

Cloning, Sequence Analysis, and Prokaryotic Expression of cDNA Encoding a Putative Non-Specific Lipid-Transfer Protein from the Bracts of Dovetree (*Davidia involucrata* Baill.)[†]

Yun-Xiang Li*

Department of Biology, China West Normal University, Nanchong 637002, Sichuan, China

The 967-bp clone *P1A5* was isolated from a suppression subtractive hybridization cDNA library of dovetree bracts (*Davidia involucrata* Baill.). A complete cDNA of 1047 bp was obtained via 5'-RACE (5'-Rapid Amplification of cDNA End) techniques, using the gene-specific primer P1A5-1. Northern blot analyses showed that the gene was predominantly expressed in the bracts, while the blotting signal from the leaves was weak. Its deduced amino acid sequence was most highly homologous to the lipid-transfer protein 3 precursor isolated from upland cotton, the lipid-transfer protein SDi-9 from the common sunflower, and the non-specific lipid-transfer protein precursor allergen from sweet cherry. It also had features in common with plant nsLTPs (non-specific lipid-transfer proteins), including eight conserved cysteine residues, a high isoelectric point (8.9), and a lack of tryptophans. The deduced amino acid sequence had two transmembrane helices -- the first from Position 5 (Gly) to Position 35 (Val); the second, from Position 28 (Ala) to Position 46 (Leu). A cleavage site for the putative signal peptide was predicted to occur between Positions 28 (Ala) and 29 (Ala). Therefore, the putative mature form of the protein would comprise 92 amino acids, with a molecular weight of 9.2 kD. All these results provide compelling evidence that the *P1A5* clone belongs to the nsLTP1 gene family, thus being named the *P1A5* putative nsLTP1 gene. This is the first nsLTP gene reported from Davidiaceae.

Keywords: 5'-RACE, dovetree, northern blot, nsLTP1, prokaryotic expression

Many proteins with in-vitro antimicrobial activity have been identified at the protein and/or DNA level (Garcia-Olmedo et al., 1995; Broekaert et al., 1997). Among them, the non-specific lipid-transfer proteins (nsLTP) in plants are well known for their ability to enhance the inter-membrane exchange and/or transfer of various polar lipids and glycolipids in vitro. These basic and cysteine-rich proteins are ubiquitous in the plant kingdom, forming a multigenic family. The nsLTPs are synthesized as precursors with N-terminal extensions that have the sequence characteristics of signal peptides. Some are secreted or bound to the cell walls (Kader, 1996; Douliez et al., 2000). Two main families have been characterized: 1) nsLTP1, with a molecular mass of about 9 kDa; and 2) nsLTP2, with a mass of about 7 kDa (Kader, 1996; Douliez et al., 2000). nsLTPs share a common structural fold, which is stabilized by four disulfide bonds and composed of four helices packed against a C-terminal arm that is formed by a series of turns (Gincel et al., 1994; Shin et al., 1995; Heinemann et al., 1996; Lee et al., 1998).

nsLTPs may actively participate in the biosynthesis of the cutin layer and surface wax by transporting acyl

monomers (Iiendricks et al., 1994). In addition, their expression can be induced by environmental stress factors, e.g., heat shock or NaCl treatment (Torres-Schumann et al., 1992). Several nsLTP genes in barley are also up-regulated in response to infection by various strains of fungal pathogens (Garcia-Olmedo et al., 1995). Their role in plant defenses has been confirmed by Kristensen et al. (2000), who have reported that nsLTPs isolated from the seeds of radish and onion, or the leaves of barley, maize, spinach, and sugar beet, exhibit antipathogenic activity in vitro.

Davidia involucrata Baill., an ornamental known as the Chinese dovetree or handkerchief tree, is a relic deciduous species of the Tertiary Period. As the sole member of the family Davidiaceae, it is endemic to southwestern China but now also widely introduced elsewhere (Brown and McDonald, 1978; Wyman, 1978; McClintock, 1991). This tree's special features include hanging flower clusters or designated head inflorescences, each with two (sometimes three) large, showy, white bracts. The bracts, unequal in length, are generally 3 to 6+ in. long, and surround the inflorescence. Even though the species

*Corresponding author; fax +86-817-2314-468
e-mail yx_li@263.net

[†]The nucleotide sequence data reported here has been released in the GenBank, EMBL, and DDBJ Nucleotide Sequence Databases, under the accession number AY059472.

was discovered in 1869 (McClintock, 1991), little is still known about the function of its bracts in evolutionary and reproduction biology, or their molecular development. The objective of the study presented here was to identify those molecular processes as a follow-up to research that established a subtractive hybridization cDNA library using the bract as tester and the leaf as driver (unpublished). Here, the complete cDNA of a new nsLTP1 clone, *P1A5*, was cloned, sequenced, and analyzed before examining its prokaryotic expression.

MATERIALS AND METHODS

Plant Materials

In February or April, the leaves and bracts of *D. involucrata* were collected after sprouting two to three days, respectively, at the National Natural Reserve of Longchi-Hongkou, Dujiangyang County, Sichuan Province, China. The tissues were immediately frozen in liquid nitrogen for a few minutes, then stored at -80°C .

Construction of Suppression Subtractive Hybridization cDNA Library of Bracts

Total RNAs were extracted from the frozen leaves and bracts with the Plant Total RNA Kit (TaKaRa). A suppression subtractive hybridization cDNA library, which took the bract as tester and leaf as driver, was constructed essentially as described in the manual for the CLONTECH PCR-Select™ cDNA Subtraction Kit (Clontech).

Full-Length cDNA Sequence Amplification

5'-RACE was carried out using the gene-specific primer P1A5-1 (5'-ATAACACTCTGGGCACCTGTGACTC-3'), according to the procedures for the SMART™ RACE cDNA Amplification Kit (Clontech). Briefly, 1 μg of total RNA was extracted from the bracts, then reverse-transcribed with 10 pmol 5'-CDS primer [5'-(T)₂₅N₁N-3', N = A, G, C, or T; N₁ = A, G, or C], 10 pmol SMART II A oligo (5'-AAGCAGTGGTATCAACGCAGAGTACCGCGGG-3'), and 10 pmol of dNTP mixture by PowerScript Reverse Transcriptase. The amplification was performed with primer P1A5-1 and UPM (containing two components, Long: 5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3'; Short: 5'-CTAATACGACTCACTATAGGGC-3') in an Amp 2400 thermal cycler (Perkin Elmer Cetus) for 25 cycles of 5 s at 94°C , 10 s at 65°C , and 3 min at 72°C . The single, ca. 380-

bp DNA product was then cloned into the pMD18-T vector (TaKaRa) and sequenced with a DYEnamic Direct dGTP Sequencing Kit (Amersham) and a 373A DNA sequencer.

Northern Blot Analysis

Hybridization analyses of RNAs were conducted according to the method of Maniatis et al. (1982). RNA was fractionated in a 1% (w/v) agarose gel containing 6% (v/v) formaldehyde. All RNA samples (10 μg each) were denatured in the presence of 1 mg ethidium bromide. Nylon membranes were pre-hybridized in a solution containing 50% (v/v) formamide, $5\times$ SSC, $2\times$ Denhardt's, 40 mM NaHPO₄ (pH 6.5), 10 mM EDTA, and 0.2 mg mL⁻¹ herring sperm DNA. After 12 h, a DIG-labeled probe was added, and hybridization proceeded for 16 h. Final washes were carried out with $2\times$ SSC, 0.1% SDS at 50°C . All RNA blots were performed twice with replicated samples from two independent extractions. A *P1A5* gene-specific probe spanned the deduced coding region, from 448 to 905 bp. The fragment was obtained by PCR amplification with two gene-specific primers: P1A5-2 (5'-ATAGAAA-TGGGTAGGTCAGGAGTGGTGAT-3') and P1A5-3 (5'-ACTCAAGTGAGCCTCGTGCTATCTCCATCT-3'). The DNA probe was labeled by random-priming with digoxigenin-dUTP, essentially following the procedure for the DIG High Prime DNA Labeling and Detection Starter Kit II (Roche Molecular Biochemicals).

Escherichia coli Expression Analysis

Prokaryotic expression was carried out according to the QIAexpressionist™ protocol for high-level expression and purification of $6\times$ His-tagged proteins (QIAGEN). Briefly, to obtain the coding region sequence, two primers with additional restriction digestion sites -- SphI and PstI -- were designed based on the coding region of the deduced amino acids (P1A5-2: 5'-ATAGCATGCCGTAGGTCAGGAGTGGTGAT-3' and P1A5-3: 5'-ACTCTGCAGAGCCTCGTGCTATCTCCATCT-3', respectively). After amplification, with the *P1A5* clone serving as template, the amplified product and the expression vector pQE30 were digested synchronously by SphI and PstI and purified. The ligated DNAs, comprising the insert DNA and the vector, were then used to transform competent cells of *E. coli* strain M15. Positive clones were preliminarily screened out by a colony-blot procedure. Those that contained the inserted DNA were induced for 5 h by 250 μM IPTG at 37°C in a liquid broth containing 100 $\mu\text{g mL}^{-1}$ ampicillin and 25 $\mu\text{g mL}^{-1}$

kanamycin. Positive strains without IPTG induction were set as the control. Total proteins were extracted from the induced and uninduced strains, and subjected to SDS-tricine-gel electrophoresis (Schagger and von Jagow, 1987) to validate expression of the inserted fragment. The SDS gel was stained with Coomassic solution (0.25% Coomassic Brilliant Blue R-250, 45% MeOH, and 9% HOAc in water). Afterward, the background was de-stained with a solution of 5% MeOH and 7.5% HOAc in water.

Sequence Analysis

Sequence analyses for the DNA and deduced amino acids were carried out using DNATools 5.0, OMIGA 2.0, the BLAST program, SignalP V2.0.b2, and TMHMM (v. 2.0).

RESULTS

Molecular Cloning of nsLTP1 cDNA from Dovetree Bracts

Twenty clones were randomly selected and sequenced from the suppression subtractive hybridization cDNA library. Among these, clone *P1A5*, which is 967 bp

long (Fig. 1, Lane 3) and having a poly(A) tail, was of particular interest because it was strongly expressed in the bracts and was highly homologous to the nsLTP1 genes. A single, 375-bp product containing primer sequences was obtained by 5'-RACE with universal primer UPM and the gene-specific primer P1A5-1 (Fig. 1, Lane 2). Correct amplification was confirmed because the fragment completely overlapped the original 967-bp fragment. Sequence analysis showed that the entire cDNA was 1047 bp long. It contained a complete ORF that encoded a polypeptide of 120 amino acids, with the initiation (start) codon ATG at Position 454 and a stop codon at Position 816. Other features included a 5'- (453-bp) and a 3'- (221-bp) untranslated sequence, and a long poly(A) tail (Fig. 2).

Northern Blot Analysis

Expression of the *P1A5* gene was examined by northern analysis of the sprouting bracts and leaves. A fragment that spanned the coding region (Fig. 1, Lane 1) was amplified and labeled with non-radioactive digoxigenin-dUTP. This was used as a probe for

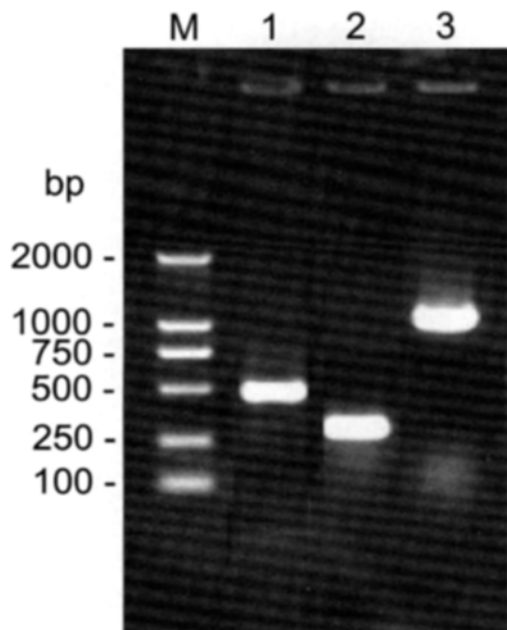


Figure 1. cDNA fragments of the *P1A5* gene from dovetree bracts. Lane 1, amplified expression fragment; Lane 2, product of 5'-RACE; Lane 3, inserted fragment of clone *P1A5* from subtractive hybridization library.

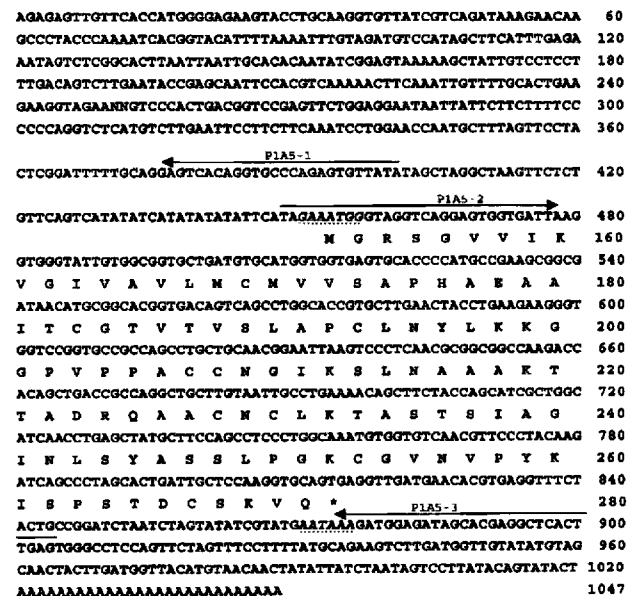


Figure 2. Nucleotide sequence of full-length cDNA of *P1A5* gene from dovetree bract, and its deduced amino acid sequence. Numbers to the right refer to nucleotide and amino acid sequences. Primers for 5'-RACE (P1A5-1) and expression analysis are indicated by arrow-headed lines. Solid, underlined letters refer to a putative polyadenylation signal, and the consensus sequence of the translation start site is indicated by dashed underline. Asterisk denotes the stop codon.

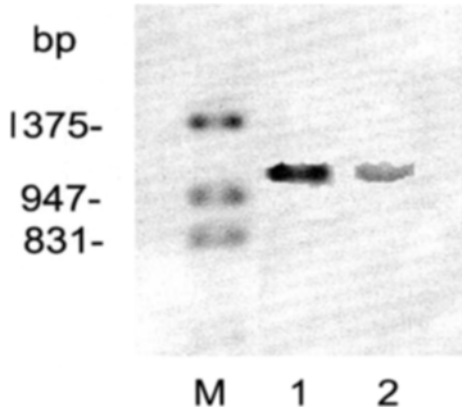


Figure 3. Northern hybridization of dovetree total RNAs, and probe of fragment spanning the coding region of the *P1A5* gene. Lane 1, bract; Lane 2, leaf.

hybridizing total RNAs from each tissue type. A single, ca. 1000-nucleotide RNA transcript was detected in both bracts (Fig. 3, Lane 1) and leaves (Fig. 3, Pane 2), with the signal intensity being much stronger in the former.

E. coli Expression Analysis

Using *E. coli*, the expression potential of the *P1A5* gene was determined in the dovetree bracts. Total proteins were fractionated according to the predicted molecular weight of the deduced peptide. A stronger band (~12 kDa) appeared in the induced strain but not in the control (Fig. 4). This indicates that the deduced polypeptide was indeed expressed in *E. coli*, although it was not the final mature form of nsLTP1s.

Homology to Non-Specific Lipid-Transfer Protein Precursors

The deduced amino acid sequence of the *P1A5* cDNA clone is 120 amino acids long, and shares homology

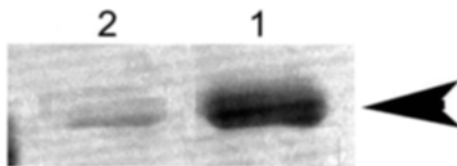


Figure 4. SDS-PAGE to examine prokaryotic expression of the *P1A5* coding region. Arrow refers to the expressed band for Lane 1. Lane 1, total protein extracted from *E. coli* strain M15 containing inserted coding region induced by IPTG; Lane 2, total proteins isolated from M15 strain containing inserted coding region but not induced.

<i>Davidia involucrate</i>	1	MGRSGVVIKVGIVAVLQCNVVSAPHAEEAITGTVTVSLAPCLMY 45
<i>Gossypium hirsutum</i>		M . ASSMSLKLACVVLQCNVVSAPLAGAVTCQVTVNSLAPCINY
<i>Prunus avium</i>		M . ACSAMTKLALVVALQCNVVSPTAGALT . CQVSSSLAPCIAY
<i>Helianthus annuus</i>		M . AKHAMVLCAGV . TCHVVGAPYTEALS . CQVSSRLAPCISY
		* *
<i>Davidia involucrate</i>	51	LKK . GGPVPPACCGIKSLNAAAKTTADROAACNCLKTASTSIA 88
<i>Gossypium hirsutum</i>		LRGSGAGAVPPGCCSGIKSLNSAAGTTPDRQAACRCCTKSAAGIT
<i>Prunus avium</i>		VRG . GGAVPPACCGIRININHLAKTTADRGATACNCLKQLSASVP
<i>Helianthus annuus</i>		LTK . GGAVPPACCGSVKSLNSAAKTTDRQAACCLKSAYNSIS
		* *
<i>Davidia involucrate</i>	101	GINLSYASSLPKCGVNVPIKISPTDCSKVQ 120
<i>Gossypium hirsutum</i>		GINFLASGLPKCGVNIPIYKISPTDCNSVK
<i>Prunus avium</i>		GVNANNAALPKCGVNVPIKISPTMCAIVK
<i>Helianthus annuus</i>		GVNAGNAASFPGKCGVSIPIYKISPTDCSKVQ
		* *

Figure 5. Comparison of the deduced amino acid sequence of the *P1A5* gene, a non-specific lipid-transfer protein precursor gene from dovetree, and three precursors of nsLTP1s from *G. hirsutum* (GenBank accession no. AAG29777; Liu et al., 2000), *P. avium* (GenBank accession no. Q9M5X8), and *H. annuus* (GenBank accession no. S71564; Ouvrard et al., 1996). Identical residues are indicated by asterisks. Gaps marked by dots have been introduced to optimize alignment.

with precursors of the nsLTP family. Its sequence was compared with the three plant nsLTP1 precursors that had the highest degree of similarity (Fig. 5), and gaps were introduced to optimize the alignment. Based on these results, *P1A5* shows 63% identity to the lipid-transfer protein 3 precursor isolated from upland cotton (*Gossypium hirsutum*; Liu et al., 2000); 59% to the lipid-transfer protein SDi-9 from common sunflower (*Helianthus annuus*; Ouvrard et al., 1996), and 57% to the non-specific lipid-transfer protein precursor ALLERGEN PRU AV3 from sweet cherry (*Prunus avium*; GenBank Accession Q9M5X8).

When the amino acid sequences of nsLTPs are compared among various plant species, only a few residues are conserved (Couthos-Thevenot et al., 1993). Among these are the 8 cysteines and 12 positions occupied by hydrophobic or aromatic residues (Broekaert et al., 1997). In the current study, the deduced amino acid sequence for the *P1A5* clone largely corresponded to this sequence homology, e.g., eight cysteines at conserved positions (Numbers 32, 42, 58, 59, 79, 81, 104, and 118 in Fig. 5), the lack of tryptophans, and a high isoelectric point (8.9).

Two strong transmembrane helices, i.e., one from Position 5 (Gly) to Position 35 (Val), and another from Position 28 (Ala) to Position 46 (Leu), were predicted by TMpred (Hofmann and Stoffel, 1993) using a predicted TM-helix length of 17 to 33. The most likely cleavage site of the signal peptide, as calculated by the SignalP V2.0.b2 program, was between Positions 28 (Ala) and 29 (Ala): AEA-AI; a similar result was obtained with the TMHMM (v. 2.0) program. The deduced

amino acid sequence had a molecular weight of 12.1 kDa. Thus, the predicted mature form, after cleavage of the signal peptide, would contain 92 amino acids, with a molecular weight of 9.2 kD. This is consistent with the data for members of the nsLTP family, meaning that *P1A5* nsLTP1 is indeed a newly cloned nsLTP1 gene predominantly expressed in dovetree bracts.

DISCUSSION

Lipid-transfer proteins (LTPs), which aid in the movement of lipid substrates between membranes *in vitro*, are widespread in the animal and plant kingdoms. In animals, LTPs can specifically bind different lipids, e.g., phosphatidylcholine and phosphatidylinositol. In contrast, those purified from plants have broad phospholipid substrate specificity that is similar to the non-specific LTPs found in animals (Wirtz, 1991). Most plant nsLTPs comprise 91 to 95 amino acid residues, and have highly conserved cysteine residue positions at 2/3-C-8-C12/15-CC-19-C-1-C-21/23-C-13-C-4/8. (N.B., "C" represents the cysteine residue, and numerals indicate the residue position of cysteine.) Although novel biological roles for plant nsLTPs have not been assigned, several have been proposed, including cutin formation, embryogenesis, defense reactions against phytopathogens, symbiosis, and adaptation to various environmental conditions (Kader, 1996).

In the study presented here, a new nsLTP1 gene was isolated from the dovetree, and was examined for its homology with members of the nsLTP1 family from other plant species. Sequence analysis revealed that the *P1A5* gene encodes a new nsLTP1 protein precursor, with a total length of 120 amino acids, and a mature nsLTP1 of 92 amino acids. In addition, prokaryotic expression was evaluated, using the putative coding region of *P1A5*. Northern blot analysis demonstrated that this gene is highly expressed in the bracts but only weakly in the leaves. Whether those expression patterns remain constant or are altered by developmental progresses requires further investigation.

The bracts of *D. involucreta* are much more vulnerable and prone to damage than are the leaves. Because the putative nsLTP mRNA was over-expressed in those bracts (presumably as a defensive protein), this activity may possibly indicate a role as a protective mechanism. A group of over-expressed mRNAs has previously been isolated from the bracts of this species, and has been characterized as part of an adversity-resistance or defense reaction. These include clone P1B4, which encodes a putative raucassicine-O- β -D-glucosidase,

and clone P2D4, encoding a putative praline-rich protein. Those patterns of over-expression may be the product of bract evolution for adapting to their inherent anatomical vulnerability.

In this laboratory, genomic walking and Southern blots are being utilized to obtain the entire gene and identify temporal and spatial patterns of expression. Future research will include identifying expression of the mature form of this LTP and examining the activity of the expressed polypeptide.

ACKNOWLEDGEMENTS

This work was supported in part by grants from Project No. 39770134 of the National Natural Science Foundation of China, and from Project 1999-164 of the Department of Education, Sichuan Province, China.

Received April 3, 2003; accepted July 21, 2003.

LITERATURE CITED

- Broekaert WF, Cammue BPA, de Bolle M, Thevissen K, de Samblanx GW, Osborn RW (1997) Antimicrobial peptides from plants. *Crit Rev Plant Sci* 16: 297-323
- Browne B, McDonald E (1978) The propagation of seed of *Davidia involucreta* Baill. Handkerchief tree. *Irish Nat J* 19: 208
- Couthos-Thevenot P, Jouenne T, Maes O, Guerbette F, Grosbios M, le Caer JP, Boulay M, Deloire A, Kader J-C, Guern J (1993) Four 9-kDa proteins excreted by somatic embryos of grapevine are isoforms of lipid-transfer proteins. *Eur J Biochem* 217: 885-889
- Douliez J-P, Michon T, Elmorjani K, Marion D (2000) Mini review: Structure, biological and technological functions of lipid transfer proteins and indolines, the major lipid binding proteins from cereal kernels. *J Cereal Sci* 32: 1-20
- Garcia-Olmedo F, Molina A, Segura A, Moreno M (1995) The defensive role of nonspecific lipid-transfer proteins in plants. *Trends Microbiol* 3: 72-74
- Gincel E, Simorre JP, Caille A, Marion D, Plak M, Vovelle F (1994) Three-dimensional structure in solution of a wheat lipid transfer protein from multidimensional 1H-NMR data - a new folding for lipid carriers. *Eur J Biochem* 226: 413-422
- Heinemann K, Andersen K, Nielsen P, Bech L, Poulsen F (1996) Structure in solution of a four helix lipid binding protein. *Protein Sci* 5: 13-23
- Hofmann K, Stoffel W (1993) TMbase - A database of membrane spanning proteins segments. *Biol Chem* 347: 166
- Iiendriks M, Meijer EA, Thoma S, Kader J-C, de Vries SC (1994) The carrot extracellular lipid transfer protein

- EP2: Quantitative aspects with respect to its putative role in cutin synthesis, *In* G Coruzzi, Puigdomenech, eds, *Plant Molecular Biology*, Springer-Verlag, Berlin, NATO ASI Series II 81, pp 85-94
- Kader J-C (1996) Lipid-transfer proteins in plants. *Annu Rev Plant Physiol Plant Mol Biol* 47: 627-654
- Kristensen AK, Brunstedt J, Nielsen KK, Roepstorff P, Mikkelsen JD (2000) Characterization of a new antifungal non-specific lipid transfer protein (nsLTP) from sugar beet leaves. *Plant Sci* 155: 31-40
- Lee JY, Min K, Cha H, Shin DH, Hwang KY, Suh SW (1998) Rice non-specific lipid transfer protein: The 1.6 Å crystal structure in the unliganded state reveals a small hydrophobic cavity. *J Mol Biol* 276: 437-448
- Liu HC, Creech RG, Jenkins JN, Ma DP (2000) Cloning and promoter analysis of the cotton lipid transfer protein gene *Ltp3*. *Biochim Biophys Acta* 1487: 106-111
- Maniatis T, Fritsch EF, Sambrook J (1982) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, NY, pp 199-201
- McClintock E (1991) *Davidia involucrata* and *Acer pentaphyllum*. *Pacific Hort* 52: 57-61
- Ouvrard O, Cellier F, Ferrare K, Tusch D, Lamaze T, Dupuis JM, Casse-Delbart F (1996) Identification and expression of water stress- and abscisic acid-regulated genes in a drought-tolerant sunflower genotype. *Plant Mol Biol* 31: 819-829
- Schagger H, von Jagow G (1987) Tricine-sodium dodecyl-sulfate polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *FEBS Lett* 3: 243-246
- Shin DH, Lee JY, Qwang KY, Kim KK, Suh SW (1995) High resolution crystal structure of the non specific lipid transfer protein from maize seedlings. *Structure* 3: 189-199
- Torres-Schumann S, Goodoy JA and Pintor-Toro JA (1992) A probable lipid transfer protein gene induced by NaCl in stems of tomato plants. *Plant Mol Biol* 18: 749-757.
- Wirtz KWA (1991) Phospholipid transfer proteins. *Annu Rev Biochem* 60: 73-99
- Wyman D (1978) The dovetree. *Amer Hort* 57: 6-7