Cellular & Molecular Biology
Cell. Mol. Biol. 2015; 61 (8): 112-117
Published online December 26, 2015 (http://www.cellmolbiol.com)
Received on October 5, 2015, Accepted on December 21, 2015.
doi : 10.14715/cmb/2015.61.8.18

Identification and characterization of *chalcone synthase* cDNAs (*NnCHS*) from *Nelumbo nucifera*

C. Dong[1], A. Q. Yu[2], M. L. Wang[2], X. W. Zheng[3], Y. Diao[2], K. Q. Xie[3], M. Q. Zhou[4] and Z. L. Hu[2]✉

[1] College of Biological Engineering, Henan University of Technology, Zhengzhou 450051, China
[2] State Key Laboratory of Hybrid Rice, College of Life Science, Wuhan University, Wuhan 430072, China
[3] Guangchang White Lotus Research Institute, Fuzhou 344900, China
[4] Lotus Center, Wuhan University, Wuhan 430072, China

Corresponding author: Z. L. Hu, State Key Laboratory of Hybrid Rice, Wuhan University, Wuhan 430072, China. Email: huzhongli@whu.edu.cn

Abstract

Chalcone synthase (CHS) catalyzes the first committed step in flavonoids biosynthetic pathway. In this study, six full-length cDNAs (*NnCHS*) encoding CHS from *Nelumbo nucifera* were successfully isolated, using rapid amplification cDNA end (RACE) assay. The obtained cDNAs were 1426 bp in size, containing a 1167 bp open reading frame coding 389 amino acids. Exons-intron architecture of *NnCHS* gene was illustrated, consisting two exons inserted by a 426 bp intron. The putative NnCHS possessed all the conserved active sites for CHS function as well as the family signature. Phylogenetic analysis revealed that NnCHS shared high homology with CHS from high plants, and the homology-based structural modeling showed that NnCHS had the typical structure of CHS. Moreover, Real-time PCR assays demonstrated that *NnCHS* mRNAs were expressed in various tissues of *N. nucifera*, with the highest expression in red flower and lowest level in the leaves. Moreover, patterns of *NnCHS* expression illustrated short-time wounding or low temperature significantly induced the up-regulation of *NnCHS* mRNA.

Key words: *Nelumbo nucifera*, Chalcone synthase, cDNA, Gene expression.

Introduction

The flower color mainly determined by betalains, carotenoids and flavonoids is the major characteristics of ornamental plants (1, 2). Unlike betalains and carotenoids, flavonoids are often present in the plant tissues to create co-pigmentation phenomenon, and their combination results in the color variety (3, 4). Flavonoids are commonly considered as one of the most important secondary metabolites, including chalcones, flavones, flavonols and anthocyanins. Flavonoids have a wide variety of biological roles such as UV protection, resistances against pathogens, auxin transport, signal molecular in plants and microbes interaction, and most importantly, they play a imperative role in flower attracting pollinators and dispersing the fruits and seeds (5-8).

Flavonoids biosynthesis has been reported in flowers, fruits and kernels for many years, because of their roles in biology and agriculture. Briefly, the biosynthesis of flavonoids starts from the condensation of one molecules of malonyl-CoA, resulting in naringenin chalcone which is catalyzed by chalcone synthase (CHS). As the best studied plants –specific type III polyketide synthase, CHS performs the first committed step of the branch of the phenylpropanoid leading to the synthesis of flavonoids (9, 10). Considering its essential role in flavonoids biosynthesis, *CHS* genes have been cloned and researched in hundreds of plant species including *Oryza sativa*, *Medicago sativa* and *Zea mays*. Analysis of amino acid sequence illustrates that CHS superfamily shares high similarity. *CHS* mRNAs are expressed in various tissues of plants (11), and induced by exposure to UV light, wounding, low temperature and pathogen invasion (12). CHS protein usually roles as a homodimer (monomer size about 43 kDa), containing the conserved Cys-His -Asn catalytic triad in the buried active sites (13). Moreover, the crystal structure and functional studies of alfalfa CHS2 have confirmed that the conserved Cys and His residues play a significant role in the catalytic function of CHS, and site-directed mutants of these residues have been investigated to elucidate the reaction mechanism of CHS (14, 15).

Lotus *(Nelumbo nucifera* L.) belongs to Nelumbonaceae, Nymphaeales, and is a perennial aquatic plant. As one of the oldest dicotyledonous plants, *N.nucifera* also obtains several monocotyledon characteristics (16), which is native to tropic and temperate regions of Asia (17). According to history record, *N. nucifera* has been cultivated as ornamental plant and vegetable since the Chow Dynasty in China (18). Moreover, almost the whole tissues of *N. nucifera* can be used in Chinese traditional medicine, such as the seed, root, leaves, flower and embryos. Notably, its seeds can survive for thousands of years, and considered as the oldest viable seeds in the world (19). Up to date, the isolation and characterization of new *CHS* cDNAs from higher plants is on-going. However, as one of the important ornamental plant, *CHS* cDNAs of *N. nucifera* (*NnCHS*) have not been reported.

In this study, the cDNAs and gene sequences of *NnCHS* were isolated, including one intron and two exons. The intron of *NnCHS* started with the nucleotides 'GT', ending with the nucleotides 'AG', following the GT-AG rule. Additionally, the putative sequences of NnCHS proteins were characterized by comparing with other known CHSs, and performed phylogenetic analysis. The 3D structure and active sites of NnCHS were illustrated by homology model. Finally, Real-time

PCR was conducted to indicate the expression profile of *NnCHS* mRNAs in various tissues, and when response to short-time wounding or low temperature treatment.

Materials and Methods

Plant material

N. nucifera "Taikonglian-36", which kept in Wuhan University, China, was used as material. All the materials were prepared and put into liquid nitrogen immediately.

Clone the cDNAs and gene sequences of NnCHS

The RNA was extracted from young flower, and reverse transcription was performed according to the method described (20). Briefly, based on analysis the conservative amino acid sequences, the degenerate primers pairs (NnCHS DF and NnCHS DR) were designed to amplify partial cDNA (Table.1). The settings for the PCR included an initial denaturing at 94°C for 5 min, followed by 35 cycles of amplification (94°C for 30 s; 50°C for 30 s; and 72°C for 1.2 min) and a finally extension at 72°C for 10 min. The fragment was ligated into pGEM-T vector (Promega) and the product of ligation was transformed into competent *E. coli* DH5α, then the fragment was sequenced. Full-length cDNA sequences of *NnCHS* were obtained by the procedures of rapid amplification cDNA ends (RACE) method, using BD SMART™ RACE cDNA Amplification Kit (BD Biosciences Clontech). For the 3'-RACE, the primer pairs for nest PCR consisted of NnCHS 3GSP1 with UPM for the first-run PCR, and NnCHS 3GSP2 with NUP for the second-run PCR. Finally, The degenerate prime NnCHS 5F1 for 5'RACE was designed according the N-terminal conserved amino acids (MVTVE) compared with other species. The RT-PCR was performed using the NnCHS 5F1 and NnCHS 5R1 for the 5'-RACE (Table.1). The complete cDNA sequences of *NnCHS* were cloned and sequenced. Once the complete cDNAs sequences were determined, *NnCHS* gene was amplified and sequenced to examine the presence of intron by following primer pairs (NnCHS GF and NnCHS GR) in Table 1, using genomic DNA as template.

Similarity and phylogenetic analysis of NnCHS

Sub-cellular location of NnCHS was evaluated by Protcomp Version 9.0 software. Multiple sequence alignment was created using CLUSTAL W. And the subsequent phylogenetic tree based on the amino acid sequences was performed by the Parsimony method using the MEGA software version 4 (21).

Homology modeling of NnCHS

The RSCB protein data bank (http://www.rcsb.org/pdb/home/home.do) was used to find the suitable structure templates for homology model. Alfalfa CHS (PDB code. 1BI5) was chosen as the template (22). The 3D models of NnCHS were constructed by the academic version 6.2 of MODELLER (23) with the default parameters that proposed loop conformations. The qualities of the models were further assessed by PROCHECK 3.5 (24). The best model was selected, and 3D model of NnCHS was shown by Swiss-pdbviewer 4.1.0.

Table 1. Names and sequences of primers used in the present study.

Primer name	Sequence (5'-3')
NnCHS DF	CCKTCHYTGGAYGCNMGRCARGAC
NnCHS DR	GGBCCRAANCCRAANARMACACC
NnCHS 3GSP1	GCATCTCCGACTGGAACTCA
NnCHS 3GSP2	TGTGTGCTGTTCATATTGGAT
NnCHS 5F1	GCATGGTNACCGTNGA
NnCHS 5R1	CCACTGGTGGTGCAGAAGAC
NnCHS GF	ATGGTGACCGTGGAAGACAT
NnCHS GR	TGGCAGCGATACTGTGAAGC
NnCHS F	CCGTGTCCTTGTCGTCT
NnCHS R	AATAACTGCGGCTGCTC
5S rRNA F	GGATGCGATCATACCAGCAC
5S rRNA R	GGGAATGCAACACGAGGACT

The expression of NnCHS mRNAs in various tissues

Fresh materials such as leaves, terminal bud, root, white and red flower were collected when the flower has developed for one week. In order to evaluate the expression of *NnCHS* mRNAs, *5S rRNA* (GenBank accession no. EF121364) was chosen as the reference gene. The special primers for *NnCHS* (NnCHS F and NnCHS R), and *5S rRNA* (*5S rRNA* F and *5S rRNA* R) were designed with primers analyzing software Primer Premier 5 (Premier) (Table. 1). Real-time PCR was carried out by DNA binding dye SYBR Green I (TOYOBO) for detection of PCR products. The amplification program consisted of one cycle at 94°C for 30 s, followed 30 cycles of 94°C for 15 s and 60°C for 20 s. According to the method published (25), the relative expression of *NnCHS* was calculated using *5S rRNA* as the reference gene.

The expression of NnCHS mRNAs in response to mechanical wounding and short-time low temperature

For wounding treatment, the leaves of *N. nucifera* were treated by pressure-stress with a needle puncher. The leaves were separately collected at 1.5 hours post wounding (hpw), 3 hpw, 4.5 hpw and 6 hpw, using the intact leaves as the control. Then the total RNAs of materials were isolated and expression of *NnCHS* was demonstrated by Real-time PCR as method described above, the primers in Table 1 were designed as our previous research (20, 25). For short-time low temperature treatment, the seeds of *N. nucifera* were cultured in water for 5 days until their plumule protrusion was about 4 cm. The germinated seedlings were incubated at 4°C, -10°C and -20°C respectively, the seedlings were collected after 4 hours treatment. The patterns of *NnCHS* expression were examined by Real-time PCR using the seedlings at room temperature as control.

Statistical analysis

Experiments were repeated at least twice to ensure reproducibility. Statistical analysis was performed according to the report (26).

Results

Sequences and structure analysis of NnCHS

Six *NnCHS* cDNAs were isolated (GenBank accession no. FJ999627, FJ999628, FJ999629, FJ999630, FJ999631 and FJ999632), with only several different nucleotides. While the protein sequence was almost the same (data not shown). So the cDNA (GenBank acces-

```
1     ATGGTGACCGTGGAAGACATCCGCAAGGCACAGAGGGCTGAAGGACCGGCCACTGTGATGGCCATTGGAACAGCCAATCCGCCCAACTGT
1      M  V  T  V  E  D  I  R  K  A  Q  R  A  E  G  P  A  T  V  M  A  I  G  T  A  N  P  P  N  C
91    GTCGACCAGAGCACATACCCTGACTACTACTTCCGGATCACCAACAGCGAGCACAAGACAGAGCTCAAAGAGAAGTTTAAGCGCATGTGC
31     V  D  Q  S  T  Y  P  D  Y  Y  F  R  I  T  N  S  E  H  K  T  E  L  K  E  K  F  K  R  M  C
181   GAGAAATCAATGATCAAGAAGGCTACATGCACTTGACCGAGGAGATCTTGAAGGAGAACCCAAACATCTGTGAGTACATGGCTTCCTCG
61     E  K  S  M  I  K  K  R  Y  M  H  L  T  E  E  I  L  K  E  N  P  N  I  C  E  Y  M  A  S  S
271   CTGGATGCTAGACAGGACATGGTGGTGGTTGAGGTGCCAAAACTGGGCAAGGAGGCTGCCACGAAGGCCATTAAGGAATGGGGACAGCCC
91     L  D  A  R  Q  D  M  V  V  V  E  V  P  K  L  G  K  E  A  A  T  K  A  I  K  E  W  G  Q  P
361   AAGTCCAAGATCACCCACCTTGTCTTCTGCACCACCAGTGGCGTCGACATGCCCGGGGCTGACTACCAGCTCACCAAGCTCCTCGGCCTT
121    K  S  K  I  T  H  L  V  F  C  T  T  S  G  V  D  M  P  G  A  D  Y  Q  L  T  K  L  L  G  L
451   CGCCCCTCCGTCAAGAGACTCATGATGTACCAACAAGGATGCTTCGCCGGAGGCACAGTCCTTCGCCTGGCCAAGGACCTTGCAGAGAAC
151    R  P  S  V  K  R  L  M  M  Y  Q  Q  G  C  F  A  G  G  T  V  L  R  L  A  K  D  L  A  E  N
541   AACAGAGGCGCCCGTGTCCTTGTCGTCTGCTCAGAGCTCACTGCTGTTACCTTCCGTGGCCCAAGTGATACTCACCTCGACAGTCTTGTA
181    N  R  G  A  R  V  L  V  V  C  S  E  L  T  A  V  T  F  R  G  P  S  D  T  H  L  D  S  L  V
631   GGCCAGGCACTCTTCGGGGATGGAGCAGCCGCAGTTATTGTGGGTGCAGACCCGGTGCCCGGTGTAGAAAAGCCTTTGTTTGAGTTGGTG
211    G  Q  A  L  F  G  D  G  A  A  A  V  I  V  G  A  D  P  V  P  G  V  E  K  P  L  F  E  L  V
721   TCGGCAGCCCAGACAATTCTCCCAGACAGCCATGGCGCCATTGACGGGCACCTGAGAGAGGTTGGACTTACCTTCCACCTGCTCAAGGAT
241    S  A  A  Q  T  I  L  P  D  S  H  G  A  I  D  G  H  L  R  E  V  G  L  T  F  H  L  L  K  D
811   GTGCCCGGGCTCATCTCAAAGAACATCGAGAAGAGCCTGGTGGAGGCATTCCAGCCTCTGGGCATCTCCGACTGGAACTCAATTTTCTGG
271    V  P  G  L  I  S  K  N  I  E  K  S  L  V  E  A  F  Q  P  L  G  I  S  D  W  N  S  I  F  W
901   ATCGCCCACCCTGGTGGTCCAGCCATCCTAGACCAAGTGGAAGAAAAGCTGGCCCTTAAGCCCGAGAAGCTAAGCCCCACACGACACATC
301    I  A  H  P  G  G  P  A  I  L  D  Q  V  E  E  K  L  A  L  K  P  E  K  L  S  P  T  R  H  I
991   CTGAGCGAGTATGGAAACATGTCAAGTGCTTGTGTGCTGTTCATATTGGATGAGATGCGGAAGAAGTCGATTGAGGATGGCCTCAAGACC
331    L  S  E  Y  G  N  M  S  S  A  C  V  L  F  I  L  D  E  M  R  K  K  S  I  E  D  G  L  K  T
1081  ACTGGAGAAGGGCTCGAGTGGGGTGTGCTGTTTGGGTTTGGACCAGGGCTCACCGTTGAGACTGTGGTGCTTCACAGTATCGCTGCCTAG
361    T  G  E  G  L  E  W  G  V  L  F  G  F  G  P  G  L  T  V  E  T  V  V  L  H  S  I  A  A  *
1171  AGAGAGATGCATGGAGGCTCCAAAATCTCTCTTAACCAAATGAATGGTTAAAATTTGCATTTTCTTCTTGTTTTCCTCCCTCTCAAGTGAA
1261  TTGAGATGTGGGTTGGTTGTTTGATCTGGTCTTGTAATTGTAGTATGCCAATGTGTCCTACCGACAAAGACATCGATCAGAGAATGTA
1351  TTACTTTATATTTAATAATACATATGGTCACCACGTTCTTTGGCAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

Figure 1. Nucleotide and deduced amino acid sequence of *NnCHS*. Six *NnCHS* cDNAs had been deposited in GenBank (GenBank accession no. FJ999627, FJ999628, FJ999629, FJ999630, FJ999631 and FJ999632), with only several different nucleotides. While the protein sequence was almost the same (data not shown). The cDNA (GenBank accession no. FJ999627) was investigated in this study. The malonyl-CoA binding motif (^{313}VEEKLALKPEKLSPTRH^{329}) and highly conserved CHS signature sequence (^{372}GFGPG^{376}) were underlined. Three conserved catalytic residues (Cys^{164}, His^{303}, and Asn^{336}) were boxed. Two important residues (Phe^{215} and Phe^{265}) determining the substrate specificity of CHS were circled. The initiation codon (ATG) was underlined, the asterisk (*) indicated the stop codon.

sion no. FJ999627) was investigated in this study. The cDNA of *NnCHS* containing 1426 bp had been isolated in young flower by RACE, which consisted the open reading frame of 1167 bp, start code ATG, stop code TAG, 259 bp 3′ untranslated region and poly(A) tails (Fig. 1). The degenerate prime NnCHS 5F1 for 5'RACE was designed according the N-terminal conserved amino acids (MVTVE). So *NnCHS* cDNAs contained the complete opening reading frame of *NnCHS* and 3'UTR, and it could not be concluded if there were 5' UTR upstream of the ATG. The putative NnCHS protein encoded 389 amino acid residues with putative molecular weight of 42.55 KD and isoelectric point 6.10. Protcomp Version 9.0 software indicated NnCHS protein sub-cellular located in cytoplasm. Consistent with previous results with other CHS, the NnCHS also contained three imperative catalytic residues (Cys^{164}, His^{303} and Asn^{336}), and two highly conserved Phe residues (Phe^{215} and Phe^{265}), which were found to be important for the formation of the active sites as well as for the substrate specificity of CHS (Fig. 1). Additionally, the malonyl-CoA binding motif (^{313}VEEKLALKPEKLSPTRH^{329}) and highly conserved CHS signature sequence (^{372}GFGPG^{376}) were also found in the NnCHS protein (Fig.1).

CHS genes in other species usually contained one or two intron. The *NnCHS* cDNA was compared with the known gene sequences from *Physcomitrella patens CHS* (*PpCHS11*, ABU87504), *Antirrhinum majus CHS* (*AmCHS*, X03710) and *Freesia hybrida CHS* (*FhCHS*, JF732898), indicating there was only one intron in *NnCHS* gene. We designed the primes according the sequences around splicing site and cloned the intron sequence of *NnCHS* gene, using genomic DNA as template. *NnCHS* gene (GenBank, accession no. FJ999634) was cloned, consisting two exons inserted by a 426 bp

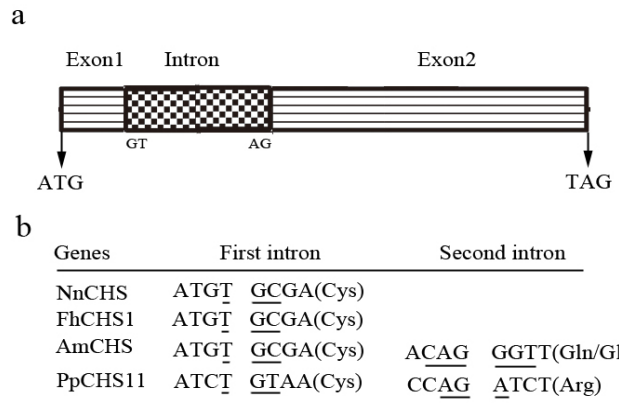

Figure 2. Common characters of *NnCHS* gene (GenBank, accession no. FJ999634). a, Exons-intron architecture of *NnCHS* gene, exons, intron, initiation codon (ATG) and terminator codon (TAG) were labeled. b, Nucleotide sequences surrounding splice sites in *NnCHS*, *Freesia hybrida CHS* (*FhCHS*, JF732898), *Antirrhinum majus CHS* (*AmCHS*, X03710) and *Physcomitrella patens CHS* (*PpCHS11*, ABU87504). Amino acid residues and codons (underlined) split by introns were shown.

intron between ^{175}ATGT^{178} and ^{179}GCGA^{182} (Fig. 2a). The intron of *NnCHS* gene started with the nucleotides 'GT' and ended with the nucleotides 'AG', following the GT-AG rule (Fig. 2b).

Similarity and phylogenetic analysis of NnCHS protein

The multiple alignment of NnCHS was performed by selecting CHS sequences of representative species from high plants including *Juglans nigra*, *Gossypium arboretum*, *Abelmoschus esculentus*, *Theobroma cacao*, *Hibiscus cannabinus*, *Malus domestica*, *Pyrus communis*, *Prunus avium*, *Vaccinium ashei*, *Clitoria ternatea*, *Mangifera indica*, *Camellia nitidissima* and *Camellia sinensis*. Compared with other CHSs, the three imperative catalytic (Cys-His -Asn) and two Phe residues were highly conserved in CHS superfamily, which were involved in formation of the active sites as well as for the substrate specificity of CHS (Fig. 3). Additionally, the

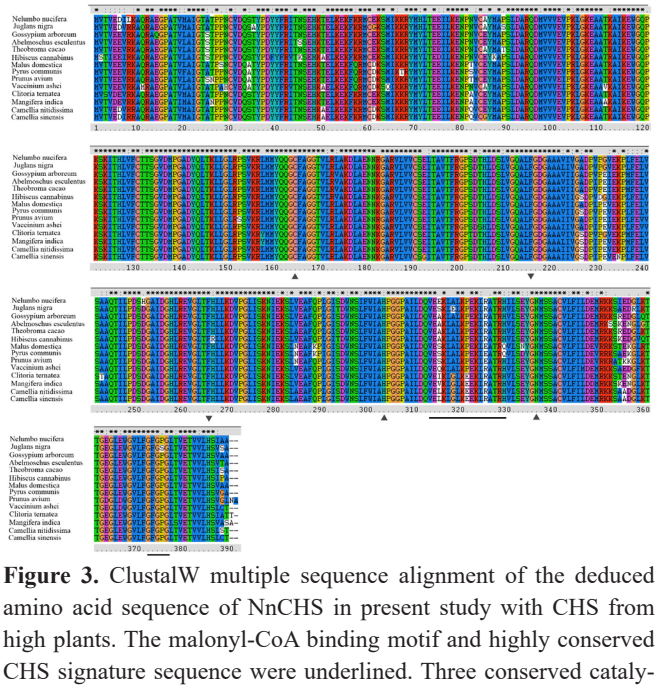

Figure 3. ClustalW multiple sequence alignment of the deduced amino acid sequence of NnCHS in present study with CHS from high plants. The malonyl-CoA binding motif and highly conserved CHS signature sequence were underlined. Three conserved catalytic residues (Cys, His and Asn) were labeled as (▲). Two highly conserved Phe residues were marketed with (▼).

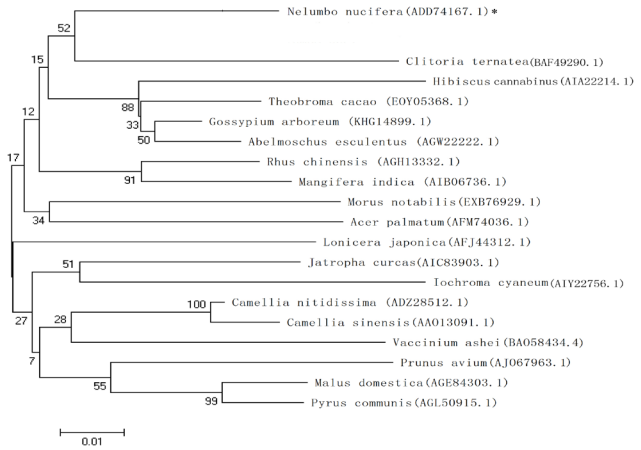

Figure 4. Phylogenetic analysis of NnCHS was performed by neighbour-joining method using MEGA4 software. Numbers in the branches represented the bootstrap values (%) from 100 replicates.

malonyl-CoA binding motif and highly conserved CHS signature sequence were also found in CHS superfamily (Fig. 3). Moreover, the phylogenetic analysis indicated NnCHS grouped together with high plants and provided evidence that CHS had been derived from a common ancestor (Fig. 4).

The homology model of NnCHS

Considering the high homology of NnCHS, the putative three-dimensional structure of NnCHS was constructed by the homology model using the known crystal structure of Alfalfa CHS (PDB code. 1BI5) in order to understand the structure-function relationship of NnCHS (Fig. 5). The NnCHS protein and template protein shared a high degree of homology (86.6 %), suggesting a high reliability of predicted structure. The 3D structure of NnCHS was generated by MODELLER software. The Ramachandran plot (data not shown) provided by PROCHECK demonstrated that most of model residues were in most favorable regions, and none of residues in generously allowed regions and disallowed regions.

Data analysis of Real-time PCR

All the products of Real-time PCR were from 75 bp through 200 bp in size, which were exactly suitable to improve the efficiency of PCR. The data indicated the R^2 values of *5S rRNA* (0.997) and *NnCHS* (0.999) were close to 1. Moreover, the amplification efficiency for *NnCHS* (0.96) was approximately equal to that of *5S rRNA* (1.00) used as reference genes. So the $2^{-\Delta\Delta Ct}$ methods could be used to calculate the relative quantity. *NnCHS* mRNAs were detected in various tissues of *N. nucifera* including leaves, terminal bud, root, white and red flower when the flower had developed for one week. The highest expression of *NnCHS* mRNA was found in the flower, while red flower kept more *NnCHS* mRNA level than white one (2 fold increase). The mRNA level of root was at moderate level. However, the lowest expression was detected in leaves compared to all other selected tissues (Fig. 6a). Additionally, Real-time PCR analysis indicated that short-time mechanical wounding significantly elevated the expression of *NnCHS* mRNA as soon as 1.5 hpw (Fig. 6b), using the intact leaves as

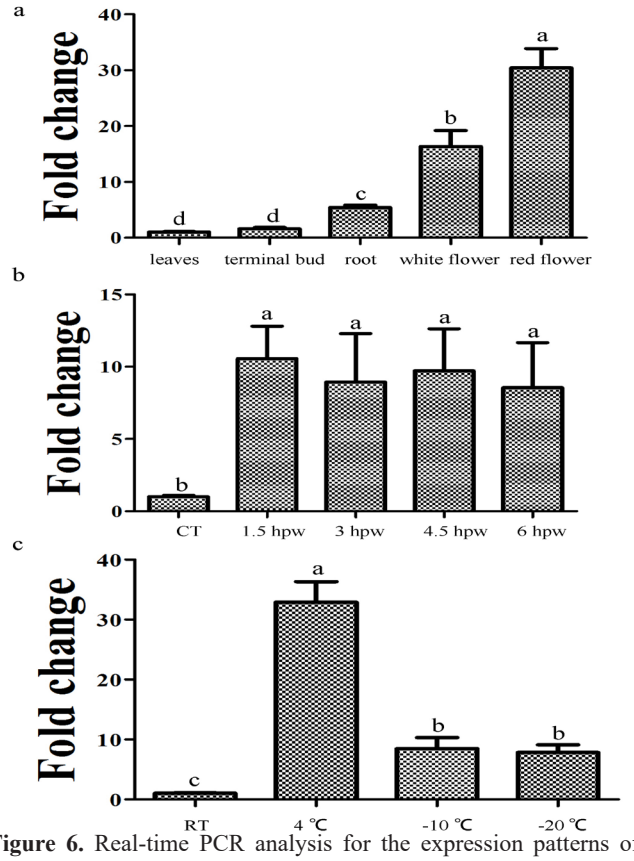

Figure 6. Real-time PCR analysis for the expression patterns of *NnCHS* mRNAs in various tissues and during short-time mechanical wounding or cold temperature. a, Relative amounts of *NnCHS* mRNAs in various tissues. Total RNAs were isolated from root, leaves, terminal bud, white and red flower, the levels of *NnCHS* mRNAs were examined by Real-time PCR using *5S rRNA* as the reference gene. *NnCHS* mRNAs were expressed in various tissues of *N. nucifera*, with the highest expression in flower and lowest level in the leaves. b, The expression of *NnCHS* mRNA was up-regulated in response to short-time mechanical wounding. Real-time PCR analysis indicated that short-time mechanical wounding significantly elevated the expression of *NnCHS* mRNA as soon as 1.5 hours post wounding (hpw), using the intact leaves as control. And mRNA level of *NnCHS* was relatively stable after 1.5 hpw. c, Short-time low temperature significantly augmented the expression of *NnCHS* mRNA. *NnCHS* mRNA reached the highest level during 4°C treatment, while the level was relatively decreased when the seeds were treated by -10°C and -20°C, which was still more than the level of *NnCHS* mRNA in room temperature. Different letters represented significant difference at $p<0.05$.

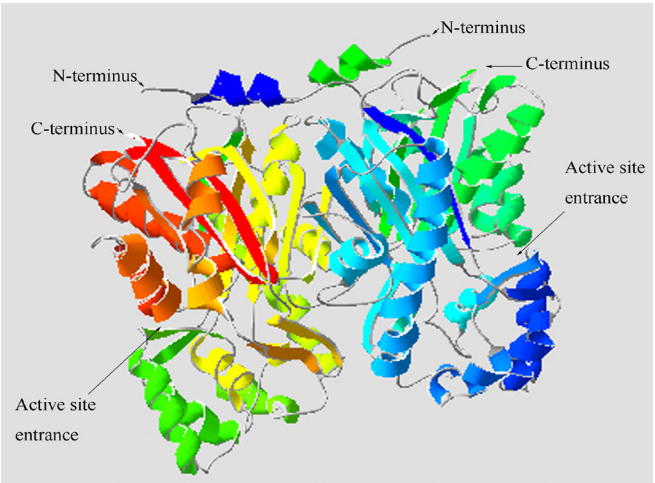

Figure 5. Functional unit of NnCHS. The homo-dimer with one of the monomer highlighting secondary structure components: α-helix, β-strand and loop were shown by different colors. The N-terminus, C-terminus and active site entrance were also labeled.

control. *NnCHS* mRNA was relatively stable after 1.5 hpw. Short-time low temperature also significantly augmented the expression of *NnCHS* mRNA of germinated seedlings (Fig. 6c). *NnCHS* mRNA reached the highest level during 4°C treatment, while the level was relatively decreased when the seeds were treated by -10°C and -20°C, which was still more than the expression of *NnCHS* mRNA in room temperature.

Discussion

The cDNAs sequences of *NnCHS* with the complete open reading frame and 3'UTR were isolated and characterized, possessing all the main characteristic amino acid residues, binding motif and conserved CHS signature sequence (Fig. 1). It was indicated most *CHS* genes of plants contained a single intron splitting the conserved Cys in the consensus sequence of (K/Q) R(M/I) C(D/E)KS (27), while few *CHS* had two introns (28, 29). Additionally, *CHS* genes containing two introns also had the first intron at the conserved Cys, and the second intron was situated within the codons for conserved amino acid residues such as Gly in *AmCHS* and Arg in *PpCHS11* (28, 29). In this study, one intron was found in *NnCHS* gene following GT-AG rule (Fig. 2). The intron was inserted in the first Cys codon as other *CHS* genes reported (27). The conserved amino acids for the active sites and the substrate specificity of CHS were found in all CHSs of high plants, and the malonyl-CoA binding motif and CHS signature sequence were also highly conserved (Fig. 3).

The phylogenetic analysis indicated NnCHS grouped together with plants CHS and provided evidence that all CHS had been derived from a common ancestor (Fig. 4). NnCHS showed high homology (86.6 %) with Alfalfa CHS. Our structural model comparison results showed that the active sites of NnCHS were conserved and matched with Alfalfa CHS (22), suggesting their functional similarity. As shown in figure 5, the 3D structural model of NnCHS usually formed the homodimer as the biological functional unit, with the each monomer of NnCHS consisted of roughly two domains with one active site located at the clef between the two domains and the dimerization interface, therefore sharing the accessible substrate binding area with the dimer partner. Protein structure modeling revealed that NnCHS monomer consisted of two structural domains. In the upper domain, an N-terminal α-helix and a protruding tight loop important for protein dimerisation were observed. Most residues that made up the dimmer interface were located within the upper domain, producing relatively flat surface in the NnCHS monomer. In the lower domain, a large substrate binding site required for chalcone formation was connected to CoA-binding tunnel and the active site was buried in the cleft formed between two domains.

To understand the spatial regulation of *NnCHS* mRNAs, tissue expression of *NnCHS* mRNAs were investigated by real-time PCR in root, terminal bud, young leaves and flower of *N.nucifera*. Different expression of *NnCHS* mRNAs was detected in various tissue types, with the highest expression in flower and lowest level in leaves (Fig. 6a). Real-time PCR illustrated relatively low expression of *NnCHS* mRNAs was observed in non-pigmented tissues including leaves and terminal bud. This indicated *NnCHS* was not flower specific. This result was also consistent with the finding of Li et al. (2010) who reported that most of flavonoids biosynthesis genes including *CHS* were highly expressed in the lower parts of the plants, especially root (30). In particular, flavonoids production was associated with plant-microbe interactions. Several reports illustrated flavonoids expressed in the root conducted as key signaling molecules for some symbiotic and pathogenic plant-microbe interactions (31). The high mRNA level of *NnCHS* in root was probably involved in the high demand of flavonoids biosynthesis for the aim of defense against microbial invasion. Moreover, red pigmented flower showed higher level of *NnCHS* mRNA than white pigmented flower, it suggested that the level of *NnCHS* mRNA might be responsible for the pigment of flower in *N. nucifera*.

In general, plants will secret a number of genes in response to biotic or abiotic stresses. Among these genes, *CHS* is commonly induced under different forms of stress like UV, wounding, and microbial invasion (12). Corresponding to this finding, the expression of *NnCHS* mRNA was stimulated in response to short-time mechanical wounding as soon as 1.5 hpw (Fig. 6b), Additionally, 4°C treatment significantly augmented the expression of *NnCHS* mRNA in germinated seedlings of *N. nucifera*, while the mRNAs levels were relatively decreased in response to -10°C or -20°C treatment (Fig. 6c). All of this results suggested *NnCHS* might play a defensive role in the initial stages of defense response and might be transcriptionally regulated. In future experiment, we will currently try to develop transgenic plants in order to understand the mechanism of *NnCHS* in regulation the color of the flower.

Acknowledgements

The authors are greatly appreciative of the reviewers for their helpful comments and detailed suggestions. This work is financially supported by the National Science and Technology Supporting Program (no. 2012BAD27B01), Foundation of Henan Educational Committee (16A180024), Doctoral Scientific Research Start-up Foundation from Henan University of Technology (no.31400855), and Fundamental Research Funds for the Henan Provincial Colleges and Universities in Henan University of Technology (no. 2015QNJH07).

References

1. Chandler, S.F., Brugliera, F., Genetic modification in floriculture. *Biotechnol. Lett*. 2011, **33**: 207–214. doi:10.1007/s10529-010-0424-4

2. Maleka, M.F, Albertyn, J., Spies, J.J., The floriculture industry and flower pigmentation. *Philos. Trans. Genet*. 2013, **2**: 55–110. doi:10.1016/j.abb. 2010.07.028.

3. Niyogi, K.K, Safety valves for photosynthesis. *Curr. Opin. Plant. Biol*. 2000, **3**: 455 –460. doi:10.1016/S1369-5266(00)00113-8

4. Grotewold, E., The genetics and biochemistry of floral pigments. *Annu. Rev. Plant. Biol.* 2006, **57**: 761–780. doi: 10.1146/annurev.arplant.57.032905.105248

5. Mol, J., Grotewold, E., Koes, R., How genes paint flowers and seeds. *Trends. Plant. Sci*. 1998, **3**: 212–217. doi:10.1016/S1360-1385(98)01242-4

6. Winkel-Shirley, B., Biosynthesis of flavonoids and effects of stress. *Curr. Opin. Plant. Biol.* 2002, **5**: 218–223. doi:10.1016/S1369-5266(02)00256-X

7. Buer, C.S., Imin, N., Djordjevic, M.A., Flavonoids: new roles for old molecules. *J. Intergr. Plant. Biol.* 2010, **52**:98–111. doi: 10.1111/j.17447909.2010. 00905.x

8. Ferreyra, M.L.F., Rius, S.P., Casati, P., Flavonoids: biosynthesis, biological functions, and biotechno-logical applications. *Front. Plant. Sci.* 2012, **3**: 1– 15. doi: 10.3389/fpls.2012.00222.

9. Winkel-Shirley, B., It takes a garden. How work on diverse plant species has contributed to an un-derstanding of flavonoid metabolism. *Plant. Physiol*. 2001, **127**: 1399 – 1404. doi:10.1104/pp.010675.

10. Pang, Y., Shen, G., Wu, W., Liu, X., Lin, J., Tan, F., et al., Characterization and expression of chalcone synthase gene from Ginkgo biloba. *Plant. Sci.* 2005, **168**: 1525–1531. doi:10.1016/j.plantsci.2005.02.003

11. Lei, W., Tang, S.H., Luo, K.M., Sun, M., Molecular cloning and expression profiling of a chalcone synthase gene from hairy root cultures of Scutellariaviscidula Bunge. *Genet. Mol. Biol.* 2010, ***33(2)***: 285-291. doi:10.1590/S1415-47572010005000031

12. Schenk, P.M., Kazan, K., Wilson, I., Anderson, J.P., Richmond, T., Somerville, S.C., Manners, J.M., Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. *Proc. Natl. Acad. Sci. USA*. 2000, **97**: 11655-11660. doi:10.1073/pnas.97.21.11655

13. Abe, I., Morita, H., Structure and function of the chalcone synthase superfamily of plant type III poly-ketide synthases. *Nat. Prod. Rep*. 2010, **27**: 809– 838. doi:10.1039/B909988N

14. Lanz, T., Tropf, S., Marner, F., Schröder, J., Schröder, G., The role of cysteines in polyketide synthases. Site-directed mutagenesis of resveratrol and chalcone synthases, two key enzymes in different plant-specific pathways. *J. Biol. Chem.* 1991, **266**: 9971 – 9976.

15. Suh, D.Y., Kagami, J., Fukuma, K., Sankawa, U., Evidence for catalytic cysteine—histidine dyad in chalcone synthase. *Biochem. Bioph. Res. Co.* 2000 **275**: 725 – 730. doi:10.1006/bbrc.2000.3368

16. Esau, K., Kosakai, H., Leaf arrangement in Nelumbo nucifera: a re-examination of a unique phyllotaxy. *Phytomorphol.* 1975, **25**: 100–112.

17. Borsch, T. Barthlott, W., Classification and distribution of the genus Nelumbo Adans (Nelumbonaceae). *Betrage zur Biologie der Pflanzen*. 1994, **68**:421–450.

18. Sridhar, K.R., Bhat, R., Lotus -A potential nutraceutical source. *J. Agric. Technol*. 2007, **3**: 143–145. doi:10.7324/JAPS.2015.50419

19. Priestley, D.A., Posthumus, M.A., Extreme longevity of lotus seeds from Pulantien. *Nature.* 1982, **299**: 148-149. doi:10.1038/299148a0.

20. Dong, C., Zheng, X.F., Li, G.L., Zhu, H.L., Zhou, M.Q., & Hu, Z.L., Molecular cloning and expression of two cytosolic copper-zinc superoxide dismutases genes from Nelumbo nucifera. *Appl. Biochem. Biotech*. 2011, **163 (5)**: 679-691. doi: 10.1007/s12010-010-9074-1

21. Tamura, K., Dudley, J., Nei, M., & Kumar, S., MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 2007, **24**: 1596-1599.doi: 10.1093/molbev/msm092

22. Ferrer, J.L., Jez, J.M., Bowman, M.E., Dixon, R.A., Noel JP., Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. *Nat. Struct. Biol*. 1999, **6(8)**:775-84. doi:10.1038/11553

23. Sanchez, R., & Sali, A., Evaluation of comparative protein structure modeling by MODELLER-3. *Proteins (Suppl)*. 1997, **1**: 50-58. doi:10.1002/(SICI)1097-0134(1997)1+%3C50::AID-PROT8%3E3.0.CO;2-S

24. Laskowski, R.A., MacArthur, M.W., Moss, D.S., & Thornton, J.M., Procheck: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr*. 1993, **26**: 283-291. doi:10.1107/S0021889892009944

25. Dong, C., Zheng, X.F., Li, G.L., Pan, C., Zhou, M.Q., & Hu, Z.L., Cloning and expression of one chloroplastic ascorbate peroxidase gene from Nelumbo nucifera. *Biochem. genet*. 2011, **49 (9)**: 656-664.doi: 10.1007/s10528-011-9440-x

26. Babic, M., Radic, S., Cvjetko, P., Roje, V., Pevalek-Kozlina, B., & Pavlica, M., Antioxidative response of Lemna minor plants exposed to thallium(I)-acetate. *Aquat. Bot*. 2009, **91**: 166-172. doi:10.1016/j.aquabot.2009.05.005

27. Sun, W., Meng, X., Liang, L., Jiang, W., Huang, Y., He, J. et al., Molecular and Biochemical Analysis of Chalcone Synthase from Freesia hybrid in Flavonoid Biosynthetic Pathway. *Plos. One.* 2015, **10 (3)**: e0119054. doi:10.1371/ journal.pone.011905

28. Koduri, P.H., Gordon, G.S., Barker, E.I., Colpitts, C.C., Ashton, N.W., Suh, D.Y., Genome-wide analysis of the chalcone synthase superfamily genes of Physcomitrella patens. *Plant. Mol. Biol.* 2010, 72: 247–263. doi:10.1007/s11103-009-9565-z

29. Ma, L.Q., Pang, X.B., Shen, H.Y., Pu, G.B., Wang, H.H., Lei, C.Y., et al., A novel type III polyketide synthase encoded by a three-intron gene from Polygonum cuspidatum. *Planta*. 2009, **229**: 457 –469. doi: 10.1007/s00425-008-0845-7

30. Li, X.H., Park, N.I., Xu, H., Woo, SH., Park, C.H., Park, S.U., Differential expression of flavonoid biosynthesis genes and accumulation of phenolic compounds in common buckwheat (Fagopyrum esculentum). *J. Agr. Food. Chem*. 2010, **58**: 12176-12181. doi:10.1021/jf103310g

31. Steinkellner, S., Lendzemo, V., Langer, I., Schweiger, P., Khaosaad, T., Toussaint, J.P., Vierheilig, H., Flavonoids and strigolactones in root exudates as signals in symbiotic and pathogenic plant-fungus interactions. *Molecules*. 2007, **12**: 1290-1306. doi: 10.3390/12071290