

Molecular cloning and expression analysis of the ethylene insensitive3 (EIN3) gene in cucumber (*Cucumis sativus*)

B.B. Bie, J.S. Pan, H.L. He, X.Q. Yang, J.L. Zhao and R. Cai

Plant Science Department, School of Agriculture and Biology,
Shanghai Jiaotong University, Shanghai, China

Corresponding author: R. Cai
E-mail: cairun@sjtu.edu.cn

Genet. Mol. Res. 12 (4): 4179-4191 (2013)

Received September 18, 2012

Accepted January 25, 2013

Published October 7, 2013

DOI <http://dx.doi.org/10.4238/2013.October.7.4>

ABSTRACT. The plant gaseous hormone ethylene regulates many aspects of plant growth, development, and responses to the environment. Ethylene insensitive3 (EIN3) is a key transcription factor involved in the ethylene signal transduction pathway. To gain a better understanding of this particular pathway in cucumber, the full-length cDNA encoding EIN3 (designated as *CsEIN3*) was cloned from cucumber for the first time by rapid amplification of cDNA ends. The full length of *CsEIN3* was 2560 bp, with an open reading frame of 1908 bp encoding 635 amino acids. Sequence alignment and phylogenetic analyses revealed that *CsEIN3* has high homology with other plant EIN3/EIL proteins that were derived from a common ancestor during evolution, and *CsEIN3* was grouped into a cluster along with melon. Homology modeling demonstrated that *CsEIN3* has a highly similar structure to the specific DNA-binding domain contained in EIN3/EIL proteins. Based on quantitative reverse transcription-polymerase chain reaction analysis, we found that *CsEIN3* was constitutively expressed in all organs examined, and was increased during flower development and maturation in both male and female flowers. Our results suggest that

CsEIN3 is involved in processes of flower development. In conclusion, this study will provide the basis for further study on the role of EIN3 in relevant biological processes of cucumber and on the molecular mechanism of the cucumber ethylene signaling pathway.

Key words: *Cucumis sativus* (cucumber); Expression patterns; Ethylene insensitive3 (EIN3); RACE

INTRODUCTION

Ethylene (C₂H₄) was the first example of a gaseous signaling molecule in biological systems, discovered more than a century ago (Neljubov, 1901). It regulates a variety of developmental and stress responses in plants, including seed germination, cell elongation, cell fate, sex determination, fruit ripening, flower senescence, leaf abscission, defense against pathogens, and responses to mechanical trauma, cold, and salt stresses (Lin et al., 2009). Components of the pathway for ethylene signal transduction were identified by genetic approaches in *Arabidopsis*, and a basic model of the ethylene signal transduction pathway has been established (Guo and Ecker, 2004).

The ethylene insensitive3/ethylene insensitive3-like (EIN3/EIL) proteins, encoded by a small multigene family, are positive regulators of the ethylene signaling pathway at the downstream position that act as transactivating factors to trigger ethylene responses, mainly via the regulation of ethylene response factor (*ERF*) genes, known to be their downstream targets. Functional domains of EIN3/EIL proteins have been determined by nuclear magnetic resonance spectroscopy. The DNA-binding domains (DBDs), a V-shaped cleft formed by five α -helices, were identified from *AtEIL3* (*Arabidopsis thaliana EIL3*). This conserved DBD is also found in all EIN3 homologs examined so far in other plant species. The DBD of EIN3/EIL binds specifically to the EIN3-binding site located in the promoter region of ethylene-regulated genes (Yamasaki et al., 2005). Some members of EIN3/EIL are regulated at the post-transcriptional level by proteasomal degradation of EIN3 Binding F-Box 1 and F-Box 2 (*EBF1* and *EBF2*) that can precipitate the degradation of EIN3/EIL proteins when ethylene is present; in contrast, exogenous ethylene does not affect the accumulation of *EIN3/EIL* mRNAs (An et al., 2010). However, some EIN3/EIL mRNAs can be regulated at the transcriptional level, based on the fact that exogenous ethylene can regulate the accumulation of mRNA (De Paepe et al., 2004; Parra-Lobato and Gomez-Jimenez, 2011).

As a key nuclear transcription factor in the ethylene signaling pathway, EIN3/EIL proteins have received more attention than other components in recent years. The previous studies indicate that EIN3/EIL proteins can affect the “triple response” of seedlings (Chao et al., 1997), leaf, and root growth (Chao et al., 1997; Lee and Kim, 2003), female flower development (Chao et al., 1997; Hibi et al., 2007), and fruit ripening (Tieman et al., 2001; Chen et al., 2004), and participate in the response to stresses involving salt (Cao et al., 2007; Zhang et al., 2011), cold (Shi et al., 2012), iron deficiency (Bauer and Blondet, 2011), and sulfur deficiency (Maruyama-Nakashita et al., 2006). In addition, EIN3/EIL proteins participate in the cross-talk between ethylene signaling and other signaling, including those of gibberellins (An et al., 2012), salicylic acid (Chen et al., 2009) jasmonic acid (Zhu et al., 2011), and glucose (Yanagisawa et al., 2003).

Cucumber is an economically important crop as well as a model system for sex determination studies, owing to its rich diversity of sex expression (Huang et al., 2009). Ethylene has been demonstrated to promote female flower development and is highly correlated with femaleness in cucumber (Rudich et al., 1972). To date, some genes that determine the ratio of male-to-female flowers in cucumber have been cloned and they all encode 1-aminocyclopropane-1-carboxylic acid synthase (ACS), a key rate-limiting enzyme in ethylene biosynthesis (Knopf and Trebitsh, 2006; Li et al., 2009). In contrast, still little is known about the role of ethylene signaling in the sex determination of flowers, mainly because the important components in the ethylene signaling pathway have not yet been separated and identified. Only a few components have been isolated from cucumber, including ETR2, ETR1, ERS (Yamasaki et al., 2000), and ERAF17 (Ando et al., 2001), and the majority of key components are still waiting to be isolated and characterized.

Previous studies have discovered that the *M* locus, a genetic locus that predominantly controls sex type in cucumber, co-segregates with an *EIL* genomic sequence (Liu et al., 2008). Recently, the *M* gene was cloned and found to encode ACS (Li et al., 2009). To better understand whether EIN3/EIL proteins act as key transcription factors in the ethylene signal transduction pathway and are indeed involved in sex determination and other biological processes in cucumber, it is imperative to isolate the full-length sequence of *EIN3/EIL* from cucumber. In the present study, we first isolated the full-length cDNA of *EIN3* from cucumber using rapid amplification of cDNA ends (RACE) and then investigated its expression pattern in various cucumber organs and in different developmental stages of both male and female flowers. This work laid the groundwork for further study of EIN3/EIL functions in cucumber.

MATERIAL AND METHODS

Plant materials and sampling

The monoecious plant line S52 (inbred line derived from a southern Chinese local cultivar) was used in all the experiments reported in this paper, and was provided by the Cucumber Research Group, School of Agriculture and Biology, Shanghai Jiaotong University, China. Plants were grown in a greenhouse at Shanghai Jiaotong University. When the plants had reached the 15 to 20 nodes stage, the taproot and branch root, main stems, mature leaves, shoot apices, and mature male flowers (MMF) and mature female flowers (MFF) on the day of anthesis, as well as both male and female flowers at five different developmental stages separated on the basis of the corolla length, were collected for the gene expression analysis. With regard to the flower samples, the whole body of the flower, including the calyx, ovary, stamen, pistil, and petal, was collected.

Isolation of RNA and synthesis of the first-strand cDNA

Total RNA was extracted from material previously frozen in liquid nitrogen and stored at -80°C , by using the TRIpure reagent (Aidlab, China) according to the kit instructions. After determining the quality and concentration of RNAs by spectrophotometry for the ratio of $\text{OD}_{260}/\text{OD}_{280}$ and gel electrophoresis, total RNA (1 μg) derived from leaves of cucumber cv. S52 was used for the synthesis of the first strand of cDNA according to the manual of the SmartTM Race cDNA Amplification Kit (Clontech, USA).

Cloning of full-length cDNA of *CsEIN3* and RACE-polymerase chain reaction (PCR) experiments

A BLASTP search, using the amino acid sequence of the *EIN3* gene from *Arabidopsis thaliana* (GenBank accession No. AEE76421), against the protein database of the Cucumber Genome DataBase (<http://cucumber.genomics.org.cn>) was conducted, and a tentative consensus sequence of *csa012885* located at chromosome 1 was identified. The *csa012885* sequence showed 76% amino acid identity to AEE76421, with an open reading frame (ORF) length of 1395 bp encoding 464 amino acids. Based on the nucleotide sequence of *csa012885*, a pair of primers covering the cDNA of *csa012885* was designed to obtain the cDNA sequence containing the ORF (see Table 1).

Table 1. Primers used in this study.

Name	Sequence (5'-3')	Utilization
CsEIN3-ORF-Fw	TCTGCTGATTCATTTTGGG	ORF cloning primers
CsEIN3-ORF-Re	CGGAAAGAATACGTCGATAAAG	
3'RACE-F	CGCCCAAGTCAAATCCACCTCC	RACE primers
5'RACE-R1	GGAAGACAAAGGCGGGCAAGAATCAGG	
5'RACE-R2	CGATCCCTTCCTTAACCTTGCTTTGCTC	Universal primers
UPM702	CTAATACGACTCACTATAGGGCGTACGGCAGCTA CAACTCGACAGCTAATACGACTCACTATAGGGC	
UPM	CTAATACGACTCACTATAGGGCAAGCAGTGGTAT CAACGCAGAGTCTAATACGACTCACTATAGGGC	
NUPA	AAGCAGTGGTATCAACGCAGAGT	Full-length cDNA cloning primers
CsEIN3-Full-Fw	GGGTTTTTGCTATGAAATTTATGTT	
CsEIN3-Full-Re	TGGATATATAATCAACAACAGGGCA	Q-PCR for <i>CsEIN3</i>
EIN3-Q-2F	TCCTGATTCTTGCCCGCCT	
EIN3-Q-2R	GGTTGGTTTCCTCTTTCCGATA	Reference gene
CsActin3-Fw	TCGTGCTGGATTCTGGTG	
CsActin3-Re	GGCAGTGGTGGTGAACAT	

The cDNA sequence containing the ORF of *CsEIN3* was amplified in a PCR reaction of 50- μ L total volume, containing 40 ng cDNA, 0.5 μ M primers CsEIN3-ORF-Fw and CsEIN3-ORF-Re (Table 1), and 25 μ L 2X GoldStar Best MasterMix (CWBiotech, China). The amplification conditions were as follows: 10 min at 95°C, 30 cycles of 30 s at 95°C, 30 s at 60°C, and 2 min 50 s at 72°C, and a final 5 min at 72°C. The amplified products were analyzed by gel electrophoresis on a 1% gel and recovered with a gel extraction kit (Sangon, China). The recovered product was cloned into the pMD18-T vector (TaKaRa, China) by the method described in the instruction manual and transformed into *Escherichia coli* strain DH5 α , followed by sequencing. The PCR amplification and sequencing of *CsEIN3* were repeated three times to avoid any PCR errors. The sequence from this article has been deposited at GenBank under the accession number JQ742004.

Then, the 5' and 3' ends of *CsEIN3* were obtained by RACE-PCR with the SmartTM Race cDNA Amplification Kit (Clontech), and *CsEIN3* sequence obtained previously was used to design the gene-specific primers (Table 1). The 3' RACE reaction was carried out with primers 3' RACE-F and UPM702 (Universal Primer A Mix), with the following amplification conditions: 5 min at 94°C; followed by 30 cycles of 30 s at 94°C, 30 s at 65°C, and 3 min at 72°C; and a final extension at 5 min at 72°C. Two 5' gene-specific primers were de-

signed for the 5' RACE. For the first cycle of amplification of the 5' end of the *CsEIN3* cDNA, primers 5' RACE-R1 and UPM702 (Universal Primer A Mix) were used, with the 5' RACE-ready cDNA as template. For the nested amplification of 5' RACE, primers 5' RACE-R2 and NUPA (Nested Universal Primer A) were used, with the products of the first amplification as templates. The first and nested PCRs were carried out under the same conditions: 5 cycles of 30 s at 94°C, 2 min at 60°; 5 cycles of 30 s at 94°C, 30 s at 55°C, and 2 min at 72°C; and 25 cycles of 30 s at 94°C, 30 s at 55°C, and 1 min at 72°C. The 3' RACE and nested 5' RACE products were purified and cloned into the pMD18-T vector (TaKaRa, China), followed by sequencing.

By aligning and assembling the core fragment, 5' RACE, and 3' RACE products, the full-length cDNA of *CsEIN3* was deduced and confirmed by RT-PCR using the primers CsEIN3-Full-Fw and CsEIN3-Full-Re.

Sequence analysis and 3-D modeling

The ORF prediction was carried out with Genscan (<http://genes.mit.edu/GENSCAN.html>). The similar sequences of other species were obtained by using the BLASTP tool at the NCBI web site (<http://www.ncbi.nlm.nih.gov/BLAST/>). Multiple sequences alignment was performed using ClustalX 2.0 and processed by GeneDoc (<http://www.nrbsc.org/gfx/genedoc/index.html>). MEGA (Molecular Evolutionary Genetics Analysis, version 2.1) was subsequently used to construct the phylogenetic tree by applying the neighbor-joining method. Homology modeling was performed with the SWISS-MODEL program on the ExpASy web server (<http://www.expasy.org/>) (Arnold et al., 2006).

Expression analysis by quantitative real-time PCR

Quantitative real-time PCR (Q-PCR) was performed to investigate the expression level of *CsEIN3* mRNA in different plant organs including roots, stem, leaves, shoot apices, and mature male and female flowers, and in different developmental stages of both male and female flowers, using the SYBR Green method and FTC-3000TM System Real-time Quantitative Thermal Cycler (Funglyn Biotech, Canada).

The cDNA for Q-PCR was synthesized from total RNA extracted from various tissues (three replicates per tissue) using the ReverTra Ace® qPCR RT Kit (Toyobo, China), according to the kit instructions, with oligo(dT) primers. The total RNA used for reverse transcription was incubated with DNase I (Fermentas, China) for 30 min at 37°C to remove the DNA contamination. The primer for Q-PCR was designed using Primer Premier 5.0, and the amplification efficiency was evaluated via standard curve analysis.

The Q-PCR reactions were performed on a 96-well plate, in triplicate, with each reaction consisting of a 25- μ L mixture that contained 12.5 μ L 2X UltraSYBR Mixture (CWBio-tech), 0.4 μ M each primer, and 20 ng cDNA. The amplification conditions were as follows: 10 min at 95°C to activate the DNA polymerase, followed by 40 cycles of 15 s at 95°C and 1 min at 60°C. A dissociation curve was performed to ensure the validity of each specific PCR product. The Q-PCR was repeated three times.

The relative expressions of the *CsEIN3* genes were calculated based on the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001), using the *CsActin3* gene as an internal standard.

RESULTS

Cloning of the full-length cDNA of *CsEIN3*

Using the above-mentioned methods, the full-length cDNA designated as *CsEIN3* was obtained. *CsEIN3* was 2560 bp long, including a 194 bp 5'-untranslated region (UTR), a 458 bp 3'-UTR, a 1908 bp ORF encoding 635 amino acids residues, and a poly(A) tail of 23 bp. In addition, a potential polyadenylation signal (AAATAA) was found 10 bp upstream from the poly(A) stretch (Figure 1).

```

GGGTTTTGCTATGAAATTTATGTTACAGGTGAGGAAGTATAGAGAATT
TTAGTAGCTTCCATATCATCTGATACAGGTGAGGCAGTAGCTGATTAGGATTTGTTTTGGTGTAAAT
CTGGTTGTGGATCTCACCCGATTTCTGCTCATCTGAGTTTACGAAATCTGCTGATTCATTTTGGGGGAA
ATGATGATGATGTTCAAGAGATGGGATTTTGTGATGATGAGATTCTTCTGCTTCATTAAGGAAGGA
M M M M F N E M G F C D D M D F L S A S I T E G
GATGCTGTAGCTCCCACTGATCTGAACTGGTGGGAGACGATTATCTGATGAAGAGATCGACATG
D A V A P P T D P E V V V E D D Y S D E E I D M
GATGACTTGAAGGAGATGTTGGAGGCAAGATGCTTCAAGCTTAAAGAGCAAGCAAGTTAAG
D E L E R R M W R D K M R L K R L K E Q S K V K
GAAGGATGATATTTGAGCAAGCAATCTCAAGACAGCTAGGAGGAGAGATGAGGAGCAGAT
E G I D I V K Q R Q S Q D Q A R R K K M S R A H
GATGGATCTTCAAGTATATGTTGAGGATTATGAACTCTGTAATGCTCAAGCTTTGTATACGGATAAT
D G I L K Y N L K I M E V C N A Q G F V Y G I I
CCTGAGAAAGGAAACAGTAAACCGGCAAGTAACTCCGAGAGTGTGAAAGCAAGTACGATTT
P E K G K P V T C A S D N L R E W W K D K V R F
GATAGAAAGGAGCAGCTGCCATGCCAGTACCAGCAGACATGCAATCTCGAGCAATGATGCGCTCT
D R N G P A A I A K Y Q A D N A I P G R N D G C
AATTCAATGGTCCAAACCCCTCACACTTGCAGAACTCAGATACCACTTAGGTTCTTTTATCAGCT
N S I G P T P H T L Q E L Q D T T L G S L L S A
CTGATGACGACTGTGACCTCTCAAGAAAGATTTCATTTGGAAAGGAGTTCTCGGCAATGTTGGGCT
L M Q H C D P P Q R R F P L E K G V P P P W W P
ACTGGATGAGGAATGTTGGCTCAGCTTGGATGGGAAAGGCAAGGCTCCGGCTCAGAAAGCTT
T G V E E W W P Q L G L P K D Q G P P P Y K K P
CATGACTTGAAGAACTTGAAGATGAGTGTTTGACTGAGTATCAAGCATATGCCCTGATATGGC
H D L K K A W K V G V L T A V I K H M S P D I A
AAGATCCGAGCTGTTAGACAAATCCAAGTGTTCGAGGCAAGATGACTGCCAAGAGAGTCTACATGG
K I R K L V R Q S K C L Q D K M T A K E S A T W
CTTGCATATTAACCAAGGAGGATCTTGGCCGAGAGCTTTATCTGATTTCTCCCGCTTTGCTTCC
L A I I N Q E E I L A R E L Y P D S C P P L S S
CGTGGGCAATGGATTGTTGGCTAATGAGTGTGATGATGATGAGGAGTCTGAGGAGAACCG
A G G N G L L V I N D C S E Y D V E G A E E E P
AGCTTTGATGCAAGATGATAAATGATAAATGATGATGATGATGATGATGATGATGATGATGATGAT
S F D V Q D R K P D N H S S F N L G I D R M R D
CGGCTGCTCCCTGAGCAACCTTATGCAATGAAAGGAGGTTACTACAACCTTGGATTTATGCGAAG
R V S L R Q P P P Y A M K G E V T T N L D F M R K
AGGAACCAACAGTATTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
R K P T S D L N M M M D Q K I Y T C E F L Q C P
TACAGCAACTTCTGTTGGTTAATGACAGGACATCCAGACAATCATCAGTTGACTGCCCATATAGA
Y S E L R L G F N D R T S R D N H Q L T C P Y R
ACTTCTCAGAAATTCAGTGGCTCAAGTTTTCATGCAATGAGTCAACAGGATTTCTCCCTCAGTGGTT
T S S E F S G S S F H V N E V K P V I P P Q S P
GCCGCCCAAGTCAATCCACTCGGTCAGTCACTCCGCTCATCTTTCAGCTGCTGACTTTAGGCGTT
A P P K S N P P P V S S V P S S F D L S T L G V
CCAGAGATGGCCAAAGCTGATCAGTGGCTCATGATGATGATGATGATGATGATGATGATGATGATGAT
P E D G Q K L I S E L M S I Y D T N I Q G N K N
AACTTAAACCTGTTAGTGGCAGGAGCAATTAATCAACCAACAGCAACAGCAACAGCAACTTCAACCA
N L N T G N S A T T E N Q N L P Q L K I Q P Q Q
GAGCAATCTCCGCAATCAAGTCTAATGATCGAGGAACTCTTGTGAGTGGATGATGATGATGATGATG
D D Y F R N Q G L M I E G N F F D G S N V S S S
CATCAATCTTCAACAAGATGAGTCAATTCAGCAGATTAAAGCAATGATACCCCTTTGAAACAC
H Q M F T R D E G Q F D R F K P M N T P F E N N
CACCATGCCACAGACAGCAACAAATACAAATACAAACAGCAACAGCAACAGCAACTTCACTGATGTT
H H H H N N N N N N N N N N N N N N N N F H L M F
AGCTTCCGTTGATTGTGGACTTCCACTACAAGAGAGATGATCCGGGTAGCAGCCATGACACCTG
S S P P D L S T P D Y K B E V S G V A A I D T L
TCGAAACAGAGGATATTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
S K Q Q D I P L W Y H *
TTCCGCAAAAGCAGCAGCAACATCAGCTGACAGTGAATTCATAGCAGTCTTTTACAGTTTGG
TATCTCTCTCTCTTATGGGTTTGGCTTTTACTAGGAAACAGTCAAGTTACTTTTGGGGTT
CCCTAAGTGGTGTATGTTACTGAAAGAAATGTTGGTGTACTACTACTTGGTCAATGAAACAGT
GATGATAATGCTATTAGCTTGTGAATATATCTTAAAGTCAATGATGATGATGATGATGATGATGATG
CAATGACTGATAAATAGTGGAGAGATCTTTAGTACTCTACTTACTCTGCTGCTGCTGCTGCTGCTG
TCCAATGATTTCTATATGAAATACTATTTCAACAAAAAATAAAAAAAAAAAAAAAAAAAAAA
    
```

Figure 1. Full-length cDNA and deduced amino acid sequences of *CsEIN3*. The start codon (ATG) is shaded. The stop codon (TAA) is also shaded and marked by an asterisk. The putative polyadenylation signal (AAATAA) is underlined.

Multiple sequence alignment and phylogenetic analysis

Multiple amino acid sequence alignments showed that *CsEIN3* had very high homology with its counterparts from other plant species (Figure 2). *CsEIN3* had identities of 67, 64, 70, and 62% to *NtEIL1*, *SlEIL1*, *VrEIL2*, and *AtEIN3*, respectively, indicating that the protein structure and functional manner were strongly conserved. Sequence alignment revealed that there was high homology in the N-terminus, including the DNA-binding domains that consisted of five conserved α -helices. The lysine 245 (K) located in α -helices-3, the key amino acid residue responsible for the DNA-binding and signal transduction activities, was found in *CsEIN3* (Yamasaki et al., 2005).

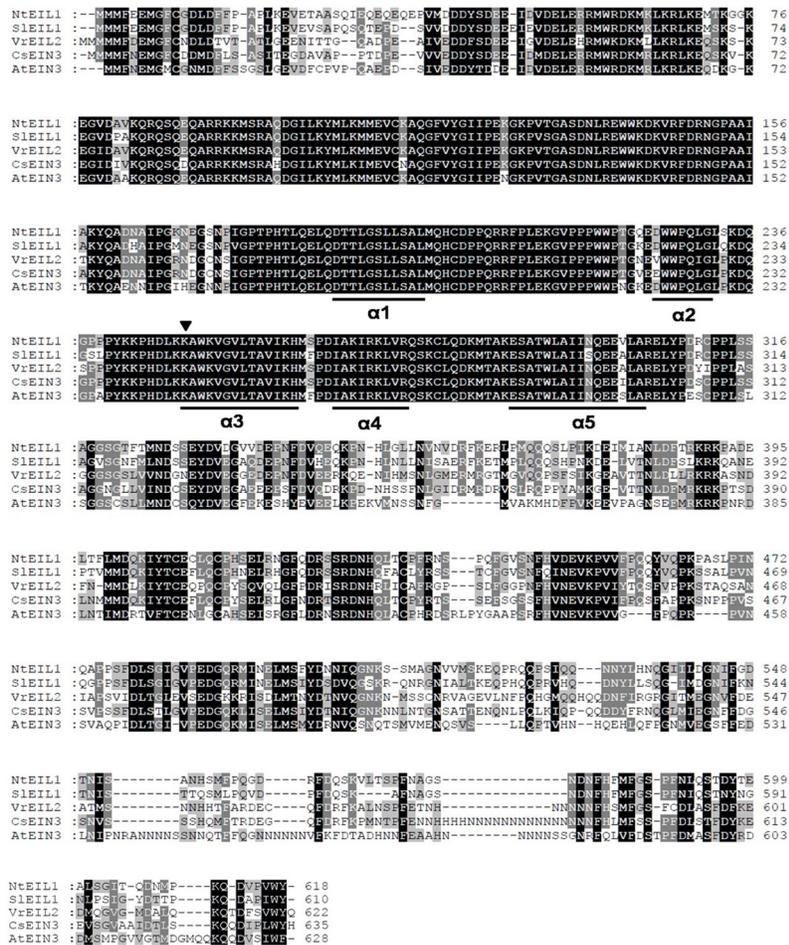


Figure 2. Multiple alignments of deduced amino acid sequence of *CsEIN3* with other known plant EIN3/EIL genes. Five EIN3/EIL genes from *Nicotiana tabacum* (*NtEIL1*, AAP03997), *Solanum lycopersicum* (*SlEIL1*, AAK58857), *Vigna radiata* (*VrEIL2*, AAL76271), *Cucumis sativus* (*CsEIN3*, AFK80347), *Arabidopsis thaliana* (*AtEIN3*, AEE76421) were used for multiple sequence alignment. Five α -helices are underlined, and the key amino acid residue responsible for the DNA-binding and signal transduction activities are indicated with black triangles.

To understand the evolutionary relationships among *CsEIN3* and other plant *EIN3/EILs*, a phylogenetic tree was constructed based on the amino acid sequences of plant *EIN3/EILs* (Figure 3). It was revealed that plant *EIN3/EILs* were derived from a common ancestor during evolution, and *CsEIN3* was grouped into a cluster along with melon, belonging to the Cucurbitaceae and paralleling their evolutionary relationships.

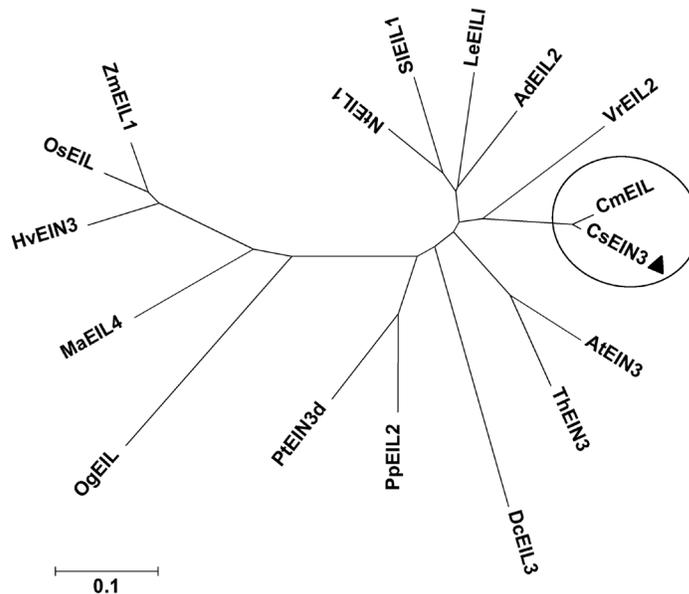


Figure 3. Phylogenetic relationships of known EIN3/EILs in plants. Amino acid sequences used for alignment include CmEIL (*Cucumis melo*, BAB64344), VrEIL2 (*Vigna radiata*, AAL76271), NtEIL1 (*Nicotiana tabacum*, AAP03997), SIEIL1 (*Solanum lycopersicum*, AAK58857), AdEIL2 (*Actinidia deliciosa*, ACJ70675), LeEIL1 (*Lithospermum erythrorhizon*, ACP56697), AtEIN3 (*Arabidopsis thaliana*, AEE76421), PpEIL2 (*Prunus persica*, ABK35086), DcEIL3 (*Dianthus caryophyllus*, AAV68141), MaEIL4 (*Musa acuminata*, ABG89103), OsEIL (*Oryza sativa*, BAB78462), ZmEIL1 (*Zea mays*, NP_001152035), OgEIL (*Oncidium Gower Ramsey*, AEK84144), PtEIN3d (*Populus trichocarpa*, XP_002310961), ThEIN3 (*Theilungiella halophila*, BAJ33951), HveIN3 (*Hordeum vulgare*, BAJ95094). The scale bar represents the estimated evolutionary distance as 0.1 amino acid substitutions per site. *CsEIN3* (indicated with black triangles) is grouped together with counterparts from melon (CmEIL), as indicated by the ellipse.

3-D modeling of *CsEIN3* protein

To better understand the protein character of *CsEIN3*, a 3-D model of the *CsEIN3* protein was established by SWISS-MODEL using the homology-modeling method. The template for modeling was the crystal structure of the EIL3 protein from *Arabidopsis thaliana* (PDB: 1wija; Kazuhiko et al, 2005). We obtained a 3-D structure of the *CsEIN3* protein fragment (178 to 301 aa) containing the DNA-binding domain. The result showed that five α -helices formed a V-shaped cleft, and the key lysine 245 (Lys245) responsible for the DNA-binding and signal transduction activities was located at a corresponding position similar to the template of AtEIL3 (Figure 4).

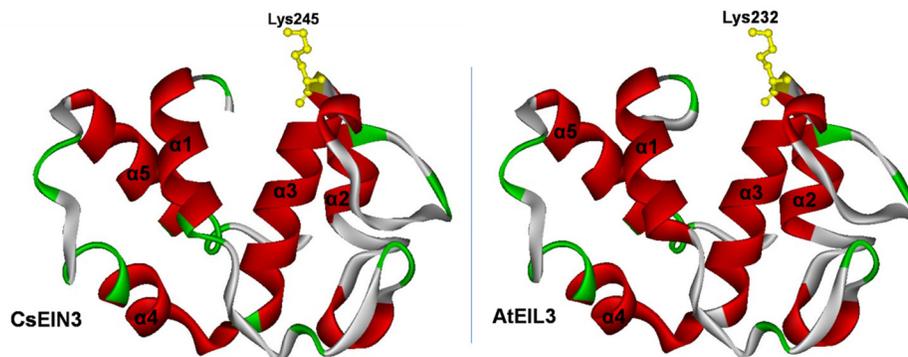


Figure 4. Three-dimensional structure of *CsEIN3* fragment (178-301 aa) contains DNA-binding domain (DBD). Five α -helices ($\alpha 1$ - $\alpha 5$) are indicated by red helices, the key amino acid residue (lysine) responsible for the DNA-binding and signal transduction activities are indicated with yellow ball.

Expression patterns of *CsEIN3* in different organs

To help us elucidate the role of *CsEIN3* in cucumber, the gene expression was analyzed in different plant organs by quantitative real-time RT-PCR, including roots, stems, leaves, shoot apices, and mature male and mature female flowers. The results showed that *CsEIN3* was expressed in all the organs tested. The highest expression of *CsEIN3* was detected in mature male flowers, mature female flowers, and stems. With regard to mature flowers, the expression level of *CsEIN3* in male flowers was higher than in female flowers (Figure 5).

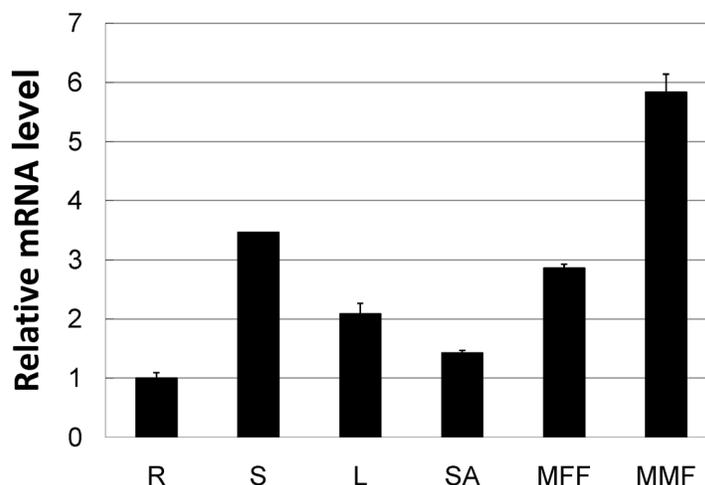


Figure 5. Expression analysis of *CsEIN3* in different organs of cucumber. Three biological replicate experiments are reported as means \pm SD. R = roots; S = stem; L = leaves; SA = shoot apices; MMF = mature male flowers; MFF = mature female flowers.

Expression patterns of *CsEIN3* in different developmental stages of flowers

To further understand whether *CsEIN3* was indeed involved in flower development, we preliminarily explored the expression patterns of *CsEIN3* in six different stages (from MF1/FF1 to MMF/MFF) throughout the development of female and male flowers (Figure 6A shows FF1-FF5 in females and MF1-MF5 in males). The results showed that accumulation of *CsEIN3* mRNA increased during both male and female flowers development and maturation on the whole, although it declined slightly at stages FF4 and FF5 in female flowers. At the early stages of flower development (MF1/FF1 and MF2/FF2), the expression of *CsEIN3* was relatively low (Figure 6B).

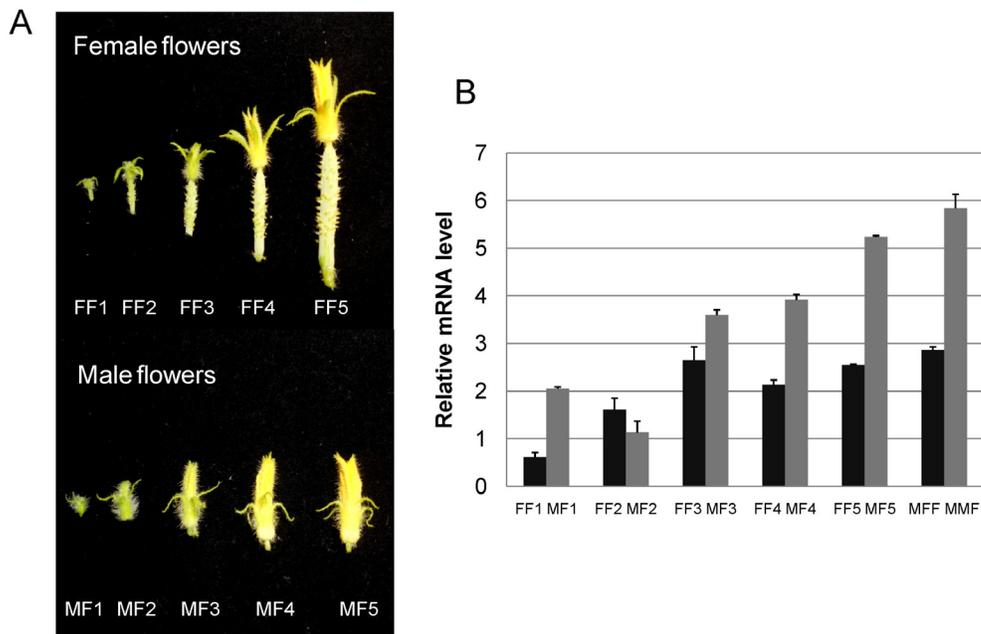


Figure 6. A. Morphology of male (MF) and female (FF) flowers at developmental stages, which were separated on the basis of the corolla length. MF1/FF1 = 2.5 ± 1 mm, MF2/FF2 = 5 ± 1 mm, MF3/FF3 = 9 ± 2 mm, MF4/FF4 = 15 ± 2 mm, MF5/FF5 = 20 ± 2 mm. B. Relative expression of *CsEIN3* in FF and MF at different developmental stages. The three biological replicate experiments are reported as means \pm SD.

DISCUSSION

Cucumber is considered a model plant for studying flower development, particularly the ethylene promotion of female flowering during unisexual flower development. Therefore, the isolation and characterization of relevant components are the foundation for study of the molecular mechanism of the cucumber ethylene signaling pathway. EIN3/EIL proteins, as key components in the ethylene signaling pathway, have still not been isolated from cucumber, thus hampering the study on their biological functions in cucumber.

In the present paper, we first reported the full-length cDNA sequence of the *EIN3* gene

in cucumber. The ORF of *CsEIN3* was 1908 bp in length, encoding a putative polypeptide of 635 amino acids. Initially, our experimental results were not in accordance with the tentative consensus sequence of csa012885 from the Cucumber Genome Database, of which the ORF was 1395 bp in length and encoding 464 amino acids. In order to clarify this inconsistent result, we aligned these two sequences and found that csa012885 lacked two fragments. To examine whether this situation was caused by a gap of genome sequencing, we next downloaded the genomic sequence corresponding to csa012885 from the database and amplified the genomic sequence of *CsEIN3*, respectively. Interestingly, the genomic sequence of *CsEIN3* did not contain an intron of which the length was also 1908 bp, like the cDNA sequence (data not shown). Based on this clue, we obtained a putative ORF of 622 amino acids, using the genomic sequence of csa012885, which was 1869 bp in length by ORF prediction; obviously, it also did not contain the intron. Subsequently, we aligned these two amino acid sequences and found that they had an identical sequence, except that there were serial 15 asparagines (Asn, N) at the C-terminus, whereas there were merely 3 asparagines in csa012885. Furthermore, we aligned the 1395 and 1869 bp sequences and found that the differential sequence conformed to the "GT-AG" rule. Therefore, the inaccuracy for the csa012885 sequence provided by the Cucumber Genome Database is most likely caused by the gene prediction algorithm that identified this fragment as an intron, which does not appear in csa012885. As for the number of serial asparagines (N) at the C-terminus, it is probably due to differences in cultivars. It is noteworthy that *CmEIL1*, which is the counterpart of *CsEIN3* in melon, also lacked the intron when we compared its genomic sequence and mRNA sequence (data not shown).

Bioinformatics analysis showed that *CsEIN3* possessed properties of *EIN3/EILs* by virtue of its highly conserved amino acid sequence and the strongly similar structure of the functional domain when compared with other plant *EIN3/EILs*. Specifically, *CsEIN3* showed high homology with its counterpart (*CmEIL1*) in melon and was grouped into the same cluster together with melon, which belongs to the same genus as cucumber. Subcellular localization prediction by WoLF PSORT (Horton et al., 2007) indicated that *CsEIN3* was localized in the nucleus (data not shown). This result was consistent with *EIN3/EILs* being nuclear transcription factors.

The previous studies have indicated that *EIN3/EILs* are expressed in almost all examined tissues; however, the expression pattern showed high diversity among different *EIN3/EIL* family members. For instance, the tomato *EIN3/EIL* family consists of four members (*LeEIL1*, *LeEIL2*, *LeEIL3*, and *LeEIL4*), all four of which were detected in all tissues examined. Interestingly, a high expression level of *LeEIL4* was detected in fruits and increased during fruit ripening, whereas the other three members were highly expressed in roots, stems, and flowers (Yokotani et al., 2003). In this study, we found that *CsEIN3* was expressed in all the plant organs examined including roots, stems, leaves, shoot apices, and male and female flowers, although the expression level was somewhat higher in flowers and stems than in the other organs. It is noteworthy that the accumulation of *CsEIN3* mRNA increased during flower development and maturation in both male and female flowers. This increase might be involved in the regulation of flower development and maturation in response to ethylene, consistent with a previous report (Iordachescu and Verlinden, 2005).

In summary, we have successfully cloned the full-length cDNA of the *EIN3* gene from cucumber and investigated its expression pattern in different plant organs and different developmental stages of flowers. This study will lay the foundation for further investigation into the role of *EIN3/EIL* protein factors in relevant biological processes of cucumber and the molecular mechanism of the cucumber ethylene signaling pathway.

ACKNOWLEDGMENTS

Research supported by the National Program on Key Basic Research Projects (The “973” Program: #2012CB113900) and the National Natural Science Foundation of China (#31071081 and #31272185).

REFERENCES

- An F, Zhao Q, Ji Y, Li W, et al. (2010). Ethylene-induced stabilization of ETHYLENE INSENSITIVE3 and EIN3-LIKE1 is mediated by proteasomal degradation of EIN3 binding F-box 1 and 2 that requires EIN2 in *Arabidopsis*. *Plant Cell* 22: 2384-2401.
- An F, Zhang X, Zhu Z, Ji Y, et al. (2012). Coordinated regulation of apical hook development by gibberellins and ethylene in etiolated *Arabidopsis* seedlings. *Cell Res.* 22: 915-927.
- Ando S, Sato Y, Kamachi S and Sakai S (2001). Isolation of a MADS-box gene (ERAF17) and correlation of its expression with the induction of formation of female flowers by ethylene in cucumber plants (*Cucumis sativus* L.). *Planta* 213: 943-952.
- Arnold K, Bordoli L, Kopp J and Schwede T (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22: 195-201.
- Bauer P and Blondet E (2011). Transcriptome analysis of ein3 eil1 mutants in response to iron deficiency. *Plant Signal. Behav.* 6: 1669-1671.
- Cao WH, Liu J, He XJ, Mu RL, et al. (2007). Modulation of ethylene responses affects plant salt-stress responses. *Plant Physiol.* 143: 707-719.
- Chao Q, Rothenberg M, Solano R, Roman G, et al. (1997). Activation of the ethylene gas response pathway in *Arabidopsis* by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. *Cell* 89: 1133-1144.
- Chen G, Alexander L and Grierson D (2004). Constitutive expression of EIL-like transcription factor partially restores ripening in the ethylene-insensitive Nr tomato mutant. *J. Exp. Bot.* 55: 1491-1497.
- Chen H, Xue L, Chintamanani S, Germain H, et al. (2009). ETHYLENE INSENSITIVE3 and ETHYLENE INSENSITIVE3-LIKE1 repress SALICYLIC ACID INDUCTION DEFICIENT2 expression to negatively regulate plant innate immunity in *Arabidopsis*. *Plant Cell* 21: 2527-2540.
- De Paepe A, Vuylsteke M, Van Hummelen P, Zabeau M, et al. (2004). Transcriptional profiling by cDNA-AFLP and microarray analysis reveals novel insights into the early response to ethylene in *Arabidopsis*. *Plant J.* 39: 537-559.
- Guo H and Ecker JR (2004). The ethylene signaling pathway: new insights. *Curr. Opin. Plant Biol.* 7: 40-49.
- Hibi T, Kosugi S, Iwai T, Kawata M, et al. (2007). Involvement of EIN3 homologues in basic PR gene expression and flower development in tobacco plants. *J. Exp. Bot.* 58: 3671-3678.
- Horton P, Park KJ, Obayashi T, Fujita N, et al. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 35: W585-W587.
- Huang S, Li R, Zhang Z, Li L, et al. (2009). The genome of the cucumber, *Cucumis sativus* L. *Nat. Genet.* 41: 1275-1281.
- Iordachescu M and Verlinden S (2005). Transcriptional regulation of three EIN3-like genes of carnation (*Dianthus caryophyllus* L. cv. Improved White Sim) during flower development and upon wounding, pollination, and ethylene exposure. *J. Exp. Bot.* 56: 2011-2018.
- Knopf RR and Trebitsh T (2006). The female-specific Cs-ACS1G gene of cucumber. A case of gene duplication and recombination between the non-sex-specific 1-aminocyclopropane-1-carboxylate synthase gene and a branched-chain amino acid transaminase gene. *Plant Cell Physiol.* 47: 1217-1228.
- Lee JH and Kim WT (2003). Molecular and biochemical characterization of VR-EILs encoding mung bean ETHYLENE INSENSITIVE3-LIKE proteins. *Plant Physiol.* 132: 1475-1488.
- Li Z, Huang S, Liu S, Pan J, et al. (2009). Molecular isolation of the M gene suggests that a conserved-residue conversion induces the formation of bisexual flowers in cucumber plants. *Genetics* 182: 1381-1385.
- Lin Z, Zhong S and Grierson D (2009). Recent advances in ethylene research. *J. Exp. Bot.* 60: 3311-3336.
- Liu S, Xu L, Jia Z, Xu Y, et al. (2008). Genetic association of ETHYLENE-INSENSITIVE3-like sequence with the sex-determining M locus in cucumber (*Cucumis sativus* L.). *Theor. Appl. Genet.* 117: 927-933.
- Livak KJ and Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25: 402-408.
- Maruyama-Nakashita A, Nakamura Y, Tohge T, Saito K, et al. (2006). *Arabidopsis* SLIM1 is a central transcriptional regulator of plant sulfur response and metabolism. *Plant Cell* 18: 3235-3251.

- Neljubov DN (1901). Über die horizontale Nutation der Stengel von *Pisum sativum* und einiger anderen Pflanzen. *Beih. Bot. Centralbh.* 10: 129-139.
- Parra-Lobato MC and Gomez-Jimenez MC (2011). Polyamine-induced modulation of genes involved in ethylene biosynthesis and signalling pathways and nitric oxide production during olive mature fruit abscission. *J. Exp. Bot.* 62: 4447-4465.
- Rudich J, Halevy AH and Kedar N (1972). Ethylene evolution from cucumber plants as related to sex expression. *Plant Physiol.* 49: 998-999.
- Shi Y, Tian S, Hou L, Huang X, et al. (2012). Ethylene signaling negatively regulates freezing tolerance by repressing expression of CBF and type-A ARR genes in *Arabidopsis*. *Plant Cell* 24: 2578-2595.
- Tieman DM, Ciardi JA, Taylor MG and Klee HJ (2001). Members of the tomato LeEIL (EIN3-like) gene family are functionally redundant and regulate ethylene responses throughout plant development. *Plant J.* 26: 47-58.
- Yamasaki K, Kigawa T, Inoue M, Yamasaki T, et al. (2005). Solution structure of the major DNA-binding domain of *Arabidopsis thaliana* ethylene-insensitive3-like3. *J. Mol. Biol.* 348: 253-264.
- Yamasaki S, Fujii N and Takahashi H (2000). The ethylene-regulated expression of CS-ETR2 and CS-ERS genes in cucumber plants and their possible involvement with sex expression in flowers. *Plant Cell Physiol.* 41: 608-616.
- Yanagisawa S, Yoo SD and Sheen J (2003). Differential regulation of EIN3 stability by glucose and ethylene signalling in plants. *Nature* 425: 521-525.
- Yokotani N, Tamura S, Nakano R, Inaba A, et al. (2003). Characterization of a novel tomato EIN3-like gene (LeEIL4). *J. Exp. Bot.* 54: 2775-2776.
- Zhang L, Li Z, Quan R, Li G, et al. (2011). An AP2 domain-containing gene, ESE1, targeted by the ethylene signaling component EIN3 is important for the salt response in *Arabidopsis*. *Plant Physiol.* 157: 854-865.
- Zhu Z, An F, Feng Y, Li P, et al. (2011). Derepression of ethylene-stabilized transcription factors (EIN3/EIL1) mediates jasmonate and ethylene signaling synergy in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 108: 12539-12544.