

# Genome-wide identification of thaumatin-like protein family genes in *Panax notoginseng* and analysis of their responses to *Fusarium solani* infection

yang yun (✉ [yang1605@user.ynnu.edu.cn](mailto:yang1605@user.ynnu.edu.cn))

Yunnan Normal University School of Energy and Environment Science <https://orcid.org/0000-0002-5028-1598>

Huilin Guan

Yunnan Normal University School of Energy and Environment Science

Fugang Wei

Wenshan Miaoxiang Notoginseng Technology Co., Ltd.

Zecheng Li

Kunming Yingwu Agricultural Science and Technology Co., Ltd.

Shaoshou Yang

Wenshan Miaoxiang Notoginseng Technology Co., Ltd.

Jingxin Huang

Yunnan Normal University School of Energy and Environment Science

---

## Research Article

**Keywords:** *Panax notoginseng*, thaumatin-like protein, *Fusarium solani*, expression patterns

**Posted Date:** August 17th, 2023

**DOI:** <https://doi.org/10.21203/rs.3.rs-3252650/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Genetic Resources and Crop Evolution on November 3rd, 2023. See the published version at <https://doi.org/10.1007/s10722-023-01736-z>.

# Abstract

*Panax notoginseng* is the most widely used Chinese medicinal herb, and one main challenge to the successful cultivation is root rot, which significantly reduces its yield and quality. Thaumatin-like proteins (TLPs) are a large and complex proteins of pathogenesis-related family 5 (PR-5), which play a important role in plant host defense and other physiological processes. The objective of this study is to identify the *TLP* gene family in *P. notoginseng* on a genome-wide scale and explore their response to *Fusarium solani* infection. In this study, a total of twenty *PnTLP* genes were identified in *P. notoginseng* genome and divided into ten subfamilies (Group I–X) by the phylogenetic analysis. Subcellular localization prediction showed that most of the *PnTLP* proteins were located in the extracellular, and few were distributed in chloroplast. Meanwhile, the gene structure and protein conserved domain were established by MEME Suite. Duplication analysis showed that segmental duplication, rather than tandem duplications promoting the expansion of the *PnTLPs* gene family. The analysis of collinear relationships reflected that there were 36 homologous genes between *P. notoginseng* and *Arabidopsis* and 9 between *P. notoginseng* and rice. Moreover, cis-acting elements of *PnTLP* gene promoters were predicted which indicated that *PnTLPs* are involved in biotic, abiotic stress, and hormone induction. Expression profile analysis revealed the possible roles of *TLPs* in plant development and fungal stress response. Furthermore, qRT-PCR analyses revealed that the *PnTLP* members were up-regulated or down-regulated in response to *F. solani* stress. The results of this study provide insight into the *PnTLP* gene family, and their potential biological functions in plants' defenses against pathogenic fungi, which would provide candidate genes for future studies of their roles in disease resistance.

## Introduction

Disease progression related proteins (PRs) are commonly used as molecular markers for plant disease resistance, enabling plants to achieve systemic resistance (Du, Amin, et al., 2023). PRs are a family of small molecular proteins induced by plants under various biotic and abiotic stresses. They play an important role in the defense mechanisms of plants (Kaur, Pati, et al., 2017). Thaumatin-like proteins (TLPs) are important members of the highly complex gene family pathogenesis-related protein group 5 (PR5), with molecular masses of 20–26 kDa (Linthorst and Van Loon, 1991). TLPs have been identified and shown to be acidic, neutral and basic in tobacco (Koiwa, Sato, et al., 1994). Most *TLP* consists of 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 (Sharma, Sharma, et al., 2021). TLPs can be divided into two types according to their molecular masses: large type (L) and small type (S). Large type *TLPs* usually have 16 conserved cysteine residues, while small type *TLPs* contain only 10 conserved cysteine residues (Liu, Cui, et al., 2020). Previous studies have reported that the plants such as *Arabidopsis thaliana*, *Oryza sativa*, *Vitis vinifera*, *Cucumis melo*, *Populus trichocarpa*, *Gossypium raimondii*, *G. arboreum*, *G. barbadense*, and *G. hirsutum*, which have been identified to 28, 31, 33, 28, 55, 50, 48, 91, and 90 *TLP* family genes, respectively (Zhang, Chen, et al., 2021).

*TLPs* are pathogenesis-related proteins, which were demonstrated to have a direct antifungal activity (Vigers, Wiedemann, et al., 1992), and specificity for different fungal pathogens species. Numerous studies have shown that plant defense-related *TLP genes* were over-expression after infection by pathogenic fungi, such as: Cotton infected by *Verticillium dahlia* (Li, Wang, et al., 2020), *Gossypium barbadense* infected by *Verticillium dahlia* (Zhang, Chen, et al., 2021), *Fragaria × ananassa* infected by *Colletotrichum gloeosporioides* (Zhang, Miao, et al., 2022), *Carya cathayensis* infected by *Botryosphaeria dothidea* (Li, Xu, et al., 2022), and expression varies among

different tissues, as well as faster *TLP* defense responses to pathogenic fungi in resistant cultivars. In recent years, *TLP* has been applied to the study of breeding disease-resistant crops.

*Panax notoginseng* (Burke) F. H. Chen, known as "Sanqi". It is widely used in China for its healing and tonic properties, such as treating cardiovascular disease, as well as for its antioxidant properties (Wang, Yang, et al., 2021). However, root rot disease is the most destructive disease that hampers industrial development. Root rot, mainly caused by *Fusarium solani*, severely reduces the yield and quality of *P. notoginseng* (Deng, Guan, et al., 2022). Nowadays, the root rot is generally controlled with fungicides, which can cause environmental pollution and leave pesticide residues. Therefore, breeding disease-resistant genotypes is the most sustainable and environmentally friendly measure to control *P. notoginseng* disease (Ning, Li, et al., 2021). The *TLPs* which have been prove to be promising candidates for breeding purposes and for plant transformation aiming a better performance under biotic and abiotic stresses condition (de Jesus-Pires, Ferreira-Neto, et al., 2020).

Previous studies have shown that five *TLP* proteins have been isolated from *P. notoginseng*, and they were demonstrated as defence-related genes in response to an *A. panax* infection (Li, Li, et al., 2020). However, it is a pity that they did not analyze the evolutionary mechanism of *TLP* gene families and identify cis-acting elements in the promoters at the genome-wide level. Fortunately, the whole genome sequence of *P. notoginseng* has been released and published (Jia, Ge, et al., 2023, Yang, Liu, et al., 2021), which allows us to conduct a detailed analysis of the *TLP* gene family in *P. notoginseng*. In the current study, we identified the *TLP* genes in the *P. notoginseng* genome, and systematically analyze the *PnTLPs* physicochemical properties, subcellular localization, gene duplication, phylogeny, gene structure, promoter cis-acting elements and RNA-Seq data, as well as the response to *F. solani* infection of a subset of *PnTLP* genes. The results provide a theoretical basis for revealing the function of *TLP* genes in *P. notoginseng*, which can be used to assist the breeding of disease-resistant *P. notoginseng* cultivars.

## Materials & Methods

### Identification of *PnTLP* genes and analysis of its characteristics

The genomic and annotation data of *P. notoginseng* was obtained in the Herbal Medicine Omics (<http://herbalplant.ynau.edu.cn>) Database (Jia, Ge, et al., 2023). The HMM profile of the thaumatin domain (PF00314) was downloaded from the Pfam website (<http://pfam.xfam.org/>), and then used the hmm-search program HMMWE3.1 to identify the HMM for the *TLP* proteins in the resulting protein database (Liu, Zheng, et al., 2022). We also downloaded *AtTLPs* and *OsTLPs* protein sequences from Ensembl Plants (<http://plants.ensembl.org>) and used as query sequence in BLASTp search against *P. notoginseng* protein sets (Zhang, Chen, et al., 2021, Zhao and Su, 2010). The e-value of  $10^{-5}$  and > 100-bit scores were kept cut off to identify putative *TLP* genes, and eventually, the BLASTp output was tabulated (Ram, Danish, et al., 2022). We compared the results of the two methods to confirm *TLP* candidate genes in these species. Subsequently, the two candidate sets were merged, the redundant proteins were removed, and their conserved domains were further identified by using NCBI-CDD (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) and SMART (<http://smart.embl-heidelberg.de/>) (Liu, Zhao, et al., 2020). The characteristic of *PnTLPs*, including protein length, isoelectric point (pI), molecular weight (MW) and grand average of hydropathicity (GRAVY) were predicted by ExPASy ProtParam (<http://web.expasy.org/protparam/>) (Liu, Cui, et al., 2020). The subcellular localization of

proteins were predicted by the Bologna Unified Subcellular Component Annotator (BUSCA) website (<http://busca.biocomp.unibo.it/>) (Du, Zhang, et al., 2022).

The detailed chromosome distribution information of each *PnTLP* genes was obtained from the *P. notoginseng* genome annotation documents, and visualization was carried out using the TBTools (<https://github.com/CJ-Chen/TBtools>) software (Chen, Mao, et al., 2022).

### **Evolutionary analysis of TLP gene family in *P. notoginseng***

To compare the evolutionary relationships and identify the subfamilies, the putative *TLPs* from *P. notoginseng*, *A. thaliana* and *O. sativa* were used to construct the phylogenetic tree using MEGA-7 with neighbor-joining (NJ) method (Dai, Dong, et al., 2022), in which 1000 replicates of bootstrap parameters were set. The online website Itol (<https://itol.embl.de/>) was used to beautify and visualize the phylogenetic tree (Xue, Hu, et al., 2023).

### **Gene structure and conserved motif analysis of the TLP gene family**

The gene structure information was extracted from the *P. notoginseng* genome data, and a gene structure map was drawn using TBtools (version 1.120) (Deng, Dong, et al., 2023). The conserved motifs of *PnTLPs* with default parameters were analyzed using MEME (<https://meme.nbcr.net/meme/cgi-bin/meme.cgi>), and the maximum discovery number of motifs set to 10 (Li, Wang, et al., 2022).

### **Gene duplication and collinearity analysis of *PnTLP* genes**

The gene duplication of *PnTLPs* in *P. notoginseng* genome was performed by MCscanX with the default parameters, and circos was used to locate all the *PnTLP* genes on the *P. notoginseng* chromosome by Tbtools version 1.120 (Du, Zhang, et al., 2022). The syntenic relationships between the *PnTLP* genes and *TLP* genes from *A. thaliana* and *O. sativa* were obtained also using McscanX and both results visualized in Tbtools (Feng, Abubakar, et al., 2023).

### **Cis-element analysis in the promoter of *PnTLP* genes**

To explore the putative cis-regulating elements in the promoter regions of *PnTLP* genes, the upstream 2000 bp sequences of the transcription initiation codon of all the *TLP* genes were extracted from the genomes of *P. notoginseng*, and cis-regulating elements in the promoter sequence were predicted and analyzed using the PlantCARE databases (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) (Tao, Jia, et al., 2022). The visualization tool TBtools was used to map the distribution of cis-regulating elements in the promoters (Wang, Cheng, et al., 2021).

### **Transcriptome sequencing analysis of *PnTLP* gene expression levels in different tissues and *Alternaria panax* infection**

In this study the RNA-Seq data of different tissues (root, stem, leaf and flower) of *P. notoginseng* were downloaded from the NCBI public database (Accession Number: SRR7427743, SRR7427744, SRR7427758, SRR7427747, SRR7427748, SRR7427750, SRR7427752, SRR7427755, SRR7427756, SRR7427751, SRR7427754, SRR7427757) (Jia, Ge, et al., 2023). The PRJNA706472 project is the transcriptome data of *P. notoginseng* leaf samples infected by the pathogen of *Alternaria panax*. The heat maps of the expression levels of the *TLP* genes were visualized using the Heatmap illustrator program in the toolkit of TBtools (Deng, Dong, et al., 2023).

Inoculation of *Fusarium solani* and qRT-PCR analysis

One-year-old *P. notoginseng* plants with consistent growth were selected and transplanted into Hoagland nutrient solution for hydroponics for two weeks. The plants were then inoculated with *Fusarium solani* by placing them into a 250mL spore suspension ( $1\times10^6$ CFU/mL). After inoculation, the plants were transplanted into the hydroponic environment and root samples were taken at 0, 24, 48, 72, and 96 hours. Three biological repetitions were performed and the samples were stored in liquid nitrogen. Total RNA was extracted using a rapid RNA extraction Kit (for plant) according to the manufacturer’s protocol (TransGen Biotech, Beijing, China). The first strand cDNA was synthesized using the TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix Kit (TransGen Biotech, Beijing, China).

The primers used for the qRT-PCR assay of *TLP* genes were listed in Table 1. Primers were designed online using the NCBI website, and the Actin gene was used as an internal reference gene, with conditions of 95°C for 3min, and.40 cycles of 95°C for 15s, 60°Cfor 1min. The relative gene expression was calculated according to the  $2^{-\Delta\Delta Ct}$  method.

Table 1  
Primers of *PnTLP* genes used for the qRT-PCR analysis

| Gene           | PCR primer sequences (5'→3') |                               |
|----------------|------------------------------|-------------------------------|
| <i>PnACT2</i>  | F: TCCAAGGGTGAATATGATGAATCG  | R: AACCTCTCCAAAGAGAATTTCTGAGT |
| <i>PnTLP01</i> | F: CCGGTTTCGCCCTTCAAAAG      | R: CTCCTGTGACGCAGGTGAAT       |
| <i>PnTLP02</i> | F: ACGTTTACGTGTACCGGAGC      | R: CTTCTCTTGGAACCCGGACC       |
| <i>PnTLP05</i> | F: GGCCAAGTTGAGTGCAATGG      | R: GCCATCCACGAGGCTTACAT       |
| <i>PnTLP06</i> | F: GATTGTGGCGGAGCCCTATT      | R: TAACACAGCCGGCATAGGTG       |
| <i>PnTLP08</i> | F: CGGCCACCATCGTTGAAATG      | R: GGAGTCATTGCCACCGGAAT       |
| <i>PnTLP10</i> | F: GCCCCTTTCCAATATGGCCT      | R: CCGTTCCAATCTCCTGGGAC       |
| <i>PnTLP14</i> | F: AGCCGAATATGCCCTGAACC      | R: GGTTATTACACCCACCCGCA       |
| <i>PnTLP15</i> | F: GCTCCTGTCCACAACGTAGG      | R: CTCGGAGCAACCCAAGACAT       |
| <i>PnTLP17</i> | F: GCACAGGATTTGAGCTCCCT      | R: CATTCCACCTGACCAGACCC       |

Results

Identification and characterization of TLP genes in *P. notoginseng*

A total of twenty *PnTLP* genes were identified and confirmed with the *TLP* domain by NCBI-CDD and SMART database, and they were named as *PnTLP01* to *PnTLP20*. We analyzed the basic characteristics of the *PnTLP* gene family members (Table 2), including coding sequence length (CDS), the number of amino acids in the protein, isoelectric point (pI), relative molecular weight (MW), grand average of hydropathicity (GRAVY), and subcellular localization. The results showed that the CDS length among all *PnTLPs* was 357 bp to 5277 bp, and the encoded protein size ranged from 118 to 1758 aa. The pI index ranged from 4.35 to 8.99, and the lowest MW value was *PnTLP19* (12.88 kDa) and the highest was *PnTLP06* (197.67 kDa). The GRAVY index ranged from –

0.618 (*PnTLP06*) to 0.41 (*PnTLP20*). The prediction of the subcellular localization results showed that most of the *PnTLP* proteins were confined in the extracellular and chloroplast, except for the *PnTLP06*, *PnTLP09*, and *PnTLP15* which were localized in nuclear, Plasma membrane, and Vacuolar, respectively.

To know whether *PnTLP* gene family has covered the whole genome or it is randomly distributed in the *P. notoginseng* genome, chromosomal distribution of *PnTLP* gene family members was determined. A total of 20 *PnTLP* genes were located on 9 chromosomes, of which Chr.2 and Chr.5 contained four genes, Chr.1 and Chr4 contained three genes, excepted for Chr.11 contained two genes, Chr3, Chr7, and Chr8 contained only one *PnTLP* genes (Fig. 1).

Table 2  
Basic characteristics of 20 *PnTLP* genes in *P. notoginseng*

| Gene ID         | Gene Name      | Locus | CDS/bp | Size/aa | pI   | MW(kDa) | GRAVY  | Subcellular Location |
|-----------------|----------------|-------|--------|---------|------|---------|--------|----------------------|
| Pno01G000321.t1 | <i>PnTLP01</i> | Chr1  | 957    | 318     | 4.81 | 32.94   | 0.01   | Extracellular        |
| Pno01G000322.t1 | <i>PnTLP02</i> | Chr1  | 978    | 325     | 5.00 | 34.55   | -0.178 | Chloroplast          |
| Pno01G006130.t1 | <i>PnTLP03</i> | Chr1  | 705    | 234     | 4.54 | 23.97   | 0.128  | Chloroplast          |
| Pno02G000923.t1 | <i>PnTLP04</i> | Chr2  | 738    | 245     | 4.35 | 25.35   | 0.107  | Chloroplast          |
| Pno02G000925.t1 | <i>PnTLP05</i> | Chr2  | 924    | 307     | 5.03 | 31.81   | 0.013  | Chloroplast          |
| Pno02G006322.t1 | <i>PnTLP06</i> | Chr2  | 5277   | 1758    | 8.09 | 197.67  | -0.618 | Nuclear              |
| Pno02G006323.t1 | <i>PnTLP07</i> | Chr2  | 738    | 245     | 8.99 | 26.38   | 0.049  | Chloroplast          |
| Pno03G000876.t1 | <i>PnTLP08</i> | Chr3  | 864    | 287     | 8.76 | 31.10   | 0.254  | Chloroplast          |
| Pno04G004238.t1 | <i>PnTLP09</i> | Chr4  | 852    | 283     | 7.75 | 30.29   | 0.147  | Plasma membrane      |
| Pno04G004995.t1 | <i>PnTLP10</i> | Chr4  | 789    | 262     | 7.82 | 28.54   | -0.144 | Chloroplast          |
| Pno04G005703.t1 | <i>PnTLP11</i> | Chr4  | 852    | 283     | 4.52 | 29.54   | 0.008  | Chloroplast          |
| Pno05G000135.t1 | <i>PnTLP12</i> | Chr5  | 891    | 296     | 4.79 | 31.30   | 0.246  | Extracellular        |
| Pno05G000239.t1 | <i>PnTLP13</i> | Chr5  | 951    | 316     | 4.56 | 32.98   | -0.02  | Extracellular        |
| Pno05G004365.t1 | <i>PnTLP14</i> | Chr5  | 726    | 241     | 7.84 | 26.47   | -0.341 | Extracellular        |
| Pno05G005586.t1 | <i>PnTLP15</i> | Chr5  | 1248   | 415     | 6.66 | 45.49   | -0.277 | Vacuolar             |
| Pno06G005644.t1 | <i>PnTLP16</i> | Chr6  | 738    | 245     | 6.84 | 25.77   | 0.17   | Extracellular        |
| Pno07G005021.t1 | <i>PnTLP17</i> | Chr7  | 1791   | 596     | 5.71 | 62.39   | -0.068 | Extracellular        |
| Pno08G002837.t1 | <i>PnTLP18</i> | Chr8  | 738    | 245     | 4.71 | 25.61   | 0.031  | Chloroplast          |
| Pno11G000360.t1 | <i>PnTLP19</i> | Chr11 | 357    | 118     | 8.28 | 12.88   | -0.042 | Chloroplast          |
| Pno11G000361.t1 | <i>PnTLP20</i> | Chr11 | 447    | 148     | 4.61 | 15.45   | 0.41   | Extracellular        |

## Phylogenetic and evolutionary analyses

In order to understand the evolutionary relationship of *TLP* genes among different species, 10 representative *Arabidopsis thaliana* and *Oryza sativa* *TLPs* sequences were used as reference sequences together with *P. notoginseng* *TLPs* in the phylogenetic analysis. Similar to the phylogenetic results of *Arabidopsis* and *Oryza* *TLPs*, the 20 *PnTLLPs* were also classified in 10 phylogenetic groups, Group I to X (Fig. 2). A maximum number (4) of *PnTLPs* were clustered in Group V, followed by Group I, Group II, Group III, Group IV, and Group VI contained two genes, all of them accounted for 85% of the *PnTLPs*. The other Groups contained the least amount of *TLP* family genes with one gene each. Each of these subgroups contained at least one *AtTLP* gene, indicating the close relationship between *PnTLPs* and those of other plants.

### Gene structures and motif analysis of *PnTLPs*

The unrooted tree of 20 *PnTLP* genes in *P. notoginseng* showed that all *PnTLPs* clearly evolved into ten main clades (Fig. 3a). To improve our understanding of gene function and evolution, we analyzed the *TLP* gene structure (Fig. 3b). Although some *PnTLPs* are very close phylogenetically, their exon number, distribution, and gene length greatly differ, especially Group V (*PnTLP02*, *PnTLP05*, *PnTLP13*, *PnTLP17*), and Group I (*PnTLP06* and *PnTLP07*). Analysis of gene structure showed that *PnTLP06* consisted of fifteen exons; *PnTLP17* consisted of five exons; followed by *PnTLP07* and *PnTLP11* consisted of three exons; *PnTLP01*, *PnTLP02*, *PnTLP05*, *PnTLP09*, *PnTLP13* and *PnTLP15* consisted of three exons, and the other *PnTLPs* consisted of one exon.

To further understand the process of *TLP* evolution, the sequence characteristics of 20 *TLP* genes in *P. notoginseng* were analyzed using MEME software, and a total of 10 conserved motifs were obtained. According to the motif analysis results, four motif patterns (motif 2, motif 3, motif 5 and motif 8) were found within 20 *PnTLP* genes (Fig. 3c). In addition, except for *PnTLP14*, *PnTLP15*, *PnTLP19*, and *PnTLP20*, all other *PnTLP* genes were contained motif 1 and motif 7. These results suggested the *P. notoginseng* *TLP* family proteins shared highly conserved structural basis. It was confirmed that the diversity of *TLP* gene evolution was affected by gene structure and motif patterns.

### Collinearity analysis and duplication events of *PnTLP* genes

The segmental and tandem duplication events are major drivers of the gene family expansion. From the analysis of fragment duplication events in the *P. notoginseng* *TLP* gene family (Fig. 4a), a total of six pairs of fragment duplication genes were identified in the 20 *PnTLPs* gene members (*PnTLP01* and *PnTLP11*, *PnTLP01* and *PnTLP13*, *PnTLP03* and *PnTLP18*, *PnTLP04* and *PnTLP11*, *PnTLP08* and *PnTLP16*, *PnTLP11* and *PnTLP13*), and each segment duplicated genes belong to the same subclass, suggesting that these chromosomal segments may have undergone duplication events during evolution and did not fully diverge, and may be functionally redundant. In addition, covariance analysis was performed between *P. notoginseng* and representative model plants of *A. thaliana* and *O. sativa*. The results showed (Fig. 4b) that the largest number of 36 homologous gene pairs existed between *P. notoginseng* and *A. thaliana*, and 9 homologous gene pairs existed with *O. sativa*. It indicated that *P. notoginseng* was more closely related to the dicotyledonous plant *A. thaliana* and more distantly related to the monocotyledonous plant *O. sativa*.

### Cis-acting element analysis of *PnTLP* gene family

The cis-acting regulatory element is a specific promoter motif to which an appropriate transcription factor binds in order to regulate gene transcription in plants. The results showed that there are a large number of core elements (TATA-box and CAAT-box) in the promoter region of the *PnTLP* family, and also contain response elements related

to light-responsive, hormone-responsive, tissue responsive, stress-responsive (biotic and abiotic stresses), metabolic responsive, cellular development, MYB-binding sites, growth and development and other functions responsive (Fig. 5). It is speculated that the expression regulation mode of *PnTLP* is relatively complex, may be affected by multiple factors. Among them, the number of light-responsive elements, stress response element and hormone response elements is the largest.

### Expression patterns analysis of *PnTLP* genes by RNA-Seq

The expression levels of members of *PnTLP* gene family in each tissue were generated in a heatmap, and the main results were analyzed as follows (Fig. 6a), most of the *TLP* genes were expressed at different levels in these four tissues, and the expression of some genes is highly tissue-specific. Interestingly, compared with other family members, the expression of *PnTLP02*, *PnTLP05*, *PnTLP06*, *PnTLP07*, *PnTLP09*, *PnTLP10*, *PnTLP11*, *PnTLP12*, *PnTLP13*, *PnTLP15*, *PnTLP16* and *PnTLP17* genes were highly expressed in stem and flowers, and general or down-regulated in leaves and roots. The expression levels of *PnTIFY1*, *PnTIFY8*, *PnTIFY19* and *PnTIFY20* genes were highly expressed in leaves, and their expression levels in other part were general or down-regulated. In addition, only the *PnTLP14* gene was highly expressed in the roots. It seemed that most of the *PnTLP* genes were only highly expressed in one specific tissue and the other tissues were expressed generally or down regulated. Generally speaking, expression patterns of *PnTLP* genes in different tissues can be used to identify the functional genes in *P. notoginseng*.

For the expression profiles of *TLPs* in response to biotic stress, the RNA-Seq data from a previous study was reanalyzed here. As seen in Fig. 6b, the expression levels of specific *TLP* genes in the leaves were found to be significantly up-regulated (*PnTLP01*, *PnTLP02*, *PnTLP10*, *PnTLP14* and *PnTLP15*) and down-regulated (*PnTLP05*, *PnTLP06*, *PnTLP07*, *PnTLP08* and *PnTLP16*). By reanalyzing the transcriptomic data, we succeeded in identifying the major biotic stress-responsive *TLPs* in *P. notoginseng*.

### Expression of *PnTLP* genes in response to *F. solani* Infection

To test whether the expression of *PnTLP* genes was influenced by challenge of *F. solani*, 9 *PnTLP* members that were random selected to analysis their responses to *F. solani* infection in *P. notoginseng* (Fig. 7). The relative expression levels of *PnTLP01*, *PnTLP02*, *PnTLP05*, *PnTLP14* and *PnTLP15* were significantly up-regulated at 48 h and 96 h following inoculation with *F. solani* compared to the control (0 h), and *PnTLP02*, *PnTLP05*, *PnTLP14* and *PnTLP15* showed increased first and then decreased, reaching the lowest level at 96 h. The relative expression levels of *PnTLP06*, *PnTLP08* and *PnTLP17* were down-regulated after 96 h infection, and *PnTLP08* were down-regulated during the whole infection process. The comprehensive analysis showed that *P. notoginseng PnTLPs* gene family members could respond to pathogen infection in different degrees.

## Discussion

Thaumatococcal proteins (*TLPs*) are vital proteins involved in plant defense, developmental processes and seed germination (Iqbal, Tripathi, et al., 2020). Several studies have confirmed that plant *TLPs* are potential candidates for plant transformation and breeding with the goal of improving performance under biotic and abiotic stresses (de Jesus-Pires, Ferreira-Neto, et al., 2020). However, there have been few reports on *TLPs* of *P. notoginseng*. In this study, we performed a systematic and comprehensive analysis of *PnTLPs* gene.

There were twenty *TLP* genes identified in *P. notoginseng* genome, which is similar to the numbers reported 19 for barley (Iqbal, Tripathi, et al., 2020), less than the numbers reported 28 for *Arabidopsis* (Liu, Cui, et al., 2020) and 37 for rice (Li, Xu, et al., 2022), but more than the numbers reported 15 for maize (Li, Wang, et al., 2022). It suggesting that *TLP* genes family members contained in different plant genomes varies widely. Subcellular localization indicated that most of the *PnTLPs* were localized to the extracellular. Similarly, the extracellular localization of the *TaTLP2-B* was also reported in an earlier study (Sharma, Sharma, et al., 2021). Studies showed that the *TaLr19TLP1* gene encodes an extracellular protein, and it plays a definitive role in wheat resistance to leaf rust fungus via mul-tiple defense signaling pathways (Zhang, Yan, et al., 2017). Therefore, we hypothesized that these extracellular *TLPs* may be related to its involvement in various defense-related functions.

Gene duplication is the main mode of gene family expansion, where whole-genome duplications was considered to be an important driver of expansion and an important source of functional diversification, and low tandem and high segmental duplications of the *TLP* gene family were widely observed in plants (Cao, Lv, et al., 2015, Ram, Danish, et al., 2022). Gene duplication analysis showed that thirteen gene duplication events in the whole-genome of watermelon, which included one tandem and twelve segmental duplications (Ram, Danish, et al., 2022). Meanwhile, twelve segmental and seven tandem duplication gene pairs that underwent purifying selection were identified in the whole-genome of melon (Liu, Cui, et al., 2020). In this study, a total of six segmentally duplicated *PnTLP* gene pairs were found in the whole-genome of *P. notoginseng*. These results indicated that gene duplication events, especially segmental duplication is the main driver of *PnTLP* gene family evolution and amplification.

The promoter cis-acting elements were involved in the regulation of gene expression, and the gene function could be preliminarily predicted by analyzing the promoter cis-acting elements of *TLP* family genes (Su, Song, et al., 2023). In the present study, we identified more than 23 putative cis-regulatory elements in promoter region of the *PnTLP* gene family including hormone response elements (methyl jasmonate, abscisic acid, gibberellin, growth hormone, salicylic acid and ethylene), adversity response elements (anaerobic induction, low temperature, defense and stress), light response elements, development responsive elements, metabolic responsive elements, and other functions responsive elements. Among the hormone response elements, MeJA and ABA response elements are numerous, and a few other hormone response components. A number of plant hormones, including ABA, SA, and JA, have a role in stress reactions by triggering the transcription of genes involved in defense (Zheng, Wan, et al., 2022). Which shown that *PnTLPs* may be susceptible to MeJA and ABA regulation and thus play a role in plant growth and development as well as in adversity stress, which is generally consistent with the studies in watermelon (Ram, Danish, et al., 2022), woodland strawberry (Li, Wang, et al., 2022), and Hickory (Li, Xu, et al., 2022).

The clustering results showed that the *TLP* gene family members of *P. notoginseng* could be grouped into 10 clusters, which is similar to those of a previous report comparing the evolutionary relationship of *Arabidopsis* and rice *TLPs* (Shatters, Boykin, et al., 2006, Zhang, Miao, et al., 2022). It is noteworthy that *PnTLP14* belongs to a subgroup with AT4G11650, it encodes a protein which participates in the signaling pathway of *A. thaliana* to respond to pathogenic and environmental stress (Zhang, Chen, et al., 2021). Further transcriptome data and qRT-PCR analysis verified that *PnTLP14* expression was significantly higher in *P. notoginseng* roots than other tissue sites. We also discovered that *PnTLP14* was significantly up-regulated by the pathogenic bacteria *A. panax* and *F. solani* under stress, which led to the hypothesis that *PnTLP14* has a similar function to AT4G11650. Therefore, *PnTLP14* could be a candidate gene for studying the disease resistance mechanism of *P. notoginseng*.

# Conclusions

In this study, a total of 20 *PnTLP* genes were identified in the *P. notoginseng* genome. They were divided into ten subfamilies based on phylogenetic relationships, gene structures and protein conserved domains. Meanwhile, cis-acting regulatory elements indicated that *PnTLPs* played important roles in biotic stress and hormone induction. Additionally, we found that segmental duplication events was involved in the number of *PnTLP* genes in the process of plant evolution. RNA-seq data showed that *PnTLP* genes had a tissue-specific expression, especially in flowers, stems, and leaves. Furthermore, analysis of quantitative expression showed that the nine representative genes showed significantly different expression levels in *P. notoginseng* following *F. solani* infection. In summary, the results showed that *PnTLP* genes play important roles in responses to stresses and can be used as a basis for further in-depth study. Together, these results strengthen the understanding of the *PnTLP* gene family in *P. notoginseng* and provide candidate genes for future studies of their roles in disease resistance.

# Declarations

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work was supported by the Yunnan Academician Expert Workstation (grant number: 202305AF150027); Major Science and Technology Special Project of Yunnan Province (grant number: 202202AE090025); Yunnan (Kunming) Academician Expert Workstation (grant number: YSZJGZZ-2021062); National Natural Science Foundation of China (grant number: 31800443); Postgraduate research Innovation Fund Project in Yunnan Normal University (grant number: YJSJJ22-A15).

## Data availability

All data and materials used in this research are publicly available. Other supporting data are provided as supplementary files with the manuscript.

# References

1. Cao, J., Y. Lv, Z. Hou, X. Li and L. Ding. 2015. Expansion and evolution of thaumatin-like protein (*TLP*) gene family in six plants. *Plant Growth Regulation* 79: 299-307.
2. Chen, X., Y. Mao, W. Chai, K. Yan, Z. Liang and P. Xia. 2022. Genome-wide identification and expression analysis of *MYB* gene family under nitrogen stress in *Panax notoginseng*. *Protoplasma*.
3. Dai, Z., S. Dong, H. Miao, X. Liu, J. Han, C. Li, et al. 2022. Genome-Wide Identification of *TIFY* Genes and Their Response to Various Pathogen Infections in Cucumber (*Cucumis sativus* L.). *Scientia Horticulturae* 295.
4. De Jesus-Pires, C., J.R.C. Ferreira-Neto, J. Pacifico Bezerra-Neto, E.A. Kido, R.L. de Oliveira Silva, V. Pandolfi, et al. 2020. Plant Thaumatin-like Proteins: Function, Evolution and Biotechnological Applications. *Curr Protein Pept Sci* 21: 36-51.

5. Deng, J., R. Guan, T. Liang, L. Su, F. Ge, X. Cui, et al. 2022. Dirigent gene family is involved in the molecular interaction between *Panax notoginseng* and root rot pathogen *Fusarium solani*. *Industrial Crops and Products* 178.
6. Deng, M., Y. Dong, S. Xu, S. Huang, X. Zhai and G. Fan. 2023. Genome-Wide Identification and Expression of the *Paulownia fortunei* MADS-Box Gene Family in Response to Phytoplasma Infection. *Genes (Basel)* 14.
7. Du, J., Q. Zhang, S. Hou, J. Chen, J. Meng, C. Wang, et al. 2022. Genome-Wide Identification and Analysis of the R2R3-MYB Gene Family in *Theobroma cacao*. *Genes (Basel)* 13.
8. Du, Y., N. Amin, N. Ahmad, H. Zhang, Y. Zhang, Y. Song, et al. 2023. Identification of the Function of the Pathogenesis-Related Protein GmPR1L in the Resistance of Soybean to *Cercospora sojina* Hara. *Genes (Basel)* 14.
9. Feng, X., A.S. Abubakar, K. Chen, C. Yu, A. Zhu, J. Chen, et al. 2023. Genome-wide analysis of R2R3-MYB transcription factors in *Boehmeria nivea* (L.) gaudich revealed potential cadmium tolerance and anthocyanin biosynthesis genes. *Front Genet* 14.
10. Iqbal, I., R.K. Tripathi, O. Wilkins and J. Singh. 2020. Thaumatin-Like Protein (*TLP*) Gene Family in Barley: Genome-Wide Exploration and Expression Analysis during Germination. *Genes (Basel)* 11.
11. Jia, J.S., N. Ge, Q.Y. Wang, L.T. Zhao, C. Chen and J.W. Chen. 2023. Genome-wide identification and characterization of members of the *LEA* gene family in *Panax notoginseng* and their transcriptional responses to dehydration of recalcitrant seeds. *BMC Genomics* 24: 126.
12. Kaur, A., P.K. Pati, A.M. Pati and A.K. Nagpal. 2017. In-silico analysis of cis-acting regulatory elements of pathogenesis-related proteins of *Arabidopsis thaliana* and *Oryza sativa*. *PLoS One* 12: e0184523.
13. Koiwa, H., F. Sato and Y. Yamada. 1994. Characterization of Accumulation of Tobacco PR-5 Proteins by IEF-Immunoblot Analysis. *Plant and Cell Physiology* 35: 821-827.
14. Li, P., Y. Xu, K. Wang, W. Guo, Y. Gu, S. Lyu, et al. 2022. Genome-Wide Identification of *TLP* Gene Family and Their Roles in *Carya cathayensis* Sarg in Response to *Botryosphaeria dothidea*. *Front Plant Sci* 13.
15. Li, S., G. Wang, L. Chang, R. Sun, R. Wu, C. Zhong, et al. 2022. Genome-Wide Identification, Characterization, and Expression Analysis of Tubby-like Protein (*TLP*) Gene Family Members in Woodland Strawberry (*Fragaria vesca*). *Int J Mol Sci* 23.
16. Li, X., S. Li, B. Qiu, Y. Zhang, X. Cui, F. Ge, et al. 2020. Thaumatin-like protein genes of *Panax notoginseng* confers resistance to *Alternaria panax*. *Physiological and Molecular Plant Pathology* 112.
17. Li, Z., X. Wang, Y. Cui, K. Qiao, L. Zhu, S. Fan, et al. 2020. Comprehensive Genome-Wide Analysis of Thaumatin-Like Gene Family in Four Cotton Species and Functional Identification of GhTLP19 Involved in Regulating Tolerance to *Verticillium dahlia* and Drought. *Front Plant Sci* 11: 575015.
18. Linthorst, H.J.M. and L.C. Van Loon. 1991. Pathogenesis-related proteins of plants. *Critical Reviews in Plant Sciences* 10: 123-150.
19. Liu, X., C. Zhao, L. Yang, Y. Zhang, Y. Wang, Z. Fang, et al. 2020. Genome-Wide Identification, Expression Profile of the *TIFY* Gene Family in *Brassica oleracea* var. capitata, and Their Divergent Response to Various Pathogen Infections and Phytohormone Treatments. *Genes (Basel)* 11.
20. Liu, Y., J. Cui, X. Zhou, Y. Luan and F. Luan. 2020. Genome-wide identification, characterization and expression analysis of the *TLP* gene family in melon (*Cucumis melo* L.). *Genomics* 112: 2499-2509.
21. Liu, Y.L., L. Zheng, L.G. Jin, Y.X. Liu, Y.N. Kong, Y.X. Wang, et al. 2022. Genome-Wide Analysis of the Soybean *TIFY* Family and Identification of GmTIFY10e and GmTIFY10g Response to Salt Stress. *Front Plant*

22. Ning, K., M. Li, G. Wei, Y. Zhou, G. Zhang, H. Huai, et al. 2021. Genomic and Transcriptomic Analysis Provide Insights Into Root Rot Resistance in *Panax notoginseng*. *Frontiers in Plant Science* 12.
23. Ram, C., S. Danish, M.S. Kesawat, B.S. Panwar, M. Verma, L. Arya, et al. 2022. Genome-wide comprehensive characterization and expression analysis of *TLP* gene family revealed its responses to hormonal and abiotic stresses in watermelon (*Citrullus lanatus*). *Gene* 844: 146818.
24. Sharma, A., H. Sharma, R. Rajput, A. Pandey and S.K. Upadhyay. 2021. Molecular Characterization Revealed the Role of Thaumatin-Like Proteins of Bread Wheat in Stress Response. *Front Plant Sci* 12: 807448.
25. Shatters, R.G., Jr., L.M. Boykin, S.L. Lapointe, W.B. Hunter and A.A. Weathersbee, 3rd. 2006. Phylogenetic and structural relationships of the PR5 gene family reveal an ancient multigene family conserved in plants and select animal taxa. *J Mol Evol* 63: 12-29.
26. Su, J., S. Song, Y. Wang, Y. Zeng, T. Dong, X. Ge, et al. 2023. Genome-wide identification and expression analysis of *DREB* family genes in cotton. *BMC Plant Biol* 23: 169.
27. Tao, J., H. Jia, M. Wu, W. Zhong, D. Jia, Z. Wang, et al. 2022. Genome-wide identification and characterization of the *TIFY* gene family in kiwifruit. *BMC Genomics* 23: 179.
28. Vigers, A.J., S. Wiedemann, W.K. Roberts, M. Legrand, C.P. Selitrennikoff and B. Fritig. 1992. Thaumatin-like pathogenesis-related proteins are antifungal. *Plant Science* 83: 155-161.
29. Wang, K., Y. Cheng, L. Yi, H. He, S. Zhan and P. Yang. 2021. Genome-wide identification of the Tubby-Like Protein (TLPs) family in medicinal model plant *Salvia miltiorrhiza*. *PeerJ* 9: e11403.
30. Wang, P., L. Yang, J. Sun, Y. Yang, Y. Qu, C. Wang, et al. 2021. Structure and Function of Rhizosphere Soil and Root Endophytic Microbial Communities Associated With Root Rot of *Panax notoginseng*. *Front Plant Sci* 12: 752683.
31. Xue, G., L. Hu, L. Zhu, Y. Chen, C. Qiu, R. Fan, et al. 2023. Genome-Wide Identification and Expression Analysis of CCO Gene Family in *Liriodendron chinense*. *Plants* 12.
32. Yang, Z., G. Liu, G. Zhang, J. Yan, Y. Dong, Y. Lu, et al. 2021. The chromosome-scale high-quality genome assembly of *Panax notoginseng* provides insight into dencichine biosynthesis. *Plant Biotechnology Journal* 19: 869-871.
33. Zhang, Y., W. Chen, X. Sang, T. Wang, H. Gong, Y. Zhao, et al. 2021. Genome-Wide Identification of the Thaumatin-like Protein Family Genes in *Gossypium barbadense* and Analysis of Their Responses to *Verticillium dahliae* Infection. *Plants (Basel)* 10.
34. Zhang, Y., L. Miao, X. Yang and G. Jiang. 2022. Genome-wide characterization and expression of the *TLP* gene family associated with *Colletotrichum gloeosporioides* inoculation in *Fragaria x ananassa*. *PeerJ* 10: e12979.
35. Zhang, Y., H. Yan, X. Wei, J. Zhang, H. Wang and D. Liu. 2017. Expression analysis and functional characterization of a pathogen-induced thaumatin-like gene in wheat conferring enhanced resistance to *Puccinia tritici*. *Journal of Plant Interactions* 12: 332-339.
36. Zhao, J.P. and X.H. Su. 2010. Patterns of molecular evolution and predicted function in thaumatin-like proteins of *Populus trichocarpa*. *Planta* 232: 949-962.
37. Zheng, L., Q. Wan, H. Wang, C. Guo, X. Niu, X. Zhang, et al. 2022. Genome-wide identification and expression of *TIFY* family in cassava (*Manihot esculenta* Crantz). *Front Plant Sci* 13: 1017840.

# Figures

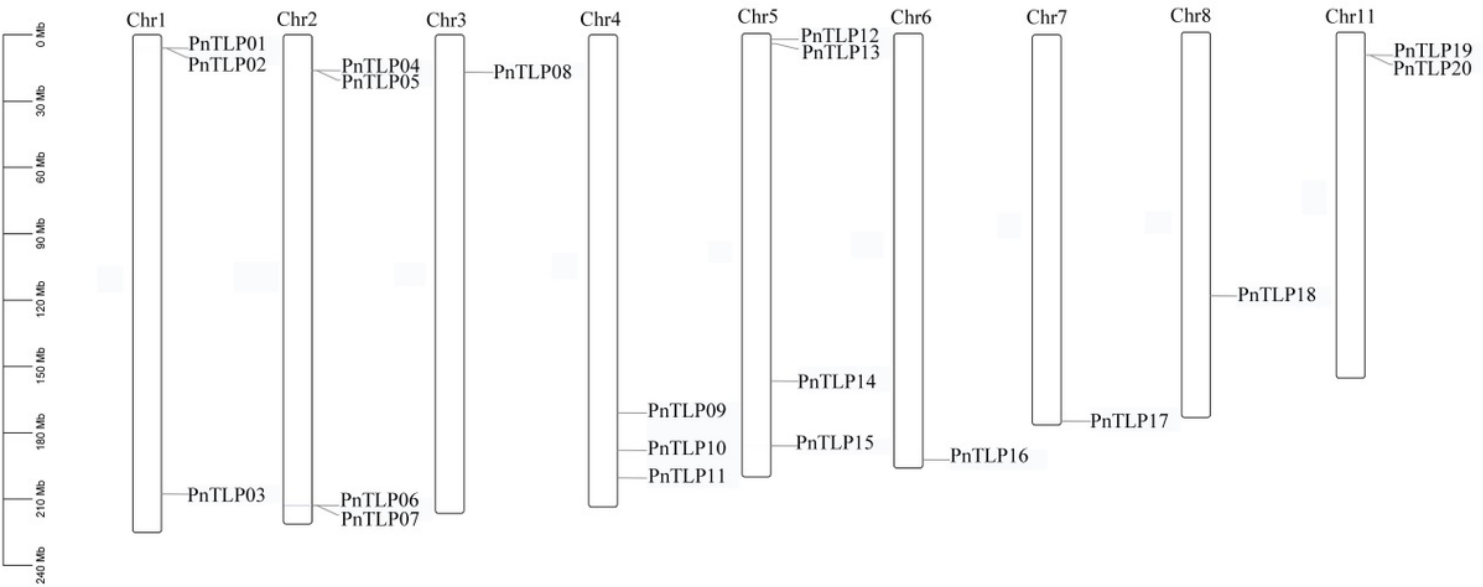


**Figure 1**

Chromosome distribution of *TLP* gene family in the genome of *P. notoginseng*.

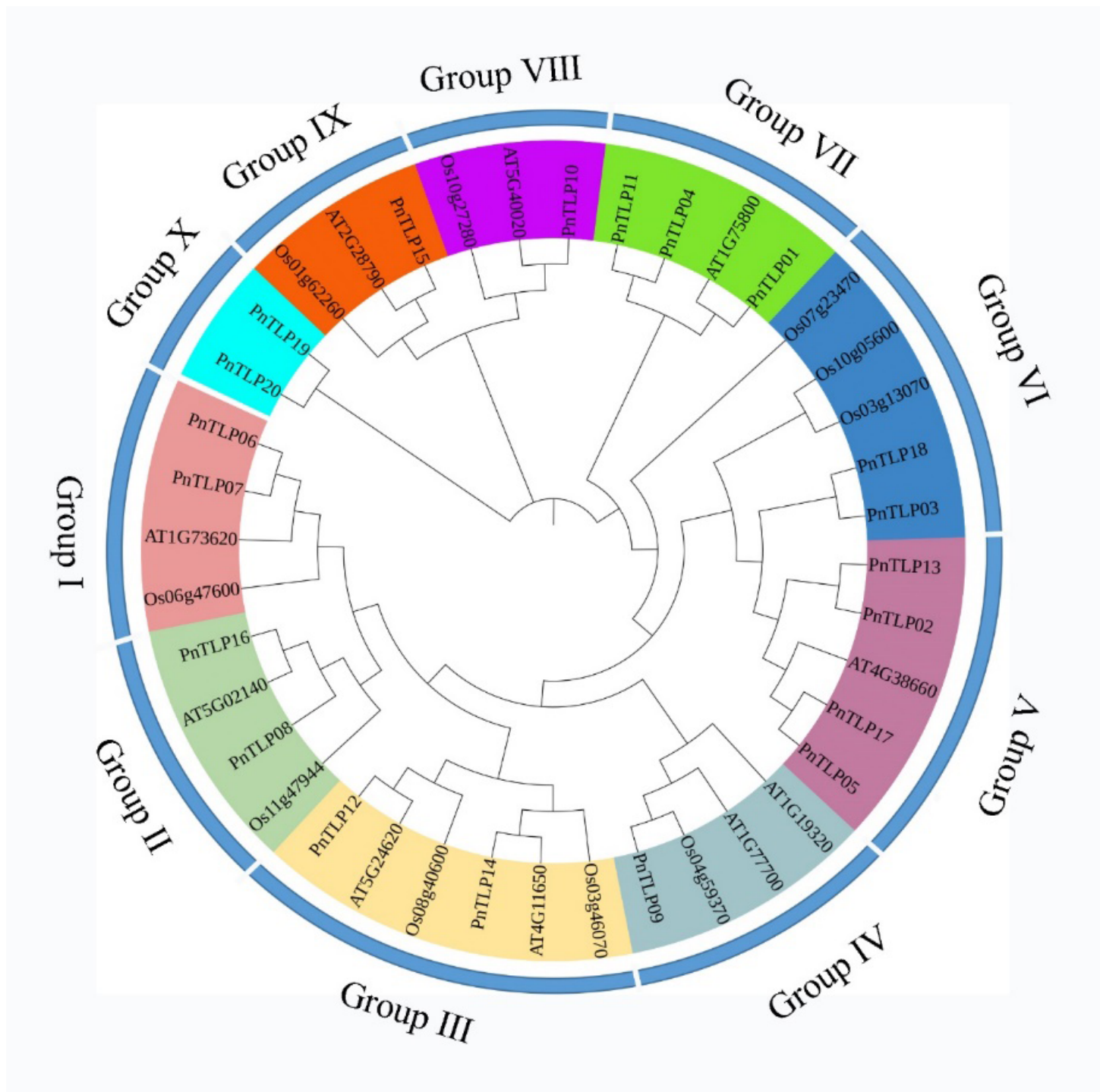


**Figure 2**

The phylogenetic classification of *TLP* proteins from *P. notoginseng*, *O. sativa* and *A. thaliana*.

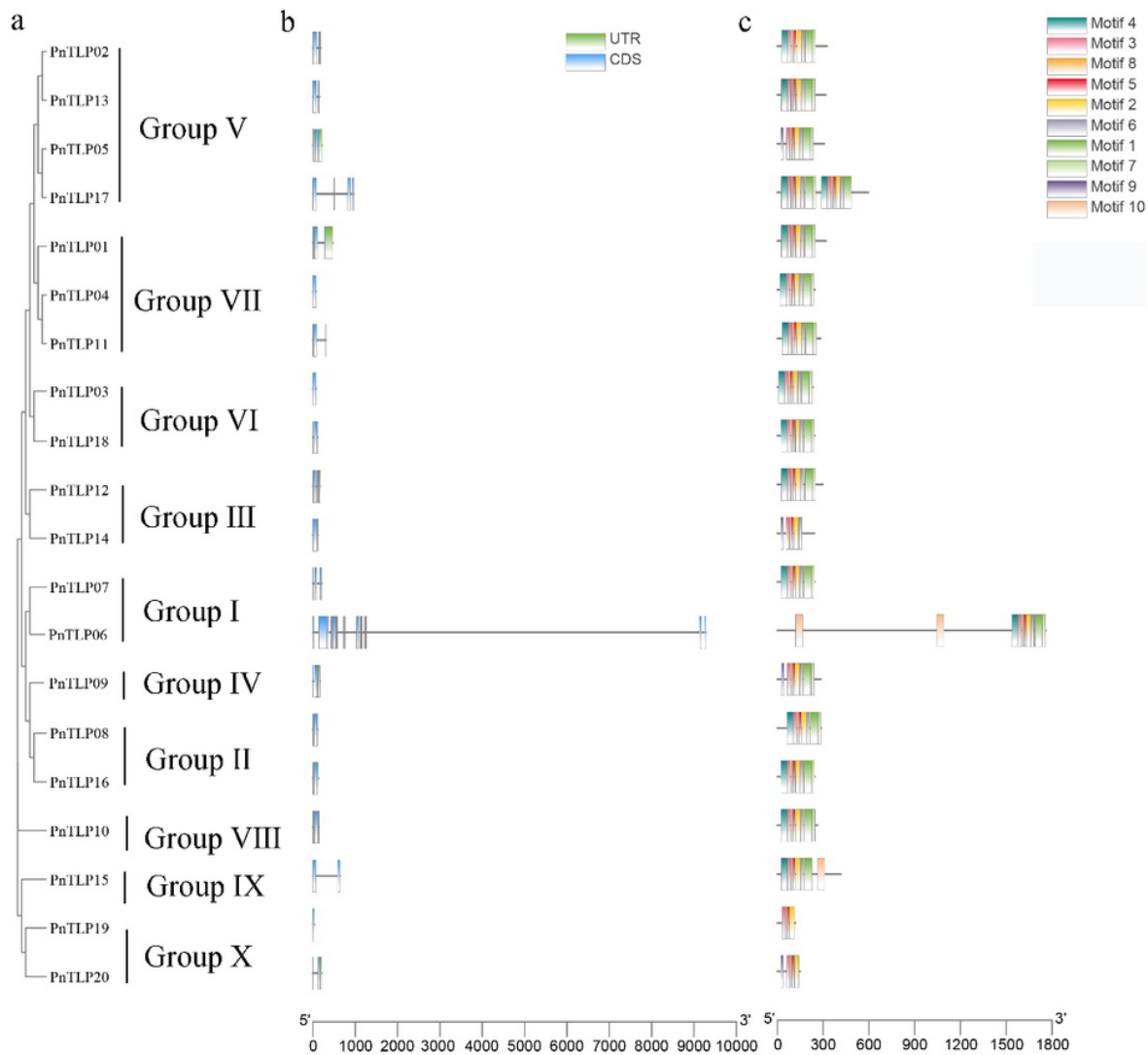


**Figure 3**

Phylogenetic clustering, gene structure, and motifs *PnTLPs*. **a** The Neighbor-Joining (NJ) tree of *PnTLPs* based on full-length protein sequences. **b** Gene structure of *PnTLPs*. **c** The ten motifs of *PnTLPs*.

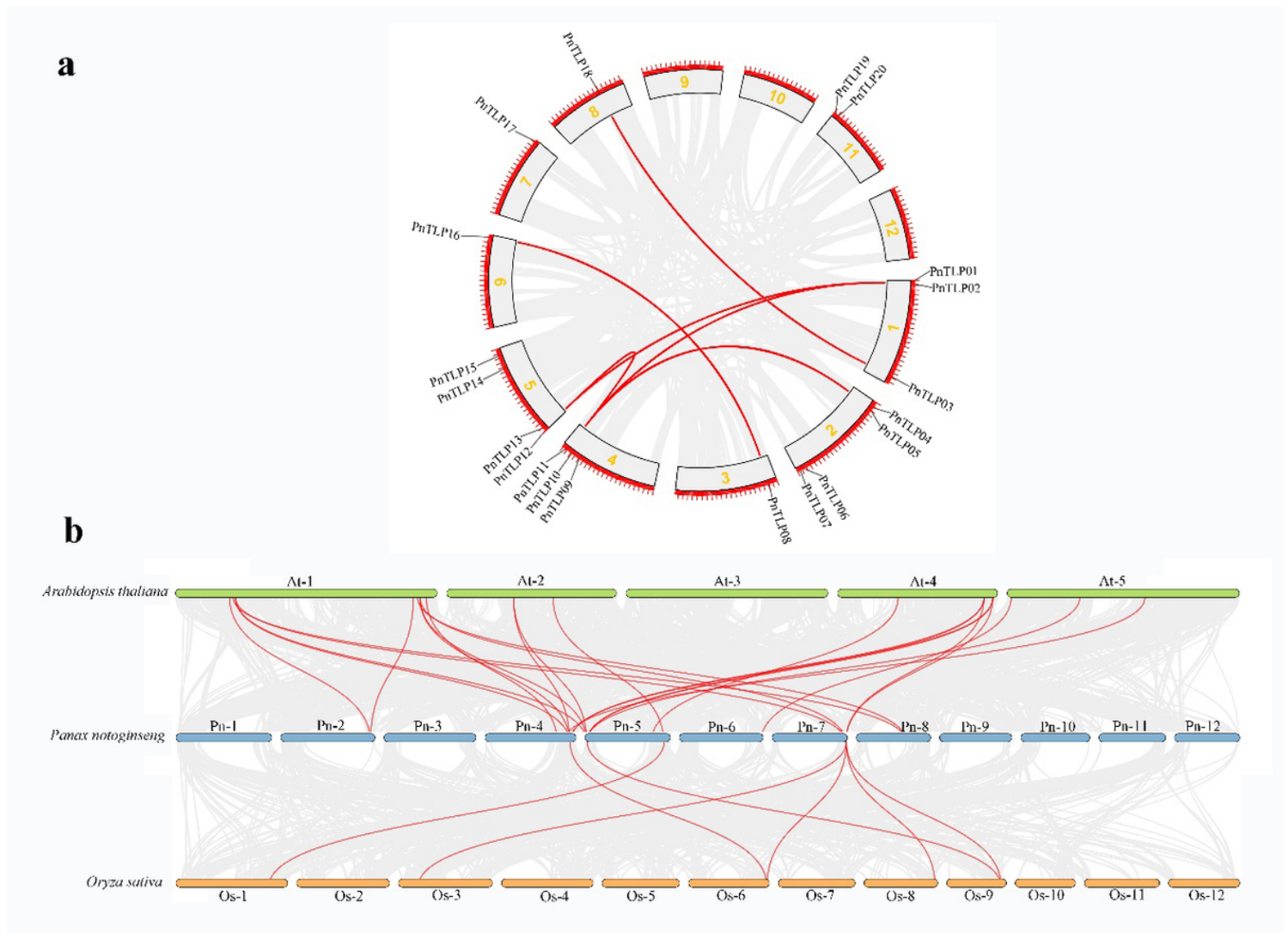


**Figure 4**

Collinearity analysis of the *TLP* gene family in the *P. notoginseng* genome. **a** Synteny analysis of the *PnTLP* genes. **b** Synteny analysis of *TLP* genes between *P. notoginseng* and *Arabidopsis thaliana*, and between *P. notoginseng* and *Oryza sativa*, respectively. The red lines indicate the duplication of the *TLP* gene pairs.



**Figure 5**

Cis-element analysis of putative *PnTLP* promoters.

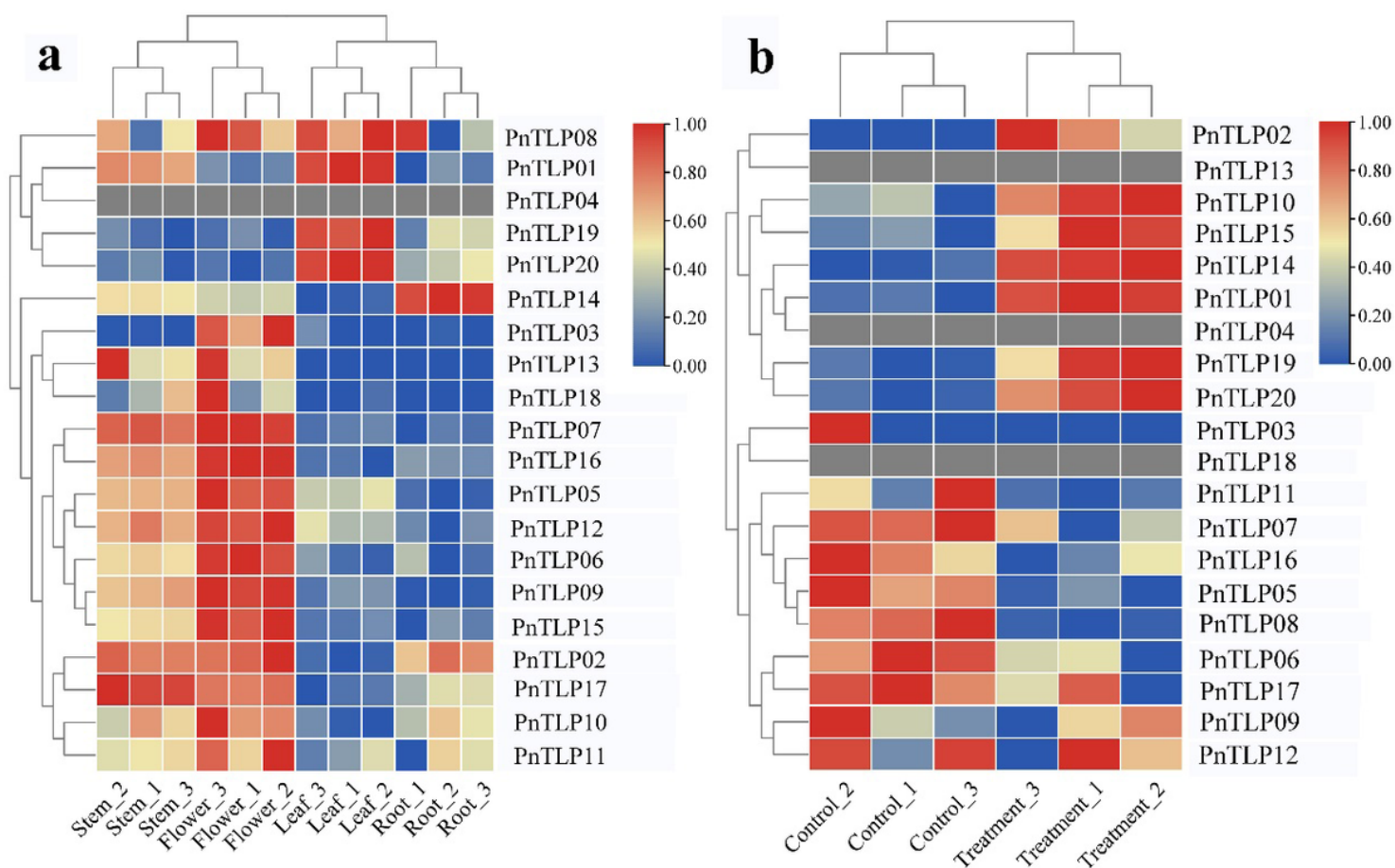


Figure 6

Expression profiles of *PnTLP* genes in different tissues (a) and infected by *A. panax* (b)

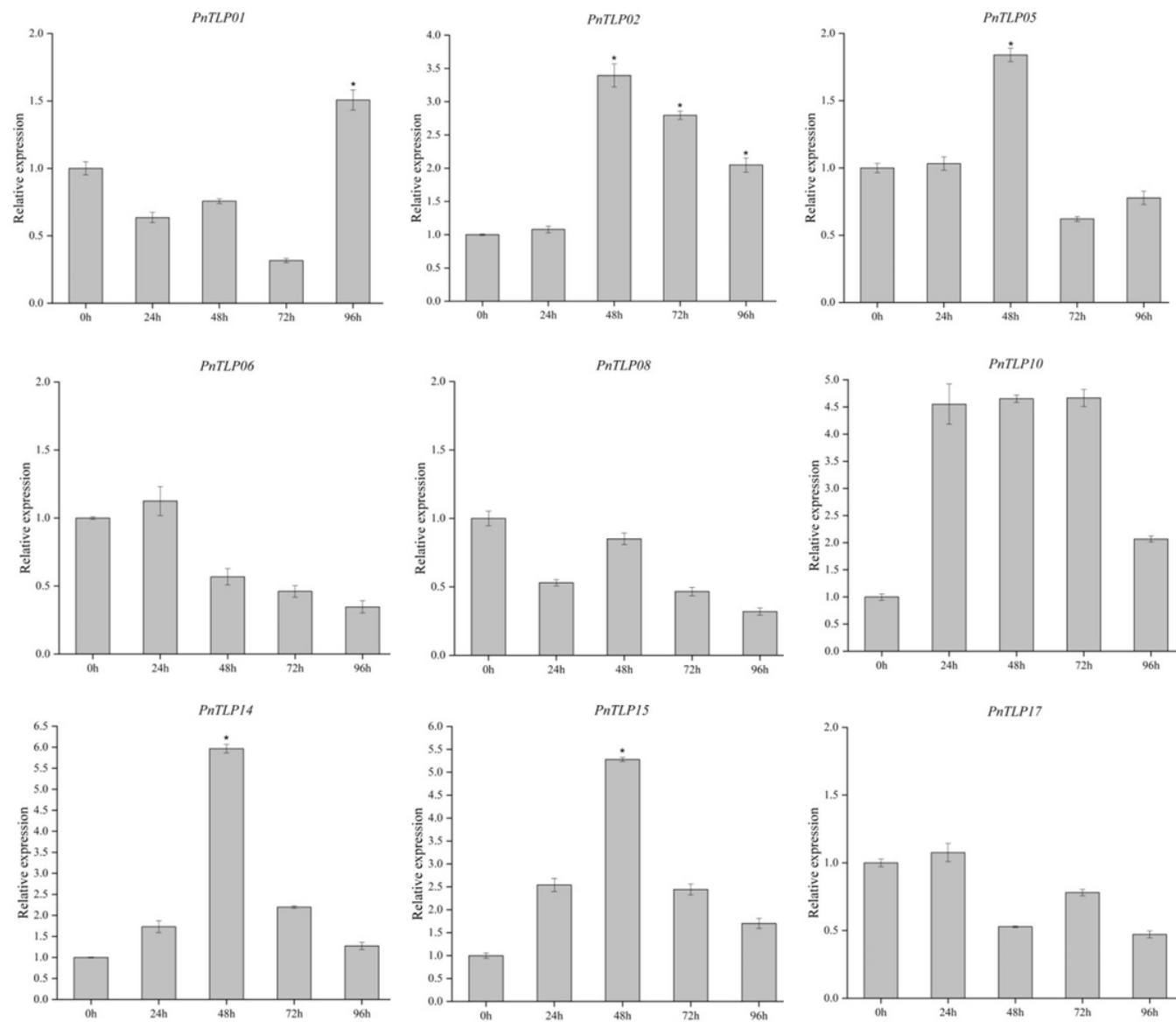


Figure 7

Gene expression of identified *PnTLP* after *F. solani* infection. Gray bars represented the relative expression of *TLP* in *P. notoginseng*. Relative expression was calculated using the comparative threshold  $2^{-\Delta\Delta CT}$  method, and the results are showed with the average values calculated from three replicates, and the statistical differences between the treatments and the control were analyzed by Student's t-test (\* $p < 0.05$ ).