GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Genome-wide identification, phylogeny and expression analysis of the lipoxygenase gene family in cucumber

S.Q. Liu[1], X.H. Liu[2] and L.W. Jiang[1]

[1]School of Sciences, Jiangxi Agricultural University, Nanchang, China
[2]Hunan Vegetable Research Institute, Hunan Academy of Agricultural Sciences, Changsha, China

Corresponding author: S.Q. Liu
E-mail: lsq_hn306@163.com

Genet. Mol. Res. 10 (4): 2613-2636 (2011)
Received January 5, 2011
Accepted September 14, 2011
Published October 25, 2011
DOI http://dx.doi.org/10.4238/2011.October.25.9

ABSTRACT. Plant lipoxygenase (LOX) is involved in growth and developmental control processes, through the biosynthesis of regulatory molecules and defense responses to pathogens, wounding and stress. To date, few LOX proteins and little tissue expression profiling have been reported in detail for cucumber (*Cucumis sativus* L.). Recent completion of the cucumber genome sequence now permits genome-wide analysis of the LOX gene family in cucumber as well as comparison with LOX in *Arabidopsis* and rice. We identified 23 candidate *LOX* genes in the cucumber genome; phylogenetic analysis indicated that these *LOX* members cluster into two groups, designated types 1 and 2, as expected from previous studies. Sequence analysis showed that five binding sites of iron, including two consensus histidines in the LOX domain, are highly conserved in the cucumber LOX proteins. Analysis of chromosomal localization and genome distribution suggested that tandem duplication and/or polyploidal duplication contributed to the expansion of the cucumber LOX gene family. Based on intron/exon structure analysis, only a few of the extant intron patterns existed in the

ancestor of monocots and eudicots. Expression data showed widespread distribution of the cucumber LOX gene family within plant tissues, suggesting that they perform different functions in different tissues.

Key words: *Cucumis sativus* L.; Lipoxygenase; Phylogenetic analysis

INTRODUCTION

Lipoxygenases (LOXs; EC 1.13.11.12) are non-heme iron-containing dioxygenases that are widely distributed in plants and animals. LOXs catalyze the regio- and stereo-specific dioxygenation of PUFAs containing a (1*Z*,4*Z*)-pentadiene system to yield the corresponding (1*S*,2*E*,4*Z*)-hydroperoxides (Liavonchanka and Feussner, 2006). Plant LOXs catalyze the incorporation of molecular oxygen into free polyunsaturated fatty acids, at either position 9 or 13, in a stereospecific manner and, therefore, are referred to as linoleate 9-LOXs or 13-LOXs (Howe and Schilmiller, 2002). According to the classification based on their primary structure and overall sequence similarity, plant LOXs can be grouped into two gene subfamilies. Those enzymes harboring no plastidic transit peptides have a high sequence similarity (>75%) to one another and are designated type 1-LOXs. However, another subset of LOXs carry a plastidic transit peptide sequence; these enzymes show only a moderate overall sequence similarity (~35%) to one another and have been classified as type 2-LOXs. The LOX reaction usually uses linoleic (18:2) and linolenic acids (18:3) as substrates, resulting in the formation of a series of biologically active molecules, collectively called oxylipins (Grechkin et al., 1991), which can be further converted to different compounds through the action of enzymes participating in at least seven pathways (Feussner and Wasternack, 2002).

LOX and the products of the LOX pathway were involved in a series of biological events such as seed germination (Feussner et al., 2001; Rudolph et al., 2010), tuber development (Kolomiets et al., 2001), sex determination (Acosta et al., 2009), and fruit ripening (Chen et al., 2004; Zhang et al., 2006). Cucumber CsLOX1 is capable of adding oxygen to the sterified fatty acids, thus generating triacylglycerol containing one, two or three 13-HPOD acid residues that play a specialized role in fat mobilization in lipid bodies during seed germination (Feussner et al., 2001). The potato *POTLX-1* gene is involved in the control of tuber growth and development. RNA hybridization analysis showed that *POTLX-1* mRNA was observed in the distal, most actively growing portion of the developing tuber, and the antisense *POTLX-1* plants displayed reduced LOX activity and a several-fold reduction in tuber yield as well as the decreased average tuber size and a disruption of tuber formation (Kolomiets et al., 2001). LOX activity involved in fruit ripening was reported in tomato and kiwifruit as well as apple (Defilippi et al., 2005) and strawberry (Perez et al., 1999). In tomato, three *LOX* genes, *TomloxA*, *TomloxB* and *TomloxC*, are active during fruit ripening; in kiwifruit, *AdLox1* and *AdLox5* are active in ripening and responsive to ethylene and likely involved in LOX-based metabolism in kiwifruit ripening (Chen et al., 2004; Zhang et al., 2006).

Another important role of plant LOX enzymes is an involvement in defense responses to wounding (Bell et al., 1995; Wang et al., 2008) and pathogens (Gao et al., 2008). *Arabidopsis* plants with co-suppressed expression of the nuclear gene coding for the chloroplastic *AtLOX2* have diminished levels of jasmonic acid (JA) and reduced expression levels of the wound-induced *vsp* gene (Bell et al., 1995). In rice, *OsLOX1* transcripts are detected at low abundance

in immature seeds and newly germinated seedlings, but accumulate rapidly and transiently in response to wounding with a peak 3 h after wounding (Wang et al., 2008). In maize, a 9-lipoxygenase ZmLOX3 controls the root-specific expression of defense genes and its activity is required for normal levels of resistance to root-knot Nematodes (Gao et al., 2008).

Cucumber (*Cucumis sativus* L.) is an economically and nutritionally important vegetable crop cultivated worldwide and belongs to the Cucurbitaceae family. Previous studies have shown that there are 23 putative *LOX* genes in the cucumber genome (Huang et al., 2009). However, only two cucumber *LOX* genes, *CsLOX1* and *CsLOX9* (corresponding to *CsLOX01* and *CsLOX09* in our data, respectively), have been characterized up to now. The expression profiling of the cucumber lipoxygenase gene family and phylogenetic relationship with other plant *LOX* members in detail remain poorly understood (Matsui et al., 1999, 2006). In this report, we carry out the genome-wide identification, and phylogenetic and expression analysis of the lipoxygenase gene family in cucumber, as well as a comparative analysis with *Arabidopsis* and rice LOXs, which will be useful in conducting future functional genomic studies to understand how the molecular roles of the lipoxygenase gene family translate into a diversity of biological functions.

MATERIAL AND METHODS

Database search for cucumber *LOX* genes

A cucumber lipoxygenase sequence (GeneBank No. U25058) as a query sequence and the National Center for Biotechnology Information (NCBI) program TBLASTN (Altschul et al., 1997) were utilized to search for lipoxygenase genes encoded in the Cucumber Genome Initiative (CuGI, http://cucumber.genomics.org.cn), which was released by the Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences (IVF-CAAS). The default parameters with the TBLASTN program, wordsize 2 and extension 11, were used to obtain sequences that were as similar as possible. The redundant sequences with the same chromosome locus were removed from our data set. In addition, we have also obtained the same sequences using Hidden Markov Model (HMM) analysis with the Pfam number PF00305 containing typical LOX domain from the Pfam HMM library (http://pfam.jouy.inra.fr/).

Based on the results of both BLASTN searches in the cucumber genome and annotation database CuGI, we obtained the information of the chromosome locations and intron distribution pattern of predicted genes.

To further confirm the accuracy of these genes, the cDNA sequences were translated into amino acid sequences, which were then searched for in the conserved domain using the Simple Modular Architecture Research Tool (SMART) (Letunic et al., 2004).

Multiple sequences alignments and conserved motif prediction

Multiple sequence alignments using Clustal X (http://www.clustal.org/) (Larkin et al., 2007) with the default parameters were performed on the CsLOX protein sequences obtained and the alignments were then adjusted manually. The similar amino acids were highlighted using the GeneDoc tool (http://www.nrbsc.org/gfx/genedoc/) (Nicholas et al., 1997). We also used the Multalin software (http://multalin.toulouse.inra.fr/multalin/multalin.html) (Corpet,

1988) as a secondary method to align sequences and to recheck the result. In addition, to search for conserved motifs within the CsLOX members, we used the MEME tool (http://meme.sdsc.edu/meme/cgi-bin/meme.cgi) (Bailey and Elkan, 1994) to find similar sequences shared by CsLOX members.

Tree building

A phylogenetic tree was constructed with aligned CsLOX protein sequences using MEGA4 (http://www.megasoftware.net/mega4/mega.html) (Tamura et al., 2007), using the neighbor joining (NJ) method with the following parameters: Poission correction, pairwise deletion, and bootstrap (1000 replicates; random seed). Meanwhile, maximum parsimony (MP) method of the PHYLIP3.69 tool (http://evolution.genetics.washington.edu/phylip.html) (Felsenstein, 1989) was also utilized to create a second phylogenetic tree with a bootstrap of 1000 replicates to test and verify the results from the NJ method. The constructed tree file was visualized by TreeView1.6.6 (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html) (Page, 1996).

Intron/exon structure, genome distribution and segmental duplication

To analyze the intron/exon structure, we downloaded the DNA and cDNA sequences corresponding to each predicted gene from the cucumber genome and annotation database CuGI and then analyzed the intron distribution pattern and splicing phase using the web-based bioinformatics tool GSDS (http://gsds.cbi.pku.edu.cn/).

To obtain information on *CsLOX* gene locations, a map with the distribution of *CsLOX* family members throughout the cucumber genome was drawn using the MapInspect tool (http://www.plantbreeding.wur.nl/UK/software_mapinspect.html). To detect the large segment duplicated events, we analyzed 100-kb DNA segments flanking each *CsLOX* gene. The region on different linkage groups containing six or more homologous pairs, which have fewer than 25 nonhomologous genes intervening, was defined as a duplicated segment. A gene pair was considered to be tandem duplicated if they were both separated by less than five intervening genes and shared ≥40% sequence similarity at the amino acid level. The BioEdit5.0.6 software (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) (Hall., 1999) was used to analyze the homologs for similarity on the phylogenetic tree of obtained *CsLOX* genes.

Expression analysis of cucumber *LOX* genes

We used semi-quantitative reverse transcription-polymerase chain reaction (RT-PCR) to detect the expression patterns of the *CsLOX* genes. The PCR primers were designed to avoid the conserved region and the information of primer sequences is shown in detail in Supplementary Table 1. The length of amplified products was about 150~450 bp. Seeds of the 'Chinese long' inbred line 9930, which is commonly used in modern cucumber breeding (Tanurdzic and Banks, 2004), were germinated and grown in trays containing a soil mixture (peat:sand:pumice, 1:1:1, v/v/v). Plants were adequately watered and grown at day/night temperatures of 24°/18°C with a 16-h photoperiod. Total RNA of root, stem, leaf, fruit, and flower of cucumber, at the stage when the number of main-stem nodes reaches 20, was isolated using the TRIzol Reagent (Invitrogen, USA) and RT-PCR was performed according to manufacturer

recommendations (Tiangen Biotech Co. Ltd., Beijing, China). Twenty-eight cycles of the PCR amplification were performed with no template as negative control, and *actin* DNA fragment (161 bp) of cucumber, as inner standard, were employed for each gene. To confirm its accuracy, all products of the expected size were sequenced (Sangon Biotech Co. Ltd., Shanghai, China). EST data came from NCBI (http://www.ncbi.nlm.nih.gov/dbEST/). An EST was considered to be corresponding to its gene if they shared ≥95% sequence similarity, the length of matching sequences ≥160 bp and E values $\leq 10^{-10}$.

RESULTS AND DISCUSSION

Identification of 23 *CsLOX* genes

To identify the *CsLOX* genes in the cucumber genome, we used criteria provided by Feussner and Wasternack (2002) to define a CsLOX protein. Initially, the LOX domain of a cucumber lipoxygenase (GeneBank No. U25058) containing a representative motif His-$(X)_4$-His-$(X)_4$-His-$(X)_{17}$-His-$(X)_8$-His, which represents a potential binding site for non-heme iron to the protein, was used as a BLAST query to identify all candidate *CsLOX* genes. Subsequently, we used the TBLASTN to remove redundant sequences of candidate *CsLOX* genes according to their corresponding chromosome locations and finally obtained 23 *CsLOX* genes (Supplementary Table 2A). The same 23 sequences were also identified using HMM analysis with PfamPF00305 containing typical LOX domain from the CuGI database. The number designation of these *CsLOX* genes was based on the order of the multiple sequence alignments (Supplementary Figure 1) and two previously reported genes, *CsLOX1* and *CsLOX9* (Matsui et al., 1999, 2006). To further verify the reliability of these candidate sequences, we performed SMART analysis on the 23 putative LOXs and found that all had a typical LOX domain. Compared with two previously confirmed members, *CsLOX1* and *CsLOX9* (Matsui et al., 1999, 2006), more lipoxygenase genes could be found in the cucumber genome. The discrepancy is probably due to the fact that no cucumber genome database was available previously.

At the same time, we identified 11 *OsLOX* genes (Supplementary Table 2B) in the version of release 5 of the TIGR pseudomolecules in rice (http://rice.plantbiology.msu.edu/) using the same methods.

Multiple sequence alignments and conserved motif prediction

To examine sequence features of these CsLOX members, we performed multiple sequence alignment of the 23 CsLOX amino acid sequences (Supplementary Figure 2). In agreement with a previous study (Schneider et al., 2007), we found the common LOX domain, which being located within the carboxy-terminal colipase binding domain, was more conserved than other parts in all CsLOX sequences, while the amino-terminal region of sequences was the most divergent region in terms of both length and composition of amino acids. In addition, we identified 158 amino acid residues that were at least 75% identical among the 23 CsLOX sequences (Supplementary Figure 2, indicated at the bottom of the alignments).

It was reported that a region rich in histidine residues was observed in the primary structure of the soybean L-1 and L-2 lipoxygenase isozymes (Shibata et al., 1987, 1988), and further investigation showed that His^{499}, His^{504}, His^{690}, Asn^{694}, and Ile^{839} were essential for bind-

ing of iron and Ile[839] as the C-terminal amino acid showing a function in binding as a monodentate ligand in the octahedral coordination sphere of the Fe iron (Steczko et al., 1992; Minor et al., 1993). In 23 CsLOX members, five such iron binding sites were also identified, and the first two His were included in the consensus LOX domain (the positions were indicated with asterisk in Supplementary Figure 3). The previous study showed that the replacement of the C-terminal isoleucine with valine resulted in an enzyme having substantial activity, whereas other replacements resulted in inactivation by modifying two murine lipoxygenase (Chen et al., 1994). Interestingly, there were 16 sequences that terminate in Ile among the 23 CsLOXs, whereas another 7 CsLOXs terminate in another amino acid, like Thr for CsLOX07, CsLOX15, and CsLOX16, which leads us to propose that these 7 CsLOXs likely show a difference in enzyme activities with the other CsLOXs.

Motif analysis showed that a representative motif of 38 amino acid residues, which was in accordance with the form His-$(X)_4$-His-$(X)_4$-His-$(X)_{17}$-His-$(X)_8$-His of a previous study (Zhang et al., 2006), was highly conserved, excluding 5 members, CsLOX03, CsLOX07, CsLOX11, CsLOX17, and CsLOX18, because of sequence truncation (Figure 1).

A

B

```
CsLOX05 : HQLISHWLNTHAVIEPFVIATNRQLSVLHPIHKLLIPH
CsLOX08 : HQLISHWLNTHATIEPFVIATNRQLSVVHPIHKLLVPH
CsLOX04 : HQLISHWLNTHAVLEPFVIATHRQLSVLHPIHKLLDPH
CsLOX02 : HQLISHWLHTHAVQEPFVIATHRQLSALHPIHKLLVPH
CsLOX03 : --------------MPFVIATHRQLSVLHPIHKLLVPH
CsLOX01 : HQLISHWLHTHAVLEPFVIATHRQLSVLHPIHKLLVPH
CsLOX10 : HQLISHWLNTHAVIEPFVIATNRQLSVLHPVFKLLHPH
CsLOX06 : HQVVSHWLNTHAVMEPFVIATNRQLSVLHPIHKLLVPH
CsLOX09 : HQLISHWLNTHAVIEPFIIGTNRQLSVVHPIYKLLHPH
CsLOX11 : ------------------------------------
CsLOX22 : HQLVNHWLRTHATLEPFILAAHRHLSAMHPIFKLLDPH
CsLOX23 : HQLVNHWLRTHASMEPYIIATHRQLSSMHPIYKLLHPH
CsLOX18 : --------------------------AMHPMYRLLHPH
CsLOX19 : HQLVSHWLRTHCCVEPYIIAANRQLSAMHPIYRLLHPH
CsLOX20 : HQLVSHWLSTHCAVEPYVIATNRQLSAMHPIYRLLQPH
CsLOX21 : HQLVSHWLRTHCVVEPYVIATNRQLSAMHPIYRLLHPH
CsLOX07 : ----------------------------------RLLHPH
CsLOX17 : ------------------------------------
CsLOX16 : HQLVVHWLRTHCCMEPYAIALNRQLSTMHPIYRLLHPH
CsLOX12 : HQLVIHWLRAHCCMESYAIATNRQLSTMHPIYRLLHPH
CsLOX13 : HQLVIHWLRAHCCMEPYAIATNRQLSTMHPIYRLLHPH
CsLOX14 : HQLVIHWLRAHCCMEPYAIATHRQLSTMHPIYRLLHPH
CsLOX15 : HQLVIHWLRAHCCMEPYAIATNRQLSTVHPIYRLLHPH
```

Figure 1. A 38-residue motif among cucumber LOX sequences. **A.** The logo was created with 23 cucumber LOX sequences. Overall height in each stack indicates the sequence conservation at that position and height of each residue letter indicates relative frequency of the corresponding amino acid residue. Below the logo is the consensus sequence. The conserved histidines (H) are marked with boldface letters. **B.** Sequence alignments of 38-residue motifs in cucumber LOX members.

Phylogenetic analysis of the *CsLOX* genes

To determine the evolutionary relationships of cucumber *LOXs* with those from *Arabidopsis* and rice, an unrooted NJ phylogenetic tree with bootstrap analysis (1000 replicates) was generated by the multiple sequence alignments of 40 LOX members including 23 CsLOX, 11 OsLOX and 6 reported *Arabidopsis* AtLOX members (Figure 2) (Bannenberg et al., 2009). The analysis result showed that the 40 LOX members were clustered into type 1 and type 2 groups according to clades with at least 50% bootstrap support, similar to the result described by Bannenberg et al. (2009) and Huang et al. (2009). There were nine cucumber *LOXs* (*CsLOX01~06*, *CsLOX08~10*) grouped in type 1 including another four rice *LOXs* (*OsLOX06~09*) and two *Arabidopsis*, *AtLOX01* and *AtLOX05*. Upon closer inspection, 7 *CsLOXs* (*CsLOX01~06*, *CsLOX08*) existed in the form of an exclusive cluster, indicating that these 7 *CsLOX* genes might diverge after the monocotyledon/dicotyledon separation (Figure 2). Two other members, *CsLOX09* and *CsLOX10*, were found close to *AtLOX05* and *AtLOX01*, respectively. The former study indicated that AtLOX05 and AtLOX01 belonged to 9*S*-lipoxygenases (Bannenberg et al., 2009), which led us to hypothesize that CsLOX09 and CsLOX10 likely were 9*S*-lipoxygenases. Similar to type 1, an exclusive cluster including 11 *CsLOX* members (*CsLOX07*, *CsLOX12~21*) existed in type 2, which was closed to *AtLOX2* and *OsLOX3*. *AtLOX2* was essential for jasmonic acid formation upon wounding and its expression was down-regulated during senescence (Melan et al., 1993; Bell et al., 1995; He et al., 2002), and *OsLOX3* was involved in herbivore-induced JA biosynthesis, which suggests that these 11 *CsLOXs* (*CsLOX07*, *CsLOX12~21*) genes may be participating in JA formation; the detailed function requiring further research and confirmation. Meanwhile, in type 2, *CsLOX22* was closely related to the cluster containing 2 *AtLOXs* (*AtLOX03* and *AtLOX04*), and *CsLOX23* related to *AtLOX06*. These 3 *AtLOX* gene encoding proteins, such as 13*S*-lipoxygenases, were reported to oxygenate linolenic acid more effectively than linoleic acid (Bannenberg et al., 2009).

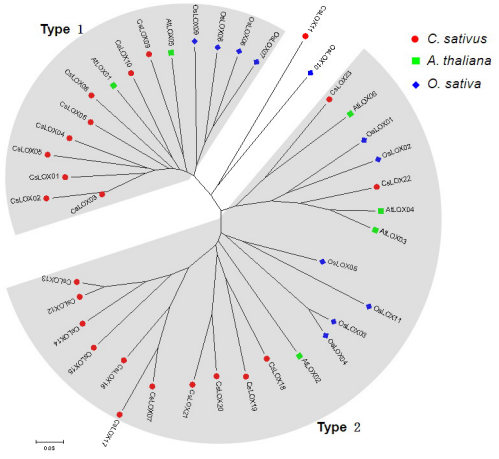


Figure 2. Phylogenetic analysis of cucumber LOXs with those from *Arabidopsis* and rice. An unrooted neighbor-joining phylogenetic tree was constructed using MEGA4 by the multiple sequence alignments of 40 LOX protein sequences including 23 CsLOX, 11 OsLOX and 6 reported *Arabidopsis* AtLOX members. Two classes are marked, type 1 and type 2, as described by Bannenberg et al. (2009) and Huang et al. (2009).

It was reported that type 2-LOXs just showed a moderate overall sequence similarity (~35%) to one another (Feussner and Wasternack, 2002); however, our analysis showed that the overall identity of the cucumber type 2-LOXs *CsLOX07* to *CsLOX17* reached 89%, obviously much higher than 35%. This discrepancy is probably because only a few *LOXs* have been compared previously.

In addition, MP analysis was also used to construct a phylogenetic tree (Supplementary Figure 4), which found that 83% of the *CsLOX* members were placed into the same locations as those in the NJ tree, indicating that the above NJ tree analysis result was reliable.

Structure and evolution of the *CsLOX* genes

The intron number of the 23 cucumber *LOX* genes ranged from zero to nine, only *CsLOX03* had no intron. Further analysis showed the presence of eight highly conserved intron positions (Supplementary Figure 5, indicated by black arrows), which can be observed in a total of 19 genes. According to the presence and positions of their introns, these 19 genes could be grouped into 11 patterns (1, 2, 5~7, 9~13, and 18). Among the 11 patterns, just pattern 1 was shared by the *LOX* genes from cucumber, *Arabidopsis* and rice. The percentages of pattern 1 in *CsLOX* and *AtLOX* members were quite close, 30% for *CsLOX* members and 33% for *AtLOX* members, whereas for *OsLOX* members, the percentage was just 9%. For the remaining 10 patterns, 8 patterns (2, 5~7, 9~11, and 13) with 80% high ratio were found just in cucumber *LOX* genes, another 2 patterns (12 and 18) were observed only in both cucumber and rice *LOX* genes, and no pattern exhibited in both cucumber and *Arabidopsis LOX* genes. The appearance of relatively few conserved patterns (1, 12 and 18) indicated that only a few of the intron patterns existed in the ancestor of monocots and eudicots.

To gain insight into the evolution of 23 *LOX* forms in cucumber, we analyzed their genomic distribution and found that they were unevenly distributed on chromosomes 2, 4~7, except for *CsLOX03* and *CsLOX04,* which lie within unassembled scaffold000337 (Figure 3; Supplementary Table 2A). On chromosome 2 and 4, a cluster with a relatively high density was observed, particularly for the cluster identified on the region close to centromeric on chromosome 4, where up to 11 *CsLOX* members (*CsLOX07*, *CsLOX12~CsLOX21*) were identified (Figure 3). Moreover, these 11 *CsLOX* members have a high sequence similarity with each other. For instance, the entire protein sequences of CsLOX07 and CsLOX17 shared 89% identities, and CsLOX20 and CsLOX21 shared 86% identities. *CsLOX23* and *CsLOX09* were merely distributed on chromosome 5 and 6, respectively, and *CsLOX11* and *CsLOX22* localized on chromosome 7. The genomic distribution indicated that some *CsLOX* members of type 1 and type 2 were located within the same small chromosomal region, whereas other members of the same group were distributed in different chromosomal regions. A similar situation also appeared in the *OsLOX* and *AtLOX* members, indicating that the *LOX* genes were distributed widely in the genome of the common ancestor of monocots and eudicots.

On average, genome duplication events are thought to have occurred throughout the process of plant genome evolution (Cannon et al., 2004). Previous research has revealed that the event of recent whole-genome duplication is absent in the cucumber genome, but a number of tandem duplications have occurred (Huang et al., 2009), which defined the members both separated by less than five intervening genes and sharing ≥40% sequence similarity at the amino acid level. However, whether a segment duplicated event existed or not in the cucum-

ber LOX gene family was not clear. Therefore, we performed segment duplication analysis on 100-kb DNA segments flanking each of *CsLOX* genes and found no gene contributed to segmental duplication. Hence, it was most likely that tandem duplication and/or polyploidy duplication have played a key role in the observed gene expansion of *LOX* genes in cucumber genome. There were 23 *LOX* genes distributed in the cucumber genome, an obviously higher number compared to that of other reported species (6 for *Arabidopsis* and 15 for *papaya*) (Huang et al., 2009), possibly resulting from the higher frequency of the tandem duplication and/or polyploidy duplication.

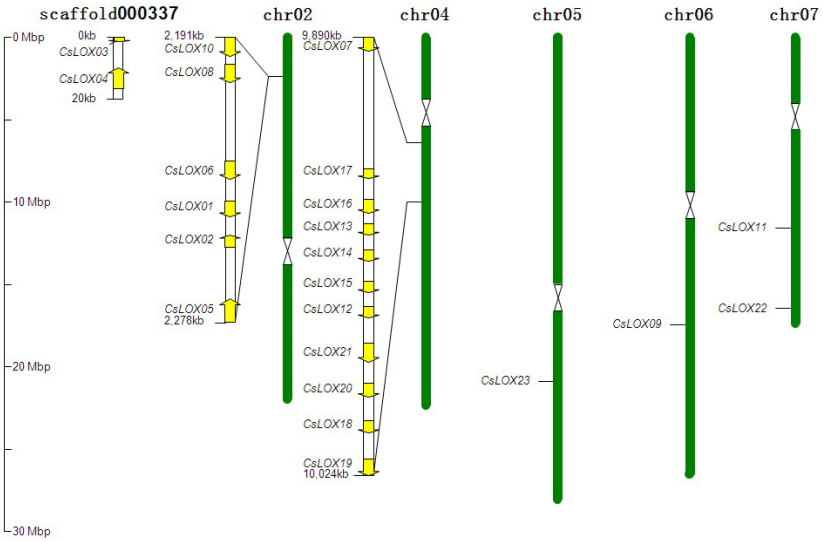


Figure 3. Chromosomal localization of 23 cucumber *LOX* genes. The scale is in megabases (Mb). The white funnels in the middle of the 5 chromosomes show the centromeric positions according to the sequencing result of the cucumber genome (Huang et al., 2009). Two genes, *CsLOX03* and *CsLOX04*, on the scaffold000337, cannot be anchored on a specific chromosome.

The expression pattern of the *CsLOX* genes

In view of the fact that gene expression pattern is often correlated with its function, we investigated the expression information of *CsLOX* members using RT-PCR analysis with RNA from root, stem, leaf, fruit, and flower as well as the EST data from NCBI. The RT-PCR result showed that a total of 13 *CsLOX* transcripts were detected in at least one tissue, 5 *CsLOX* genes (*CsLOX02*, *CsLOX10*, *CsLOX19*, *CsLOX22*, and *CsLOX23*) expressed in all tissues investigated, and the expression signal for *CsLOX22* was at a lower abundance. Three *CsLOX* genes (*CsLOX01*, *CsLOX06*, and *CsLOX20*) had similar expression patterns in cucumber stem, leaf, flower, and fruit, but *CsLOX01* appeared to have a higher transcript abundance in stem. *CsLOX17* exhibited an obvious transcript signal in cucumber root, leaf, flower, and fruit, but no expression signal was detected in the stem. For the remaining four genes, *CsLOX04*, *CsLOX08*, *CsLOX09*, and *CsLOX16*, the first 3 genes expressed in both flower and fruit while *CsLOX16* just expressed in the flower. In addition, *CsLOX08* and *CsLOX16* had an obvious

expression signal in root, and *CsLOX09* in leaf. No obvious *CsLOX04* expression signal was observed in the 3 investigated vegetative tissues. Subsequently, we searched the EST information on *CsLOX* genes from NCBI and found EST data for 6 *CsLOXs* (*CsLOX02*, *CsLOX03*, *CsLOX07*, *CsLOX10*, *CsLOX17*, and *CsLOX19*). Except for *CsLOX03* and *CsLOX07*, the other 4 genes had positive RT-PCR results (Figure 4). This discrepancy may be the result of weak expression that was barely detected by RT-PCR. After integrating, a total of 8 *CsLOX* genes (*CsLOX05*, *CsLOX11~15*, *CsLOX18* and *CsLOX21*) were not detectably expressed according to the results of RT-PCR and EST data, which suggested that these genes might be pseudogenes, or expressed at specific developmental stages or under special conditions. Another reason might be relatively little cucumber EST data available in NCBI.

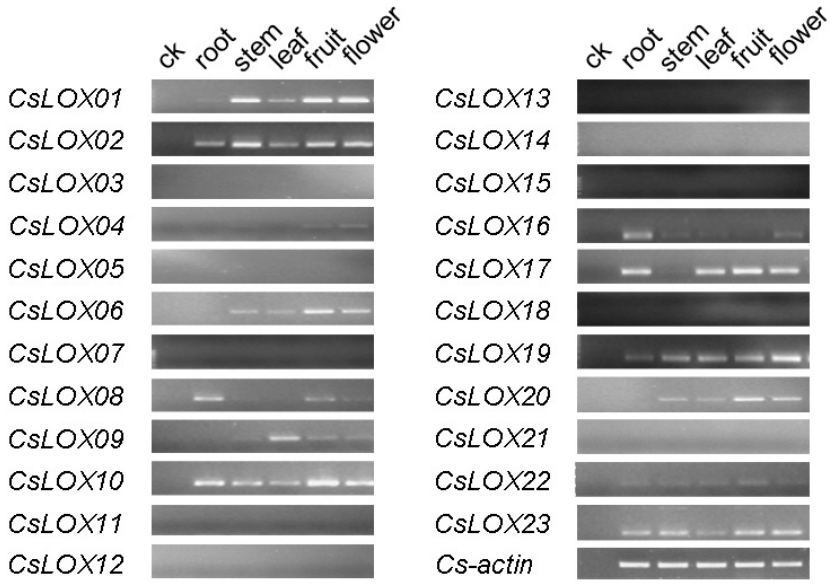


Figure 4. Expression analysis of 23 cucumber *LOX* genes in different tissues using RT-PCR. RT-PCR was performed using the primers specific to the 23 *CsLOX* genes. PCR products from 28 cycles were run on 1.5% agarose gels. *Cs-actin* primers giving a 161-bp product was used as inner standard for each gene. ck means a negative control with deionized water as template. The sample identities are as follows: root, stem, leaf, fruit, and flower.

Among the 15 expressed *CsLOX* genes, some genes such as *CsLOX02*, *CsLOX10*, *CsLOX19*, and *CsLOX23* were expressed in both vegetative and reproductive tissues, suggesting that these *CsLOX* genes played regulatory roles in different tissues in cucumber; however, some members have preferential expression that is tissue specific, including *CsLOX16* with expression in root and *CsLOX04* in reproductive tissues, suggesting that members of this family might take part in specific biological processes in cucumber.

Two cucumber *LOX* genes, *CsLOX01* and *CsLOX09*, have been characterized. *CsLOX1* (corresponding to *CsLOX01* in our data) was a major lipid body protein and played a specialized role in fat mobilization in lipid bodies during seeding growth (Matsui et al., 1999). The transcript of *CsLOX01* was detected in stem, leaf, fruit, and flower. Whether it also functioned

with a similar mechanism or not in these tissues deserves further research. For *CsLOX9* (corresponding to *CsLOX09* in our data), the transcript was detected in leaf, flower and fruit, similar to a previous study (Matsui et al., 2006). Matsui research suggested that CsLOX9 might provide HPL (hydroperoxide lyase) with fatty acid 9-hydroperoxides as substrates, and that the HPL activity was developmentally regulated and induced by mechanical wounding. In addition, HPL-derived product C9-aldehydes were formed rapidly after disruption of the tissues and showed fungicidal activities against the fungal pathogens *Botrytis cinerea* and *Fusarium oxysporum* (Matsui et al., 2006). Although the specific functions of most *CsLOX* genes are unknown, the expression analysis will provide a beneficial foundation for future functional studies.

In summary, we have performed extensive analyses of the 23 cucumber *LOX* genes and compared them with 6 *Arabidopsis* and 11 rice *LOX* genes. The 40 LOX members clustered into type 1 and type 2 groups, which were in general agreement with previous studies (Bannenberg et al., 2009; Huang et al., 2009). Chromosomal localization and genome distribution of *CsLOX* genes have revealed that tandem duplication and/or polyploidy duplication may be contributing to the expansion of *CsLOX* genes. Expression data showed widespread distribution of the *CsLOX* gene family within plant tissues, which will facilitate the selection of candidate *CsLOX* genes during plant development, and abiotic or biotic stress responses together with the result of phylogenetic analysis.

ACKNOWLEDGMENTS

The authors thank Sanwen Huang for help with expression analysis, Zhonghua Zhang for help with statistical analysis and Lifang Hu for useful comments and revision of the manuscript. Research supported by the National Natural Science Foundation of China (#31060262), the Natural Science Foundation of Jiangxi, China (#2009GQN0034) and the Foundation of Jiangxi Agricultural University, Jiangxi, China (#2009).

REFERENCES

Acosta IF, Laparra H, Romero SP, Schmelz E, et al. (2009). *Tasselseed1* is a lipoxygenase affecting jasmonic acid signaling in sex determination of maize. *Science* 323: 262-265.

Altschul SF, Madden TL, Schaffer AA, Zhang J, et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.

Bailey TL and Elkan C (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2: 28-36.

Bannenberg G, Martinez M, Hamberg M and Castresana C (2009). Diversity of the enzymatic activity in the lipoxygenase gene family of *Arabidopsis thaliana*. *Lipids* 44: 85-95.

Bell E, Creelman RA and Mullet JE (1995). A chloroplast lipoxygenase is required for wound-induced jasmonic acid accumulation in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 92: 8675-8679.

Cannon SB, Mitra A, Baumgarten A, Young ND, et al. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol.* 4: 10.

Chen G, Hackett R, Walker D, Taylor A, et al. (2004). Identification of a specific isoform of tomato lipoxygenase (TomloxC) involved in the generation of fatty acid-derived flavor compounds. *Plant Physiol.* 136: 2641-2651.

Chen XS, Kurre U, Jenkins NA, Copeland NG, et al. (1994). cDNA cloning, expression, mutagenesis of C-terminal isoleucine, genomic structure, and chromosomal localizations of murine 12-lipoxygenases. *J. Biol. Chem.* 269: 13979-13987.

Corpet F (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 16: 10881-10890.

Defilippi BG, Dandekar AM and Kader AA (2005). Relationship of ethylene biosynthesis to volatile production, related enzymes, and precursor availability in apple peel and flesh tissues. *J. Agric. Food Chem.* 53: 3133-3141.

Felsenstein J (1989). PHYLIP: Phylogeny Inference Package (Version 3.2). Cladistics, 164-166.
Feussner I and Wasternack C (2002). The lipoxygenase pathway. *Annu. Rev. Plant Biol.* 53: 275-297.
Feussner I, Kühn H and Wasternack C (2001). Lipoxygenase-dependent degradation of storage lipids. *Trends Plant Sci.* 6: 268-273.
Gao X, Starr J, Gobel C, Engelberth J, et al. (2008). Maize 9-lipoxygenase ZmLOX3 controls development, root-specific expression of defense genes, and resistance to root-knot nematodes. *Mol. Plant Microbe Interact.* 21: 98-109.
Grechkin AN, Kuramshin RA, Latypov SK, Safonova YY, et al. (1991). Hydroperoxides of alpha-ketols. Novel products of the plant lipoxygenase pathway. *Eur. J. Biochem.* 199: 451-457.
Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser. 41: 95-98.* 41: 95-98.
He Y, Fukushige H, Hildebrand DF and Gan S (2002). Evidence supporting a role of jasmonic acid in *Arabidopsis* leaf senescence. *Plant Physiol.* 128: 876-884.
Howe GA and Schilmiller AL (2002). Oxylipin metabolism in response to stress. *Curr. Opin. Plant Biol.* 5: 230-236.
Huang S, Li R, Zhang Z, Li L, et al. (2009). The genome of the cucumber, *Cucumis sativus* L. *Nat. Genet.* 41: 1275-1281.
Kolomiets MV, Hannapel DJ, Chen H, Tymeson M, et al. (2001). Lipoxygenase is involved in the control of potato tuber development. *Plant Cell* 13: 613-626.
Larkin MA, Blackshields G, Brown NP, Chenna R, et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.
Letunic I, Copley RR, Schmidt S, Ciccarelli FD, et al. (2004). SMART 4.0: towards genomic data integration. *Nucleic Acids Res.* 32: D142-D144.
Liavonchanka A and Feussner I (2006). Lipoxygenases: occurrence, functions and catalysis. *J. Plant Physiol.* 163: 348-357.
Matsui K, Hijiya K, Tabuchi Y and Kajiwara T (1999). Cucumber cotyledon lipoxygenase during postgerminative growth. Its expression and action on lipid bodies. *Plant Physiol.* 119: 1279-1288.
Matsui K, Minami A, Hornung E, Shibata H, et al. (2006). Biosynthesis of fatty acid derived aldehydes is induced upon mechanical wounding and its products show fungicidal activities in cucumber. *Phytochemistry* 67: 649-657.
Melan MA, Dong X, Endara ME, Davis KR, et al. (1993). An *Arabidopsis thaliana* lipoxygenase gene can be induced by pathogens, abscisic acid, and methyl jasmonate. *Plant Physiol.* 101: 441-450.
Minor W, Steczko J, Bolin JT, Otwinowski Z, et al. (1993). Crystallographic determination of the active site iron and its ligands in soybean lipoxygenase L-1. *Biochemistry* 32: 6320-6323.
Nicholas KB, Hugh BN Jr and David WD (1997). GeneDoc: analysis and visualization of genetic variation. *Embnew News* 4: 14.
Page RD (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12: 357-358.
Perez AG, Sanz C, Olias R and Olias JM (1999). Lipoxygenase and hydroperoxide lyase activities in ripening strawberry fruits. *J. Agric. Food Chem.* 47: 249-253.
Rudolph M, Schlereth A, Korner M, Feussner K et al. (2010). The lipoxygenase-dependent oxygenation of lipid body membranes is promoted by a patatin-type phospholipase in cucumber cotyledons. *J. Exp. Bot.* 62: 749-760.
Schneider C, Pratt DA, Porter NA and Brash AR (2007). Control of oxygenation in lipoxygenase and cyclooxygenase catalysis. *Chem. Biol.* 14: 473-488.
Shibata D, Steczko J, Dixon JE, Hermodson M, et al. (1987). Primary structure of soybean lipoxygenase-1. *J. Biol. Chem.* 262: 10080-10085.
Shibata D, Steczko J, Dixon JE, Andrews PC, et al. (1988). Primary structure of soybean lipoxygenase L-2. *J. Biol. Chem.* 263: 6816-6821.
Steczko J, Donoho GP, Clemens JC, Dixon JE, et al. (1992). Conserved histidine residues in soybean lipoxygenase: functional consequences of their replacement. *Biochemistry* 31: 4053-4057.
Tamura K, Dudley J, Nei M and Kumar S (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596-1599.
Tanurdzic M and Banks JA (2004). Sex-determining mechanisms in land plants. *Plant Cell* 16 Suppl: S61-S71.
Wang R, Shen W, Liu L, Jiang L, et al. (2008). A novel lipoxygenase gene from developing rice seeds confers dual position specificity and responds to wounding and insect attack. *Plant Mol. Biol.* 66: 401-414.
Zhang B, Chen K, Bowen J, Allan A, et al. (2006). Differential expression within the LOX gene family in ripening kiwifruit. *J. Exp. Bot.* 57: 3825-3836.
Zhou G, Qi J, Ren N, Cheng J, et al. (2009). Silencing OsHI-LOX makes rice more susceptible to chewing herbivores, but enhances resistance to a phloem feeder. *Plant J.* 60: 638-648.

SUPPLEMENTARY MATERIAL

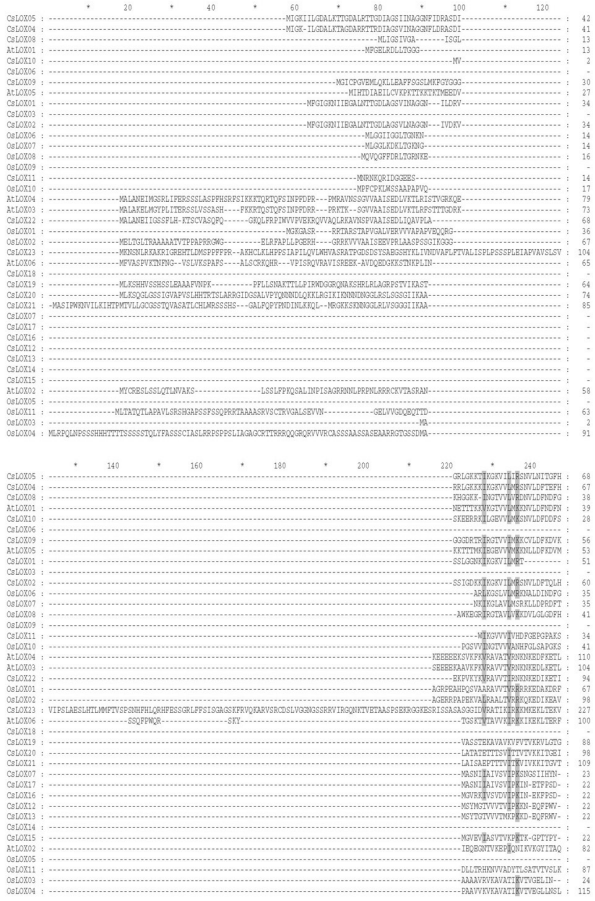
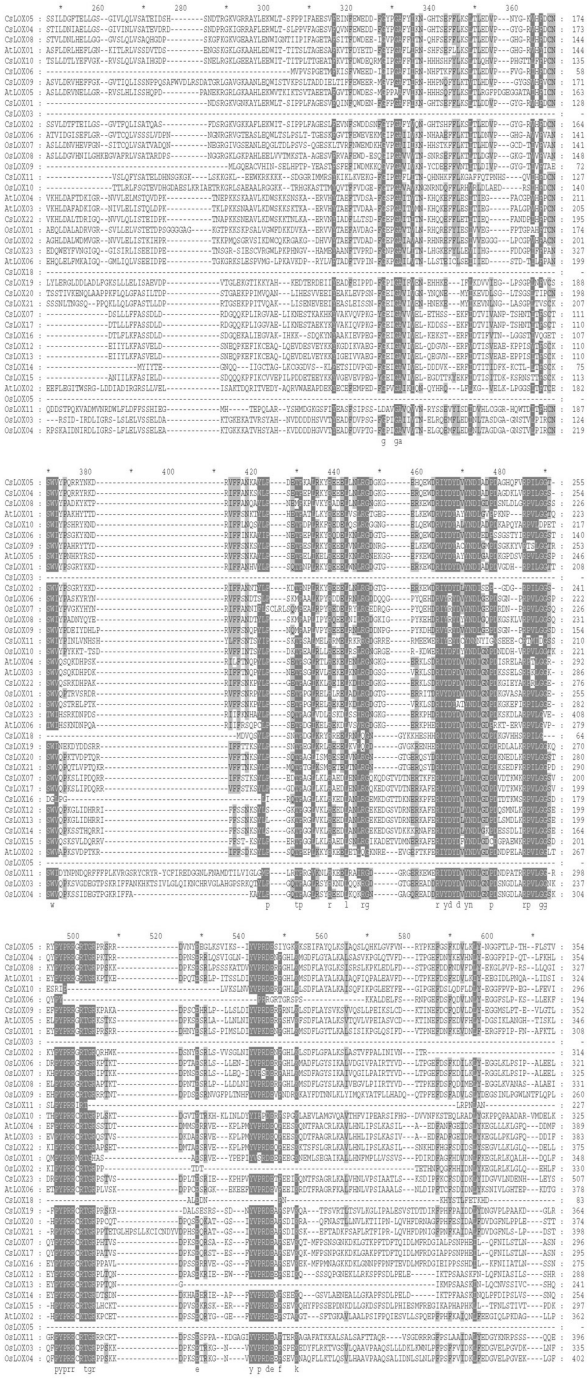


Figure 1. Multiple sequence alignments of cucumber LOXs with those from *Arabidopsis* and rice. Forty LOX protein sequences including 23 CsLOX, 11 OsLOX and 6 reported *Arabidopsis* AtLOX members were aligned using the Clustal X software. The consensus amino acid residues are marked with letters at the bottom.

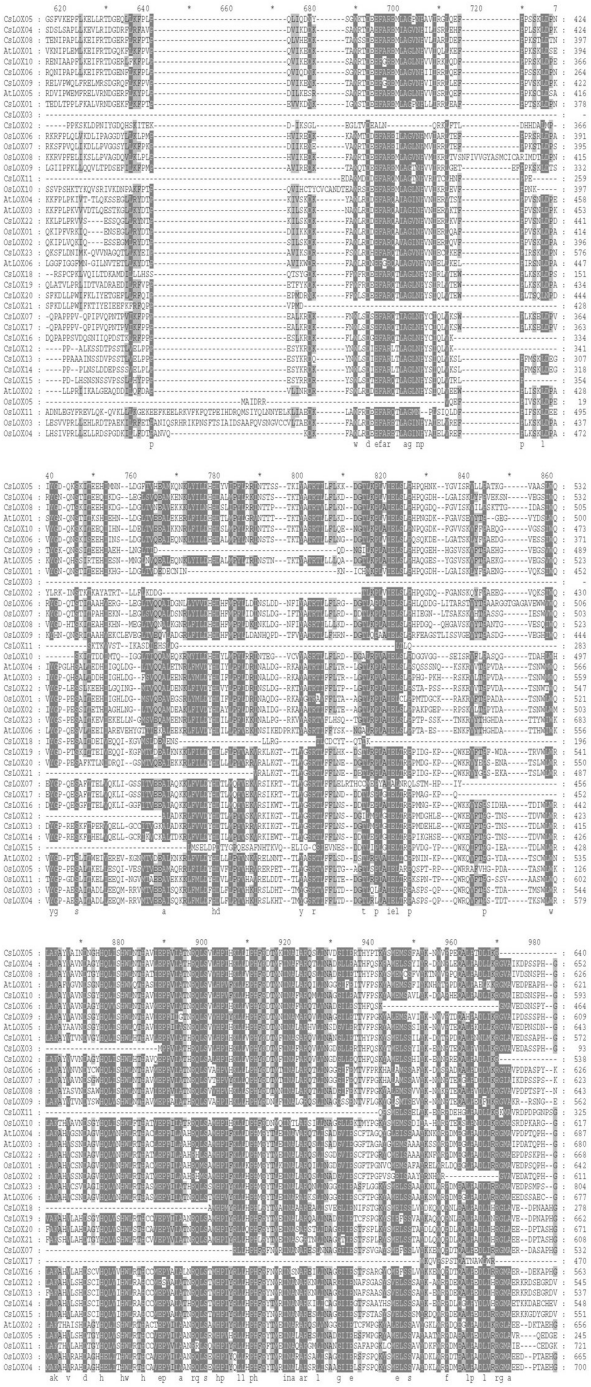
Continued on next page

Figure 1. Continued.



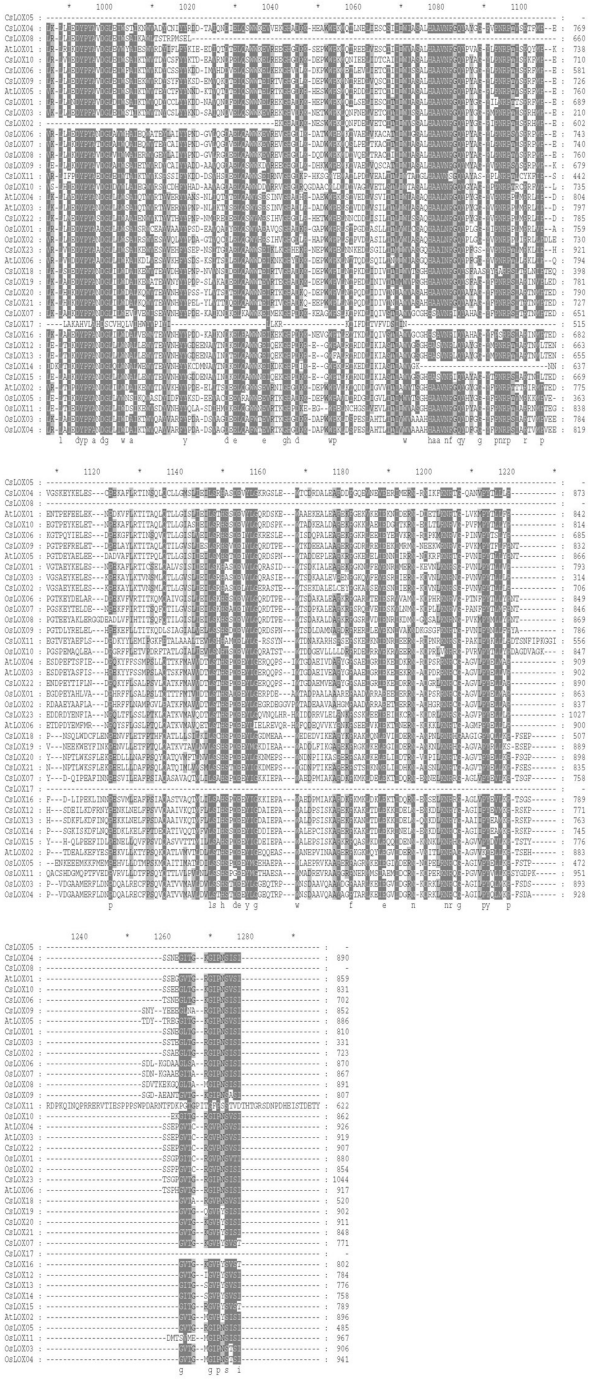
Continued on next page

Figure 1. Continued.



Continued on next page

Figure 1. Continued.



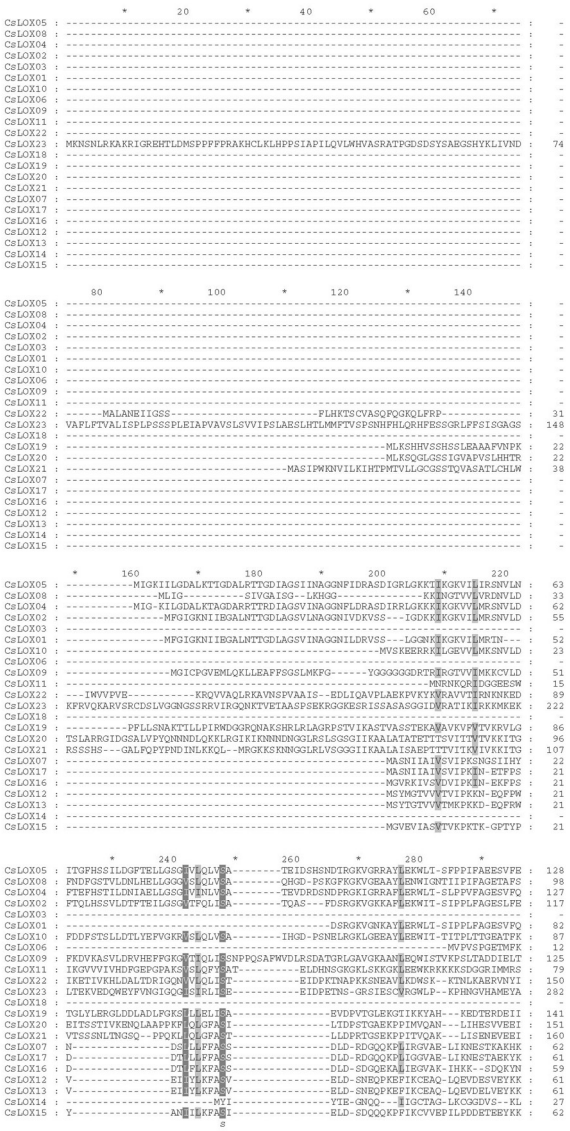
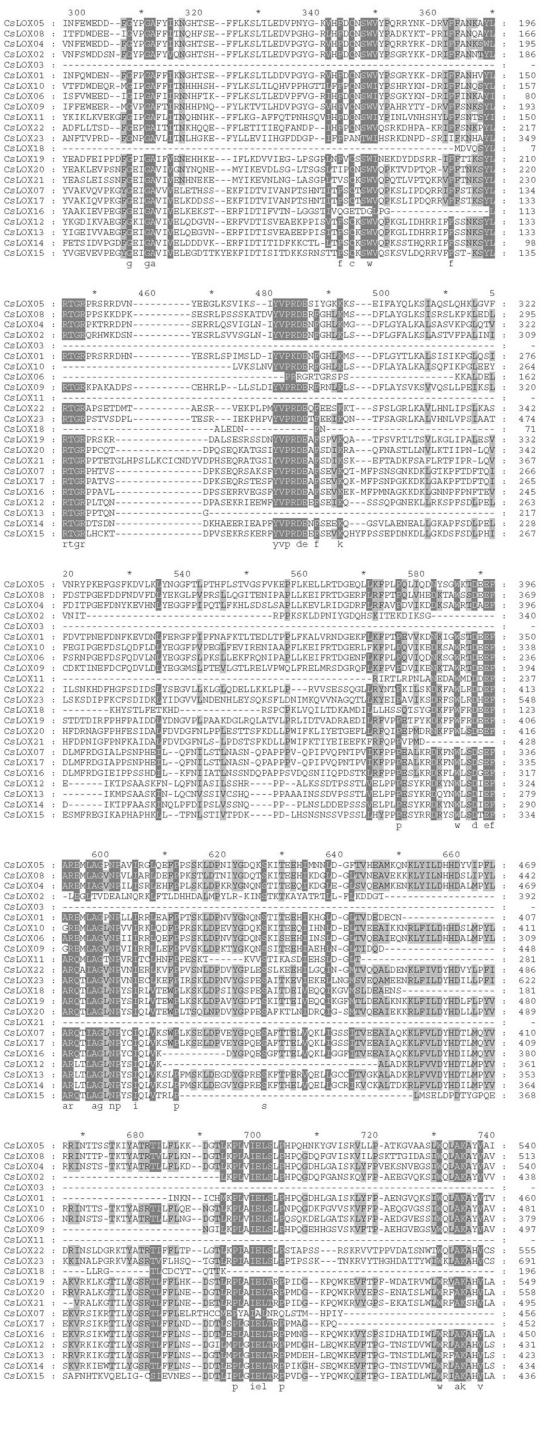


Figure 2. Multiple sequence alignments of 23 cucumber LOX proteins. The protein sequences were aligned using the Clustal X software. The consensus amino acid residues are marked with letters at the bottom.

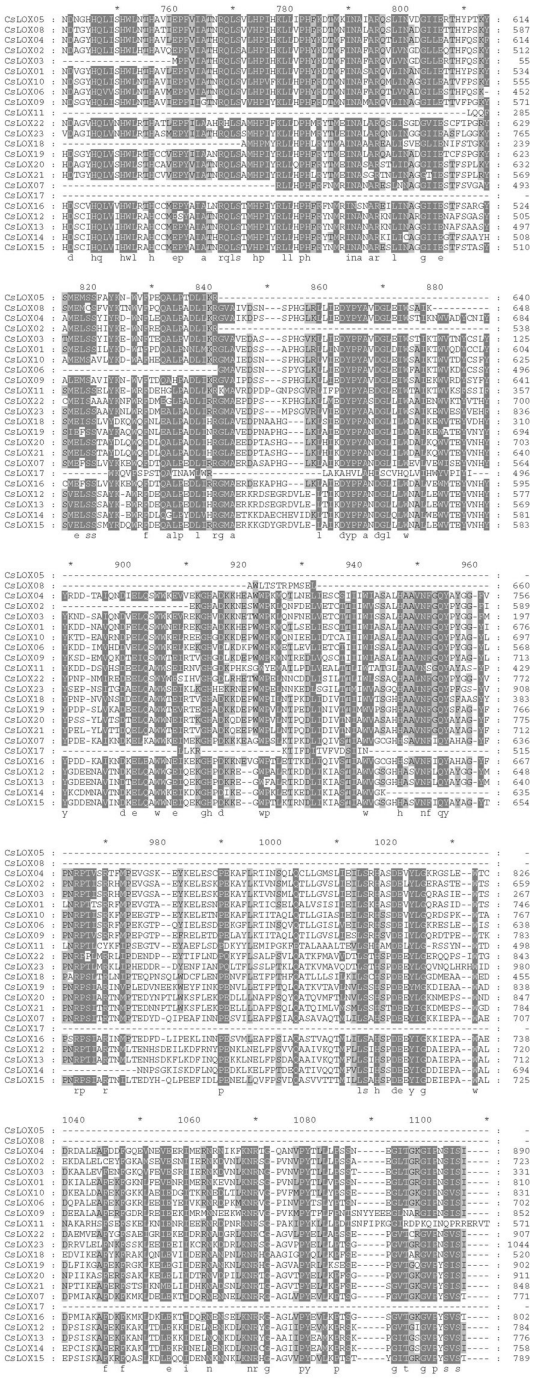
Continued on next page

Figure 2. Continued.



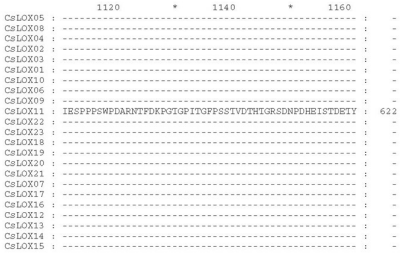
Continued on next page

Figure 2. Continued.



Continued on next page

Figure 2. Continued.



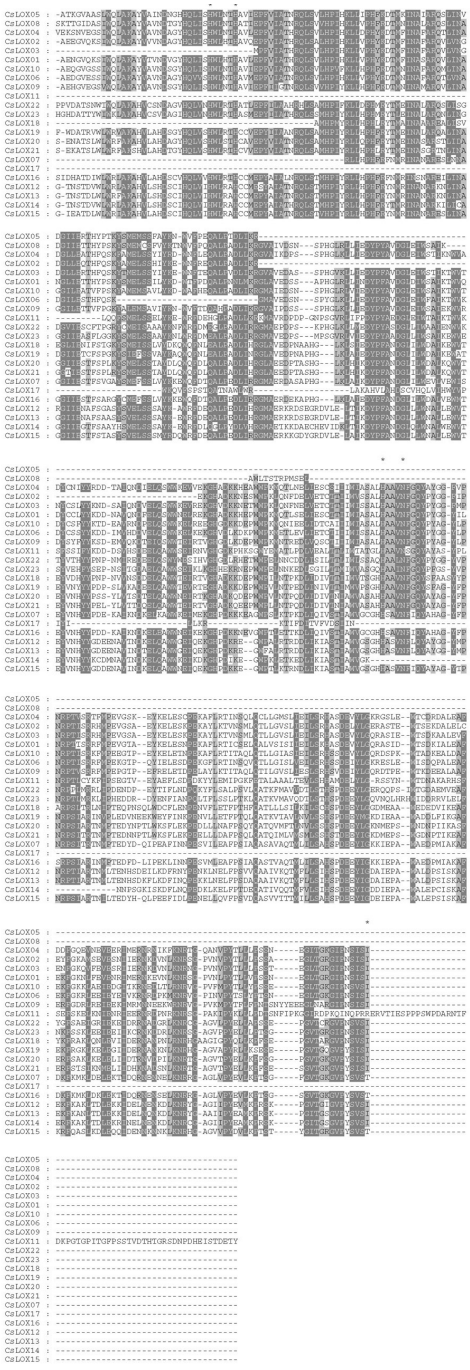


Figure 3. Multiple sequence alignments of LOX domains of 23 cucumber LOX proteins. The protein sequences were aligned using the Clustal X software. The putative five iron binding sites are identified and the positions are marked with asterisks.

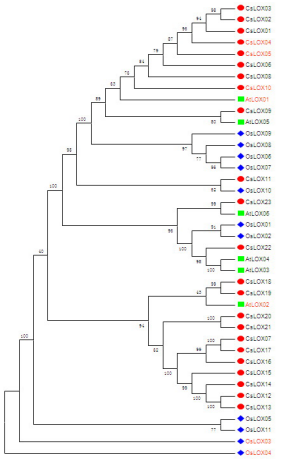


Figure 4. Phylogenetic analysis of cucumber LOXs with those from *Arabidopsis* and rice. A maximum parsimony tree was constructed using PHYLIP3.69 by the multiple sequence alignments of 40 LOX protein sequences including 23 CsLOX, 11 OsLOX and 6 reported *Arabidopsis* AtLOX members and 83% of the *CsLOX* members were placed into the same locations as those in the neighbor joining tree. The names of discrepant proteins are marked in red.

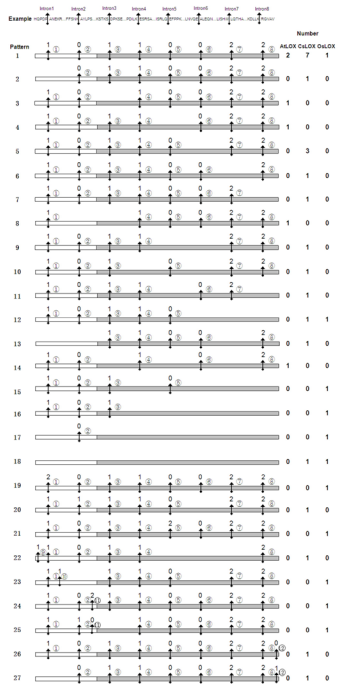


Figure 5. Scheme of the intron distribution patterns (designated 1~27) of cucumber, *Arabidopsis* and rice *LOX* genes. The gray rectangles mean the LOX domain. The upward arrows indicate the position of the intron coinciding with the example. The numbers above the arrows indicate the splicing phases: 0 refers to phase 0, 1 to phase 1, and 2 to phase 2. The markers 1~12 beside the arrows correspond to the different positions of the introns. The number of proteins with each pattern is indicated on the right. Here the position of introns in the variable region has been adjusted manually to make them more contracted.

Supplementary Table 1. Information of primers with RT-PCR.

Primer name	Sequence (5' to 3')	Size (bp)
CsLOX01_F	TGAGCGTGCTTCATCCAATCC	435
CsLOX01_R	CATGTTTCTTGTCAGCGTGGC	
CsLOX02_F	CATGAAGGGCTGAAGATACCC	244
CsLOX02_R	GTCCTTGTTAATGGTGATGGTC	
CsLOX03_F	TGGTGGAAAGAGGTTAGA	202
CsLOX03_R	TATGCCTACGACTTATTG	
CsLOX04_F	ATTTGCTATGGAGTTGTCTTC	312
CsLOX04_R	CATGCTTCATGTTTCTTGTCA	
CsLOX05_F	ACACCCTCAACATAACAA	304
CsLOX05_R	CTCAATAATGCCATCAAC	
CsLOX06_F	AACCAAATCTCAAGTCCATCA	215
CsLOX06_R	CTTCATCCAATTCACAAGCTG	
CsLOX07_F	CGTATCAACGCAAATGCT	231
CsLOX07_R	ATTGGCGAAAGGGTAGTC	
CsLOX08_F	GCTGATGGGATTATTGAAAC	257
CsLOX08_R	TCACTCATTGGACGAGTTGA	
CsLOX09_F	GAGCCATTCATCATTGGAACA	320
CsLOX09_R	CATCGACTGCGTAAGGGTAAT	
CsLOX10_F	TCATTCCGTACTGCTTCATCG	348
CsLOX10_R	GCTGCTCCATCCTCACTTTCG	
CsLOX11_F	GAAATCGAGATCCCAACC	257
CsLOX11_R	TCTTCCCGTATGAGTGTC	
CsLOX12_F	CGTAAGAACCTCATCAATGCCA	323
CsLOX12_R	GCTTGCAGCTCTTTATCATTTG	
CsLOX13_F	CCACCAAGCTTGTAGCTCTG	332
CsLOX13_R	CTCGTAAGAACCTTATCAATGCT	
CsLOX14_F	GCTTGCAGCTCTTTATCATTAG	304
CsLOX14_R	GTGGCATCATTGAAGGAACAT	
CsLOX15_F	TCAAGAACATCACGACCGTAGT	274
CsLOX15_R	CAAATAGGCAATTGAGTACGG	
CsLOX16_F	GGCGAAAGGGTAGTCTTCTATG	353
CsLOX16_R	AACTCGTTGTTCACTGGCTTAG	
CsLOX17_F	AAGCCTTTCACCGACTTTAC	344
CsLOX17_R	GCTGATTCTTGAGGACCATA	
CsLOX18_F	ACAACATTTGGGTTGGGATAG	402
CsLOX18_R	GCGTGGAACCTTATGTGATTG	
CsLOX19_F	AGGGTAGTCTTCAATAGCAAGC	305
CsLOX19_R	CTTACATCATTGCAGCAAACAG	
CsLOX20_F	TGCTGGTCTTAACCCATACAG	286
CsLOX20_R	CCTAAGTGTTCCATCCTCGTT	
CsLOX21_F	GAATTTAAGTTCCGCTTTCAG	152
CsLOX21_R	CCATCAATAGGTGGTCGTGTC	
CsLOX22_F	ACGCCACAAGCAACTGGACAT	302
CsLOX22_R	TTTGTAAGCTGCGGCACTGAT	
CsLOX23_F	TGCCTCCAACACCTTCTTCAA	381
CsLOX23_R	CTTCCATATCAAATCGCCACA	
Cs-actin_F	TCGTGCTGGATTCTGGTG	161
Cs-actin_R	GGCAGTGGTGGTGAACAT	

Supplementary Table 2. The *CsLOX* and *OsLOX* genes identified in this study are listed according to their CsLOX and OsLOX numbers determined by the multiple sequence alignments in Supplementary Figure 1.

A. Information of the *CsLOX* genes.

Serial No.	Gene name	Gene ID	CuGI (5'-3')	Length of chr (bp)	Chromosomal location	Length (a.a.)	No. of introns
1	*CsLOX01*	Csa006735	2240279-2245149	22000590	2	810	8
2	*CsLOX02*	Csa006736	2254250-2250605	22000590	2	723	6
3	*CsLOX03*	Csa022479	2-1249	18980	Scaffold000337	331	0
4	*CsLOX04*	Csa022478	15160-8798	18980	Scaffold000337	890	8
5	*CsLOX05*	Csa013924	2277497-2270183	22000590	2	640	7
6	*CsLOX06*	Csa006734	2227936-2233502	22000590	2	702	7
7	*CsLOX07*	Csa010334	9890482-9894753	22434855	4	771	7
8	*CsLOX08*	Csa006732	2198766-2204394	22000590	2	660	9
9	*CsLOX09*	Csa007837	17399256-17394225	26547709	6	852	8
10	*CsLOX10*	Csa006731	2190914-2196777	22000590	2	831	8
11	*CsLOX11*	Csa019335	11655143-11663841	17370973	7	622	8
12	*CsLOX12*	Csa010340	9971732-9975311	22434855	4	784	7
13	*CsLOX13*	Csa010337	9946894-9950404	22434855	4	763	8
14	*CsLOX14*	Csa010338	9955445-9958917	22434855	4	758	8
15	*CsLOX15*	Csa010339	9963992-9967539	22434855	4	789	8
16	*CsLOX16*	Csa010336	9939209-9943434	22434855	4	802	7
17	*CsLOX17*	Csa010335	9930238-9933070	22434855	4	515	5
18	*CsLOX18*	Csa010343	10007011-10010741	22434855	4	520	5
19	*CsLOX19*	Csa010344	10018491-10023531	22434855	4	902	8
20	*CsLOX20*	Csa010342	9994839-9999134	22434855	4	911	7
21	*CsLOX21*	Csa010341	9983100-9989131	22434855	4	848	6
22	*CsLOX22*	Csa009893	16491370-16487281	17370973	7	907	6
23	*CsLOX23*	Csa000832	20858846-20863032	28091742	5	1044	8

The *CsLOX01* and *CsLOX09* are coincident with *CsLOX1* and *CsLOX9*, respectively, which are identified by Matsui et al. (1999, 2006).

B. Information of the *OsLOX* genes.

Serial No.	Gene name	Locus ID	Coordinates (5'-3')	Length of chr (bp)	Chromosomal location	Length (a.a.)	No. of introns
1	*OsLOX01*	LOC_Os04g37430	22085990-22090246	35498469	4	880	7
2	*OsLOX02*	LOC_Os03g08220	4166275-4172656	36192742	3	854	8
3	*OsLOX03*	LOC_Os08g39840	25083299-25091237	28434780	8	924	5
4	*OsLOX04*	LOC_Os08g39850	25108063-25117436	28434780	8	941	5
5	*OsLOX05*	LOC_Os02g10120	5276617-5282623	35954743	2	485	0
6	*OsLOX06*	LOC_Os03g52860	30266016-30269529	36192742	3	870	3
7	*OsLOX07*	LOC_Os03g49260	27998615-28002905	36192742	3	867	9
8	*OsLOX08*	LOC_Os03g49380	27906867-27913264	36192742	3	891	8
9	*OsLOX09*	LOC_Os11g36719	21173710-21183442	28386948	11	807	1
10	*OsLOX10*	LOC_Os05g23880	13656472-13662485	29737217	5	862	8
11	*OsLOX11*	LOC_Os12g37260	22821051-22826640	27566993	12	967	8

The *OsLOX03* is coincident with the *OsHI-LOX*, which was identified by Zhou et al. (2009). Similarly, *OsLOX06*, *OsLOX07* and *OsLOX08* correspond to *L-2*, *rg-LOX1*, *OsLOX1*, respectively.