Theor Appl Genet (2012) 124:613–622
DOI 10.1007/s00122-011-1733-2

ORIGINAL PAPER

Molecular sequence variations of the lipoxygenase-2 gene in soybean

Jin Hee Shin · Kyujung Van · Kyung Do Kim · Yeong-Ho Lee · Tae-Hwan Jun · Suk-Ha Lee

Received: 23 May 2011 / Accepted: 14 October 2011 / Published online: 9 November 2011

Abstract Soybean lipoxygenase genes comprise a multigene family, with the seed lipoxygenase isozymes LOX1, LOX2, and LOX3 present in soybean seeds. Among these, the LOX2 isozyme is primarily responsible for the "beany" flavor of most soybean seeds. The variety, Jinpumkong 2, having null alleles (*lx1*, *lx2*, and *lx3*) lacks the three seed lipoxygenases; so, sequence variations between the lipoxygenase-2 genes of Pureunkong (*Lx2*) and Jinpumkong 2 (*lx2*) cultivars were examined. One indel, four single nucleotide polymorphisms (SNPs), a 175-bp fragment in the 5′-flanking sequence, and a missense mutation within the coding region were found in Jinpumkong 2. The distribution of the sequence variations was investigated among 90 recombinant inbred lines (RILs) derived from a cross of Pureunkong × Jinpumkong 2 and in 480 germplasm accessions with various origins and maturity groups. Evidence for a genetic bottleneck was observed: the 175-bp fragment was rare in *Glycine max*, but present in the majority of the *G. soja* accessions. Furthermore, the 175-bp fragment was not detected in the 5′ upstream region of the *Lx2* gene on chromosome (Chr) 13 in Williams 82; instead, a similar 175-bp fragment was positioned in the homeologous region on Chr 15. The findings indicated that the novel fragment identified was originally present in the *Lx2* region prior to the recent genome duplication in soybean, but became rare in the *G. max* gene pool. The missense mutation of the conserved histidine residue of the *lx2* allele was developed into a single nucleotide-amplified polymorphism (SNAP) marker. The missense mutation showed a perfect correlation with the LOX2-lacking phenotype, so the SNAP marker is expected to facilitate breeding of soybean cultivars which lack the LOX2 isozyme.

Communicated by H. T. Nguyen.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-011-1733-2) contains supplementary material, which is available to authorized users.

J. H. Shin · K. Van · K. D. Kim · Y.-H. Lee · S.-H. Lee (✉)
Department of Plant Science and Research Institute for Agriculture and Life Sciences, Seoul National University, San 56-1, Sillim-dong, Gwanak-gu, Seoul 151-921, The Republic of Korea
e-mail: sukhalee@snu.ac.kr

T.-H. Jun
Department of Entomology, The Ohio State Univ./OARDC, 1680 Madison Avenue, Wooster, OH 44691, USA

S.-H. Lee
Plant Genomics and Breeding Research Institute, Seoul National University, Seoul 151-921, Korea

Introduction

Lipoxygenases (LOXs) are lipid-oxidizing enzymes which are ubiquitous in the plant and animal kingdoms (Start et al. 1986). They may be present in all plant tissues, but concentrations are thought to depend on the developmental stage and organ. LOXs are involved in various physiological processes such as growth, development, and responses to biotic and abiotic stresses (Bell and Mullet 1993; Veronesi et al. 1996; Gardner 1998; Kolomiets et al. 2000).

Soybean LOX1, 2, and 3 have been extensively studied because their activity in the seed results in an unpleasant beany flavor (Kitamura et al. 1985). Several treatments have been proposed to resolve the problem of the undesirable flavor associated with soybean products, but they are expensive and not entirely satisfactory (Nishiba et al. 1995). Several soybean lines showing deficiencies in lipoxygenase

isozymes in mature seeds have been previously reported (Kim et al. 1997; Kitamura 1984; Reinprecht et al. 2011). Genetic studies of these lines have demonstrated that the absence of LOX1, LOX2, and LOX3 in mature seeds is due to the single recessive alleles *lx1*, *lx2*, and *lx3*, respectively (Hajika et al. 1992). LOX2 is the main isozyme associated with the undesirable flavor, and soymilk made from LOX2-free soybeans has a less beany flavor (Davies et al. 1987). Similarly, soy products made from soybeans lacking seed lipoxygenases have shown improved flavor characteristics (Narvel et al. 1998).

The amounts and activity of lipoxygenase have been measured using several methods, including sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of the isozymes, assays that measure substrate consumption, and carotenoid bleaching reactions (Axelrod et al. 1981; Kitamura 1984; Romero and Barrett 1997). A SNP in the intron of *lx2* was developed into a SNAP marker and used for mapping (Kim et al. 2004). In addition, genotyping assays based on the melting properties of gene-specific PCR products were developed for the *Lx1*, *Lx2*, and *Lx3* genes (Lenis et al. 2010). The development of lipoxygenase-targeted markers will offer advantages over existing assays for screening of soybeans to select plants or lines lacking LOXs. Moreover, molecular markers are good tools for measuring genetic distance due to their selective neutrality (Gepts 2006).

Soybean is a diploidized paleopolyploid whose genome has gone through two rounds of whole-genome duplication (WGD) (Blanc and Wolfe 2004; Schlueter et al. 2004). As a result, four homeologous chromosomal regions or genes are expected in soybean (Innes et al. 2008; Shin et al. 2008; Kim et al. 2009). In a previous study, we reported that the lipoxygenase gene family was expanded by two WGDs and extensive tandem gene duplication (Shin et al. 2008). Four genomic regions containing 19 *Lx* genes were analyzed, with the results showing that these regions contain two similar yet distinctive pairs of homeologous regions (GmA and GmA′, and GmB and GmB′) (Shin et al. 2008). The three seed lipoxygenase genes (*Lx1*, *Lx2*, and *Lx3*) were neighbored by previously reported vegetative *Lx* genes (*Lx4*, *Lx5* and *Lx7*) in a pair of homeologous regions (GmA and GmA′) duplicated by a recent WGD event.

In the present study, sequence variation between the lipoxygenase-2 gene from Pureunkong (*Lx2*) and Jinpumkong 2 (*lx2*) was investigated in the 5′-UTR region and in the eighth of the nine exons. The objective of this study was to investigate the distribution of the sequence variation on lipoxygenase-2 gene among 90 recombinant inbred lines (RILs) developed from Pureunkong × Jinpumkong 2 and in 480 other *G. max* and *G. soja* germplasm accessions. The expression pattern of the *Lx2* gene was compared to that of other duplicated lipoxygenase genes in two homeologous regions, and a conserved regulatory element in the fragment was detected. A SNP in the eighth exon of the *lx2* allele was determined to be the causative mutation for the LOX2-lacking phenotype, and it was developed into a SNAP marker for a simple genotyping assay based on PCR and agarose gel electrophoresis.

Materials and methods

Plant materials

Pureunkong produces small seeds with a grassy–beany flavor, and the *Lx2* gene is dominant in this genotype. Jinpumkong 2 for soybean curd production lacks the three seed lipoxygenase isozymes in mature seeds, and has the single recessive alleles *lx1*, *lx2*, and *lx3* (Kim et al. 1997). Ninety RILs developed by single-seed descent from Pureunkong × Jinpumkong 2 were used in this study (Kim et al. 2004). In addition to the RILs and their parents, 376 *G. max* and 104 *G. soja* accessions (480 total) with various origins and maturities were screened (Table 1 and Supplementary Table 1). These germplasm accessions were provided by the USDA Soybean Germplasm Collection and the Korean National Institute of Crop Science, Rural Development Administration in Korea (http://www.nics.go.kr/).

Sequence analysis, SNP discovery, and primer design

To obtain nucleotide sequences immediately 5′ upstream of the *Lx2* gene in Pureunkong (GenBank accession No. EU028326) and Jinpumkong 2 (GenBank accession No. EU028327), three primer sets were designed using the GENERUNR program based on the 5′-upstream sequence of *Glyma13g42310* available from the Soybean Genome Project, DOE Joint Genome Institute (http://www.phytozome.com/soybean.php). Primer sequences are shown in Supplementary Table 2. The presence of the 175-bp fragment was investigated using a Set B primer, which showed an amplicon length polymorphism between Pureunkong and

Table 1 Distribution of the 175-bp fragment in various soybean germplasms

	NPJ	No. of accessions (%)	
	No. of lines	*G. max*	*G. soja*
Presence of 175 bp	52	65 (17)	78 (75)
Hetero	2	9 (2)	8 (8)
Absence of 175 bp	36	296 (79)	18 (17)
No band	–	6 (2)	0 (0)
Total	90	376 (100)	104 (100)

NPJ is a RIL population from a cross between Pureunkong (*Lx2*) and Jinpumkong 2 (*lx2*)

Jinpumkong 2. The amplification reaction was carried out in a total volume of 11 µl and contained 100 ng of each genomic DNA, 15 pmol of each forward and reverse primer, 10.0 mM of dNTP mix, 1.0 µl of 10× buffer, 6.9 µl of ddH_2O, and 0.2 units of *Taq* DNA polymerase (Vivagen, Sungnam, Korea). PCR was performed using a DNA Engine Gradient Cycler from MJ Research (Watertown, MA, USA). The PCR conditions were 94°C for 2 min, 35 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 30 s, and a final extension of 2 min at 72°C. The PCR products were purified using 2 U exonuclease I and 1 U shrimp alkaline phosphatase for 1 h at 37°C and 15 min at 75°C. Sequencing was done using the BigDye terminator cycle sequencing kit (Applied Biosystems, Foster City, CA, USA) and an ABI 3730 sequencer. The sequences were aligned using SeqScape version 2.0 (Applied Biosystems) to detect SNPs.

To analyze the 5′ upstream region of the *Lx2* gene, sequences were BLASTN searched against the soybean genome database (http://www.phytozome.net/soybean.php). The homologous region in Williams 82 was compared to the region around *Lx2* with the BLASTZ program (Schwartz et al. 2003) and visualized using SynBrowse (http://www.synbrowse.org) and Gbrowse (http://www.gmod.org/ggb/gbrowse.shtml). The 175-bp fragment and 175 bp-like sequences were compared using BOXSHADE 3.21 (http://www.ch.embnet.org/software/BOX_form.html).

A missense mutation of T to A in the *lx2* allele has been reported to result in a change of one of the conserved histidines (CAT) to a glutamine (CAA) (Wang et al. 1994). To detect the previously reported SNP from Pureunkong and Jinpumkong 2, primers were designed based on D13949, *G. max lx2* mRNA for lipoxygenase-2. As the amino acid identity between LOX1 and LOX2 was high (81.8%) and the 5′ sequence of the missense mutation was highly conserved between *Lx1* and *Lx2*, allele-specific primers were designed from the 3′ region downstream of the mutation. The primers were BLASTN searched against the soybean genome database to prevent cross-hybridization among other *Lx* genes. The amplification reaction and conditions were the same as those described above. To find the optimum PCR conditions for each primer, gradient PCR with various annealing temperatures from 55 to 70°C was performed. The PCR products were electrophoresed on 1.0% agarose gels and visualized by ethidium bromide staining. PCR for genotyping was conducted with the annealing temperature resulting in a single band. Genotyping results were confirmed by sequencing as described above.

RNA extraction and reverse transcription (RT)-PCR

Two-week-old leaves and mid-maturation stage seeds collected 35–40 days after anthesis were harvested from the Pureunkong and Jinpumkong 2 cultivars and immediately frozen in liquid nitrogen. The Williams 82 cultivar was grown in a growth chamber at 25–28°C with a 12-h light/12-h dark cycle. Tissues were sampled as described by Schlueter et al. (2006) with slight modifications. Cotyledon tissues were collected 3 and 7 days after emergence (DAE), roots at 3 and 8 DAE, furled unifoliate at 3 DAE, unfurled unifoliate at 4 DAE, furled trifoliolate at 11 DAE, unfurled trifoliolate at 15 DAE, hypocotyls at 3 and 8 DAE, and flowers at 35 DAE. Seeds were collected from 14 to 35 days after flowering (DAF) and mixed to mitigate the effect of developmental stages. Total RNAs were extracted using Trizol (Invitrogen, Carlsbad, CA, USA) and treated with DNase (Promega, Madison, WI, USA) prior to RT-PCR to remove genomic DNA contamination. RT-PCR was performed using One-Step RT-PCR (Invitrogen).

Phylogenetic analysis and motif search

Sequences of 19 soybean and two *M. truncatula* lipoxygenases were compiled and aligned using ClustalX, and the sequence overhangs at the 5′- and 3′-ends of alignments were removed. A parsimony tree was generated using PAUP* 4.0 (Swofford 2003) by heuristic search with 1,000 bootstrap replicates. The tree was rooted with *Arabidopsis* and rice lipoxygenase genes as outgroups. The 1-kb upstream sequences were downloaded from Phytozome (http://www.phytozome.net/) and conserved regulatory elements were identified using FootPrinter (Blanchette and Tompa 2003) with motif sizes of 7, 9 and 11 bases. Predicted motif sequences were searched in the PLACE database for the presence of known *cis*-acting elements (Higo et al. 1999).

Results

Sequence variations in the 5′-flanking region of lipoxygenase-2

Sequence variations in the 5′ upstream of *Lx2* and *lx2* alleles between Pureunkong and Jinpumkong 2 were investigated (Fig. 1). Three primer sets (Set A, Set B, and Set C) were designed to amplify the 5′ flanking region of *Lx2* genes (Supplementary Table 2).

There was no polymorphism in the length of the PCR products amplified with Set A and Set C primers. However, amplification with the Set B primer pairs produced amplicons of different sizes for Pureunkong (594 bp) and Jinpumkong 2 (769 bp). Analysis of the 5′ flanking regions of *Lx2* genes from Pureunkong (GenBank accession No. EU028326) and Jinpumkong 2 (GenBank accession No. EU028327) revealed one indel, four SNPs, and a 175-bp fragment in Jinpumkong 2. The 175-bp fragment was not

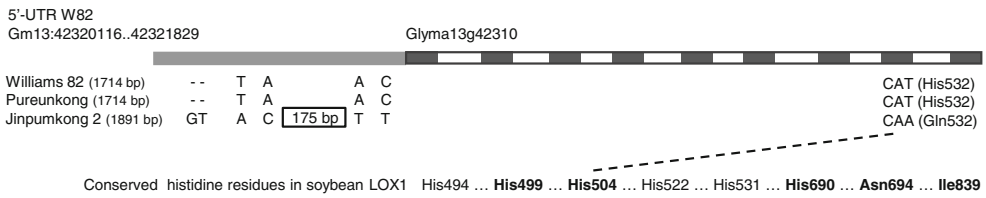

Fig. 1 Sequence variations in the 5′-upstream and in the eighth exon of lipoxygenase-2 gene

Table 2 Sequence of primers specific to the T/A SNP between *Lx2* and *lx2*

Primer	Direction	Primer sequence (5′ → 3′)
Lx2-specific primer	Forward	CCAGTAAGATAACAGCAGATGCC
	Reverse	AATGGCTCAATCACCGC**A**
lx2-specific primer	Forward	AAACCAGTAAGATAACAGCAGATG
	Reverse	AATGGCTCAATCACCGC**T**

present in the Williams 82 genomic sequence (http://www.phytozome.net/soybean.php). A seed lipoxygenase-free line, OX948 (GenBank accession No. GU942744), was found to have the 175-bp fragment that shares 100% sequence identity with that of Jinpumkong 2. The OX948 sequence shared four SNPs and the 175-bp insertion with the Jinpumkong 2 sequence, and additionally it had five deletions and one SNP in the 5′ UTR of *lx2* allele (Reinprecht et al. 2011).

Distribution of the 175-bp fragment among soybean genotypes

To test for the presence of the 175-bp fragment among 90 RILs and 480 accessions, Set B primer pairs were used to amplify the region around the fragment. Among 90 RILs, 36 and 52 lines showed the same band pattern as Pureunkong (594 bp) and Jinpumkong 2 (769 bp), respectively (Table 1 and Supplementary Table 2). Two lines exhibited double bands, indicating heterozygosity.

The distribution of the 175-bp fragment among the 480 germplasm accessions showed that the majority of the *G. max* accessions did not contain the fragment within the 5′ region of the *Lx2* gene. In contrast, this fragment was common in *G. soja* with 78 out of 105 wild soybean accessions containing the 175-bp fragment. Additionally, a wild soybean accession with low heterozygosity IT182932 also had the fragment, and the fragment sequence (GenBank accession No. HM591295) was 100% identical to that of Jinpumkong 2.

The 175 bp-like fragment in the homeologous region

The 5′ sequence of Pureunkong and Jinpumkong 2 was BLASTN searched against the soybean genome database and was found to be located on chromosome (Chr) 13 of the Williams 82 reference sequence (Fig. 2) (Schmutz et al. 2010). However, the 175-bp fragment from Jinpumkong 2 was not found on Chr 13 of Williams 82. Instead, a region similar to the 175-bp sequence was detected in Chr 15 of Williams 82. This region was found to be homeologous to the *Lx2* region on Chr 13 that was produced by a genome duplication approximately 13 million years ago (Fig. 2) (Schlueter et al. 2004; Shin et al. 2008).

The size of the duplicated blocks was approximately 3.6 Mbp. They were labeled GmA (Gm15: 180,263–3,799,985 bp, 3.619 Mbp) and GmA′ (Gm13: 40,852,483–44,408,124 bp, 3.556 Mbp) (http://www.soybase.org). Most genes and their order were conserved between these regions, and seven *Lx* genes were identified. The detailed structure of the regions around the *Lx* genes is represented by BLASTZ (Fig. 2). The orientation and the length of the genes are represented by arrows, and homologous sequences are connected by gray boxes. Seven *Lx* genes are highlighted with yellow and green boxes depending on the level of similarity of their proteins. *Lx* gene names and identifiers in the whole soybean genome assembly (http://www.phytozome.net/soybean) are denoted. A detailed description of *Lx* genes is provided in Supplementary Table 3. Previously reported *Lx1* (*Glyma13g42320*), *Lx2* (*Glyma13g42310*), *Lx3* (*Glyma15g03030*), *Lx4* (*Glyma13g42330*, *vlxA*), *Lx5* (*Glyma15g03050*, *vlxB*) and *Lx7* (*Glyma13g42340*, *vlxE*) genes were located in these regions. One additional *Lx* gene (*Glyma15g03040*) with a high homology to *Lx7* was identified between the *Lx3* and *Lx5* genes.

Sequence analysis of the 5′-upstream sequence of lipoxygenase-2

To assess whether the presence of the 175-bp fragment in the 5′ region of the *Lx2* gene has an effect on transcription, RT-PCR was performed. The *Lx2* and *lx2* transcripts were not detected in 2-week-old soybean leaves collected from the two genotypes having either the *Lx2* or *lx2* genes. In contrast, the gene was preferentially expressed in the seeds of both cultivars, despite the absence of the LOX2 enzyme in Jinpumkong 2 seeds (Supplementary Fig. 1a). The sequences of the RT-PCR products were found to be identical to the *Lx2* and *lx2* sequences. These results

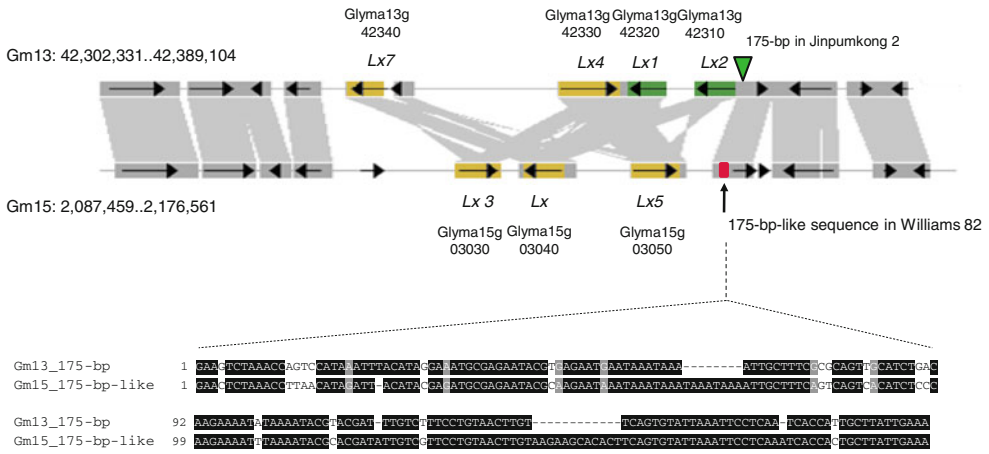

Fig. 2 Clusters of lipoxygenase genes among two homeologous regions. Homeologs are connected with *gray boxes* and genes are represented by *arrows*. Duplicated *Lx* genes are highlighted in *yellow* (**a**) and *green* (**b**) depending on their protein similarity (Fig. 3a). The 175-bp fragment in Jinpumkong 2 (denoted as a *green inverted triangle*) was not found in the 5′ upstream region of the *Lx2* gene on Chr 13 of Williams 82. However, a 175-bp like sequence (a *magenta rectangle*) was detected from the homeologous region on Chr 15 produced by recent genome duplication of soybean

suggest that the 175-bp fragment does not control the transcription of the *Lx2* gene under normal conditions.

The expression patterns and levels of *Lx1*, *Lx2*, *Lx3*, *Lx4,* and *Lx5* have been reported previously (Kato et al. 1992; Wang et al. 1999; Bunker et al. 1995); however, most of these studies focused on gene expression in the seed. In this study, the expression pattern of *Lx2* was compared to other *Lx* genes in the four homeologous regions by RT-PCR and by collecting data from the RNA-Seq Atlas (http://www.soybase.org). A soybean *tubulin* gene, amplified with primers 5′-AACCTCCTCCTCATCGTACT-3′ and 5′-GACAGCATCAGCCATGTTCA-3′, was used as a control. The primer sequences specific to six lipoxygenase genes in the two homeologous regions are shown in Supplementary Table 4. Each primer was designed to span an intron between exons in order to distinguish between amplicons derived from cDNA and those derived from genomic DNA contamination. Each primer was tested prior to RT-PCR by sequencing and a BLASTN search was conducted to confirm their specificity in amplifying only the designated isoforms.

The *Lx1* (*Glyma13g42320*) and *Lx2* (*Glyma13g42310*) sequences were predominantly expressed in the seed, while three *Lx* genes (*Glyma13g42330*, *Glyma15g03050*, *Glyma15g03050*) were ubiquitous in all tissues examined (Supplementary Fig. 1b). *Lx3* (*Glyma15g03030*) showed similar patterns of seed-specific expression, but its transcripts were also weakly detected in the stem, root, and flower. *Lx3* transcripts were not detected in the stem or root in the RNA-Seq Atlas, possibly due to differences between the developmental stage of the root and stem. A detailed description of the 19 *Lx* genes in the four homeologous regions is provided in Supplementary Table 3. Six out of 7 *Lx* genes in the GmA and GmA′ regions have been previously cloned and registered in the GenBank database, whereas only 3 out of 12 *Lx* genes in the GmB and GmB′ regions were reported. This is further supported by the gene expression data provided by the RNA-Seq Atlas (http://www.soybase.org): *Lx* genes registered in GenBank show a higher level of expression than other *Lx* genes in the four homeologous regions analyzed.

Phylogenetic analysis and a motif search were performed to determine if there were known regulatory elements conserved in the 175-bp fragment. A parsimony tree using 19 soybean LOXs was constructed, and two *M. truncatula* LOXs representing two distinct LOX types (α and β) in the most recent common ancestor legume were included (Fig. 3a). Two seed soybean LOXs (GmLOX1 and GmLOX2) were classified into one group, whereas one seed LOX (GmLOX3) was classified into another clade, together with vegetative lipoxygenase proteins.

A total of 22 motifs were identified in the 175-bp fragment using a database of plant *cis*-acting regulatory elements (PLACE) (Higo et al. 1999). Among those, a motif size of seven bases (TGAAAAA) was found to be conserved by FootPrinter (Fig. 3b). The results with different motif sizes, of 9 and 11 bases, are described in Supplementary Fig. 2. The conserved motif was the GT-1 *cis*-element found in the promoter of a soybean CaM isoform (SCaM-4), which plays a role in inducing the expression of the *SCaM-4* gene under pathogen and NaCl stresses (Park et al. 2004). Sequence analysis of the upstream regions suggested diversification of the regulatory sequences among the duplicated *Lx* genes (Fig. 3b).

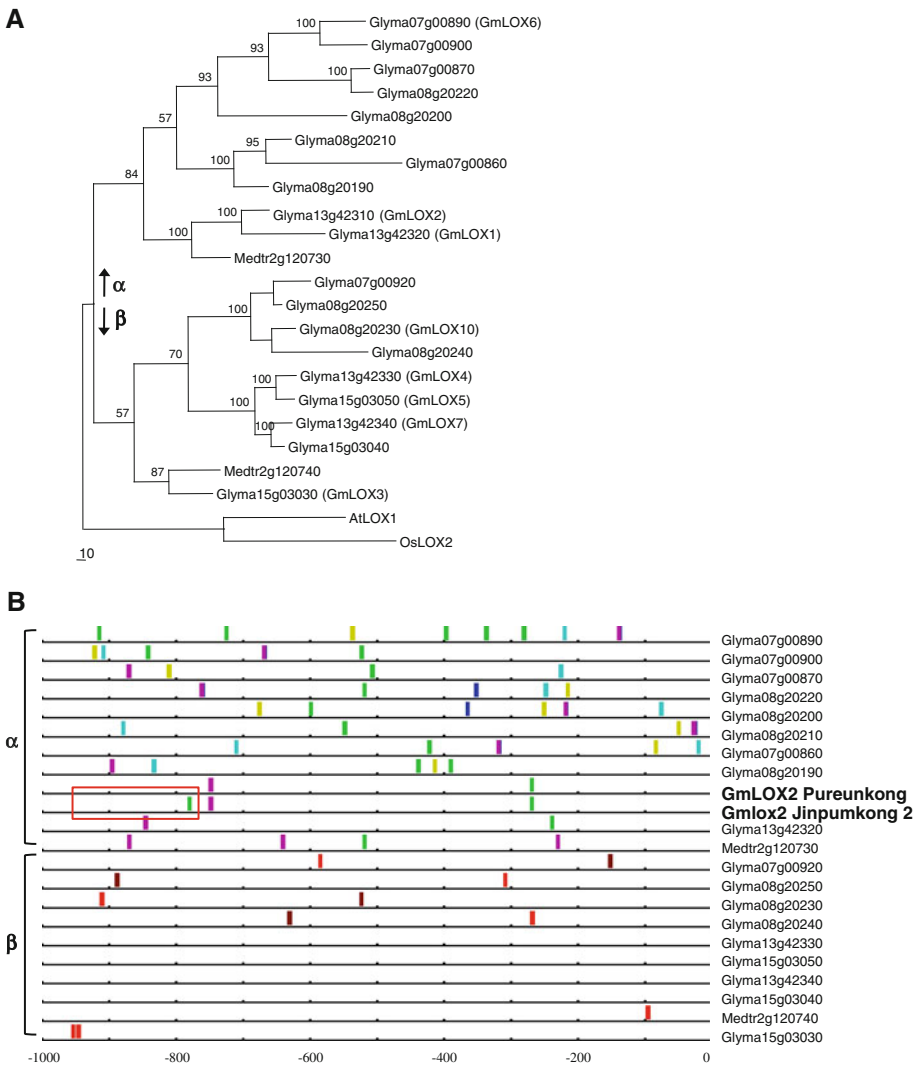

Fig. 3 **a** Phylogenetic analysis, RT-PCR, and the conserved motif search of duplicated *Lx* genes. Phylogenetic analysis was conducted with 19 soybean LOXs. Two *M. truncatula* genes representing two distinct *Lx* genes in the most common ancestor legume were included and the tree was rooted using *Arabidopsis thaliana* LOX1 (AtLOX1) and rice LOX1 (OsLOX1) as outgroups. **b** Conserved regulatory motifs in the 1-kb upstream of *Lx* genes were identified by FootPrinter with motif size of seven bases. Predicted motif sequences were searched in the PLACE database for the presence of known *cis*-acting elements. The location of the 175-bp fragment is represented by a *red box*. A motif which was not contained in Pureunkong (*green bar* in the *red box*) was detected in the 175-bp fragment from Jinpumong 2. This motif was the GT-1 *cis*-element, reported to be induced by pathogen and NaCl stresses

Sequence variation within the coding region of lipoxygenase-2

Two allele-specific primers for the T/A single nucleotide polymorphism (SNP) were designed (Table 2, Fig. 1). The allele-specific reverse primers had T or A at the 3′-termini to ensure that the 3′-terminal nucleotide could be complementary to A (*lx2*) or T (*Lx2*), respectively (Table 2). When gradient PCR was performed at a sufficiently high annealing temperature, only Pureunkong was amplified with a T-specific primer and only Jinpumkong 2 was amplified with an A-specific primer. Thus, the optimum annealing temperature for T/A SNP detection was determined to be 69–70°C. The A-specific primer amplified 52 out of 90 RILs derived from Pureunkong and Jinpumkong 2, and those RILs were previously reported to lack the LOX2 enzyme (Kim et al. 2004). The 175-bp fragment co-segregated with the A allele (*lx2*) in the exon region among 90 RILs. Among the 480 germplasm accessions, only Jinpumkong and Jinpumkong 2 had the A allele (Supplementary Table 1). This result was confirmed by sequencing of the mutation site. These two soybean varieties were previously reported to lack LOX2 (Kim et al. 1997, 2004). These findings suggest that the single point mutation (T to A) in the *lx2* allele is related to the LOX2-free phenotype.

Discussion

Possible role and the origin of the 175-bp fragment, and the expression of lipoxygenase-2

The presence of 175-bp fragment variation in the 5′-untranslated regions of *Lx2* and *lx2* between Pureukong and Jinpumkong 2 co-segregated with the mutant phenotype among the 90 RILs. As a result, the inheritance of the *lx2* allele, the presence of the 175-bp fragment, and the

LOX2-free phenotype were completely linked (52 out of 52) among the 90 RILs of the segregating population derived from Pureunkong and Jinpumkong 2. However, about 30% of the germplasm lines we evaluated had the fragment, suggesting that the mutant phenotype was not a consequence of the presence of the 175-bp fragment in the 5′-UTR (Table 1 and Supplementary Table 1). These results differed from the findings about pea seed lipoxygenase-2, which showed that SNPs within the promoter of the LOX-2 null mutant were responsible for the null phenotype (Forster et al. 1999). In this study, the *lx2* gene is still transcribed in Jinpumkong 2 regardless of the presence of the 175-bp fragment. A known regulatory element, the GT-1 *cis*-element present in the 175-bp fragment, has been shown to play a role in inducing the expression of the *SCaM* gene by disease pathogens and NaCl stress (Park et al. 2004). Further studies are needed to determine the function of the fragment.

The expression pattern of the *Lx2* gene in various tissues was investigated and compared to that of other duplicated *Lx* genes. The expression pattern of the *Lx2* gene was most similar to that of the *Lx1* gene among the 19 *Lx* genes in the four homeologous regions. In a previously reported phylogenetic analysis, 19 LOX proteins from four homeologous regions of the soybean genome were grouped into two major groups (Shin et al. 2008). Interestingly, most of the *Lx* genes previously reported (*Lx1*, *Lx2*, *Lx3*, *Lx4*, *Lx5*, and *Lx7*) were clustered into one pair of homeologous regions (GmA and GmA′). Furthermore, the higher density of *Lx* gene transcripts from a pair of homeologous regions (GmA and GmA′) compared to that from the other pair of regions (GmB and GmB′) was observed from the RNA-Seq Atlas of *G. max*. The different transcription patterns between the tandemly duplicated *Lx* genes in the two homeologous regions reflected the diversification of regulatory sequences and the preferential expression of homeologs. In addition, the differences in optimal pH, substrate specificity, and subcellular localization of the various lipoxygenase isozymes suggest functional diversification (Kato et al. 1992; Dubbs and Grimes 2000; Fuller et al. 2001). The distinct biological functions of vegetative soybean LOXs were substantiated at the molecular, biochemical, metabolic, and structural levels (Fischer et al. 1999). The diverse expression patterns and protein functions of the duplicated lipoxygenase gene family could be indicative of a process of subfunctionalization (Moore and Purugganan 2005). This observation is consistent with the observation that the majority of rice gene families show highly diversified expression patterns compared to homeologs produced by whole-genome duplication (Jang et al. 2007, 2008; Yim et al. 2009).

The majority of the duplicated genes were conserved between the two homeologous regions, and lipoxygenase genes were present as tandem duplications (Shin et al. 2008). In addition, the 175-bp fragment-like sequence was conserved in the homeologous region of Williams 82, despite its absence in the 5′ region of the *Lx2* gene. Moreover, the presence of the fragment in the majority of the wild soybeans (75%) indicates that the fragment was originally present in soybean because *G. soja* is believed to be the wild progenitor of *G. max* (Fig. 4). In contrast, only 17% of the *G. max* cultivars contained the fragment. This fragment may have become rare in current soybean cultivars because of their narrow genetic base and the small number of ancestral lines used for soybean breeding and domestication (Li and Nelson 2002; Hyten et al. 2006). Genetic bottlenecks are, therefore, responsible for causing differences in allele frequency between *G. max* and *G. soja*. The fragment could have been present in the *Lx2* region prior to the recent genome duplication event in soybean and would have then undergone chromosomal duplication. Later, one of the duplicated fragments became rare in cultivated soybeans.

Three possible scenarios can be proposed to surmise the origin of the 175-bp fragment prior to a recent duplication in soybean (Fig. 4). In scenario A, the fragment in the ancestor genome was duplicated by autopolyploidy and deleted from one of the duplicated genomes. Scenario B assumes that the fragment was inserted into one of the genomes after autopolyploidization of an ancestral genome that did not contain the fragment. In a third possible sequence of events, scenario C describes an allopolyploidy state involving two genomes that differed regarding the presence of the fragment. Recently, Gill et al. (2009) provided evidence of two subgenomes by showing two different types of centromeric repeats in the soybean genomes. Although the present data are insufficient to determine the true sequence of events with certainty, scenario C appears to be the most likely because it does not include the insertion or deletion of the fragment.

A missense mutation in the lipoxygenase-2

Plant lipoxygenases consist of two domains, an N-terminal PLAT domain (pfam01477) and a C-terminal lipoxygenase domain (pfam00305). Six histidine residues in the C-terminal lipoxygenase domain are highly conserved across plant and mammalian lipoxygenase sequences, and replacement of conserved histidine residues results in loss of enzyme activity (Steczko et al. 1992; Weaver 2000; Segraves et al. 2006). At least two of these, His 499 and His 504 in LOX1, are required for lipoxygenase activity (Weaver 2000). LOX2 contains 865 amino acid residues, in which His 532 corresponds to His 504 in LOX1 (Fig. 1). A missense mutation at position His 532 of LOX2 was found in Kyushu 111 and may be responsible for the loss of

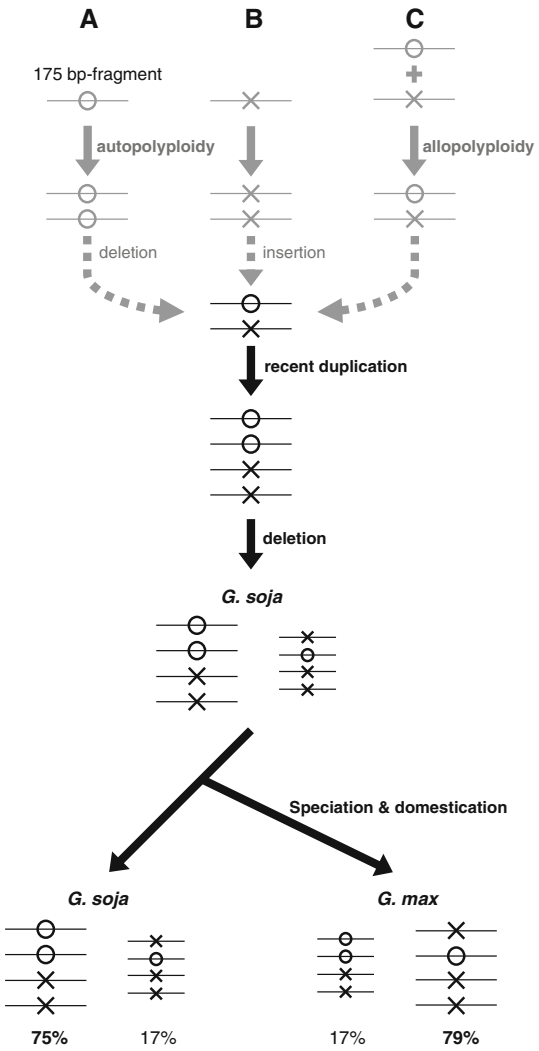

Fig. 4 Possible scenarios explaining the appearance of the novel sequence fragment. The fragment was found to be present in the soybean genome prior to the recent duplication. The distribution of the fragment among various soybean varieties was different between *G. max* and *G. soja*. Major genotypes and their corresponding percentages are represented by *larger* figures. The majority of *G. soja* (75%), the proposed wild progenitor of soybean, contained the fragment (Table 1). On the other hand, 79% of *G. max* accessions did not contain the fragment and only 17% of the soybean varieties were found to have the fragment. This fragment may have become rare in modern soybean populations because of the narrow genetic backgrounds of modern soybean cultivars and the small number of founder lines used for soybean breeding and domestication. In scenarios A and B, the soybean ancestor genome was duplicated by autopolyploidy and the fragment was deleted or inserted from the genome. In scenario C, allopolyploidy occurred between two different genomes containing and not containing the fragment. The *gray dotted lines* indicate that the real scenario cannot be determined in the present study

enzyme activity due to alterations in structure or stability (Wang et al. 1994). A single nucleotide polymorphism from T to A in the Jinpumkong 2 cultivar changed a histidine (CAT) into a glutamine (CAA) in LOX2. In the present study, allele-specific primers with 3′-terminal nucleotides complementary to the T/A SNP that were designed from the conserved histidine site successfully distinguished between wild types and mutants among 90 RILs and 480 germplasm accessions. The fact that only two out of 480 germplasm accessions had the glutamine (CAA) instead of histidine (CAT) suggests that this substitution is a rare mutation. Moreover, those two soybeans have been previously reported to lack LOX2. This rare occurrence of the mutation only found in LOX2 null cultivars substantiated that the substitution is responsible for the null LOX2 phenotype.

Inheritance of the A allele corresponded entirely with the null LOX-2 phenotype among the 90 RILs, and the presence of this allele was detected by the SNAP marker developed in this study. This result demonstrates that this marker would facilitate breeding of elite cultivars with the lipoxygenase-free seed trait by allowing early selection of LOX2 lacking lines among segregating populations. The SNAP marker detects the mutation for *lx2* by using a simple and rapid method requiring only agarose gel electrophoresis after PCR. In addition, the identified marker can be classified as a functional marker because the causative mutation was consistently linked to the null LOX2 phenotype. Therefore, the molecular marker developed in the present study would be a useful tool for high-throughput marker-assisted selection to develop soybeans with null LOX2 in a breeding program.

Acknowledgments This research was supported by a grant from the Next Generation BioGreen 21 Program (Code No. PJ0080602011), Rural Development Administration, the Republic of Korea.

References

Axelrod B, Cheesbrough TM, Laakso S (1981) Lipoxygenases from soybeans. Methods Enzymol 71:441–451

Bell E, Mullet JE (1993) Characterization of an *Arabidopsis* lipoxygenase gene responsive to methyl jasmonate and wounding. Plant Physiol 103:1133–1137

Blanc G, Wolfe KH (2004) Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. Plant Cell 16:1667–1678

Blanchette M, Tompa M (2003) FootPrinter: a program designed for phylogenetic footprinting. Nucleic Acids Res 31:3840–3842

Bunker TW, Koetje DS, Stephenson LC, Creelman RA, Mullet JE, Grimes HD (1995) Sink limitation induces the expression of multiple soybean vegetative lipoxygenase mRNAs while the endogenous jasmonic acid level remains low. Plant Cell 7:1319–1331

Davies CS, Nielsen SS, Nielsen NC (1987) Flavor improvement of soybean preparations by genetic removal of lipoxygenase-2. J Am Oil Chem Soc 64:1428–1433

Dubbs WE, Grimes HD (2000) The mid-pericarp cell layer in soybean pod walls is a multicellular compartment enriched in specific lipoxygenase isoforms. Plant Physiol 123:1281–1288

Fischer AM, Dubbs WE, Baker RA, Fuller MA, Stephenson LC, Grimes HD (1999) Protein dynamics, activity and cellular localization of soybean lipoxygenases indicate distinct functional roles for individual isoforms. Plant J 19:543–554

Forster C, North H, Afzal N, Domoney C, Hornostaj A, Robinson DS, Casey R (1999) Molecular analysis of a null mutant for pea (*Pisum sativum* L.) seed lipoxygenase-2. Plant Mol Biol 39:1209–1220

Fuller MA, Weichert H, Fischer AM, Feussner I, Grimes HD (2001) Activity of soybean lipoxygenase isoforms against esterified fatty acids indicates functional specificity. Arch Biochem Biophys 388:146–154

Gardner HW (1998) 9-Hydroxy-traumatin, a new metabolite of the lipoxygenase pathway. Lipids 33:745–749

Gepts P (2006) Plant genetic resources conservation and utilization: the accomplishments and future of a societal insurance policy. Crop Sci 46:2278–2292

Gill N, Findley S, Walling JG, Hans C, Ma JX, Doyle J, Stacey G, Jackson SA (2009) Molecular and chromosomal evidence for allopolyploidy in soybean. Plant Physiol 151:1167–1174

Hajika M, Kitamura K, Igita K, Nakazawa Y (1992) Genetic relationship among the genes for lipoxygenase-1,-2, and -3 isozymes in soybean [*Glycine max* (L.) Merril] seed. Jpn J Breed 42:787–792

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res 27:297–300

Hyten DL, Song QJ, Zhu YL, Choi IY, Nelson RL, Costa JM, Specht JE, Shoemaker RC, Cregan PB (2006) Impacts of genetic bottlenecks on soybean genome diversity. Proc Natl Acad Sci USA 103:16666–16671

Innes RW, Ameline-Torregrosa C, Ashfield T, Cannon E, Cannon SB, Chacko B, Chen NWG, Couloux A, Dalwani A, Denny R, Deshpande S, Egan AN, Glover N, Hans CS, Howell S, Ilut D, Jackson S, Lai HS, Mammadov J, del Campo SM, Metcalf M, Nguyen A, O'Bleness M, Pfeil BE, Podicheti R, Ratnaparkhe MB, Samain S, Sanders I, Segurens B, Sevignac M, Sherman-Broyles S, Thareau V, Tucker DM, Walling J, Wawrzynski A, Yi J, Doyle JJ, Geffroy V, Roe BA, Maroof MAS, Young ND (2008) Differential accumulation of retroelements and diversification of NB-LRR disease resistance genes in duplicated regions following polyploidy in the ancestor of soybean. Plant Physiol 148:1740–1759

Jang CS, Jung JH, Yim WC, Lee BM, Se YW, Kim W (2007) Divergence of genes encoding non-specific lipid transfer proteins in the Poaceae family. Mol Cells 24:215–223

Jang CS, Yim WC, Moon JC, Jung JH, Lee TG, Lim SD, Cho SH, Lee KK, Kim W, Seo YW, Lee BM (2008) Evolution of non-specific lipid transfer protein (nsLTP) genes in the Poaceae family: their duplication and diversity. Mol Genet Genomics 279:481–497

Kato T, Ohta H, Tanaka K, Shibata D (1992) Appearance of new lipoxygenases in soybean cotyledons after germination and evidence for expression of a major new lipoxygenase gene. Plant Physiol 98:324–330

Kim SD, Kim YH, Park KY, Kim HT, Lee YH, Lee SH, Seung YK, Kim HS, Hong EH, Kim YS (1997) A new beany tasteless soybean variety "Jinpumkong 2" with good quality. RDA J Crop Sci 30:112–115

Kim MY, Ha BK, Jun TH, Hwang EY, Van K, Kuk YI, Lee SH (2004) Single nucleotide polymorphism discovery and linkage mapping of lipoxygenase-2 gene (*Lx2*) in soybean. Euphytica 135:169–177

Kim KD, Shin JH, Van K, Kim DH, Lee SH (2009) Dynamic rearrangements determine genome organization and useful traits in soybean. Plant Physiol 151:1066–1076

Kitamura K (1984) Biochemical-characterization of lipoxygenase lacking mutants, L-1-less, L-2-less, and L-3-less soybeans. Agric Biol Chem Tokyo 48:2339–2346

Kitamura K, Kumagai T, Kikuchi A (1985) Inheritance of lipoxygenase-2 and genetic-relationships among genes for lipoxygenase-1, lipoxygenase-2 and lipoxygenase-3 isozymes in soybean seeds. Jpn J Breed 35:413–420

Kolomiets MV, Chen H, Gladon RJ, Braun EJ, Hannapel DJ (2000) A leaf lipoxygenase of potato induced specifically by pathogen infection. Plant Physiol 124:1121–1130

Lenis JM, Gillman JD, Lee JD, Shannon JG, Bilyeu KD (2010) Soybean seed lipoxygenase genes: molecular characterization and development of molecular marker assays. Theor Appl Genet 120:1139–1149

Li ZL, Nelson RL (2002) RAPD marker diversity among cultivated and wild soybean accessions from four Chinese provinces. Crop Sci 42:1737–1744

Moore RC, Purugganan MD (2005) The evolutionary dynamics of plant duplicate genes. Curr Opin Plant Biol 8:122–128

Narvel JM, Fehr WR, Welke GA (1998) Agronomic and seed traits of soybean lines lacking seed lipoxygenases. Crop Sci 38:926–928

Nishiba Y, Furuta S, Hajika M, Igita K, Suda I (1995) Hexanal accumulation and DETBA value in homogenate of soybean seeds lacking 2 or 3 lipoxygenase isozymes. J Agric Food Chem 43:738–741

Park HC, Kim ML, Kang YH, Jeon JM, Yoo JH, Kim MC, Park CY, Jeong JC, Moon BC, Lee JH, Yoon HW, Lee SH, Chung WS, Lim CO, Lee SY, Hong JC, Cho MJ (2004) Pathogen- and NaCl-induced expression of the SCaM-4 promoter is mediated in part by a GT-1 box that interacts with a GT-1-like transcription factor. Plant Physiol 135:2150–2161

Reinprecht Y, Luk-Labey SY, Yu K, Poysa VW, Rajcan I, Ablett GR, Peter Pauls K (2011) Molecular basis of seed lipoxygenase null traits in soybean line OX948. Theor Appl Genet 122(7):1247–1264

Romero MV, Barrett DM (1997) Rapid methods for lipoxygenase assay in sweet corn. J Food Sci 62:696–700

Schlueter JA, Dixon P, Granger C, Grant D, Clark L, Doyle JJ, Shoemaker RC (2004) Mining EST databases to resolve evolutionary events in major crop species. Genome 47:868–876

Schlueter JA, Scheffler BE, Schlueter SD, Shoemaker RC (2006) Sequence conservation of homeologous bacterial artificial chromosomes and transcription of homeologous genes in soybean (*Glycine max* L. Merr.). Genetics 174:1017–1028

Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178–183

Schwartz S, Kent WJ, Smit A, Zhang Z, Baertsch R, Hardison RC, Haussler D, Miller W (2003) Human–mouse alignments with BLASTZ. Genome Res 13:103–107

Segraves EN, Chruszcz M, Neidig ML, Ruddat V, Zhou J, Wecksler AT, Minor W, Solomon EI, Holman TR (2006) Kinetic, spectroscopic, and structural investigations of the soybean lipoxygenase-1 first-coordination sphere mutant, Asn694Gly. Biochemistry 45:10233–10242

Shin JH, Van K, Kim DH, Kim KD, Jang YE, Choi BS, Kim MY, Lee SH (2008) The lipoxygenase gene family: a genomic fossil of shared polyploidy between *Glycine max* and *Medicago truncatula*. BMC Plant Biol 8:133

Start WG, Ma Y, Polacco JC, Hildebrand DF, Freyer GA, Altschuler M (1986) Two soybean seeds nulls accumulate reduced levels of lipoxygenase transcripts. Plant Mo1 Biol 7:11–23

Steczko J, Donoho GP, Clemens JC, Dixon JE, Axelrod B (1992) Conserved histidine residues in soybean lipoxygenase: functional consequences of their replacement. Biochemistry 31: 4053–4057

Swofford DL (2003) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland

Veronesi C, Rickauer M, Fournier J, Pouenat ML, Esquerre-Tugaye MT (1996) Lipoxygenase gene expression in the tobacco-*Phytophthora parasitica nicotianae* interaction. Plant Physiol 112:997–1004

Wang WH, Takano T, Shibata D, Kitamura K, Takeda G (1994) Molecular basis of a null mutation in soybean lipoxygenase 2: substitution of glutamine for an iron-ligand histidine. Proc Natl Acad Sci USA 91:5828–5832

Wang C, Croft KP, Jarlfors U, Hildebrand DF (1999) Subcellular localization studies indicate that lipoxygenases 1 to 6 are not involved in lipid mobilization during soybean germination. Plant Physiol 120:227–236

Weaver TD (2000) The pi-helix translates structure into function. Protein Sci 9:201–206

Yim WC, Lee BM, Jang CS (2009) Expression diversity and evolutionary dynamics of rice duplicate genes. Mol Genet Genomics 281:483–493