Plant Mol Biol Rep (2011) 29:88–97
DOI 10.1007/s11105-010-0214-0

Characterization of Expressed Sequence Tags from Flower Buds of Alpine *Lilium formosanum* using a Subtractive cDNA Library

Wei-Kuang Wang · Chia-Chin Liu · Tzen-Yuh Chiang · Ming-Tse Chen · Chang-Hung Chou · Ching-Hui Yeh

Published online: 21 May 2010

Abstract Formosan lily (*Lilium formosanum*), a species endemic in Taiwan, is characterized by showy and fragrant flowers. To understand the gene expression at its reproductive phase, we constructed a suppression subtractive cDNA library of immature flower buds, from which 1,324 expressed sequence tags (ESTs) were randomly selected and sequenced. These EST sequences were clustered into 974 nonredundant sequences. Based on BLAST searching, functions of 376 sequences (39%) were determined, and 80 sequences showed high similarities to genes encoding hypothetical proteins without known functions. Another 518 sequences did not show significant homology to any known sequences and were therefore classified as novel sequences. Further analyses of the 376 ESTs sequences revealed high abundance of stress-related and flower-development genes. The highly expressing stress-related transcripts include 39 with high similarities to lipid transfer proteins, five to ascorbate peroxidases, and five to heat shock proteins 70. Using real-time quantitative RT-PCR analysis, we further revealed the expression of these three genes in the immature flower buds and in the pistils or stamens of the blooming flower of Formosan lily collected from alpine regions. These results suggest that the flower of *L. formosanum* possesses a significantly elevated level of stress genes in response to alpine environment and the ESTs analyzed here represent a valuable resource for studying a resistance mechanism of the reproductive organs of Formosan lily.

Keywords *Lilium formosanum* · Expressed sequence tags · Lipid transfer protein · Ascorbate peroxidase · Heat shock protein

Wei-Kuang Wang, Chia-Chin Liu, and Tzen-Yuh Chiang contributed equally to this work.

W.-K. Wang · M.-T. Chen · C.-H. Yeh (✉)
Department of Life Science, National Central University, Taoyuan, Taiwan
e-mail: yehch63@cc.ncu.edu.tw

C.-C. Liu
Department of Life Sciences, Tzu Chi University, Hualien, Taiwan

W.-K. Wang · T.-Y. Chiang
Department of Life Sciences, National Cheng Kung University, Tainan, Taiwan

C.-H. Chou (✉)
Research Center for Biodiversity, China Medical University, Taichung, Taiwan
e-mail: choumasa@mail.cmu.edu.tw

Introduction

Flowering is the developmental turning point from the vegetative phase to the reproductive phase. Molecular genetic analyses have been used to identify genes expressed during the floral formation, and provide much information about the genetic control of the floral development. However, the complexity of flower development remains to be explored. From a fitness standpoint, the success of flowering and subsequent fruiting will enable plants to transfer the genetic information from one generation to the next. Flowering relies on a combination of integrating effects of endogenous and external signals (Mouradov et al. 2002; Simpson and Dean 2002). However, it has been known that plants exposed to environmental stresses, such as low water availabilities, temperature fluctuations, high irradiation, pathogen infection, and nutrient deprivation, usually exhibit flowering irregularity and low seed production. To avoid damage in the reproductive organs caused by

environmental stresses, plants adapt highly efficient resistance mechanisms.

Lilies, monocotyledonous ornamental plants, belong to the *Liliaceae*, and are one of the three major flower bulbs in the floriculture industry (Robinson and Firoozababy 1993). Taxonomcially, the genus *Lilium* comprises about 96 species, which are distributed in Europe, Asia, and North America. Most of the lilies are bred and cultivated in moderate climates; only a few commercial cultivars can grow in tropical or subtropical areas. Due to slow growth and low quality of the cut flowers, the lily industry is somewhat hampered in subtropical areas. For the lily breeding, introduction of traits from wild species that adapt to the subtropical climates into the commercial cultivars will ensure the flourishing of the business.

Two species originating from subtropical areas and widely cultivated around the world are *Lilium formosanum* (Formosan lily) from Taiwan and *L. longiflorum* from south Japan and Taiwan (Shii 1983). *L. formosanum* grows across a large range of altitudes, from the coastal sand dunes to mountains of 2,500 m in elevations in Taiwan, displaying a marked geographic cline (geographically continuous variation), whereas *L. longiflorum* is geographically distributed in the coastal lands of Taiwan and the Ryukyu (Shii 1983). In contrast to the imported lily *L. hybridium* with high temperature and pathogen sensitivities, Formosan lily shows high resistance to drought, pathogens, and climatic fluctuations. *L. formosanum* is a long day plant. Nevertheless, flowering time heterogeneities were observed in wild populations. Plants growing at the plains generally flower from March to June, whereas in the high mountains, they do not flower until July to October. Besides, environmental factors, such as temperature, light intensity, and nutrients of the growth environment can all affect the flowering of wild *L. formosanum* (Shii 1983).

Formosan lilies are characterized by showy and fragrant flowers. Under disturbed environments, they usually show higher resistance than *L. longiflorum*, and can adjust flowering time to avoid the unfavorable effects on the reproductive organs (Hiramatsu et al. 2002). However, few studies have been conducted to investigate the gene expression patterns and functions during the reproductive phase of *L. formosanum*, mostly because of the large genome of lilies and the inefficient transformation system. To examine these characteristics of Formosan lily flower in the field, we used suppression subtraction hybridization to identify and isolate differentially expressed genes in the immature flower buds, including those associated with developing microspores and tapetum cells, of *L. formosanum*. We determined the partial sequences of 1,324 randomly selected cDNAs from the expressed sequence tag (EST) library. Here, we aimed to discover the genes expressed in reproductive organs of *L. formosanum*.

Materials and Methods

Plant Materials and cDNA Library Construction

Plants of *L. formosanum* were collected from high mountains at elevations above 2,500 m in Taiwan (average temperature: 7~15°C; average rainfall: 2,100 mm) and were kept in liquid nitrogen. Flower buds 1~2-cm long and young, healthy leaves were collected for total RNA preparation. Seeds collected from the same area were allowed to germinate and grow in a greenhouse (14-h light/10-h dark; 20~28°C). Samples were ground in liquid nitrogen. Total RNA was extracted from lily immature flower buds and leaves with use of the SV total RNA isolation system (Cat. No. Z3100, Z3101 & Z3105, Promega, Madison, WI, USA), and poly $(A)^+$ RNA was isolated by use of an Oligotex kit following the manufacturer's instructions (Cat. No. 70042, Qiagen, Chatsworth, CA, USA). Suppression subtractive hybridization involved use of the polymerase chain reaction (PCR)-Select cDNA subtraction kit (Cat. No. 70042, Clontech, Palo Alto, CA, USA). Tester cDNA was synthesized from mRNA of immature flower buds of *L. formosanum*, and driver cDNA was from mRNA of leaves. Products from the secondary PCR were cloned into a pGEM-T Easy vector (Promega, Madison, WI, USA) and transformed into *Escherichia coli* JM109 cells.

DNA Sequencing and Bioinformatic Analyses

Sequencing of cDNA involved use of the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction kit (PerkinElmer Life Sciences, Foster City, CA, USA) with the SP6 vector primers on an ABI 377 automated sequencer (Perkin-Elmer Applied Biosystems, Foster City, CA, USA). All cDNA and deduced amino acid sequences underwent a BLAST search of the National Center for Biotechnology Information (NCBI) database for similarities to known sequences (Altschul et al. 1997). To identify nonredundant EST sequences, we clustered the EST sequences. The EST sequences were aligned by use of the program Genetics Computer Group (GCG) Wisconsin Package (v10.0; Madison, WI, USA). The sequences showing greater than 95% identity over 100 bp were categorized as the same gene. Each EST sequence was used for a similarity search of the nonredundant database provided by NCBI with use of the BLASTx program. Sequence similarity was considered significant at an expectation value (E) lower than 1×10^{-5}.

Analysis of Transcript Abundance

Samples were frozen in liquid nitrogen and ground into powder by use of a mortar and pestle. Total RNA was

extracted as described by Chang et al. (1993). The quantity of total RNA was determined by optical density measurement with use of an Eppendorf BioPhotometer and verified on a 1% agrose gel. To remove genomic DNA contamination, total RNA used in cDNA synthesis was treated with DNase I (Promega, Madison, WI, USA) for 30 min at 37°C before cDNA synthesis. The first-strand cDNA was synthesized from 2 μg total RNA in a 20-μl reaction volume with use of the SuperScriptIII First-Strand Synthesis system (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol.

Transcript levels were analyzed by real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR) by use of the iCycler- iQ5 Multicolor Real time PCR Detection System and iQ SYBR Green Hot-start Supermix (Bio-Rad, Hercules, CA, UA). The primers used in this study were designed by use of Beacon Designer (Premier, Palo Alto, CA, USA) and are listed in Table 1. Real time PCR reaction was carried out using the prepared cDNA (60 ng) with each set of primer and probe and iQ™ SYBR® Green Supermix (Cat. No. 170-8882, Bio-Rad, Hercules, CA, USA). PCR cycling was at 95°C (10 s), 56°C (30 s), and 72°C (20 s). Three independent replicates were performed for each sample. The comparative C_T method was used to determine the relative amount of *LfLTP*, *LfAPX*, and Lf*HSP70* in *L. formosanum* plants, with the expression of *18S rRNA* used as an internal control.

Isolation of Full-Length cDNA

Rapid amplification of cDNA ends (RACE) was used to isolate the complete sequence of the lipid transfer protein gene. One microgram of mRNA isolated from the flower buds was converted into 5′- and 3′-RACE-ready cDNAs with the 5′ and 3′ CDS primers by use of the SMART RACE cDNA amplification kit (Clontech, Palo Alto, CA, USA). According to the partial sequence of the lipid transfer protein gene (*LfLTP*) of the EST clone, specific primers ltp5f (5′- TCTAGCATAGCCAAGGCAGGA-3′) and ltp3r (5′- CTGCTATCGTCGCCGGCATC-3′) were designed for amplification of the 5′ and 3′ ends, respectively. All PCR products were cloned into pGEMT-T Easy vector. Plasmid DNAs purified from overnight cultures of three independent clones were sequenced for each transformation, and all resulting sequences were aligned with the partial cDNA sequence by use of the GCG program.

Phylogenetic Analysis

Nucleotide sequences were aligned with use of the program ClustalW (Thompson et al. 1994). Neighbor-joining analyses of amino acid sequences involved use of MEGA4 (Tamura et al. 2007) by calculating genetic distance based on Poisson correction model. Confidence of the reconstructed clades was tested by bootstrapping (Felsenstein 1985), re-sampling with 1,000 replicates. As a rule, nodes with bootstrap values greater than 70 are significantly supported with 95% probability (Hillis and Bull 1993). The Genbank accession numbers for amino acid sequences of the *LTP* gene in other plants used are almond (ACH58427), peach (AAV40850), pear (AAF26451), strawberry (AAY83341), grapevine (ABA29446), upland cotton (ACI26696), *Arabidopsis thaliana* (NP183188), broccoli (AAA32995), rough lemon (BAH03575), carrot (AAB96834), *Solanum* (ABH03042), tobacco (AAM74206), *Capsicum* (ACB05670), potato (ABU49730), tomato (AAB42069), *Salvia* (ABP01768), *Lilium longiflorum* (AAD46683), maize (AAB06443), wheat (ABF14725), rice (AAC18567), barley (AAV49759), and *Bromus* (AAL23748).

Results

Characterization of EST Sequences

Here, we used cDNA sequencing of ESTs to identify genes expressed in the flowers of *L. formosanum*. A total of 2,500 cDNA colonies were randomly selected. The 5′ ends of the 1,324 clones were obtained for further analysis after poor quality sequence data were eliminated. The average length of the EST sequences was 300 bp. Sequence comparison of the 1,324 ESTs by use of the GCG program generated 974 nonredundant genes, so about 26% of the ESTs are redundant, a result similar to those of previous

Table 1 Oligonucleotide primers used in real-time quantitative RT-PCR

Gene	Primer	Sequence	Products (bp)
LfLTP	Forward Reverse	5′-GCTGGTACTACGCATCTG-3′ 5′-CCACCATAGAAGATAATGAGTC-3′	169 bp
LfAPX	Forward Reverse	5′-ACTTTTGGGTGGGGAGAAGGAAGG-3′ 5′-GTGAGCACCCAAGATTACCAGAGC-3′	245 bp
LfHSP70	Forward Reverse	5′-GTACAAGTCTGAGGAT-3′ 5′-GCTCCTTCATCTTGTC-3′	224 bp
18S	Forward Reverse	5′-GTGACGGGTGACGGAGAATTA-3′ 5′-ACACTAAAGCGCCCGGTATTG-3′	148 bp

studies of the EST library from mature and immature flowers (25% and 20% redundancy, respectively; Endo et al. 2000, 2002). All 1,324 EST sequences were deposited in the dbEST division of GeneBank (Accession numbers: GW589879–GW591202). Furthermore, the EST sequences underwent a similarity search by use of the BLASTx program against the nonredundant database provided by NCBI, and 698 of the 1,324 ESTs (53%) were identified. Among them, 117 showed high similarities with putative protein or predicted mRNA, the functions of both not yet determined. Most unique genes (764; ~58% of total ESTs) were found as singletons, which indicate the low redundancy of the libraries and suggests that lily buds are a rich source for gene mining. The level of redundancy varied, from 2 to 4 ESTs per contig, 148 contigs detected; to 5-7 ESTs per contig, 11 contigs found; to ≥8 ESTs per contig, five contigs found.

Putative Functional Categorization of Unique Transcripts of ESTs

According to the functional catalog database (http://mips.gsf.de) of the Munich Information Center for Protein Sequences, the sequences of immature flowers of Formosan lily bud ESTs were classified into 17 groups (Fig. 1). The putative functions of 39% of the 974 unique genes could be assigned on the basis of similarity to plant ESTs and the annotated *Arabidopsis* genes. Detailed categorization of these genes is in Fig. 1. The results showed that 7% of EST sequences were represented by proteins putatively involved in plant metabolism. The genes encode proteins that are putatively involved in protein synthesis (5%); protein destination (4%); cell rescue and defense (3%); cellular organization (3%); energy (3%); transcript facilitation (3%); signal transduction (2%); transcription (2%); cellular transport and transport mechanisms (2%); cell growth, division and DNA synthesis (1%); cellular biogenesis (1%); development (1%); and ionic homeostasis (1%); or are plasmid proteins (1%). However, the remaining 61% showed similarity to unclassified proteins (8%) and proteins with unknown function (53%). Thus, 376 EST sequences (39%) showed sequence similarity to genes encoding proteins with identified function. The high proportion of unknown sequences suggests that the immature flower bud of Formosan lily is an intriguing source of novel genes.

Table 2 lists the number of EST clones more than five in our library. The most frequent redundant contig in the database contains 39 ESTs in the library, mostly corresponding to LTP, which was found to participate in cutin formation and the adaptation of plants to various

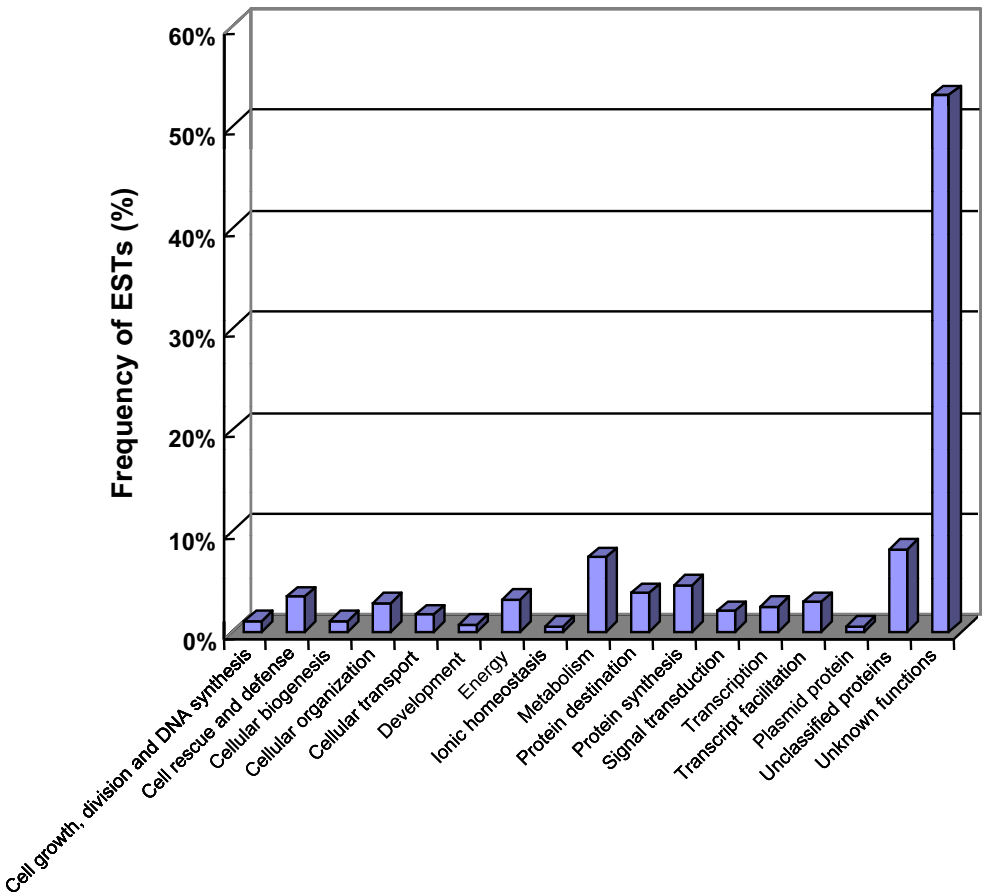

Fig. 1 Functional categorization of the BLAST results of the subtractive floral ESTs from *Lilium formosanum*. Relative frequencies of ESTs assigned to predicted functions were expressed as the percentage of sequenced ESTs

Table 2 ESTs recovered from multiple clones of a suppression subtractive cDNA library derived from immature flower bud tissue of alpine *L. formosanum*

Putative Function	Reference organism	GI number	*E* value	ESTs
Lipid transfer protein	*Lilium longiflorum*	5670318	1.0E-173	39
Pectate lyase	*Lilium longiflorum*	19450	1.0E-166	16
MADS domain transcription factor	*Gnetum gnemon*	257165011	2.0E-24	15
MADS-box protein	*Capsicum annuum*	8574456	1.0E-15	14
Pollen-specific LLP-B3 protein	*Lilium longiflorum*	25990490	1.0E-166	10
Elongation factor-1	*Lilium longiflorum*	4680248	1.0E-162	7
Shaggy-like kinase 59	*Nicotiana tabacum*	2598602	5.0E-21	7
70 kDa heat shock protein mRNA	*Sandersonia aurantiaca*	19172402	1.0E-51	5
Cytosolic ascorbate peroxidase	*Nicotiana tabacum*	1389653	4.0E-24	5
Mitochondrial half-ABC transporter gene	*Arabidopsis thaliana*	9187882	2.0E-15	5
Methionine sulfoxide reductase	*Fragaria x ananassa*	11342532	3.0E-11	5

environmental stresses. A total of 69 stress-related ESTs (16% of known function proteins) were further collected and are listed in Table 3. In addition to finding many stress-related ESTs in the subtractive cDNA library, we identified 2 EST contigs as putative MADS-box genes encoding the transcription factors for controlling a variety of flower developmental processes (Ng and Yanofsky 2001). Both ESTs had high redundancy in the EST library: 15 and 14, respectively (Table 2). In angiosperms, the floral organs are typically arranged in four separate whorls—sepals, petals, stamens, and carpels—identified by the floral homeotic genes of classes A, A+B, B+C, and C, respectively, in the well-known ABC model; most of the genes belong to the MADS box family of genes. The two MADS-box genes we screened were identified as members of the B-class genes.

Expression Patterns of Stress-related Genes in the Flower Bud of Formosan Lily

Three stress-related sequences—*LfLTP*, cytosolic ascorbate peroxidase (*LfAPX*; 5 ESTs), and 70-kDa heat shock protein (*LfHSP70*; five ESTs)—were selected as probes for studying their expression levels in Formosan lily plants that live in the high mountains at elevations above 2,500 m. Real-time qRT-PCR analysis indicated that the transcript levels of *LfLTP*, *LfAPX*, and *LfHSP70* in the immature flower bud were 6.5~39.0-fold, 3.2~5.0-fold, and 3.9~12.2-fold higher than those in root, stem, and leaf, respectively (Fig. 2). From these results of real-time qRT-PCR analyses, we can conclude that *LfLTP*, *LfAPX*, and *LfHSP70* are highly expressed in reproductive tissue of Formosan lily plants grown in high mountains.

To further understand the roles of stress-related genes in the flowers of Formosan lily, we used real-time qRT-PCR to compare the transcript levels of *LfLTP*, *LfAPX*, and *LfHSP70* in blooming flowers (5~8 cm) of alpine Formosan lilies collected from high mountains and the greenhouse. The effect of cold treatment (12°C for 24 h) on induction of these three stress genes in greenhouse-grown alpine Formosan lily flowers was also examined. As shown in Fig. 3A, the expression of *LfLTP* in the pistil of high mountain-grown alpine lily flowers was 74-fold higher than

Table 3 Possible stress-related ESTs recovered from a suppression subtractive cDNA library derived from immature flower bud tissue of alpine *L. formosanum*

Stress related ESTs	Reference organism	GI number	*E* value	ESTs
Lipid transfer protein	*Lilium longiflorum*	5670318	1.0E-173	39
Molecular chaperone Hsp90	*Lycopersicon esculentum*	38154492	4.0E-5	1
70 kDa heat shock protein mRNA	*Sandersonia aurantiaca*	19172402	1.0E-51	5
Water channel-like protein	*Arabidopsis thaliana*	20148528	2.0E-51	2
Cytosolic ascorbate peroxidase	*Nicotiana tabacum*	1389653	4.0E-24	5
Abscisic stress ripening protein	*Prunus armeniaca*	2677823	6.0E-20	3
Methionine sulfoxide reductase	*Fragaria x ananassa*	11342532	3.0E-11	4
Plasma membrane H^+ ATPase	*Lilium longiflorum*	13641332	1.0E-114	4
Calmodulin	*Lilium longiflorum*	308899	2.0E-37	1
Alcohol dehydrogenase 1	*Vitis vinifera*	9885271	2.0E-12	2
Metallothionein-like protein	*Prunus persica*	5931762	1.0E-8	1
Amino acid selective channel protein	*Hordeum vulgare*	3758826	1.0E-14	1
Chromatin remodeling complex ATPase	*Arabidopsis thaliana*	14334971	4.0E-13	1

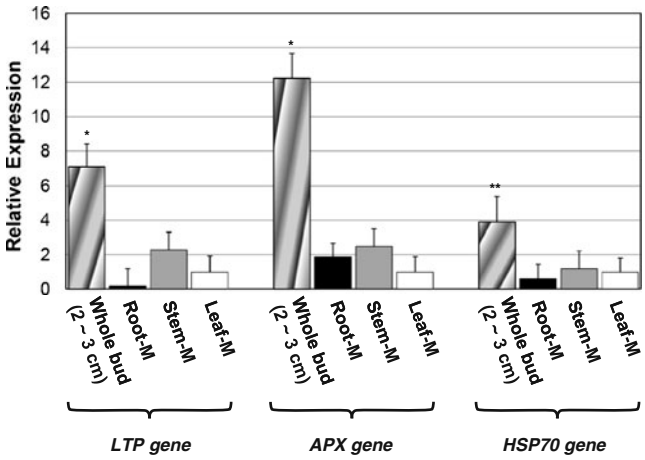

Fig. 2 Expression levels of highly abundant stress-related transcripts in the premature floral bud of *L. formosanum*. Total RNA was isolated from ~2 to ~3-cm floral bud, root, stem, and leaf of high mountain-grown (M) alpine *L. formosanum*. Real-time quantitative RT-PCR was performed to assay the gene expression of *LfLTP*, *LfAPX*, and *LfHSP70* genes as indicated. The resulting mean values are relative to the expression of *18S* rRNA. *Single asterisk* and *double asterisk* represent data points significantly different from their corresponding control ($P < 0.01$, 0.05, respectively, student's *t* test; $n=6$)

in pistils of plants grown in the greenhouse. Elevated *LfLTP* expression was also detected in the perianth of alpine-grown lily plants. In contrast, the expression of *LfAPX* and *LfHSP70* was higher (13.8-fold and 4-fold induction, respectively) in stamens of flowers grown in alpine regions than in stamens of those grown in the greenhouse (Fig. 3B and C). Regarding the effect of cold stress on induction of *LfLTP*, *LfAPX*, and *LfHSP70*, low temperature was shown to induce the accumulation of *LfLTP* transcripts in the pistil (21.5-fold) and *LfAPX* transcripts in the stamen (4.4-fold; Fig. 3A and B) but had less effect on induction of *LfHSP70* (Fig. 3C). Compared with the expression of *LfLTP*, *LfAPX*, and *LfHSP70* in the flowers of alpine Formosan lily plants, the expression of these three stress-related genes was detected less in root, stem, and leaf from alpine- and greenhouse-grown plants. Thus, the transcripts of *LfLTP*, *LfAPX*, and *LfHSP70* mainly accumulated in pistil or stamen during the reproductive phase of alpine Formosan lily.

Phylogenetic Analysis of DNA Sequence and Deduced Amino Acid Sequence of *LfLTP*

The highest expressing transcript in our EST database was homologous to lipid transfer proteins. For further analysis, we thus screened the complete sequence of *LfLTP* from the Formosan lily cDNA library by RACE. The putative transcription start and stop sites of *LfLTP* were amplified with gene-specific primers designed from the partial cDNA sequence. The transcripts defined the full-length cDNA sequence with putative open reading frames of 342 bp, encoding predicted proteins of 113 amino acids (Fig. 4). The predicted molecular mass and pI value of LfLTP were 11.6 kDa and 8.98, respectively. The sequence data have been deposited at Genbank under Accession number FJ883019.

The DNA sequence and deduced amino acid sequence of Formosan lily LTP showed similarities to members of the protease inhibitor/seed storage/LTP family from *Arabidopsis*, rice and other plant species. Formosan lily LTP contains the plant LTP signature domain (amino acids 91–112: IPakCg. Vn.IpYp.ISmqtDCnkV) as detected by the Motif scan program implemented in the PROSITE database (http://myhits.isb-sib.ch/cgi-bin/motif_scan; Falquet et al. 2002). A signal peptide of 22 amino acids was predicted by SignalP (http://www.cbs.dtu.dk/services/SignalP-2.0/), which is consistent with signal peptides found in LTPs from other species (Kader 1996; Nielsen et al. 1997). Similar to other plant LTPs, Formosan lily LTP has a conserved motif of eight cysteine residues that were suggested to form intra-molecular disulfide bridges, which, in turn, are important for LTP structure and function (Takishima et al. 1988; Liu et al. 2005; Fig. 4). The mature predicted protein without the signal peptide was 90 amino acids long, which corresponds to a molecular mass of 9.3 kDa.

We constructed an amino acid neighbor-joining tree for *LfLTP* (Fig. 5). Twenty-two amino acid sequences of various plants were acquired from the Genbank database. All of these proteins have an AAI_LTSS domain, for *A*lpha *A*mylase *I*nhibitor and *L*ipid *T*ransfer and *S*eed *S*torage proteins. The AAI_LTSS domain is found in plant trypsin-alpha amylase-inhibitor, plant lipid-transfer and seed-storage proteins. The LTPs facilitate the transfer of lipids between natural or artificial membranes. We performed a pairwise comparison of the degree of amino acid identity for the LTPs. The identity ranged from 54% (for the barley–pear, rice–broccoli, and rice–broccoli pairs) to 98% (for the *L. formosanum–L. longiflorum* pair), which shows high conservation of LTPs in plants. In the LTP amino acid tree, the *LTP* of *L. formosanum* (*LfLTP*) is closely related to *L. longiflorum* (*LlLTP*), with high bootstrap value. The *LTP* of lily clustering with other monocots is a sister to a clade of other dicots. The phylogeny of the LTP group suggests that a single-class LTP gene existed before the divergence of monocots and dicots (Springer and Kaeppler 2005).

Discussion

Growing at habitats from low lands to mid-elevation mountains in Taiwan, Formosan lily provides a perfect material for analyzing gene and environment interactions. In contrast to low land populations, plants growing at mountain areas usually face severe climatic conditions, and

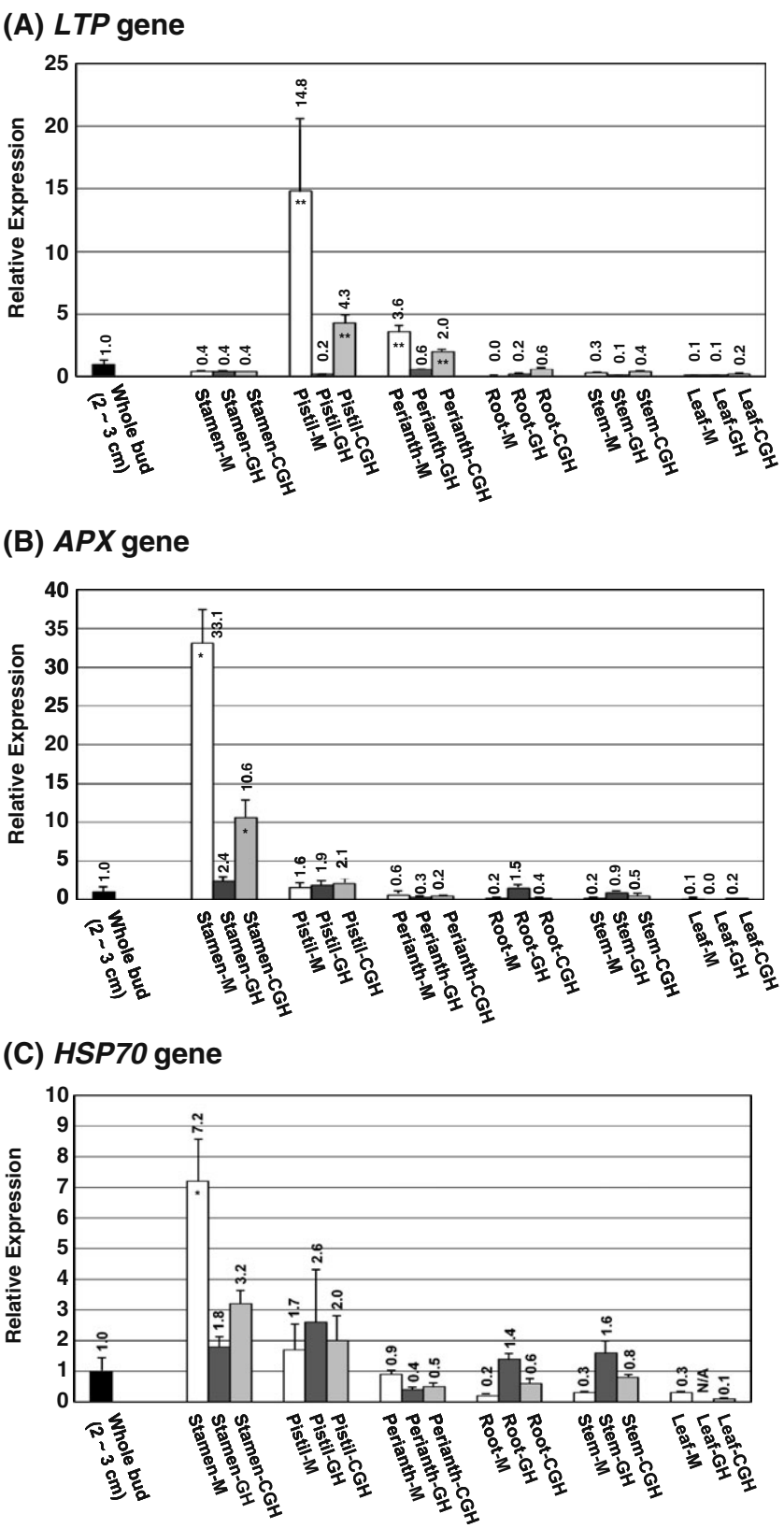

Fig. 3 Expression levels of highly abundant stress-related transcripts in blooming flowers of *L. formosanum*. Total RNA was isolated from 5~8-cm flower, root, stem, and leaf of high mountain (*M*), greenhouse (*GH*), and cold-treated greenhouse-grown (*CGH*) alpine *L. formosanum*. Real-time qRT-PCR was performed to assay the gene expression of *LfLTP* (**A**), *LfAPX* (**B**), and *LfHSP70* (**C**) genes as indicated. The resulting mean values are relative to the expression of *18S* rRNA signal value. *Asterisk* and *double asterisk* represent data points significantly different from their corresponding control ($P < 0.01$, 0.05, respectively, student's *t* test; $n=6$)

```
1   ATGGCTCGCTCCTCCGCCGTCTGCTTCCTCCTCCTCCTCGCCTTCCTTATTGGCACA
      M  A  R  S  S  A  V  C  F  L  L  L  L  A  F  L  I  G  T
    ----------------------------------------------------------
58  GCCTCGGCAATCACCTGTGGTCAGGTTGACTCTGACCTCACCTCCTGCCTTGGCTAT
      A  S  A  I  T  C  G  Q  V  D  S  D  L  T  S  C  L  G  Y
    -----
115 GCTAGAAAGGGCGGGGTCATCCCACCGGGCTGCTGCGCGGGTGTGAGGACCCTTAAC
      A  R  K  G  G  V  I  P  P  G  C  C  A  G  V  R  T  L  N
172 AACTTAGCCAAGACGACTCCTGATCGCCAGACTGCATGCAACTGCCTCAAGTCTCTG
      N  L  A  K  T  T  P  D  R  Q  T  A  C  N  C  L  K  S  L
229 GTGAACCCCAGCCTTGGCCTCAATGCTGCTATCGTCGCCGGCATCCCCGCCAAGTGC
      V  N  P  S  L  G  L  N  A  A  I  V  A  G  I  P  A  K  C
286 GGCGTCAACATCCCCTACCCGATCAGCATGCAGACTGATTGCAACAAGGTGAGGTAA
      G  V  N  I  P  Y  P  I  S  M  Q  T  D  C  N  K  V  R
```

Fig. 4 Analysis of cDNA sequence encoding a lipid transfer protein isolated from Taiwan lily, *L. formosanum*. Nucleotide and deduced amino acid sequences of the lipid transfer protein. Putative N-terminal signal sequence underlined and likely cleavage site indicated by arrow. The eight cysteine residues characteristic of CcLTPs are marked by bold italics

therefore have adapted to combat the combined effects of cold, heat, drought, and high irradiation. *L. formosanum* plants present acclimations to the extremes in the natural habitats by developing efficient protective mechanisms, especially at the reproductive phase (Streb et al. 1997). Therefore, researches showing the high stress tolerance of Formosan lily would provide a wealth of information about phenotypic variation for the lily breeding.

In this study, we used genomics and bioinformatics approaches to identify stress-related genes that are expressed specifically in reproductive organs of *L. formosanum*. A suppression subtraction hybridization EST library was generated from immature flower buds. These analyses revealed a great number of stress-related genes and flower development genes in the constructed ESTs (Tables 2 and 3). Furthermore, the highly expressed stress-related genes (*LfLTP*, *LfAPX*, and *LfHSP70*) were further identified. All of them are physiologically specific to reproductive tissues, especially in pistils or stamens (Figs. 2 and 3). Here our results show that *L. formosanum* may provide many genes useful for improving environmental resistance of cultivated lily.

Base on the results of real-time qRT-PCR analyses, we found that the transcript levels of three stress-related genes in the immature flower bud of alpine Taiwan native lily were higher than those in root, stem, or leaf. Among these genes, LTPs, widely found in higher plants, are characterized by their abilities to bind and transfer lipid molecules between membranes in vitro with low specificities (Kader 1996; Douliez et al. 2000). They have been implicated in a plant defense mechanism against environmental stresses such as high and low temperature or drought stress (Molina et al. 1996; Trevino and Oconnell 1998; Gaudet et al. 2003) and various pathogens (Jung et al. 2003; Park et al. 2003). LTPs, found in rice, strawberry, and sunflower, are induced by abscisic acid or salicylic acid, which mediates plant responses to abiotic stresses (Ouvrard et al. 1996; Garcia-Garrido et al. 1998; Yubero-Serrano et al. 2003). However, their biological function in plants is still unclear. Cytosolic APX plays a key role in protecting chloroplasts against UV damage and providing cross-compartment protection for thylakoid and stromal/mitochondrial APX in *Arabidopsis* plants during light stress (Shigeoka et al. 2002; Davletova et al. 2005). Suppression of cytosolic APX activates the protection mechanism for heat and salinity in tobacco cells

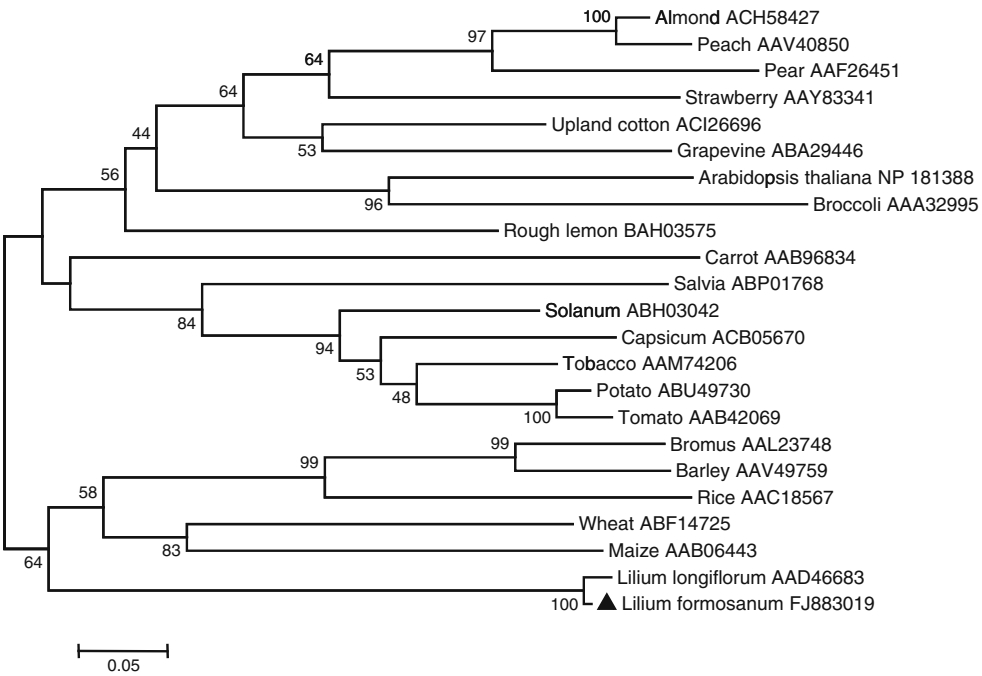

Fig. 5 Phylogenetic tree of *L. formosanum* LTP with 22 other plant LTPs based on deduced amino acid sequences. The alignment involved use of the ClustalW program. A phylogenetic tree was created from the aligned sequences by the Neighbor-joining method with the MEGA program. The bootstrap values are shown on the node. Only bootstrap values greater than 40 are shown

(Ishikawa et al. 2005). As well, we found the ESTs of a 70-kDa heat shock protein (HSP) and molecular chaperone HSP90 but no small HSP (sHSP) genes, although sHSPs were found in a rose flower-related EST library constructed without suppression subtraction hybridization (Guterman et al. 2002). Therefore, the expression of *sHSP* genes might be nearly equal in the flower buds and leaves of lily. Furthermore, as compared with the expression of *LfLTP*, *LfAPX*, and *LfHSP70* in alpine lily roots, stems, and leaves and in greenhouse lily flowers, roots, stems, and leaves, their expression in mountain-grown Formosan lily pistils or stamens is high during the reproductive stage. Similar induction patterns of these three stress-related genes were also found in cold-treated greenhouse-grown Formosan lily flowers, with the accumulation levels less than those of high mountain-grown Formosan lily flowers. The high expression of *LfLTP*, *LfAPX*, and *LfHSP70* may have been caused by not only cold stress but also the combined effects of varied environmental stresses (drought, temperature, UV, and pathogens) in the high mountain regions. Phylogenetic analysis indicated that *LfLTP* was closely related to *LlLTP* (Fig. 5) This result coincides with that *L. formosanum* and *L. longiflorum* are shown to be genetically close (Hiramatsu et al. 2002). However, *L. formosanum* is characterized by its high resistance to environmental adverse conditions and widespread distribution across an altitudinal gradient. Therefore, the expression of these stress-related genes may play pivotal roles in the adaption of Formosan lily in alpine regions.

Floral organ identity genes are divided into three classes of homeotic genes (A, B, and C class genes), depending on which organs they affect (Coen and Meyerowitz 1991). They act alone or together to give rise to sepals, petals, stamens, and carpels. Morphologically, the two outer whorls of the lily flower are similar and generate a combination of organ tepals instead of sepals and petals. This structure similarity could be ascribed to a modification in the ABC model, thus leading to an extension of class B gene function towards the first whorl in *Liliaceae* species (Theissen et al. 2000). The two putative lily class B genes are therefore good candidates for the study of morphological and developmental modification (Table 2). As well, we found putative genes for floral scent, like *O-methyltransferase* genes, which could be involved in biosynthesis of rose flower scent compounds (Wu et al. 2003). These results suggest that this library is a good source for studying useful characters in the lily breeding.

We generated 974 *L. formosanum* ESTs that provide information on gene expression patterns involved in the plant's flower development and resistance. From a functional categorization analysis of ESTs, followed by real-time qRT-PCR analysis, we found a high expression of genes involved in the stress response and flower development. The cDNAs identified from the subtractive cDNA library include three stress-related genes (*LfLTP*, *LfAPX*, and *LfHSP70*) and two MADS-box genes, all of which appear to be promising candidates for playing key roles in the development and resistance of flowering in alpine *L. formosanum* plants.

Acknowledgments We are grateful to Dr. THD Ho for comments and suggestions on the manuscript. We also thank KC Tsai for help with preparing the manuscript. This work was supported in part by a grant from the R.O.C. Agriculture Council to CHC and a grant from the R.O.C. National Science Council NSC94-2311-B-008-004 to CHY.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. Plant Mol Biol Reporter 11:113–116

Coen ES, Meyerowitz EM (1991) The war of the whorls: genetic interactions controlling flower development. Nature 353:31–37

Davletova S, Rizhsky L, Liang H, Shengqiang Z, Oliver DJ, Coutu J, Shulaev V, Schlauch K, Mittler R (2005) Cytosolic ascorbate peroxidase 1 is a central component of the reactive oxygen gene network of *Arabidopsis*. Plant Cell 17:268–281

Douliez JP, Jégou S, Pato C, Larré C, Mollé D, Marion D (2000) 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

Endo M, Hakozaki H, Kokubun T, Masuko H, Takahata Y, Tsuchiya T, Higashitani A, Tabata S, Watanabe M (2002) Generation of 919 expressed sequence tags from immature flower buds and gene expression analysis using expressed sequence tags in the model plant *Lotus japonicas*. Genes Genet Syst 77:277–282

Endo M, Kokubun T, Takahata Y, Higashitani A, Tabata S, Watanabe M (2000) Analysis of expressed sequence tags of flower buds in *Lotus japonicas*. DNA Res 7:213–216

Falquet L, Pagni M, Bucher P, Hulo N, Sigrist CJA, Hofmann K, Bairoch A (2002) The PROSITE database, its status in 2002. Nucleic Acids Res 30:235–238

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Garcia-Garrido JM, Menossi M, Puigdomenech P, Martinez-Izquierdo JA, Delseny M (1998) Characterization of a gene encoding an abscisic acid-inducible type-2 lipid transfer protein from rice. FEBS Lett 428:193–199

Gaudet DA, Laroche A, Frick M, Huel R, Puchalski B (2003) Cold-induced expression of plant defensin and lipid transfer protein transcripts in winter wheat. Physiol Plant 117:195–205

Guterman I, Shalit M, Menda N, Piestun D, Dafny-Yelin M, Shalev G, Bar E, Davydov O, Ovadis M, Emanuel M, Wang J, Adam Z, Pichersky E, Lewinsohn E, Zamir D, Vainstein A, Weiss D (2002) Rose scent: genomics approach to discovering novel floral fragrance-related genes. Plant Cell 14:2325–2338

Hillis DM, Bull JJ (1993) An empirical test of bootstrapping as a method assessing confidence in phylogenetic analysis. Syst Biol 41:182–192

Hiramatsu MH, Okubo H, Yoshimura K, Huang KL, Huang CW (2002) Biogeography and origin of *Lilium longiflorum* and *L. formosanum* II - Intra- and Interspecific variation in leaf morphology, and in flowering rate and individual net production during the first year seedling growth. Acta Hortic 570:331–334

Ishikawa T, Morimoto Y, Madhusudhan R, Sawa Y, Shibata H, Yabuta Y, Nishizawa A, Shigeoka S (2005) Acclimation to diverse environmental stresses caused by a suppression of cytosolic ascorbate peroxidase in tobacco BY-2 cells. Plant Cell Physiol 46:1264–1271

Jung HW, Kim W, Hwang BK (2003) Three pathogen-inducible genes encoding lipid transfer protein from pepper are differentially activated by pathogens, abiotic, and environmental stresses. Plant Cell Environ 26:915–928

Kader JC (1996) Lipid transfer proteins in plants. Annu Rev Physiol Plant Mol Biol 47:627–654

Liu QS, Pu L, Poo MM (2005) Repeated cocaine exposure in vivo facilitates LTP induction in midbrain dopamine neurons. Nature 437:1027–1031

Molina A, Diaz I, Vasil IK, Carbonero P, Garcia-Olmedo F (1996) Two cold-inducible genes encoding lipid transfer protein LTP4 from barley show differential responses to bacterial pathogens. Mol Gen Genet 27:162–168

Mouradov A, Cremer F, Coupland G (2002) Control of flowering time: interacting pathways as a basis for diversity. Plant Cell 14: S111–S130

Ng M, Yanofsky MF (2001) Function and evolution of the plant MADS-box gene family. Nat Rev Genet 2:186–195

Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1–6

Ouvrard O, Cellier F, Ferrare K, Tousch 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

Park CJ, Shin R, Park JM, Lee GJ, You JS, Paek KH (2003) Expression studies of SCA in lily and confirmation of its role in pollen tube adhesion. Plant Mol Biol 51:183–189

Robinson K, Firoozababy E (1993) Transformation of floriculture crops. Sci Hortic 55:83–99

Shii CT (1983) The distribution and variation of *Lilium formosanum* and *L. longiflorum* in Taiwan. Lily Yearbook of the North Amer. Lily Soc 36:48–51

Shigeoka S, Ishikawa T, Tamoi M, Miyagawa Y, Takeda T, Yabuta Y (2002) Regulation and function of ascorbate peroxidase isoenzymes. J Exp Bot 53:1305–1319

Simpson GG, Dean C (2002) *Arabidopsis*, the *Rosetta* stone of flowering time. Science 296:285–289

Springer NM, Kaeppler SM (2005) Evolutionary divergence of monocot and dicot methyl-CpG-binding domain proteins. Plant Physiol 138:92–104

Streb P, Feierabend J, Bligny R (1997) Resistance to photoinhibition of photosystem II and catalase and antioxidative protection in high mountain plants. Plant Cell Environ 20:1030–1040

Takishima K, Watanabe S, Yamada M, Suga T, Mamiya G (1988) Amino acid sequences of two nonspecific lipid-transfer proteins from germinated castor bean. Eur J Biochem 177:241–249

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Munster T, Winter KU, Saedler H (2000) A short history of MADS-box genes in plants. Plant Mol Biol 42:115–149

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Trevino MB, Oconnell MA (1998) Three drought-responsive members of the nonspecific lipid-transfer protein gene family in *Lycopersicon pennellii* show different developmental patterns of expression. Plant Physiol 116:1461–1468

Wu S, Watanabe N, Mita S, Ueda Y, Shibuya M, Ebizuka Y (2003) Two o-methyltransferases isolated from flower petals of *Rosa chinensis var. spontanea* involved in scent biosynthesis. J Biosci Bioeng 96:119–128

Yubero-Serrano EM, Moyano E, Medina-Escobar N, Munoz-Blanco J, Caballero JL (2003) Identification of a strawberry gene encoding a non-specific lipid transfer protein that responds to ABA, wounding and cold stress. J Exp Bot 54:1865–1877