Genome-wide investigation of CCCH zinc finger family in longan (*Dimocarpus longan* Lour): characteristic identification and expression profiles in longan somatic embryo

Liyao Su
Fujian Agriculture and Forestry University
Mengqi Jiang
Fujian Agriculture and Forestry University
Shuqi Huang
Fujian Agriculture and Forestry University
Xiaodong Xue
Fujian Agriculture and Forestry University
Xue Li
Fujian Agriculture and Forestry University
Zhongxiong Lai
Fujian Agriculture and Forestry University
Yuling Lin (buliang84@163.com)
Fujian Agriculture and Forestry University

Research article

Keywords: Dimocarpus longan Lour, CCCH Zinc finger, characteristic identification, expression profile

Posted Date: January 16th, 2020

DOI: https://doi.org/10.21203/rs.2.21020/v1

Abstract

Background: CCCH Zinc finger (Znf) transcription factors (TF), as a novel type Znf genes, regulate genes expression by binding on their mRNA and play important roles in plant abiotic stress, growth and development. However, no overall genome-wide analysis or expression profiling of CCCH (C3H) gene family in *Dimocarpous longan* , especially during the early stages of somatic embryo in longan has been studied. Longan is a tropical/subtropical fruit tree of great economic importance in Southeast Asia,and longan embryogenesis is the main factor affecting fruit quality and yield.

Results: In this study, a comprehensive analysis of longan C3H DlC3H gene family was carried out. 49 DlC3H genes were identified from longan genome database,which divided into 3 clades. Besides, genes characteristics, phylogenetic tree, gene structure, motif composition were comprehensively analyzed. The analysis of alternative splicing events (AS) suggested that AS events of DlC3H genes were related to longan non-embryonic and embryonic callus transformation. Promoter analysis indicted that most of DlC3H genes included cis -elements associated with hormones and stress response. Quantitative real-time PCR analysis indicated that 26 DlC3Hs ,which possess MeJA and ABA responsive cis -elements, showed different expression patterns and may involved into ABA and MeJA signaling pathway. The expression profiles of 17 DlC3Hs were performed in four stages of longan, the results showed that only DlC3H01/07/14/16/38 was consistent with the data in the transcriptome. DlC3H 07/14/16/36/49 were highly expressed in EC and only DlC3H 04/38 was in GE , suggesting that they have different functions in embryonic development. Finally, sRNAs were verified involved into regulating 6 DlC3Hs .

Conclusion: This study provides the first systematic analysis of CCCH protein in longan somatic embryo. Particularly, CCCH genes may be involved in hormone and stress respond, and somatic embryogenesis. Our results presented here may provide a insight into the characteristics and functions of this family in somatic embryogenesis.

Background

Transcription factors (TFs), as a gene widely distributed in plants, play an important role in the growth and development of plants and morphogenesis[1, 2]. Zinc finger (Znf) transcription factors is one of the largest TF families containing RING-finger[3], LIM[4], WRKY[5] and DOF[6] gene families which regulate gene expression through DNA-binding and protein binding proteins. However, recent evidence suggests that CCCH (C3H) as a novel type Znf transcription factors regulate gene expression by binding on targets genes mRNA[7-9]. To verify its characteristics, based on the genome-data, the comprehensive analysis of *C3H* Znf family was performed in dicotyledonous *Arabidopsis thaliana*[9], *Medicago truncatula*[10], *Populus trichocarpa*[11], *Clementine mandarin*[12], *Vitis vinifera*[13], *Cicer arietinum*[14], monocotyledonous *Musa acuminata*[15], *Zea mays*[16], and *Oryza sativa*[9] , showed that *C3H* Znf family widely involved in biotic and abiotic pathways. Besides, *C3H* Znf members have various functions. In *Arabidopsis*, *ATSZF1* and *ATSZF2* negatively regulate the expression of salt responsive genes[17]. Cotton *GHZFP1* can be induced by salt, drought and SA, and its transgenic tobacco shows resistance to fungal diseases[18]. Rice *OsDOS* controls leaf senescence through jasmonate (JA) pathway[19]. *Arabidopsis* HUA1 is confirmed a regulator

for flower morphogenesis[20]. Moreover, *C3H* Znf genes is an essential regulator for plants somatic embryogenesis. Previous studies have reported that PEI1, as an embryo-specific expression *C3H* Znf gene, directly regulate the heart-shaped embryo development in *Arabidopsis*[21]. Cucumber *CsSEF1* shows the importance of controlling cell polarity, and marks the cotyledon primordia and procambium tissues in later developmental stages[22]. The above studies indicate that *C3H* Znf family plays a significant role in plant abiotic stress, growth and development and somatic embryo morphogenesis.

Sapindaceae plants are widely distributed in tropical and subtropical areas, including important tropical fruit trees such as longan, lychee, *Nephelium lappaceum*, and well-know oil plants such as *Sapindus mukurossi*, *Xanthoceras sorbifolium*. To date, many draft genome sequence of plants have been identified, which has greatly promoted the research of corresponding plants. Longan, as the first Sapindaceae plant completeing the genome sequencing[23], provides a reference for studying the molecular genetic characteristics of Sapindaceae plants. So far, comprehensive analysis of gene family in Sapindaceae plants is still limited. Only longan *WRKY*[24], Ubiquitin-conjugating enzymes[25], *Laccase*[26] families have been comprehensively analyzed. As an important fruit tree, fruit quality of longan is closely related to economic effects. Embryonic development as one of the main factors regulate longan fruit quality. Therefore, understanding the mechanism of longan embryonic development is critical for improving longan fruit quality.

Despite the *CCCH* Znf gene is great significance in plants, the comprehensive analysis of *CCCH* gene family in plants embryo has not been performed. Longan genome sequencing provides an opportunity to reveal the function at the genome-wide level[23]. The 49 *DlC3Hs* were identified from longan genome database. We further analyzed the gene characteristics, phylogenetic tree, gene structure, motif composition, alternative splicing events and promoter *cis*-elements. Additionally, the expression profiles of 26 *DlC3Hs* were carried out by RT-qPCR to explore their responses to methyl jasmonate (MeJA), abscisic acid (ABA) and their endogenous inhibitor (Salicylhydroxamic acid, SHAM, Sodium Tungstate Dihydrate, STD) treatment. According to the transcriptome data, 17 *DlC3Hs* were selected to analyze the expression levels in longan embryogenic callus（EC）, incomplete compact pro-embryogenic cultures（ICpEC）, globular embryos（GE）, non-embryonic callus（NEC）and their cleavage sites were verified . Our preliminary results might provide valuable clues for researching the function of the *CCCH* Znf gene family in plant embryonic development.

Results

Analysis the characteristics of longan *C3H* gene family

According to annotation files of InterPro software, the 68 candidate longan *CCCH* Znf family members were found in longan genome database. Then, the BLASTP program and CD search were performed. A total 49 non-redundant *CCCH* Znf genes were confirmed in longan, then we named them *DlC3H01* to *DlC3H49.* Gene characteristics, including the *Arabidopsis* orthologs locus, number of exons, length of CDS, molecular weight (kD), isoelectric point (PI), number of CCCH motif and subcellular localization were showed in Table 1. Among the 49 *DlC3H* genes, the *DlC3H41* was identified to be the smallest protein with

136 amino acid, whereas the *DlC3H27* was largest with 1811 amino acid. The number exons of the genes range from 1 to 14, the kD range from 14.46 (*DlC3H41*) to 198.20 (*DlC3H27*), and the PI range from 4.90 (*DlC3H33*) to 9.50 (*DlC3H28*). In addition, the number of CCCH motif of *DlC3H* gene family was the same as that in *Arabidopsis* and rice, which was range from 1 to 6. Finally, the subcellular location showed that 9 of *DlC3H* members were located in cytoplasm, 35 *DlC3Hs* members were located in nucleus, and the rest was secreted protein.

Phylogenetic analysis and conserved motif multiple sequence alignment

A phylogenetic tree of longan and *Arabidopsis* was constructed by maximum likelihood (ML) method based the full length of protein sequence. The phylogenetic analysis showed that DlC3Hs and AtC3Hs gene family was divided into 3 clades contained 21, 39 and 55 members, respectively (figure 1). DlC3Hs had 9, 19 and 21 members in each of the three clades. In the first clade, four AtC3Hs members were not classified, and all members of the longan were classified. These results indicted that DlC3Hs had three different evolutionary directions. Such as AtC3H51 (PEI1), as a key protein for plant embryogenesis, was clustered with DlC3H01, speculated that they had similar function. The longan CCCH Znf domains were further multiple aligned according to the phylogenetic tree. The AtC3Hs (AtC3H01, AtC3H51, AtC3H08) in each clades were selected for a representatives. The results showed that longan CCCH Znf domain sequences were highly conserved in each clades with the length range from 19 to 27 amino acids (figure 2). And it basically belonged to C-X_8-C-X_5-C-X_3-H and C-X_7-C-X_5-C-X_3-H types, suggesting that these two types were parallel evolutionary. Besides, the conservation of clade was the worst. There are three different types domain in clade belonging to DlC3H15-1 (C-X_9-C-X_5-C-X_3-H), DlC3H21 (C-X_7-C-X_4-C-X_3-H) and DlC3H25 (C-X_9-C-X_4-C-X_3-H).

Gene structure and motif composition of *DlC3Hs*

The introns and exons of all 49 *DlC3Hs* were identified for better understanding the evolution of *DlC3Hs*. As shown in figure 3B, among the 49 *DlC3Hs*, the number of exons were range from 1 to 14 (eight with one exons, seven with two exons, six with three and four exons, two with five exons, one with six exons, nine with seven exons, three with eight exons, one with nine exons, three with ten exons, one with 11 exons, one with 12 exons and one with 14 exons). In the same class, genes usually had the same structure, such as class , except *DlC3H08*, they all contained one intron. All class e/f members had no intron, except *DlC3H26/15*. Besides, within the same class, the intron structure were highly consistent. Although the gene structure and the introns phase were similar with phylogenetic relationship, the different between classes were significant.

The conserve motif was identified by CDD. Comparing the previous researches, the motif found in longan C3H family was the most containing 25 types (figure 3C). Only one C3H domain was observed in 16 *DlC3Hs*. The rest genes all possessed 2 to 5 domains. The cluster genes (*DlC3H23/27*, *DlC3H13/20*,

DlC3H45/10) had consistent motif composition indicating functional similarity in longan. In addition, some motifs were unique to one group, for example, motif 6, motif 7 and motif 23 were special to class Ⅱd, Ⅱa and Ⅱf, respectively. The differentiation of motifs between different members reflected the functional diversification of *DlC3Hs*, and the function of motifs needed further verification. Overall, *DlC3Hs* members consisting of the same gene structure and motif composition were clustered into one branch of phylogenetic tree implying it's highly conserved.

Analysis the AS events of *DlC3Hs* in longan non-embryonic and embryonic cultures

According to the RNA-seq analysis of longan NEC, EC, ICpEC and GE, the alternative splicing events of *DlC3Hs* were identified. A total of 445 AS events, including alternative 3' splice site (A3'S), alternative 5' splice site (A5'S), intron retaintion (IR) and exon skipping (ES), were detected from 29 *DlC3Hs*. The type of AS event and the statistics of AS events in 29 *DlC3Hs* was showed in Table 2. A3'S events (26.17%) were more frequent than A5'S events(18.30%). IR events were the most frequent with 45.17% (Table 2). This result was the same with previous studies which considered IR events were the most frequent events of AS in plants. Furthermore, the number of genes that with A3'S, A5'S and IR events was basically the same (Table 2). In addition, as the figure 4A shown, AS events might play a key role in longan somatic embryo morphological. For example, in EC stage, IR event sharp decrease and A3'S/A5'S marked increase. The ES event slight rise in ICpEC and GE stages.Meantime, we counted the number of AS events in longan NEC, EC, ICpEC and GE. The results shown that the AS events occur most frequently in the NEC stage and least frequently in the EC stage.(figure 4B). This result suggested that the AS events in *DlC3Hs* was related with longan somatic embryogenesis.

Stress and hormone related *cis*-elements in *DlC3Hs* promoter

To further explore the potential regulatory mechanism of *DlC3Hs* during external stress, the promoters regions, which were up-stream 2Kb sequences of *DlC3H* genes translation starts site, were submitted into PlantCARE database to search *cis*-elements. A total of 559 *cis*-elements related to hormone and stress were detected in *DlC3H* genes (Figure 7). Among them, except *DlC3H07*, *DlC3H40* and *DlC3H49*, the rest genes contained at least 1 anaerobic induction element. Meanwhile, drought and low-temperature related *cis*-elememts possessed in 25 and 16 *DlC3Hs*, respectively. This result showed that *DlC3H* family might response these abiotic stress. In addition, 36 *DlC3H* genes contained 164 MeJA responsive *cis*-elements and 31 *DlC3Hs* possessed 88 abscisic acid responsive element indicating that MeJA and ABA play a key role in *DlC3Hs* regulatory. Furthermore, 34 auxin-responsive elements existed in 20 *DlC3Hs* and 38 gibberellin-responsive elements were found in 23 *DlC3Hs*. 29 salicylic acid responsive element was located in 20 *DlC3Hs*. On the whole, the *cis*-element analysis suggested that *DlC3Hs* family could involved into abiotic stress and hormone responsive.

Expression patterns of *DlC3H* genes after ABA, MeJA and theirs endogenous inhibitor treatments

According to the potential *cis*-elements analysis above, 26 *DlC3H* members, which possessed MeJA and ABA responsive *cis*-element, were selected from 49 *DlC3H* genes. The qPCR was performed to analyze their expression patterns after the identical concentration of MeJA, ABA and their endogenous inhibitor treatments. In ABA treatment, among the 26 *DlC3Hs*, 10 were up-regulated, 8 were down regulated and 8 *DlC3Hs* were no changed (Figure 6). STD was the inhibitor of endogenous ABA. In STD treatment, among the 26 *DlC3Hs*, 4 were up-regulated, 13 were down regulated and 9 were no changed (Figure 6). Some of *DlC3Hs* showed the opposite trends in ABA and STD treatment, such as *DlC3H10/ 24/ 28/ 37/ 45/46* (Figure 6). Most of *DlC3Hs* signal significantly up-regulated responded MeJA. However, in SHAM treatment, the expression of *DlC3Hs* was almost invariant compared the control. In addition, several *DlC3Hs* (*DlC3H09/24/26/28/30/33/37/46*) were up-regulated in MeJA treatment, and down-regulated in SHAM treatment (Figure 7). This results implying that *DlC3Hs* were involved into ABA and MeJA signaling pathway.

Expression profiling of *DlC3Hs* with RNA-seq and qPCR in longan non-embryonic and embryonic cultures

The expression patterns of longan CCCH family in the longan NEC, EC, ICpEC and GE transcriptomes were investigated in this study (Transcriptome datas of *DlC3H02*, *DlC3H08*, *DlC3H28*, *DlC3H29*, *DlC3H30* and *DlC3H32* were absent.). As the figure 8 showed that the expression of 43 *DlC3Hs* was divided into 2 group. In the group Ⅰ, they were at a low expression levels in NEC stage, and high expression between EC stage and GE stage indicating that these genes were related to embryonic of longan somatic embryo. Moreover, 3 *DlC3H* genes were specific in GE stage and 2 *DlC3H* genes in ICpEC stage. 12 *DlC3Hs* highly expressed in NEC and EC stage, which were clustered at group Ⅱ. This results implied that these genes which highly expressed in specific stage might involved into their morphogenesis .

To further confirm whether the specific expression of *DlC3Hs* could regulate longan somatic embryo morphogenesis of specific stage, 17 *DlC3Hs* which highly expressed in a special stage were selected to study. Then, the qPCR was carried out to verify the expression patterns of these *DlC3Hs* in longan early SE. The results showed that only *DlC3H01/07/14/16/38* was consistent with the data in the transcriptome. *DlC3H05*, *DlC3H31*, *DlC3H39*, *DlC3H43* and *DlC3H47* were down regulated during longan SE, and *DlC3H38* and *DlC3H41* showed the reverse trend,suggesting that members of the *DlC3Hs* gene family may have different functions in embryonic development. Whilst, 6 *DlC3Hs* (*DlC3H07/11/14/16/36/49*) were highly expressed in EC, and there were lower expression level of most *DlC3Hs* in ICpEC and GE than NEC and EC (Figure 9).

Small RNA involved into *DlC3Hs* transcription

Small RNAs played an important role in plant growth and development .These regulatory small RNAs (mainly include miRNAs and ta-siRNAs, sic passim) negatively regulate gene expression at post-transcriptional level by directing the cleavage of target transcript (mRNA)[30].Li Yiqun reperted that the MulZF1 which s a zinc finger protein containing CCCH domain is the target gene of mul-miRn26 in *Morus alba* L[31].To understand whether the *DlC3Hs* were regulated by sRNA in longan, the modified RLM-RACE was carried out to verified the cleavage site of 17 *DlC3Hs* which highly expressed in a special stage.As the figure 10 shown, among the 17 *DlC3Hs*, the fragments of 6 *DlC3Hs* (*DlC3H01/03/05/11/19/39*) were detected. The 6 *DlC3Hs* had 1 to 5 cleavage sites. Meantime, the longan small RNA (sRNA) database was used to predict the potential sRNA that could cleaved the 6 *DlC3Hs*. As the results shown, the 14 cleavage sites of 6 *DlC3Hs* were identified as the putative cleavage site for 131 sRNAs (Figure 10, Additional file 2 to 7). This implied that sRNAs could widely involve into *DlC3Hs* pathway. For example, each of three cleavage sites of *DlC3H01* could be combined with 4, 5 and 17 sRNA, respectively. Among these sRNA, 21 sRNA had been registered in miRBase database. It is suggested that miRNA could regulate *DlC3Hs* in longan somatic embryogenesis. Meantime, a larger number animal origin miRNAs were found in these sRNAs, indicating that the *C3H* family might conserved between plant and animal in terms of the formation principle of miRNA. Furthermore, the rest 5 sRNA had no similar in miRBase database and their had a reliable E value (one with 1.5, one with 2.5, three with 3.0). Thus, we speculated that they might be siRNA or piRNA.

Discussion

Evolutionary conservation and divergence of the *DlC3H* gene family of longan

In plants, there are many studies about *CCCH* Znf family. Because of lacking draft genome sequence, *CCCH* Znf family has not been reported in Sapindaceae plants. The publication of longan draft genome sequence provides a solid foundation for our study[23]. We identify 49 *CCCH* Znf genes from longan genome, the number of is less than that of *Arabidopsis* (68)[9], rice (69)[9], maize (68)[16], *Vitis vinifera* (69)[13], *Clementine mandarin* (62)[12], more than *Medicago truncatula* (34)[10]. The size of longan genome is 445 Mb[23], which is greater than *Arabidopsis* (125 Mb)[32], rice (375 Mb)[33], *Clementine mandarin* (367 Mb)[34], smaller than *Vitis vinifera* (125 Mb)[35], maize (2300 Mb)[36], *Medicago truncatula* (500 Mb)[37]. This results show that the number of plants CCCH Znf genes is may independent with the plants genome size, implying that it is related to species evolution. In addition, the C-X_{7-8}-C-X_5-C-X_3-H types of CCCH Znf domain is conserved in plants. In longan, it occupies 96.62%, which outclass *Populus trichocarpa* (72%)[11] , *Arabidopsis* (83%)[9] , *Vitis vinifera* (79.8%)[13] , maize (79.4%)[16] and *Cicer arietinum* (82.3%)[12] . This indicates that longan CCCH Znf domain is highly conserved. Furthermore, there are three nonconservative CCCH Znf domains in longan. Among them, C-X_9-C-X_5-C-X_3-H and C-X_7-C-X_4-C-X_3-H domain have been reported in *Musa acuminata*[15], maize[16], rice[9], *Arabidopsis*[9]. The domain of C-

X_9-C-X_4-C-X_3-H type is first identified in longan speculated that *DlC3H25* has Special biological functions and regulatory pathways.

We further study the gene structure and motif composition of longan *CCCH* Znf family. DlC3Hs within the same groups share similar intron/exon conposition, intron phase and conserved domain. This suggest that longan CCCH Znf members are functionally conservative during the evolution. In addition, there are a larger number of conserved domain except CCCH Znf domain, which is a functional region of the protein. 15 conserved domains are identified in *Populus trichocarpa*[11], 13 in *Arabidopsis*[9], 8 in *Medicago truncatula*[10], *6 in Vitis vinifera*[13], 10 in maize[16], 16 in *Clementine mandarin*[12]. There are 25 conserved domains are found in longan implying functions diversification of longan *CCCH* Znf family. The above plants share many domains such as ANK, RRM, WD-40, KH, etc, which play important roles in various life activities such as particle transport, signal transduction, RNA/DNA recognition, RNA binding and protein interaction[38-42], suggesting the similarity and importance of *CCCH* genes function among different plants. Meanwhile, longan *CCCH* Znf members have a lot specific conserved domains for example Torus super family, PRK12678 super family, DNA_pol3_delta2 super family, etc. The results show that longan *CCCH* Znf family may involved in a wider range of life activities.

DlC3H genes may response to plant stress and hormone response

Promoter, as a non-coding region upstream of coding gene, plays a key role in regulating gene expression. In this study, identification of a large number of *cis*-acting elements associated with biotic and abiotic stresses, including MeJA (29.34%), ABA (15.20%), SA (5.19%), Auxin (6.10%), GA (6.80%), drought (6.40%), anaerobic induction (27.01%) and low-temperature responsive elements. It is suggested that the longan *CCCH* Znf family may involve in these signaling pathways. In MeJA and ABA treatment, longan *CCCH* Znf members can increase or decrease the expression in response to hormone signals. This results are similar with previous research in *Arabidopsis*[9][43], *Populus trichocarpa*[11], maize[16] and *Medicago truncatula*[10]*.* Although many studies have shown that CCCH Znf can respond to exogenous hormone signals, the effects of endogenous hormones on its expression remain unknown. In endogenous inhibitor of MeJA and ABA treatment, the effect of endogenous hormone inhibitors on some *DlC3Hs* are more obvious than exogenous hormones. Besides, some *DlC3Hs* have opposite expression trends to hormone and their endogenous inhibitors, such as *DlC3H10/ 24/ 28/ 37/ 45/46* to ABA and STD *DlC3H09/24/26/28/30/33/37/46* to MeJA and SHAM. These results demonstrate the importance of the CCCH Znf transcription factor family in plant stress and hormone response.

Potential roles of DlC3H genes during plants somatic embryo

Comprehensive analysis of the *CCCH* Znf family has not been reported in plants somatic embryo. Combined with longan transcriptome data and RT-qPCR analysis, only *DlC3H01/07/14/16/38* was

consistent with the data in the transcriptome. *DlC3H07/11/14/16/19/36//38/49* were highly expressed in single stage suggesting that this members can participate in specific stage of longan somatic embryo morphogenesis. Most of them are highly expressed in EC, indicating that these *DlC3Hs* play an important role in the formation of longan embryonic cells. Some *DlC3Hs* up-regulated (*DlC3H38/41*) or down-regulated (*DlC3H05/31/39/43/47*) during NEC to GE indicate that up- or down-regulation of these genes can promote the formation and differentiation of longan somatic cells. These result are similar to the function of *AtPEI1*[21] and *CsSEF1*[22] in somatic embryogenesis. Moreover, sRNA is an endogenous non-coding small molecule regulator, and many studies have shown that sRNA is of importance role to longan somatic embryogenesis[44-46]. While *DlC3Hs* is involved in plant growth and stress response, it is also regulated by sRNA. 14 cleavage sites of 6 *DlC3Hs* may be regulated by 131 presumed sRNA. Moreover, a large number of animal-derived sRNA also reflects the conservatism of *CCCH* Znf family in the animals and plants.

Conclusions

In conclusion,49 *DlC3H* genes were identified in the longan genome,which divided into 3 clades. The results of a comprehensive analysis demonstrate the importance of CCCH zinc fnger genes in the regulation of plant somatic differentiation and in response to hormone and stresses. A systematic and comprehensive analysis of longan *CCCH* Znf family is conducive to further screening *DlC3Hs* for functional identification, as well as to improving longan fruit quality and enhancing genetic improvement against stress.

Methods

Plant materials and treatments

The ‘HHZ’ longan friable-embryogenic callus (EC) that preserved by Institute of Horticultural Biotechnology, fujian agriculture and forestry university was used in this experiment. For hormone treatments, the EC was cultivated in liquid MS with 20 g/L sucrose and 50 mg/L MeJA, ABA and theirs endogenous inhibitor (Salicylhydroxamic acid, SHAM, Sodium Tungstate Dihydrate, STD) for 24h. The EC, ICpEC (incomplete compact pro-embryogenic cultures), GE (globular embryos) were induced in solid MS with 1.0 mg/L, 0.5mg/L and 0.1mg/L 2,4-D, respectively. The NEC (non-embryonic callus) was induced from longan mature embryo in solid MS medium.

Identification the longan *CCCH* Znf family

The longan protein-coding DNA sequences and protein sequences were download from GigaScience Database (http://dx.doi.org/10.5524/100276). The annotation files of InterPro software of all longan genes downloaded in longan genome database. The 68 genes corresponding to IPR000571 (Zinc finger, CCCH-type) was obtained. Then, all obtained genes was further verified by NCBI BLASTP (Basic Local

Alignment Search Tool Protein) and CD Search (Conserved Domain Search Service). So far, 49 *DlC3Hs* were identified from longan genome database, and named from *DlC3H01* to *DlC3H49* according to their position on pseudo molecules.

Gene sequence characteristics analyze

The ExPasy website (https://web.expasy.org/protparam/) was used to identify the length of sequence, molecular weight and isoelectric points of longan CCCH protein members. The local BLAST was performed to prediction homologous gene of longan *CCCH* genes in *Arabidopsis* genome database. In addition, subcellular location of longan CCCH protein members was predicted by LocTrees3 (https://rostlab.org/services/loctree3/). The number of exon and CCCH Znf domain were obtain from longan genome database and NCBI conserved domain search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). The data of Alternative splicing（AS） events of 49 *DlC3Hs* were extracted from longan non-embryonic callus, embryonic callus, incomplete compact pro-embryogenic cultures and globular embryo transcriptome（SRA050205）.

Phylogenetic analysis and multiple alignment of CCCH domain

The CCCH protein sequences of *Arabidopsis* were downloaded from PlantTF database (http://planttfdb.cbi.pku.edu.cn/). All acquired protein sequences were aligned and constructed ML (maximum likelihood) phylogenetic tree by BioEdit software with default parameters and 1000 bootstrap. The conserved domain amino sequence of longan CCCH members and selected *Arabidopsis* member were aligned by GeneDoc.

The *cis*-elements analysis of *DlC3Hs*' promoters

The up-stream sequences (2K) of *DlC3Hs* CDS (coding sequences) were obtained by TBtools Gtf/Gff3 sequence extractor. Then we deleted the base N found in the promoter of *DlC3H01*, *DlC3H12*, *DlC3H28* and *DlC3H30*. Next, the sequences were submitted to PlantCARE database to predict *cis*-elements.

RNA extraction and expression levels analyses of *DlC3Hs*

Total RNA was extracted by Trzol Reagent kit according to the protocol. The cDNA for quantitative PCR was synthesized by using PrimerScript RT Reagen Kit (Takara). Quantitative PCR was preformed with Roche LightCyclers 480 instrument using SYBR Prumix EX Taq^{TM} Ⅱ (Takara). The 20 μL qPCR reaction was carried out containing 10 μL SYBR Prumix EX Taq^{TM} Ⅱ, ddH_2O 6.4 μL, 1 μL each primer (10 μM), cDNA 2 μL. To acquire reliable results, three biological repeats and three technical repeats were preformed. The reference genes *FSD*, *EF-1α* and *EIF-4α*[27, 28] were used as the internal control. We obtained the relative expression

of *DlC3Hs* according to the $2^{-\Delta Ct}$ method, and results were shown as mean and standard deviation (SD). All the primers used in this study were listed in additional file 1.

Small RNA cleaved verification of a part of *DlC3Hs*

Small RNAs can regulate gene expression by directing the cleavage of target transcript.To understand whether the *DlC3Hs* were regulated by sRNA in longan,17 *DlC3Hs* which highly expressed in a special stage were choosed to verified the cleavage site. The mixture of longan EC, ICpEC and GE total RNA was used to synthesized the cDNA for modified RLM-RACE followed the GeneRacer^{TM} Kit instruction. Using DNAMAN, two gene special primers were designed for modified RLM-RACE. Then, the potential cleavage sites of a part of *DlC3Hs* were predicted by psRNAtarget software against longan sRNA database[29] with default parameters and a maximum expectation value of 3.5 (except *DlC3H36*). All the primers used in this study were listed in additional file 1.

Abbreviations

C3H-Znf: CCCH Zinc finger transcription factors; Longan: *Dimocarpus longan* Lour.; *DlC3H:* longan CCCH Zinc finger gene; AS: alternative splicing events; RT-PCR: real-time PCR; qPCR: quantitative real-time PCR; MeJA: methyl jasmonate; SHAM: Salicylhydroxamic acid.; STD: Sodium Tungstate Dihydrate; ABA: abscisic acid; EC: embryogenic callus.; ICpEC: incomplete compact pro-embryogenic cultures; GE: globular embryos; NEC: non-embryonic callus.

Declarations

Ethics approval and consent to participate

Experimental materials provided by the Institute of Horticultural Biotechnology, Fujian Agriculture and Forestry University, Fuzhou, Fujian, China. The 'HHZ' cultivar used in this study were planted and grown in Fujian Agriculture and Forestry University, Fujian Province, China. No specific permits were required for plant collection. The study did not require ethical approval or consent as no endangered or protected plant species were involved.

Consent for publication

Not applicable.

Availability of data and materials

All data presented in this study are provided either in the manuscript or additional files.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was funded for the National Natural Science Foundation of China (31672127, 31572088), the Scientific Research Foundation of Horticulture College of Fujian Agriculture and Forestry University (2018S02) and the Construction of Plateau Discipline of Fujian Province (102/71201801101) .The funding body did not play a role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Authors’ contributions

ZXL and YLL designed and coordinated the research, and helped to draft the manuscript. YLS participated in its design, carried out the experimental work and wrote the manuscript. MQJ helped to draft the manuscript. SQH, XDX and XL prepared the materials. YLL revised the paper. All authors read and approved the final version of the manuscript. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Ackownledgments

The authors thank to all the research workers at the Institute of Horticultural Biotechnology for their contribution to our study and appreciate those contributors who make related genome and transcriptome datasets accessible in public databases.

Authors' information

1Institute of Horticultural Biotechnology, Fujian Agriculture and Forestry University, Fuzhou 350002, Fujian, China

Additional Files

Additional file 1 :The primers used in this study.

Additional file 2 :The potential sRNAs cleave DlC3H01.

Additional file 3 :The potential sRNAs cleave DlC3H04.

Additional file 4 :The potential sRNAs cleave DlC3H05.

Additional file 5 :The potential sRNAs cleave DlC3H11.

Additional file 6 :The potential sRNAs cleave DlC3H19.

Additional file 7 :The potential sRNAs cleave DlC3H36.

References

1. Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. 2000; 290,5499:2105-10.
2. Wray GA, Hahn MW, Ehab A, Balhoff JP, Margaret P, Rockman MV, Romano LA. The evolution of transcriptional regulation in eukaryotes. Mol Biol Evol.2003;20,9:1377-1491.
3. Kosarev P, Mayer KF, Hardtke CS. Evaluation and classification of RING-finger domains encoded by the Arabidopsis genome. Genome Biol. 2002, 3,4:1-12.
4. Arnaud D, Dejardin A , Leple JC, Lesage-Descauses MC, Pilate G. Genome-wide analysis of LIM gene family in *Populus trichocarpa*, *Arabidopsis thaliana*, and *Oryza sativa*. Dna Res. 2007; 14,3,103-16.
5. Bekir Ülker, Imr ES. WRKY transcription factors: from DNA binding towards biological function. Cerr Opin Plant Biol. 2004; 7,5,491-8.
6. Shuichi Y,Dof domain proteins: plant-specific transcription factors associated with diverse phenomena unique to plants. Plant Cell Physiol. 2004; 45,4,386-91.
7. Carine B, Luc P, H.Beverley O. AU-rich elements and associated factors: are there unifying principles?. 2005;33,22,7138—50.
8. Tala B, Williams BRG, Khabar KSA. ARED 3.0: the large and diverse AU-rich transcriptome. Nucleic Res.2006; 34,D111-4.
9. Wang D, Guo Y, Wu C, Yang G, Li Y, Zheng C: Genome-wide analysis of CCCH zinc finger family in Arabidopsis and rice. BMC Genomics. 2008; 9:44.
10. Zhang C, Zhang H, Zhao Y, Jiang H, Zhu S, Cheng B, Xiang Y. Genome-wide analysis of the CCCH zinc finger gene family in Medicago truncatula. Plant Cell Rep. 2013; 32,10,1543-55.
11. Chai G. Comprehensive analysis of CCCH zinc finger family in poplar (Populus trichocarpa). BMC Genomics. 2012; 13:253.
12. Liu S, Khan MRG, Li Y, Zhang J, Hu C. Comprehensive analysis of CCCH-type zinc finger gene family in citrus (Clementine mandarin) by genome-wide characterization. Mol Genet Genomics. 2014; 289,5,855-72.
13. Wang XL, Zhong Y, Cheng ZM. Evolution and Expression Analysis of the CCCH Zinc Finger Gene Family in Vitis vinifera. Plant Genome-US. 2014; 7,3,16-25.
14. Pradhan S, Kant C, Verma S, Bhatia S. Genome-wide analysis of the CCCH zinc finger family identifies tissue specific and stress responsive candidates in chickpea (*Cicer arietinum L.*). PLOS One. 2017; 12:7.
15. Mazumdar P, Lau SE, Wei YW, Singh P, Harikrishna JA. Genome-wide Analysis of the CCCH Zinc-Finger Gene Family in Banana (*Musa acuminata*): An Insight Into Motif and Gene Structure Arrangement, Evolution and Salt Stress Responses. Trop Plant Biol. 2017;8,1-17.
16. Peng XJ , Zhao Y, Cao JG , Zhang W , Jiang HY , Li XY , Ma Q , Zhu SW , Cheng BJ. CCCH-type zinc finger family in maize: genome-wide identification, classification and expression profiling under abscisic acid and drought treatments. PLOS One.2012; 7:7.

17. Sun JQ, Jiang HL, Xu YX, Li HM, Wu XY, Xie Q, Li Chuanyou. The CCCH-type zinc finger proteins AtSZF1 and AtSZF2 regulate salt stress responses in Arabidopsis. Plant Cell Physiol. 2007; 48,8,1148-58.
18. Guo YH, Yu YP, Wang D, Wu CA, Yang GD, Huang JG,Zheng CC. GhZFP1, a novel CCCH-type zinc finger protein from cotton, enhances salt stress tolerance and fungal disease resistance in transgenic tobacco by interacting with GZIRD21A and GZIPR5. New Phutol. 2009; 183,1,62-75.
19. Kong ZS, Li MN, Yang WQ , Xu WY , Xue YB. A novel nuclear-localized CCCH-type zinc finger protein, OsDOS, is involved in delaying leaf senescence in rice (*Oryza sativa L.*). Plant Physiol.2006; 14,1376-88.
20. Li JJ , Jia DX ,Chen XM. HUA1, a regulator of stamen and carpel identities in Arabidopsis, codes for a nuclear RNA binding protein. Plant Cell. 2001; 13,10,2269-81.
21. Li ZS, Terry LT. PEI1, an embryo-specific zinc finger protein gene required for heart-stage embryo formation in Arabidopsis. Plant Cell. 1998;10,3,383-98.
22. Agnieszka G, Anita W, Norikazu T, Stefan M, Marcin F. Characterization of CsSEF1 gene encoding putative CCCH-type zinc finger protein expressed during cucumber somatic embryogenesis. J Plant Physiol. 2009; 166,3,310-23.
23. Lin YL, Min JM, Lai RL, Wu ZY , Chen YK, Yu LL, Cheng CZ, Jin YC, Tian QL , Liu Q.Genome-wide sequencing of longan (*Dimocarpus longan Lour.*) provides insights into molecular basis of its polyphenol-rich characteristics. GigaScience. 2017; 6,5,1-14.
24. Jue DW, Sang XL, Liu LQ, Shu B, Wang YC , Liu CM, Xie JH,Shi SY. Identification of WRKY Gene Family from Dimocarpus longan and Its Expression Analysis during Flower Induction and Abiotic Stress Responses. Int J Mol Sci. 2018;19:8.
25. Jue DW, Sang XL, Liu LQ, Shu B, Wang YC , Liu CM, Xie JH,Shi SY. The Ubiquitin-Conjugating Enzyme Gene Family in Longan (*Dimocarpus longan Lour.*): Genome-Wide Identification and Gene Expression during Flower Induction and Abiotic Stress Responses. Molecules. 2018; 23:326.
26. Xu XP, Chen XH, Lv KL, Chen X, Chen YK, Lin YL, Lai ZX. Genome-wide identification and function analysis of the Laccase gene family in Dimocarpus longan Lou. Chinese Journal of Applied and Envrionmental. 2018; 24,04,833-44.
27. Lin YL, Lai ZX. Reference gene selection for qPCR analysis during somatic embryogenesis in longan tree. Plant Sci. 2010; 178,4,359-65.
28. Lin YL, Lai ZX. Evaluation of suitable reference genes for normalization of microRNA expression by real-time reverse transcription PCR analysis during longan somatic embryogenesis. Plant Physiol Bioch. 2013; 66,20-25.
29. Lin YL, Lai ZX. Comparative Analysis Reveals Dynamic Changes in miRNAs and Their Targets and Expression during Somatic Embryogenesis in Longan (*Dimocarpus longan Lour.*). PLOS One. 2013; 8:4.
30. Bartel DP. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell. 2004;16,2,281-97.

31. Li YQ. Analysis of Biological Functioon of the Phytoplasma-responsive Novel MicroRNAs in Mulberry.Shandong Taian of China,Shandong Agricultural University; 2014.
32. Initiative AG. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature. 2000; 408,6814,796-815.
33. Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, Wu Z, Li M, Zheng T, Fuentes RR, Zhang F. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. Nature. 2018; 557,43-49.
34. Xu Q, Chen LL, Ruan XA, Chen DJ, Zhu AD, Chen CL, Denis B, Jiao WB, Hao BH, Matthew P L. The draft genome of sweet orange (*Citrus sinensis*). Nat Gene. 2013;45,1,59-92.
35. The FIPC, Jaillon O, Aury J, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;449,7161,463-7.
36. Schnable PS, Doreen W, Fulton RS, Stein JC, Fusheng W, Shiran P, Chengzhi L, Jianwei Z, Lucinda F, Graves TA. The B73 maize genome: complexity, diversity, and dynamics. Science. 2015;326,5956,1112-5.
37. Antoine B, Paape TD, Peng Z, Roman B, Farmer AD, Joann M, Bharti AK, Woodward JE, May GD, Laurent G. Whole-genome nucleotide diversity, recombination, and linkage disequilibrium in the model legume Medicago truncatula. Proc Natl Acad Sci-U S A. 2011, 108,42,864-70.
38. Mosavi LK, Minor DL, Peng ZY . Consensus-derived structural determinants of the ankyrin repeat motif. Proc Natl Acad Sci-U S A. 2002; 99,25,16029-34.
39. Cristian B, Torben J, Pere P, Vicient CM. Ankyrin repeat-containing proteins in Arabidopsis: characterization of a novel and abundant group of genes coding ankyrin-transmembrane proteins. Gene. 2004;340,1,111-21.
40. Lorkovi ZJ, Andrea B. Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant Arabidopsis thaliana. Nucleic Acids Res. 2002;30,3,623-635.
41. Couture JF, Collazo E, Trievel RC: Molecular recognition of histone H3 by the WD40 protein WDR5. Nat Struct Mol Biol.2006; 13,8,698–703.
42. Nocker SV, Ludwig P. The WD-repeat protein superfamily in Arabidopsis: conservation and divergence in structure and function. BMC Genomics. 2003; 4:50.
43. Marcelo P, Lin PC, John F, Jang JC. AtTZF gene family localizes to cytoplasmic foci. Plant Signal Behav. 2010;5,2,190-2.
44. Lin YL, Lai ZX, Lin LX, Lai RL, Tian QL, Ye W,Zhang DM, Yang MM, Chen YK, Zhang ZH. Endogenous target mimics, microRNA167, and its targets ARF6 and ARF8 during somatic embryo development in Dimocarpus longan Lour. Mol Breeding. 2015;35:227.
45. Lin YL, Lai ZX, Tian QL, Lin LX, Lai RL, Yang MM, Zhang DM, Chen YK, Zhang ZH. Endogenous target mimics down-regulate miR160 mediation ofARF10,-16, and-17cleavage during somatic embryogenesis inDimocarpus longanLour. Front PlantL Sci. 2015, 6:956.

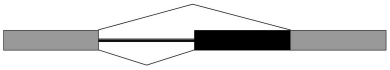
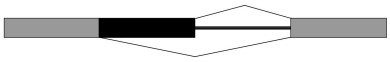
46. 46. Lin YL, Lin LX, Lai RL, Liu WH, Chen YK, Zhang ZH, Xu XH, Lai Z MicroRNA390-Directed TAS3 Cleavage Leads to the Production of tasiRNA-ARF3/4 During Somatic Embryogenesis in Dimocarpus longanLour. Front Plant Sci. 2015, 6:127.

Tables

Table 1 The characteristics of longan ZF_C3H gene family

Gene name	Locus name	*Arabidopsis* orthologs locus	Exons	Protein			Number of CCCH motif	Subcellular location
				Length (aa)	Wt (kD)	PI		
DlC3H01	Dlo_001632.2	AT3G02830.1	8	440	48.10	8.86	5	nucleus
DlC3H02	Dlo_001685.1	/	2	256	29.32	7.99	1	cytoplasm
DlC3H03	Dlo_002620.1	AT5G12440.1	7	639	70.50	5.85	1	nucleus
DlC3H04	Dlo_003675.1	/	7	183	20.35	8.64	1	secreted
DlC3H05	Dlo_005201.1	AT2G41900.1	1	739	80.41	5.73	1	nucleus
DlC3H06	Dlo_006391.1	AT5G12850.1	1	730	79.29	6.10	1	nucleus
DlC3H07	Dlo_007322.1	AT5G51980.1	9	405	44.32	8.12	1	cytoplasm
DlC3H08	Dlo_008757.1	/	3	274	30.81	7.12	1	nucleus
DlC3H09	Dlo_009034.1	AT3G12130.1	3	293	30.64	9.36	3	secreted
DlC3H10	Dlo_009361.1	AT2G29580.1	4	499	56.30	8.15	1	nucleus
DlC3H11	Dlo_010735.1	AT2G32930.1/AT2G32930.2	7	448	48.40	7.19	5	nucleus
DlC3H12	Dlo_011815.1	AT2G10553.1 AT3G18640.1	3	949	106.69	8.62	2	nucleus
DlC3H13	Dlo_012862.1	AT3G21100.1	7	549	62.19	6.66	1	nucleus
DlC3H14	Dlo_013962.1	AT1G68200.2	2	343	37.99	6.99	1	cytoplasm
DlC3H15	Dlo_014492.1	AT2G02160.1	3	721	79.80	5.30	2	nucleus
DlC3H16	Dlo_015528.1	AT2G24830.1	4	504	56.79	5.22	1	nucleus
DlC3H17	Dlo_015573.1	AT2G29580.1	5	413	46.12	8.26	1	nucleus
DlC3H18	Dlo_016726.1	AT1G32360.1	2	379	40.01	7.12	3	secreted
DlC3H19	Dlo_017697.1	AT2G47850.1/AT2G47850.2/AT2G47850.3	7	473	50.84	9.17	5	nucleus
DlC3H20	Dlo_017983.1	AT3G52980.1	7	570	64.02	5.97	1	nucleus
DlC3H21	Dlo_018202.2	/	11	415	46.13	7.23	1	nucleus
DlC3H22	Dlo_018738.1	AT3G27700.1/AT3G27700.2	4	931	101.64	6.52	1	cytoplasm
DlC3H23	Dlo_019035.1	AT4G29190.1	1	388	42.08	7.86	1	nucleus
DlC3H24	Dlo_019408.1	AT3G19360.1	2	333	37.37	8.68	2	secreted
DlC3H25	Dlo_019413.1	AT2G02160.1	3	502	57.28	8.69	1	nucleus
DlC3H26	Dlo_020091.2	AT5G44260.1	2	601	67.27	7.22	1	nucleus
DlC3H27	Dlo_021455.2	AT3G51120.1	10	1811	198.20	5.77	1	nucleus
DlC3H28	Dlo_021549.1	/	2	212	23.66	9.50	1	cytoplasm
DlC3H29	Dlo_021686.1	AT2G20280.1	6	246	27.76	5.95	1	cytoplasm
DlC3H30	Dlo_021827.1	AT3G12130.1	3	293	30.67	9.36	3	secreted
DlC3H31	Dlo_023454.1	AT3G06410.1	7	481	50.67	8.53	5	nucleus
DlC3H32	Dlo_023553.1	AT2G47850.1/AT2G47850.2/AT2G47850.3	8	479	51.54	9.18	5	nucleus
DlC3H33	Dlo_024382.1	AT2G16485.2	10	1568	171.01	4.90	1	nucleus
DlC3H34	Dlo_026594.1	AT2G28450.1/AT2G28450.2	14	819	90.12	5.63	1	nucleus
DlC3H35	Dlo_026957.3	AT5G63260.1/AT5G63260.2	7	367	41.44	7.54	4	nucleus
DlC3H36	Dlo_028391.1	AT2G40140.1/AT2G40140.2	1	604	65.56	8.07	1	nucleus
DlC3H37	Dlo_028442.1	AT4G29190.1	1	366	40.62	7.04	2	nucleus
DlC3H38	Dlo_028626.1	AT3G51950.1	8	743	80.26	6.15	1	nucleus
DlC3H39	Dlo_028726.1	AT5G12850.1	1	721	78.44	6.38	2	nucleus
DlC3H40	Dlo_028971.1	AT2G20280.1	7	364	41.53	5.26	1	cytoplasm
DlC3H41	Dlo_029035.1	/	2	136	14.46	9.21	2	cytoplasm
DlC3H42	Dlo_029077.1	AT3G12680.1	12	494	54.08	6.99	6	nucleus
DlC3H43	Dlo_029354.1	AT5G58620.1	1	708	76.56	5.69	1	nucleus
DlC3H44	Dlo_029383.1	AT3G47120.1	4	357	41.75	8.30	1	nucleus
DlC3H45	Dlo_029440.1	/	4	498	55.98	7.58	1	cytoplasm
DlC3H46	Dlo_029632.1	AT1G19860.2	4	471	50.49	6.73	1	nucleus
DlC3H47	Dlo_030292.1	AT5G56900.1	10	598	66.35	7.81	1	nucleus
DlC3H48	Dlo_033257.1	AT5G07500.1	1	298	33.74	6.59	1	nucleus
DlC3H49	dlo_035320.1	AT5G26749.1	5	158	17.54	7.66	1	nucleus

Table 2 Classification of AS events in longan somatic embryo

Type	Structure	Events	Genes
A3’S		84 (26.17%)	22 (30.14%)
A5’S		59 (18.30%)	18 (24.66%)
IR		145(45.17%)	21 (28.77%)
ES		33 (10.28%)	12 (16.44%)
Total		321	73

Figures

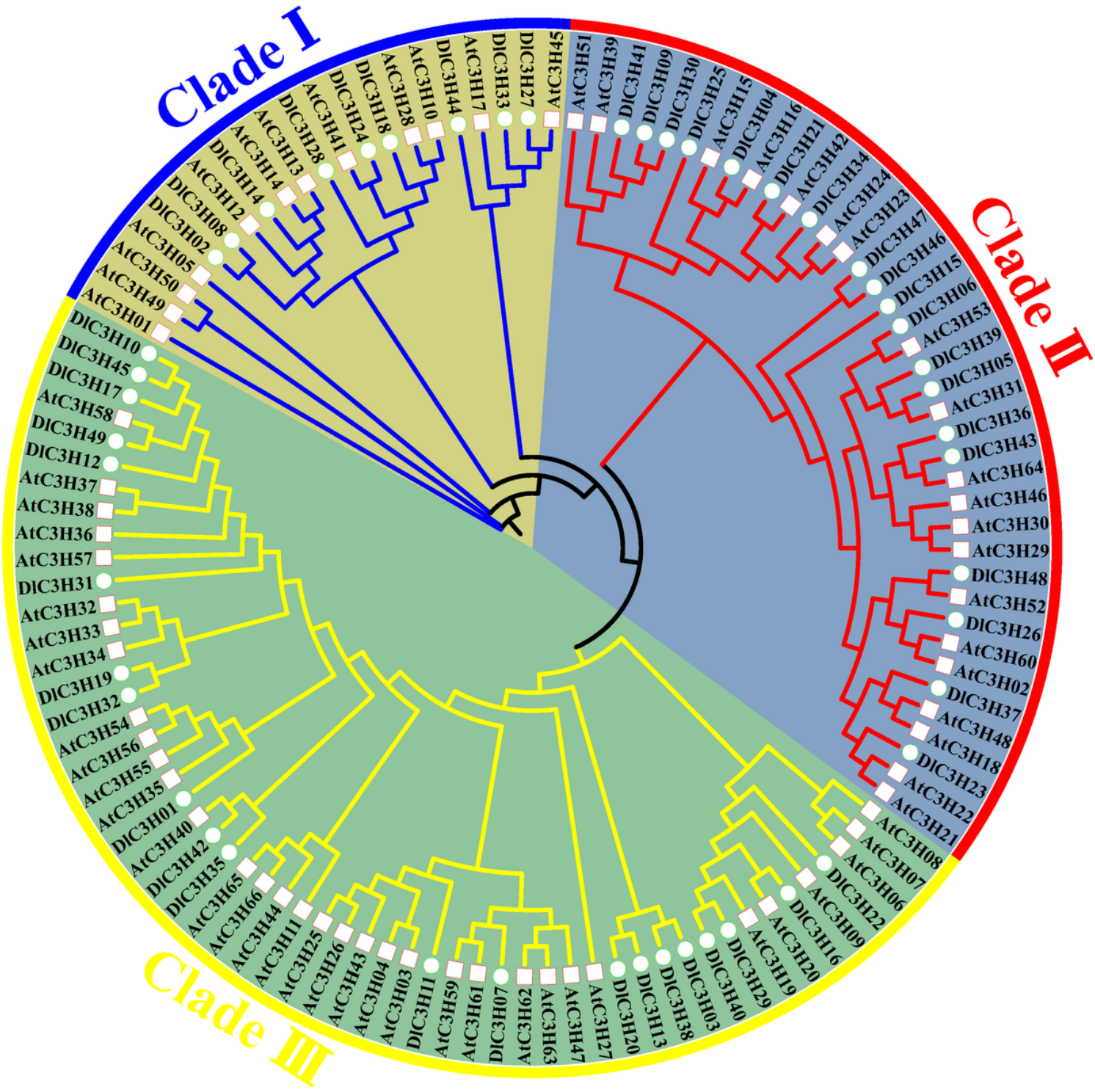


Figure 1

The phylogenetic tree of C3H genes from longan and Arabidopsis.

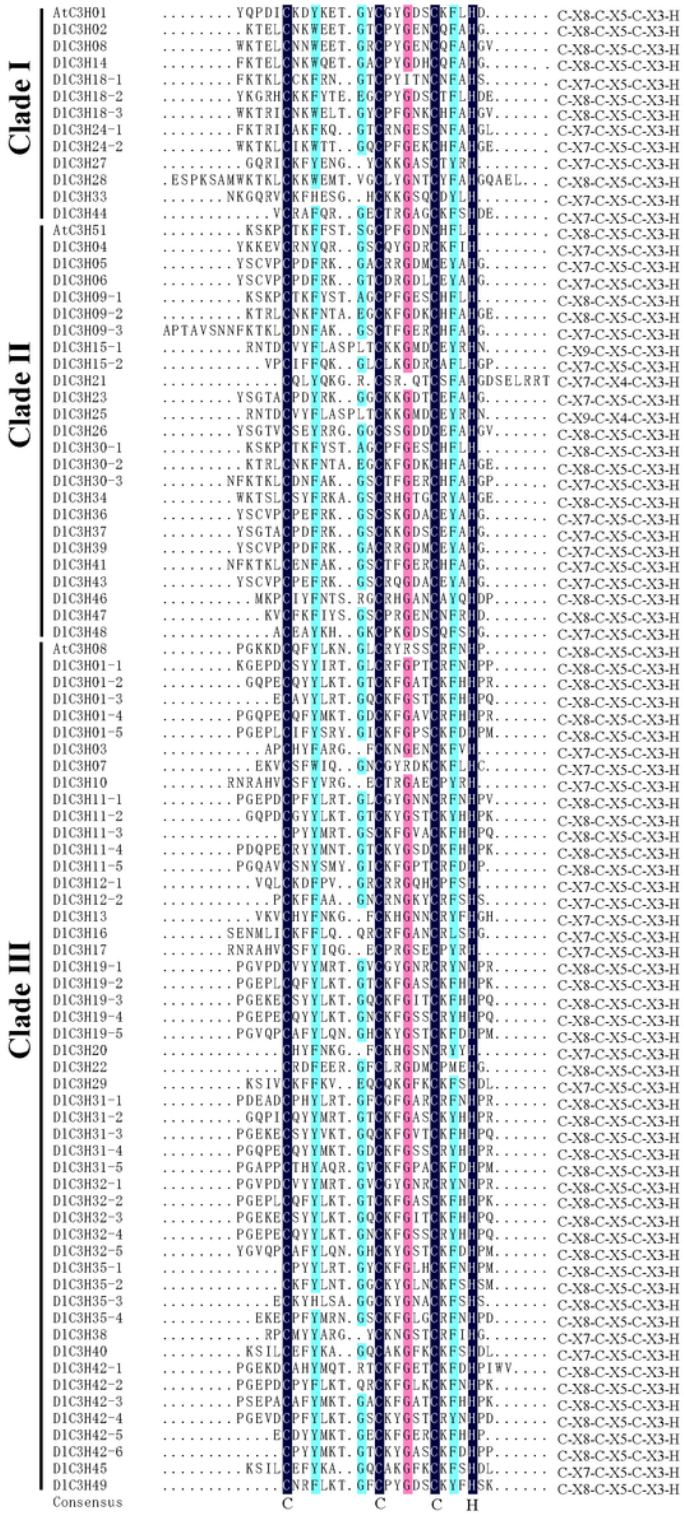


Figure 2

The multiple alignment of DIC3Hs and selected AtC3Hs ZF-CCCH domain amino acid sequences.

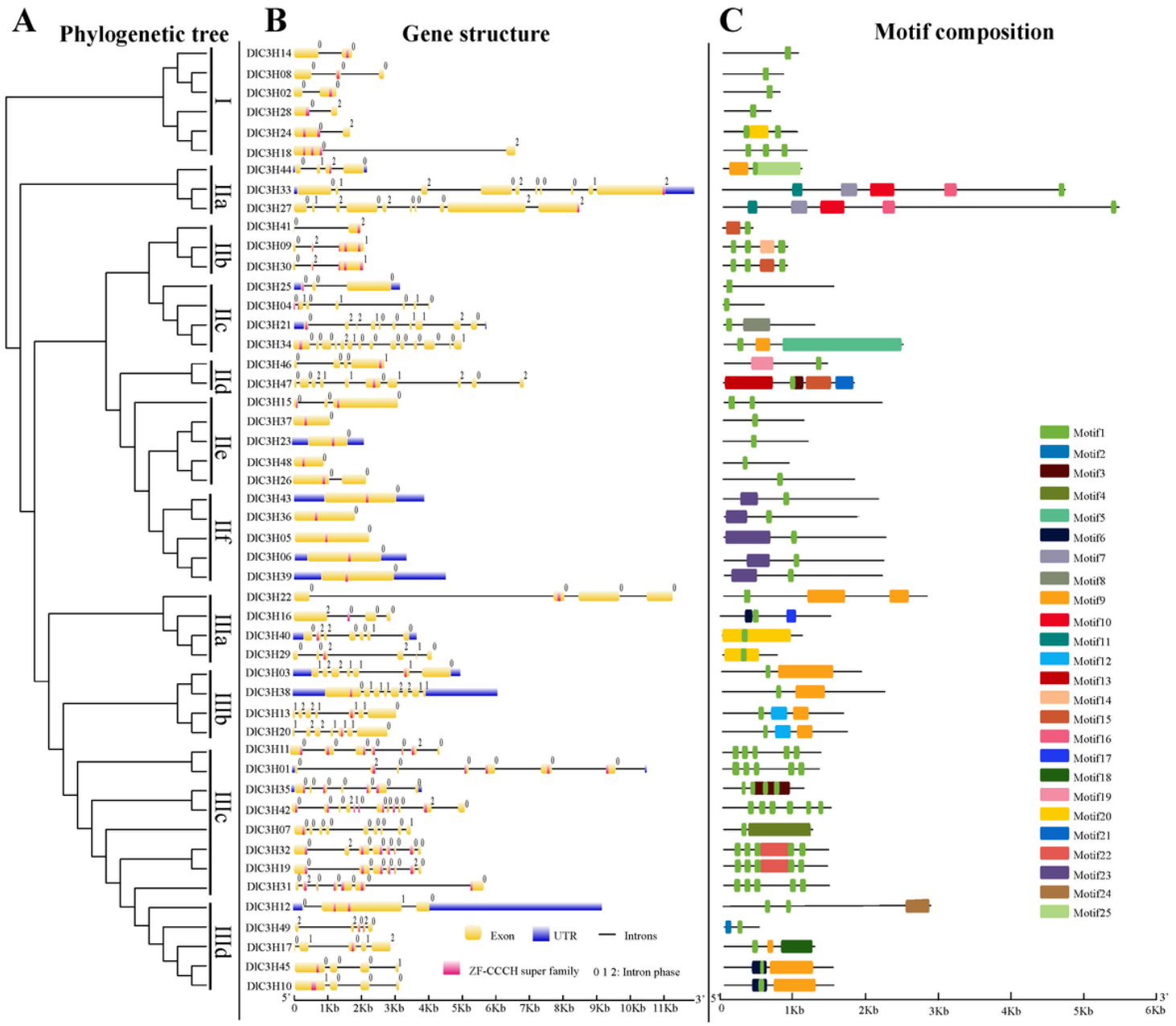


Figure 3

The evolution relationship, gene structure and conserve motif in longan C3H gene family. A The phylogenetic tree of DIC3Hs. B Gene structure of longan C3H genes. Yellow box indicate the coding sequence; Blue box is the untranslation 3’ and 5’ region. Black line represent intron; Red box is the ZF-CCCH conserved domain. The numbers of 0, 1, 2 were the phase of corresponding introns. The sequence length can be inferred by bottom scale. C Distribution of conserved domain in longan C3H genes. The 25 motifs are display with different color and the length of protein can be estimated by bottom scale. Motif1: ZF-CCCH super family; Motif 2: zf-U1; Motif3: YTH1 super family; Motif4: WD40 super family; Motif5: TrmA; Motif6: Torus super family; Motif7: SWIB; Motif8: SMC_N super family; Motif9: RRM super family; Motif10: Plus3; Motif11: PHD5_NSD; Motif12: OST-HTH; Motif13: MPP_CWF19_N; Motif14: KH_1 super family; Motif15: HIT_like; Motif16: GYF; Motif17: G-patch; Motif18: FtsK super family; Motif19: DNA_pol3_delta2

super family; Motif20: DFRP_C; Motif21: CwfJ_C_2; Motif22: Atrophin-1 super family; Motif23: ANK; Motif24: Amino_oxidase super family; Motif25: PRK12678 super family

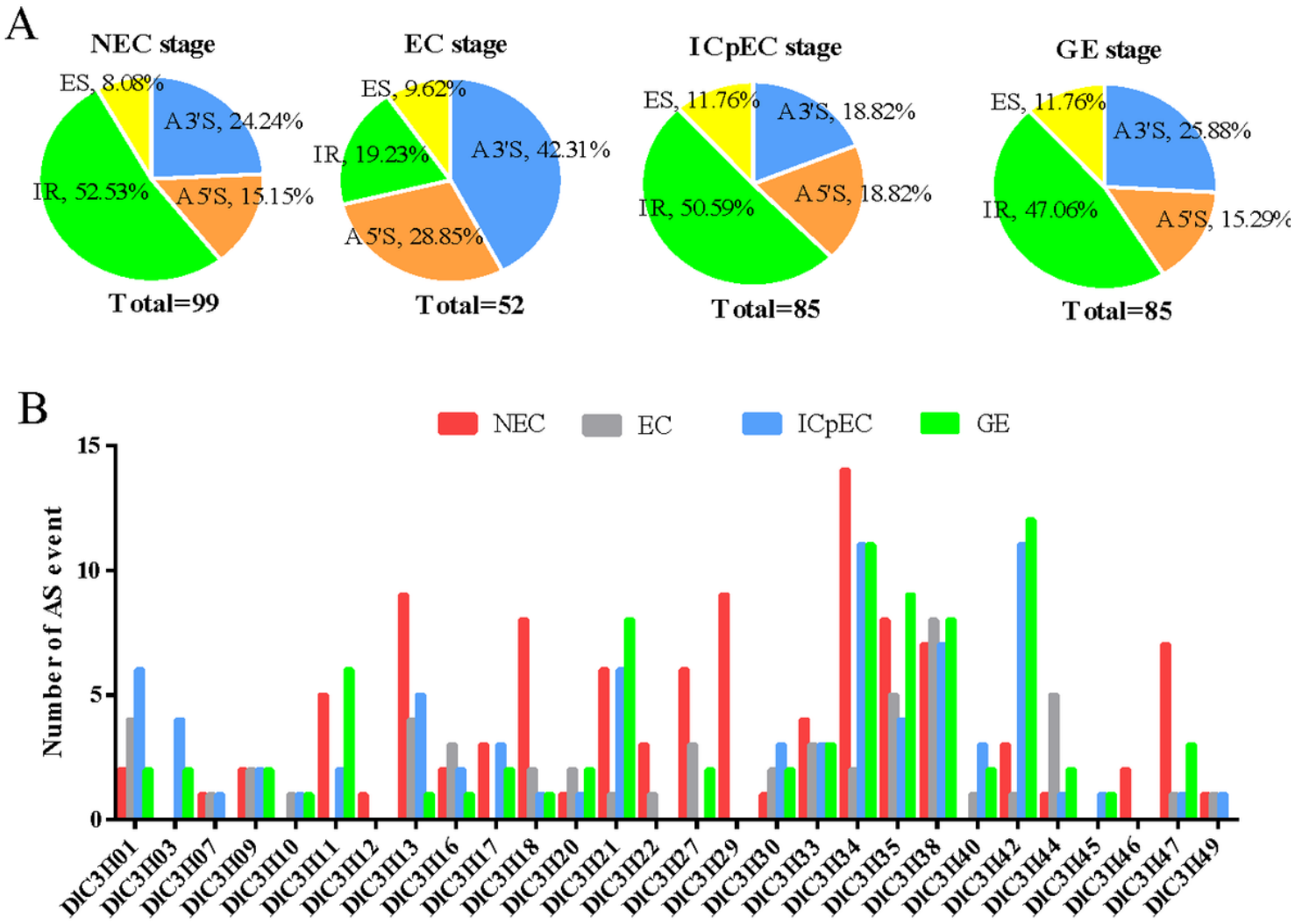


Figure 4

The AS event distribution of DIC3H genes. A The percentage of four AS events find in NEC, EC, ICpEC and GE stages of longan somatic. A3'S: alternative 3' splice site; A5'S: alternative 5' splice site; IR: intron retaintion; ES: exon skipping. B The number distribution of AS event in DIC3Hs in four stages of longan somatic embryo.

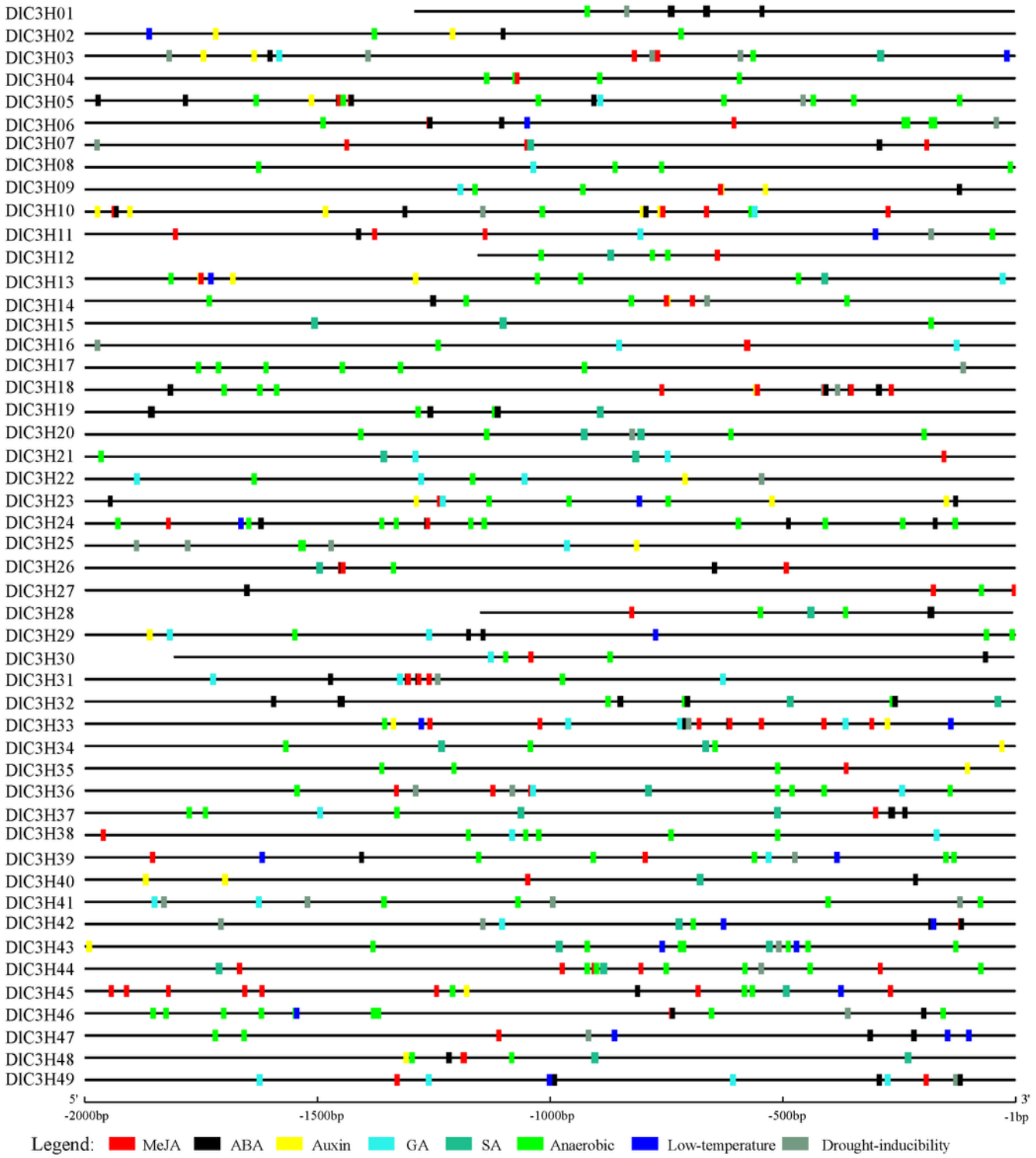


Figure 5

The potential cis-elements of DlC3H promoters. The promoter sequences (up-stream 2000Kb) of 49 DlC3H genes are analyzed by PlantCARE database. A part of promoter region of DlC3H01, DlC3H12, DlC3H28 and DlC3H30 are absent. The length of sequences can be inferred by the bottom scale.

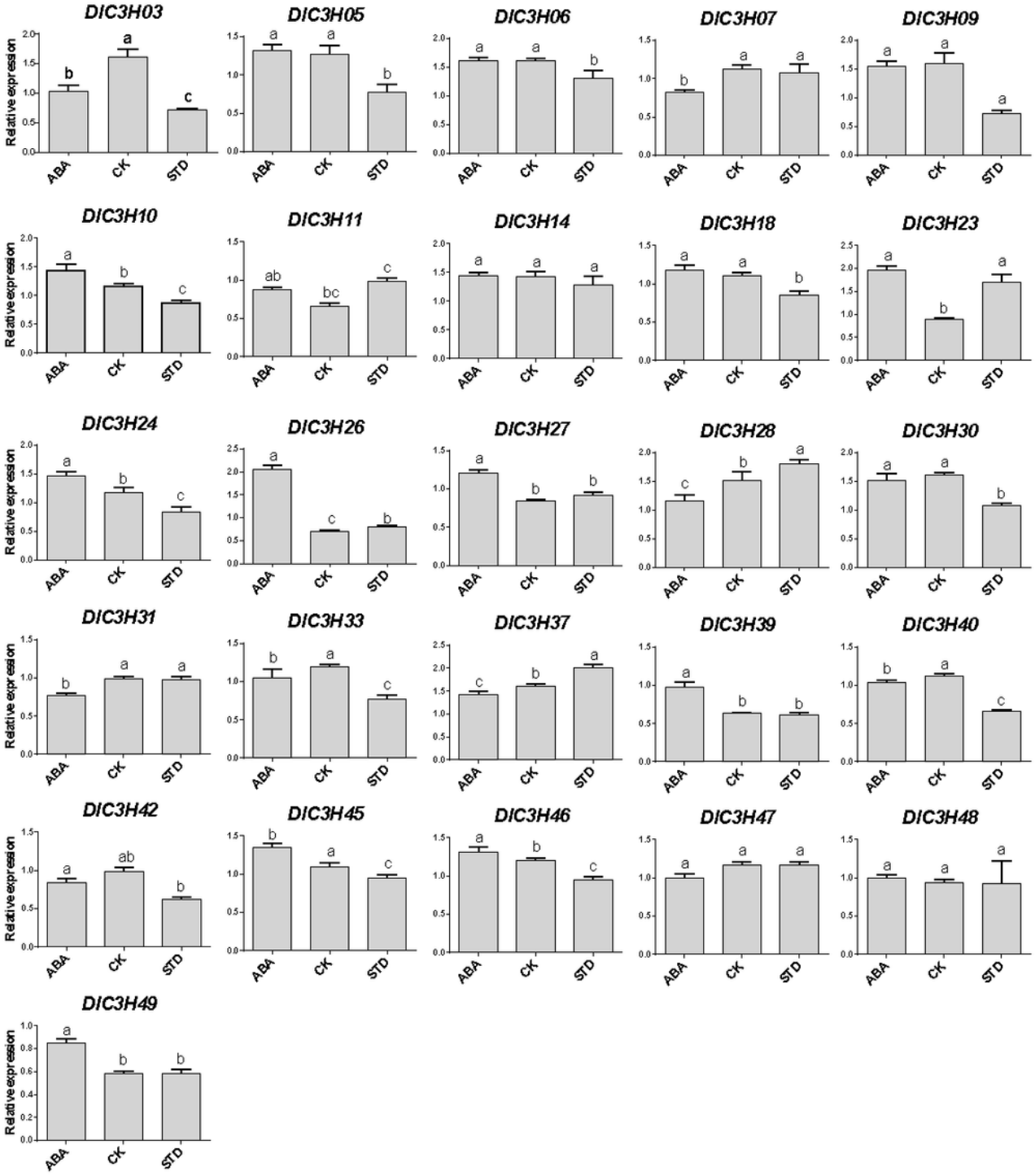


Figure 6

Expression levels of 26 selected DIC3Hs after identical concentration of ABA and STD treatment.

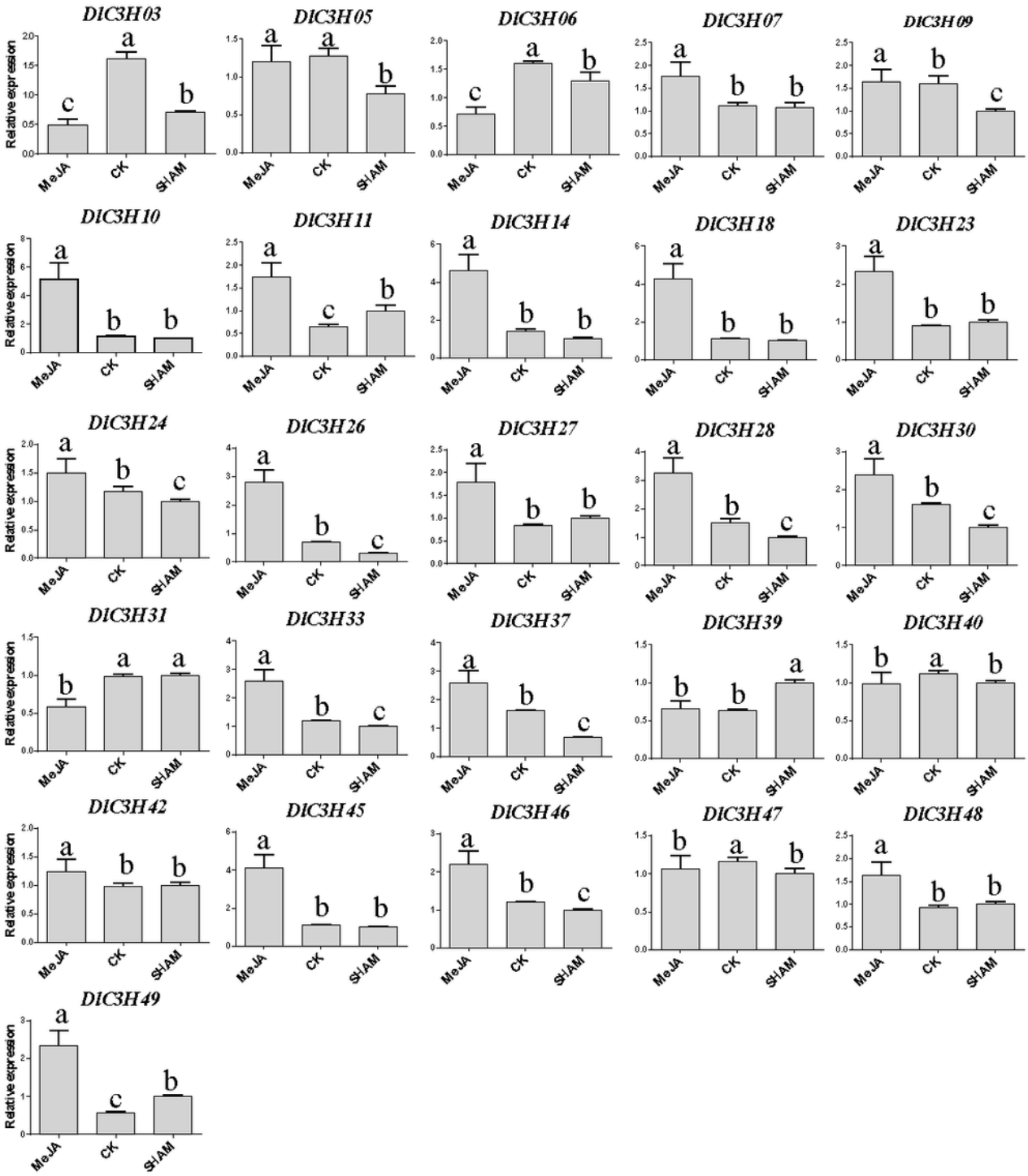


Figure 7

Expression levels of 26 selected DlC3Hs after identical concentration of MeJA and SHAM treatment.

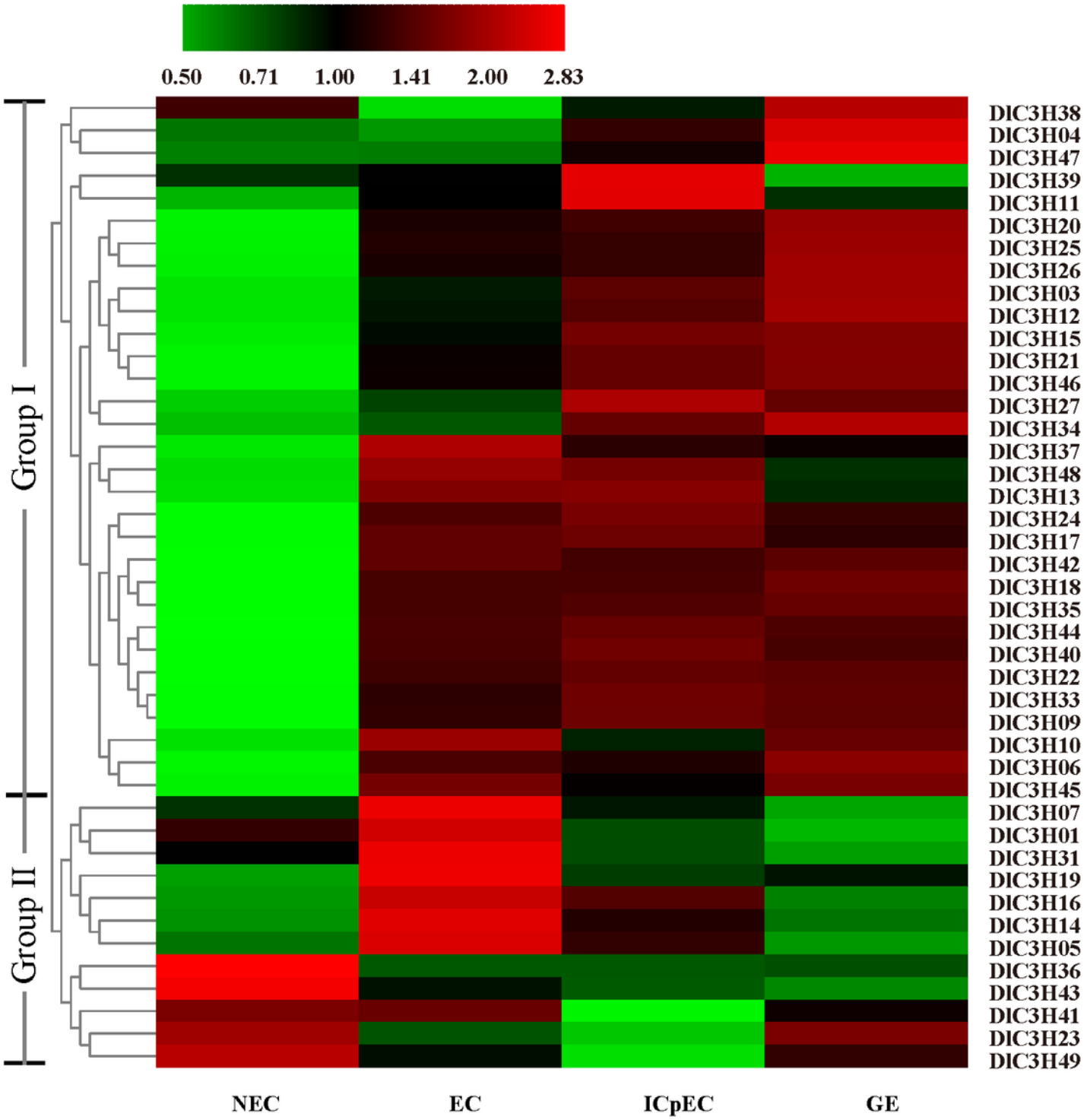


Figure 8

Expression profile of DlC3Hs in different stages of longan somatic embryo.

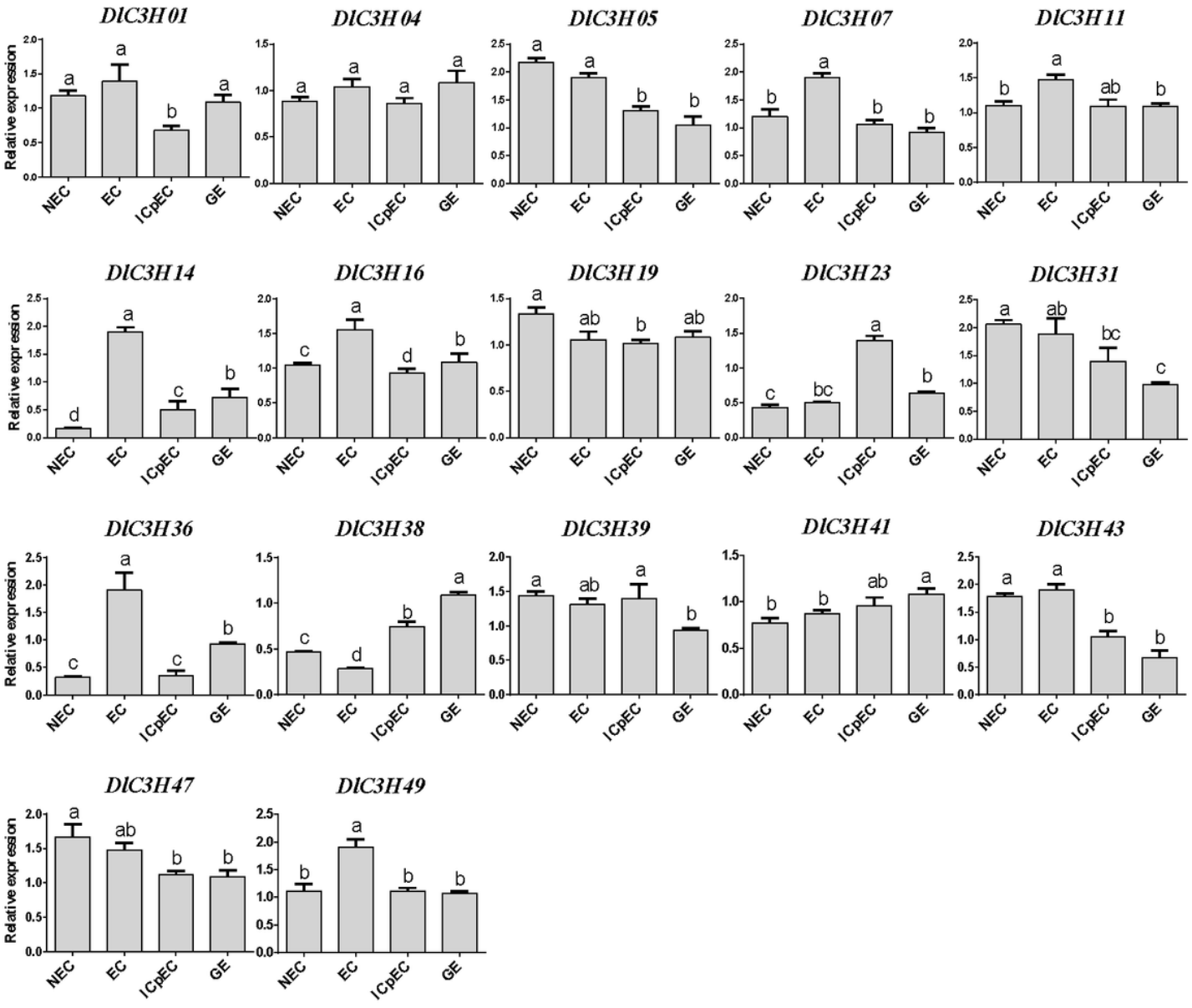


Figure 9

Expression patterns of 17 DlC3Hs in longan early somatic embryo. NEC: non-embryonic callus; EC: embryonic callus; ICpEC: incomplete compact pro-embryogenic cultures; GE: globular embryos

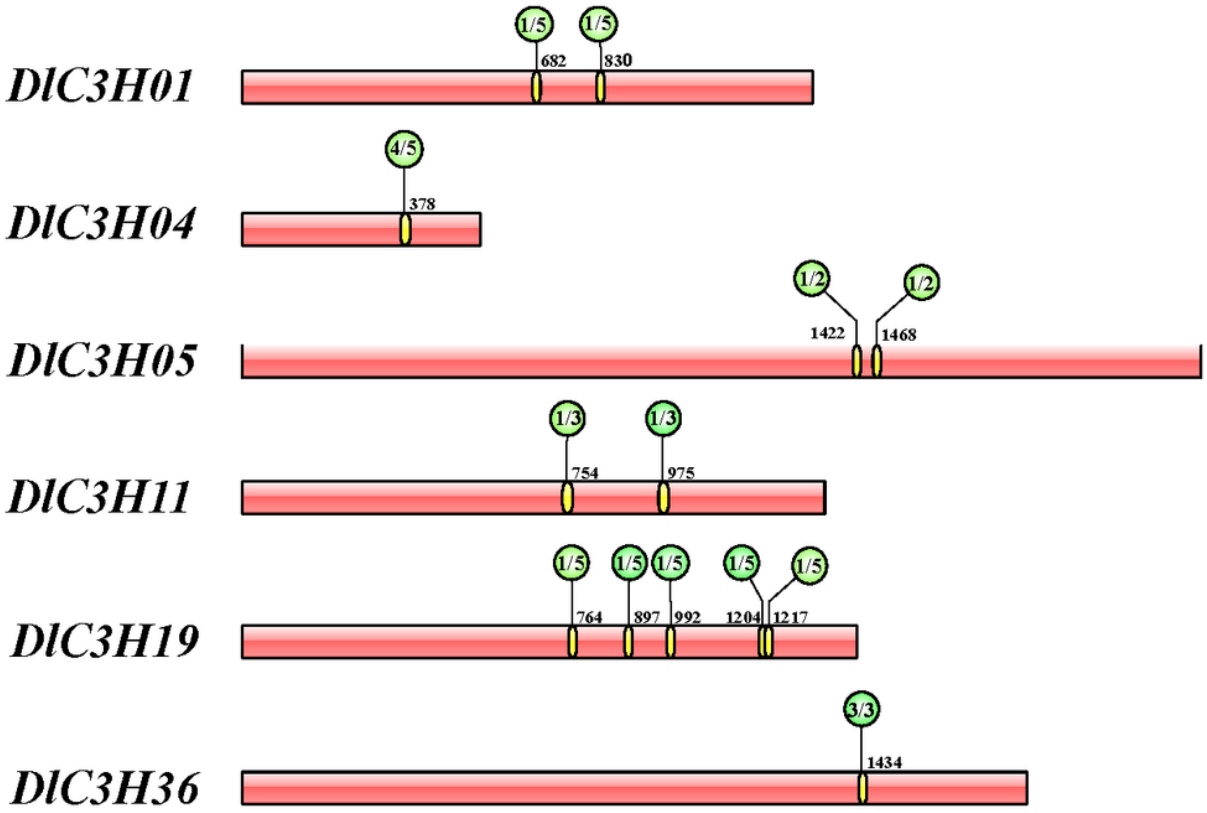


Figure 10

Analysis of the sRNAs cleavage site in longan C3Hs.The yellow area is the sRNA binding site,the number above is the cleavage position;the number in the green area is the number of cleavage.

Supplementary Files

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

- Additionalfile2ThepotentialsRNAscleaveDlC3H01.xls
- Additionalfile7ThepotentialsRNAscleaveDlC3H36.xls
- Additionalfile6ThepotentialsRNAscleaveDlC3H19.xls
- Additionalfile3ThepotentialsRNAscleaveDlC3H04.xls
- Additionalfile4ThepotentialsRNAscleaveDlC3H05.xls
- Additionalfile5ThepotentialsRNAscleaveDlC3H11.xls
- Additionalfile1Theprimersusedinthisstudy.xls