Medicinal Plant Biology

ARTICLE

Open Access

https://doi.org/10.48130/mpb-0025-0030

*Medicinal Plant Biology* **2025**, 4: e032

# The basic helix-loop-helix (bHLH) gene family in *Clinacanthus nutans*: insights into tissue-specific expression, stress responses, and methyl jasmonate (MeJA) responsiveness

Chang An[1], Denghang Wu[1], Ningwei Li[1], Yixin Yao[2], Lin Lu[1], Yan Cheng[1], Zongshen Zhang[3], Ping Zheng[1*] and Yuan Qin[1*]

[1] *Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, Center for Genomics, College of Life Science, Fujian Agriculture and Forestry University, Fuzhou 350002, China*

[2] *State Key Laboratory of Quality Research in Chinese Medicine Institute of Chinese Medical Sciences, University of Macau, Macau 519000, China*

[3] *Laboratory of Pharmaceutical Plant Cell Culture Research, School of Biological Engineering, Dalian Polytechnic University, Dalian 116034, China*

* Corresponding authors, E-mail: zhengping13@mails.ucas.ac.cn; yuanqin@fafu.edu.cn

## Abstract

*Clinacanthus nutans* is valued for its medicinal properties, largely attributed to bioactive compounds such as flavonoids, polyphenols, and other secondary metabolites. These compounds accumulate differentially across tissues and are influenced by environmental cues and phytohormones like methyl jasmonate (MeJA). Transcription factors play pivotal roles in mediating these responses, with the basic helix-loop-helix (bHLH) gene family emerging as key regulators of hormone- and stress-responsive gene expression involved in secondary metabolism. In this study, 182 *bHLHs* were identified in *C. nutans*, labeled *CnbHLH001* to *CnbHLH182* according to their chromosomal locations. Phylogenetic analysis grouped them into 26 subfamilies, with 10 classified as orphans. Structural characterization revealed subfamily-specific motif architectures and exon–intron structures. Gene duplication analysis indicated that segmental duplication was a major driver of CnbHLH family expansion, and expression analysis suggested both conserved and divergent expression among duplicated gene pairs. Tandemly duplicated genes such as *CnbHLH044* to *CnbHLH047* showed similar expression patterns. Synteny analysis showed that *CnbHLHs* in Acanthaceae species are more similar to each other than to genes in other species, suggesting a possible Acanthaceae-specific expansion. Furthermore, this study provides the first genome-wide analysis of *CnbHLHs* expression in response to abiotic stresses. Many *CnbHLHs* showed tissue-specific expression and responded differentially to MeJA and stress treatments. These findings provide a foundation for future research to explore the regulatory mechanisms of *CnbHLHs* and their potential to improve the medicinal properties of *C. nutans*.

**Citation:** An C, Wu D, Li N, Yao Y, Lu L, et al. 2025. The basic helix-loop-helix (bHLH) gene family in *Clinacanthus nutans*: insights into tissue-specific expression, stress responses, and methyl jasmonate (MeJA) responsiveness. *Medicinal Plant Biology* 4: e032 https://doi.org/10.48130/mpb-0025-0030

## Introduction

*Clinacanthus nutans,* a medicinal plant in the Acanthaceae family[1], is found throughout tropical and subtropical regions. It has significant applications in indigenous healthcare systems in Southeast Asia and southwestern China. *C. nutans* has various medicinal properties, including anti-inflammatory, antiviral, antioxidant, and anticancer activities, which have been largely attributed to its abundant accumulation of bioactive secondary metabolites such as polyphenols, triterpenes, and flavonoids[2]. Notably, the amount of these compounds varies across different tissues and is strongly affected by environmental factors[3]. For example, *C. nutans* grown in cooler temperatures has higher cytotoxicity in leaves compared with plants grown in warmer conditions[4]. Furthermore, changes in environmental conditions can lead to different morphological traits and physiological responses in *C. nutans*, which affect polyphenol and flavonoid content[5]. Plant hormones, such as methyl jasmonate (MeJA), are also known to play a key role in regulating the biosynthesis of secondary metabolites, including flavonoids, in medicinal plants[6]. Identifying candidate genes involved in the biosynthetic of active compounds like flavonoids and polyphenols in *C. nutans* is crucial for advancing molecular breeding and synthetic biology research in this species.

In plants, transcription factors (TFs) from several families, including MYB, bHLH, ERF, WRKY, NAC, bZIP, GRAS, and SPL, regulate gene expression related to stress tolerance and secondary metabolism. These TFs help plants adapt to environmental stressors and support the synthesis and accumulation of secondary metabolites that improve plant survival[7]. Among these, the bHLH family is one of the largest and plays crucial roles in plant growth and development, stress responses, and secondary metabolite biosynthesis. The basic helix-loop-helix (bHLH) is characterized by a conserved domain of about 60 amino acids[8]. This domain includes two regions: a basic region and a helix-loop-helix (HLH) region[9]. The basic region, located at the N-terminus, consists of 13–17 amino acids, typically containing around six basic residues necessary for DNA binding[10]. The HLH region, located at the C-terminus, contains approximately 40–50 amino acids, including two amphipathic alpha helices connected by a hydrophobic loop, which facilitates protein dimerization and promotes protein–protein interactions[11]. In plants, bHLHs regulate the transcription of target genes by binding to specific promoter sites, influencing processes like photomorphogenesis, flowering, and secondary metabolite biosynthesis, all of which are essential for plant adaptation and stress tolerance[12]. In *Arabidopsis thaliana*[13], AtICE1/AtbHLH116 binds to the CBF promoter region under cold stress, enhancing cold tolerance in plants overexpressing *AtICE1/AtbHLH116*. In *Vitis vinifera*[14], overexpression of *VvbHLH1* increases flavonoid accumulation and Abscisic Acid (ABA) signaling in transgenic *Arabidopsis*, conferring salt and drought resistance. Accumulating evidence shows that bHLHs are

key regulators that connect jasmonic acid (JA) signaling with downstream metabolic pathways. They play an important role in the specialized metabolism of medicinal plants and crops. In *Artemisia annua*[15], the AabHLH1 binds to E-box cis-elements in the promoters of *ADS* and *CYP71AV1*, regulating artemisinin biosynthesis. Furthermore, AabHLH113 integrates JA and ABA signaling, positively affecting artemisinin synthesis[16]. In *Catharanthus roseus*, CrMYC2, CrBIS1, and CrRMT1 regulate different branches of the monoterpenoid indole alkaloid (MIA) pathway, influencing MIA biosynthesis and accumulation[17]. In *Cucumis melo*, *CmeBrp* has been shown to regulate the biosynthesis of triterpenoid cucurbitacin compounds in roots[18]. These studies provide valuable insights into the functional roles of bHLHs in plants.

In recent years, bHLHs have been identified and analyzed on a genome-wide scale in medicinal plants such as *Salvia miltiorrhiza*[19], *Andrographis paniculata*[20], and *Gynostemma pentaphyllum*[21]. However, most research has primarily focused on model plants and crops[22]. To further the study of the bHLHs in medicinal plants, the first genome-wide identification of *bHLH* gene family members (*CnbHLHs*) in the genome of *C. nutans* was conducted. Using bioinformatics tools, the *CnbHLHs* were analyzed and visualized, examining their phylogenetic relationships, sequence features, conserved motifs, gene structure, cis-regulatory elements, and evolutionary history. Additionally, transcriptome data were used to study the expression profiles of *CnbHLHs* in different tissues and under MeJA treatment to understand their functional differentiation and regulatory patterns across developmental stages. Finally, qRT-PCR was performed to investigate the expression profiles of 20 *CnbHLHs*, providing insights into their roles under abiotic stress. This study offers a thorough analysis of *CnbHLHs*, presenting candidate genes for exploring the molecular mechanisms of abiotic stress responses and developmental processes in this species.

## Materials and methods

### Identification and characterization of *CnbHLHs* in the *C. nutans* genome

The complete genome sequence of *C. nutans* was generated by the current research group. Two approaches were used to identify all potential bHLH genes. First, the hidden Markov model (PF00010) of the bHLH domain was downloaded from the Pfam database[23] (http://pfam.xfam.org) and the HMMER web server[24] was employed to search candidate sequences within the *C. nutans* protein database, setting an E-value threshold of $10^{-5}$. Second, the AtbHLH protein sequences of *A. thaliana* were retrieved from TAIR (www.arabidopsis.org) to construct a BLAST comparison database, and a second round of candidate sequence screening conducted using the BLASTP program[25], setting the E-value at $10^{-5}$. All candidate protein sequences were subsequently merged, and redundancies were removed. To confirm that these candidate genes belong to the bHLH family, all candidate sequences were further analyzed using the NCBI Conserved Domain Database[26] (www.ncbi.nlm.nih.gov/cdd), and the SMART database[27] (http://smart.embl-heidelberg.de), ultimately identifying 182 non-redundant sequences as putative bHLH proteins of *C. nutans* (Supplementary Table S1).

### Classification and phylogenetic analysis

All identified CnbHLHs were renamed sequentially according to their chromosomal distribution in *C. nutans*. Protein physicochemical properties, including molecular weight, isoelectric point, and grand average of hydropathicity (GRAVY), were predicted using the Peptides package[28]. Subfamily classification of CnbHLHs followed the methodology proposed by Pires[29], which organizes 544 representative bHLH sequences from major plant evolutionary lineages into 26 subfamilies. The combined dataset of 182 CnbHLHs and 162 AtbHLHs was used for phylogenetic inference. Sequence alignment of the combined protein dataset was performed using MUSCLE[30], and phylogenetic tree construction was conducted using IQ-TREE[31] with the maximum likelihood (ML) method under default parameters. Visualization and annotation of the phylogenetic tree were completed using the Evolview v3 online platform[32].

### Gene structure, conserved motif, and cis-regulatory elements (CREs) analysis

The conserved motifs of CnbHLHs were analyzed using the MEME online suite[33] (http://meme-suite.org), with optimized parameters allowing any number of repeats, a maximum of 10 motifs, and an optimal motif width ranging from six to 200 amino acid residues. The intron-exon structural organization of *C. nutans* bHLH genes was annotated based on the GFF file of the *C. nutans* genome. The 2000 bp upstream sequences of each *CnbHLH* were extracted from the full-length genomic DNA using SeqKit[34], and subsequently uploaded to the PlantCARE database for cis-regulatory element prediction[35]. The prediction results were categorized in Excel, and visualized using R[36].

### Chromosomal location and gene duplication analysis

Based on the physical location information from the *C. nutans* genome database, the anchoring relationships between all *CnbHLHs* and chromosomes were visualized using the ChromoMap package. To illustrate synteny both within *C. nutans* and across species, homologous bHLHs were identified from genomic data of *C. nutans* and other selected species. The analysis included four species: *A. thaliana* (www.arabidopsis.org), *V. vinifera* (http://plants.ensembl.org), *A. paniculata* (www.ncbi.nlm.nih.gov), and *Strobilanthes cusia* (http://indigo-plant.iflora.cn), with genomic data obtained from the corresponding database websites. Gene duplication events were analyzed using MCScanX with default parameters[37], and the results were visualized in the Advanced circos module of TBtools[38]. For Ka/Ks analysis, nonsynonymous substitution rate (Ka), synonymous substitution rate (Ks), and Ka/Ks ratios between homologous gene pairs in *C. nutans* and other species were calculated using the KaKs calculator[39].

### Prediction of protein–protein interaction network and homology modeling analysis

All submitted CnbHLH protein sequences were uploaded to the STRING website[40] (http://string-db.org), and Arabidopsis orthologs were chosen as reference sequences. After the BLAST analysis, the highest-scoring gene (bitscore) was used to construct the network, and genes without interactions with others were eliminated. Protein models homologous to the CnbHLHs were retrieved from the PDB database[41] (www.rcsb.org). Protein secondary structures were then predicted through NetSurfP-3.0[42] (https://services.healthtech.dtu.dk/services/NetSurfP-3.0), and SWISS-MODEL (www.swissmodel.expasy.org) was used for homology modeling to predict the tertiary structure[43]. Additionally, ConSurf (https://consurf.tau.ac.il) was employed to examine conserved regions. Finally, the predicted model structures were visualized and refined using PyMOL[44].

### Plant culture and stress treatments

The *C. nutans* utilized in this study were cultivated at Fujian Agriculture and Forestry University (FAFU) (Fujian, China). Approximately 30 cm-long stem cuttings were obtained from tall and vigorous mother plants and propagated through stem cuttings. Once rooted and showing vigorous growth, each plant was individually transplanted into pots filled with a soil and peat mixture at a ratio of 2:1. Plants were grown under greenhouse conditions for 60 d (28 °C,

75% relative humidity, and a photoperiod of 16 h light/8 h dark). At the beginning of the treatments, plants of uniform growth status were randomly grouped for stress and hormone treatments, with three biological replicates per group. The experimental design included five treatment conditions: drought, salinity, cold, heat, and MeJA. Drought stress was simulated using 150 mmol/L mannitol solution, and salinity stress was induced by 150 mmol/L NaCl solution. For temperature stress, plants were exposed to 8 °C (cold) and 45 °C (heat). MeJA treatment was performed by foliar spraying with 200 μmol/L MeJA solution until leaf surfaces were fully covered and dripping. Samples were collected from roots and leaves at four time points (0, 24, 48, and 72 h) post-treatment. Additionally, organ-specific samples (root, stem, young leaf, mature leaf, and flower) were collected from flowering plants under normal growth conditions. All samples were collected in three biological replicates, immediately frozen in liquid nitrogen, and stored at −80 °C for further analysis.

## Transcriptomic and qRT-PCR analyses

Transcriptome sequencing was performed on organ-specific samples (root, stem, young leaf, mature leaf, and flower) as well as MeJA-treated root and leaf samples collected at 0, 24, 48, and 72 h. All RNA-seq experiments, including RNA extraction, library construction, and sequencing, were conducted by Biomarker Technologies Co., Ltd (Beijing, China). Approximately 6.7 Gb of raw sequencing data were generated per sample and subsequently assessed using FastQC. Low-quality reads and adapter sequences were removed using fastp[45], resulting in high-quality clean reads. Clean reads were aligned to the *C. nutans* reference genome using HISAT2[46]. Transcript abundance was quantified using the featureCounts program in the Subread package[47]. Expression levels of *CnbHLHs* were normalized as TPM (Transcripts Per Kilobase Million), and a heatmap based on $\log_2$(TPM + 1) was generated using R. All RNA-seq data have been deposited in the China National GeneBank DataBase (CNGBdb) under accession number: CNP0006103.

To assess the expression of *CnbHLHs* under abiotic stress, qRT-PCR was performed using samples treated with cold (4 °C), heat (45 °C), drought (150 mmol/L mannitol), and salinity (150 mmol/L NaCl). Untreated plants served as the control group. Samples were collected from roots and leaves at 0, 24, 48, and 72 h post-treatment. Total RNA was extracted using the Trizol method (Invitrogen, Carlsbad, CA, USA), and first-strand cDNA synthesis was carried out using the ThermoScript RT-PCR Kit (Thermo Fisher Scientific, USA). Gene-specific primers were designed using Primer-BLAST, and primer sequences are listed in Supplementary Table S2. The qRT-PCR program was: 95 °C for 30 s; followed by 40 cycles of 95 °C for 5 s, 60 °C for 34 s, and 95 °C for 15 s. The *CnRPL* was used as an internal control to normalize expression levels[48]. The fold change in gene expression was calculated using the $2^{-\Delta\Delta C_T}$ method.

# Results

## Identification, characterization, and phylogenetic classification of bHLHs in *C. nutans*

A total of 182 putative bHLH family members were identified in the *C. nutans* genome, which were sequentially named from *CnbHLH001* to *CnbHLH182* based on their chromosomal positions. To gain a comprehensive understanding of the physicochemical properties of the CnbHLH proteins, the predicted characteristics of the proteins encoded by each *CnbHLH* gene were analyzed (Fig. 1, Supplementary Table S3). The results indicated that the predicted protein lengths of all *CnbHLH*s ranged from 82 (*CnbHLH118*) to 1,074 (*CnbHLH078*) amino acids, with molecular weights spanning 9.39 to 118.90 kDa. The isoelectric points (pI) varied from 4.47 (*CnbHLH020*) to 12.29 (*CnbHLH100*); of these, 88 members predominantly contained basic amino acids (pI > 7), while 93 members were primarily composed of acidic amino acids (pI < 7), with *CnbHLH175* exhibiting a pI of 7, representing a neutral protein. The GRAVY values ranged from −0.94 to 0.1, with the majority being hydrophilic proteins; only *CnbHLH088* (0.02) and *CnbHLH136* (0.1) exhibited a tendency toward hydrophobicity. The Instability Index values ranged from 27.12 (*CnbHLH108*) to 98.91 (*CnbHLH087*), with 166 (91.21%) members exceeding a value of 40, indicating that most proteins are relatively unstable and may be more susceptible to structural changes or degradation *in vitro*.

An unrooted maximum likelihood (ML) phylogenetic tree, based on the complete protein sequences of 182 *CnbHLHs* from *C. nutans* and 162 *AtbHLHs* from *A. thaliana,* clearly depicted the evolutionary relationships among the bHLH proteins. A total of 172 *CnbHLH*s from *C. nutans* were classified into 26 subfamilies, and the remaining 10 were grouped as Orphans, showing no significant similarity or evolutionary relationship with other family members (Fig. 2). The tree's topology revealed considerable diversity and variation in the distribution of *CnbHLHs* across subfamilies. Subfamily XII with 23 genes, accounting for 12.64% of all *CnbHLHs*. Subfamilies Ib(2) and X had 10 or more members, while subfamilies VIIIa, VIIIc(1), and XIV each had only one member. Most subfamilies (20, comprising 76.92% of all subfamilies) had between two and 10 members.

Chromosomal localization analysis showed that *CnbHLHs* were distributed across all 18 chromosomes (Chr) of *C. nutans*. Visualization showed an uneven distribution of *CnbHLHs* on these chromosomes, with a tendency for clustering at the ends (Supplementary Fig. S1a). Ten chromosomes contained 10 or more *CnbHLHs*, and the remaining eight chromosomes had fewer than 10 members. Chr7 had the highest number, with 16 *CnbHLHs*, whereas Chr8 and Chr12 each contained only four members. Also, certain chromosome pairs shared the same number of *CnbHLHs*, including Chr1 and Chr18, Chr2 and Chr16, Chr3 and Chr14, Chr4 and Chr6, Chr8 and Chr12, Chr9 and Chr15, and Chr10 and Chr17 (Supplementary Table S3). Considering the subfamily classifications of *CnbHLHs*, no specific subfamily members were found to cluster on a single chromosome (Supplementary Fig. S1b), indicating a random distribution that may reflect their roles across different biological functions and regulatory networks.

## Conserved domains and motifs, gene structure of *CnbHLHs*

To further explore gene structure in different subgroups, *CnbHLHs* gene structures were analyzed and visualized, including conserved domains, motifs, and exon–intron compositions. The bHLH domains in *C. nutans* ranged from 37 to 88 residues, with gaps inserted to optimize alignment across conserved regions. Multiple sequence alignment identified 113 conserved positions in the bHLH domain (Fig. 3a, Supplementary Table S4). Positions 1–13 corresponded to the basic amino acid region, 14–55 to the helix1 region, 56–71 to the loop region, and 72–113 to the helix2 region. Conserved residues, such as H5/Q5, E9/A9, and 13R, essential for E-box binding, matched previous reports[49] (Fig. 3b). Residues 10R and 12R showed significant conservation, with alignment rates of 75.80% and 86.30%, respectively. In helix1, P55—a critical residue for terminating the first helix and linking to the variable-length loop—was conserved (86.30%). Several hydrophobic residues, including I31, R34, L38, L41, and V54, were also conserved, contributing to the stability of the bHLH dimeric structure (Fig. 3c). In the loop region, two conserved residues, K69 and D71, were found. In helix2, several hydrophobic residues (A73, L76, A79, I86, L92, V96, and L113) were conserved,

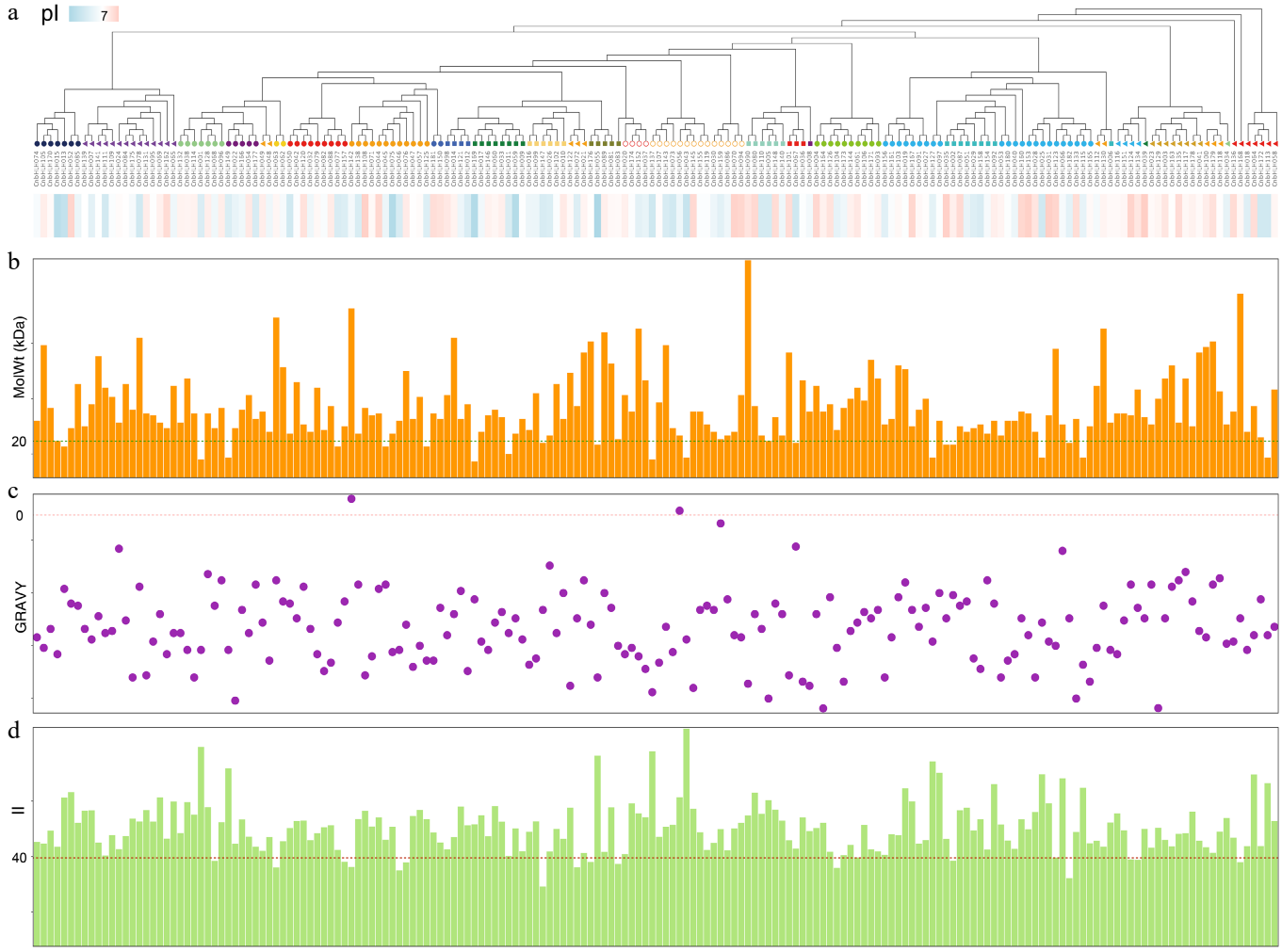


**Fig. 1** Physicochemical properties of *CnbHLHs*. (a) Phylogenetic tree of the CnbHLHs constructed based on full-length amino acid sequences, with subfamily clades color-coded. The bar below indicates the isoelectric point (pI) values of each protein, ranging from acidic (blue) to basic (red). (b) Bar plot showing the molecular weight (MolWt) distribution of CnbHLHs in kilodaltons (kDa). (c) GRAVY (Grand Average of Hydropathy) index of each CnbHLH. Most proteins exhibited negative GRAVY values, indicating hydrophilicity. (d) Bar chart of instability index (II) values. The dashed line at II = 40 denotes the threshold above which proteins are considered potentially unstable.

along with D71, S74, E78, and Y88, highlighting the structural stability of these domains.

Motifs are important for mediating interactions and signal transduction between different modules during gene transcription[50]. To investigate these motifs, MEME was used to analyze the conserved motifs in *CnbHLHs*. The composition patterns of the 10 identified motifs were largely consistent with the results of the present phylogenetic tree and gene structure analysis, showing high similarity within the same subfamily but significant differences between subfamilies (Supplementary Fig. S2b, Supplementary Table S5). The number of motifs in the 182 *CnbHLHs* ranged from one (*CnbHLH100*, *CnbHLH112*, *CnbHLH125*) to seven (*CnbHLH060*, *CnbHLH098*, *CnbHLH121*). For motif distribution, motif 1 and motif 2 were present in 97.25% and 98.90% of *CnbHLHs*, respectively. Except for five orphan genes (*CnbHLH130*, *CnbHLH048*, *CnbHLH049*, *CnbHLH100*, *CnbHLH112*), all other *CnbHLHs* contained motif 1. *CnbHLH006* and *CnbHLH125* were the only two genes lacking motif 2. Additionally, all members of subfamily XII contained motif 3, and all members of subfamily XIII contained motif 6. Notably, certain motifs were exclusive to specific groups; for example, motif 10 was found only in group VIIIb, while the combination of motif 7 and motif 8 was combined only in group III. This pattern reflects the conservation and specificity of motifs within particular subfamilies.

The diversity of exon-intron structures is critical to the evolution of multigene families[51]. The number of introns and exons in each gene were analyzed and subsequently visualized the gene structural characteristics of the *CnbHLHs* (Supplementary Fig. S2c, Supplementary Table S5). Among the 182 *CnbHLHs*, the number of exons ranges from one to 22, with 167 *CnbHLHs* (91.76%) containing multiple exons, and 30 genes (16.48%) with three exons, representing the highest proportion. Additionally, 15 CnbHLHs (8.24%) were intron-less, predominantly found in subfamilies III(d + e) and VIIIb, such as *CnbHLH059*, *CnbHLH017*, and *CnbHLH100*. Structural diversity was evident among members of different subfamilies, including variations in intron and exon counts and their relative positions. Conversely, genes within the same subfamily exhibit highly similar intron/exon patterns; for instance, members of subfamily Ia contain 2–3 exons, while those in subfamily III(a + c) contain 4–5 exons, and nearly all members (22/23) of subfamily XII have 5–9 exons. These findings align with the general structural characteristics observed in most plant bHLH family genes.

## Cis-regulatory elements (CREs) analysis of *CnbHLHs*

Conserved motifs located in gene promoter regions act as recognition and binding sites for proteins. Using PlantCARE, the putative promoter sequences of 182 *CnbHLHs* were analyzed, identifying

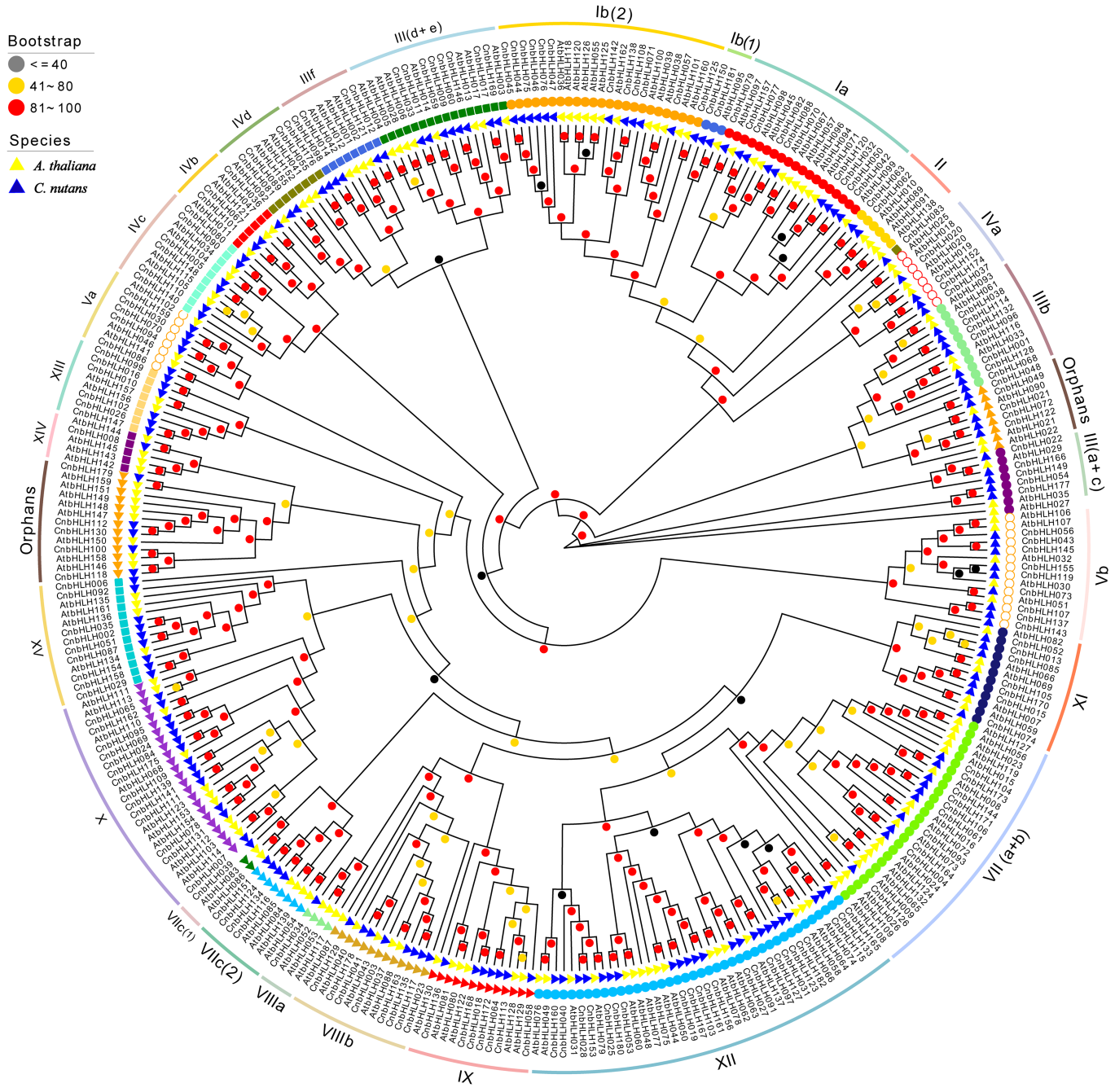


**Fig. 2** Phylogenetic relationship of bHLHs between *C. nutans* and *A. thaliana*. The unrooted evolutionary tree was constructed using IQTree with the maximum likelihood (ML) method. The 182 CnbHLHs are classified into 26 subfamilies and one Orphan group, following the classification method used of AtbHLHs, represented by different colored subgroups within the tree. The bootstrap value is displayed a circle on each branch, with each color corresponding to a specific bootstrap value range.

4,672 CREs belonging to 42 types (Supplementary Table S6). These CREs were divided into three main groups based on function: plant growth and development (26 types, 2,765 CREs), plant hormone response (nine types, 1,282 CREs), and stress response (seven types, 625 CREs) (Fig. 3d). Among growth-related elements, light-responsive elements were the most common (2,504, accounting for 90.56%), widely distributed across nearly all *CnbHLHs*' promoter region (e.g., CAG-motif, GT1-motif, GATA-motif, G-box, ACE, and Box I). For hormone response elements, MeJA-responsive elements (508, 39.63%) and abscisic acid (ABA)-responsive elements (437, 34.09%) were the most abundant, showing enrichment in subfamily XII. In stress response elements, anaerobic induction elements (293, 46.88%) were the most frequent, with ARE elements being the most prominent in all subfamilies (Fig. 3e). Overall, the composition and number of CREs of different *CnbHLHs* vary significantly between and within subfamilies, suggesting that the expression of *CnbHLHs* is regulated by a diverse set of CREs associated with hormones, growth, development, and stress responses.

## Gene duplication and synteny analysis of *CnbHLHs*

Gene duplication events are important for gene family evolution, with family expansion usually driven by tandem and segmental duplications. To understand the amplification mechanisms of *CnbHLHs*, gene duplication events in *C. nutans* were analyzed, revealing 97 gene duplication pairs (Fig. 4a, Supplementary Table

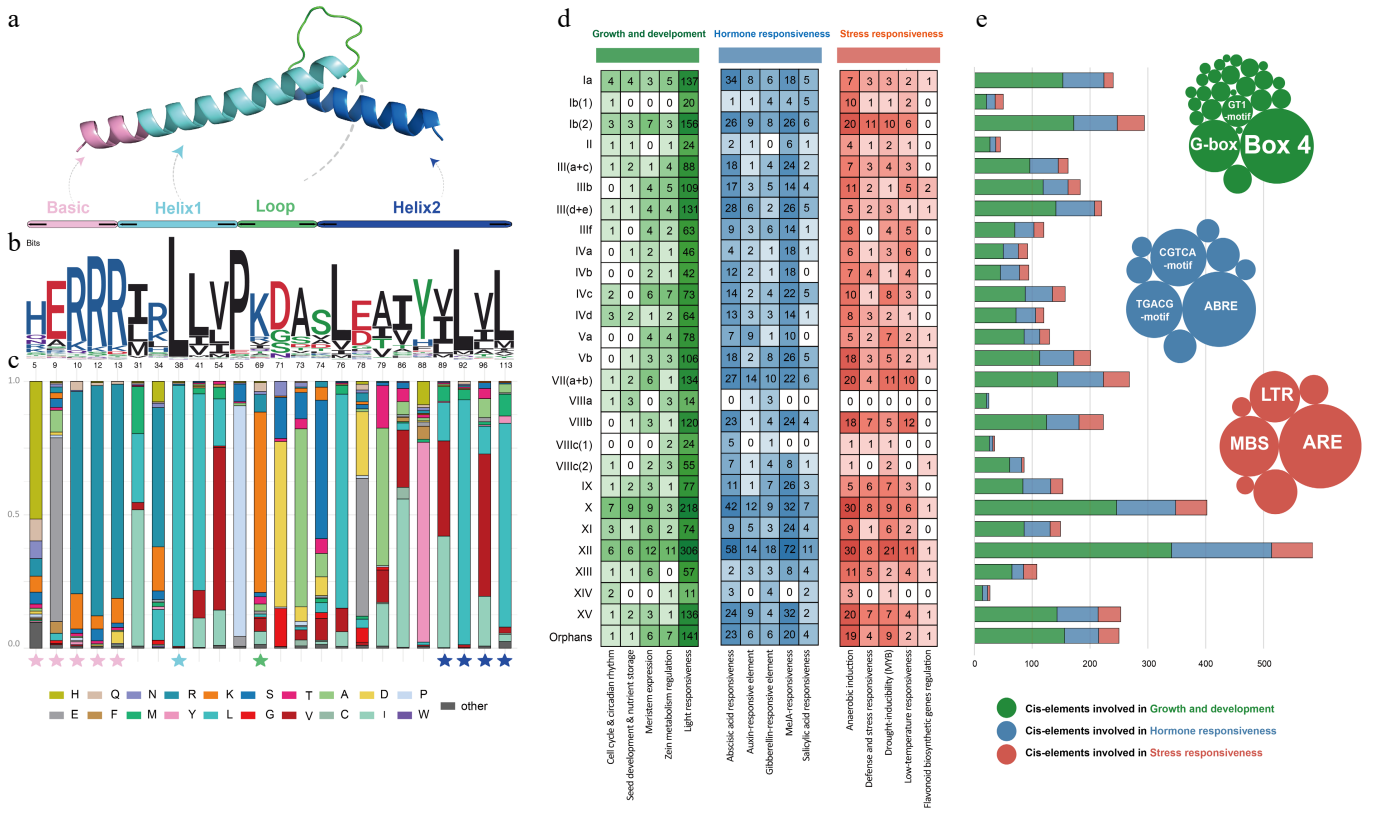


**Fig. 3** Structural characteristics and CRE feature of CnbHLHs. (a) Three-dimensional protein structure of the bHLH domain. (b) Sequence logo of the CnbHLHs domain. The numbers at the bottom represent the positions of residues in the alignment. (c) Distribution of amino acids in the CnbHLHs conserved sites. The pentagrams below the bar chart indicate the positions of highly conserved residues. (d) The identified CREs were grouped into three functional categories, shown in green, blue, and red shades. The numbers show the quantity of corresponding CREs in the promoter regions of each subfamily of CnbHLHs. (e) Statistics for the total number of CREs. The bubble chart displays the number of all elements under each functional category, with the top three most abundant elements labeled.

S7). Among them, five pairs were tandem duplications. *CnbHLH044–CnbHLH047* formed three consecutive tandem duplication pairs on Chr4, with one additional pair on Chr6 and Chr7. All these duplicated genes belong to the same subfamily. The other 92 pairs were segmental duplications, suggesting that segmental duplication plays a key role in *CnbHLHs* expansion and evolution. Notably, only three segmental duplication pairs are on the same chromosome: *CnbHLH139/CnbHLH141* (Chr15), *CnbHLH154/CnbHLH158* (Chr16), and *CnbHLH156/CnbHLH161* (Chr16). These inter-chromosomal segmental duplications suggest recent chromosomal rearrangements in this species, defining potential hotspots for further rearrangements.

To investigate the evolutionary trajectory of family members, four representative species were selected for synteny analysis with *C. nutans* (Fig. 4b, Supplementary Table S8): *A. thaliana* (Brassicacea), *V. vinifera* (Vitaceae), and two closely related species, *A. paniculata* and *S. cusia*. A total of 138 *CnbHLHs* exhibited synteny with *A. paniculata*, followed by 129 *CnbHLHs* with *S. cusia*, 106 *CnbHLHs* with *A. thaliana*, and 82 *CnbHLHs* with *V. vinifera*. Notably, 37 *CnbHLHs* displayed syntenic relationships with all four species. Among these, 102 syntenic genes were shared across the three Acanthaceae species, while only 53 and 37 genes showed synteny with *V. vinifera* and *A. thaliana*, respectively. This suggests possible lineage-specific expansions within Acanthaceae, potentially linked to specific evolutionary or adaptive needs of the family. Furthermore, to explore the forces driving gene evolution, a Ka/Ks analysis of homologous genes between *C. nutans* and the four species was performed (Fig. 4c, Supplementary Table S9). The frequency of nonsynonymous mutations was lower than that of synonymous mutations, with Ka/Ks ratios below 1, indicating that these *CnbHLHs* underwent strong purifying selection during evolution.

## Protein–protein interaction networks analysis

bHLH proteins typically function by forming homodimers or heterodimers, either among bHLH proteins or between bHLH and non-bHLH proteins. A protein interaction network was constructed for CnbHLHs based on collinearity, phylogenetic analysis, and alignment scores, using homologous proteins from *A. thaliana* as references (Fig. 5a, Supplementary Table S10). This network includes 408 interactions, with 90 showing high reliability (combined score > 0.7). These 408 interactions form 59 nodes. Among these nodes, only one protein pair shows a single interaction mode (CnbHLH053 and CnbHLH093), while other nodes have numerous connections (> 2) to various nodes, indicating complex multi-gene interactions. CnbHLH038 has the most interaction partners (16), though only three interactions have high binding scores: CnbHLH038–CnbHLH058 (0.852), CnbHLH038–CnbHLH079 (0.761), and CnbHLH038–CnbHLH153 (0.852). The CnbHLH168 node shows the highest interaction reliability, with five of eight interactions having high reliability: CnbHLH168–CnbHLH059 (0.72), CnbHLH168–CnbHLH175 (0.758), CnbHLH168–CnbHLH139 (0.881), CnbHLH168–CnbHLH001 (0.719), and CnbHLH168–CnbHLH121 (0.777). CnbHLH075, CnbHLH155, CnbHLH176, CnbHLH139, CnbHLH141, CnbHLH175, CnbHLH008, CnbHLH036, CnbHLH090, and CnbHLH140 each has more than 10 interactions, suggesting they may play key roles in specific functions or pathways. These findings further highlight the functional diversity of CnbHLHs.

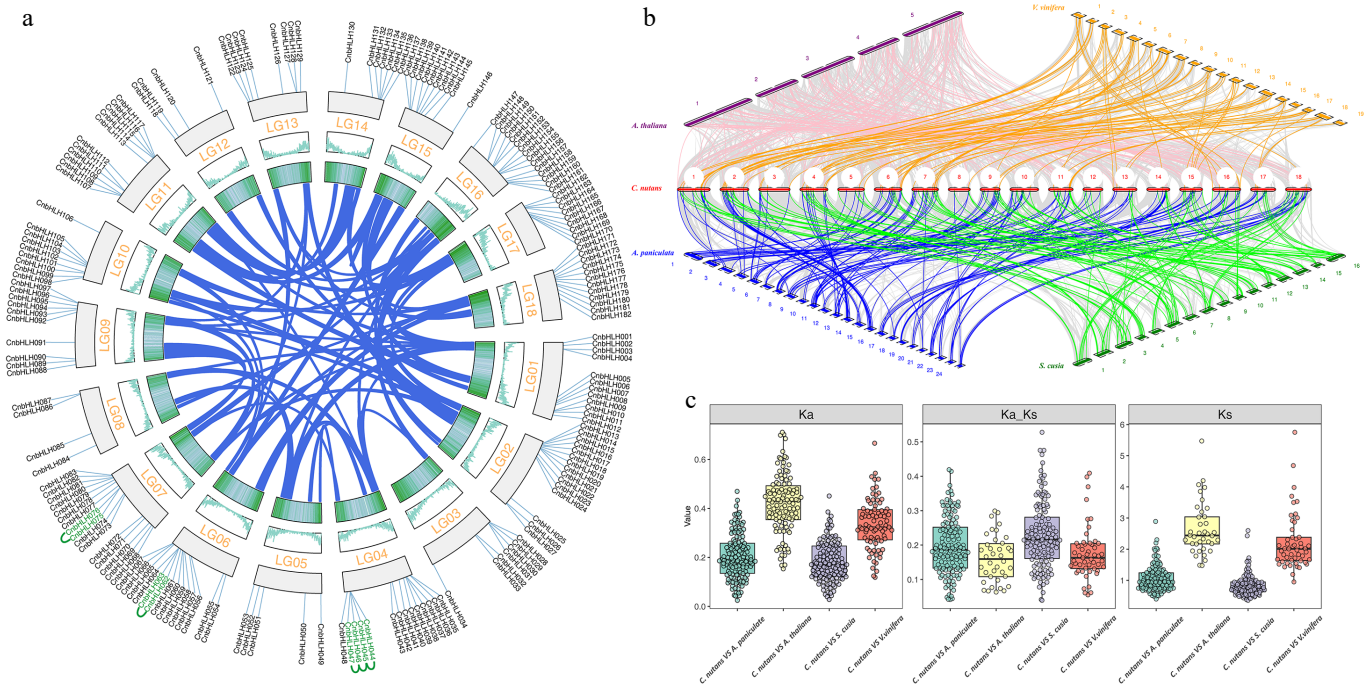


**Fig. 4** Evolutionary analysis of the *bHLH* gene family between *C. nutans* and other species. (a) Collinearity analysis of *CbHLHs*. (b) *bHLHs* collinearity between *C. nutans* and other species (*A. thaliana*, *V. vinifera*, *A. paniculata*, *S. cusia*) genomes. (c) Ka/Ks analysis of *bHLHs* between *C. nutans* and closely related species (*A. paniculata*, *S. cusia*).

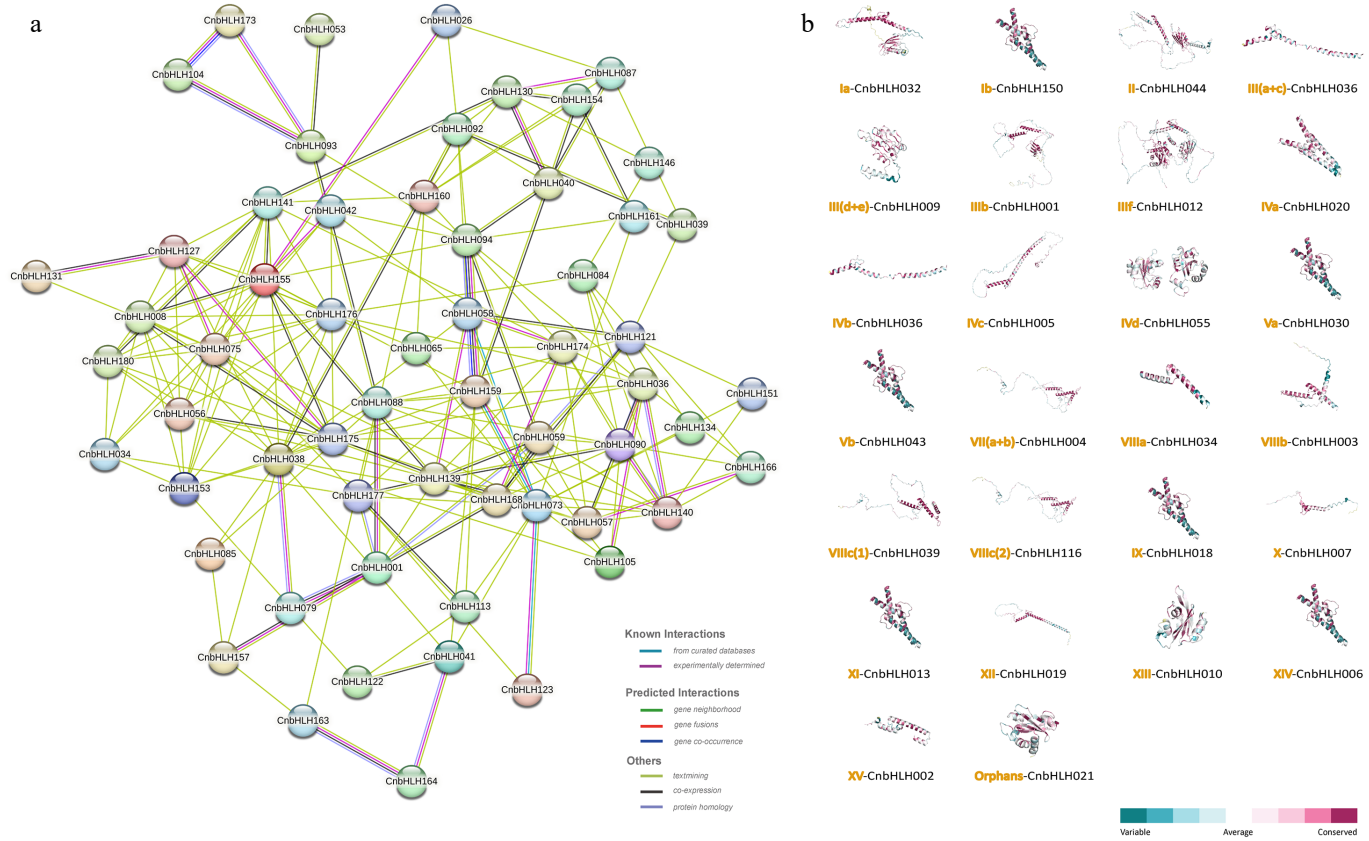


**Fig. 5** Protein interaction network and predicted 3D structure of CnbHLHs. (a) Protein–protein interaction network of CnbHLHs based on homologous proteins in *A. thaliana*. (b) Predicted 3D structures of CnbHLHs, showing only the member with the highest GMQE value from each subfamily.

Secondary structure prediction results show that random coil structures dominate in CnbHLHs (35.19%–91.90%), while $\alpha$-helix (7.48%–56.64%), and $\beta$-sheet (0.13%–20.24%) contents are relatively lower (Supplementary Table S11). The disorder scores of residues in these proteins range from 11.52 to 86.57, suggesting that CnbHLHs have structural flexibility while maintaining stability. Moderate Relative Solvent Accessibility (RSA) values (36.46–64.68) suggest that these proteins are involved in functional interactions.

Subsequently. Homology modeling of CnbHLHs was performed using the SWISS-MODEL database, predicting three-dimensional structures for 182 CnbHLHs. CnbHLHs in different subfamilies showed diverse three-dimensional structures, while proteins in closely related subfamilies had similar 3D structures. The structure with the highest GMQE and QMean scores in each subfamily was chosen as the representative model for CnbHLHs (Fig. 5b, Supplementary Table S12). Subfamily XV showed the highest GMQE value (0.8), indicating high model reliability. Subfamilies Ib(1) and III(a + c) had multiple genes with GMQE values below 0.1, suggesting lower model reliability and the need for further optimization and validation. Overall, these homology models provide a basis for understanding the molecular functions of CnbHLHs.

## Expression analysis

To investigate the functions of *CnbHLH*s, the expression profiles of 182 *CnbHLH*s across different tissues of *C. nutans*, including root, stem, young leaf, mature leaf, and flower, were analyzed using RNA-seq data (Fig. 6a, Supplementary Table S13). After filtering out genes with low expression levels (TPM < 30), 89 *CnbHLH*s were retained for further analysis. Many of these genes showed tissue-specific expression, suggesting that *CnbHLH*s are involved in tissue- or developmental stage-specific functions. Hierarchical clustering grouped the *CnbHLH*s into five blocks (1–5), with genes in each block showing unique expression patterns: (1) *CnbHLH*s in block a were mainly expressed in the root, with *CnbHLH015*, *CnbHLH074*, *CnbHLH030*, and *CnbHLH017* also showing high expression in other tissues. (2) Most *CnbHLH*s in block b were highly expressed in mature leaves, while *CnbHLH011*, *CnbHLH172*, *CnbHLH061*, and *CnbHLH008* exhibited high expression in all tissues. (3) *CnbHLH*s in block c were mainly expressed in young leaves, with *CnbHLH101* and *CnbHLH059* also showing higher expression in other tissues. (4) *CnbHLH*s in block d were mainly expressed in flowers, with *CnbHLH148*, *CnbHLH110*, *CnbHLH126*, *CnbHLH130*, and *CnbHLH140* also showing high expression in other tissues. (5) *CnbHLH*s in block e were highly expressed in the stem, with more than half also showing abundant expression in the leaf. These results also highlight that different *CnbHLH*s have diverse expression profiles within and between subfamilies.

Studies show that MeJA treatment can modulate plant metabolic processes, influencing the synthesis and accumulation of bioactive compounds like flavonoids and triterpenes in medicinal plants[52]. Previous studies show that bHLHs play important roles in jasmonic acid (JA) signaling pathways[53]. To explore the roles of *CnbHLH*s in *C. nutans* in response to MeJA, the expression patterns of these genes in the leaves and roots were analyzed at different time points after MeJA treatment using transcriptome data. After filtering out genes with low expression, 39 genes remained in the leaf group (Fig. 6b), and 44 genes remained in the root group (Fig. 6c). Notably, the gene composition in the two groups differed, with only 19 genes in both tissues, suggesting distinct *CnbHLH*s responses to MeJA treatment in each tissue. The remaining *CnbHLH*s that were differentially expressed in the leaves under MeJA treatment were grouped into four groups (blocks I–IV) based on hierarchical clustering (Fig. 6b). Genes in block I were highly expressed at the pre-treatment (0 h) and early treatment (24 h) stages, with expression decreasing over time. In contrast, most genes in block II showed high expression across all time points, with *CnbHLH011* and *CnbHLH093* mainly expressed before treatment, while *CnbHLH126*, *CnbHLH101*, and *CnbHLH148* showed high expression at both the pre-treatment and final time points (72 h), following a valley-shaped pattern. Genes in block III were mainly expressed during the later stages of treatment (48 and 72 h), while genes in block IV peaked at the mid-treatment stages (24 and 48 h). In the roots, the remaining *CnbHLH*s were grouped into four clusters (blocks i–iv) (Fig. 6c). Genes in block i were upregulated by MeJA and showed higher expression at the later time points (48 and 72 h). Genes in block ii were mainly expressed during the mid-treatment period (24 and 48 h), showing a peak-shaped expression pattern. Genes in block iii showed high expression at the pre-treatment (0 h), and early treatment (24 h) time points, with expression decreasing later. Most genes in block iv were mainly expressed before treatment (0 h), with their expression suppressed after MeJA induction; however, genes like *CnbHLH038*, *CnbHLH130*, *CnbHLH177*, *CnbHLH101*, *CnbHLH050*, and *CnbHLH052* also showed higher expression later. A small subset of genes in this block, including *CnbHLH030*, *CnbHLH047*, and *CnbHLH132*, were mainly expressed at the later time points. These findings show that the *CnbHLH*s in the leaves and roots of *C. nutans* have diverse expression patterns in response to MeJA treatment, suggesting that different *CnbHLH*s may play distinct roles at various time points during treatment.

## Expression patterns of *CnbHLH*s under abiotic stress

Numerous studies have shown that bHLHs play a critical role in abiotic stress response pathways, influencing the accumulation of

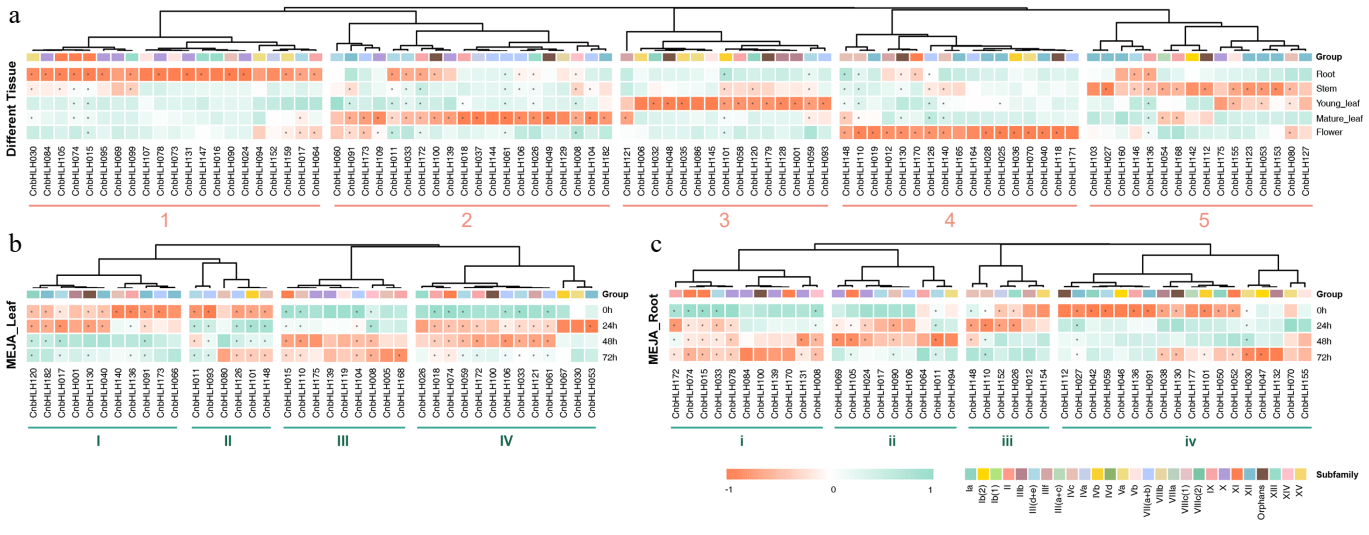


**Fig. 6** The heat maps of gene expression of *CnbHLH*s in different tissues and the MeJA treatment. (a) The heat map of *CnbHLH*s expressionin different tissues. (b) The heat map of *CnbHLH*s expression in leaves after MeJA treatment. (c) The heat map of *CnbHLH*s expression in roots after MeJA treatment.

secondary metabolites[12,54]. To explore the roles of *CnbHLHs* in response to various stress conditions in *C. nutans*, qRT-PCR was used to analyze the expression patterns of 20 representative *CnbHLHs* at different time points after exposure to cold (8 °C), heat (45 °C), salt (150 mM NaCl), and drought (150 mM mannitol) (Fig. 7, Supplementary Table S14). The results showed gene members from the same subfamily exhibited different response patterns to the same stress. Additionally, significant differences were observed in the expression patterns of the same genes between leaves and roots under identical treatment conditions, suggesting distinct responses in aboveground and underground tissues. Under cold stress, 15 out of the 20 *CnbHLHs* were significantly upregulated in the leaves, with expression initially increasing, followed by a decrease, and then an eventual increase. Several *CnbHLHs* showed significant upregulation at specific time points during cold stress, including *CnbHLH148* (15-fold), *CnbHLH18* (14-fold), *CnbHLH104* (28-fold), *CnbHLH126* (8-fold), *CnbHLH173* (12-fold), and *CnbHLH8* (14-fold). Under the same cold stress, 12 out of the 20 *CnbHLHs* showed decreased expression in the roots, while *CnbHLH126* and *CnbHLH173* were upregulated, showing an increasing expression trend. These genes were also upregulated in the leaves under cold stress, suggesting that they may play important roles in *C. nutans* response to cold stress. Under heat stress treatment, 14 out of the 20 *CnbHLHs* exhibited decreased expression in the leaves, while 15 *CnbHLHs* showed reduced expression in the roots, with variations between the two tissues. Notably, *CnbHLH126* was initially repressed under heat stress but was upregulated over time in the root. Under salt stress treatment, 12 out of the 20 *CnbHLHs* showed decreased expression in the leaves, with only *CnbHLH78* gradually upregulated. Interestingly, in the roots, only five *CnbHLHs* were downregulated post salt stress treatment, while most genes were upregulated at different time points. Under drought stress, 14 out of the 20 *CnbHLHs* showed decreased expression in the leaves, while 14 *CnbHLHs* were upregulated in the roots. Notably, the expression of *CnbHLH11*, *CnbHLH17*, *CnbHLH30*, *CnbHLH110*, *CnbHLH140*, *CnbHLH148*, *CnbHLH130*, *CnbHLH61*, *CnbHLH104*, and *CnbHLH173* increased at all time points post-treatment, with *CnbHLH11*, *CnbHLH140*, and *CnbHLH104* showing sustained upregulation throughout the treatment. This analysis shows that *CnbHLHs* in *C. nutans* exhibit tissue-specific and dynamic responses to various abiotic stresses, suggesting their role in regulating stress adaptation mechanisms.

A diagram was made to show the expression patterns of *CnbHLHs* in different tissues and after MeJA or stress treatments (Fig. 8). The expression patterns of *CnbHLHs* was different between tissues, which may mean they help control various growth and development processes. Hormones like MeJA help control the accumulation of secondary metabolites in plants[6]. The present analysis showed big differences in how *CnbHLHs* in the leaves and roots reacted to MeJA treatment, with only a few showing similar patterns in both tissues. This suggests that *CnbHLHs* may regulate metabolic pathways in different tissues. Stress conditions cause production and accumulation of secondary metabolites in plants[55]. qRT-PCR was used to check how 20 representative *CnbHLHs* were expressed under different stress conditions. The results showed that, for the same stress, the same genes were expressed differently in the roots and leaves, meaning stress may affect the metabolism of the above- and below-ground parts of the plant differently. The medicinal properties of different parts of *C. nutans* are related to differences in the accumulation of active compounds[56]. These results show that different *CnbHLHs* may regulate different metabolic processes in the above- and below-ground tissues. The results provide valuable insights for further functional studies on *CnbHLHs* and suggest potential gene targets for molecular breeding to enhance the accumulation of active compounds in *C. nutans*.

## Discussion

*C. nutans* has gained attention for its potential medicinal compounds, such as flavonoids, polyphenols, and other secondary metabolites[57]. The accumulation of these metabolites vary in different tissues and are influenced by environmental factors and hormones like MeJA[6]. bHLHs regulate key processes, including plant growth and development, stress and hormone responses, and secondary metabolism[58]. However, the role of bHLHs in *C. nutans* remains unexplored. In this study, 182 bHLH genes in *C. nutans* were identified and named *CnbHLH001–CnbHLH182* based on their chromosomal positions. These genes were unevenly distributed across the 18 chromosomes of *C. nutans*, with noticeable clustering at the chromosome ends (Supplementary Fig. S1). The number of *CnbHLHs* (n = 182) was similar to that in rice (n = 183)[22], but higher than in *A. thaliana* (n = 162)[59], *Prunus mume* (n = 100)[60], and *Passiflora edulis* (n = 138)[61]. The differences in *bHLH* gene numbers across species may be linked to genome size or gene duplication events[62]. The phylogenetic tree of *C. nutans* and *Arabidopsis* bHLH proteins further classified the 182 *CnbHLHs* into 26 subfamilies, with 10 additional *CnbHLHs* classified as orphans (Fig. 2). Conserved motifs and exon/intron analyses showed structural diversity among *CnbHLHs* from different subfamilies, while members within the same subfamily were more similar (Supplementary Fig. S2), confirming the accuracy of subfamily classification. Motif 1 and motif 2 were conserved in almost all CnbHLHs and were major components of the bHLH conserved domain, indicating their key role in the functional specificity of these TFs. Other motifs showed varying distributions across CnbHLHs, but many proteins in the same subfamily shared similar conserved motifs. For example, motif 7 and motif 8 were found only in subfamily III, and motif 10 was specific to group VIIIb. The number of exons in the *CnbHLHs* ranged from 1 to 22, reflecting the structural complexity of the bHLH family. Similarly, genes within the same subfamily had similar exon/intron patterns, such as subfamily Ia members having 2–3 exons, and subfamily III(a + c) members having 4–5 exons. Notably, some *CnbHLHs* were intron-less, mainly in subfamilies III(d + e) and VIIIb, including *CnbHLH059*, *CnbHLH017*, and *CnbHLH100* (Supplementary Fig. S2, Supplementary Table S5). This feature might give these genes an advantage in rapidly responding to biological or environmental signals[63]. Transcriptome data showed that these genes were rapidly induced in response to MeJA treatment in leaves, with *CnbHLH017* also showing a rapid response in roots (Fig. 7). Studies in *Arabidopsis* and maize have shown that intron-less genes exhibit rapid expression responses to light and temperature stimuli[64], suggesting that such *CnbHLHs* might help *C. nutans*'s adapt quickly to changing environments. The 3D structural predictions reveal varied features across subfamilies (Fig. 5b), providing a basis for the functional diversity of family members while maintaining core roles. Analysis of physicochemical properties also showed significant differences among CnbHLHs (Fig. 1, Supplementary Table S3), suggesting that these proteins may have distinct roles in their microenvironments[65].

Gene duplication is prevalent in plant genomes and contributes significantly to evolution and adaptation in response to environmental changes[62]. In *C. nutans*, 97 gene duplication pairs were found(Fig. 4a, Supplementary Table S7), which include five tandem duplicates and 92 segmental duplicates. These results suggested that *CnbHLHs* expansion mainly occurred through segmental duplication, similar to other species[66]. Some duplicated genes with similar structures and expression patterns may have functional redundancy. For instance, the segmental duplicated gene pair *CnbHLH061* and *CnbHLH106*, which were clustered in subfamily VII(a + b), both have motif 1, motif 2, and six exons. Tissue-specific expression


**Fig. 7** qRT-PCR analysis of 20 *CnbHLHs* under various stress conditions, including cold (4 °C), heat (45 °C), NaCl (150 mM), and drought (mannitol 150 mM). All experiments were conducted independently three times. Error bars denote the standard deviation of triplicate samples. Asterisks denote statistically significant differences in transcript levels when compared with the blank control (25 °C, 0 h, 0 mM). (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). (a) Relative expression levels of selected *CnbHLHs* in leaf tissues. (b) Relative expression levels of selected *CnbHLHs* in root tissues.

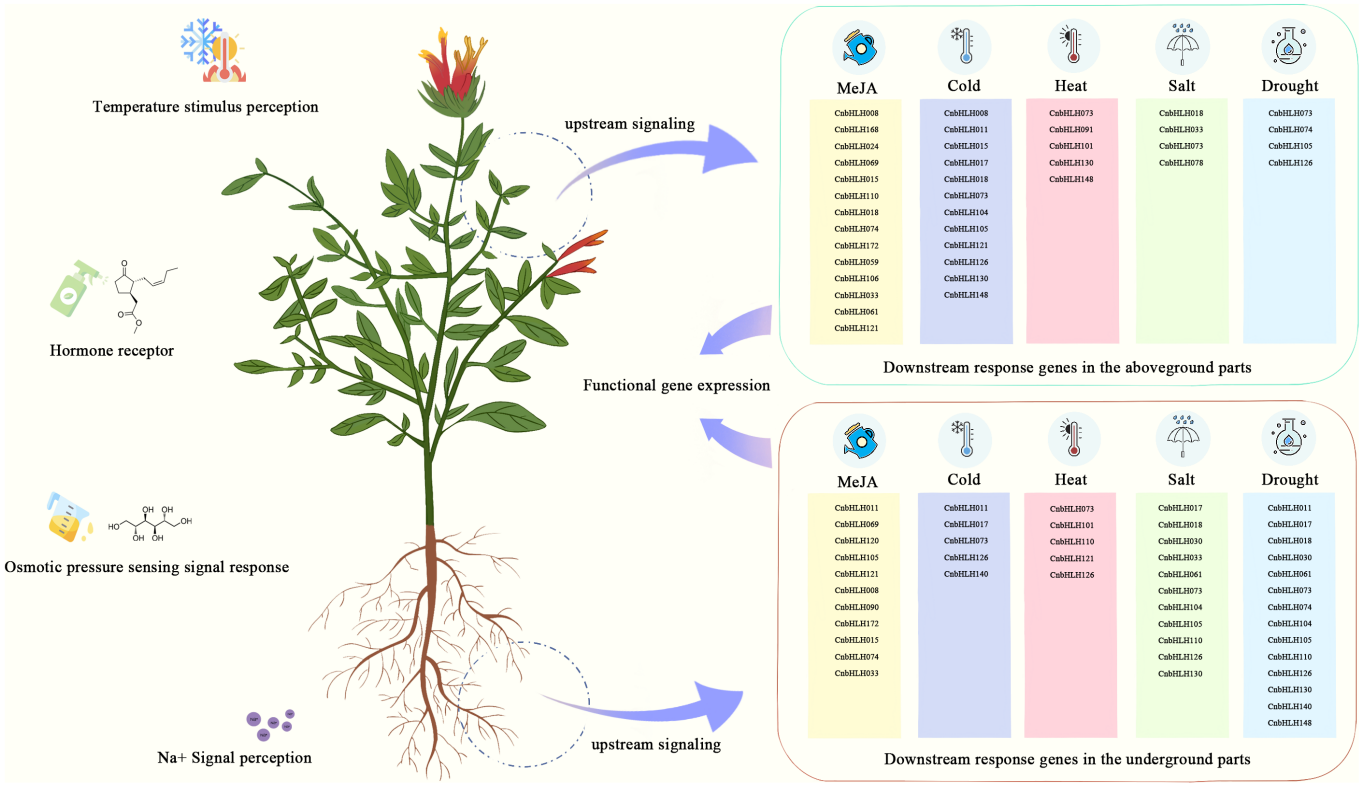


**Fig. 8** Diagram showing the response pathways to different abiotic stresses mediated by various bHLHs in *C. nutans*.

analysis showed that both genes were mainly expressed in mature leaves. Their expression in leaves both increased initially and then decreased after MeJA treatment (Fig. 6). However, many duplicated gene pairs showed different expression patterns. For example, *CnbHLH026* and *CnbHLH147*, which are in subfamily XIII, are segmentally duplicated genes. *CnbHLH026* was specifically highly expressed in mature leaves and showed a transient increase in expression in both leaves and roots following short-term MeJA induction. In contrast, *CnbHLH147* was specifically highly expressed in roots, but its expression in MeJA-treated leaves and roots was relatively low (Fig. 6). In addition, the segmentally duplicated gene pair *CnbHLH140* and *CnbHLH005*, both from subfamily IVc, were preferentially expressed in leaves before MeJA treatment (0 h), and at the final time point (72 h), respectively. These gene duplication events in *C. nutans* may have provided additional genetic material with diversified functions for adapting to changing environments. In addition, a *bHLH* gene cluster consisting of the tandem duplicated gene members including *CnbHLH044*, *CnbHLH045*, *CnbHLH046,* and *CnbHLH047* was identified, with all genes belonging to the same subfamily Ib(2). All of these genes exhibited higher expression in the roots compared with the leaves and responded to MeJA treatment. In *Catharanthus roseus*, three clustered *CrbHLHs* (named *BIS1*, *BIS2*, and *BIS3*) collectively regulate the biosynthesis of phenylethylamine in the terpene alkaloid biosynthesis pathway[18]. These clustered *CnbHLHs* might also have overlapping or synergistic regulatory functions in root development and/or MeJA response in *C. nutans*. Synteny analysis showed that 138 *CnbHLHs* had syntenic relationships with *A. paniculata*, followed by 129 with *S. cusia*, 106 with *A. thaliana*, and 82 with *V. vinifera* (Fig. 4b, Supplementary Table S8), suggesting large-scale gene expansion occurred before these species diverged. Thirty-seven *CnbHLHs* exhibited synteny across all five species, indicating that these genes within the bHLH family played a crucial role in evolutionary processes. The extent of synteny among Acanthaceae species (*C. nutans*, *A. paniculata*, and *S. cusia*) was much higher than with other species, suggesting that *CnbHLHs* may have expanded specifically in Acanthaceae to meet the unique evolutionary or adaptive needs of this plant family. The low Ka/Ks ratios of *CnbHLHs* compared with their homologs in other species suggest that these gene family members are under strong purifying selection (Fig. 4c, Supplementary Table S9), with harmful mutations being eliminated to maintain functional stability. This high level of functional conservation shows the importance of these genes in environmental adaptation and maintaining physiological stability across species. The CnbHLHs interaction network, consisting of 408 interactions and 59 nodes (Fig. 5a), showed that interactions occurred among different CnbHLHs. Notably, the CnbHLH168 node had the highest interaction reliability with several CnbHLHs, such as CnbHLH059, CnbHLH175, CnbHLH139, and CnbHLH121. The gene expression patterns of these interacting proteins were similar to CnbHLH168. *CnbHLH168* was mainly expressed in leaves and stems, while its interacting partners, *CnbHLH059* and *CnbHLH175*, were mainly expressed in young leaves (Fig. 6a, block b), and *CnbHLH139* and *CnbHLH121* were mainly expressed in mature leaves (Fig. 6a, block c). Additionally, *CnbHLH168* and these highly reliable interacting *CnbHLHs* were all preferentially expressed in the leaves after MeJA treatment (Fig. 6b, block I), suggesting these gene pairs may jointly regulate leaf development or the response to MeJA. However, many other interacting *CnbHLH* pairs with low interaction reliability showed different expression patterns. For example, *CnbHLH038* was highly expressed in the roots at both 0 and 72 h after MeJA treatment. The interacting genes showed different expression patterns: *CnbHLH058* was highly expressed in young leaves, while *CnbHLH153* had the highest expression in the stem. Whether these gene pairs interact to regulate certain biological processes together needs further validation.

Increasing evidence shows that bHLHs are crucial in plant growth, development, stress and hormone responses[58]. Transcriptomic analysis showed that most *CnbHLHs* had tissue-specific expression

patterns (Fig. 6a), suggesting their broad involvement in growth and developmental processes of *C. nutans*. MeJA plays a crucial role in regulating plant secondary metabolism, with many TFs, including *bHLH* and *MYB*, involved in this process[53]. The present results showed that after filtering out low-expressed genes, 39 *CnbHLHs* were identified in the leaves (Fig. 6b), and 44 in the roots (Fig. 6c) in response to MeJA treatment. Only 19 genes responded to MeJA in both tissues, indicating notable differences in the *CnbHLHs* responding to MeJA in the root and leaves. In both the roots and leaves, *CnbHLHs* expression in response to MeJA showed varied patterns during treatment, with different *CnbHLHs* involved at various time points after MeJA treatment. This suggests that *CnbHLHs* help fine-tuning the plant's response to MeJA, with different members active at various stages of signal transduction or cellular processes. In many medicinal plants, the accumulation of bioactive compounds such as flavonoids and terpenoids is a result of the plant's response to environmental stresses[67]. qRT-PCR was employed to analyze the expression patterns of 20 representative *CnbHLH* genes under different stress conditions, including cold, heat, salt, and drought (Fig. 7). The results showed that the expression patterns of the same gene in response to the same stress were markedly different between the roots and leaves, suggesting that different stress conditions may affect the metabolism of the aerial and underground parts differently. Specifically, most *CnbHLHs* in the leaves showed an initial increase in expression followed by a decrease, and then another increase during cold stress treatment. This phenomenon suggests that the leaves may initially adjust the expression of these genes to cope with low temperatures, and then gradually recover or enhance expression to adapt to prolonged stress. For example, *CnbHLH148* showed a more than 15-fold upregulation in the leaves after 72 h of cold stress treatment. Its homolog, *AtILR3*, has been reported to improve the activities of superoxide dismutase, peroxidase, and catalase, thereby enhancing stress tolerance[68]. In contrast, more than half of the *CnbHLHs* in the roots showed reduced expression under cold stress, suggesting that the roots may use a different mechanism to cope with cold stress. Moreover, most *CnbHLHs* in the roots showed increased expression after salt and drought stress. Some *CnbHLHs* were highly expressed in both salt and drought stress, such as *CnbHLH30*, which showed increased expression at various time points under both treatments in the roots. The homolog of *CnbHLH30*, *AtbHLH102* (*BIM2*), works with TFs like *HDA15* to regulate anthocyanin accumulation and growth responses under salt stress[69]. However, most *CnbHLHs* were more sensitive to one of these stresses. For instance, *CnbHLH011* and *CnbHLH017* showed consistently increased expression at all time points during drought stress in the roots. Their homolog, *AtMYC2* (*AtbHLH006*), expressed in the roots, has been identified as a transcriptional activator induced by ABA during drought stress[70]. *CnbHLH121* showed upregulated expression at various time points after salt stress in the roots. Its homologs, AtbHLH002 and AtbHLH042 from subfamily IIIf interacts with MYBs to regulate the biosynthesis of flavonoids and anthocyanins, enhancing stress tolerance[71]. Additionally, *CnbHLH105* showed significantly increased expression after prolonged salt stress in the roots. Its homolog, *AtbHLH059*, is induced by salt stress and contributes to physiological processes like $Na^+/K^+$ balance, redox homeostasis, and transpiration, thus improving salt tolerance[72]. Conversely, in the leaves, most *CnbHLHs* were reduced by salt and drought stress, which may indicate different adaptation strategies of the root and leaf tissues to these stresses. Only a few *CnbHLH* genes exhibited similar expression patterns in both roots and leaves under stress conditions. For example, *CnbHLH078* was induced to express higher levels after prolonged salt stress in both the roots and leaves. Its homolog, *AtbHLH112*, from subfamily X, binds to E-box and GCG-box elements in target gene promoters to regulate stress-responsive genes, enhancing stress resistance. Overexpression of *AtbHLH112* increases tolerance, while its functional loss renders plants more susceptible to stress[73]. Collectively, these findings highlight the tissue-specific and dynamic expression patterns of *CnbHLHs* under various stress conditions. *CnbHLHs* might function similarly to their homologs by regulating key processes such as antioxidant enzyme activity and secondary metabolism in response to environmental stressors.

## Conclusions

In this study, 182 *bHLHs* were identified and analyzed in *C. nutans*, revealing their potential roles in regulating important processes such as growth, development, and responses to hormones and stress. The expansion of *CnbHLHs* mainly happened through segmental duplication, with many genes showing functional redundancy, while others exhibited different expression patterns, indicating functional diversification. A group of tandemly duplicated genes, including *CnbHLH044–CnbHLH047*, which shared similar expression patterns, was identified, suggesting their possible cooperative functions in root development and MeJA responses. Synteny analysis offered insight into the evolutionary history of *CnbHLHs*, revealing higher synteny within Acanthaceae species, suggesting species–specific expansions. Tissue-specific expression patterns and varied responses to MeJA and environmental stress further emphasized the dynamic roles of CnbHLHs in regulating metabolic processes in different tissues. These findings provide new insights into the functional roles of CnbHLHs, offering a valuable resource for advancing molecular breeding and enhancing stress resilience in medicinal plants. However, this study has certain limitations. The current analysis was conducted based on a limited sample scope, and key regulatory mechanisms inferred from *in silico* predictions lack direct experimental validation. In addition, the integration of multi-omics data remains relatively shallow, which may limit the depth of functional interpretation. Future work should aim to perform molecular assays to verify predicted interactions, and incorporate epigenetic or metabolomic data. These efforts will help to achieve a more precise and holistic understanding of the roles of bHLHs in *C. nutans*.

## Author contributions

The authors confirm their contributions to the paper as follows: resources, data curation, formal analysis, writing—original draft, writing—review and editing: An C; software: Wu D; software, methodology: Li N; conceptualization, supervision: Yao Y; software: Lu L; supervision: Zhang Z; supervision: Cheng Y; supervision, project administration: Zheng P; Funding acquisition, conceptualization: Qin Y. All authors reviewed the results and approved the final version of the manuscript.

## Data availability

The raw transcriptome sequencing data used in this study has been deposited in CNGBdb under the accession number CNP0006103. The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

## Acknowledgments

We would like to express our sincere gratitude to the reviewers for their valuable comments and suggestions, which greatly improved the quality of our manuscript. This work was supported by the Fund of Fujian Key Laboratory of Island Monitoring and Ecological Development (Island Research Center, MNR) (2022ZD06), the Science and Technology Innovation Project of Fujian Agriculture and Forestry University (CXZX2019144G, CXZX2020086A), the National Natural Science Foundation of China (31970333 and 32100168), the Basic Discipline Research Enhancement Program of Fujian University of Traditional Chinese Medicine (XJC2023008), Construction Funds for Cross-disciplinary Synthetic Biology at the Institute of Future Technology, Fujian Agriculture and Forestry University (133-712023010), and the Scientific Research Start-up Fund of Fujian Agriculture and Forestry University (Yuan Qin).

## Conflict of interest

The authors declare that they have no conflict of interest.

**Supplementary information** accompanies this paper at (https://www.maxapress.com/article/doi/10.48130/mpb-0025-0030)

## Dates

Received 23 April 2025; Revised 8 August 2025; Accepted 25 August 2025; Published online 29 September 2025

## References

1. Chia TY, Gan CY, Murugaiyah V, Hashmi SF, Fatima T, et al. 2022. A narrative review on the phytochemistry, pharmacology and therapeutic potentials of *Clinacanthus nutans* (Burm. f. ) lindau leaves as an alternative source of future medicine. *Molecules* 27:139
2. Khoo LW, Mediani A, Zolkeflee NKZ, Leong SW, Ismail IS, et al. 2015. Phytochemical diversity of *Clinacanthus nutans* extracts and their bioactivity correlations elucidated by NMR based metabolomics. *Phytochemistry Letters* 14:123–33
3. Chiangchin S, Thongyim S, Pandith H, Kaewkod T, Tragoolpua Y, et al. 2023. *Clinacanthus nutans* genetic diversity and its association with anti-apoptotic, antioxidant, and anti-bacterial activities. *Scientific Reports* 13:19566
4. Fong SY, Piva T, Dekiwadia C, Urban S, Huynh T. 2016. Comparison of cytotoxicity between extracts of *Clinacanthus nutans* (Burm. f.) Lindau leaves from different locations and the induction of apoptosis by the crude methanol leaf extract in D24 human melanoma cells. *BMC Complementary and Alternative Medicine* 16:368
5. Bechtold U, Field B. 2018. Molecular mechanisms controlling plant growth during abiotic stress. *Journal of Experimental Botany* 69(11):2753–58
6. Khan V, Jha A, Princi, Seth T, Iqbal N, et al. 2024. Exploring the role of jasmonic acid in boosting the production of secondary metabolites in medicinal plants: Pathway for future research. *Industrial Crops and Products* 220:119227
7. Zheng H, Fu X, Shao J, Tang Y, Yu M, et al. 2023. Transcriptional regulatory network of high-value active ingredients in medicinal plants. *Trends in Plant Science* 28:429–46
8. Atchley WR, Zhao J. 2007. Molecular architecture of the DNA-binding region and its relationship to classification of basic helix-loop-helix proteins. *Molecular Biology and Evolution* 24:192–202
9. Ledent V, Vervoort M. 2001. The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis. *Genome Research* 11:754–70
10. Sun X, Wang Y, Sui N. 2018. Transcriptional regulation of bHLH during plant response to stress. *Biochemical and Biophysical Research Communications* 503:397–401
11. Klein ES, Simmons DM, Swanson LW, Rosenfeld MG. 1993. Tissue-specific RNA splicing generates an ankyrin-like domain that affects the dimerization and DNA-binding properties of a bHLH protein. *Genes & Development* 7:55–71
12. Wang K, Liu H, Mei Q, Yang J, Ma F, et al. 2023. Characteristics of bHLH transcription factors and their roles in the abiotic stress responses of horticultural crops. *Scientia Horticulturae* 310:111710
13. Chinnusamy V, Ohta M, Kanrar S, Lee BH, Hong X, et al. 2003. ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. *Genes & Development* 17:1043–54
14. Wang F, Zhu H, Chen D, Li Z, Peng R, et al. 2016. A grape bHLH transcription factor gene, *VvbHLH1*, increases the accumulation of flavonoids and enhances salt and drought tolerance in transgenic *Arabidopsis thaliana*. *Plant Cell, Tissue and Organ Culture (PCTOC)* 125:387–98
15. Ji Y, Xiao J, Shen Y, Ma D, Li Z, et al. 2014. Cloning and characterization of AabHLH1, a bHLH transcription factor that positively regulates artemisinin biosynthesis in *Artemisia annuac*. *Plant & Cell Physiology* 55:1592–604
16. Yuan M, Shu G, Zhou J, He P, Xiang L, et al. 2023. AabHLH113 integrates jasmonic acid and abscisic acid signaling to positively regulate artemisinin biosynthesis in *Artemisia annua*. *New Phytologist* 237:885–99
17. Patra B, Pattanaik S, Schluttenhofer C, Yuan L. 2018. A network of jasmonate-responsive bHLH factors modulate monoterpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *New Phytologist* 217:1566–81
18. Xu Y, Zhang H, Zhong Y, Jiang N, Zhong X, et al. 2022. Comparative genomics analysis of *bHLH* genes in cucurbits identifies a novel gene regulating cucurbitacin biosynthesis. *Horticulture Research* 9:uhac038
19. Zhang X, Luo H, Xu Z, Zhu Y, Ji A, et al. 2015. Genome-wide characterisation and analysis of bHLH transcription factors related to tanshinone biosynthesis in *Salvia miltiorrhiza*. *Scientific Reports* 5:11244
20. Xu J, Xu H, Zhao H, Liu H, Xu L, et al. 2022. Genome-wide investigation of *bHLH* genes and expression analysis under salt and hormonal treatments in *Andrographis paniculata*. *Industrial Crops and Products* 183:114928
21. Qin Y, Li J, Chen J, Yao S, Li L, et al. 2024. Genome-wide characterization of the bHLH gene family in *Gynostemma pentaphyllum* reveals its potential role in the regulation of gypenoside biosynthesis. *BMC Plant Biology* 24:205
22. Wei K, Chen H. 2018. Comparative functional genomics analysis of bHLH gene family in rice, maize and wheat. *BMC Plant Biology* 18:309
23. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, et al. 2021. Pfam: the protein families database in 2021. *Nucleic Acids Research* 49:D412–D419
24. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, et al. 2018. HMMER web server: 2018 update. *Nucleic Acids Research* 46:W200–W204
25. McGinnis S, Madden TL. 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research* 32:W20–W25
26. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, et al. 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Research* 43:D222–D226
27. Letunic I, Khedkar S, Bork P. 2021. SMART: recent updates, new developments and status in 2020. *Nucleic Acids Research* 49:D458–D460
28. Osorio D, Rondón-Villarreal P, Torres R. 2015. Peptides: a package for data mining of antimicrobial peptides. *The R Journal* 7(1):4–14
29. Pires N, Dolan L. 2010. Origin and diversification of basic-helix-loop-helix proteins in plants. *Molecular Biology and Evolution* 27:862–74
30. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32:1792–97
31. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, et al. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* 37:1530–34
32. Subramanian B, Gao S, Lercher MJ, Hu S, Chen WH. 2019. Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Research* 47:W270–W275
33. Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME suite. *Nucleic Acids Research* 43:W39–W49

34. Shen W, Sipos B, Zhao L. 2024. SeqKit2: a Swiss army knife for sequence and alignment processing. *iMeta* 3:e191
35. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, et al. 2002. PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Research* 30:325–27
36. Horton NJ, Kleinman K. 2015. *Using R and RStudio for data management, statistical analysis, and graphics*. 2nd Edition. New York: Chapman & Hall/CRC Press. doi: 10.1201/b18151
37. Wang Y, Tang H, DeBarry JD, Tan X, Li J, et al. 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research* 40:e49
38. Chen C, Wu Y, Li J, Wang X, Zeng Z, et al. 2023. TBtools-II: a 'one for all, all for one' bioinformatics platform for biological big-data mining. *Molecular Plant* 16:1733–42
39. Zhang Z, Li J, Zhao XQ, Wang J, Wong GK, et al. 2006. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. *Genomics, Proteomics & Bioinformatics* 4:259–63
40. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, et al. 2015. STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Research* 43:D447–D452
41. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. 2000. The Protein Data Bank. *Nucleic Acids Research* 28:235–42
42. Høie MH, Kiehl EN, Petersen B, Nielsen M, Winther O, et al. 2022. NetSurfP-3.0: accurate and fast prediction of protein structural features by protein language models and deep learning. *Nucleic Acids Research* 50:W510–W515
43. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, et al. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research* 46:W296–W303
44. DeLano WL. 2002. *PyMOL: an open-source molecular graphics tool*. CCP4 Newsletter on Protein Crystallography 40. pp. 82–92 https://legacy.ccp4.ac.uk/newsletters/newsletter40/11_pymol.pdf
45. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890
46. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* 37:907–15
47. Liao Y, Smyth GK, Shi W. 2013. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Research* 41:e108
48. An C, Lu L, Yao Y, Liu R, Cheng Y, et al. 2025. Selection and validation of reference genes in clinacanthus nutans under abiotic stresses, MeJA treatment, and in different tissues. *International Journal of Molecular Sciences* 26:2483
49. De Masi F, Grove CA, Vedenko A, Alibés A, Gisselbrecht SS, et al. 2011. Using a structural and logics systems approach to infer bHLH-DNA binding specificity determinants. *Nucleic Acids Research* 39:4553–63
50. Kato M, Hata N, Banerjee N, Futcher B, Zhang MQ. 2004. Identifying combinatorial regulation of transcription factors and binding motifs. *Genome Biology* 5:R56
51. Xu G, Guo C, Shan H, Kong H. 2012. Divergence of duplicate genes in exon-intron structure. *Proceedings of the National Academy of Sciences of the United States of America* 109:1187–92
52. Zhang W, Zhang J, Fan Y, Dong J, Gao P, et al. 2024. RNA sequencing analysis reveals *PgbHLH28* as the key regulator in response to methyl jasmonate-induced saponin accumulation in *Platycodon grandiflorus*. *Horticulture Research* 11:uhae058
53. Goossens J, Mertens J, Goossens A. 2017. Role and functioning of bHLH transcription factors in jasmonate signalling. *Journal of Experimental Botany* 68:1333–47
54. Zhou Z, Wu M, Sun B, Li J, Li J, et al. 2024. Identification of transcription factor genes responsive to MeJA and characterization of a LaMYC2 transcription factor positively regulates lycorine biosynthesis in Lycoris aurea. *Journal of Plant Physiology* 296:154218
55. Upadhyay R, Saini R, Shukla PK, Tiwari KN. 2025. Role of secondary metabolites in plant defense mechanisms: a molecular and biotechnological insights. *Phytochemistry Reviews* 24:953–83
56. Chang A, Lin L, Kangzhuo Y, Shengzhen C, Bingrui W, et al. 2023. 药用植物鳄嘴花 (忧遁草) 的研究概况及质量标志物预测分析 [ Comprehensive review on research of Ezuihua (Clinacanthus nutans) and predictive analysis on quality marker (Q-marker)]. *中华中医药学刊 [Chinese Archives of Traditional Chinese Medicine]* 41:25–35
57. Lin CM, Chen HH, Lung CW, Chen HJ. 2023. Antiviral and immunomodulatory activities of *Clinacanthus nutans* (Burm. f.) Lindau. *International Journal of Molecular Sciences* 24:10789
58. Lei P, Jiang Y, Zhao Y, Jiang M, Ji X, et al. 2024. Functions of basic helix-loop-helix (bHLH) proteins in the regulation of plant responses to cold, drought, salt, and iron deficiency: a comprehensive review. *Journal of Agricultural and Food Chemistry* 72:10692–709
59. Gao F, Dubos C. 2024. The arabidopsis bHLH transcription factor family. *Trends in Plant Science* 29:668–80
60. Wu Y, Wu S, Wang X, Mao T, Bao M, et al. 2022. Genome-wide identification and characterization of the *bHLH* gene family in an ornamental woody plant *Prunus mume*. *Horticultural Plant Journal* 8:531–44
61. Liang J, Fang Y, An C, Yao Y, Wang X, et al. 2023. Genome-wide identification and expression analysis of the *bHLH* gene family in passion fruit (*Passiflora edulis*) and its response to abiotic stress. *International Journal of Biological Macromolecules* 225:389–403
62. Flagel LE, Wendel JF. 2009. Gene duplication and evolutionary novelty in plants. *New Phytologist* 183:557–64
63. Jain M, Tyagi AK, Khurana JP. 2008. Genome-wide identification, classification, evolutionary expansion and expression analyses of homeobox genes in rice. *The FEBS Journal* 275:2845–61
64. Liu H, Lyu HM, Zhu K, Van de Peer Y, Max Cheng ZM. 2021. The emergence and evolution of intron-poor and intronless genes in intron-rich plant gene families. *The Plant Journal* 105:1072–82
65. Wu Y, Wen J, Xia Y, Zhang L, Du H. 2022. Evolution and functional diversification of R2R3-MYB transcription factors in plants. *Horticulture Research* 9:uhac058
66. Ye D, Liu J, Tian X, Wen X, Zhang Y, et al. 2024. Genome-wide identification of bHLH gene family and screening of candidate gene in response to salt stress in kiwifruit. *Environmental and Experimental Botany* 222:105774
67. He J, Yao L, Pecoraro L, Liu C, Wang J, et al. 2023. Cold stress regulates accumulation of flavonoids and terpenoids in plants by phytohormone, transcription process, functional enzyme, and epigenetics. *Critical Reviews in Biotechnology* 43:680–97
68. Jiang M, Niu Y, Wen G, Zhao C, Gao Q, et al. 2024. Evolutionary studies of the basic helix–loop–helix (bHLH) IVc gene family in plants and the role of *AtILR3* in *Arabidopsis* response to ABA stress. *Physiologia Plantarum* 176:e14128
69. Liao HS, Yang CC, Hsieh MH. 2022. Nitrogen deficiency-and sucrose-induced anthocyanin biosynthesis is modulated by HISTONE DEACETYLASE15 in *Arabidopsis*. *Journal of Experimental Botany* 73:3726–42
70. Le Hir R, Castelain M, Chakraborti D, Moritz T, Dinant S, et al. 2017. *AtbHLH68* transcription factor contributes to the regulation of ABA homeostasis and drought stress tolerance in *Arabidopsis thaliana*. *Physiologia Plantarum* 160:312–27
71. Kim D, Jeon SJ, Yanders S, Park S, Kim HS, et al. 2022. MYB3 plays an important role in lignin and anthocyanin biosynthesis under salt stress condition in *Arabidopsis*. *Plant Cell Reports* 41:1549–60
72. He Z, Wang Z, Nie X, Qu M, Zhao H, et al. 2022. UNFERTILIZED EMBRYO SAC 12 phosphorylation plays a crucial role in conferring salt tolerance. *Plant Physiology* 188:1385–401
73. Chen HC, Hsieh-Feng V, Liao PC, Cheng WH, Liu LY, et al. 2017. The function of *OsbHLH068* is partially redundant with its homolog, *AtbHLH112*, in the regulation of the salt stress response but has opposite functions to control flowering in *Arabidopsis*. *Plant Molecular Biology* 94:531–48