction between these two genes, coordinating their physiological state, especially under abiotic stress conditions and the vegetative periods. The coordinating action of these genes may be crucial for regulating physiological processes, especially in response to abiotic stress and during vegetative development. In conclusion, these results showed that particular justification of *AcbHLH* gene are integral components of gene regulatory networks.

4. Methods

4.1. Identification of AcbHLH genes in A. catechu

The complete genome sequence of *A. catechu* L. was obtained from the National Center for Biotechnology Information (NCBI) using the following (ID: JAHSVC000000000; BioSample: SAMN19591864; Accession: PRJNA735650). The *bHLH* protein domain (Pfam ID: PF00010) were retrieved via the Pfam database. The conserved *bHLH* protein domain (Pfam ID: PF00010) was retrieved from the Pfam database. A two-step approach was employed to identify *AcbHLH* within *A. catechu*. The first approach, the protein sequences of *Arabidopsis bHLH* used as query sequences downloaded from TAIR (http://www.arabidopsis.org) apply BLASTp with a score value of at least 100 and an E-value of less than or equal to 1e-10. In second approach, The bHLH candidates were initially identified within the *A. catechu* genome through a hidden Markov model (HMM) search, employing e-value lower than 10^{-5} based on a previously used method [52]. Finally, Unidentified conserved sequence motifs were manually removed from the dataset. To validate the presence of *bHLH* domains, HMMER R3.0 (http://plants.ensembl.org/hmmer/index.html) [53] with default parameters and a 0.01 cutoff was used for protein homology, and confirmation using the SMART domain database (http://smart.embl-heidelberg.de/)."

4.2. Analysis of Physicochemical Properties of AcbHLH

The ExPasy ProtParam server (http://web.expasy.org/protparam/) was utilized to characterize the fundamental characteristics of *AcbHLH* tri-helix proteins, sequence length, isoelectric point (pI), and molecular mass within the *AcbHLH* gene family. A protein subcellular localization prediction tool (PSORT) was used to determine the likely cellular location of the predicted proteins (https://www.genscript.com/psort.html) [54].

 doi:10.20944/preprints202410.1903.v1

4.3. The bHLH Gene Structure, Conserved Motif and Cis-Acting Elements

The characterized *AcbHLH* proteins were aligned using ClustalW, and GeneDoc software was used to manually made and improve the quality of the alignment. An in-depth analysis of the intron-exon structure of *AcbHLH* genes in *A. catechu* was achieved the Gene Structure Display Server (GSDS) (http://GSDS.cbi.pku.edu.cn/) by leveraging the GFF3 annotation data. To identify motifs of *AcbHLH* proteins online tool MEME were used (http://meme-suite.org/tools/meme). Promoter sequences of *AcbHLH* genes were retrieved from the *A. catechu* genome (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Upstream regulatory sequences of these genes were identified as the 2,000 base pairs preceding the ATG codon. The distribution of cis-elements within the promoter regions was analyzed using the PlantCARE database, (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

4.4. Chromosomal distribution and gene duplication

Chromosomal localization of *AcbHLH* was determined by mapping their physical location to the *A. catechu* genomic database and Circular Genome Data Visualization tool (Circos).To analyze gene duplication events, the Multiple Collinearity Scan toolkit (MCScanX) was employed with its parameters [55]. Dual Synteny Plotter were used to analysis gene homogeneity among *A.catechu* and *A.thaliana* (https://github.com/CJ-Chen/TBtools). Nonsynonymous and synonymous substitution rates were calculated for each duplicated bHLH gene using the Ka/Ks Calculator 2.0 [56].

4.5. Evolutionary Relationships and classification of AcbHLH gene family

All the recognized genes *AcbHLH* were classified into categories on *AtbHLH* categorisation. The phylogenic tree was created utilizing the NJ method in MEGA-X software using 1000 bootstrap replicates and the neighbour-joining method, the phylogenetic tree was created and visualized using iTOL [56]. For phylogenetic analysis, The full-length amino acid sequences of *A. thaliana bHLH* proteins were utilize, ((https://itol.embl.de/) (Additional file 1: Table S1).

4.6. RNA-seq data analysis

The expression profile study of *AcbHLH* genes was conducted using Illumina Hisequence 4000 RNA-sequence datas submitted to NCBI database (accession number: PRJNA767949). Genes exhibiting a log2FC fold change greater than 1 and a false discovery rate (FDR) of 0.05, and differential expression were identified using a significant threshold of p-value<0.05. The expression pattern of *AcbHLH* genes was plotted using TBtools. Gene expression studies were conducted using the BMK Cloud platform available at (https://.biocloud.net).

4.7. Plant materials, growth conditions, and abiotic stress in A. catechu

The research utilized seedling of *A.catechu* L., (Reyan No. 1) from the Coconut Research Institute of the Chinese Academy of Tropical Agricultural Sciences, located in Wenchang, Hainan province, China. Seedlings grown in pots (size: 12 cm × 12 cm) at 14/10 hours day/night at 28 /25 °C. The plants were treated including drought (25% PEG6000), and salt (5% NaCl) [30]. Leaf and root samples were collected at 0, 7, 14, 21 and 28 days for each treatment, and preserved at -80 °C as subsequent analysis.

4.8. Total RNA extraction, cDNA and qRT-PCR

The extraction of total RNA from plant samples was performed using a specialized extraction kit (TIANGEN, Beijing, China). RNA concentration and purity were checked using NanoDrop 2000 (KAIAO, Beijing, China). First-strand cDNA synthesed was performed using the TIANScript RT Kit (TIANGEN, Beijing, China).

To analyze Gene expression quantitative real-time PCR (qRT-PCR) was performed using Vazyme Master Mix (Vazyme, Nanjing, China). PCR conditions were set according to the company manual. Primers were designed via Primer 6.0 software, and *Actin* was used as an internal control

 doi:10.20944/preprints202410.1903.v1

for gene expression analysis (**Table S2**). The $2^{-\Delta\Delta CT}$ method was utilized to analyze qRT-PCR data and calculate relative gene expression levels.

4.9. Statistical analysis

Analysis of variance (ANOVA) were analyzed by Statistics 8.1, software, followed by the Tukey LSD test, and compared with least significant difference (LSD) ($P \geq 0.05$). Histogram drawn with Origin 8.0 software (OriginLab).

5. Conclusions

In this research, we represents the first comprehensive and systematic evaluation of the *AcbHLH* gene family within the *A. catechu* genome, encompassing a wide range of analytical approaches and verifications. A total of 76 *AcbHLH* genes-proteins were identified and categorized into 24 distinct subgroups based on their protein domain profiles and gene structure, which carry the accuracy of the categorization based on phylogenic analysis and irregularly distributed on 16 chromosomes (Chr00). The distribution of AcbHLH genes across the chromosomes was not evenly distributed. Some *AcbHLH* genes are involved in gene replication events, and fragment repeat contributes more favourably than tandem duplication. A significant homologous phenomenon was discovered, revealing the homology between one *AcbHLH* gene and multiple *bHLH* genes in *A. Catechu*. The *AcbHLH* genes from *A. catechu* are the most closely related genes among the six representative plant species, as determined by sequence comparison. Finally, qRT-PCR analysis revealed that 9 *AcbHLH* genes were differentially expressed in response to abiotic stress conditions, expression mechanism of these genes was investigated throughout the developmental stages, *AcbHLH22* and *AcbHLH144* demonstrated a notable impact on abiotic stress resistance. It is speculated that *AcbHLH22* and *AcbHLH62* are essential components of the *A. catechu* life cycle, contributing to its viability and development. In conclusion, the result of this research offers valuable insights into the biological functions and development of *AcbHLH* in *A. catechu* and will facilitate future research on the functions of *AcbHLHs*.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Author Contributions: A.A and N.M.K conceived and designed this project. A.A performed the analyses. J.Y.Q participated in Areca seedling's stress and the data analyses. A.A carried out the experimental study and wrote the draft manuscript. G.Z contributed to its improvement and provided literature support. Y.W financially support, checked and supervised this project. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (31960064).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data generated and analyzed in this study are available in the article.

Acknowledgments: We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript. We thank BMKCloud (www.biocloud.net) for gene expression analysis assistance in this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Xu, Y.; Zhang, H.; Zhong, Y.; Jiang, N.; Zhong, X.; Zhang, Q.; Chai, S.; Li, H.; Zhang, Z. Comparative Genomics Analysis of BHLH Genes in Cucurbits Identifies a Novel Gene Regulating Cucurbitacin Biosynthesis. *Hortic. Res.* **2022**, *9*, uhac038.
2. Li, J.; Li, X.; Han, P.; Liu, H.; Gong, J.; Zhou, W.; Shi, B.; Liu, A.; Xu, L. Genome-Wide Investigation of BHLH Genes and Expression Analysis under Different Biotic and Abiotic Stresses in Helianthus Annuus L. *Int. J. Biol. Macromol.* **2021**, *189*, 72–83.

3. Lai, X.; Vega-Léon, R.; Hugouvieux, V.; Blanc-Mathieu, R.; van Der Wal, F.; Lucas, J.; Silva, C. S.; Jourdain, A.; Muino, J. M.; Nanao, M. H. The Intervening Domain Is Required for DNA-Binding and Functional Identity of Plant MADS Transcription Factors. *Nat. Commun.* **2021**, *12* (1), 4760.
4. Pires, N.; Dolan, L. Origin and Diversification of Basic-Helix-Loop-Helix Proteins in Plants. *Mol. Biol. Evol.* **2010**, *27* (4), 862–874.
5. Zhang, Z.; Chen, J.; Liang, C.; Liu, F.; Hou, X.; Zou, X. Genome-Wide Identification and Characterization of the BHLH Transcription Factor Family in Pepper (Capsicum Annuum L.). *Front. Genet.* **2020**, *11*, 570156.
6. Grove, C. A. A Multiparameter Network Reveals Extensive Divergence Between C. Elegans BHLH Transcription Factors: A Dissertation. 2009.
7. Qin, Y.; Li, J.; Chen, J.; Yao, S.; Li, L.; Huang, R.; Tan, Y.; Ming, R.; Huang, D. Genome-Wide Characterization of the BHLH Gene Family in Gynostemma Pentaphyllum Reveals Its Potential Role in the Regulation of Gypenoside Biosynthesis. *BMC Plant Biol.* **2024**, *24* (1), 205. https://doi.org/10.1186/s12870-024-04879-y.
8. Gonzalez, A.; Mendenhall, J.; Huo, Y.; Lloyd, A. TTG1 Complex MYBs, MYB5 and TT2, Control Outer Seed Coat Differentiation. *Dev. Biol.* **2009**, *325* (2), 412–421. https://doi.org/10.1016/j.ydbio.2008.10.005.
9. Chang, G.; Ma, J.; Wang, S.; Tang, M.; Zhang, B.; Ma, Y.; Li, L.; Sun, G.; Dong, S.; Liu, Y.; Zhou, Y.; Hu, X.; Song, C.-P.; Huang, J. Liverwort BHLH Transcription Factors and the Origin of Stomata in Plants. *Curr. Biol.* **2023**, *33* (13), 2806-2813.e6. https://doi.org/10.1016/j.cub.2023.05.050.
10. Qiu, Y.; Tao, R.; Feng, Y.; Xiao, Z.; Zhang, D.; Peng, Y.; Wen, X.; Wang, Y.; Guo, H. EIN3 and RSL4 Interfere with an MYB–BHLH–WD40 Complex to Mediate Ethylene-Induced Ectopic Root Hair Formation in Arabidopsis. *Proc. Natl. Acad. Sci.* **2021**, *118* (51). https://doi.org/10.1073/pnas.2110004118.
11. Groszmann, M.; Paicu, T.; Alvarez, J. P.; Swain, S. M.; Smyth, D. R. SPATULA and ALCATRAZ, Are Partially Redundant, Functionally Diverging BHLH Genes Required for Arabidopsis Gynoecium and Fruit Development. *Plant J.* **2011**, *68* (5), 816–829. https://doi.org/10.1111/j.1365-313X.2011.04732.x.
12. Leivar, P.; Monte, E.; Oka, Y.; Liu, T.; Carle, C.; Castillon, A.; Huq, E.; Quail, P. H. Multiple Phytochrome-Interacting BHLH Transcription Factors Repress Premature Seedling Photomorphogenesis in Darkness. *Curr. Biol.* **2008**, *18* (23), 1815–1823. https://doi.org/10.1016/j.cub.2008.10.058.
13. Liang, J.; Fang, Y.; An, C.; Yao, Y.; Wang, X.; Zhang, W.; Liu, R.; Wang, L.; Aslam, M.; Cheng, Y.; Qin, Y.; Zheng, P. Genome-Wide Identification and Expression Analysis of the BHLH Gene Family in Passion Fruit (Passiflora Edulis) and Its Response to Abiotic Stress. *Int. J. Biol. Macromol.* **2023**, *225*, 389–403. https://doi.org/10.1016/j.ijbiomac.2022.11.076.
14. Su, L.; Zhang, Y.; Yu, S.; Geng, L.; Lin, S.; Ouyang, L.; Jiang, X. RcbHLH59-RcPRs Module Enhances Salinity Stress Tolerance by Balancing Na+/K+ through Callose Deposition in Rose (Rosa Chinensis). *Hortic. Res.* **2023**, *10* (3). https://doi.org/10.1093/hr/uhac291.
15. Xue, L.; Wei, Z.; Zhai, H.; Xing, S.; Wang, Y.; He, S.; Gao, S.; Zhao, N.; Zhang, H.; Liu, Q. The <scp>IbPYL8–IbbHLH66–IbbHLH118</Scp> Complex Mediates the Abscisic Acid-dependent Drought Response in Sweet Potato. *New Phytol.* **2022**, *236* (6), 2151–2171. https://doi.org/10.1111/nph.18502.
16. Zhang, L.; Xiang, Z.; Li, J.; Wang, S.; Chen, Y.; Liu, Y.; Mao, D.; Luan, S.; Chen, L. BHLH57 Confers Chilling Tolerance and Grain Yield Improvement in Rice. *Plant. Cell Environ.* **2023**, *46* (4), 1402–1418. https://doi.org/10.1111/pce.14513.
17. Singh, S. K.; Patra, B.; Paul, P.; Liu, Y.; Pattanaik, S.; Yuan, L. BHLH IRIDOID SYNTHESIS 3 Is a Member of a BHLH Gene Cluster Regulating Terpenoid Indole Alkaloid Biosynthesis in Catharanthus Roseus. *Plant Direct* **2021**, *5* (1). https://doi.org/10.1002/pld3.305.
18. Li, X.; Cao, L.; Jiao, B.; Yang, H.; Ma, C.; Liang, Y. The BHLH Transcription Factor AcB2 Regulates Anthocyanin Biosynthesis in Onion (Allium Cepa L.). *Hortic. Res.* **2022**. https://doi.org/10.1093/hr/uhac128.
19. Liu, S.; Wang, Y.; Shi, M.; Maoz, I.; Gao, X.; Sun, M.; Yuan, T.; Li, K.; Zhou, W.; Guo, X.; Kai, G. SmbHLH60 and SmMYC2 Antagonistically Regulate Phenolic Acids and Anthocyanins Biosynthesis in Salvia Miltiorrhiza. *J. Adv. Res.* **2022**, *42*, 205–219. https://doi.org/10.1016/j.jare.2022.02.005.
20. Mohammad; Hurrah, I. M.; Kumar, A.; Abbas, N. Synergistic Interaction of Two <scp>bHLH</Scp> Transcription Factors Positively Regulates Artemisinin Biosynthetic Pathway in *Artemisia Annua* L. *Physiol. Plant.* **2023**, *175* (1). https://doi.org/10.1111/ppl.13849.
21. Hao, Y.; Zong, X.; Ren, P.; Qian, Y.; Fu, A. Basic Helix-Loop-Helix (BHLH) Transcription Factors Regulate a Wide Range of Functions in Arabidopsis. *Int. J. Mol. Sci.* **2021**, *22* (13), 7152.
22. Li, X.; Duan, X.; Jiang, H.; Sun, Y.; Tang, Y.; Yuan, Z.; Guo, J.; Liang, W.; Chen, L.; Yin, J.; Ma, H.; Wang, J.; Zhang, D. 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.
23. Wang, J.; Hu, Z.; Zhao, T.; Yang, Y.; Chen, T.; Yang, M.; Yu, W.; Zhang, B. Genome-Wide Analysis of BHLH Transcription Factor and Involvement in the Infection by Yellow Leaf Curl Virus in Tomato (Solanum Lycopersicum). *BMC Genomics* **2015**, *16* (1), 1–14. https://doi.org/10.1186/s12864-015-1249-2.

24. 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), 1–14. https://doi.org/10.1186/s12870-018-1441-z.
25. Wang, L.; Xiang, L.; Hong, J.; Xie, Z.; Li, B. Genome-Wide Analysis of BHLH Transcription Factor Family Reveals Their Involvement in Biotic and Abiotic Stress Responses in Wheat (Triticum Aestivum L.). *3 Biotech* **2019**, *9*, 1–12.
26. Zhao, K.; Li, S.; Yao, W.; Zhou, B.; Li, R.; Jiang, T. Characterization of the Basic Helix–Loop–Helix Gene Family and Its Tissue-Differential Expression in Response to Salt Stress in Poplar. *PeerJ* **2018**, *6*, e4502. https://doi.org/10.7717/peerj.4502.
27. Wang, R.; Zhao, P.; Kong, N.; Lu, R.; Pei, Y.; Huang, C.; Ma, H.; Chen, Q. Genome-Wide Identification and Characterization of the Potato BHLH Transcription Factor Family. *Genes (Basel).* **2018**, *9* (1), 54.
28. Salehi, B.; Konovalov, D. A.; Fru, P.; Kapewangolo, P.; Peron, G.; Ksenija, M. S.; Cardoso, S. M.; Pereira, O. R.; Nigam, M.; Nicola, S. Areca Catechu—From Farm to Food and Biomedical Applications. *Phyther. Res.* **2020**, *34* (9), 2140–2158.
29. Zhou, G.; Yin, H.; Chen, F.; Wang, Y.; Gao, Q.; Yang, F.; He, C.; Zhang, L.; Wan, Y. The Genome of Areca Catechu Provides Insights into Sex Determination of Monoecious Plants. *New Phytol.* **2022**, *236* (6), 2327–2343.
30. Xue, G.; Fan, Y.; Zheng, C.; Yang, H.; Feng, L.; Chen, X.; Yang, Y.; Yao, X.; Weng, W.; Kong, L. BHLH Transcription Factor Family Identification, Phylogeny, and Its Response to Abiotic Stress in Chenopodium Quinoa. *Front. Plant Sci.* **2023**, *14*, 1171518.
31. Li, K.; Duan, L.; Zhang, Y.; Shi, M.; Chen, S.; Yang, M.; Ding, Y.; Peng, Y.; Dong, Y.; Yang, H. Genome-Wide Identification and Expression Profile Analysis of Trihelix Transcription Factor Family Genes in Response to Abiotic Stress in Sorghum [Sorghum Bicolor (L.) Moench]. *BMC Genomics* **2021**, *22* (1), 1–17.
32. Toledo-Ortiz, G.; Huq, E.; Quail, P. H. The Arabidopsis Basic/Helix-Loop-Helix Transcription Factor Family. *Plant Cell* **2003**, *15* (8), 1749–1770.
33. Fan, Y.; Lai, D.; Yang, H.; Xue, G.; He, A.; Chen, L.; Feng, L.; Ruan, J.; Xiang, D.; Yan, J. Genome-Wide Identification and Expression Analysis of the BHLH Transcription Factor Family and Its Response to Abiotic Stress in Foxtail Millet (Setaria Italica L.). *BMC Genomics* **2021**, *22*, 1–18.
34. Sun, W.; Jin, X.; Ma, Z.; Chen, H.; Liu, M. Basic Helix–Loop–Helix (BHLH) Gene Family in Tartary Buckwheat (Fagopyrum Tataricum): Genome-Wide Identification, Phylogeny, Evolutionary Expansion and Expression Analyses. *Int. J. Biol. Macromol.* **2020**, *155*, 1478–1490.
35. Sun, H.; Fan, H.-J.; Ling, H.-Q. Genome-Wide Identification and Characterization of the BHLH Gene Family in Tomato. *BMC Genomics* **2015**, *16*, 1–12.
36. Chen, F.; Hu, Y.; Vannozzi, A.; Wu, K.; Cai, H.; Qin, Y.; Mullis, A.; Lin, Z.; Zhang, L. The WRKY Transcription Factor Family in Model Plants and Crops. *CRC. Crit. Rev. Plant Sci.* **2017**, *36* (5–6), 311–335.
37. Li, J.; Jia, X.; Yang, Y.; Chen, Y.; Wang, L.; Liu, L.; Li, M. Genome-Wide Identification of the DOF Gene Family Involved in Fruitlet Abscission in Areca Catechu L. *Int. J. Mol. Sci.* **2022**, *23* (19), 11768.
38. Song, M.; Wang, H.; Wang, Z.; Huang, H.; Chen, S.; Ma, H. Genome-Wide Characterization and Analysis of BHLH Transcription Factors Related to Anthocyanin Biosynthesis in Fig (Ficus Carica L.). *Front. Plant Sci.* **2021**, *12*, 730692.
39. Gao, C.; Sun, J.; Wang, C.; Dong, Y.; Xiao, S.; Wang, X.; Jiao, Z. Genome-Wide Analysis of Basic/Helix-Loop-Helix Gene Family in Peanut and Assessment of Its Roles in Pod Development. *PLoS One* **2017**, *12* (7), e0181843.
40. Qiu, J. R.; Huang, Z.; Xiang, X. Y.; Xu, W. X.; Wang, J. T.; Chen, J.; Song, L.; Xiao, Y.; Li, X.; Ma, J.; Cai, S. Z.; Sun, L. X.; Jiang, C. Z. MfbHLH38, a Myrothamnus Flabellifolia BHLH Transcription Factor, Confers Tolerance to Drought and Salinity Stresses in Arabidopsis. *BMC Plant Biol.* **2020**, *20* (1), 1–14. https://doi.org/10.1186/s12870-020-02732-6.
41. Heim, M. A.; Jakoby, M.; Werber, M.; Martin, C.; Weisshaar, B.; Bailey, P. C. The Basic Helix-Loop-Helix Transcription Factor Family in Plants: A Genome-Wide Study of Protein Structure and Functional Diversity. *Mol. Biol. Evol.* **2003**, *20* (5), 735–747. https://doi.org/10.1093/molbev/msg088.
42. Rivera, A. M.; Swanson, W. J. The Importance of Gene Duplication and Domain Repeat Expansion for the Function and Evolution of Fertilization Proteins. *Front. Cell Dev. Biol.* **2022**, *10*. https://doi.org/10.3389/fcell.2022.827454.
43. Wray, G. A.; Hahn, M. W.; Abouheif, E.; Balhoff, J. P.; Pizer, M.; Rockman, M. V; Romano, L. A. The Evolution of Transcriptional Regulation in Eukaryotes. *Mol. Biol. Evol.* **2003**, *20* (9), 1377–1419.
44. Kavas, M.; Baloğlu, M. C.; Atabay, E. S.; Ziplar, U. T.; Daşgan, H. Y.; Ünver, T. Genome-Wide Characterization and Expression Analysis of Common Bean BHLH Transcription Factors in Response to Excess Salt Concentration. *Mol. Genet. genomics* **2016**, *291*, 129–143.
45. Liu, M.; Ma, Z.; Sun, W.; Huang, L.; Wu, Q.; Tang, Z.; Bu, T.; Li, C.; Chen, H. Genome-Wide Analysis of the NAC Transcription Factor Family in Tartary Buckwheat (Fagopyrum Tataricum). *BMC Genomics* **2019**, *20*, 1–16.

46. Nakashima, K.; Ito, Y.; Yamaguchi-Shinozaki, K. Transcriptional Regulatory Networks in Response to Abiotic Stresses in Arabidopsis and Grasses. *Plant Physiol.* **2009**, *149* (1), 88–95.
47. Jiang, Y.; Yang, B.; Deyholos, M. K. Functional Characterization of the Arabidopsis BHLH92 Transcription Factor in Abiotic Stress. *Mol. Genet. Genomics* **2009**, *282*, 503–516.
48. Liu, W.; Tai, H.; Li, S.; Gao, W.; Zhao, M.; Xie, C.; Li, W. B HLH 122 Is Important for Drought and Osmotic Stress Resistance in A Rabidopsis and in the Repression of ABA Catabolism. *New Phytol.* **2014**, *201* (4), 1192–1204.
49. Yang, J.; Gao, M.; Huang, L.; Wang, Y.; van Nocker, S.; Wan, R.; Guo, C.; Wang, X.; Gao, H. Identification and Expression Analysis of the Apple (Malus× Domestica) Basic Helix-Loop-Helix Transcription Factor Family. *Sci. Rep.* **2017**, *7* (1), 28.
50. Kurt, F.; Filiz, E. Genome-Wide and Comparative Analysis of BHLH38, BHLH39, BHLH100 and BHLH101 Genes in Arabidopsis, Tomato, Rice, Soybean and Maize: Insights into Iron (Fe) Homeostasis. *BioMetals* **2018**, *31* (4), 489–504. https://doi.org/10.1007/s10534-018-0095-5.
51. Zhang, C.; Feng, R.; Ma, R.; Shen, Z.; Cai, Z.; Song, Z.; Peng, B.; Yu, M. Genome-Wide Analysis of Basic Helix-Loop-Helix Superfamily Members in Peach. *PLoS One* **2018**, *13* (4), e0195974.
52. Zhan, H.; Liu, H.; Ai, W.; Han, X.; Wang, Y.; Lu, X. Genome-Wide Identification and Expression Analysis of the BHLH Transcription Factor Family and Its Response to Abiotic Stress in Mongolian Oak (Quercus Mongolica). *Curr. Issues Mol. Biol.* **2023**, *45* (2), 1127–1148.
53. Letunic, I.; Bork, P. 20 Years of the SMART Protein Domain Annotation Resource. *Nucleic Acids Res.* **2018**, *46* (D1), D493–D496.
54. Xie, T.; Chen, C.; Li, C.; Liu, J.; Liu, C.; He, Y. Genome-Wide Investigation of WRKY Gene Family in Pineapple: Evolution and Expression Profiles during Development and Stress. *BMC Genomics* **2018**, *19*, 1–18.
55. Wang, Y.; Tang, H.; DeBarry, J. D.; Tan, X.; Li, J.; Wang, X.; Lee, T.; 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–e49.
56. Sudhakar Reddy, P.; Srinivas Reddy, D.; Sivasakthi, K.; Bhatnagar-Mathur, P.; Vadez, V.; Sharma, K. K. Evaluation of Sorghum [Sorghum Bicolor (L.)] Reference Genes in Various Tissues and under Abiotic Stress Conditions for Quantitative Real-Time PCR Data Normalization. *Front. Plant Sci.* **2016**, *7*, 529.