

PROCEEDINGS

Open Access

# Construction and analysis of a plant non-specific lipid transfer protein database (nsLTPDB)

Nai-Jyuan Wang<sup>1†</sup>, Chi-Ching Lee<sup>1,2†</sup>, Chao-Sheng Cheng<sup>1</sup>, Wei-Cheng Lo<sup>3</sup>, Ya-Fen Yang<sup>1</sup>, Ming-Nan Chen<sup>1</sup>, Ping-Chiang Lyu<sup>1,4,5\*</sup>

From The Tenth Asia Pacific Bioinformatics Conference (APBC 2012)  
Melbourne, Australia. 17-19 January 2012

## Abstract

**Background:** Plant non-specific lipid transfer proteins (nsLTPs) are small and basic proteins. Recently, nsLTPs have been reported involved in many physiological functions such as mediating phospholipid transfer, participating in plant defence activity against bacterial and fungal pathogens, and enhancing cell wall extension in tobacco. However, the lipid transfer mechanism of nsLTPs is still unclear, and comprehensive information of nsLTPs is difficult to obtain.

**Methods:** In this study, we identified 595 nsLTPs from 121 different species and constructed an nsLTPs database – nsLTPDB – which comprises the sequence information, structures, relevant literatures, and biological data of all plant nsLTPs <http://nsltpdb.life.nthu.edu.tw/>.

**Results:** Meanwhile, bioinformatics and statistics methods were implemented to develop a classification method for nsLTPs based on the patterns of the eight highly-conserved cysteine residues, and to suggest strict Prosite-styled patterns for Type I and Type II nsLTPs. The pattern of Type I is C X<sub>2</sub> V X<sub>5-7</sub> C [V, L, I] × Y [L, A, V] X<sub>8-13</sub> CC × G X<sub>12</sub> D × [Q, K, R] X<sub>2</sub> CXC X<sub>16-21</sub> P X<sub>2</sub> C X<sub>13-15</sub>C, and that of Type II is C X<sub>4</sub> L X<sub>2</sub> C X<sub>9-11</sub> P [S, T] X<sub>2</sub> CC X<sub>5</sub> Q X<sub>2-4</sub> C[L, F]C X<sub>2</sub> [A, L, I] × [D, N] P X<sub>10-12</sub> [K, R] X<sub>4-5</sub> C X<sub>3-4</sub> P X<sub>0-2</sub> C. Moreover, we referred the Prosite-styled patterns to the experimental mutagenesis data that previously established by our group, and found that the residues with higher conservation played an important role in the structural stability or lipid binding ability of nsLTPs.

**Conclusions:** Taken together, this research has suggested potential residues that might be essential to modulate the structural and functional properties of plant nsLTPs. Finally, we proposed some biologically important sites of the nsLTPs, which are described by using a new Prosite-styled pattern that we defined.

## Background

Lipids are hydrocarbons insoluble in water but soluble in organic solvents. They are commonly translocated among subcellular membranes to enable various metabolic activities [1]. Lipid transfer proteins (LTPs) have been found in animals, plants and some fungi, and they exist in many tissues with various sizes and functions [1-3]. LTPs play an important role not only in plant but also in human. It mediates *in vitro* the transfer of all

common phospholipids, cholesterol and gangliosides between membranes [4,5]. The term plant “nonspecific lipid transfer proteins” indicates that LTPs can associate with various phospholipids with broad specificity [6]. The first known plant lipid transfer protein was isolated from potato tuber in 1975 by Kadar [7]. At present, much more nsLTPs have been found in monocots, dicots and gymnosperms, etc [3]. Plant nsLTPs are a kind of small (usually 6.5 to 10.5 kDa), basic (isoelectric point, or pI, usually falls between 8.5 and 12) and stable (with four conserved disulfide bonds) proteins. They can be isolated from various plants, e.g., Arabidopsis, rice, barely, wheat, maize, castor bean, and spinach leaf [8-15]. All the identified nsLTPs show high homology in

\* Correspondence: [pclyu@mx.nthu.edu.tw](mailto:pclyu@mx.nthu.edu.tw)

† Contributed equally

<sup>1</sup>Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu, Taiwan

Full list of author information is available at the end of the article

protein sequence and share similar characteristics. NsLTPs are stabilized by eight conserved cysteine residues forming four disulfide bonds and they usually contain signal peptides in the N-terminus [1]. Previous studies showed that nsLTPs can be divided into two main groups according to their molecular weight: nsLTP1 (9 kDa) and nsLTP2 (7 kDa) [16]. These two groups exhibit different disulfide bond patterns. The disulfide bond linkages of nsLTP1 at Cys<sub>1</sub>-Cys<sub>6</sub> and Cys<sub>5</sub>-Cys<sub>8</sub> differ from those of nsLTP2 at Cys<sub>1</sub>-Cys<sub>5</sub> and Cys<sub>6</sub>-Cys<sub>8</sub>. The major difference is observed at the C<sub>6</sub>-X-C<sub>8</sub> motif. For the CXC motif in nsLTP1, X is a hydrophilic residue, for example asparagine; however, in nsLTP2, a hydrophobic residue, such as leucine or phenylalanine, was found at the X position. These conserved hydrophilic or hydrophobic residues may play important roles in the biological functions of nsLTPs [17].

Several plant nsLTP structures have been determined. Three dimensional structures of either ligand-free and ligand-bound forms of nsLTPs are available [9,10,12,18]. The structure of nsLTP1 is composed of four alpha helices and a flexible stretching C-terminus [19,20]. The four alpha helices are connected by flexible loops and stabilized by the four disulfide bonds [10,12,18]. A typical characteristic of nsLTPs is the existence of an internal hydrophobic cavity running through the molecule. The cavity allows the binding of one or two monoacyl lipids, diacylated lipids, or some hydrophobic molecules [18,21]. The hydrophobic cavity in nsLTP1 shows a tunnel-like conformation, and nsLTP2 exhibits a triangular conformation [19]. The major structural difference between nsLTP1 and nsLTP2 is the size of the hydrophobic cavity; the cavity of an nsLTP1 is usually larger than that of an nsLTP2 protein [10,18,22]. In recent years, increasing studies have reported that plant nsLTPs are involved in many crucial biological functions but the mechanisms responsible for these functions are unclear yet.

Several biological functions of plant nsLTPs have been identified, inclusive of mediating phospholipid transfer, involving in plant defence activity against bacterial and fungal pathogens, and participating in the assembly of hydrophobic protective layers of surface polymers such as the formation of cutin [23,24]. NsLTPs were found accumulated at the surface of certain tissues at a high concentration [25], which may be correlated with the adaptation to different environmental stresses [26]. Several studies pointed out that the expression of nsLTPs can be induced by environmental stresses like extreme temperatures, osmosis pressures and drought [27]. Furthermore, nsLTPs exhibit defence activities toward blight or pathogens because of their high thermal stability and resistance to proteases [23,28]. In addition, nsLTPs are involved in the formation of beer foam [29]

and in food allergy to processed fruits [30]. Jeroen Nieuwland et al. postulated that nsLTPs can associate with hydrophobic cell wall compounds and disrupt the cell wall or facilitate the extension of cell wall [31]. These features of nsLTPs suggest that their functions are very diverse, and these features may exist because of their ability to bind and/or carry hydrophobic molecules such as fatty acid or fatty acid derivatives [25].

There is no golden standard for the identification and classification of nsLTPs because of their unclear lipid transfer mechanisms and the insufficiency of publicly available data. In the last twenty years, nsLTPs were mainly categorized into two subfamilies based on their molecular weights, nsLTP1 (~9 kDa) and nsLTP2 (~7 kDa) [3]. Nevertheless, this method is inadequate for categorizing many newly identified nsLTPs [7]. In 2008, Boutrot et al. proposed a new classification for nsLTPs using the putative mature form of rice, wheat, and *Arabidopsis thaliana*. The authors divided nsLTPs into nine types (from I to IX) according to their sequence similarities (see Additional File 1) [32]. Some recent papers minorly modified this classification system using a very limited number of sequences (see Additional File 2) [33,34].

Plant non-specific lipid transfer proteins are one of the most well-known proteins that are widely distributed in the plant kingdom. Our wet-lab laboratory has been studying nsLTPs for years, but there is still much unknown space left about these sequence highly-diverse proteins. Importantly, there is no nsLTPs database systematically collecting and organizing relevant data about nsLTPs. Boutrot et al. had identified and classified 267 nsLTPs sequences in 2008 [32], but their method still failed to classify many nsLTPs (see Table 1) [32]. This work aimed to establish an nsLTPs database, develop a robust classification method for nsLTPs and formulate Prosite signature patterns for the identification of nsLTPs as well as the key residues for the structural stability or the lipid binding ability of nsLTPs.

## Methods

### Databases and web-based tools utilized

**NCBI** <http://www.ncbi.nlm.nih.gov/>

The National Center for Biotechnology Information provides many public databases and tools relating to biotechnology. First, we established a non-redundant protein sequence dataset by retrieving data from NCBI RefSeq and Genbank; then, we used BLAST [35] to search for homologous sequences for each sequence in our dataset.

**ExpASY** <http://expasy.org/>

The ExpASY (Expert Protein Analysis System) database is established by Swiss Institute of Bioinformatics (SIB) and European Bioinformatics Institute (EBI), such as

**Table 1 Comparison of several published nsLTP datasets**

Authors [reference number]	Year of publication	Species	Total number of nsLTP sequences	Number of non-redundant nsLTP sequences
Gautier, et al. [32]	2008	3	251	131
Liu, et al. [33]	2010	1	135	75
Edstam, et al. [52]	2011	5	149	81
This study	2011	121	1,395	595

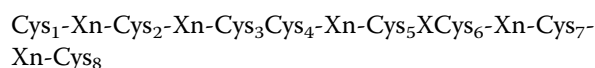
Swiss-Prot, UniProtKB and TrEMBL. Swiss-Prot and TrEMBL provide many information relating to protein sequence, structure and function (e.g., domains structure, post-translational modifications, variants) [36].

**TARGETP** <http://www.cbs.dtu.dk/services/TargetP>

The TargetP 1.1 Server is a web based tool to predict the subcellular location of eukaryotic proteins [37].

**SignalP 3.0 Server** <http://www.cbs.dtu.dk/services/SignalP/>

The SignalP 3.0 Server is a sequence prediction server that allows user to submit the sequence query and receive the result about presence and location of signal peptide cleavage sites in amino acid sequences from different organisms. In this study, all identified nsLTPs were analyzed for presence of potential signal peptide cleavage sites by using this tool. After removing the signal peptide of all nsLTPs in our database, we got putative mature-form nsLTPs sequences. Each putative mature nsLTP sequence was validated through the analysis of the 8-cysteine residue motif (8-Cys motif):



After removing proteins improbable to be nsLTPs, we identified 1,395 putative nsLTP sequences. Then we constructed a database and a web-based user interface collecting all these putative nsLTPs and relevant information. Additionally, in order to make our results more reliable, we deleted any redundant sequences with 100% sequence identities and finally got 595 putative nsLTP sequences; these sequences were employed for subsequent protein analysis and evolutionary study.

**Protein Data Bank** <http://www.pdb.org/>

The protein structure files utilized in this work were obtained from the Protein Data Bank (PDB) [38].

**Prosite database** <ftp://ftp.expasy.org/databases/prosite/>

The Prosite database is a collection of annotated motif descriptors from protein families and domains [39-42]. These descriptors, or patterns, are extracted from SWISS-PROT protein databases. Each pattern is recorded with two files: PROSITE.DAT is a computer readable text file providing all information necessary to programs that will scan sequences with patterns and/or matrices, and PROSITE.doc contains textual information and the documentation of patterns listed in PROSITE.DAT.

The version of Prosite we used was 20.8. After careful tests, we found that, although this version of Prosite possess 1,331 patterns, few of them are related with nsLTPs and most mature nsLTP sequences could not be recognized by those patterns. In this study, we examined the eight well-conserved cysteine region of collected nsLTPs and finally proposed new Prosite-styled patterns for Type I and Type II nsLTPs.

#### Data mining and Hidden Markov model

The standalone BLAST (version 2.2.17) [33] was utilized as the search engine, by using which we searched all known plant nsLTP sequences against the SwissProt protein sequence database. For every known nsLTP, homologous sequences from plant organisms with sequence identities >15% were considered as candidate nsLTP sequences. Then, candidates without 8-Cys motif were filtered out. After further removing redundant homologous sequences with 100% sequence identities, we manually examined every remaining candidate nsLTP sequence and thus identified 595 nsLTPs.

#### Sequence alignment and phylogenetic tree reconstruction

To examine the phylogenetic relationships of the nsLTPs identified in this study, we used ClustalW (version 2.0.12), a well-known multiple sequence alignment method, to obtain all the pairwise sequence similarities between nsLTP sequences. After refining the alignment results manually, we utilized the PHYLIP package v3.67 [43] to construct the phylogenetic tree of nsLTPs by using the UPGMA (Unweighted Pair Group Method with Arithmetic Mean) [44] and the neighbor-joining [45] clustering methods. Finally, MEGA4 [46] and Dendroscope [47] software packages were recruited to draw the tree graphs. In the web interface of our database, BioEdit (v7.0.9.0) was also utilized to compute amino acid identities and visualize sequence alignments.

## Results and discussion

### Classification based on sequence similarities

In this study, we characterized 595 nsLTPs. The presence of signal peptide for each protein was predicted by using the SignalP 3.0 program, and we found that 98% of the nsLTP precursors were initially synthesized with a signal peptide of 7-49 amino acids. The main

characteristic of plant nsLTPs was the presence of eight cysteine residues at highly conserved positions, the spanning of which forms a common sequence pattern:

C-Xn-C-Xn-CC-Xn-CXC-Xn-CXn-C (8-Cys motif).

This 8-Cys motif is consensus in nsLTPs, but it could not be used to classify nsLTPs. To classify nsLTPs, we modified Boutrot's nine-type classification into a five-type system (see Table 2 and Table 3). After analyzing the classified nsLTPs, we found that (1) Types I and II are shared by all the species that we identified to possess nsLTPs; (2) Type III is only found in *Oryza sativa* and *Arabidopsis*; (3) Types IV and V are shared by *Triticum aestivum*, *Oryza sativa*, *Sorghum bicolor* and *Arabidopsis*.

After making the above classification, we further analyzed the pI (isoelectric point) values, Mw (molecular weight) values, charges and the CXC motifs of all available nsLTPs. As shown in Additional File 3A, Type I and Type III were mostly 9 kDa proteins and Type II nsLTPs were 7 kDa proteins; the Mw of Type IV and Type V was much higher than that of Types I-III. Judging from the pI values, Types I, II and III are mostly alkaline proteins. Type IV nsLTPs are weakly alkaline and most Type V nsLTPs are acidic (see Additional File 3B). As for the CXC motif, most residues at the × position in Type I nsLTPs were hydrophilic, while in Type II, III, IV, and V nsLTPs, the × position is usually occupied by a hydrophobic residue (see Additional File 3C). There is no obvious difference in the distribution of net charge among all types of nsLTP (see Additional File 3D).

#### Phylogenetic analysis of nsLTPs

In order to analyze the phylogenetic relationships of the nsLTP families, we performed multiple sequence

alignments for mature-form nsLTP sequences by using the ClustalW program. Unrooted phylogenetic trees were generated with the UPGMA clustering method implemented in the Phylip package. Based on the number of residues that intervene the eight conserved cysteine residues, the 595 nsLTPs were clustered into 5 different groups. The results of our phylogenetic analyses supported our classification results. As shown in Additional File 4 and 5, the 5 types of nsLTPs could be fully separated in the phylogenetic trees.

#### Strategies for defining new Prosite-styled patterns for nsLTPs

Functional sites of proteins collected in the Prosite database are expressed as regular expressions. By querying Prosite with the nsLTP sequences that have assigned Uniprot IDs, we noticed that many (*i.e.* 86 cases) of them shared in common the pattern PS00597, that is,

Prosite entry PLANT\_LTP: [LIVM]-[PA]-x(2)-C-x-[LIVM]-x-[LIVM]-x-[LIVMFY]-x-[LIVM]-[ST]-x(3)-[DN]-C-x(2)-[LIVM].

However, this regular expression pattern failed to recognize most of the other nsLTP sequences we collected. Therefore, we would like to define new Prosite-styled patterns that are feasible to identify a broad scope of nsLTPs. We have previously observed, in our multiple sequence alignment results, that several positions in the nsLTP sequence are moderately or even highly conserved. Here we computed the occurrence of 20 amino acids at every position. Notably, in both nsLTP1 and nsLTP2, we found that some positions are always occupied by the same amino acids or amino acids with the same physiochemical properties. Two new patterns for nsLTP Type I and Type II were then defined according to the amino acid occurrence at various positions; they

**Table 2 Diversity of the eight cysteine motifs of the nsLTPs recorded in the nsLTPDB**

8 CM and number of flanking amino acid residues													
Type		1		2		3,4		5 6		7		8	
I	X <sub>0-12,16</sub>	C	X <sub>8-10</sub>	C	X <sub>12-17</sub>	CC	X <sub>18-20,29</sub>	CXC	X <sub>19-24</sub>	C	X <sub>7,13-15</sub>	C	X <sub>26,37,48</sub>
II	X <sub>0-20</sub>	C	X <sub>7</sub>	C	X <sub>13,15</sub>	CC	X <sub>8-10</sub>	CXC	X <sub>16,21,24</sub>	C	X <sub>4-7</sub>	C	X <sub>0-2</sub>
III	X <sub>0-7</sub>	C	X <sub>9,10</sub>	C	X <sub>12,15,17</sub>	CC	X <sub>9</sub>	CXC	X <sub>21-24,2</sub>	C	X <sub>6-10,13</sub>	C	X <sub>0-5,10</sub>
IV	X <sub>2-5,10</sub>	C	X <sub>14</sub>	C	X <sub>14</sub>	CC	X <sub>11-13</sub>	CXC	X <sub>24</sub>	C	X <sub>10</sub>	C	X <sub>6,10,12</sub>
V	X <sub>2-17</sub>	C	X <sub>10</sub>	C	X <sub>16,17</sub>	CC	X <sub>9</sub>	CXC	X <sub>22,23</sub>	C	X <sub>7,9</sub>	C	X <sub>5-12</sub>
Cysteine residue numbers are missing													
Type		1		2		3,4		5 6		7		8	
A			X <sub>1-13,18,29</sub>	C	X <sub>8,13-16</sub>	CC	X <sub>8,9,16-19</sub>	CXC	X <sub>22,23</sub>	C	X <sub>6-14</sub>	C	X <sub>1-6</sub>
B	X <sub>1-3,11</sub>	C	X <sub>7,9</sub>	C	X <sub>13-16</sub>	CC	X <sub>8,12,19</sub>	CXC	X <sub>21-24,36</sub>	C	X <sub>4,10,13,33,64</sub>		
C			X <sub>4,9</sub>	C	X <sub>13</sub>	CC	X <sub>15,19</sub>	CXC	X <sub>22,45</sub>	C	X <sub>1,11,12</sub>		
D				C	X <sub>1-11</sub>	CC	X <sub>19</sub>	CXC	X <sub>21-23</sub>	C	X <sub>4,5,10,12</sub>		

**Table 3 Molecular weights and pI values of the nsLTPs recorded in the nsLTPDB**

Type	Number of species	Number of members	Molecular weight (Mw)	Isoelectric point (pI)
I	88	391	7,995-15,000	3.67-12.30
II	23	102	6,130-9,270	4.50-12.02
III	2	9	7,220-8,918	4.50-6.73, 9.52-10.17
IV	4	11	9,348-10,506	8.48-12.31
V	3	4	9,282-10,816	4.75-5.27, 9.82

are available in Figure 1. Type III, IV and V were omitted in this experiment because of the small number of cases.

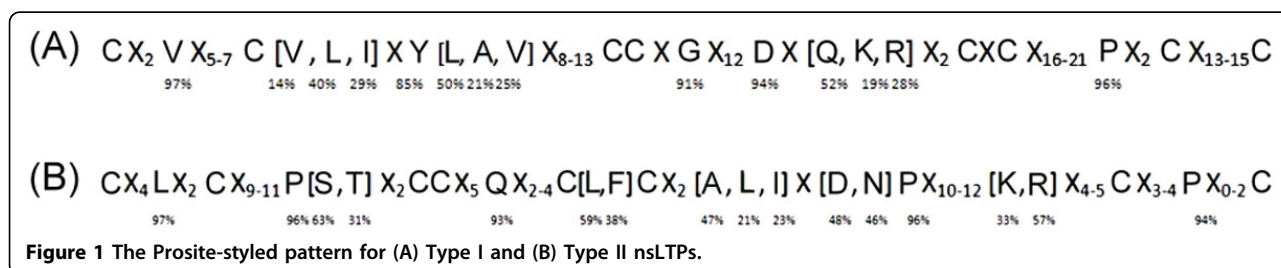
**The mungbean nsLTP1**

To evaluate our Prosite-styled pattern for the Type I nsLTPs, we referred to some mutagenesis experimental data and computational results. By using alanine scanning, we have previously identified that Asn9, Leu10, Cys13, Leu17, Leu35, Arg44, Val47, Ala66, Leu69, and Tyr79 are important to the lipid transfer activity of the LTP1 from mungbean because alanine substitutions at these residues increased the lipid transfer activity (Figure 2) [48]. For Leu10, Val31, Ile34, Arg44, Leu51, Leu69 and Val75, which are located in the hydrophobic cavity, this might be because the substituting alanine decreased the hydrophobic stack of the cavity and thus make the structure slightly loosed, creating more space to accommodate the lipid molecules (Figure 2A). Consistently, according to our new Prosite-styled pattern for the Type I nsLTPs, to which the LTP1 of mungbean belongs, the extents of sequence conservation for these functionally important positions were quite high. The occurrence probability for [Leu, Ile or Val] at the position 10, [Ile, Val or Leu] at the position 14, [Val, Ile or Leu] at the position 31, [Ile or Leu] at the position 34, Asp at the position 43, Arg at the position 44, [Leu, Ile, Phe or Val] at the position 51, [Leu or Ile] at the position 69, Val at the position 75, Tyr at the position 79, and Ile at the position 81 are mostly ≥86%. Note that most of the conserved residues are hydrophobic. We supposed that these highly conserved residues may play important roles in nsLTPs.

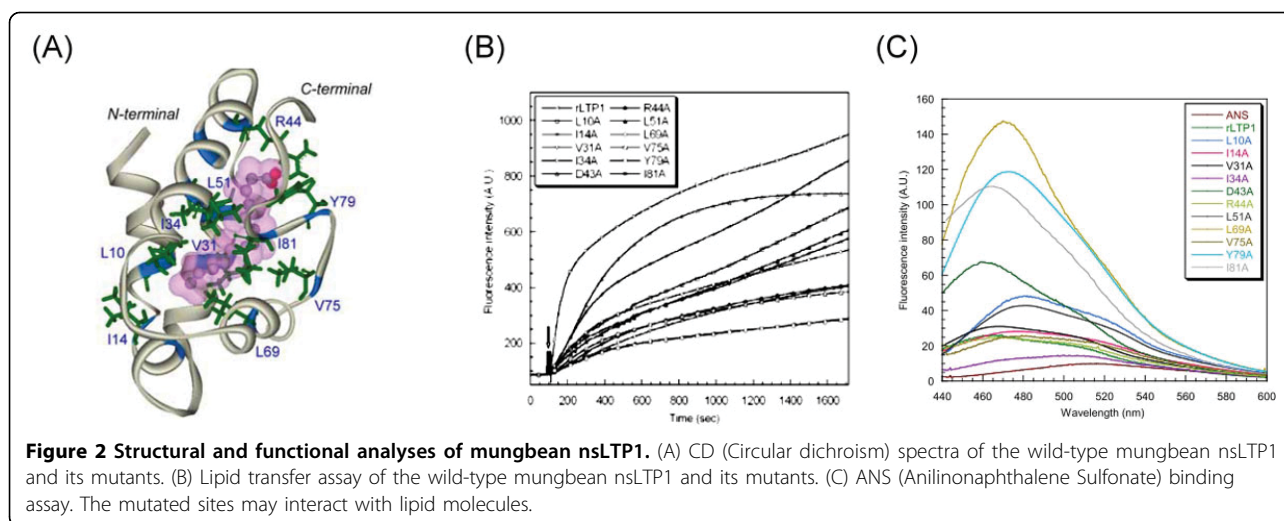
**The rice nsLTP2**

We have used the protein-protein docking model (Auto-dock) [49] to investigate the importance of the conserved hydrophobic residues, e.g., Leu8, Phe36, Phe39, Tyr45, Tyr48 and Val49, around the binding cavity of rice nsLTP2 (Figure 3A) [50]. The results indicated that changing a single residue of Leu8, Phe36, or Val49 to alanine was sufficient to destroy the integrity of the cavity. Other mutant proteins (i.e., F39A, Y45A, and Y48A) typically had native-like structure but were less stabilized compared with the wild type nsLTP2 (Figure 3B and 3C). According to our Prosite-styled pattern for the Type II nsLTPs, to which the rice LTP2 belongs, the sequence of these structurally important residues are highly conserved. The occurrence frequencies for Leu at the position 8, [Phe or Leu] at the position 36, [Phe or Tyr] at the position 39 position, [Tyr or Phe] at position 48, and [Val or Ile] at the position 49 are all ≥92%. The occurrence of [Phe or Leu] at position 45 is also high (75%). Thus, residues with higher conservation may play an important role in structural stability or lipid binding ability.

These results revealed that our Prosite-styled patterns can provide potential residues that are important to the structural and functional properties of plant nsLTPs. Interestingly, we also noticed that, in the structures of the Type I and Type II nsLTPs, there are several highly conserved (the occurrence frequency of the major amino acids > 90%) positions never studied in previous researches (see Figure 4), and most of these residues are located in alpha-helices or close to the binding cavity for lipids. We supposed that these residues may be good targets for future studies on nsLTPs.



**Figure 1 The Prosite-styled pattern for (A) Type I and (B) Type II nsLTPs.**

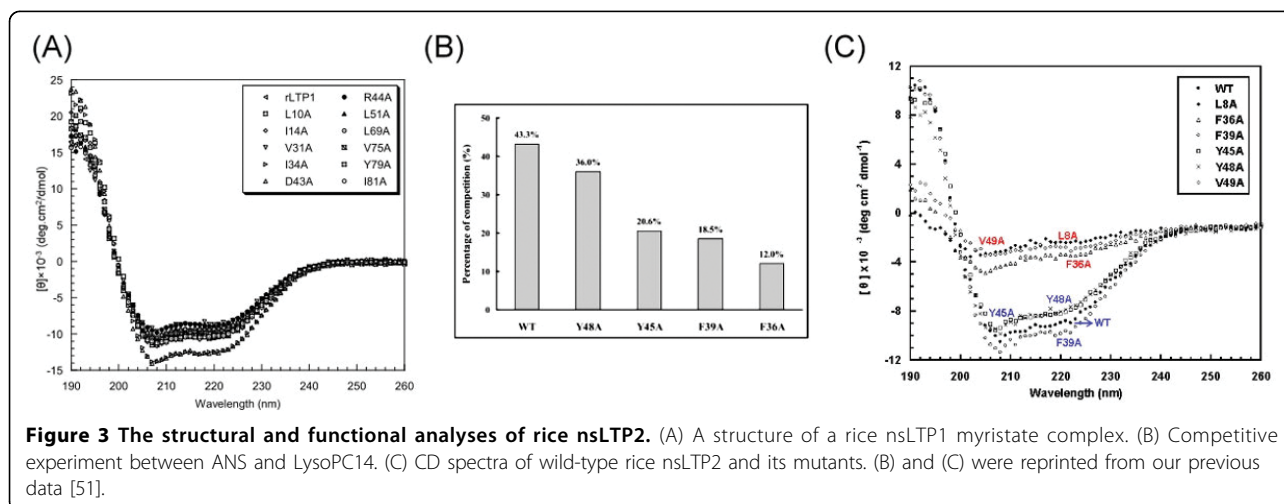


### The nsLTPDB

This database is composed of a web-based user interface collecting all nsLTPs sequence and related information (Figure 5). Our database was made up of five parts including Homepage, Species Browsing, Structure Browsing, Related References, and some useful tools. In our database, there are currently 1,395 putative non-specific lipid transfer protein sequences and 32 PDB structures. Each part is easily accessible by clicking on the hyperlink shown at the left side of the browser window. In addition, the web-based molecular viewer Chem3D <http://accelrys.com/products/informatics/cheminformatics/chime/no-fee.php> is provided to display the protein conformation. This program allows users to view and manipulate images of molecules structure in three dimensions.

### Conclusions

1. We have constructed an nsLTPs database, which provides the information of sequences, structures, literatures as well as biological data of all plant nsLTPs <http://nsltpdb.life.nthu.edu.tw/>. There are 595 nsLTPs contained in this database.
2. The phylogenetic tree of the identified nsLTPs was constructed using UPGMA and neighbor-joining clustering algorithms.
3. The 595 nsLTPs were clustered into five different types based on the sequence similarity matrix of them and the properties of their 8-cysteine motifs.
4. We compared the Prosite results with experimental mutagenesis data and found that highly conserved residues in the nsLTP sequence may play an





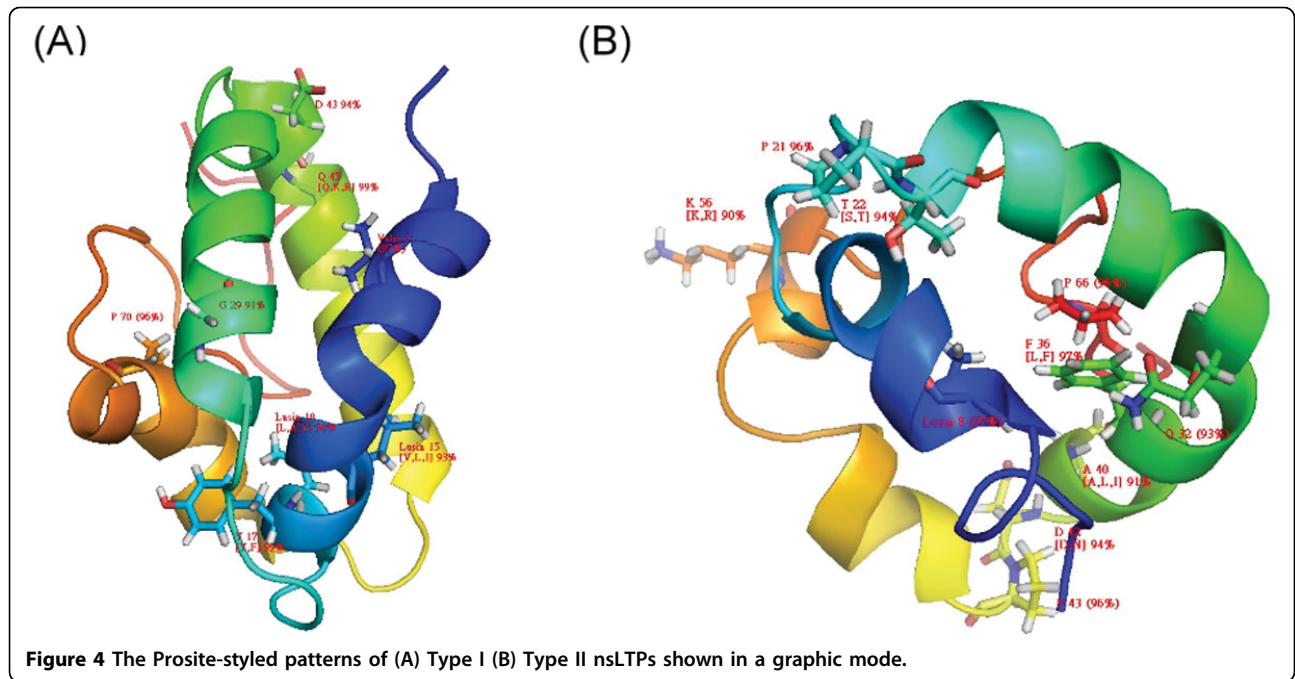


Figure 4 The Prosite-styled patterns of (A) Type I (B) Type II nsLTPs shown in a graphic mode.

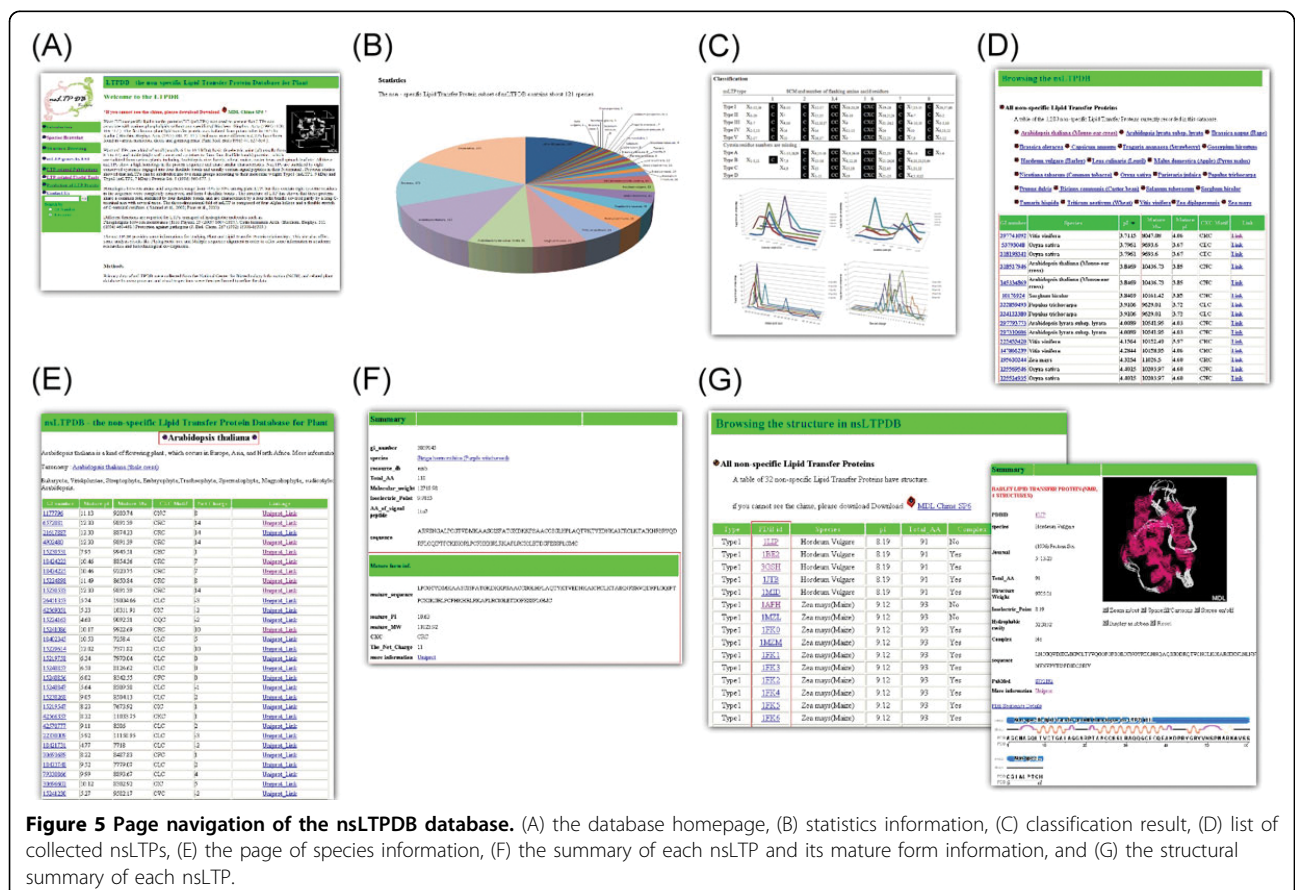


Figure 5 Page navigation of the nsLTPDB database. (A) the database homepage, (B) statistics information, (C) classification result, (D) list of collected nsLTPs, (E) the page of species information, (F) the summary of each nsLTP and its mature form information, and (G) the structural summary of each nsLTP.

important role in structural stability and/or lipid binding ability of nsLTP.

5. We created Prosite-styled patterns for nsLTPs, which are supposed useful for future identifications and studies of nsLTPs.

## Additional material

**Additional file 1: The difference between nsLTP1 and nsLTP2.** This file is in PDF format and contains schematic representations of difference between nsLTP1 and nsLTP2. Molecular weight, disulfide bond, 8-Cys patterns, the size of hydrophobic cavity, and the structures are indicated in this file.

**Additional file 2: Diversity of the eight cysteine motif in various types of nsLTPs.** This file is in PDF format and contains the diversity of eight cysteine motif of nsLTPs.

**Additional file 3: Distribution of (A) Mw, (B) pl, (C) CXC, and (D) net charge of the five types of nsLTPs defined in this work.** This file is in PDF format. Note that in figure (C) × represents a number of intervening residues between two conserved cysteines.

**Additional file 4: The unrooted phylogenetic tree constructed with the UPGMA clustering algorithm.** This file is in PDF format and contains phylogenetic tree of 595 nsLTPs.

**Additional file 5: The unrooted phylogenetic tree constructed with the neighbor-joining clustering algorithm.** This file is in PDF format and contains phylogenetic tree of 595 nsLTPs.

## Acknowledgements

This article has been published as part of *BMC Genomics* Volume 13 Supplement 1, 2012: Selected articles from the Tenth Asia Pacific Bioinformatics Conference (APBC 2012). The full contents of the supplement are available online at <http://www.biomedcentral.com/1471-2164/13?issue=S1>.

## Author details

<sup>1</sup>Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu, Taiwan. <sup>2</sup>Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan. <sup>3</sup>Institute of Bioinformatics and Systems Biology, National Chiao Tung University, Hsinchu, Taiwan. <sup>4</sup>Department of Medical Science, National Tsing Hua University, Hsinchu, Taiwan. <sup>5</sup>Graduate Institute of Molecular Systems Biomedicine, China Medical University, Taichung, Taiwan.

## Authors' contributions

NJW contributed to data collecting, experimental design, data analysis, and database construction. YFY, MNC, and CSC performed wet-lab experiments and analyzed the data. CCL contributed to the design of the web interface of our database and helped analyze computational data. WCL helped review the manuscript. PCL conceived and coordinated the study.

## Competing interests

The authors declare that they have no competing interests.

Published: 17 January 2012

## References

- Tai SP, Kaplan S: Phospholipid transfer proteins in microorganisms. *Chem Phys Lipids* 1985, **38**(1-2):41-50.
- Crain RC, Zilversmit DB: Two nonspecific phospholipid exchange proteins from beef liver. I. Purification and characterization. *Biochemistry* 1980, **19**(7):1433-1439.
- Kader JC: Lipid-transfer proteins in plants. *Annu Rev Plant Physiol Plant Mol Biol* 1996, **47**:627-654.
- Seedorf U, Scheek S, Engel T, Steif C, Hinz HJ, Assmann G: Structure-activity studies of human sterol carrier protein-2. *J Biol Chem* 1994, **269**(4):2613-2618.
- Wilmanns M, Stanley WA, Filipp FV, Kursula P, Schuller N, Erdmann R, Schliebs W, Sattler M: Recognition of a functional peroxisome type 1 target by the dynamic import receptor Pex5p. *Mol Cell* 2006, **24**(5):653-663.
- Ostergaard J, Vergnolle C, Schoentgen F, Kader JC: Acyl-binding lipid-transfer proteins from rape seedlings, a novel category of proteins interacting with lipids. *Biochim Biophys Acta* 1993, **1170**(2):109-117.
- Kader JC: Proteins and intracellular exchange of lipids. I. Stimulation of phospholipid exchange between mitochondria and microsomal fractions by proteins isolated from potato-tuber. *Biochim Biophys Acta* 1975, **380**(1):31-44.
- Jose-Estanyol M, Gomis-Ruth FX, Puigdomenech P: The eight-cysteine motif, a versatile structure in plant proteins. *Plant Physiol Biochem* 2004, **42**(5):355-365.
- Jégou S, Douliez JP, Mollé D, Boivin P, Marion D: Purification and structural characterization of LTP1 polypeptides from beer. *J Agric Food Chem* 2000, **48**(10):5023-5029.
- Lee JY, Min K, Cha H, Shin DH, Hwang KY, Suh SW: Rice non-specific lipid transfer protein: the 1.6 Å crystal structure in the unliganded state reveals a small hydrophobic cavity. *J Mol Biol* 1998, **276**(2):437-448.
- Douliez JP, Jégou S, Pato C, Larré C, Mollé D, Marion D: Identification of a new form of lipid transfer protein (LTP1) in wheat seeds. *J Agric Food Chem* 2001, **49**(4):1805-1808.
- Shin DH, Lee JY, Hwang KY, Kim KK, Suh SW: High-resolution crystal-structure of the nonspecific lipid-transfer protein from maize seedlings. *Structure* 1995, **3**(2):189-199.
- Tchang F, This P, Stiefel V, Arondel V, Morch MD, Pages M, Puigdomenech P, Grellet F, Delseny M, Bouillon P, et al: Phospholipid transfer protein: full-length cDNA and amino-acid sequence in maize. Amino-acid sequence homologies between plant phospholipid transfer proteins. *J Biol Chem* 1988, **263**(32):16849-16855.
- Takishima K, Watanabe S, Yamada M, Suga T, Mamiya G: Amino acid sequences of two nonspecific lipid-transfer proteins from germinated castor bean. *Eur J Biochem* 1988, **177**(2):241-249.
- Bouillon P, Drischel C, Vergnolle C, Duranton H, Kader JC: The primary structure of spinach-leaf phospholipid-transfer protein. *Eur J Biochem* 1987, **166**(2):387-391.
- Marion D, Douliez JP, Michon T, Elmorjani K: Structure, biological and technological functions of lipid transfer proteins and indolines, the major lipid binding proteins from cereal kernels. *J Cereal Sci* 2000, **32**(1):1-20.
- Lehmann MS, Pebayproula E, Cohenaddad C, Odani S: Crystallographic data for soybean hydrophobic protein. *J Mol Biol* 1989, **210**(1):235-236.
- Han GW, Lee JY, Song HK, Chang C, Min K, Moon J, Shin DH, Kopka ML, Sawaya MR, Yuan HS, et al: Structural basis of non-specific lipid binding in maize lipid-transfer protein complexes revealed by high-resolution X-ray crystallography. *J Mol Biol* 2001, **308**(2):263-278.
- Samuel D, Liu YJ, Cheng CS, Lyu PC: Solution structure of plant nonspecific lipid transfer protein-2 from rice (*Oryza sativa*). *J Biol Chem* 2002, **277**(38):35267-35273.
- Pons JL, de Lamotte F, Gautier MF, Delsuc MA: Refined solution structure of a liganded type 2 wheat nonspecific lipid transfer protein. *J Biol Chem* 2003, **278**(16):14249-14256.
- Breiteneder H, Mills C: Nonspecific lipid-transfer proteins in plant foods and pollens: an important allergen class. *Curr Opin Allergy Clin Immunol* 2005, **5**(3):275-279.
- Lerche MH, Kragelund BB, Bech LM, Poulsen FM: Barley lipid-transfer protein complexed with palmitoyl CoA: the structure reveals a hydrophobic binding site that can expand to fit both large and small lipid-like ligands. *Structure* 1997, **5**(2):291-306.
- Lindorff-Larsen K, Winther JR: Surprisingly high stability of barley lipid transfer protein, LTP1, towards denaturant, heat and proteases. *FEBS Lett* 2001, **488**(3):145-148.
- Hendriks TME, Thoma S, Kader J-C, De Vries SC: The carrot extracellular lipid transfer protein EP2: quantitative aspects with respect to its putative role in cutin synthesis. In *NATO ASI Series, Plant Molecular Biology. Volume H 81*. Springer-Verlag Berlin Heidelberg;Coruzzi G, Puigdomenech P 1994:85-94.



25. Garcia-Ormedo F, Molina A, Segura A, Moreno M: **The defensive role of nonspecific lipid-transfer proteins in plants.** *Trends Microbiol* 1995, **3**(2):72-74.
26. Cameron KD, Teece MA, Smart LB: **Increased accumulation of cuticular wax and expression of lipid transfer protein in response to periodic drying events in leaves of tree tobacco.** *Plant Physiol* 2006, **140**(1):176-183.
27. Liu Q, Zhang Y, Chen SY: **Plant protein kinase genes induced by drought, high salt and cold stresses.** *Chinese Sci Bull* 2000, **45**(13):1153-1157.
28. Sancho AI, Rigby NM, Zuidmeer L, Asero R, Mistrello G, Amato S, Gonzalez-Mancebo E, Fernandez-Rivas M, Ree R, Mills ENC: **The effect of thermal processing on the IgE reactivity of the non-specific lipid transfer protein from apple, Mal d 3.** *Allergy* 2005, **60**(10):1262-1268.
29. Sorensen SB, Bech LM, Muldbjerg M, Beenfeldt T, Breddam K: **Barley lipid transfer protein 1 is involved in beer foam formation.** *Tech Q Master Brew Assoc Am* 1993, **30**(4):136-145.
30. Hoffmann-Sommergruber K: **Plant allergens and pathogenesis-related proteins. What do they have in common?** *Int Arch Allergy Immunol* 2000, **122**(3):155-166.
31. Nieuwland J, Feron R, Huisman BAH, Fasolino A, Hilbers CW, Derksen J, Mariani C: **Lipid transfer proteins enhance cell wall extension in tobacco.** *Plant Cell* 2005, **17**(7):2009-2019.
32. Boutrot F, Chantret N, Gautier MF: **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* 2008, **9**:86.
33. Liu W, Huang D, Liu K, Hu S, Yu J, Gao G, Song S: **Discovery, identification and comparative analysis of non-specific lipid transfer protein (nsLtp) family in Solanaceae.** *Genomics Proteomics Bioinformatics* 2010, **8**(4):229-237.
34. Birney E, Clamp M, Durbin R: **GeneWise and Genomewise.** *Genome Res* 2004, **14**(5):988-995.
35. McGinnis S, Madden TL: **BLAST: at the core of a powerful and diverse set of sequence analysis tools.** *Nucleic Acids Res* 2004, **32**:W20-W25.
36. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A: **ExpPASy: the proteomics server for in-depth protein knowledge and analysis.** *Nucleic Acids Res* 2003, **31**(13):3784-3788.
37. Nielsen H, Engelbrecht J, Brunak S, von Heijne G: **Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites.** *Protein Eng* 1997, **10**(1):1-6.
38. Berman HM: **The Protein Data Bank: a historical perspective.** *Acta Crystallogr A* 2008, **64**:88-95.
39. Sigrist CJA, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V, Bairoch A, Hulo N: **PROSITE, a protein domain database for functional characterization and annotation.** *Nucleic Acids Res* 2010, **38**:D161-D166.
40. Bairoch A: **Prosite - a dictionary of sites and patterns in proteins.** *Nucleic Acids Res* 1991, **19**:2241-2245.
41. de Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N: **ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins.** *Nucleic Acids Res* 2006, **34**:W362-W365.
42. Sigrist CJA, De Castro E, Langendijk-Genevaux PS, Le Saux V, Bairoch A, Hulo N: **ProRule: a new database containing functional and structural information on PROSITE profiles.** *Bioinformatics* 2005, **21**(21):4060-4066.
43. Lorenzen S, Sieg J: **Phylip, Paup, and Hennig-86 - how reliable are computer parsimony programs used in systematics?** *J Zool Syst Evol Res* 1991, **29**(5-6):466-472.
44. Sokal RMC: **A statistical method for evaluating systematic relationships.** *University of Kansas Science Bulletin* 1958, **38**:1409-1438.
45. Saitou N, Nei M: **The neighbor-joining method - a new method for reconstructing phylogenetic trees.** *Jpn J Genet* 1986, **61**(6):611-611.
46. Kumar S, Tamura K, Dudley J, Nei M: **MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0.** *Mol Biol Evol* 2007, **24**(8):1596-1599.
47. Huson DH, Richter DC, Rausch C, DeZulian T, Franz M, Rupp R: **Dendroscope: an interactive viewer for large phylogenetic trees.** *BMC Bioinformatics* 2007, **8**:460.
48. Yang Y-F: **Site-directed mutagenesis studies of mung bean nonspecific lipid transfer protein 1.** *Unpublished Master's thesis, National Tsing Hua University, Taiwan, ROC* 2006.
49. **New, faster AutoDock (TM) 3.0 offers pharmaceutical scientists major enhancements to speed structure-based drug design.** *J Chem Educ* 2000, **77**(3):319.
50. Chen M-N: **Mutagenesis study on the structure conformation and binding mechanism of rice nonspecific lipid transfer protein 2.** *Unpublished Master's thesis, National Tsing Hua University, Taiwan, ROC* 2005.
51. Cheng CS, Chen MN, Lai YT, Chen T, Lin KF, Liu YJ, Lyu PC: **Mutagenesis study of rice nonspecific lipid transfer protein 2 reveals residues that contribute to structure and ligand binding.** *Proteins* 2008, **70**(3):695-706.
52. Edstam MM, Viitanen L, Salminen TA, Edqvist J: **Evolutionary history of the non-specific lipid transfer proteins.** *Mol Plant* 2011, **4**:947-964.

doi:10.1186/1471-2164-13-S1-S9

**Cite this article as:** Wang et al.: Construction and analysis of a plant non-specific lipid transfer protein database (nsLTPDB). *BMC Genomics* 2012, **13**(Suppl 1):S9.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

