GENOME-WIDE IDENTIFICATION AND EXPRESSION ANALYSIS OF *TLPS* GENE FAMILY UNDER ABIOTIC STRESSES IN *CAMELINA SATIVA* USING BIOINFORMATICS METHODS

LIU, Y. X.[1] – LIU, X.[1] – WU, B. X.[2] – PAN, Y. J.[1] – WU, L. K.[3] – LI, Y. Y.[1] – WANG, Z.[1] – PANG, H. B.[1*]

[1]*College of Life Sciences, Shenyang Normal University, Shenyang 110034, China*

[2]*College of Life Science, Liaoning Normal University, Dalian 116029, China*

[3]*College of Grain Science and Technology, Shenyang Normal University, Shenyang 110034, China*

[*]*Corresponding author*
e-mail: panghongbo800206@163.com; phone: +86-186-2400-6083

(Received 10th Jun 2023; accepted 16th Nov 2023)

Abstract. The membership of *Camelina sativa* in the Brassicaceae family has generated substantial interest in its potential as a high-value crop for food, feed, and fuel. Despite the critical regulatory roles played by Thaumatin-like proteins (*TLPs*) in plant growth and development, their presence in *C. sativa* has not been previously reported. Through a comprehensive analysis of the *C. sativa* genome, 34 *CsTLP* genes have been identified. The chromosomal localization and gene duplication of *CsTLPs* suggest that the expansion of the gene family was significantly influenced by segmental duplication, while the evolution of *CsTLPs* was primarily driven by purification selection. Synteny analysis indicates that *C. sativa* shares a greater number of homologous genes with *Arabidopsis thaliana*. Additionally, RNA-seq analysis suggests that numerous members of the *CsTLPs* gene family may potentially participate in abiotic stress. Significantly, the expression of the *CsTLP29* gene exhibited a remarkable decrease in response to four distinct stress conditions. The comprehensive examination of the *CsTLPs* gene family suggests that these genes may have a crucial function in the adaptation of *C. sativa* to stress and establishes a foundation for further exploration of the *TLPs* gene family in *C. sativa*.
Keywords: *thaumatin-like proteins, phylogenetic analysis, motif, Cis-acting element, RNA-seq*

Introduction

Throughout the growth and development phases, plants encounter a range of biotic and abiotic stressors. In response, organisms have developed diverse mechanisms to acclimate to environmental fluctuations (Hirayama and Shinozaki, 2010), including the production of specialized functional proteins, such as antioxidant proteins, late embryogenesis-abundant proteins (LEA), and pathogenesis-related proteins (PR). These proteins serve a crucial role in enhancing plant stress tolerance and bolstering plant disease resistance (Rubio et al., 2009; Chen et al., 2015; Lakhssassi et al., 2020). Thaumatin-like proteins (TLPs) are a member of the pathogenesis-related protein families (PR-5), consisting of a polypeptide chain of approximately 200 amino acid residues. TLPs have garnered significant attention due to their structural and functional similarities to the natural sweetener thaumatin (van Kan et al., 1989; Breiteneder, 2004).

The TLP protein sequence is characterized by a highly conserved motif, known as the thaumatin domain, which is comprised of G-X-[GF]-X-C-X-T-[GA]-D-C-X(1,2)-G-X-(2,3)-C (Jami et al., 2007; Tachi et al., 2009). Based on molecular weight and sequence homology, TLPs can be classified into two types, with the majority containing 16

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

conserved cysteine residues that can form eight disulfide bonds. The L-type structure of Thaumatin-like proteins (TLPs) serves a dual purpose of maintaining protein stability and aiding in plant resistance against extreme pH levels, proteolytic hydrolysis, and heat-induced denaturation (Ghosh and Chakrabarti, 2008). Conversely, the S-type TLPs, which lack certain peptide segments, contain only 10 conserved cysteine residues (Liu et al., 2015).

TLPs are widely distributed in various plant species, including rice (*Oryza sativa* L.), barley (*Hordeum vulgare* L.), melon (*Cucumis melo* L.), watermelon (*Citrullus lanatus* L.), European grape (*Vitis vinifera* L.), and bread wheat (*Triticum aestivum* L.) (Liu et al., 2010, 2020; Yan et al., 2017; Iqbal et al., 2020; Sharma et al., 2021; Ram et al., 2022). Extensive research has been conducted on *TLPs*, revealing their diverse biological functions in plants. *TLPs*, despite their initial association with sweetness, serve a crucial function in plant disease defense, growth and developmental regulation, and adaptation to adverse conditions. According to the findings of Naseri et al. (2012), the utilization of *TLPs* can enhance the resistance of *O. sativa* against rice blast disease. The heterologous expression of the *TLPs* gene in *A. thaliana* resulted in an increased resistance to *Pseudomonas syringae* and powdery mildew (Yan et al., 2017). Moreover, the purified BanTLPs protein in banana (*Musa acuminate* L.) demonstrated antifungal activity by inducing fungal cell membrane disruption and cell wall disassembly (Irigoyen et al., 2020). Additionally, the ectopic expression of the peanut (*Arachis hypogaea* L.) *TLPs* gene in tobacco plants has been found to enhance the tolerance of tobacco seedlings to both salt and oxidative stress (Singh et al., 2013). Brewer's yeast exhibiting elevated expression of the wheat *TaTLP2* gene demonstrated heightened resistance to cold, heat, osmotic, and salt stresses relative to control groups (Sharma et al., 2021). These findings indicate that the quantity of sweet-like protein family constituents may differ across plant species, as well as their capacity to govern adversity adaptation and pathogenic defense mechanisms. Consequently, exploring the structure and biology of sweet-like proteins in diverse species holds significant scientific merit in comprehending the structure and function of the sweet-like protein family and investigating adaptive mechanisms of plants to environmental stresses.

Camelina sativa, a plant species closely related to *A. thaliana* and *Brassica napus*, played a significant role in human diet during the Bronze Age (Acamovic et al., 1999). Presently, it has garnered scientific interest due to its potential as a novel oilseed crop, high disease resistance, broad growth potential, and nutritional value (Yuan et al., 2017; Song et al., 2020; Sun et al., 2022). Despite some research of the *LTPs* gene family in various species, its presence in *C. sativa* has yet to be reported. The investigation has successfully identified a total of 34 *CsLTPs* genes in *C. sativa* at the genome-wide level. A comprehensive analysis was conducted on their gene structure, chromosomal positioning, conserved motifs, phylogeny, gene duplication, and cis-element roles. Furthermore, the study investigated changes in their expression patterns under different adversity conditions. The results of this study establish a fundamental basis for future inquiries into the biological mechanisms underlying the *LTPs* gene family.

Materials and methods

Identification of members of the TLP gene family of Camelina sativa

Firstly, the *C. sativa* genome and gene structure annotation information file were obtained from the Ensemble plants database (http://plants.ensembl.org/). To determine

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

potential members of the *C. sativa* TLP family, the amino acid sequence of the *A. thaliana* TLP was retrieved from the TAIR database (http://www.arabidopsis.org/). Blast comparisons were performed using TBtools (v1.106) with a threshold set to 1e-5 (Chen et al., 2020). To validate the membership of these sequences in the TLP gene family, the NCBI Conserved Domain Database (CDD) (https://www.ncbi.nlm.nih.gov/cdd/) and the Pfam database (https://pfam.xfam.org/) were utilized. The physicochemical properties, including isoelectric point (pI) and relative molecular mass (Mw), of all CsTLP members were evaluated using the ExPASy website (http://web.expasy.org/). Lastly, the subcellular localization of the identified CsTLPs was predicted using the YLoc website (https://abi-services.informatik.uni-tuebingen.de/yloc/).

Multiple sequence alignment and phylogenetic analysis

Multiple sequence alignment and phylogenetic analysis were conducted using MEGA7.0 software. The amino acid sequences of CsTLPs in *C. sativa, A. thaliana*, and *O. sativa* were compared with the MUSCLE program. The neighbor-joining (NJ) method was then employed to construct the phylogenetic analysis while the Jones-Taylor-Thornton (JTT) and pairwise deletion models were selected, and the bootstrap value was set to 1000. Finally, an aesthetically pleasing phylogenetic tree was generated using the iTol software available at https://itol.embl.de/.

Gene structure, chromosome positioning and gene duplication analysis

The MEME tool was used to analyze conserved motifs of the *CsTLPs* gene, with a motif value set to 10 and default values set for the remaining parameters. The exon-intron organization of *CsTLPs* was visualized, and the chromosomal localization of its members was examined using the TBtools software. To investigate segmental duplication events in the genome of *C. sativa*, MCScanX, a Multiple Collinearity Scan toolkit, was employed. Additionally, MCScanX was used to analyze the synteny relations between *CsTLPs* and *TLPs* of rice and *A.thaliana*. Furthermore, TBtools software was utilized to examine the synonymous substitution (Ks) and non-synonymous substitution (Ka) rates among gene duplicates. The Ka/Ks ratio was calculated to determine the level of natural selection pressure. Finally, *c*is-acting elements and adversity expression profiles were analyzed.

Analysis of cis-acting elements and stress-responsive gene expression profiles

The 2000 bp sequence upstream of the start codon of *CsTLPs* genes was extracted using the TBtools software. This promoter sequence was subsequently analyzed and predictions were made using the PlantCARE web tool (http://bioinformatics.psb.ugent.be/webtools/plantcare/). Raw RNA-seq data for *C. sativa* under various abiotic stresses (cold, drought, salt, and heavy metal cadmium) were retrieved from the NCBI website (SRA) using the following accession numbers: PRJNA 210879-210882. The TBtools software was used to transform and quantify transcript expression, following which the DESeq2 plug-in was employed to analyze the count values for significant differentially expressed genes. A threshold of |log2fc|>=1 and padjust<=0.05 was applied for identification of differentially expressed genes. To evaluate the transcript levels of all members of the gene family, TMP (Transcripts per million) values were calculated and subsequently converted to Log2 normalized values

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

using the TBTools software. Heatmaps were generated using the normalized values to visualize gene expression profiles.

Results

Identification and analysis of the CsTLPs gene family

Through bioinformatics analysis, a total of 34 *TLPs* genes were identified within the *C.sativa* genome database. These genes have been designated as *CsTLP1-CsTLP34* in accordance with their chromosomal position and are presented in *Table S1* alongside their respective transcript IDs. Furthermore, *Table S1* provides information pertaining to the amino acid length, protein physicochemical properties (including relative molecular mass, isoelectric point, and subcellular localization) of each gene. The amino acid length of CsTLPs ranged from 183 to 524, exhibiting a broad range of variation with an average length of 384 amino acids. The CsTLPs exhibit a range of relative molecular masses, with CsTLP33 having the largest mass at 51,707.91 KD and CsTLP25 having the smallest mass at 23,356.80 KD. The isoelectric points of all CsTLPs fall within the basic range, with values ranging from 9.1 (CsTLP19) to 10.2 (CsTLP15). Subcellular localization analysis revealed that the majority of CsTLPs (18 members) are located in the cytoplasm, while 8 and 7 members are localized in the nucleus and endoplasmic reticulum, respectively. Only one member (CsTLP30) was located in the mitochondria.

Phylogenetic analysis of CsTLPs genes

In order to examine the evolutionary relationships within the *CsTLPs* family, a phylogenetic tree was constructed utilizing the amino acid sequences of three plant TLPs from *C. sativa*, *A. thaliana*, and *O. sativa* (*Fig. 1A*), encompassing a total of 34 CsTLPs, 14 OsTLPs, and 11 AtTLPs. These TLPs were subsequently categorized into five distinct groups, denoted as Group I through V. Notably, Group V exhibited the highest number of CsTLPs members, with a total of 12, representing nearly one-third of the CsTLPs. This was followed by Group II, which contained 9 CsTLPs, and Group III, which had the fewest members, with only 3 CsTLPs. Among the five subgroups, Groups II, III, IV, and V exhibit TLPs family members from all three plants, while Group I exclusively contains TLPs family members from *A. thaliana* and *C. sativa*. It is plausible that members of Group I diverged in monocotyledons, possibly due to novel gene duplication events in dicotyledons or the loss of duplication genes in monocotyledons. Furthermore, the *CsTLPs* gene was notably more abundant than in *O. sativa* and *A. thaliana*, indicating that the *TLP* gene family underwent multiple gene duplication events and had a significantly greater number of members in *C. sativa*. In order to conduct a more comprehensive examination of the homology between *C. sativa* and the *TLPs* gene families of *A. thaliana* and *O. sativa*, this study utilized TBtools to analyze the synteny relationship between these three species (*Fig. 1B*). The study's results indicate that there are 29 *CsTLPs* genes in *C. sativa* and 10 *AtTLPs* genes in *A. thaliana* that exhibit syntenic *TLPs* gene pairs. In contrast, only four syntenic *TLPs* gene pairs were identified in *C. sativa* and *O. sativa*.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

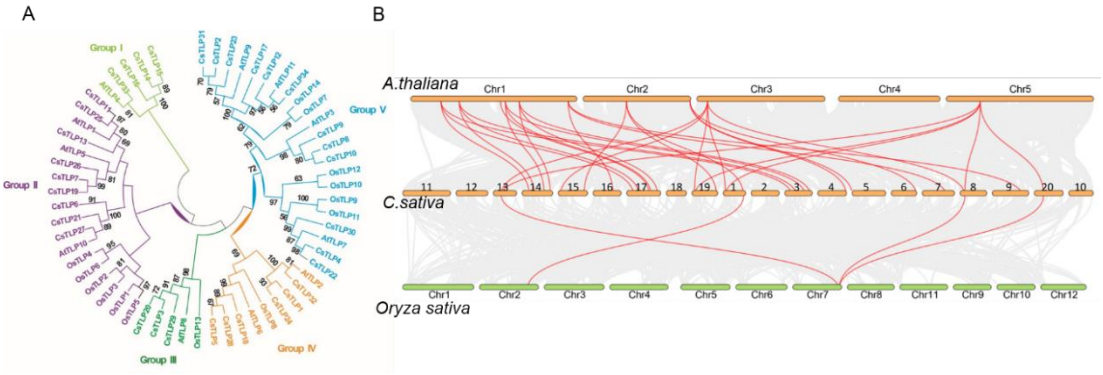

Figure 1. *Phylogenetics and synteny analysis of TLPs among C. sativa, A. thaliana and O. sativa*

Conserved motifs and gene structure of CsTLPs

The investigation of the structural diversity of *CsTLPs* genes involves an analysis of the exon-intron structure of *CsTLPs* and their distribution in the phylogenetic tree (*Fig. 2A*) to identify the potential impact of different combinations of exons and introns on gene function divergence. The family members exhibit a range of 2-8 exons and varying numbers of introns, with *CsTLP3*, *CsTLP14*, *CsTLP16*, *CsTLP20*, and *CsTLP29* having up to 8 introns, while *CsTLP33* has only 2 introns. The remaining members generally possess 3-5 introns. As an illustration, the genetic structure distribution characteristics of the three *CsTLPs* genes (*CsTLP3*, *CsTLP20*, and *CsTLP29*) belonging to group III exhibit a high degree of similarity, thereby reinforcing the credibility of the phylogenetic grouping outcomes and indicating that genes within the same group possess comparable structures.

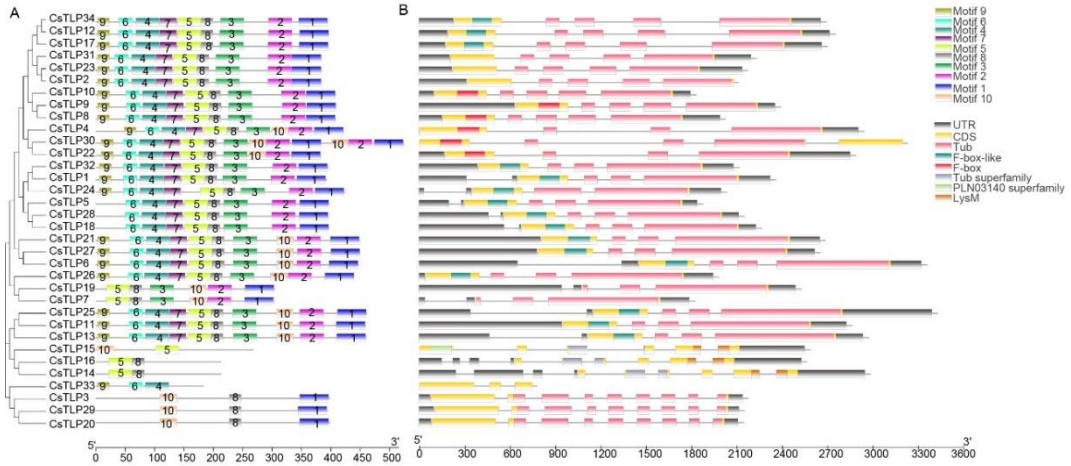

Figure 2. *Phylogenetic tree, conserved protein motifs analysis, and gene structure of TLPs in C. sativa. A. The Neighbor joining method was utilized to construct the phylogenetic tree of C. sativa, with bootstrap values set to 1000. The Conserved protein motifs of CsTLP were also incorporated, with each motif being represented by a distinct color. B. Exon-intron structure and functional domains of the TLPS gene (yellow boxes represent exons, black lines represent introns, and specific colors represent different functional domains)*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

Protein function is significantly influenced by protein structural domains, particularly those containing highly conserved amino acid motifs. In the present study, the conserved motifs of CsTLPs were subjected to analysis using the MEME online website, resulting in the identification of a total of 2-10 conserved motifs (*Fig. 2A*), designated as Motif 1-Motif 10. The majority of these motifs exhibited a similar distribution pattern, namely Motif 9-6-4-7-5-8-3-10-2-1, indicating a relatively conserved protein structure of CsTLPs. The protein structure of the *TLPs* gene family was analyzed quantitatively, revealing that 10 members (CsTLP4, 6, 11, 13, 21, 22, 25, 26, 27, 30) possess all 10 conserved motifs and are evenly distributed. Conversely, 12 members (CsTLP1, 2, 8, 9, 10, 12, 17, 23, 24, 31, 32, 34) lack only one Motif 10, suggesting that Motifs 1-9 are relatively conserved and potentially crucial to the function of the TLPs gene family. In Group IV, CsTLP5, CsTLP18, and CsTLP28 were found to lack Motif 9 and Motif 10. Similarly, CsTLP7 and CsTLP19 in Group II were observed to lack Motif 4, Motif 6, Motif 7, and Motif 9. In contrast, CsTLP14, 15, and 16 in Group I exhibited the lowest number of Motifs, with only two present. Despite the variation in motif types across groups, members within the same group displayed comparable motif patterns, as evidenced by CsTLP3, CsTLP19, and CsTLP20, indicating a high degree of homology among these genes. It is plausible that the proteins encoded by these genes may also possess similar functions.

The examination of the structural domains of CsTLPs through the utilization of the NCBI website CDD-search demonstrated that all protein sequences, with the exception of CsTLP33, exhibited the conventional Tub structural domains. Notably, certain proteins (CsTLP2, CsTLP7, CsTLP19, CsTLP31, CsTLPs32 and CsTLP33) were devoid of the F-box structural domains (*Fig. 2B*). Despite the presence of comparable structural domains in the majority of TLPs proteins, variations in number and type exist among members. For instance, CsTLPs14,15 exhibit LysM structural domains, suggesting that while CsTLPs share common functions, they may also demonstrate functional preference and redundancy across different members.

Chromosomal localization and gene duplication analysis of CsTLPs genes

The present study utilized chromosomal localization and synteny analysis to enhance the visualization of the *CsTLPs* gene family distribution across chromosomes and to investigate the impact of gene duplication on the amplification of the *CsTLPs* genes (*Fig. 3*). The distribution of the 34 *CsTLPs* genes was found to be scattered and uneven across 15 of the 20 chromosomes. Chromosomes 3, 14, and 17 exhibited the highest number of genes, each containing five, followed by chromosome 9 with four. Chromosomes 1, 15, 19, and 20 contained two genes each, while chromosomes 4-8, 13, and 16 contained only one gene each. Notably, five chromosomes (2, 10, 11, 12, and 18) did not contain any *CsTLPs* genes. Interestingly, *CsTLP14* and *CsTLP16* were found to occupy the same position, differing only in the untranslated region, suggesting a possible amplification by tandem repeats. Furthermore, the identification of 39 pairs of segmental duplicates (*Fig. 3*) was accompanied by the calculation of Ka, Ks, and Ka/Ks values to examine the evolutionary functional constraints in *C.sativa* (*Table S2*). The range of Ka values for each gene pair was 0-0.173, while Ks values ranged from 0.03-1.17. Notably, all duplicate gene pairs exhibited Ka/Ks values below 1, indicating that purifying selection was the primary force driving the evolution of *CsTLPs* in *C.sativa*.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

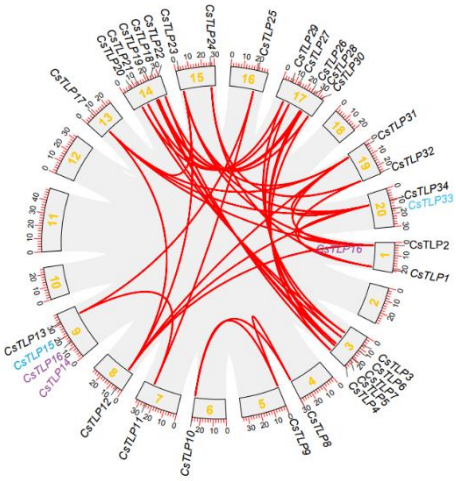

Figure 3. *Chromosomal location and duplicated genes among 34 CsTLPs genes*

Cis-acting elements and expression pattern of CsLTPS genes under abiotic stress

The promoter region of the *CsTLPs* gene family is replete with numerous *cis*-acting elements that are linked to hormones, light signalling, stress, growth, and development, as depicted in *Fig. 4*. Specifically, 25, 24, 17, 10, 9, and 6 members of the gene family possess cis-acting elements that are associated with methyl jasmonate (MeJA), salicylic acid (SA), gibberellic acid (GA), abscisic acid (ABA), and growth hormone (Auxin, IAA), respectively. Notably, each member of the gene family contains at least one phytohormone-related element, with some members having multiple *cis*-acting elements concurrently. The study revealed that the highest number of elements associated with methyl jasmonate (MeJA) was 72, followed by abscisic acid (ABA) with 47 elements. Furthermore, all members exhibited light response-related elements, with *CsTLP17* exhibiting the greatest number of such elements at 23. Additionally, the promoters of members 33, 13, 11, 16, 14, and 13 contained response elements for adversity stress, including anaerobic induction (ARE), defense and stress (TC-rich repeats), drought-induced related MYB binding site (MBS), low temperature (LTR), and injury (WUN-motif). Furthermore, the promoters encompass a diverse array of components associated with the growth and development of plants.

In order to gain a deeper comprehension of the potential roles of *CsTLPs* genes, we conducted an analysis of the alterations in the expression of 34 *CsTLPs* genes subjected to stress. This analysis was based on RNA-seq data previously obtained from diverse adversity conditions, including cold, drought, salt, and heavy metal Cd. Our findings indicate that the expression of all 33 *CsTLPs* genes exhibited some degree of change in response to the four adversity stresses, with the exception of *CsTLP33*, which displayed no expression (*Fig. 5A*). Following exposure to cold stress, the expression of *CsTLP20* and *CsTLP29* was notably down-regulated, whereas the expression of *CsTLP12*, *CsTLP21*, *CsTLP31*, and *CsTLP34* was significantly up-regulated. Similarly, under drought stress conditions, the expression of *CsTLP3*, *CsTLP20*, and *CsTLP29* exhibited a significant downward trend, while the expression of *CsTLP19*, *CsTLP21*, *CsTLP26*, and *CsTLP28* displayed a significant increase in expression after stress. Under salt stress, the expression of *CsTLP3*, *CsTLP20*, and *CsTLP29* was significantly reduced, while the expression of *CsTLP19* and *CsTLP26* was significantly increased. Conversely, under

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

heavy metal Cd stress, only the expression of *CsTLP29* was significantly down-regulated. Notably, *CsTLP3*, *CsTLP20*, and *CsTLP29* were the genes that exhibited significant down-regulation. *CsTLP29* was down-regulated significantly under all four stresses, whereas *CsTLP20* was down-regulated significantly under cold, drought, and salt stress. Finally, *CsTLP3* was down-regulated significantly under drought and salt stress, as depicted in *Fig. 5B.*

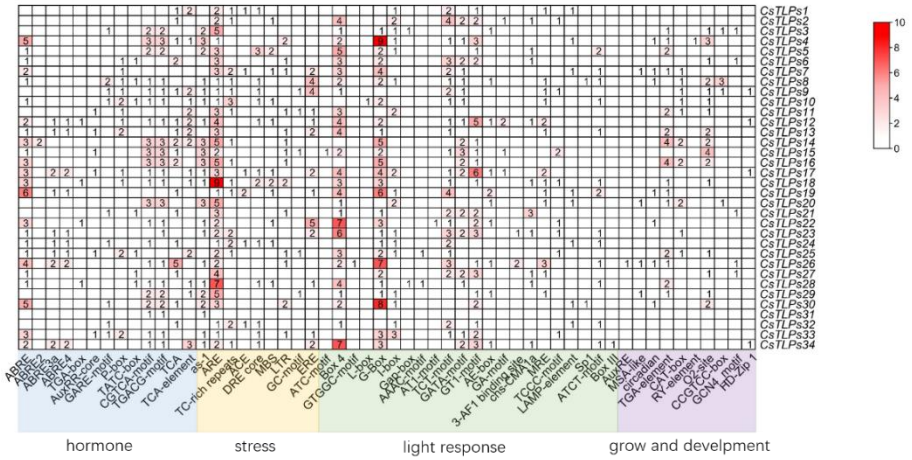

Figure 4. *Cis-elements in putative promoter regions of TLPs gene in C. sativa*

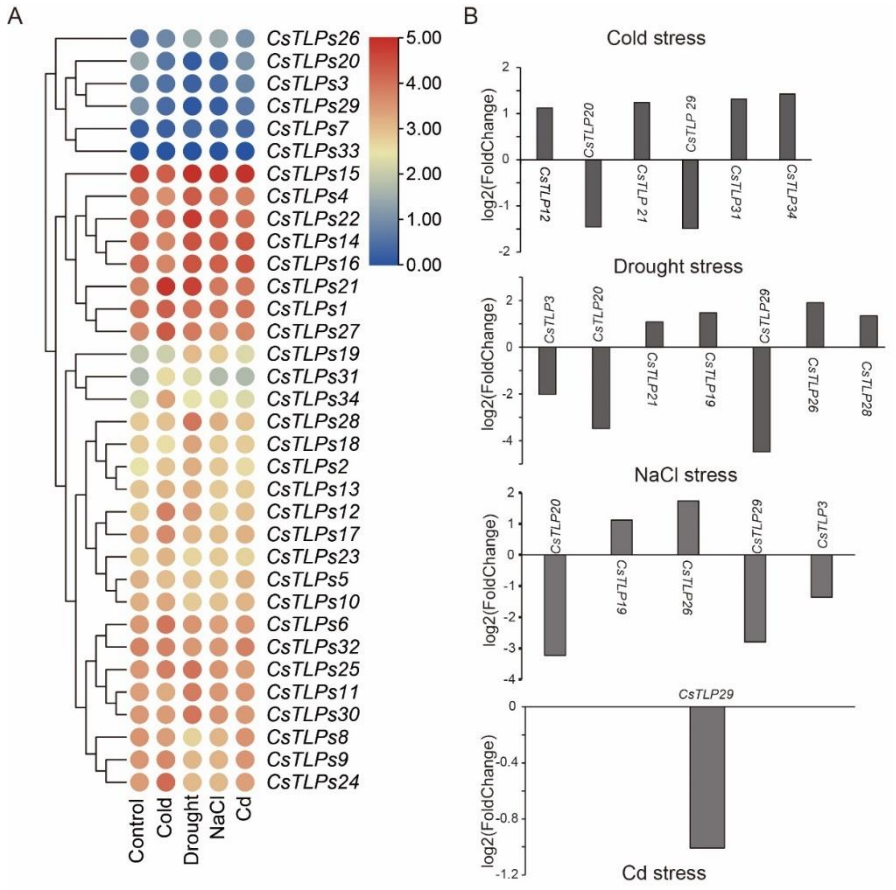

Figure 5. *The expression patterns of CsTLPs gene family under the abiotic stresses*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

Discussion

Camelina sativa is a significant crop that finds extensive application in textiles, construction materials, pharmaceuticals, and various other domains, and it ranks among the most extensive agricultural cultivars globally (Shonnard et al., 2010; Berti et al., 2016). A salient feature of *Camelina sativa* is its robust stress resistance. Plants encounter diverse abiotic and biotic stresses, such as drought, high and low temperatures, salinity, and heavy metals, under varying environmental and climatic conditions, which can impede their growth, development, yield, and quality. Hence, an imperative aspect is to investigate the biological attributes and adaptive strategies of plants when subjected to stress conditions. In the case of *C. sativa*, a plant species that exhibits remarkable stress resistance, an in-depth analysis of its stress tolerance mechanisms can serve as a benchmark for breeding and stress adaptation in other economically significant crops.

TLPs, a class of gene families prevalent in eukaryotes, have been extensively researched in various plant species. *A. thaliana*, *O. sativa*, and wheat (*Triticum aestivum* L.) exhibit 11, 14, and 23 members of the *TLPs* gene family, respectively (Liu et al., 2010; Sharma et al., 2021). Notably, soybean and walnut genomes exhibit the highest number of *TLP* genes, with 62 and 66 members, respectively, which is twice the number observed in *A. thaliana*, this observation can be attributed to the modern soybean being a diploid species that evolved from ancient tetraploids and the high-quality assembled genome of walnuts (Schmutz et al., 2010; Marrano et al., 2020). The variation in the number of *CsTLPs* gene family members across plant species is attributed to gene duplication, which is considered a primary mechanism for the expansion and diversification of gene families during biological evolution (Lynch and Conery, 2000). Gene duplication encompasses various types, such as whole genome duplication, tandem duplication, and segmental duplications. The extent and nature of gene duplication differ across gene families, which is indicative of the heterogeneity and fluctuation of gene families with respect to the evolutionary trajectories of distinct organisms. Moreover, gene duplication constitutes a significant impetus for the ongoing augmentation and diversification of gene families. Segmental duplication is a frequently observed mechanism of gene duplication that transpires in the intergenic region, wherein one gene's fragments are replicated into the other gene. In contrast to whole genome duplication and tandem duplication, segmental duplication is a more intricate process that entails diverse genetic regulation and regulatory pathways. It is frequently linked with genetic events, such as recombination and translocation, which augment genetic diversity and variation, thereby promoting the necessity for rapid evolution and adaptation of species to varying environmental pressures (Flagel and Wendel, 2009). The expansion of the *CsTLPs* gene family is primarily attributed to segmental replication, whereas in wheat, whole genome duplication is the predominant mechanism, and in *O. sativa*, tandem replication is the primary mode. Additionally, our analysis indicates that purifying selection has played a significant role in shaping the evolution of the *CsTLPs* gene family. In conclusion, the *TLPs* gene family exhibits variations in terms of its size and composition across different plant species and serves as a crucial contributor to plant genetic and functional diversity.

The *TLPs* gene family exhibits conserved protein structural domains and shares material properties akin to glycoproteins and cyclodextrins, which are commonly classified as sweet-tasting proteins in plants and are known to serve diverse biological functions, notably in fungal resistance and stress response. In a study conducted by Sharma et al. (2021), it was observed that the overexpression of the *TaTLP2-B* gene in yeast cells resulted in a significant improvement in their ability to withstand various

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

abiotic stress conditions. The aforementioned findings align with prior research indicating that the overexpression of *TLP* genes can enhance tolerance to diverse abiotic stressors in tobacco (*Nicotiana tabacum* L.), cotton (*Gossypium hirsutum* L.), *A. thaliana*, and cauliflower (*Brassica oleracea* L.) (Rajam et al., 2007; Misra et al., 2016; Li et al., 2020; He et al., 2021). Li et al. (2020) have demonstrated that the *GhTLP19* gene in cotton exerts a favorable regulatory influence on drought tolerance. Furthermore, *GbTLP1* in sea island cotton and *ObTLP1* in *Ocimum basilicum* have been identified as efficacious in mitigating the effects of drought and salt stresses (Munis et al., 2010; Misra et al., 2016). The examination of cis-acting elements has demonstrated that *CsTLPs* present in *C. sativa* possess numerous hormonal regulatory elements, including the TGACG motif that participates in the MeJA response and the TCA element that is involved in the SA response. Plant hormones, particularly SA and JA, play a crucial role in the molecular signaling of plant defense responses against pathogen attacks (Pieterse and van Loon, 1999; Park and Kim, 2021), whereas the ABA signaling pathway is fundamental in regulating abiotic stresses such as drought and salt. The findings of a recent study indicate that thaumatin mutant plants experienced modifications in the ABA signalling pathway, resulting in heightened vulnerability to abiotic stresses (Reymond and Farmer, 1998). Several stress-related elements were identified in *CsTLPs* from *C.sativa*, including ABRE elements associated with ABA signaling. The expression patterns of *TLPs* family members were also investigated under diverse stress conditions using available transcriptome data. The results indicated that the expression of all family members, with the exception of *CsTLP33*, was modified, particularly *CsTLP29*, which was significantly down-regulated under all four stress conditions.

Conclusion

In the *Camelina sativa* genome, 34 members of the *CsTLPs* gene family were identified, which were distributed across 15 chromosomes. Segmental duplication was found to be the primary mechanism responsible for the amplification of *TLPs* genes. The promoter region of these genes contained numerous hormone and stress response elements, as well as transcription factor binding sites. Furthermore, distinct members of the *CsTLPs* gene family exhibited varying expression patterns in response to low temperature, drought, salt, and heavy metal Cd stresses.

Funding. This work was financially supported by Liaoning Provincial Natural Science Foundation (2022-MS-309), Major Project Incubation Project of Shenyang Normal University (ZD202104) and Shenyang Normal University "Hundred Talents Programme" Top Talents Project.

REFERENCES

[1] Acamovic, T., Gilbert, C., Lamb, K., Walker, K. C. (1999): Nutritive value of *Camelina sativa* meal for poultry. – British Poultry Science 40: S27.
[2] Berti, M., Gesch, R., Eynck, C., Anderson, J., Cermak, S. (2016): Camelina uses, genetics, genomics, production, and management. – Industrial Crops and Products 94: 690-710.
[3] Breiteneder, H. (2004): Thaumatin-like proteins - a new family of pollen and fruit allergens. – Allergy 59: 479-81.
[4] Chen, Y. S., Lo, S. F., Sun, P. K., Lu, C. A., Ho, T. H., Yu, S. M. (2015): A late embryogenesis abundant protein HVA1 regulated by an inducible promoter enhances root

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

growth and abiotic stress tolerance in rice without yield penalty. – Plant Biotechnology Journal 13: 105-116.

[5] Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., Xia, R. (2020): TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. – Molecular Plant 13: 1194-1202.

[6] Flagel, L. E., Wendel, J. F. (2009): Gene duplication and evolutionary novelty in plants. – New Phytologist 183: 557-564.

[7] Ghosh, R., Chakrabarti, C. J. P. (2008): Crystal structure analysis of NP24-I: a thaumatin-like protein. – Planta 228: 883-890.

[8] He, L., Li, L., Zhu, Y., Pan, Y., Zhang, X., Han, X., Li, M., Chen, C., Li, H., Wang, C. (2021): *BolTLP1*, a Thaumatin-like Protein Gene, Confers Tolerance to Salt and Drought Stresses in Broccoli (*Brassica oleracea* L. var. *Italica*). – International Journal of Molecular Sciences 22(20): 11132.

[9] Hirayama, T., Shinozaki, K. (2010): Research on plant abiotic stress responses in the post-genome era: past, present, and future. – Plant Journal 61: 1041-52.

[10] Iqbal, I., Tripathi, R. K., Wilkins, O., Singh, J. (2020): Thaumatin-Like Protein (*TLP*) Gene Family in Barley: Genome-Wide Exploration and Expression Analysis during Germination. – Genes (Basel) 11(9): 1080.

[11] Irigoyen, M. L., Garceau, D. C., Bohorquez-Chaux, A., Lopez-Lavalle, L. A. B., Perez-Fons, L., Fraser, P. D., Walling, L. L. (2020): Genome-wide analyses of cassava Pathogenesis-related (PR) gene families reveal core transcriptome responses to whitefly infestation, salicylic acid and jasmonic acid. – BMC Genomics 21: 1-18.

[12] Jami, S. K., Anuradha, T. S., Guruprasad, L., Kirti, P. B. (2007): Molecular, biochemical and structural characterization of osmotin-like protein from black nightshade (*Solanum nigrum*). – Journal of Plant Physiology 164: 238-252.

[13] Lakhssassi, N., Piya, S., Bekal, S., Liu, S., Zhou, Z., Bergounioux, C., Miao, L., Meksem, J., Lakhssassi, A., Jones, K., Kassem, M. A., Benhamed, M., Bendahmane, A., Lambert, K., Boualem, A., Hewezi, T., Meksem, K. (2020): A pathogenesis-related protein GmPR08-Bet VI promotes a molecular interaction between the GmSHMT08 and GmSNAP18 in resistance to Heterodera glycines. – Plant Biotechnol Journal 18: 1810-1829.

[14] Li, Z., Wang, X., Cui, Y., Qiao, K., Zhu, L., Fan, S., Ma, Q. (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. – Frontiers in Plant Science 11: 575015.

[15] Liu, J.-J., Sturrock, R., Ekramoddoullah, A. K. (2010): The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. – Plant Cell Reports 29: 419-436.

[16] Liu, H., Liu, J., Zhao, M.-M., Chen, J. (2015): Overexpression of *ShCHLP* in tomato improves seedling growth and increases tolerance to salt, osmotic, and oxidative stresses. – Journal of Plant Growth Regulation 77: 211-221.

[17] Liu, Y., Cui, J., Zhou, X., Luan, Y., Luan, F. (2020): Genome-wide identification, characterization, and expression analysis of the *TLP* gene family in melon (*Cucumis melon* L.). – Genomics 112: 2499-2509.

[18] Lynch, M., Conery, J. S. (2000): The Evolutionary Fate and Consequences of Duplicate Genes. – Science 290: 1151-1155.

[19] Marrano, A., Britton, M., Zaini, P. A., Zimin, A. V., Workman, R. E., Puiu, D., Bianco, L., Pierro, E. A. D., Allen, B. J., Chakraborty, S., Troggio, M., Leslie, C. A., Timp, W., Dandekar, A., Salzberg, S. L., Neale, D. B. (2020): High-quality chromosome-scale assembly of the walnut (*Juglans regia* L.) reference genome. – GigaScience 9(5): giaa050.

[20] Misra, R. C., Sandeep, Kamthan, M., Kumar, S., Ghosh, S. (2016): A thaumatin-like protein of *Ocimum basilicum* confers tolerance to fungal pathogen and abiotic stress in transgenic Arabidopsis. – Scientific. Reports 6: 25340.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

[21] Munis, M. F., Tu, L., Deng, F., Tan, J., Xu, L., Xu, S., Long, L., Zhang, X. (2010): A thaumatin-like protein gene involved in cotton fiber secondary cell wall development enhances resistance against *Verticillium dahliae* and other stresses in transgenic tobacco. – Biochemical and Biophysical Research Communications 393: 38-44.
[22] Naseri, G., Sohani, M. M., Pourmassalehgou, A., Allahi, S. (2012): In planta transformation of rice (*Oryza sativa*) using thaumatin-like protein gene for enhancing resistance to sheath blight. – African Journal of Biotechnology 11: 7885-7893.
[23] Park, E. J., Kim, T. H. (2021): Thaumatin-like genes function in the control of both biotic stress signaling and ABA signaling pathways. – Biochemical and Biophysical Research Communications 567: 17-21.
[24] Pieterse, C. M., van Loon, L. C. (1999): Salicylic acid-independent plant defence pathways. – Trends in Plant Science 4: 52-58.
[25] Rajam, M. V., Chandola, N., Saiprasad Goud, P., Singh, D., Kashyap, V., Choudhary, M. L., Sihachakr, D. (2007): Thaumatin gene confers resistance to fungal pathogens as well as tolerance to abiotic stresses in transgenic tobacco plants. – Biology Plantarum 51: 135-141.
[26] Ram, C., Danish, S., Kesawat, M. S., Panwar, B. S., Verma, M., Arya, L., Yadav, S., Sharma, V. (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.
[27] Reymond, P., Farmer, E. E. (1998): Jasmonate and salicylate as global signals for defense gene expression. – Current Opinion in Plant Biology 1: 404-11.
[28] Rubio, M. C., Bustos-Sanmamed, P., Clemente, M. R., Becana, M. (2009): Effects of salt stress on the expression of antioxidant genes and proteins in the model legume *Lotus japonicus*. – New Phytologist 181: 851-859.
[29] Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D. L., Song, Q., Thelen, J. J., Cheng, J., Xu, D., Hellsten, U., May, G. D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M. K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.-C., Shinozaki, K., Nguyen, H. T., Wing, R. A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R. C., Jackson, S. A. (2010): Genome sequence of the palaeopolyploid soybean. – Nature 463: 178-183.
[30] Sharma, A., Sharma, H., Rajput, R., Pandey, A., Upadhyay, S. K. (2021): Molecular Characterization Revealed the Role of Thaumatin-Like Proteins of Bread Wheat in Stress Response. – Frontiers in Plant Science 12: 807448.
[31] Shonnard, D. R., Williams, L., Kalnes, T. N., Energy, S. (2010): Camelina-derived jet fuel and diesel: Sustainable advanced biofuels. – Environmental Progress and Sustainable Energy 29(3): 382-392.
[32] Singh, N. K., Kumar, K. R. R., Kumar, D., Shukla, P., Kirti, P. B. (2013): Characterization of a pathogen induced thaumatin-like protein gene AdTLP from Arachis diogoi, a wild peanut. – PLoS One 8: e83963.
[33] Song, Y., Cui, H., Shi, Y., Xue, J., Ji, C., Zhang, C., Yuan, L., Li, R. (2020): Genome-wide identification and functional characterization of the *Camelina sativa* WRKY gene family in response to abiotic stress. – BMC Genomics 21: 786.
[34] Sun, D., Quan, W., Wang, D., Cui, J., Wang, T., Lin, M., Wang, Y., Wang, N., Dong, Y., Li, X., Liu, W., Wang, F. (2022): Genome-Wide Identification and Expression Analysis of Fatty Acid Desaturase (*FAD*) Genes in *Camelina sativa* (L.) Crantz. – International Journal of Molecular Sciences 23(23): 14550.
[35] Tachi, H., Fukuda-Yamada, K., Kojima, T., Shiraiwa, M., Takahara, H. (2009): Molecular characterization of a novel soybean gene encoding a neutral PR-5 protein induced by high-salt stress. – Plant Physiology and Biochemistry 47: 73-79.
[36] van Kan, J. A., van de Rhee, M. D., Zuidema, D., Cornelissen, B. (1989): Structure of tobacco genes encoding thaumatin-like proteins. – Plant Molecular Biology 12: 153-155.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

[37] Yan, X., Qiao, H., Zhang, X., Guo, C., Wang, M., Wang, Y., Wang, X. (2017): Analysis of the grape (*Vitis vinifera* L.) thaumatin-like protein (*TLP*) gene family and demonstration that *TLP29* contributes to disease resistance. – Scientific Reports 7: 4269.
[38] Yuan, L., Mao, X., Zhao, K., Ji, X., Ji, C., Xue, J., Li, R. (2017): Characterisation of phospholipid: diacylglycerol acyltransferases (PDATs) from *Camelina sativa* and their roles in stress responses. – Biology Open 6: 1024-1034.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

APPENDIX

Table S1. *The physicochemical characteristics of the 34 CsTLPs gene family members in C. sativa*

Gene name	Transcripts ID	Amino acid Length	Mr	PI	Localization	Group	Position of chromosome	Gene Localization
CsTLPs1	transcript:Csa01g041410.1	393.00	43626.01	9.36	Cytoplasm	4	1	21,149,754-21,152,078bp
CsTLPs2	transcript:Csa01g007400.1	384.00	42872.20	9.66	Cytoplasm	5	1	2,332,503-2,334,581bp
CsTLPs3	transcript:Csa03g019880.1	398.00	44340.74	9.60	Cytoplasm	3	3	7,271,964-7,274,105bp
CsTLPs4	transcript:Csa03g060570.1	422.00	46904.27	9.48	Chloroplast	5	3	26,918,242-26,921,150bp
CsTLPs5	transcript:Csa03g050980.1	397.00	44286.19	9.72	Chloroplast	4	3	21,429,472-21,431,314bp
CsTLPs6	transcript:Csa03g029880.1	447.00	50189.30	9.54	Nucleus	2	3	12,225,890-12,229,214bp
CsTLPs7	transcript:Csa03g046620.1	303.00	33391.07	9.36	Chloroplast	2	3	19,366,580-19,368,370bp
CsTLPs8	transcript:Csa04g067970.1	408.00	45355.88	9.49	Cytoplasm	5	4	30,031,578-30,033,570bp
CsTLPs9	transcript:Csa05g001290.1	409.00	45430.93	9.48	Cytoplasm	5	5	137,125-139,482bp
CsTLPs10	transcript:Csa06g054410.1	408.00	45263.74	9.48	Cytoplasm	5	6	26,332,630-26,334,427bp
CsTLPs11	transcript:Csa07g048470.1	459.00	51472.67	9.38	Nucleus	2	7	24,972,277-24,975,107bp
CsTLPs12	transcript:Csa08g011240.1	397.00	43687.01	9.73	Cytoplasm	5	8	4,810,976-4,813,693bp
CsTLPs13	transcript:Csa09g082610.1	460.00	51632.83	9.35	Nucleus	2	9	31,038,552-31,041,490bp
CsTLPs14	transcript:Csa09g040900.1	213.00	23356.80	9.84	Cytoplasm	1	9	15,134,637-15,137,588bp
CsTLPs15	transcript:Csa09g051950.1	268.00	29320.62	10.20	Nucleus	1	9	20,020,636-20,023,187bp
CsTLPs16	transcript:Csa09g040900.2	213.00	23356.80	9.84	Cytoplasm	1	9	15,134,637-15,137,588bp
CsTLPs17	transcript:Csa13g021600.1	396.00	43604.02	9.64	Cytoplasm	5	13	8,170,317-8,172,982bp
CsTLPs18	transcript:Csa14g051130.1	397.00	44255.22	9.72	Chloroplast	4	14	23,282,626-23,284,857bp
CsTLPs19	transcript:Csa14g048860.1	304.00	33593.27	9.10	Chloroplast	2	14	21,677,418-21,679,910bp
CsTLPs20	transcript:Csa14g020200.1	398.00	44314.60	9.62	Cytoplasm	3	14	7,658,927-7,661,038bp
CsTLPs21	transcript:Csa14g034200.1	449.00	50331.50	9.58	Nucleus	2	14	13,141,731-13,144,384bp
CsTLPs22	transcript:Csa14g063860.1	383.00	42417.18	9.57	Chloroplast	5	14	29,754,278-29,757,131bp
CsTLPs23	transcript:Csa15g007940.1	384.00	42685.00	9.66	Cytoplasm	5	15	2,505,555-2,507,693bp
CsTLPs24	transcript:Csa15g076840.1	424.00	47304.54	9.24	Cytoplasm	4	15	27,108,989-27,110,989bp
CsTLPs25	transcript:Csa16g041080.1	461.00	51707.91	9.28	Nucleus	2	16	21,017,870-21,021,264bp
CsTLPs26	transcript:Csa17g069490.1	440.00	48956.26	9.65	Nucleus	2	17	24,065,982-24,067,929bp
CsTLPs27	transcript:Csa17g035280.1	450.00	50434.58	9.57	Nucleus	2	17	13,049,061-13,051,680bp
CsTLPs28	transcript:Csa17g072750.1	396.00	44270.21	9.72	Chloroplast	4	17	25,850,671-25,852,787bp
CsTLPs29	transcript:Csa17g021660.1	395.00	44279.67	9.60	Cytoplasm	3	17	7,414,545-7,416,663bp

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263

Gene name	Transcripts ID	Amino acid Length	Mr	PI	Localization	Group	Position of chromosome	Gene Localization
CsTLPs30	transcript:Csa17g093660.1	524.00	57952.88	9.37	Mitochondrion	5	17	32,700,140-32,703,336bp
CsTLPs31	transcript:Csa19g010040.1	384.00	42866.20	9.66	Cytoplasm	5	19	2,640,249-2,642,446bp
CsTLPs32	transcript:Csa19g057880.1	396.00	43993.46	9.41	Cytoplasm	4	19	25,785,982-25,788,063bp
CsTLPs33	transcript:Csa20g046380.1	183.00	21296.58	9.53	Cytoplasm	1	20	16,408,699-16,409,443bp

Table S2. *Synonymous and nonsynonymous substitutions of CsTLPs gene pairs in C. sativa*

Seq_1	Seq_2	Ka	Ks	Ka_Ks	Selection pressure
CsTLP21	*CsTLP27*	9.70E-04	0.099007165	0.009799755	Purifying selection
CsTLP18	*CsTLP28*	0.002203858	0.084218402	0.02616837	Purifying selection
CsTLP27	*CsTLP6*	0.002925406	0.103433852	0.028282868	Purifying selection
CsTLP21	*CsTLP6*	0.001948053	0.060435365	0.032233661	Purifying selection
CsTLP28	*CsTLP5*	0.003308829	0.096409817	0.034320456	Purifying selection
CsTLP18	*CsTLP5*	0.003300941	0.075870067	0.043507812	Purifying selection
CsTLP9	*CsTLP10*	0.004308036	0.076198407	0.056537081	Purifying selection
CsTLP34	*CsTLP12*	0.007916387	0.135664664	0.058352607	Purifying selection
CsTLP8	*CsTLP9*	0.003224945	0.042754228	0.075429838	Purifying selection
CsTLP1	*CsTLP24*	0.006717716	0.083366487	0.080580535	Purifying selection
CsTLP11	*CsTLP13*	0.009571034	0.117229738	0.081643399	Purifying selection
CsTLP17	*CsTLP34*	0.010196532	0.119209449	0.085534593	Purifying selection
CsTLP23	*CsTLP31*	0.010422866	0.120023688	0.086840075	Purifying selection
CsTLP2	*CsTLP31*	0.006927789	0.07934413	0.08731319	Purifying selection
CsTLP25	*CsTLP11*	0.009568745	0.095866213	0.099813526	Purifying selection
CsTLP1	*CsTLP32*	0.007852741	0.075086098	0.104583151	Purifying selection
CsTLP25	*CsTLP13*	0.010542463	0.099072088	0.106412035	Purifying selection
CsTLP20	*CsTLP3*	0.006591638	0.06016378	0.10956157	Purifying selection
CsTLP24	*CsTLP32*	0.008974586	0.079244234	0.113252226	Purifying selection
CsTLP8	*CsTLP10*	0.006469624	0.053674163	0.120535164	Purifying selection
CsTLP31	*CsTLP34*	0.154091598	1.098941223	0.140218234	Purifying selection
CsTLP17	*CsTLP23*	0.150486337	1.046030202	0.143864237	Purifying selection
CsTLP2	*CsTLP23*	0.01274531	0.087468426	0.145713266	Purifying selection
CsTLP31	*CsTLP12*	0.164720597	1.124320781	0.146506762	Purifying selection
CsTLP23	*CsTLP34*	0.162786538	1.108646387	0.146833598	Purifying selection
CsTLP23	*CsTLP12*	0.173184246	1.169327199	0.14810589	Purifying selection
CsTLP2	*CsTLP34*	0.15535431	1.04032929	0.149331862	Purifying selection
CsTLP30	*CsTLP4*	0.009204334	0.061376083	0.149966139	Purifying selection
CsTLP2	*CsTLP12*	0.16347525	1.067605023	0.153123343	Purifying selection
CsTLP17	*CsTLP31*	0.166935939	1.072339635	0.155674502	Purifying selection
CsTLP2	*CsTLP17*	0.164226389	1.033530499	0.158898445	Purifying selection
CsTLP29	*CsTLP3*	0.012201027	0.073017619	0.167097026	Purifying selection
CsTLP20	*CsTLP29*	0.012211186	0.072807602	0.167718551	Purifying selection
CsTLP22	*CsTLP30*	0.014442965	0.079377688	0.181952455	Purifying selection
CsTLP22	*CsTLP4*	0.006899244	0.033797959	0.204131965	Purifying selection
CsTLP17	*CsTLP12*	0.011277348	0.049619508	0.227276498	Purifying selection
CsTLP19	*CsTLP26*	0.023624	0.087477678	0.270057469	Purifying selection
CsTLP19	*CsTLP7*	0.019099955	0.067569783	0.282670065	Purifying selection
CsTLP26	*CsTLP7*	0.031138646	0.087691004	0.3550951	Purifying selection

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2201_249263