

# Genome-wide identification and characterization of melon bHLH transcription factors in regulation of fruit development

**Seling Bai**

School of Life Science, Inner Mongolia University

**Chao Tan**

Inner Mongolia University

**Yunyun Tian**

Inner Mongolia University

**Ming Ma**

Inner Mongolia University

**Agula Hasi** (✉ [hasind@sina.com](mailto:hasind@sina.com))

---

## Research article

**Keywords:** bHLH, Fruit ripening, bHLH93, Melon

**Posted Date:** February 6th, 2020

**DOI:** <https://doi.org/10.21203/rs.2.22819/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** The basic helix-loop-helix (bHLH) transcription factor family is one of the largest transcription factor families in plants, and plays crucial roles in plant development. Melon is one of an important horticulture plants, and is an attractive model plant for studying fruit ripening. However, the bHLH gene family of melon has not been identified yet, and functions in fruit growth and ripening are seldom researched.

**Results:** In this study, 118 bHLH genes were identified in the genome of melon. Phylogenetic analysis illustrated that these CmbHLHs could be classified into 16 subfamilies. Intron distribution pattern analysis of bHLH domain found 13 intron distribution patterns in CmbHLHs. CmbHLH genes were unevenly distributed on chromosomes 1 to 12 of the melon genome, and five CmbHLH s were tandem repeat on chromosomes 4 and 8. Expression characters of CmbHLH genes were studied using the transcriptome data. Tissue analysis of indicated CmbHLH32 high expressed in female flowers and early fruit growth stage. Transgenic plant lines of overexpression of CmbHLH32 were constructed, and overexpression of CmbHLH32 result in early fruit ripening compared to the wild type fruit.

**Conclusions:** The bHLH transcription factor family was identified and analyzed for the first time in the melon, overexpression of CmbHLH32 will affect the ripening time of melon fruit, these findings laid a theoretical foundation for further study on the role of bHLH family members in the growth and development of melon .

## Background

Transcription factors (TFs) play important roles in regulating plant growth, development, stress response and signaling transduction [1–4]. Basic helix-loop-helix (bHLH) TFs is one of the largest TF superfamily in plant [5]. Studies on the bHLH gene family in various species will increase our understanding of their evolution and functions. So far, comprehensive identification of the bHLH gene family has performed in range of plant species. Such as, *Arabidopsis*, *Brachypodium distachyon*, *Solanum lycopersicum*, *Arachis hypogae*, *Malus pumila* [6–10]. Plant bHLH is characterized by a basic helix-loop-helix domain and the domain is highly conserved in evolution [11]. The bHLH domain contains 50–60 amino acids and can be separated into two regions: the region at N-terminal end is a DNA binding domain, comprises about 13–17 amino acid [5]; the C-terminal end of the region is HLH domain, containing two amphipathic  $\alpha$ -helices connected by a loop region of variable length, which helps to form a dimerization domain, and allows the formation of homo- or hetero-dimeric complexes [12]. With regard to the bHLH proteins domain, 19 amino acids were conserved and functional for DNA binding or dimerization formation. A highly conserved HER-motif (His 5-Glu 9-Arg 13) was considered important for binding to specific DNA sequences, however some atypical bHLH lack of the N-terminal binding region [13].

Phylogenic analysis of 544 plant bHLH shows that plant bHLH proteins formed 26 distinct subfamilies, and these subfamilies are highly conserved throughout plant evolution; among them, 20 subfamilies were

present in early land plants 443 Ma [14]. However, the plant bHLH subfamilies lack of discernible phylogenetic relationships and possibly plant bHLH proteins are monophyletic [5].

Plant bHLH proteins involved in a wide arrange of biological processes, and participate in the regulation of plant growth and development, abiotic/biotic stress response, hormone signaling, iron homeostasis and secondary metabolism [15–21]. Nevertheless, researches on bHLH genes regulation in fruit growth and development were limited. Fruit growth is triggered by the fertilization of ovule, and changed drastically, these processes including cell division and expansion, secondary metabolites accumulation, carbohydrate biosynthesis, and involves many genes transcription and regulation [22, 23]. To the climacteric fruit like melon, the ripening process also accompanies with increasing in ethylene emission and burst of respiratory climacteric [24]. Recent research on tomato reveals that overexpression of SlbHLH22 resulting in early flowering time, accelerates fruit ripening, accumulation of carotenoid by activation of the SISFT or SILFY genes, and exogenous ACC, IAA, ABA, and ethephon would upregulate the expression of SlbHLH22 [25].

Melon (*Cucumis mleo* L.) is an important horticultural crop worldwide, in 2017 whole world production of melon was more than 49 million tones, and China produced over one third of the melon (FAO). Melon is also an attractive model plant of the Cucurbitaceae family for studying the fruit development and ripening, especially for respiratory climacteric [26]. Although The bHLH genes have been suggested to being involved in a wide arrange of metabolic, physiological, and developmental processes in plant, very few studies efforts on bHLH genes in the regulation on fruit ripening. In this study, we identified the melon bHLH gene family, and using high-throughput sequencing data to research the gene expression of melon fruit growth, ripening, climacteric, and post-climacteric stage. Transgenic lines that overexpression of CmbHLH32 was generated to study the effect on fruit ripening, and yeast two hybrid was used to research the transcription activation activity of CmbHLH32. Our findings shed light on the molecular properties and evolution patterns of the CmbHLH gene family and demonstrate biological function of CmbHLH32 gene in melon fruit growth and ripening.

## Results

### Identification of CmbHLH proteins and conserved domain alignment

A total of 169 CmbHLH candidate protein sequences were obtained by HMM analysis, 213 protein sequences were found using blast method, then repetitive sequences were removed. The remaining sequences were searched against the CmbHLH proteins of the PlantTFDB database. After that 214 sequences were reserved and submitted to CDD domain search; then 159 sequences were found with a bHLH conserved domain above the minimum domain hit, and the redundant sequences of the 159 proteins were removed, finally, 118 sequences were left as the *CmbHLHs* gene models for the analysis and renamed based on their chromosome localization (**Supplementary Table S1**).

These putative CmbHLH lengths varied from 84 to 707 aa, the molecular weight ranges from 9.48 kDa to 75.97 kDa, and the theoretical isoelectric points (PI) range from 4.52 to 10.27 (**Supplementary Table S1**).

The CmbHLH proteins density (0.442) was lower than that in *Cucumis sativus* (0.527), which is also belongs to the Cucurbitaceae plant. The reason may be the smaller genome size of *Cucumis sativus* (203.0 Mb) compared to *Cucumis melo* (364.0 Mb) [27]. And the CmbHLH density was a little higher than that in *Citrus sinensis* (0.420) which has a similar genome size (*Citrus sinensis*, 367.0 Mb).

Multiple sequences alignment of the bHLH domain of CmbHLH proteins shows that 24 amino acid residues in their bHLH domains were conserved with more than 50% consensus ratio. The bHLH domain is highly conserved and comprises of two functionally distinct regions. The basic region of the bHLH domain determines the DNA-binding activity of target genes, figure 1 shows the bHLH domain logo of CmbHLH (**Fig. 1**), the basic region of CmbHLH proteins contains 5 conserved amino acids, and the HLH has 19 conserved amino acids. Previously described a classification of bHLH proteins that classified the bHLH proteins into four groups A, B, C and D. The classification was based on DNA-binding specificity as well as conservation of amino acids at certain positions [13]. According to the criterion, 8 CmbHLH classified into A group, 69 located in B group, 16 and 33 CmbHLH belonged to C and D group respectively.

Domain analysis of CmbHLH illustrated that there were two kinds of domains were found in CmbHLH except the bHLH domain (**Fig. 2**). One domain is bHLH\_MYC\_N, was found in 11 CmbHLH. All of the CmbHLHs bHLH\_MYC\_N domain located in the N-terminal of the proteins and bHLH domain in the C-terminal of proteins. Apart from the bHLH and bHLH\_MYC\_N domains, ACT (Aspartokinase, Chorismate, and TyrA) domain was also identified in eight CmbHLH (CmbHLH32, CmbHLH37, CmbHLH56, CmbHLH60, CmbHLH68, CmbHLH97, CmbHLH100, CmbHLH114), and all of the ACT domain located in the C terminal of the bHLH domain.

### Phylogenetic, Motif Analysis and Gene Structure of CmbHLH

To evaluate the evolutionary relationships of the CmbHLH, we conducted a phylogenetic analysis based on full-length of protein sequences. Applying the ML method, we assigned the *CmbHLH* genes into 16 subfamilies and 4 orphan genes (**Fig. 3**). Subfamilies A, D and J were the largest groups, the smallest subfamily (L) had only 2 members. According to the phylogenetic tree, the *CmbHLH* binding activities was phylogenetically clustered which were consistent with the previously report. For example, nine subfamilies members (A, B, C, H, I, J, K, M and O) were belonged to B group protein, four subfamilies members (E, F, G and M) classified into group D protein.

The evolutionary relationships of these CmbHLH proteins were also determined by conserved motifs. A total of 10 conserved motifs were characterized from CmbHLH proteins (**Fig. 4A, B**). Among these motifs, motif 1 and 2 were annotated to bHLH domain (IPR011598, IPR036638), motif 7 and 10 were annotated to transcription factor bHLH-MYC-N-terminal (IPR025610). The subfamily L contains the highest number of motifs (six motifs). CmbHLH14 and CmbHLH94 possess two and three motif-2 respectively. The motifs distribution and construction pattern exhibited similar model within subfamilies.

The exon-intron organizations of *CmbHLH* were examined to gain more insight into the evolution of the bHLH family in melon. The exon number of *CmbHLH* varied from 1 to 11 (**Fig. 4C**). Whereas, the exon-

intron organizations were phylogenetically related. For example, The *CmbHLH* with one exon were clustered in two subfamilies (D and K), all the members of subfamily I have two exons. Intron distribution analysis of bHLH domain within all *CmbHLH* proteins exhibit 13 intron distribution patterns, and this pattern strongly related to the subfamilies of *CmbHLH* (**Fig. 4D**). As shown in figure 4, 85% of *CmbHLH* have intron insertion in their bHLH domain sequence region. Although the intron positions and lengths were varied, only five intron insertion positions of *CmbHLH* were unconserved. Overall, the conserved motif arrangement and composition and the gene structure of *CmbHLH* genes, together with the phylogenetic analysis results, could strongly support the reliability of the classification.

### Chromosomal distribution and collinearity analysis of *CmbHLH*

*CmbHLH* genes were distributed unevenly among twelve chromosomes of melon (**Fig. 5**). However, the distribution of *CmbHLH* genes did not show either a chromosome length correlation or a phylogenetic correlation. Gene tandem duplication may involve in gene family enlargement and maintains of gene copy numbers. Thus, we analyzed the tandem duplication events of *CmbHLH*. Five genes were confirmed to be tandem duplicated genes. Three of them (*CmbHLH81*, *CmbHLH82* and *CmbHLH83*) located on chromosome 8, and two of tandem duplicated genes (*CmbHLH45* and *CmbHLH44*) located on chromosome 4.

To further infer the origin and phylogenetic relationships of bHLH genes, comparative collinearity analysis between *Cucumis melo* and other cucurbit species were conducted. Figure 6 displayed the collinearity relationship of *CmbHLH* genes with those in bottle gourd, cucumber, watermelon and *Cucurbita maxima* (Rimu) (**Fig. 6**). A total of 115 *CmbHLH* genes have orthologous in the four species, among them 95 *CmbHLH* genes were common in the four species. Interestingly, 43 *CmbHLH* genes have at least two orthologous in Rimu, and these genes spread out on the 12 chromosomes of melon. The reason maybe that Rimu genome underwent a whole-genome duplication (WGD) event, which was not observed in other four cucurbits (cucumber, melon, watermelon, and bitter melon) [28]. Gene duplication events of *CmbHLH* in melon were also studied. Results show 38 *CmbHLH* genes duplicated among *CmbHLH* genes (**Supplementary Figure S1**). Except chromosome 9, all the other chromosomes have duplication genes of *CmbHLH*. Most of the duplication genes were on chromosome 2, and 1 *CmbHLH* locates on chromosome 6, illustrating an uneven distribution of the duplication genes.

### Expression pattern of *CmbHLH* in melon fruit development

Analysis of the expression data of PRJNA543288 exhibits 161 transcripts of 98 *CmbHLH* genes effectively expressed (expressed at least in two replicate libraries) in melon fruit Growth stage (G), Ripening stage (R), Climacteric stage (C), and Post-climacteric stage (P) samples. However, most of them were low expressed, 45 *CmbHLH* genes have an average expression higher than 10 FPKM, only 5 genes (*CmbHLH23*, *CmbHLH32*, *CmbHLH41*, *CmbHLH67* and *CmbHLH79*) expression higher than 100 FPKM (**Fig. 7A**). Differential expression analysis shows 32 *CmbHLH* genes differentially expressed in G vs R, R vs C and C vs P stage samples. A total of 21 *CmbHLH* genes differentially expressed in G vs R samples (7 genes upregulated and 14 genes downregulated), *CmbHLH32* was the highest expressed among these

differential expression genes; 26 *CmbHLH* genes differentially expressed in R vs C samples, only 2 genes (*CmbHLH9* and *CmbHLH114*) were up regulated, the others were down regulated. There were six *CmbHLH* genes differentially expressed in C vs P stage samples. Taken together, the expression of *CmbHLH* genes exhibits a down regulation trend from melon fruit G to P stage samples, suggesting most of the *CmbHLH* genes may function in early fruit developmental stage. Whereas, two genes (*CmbHLH9* and *CmbHLH114*) up regulated in R vs C stage samples, indicating they may be involved in the regulation of fruit Climacteric.

### Overexpression of *CmbHLH32* leading to early ripen in melon fruit

To further investigate the function of *CmbHLH* genes in fruit ripening, we generated the transgenic plant lines of overexpression *CmbHLH32* (*CmbHLH32*-OE). The reasons for studying *CmbHLH32* gene were: first, *CmbHLH32* was one of the highest expressed *CmbHLH* genes, second; *CmbHLH32* was highest expressed among differential expression genes in G vs R stage samples; third, result in tissue expression analysis of *CmbHLH32* illustrates *CmbHLH32* was high expressed in female flower and early developmental stage of fruit (**Fig. 7B**); fourth, *CmbHLH32* was identified homolog to *AtbHLH93*, which was proved to control *Arabidopsis* flowering by repressing *MAF5* [29]. However, blast analysis was failed to find a homolog gene of *MAF5* in melon, and *CmbHLH32* also contains an ACT domain which is not find in *AtbHLH93*, suggesting *CmbHLH32* may have a different function in melon flowering.

Transgenic T1 seeds that overexpression of *CmbHLH32* was generated by the ovary injection method. Fruit ripening related phenotype observation of *CmbHLH32*-OE T1 plant that were PCR detection positive indicates that overexpression of *CmbHLH32* results in early fruit ripening compared to the wild type (WT) melon fruit (**Fig. 7C**). The fruit ripening of *CmbHLH32*-OE line (about  $38.7 \pm 1.1$  DAP) is in average 4 days earlier than that of WT melon fruit (about  $42.6 \pm 0.8$  DAP). Quantitative RT-PCR analysis exhibits that the expression level of *CmbHLH32* gene has increased an average about 5.5 times than WT fruit (**Fig. 7D**). Transcriptional activity of *CmbHLH32* was also studied by yeast two hybrid, however, neither monomer nor homodimerizes of *CmbHLH32* shows transcription activation activity in yeast (**Fig. 8A**).

Analysis of fruit weight, length, width and fruit soluble solids content shows no different between WT melon fruit and *CmbHLH32*-OE transgenic line fruits. However, *CmbHLH32*-OE transgenic plant fruits exhibit less firmness than WT melon fruits (**Table 1**). Expression correlation networks analysis using the transcriptome data suggesting 94 genes correlated with *CmbHLH32* (**Supplementary Table S2**). Further Gene Ontology (GO) analysis of correlation genes revealed GO term plant-type cell wall biogenesis (GO:0009832) was enriched (**Fig. 8B**), suggesting *CmbHLH32* function may affect fruit softening through plant cell wall synthesis.

Table 1 Phenotypes of wild type (WT), *CmbHLH32*-overexpressing (*CmbHLH32*-OE) transgenic melon plants.

Parameter	WT	<i>CmbHLH32</i> -OE L1	<i>CmbHLH32</i> -OE L2
fruit weight (g)	961.8±155.3	954.1±169.4	890.5±113.1
Horizontal diameters of fruit (cm)	12.7±0.9	12.7±0.9	12.6±0.7
Vertical diameters of fruit (cm)	13.6±1.0	13.6±0.7	13.1±0.7
soluble solids content (%)	13.7±1.0	12.7±2.9	11.6±1.9
firmness of fruit (Kg)	5.5±0.6a	4.5±0.4b	4.0±0.2b

Values are means of 3-10 plants, ±SE. The statistical significance of mean differences was analyzed using Student' *t*-test, *P* < 0.05

## Discussion

### Characteristics of *CmbHLH* genes in melon

Although the bHLH family has been widely studied in a diverse range of plants, the present study are the first to report the identification and characterization of bHLH transcription factors based on the entire genome sequence of *Cucumis melo*. In this study, 118 *CmbHLH* genes were characterized and classified into 16 subfamilies and 4 orphan genes. Previous studies reveal that most of the bHLH proteins contained 19 conserved amino acid sites, for example in *Arabidopsis* and Chinese cabbage [6, 27]. However, in melon *CmbHLH* proteins contained 24 conserved amino acid sites, the conservation pattern was similar to the bHLH proteins in tomato. They had more conserved amino acid in the second helix motif in the *CmbHLH* and *SlbHLH* proteins.

Proteins containing the HLH motif often form homo- or heterodimers with other bHLH proteins. Leu-23 and Leu-52 (Leu-60 in melon) residue of helix 1 and 2, respectively, are structurally necessary for dimer formation of plant bHLH and have been identified as the most highly conserved residues across plant bHLH [11, 30]. Interestingly, in melon five *CmbHLH* have mutant Leu-23 or Leu-60 or both. Leu-23 was mutated to Phe-23 and Ile-23 in *CmbHLH31* and *CmbHLH58* respectively. Leu-60 mutant to Met-60 in *CmbHLH57* and *CmbHLH91*. In *CmbHLH62*, Leu-23 and Leu-60 was mutated to Val-23 and Glu-60 respectively. In *Arabidopsis* mutation on the two Leu site significantly affect the dimerization of the bHLH [31]. This result indicating a different functional mechanism may exist at least on these five *CmbHLH*.

Domain analysis of the *CmbHLH* illustrated eight proteins possessed an ACT domain. ACT domain is small regulatory domain involved in amino acid or purine metabolism and can bind to variety of ligands [32]. Mas-Droux et.al found a Lys and S-adenosylmethionine–sensitive Asp kinase isoform form a dimeric structure through ACT domain in *Arabidopsis* [33]. Kong et.al reveals that bHLH DNA-binding activity is suppressed if the C-terminal ACT domain is licensed to homodimerize, and this protein-protein interaction domain is important for the regulation of anthocyanin pigment biosynthesis in maize [34]. Eleven MYC type *CmbHLH* were found. In plant, MYC genes participate in growth and development, and

are also key regulators of the jasmonic acid (JA) signaling pathway [35, 36]. This result indicates CmbHLH may involve in a more complex regulation network.

### Potential roles of CmbHLH genes in melon fruit development

In the past few decades, characterization and function of bHLH family on several species have been widely and extensively investigated. As one of the largest transcription factor family, bHLH functions involved in the regulation of plant growth and development, abiotic/biotic stress respond, hormone signaling, iron homeostasis, secondary metabolism. For example, in *Arabidopsis*, brassinosteroid (BR) and gibberellin can promote cell elongation by inhibiting an atypical bHLH transcription factor INCREASED LEAF INCLINATION1 BINDING bHLH1 (IBH1), and ectopic accumulation of IBH1 causes a dwarf phenotype in *Arabidopsis* [37]. In soybean, *GmORG3-like* gene enhances cadmium tolerance by increasing iron and reducing cadmium uptake and are transported from roots to shoots [38]. Whereas, studies on bHLH regulating fruit growth and development were less. Waseem et.al found that overexpression of *SlbHLH22* leading to earlier fruit ripening, and produce more ethylene-producing phenotypes in tomato [25]. Zhao et.al used white-flesh mutant strawberry identified seven *FabHLH* genes that are responsive to the fruit anthocyanin biosynthesis and hormone signaling transduction [39].

Using the transcriptome data, we studied the expression of *CmbHLH* in fruit different developmental stages. Three *CmbHLH* (*CmbHLH14*, *CmbHLH32* and *CmbHLH41*) genes were high expressed (>100 FPKM) in G stage and down regulated in R stage. *CmbHLH14* is 74% identity to UNE12 of *Arabidopsis*, which might be involved in the regulation of the specific processes required for fertilization, and as a temperature-responsive SA immunity regulator [40, 41]. *CmbHLH23* was upregulated in R stage, and was homolog genes of AtbHLH68 were proposed to regulate lateral root elongation, and in the response to drought stress, likely through an ABA-dependent pathway in *Arabidopsis* [42]. *CmbHLH59* showed significantly down regulation in R vs C stage samples. *CmbHLH59* was an MYC type protein and differentially expressed from development to ripening stage in melon fruit. *CmbHLH59* sequence is similar to ATbHLH13 in *Arabidopsis*, functions of ATbHLH13 are repressing *Arabidopsis* defense responses and regulating anthocyanin biosynthesis through JA signaling pathway [43, 44]. *CmbHLH114* was upregulated in R vs C stage samples, and *CmbHLH114* is annotated to ICE1-like gene. in *Arabidopsis*, *ICE1* is a multifunctional gene that responsive to cold stress, ABA signaling regulation, determination of seed dormancy and so on [45, 46]. In banana, MaICE1 targeting to MaNAC1 during fruit cold storage and enhance fruit cold tolerance [47]. The expression of *CmbHLH67* was down regulated in R vs C stage and upregulated in C vs P stage. *CmbHLH67* is 59% identity to MYC2 of *Arabidopsis*, and response to a variety of JA-dependent functions including secondary metabolism, insect resistance and stress tolerance [48, 49]. In apple fruit ripening, MdMYC2 is required for JA-induced ethylene biosynthesis, besides, MdMYC2 was also found interacting with MdERF2, which is a downstream molecular of ethylene signaling [50].

In this study, we observed that overexpression of *CmbHLH32* leading to early fruit ripening. This result suggesting *CmbHLH32* may participate in growth and ripening of melon fruit, however, detailed study is



still needed. In apple leaf senescence and the expression of senescence related genes is promoted by MdbHLH93 (homolog to *AtbHLH93*), and leaf senescence was delayed when an ABA-responsive protein, MdbT2, interacted directly with MdbHLH93 [51]. In *Arabidopsis*, double mutants of *bHLH93* failed to flower under short day (SD) condition and *bHLH93* plays a major role in regulation *Arabidopsis* SD flowering [29]. Tissue expression analysis illustrates high expression level of *CmbHLH32* in female flower, suggesting *CmbHLH32* may play a regulation role in flowering. However, melon is day-neutral plant, suggesting *CmbHLH32* may have different function of *bHLH93* in *Arabidopsis*. Yeast two hybrid analysis indicates transcriptional activation activity of CmbHLH32 may depend on other proteins. Melon CmbHLH32 contains an ACT domain, ACT domain was first identified as a ligand-binding domain, and was found to suppress bHLH DNA-binding activity [34, 52]. This could explain homodimerizes of CmbHLH32 did not show transcriptional activation activity in vitro.

## Conclusions

This study was focused on the identification of melon bHLH gene family, and function of bHLH genes in fruit growth and ripening. We identified the CmbHLH gene family in melon. We demonstrated the CmbHLH classification, the bHLH domain characters and intron patterns, and phylogenetic relations of CmbHLH and other four cucurbit species. Expression characters of CmbHLH genes using the transcriptome data suggesting most of these CmbHLH genes were low/ no expressed in fruit development. And CmbHLH genes tend to expressed in early stage of fruit development, CmbHLH32 was the most prominent among them. Expression of CmbHLH32 was high in female flowers and early fruit growth stage. Transgenic plant lines of overexpression of CmbHLH32 result in early fruit ripening compared to the wild type fruits. These findings clarify the members and characters of CmbHLH gene family, and provide new insights into the role of melon bHLH genes in regulating fruit development.

## Methods

### Sequence retrieval and identification of bHLH proteins in melon

The Hidden Markov Model method and blast method were used to identify the melon bHLH protein sequences. Melon protein sequences (CM3.5.1\_protein) were downloaded from the Cucurbit Genomics Database (CuGenDB, <http://cucurbitgenomics.org/>) [53]. The protein sequences of *A. thaliana* bHLH (AtbHLH, Araport11\_genes.201606.pep.fasta) were retrieved from The Arabidopsis Information Resource (TAIR), ([https://www.arabidopsis.org/download/index-auto.jsp?dir=/download\\_files/Proteins](https://www.arabidopsis.org/download/index-auto.jsp?dir=/download_files/Proteins)), on [www.arabidopsis.org](http://www.arabidopsis.org). The Hidden Markov Model (HMM) file of the HLH domain (PF00010) was downloaded from the Pfam database (version 32.0; <http://pfam.xfam.org/>) [54]. The HMM software (version3.2.1; <http://www.hmmer.org/>) was used to search against the melon protein sequence data using default parameters [55]. The *Arabidopsis* bHLH protein sequences were used to blast against the melon protein sequences using Blast software by default parameters [56]. Then sequences were compared to the CmbHLH proteins downloaded from PlantTFDB ([http://planttfdb\\_v4.cbi.pku.edu.cn/](http://planttfdb_v4.cbi.pku.edu.cn/)), and repeated sequences were removed. After that these protein sequences were submitted to the online

Batch CD-search tool (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) to verify the existence and integrity of the conserved bHLH domain [57, 58]. The sequences above the minimum threshold bit score were kept, and the redundant sequences were removed. These representative sequences of putative *CmbHLHs* were named based on their chromosome location. Transcriptome data using previously published data from our laboratory (PRJNA543288).

### **Identification of conserved motifs and tandem repeat genes**

The molecular weight (Mw) and theoretical isoelectric point (pI)-values for these CmbHLH protein sequences were determined by the Compute pI/Mw tool on the ExPASy server ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/)) [59]. Conserved motifs of CmbHLHs were identified by MEME (<http://meme-suite.org/>) server with maximum number of motifs set at 10 and optimum width of motifs from 6 to 70 amino acids [60]. MCscanX software was used to identify the tandem repeat genes of the *CmbHLHs* using the default parameters. And the results were displayed by the TBtools software (version 0.66682) [61, 62].

### **Gene structure and collinearity analysis**

Gene structure (intron-exon) information of these putative *CmbHLHs* was obtained from the GFF (CuGenDB, CM3.5.1\_pseudomol\_gene\_model.gff) file and displayed by the TBtools software. Intron distribution pattern within bHLH domain of CmbHLHs was analyzed manually. The cds sequences and gff file of cucumber (Chinese long v2), watermelon (Charleston Gray v2), *cucurbita maxima* (Rimu v1.1) and Bottle gourd (*Lagenaria siceraria* cv. USVL1VR-Ls v1) were downed from CuGenDB [28, 63-65]. Collinearity analysis of the four species and melon was performed by the MCscanX software.

### **Phylogenetic analyses of CmbHLHs**

Multiple alignment of the putative melon CmbHLHs proteins were used MEGA X software (version 10.0.5) by MUSCLE program with default options, the maximum likelihood (ML) method were used to construct phylogenetic trees by PhyML 3.0 software, with VT (variable time) amino acid substitution model and 1000 replications of bootstrap values [66, 67].

### **Expression analysis and GO annotation**

Expression value (FPKM) of CmbHLHs transcripts were extracted from the transcriptome expression data sets, and genes expression correlation network were constructed by Expression correlation plugin (version 1.1.0) in Cytoscape software (version 3.7.1) [68], genes with correlation coefficient < -0.99 or > 0.99 and E-box (CANNTG) sequence in their promoter sequence (1000 bp) were kept for later analysis. GO annotation were conducted by online website Metascape (<http://metascape.org/gp/index.htm>) [69].

### **Plant material and growth condition**

Melon (*Cucumis melo* cv. Hetao) plants were grown in green house at Dengkou (N40°19'46.07", E107°0'11.46"), Inner Mongolia Autonomous Region, China. Hetao melon is a conventional melon variety planted in the western part of Inner Mongolia. It has been cultivated for more than 70 years in the local area. The fruit of Hetao melon is respiratory climacteric fruit, small fruit and medium early maturing variety, which can be cultivated in field or in greenhouse. The melon plants that we used to construct transgenic plant were using the above-mentioned cultivars, after our laboratory 11 generations of self-pollination breeding selection and obtained typical and stable inbred lines of Hetao melon. The flowers were self-pollinated, the pollination time was recorded, and the date of pollination was recorded as 0 days after pollination (DAP). Only one fruit was kept for each plant.

The wild-type melon seeds, which were from the above mentioned stable inbred lines of Hetao melon, sterilized by mercuric chloride were planted in 1/2MS medium at 25 °C, light (1000lx) for 16 hours, dark treatment for 8 hours, and cultured for 30 days to obtain the root, stem and leaf tissues of melon. The male and female flowers of wild type melon without pollination which were planted in greenhouse were collected and preserved immediately in liquid nitrogen.

### Gene cloning and transgenic plant generation

An open reading frame of 1023bp of *CmbHLH32* (*MELO3C011110*) gene was amplified by RT-PCR from melon total RNA using upstream primer: 5'-atggagagctcacgaacgtctt-3' with a *Kpn*I restriction site and downstream primer: 5'-ctacagcagcttctctacac-3' with a *Xba*I restriction site. The PCR products were cloned into the vector of pMD-19T (Takara Bio, Shiga, Japan) by TA cloning. Subsequently, the fragment was inserted into the overexpression vector pPZP221 applying *Kpn*I and *Bam*HI sites. The plasmid pPZP221-*CmbHLH32* was diluted with 2×SSC (pH 7.0) solution to 100ng/μl. After 7 hours of artificial pollination, 100ng/μl of plasmid solution was injected into the pistil ovary by ovary injection method to obtain transgenic plant.

T1 transgenic plants were identified by PCR test using two primer sets, one is CaMV35S upstream primer: 5'-CAGAAAGAATGCTAACCC-3' and downstream primer: 5'-TTCTTCTTGTCATTGAGTCGTA-3'; another is upstream primer: 5'-TTTCGGTCGTGAGTTCGGAG-3' and downstream primer: 5'-CACTTCTTCCCGTATGCCCA-3'. Only if the plant that was detected by both of these primers will be preserved for later observations.

### Quantitative real-time PCR analysis

A quantitative real-time PCR (qRT-PCR) assay was performed to validate the expression of *CmbHLHs*. Total mRNAs were reverse transcribed using PrimeScript™ RT reagent Kit with gDNA Eraser kit (Takara Bio, Shiga, Japan) following the manufacturer. The qRT-PCRs were performed using SYBR® Premix Ex Taq™ II (Takara Bio, Shiga, Japan) by a 96-well Chromo4 Real-Time PCR system. The qRT-PCR conditions were as follows: a predenaturation of 30 s at 95°C, followed by 35 cycles of 5 sec at 95°C and 30 sec at 60°C. The  $2^{-\Delta\Delta CT}$  method was used to analyze the relative mRNA expression level.

## Transcriptional activation activity analysis in yeast

Full length of *CmbHLH32* was cloned and inserted into pGBKT7 and pGADT7 vectors (TaKaRa) to obtain pGBKT7-CmbHLH32 and pGADT7- CmbHLH32 vectors. The pGBKT7-CmbHLH32 was transformed into the yeast strain AH109 (TaKaRa) with pGADT7-T vector for transcriptional activation activity test. Homodimerizes transcriptional activation activity of *CmbHLH32* used cotransformation of pGBKT7-CmbHLH32 and pGADT7- CmbHLH32 vectors. Cotransformation of pGBKT7-53 + pGADT7-T and pGBKT7-Lam + pGADT7-T were used as positive control and negative control respectively. The transformed yeast cells were streaked on SD/-Trp/-His/Ade/3-AT solid medium in different dilutions ( $10^{-1}$ ,  $10^{-2}$  and  $10^{-3}$ ). After 3-day incubation at 30 °C, the transcriptional activation activity of *CmbHLH32* were evaluated according to growth status of transformed yeast cells.

## Abbreviations

bHLH: Basic helix–loop–helix; TF: Transcription factor; DAP: days after pollination; SFT: SINGLE FLOWER TRUESS; LFY: LEAFY; ACC: 1-Aminocyclopropane-1-carboxylate; IAA: auxin indole-3-acetic acid. MAF5: FLOWERING LOCUS M (FLM)/MADS AFFECTING 5.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Availability of data and materials

The datasets analyzed during the current study are available in the SRA (<https://www.ncbi.nlm.nih.gov/sra>), SRA accession: PRJNA543288.

### Competing interests

The research was conducted in the absence of any potential conflict of interest.

### Funding

This work was supported by the National Natural Science Foundation of China (No.31560561) and the Inner Mongolia Natural Science Foundation of China (No.2017ZD05). The funding bodies were not involved in the design of the study, data collection, interpretation of data, or in writing the manuscript.

### Authors' contributions

AH designed experiments; the CmbHLH gene family identification and characterization was analyzed by SLGB, CT and MM; qRT-PCR, transgenic plant lines and yeast experiments were performed by CT, SLGB and YYT; SLGB and CT wrote the manuscript. All authors have read and approved the manuscript.

## Acknowledgments

Not applicable.

## References

1. Wu S, Gallagher KL. Transcription factors on the move. *Current opinion in plant biology* 2012, 15(6):645-651.
2. Liang W, Ma X, Wan P, Liu L. Plant salt-tolerance mechanism: A review. *Biochemical and biophysical research communications* 2018, 495(1):286-291.
3. Ng D, Abeyasinghe J, Kamali M. Regulating the regulators: The control of transcription factors in plant defense signaling. *International journal of molecular sciences* 2018, 19(12):3737.
4. Kavas M, Baloğlu MC, Atabay ES, Ziplar UT, Daşgan HY, Ünver T. Genome-wide characterization and expression analysis of common bean bHLH transcription factors in response to excess salt concentration. *Molecular genetics and genomics* 2016, 291(1):129-143.
5. Feller A, Machemer K, Braun EL, Grotewold E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *The Plant Journal* 2011, 66(1):94-116.
6. Toledo-Ortiz G, Huq E, Quail PH. The Arabidopsis basic/helix-loop-helix transcription factor family. *The Plant Cell* 2003, 15(8):1749-1770.
7. Niu X, Guan Y, Chen S, Li H. Genome-wide analysis of basic helix-loop-helix (bHLH) transcription factors in *Brachypodium distachyon*. *BMC genomics* 2017, 18(1):619.
8. 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):9.
9. 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.
10. Mao K, Dong Q, Li C, Liu C, Ma F. Genome wide identification and characterization of apple bHLH transcription factors and expression analysis in response to drought and salt stress. *Frontiers in plant science* 2017, 8:480.
11. Carretero-Paulet L, Galstyan A, Roig-Villanova I, Martínez-García JF, Bilbao-Castro JR, Robertson DL. Genome-wide classification and evolutionary analysis of the bHLH family of transcription factors in *Arabidopsis*, poplar, rice, moss, and algae. *Plant physiology* 2010, 153(3):1398-1412.
12. Jones S. An overview of the basic helix-loop-helix proteins. *Genome biology* 2004, 5(6):226.

13. Atchley WR, Fitch WM. A natural classification of the basic helix–loop–helix class of transcription factors. *Proceedings of the National Academy of Sciences* 1997, 94(10):5172-5176.
14. Pires N, Dolan L. Origin and diversification of basic-helix-loop-helix proteins in plants. *Molecular biology and evolution* 2009, 27(4):862-874.
15. Bao M, Bian H, Zha Y, Li F, Sun Y, Bai B, Chen Z, Wang J, Zhu M, Han N. miR396a-mediated basic helix–loop–helix transcription factor bHLH74 repression acts as a regulator for root growth in *Arabidopsis* seedlings. *Plant and Cell Physiology* 2014, 55(7):1343-1353.
16. Babitha K, Vemanna RS, Nataraja KN, Udayakumar M. Overexpression of EcbHLH57 transcription factor from *Eleusine coracana* L. in tobacco confers tolerance to salt, oxidative and drought stress. *PLoS One* 2015, 10(9):e0137098.
17. Yin J, Chang X, Kasuga T, Bui M, Reid MS, Jiang C-Z. A basic helix-loop-helix transcription factor, PhFBH4, regulates flower senescence by modulating ethylene biosynthesis pathway in petunia. *Horticulture research* 2015, 2:15059.
18. Tanabe N, Noshi M, Mori D, Nozawa K, Tamoi M, Shigeoka S. The basic helix-loop-helix transcription factor, bHLH11 functions in the iron-uptake system in *Arabidopsis thaliana*. *Journal of plant research* 2019, 132(1):93-105.
19. Gao F, Robe K, Gaymard F, Izquierdo E, Dubos C. The transcriptional control of iron homeostasis in plants: a tale of bHLH transcription factors? *Frontiers in plant science* 2019, 10:6.
20. Lim S-H, Kim D-H, Kim JK, Lee J-Y, Ha S-H. A radish basic helix-loop-helix transcription factor, RsTT8 acts a positive regulator for anthocyanin biosynthesis. *Frontiers in plant science* 2017, 8:1917.
21. Quattrocchio F, Verweij W, Kroon A, Spelt C, Mol J, Koes R. PH4 of *Petunia* is an R2R3 MYB protein that activates vacuolar acidification through interactions with basic-helix-loop-helix transcription factors of the anthocyanin pathway. *The Plant Cell* 2006, 18(5):1274-1291.
22. Barry CS. Factors influencing the ripening and quality of fleshy fruits. *Annual Plant Reviews online* 2018:296-325.
23. Forlani S, Masiero S, Mizzotti C. Fruit ripening: the role of hormones, cell wall modifications, and their relationship with pathogens. *Journal of experimental botany* 2019, 70(11):2993-3006.
24. Saladié M, Cañizares J, Phillips MA, Rodriguez-Concepcion M, Larrigaudière C, Gibon Y, Stitt M, Lunn JE, Garcia-Mas J. Comparative transcriptional profiling analysis of developing melon (*Cucumis melo* L.) fruit from climacteric and non-climacteric varieties. *BMC genomics* 2015, 16(1):440.
25. Waseem M, Li N, Su D, Chen J, Li Z. Overexpression of a basic helix–loop–helix transcription factor gene, SlbHLH22, promotes early flowering and accelerates fruit ripening in tomato (*Solanum lycopersicum* L.). *Planta* 2019, 250(1):173-185.
26. Garcia-Mas J, Benjak A, Sanseverino W, Bourgeois M, Mir G, González VM, Hénaff E, Câmara F, Cozzuto L, Lowy E. The genome of melon (*Cucumis melo* L.). *Proceedings of the National Academy of Sciences* 2012, 109(29):11872-11877.
27. Song X-M, Huang Z-N, Duan W-K, Ren J, Liu T-K, Li Y, Hou X-L. Genome-wide analysis of the bHLH transcription factor family in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Molecular genetics*

- and genomics 2014, 289(1):77-91.
28. Sun H, Wu S, Zhang G, Jiao C, Guo S, Ren Y, Zhang J, Zhang H, Gong G, Jia Z. Karyotype stability and unbiased fractionation in the paleo-allotetraploid *Cucurbita* genomes. *Molecular plant* 2017, 10(10):1293-1306.
  29. Sharma N. Role of bHLH93 in controlling flowering time in *Arabidopsis thaliana*. 2011.
  30. Brownlie P, Ceska T, Lamers M, Romier C, Stier G, Teo H, Suck D. The crystal structure of an intact human Max–DNA complex: new insights into mechanisms of transcriptional control. *Structure* 1997, 5(4):509-520.
  31. Heim MA, Jakoby M, Werber M, Martin C, Weisshaar B, Bailey PC. The basic helix–loop–helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. *Molecular biology and evolution* 2003, 20(5):735-747.
  32. Grant GA. The ACT domain: a small molecule binding domain and its role as a common regulatory element. *Journal of Biological Chemistry* 2006, 281(45):33825-33829.
  33. Mas-Droux C, Curien G, Robert-Genthon M, Laurencin M, Ferrer J-L, Dumas R. A novel organization of ACT domains in allosteric enzymes revealed by the crystal structure of *Arabidopsis* aspartate kinase. *The Plant Cell* 2006, 18(7):1681-1692.
  34. Kong Q, Pattanaik S, Feller A, Werkman JR, Chai C, Wang Y, Grotewold E, Yuan L. Regulatory switch enforced by basic helix-loop-helix and ACT-domain mediated dimerizations of the maize transcription factor R. *Proceedings of the National Academy of Sciences* 2012, 109(30):E2091-E2097.
  35. Chen S, Zhao H, Luo T, Liu Y, Nie X, Li H. Characteristics and Expression Pattern of MYC Genes in *Triticum aestivum*, *Oryza sativa*, and *Brachypodium distachyon*. *Plants* 2019, 8(8):274.
  36. Fernández-Calvo P, Chini A, Fernández-Barbero G, Chico J-M, Gimenez-Ibanez S, Geerinck J, Eeckhout D, Schweizer F, Godoy M, Franco-Zorrilla JM. The *Arabidopsis* bHLH transcription factors MYC3 and MYC4 are targets of JAZ repressors and act additively with MYC2 in the activation of jasmonate responses. *The Plant Cell* 2011, 23(2):701-715.
  37. Zhiponova MK, Morohashi K, Vanhoutte I, Machemer-Noonan K, Revalska M, Van Montagu M, Grotewold E, Russinova E. Helix–loop–helix/basic helix–loop–helix transcription factor network represses cell elongation in *Arabidopsis* through an apparent incoherent feed-forward loop. *Proceedings of the National Academy of Sciences* 2014, 111(7):2824-2829.
  38. Xu Z, Liu X, He X, Xu L, Huang Y, Shao H, Zhang D, Tang B, Ma H. The soybean basic helix-loop-helix transcription factor ORG3-Like enhances cadmium tolerance via increased iron and reduced cadmium uptake and transport from roots to shoots. *Frontiers in plant science* 2017, 8:1098.
  39. Zhao F, Li G, Hu P, Zhao X, Li L, Wei W, Feng J, Zhou H. Identification of basic/helix-loop-helix transcription factors reveals candidate genes involved in anthocyanin biosynthesis from the strawberry white-flesh mutant. *Scientific reports* 2018, 8(1):2721.
  40. Pagnussat GC, Yu H-J, Ngo QA, Rajani S, Mayalagu S, Johnson CS, Capron A, Xie L-F, Ye D, Sundaresan V. Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. *Development* 2005, 132(3):603-614.

41. Bruessow F, Bautor J, Hoffmann G, Parker JE. Arabidopsis thaliana natural variation in temperature-modulated immunity uncovers transcription factor UNE12 as a thermoresponsive regulator. *bioRxiv* 2019:768911.
42. Le Hir R, Castelain M, Chakraborti D, Moritz T, Dinant S, Bellini C. At bHLH68 transcription factor contributes to the regulation of ABA homeostasis and drought stress tolerance in Arabidopsis thaliana. *Physiologia plantarum* 2017, 160(3):312-327.
43. Huang H, Gao H, Liu B, Fan M, Wang J, Wang C, Tian H, Wang L, Xie C, Wu D. bHLH13 Regulates Jasmonate-Mediated Defense Responses and Growth. *Evolutionary Bioinformatics* 2018, 14:1176934318790265.
44. Zhou M, Memelink J. Jasmonate-responsive transcription factors regulating plant secondary metabolism. *Biotechnology Advances* 2016, 34(4):441-449.
45. Hu Y, Han X, Yang M, Zhang M, Pan J, Yu D. The transcription factor INDUCER OF CBF EXPRESSION1 interacts with ABSCISIC ACID INSENSITIVE5 and DELLA proteins to fine-tune abscisic acid signaling during seed germination in Arabidopsis. *The Plant Cell* 2019, 31(7):1520-1538.
46. Miura K, Jin JB, Lee J, Yoo CY, Stirm V, Miura T, Ashworth EN, Bressan RA, Yun D-J, Hasegawa PM. SIZ1-mediated sumoylation of ICE1 controls CBF3/DREB1A expression and freezing tolerance in Arabidopsis. *The Plant Cell* 2007, 19(4):1403-1414.
47. Shan W, KUANG JF, LU WJ, CHEN JY. Banana fruit NAC transcription factor MaNAC 1 is a direct target of MalCE 1 and involved in cold stress through interacting with MaCBF 1. *Plant, cell & environment* 2014, 37(9):2116-2127.
48. Schweizer F, Fernández-Calvo P, Zander M, Diez-Diaz M, Fonseca S, Glauser G, Lewsey MG, Ecker JR, Solano R, Reymond P. Arabidopsis basic helix-loop-helix transcription factors MYC2, MYC3, and MYC4 regulate glucosinolate biosynthesis, insect performance, and feeding behavior. *The Plant Cell* 2013, 25(8):3117-3132.
49. Dombrecht B, Xue GP, Sprague SJ, Kirkegaard JA, Ross JJ, Reid JB, Fitt GP, Sewelam N, Schenk PM, Manners JM. MYC2 differentially modulates diverse jasmonate-dependent functions in Arabidopsis. *The Plant Cell* 2007, 19(7):2225-2245.
50. Li T, Xu Y, Zhang L, Ji Y, Tan D, Yuan H, Wang A: The jasmonate-activated transcription factor MdMYC2 regulates ETHYLENE RESPONSE FACTOR and ethylene biosynthetic genes to promote ethylene biosynthesis during apple fruit ripening. *The Plant Cell* 2017, 29(6):1316-1334.
51. An JP, Zhang XW, Bi SQ, You CX, Wang XF, Hao YJ. Mdb HLH 93, an apple activator regulating leaf senescence, is regulated by ABA and Md BT 2 in antagonistic ways. *New Phytologist* 2019, 222(2):735-751.
52. Chipman DM, Shaanan B. The ACT domain family. *Current opinion in structural biology* 2001, 11(6):694-700.
53. Zheng Y, Wu S, Bai Y, Sun H, Jiao C, Guo S, Zhao K, Blanca J, Zhang Z, Huang S. Cucurbit Genomics Database (CuGenDB): a central portal for comparative and functional genomics of cucurbit crops. *Nucleic acids research* 2018, 47(D1):D1128-D1136.



54. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A. The Pfam protein families database in 2019. *Nucleic acids research* 2018, 47(D1):D427-D432.
55. Johnson LS, Eddy SR, Portugaly E. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC bioinformatics* 2010, 11(1):431.
56. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC bioinformatics* 2009, 10(1):421.
57. Jin J, Tian F, Yang D-C, Meng Y-Q, Kong L, Luo J, Gao G. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic acids research* 2016:gkw982.
58. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic acids research* 2016, 45(D1):D200-D203.
59. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. In: *The proteomics protocols handbook*. Springer; 2005: 571-607.
60. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in bipolymers. 1994.
61. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee T-h, Jin H, Marler B, Guo H. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic acids research* 2012, 40(7):e49-e49.
62. Chen C, Xia R, Chen H, He Y. TBtools, a Toolkit for Biologists integrating various biological data handling tools with a user-friendly interface. *BioRxiv* 2018:289660.
63. Huang S, Li R, Zhang Z, Li L, Gu X, Fan W, Lucas WJ, Wang X, Xie B, Ni P. The genome of the cucumber, *Cucumis sativus* L. *Nature genetics* 2009, 41(12):1275.
64. Wu S, Wang X, Reddy U, Sun H, Bao K, Gao L, Mao L, Patel T, Ortiz C, Abburi VL. Genome of 'Charleston Gray', the principal American watermelon cultivar, and genetic characterization of 1,365 accessions in the US National Plant Germplasm System watermelon collection. *Plant biotechnology journal* 2019.
65. Wu S, Shamimuzzaman M, Sun H, Salse J, Sui X, Wilder A, Wu Z, Levi A, Xu Y, Ling KS. The bottle gourd genome provides insights into Cucurbitaceae evolution and facilitates mapping of a Papaya ring-spot virus resistance locus. *The Plant Journal* 2017, 92(5):963-975.
66. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution* 2018, 35(6):1547-1549.
67. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology* 2010, 59(3):307-321.
68. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks.

Genome research 2003, 13(11):2498-2504.

69. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature communications 2019, 10(1):1523.

## Figures

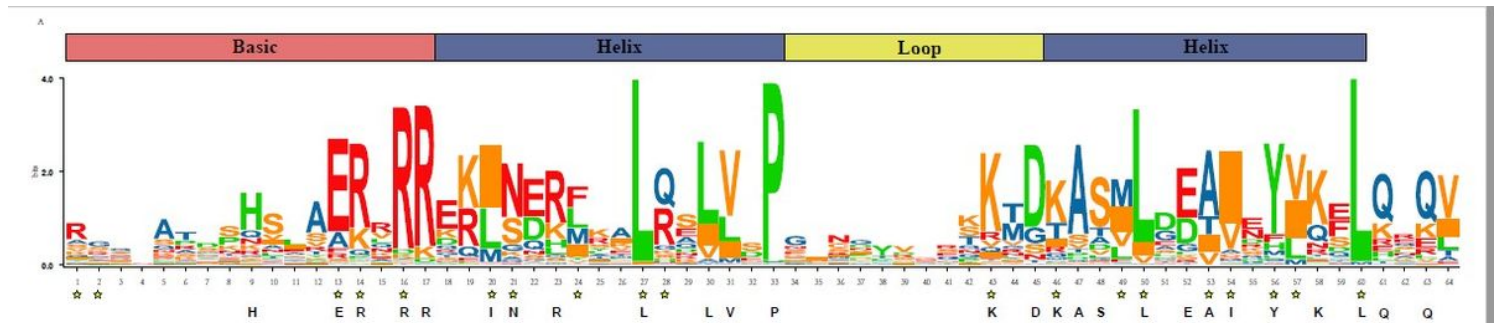


Figure 1

Sequence logo of bHLH domain of CmbHLH. The high of letter represent the conservation of the sequence at that site. Stars represent the conservation sites of bHLH in Arabidopsis. Bold letters represent the conservation sites of the bHLH domain that higher than 50% consensus ratio.

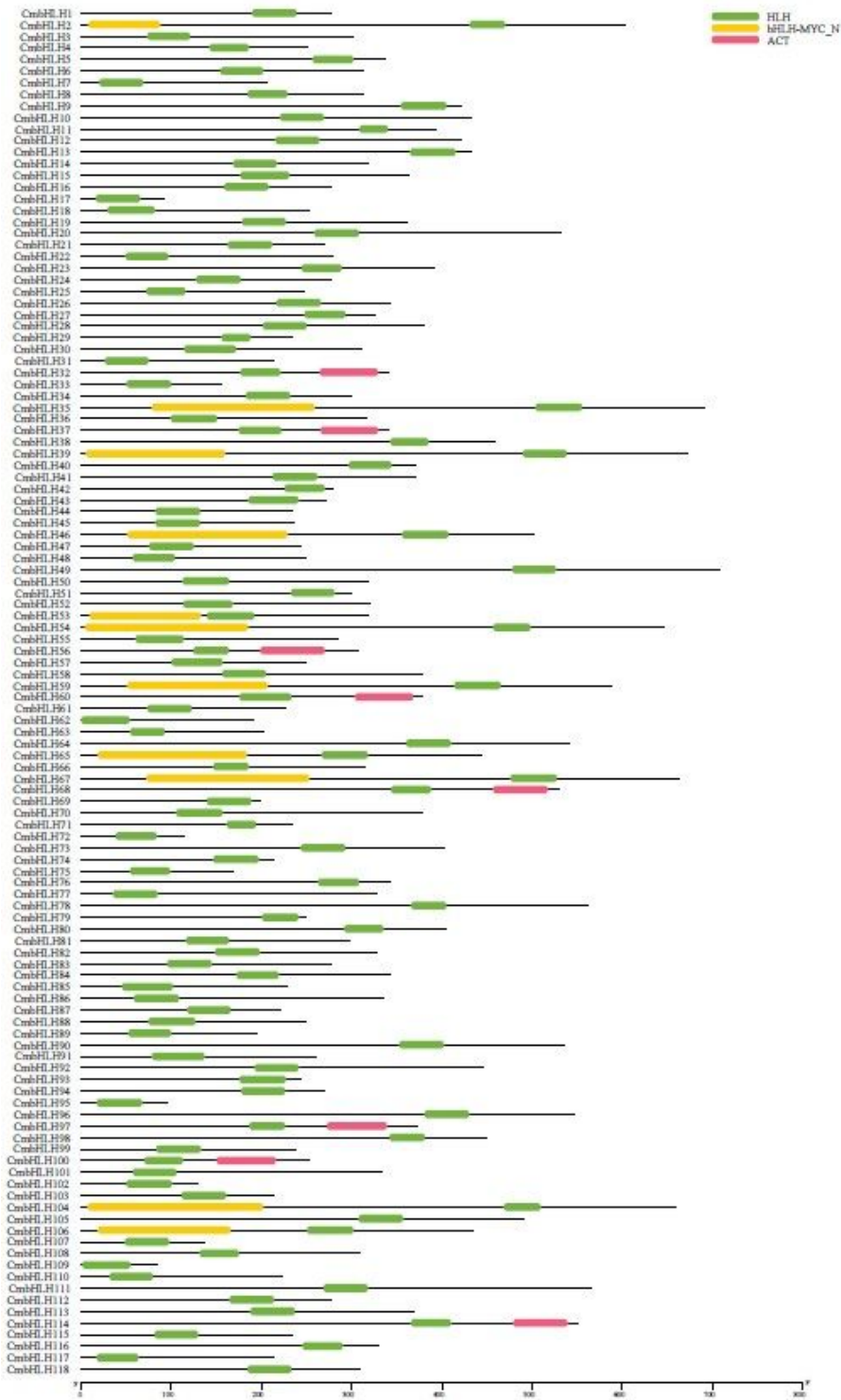
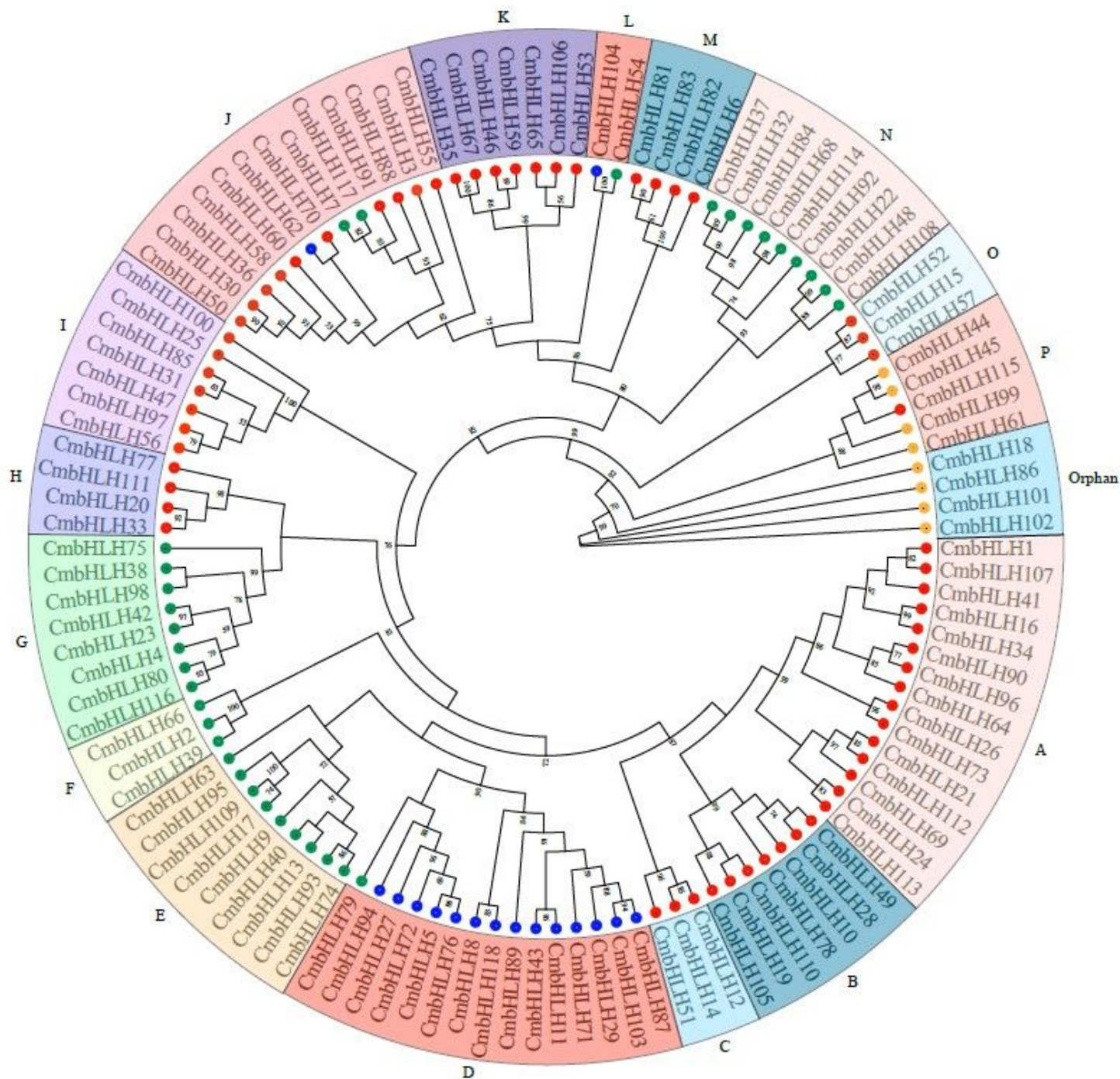


Fig 2 Scheme of CmbHLH domain distribution. Different domains were indicated by different colored rectangles, lines represent CmbHLHs protein length

## Figure 2

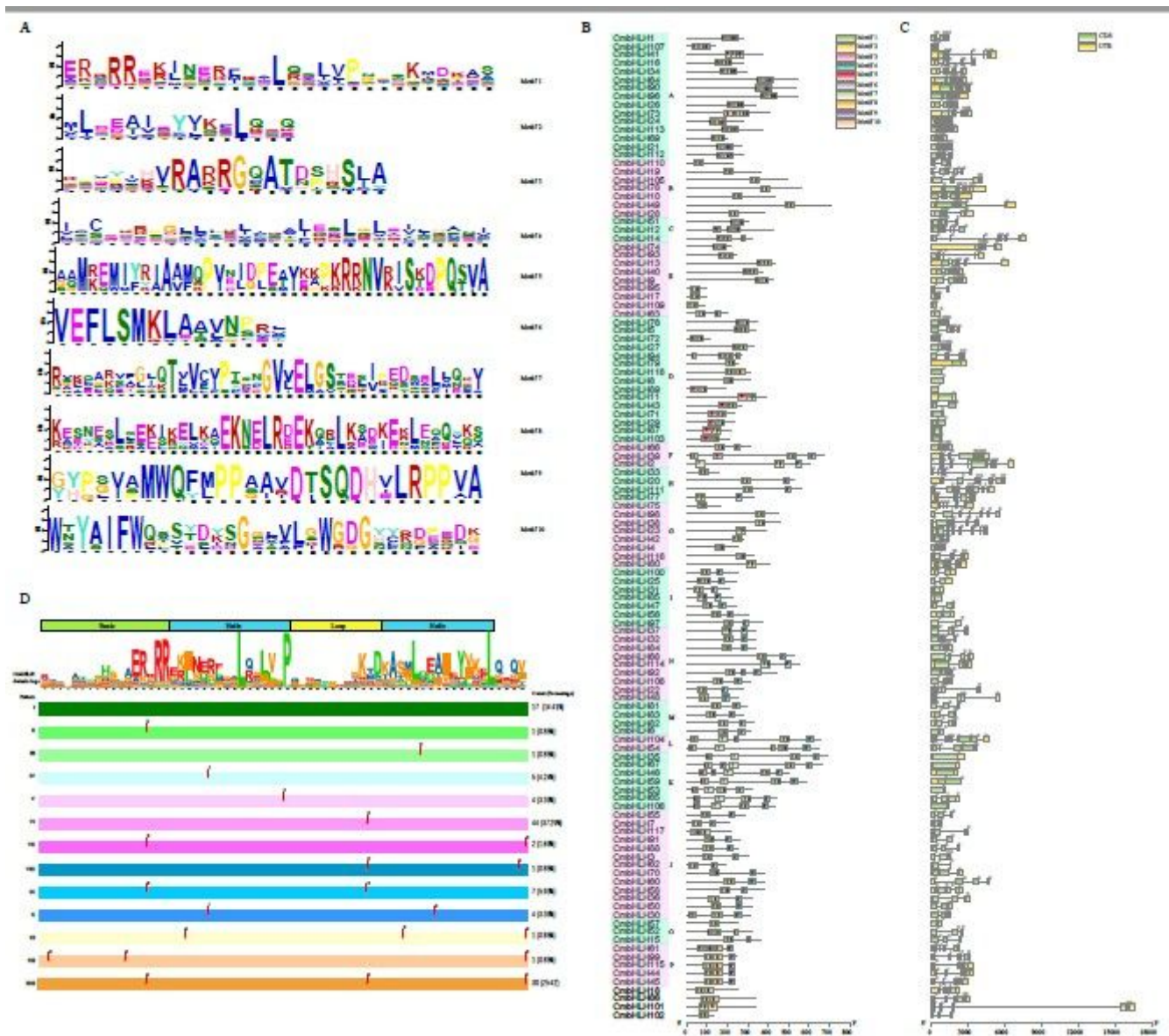
The domain distribution and types of CmbHLH.



**Figure 3**

ML phylogenetic tree of the CmbHLH with the predicted DNA-binding activities. The tree shows the 16 phylogenetic subfamilies marked with different colored background. The different colored dot indicated A (yellow), B (red), C (blue), D (green) four bHLH DNA binding activities groups according to Atchley.

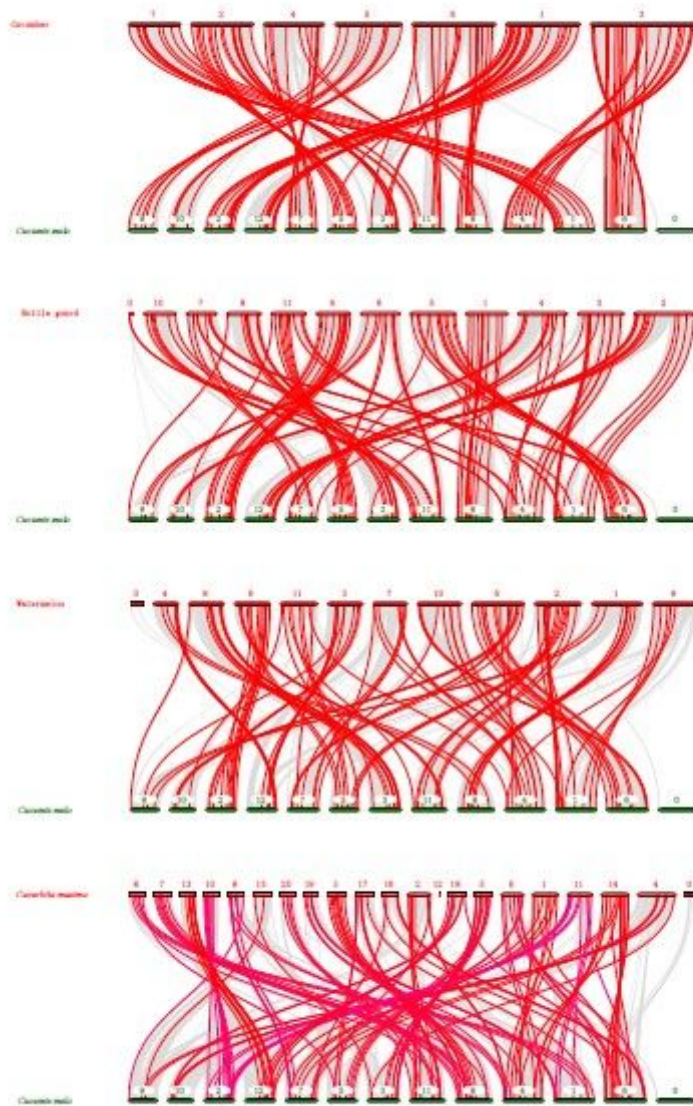




**Figure 4**

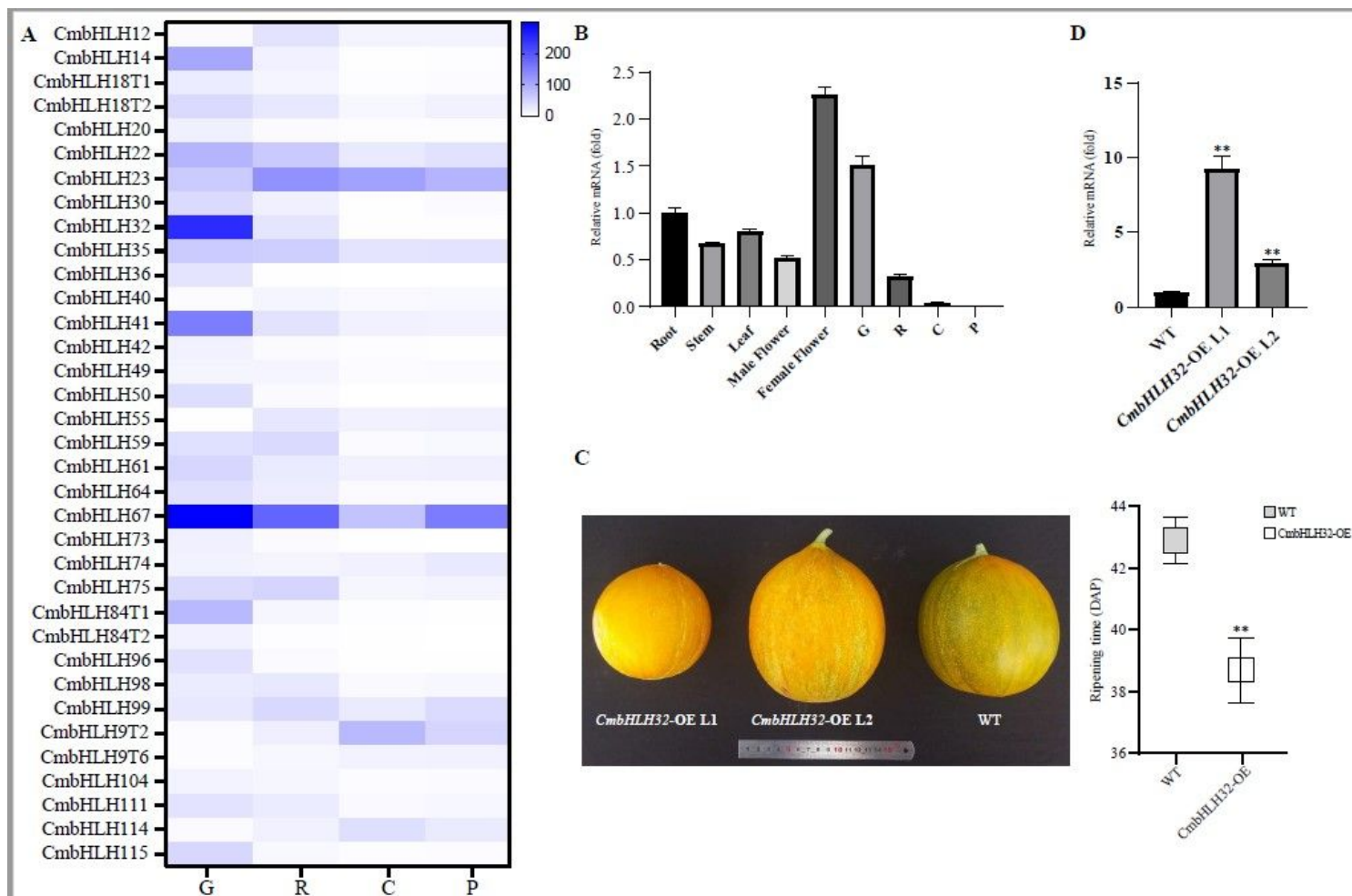
Sequence logo and arrangement of conserved motifs, gene structure and intron pattern of CmbHLH. (A) Sequence logo of 10 conserved motifs of CmbHLH. (B) Arrangement of conserved motifs. Motifs labeled with 1 to 10 in different colores. (C) Intron-exon structure of CmbHLH. The exon, intron and UTR were represented by green rectangle, lines and yellow rectangle, respectively; the numbers (0-3) stand for intron phase. (D) Intron distribution pattern of bHLH of CmbHLH. The bars (labeled from I to XIII) illustrate the intron distribution pattern of coding sequence of bHLH domain within CmbHLHs. The arrows indicate the intron positions, the numbers (0-3) above the arrows represent the intron phase. The count and percentage of genes in each pattern are shown on the right.





**Figure 6**

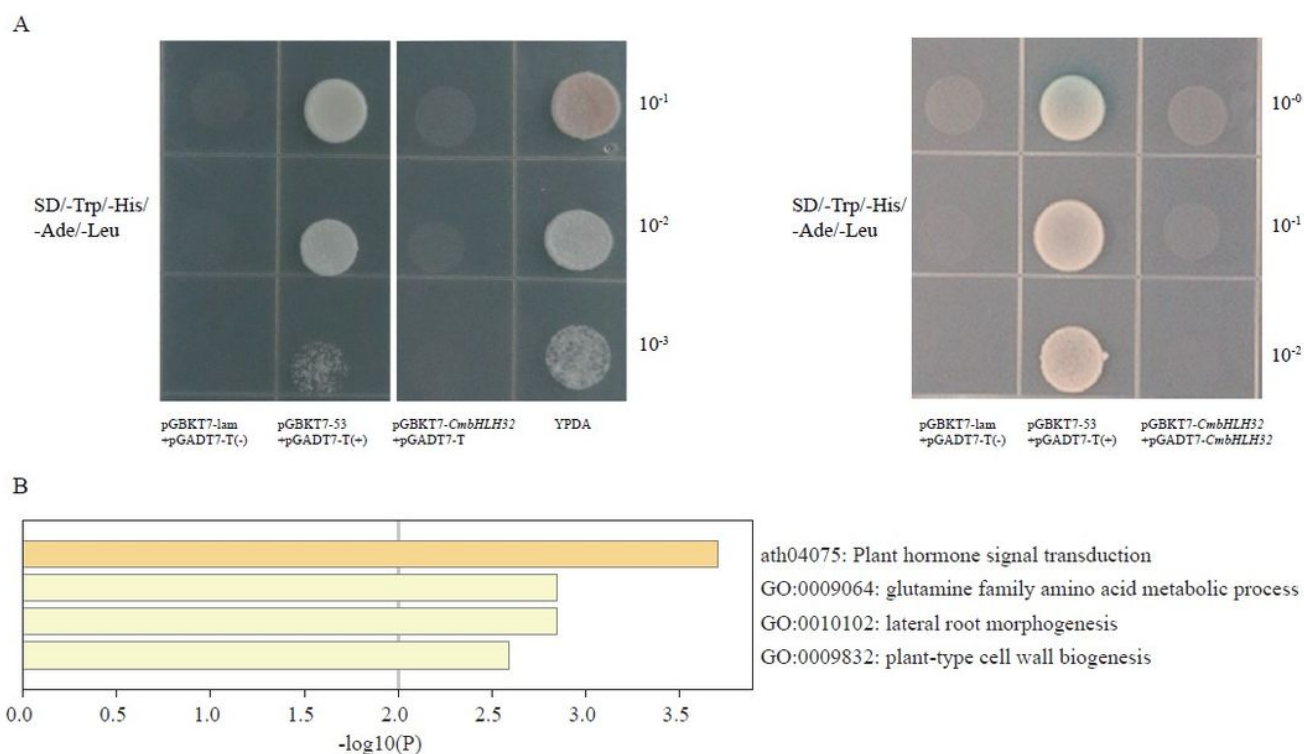
Collinearity analysis of CmbHLH between melon and four Cucurbitaceae plants. The grey lines indicate the collinear blocks within melon and other plant genomes, the numbers represent the chromosome numbers of plant genomes.



**Figure 7**

The differential expression profile of CmbHLH, different tissues and CmbHLH32-OE transgenic lines expression of CmbHLH32. (A) Heatmap of CmbHLH differential expression genes. (B) Quantitative PCR analysis of different tissues expression of CmbHLH32. The root was normalized to 1. (C) The relative expression level of CmbHLH32-OE transgenic lines and WT fruit. (D) Fruit ripening of CmbHLH32-OE transgenic lines was earlier than WT. Quantitative PCR expression level was calculated by  $2^{-\Delta\Delta CT}$  method, the results were represented by means  $\pm$  standard deviations, \*\* p < 0.01, N > 3.





**Figure 8**

Transcriptional activation activity of CmbHLH32 and GO annotation of correlation expression genes of CmbHLH32. (A) Transcriptional activation activity of CmbHLH32. Full length of CmbHLH32 fused to GAL4 domain in pGADT7-T vector and cotransformed with pGBKT7-53 vector for monomer transcriptional activation activity of CmbHLH32. For homodimerizes transcriptional activation activity tests, Full length of CmbHLH32 fused into pGADT7 and pGBKT7 were cotransformed into yeast AH109 strain. The growth performances of the transformed yeast cells were exhibited in different dilutions on SD/-Trp/-His/-Ade/3-AT solid medium. (B) GO annotation of correlation expression genes of CmbHLH32 genes.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTableS2.xlsx](#)
- [SupplementaryFigureS1.pdf](#)
- [SupplementaryTableS1.xlsx](#)