SCIENTIA PLENA
www.scientiaplena.org.br
VOL. 20, NUM. 09 2024
doi: 10.14808/sci.plena.2024.095201

Genome-Wide Identification and Characterization Thaumatin-like Protein (TLP) Genes in Wild Olive (*Olea europaea* var. *sylvestris*)

Identificação e Caracterização de Genes Thaumatin-Like Proteins (TLPs) em Todo Genoma da Azeitona Selvagem (*Olea europaea* var. *sylvestris*)

M. C. S. Alves[1]*; R. da S. de Souza[1]; R. C. C. da Silva[2]

[1]*Department of Genetics, Biosciences Center, Federal University of Pernambuco, 50670-423, Recife - PE, Brazil*
[2]*Biological Sciences Research Center, Federal University of Ouro Preto, 35400-000, Ouro Preto - MG, Brazil*

**cidinaria.alves@ufpe.br*
(Recebido em 29 de novembro de 2023; aceito em 27 de agosto de 2024)

The wild olive tree (*Olea europaea* var. sylvestris) is the ancestor of all olive varieties, displaying remarkable adaptability to diverse climatic conditions and soil types. This iconic tree of the Mediterranean Basin stands out for its ability to withstand environmental stresses, such as water scarcity and temperature variations, playing a crucial role in the ecology, economy, and culture of the region. In this study, 42 TLP genes were characterized in the wild olive tree, distributed across 14 chromosomes and nine scaffolds. In silico analysis revealed the presence of domains associated with antifungal activity in some sequences, indicating their potential role in pathogen resistance. The structural diversity of OeTLP genes was explored through multiple alignments, conserved motifs and phenetic analysis. Correlation was observed between phenetic diversity, genetic structure, and motif patterns, suggesting different functions among OeTLP groups. Additionally, the distribution of exons and introns in the pheneticc tree provided further insights into the evolution of these genes. Comparative modeling of OeTLPs unveiled three-dimensional models with good structural quality and distinctive surface charge characteristics. This study provides a comprehensive understanding of the physicochemical, structural, and phenetic features of TLP genes in Wild Olive, contributing to knowledge about this species' response to environmental stresses and pathogens.
Keywords: genome, antimicrobial peptides, thaumatin.

A oliveira selvagem (*Olea europaea* var. sylvestris) é a ancestral de todas as variedades de oliveira, exibindo notável adaptabilidade a diversas condições climáticas e tipos de solo. Essa árvore icônica da Bacia do Mediterrâneo destaca-se por sua capacidade de resistir a estresses ambientais, como escassez hídrica e variações de temperatura, desempenhando um papel crucial na ecologia, economia e cultura da região. Nesse estudo, foram caracterizados 42 genes de TLPs na oliveira selvagem, distribuídos em 14 cromossomos e nove scaffolds. A análise *in silico* revelou a presença de domínios associados a atividade antifúngica em algumas sequências, indicando seu potencial papel na resistência a patógenos. A diversidade estrutural dos genes OeTLP foi explorada por meio de alinhamentos múltiplos, motivos conservados e análise filogenética. Foi observada correlação entre diversidade filogenética, estrutura genética e padrões de motivos, sugerindo diferentes funções entre grupos de OeTLPs. Além disso, a distribuição de éxons e íntrons na árvore filogenética proporcionou *insights* adicionais sobre a evolução desses genes. A modelagem comparativa das OeTLPs revelou modelos tridimensionais com boa qualidade estrutural e características distintas de carga superficial. Este estudo oferece uma compreensão abrangente das características físico-químicas, estruturais e filogenéticas dos genes TLP em Oliveira Selvagem, contribuindo para o conhecimento sobre a resposta dessa espécie a estresses ambientais e patógenos.
Palavras-chave: genoma, peptídeos antimicrobianos, taumatina.

1. INTRODUCTION

Pathogenesis-related proteins (PR) are renowned for their relevance in conferring resistance to environmental stresses and are classified into 17 families (PR-1 to PR-17). Thaumatin-like Proteins (TLPs), belonging to the PR-5 family, are particularly notable due to their similarity to thaumatin found in the plant *Thaumatococcus daniellii*, native to West Africa [1, 2].

The TLP family consists of proteins with an average sequence length of about 207 amino acids, a molecular mass ranging between 18 kDa and 26 kDa, and contains eight disulfide linkages due to the presence of 16 cysteine (Cys) residues [3]. These characteristics render TLPs stable under various stress conditions, including high temperatures and pH variations [3, 4]. Additionally, they feature a conserved motif known as the acid cleft (REDDD), composed of conserved amino acids such as arginine, glutamic acid, and three aspartic acid residues, playing a crucial role in binding to specific receptors when TLPs are involved in antifungal activity [3-5]. Furthermore, TLPs exhibit a conserved signature motif, G-X-[GF]-X-C-X-T-[GA]-D-C-X(1,2)-[GQ]-X(2,3)-C, present in most TLPs [3, 6].

Overexpression of TLPs in plants results in resistance to various pathogens such as *Phytopthora infestans* and *Macrophomina phaseolina* [7], as well as conferring resistance to abiotic stresses such as salinity [8] and drought [9, 10]. Further research has shown that TLPs play significant roles in plant physiology, including fruit ripening, seed germination e antimicrobial activity [3, 11, 12].

TLPs have been identified in various plant species, including *Arabidopsis thaliana* [13], *Arachis diogoi* [14], *Cucumis melo* (L.) [15], *Citrullus lanatus* (Thunb.) [16], *Glycine max* (L.) [17], *Hordeum vulgare* [18], *Oryza sativa* (L.) [13], and *Vitis vinifera* (L.) [19]. However, TLPs exhibit significant diversity among different plant species and in response to different stresses, suggesting that their functions still require further investigation.

Given this, to explore the diversity of TLPs, the present study investigates the TLPs of the wild olive tree (*Olea europaea* L.). This species represents one of the most iconic trees in the Mediterranean Basin, with significant social, economic, and ecological implications [20, 21]. The wild olive (var. *sylvestris*) is the ancestor of all olive varieties. This tree has developed a remarkable ability to adapt to a wide range of climatic conditions, varied altitudes, and soil types [22]. Additionally, it demonstrates a notable capacity to withstand water scarcity and distinct temperature variations [23].

Drastic climate changes can adversely impact olive cultivation, with predictions of warming and decreased precipitation in the Mediterranean in the coming decades, leading to serious economic and ecological implications [24, 25]. These changes are likely to result in a significant reduction in the suitable area for olive cultivation, given the adaptation of widely cultivated varieties to specific climatic conditions [21, 23]. For a better understanding of strategies favoring the species' adaptation, it is essential to study genes that participate, especially under adverse conditions, in various situations within the plant's genetic framework.

The aim of this study was to characterize TLP genes in Wild Olive (*Olea europaea* var. *sylvestris*), analyzing their physicochemical characteristics, chromosomal location, gene structure, phenetic tree, and three-dimensional structure.

2. MATERIAL AND METHODS

2.1. Identification of TLPs

The genome of *Olea europaea* var. *sylvestris* was obtained from NCBI (https://www.ncbi.nlm.nih.gov/, RefSeq: GCF_002742605.1). The identification of OeTLPs was carried out using the characterization method suggested by dos Santos-Silva et al. (2020) [26]. Was conducted by aligning seed sequences of TLPs available in the public UniProt database (https://www.uniprot.org/) with the wild olive genome database using the BLAST tool.

The identified candidate sequences were translated using the Transdecoder tool and then subjected to Batch CD-Search (http://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi) and InterProScan (https://www.ebi.ac.uk/interpro/search/sequence/) for the identification of conserved domains characteristic of the TLP class. The potential presence of transmembrane domains (TM) and signal peptides was predicted using the TMHMM 2.0 [27] and SignalP 5.0 [28] tools, respectively.

2.2. Characterization of Sequences

The sequences that possessed the complete domain were subjected to a comprehensive analysis, including the assessment of protein size, isoelectric point (pI), molecular weight (MW), and grand average of hydropathicity (GRAVY), based on predictions provided by the ExPASy ProtParam service (http://web.expasy.org/protparam/). Additionally, subcellular localization was predicted using the ProtComp 9.0 program from Softberry Inc. (Mount Kisco, NY, USA).

2.3. Multiple Sequence Alignment and Phenetic Analysis

To identify conserved patterns, a multiple sequence alignment was performed using the Clustal Omega tool [29]. The secondary structure was obtained from the JNet Secondary Structure Prediction package [30]. A phenetic tree was constructed using the Neighbor-Joining (NJ) method with the assistance of the MEGA v.7 program [31]. The results were subsequently visualized on iTOL (Interactive Tree of Life) [32].

2.4. Gene Structure, Conserved Motif, and Chromosome Location Analysis

The Gene Structure Display Server 2.0 (http://gsds.cbi.pku.edu.cn/) software was used to investigate the exon-intron organizations of TLP genes. For the identification of conserved motifs, the MEME 5.0.3 tool was employed. Chromosome location analysis was performed using the MG2C software (v2.1).

2.5. Three-Dimensional Modeling

Six OeTLPs were selected for generating 3D models. Initially, the homolog with the highest identity for each OeTLP was assessed in AlphaFold [33]. One hundred theoretical models were generated for each TLP sequence using the Modeller software (v.10.1) [34]. Models that best solved the minimum energy by consensus of solutions that best met the special modeling constraints were selected based on the DOPE-Score (Discrete Optimized Protein Energy) energy metric [35], which evaluates the likely structure of the model.

The DOE-MBI/SAVES6.0 server was used via PROCHECK to investigate critical regions and stereochemical conflicts of the models [36]. Subsequently, the PROSA-Web: Protein Structure Analyses server was used to evaluate the folding of the models using the Z-score metric [37, 38]. The QMEANDisCo tool was used to assess the overall quality estimate of the theoretical models [39].

The physicochemical characterization of biomolecular solvent-solute interactions was performed using the APBS-PDB2PQR software suite (http://server.poisonboltzmann.org). The PDB2PQR algorithm was used for model preparation, and pKa parameters were calculated using the PROPKA algorithm for protonation state signatures at pH 7, with the Charmm force field [40]. The solvation energy estimate was calculated using the Poisson-Boltzmann equation, which simplifies the model for polar solvation energies. Structural features of TLP models based on conserved sites were accessed by comparing the 3D models using the PyMOL software (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

3. RESULTS

3.1. In Silico Identification of TLP Genes Across the Genome

In the wild olive genome, 42 TLP transcripts were identified (Table 1). The 42 TLP genes identified were mapped onto chromosomes or scaffolds. The TLP genes were distributed across 14 chromosomes and nine scaffolds, with 29 and 13 TLP genes, respectively (Figure 1). Chromosome 8 (Chr8) harbored the highest number of TLP genes, with 4, followed by

chromosomes Chr2, Chr5, and Chr7, each with 3 genes, along with scaffold6670_v1. Chr10, Chr13, Chr15, and scaffold894 had 2 genes each, while the remaining chromosomes and scaffolds contained only one TLP gene.

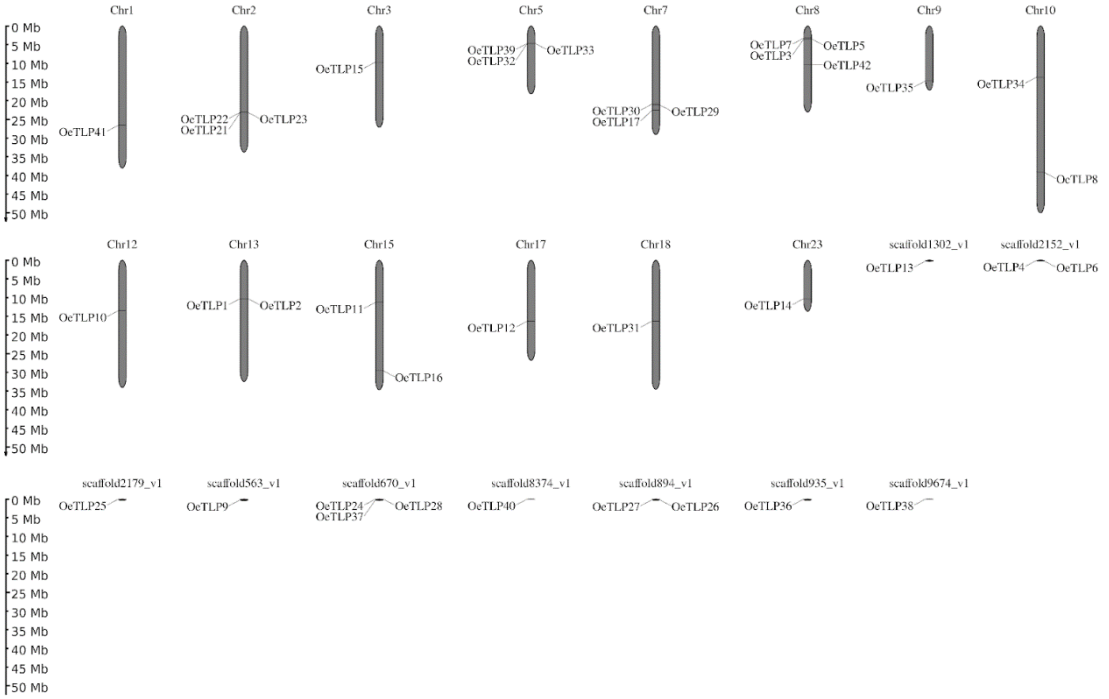

Figure 1: Distribution of OeTLP Genes on Chromosomes and Scaffolds. The scale represents megabases (Mb). Chromosome and scaffold numbers are indicated above each vertical bar.

Out of the 42 genes, 7 sequences presented the GH64-TLP-SF domain, which includes glycoside hydrolases. Additionally, 7 of these sequences contained the Thaumatin domain, which is related to antifungal activity. Finally, 28 transcripts exhibited the TLP-PA domain, similar to the thaumatin domain, found in proteins with antifungal and allergenic properties.

The majority of OeTLPs showed the presence of a N-terminal signal peptide, except for OeTLPs 25, 38, and 39. Out of the total 42, 34 were found in the extracellular region, while 4 were located in the vacuolar region and 4 in the vacuolar membrane-associated region.

Regarding OeTLPs, the absence of transmembrane domains was observed in 48% of them, while 18 OeTLPs presented one transmembrane domain, and 4 OeTLPs exhibited two of these domains. Mature proteins, excluding the signal peptide, ranged in length from 180 (OeTLP39) to 228 (OeTLP36) amino acids (aa), with an average size of 215 aa. The molecular mass of the proteins varied from 18.70 kDa (OeTLP39) to 24.22 kDa (OeTLP40), and the isoelectric point (pI) ranged between 4.12 (OeTLP30) and 9.13 (OeTLP12 to 23) (Table 1).

Table 1: Domain, Protein size (aa), molecular weight (MW), isoelectric point (pI), average hydropathicity index (GRAVY), transmembrane domains (TM), and subcellular localization of OeTLPs.

Gene	Gene-ID	Domain	TM	Length (aa)	pI	MW (kDa)	GRAVY	Subcellular Localization
OeTLP1	XP_022892831.1	GH64-TLP-SF	0	202	21.95	4.76	-0.430	Membrane bound Vacuolar
OeTLP2	XP_022892832.1	GH64-TLP-SF	1	202	22.19	8.17	-0.514	Membrane bound Vacuolar
OeTLP3	XP_022881237.1	GH64-TLP-SF	1	201	21.68	8.49	-0.414	Membrane bound Vacuolar
OeTLP4	XP_022859839.1	GH64-TLP-SF	1	201	21.59	8.15	-0.409	Membrane bound Vacuolar
OeTLP5	XP_022881236.1	GH64-TLP-SF	0	204	21.75	4.79	-0.252	Vacuolar
OeTLP6	XP_022859834.1	GH64-TLP-SF	1	200	21.69	8.58	-0.260	Vacuolar
OeTLP7	XP_022881229.1	GH64-TLP-SF	1	205	22.43	8.32	-0.290	Vacuolar
OeTLP8	XP_022885825.1	Thaumatin	0	213	22.90	7.55	-0.175	Extracellular
OeTLP9	XP_022869736.1	TLP-PA	0	213	22.62	6.72	-0.090	Extracellular
OeTLP10	XP_022890429.1	TLP-PA	2	216	22.45	4.43	-0.085	Extracellular
OeTLP11	XP_022896086.1	TLP-PA	1	216	22.16	4.19	-0.081	Extracellular
OeTLP12	XP_022841482.1	TLP-PA	0	219	23.19	8.92	-0.143	Extracellular
OeTLP13	XP_022853410.1	TLP-PA	2	216	22.28	4.24	-0.097	Extracellular
OeTLP14	XP_022850376.1	TLP-PA	0	219	23.03	8.76	-0.102	Extracellular
OeTLP15	XP_022865367.1	TLP-PA	1	217	22.76	4.40	-0.234	Extracellular
OeTLP16	XP_022896985.1	TLP-PA	1	212	22.71	5.31	-0.259	Vacuolar
OeTLP17	XP_022880043.1	TLP-PA	0	223	23.12	4.55	-0.114	Extracellular
OeTLP18	XP_022842766.1	TLP-PA	0	219	23.34	9.13	-0.216	Extracellular
OeTLP19	XP_022842750.1	TLP-PA	0	219	23.34	9.13	-0.216	Extracellular
OeTLP20	XP_022842743.1	TLP-PA	0	219	23.34	9.13	-0.216	Extracellular
OeTLP21	XP_022842761.1	TLP-PA	0	219	23.34	9.13	-0.216	Extracellular
OeTLP22	XP_022842734.1	TLP-PA	0	219	23.34	9.13	-0.216	Extracellular
OeTLP23	XP_022842726.1	TLP-PA	0	219	23.34	9.13	-0.216	Extracellular
OeTLP24	XP_022871089.1	TLP-PA	1	223	23.27	4.53	-0.033	Extracellular
OeTLP25	XP_022860008.1	TLP-PA	1	220	22.77	4.67	-0.057	Extracellular
OeTLP26	XP_022873208.1	Thaumatin	1	223	23.11	4.37	0.020	Extracellular
OeTLP27	XP_022873207.1	Thaumatin	1	223	23.11	4.37	0.020	Extracellular
OeTLP28	XP_022871091.1	TLP-PA	1	223	23.28	4.46	-0.033	Extracellular
OeTLP29	XP_022879913.1	TLP-PA	1	221	22.73	4.43	-0.152	Extracellular
OeTLP30	XP_022880476.1	TLP-PA	1	221	22.53	4.12	-0.081	Extracellular
OeTLP31	XP_022844303.1	TLP-PA	0	223	23.12	4.77	-0.125	Extracellular
OeTLP32	XP_022876048.1	TLP-PA	2	221	23.03	4.68	-0.181	Extracellular
OeTLP33	XP_022876047.1	TLP-PA	2	221	23.03	4.68	-0.181	Extracellular
OeTLP34	XP_022884720.1	TLP-PA	1	221	22.95	4.34	-0.145	Extracellular
OeTLP35	XP_022882974.1	Thaumatin	1	222	23.45	7.87	0.022	Extracellular
OeTLP36	XP_022873609.1	Thaumatin	0	228	24.20	7.54	-0.179	Extracellular
OeTLP37	XP_022871100.1	TLP-PA	1	221	23.09	4.31	-0.238	Extracellular
OeTLP38	XP_022873877.1	TLP-PA	0	220	22.98	4.18	-0.149	Extracellular
OeTLP39	XP_022875687.1	Thaumatin	0	180	18.70	4.58	-0.365	Extracellular
OeTLP40	XP_022872855.1	TLP-PA	0	220	24.22	8.09	-0.217	Extracellular
OeTLP41	XP_022880772.1	Thaumatin	0	220	23.44	7.80	-0.038	Extracellular
OeTLP42	XP_022881590.1	TLP-PA	0	223	23.93	7.32	-0.121	Extracellular

3.2. Multiple Alignment and Conserved Motifs of OeTLPs

The alignment results revealed that OeTLPs exhibited 16 conserved cysteine (C) residues, with the exception of OeTLP39, which had 15 cysteine residues due to the protein length limitation (Figure 2). Additionally, the REDDD motif [Arginine (R), glutamic acid (E), and three aspartic acid (D) residues] was also conserved in the majority of OeTLPs. OeTLP6 and OeTLP7 showed substitutions in arginine to lysine (K) and in the second aspartic acid to lysine (K) and histidine (H). In OeTLP38, the arginine region was replaced by histidine (H). In OeTLP42, the α-helix region of glutamic acid was replaced by glutamine (Q), and the first aspartic acid was replaced by glycine (G). Sequences OeTLP5 and OeTLP40 exhibited the substitution of the third aspartic acid by glycine and glutamic acid.

The alignment also presented the typical signature motif of the TLP family. The prediction of the secondary structure of OeTLPs indicated that the sequences had nine to fourteen β-sheets, along with one to three α-helices (Figure 2).

To explore the structural diversity of OeTLP genes, we conducted an analysis and mapping of the distribution of exons and introns on the phenetic tree (Figure 3B). We observed that five TLP genes (OeTLP1, OeTLP4, OeTLP6, OeTLP25, and OeTLP42) did not have introns, while twelve genes contained only one intron, such as OeTLP2, OeTLP3, and OeTLP5. Thirteen genes, including OeTLP8, OeTLP10, and OeTLP11, exhibited two introns, while the remaining twelve genes (OeTLP9, OeTLP13, and OeTLP20) had three introns.

To deepen our understanding of the structural diversity and functional characteristics of TLPs, we investigated motif patterns among OeTLPs (Figure 3C). Our findings revealed a total of 18 motifs, with 10 of them being highly conserved. All OeTLPs contained motifs 1, 2, 3, 6, and 7, with 35 of them also presenting motif 4. Motifs 5, 8, 9, and 10 were frequently identified in most proteins. Additionally, we noticed that some groups exhibited unique motifs, such as motif 15 in group II and motif 17 in proteins OeTLP8 and OeTLP9 of group I, along with motifs 11 and 18 in group IV.

While we observed differences in motif types among groups, members belonging to the same group tended to share similar motif patterns. For example, OeTLP18 to OeTLP23, OeTLP11 and OeTLP13, and OeTLP1 to OeTLP7 exhibited similar motif patterns, suggesting the possibility of highly similar functions among these members. It became evident that phenetic diversity was correlated with genetic structure and motif patterns (Figure 3A).

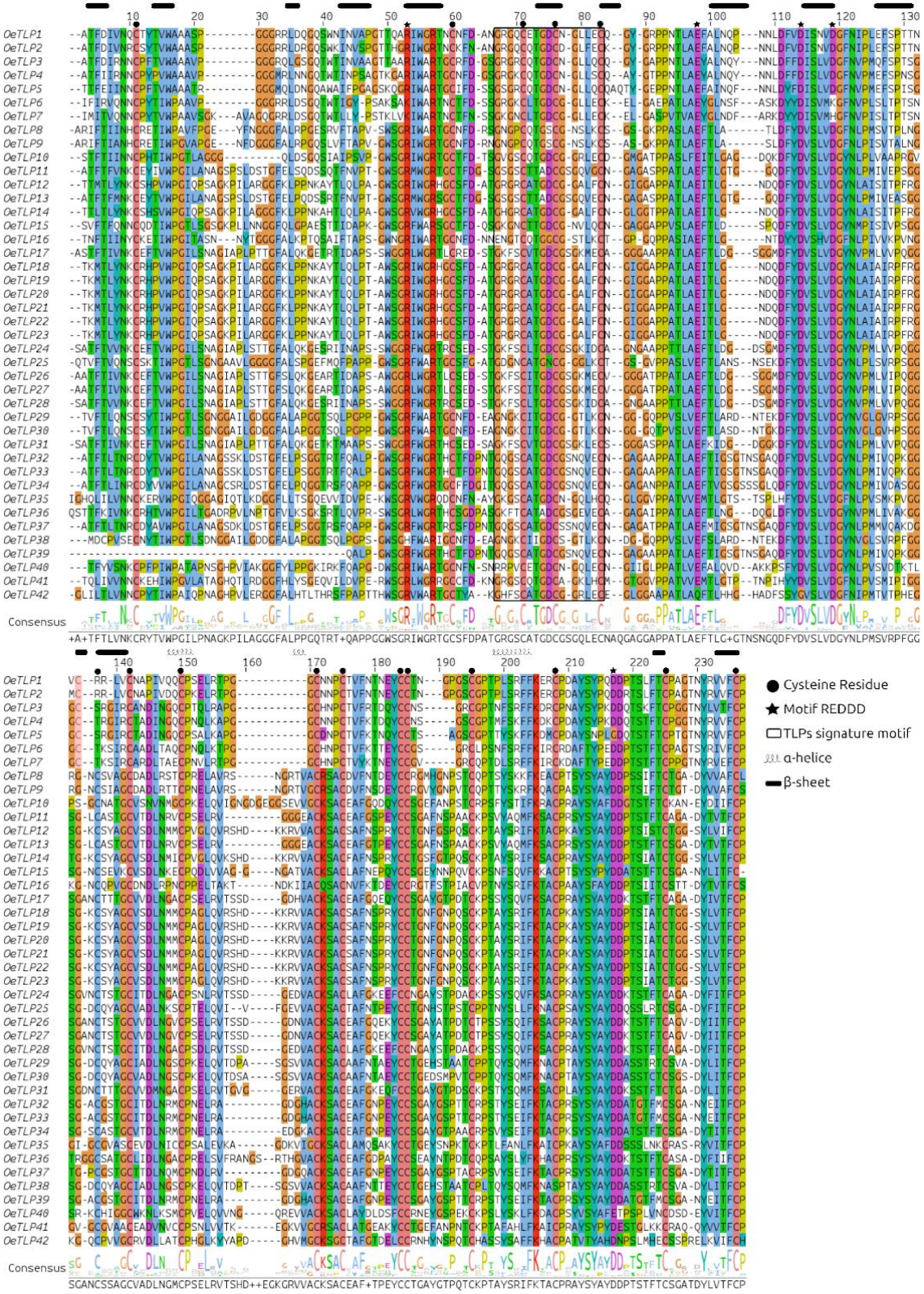

Figure 2: Multiple sequence alignment and prediction of the secondary structure of Thaumatin amino acid sequences and OeTLPs.

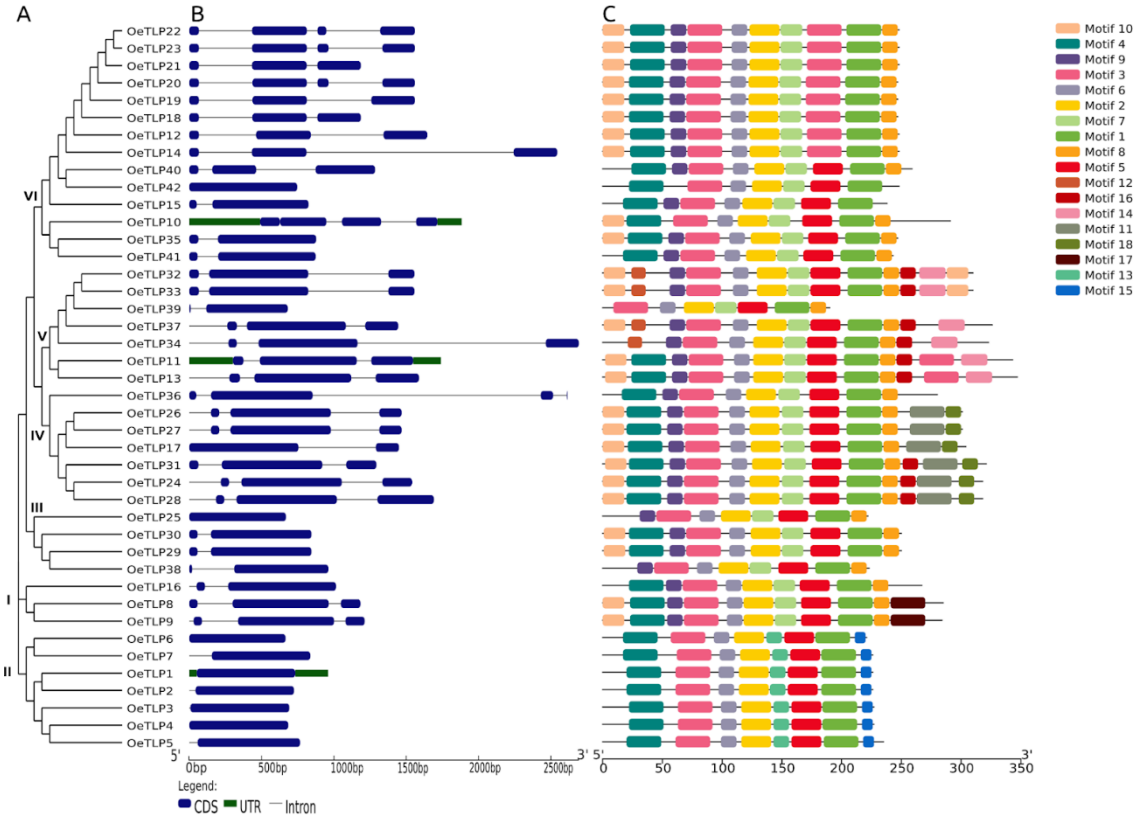

Figure 3: Cluster analysis, genetic structure, and domain analysis of the OeTLP gene family. A. Phenetic tree of OeTLPs formed by the NJ method (bootstrap test: 1000 replicates). B. Exon-intron structures of OeTLP genes. C. Conserved motif patterns of OeTLPs. The 18 motifs are arranged by colors.

3.3. Phenetic Analysis of OeTLPs

The phenetic tree was divided into 10 paraphyletic groups based on their phenetic relationship, named as groups I–X (Figure 4). The highest number of TLPs was found in group X, followed by group VI and group IV, with 14, 10, and 7 OeTLPs, respectively. Groups I, II, and IX did not group any OeTLP. In groups with OeTLPs, the presence of other legume species is also observed.

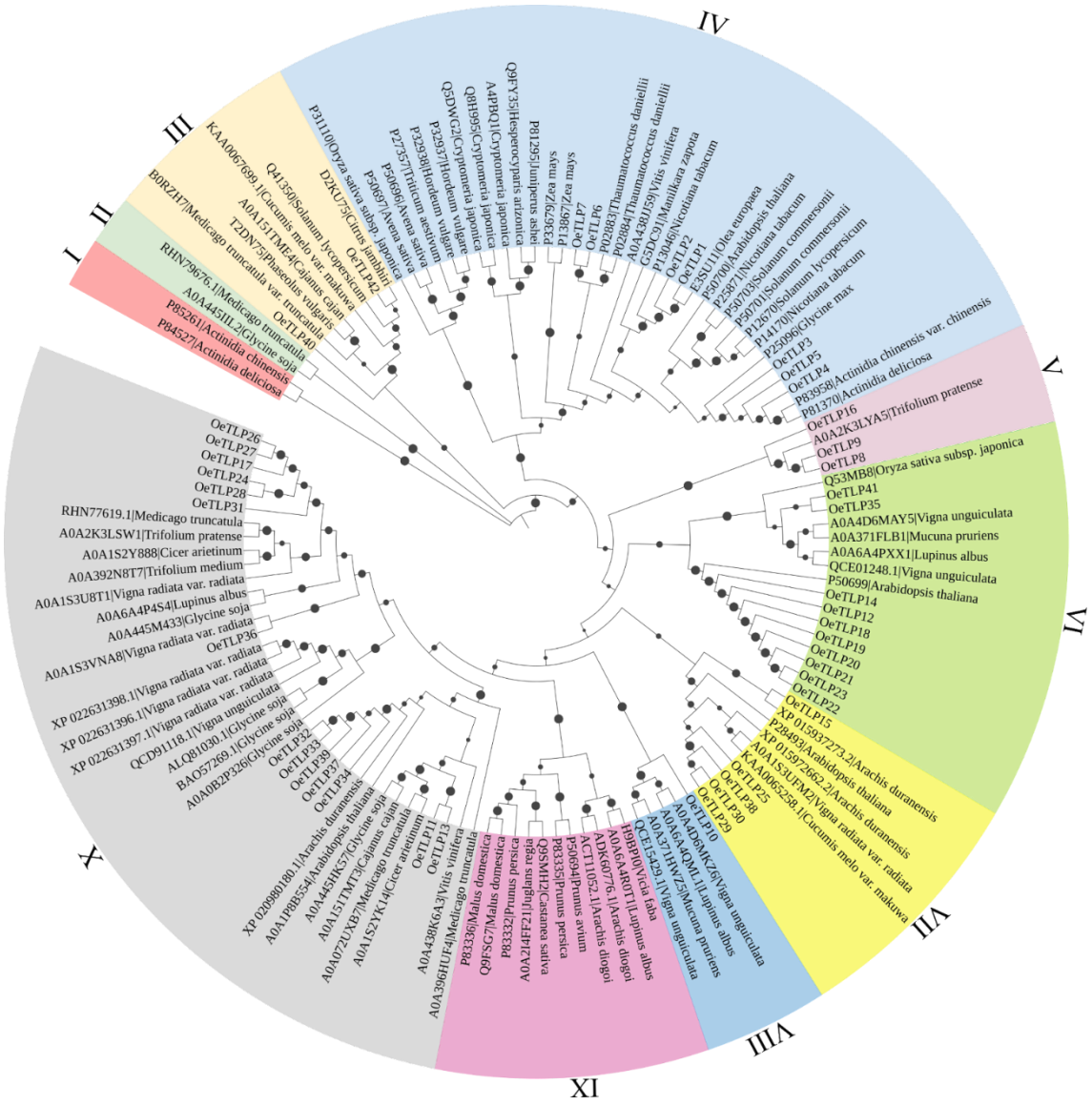

Figure 4: Phenetic analysis generated through the NJ method (bootstrap test: 2000 replicates) of OeTLPs compared to TLPs from the NCBI and UniProt databases. Proteins are represented by their accession codes followed by the plant species. Ten groups were classified.

3.4. Comparative Modeling of OeTLPs

For modeling, templates with complete coverage were selected. For OeTLP4 and OeTLP10, the best models were from *Glycine soja* with 82.59% identity and *Citrus unshiu* with 83.26%. Meanwhile, for OeTLP15, a model with 71.89% identity was obtained from a TLP of *Artemisia annua*. Templates from *Davidia involucrata* were chosen for OeTLP23 and OeTLP31, with identities of 89.04% and 80.27%, respectively. The best-selected model for OeTLP42 was a TLP structure from *Theobroma cacao* with 82.06% identity.

After stereochemical analysis of amino acid residue sequences, it was observed that over 90.1% of amino acid residues were in favorable positions for all models (Table 2). Regarding the folding quality of the models, Z-Score values were observed to be approximately similar to those of experimentally solved models of the same size, ranging between -7.90 and -6.13. The estimation of global energy function by the QmeanDisCo algorithm showed global energy between 0.73 and 0.86, with energy peaks in loop regions.

Table 2: Results of the algorithms for modeling and validating the theoretical models of OeTLPs.

TLP	Template	Organism	Coverage	Identity	Ramachandran	QMEANDisCo	Z-score
OeTLP4	A0A445KRA3	*Glycine soja*	Total	82.59%	90.10%	0.86 ± 0.06	-6.52
OeTLP10	A0A2H5NR19	*Citrus unshiu*	Total	83.26%	91.20%	0.77 ± 0.06	-6.44
OeTLP15	A0A2U1L7J3	*Artemisia annua*	Total	71.89%	91.55%	0.80 ± 0.06	-7.90
OeTLP23	A0A5B6Z1K0	*Davidia involucrata*	Total	89.04%	90.72%	0.77 ± 0.06	-6.99
OeTLP31	A0A5B6YK31	*Davidia involucrata*	Total	80.27%	90.90%	0.80 ± 0.06	-7.42
OeTLP42	A0A061GY18	*Theobroma cacao*	Total	82.06%	92.10%	0.73 ± 0.06	-6.13

The obtained models exhibit unique compactation and fluctuation characteristics, strongly influencing structural changes (Figure 5A). Through the characterization of the electrostatic potential map (Figure 5B), it was observed that OeTLP4 and OeTLP23 predominantly exhibited a basic physicochemical character (blue) on the surface, compared to models OeTLP10, OeTLP15, and OeTLP31, which had densely exposed acidic amino acid residues (red) on the electronegative cleft surface of the REDDD domain. Additionally, the OeTLP42 model showed a predominantly neutral physicochemical character (white), with both acidic and basic regions.

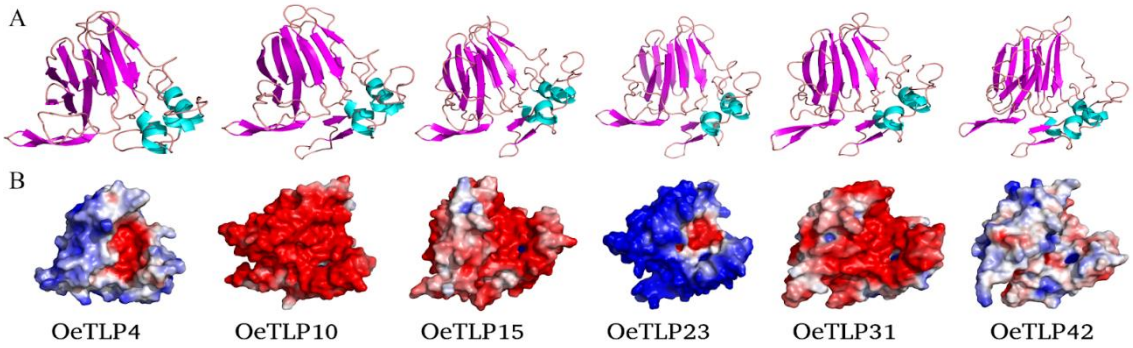

Figure 5: Three-dimensional models of OeTLPs. A. Structure before (green) and after (magenta) molecular dynamics simulation with the respective distance (in nanometers) between the models and the position of domains represented by I, II, and III. B. The electrostatic surface potential for theoretical three-dimensional models of OeTLPs, where the red color indicates acidic charge, and the blue color indicates basic charge.

4. DISCUSSION

The impact of abiotic stress is evident in the environmental factors that most affect the growth and development of plants, triggering noteworthy physiological, biochemical, and molecular changes [41]. These include the accumulation of reactive oxygen species, loss of photosynthetic efficiency, membrane damage, among others [9, 41, 42].

In response to these stresses, plants commonly exhibit a systemic response, mobilizing genes present in different organs and tissues [42, 43]. Studies indicate the activation of various signaling pathways in the regulation of PR (pathogenesis-related) gene expression following abiotic stress, including drought and salinity conditions [44, 43]. Among the PR-5 proteins, Thaumatin-like proteins (TLPs) stand out for playing a crucial role in plant defense against various stresses [45, 46].

In the genome of wild olive, we identified 42 Thaumatin-like proteins (TLPs), compared to 19, 37, 28, and 35 TLPs in barley, *Oryza sativa*, *Brachypodium*, and sorghum, respectively [18];

28 TLPs in *Cucumis melo* [15] and *Arabidopsis* [13]; 29 in *Brassica napus* [47]; 33 in *Vitis vinifera* [19]; 55 in *Populus trichocarpa* [6], and 56 in *Vigna unguiculata* [48]. This variation in the number of TLP-encoding genes is expected due to the significant roles of TLPs associated with stress response [48] and indicates considerable variation in the expansion of TLP genes among different plant groups [49].

Wild olive trees, known for their tolerance to water deficit, exhibit defense mechanisms triggered by TLPs during periods of stress, such as increased antioxidant enzyme activity and cuticle thickening to prevent transpiration, ultimately enhancing their tolerance to pathogen infections like anthracnose [50].

Previous studies highlight common characteristics of TLPs, such as acidic or basic isoelectric points and low molecular mass [3, 4, 51]. The identified TLPs in this study have an average length of 211 amino acids, with isoelectric points ranging from 4.12 to 9.13 and molecular masses between 18.70 kDa and 24.22 kDa.

PR proteins, including TLPs, are distributed in various plant organs, with a predominance in floral buds, epidermis, roots, and leaves. However, the localization of TLPs may vary, being predicted in the extracellular region [52] or accumulated in the vacuole in response to pathogenic attacks [53]. In this study, most TLPs were predicted in the extracellular region, while about 19% were predicted in the vacuole. Similar findings were reported by Tachi et al. (2009) [54], evaluating TLP expression in soybean (*Glycine max*) roots under saline stress, where the majority of TLPs were predicted in the extracellular region (while others remained in the vacuole), as well as in upland cotton (*Gossypium hirsutum*) under drought [55].

Most OeTLPs presented a N-terminal signal peptide and at least one transmembrane domain, as observed in cotton [55] and melon [15]. Such sequences are present in most TLPs and are crucial for directing them to secretory pathways, indicating their role in defense as extracellular proteins [14, 47].

The identified OeTLPs exhibited conserved domains Thaumatin, TLP-PA, and the GH64-TLP-SF superfamily, showing high similarity to thaumatin, as reported in *Populus szechuanica* under infection by the fungus *Melampsora laricis-populina* [56].

Alignment of OeTLPs revealed 16 highly conserved cysteine residues in most sequences, responsible for disulfide bridge formation, allowing correct folding and conferring stability to the protein [57]. The alignment also showed amino acids highly conserved and corresponding to the signature motif (G-X-[GF]-X-C-X-T-[GA]-D-C-X(1,2)-[GQ]-X(2,3)-C) present throughout the PR-5 protein family [48, 54].

The REDDD motif (arginine (R), glutamic acid (E), and three aspartic acid residues (D)) was present in most OeTLPs. Some OeTLPs showed substitutions in the REDDD motif, resulting in variations in protein surface topology and electrostatic potential, indicating functional specificity of these TLPs. This motif is essential for TLP stability under stress conditions [58-60]. Several studies have reported the presence of such motifs in TLPs from various species, such as *Arabidopsis* and rice [61], melon [15], and grape [19].

The secondary structure of OeTLPs revealed up to three α-helices and nine to fourteen β-sheets. In melon, TLPs had up to four α-helices and fourteen β-sheets [15]. Typically, TLPs have this structure with a predominance of β-sheets, which are crucial for their function and stability [3].

Due to the composition of their sequences, TLP proteins are paraphyletic, originating from 10 common ancestors before the differentiation of monocots and dicots [13, 62]. In this study, OeTLPs and TLPs from other species were distributed into 10 groups, relatively dispersed, similar to results obtained in Arabidopsis [3], cotton [55], and melon [15].

The dispersed distribution of OeTLPs in groups indicates structural and functional similarities with TLPs from other species, as described by Wang et al. (2011) [47]. This variation in the distribution of OeTLPs occurs due to the differential duplication of TLP family genes [63] and suggests that TLPs may play different functions under different types of stresses and plant growth stages [49].

The absence of OeTLPs in groups I, II, and IX suggests that TLPs from *Olea europaea* var. *sylvestris* share 7 out of 10 common ancestral genes with other plants [13]. The absence of the TLP gene in certain groups, such as groups 1 and 2, can be attributed to genetic deletion or

insertion events, as suggested by the evolutionary study of gene families. The presence of only two species in these groups may indicate a significant evolutionary distance from other species, leading to divergence over time and potentially resulting in gene loss in some lineages or independent gene acquisition in others [64]. The evolutionary analyses also suggest that TLPs have evolved under purifying selection pressure, contributing to their expansion and diverse functions in stress responses in plants [65]. All groups composed of OeTLPs also included TLPs from legumes, indicating high similarity among TLPs in this family.

The amino acid sequences of OeTLPs revealed 10 highly conserved motifs (out of 18 predicted), and some sequences had exclusive motifs. These sequence specificities are commonly found in plant TLPs and demonstrate that some amino acids undergo diversification of selection [66]. Thus, due to the abundance of cysteines and intense preservation of the REDDD motif, OeTLPs exhibit high conservation, indicating functional similarity with TLPs, while molecule modifications suggest functional specificity [59, 60].

Most OeTLP models exhibited surface charge predominately composed of anionic or cationic residues, as observed in *Musa acuminata* [67] and *Zea mays* [58]. However, due to substitutions in regions of the REDDD motif, OeTLP42 exhibits a surface charge predominantly composed of neutral amino acid residues, suggesting a distinct function compared to other OeTLPs. Additionally, certain TLPs have electrically charged amino acids on their surfaces, providing them with specific electrostatic properties that facilitate interactions with other charged molecules [59, 60].

5. CONCLUSION

The comprehensive analysis of Thaumatin-like proteins (TLPs) in *O. europaea* var. *sylvestris* has revealed significant insights into their genomic diversity, molecular structure, and potential functions. The presence of 42 TLPs in the genome of this species, compared to other plants, underscores the complexity and specific adaptation of this protein family in response to adverse environmental conditions.

This study substantially contributes to the understanding of genetic regulation related to plant defense. The information obtained may be pivotal for the development of agricultural and biotechnological strategies aimed at enhancing plant resistance to adverse conditions, leading to significant advancements in sustainable agriculture and food security. Furthermore, it paves the way for future investigations into the molecular interactions and signaling mechanisms associated with TLPs in *Olea europaea* var. *sylvestris*.

6. REFERENCES

1. Christensen AB, Cho BH, Næsby M, Gregersen PL, Brandt J, Madriz-Ordeñana 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(3):135-44. doi: 10.1046/j.1364-3703.2002.00105.x
2. van Loon LC, Rep M, Pieterse CMJ. Significance of inducible defense-related proteins in infected plants. Annu Rev Phytopathol. 2006;44(1):135-62. doi: 10.1146/annurev.phyto.44.070505.143425
3. Liu J-J, Sturrock R, Ekramoddoullah AKM. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. Plant Cell Rep. 2010;29(5):419-36. doi: 10.1007/s00299-010-0826-8
4. Breiteneder H. Thaumatin-like proteins – a new family of pollen and fruit allergens. Allergy. 2004;59(5):479-81. doi: 10.1046/j.1398-9995.2003.00421.x
5. Ghosh R, Chakrabarti C. Crystal structure analysis of NP24-I: a thaumatin-like protein. Planta. 2008;228(5):883-90. doi: 10.1007/s00425-008-0790-5
6. Zhao JP, Su XH. Patterns of molecular evolution and predicted function in thaumatin-like proteins of *Populus trichocarpa*. Planta. 2010;232(4):949-62. doi: 10.1007/s00425-010-1218-6
7. 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(2):609-22. doi: 10.1007/s12033-012-9603-y

8. 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(6):1909-25. doi: 10.1007/s00425-012-1733-8
9. He C, Du Y, Fu J, Zeng E, Park S, White F, et al. Early drought-responsive genes are variable and relevant to drought tolerance. G3 Genes Genomes Genetics. 2020;10(5):1657-70. doi: 10.1534/g3.120.401199
10. Liu W-X, Zhang F-C, Zhang W-Z, Song L-F, Wu W-H, Chen Y-F. Arabidopsis Di19 functions as a transcription factor and modulates PR1, PR2, and PR5 expression in response to drought stress. Mol Plant. 2013;6(5):1487-502. doi: 10.1093/mp/sst031
11. 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 grape1. Plant Physiol. 1998;117(2):465-72. doi: 10.1104/pp.117.2.465
12. 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(3):334-44. doi: 10.1093/pcp/pcn011
13. 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(1):12-29. doi: 10.1007/s00239-005-0053-z
14. 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(12):e83963. doi: 10.1371/journal.pone.0083963
15. 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(3):2499-509. doi: 10.1016/j.ygeno.2020.02.001
16. Zhang M, Xu J, Liu G, Yang X. Antifungal properties of a thaumatin-like protein from watermelon. Acta Physiol Plant. 2018;40(11):186. doi: 10.1007/s11738-018-2759-8
17. 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(9):1679-89. doi: 10.1093/pcp/pcu099
18. Iqbal I, Tripathi RK, Wilkins O, Singh J. Thaumatin-Like Protein (TLP) gene family in barley: Genome-wide exploration and expression analysis during germination. Genes. 2020;11(9):1080. doi: 10.3390/genes11091080
19. 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(1):4269. doi: 10.1038/s41598-017-04105-w
20. Carrión Y, Ntinou M, Badal E. *Olea europaea* L. in the North Mediterranean Basin during the Pleniglacial and the Early–Middle Holocene. Quat Sci Rev. 2010;29(7):952-68. doi: 10.1016/j.quascirev.2009.12.015
21. Fanelli V, Mascio I, Falek W, Miazzi MM, Montemurro C. Current status of biodiversity assessment and conservation of wild olive (*Olea europaea* L. subsp. *europaea* var. *sylvestris*). Plants. 2022;11(4):480. doi: 10.3390/plants11040480
22. Belaj A, Gurbuz Veral M, Sikaoui H, Moukhli A, Khadari B, Mariotti R, et al. Olive genetic resources. In: Rugini E, Baldoni L, Muleo R, Sebastiani L, organizers. The olive tree genome. Cham: Springer International Publishing; 2016. p. 27-54. (Compendium of Plant Genomes). doi: 10.1007/978-3-319-48887-5_3
23. Arenas-Castro S, Gonçalves JF, Moreno M, Villar R. Projected climate changes are expected to decrease the suitability and production of olive varieties in southern Spain. Sci Total Environ. 2020;709:136161. doi: 10.1016/j.scitotenv.2019.136161
24. Gualdi S, Somot S, Li L, Artale V, Adani M, Bellucci A, et al. The CIRCE simulations: Regional climate change projections with realistic representation of the Mediterranean Sea. Bull Am Meteorol Soc. 2013;94(1):65-81. doi: 10.1175/BAMS-D-11-00136.1
25. Kassa A, Konrad H, Geburek T. Molecular diversity and gene flow within and among different subspecies of the wild olive (*Olea europaea* L.): A review. Flora. 2019;250:18-26. doi: 10.1016/j.flora.2018.11.014
26. dos Santos-Silva CA, Zupin L, Oliveira-Lima M, Vilela LMB, Bezerra-Neto JP, Ferreira-Neto JR, et al. Plant antimicrobial peptides: State of the art, in silico prediction and perspectives in the omics era. Bioinforma Biol Insights. 2020;14:1177932220952739. doi: 10.1177/1177932220952739
27. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. J Mol Biol. 2001;305(3):567-80. doi: 10.1006/jmbi.2000.4315

28. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat Biotechnol. 2019;37(4):420-3. doi: 10.1038/s41587-019-0036-z
29. Sievers F, Higgins DG. Clustal omega, accurate alignment of very large numbers of sequences. In: Russell DJ, organizer. Multiple sequence alignment methods. Totowa (US): Humana Press; 2014. p. 105-16. (Methods in Molecular Biology). doi: 10.1007/978-1-62703-646-7_6
30. Cole C, Barber JD, Barton GJ. The Jpred 3 secondary structure prediction server. Nucleic Acids Res. 2008;36(suppl_2):W197-201. doi: 10.1093/nar/gkn238
31. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33(7):1870-4. doi: 10.1093/molbev/msw054
32. Letunic I, Bork P. Interactive tree of life (iTOL) v3: An online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44(W1):W242-5. doi: 10.1093/nar/gkw290
33. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, et al. AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 2022;50(D1):D439-44. doi: 10.1093/nar/gkab1061
34. Webb B, Sali A. Comparative protein structure modeling using MODELLER. Curr Protoc Bioinforma. 2014;47(1):5.6.1-32. doi: 10.1002/0471250953.bi0506s47
35. Shen M, Sali A. Statistical potential for assessment and prediction of protein structures. Protein Sci. 2006;15(11):2507-24. doi: 10.1110/ps.062416606
36. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: A program to check the stereochemical quality of protein structures. J Appl Crystallogr. 1993;26(2):283-91. doi: 10.1107/S0021889892009944
37. Sippl MJ. Boltzmann's principle, knowledge-based mean fields and protein folding. An approach to the computational determination of protein structures. J Comput Aided Mol Des. 1993;7(4):473-501. doi: 10.1007/BF02337562
38. Wiederstein M, Sippl MJ. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007;35(suppl_2):W407-10. doi: 10.1093/nar/gkm290
39. Studer G, Rempfer C, Waterhouse AM, Gumienny R, Haas J, Schwede T. QMEANDisCo—distance constraints applied on model quality estimation. Bioinformatics. 2020;36(6):1765-71. doi: 10.1093/bioinformatics/btz828
40. Brooks BR, Brooks III CL, Mackerell Jr. AD, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: The biomolecular simulation program. J Comput Chem. 2009;30(10):1545-614. doi: 10.1002/jcc.21287
41. Rajput VD, Minkina T, Kumari A, Harish, Singh VK, Verma KK, et al. Coping with the challenges of abiotic stress in plants: New dimensions in the field application of nanoparticles. Plants. 2021;10(6):1221. doi: 10.3390/plants10061221
42. Bartlett MK, Klein T, Jansen S, Choat B, Sack L. The correlations and sequence of plant stomatal, hydraulic, and wilting responses to drought. Proc Natl Acad Sci. 2016;113(46):13098-103. doi: 10.1073/pnas.1604088113
43. Islam MM, El-Sappah AH, Ali HM, Zandi P, Huang Q, Soaud SA, et al. Pathogenesis-related proteins (PRs) countering environmental stress in plants: A review. South Afr J Bot. 2023;160:414-27. doi: 10.1016/j.sajb.2023.07.003
44. Wu J, Kim SG, Kang KY, Kim J-G, Park S-R, Gupta R, et al. Overexpression of a pathogenesis-related protein 10 enhances biotic and abiotic stress tolerance in rice. Plant Pathol J. 2016;32(6):552-62. doi: 10.5423/PPJ.OA.06.2016.0141
45. An M, Tong Z, Ding C, Wang Z, Sun H, Xia Z, et al. Molecular characterization of the thaumatin-like protein PR-NP24 in tomato fruits. J Agric Food Chem. American Chemical Society. 2019;67(47):13001-9. doi: 10.1021/acs.jafc.9b05256
46. 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(3):299-307. doi: 10.1007/s10725-015-0134-y
47. 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(2):e16930. doi: 10.1371/journal.pone.0016930
48. de Jesús-Pires C, Ferreira-Neto JRC, Bezerra-Neto JP, Kido EA, Silva RLO, Pandolfi V, et al. Plant Thaumatin-like Proteins: Function, evolution and biotechnological applications. Curr Protein Pept Sci. 2020;21(1):36-51.
49. Su L, Zhao X, Geng L, Fu L, Lu Y, Liu Q, et al. Analysis of the thaumatin-like genes of *Rosa chinensis* and functional analysis of the role of RcTLP6 in salt stress tolerance. Planta. 2021;254(6):118. doi: 10.1007/s00425-021-03778-y

50. Innamorato APC, Cambón JOB, de León IP. Olive-tree physiological response to biotic and abiotic stress fruit yield, oil quality and tolerance to Anthracnose: Doctoral thesis abstract. Agrociencia Urug. 2024;28(Supplement theses):e1474-e1474. doi: 10.31285/AGRO.28.1474
51. 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(1):27-38. doi: 10.1111/j.1399-3054.2009.01338.x
52. Samac DA, Peñuela S, Schnurr JA, Hunt EN, Foster-Hartnett D, Vandenbosch KA, et al. Expression of coordinately regulated defence response genes and analysis of their role in disease resistance in *Medicago truncatula*. Mol Plant Pathol. 2011;12(8):786-98. doi: 10.1111/j.1364-3703.2011.00712.x
53. Jami SK, Swathi Anuradha T, Guruprasad L, Kirti PB. Molecular, biochemical and structural characterization of osmotin-like protein from black nightshade (*Solanum nigrum*). J Plant Physiol. 2007;164(3):238-52. doi: 10.1016/j.jplph.2006.01.006
54. Tachi H, Fukuda-Yamada K, Kojima T, Shiraiwa M, Takahara H. Molecular characterization of a novel soybean gene encoding a neutral PR-5 protein induced by high-salt stress. Plant Physiol Biochem. 2009;47(1):73-9. doi: 10.1016/j.plaphy.2008.09.012
55. Li Z, Wang X, Cui Y, Qiao K, Zhu L, Fan S, et al. 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. 2020;11:1-15. doi: 10.3389/fpls.2020.575015
56. Chen ZJ, Cao ZM, Yu ZD, Yu D. Cloning and characterization of defense-related genes from Populus szechuanica infected with rust fungus *Melampsora larici-populina*. Genet Mol Res GMR. 2016;15(1):1-15. doi: 10.4238/gmr.15017314
57. 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(3):646-54. doi: 10.1080/14756360802321831
58. Batalia MA, Monzingo AF, Ernst S, Roberts W, Robertus JD. The crystal structure of the antifungal protein zeamatin, a member of the thaumatin-like, PR-5 protein family. Nat Struct Biol. 1996;3(1):19-22. doi: 10.1038/nsb0196-19
59. Faillace GR, Caruso PB, Timmers LFSM, Favero D, Guzman FL, Rechenmacher C, et al. Molecular Characterisation of soybean osmotins and their involvement in drought stress response. Front Genet. 2021;632685(12):1-16. doi: 10.3389/fgene.2021.632685
60. Min K, Ha SC, Hasegawa PM, Bressan RA, Yun D-J, Kim KK. Crystal structure of osmotin, a plant antifungal protein. Proteins Struct Funct Bioinforma. 2004;54(1):170-3. doi: 10.1002/prot.10571
61. Kaur A, Pati PK, Pati AM, Nagpal AK. Physico-chemical characterization and topological analysis of pathogenesis-related proteins from *Arabidopsis thaliana* and *Oryza sativa* using in-silico approaches. PLOS ONE. 2020;15(9):e0239836. doi: 10.1371/journal.pone.0239836
62. Kohler A, Rinaldi C, Duplessis S, Baucher M, Geelen D, Duchaussoy F, et al. Genome-wide identification of NBS resistance genes in *Populus trichocarpa*. Plant Mol Biol. 2008;66(6):619-36. doi: 10.1007/s11103-008-9293-9
63. Sharma A, Sharma H, Rajput R, Pandey A, Upadhyay SK. Molecular characterization revealed the role of thaumatin-like proteins of bread wheat in stress response. Front Plant Sci. 2022;12:807448. doi: 10.3389/fpls.2021.807448
64. Chen H, Zhang Y, Feng S. Whole-genome and dispersed duplication, including transposed duplication, jointly advance the evolution of TLP genes in seven representative Poaceae lineages. BMC Genomics. 2023;24(1):290. doi: 10.1186/s12864-023-09389-z
65. Sharma A, Madhu, Kaur A, Kumar Upadhyay S. Chapter 11 - Thaumatin-domain containing receptor-like kinases in plants. In: Upadhyay SK, Shumayla, organizers. Plant receptor-like kinases. Academic Press; c2023. p. 207-21. doi: 10.1016/B978-0-323-90594-7.00011-9
66. 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(1):33. doi: 10.1186/1471-2229-11-33
67. 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(1):45-52. doi: 10.1016/j.biochi.2005.07.001