# Thaumatin-like proteins: Molecular characterization, evolutionary analysis and expression profiling in major cereal crops

**Alok Sharma[1], Himanshu Sharma[1, 2] and Santosh Kumar Upadhyay[*1]**

1. Department of Botany, Panjab University, Chandigarh, India-160014

2. I.K. Gujral Punjab Technical University, Kapurthala, Jalandhar, India-144603

*Corresponding Author

***Corresponding Author: Dr. Santosh Kumar Upadhyay**, Department of Botany, Panjab University, Chandigarh, India-160014. Email: skupadhyay@pu.ac.in, Tel: +91-172-2534001, Fax: +91-172-2779510

# Abstract

**Background**

PR superfamily, pathogenesis related (PR) proteins divided into P1-PR17 subfamilies, is involved in the defense mechanism of plants. Thaumatin-like proteins (TLPs) are part of PR-5 family that facilitates resistance against abiotic and biotic stress responses.

**Results**

In the present study, a total of 26, 27, 39, 93 and 37 TLPs with conserved signature-motif were identified in *Brachypodium distachyon*, *Oryza sativa*, *Sorghum bicolor*, *Triticum aestivum* and *Zea mays*, respectively. They were phylogenetically divided into 10 different clades. Majority of genes possessed one to three exons. Physicochemical analysis suggested significant differences between small and long TLPs. Majority consisted of signal peptide and localized in extracellular region. Evolutionary analyses suggested the role of duplication events in the expansion of this gene family. Tissue-specific expression of certain genes suggested their role in development, while differential expression revealed their putative roles in stress-response. The occurrence of assorted set of *cis*-regulatory elements further suggested their diverse functions. Co-expression analysis depicted their interaction with several stress responsive genes.

**Conclusions**

The detailed characterization of TLP proteins in cereals depicted their importance in combating various biotic and abiotic challenges. This study would be useful in further characterization of individual *TLP* genes in each plant, and various crop improvement programs in future studies.



**Keywords:** Abiotic stress, Biotic stress, Expression, Thaumatin-like proteins, *Triticum aestivum*

## Introduction

Plants have been facing numerous challenges over the course of evolutionary period. Various features have been acquired by plants to tackle the challenging conditions such as climatic change, pathogenic attacks, tissue injury, salinity etc. Growth and development of plants are strongly affected by such stress conditions. Stress conditions including abiotic such as heat, drought, salinity, cold etc. On the other hand, biotic stresses are caused by living organisms such as bacteria, fungi, nematode etc. Plants are adapting and gaining ways of mechanisms to restrain these stress causing agents. For example, production of phytoalexins, proline, sugars, late embryogenesis abundant (LEA) proteins, Pathogenesis related (PR) proteins etc. [1, 2].

PR proteins are a superfamily of defense related proteins named due to their involvement in pathogen stress responses. PR proteins superfamily includes PR1-PR17 sub families, classified on the basis of their different roles and mechanisms [3].

PR5, also known as thaumatin-like proteins (TLPs) family, is one of the important classes of PR superfamily [3, 4]. These proteins were nomenclatured as TLPs due to their significant similarity with a ~23kDa sweet-tasting protein known as thaumatin [5]. Thaumatin protein was first isolated from an African shrub known as *Thaumatococcus danielli* [6]. Over the period of time, TLPs have been reported in a diverse group of organisms including insects, nematodes, fungi and plants [7-13].

In plants, TLPs are found to be present in algae, bryophytes, gymnosperm, and angiosperm [12-15]. The TLP proteins are grouped into two different forms, long TLPs and small TLPs, based on their molecular weight (MW). Unlike to the cosmopolitan nature of long TLPs, the small TLPs were found to be confined in gymnosperm and monocots only [12, 15, 16]. Furthermore, the long TLPs have a molecular weight of ~21-26 kDa, while Small TLPs have MW of ~16-17 kDa. A total of 16 and 10 conserved cysteine residues are present in long and Small TLPs, respectively. These cysteine residues form eight disulfide linkages in L-type and five disulfide linkages in S-type TLPs. These disulfide bonds provide resistance to TLPs against extreme pH, heat and protease degradation [17, 18]. Moreover, structural properties revealed that TLPs have a conserved REDDD motif and a thaumatin signature motif G-X-[GF]-X-C-X-T- [GA]-D-C-X-(1,2)-G-X-(2,3)-C. The REDDD motif is known to be involved in receptor binding for their antifungal action [12, 19]. TLPs are known for their diverse roles in abiotic and biotic stress tolerance in plants. Various research groups have reported that overexpression of TLPs increases

resistance against various fungus in both dicot and monocot plants including tobacco, soybean, potato, grapes, rice, wheat etc. [20-25]. The mechanism of TLPs in fungal resistance is ambiguous; however, these are assumed to work by degradation and permeabilization of the fungal cell walls [26, 27]. In addition, TLPs are known to be involved in antifreeze action and abiotic stress resistance in plants [21, 28-30]. Furthermore, TLPs are reported to be involved in flowering, fruit ripening and seed germination [31-33].

TLP is an important gene family because of its diverse roles in plant development and defense mechanisms. A detailed genome-wide analysis in crops including *B. distachyon, S. bicolor* and *T. aestivum* has not been carried out till date. Furthermore, a detailed *cis*-regulatory, physicochemical, expression and coexpression analysis of TLPs in *O. sativa* and *Z. mays* has been carried out in this study, in addition to the previous work [13]. In the present study, we have performed the comprehensive genomic analysis of the TLP protein family in five crops including *B. distachyon*, *O. sativa*, *S. bicolor*, *T. aestivum* and *Z. mays*. TLPs were studied in detail to explore their chromosomal distribution, phylogeny, physicochemical properties, gene duplication events, *cis*-regulatory elements and structural configuration. Moreover, to decipher their putative roles in plant development and under abiotic and biotic stresses, their expression and coexpression analysis was performed using high throughput RNA-seq data. The study will enhance the knowledge about the TLP protein family in five studied cereals. It will shed light on the putative roles of TLPs in various stress and developmental processes, which eventually provides a base for detailed functional studies and various crop improvement programs.

## Results

### Identification and Chromosomal localization of *TLP* genes

An extensive genomic search of the *TLPs* in *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays* resulted in the identification of 26, 27, 39, 93 and 37 *TLP* genes, respectively. All of the identified genes were having a complete thaumatin signature motif (Additional file 1: Table S1). The genes lacking or having incomplete thaumatin signature motifs were excluded from the study. In this study, four *TLPs* were identified as small *TLPs (sTLPs)* in each *B. distachyon, O. sativa* and *Z. mays*. A total of 10 and 20 genes were categorized into *sTLPs* in *S. bicolor* and *T. aestivum*, respectively (Additional file 1: Table S1).

The chromosomal localization analysis suggested the scattered distribution pattern of *TLPs* on most of the chromosomes of the respective crops (Fig. 1). In *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays, TLP* genes were localized on 5/5, 11/12, 9/10, 21/21 and 9/10 chromosomes. In *B. distachyon,* chromosome number 4 possessed the highest 10 *TLPs* (both long and small *TLPs*), while chromosome numbers 12, 8, and 1 were having the highest number of *TLPs* in *O. sativa, S. bicolor* and *Z. mays,* which were 6, 11 and 13, respectively. In case of *T. aestivum,* the maximum number of *TLPs* (i.e. 13) were present on chromosome number 5A, while the location of two genes were unknown. Unlike long *TLP* genes, *sTLPs* were localized specifically on chromosome numbers 12, 8 and 1 in *O. sativa*, *S. bicolor* and *Z. mays*, respectively. In *B. distachyon*, *sTLPs* were specifically present on chromosome numbers 1 and 4; however, *sTLPs* in *T. aestivum* were observed on all the chromosomes except chromosome number 1.

The nomenclatures of *TLP* genes in *B. distachyon, O. sativa, S. bicolor* and *Z. mays* were allotted on the basis of their specific location on their respective chromosomes. In case of *T. aestivum*, homeologous grouping resulted in 32 homeologous groups, which were named according to the procedure described in methods.

**Multiple sequence alignments and tertiary structure of TLPs**

TLPs have various characteristic and conserved residues, which are important for their functional activities. To explore these conserved regions and their similarity with thaumatin protein (P02883.2|THM1_THADA), TLPs of these five cereals along with thaumatin protein were aligned and studied (Fig. 2; Additional file 2: Figure S1). It was found that most of the long TLPs have 16, while all the small TLPs possessed 10 conserved cysteine residues in all the five crops. In addition, 7-5 cysteine residues were predicted in some of the TLPs from *O. sativa, S. bicolor, T. aestivum* and *Z. mays*. REDDD motif residues were conserved in most of the TLPs; however, some of these residues were replaced by amino acids with equivalent or different properties in a few TLPs. Thaumatin signature motif is a characteristic feature of thaumatin-like proteins, which was well conserved in these TLPs. Some other regions like FF hydrophobic motif and bottom of acidic cleft forming amino acids were found well conserved in most of the TLPs.

A comparative study of the structural configuration of one representative of each small TaTLPs (TaTLP6-A) and long TaTLPs (TaTLP2-A) was carried out in the context of their secondary and tertiary structures. Secondary structure analysis predicted the highest percentage of coil region followed by β strands and α helices in both long and small TLPs. The tertiary structure revealed the three complete domains (I-III) in the long TLP, while small TLP protein was lacking a major region of domain II (Fig. 3a, b). In TaTLP2-A, domain I was comprised of 9 β strands and was the largest one. Further, Domain II was mainly consisted of three small and a large α helix structure, while the domain III was the smallest one and constituted of two β strands and a long loop region. In case of TaTLP6-A, sTLP, domain II was comprised of only two α helices, while domain I and III configuration were found similar in both small and long TLP. The cysteine residues were conserved in both small and long TLP proteins. Moreover, eight and five disulfide bonds were present in TaTLP2-A and TaTLP6-A, respectively. Furthermore, the bioinformatics-based ligand prediction showed iodine, β-D glucose, dipicolinic acid, ascorbic acids, and meso-tartaric acid as probable ligands for both TaTLPs.

**Phylogenetic analysis**

To decipher the evolutionary relationship of the TLP family among five cereals and *A. thaliana*, a phylogenetic tree was built using full-length peptide sequences. These TLPs were clustered into 10 different clades based on their phylogenetic relatedness, named as groups I to X (Fig. 4). The highest number of TLPs were found in group X followed by group II, while group IV was smallest with only five genes. Interestingly, all the sTLPs were specific to group X only. Furthermore, it was observed that groups I, II, VII, VIII and IX consisted of TLPs from all the five cereals, while Group III, V and VI lack the members of one or more plant species. However, group IV was composed of only three TaTLP and two AtTLP proteins. Additionally, it was found that, with few exceptions, most of the TLPs of Group I, III, IV, IX and X were with 3, 1, 3, 2 and 1 exons, respectively. Moreover, the ranges of exon varied from 1-3, 2-3, 2-11, 2-4 and 1-4 in groups II, V, VI, VII and VIII, respectively. The homeologous *TLPs* of *T. aestivum* were found together in close proximity.

## Duplication events investigation

We have performed the duplication event (DE) analysis of *TLPs* of all five crops. DEs were studied to understand their roles in the evolution and expansion of *TLP* gene family. A total of one, four, three, six and eight DEs were found in *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z mays* (Fig. 1; Additional file 3: Table S2). All the paralogous genes in *B. distachyon* (*BdTLP1-BdTLP2*), *O. sativa* (*OsTLP5-OsTLP7, OsTLP22-OsTLP24, OsTLP23-OsTLP25, OsTLP26-OsTLP27*) and *S. bicolor* (*SbTLP28-SbTLP29, SbTLP31-SbTLP34, SbTLP32-SbTLP33*) were due to the tandem duplication events. In the case of *T. aestivum*, five out of six DEs were segmental (*TaTLP1-A-TaTLP24-D, TaTLP1-B-TaTLP24-B, TaTLP12-B-TaTLP13-D, TaTLP15-A-TaTLP18-B, TaTLP15-B2-TaTLP18-A*) and one was tandem DE (*TaTLP16-A-TaTLP17-A*). In *Z. mays*, seven segmental (*ZmTLP1-ZmTLP35, ZmTLP2-ZmTLP34, ZmTLP3-ZmTLP22, ZmTLP9-ZmTLP12, ZmTLP13-ZmTLP37, ZmTLP16-ZmTLP33, ZmTLP20-ZmTLP25*) and one tandem (*ZmTLP6-ZmTLP7*) DEs were found. In the context of paralogous genes, it was observed that all the duplicated gene pairs of each cereal crop were tightly clustered in the same clade, for instance, BdTLP1 and BdTLP2, OsTLP5 and OsTLP7, SbTLP28 and SbTLP29, TaTLP1-B and TaTLP24-B, ZmTLP16 and ZmTLP33.

## Ka/Ks and Tajima's relative rate test

Over the course of evolution, various evolutionary forces and natural pressures affect the duplicated genes [34]. To understand the evolutionary divergence between the paralogous gene pairs of these five cereals, we carried out the Ka/Ks analysis. The value of more than one and less than one of Ka/Ks ratio suggests the positive (Purifying) and negative (non-purifying) selection pressure, respectively. All the paralogous genes showed the Ka/Ks value lesser than one, which suggested the negative or purifying selection on *TLP* duplicated genes (Table 1). However, Ka/Ks analysis for *ZmTLP6* and *ZmTLP7* was not performed due to their 100% similarity. Additionally, the divergence time of DEs was also calculated using the Ks value and previously described methods [35, 36]. The divergence time of duplicated genes was calculated as 83 million years ago (MYA) in *BdTLPs*, while the range varied from 52-108, 6-137, 63-99 and 1-63 MYA in *O. sativa, S. bicolor, T. aestivum* and *Z. mays*, respectively (Table 1).

The Tajima's relative rate test found the insignificant $\chi 2$ value for all the duplicated gene pairs at $P > 0.05$ (Table 2). The P value of more than 0.05 depicted their acceptance of the molecular evolutionary clock hypothesis [37].

**Gene architecture, domain and motif analyses**

Gene structural organization analysis includes exon/intron arrangements and phase distribution in *TLPs*. The number of exons varied from one to three in *O. sativa,* while one to four in *B. distachyon, S. bicolor* and *T. aestivum*. In case of *Z. mays,* most of the *TLPs* had a range of one to three exons, while *ZmTLP16, ZmTLP20, ZmTLP25* and *ZmTLP33* consisted of eleven, seven, eight and ten exons, respectively. A total of 98 out of 222 *TLPs* of five crops were intronless. Intriguingly, except *TaTLP-6A* and *TaTLP20-A1*, all the *sTLPs* were intronless. Furthermore, the intron phase analysis revealed that the maximum number of introns were in phase 1, followed by phase 2, while the least number of introns were in phase zero (Additional file 4: Figure S2).

Functional nature of a protein depends on the type of domains present in it. All of the TLPs from these five cereals possess a thaumatin domain (PF00314). The average domain sizes of long TLP in *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays* were approx. 217, 207, 216, 214 and 202 AA, respectively. On the other hand, average thaumatin domain sizes of sTLPs were 154, 151, 147,134 and 148 amino acids (AA) in *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays*, respectively. In addition to the thaumatin domain, ZmTLP16, ZmTLP20, ZmTLP25 and ZmTLP33 also had a nuclear protein 96 (NUP96) domain. The NUP96 domain of size ~211 AA was present on the cytosolic terminal of these proteins (Additional file 5: Table S3).

Motif investigation revealed the 10 most conserved motifs in TLPs. Motifs 1-9 were the parts of the thaumatin domain, while motif 10 was unknown. Furthermore, motifs 5, 6 and 8 were the most conserved in TLP proteins. The thaumatin signature sequence was present in motif 8 (Additional file 4: Figure S2).

**Physicochemical properties**

To understand their several important features, analysis of several physicochemical properties of identified TLPs were studied. TLPs were analyzed for their MW, peptide length, pI, subcellular localization, transmembrane helix and signal peptide (Additional file 6: Table S4). The average peptide lengths of long TLPs were 299, 289, 294, 284 and 334 AA and of small TLPs were 185,

178, 178, 177 and 175 AA in *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays*, respectively. The average MW were 30.8, 29.8, 30.0, 29.1 and 35 kDa for long TLPs, while 19.5, 18.1, 18.3, 18.2 and 17.7 kDa for small TLPs in *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays*, respectively. The average isoelectric points were 6.18, 6.95, 6.05, 6.18 and 5.89 for long TLPs and 6.82, 6.06, 5.74, 6.17 and 4.90 for sTLPs in *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays*, respectively (Additional file 6: Table S4). The higher average MW and peptide length of ZmTLPs with respect to other crops were due to the presence of four TLPs (ZmTLP16, ZmTLP20, ZmTLP25 and ZmTLP33) with an additional domain.

Transmembrane (TM) helix analysis predicted the 21/26, 21/27, 29/39, 84/93 and 30/37 TLPs were with no TM region, while 5/26, 5/27, 10/39, 7/93 and 7/37 TLPs consisted of single TM region in *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays*, respectively. However, OsTLP18, TaTLP4-D and TaTLP5-D were comprised of two TM helices each (Additional file 6: Table S4). The analysis of subcellular localization predicted that most of the TLPs were localized in the extracellular region, while three TLPs from *O. sativa* (OsTLP10, OsTLP15 and OsTLP16) shown nuclear and BdTLP16, TaTLP21-A, ZmTLP20, ZmTLP25 and ZmTLP33 shown plasma membrane localization (Additional file 6: Table S4). Signal peptides were found in 23/26, 21/27, 36/39, 86/93 and 33/37 TLPs of *B. distachyon, O. sativa,* and *Z. mays*, respectively (Additional file 6: Table S4). All the sTLPs were devoid of the transmembrane helix, and one signal peptide was predicted in all of them.

**Promoter region analyses**

Based on their putative functions, the identified *cis*-regulatory elements were segregated into four types including growth and development, light, hormone and stress-responsive elements (Additional file 7: Table S5). The promoter regions comprised of a number of growth-related elements including TE2F2NTPCNA, EBOX, RYREPEAT4, RAV1AAT, POLLEN1LEAAT52 etc. In case of light-responsive elements, GT1CONSENSUS, GATA box, GBOXLERBCS, BOXII, IBOXCORE were some common *cis*-regulatory elements. However, under the hormonal responsive category, DPBFCOREDCD3, ARFAT, ARR1, ERELEE4 and GARE10OSREP1 were abscisic acid, auxin, cytokinin, ethylene and gibberellic acid responsive, respectively. Moreover, various abiotic and biotic stress-responsive elements were also found in most of the

studied TLPs, for instance, DRE2COREZMRAB17, ASF1, BIHD1, CBFHV, CGG-box, GCCCORE, LTRECOREATCOR15, MYB, MYC, WRKY1, W-box, T/G-box etc.

**Expression profiling of *TLP* genes under different tissues and developmental stages**

Expression analysis of genes is a significant way to understand their involvement in various developmental and physiological processes. *TLP* genes in plants are known to play roles in various tissues and their development pathways. To decipher the role of *TLPs* in *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays,* an expression study was carried out using high throughput RNA-seq data [38-40]. Expression of *TLPs* was very low, hence we have selected only the genes with FPKM value ≥ 0.2 in at least one of the tissue developmental stage for further analysis. This filter criterion led to the selection of 23, 27, 37, 84 and 29 *TLPs* in *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays,* respectively. In each plant species, the genes in heat maps formed the different cluster groups on the basis of their expression patterns (Fig. 5a-e). In case of *B. distachyon*, the genes were clustered into 1-3 groups. *BdTLPs* of group 1 were highly expressed in early and emerging inflorescence stages, while *BdTLP9, BdTLP12* and *BdTLP23, BdTLP28* were upregulated in endosperm and embryo, respectively. Group 2 genes depicted upregulated expression in plant embryo, which was followed by early and emerging inflorescence tissues. Moreover, group 2 genes showed their varied expression in leaf, pistil, embryo and endosperm tissues. The group 3 genes were maximum upregulated in 10 days after pollination (DAP) of seeds, while *BdTLP8* and *BdTLP18* showed their upregulated expression in pistil tissue as well (Fig. 5a). In *B. distachyon, TLPs* showed their expression in all tissues and their stages, which suggested their involvement in vegetative and reproductive growth and development. In *O. sativa, TLPs* were clustered into five groups (Fig. 5b). All of the group 1 *OsTLPs* showed upregulated expression in seed except *OsTLP27,* which showed upregulation in root tissue. The majority of group 2 *OsTLPs* showed high expression in callus. Group 3 genes showed their maximum expression in the shoot; however, *OsTLP14* and *OsTLP18* were also upregulated in the early inflorescence. *OsTLPs* belonging to group 4 showed upregulation in early inflorescence, while *OsTLP12* and *OsTLP19* were upregulated in carpel tissue as well. The result depicted their roles in tissues including both vegetative and reproductive. In *S. bicolor,* four groups were formed, in which group 1 genes showed differential expression in root tissues except *SbTLP8, SbTLP23* and *SbTLP27,* which were expressed highly

in anther tissue. In group 2, most of the genes showed high expression in the pistil, while *SbTLP4*, *SbTLP10* and *SbTLP26* were differentially expressed in seed as well. Similarly, *SbTLP13*, *SbTLP17*, *SbTLP38* and *SbTLP39* were shown upregulation in plant embryo. *SbTLP17* exhibited upregulated expression in all tissues except leaf and anther. In group 3, half of the genes showed their highest expression in the endosperm, while another half showed in flower. In case of group 4, the maximum genes were showing differential expression in leaf (Fig. 5c). In *Z. mays*, genes were clustered into four different groups (Fig. 5e). Most of the genes of group 1 showed significant expression in all stages of tissue development; however, half of the genes were seen to be significantly upregulated in unpollinated silk while the other halves were expressing in pericarp and aleurone layer. Majority of group 2 genes were specifically upregulated to root tissues, while some genes also showed upregulation in internode tissue, namely *ZmTLP17*, *ZmTLP19* and *ZmTLP28*. Group 3 genes were mostly downregulated but showed significant expression in the embryo. Group 4 genes showed upregulated expression in most of the tissues; however, the maximum expression of these genes observed in leaf followed by internode tissue.

In *T. aestivum*, the genes were high in number and were clustered into four large groups (Fig. 5d). Group 1 was the largest group and most of the genes showed differential expression in root tissue followed by shoot. In addition, genes like *TaTLP29-A, B, D* and *TaTLP4-D* showed their upregulated expression in spike tissue as well. Group 2 genes showed highest upregulation in spike and grain; however, some genes such as *TaTLP2-D1* and *TaTLP1*-A also showed varied expression in root tissue. In group 3, most of the genes exhibited upregulated expression in grain, while *TaTLP3-A, U, TaTLP13-D, TaTLP32-A, B,* and *D* were highly expressing in leaf, followed by root. Group 4 *TaTLPs* were expressing higher in grain, while *TaTLP7-A, B, D* and *TaTLP30-A* were also upregulated in spike. Overall *TaTLPs* were highly expressed in root, stem and grain tissues, while some genes were also upregulated in spike and leaf.

**Expression profiling of *TaTLP* genes under biotic and abiotic stress conditions**

TLPs belong to defense-related protein family and are well known for their stress-responsive behavior. The expression analysis of *TaTLP* genes using RNA-seq data was performed under various biotic and abiotic stress conditions to understand their roles in stress resistance [41-43]. In biotic stress, RNA-seq data generated at 24, 48 and 72 h after infestation of two important

fungal pathogens, namely, *Blumeria graminis f. sp. tritici* (Bgt) and *Puccinia striiformis f. sp. tritici* (Pst) was used for expression pattern analysis [41]. These fungal pathogens are causing agent of powdery mildew and stripe rust diseases, respectively. In *T. aestivum*, a total of 50 genes were analyzed because of their differential expression in these stress conditions and clustered into four different groups based on their expression pattern (Fig. 6a). In group 1, most of the genes were expressed highly in Bgt infection, while these genes were downregulated or showed mild expression after Pst infestation. However, *TaTLP23-A* was upregulated in all the conditions. In groups 2 and 3, most of the genes were low expressing and showed their expression mainly in Pst infection. However, *TaTLP27-D* showed upregulation under both Bgt and Pst infection. The group 4 genes were upregulated at the 24 days after Pst infestation. In Bgt infection, *TaTLP3-A2*, *TaTLP12-A*, *TaTLP32-A* were most differentially upregulated genes with 37, 36 and 35 folds, respectively. However, in comparison to Bgt infection, under Pst infection, a small number of genes, for instance, *TaTLP10-A* (51 fold up), *TaTLP27-D* (17 fold up) etc. showed upregulated expression.

Abiotic stresses are another major threat to plants, which affect the various physiological and biological processes of plants. To understand the putative role of *TaTLPs* in abiotic stress conditions, expression analysis was carried out under drought (DS), heat (HS) and combined stress of heat and drought (HD) stress using online available RNA-seq data [42]. A total of 52 *TaTLPs* were analyzed which formed four different clusters in the heatmap (Fig. 6b). Most of the genes in group 1 were downregulated in DS, HS and HD stresses. However, *TaTLP15-B3* and *TaTLP27-B* exhibited significantly upregulated expression under HS and HD. In group 2, only three genes namely, *TaTLP5-D*, *TaTLP9-A* and *TaTLP9-D* were highly expressing in all three stresses. Most of the *TaTLPs* from group 3 exhibited higher expression at 6 h of DS, while few genes of group 3 such as *TaTLP14-A1, B1*, *TaTLP25-D* etc., showed differential expression in all the stress conditions. In group 4, *TaTLP7-D* showed upregulated expression at 1 h of HS and HD stress, while *TaTLP28-A* was upregulated at both 1 and 6 h after HS treatment.

Expression analysis of *TaTLPs* under salt stress was done using online available RNA-seq data [43]. Out of 93 genes, 63 genes were found to be differentially expressing and used for further expression analysis. These 63 genes were clustered into 4 groups, based on their expression patterns (Fig. 6c). In group 1, genes were downregulated in all the salt treatment conditions, while group 2 genes showed similar expression as control conditions. However, *TaTLP2-B*

exhibited 6 fold upregulated expression in initial hours of salt treatment. In group 3, *TaTLPs* exhibited upregulated expression at 24 and 48 h of salt treatment. Precisely, *TaTLP3-U*, *TaTLP14-B1*and *B5* showed upregulated expression at all the time intervals of salt stress with most 18, 9 and 10 fold up change at salt 48 h, respectively. Group 4 genes were upregulated at 12 and 24 h and showed downregulation at 48 h of salt treatment.

**Co-expression networking of *TLPs***

To understand the interacting network of *TLP* genes, a co-expression study was carried out under various tissue developmental stages of each studied cereal (Fig. 7a-e). Additionally, in *T. aestivum*, the co-expression study was also performed under a biotic and abiotic stress conditions (Fig. 8a-c). In *B. distachyon*, a total of 64 transcripts were co-expressing with three *BdTLP* genes. Out of these three genes, *BdTLP17* exhibited co-expression with maximum 51 transcripts. These co-expressed gene encode hydrolases, chitinases, kinases, defense-related, DNA binding, etc. proteins (Fig. 7a; Additional file 8: Table S6). In the case of *O. sativa*, seven *OsTLPs* were co-expressed with a total of 101 transcripts. *OsTLP4* and *OsTLP1* exhibited co-expression with 29 and 25 transcripts, respectively. Various proteins, for instance, PR-1, hydrolases, ion binding, proteases, kinases etc. were encoded by these co-expressed transcripts (Fig. 7b; Additional file 8: Table S6). In the *S. bicolor*, nine *SbTLP*s showed co-expression with a total of 137 transcripts. *SbTLP15*, *SbTLP36* and *SbTLP5* showed maximum co-expression with 38, 32 and 29 transcripts, respectively. Co-expressed transcripts included various proteins responsible for ion binding, protein degradation, stress-related, development-related etc. (Fig. 7c; Additional file 8: Table S6). In *Z. mays*, a total of four *ZmTLPs* exhibited co-expression with 97 transcripts. The *ZmTLP29* showed co-expression with 68 transcripts encoding responsive proteins for drought, hormone, oxidative stress, biotic stress etc. (Fig. 7e; Additional file 8: Table S6).

In case of *T. aestivum*, under tissue and developmental stages, a total of 10 *TaTLPs* revealed their co-expression with 121 transcripts. Most co-expressed genes were *TaTLP31-B* and *TaTLP20-A6* with 53 and 51 transcripts, respectively. These co-expressed transcripts translate different proteins, like the transporter, kinase, ATP synthase, ethylene-responsive etc. (Fig. 7d; Additional file 8: Table S6). However, under biotic stress, most of the co-expressed genes were interacted with *TaTLP18-B* and these interacted genes encode proteins related to signaling, transport, DNA repair, disease resistance pathways (Fig. 8a; Additional file 9: Table S7).

Moreover, in case of HD and salt, 15 and 35 *TaTLPs* were co-expressed with 109 and 102 transcripts, respectively. However, *TaTLP14-B5* and *TaTLP10-A* exhibited co-expression with most of the transcripts in HD and salt stress, respectively. Annotation revealed that co-expressed transcripts encode various proteins for stress-resistant, hormone-responsive, kinase, transporters etc. (Fig. 8b, c; Additional file 9: Table S7). Furthermore, gene ontology study of co-expressed transcripts also predicted their involvement in antifungal response (GO:0009817), antibacterial stress response (GO:0042742), drought stress response (GO:0009414), salt stress response (GO:0005634), oxidative stress response (GO:0006979), signaling transduction (GO:0007165), calcium ion transport activity (GO:0070588), metal ion binding activity (GO:0046872), hydrolase activity (GO:0016787), peroxidase activity (GO:0005777), plant development (GO:0009960) etc.

**Comparative expression profiling of *TaTLP* paralogous genes**

To understand the functional relationship between the paralogous genes, a comparative expression profiling of each paralogue pair was done. Generally, on the basis of their expression pattern, duplicated genes could categorize into three different types; retention of function, pseudo-functionalization and neo-functionalization. Out of six paralogous genes, three pairs including *TaTLP1-A-TaTLP24-D, TaTLP1-B-TaTLP24-B, TaTLP15-A-TaTLP18-B* showed a similar trend of expression, which suggested these as retention of function type (Fig. 9a-c). However, two duplicated pairs of *TaTLPs* (*TaTLP12-D-TaTLP13-D, TaTLP16-A-TaTLP17-A*) had insignificant expression in one of the gene under most of the tissues, which kept these pairs under pseudo-functional category (Fig. 9d, e). *TaTLP15-B2* and *TaTLP18-A* lack the expression data in most of the tissue; therefore, they were excluded for comparative analysis. Maximum number of retention of function type and lack of neo-functionalization kind suggested the conserved nature of *TLPs* during the evolution.

## Discussion

TLPs are member of defense proteins and these are known to play pivotal roles in both abiotic and biotic stresses [44-46]. Moreover, these proteins are also found to be involved in plant development [31-33]. Therefore, we have studied TLPs in five major cereals including *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays*. In the present study, we have

identified a total of 26, 27, 39, 93 and 37 *TLPs* in the genome of *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays*, respectively. Previous studies reporting different numbers of *TLPs* in Arabidopsis (24), false brome (26), field mustard (20), grape (33), poplar (49), rice (49), sorghum (36), maize (49), moss (6) and white pine (10) [13-15, 47-49] suggests the presence of a varied number of TLPs in different plants. The higher number of *TLPs* in *T. aestivum* as compare to other species might be attributed to their complex allohexaploid genome configuration (AABBDD) [50]. Furthermore, the different extent of expansion of *TLPs* among different plant species might be attributed to the number of duplication events. The duplication events are responsible for the expansion of gene families and these are the major forces to attain neo functionality by causing genetic variability [51-53]. In DEs analysis, our results suggested that tandem DEs are a major factor in *B. distachyon, O. sativa* and *S. bicolor* and segmental DEs in *T. aestivum* and *Z. mays* behind the gene expansion of *TLP* family. The higher number of paralogous genes in *T. aestivum* and *Z. mays* might be attributed to their larger genome size and the presence of more transposable elements [53, 54]. In the year 2016, Cao et al., also found the role of tandem and segmental types of duplication in the evolution of *TLP* gene family in Arabidopsis, rice, poplar, and maize [13]. Furthermore, the divergence time of monocots from eudicots was estimated as ~170-235 Mya, which were further diverged into grasses before ~77 Mya [54, 55]. Time calculation of paralogous gene formation in *B. distachyon* in our study showed the occurrence of DE before the divergence of grasses. The main duplication incidence in *S. bicolor* and *O. sativa* genome was estimated around ~70 Mya [56]. Most of the paralogous *TLPs* of *O. sativa* and *S. bicolor* suggested their formation after the main DE occurrence, while few pair showed their emergence before the grass divergence during the analysis. In *Z. mays*, paralogous genes depicted their time of occurrence near to divergence time of maize from sorghum which was predicted to be around 11-28 Mya [56]. However, two duplicated pairs, having NUP 96 domain showed their recent incidence of DEs. In *T. aestivum*, paralogous genes showed their occurrence much earlier than hybridization event of A, B and D sub genome [50]. Duplicated genes face various environmental forces and natural pressures during the process of evolution [34]. In our study, all the paralogous gene pairs showed purifying selection as a major force of selection. The similar results were found in the evolutionary study of six plant species and melon in which purifying selection of *TLPs* was reported, while the positive selection was experienced by only few amino acid residues [13, 57]. However, in poplar, positive selection was

also experienced in 10 *TLPs*, which suggested the varied natural selection procedure in *TLPs*' evolution [14]. Furthermore, Tajima's relative test found the equal rate of evolution in paralogous genes, which again suggested the conserved nature of *TLP* genes.

TLPs were found to be paraphyletic in nature and were supposed to be originated from 10 ancestral genes [11]. Similarly, in our study, phylogenetic analysis showed the clustering of TLPs in 10 different clades. The uneven pattern of genes in different clades was observed, where maximum genes were present in the group number 10. The high number of paralogous genes in group 10 suggested a higher rate of expansion in this group. However, specific plant TLPs in some groups suggested the species-specific expansion of these groups. Intron-exon numbers were found to be equal in most of the members of the same clades, while half of the clades showed varied intron-exon numbers amongst them. The intron-less nature of group 10 showed their evolved nature, which was well supported by existence of all the sTLPs exclusively in this group. The similar results with uneven distribution and paraphyletic nature of TLPs have also been reported by previous studies [13, 57]. The close clustering of paralogous genes and homeologous groups further confirmed the grouping of TLPs in the phylogenetic tree.

The organization of introns and exons of *TLPs* have been reported to be in the range of one to ten exons [12, 13, 15, 49]. On similar patterns, our results suggest comparable findings with one to four exons in most of the *TLPs*. However, most of the monocots' *TLPs* have intron-less nature, which was also found in our analysis [13]. Additionally, our results with respect to the protein lengths, molecular weights and isoelectric points of long and small TLP proteins, found in accordance with previously studied TLPs [12, 13, 15]. Most of the TLPs have N-terminal signal peptides and are localized to the secretory pathways [12, 13]. Similar results in respect to localization and signal peptide were obtained in our study. Furthermore, lack of transmembrane helix in most of the TLPs, confirm their extracellular location and non-membranous nature. In addition, the presence of thaumatin-like domain in all the TLP proteins, make our studies similar to previously reported findings. However, four ZmTLPs (ZmTLP16, ZmTLP20, ZmTLP25, and ZmTLP33) have an additional NUP 96 domain. The function of the NUP 96 domain was found to be involved in plant development and stress resistance against fungal and bacterial pathogens, which suggest the role of TLPs in signaling cascade [58, 59]. Most of the motifs were conserved in all the TLPs, which again suggested their similar functional roles and conservation among them. The presence of conserved residues and their similarity with thaumatin protein was

explored with multiple sequence alignment. In multiple sequence alignment, all the 10 and 16 cysteine residues in most of the small and long TLPs were found to be conserved, respectively. The REDDD, FF hydrophobic, thaumatin signature motifs were found to be conserved among most of the TLPs, which suggested their functional and structural conserved nature with previously studied TLPs from different plants [12, 14, 15, 48, 49, 57].

The secondary and tertiary structural modeling of TaTLP2-A and TaTLP6-A provided an overview of the three-domain structural aspects of the TLPs. The three-domain structure of both sTLPs and long TLPs showed significant differences in the domain architecture. Partial presence of domain II in TaTLP6-A might be important; however, it is yet unexplored. Further detailed analysis is required to unveil the impact of the absence of domain II in small TLPs. Conserved disulfide bonds in both representatives of small and long TLPs might be responsible for the structural integrity of the protein. In the year 2008, Ghosh and Chakrabarti reported the involvement of conserved cysteine residues in the structural integrity of TLPs in tomato [17]. Furthermore, a thaumatin-like xylanase inhibitor of *T. aestivum*, TLX1, has been reported to possess five disulfide bonds and three β sheets [16]. In addition, the presence of, three domains pattern, a varied number of β sheets, alpha helices, and disulfide bonds have also been reported in tobacco, banana, tomato, grape and sodom apple TLPs [17, 60-63]. More than one ligand was predicted in a bioinformatic based analysis of both TLP proteins. In a study, the binding interactions of TLPs to carbohydrates like (1,3)□β□D□glucans and their role in malting has been reported in barley [26, 64]. One such predicted ligand, β-D glucose, may have similar roles in the TaTLPs. The binding of ascorbic acid with TaTLPs suggested their possible roles in oxidative stress tolerance. However, experimental functional analysis is required to explore the possibilities.

Expression pattern of genes is important to depict their functional roles [65]. Expression profiling of *TLP* genes in five cereals, suggested their involvement in various tissues and their developmental stages. The expression results suggested the divergent expression of *TLP* genes in various tissues including, root, shoot, leaf, inflorescence, endosperm, embryo etc. This expression pattern depicted their roles in both vegetative and reproductive development. In *T. aestivum*, expression profiling also suggested the divergent expression of *TaTLPs* in the root, leaf, stem, spike and grain, which depicted their functional involvement in these plant organs. However, the upregulation of some genes under specific tissues hinted towards their role in the

development of the tissues. Previously, a similar trend of *TLP* expression was reported in rice [13]. Moreover, TLPs were also found to be involved in fruit ripening, seed germination, and nodule specificity in different plants [33, 66, 67].

The expression studies of *TaTLP* genes under various abiotic stress conditions such as heat, drought, combined heat and drought and salt stress showed their roles in these stress conditions. In case of DS, most of the *TaTLP* genes were upregulated in initial hours; however, they showed downregulation afterward. Under HS and HD stress, *TaTLPs* showed differential expression at later stages of stress treatments. These results suggested the involvement of *TaTLPs* at different time intervals of these stresses. The salinity stress treatment unveiled the participation of a few *TaTLP* genes in 48 and 72 hours of salt stress. Our results suggest possible roles of *TLPs* against different abiotic stress conditions in *T. aestivum.* In a previous study, the role of *TLPs* as heat shock protein was suggested in rice [13]. In another study upregulated expression of *TLPs* at different hours of drought, salt and cold treatment in *Brassica rapa* was reported [48]. Moreover, over-expression of *TLPs* was found to increase the resistance against various abiotic stresses in tobacco, cotton, cynanchum, peanut, Arabidopsis etc. [20, 44, 45, 68, 69]. Furthermore, *TLPs* are well studied with respect to their antifungal properties. Expression profiling of *TaTLPs* depicted their upregulated expression in response to Bgt and Pst infection. Half of the *TaTLPs* were highly expressed upon Bgt infestation and showed their downregulated expression in Pst, while other half showed a reverse expression pattern in Bgt and Pst. There are studies, which showed the involvement of PR5 genes or specifically *TLPs* of *T. aestivum* in response to fungal attacks. For instance, the induction of *TaPR5* by stripe rust and *TaLr35PR5* by leaf rust is reported [25, 46, 70]. Furthermore, various other studies also found the role of *TLP*s to increase the resistance of wheat against *Fusarium graminearum* [71-73], *Microdochium nivale* [74] and *Stagonospora nodorum* [75]. In addition to wheat, some other plants' TLPs have also been reported to be involved against various fungal pathogens, for instance, *Alternaria sp*. [76, 77], *Colletotrichum gloeosporioides* [78], *Fusarium sp*. [79], *Podosphaera xanthii* [57], *Rhizoctonia solani* [45, 80], *Sclerotinia sclerotiorum* [81], *Verticillium dahliae* [68] etc. The results of expression analysis of studied five cereals suggest the putative roles of *TLPs* in various abiotic and biotic stress conditions.

Promoter elements are known to be necessary for the regulation of gene expression under different conditions. The presence of *cis*-regulatory elements responsible for various important

functions suggests the possible involvement of *TLPs* in different processes. The expression of *TLP* genes were reported to be regulated by various factors like hormones, wound, pathogens, osmotic stress [5]. The expression profiling of *TaTLPs* revealed the significant upregulated expression of *TaTLP3-A2*, *TaTLP12-A* and *TaTLP32-A* under biotic stress, which is well supported by the occurrence of various diseases and wound responsive elements such as W-box elements in these genes. Earlier, the involvement of W-box elements in fungal response has been reported in a rice *TLP* (*Rtlp1*) [82]. Similarly, the presence of drought, salt and cold (MYB, MYC, DRE2COREZMRAB17, DRECRTCOREAT, LTRECOREATCOR15) responsive elements in *TLPs* confirm their involvement in abiotic stress tolerance. The promoter sequence analysis has suggested that *TLPs* might be involved in the defense mechanism as well as the developmental processes of plants.

During co-expression analysis, the majority of *TLPs* made an interaction network with chitinases, kinases, hydrolases, transporters etc. These co-expressed proteins were earlier found to be involved in disease resistance mechanisms [83-86]. Our study suggests the possible roles of TLPs in signaling, transport, fugal cell wall lysis etc., which was well supported by the earlier findings [13, 87]. However, to functionally validate the *TLP* genes in various stress conditions, functional characterization of each candidate gene is needed to be done in future studies.

## Conclusions

Thaumatin-like proteins (TLPs), a part of PR5 protein family, are known to be involved in various biotic and abiotic stress conditions. The curiosity to unravel the various possibilities led us to detailed in silico characterization of a total of 222 *TLP* genes in five cereals. Factors like heat, drought, cold, pathogenic attacks, tissue injury have made a negative impact on the worldwide production of cereals. Here in this study, we have analyzed the physical and chemical characteristics of TLP proteins along with the analyses such as evolutionary, domain structure, transmembrane, *cis*-regulatory element, multiple sequence alignments, sub-cellular localization, and secondary and tertiary structure analysis. Furthermore, expression analysis has given us strong indications about the important roles of *TLPs* in combating abiotic and biotic stresses along with playing roles in various plant development stages. The co-expression study has provided us the prediction about the interaction of *TLP* genes with other stress related genes of respective cereals. The detailed analysis will lay down a path for various crop improvement

programs for important cereals including *T. aestivum*, *O. sativa*, *Z. mays*, *S. bicolor*, *B. distachyon*.

## Methods

### Identification and nomenclature of *TLP* genes

Extensive BLAST searches were used to identify the TLP proteins in five different cereals including *B. distachyon*, *O. sativa*, *S. bicolor*, *T. aestivum* and *Z. mays*. Arabidopsis TLP sequences were used as query sequences against the protein model sequences of each crop for separate BLAST searches at Ensembl Plants [15, 87]. Presence of the thaumatin (PF00314) domain was analyzed using Hidden Markov Model (HMM) and Pfam blast searches at e-value $10^{-10}$. Identified TLP sequences were further subjected to NCBI Conserved Domain Database (CDD) BLAST for the confirmation of the presence of the thaumatin domain. In terms of splice variants, only the longest transcripts were selected in each crop. *TLP* genes with complete thaumatin family signature were selected for further analysis. Furthermore, TLPs having kinase were also removed as they have already been characterized [87]. The sequences of crops were named on the basis of their chromosomal distribution with the exception of *T. aestivum*. The international rules for gene symbolization of *T. aestivum* (http://wheat.pw.usda.gov/ggpages/wgc/98/Intro.htm) were followed for the nomenclature of *TaTLPs*.

### Chromosomal localization and duplication events

Plant Ensemble (http://plants.ensembl.org/) was used for gathering the chromosomal and sub-genomic location of *TLPs* of all five crops. Furthermore, homeologous groups in *T. aestivum* with ≥90% sequence similarity and their presence on the related chromosomes were predicted using the best bidirectional blast hit approach. MapInspect software was used for the mapping of the *TLP* genes on their respective chromosomes (http://www.plantbreeding.wur.nl/uk/software_mapinspect.html.2012). Duplication events were identified using a bidirectional blast hit approach with sequence similarity ≥80%. Duplication events (tandem and segmental) were differentiated on the basis of distance as per the previous studies [88].

### Multiple sequence alignments, phylogeny and structural analysis

Softwares including Muscle and Multalin were used for the multiple sequence alignments of all the TLPs with a thaumatin protein (P02883.2|THM1_THADA) to find the conserved residues

[89, 90]. Evolutionary analysis was done using MEGAX software [91]. Maximum likelihood method with a bootstrap value of 1000 was used to build an evolutionary tree.

Selected TLPs from *T. aestivum* were studied for their secondary and tertiary structure analysis using Iterative Threading ASSEmbly Refinement (I-TASSER) server. The probable ligands and their binding sites were also predicted [92].

**Ka/Ks and Tajima's relativity test**

The alignment of the protein and nucleotide sequences of the paralogous gene pairs was done using ClustalOmega server (https://www.ebi.ac.uk/Tools/msa/clustalo). Further, the synonymous substitution per synonymous site (Ks) and non-synonymous substitution per non-synonymous site (Ka), and the Ka/Ks ratio was calculated using PAL2NAL programme [93]. The calculation of divergence time of each pair of duplicated genes was done using the formula T=Ks/2r, where T represents the divergence time and r represents the divergence rate. The divergence rate was assumed to be $6.5 \times 10^{-9}$ for cereals [94]. Moreover, Tajima's relativity test was performed to find out the evolutionary rate between paralogous genes [37].

**Gene structure organization**

The genomic and CDS sequences of identified *TLPs* were used for the analysis of gene structure in terms of exon-intron organization and intron phases. GSDS 2.0 server was used to prepare the pictorial display of the exon-intron organization for the sequences of each crop [95].

**Physicochemical properties of TLPs**

Various characteristics including peptide length, molecular weight and isoelectric point of each TLP were analyzed using the ExPasy tool, which were further confirmed by Ensemble plants and Sequence Manipulation Suite [96-98]. Moreover, tools including CELLO v.2.5, ngLOC, ProtComp9 and WoLF PSORT were used for the prediction of subcellular localization of TLPs [99-102]. Tools such as Phobius and DAS-Tmfilter were used for the prediction of transmembrane regions in each TLP protein [103, 104]. Further, the signal peptides were detected using the tool Phobius and SignalP [103, 105]. SMART server was used for the domain analysis, whereas, motifs were analyzed using MEME v.4.11.4 [106, 107].

***Cis*-regulatory element analysis**

To find out the promoters, 15 kb upstream sequence from initiation codon was selected in genomic sequences of each cereal. These promoter region elements were analyzed using PLACE

software [108]. Furthermore, the *cis*-regulatory elements were categorized on the basis of their functions.

**Expression profiling**

Expression analysis in various tissues of each crop was done using the high throughput RNA-seq data retrieved from URGI database (wheaturgi.versailles.inra.fr/files/RNASeqWheat/) and Expression ATLAS [38-40]. In *T. aestivum*, the expression data generated in replicates for various tissue developmental stages, under biotic (fungal pathogen) and abiotic (heat, drought and salt) were used [41-43]. Data generated for three developmental stages of root, stem, leaf, spike and grain was used to analyze the tissue-specific expression in wheat. Further, RNA-seq data (PRJNA243835) generated in triplicates after the 24, 48 and 72 hour (h) of inoculation of *Blumeria graminisf. sp. tritici* (Bgt) and *Puccinia striiformisf. sp.tritici* (Pst) in seven days old seedlings was used for the *TaTLP* expression analysis under biotic stress [41]. Under abiotic stress conditions, duplicate RNA-seq data (SRP045409) for 1 and 6 hours treatments of heat (40º C), drought (20% PEG 6000) and the combination of both heat and drought stresses was used to study the *TaTLP* expression [42]. Furthermore, root RNA-seq data (SRP062745) available in triplicates under the treatment of 150mM NaCl at 6, 12, 24 and 48 treatment were used to analyze the expression of *TaTLPs* under salt stress [43]. The RNA-seq datasets were further validated using the WheatExp server. Trinity package was used to calculate the expression value in terms of FPKM value [109]. Hierarchical Clustering Explorer 3.5 was used to generate heat maps of differentially expressed genes, which were clustered using Euclidean distance method [110].

**Co-expression and interactome analysis**

Whole expression data of each crop were used to perform co-expression analysis of *TLPs* in all the five cereals [38-40]. In addition, the co-expression analysis of *TaTLPs* under abiotic and biotic stress was performed using complete expression data of *T. aestivum*. co-expression analyses were done using the software CoExpress v.1.5 with the threshold value ≥0.9, correlation power 1 with 100 bootstraps [111]. Further, Blast2GO software was used for the functional annotation of co-expressed genes. The pictorial representation of interacting genes was made by the tool Gephi 0.9.1 [112].

**Abbreviations**

**Bgt:** *Blumeria graminis*

**PR:** Pathogenesis related

**DS:** Drought stress

**HD:** Heat and drought stress

**HS:** Heat stress

**Pst:** *Puccinia striiformis*

**sTLP:** Small thaumatin-like protein

## Declarations

### Ethics approval and consent to participate:

Not applicable

### Consent for publication:

All the authors have agreed for publication.

### Availability of data and materials

All the analyzed data of this work are included in this manuscript and its additional files. The genomic data used in this study are available at (https://plants.ensembl.org/index.html).The RNA-seq data of *B. distachyon*, *O. sativa*, *S. bicolor*, *Z. mays* and *T. aestivum* is available at Expression ATLAS and URGI database (wheaturgi.versailles.inra.fr/files/RNASeqWheat/). The RNA-seq data used for expression analyses in *T. aestivum* is online available at NCBI database with accession numbers PRJNA243845, SRP045409 and SRP062745 (https://www.ncbi.nlm.nih.gov/).

### Competing interests

The authors declare that they have no competing interests

### Funding

No Funding.

### Author's contribution

SKU conceived the idea. AS and SKU designed the experiments. AS performed the experiments. AS and HS analyzed the data. AS and HS wrote the manuscript, and AS, HS and SKU finalized the manuscript.

**Acknowledgement**

Authors are grateful to Panjab University, Chandigarh, India for research facilities, Ensembl Plants, URGI and NCBI for data availability. AS and HS is grateful to CSIR for the senior research fellowship. HS is also thankful to IKGPTU, Jalandhar for Ph.D registration.

## References


1. Kombrink E, Somssich IE. Pathogenesis-Related Proteins and Plant Defense. In: Plant Relationships. Berlin, Heidelberg: Springer Berlin Heidelberg; 1997. p. 107–28. doi:10.1007/978-3-662-10370-8_7.
2. Chadha P, Das RH. A pathogenesis related protein, AhPR10 from peanut: an insight of its mode of antifungal activity. Planta. 2006;225:213–22. doi:10.1007/s00425-006-0344-7.
3. van Loon LC, Rep M, Pieterse CMJ. Significance of Inducible Defense-related Proteins in Infected Plants. Annu Rev Phytopathol. 2006;44:135–62. doi:10.1146/annurev.phyto.44.070505.143425.
4. Christensen AB, Cho BH, Naesby M, Gregersen PL, Brandt J, Madriz-Ordenana K, et al. The molecular characterization of two barley proteins establishes the novel PR-17 family of pathogenesis-related proteins. Mol Plant Pathol. 2002;3:135–44. doi:10.1046/j.1364-3703.2002.00105.x.
5. Velazhahan R, Datta SK, Muthukrishnan S. The PR-5 family: Thaumatin-like proteins. In: Pathogenesis-Related Proteins in Plants. CRC Press; 1999. p. 107–29. doi:10.1201/9781420049299.
6. van der Wel H, Loeve K. Isolation and characterization of thaumatin I and II, the sweet-tasting proteins from *Thaumatococcus daniellii* Benth. Eur J Biochem. 1972; 31:221–5. doi:10.1111/j.1432-1033.1972.tb02522.x.
7. Kitajima S, Sato F. Plant pathogenesis-related proteins: molecular mechanisms of gene expression and protein function. J Biochem. 1999; 125:1–8. http://www.ncbi.nlm.nih.gov/pubmed/9880788. Accessed 25 Feb 2019.
8. Brandazza A, Angeli S, Tegoni M, Cambillau C, Pelosi P. Plant stress proteins of the thaumatin-like family discovered in animals. FEBS Lett. 2004; 572:3–7. doi:10.1016/j.febslet.2004.07.003.
9. Grenier J, Potvin C, Trudel J, Asselin A. Some thaumatin-like proteins hydrolyse polymeric beta-1,3-glucans. Plant J. 1999;19:473–80. doi:10.1046/j.1365-313X.1999.00551.x.
10. Sakamoto Y, Watanabe H, Nagai M, Nakade K, Takahashi M, Sato T. *Lentinula edodes tlg1* Encodes a Thaumatin-Like Protein That Is Involved in Lentinan Degradation and

Fruiting Body Senescence. PLANT Physiol. 2006;141:793–801. doi:10.1104/pp.106.076679.

11. Shatters RG, Boykin LM, Lapointe SL, Hunter WB, Weathersbee AA. Phylogenetic and Structural Relationships of the PR5 Gene Family Reveal an Ancient Multigene Family Conserved in Plants and Select Animal Taxa. J Mol Evol. 2006;63:12–29. doi:10.1007/s00239-005-0053-z.
12. Liu J-J, Sturrock R, Ekramoddoullah AKM. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. Plant Cell Rep. 2010a;29:419–36. doi:10.1007/s00299-010-0826-8.
13. Cao J, Lv Y, Hou Z, Li X, Ding L. Expansion and evolution of thaumatin-like protein (TLP) gene family in six plants. Plant Growth Regul. 2016;79:299–307. doi:10.1007/s10725-015-0134-y.
14. Liu J-J, Zamani A, Ekramoddoullah AKM. Expression profiling of a complex thaumatin-like protein family in western white pine. Planta. 2010b;231:637–51. doi:10.1007/s00425-009-1068-2.
15. Petre B, Major I, Rouhier N, Duplessis S. Genome-wide analysis of eukaryote thaumatin-like proteins (TLPs) with an emphasis on poplar. BMC Plant Biol. 2011;11:33. doi:10.1186/1471-2229-11-33.
16. Fierens E, Gebruers K, Voet ARD, De Maeyer M, Courtin CM, Delcour JA. Biochemical and structural characterization of TLXI, the *Triticum aestivum* L. thaumatin-like xylanase inhibitor. J Enzyme Inhib Med Chem. 2009;24:646–54. doi:10.1080/14756360802321831.
17. Ghosh R, Chakrabarti C. Crystal structure analysis of NP24-I: a thaumatin-like protein. Planta. 2008;228:883–90. doi:10.1007/s00425-008-0790-5.
18. Smole U, Bublin M, Radauer C, Ebner C, Breiteneder H. Mal d 2, the Thaumatin-Like Allergen from Apple, Is Highly Resistant to Gastrointestinal Digestion and Thermal Processing. Int Arch Allergy Immunol. 2008;147:289–98. doi:10.1159/000144036.
19. Zhao JP, Su XH. Patterns of molecular evolution and predicted function in thaumatin-like proteins of *Populus trichocarpa*. Planta. 2010;232:949–62. doi:10.1007/s00425-010-1218-6.

20. Rajam MV, Chandola N, Saiprasad Goud P, Singh D, Kashyap V, Choudhary ML, et al. Thaumatin gene confers resistance to fungal pathogens as well as tolerance to abiotic stresses in transgenic tobacco plants. Biol Plant. 2007;51:135–41. doi:10.1007/s10535-007-0026-8.
21. Subramanyam K, Arun M, Mariashibu TS, Theboral J, Rajesh M, Singh NK, et al. Overexpression of tobacco osmotin (Tbosm) in soybean conferred resistance to salinity stress and fungal infections. Planta. 2012;236:1909–25. doi:10.1007/s00425-012-1733-8.
22. Acharya K, Pal AK, Gulati A, Kumar S, Singh AK, Ahuja PS. Overexpression of *Camellia sinensis* Thaumatin-Like Protein, CsTLP in Potato Confers Enhanced Resistance to *Macrophomina phaseolina* and *Phytophthora infestans* Infection. Mol Biotechnol. 2013;54:609–22.
23. He R, Wu J, Zhang Y, Agüero CB, Li X, Liu S, et al. Overexpression of a thaumatin-like protein gene from *Vitis amurensis* improves downy mildew resistance in *Vitis vinifera* grapevine. Protoplasma. 2017;254:1579–89. doi:10.1007/s00709-016-1047-y.
24. Datta K, Velazhahan R, Oliva N, Ona I, Mew T, Khush GS, et al. Over-expression of the cloned rice thaumatin-like protein (PR-5) gene in transgenic rice plants enhances environmental friendly resistance to *Rhizoctonia solani* causing sheath blight disease. Theor Appl Genet. 1999;98:1138–45. doi:10.1007/s001220051178.
25. Wang X, Tang C, Deng L, Cai G, Liu X, Liu B, et al. Characterization of a pathogenesis-related thaumatin-like protein gene *TaPR5* from wheat induced by stripe rust fungus. Physiol Plant. 2010;139:27–38. doi:10.1111/j.1399-3054.2009.01338.x.
26. Osmond RI, Hrmova M, Fontaine F, Imberty A, Fincher GB. Binding interactions between barley thaumatin-like proteins and (1,3)-beta-D-glucans. Kinetics, specificity, structural analysis and biological implications. Eur J Biochem. 2001;268:4190–9. http://www.ncbi.nlm.nih.gov/pubmed/11488912. Accessed 25 Feb 2019.
27. Abad LR, D'Urzo MP, Liu D, Narasimhan ML, Reuveni M, Zhu JK, et al. Antifungal activity of tobacco osmotin has specificity and involves plasma membrane permeabilization. Plant Sci. 1996;118:11–23.
28. Yu X-M, Griffith M. Antifreeze proteins in winter rye leaves form oligomeric complexes. Plant Physiol. 1999;119:1361–70. http://www.ncbi.nlm.nih.gov/pubmed/10198095. Accessed 25 Feb 2019.

29. Zhu B, Chen TH, Li PH. Expression of three osmotin-like protein genes in response to osmotic stress and fungal infection in potato. Plant Mol Biol. 1995;28:17–26. http://www.ncbi.nlm.nih.gov/pubmed/7787181. Accessed 25 Feb 2019.
30. Zhang Y, Shih DS. Isolation of an osmotin-like protein gene from strawberry and analysis of the response of this gene to abiotic stresses. J Plant Physiol. 2007;164:68–77. doi:10.1016/j.jplph.2006.02.002.
31. Neale AD, Wahleithner JA, Lund M, Bonnett HT, Kelly A, Meeks-Wagner DR, et al. Chitinase, [beta]-1,3-Glucanase Osmotin, and Extensin Are Expressed in Tobacco Explants during Flower Formation. The Plant Cell. 1990;2:673–84. doi:10.1105/tpc.2.7.673.
32. Salzman RA, Tikhonova I, Bordelon BP, Hasegawa PM, Bressan RA. Coordinate accumulation of antifungal proteins and hexoses constitutes a developmentally controlled defense response during fruit ripening in grape. Plant Physiol. 1998;117:465–72. http://www.ncbi.nlm.nih.gov/pubmed/9625699. Accessed 25 Feb 2019.
33. Seo PJ, Lee A-K, Xiang F, Park C-M. Molecular and Functional Profiling of Arabidopsis Pathogenesis-Related Genes: Insights into Their Roles in Salt Response of Seed Germination. Plant Cell Physiol. 2008;49:334–44. doi:10.1093/pcp/pcn011.
34. Ellegren H. Comparative genomics and the study of evolution by natural selection. Molecular Ecology. 2008;17:4586–96. doi:10.1111/j.1365-294X.2008.03954.x.
35. Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, et al. Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum*/*Aegilops* complex and the evolutionary history of polyploid wheat. Proc Natl Acad Sci U S A. 2002;99:8133–8. doi:10.1073/pnas.072223799.
36. Sharma M, Singh A, Shankar A, Pandey A, Baranwal V, Kapoor S, et al. Comprehensive expression analysis of rice Armadillo gene family during abiotic stress and development. DNA Res. 2014;21:267-83. doi: 10.1093/dnares/dst056.
37. Tajima F. Simple methods for testing the molecular evolutionary clock hypothesis. Genetics. 1993;135.
38. Choulet F, Alberti A, Theil S, Glover NM, Barbe V, Daron J, et al. Structural and Functional Partitioning of Bread Wheat Chromosome 3B. doi:10.1126/science.1249721ï.

39. Pingault L, Choulet F, Alberti A, Glover N, Wincker P, Feuillet C, et al. Deep transcriptome sequencing provides new insights into the structural and functional organization of the wheat genome. Genome Biol. 2015;16:29. doi:10.1186/s13059-015-0601-9.
40. Papatheodorou I, Fonseca NA, Keays M, Tang YA, Barrera E, Bazant W, et al. Expression Atlas: gene and protein expression across multiple studies and organisms. Nucleic Acids Res. 2018;46:D246-D251. doi:10.1093/nar/gkx1158.
41. Zhang H, Yang Y, Wang C, Liu M, Li H, Fu Y, et al. Large-scale transcriptome comparison reveals distinct gene activations in wheat responding to stripe rust and powdery mildew. BMC Genomics. 2014;15:898. doi:10.1186/1471-2164-15-898.
42. Liu Z, Xin M, Qin J, Peng H, Ni Z, Yao Y, et al. Temporal transcriptome profiling reveals expression partitioning of homeologous genes contributing to heat and drought acclimation in wheat (*Triticum aestivum* L.). BMC Plant Biol. 2015;15:152. doi:10.1186/s12870-015-0511-8.
43. Zhang Y, Liu Z, Khan AA, Lin Q, Han Y, Mu P, et al. Expression partitioning of homeologs and tandem duplications contribute to salt tolerance in wheat (*Triticum aestivum* L.). Sci Rep. 2016;6:21476. doi:10.1038/srep21476.
44. Wang Q, Li F, Zhang X, Zhang Y, Hou Y, Zhang S, et al. Purification and Characterization of a CkTLP Protein from *Cynanchum komarovii* Seeds that Confers Antifungal Activity. PLoS One. 2011;6:e16930. doi:10.1371/journal.pone.0016930.
45. Singh NK, Kumar KRR, Kumar D, Shukla P, Kirti PB. Characterization of a Pathogen Induced Thaumatin-Like Protein Gene AdTLP from *Arachis diogoi*, a Wild Peanut. PLoS One. 2013;8:e83963. doi:10.1371/journal.pone.0083963.
46. Zhang J, Wang F, Liang F, Zhang Y, Ma L, Wang H, et al. Functional analysis of a pathogenesis-related thaumatin-like protein gene *TaLr35PR5* from wheat induced by leaf rust fungus. BMC Plant Biol. 2018;18:76. doi:10.1186/s12870-018-1297-2.
47. Faillace GR, Turchetto-Zolet AC, Guzman FL, de Oliveira-Busatto LA, Bodanese-Zanettini MH. Genome-wide analysis and evolution of plant thaumatin-like proteins: a focus on the origin and diversification of osmotins. Mol Genet Genomics. 2019;294:1137–57. doi:10.1007/s00438-019-01554-y.

48. Ahmed NU, Park J-I, Jung H-J, Kang K-K, Lim Y-P, Hur Y, et al. Molecular characterization of thaumatin family genes related to stresses in *Brassica rapa*. Sci Hortic. 2013;152:26–34. doi:10.1016/j.scienta.2013.01.007.
49. Yan X, Qiao H, Zhang X, Guo C, Wang M, Wang Y, et al. Analysis of the grape (*Vitis vinifera* L.) thaumatin-like protein (TLP) gene family and demonstration that TLP29 contributes to disease resistance. Sci Rep. 2017;7:4269. doi:10.1038/s41598-017-04105-w.
50. Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer M, Jakobsen KS, et al. Ancient hybridizations among the ancestral genomes of bread wheat. Science. 2014;345:1250092–1250092. doi:10.1126/science.1250092.
51. Kondrashov FA. Gene duplication as a mechanism of genomic adaptation to a changing environment. Proc R Soc B Biol Sci. 2012;279:5048–57. doi:10.1098/rspb.2012.1108.
52. Magadum S, Banerjee U, Murugan P, Gangapur D, Ravikesavan R. Gene duplication as a major force in evolution. J Genet. 2013;92:155–61. http://www.ncbi.nlm.nih.gov/pubmed/23640422. Accessed 28 Jul 2019.
53. Appels R, Eversole K, Stein N, Feuillet C, Keller B, Rogers J, et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science. 2018;361:eaar7191. doi:10.1126/science.aar7191.
54. Gaut BS. Evolutionary dynamics of grass genomes. New Phytologist. 2002;154:15–28. doi:10.1046/j.1469-8137.2002.00352.x.
55. Yang YW, Lai KN, Tai PY, Li WH. Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between Brassica and other angiosperm lineages. J Mol Evol. 1999;48:597–604. doi:10.1007/PL00006502.
56. Paterson AH, Bowers JE, Chapman BA. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Acad Sci U S A. 2004;101:9903–8.
57. Liu Y, Cui J, Zhou X, Luan Y, Luan F. Genome-wide identification, characterization and expression analysis of the TLP gene family in melon (*Cucumis melo* L.). Genomics. 2020;112:2499–509.

58. Parry G, Ward S, Cernac A, Dharmasiri S, Estelle M. The Arabidopsis suppressor of auxin resistance proteins are nucleoporins with an important role in hormone signaling and development. Plant Cell. 2006;18:1590–603. doi:10.1105/tpc.106.041566.
59. Zhang Y, Li X. A putative nucleoporin 96 is required for both basal defense and constitutive resistance responses mediated by suppressor of npr1-1, constitutive 1. Plant Cell. 2005;17:1306–16. doi:10.1105/tpc.104.029926.
60. Koiwa H, Kato H, Nakatsu T, Oda J, Yamada Y, Sato F. Crystal structure of tobacco PR-5d protein at 1.8 Å resolution reveals a conserved acidic cleft structure in antifungal thaumatin-like proteins 1 1Edited by R. Huber. J Mol Biol. 1999;286:1137–45. doi:10.1006/jmbi.1998.2540.
61. Leone P, Menu-Bouaouiche L, Peumans WJ, Payan F, Barre A, Roussel A, et al. Resolution of the structure of the allergenic and antifungal banana fruit thaumatin-like protein at 1.7-Å. Biochimie. 2006;88:45–52. doi:10.1016/j.biochi.2005.07.001.
62. Marangon M, Van Sluyter SC, Waters EJ, Menz RI. Structure of Haze Forming Proteins in White Wines: *Vitis vinifera* Thaumatin-Like Proteins. PLoS One. 2014;9:e113757. doi:10.1371/journal.pone.0113757.
63. Ramos M V., de Oliveira RSB, Pereira HM, Moreno FBMB, Lobo MDP, Rebelo LM, et al. Crystal structure of an antifungal osmotin-like protein from *Calotropis procera* and its effects on *Fusarium solani* spores, as revealed by atomic force microscopy: Insights into the mechanism of action. Phytochemistry. 2015;119:5–18. doi:10.1016/j.phytochem.2015.09.012.
64. Singh S, Tripathi RK, Lemaux PG, Buchanan BB, Singh J. Redox-dependent interaction between thaumatin-like protein and β-glucan influences malting quality of barley. Proc Natl Acad Sci U S A. 2017;114:7725–30. doi:10.1073/pnas.1701824114.
65. Durick K, Mendlein J, Xanthopoulos KG. Hunting with traps: Genome-wide strategies for gene discovery and functional analysis. Genome Research. 1999;9:1019–25. doi:10.1101/gr.9.11.1019.
66. Fils-Lycaon BR, Wiersma PA, Eastwell KC, Sautiere P. A cherry protein and its gene, abundantly expressed in ripening fruit, have been identified as thaumatin-like. Plant Physiol. 1996;111:269–73. doi:10.1104/pp.111.1.269.

67. Hayashi M, Shiro S, Kanamori H, Mori-Hosokawa S, Sasaki-Yamagata H, Sayama T, et al. A Thaumatin-Like Protein, Rj4, Controls Nodule Symbiotic Specificity in Soybean. Plant Cell Physiol. 2014;55:1679–89. doi:10.1093/pcp/pcu099.
68. Munis MFH, Tu L, Deng F, Tan J, Xu L, Xu S, et al. A thaumatin-like protein gene involved in cotton fiber secondary cell wall development enhances resistance against *Verticillium dahliae* and other stresses in transgenic tobacco. Biochem Biophys Res Commun. 2010;393:38–44. doi:10.1016/j.bbrc.2010.01.069.
69. Misra RC, Sandeep, Kamthan M, Kumar S, Ghosh S. A thaumatin-like protein of *Ocimum basilicum* confers tolerance to fungal pathogen and abiotic stress in transgenic Arabidopsis. Sci Rep. 2016;6:25340. doi:10.1038/srep25340.
70. Li XY, Gao L, Zhang WH, Liu JK, Zhang YJ, Wang HY, et al. Characteristic expression of wheat PR5 gene in response to infection by the leaf rust pathogen, *Puccinia triticina*. J Plant Interact. 2015;10:132–41. doi:10.1080/17429145.2015.1036140.
71. Chen WP, Chen PD, Liu DJ, Kynast R, Friebe B, Velazhahan R, et al. Development of wheat scab symptoms is delayed in transgenic wheat plants that constitutively express a rice thaumatin-like protein gene. Theor Appl Genet. 1999;99:755–60. doi:10.1007/s001220051294.
72. Anand A, Zhou T, Trick HN, Gill BS, Bockus WW, Muthukrishnan S. Greenhouse and field testing of transgenic wheat plants stably expressing genes for thaumatin-like protein, chitinase and glucanase against *Fusarium graminearum*. J Exp Bot. 2003;54:1101–11. http://www.ncbi.nlm.nih.gov/pubmed/12598580. Accessed 26 Feb 2019.
73. Mackintosh CA, Lewis J, Radmer LE, Shin S, Heinen SJ, Smith LA, et al. Overexpression of defense response genes in transgenic wheat enhances resistance to Fusarium head blight. Plant Cell Rep. 2007;26:479–88. doi:10.1007/s00299-006-0265-8.
74. Kuwabara C, Takezawa D, Shimada T, Hamada T, Fujikawa S, Arakawa K. Abscisic acid- and cold-induced thaumatin-like protein in winter wheat has an antifungal activity against snow mould, *Microdochium nivale*. Physiol Plant. 2002;115:101–10. doi:10.1034/j.1399-3054.2002.1150112.x.
75. Jayaraj J, Muthukrishnan S, Liang GH, Velazhahan R. Jasmonic acid and salicylic acid induce accumulation of β-1,3-glucanase and thaumatin-like proteins in wheat and

enhance resistance against *Stagonospora nodorum*. Biol Plant. 2004;48:425–30. doi:10.1023/B:BIOP.0000041097.03177.2d.

76. Velazhahan R, Muthukrishnan S. Transgenic Tobacco Plants Constitutively Overexpressing a Rice Thaumatin-like Protein (PR-5) Show Enhanced Resistance to *Alternaria alternata*. Biol Plant. 2004;47:347–54. doi:10.1023/B:BIOP.0000023876.55053.5e.
77. Safavi K, Zareie R, Tabatabaei BES. Constitutive expression of thaumatin-like protein (TLP-3) in transgenic tobacco plants leads to enhance resistance to *Alternaria alternata*. Arch Phytopathol Plant Prot. 2012;45:161–9. doi:10.1080/03235408.2010.507947.
78. Odeny Ojola P, Nyaboga EN, Njiru PN, Orinda G. Overexpression of rice thaumatin-like protein (*Ostlp*) gene in transgenic cassava results in enhanced tolerance to *Colletotrichum gloeosporioides* f. sp. manihotis. J Genet Eng Biotechnol. 2018;16:125–31. doi:10.1016/J.JGEB.2017.12.002.
79. Mahdavi F, Sariah M, Maziah M. Expression of Rice Thaumatin-Like Protein Gene in Transgenic Banana Plants Enhances Resistance to *Fusarium Wilt*. Appl Biochem Biotechnol. 2012;166:1008–19. doi:10.1007/s12010-011-9489-3.
80. Naseri G, Sohani MM, Pourmassalehgou A, Allahi S. In planta transformation of rice (*Oryza sativa*) using thaumatin-like protein gene for enhancing resistance to sheath blight. African J Biotechnol. 2012;11:7885–93. doi:10.5897/AJB11.3331.
81. Liu D, He X, Li W, Chen C, Ge F. Molecular cloning of a thaumatin-like protein gene from *Pyrus pyrifolia* and overexpression of this gene in tobacco increased resistance to pathogenic fungi. Plant Cell, Tissue Organ Cult. 2012;111:29–39. doi:10.1007/s11240-012-0167-0.

82. Hiroyuki K, Terauchi R. Regulation of expression of rice thaumatin-like protein: Inducibility by elicitor requires promoter W-box elements. Plant Cell Rep. 2008;27:1521–8. doi:10.1007/s00299-008-0536-7.
83. Woloshuk CP, Meulenhoff JS, Sela-Buurlage M, van den Elzen PJ, Cornelissen BJ. Pathogen-induced proteins with inhibitory activity toward *Phytophthora infestans*. Plant Cell. 1991;3:619–28. doi:10.1105/tpc.3.6.619.

84. Chen K, Fan B, Du L, Chen Z. Activation of hypersensitive cell death by pathogen-induced receptor-like protein kinases from Arabidopsis. Plant Mol Biol. 2004;56:271–83. doi:10.1007/s11103-004-3381-2.
85. Takenaka Y, Nakano S, Tamoi M, Sakuda S, Fukamizo T. Chitinase gene expression in response to environmental stresses in *Arabidopsis thaliana*: Chitinase inhibitor allosamidin enhances stress tolerance. Biosci Biotechnol Biochem. 2009;73:1066–71. doi:10.1271/bbb.80837.
86. Ng G, Seabolt S, Zhang C, Salimian S, Watkins TA, Lu H. Genetic dissection of salicylic acid-mediated defense signaling networks in Arabidopsis. Genetics. 2011;189:851–9. doi:10.1534/genetics.111.132332.
87. Sharma A, Shumayla, Tyagi S, Alok A, Singh K, Upadhyay SK. Thaumatin-like protein kinases: Molecular characterization and transcriptional profiling in five cereal crops. Plant Sci. 2020;290:110317.
88. Shumayla, Tyagi S, Sharma A, Singh K, Upadhyay SK. Genomic dissection and transcriptional profiling of Cysteine-rich receptor-like kinases in five cereals and functional characterization of *TaCRK68-A*. Int J Biol Macromol. 2019;134:316–29. doi:10.1016/j.ijbiomac.2019.05.016.
89. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792-97. doi:10.1093/nar/gkh340.
90. Corpet F. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. 1988;16:10881–90. http://www.ncbi.nlm.nih.gov/pubmed/2849754. Accessed 31 Aug 2017.
91. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol. 2018;35:1547-1549. doi: 10.1093/molbev/msy096.
92. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Res. 2015;43:W174-181. https://academic.oup.com/nar/article/43/W1/W174/2467872. Accessed 17 Jul 2020.
93. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 2006;34:W609-

12. https://academic.oup.com/nar/article/34/suppl_2/W609/2505720. Accessed 17 Jul 2020.

94. Gaut BS, Morton BR, McCaig BC, Clegg MT. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL. Proc Natl Acad Sci USA. 1996;93:10274. doi:10.1073/PNAS.93.19.10274.

95. Hu B, Jin J, Guo A-Y, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2015;31:1296–7. doi:10.1093/bioinformatics/btu817.

96. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: The proteomics protocols handbook. Totawa: Humana Press; 2005. p. 571-607. doi:10.1385/1-59259-890-0:571.

97. Stothard P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques. 2000;28:1102, 1104. doi:10.2144/00286ir01.

98. Nussbaumer T, Martis MM, Roessner SK, Pfeifer M, Bader KC, Sharma S, et al. MIPS PlantsDB: a database framework for comparative plant genome research. Nucleic Acids Res. 2013;41:D1144-51. doi:10.1093/nar/gks1153.

99. Yu CS, Chen YC, Lu CH, Hwang J-K. Prediction of protein subcellular localization. Proteins Struct Funct Bioinforma. 2006;64:643-51. doi:10.1002/prot.21018.

100. King BR, Guda C. ngLOC: an n-gram-based Bayesian method for estimating the subcellular proteomes of eukaryotes. Genome Biol. 2007;8:R68. doi:10.1186/gb-2007-8-5-r68.

101. Emanuelsson O, Nielsen H, Brunak S, Von Heijne G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol. 2000;300:1005–16. doi:10.1006/jmbi.2000.3903.

102. Horton P, Park K-J, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, et al. WoLF PSORT: protein localization predictor. Nucleic Acids Res. 2007;35:W585–7. doi:10.1093/nar/gkm259.

103. Käll L, Krogh A, Sonnhammer EL. A Combined Transmembrane Topology and Signal Peptide Prediction Method. J Mol Biol. 2004;338:1027–36. doi:10.1016/j.jmb.2004.03.016.
104. Cserzo M, Eisenhaber F, Eisenhaber B, Simon I. TM or not TM: transmembrane protein prediction with low false positive rate using DAS-TMfilter. Bioinformatics. 2004;20:136–7. doi:10.1093/bioinformatics/btg394.
105. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 2011;8:785–6. doi:10.1038/nmeth.1701.
106. Letunic I, Doerks T, Bork P. SMART: recent updates, new developments and status in 2015. Nucleic Acids Res. 2015;43:D257-60. doi:10.1093/nar/gku949.
107. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37:W202–8. doi:10.1093/nar/gkp335.
108. Higo K, Ugawa Y, Iwamoto M, Higo H. PLACE: a database of plant *cis* -acting regulatory DNA elements. Nucleic Acids Res. 1998;26(1):358–9.
109. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013;8:1494–512. doi:10.1038/nprot.2013.084.
110. Seo J, Gordish-Dressman H, Hoffman EP. An interactive power analysis tool for microarray hypothesis testing and generation. Bioinformatics. 2006;22:808–14. doi:10.1093/bioinformatics/btk052.
111. Nazarov PV, Muller A, Khutko V. Co-expression analysis of large microarray data sets using CoExpress software tool. Conference: WCSB 2010.
112. Bastian M, Heymann S, Jacomy M. Gephi□: An Open Source Software for Exploring and Manipulating Networks Visualization and Exploration of Large Graphs. Proc. Third Int. SM Conf. 2009. www.aaai.org. Accessed 28 Jul 2019.

## Figure legends

**Fig. 1** Chromosomal distribution and paralogous gene analysis of *TLPs* in five cereal crops. Image depicts the scattered localization of *TLP* genes of **(a)** *B. distachyon*, **(b)** *O. sativa*, **(c)** *S. bicolor*, **(d)** *T. aestivum* and **(e)** *Z. mays* on their respective chromosomes. Red lines and blue lines denote the tandem and segmental duplication events, respectively.

**Fig. 2** Multiple sequence alignment of few randomly selected long and small TLPs of *B. distachyon, O. sativa, S. bicolor, T. aestivum and Z. mays* with thaumatin protein sequence. Figure shows the alignment of long (OsTLP5, SbTLP2, BdTLP13, ZmTLP13 and TaTLP1-A) and small (BdTLP10, OsTLP25, ZmTLP6, SbTLP29 and TaTLP6-B) TLPs with a thaumatin (P02883.2|THM1_THADA) protein sequence. The characteristic thaumatin signature motif is shown in sky blue box. All the conserved cysteine residues were marked with asterisk. The REDDD motif was marked with red circle. The FF hydrophobic motif and bottom of acidic cleft forming regions are marked with green triangle and brown box, respectively. The conserved regions of thaumatin domain are marked with black line.

**Fig. 3** The comparative tertiary configuration analysis of a small and a long TLP of *T. aestivum.* The image shows the tertiary structure of a small (TaTLP2-A) and a long (TaTLP6-A) TaTLPs **(a, b)**. Figure **(a)** illustrate the TaTLP2-A and shows the three complete domains, 13 β strands, four α helix and eight disulphide bonds. Figure **(b)** is of TaTLP6-A and shows the 11 β strands, two α helix, five disulphide bond and deletion of domain II. The coils, α helices, β strands and disulphide bridges are coloured with green, red, yellow and blue colours, respectively. The ligand binding position are marked with the purple coloured sticks.

**Fig. 4** Phylogenetic tree analysis of *A. thaliana, B. distachyon, O. sativa. S, bicolor, T. aestivum and Z. mays* using full length peptide of TLPs. Figure shows the phylogenetic tree of TLPs, built by maximum likelihood approach with 1000 bootstrap values using MEGAX software. Phylogenetic tree displays the clustering of TLPs of six plant species in I-X different groups, each group is colored differently. The TLPs from each plant are marked with different colour.

**Fig. 5** Expression analysis of *TLP* genes of *B. distachyon* (*BdTLP*), *O. sativa* (*OsTLP*). *S. bicolor* (*SbTLP*), *T. aestivum* (*TaTLP*) and *Z. mays* (*ZmTLP*) under various tissue and developmental stages. Heatmaps **(a-e)** shows the clustering and expression profiling of *TLP* genes in *B. distachyon*, **(b)** *O. sativa*, **(c)** *S. bicolor*, **(d)** *T. aestivum* and **(e)** *Z. mays*. The different clusters were formed on the basis of expression values of *TLP* genes in respective crop and represents with different colour. The colour bar depict the upregulated and downregulated expression with red and green colours, respectively.

**Fig. 6** Expression profiling of *TLP* genes of *T. aestivum* under biotic and abiotic stress conditions. Heat maps **(a-c**) shows the clustering and expression profiling of *TaTLP* genes under **(a)** *Blumeria graminis* (Bgt) and *Puccinia striiformis* (Pst), **(b)** HS, DS and HD and **(c)** salt stress conditions. The clustering of *TaTLP* genes in each heat map is based on their expression pattern in respective stress condition. The colour bar shows the high and low expression with red and green colours, respectively.

**Fig. 7** Co-expression network assembly of *TLP* genes under various tissue and their development stages. Figure **(a-e)** shows the co-expression web of **(a)** *B. distachyon*, **(b)** *O. sativa*, **(c)** *S. bicolor*, **(d)** *T. aestivum* and **(e)** *Z. mays TLPs* with their interacting partners of respective genome.

**Fig. 8** Co-expression analysis of *TLP* genes of *T. aestivum* under biotic and abiotic stress conditions. Figure **(a-c)** depict the interaction web of *TaTLP* genes under **(a)** biotic **(b)** HD and **(c)** salt stress with the other co-expressed transcripts.

**Fig. 9** Comparative expression analysis of duplicated gene pairs in *T. aestivum*. Graphs **(a-e)** represents the comparative expression pattern of *TaTLPs* under various tissue and developmental stages. Based on their expression trend, *TaTLPs* categorized into retention of function **(a-c)** and pseudo functionalization types **(d, e)**.

## Additional files

**Additional file 1:**

**Table S1.** List of identified *TLP* genes with their proposed nomenclature in *B. distachyon, O. sativa S, bicolor, T. aestivum* and *Z. mays*. The Red colour represents the *sTLPs* of each crop. (XLSX 16 kb)

**Additional file 2:**

**Figure S1.** Multiple sequence alignment of 222 TLPs of *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays*. Conserved cysteine residues are marked with asterisk. Red dot denotes REDDD motif. The FF hydrophobic motif is highlighted with green triangle. The thaumatin signature motif, conserved domain, amino acids forming the bottom of acidic cleft are marked with sky blue, red and brown lines, respectively. (JPG 15,782 kb)

**Additional file 3:**

**Table S2.** List of paralogous *TLP* genes in *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays*. (XLXS 12 kb)

**Additional file 4:**

**Figure S2.** Intron/exon organization and conserved motifs of TLPs of *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays.* (JPG 883 kb)

**Additional file 5**:

**Table S3**. Domain organization of TLP proteins of *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays*. (XLXS 22 kb)

**Additional file 6**:

**Table S4**. The characterization table of *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays* TLPs. (XLXS 33 kb)

**Additional file 7:**

**Table S5.** List of predicted *cis*-regulatory elements in the promoter region of *TLP* genes of *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z, mays*. (DOCX 61 kb)

**Additional file 8:**

**Table S6.** Annotation and abbreviation of co-expressed genes of *B. distachyon, O. sativa, S. bicolor, T. aestivum* and *Z. mays* under tissue and developmental stages. (XLXS 43 kb)

**Additional file 9:**

**Table S7.** Annotation and abbreviation of co-expressed genes of *T. aestivum* under abiotic and biotic stress conditions. (XLXS 29 kb)

**Table 1** The Ka/Ks ratio and divergence time of duplicated *TLP* gene pairs.

| Gene A | Gene B | Ka | Ks | Ka/Ks | Selection pressure | T = Ks/2r |
|---|---|---|---|---|---|---|
| *BdTLPK1* | *BdTLPK2* | 0.1102 | 1.0810 | 0.1020 | Purifying | 83.2 |
| *OsTLP5* | *OsTLP7* | 0.1014 | 0.6788 | 0.1494 | Purifying | 52.2 |
| *OsTLP22* | *OsTLP24* | 0.0697 | 0.4134 | 0.3190 | Purifying | 27.7 |
| *OsTLP23* | *OsTLP25* | 0.1149 | 0.3603 | 0.31646 | Purifying | 29.4 |
| *OsTLP26* | *OsTLP27* | 0.1569 | 1.3980 | 0.1122 | Purifying | 107.6 |
| *SbTLP28* | *SbTLP29* | 0.0911 | 0.6808 | 0.1338 | Purifying | 52.4 |
| *SbTLP31* | *SbTLP34* | 0.1388 | 1.7816 | 0.0779 | Purifying | 137 |
| *SbTLP32* | *SbTLP33* | 0.0055 | 0.0716 | 0.0771 | Purifying | 5.5 |
| *TaTLP1-A* | *TaTLP24-D* | 0.1224 | 0.8739 | 0.1401 | Purifying | 67.2 |
| *TaTLP1-B* | *TaTLP24-B* | 0.0937 | 0.8163 | 0.1148 | Purifying | 62.8 |
| *TaTLP12-B* | *TaTLP13-D* | 0.1193 | 1.1836 | 0.1008 | Purifying | 91 |
| *TaTLP15-A* | *TaTLP18-B* | 0.1368 | 0.9879 | 0.1384 | Purifying | 76 |
| *TaTLP15-B2* | *TaTLP18-A* | 0.1423 | 0.9582 | 0.1486 | Purifying | 73.7 |
| *TaTLP16-A* | *TaTLP17-A* | 0.1456 | 1.2825 | 0.1135 | Purifying | 98.7 |
| *ZmTLP1* | *ZmTLP35* | 0.0955 | 0.3929 | 0.2431 | Purifying | 30.2 |
| *ZmTLP2* | *ZmTLP34* | 0.1044 | 0.8149 | 0.1281 | Purifying | 62.7 |
| *ZmTLP3* | *ZmTLP22* | 0.0495 | 0.3036 | 0.1630 | Purifying | 23.4 |
| *ZmTLP9* | *ZmTLP12* | 0.0767 | 0.3783 | 0.2027 | Purifying | 29.1 |
| *ZmTLP13* | *ZmTLP37* | 0.0523 | 0.3200 | 0.1634 | Purifying | 24.7 |
| *ZmTLP16* | *ZmTLP33* | 0.0133 | 0.0141 | 0.9483 | Purifying | 1.1 |
| *ZmTLP20* | *ZmTLP25* | 0.0797 | 0.0889 | 0.8965 | Purifying | 6.9 |

**Ka**- non-synonymous substitutions per non-synonymous site, **Ks**- synonymous substitutions per synonymous site, **T**- Divergence time

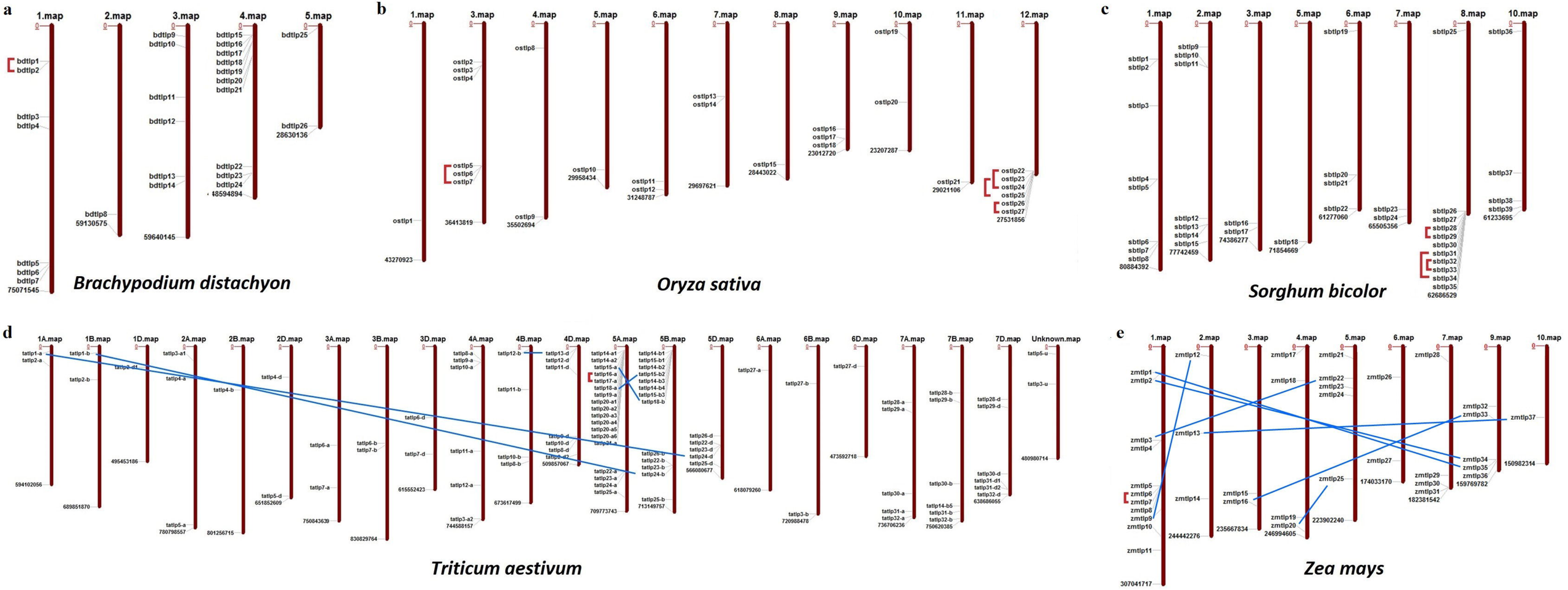
**Table 2** Tajima's relative rate test of paralogous genes.

| Group A | Group B | Out group | Nt | Na | Nb | $\chi^2$ | P |
|---|---|---|---|---|---|---|---|
| *BdTLP1* | *BdTLP2* | *BdTLP11* | 403 | 47 | 42 | 0.28 | 0.59611 |
| *OsTLP5* | *OsTLP7* | *OsTLP18* | 401 | 28 | 29 | 0.02 | 0.89463 |
| *OsTLP22* | *OsTLP24* | *OsTLP18* | 333 | 15 | 17 | 0.13 | 0.72367 |
| *OsTLP23* | *OsTLP25* | *OsTLP18* | 313 | 20 | 27 | 1.04 | 0.30723 |
| *OsTLP26* | *OsTLP27* | *OsTLP18* | 411 | 34 | 34 | 0 | 1 |
| *SbTLP28* | *SbTLP29* | *SbTLP5* | 316 | 19 | 30 | 2.47 | 0.11608 |
| *SbTLP31* | *SbTLP34* | *SbTLP5* | 312 | 27 | 32 | 0.42 | 0.51508 |
| *SbTLP32* | *SbTLP33* | *SbTLP5* | 347 | 2 | 2 | 0 | 1 |
| *TaTLP1-A* | *TaTLP24-D* | *TaTLP4-D* | 483 | 37 | 30 | 0.73 | 0.39245 |
| *TaTLP1-B* | *TaTLP24-B* | *TaTLP4-D* | 486 | 36 | 31 | 0.37 | 0.5413 |
| *TaTLP12-B* | *TaTLP13-D* | *TaTLP4-D* | 321 | 24 | 33 | 1.42 | 0.23323 |
| *TaTLP15-A* | *TaTLP18-B* | *TaTLP4-D* | 391 | 36 | 38 | 0.05 | 0.81615 |
| *TaTLP15-B2* | *TaTLP18-A* | *TaTLP4-D* | 260 | 28 | 27 | 0.02 | 0.89274 |
| *TaTLP16-A* | *TaTLP17-A* | *TaTLP4-D* | 291 | 36 | 31 | 0.37 | 0.5413 |
| *ZmTLP1* | *ZmTLP35* | *ZmTLP28* | 551 | 36 | 33 | 0.13 | 0.71798 |
| *ZmTLP2* | *ZmTLP34* | *ZmTLP28* | 541 | 34 | 25 | 1.37 | 0.24132 |
| *ZmTLP3* | *ZmTLP22* | *ZmTLP28* | 560 | 11 | 14 | 0.36 | 0.54851 |
| *ZmTLP9* | *ZmTLP12* | *ZmTLP28* | 372 | 6 | 12 | 2 | 0.1573 |
| *ZmTLP13* | *ZmTLP37* | *ZmTLP28* | 464 | 19 | 22 | 0.22 | 0.63941 |
| *ZmTLP16* | *ZmTLP33* | *ZmTLP28* | 551 | 5 | 7 | 0.33 | 0.5637 |
| *ZmTLP20* | *ZmTLP25* | *ZmTLP28* | 525 | 17 | 23 | 0.9 | 0.34278 |

**Nt-** Identical sites in all three sequences**, Na**-Unique differences in Sequence A, **Nb-** Unique differences in Sequence B

**Table 3** Physicochemical properties of long and small TLP proteins.

| Properties | Type | *Brachypodium distachyon* | *Oryza sativa* | *Sorghum bicolor* | *Triticum aestivum* | *Zea mays* |
|---|---|---|---|---|---|---|
| **Average Protein length** | Long TLPs | 299 | 289 | 294 | 284 | 334 |
| | Small TLPs | 185 | 178 | 178 | 177 | 175 |
| **Average Molecular weight (kDa)** | Long TLPs | 31 | 30 | 30 | 29 | 35 |
| | Small TLPs | 19.5 | 18.1 | 18.3 | 1822 | 17.7 |
| **Average pI** | Long TLPs | 6.18 | 6.95 | 6.05 | 6.18 | 5.88 |
| | Small TLPs | 6.82 | 6.06 | 5.74 | 6.17 | 4.9 |


a
1.map
2.map
3.map
4.map
5.map
bdtip1
bdtip2
bdtip3
bdtip4
bdtip5
bdtip6
bdtip7
75071545
bdtip8
59130575
bdtip9
bdtip10
bdtip11
bdtip12
bdtip13
bdtip14
59640145
bdtip15
bdtip16
bdtip17
bdtip18
bdtip19
bdtip20
bdtip21
bdtip22
bdtip23
bdtip24
48594894
bdtip25
bdtip26
28630136
Brachypodium distachyon
b
1.map
3.map
4.map
5.map
6.map
7.map
8.map
9.map
10.map
11.map
12.map
ostip1
43270923
ostip2
ostip3
ostip4
ostip5
ostip6
ostip7
36413819
ostip8
ostip9
35502694
ostip10
29958434
ostip11
ostip12
31248787
ostip13
ostip14
29697621
ostip15
28443022
ostip16
ostip17
ostip18
23012720
ostip19
ostip20
23207287
ostip21
29021106
ostip22
ostip23
ostip24
ostip25
ostip26
ostip27
27531856
Oryza sativa
c
1.map
2.map
3.map
5.map
6.map
7.map
8.map
10.map
sbtip1
sbtip2
sbtip3
sbtip4
sbtip5
sbtip6
sbtip7
sbtip8
80884392
sbtip9
sbtip10
sbtip11
sbtip12
sbtip13
sbtip14
sbtip15
77742459
sbtip16
sbtip17
74386277
sbtip18
71854669
sbtip19
sbtip20
sbtip21
sbtip22
61277060
sbtip23
sbtip24
65505356
sbtip25
sbtip26
sbtip27
sbtip28
sbtip29
sbtip30
sbtip31
sbtip32
sbtip33
sbtip34
sbtip35
62686529
sbtip36
sbtip37
sbtip38
sbtip39
61233695
Sorghum bicolor
d
1A.map
1B.map
1D.map
2A.map
2B.map
2D.map
3A.map
3B.map
3D.map
4A.map
4B.map
4D.map
5A.map
5B.map
5D.map
6A.map
6B.map
6D.map
7A.map
7B.map
7D.map
Unknown.map
tatip1-a
tatip2-a
594102056
tatip1-b
tatip2-b
689851870
tatip2-d1
495453186
tatip3-a1
tatip4-a
tatip5-a
780798557
tatip4-b
801256715
tatip4-d
tatip5-d
651852609
tatip6-a
tatip7-a
750843639
tatip6-b
tatip7-b
830829764
tatip6-d
tatip7-d
615552423
tatip8-a
tatip9-a
tatip10-a
tatip11-a
tatip12-a
tatip3-a2
744588157
tatip11-b
tatip10-b
tatip8-b
673617499
tatip12-b
tatip13-d
tatip12-d
tatip11-d
tatip9-d
tatip10-d
tatip8-d
tatip2-d2
509857067
tatip14-a1
tatip14-a2
tatip15-a
tatip16-a
tatip17-a
tatip18-a
tatip19-a
tatip20-a1
tatip20-a2
tatip20-a3
tatip20-a4
tatip20-a5
tatip20-a6
tatip24-a
tatip22-a
tatip23-a
tatip24-a
tatip25-a
709773743
tatip14-b1
tatip15-b1
tatip14-b2
tatip15-b2
tatip14-b3
tatip14-b4
tatip15-b3
tatip18-b
tatip26-b
tatip22-b
tatip23-b
tatip24-b
tatip25-b
713149757
tatip26-d
tatip22-d
tatip23-d
tatip24-d
tatip25-d
566080677
tatip27-a
618079260
tatip27-b
tatip3-b
720988478
tatip27-d
473592718
tatip28-a
tatip29-a
tatip30-a
tatip31-a
tatip32-a
736706236
tatip28-b
tatip29-b
tatip30-b
tatip14-b5
tatip31-b
tatip32-b
750620385
tatip28-d
tatip29-d
tatip30-d
tatip31-d1
tatip31-d2
tatip32-d
638686055
tatip5-u
tatip3-u
480980714
Triticum aestivum
e
1.map
2.map
3.map
4.map
5.map
6.map
7.map
9.map
10.map
zmtip1
zmtip2
zmtip3
zmtip4
zmtip5
zmtip6
zmtip7
zmtip8
zmtip9
zmtip10
zmtip11
307041717
zmtip12
zmtip13
zmtip14
244442276
zmtip15
zmtip16
235667834
zmtip17
zmtip18
zmtip19
zmtip20
246994605
zmtip21
zmtip22
zmtip23
zmtip24
zmtip25
223902240
zmtip26
zmtip27
174033170
zmtip28
zmtip29
zmtip30
zmtip31
182381542
zmtip32
zmtip33
zmtip34
zmtip35
zmtip36
159769782
zmtip37
150982314
Zea mays

```
THM1_THADA  NRCSYTVWAAASKGDAA--LDAGGRQLNSGESWTINVEPGTN-GCKIWARTDCYF----DDSGSGI--CKTGDCG-GLLRCNRFG-RPFTTLAEFSLN---QY-GKDYIDISNIKGFN
   OsTLP5   NRCQYTVWPAAVPSG-------GGTKLDPGQTWTINVPAGTT-GCRVWARTGCGF----DGSGNGQ--CQTGDCG-GKLRCTAYG-AAFNTLAEFALN---QWNNLDFFDISLIDGFN
   SbTLP2   NKCQYTVWAAAVP-G-------GGQQLDNGQTWSINVPAGTT-GCRVWARTGCSF----DGSGNGQ--CQTGDCG-GYLRCTQYG-QPFNTLAEFGLN---QYQGLDFIDISLVDGFN
  BdTLP13   NHCSHTIWPGTMAGAGTPQLPTTGFRLDPGQT--VQIPAPAGWSCRIWARTGCNFSADGAGDPAGAYACQTGDCGGGHMECCGTGGKPFATLFEYTLGKGTTADDQDYYDVSLVDGYN
  ZmTLP13   NACKHAVWVGALHGASSPPLARSGFYLAPSATSSMDAPSSGTWSCTLWARTGC-----VADPATGRFSCATADCGTGDVACCGRGPAPFVSLVEYTLA-APGSGRPDFYDVSLVDGFN
 TaTLP1-A   NECSHPVWVGVLSGATSPQLPRTGFYLAPGGTSSLAAPSSGAWSCNFWARTGC-----AVDPGTGRLACATADCGSGDVACCGRGPAPFVTLAEFTLA-APGSGGLDFYDVSLVDGFN
  BdTLP10   NQCGFPVWPAAIPV--G-----GGRQLNRGETWNLEVPAGTS-SARIWGRTGCSF---N-GDGRGS--CATGDCG-GALSCSLSG-QPFATLAEFTLG-----GAQDFYDISVIDGFN
  OsTLP25   NRCSFTVWPAATPV--G-----GGTQLNPGQTWTINVPAGTS-SCRVWGRTGCSF---D-GDGRGR--CATGDCG-GALSCKLSG-QPFLTLAEFTLGTS--GGNRDFYNLSVIDGYN
   ZmTLP6   NNCGSTVWPAATPV--G-----GGTQLDPGQTWTVDVPAGTQ-SCRVWGRTGCSF---N-G-GSGS--CQTGDCG-GALSCTLSG-QPFMTLAEFTIG-G--GGSQDFYDISVIDGYN
  SbTLP29   NNCSSTVWPAAIPV--G-----GGTQLDPGQTWTVDVPAGT--TCRFWGRTNCSF---V-G-GNGH--CGTGDCA-GALNCTVSG-QPFTTLAEFSLGGGSAGGNQDFYDMSVIEGYN
 TaTLP6-B   NRCPFPVCAAAMPVSGG-----GGAQLAPGETWTLTFAAGGT-SCRIWPRMGCAF---HPATGRGG--CQTGDCG-GALRCVLSG-KPFATLAEFNLAN---GGAPGRYGISVVDGFT
```

```
THM1_THADA  VPMDFSP--TTRG--CR--GVRCAADIVGQCPAKLK---APGG----GCNDACTVFQTSEYCCTT--G---KCGPTEYSRFFKRLCPDFFSYVLDKPTTV-TCPGSSNYRVTFCP
   OsTLP5   VPMAFLP--AGSGAGCPKGGPRCATAITPQCPSELR---APGG-----CNNACTVFRQDRYCCTG--SAANSCGPTNYSEFFKRLCPDFYSYPKDDASSTYTCPAGTNYQVVFCP
   SbTLP2   VPMDFLP--AGDGSGCPKGGPRCDADVTSQCPAALK---ATGG-----CNNPCTVFKTDEYCCTG--SAANTCGPTDYSKFFKGLCPDFYSYPKDDATSTYTCPGGTNYNVVFCP
  BdTLP13   LPVVAVPRAQQGGGSCN--ATGCAADLNLSCPKELQVEDGIGGSGTVACQSACEAFGKDKYCCSGAYATPTTCSPTAYSSVFKSACPRFYSYAYDDGSSLFTCNA-IDYTIAFCP
   ZmTLP13  VPVRVAP--SGGGGDCR--PAACAGDVNAVCPGDLRVVASSGG-GVVACKSACIAYGSARYCCTGQYGTPAMCGPTNYSQVFKSACPSFYSYAYDDATSTFTCSGPSSYDVTFC
 TaTLP1-A   LPVSIAP--AHGG--CQ--AAACAGDVNAVCPSDLSVASGSSG-EVVACKSACNAYGSARYCCSGDYGTPAACGPTSYSQVFKAACPAFYSYAYDDASSTFTCSAAATYHITFC
  BdTLP10   VAMDF----------------SCST---G---DALR------------CR-----------------------DAG---------CPDFYHHP-NDVKTH-ACSGNRSFRVVFCP
  OsTLP25   VAMSF----------------SCSS---G---VTLT------------CR-----------------------ERS---------CPDFYQYPSDDSKLR-SCNGNSNYRVVFCP
   ZmTLP6   LAMLF----------------SCST---G---VTLN------------CG-----------------------GPS---------CPDFYLFPSDNTKTH-ACSGNSNYQVTFCP
  SbTLP29   LPMAF----------------SCST---G---VGLV------------CT-----------------------SPT---------CPAFCLFASDDCRVH-ACSSNSNYQVIFCP
 TaTLP6-B   LPMDF----------------SCSGAGDG---GVIR------------CR-----------------------DPG---------CPDCNHRP-GDGKYR-ACPAYSDYKVVFCP
```

Cysteine residue

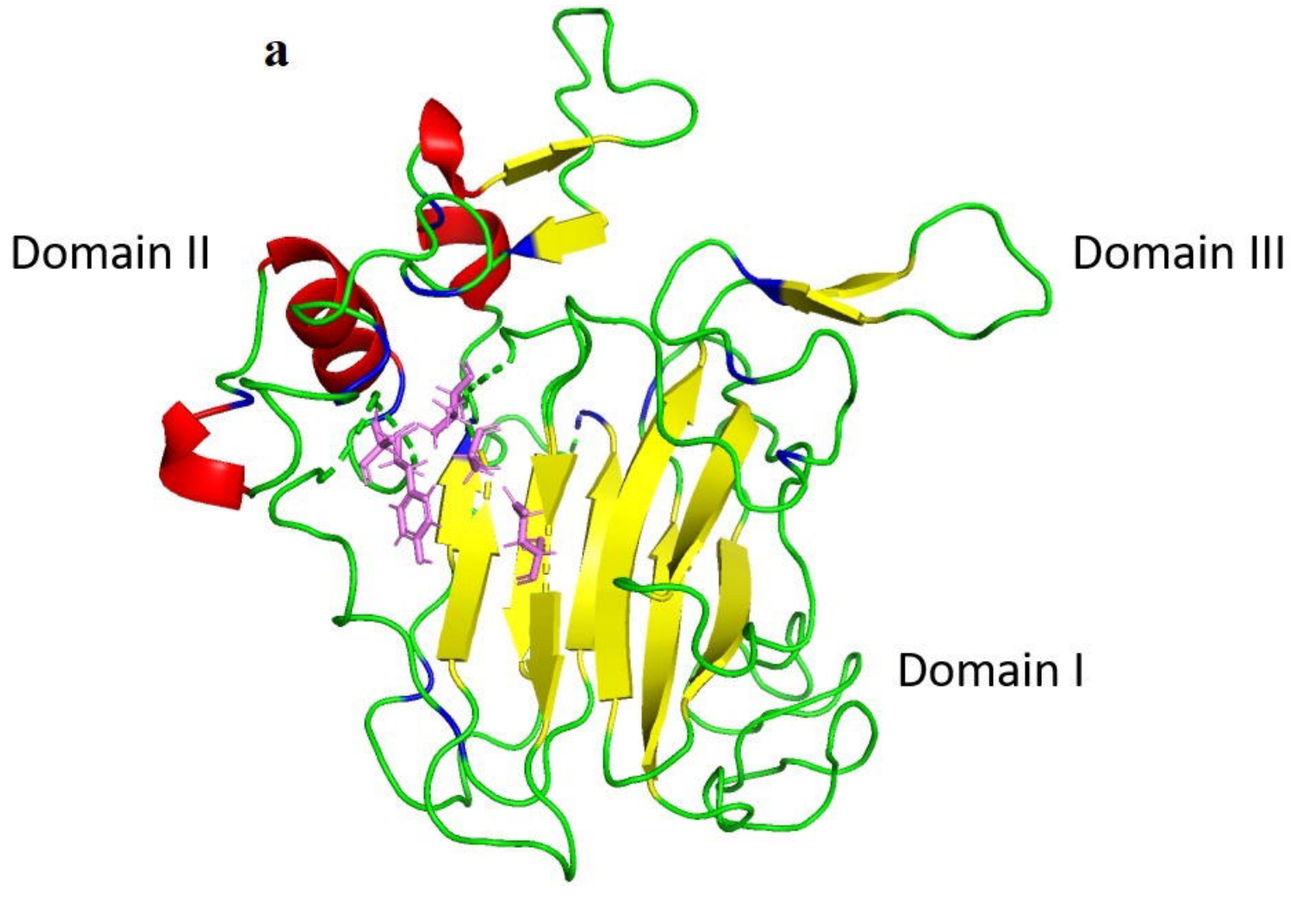
REDDD motif

FF hydrophobic motif

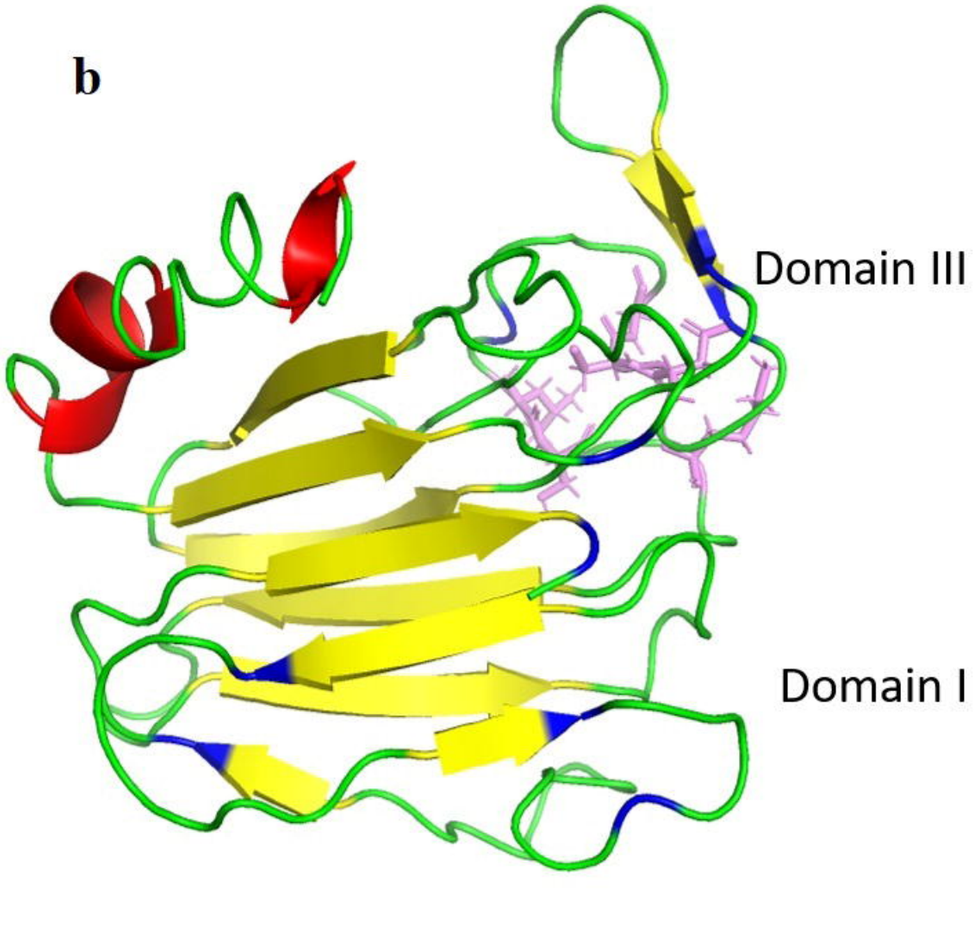
Bottom of acidic cleft forming amino acids

Signature motif

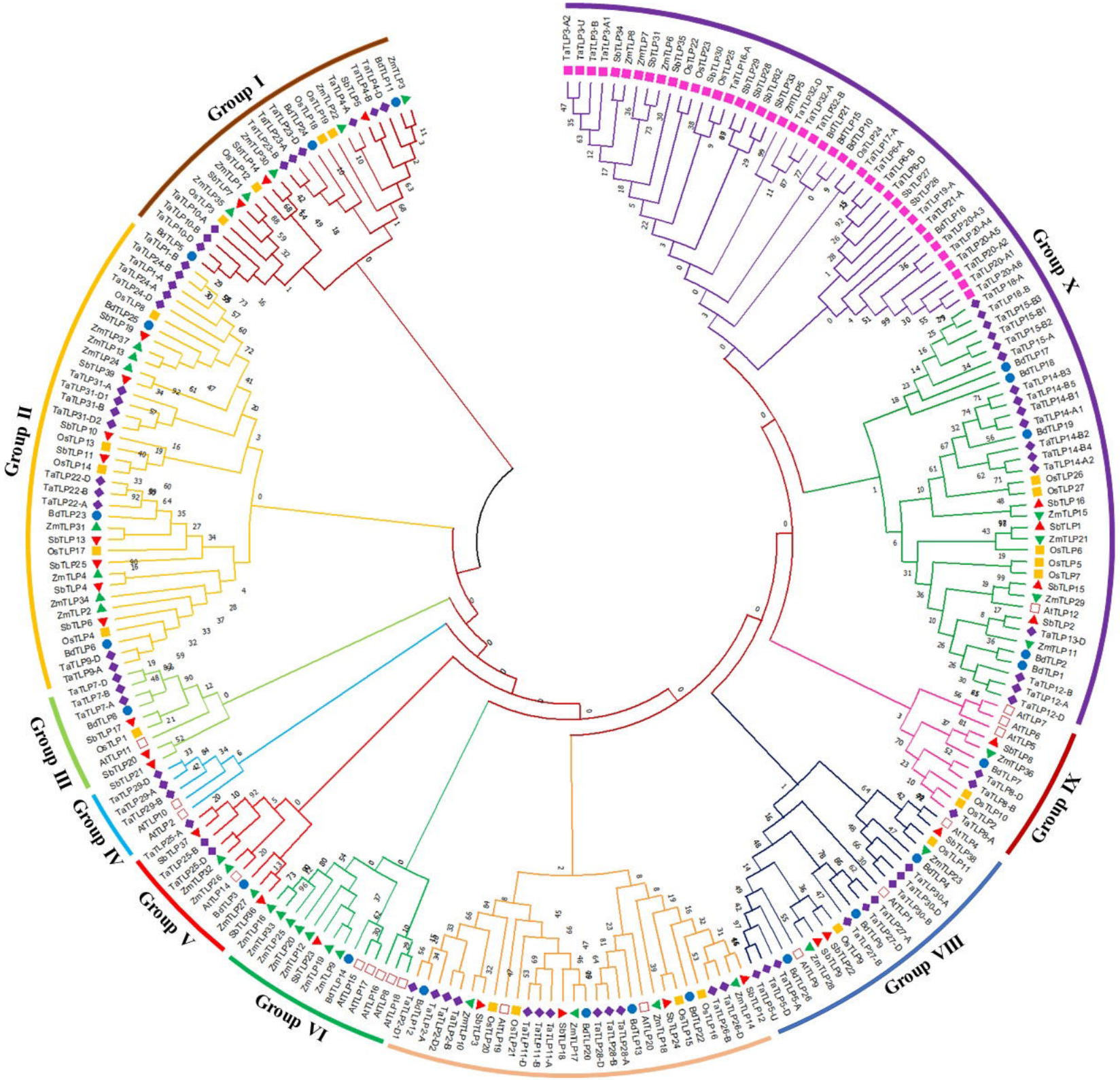
Thaumatin conserverd domain

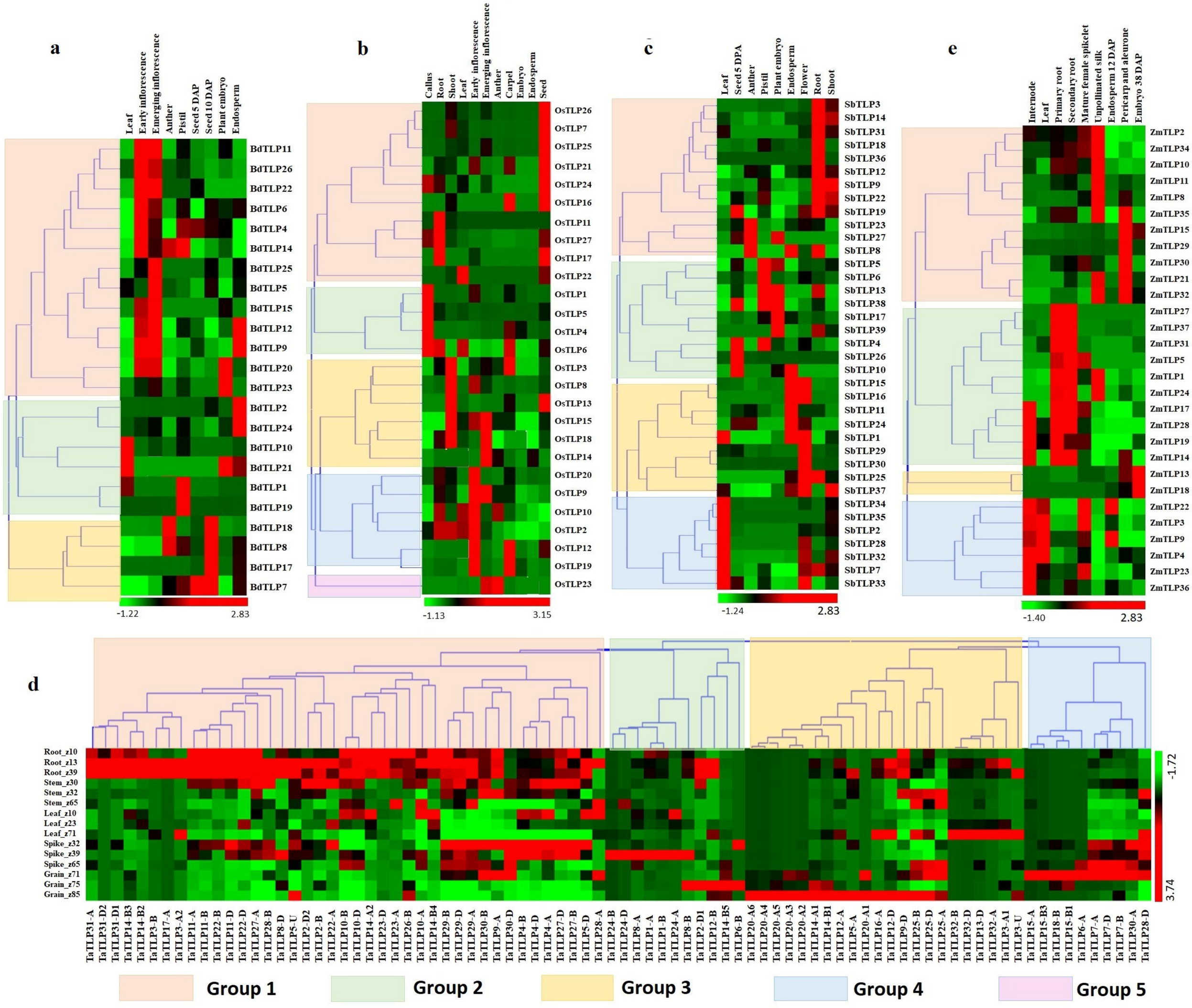


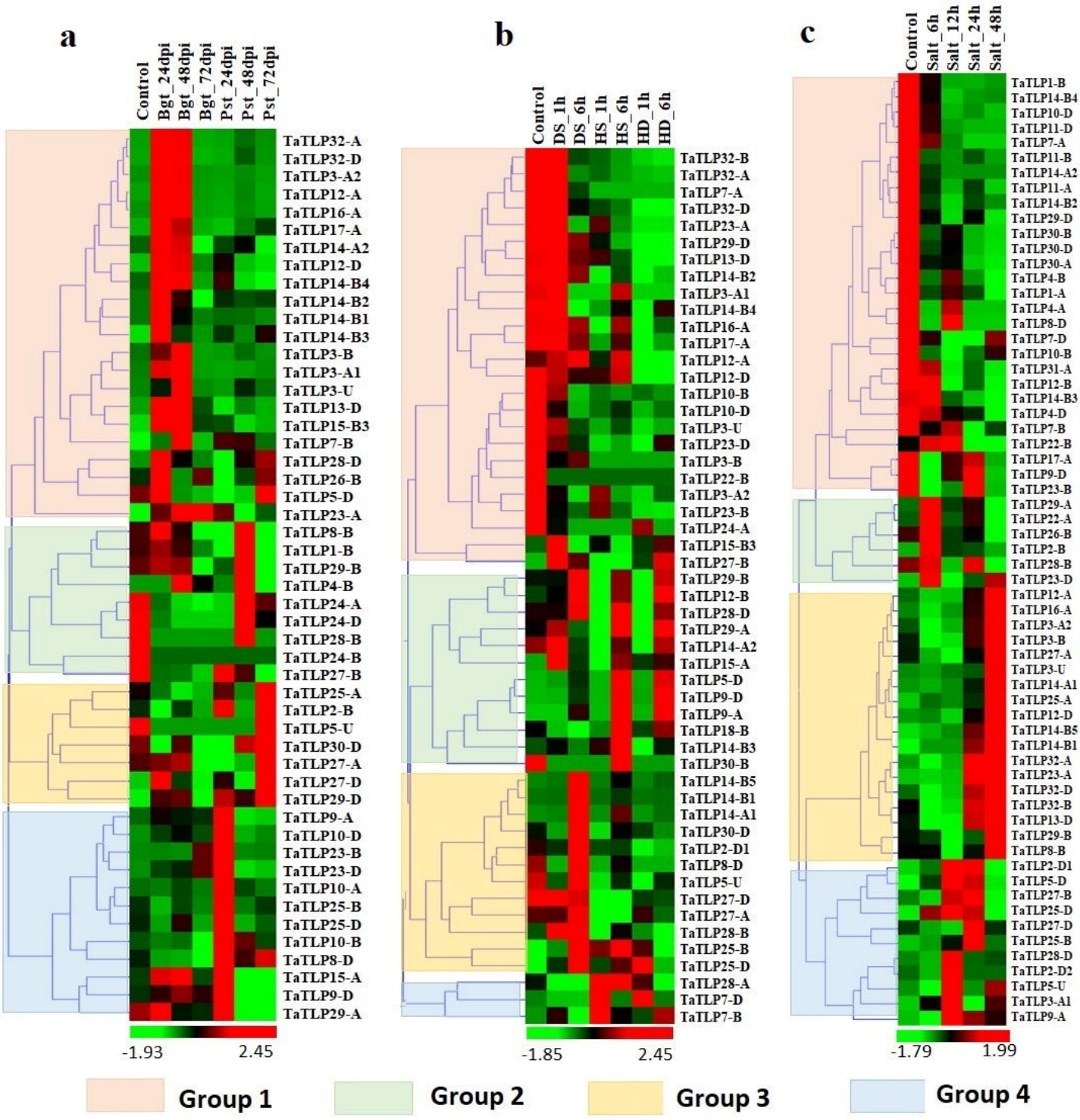

TaTLP2-A

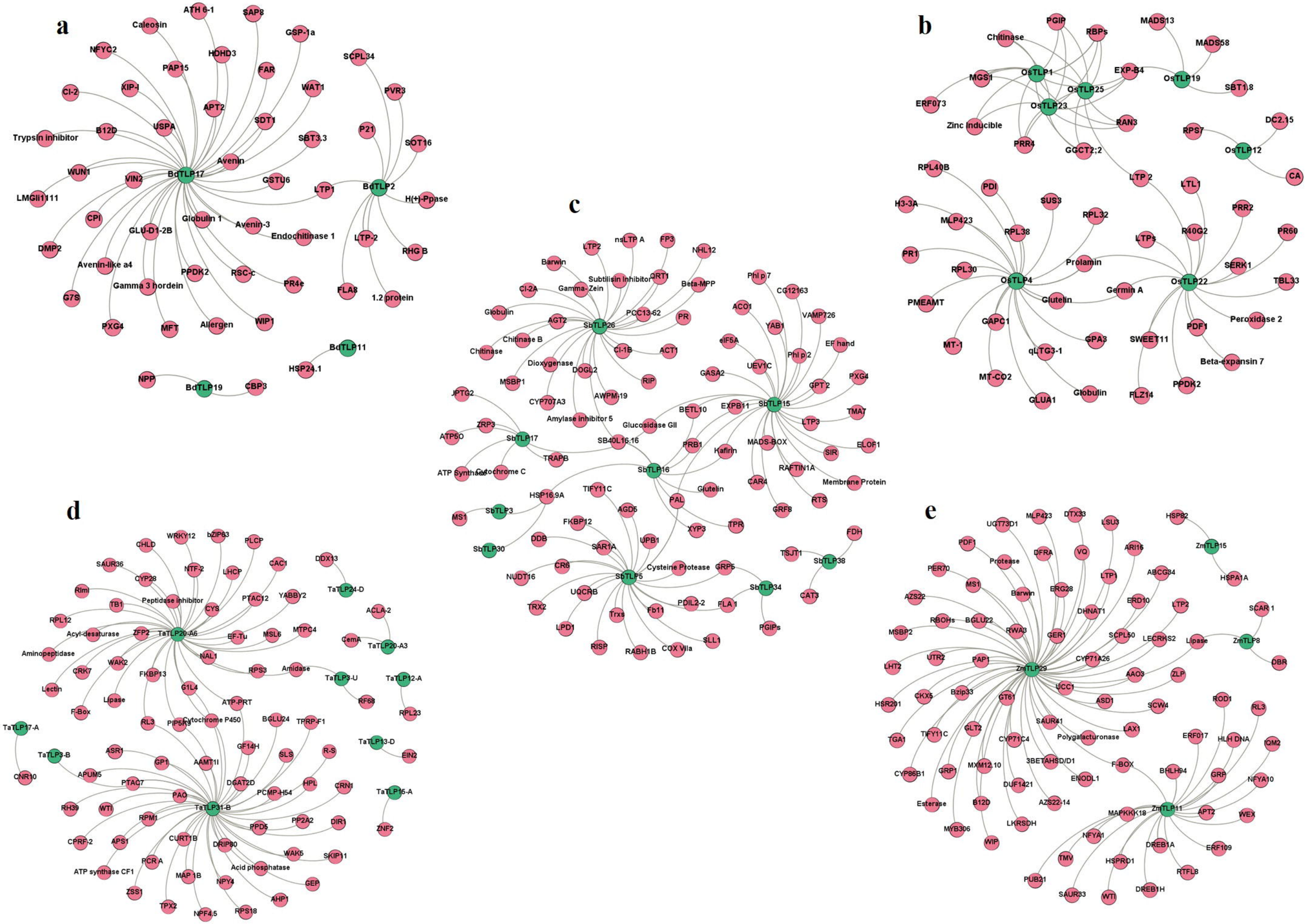


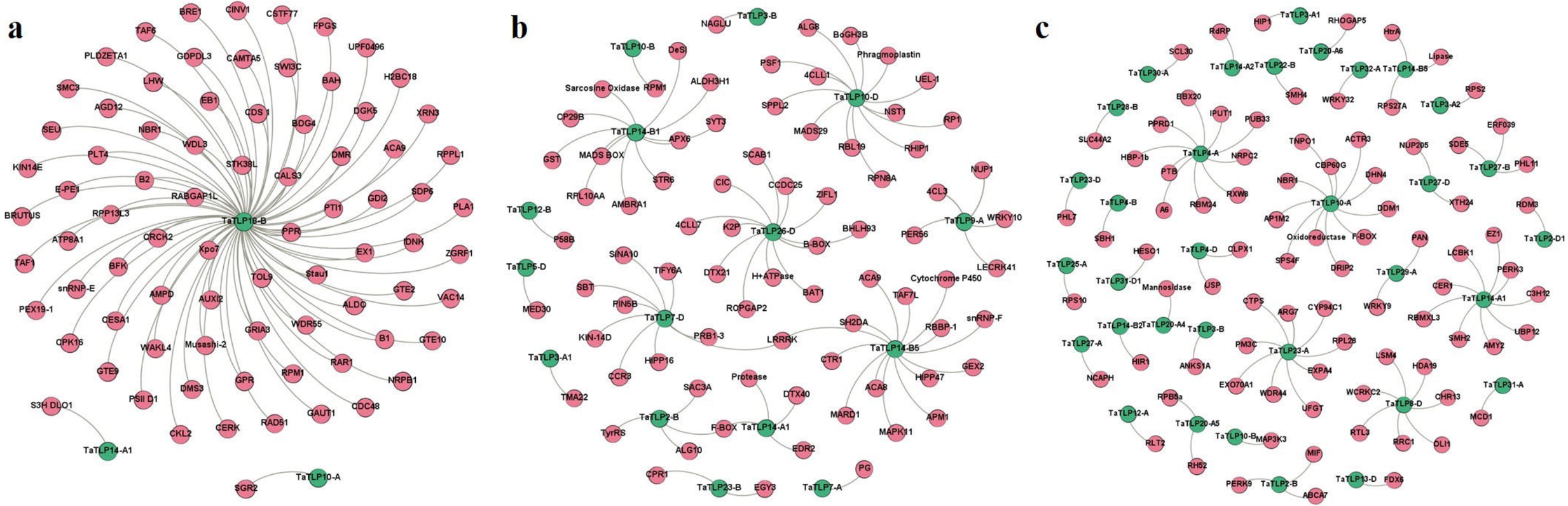

TaTLP6-A

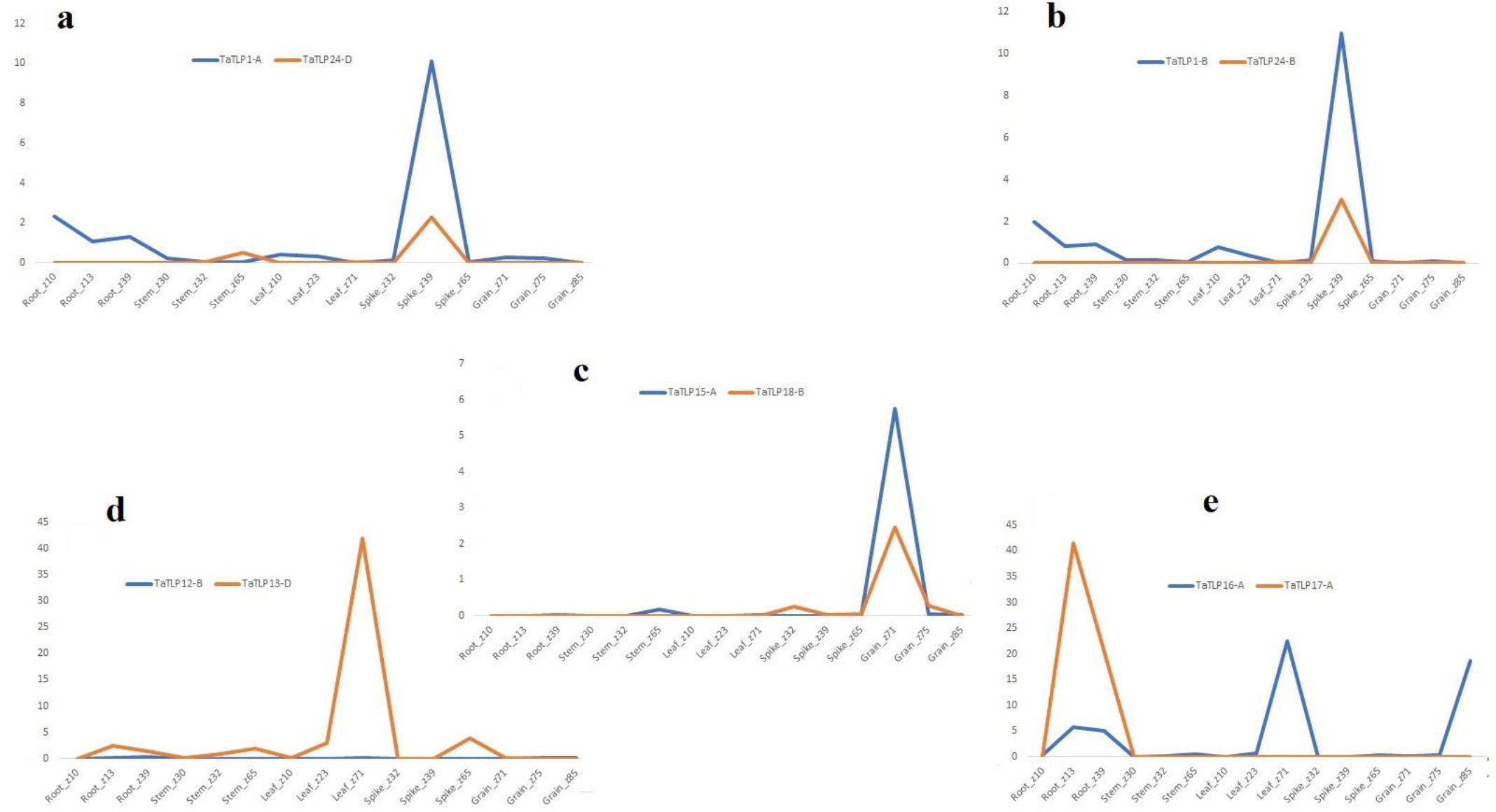

Group I
Group II
Group III
Group IV
Group V
Group VI
Group VII
Group VIII
Group IX
Group X


a
Leaf
Early inflorescence
Emerging inflorescence
Anther
Pistil
Seed 5 DAP
Seed 10 DAP
Plant embryo
Endosperm
BdTLP11
BdTLP26
BdTLP22
BdTLP6
BdTLP4
BdTLP14
BdTLP25
BdTLP5
BdTLP15
BdTLP12
BdTLP9
BdTLP20
BdTLP23
BdTLP2
BdTLP24
BdTLP10
BdTLP21
BdTLP1
BdTLP19
BdTLP18
BdTLP8
BdTLP17
BdTLP7
-1.22
2.83
b
Callus
Root
Shoot
Leaf
Early inflorescence
Emerging inflorescence
Anther
Carpel
Embryo
Endosperm
Seed
OsTLP26
OsTLP7
OsTLP25
OsTLP21
OsTLP24
OsTLP16
OsTLP11
OsTLP27
OsTLP17
OsTLP22
OsTLP1
OsTLP5
OsTLP4
OsTLP6
OsTLP3
OsTLP8
OsTLP13
OsTLP15
OsTLP18
OsTLP14
OsTLP20
OsTLP9
OsTLP10
OsTLP2
OsTLP12
OsTLP19
OsTLP23
-1.13
3.15
c
Leaf
Seed 5 DPA
Anther
Pistil
Plant embryo
Endosperm
Flower
Root
Shoot
SbTLP3
SbTLP14
SbTLP31
SbTLP18
SbTLP36
SbTLP12
SbTLP9
SbTLP22
SbTLP19
SbTLP23
SbTLP27
SbTLP8
SbTLP5
SbTLP6
SbTLP13
SbTLP38
SbTLP17
SbTLP39
SbTLP4
SbTLP26
SbTLP10
SbTLP15
SbTLP16
SbTLP11
SbTLP24
SbTLP1
SbTLP29
SbTLP30
SbTLP25
SbTLP37
SbTLP34
SbTLP35
SbTLP2
SbTLP28
SbTLP32
SbTLP7
SbTLP33
-1.24
2.83
e
Internode
Leaf
Primary root
Secondary root
Mature female spikelet
Unpollinated silk
Endosperm 12 DAP
Pericarp and aleurone
Embryo 38 DAP
ZmTLP2
ZmTLP34
ZmTLP10
ZmTLP11
ZmTLP8
ZmTLP35
ZmTLP15
ZmTLP29
ZmTLP30
ZmTLP21
ZmTLP32
ZmTLP27
ZmTLP37
ZmTLP31
ZmTLP5
ZmTLP1
ZmTLP24
ZmTLP17
ZmTLP28
ZmTLP19
ZmTLP14
ZmTLP13
ZmTLP18
ZmTLP22
ZmTLP3
ZmTLP9
ZmTLP4
ZmTLP23
ZmTLP36
-1.40
2.83
d
Root_z10
Root_z13
Root_z39
Stem_z30
Stem_z32
Stem_z65
Leaf_z10
Leaf_z23
Leaf_z71
Spike_z32
Spike_z39
Spike_z65
Grain_z71
Grain_z75
Grain_z85
-1.72
3.74
TaTLP31-A
TaTLP31-D2
TaTLP31-D1
TaTLP14-B3
TaTLP14-B2
TaTLP3-B
TaTLP17-A
TaTLP3-A2
TaTLP11-A
TaTLP11-B
TaTLP22-B
TaTLP11-D
TaTLP22-D
TaTLP27-A
TaTLP28-B
TaTLP8-D
TaTLP5-U
TaTLP2-D2
TaTLP2-B
TaTLP22-A
TaTLP10-B
TaTLP10-D
TaTLP14-A2
TaTLP23-D
TaTLP23-A
TaTLP26-B
TaTLP10-A
TaTLP14-B4
TaTLP29-B
TaTLP29-D
TaTLP29-A
TaTLP30-B
TaTLP9-A
TaTLP30-D
TaTLP4-B
TaTLP4-D
TaTLP4-A
TaTLP27-D
TaTLP27-B
TaTLP5-D
TaTLP28-A
TaTLP24-B
TaTLP24-D
TaTLP8-A
TaTLP1-A
TaTLP1-B
TaTLP24-A
TaTLP8-B
TaTLP2-D1
TaTLP12-B
TaTLP14-B5
TaTLP6-B
TaTLP20-A6
TaTLP20-A4
TaTLP20-A5
TaTLP20-A3
TaTLP20-A2
TaTLP14-A1
TaTLP14-B1
TaTLP12-A
TaTLP5-A
TaTLP20-A1
TaTLP16-A
TaTLP12-D
TaTLP9-D
TaTLP25-B
TaTLP25-D
TaTLP25-A
TaTLP32-B
TaTLP32-D
TaTLP13-D
TaTLP32-A
TaTLP3-A1
TaTLP3-U
TaTLP15-A
TaTLP15-B3
TaTLP18-B
TaTLP15-B1
TaTLP6-A
TaTLP7-A
TaTLP7-D
TaTLP7-B
TaTLP30-A
TaTLP28-D
Group 1
Group 2
Group 3
Group 4
Group 5


a
Control
Bgt_24dpi
Bgt_48dpi
Bgt_72dpi
Pst_24dpi
Pst_48dpi
Pst_72dpi
TaTLP32-A
TaTLP32-D
TaTLP3-A2
TaTLP12-A
TaTLP16-A
TaTLP17-A
TaTLP14-A2
TaTLP12-D
TaTLP14-B4
TaTLP14-B2
TaTLP14-B1
TaTLP14-B3
TaTLP3-B
TaTLP3-A1
TaTLP3-U
TaTLP13-D
TaTLP15-B3
TaTLP7-B
TaTLP28-D
TaTLP26-B
TaTLP5-D
TaTLP23-A
TaTLP8-B
TaTLP1-B
TaTLP29-B
TaTLP4-B
TaTLP24-A
TaTLP24-D
TaTLP28-B
TaTLP24-B
TaTLP27-B
TaTLP25-A
TaTLP2-B
TaTLP5-U
TaTLP30-D
TaTLP27-A
TaTLP27-D
TaTLP29-D
TaTLP9-A
TaTLP10-D
TaTLP23-B
TaTLP23-D
TaTLP10-A
TaTLP25-B
TaTLP25-D
TaTLP10-B
TaTLP8-D
TaTLP15-A
TaTLP9-D
TaTLP29-A
-1.93
2.45
b
Control
DS_1h
DS_6h
HS_1h
HS_6h
HD_1h
HD_6h
TaTLP32-B
TaTLP32-A
TaTLP7-A
TaTLP32-D
TaTLP23-A
TaTLP29-D
TaTLP13-D
TaTLP14-B2
TaTLP3-A1
TaTLP14-B4
TaTLP16-A
TaTLP17-A
TaTLP12-A
TaTLP12-D
TaTLP10-B
TaTLP10-D
TaTLP3-U
TaTLP23-D
TaTLP3-B
TaTLP22-B
TaTLP3-A2
TaTLP23-B
TaTLP24-A
TaTLP15-B3
TaTLP27-B
TaTLP29-B
TaTLP12-B
TaTLP28-D
TaTLP29-A
TaTLP14-A2
TaTLP15-A
TaTLP5-D
TaTLP9-D
TaTLP9-A
TaTLP18-B
TaTLP14-B3
TaTLP30-B
TaTLP14-B5
TaTLP14-B1
TaTLP14-A1
TaTLP30-D
TaTLP2-D1
TaTLP8-D
TaTLP5-U
TaTLP27-D
TaTLP27-A
TaTLP28-B
TaTLP25-B
TaTLP25-D
TaTLP28-A
TaTLP7-D
TaTLP7-B
-1.85
2.45
c
Control
Salt_6h
Salt_12h
Salt_24h
Salt_48h
TaTLP1-B
TaTLP14-B4
TaTLP10-D
TaTLP11-D
TaTLP7-A
TaTLP11-B
TaTLP14-A2
TaTLP11-A
TaTLP14-B2
TaTLP29-D
TaTLP30-B
TaTLP30-D
TaTLP30-A
TaTLP4-B
TaTLP1-A
TaTLP4-A
TaTLP8-D
TaTLP7-D
TaTLP10-B
TaTLP31-A
TaTLP12-B
TaTLP14-B3
TaTLP4-D
TaTLP7-B
TaTLP22-B
TaTLP17-A
TaTLP9-D
TaTLP23-B
TaTLP29-A
TaTLP22-A
TaTLP26-B
TaTLP2-B
TaTLP28-B
TaTLP23-D
TaTLP12-A
TaTLP16-A
TaTLP3-A2
TaTLP3-B
TaTLP27-A
TaTLP3-U
TaTLP14-A1
TaTLP25-A
TaTLP12-D
TaTLP14-B5
TaTLP14-B1
TaTLP32-A
TaTLP23-A
TaTLP32-D
TaTLP32-B
TaTLP13-D
TaTLP29-B
TaTLP8-B
TaTLP2-D1
TaTLP5-D
TaTLP27-B
TaTLP25-D
TaTLP27-D
TaTLP25-B
TaTLP28-D
TaTLP2-D2
TaTLP5-U
TaTLP3-A1
TaTLP9-A
-1.79
1.99
Group 1
Group 2
Group 3
Group 4


a
ATH 6-1
SAP8
Caleosin
GSP-1a
NFYC2
HDHD3
SCPL34
PAP15
FAR
XIP-I
WAT1
PVR3
CI-2
APT2
SDT1
USPA
P21
B12D
SBT3.3
SOT16
Trypsin inhibitor
Avenin
WUN1
VIN2
BdTLP17
GSTU6
LTP1
BdTLP2
H(+)-Ppase
LMGli1111
CPI
Globulin 1
Avenin-3
GLU-D1-2B
Endochitinase 1
LTP-2
RHG B
DMP2
Avenin-like a4
PPDK2
RSC-c
PR4e
Gamma 3 hordein
FLA8
1.2 protein
G7S
PXG4
MFT
Allergen
WIP1
BdTLP11
HSP24.1
NPP
BdTLP19
CBP3
b
PGIP
MADS13
Chitinase
RBPs
MADS58
MGS1
OsTLP1
EXP-B4
OsTLP19
OsTLP25
SBT1.8
ERF073
OsTLP23
RAN3
DC2.15
Zinc Inducible
RPS7
PRR4
GGCT2;2
OsTLP12
RPL40B
CA
LTP 2
LTL1
PDI
H3-3A
SUS3
PRR2
MLP423
RPL32
RPL38
R40G2
PR60
LTPs
PR1
Prolamin
SERK1
RPL30
OsTLP4
OsTLP22
TBL33
Germin A
PMEAMT
Glutelin
Peroxidase 2
GAPC1
PDF1
GPA3
SWEET11
MT-1
qLTG3-1
Beta-expansin 7
MT-CO2
PPDK2
Globulin
FLZ14
GLUA1
c
LTP2
nsLTP A
FP3
NHL12
Barwin
Subtilisin Inhibitor
QRT1
Beta-MPP
Phl p 7
CI-2A
Gamma-Zein
CG12163
Globulin
PCC13-62
ACO1
AGT2
SbTLP26
PR
VAMP726
YAB1
Chitinase B
eIF5A
EF hand
Chitinase
CI-1B
ACT1
Phl p 2
Dioxygenase
DOGL2
UEV1C
GASA2
PXG4
MSBP1
RIP
GPT 2
JPTG2
AWPM-19
CYP707A3
BETL10
EXPB11
SbTLP15
TMA7
Amylase inhibitor 5
Glucosidase GII
LTP3
ZRP3
ATP5O
SbTLP17
SB40L16.16
PRB1
MADS-BOX
ELOF1
Kafirin
SIR
TRAPB
ATP Synthase
Cytochrome C
SbTLP16
RAFTIN1A
CAR4
Membrane Protein
HSP16.9A
TIFY11C
Glutelin
RTS
PAL
MS1
SbTLP3
AGD5
GRF8
FKBP12
XYP3
TPR
FDH
DDB
UPB1
SbTLP30
SAR1A
TSJT1
SbTLP38
CR6
Cysteine Protease
GRP5
NUDT16
SbTLP5
SbTLP34
UQCRB
CAT3
TRX2
PDIL2-2
FLA 1
Trxs
Fb11
LPD1
PGIPs
SLL1
RISP
RABH1B
COX VIIa
d
WRKY12
bZIP63
PLCP
CHLD
SAUR36
NTF-2
LHCP
CAC1
DDX13
CYP28
TaTLP24-D
RImi
Peptidase inhibitor
PTAC12
YABBY2
TB1
CYS
ACLA-2
RPL12
Acyl-desaturase
ZFP2
TaTLP20-A6
EF-Tu
MSL6
MTPC4
CemA
TaTLP20-A3
Aminopeptidase
NAL1
WAK2
CRK7
FKBP13
RPS3
Amidase
TaTLP3-U
TaTLP12-A
Lectin
G1L4
Lipase
ATP-PRT
RF68
F-Box
RPL23
RL3
PIP5K9
Cytochrome P450
BGLU24
TPRP-F1
TaTLP17-A
TaTLP13-D
GF14H
ASR1
SLS
R-S
EIN2
TaTLP3-B
GP1
AAMT1I
CNR10
APUM5
DGAT2D
HPL
CRN1
PTAC7
PCMP-H54
TaTLP15-A
PAO
RH39
WTI
TaTLP31-B
RPM1
PP2A2
DIR1
PPD5
ZNF2
CPRF-2
APS1
CURT1B
DRIP80
WAK5
SKIP11
PCR A
Acid phosphatase
ATP synthase CF1
MAP 1B
NPY4
GEP
ZSS1
AHP1
TPX2
NPF4.5
RPS18
e
MLP423
DTX33
LSU3
HSP82
UGT73D1
PDF1
DFRA
VQ
ARI16
ZmTLP15
Protease
PER70
LTP1
ABCG34
HSPA1A
MS1
Barwin
ERG28
AZS22
ERD10
LTP2
SCAR 1
DHNAT1
RBOHs
BGLU22
MSBP2
RWA3
GER1
SCPL50
LECRKS2
Lipase
ZmTLP8
UTR2
PAP1
CYP71A26
DBR
LHT2
ZmTLP29
AAO3
ZLP
UCC1
CKX5
Bzip33
GT61
ASD1
SCW4
ROD1
HSR201
RL3
SAUR41
LAX1
TGA1
TIFY11C
GLT2
Polygalacturonase
ERF017
HLH DNA
CYP71C4
IQM2
3BETAHSD/D1
F-BOX
CYP86B1
GRP1
MXM12.10
BHLH94
GRP
ENODL1
NFYA10
DUF1421
Esterase
B12D
AZS22-14
MAPKKK18
ZmTLP11
APT2
WEX
LKRSDH
MYB306
NFYA1
WIP
DREB1A
ERF109
TMV
HSPRO1
RTFL8
PUB21
DREB1H
SAUR33
WTI


a
TaTLP18-B
BRE1
CINV1
CSTF77
FPGS
TAF6
UPF0496
PLDZETA1
GDPDL3
CAMTA5
SWI3C
LHW
BAH
H2BC18
SMC3
EB1
AGD12
CDS 1
DGK5
XRN3
SEU
NBR1
BDG4
WDL3
ACA9
PLT4
DMR
RPPL1
KIN14E
STK38L
CALS3
B2
E-PE1
RABGAP1L
GDI2
SDP6
BRUTUS
RPP13L3
PTI1
PLA1
ATP8A1
CRCK2
PPR
IDN4
TAF1
Xpo7
EX1
ZGRF1
BFK
Stau1
snRNP-E
TOL9
GTE2
VAC14
AMPD
AUXI2
ALDO
PEX19-1
CESA1
WDR55
GRIA3
B1
GTE10
CPK16
WAKL4
Musashi-2
RAR1
GTE9
RPM1
NRPB1
DMS3
GPR
PSII D1
S3H DUO1
CDC48
GAUT1
RAD51
CERK
CKL2
TaTLP14-A1
TaTLP10-A
SGR2
b
TaTLP3-B
NAGLU
ALG8
BoGH3B
TaTLP10-B
DeSI
Phragmoplastin
PSF1
4CLL1
UEL1
Sarcosine Oxidase
RPM1
ALDH3H1
TaTLP10-D
SPPL2
NST1
CP29B
SYT3
RP1
TaTLP14-B1
APX6
MADS29
GST
MADS BOX
RBL19
RHIP1
SCAB1
STR6
RPN8A
NUP1
CIC
CCDC25
RRL10AA
ZIFL1
4CL3
AMBRA1
TaTLP12-B
TaTLP9-A
WRKY10
4CLL7
K2P
TaTLP26-D
BHLH93
PER56
P58B
B-BOX
SINA10
LECRK41
TaTLP5-D
TIFY6A
DTX21
H+ATPase
ACA9
Cytochrome P450
SBT
BAT1
TAF7L
MED30
PIN5B
ROPGAP2
TaTLP7-D
SH2DA
snRNP-F
KIN-14D
PRB1-3
LRRRK
RBBP-1
TaTLP14-B5
TaTLP3-A1
CTR1
GEX2
HIPP16
CCR8
Protease
HIPP47
SAC3A
DTX40
ACA8
TMA22
MARD1
APM1
TaTLP2-B
F-BOX
TaTLP14-A1
TyrRS
MAPK11
ALG10
EDR2
PG
CPR1
TaTLP23-B
EGY3
TaTLP7-A
c
TaTLP3-A1
HIP1
RHOGAP5
RdRP
HtrA
SCL30
TaTLP20-A6
Lipase
TaTLP30-A
TaTLP14-A2
TaTLP22-B
TaTLP32-A
TaTLP14-B5
SMH4
RPS2
BBX20
WRKY32
RPS27A
TaTLP3-A2
TaTLP28-B
IPUT1
PUB33
PPRD1
ERF039
SLC44A2
TNPO1
ACTR3
NUP205
SDE5
HBP-1b
TaTLP4-A
NRPC2
CBP60G
TaTLP27-B
PHL11
PTB
DHN4
TaTLP23-D
NBR1
TaTLP27-D
RXW8
XTH24
TaTLP4-B
RBM24
TaTLP10-A
RDM3
A6
DDM1
PHL7
AP1M2
Oxidoreductase
F-BOX
PAN
EZ1
TaTLP2-D1
SBH1
HESO1
TaTLP4-D
CLPX1
LCBK1
SPS4F
TaTLP25-A
DRIP2
TaTLP29-A
PERK3
TaTLP31-D1
Mannosidase
USP
CER1
RPS10
CTPS
TaTLP14-A1
C3H12
ARG7
CYP94C1
WRKY9
RBMXL3
TaTLP14-B2
TaTLP20-A4
TaTLP3-B
UBP12
SMH2
TaTLP27-A
PM3C
RPL28
TaTLP23-A
AMY2
HIR1
ANKS1A
LSM4
HDA19
NCAPH
EXPA4
EXO70A1
TaTLP31-A
WDR44
WCRKC2
RPB5a
CHR13
UFGT
TaTLP6-D
TaTLP12-A
MCD1
TaTLP20-A5
RTL3
TaTLP10-B
MAP3K3
RRC1
OLI1
RLT2
MIF
RH52
TaTLP13-D
FDX6
PERK9
TaTLP2-B
ABCA7


a
TaTLP1-A
TaTLP24-D
b
TaTLP1-B
TaTLP24-B
c
TaTLP15-A
TaTLP18-B
d
TaTLP12-B
TaTLP13-D
e
TaTLP16-A
TaTLP17-A
Root_z10
Root_z13
Root_z39
Stem_z30
Stem_z32
Stem_z65
Leaf_z10
Leaf_z23
Leaf_z71
Spike_z32
Spike_z39
Spike_z65
Grain_z71
Grain_z75
Grain_z85