Available online at www.sciencedirect.com
ScienceDirect
www.sciencedirect.com/science/journal/16720229
GENOMICS PROTEOMICS & BIOINFORMATICS

Article

Discovery, Identification and Comparative Analysis of Non-Specific Lipid Transfer Protein (nsLtp) Family in Solanaceae

Wanfei Liu[1,2], Dawei Huang[1,2], Kan Liu[1,2], Songnian Hu[1], Jun Yu[1], Gang Gao[3*], and Shuhui Song[1*]

[1]*Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029, China;*
[2]*Graduate University of Chinese Academy of Sciences, Beijing 100049, China;*
[3]*College of Life Sciences, Shanxi Normal University, Linfen 041004, China.*

Genomics Proteomics Bioinformatics 2010 Dec; 8(4): 229-237 DOI: 10.1016/S1672-0229(10)60024-1

Abstract

Plant non-specific lipid transfer proteins (nsLtps) have been reported to be involved in plant defense activity against bacterial and fungal pathogens. In this study, we identified 135 (122 putative and 13 previously identified) Solanaceae nsLtps, which are clustered into 8 different groups. By comparing with Boutrot's nsLtp classification, we classified these eight groups into five types (I, II, IV, IX and X). We compared Solanaceae nsLtps with *Arabidopsis* and Gramineae nsLtps and found that (1) Types I, II and IV are shared by Solanaceae, Gramineae and *Arabidopsis*; (2) Types III, V, VI and VIII are shared by Gramineae and *Arabidopsis* but not detected in Solanaceae so far; (3) Type VII is only found in Gramineae whereas type IX is present only in *Arabidopsis* and Solanaceae; (4) Type X is a new type that accounts for 52.59% Solanaceae nsLtps in our data, and has not been reported in any other plant so far. We further built and compared the three-dimensional structures of the eight groups, and found that the major functional diversification within the nsLtp family could be predated to the monocot/dicot divergence, and many gene duplications and sequence variations had happened in the nsLtp family after the monocot/dicot divergence, especially in Solanaceae.

Key words: nsLtp family, Solanaceae, phylogenetic analysis, three-dimensional structure

Introduction

Plant non-specific lipid transfer protein (nsLtp) was first purified and characterized from spinach leaves and named for its capability of transferring phospholipids from liposome to mitochondria or chloroplasts (*1, 2*). Two nsLtps have been discovered from germinated castor bean seeds (*3*), and Tsuboi *et al* demonstrated organ-specific occurrence and expression of the nsLtp isoforms in castor bean seedlings (*4*). NsLtps are encoded by a multigene family in many plant species and most nsLtps have been identified on the basis of sequence homology from cDNA clones or their sequences. For instance, 52, 49 and 122 putative nsLtps have been identified in rice (*Oryza sativa*), *Arabidopsis thaliana* and wheat (*Triticum aestivum*), respectively (*5*).

Plant nsLtps are involved in many useful physiological functions, such as mediating phospholipid transfer, involving in plant defense activity against

*Corresponding authors.
E-mail: ggniche@126.com; songshh@big.ac.cn

bacterial and fungal pathogens, participating in the assembly of hydrophobic protective layers of surface polymers like cutin, and associating with abiotic stresses (*1, 6-13*). In addition, several members of the nsLtp family have been identified as relevant allergens in plant foods and pollens (*11, 14-16*). NsLtps display a complex tissue- and development-specific expression pattern (*17, 18*). NsLtp expression is also mediated by Ca^{2+} and calmodulin, which interact specifically with nsLtp (*19, 20*).

All known plant nsLtp precursors are synthesized with an N-terminal signal peptide. Plant nsLTPs are small (usually 6.5 to 10.5 kDa except for type VII nsLTPs, which are about 15 kDa) and their isoelectric point (pI) is usually in the range of 8.5 to 12 (some are below 5). NsLtps have eight cysteine motifs (8CM) in a backbone sequence: C-Xn-C-Xn-CC-Xn-CXC-Xn-C-Xn-C (*21*). According to the sequence similarity, Boutrot *et al* divided nsLtps into nine types (I, II, III, IV, V, VI, VII, VIII and IX) (*5*).

nsLtps are predicted to be located in the extracellular matrix and possess a broad lipid-binding specificity, which is defined mostly based on their three-dimensional (3D) structures. The typical nsLtp fold, which is characterized by an α-helical compact domain composed of four α-helices (H1-H4), is connected by short loops (L1-L3) and a non-structured C-terminal tail. The compact domain is firmly held by a network of four conserved disulphide bridges [Pattern 1: C1-C6, C2-C3, C4-C7 and C5-C8 for type I nsLtps (*22*); Pattern 2: C1-C5, C2-C3, C4-C7 and C6-C8 for type II (*23*) and type IV (*24*)], such as *LTP1* gene, which belongs to pattern 1 and type X (**Figure 1**) (*2*). Additionally, a large number of intra-molecular H-bonds contribute to the stabilization of their 3D structures. This particular folding structure presents a large internal tunnel-like cavity that accommodates different types of lipids and also exhibits unusual stability against thermal processing and digestion (*25, 26*).

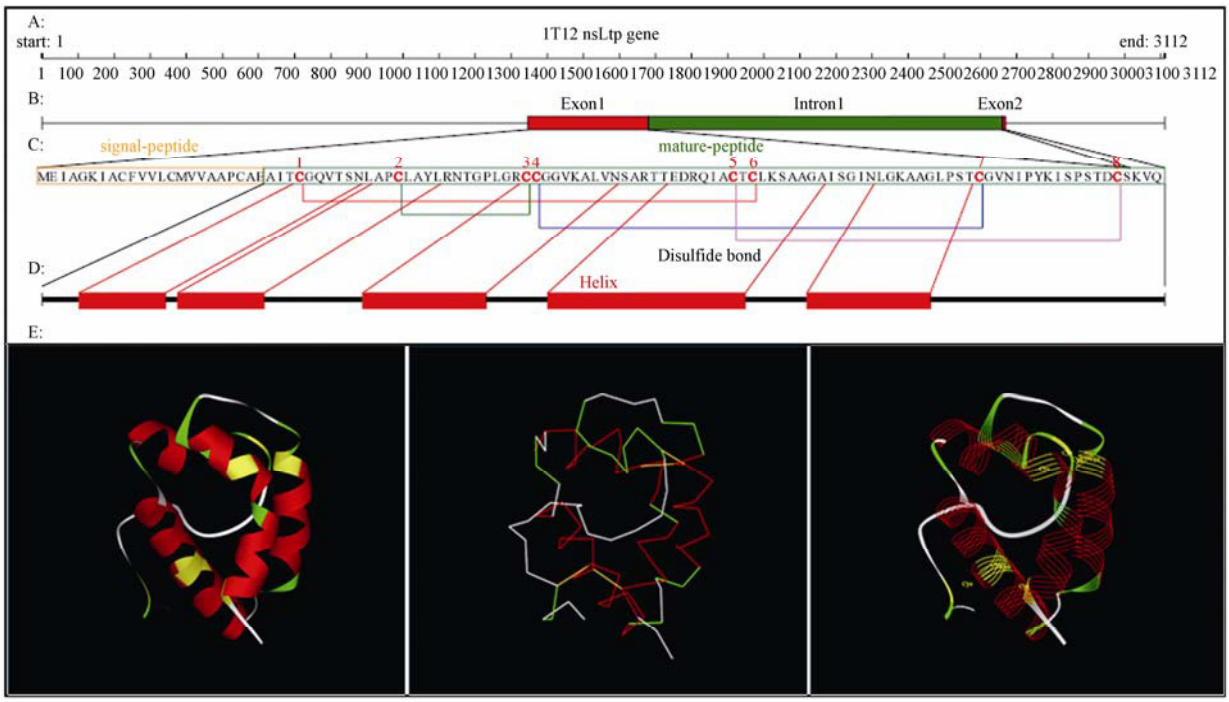

Figure 1 The 1T12 nsLtp structure of tobacco (*Nicotiana tabacum*). **A**. The sequence length, start and end position of 1T12 *nsLtp* gene in NCBI genomic DNA sequence. **B**. The exon and intron positions in 1T12 gene. The start of first exon and the end of second exon are not clear, so we use protein translation start and end site instead. **C**. The protein sequence of 1T12 and its position in the gene sequence. Signal peptide is yellow and mature peptide is green. **D**. The secondary structure of 1T12 and its position in protein sequence. **E**. The 3D structures of 1T12 in three styles shown from left to right. The first style shows the eight Cys residues in yellow color. The second style shows the four disulfide bonds in yellow color. The third style shows the eight Cys residues in yellow color and labels the Cys amino acids.

Solanaceae plants are major economic crops, such as potato (*Solanum tuberosum*), tomato (*Solanum lycopersicum*), tobacco (*Nicotiana tabacum*), and pepper (*Capsicum annuum*). Moreover, some Solanaceae plants have pharmacologic and horticultural value, such as Chinese wolfberry (*Lycium barbarum*) and Garden Petunia (*Petunia hybrida* Vilm). Therefore, it is not surprising to see 11 genome projects on Solanaceae plants at NCBI (http://www.ncbi.nlm.nih.gov/genomeprj). However, the Solanaceae nsLtp family has not been thoroughly annotated. In our study, we classified Solanaceae (the six Solanaceae species analyzed: *S. tuberosum*, *S. lycopersicum*, *N. tabacum*, *Nicotiana benthamiana*, *C. annuum* and *P. hybrida*) nsLtps by data mining, phylogenetic analysis, and 3D structure building. We also compared Solanaceae nsLtps with those of *Arabidopsis* and Gramineae (rice and wheat). Our analysis provides a detailed picture and novel insights on Solanaceae *nsLtp* gene family to support further functional and evolutionary research on *nsLtp* gene family among plants.

Results

Discovery and identification of *nsLtp* genes based on EST data mining in six Solanaceae species

In order to identity the non-redundant Solanaceae nsLtp sequences, we downloaded the identified sequences of nsLtp mRNAs, genes and proteins of Solanaceae*, O. sativa* and *A. thaliana* from NCBI. We found 22, 6 and 23 mRNAs; 5, 1 and 6 genes; and 13, 8 and 11 proteins in Solanaceae, *O. sativa* and *A. thaliana*, respectively (Table S1). Since none of the Solanaceae plant genome sequences have been finished until September, 2008, for our analysis we used 61,372, 46,849, 33,522, 16,127, 14,249 and 8,729 tentative consensus (TC) sequences from six Solanaceae species (including *S. tuberosum*, *S. lycopersicum*, *N. tabacum*, *N. benthamiana*, *C. annuum* and *P. hybrida*) from the TGI database of DFCI, respectively.

In order to identify all putative Solanaceae *nsLtp* genes, we first blasted the identified mRNA, gene and protein sequences of Solanaceae, *O. sativa* and *A. thaliana* nsLtp against the Solanaceae TC sequence databases. We obtained 246 blast hits at E-values less than 10^{-10} and performed alignment using the Est-wisedb program to validate the blast hits, yielding 235 putative *nsLtp* genes; we then used the CLC Sequence Viewer to translate all putative *nsLtp* genes into protein sequences and examined the 8CM motif manually and aligned all the nsLtp sequences using ClustalX program (*27, 28*). In order to reduce the sequence redundancy because of SNP and EST sequence error, we just considered the difference at protein level. After removing those redundant nsLtp sequences according to the alignment result, we identified 122 new Solanaceae putative nsLtp sequences (22 members have been identified in *S. tuberosum*, 24 in *S. lycopersicum*, 30 in *N. tabacum*, 17 in *N. benthamiana*, 19 in *C. annuum* and 10 in *P. hybrida*) based on EST data mining in the TGI database of DFCI (**Table 1**). Importantly, we first reported the putative nsLtp in *N. benthamiana*, *C. annuum* and *P. hybrid*. To distinguish those different species of Solanaceae, we named *S. tuberosum*, *S. lycopersicum*, *N. tabacum*, *N. benthamiana*, *C. annuum* and *P. hybrida* nsLtp as StLtp, SlLtp, NtLtp, NbLtp, CaLtp and PhLtp, respectively. These 122 newly identified and 13 previously identified protein sequences were employed for subsequent protein analysis and phylogenetic study.

Characteristics of Solanaceae nsLtps

We summarized the characteristics of 135 (122 new identified and 13 previously identified) Solanaceae nsLtps in Table S2. All protein sequences were predicted for the presence of the signal peptides using SignalP 3.0 program (*29, 30*). 97% of the Solanaceae nsLtp precursors are synthesized with a signal peptide of 15 to 33 amino acids, except for 4 proteins (CaLtpIb.1, CaLtpId.1, NbLtpX.7 and StLtpX.14), in which we did not find signal peptides. We believe that these sequences may not be complete or they may just simply not have signal peptides. NtLtpX.1 and NtLtpX.2 have different signal peptides, but their mature proteins are identical. This phenomenon was also reported in rice and wheat. We further predicted the subcellular targeting of 131 Solanaceae nsLtps using TargetP 1.1 Server program to find their secretory pathway (*31*). CaLtpIX.1, StLtpIc.1, SlLtpIX.1 and NbLtpIV.1 have been predicted to contain a chloro-

plast targeting peptide and others are secreted into the extracellular space. Solanaceae nsLtps are small and their molecular masses usually range from 7,253 Da to 10,835 Da. The average molecular mass and the theoretical pI are 9,049 Da and 10.06 pI, respectively. Some Solanaceae *nsLtp* gene sequences, which have genomic DNA sequence (such as AF525363, AF525362, U66465, U66466 and X62395), have one intron and two exons (Figure 1).

The main characteristic of plant nsLtps is the presence of eight cysteine (Cys) residues in a highly conserved position C-Xn-C-Xn-CC-Xn-CXC-Xn-CXn-C. The consensus sequence from mature protein sequence alignments of Solanaceae nsLtp showed the 8CM features except for type Ic, which has nine Cys residues (**Table 2**). So we predicted that most of Solanaceae nsLtps have the typical nsLtp structure. The mature protein sequence alignments of Solanaceae nsLtp and some sequences of *O. sativa* and *A. thaliana* are shown in Figure S1.

Table 1 Categorization of Solanaceae nsLtp

Species	Total	Previously identified	Newly identified	Type Ia	Type Ib	Type Ic	Type Id	Type II	Type IV	Type IX	Type X
ST	28	6	22	2	0	1	0	2	3	1	19
SL	28	4	24	6	2	0	0	1	2	1	16
NT	33	3	30	4	10	2	2	3	3	1	8
NB	17	**0**	**17**	4	0	1	0	1	1	0	10
CA	19	**0**	**19**	3	2	0	1	0	1	1	11
PH	10	**0**	**10**	3	1	0	0	0	0	0	6
Total	135	13	122	22	15	4	3	7	10	4	70

Note: ST, *Solanum tuberosum*; SL, *Solanum lycopersicum*; NT, *Nicotiana tabacum*; NB, *Nicotiana benthamiana*; CA, *Capsicum annuum*; and PH, *Petunia hybrida*. The bold numbers show the species in which nsLtps were never identified before.

Table 2 The consensus motif of each nsLtp type

NsLtp type	8CM and number of flanking amino acid residues												
Cysteine order		1		2		3 4		5 6		7		8	
Type Ia (Group 1)	X3,4,6,14,15	C	X9	C	X13-16	CC	X19	CXC	X22,23	C	X13,14,10	C	X4,7,1
Type Ib (Group 2)	X2,3,10,12	C	X9	C	X13	CC	X19	CXC	X22	C	X13	C	X4,5
Type Ic (Group 3)	X3	C	X9	C	X14	CC	**X2CX16**	CXC	X**21**	C	X13	C	X4
Type Id (Group 4)	X4	C	X9	C	X13,16,17	CC	X19	CXC	X22,23	C	X14	C	X6,15
Type II (Group 5)	X2	C	X**7**	C	X13	CC	X**8**	CXC	X23	C	X6	C	X0
Type IV (Group 6)	X3	C	X9	C	X15	CC	X9	CXC	X**24**	C	X**7**	C	X0
Type IX (Group 7)	X2	C	X**13**	C	X15	CC	X9	CXC	X22	C	X6	C	X4
Type X (Group 8)	X2,3,12	C	X9	C	X**12**,14	CC	X6,19,20	CXC	X22,25	C	X13,14	C	X4,6,7,8,14,17

Note: The specific elements of consensus motif that could be used to identify nsLtp types are shown in bold. The consensus motif of each nsLtp type is deduced from the analysis of the mature sequences of Solanaceae nsLtps or putative nsLtps presented in Table 1.

Phylogenetic analysis of the Solanaceae nsLtps

In order to analyze the phylogenetic relationship of the Solanaceae nsLtp families, we performed multiple sequence alignments with Solanaceae (135), *O. sativa* (49), *A. thaliana* (45) and *T. aestivum* (122) nsLtp mature sequences using the ClustalX program (Figure S2). We used AtI.1 and OsI.1 instead of AtLTPI.1 and OsLTPI.1. The rest had been done in the same manner. The unrooted phylogenetic trees were constructed using the neighbor-joining method in Phylip package and drawn by MEGA4 software (Figure S3) (*32*).

Based on the identity matrix of the alignments and the protein sequence 8CM property, the 135 Solanaceae nsLtps were clustered into 8 different groups (Figure S4). We compared the result with Freddy Boutrot's nsLtp classification and classified these eight groups into five types (group 1 in type Ia, group 2 in type Ib, group 3 in type Ic, group 4 in type Id, group 5 in type II, group 6 in type IV, group 7 in type

IX and group 8 in type X). The type X is the largest one, which has 70 Solanaceae nsLtps and is unique to Solanaceae. Types I, II, IV and IX are consisted of 44, 7, 10 and 4 members, respectively. Type I is further classified into four subtypes (type Ia, type Ib, type Ic and type Id) according to their sequence similarity and phylogeny, and we named them as LtpIa, LtpIb, LtpIc and LtpId, respectively (Table S2). The Solanaceae nsLtps are not homogeneously distributed in each type, and such a distribution may be related to their functional differentiation and complementarities.

Molecular modeling and comparison

We chose eight typical members (NtLtpIa.4, type Ia; NtLtpIb.4, type Ib; NtLtpIc.2, type Ic; NtLtpId.2, type Id; NtLtpII.3, type II; NtLtpIV.3, type IV; SlLtpIX.1, type IX and NtLtpX.1, type X) from each type or subtype, and performed molecular modeling using the Modeller program (*33*-*38*). The atomic coordinates of nsLtps (1T12, 1FK5, 1T12, 2ALG, 1L6H, 2RKN, 2RKN and 1T12) were used to build the 3D models according to sequence identity and resolution of 3D structure. The electric charge and water environment are considered by using Amber 9 program, and the final models were validated by PROCHECK (*39*). We viewed the primary and the energy minimized structure of these models (**Figure 2**) using VMD and Discovery Studio Visualizer software and the final structures are listed in Table S3 (*40*).

We compared the 8 nsLtp models with 1T12 structure using MultiProt program, and the alignment size is only 22 because of the non-conserved residues of every type (*41*). The alignment size is 81 for NtLtpIa.4 (type Ia), NtLtpIb.4 (type Ib), NtLtpIc.2 (type Ic), NtLtpId.2 (type Id) and NtLtpX.1 (type X), and 71 for NtLtpIV.3 (type IV) and SlLtpIX.1 (type IX). We did not show NtLtpII.3 (type II) alignment result since it is too short. We further compared the eight structures with that of 1T12 based on the combinatorial extension (CE) method (**Figure 3**) (*42*). The sequence identities according to the structures are shown in **Table 3**. The result is in agreement with the phylogenetic analysis.

Discussion

In this study, we carried out an analysis on *nsLtp* gene family in six Solanaceae species, which enabled us to identify 122 novel putative *nsLtp* genes. We have also clustered them into five types according to their phylogenies, where type X is a new type identified in this work. Most of the different types are monophyletic and supported by convincing bootstrap values except for type Ia (the bootstrap values are 53 and 43 for the two type Ia groups), which is largely due to their sequence diversity.

Comparing Solanaceae nsLtps in six Solanaceae species with those of *Arabidopsis* and Gramineae (rice and wheat), we found that: (1) Types I, II and IV are shared by Solanaceae, Gramineae and *A. thaliana*; (2) Types III, V, VI and VIII are not identified in six Solanaceae species so far, but exist in *Arabidopsis* and Gramineae; (3) Type VII was only found in Gramineae, whereas type IX was found in Solanaceae and *A. thaliana*, suggesting that type VII and IX may be unique to monocots or dicots but not both (*5*). The existence of linage-specific types suggests that they may evolve before the monocot-dicot split. Type X is a new group, as it has not been reported in any other plants so far. It accounts for 52.59% of total nsLtp among the six Solanaceae species. Based on the identity matrix of the multiple sequence alignments, type X may be a subtype of type I. However, according to their positions in the phylogenetic trees and the sequence 8CM property, these sequences probably do not share the same common ancestor with other type I members. The major conclusion we could draw from these observations is that the ancestral *nsLtp* gene family already evolved to three types before the separation of monocots and dicots, and subsequently, the family expanded specifically in each species by gene duplication.

Our phylogenetic result turned out to be more informative about the evolutionary relationships and amended the previous results (*5*). We classified types not only according to the percentage identity, but also the property of 8CM sequence and the comparison of 3D structures. By comparing the 3D structures of the eight groups in Solanaceae, we found

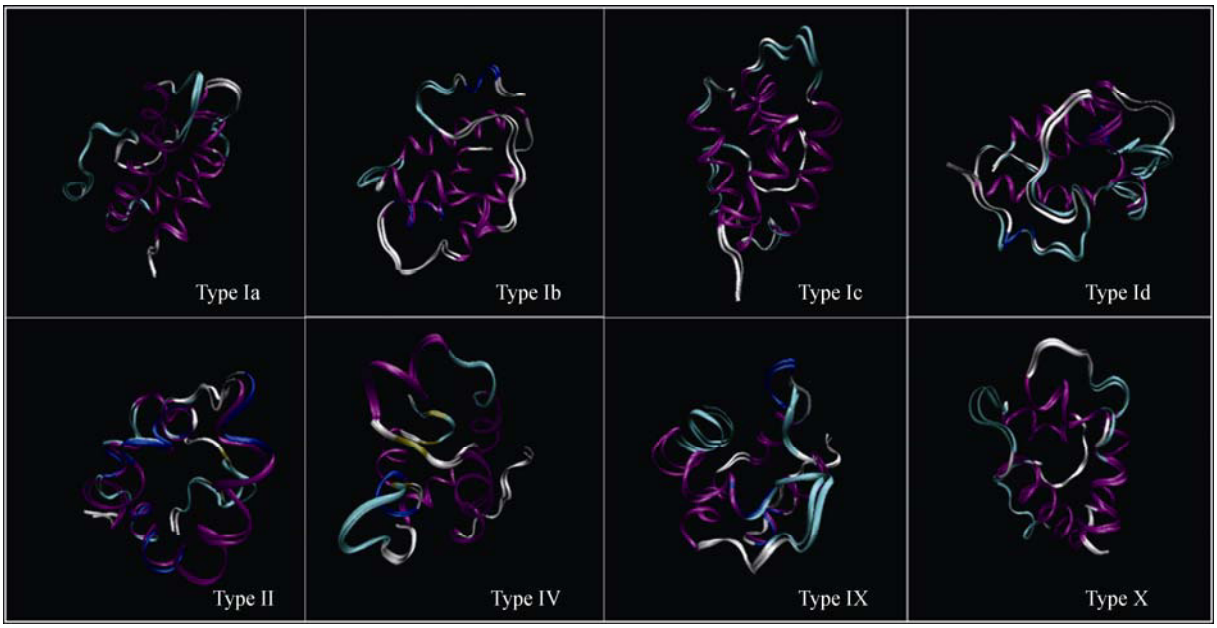

Figure 2 Comparison of the structural models before and after energy minimization.

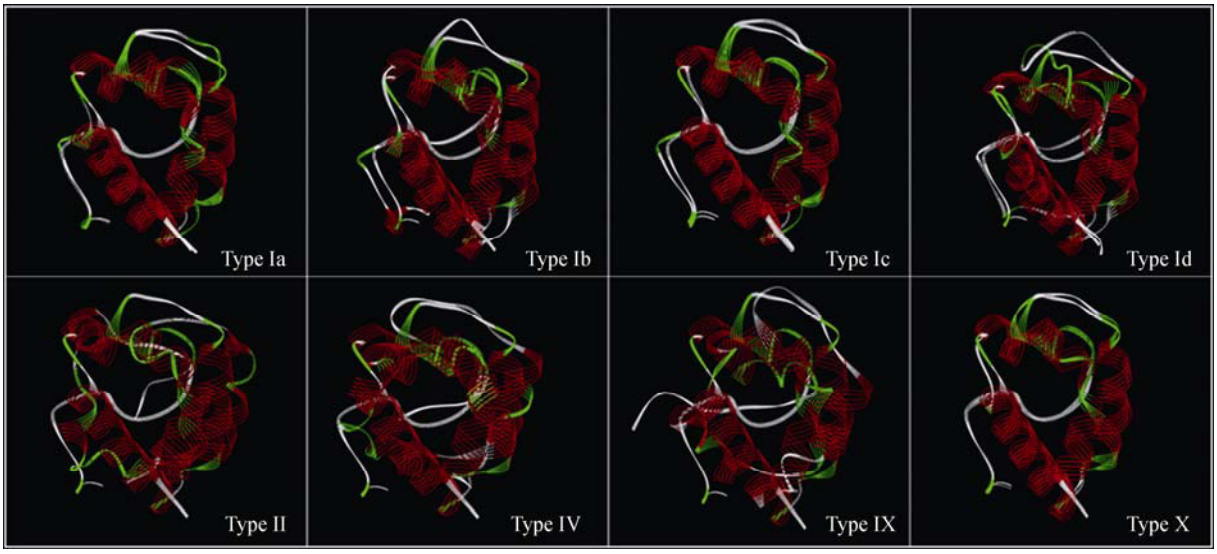

Figure 3 Comparison of eight protein models with 1T12 based on combinatorial extension calculation of the two chains.

Table 3 The RMSD values and the sequence identities of the eight typical nsLtps

	NtLtpIa.4 (type Ia)	NtLtpIb.4 (type Ib)	NtLtpIc.2 (type Ic)	NtLtpId.2 (type Id)	NtLtpII.3 (type II)	NtLtpIV.3 (type IV)	SlLtpIX.1 (type IX)	NtLtpX.1 (type X)
RMSD value	0.6856	0.6960	0.8163	0.9116	0.9403	0.5881	0.8373	0.6486
sequence identity	54.9%	31.1%	42.2%	26.7%	8.2%	14.3%	11.3%	64.4%

that type X is derived from type I based on protein structural features (for example, the alignment size between type X and type I). The root mean square deviation (RMSD) values between the eight structures and 1T12 structure are also accordant with their sequence homology (Table 3).

Conclusion

In this study we identified four known types and a new type (type X) for nsLtp, yielding a total of ten types in plants. Phylogenetic analysis and 3D structural modeling further confirmed that type X is a novel type. The diversity and complexity of type X *nsLtp* genes are the evolutionary consequence of adaptation to Solanaceae-specific functions or various environmental changes. Our work provides essential information on the classification of the Solanaceae *nsLtp* gene family and gives novel insights into their functional and evolutionary studies.

Materials and Methods

Sequence retrieval

Previously identified nsLtp sequences of Solanaceae, *O. sativa* and *A. thaliana* were retrieved from NCBI according to their annotations (**Table 4**). In order to identify all the other putative Solanaceae *nsLtp* genes, TC sequences of six Solanaceae plants (*S. tuberosum*, *S. lycopersicum*, *N. tabacum*, *N. benthamiana*, *C. annuum* and *P. hybrida*) were downloaded from the TGI databases in DFCI (http://compbio.dfci.harvard.edu/tgi/), which come from the Gene Index Project, and installed on our local server.

In silico discovery and identification of Solanaceae *nsLtp* genes by EST data mining

We performed BLAST (version 2.2.17) search using all known nsLtp molecular sequences of Solanaceae, *O. sativa* and *A. thaliana* against each Solanaceae TC sequences database. Sequences that satisfied E-value less than 10^{-10} and score-value more than 100 were considered as new candidate *nsLtp* genes. For validating the results of blast, we further compared the 32 known nsLtp sequences with all blast results using Estwisedb (version 2.2.0) (*27*). Sequences with more than 20 score-value were identified as putative *nsLtp* gene homologues. We translated all the TC sequences of putative *nsLtp* genes into proteins using CLC Sequence Viewer 5.0.1, and multiple alignment was then performed using ClustalX (version 2.0.6) (*28*). A protein was considered as a new putative nsLtp if at least one mutation was observed.

All identified nsLtps were analyzed for presence of potential signal peptide cleavage sites using SignalP 3.0 (*29, 30*). After removing the signal peptide of all nsLtps, we got putative mature nsLtp sequences. Each of the putative mature nsLtp sequences was manually validated through the analysis of the Cys residue pattern: the 8CM in a strongly conserved position (Cys1-Xn-Cys2-Xn-Cys3Cys4-Xn-Cys5XCys6-Xn-Cys 7-Xn-Cys8). After removing those sequences that did not have the 8CM (except for the sequences that lose Cys residue because of losing some sequences or having some sequence errors), we identified 122 new Solanaceae putative nsLtp sequences in total. These protein sequences were employed for subsequent protein analysis and evolutionary study.

Sequence alignment and phylogenetic tree reconstruction

Solanaceae, *O. sativa* and *A. thaliana* nsLtp sequences were aligned with ClustalX (version 2.0.6). After refining the alignment results manually, we built the phylogenetic tree using neighbor-joining method in Phylip package and drew it by MEGA4 software package (*32*). We used AtI.1 and OsI.1 instead of AtLTPI.1 and OsLTPI.1. The rest had been done in the same manner. The confidence level of each node was estimated by bootstrap procedure using 1,000 re-sampling repetitions of the data.

Table 4 The numbers of nsLtp mRNA, gene and protein sequence

	S. tuberosum	*S. lycopersicum*	*N. tabacum*	*C. annuum*	*O. sativa*	*A. thaliana*	Total
No. of *nsLtp* mRNA sequences	6	4	5	7	6	23	51
No. of *nsLtp* gene sequences	2	2	1	0	1	6	12
No. of nsLtp protein sequences	6	4	3	0	8	11	32

Molecular modeling and comparison

We constructed the models of the eight typical Solanaceae nsLtps with Modeller program using known nsLtp 3D structures [PDB ID: 1T12 (*2*), 1FK5 (*43*), 1T12, 2ALG (*44*), 1L6H (*45*), 2RKN (*24*), 2RKN and 1T12, respectively] according to sequence identity and resolution of 3D structure. The lowest sequence identity between the target protein and the template protein is 33%. A total of 50 models were built for each nsLtp molecule, and the model quality was evaluated by DOPE assessment score, GA341 assessment score and PROCHECK program (*39*). The model selection was based on the model quality result. At last, we performed the model optimization using the Amber 9 program by adding Na^+ or Cl^- and water molecular to the atomic coordinate of models. The structure alignment between eight final Solanaceae nsLtp models and 1T12 structure was performed by MultiProt program, and CE program calculates the two chains (*41, 42*).

Acknowledgements

We are grateful to our bioinformatics team members for providing critical reading of the manuscript and many constructive discussions. We also acknowledge Dr. Zhang Zhang's critical reading and editing and the help of Tongwu Zhang and Jingfa Xiao in protein structure field. This work was supported by the National Natural Science Foundation of China (Grant No. 30900831).

Authors' contributions

WL and SS conceived and carried out the study design, performed the bioinformatics analysis, interpreted the analysis results and drafted the manuscript. DH, KL, SH, JY and GG participated in the study design and discussion, and helped in revising and editing the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors have declared that no competing interests exist.

References

1 Kader, J.C., *et al.* 1984. Purification and characterization of a spinach-leaf protein capable of transferring phospholipids from liposomes to mitochondria or chloroplasts. *Eur. J. Biochem.* 139: 411-416.
2 Da Silva, P., *et al.* 2005. Solution structure of a tobacco lipid transfer protein exhibiting new biophysical and biological features. *Proteins* 59: 356-367.
3 Takishima, K., *et al.* 1988. Amino acid sequences of two nonspecific lipid-transfer proteins from germinated castor bean. *Eur. J. Biochem.* 177: 241-249.
4 Tsuboi, S., *et al.* 1991. Organ-specific occurrence and expression of the isoforms of nonspecific lipid transfer protein in castor bean seedlings, and molecular cloning of a full-length cDNA for a cotyledon-specific isoform. *J. Biochem.* 110: 823-831.
5 Boutrot, F., *et al.* 2008. Genome-wide analysis of the rice and *Arabidopsis* non-specific lipid transfer protein (nsLtp) gene families and identification of wheat nsLtp genes by EST data mining. *BMC Genomics* 9: 86.
6 Kristensen, A.K., *et al.* 2000. Characterization of a new antifungal non-specific lipid transfer protein (nsLTP) from sugar beet leaves. *Plant Sci.* 155: 31-40.
7 Ge, X.C., *et al.* 2002. Expression, purification and function of rice nonspecific lipid transfer protein. *Acta Biochimica et Biophysica Sinica* 34: 83-87.
8 Gomes, E., *et al.* 2003. Nonspecific lipid-transfer protein genes expression in grape (Vitis sp.) cells in response to fungal elicitor treatments. *Mol. Plant Microbe. Interact.* 16: 456-464.
9 Yang, X., *et al.* 2008. Characterization and expression of an nsLTPs-like antimicrobial protein gene from motherwort (*Leonurus japonicus*). *Plant Cell Rep.* 27: 759-766.
10 Han, G.W., *et al.* 2001. Structural basis of non-specific lipid binding in maize lipid-transfer protein complexes revealed by high-resolution X-ray crystallography. *J. Mol. Biol.* 308: 263-278.
11 Salcedo, G., *et al.* 2007. Plant non-specific lipid transfer proteins: an interface between plant defence and human allergy. *Biochim. Biophys. Acta* 1771: 781-791.
12 Ouvrard, O., *et al.* 1996. Identification and expression of water stress- and abscisic acid-regulated genes in a drought-tolerant sunflower genotype. *Plant Mol. Biol.* 31: 819-829.
13 Hong, S.T., *et al.* 2001. Molecular cloning of the anther-preferential nonspecific lipid transfer protein cDNA in hot pepper by mRNA differential display. *Mol. Cells* 11: 241-249.
14 Borges, J.P., *et al.* 2006. The lipid transfer proteins (LTP) essentially concentrate in the skin of Rosaceae fruits as cell surface exposed allergens. *Plant Physiol. Biochem.* 44: 535-542.

15 Breiteneder, H. and Mills, C. 2005. Nonspecific lipid-transfer proteins in plant foods and pollens: an important allergen class. *Curr. Opin. Allergy Clin. Immunol.* 5: 275-279.
16 Hartz, C., *et al.* 2007. Molecular characterisation of Lac s 1, the major allergen from lettuce (*Lactuca sativa*). *Mol. Immunol.* 44: 2820-2830.
17 Boutrot, F., *et al.* 2005. Wheat non-specific lipid transfer protein genes display a complex pattern of expression in developing seeds. *Biochim. Biophys. Acta* 1730: 114-125.
18 Horvath, B.M., *et al.* 2002. Expression analysis of a family of nsLTP genes tissue specifically expressed throughout the plant and during potato tuber life cycle. *Plant Physiol.* 129: 1494-1506.
19 Gao, G., *et al.* 2009. The potato StLTPa7 gene displays a complex Ca-associated pattern of expression during the early stage of potato-Ralstonia solanacearum interaction. *Mol. Plant Pathol.* 10: 15-27.
20 Xie, W.Q., *et al.* 2006. The effects of calmodulin on the lipid-binding activity of CaM-binding protein-10 and maize non-specific lipid transfer protein. *J. Plant Physiol. Mol. Biol*. 32: 679-684.
21 Jose-Estanyol, M., *et al.* 2004. The eight-cysteine motif, a versatile structure in plant proteins. *Plant Physiol. Biochem.* 42: 355-365.
22 Gincel, E., *et al.* 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.
23 Pons, J.L., *et al.* 2003. Refined solution structure of a liganded type 2 wheat nonspecific lipid transfer protein. *J. Biol. Chem.* 278: 14249-14256.
24 Lascombe, M.B., *et al.* 2008. The structure of "defective in induced resistance" protein of *Arabidopsis thaliana*, DIR1, reveals a new type of lipid transfer protein. *Protein Sci.* 17: 1522-1530.
25 Gaier, S., *et al.* 2008. Purification and structural stability of the peach allergens Pru p 1 and Pru p 3. *Mol. Nutr. Food Res.* 52: S220-229.
26 Scheurer, S., *et al.* 2004. Strong allergenicity of Pru av 3, the lipid transfer protein from cherry, is related to high stability against thermal processing and digestion. *J. Allergy Clin. Immunol.* 114: 900-907.
27 Birney, E., *et al.* 2004. GeneWise and genomewise. *Genome Res.* 14: 988-995.
28 Larkin, M.A., *et al.* 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.
29 Bendtsen, J.D., *et al.* 2004. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 340: 783-795.
30 Nielsen, H., *et al.* 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 10: 1-6.
31 Emanuelsson, O., *et al.* 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* 300: 1005-1016.
32 Tamura, K., *et al.* 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596-1599.
33 Eswar, N., *et al.* 2007. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* Chapter 2, Unit 2.9.
34 Fiser, A., *et al.* 2000. Modeling of loops in protein structures. *Protein Sci.* 9: 1753-1773.
35 Marti-Renom, M.A., *et al.* 2000. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29: 291-325.
36 Sali, A. and Blundell, T.L. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234: 779-815.
37 Sanchez, R. and Sali, A. 1997. Evaluation of comparative protein structure modeling by MODELLER-3. *Proteins* Suppl 1: 50-58.
38 Sanchez, R. and Sali, A. 2000. Comparative protein structure modeling. Introduction and practical examples with modeller. *Methods Mol. Biol.* 143: 97-129.
39 Laskowski, R.A., *et al.* 1993. Procheck: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26: 283-291.
40 Humphrey, W., *et al.* 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14: 27-28, 33-38.
41 Shatsky, M., *et al.* 2004. A method for simultaneous alignment of multiple protein structures. *Proteins* 56: 143-156.
42 Shindyalov, I.N. and Bourne, P.E. 1998. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* 11: 739-747.
43 Han, G.W., *et al.* 2001. Structural basis of non-specific lipid binding in maize lipid-transfer protein complexes revealed by high-resolution X-ray crystallography. *J. Mol. Biol.* 308: 263-278.
44 Pasquato, N., *et al.* 2006. Crystal structure of peach Pru p 3, the prototypic member of the family of plant non-specific lipid transfer protein pan-allergens. *J. Mol. Biol.* 356: 684-694.
45 Samuel, D., *et al.* 2002. Solution structure of plant nonspecific lipid transfer protein-2 from rice (*Oryza sativa*). *J. Biol. Chem.* 277: 35267-35273.

Supplementary Material

Figures S1-S4; Tables S1-S3

DOI: 10.1016/S1672-0229(10)60024-1